CS 445/545: Final project Stage 1

For the first stage of your final project, your task is to select a dataset that is both interesting and sufficiently rich to support detailed analysis. Your dataset should be relevant to your chosen project theme and capable of yielding meaningful insights when visualized. Once you have chosen your dataset, you will prepare a one-page report that introduces it in a clear and organized manner.

Your report should include the following elements:

- **Introduction to the Dataset:**
  Provide a brief overview of the dataset. Explain where the dataset comes from (e.g., a public repository, a government website, or an organizational database) and describe its primary purpose. For example, if the dataset is about global economic indicators, mention that it originates from the World Bank and is used to monitor economic trends across countries.
- **Attribute Descriptions:**
  List and describe each attribute (or variable) in the dataset. For each attribute, include:
    - **Type:** Specify the data type (e.g., numerical, categorical, text). This helps in understanding how the attribute can be analyzed and visualized later.
    - **Length/Range:** Provide details about the size or range of the attribute. This might include the number of characters for text data, the range of values for numerical data, or the categories available for categorical data.
- **Potential Significance and Insights:**
  Discuss why this dataset is interesting and what potential insights it might offer. Consider questions like:
    - What patterns or trends might you be able to uncover through visualization?
    - Are there any relationships between attributes that could be significant?

Deliverable Breakdown

- Team name, member name and logo in your report
- Datalink (could be your GitHub Link) in your report
- Your Report

Sample Project Ideas

To inspire your project direction, consider these sample ideas:

- **Global Economic Indicators:**
  Analyze datasets from sources like the World Bank to explore relationships between GDP, education levels, health metrics, and economic growth.
- **Urban Development Trends:**
  Use demographic and geographic data to investigate urbanization patterns, population density, and infrastructure development in cities.

- **Environmental Impact Studies:**
  Examine climate data, such as temperature trends, CO_2 emissions, and air quality indices, to study environmental changes and sustainability efforts.

Choose a dataset that not only interests your team but also allows you to integrate various visualization techniques learned throughout the course.

Effective Communication

When writing your one-page report, keep the following in mind:

Detail: Ensure every attribute is described thoroughly (including its meaning, range, and type).

Insight: Explain the potential insights the dataset might offer and outline preliminary research questions or directions for further analysis.

Visual Appeal: Although it's a written report, formatting and readability count. Use bullet points or numbered lists where appropriate to enhance clarity.

Rubric

Your submission will be evaluated based on the following criteria:

Completeness:
All required sections (Introduction, Dataset Description, and Potential Significance) are included and fully developed.

Depth of Analysis:
Each attribute is clearly described, and potential insights are thoughtfully discussed.

Helpful Links

Here is an excellent article which goes over the basics of a data visualization projects:
https://www.knowledgehut.com/blog/business-intelligence-and-visualization/data-visualization-projects
An article about top 10 data visualization project ideas in 2021:
https://365datascience.com/trending/data-visualization-project-ideas/
Another article with some sample data visualization project ideas:
https://www.projectpro.io/article/data-visualization-projects-ideas/471

Simplified Submission Template

Title:
Provide a concise title that reflects your project theme.

Team name:

Member name:

Logo:

1. Introduction to the Dataset:

- **Source:**
  e.g., "Data obtained from [source name]."
- **Purpose:**
  e.g., "This dataset is used to analyze [topic]."

2. Dataset Description:
For each attribute:

- **Attribute Name: xxxx**
  - *Type:* [Numerical, Categorical, Text, etc.]
  - *Length/Range:* [Details about value range or size]

3. Potential Significance and Insights:

- *1. Discuss potential research questions and the insights you hope to gain from this data.*
- *2. Discuss potential research questions and the insights you hope to gain from this data.*
- *3. Discuss potential research questions and the insights you hope to gain from this data.*