CS 6635/5635 Spring Semester 2021

Date: Monday, February 1, 2021

Due Date: Wednesday, February 10, 2021 by Midnight MST

Assignment 1

Plotting and Graphing using Python and/or Matlab

Use either Matplotlib in Python[1] or Matlab to perform the following plotting:

Part 1: Generate your own data and visualize it [20 pts]

1. [4 pts] Create an array with 200 elements from 1 to 200 in order. Create a **box plot** for visualization of your data.
2. [4 pts] Create an array with 10,000 random numbers. Create a **histogram** of the data using 20 bins.
3. [6 pts] Write a program to generate 100 random numbers Gaussian distributed between 1 and 100. Write the numbers out to a binary file and use a **line graph** to draw the 100 numbers.
    3.a You will need to find an appropriate mean and standard deviation for the Gaussian. It is okay if just a few of the numbers generated are outside the [1,100] range.
4. [6 pts] Write a program to read the binary file back, divide the range between 1 and 100 into 7 intervals, and calculate the frequency for each interval: Display **a histogram** of your result.

Part 2: Interesting data sets for visualization [36 pts]

1. [6 pts] Download the NOAA Land Ocean Temperature Anomalies Data Set: https://my.eng.utah.edu/~cs6635/NOAA-Temperatures.csv. Create a bar plot of the data. Include a label called "Year" along the x-axis and a label called Degrees F +/- From Average along the y-axis. Describe trends in the data.
   [6 pts] Download the member of Congress by Age data set: https://git.io/Jt45w
2. Create a **Star Plot** of the data and create a **Parallel Coordinates** Plot of the data. Describe the trends in the data.
3. [6 pts] Download the U.S. Birth data set: https://git.io/Jt45X. What day of the month had the highest number of births? What day of the month had the lowest number of births? Are there any interesting trends in the data, i.e. more births in Summer or Winter? What about births on Friday the 13th?
4. [18 pts] Five Thirty Eight maintains a sever with many interesting datasets: https://github.com/fivethirtyeight/data . Choose 3 different data sets to visualize. Visualize each data set in a different way. Describe the trends in the data.

---

1 Other Python visualization libraries, such as Seaborn, Pandas, and Plotly are also acceptable, but note that, for the purposes of plotting in Python, the TAs are most familiar with Matplotlib.

Part 3: Questions on The Value of Visualization Paper (https://my.eng.utah.edu/~cs6635/Value-of-Visualization.pdf) [24 pts]

5 [6 pts] Why is assessing value of visualizations important? What are the two measures for deciding the value of visualizations?
6 [6 pts] Briefly describe a mathematical model for the visualization block shown in Fig. 1.
7 [6 pts] State four parameters that describe the costs associated with any visualization technique.
8 [6 pts] What are the pros and cons of interactivity of visualizations?

Part 4: 3D scalar volume data sets (Only for CS 6635) [20 pts]

MATLAB/Python also can be used for analysis and visualization of 3D volume data sets, such as brain MRI images. Download the brain MRI data set from https://pubweb.eng.utah.edu/~cs6635/T2.nii.gz . The data format is .nii with 320 x 320 x 256 di mensions. Load data in MATLAB/Python. Extract one slice from the volume and save it as an i mage.

- One library that can be used for this in Python is NiBabel: https://nipy.org/nibabel/

Extra Credit: Implementation of Part 1 [5 pts], Part 2 [10 pts] , and Part 4 [10 pts] (Only for CS 6635) in BOTH Python and Matlab.

What to turn in:

Write a *short* report documenting your results, including any necessary plots/figures, and answering any questions asked above. Be sure to explain any figures you submit and to write a conclusion at the end of your report. Your homework is primarily graded upon your report. Please submit your report on Canvas in PDF format.

If you wish, you may create a Jupyter Notebook[2] with prose documenting/explaining your results, export that to PDF, and turn the PDF in as your report. This is an option for those who like using Jupyter, but it is not required.

2 See https://www.datacamp.com/community/tutorials/tutorial-jupyter-notebook for a tutorial