Title: MM802 - Visualization Mini-project

Date: March 19, 2024

Name: Soumya Kulal, Tobias

Project Title: Canadian Post Secondary Institution Data Visualization

1 Abstract

This report underscores the critical significance of the "Canadian Post Secondary Institution Data Visualization" project, addressing a notable gap in comprehensive data visualization within the Canadian post-secondary education sector. With Canada's education services valued at an astounding 15.5 billion Canadian Dollars annually, the urgency to provide accessible and informative platforms for decision-making is paramount. Leveraging data sourced from the Canadian government's official website, the project aims to create an interactive visualization platform using modern web technologies like HTML, CSS, JavaScript, and D3.js, along with PHP for server-side connections, promising a user-friendly experience.

Beyond data visualization, the project holds the potential to benefit various stakeholders, including prospective international students seeking suitable universities, education agencies facilitating student placements, and the Canadian government aiming to bolster the competitiveness of its education sector. By facilitating informed decision-making and fostering transparency, this initiative has the capacity to contribute significantly to the growth and sustainability of the Canadian economy. In conclusion, this project represents a pivotal advancement towards enhancing accessibility and transparency in the Canadian education landscape.

2 Introduction

Education services have a major impact on the Canadian economy, valued at over 15.5B Canadian Dollar annually according to the EduCanada 2017 Report. However, despite its tremendous impact, the data visualization of this sector is often neglected. With numerous post-secondary institutions across Canada, it is challenging for international students to choose the most suitable university and program that aligns with their needs. By visualizing post-secondary institution information across Canada and providing detailed information and interactive displays, it can be highly beneficial for multiple parties, including prospective international students, education agencies, and the Canadian government.

In this project, we are visualizing a dataset from the Canadian government's official website for the list of post-secondary academic institutions and programs validated by the Public Service Commission [1]. This dataset includes the list of all post-secondary institutions in Canada as per 2024, along with detailed information such as institution name, academic level, area of study, program of study, program length, duration of term, as well as the institution's contact person and their email address.

With this comprehensive dataset, our aim is to create interactive visualizations that are easily understandable by users. Leveraging current web-based technology, we are using Hypertext Markup Language (HTML) as the basis to create the website, with Cascading Style Sheets (CSS) to enhance the visual layout, JavaScript to make it interactive, and Hypertext Preprocessor (PHP) for the server-side to ensure all data used can be connected to the frontend accordingly. Additionally, we aim to enhance user experience by utilizing jQuery library animations. For the main data visualization development, we are leveraging one of the most popular data visualization libraries, D3.js. This JavaScript data visualization library offers various layout options such as bar charts, pie charts, and more, making it a perfect option for this project due to its variety, flexibility, and widespread use in current data visualization practices.

3 Project Status

As of the present moment, our project has successfully acquired the requisite dataset from the official Canadian government website [1]. This dataset contains comprehensive information regarding post-secondary institutions across Canada, encompassing crucial details such as institution names, academic levels, areas of study, program specifics, and contact information. We are currently in the process of integrating this dataset into our platform to facilitate client-side data visualization.

Given the constraints imposed by time considerations, our immediate focus is on implementing interactive client-side data visualization features. This approach enables users to engage dynamically with the data, fostering a more immersive and informative experience. However, it is important to note that our ambitions extend beyond the current scope. We aspire to develop a server-side portal that empowers each post-secondary institution in Canada to directly input and manage their institutional data. This future enhancement will not only streamline the data management process but also ensure the most up-to-date information is available for visualization.

Furthermore, it is worth highlighting that rigorous testing has been conducted on the project's code to ensure robust functionality and performance. Testing has been carried out across both local and live server environments, validating the integrity and reliability of our solution. Interested stakeholders can access and explore our project's progress via the provided link below: . This demonstration serves as a testament to our commitment to transparency and accountability in delivering a high-quality data visualization solution.

4 Development Environment

Network Configuration

- **Server:** XAMPP provides a local Apache server for hosting web applications. This server is accessed via localhost or 127.0.0.1 on the development machine.
- **Clients:** The development machine acts as the client, accessing the web applications hosted on the local Apache server.
- **Mobile Devices:** Mobile devices can also access the locally hosted web applications for testing purposes. Developers can use features like port forwarding to access the local server

from mobile devices connected to the same network.

Software Used

- **XAMPP**: XAMPP is the primary software tool used for creating a local web server environment. It includes Apache as the web server, MySQL/MariaDB as the database server, PHP, and Perl.
- **Text Editor/IDE**: Developers typically use text editors or IDEs like Visual Studio Code, Sublime Text, or PhpStorm for writing code and we used the Notepad as our text editor.
- **Web Browsers**: Standard web browsers such as Google Chrome, Mozilla Firefox, or Microsoft Edge are used for testing and debugging web applications hosted on the local server and we used Google Chrome as our browser.
- **Database Management Tools**: Tools like phpMyAdmin, included in XAMPP, are used for managing MySQL/MariaDB databases so we used the same.
- **Command Line Interface (CLI)**: Developers use the command line interface for tasks like starting/stopping the Apache server, managing databases, and running scripts.

5 Development Work

Software Libraries and Tools Used:

- **D3.js:** Used for data manipulation and DOM manipulation in JavaScript. It's primarily utilized for reading data from CSV files (institution.csv) and populating dropdown menus dynamically based on the data.
- **Plotly.js:** Utilized for creating interactive data visualizations such as pie charts and bar charts. It's used in the `secondpage.html` to generate the program type pie chart and the duration of term bar chart.
- **XAMPP:** The development environment includes XAMPP for setting up a local Apache server environment. XAMPP provides Apache as the web server for hosting web applications.

How to Use the Project:

- **First Page (`index.html`):** Users can select an institution, academic level, and program length from the dropdown menus. Upon making a selection, they can click the "Go to Second Page" button to navigate to the second page (`secondpage.html`).

- **Second Page (`secondpage.html`):** This page displays visualizations based on the selections made on the first page. Users can view the program type pie chart and the duration of term bar chart corresponding to their selections.

Sample Output Figures:

- The sample output figures would include screenshots or visual representations of the dropdown menus populated with options based on the data from `institution.csv`.

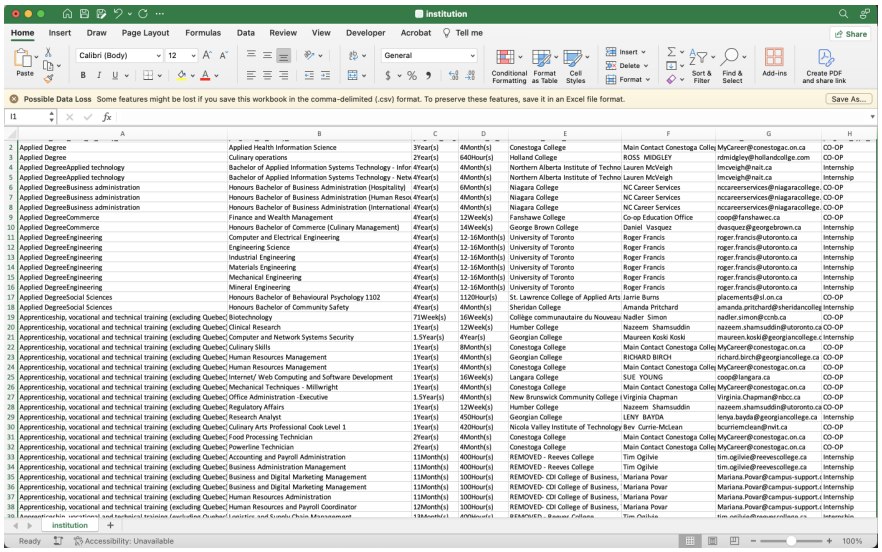

Figure 1: Institution list CSV dataset

- Additionally, sample output figures would include screenshots of the visualizations generated on the second page, such as the program type pie chart and the duration of term bar chart.

Limitations:

- **Local Environment Dependency:** The project relies on a local server environment provided by XAMPP. Therefore, it may not be suitable for deployment in a production environment without appropriate adjustments.
- **Limited Data Interaction:** While the project provides dynamic dropdown menus based on the data from `institution.csv`, it may not handle large datasets efficiently, leading to potential performance issues.
- **Solely Frontend Implementation:** The project appears to be primarily a frontend implementation with no backend functionality. As a result, it may not support dynamic data fetching or real-time updates without additional backend integration.

Overcoming Limitations:

- **Local Environment Dependency:** Deploy to a cloud-based server or use Docker for containerization.
- **Limited Data Interaction:** Implement server-side scripts or APIs, and utilize database systems for data management.
- **Solely Frontend Implementation:** Integrate a backend framework, implement RESTful APIs or GraphQL, and utilize modern frontend frameworks.

6 Concluding Remarks

Accomplishments:

- Developed a web application with dynamic dropdown menus and interactive data visualizations using D3.js and Plotly.js.
- Established a local server environment using XAMPP, enabling data manipulation and visualization on the client-side.
- Enhanced user experience by providing a seamless navigation flow between the first page and the second page of the application.

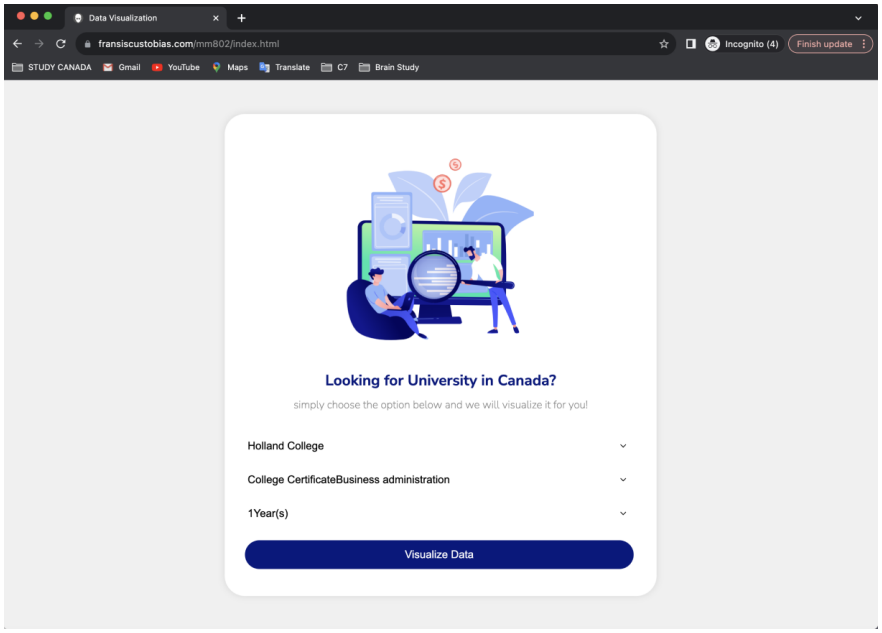

Figure 2: Institution search homepage

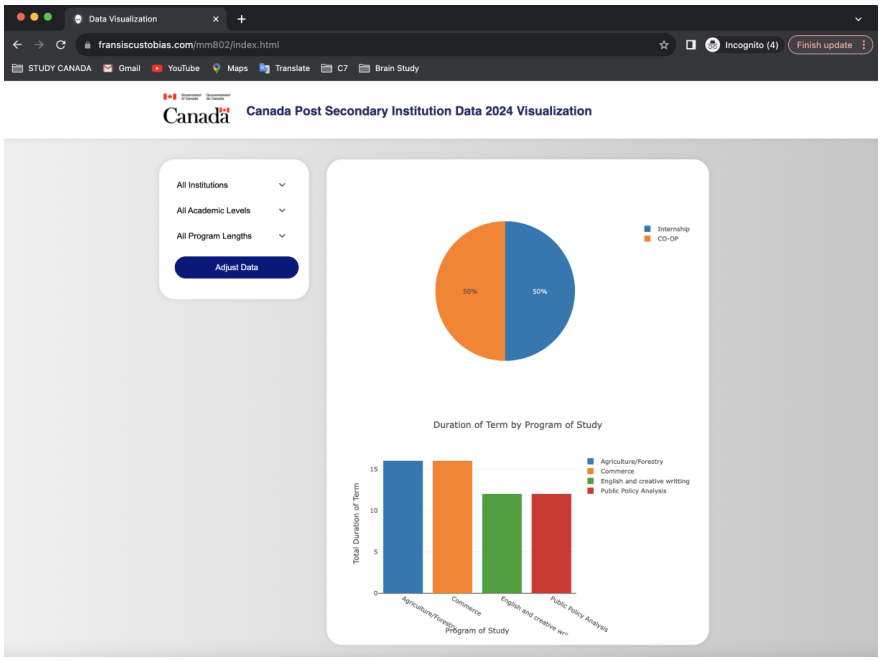

Figure 3: Institution data visualization

Lessons Learned:

- Understanding the importance of frontend-backend integration for dynamic data handling and real-time updates.
- Gained proficiency in utilizing JavaScript libraries like D3.js and Plotly.js for data visualization tasks.
- Learned techniques for populating dropdown menus dynamically based on external data sources.

Possible Future Extensions:

- Implementing a backend server using frameworks like Node.js or Flask to enhance data interaction capabilities.
- Integrating a database system for persistent data storage and retrieval, enabling scalability for larger datasets.
- Enhancing user interactivity with features like filtering, sorting, and searching within the visualizations.
- Adding authentication and user management functionalities to secure the application and personalize user experiences.

Notes

The project utilizes various tools and libraries such as D3.js, Plotly.js, and XAMPP.

References

[1] Government of Canada. List of post-secondary academic institutions and programs validated by the Public Service Commission. https://open.canada.ca/data/dataset/ab36a4ab-9b69-49b1-8be6-6faa3f0d0c67/resource/5d42f06e-2484-42e0-8681-4a360198202c/download/cfppsc$_{c}oop02 - bom - 2024_{0}3_{0}1.csv, 2024.$