CDS6324: Data Visualization
Trimester 2, 2023/2024
Project: Creating Interactive Visualization

In this project, you will design and implement an interactive visualization application. You can choose the data domain as well as the techniques you wish to implement. For example, the ZipDeCode (https://benfry.com/zipdecode/) and Name Voyager (https://namerology.com/baby-name-grapher/) examples shown in class apply the interactive technique of dynamic queries – to the problem of uncovering patterns in zip codes and baby names. Similarly, treemap was originally developed to display hierarchical file structure, and were later adapted to the problem of interactively depicting a map of the stock market. We have seen many other examples of interactive visualization techniques in class. The goal with this project is not only for you to gain hands-on experience implementing a visualization technique, but also for you to think about the effectiveness of the specific visualization techniques you reimplement in the context of the data domain you work with.

This is a group project of 2-3 students. Your implementation should contain at least six meaningful visualization charts. The visualization charts in this project should be completed using D3js but you are allowed to combine the charts with other libraries such as Javascript front-end library React to make the interface of your dashboard more captivating and interactive.

Project Proposal Due: 30th May 2024, 11:59pm
Final Project Due:21st June 2024, 11:59pm

Deliverables

For this project you are responsible for the following:

1. **Describe the data domain and storyboard the interaction techniques you will use.**

   Start by choosing the data domain and interactive visualization techniques you will implement. Think about why the domain and the techniques are a good match for one another. Then write a description of the data domain and the interactive visualization application you will build. The description should include a storyboard of the visualization charts/dashboard/interface you will create. Be sure to explain the features of your application. Most importantly you should explain why the interaction techniques you will implement will be effective in the context of your data domain. The goal of this exercise is to think through the various concerns that go into the implementation of the visualization and interaction techniques. This is why it is important that you perform this task first, before building the visualization. As part of your prototyping, you may find it helpful to use existing tools (e.g., Tableau) to explore the data and test multiple visualization strategies.

2. **Implement your visualization design with D3.**

   Note: while you are free to use non-programming tools (e.g., Tableau) to explore your data set and try out design ideas, you must program the final application in D3 and submit your code.

3. **Produce a final writeup.** Your final submission should include:

    1. **Project Report**
        - An overview of your interactive visualization application, including the choice and description of dataset(s) and the associated attributes, the theme, and the questions your visualization attempts to answer.
        - Describe the selected dataset(s) and the data transformation trend, pattern, or relationship you found within your data.
        - Describe your visualization charts/dashboards. Provide an explanation of changes between the storyboard in the proposal and the final implementation. Highlight the valuable findings or insights.
        - Provide details on the development processes. Include a breakdown of how the work was split among the group members. Include a commentary on the development process, including answers to the following questions: Roughly how much time did you spend developing your application (in man-hour)? What aspects took the most time?
        - State the Data Visualization design principles (i.e. Shaffer's 4Cs or Cole Nussbaumer's 4As) you used as a guide for your visualization and describe how the visualizations of your data story fulfil each of the design principles.

    2. **Source Code**

        - The source code for your application. Please ensure that the code submitted is in working order. If any special instructions are needed for building or running your software, please include them in the write-up or provide a readme file. Marks for coding-related marking criteria will not be allocated for code that does not work.

Data

You can choose one or more datasets that fulfills one of the following themes:

- **Earth and Nature** (e.g. climate change, air pollution, endangered species)
- **Economy and Growth** (e.g. Covid-19 socio-economic impact, GenZ employment analysis)
- **Science and Technology** (e.g. electronic vehicles, rise of LLMs)

The dataset(s) should contain at least **10 attributes and 3,000 records**.

Here are some data resources that you can explore to choose and appropriate dataset for your project:

- World Bank Open Data: https://data.worldbank.org/indicator
- The Humanitarian Data Exchange: https://data.humdata.org/
- Dataportal Asia: https://dataportal.asia/home
- DOSM Malaysia: https://open.dosm.gov.my/data-catalogue
- Kaggle: https://www.kaggle.com/

Important Note: There should not be more than 2 groups using the same dataset(s).

Milestones

Project Proposal: You should first prepare a project proposal and submit it to MMLS by **30th May 2024, 11:59pm**. Your project proposal should include the following details:

- **Project Members**
- **Project Title**
- **Proposed Dataset and Its Attributes**
- **Project Description**: Storyboard of the visualization charts/dashboards/interface you will create. Explain the features of your application and why the interaction techniques you will implement will be effective in the context of your data domain.

Final Project: You should submit a zip file that contains the report and source code of your visualization project to the Google Classroom by **21st June 2024, 11:59pm**. A presentation session for the project will be arranged on Week 14.

Project Rubrics

PROPOSAL	
Criteria	**Marks**
Data used is appropriate and sufficient for analysis	**1**
Clearly defined topic that answers a specific question or facilitates decision making (Does the visualization tells a good story?).	**1**
Description and design of proposed visualization	**3**
	5

PROJECT	
Criteria	**Marks**
Visualization Design	
Applies appropriate graphic variable types (visual encoding) for the data type and scale.	**2**
Appropriate choice of charts	**2**
Legends should describe and explain every graphic variable type employed.	**1**
Animation	**2**
Interactivity - the visualization is usable and actionable	**4**
Dashboard design	**2**
Design Principles and Impact	
Everything in the visualization conveys some information to the viewer (Data-Ink ratio)	**2**
Graphics integrity, and impact of the findings and analysis	**2**
Basic visualization rules and best practices are consistently applied (Shaffer's 4Cx or Cole Nussbaumer's 4As)	**4**
Data Storytelling	
Data used is appropriate and sufficient for analysis	**1**
Clearly defined topic that answers a specific question or facilitates decision making (Does the visualization tells a good story?).	**1**
Design of Data Story	**2**
Aesthetics	
Aesthetics appeal of visualization	**2**
Report	
Quality and Comprehensiveness of Report	**3**
TOTAL	**30**

PRESENTATION	
Criteria	**Marks**
Organization	**1.5**
Delivery	**2**
Preparedness / Participation / Group dynamic	**1.5**
TOTAL	**5**