Cloudera Data Visualization 8.1.7

Visualization

Date published: 2020-10-30
Date modified: 2026-10-30

CLOUDERA

https://docs.cloudera.com/

Legal Notice

© Cloudera Inc. 2026. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Creating a visual 7

Editing a visual 10

Cloning a visual 10

Linking visuals 11

Unlinking visuals 12

Emailing a visual 13

Exporting a visual 16

Importing a visual 16

Downloading a visual 16

- Downloading a visual as a PNG file 16
- Downloading a visual as a PDF file 18
- Downloading a visual as a PPTX file 19
- Downloading a visual as a CSV file 20
- Downloading a visual as an Excel file 21

Removing a visual from a dashboard 22

Deleting a visual 23

Adding filters to a visual 24

Creating annotation for visuals 26

Commenting on visuals and dashboards [Technical Preview] 31

Exploring visual options 35

Visual type catalog.....37
Table.....39
Shelves for tables.....46
Saving tables.....47
Cross tabulation.....52
Displaying totals with data expansion.....57
Shelves for cross tabulation.....64
Bars.....64
Filtering 1: Bar chart by year.....71
Filtering 2: bar chart by region.....73
Horizontal bars.....76
Stacked bars.....78
Labelled bars.....80
Shelves for bars.....83
Lines.....83
Lines with dual axes.....88
Shelves for lines.....90
Combined bar/line visuals.....91
Creating basic combos.....92
Assigning a measure to a secondary axis.....95
Change line to area.....97
Shelves for combos.....97
Areas.....98
Using the trellis option with areas.....100
Using the trellis option with independent scale.....101
Shelves for areas.....102
Grouped bars.....103
Shelves for grouped bars.....107
KPI.....108
Shelves for KPIs.....112
Gauge.....113
Shelves for gauges.....121
Row listing.....122
Shelves for row listing visuals.....124
Bullet.....124
Shelves for bullets.....132
Box plot.....132
Shelves for box plots.....136
Packed bubbles.....137
Shelves for packed bubbles.....141
Word cloud.....142
Trellised word clouds.....145
Shelves for word clouds.....145
Scatter.....146
Creating a basic scatter visual.....147
Adding size to scatter visuals.....151
Adding transition animation to scatter visuals.....154
Adding trellis to scatters.....158
Shelves for scatters.....160
Pie.....161
Shelves for pies.....165
Radial.....166
Trellised radial.....170
Shelves for radials.....171

Chord....172
Shelves for chords....175
Network....176
Shelves for networks....182
Advanced options for networks....183
Dendrograms....191
Shelves for dendrograms....195
Treemap....195
Shelves for treemaps....199
Correlation heatmap....200
Shelves for correlation heatmaps....203
Correlation flow....204
Shelves for correlation flows....206
Calendar heatmaps....207
Shelves for calendar heatmaps....211
Maps....211
Choropleth country map....212
Country map with bubbles....216
Choropleth US state map....220
US state map with bubbles....223
Choropleth US county map....225
US county map with bubbles....226
US zipcode map with bubbles....228
Latitude and longitude map with bubbles....231
Choropleth DMA map....232
Choropleth Canadian province map....235
Canadian province map, bubbles....238
Choropleth Canadian postal code map....241
Canadian postal code map, bubbles....244
Shelves for maps....247
Interactive maps....248
Basic interactive map....249
Choropleth interactive maps....255
Interactive map with multiple dimensions....261
Interactive map with multiple measures....262
Changing the map server for interactive maps....264
Changing layer options for interactive maps....265
Plotting routes on interactive maps....266
Using alphabetic values in interactive maps....268
Segmenting data qualitatively in interactive maps....271
Shelves for interactive maps....274
Sparklines....275
Using sparklines....276
Shelves for sparklines....278
AI visual....278
Creating an AI visual....279
Using an AI visual....283
Shelves for AI visual....285
Action....286
Search....291
Timeline....291
Shelves for timelines....294
External link visuals....295
Setting up dynamic external links....296
Shelves for external links....299
Histograms....299

Specifying the bucket count..... 302
Specifying data range..... 304
Normalizing the histogram..... 307
Using the cumulative option..... 308
Shelves for histograms..... 310
Extensions..... 311
Examples of extensions..... 314
Rich texts..... 315
Inserting URL links..... 318
Inserting tables..... 319
Inserting images from local machine..... 319
Inserting images from static assets..... 321
Resizing images..... 322
Inserting videos..... 322
Field statistics..... 323
SQL queries..... 326

Creating a visual

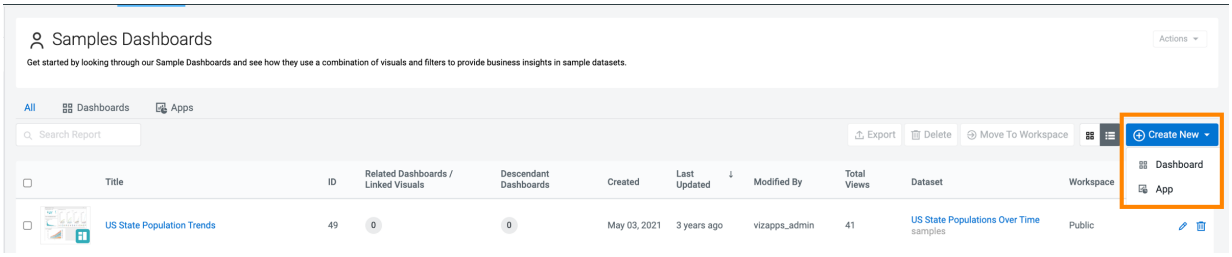
You can easily create new visual representations for your data in Cloudera Data Visualization.

About this task

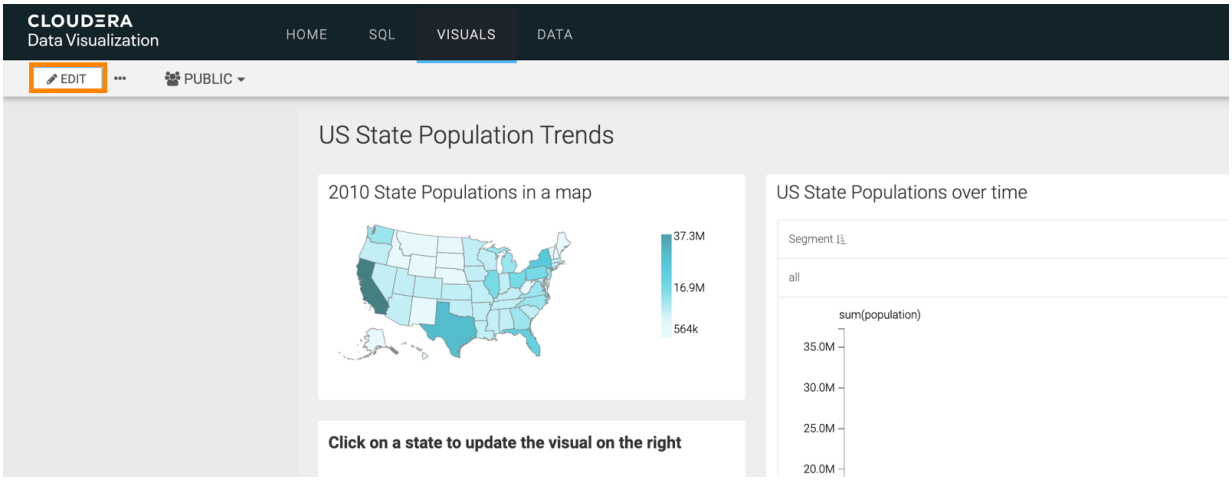
The following steps demonstrate how to create a new visual.

Procedure

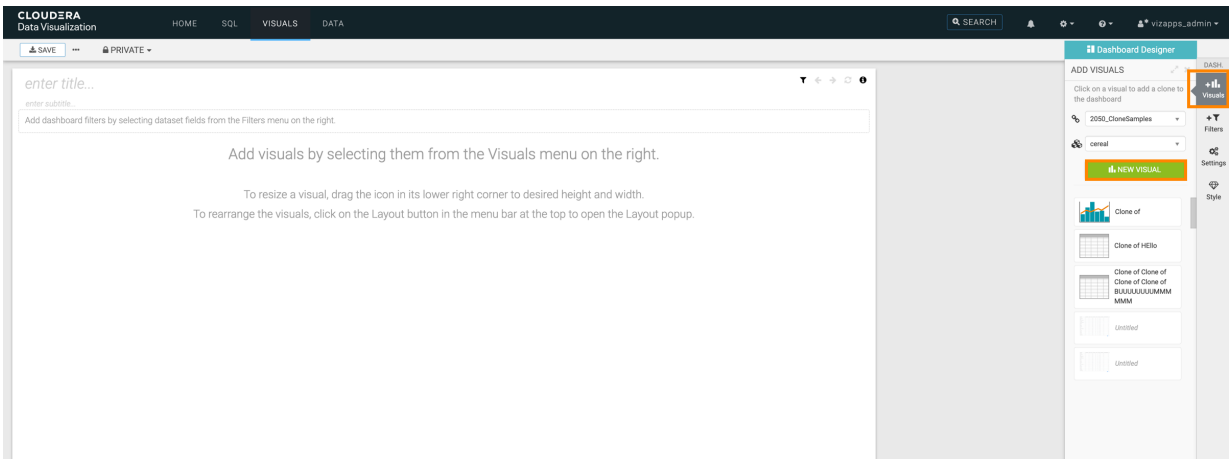
1. On the main navigation bar, click VISUALS.
2. Click Create New Dashboard to open a new, blank dashboard.



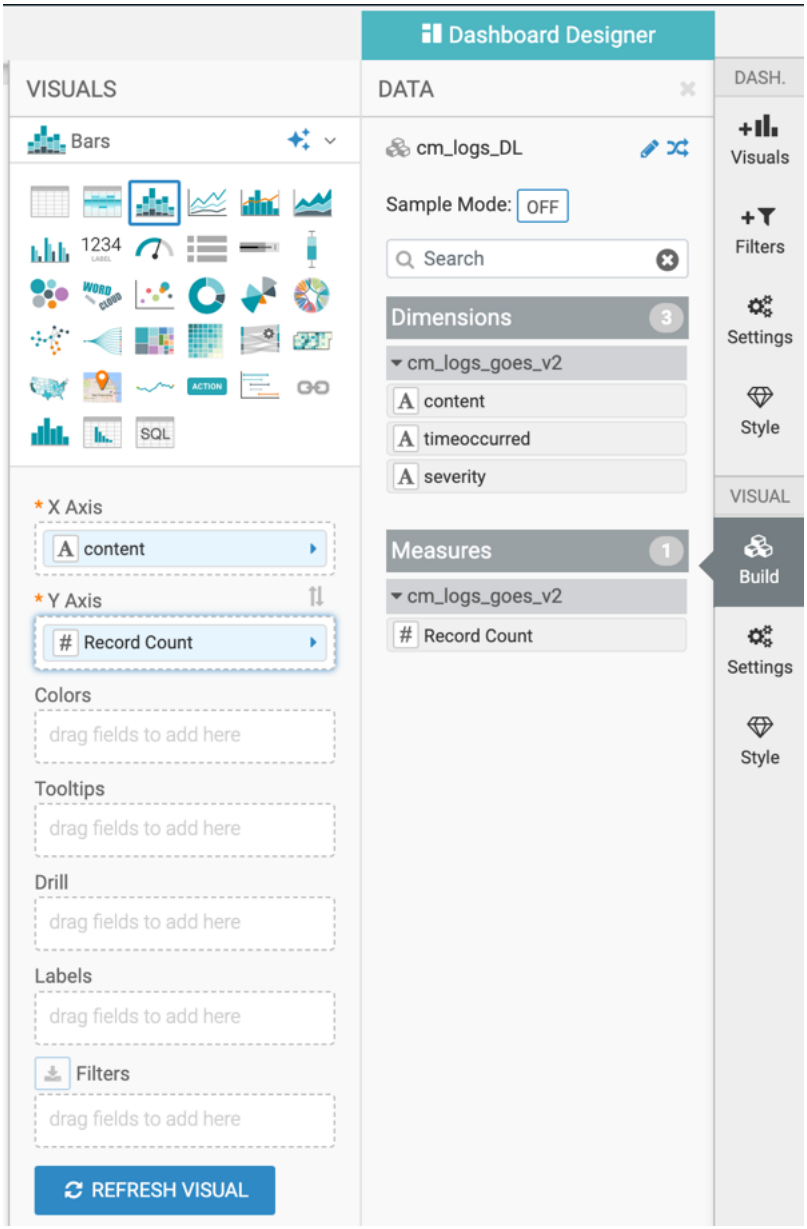
If you already have an existing dashboard and you just want to add a new visual, open the dashboard and click EDIT to make changes to it.



3. In the Dashboard Designer interface, open the Visuals menu from the side menu bar and click NEW VISUAL.

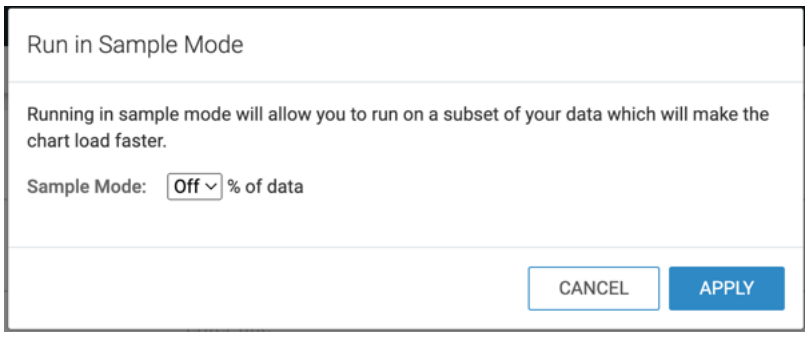


4. The Visual Builder appears.
   By default, a table visual is selected. For more information, see the *Visual Builder* overview.
5. Choose the visual type you want to build.
   In this example, the Bars visual has been selected.



6. You can switch on Sample Mode for the visual you are building.
   With sample mode, you create the visual on a subset of your data to test it. You can define this subset as a percentage of the original data.

   Note: This feature is only available on Impala, Hive, and Spark connections.



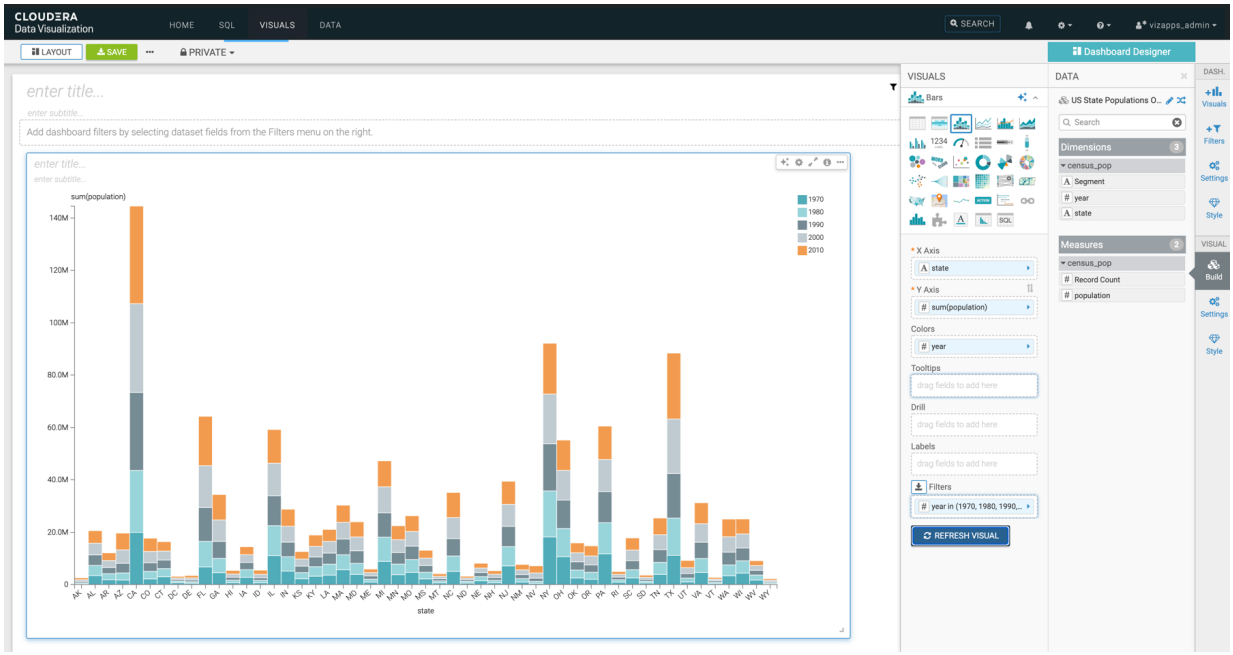
7. Populate the shelves of the visual from the Dimensions and Measures fields:
   In this example, the following fields have been added:
   - state dimension on the X Axis shelf
   - population measure on the Y Axis shelf

     Note: The population measure appears with its default aggregation, as sum(population).

   - year measure on the Colors shelf
   - year measure on the Filters shelf

     Note: The Filter for year module opens automatically. Years 1970, 1980, 1990, 2000, and 2010 have been selected.

8. Click REFRESH VISUAL.



9. Click the enter title... field and enter a name for the visual.
   - In this example the title 'Recent US Census Results by State' has been added.
   - You can also click the enter subtitle... field below the title of the visualization to add a brief description of the visual.

10. Click SAVE at the top left corner of the Dashboard Designer.

Related Information

Visual Builder

Editing a visual

In Cloudera Data Visualization, you can edit an existing visual to update its configuration, data, or appearance.

Before you begin

You must have Edit access to the dashboard that contains the visual to be able to edit it.

Procedure

1. Open the dashboard that contains the visual in Edit mode.
2. Click the visual to select it.
   The VISUAL menu appears in the right-hand panel.
3. Click Build to open the Visual Builder panel.
4. Make the required changes.
5. Click REFRESH VISUAL to preview your changes while making your edits.
6. When you are satisfied with your edits, click SAVE in the top-left corner of the Dashboard Designer.

Cloning a visual

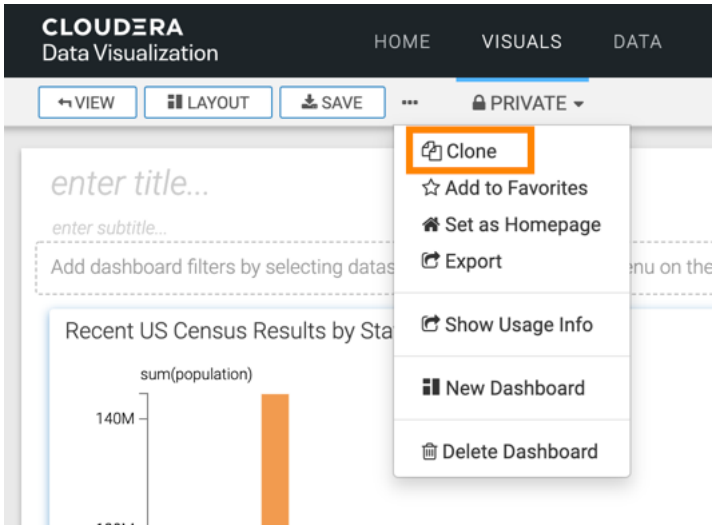
In Cloudera Data Visualization, you can create a new visual that is a clone of an existing visual.

About this task

The following steps demonstrate how to clone an existing visual.

Procedure

In the Dashboard Designer, click the (ellipsis) icon at the top of the interface, and select Clone.



You should briefly see a green success notification. The main area of the screen now has a new visual titled Clone of...[Source Visual].

The SAVE button is not highlighted, because the process adds a cloned, unlinked version of the visual to the bottom of the dashboard. This cloned visual includes any setting or shelf changes you made before clicking Clone.

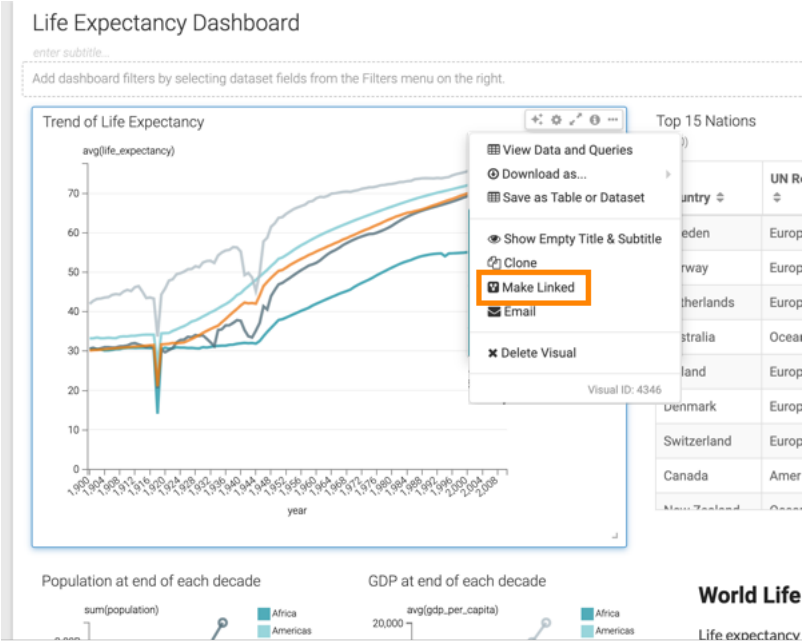
After working on the visual, you can rename and save it.

Linking visuals

In Cloudera Data Visualization, you can link visuals and these linked visuals can be shared across various dashboards and sheets within one dashboard. The following steps demonstrate how to make an existing visual as linked.

Procedure

1. Open a dashboard in Edit mode and hover over a visual until its control widgets appear.
2. Click the (ellipsis) icon at the top of the visual to open the supplemental menu.
3. Select Make Linked.

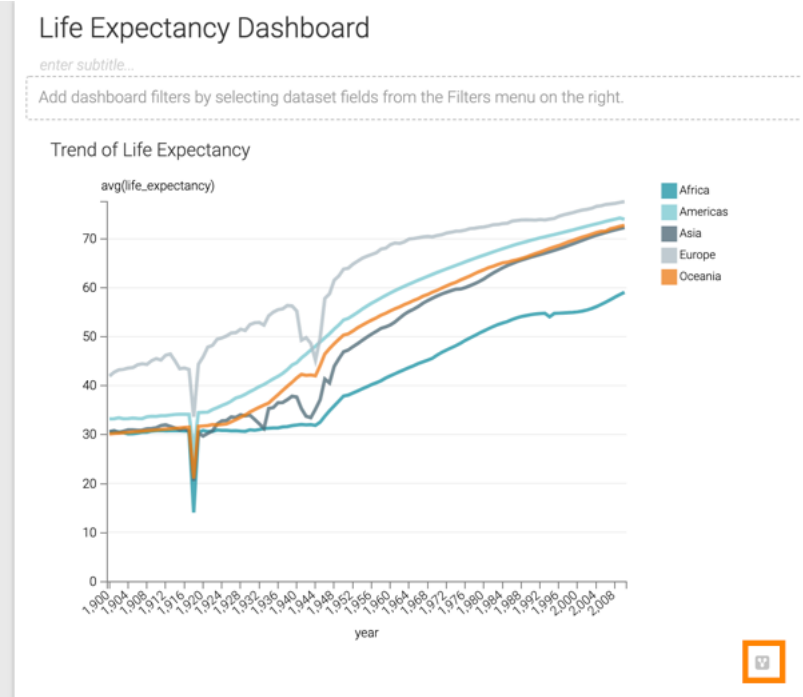


Note: Linked visuals cannot be deleted when they are dropped from all dashboards. They are still available to be added in future work.

4. Confirm the operation by clicking LINK.

Results

The visual can now be used in other dashboards. You can see a small icon in the lower right corner of the visual showing the linked designation.



For instructions on how to add a linked visual to a dashboard, see *Adding visuals to a sheet*.

For information on how to unlink a visual, see *Unlinking visuals*.

Unlinking visuals

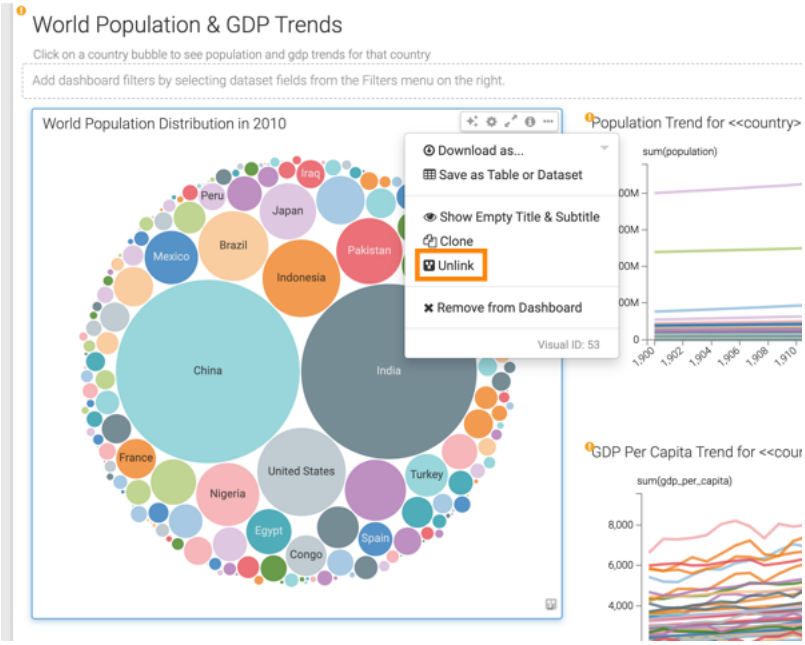
In Cloudera Data Visualization, you can unlink linked visual to break the sharing across dashboards.

About this task

The following steps demonstrate how to unlink a visual.

Procedure

1. In the Dashboard Desginer interface, hover over the visual until its control widgets appear.
2. Click the (ellipsis) icon at the top of the visual to open the supplemental menu.
3. Select Unlink.



4. Confirm the operation by clicking UNLINK.

   When a visual appears in multiple dashboards, it cannot be unlinked. You can clone it instead. Confirm the operation by clicking CLONE.

   After you unlinked the visual by cloning it, it no longer has a linked designation. it is completely independent of the original visual, and its title is Clone of …[Original Name].
5. You can change the title of the visual to the original title, or use something different.
6. At the top left corner of the Dashboard Designer, click SAVE.

Results

The visual no longer has the small icon in the lower right corner showing the linked designation.

Emailing a visual

Cloudera Data Visualization enables you to share visuals in an email. You can send them immediately, schedule them at specific times and intervals, or through a triggered response when reaching a specific monitored threshold.

Before you begin

Before using the email function, you must enable the feature in site settings. For instructions, see *Email settings*.

Note: This setting is only available for users with administrative privileges.

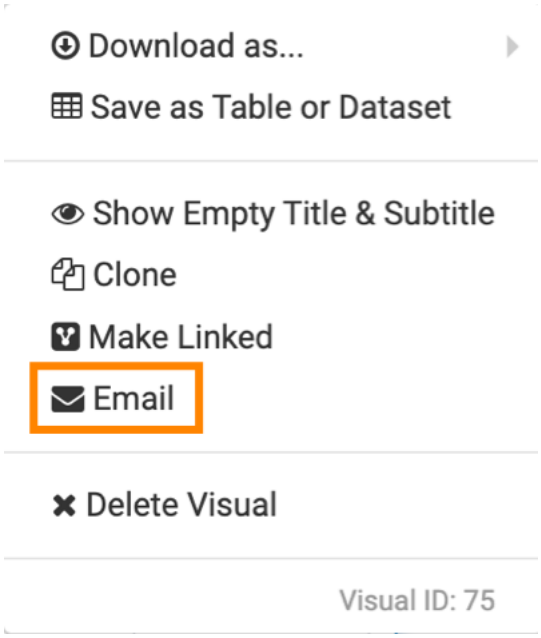
Complete these prerequisites in your system configuration:

- Enable screenshot and thumbnail downloads. For instructions, see *Snapshots*.
- Enable email functionality. For instructions, see *Email settings*.
- Enable scheduled jobs. For instructions, see *Enabling scheduled jobs*.

Procedure

1. Find the visual you want to send in an email and open it in Edit or in View mode.

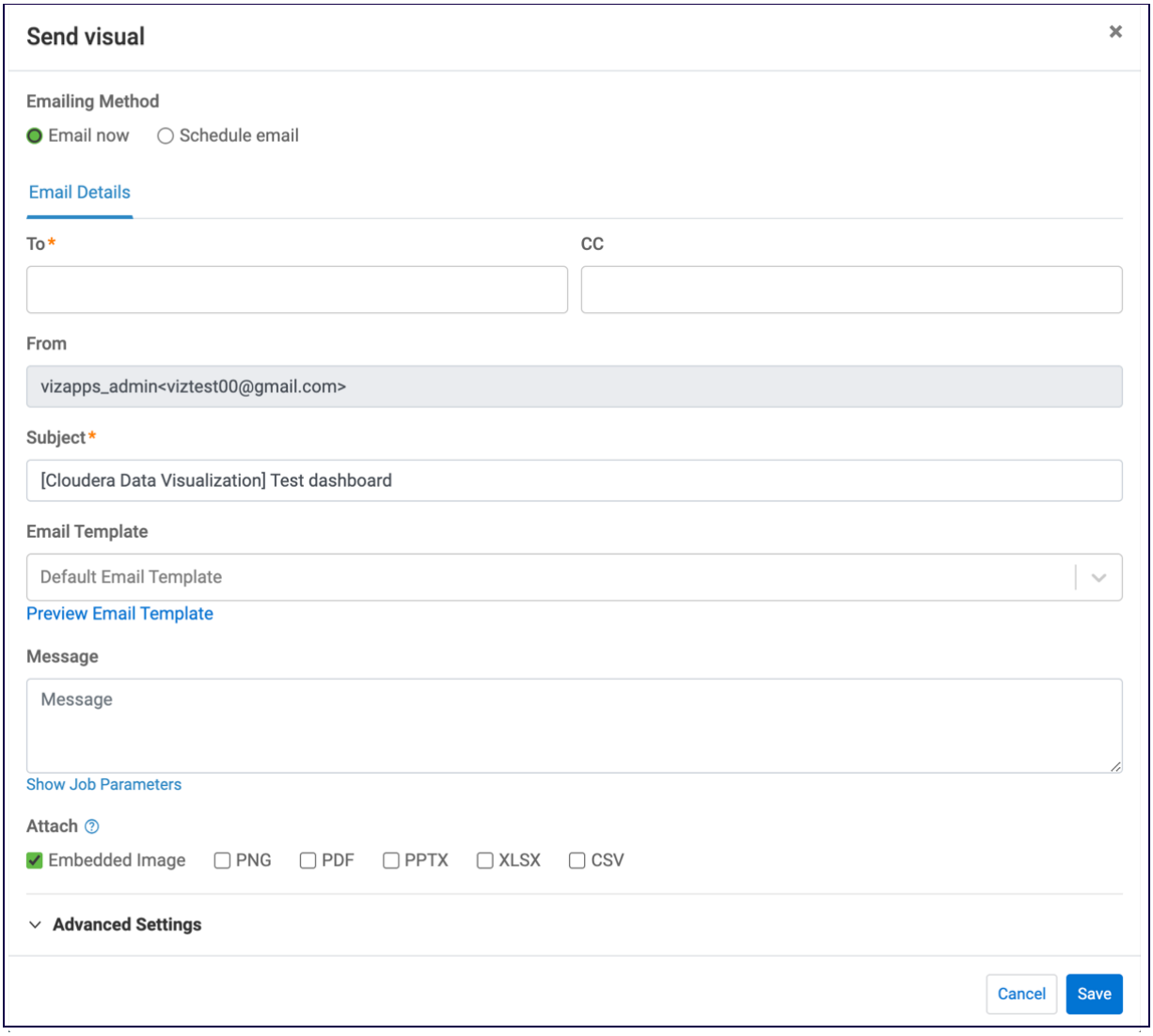
2. Click the ••• icon to open the supplemental menu in the top right corner of the visual and select Email from the drop-down.



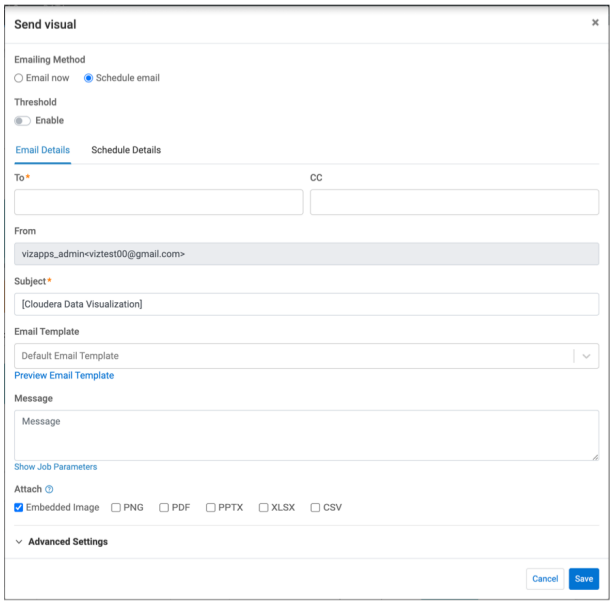
The Send visual modal window appears.

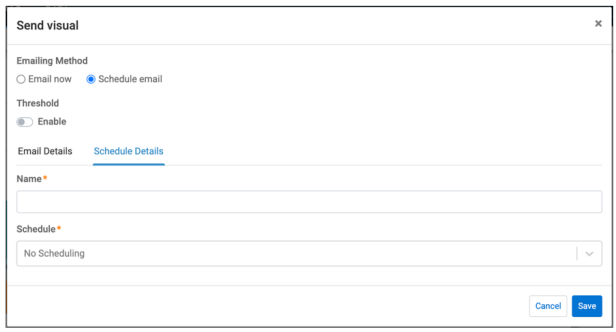
3. Choose one of the two options for sending email.

   Email now – For instructions on how to provide email details, see *Immediate email*.



Schedule email – For instructions on how to provide email and schedule details, see *Schedule-based email*.





4. After creating emails, you can monitor them in the Jobs interface, review their details, rerun, cancel, and so on. For more information, see *Managing jobs*.

Related Information

Email settings
Snapshots
Enabling scheduled jobs
Immediate email
Schedule-based email
Managing jobs

Exporting a visual

In Cloudera Data Visualization, you can export dashboards and applications, but you cannot export individual visuals on their own. Visuals are exported only as part of a dashboard or application.

Related Information

Exporting and importing visual artifacts
Exporting a dashboard
Exporting an application

Importing a visual

In Cloudera Data Visualization, you can import dashboards and applications, but you cannot import individual visuals on their own. Visuals are imported only as part of a dashboard or application.

Related Information

Exporting and importing visual artifacts
Importing a dashboard
Importing an application

Downloading a visual

Cloudera Data Visualization supports downloading visuals and their underlying data in multiple file formats for sharing, offline access, or further analysis.

Downloading a visual as a PNG file

You can download a visual as a PNG image in Cloudera Data Visualization for easy sharing or offline viewing. The downloaded file name includes the visual's name and a timestamp indicating when the file was generated.

About this task

This feature is available for all visual types. When you download a visual as a image, Cloudera Data Visualization captures it as a PNG file that includes the visual, its title, and subtitle. If the visual contains annotations, they are also included alongside the visual.

Note:

Because visuals are captured as screenshots, portions of large visuals may be truncated in the resulting PNG file.

Additional details:

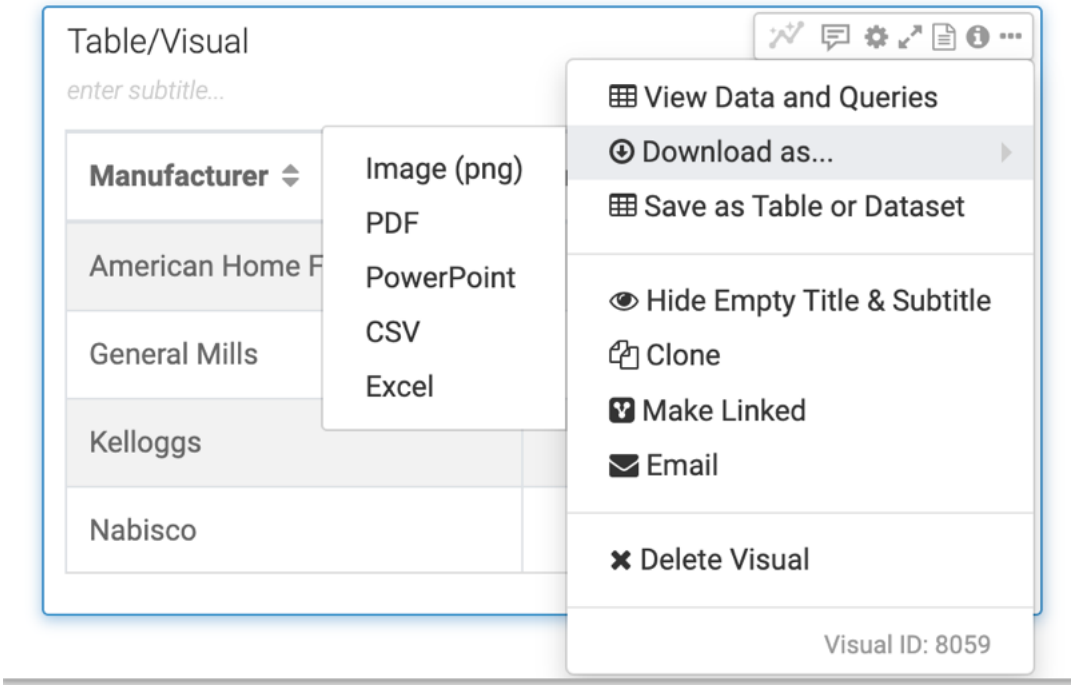
- Custom JavaScript visual snapshots are supported only only if written in ECMAScript 5.
- PNG download is not supported on Windows platforms.

Before you begin

Before downloading a visual as a PNG file, ensure the download feature is enabled in the site settings. For instructions, see Snapshots in *Managing dashboard site settings*.

Procedure

1. Click the ••• icon in the top right corner of the visual and select Download as... Image (png) .



   A message appears: Generating the requested file and the system saves the visual as a PNG file.
2. When the download is complete, open your browser's default download folder.

   In this example, the generated file name, Table_Visual_20250714093845.png, has the following components:

   - Table_Visual: Same as the visual name
   - 20250714: Date of download in YYYYMMDD format
   - 093845: Time of download in hhmmss format

   If the visual name contains characters not supported by your file system (for example slashes), Cloudera Data Visualization replaces them automatically. For example:

   - Original visual name: Table/Visual
   - Generated file name: Table_Visual_20250714093845.png (Forward slash / is replaced with an underscore _)

Downloading a visual as a PDF file

You can download a PDF snapshot of a visual in Cloudera Data Visualization for sharing or offline viewing. The downloaded file name reflects the visual's name and includes a timestamp indicating when it was created.

About this task

This feature is available for all visual types. When you download a visual as a PDF, Cloudera Data Visualization captures it as a single-page document that includes the visual, its title, and subtitle. If the visual contains annotations, they are also included on the page alongside the visual.

Note:

Because visuals are captured as screenshots, portions of large visuals may be truncated in the resulting PDF file.

Additional details:

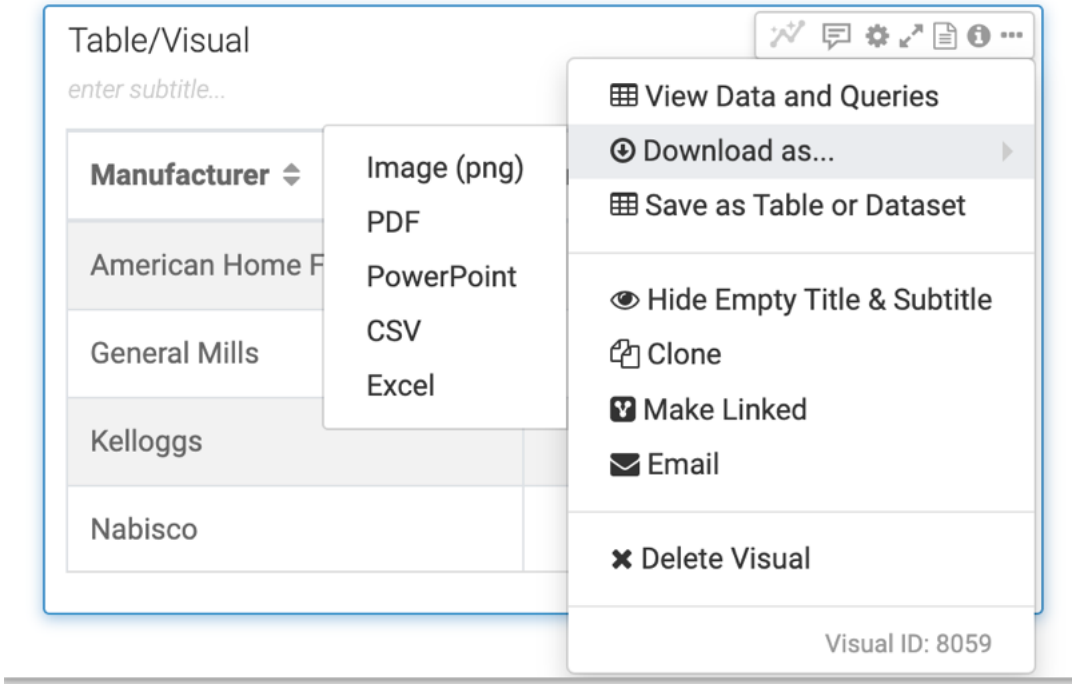
- Custom JavaScript visual snapshots are supported only only if written in ECMAScript 5.
- PNG download is not supported on Windows platforms.

Before you begin

Before you download the visual as a PDF file, make sure the download feature is enabled in site settings. For instructions, see Snapshots in *Managing dashboard site settings*.

Procedure

1. Click the ••• icon in the top-right corner of the visual and select Download as... PDF .



   A message appears: Generating the requested file and Cloudera Data Visualization saves the visual as a PDF file.
2. When the download is complete, open your browser's default download folder.

   In this example, the generated file name, Table_Visual_20250714093845.pdf, has the following components:

   - Table_Visual: Same as the visual name

- 20250714: Date of download in YYYYMMDD format
- 093845: Time of download in hhmmss format

If the visual name contains characters not supported by your file system (for example slashes), Cloudera Data Visualization replaces them automatically. For example:

- Original visual name: Table/Visual
- Generated file name: Table_Visual_20250714093845.pdf (Forward slash / is replaced with an underscore _)

Downloading a visual as a PPTX file

You can download individual visuals as PowerPoint (PPTX) files in Cloudera Data Visualization. This feature allows you to easily share insights or present visual content outside the platform.

About this task

This feature is available for all visual types. When you download a visual as a PowerPoint file, Cloudera Data Visualization generates a single slide containing the visual's title, subtitle, and a screenshot of the visual. If annotations are available, they are included alongside the visual on the slide.

Note:

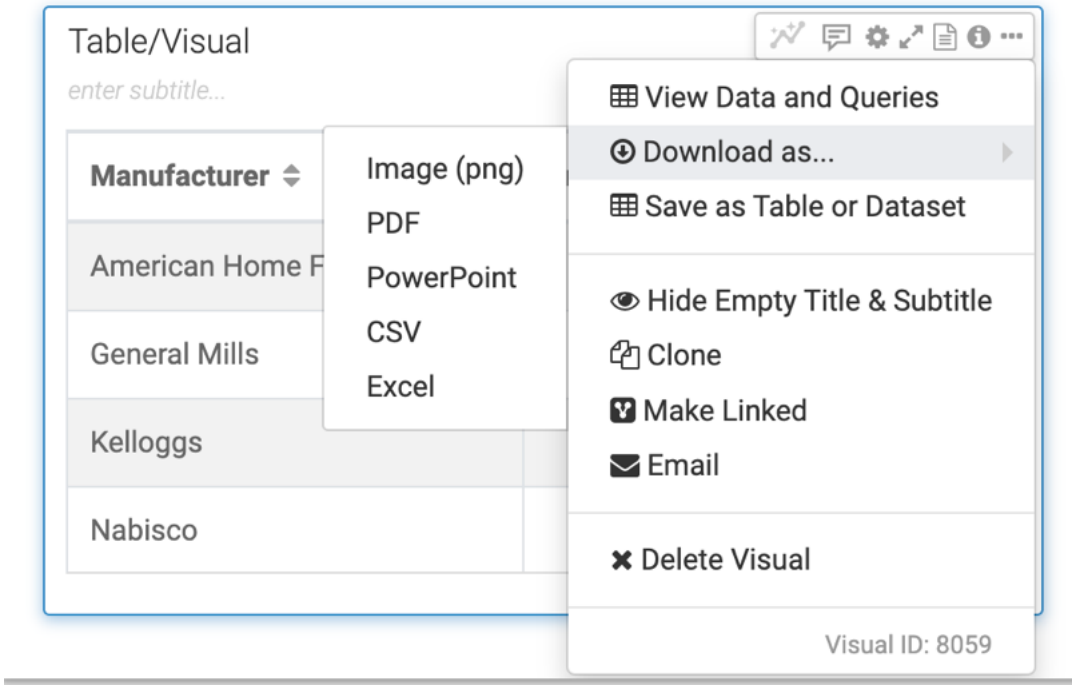
Visuals are captured as screenshots, so very large visuals may be truncated in the PowerPoint slides.

Before you begin

Before downloading the visual as a PPTX file, ensure the download feature is enabled in the site settings. For instructions, see Snapshots in *Managing dashboard site settings*.

Procedure

1. Click the ••• icon in the top-right corner of the visual and select Download as... PowerPoint .



A message appears: Generating the requested file and Cloudera Data Visualization saves the visual as a PPTX file.

2. When the download is complete, open your browser's default download folder to locate the generated PowerPoint file.

   Example:

   - Generated file name: Table_Visual_20251026103845.pptx
   - File name components:

     a. Table_Visual: Same as the visual name
     b. 20251026: Date of download in YYYYMMDD format
     c. 103845: Time of download in hhmmss format

   If the visual's title or subtitle contains unsupported characters, Cloudera Data Visualization replaces them automatically.

   The downloaded PowerPoint file has a single slide that includes:

   - Title and subtitle of the visual
   - Annotations, when present, displayed next to the visual

     Note: There is no option to prevent the annotation from being included by default, but they can be edited or removed in the Powerpoint file

   - Username and timestamp for the download

Downloading a visual as a CSV file

You can save the underlying data used in a visual to a CSV file in Cloudera Data Visualization. This is useful for further analysis or offline data sharing. The downloaded file name reflects the visual's name and includes a timestamp indicating when it was created.

About this task

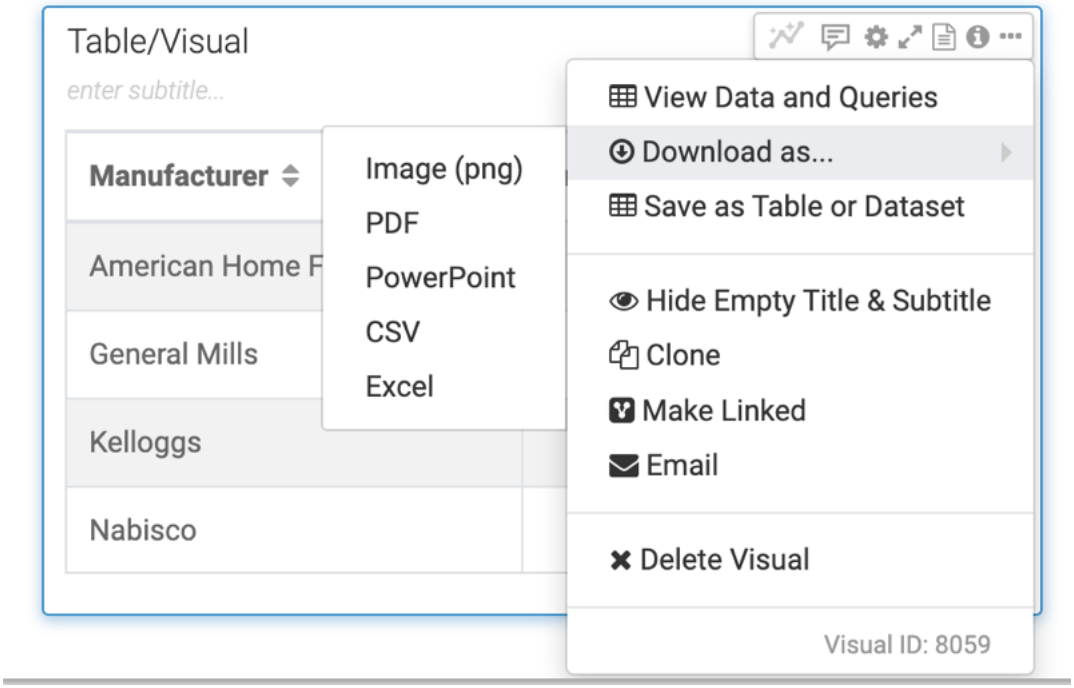
This feature is available for all visual types. When you download a visual as a CSV file, Cloudera Data Visualization captures it as a single-page document that includes the visual. The downloaded CSV file preserves display formatting, allowing you to sort and calculate data immediately, without additional cleanup.

Before you begin

Before you download the visual as CSV file, make sure the Enable "Download as CSV/Excel" by default feature is enabled in site settings. For instructions, see Visuals in *Managing dashboard site settings*.

Procedure

1. Click the ••• icon in the top-right corner of the visual and select Download as... CSV .



A message appears: Generating the requested file and Cloudera Data Visualization saves the visual to a CSV file.

2. When the download is complete, open your browser's default download folder.

   In this example, the generated file name, Table_Visual_20250714093845.csv, has the following components:

   - Table_Visual: Same as the visual name
   - 20250714: Date of download in YYYYMMDD format
   - 093845: Time of download in hhmmss format

   If the visual name contains characters not supported by your file system (for example slashes), Cloudera Data Visualization replaces them automatically. For example:

   - Original visual name: Table/Visual
   - Generated file name: Table_Visual_20250714093845.csv (Forward slash / is replaced with an underscore _)

Downloading a visual as an Excel file

You can save data from a visual into an Excel (XLSX) file in Cloudera Data Visualization. The downloaded file includes both the data displayed in the visual and a separate sheet listing any defined parameters. The file name reflects the visual's name and includes a timestamp showing when it was created.

About this task

This feature is available for all visual types. When you download a visual as an Excel file, Cloudera Data Visualization exports the visual's underlying data in a structured format suitable for further analysis.

The downloaded Excel file:

- Preserves display formatting, allowing you to sort, filter, and calculate data immediately, without additional cleanup
- Exports numbers as numeric values with applied formatting (for example, currency or percentage), rather than as plain text
- Safely handles string values that begin with = by exporting them as text to prevent unintended formula evaluation in Excel.

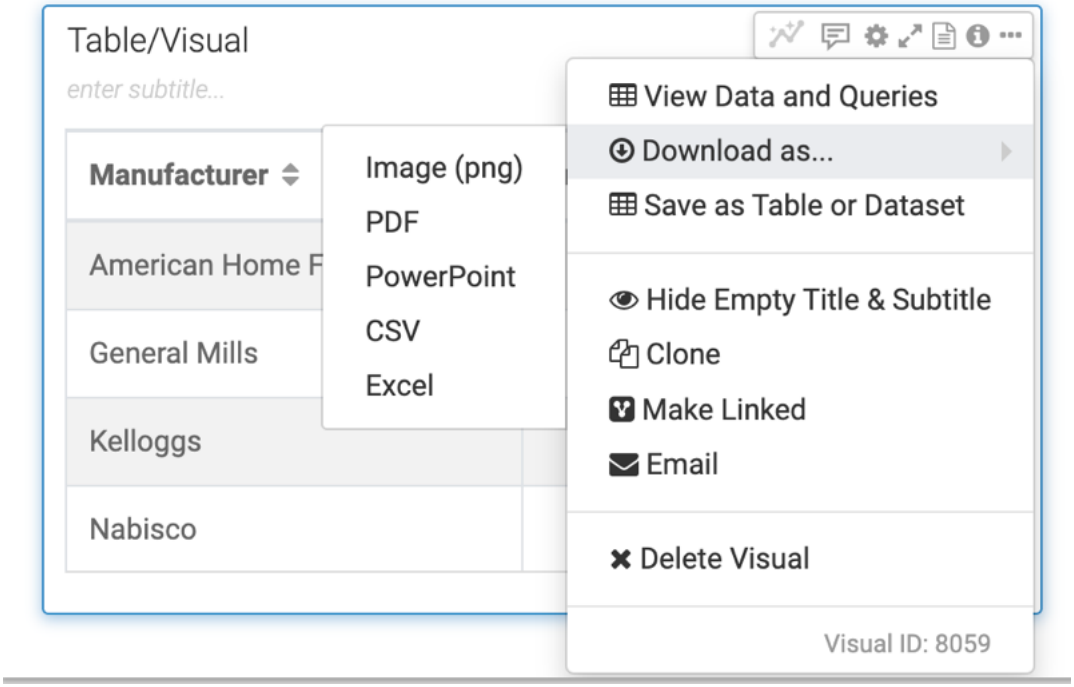
- Includes column totals, if totals are enabled in the visual. To enable totals in the visual, see *Showing column totals*.

Before you begin

Before downloading a visual as an Excel file, ensure that the Enable "Download as CSV/Excel" by default option is turned on in the site settings. For instructions, see Visuals in *Managing dashboard site settings*.

Procedure

1. Click the ••• icon in the top-right corner of the visual and select Download as... Excel .



   A message appears: Generating the requested file and Cloudera Data Visualization exports the visual to an XLSX file.
2. When the download is complete, open your browser's default download folder and locate the generated Excel file.

   In this example, the generated file name, Table_Visual_20251026150859.xlsx includes the following components:

   - Table_Visual: Same as the visual name
   - 20251026: Date of download in YYYYMMDD format
   - 150859: Time of download in hhmmss format

   Totals appear in the table only if they are enabled, as described in Showing column totals. If the visual includes filters, you can view the Params sheet in the Excel file to see the parameter definitions that were applied when the visual was created.

Related Information

Showing column totals

Removing a visual from a dashboard

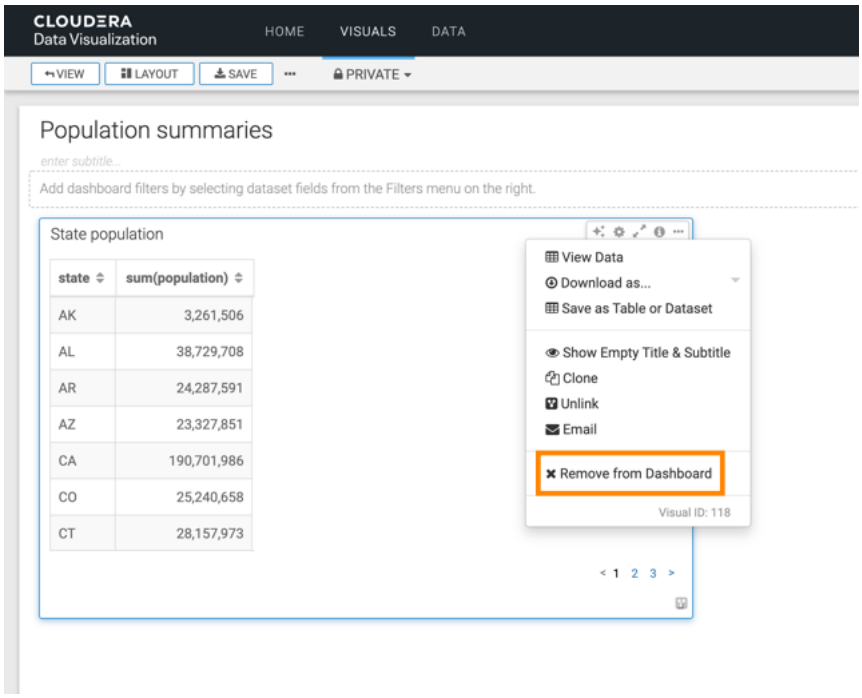
In Cloudera Data Visualization, you can remove a linked visual from a dashboard without deleting it from the system. This is useful when you no longer need the visual on the current dashboard but want to keep it available for reuse.

About this task

This section explains how to use the Remove action. Use Remove for a linked visual if you want to remove it only from the current dashboard and keep it available for use on other dashboards. If you want to remove a visual permanently from the system, use the Delete action. For more information, see *Deleting visuals*.

Procedure

1. Open the dashboard in Edit mode in the Dashboard Designer.
2. Click the (ellipsis) icon at the top right corner of the linked visual.
3. Select Remove from Dashboard.



Deleting a visual

In Cloudera Data Visualization you can delete an unlinked visual from the system.

About this task

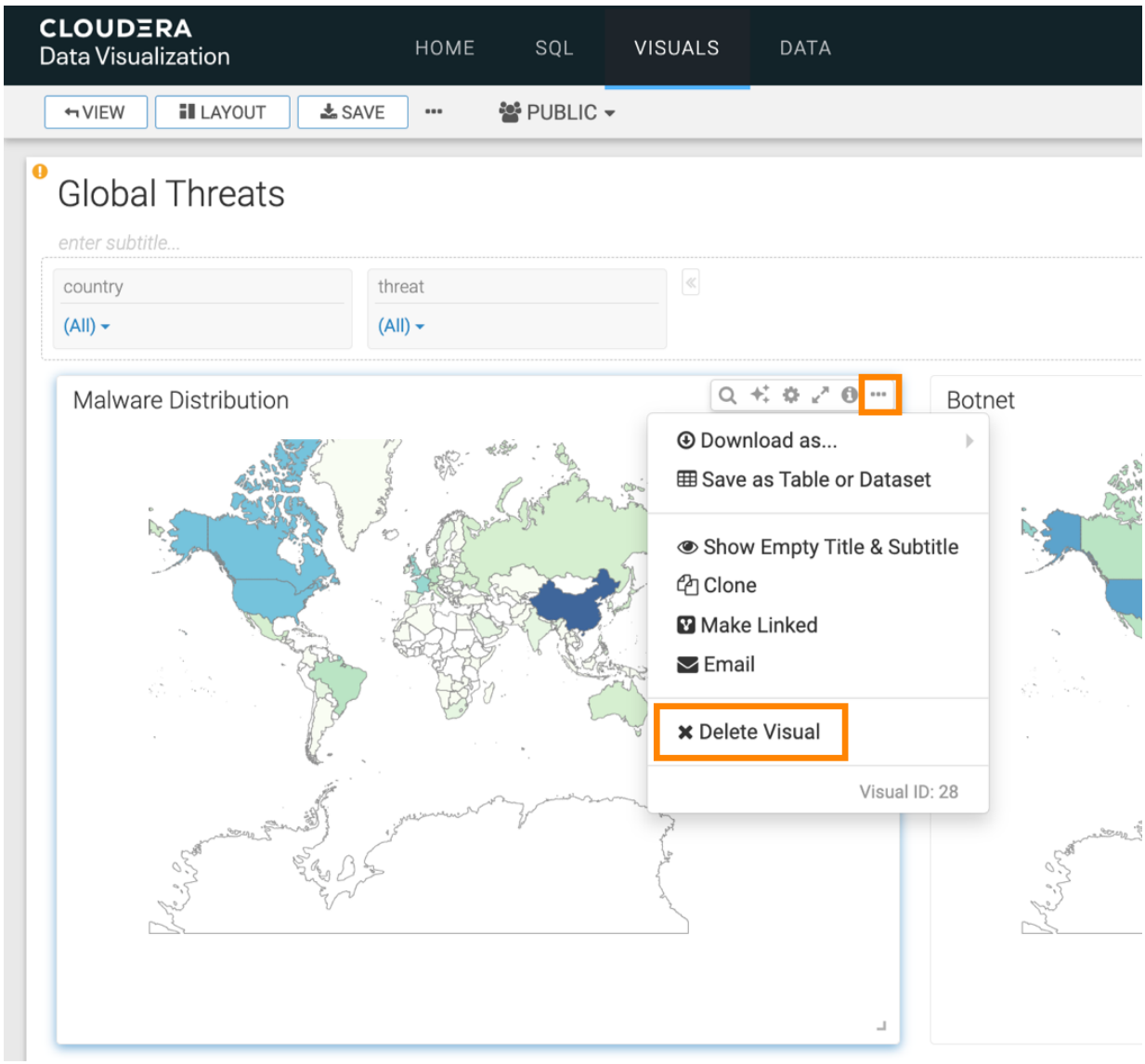
These instructions show you how to use the Delete action. Use Delete for unlinked visuals to permanently remove from the system. If you want to remove a visual only from the current dashboard and keep it available for use on other dashboards, use the Remove action. For more information, see *Removing a visual*.

Before you begin

To delete a visual, you must have Edit access to the dashboard that contains it.

Procedure

1. Open the dashboard in Edit mode.
2. Hover over the visual you want to delete and click the Supplemental menu in the top-right corner the visual.
3. Click Delete Visual.



Adding filters to a visual

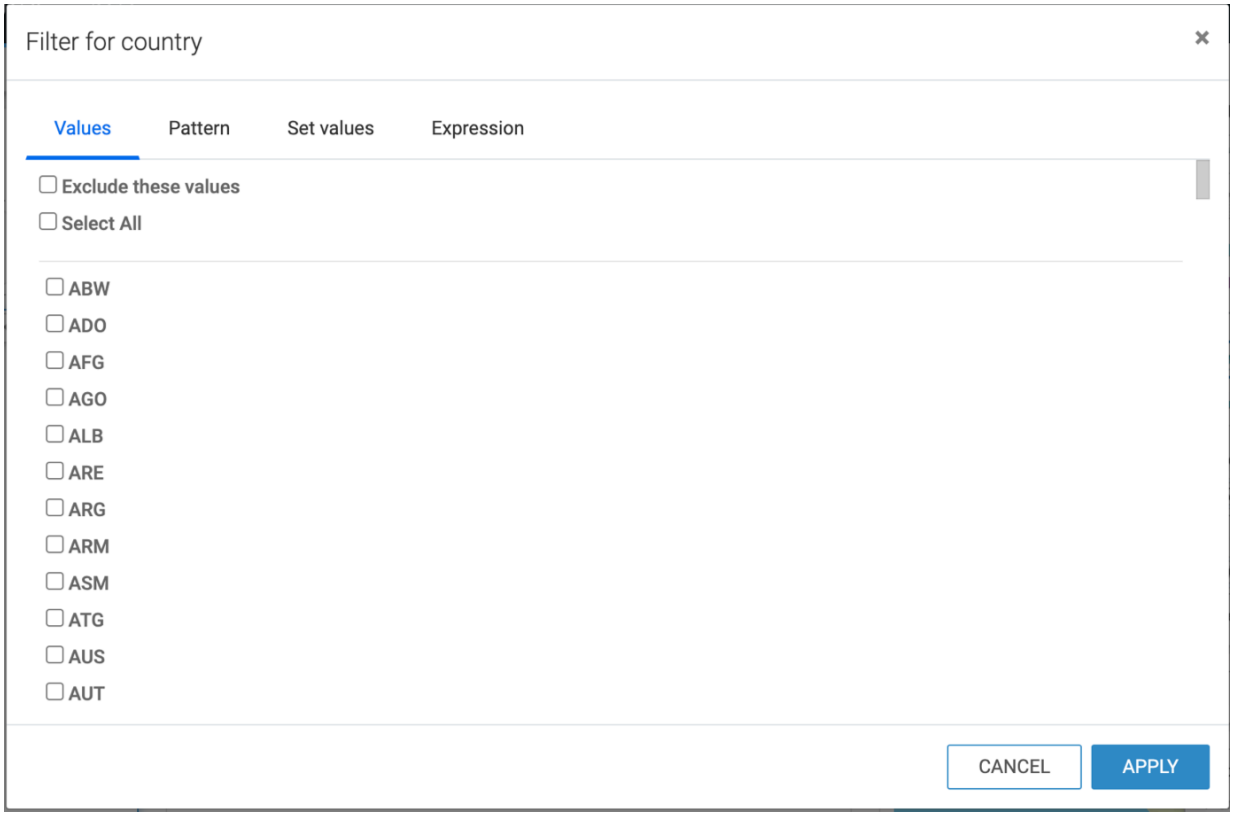
Cloudera Data Visualization enables you to add filters to a visual if the visual has a Filters shelf available.

Procedure

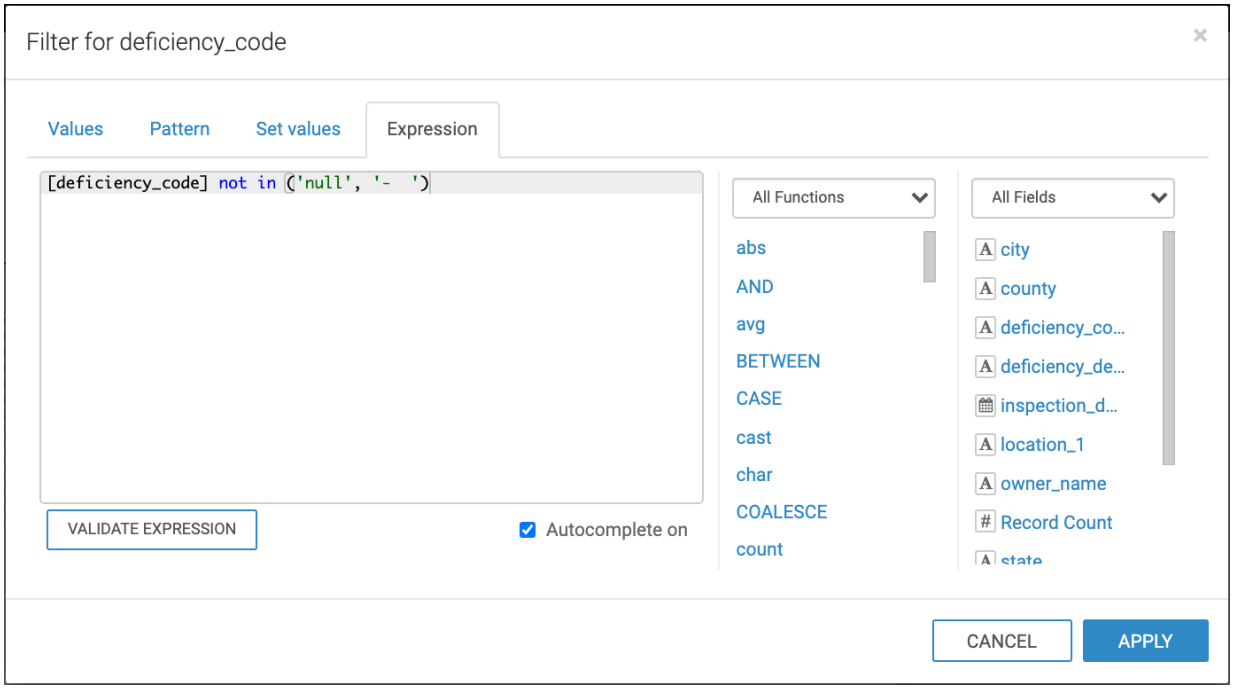
1. Open a visual of your choice in Edit mode.

   In this example, a visual called Deficiency date correlation overview is used.
2. Populate the Filters shelf from the available fields (Dimensions, Measures) in the DATA menu.

   In this example, the country field is added to the filter shelf. A modal window pops up where you can set the values you want to filter for. You can choose discrete values, ranges, patterns, or use the Select All option for a bulk selection of the shown filter values.



You can also create your own custom filter expressions by using the Expression tab. In this example, the deficiency_code field is added to the filter shelf with a custom filter expression.



You can use the filter with its default settings, or you can open the Settings menu from the side menu bar to configure your filter.

Creating annotation for visuals

Annotations are brief summaries or description that you can associate with a visual or dashboard. You can use them to clarify data details, highlight outlier values, or provide additional context for visual consumers. The annotations feature supports both manually written descriptions and AI-generated summaries based on the data displayed in a visual.

How annotations are displayed

If a visual has no annotation:

- In View mode, the Annotation icon ( ) is hidden.
- In Edit mode, the Annotation icon is visible. Clicking it displays a message indicating that no annotations have been added yet, and provides the Edit option for adding an annotation.

If a visual has an annotation:

- In View mode, the Annotation icon is visible. Clicking it or hovering over it displays the annotation.
- In Edit mode, the Annotation icon is visible. Clicking it displays the annotation, and provides the Edit option for editing the annotation.

Before you begin

- You can create or edit annotations only in Edit mode.
- To use AI-powered summarization, enable AI features and AI summary in Site Settings. If these are not enabled, only manual annotations can be created. For more information on these settings, see Managing AI settings.

Creating an AI-powered annotation

Cloudera Data Visualization allows you to generate AI-powered summaries directly from your dashboard visuals. These summaries can help you quickly identify patterns, trends, and outliers without manually analyzing each visual. Using generative AI, the system automatically analyzes the underlying data and produces concise, insightful summaries that make complex visuals easier to interpret.

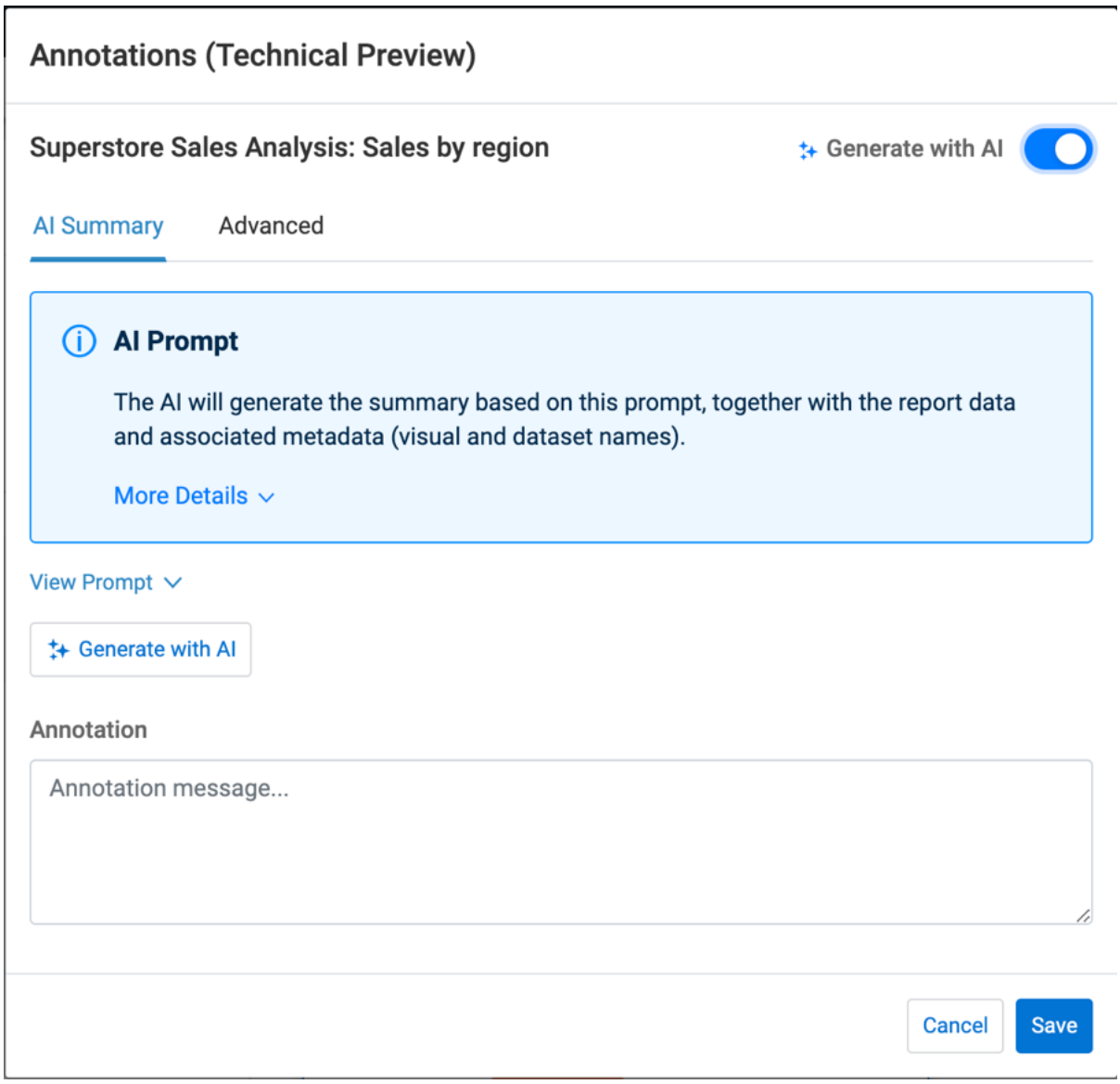
1. Open a dashboard in Edit mode.
2. Click in the top-right corner of the visual you want to annotate and click Edit.

   The Annotations modal is displayed.
3. Enable the Generate with AI toggle.

   Two tabs are displayed:

   - AI Summary – where you can configure your prompt and view your AI-generated annotation
   - Advanced – where you can review metadata, dataset information, and the full AI response
4. Click View Prompt to review the default AI prompt text.

   You can edit the suggested prompt or restore the default version if needed.



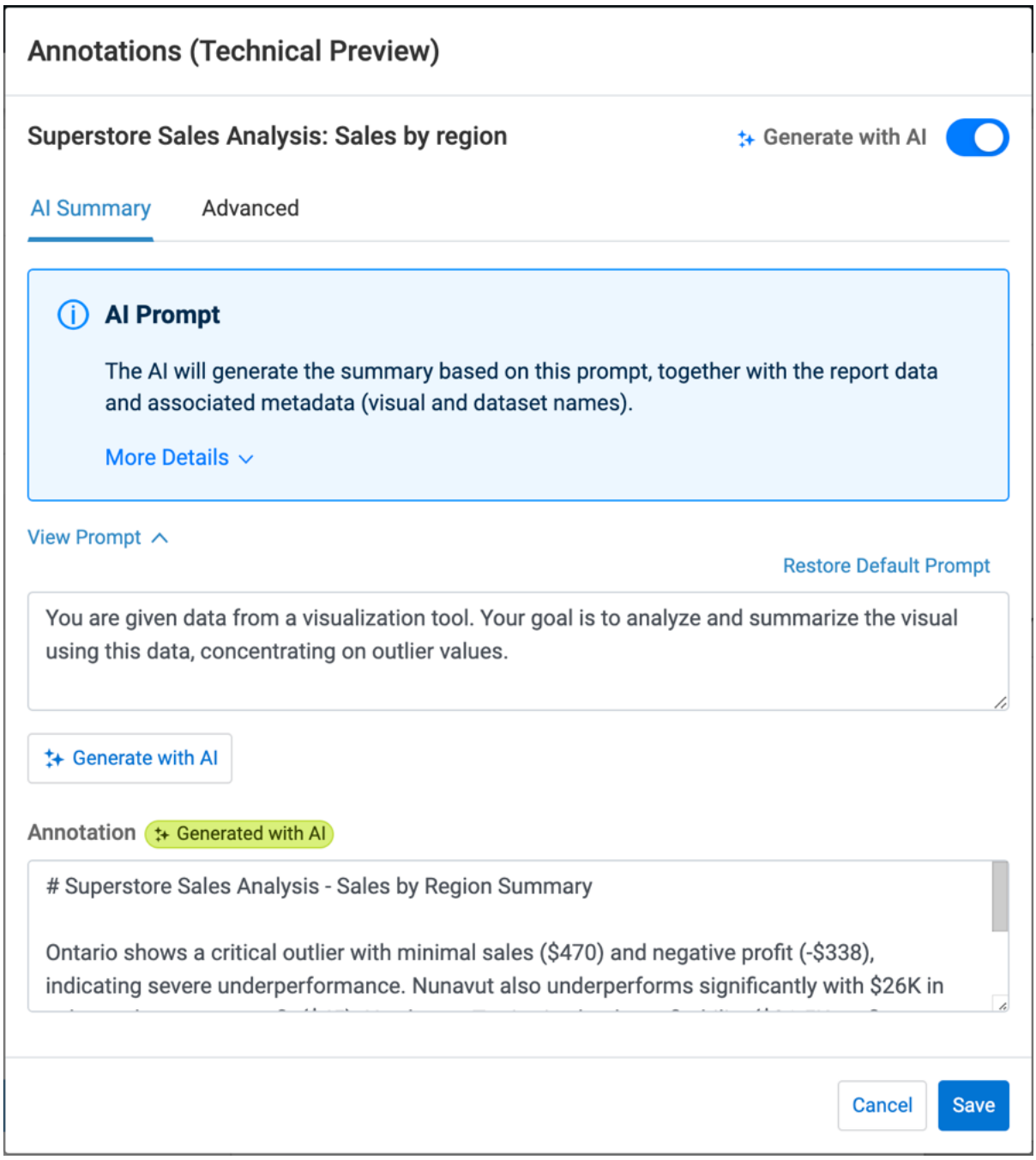
5. Click Generate with AI to run the prompt against the visual's dataset.

   The AI processes the data, generates a short summary in the Annotation field, and adds a Generated with AI label.

   Important:

   AI-generated content may be inaccurate or misleading. Always review and verify the output.

   If the dataset contains sensitive fields, a warning appears. Sensitive fields are not sent to the AI model, which may affect the annotation's accuracy.



6. Review the generated annotation for accuracy.

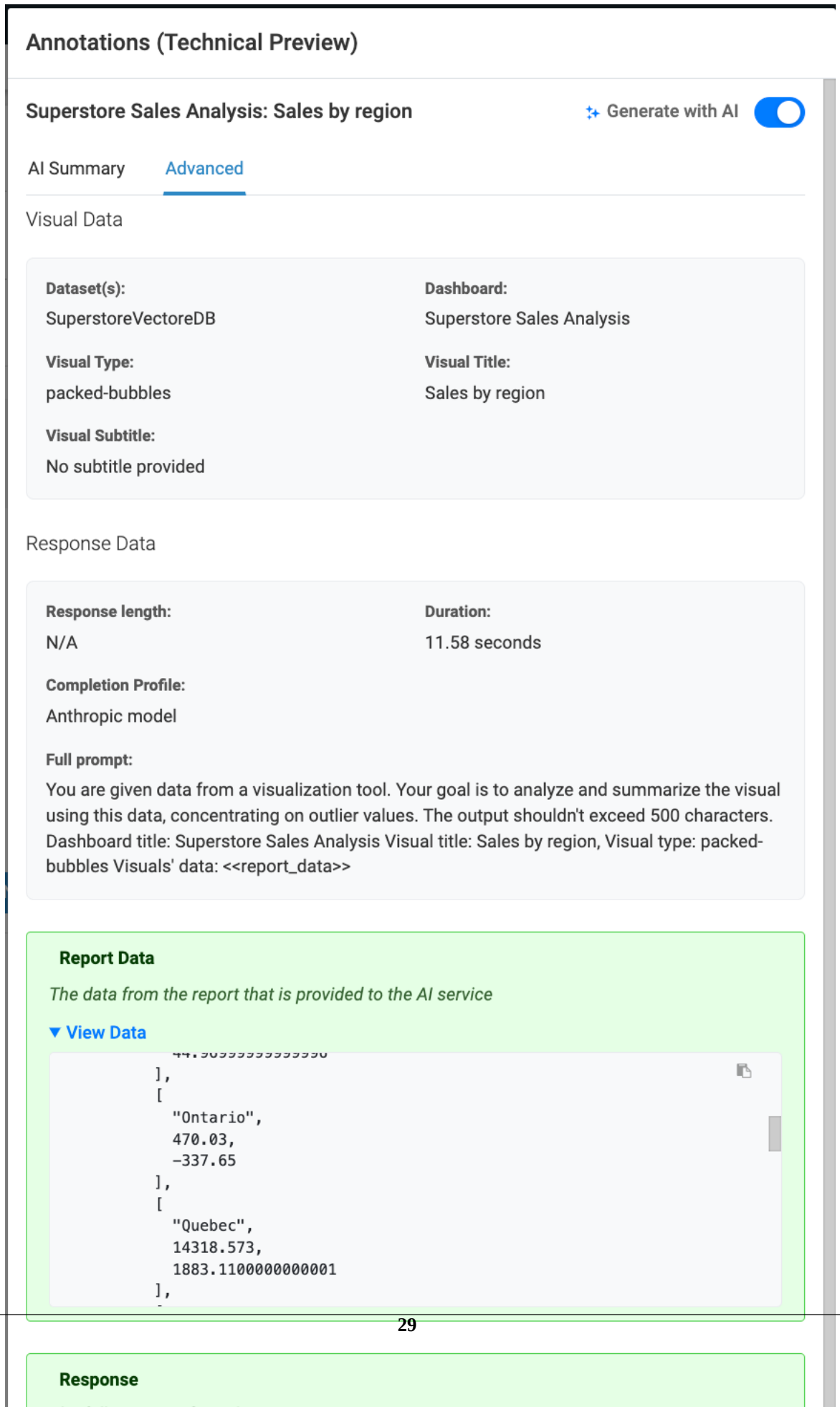
   You can edit the text before saving, ensuring that the clarity and tone of the annotation meet your needs.

   Note: If you edit the text, the Generated with AI label is removed.

7. [Optional] Review technical details.

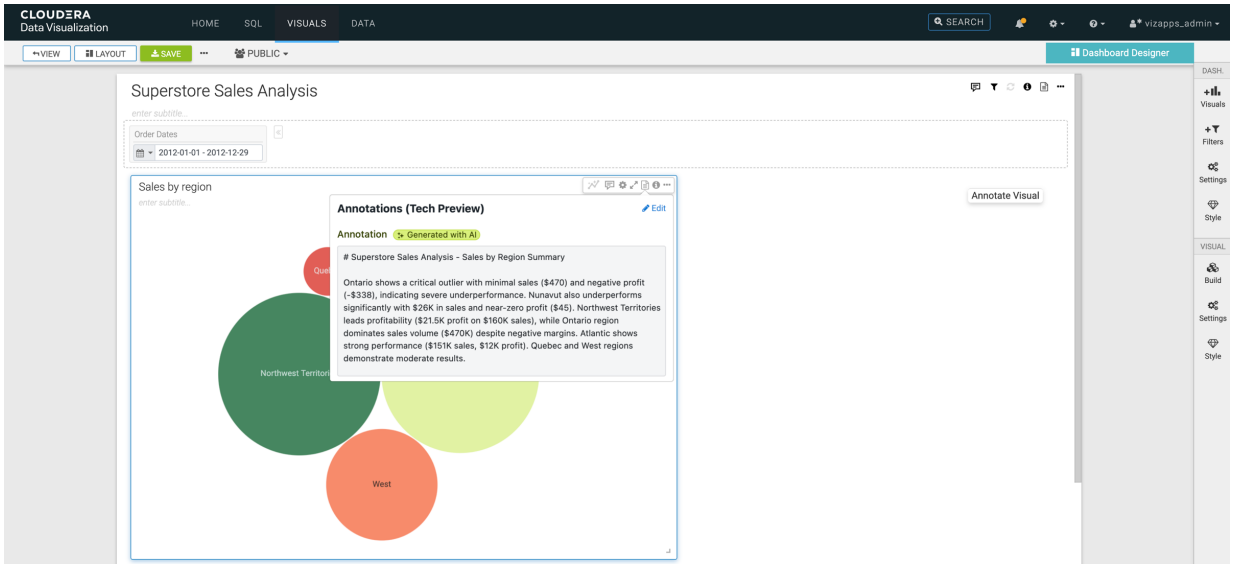
   Switch to the Advanced tab to see the underlying details of the generation process, useful for validation or debugging.

   - Visual Data – Dataset, dashboard, and visual metadata
   - Response Data – Full prompt, report data, and AI response details


Annotations (Technical Preview)
Superstore Sales Analysis: Sales by region
Generate with AI
AI Summary
Advanced
Visual Data
Dataset(s):
SuperstoreVectoreDB
Dashboard:
Superstore Sales Analysis
Visual Type:
packed-bubbles
Visual Title:
Sales by region
Visual Subtitle:
No subtitle provided
Response Data
Response length:
N/A
Duration:
11.58 seconds
Completion Profile:
Anthropic model
Full prompt:
You are given data from a visualization tool. Your goal is to analyze and summarize the visual using this data, concentrating on outlier values. The output shouldn't exceed 500 characters. Dashboard title: Superstore Sales Analysis Visual title: Sales by region, Visual type: packed-bubbles Visuals' data: <<report_data>>
Report Data
The data from the report that is provided to the AI service
▼ View Data
],
[
"Ontario",
470.03,
-337.65
],
[
"Quebec",
14318.573,
1883.1100000000001
],
Response

8. Click Save.

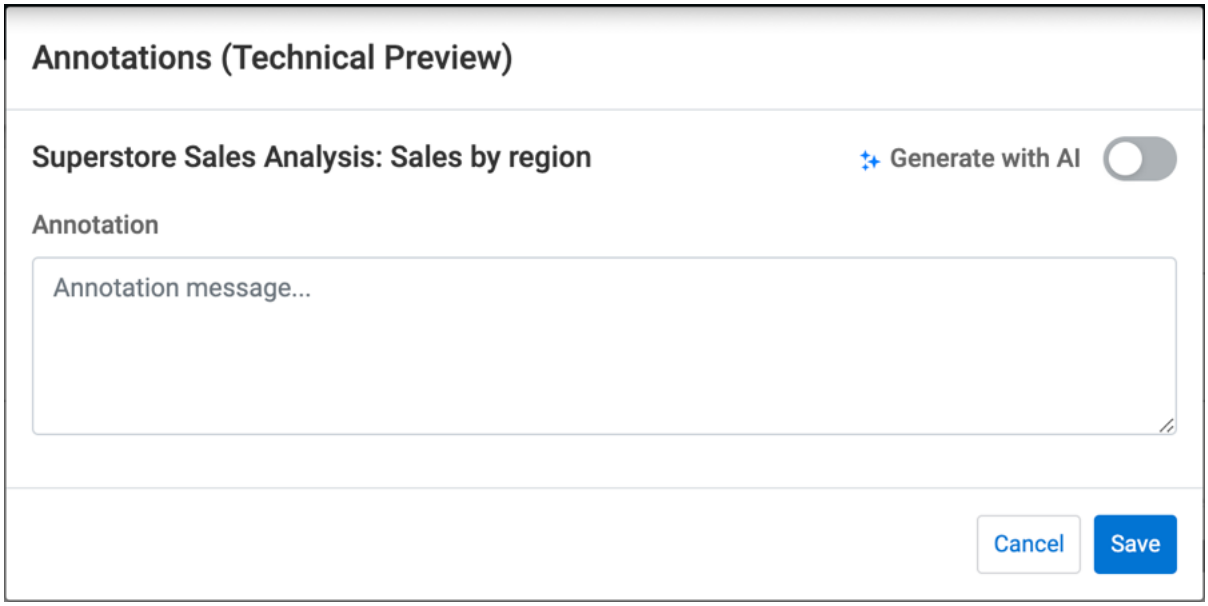
   The annotation is attached to your visual.



Creating a manual annotation

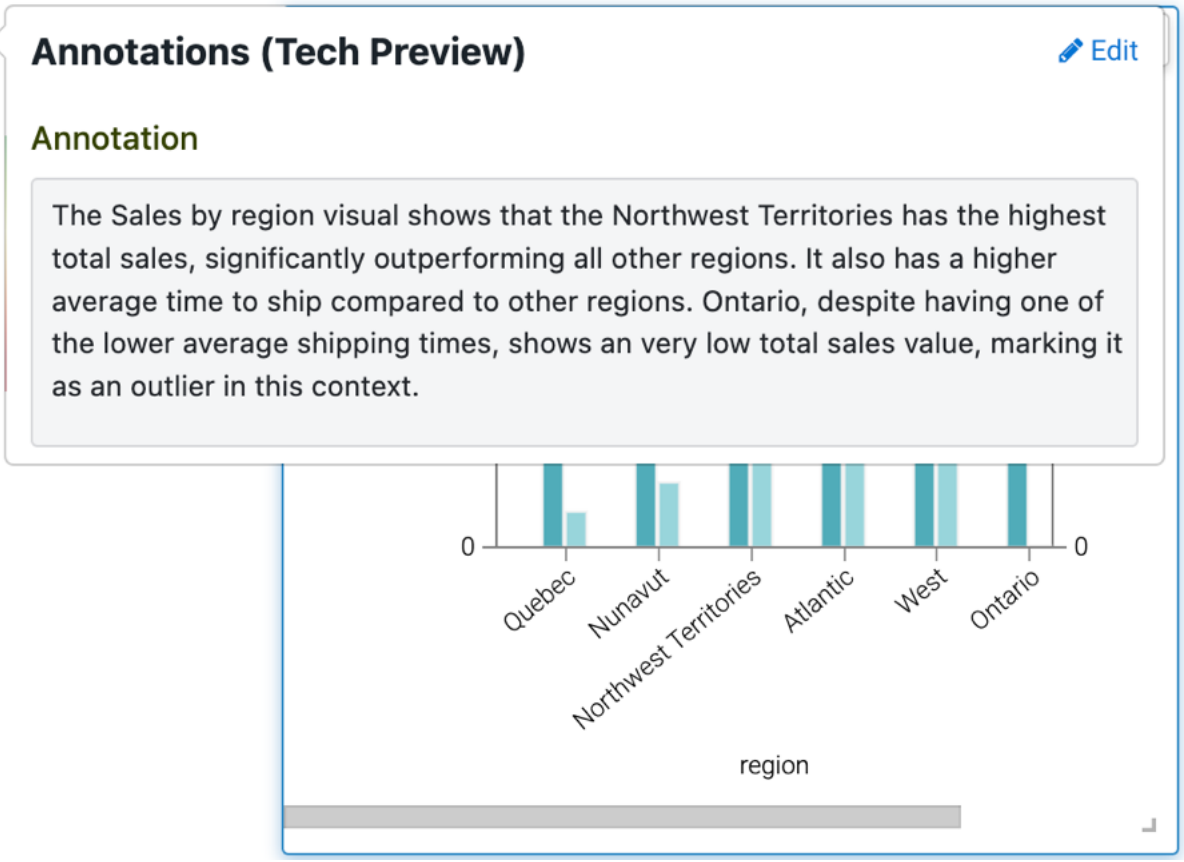
1. Open a dashboard in Edit mode.
2. Click in the top-right corner of the visual that you want to annotate and click Edit.

   The Annotations modal is displayed.



3. Enter your annotation text in the Annotation field.
4. Click Save.

   The annotation is attached to your visual.



Commenting on visuals and dashboards [Technical Preview]

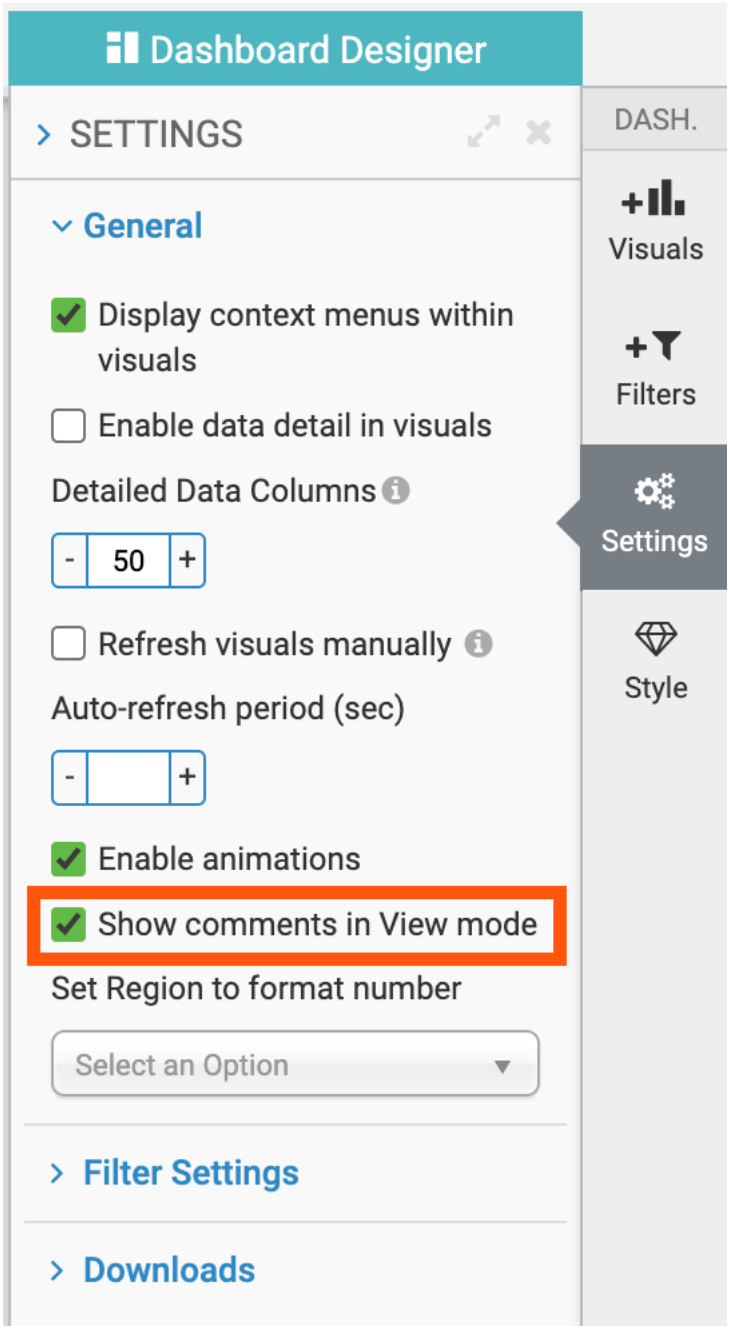
The Comments panel in Cloudera Data Visualization allows you to add feedback or insights and ask questions directly within dashboards and visuals. Comments can be added at either the dashboard level or the visual level, making it easier to share context-specific insights.

About this task

Commenting makes collaboration with other users easier by keeping discussions in context, without switching to external tools. Each comment includes the author and a timestamp, ensuring traceability and accountability. You can reply to the comments added by other users, so discussion threads keep conversations organized and easy to follow.

By default, commenting on visuals and dashboards is disabled. To use this feature, an administrator must first enable it in Site Settings. For details, see Enable commenting.

When commenting is enabled, you can see and add comments in View mode or Edit mode. You can turn off showing comments in View mode in the general settings for dashboards. When comments are hidden, you cannot see existing comments or add new ones in View mode.



Viewing comments

The Comments panel displays all comments associated with the selected dashboard or visual. If no comments exist for the dashboard or the visual, the panel shows No Comments Yet.

To review comments for a dashboard:

1. Open the dashboard in View or Edit mode.
2. Click located in the top-right corner of the dashboard to open the Comments panel.

   By default, dashboard-level comments are shown.

   a. To comment on the dashboard, click located at the top-right corner of the dashboard.

b. To comment on a specific visual, hover over the visual's top-right corner and click .

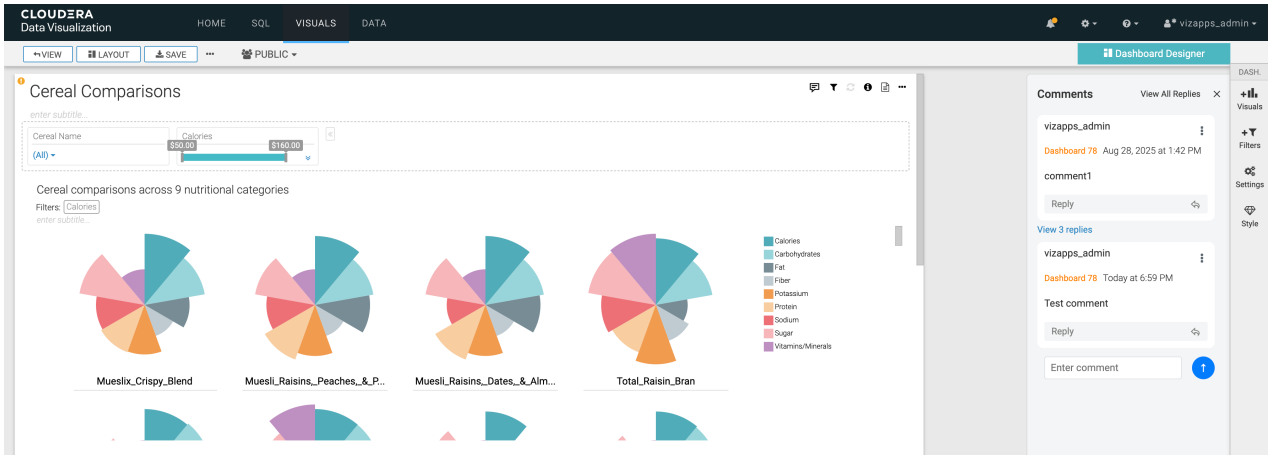
3. [Optional] Click Filter to open the dropdown filter menu and choose what comments you want to display.

   Setting the filter to All comments provides a complete overview of both dashboard-level and visual-level comments. In this view, each comment is clearly labeled to indicate whether it applies to the dashboard or to a specific visual.

To review comments for a specific visual:

Hover over the visual's top-right corner and click , or the dashboard-level Comments panel filter to select the specific visual. Both methods display only the comments that were added to the selected visual.

The Comments panel remains open until you manually close it.



Filtering comments at dashboard level

The dashboard Comments panel provides a consolidated view of all feedback associated with a dashboard. It displays both dashboard-level comments and visual-level comments from the visuals included in that dashboard.

A comments filter allows you to control which comments are shown, helping you focus on the most relevant feedback.

Using the comments filter

A dropdown filter control is available at the top of the dashboard Comments panel. You can choose from the following options:

Dashboard comments

Shows only comments that apply to the entire dashboard.

Comments for specific visuals

Allows you to select one or more visuals whose comments you want to view. Comments are grouped by the visual they belong to, showing the visual's ID.

Dashboard comments and comments for specific visuals

Displays dashboard-level comments together with comments for the selected visuals in a combined view.

All comments

Shows all comments associated with the dashboard, including comments added directly to visuals.

Clear

Allows you to clear all selections and reset the view.

The filter selection determines which comments appear in the dashboard Comments panel.

Adding a comment to a dashboard or a visual

1. Open the Comments panel from the dashboard or the visual.
2. Type your message in the Enter comment field.

   When you start typing, the comment field expands into a larger text area, and a short markdown reference appears below the comment box, showing common formatting options. You can use basic markdown syntax (for example: **bold**, _italic_, `code`, - bullet list item) to format your comment text. You can use this formatting to highlight important information, add emphasis, or present short code snippets.

   Tip: Press Shift+Enter to insert a new line within the same comment without posting it.

   Note: Other markdown elements may render, but comments are designed for short notes with basic emphasis, lists, and inline code. Images and complex layouts are not supported.

3. Click to submit the comment.

   The comment is displayed. If you used markdown, the markdown formatting is applied.

Each added comment shows:

- Author's username
- Related context (dashboard or visual ID)
- Timestamp

Replying to comments

1. Click Reply below the comment you want to respond to.
2. Enter your response in the text box.

   Markdown support is available. For more information, see the *Adding a comment to a dashboard or a visual* section.

3. Click to post your reply.

- By default, only the most recent reply is shown in each comment thread.
- If a comment has multiple replies, click View X replies to expand the thread and view all replies, or click Hide X replies to collapse the thread.
- To expand or collapse all comment threads at once, click View All Replies or Hide All Replies.

Managing comments

From the # (more options) menu next to a comment, you can:

- Edit – Update the content of your own comment.
  - Comments and replies show an Edited badge and updated timestamp when modified, making it clear when feedback has changed.
  - Markdown support is available when writing comments. When editing a formatted comment, you see and edit the raw markdown syntax in the textarea. For more information on using Markdown in comments, see the *Adding a comment to a dashboard or a visual* section.
- Mark as resolved – Mark the selected comment as resolved.
  - Only that specific comment is marked as resolved. Replies remain open and can be resolved individually.
  - Resolved comments cannot be edited or deleted, but they can be reopened.

    - Resolved comments are automatically moved to the bottom of the thread, keeping active discussions visible. If a resolved comment is reopened, it returns to its original position.
- Copy comment link – Copy a direct link to a comment or reply in a comment thread. Pasting this link in your browser opens the dashboard, automatically displays the comments panel, and highlights the specific comment or reply.
- Delete – Permanently remove a comment.
    - When you delete a comment, all replies are deleted as well.

Important: Only the comment owner or administrators can edit or delete comments.

Exploring visual options

Cloudera Data Visualization helps you to develop valuable insights into your data by suggesting new visual types and options based on the information you already use in your existing visuals.

Before you begin

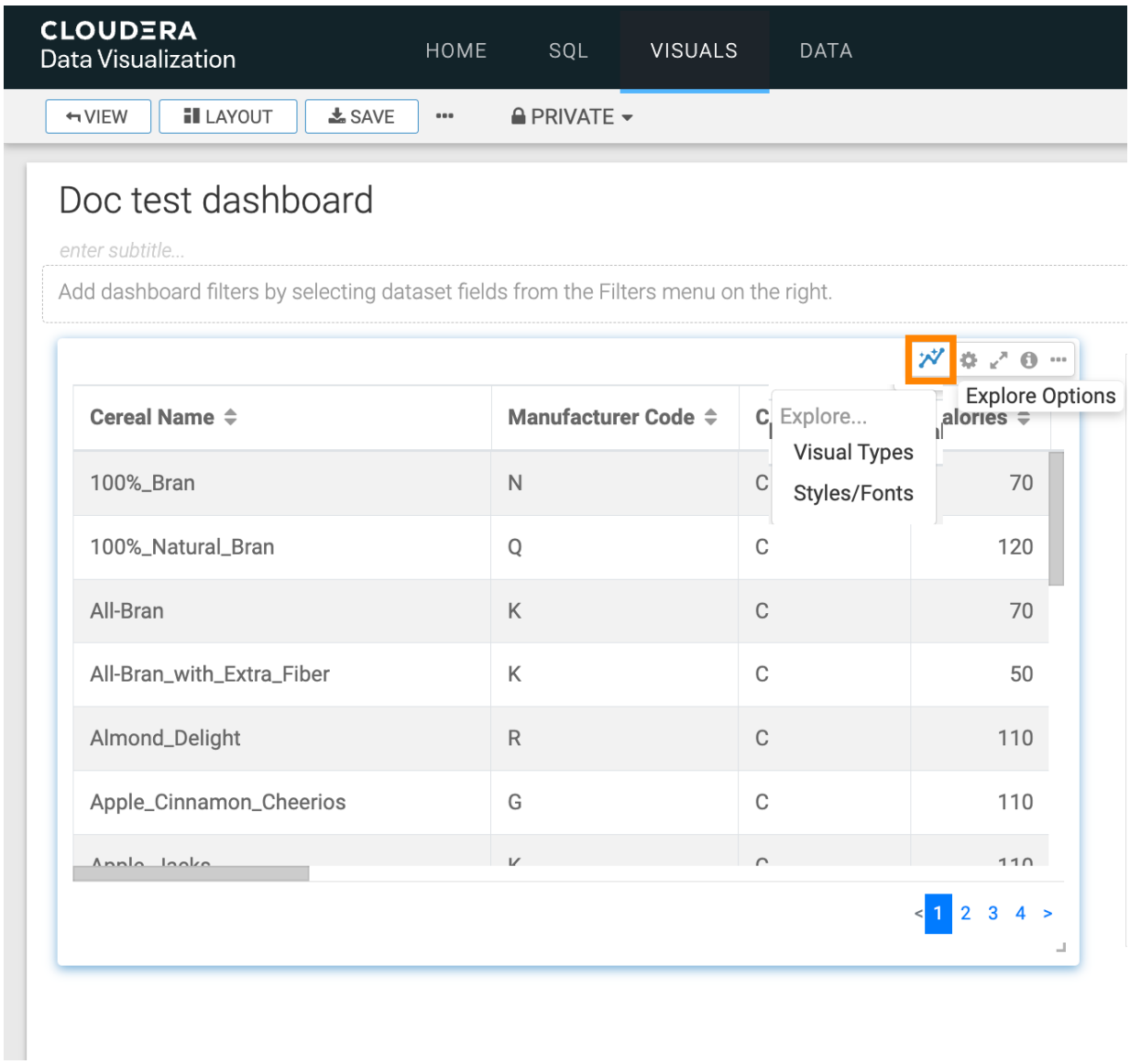
To view visual type and option recommendations, you must activate Enable Explore in the Site Settings Dashboards Visuals menu. For more information, see *Enabling visual options*.

Procedure

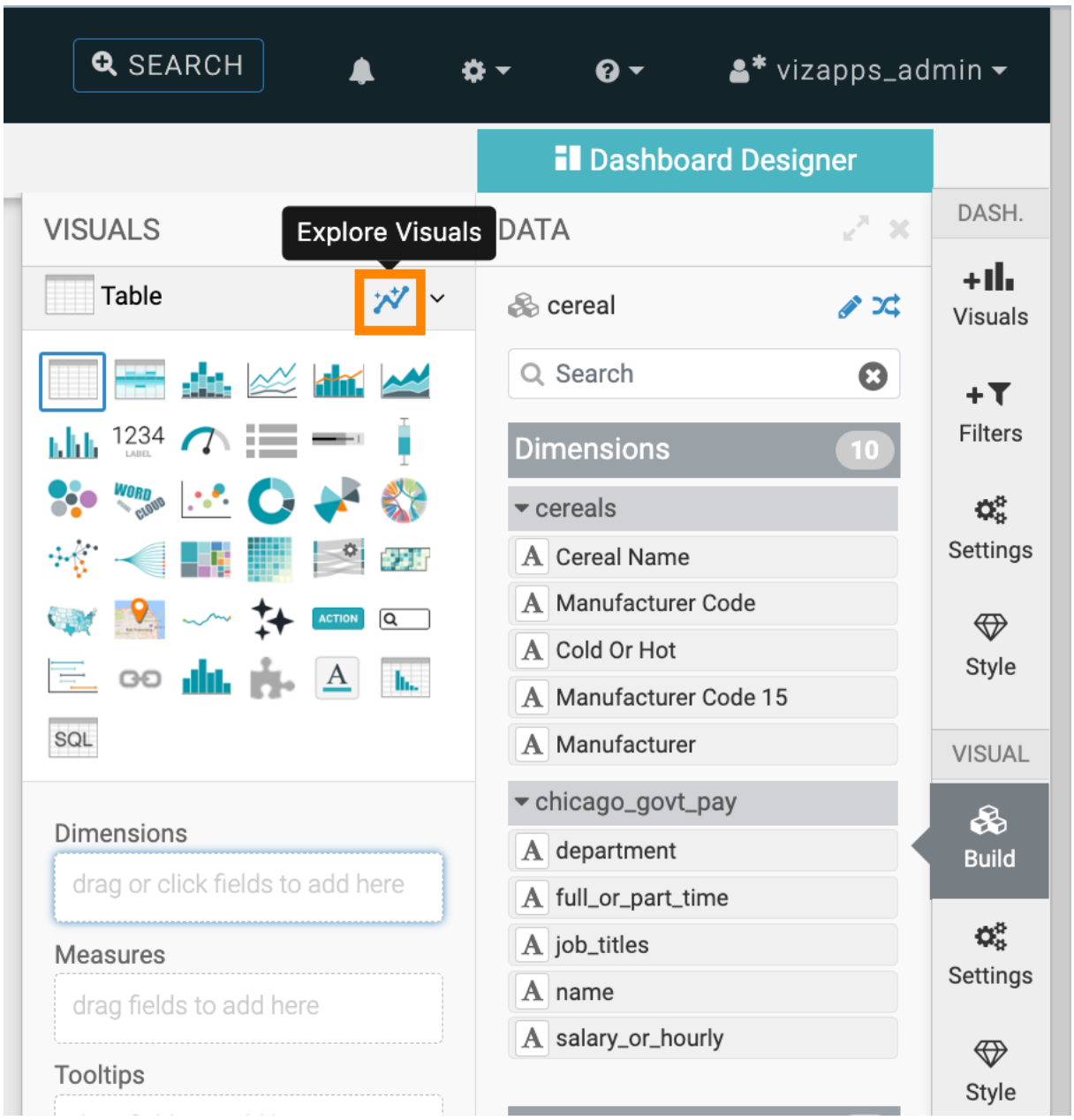
1. Open the dashboard in Edit mode.
2. Select an existing visual and hover over its top right corner.

   Click Explore Options and choose one from the available recommendations. The options presented depend on the visual type.

    - Visual Types
    - Colors
    - Styles/Fonts
    - Settings



3. You can also select the Explore Visuals icon in the Dashboard Designer.



Visual type catalog

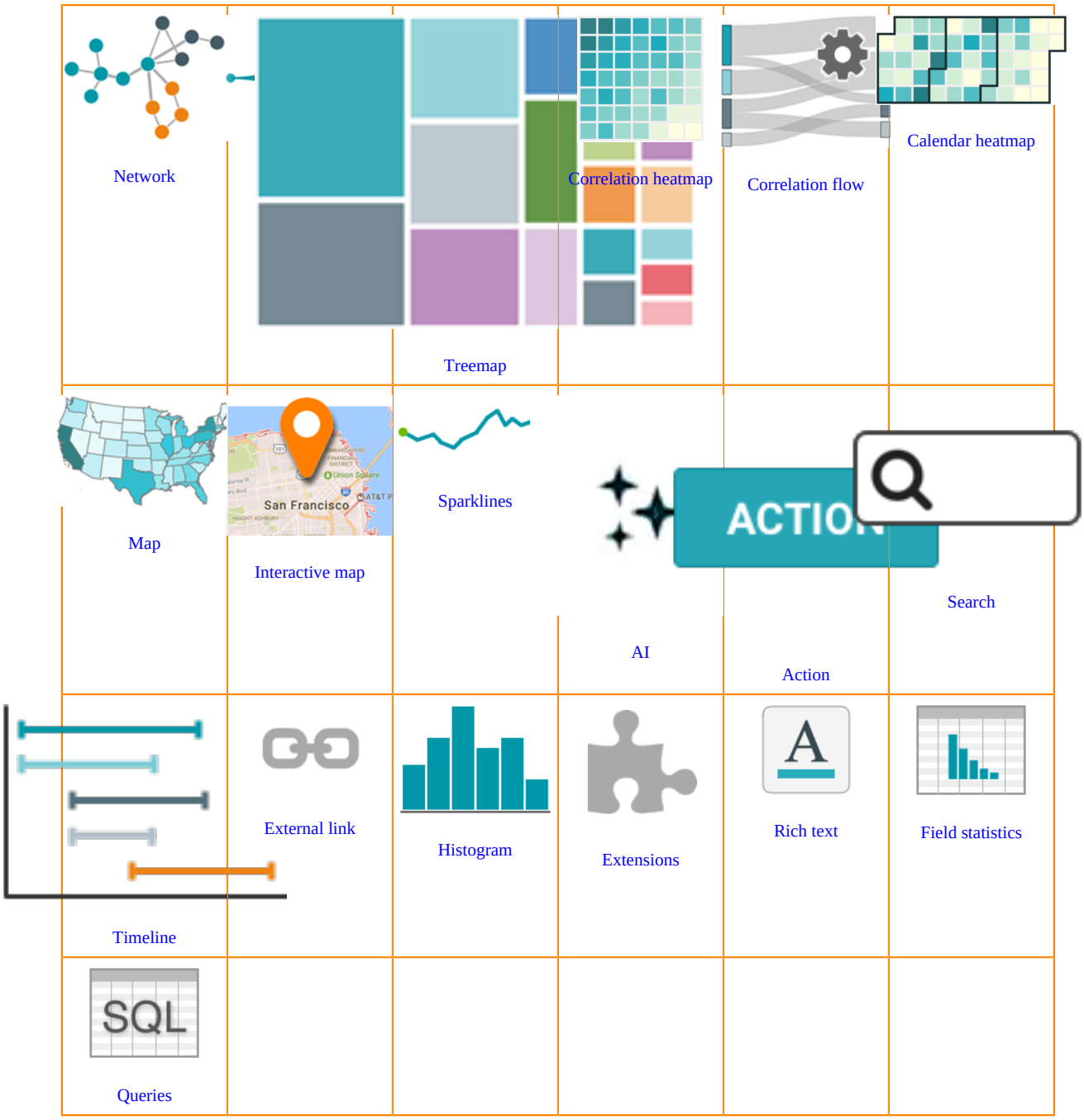
Cloudera Data Visualization has a rich offering of visualization types to assist you in analyzing data.

Click each visual to see how to build and use the visual type.

availability:

The following visual types support smooth rendering of streaming data: *Lines*, *Bars*, *Areas*, *Stacked bars*, *Combo*, *Calendar heatmap*, *Correlation heatmap*, *Packed bubbles*, *Pie*, and *Scatter*.

Table
Cross tabulation
Bars
Lines
Combo
Areas
Grouped bars
Box plot
Bullet
Row listing
Packed bubbles
WORD
VISUAL
CLOUD
Word cloud
Scatter
Pie
Radial
Chord



Table

A table visual displays dataset data in a tabular format, making it easy to view, compare, and analyze individual records and values. You can sort the data by any displayed measure or dimension by clicking a column header. Each click cycles through ascending, descending, and unsorted order. Table cells can also be formatted using format strings.

About this task

The following example demonstrates how to create a new tabular visual representation on the World Life Expectancy dataset. For information about the shelves available for this visual type, see *Shelves for tables*.

Tip:

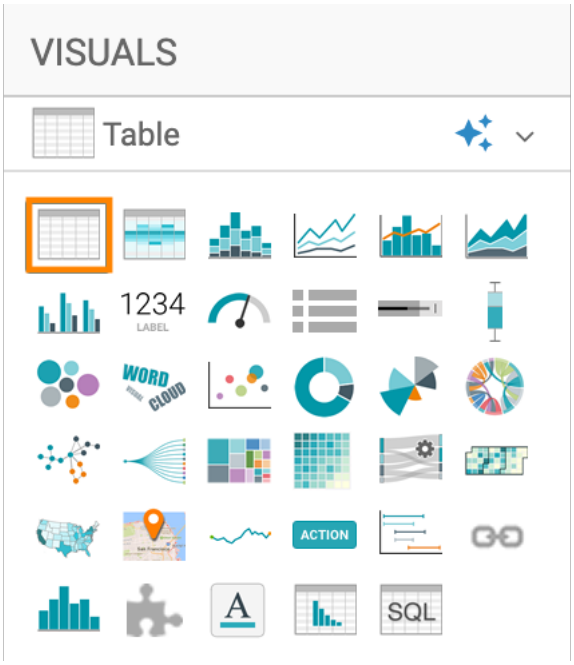
In datasets that contain joins, the default table visual displays columns only from the first table in the dataset, up to the maximum number of columns specified in the Table Style menu.

- To change the maximum number of displayed columns, see *Specifying maximum number of columns*.
- To show columns from other tables of the dataset, add them to the shelves of visual from the Dimensions and Measures of the DATA menu.

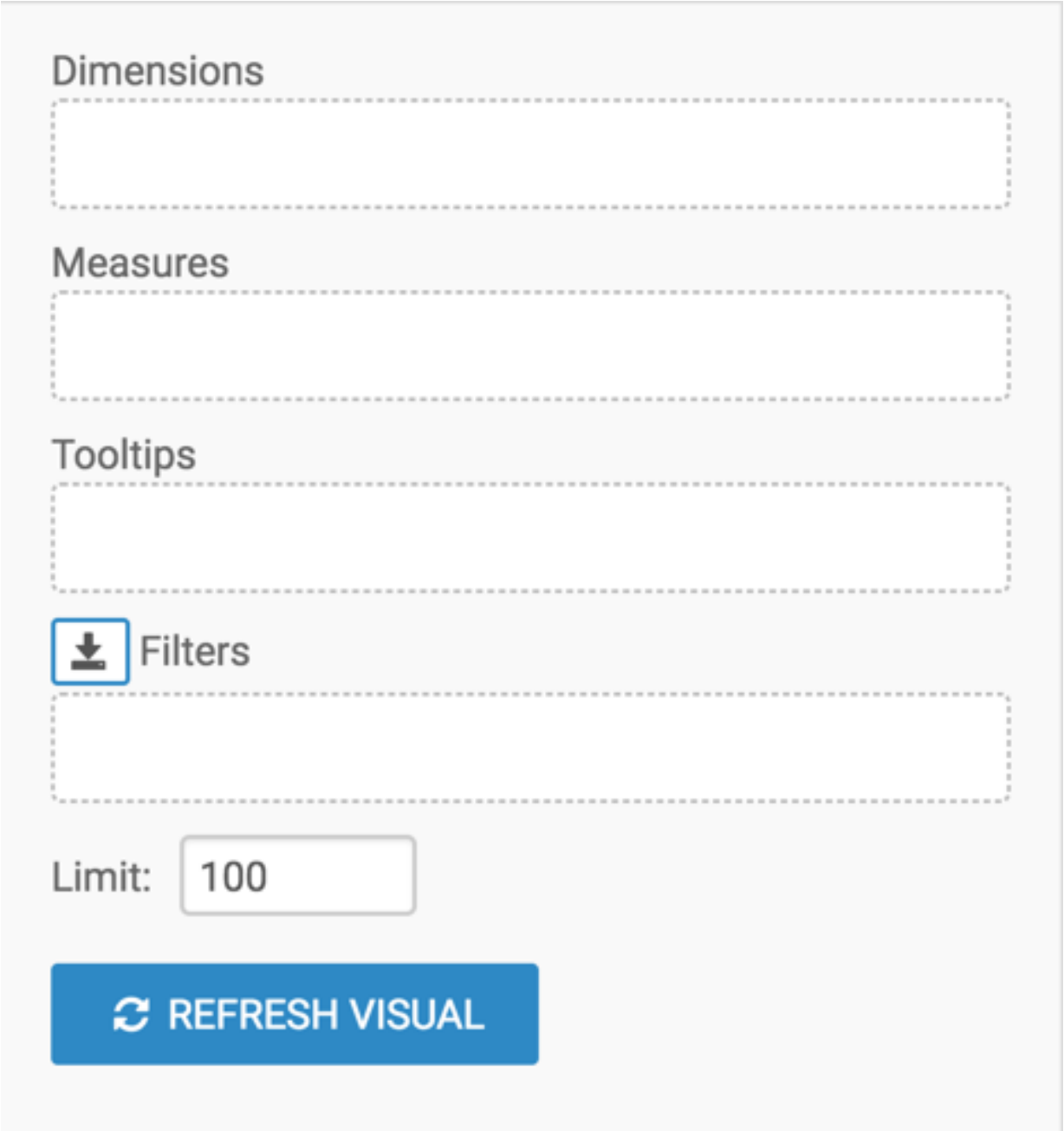
Procedure

1. Create a new visual based on the World Life Expectancy dataset.

   For instructions, see *Creating a visual*. By default, the table shows all fields in the dataset. If your starting point is another visual type, select Table in the VISUALS menu.

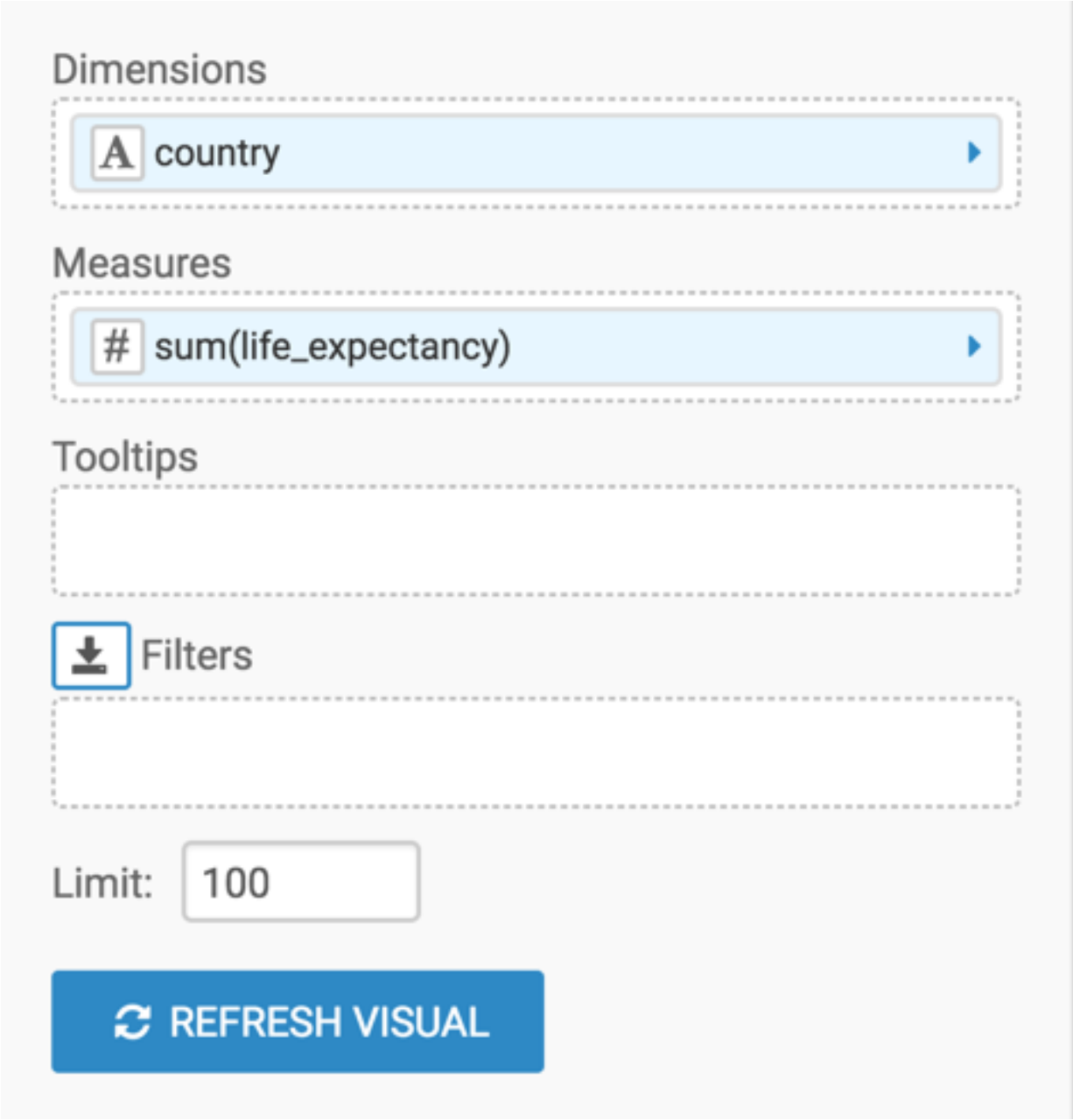


   By default, the table visual displays all available dataset fields.
2. Review the available shelves.
   - Dimensions
   - Measures
   - Tooltips
   - Filters
   - Limit

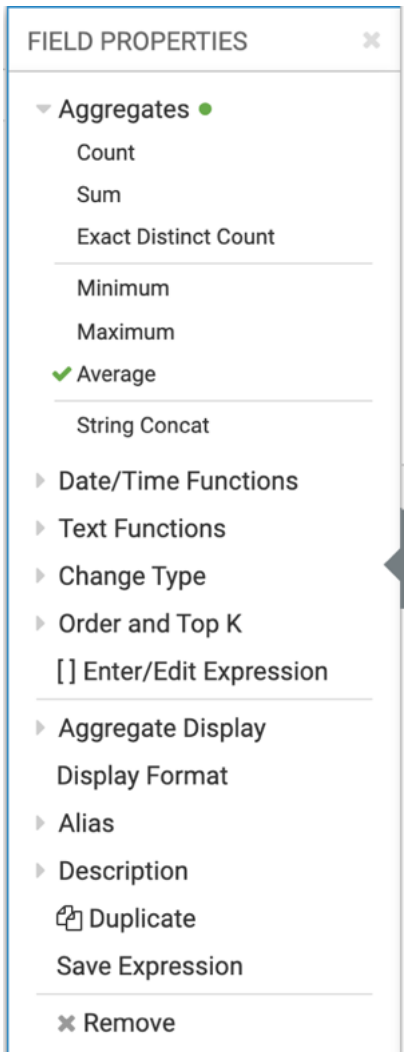


None of these shelves are required. If you do not add fields to any shelf, all available columns are displayed.

3. To show only specific items, populate the shelves from the fields available in the DATA menu.
   - Under Dimensions, select 'country' and drag it to the Dimensions shelf.
   - Under Measures, select 'life_expectancy' and drag it to the Measures shelf.



4. Change the aggregation from 'life_expectancy' field from 'sum(life_expectancy)' to 'avg(life_expectancy)'.
   a. On the Measures shelf of the visual, click the caret-right icon to the right of the 'sum(life_expectancy)' field.
   b. In the FIELD PROPERTIES menu, click the caret-right icon next to Aggregates to open a drop-down menu.
   c. From the list of aggregate functions, select Average.



d. Click the x icon at the top of the FIELD PROPERTIES menu to close it.

The shelf now contains the modified field with the avg() aggregation function.

Alternatively, you can change the aggregation at the level of the dataset, as described in *Changing field aggregation*.

5. Click REFRESH VISUAL.

The updated table visual appears. It displays the average life expectancy by country.

Country	Life Expectancy
Afghanistan	35.4
Albania	54.7
Algeria	45.3
American Samoa	
Andorra	
Angola	33.8
Anguilla	
Antigua and Barbuda	55.4
Argentina	60.9
Armenia	56.2
Aruba	57.0
Australia	68.6

Tip:

If no display format is specified for a numeric field, comma-separated formatting is applied by default. For more information, see, *Customizing display format*.

6. To improve the appearance of the visual, alias the 'country' field as 'Country' and the 'life_expectancy' field as 'Life Expectancy'.
   a. On the shelf of the visual, click the caret-right icon to the right of the field.
   b. In the FIELD PROPERTIES menu, click the caret-right icon next to Alias.
   c. In the text box below Alias, enter the alias name of the column, as it should appear in the visual.
   d. Click the x icon at the top of the FIELD PROPERTIES menu to close it.

   The shelves now show the columns with their alias names.
7. On the Filters shelf, you can add several Dimensions and Measures from the DATA menu.

   This enables you to dynamically control the data input and discover the data at a more granular level.

   For example, from Dimensions, drag the 'un_region' and 'year' fields to the Filters shelf.

Note: The filter selection wizard appears each time you add a field to the Filters shelf. You can specify filter values at this time, or click CANCEL and select them later.

In this example, 'un_region' is set to 'Americas'. It should appear as un_region in ('Americas') on the Filters shelf. 'year' is set to '2001 through 2010'. It should appear as year in (2001,2002,…,2010) on the Filters shelf.

8. Change the value on the Limit shelf to another number.

   In this example, the new value is '12'.

9. Click REFRESH VISUAL.

   The updated table visual appears.

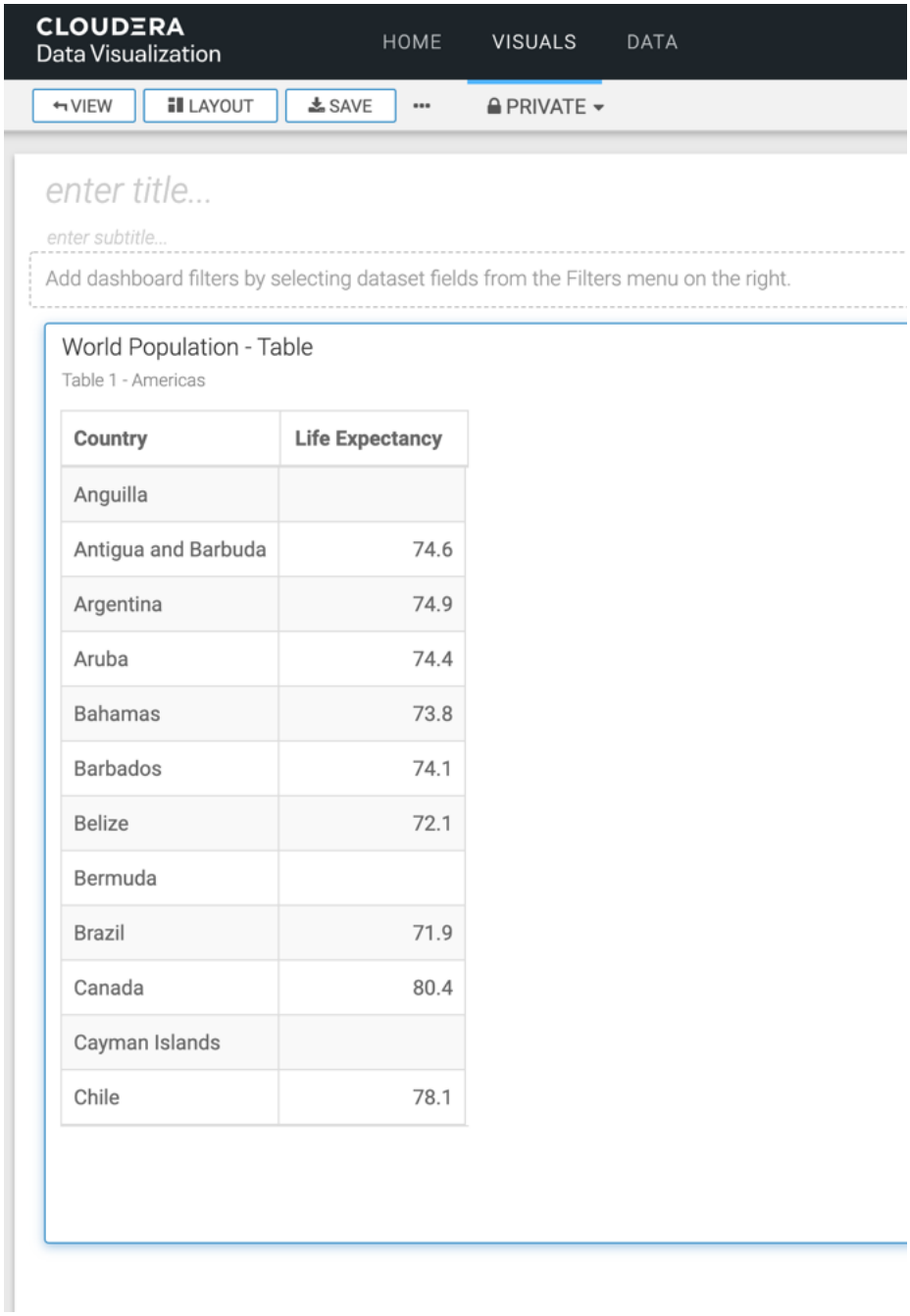
Country	Life Expectancy
Anguilla	
Antigua and Barbuda	74.6
Argentina	74.9
Aruba	74.4
Bahamas	73.8
Barbados	74.1
Belize	72.1
Bermuda	
Brazil	71.9
Canada	80.4
Cayman Islands	
Chile	78.1

10. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

    You can also add a brief description of the visual as a subtitle below the title of the visualization.

11. At the top left corner of the Visual Builder, click SAVE.

Results



Related Information

Specifying maximum number of columns
Creating a visual
Changing field aggregation
Customizing display format

Shelves for tables

Overview of shelves for Cloudera Data Visualization Table visuals.

A Table visual has the following shelves:

Dimensions

Add a Dimension field from the DATA menu to include a dimension column in the table.

This shelf accepts multiple fields.

Measures

Add a Measure field from the DATA menu to include an aggregate column in the table.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Limit

This shelf specifies the number of records displayed in the table.

The default value is 100.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

This field is necessary to see measures that do not appear in the table.

Saving tables

Cloudera Data Visualization enables you to save tabular data as its own table, and writes it back to the database. It also enables you to produce a dataset based on data exposed in a visual, which may be used for self-referenced application stacking.

About this task

This feature is available only when:

- Examining data in either table or query visuals
- Using Impala connections
- User has Data Import/Load privileges

Saving a table as a materialized table

Cloudera Data Visualization allows you to save a table as pre-computed (materialized) data. Use a materialized table if you want to permanently store your query results.

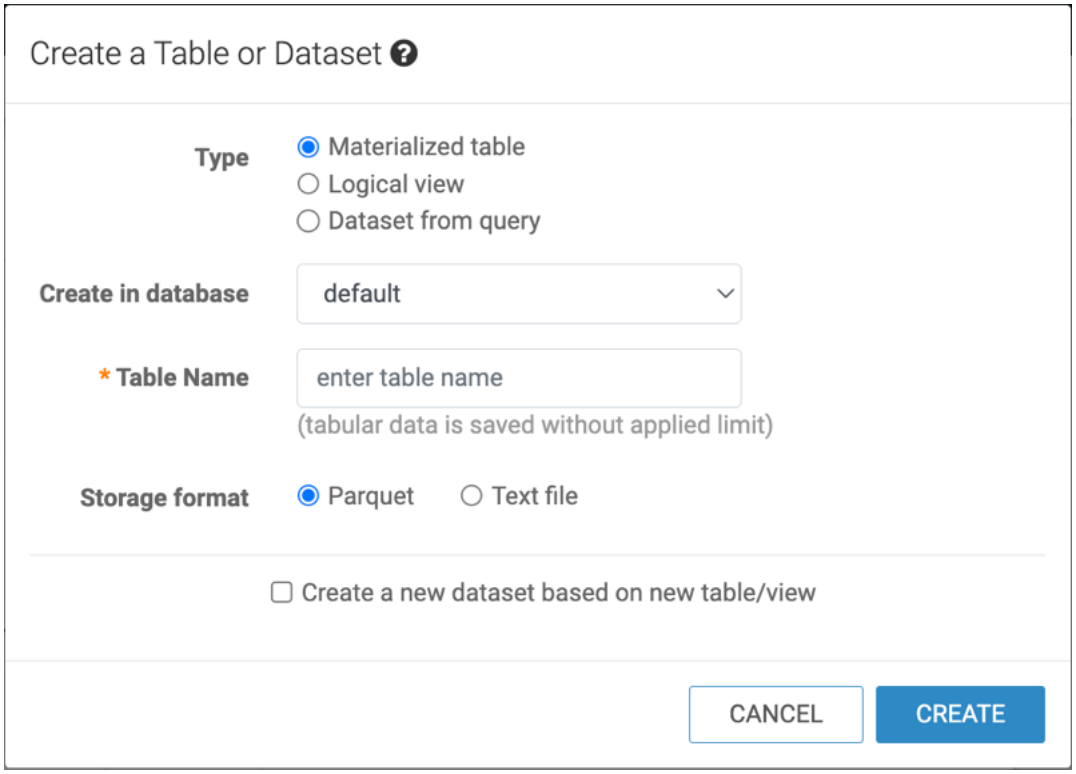
Before you begin

You have a table visual that you want to save as a materialized table.

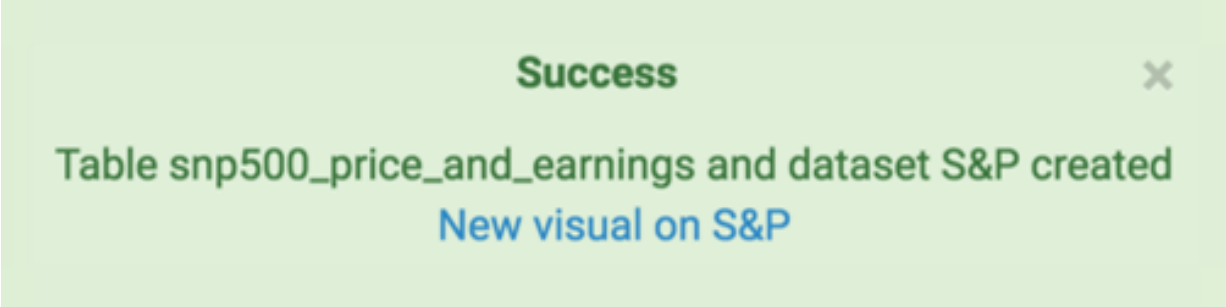
Procedure

1. On the top right menu, click the (ellipsis) icon and select Save Table from the menu.
2. In the Create a Table or Dataset modal window, fill out these fields:
   a) Choose Materialized table for Type to permanently save query results.
   b) Under Create in database, choose an existing database where you want to save the table.
   c) Enter the name of the new table.
   d) Choose Parquet or Text file as Storage format.

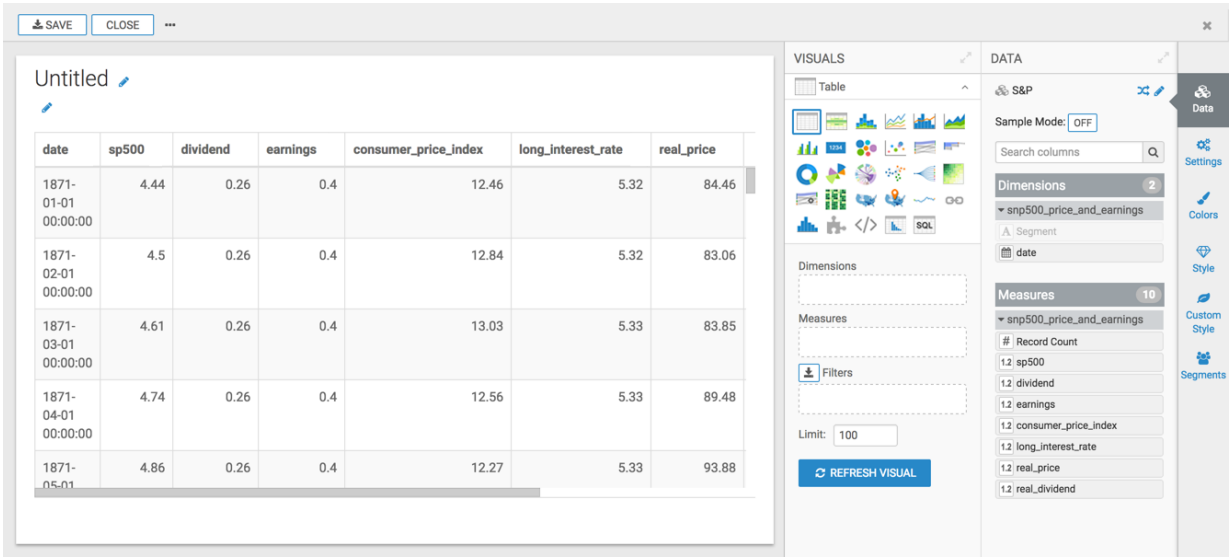
e) When creating a dataset as part of this operation, select the Create a new dataset based on new table/view option and enter the Dataset name.
f) Click CREATE.



A success confirmation modal appears, with a link to the new default visual on the new dataset.



3. Click the link to open a new default visual on the new dataset in Edit mode.



Saving a table as a logical view

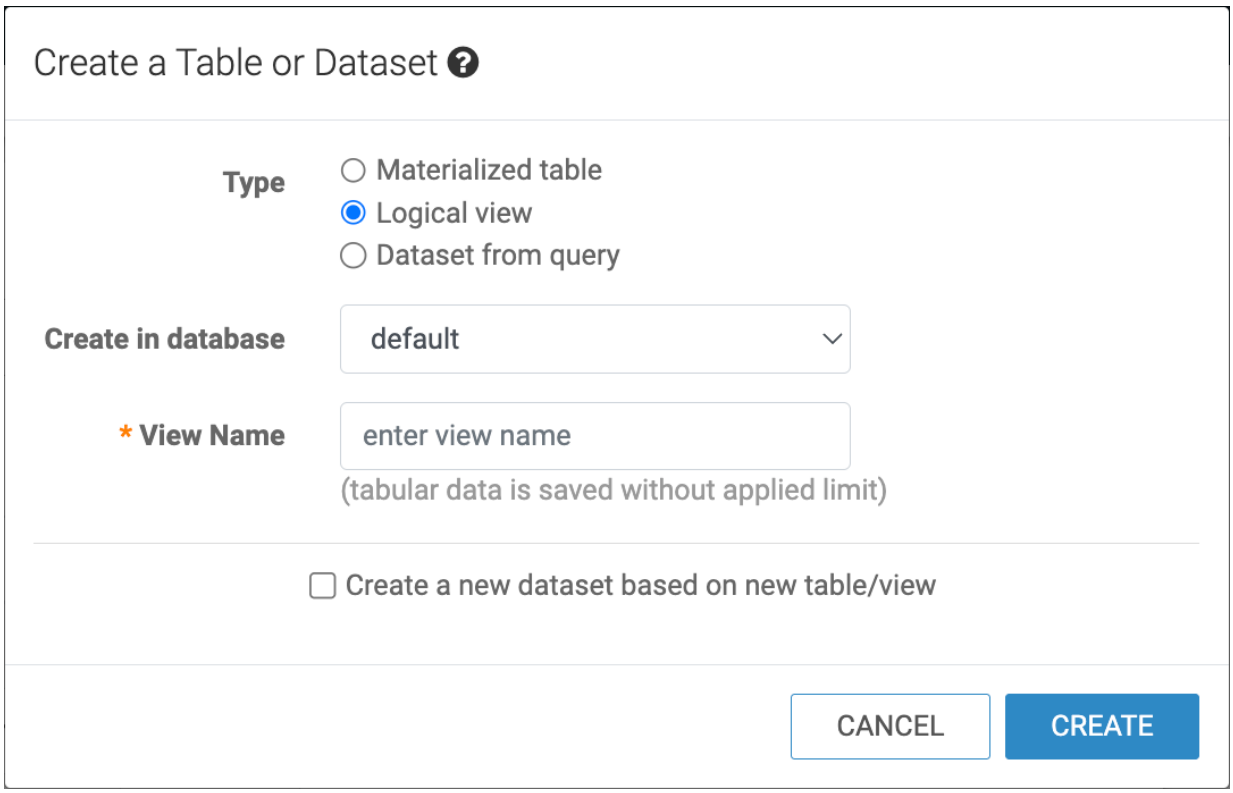
Cloudera Data Visualization allows you to create logical tables. Use a logical view if you want updated query results every time you run the visualization.

Before you begin

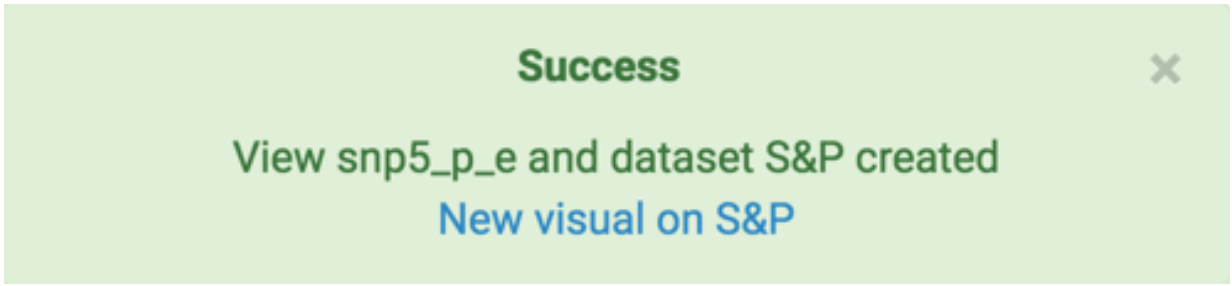
You have a table visual that you want to save as a logical view.

Procedure

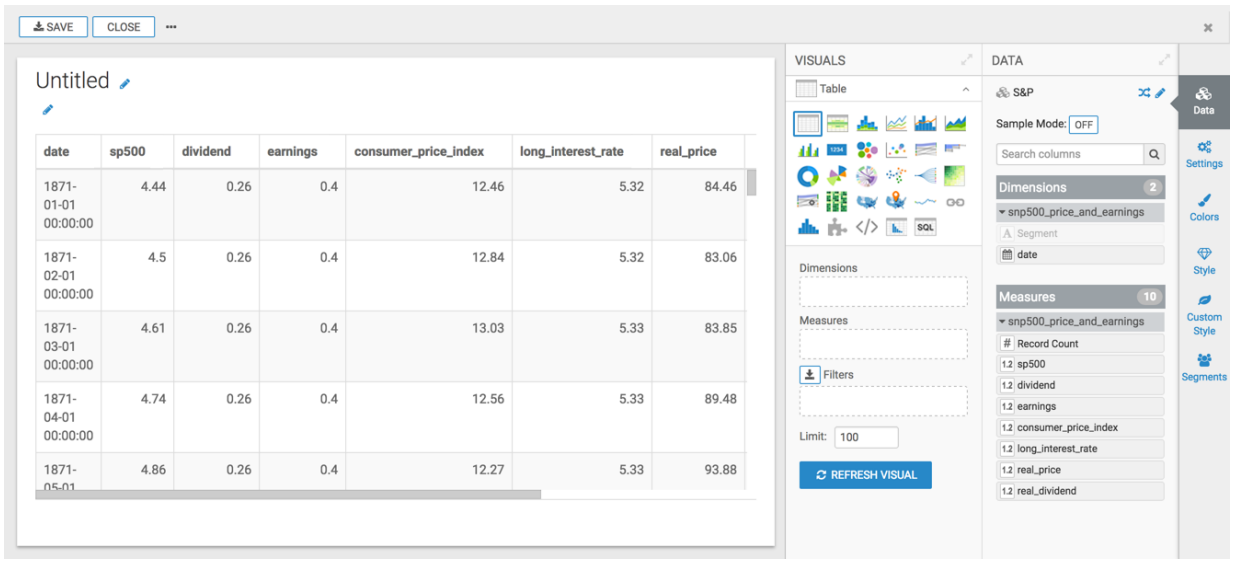
1. On the top right menu, click the (ellipsis) icon and select Save Table from the menu.
2. In the Create a Table or Dataset modal window, fill out these fields:
   a) Choose Logical view for Type to generate updated query results every time the visualization is loaded or refreshed.
   b) Under Create in database, choose an existing database where you want to save the table.
   c) Enter the name of the new logical view.
   d) When creating a dataset as part of this operation, select the Create a new dataset based on new table/view option and enter the Dataset name.



A success confirmation modal appears, with a link to the new default visual on the new dataset.



3. Click the link to open a new default visual on the new dataset, in Edit mode.



Saving a table as a dataset based on the query

Cloudera Data Visualization allows you to save a query result as a separate dataset. Unlike a logical view, which is created on the server, this method uses the current query as the basis for the dataset. This feature is particularly useful when working with subsets of data, specific sections, or when constructing datasets with self-joins.

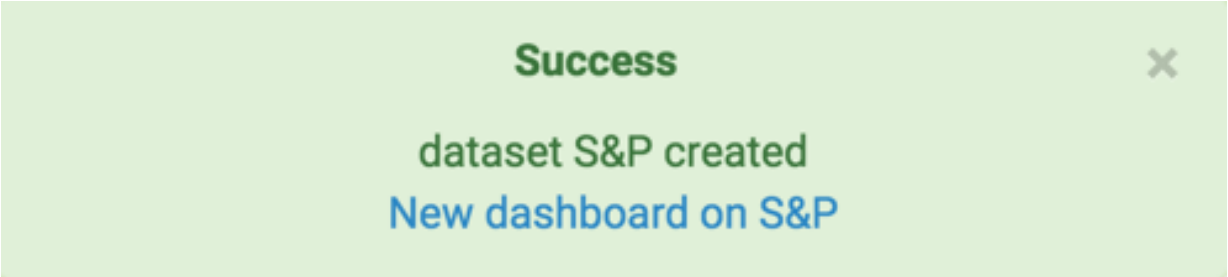
Before you begin

You have a table visual that you want to save as a dataset.

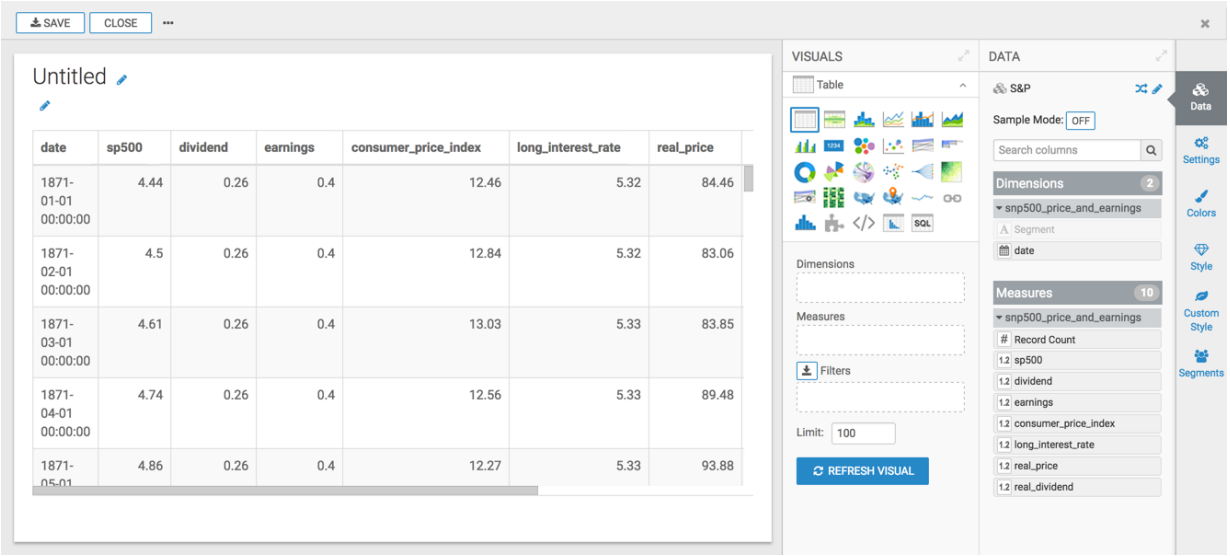
Procedure

1. Add the necessary fields on the table shelves, and apply functions, for example sort, order, alias, or filters.
2. On the top right menu, click the (ellipsis) icon and select Save Table from the menu.
3. In the Create a Table or Dataset modal window, fill out these fields:
   a) Choose Dataset from query for Type.
   b) Enter the name of the new dataset.
   c) In the SQL for dataset creation field, you can edit the SQL expression used for creating the dataset.
   d) Click CREATE.

A success confirmation modal appears, with a link to the new default visual on the new dataset.



4. Click the link to open a new default visual on the new dataset, in Edit mode.



Cross tabulation

Cross Tabulation visual is a tool that allows you to compare the relationship between two variables in table format. Cloudera Data Visualization enhances the understanding of relative value size by assigning more saturated colors to the table fields with larger values.

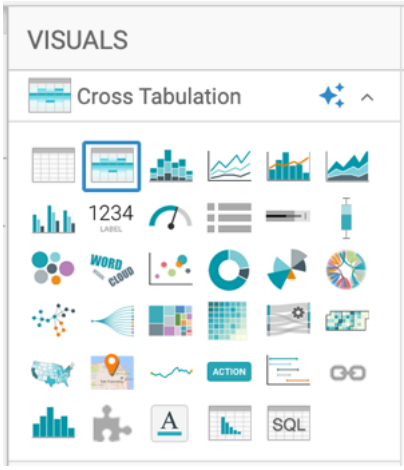
About this task

The following steps demonstrate how to create a new Cross Tabulation visual on the World Life Expectancy dataset. For an overview of shelves that specify this visual, see *Shelves for cross tabulation*.

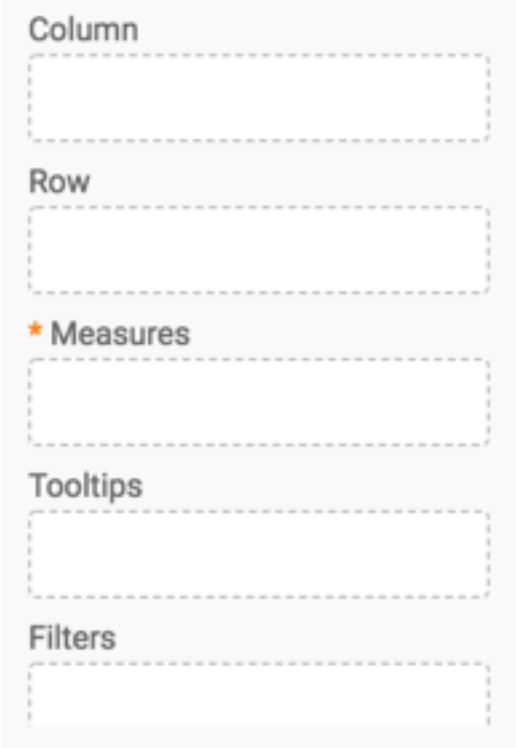
Procedure

1. Start a new visual based on the World Life Expectancy dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Cross Tabulation.

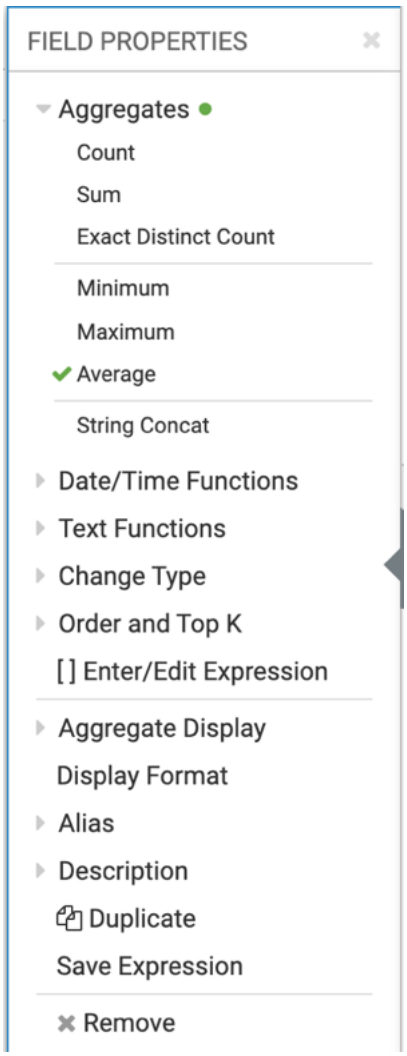


The shelves of the visual changed. They are now Column, Row, Measures, Tooltips, and Filters. The Measures shelf is mandatory.



3. To show specific items, populate the shelves from the available fields in the DATA menu.
   - Under Dimensions, select 'country' and drag it to the Column shelf.
   - Under Dimensions, select 'year' and drag it to the Row shelf.
   - Under Measures, select 'life_expectancy' and drag it to the Measures shelf.

     By default, the sum aggregate of the field appears on the shelf, as 'sum(life_expectancy)'.
4. Change the aggregate of the 'life_expectancy' field from 'sum(life_expectancy)' to 'avg(life_expectancy)'.
   a. On the Measures shelf of the visual, click the caret-right icon to the right of the field.
   b. In the FIELD PROPERTIES menu, click the caret-right icon next to Aggregates to open a drop-down menu..
   c. From the list of aggregate functions, select Average.



d. Click the x icon at the top of the FIELD PROPERTIES menu to close it.

The shelf now contains the modified field with the 'avg()' aggregation function.

Alternatively, you can change the aggregation at the level of the dataset, as described in *Changing field aggregation*.

5. Click REFRESH VISUAL.

The updated Cross Tabulation visual appears. It shows life expectancy information for each row of the dataset. By default, Cross Tabulation visuals show smaller values in light colored cells, and larger values in darker colored cells.

	country													
	Afghanistan	Albania	Algeria	American Samoa	Andorra	Angola	Anguilla	Antigua and Barbuda	Argentina	Armenia	Aruba	Australia	Austria	Azerbaijan	Bahamas
year	avg(life_expectancy)														
1900	27.2	35.4	28.9			27		33.5	36.8	35.9	34.4	49.9	41.5	31.1	35.2
1901	27.2	35.4	28.9			27		33.5	37.4	36.1	34.4	50.5	42	31.3	35.2
1902	27.2	35.4	28.9			27		33.5	38	36.4	34.4	51	41.1	31.6	35.2
1903	27.2	35.4	29.9			27		33.5	38.5	36.6	34.4	51.5	40.1	31.8	35.2
1904	27.1	35.4	24.1			27		33.5	39.1	36.9	34.4	52	40.7	32.1	35.2
1905	27.1	35.4	26.7			27		33.5	39.7	37.1	34.4	52.6	41.4	32.3	35.2
1906	27.1	35.4	28			27		33.5	40.5	37.4	34.4	53.1	42	32.6	35.2
1907	27.1	35.4	27.9			27		33.5	41.2	37.6	34.4	53.6	42.6	32.8	35.2
1908	27.1	35.4	27.8			27		33.5	42	37.9	34.4	54.2	43.2	33.1	35.2
1909	27.1	35.4	29.3			27		33.5	42.7	38.1	34.4	54.7	43.9	33.3	35.2
1910	27.1	35.4	30.8			27		33.5	43.5	38.4	34.4	55.2	44.5	33.6	35.2
1911	27.1	35.4	30.5			27		33.5	44	41.4	34.4	55.7	45.1	36.5	35.2
1912	27.1	35.4	31.8			27		33.5	44.6	41.8	34.4	56.3	45.7	36.9	35.2
1913	27.1	35.4	29.7			27		33.5	45.1	39.6	34.4	56.8	46.2	34.8	35.2
1914	27	35.4	29.2			27		33.5	45.7	39.2	34.4	57.3	46.8	34.4	35.2
1915	27	35.4	28.7			27		33.5	46.2	38.8	34.4	57.9	47.4	34	35.2
1916	27	35.4	28.3			27		33.5	47	38.8	34.4	58.4	47.9	34	35.2
1917	27	35.4	28.3			27		33.5	47.8	35.8	34.4	58.9	48.5	31	35.2
1918	7	19.4	22.1			10.4		21.8	41.8	27	22.3	54.8	32.4	22	22.8
1919	27	35.4	28.3			27		33.5	49.4	37	34.4	60	49.6	32	35.2
1920	27	35.4	27.5			27		33.5	50.2	28	34.4	60.5	50.2	30	35.2
1921	27	35.4	27.5			27		33.5	50.7	38	34.4	61	50.8	33	35.2

6. On the Filters shelf, you can add several Dimensions and Measures from the DATA menu.

   This enables you to dynamically control the data input, and discover the data at a more granular level.

   - For example, from Dimensions, drag 'un_subregion' to the Filters shelf.

     Note: The filter selection wizard appears each time you add a field to the Filters shelf. You can specify filter values at this time, or click CANCEL and select them later.

     In this example, the filter value is set to 'Middle Africa'.
   - To limit the time interval, add the 'year' field to the Filters shelf.

     In the Filter for year modal window, click the Range tab, select the last 30 years of the dataset, 1981 through 2010, and lick APPLY.
7. To improve the appearance of the visual, alias the 'country' field as 'Country' and the 'life_expectancy' field as 'Life Expectancy'.

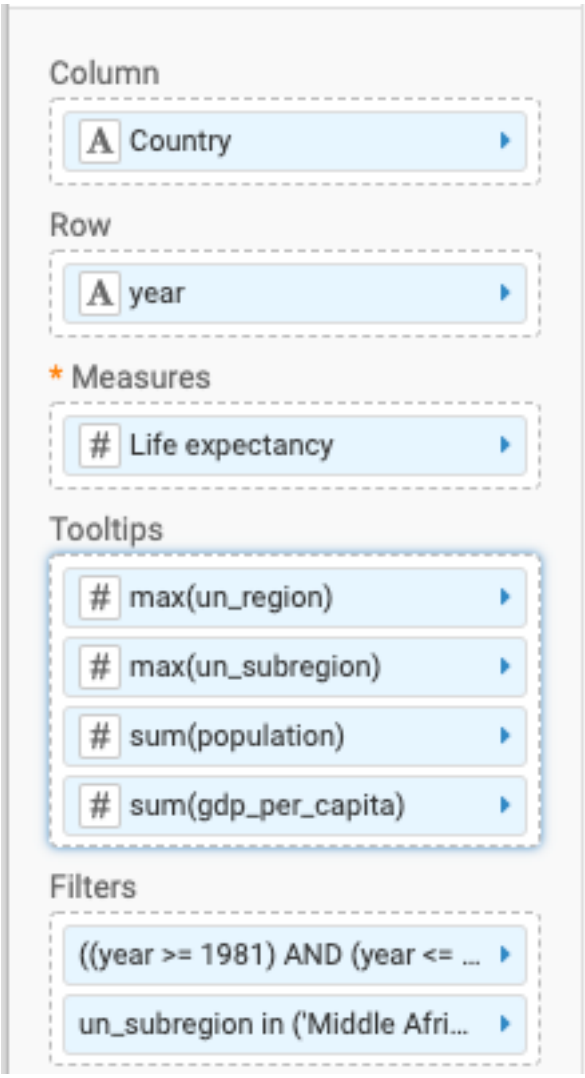
   a. On the shelf of the visual, click the caret-right icon to the right of the field.
   b. In the FIELD PROPERTIES menu, click the caret-right icon next to Alias.
   c. In the text box below Alias, enter the alias name of the column, as it should appear in the visual.
   d. Click the x icon at the top of the FIELD PROPERTIES menu to close it.

   The shelves now show the columns with their alias names.
8. Adjust the display format for the 'Life Expectancy' field to use a comma thousands separator, and to carry zero decimal places.

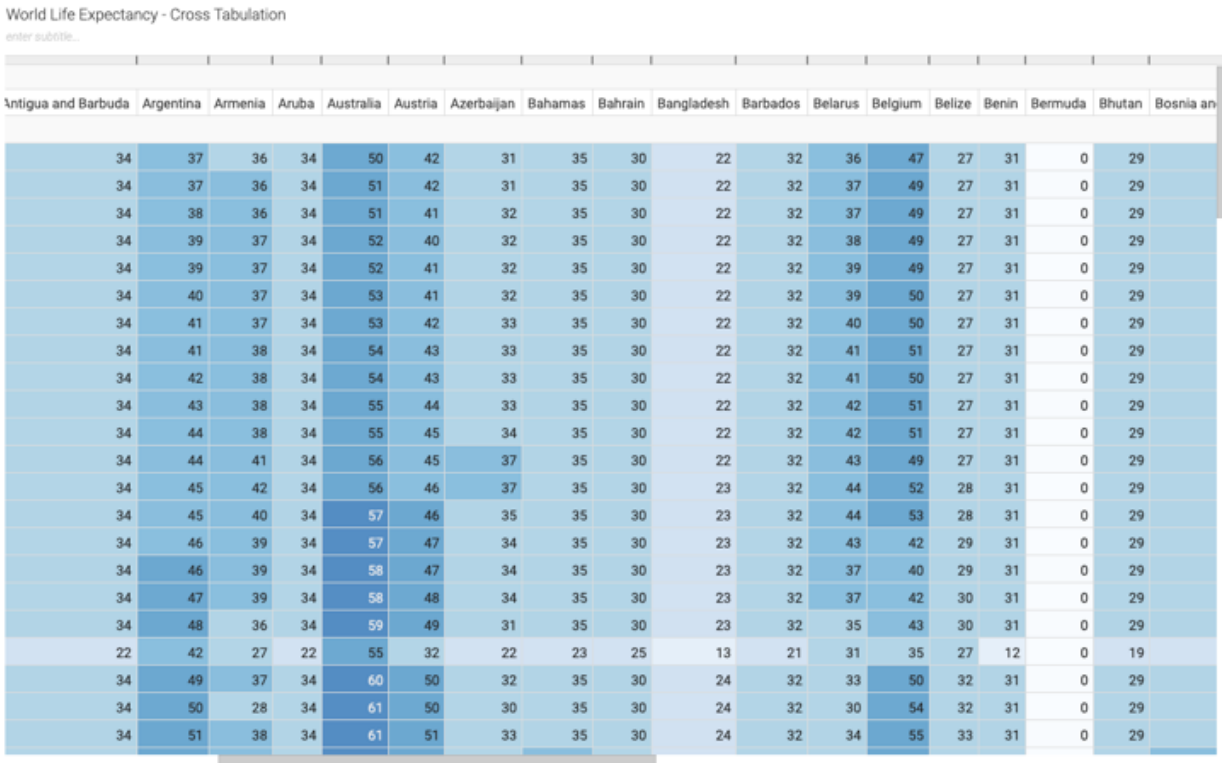
   a. On the Measures shelf of the visual, click the caret-right icon to the right of the field.
   b. In the FIELD PROPERTIES menu, click Display Format.
   c. In the Display Format modal window, select the Category 'Real Number'.
   d. In the Decimal Places text box, set '0'.
   e. Select the Use 1000 separator option, and click SAVE.
   f. Click x icon at the top of the FIELD PROPERTIES menu to close it.
9. To provide more detailed information, add some fields to the Tooltips shelf.

For example, drag the Dimensions 'un_region' and 'un_subregion', and the Measures 'population' and 'gdp_per_capita' to the Tooltips shelf.

10. Click REFRESH VISUAL.



The updated Cross Tabulation visual appears. It demonstrates the relative size of populations by assigning a progressively more saturated color to represent increasingly larger values.



11. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

In this example, the title is changed to 'World Life Expectancy - Cross Tabulation'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

12. At the top left corner of the Visual Builder, click SAVE.

Related Information

Shelves for cross tabulation

Creating a visual

Changing field aggregation

Displaying totals with data expansion

When performing granular data analysis in Cross Tabulation visuals, Cloudera Data Visualization enables you to display both subtotals and combined total. These summations operate on the dimensions in the Row shelf. When you enable expansion, you can expand rows to view details of some of the dimensions.

Check out the instructions below on how to enable expansion and display subtotals and totals in Cross Tabulation visuals.

For more information about enabling expansion totals, see *Showing expansion totals above expansion rows*.

Related Information

Showing expansion totals above expansion rows

Configuring a cross tabulation visual for expansion

Follow these steps to enable expansion for a Cross Tabulation visual.

About this task

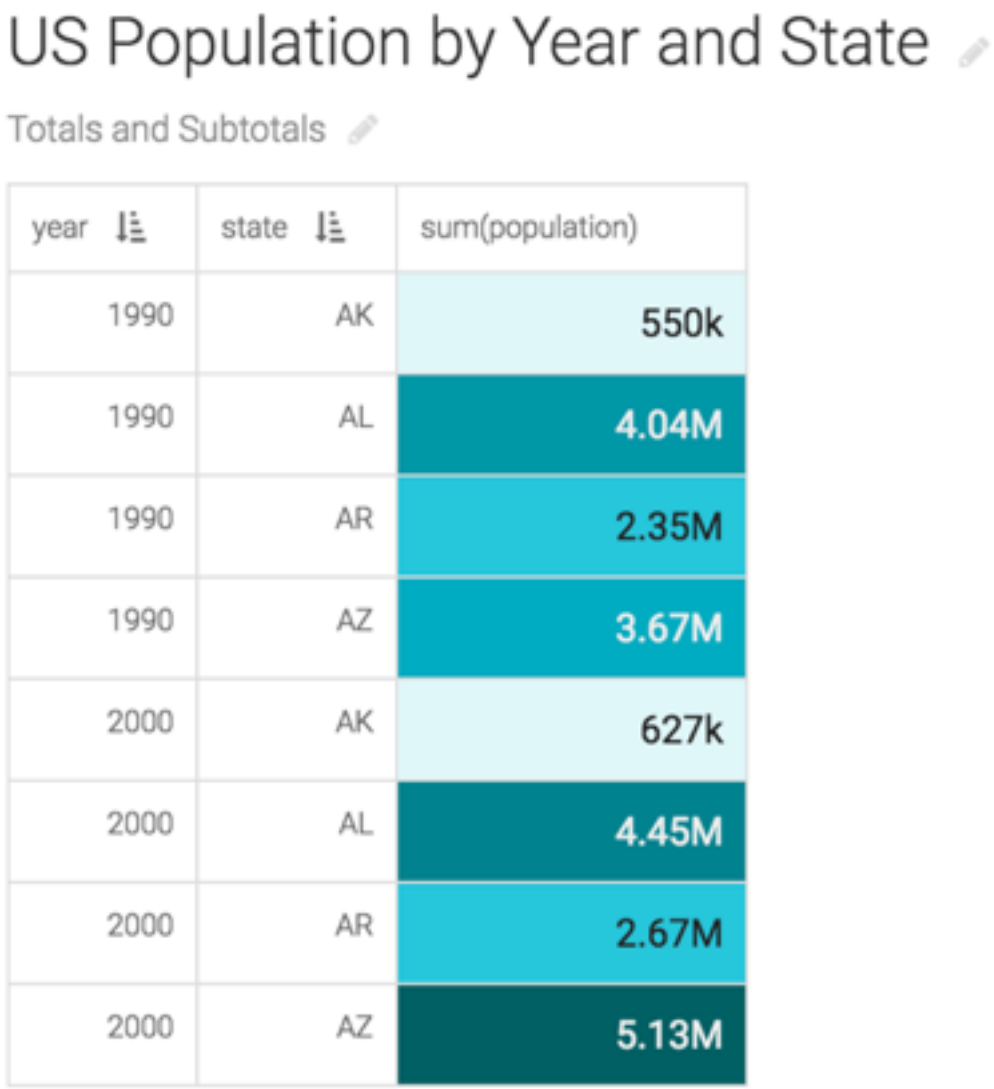
This example shows how to enable expansion on the year and state fields.

Procedure

1. Start a new Cross Tabulation visual based on the US State Populations Over Time dataset.

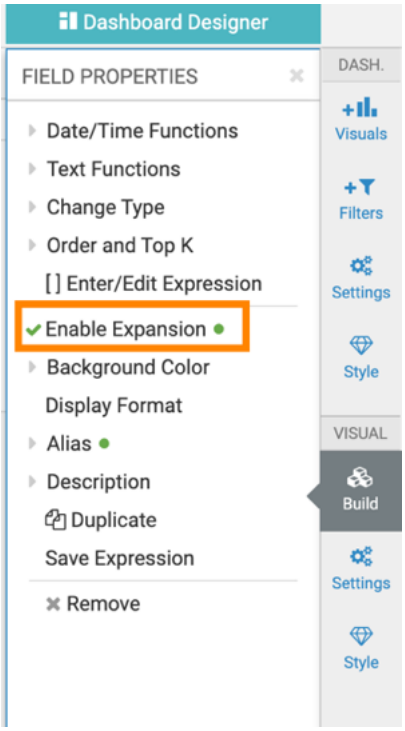
   For instructions, see *Creating a visual*.
2. Populate the shelves from the available Fields (Dimensions, Measures) in the DATA menu.
   a) Add year and state to the Row shelf.
   b) Add population to the Measures shelf. Ensure that the default aggregation is sum(population).
   c) Add year to the Filters shelf, set it to 1990 and 2000, and click APPLY.
   d) Add state to the Filters shelf, select values AK, AL, AR, and AZ, then click APPLY.
3. Click REFRESH VISUAL.

   When the Cross Tabulation visual appears, it does not have subtotals and the combined total in the rows and columns.



4. On the Row shelf, click year.
5. In the FIELD PROPERTIES menu, click Enable Expansion.

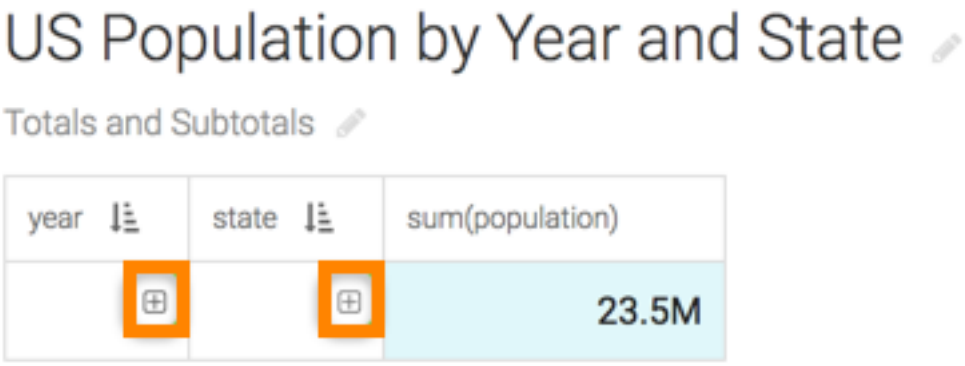
   A green tick appears to the left of the option and a green dot appears to the right of the option.



6. Click state on the Row shelf and enable expansion on that field.
7. Click REFRESH VISUAL to see the changes.

   The visual appears with a + icon in the year and state columns. This indicates that you can expand the data.

   The sum(population) column shows the combined population for both years.

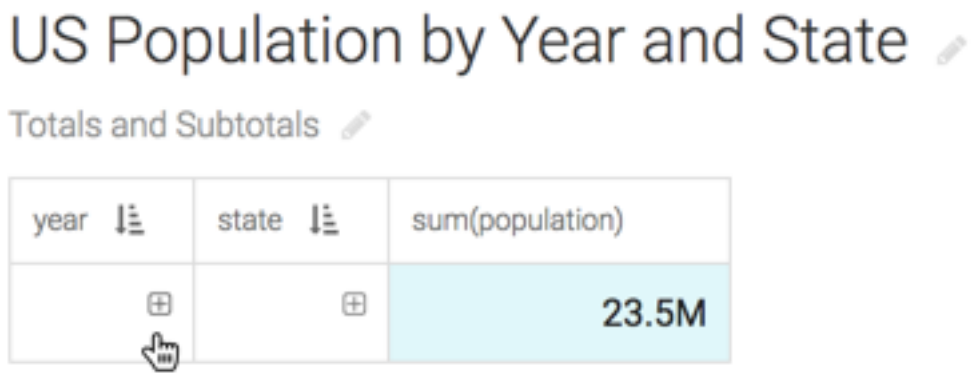


Displaying subtotals and totals

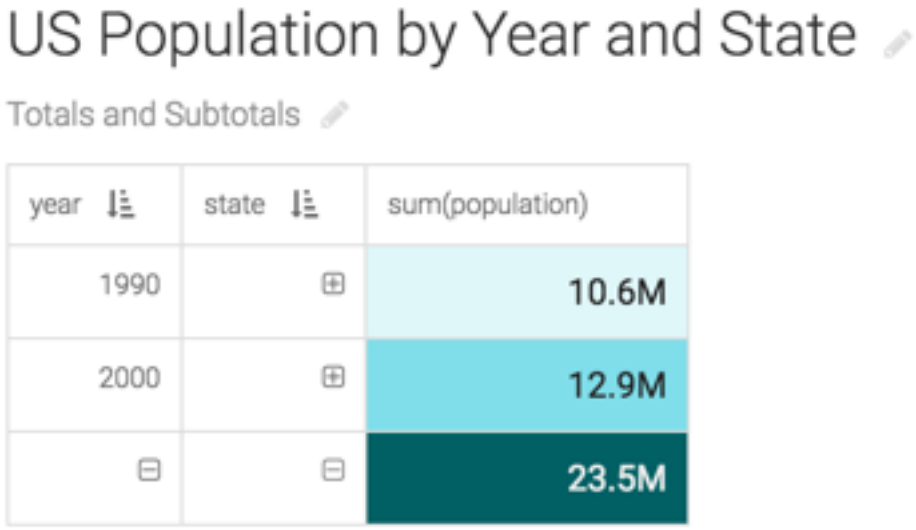
You can expand and shrink the data rows, and display only subtotals and combined total in a Cross Tabulation visual.

Procedure

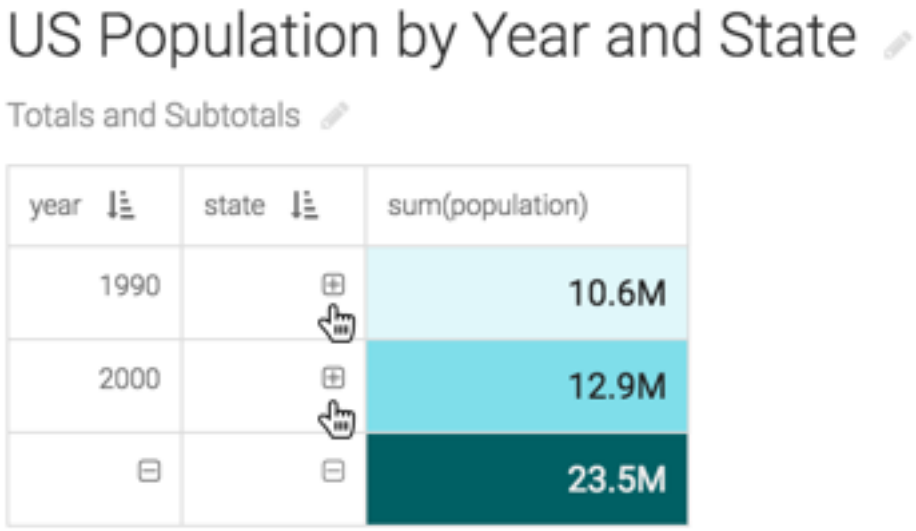
1. Click the + icon in the year column.



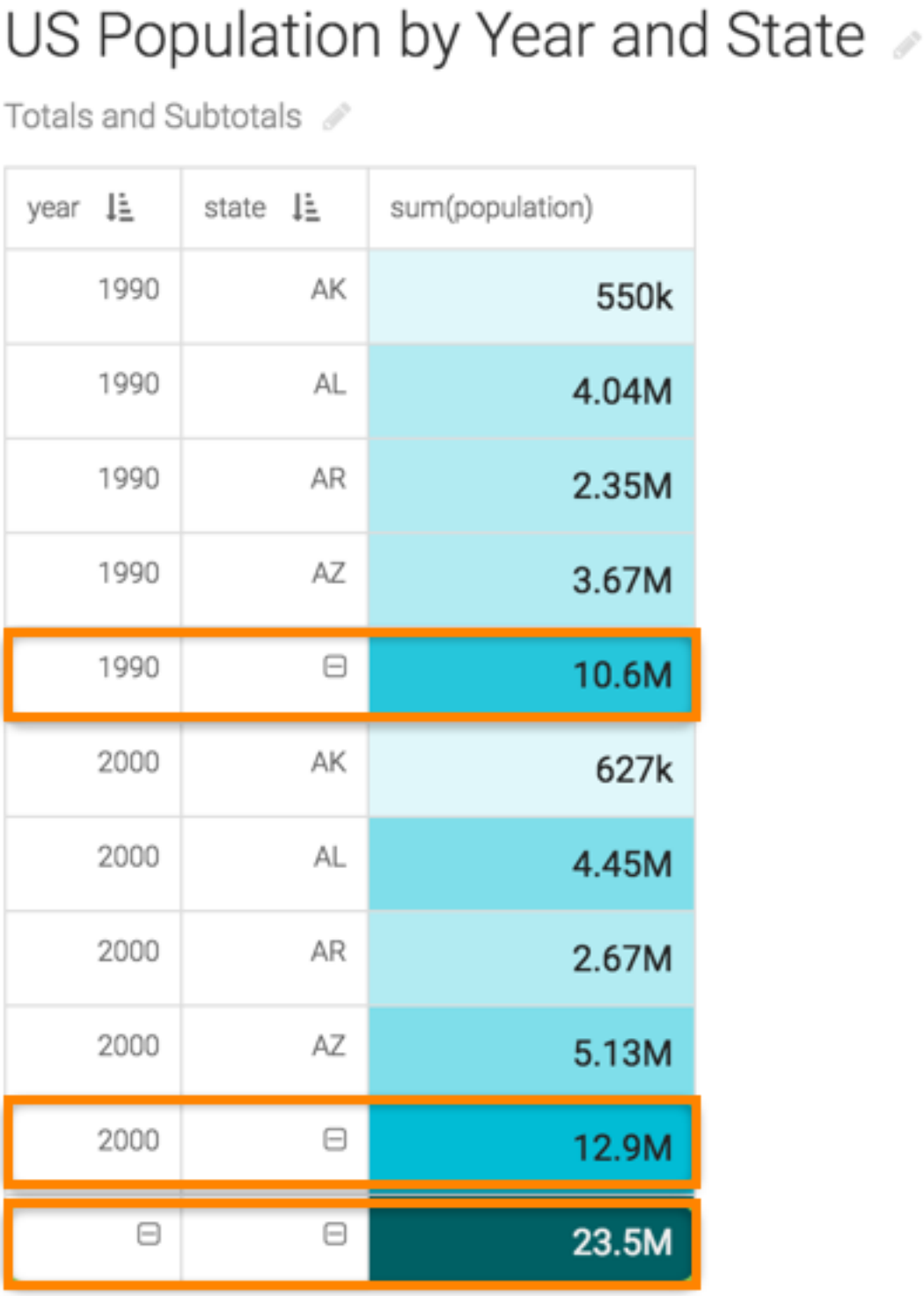
The visual expands the data to show subtotals of each year, and combined total for both years.



2. Click the two + (plus) icons in the state column.



The visual expands the data to show population in each state, and the subtotals and combined population for both years. The subtotals and combined total appear at the bottom of expanded rows. To move the subtotals and totals over the rows, see *Showing expansion totals above expansion rows*.



What to do next

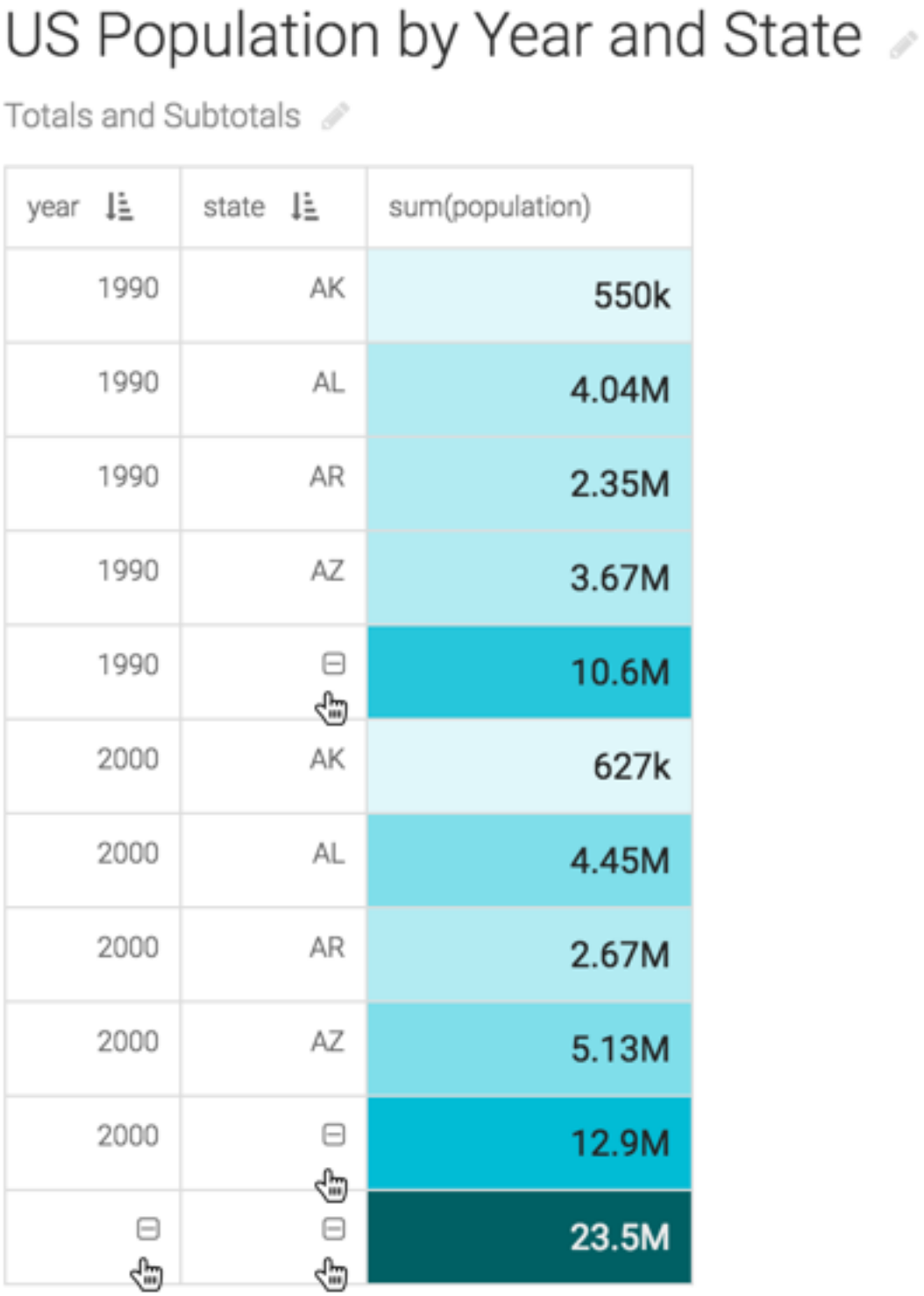
To disable subtotals and combined totals, make changes in the FIELD PROPERTIES menu by de-selecting the Enable Expansion option.

Hiding data and subtotal rows

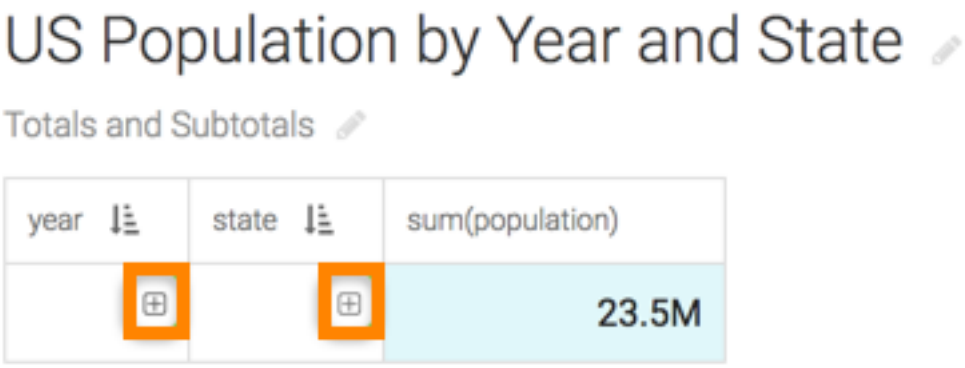
To hide the expanded rows and subtotals, follow these steps.

Procedure

1. Click the - (minus) icon in the year and state columns.



2. The expanded rows and subtotals no longer appear in the visual. Note that a (plus) icon appears again in the year and state columns.



Shelves for cross tabulation

Overview of shelves for Cloudera Data Visualization Cross Tabulation visuals.

A Cross Tabulation visual has the following shelves:

Column

Add a Dimension field from the DATA menu to break aggregations into columns.

This shelf accepts multiple fields.

Row

Add a Dimension field from the DATA menu to break aggregations into rows.

This shelf accepts multiple fields.

Measures

Add a Measure field from the DATA menu to specify the values aggregated across columns and rows.

This shelf accepts multiple fields.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Bars

In Cloudera Data Visualization, bar charts use either vertical or horizontal bars to show comparisons among categories. One axis shows specific categories (typically, dimensions), while the other axis represents discrete values (measures).

About this task

Bar charts are good for avoiding clutter when data labels are long, or when comparing a large number of items, typically more than ten. Bar charts can also display negative numbers.

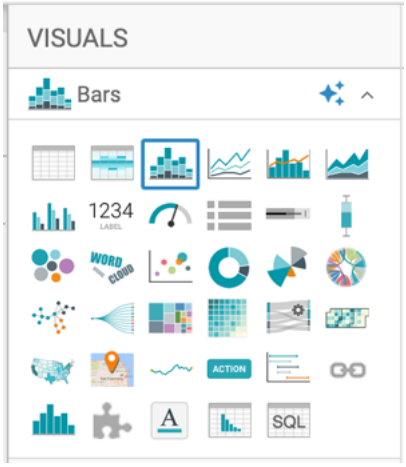
Note: This visual type supports smooth rendering of streaming data.

The following steps demonstrate how to create a new Bars visual on the World Life Expectancy dataset.

For more information, see *Adding data* and *Creating datasets*. For an overview of shelves that specify this visual, see *Shelves for bars visuals*.

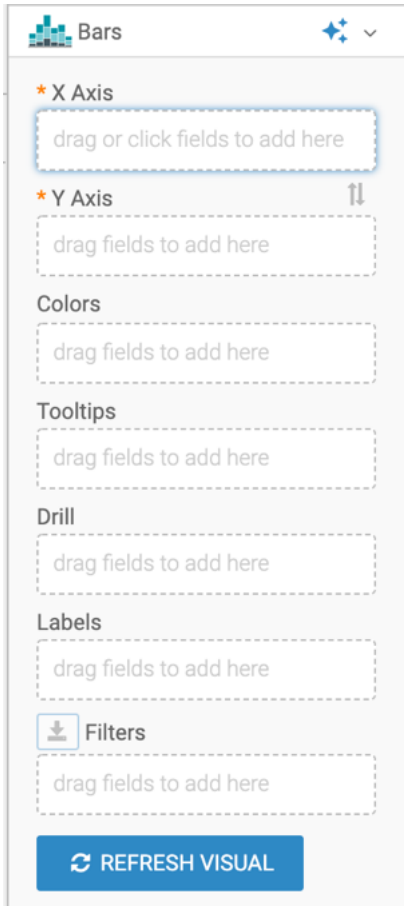
Procedure

1. Start a new visual based on the World Life Expectancy dataset. For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Bars.

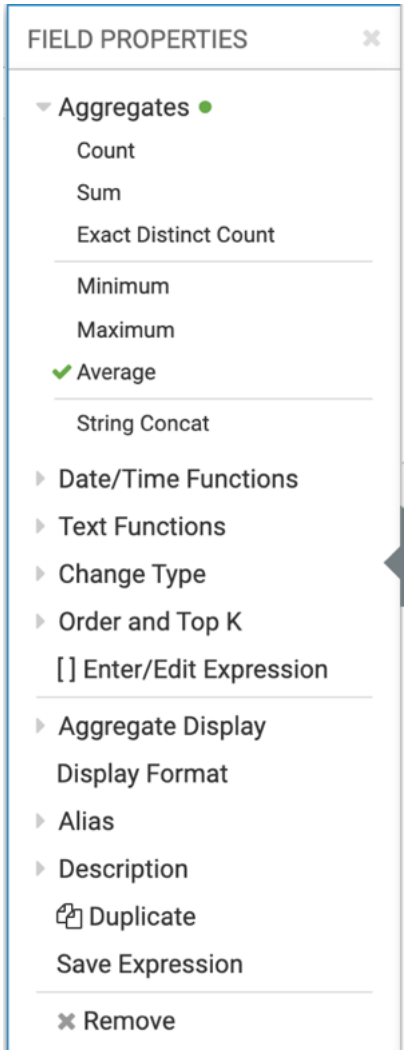


The shelves of the visual changed. They are now X Axis, Y Axis, Colors, Tooltips, Drill, Labels, and Filters.

The mandatory shelves are X Axis and Y Axis. Note also that the fields placed on these two shelves may be easily swapped by switching X and Y.

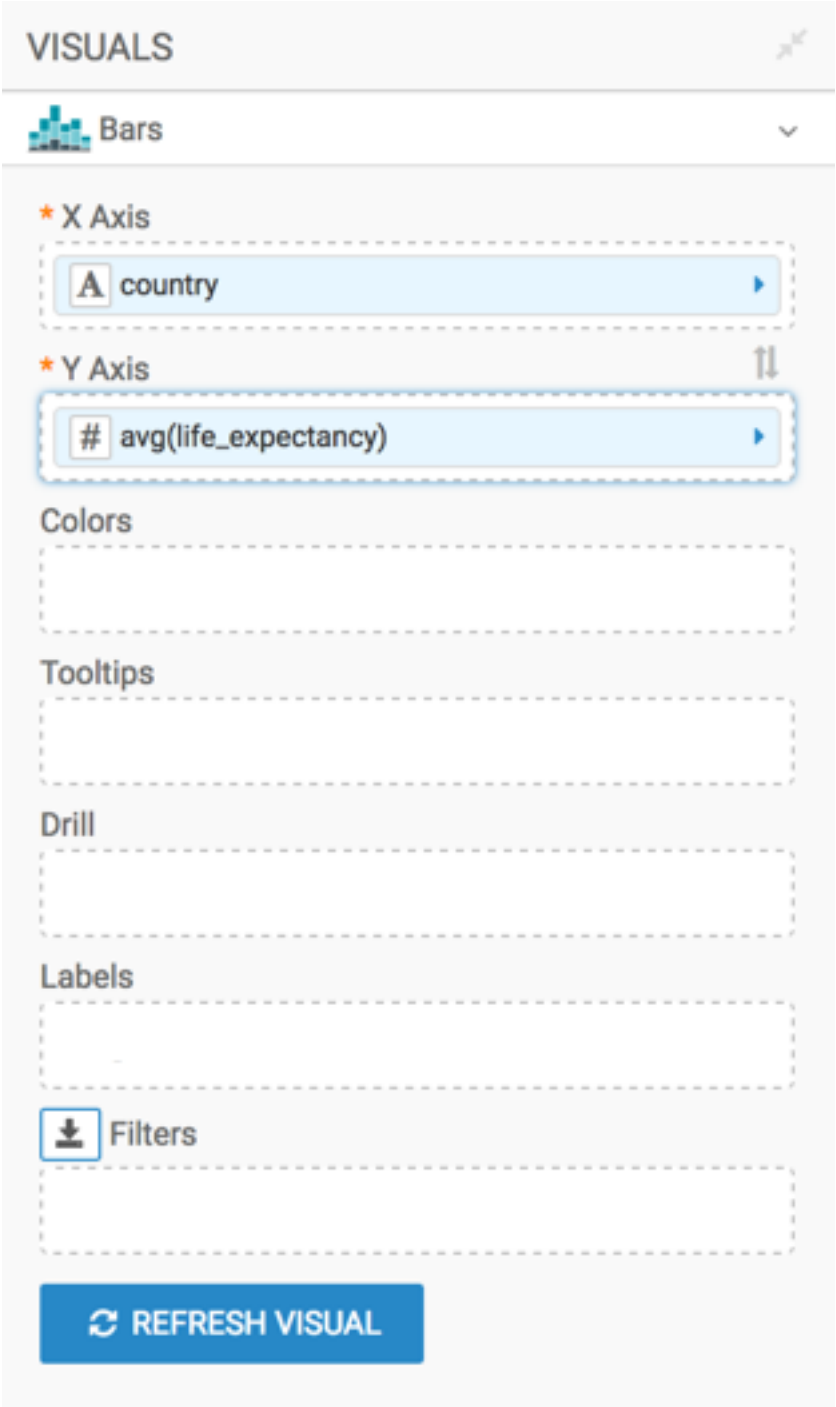


3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select country and drag it to the X Axis shelf .
   b) Under Measures, select life_expectancy and drag it to the Y Axis shelf .
   c) Change the aggregate of the life_expectancy field from sum(life_expectancy) to avg(life_expectancy):
      - Click the field on the shelf of the visual to open the FIELD PROPERTIES menu.
      - Open the Aggregates menu and select Average from the list of aggregate functions.



- Click the x icon at the top of the FIELD PROPERTIES menu to close it.

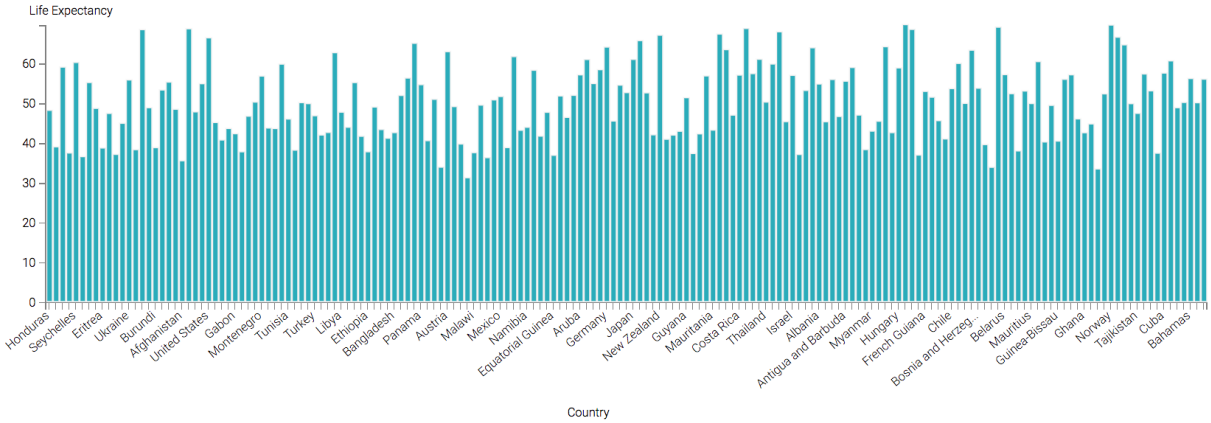
  The shelf now contains the modified field with avg() aggregation function.



4. Alias the country field as Country, and the life_expectancy field as Life Expectancy:
   - Click the field on the shelf of the visual to open the FIELD PROPERTIES menu.
   - Open the Alias menu and enter the alias name of column in the textbox as it should appear in the visual.
   - Click the x icon at the top of the FIELD PROPERTIES menu to close it.

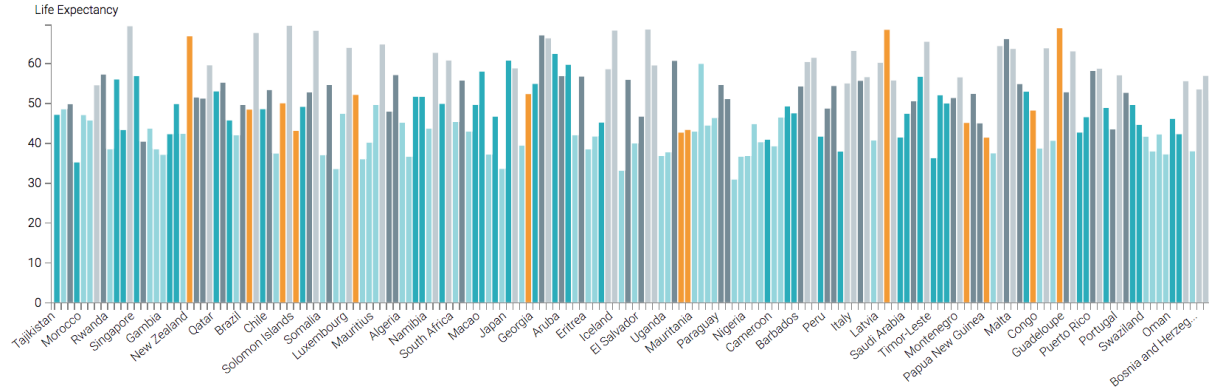
     Note that the shelf now shows the column with its alias name.
5. Click REFRESH VISUAL.

   The Bars visual appears.



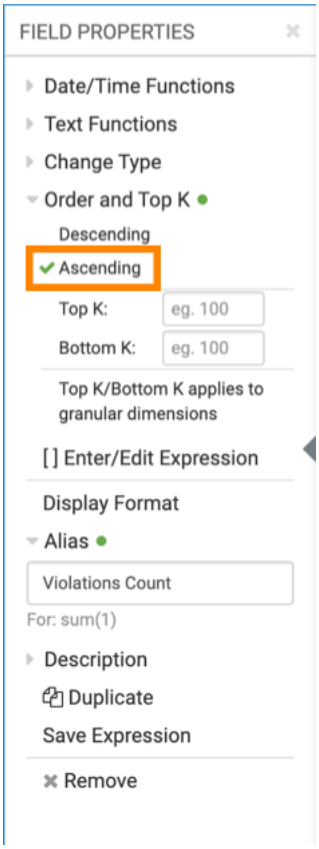
6. To see the bars as different colors, you can place a dimension to the Colors shelf.

Drag the un_region field to the Colors shelf and click REFRESH VISUAL to see the changes.



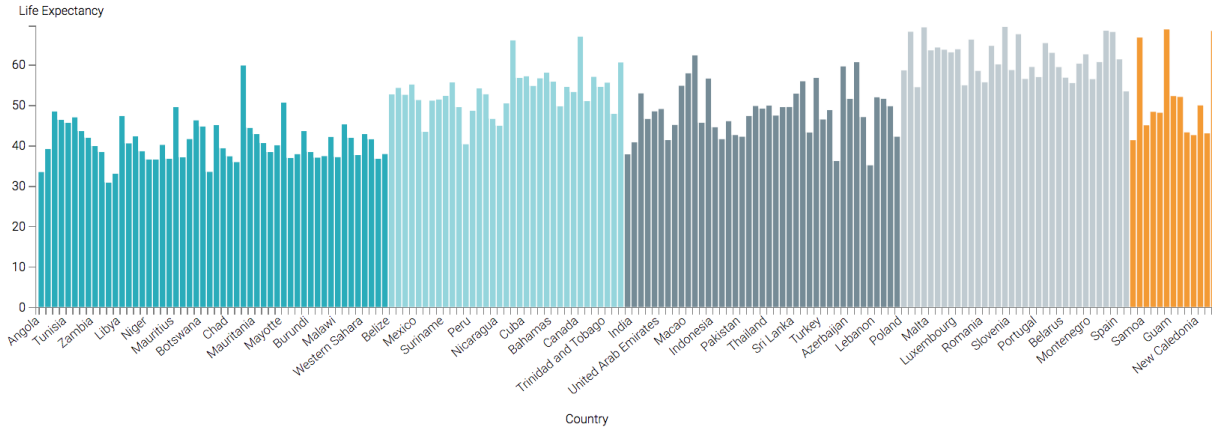
7. To group all countries of UN Regions together, change the order for the un_region field on the Colors shelf:

- Click the field on the shelf of the visual to open the FIELD PROPERTIES menu.
- Open the Order and Top K menu and select Ascending from the list of options.
-



- Click the x icon at the top of the FIELD PROPERTIES menu to close it.

8. Click REFRESH VISUAL to see changes.



9. On the Filters shelf, add several Dimensions and Measures from the DATA menu.

   This enables you to dynamically control the data input, and discover the data at a more granular level.

   For example, from Dimensions, drag country, un_subregion, un_region, and year onto the Filters shelf. The filter selection wizard appears each time you place a field on a Filters shelf. You can specify filter values at this time, or click CANCEL and select them later.

10. Change the title to 'World Population - Bars'.

    - Click (pencil) icon next to the title of the visualization to edit it and enter the new name.
    - [Optional] Click (pencil) icon below the title of the visualization to add a brief description of the visual.

11. At the top left corner of the Dashboard Designer, click SAVE.
12. At the top left corner of the Dashboard Designer, click CLOSE.

What to do next

After learning how to create a basic bar chart, you can explore the following additional uses.

Related Information

Changing the min/max of the aggregate scale

Filtering 1: Bar chart by year

When working with a very large set of data, it is necessary to quickly access and examine a specific subset of information. Cloudera Data Visualization gives you the ability to filter information on all fields of the dataset.

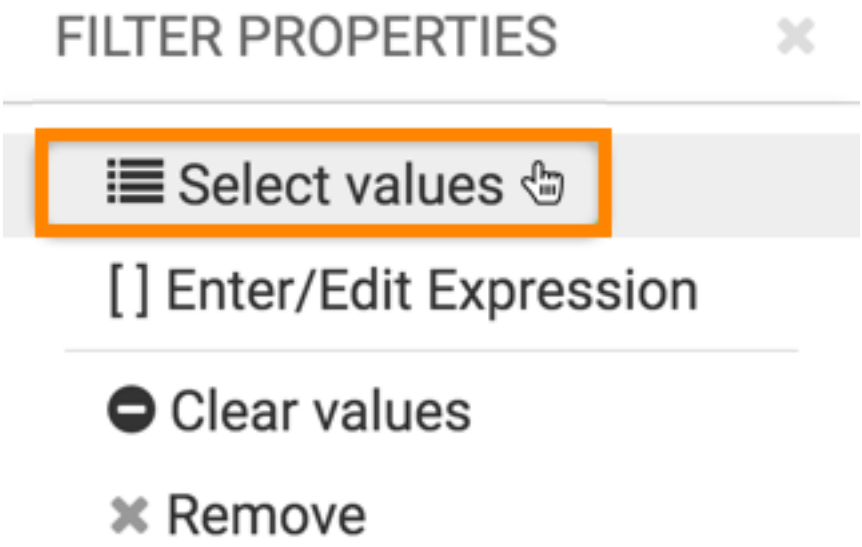
About this task

This example shows how to use such a filter to limit the result set and visualization to one year. It uses the visual previously created in Bars. Clone it and make the changes to the duplicate visual.

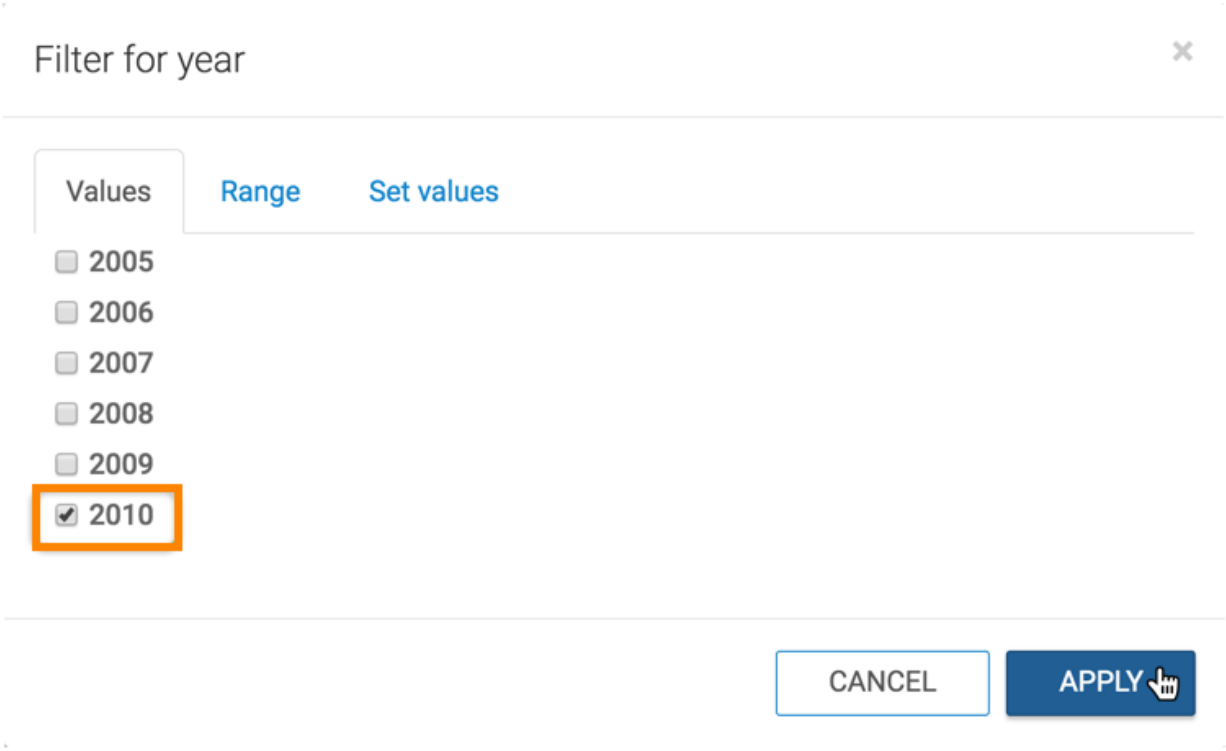
To use this visual to see the bar chart for a specific year such as 2010, follow these steps:

Procedure

1. Set the filter value for year to 2010.
   - Click the field on the shelf of the visual to open the FIELD PROPERTIES menu.
   - Click Select values.

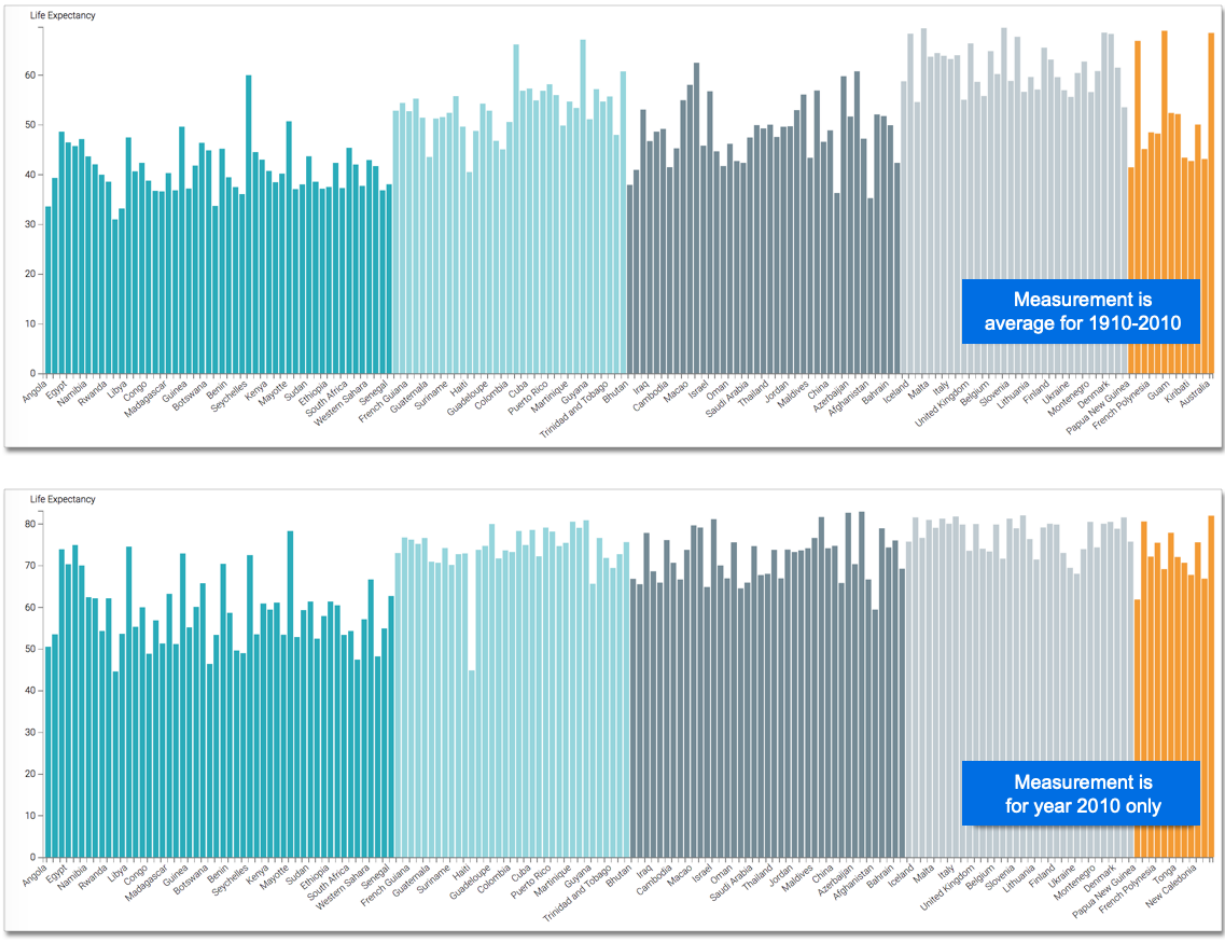


   - In the Filter for year modal window, select 2010 and click APPLY.



2. Click REFRESH VISUAL.

   The Bars visual for year 2010 has changed compared to the previous display, which was averaged over the 1900 - 2010 time span.



3. Rename the visual.
4. At the top left corner of the Dashboard Designer, click SAVE.

Filtering 2: bar chart by region

When working with a very large set of data, it is necessary to quickly access and examine a specific subset of information. Cloudera Data Visualization gives you the ability to filter information on all fields of the dataset.

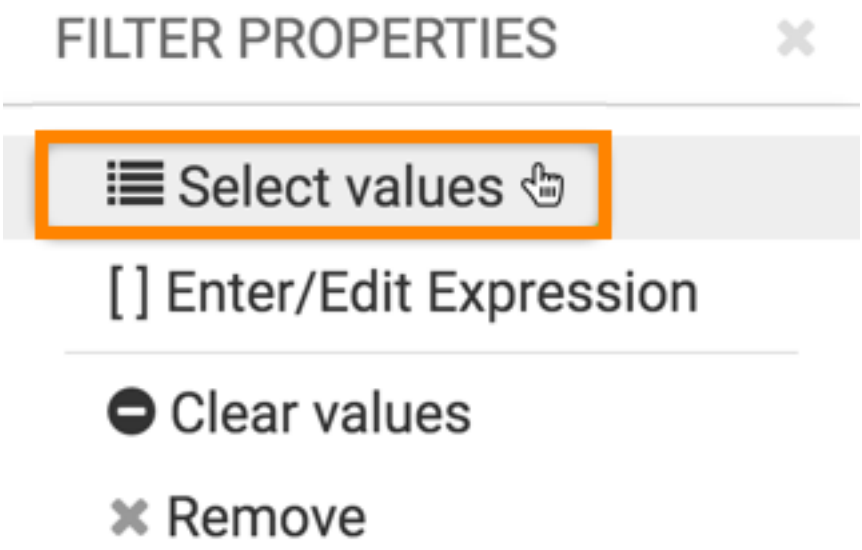
About this task

This example shows how to use such a filter to limit the result set and visualization to a particular geographic region. It uses the visual previously created in Bars. Clone it and make the changes to the duplicate visual.

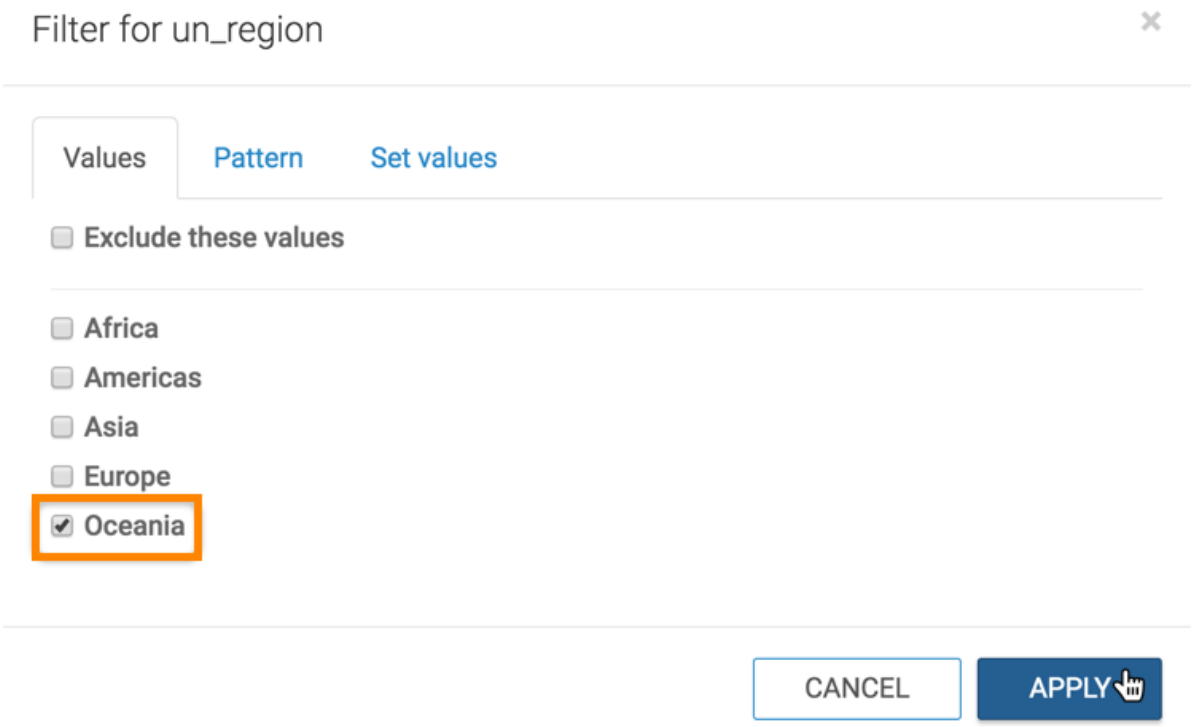
To use this visual to see a bar chart for a region, such as Oceania, follow these steps:

Procedure

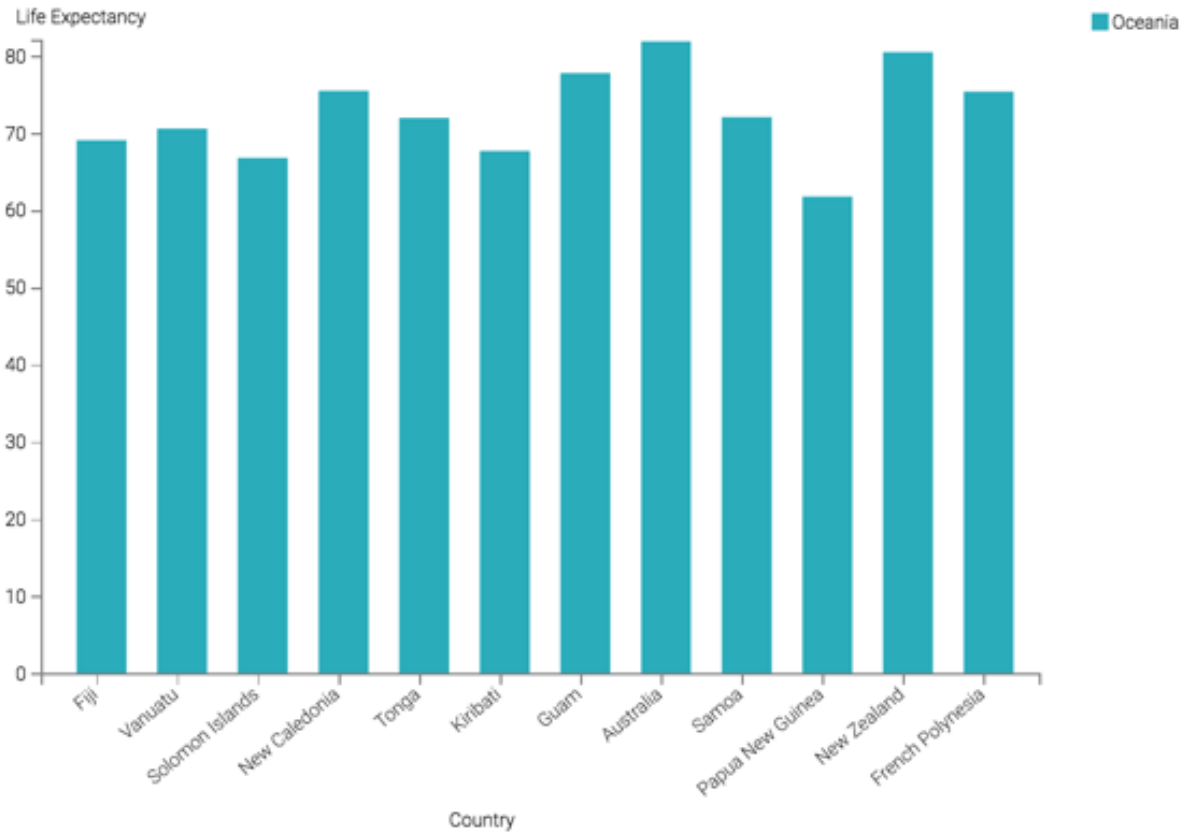
1. Add the field un-region to the Filters shelf, and set its value to Oceania.
   a) Click the field on the shelf of the visual to open the FIELD PROPERTIES menu.
   b) Click Select values.



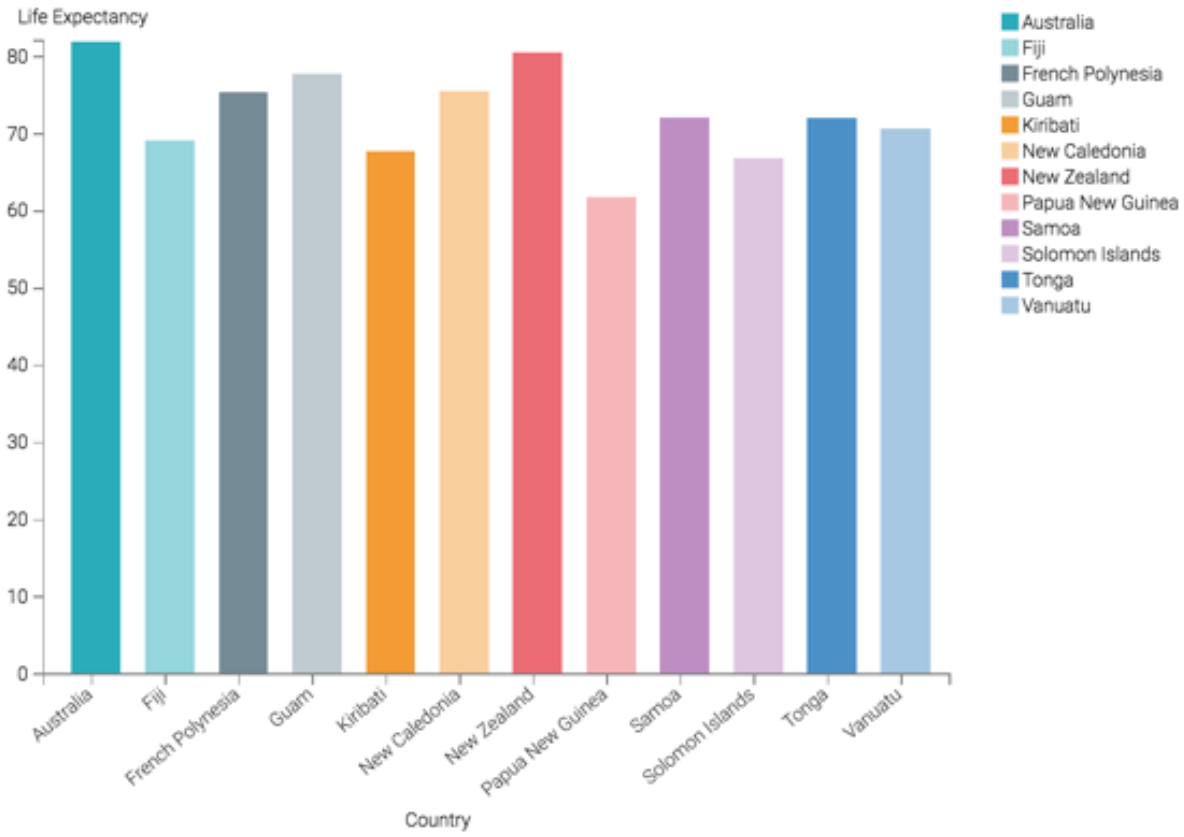
c) In the Filter for un_region modal window, select Oceania and click APPLY.



2. Click REFRESH VISUAL.
   The Bars visual changed to show only the countries in the Oceania UN Region.



3. Remove un_region from the Colors shelf, and replace it with the country field.
4. Click REFRESH VISUAL.



5. Rename the visual.
6. At the top left corner of the Dashboard Designer, click SAVE.

Horizontal bars

In general, dimensions of the dataset represent categories that are placed on the horizontal axis (and therefore the X shelf). The result set is rendered as a vertical bar chart. If your business requirements include the use of a horizontal bar chart, Cloudera Data Visualization makes it very easy to switch from an existing vertical bar chart to a horizontal one.

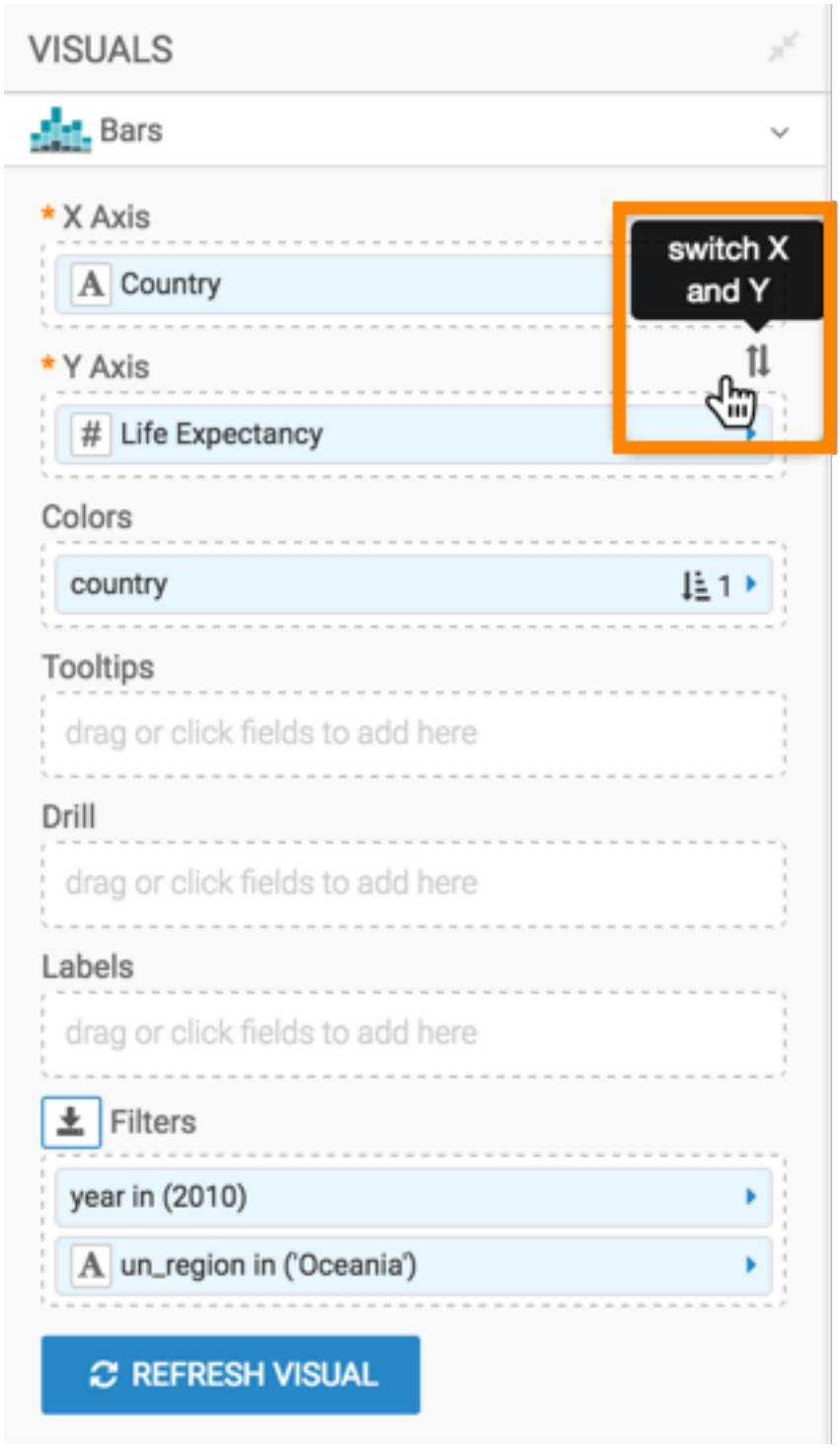
About this task

This example uses the visual previously created in Bars. Clone it, and make the changes to the duplicate visual.

To use this visual to create a horizontal bar chart, follow these steps.

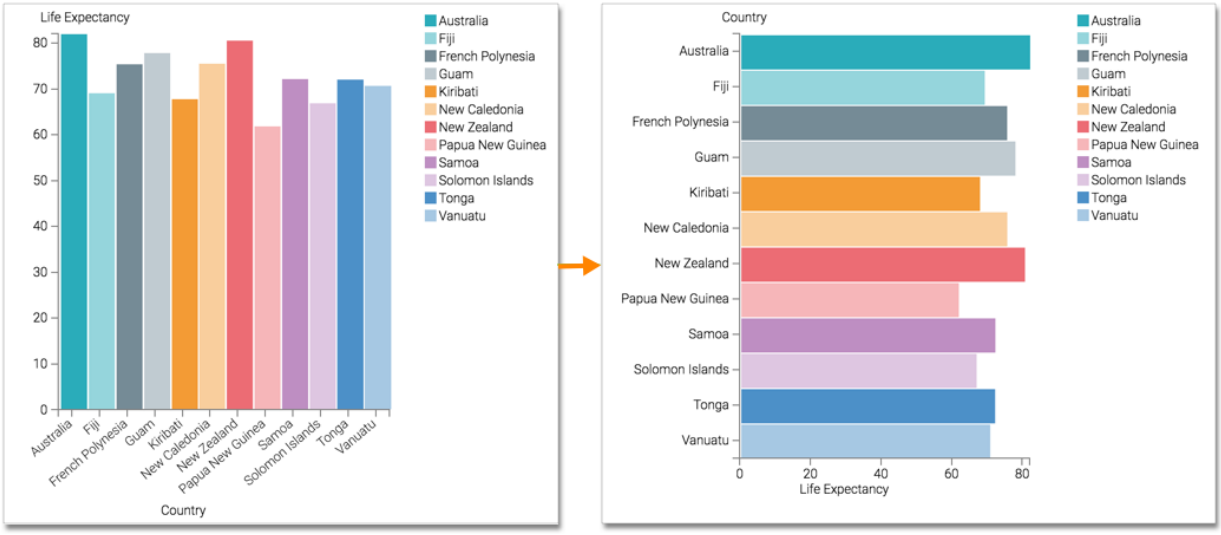
Procedure

1. Click the (switch) icon between the X Axis and Y Axis shelves.



2. Click REFRESH VISUAL.

   The graph has changed, with country categories on the Y axis and horizontal bars.



3. Rename the visual.
4. At the top left corner of the Dashboard Designer, click SAVE.

Stacked bars

While bar charts show comparisons among categories, Cloudera Data Visualization stacked bar charts show the bars divided into sub-parts. This demonstrates the cumulative effect of the component parts of a particular category.

About this task

This example uses the visual previously created in Bars. Clone it and make the changes to the duplicate visual.

To develop a stacked bar representation of European population divided by its subregions, follow these steps.

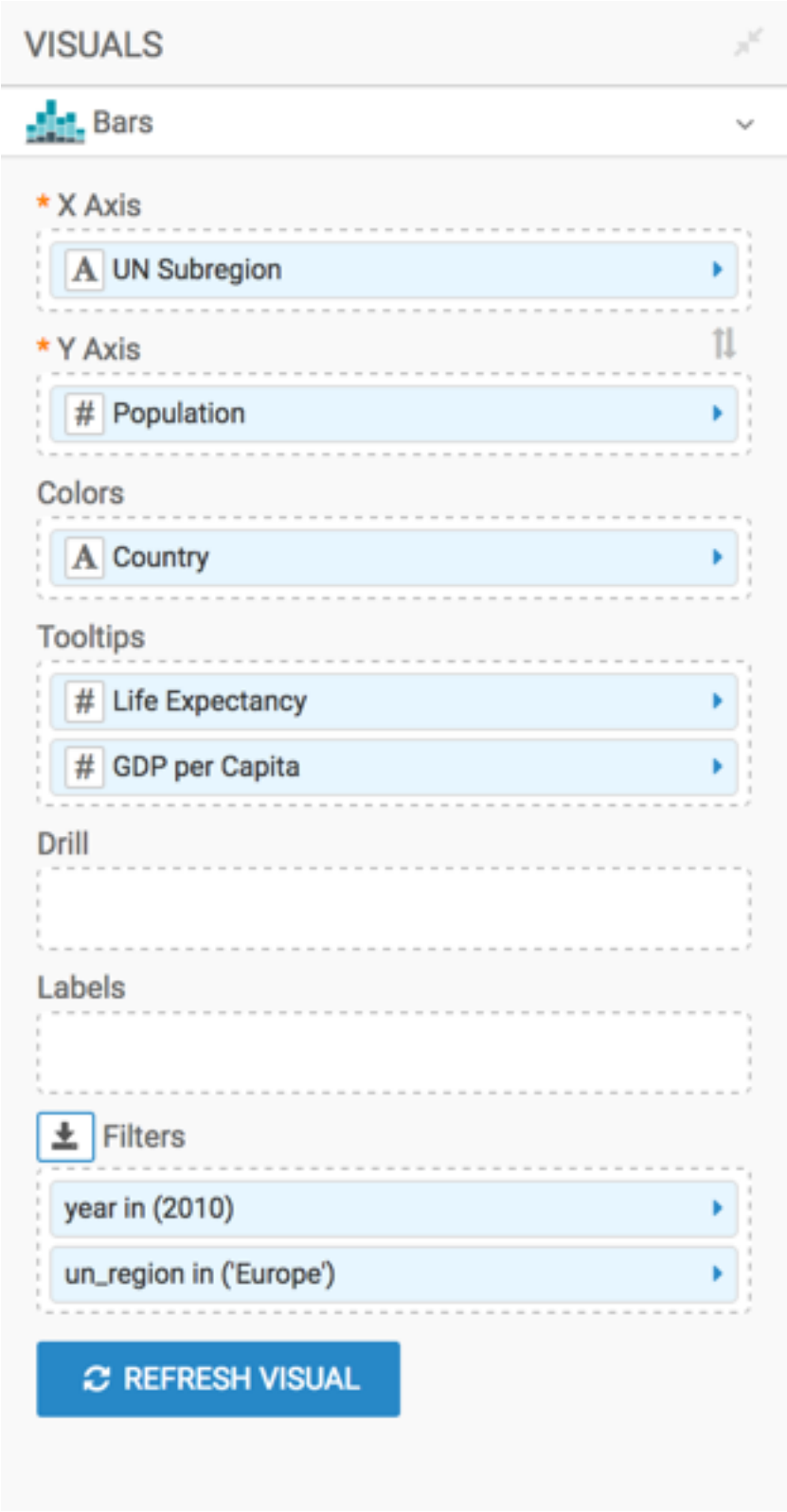
Procedure

1. Remove Life Expectancy from the Y Axis shelf and un_region from the Colors shelf.
2. From Dimensions in the DATA menu, drag un_subregion to the X Axis shelf.
3. Under Measures, select population and drop it to the Y Axis shelf.
4. Pull Country from the X Axis shelf to the Colors shelf.
5. On the Filters shelf, set year to 2010 and un_region to Europe.
6. Add all fields of interest to the Tooltips shelf.

   For example, add life_expectancy and gdp_per_capita.
7. Adjust Display Format for the values on the main shelves of the visual and in the Tooltip shelf.

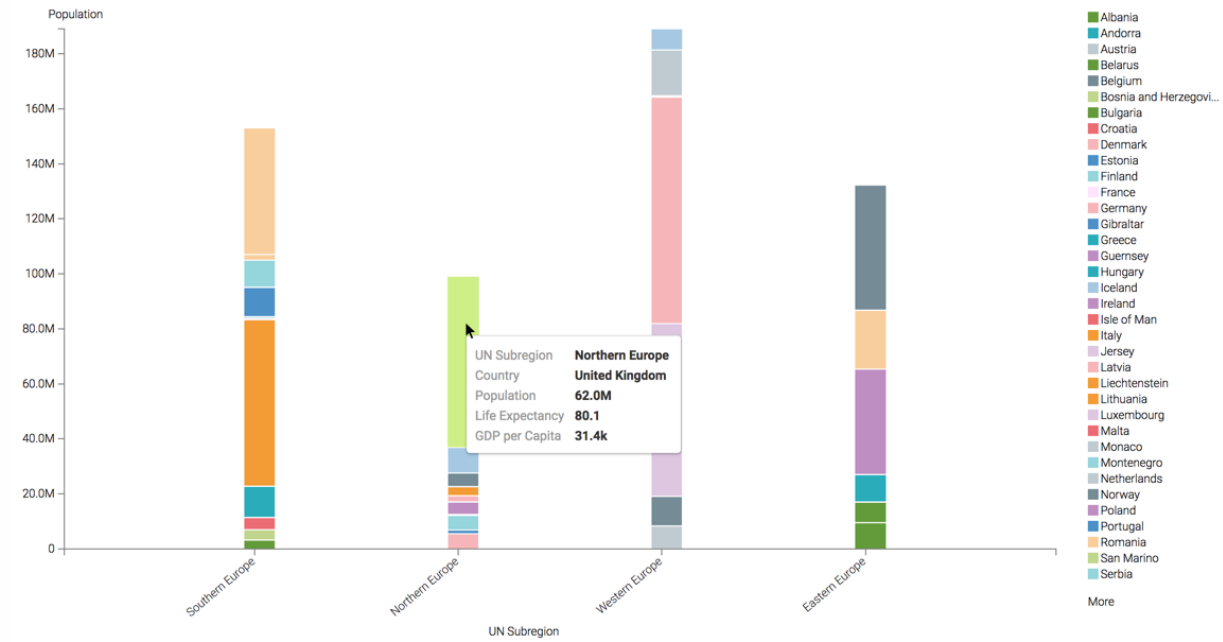
   For example, set the gdp_per_capita to Currency, with $ symbol.
8. Alias the fields on the shelves as UN Subregion, Population, GDP per Capita, and so on:
   a) Click the field on the shelf of the visual to open the FIELD PROPERTIES menu.
   b) Open the Alias menu and enter the alias name of column in the textbox as it should appear in the visual.
   c) Click the x icon at the top of the FIELD PROPERTIES menu to close it.

   The shelf now shows the column with its alias name.
9. Click REFRESH VISUAL.



The visual has changed. You can see the relative size of total populations for the subregions, and the component countries for all subregions.

10. If you hover the mouse over the graph, you can see the tooltip text for each country.



11. Rename the visual.

12. At the top left corner of the Dashboard Designer, click SAVE.

Labelled bars

In Cloudera Data Visualization, you can use the Labels shelf to display the measurement value. This way you can improve the comparison of relative measures on Bar visuals.

About this task

This example uses the visual previously created in Stacked bars. Clone it and make the changes to the duplicate visual.

To add labels to bar charts, follow these steps.

Procedure

1. Click inside the Labels shelf.

2. In the DATA menu, select the field for the label.

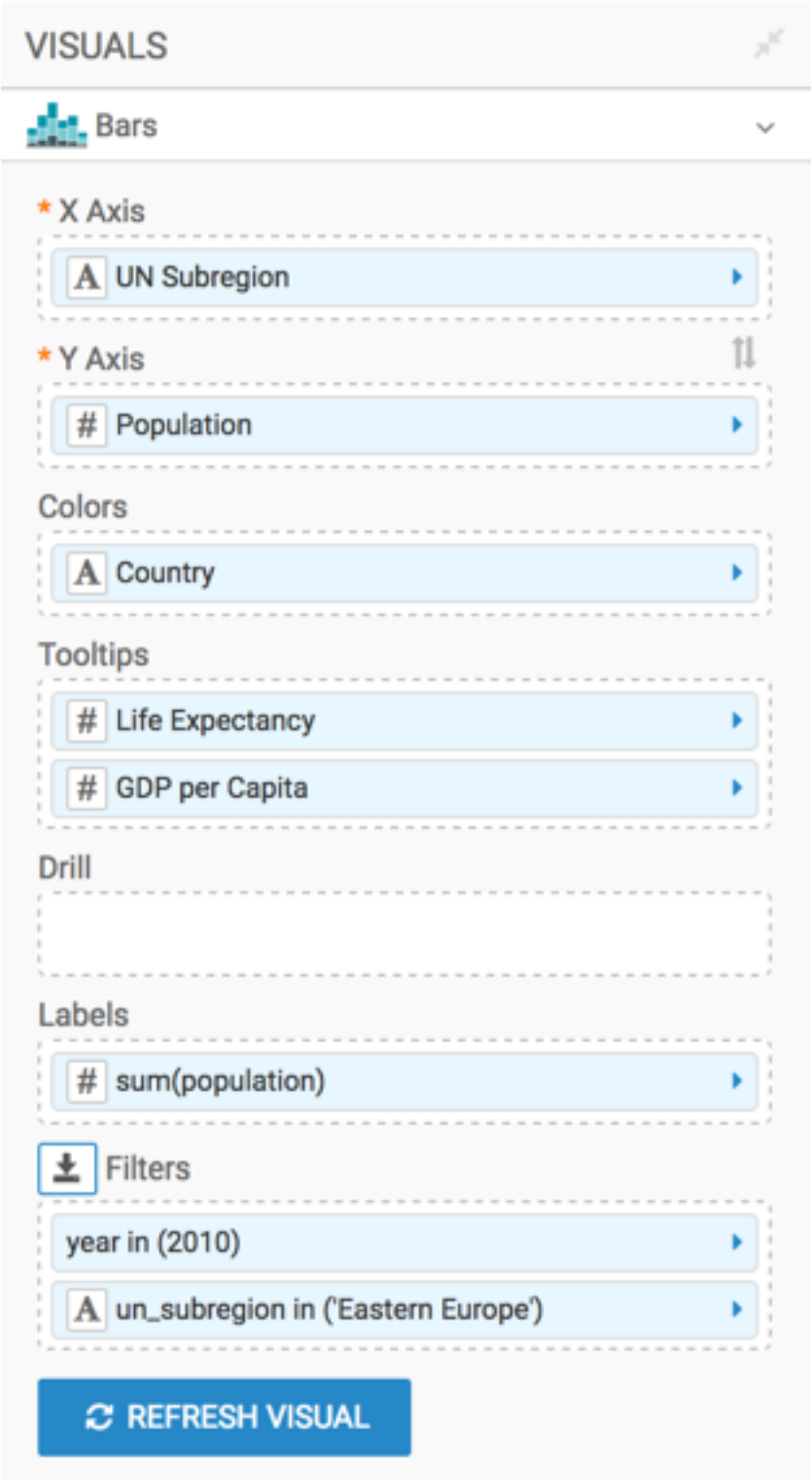
This example uses the population field.

3. Set the Filter shelf to hold the following fields:

- year in (2010)
- un_subregion in ('Eastern Europe')

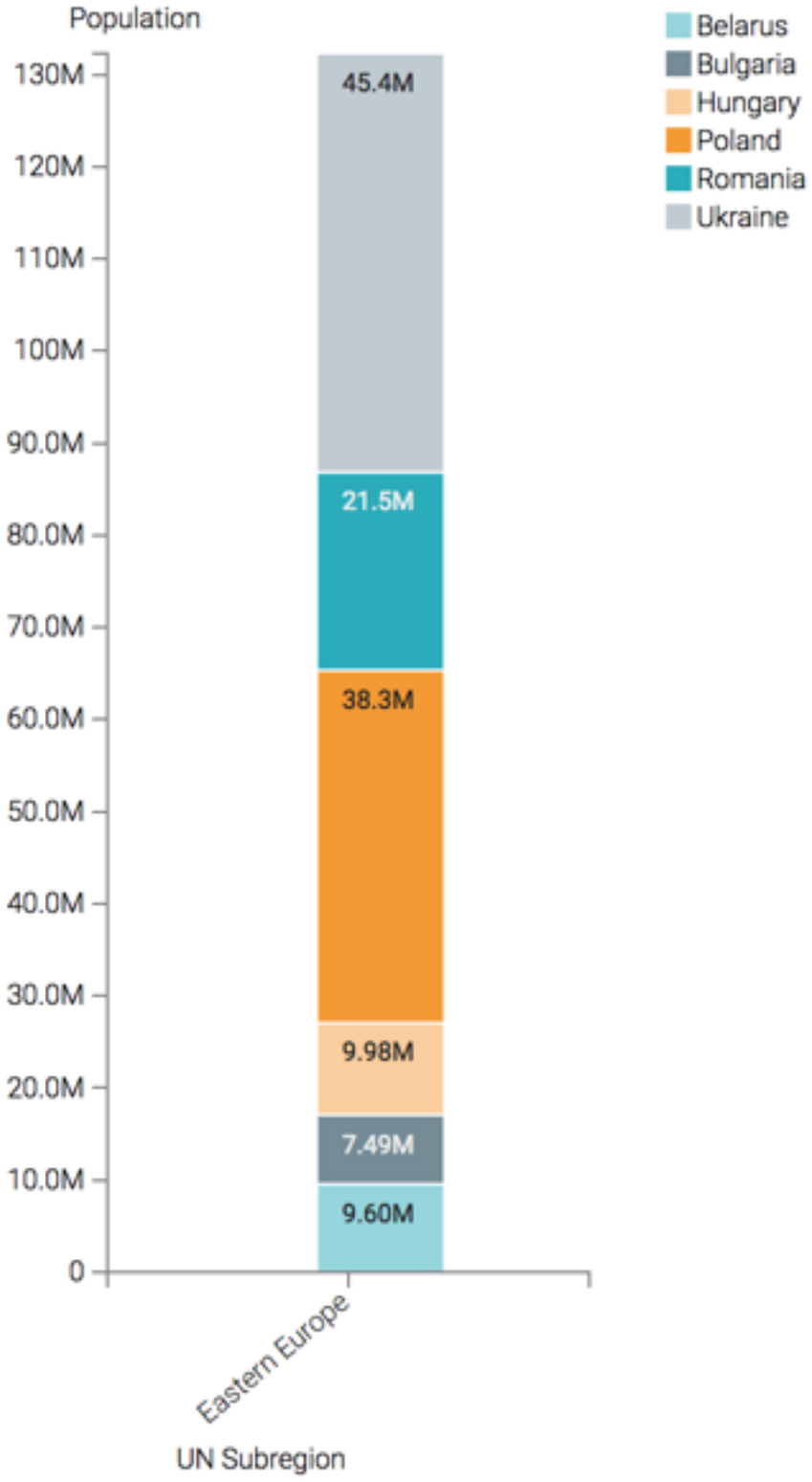
4. Increase the width of the bars as described in *Bar size range*.

Your shelf setup should look something like the following:



5. Click REFRESH VISUAL.

   The visual changed. You can see the relative size of total populations for the UN Subregion on one bar, and read off the value.



6. Rename the visual.
7. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for bars

Overview of shelves for Cloudera Data Visualization Bars visuals.

A Bars visual has the following shelves:

X Axis

This is conceptually the X axis. Use a Dimension for horizontal bars or a Measure for vertical bars.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in columns.

Y Axis

This is conceptually the Y axis. Use a Measure for horizontal bars or a Dimension for vertical bars.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in rows.

Colors

To stack bars by category, add a Dimension field to this shelf.

This shelf displays only one field on the graph. Additional fields appear only in the Tooltip.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

Drill

Add Dimension fields from the DATA menu to this shelf.

When Drill Into context menu is enabled but no fields are specified, you can drill into any dimension. Specifying the fields on this shelf limits drilling to these dimensions.

This shelf accepts multiple fields.

Labels

Additional data labeling.

This shelf accepts at most one field.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Lines

In Cloudera Data Visualization, a line chart displays information as a series of data points, or markers, connected by straight line segments. It reveals trends or change over time. Line charts show relationships within a continuous data set, and can be applied to a wide variety of categories. While line charts are similar to scatter plots, the line chart orders the measurements along the primary axis and it is often used to visualize a time series.

About this task

Note:

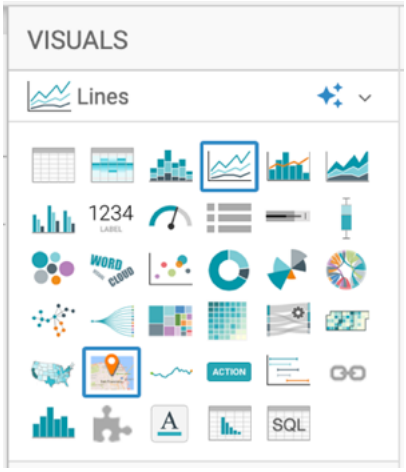
This visual type supports smooth rendering of streaming data.

The following steps demonstrate how to create a new Lines visual on the World Life Expectancy dataset.

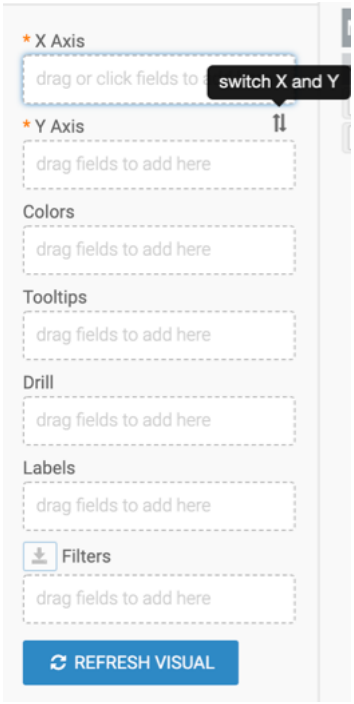
For an overview of shelves that specify this visual, see *Shelves for lines visuals*. For customization options, see *Settings for lines visuals*.

Procedure

1. Start a new visual based on dataset World Life Expectancy.
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Lines.

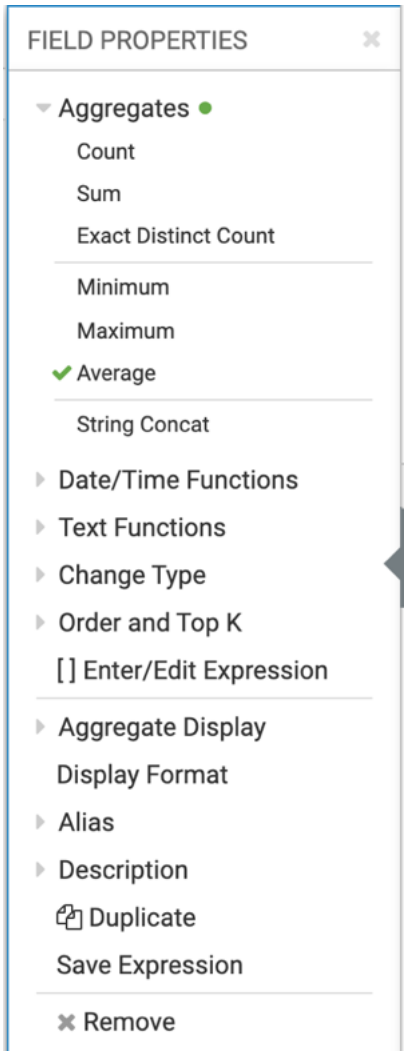


   The shelves of the visual changed. The mandatory shelves are X Axis and Y Axis.



   The fields placed on these two shelves may be easily swapped by switching X and Y.
3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   - Under Dimensions, select year and drag it to the X Axis shelf .
   - Click the Y Axis shelf and select the life_expectancy field from the DATA menu by clicking it. This adds the field to the shelf.
4. Click the field on the X Axis shelf, and in the FIELD PROPERTIES menu, select Order and Top K Ascending .

5. If the field on the Y Axis appears as a sum aggregate, change the aggregate to avg(life_expectancy):
   a) Click the field on the Y Axis shelf to open the FIELD PROPERTIES menu.
   b) Select Aggregates Average .

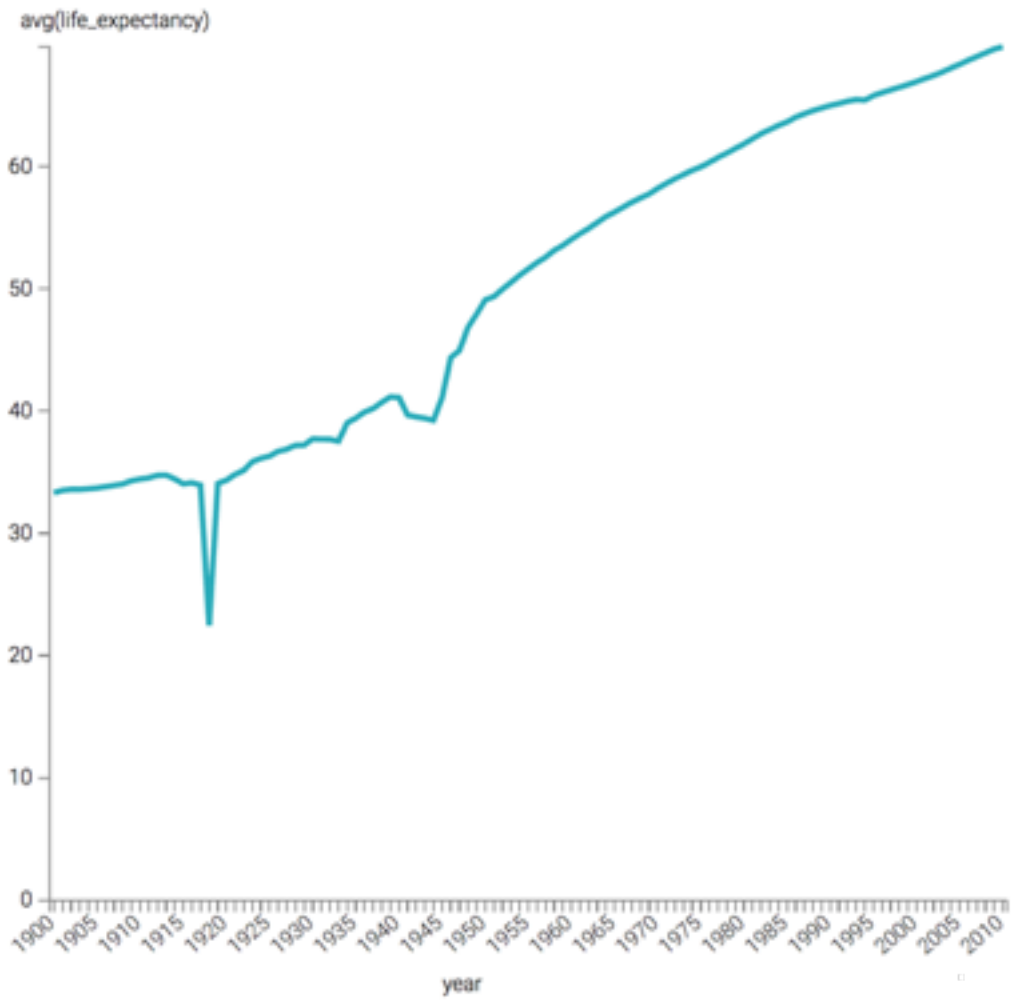


   c) Click the x icon at the top of the FIELD PROPERTIES menu to close it.

   The shelf now contains the modified field with avg() aggregation function.

   Alternatively, you can change the aggregation at the level of the dataset, as described in *Changing field aggregtion*.
6. Click REFRESH VISUAL.

   The Lines visual appears.



Note: You can change line thickness in the Settings Marks menu using the Line width (px) field. For more information, see

7. Alias the year field as Year, and the life_expectancy field as Life Expectancy:
   a) Click the field to open the FIELD PROPERTIES menu.
   b) Select Alias and enter the alias name of the column in the text box as it should appear in the visual.
   c) Click the x icon at the top of the FIELD PROPERTIES menu to close it.

   The shelf now shows the column with its alias name.
8. To see the individual countries in different colors, place the country field on the Colors shelf.

   To make the visual easier to understand, you can apply an *Alias* to it.
9. Click REFRESH VISUAL again.

   The single line is broken out into country components, and Cloudera Data Visualization automatically provides a color legend.

10. On the Tooltips shelf, add several fields from Dimensions and Measures.

This enables you to see specific descriptive information in your visuals, such as input values, segment affiliation, and calculations.

11. On the Filters shelf, add several Dimensions and Measures from the DATA menu.

This enables you to dynamically control the data input and discover the data at a more granular level.

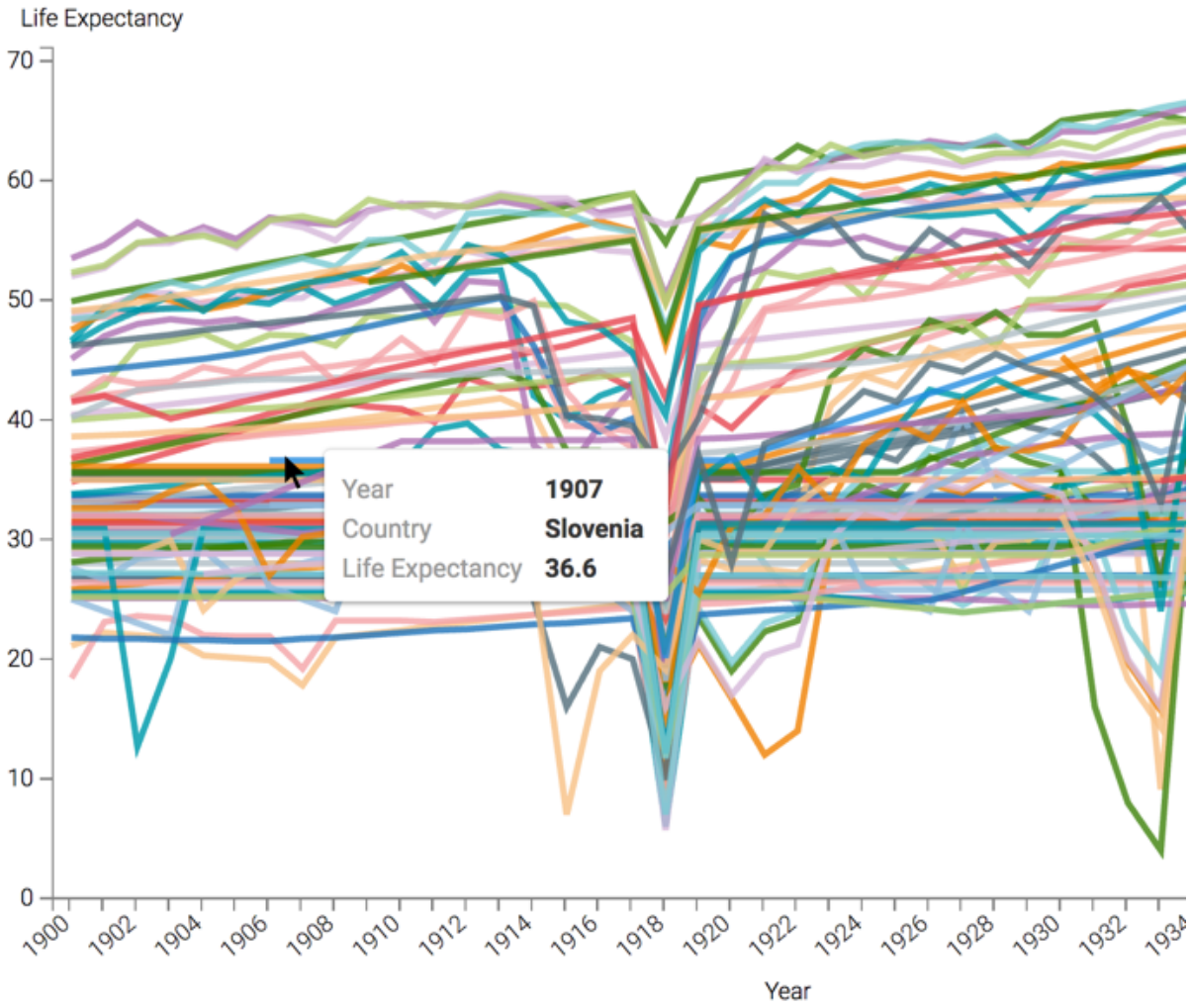
For example, from Dimensions, drag country, un_subregion, un_region, year, and population onto the Filters shelf.

12. On the Filters shelf:

a) Click the population field to open the FIELD PROPERTIES menu.
b) Select []Enter/Edit Expression.
c) Enter [population]>10000000.
d) Click VALIDATE & SAVE.

13. Click REFRESH VISUAL.

The updated Lines visual appears. Some of the lines do not originate at the beginning of the time series. This is because their population were lower than the 10,000,000 threshold that you set in the population filter.



14. To limit the dimensions available to the Drill Into feature (it appears on dashboards that include the visual), specify the fields that can remain visible by placing them on the Drill shelf.

15. Click the (pencil) icon next to the title of the visual to change the title to 'World Population - Lines'.

16. At the top left corner of the Dashboard Designer, click SAVE.

Related Information

Changing the min/max of the aggregate scale

Lines with dual axes

It is often enlightening to see related dimensions on the same line chart. It clearly shows both correlation and pattern. However, values may be measuring completely different things, or the range of values (scale) is too different to show both lines clearly. Cloudera Data Visualization offers a dual axis option for situations like these.

About this task

The following steps demonstrate how to create a Lines visual that uses a double axis. The example is based on the World Life Expectancy dataset.

Procedure

1. Start a new visual based on the World Life Expectancy dataset.

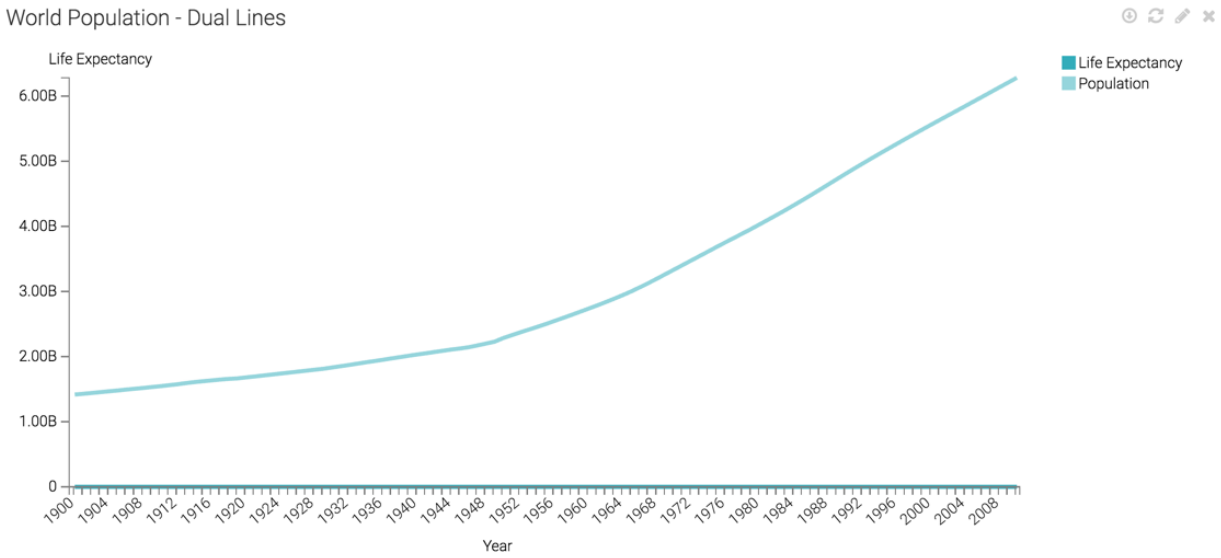
For instructions, see *Creating a visual*.

2. In the VISUALS menu, select the Lines visual type.
3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select year and place it on the X Axis shelf. Specify ascending order and alias the field as 'Year'.
   b) Under Measures, select life_expectancy and place it on the Y Axis shelf, and add an alias to the field.
   c) Under Measures, select Population and place it on the Colors shelf, and add an alias to the field.
4. Click REFRESH VISAL.

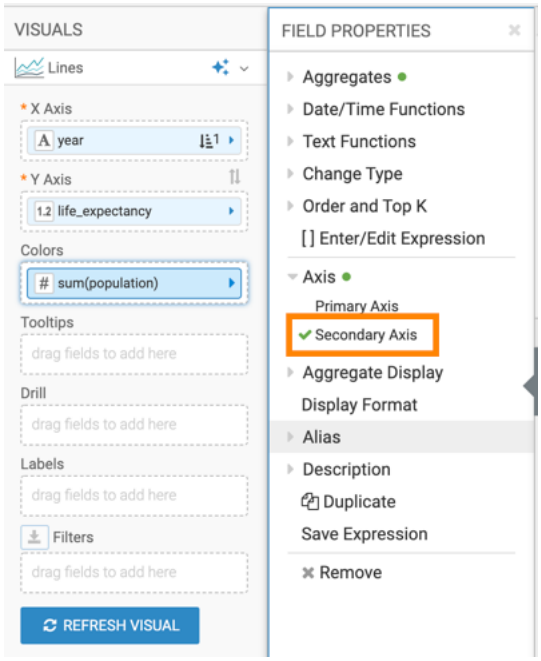
   The Lines visual appears. You can see the linear plot of population and life expectancy, on the same axis.

   It is impossible to see the values on the life expectancy line because the population line sets the scale for the Y axis.

   In this case, the label does not correspond to the scale shown, because this information comes from the Y axis shelf.



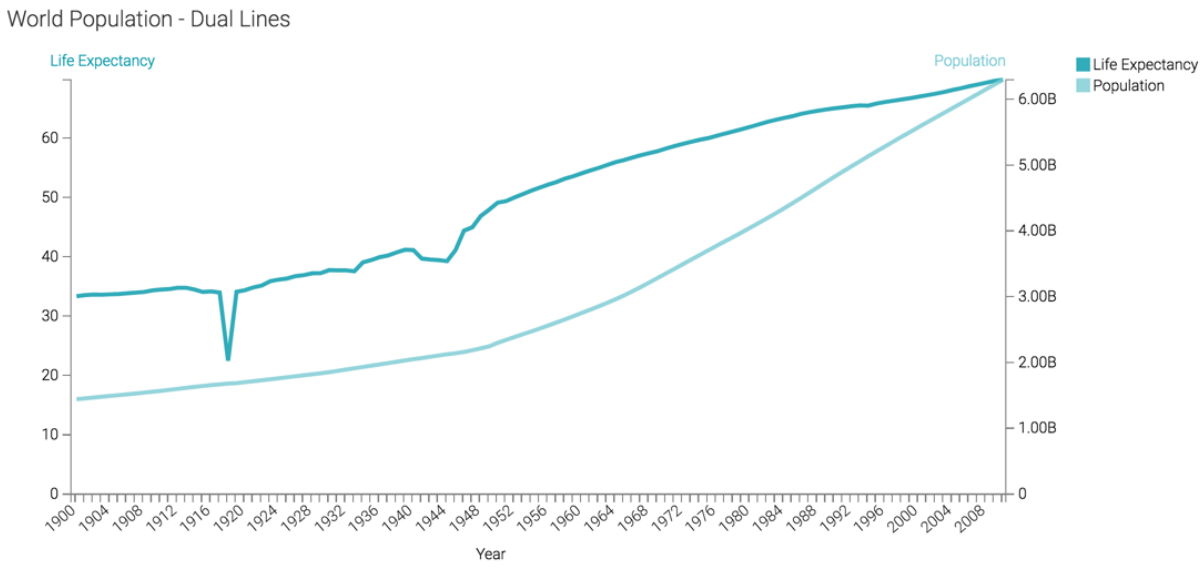
5. On the Colors shelf, click the sum(population) field to open the FIELD PROPERTIES menu.
6. Select Axis Secondary Axis .



7. Click REFRESH VISUAL.

   The two lines can now be seen clearly. There is a vertical axis both on the left and the right side of the visual. The linear plot of population and life expectancy are visible each on its own axis.

   For convenience, the colors of the axis labels match the line.



8. Save the visual as World Population - Dual Lines.

What to do next

Compare this approach with plotting multiple measures as trellised charts described in *Trellis on measures*.

Shelves for lines

Overview of shelves for Cloudera Data Visualization Line visuals.

A Lines visual has the following shelves:

X Axis

This is conceptually the X axis. Use a Dimension for a horizontal line or a Measure for a vertical line.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in columns.

Y Axis

This is conceptually the Y axis. Use a Measure for a horizontal line or a Dimension for a vertical line.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in rows.

Colors

To stack lines by category, add a Dimension field to this shelf.

This shelf displays only one field on the graph. Additional fields appear only in the Tooltip.

Tooltips

Add Measure fields from the DATA menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

Drill

Add Dimension fields from the DATA menu to this shelf.

When Drill Into context menu is enabled but no fields are specified, you can drill into any dimension. Specifying the fields on this shelf limits drilling to these dimensions.

This shelf accepts multiple fields.

Labels

Additional data labeling.

This shelf accepts at most one field.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Combined bar/line visuals

In Cloudera Data Visualization, Combined Bar/Line (Combo) visuals unite bar and line plots to enable easy comparison of trends and measures across the same dimensions.

Note: This visual type supports smooth rendering of streaming data.

The bars and the line can share the same vertical axis, or use a secondary axis. For more information, see Creating basic combo visuals and Assigning a measure to a secondary axis.

You can optionally render the line portion of the visual as an area. For more details, see Change the line to an area.

Related Information

Changing the min/max of the aggregate scale

Creating basic combos

Cloudera Data Visualization enables you to create Combined Bar/Line (Combo) visuals.

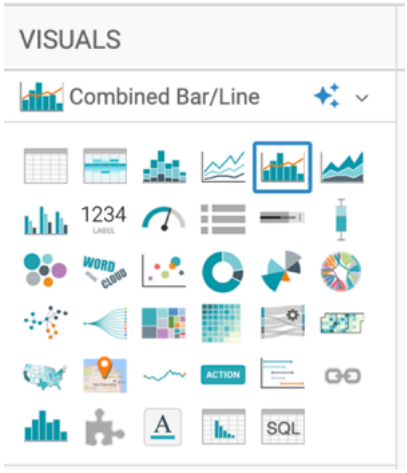
About this task

The following steps demonstrate how to create a new combined bar and line chart .

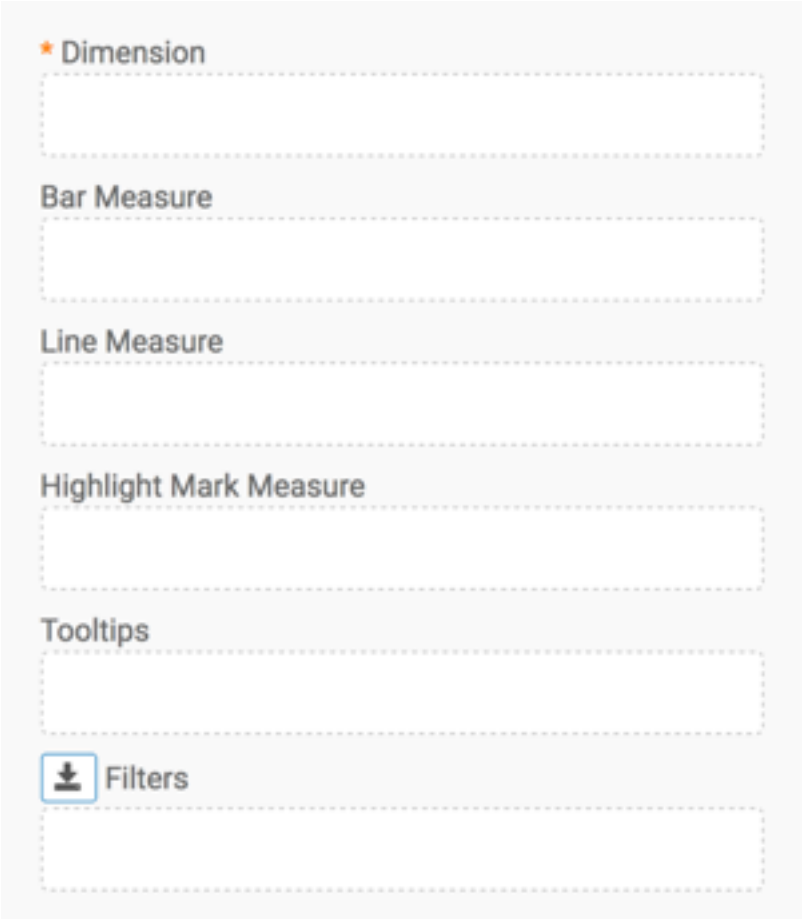
For an overview of shelves that specify this visual, see *Shelves for combos*.

Procedure

1. Start a new visual based on the World Life Expectancy dataset.
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Combined Bar/Line.



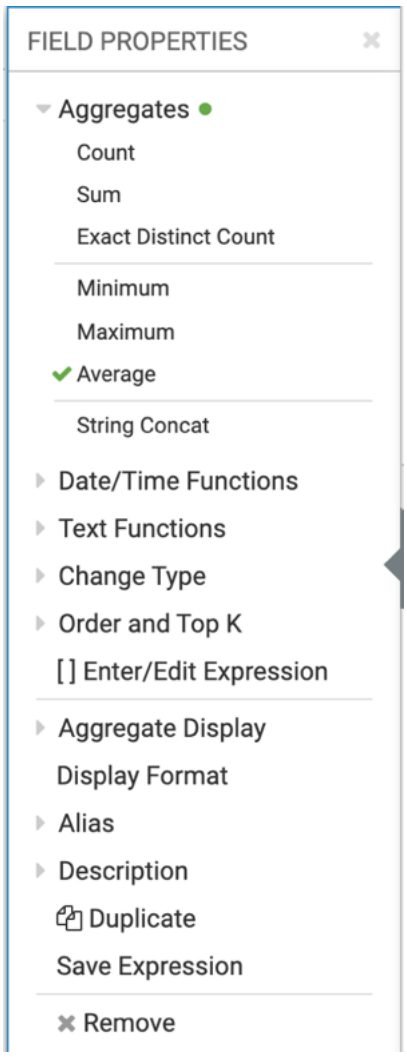
The shelves of the visual changed.



The mandatory shelf for Combined Bar/Line visuals is Dimension that maps to the X-axis.

Bar Measure and Line Measure both map to the Y-axis.

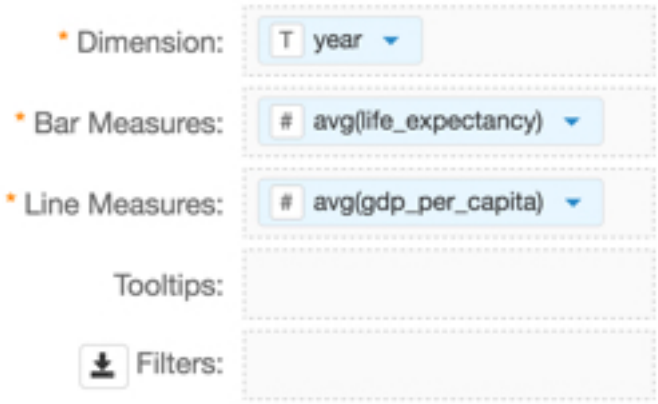
3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select year and drag it to the Dimension shelf.
   b) Under Measures, select life_expectancy and drag it to the Bar Measures shelf.
   c) Under Measures, select gdp_per_capita and drag it to the Line Measures shelf.
4. Change the aggregate of the life_expectancy field from sum(life_expectancy) to avg(life_expectancy):
   a) Click the life_expectancy field to open the FIELD PROPERTIES menu.
   b) Select Aggregates Average .



c) Click the x icon at the top of the FIELD PROPERTIES menu to close it.

5. Similarly, change the aggregate of the gdp_per_capita field from sum(gdp_per_capita) to avg(gdp_per_capita):
   a) Click the gdp_per_capita field to open the FIELD PROPERTIES menu.
   b) Select Aggregates Average .
   c) Click the x icon at the top of the FIELD PROPERTIES menu to close it.

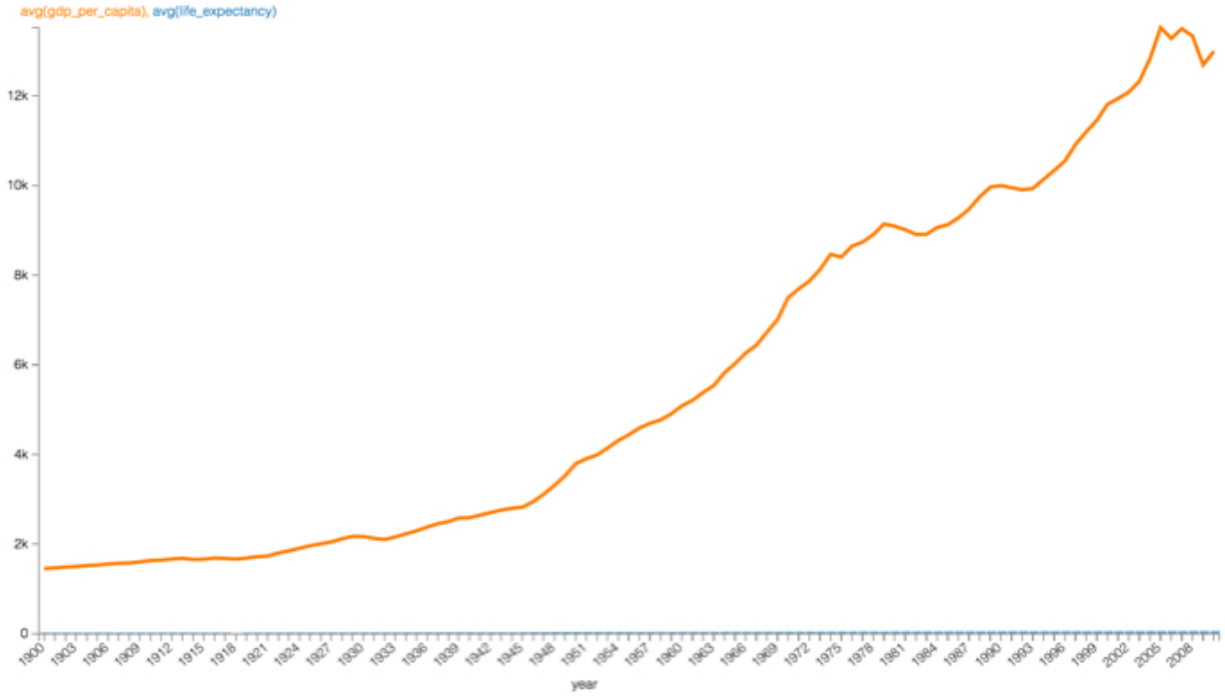
   The shelves of your visual should now look similar to the following image:



6. Click REFRESH VISUAL.

The combined bar and line visual appears.

The two legends for the Y axis match the color of the plot: the orange line represents sum(gdp_per_capita) and the blue bars represent avg(life_expectancy). However, you may not be able to see the bars because the values of that measure are much smaller than the values for the line measure. To see the bars, *assign a measure to a secondary axis*.



7. Alias the life_expectancy field as Life Expectancy, and the gdp_per-capita field as GDP Per Capita:
   a) Click the fields to open the FIELD PROPERTIES menu.
   b) Select Alias and enter the alias name of the column in the text box as it should appear in the visual.
   c) Click the x icon at the top of the FIELD PROPERTIES menu to close it.
8. Click REFRESH VISUAL again.

Assigning a measure to a secondary axis

Sometimes plotting two measures on the same axis does not show the bars because of their relative values, or because of their incompatible units of measurement. To be able to clearly see both a line and the bars in a Combined Bar/Line (Combo) visual, Cloudera Data Visualization enables you to specify a separate axis for each of the measures in the combined visual.

Before you begin

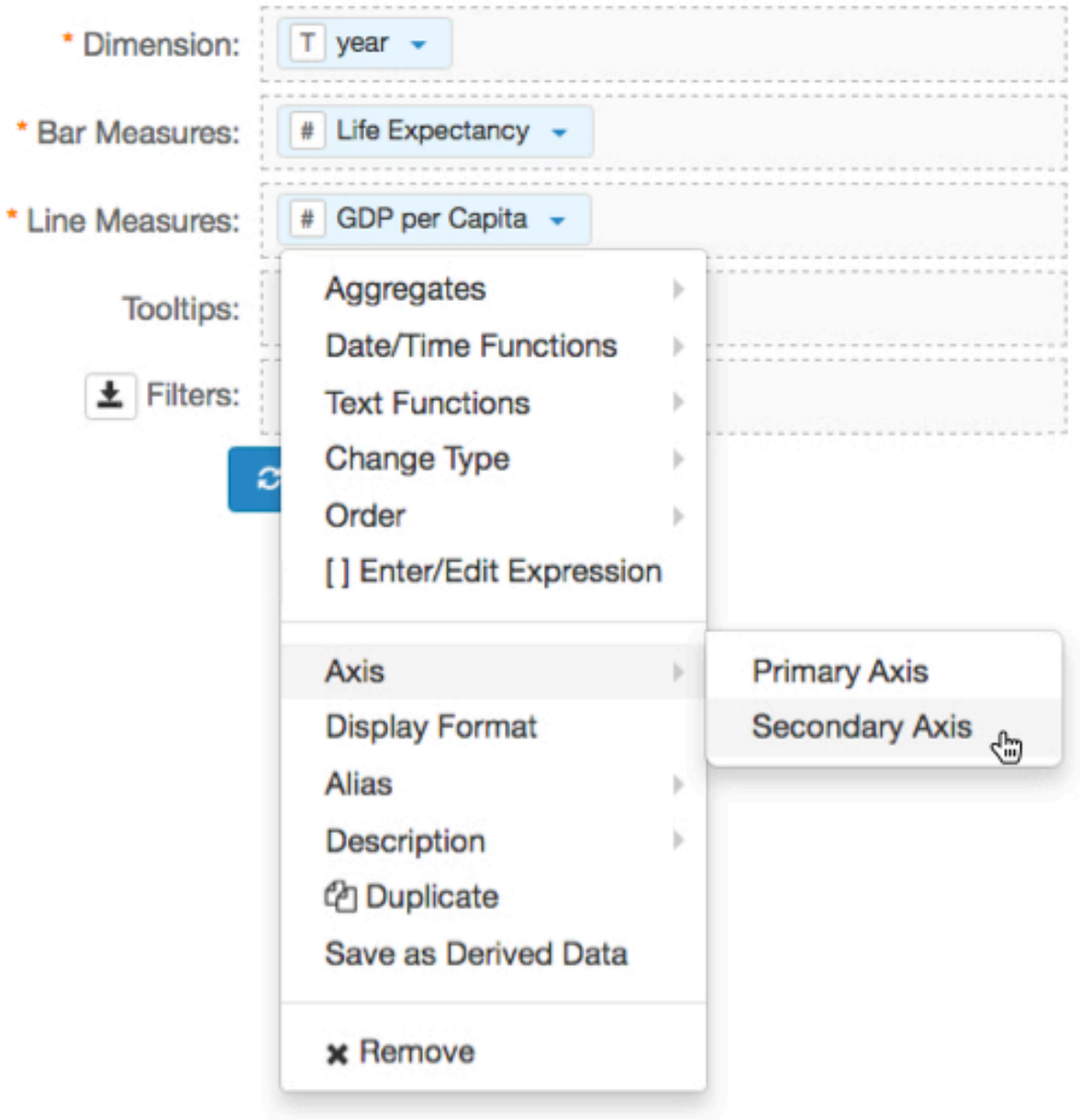
You have a already created a basic Combo visual based on the World Life Expectancy dataset.

About this task

Follow these steps to specify a different axis for each of the measures in your visual:

Procedure

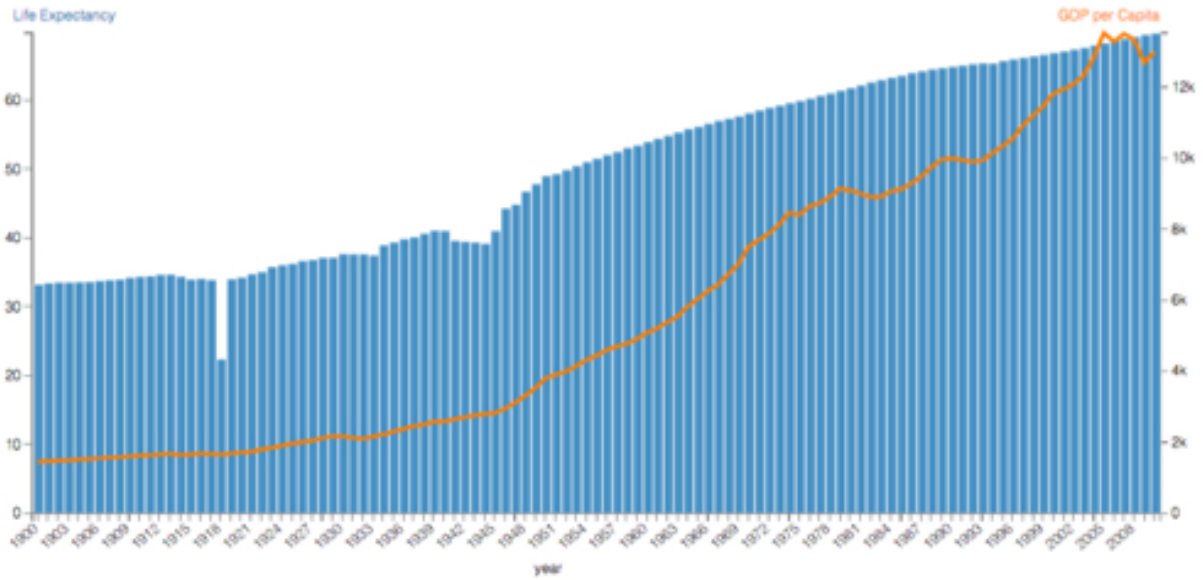
1. Assign the 'GDP per Capita' field to the secondary axis:
   a) Click the field to open the FIELD PROPERTIES menu.
   b) Select Axis Secondary Axis .



2. Click REFRESH VISUAL.

   Both the blue bars and the orange line are now clearly visible.

   The graph lists Life Expectancy on the left-hand vertical axis, in the range of 0 to 70, and the GDP per Capita on the right-hand vertical axis, in the range of 0 to 14k.



3. Click the (pencil) icon next to the title of the visual to edit it, and change the title to 'World Population - Dual Bar and Line'.
4. At the top left corner of the Dashboard Designer, click SAVE.

Change line to area

Cloudera Data Visualization allows you to convert the line to a semi-transparent area in Combined Bar/Line (Combo) visuals for better representation of data.

The default set-up is a combination of a bar measurement and a line measurement. To convert the line to a semi-transparent area that allows you to see the bar chart too, use the VISUAL SETTINGS interface, as described in *Converting line to area visual*.

Shelves for combos

Overview of shelves for Cloudera Data Visualizationn Combined Bar/Line (Combo) visuals.

A Combo visual has the following shelves:

Dimension

This is conceptually the X axis. Use a Dimension field to create bars and lines/areas for aggregate values in this dimension.

This shelf accepts a single field.

Bar Measure

This is conceptually the Y axis. Use a Measure field from the DATA menu to display aggregate values as bars.

This shelf accepts a single field.

Line Measure

This is conceptually the Y axis. Use a Measure field from the DATA menu to display aggregate values as lines or areas.

This shelf accepts a single field.

Highlight Mark Measure

This is conceptually the Y axis. Use a Measure field from the DATA menu to highlight aggregates of this measure.

This shelf accepts a single field.

Tooltips

Add Measure fields from the DATA menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Areas

An Area visual shows how values change over time. Use it to show how several series contribute to a total, or to compare the values and trends of individual series over time. For example, you can compare population trends across countries or see how those countries contribute to a combined total.

About this task

When you add a field to the Colors shelf, the visual creates a separate area for each value of that field. You can choose how those areas are displayed:

- **Stacked** (default): The areas sit on top of one another. The Y-axis shows their cumulative total, making it easier to see how each series contributes to the whole.
- **Unstacked**: Each area starts from the same zero baseline and can overlap with other areas. The Y-axis and tooltips can show each series' individual value, rather than a cumulative total.

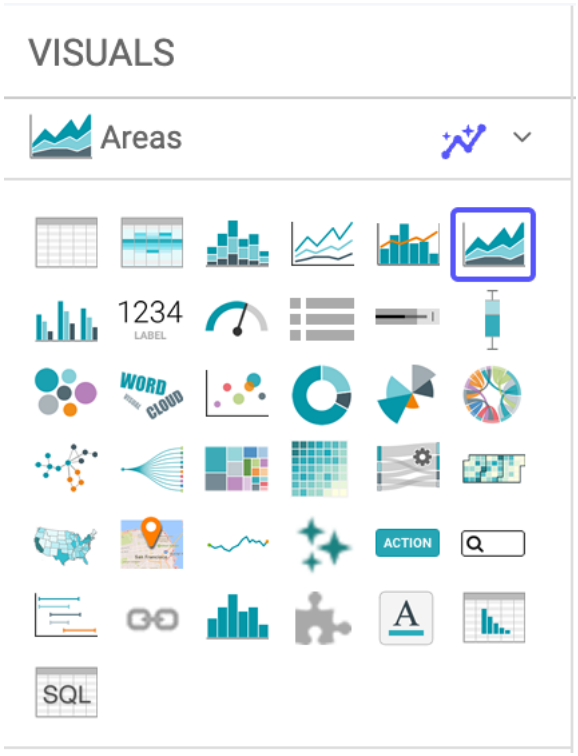
Note: Areas visuals support smooth rendering of streaming data.

The following steps show how to create an Areas visual. This example uses the World Life Expectancy from the samples connection. For an overview of shelves that specify this visual, see *Shelves for areas*.

Procedure

1. Create a new visual.

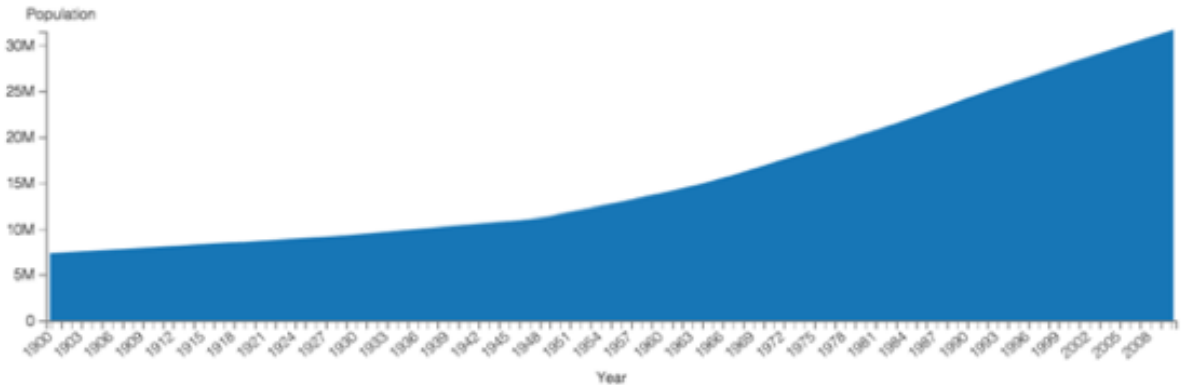
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, select Areas.



The shelves of the visual changed. The mandatory shelves are X Axis and Y Axis. To swap the fields on these shelves, click ⇅

3. Populate the shelves from the Dimensions and Measures fields available in the DATA menu.
   a) Drag year from Dimensions to the X Axis.
   b) Drag population from Measures to the Y Axis.
4. Set the aggregation of population to Average:
   a) Click population on the Y Axis to open FIELD PROPERTIES.
   b) Select Aggregates Average .
   c) Close FIELD PROPERTIES and click REFRESH VISUAL.

   The Areas visual appears.

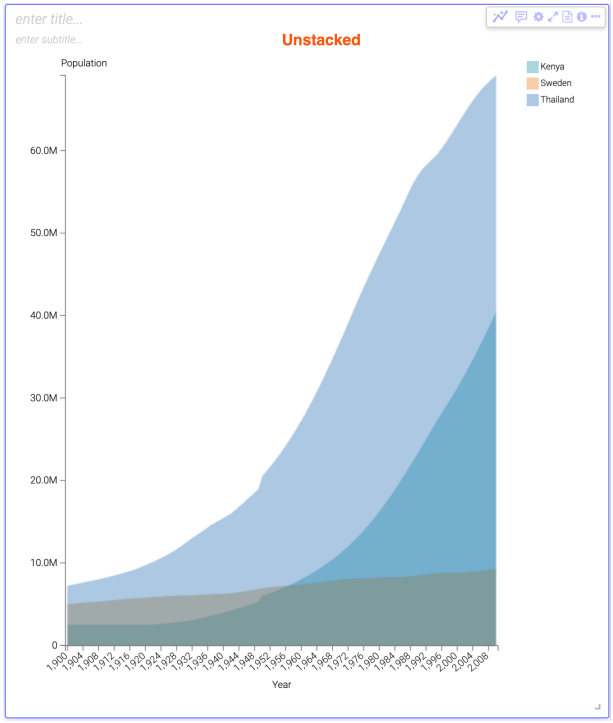


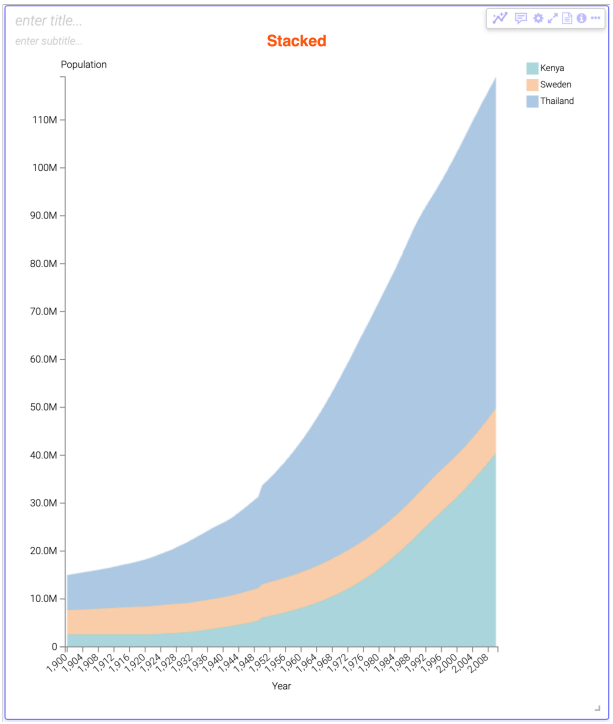
5. Add clearer field labels.
   a) Click year to open FIELD PROPERTIES.
   b) Under Alias, enter Year.
   c) Repeat for population on the Y Axis, entering Population.

d) Close FIELD PROPERTIES and click REFRESH VISUAL.

The shelves now show the fields with their alias names.

6. Drag country from Dimensions to the Colors shelf to create a separate area for each country.
7. To show only selected countries, drag country from Dimensions to the Filters shelf and set the filter settings to include only Kenya, Sweden, and Thailand.
8. Under Visual Settings Marks Area display mode , select Display areas unstacked to display series as overlapping areas on a shared zero baseline. Leave it unselected to stack the areas and show their cumulative total.
9. Click REFRESH VISUAL to display the updated chart.





10. Enter a title for the visual.

    You can also add a subtitle.
11. In the upper-left corner of the Dashboard Designer, click SAVE.

Related Information

Changing the min/max of the aggregate scale

Using the trellis option with areas

In Areas visuals, Cloudera Data Visualization enables you to create Trellis charts that show different partitions (segments) of the dataset alongside one other.

Before you begin

Create the World Population - Area visual as described in *Areas*.

About this task

This example uses the visual World Population - Area from *Areas* to create trellis charts.

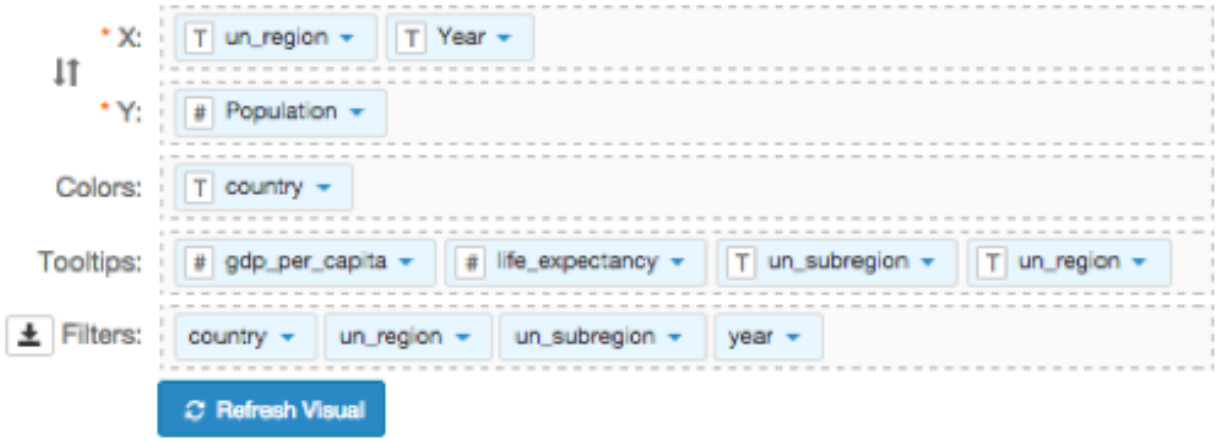
The following steps demonstrate how to create a new trellised Area visual on the World Life Expectancy dataset.

Procedure

1. Open the World Population - Area visual in Edit mode.

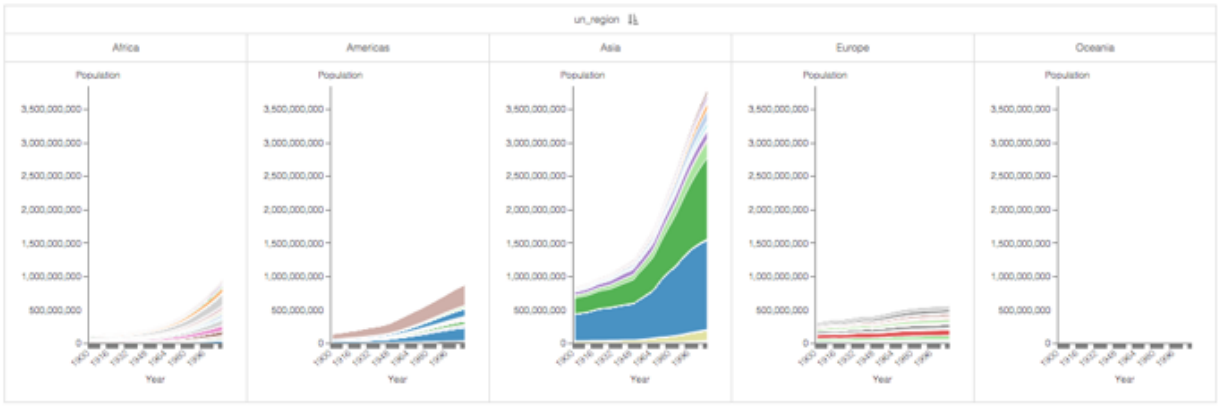
2. Clone the visual as described in *Cloning a visual*.
3. Under Dimensions, select un_region and drag it to the X shelf. If it is not the first (left-most) field on the shelf, drag it to the left of the field Year.

   The first element on the shelf controls the category of the trellis.



4. On the Filters shelf, clear any existing settings.
5. Click REFRESH VISUAL.

   The new trellised area chart appears.



6. Click the (pencil) icon next to the title of the visual to edit it, and change the title to World Population - Area Trellis.
7. At the top left corner of the Dashboard Designer, click SAVE.

Using the trellis option with independent scale

Cloudera Data Visualization enables you to create a new trellised Areas visual with independent aggregate axes.

Before you begin

- Create the World Population - Area visual as described in *Areas*.
- Create the World Population - Area Trellis as described in *Enabling the trellis option*.

About this task

Trellis charts are very powerful visualization tools. However, they do not work well when one chart contains values that are several orders of magnitude larger than in the other charts. The charts that represent segments with smaller values are very difficult to interpret. You can mitigate this problem by enabling independent axis.

This example uses the visual World Population - Area Trellis from *Enabling the trellis option*, to create trellis charts with independent scale.

The following steps show you how to create a new trellised area chart with independent aggregate axes on the World Life Expectancy dataset.

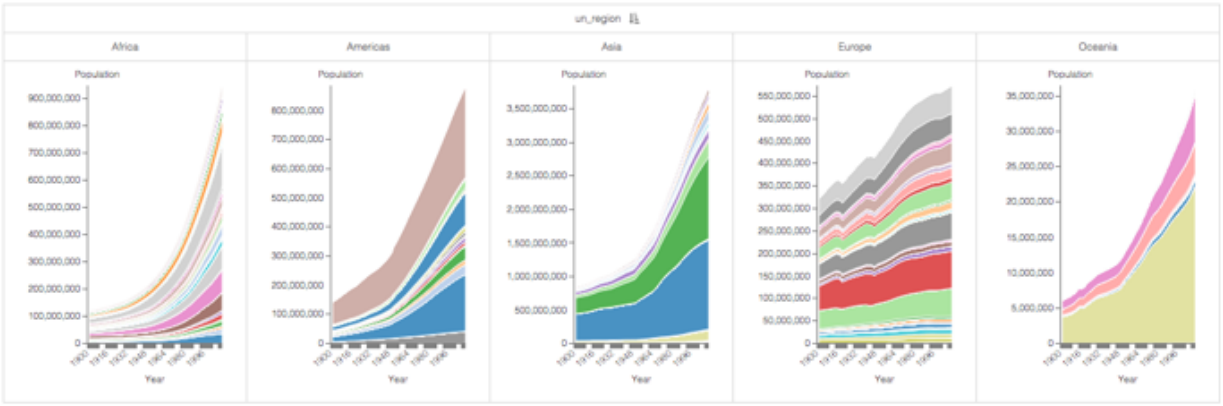
Procedure

1. Open the World Population - Area Trellis visual in Edit mode.
2. Clone the visual as described in *Cloning a visual*.

   While the chart for Asia shows clear growth, the one of Oceania does not show any color. This is because by default, trellis charts are generated to the same scale. You should set Oceania to its own scale to be visible.

   Cloudera Data Visualization provides an override to this default setting. You have an option to set the trellis charts to their own axes by customizing the axes, and selecting Independent Aggregate Scale option, as described in *Customizing axes*.
3. Click REFRESH VISUAL.

   The updated trellised area chart appears.



4. Click the (pencil) icon next to the title of the visualization to edit it, and change the title to World Population - Area Trellis, Independent Axes.
5. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for areas

Overview of shelves for Cloudera Data Visualization Areas visuals.

An Areas visual has the following shelves:

X Axis

This is conceptually the X axis. Use a Dimension for a horizontaly-oriented area graph or a Measure for a vertically-oriented area graph.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in columns.

Y Axis

This is conceptually the Y axis. Use a Measure for a horizontally-oriented area graphor a Dimension for a vertically-oriented area graph.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in rows.

Colors

To stack areas by category, add a Dimension field to this shelf.

This shelf displays only one field on the graph. Additional fields appear only in the Tooltip.

Tooltips

Add Measure fields from the DATA menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

Drill

Add Dimension fields from the DATA menu to this shelf.

When Drill Into context menu is enabled but no fields are specified, you can drill into any dimension. Specifying the fields on this shelf limits drilling to these dimensions.

This shelf accepts multiple fields.

Labels

Additional data labeling.

This shelf accepts at most one field.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Grouped bars

Cloudera Data Visualization enables you to create Grouped Bar visuals, which show clustered groups of related measurements as bars, enhancing comparisons among categories. While bar charts show comparisons among categories, grouped bar charts show clustered groups of related measurements as bars.

About this task

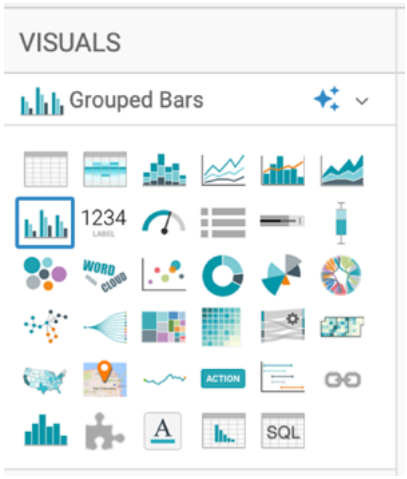
The following steps demonstrate how to create a Grouped Bars visual representation of ethnic distribution of population in some US counties. The example is based on the US Counties dataset. For more information, see *Adding data* and *Creating a dataset*.

For an overview of shelves that specify this visual, see *Shelves for grouped bars*.

Procedure

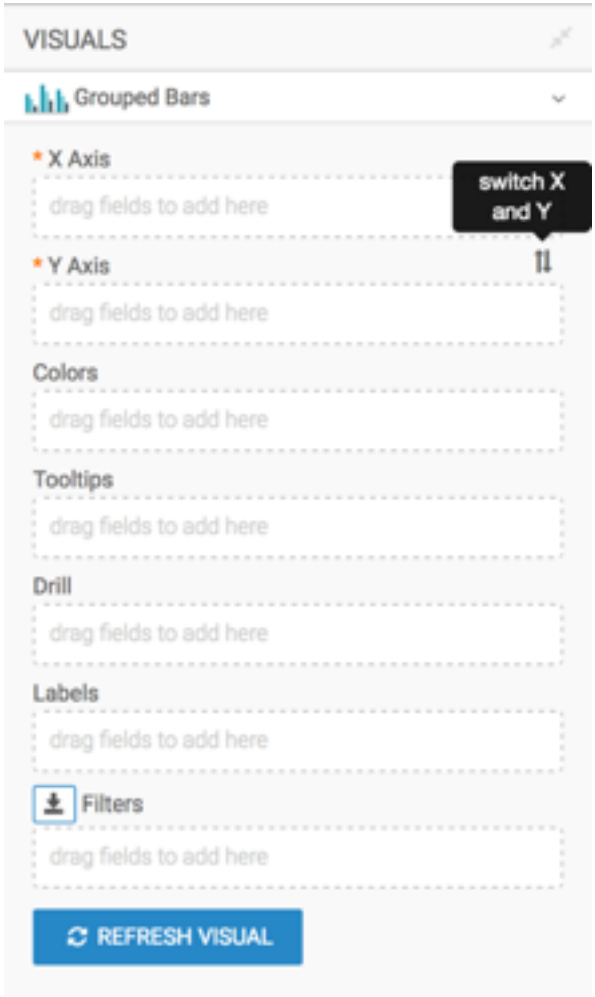
1. Create a new visual based on the US Counties dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Grouped Bars.



You can see that the shelves of the visual changed.

The mandatory shelves are X Axis and Y Axis. Fields placed on these shelves can be swapped to produce a horizontal chart.



3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select the Ctyname field and drag it to the X Axis shelf.
      At this point, it is recommended that you alias the field as County. For more information, see *Alias*.
   b) Select the Wa male field and drag it to the Y Axis shelf.

This first measurement provides the axis definition.

c) Add the other measurements to the Colors shelf.

Together with the measurement on the Y Axis, this creates the 'group' Cloudera Data Visualization uses to draw grouped bars.

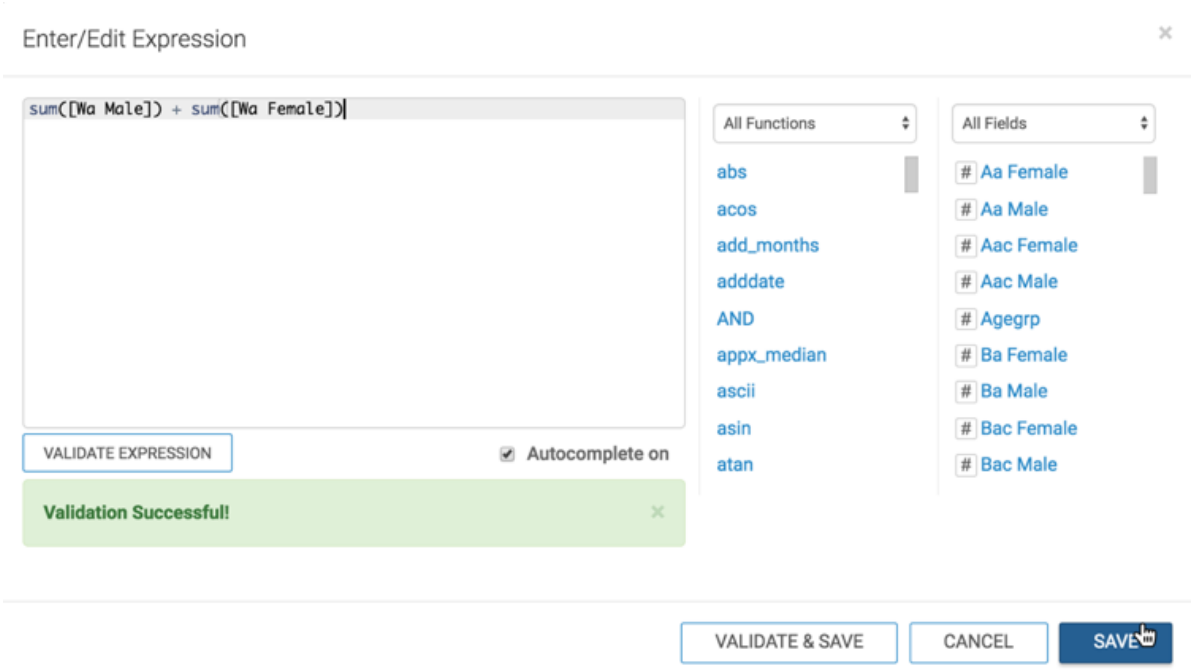
For this example visual, add Ba Male, Ia Male, Aa Male, Na Male, and Tom Male.

d) On the Y Axis shelf, click the field, and from the FIELD PROPERTIES menu, select [ ] Enter/Edit Expression.

e) In the Enter/Edit Expression modal window, edit the expression to the following:

```
sum([Wa Male]) + sum([Wa Female])
```

Click VALIDATE EXPRESSION, and then click SAVE.

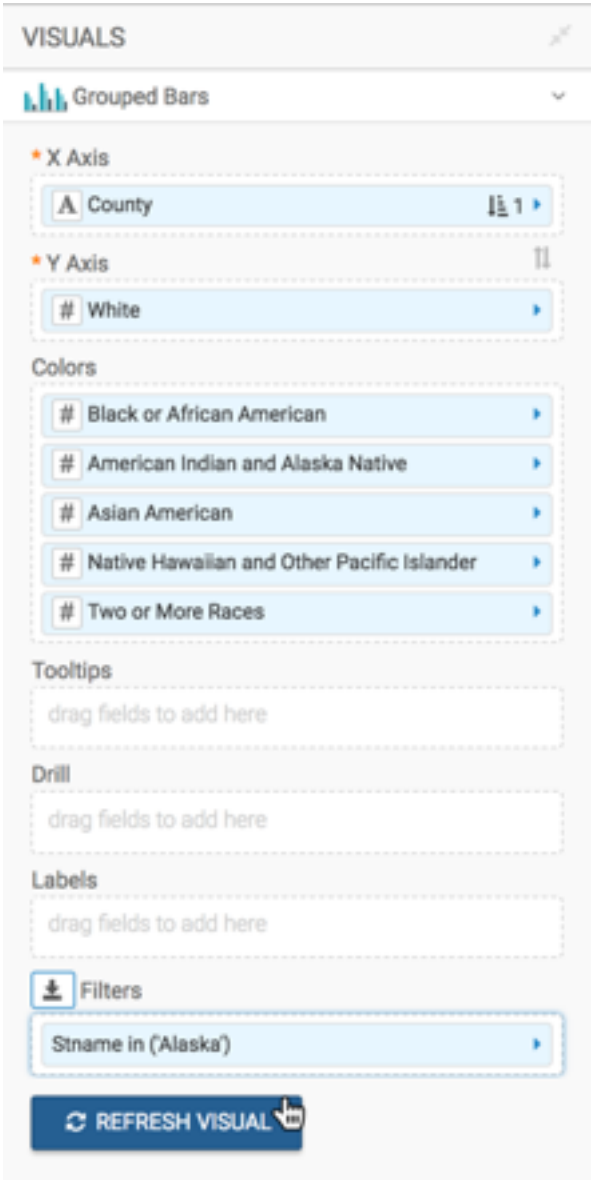


f) Repeat the change in the expression on all the fields placed on the color shelves.

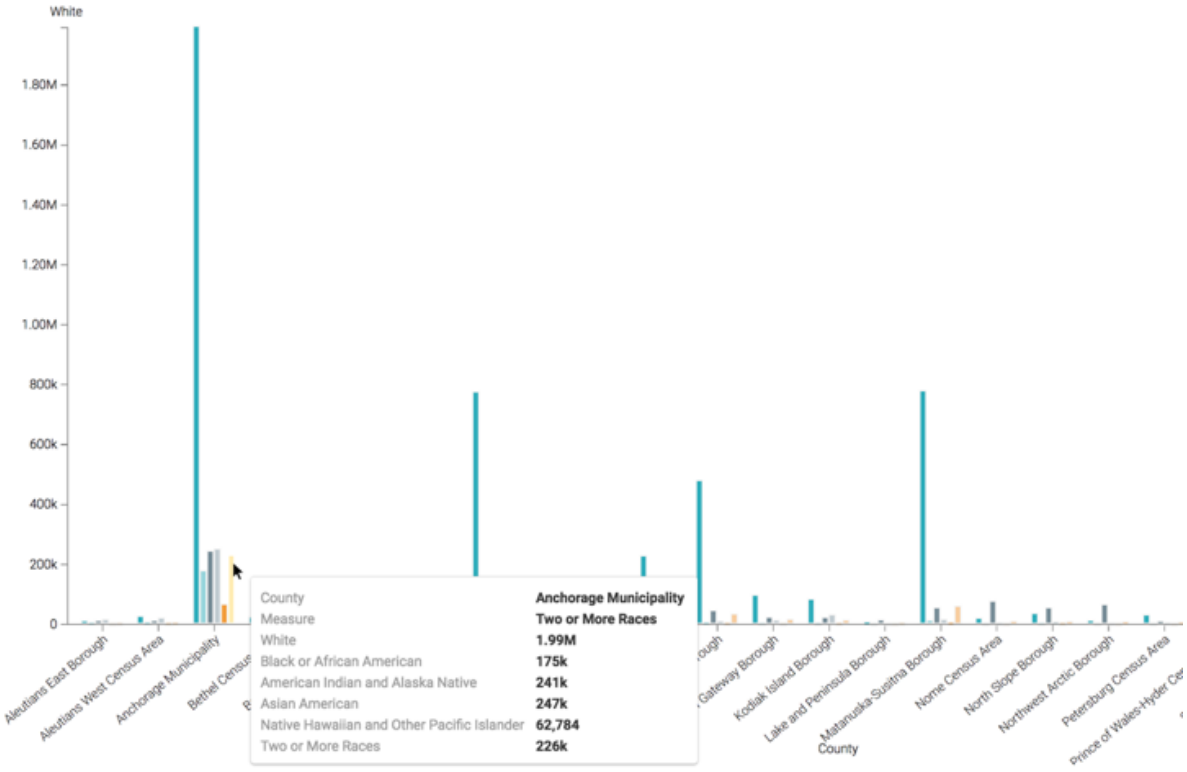
```
sum([Ba Male]) + sum([Ba Female])
sum([Ia Male]) + sum([Ia Female])
sum([Aa Male]) + sum([Aa Female])
sum([Na Male]) + sum([Na Female])
sum([Tom Male]) + sum([Tom Female])
```

g) Alias the field on the Y Axis shelf as 'White', and alias the fields on the Colors shelf as 'Black or African American', 'American Indian and Alaska Native', 'Asian American', 'Native Hawaiian and Other Pacific Islander', and 'Two or More Races'.

h) Add the field Stname to the Filters shelf and select the value 'Alaska'.



4. Click REFRESH VISUAL to generate the Grouped Bars visual.



5. Click the (pencil) icon next to the title of the visualization to edit it, and change the title to World Population - G rouped Bars.
6. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for grouped bars

Overview of shelves for Cloudera Data Visualization Grouped Bars visuals.

A Grouped Bars visual has the following shelves:

X Axis

This is conceptually the X axis. Use a Dimension for horizontal bars or a Measure for vertical bars.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in columns.

Y Axis

This is conceptually the Y axis. Use a Measure for horizontal bars or a Dimension for vertical bars.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in rows.

Colors

To stack bars by category, add a Dimension field to this shelf.

This shelf displays only one field on the graph. Additional fields appear only in the Tooltip.

Tooltips

Add Measure fields from the DATA menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

Drill

Add Dimension fields from the DATA menu to this shelf.

When Drill Into context menu is enabled but no fields are specified, you can drill into any dimension. Specifying the fields on this shelf limits drilling to these dimensions.

This shelf accepts multiple fields.

Labels

Additional data labeling.

This shelf accepts at most one field.

Filters

Add Dimension or Measure fields from the Data menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

KPI

Cloudera Data Visualization enables you to create Key Performance Indicator (KPI) visuals. KPI visuals return values to emphasize particular areas of dashboards or applications. They can be used to provide section titles, totals for segmented data, draw attention to aggregation results, and so on.

About this task

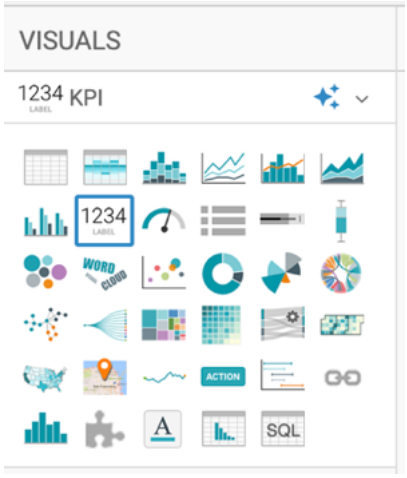
The following steps demonstrate how to create a new KPI visual representation on the Customer Value Analysis dataset. This dataset is based on sample data previously imported into Cloudera Data Visualization.

For information on customization options, see *KPI settings*. For an overview of shelves that specify this visual, see *Shelves for KPI visuals*.

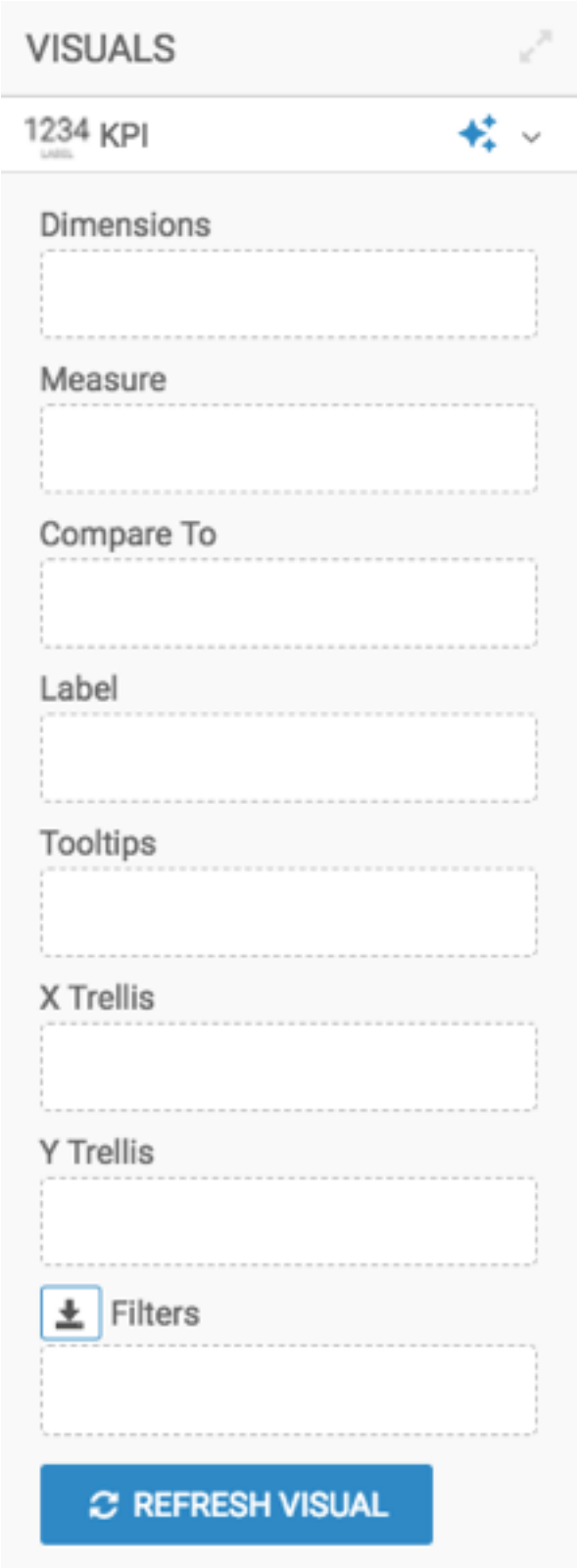
Procedure

1. Start a new visual based on the Customer Value Analysis dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click KPI.



The shelves of the visual changed. There are no mandatory shelves for KPI visuals. Instead, the primary measure appears either on the Dimension or the Measure shelves.



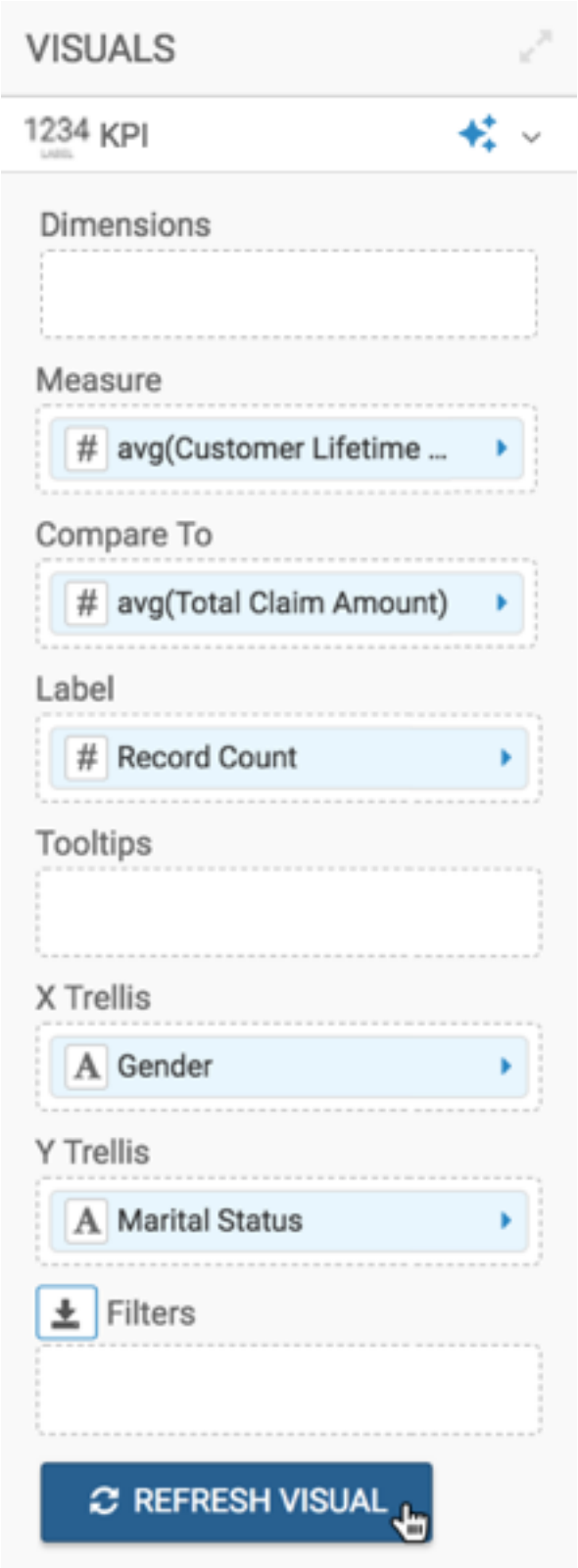
3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) From Measures, select Customer Lifetime Value and place it on the Measure shelf. Ensure that the aggregate is set to Average.

b) From Measures, select Total Claim Amount and place it on the Compare To shelf. Ensure that the aggregate is set to Average.
c) From Measures, select Record Count and place it on the Label shelf.
d) Change the expression on the shelf to the following calculation:

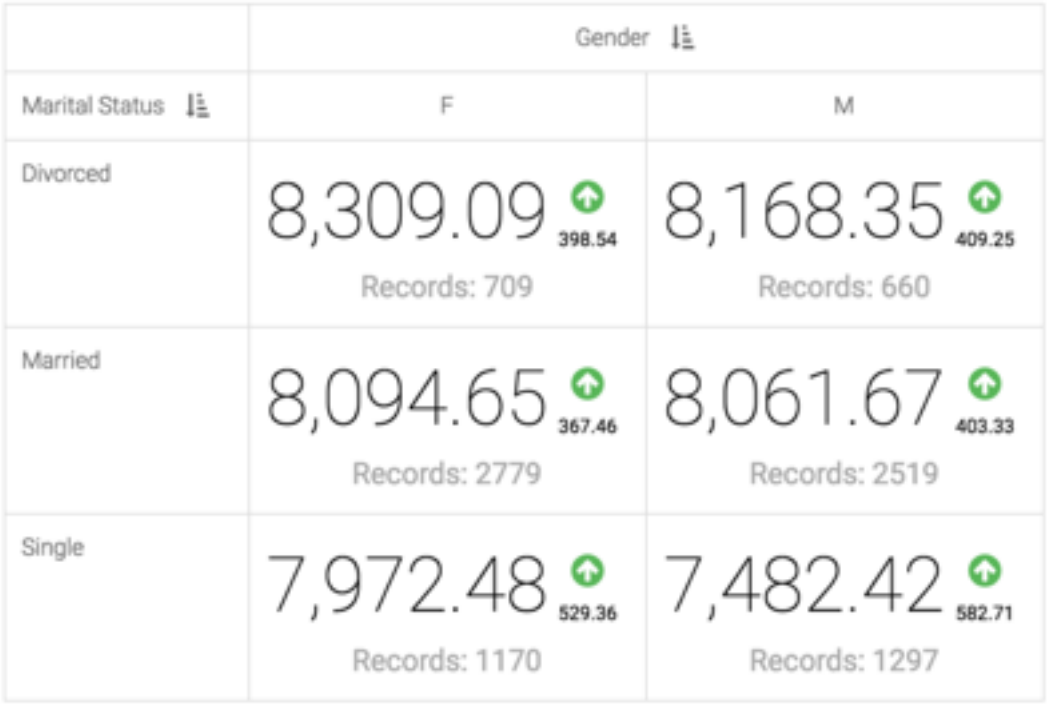
```
concat("Records: ", cast(sum(1) as string)) as 'Record Count'
```

The exact expression may be different, depending on the target data repository.
e) Place the fields Gender and Marital Status (both under Dimensions) on the X Trellis and Y Trellis shelves.

4. Click REFRESH VISUAL.



The KPI visual appears.



5. Click the (pencil) icon next to the title of the visualization to edit it, and change the title to KPI by Gender and Ma rital Status.
6. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for KPIs

Overview of shelves for Cloudera Data Visualization KPI visuals.

A KPI visual has the following shelves:

Dimensions

Add a Dimension field from the DATA menu to display a specific dimension value.

This shelf accepts multiple fields, but the visual displays only the first dimension or measure. Some additional information about the shelf:

- If something is placed on this shelf while the Measure shelf is blank, the visual displays the first value in the dimension as the main indicator.
- If you order this shelf, it displays the first value based on the order parameters.
- If you want to display a dimension sorted by another dimension, place a second dimension on the shelf and sort by it. The visual displays only the first dimension value on the dimension shelf, but the sorted specification on the second dimension determines the order.
- If you want to display a measure value sorted by a dimension (or a dimension sorted by a measure), place the fields on the matching shelves, and sort according to your business case.

For more information, see *Displaying first measure as main indicator*.

Measure

Add a Measure field from the DATA menu to display an aggregate.

This shelf accepts a single field.

The measure shelf either provides the main indicator, or it can be used for sorting. If a measure is used, the display format on this shelf dictates the appearance of the main indicator.

Compare To

Add a Measure field from the DATA menu to show a comparison aggregate as a target value.

This shelf accepts a single field.

If the Dimensions/Measure and Compare To shelves hold numeric values, this shelf specifies both the small value to the right side of the main measure (default placement), and an arrow.

Some additional information about the arrow:

- If the primary measure indicator is higher than the Compare To value, the arrow points up and it has the color specified for the up arrow in the color settings of the visual.
- If the compare value is higher, than the arrow points down, in the color specified for the down arrow in the color settings of the visual.
- If the two values are identical, the arrow is 'flat'. It means that it is horizontal and points to the right.

It is possible to switch the comparison direction by *flipping the comparison* setting.

Label

Add one field to this optional shelf to provide a label under the main indicator (default placement). You can *change label position*.

This shelf accepts 0 or 1 fields.

Tooltips

Add Measure fields from the DATA menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use fields from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use fields from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Gauge

In Cloudera Data Visualization, Gauge visuals are powerful tools for communicating performance against set targets, projected trends, and a qualitative assessment of the measurement.

About this task

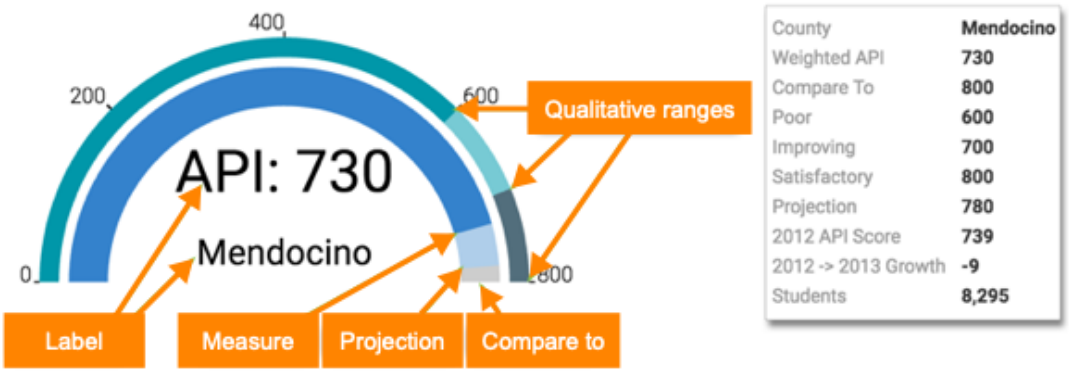
Gauges show minimum and maximum values on an evenly distributed scale, and qualitative ranges for classifying the primary measurement as different colors, which is very useful for visually distinguishing them from one another.

Gauges are semi-circular by default, and include many setting adjustments.

You can review the following three examples of gauge charts to see their structure:

Example 1: Measure < Projection < Compare To

In this scenario, both the measurement and its projected future value are less than the target value.

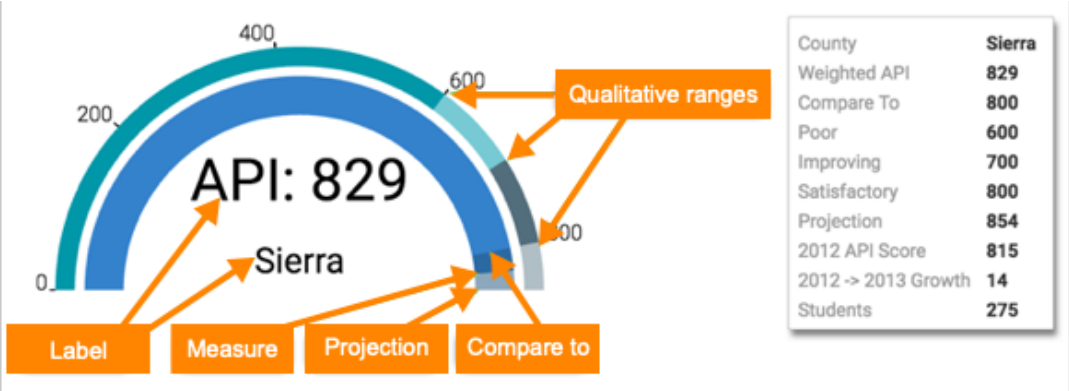


Example 2: Measure < Compare To < Projection

In this scenario, the measurement is less than the target value, but its projection is expected to exceed the target.

Example 3: Compare To < Measure < Projection

In this scenario, the measurement has already exceeded the target value, and the projection is even greater.



- The long dark inner arc of the visual represents the value on the Measure shelf. It is the feature measurement of the visual. In Example 3, the last section appears in dark blue because the value of Measure is greater than the Compare To value.
- The extension of the same arc is the expected value of the measure at a future date, supplied by the Projection shelf of the visual. The overlay of colors and their opacity leads to different effects depending on the relative values of Measure, Projection, and Compare To shelves:
  - In Example 1, this appears in pale blue and does not terminate the arc. This is because the value on the Projection shelf is less than the Compare To value.
  - In Example 2, Projection appears in darker grey as extension beyond the lower Compare To value (in pale blue).
  - In Example 3, Projection again appears in darker grey, but this time as an extension beyond the darker blue arc that shows by how much the Measure exceeds the lower Compare To value.
- The Compare To shelf shows the target value for the measurement. In these examples, it has a value of 800 for all gauges.
- The secondary outer arc shows the Qualitative Ranges that form the upper limits of the 'goodness' for the current measurement. From the origin to the upper value, they specify relative levels of performance success. You may define anywhere from 0 to 4 qualitative ranges.
- The Label shelf overrides default labeling taken from the Measure shelf. This is an opportunity to clarify what is being measured, as well as the units of measurement. This shelf accepts a maximum of two values.

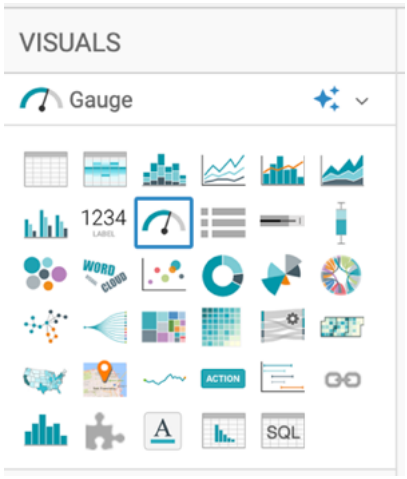
Like other Cloudera Data Visualization visuals, the Gauge visual may be trellised using X and Y shelves, have additional fields on the Tooltips shelf, and use Filters. For more information, see *Shelves for gauge visuals*.

We are working with the District Performance dataset, built on a datafile that contains the California School District APIs from 2012 and 2013. The range of possible scores varies from a low of 200 to a high of 1000. However, the statewide API performance target is 800.

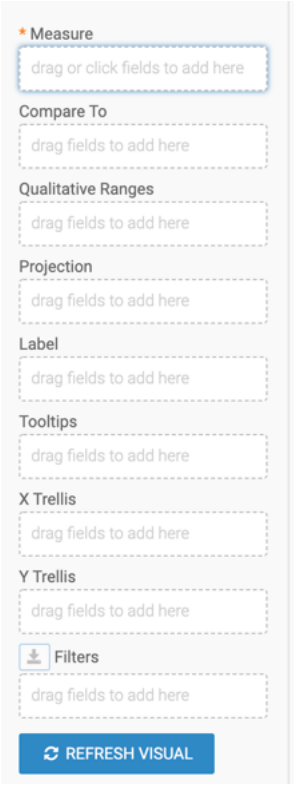
Procedure

1. Start a new visual based on dataset District Performance.

   For more information, see *Adding data*, *Creating datasets*, and *Creating a visual*.
2. In the visuals menu, find and click Gauge.



The shelves of the visual changed. The Measures shelf is mandatory.



3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) For the Measure shelf, add the field Api13.
   b) Customize the expression to calculated the waited average of the field, based on the number of students tested in each district (field Valid).
      1. Click the (right-arrow) icon to the left of the field, and select the [ ] Enter/Edit Expression option.
      2. In the Enter/Edit Expression modal window, change the expression to the following formula:

         ```
         floor(sum([Api13]*[Valid])/sum([Valid]))
         ```

      3. Click VALIDATE & SAVE to ensure that everything works.
   c) For the Compare To shelf, add the field Api13.
   d) Customize the expression to change it to the state-wide target of API=800.

1. Click the (right-arrow) icon to the left of the field, and select the [ ] Enter/Edit Expression option.
2. In the Enter/Edit Expression modal window, change the expression to the scalar value of 800.
3. Click VALIDATE & SAVE to ensure that everything works.

e) For the Qualitative Ranges shelf, add the field Api13.

f) Specify the equation that returns the upper limit of the first qualitative range. For this visual, the example uses simple scalar values.

1. Click the (right-arrow) icon to the left of the field, and select the [ ] Enter/Edit Expression option.
2. In the Enter/Edit Expression modal window, change the expression to the scalar value of 600.
3. Click VALIDATE & SAVE to ensure that everything works.

g) To create other ranges:

1. Click the (right-arrow) icon to the left of the field, and select the Duplicate option.
2. Open the Enter/Edit Expression modal window for the duplicate and change the value.
3. Click VALIDATE & SAVE.

In this example, we defined three ranges: at 600, 700. and 800, and we aliased the qualitative ranges as Poor, Improving, and Satisfactory, in that order.

An alternative to specifying qualitative ranges through the shelf is to calculate them as percentages of the Compare To field. To do that, follow the instructions in.*Specifying qualitative range as percentages*.

h) For the Projection shelf, add the field Api13.

i) Specify the equation that projects the current measurement to some future date by clicking the (right-arrow) icon to the left of the field, selecting the [ ] Enter/Edit Expression option, and changing the expression to the following SQL clause:

```
if(  sum([Api13]*[Valid])/sum([Valid])>800,
     sum([Api13]*[Valid])/sum([Valid])+25,
     sum([Api13]*[Valid])/sum([Valid]) + 50)
```

j) For the Label shelf, add the field Api13, and convert the expression to the following SQL clause:

```
concat("API: ", cast(floor(sum([Api13]*[Valid])/sum([Valid])) as str
ing))
```

You can add the field Cname under the first field and change its aggregation to max([Cname]).

k) For the Tooltips shelf, add more fields here to provide additional information.

In this example, the following tooltips are used:

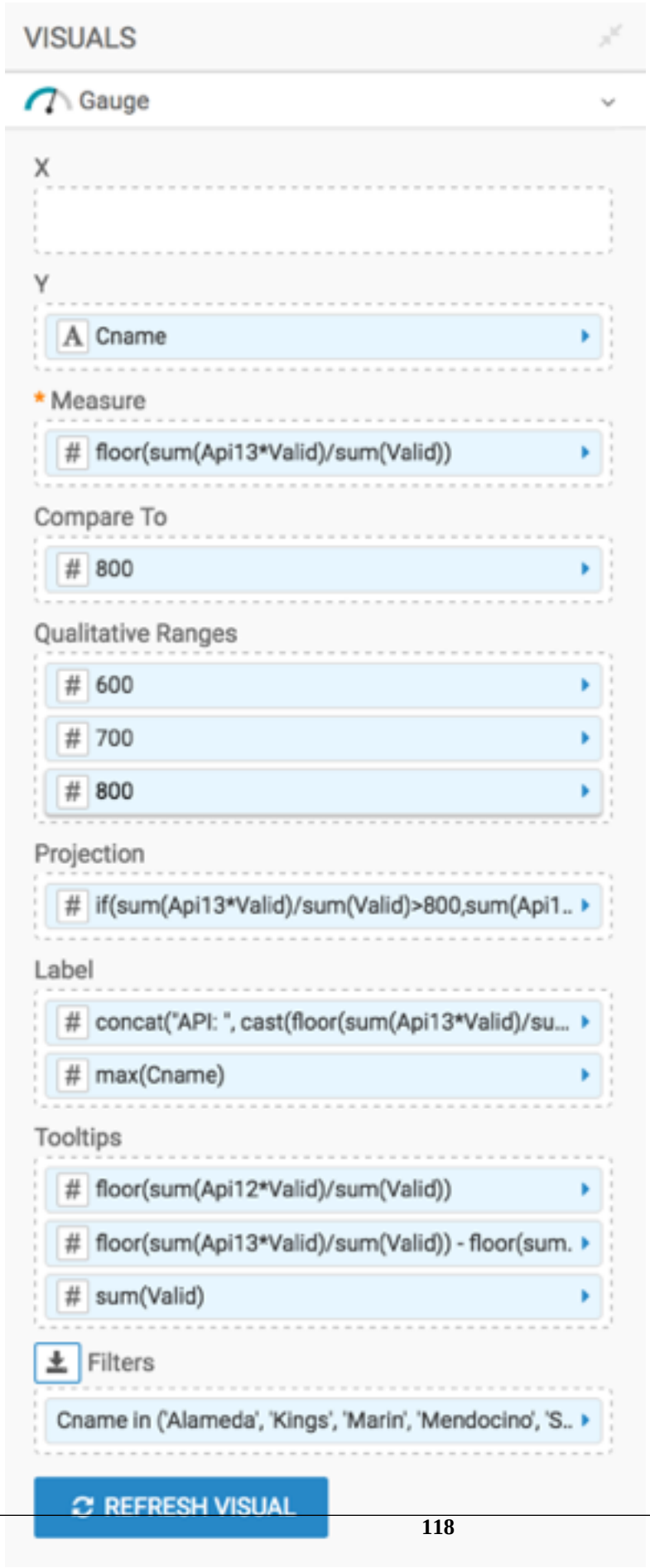
- floor(sum([Api12]*[Valid])/sum([Valid]))
- floor(sum([Api13]*[Valid])/sum([Valid])) - floor(sum([Api12]*[Valid])/sum([Valid]))
- sum([Valid])

It is good practice to apply an alias to these fields. In this example, the following aliases are used: 2012 API Score, 2012 -> 2013 Growth, and Students.

l) For the Filters shelf, add the filters youo need.

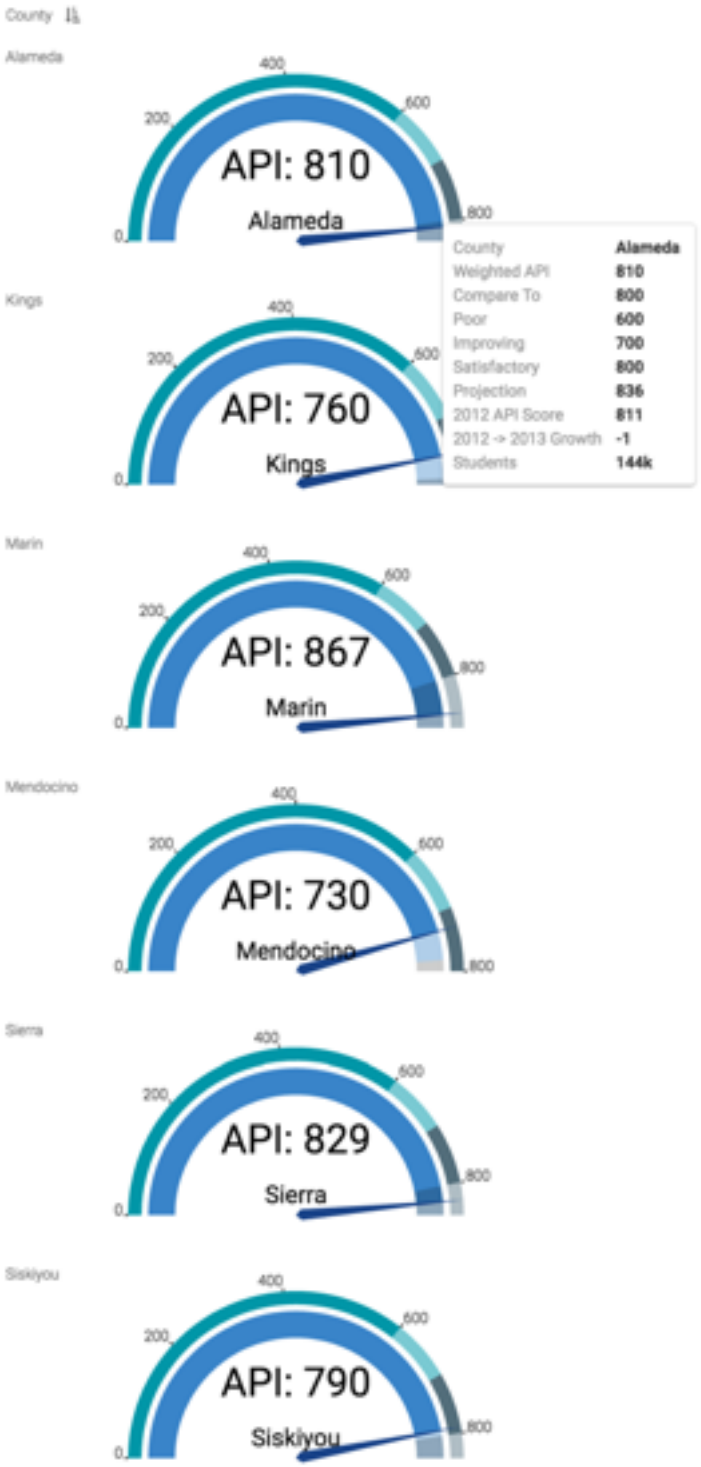
In this example the following filters are used: Cname for name of the County, selecting the values Alameda, Kings, Marin, Mendocino, Sierra, and Siskiyou.

m) For the Y shelf, add the field Cname (county name) to trellis the visual, and lias it as County.


VISUALS
Gauge
X
Y
Cname
* Measure
floor(sum(Api13*Valid)/sum(Valid))
Compare To
800
Qualitative Ranges
600
700
800
Projection
if(sum(Api13*Valid)/sum(Valid)>800,sum(Api1..
Label
concat("API: ", cast(floor(sum(Api13*Valid)/su...
max(Cname)
Tooltips
floor(sum(Api12*Valid)/sum(Valid))
floor(sum(Api13*Valid)/sum(Valid)) - floor(sum.
sum(Valid)
Filters
Cname in ('Alameda', 'Kings', 'Marin', 'Mendocino', 'S..
REFRESH VISUAL

4. Click the Settings menu, select Trellis, and de-select the Show trellis borders option.
5. Click the Settings menu, select Marks, and select the Include a gauge line option. For more information, see *Displaying the gauge needle*.
6. Click the Colors menu, and select different options for Foreground Color, Pointer Color, and Background Color from the color palette menu. You may change the Projection Opacity and Background Opacity options too. In this example, they haven't been changed.
7. Alias the fields. In this example, the names of the shelves are used for each alias.
8. Click REFRESH VISUAL.

   The new Gauge visual appears.



9. Change the title to District Performance, by County - Gauge.

- Click (pencil icon) next to the title of the visualization to edit it, and enter the new name.
- You can click (pencil icon) below the title of the visualization to add a brief description of the visual.

10. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

To adjust the gauge display, check all the available settings for this visual.

Shelves for gauges

Overview of shelves for Cloudera Data Visualization Gauge visuals.

A Gauge visual has the following shelves:

Measure

Add a Measure field from the Data menu to display an aggregate.

This is the feature measurement, such as the current value.

This shelf accepts a single field.

Compare To

Add a Measure field from the Data menu to show a comparison aggregate as a target value.

This shelf accepts a single field.

Qualitative Ranges

Add Measure fields from the Data menu to specify a range of values, such as poor, satisfactory, good, and excellent.

This shelf accepts multiple fields.

Projection

Add a Measure field from the Data menu to show a projected value based on current trends.

This shelf accepts a single field.

Label

Add as many as two Measure or Dimension fields from the Data menu to display label(s) for each graph. This overrides the default value provided by the field on the Measure shelf.

This shelf accepts 0 to 2 fields.

Tooltips

Add Measure fields from the Data menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the Data menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the Data menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the Data menu to trellis the visual into rows.

This shelf accepts multiple fields.

Row listing

Cloudera Data Visualization allows you to create Row Listing visuals. Row Listing is useful when you want to display raw data in a structured format.

About this task

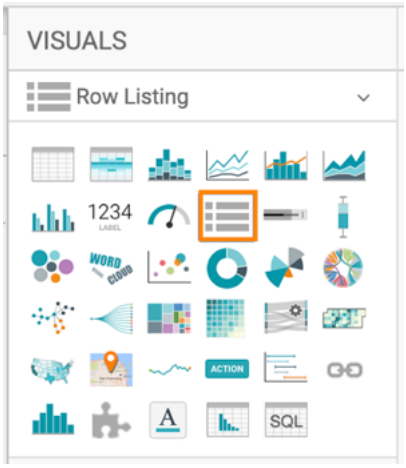
The following steps demonstrate how to create a Row Listing visual based on the Cereals dataset. For an overview of the shelves that specify this visual, see *Shelves for row listing*.

Procedure

1. Start a new visual based on the Cereals dataset.

   By default, a table visual is displayed that shows all fields in the dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click the Row Listing icon.

   The shelves of the visual change.
3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.

   Make the following selections to recreate the example Row Listing visual:

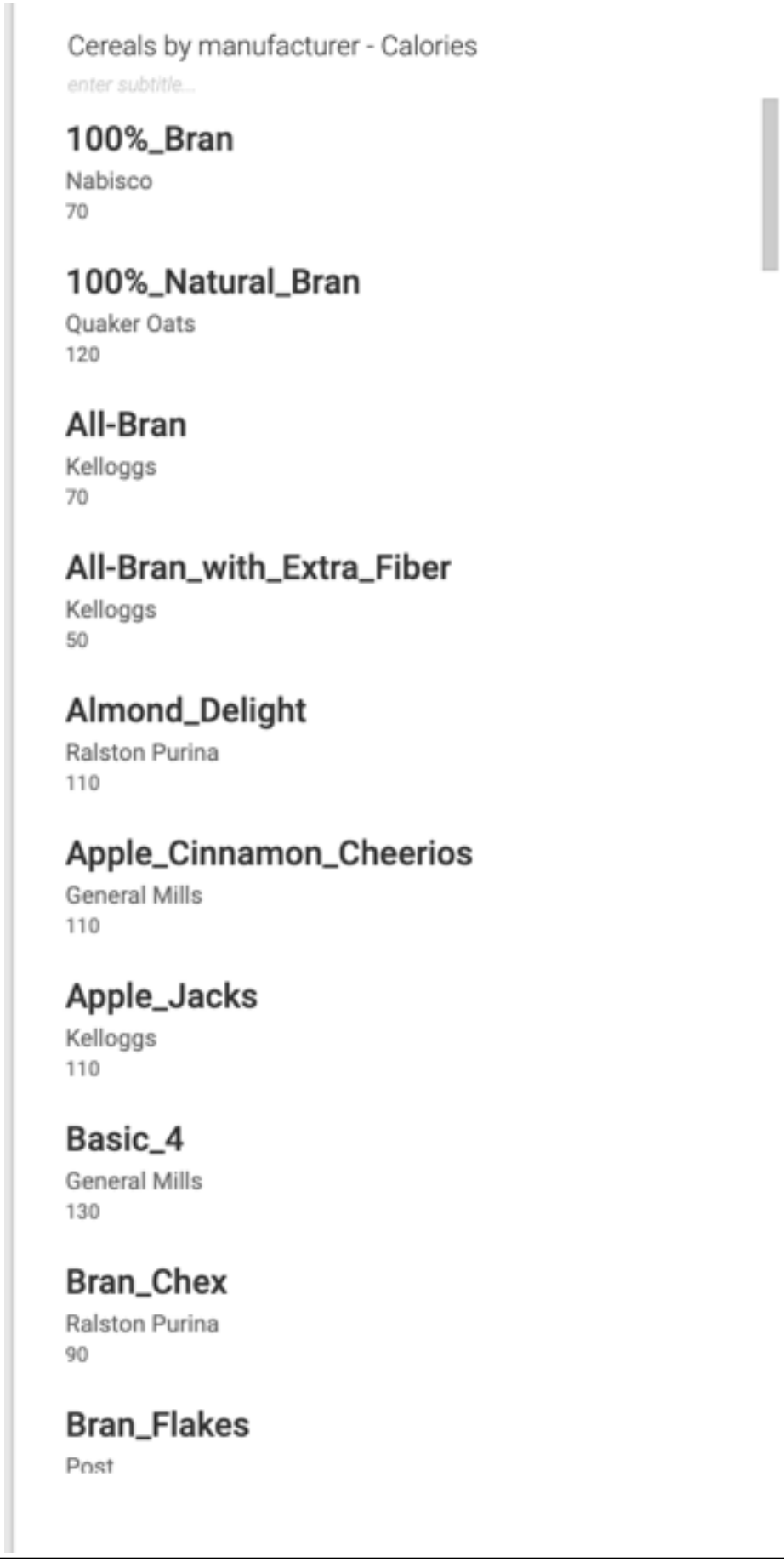
   - Under Dimensions, select ‘cereal_name’ and ‘manufacturer’ and add them to the Dimensions shelf.
   - Under Measures, select ‘calories’ and add it to the Measures shelf.

   For further information about the shelves of the visual, see *Shelves for row listing*.
4. Click REFRESH VISUAL.

   The updated Row Listing visual appears.
5. At the top left corner of the Visual Builder, click SAVE.

Example

In this example, you can see a list of different cereals together with their manufacturer and calorie information.


Cereals by manufacturer - Calories
enter subtitle...
100%_Bran
Nabisco
70
100%_Natural_Bran
Quaker Oats
120
All-Bran
Kelloggs
70
All-Bran_with_Extra_Fiber
Kelloggs
50
Almond_Delight
Ralston Purina
110
Apple_Cinnamon_Cheerios
General Mills
110
Apple_Jacks
Kelloggs
110
Basic_4
General Mills
130
Bran_Chex
Ralston Purina
90
Bran_Flakes
Post

Tip: If no default display format is specified, comma separation will be applied. See, Customizing display format.

Related Information

Creating a visual

Customizing display format

Shelves for row listing visuals

Cloudera Data Visualization Row Listing visuals have the following shelves.

Dimensions

This shelf is required and it accepts multiple fields.

You can add a dimension field from the DATA menu to display a specific dimension value. The visual displays all the added dimensions. If several fields are added, the visual displays the first value in the dimension as the main indicator. If you want to display a dimension sorted by another dimension, place a second dimension on the shelf and sort by it.

Measures

This shelf is optional and it accepts multiple fields.

You can add a measure field from the DATA menu to display an aggregate.

The measure shelf either provides the main indicator, or it can be used for sorting. To display a measure value sorted by a dimension (or a dimension sorted by a measure), place the fields on the matching shelves, and sort according to your business case.

Tooltips

This shelf is optional and it accepts multiple fields.

You can add measure fields from the DATA menu to provide additional information in the visual's tooltip.

Filters

This shelf is optional and it accepts multiple fields.

You can add dimension or measure fields from the DATA menu to include or exclude categorical or value criteria.

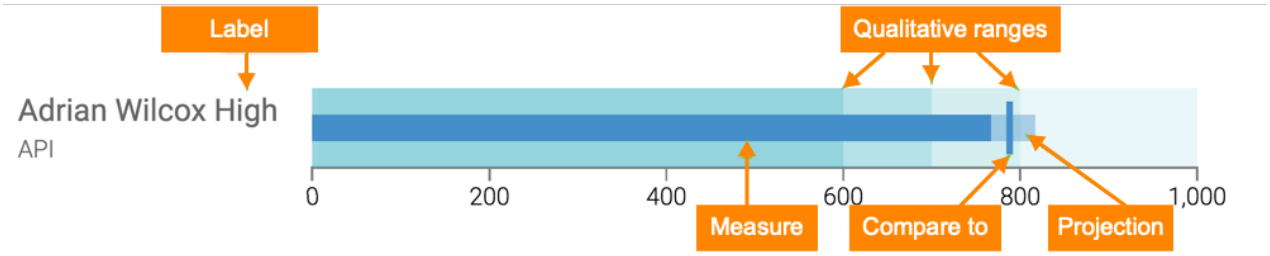
Bullet

Cloudera Data Visualization enables you to create Bullet visuals.

About this task

Bullet visuals are a variation of bar charts, customized for showing progress towards a predefined goal. They are developed specifically for financial data, but may be used in many other situations. They are an alternative to gauges.

This is an example of a Bullet visual for a high schools. It illustrates the various elements of the bullet chart.



- The dark bar in the middle of the visual represents the value on the Measure shelf. It is the feature measurement of the visual.
- The lighter-shaded extension of the same bar is the expected value of the measure at a future date. This value is on the Projection shelf of the visual.
- The vertical bar is the target value of the measurement. Place this calculation on the Compare To shelf.
- Calculations on the Qualitative Ranges shelf determine the upper limits of the 'goodness' for the current measurement. From the origin to the upper value, they specify poor, better, good, and excellent performance. You may define anywhere from 0 to 4 qualitative ranges.
- The Label shelf overrides default labeling taken from the Measure shelf. This is an opportunity to clarify what is being measured, as well as the units of measurement. This shelf accepts a maximum of two values.

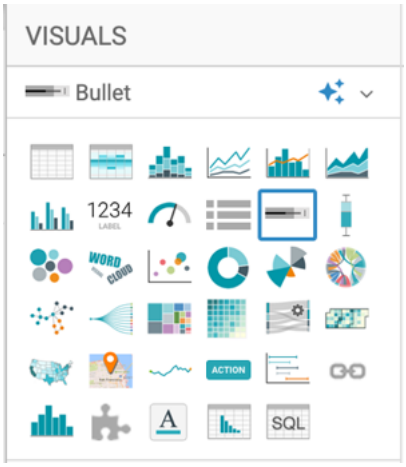
Just as other Cloudera Data Visualization visuals, the Bullet visual may be trellised using X and Y shelves, have additional fields on the Tooltips shelf, and you can use Filters with it. It also includes additional setting adjustments. For more information, see *Shelves for bullet visuals*.

The following example uses the 2012 and 2013 API Scores for California Schools. The School Performance dataset is used, which was built on a datafile that contains the California School APIs from 2012 and 2013. The range of possible scores varies from a low of 200 to a high of 1000. The statewide API performance target for all schools is 800. However, each school that has not achieved that target may have a more realistic score set for it by the Department of Eduction, based on a number of factors we will not discuss here.

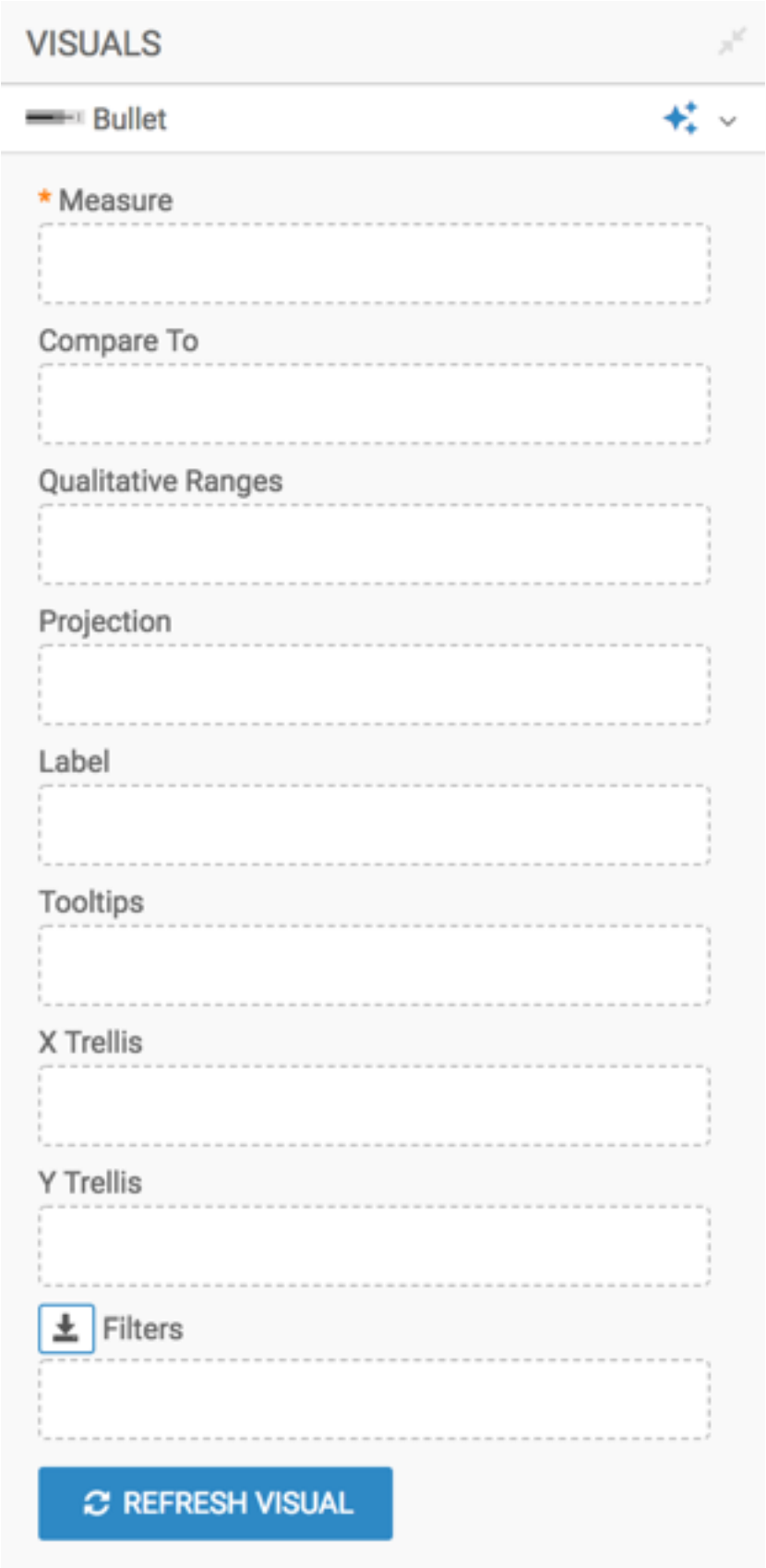
Procedure

1. Start a new visual based on the School Performance dataset.

   For instructions, see *Adding data*, *Creating datasets*, and *Creating a visual*.
2. In the VISUALS menu, find and click Bullet.



   The shelves of the visual changed. They are now Measure (mandatory), Compare To, Qualitative Ranges, Projection, Label, Tooltips, X Trellis, Y Trellis, and Filters.



3. Populate the shelves from the available fields (in Dimensions and Measures) in the DATA menu.
   a) Measure - Add the field Api13, and change the aggregation, so that the shelf contains max([Api13]).

[Optional] Limit the output to the top 10 performers:

1. Click the (right-arrow) icon to the left of the field, and select the Order and Top K option.
2. After the menu expands, enter 10 in the Top K option.

[Optional] Alias the field as the name of the measurement. In this case, use API.

Note: The Measure shelf supports only positive values.

b) Compare To - Add the field Api12. Then customize the expression to reflect whether the school reached the state-wide target of API=800.

1. Click the (right-arrow) icon to the left of the field, and select the [ ] Enter/Edit Expression option.
2. In the Enter/Edit Expression modal window, change the expression to the following formula:

```
Max(if([Api12] < 800, [Api12] + [Target], 800))
```

3. Click VALIDATE EXPRESSION to ensure that everything works.
4. Click SAVE.

[Optional] Alias the field as Compare To.

c) Qualitative Ranges - Add the field Api12. Then, specify the equation that returns the upper limit of the first qualitative range. For this visual, simple scalar values are used.

1. Click the (right-arrow) icon to the left of the field, and select the [ ] Enter/Edit Expression option.
2. In the Enter/Edit Expression modal window, change the expression to the scalar value 600, and click SAVE.
3. To create other ranges, simply click the (right-arrow) icon to the left of this field, and select the Duplicate option. Then open the Enter/Edit Expression modal window for the duplicate, change the value, and click SAVE.

   In this example, four ranges are defined: at 600, 700. 800, and 1000.
4. [Optional] Alias the ranges as Range 1, Range 2, Range 3, and Range 4.

d) Projection - Add the field Api13. Then, specify the equation that projects the current measurement to some future date.

1. Click the (right-arrow) icon to the left of the field, and select the [ ] Enter/Edit Expression option in the FIELD PROPERTIES menu.
2. In the Enter/Edit Expression modal window, change the expression to the following formula and click SAVE:

```
if(max([Api13]) > 800, max([Api13]), max([Api13]) + 2 * max([Target]
))
```

3. [Optional] Alias the field as Projection.

e) Label - Add the field Api13, and convert the aggregation to maximum, for max([Api13]).

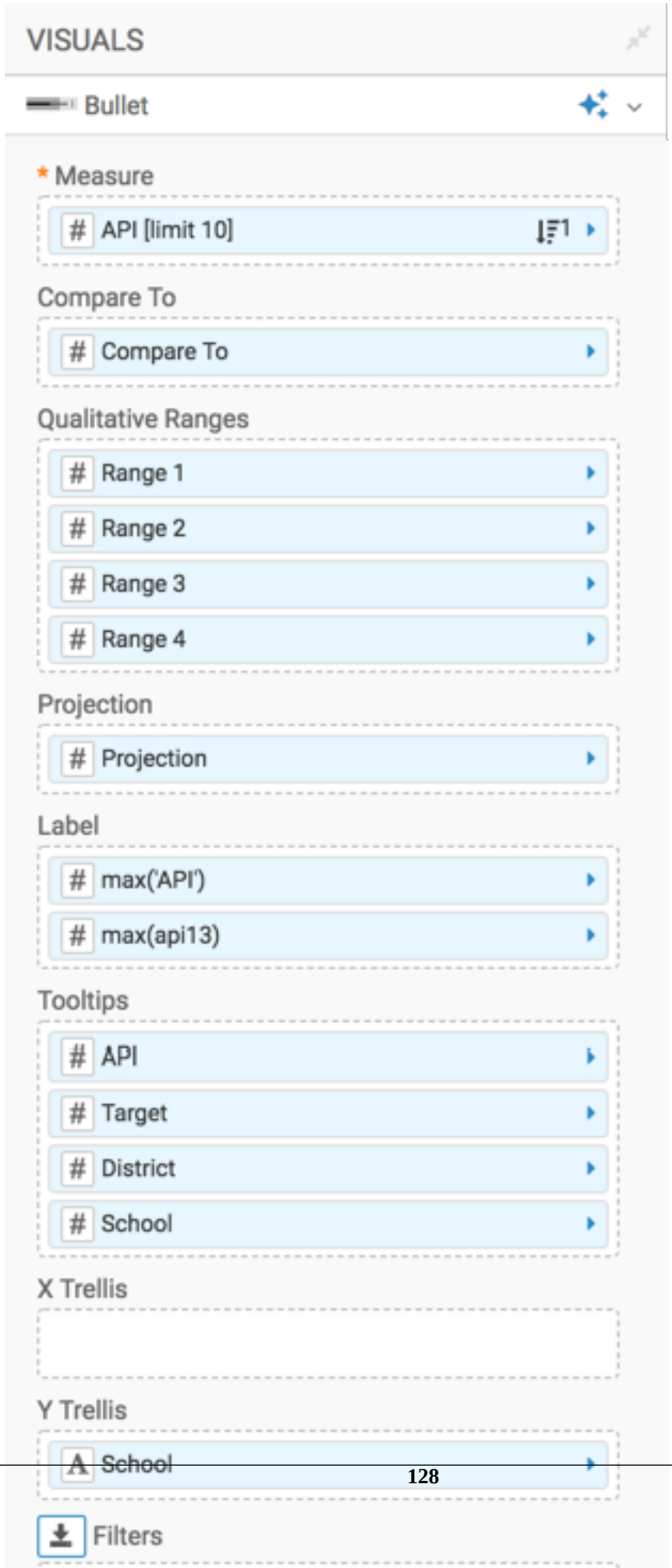
[Optional] Duplicate the field, and change its definition to max('API'), using the Enter/Edit Expression option in the FIELD PROPERTIES menu.. Move this field over the other field.

f) Tooltips - Add more fields here, such as Api12, Target, Dname, Sname, and so on. You can apply an Alias to these fields.

g) Y Trellis - Add the field Sname, and Alias it as School.

h) Filters - Add the filters here. Cname is used in this example for name of the County (selecting Santa Clara), and Stype for school type (choosing H for high school).

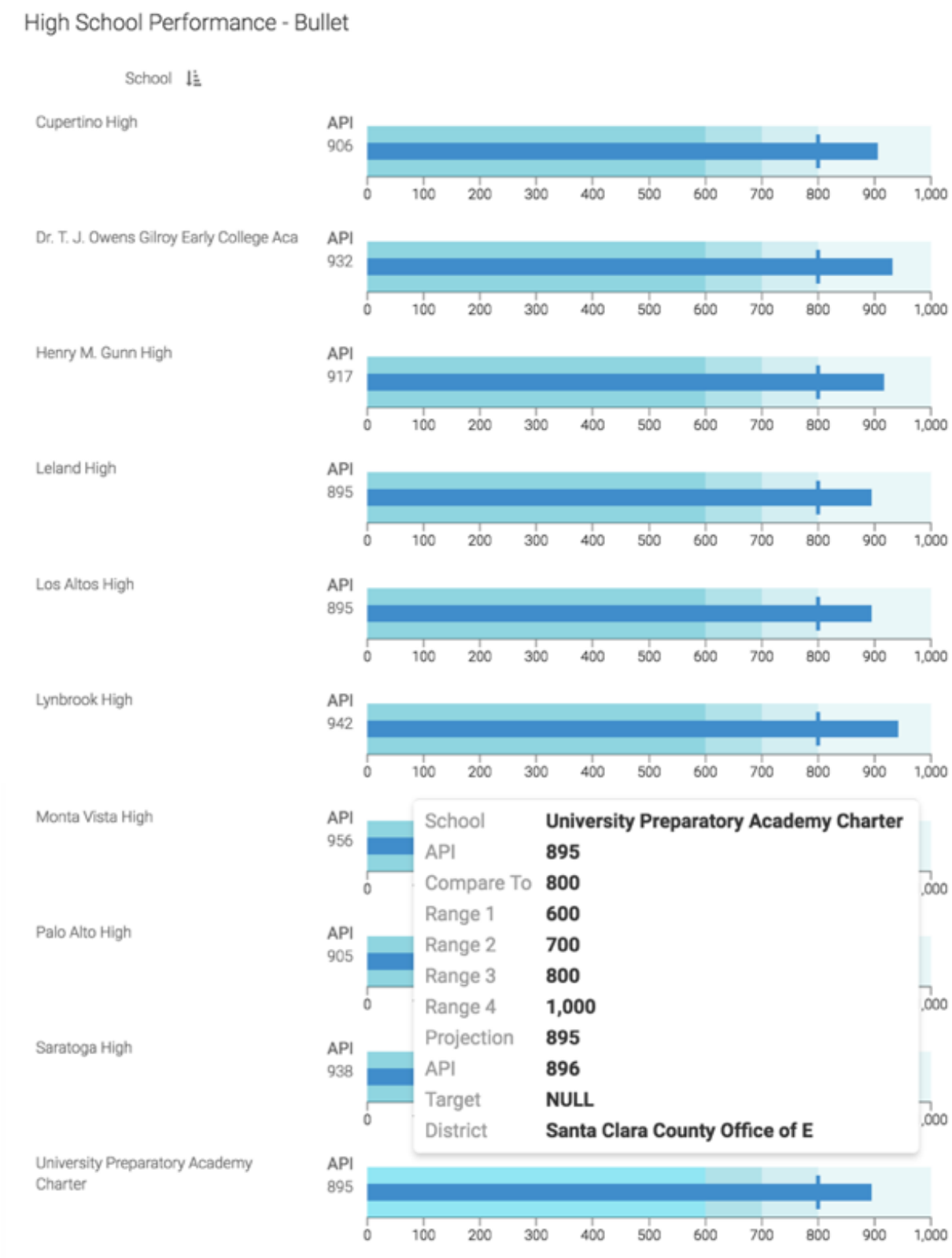
i) Y shelf - Ad the field Sname (school name) to trellis the visual.


VISUALS
Bullet
* Measure
# API [limit 10]
Compare To
Compare To
Qualitative Ranges
Range 1
Range 2
Range 3
Range 4
Projection
Projection
Label
max('API')
max(api13)
Tooltips
API
Target
District
School
X Trellis
Y Trellis
A School
Filters

4. [Optional] Click the Settings menu, select Trellis, and de-select the Show trellis borders option.
5. [Optional] Click the Colors menu, and select different options for Background Color and Foreground Color from the color palette menu. You may change the Projection Opacity too.
6. Click REFRESH VISUAL.

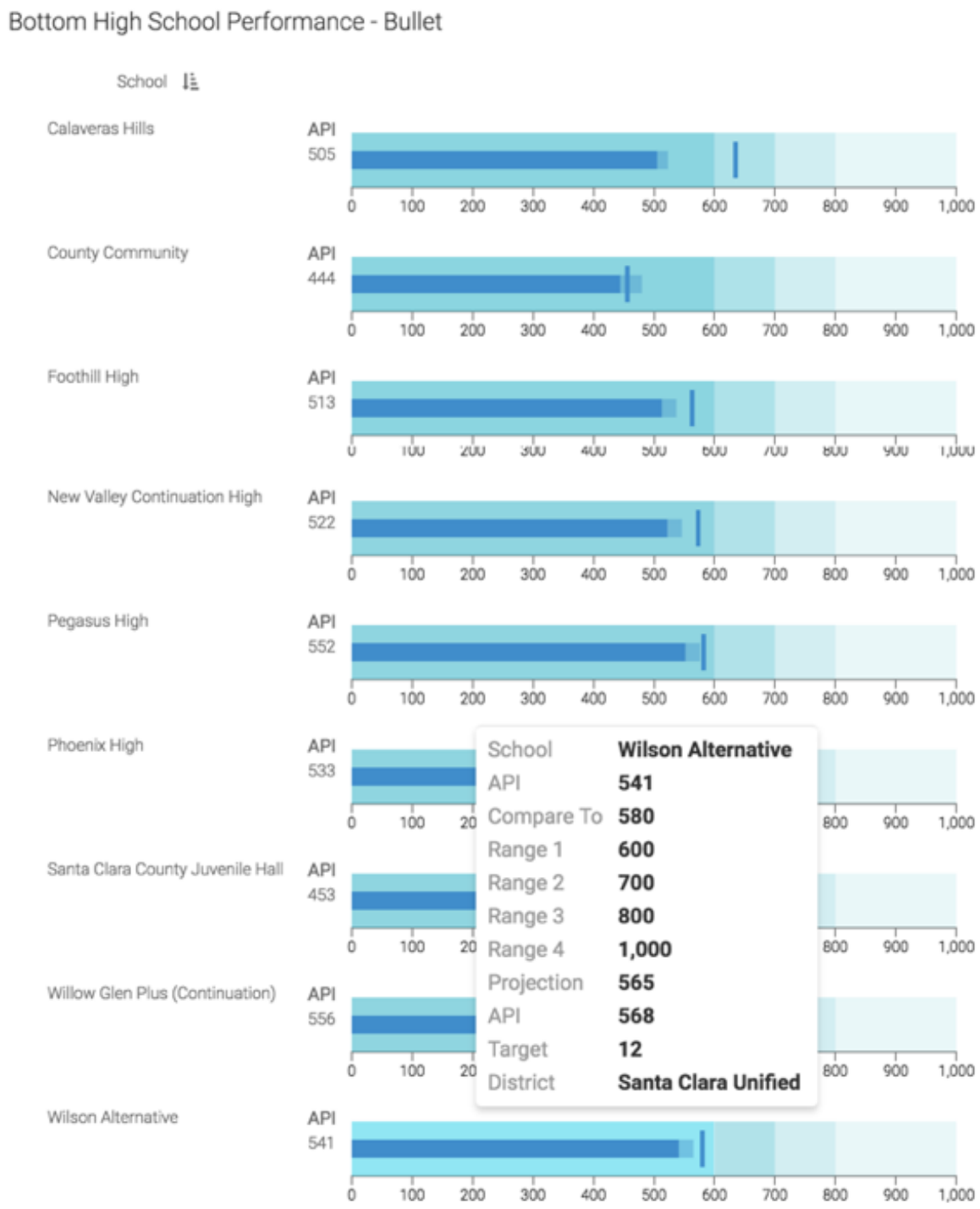
   The new Bullet visual appears. Note that it is trellised, which is why it is sorted on the Sname field (school name), not on the measure Api13.

   Additionally, the Projection value is not visible; this is because it is relatively low for high-performing schools.



7. Change the order property of the Measure shelf by selecting the Bottom K option with a value of 10, and click REFRESH VISUAL.

   Note that in all cases here, the projection value is visible because it is relatively large compared to the projected improvement for schools that already reached the 800 API threshold.



8. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'High School Performance - Bullet. You can also add a brief description of the visual as a subtitle below the title of the visualization.
9. At the top left corner of the Dashboard Designer, click SAVE.
10. At the top left corner of the Dashboard Designer, click CLOSE.

What to do next

To adjust the Bullet visual, check all the available settings for this visual.

Shelves for bullets

Overview of shelves for Cloudera Data Visualization Bullet visuals.

A Bullet visual has the following shelves:

Measure

Add a Measure field from the DATA menu to display an aggregate.

This is the feature measurement, such as the current value.

This shelf accepts a single field.

Compare To

Add a Measure field from the DATA menu to show a comparison aggregate as a target value.

This shelf accepts a single field.

Qualitative Ranges

Add Measure fields from the DATA menu to specify a range of values, such as poor, satisfactory, and good. This determines the background shading intensity.

This shelf accepts multiple fields.

Projection

Add a Measure field from the DATA menu to show a projected value based on current trends.

This shelf accepts a single field.

Label

Add as many as two Measure or Dimension fields from the DATA menu to display label(s) for each graph. This overrides the default value provided by the field on the Measure shelf.

This shelf accepts 0 to 2 fields.

Tooltips

Add Measure fields from the DATA menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Box plot

Cloudera Data Visualization enables you to create Box Plot visuals.

About this task

A Box Plot visual displays the distribution of data values. The box represents the bulk of the data (25th percentile through 75th percentile), and whisker lines represent statistical minimum and maximum values.

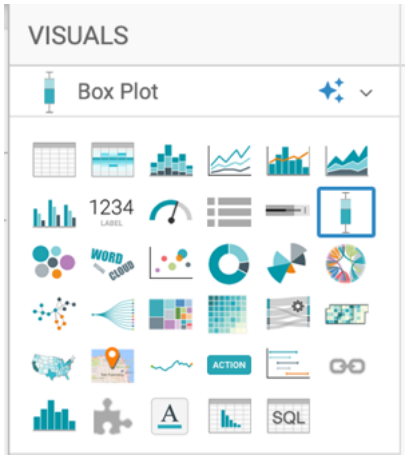
The following example uses the 2012 and 2013 API Scores for California Schools. The School Performance dataset is used, which is built on a datafile that contains the California School APIs from 2012 and 2013. The range of possible scores varies from a low of 200 to a high of 1000. The statewide API performance target for all schools is 800.

For an overview of shelves that specify this visual, see *Shelves for box plots*.

Procedure

1. Start a new visual based on the NYC Taxi Data dataset.

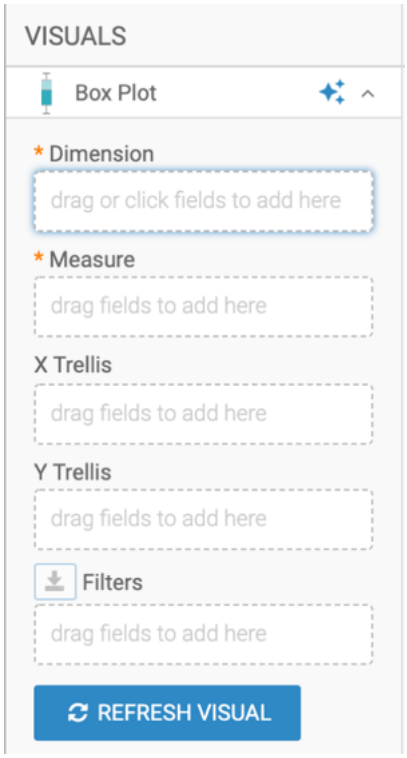
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Box Plot.



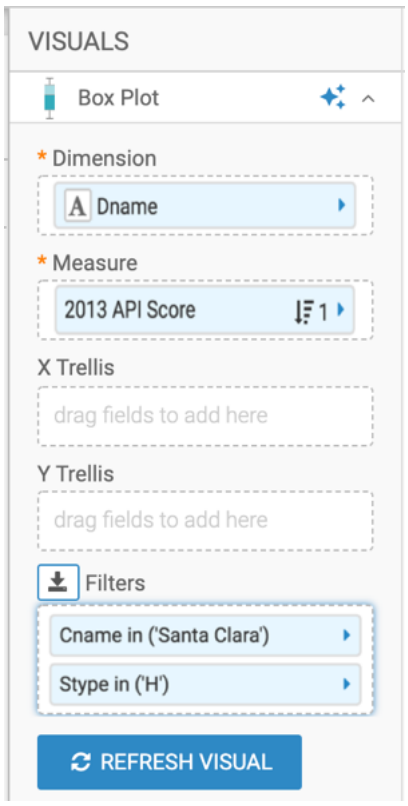
The shelves of the visual changed. They are now Dimension, Measure, X Trellis, Y Trellis, and Filters.

The mandatory shelves are:

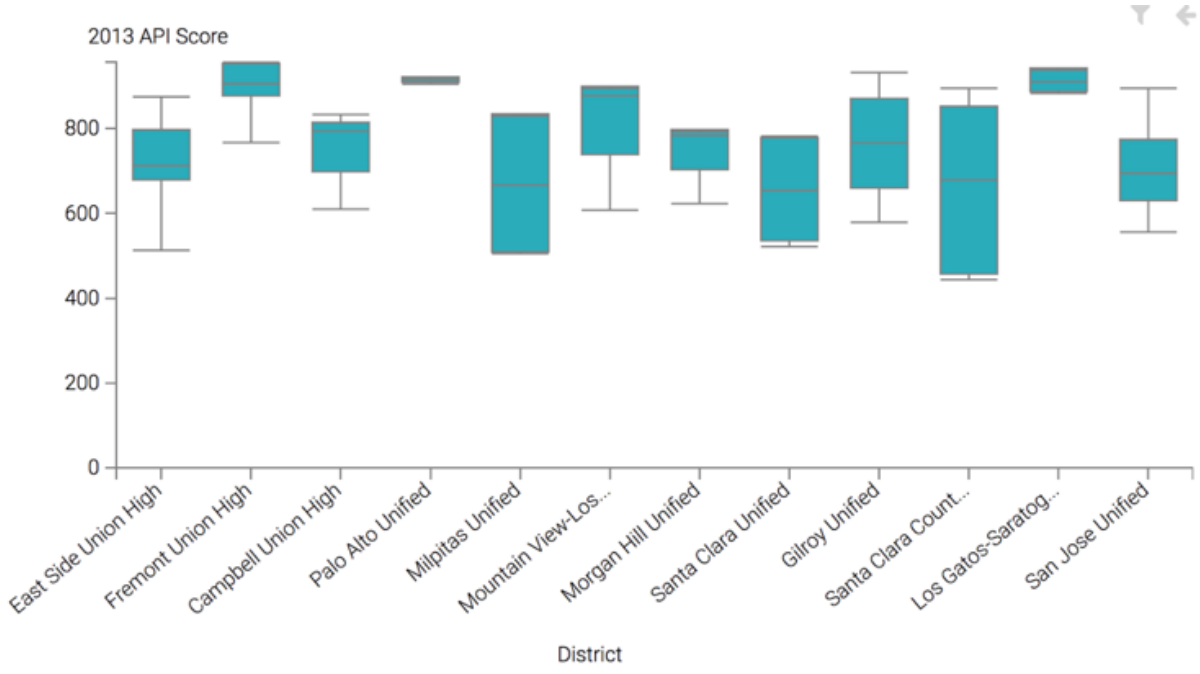
- Dimension: you can specify one dimension for grouping the box plot
- Measure: you can specify the measurement for box plot calculations



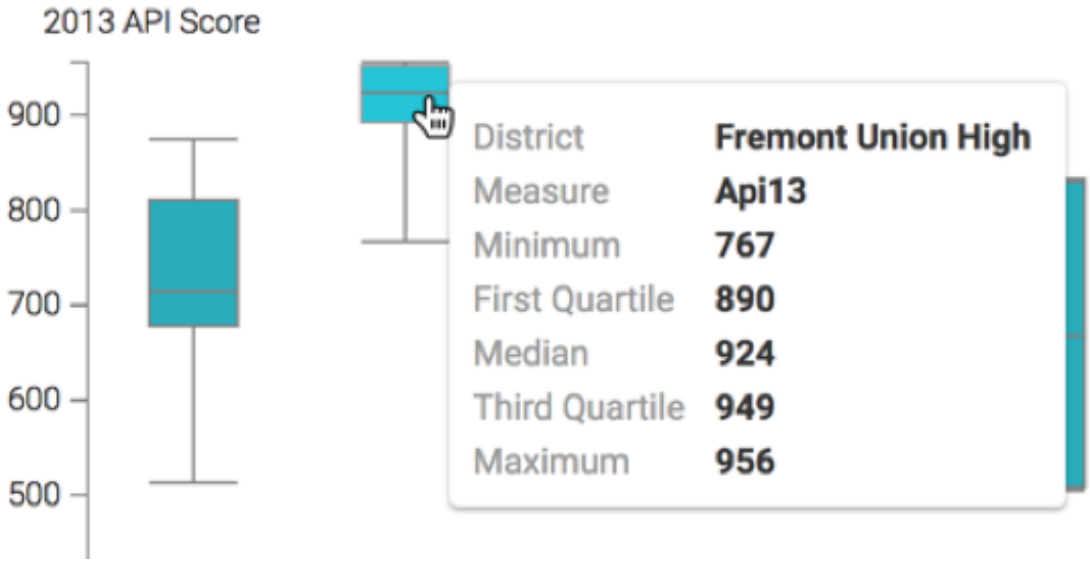
3. Populate the shelves:
   a) Add the Dname (district name) field to the Dimension shelf.
   b) Add the Api13 (API obtained in 2013)field to the Measure shelf.
   c) Add the Cname (county name) field to the Filters shelf, and select a single value. In this example, Santa Clara is used.
   d) Add the Stype (school type) field to the Filters shelf, and select the value H, for Highschool.
4. Alias the shelves.
5. Click REFRESH VISUAL.



The new Box Plot visual appears.



6. Hovering over a particular box plot distribution rectangle opens a Tooltip modal that contains the statistical measures of that grouping, including the minimum, the first quartile, the median/second quartile, the third quartile, and the maximum.



7. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'High School Performance - Box Plot'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
8. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

If you wish to generate a horizontal Box Plot visual, see the instructions in *Making box plots horizontal*.

To adjust the Box Plot display, check all the available settings for this visual.

Related Information

Creating a visual

Making box plots horizontal

Shelves for box plots

Overview of shelves for Cloudera Data Visualization Box Plot visuals.

A Box Plot visual has the following shelves:

Dimension

Add a Dimension field from the DATA menu to specify the grouping of each box plot.

This is the feature measurement, such as the current value.

This shelf accepts a single field.

Measure

Add a Measure field from the DATA menu; this is the measurement that determines the shape of the box plot.

We optimized Box plot statistics for Impala, Hive, and Spark SQL connections.

This shelf accepts a single field.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Packed bubbles

Cloudera Data Visualization enables you to create Packed Bubbles visuals.

About this task

Packed Bubbles visuals compare the relative values in close physical proximity. Two main series of data are used: the field on the Dimensions shelf names the bubble and the field on the Measures shelf determines the relative area of the bubble.

This is different from the traditional approach, where bubbles represent three dimensions of data — one each for horizontal and vertical axes, and one for the area of bubble. This type of visualization is part of flexibility built into Scatter visuals available in Cloudera Data Visualization.

Note:

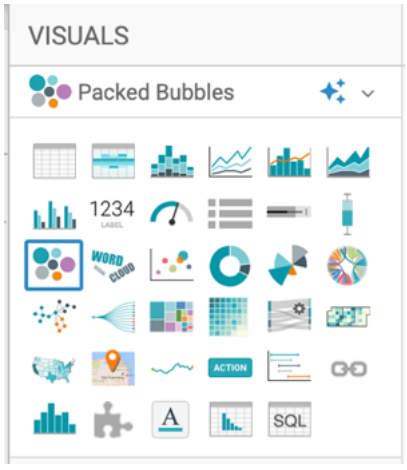
This visual type supports smooth rendering of streaming data.

The following steps demonstrate how to create a new Packed Bubbles visual on the World Life Expectancy dataset [data source samples.world_life_expectancy]. For an overview of shelves that specify this visual, see *Shelves for packed bubbles visuals*.

Procedure

1. Start a new visual based on the World Life Expectancy dataset.

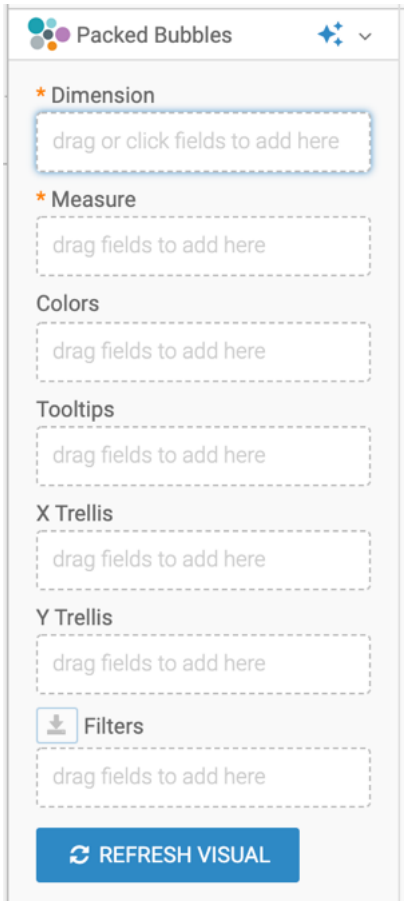
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Packed Bubbles.



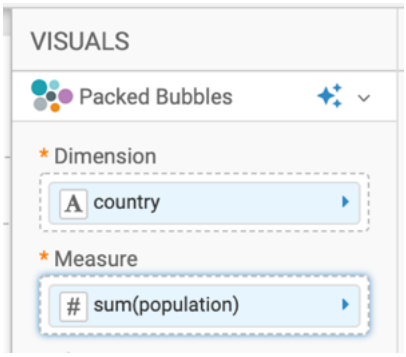
   The shelves of the visual changed.

   They are now Dimension, Measure, Color, Tooltips, X Trellis, Y Trellis, and Filters.

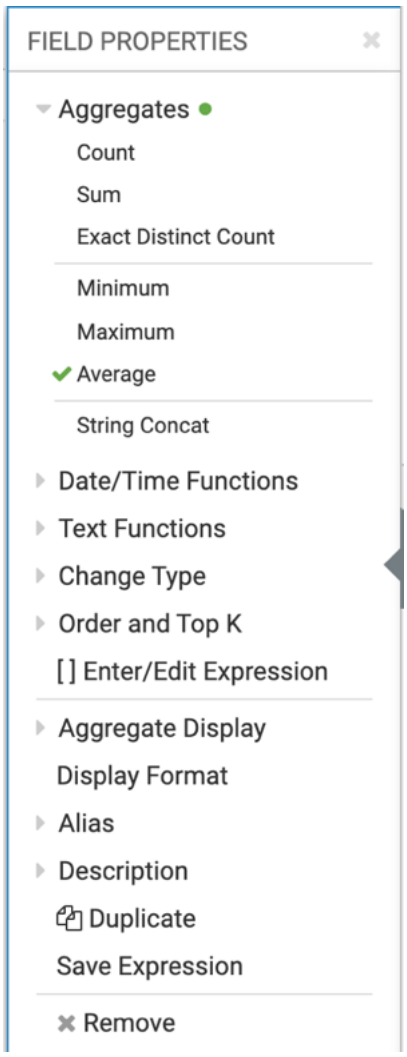
   Both Dimension and Measure are mandatory.



3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select country and drag it to the Dimension shelf.
   b) Under Measures, select population and drag it to the Measure shelf.



4. Change the aggregate of the population field from sum(population) to avg(population):
   a) On the shelf of the visual, click the caret-right icon to the right of the field.
   b) In the FIELD PROPERTIES menu, click Aggregates Average .



c) Click the x icon at the top of the Column Properties menu to close it.

5. Click REFRESH VISUAL.

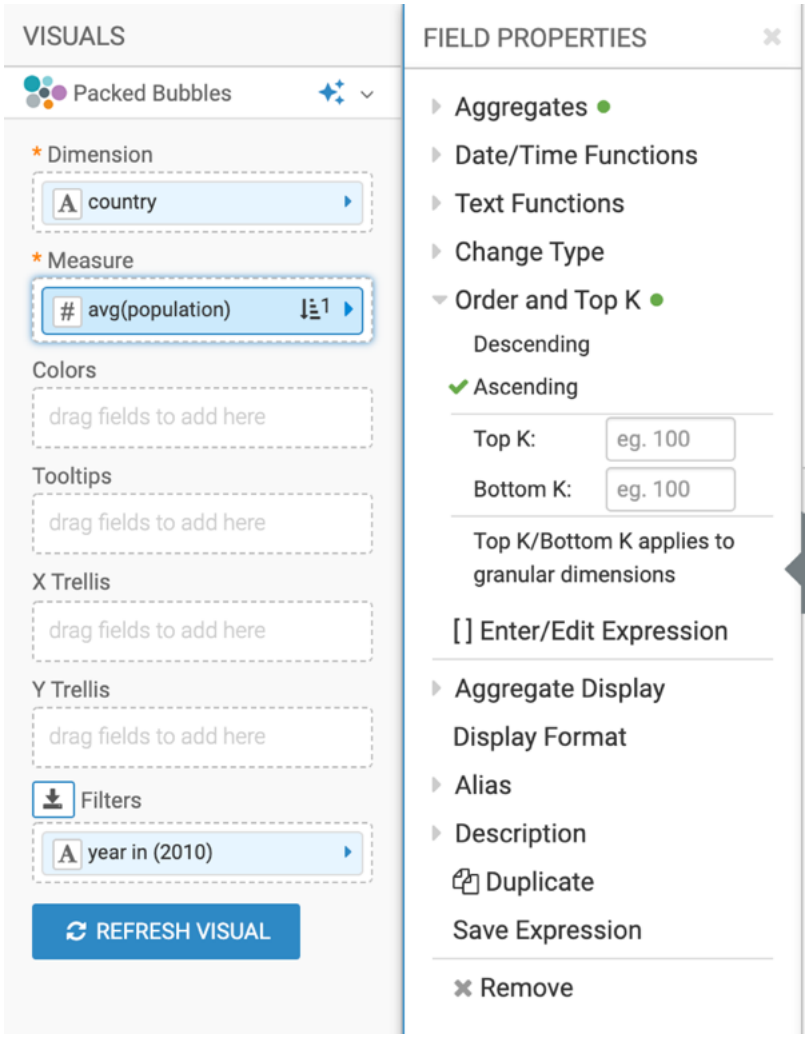
The Packed Bubbles visual appears.

6. On the Filters shelf, add further Dimensions and Measures from the DATA menu.

   This enables you to dynamically control the data input, and discover the data at a more granular level.

   For example, from Dimensions, drag year to the Filters shelf.
7. On the Filter shelf, click the (down arrow) icon on the year field, and select Pick values from a list.

   The Filter for year modal window appears.
8. In the Values tab, select 2010, and click APPLY.
9. On the Measures shelf, click the (down arrow) icon on the avg(population) field, select Order and Top K, and select Ascending.



10. Click REFRESH VISUAL to see the updated Packed Bubbles chart.

11. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

In this example, the title is changed to 'World Population - Packed Bubbles'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

12. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for packed bubbles

Overview of shelves for Cloudera Data Visualization Packed Bubbles visuals.

A Packed Bubbles visual has the following shelves:

Dimensions

Add a Dimension field from the DATA menu to determine the number of bubbles to plot, and their names.

This is a mandatory shelf, and it accepts a single field.

Measures

Add a Measure field from the DATA menu to determine the size of the individual bubbles.

This is a mandatory shelf.

This shelf displays only one field on the visual. Additional fields appear only in the Tooltip.

Colors

To visually differentiate bubbles by category, add a Measure field to this shelf.

This shelf accepts at most one field.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Word cloud

Cloudera Data Visualization allows you to create Word Cloud visuals that aggregate text data and highlight the most frequently occurring words. This visual type is useful for emphasizing prominent terms or phrases in textual datasets. It offers a variety of customizable display settings.

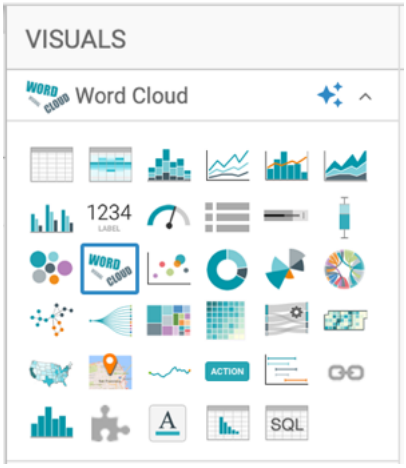
About this task

The following example uses the Shakespeare's Sonnets dataset. To learn how to create datasets, see *Creating a dataset*. For an overview of the shelves that specify this visual, see *Shelves for word clouds*.

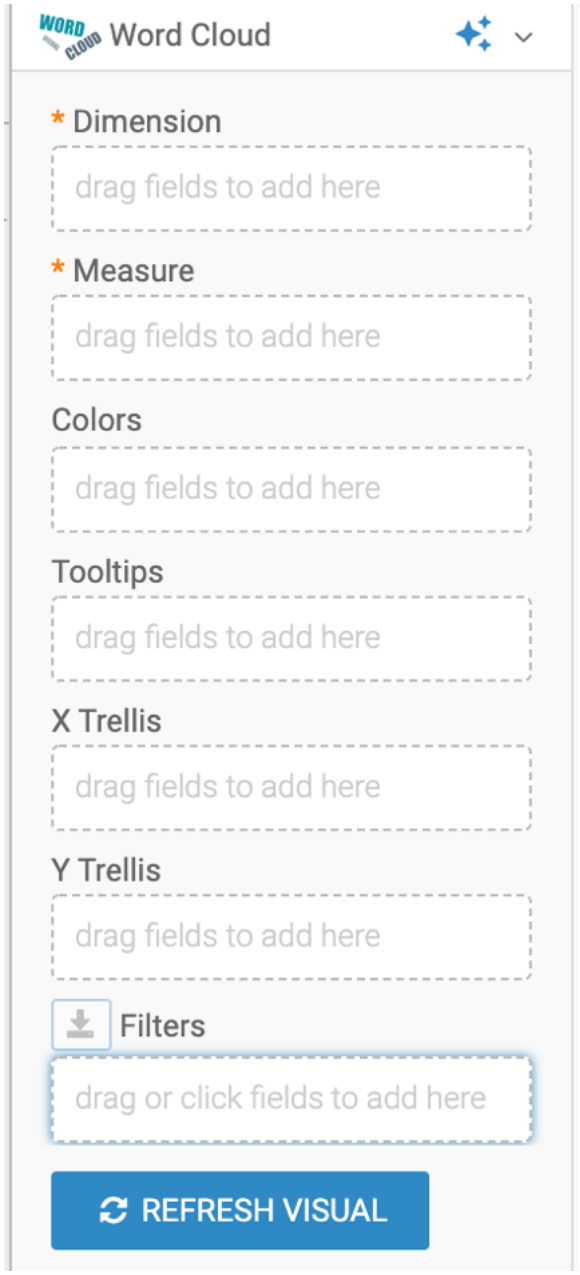
Procedure

1. Start a new visual based on the Shakespeare's Sonnets dataset.

   For detailed instructions on how to get started, see *Creating a visual*.
2. In the VISUALS menu, select Word Cloud.



The shelves for this visual type are: Dimension, Measure, Colors, Tooltips, X Trellis, Y Trellis, and Filters. The mandatory shelves are Dimension and Measure.



3. Populate the shelves using the available dataset fields:
   a) Drag Text to the Dimension shelf.
   b) Drag Record Count to the Measure shelf.
4. Click REFRESH VISUAL.
5. Click the icon next to the visual title to enter a name for it.

   In this example, the title is changed to 'Shakespeare Sonnets - Word Cloud'. Optionally, add a subtitle for a brief description.
6. In the top left corner of the Dashboard Designer, click SAVE.

Results

The resulting visual displays a word cloud where more frequent words appear larger.

What to do next

Customizing the visual

You can adjust how the Word Cloud is displayed using the available settings for this visual type. Explore the settings panel for font styles, color options, word rotation, and more.

Comparing texts

To compare multiple texts visually, consider using trellised word clouds. For more details, see *Trellised word clouds*

Related Information

Creating a dataset

Creating a visual

Trellised word clouds

Cloudera Data Visualization allows you to create trellised Word Cloud visuals to compare text patterns across different categories.

About this task

In the following example, you can compare the text profiles of selected plays from William Shakespeare using trellised word clouds. The example uses the Shakespeare's Plays dataset, created from the shakespeare_play_text.csv file.

Procedure

1. Start a new visual based on the Shakespeare's Plays dataset.

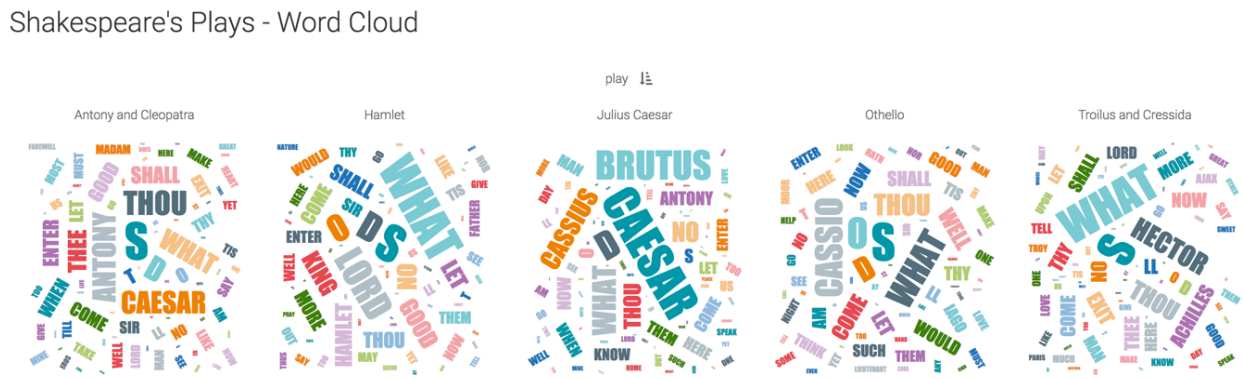
   For detailed instructions on how to get started, see *Creating a visual*.
2. In the VISUALS menu, select Word Cloud.
3. Populate the visual shelves using the following fields:
   a) Add Text to the Dimension shelf.
   b) Add Record Count to the Measure shelf.
   c) Add Play to the X Trellis shelf to trellis the visual.
   d) Add Play to the Filters shelf, select a few plays for comparison.

      In this example the visual is filtered on Anthony and Cleopatra, Hamlet, Julius Caesar, Othello, and Troilus and Cressida.
4. Remove the trellis borders for a cleaner look.

   See *Displaying trellis borders* for instructions.
5. Click REFRESH VISUAL to generate the trellised word clouds.
6. Click SAVE.

Results

You now see individual Word Cloud visuals for each selected play, allowing you to easily compare their language. Prominent words stand out in each visual.



Related Information

Creating a visual

Displaying trellis borders

Shelves for word clouds

Overview of shelves for Cloudera Data Visualization Word Cloud visuals.

A Word Cloud visual has the following shelves:

Dimension

Add a Dimensions field from the DATA menu to determine the number of individual words in the word bubble graph.

This shelf accepts a single field.

Measure

Add a Measures field from the DATA menu to determine the size of the individual words.

This shelf accepts a single field.

Colors

To visually differentiate words by color gradient based on the field value, add a Measures field to this shelf.

This shelf accepts at most one field.

Tooltips

Add Measures fields from the DATA menu for additional information in the graph's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimensions field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimensions field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimensions or Measures fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Scatter

Cloudera Data Visualization enables you to create Scatter visuals. Scatter charts (or scatter plots) have points that show the relationship between two sets of data. They are great for detecting groupings or patterns in a large set of data points.

Scatter visuals have the following additional features in Cloudera Data Visualization:

- You can use color as a third 'dimension'. For more information, see *Creating basic scatter visuals*.
- You can control the size of the marks that represent individual data points, as shown in *Adding size to scatter visuals*.
- You can add a regular increment, such as time, to create an animated representation of data. For more information, see *Adding transition animation to scatter visuals*.
- You can divide the data along natural categories to create side-by-side charts to see individual data points more clearly. For instructions, see *Adding trellises to scatter visuals*.

Note: This visual type supports smooth rendering of streaming data.

For an overview of shelves that specify this visual, see *Shelves for scatter visuals*.

Creating a basic scatter visual

Cloudera Data Visualization enables you to create a basic Scatter visual.

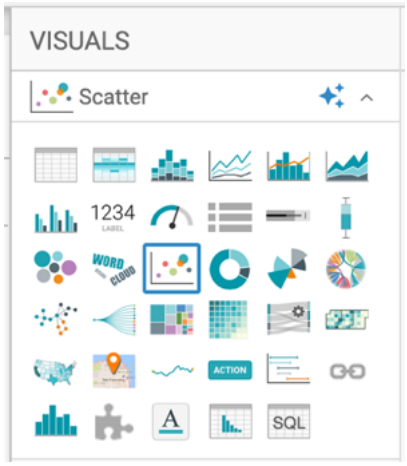
About this task

The following steps demonstrate how to create a new Scatter visual representation, showing the trend of life-expectancy and GDP per capita for selected countries.

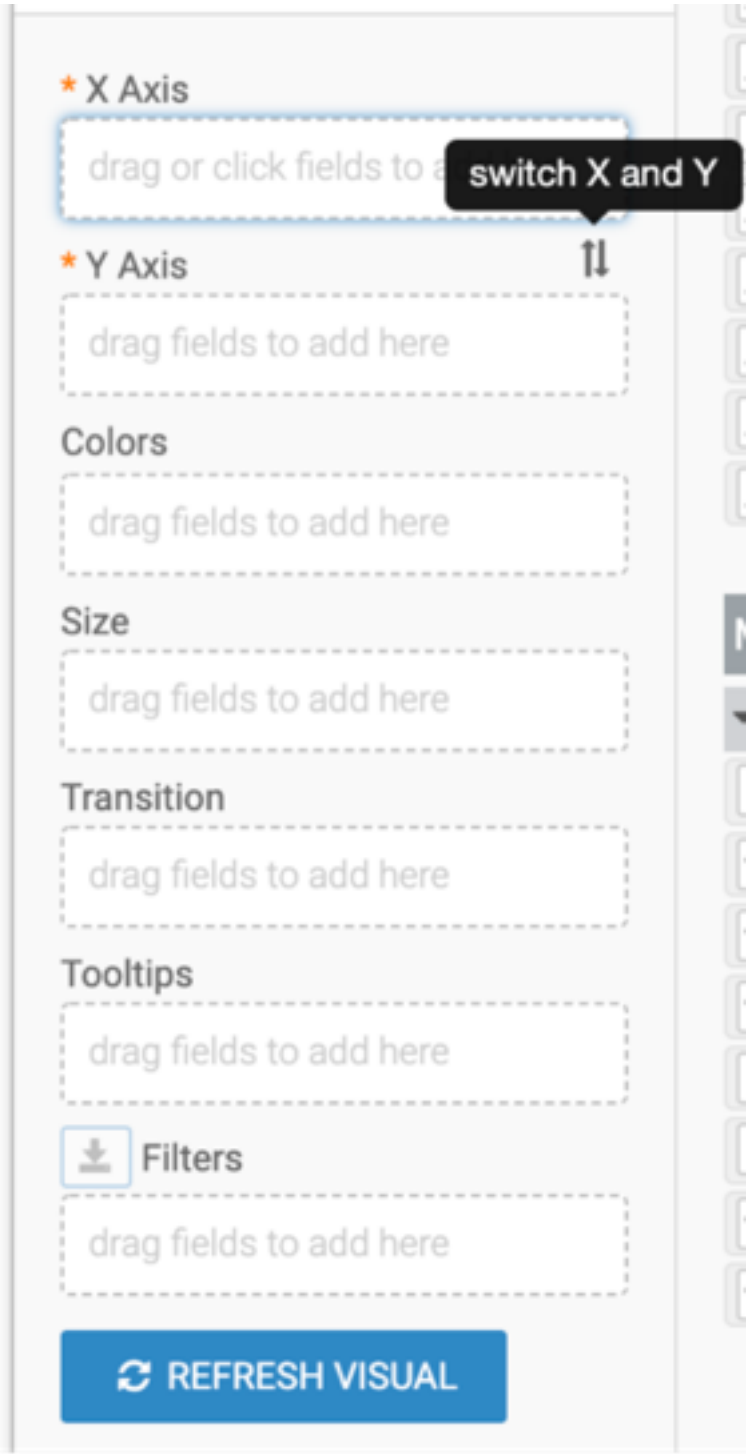
Procedure

1. Start a new visual based on the World Life Expectancy dataset. The data source for this is samples.world_life_expectancy.

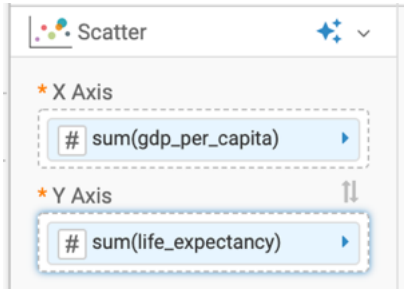
   For instructions, see *Creating a visual*.
2. In the visuals menu, find and click Scatter.



The shelves of the visual changed. They are now X Axis, Y Axis, Colors, Size, Transition, Tooltips, and Filters. The mandatory shelves for Scatter visuals are X Axis and Y Axis. The fields placed on these two shelves may be easily swapped by switching X Axis and Y Axis.

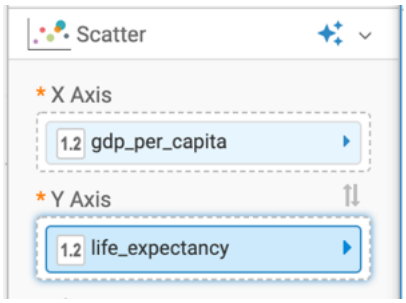


3. Populate the shelves from the available fields of Dimensions and Measures in the DATA menu.
   a) Under Measures, select gdp_per_capita and drag it to the X shelf.
   b) Under Measures, select life_expectancy and drag it to the Y shelf.



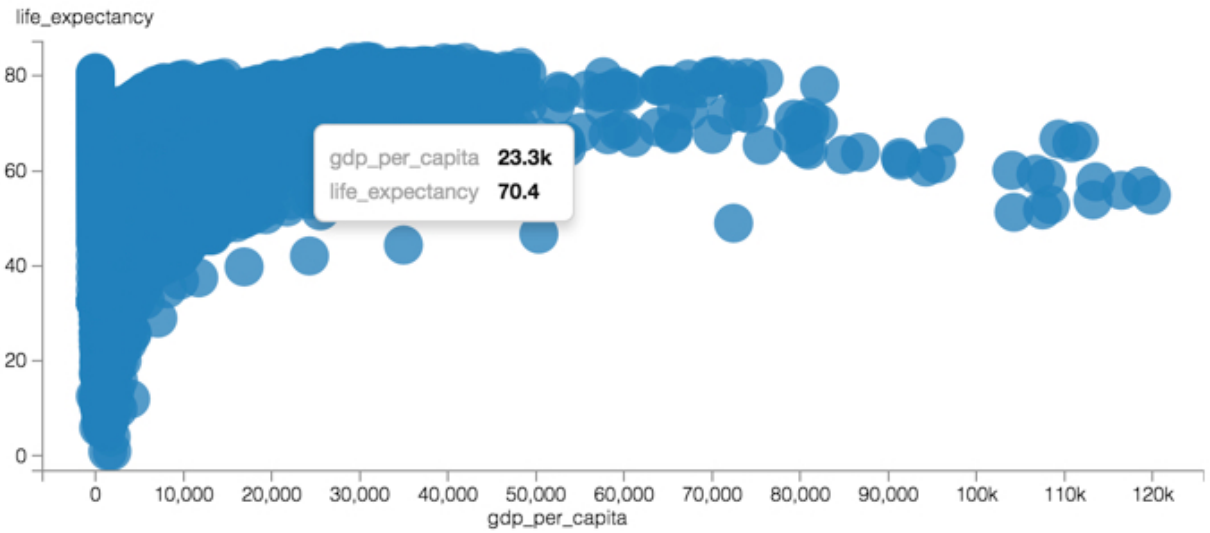
4. In both cases, remove the aggregate to see individual data points.
   a) On the shelf, on the sum(gdp_per_capita) field, click the (right arrow) icon to open the FIELD PROPERTIES menu.
   b) Expand Aggregates and click the check mark next to Sum to remove the aggregate.
   c) Repeate it for the sum(life_expectancy) field

   The shelves now contain the modified fields.



5. Click REFRESH VISUAL.

   The Scatter visual appears.

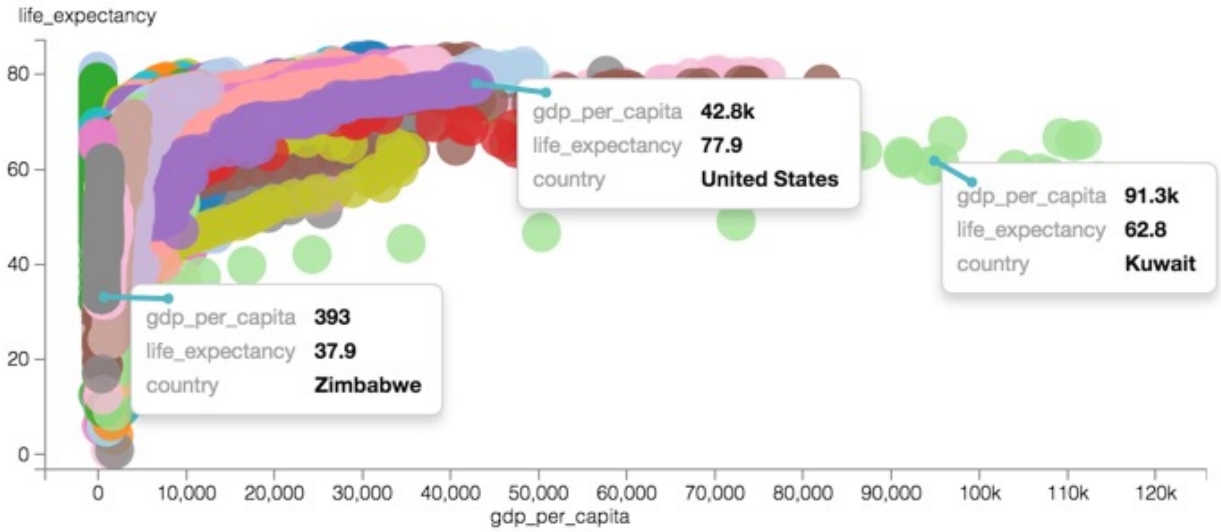


   While you can see the general shape of data and a few outliers, there is too little distinguishing information to help us understand the trends.

6. Use the Colors shelf to see if a pattern emerges.

   Under Dimensions, select country and drag it to the Colors shelf .

7. Click REFRESH VISUAL and examine the resulting graph.

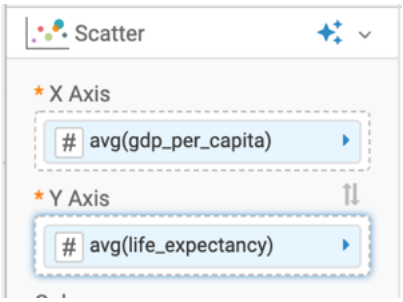


You can now see the trend for some of the countries very clearly, over time. The visualization is less than ideal, yet obvious details show:

- Zimbabwe, in gray, shows an appreciable increase in life expectancy, but relative to other countries, very little economic improvement (measured as GDP per capita).
- United States, in purple, shows the expected improvement in both longevity and in GDP.
- Kuwait, in light green, is a true outlier that shows phenomenal increase in GDP per capita, but suffers a noticeable decrease in life expectancy from about the mid-century to present day.

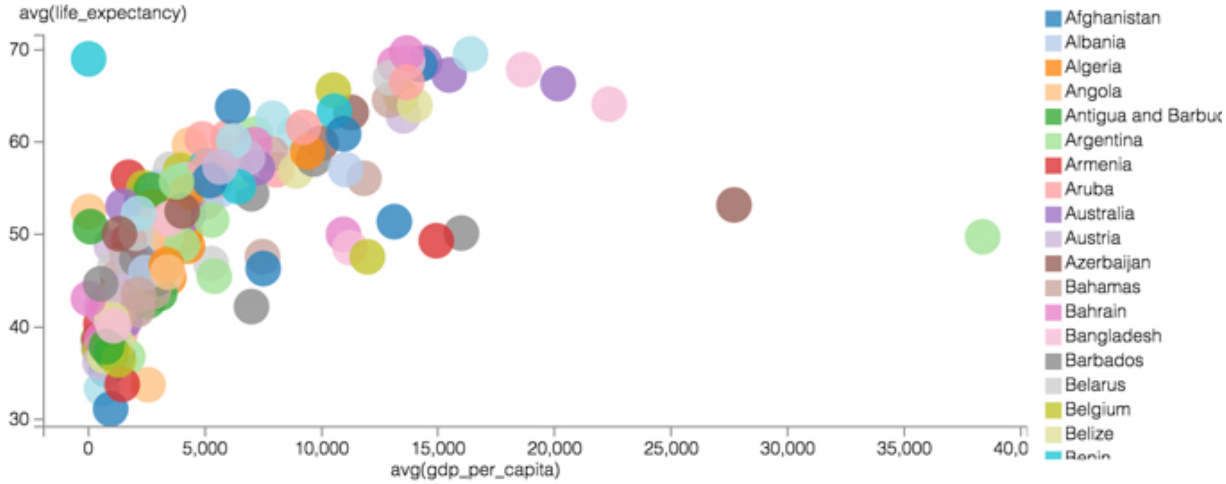
8. Add aggregation back to the measures on the X Axis and Y Axis shelves. This time, you want to see the average of the dimensions.
   a) On the shelf, on the gdp_per_capita field, click the (right arrow) icon to open the FIELD PROPERTIES menu.
   b) Expand Aggregates and select Average.
   c) Repeate it for the life_expectancy field

   The shelves now contain the modified fields.



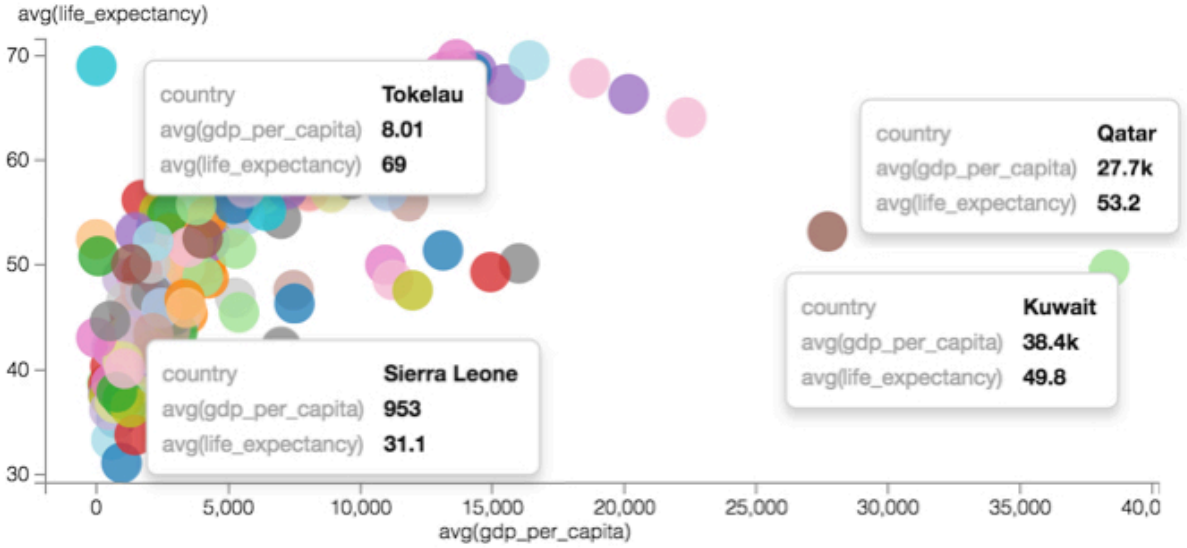
9. Turn on the *Changing legend style and removing the legend* feature.
10. Click REFRESH VISUAL.

    The Scatter visual appears.



The distribution is an average of ALL years covered by the dataset, from 1990 through 2010. Still, some outliers are already clearly visible: Qatar and Kuwait for exceptionally high GDP per capita, Tokelau for high life expectancy at very low GDP per capita, and Sierra Leone, with the lowest life expectancy in the world.

Figure 1: Contrasting Average Life Expectancy and GDP Per Capita, World-Wide



11. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

In this example, the title is changed to 'World Population - Scatter'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

12. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

It is useful to remember at this point that GDP per capita is actually influenced by the population of the country.

As a next step, look at how you can show the population variation on this graph, by *Adding size to a scatter visual*.

Adding size to scatter visuals

In Scatter visuals, you can control the size of the individual bubbles to show variation of a measure more clearly.

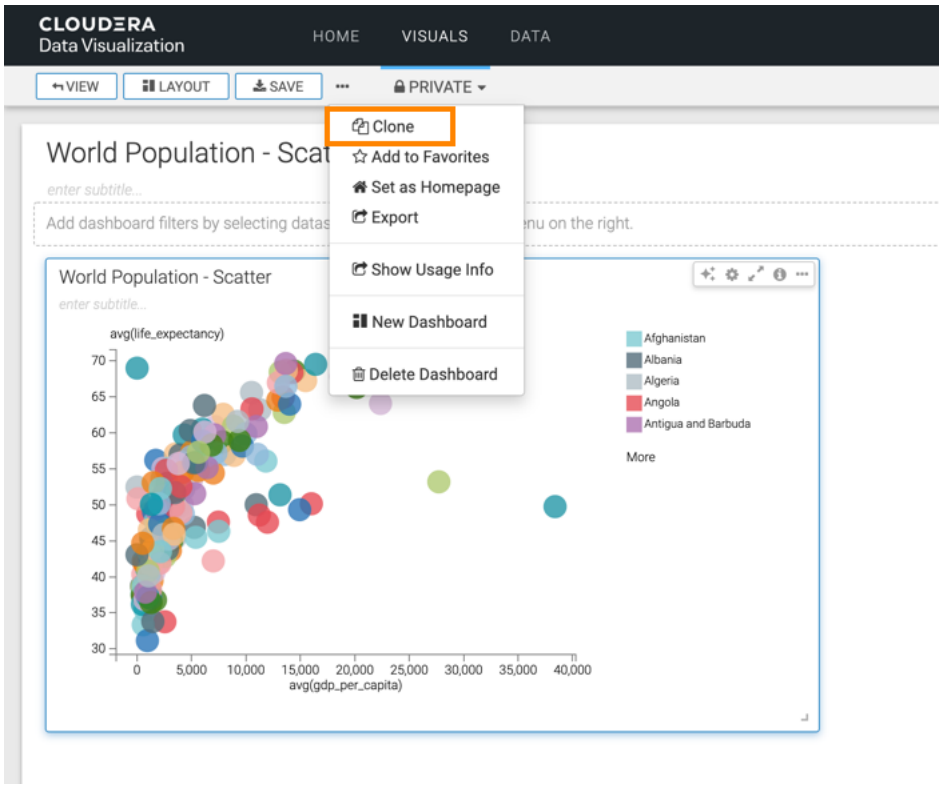
About this task

The World Life Expectancy dataset has several other dimensions and measurements that you can use to enrich the scatter visual. Consider how using the Size shelf, which controls the size of the individual bubbles, changes our immediate understanding of the dataset and represents the relative population size of each country more clearly.

The starting point of the following example is the 'World Population - Scatter' visual developed in *Creating basic scatter visuals*.

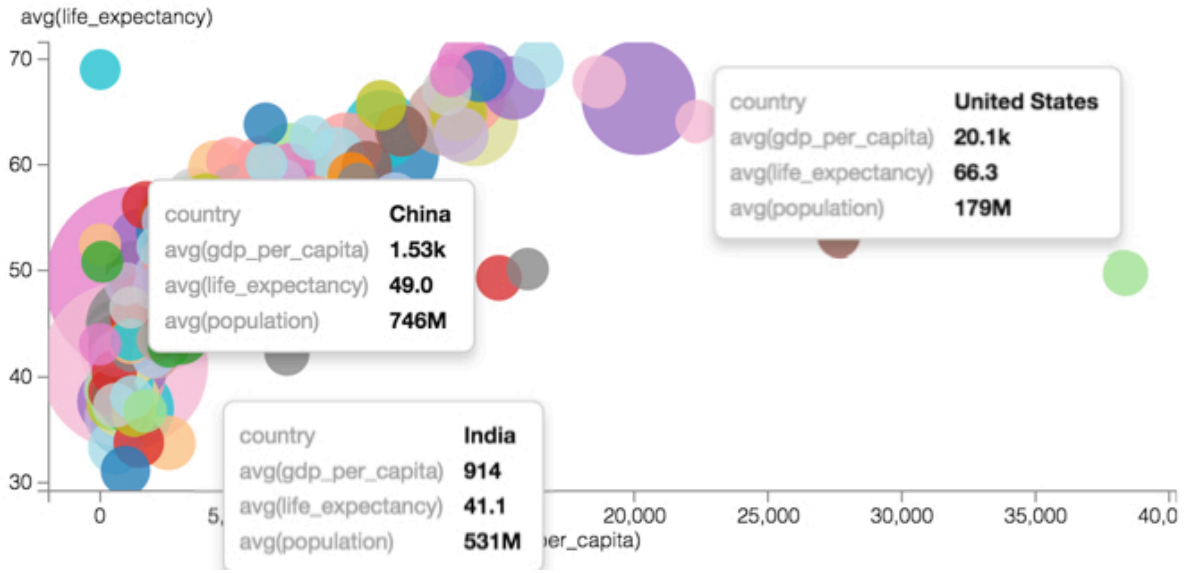
Procedure

1. Click Clone to create a new visual that has the same properties as the original.



2. Under Measures, select population and drag it to the Size shelf.
3. On the sum(population) field, click the (right arrow) icon to open the FIELD PROPERTIES menu.
4. Expand Aggregates and select Average.
5. Click REFRESH VISUAL.

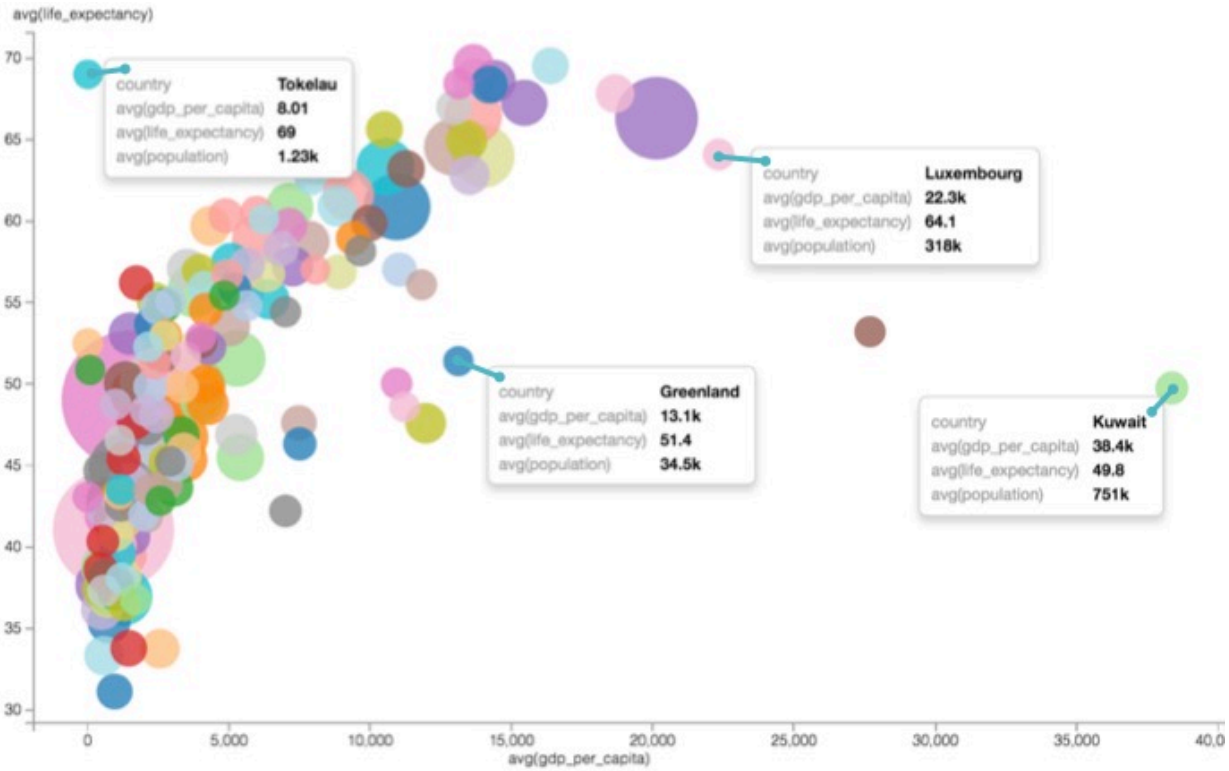
   The Scatter visual appears with its own visible population size 'outliers': China, India, and the United States.



6. It is important to realize that the relative size of the bubbles is not linearly proportional to their values.

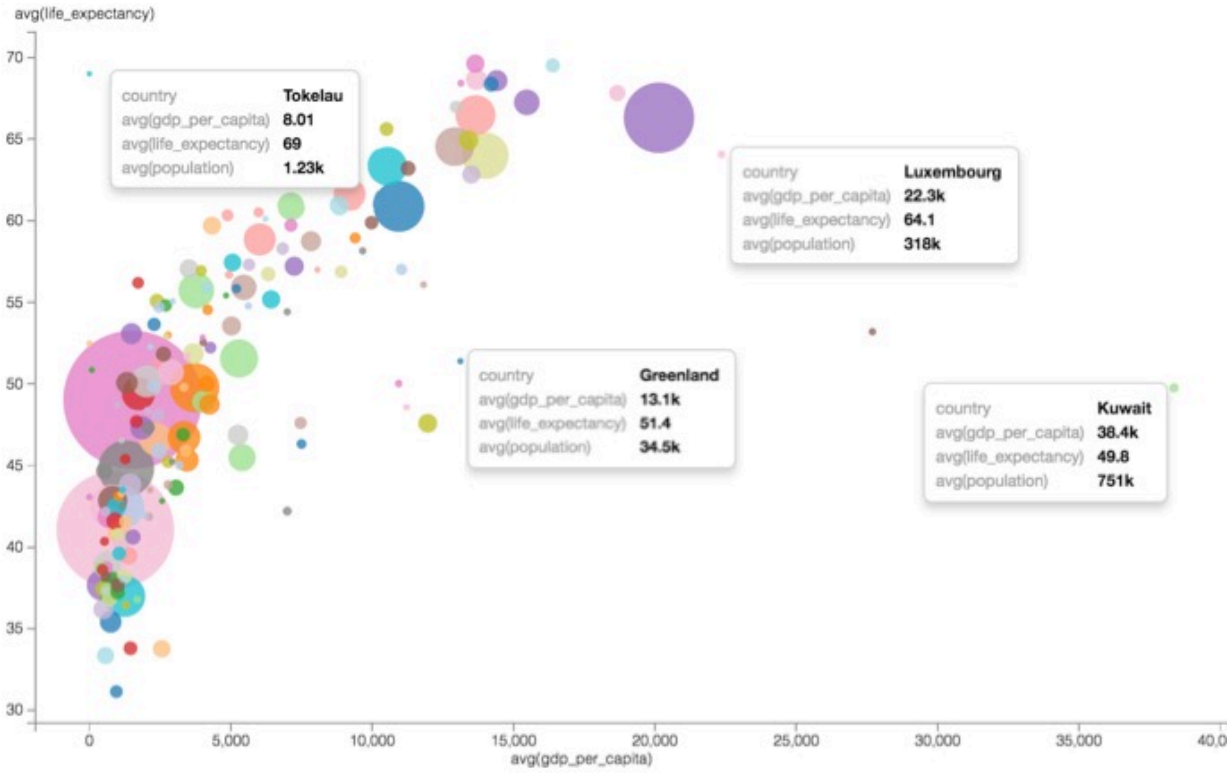
   At the lower range of population levels, the visual shows the same size marks for countries with very different average populations, such as Tokelau, where avg(population)=1.23K, Greenland, with avg(population)=34.5K, Luxembourg, with avg(population)=318K, and Kuwait, with avg(population)=751K. This is because by default, the size of marks ranges from 10 to 50 pixels. Divided equally across the range of values, this squeezes too many disparate measurements on the lower end to use the same mark size 'bucket'.

Figure 2: World Population Figures with Mark Size Range of 10 - 50 Pixels



It is easy to somewhat alleviate this problem by changing the range of sizes of the bubbles, as described in *Changing the mark size range*. This graph shows the size range of 1-50.

Figure 3: World Population Figures with Mark Size Range of 1 - 50 Pixels



Incidentally, and obviously, the increased size range generates a more realistic representation of relative population size.

7. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'World Population - Scatter with Size'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
8. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

While you can see the average data for individual countries more clearly, these measurements do not help you to see the changes that occurred over the entire time domain of the dataset.

As a next step, look at how you can add the concept of time back into the visual by animating the scatter plot, as described in *Adding transition animation to scatters*.

Adding transition animation to scatter visuals

Scatter visuals have a powerful feature called transition animation.

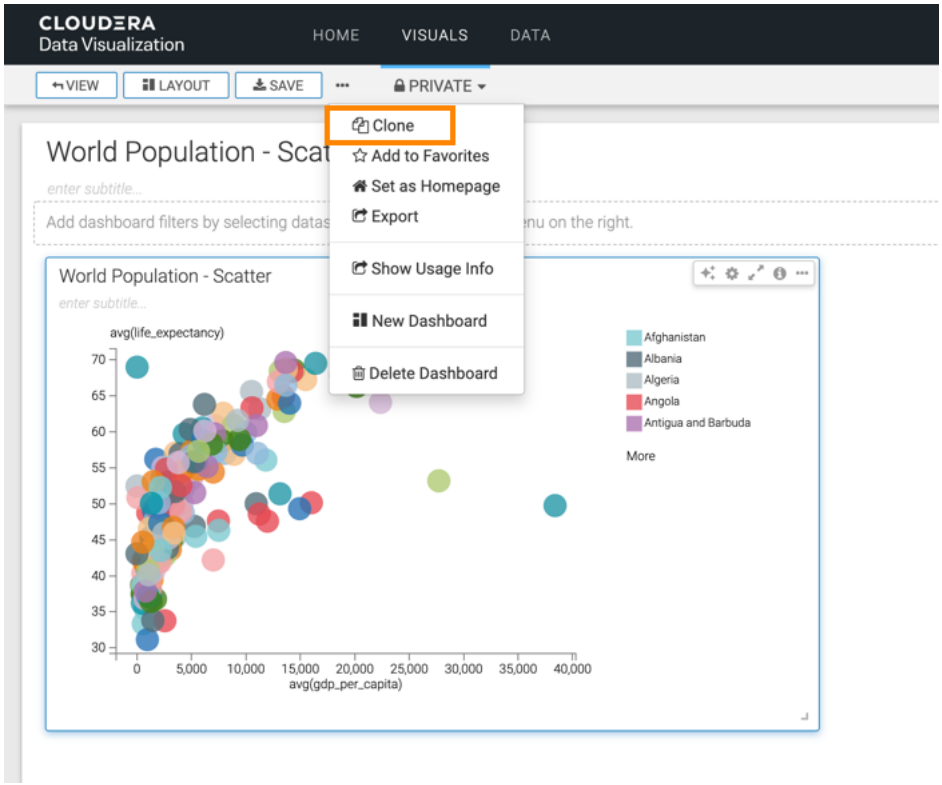
About this task

In the World Life Expectancy dataset, the time is used as the transition 'axis' to show how GDP per capita, life expectancy, and population size changes over time, from year 1990 through year 2010.

The starting point of the following example is the 'World Population - Scatter with Size' visual developed in *Adding size to scatters*.

Procedure

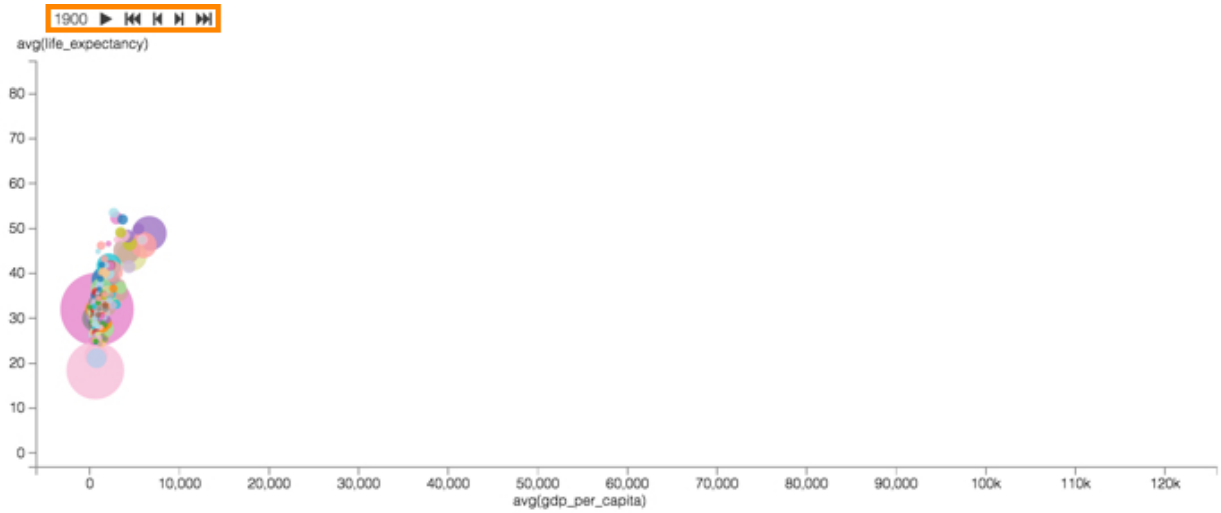
1. Click Clone to create a new visual that has the same properties as the original.



2. Under Dimensions, select year and drag it to the Transition shelf.
3. On the sum(population) field, click the (right arrow) icon to open the FIELD PROPERTIES menu.
4. Expand Aggregates and select Average.
5. Click REFRESH VISUAL.

   The Scatter visual refreshes, but this time you can see that it has animation controls at the very top, over the vertical axis legend. The controls common to all video animations: 'play' (which changes to 'pause' when the animation is active), 'back to start', 'back-up', 'advance', and 'to end'.

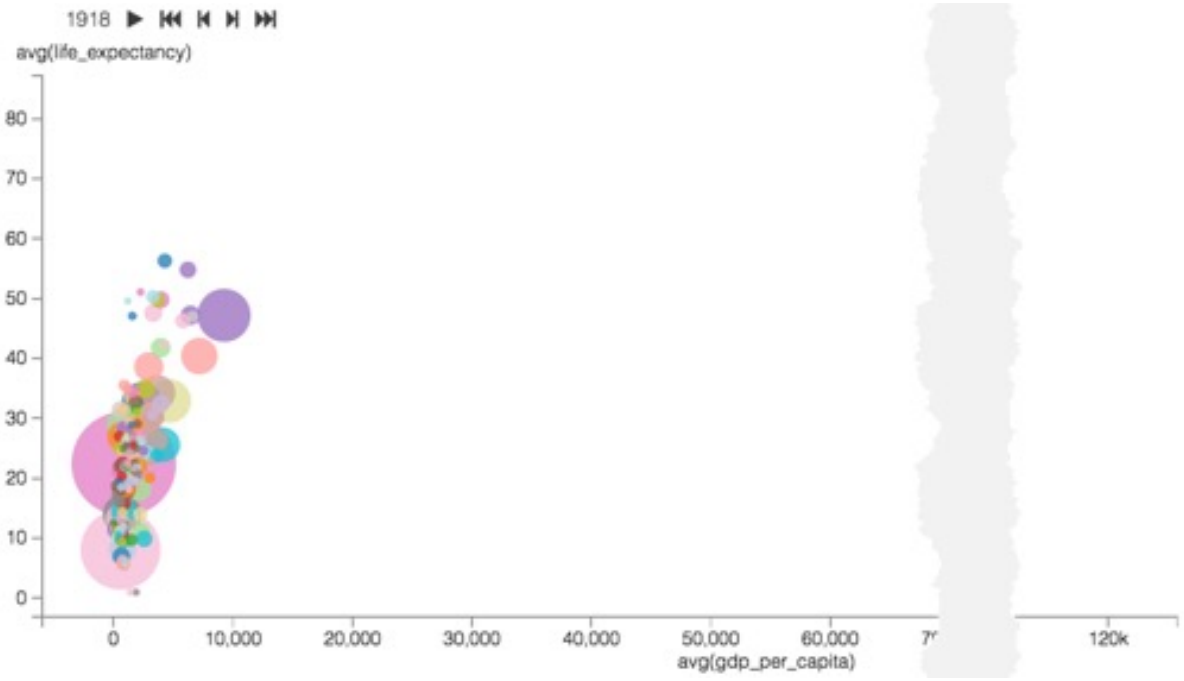
   You can also notice that both axes extend far beyond the data points for the initial conditions, year=1990. This is because the cover the entire range of measurements for all data.



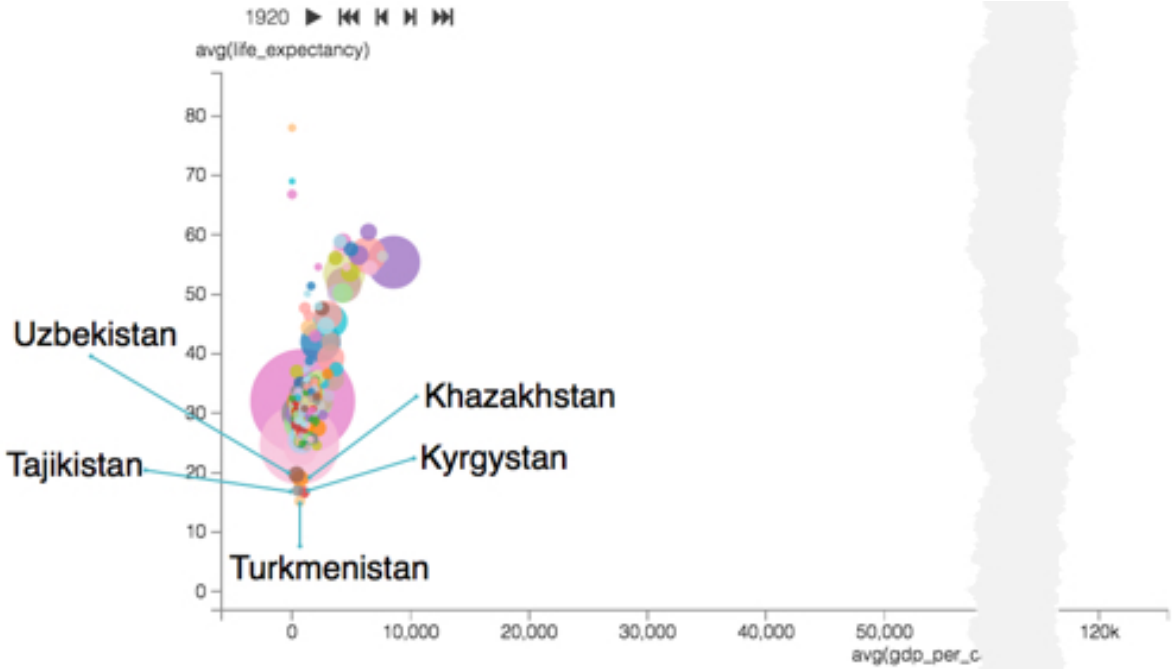
6. Turn on the gridline option, as described in *Enabling gridlines*.
7. Adjust the speed of animation, as described in *Changing transition delay*.

8. Press play and see the changes in data.
   Here are some highlights:

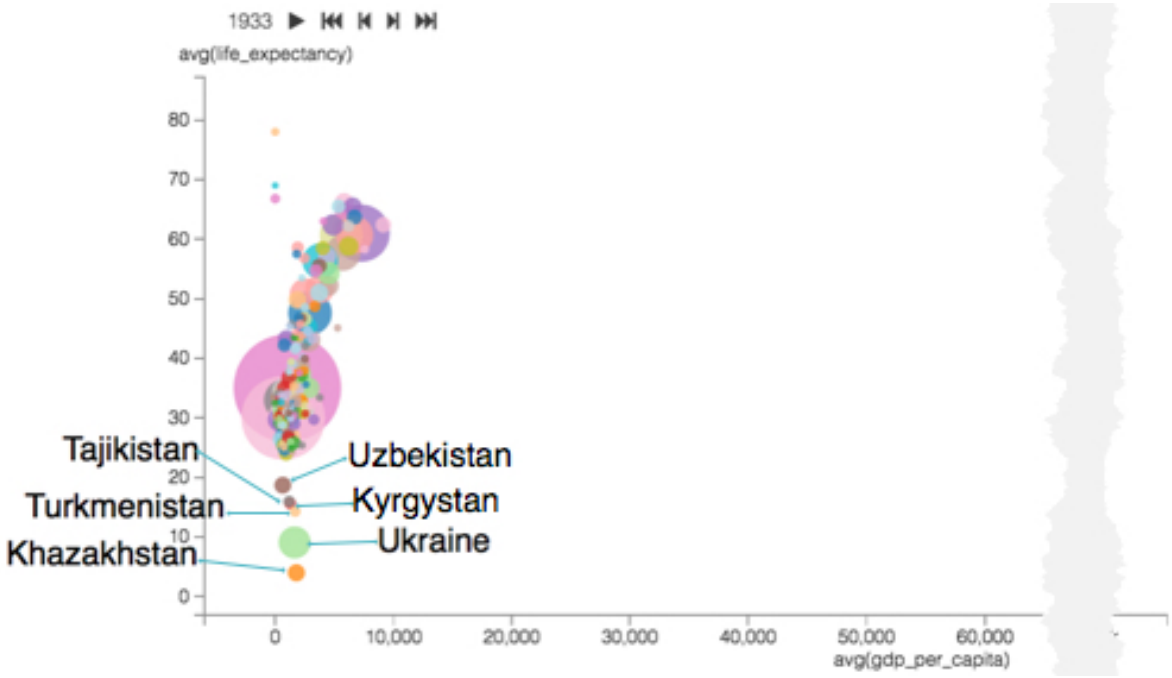
   - The World-Wide downward bounce in life expectancy centered on 1918, due to the Spanish Flu pandemic.



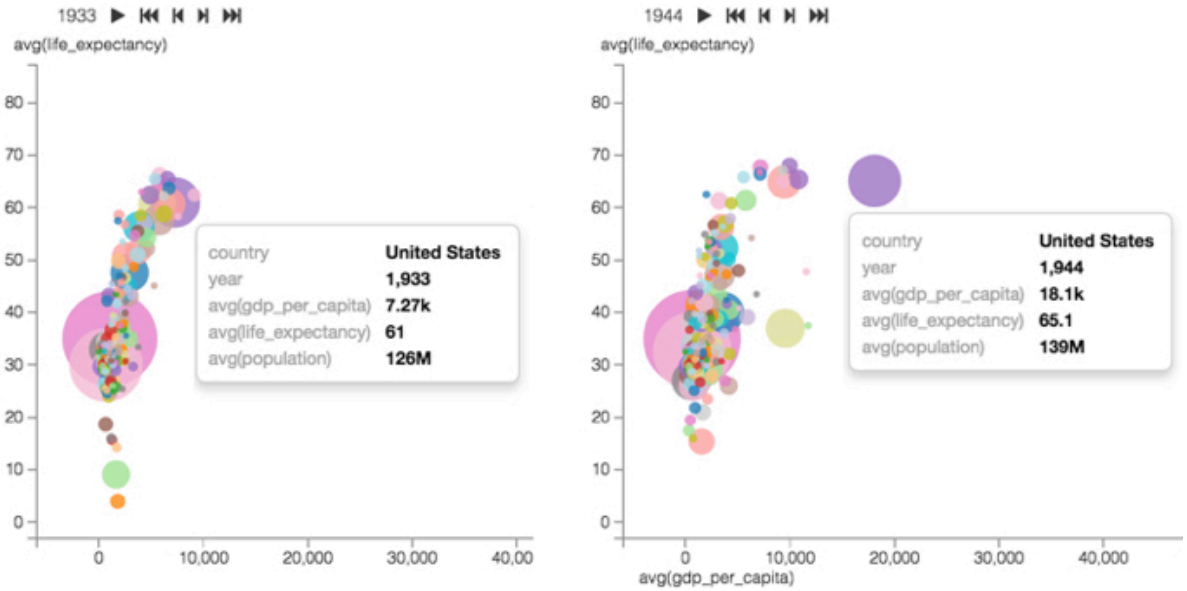
   - The Soviet Union secularization efforts and agricultural policies in its Asian territories resulted in war and famine that claimed many lives in Uzbekistan, Kazakhstan, Tajikistan, Kyrgyzstan, and Turkmenistan. This is reflected in the precipitous decrease in life expectancy for these countries in 1919-1923, with heaviest losses in 1920.



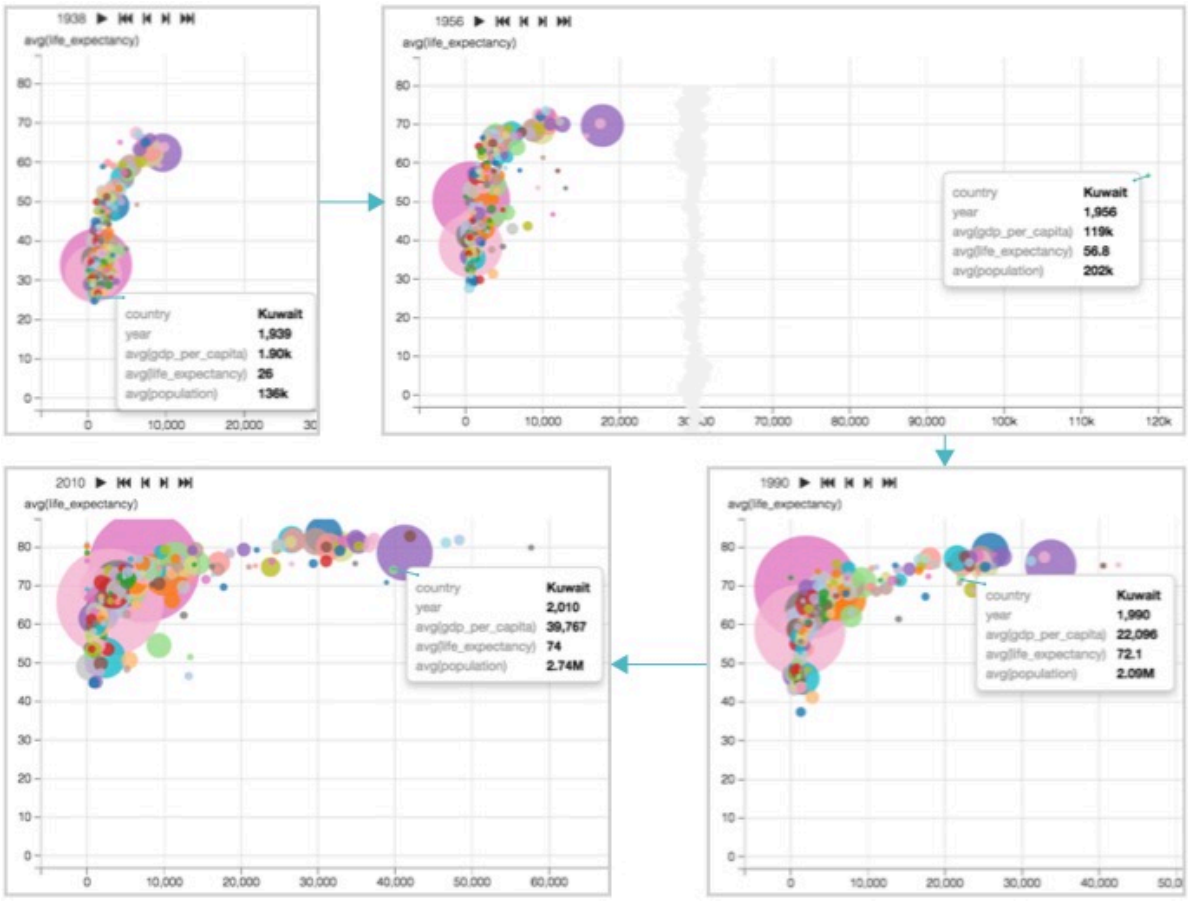
   - There was another significant decrease in life expectancy due to the loss of as many as 8 million lives in the Soviet famine of 1932–33 and more specifically, the Hlodomor. Soviet Union farm collectivization in the Russia and Ukraine withheld grain from farmers and collectivized farms until unrealistic production quotas were be met.



- The United States economy grew rapidly from mid-1930s through mid-1940s, because New Deal's economic relief and work creation policies stabilized the country after the Great Depression, while the World War II material support of the Allies and its own involvement in Europe, Africa, and the Pacific meant even greater production of materials, armaments, and two new naval fleets. You can see that GDP per capita increased nearly 3-fold, while the population increase and life expectancy were relatively modest.



- The exceptional increase in GDP in a very modest country of Kuwait after the discovery of the world's sixth largest oil deposits in 1938, through 1956 peak levels of GDP per capita growth, and large scale modernization from 1946 through 1982. After a period of international instability in 1980s, the 1990 invasion by Iraq destroyed much of the infrastructure. After 1991 US intervention, the country has been rebuilding its infrastructure.



9. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'World Population - Scatter with Transition'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

10. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

While we can see the data for individual countries more clearly, it is often useful to contrast well-defined logical segments of data.

As a next step,you can compare and contrast GDP per capita and life expectancy for different UN Regions in *Adding trellises to scatter visuals*.

Adding trellis to scatters

In Scatter visuals, you can add trellises for comparing and contrasting data. To discover variation within the dataset, to see both similarities and differences among categorical subsets of data, it is very useful to see visuals for these categories of data side by side. In Cloudera Data Visualization, this is very easy to do with the trellis option, which is essentially a doubling down on one or both of the axes.

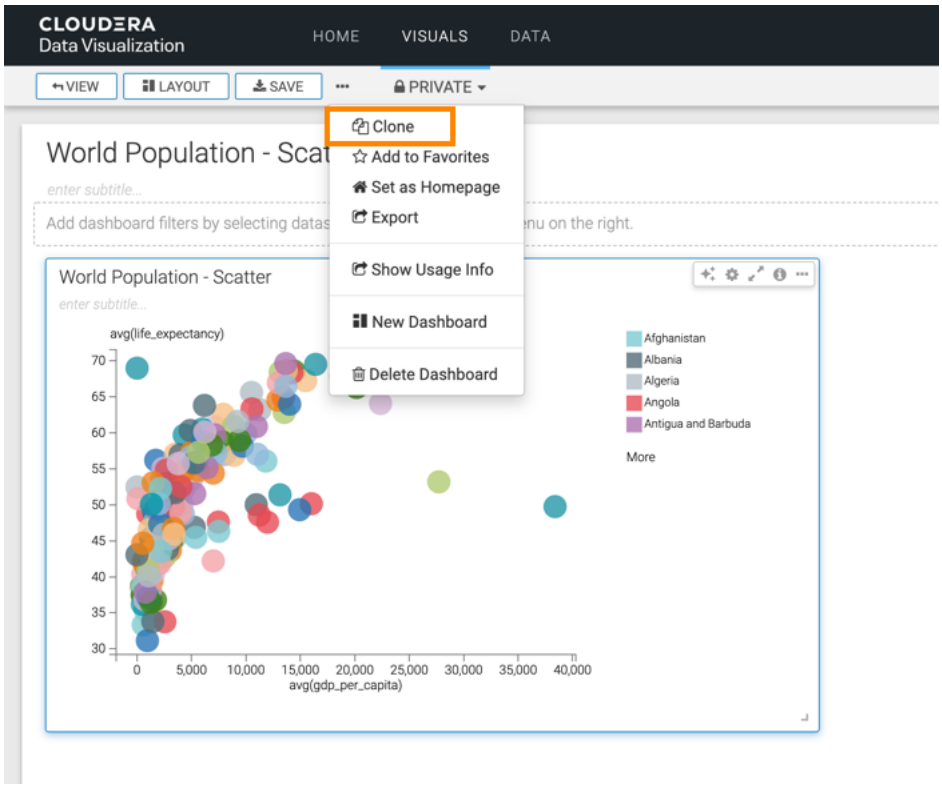
About this task

In the World Life Expectancy dataset, you can use un_region to compare how life expectancy and GDP per capita vary in different geographic zones. For an even greater degree of detail, you could use un_subregion.

The starting point of the following example is the 'World Population - Scatter with Transition' visual developed in *Adding transition animation to scatters*.

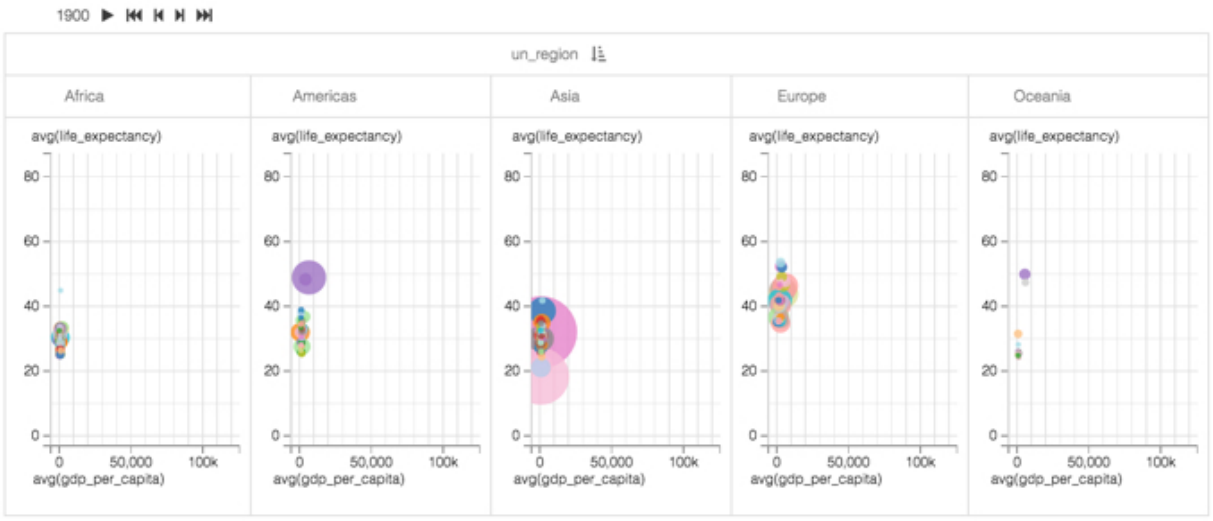
Procedure

1. Click Clone to create a new visual that has the same properties as the original.



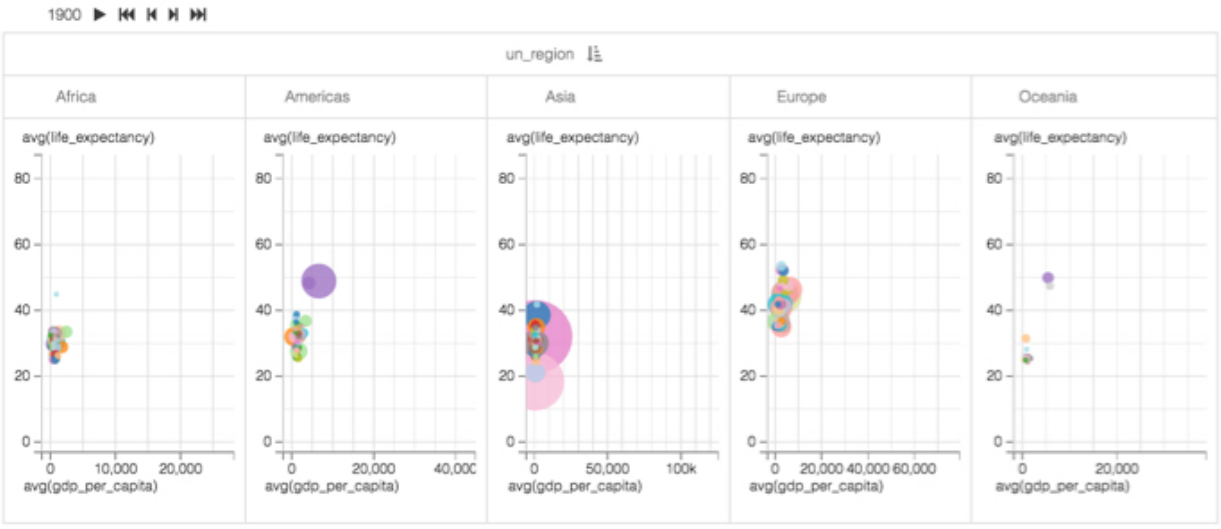
2. Under Dimensions, select un_region and drag it to the X Axis shelf.
3. Click REFRESH VISUAL.

   The Scatter visual refreshes, but this time you can see that it is divided into five separate graphs, that correspond to the identified UN Regions: Africa, America, Asia, Europe, and Oceania. There is a single animation control at the top, so when you click Play, all UN Regions animate simultaneously.

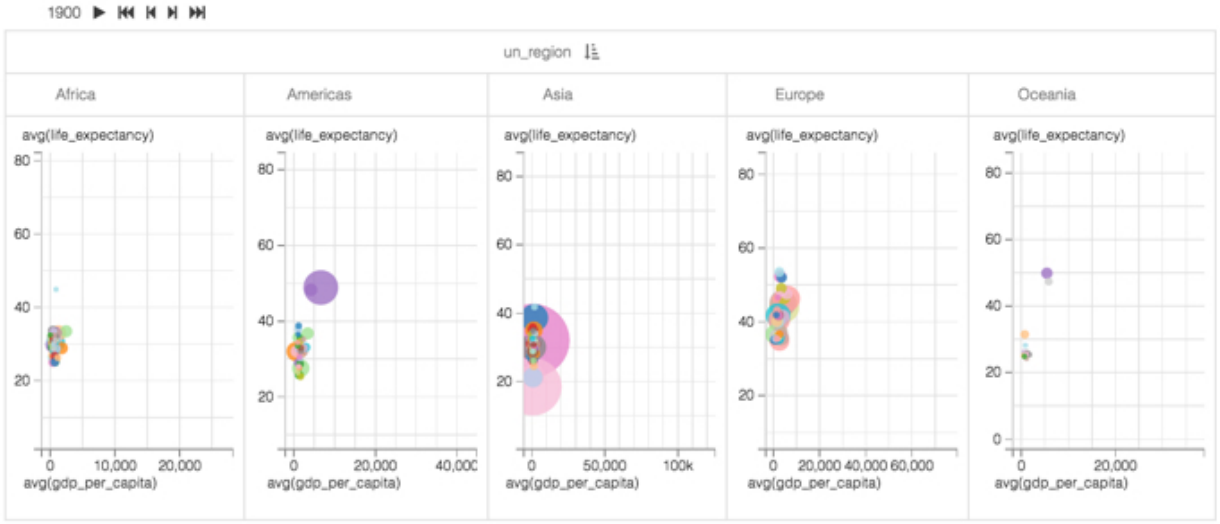


   After running the animation, you will notice that the horizontal axes extend far beyond the necessary space for some of the areas. In particular, the GDP per Capita scale seems too large for Africa.
4. To have each graph use its own optimized scale for the horizontal axis, enable the independent x-axis option, as described in *Enabling independent axis scale*, and click REFRESH VISUAL.

You can see that each area has its horizontal scale.



5. Similarly, choose the independent Y axis option. The differences in vertical scaling are slight, but still noticeable.



6. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'World Population - Scatter with Transition and Trellis'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

7. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for scatters

Overview of shelves for Cloudera Data Visualization Scatter visuals.

A Scatter visual has the following shelves:

X Axis

This is conceptually the X axis. Use a Dimensions field from the DATA menu to plot the scatter horizontally, or a Measures field to plot the scatter vertically.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in columns.

Y Axis

This is conceptually the Y axis. Use a Measures field from the DATA menu to plot the scatter horizontally, or a Dimensions field to plot the scatter vertically.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in rows.

Colors

To visually differentiate bubbles by category, add a Dimensions field to this shelf.

This shelf accepts at most one field.

Size

Add a Measures field from the DATA menu to show an additional measurement as a relative size of the bubbles on the scatter graph.

This shelf accepts a single field.

Transition

Add Dimensions or Measures fields from the DATA menu to animate the graph based on the sequence defined by this field.

A good example would be time fields, such as timestamp, day, month, and so on. Cloudera recommends that you apply a sequential sort to ensure a natural transition.

This shelf accepts a single field.

Tooltips

Add Measures fields from the DATA menu for additional information in the visual's tooltip.

This shelf accepts multiple fields.

Filters

Add Dimensions or Measures fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Pie

Cloudera Data Visualization enables you to create Pie charts.

About this task

Pie charts are circular statistical graphics that show numerical proportion through relative size of slices; the arc length of each slice is proportional to the quantity it represents. In Cloudera Data Visualization, pie charts use a default 'doughnut' representation. However, you can change them to the more traditional closed pie configuration by controlling the arc width of the circle, as described in *Changing arc width*.

The following steps demonstrate how to create a new Pie visual on the Capital Expenditure by Industry dataset. You must first create this dataset, following the steps described in *Creating datasets*.

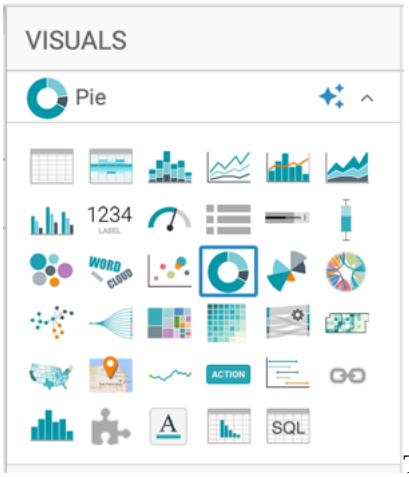
For an overview of shelves that specify this visual, see *Shelves for pie visuals*.

Note: This visual type supports smooth rendering of streaming data.

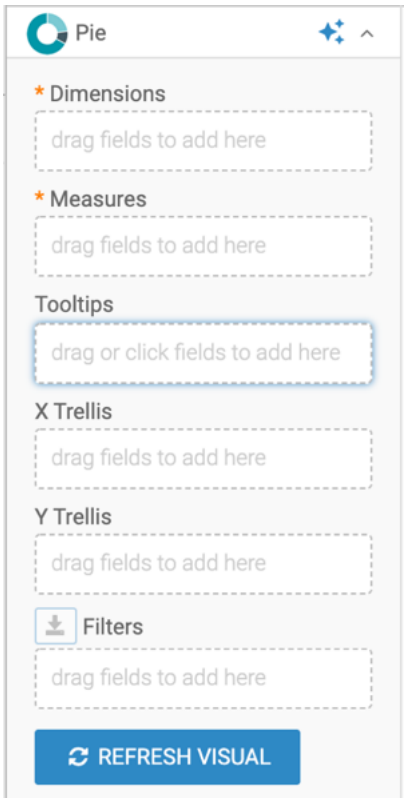
Procedure

1. Start a new visual based on the Capital Expenditure by Industry dataset.
   For instructions, see *Creating a visual*.

2. In the VISUALS menu, find and click Pie.

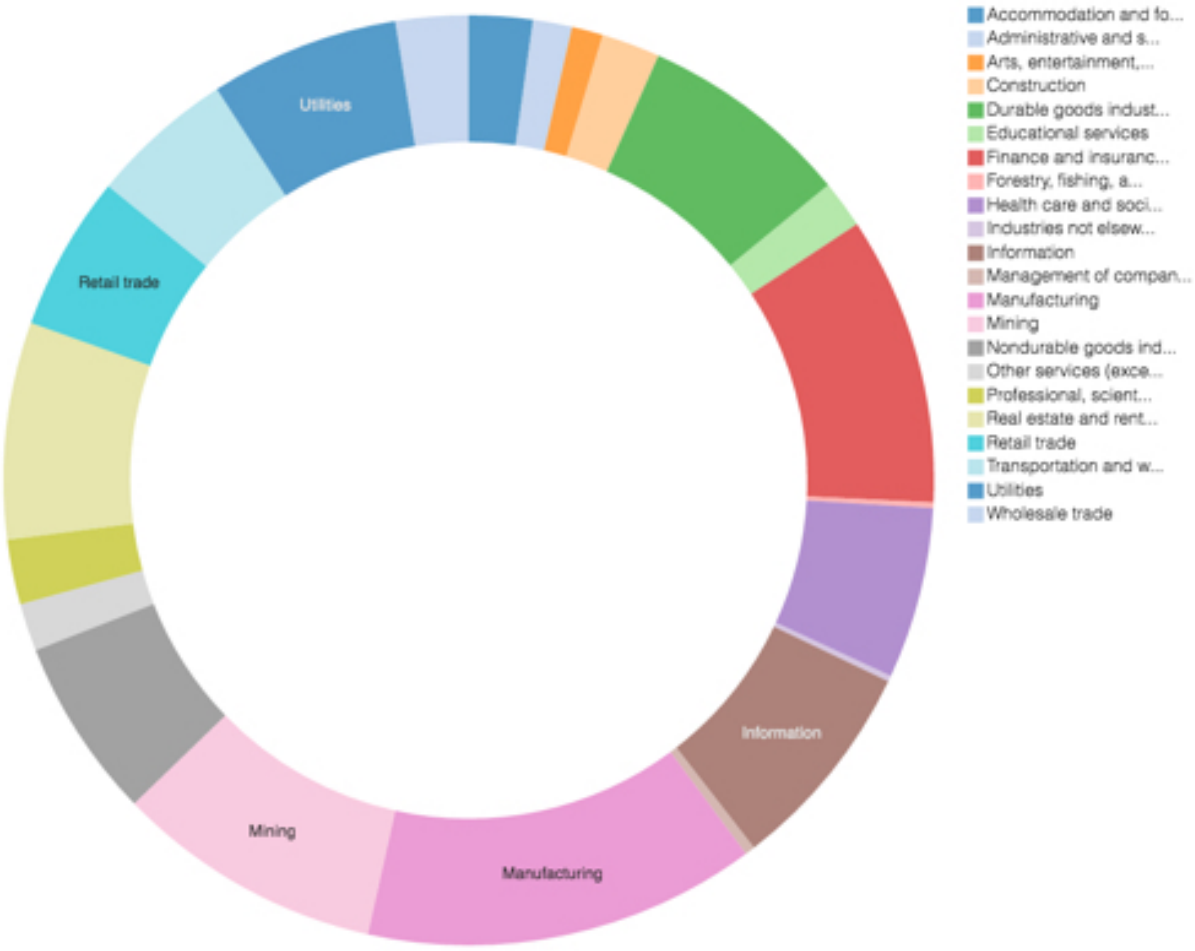


The shelves of the visual changed. They are now Dimensions, Measures, Tooltips, X Trellis, Y Trellis, and Filters. Both Dimensions and Measures are mandatory.



3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Dimensions, select industry and drag it to the Dimensions shelf.
   b) Under Measures, select expenditure and drag it to the Measures shelf.
4. Click REFRESH VISUAL.

   The Pie visual appears.

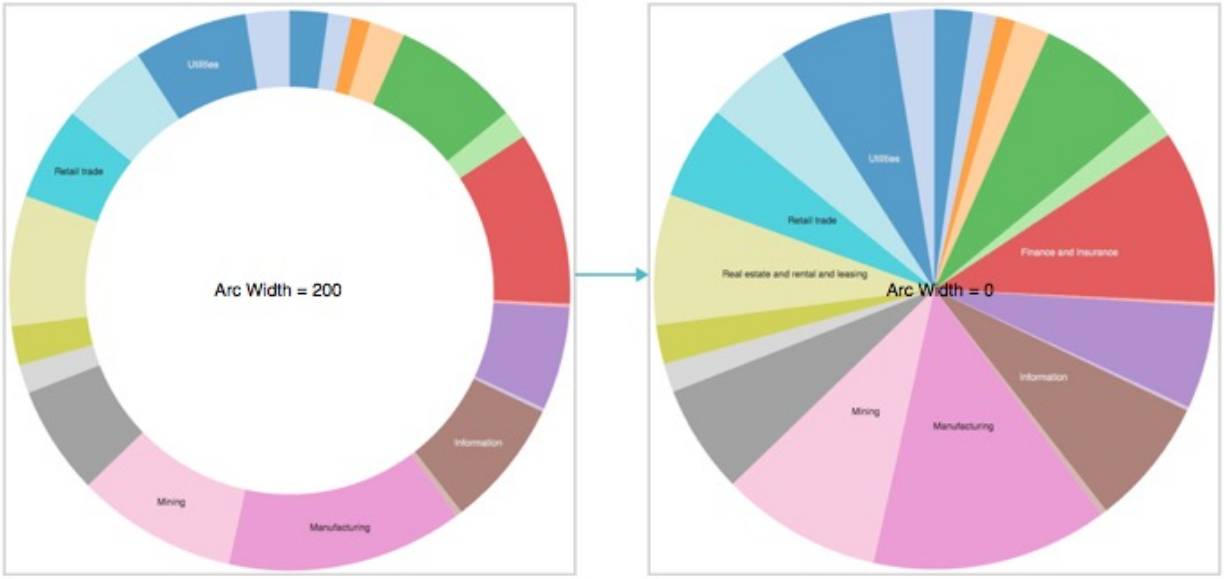


5. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'Capital Expenditures - Pie'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
6. At the top left corner of the Dashboard Designer, click SAVE.
7. Adjust the arc width to close the hole in the middle of the 'doughnut'.
   a) From the visual Settings menu, select the Marks tab, and set Arc Width to 0.

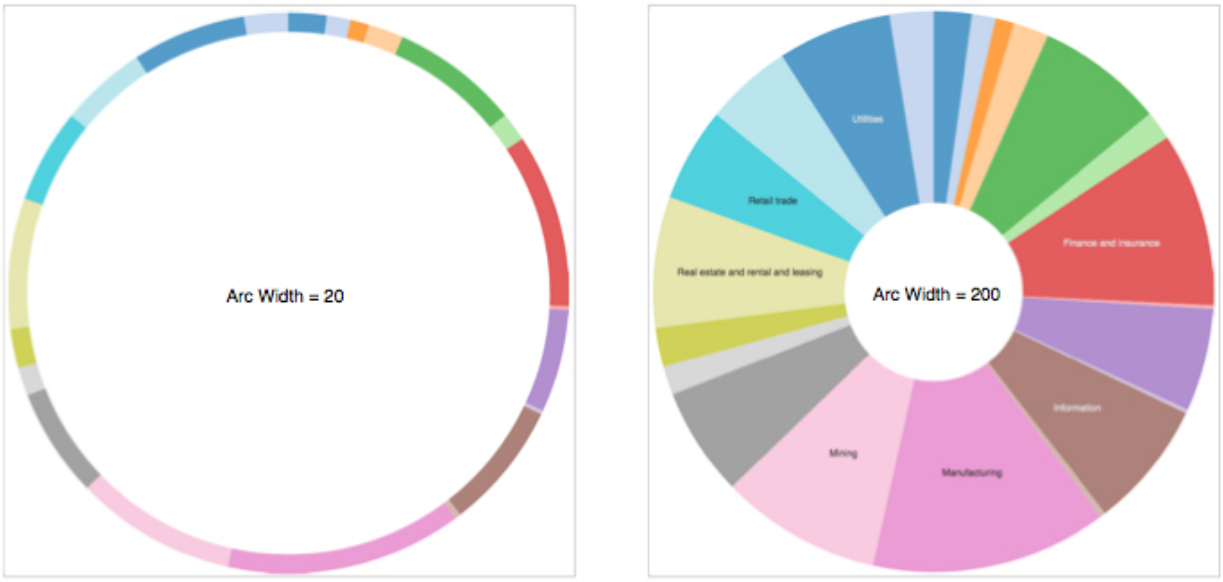
   The default arc width is 80.

Figure 4: Closing the Hole; Changing Arc Width from 80 to 0



You can also experiment with changing the thickness of the 'doughnut'. See *Changing arc width* for details on how to adjust the thickness of the 'doughnut'.

Figure 5: Contrasting Arc Widths: 20 and 200



8. You can adjust the default percentage reporting inside the Tooltip.
   a) From the visual Settings menu, select the Tooltip.
   b) Change the Label for Percentage to 'Percent of Total' and the Decimals included in percentage to '3'.

Figure 6: Changing Tooltip Percentage Reporting Options

To remove the default percentage reporting, unselect the default Include percentage of total in tooltip.

Figure 7: Removing Percentage Reporting from Tooltip

For more information, see *Customizing tooltip*.

Shelves for pies

Overview of shelves for Cloudera Data Visualization Pie visuals.

A Pie visual has the following shelves:

Dimensions

Add a Dimension field from the DATA menu to create Pie chart segments for each value in that dimension.

This shelf accepts a single field.

Measures

Add a Measure field from the DATA menu to aggregate this measurement for each segment of the Pie chart.

This shelf displays only one field on the visual. Additional fields appear only in the Tooltip.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Radial

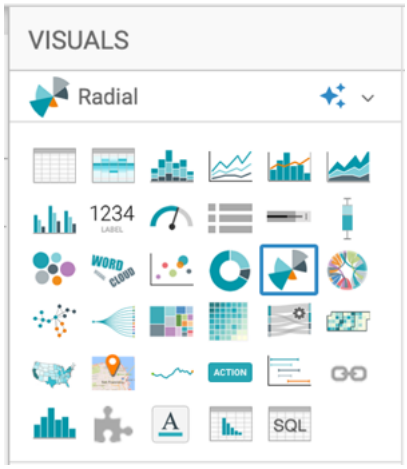
In Cloudera Data Visualization, Radial charts are a good approach for comparing the relative values of several measures. These measures are assigned an equal angle of rotation, and the relative values are represented through the areas of each wedge.

About this task

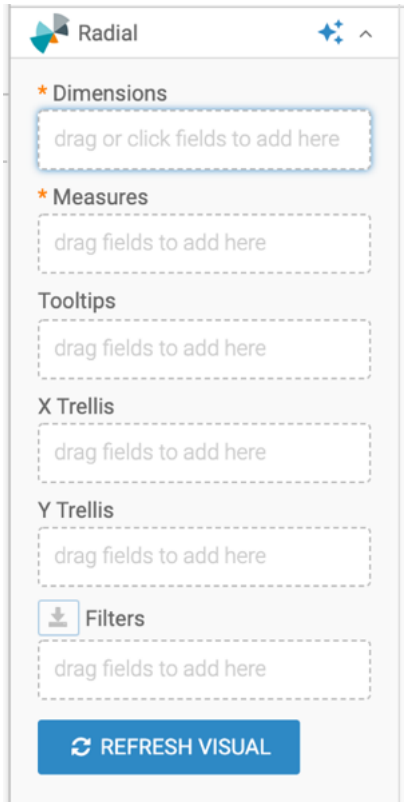
The following steps demonstrate how to create a new Radial visual representation on the US County Population dataset, for which the data source is samples.us_counties]. For an overview of shelves that specify this visual, see *Shelves for radials*.

Procedure

1. Start a new visual based on the US County Population dataset (data source samples.us_counties).
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Radial.



The shelves of the visual changed. They are now Dimensions, Measures, Tooltips, X Trellis, Y Trellis, and Filters. Both Dimensions and Measures are mandatory.

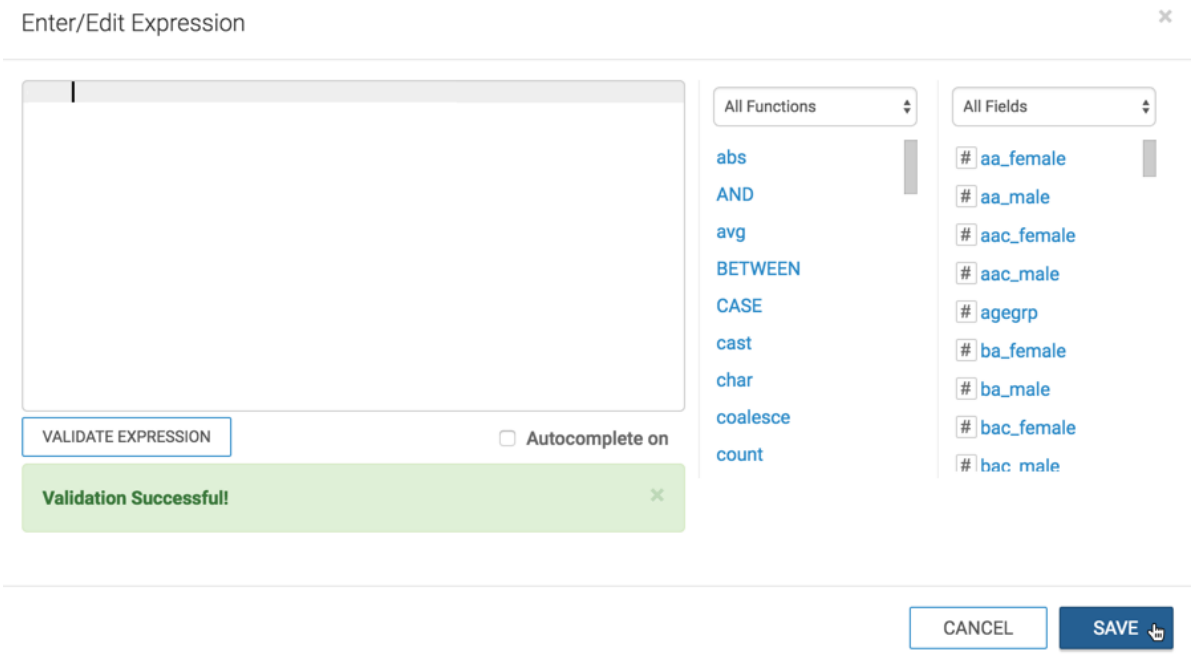


3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Dimensions, select ctyname, and add it to the Dimensions shelf.
   b) Under Measures, select the following and add them all to the Measures shelf: wa_male, ba_male, ia_male, aa_male, na_male and tom_male.

   Now all fields on the Measures shelf show the sum of male populations.
4. To show both male and female (total) in the group, change the calculations for the fields: change the expression sum([wa_male] to (sum([wa_male])+sum([wa_female])), and so on.

   The following steps demonstrate how to use the Expression Editor:

   a) On the shelf of the visual, click the (caret-right) icon to the right of the field.
   b) In the FIELD PROPERTIES menu, click [ ] Enter/Edit Expression.
   c) In the Enter-Edit Expression modal window, edit the text manually or pull in the appropriate functions and fields.
   d) Click VALIDATE EXPRESSION to ensure that the computation works, and then click SAVE.



e) Click the x icon at the top of the FIELD PROPERTIES menu to close it.

The column in the shelf now appears to hold the expression.

5. To improve the appearance of the visual, alias ctyname as County on the Dimensions shelf, and all the fields on the Measures shelf according to the following:

   - (sum([wa_male])+sum([wa_female])) as White
   - (sum([ba_male])+sum([ba_female])) as Black or African American
   - (sum([ia_male])+sum([ia_female])) as American Indian and Alaska Native
   - (sum([aa_male])+sum([aa_female])) as Asian
   - (sum([na_male])+sum([na_female])) as Native Hawaiian and Other Pacific Islander
   - (sum([tom_male])+sum([tom_female])) as Two or More Races

   Follow the steps to create an alias:

   a) On the shelf of the visual, click the caret-right icon to the right of the field.
   b) In the FIELD PROPERTIES menu, click the caret-right icon next to Alias.
   c) In the text box below Alias, enter the alias name of column, as it should appear in the visual.
   d) Click the x icon at the top of the FIELD PROPERTIES menu to close it.

   The shelf now shows the column with its alias name.

6. Click REFRESH VISUAL.
7. Note all the information provided by the visual, for each county. Also, note how long the visual is, as it lists every county in the nation.

8. Filter on the stname field to see a smaller visual, or trellis the visual on the stname field.

   For instructions, see *Trellised radial*.

9. Adjust the axes depending on what your visual demonstrates: comparison of measures within same dimension, relative values of measure compared to same measure across all dimensions, or relative values of measures compared to all measures across all dimensions.

   For instructions, see *Changing the axes scale for radial graphs*.

10. You can replace the default 'wedge' representation of the radial chart with complete circles by enabling the overlap option on the Marks menu.

    For instructions, see *Changing display to overlap*.

11. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

    In this example, the title is changed to 'Populations, by County - Radial'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

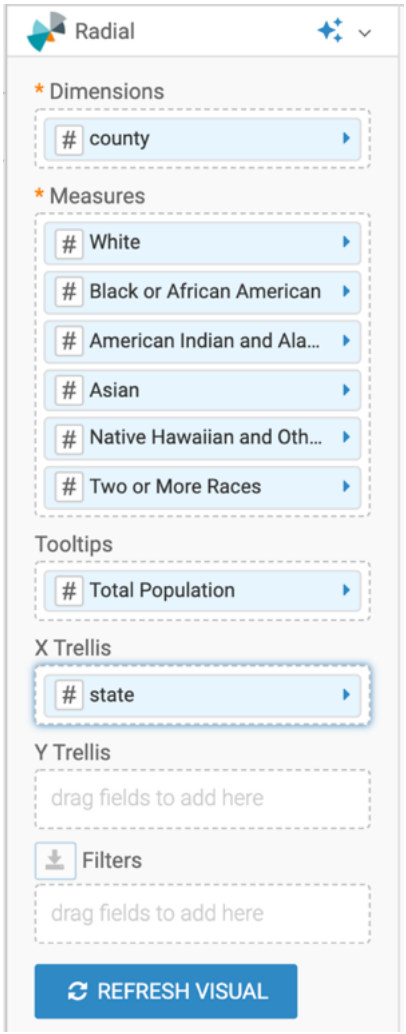
12. At the top left corner of the Dashboard Designer, click SAVE.

Trellised radial

In Radial visuals, Cloudera Data Visualization enables you to add trellises for comparing and contrasting data.

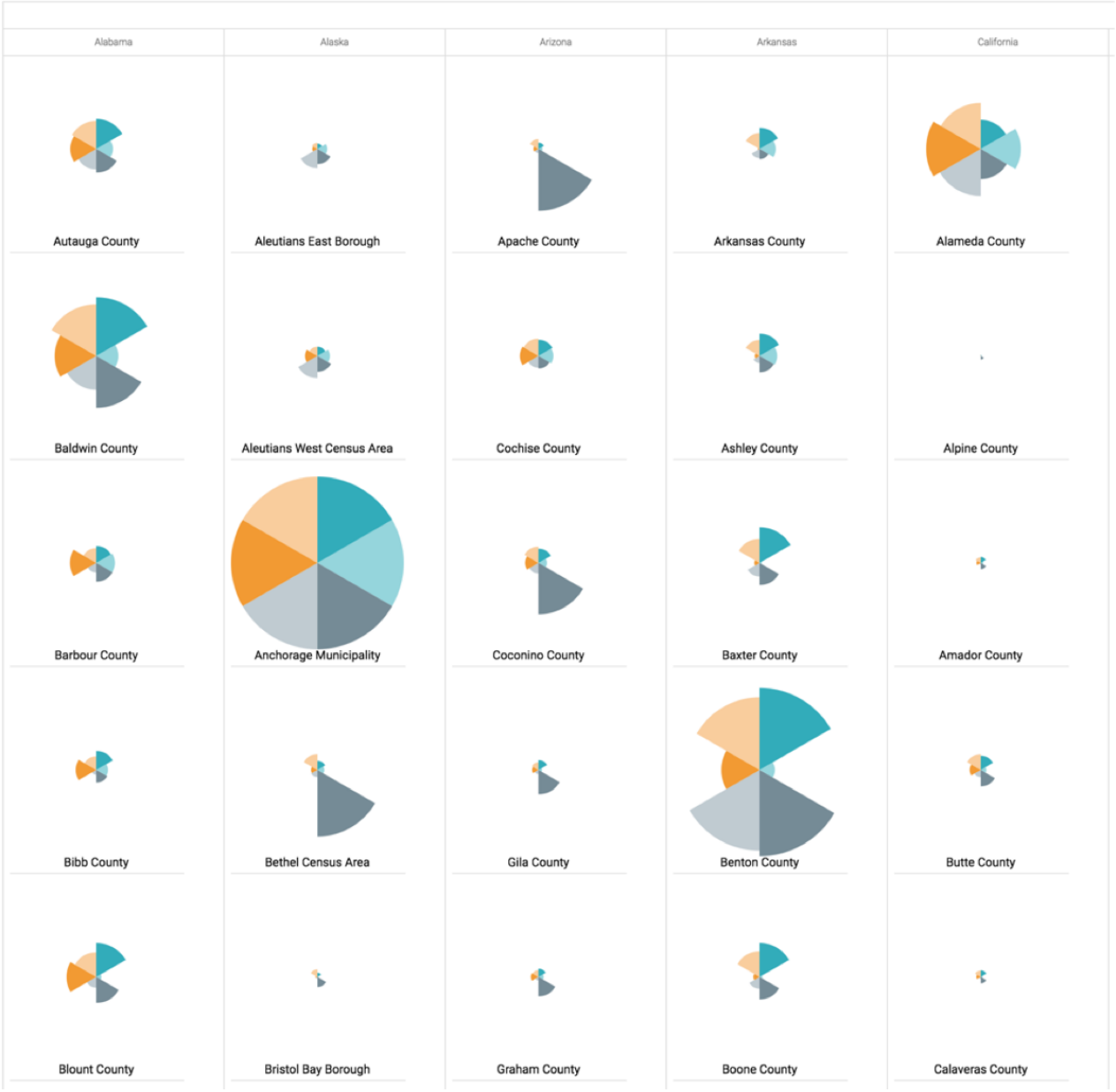
In the following example in the US County Population dataset, the field stname (aliased as State) is used as a trellising partition. Placed on the X Trellis shelf , it creates a trellised version of a radial visual you developed in *Radials*.

This is the shelves set-up for the trellised visual:



The trellising organizes the individual radial charts into columns, by State. Within each column, the charts appear by alphabetical name of the County.

Figure 8: Trellised Radial Visual



Shelves for radials

Overview of shelves for Cloudera Data Visualization Radial visuals.

A Radial visual has the following shelves:

Dimensions

Add a Dimensions field from the DATA menu to create radials for each value in that dimension.

This shelf accepts a single field.

This is a mandatory shelf.

Measures

Add a Measures field from the DATA menu to aggregate this measurement for each radial of the visual.

This shelf displays only one field on the visual. Additional fields appear only in the Tooltip.

This is a mandatory shelf.

Tooltips

Add Measures fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimensions field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimensions field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimensions or Measures fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Chord

Cloudera Data Visualization enables you to create Chord visuals. These visuals expose the inter-relationships between entities, making them an ideal choice for comparing the similarities within a dataset, or between different groups of data. The nodes/dimensions are on the circumference of a circle, and the relationships between them are the arcs inside the circle. The size of the node is proportional to the measurement. Cloudera Data Visualization uses color to help distinguish the overlapping relationship arcs.

About this task

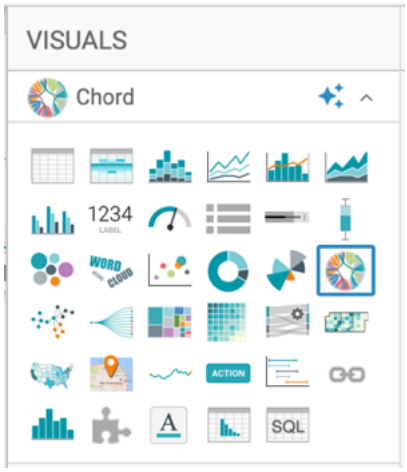
Chord visuals are similar to Correlation Heatmaps because they both compare exactly two dimensions. Chords, however, plot dimensions that cover the same value domain, and are better at pointing out the presence of a relationship between nodes than the relative magnitude of that relationship.

The following steps demonstrate how to create a new chord visual on the SF Cabrides dataset. This dataset is based on data previously imported into Cloudera Data Visualization from a datafile. Each distinct line segment between two dimensions reports the correlation data. Because the values for pick-up and drop-off areas, srcarea and tgtarea , are in the same domain, it makes the dataset ideal for demonstrating chord visuals.

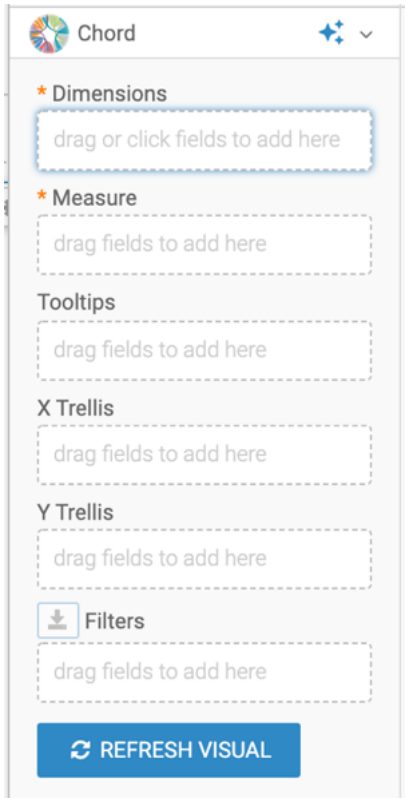
For an overview of shelves that specify this visual, see *Shelves for chord visuals*.

Procedure

1. Start a new visual based on the SF Cabrides dataset.
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Chord.



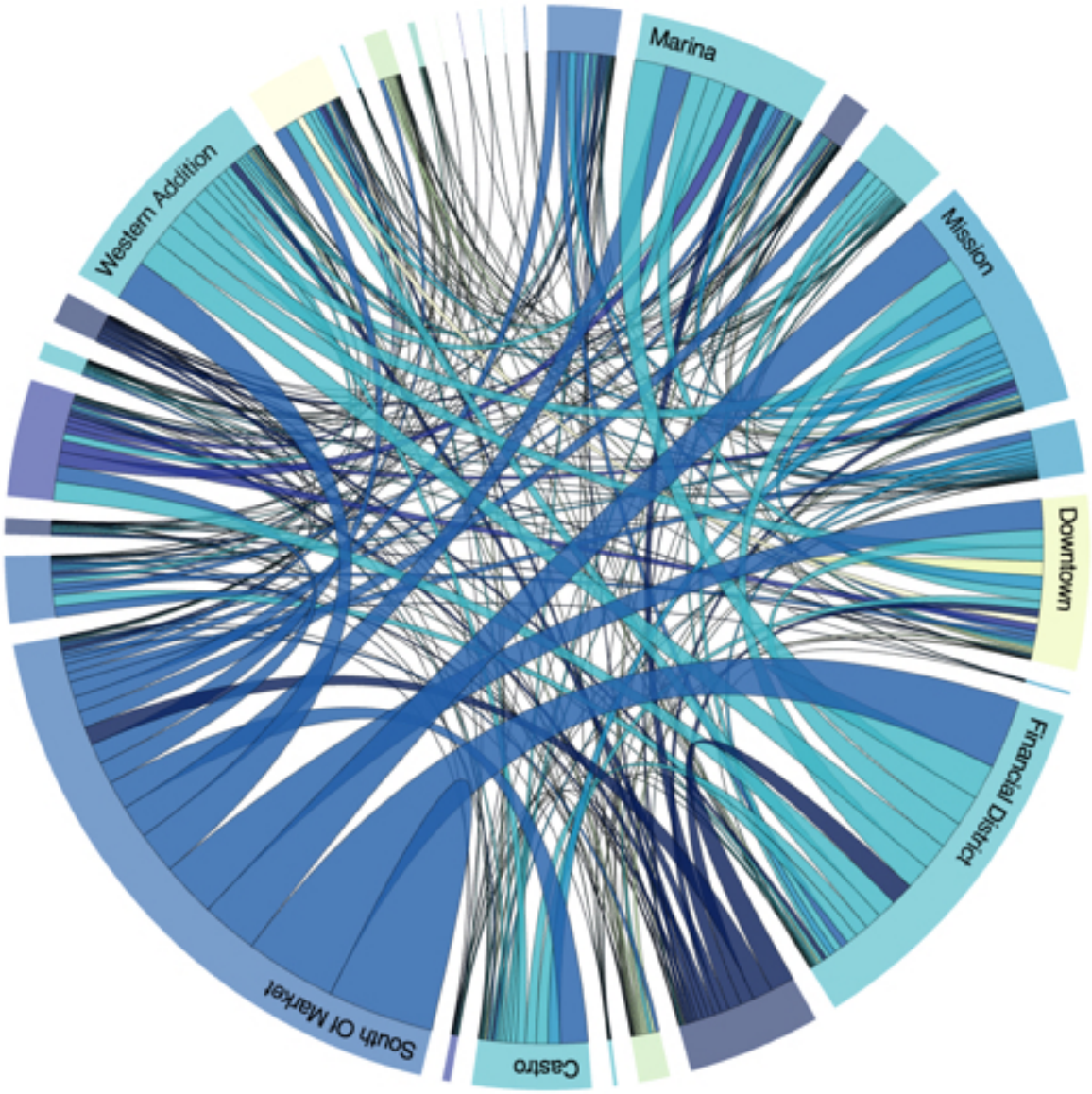
The shelves of the visual changed. They are now Dimensions, Measure, Tooltips, X Trellis, Y Trellis and Filters. Both Dimensions and Measure are mandatory.



3. Populate the shelves from the fields available in the DATA menu:
   a) From Dimensions, select srcarea and trgarea, and drag them to the Dimensions shelf.
   b) From Measures, select rides and place it to the Measure shelf.
4. Click REFRESH VISUAL.

   The Chord visual appears.

Figure 9: Chord Visual

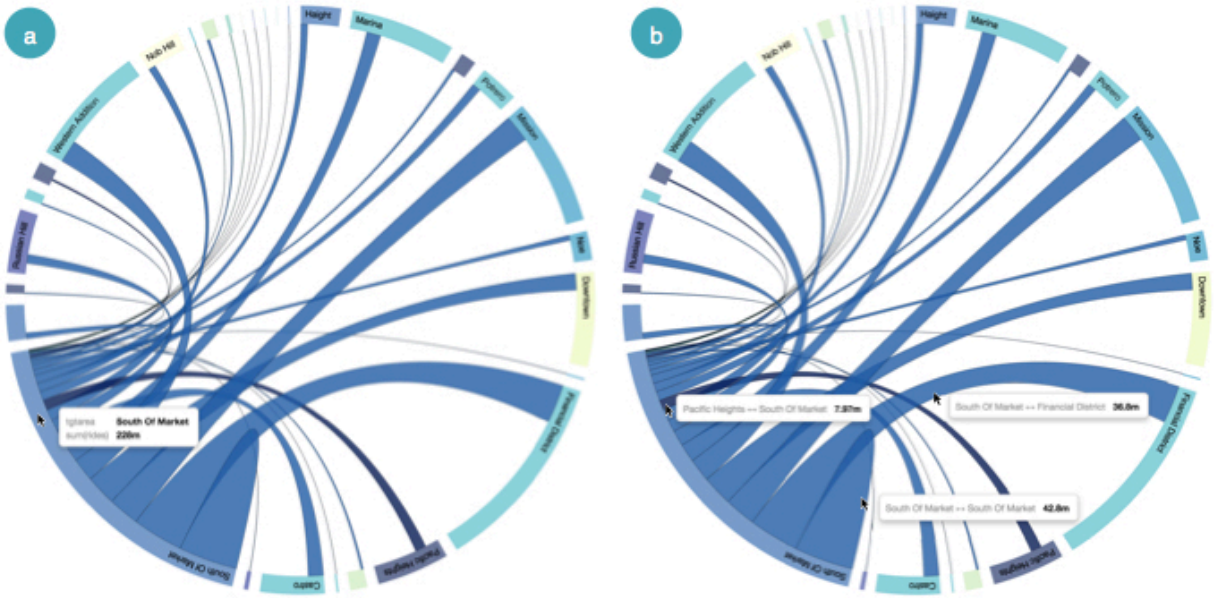


5. Check the details of the individual dimensions:
   a) Hover the pointer over one of the nodes.

      The visual shows the connections that originate or terminate with this specific value, as well as the tooltip. In the following image, you can see all connections for 'South of Market'.
   b) Hover the pointer over one of the connections.

      The tooltip shows the detail for that connection: starting and ending notes, and the measurement.

Figure 10: Viewing Details of a Chord Visual



6. Click EDIT next to the title of the visualization, and change the title to 'San Francisco Cab Rides - Chord'.
7. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for chords

Overview of shelves for Cloudera Data Visualization Chord visuals.

A Chord visual has the following shelves:

Dimensions

Add two Dimensions fields from the DATA menu to define the start and end points of each connection in the plot.

This shelf accepts exactly two fields.

This is a mandatory shelf.

Measure

Add a Measures field from the DATA menu to aggregate this measurement for each chord of the chart. The aggregate measurement determines the thickness of the chord connector between each pair of dimensions.

This shelf accepts a single field.

This is a mandatory shelf.

Tooltips

Add Measures fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimensions field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimensions or Measures fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Network

In Cloudera Data Visualization, Network visuals can be used to represent many systems and structures. The implementation is dynamic, with ongoing optimization adjustments that implement a force layout based on the Barnes–Hut approximation. It also includes many manual setting adjustments.

About this task

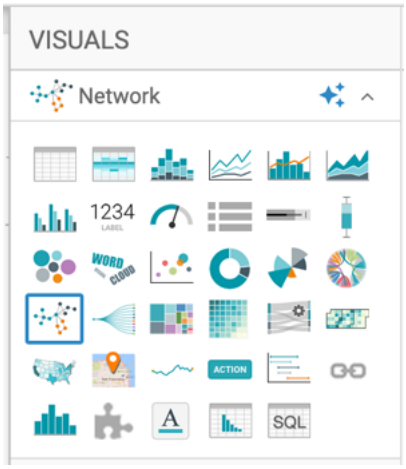
In the following example, the NCAA Basketball dataset is used for years 1985 through 2018.

For an overview of shelves that specify this visual, see *Shelves for networks*.

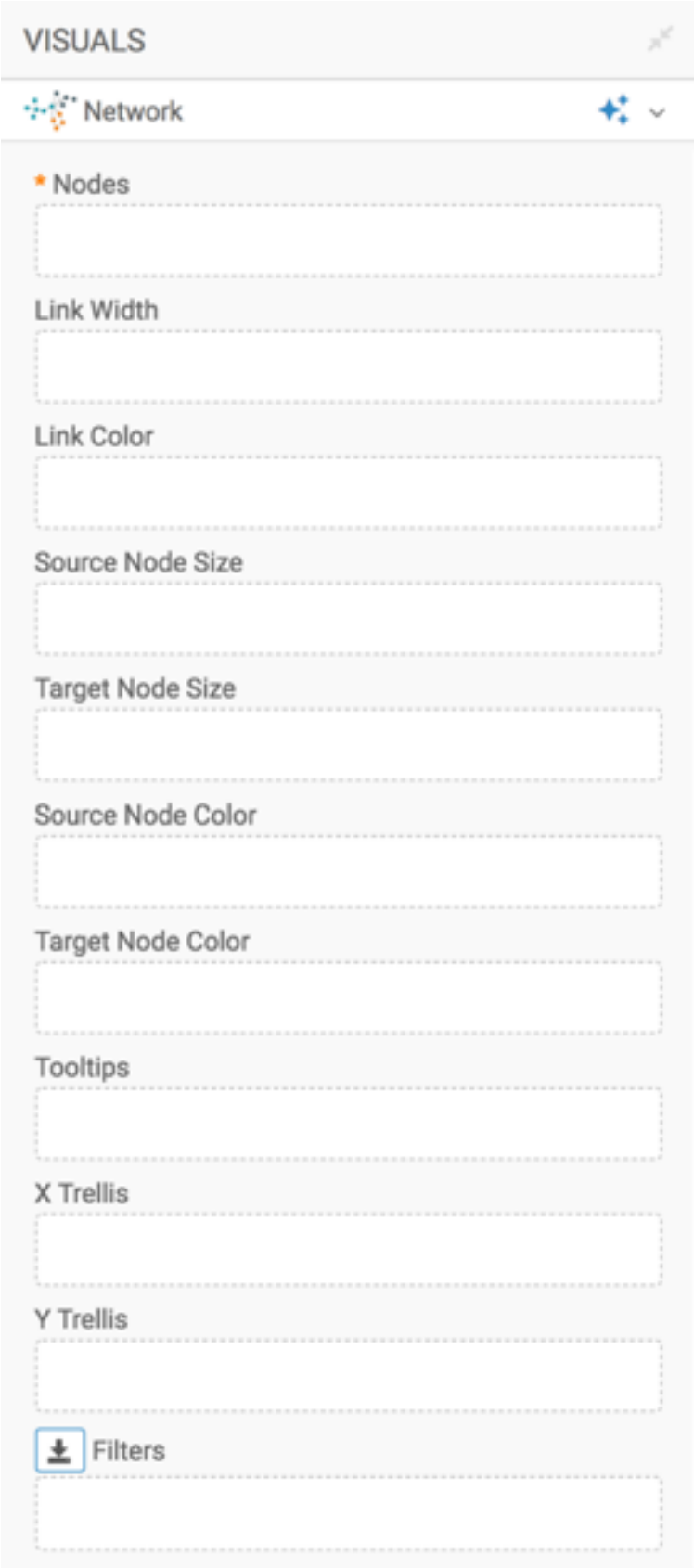
Procedure

1. Start a new visual based on the NCAA Basketball dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Network.



The shelves of the visual changed. They are now Nodes, Link Width, Link Color, Source Node Size, Target Node Size, Source Node Color, Target Node Color, Tooltips, X Trellis, Y Trellis, and Filters. There is only one mandatory shelf: Nodes, which adds two dimensions from the data menu to specify the start and end of each network connection in the plot.

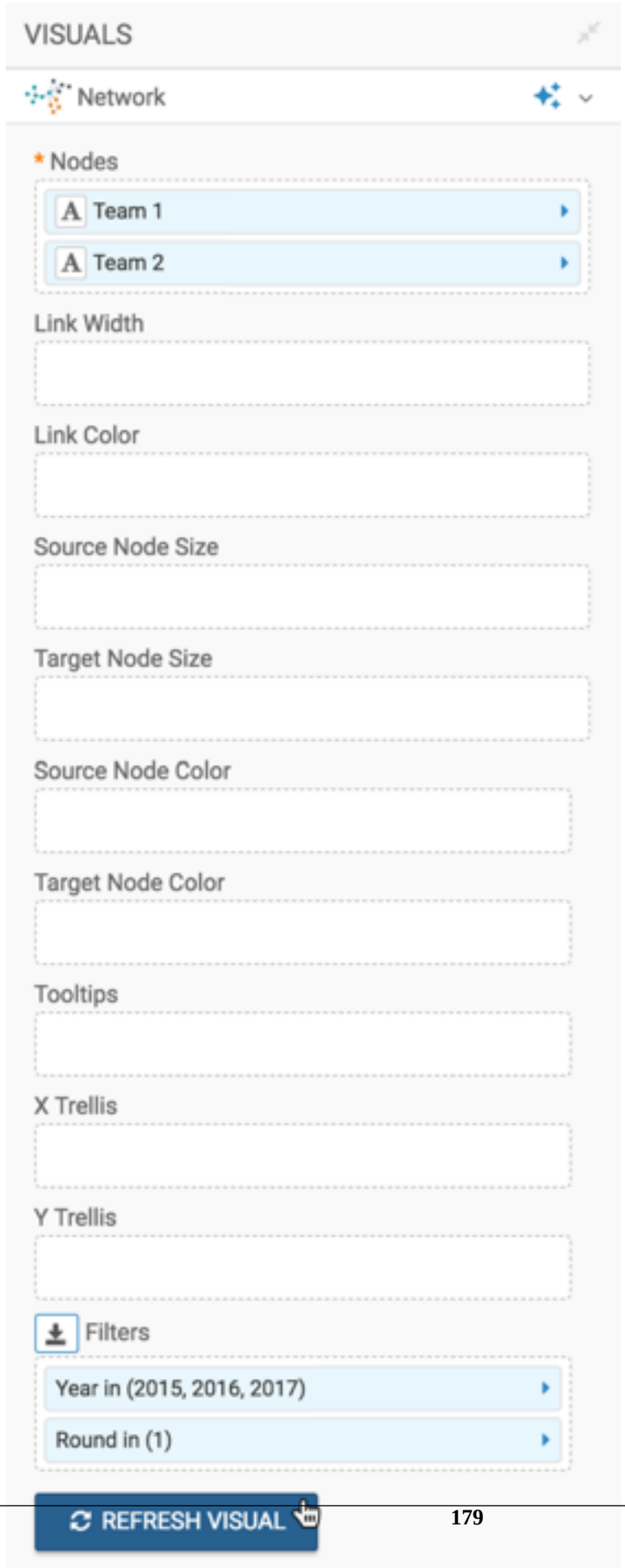


To learn more about the shelves of this visual, see *Shelves for networks*.

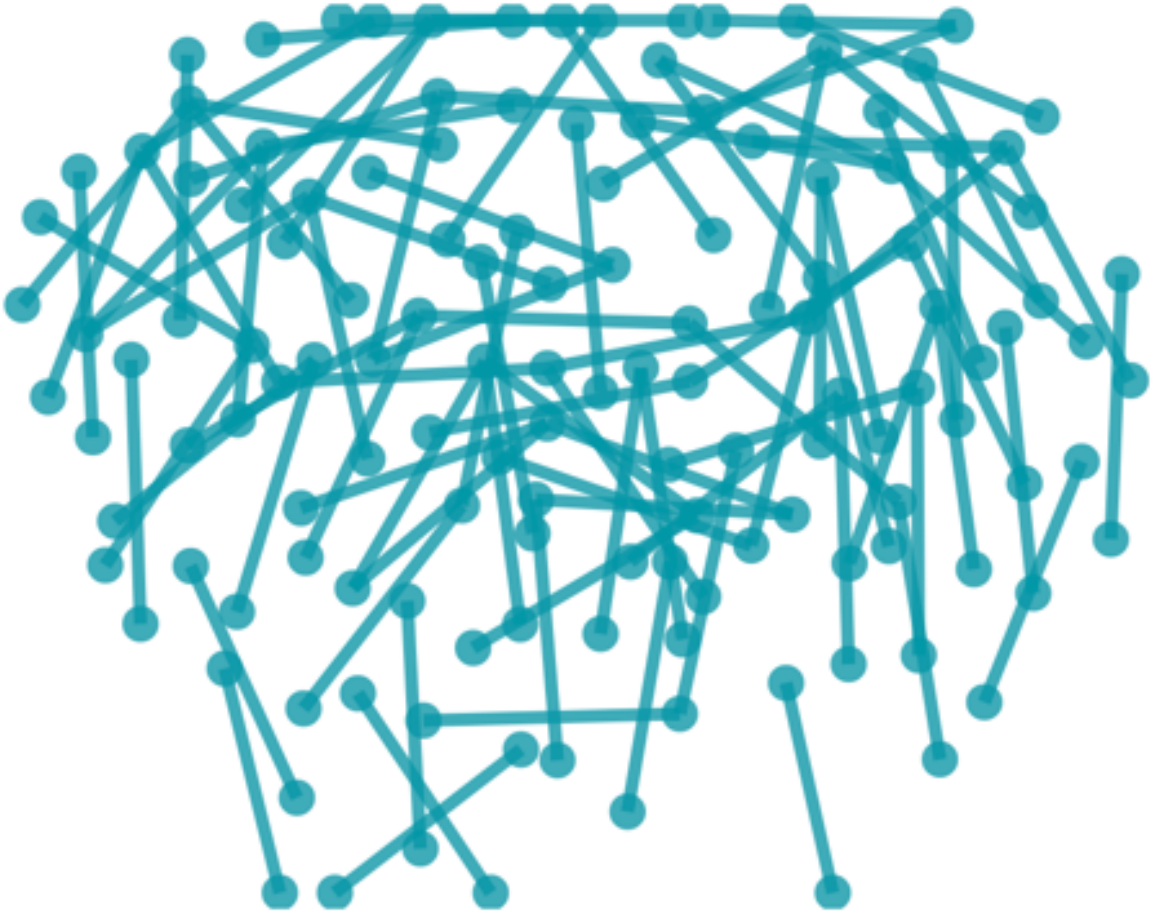
3. Populate the shelves from the fields available in the DATA menu:

a) Under Dimensions, select Team 1 and Team 2, and add them to the Nodes shelf.
b) Under Measures, select Year and add it to the Filters shelf. Select the years 2015, 2016, and 2017.
c) Add Round to the Filters shelf, and select 1.

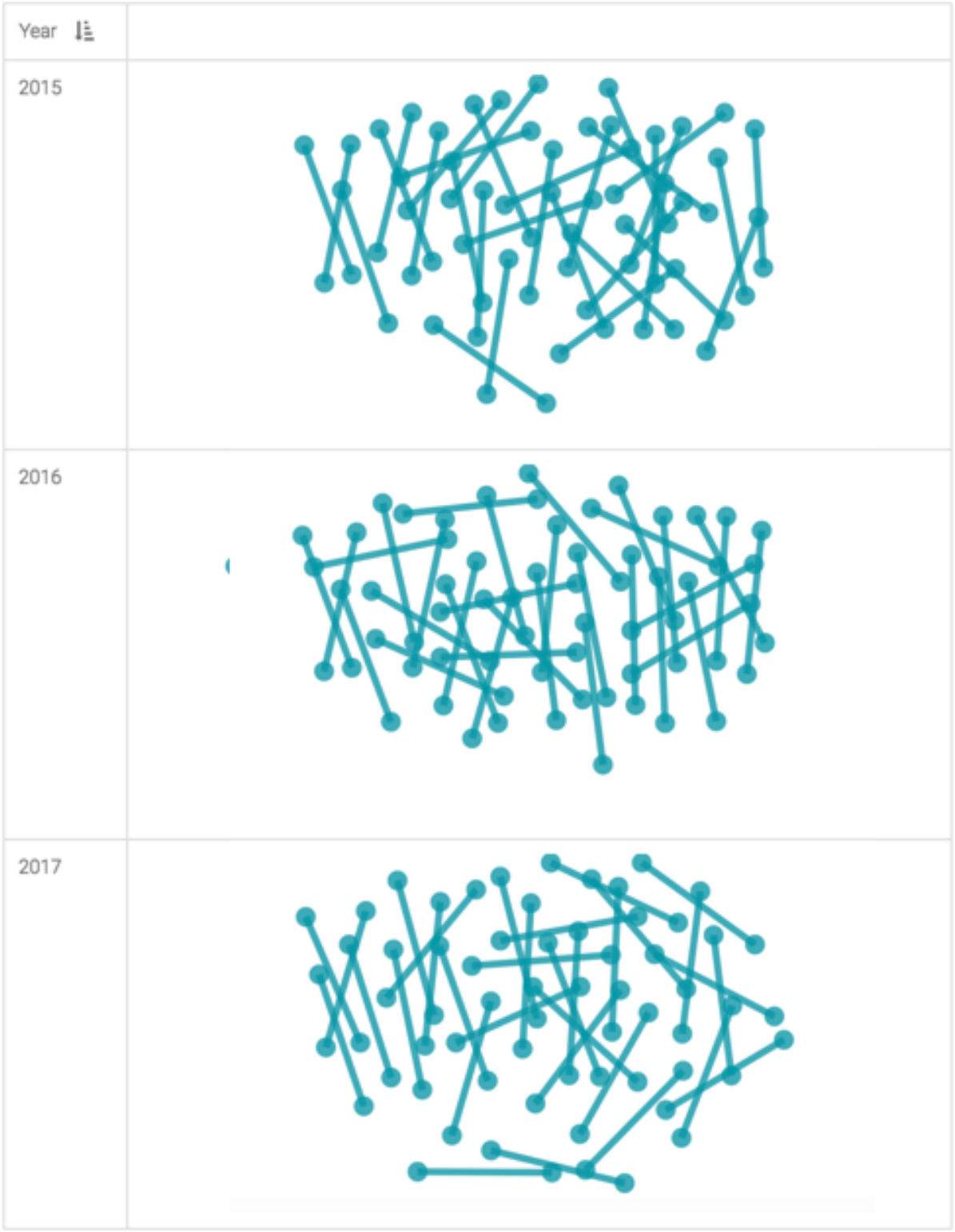
4. Click REFRESH VISUAL.


VISUALS
Network
* Nodes
Team 1
Team 2
Link Width
Link Color
Source Node Size
Target Node Size
Source Node Color
Target Node Color
Tooltips
X Trellis
Y Trellis
Filters
Year in (2015, 2016, 2017)
Round in (1)
REFRESH VISUAL

The new network visual appears. You can see that this visual shows 1:1 pairing, because round 1 of play represents 1:1 elimination.



5. To show the years separately, place the Year field on the Y Trellis shelf and click REFRESH VISUAL. The trellised version of the network visual appears.



6. Change the title to Round 1, 2015-2017.
   - Click (pencil icon) next to the title of the visualization to edit it, and enter the new name.
   - [Optional] Click (pencil icon) below the title of the visualization to add a brief description of the visual.
7. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

In this example, the title is changed to 'Round 1, 2015-2017'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

8. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

To adjust the network display, check all the available settings for this visual.

For more information, see *Network visuals: same source and target data values* and *Advanced options for network visuals*. These topics demonstrate some of the many customizations that make the network visual an extremely versatile analysis tool.

Shelves for networks

Overview of shelves for Cloudera Data Visualization Network visuals.

A Network visual has the following shelves:

Nodes

Add two Dimensions fields from the DATA menu to define the start- and end-points of each connection in the plot.

This shelf accepts exactly two fields.

This is a mandatory shelf.

Link Width

Add a Measures field from the DATA menu to aggregate this measurement for each connection of the chart. The aggregate measurement determines the thickness of the connector between each pair of dimensions.

This shelf accepts a single field.

Link Color

Add a Measures or Dimensions field from the DATA menu to color the links of the plot.

This shelf accepts a single field.

Source Node Size

Add a Measures field from the DATA menu to show an additional measurement as a relative size of the source node.

This shelf accepts a single field.

Target Node Size

Add a Measures field from the DATA menu to show an additional measurement as a relative size of the target node.

This shelf accepts a single field.

Source Node Color

Add a Measures or Dimensions field from the Data menu to color the source node of the plot.

This shelf accepts a single field.

Target Node Color

Add a Measures or Dimensions field from the DATA menu to color the target node of the plot.

This shelf accepts a single field.

Tooltips

Add Measures fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimensions field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimensions field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimensions or Measures fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Advanced options for networks

Network visuals enable you to show a large amount of supplemental information directly by using the optional shelves of the visual, and the various settings for Marks, Nodes, and Links.

Networks: same source and target data values

Network visuals can show the same node once, or several times on the same chart, depending on node settings.

About this task

In the following example the NCAA Basketball dataset is used, and the visual is based on a visual developed in *Networks*. The original visual shows only Round 1 play, or 1:1 match-ups.

Procedure

1. Clone the visual, and save it as All Rounds, 2015-2018.

   For instructions, see *Cloning a visual*.
2. Add Region Name to the X Trellis shelf.

   This shows all regions separately, and makes it easier to pick out the nodes that have multiple links.
3. Because this dataset defines the 'Final 4' and 'Championship' play as regions, add Region Name to the Filters shelf, and exclude these values.
4. To choose all rounds of play, remove the Round field from the Filters shelf.
5. Click Settings Nodes and select Source & target node data come from the same base values.

   For more information, see *Showing the same source and target values*.
6. Click REFRESH VISUAL.

   You can see the difference between the two visuals: the one showing all rounds is much denser, and shows multi-branching from some of the same nodes.

Networks: link color menu

Network visuals have many color palette options for links.

About this task

To get to the link color options, follow these steps:

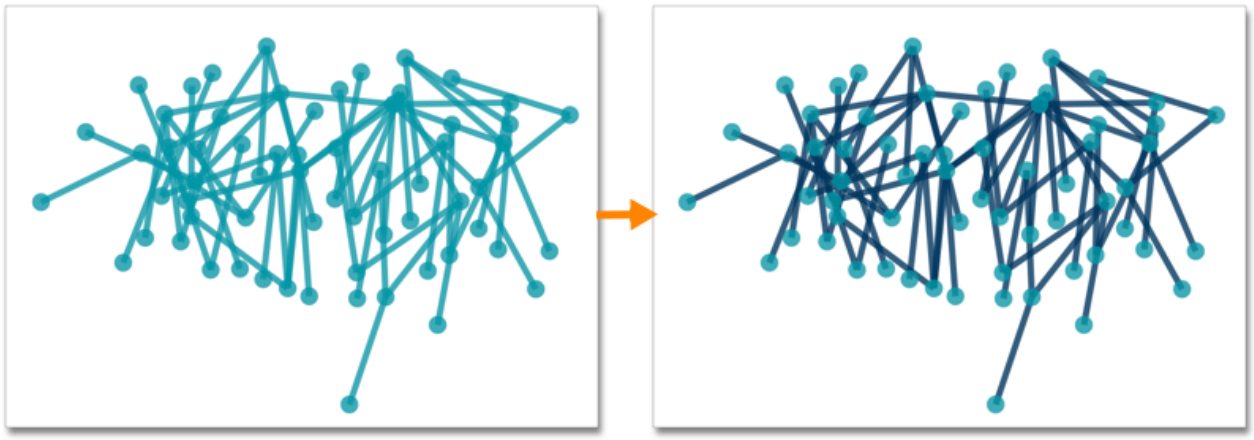
Procedure

1. On the right side of the Visual Designer, click the Style menu.
2. Open the Link Colors menu and choose a color palette.
   - You can select Reverse color palette of visuals if you want to change the order of the colors.
   - You can click on an individual color swatch in the palette to change the first color within palette.
   - You can also change the color of the Direction Arrow.

Example

In this example the links palette is changed so links are more clearly distinguished from the nodes. For this visual, all rounds and all regions of the 2017 season are shown:

Figure 11: Changing color palette for links in network visual



Networks: link color shelf

By default, the links of the network visual have the same thickness and length. Network visual has a data shelf called Link color that uses a specified palette to plot different link colors.

About this task

To specify the link color categories, follow these steps:

Procedure

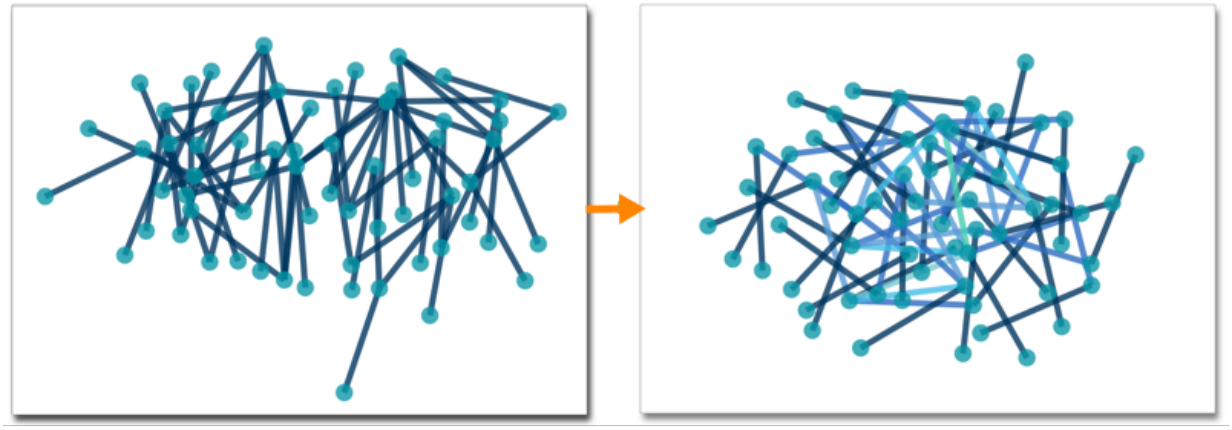
1. Choose the appropriate palette, as described in *Networks: link color menu*.
2. Add the round field to the Link Color shelf.
3. Click the round field on the shelf to see the FIELD PROPERTIES menu.
4. In the FIELD PROPERTIES menu, select [ ] Enter/Edit Expression.
5. In the expression editor, specify the max() aggregation function, with an alias:

   ```
   max([Round]) as 'Round'
   ```

6. Click VALIDATE & SAVE.
7. Click REFRESH VISUAL.

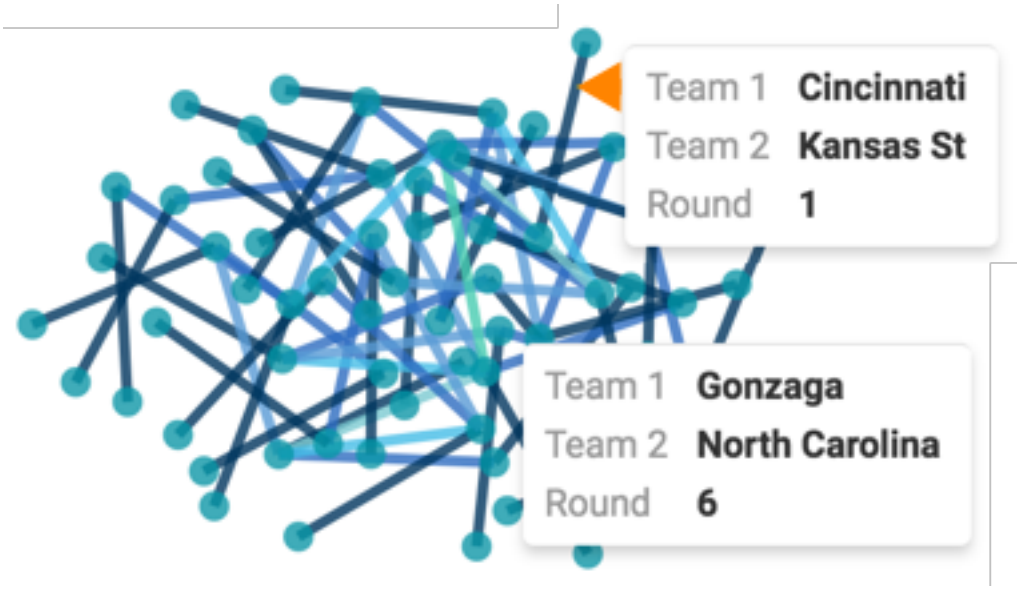
   You can see the change to the appearance of the visual:

   Figure 12: Specifying the link color shelf in a network visual

8. If you hover over a specific link, you can see the detail information. For example, the dark-colored links represent Round 1 games, and the lightest link in the middle is the final game, Round 6.

   Figure 13: Tooltips on a network visual with active link color shelf



Networks: link width shelf

By default, the links of the network visual have the same thickness and length. Network visual has a data shelf called Link width that produces a chart with link thickness that depends on specified values.

About this task

To specify the width of the link, follow these steps:

Procedure

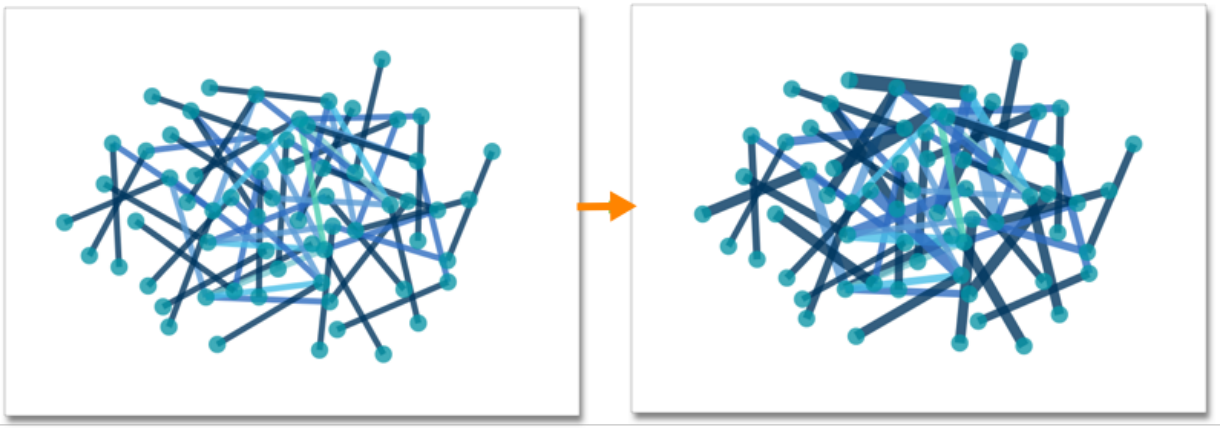
1. Add the Score 1 field to the Link Width shelf.
2. Click the Score 1 field on the shelf, to see the FIELD PROPERTIES menu.
3. In the FIELD PROPERTIES menu, select [ ] Enter/Edit Expression.
4. In the expression editor, specify a new calculation as a max() aggregation function, with an alias:

   ```
   max(abs([Score 1] - [Score 2])) as 'Score Differential'
   ```

5. Click VALIDATE & SAVE.
6. Click REFRESH VISUAL.

   You can see the change to the appearance of the visual:

   Figure 14: Specifying the link width shelf in a network visual

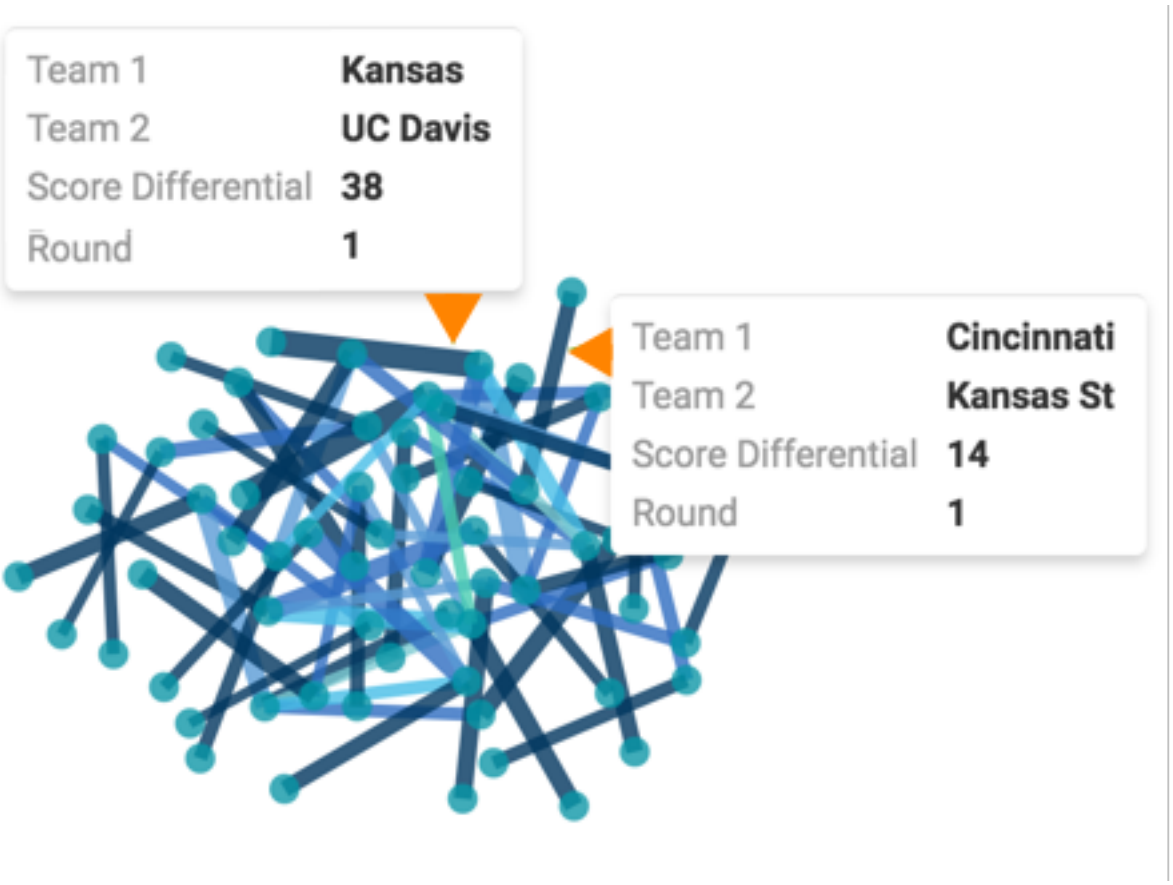


7. If you hover over a specific link, you can see the detail information.

   For example, compare the following two Round 1 games:

   - Kansas vs. UC Davis that has a 38 point score differential, and the thickest link in the chart
   - Cincinnati vs. Kansas State that has a 14 point score differential, and a comparatively thinner link

   Figure 15: Tooltips on a network visual with active link color shelf



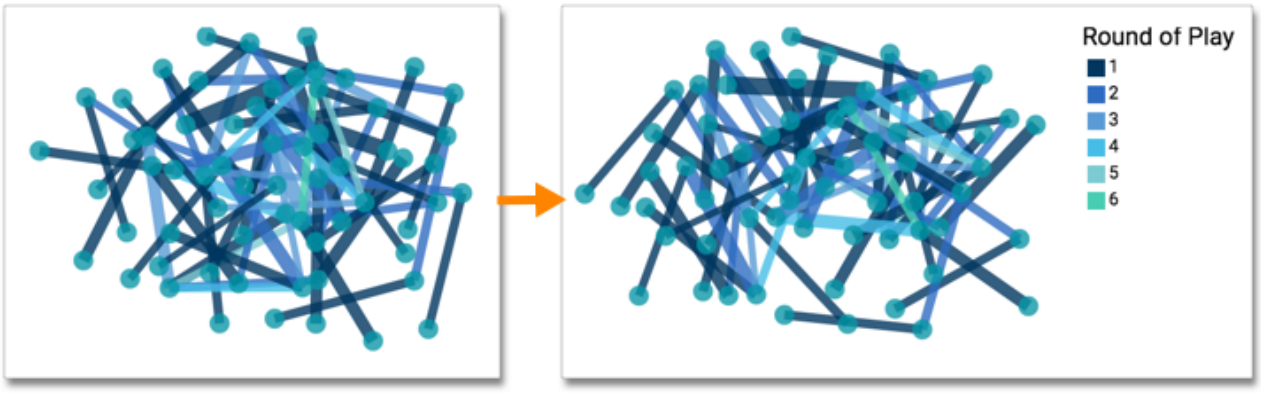
Networks: link legend

In network visuals, the legend for link colors helps you to visually identify the correct values.

To enable the link legend in network visuals, see the instructions in *Enabling the link legend*.

In the following example, a link legend is added to the visual developed in *Networks: link width shelf*.

Figure 16: Adding a legend for links in a network visual



Networks: node size shelves

Network visuals enable you to change the size of the nodes.

About this task

By default, the nodes of a network visual have the same size. This visual has two data shelves, Source node size and Target node size, that produce a chart with nodes that represent values in the dataset.

To specify the size of the nodes, follow these steps:

Procedure

1. Add the Seed 1 field to the Source Node Size shelf.
2. Click the Score 1 field on the shelf, to see the FIELD PROPERTIES menu.
3. In the FIELD PROPERTIES menu, select [ ] Enter/Edit Expression.
4. In the expression editor, specify a new calculation as a max() aggregation function, with an alias:

   ```
   max([Seed 1]) as 'Seed'
   ```

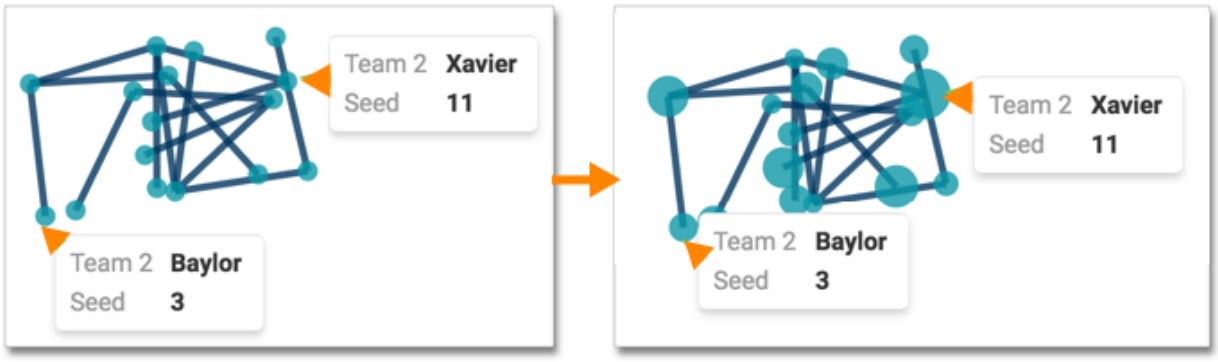
5. Click VALIDATE & SAVE.
6. Add the Seed 2 field to the Target Node Size shelf.
7. Click the Score 2 field on the shelf, to see the FIELD PROPERTIES menu.
8. In the FIELD PROPERTIES menu, select [ ] Enter/Edit Expression.
9. In the expression editor, specify a new calculation as a max() aggregation function, with an alias:

   ```
   max([Seed 2]) as 'Seed'
   ```

10. Click VALIDATE & SAVE.
11. Click REFRESH VISUAL.

    You can see the change to the appearance of the visual: the tooltips for the visual to point out the size difference between Baylor (seed 3) and Xavier (seed 11):

    Figure 17: Specifying node size on a network visual



Networks: nodes color menu

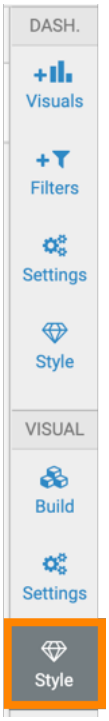
Cloudera Data Visualization provides many color palette options for nodes in a Network visual.

About this task

To get to the options in the Colors menu, follow these steps:

Procedure

1. On the right side of the Visual Designer, click the Style menu.



2. Open the Node Colors menu and choose a color palette. select one of the following options:
   - You can select Reverse color palette of visuals if you want to change the order of the colors.
   - You can click on an individual color swatch in the palette to change the first color within palette.
   - You can select Fix colors for specified dimension values.
   - You can add dimension values.

Networks: node color shelves

By default, the nodes of a network visual have the same color. This visual has two data shelves, Source node color and Target node color, that produce a where all nodes for a category have the same color value.

About this task

To specify the color of the nodes, follow these steps:

Procedure

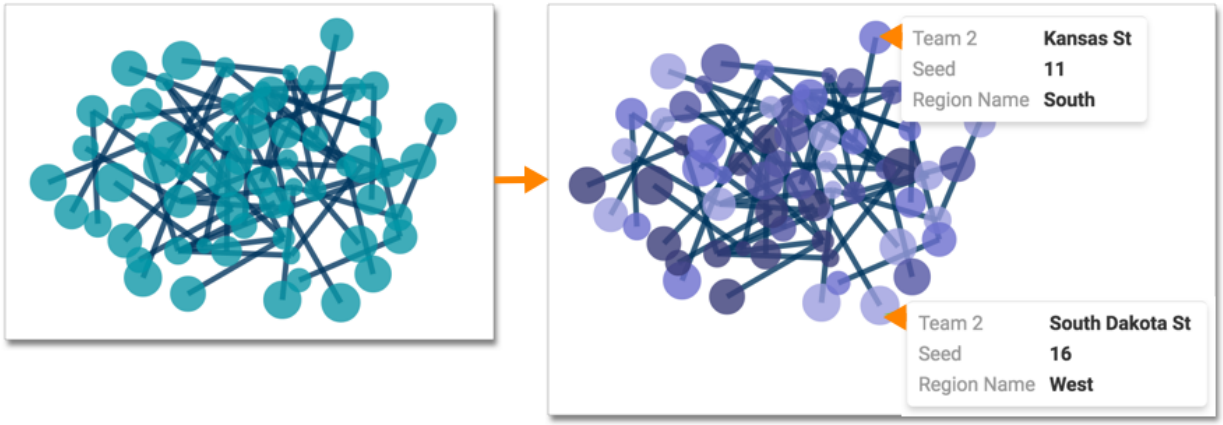
1. Add the Region Name field to the Source Node Color shelf.
2. Add the Region Name field to the Target Node Color shelf.
3. For both shelves, ensure that the aggregation function is max().
4. Alias the fields on the shelves.

   In this example, the fields are named as Region Name to hide the necessary aggregation information.
5. Remove the round field from the Filter shelf.

   In this example, only one filter (Year in (2017)) is left in place.
6. Click VALIDATE & SAVE.
7. Click REFRESH VISUAL.

   You can see how much easier it is to distinguish between the nodes that represent different regions. The following example shows the tooltips for a South and West region:

Figure 18: Specifying node color on a network visual



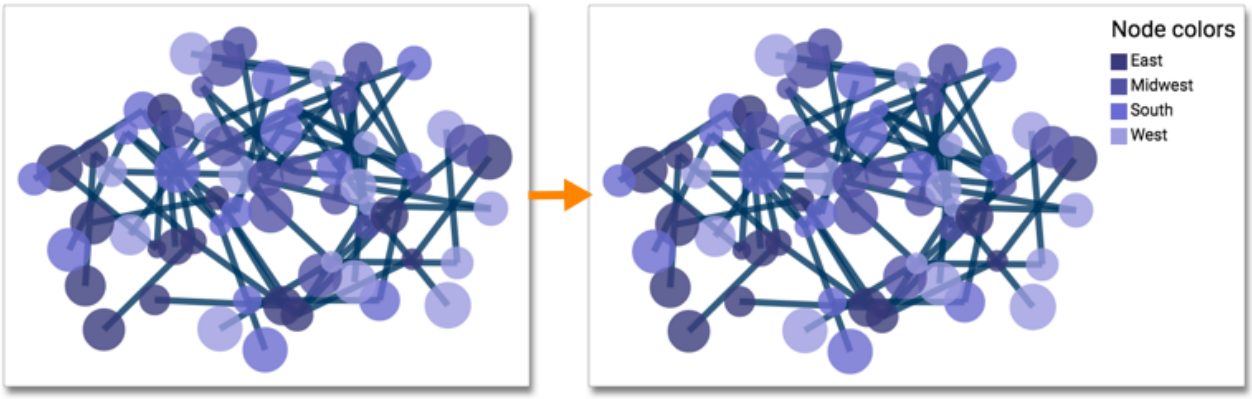
Networks: node legend

In network visuals, the legend for node colors helps you to visually identify the correct values.

To enable the node legend in network visuals, see the instructions in *Enabling the nodes legend*.

In this example, the node legend is added to the visual developed in *Networks: node color shelves*.

Figure 19: Displaying a Legend for Nodes in Network Visual



Dendrograms

Cloudera Data Visualization enables you to create Dendrogram visuals. Dendrograms are good at expressing hierarchical relationships in data. As the name derivation suggests (Greek 'dendro' for 'tree' and 'gramma' for 'drawing'), dendrograms are tree diagrams.

About this task

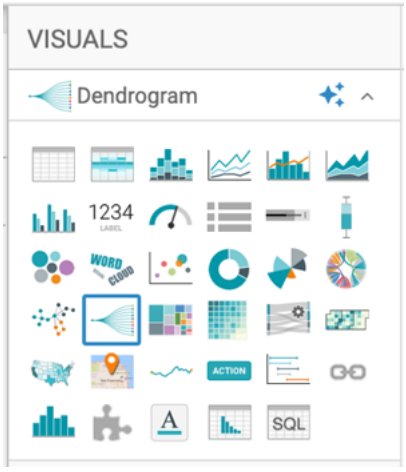
In the following example the NYC Taxi Data for year 2013 dataset is used. You can download a subset of this data for trips made within *Staten Island boro*, upload it, and create the dataset NYC Taxi Data.

For an overview of shelves that specify this visual, see *Shelves for dendogram visuals*.

Procedure

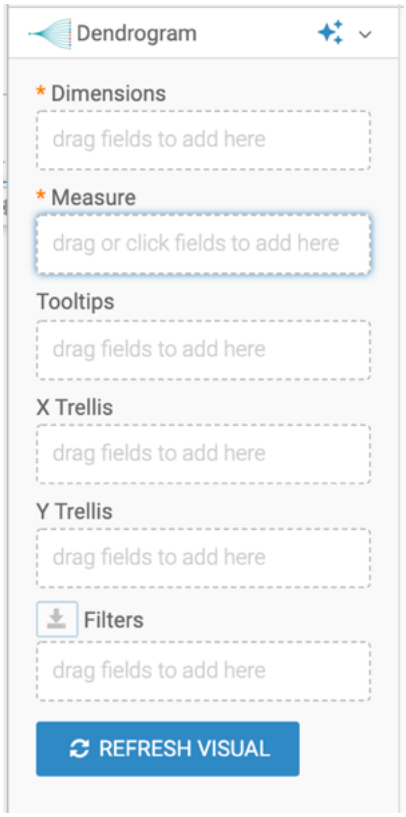
1. Start a new visual based on dataset NYC Taxi Data.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Dendrogram.

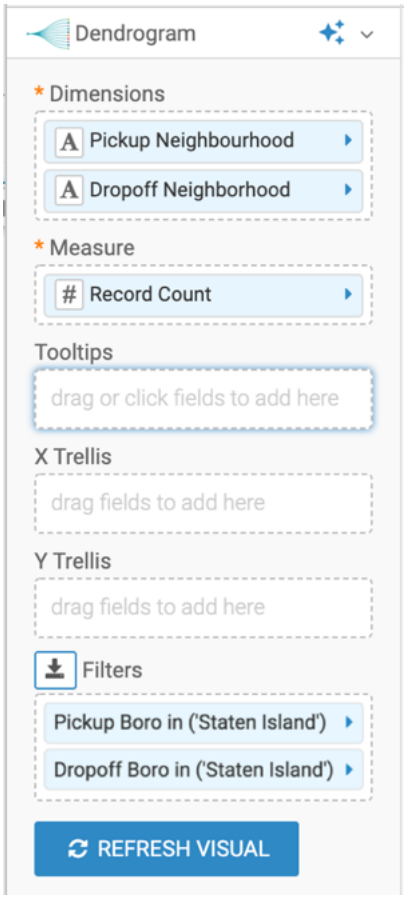


The shelves of the visual changed. They are now Dimensions, Measure, Tooltips, X Trellis, Y Trellis and Filters. The mandatory shelves are:

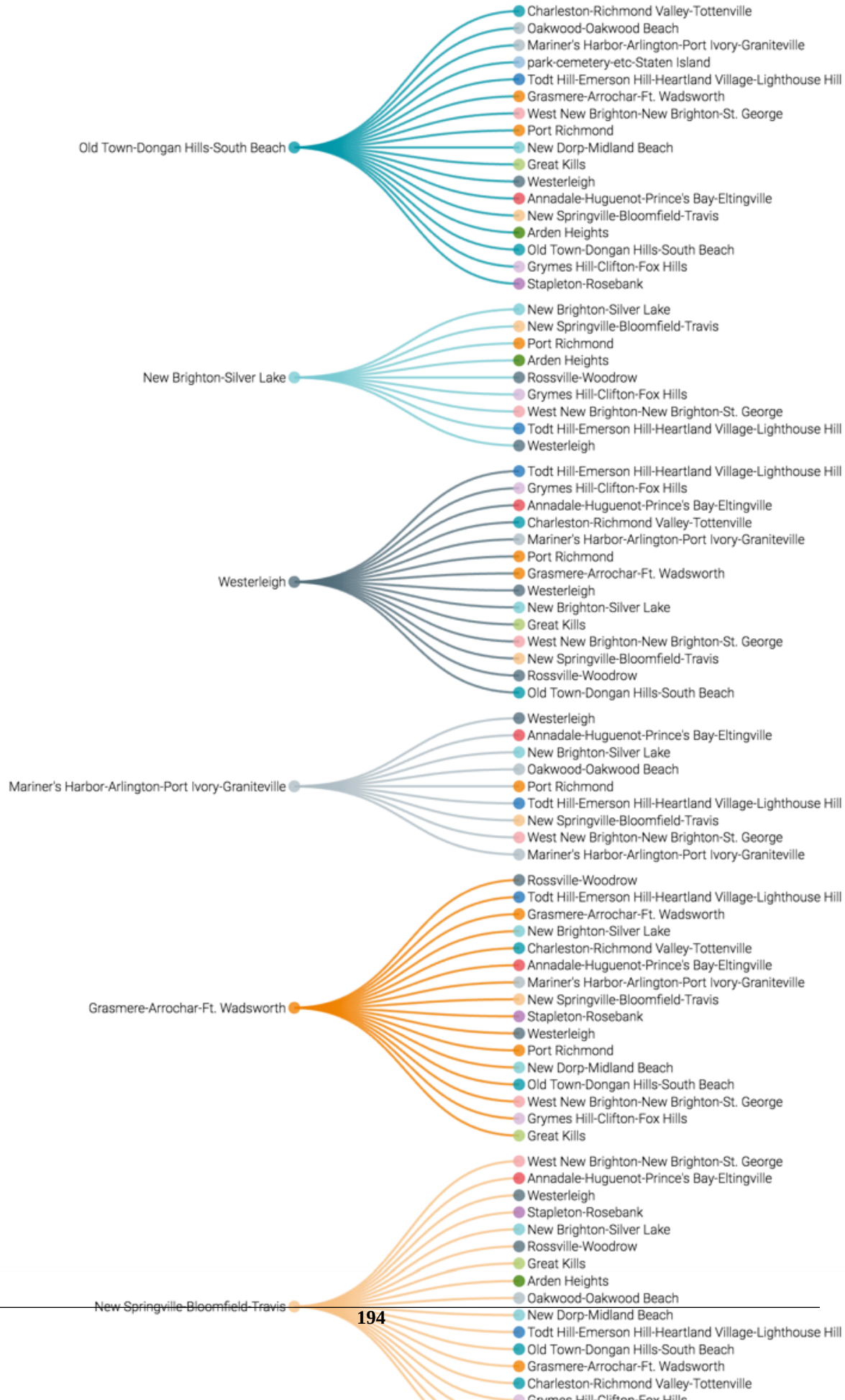
- Dimensions: specify two dimensions, because there are at least two nodes on each branch
- Measure; specify one measurement plotted on the branch between nodes



3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Dimensions, select Pickup Neighborhood and Dropoff Neighborhood, and add them to the Dimensions shelf.
   b) Under Measures, select Record Count and add it to the Measure shelf.
4. To examine only the rides within Staten Island, add Pickup Boro and Dropoff Boro to the Filters shelf, and set them to a single value 'Staten Island'.



5. Click REFRESH VISUAL.
   The new dendrogram visual appears.


Old Town-Dongan Hills-South Beach
Charleston-Richmond Valley-Tottenville
Oakwood-Oakwood Beach
Mariner's Harbor-Arlington-Port Ivory-Graniteville
park-cemetery-etc-Staten Island
Todt Hill-Emerson Hill-Heartland Village-Lighthouse Hill
Grasmere-Arrochar-Ft. Wadsworth
West New Brighton-New Brighton-St. George
Port Richmond
New Dorp-Midland Beach
Great Kills
Westerleigh
Annadale-Huguenot-Prince's Bay-Eltingville
New Springville-Bloomfield-Travis
Arden Heights
Old Town-Dongan Hills-South Beach
Grymes Hill-Clifton-Fox Hills
Stapleton-Rosebank
New Brighton-Silver Lake
New Brighton-Silver Lake
New Springville-Bloomfield-Travis
Port Richmond
Arden Heights
Rossville-Woodrow
Grymes Hill-Clifton-Fox Hills
West New Brighton-New Brighton-St. George
Todt Hill-Emerson Hill-Heartland Village-Lighthouse Hill
Westerleigh
Westerleigh
Todt Hill-Emerson Hill-Heartland Village-Lighthouse Hill
Grymes Hill-Clifton-Fox Hills
Annadale-Huguenot-Prince's Bay-Eltingville
Charleston-Richmond Valley-Tottenville
Mariner's Harbor-Arlington-Port Ivory-Graniteville
Port Richmond
Grasmere-Arrochar-Ft. Wadsworth
Westerleigh
New Brighton-Silver Lake
Great Kills
West New Brighton-New Brighton-St. George
New Springville-Bloomfield-Travis
Rossville-Woodrow
Old Town-Dongan Hills-South Beach
Mariner's Harbor-Arlington-Port Ivory-Graniteville
Westerleigh
Annadale-Huguenot-Prince's Bay-Eltingville
New Brighton-Silver Lake
Oakwood-Oakwood Beach
Port Richmond
Todt Hill-Emerson Hill-Heartland Village-Lighthouse Hill
New Springville-Bloomfield-Travis
West New Brighton-New Brighton-St. George
Mariner's Harbor-Arlington-Port Ivory-Graniteville
Grasmere-Arrochar-Ft. Wadsworth
Rossville-Woodrow
Todt Hill-Emerson Hill-Heartland Village-Lighthouse Hill
Grasmere-Arrochar-Ft. Wadsworth
New Brighton-Silver Lake
Charleston-Richmond Valley-Tottenville
Annadale-Huguenot-Prince's Bay-Eltingville
Mariner's Harbor-Arlington-Port Ivory-Graniteville
New Springville-Bloomfield-Travis
Stapleton-Rosebank
Westerleigh
Port Richmond
New Dorp-Midland Beach
Old Town-Dongan Hills-South Beach
West New Brighton-New Brighton-St. George
Grymes Hill-Clifton-Fox Hills
Great Kills
New Springville-Bloomfield-Travis
West New Brighton-New Brighton-St. George
Annadale-Huguenot-Prince's Bay-Eltingville
Westerleigh
Stapleton-Rosebank
New Brighton-Silver Lake
Rossville-Woodrow
Great Kills
Arden Heights
Oakwood-Oakwood Beach
New Dorp-Midland Beach
Todt Hill-Emerson Hill-Heartland Village-Lighthouse Hill
Old Town-Dongan Hills-South Beach
Grasmere-Arrochar-Ft. Wadsworth
Charleston-Richmond Valley-Tottenville

6. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'NYC Taxi - Dendrogram'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
7. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

To adjust the dendrogram display, check all the available settings for this visual.

Shelves for dendrograms

Overview of shelves for Cloudera Data Visualization Dendogram visuals.

A Dendrogram visual has the following shelves:

Dimensions

Add two Dimensions fields from the DATA menu to define the start and end points of each connection in the plot.

This shelf accepts exactly two fields.

This is a mandatory shelf

Measure

Add a Measure field from the DATA menu. The aggregate determines the size of the connection.

This shelf accepts a single field.

This is a mandatory shelf.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Treemap

Cloudera Data Visualization enables you to create Treemap visuals. Treemaps display hierarchical data as a set of nested rectangles. Each 'branch' of the tree is a rectangle that corresponds to a particular dimension, with smaller rectangles that represent secondary dimensions. All rectangle areas are proportional to the aggregate measure that corresponds to the dimension.

About this task

Using color and size dimensions in the tree structure often brings to light new insights and correspondence patterns. Because of the highly efficient use of space, treemaps legibly display very large number of items simultaneously.

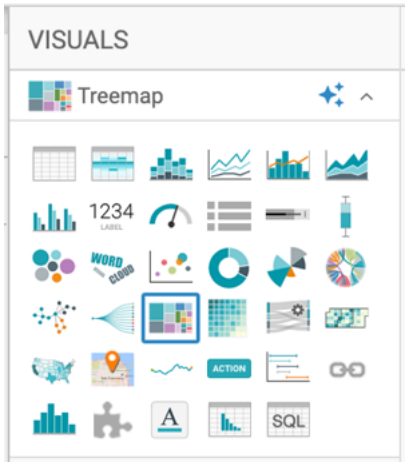
In the following example the NYC Taxi Trips dataset is used, which is based on the data in the ny-trips.csv file.

For an overview of shelves that specify this visual, see *Shelves for treemaps*.

Procedure

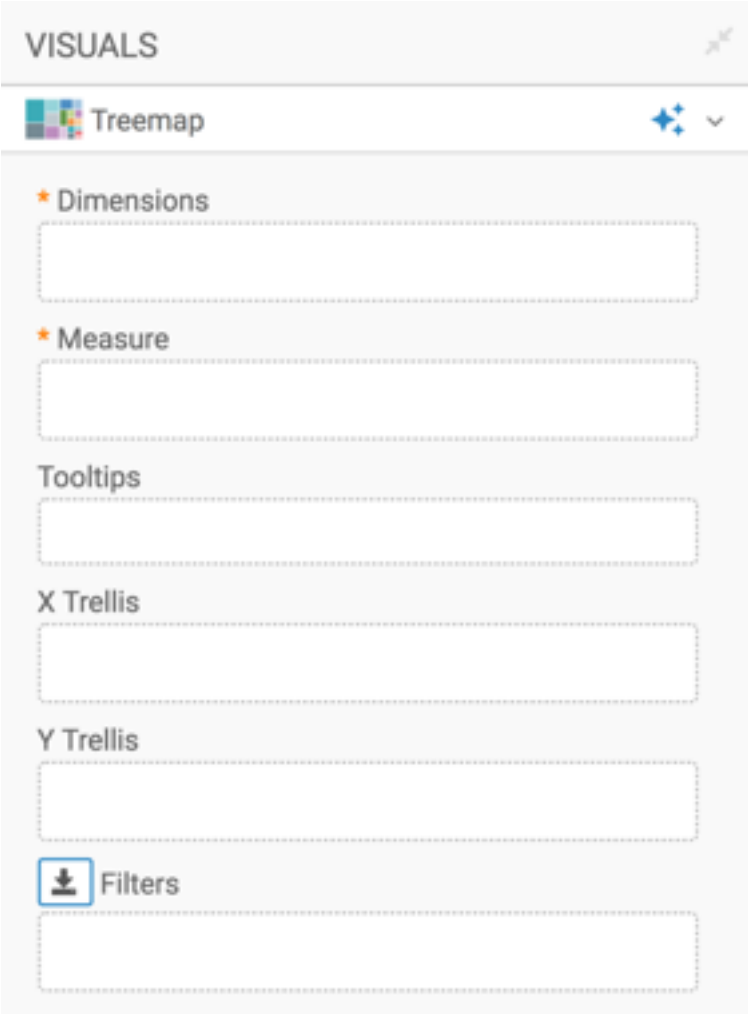
1. Start a new visual based on the NYC Taxi Trips dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Treemap.

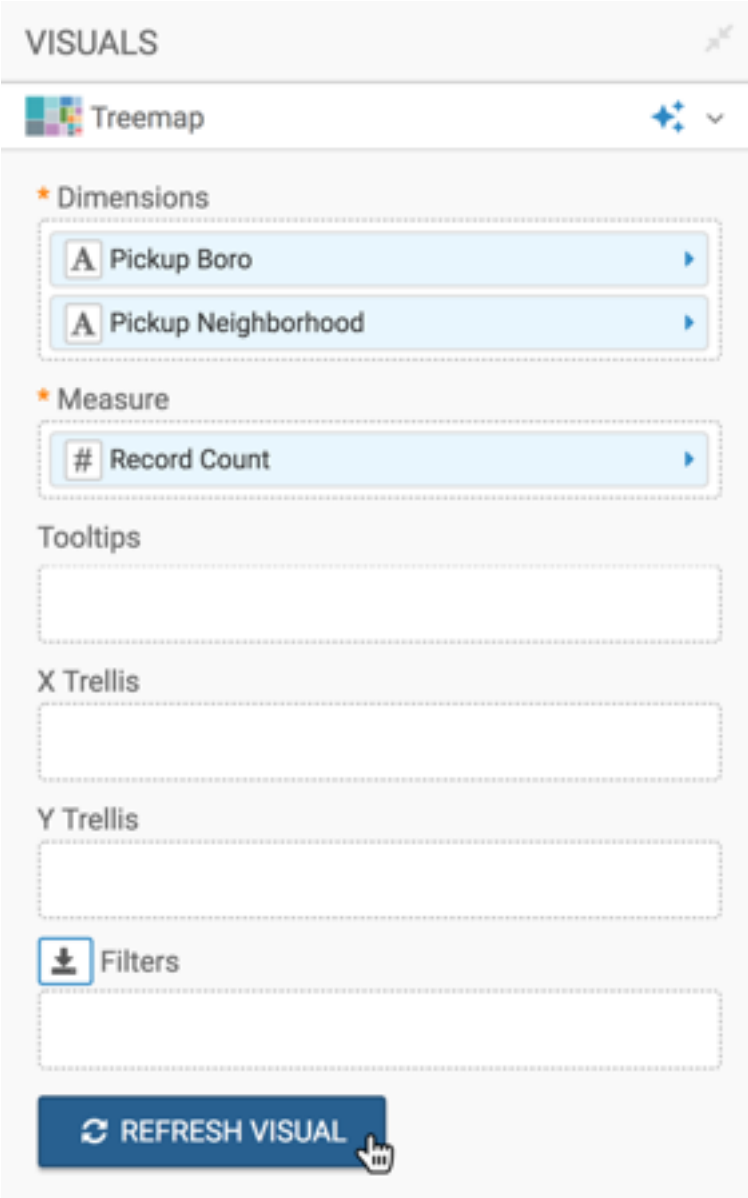


The shelves of the visual changed. They are now Dimensions, Measure, Tooltips, X Trellis, Y Trellis, and Filters. The mandatory shelves are:

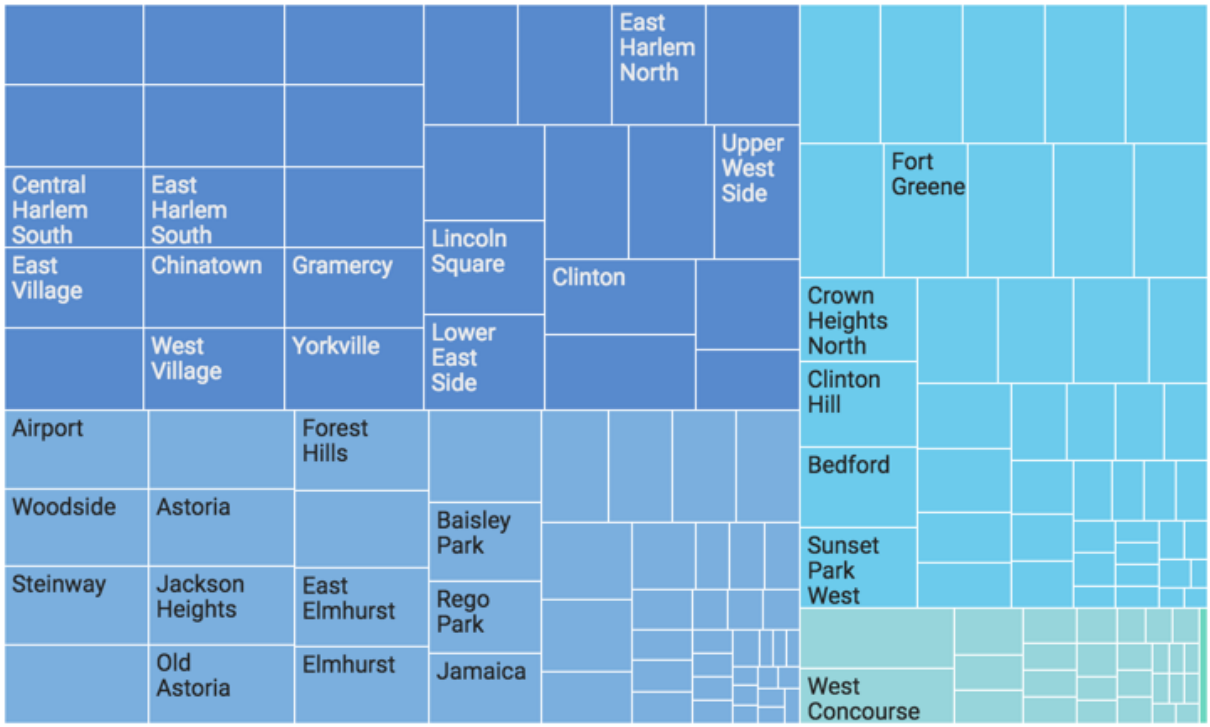
- Dimensions: specify one or two dimensions to define the branches and leaf nodes of the treemap.
- Measure: specifies one measurement that determines the area of the leaf node.



3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Dimensions, select Pickup Boro and Pickup Neighborhood, and add them to the Dimensions shelf, in that order.
   a) Under Measures, select Record Count and add it to the Measure shelf.



4. Click REFRESH VISUAL.
   The new treemap visual appears.



5. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'NYC Taxi - Treemap'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
6. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

To adjust the treemap display, check all the available settings for this visual.

Shelves for treemaps

Overview of shelves for Cloudera Data Visualization Treemap visuals.

A Treemap visual has the following shelves:

Dimensions

Add two Dimension fields from the DATA menu to define the branches of the treemap (primary rectangular divisions) and leaves of the treemap (the smaller nested divisions within each larger rectangular area).

This shelf accepts exactly two fields.

This is a mandatory shelf.

Measure

Add a Measure field from the DATA menu. The aggregate determines the size of the rectangular areas.

This shelf accepts a single field.

This is a mandatory shelf.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Correlation heatmap

Cloudera Data Visualization enables you to create Correlation Heatmap visuals. Correlation Heatmaps use colored cells, typically in a monochromatic scale, to show a 2D correlation matrix (table) between two discrete dimensions or event types. The values of the first dimensions appear as rows of the table, while the values of the second dimension are represented by the columns of the table. The color value of the cells is proportional to the number of measurements that match the dimensional values. This enables you to quickly identify incidence patterns, and to recognize anomalies.

About this task

Correlation Heatmap visuals are similar to Chords because they both compare exactly two dimensions. Correlation heatmaps are ideal for comparing the measurement for each pair of dimension values.

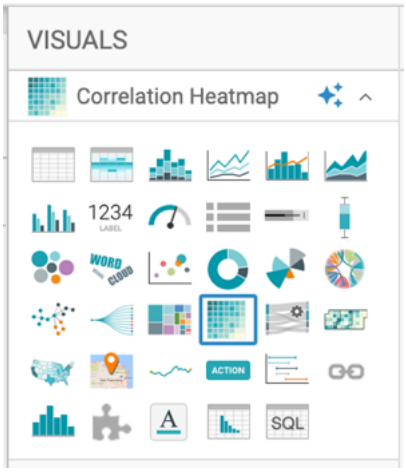
The following steps demonstrate how to create a new correlation visual on the SFPD Incidents dataset, which is based on data previously imported into Cloudera Data Visualization from the sfpd_incidents.csv datafile [data source default.sfpd_incidents].

For an overview of shelves that specify this visual, see *Shelves for correlation heatmaps*.

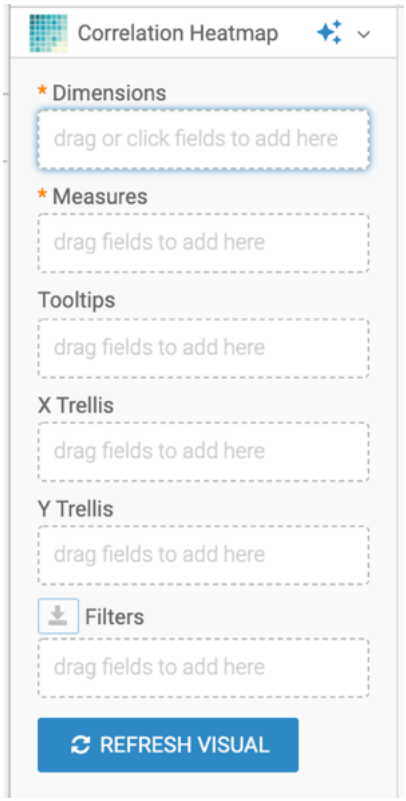
Note: This visual type supports smooth rendering of streaming data.

Procedure

1. Start a new visual based on the SFPD Incidents dataset.
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Correlation Heatmap.



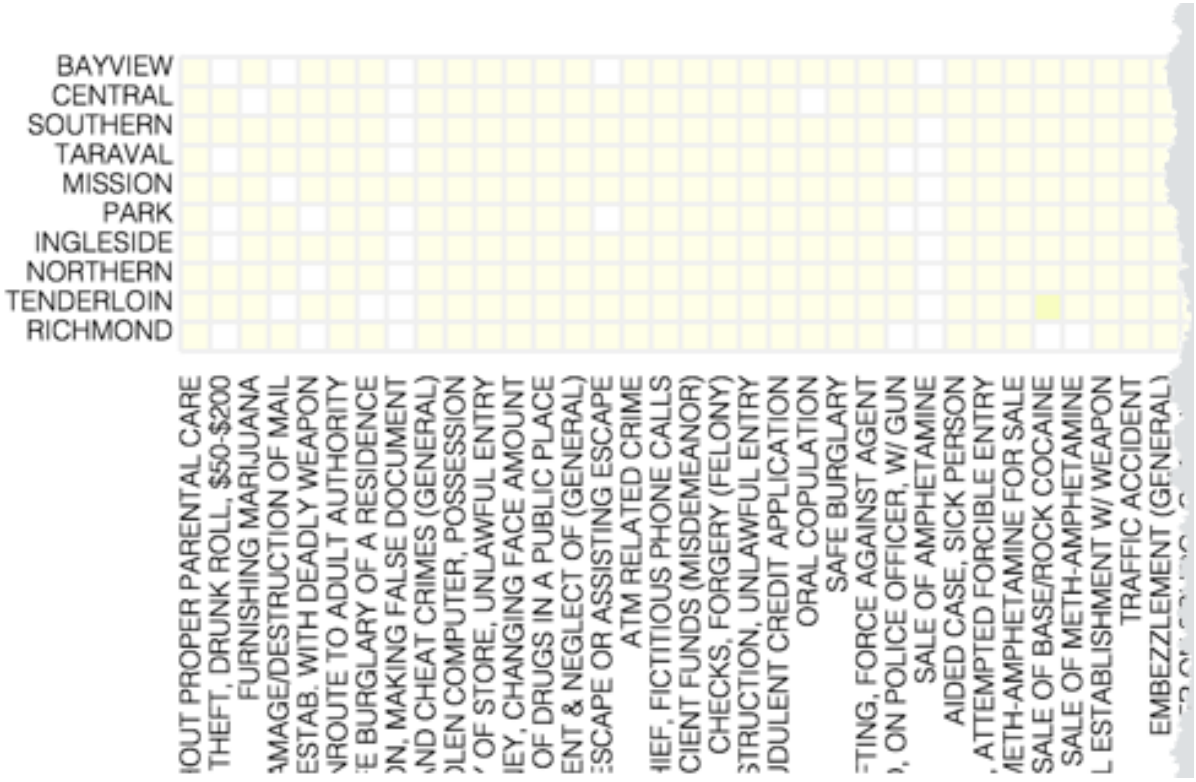
The shelves of the visual changed. They are now Dimensions, Measures, Tooltips, X Trellis, Y Trellis, and Filters. Both Dimensions and Measures are mandatory.



3. To show specific items, populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Dimensions, select pddistrict and drag it to the Dimensions shelf.
   b) Under Dimensions, select descript and drag it to the Dimensions shelf.
   c) Under Measures, select Record Count and drag it to the Measures shelf.

      Record Count is defined as a sum of events. If you hover over it with your mouse, you can see a black detail bubble with sum(1) contents.
4. Click REFRESH VISUAL.

   The default correlation heatmap visual appears.

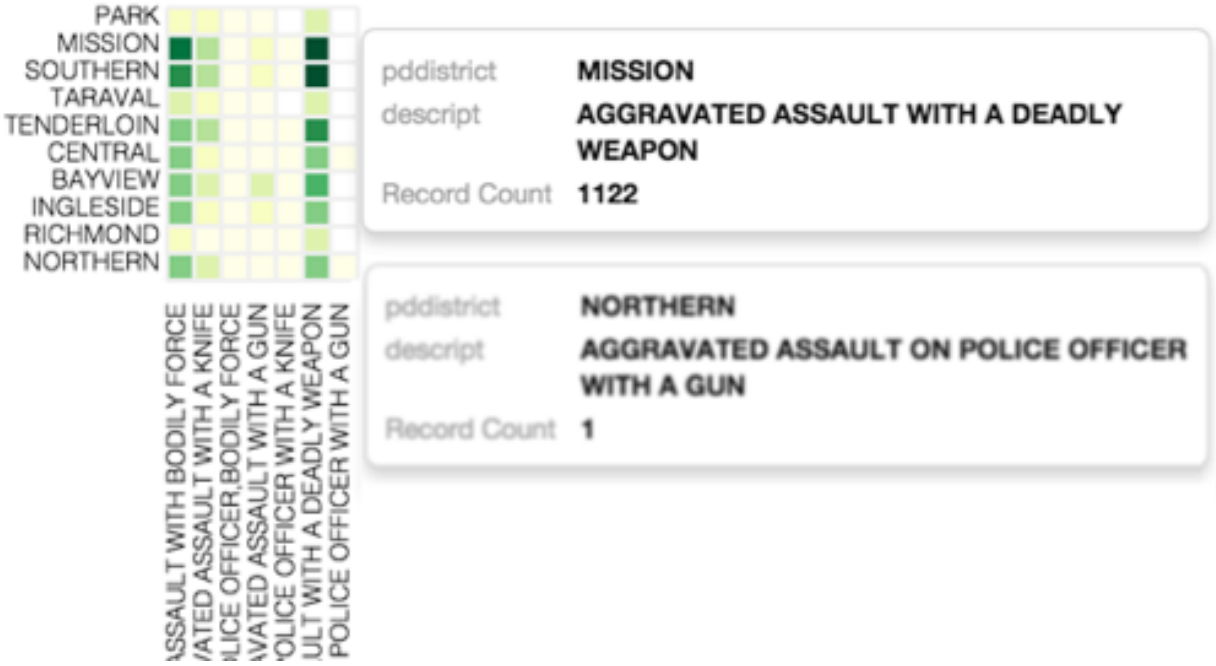


You can see that this dataset has a very large number of possible values that represent the columns of the table. If you scroll to the right, you will see some cells rendered in dark shades of green.

5. To examine a shorter list of categories, you can add some filtering to the visual.
   a) Under Dimensions, select datetime and drag it to the Filters shelf .
   b) Repeat it with category and descript.
   c) On the Filters shelf, click the (down arrow) on the descript field, and then click Pick values from a list.
   d) Select a number of values.

      In this example, 7 distinct options were picked.
6. Click REFRESH VISUAL.

   You can see that this smaller matrix also shows the entire range of color values.



7. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'SFPD Incidents - Correlation Heat Map'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
8. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for correlation heatmaps

Overview of shelves for Cloudera Data Visualization Correlation Heatmap visuals.

A Correlation Heatmap visual has the following shelves:

Dimensions

Add two Dimension fields from the DATA menu. The first dimension defines the rows and the second dimension defines the columns of the plot.

This shelf accepts exactly two fields.

This is a mandatory shelf.

Measures

Add a Measure field from the DATA menu to specify the values aggregated across columns and rows.

This shelf accepts a single field.

This is a mandatory shelf.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Correlation flow

Cloudera Data Visualization enables you to create Correlation Flow visuals. Correlation Flows enable quick visual analysis of co-occurrence of two or more distinct dimensions. In simple form, a Correlation Flow visual is a linear representations of a two-dimensional Correlation Heatmap matrix. Where a Correlation Heatmap represents correlation values through color intensity, the Correlation Flow plots thicker lines and taller dimension boxes to represent greater correlation.

About this task

The following steps demonstrate how to create a new Correlation Flow visual on the Customer Value Analysis dataset. This dataset is based on data previously imported into Cloudera Data Visualization from the customer-value-analysis.csv datafile.

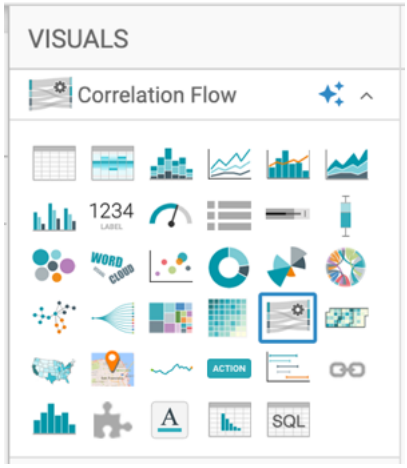
For an overview of shelves that specify this visual, see *Shelves for correlation flows*.

Note: Correlation flow is not limited to two dimensions. You can model a large number of correlation 'dimensions'. Ordering of the dimensions on the shelves of the visual provides different insights into the dataset.

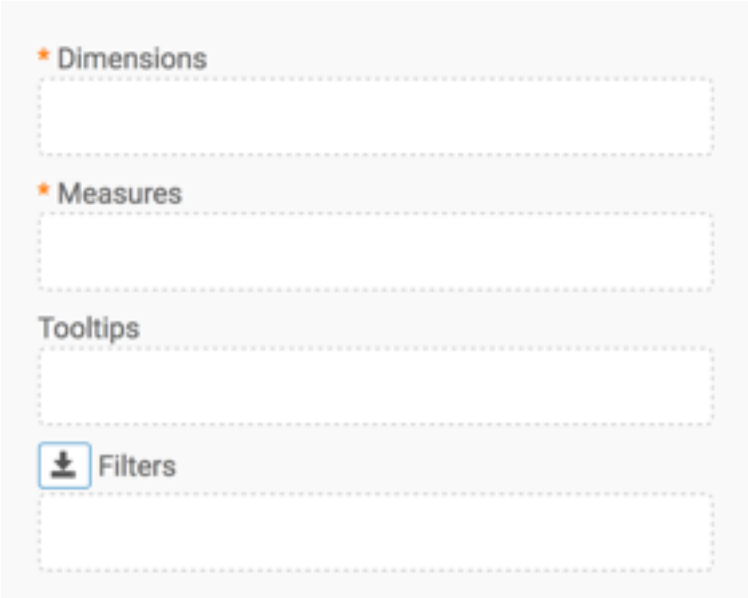
Procedure

1. Start a new visual based on the Customer Value Analysis dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Correlation Flow.



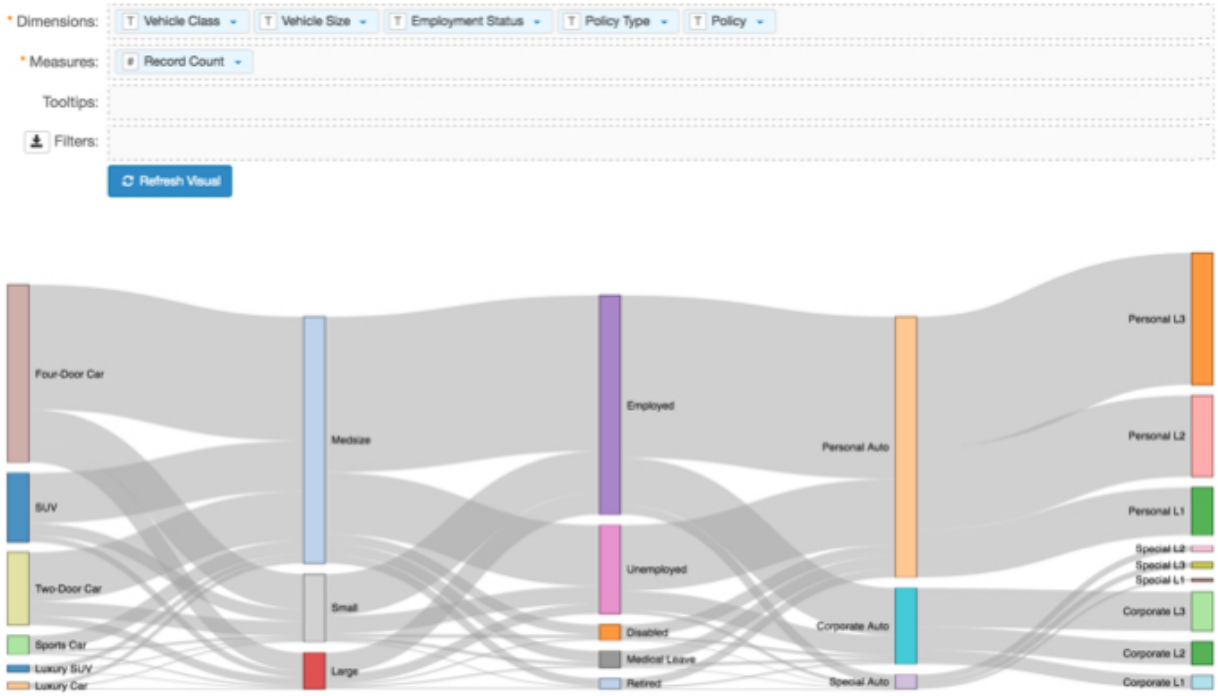
The shelves of the visual changed. They are now Dimensions, Measures, Tooltips, and Filters. Both Dimensions and Measures are mandatory.



3. Populate the shelves from the available fields:
   a) Under Dimensions, select Vehicle Class and drag it to the Dimension shelf.
   b) Repeat it for Vehicle Size, Employment Status, Policy Type, and Policy.
   c) Under Measures, select Record Count and drag it to the Measures shelf.
4. Click REFRESH VISUAL.

   The correlation flow visual appears.

Figure 20: Correlation Flow Visual



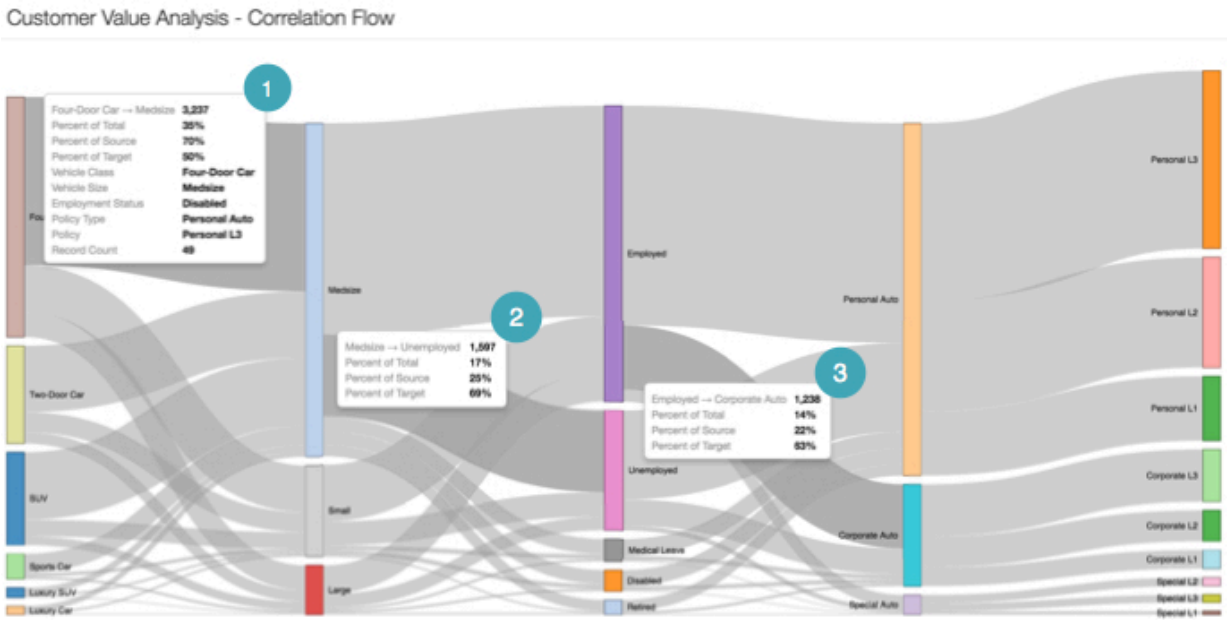
You can see that each distinct line segment between two dimensional events reports its own correlation data, as demonstrated by the Tooltips in the following chart.

- Tooltip 1 demonstrates the correlation between Four-Door and Medsize vehicles. The dataset has 3,237 Four-Door cars, and they represent 36% of the total. Additionally, 70% of these cars are considered Medsize.

Viewing this from the other direction of the correlation relationship, 50% of all Medsize cars are four-door cars.

- Tooltip 2 demonstrates the correlation between Medsize vehicles and Unemployed owners. It shows that 1,597 people are in both categories, and this segment represents 17% of all records. 25% of Medsize cars are owned by people who are unemployed, while 69% of unemployed people own a medsize vehicle.
- Tooltip 3 demonstrates the correlation between Employed owners and the ones who have a Corporate Auto vehicle policy. It shows that 1,238 people are in these two categories, which is 14% of the total. 22% of employed owners have a corporate auto policy, while 63% of people with a corporate auto policy are actively employed.

Figure 21: Correlation Flow Visual



5. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'Customer Value Analysis - Correlation Flow'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
6. At the top left corner of the Dashboard Designer, click SAVE.

Shelves for correlation flows

Overview of shelves for Cloudera Data Visualization Correlation Flow visuals.

A Correlation Flow visual has the following shelves:

Dimensions

Add two Dimension fields from the DATA menu to define the start and end of each connection of the plot.

This shelf accepts exactly two fields.

This is a mandatory visual

Measures

Add a Measure field from the DATA menu. The aggregate measurement determines the thickness of the connector between each pair of dimensions.

This shelf accepts a single field.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Calendar heatmaps

Cloudera Data Visualization enables you to create Calendar Heatmap visuals. A Calendar Heatmap uses colored cells, typically in a monochromatic scale, to show relative number of events for each day in a calendar view. Days are arranged into columns by week and grouped by month and years. This enables you to quickly identify daily and weekly patterns, and to recognize anomalies.

About this task

The following steps demonstrate how to create a new calendar visual on the SFPD Incidents dataset, based on data previously imported into Cloudera Data Visualization from the sfpd_incidents.csv.zip datafile.

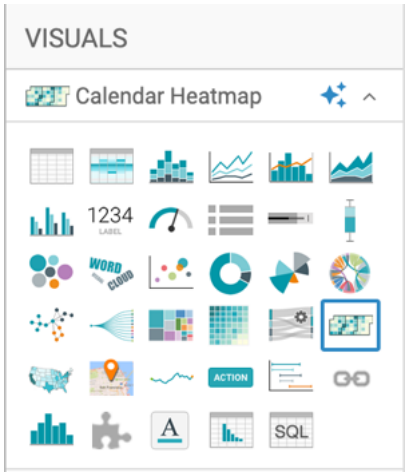
For an overview of shelves that specify this visual, see *Shelves for calendar heat maps*.

Note: This visual type supports smooth rendering of streaming data.

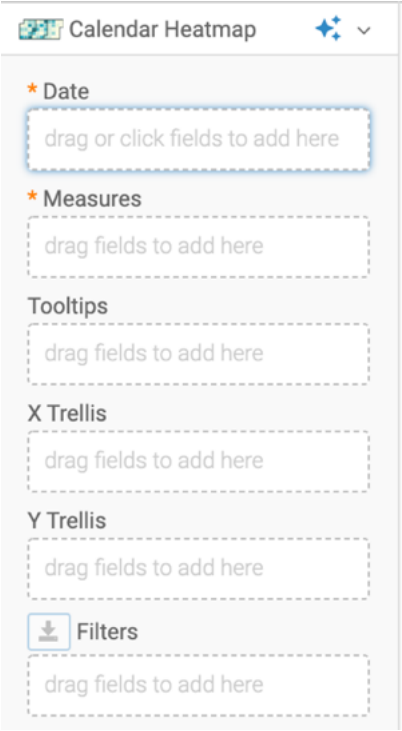
Procedure

1. Start a new visual based on the SFPD Incidents dataset.

   For instructions, see *Creating a visual*.
2. Select Calendar Heatmap in the VISUALS menu.

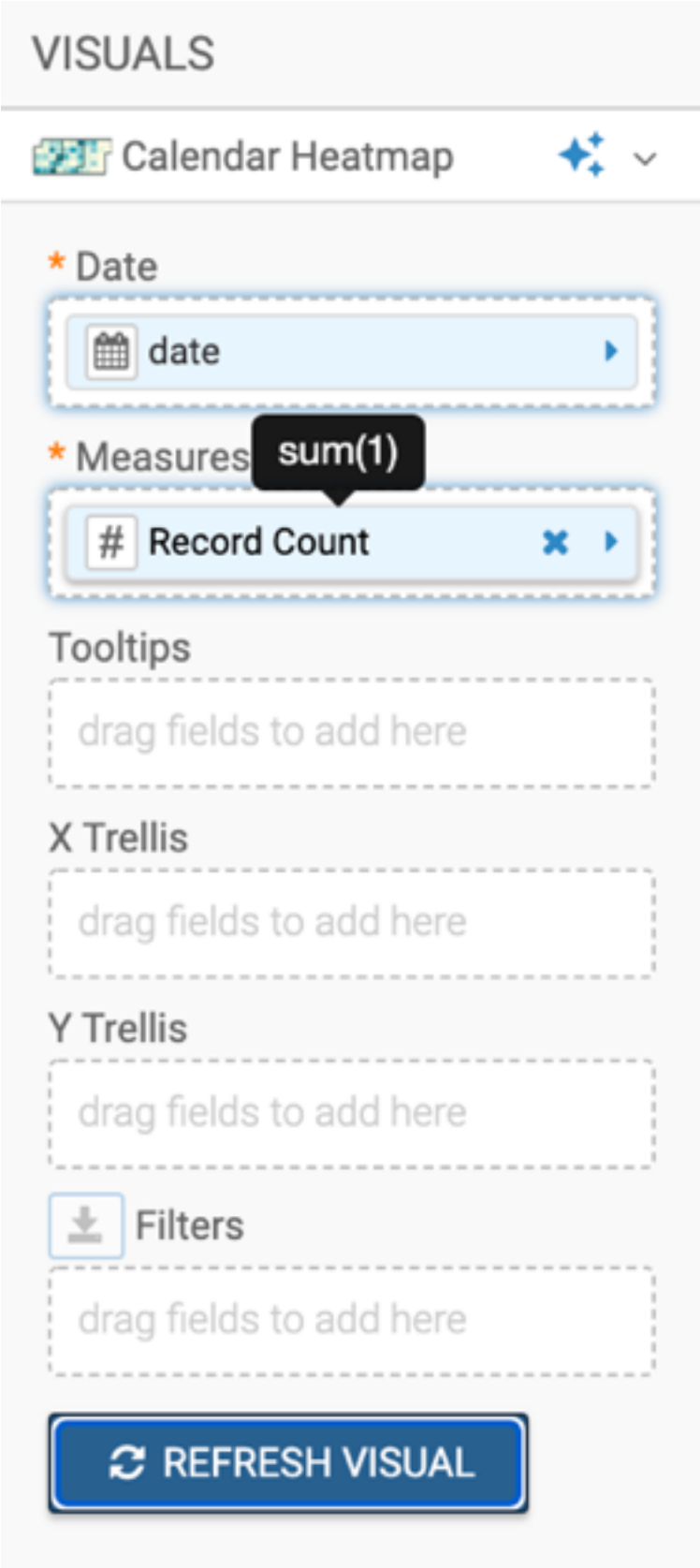


The shelves of the visual changed. They are now Date (mandatory), Measures (mandatory), Tooltips, X Trellis, Y Trellis, and Filters.



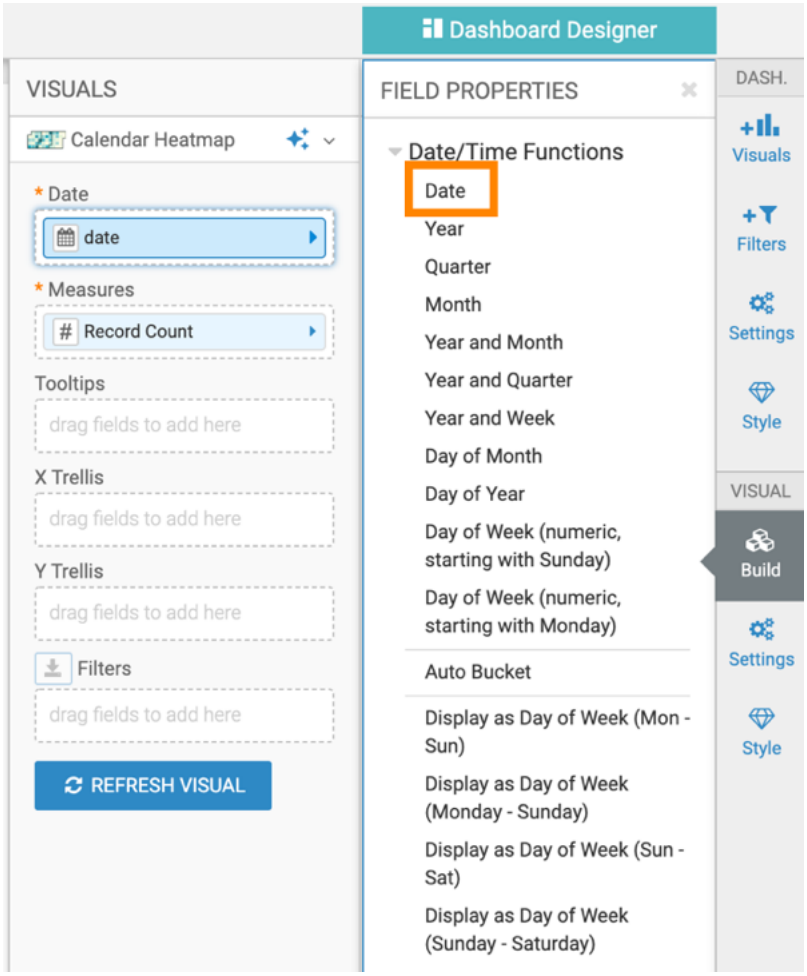
3. To show specific items, populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select date and drag it to the Date shelf.
   b) Under Measures, select Record Count and drag it to the Measures shelf.

   Note: Record Count is defined as a sum of events. If you hover over it with your mouse, you can see a black detail bubble with sum(1) contents.



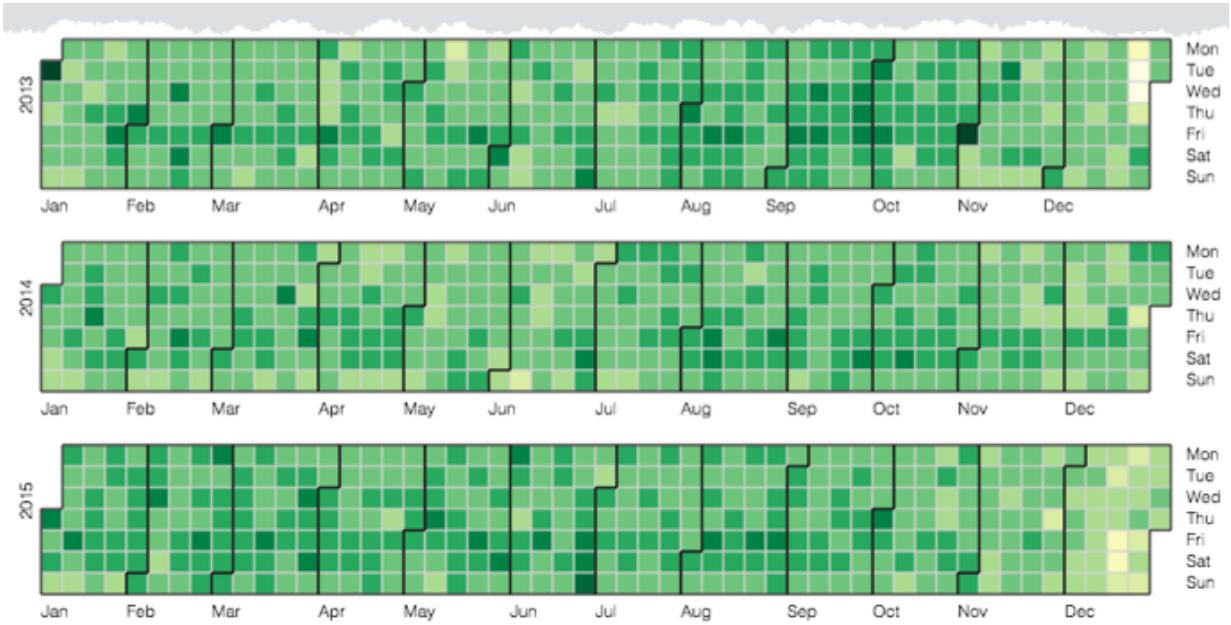
4. The field on the Date shelf should have a calendar icon to indicate that it is in the Date/Time category of data types. If the field is not automatically recognized as a date data type, convert it.

   On the Date shelf of the visual, click the date field, select Date/Time Function Date .



5. Click REFRESH VISUAL.

   The default calendar heat map visual appears.



6. Change the title of the visual to SFPD Incidents - Calendar Heat Map.
7. Click SAVE at the top left corner of the Dashboard Designer.

Shelves for calendar heatmaps

Overview of shelves for Cloudera Data Visualization Calendar Heatmap visuals.

A Calendar Heatmap visual has the following shelves:

Date

Add a date field from the DATA menu to determine the range of the calendar.

This shelf accepts a single field.

This is mandatory shelf.

Measures

Add a Measure field from the DATA menu to specify the values aggregated across the dates of the calendar.

This shelf displays only one field on the visual. Additional fields appear only in the Tooltip.

This is mandatory shelf.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Maps

Cloudera Data Visualization enables you to create Map visuals.

A geomap of a country, continent, or region superimposes measurement values onto a geographic map.

Cloudera Data Visualization has two basic geomap representation options:

- A filled or choropleth map (Greek #### for area or region and ###### for multitude) shows geographical areas that are shaded in proportion to the value of the measurement that is displayed on the map.

  Tip: Consider the following when building choropleth maps:

  - Use choropleth maps when your data are attached to enumeration units, such as counties, countries, zipcodes, and so on.
  - Standardize data to show rates and ratios instead of raw data counts. For example, instead of plotting gross GDP, plot GDP per Capita.
  - Use choropleth maps when you have data with continuous statistical surface. For example, Population is a count and representing it in colored map may be misleading. Population density is a ratio with a continuous statistical surface (even where 0 represents unpopulated areas); so it can be used in choropleth maps.

- A geomap with symbols or marks, such as bubbles, superimposes these symbols onto a geographical map. By default, the area of the marks corresponds to the relative size of the measurement.

Cloudera Data Visualization supports the following geographical specifications by automatically recognizing them as qualified GEO dimensions inside the dataset:

- country
- state or province for United States and Canada, respectively
- county, for United States only
- zipcode or postal code for United States and Canada, respectively
- longitude and latitude
- DMA, for Nielsen Designated Market Area Regions, in United States

If the relevant field in your dataset is not named correctly, be sure to adjust it using the Change Type Geo Option, described in *Change type*. Note also that you must use Geo Type fields when limiting the Drill Into options. For more information, see *Limiting drill into options*.

To build the two types of map visuals with specified geographical units, follow the instructions for each option. Additionally, see the description on how to build a *Trellised map*.

For an overview of shelves that specify this visual, see *Shelves for maps*.

Choropleth country map

Map visuals can show relative measurement values for countries as colored areas. A world map can be used to show a comparison of measurement values among countries. With data that meets ratio and statistical continuity guidelines, a choropleth map is a good chart choice.

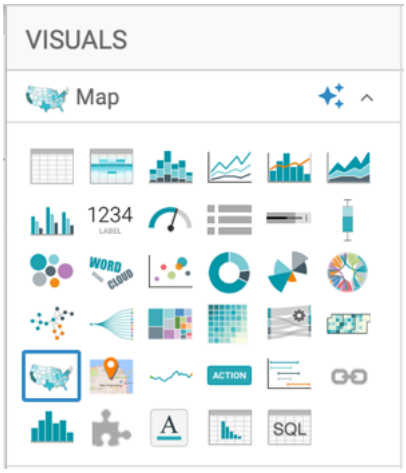
About this task

The following steps demonstrate how to create a new map visual on the dataset World Life Expectancy [data source samples.world_life_expectancy]. It produces colored regions for the measurements that correspond to the country field.

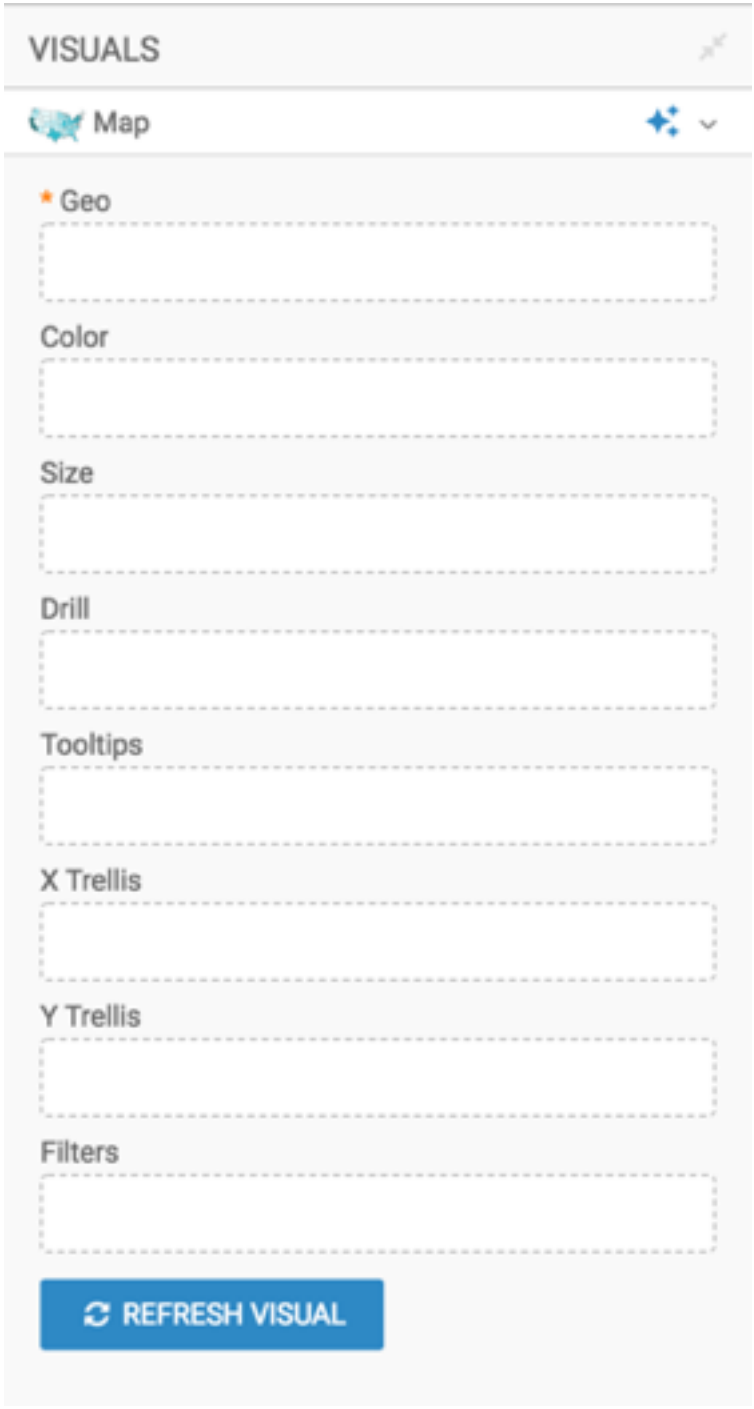
Procedure

1. Start a new visual based on the World Life Expectancy dataset.

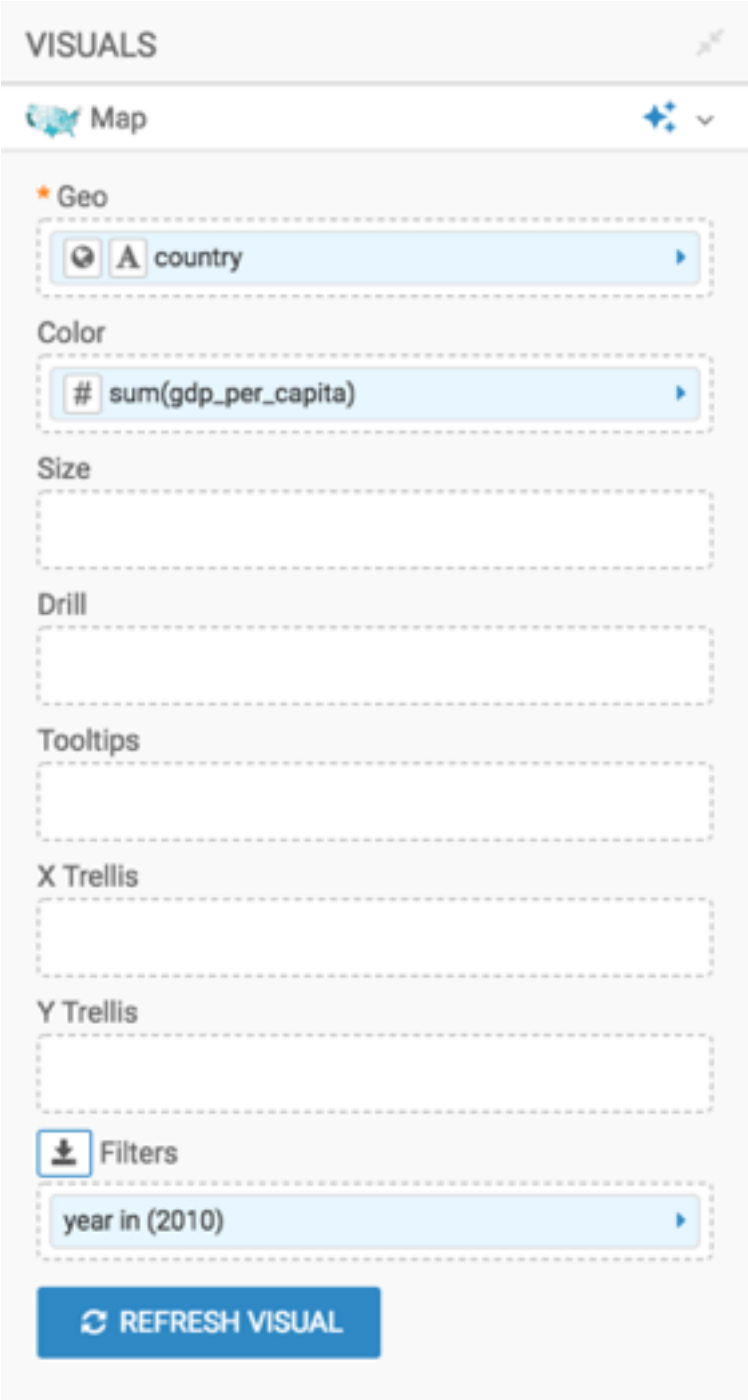
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Map.



   The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Dimensions, select country and add it to the Geo shelf.
   b) Under Measures, select gdp_per_capita and add it to the Color shelf.
   c) Under Measures, select year and add it to the Filters shelf. When the Filter for year window modal appears, select 2010 in the Values tab, and click Apply.



4. Click REFRESH VISUAL.
   The map visual graph appears.

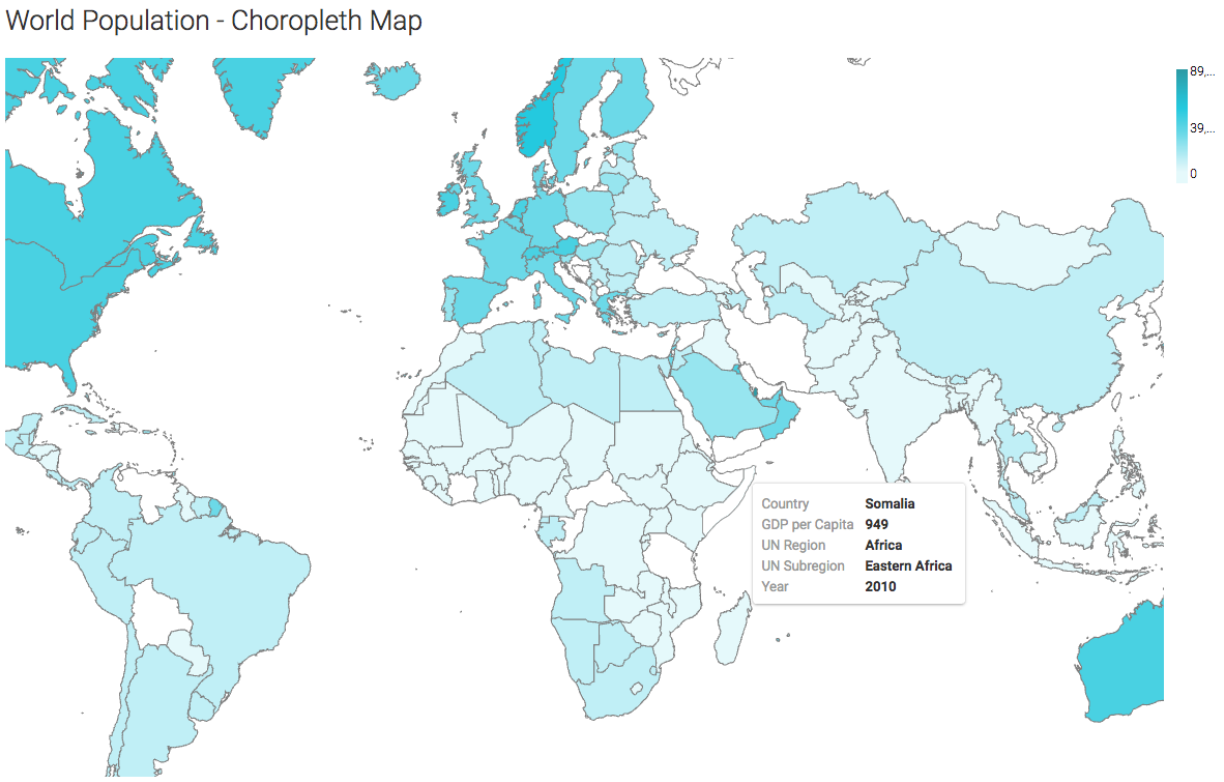
5. On the Tooltips shelf, add several fields from Dimensions and Measures.

   This enables you to see specific descriptive information in your visuals, such as input values, segment affiliation, and calculations.

   For example, from Dimensions, add un_subregion, un_region, and year onto the Tooltip shelf.
6. Alias the fields on the shelves to improve the readability of the visual and its tooltip. See *Alias*.
   - gdp_per_capita as GDP per Capita
   - un_subregion as UN Subregion
   - un_region as UN Region
   - year as Year
7. Format the number reported by the GDP per Capita field to include a thousands separator, and format Year not to use the thousands separator.

   For more information, see *Real number display format* and *Integer display format*.
8. [Recommended] Enable panning and zooming option, as described in *Customizing zoom*.
9. Change the color palette to blue, as described in *Change color palette of visuals*.
10. Click REFRESH VISUAL.

    The map visual appears. You can move the visual (pan), zoom in and out, and see the details in the Tooltips text.



11. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

In this example, the title is changed to 'World Population - Choropleth Map'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

12. At the top left corner of the Dashboard Designer, click SAVE.

Country map with bubbles

Cloudera Data Visualization Map visuals can use bubbles to show relative measurement values for countries. You can display two measurements at the same time: the first measurement as the color of the bubble, and the second one as the relative size of the bubble.

About this task

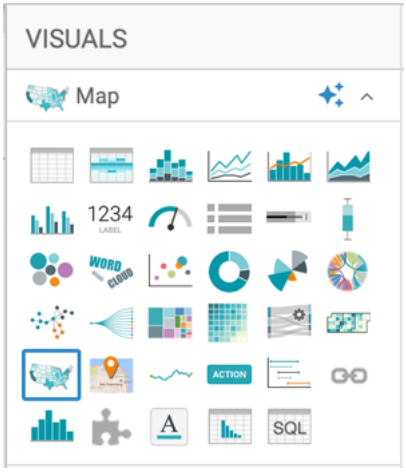
The following steps demonstrate how to create a new map visual on the dataset World Life Expectancy [data source samples.world_life_expectancy]. It produces bubbles with area that corresponds to the relative population in each country.

Procedure

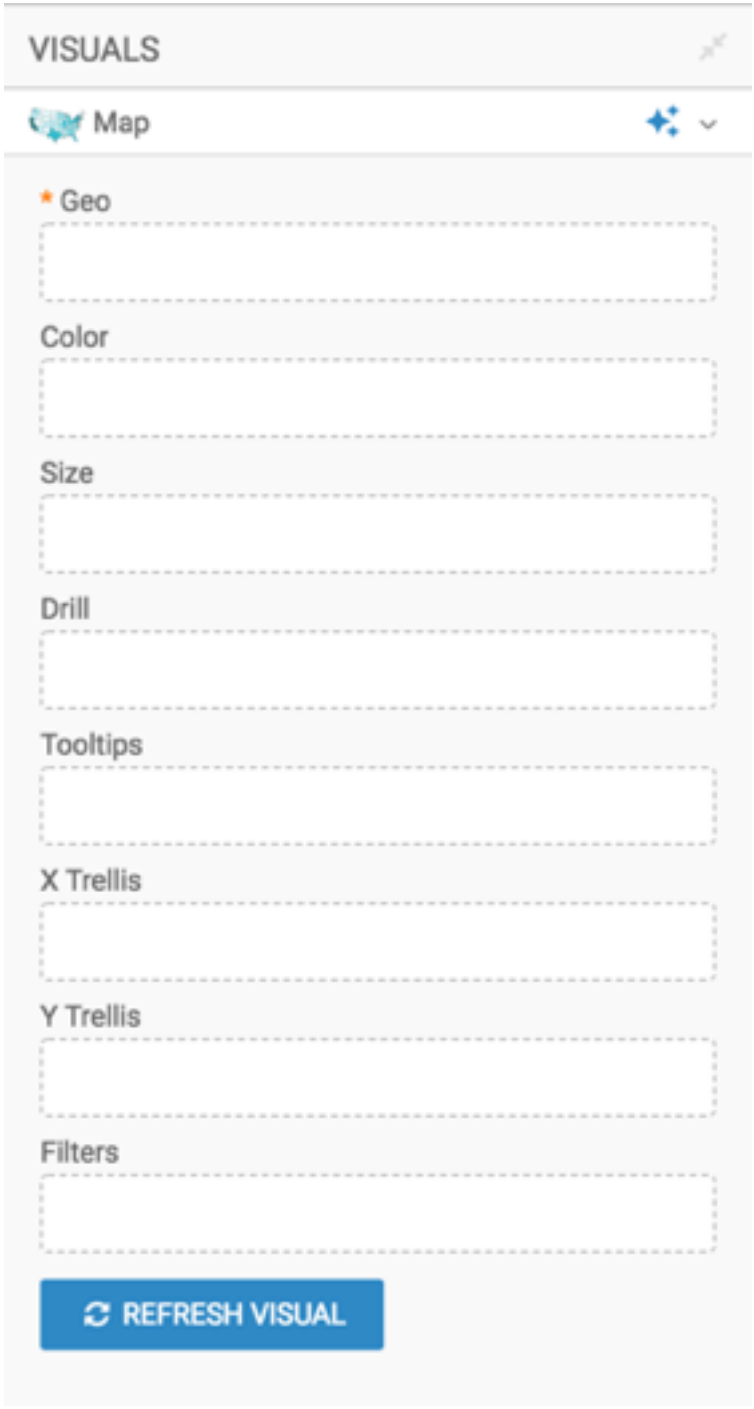
1. Start a new visual based on the World Life Expectancy dataset.

For instructions, see *Creating a visual*.

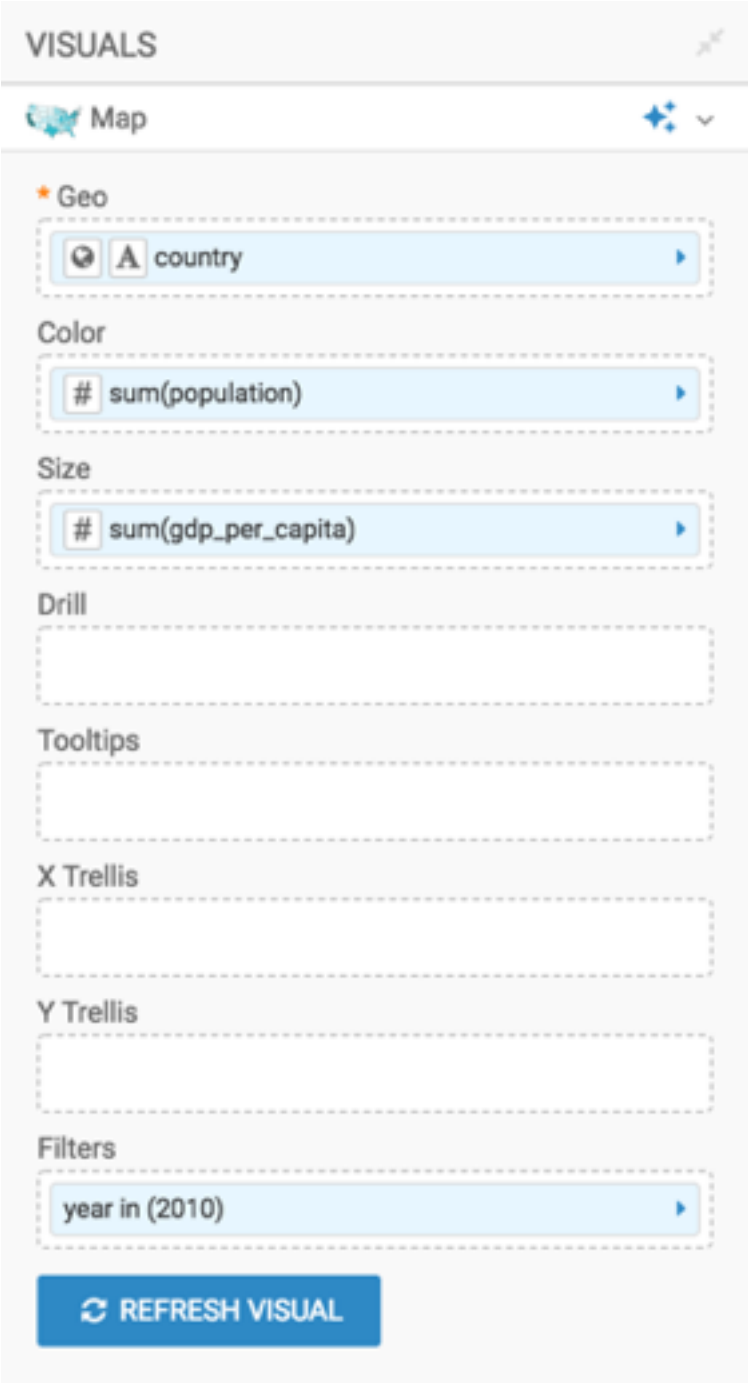
2. In the VISUALS menu, find and click Map.



The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Dimensions, select country and add it to the Geo shelf.
   b) Under Measures, select population and add it to the Color shelf.
   c) Under Measures, select gdp_per_capita and add it to the Size shelf.
   d) Under Measures, select year and add it to the Filters shelf. When the Filter for year window modal appears, select 2010 in the Values tab, and click APPLY.



4. Change the map from choropleth to bubble marks, as described in *Displaying bubbles on maps*.
5. Adjust the size range of the marks to 1-10, as described in *Changing the mark size range*.
6. Click REFRESH VISUAL.

   The map visual graph appears.
7. On the Tooltips shelf, add several fields from Dimensions and Measures.

   This enables you to see specific descriptive information in your visuals, such as input values, segment affiliation, and calculations.

   For example, from Dimensions, drag un_subregion, un_region, and year to the Tooltips shelf.
8. Alias the fields on the shelves to improve the readability of the visual and its tooltip.

- population as Population
- un_subregion as UN Subregion
- un_region as UN Region
- year as Year

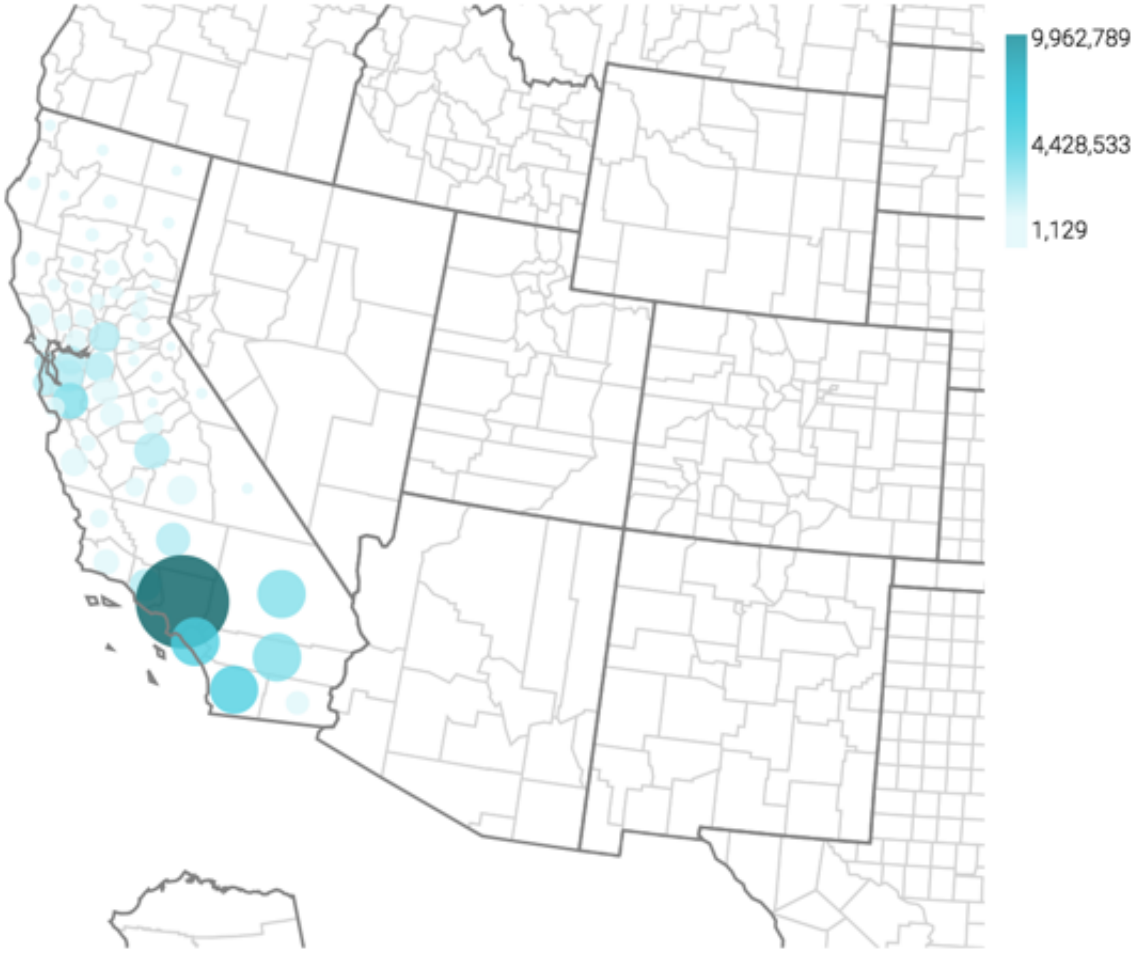
For instructions, see *Alias*.

9. Format the number reported by the GDP per Capita field to include a thousands separator, and format Year not to use the thousands separator.

   For more information, see *Real number display format* and *Integer display format*.

10. [Recommended] Enable panning and zooming option, as described in *Customizing zoom*.
11. Change the color palette to blue, as described in *Change color palette visuals*.
12. Click REFRESH VISUAL.

    You can move the visual (pan), zoom in and out, and see the details in the Tooltips text.



13. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

    In this example, the title is changed to 'World Population - Map with Bubbles'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

14. At the top left corner of the Dashboard Designer, click SAVE.

Choropleth US state map

Cloudera Data Visualization Map visuals let you plot measurements, as areas, over US State regions.

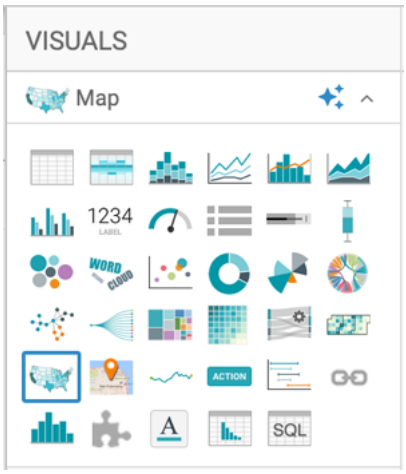
About this task

A map of the United States shows a comparison of measurement values among US States. The following steps demonstrate how to create a new map visual on the dataset US County Population [data source samples.us_counties]. It produces colored regions for the measurements that correspond to the state field.

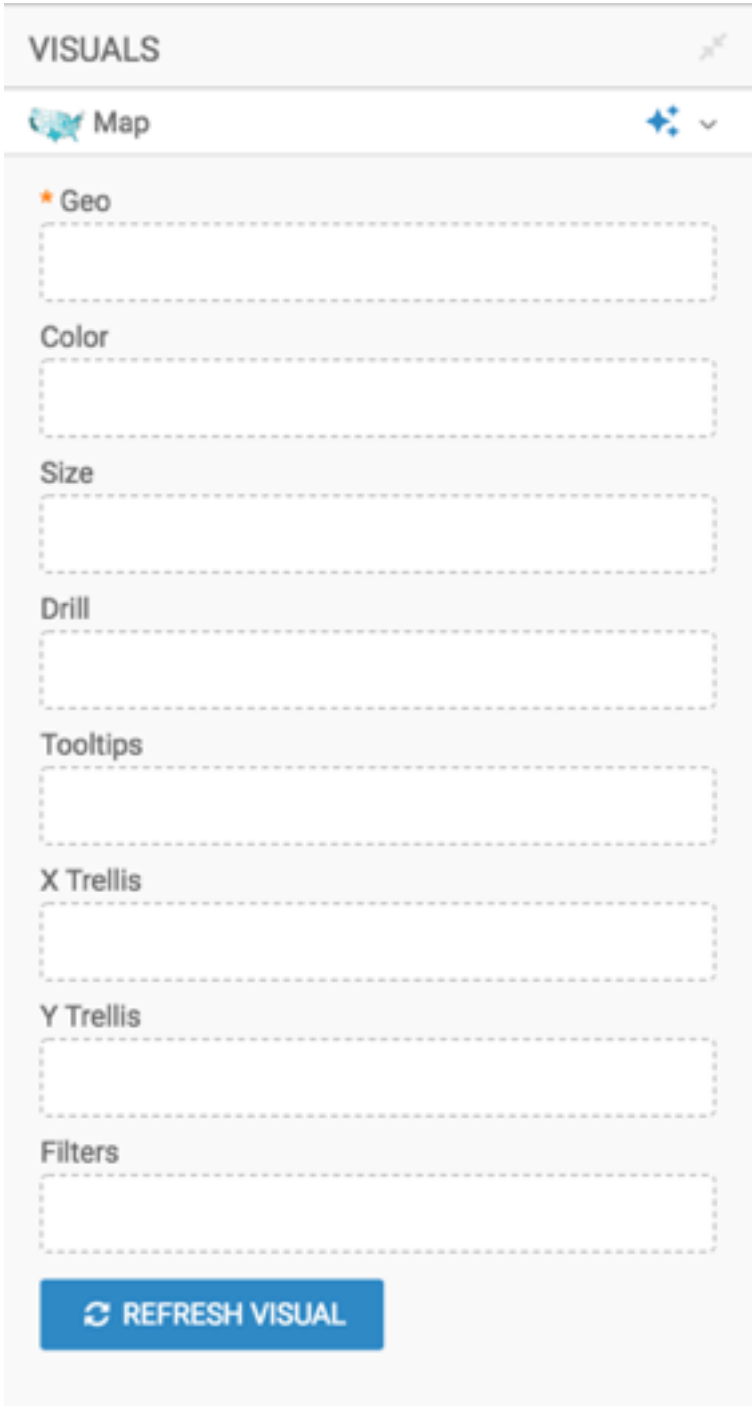
Procedure

1. Start a new visual based on the US County Population dataset.

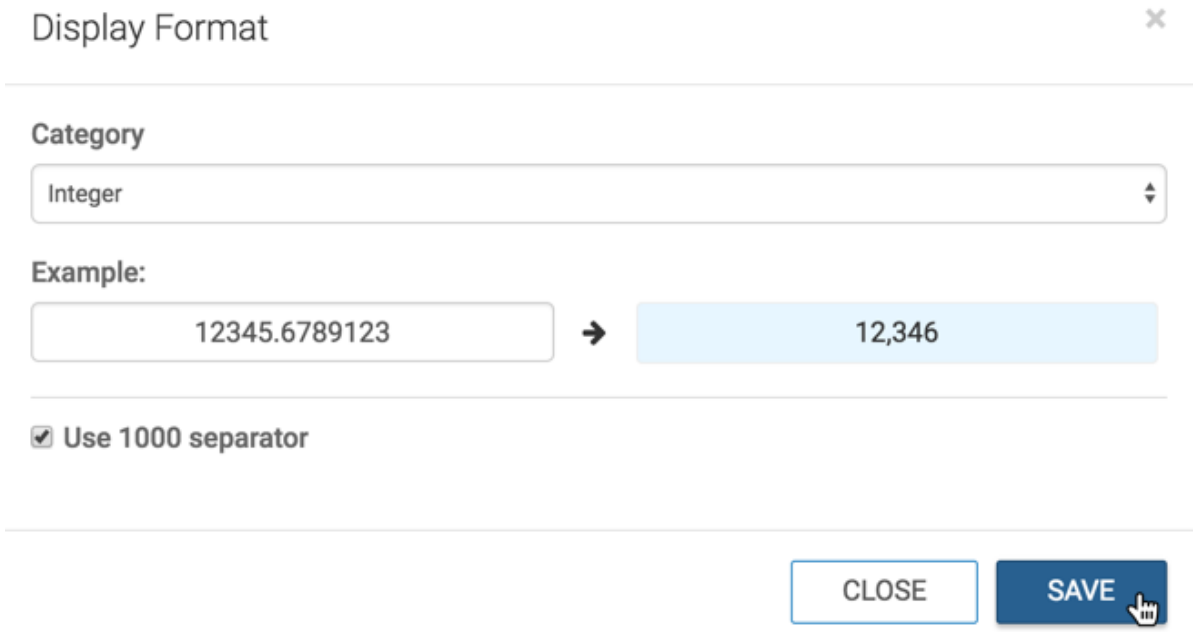
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Map.



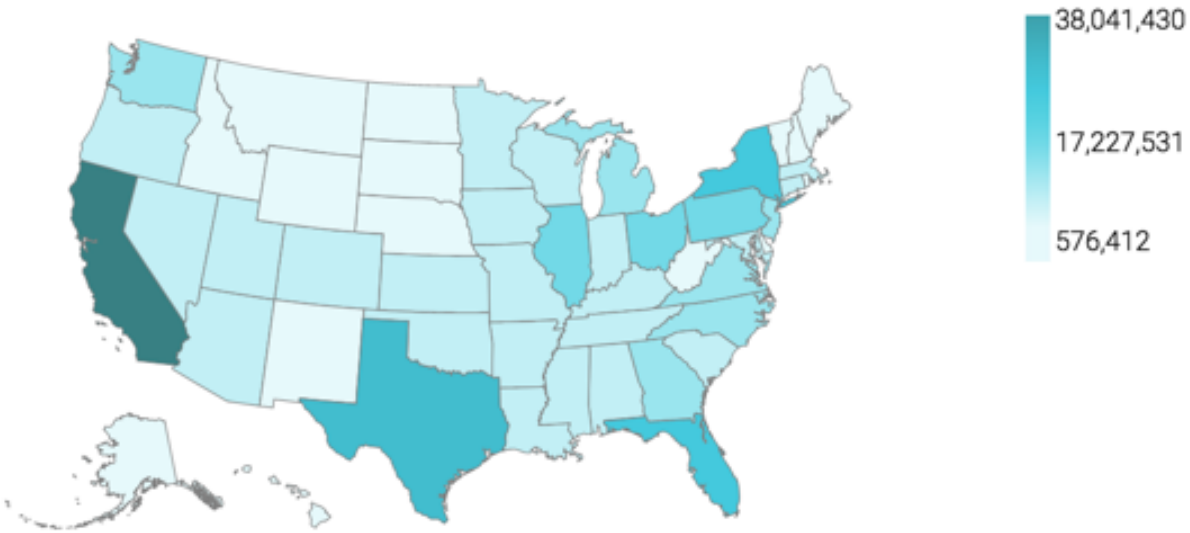
The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select stname and add it to the Geo shelf.
   b) Click the field to open the FIELD PROPERTIES menu, expand the Change Type option, and select the Geo Type State.
   c) Under Measures, select tot_pop and add it to the Colors shelf.
   d) Click the field to open the Field Properties menu, select Display Format,Category Integer, and select the thousands separator option. Click Save.



4. Alias the field stname as State, and sum(tot_pop) as Population; see *Alias*.
5. Enable panning and zooming option, as described in *Customizing zoom*.
6. Change the color palette, as described in *Change color palette of visuals*.
7. Click REFRESH VISUAL.
   The map visual appears.



8. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.
   In this example, the title is changed to 'US State Population - Choropleth Map'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
9. At the top left corner of the Dashboard Designer, click SAVE.

US state map with bubbles

Cloudera Data Visualization Map visuals let you plot measurements, as bubbles, over US State regions. Bubble maps display up to two measurements simultaneously: as the color of the bubble, and as the relative size of the bubble.

About this task

A map of the United States shows a comparison of measurement values among US States. The following steps demonstrate how to create a new map visual on the dataset US County Population [data source samples.us_counties]. This produces a map with bubble marks that correspond to the total population of each state.

Procedure

1. Clone the visual that you developed in *State map, choropleth*.

   For instructions, see *Cloning a visual*.
2. Add a measure to the Size shelf.

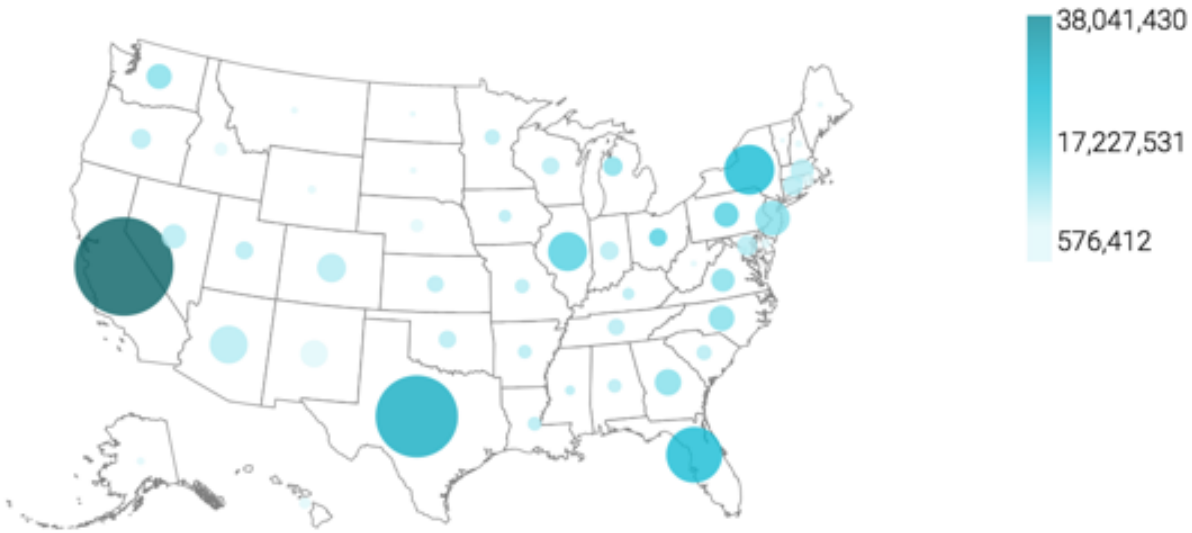
   For example, to show the relative size of the Hispanic population across the country, use the following expression in the Enter/Edit Expression editor:

   ```
   sum([h_female]+[h_male])
   ```

   For more information, see *Enter or edit expressions*.
3. Change the map from choropleth to bubble marks, as described in *Displaying bubbles on maps*.
4. Adjust the size range of the marks to 1-10, as described in *Changing the mark size range*.
5. Alias all the fields.

   For instructions, see *Alias*.
6. Change the color palette or simply reverse it, as described in *Change color palette of visuals* and *Reverse color palette of visuals*.
7. Enable panning and zooming option, as described in *Customizing zoom*.
8. Click REFRESH VISUAL.

   The map visual appears.



9. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'US State Population - Bubbles'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
10. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

You can move the visual (pan), zoom in and out, and see the details in the Tooltips text.

Choropleth US county map

In Cloudera Data Visualization Map visuals, you can plot measurements as areas over US County regions.

About this task

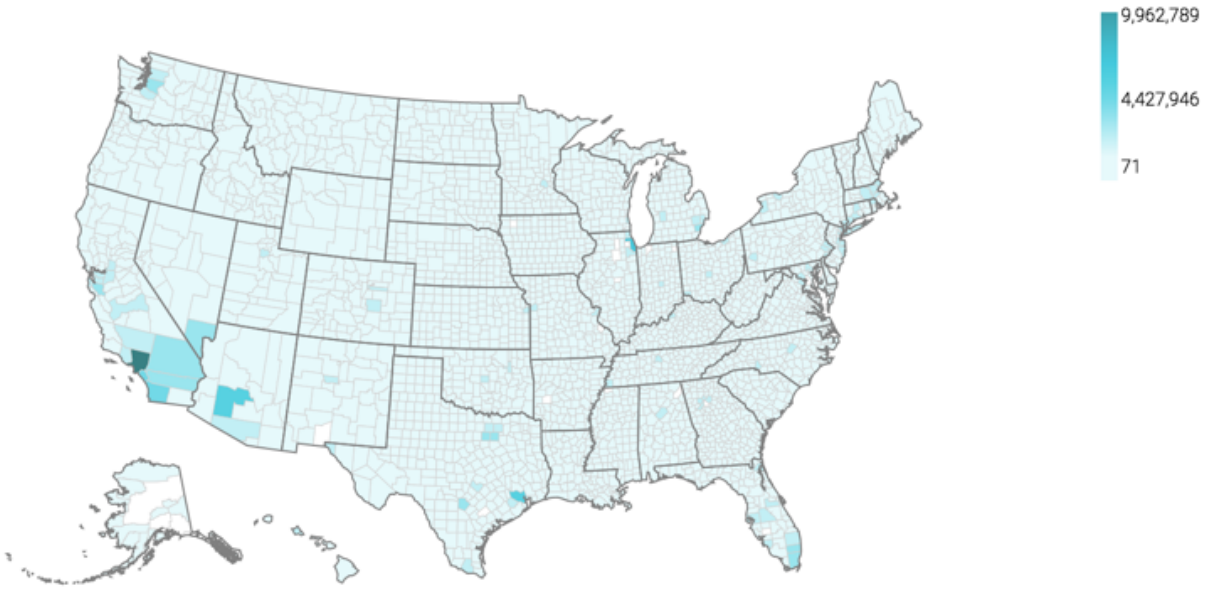
A county map of the United States shows a comparison of measurement values across US States. The following steps extend the work completed in *State map, choropleth*, and demonstrate how to create a map visual on the dataset US County Populations [data source samples.us_counties]. It produces colored regions for the measurements that correspond to the county field.

Procedure

1. Clone the US State Population - Choropleth Map visual that you developed in *State map, choropleth*.
2. Under Dimensions, select ctyname and add it to the Geo shelf.
3. Click the field to open the FIELD PROPERTIES menu, expand the Change Type option, and select the Geo Type County.
4. Alias the field ctyname as County.

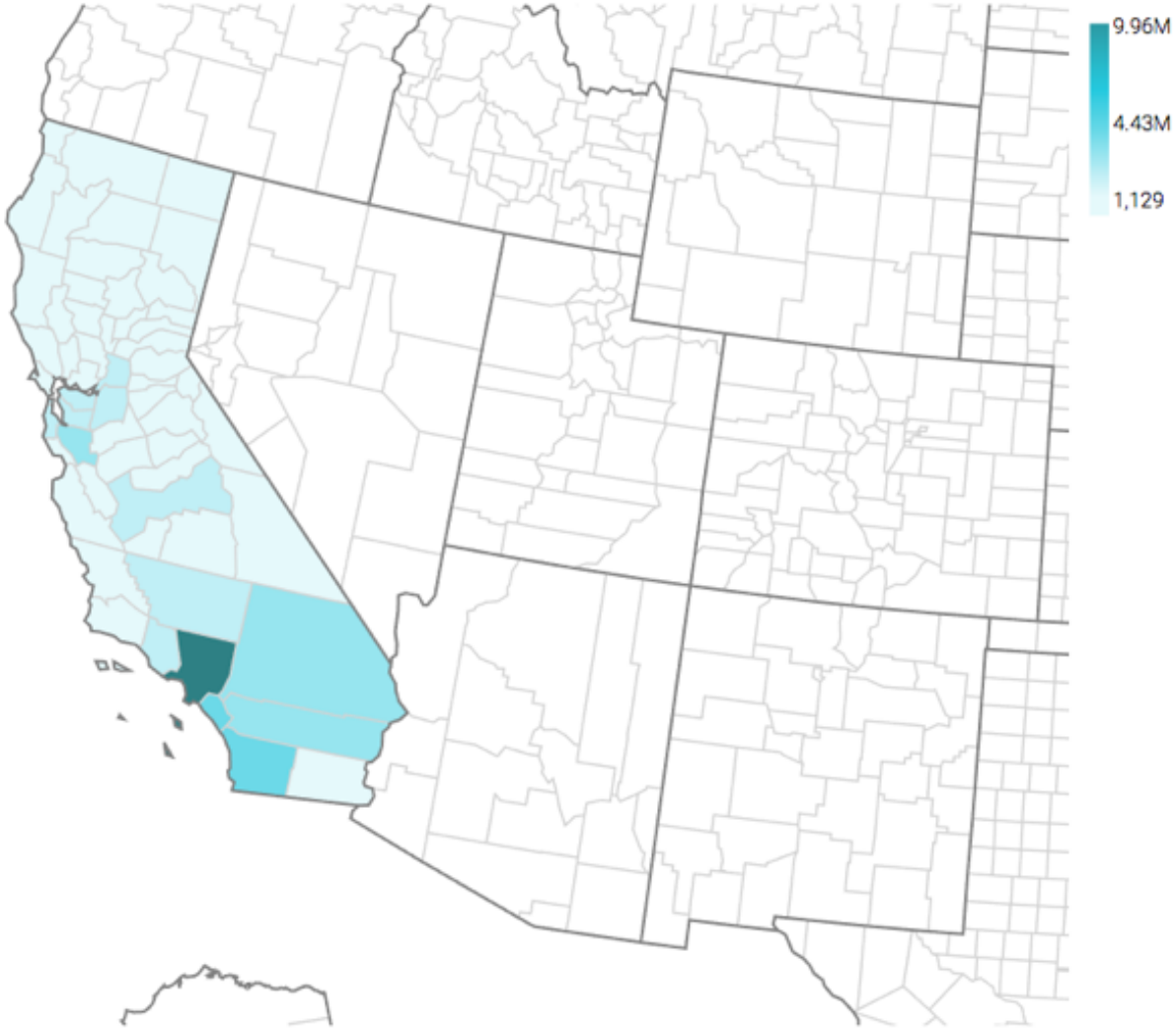
   For instructions, see *Alias*.
5. Click REFRESH VISUAL.

   The map visual appears.



6. If not already selected, select the Enable pan and Zoom and Zoom to active area options, as described in *Customizing zoom*.
7. Under Dimensions, select stname and add it to the Filters shelf. When the Filter for stname window modal appears, select California in the Values tab, and click APPLY.
8. Click REFRESH VISUAL.

   The map automatically zooms in to show detail of California.



9. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'US County Population - Choropleth Map'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

10. At the top left corner of the Dashboard Designer, click SAVE.

US county map with bubbles

Cloudera Data Visualization Map visuals let you plot measurements, as bubbles, over US County regions. Bubble maps display up to two measurements simultaneously: as the color of the bubble, and as the relative size of the bubble.

About this task

A county map of the United States shows a comparison of measurement values across US States. The following steps extend the work completed in *State map with bubbles* and demonstrate how to create a map visual on the dataset US County Populations [data source samples.us_counties]. This produces a map with bubble marks that correspond to the total population of each county.

Procedure

1. Clone the visual that you developed in *Country map, choropleth*.

   For instructions, see *Cloning a visual*.

2. Under Dimensions, select ctyname and add it to the Geo shelf.
3. Click the field to open the Field Properties menu, expand the Change Type option, and select the Geo Type County.
4. Alias the field ctyname as County.

   For instructions, see *Alias*.
5. If the Size shelf is empty, add a measure to it.

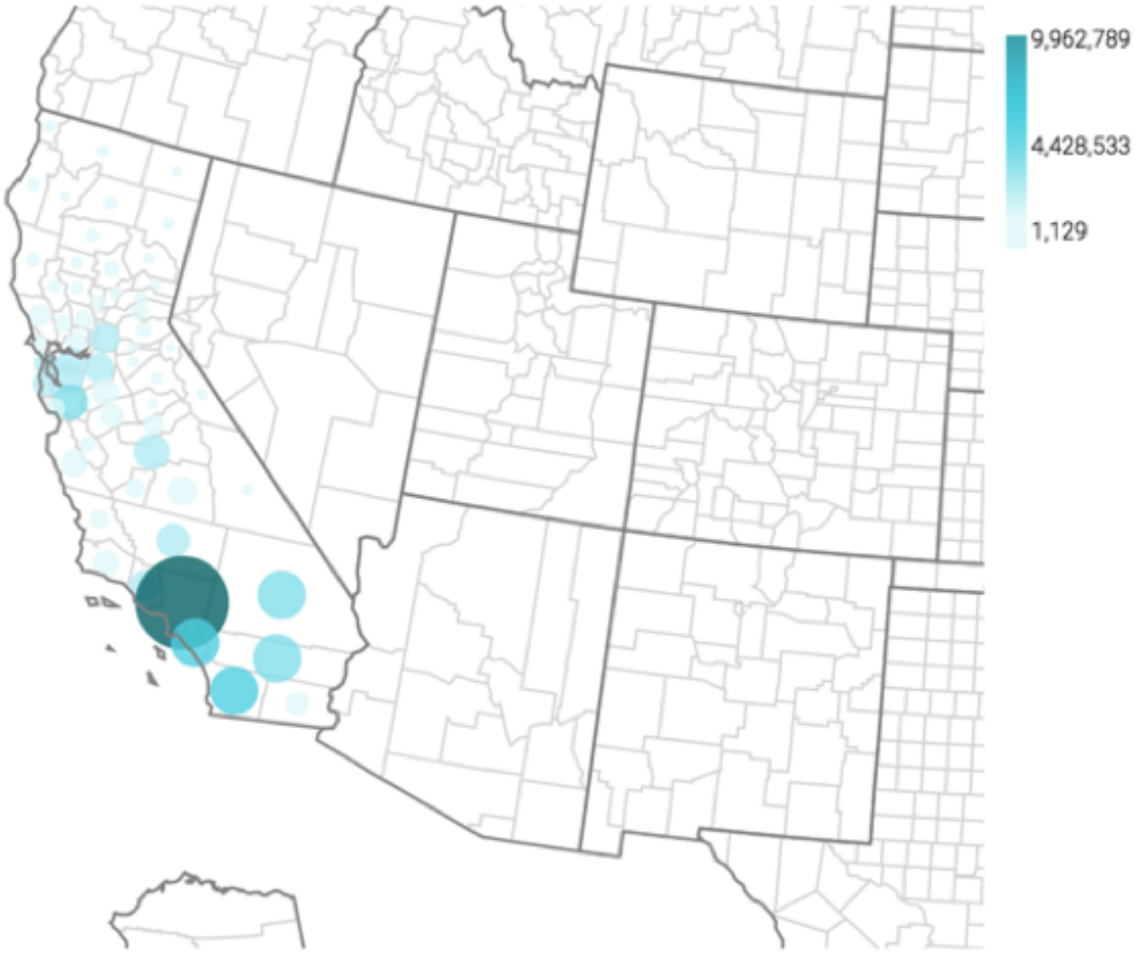
   In this example the goal is to show the relative size of the Hispanic population across the country, so the following expression is used in the Enter/Edit Expression editor:

   ```
   sum([h_female]+[h_male])
   ```

   For more information, see *Enter or edit expressions*.
6. Alias all the fields.

   For instructions, see *Alias*.
7. Change the color palette or simply reverse it, as described in *Change color palette of visuals* and *Reverse color palette of visuals*.
8. Click REFRESH VISUAL.

   The map visual appears. Note that if you have Zoom enabled, the map automatically zooms in to show detail of California. For more information, see *Customizing zoom*.



9. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

In this example, the title is changed to 'US County Population - Bubbles'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

10. At the top left corner of the Dashboard Designer, click SAVE.

US zipcode map with bubbles

In Cloudera Data Visualization Map visuals, you can plot measurements based on zipcode data. Bubble maps display up to two measurements simultaneously: as the color of the bubble, and as the relative size of the bubble.

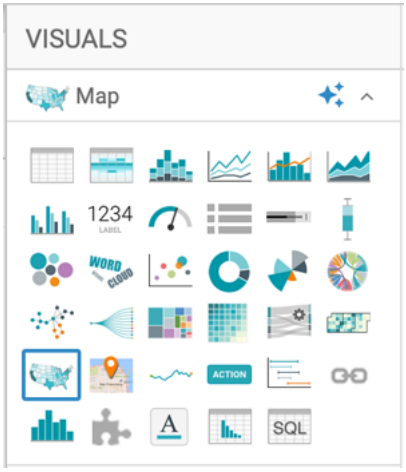
About this task

A zip code map of the United States shows a comparison of measurement values across US States. Zipcode maps only work with bubble options. The Income Tax Returns dataset is used in this example, which is built on the tax_returns.csv datafile that contains statistical information about Personal Income Tax Returns for 2016 (zip_tax_returns_2016.csv), summarized by zipcode.

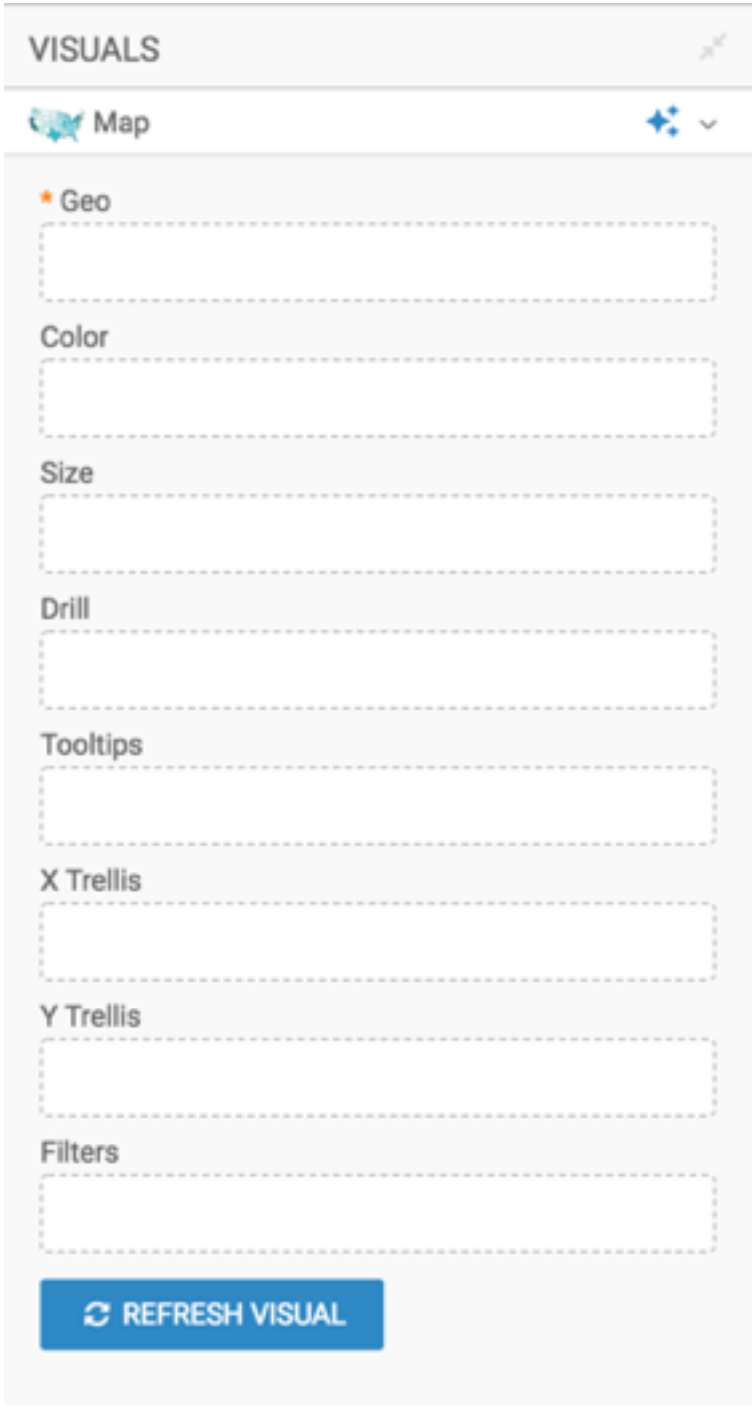
Procedure

1. Start a new visual based on the Income Tax Returns dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Map.

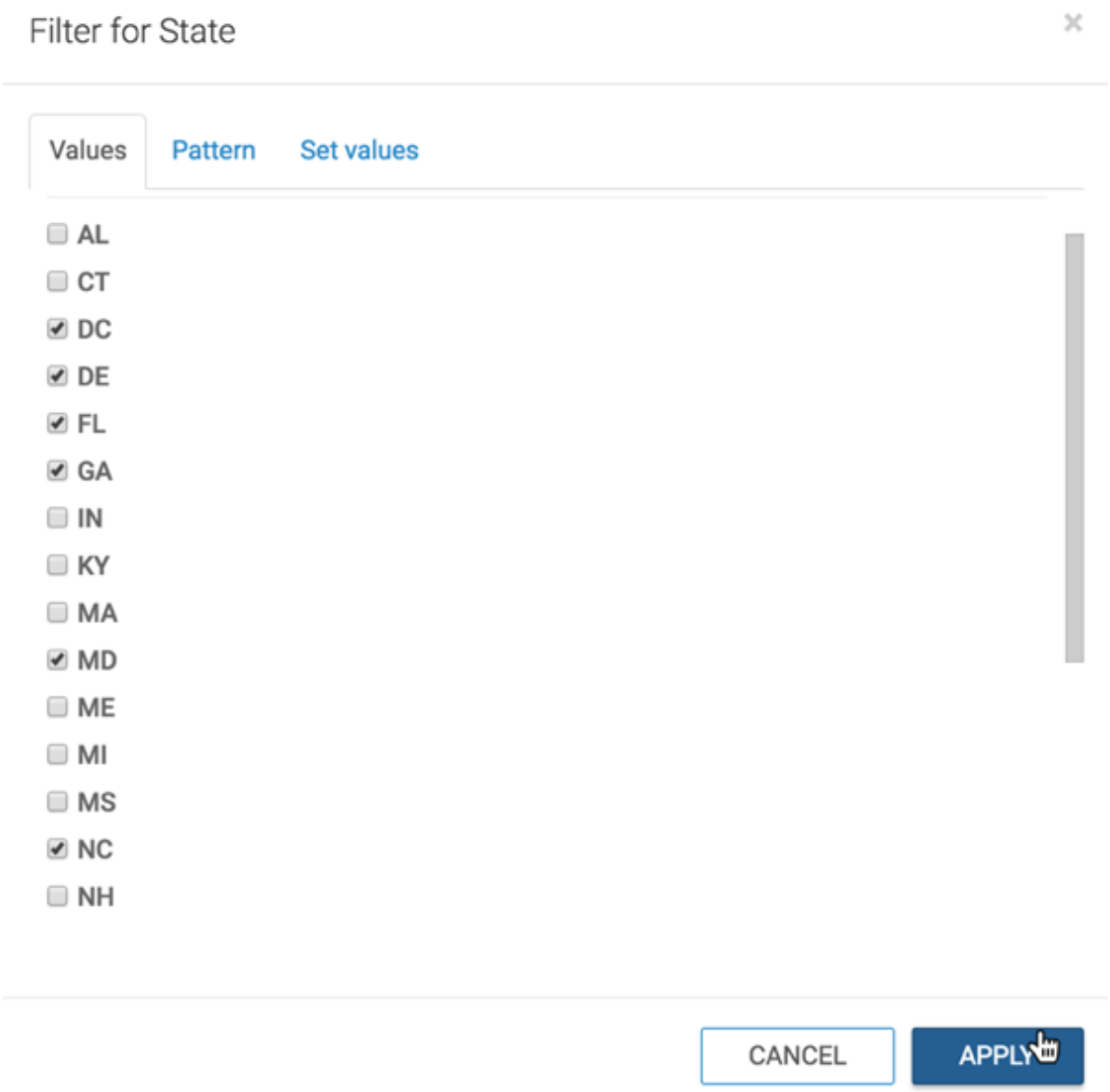


The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



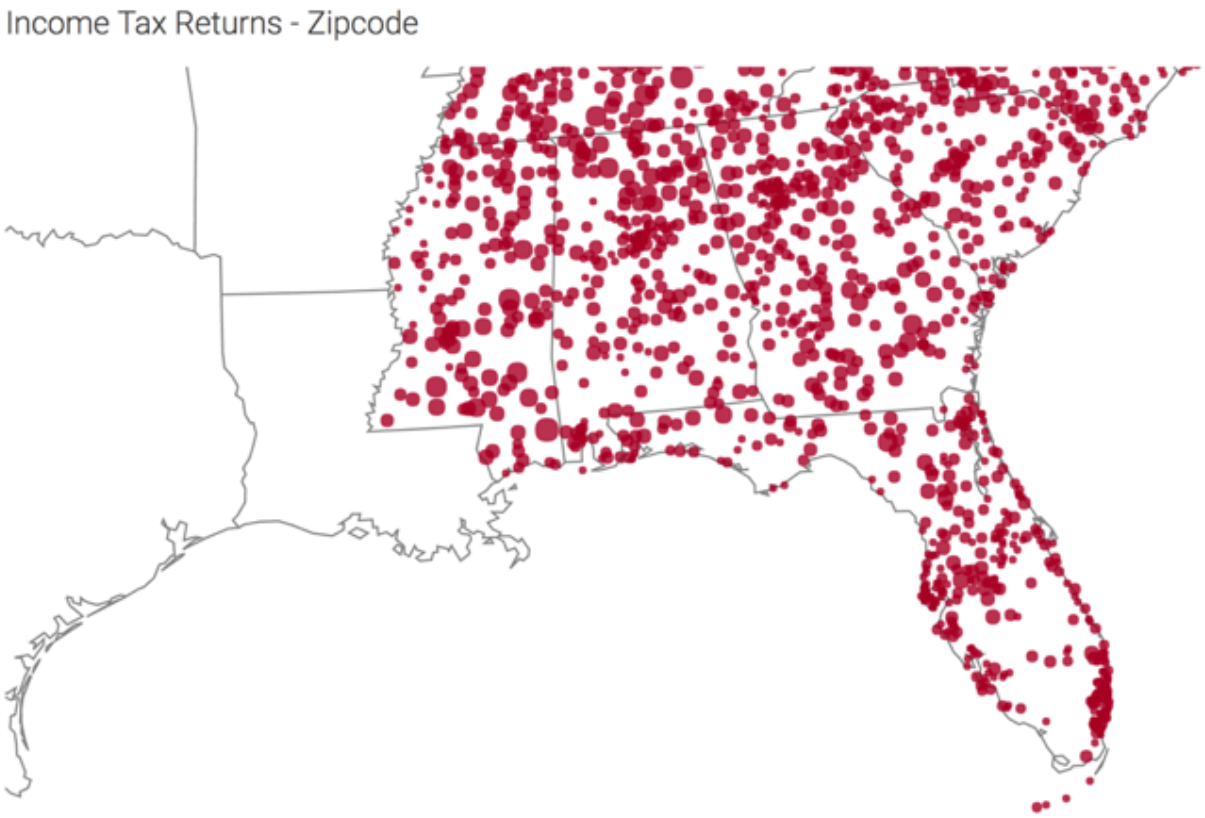
3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Select Zipcode and add it to the Geo shelf.
      The field is automatically accepted as a Geo Type.
   b) Apply the display format for integers, removing the 1000s separator.
      For instructions, see *Integer display format*.
   c) In the DATA menu, enter the term 'schf' in the Search box, and add the field to the Size shelf.
   d) Alias the field as Farm Returns.
   e) Select State and add it to the Filters shelf. When the Filter for State window modal appears, select the South Atlantic region states under the Values tab: DC (District of Columbia), DE (Delaware), FL (Florida), GA

(Georgia), MD (Maryland), NC (North Carolina), SC (South Carolina), VA (Virginia), and WV (West Virginia). Click Apply.



4. Adjust the size range of the marks to 1-3, as described in *Changing the mark size range*.
5. Change the color palette or simple reverse it, as described in *Change color palette visuals* and *Reverse color palette visuals*.
6. Enable panning and zooming option, as described in *Customizing zoom*.
7. Click REFRESH VISUAL.

   The map visual appears. You can adjust the size and focus of the visual to show the area of interest.



8. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'Income Tax Returns - Zipcode'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
9. At the top left corner of the Dashboard Designer, click SAVE.

Latitude and longitude map with bubbles

Cloudera Data Visualization Map visuals support measurement plots over a geographic map, based on latitude and longitude. Bubble maps display up to two measurements simultaneously: as the color of the bubble, and as the relative size of the bubble.

About this task

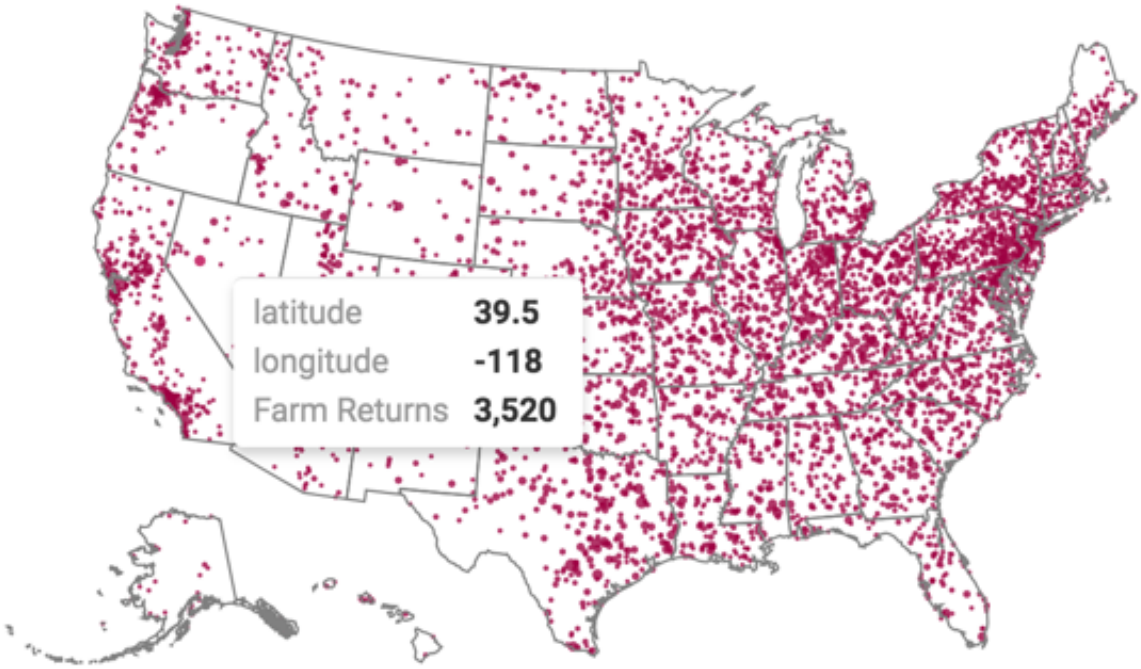
A geographic coordinate map of the United States, based on degrees of latitude and longitude, shows a comparison of measurement values across US States. The following steps extend the work completed in *US zipcode map with bubbles*, and demonstrate how to create a map visual on the Income Tax Returns dataset. It produces maps with bubble marks that correspond to the relative size of the measurement at each latitude and longitude pair.

Before starting this exercise, supplement the Income Tax Returns dataset by adding data from zipcode-lat-long.csv into your database, and making a simple joins between zip_tax_returns_2016.zipcode and zipcode_lat_long.zipcode. For instructions, see *Data model: joins*.

Procedure

1. Clone the Income Tax Returns - Zipcode visual.
2. Change the fields on the shelves:
   a) Remove all the fields from the Filters shelf.
   b) Remove the field Zipcode from the Geo shelf.
   c) In the DATA menu, enter the term 'latitude' in the Search box, and add the Latitude field to the Geo shelf.
   d) Search for term 'longitude', and add the Longitude field to the Geo shelf.

3. Open the Settings menu, expand the Map menu, and select the United Sates Map View option.
   For more information, see *Map view*.
4. Click REFRESH VISUAL.
   The map visual appears. You can move the visual (pan), zoom in and out, and see the details in the Tooltips text.



5. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.
   In this example, the title is changed to 'Income Tax Returns - Lat and Long'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
6. At the top left corner of the Dashboard Designer, click SAVE.

Choropleth DMA map

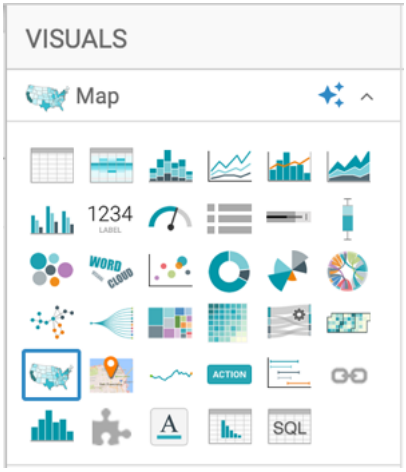
In Cloudera Data Visualization Map visual, you can visualize measurements on a map of DMA regions. A DMA map shows various measurements plotted relative to Nielsen Designated Market Area Regions.

About this task

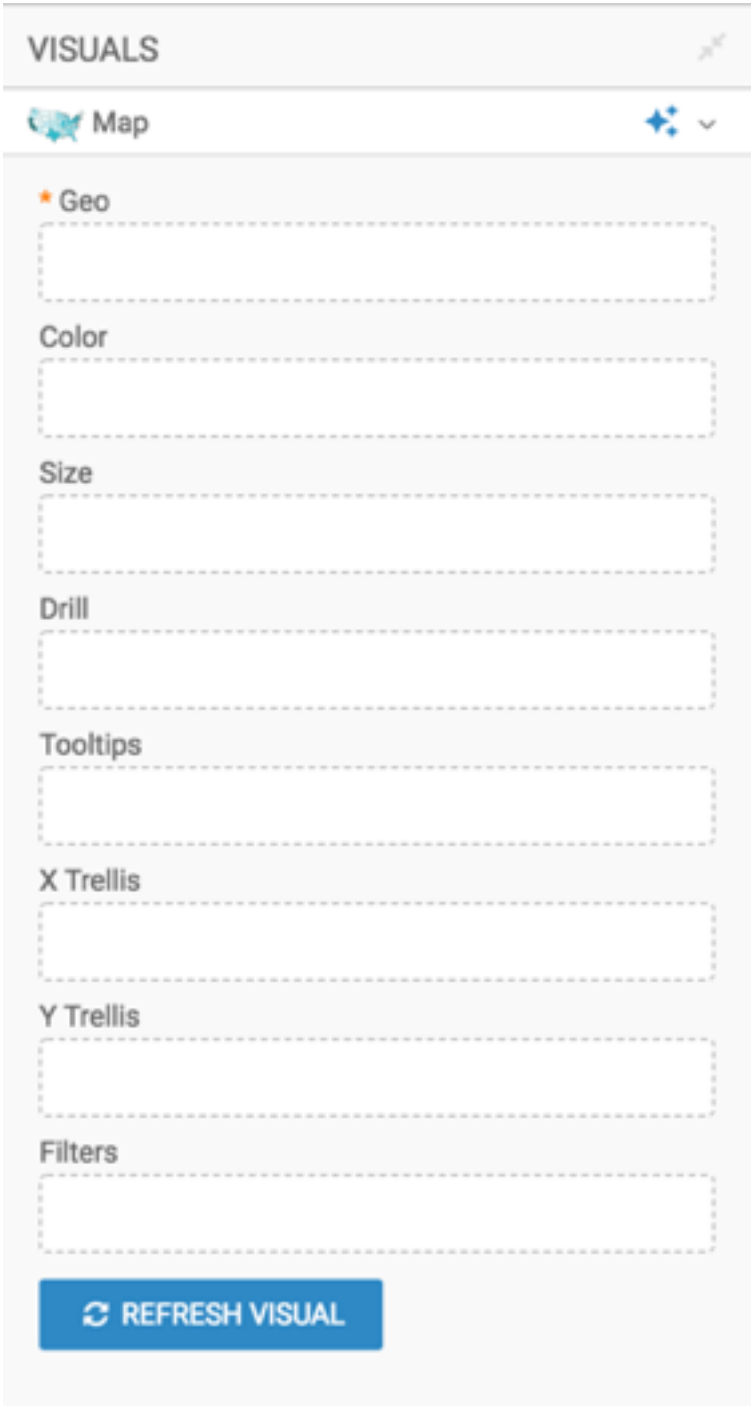
In this example, the DMA TV Home Viewing is used, which is built on the dma-tv-home-viewing.csv datafile.

Procedure

1. Start a new visual based on the DMA TV Home Viewing dataset.
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Map.

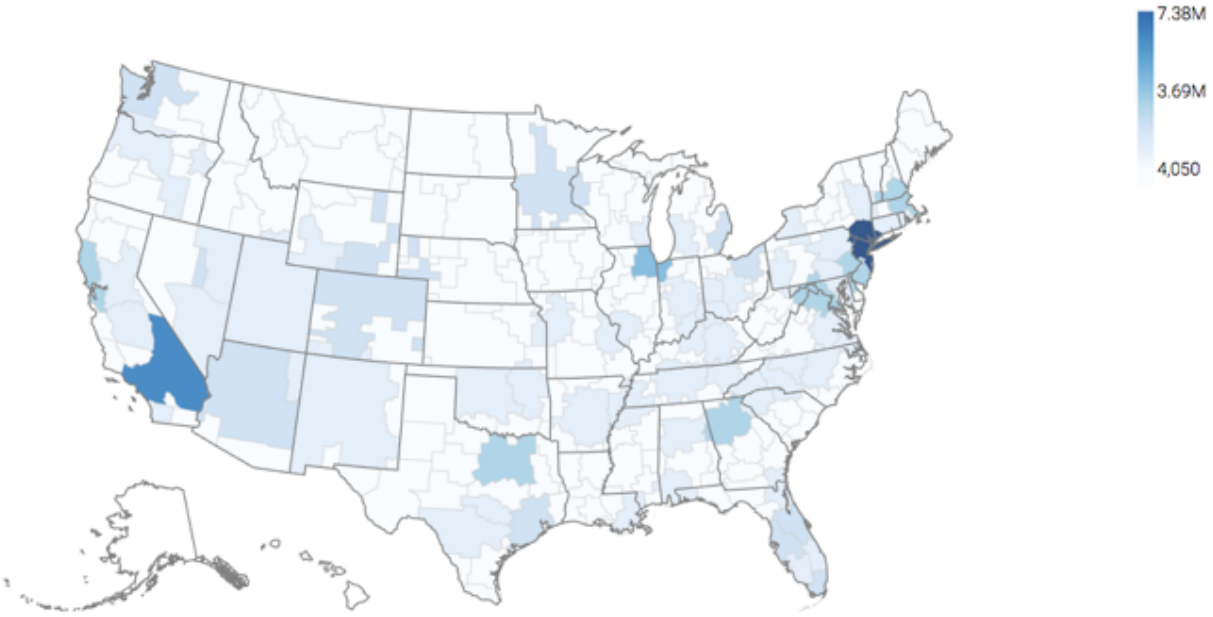


The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   - Under Dimensions, select dma and add it to the Geo shelf.
   - Under Measures, select homes and add it to the Color shelf.
4. Alias the field dma as DMA, and sum(homes) as Number of Homes.
   For instructions, see *Alias*.
5. Add the dimension dma_name to the Tooltips shelf, and then alias it as DMA Name.
6. Enable panning and zooming option, as described in *Customizing zoom*.
7. Change the color palette, as described in *Change color palette visuals*.
8. Click REFRESH VISUAL.

The map visual appears. You can move the visual (pan), zoom in and out, and see the details in the Tooltips text.



9. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'US DMA TV Viewing Map'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

10. At the top left corner of the Dashboard Designer, click SAVE.

Choropleth Canadian province map

Cloudera Data Visualization Map visuals support choropleth mapping of measurements over a map of Canadian Provinces and Territories.

About this task

A map of Canada shows a comparison of measurement values among Canadian Provinces and Territories. The following steps demonstrate how to create a new choropleth map of Canada. This example uses the Canadian Census dataset, constructed from the canada_census_population_dwellings.csv and ca-postal-codes.csvfiles, and joined of the fields fsa and Postal Code, respectively. For more information, see *Canadian census*, *Adding data*, and *Creating datasets*.

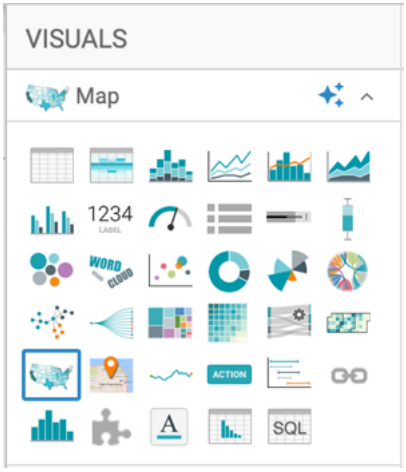
Tip: Remember that to map province information, the dimension must be recognized as a Geo Type. For more information, see *Specifying geographic fields*.

Procedure

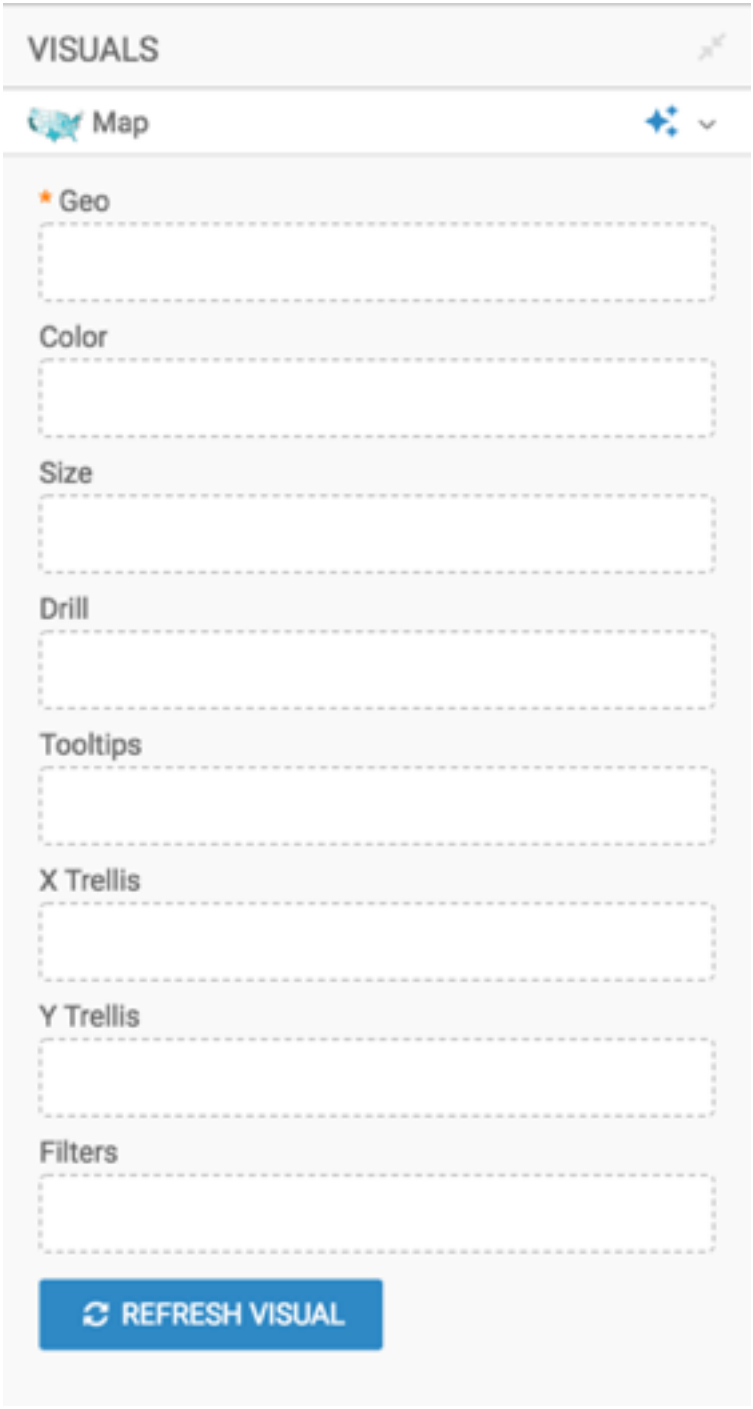
1. Start a new visual based on the Canadian Census dataset.

   For instructions, see *Creating a visual*.

2. In the VISUALS menu, find and click Map.



The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the DATA menu:
   a) Under Dimensions, select Province from the table ca_postal_codes and add it to the Geo shelf.
   b) Under Measures, select Population 2011 from table canada_census_population_dwellings and add it to the Colors shelf.
4. Specify the use of Canada Map View in the Settings Map interface.
   For instructions, see *Map view*.
5. Click REFRESH VISUAL.
6. Alias the fields as necessary; see *Alias*.
7. Enable panning and zooming option, as described in *Customizing zoom*.
8. Change the color palette, as described in *Changing color palette of visuals*.

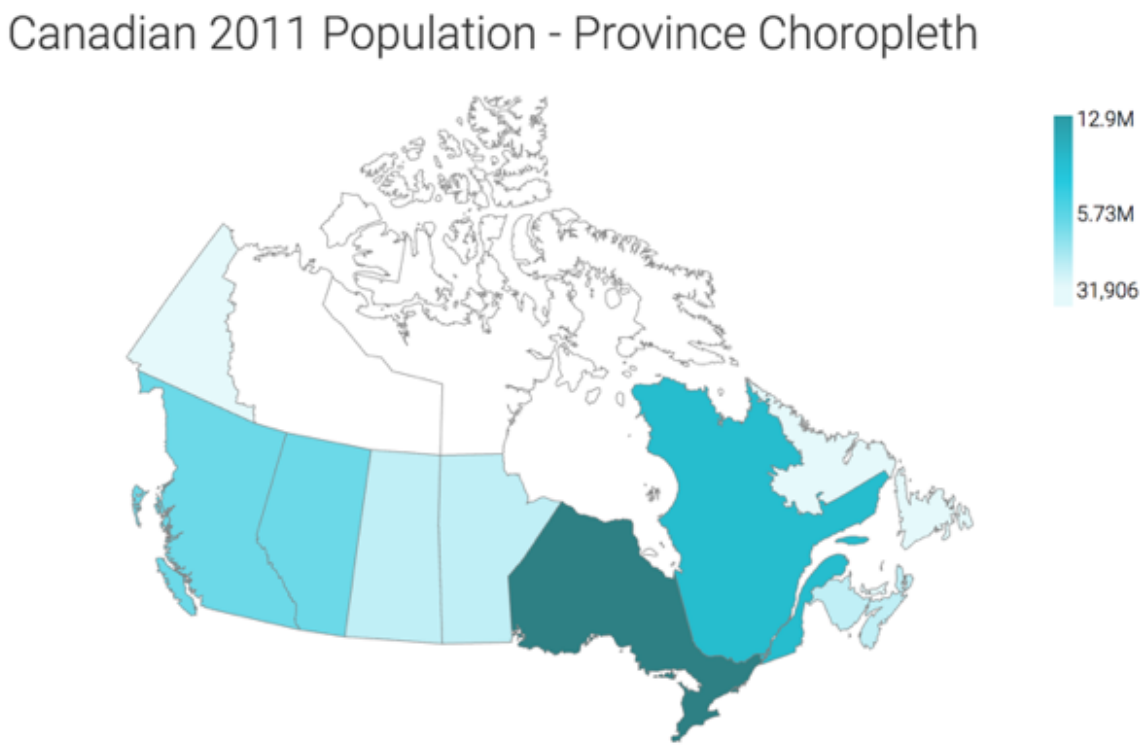
9. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'Canadian 2011 Population - Province Choropleth'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

10. At the top left corner of the Dashboard Designer, click SAVE.

Results

Your visual should look something like this:



Canadian province map, bubbles

Cloudera Data Visualization Map visuals enable you to show measurements, as bubbles, over a map of Canadian Provinces and Territories. Bubble maps display up to two measurements simultaneously: as the color of the bubble, and as the relative size of the bubble.

About this task

A map of Canada shows a comparison of measurement values among Canadian Provinces and Territories. The following steps demonstrate how to create a new choropleth map of Canada. This example uses the Canadian Census dataset, constructed from the canada_census_population_dwellings.csv and ca-postal-codes.csvfiles, and joined of the fields fsa and Postal Code, respectively. For more information, see *Canadian census*, *Adding data*, and *Creating datasets*.

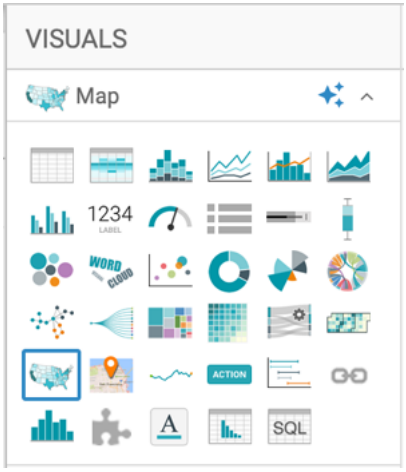
Tip: Remember that to map province information, the dimension must be recognized as a Geo Type. See *Specifying geographic fields*.

Procedure

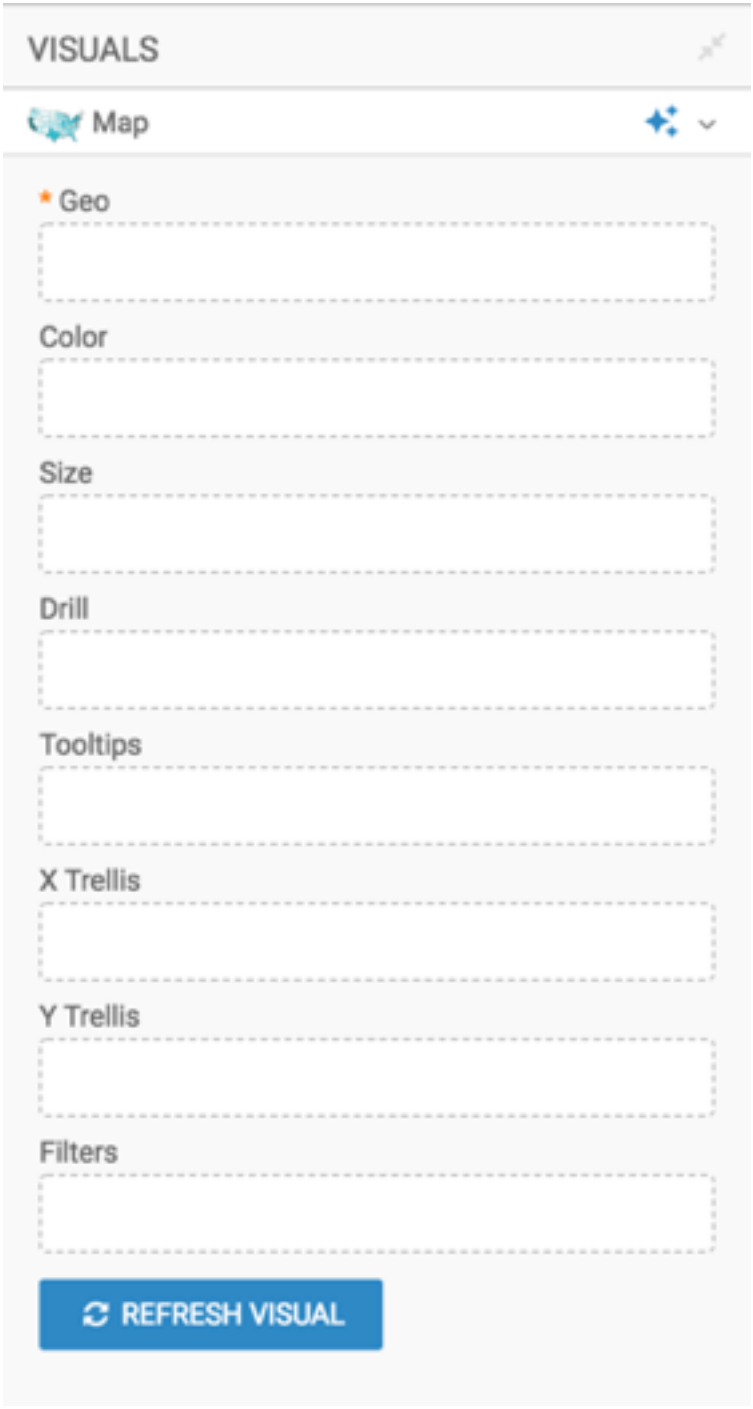
1. Start a new visual based on the Canadian Census dataset.

   For instructions, see *Creating a visual*.

2. In the VISUALS menu, find and click Map.



The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the DATA menu:
   a) Under Dimensions, select Province from the table ca_postal_codes, and add it to the Geo shelf.
   b) Under Measures, select Population 2011 from table canada_census_population_dwellings, and add it to the Measures shelf.
4. Specify the use of Canada Map View in the Settings Map interface.
   For instructions, see *Map view*.
5. Click REFRESH VISUAL.
6. Change the display of marks to bubbles as described in *Displaying bubbles on maps*.
7. Change the mark size range to 1-30 as described in *Changing the mark size range*.
8. You can change or simply reverse the color palette to suit your needs.

For instructions, see *Change color palette of visuals* and *Reverse color palette of visuals*.

9. Alias the fields as necessary.

   For instructions, see *Alias*.

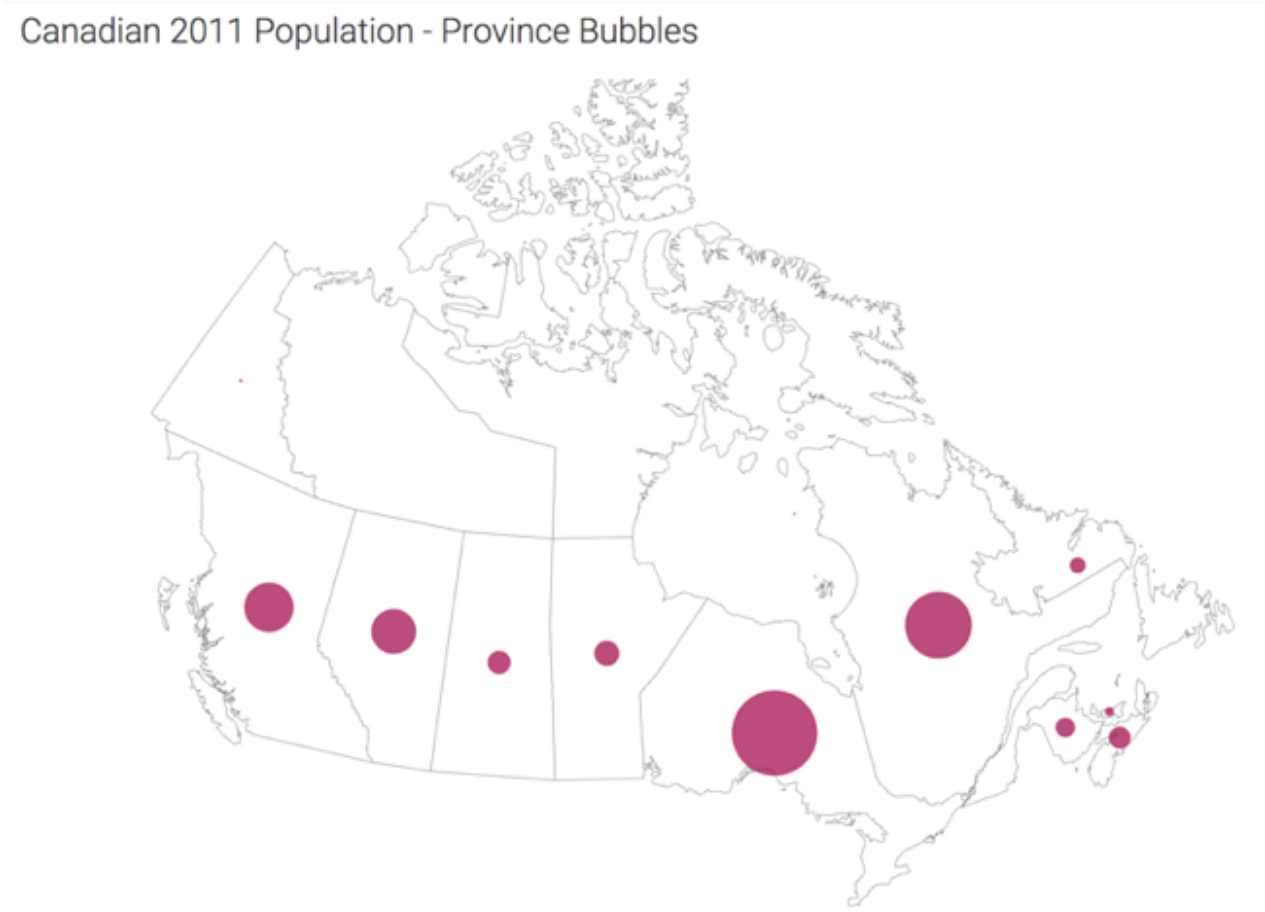
10. Enable panning and zooming option, as described in *Customizing zoom*.
11. Change the color palette, as described in *Change color palette of visuals*.
12. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

    In this example, the title is changed to 'Canadian 2011 Population - Province Bubbles'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

13. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

Your visual should look something like this:



Choropleth Canadian postal code map

Cloudera Data Visualization Map visuals support choropleth mapping of measurements over a map of Canadian FSA and Postal Codes.

About this task

A map of Canada shows a comparison of population density, at the level of the FSA (1st 3 digits of the Postal Code). The following steps demonstrate how to create a new choropleth map of Canada. This example uses the Canadian Census dataset, constructed from the canada_census_population_dwellings.csv and ca-postal-codes.csvfiles, and

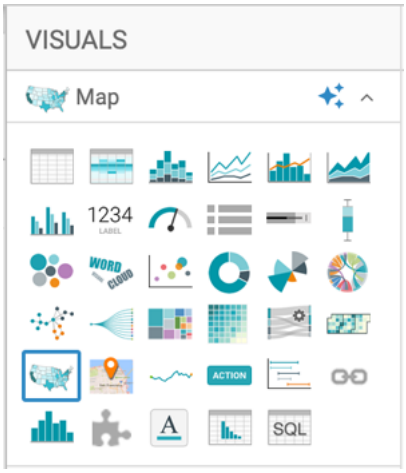
joined of the fields fsa and Postal Code, respectively. For more information, see *Canadian census*, *Adding data*, and *Creating datasets*.

Tip: Remember that to map province information, the dimension must be recognized as a Geo Type. See *Specifying geographic fields*.

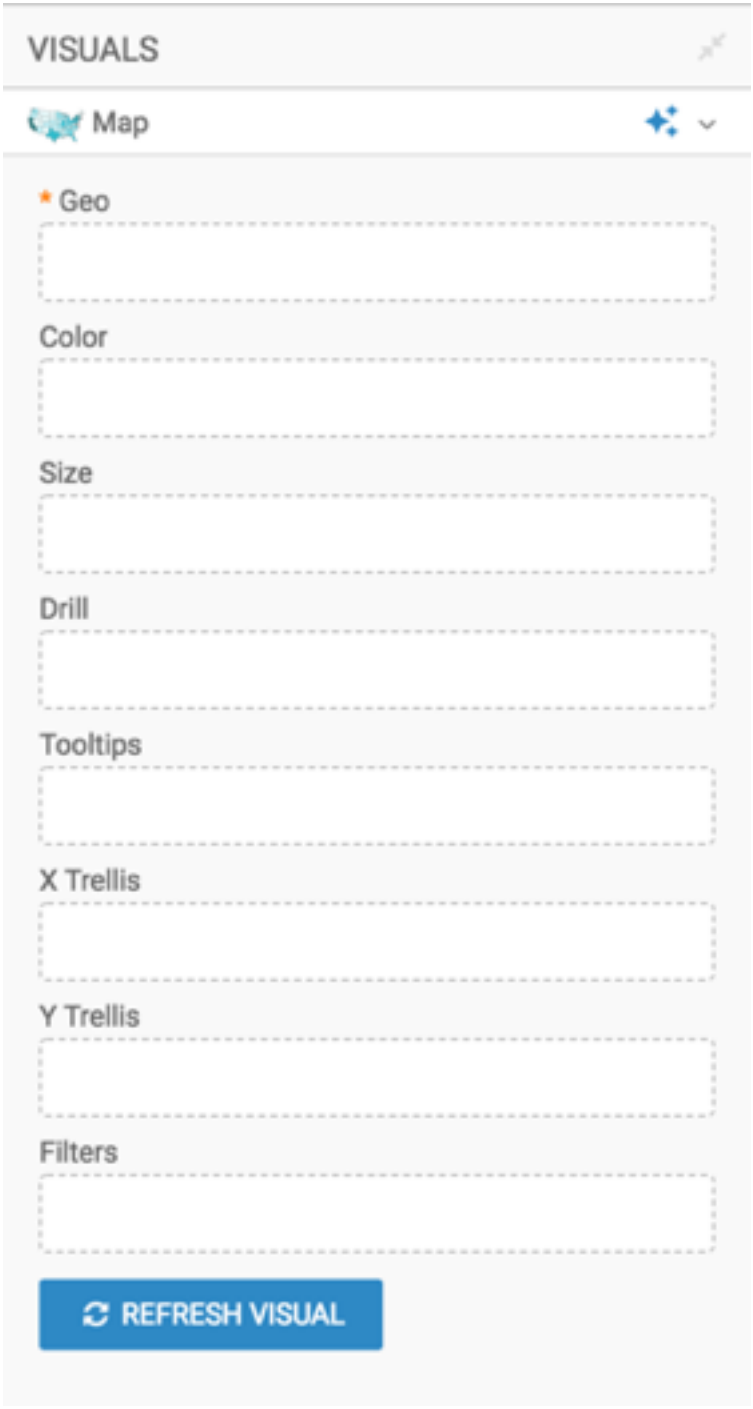
Procedure

1. Start a new visual based on the Canadian Census dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Map.



The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



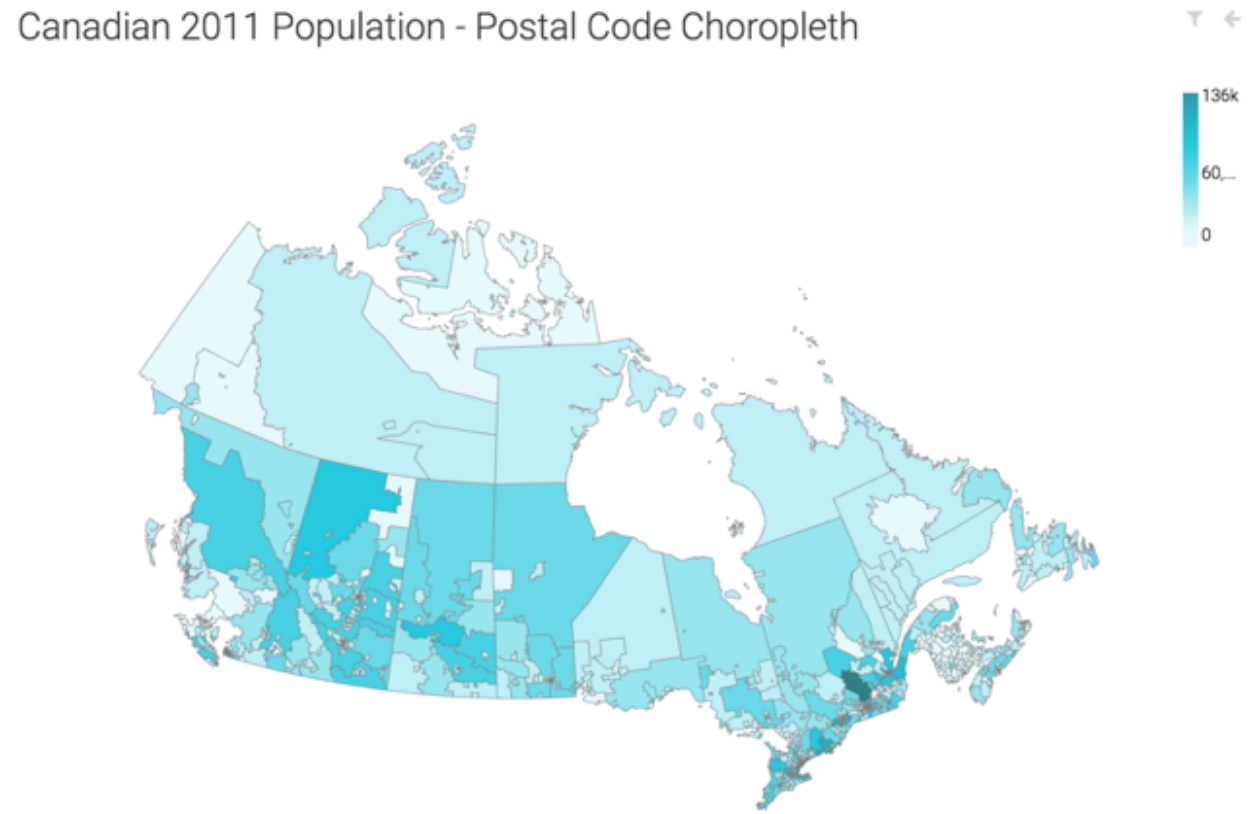
3. Populate the shelves from the DATA menu:
   a) Under Dimensions, select Postal Code from the table ca_postal_codes, and add it to the Geo shelf.
   b) Under Measures, select Population 2011 from table canada_census_population_dwellings, and add it to the Color shelf.
4. Specify the use of Canada Map View in the Settings Map interface.
   For instructions, see *Map view*.
5. Click REFRESH VISUAL.
6. Alias the fields as necessary.
   For instructions, see *Alias*.
7. Enable panning and zooming option, as described in *Customizing zoom*.

8. Change the color palette, as described in *Change color palette of visuals*.
9. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'Canadian 2011 Population - Postal Code Choropleth'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
10. At the top left corner of the Dashboard Designer, click SAVE.

Results

Your visual should look something like this:



Canadian postal code map, bubbles

Cloudera Data Visualization Map visuals support choropleth mapping of measurements over a map of Canadian FSA and Postal Codes. Map visual provides support for mapping measurements, as bubbles, over a map of Canadian FSA and Postal Codes. Bubble maps display up to two measurements simultaneously: as the color of the bubble, and as the relative size of the bubble.

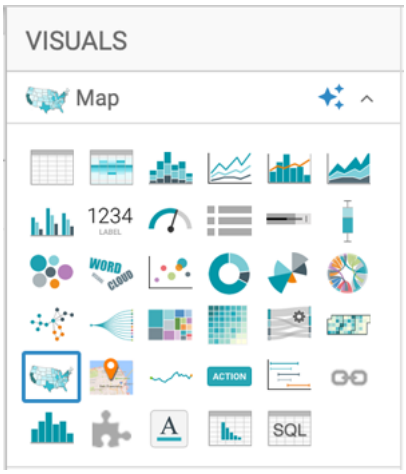
About this task

A map of Canada shows a comparison of population density, at the level of the FSA (1st 3 digits of the Postal Code). The following steps demonstrate how to create a new choropleth map of Canada. This example uses the Canadian Census dataset, constructed from the canada_census_population_dwellings.csv and ca-postal-codes.csvfiles, and joined of the fields fsa and Postal Code, respectively. For more information, see *Canadian census*, *Adding data*, and *Creating datasets*.

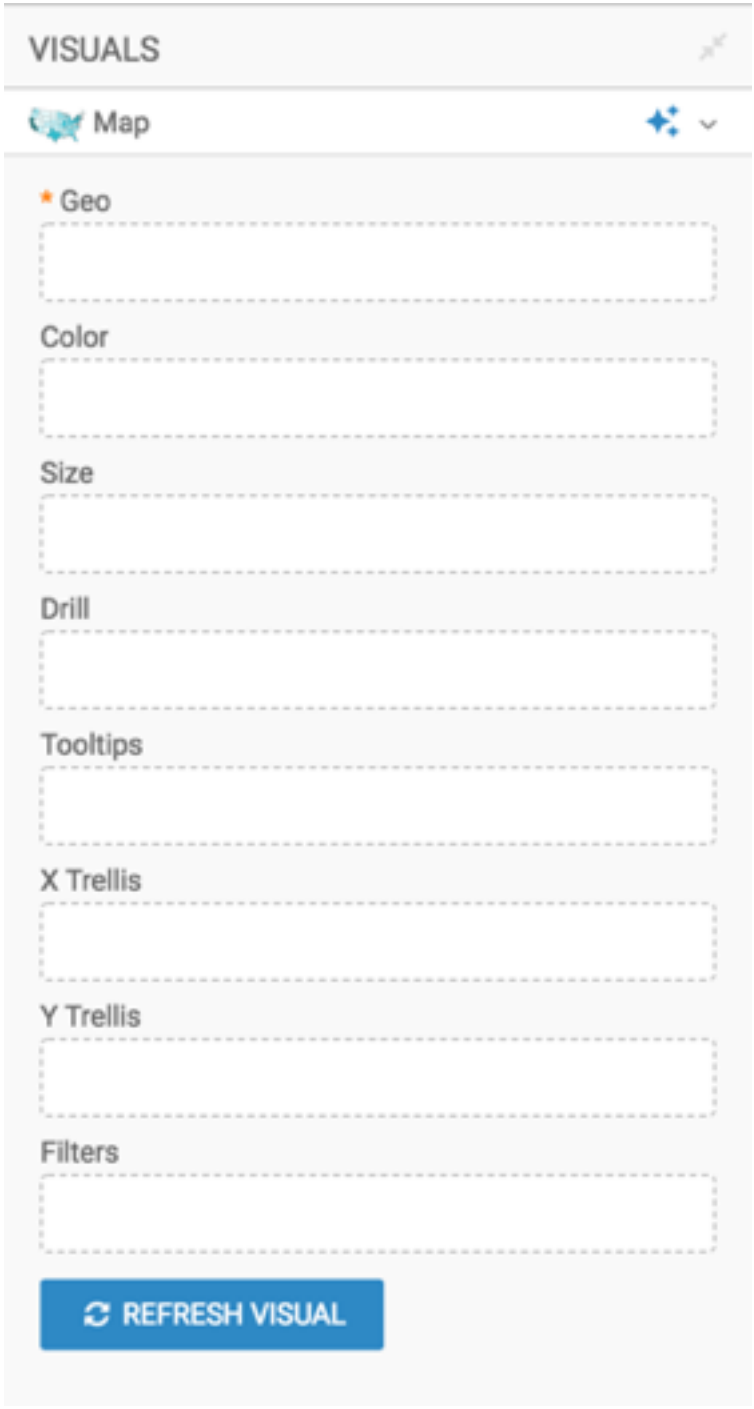
Tip: Remember that to map province information, the dimension must be recognized as a Geo Type. See *Specifying geographic fields*.

Procedure

1. In the VISUALS menu, find and click Map.



The shelves of the visual changed. They are now Geo, Color, Size, Drill, Tooltips, X Trellis, Y Trellis, and Filters. The only mandatory shelf for map visuals is Geo.



2. Populate the shelves from the DATA menu:
   a) Under Dimensions, select Postal Code from the table ca_postal_codes, and add it to the Geo shelf.
   b) Under Measures, select Population 2011 from table canada_census_population_dwellings, and add it to the Color shelf.
3. Specify the use of Canada Map View in the Settings Map interface.
   For instructions, see *Map view*.
4. Click REFRESH VISUAL.
5. Change the display of marks to bubbles as described in *Displaying bubbles on maps*.
6. Change the mark size range to 1-10 as described in *Changing the mark size range*.
7. You can change or simply reverse the color palette to suit your needs.

For instructions, see *Change color palette of visuals* and *Reverse color palette of visuals*.

8. Alias the fields as necessary.

   For instructions, see *Alias*.

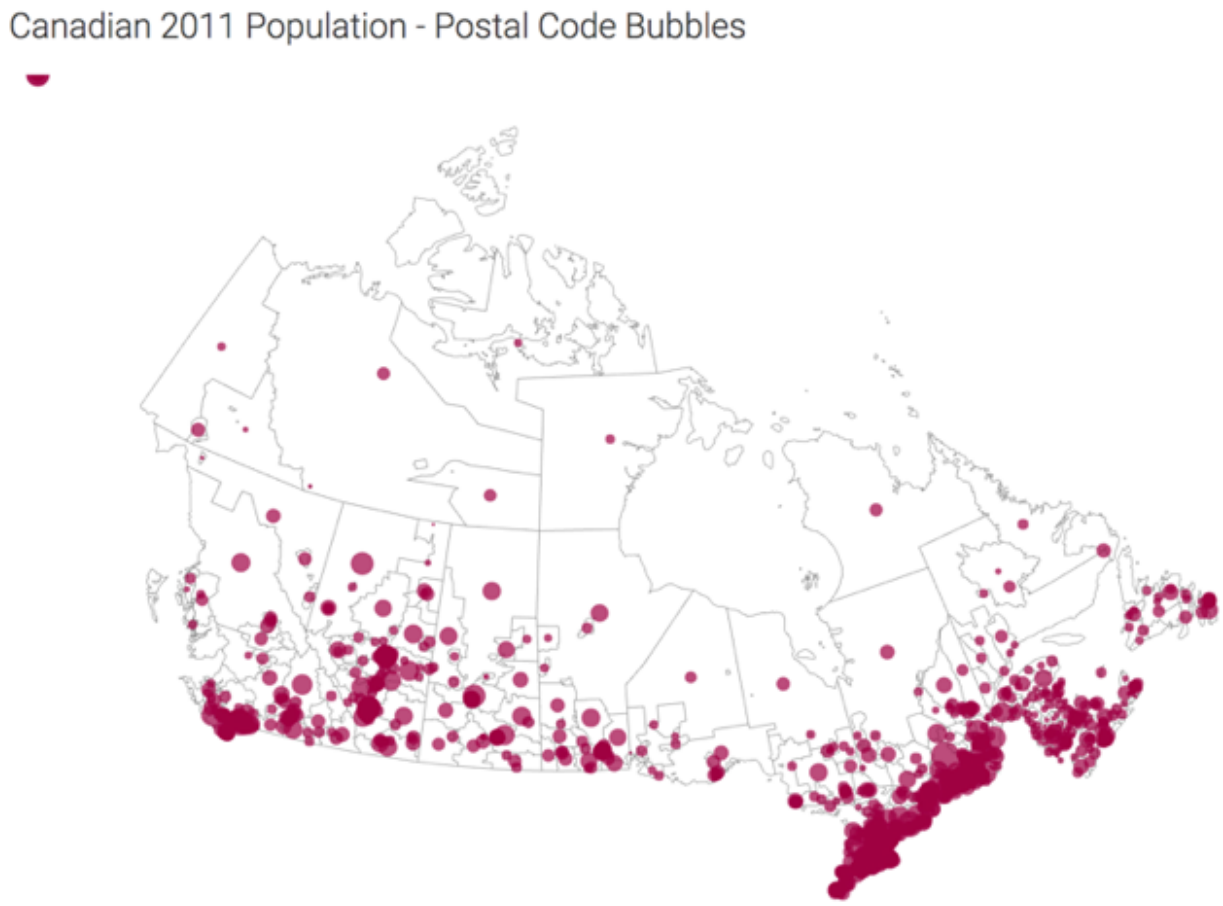
9. Enable panning and zooming option, as described in *Customizing zoom*.
10. Change the color palette, as described in *Change color palette of visuals*.
11. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

    In this example, the title is changed to 'Canadian 2011 Population - Postal Code Bubbles'. You can also add a brief description of the visual as a subtitle below the title of the visualization.

12. At the top left corner of the Dashboard Designer, click SAVE.

Results

Your visual should look something like this:



Shelves for maps

Overview of shelves for Cloudera Data Visualization Map visuals.

A Map visual has the following shelves:

Geo

Add a Geo-type field from the DATA menu to determine the location on the map.

Geo types include country, state, county, DMA, zipcode or postal code, or latitude and longitude. If the field is not automatically recognized as a qualified Geo type, select Change Type from the FIELD PROPERTIES menu.

This shelf accepts multiple fields, but plots a map with the highest possible resolution.

This is mandatory shelf.

Color

Add a Measure field from the DATA menu to aggregate values across locations.

This shelf displays only one field on the graph. Additional fields appear only in the Tooltip.

Size

Add a Measure field from the DATAmenu to aggregate values across locations.

This shelf displays only one field on the graph. Additional fields appear only in the Tooltip.

Drill

Add Dimension fields from the DATA menu to this shelf; use Geo-type fields only.

When Drill Into context menu is enabled but no fields are specified, you can drill into any dimension. Specifying the fields on this shelf limits drilling to these dimensions.

This shelf accepts multiple fields.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Interactive maps

An interactive map paints geographic details over a basic map tile layer provided by a third-party provider. Cloudera Data Visualization offers several overlay layers for data display. It is an excellent choice for displaying large amounts of geo-based information in relevant detail.

- The base Map Servers include either Google Maps or Mapbox.
- The overlay layers that display data include the Heatmap, Cluster, Circles, and Routes and pins options.

Note:

- To use this feature, an Admin user must first specify, through the Site Settings interface, the API keys granted to your organization. This is described in *Keys for interactive visuals*.
- This visual type supports zero, one, or two field aggregations on the measures shelf. If no measure is supplied, the map will paint with an assumed value of 1. With one measure, all options use that measure. With two measures, the Heatmap, Cluster and Routes & Pins use the first measure only, while the Circles display both measures.
- Remember that to place information onto a map or interactive map visual, the records must contain geographical information, such as valid latitude and longitude pairs, zipcodes, states, and so on. To learn how to cast dataset fields into usable type, see *Change type* and *Geo type*.

To learn how to use the interactive maps in their various forms, read the following topics.

Basic interactive map

Cloudera Data Visualization enables you to create a basic Interactive Map visual.

About this task

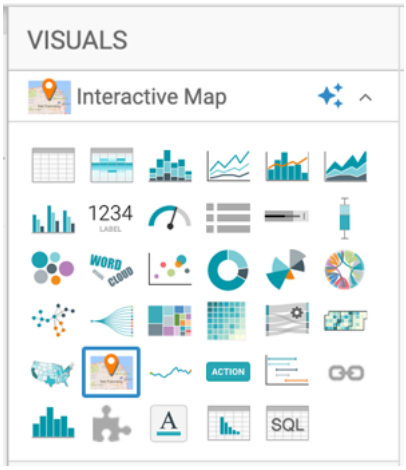
The following steps demonstrate how to create an interactive map visual on the National Geographic Features dataset. The data for this dataset comes from the United States Board on Geographic Names based on the AllStates zip file. Additionally, you can use Feature Class Definitions and State Abbreviations to supplement the dataset.

For an overview of shelves that specify this visual, see *Shelves for interactive visuals*.

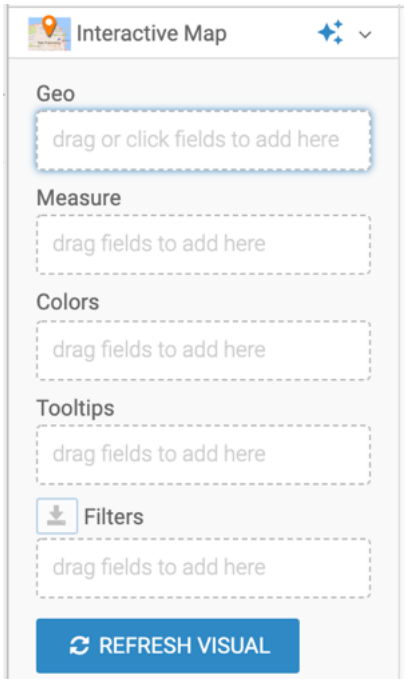
Procedure

1. Start a new visual based on the National Geographic Features dataset.

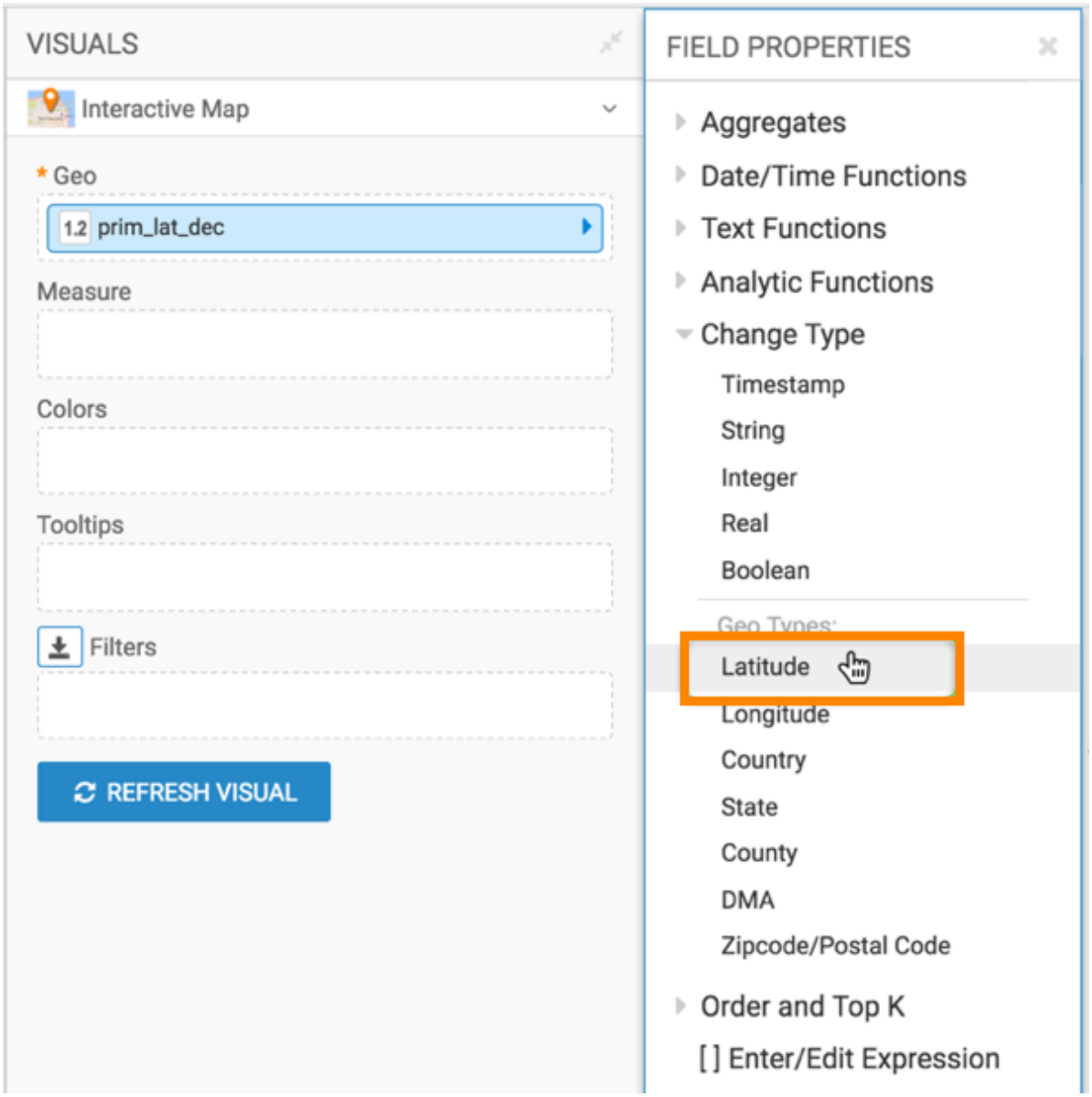
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Interactive Map.



The shelves of the visual changed. They are now Geo, Measure, Colors, Tooltips, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Under Measures, select prim-lat-dec and drag it to the Geo shelf.
   b) Cast this to the appropriate Geo Type: click the filed on the shelf, and in the FIELD PROPERTIES menu and select Change Type Latitude .



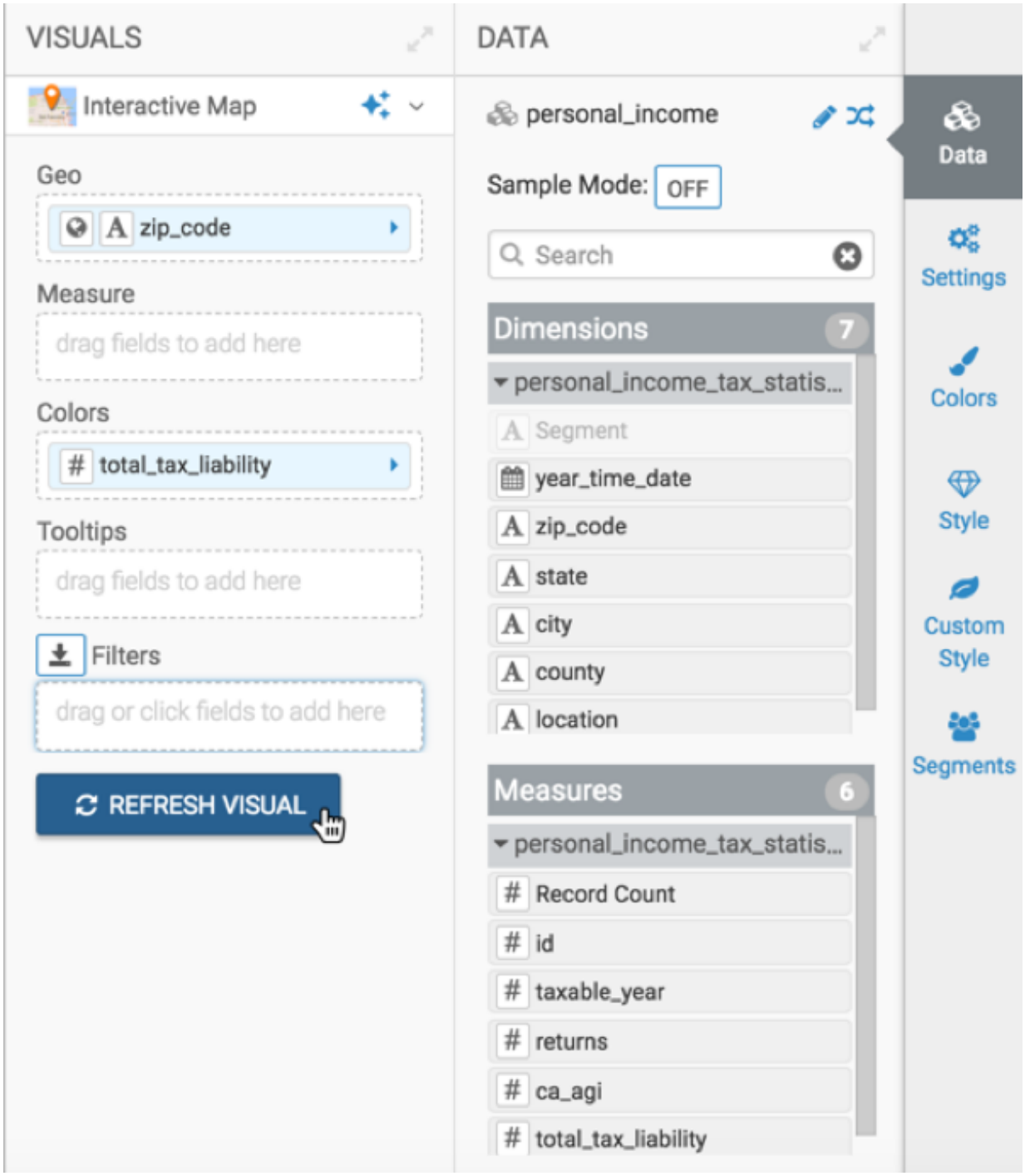
c) Under Measures, select prim_long_dec, and drag it to the Geo shelf on the main part of the screen.
d) Cast this to the appropriate Geo Type: click the filed on the shelf, and in the FIELD PROPERTIES menu and select Change Type Longitude .

The two converted measurements have a (globe) icon tag. This indicates that these fields can be 'understood' as geographic fields. For more information, see *Change type* and *Geo data type*.

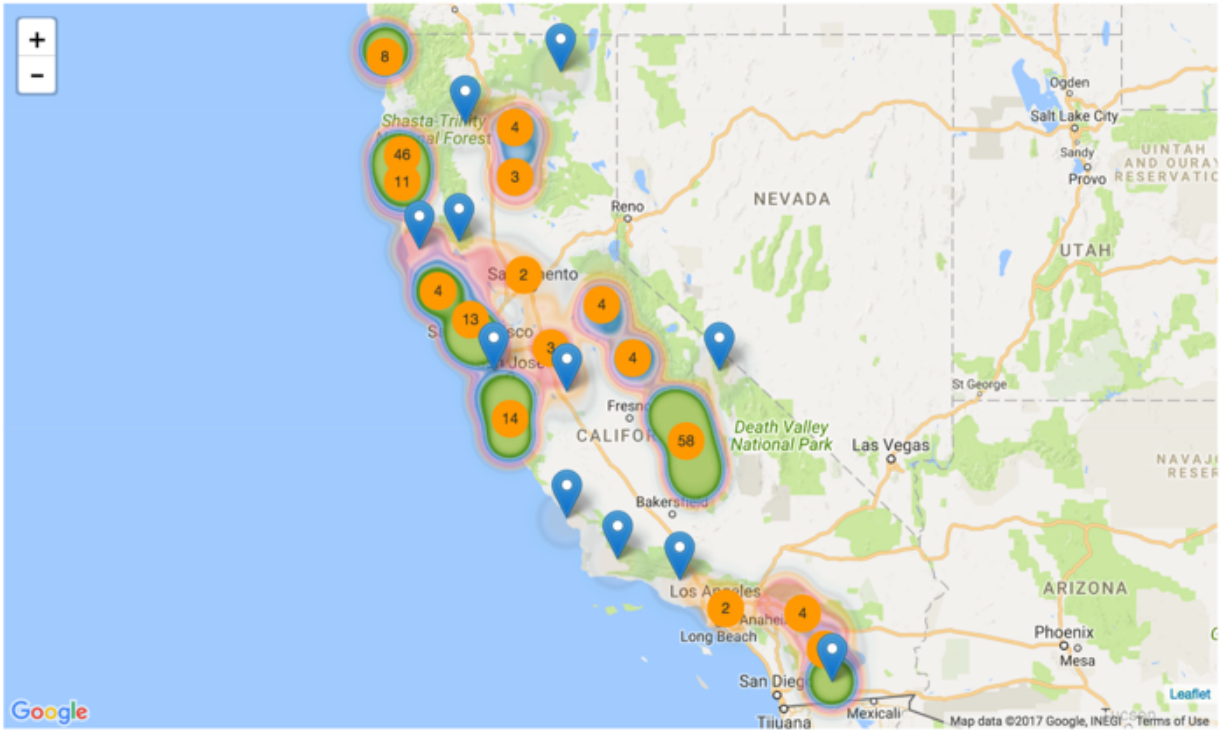
4. Under Dimensions, select feature_id, and drag it to the Measures shelf.

Note: If you leave the Measures shelf empty, the system uses the value of 1, and does not require an aggregate.

5. Change the aggregation from sum(feature Id) to count(Feature Id), by selecting Aggregates Count.
6. Add the following fields to the Filters shelf:
   - Field feature_class, set to Woods
   - Field sate_alpha, set to CA
7. Click REFRESH VISUAL.



The visual appears. In this example, you can see a Google INEGI map.

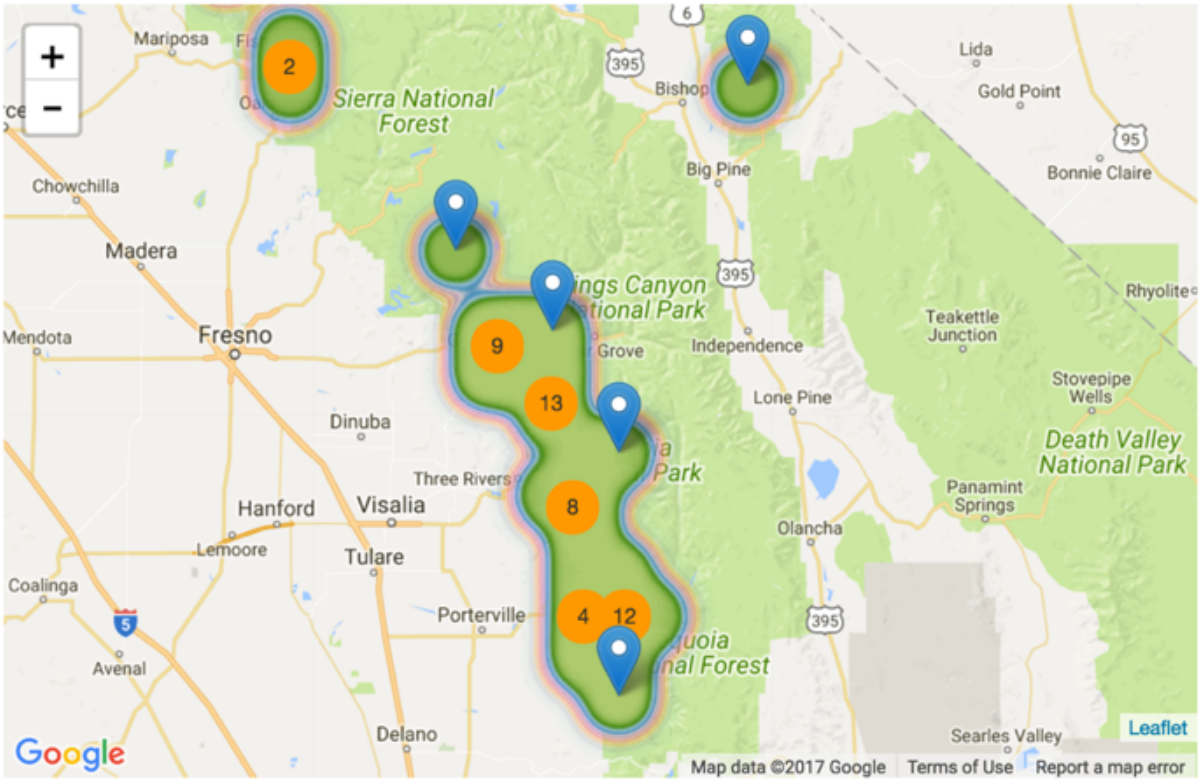


8. In this representation, there is noticeable overlap on the Heatmap. However, you can see the boundaries of each region by hovering the pointer over a section of the map.

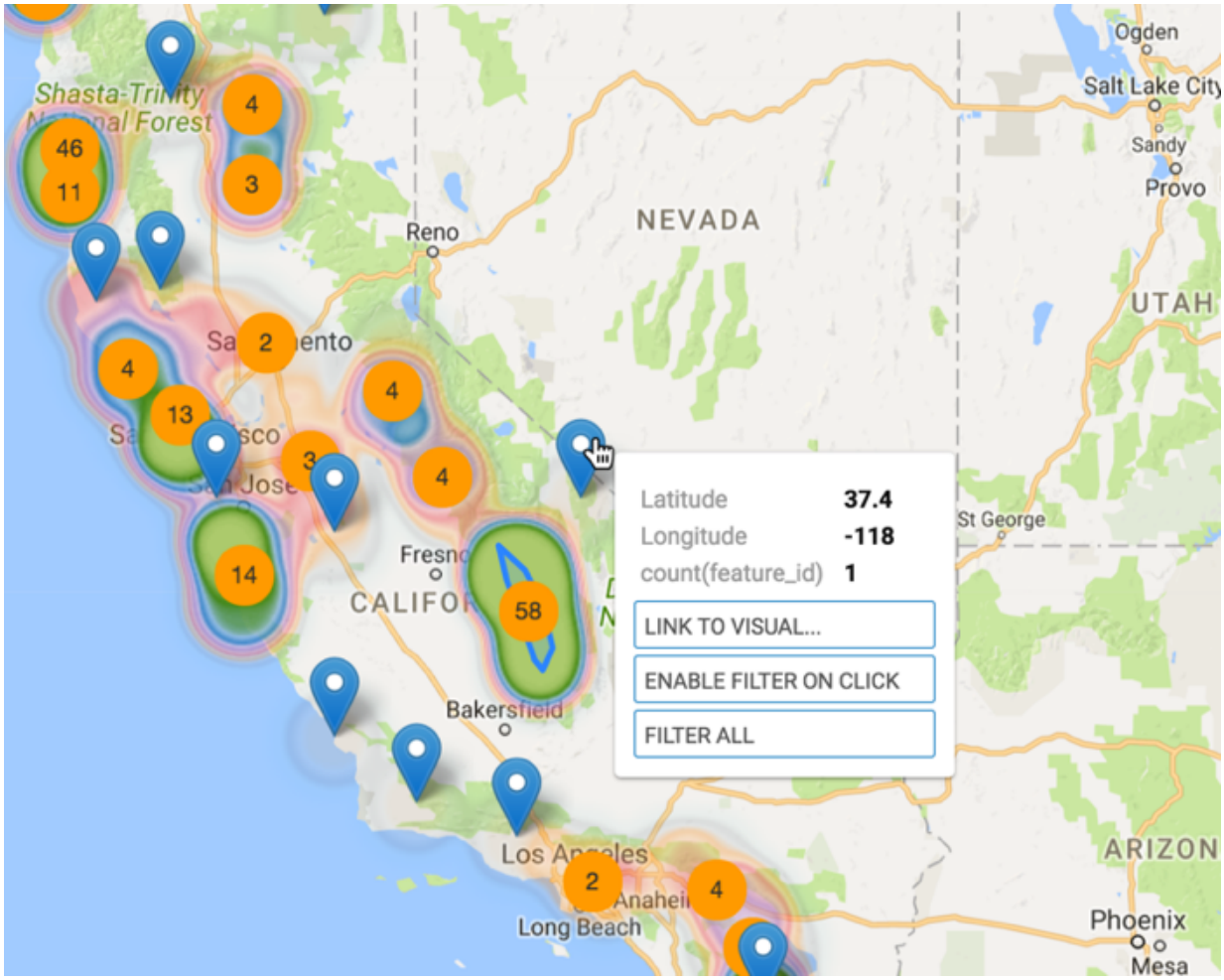
   The total count of measurements in each region is in the circle inside the region. This comes from the Cluster. Both Heatmap and Cluster are on my default.

9. To see more detail, manipulate the map through magnification changes. You can also click on the map and manually 'pan' it to the area of interest.

   The following image shows that the regions automatically adjust into smaller sections, and you can see both greater detail and more individual 'totals'.



10. The 'interactive' part of this visual is that after clicking on a region to zoom-in and reposition, to the level of an individual marker, you can enable click behavior.



11. Change the title to California Woods.

12. At the top left corner of the Dashboard Designer, click SAVE.

What to do next

There are many interesting visualization options. To change the map server, or the base map from Google to Mapbox, see *Changing the map server to interactive maps*. To configure the various options, see these topics:

- *Map server*
- *Map style*
- *Choropleth*
- *Heatmap*
- *Cluster*
- *Circles*
- *Routes and pins*

Choropleth interactive maps

In an Interactive Map visual, Cloudera Data Visualization enables you to create choropleth maps. A choropleth map shows geographical areas that are shaded in proportion to the value of the measurement that is displayed on the map.

About this task

Tip: Consider the following when building choropleth maps:

- Use choropleth maps when your data is attached to enumeration units, such as counties, countries or zip codes.
- Standardize data to show rates and ratios instead of raw data counts. For example, instead of plotting gross GDP, plot GDP per capita.
- Use choropleth maps when you have data with continuous statistical surface. For example, Population is a count and representing it in colored map may be misleading. Population density is a ratio with a continuous statistical surface (even where 0 represents unpopulated areas); so it can be used in choropleth maps.

In this example, you will create a choropleth map that colors the average tax liability in California by zip code. The choropleth Interactive Map visual is based on the Personal Income Tax Statistics by Zip Code dataset. For an overview of shelves that specify this visual, see *Shelves for interactive maps*.

Before you begin

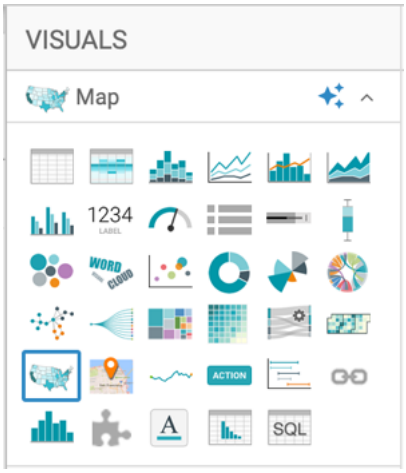
To create a choropleth map, you must specify the following for each visual:

- Custom mapbox style, which has a choropleth shape file in it. For more information, see *Adding a custom mapbox style*.
- Layer ID, which is the name of a unique layer within the Custom Mapbox style that identifies the shapes that you would like to color by a measure in your data. For more information, see *Specifying a choropleth layeer ID*.
- Geo-type as Shape ID: The dataset must have at least one field in the Geo shelf to determine the location on the map. You must change the Geo-type to Shape ID. A Shape ID is a key that joins the shapes in the Custom mapbox style and the field in the Cloudera Data Visualization dataset. This helps in identifying the measure that maps to the shape. This Geo field must be in the dataset that was used to create the custom mapbox style.

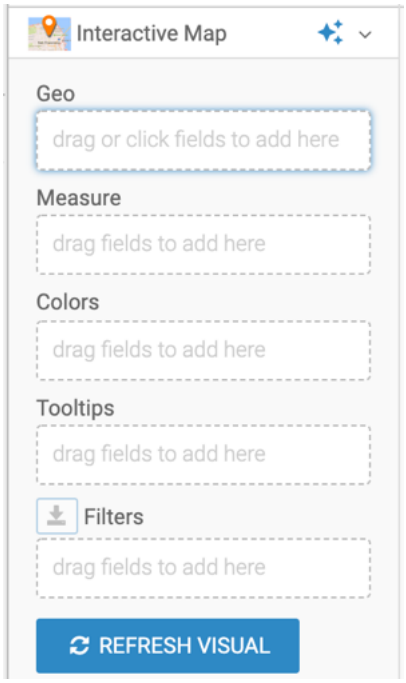
Procedure

1. Start a new visual based on the National Geographic Features dataset.

   For instructions, see *Creating a visual*.
2. In the VISUALS menu, find and click Interactive Map.



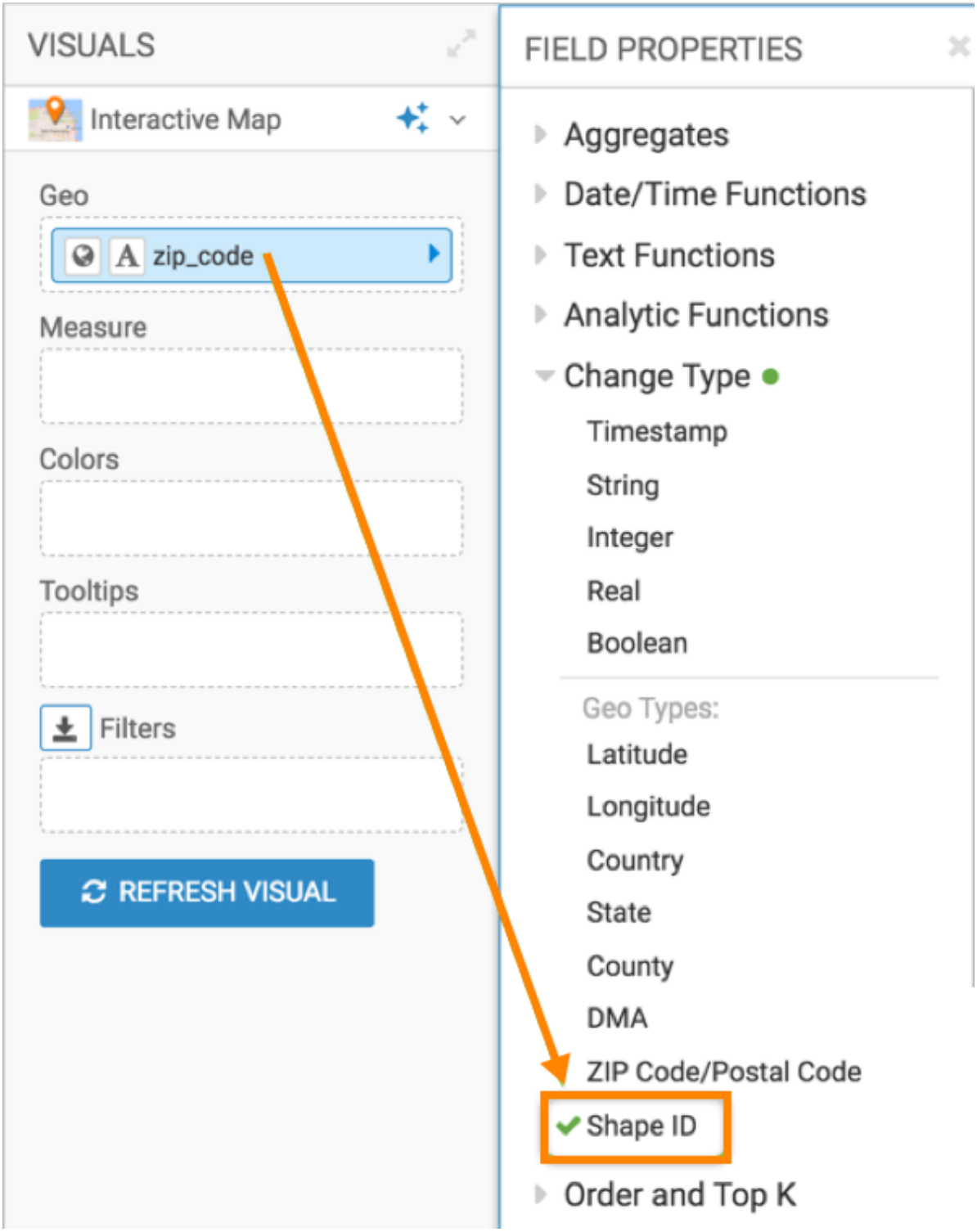
The shelves of the visual changed. They are now Geo, Measure, Colors, Tooltips, and Filters. The only mandatory shelf for map visuals is Geo.



3. Populate the shelves from the available fields in Dimensions and Measures in the DATA menu.

   Note: To create a choropleth map, the dataset must have at least one field in the Geo shelf with Geo-type Shape ID.

   a) Select the zip_code field and drag it to the Geo shelf to identify the shapes by zip code.
   b) Cast this to the appropriate Geo Type: click the field on the shelf, and in the Field Properties menu, select Change Type, and select Shape ID. This helps in identifying the measure that maps to the shape.

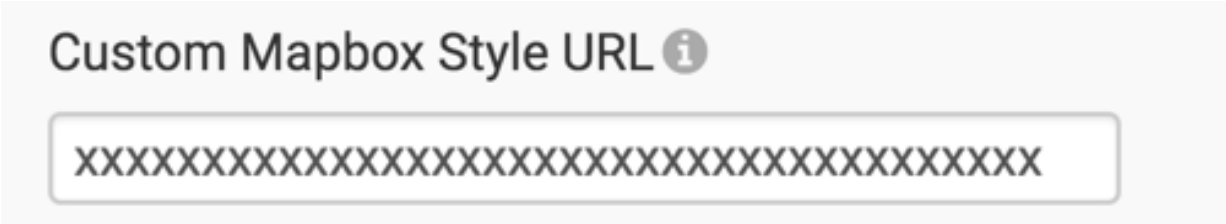


c) Under Measures, select the avg(total_tax_liability) field, and drag it to the Colors shelf.

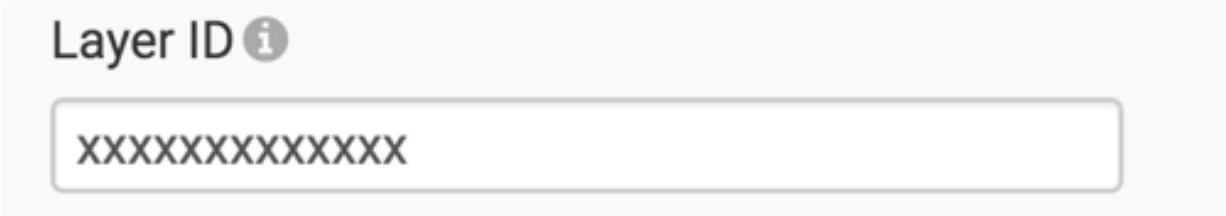
d) Change the aggregation from total_tax_liability, by selecting Aggregates Average

This colors the average tax liability in California by zip code.

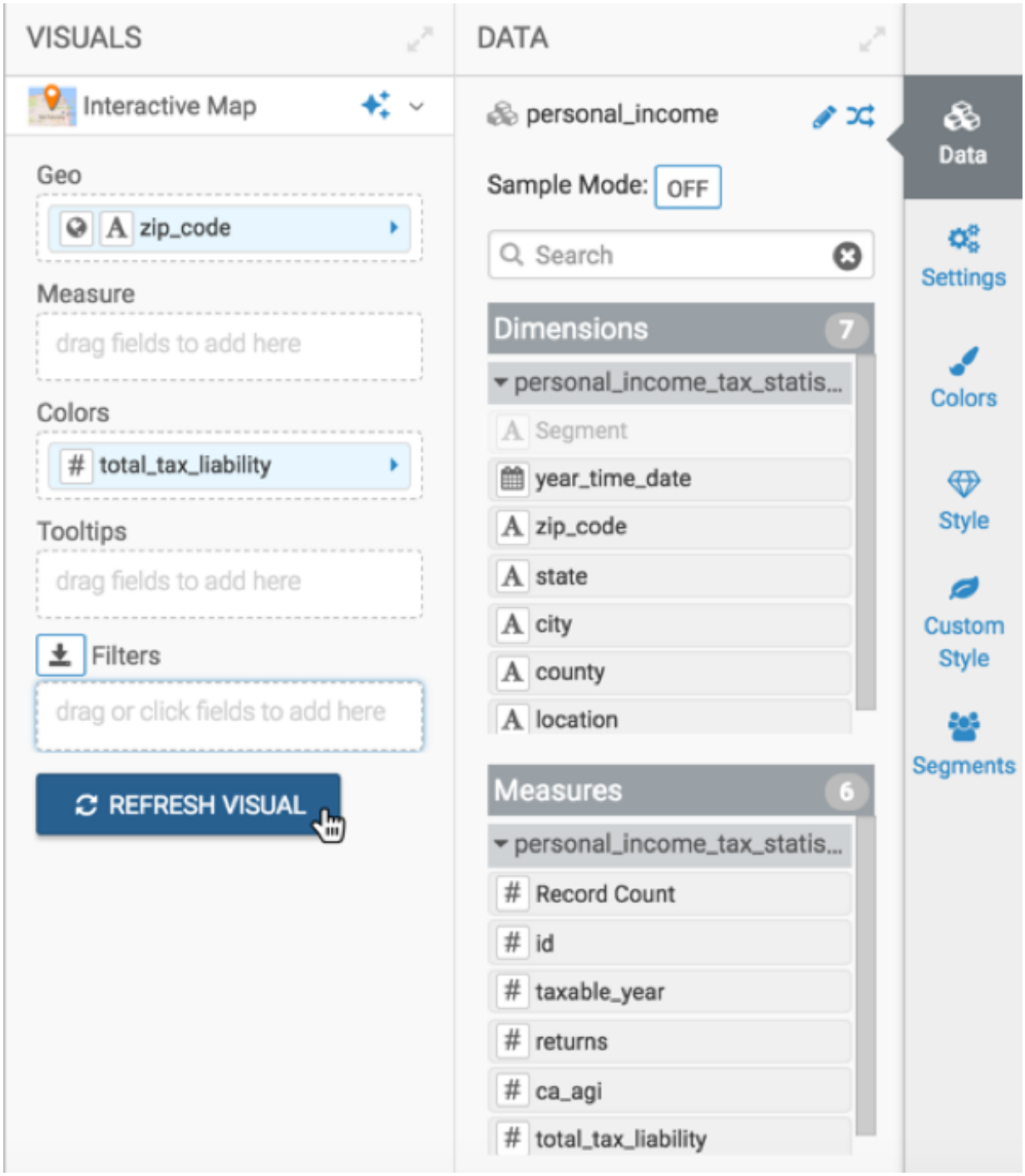
4. Navigate to the Map Style menu under Settings, and enter a Custom mapbox style that you have obtained from Mapbox. Make sure that this style has a choropleth shape file in it.



5. Navigate to the Choropleth menu under Settings, and enter a Layer ID.
   Choropleth Layer ID identifies the shapes that you would like to color by a measure in your data.

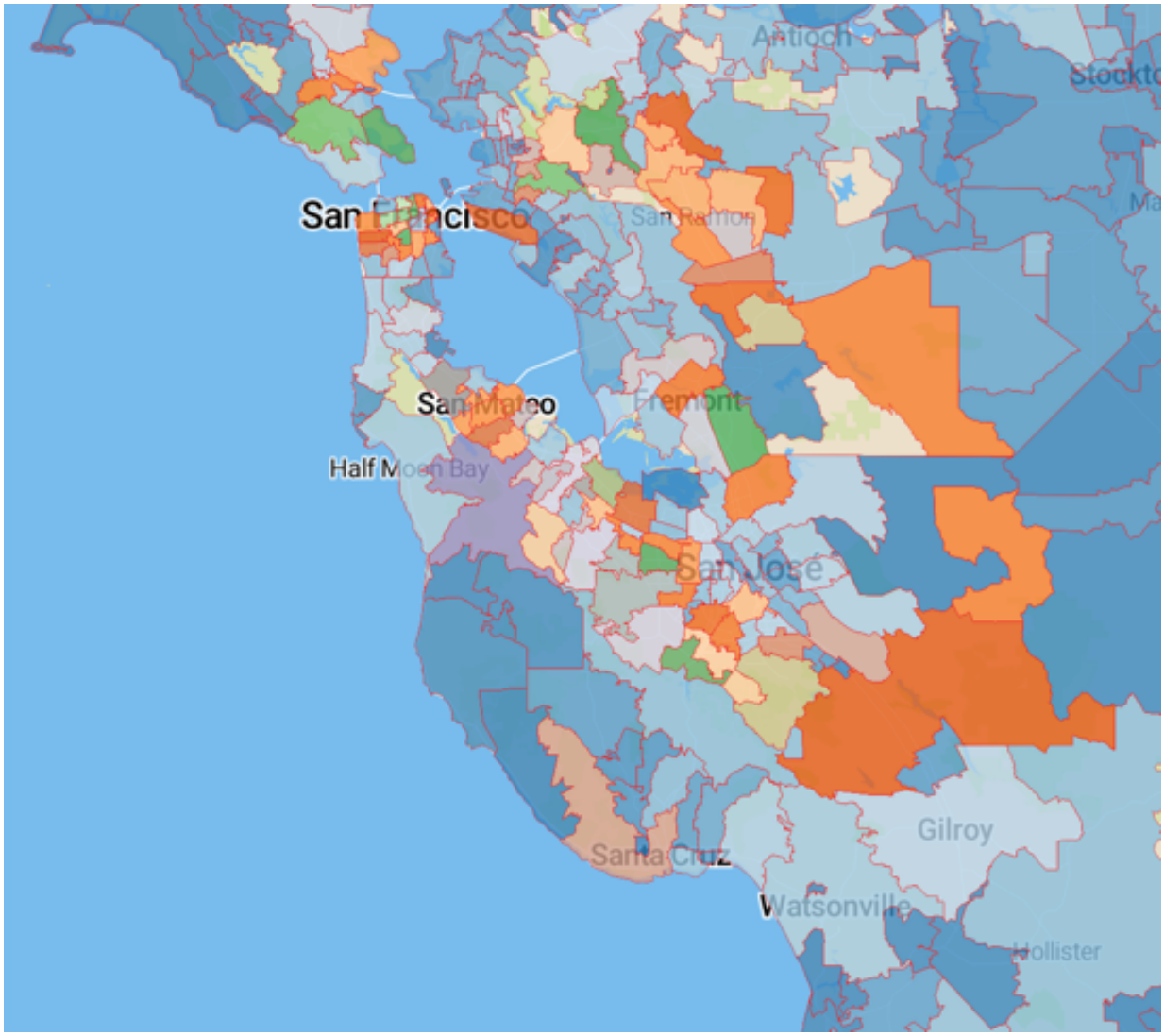


6. Click REFRESH VISUAL.



The choropleth map visual appears.

7. If the visual does not display a choropleth map, check that the Geo-type of the zip_code field is set to Shape ID.



What to do next

Proceed to *Customizing* the choropleth map.

Interactive map with multiple dimensions

Cloudera Data Visualization enables you to display multiple dimensions in Interactive Map visuals.

About this task

In Interactive Map visuals you can display two distinct dimensions in the Circle Layer by placing the secondary dimension on the Colors shelf. The following steps extend the work completed in *Interactive maps*, and demonstrate how to visualize two separate measures on the same map.

Note: To use this feature, an Admin user must specify, through the Site Settings interface, the API keys granted to your organization. This is described in *Enabling API keys for interactive maps*.

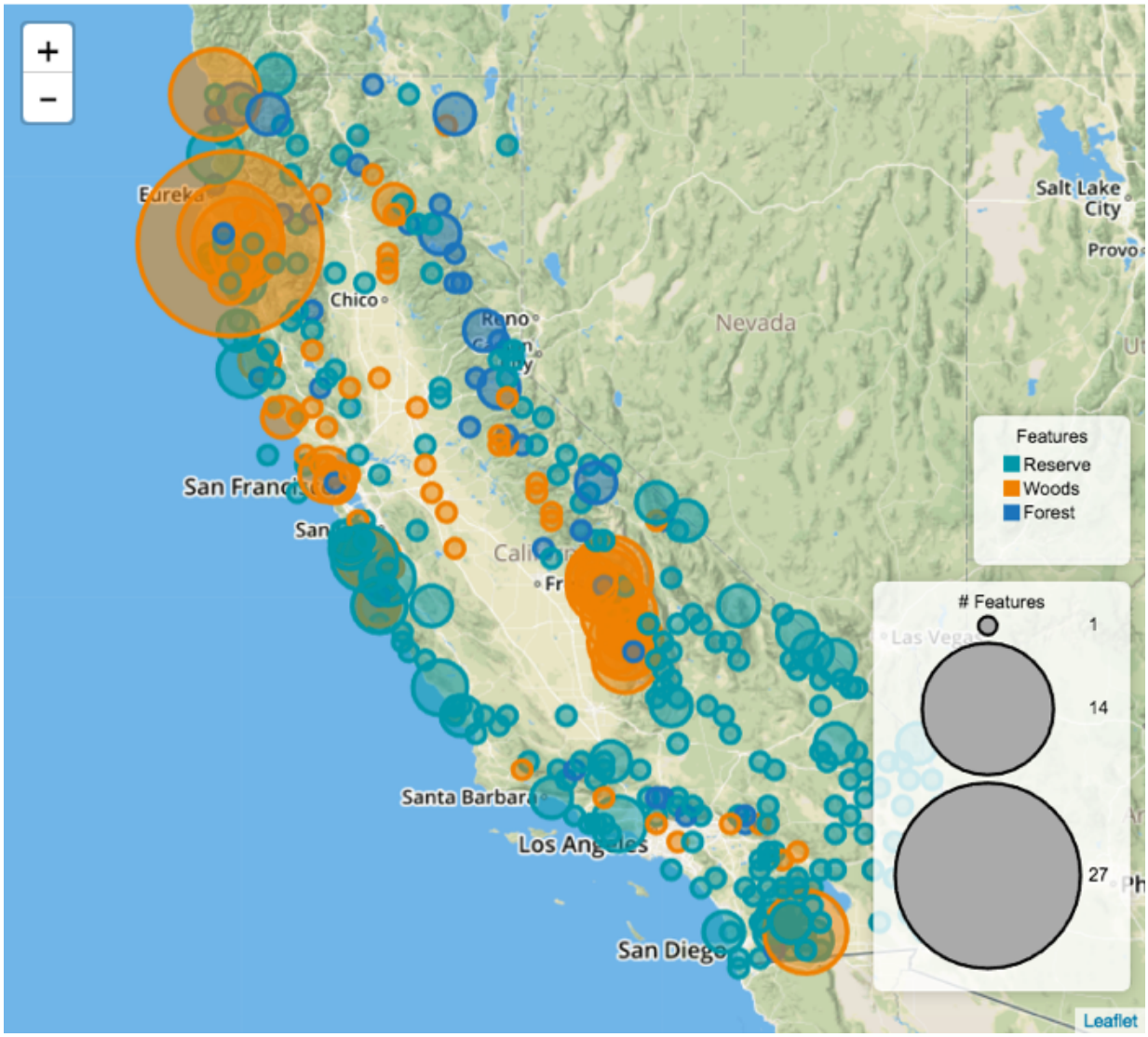
Procedure

1. Clone the visual California Woods, previously built in *Interactive maps*.

   For instructions on how to clone visuals, see *Cloning a visual*.

2. Select another dimension to add to the visual, and add it to the Colors shelf.

   In this example, feature_class is selected, and the Alias Features is applied.
3. Change the definition of the feature_class filter (already on the Filters shelf) to add the Reserve and Woods values.
4. In the Settings menu, turn off the Heatmap and the Cluster.
5. In the Settings menu, enable the Circles, and turn on both the Area legend and the Color legend.
6. Click REFRESH VISUAL to see two measures on the same map.



7. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'California Woods - Mapbox'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
8. At the top left corner of the Dashboard Designer, click SAVE.

Interactive map with multiple measures

Cloudera Data Visualization enables you to display multiple measures in Interactive Map visuals.

About this task

You can display two distinct measurements in the Circle Layer by placing the secondary measurement on the Colors shelf. The following steps extend the work completed in *Interactive maps*, and demonstrate how to visualize two separate measures on the same map.

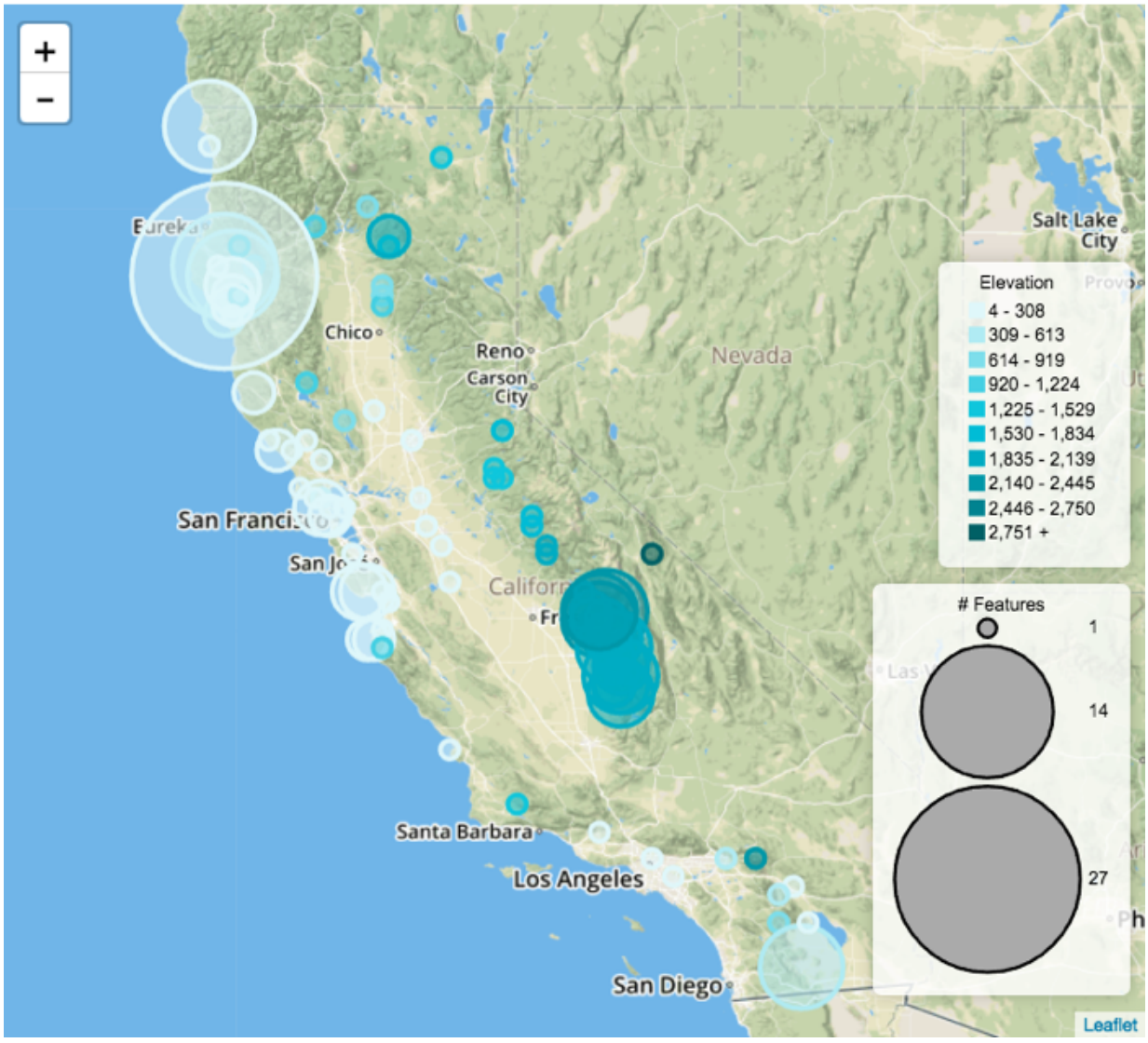
Note: To use this feature, an Admin user must specify, through the Site Settings interface, the API keys granted to your organization. This is described in *Enabling API keys for interactive maps*.

Procedure

1. Clone the California Woods visual, previously built in *Interactive maps*.

   For instructions, see *Cloning visuals*.
2. Select another measurement to add to the visual, and add it to the Colors shelf.

   In this example, elev_in_m is selected, the aggregation function is changed to max(elev_in_m), and the Elevation alias has been applied
3. In the Settings menu, turn off Heatmap, Cluster, and Routes.
4. In the Settings menu, enable Circles, and turn on both the Area legend and the Color legend options.
5. Click REFRESH VISUAL to see two measures on the same map.



6. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'California Woods - Two Dimensions'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
7. At the top left corner of the Dashboard Designer, click SAVE.

Changing the map server for interactive maps

Cloudera Data Visualization uses two third-party options for its Interactive Map visuals: Google and Mapbox.

About this task

The Map Server setting is used to specify the mapping choice.

Note: To use this feature, an Admin user must specify, through the Site Settings interface, the API keys granted to your organization. This is described in *Enabling API keys for interactive maps*.

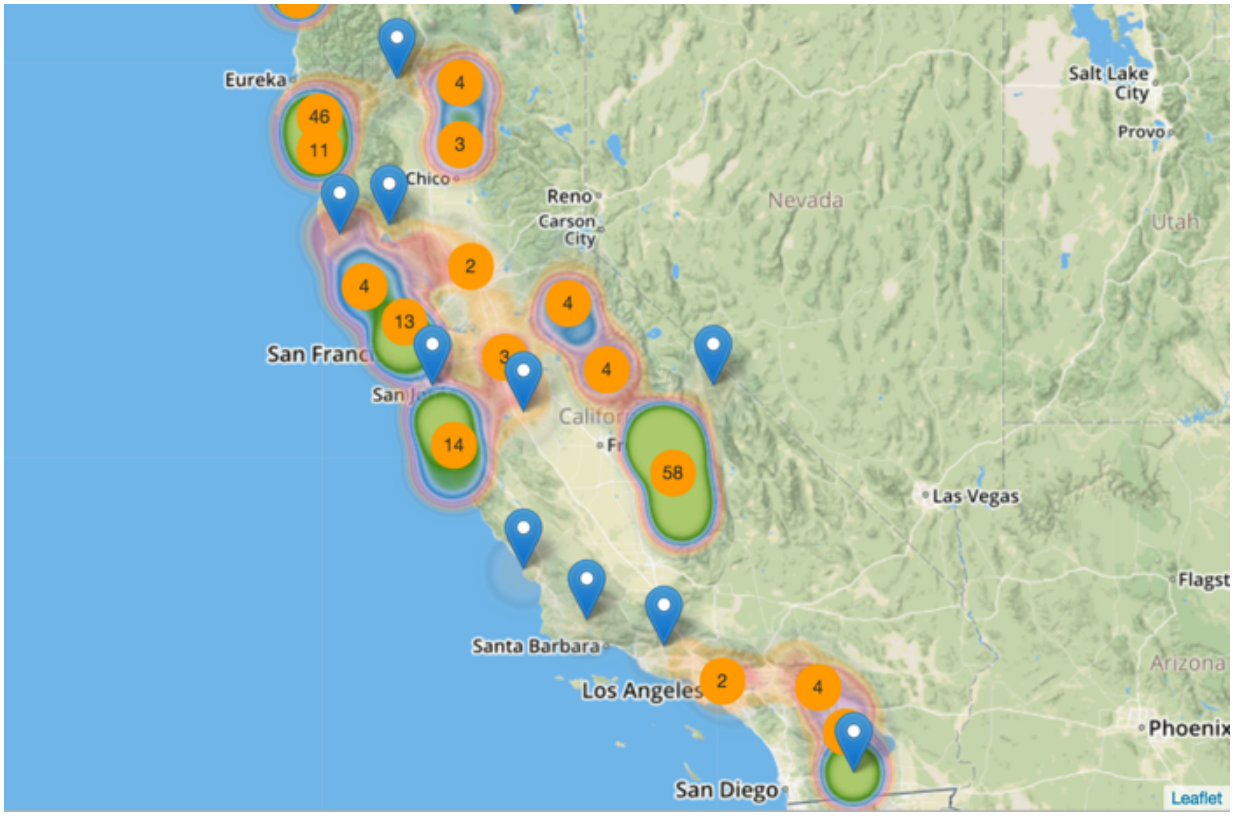
The following steps demonstrate how to switch between Google and Mapbox interactive map representation.

Procedure

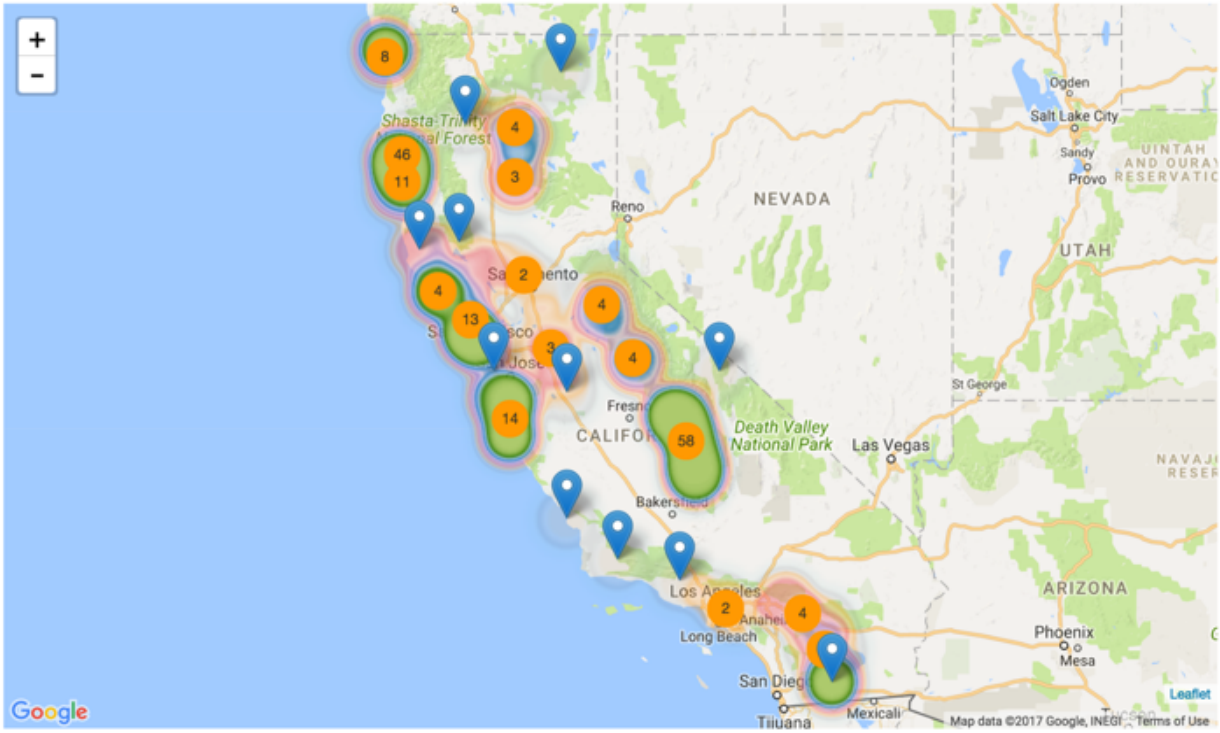
1. Clone the visual previously built in *Interactive maps*.

   For instructions, see *Cloning visuals*.
2. Follow the instructions in *Changing the map server* to switch between Google and Mapbox map styles.
3. Click REFRESH VISUAL to see the map rendered with Mapbox option.

   See the *Mapbox* documentation for more details.



4. Compare with the similar map rendered on a Google INEGI map.



5. Click the pencil/edit icon next to the title of the visualization to enter a name for the visual.

   In this example, the title is changed to 'California Woods - Mapbox'. You can also add a brief description of the visual as a subtitle below the title of the visualization.
6. At the top left corner of the Dashboard Designer, click SAVE.

Changing layer options for interactive maps

Cloudera Data Visualization enables you to make adjustments to all layers of an Interactive Map visual.

About this task

By default, interactive maps render with a simple heatmap and a cluster. However, Cloudera Data Visualization has several adjustments to all the layering options: Heatmap, Cluster, and Circles.

Note: To use this feature, an Admin user must specify, through the Site Settings interface, the API keys granted to your organization. This is described in *Enabling API keys for interactive maps*.

The following steps demonstrate how to make changes to layer specifications in interactive maps.

Procedure

1. Clone the visual previously built in *Interactive maps*.

   For instructions, see *Cloning visuals*.
2. In the Settings interface, select the appropriate option for making customizations to interactive visuals:
   - Heatmap
   - Cluster
   - Circles
3. Make the necessary changes, and click REFRESH VISUAL.
4. Change the title of the visual, and SAVE it.

Plotting routes on interactive maps

Cloudera Data Visualization enables you to plot routes over an Interactive Map.

About this task

You can plot routes over a map by specifying distinct IDs on the Color shelf, a time series on the Tooltips shelf, and enabling routes as described in *Enabling routes*.

The following steps demonstrate how to plot routes on an interactive map. In the following example, truck movement data is used.

Procedure

1. Select the appropriate dataset, and create a new dashboard.
2. In the dashboard, open the default visual and change the visual type to Interactive Map.
3. Enable the Routes, as described in *Enabling routes*
   a) Click the Settings menu.
   b) Select Routes/Pins Enable Routes .
4. Disable heatmap and cluster. It is recommended that you leave the circles in place, so it shows the individual elements of the route. If you choose to hide the circles, the graph only shows the route connectors.
5. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Geo Shelf - Under Measures, select latitude and longitude, and add them to the Geo shelf.

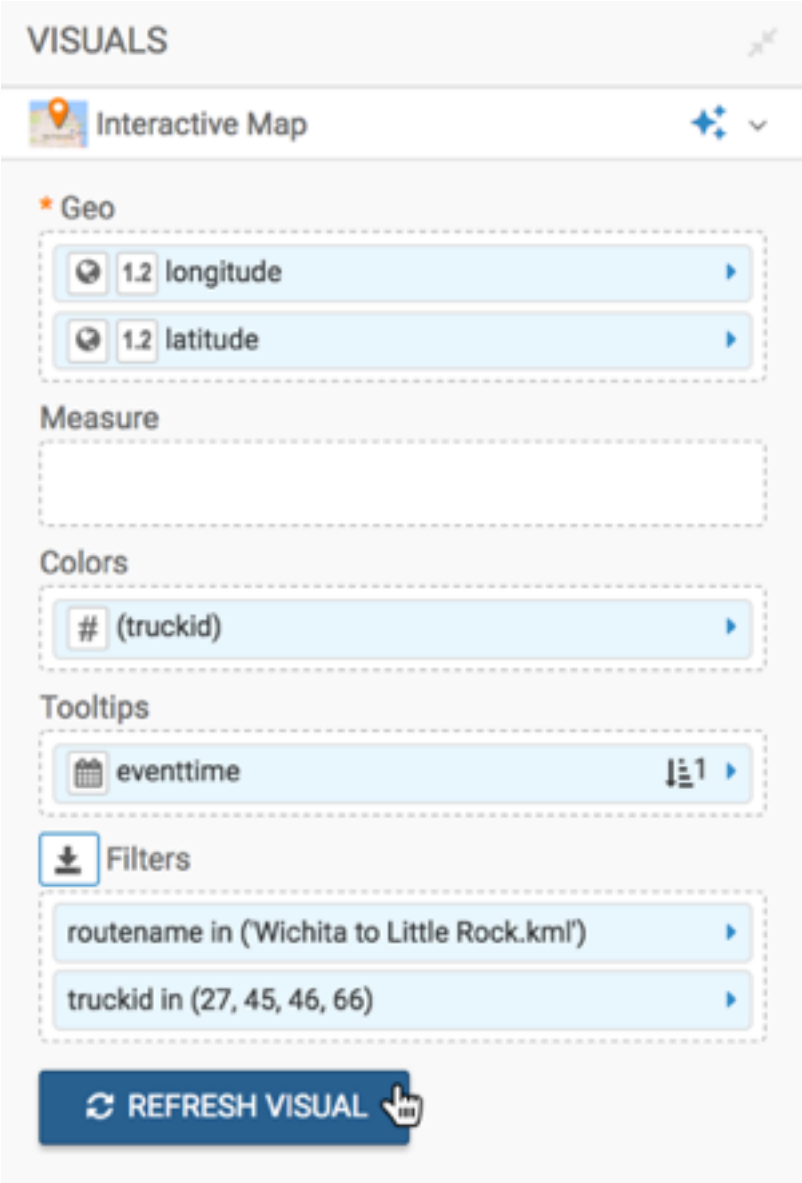
      If your longitude and latitude measurements are not automatically recognized (appear with a (globe) icon on the Geo shelf), cast them to the appropriate Geo Type.

      1. Click the field on the shelf to open the FIELD PROPERTIES menu.
      2. Select Change Type and choose Latitude or Longitude as appropriate.

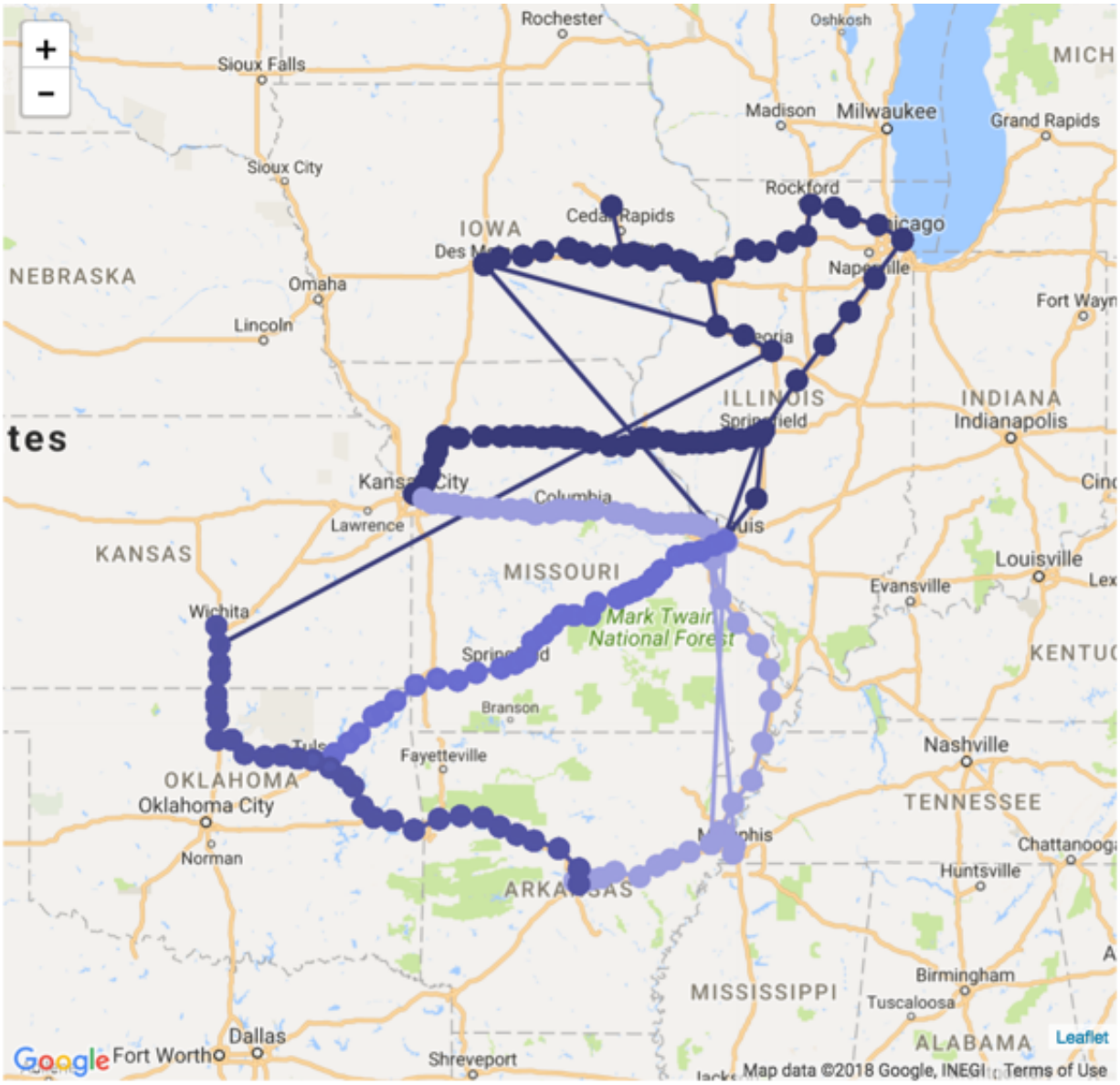
      For more information, see *Change type* and *Geo data type*.
   b) Colors Shelf - Under Measures, select truckid, and place it on the Colors shelf.

      You may need to remove the aggregation from the field on the shelf.
   c) Tooltips Shelf - Under Dimensions, find the eventtime field, and add it to the shelf.

      1. Click the field on the shelf to open the FIELD PROPERTIES menu.
      2. Click Order and Top K.
      3. Choose the Ascending order.
   d) Filters Shelf - [Optional] To restrict the number of routes, you can choose the filter on the routename field. To restrict the number of trucks displayed, you may choose to restrict on truckid field, and so on.
   e) Measure Shelf - [Optional] You can display a measure here.
6. Change the color palette to ensure that the routes stand out sufficiently over the regular features of the map: click Colors and select an appropriate palette.
7. Click REFRESH VISUAL.



The visual is displayed.



8. Change the title of the visual, and SAVE it.

Using alphabetic values in interactive maps

Cloudera Data Visualization Interactive Map visuals support both numeric and non-numeric values on the Color shelf. As a result, aspects of named categories can be viewed simultaneously on the same visual as a distinct series.

About this task

In the following example, truck movement data is used to demonstrate how to plot incident counts for several drivers, and differentiate them by the color.

Procedure

1. Select the appropriate dataset, and create a new dashboard.
2. In the dashboard, open the default visual and change the visual type to Interactive Map.
3. Enable the Circles, as described in *Displaying circles*.
   a) Click the Settings menu.

b) Select Circles Enable Circles .

4. Enable both the area and color legends of the Circles option by selecting the Add Circle Area Legend and Add Circle Color Legend options.
5. Disable all the other layers: heatmap, cluster, marker, and route.
6. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Geo Shelf - Under Measures, select latitude and longitude, and add them to the Geo shelf.

      If your longitude and latitude measurements are not automatically recognized (appear with a (globe) icon on the Geo shelf), cast them to the appropriate Geo Type.

      1. Click the field on the shelf to open the FIELD PROPERTIES menu.
      2. Select Change Type and choose Latitude or Longitude as appropriate.

      For more information, see *Change type* and *Geo data type*.

   b) Measures Shelf - Under Measures, select eventtype and place it on the measures shelf.

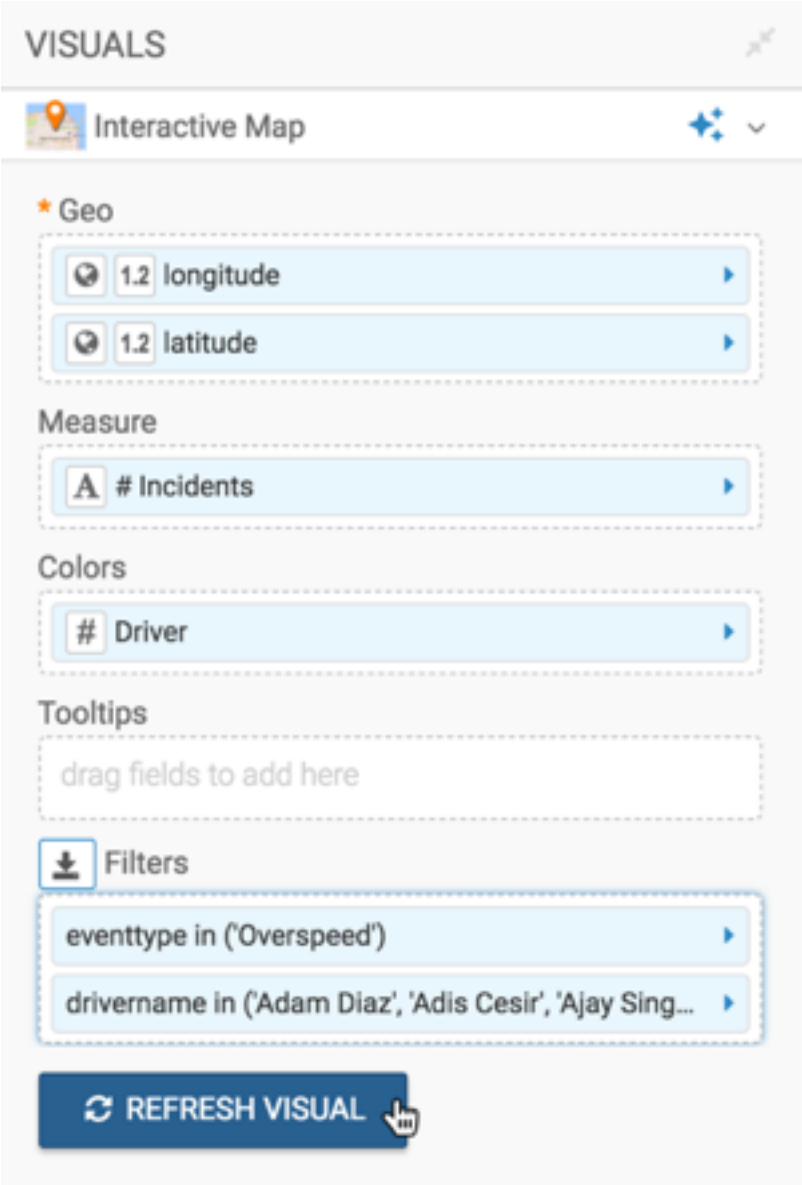
      1. Change the aggregation function to count().
      2. Alias the field as # Incidents.

   c) Colors Shelf - Under Measures, select drivername, and place it on the Colors shelf.

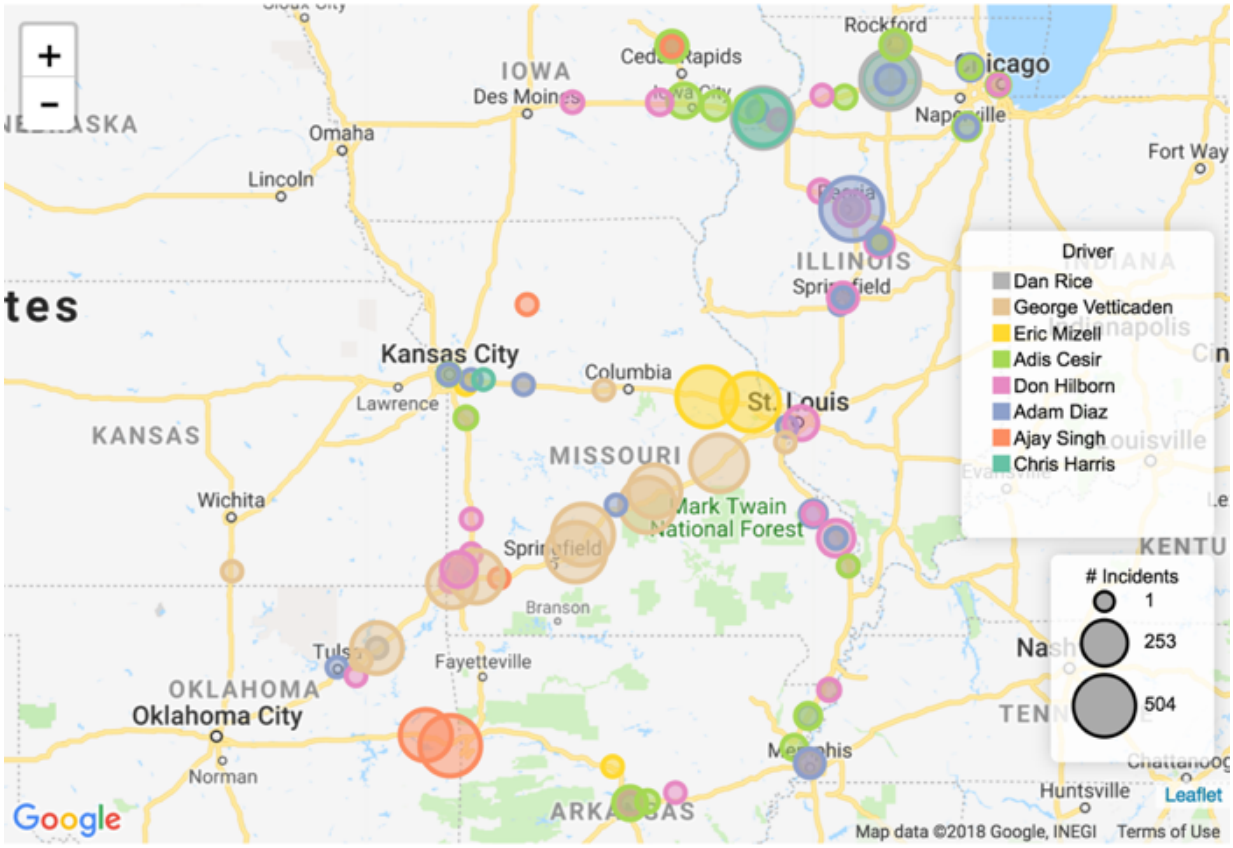
      You may need to remove the default field aggregation from the field on the shelf. Alias the field as Driver.

   d) Filters Shelf - Add the field eventtype to this shelf, and select the single value 'Overspeed'.

      1. Add the field drivername to the filters shelf.
      2. Select a small number of drivers to see details.
7. Change the color palette to ensure that the routes stand out sufficiently over the regular features of the map: Click Colors, and select an appropriate palette.
8. Click REFRESH VISUAL.



The visual is displayed. The colors are defined by the names of the drivers, which are non-numeric.



9. Change the title of the visual, and SAVE it.

Segmenting data qualitatively in interactive maps

Cloudera Data Visualization enables you to plot qualitative ranges of data by specifying the data segmentation on the Color shelf of an Interactive Map visual.

About this task

The following steps demonstrate how to plot incident counts on an Interactive Map. Truck movement data is used in the example.

Procedure

1. Select the appropriate dataset, and create a new dashboard.
2. In the dashboard, open the default visual and change the visual type to Interactive Map.
3. Enable Circles, as described in *Displaying circles*.
   a) Click the Settings menu.
   b) Select Circles Enable Circles .
4. Enable both the area and color legends of Circles by selecting the Add Circles Area Legend and Add Circles Color Legend options.
5. Disable all the other layers: heatmap, cluster, marker, and route.
6. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.
   a) Geo Shelf - Under Measures, select latitude and longitude, and add them to the Geo shelf.

      If your longitude and latitude measurements are not automatically recognized (appear with a (globe) icon on the Geo shelf), cast them to the appropriate Geo Type.

      1. Click the field on the shelf to open the FIELD PROPERTIES menu.

2. Select Change Type and choose Latitude or Longitude as appropriate.

   For more information, see *Change type* and *Geo data type*.

b) Measures Shelf - Under Measures, select eventtype and place it on the measures shelf.

1. Change the aggregation function to count().
2. Alias the field as # Incidents.

c) Colors Shelf - Under Measures, select eventtype and place it on the Colors shelf.

1. Click the field on the shelf, to open the FIELD PROPERTIES menu.
2. Select the [] Enter/Edit Expression option.
3. Change the field expression to split the data into two qualitative ranges, High and Low:

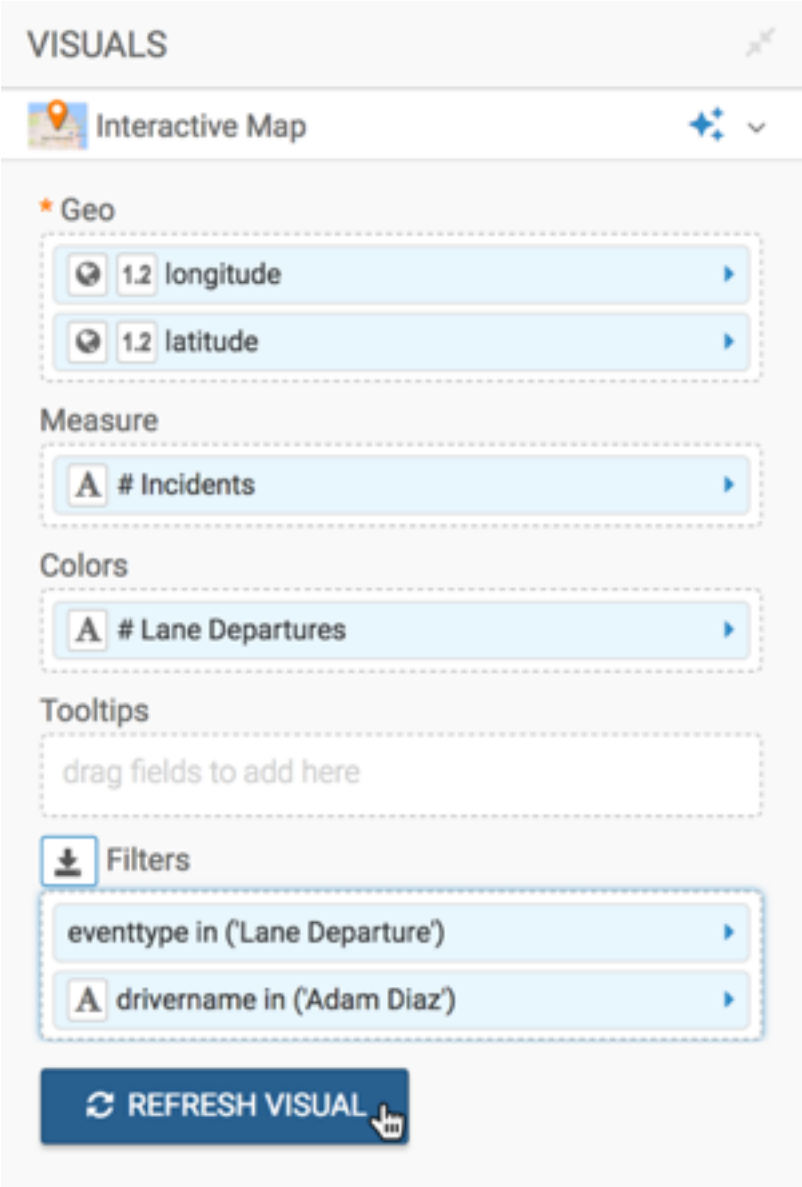
```
case when count([eventtype]) > 20 then 'High' else 'Low (20 or less)'
 end
```

4. Change the field Alias to # Lane Departures.

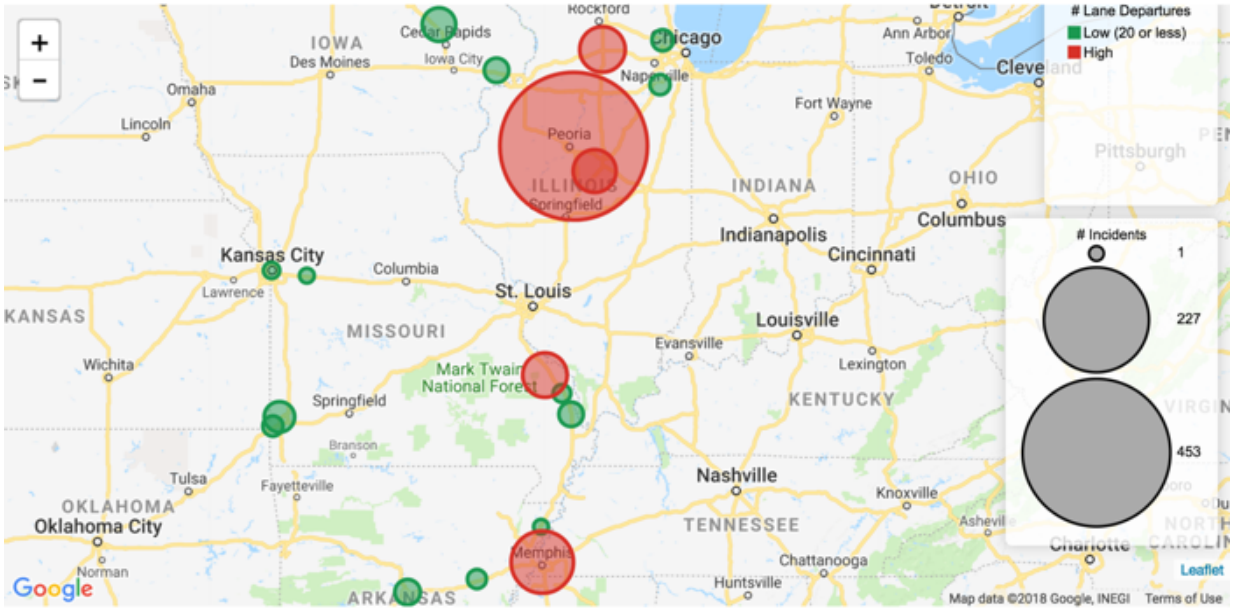
d) Filters Shelf - Add the field eventtype to this shelf, and select the single value 'Lane Departures'.

   [Optional] Add the field drivername to the filters shelf, and select a small number of drivers to see details.

7. Change the color palette to ensure that the colors are distinct: Click Colors, and select an appropriate palette.
   In this example, because the data is divided into two distinct groups, a two-color palette is chosen.
8. Click REFRESH VISUAL.



The visual is displayed.



9. Change the title of the visual, and SAVE it.

Shelves for interactive maps

Overview of shelves for Cloudera Data Visualization Interactive Map visuals.

An Interactive Map visual has the following shelves:

Geo

Add a Geo-type field from the DATA menu to determine the location on the map.

This visual supports zipcode or latitude/longitude Geo types . If the field is not automatically recognized as a qualified Geo type, select Change Type from the FIELD PROPERTIES menu.

This shelf accepts multiple fields, but plots a map with the highest possible resolution.

This is a mandatory shelf.

Measure

Add a Measure field from the DATA menu to compute heatmap intensity across locations, values of cluster aggregations, and the size of circle markers.

This shelf accepts a single field, or uses the value of 1 if no measurement is specified.

Colors

Add a Dimension or Measure field from the DATA menu to specify the color of the circle markers, or distinct routes by common identifier.

This shelf accepts a single field, and uses both numeric and non-numeric values.

Tooltips

Add Measure fields from the DATA menu for additional information in the visual's Tooltip.

This shelf accepts multiple fields.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

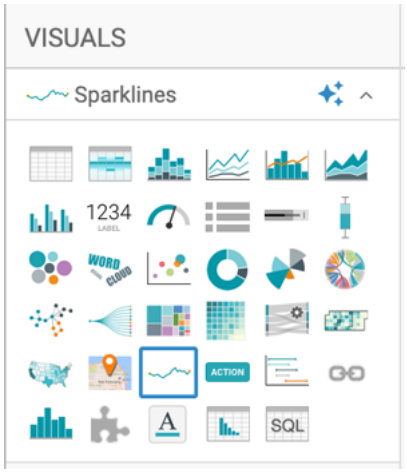
This shelf accepts multiple fields.

Sparklines

Cloudera Data Visualization enables you to create Sparklines visuals. A Sparklines chart shows the general shape of the variation in measurement (usually over time) in a simple, condensed, and memorable manner. Sparklines are small and succinct, and can be easily located with the text where they are discussed. You can contrast Sparklines with typical X/Y charts that are designed to show as much data as possible.

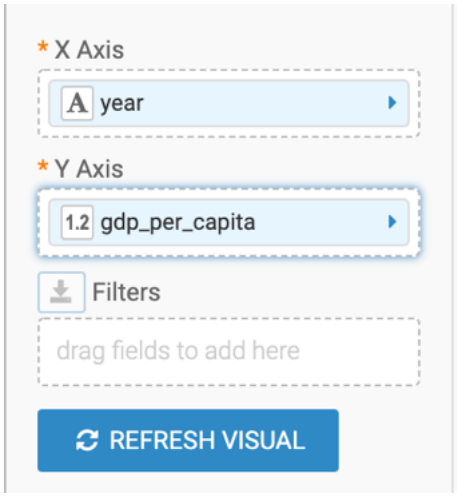
Procedure

1. Start a new visual based on the World Life Expectancy dataset.
2. Select Sparklines in the VISUALS menu.

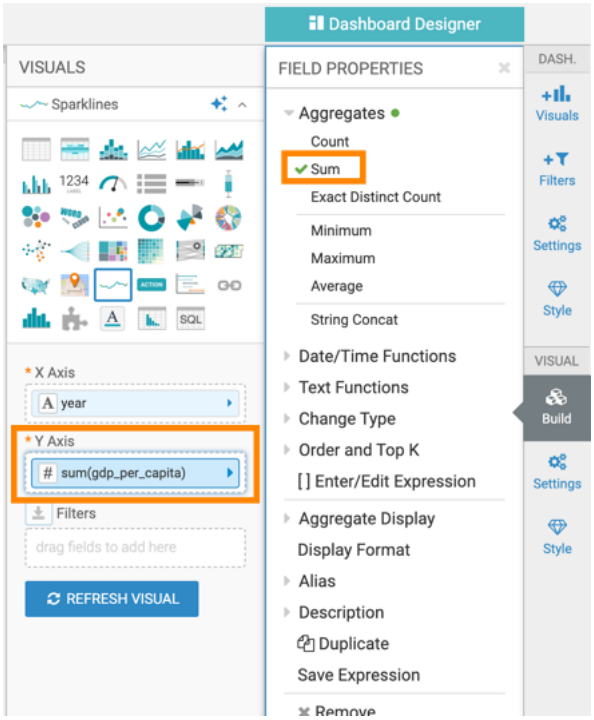


The shelves of the visual changed. They are now X Axis (mandatory), Y Axis (mandatory), and Filters.

3. Populate the shelves from the available fields (Dimensions, Measures) in the DATA menu.
   a) Under Dimensions, select year and drag it to the X Axis.
   b) Under Measures, select gdp_per_capita and drag it to the Y Axis.



4. On the gdp_per_capita field of the Y Axis shelf, click the (arrow) icon and select Aggregates Sum .
   This is necessary because the Sparklines visual type uses aggregated measurements.



The Y Axis shelf now contains the modified field sum(gdp_per_capita).

5. Click REFRESH VISUAL.

   The Sparklines visual graph appears.

6. On the Filters shelf, add several Dimensions and Measures from the DATA menu.

   This enables you to dynamically control the data input and to discover the data at a more granular level.

   For example, from Dimensions, drag country, un_subregion, and un_region to the Filters shelf.

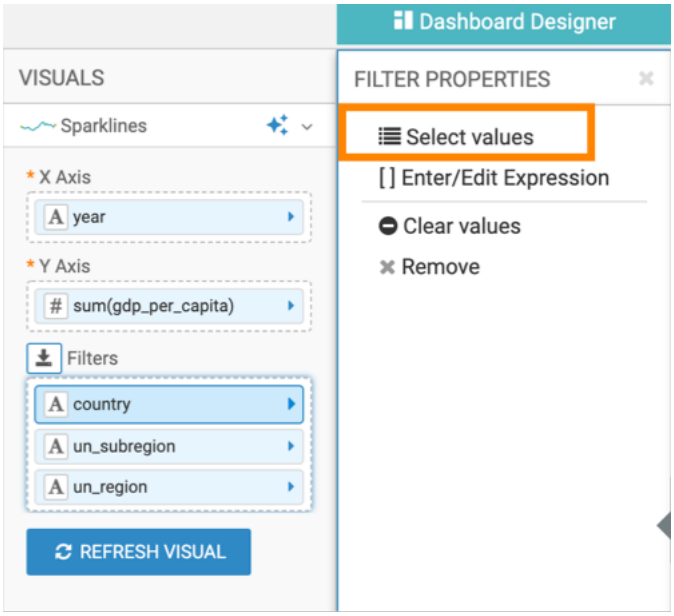
7. Click the (pencil) icon next to the title of the visualization to edit it, and enter World Population - Sparklines as the new name.
8. Click SAVE at the top left corner of the Dashboard Designer.

Using sparklines

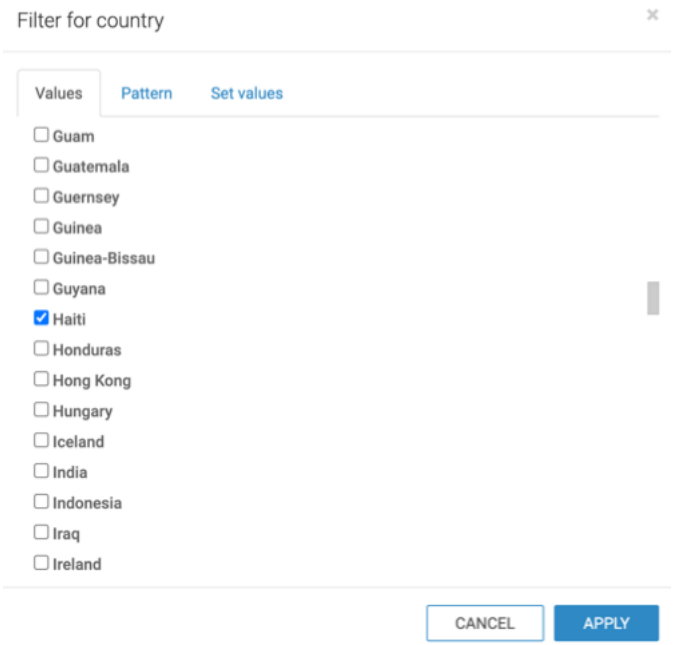
Learn about using Sparkline visuals.

Procedure

- Follow these steps to create a Sparklines chart for a specific country:
  a) On the Filters shelf, click country and click Select values.



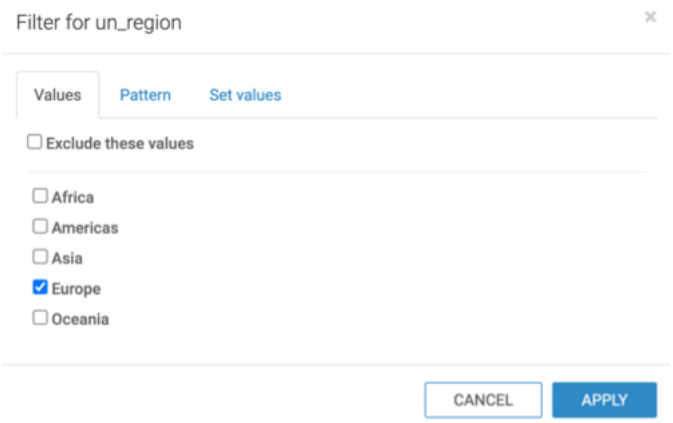
b) In the Filter for country modal window, select Haiti and click APPLY.



c) Click REFRESH VISUAL.

You can see how the line has changed compared to the previous display (world-wide).

- Follow these steps to create a Sparklines for a region, such as Europe:
  a) Clear the selection in the existing filters.
  b) From Dimensions, drag un_region to the Filters shelf.
  c) In the Filter for un_region modal window, select Europe and click APPLY.



d) Click REFRESH VISUAL.

You can see how the line has changed.

Shelves for sparklines

Overview of shelves for Cloudera Data Visualization Sparklines visuals.

A Sparklines visual has the following shelves:

X Axis

This is conceptually the X axis. Use a Dimension for a horizontally-oriented plot or a Measure for a vertically-oriented plot.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in columns.

Y Axis

This is conceptually the Y axis. Use a Measure for a horizontally-oriented plot or a Dimension for a vertically-oriented plot.

When using multiple fields, the last field defines the axis, while the preceding fields trellis the graph in rows.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

AI visual

Cloudera Data Visualization enables you to embed interactive and intelligent components into dashboards and applications.

The AI visual provides a natural language interface for exploring datasets through text or speech. It enables you to ask questions in everyday language and receive contextual, data-driven responses directly within dashboards. By

combining conversational input with flexible query approaches, the AI visual makes it easier to analyze data and discover insights without requiring technical expertise.

When added to a dashboard, the AI visual connects to the selected dataset and answers user questions using the configured AI Approach. Depending on the settings and the query, it may:

- Use the ongoing conversation context – Reference previous exchanges to provide more relevant, contextual answers.
- Run a SQL query – Translate the user question into SQL, run it on the dataset, and summarize the results.
- Perform a similarity search – Run a query against the vectorized dataset and return matching records.

Important:

The option to select between AI approaches is available starting with Cloudera Data Visualization 8.0.6. If you are using AI visual in an earlier product version, you can only view visuals on embeddings-supported connections and use the similarity search feature.

The retrieved information is then passed to the large language model, which generates a natural language response displayed within the visual.

The AI visual can apply smart filters to your dashboard based on your questions. For example, asking about a specific topic will automatically adjust other visuals to highlight relevant data. This helps validate the AI's response and offers a more intuitive way to explore insights, without the need to manually search through filters or menus.

Creating an AI visual

You can create an AI visual in Cloudera Data Visualization to let users query data through natural language. Before creating the visual, you must configure the required AI engine and data settings, and ensure the dataset meets the requirements for the chosen AI approach (SQL query or similarity search).

Before you begin

Before you start creating the AI visual, make sure the following prerequisites are met.

Configuration requirements:

- Enable the AI visual in Site Settings AI Settings General . For more information, see Managing AI settings.
- Select and configure the AI engine. For more information, see Managing AI settings.
- Enable data profiling in Site Settings Data . For more information, see Managing data related site settings.
- [optional] Enable speech detection in Site Settings Site configuration to support voice input. For more information, see Managing site configuration.

Data requirements:

Important:

The option to select between AI approaches is available starting with Cloudera Data Visualization 8.0.6. If you are using AI visual in an earlier product version, you can only view visuals on embeddings-supported connections and use the similarity search feature.

For SQL query AI approach

There are no special data requirements if you are using the SQL query AI approach.

For Similarity search AI approach

If you want to use similarity search as the AI approach, you must connect to a data source that contains a vector column with pre-created embeddings. This can be done in different ways depending on how your data is stored and ingested.

One common method is to upload a CSV file into a supported vector database (Solr or SQLite). For instructions, see Importing data in CSV format .

Note: The CSV file must include a vector column with pre-generated embeddings.

Another option is to use continuous data ingestion. If your dataset is updated continuously (for example stored in an Apache Iceberg table), you can use a data pipeline to:

- Convert new records into vector embeddings
- Write them into the connected vector database

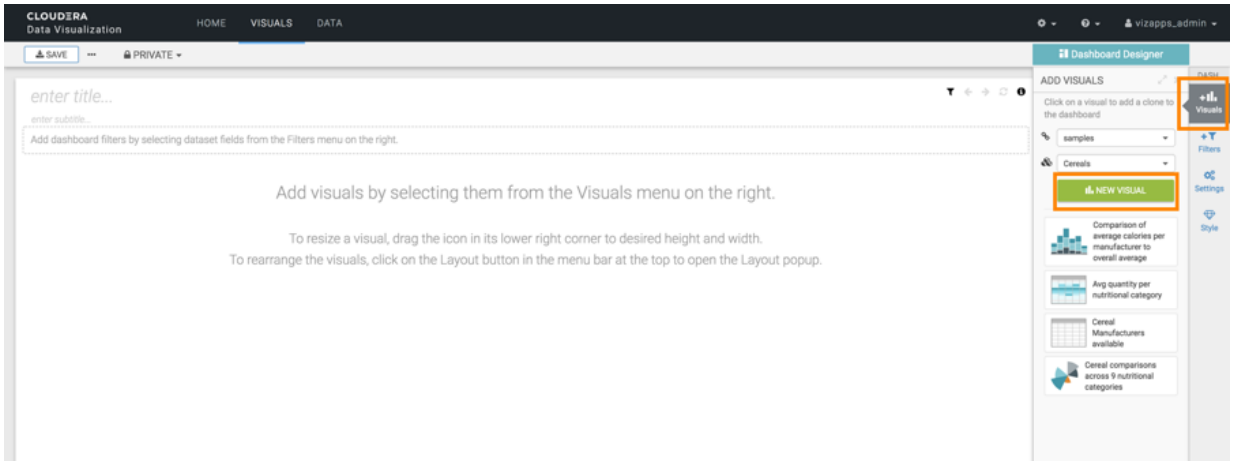
Continuous data ingestion allows you to keep the dataset used by the AI visual up to date and relevant with minimal manual effort. Follow the below steps to connect to the required data source.

1. Ensure the data you plan to use with the AI visual is vectorized and available in a supported vector database (SQLite, Solr 9+, or CSS). This can be from a CSV upload or a continuously updated data source.
2. Connect to the vector database that contains your dataset with vector embeddings.
3. Create a dataset from the vectorized data. For instructions, see Creating a dataset.

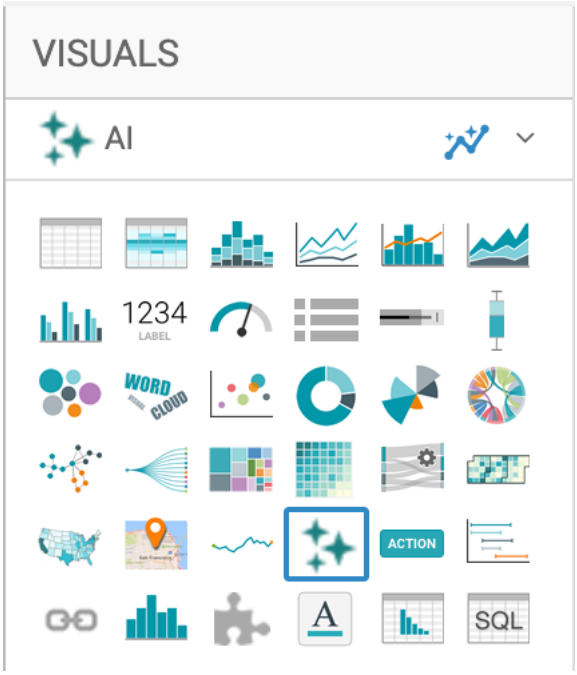
Note: Solr integration requires Solr 9. Earlier versions, including Solr 8, are not supported.

Procedure

1. On the main navigation bar, click VISUALS.
   - To add the AI visual to an existing dashboard: open the dashboard and click EDIT.
   - To start fresh: click Create New Dashboard to open a new, blank dashboard.
2. In the Dashboard Designer interface, open the Visuals menu from the side menu bar and click NEW VISUAL.



   The Visual Builder appears. By default, a table visual is created. For more information, see *Visual Builder overview*.
3. In the VISUALS menu, click the AI visual icon.



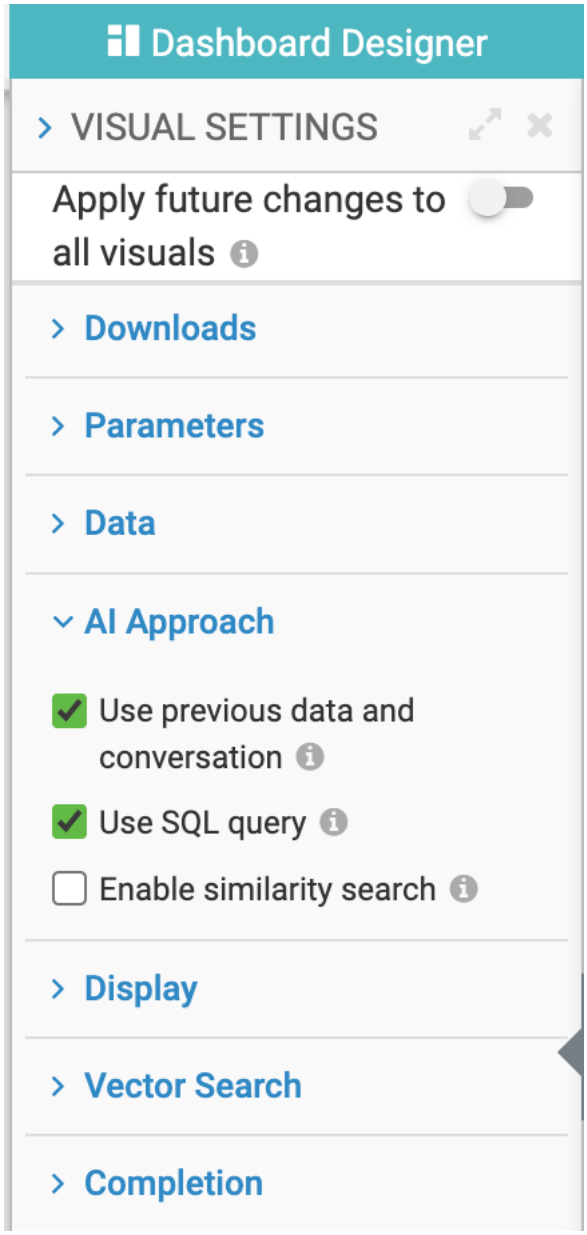
A blank AI visual is added to the dashboard.

The screen also displays the current AI approach (SQL query by default) and the default completion model. You can adjust the AI approach and other settings in the VISUAL Settings panel.

4. Click Settings AI Approach from the right-side VISUAL menu and choose how user queries will be answered.

   The following options are available:

   - Use previous data and conversation – Incorporates previous context in the conversation to refine or expand responses to provide more contextual answers. This option is enabled by default.
   - Use SQL query – Generates SQL queries against the connected dataset and summarizes the results. It is enabled by default, and when enabled, it disables similarity search.
   - Enable similarity search – Performs a vector-based search to find the most relevant records to answer the query. When enabled, SQL query is disabled automatically. This option is only available for SQlite, Solr 9+, and CSS datasets.



To use the AI visual, you must select either SQL query or similarity search. With both options, you can also include previous data and conversation context.

If you are using SQL query, continue with Step 5.

If you are using similarity search, first you need to populate the shelves of the AI visual.

a. Switch back to the Build menu and populate the AI visual shelves from the fields listed in the DATA pane.

- Embeddings: Add vector fields containing embeddings that you want the AI model to analyze in a semantic way.
- Context Dimensions: Add fields to group or segment embedding results into categories.
- Context Aggregates: Add aggregated fields that summarize data for similar records in the AI visual. This allows the AI visual to better answer questions related to quantities, distributions, and comparisons.
- Tooltip: Define the source information to be included in the response tooltip. This information appears when the user hovers over the response's Info icon, but it is not sent to the completion service.
- Limit: Define the maximum number of data rows retrieved from the vector database and analyzed.

b. Click REFRESH VISUAL.

c. [Optional] If you are using similarity search, you can configure embedding and completion related settings. Click Settings Vector Search from the right-side VISUAL menu to:

- Set the maximum number of tokens. You can use 0 or a negative value to disable this setting
- Change the default profile
- Provide a template for vector search responses

Note: If you are using SQL query as the AI approach, the Vector Search menu is disabled.

5. Click Settings Completion from the right-side VISUAL menu to configure completion.

You can:

- Change the selected profile.
- Set the maximum number of tokens. If exceeded, the request is not sent to the service and an error message is displayed.
- Choose a context overflow policy that defines how to handle situations where the length of generated tokens exceeds the context window size:
  - Throw an error: Raises an error when the conversation length exceeds the defined context window.
  - Remove old conversation first: Automatically manages conversation length by removing old tokens.
  - Truncate embedding context first: Truncates the embedding context, retaining the latest tokens.
- Specify the context placeholder text that is used in the search results when data is missing or empty. If set, the text replaces empty values, if left blank, empty values are removed.
- Specify the question prompt, which is the formatting of the user question.
- If you are using SQL query, you also have the option to provide custom instructions for formatting SQL query results.These instructions guide how data analysis results are presented to the users.

Note: If you are using similarity search, the SQL result formatting instructions field is not available.

6. Click Settings Display from the right-side VISUAL menu to configure how the visual is displayed.

- Customize the chat welcome message, which is displayed as the first chat message from the AI visual.
- Set the display format for the conversation: plain text or markdown

7. Click SAVE.

Using an AI visual

The AI visual allows you to explore datasets through natural language queries. You can ask questions by typing or speaking, review detailed execution steps to understand how responses are generated, and share parameters across visuals to keep dashboards synchronized.

Interacting with the AI visual

After configuring the AI visual, you can begin exploring your data through conversational queries. Ask questions about your data using voice or text. The AI visual analyzes the connected dataset and generates responses using natural language. Responses are streamed into the conversation as they are generated.

Adding text input:

1. Type your question into the input field.
2. Click SEND to submit the query.

The AI visual processes your question and returns a relevant response based on the dataset. When the response begins streaming, the answer appears progressively in the conversation.

Adding speech input:

1. Click to ask a question using your voice.
2. Click SEND to submit the query.

   The AI visual processes your question and returns a relevant response based on the dataset. When the response begins streaming, the answer appears progressively in the conversation.

Adding follow-up queries:

You can continue the conversation by asking follow-up questions to refine or extend the previous query, using either text or voice input.

Canceling a response:

After you ask a question, the SEND button changes to STOP while the question is being processed.

Click STOP to:

- Stop displaying any additional streamed text immediately
- Cancel the in-progress request
- Return to idle state so you can ask another question right away.

If it was a new question, the pending question is removed from the chat and restored in the input box so you can edit and resend it.

If your question was a retry of a failed response, the original question remains in the chat history, and the retry is canceled.

Viewing response details

You can review how the AI visual generated a response. Execution Details provide a transparent, step-by-step explanation of how the query was processed and which data was used. This information can help you:

- understand how the response was generated
- verify which data was used
- troubleshoot unexpected results

To view execution details, click next to the response. The Execution Details modal includes the following sections:

Overview

Displays summary metadata for the response, including:

- AI approach (SQL / Vector / Previous Data) with description
- Completion profile
- Completion model details:
  - ID
  - name (when available)
  - engine
- Embedding profile
- Embedding model details (vector approach only):
  - ID
  - name
  - engine
- Execution summary (completed versus total steps, including failures)
- Execution duration
- Data summary (rows × columns)

- Dataset name
- Number of applied filters
- Retry indicator (if the response is a retry, including failed step)
- Original question
- Response length (characters)

Important:

The option to select between AI approaches is available starting with Cloudera Data Visualization 8.0.6. In earlier versions, you can only view visuals on embeddings-supported connections and use the similarity search feature.

Execution Steps

Shows a breakdown of how the response was generated:

- Checking Previous Data determines whether earlier queries already contain the required information.
- Collecting Dataset Information and Filters shows dataset details and applied dashboard filters.
- Generating and Validating SQL shows the generated query and visual configuration.
- Analyzing Results provides the natural language summary of the query results.
- Vector Search / Generating Response performs embedding-based retrieval and generates the final answer. (vector or text-only paths)

You can expand each step to view the associated data, dataset information, filters, or query results.

To save the execution details, click Download.

Sharing parameters

When querying a dataset with an AI visual, the generated parameters can be shared with other visuals on the dashboard to enable cross-filtering.

Sharing parameters allows you to explore relationships between different data points more efficiently. When a shared parameter changes, all dependent visuals on the dashboard update automatically to reflect the same subset of data. This eliminates the need to adjust each visual individually and helps maintain consistency across the dashboard.

1. Click next to the response.

   The Share Parameters modal lists the available parameters with column and values information.
2. Select the parameters you want to share.

   You can use Select All or Select None if multiple parameters are available.
3. [Optional] Enable the toggle to include null or empty values in shared parameters.

   If enabled, null or empty results from the AI query are included in the shared parameters, allowing other visuals to filter for missing or empty data.
4. Click SHARE PARAMETERS to apply the changes.

   You can click Cancel to close the modal without applying parameter sharing.

Shelves for AI visual

Overview of shelves for Cloudera Data Visualization AI visual.

The AI visual supports the following shelves:

Embeddings

The Embeddings shelf is designed for the vector database comparisons. The user input is vectorized and compared to the defined embeddings to retrieve rows containing the most similar embeddings. The first field that contains results is returned.

Add a Dimension or Measures vector field that contains embeddings.

This is a mandatory shelf that accepts multiple fields.

Embedding Context

The Embedding Context shelf allows you to add fields that you want to send as part of the chat prompt in a formatted manner, incorporating all embedding context data.

This is a mandatory and sensitive shelf that accepts multiple fields. For information on how to mark a dataset field as 'Sensitive' to restrict its usage in specific visual shelves, see *Editing dataset fields*.

Context Aggregates

The Context Aggregates shelf allows you to include aggregated fields that summarize data for similar records in the AI visual. You can use this shelf to enhance data insights by providing more accurate responses to questions involving quantities, distributions, or comparisons.

Note:

Only aggregated fields are supported. Adding non-aggregated fields will result in an error when refreshing the visual or sending a question.

Aggregation is performed on rows that share the same Dimension, Tooltip, and Vector fields. By including Context Aggregates, you ensure that numeric values (such as counts, sums, or averages) are calculated across similar records, rather than returning individual row-level results that may not provide meaningful insights.

Tooltip

The Tooltip shelf facilitates adding a field for source information to be shown along the displayed results.

Note: The Tooltip shelf is limited to one field for the AI visual type.

Limit

This setting defines the maximum number of data rows that are processed by the visual.

Related Information

Editing dataset fields

Action

Cloudera Data Visualization enables you to include interactive, actionable elements in your dashboards and applications. You can use the Action visual to provide access to external sources of information or to internal search within your dataset.

About this task

The three main use cases of the Action visual are:

- **Linking**

  You can use Action to navigate to an external website or another application, and to move between different parts of your dashboard.
- **Search**

  You can provide a shortcut to the search pane of your current dataset(s).

  Note: Before using search in Cloudera Data Visualization, you must enable the feature in the Site Settings. For instructions, see *Enabling NLS in Site Settings*.
- **Email**

You can share your visuals through email by sending out your dashboard in an email.

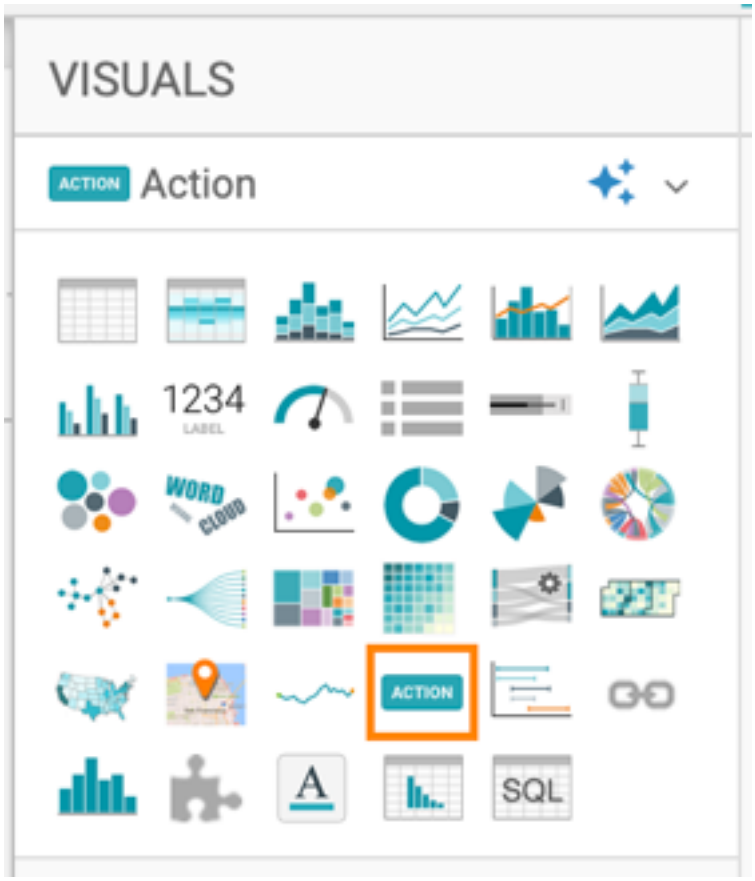
Note: Before using this function, you must define your email settings in the Site Settings to allow users to share visuals through email. For instructions, see *Email settings*.

The following steps demonstrate how to create an Action visual based on the Restaurant Inspection SF dataset.

Procedure

1. Start a new visual based on the Restaurant Inspection SF dataset.
   For instructions, see *Creating a visual*.

   By default, a table visual is displayed that shows all fields in the dataset.
2. In the VISUALS menu, find and click the Action icon.



   The shelves of the visual change.
3. From the Type drop-down list, choose the action element you want to add to your dashboard.
   You can select Button or Image.

   - If you choose Button, the remaining shelves are the following:
     Label

     Add the text you want to show on the button. The default value is ‘Click me!’.

     Action

     Select the action you want to invoke with the button. The options available in the drop-down list are Open URL, Search Dataset, and Send email.

     - If you choose Open URL, you can add the address of the webpage in the URL to open field.
     - If you choose Search Dataset, you can select an additional checkbox whether you want to include all datasets on the sheet.

- If you choose Send email, you do not have to fill in additional fields.

Button Type

Select the button type that you want to add. The options available in the drop-down list have different designs. The default value is ‘Primary’.

Button Size

Select the size of the button from the drop-down list. The default value is ‘Normal’.

- If you choose Image, the remaining shelves are the following:

Image URL

Add the URL of the image file that you want to add to your Action visual.

Action

Select the action you want to invoke with the button. The options available in the drop-down list are Open URL, Search Dataset, and Send email.

- If you choose Open URL, you can add the address of the webpage in the URL to open field.
- If you choose Search Dataset, you can select an additional checkbox whether you want to include all datasets on the sheet.
- If you choose Send email, you do not have to fill in additional fields.

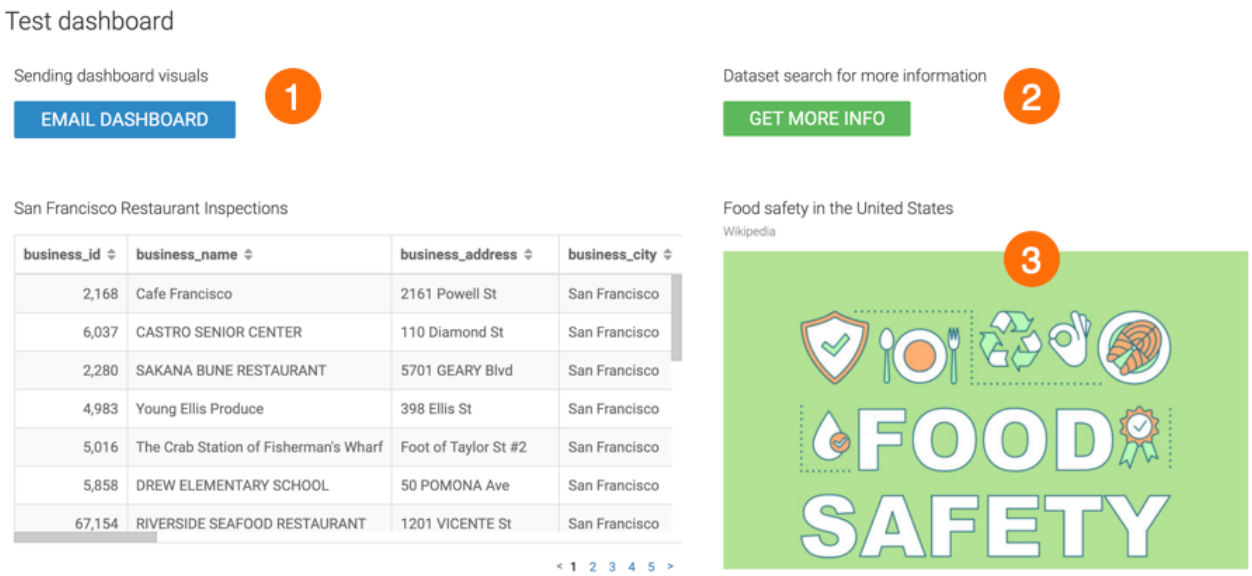
4. Click REFRESH VISUAL.

   The updated Action visual appears.

5. At the top left corner of the Dashboard Designer, click SAVE.

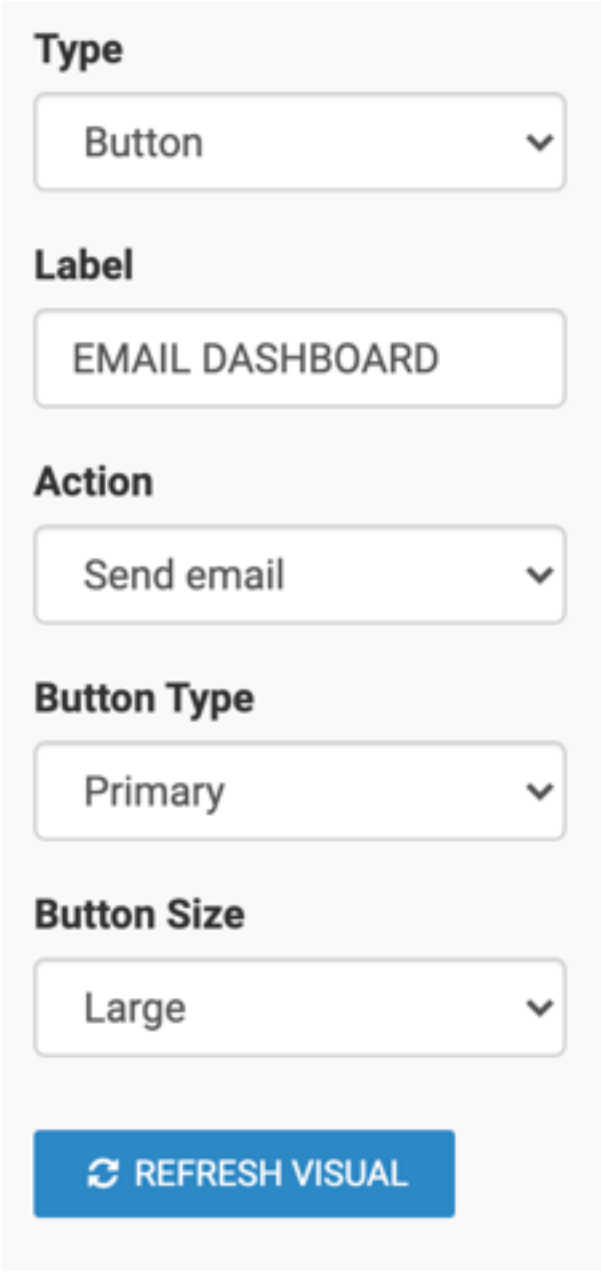
Example

In this example, you can see a dashboard with three action visuals and a table visual:



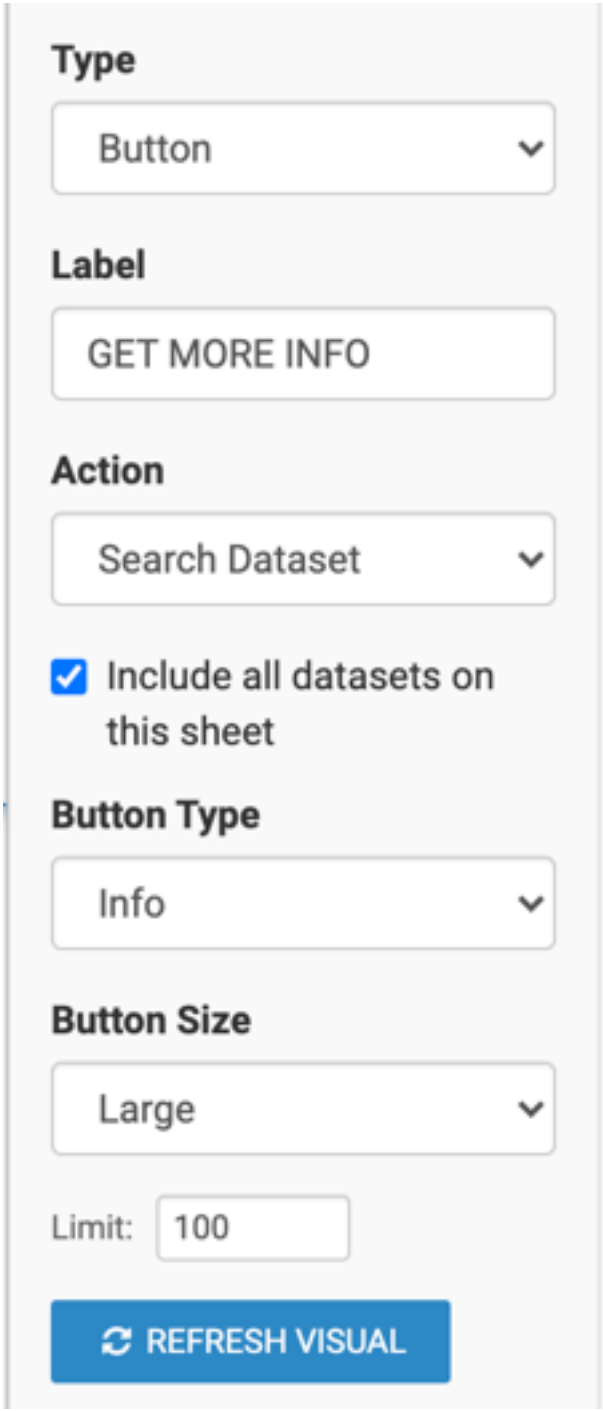
1. Action button for email

   The blue Action button is a large Primary button that you can use for sending the dashboard to an email address:



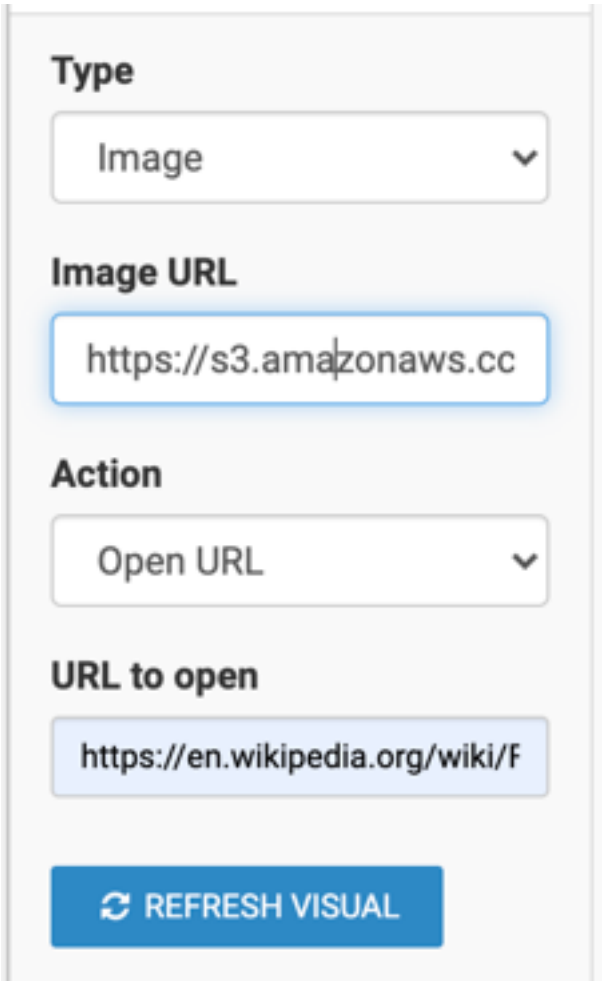
2. Action button for search

   The green Action button is a large Info button that you can use for searching the dataset:



3. Action clickable image with link

   The Action image is clickable and it points to a Wikipedia page about food safety in the US:



Related Information

Enabling NLS in Site Settings
Email settings
Creating a visual

Search

The purpose of the Search visual is to include search functionality in embedded dashboards. It enables you to place a data search button directly on the dashboard.

For instructions on how to create a Search visual, see *Creating a search visual*.

For more information, see *Using the search visual in a dashboard*.

Related Information

Creating a search visual
Using the search visual in a dashboard

Timeline

Cloudera Data Visualization allows you to create Timeline visuals. Timelines are graphical representations of activities/events displayed against time.

About this task

In a Timeline visual, each activity/event is represented by a bar of a time period along a suitable time scale. You need two different timestamps to define the projection of the time period. The starting point of the bar reflects the start date, the end point reflects the end date, and the length of the bar corresponds to the duration of the activity.

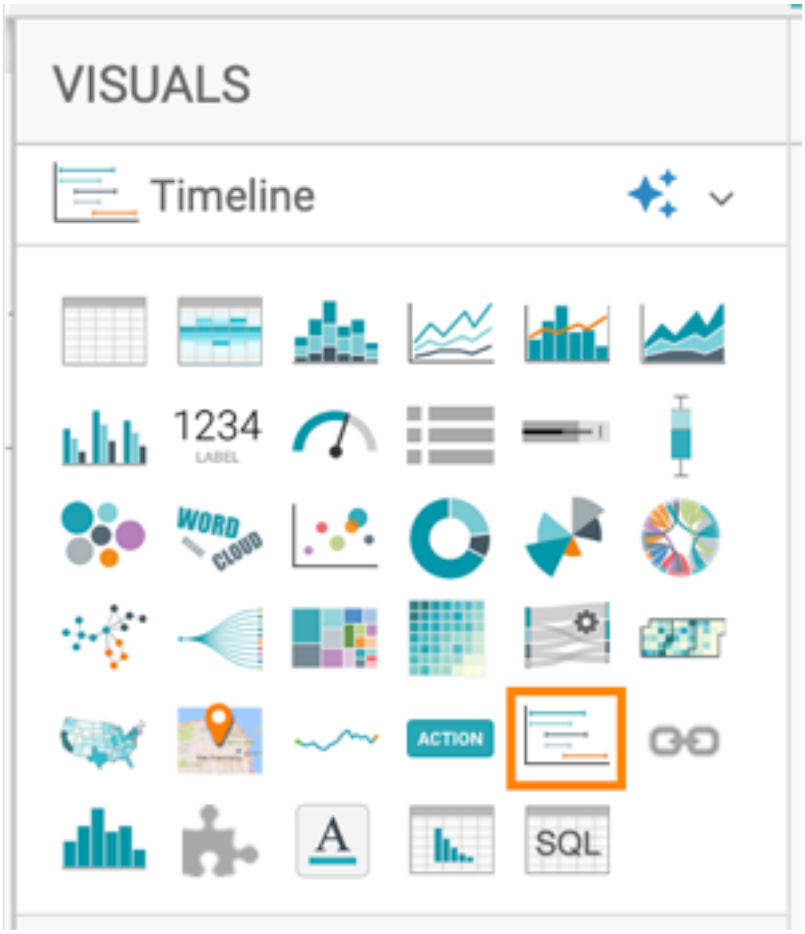
The following steps demonstrate how to create a Timeline visual based on the Earthquake Data January 2019 dataset. You can create simple linear representations or more complex visualizations. For an overview of the shelves that specify this visual, see *Shelves for Timelines*.

Procedure

1. Start a new visual based on the Earthquake Data January 2019 dataset.

   For instructions, see *Creating a visual*.

   By default, a table visual is displayed that shows all fields in the dataset.
2. In the VISUALS menu, find and click the Timeline icon.

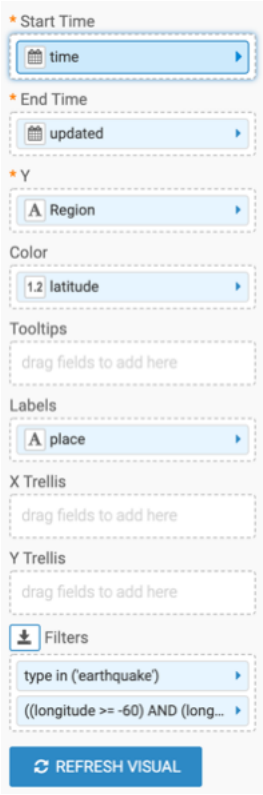


   The shelves of the visual change.
3. Populate the shelves from the available fields (Dimensions and Measures) in the DATA menu.

   Make the following selections to recreate the example Timeline visual:

   - Add ‘time’ to the Start Time shelf.
   - Add ‘updated’ to the End Time shelf.
   - Add ‘Region’ to the Y axis.
   - Add ‘latitude’ to the Color shelf.
   - Specify ‘place’ as a label on the Labels shelf.
   - Define a filter for ‘type’. Select ‘earthquake’.

- Define a filter for ‘longitude’. Only include earthquakes between -60° and 0°.



For further information about the shelves of the visual, see *Shelves for Timelines*.

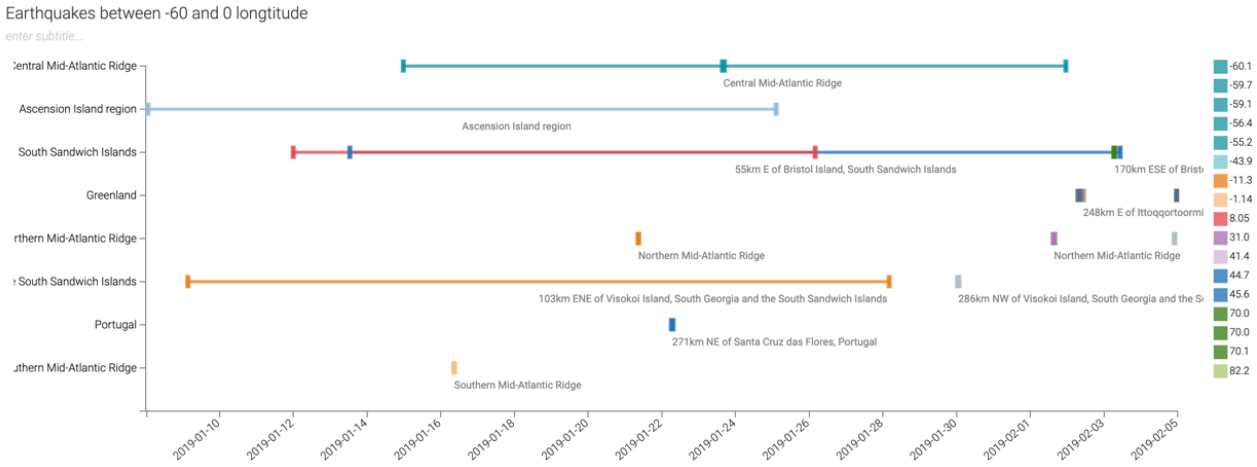
4. Click REFRESH VISUAL.

   The updated Timeline visual appears.

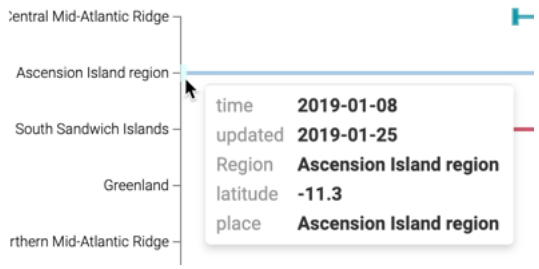
   Legend is on by default, and it appears on the right side of the visuals. If you want to change the legend style or remove it from the visual, see Adding and removing legend.
5. Add the title Earthquakes between -60 and 0 longtitude.
6. At the top left corner of the Dashboard Builder, click SAVE.

Results

Your visual should look like the following:



Details of a particular timeline are shown in tooltips if you hoove over:



Related Information

Creating a visual

Adding and removing legend

Shelves for timelines

Cloudera Data Visualization Timeline visuals have the following shelves.

Start Time

This shelf is required and it is limited to one field.

You can add a time dimension from the DATA menu to specify the starting point of the analyzed timescale.

Note: The Display Format field property defined for the Start Time shelf will be used for the End Time shelf as well.

End Time

This shelf is required and it is limited to one field.

You can add a time dimension from the DATA menu to specify the ending point of the analyzed timescale.

Y

This shelf is required and it accepts multiple fields.

You can add one or more dimensions that uniquely identify the elements on the Y axis.

Color

This shelf is optional and it is limited to one field.

You can define how colors are used in the visual.

Tooltips

This shelf is optional and it accepts multiple fields.

You can add measures from the DATA menu for additional information in the graph's tooltips.

Labels

This shelf accepts multiple fields.

You can define what information should appear in the graph as data labeling.

X Trellis

This shelf is optional and it accepts multiple fields.

You can use a dimension field from the DATA menu to trellis the visual into columns.

Y Trellis

This shelf is optional and it accepts multiple fields.

You can use fields from the DATA menu to trellis the visual into rows.

Filters

This shelf is optional and it accepts multiple fields.

You can add measures and dimensions from the DATA menu to set exclude or include criteria for the visual.

External link visuals

Cloudera Data Visualization enables you to create External Link visuals.

About this task

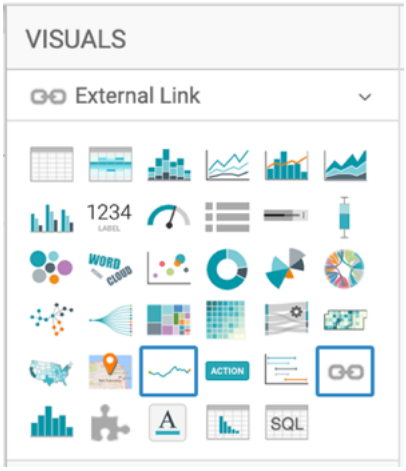
A link visual enhances your applications by provding access to external sources of information. These sources may either be static URL addresses, or they can be dynamic serach results to which you pass a name of a variable used in the dataset.

The following steps demonstrate how to create a new External Link visual for the World Life Expectancy dataset. In this exercise, the variable country is passed to the Wikipedia site.

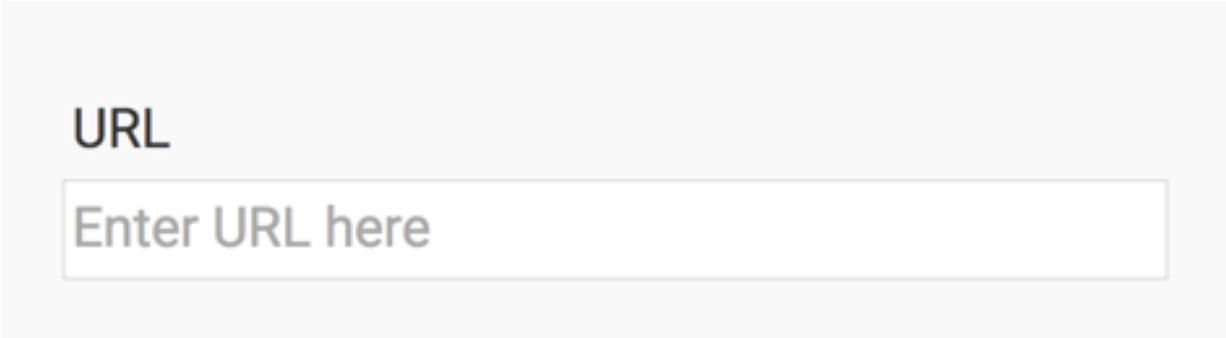
For an overview of shelves that specify this visual, see *Shelves for external links*.

Procedure

1. Start a new visual based on the World Life Expectancy dataset.
   For instructions, see *Creating a visual*.
2. Select Link in the VISUALS menu.



The shelves of the visual changed. The only shelf is URL.

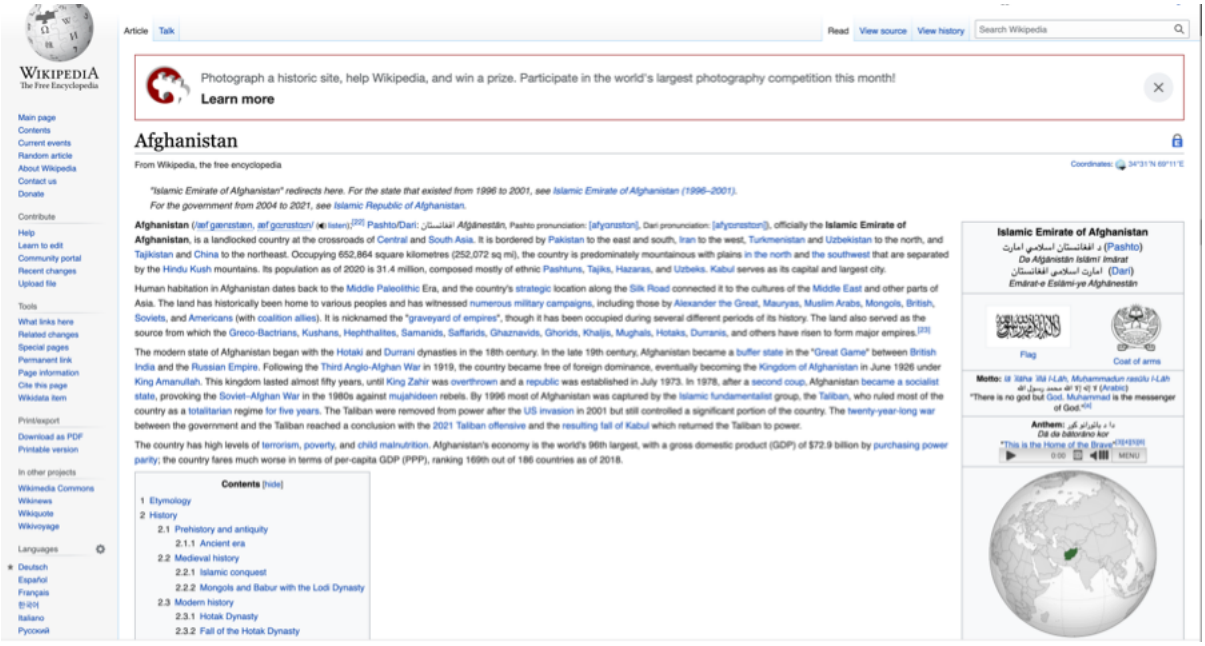


3. In the URL field, enter the following text:

```
https://en.wikipedia.org/wiki/Afghanistan
```

4. Click REFRESH VISUAL.

   The page specified in the URL appears.



5. Click the (pencil icon) next to the title of the visualization to edit it, and change the title to World Population - C ountry UTL.
6. At the top left corner of the Dashboard Designer, click SAVE.

Setting up dynamic external links

You need to perform extra set-up steps if you want to create a visual with dynamic links.

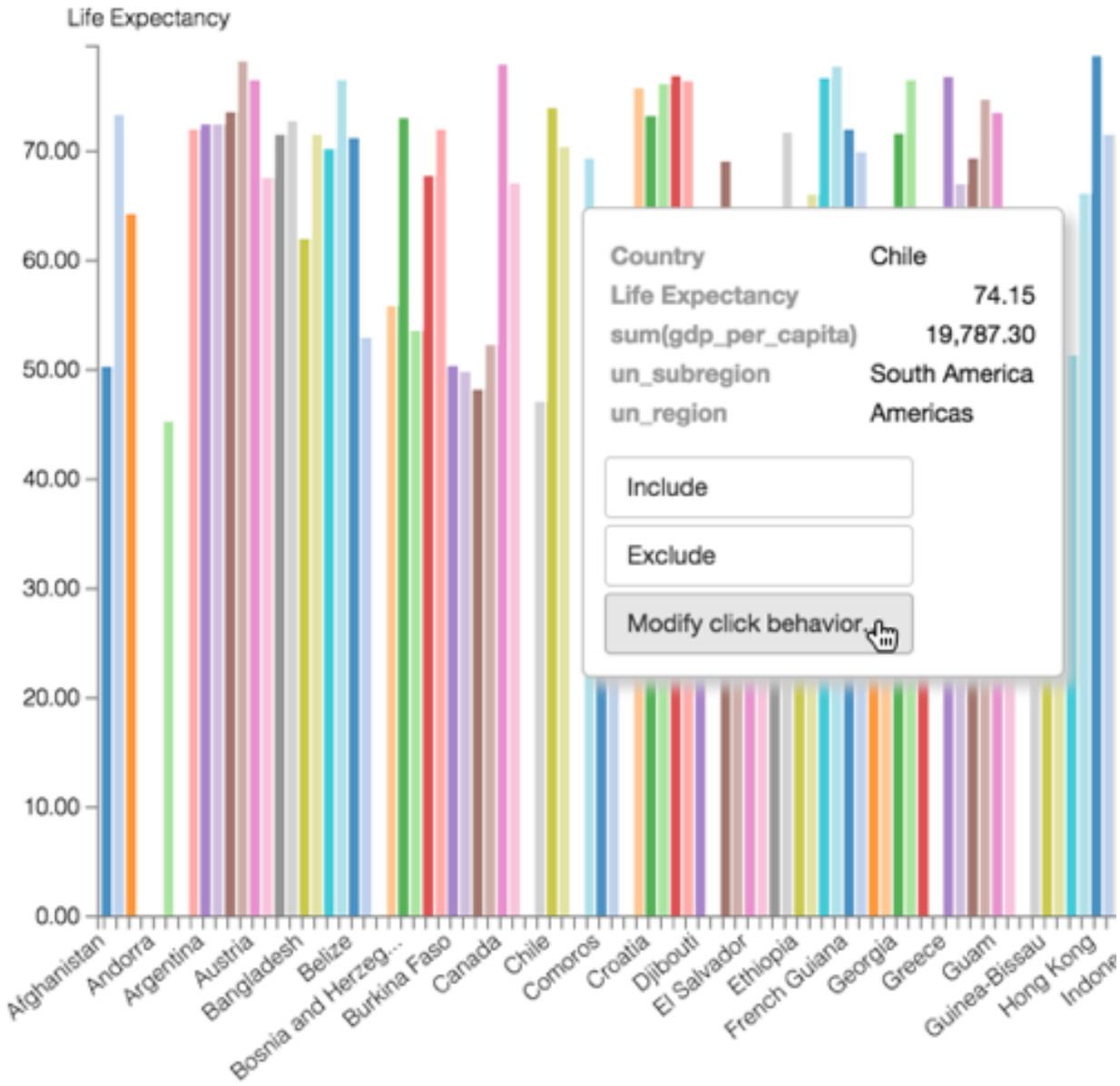
About this task

To use a visual to see the Wikipedia page for a specific country, follow these steps:

Procedure

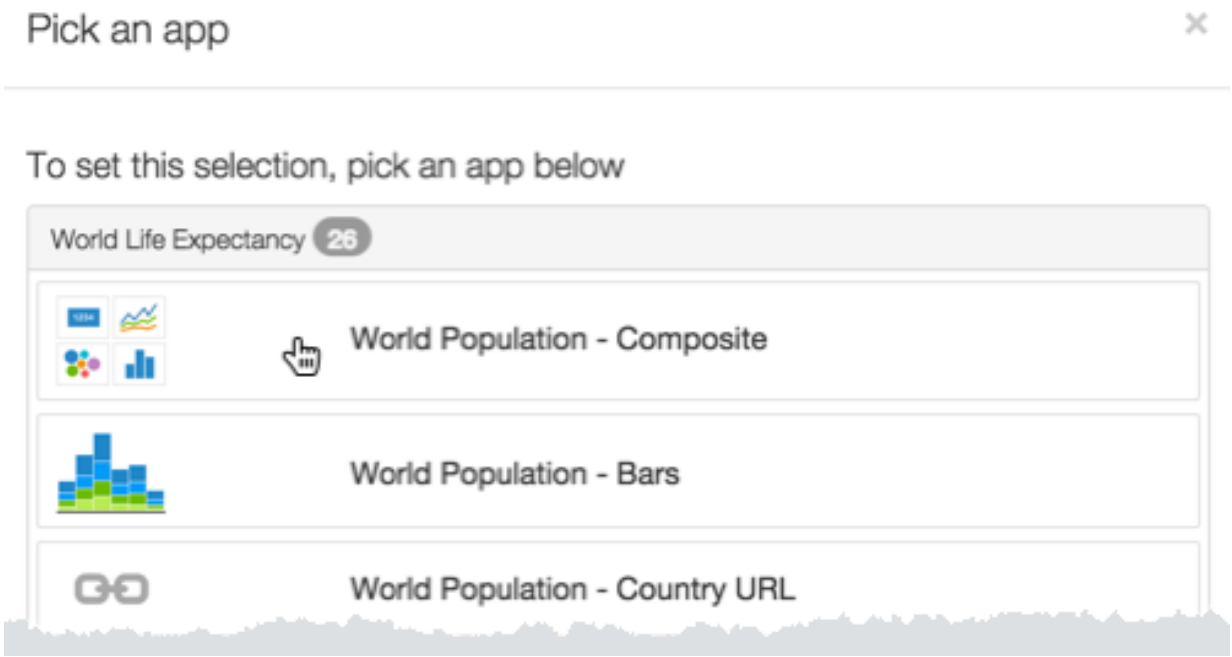
1. Create a new application for the World Life Expectancy dataset.
2. Add the World Population - Bars Bars visual that you created in *Bars*.
3. Add World Population - Country UTL as the title of the application.
4. Add World Population as the title of the active menu tab.

5. At the top left corner of the App Designer, click SAVE.
6. In the application, click the Bars visual to open it in a new tab.
7. Click EDIT to open the visual in Edit mode.
8. Click the visual and select Modify click behavior.



9. In the Pick an app modal window, under To set this selection, pick an app below, select World Life Expectancy.

10. After the modal window changes, select World Population - Composite.



11. At the top left corner of the Dashboard Designer, click SAVE.

12. While in Edit mode, ensure that the settings work:

a) Left-click the visual.

b) Select Test click behavior.

Results

In the World Population - Composite app, when you click on the specific bars that represent a country, the Wikipedia page changes to match the country.

Shelves for external links

Overview of shelves for Cloudera Data Visualization External Link visuals.

An External Link has only one shelf:

URL

Add a URL address of an external site.

This shelf accepts multiple fields, but plots a map with the highest possible resolution.

Histograms

You can create Histogram visuals in Cloudera Data Visualization. Histograms enable the quick visual analysis of numerical data distribution. They provide an estimate of the frequency distribution of a continuous quantitative variable, accomplished by splitting it into consecutive, non-overlapping intervals, referred to as "buckets".

About this task

Note: Histograms can only be plotted for bare metrics, not aggregates. For example, you can plot a histogram for population, not for avg(population).

To work around this limitation, first save the calculated metric as derived data, and then use this field as a measure in a histogram. See *Derived data*.

When you change the number of buckets (the default is 10), Cloudera Data Visualizationn determines the data range and splits it equally among the buckets, groups the metric into these buckets, and plots vertical bars to represent it.

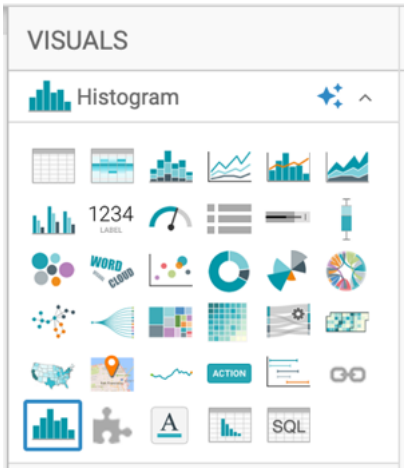
Histograms support both normalized and cumulative forms, and may be used in a trellis formation.

The following steps demonstrate how to create a new histogram visual on the World Life Expectancy dataset.

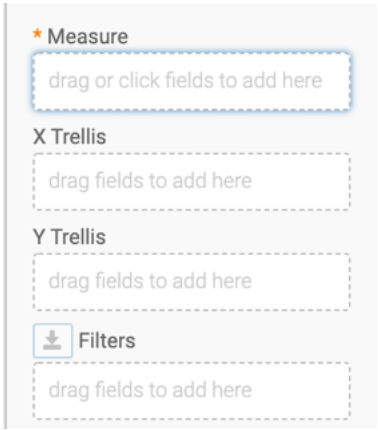
For an overview of shelves that specify this visual, see *Shelves for histograms*.

Procedure

1. Start a new visual based on the World Life Expectancy dataset.
   For instructions, see *Creating a visual*.
2. In the VISUALS menu, select Histogram.

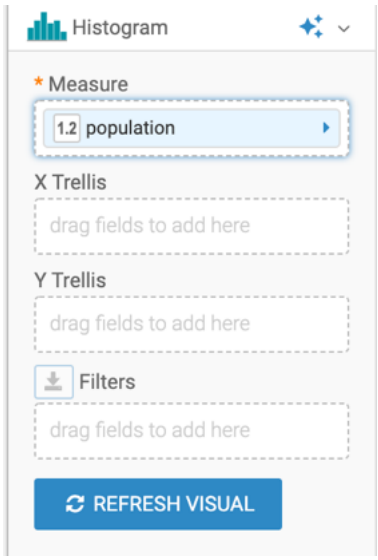


The shelves of the visual changed. You can see: Measure (mandatory), X Trellis, Y Trellis, and Filters.



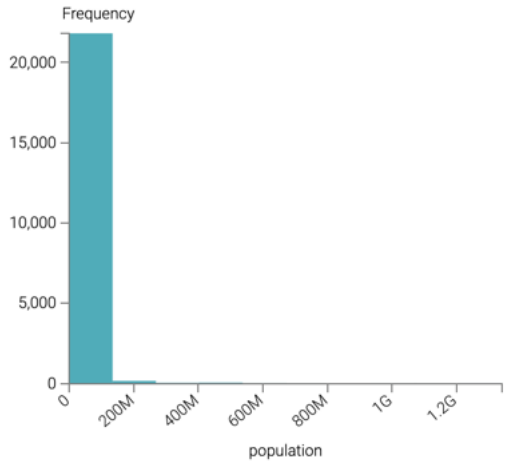
3. To show specific items, populate the shelves from the available fields (X Trellis, Y Trellis) in the DATA menu.

   Under Measures, select population and drag it to the Measures shelf.



4. Click REFRESH VISUAL.

   The default histogram visual appears, dividing the data into 10 buckets. As you can see, most the data is in the first bucket, and this visual is not very enlightening.



5. On the Filters shelf, add several Dimensions and Measures from the DATA menu.

   This enables you to dynamically control the data input and discover the data at a more granular level.

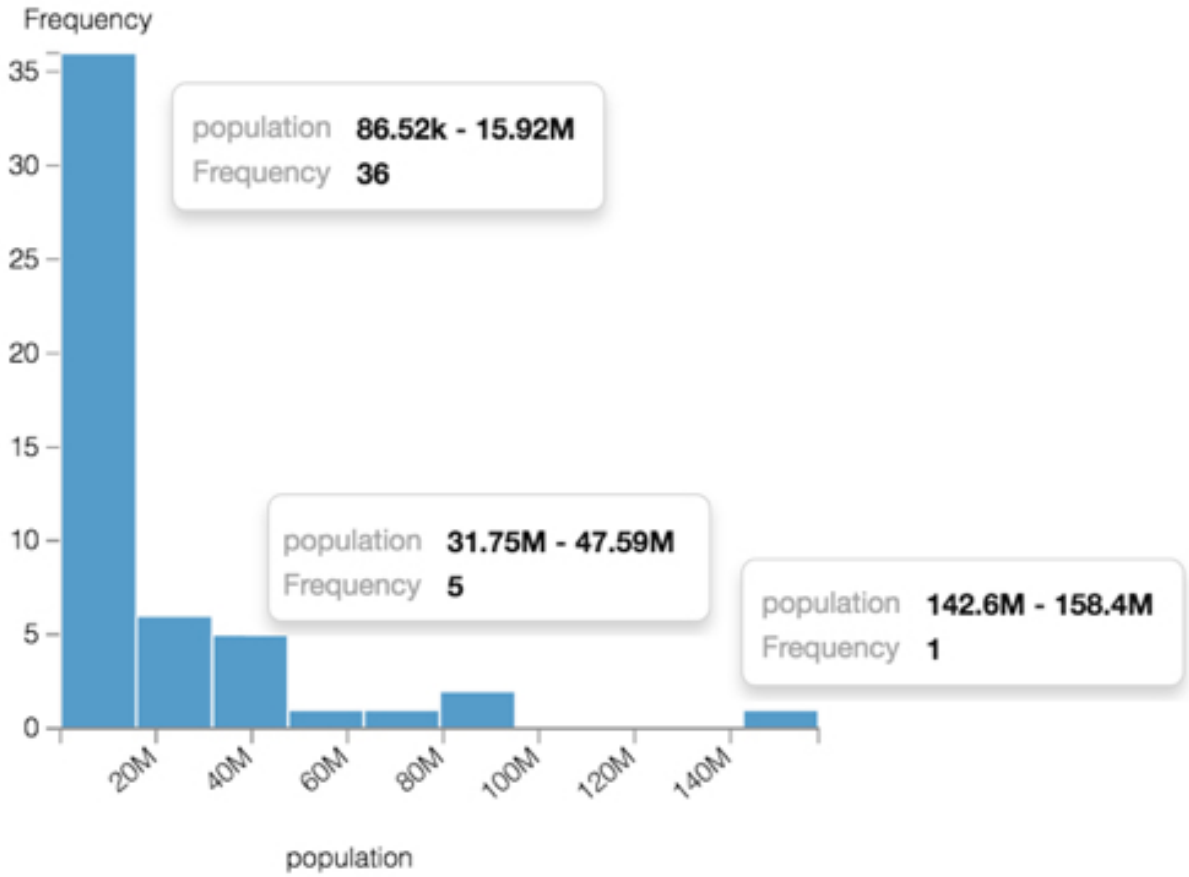
   For example, from Dimensions, drag year to the Filters shelf and select the year 2010.

   You can also add un_region to the Filters shelf and select Africa.

6. Click REFRESH VISUAL.

   The shape of the diagram has changed. Hovering the mouse over the histogram shows the data ranges and frequency for each bucket of the histogram.

   You can see that the filtering options significantly reduced result set.



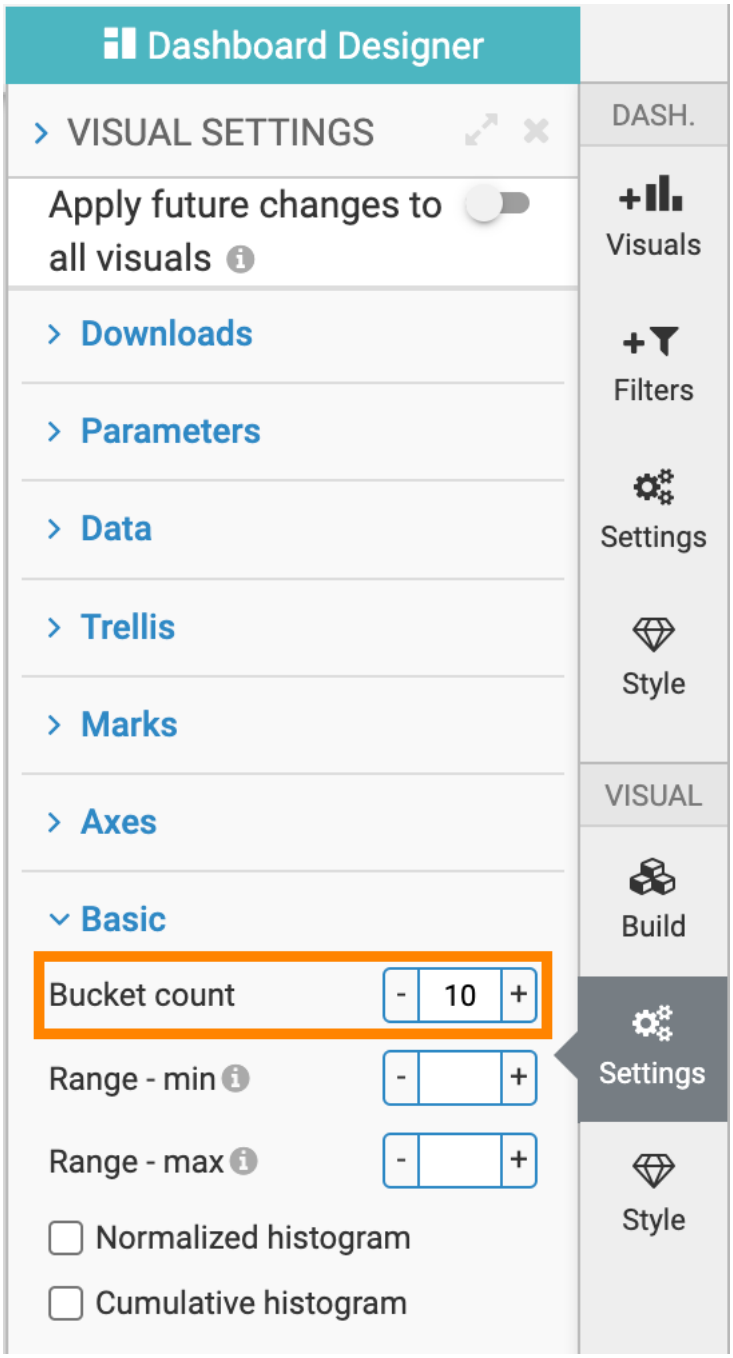
7. Change the title to World Population - Histogram.
   [Optional] You can also add a brief description of the visual as a subtitle below the title of the visualization.
8. Click SAVE at the top left corner of the Dashboard Designer.

Specifying the bucket count

The default histogram displays data in 10 buckets. You can easily adjust this setting to control the granularity of your data display.

Procedure

1. On the right side of the Dashboard Designer, go to the VISUAL menu and click Settings.
2. In the Settings menu, click Basic to expand its options.
3. Find the Bucket count field and use the selector to increase or decrease the number of buckets as desired.



Example

Increasing the bucket count results in narrower buckets, meaning each bucket represents a smaller range of values. This can provide a more detailed view of the data distribution. Conversely, reducing the bucket count widens each bucket, simplifying the visualization and giving a broader overview.

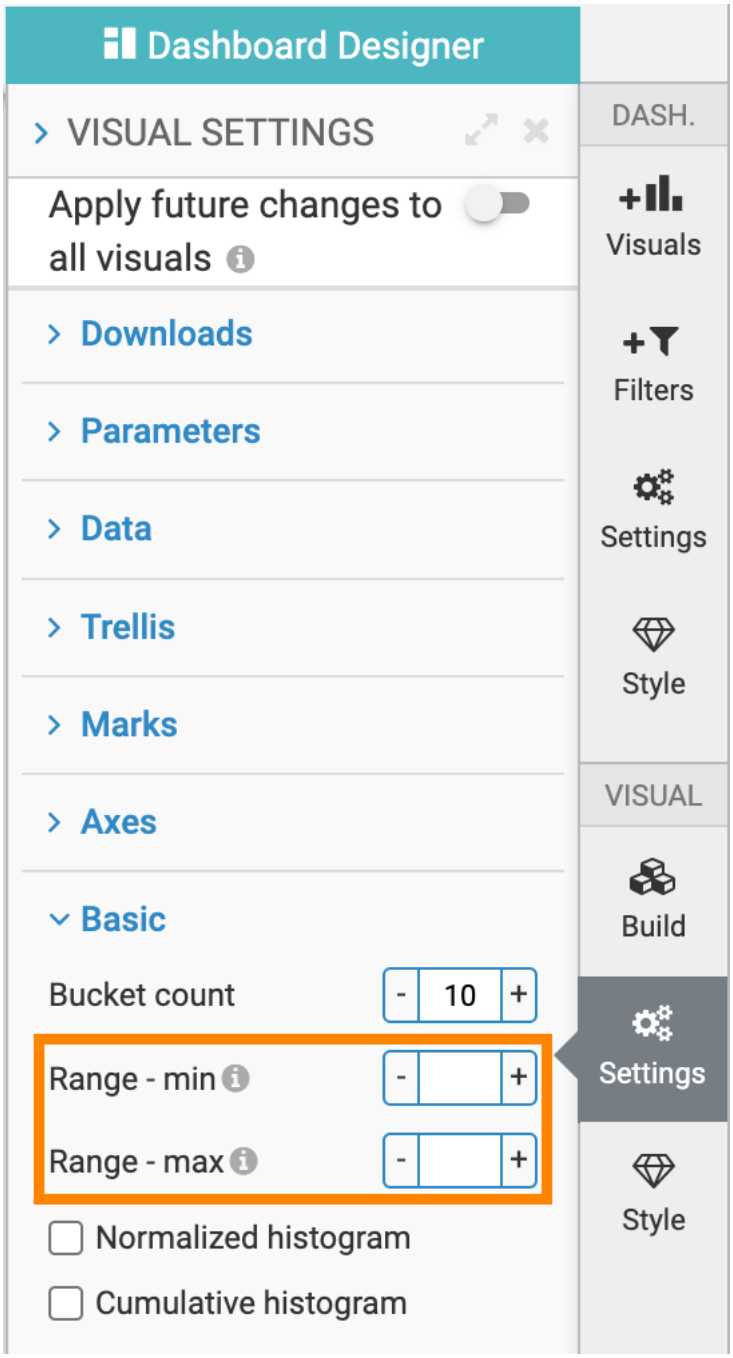
The following example shows a histogram with 10 buckets compared to a histogram with 20 buckets. You can see how doubling the number of buckets reduces the range of values in each bucket by half, providing a more granular look at data distribution.

Specifying data range

You can specify the exact portion of your dataset to visualize, allowing you to focus on specific data intervals. This is especially useful for narrowing in on relevant subsets without including outliers or unnecessary values.

Procedure

1. On the right side of the Dashboard Designer, go to the VISUAL menu and click Settings.
2. In the Settings menu, click Basic to expand its options.
3. Define the range of values to focus on for your histogram.
   a) In the Range - min field, enter the minimum value for your range.
      This will be the starting point for the histogram's x-axis.
   b) In the Range - max field, enter the maximum value for your range.
   This will define the endpoint for the histogram's x-axis.



After setting these values, the histogram will automatically update to display only the data within the specified range. You can adjust the values as needed to explore different parts of your dataset.

Example

Using the range setting, you can easily control the focus of your histogram, allowing you to isolate and analyze specific intervals of your data.

The following example shows a histogram with a wider range compared to a histogram with narrower range. You can see how adjusting the range changes the histogram's appearance by displaying more or fewer data points.

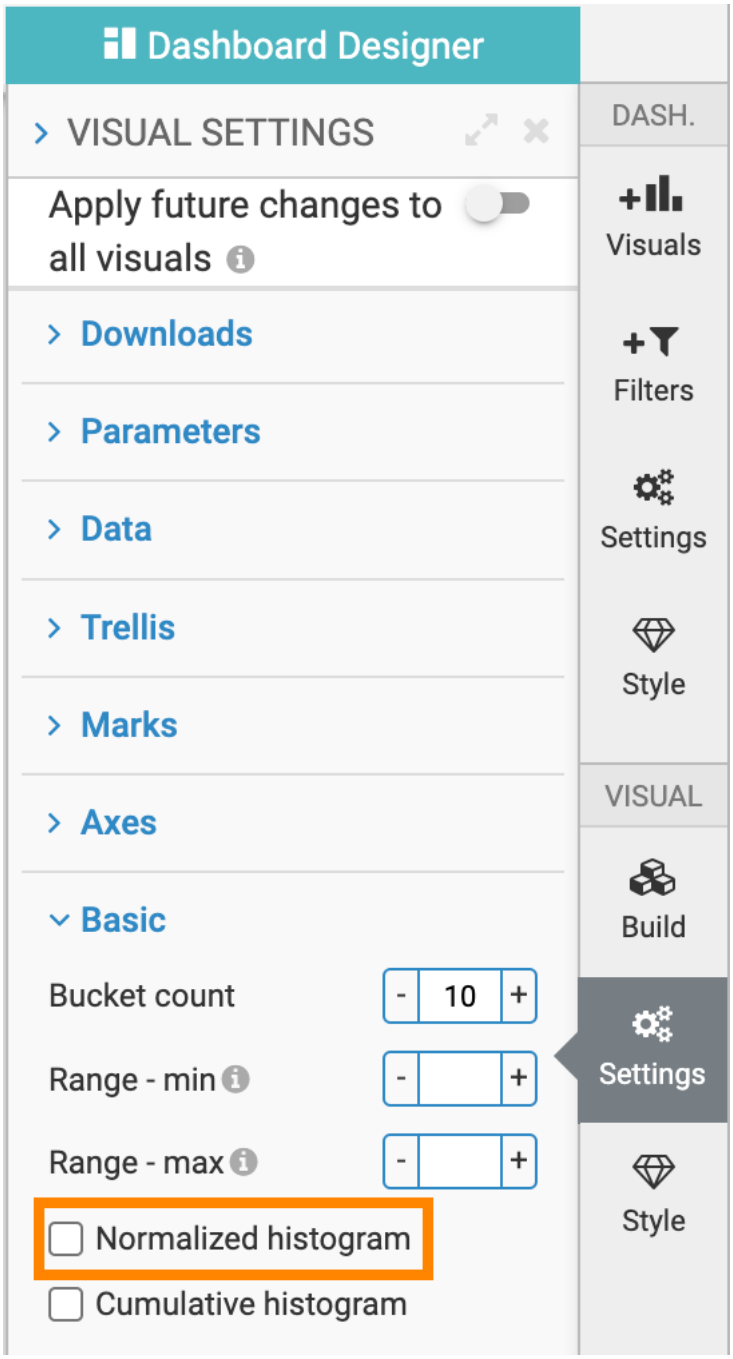
Calories across manufacturers
enter subtitle...
Frequency
20
10
0
50
60
70
80
90
100
110
120
130
140
150
160
Calories
Calories across manufacturers
enter subtitle...
Frequency
20
10
0
80
90
100
110
120
130
140
150
160
Calories

Normalizing the histogram

By default, the vertical axis of a histogram represents the count of values that map to a particular bucket. To represent these values as percentages of the total, you can normalize the histogram. This adjustment changes the histogram counts so they sum to 1, with each bar representing the proportionate frequency.

Procedure

1. On the right side of the Dashboard Designer, go to the VISUAL menu and click Settings.
2. In the Settings menu, click Basic to expand its options.
3. Select the Normalized histogram option.



Example

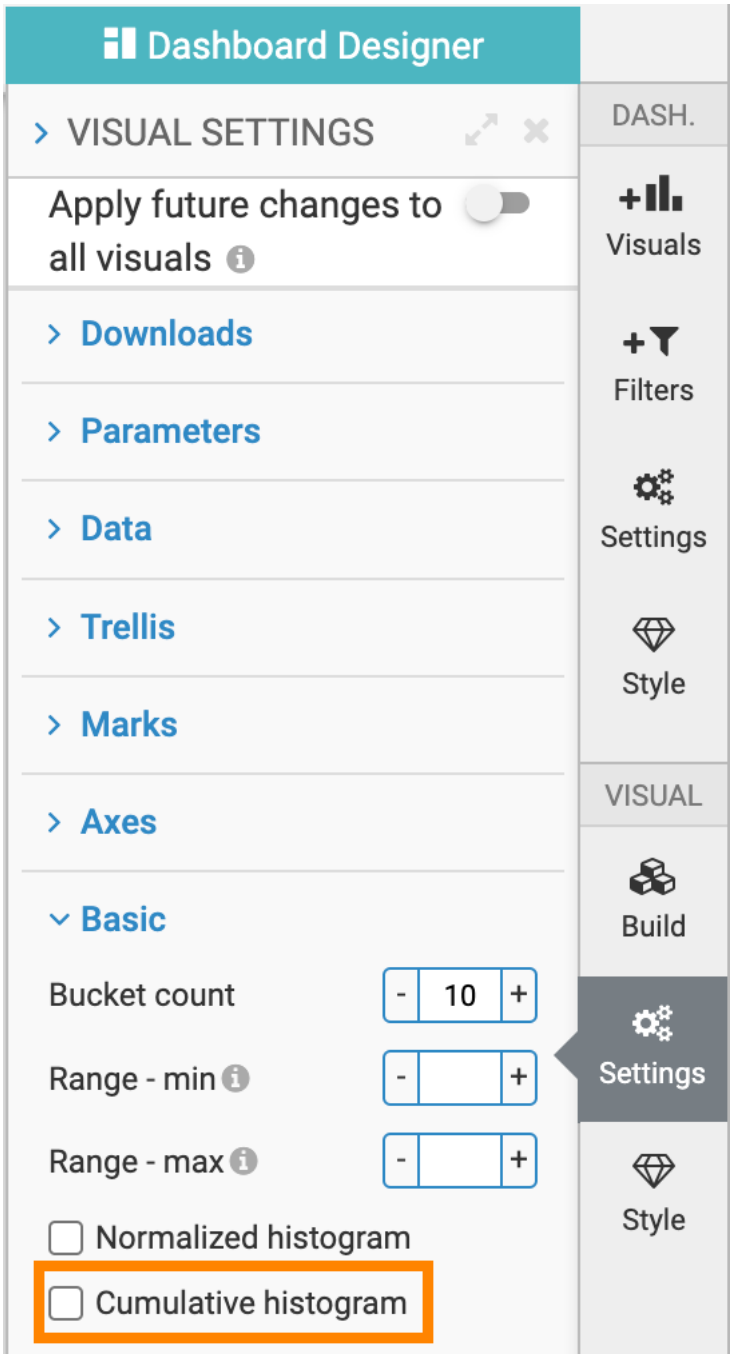
The following image compares two versions of a histogram. In the standard histogram (top), the vertical axis and tooltips display raw counts. In the normalized histogram (bottom), the vertical axis and tooltips show values as percentages, indicating the relative frequency of each bucket.

Using the cumulative option

By default, a histogram displays the count or frequency for each bucket individually. Enabling the cumulative option adjusts the histogram so that each bucket's count or frequency is added to the previous total. As a result, the right-most bucket shows the total count or 1 (100%) if the histogram is normalized.

Procedure

1. On the right side of the Dashboard Designer, go to the VISUAL menu and click Settings.
2. In the Settings menu, click Basic to expand its options.
3. Select the Cumulative histogram option.



Example

The following image shows the difference in the vertical axes. In the lower image, the cumulative values build up towards 1 (or 100%) on the right side of the graph, reflecting the running total.

Shelves for histograms

Overview of shelves for Cloudera Data Visualization Histogram visuals.

A Histogram visual has the following shelves:

X Trellis

Use a Dimension field from the DATA menu to trellis the visual into columns.

This shelf accepts multiple fields.

Y Trellis

Use a Dimension field from the DATA menu to trellis the visual into rows.

This shelf accepts multiple fields.

Measure

Add a Measure field from the DATA menu to generate a histogram of its values.

This shelf accepts a single field.

Filters

Add Dimension or Measure fields from the DATA menu to include or exclude categorical or value criteria.

This shelf accepts multiple fields.

Extensions

In Cloudera Data Visualization, you can use extension visuals to add supporting textual, image, or behavioral aspects to your dashboard.

About this task

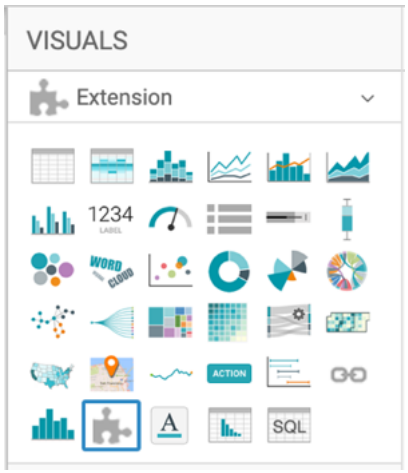
The following steps demonstrate how to create a simple extension visual with text content and minor HTML formatting.

Note: You must have system-level Manage styles and settings privilege to create Extension visuals. See *Setting system privileges* for details.

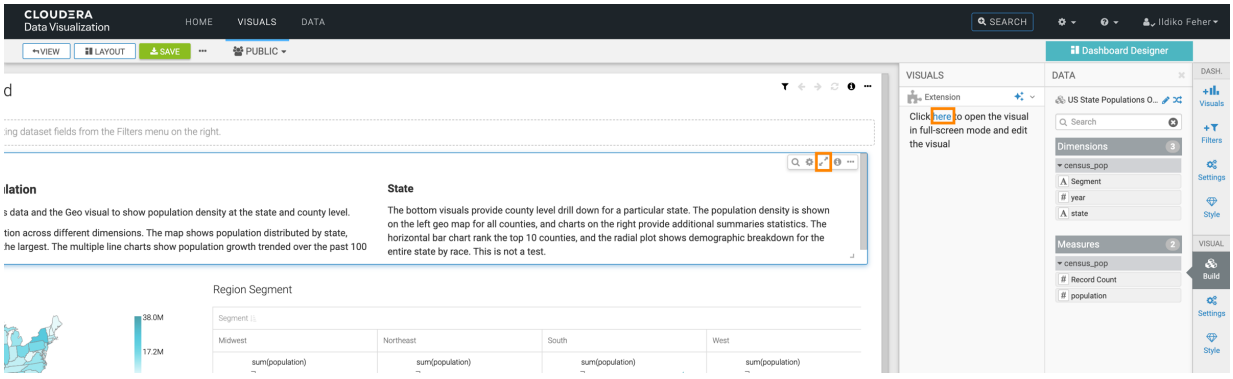
Procedure

1. Start a new visual based on the US State Populations Over Time dataset.

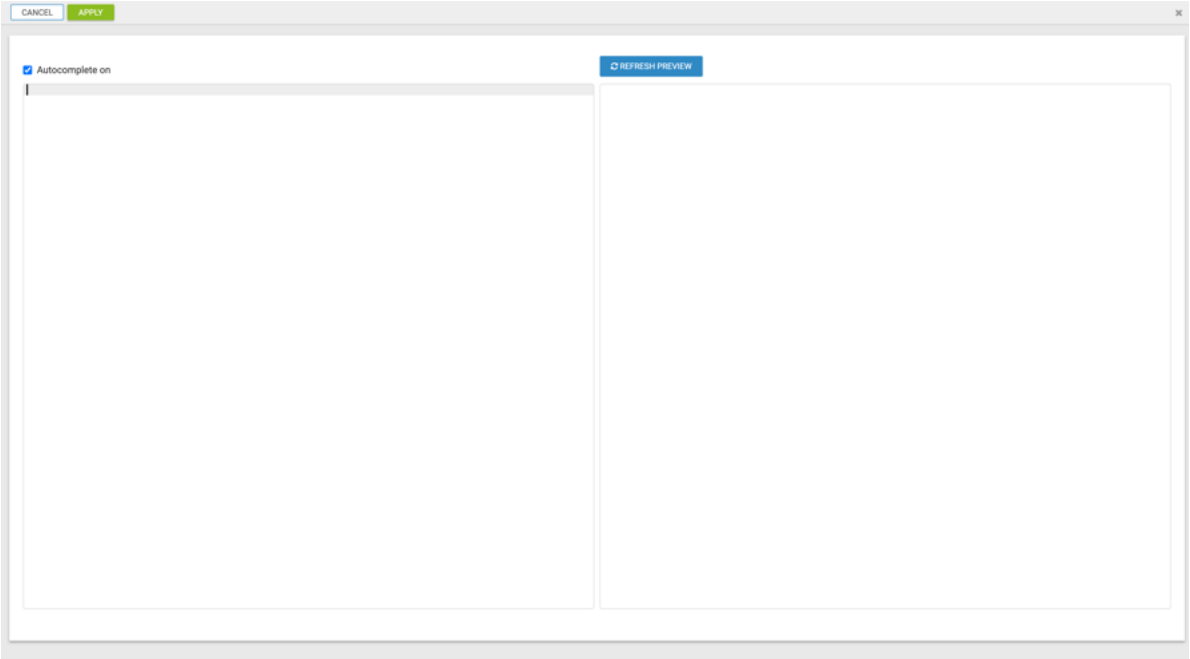
   For instructions, see *Creating a visual*.
2. Select Extension in the VISUALS menu.



   The shelves of the have visual changed. This visual type has a single text box for entering and editing text or markup.
3. Click the (expand) icon or the link in the Dashboard Designer to open the HTML editor.



   The editor modal window has two columns: one for code and the other one for preview.



4. Paste a code block in the text box for code.

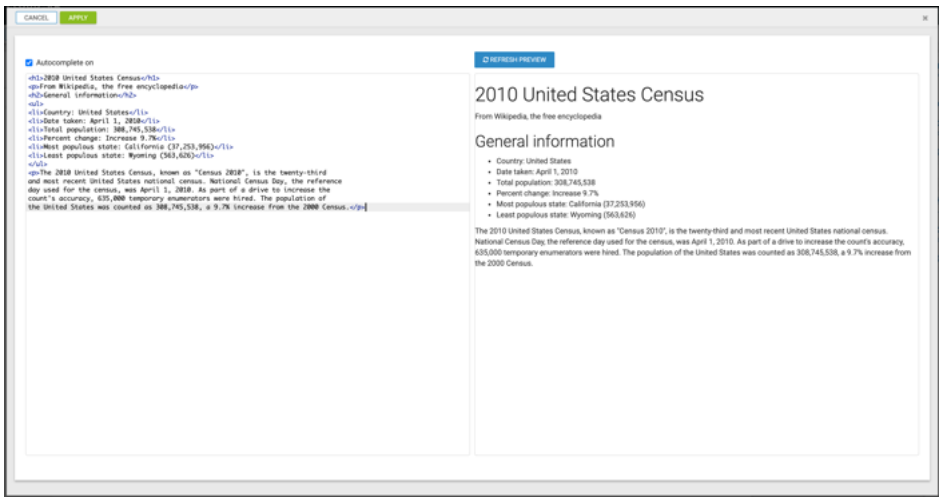
   In this example, the following code block is used:

```
<h1>2010 United States Census</h1>
<p>From Wikipedia, the free encyclopedia</p>
<h2>General information</h2>
<ul>
<li>Country: United States</li>
<li>Date taken: April 1, 2010</li>
<li>Total population: 308,745,538</li>
<li>Percent change: Increase 9.7%</li>
<li>Most populous state: California (37,253,956)</li>
<li>Least populous state: Wyoming (563,626)</li>
</ul>
<p>The 2010 United States Census, known as "Census 2010", is the twenty-
third
and most recent United States national census. National Census Day, the r
eference
day used for the census, was April 1, 2010. As part of a drive to increase
 the
count's accuracy, 635,000 temporary enumerators were hired. The population
 of
the United States was counted as 308,745,538, a 9.7% increase from the
2000 Census.</p>
```

5. Click REFRESH PREVIEW to view the changes in the preview column.

   The formatted text appears in the preview area of the HTML editor. You can make further edits to the HTML code and check the changes by refreshing the preview again.

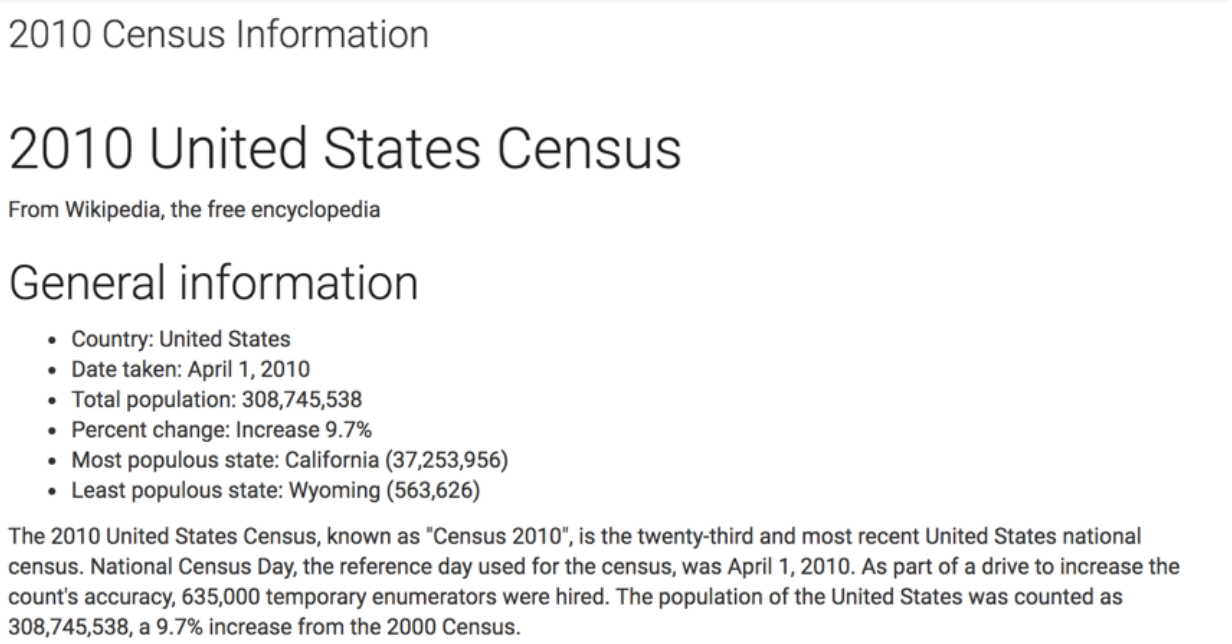
   If you do not need the autocomplete function, uncheck the Autocomplete on option.



6. Click APPLY to save the changes.
   If you do not want to apply your changes, click CANCEL or just close the modal window.
7. Change the title to 2010 Census Information.
   [Optional] You can also add a brief description of the visual as a subtitle below the title of the visualization.
8. Click SAVE at the top left corner of the Dashboard Designer.

Results

Your visual should look something like this:



What to do next

For more sophisticated uses of the Extension visual, see *Examples of extensions*.

Related Information

Setting system privileges
Creating a visual

Examples of extensions

See additional examples for Extension visuals.

In the following examples, the definitions for *Print dashboard* and *Download dashboard as PDF* show how to perform these actions through an Extension visual. They have been implemented together, creating a visual that serves as an action panel within the dashboard.

Print dashboard

The following extension definition builds a (print) button with an icon. You can use this button to print the dashboard in View mode.

```
<div id="extension-print-btn">
  <button class="btn btn-xs btn-primary"
  onclick="window.print();return false;">
  <span class="fa fa-print"></span> Print</button>
</div>
```

Download dashboard as PDF

The following extension definition builds a (download PDF) button with an icon. You can use this button to save the dashboard in PDF format from View mode.

```
<div id="extension-pdf-btn">
  <button class="pdf-download btn btn-xs btn-primary">
  <span class="fa fa-download"></span> Download PDF</button>
  <div class="image-dropdown-submenu"/>
</div>
```

Print dashboard and download PDF

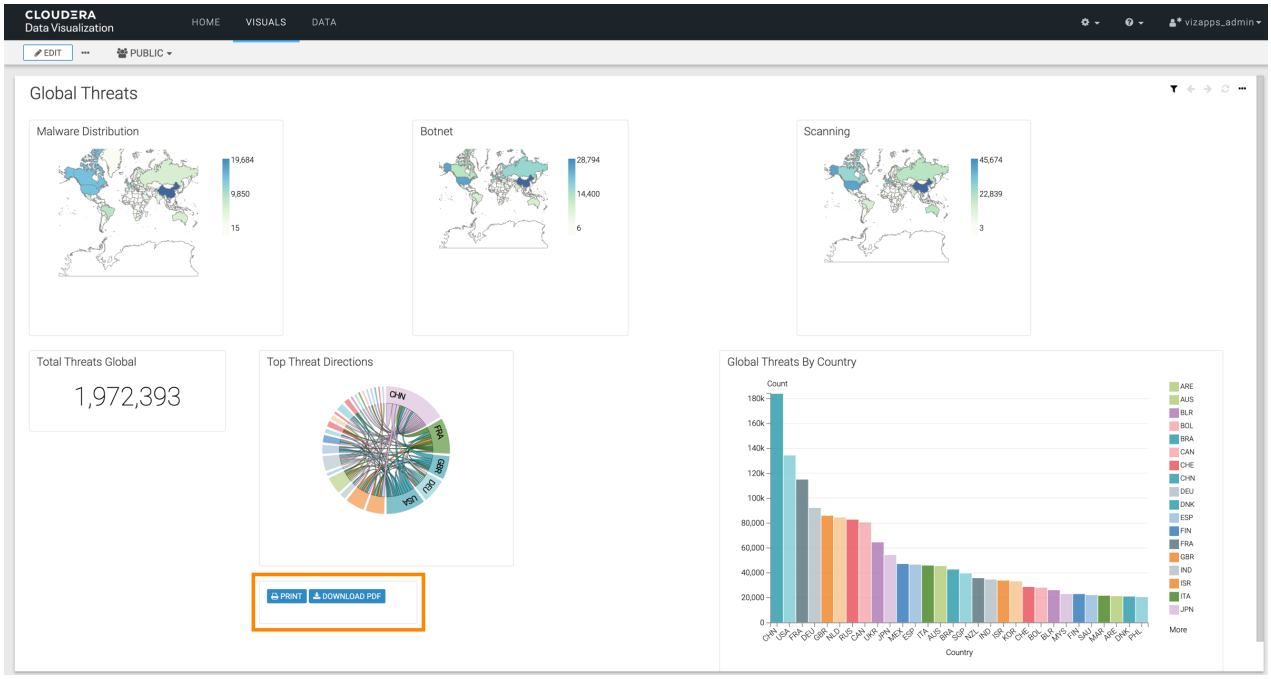
The following definition creates two buttons in an Extension visual:

- Print button with an icon
- Download PDF button with an icon

```
<div id="extension-print-and-pdf-btn">
  <button class="btn btn-xs btn-primary"
  onclick="window.print();return false;">
  <span class="fa fa-print"></span> Print</button>

  <button class="pdf-download btn btn-xs btn-primary">
  <span class="fa fa-download"></span> Download PDF</button>
  <div class="image-dropdown-submenu"/>
</div>
```

When you add this Extension visual to a dashboard, the dashboard will display print and download buttons:



Rich texts

Cloudera Data Visualization enables you to create Rich Text visuals.

About this task

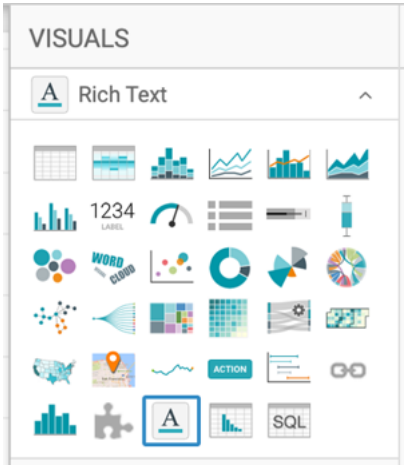
You can use markup code to add supporting textual or visual information to your dashboard. You can insert styled paragraphs, hyperlinks, tables, images, and other elements. These insterted elements do not use any information from the dataset.

Note: You must have system-level Manage styles and settings privilege to create Rich text visuals. For more information, see *Setting system privileges*.

The following steps demonstrate how to create a visual with rich text formatting.

Procedure

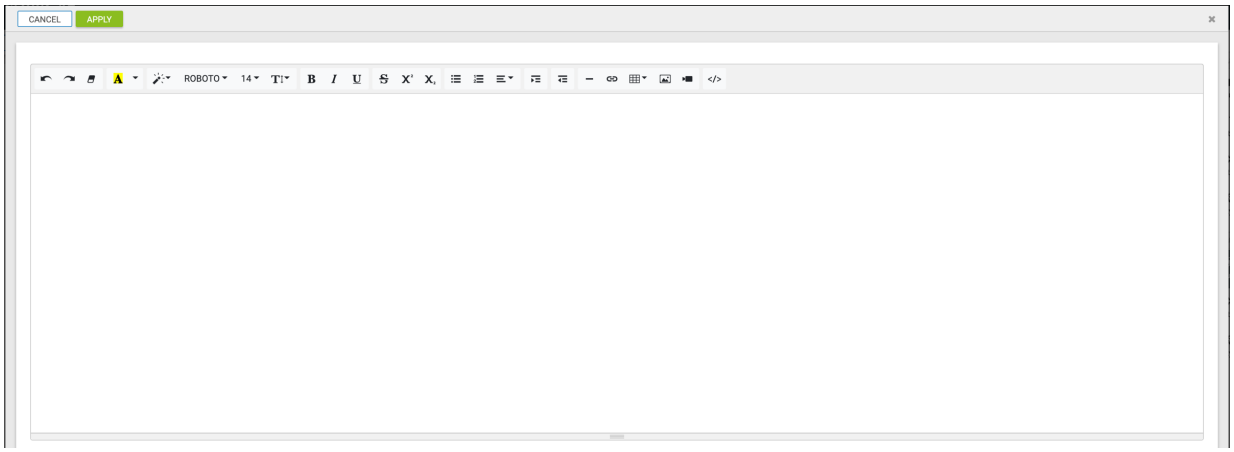
1. Start a new visual based on the World Life Expectancy dataset.
   For instructions, see *Creating a visual*.
2. Select Rich Text in the VISUALS menu.



The shelves of the have visual changed. This visual type does not have any shelves associated with it. Instead, it has a single text box for entering and editing text or markup.

3. Click the (expand) icon to open the Rich Text Editor.

The editor modal window contains a large text box and an editing tool bar.



4. Paste the following text in the text box.

```
In the early 1800's, life expectancy at birth began to increase signific
antly in the countries that industrialized early, but it remained low in
 the rest of the world. Industrialization levels correlated strongly wit
h life expectancy, and lead to notable inequalities in global health dis
tribution. Good health was prevalent in the rich countries, where improv
ements in sanitation, infant/maternal care, and food quantity lead to lo
nger life. Meanwhile, poor health persisted in poorer countries.

This global inequality decreased over the last half of the 20th century. C
ountries that relatively recently suffered from bad health are catching up
 rapidly. Since 1900, the global average life expectancy more than doubled
, increasing from 32.6 years for all reporting countries to 70.0 years.
```

5. You can format the text that you have added to the visual. Enhance the basic text according to your requirements.

   Formatting controls include the following options (from left to right on the editing toolbar):

   - Undo and Redo
   - Remove Font Style
   - Recent Color and More Color selector for Background Color and Text Color

- Style selector for Normal, Blockquote, Code, Header 1 - Header 6, Font Family menu, Font Size menu, and Line Height menu
- Character formatting: Bold, Italic, Underline, Strikethrough, Superscript, and Subscript
- Paragraph formatting: Unordered list, Ordered list, and Paragraph alignment: left, center, right, justified, outdent, and indent
- Indentation options: Indent and Outdent

Embedding controls include the following options:

- Horizontal Rule
- Link
- Table
- Picture
- Video

6. You can also edit your visual in Code view.

   Rich text visuals support the Code View option, as a simple toggle between Text view and Code view.

   For example, you can override the width settings on a previously inserted image. Editing in Code view enables you to specify the width in terms of percentages:

   ```
   <img src="/arc/reports/staticasset/raw/5.png" style="width: 50%;">
   ```

   or in terms of pixels:

   ```
   <img src="/arc/reports/staticasset/raw/5.png" style="width: 350px;">
   ```

   You can also remove table borders or format custom email notification templates.
7. Click APPLY to save your changes and return to the dashboard.

   If you do not want to apply your changes, click CANCEL or just close the modal window.
8. Change the title to World Population - Rich Text.

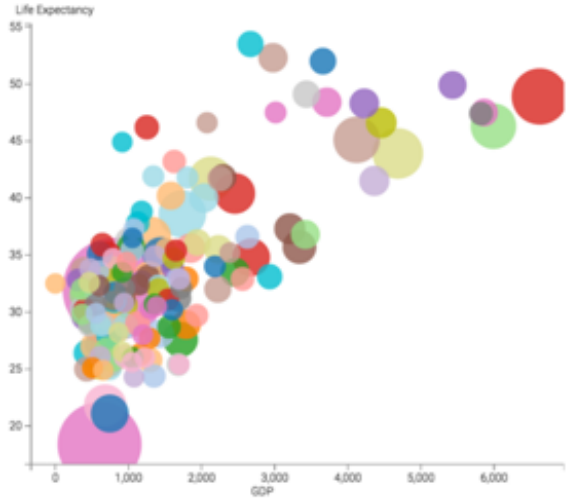
   [Optional] You can also add a brief description of the visual as a subtitle below the title of the visualization.
9. Click SAVE at the top left corner of the Dashboard Designer.

Results

In this example, the following changes have been implemented: two tables with images in the left column and text in the right column, 2 levels of headings, and links for sources. Your visual should look something like this:

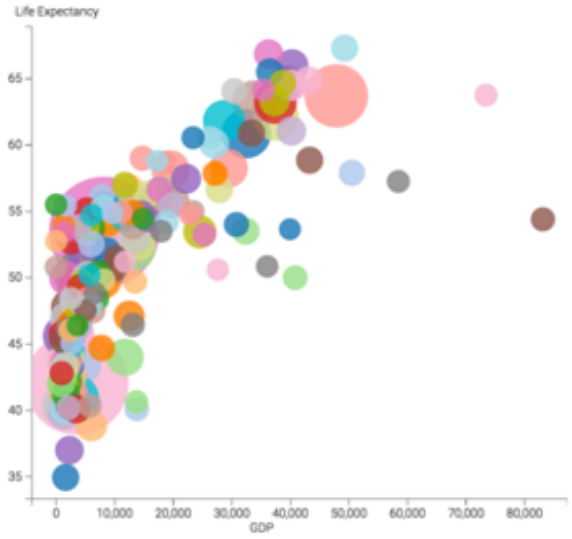
World Population - Rich Text

A Global Imbalance



In the early 1800's, life expectancy at birth began to increase significantly in the countries that industrialized early, but it remained low in the rest of the world. Industrialization levels correlated strongly with life expectancy, and lead to notable inequalities in global health distribution. Good health was prevalent in the rich countries, where improvements in sanitation, infant/maternal care, and food quantity lead to longer life. Meanwhile, poor health persisted in poorer countries.

Recent Improvements



This global inequality decreased over the last half of the 20th century. Countries that relatively recently suffered from bad health are catching up rapidly. Since 1900, the global average life expectancy more than doubled, increasing from 32.6 years for all reporting countries to 70.0 years.

Sources

1. Hans Rosling
2. Gapminder World
3. Life Expectancy; World Health Organization

Related Information

Setting system privileges

Creating a visual

Inserting URL links

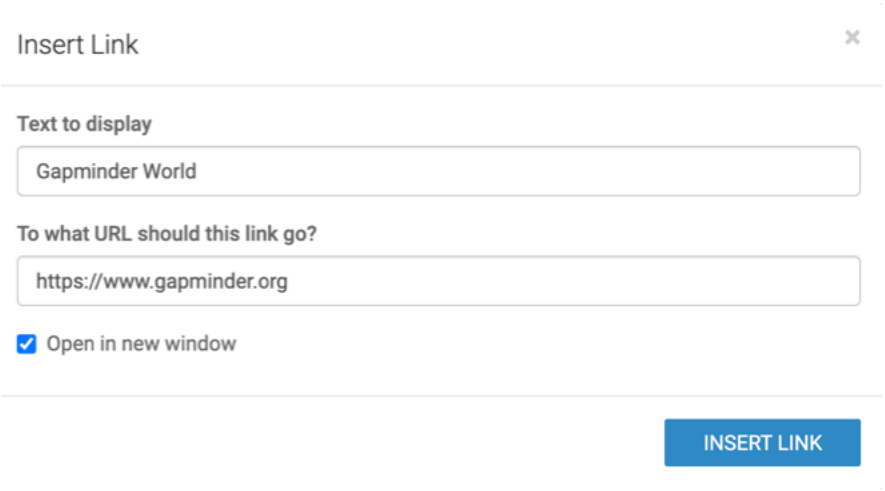
Cloudera Data Visualization enables you to insert URL links in Rich text visuals.

Procedure

1. Click the Link icon in the Rich Text Editor.

This opens the Insert Link modal window.

2. Make the following specifications:
   a) In the Text to display field, add the display text.
      In this example, 'Gapminder World' has been used as display text.
   b) In the To what URL should this link go? field, add the URL address.
      In this example, https://www.gapminder.org has been used.
   c) Select the Open in new window option.



3. Click INSERT LINK to add the link to your visual.

Inserting tables

Cloudera Data Visualization enables you to insert tables in Rich text visuals.

In the Rich-Text Editor, click the Table icon and select the number of rows/columns for the new table.

The table is automatically added to your visual.

Inserting images from local machine

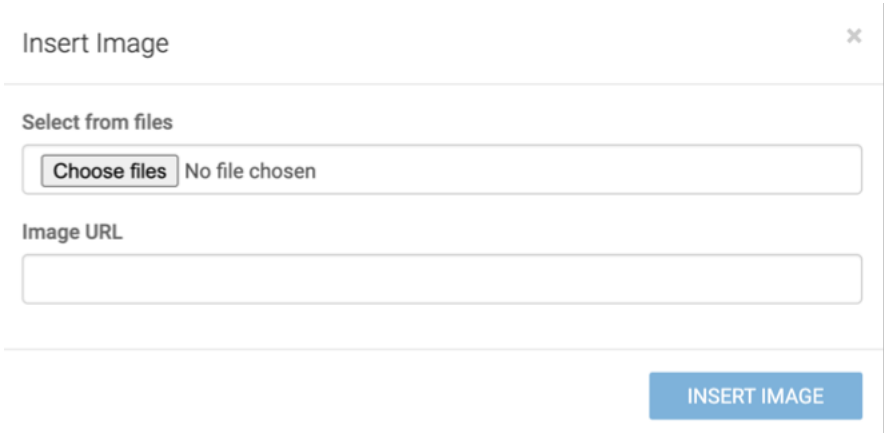
Cloudera Data Visualization enables you to insert an image into a Rich text visual from a local directory.

Procedure

1. Click the Picture icon in the Rich Text Editor.

This opens the Insert Image modal window.

2. Under Select from files, click Choose Files and select an image file from your file system.



3. Click INSERT IMAGE to add the picture to your visual.

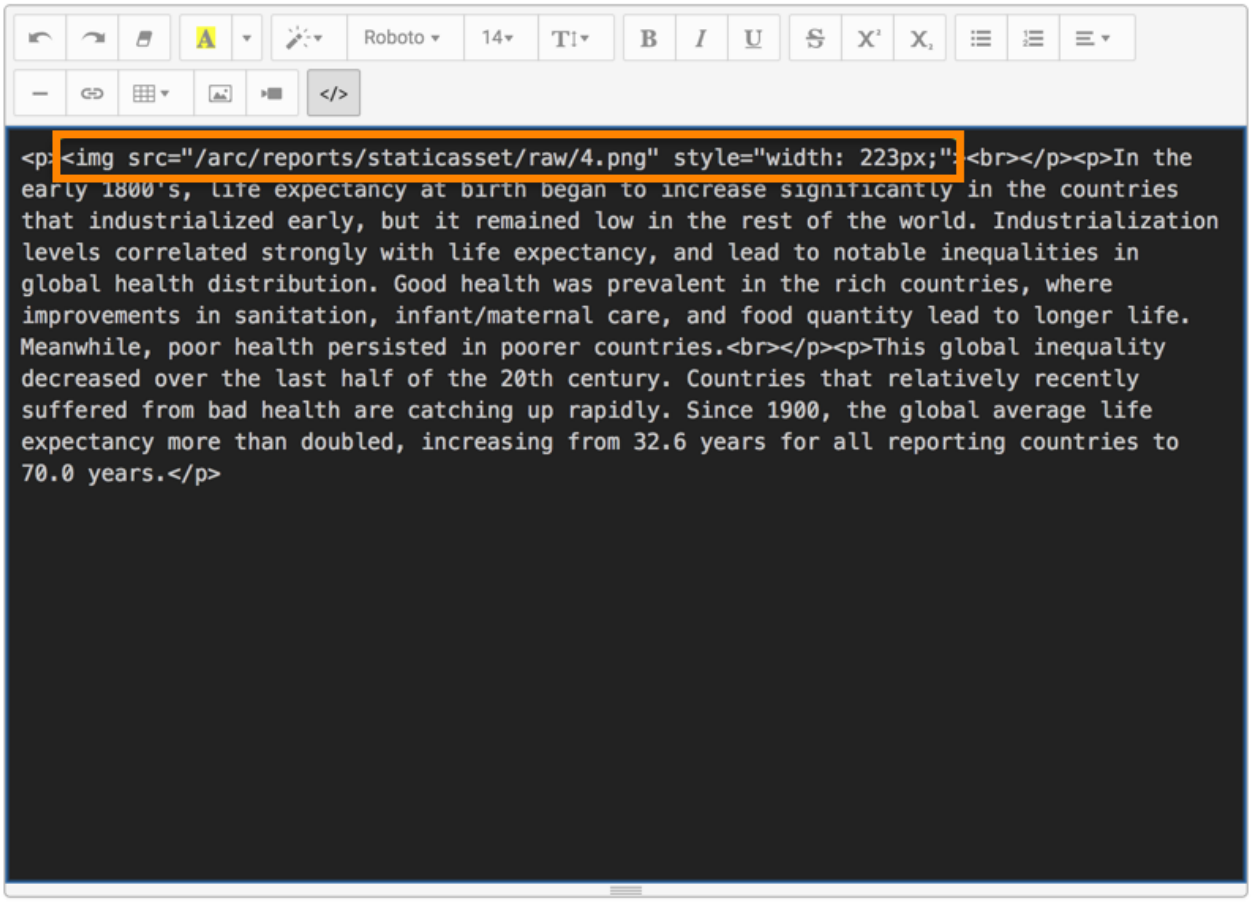
Note: Cloudera Data Visualization automatically converts the image into a static asset and adds it to the set of static assets it manages.

If you click the Code icon, it shows a link to the image inside the staticasset database.

Example

In this example, the link appears as asset_id.image_type, where asset_id is assigned automatically based on the order of upload, and image_type is consistent with the format of the uploaded file, such as png.

```
<img src="/arc/reports/staticasset/raw/4.png" style="width: 223px;">
```



Inserting images from static assets

Cloudera Data Visualization enables you to insert an image that has been previously uploaded to the system as a static asset.

Procedure

1. Click the Picture icon in the Rich-Text Editor.

   This opens the Insert Image modal window.
2. Provide the address of the image in the Image URL field.

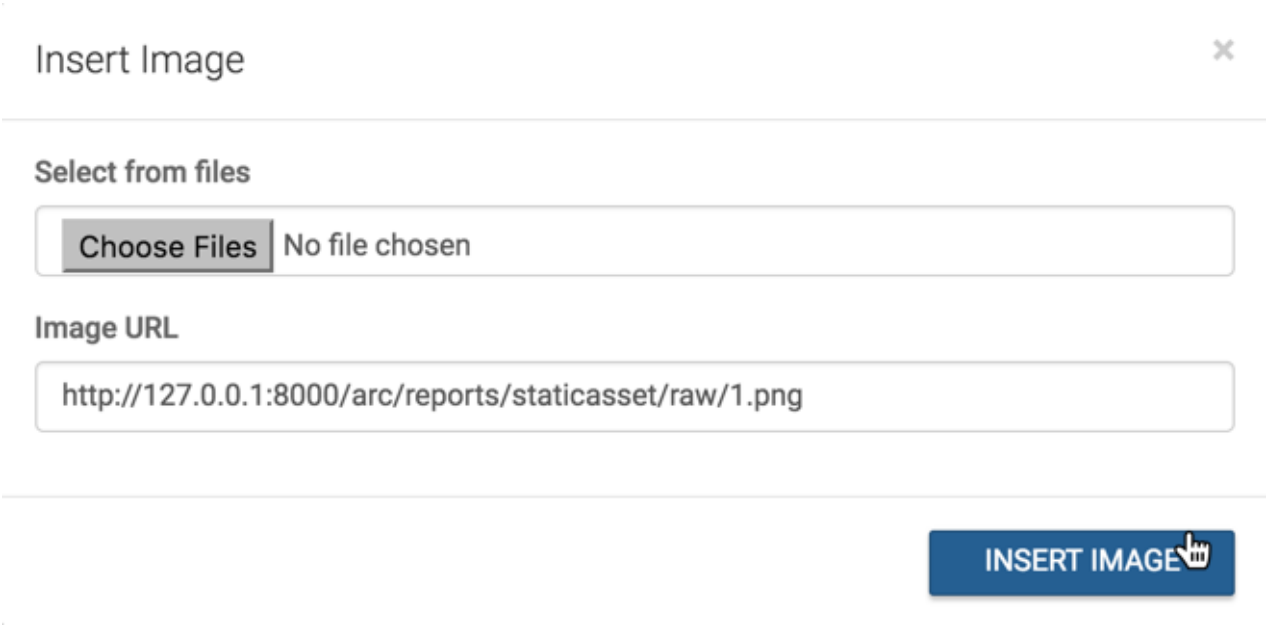
   The format for inserting existing static assets is the following:

   ```
   ip-address:port-address/arc/reports/staticassets/raw/asset-id.asset-type
   ```

   Where asset_id is the ID of an asset already imported into the system, and asset-type is the file extension.
3. Click INSERT IMAGE.

Example

In this example, asset 1.png is loaded.



Resizing images

Cloudera Data Visualization enables you to resize an image inside a rich text visual.

Procedure

1. Open the Rich text visual in Edit mode.
2. Click the image and select one of the magnification controls: 100%, 50%, or 25%.
   You also have the option to delete the image from your visual.

Inserting videos

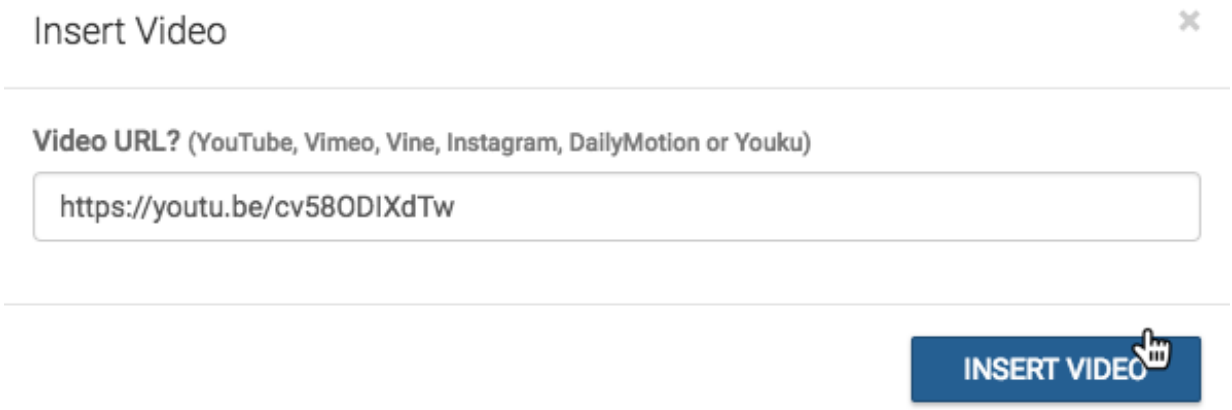
Cloudera Data Visualization enables you to insert videos in Rich text visuals.

Procedure

1. Click the Video icon in the Rich Text Editor.

   This opens the Insert Video modal window.
2. Enter the URL address of the video.

   In this example, the following video is used: https://youtu.be/cv58ODIXdRw.



3. Click INSERT VIDEO.

   Cloudera Data Visualization automatically converts the URL into an embedded object. When you examine the visual in Code view, the video appears in an <iframe> element:

```
<iframe frameborder="0" src="//www.youtube.com/embed/cv58ODIXdTw" width=
"640" height="360" class="note-video-clip"></iframe>
```

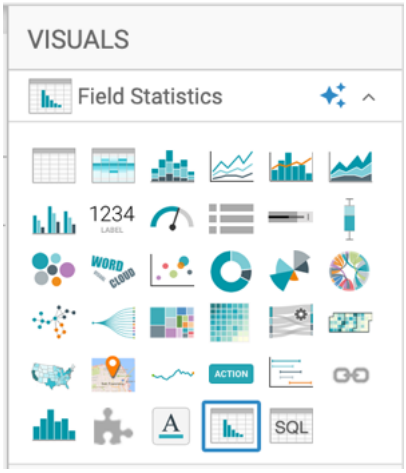
   In View mode, the visual contains the embedded video object, together with the controls specific to the hosting site.

Field statistics

Cloudera Data Visualization enables you to create Field statistics visuals. Field statistics visuals describe the dataset in statistical terms. They work across all data connections, and on datasets that contain table joins.

Procedure

1. Start a new visual based on the US County Population dataset.
   For instructions, see *Creating a visual*.
2. Select Field Statistics in the VISUALS menu.

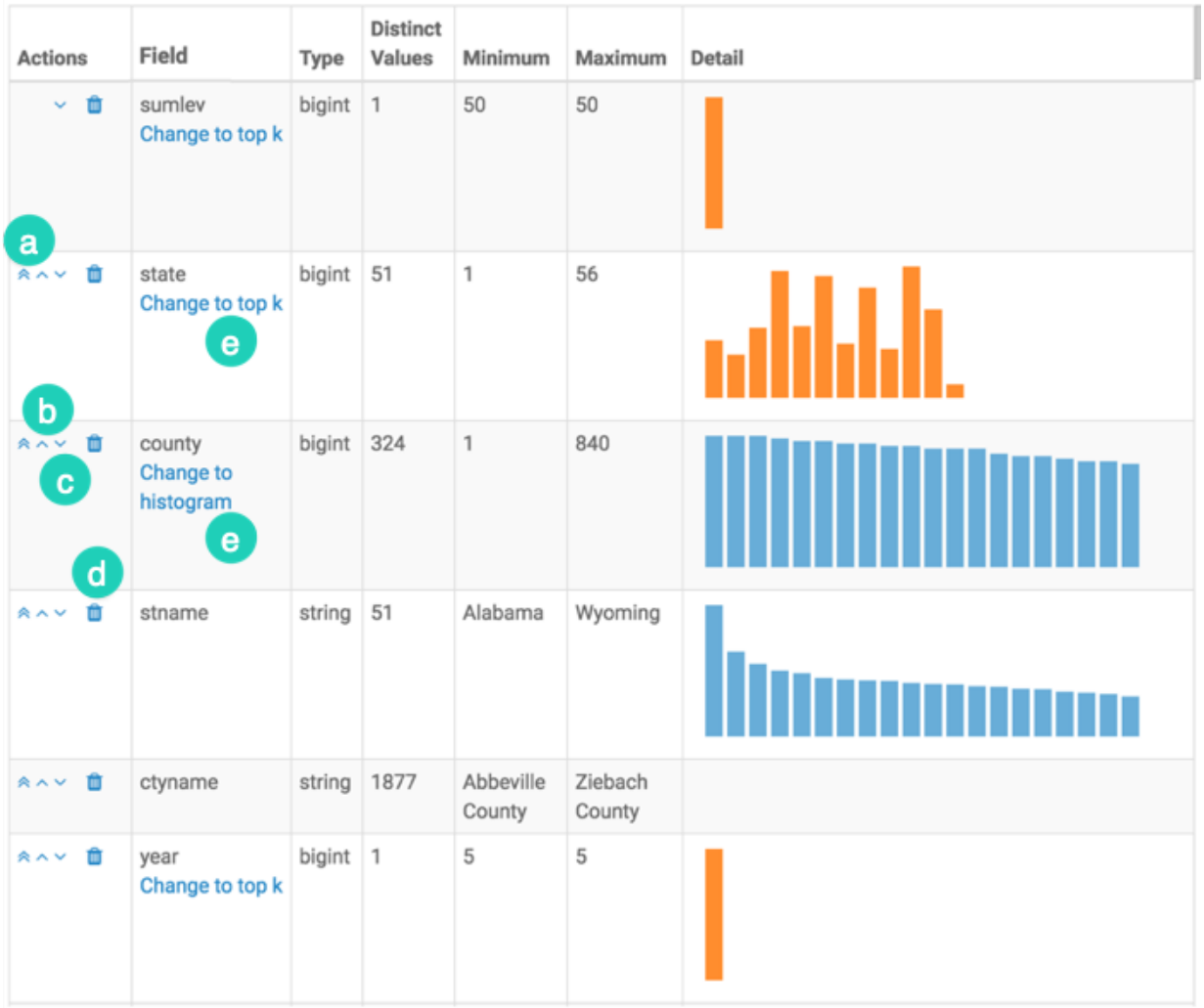


   The visual appears as a table that shows the following statistical information for each field in the dataset.

   - Actions column contains some widgets for making changes to the appearance of the table.
   - Field column indicates the name of the dataset field, and provides an option to include only a limited number of top values, when applicable.
   - Type column indicates the data type of the dataset field.
   - Distinct Values column reports the number of distinct values for this field. This is typically a good indication of the field's suitability as a dimension.
   - Minimum column reports the smallest (lowest) value in the dataset field. For strings, this would be the 'least' in alphabetical sort.

     In this example, for field stname (state name), this would be Alabama, and for field ctyname (county name), this would be Abbeville County.
   - Maximum column reports the largest (highest) value in the dataset field. For strings, this would be the 'greatest' in alphabetical sort.

     In this example, for field stname (state name), this would be Wyoming, and for field ctyname (county name), this would be Zieback County.
   - Detail column holds a graphic that is either a bar chart of top K values, or a distribution histogram of these values, for each column.
3. You can make changes to the visual through the following actions:
   a) In the Actions column, click the (double up) icon to move the field to the very top of the table.
   b) In the Actions column, click the (up) icon to move the field up one row in the table.
   c) In the Actions column, click the (down) icon to move the field down one row in the table.
   d) In the Actions column, click the (trash) icon to remove the field from the table.
   e) In the Field column, toggle between:
      - Change to top k to change the Detail column to show a bar representation of the top 20 values (default) for the dataset field
      - Change to histogram to change the Detail column to show a distribution histogram for the dataset field



4. Change the title to US County Populations - Field Statistics.
   [Optional] You can also add a brief description of the visual as a subtitle below the title of the visualization.
5. Click SAVE at the top left corner of the Dashboard Designer.

Results

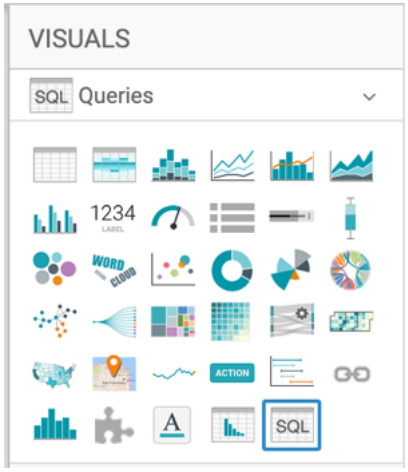
In View mode, the Field statistics visual does not have actions, or it does not contain the Actions column.

SQL queries

In Cloudera Data Visualization, SQL query visuals enable you to process SQL language queries using an enhanced visual editor for SQL authoring, including autocomplete.

Procedure

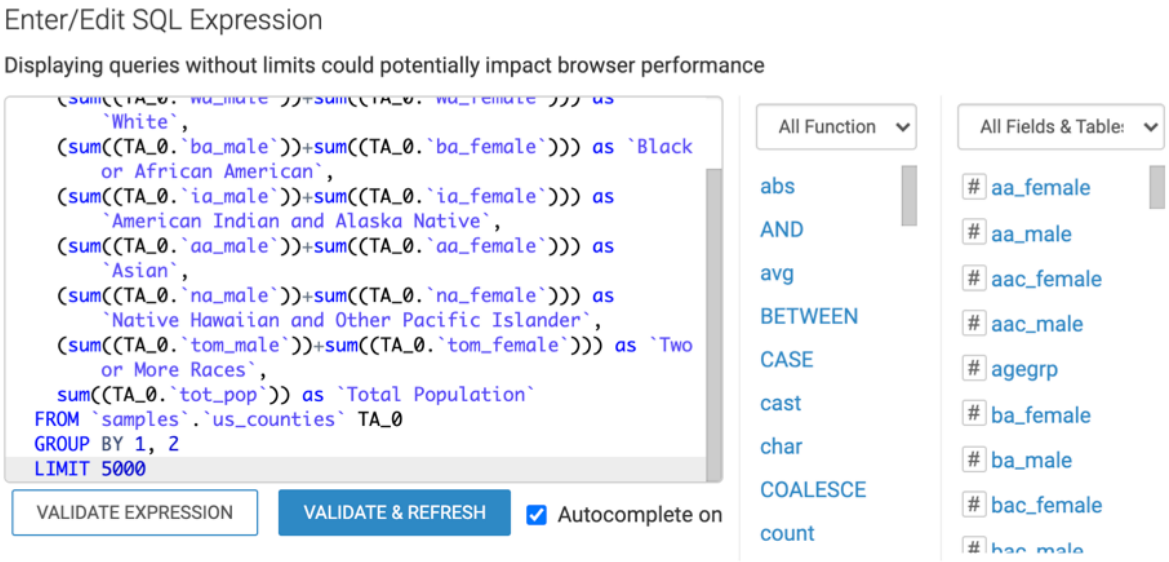
1. Start a new visual based on the US County Population dataset.
   For instructions, see *Creating a visual*.
2. Select SQL Queries in the VISUALS menu.



The shelves of the visual changed. You can find a link in the shelf area that leads you to the visual in full-screen mode, where you can view/edit the SQL statement in the SQL editor.

3. Enter the following text in the Enter/Edit SQL Expression modal window:

```
SELECT
    TA_0.`stname` as `State`,
    TA_0.`ctyname` as `County`,
    (sum((TA_0.`wa_male`))+sum((TA_0.`wa_female`))) as `White`,
    (sum((TA_0.`ba_male`))+sum((TA_0.`ba_female`))) as `Black or African
 American`,
    (sum((TA_0.`ia_male`))+sum((TA_0.`ia_female`))) as `American Indian an
d Alaska Native`,
    (sum((TA_0.`aa_male`))+sum((TA_0.`aa_female`))) as `Asian`,
    (sum((TA_0.`na_male`))+sum((TA_0.`na_female`))) as `Native Hawaiian
and Other Pacific Islander`,
    (sum((TA_0.`tom_male`))+sum((TA_0.`tom_female`))) as `Two or More Race
s`,
    sum((TA_0.`tot_pop`)) as `Total Population`
  FROM `samples`.`us_counties` TA_0
  GROUP BY 1, 2
  LIMIT 5000
```



4. Click VALIDATE & REFRESH.

The visual is updated in the preview area. You can make further edits to the SQL code and check the changes by validating and refreshing the expression again.

If you do not need the autocomplete function, uncheck the Autocomplete on option.

5. Click APPLY to save the changes.

   If you do not want to apply your changes, click CANCEL or just close the modal window.

6. Change the title to County Populations, by State /County - Query.

   [Optional] You can also add a brief description of the visual below the title of the visualization.

7. Click SAVE at the top left corner of the Dashboard Designer.

Results

The table created from the query should look similar to this:

County Populations, by State /County - Query

State	County	White	Black or African American	American Indian and Alaska Native	Asian	Native Hawaiian and Other Pacific Islander	Two or More Races	Total Population
Alabama	Autauga County	43,568	10,211	289	566	48	832	55,514
Alabama	Baldwin County	166,616	18,283	1,399	1,548	125	2,819	190,790
Alabama	Barbour County	13,735	12,884	181	122	50	229	27,201
Alabama	Bibb County	17,229	5,030	82	31	30	195	22,597
Alabama	Blount County	55,627	994	348	155	64	638	57,826
Alabama	Bullock County	2,848	7,304	98	22	89	113	10,474
Alabama	Butler County	11,053	8,843	71	195	8	137	20,307
Alabama	Calhoun County	89,177	24,597	574	950	111	1,887	117,296
Alabama	Chambers County	20,076	13,319	72	194	22	381	34,064
Alabama	Cherokee County	24,239	1,207	135	66	8	366	26,021
Alabama	Chilton County	38,422	4,451	219	171	83	473	43,819
Alabama	Choctaw County	7,625	5,887	23	33	0	65	13,633
Alabama	Clarke County	13,662	11,117	111	86	4	181	25,161
Alabama	Clay County	11,166	1,956	72	29	1	211	13,435
Alabama	Cleburne County	14,033	546	55	26	10	162	14,832
Alabama	Coffee County	39,371	8,947	736	709	208	1,281	51,252
Alabama	Colbert County	44,096	8,920	300	273	23	834	54,446
Alabama	Conecuh County	6,741	5,999	47	18	3	173	12,981
Alabama	Coosa County	7,358	3,378	66	19	17	128	10,966
Alabama	Covington County	32,128	4,828	239	196	1	563	37,955

< 1 2 3 ... 158 >

As an example of using this visual, sort on the column Native Hawaiian and Pacific Islander to see where most people of that ethnicity live. You can notice that there is a very large population in Los Angeles, San Diego, and in the Salt Lake area.

County Populations, by State /County - Query

State	County	White	Black or African American	American Indian and Alaska Native	Asian	Native Hawaiian and Other Pacific Islander	Two or More Races	Total Population
Hawaii	Honolulu County	219,187	26,894	3,271	422,319	93,362	211,339	976,372
California	Los Angeles County	7,132,409	923,555	149,382	1,440,604	37,507	279,332	9,962,789
Hawaii	Hawaii County	64,697	1,465	1,089	42,527	23,682	55,731	189,191
California	San Diego County	2,436,930	177,261	42,434	369,491	18,898	132,049	3,177,063
Utah	Salt Lake County	947,370	20,445	13,746	38,414	17,312	26,555	1,063,842
California	Sacramento County	947,022	157,728	22,687	221,624	16,804	84,256	1,450,121
Hawaii	Maui County	56,817	1,182	815	46,075	16,688	36,649	158,226
Washington	King County	1,431,673	131,316	20,944	311,876	16,371	95,260	2,007,440
Nevada	Clark County	1,464,211	227,061	23,540	188,017	16,260	81,670	2,000,759
California	Alameda County	813,510	198,719	18,177	429,695	15,030	79,589	1,554,720
California	Orange County	2,300,708	63,299	33,746	582,911	11,916	97,552	3,090,132
California	San Mateo County	471,664	22,949	6,547	194,917	11,796	31,438	739,311
Washington	Pierce County	623,549	58,458	13,326	51,358	11,488	53,502	811,681
Arizona	Maricopa County	3,351,239	218,894	104,816	154,297	10,681	102,242	3,942,169
California	Riverside County	1,833,979	158,919	42,432	149,237	9,517	74,699	2,268,783
California	San Bernardino County	1,615,386	200,072	42,224	145,245	9,476	68,910	2,081,313
California	Santa Clara County	1,058,855	53,405	25,396	618,537	8,927	72,384	1,837,504
Alaska	Anchorage Municipality	199,967	18,441	24,300	25,918	6,670	23,314	298,610
Hawaii	Kauai County	22,818	364	332	21,671	6,336	16,913	68,434
California	Contra Costa County	736,866	104,138	10,680	168,815	6,163	52,935	1,079,597

< 1 2 3 ... 158 >