CS590V: Data Visualization and Exploration – Spring 2017
Homework 2&3: Visualizations and Interactions
Submitted by: Ankita Mehta (30450796)

Datasets:

Four datasets have been considered here: Facebook, White Wine, Red Wine and Both(White-Red) wine. The main page gives you an option to select a Dataset and on every page there is a 'prev' button to go back to the main page.

Pre-processing:

1. Missing values in Facebook dataset have been filled with zero using python script.
2. Red Wine and White wine datasets are combined manually.

Facebook Dataset Visualizations and Interactions:

1. Various charts have been explored such as Scatterplot, Line chart, Histogram, Pie Chart, Heat map and Table.
2. Every chart has selection and probing functionality.
3. Every chart has common drop down menu for changing Y axis and the color attribute.
4. Every chart has reset button attached to it.

Scatterplot:

Selection: In the scatterplot, you can select the data by hovering over the legends.

Probing: When you take the cursor over any data point, a box will appear showing its related information.

X and Y axis: Its X axis can be changed from the first drop down menu and Y axis from the third menu.

Color attribute: You can change its color attribute as well from the 4th drop down menu.

Line chart:

Selection: In the line chart, you can select the particular trending line by hovering over the legends.

Probing: When you take the cursor over the line, a box will appear showing its related information.

X and Y axis: Its X axis can be changed from the second drop down menu and Y axis from the third menu.

Color attribute: You can change its color attribute as well from the 4th drop down menu.

Histogram:

Selection: In the Histogram, you can click the particular bar and the related data will be shown in other charts as well.

Probing: When you take the cursor over the bar, a box will appear showing its related information.

X and Y axis: Its X axis can be changed from the second drop down menu and Y axis from the third menu.

Pie chart:

Selection: In the pie chart, you can click the particular track and the related data will be shown in other charts as well.

Probing: When you take the cursor over the pie, a box will appear showing its related information.

Color attribute: You can change its color attribute from the 4th drop down menu.

Heat map:

Selection: In the line chart, you can click the particular heat box and the related data will be shown in other charts as well.

Probing: When you take the cursor over any heat map, a box will appear showing its related information.

X and Y axis: Its X axis can be changed from the second drop down menu and Y axis from the third menu. X – axis will be shown on the horizontal part of heat map. Grouping of information has been according to the Y – axis.

Color attribute: You can change its color attribute as well from the 4th drop down menu.
It will be shown as the vertical part of heat map.

Table:

Performance table has been created for Facebook dataset for all 500 rows. This table can be accessed using 'Another Visualisation' button on Facebook Visualisations webpage.

You can use 'next' and 'last' button to scroll between pages.

There is 'prev' button which will take you to the previous Facebook page.

White/Red Wine Dataset Visualizations and Interactions:

1. **Scatterplot:** The default scatterplot is showing the positive correlation between chloride and fixed acidity chemicals of white wine for various quality levels. They have strong correlation for 5,6 and 7 quality levels.
   Similar correlations can be seen in between various chemicals by changing the X and Y attributes.

2. **Line chart:** The default line chart is showing the variation of fixed acidity over the increase in chloride content for various quality levels. Fixed acidity is higher for quality level 6 and when chloride concentration is 0.5. Similar relationships can be seen for various chemicals by changing X and Y attributes.

3. **Histogram:** The default histogram is showing the chemical distribution for white wine. It shows somewhat normal distribution for chloride. Similar distributions can be seen for other chemicals by changing X axis attribute.

4. **Pie Chart:** Pie chart is showing the percentage distribution of various quality levels. It can be inferred that quality level 6 has maximum number of samples which is 45%.

Both Wines Dataset Visualizations and Interactions:

1. ***Bar charts:*** It is showing the counts of various quality levels for both Red and White Wine. White wine has quality samples of level 9 as well whereas Red wine doesn't have. Red wine has maximum number of level 5 samples whereas White Wine has of 6 level.

2. **Scatterplot:** It is showing the total pH content for various alcohol content for both Red and White wine. It can be inferred that white wine has large pH content when alcohol present is 9.4

3. **Pie Chart:** It is showing total percentage of samples for various quality levels for both Red and White wine.