

Apache Zeppelin 3

Apache Zeppelin Overview

Date of Publish: 2018-04-01

<http://docs.hortonworks.com>

Contents

Overview.....	3
----------------------	----------

Overview

Apache Zeppelin is a web-based notebook that supports interactive data exploration, visualization, and collaboration.

Zeppelin supports a growing list of programming languages and interfaces, including Python, Scala, Hive, SparkSQL, shell, AngularJS, and markdown.

Apache Zeppelin is useful for working interactively with long workflows: developing, organizing, and running analytic code and visualizing results.

The screenshot displays the Apache Zeppelin Notebook interface for an IoT Data Analysis workflow. At the top, there is a search bar and a user profile indicator. Below the title, a table lists various event types and their associated metrics. The table data is as follows:

Event Type	Count	Unit	Value 1	Value 2	Value 3	Value 4	Value 5	Value 6
Normal	N	miles	90	4,300	-89.52	40.7	0	0
Normal	N	miles	90	4,300	-91.05	41.72	0	1
Normal	N	miles	90	4,300	-91.47	41.74	0	0
Lane Departure	N	miles	90	4,300	-91.59	41.7	1	0
Unsafe following distance	N	miles	90	4,300	-88.77	40.76	0	1

Below the table, three KSQL queries are executed and visualized:

- Query 1:** A bar chart showing the count of occurrences for 'Overspeed' and 'Unsafe following distance' events. The 'Unsafe following distance' event has a significantly higher count (around 1,100) compared to 'Overspeed' (around 100).
- Query 2:** A pie chart showing the distribution of violations by certification status. The 'N' (Not Certified) category represents approximately 70% of the violations, while the 'Y' (Certified) category represents approximately 30%.
- Query 3:** A line chart showing the number of hours driven for different event types, grouped by certification status. The chart shows a significant spike in hours driven for the 'Unsafe following distance' event type, particularly for the 'N' (Not Certified) group.