DATA VISUALIZATION TECHNIQUES

21CS3051A

STUDENT ID: ACADEMIC YEAR: 2023-24
STUDENT NAME:

Table of Contents

A.Y. 2023-24 LAB CONTINUOUS EVALUATION

S.No	Date	Experiment Name	Pre-Lab (10M)	In-Lab (25M)			Post-Lab (10M)	Viva Voce (5M)	Total (50M)	Faculty Signature
				Program/ Procedure (5M)	Data and Results (10M)	Analysis & Inference (10M)				
1.		Introductory Session				-NA-				
2.		Visualization of Spreadsheet Models								
3.		Visualization of Semi-Structured Data								
4.		Interactive Plots using Python								
5.		Data visualization Using Tableau								
6.		Aggregation Methods in Tableau								
7.		Visual Encodings and Basic Dashboards in Tableau								
8.		Hierarchical and Topographical Data Visualizations in Tableau								
9.		Calendar Heatmaps and Flow Data Visualizations in Python								
10.		Time Series Data Visualization in Python								
11.		Dashboards, Actions and Story Telling in Tableau								
12		Oracle Database Connectivity using Python								
13.		Tabpy -Integrate Tableau with Python								

S.No	Date	Experiment Name	Pre-Lab (10M)	In-Lab (25M)			Post-Lab (10M)	Viva Voce (5M)	Total (50M)	Faculty Signature
				Program/ Procedure (5M)	Data and Results (10M)	Analysis & Inference (10M)				
14.		Geospatial Data Visualization in Tableau (Adv/Peer)								
15.		Interactive Web-based Data Visualizations using JavaScript								
16.		Network Analysis and Visualizations (Adv/Peer)								
17.		Text Data Visualization and Sentiment Analysis#								
18.		Interactive Dashboard Development in Tableau								
19.		Data Storytelling Visualization best practices								
20.		Big Data Visualization and Scalability								
21		Ethical Considerations in Data Visualization								

Introductory Session

Organization of the Student Lab Workbook

The laboratory framework includes a creative element but shifts the time-intensive aspects outside of the Two-Hour closed laboratory period. Within this structure, each laboratory includes three parts: Prelab, In-lab, and Post-lab.

a. Pre-Lab

The Prelab exercise is a homework assignment that links the lecture with the laboratory period - typically takes 2 hours to complete. The goal is to synthesize the information they learn in lectures with material from their textbook to produce a working piece of software. Prelab Students attending a two-hour closed laboratory are expected to make a good-faith effort to complete the Prelab exercise before coming to the lab. Their work need not be perfect, but their effort must be real (roughly 80 percent correct).

b. In-Lab

The In-lab section takes place during the actual laboratory period. The First Hour of the laboratory period can be used to resolve any problems the students might have experienced in completing the Prelab exercises. The intent is to give constructive feedback so that students leave the lab with working Prelab software a significant accomplishment on their part. During the second hour, students complete the In-lab exercise to reinforce the concepts learned in the Prelab. Students leave the lab having received feedback on their Prelab and In-lab work.

c. Post-Lab

The last phase of each laboratory is a homework assignment that is done following the laboratory period. In the Post-lab, students analyze the efficiency or utility of a given system call. Each Post-lab exercise should take roughly 120 minutes to complete.

Software Prerequisites

1. Python Installation Link: https://www.python.org/downloads/

2. Anaconda Installation Link: https://www.anaconda.com/products/individual

3. Jupyter Notebook Online

   https://hub.gke2.mybinder.org/user/jupyterlab-jupyterlab-demo-5inargwu/lab/tree/demo

4. Install Tableau Desktop – version 2021.1
   https://www.tableau.com/support/releases/desktop/2021.1

Experiment Title: Visualization of Spreadsheet Models

Aim/Objective:

Explore and develop effective methods for visually representing spreadsheet models.

Pre-Requisites:

- Basic idea on Data Visualization.
- Basic Idea on Implementation of Pandas Library

Pre-Lab:

1. Why is data visualization necessary in data science?

2. Mention few data visualization tools in Python.

3. How can you visualize more than three dimensions in a single chart?

4. Write the syntax and describe the parameters used for the following:
   - Facet Plot
   - Violin Plot
   - Area Chart

5. Write the syntax and describe the parameters used for the following:
   - Box Plot
   - Scatter Plot
   - Histograms
   - Pie Chart

In-Lab:

1. For the given dataset that contains immigration details to Canada from 1980 to 2013,
   i. Create an area plot for top 6 immigrant countries from 1990 to 2013
   ii. Create and year-wise immigrant bar chart from India to Canada during the period of 1980 to 2013.
   iii. Create a boxplot for Indian, Phillipins and China immigrants.
   iv. Show the total no. of immigrants from India and France countries using Area Chart and Pie chart.
   v. Create a scatter Histogram for the immigrants from Fiji and Singapore in the year 2013.

2. For the given data set that contains the data of flights that were on time in January for the years 2019 and 2020. Using the two data sets visualize the data using matplot and plotly libraries to depict the following:

   - Show the difference in statistics for distance for both the years using the appropriate plotting technique.
   - Visualize the no. of flights whose destination airport id is 11778 and 11267 using a bar plot or bar chart.
   - Create a Sunburst Plot for both the years depicting the difference among them.

- **Procedure/Program:**

Experiment #	<TO BE FILLED BY STUDENT>	Student ID	<TO BE FILLED BY STUDENT>
Date	<TO BE FILLED BY STUDENT>	Student Name	<TO BE FILLED BY STUDENT>

Experiment #	<TO BE FILLED BY STUDENT>	Student ID	<TO BE FILLED BY STUDENT>
Date	<TO BE FILLED BY STUDENT>	Student Name	<TO BE FILLED BY STUDENT>

- **Data and Results:**

Experiment #	<TO BE FILLED BY STUDENT>	Student ID	<TO BE FILLED BY STUDENT>
Date	<TO BE FILLED BY STUDENT>	Student Name	<TO BE FILLED BY STUDENT>

- **Analysis and Inferences:**

Sample VIVA-VOCE Questions (In-Lab):

1. Give a description about the syntax along with parameters and functionalities for the following visualization techniques:

    a. Network

    b. Animation

    c. Cheat Sheets

    d. Data Art

    e. Colour

    f. 3D

    g. Word Cloud

    h. Density

Post-Lab:

1. Visualize the given Placement Data Full Class dataset that contains details about Campus Recruitment using the below techniques for appropriate dimensions and differentiate between the two techniques:

   - Histogram and Bar Chart [For histogram let no. of bins = 10]
   - Facet Plot and Pair Plot
   - Area Chart and Pie Chart [For yes or no data]

- **Procedure/Program:**

Experiment #	<TO BE FILLED BY STUDENT>	Student ID	<TO BE FILLED BY STUDENT>
Date	<TO BE FILLED BY STUDENT>	Student Name	<TO BE FILLED BY STUDENT>

Experiment #	<TO BE FILLED BY STUDENT>	Student ID	<TO BE FILLED BY STUDENT>

Experiment #	<TO BE FILLED BY STUDENT>	Student ID	<TO BE FILLED BY STUDENT>
Date	<TO BE FILLED BY STUDENT>	Student Name	<TO BE FILLED BY STUDENT>

- **Data and Results:**

Experiment #	<TO BE FILLED BY STUDENT>	Student ID	<TO BE FILLED BY STUDENT>
Date	<TO BE FILLED BY STUDENT>	Student Name	<TO BE FILLED BY STUDENT>

- **Analysis and Inferences:**

(For Evaluator's use only)

Evaluator Remark (if Any):	
	Marks Secured: _____out of 50
	Signature of the Evaluator with Date

Evaluator MUST ask Viva-voce prior to signing and posting marks for each experiment.