Homework Assignment: Visualizing Real Data Sets using Matplotlib

Objective

Utilize your Matplotlib programming skills to work with real-world data sets and create meaningful visualizations. This assignment is designed to reinforce your understanding of Python's data handling and visualization libraries.

Link: https://archive.ics.uci.edu/datasets

Instructions

Data Set Selection

- **Choose one data set** from the available list (do **not** use the Iris data set).

For each data set, perform the following steps:

1. **Data Set Explanation and Preview**
   - Provide a brief explanation of what the data set is about (its origin, purpose, and key features).
   - Print a portion of the raw data (e.g., using head() or equivalent) to give an insight into the dataset structure.
2. **Bar Graph Visualization**
   - Create a **bar graph** that reasonably visualizes an aspect of the data.
   - Include cell text (or Markdown cells if using a notebook) to explain what the bar graph represents and interpret the visualization results.
3. **Line Graph Visualization**
   - Create a **line graph** that represents another aspect of the data.
   - Use cell text (observation) to provide a clear explanation of the line graph and discuss what insights can be drawn from it.
4. **Pie Chart Visualization**
   - Create a **pie chart** to visualize a different dimension of the data.
   - Add cell text (observation) that explains the pie chart's purpose and interprets the results effectively.
5. **Box Chart Visualization**
   - Create a **box chart** to visualize a different dimension of the data.
   - Add cell text (observation) that explains the box chart's purpose and interprets the results effectively.
6. **Saving Visualizations**
   - Save each visualization as an image file.
   - Name each image using the format: <your name + title>. For example, if your name is John Doe and the graph title is "SalesBarChart," the image should be named: JohnDoe_SalesBarChart.png.
7. **Code comments**

- Include necessary code comments throughout your scripts to explain your logic, steps, and any assumptions made during the visualization process.
- All graphs should include an x-axis label, a y-axis label, and a title. Please ensure that the font sizes for these elements are explicitly set to enhance clarity and visual appeal.

Submission

- Upload all your results, including code files and visualization images, to a GitHub repository.
- Ensure that your GitHub repository is well-organized and includes a README file that describes the project and provides instructions for reproducing your work.
- Submit your GitHub link to Canvas.