

Data Visualization Assignment

A survey was conducted to gauge an audience interest in different data science topics, namely:

Big Data (Spark / Hadoop) Data Analysis / Statistics Data Journalism Data Visualization Deep Learning Machine Learning The participants had three options for each topic: Very Interested, Somewhat interested, and Not interested. 2,233 respondents completed the survey.

- 1. The survey results have been saved in a csv file and can be accessed through this link: https://cocl.us/datascience_survey_data.*
- 2. View top 5 rows of the data.*
- 3. View the dimensions of the data.*
- 4. View additional information of the data.*

Use the artist layer of Matplotlib to replicate the bar chart below to visualize the percentage of the respondents' interest in the different data science topics surveyed.

To create this bar chart, you can follow the following steps:

- 1. Using Style (ggplot)*
- 2. Sort the data frame in descending order of Very interested.*
- 3. Convert the numbers into percentages of the total number of respondents. Recall that 2,233 respondents completed the survey. Round percentages to 2 decimal places.*
- 4. As for the chart:*
 - use a figure size of (20, 8).*
 - bar width of 0.8.*
 - use color #5cb85c for the Very interested bars, color #5bc0de for the Somewhat interested bars and color #d9534f for the Not interested bars.*
 - use font size 14 for the bar labels, percentages, and legend.*
 - use font size 16 for the title.*
 - display the percentages above the bars as shown above, and remove the left, top, and right borders. Once you are satisfied with your chart, please upload a screenshot of your plot.*