FAT LAB EXAM WINTER 2020-21

Slot: L13+L14-Wednesday

Date: 2-6-2021

1. Consider the below data set and plotly/ggplot2 package.

Course	% of Students Who Scored Above 90%	% of Students Who Scored Less than 50%	Class Average
A	6.25	23.44	63.92
B	8.20	24.59	58.46
C	1.56	28.13	53.38
D	3.13	35.94	53.91
E	0.00	52.38	43.14
F	3.33	41.67	54.60
G	3.45	8.62	66.52

- Create scatter plot, histogram and box plot by taking the necessary attribute.
- Generate single image for above three charts.
- Map the color of the points to courses, make the points bigger by setting size to 7, add an appropriate title and label for x/y-axis to the plot
- Highlight the courses – class average between 55% to 60% in the scatter plot
- Create scatterplot matrices for dataset.

2. Consider the airquality data set in the plotly package.
- Create a scatter plot comparing the Temp and Ozone variables. Apply hover for highest temp,
- Create a histogram of the Temp variable. Can you adjust the binning so that there are (approximately) 25 bins?
- Plot the frequency of observations in each Month.
- Plot a Scatterplot Matrix(SPLOM) for the dataset.
- Create a boxplot to view the distribution of Ozone for each month.

3. Titanic dataset
   - Find percentage of survived female who age is above 30.
   - Count the number of passengers in each class.
   - What was the name of the youngest surviving female?
   - In what class was the oldest surviving male?
   - Plot stacked bar plot for appropriate attributes.
   - Plot box plot for appropriate attributes.

4. Goods.csv dataset
   - Generate word cloud by taking segment description attribute
   - Apply black-white and color theme with different shapes- Star, pentagon, and Rectangle.
   - Generate pie chart by taking the answer, segment description and segment type attribute.
   - Generate single image for above three charts.

5. Creating a Network (reg-flow and reg-plot)
   - Labelling the nodes and Colouring Vertices
   - Remove duplicates/loops
   - Find the number of regions
   - Highlight (colour) the path from "western asia" in the network
   - Show shorts path for any 4 regions