**ENSF 592 Project Report**

**Aron Saengchan and Timothy Mok**

**Summary**

The project focuses on a dataset of statistics for Ontario public libraries made available by the Government of Canada. Approximately 380 libraries self-report the data, providing general information like the library name, code and address, as well as library specific statistics like number of cardholders and the number of English and French titles available at the library.

The program opens with a main menu where the user can choose to get general information for a library or code, locate a library close to them, or access yearly statistics. After each action, the user has the choice to perform that action again on a different library or year or to return to the main menu. The main menu also has a quit option for when the user is finished.

Within the first option – Branch Information Search, the user is prompted to enter a library name or code and the program will provide them with the address, website, and the number of print and electronic resources available.

The second option – Library Locator, prompts the user to enter their postal code. If there are any libraries in their area, the program will ask whether they want a library with more resources or one that would be better for studying or working. The program will then list the top five libraries for their selection and allow them to use Branch Information Search to print information for that library.

Finally, the Access Yearly Archives option prompts the user to enter a year between 2017 and 2019 and provides them with various statistics for that year, from statistics for all the libraries, to which library had the most of each kind of resource that year.

**Code Specifications**

The concatenate method was used to compile the three data sheets together. A merge method was used to add the new columns to the dataset, and also to merge data used to create the plot. The merged dataset has a hierarchical index of Library Full Name, Library Code then the Year which it is then sorted by. Null values in the Street Address column found with a mask and filled in with forward fill.

Within the Yearly Archives section, the describe method is used to display general statistics for a given year. The groupby method and aggregation computation sum were used on each column which was then concatenated to the table. A pivot table was created to display the total resources per cardholder based on service region and type. Finally, Matplotlib is used to create subplots based on the total English and French resources available each year.

**References**

*Ontario public library statistics, 2017-2019,* Government of Canada, June 2021. [Online]. Available: <https://open.canada.ca/data/en/dataset/363fff31-6a07-41eb-9922-e9b64192b08b>