Introduction to D3

Practice Exercise

Objective

The purpose of this exercise is to learn visualization through D3.js. You are free to choose your own visualizations that you think would be the best suited for the data and tasks listed in sections below.

Data

There is a data file provided with this exercise **<results.csv>**.

About the dataset

● A team has performed several experiments to test the efficiency of video streaming methods. For better validation of results, tests are conducted using **4** *network conditions/profiles* with **7** *video samples* under **3** *buffer capacity configurations.*

● In the first configuration, the total duration of the video that can be stored in the buffer is either 30 seconds or 60 seconds based on the duration of the video. If the video is more than 10 minutes, the buffer capacity is 60 seconds, else, it is 30 seconds.

● In the other two configurations, the buffer capacity is fixed and not dependent on video duration, i.e., 120 seconds and 240 seconds.

● Video sample V1, V4, and V5 are less than 10 minutes and remaining (V2, V3, V6, and V7) are more than 10 minutes. The results are stored in the “result.csv” file.

● The **result.csv** file has the following attributes: **<profile, sample, method, quality, change, inefficiency, stall, numStall, avgStall, overflow, numOverflow, qoe, bufSize>**. The attributes are defined as follows:

- **profile:** Network profile used for testing.
- **sample:** Video sample on which test is conducted.
- **method:** Method used for streaming.

- **quality:** Average quality played (in Kbps).
- **change:** Changes in quality during the playback.
- **inefficiency:** Inefficiency of the method to fully utilize the available bandwidth.
- **stall:** Total stall duration (in seconds) during the playback.
- **numStall:** Number of stalls happened during the playback.
- **avgStall:** Average stall duration (in seconds) during the playback.
- **overflow:** The duration (in seconds) for which buffer was full.
- **numOverflow:** Number of times when buffer was full.
- **qoe:** Quality of experience during the playback.
- **bufSize:** Buffer configuration, i.e., the maximum content that can be buffered/ buffer capacity (in seconds). Do note that there are three configurations: 30 seconds and 60 seconds come under the same buffer configuration category. The other two configurations are 120 seconds and 240 seconds.

You can process the data to derive new values.

Task

Based on the description given above, plot the visualization for the following queries. Feel free to use your choice of the plot (bar plot, histogram, stacked plot, X-Y plot, etc.) to clearly visualize the results:

1. Draw plots to show the Average quality and number of changes in quality for a method. Each buffer configuration should have a separate plot.

2. Draw a single plot to show the Average QoE for a method grouped by different buffer configurations. The plot should be in such a way that we can compare the performance of the method with itself for different buffer configuration as well as with different methods within a buffer configuration.

3. Draw plots to show the correlation between inefficiency and quality for all methods in different buffer configurations.

4. What are the methods which have the minimum number of stalls for video V7 under different network profiles? Draw an appropriate plot for it.

In the next section, you can find a few sample visualizations for reference.

Sample Visualizations

- Pie chart - https://bl.ocks.org/mbostock/3887235
- Heat map - http://bl.ocks.org/tjdecke/5558084
- Grouped bar chart - https://bl.ocks.org/mbostock/3887051
- Line graph - https://bl.ocks.org/mbostock/3883245
- Bubble Chart - https://bl.ocks.org/mbostock/4063269
- Stacked Area Chart - https://bl.ocks.org/mbostock/3885211