CSCE 5300 Introduction to Big data and Data Science
ICE-9

Lesson Title: Data Visualization

Lesson Description: Various Plots

You can use google colab (https://colab.research.google.com) or Jupiter notebook.

Use Matplotlib library in python to visualize data.

1. Create line plot with the provided information in the source code attached on the Canvas.

  A. Create line plot for Sin(x) and Cos(x)
  B. Fit Sin(x) and Cos(x) in the same plot. Label Them as ICE_DV
  C. Sin(x) should be in green, and cos(x) should be in blue.
  D. Discuss the strengths and weaknesses of this plot.
  E. Identifying scenarios in this plot is most useful. Also state two real world applications of this plot.

2. Create scatter plot.

  A. Create the scatter plot of generated data in the source code.
  B. Plot a line of best fit. It refers to a line that defines the optimal relationship of the x-axis and y-axis coordinates of the data points plotted as a scatter plot on the graph.
  C. Insert title as Scatter Plot.
  D. Label x axis as X values and y axis as Y values.
  E. Discuss the strengths and weaknesses of this plot.
  F. Identifying scenarios in this plot is most useful. Also state two real world applications of this plot.

3. Using the data from question 1.

  A. Write code to plot a bar chart.
  B. Explain the trend that is depicted in the bar chart.
  C. Discuss the strengths and weaknesses of this plot
  D. Identify scenarios this plot is most useful. Also state two real world applications of this plot.

4. Using the data from question 1

  A. Write code to plot a histogram.
  B. Explain the trend depicted in the histogram.
  C. Discuss the strengths and weaknesses of this plot.
  D. Identify scenarios this plot is most useful. Also state two real world applications of this plot.

ICE Submission Guidelines

1. ICE Submission is individual.

2. ICE code must be properly commented.

3. The documentation should include screenshots of your code/queries and results.

4. Provide the explanation of the exercise for each question as per your understanding.

5. The similarity score for your document should be less than 15%.

6. Submit the source code (if any) properly commented and documentation (.pdf/.doc) with explanation and screenshot of source code/queries having input logic and output results.

7. Submission after the deadline is considered as late submission.

References:

https://matplotlib.org/stable/gallery/index.html