****

**Geospatial Operational Insights**

**for National Library Board (NLB)**

**Interim Progress Report**

**ANLY482 – Analytics Practicum AY16/17 Term 2**

**Prepared by:**

*Team Bookworms*

GAURI BHATNAGAR

THAVANESAN S/O SIVANANTHAN

WANG TIANTONG

**Supervised by:**

Prof KAM Tin Seong

*Associate Professor of Information Systems*

Table of Contents

[Project Recap 3](#_Toc475917978)

[Background Information 3](#_Toc475917979)

[Project Sponsor & Liaison Information 3](#_Toc475917980)

[Project Motivation & Problem Statement 3](#_Toc475917981)

[Project Objective 4](#_Toc475917982)

[Desired Outcome 4](#_Toc475917983)

[Data Exploration and Preprocessing 4](#_Toc475917984)

[Data provided by NLB 4](#_Toc475917985)

[Summary of Data cleaning 4](#_Toc475917986)

[Additional Data 7](#_Toc475917987)

[Surrounding Facility Dataset: 8](#_Toc475917988)

[Transportation Accessibility Dataset: 8](#_Toc475917989)

[Geographical Dataset: 8](#_Toc475917990)

[Field Research Data – Internal Facilities and Patron Preference 9](#_Toc475917991)

[Improvements to be made 9](#_Toc475917992)

[RFM Analysis 10](#_Toc475917993)

[Variables used for Huff’s Model 10](#_Toc475917994)

[MRT Centrality 10](#_Toc475917995)

[Type of Collection in a Library 11](#_Toc475917996)

[Presence of Internal Library Facilities 13](#_Toc475917997)

[Presence of Carpark Accessibility 13](#_Toc475917998)

[Rshiny Dashboard Additions 13](#_Toc475917999)

[Change of Template used for Dashboard 14](#_Toc475918000)

[Addition of new tabs under “Catchment Area Analysis” 15](#_Toc475918001)

[“Overview Analysis of Libraries” Tab 15](#_Toc475918002)

[“Individual Library Analysis” Tab 18](#_Toc475918003)

[Adaptation of the Multiplicative Competitive Interaction Model 20](#_Toc475918004)

[Revision of Methodology 21](#_Toc475918005)

[Literature Review 21](#_Toc475918006)

[For Data Preparation & Model Calibration 22](#_Toc475918007)

[R Package - Multiplicative Competitive Interaction (MCI) Model 22](#_Toc475918008)

[R Programming Language 22](#_Toc475918009)

[JMP Pro 13 22](#_Toc475918010)

[For Dashboard Visualization 23](#_Toc475918011)

[Tableau 23](#_Toc475918012)

[R Shiny & R Markdown 23](#_Toc475918013)

[Javascript 23](#_Toc475918014)

[Leaflet 23](#_Toc475918015)

[Revised Scope of Work 23](#_Toc475918016)

[Revised Work Plan 24](#_Toc475918017)

[References 25](#_Toc475918018)

Project Recap

## Background Information

The National Library Board (NLB) is a statutory board of the Ministry of Communications and Information in Singapore, tasked to manage the public libraries and lead them into the information age where non-print resources are making their mark. With a substantial amount of title collections in the regional and public libraries which it manages, the NLB aims to aid as a point source for Singaporeans to connect with the precious archives of the past. The NLB has managed to disperse its 27 Regional and Community Libraries strategically across Singapore.

## Project Sponsor & Liaison Information

Our project sponsor is the National Library Board.

## Project Motivation & Problem Statement

In this age of information, we see an increase in necessity for people and businesses to have a greater access to space and resources to further their personal and corporate needs. Libraries in Singapore are no exception to this trend.

To meet this necessity, NLB may not only have to find ideal locations within districts for additional libraries, but it also must manage their resources wisely to meet public demand with respect to each location. Currently there exists this difficulty in measuring the operational readiness of the libraries given that the measure of public demand is not in dollars and cents; unlike that of typical corporations and organizations.

Constant change management in the location and resources within the libraries have led to the dire need of a reliable and standardized method which can measure the effectiveness of past policies, as well as an accurate predictive model to conduct what-if analyses for future. Thus, a user-friendly system which displays geospatial information based on significant data evidence would provide operational insights valuable to the NLB.

Besides taking geospatial information into consideration, NLB is also motivated to leverage on an effective model (which uses relevant metrics regarding book collections and user transaction information) to accurately measure the attractiveness of existing and potential libraries.

## Project Objective

Develop a geospatial dashboard that would facilitate the NLB’s evaluation of different potential library siting locations through better prediction of the catchment and number of patrons it can potentially attract.

Assess earlier models which have been used to evaluate the attractiveness of libraries and refine them to increase the accuracy and reliability of results.

## Desired Outcome

The geospatial dashboard will potentially contribute to:

* the planning/promotion of content and services that are localised to the specific population to be served
* the evaluation of the attractiveness of existing and future NLB library locations

# Data Exploration and Preprocessing

## Data provided by NLB

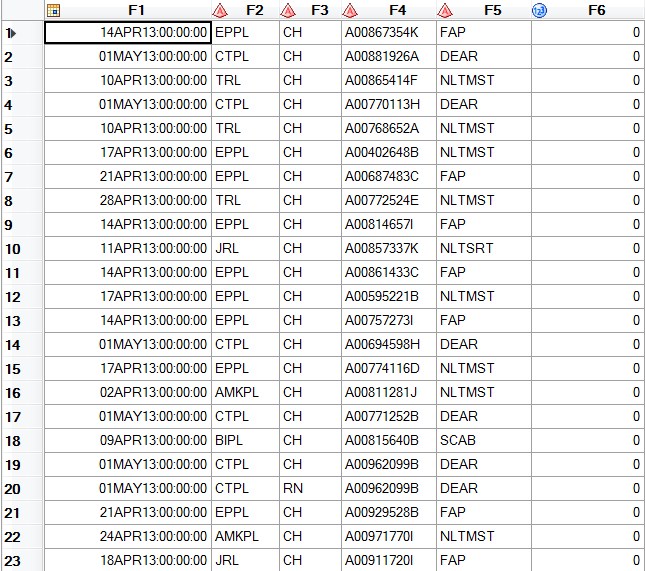
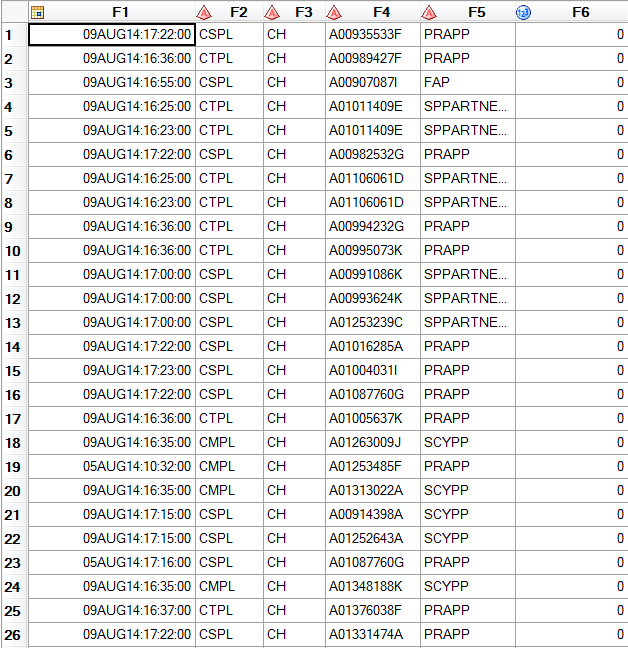
Currently we have received the raw data provided by our sponsor including the following files:

* Collection\_Dataset\_FY13 and FY14.xlsx
* Patron\_Headers.csv
* Patron\_Dataset\_FY13.csv & Patron\_Dataset\_FY14.csv
* TXN\_Headers.csv
* TXN\_FY13.csv & TXN\_FY14.csv

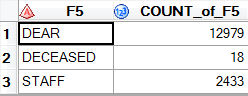
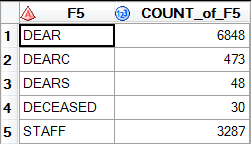
## Summary of Data cleaning

Our team performed data cleaning on the raw data provided to us by NLB and listed our observations below:

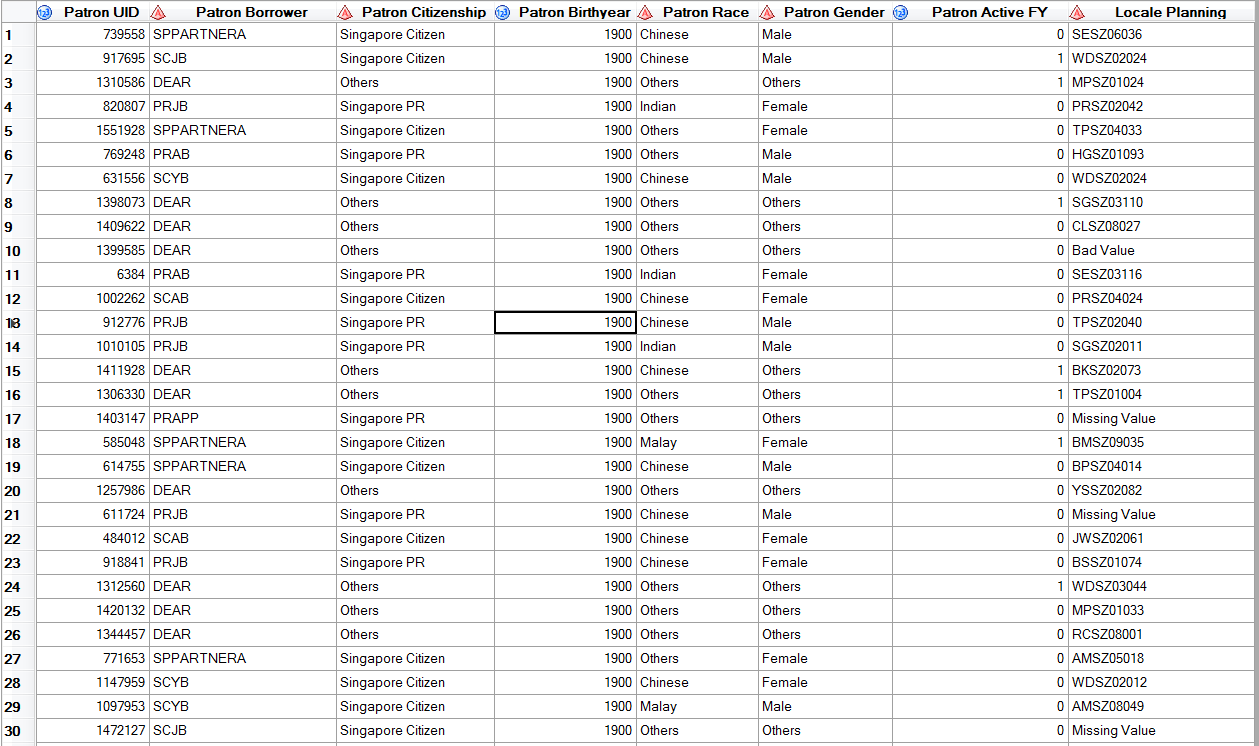
1. Both the TXN\_Y13 & TXN\_Y14 datasets displayed that the transaction’s timeline from each year’s April to the next year’s April. (12 months)
2. In both the TXN\_Y13 & TXN\_Y14 datasets, there were rows where the Patron UID was set to “0”, whereas both the Patron\_Dataset\_FY13 & Patron\_Dataset\_FY14 datasets do not contain any data of a Patron with UID “0”. After exploring the data, 1335 counts & 65 counts of rows of Data contained Patron UID set to “0” in the TXN\_Y13 & TXN\_Y14 datasets respectively.

1. In the TXN\_Y13 dataset, there were Patron Borrower Category Code types such as “DEAR”, “DEARC”, “DEARS” & “Deceased”. Based on the previous team’s data cleaning we discovered that “DEAR” should be removed from the dataset completely as it refers to institutional partnership programmes and NLB suggested to remove all records with Patron Borrower Category Code set to “DEAR” from further analysis. However, there was no mention of the other codes “DEARC”, “DEARS” & “Deceased”, thus for the time being our team has decided to remove those records as well. The frequency of these codes are displayed below for TXN\_Y13 & TXN\_Y14 datasets respectively.



1. In the TXN\_Y13 Transaction Dataset, 3 records have Branch Code set to “Bad Value” and 3,065 records set to “Missing Value”. In the TXN\_Y14 Transaction Dataset, 9,467 records have Branch Code set to “Bad Value” and 1,600 records set to “Missing Value”.
2. Based on the updated Collection dataset, it was confirmed that there are currently 26 active libraries for NLB and therefore transactions in both the TXN\_Y13 & TXN\_Y14 datasets which did not belong to these 26 branch codes were removed, such as
   1. 07LKCRL
   2. 08LKCRL
   3. 11LKCRL
   4. ORN
   5. RN
   6. RU
3. Removed any transactions in both the TXN\_Y13 & TXN\_Y14 datasets that stated that the birthyear of the Patrons is “1900”, “2015”,”2016” as that referred to values being set to the year that the institutional programme was set up.



## Additional Data

Based on the senior’s group’s work, our team will conduct the model building and calibration using geospatial data derived from online sources (e.g. https://data.gov.sg) to evaluate and update the corresponding attractiveness measurements. Our team will also do field research to collect relevant data on library layout design, collection availability, library facilities and patron preferences in person. At the first stage, the data can be categorized into 4 categories elaborated below.

### Surrounding Facility Dataset:

• Geographical location of Shopping Malls/ Plazas

As indicated in the senior’s group’s report, there is “positive inter-store externalities generated by the shopping malls that operate near the library (Brueckner, 2011), as more consumers visit the shopping malls, the patronage level of the nearby library will likely follow a similar increase.” Therefore, our team will keep studying the significant effect on the patronage of the libraries from the distribution of various shopping malls.

• Geographical location of Primary Schools/ Secondary Schools/ Junior Colleges

As one of the largest groups visiting libraries, students are nonnegligible given that they are likely to spend time in the libraries after school hours and during examination period. Hence, our team will also have a deep look at the impact on the patronage of the libraries based on the location distribution of nearby educational institutions (primary schools, secondary school, junior colleges) using the data derived online.

### Transportation Accessibility Dataset:

• Geographical location of MRT Stations (A greater weight will be assigned to MRT interchanges in the analyses)

• Geographical location of Bus Stops & No. of Bus Services Provided

In order to evaluate the likelihood for a patron to visit a library, the accessibility of transportation also plays an important role. With an easily accessed public transport network connected to a library, there will be less hindrance and thus a higher probability for a patron to visit the library. To analyze more deeply, the impact of public transportation may also vary between different neighborhoods where people are of different social and financial levels. Therefore, our team will embrace the available transportation dataset (MRT and bus stops) in our model with weight assigned to better measure and predict the attractiveness of the libraries.

### Geographical Dataset:

• Building within costaloutline.shp

As mentioned above, although the subzone clustering analysis conducted by the senior’s team returned a relatively executable model, our team aims to build up on the next level and present a more precise and accurate analysis. In terms of the geographical dataset, subzones no longer meet our demand due to the wide coverage of each subzone and the inequality analysis on patrons from different parts within the same subzone. Therefore, our team will utilize the geospatial data at HDB level (after transformation) and link it to the post-geocoding patron’s data so as to better analyze the patronage of each library.

### Field Research Data – Internal Facilities and Patron Preference

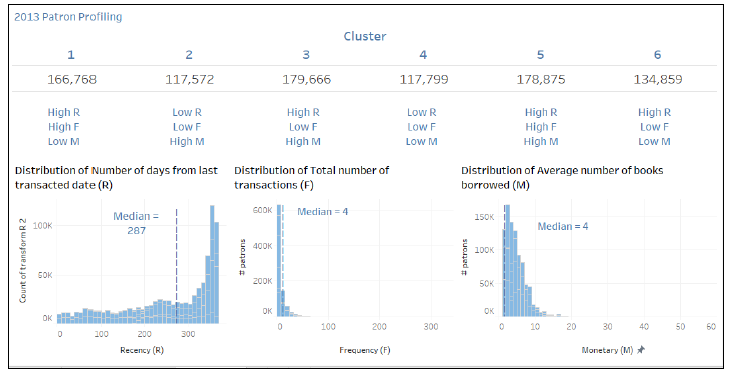
In real-world analytics, complete data may not be available for our target features so we need to do some filed research to collect useful data ourselves. In this project, our team believes that beside surrounding facilities and public transportations, the libraries’ layout design and their internal facilities are also part of their attractiveness. Based on our own research, patrons may prefer one library just due to its availability of some specific types of facilities, like quite areas and large study space etc. What’s more, in addition to the collection size, the collection categories and the accessibility of the collections also contribute to the attractiveness based on patrons’ preferences. Therefore, our team is planning to conduct more and different field researches and interviews to collected relevant data regarding this aspect.

# Improvements to be made

After taking over the previous team and looking through their report, we discovered that though they had done a lot of analysis of the data, however they were unable to match it back towards the prototype of the dashboard.

## RFM Analysis

After conducting clustering on the patron data set using the three attributes, Recency, Frequency, and Monetary (denoted as R, F, and M respectively), the previous team was able to have obtained 6 clusters, shown below.



Our team believes that though the analysis is very insightful for someone with analytical background, however in order for this to be meaningful to the NLB staff, there needs to be thorough explanation of what each of these clusters signify. Also, incorporation of these clusters in the dashboard with explanations would be much more user-friendly as well for the NLB staff members to visualize their patron distribution across different libraries.

## Variables used for Huff’s Model

In the end the previous team decided to use the following factors to be used in the Huff’s model: Number of MRT stations within 1KM of the library, Number of tuition centres within 1KM of the library, Number of shopping malls within 1KM of the library, Collection size of the library, Distance of subzones to library. However, our team believes that there are a few other factors that can affect the attractiveness of a given library such as the sum of the MRT Centralities within 1KM of the library, the type of collection a library holds, presence of internal library facilities and the presence of carpark accessibility near the library.

### MRT Centrality

Centralities in social network analysis can be used to quantify various behaviors of nodes in a network. In the context of MRTs, centralities can be assigned to them to understand which MRT would result in a higher attractiveness for the library if it were located it. There are a few types of centralities that our group intends to incorporate, such as Degree, Betweenness, Closeness and Eigenvector centrality. Our current assumption is that a lower Closeness centrality and higher centrality for the rest would lead to a higher number of patrons visiting the library. The explanations of the types of centralities are written below.

1. Degree Centrality: Number of other MRTs that are connected to a given station

* By measuring degree centrality in the project context, we are able to evaluate the accessibility of a specific library based on its nearby MRT, which might drive a higher patronage.

1. Betweenness Centrality: Number of times a MRT acts as a bridge along the shortest path between two other MRT stations

* By measuring betweenness centrality, we are able to evaluate the easiness level for a potential patron, who are not from the same MRT line as a specific library, to visit the library. i.e. the higher the betweenness level, (we assume) the more likely that people are visiting the library

1. Closeness Centrality: Average length of the shortest path between a MRT and all other MRT stations in the graph

* By measuring closeness centrality, we are able to evaluate the easiness level for a potential patron to visit a specific library, i.e. for the MRT station near the library, the higher the closeness centrality is, the average time and cost for a potential patron to travel to the library is lower, which may result in a higher patronage

1. Eigenvector Centrality: Assigns relative scores to all MRTs based on the concept that connections to high-scoring MRTs contribute more to the score of the MRT in question than equal connections to low-scoring MRT

* Weighted measurement combining the above

Approach would be to analyze the cumulative centralities of the MRTs within a 1KM buffer of a library which would be used in the MCI model.

### Type of Collection in a Library

Our group believes that the collection type and proportion of books a library holds of each genre can greatly affect the patron influx in a given library. Therefore, by analyzing the borrowed items’ type with corresponding borrowers’ profile, we expect to find possible correlation between patron profile (race, sex, age group), favored collection types, which may be helpful for our sponsor. However, we currently do not have any data from NLB which we can use to identify the collection type therefore this analysis is just a possibility for now.

Presence of Internal Library Facilities

An American Library Association Survey ranked entertainment and study spaces as the top 2 reasons for individuals visiting a library. (35% and 28% of survey participants respectively). Therefore, it can be said that the attractiveness of a given library is not limited to just external factors such as transportation, presence of MRTs/Tuition Centers, it is also dependent on the internal facilities provided by a given library. Therefore, our team plans on testing this theory by including this factor in the calculation for attractiveness of a library.

### Presence of Carpark Accessibility

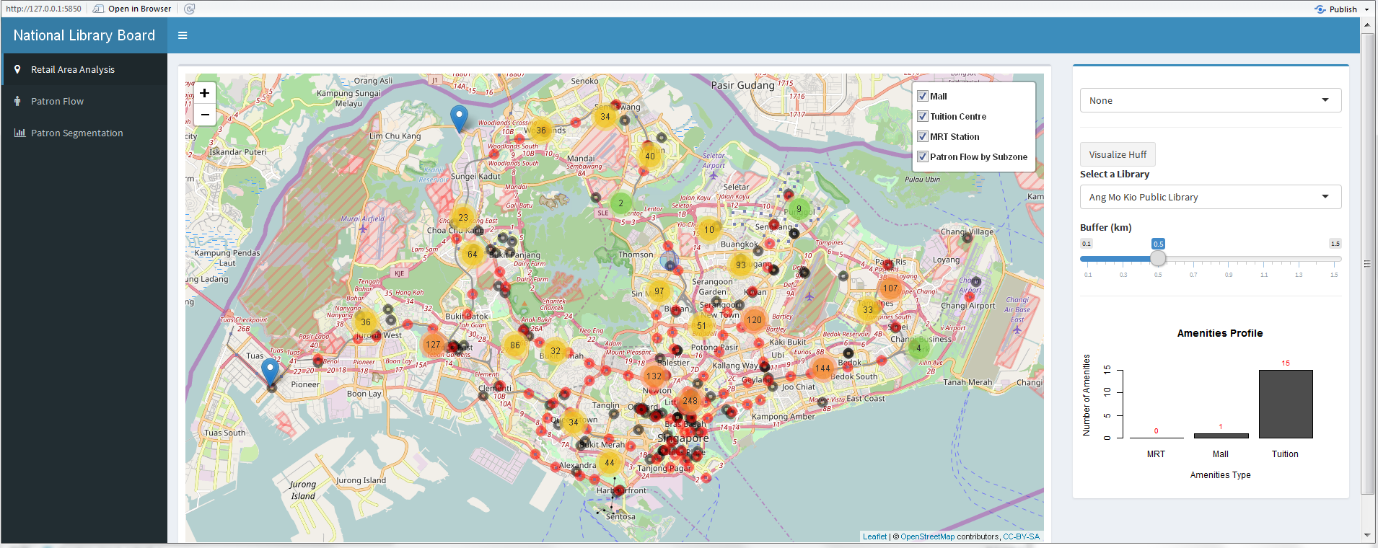
The libraries of NLB have three main library types, libraries that are located in Malls, those which are Stand-Alone libraries and lastly Regional libraries. Based on superficial analysis, it was found that generally there are more books borrowed from Regional libraries followed by Stand-Alone libraries and lastly by libraries located in Malls. There may be many different reasons which can lead to this behavior, however a potential factor worth looking deeper into would be the presence of carpark availability. This could possibly increase the attractiveness of visiting a library as well. This is what our team intends on looking into as well.

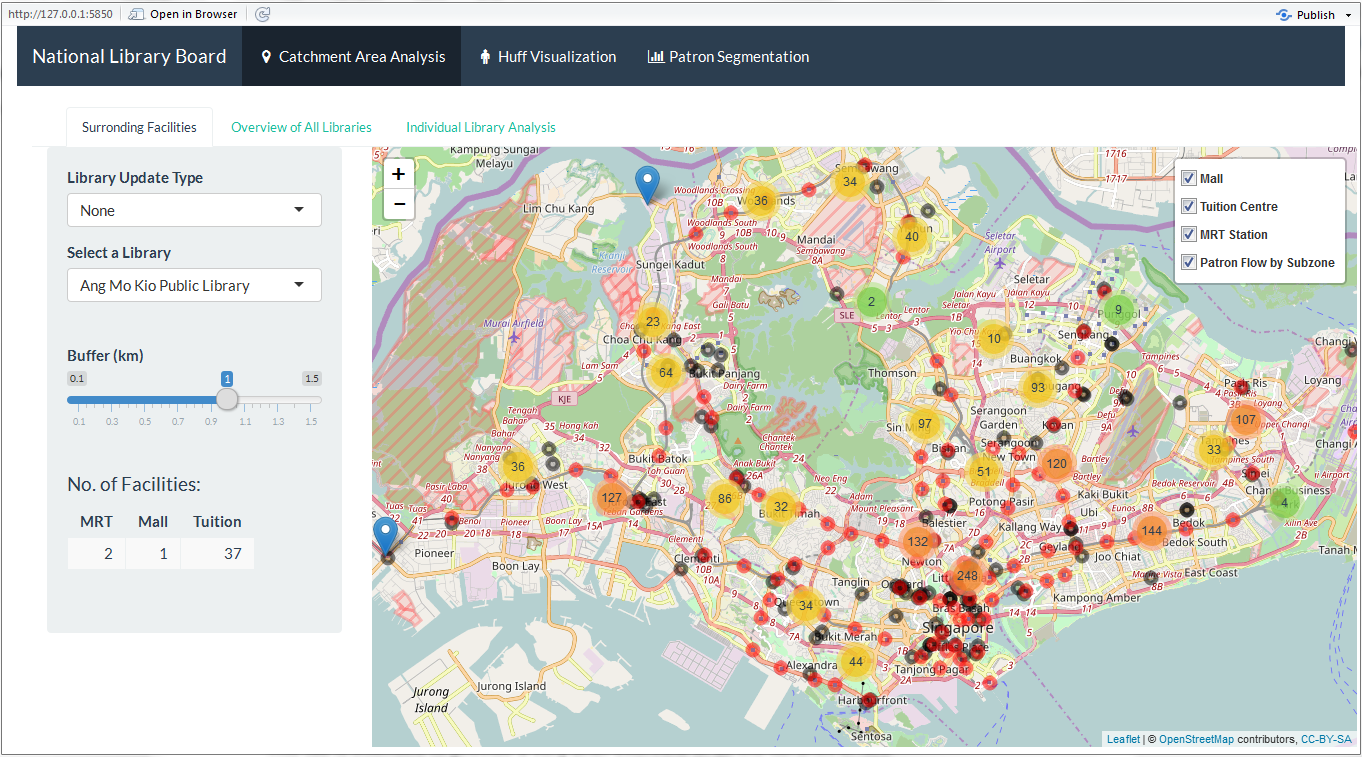
# Rshiny Dashboard Additions

After looking at the previous team’s dashboard, our team made a few UI changes to the dashboard to make it more user-friendly and utilize more space on the dashboard. All the UI changes are listed below.

## Change of Template used for Dashboard

Initial Dashboard Template:

Current Dashboard Template:



Changes include

1. Renaming of “Retail Area Analysis” to “Catchment Area Analysis” due to the incorrect word usage
2. Disabling of Zoom function on the patron flow map, so that users can have an overview of the patronal flow across the whole island without the need to move and zoom in/out the map
3. Removal of multiple layers on patron flow map to remove duplicated functions
4. Additions of many visualisation functionalities

## Addition of new tabs under “Catchment Area Analysis”

The objective of the dashboard requested by NLB was to be able to evaluate the attractiveness of existing and future NLB library locations, this would be most apt for the strategic team of NLB. Our team thought that the dashboard should not only be able to serve the strategic team of NLB, but also the common library branch heads, where they can use the dashboard to visualize their library statistics. These visualizations can not only help in understanding simple day to day behaviors of patrons but can also aid in minor changes the library might decide to take for existing libraries in order to reach certain KPIs. The addition of the next two tabs is the current progress of the team, however in the near future the goal would be to work towards the mock-up of the dashboard shown previously which would be a one stop solution.

### “Overview Analysis of Libraries” Tab

This tab is intended to provide a very brief overview of all the libraries in NLB, it includes interactive graphs which allow the users to drag and select points within the graph to view more details about a library. The three main type of visualizations provided currently are, Collection Size vs No. of Books Borrowed, Floor Size vs No. of Books Borrowed and Identification of Seasonal Months.

#### Collection Size vs No. of Books Borrowed

This visualization plots the collection size of each library by the number of books that are borrowed in total from that given library, and color coded by the library type (Mall, Stand-Alone or Regional Library). Overall it also has mean lines plotted for each axis to divide the graph in four different quadrants.

Each quadrant reflects a special characteristic of the libraries within that quadrant. The top right corner would represent the libraries that have a very large collection size and also have more than average number of books borrowed. Whereas the top left corner would be the set of libraries that have a rather smaller collection size and yet still have a high number of book borrowed. Bottom right would be libraries with high collection size and yet lower number of libraries borrowed and lastly bottom left would be those with smaller collection sizes and also lower number of books borrowed.

Such a visualization can be extremely insightful to have an overview of all the libraries and serves as a first step of understanding the characteristics of a library.



#### Floor Size vs No. of Books Borrowed

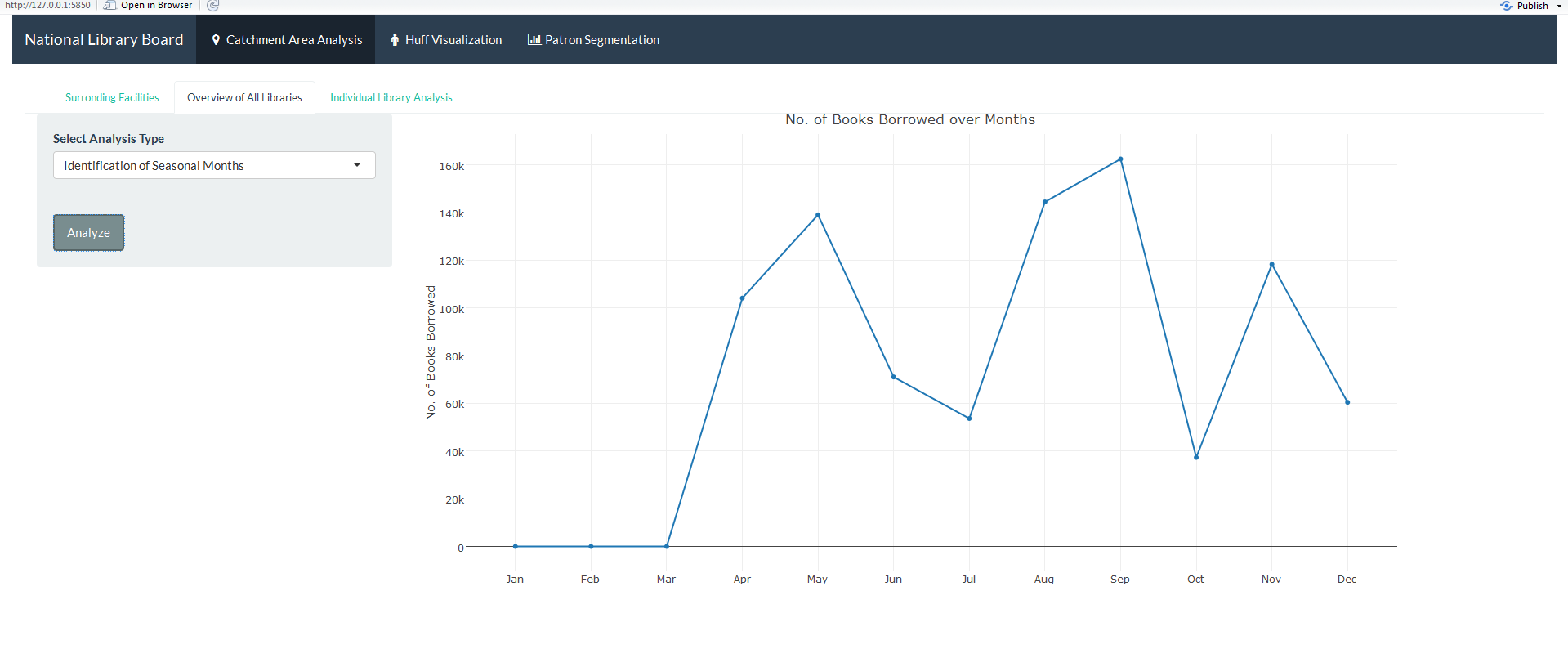
This visualization plots the floor size of each library by the number of books that are borrowed in total from that given library, and color coded by the library type (Mall, Stand-Alone or Regional Library) as well. Overall it also has mean lines plotted for each axis to divide the graph in four different quadrants.

This visualization is very similar to the collection size vs No. of Books borrowed, however this visualization shows if there is correlation between how large a library is versus the total number of books borrowed, and if there is then how well a library fares against all the other libraries. A screenshot of how the visualization looks like is shown below.



#### Identification of Seasonal Months

This visualization is intended to show if there exists a trend between the total number of books borrowed over time (months) in hopes to identify any seasonal months the library faces.

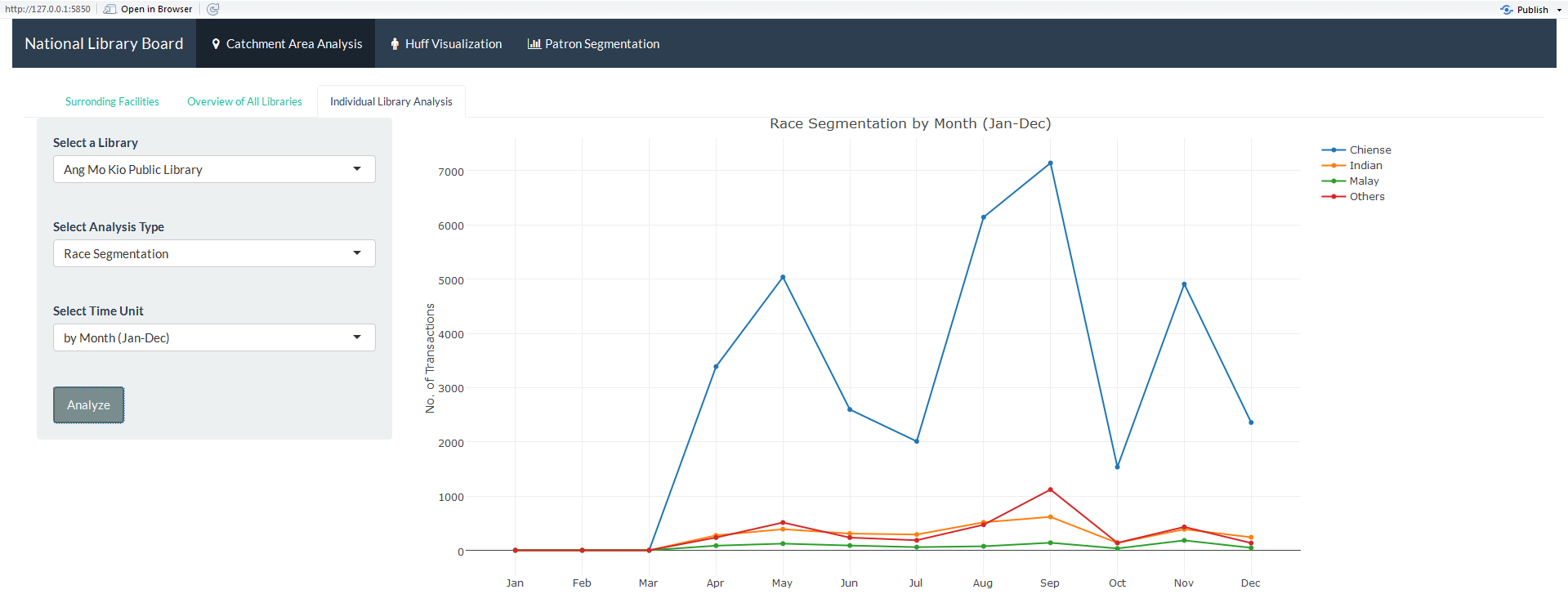


### “Individual Library Analysis” Tab

Along with the “Overview of all Libraries” tab, our team thought that having another tab for the “Individual Library Analysis” would be appropriate as well in case the users of the dashboard want to investigate further and determine different traits of each library and its demographics of patrons. The type of visualizations offered are, Race Segmentation, Age Distribution, Gender Distribution and Total No of Books Borrowed, all over different units of time (by Hour/weekday/month) for each library OR for all libraries. For the moment these tabs are completely dynamic which means that there is no hard-coding of any information for the construction of the graphs.

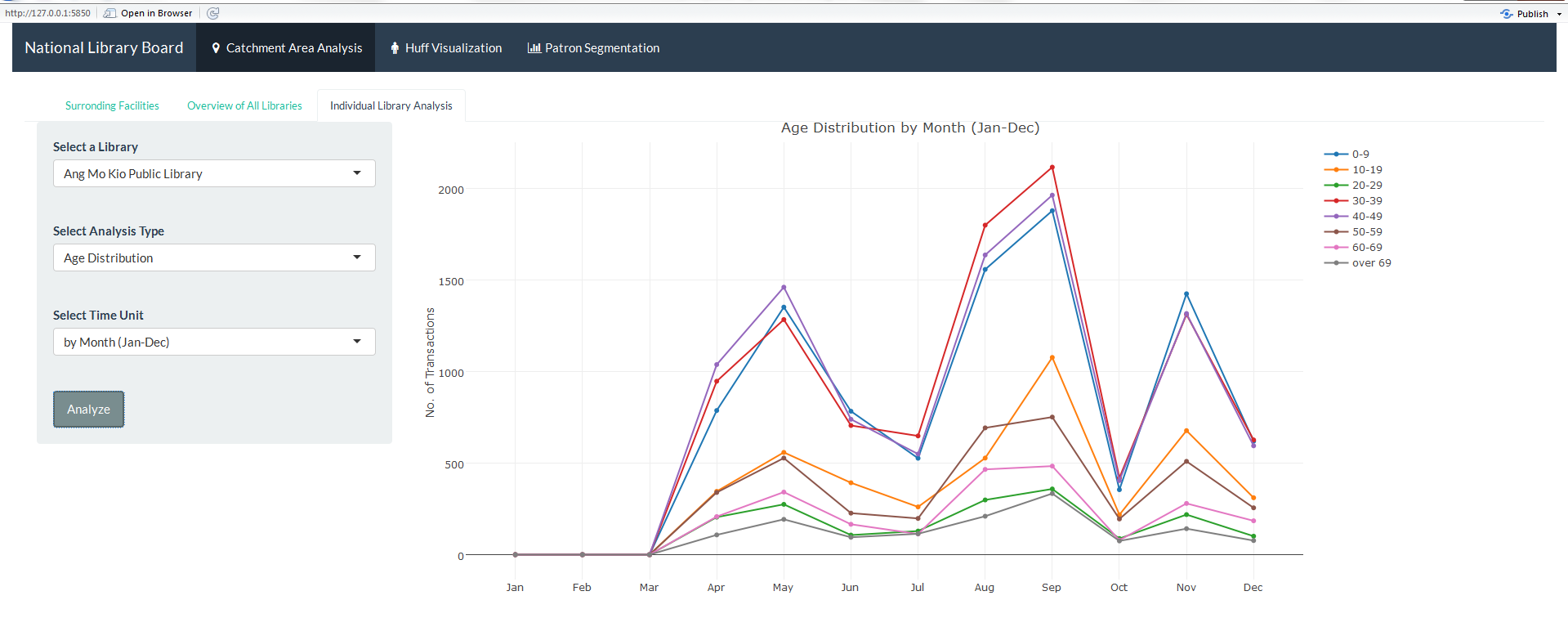
#### Race Segmentation

This is a simple line graph representation of the different absolute number of books borrowed by each race over time. This can be used to identify if a given race favors a certain period of time/month for any given library, or if a certain library experiences a higher number of patrons of a certain race. A snapshot of how visualization is shown below for Ang Mo Kio Public Library.



#### Age Distribution

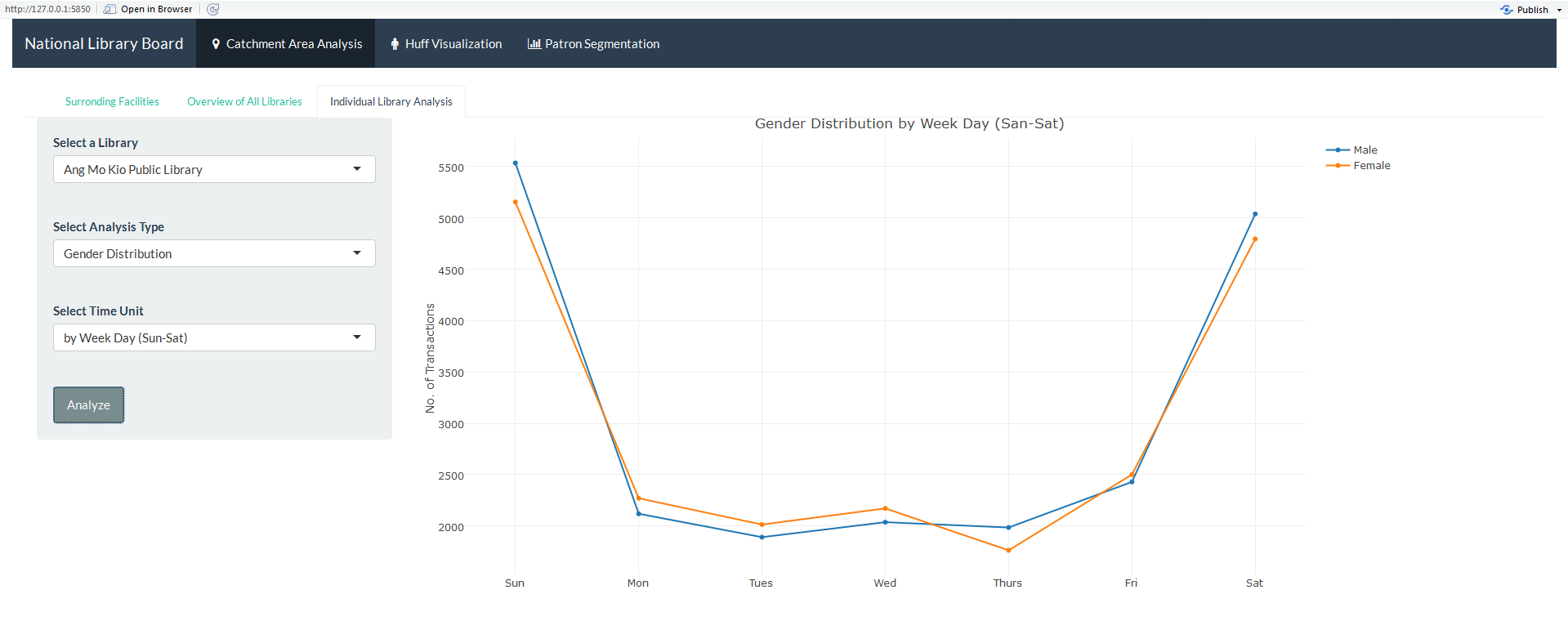
Similar to the Race Segmentation, the age distribution is intended to represent the distribution of the patron ages of a given library, this can be useful in identifying the prominent age groups that exist in a given library. The snapshot is shown below.



#### Gender Distribution

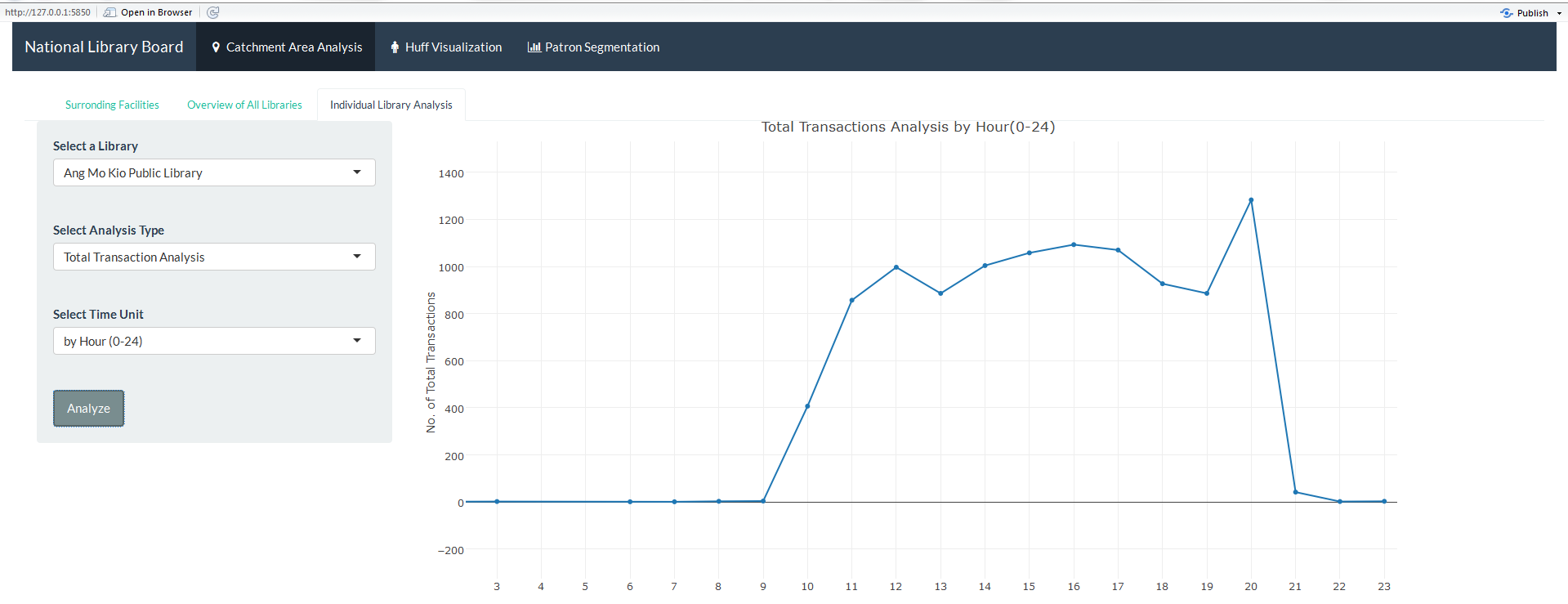
One other visualization that is present is the gender distribution across all libraries/for each library, as shown below.

The snapshot below uses weekdays as the time unit.



#### Total No. of Books Borrowed

Identifying if for a certain library a given time has more number of transactions is also an interesting visualization that our group wanted to display. This can in the future aid the library staff members in choosing to promote certain new books during a given time when they know they have a high influx of patrons. The distribution graph is shown below.



Our team understands that this is a very amateur method of visualizing the data however we intend to improve on it in the next iteration of our project with inclusion and comparison of both 2013 & 2014 data (at least three years’ data if available) for a more accurate seasonal trend visualization and analysis.

## Adaptation of the Multiplicative Competitive Interaction Model

Based on multiple sessions with our supervisor, our team has decided to use the MCI package in R to conduct the analysis of attractiveness using Huff’s model. The Multiplicative Competitive Interaction (MCI) Model (Nakanishi/Cooper 1974), is an “econometric model for analyzing market shares in a competitive environment where the market is divided in i submarkets (e.g. groups of customers, time periods or geographical regions) and served by j suppliers (e.g. firms, brands or locations)”. Resulting market share of the suppliers (libraries in our case), this model also analyzes the attraction/utility of the alternatives in the submarket.

Different from the senior group’s model, MCI model is nonlinear but can be transformed via Ordinary Least Square (OLS) regression using the multi-step log-centering transformation and our team will also re-arrange the raw data in an interaction matrix to fix into the model.

The purpose of this is to first use the MCI model along with the different new variables for attractiveness to assign weights to the variables. This would tell us which variables actually contribute towards the attractiveness of a library. After which we would then use the Huff’s model function in the MCI package to give us the probability (𝑷\_𝒊𝒋) that a Patron from a given Subzone ***i*** would visit a given Library ***j.*** One of the main motivators of shifting towards the MCI model is so that there is no longer a need to store the hard-coded Huff’s model calibrations within the dashboard, which is how the previous team had kept their results. With the implementation of the MCI model, the users of the dashboard can then recalibrate the model within the dashboard to produce new Huff’s model attractiveness based on the buffer stated by the users.

# Revision of Methodology

## Literature Review

Prior to the application of the Huff’s model as well as the MCI (multiplicative competitive interaction) model in our project, our group decided to explore the contexts with which those models have been applied elsewhere in the real world. This is so that we can understand the models better and also figure if we can learn anything which could be applied in the context of our project eventually.

Huff’s model which was initially discovered by David Huff in 1962 (Huff D.L, 1962) stated that the probability of any individual choosing a store is a ratio of the utility of that store to the sum of utilities of all other stores that the individual considers. In the context of the library, the utility/attractiveness of a store was determined by the size of the library and the transportation costs involved.

Huff’s basic model was then modified to attempt to include additional determinants of attractiveness for a conclusive picture of utility of a retail area. (Thomas, 1976) Reputation of the library would be one such example of a determinant of attractiveness.

The MCI model eventually got introduced. It extended the Huff’s Model and basically stated that attractiveness should capture the essence of competitive interactions. (Nakanishi & Cooper, 1974). In the context of the library, this meant that some of the factors included in the attractiveness index could be considered as substitutes for the library and can eventually cannibalise patrons from going to the public libraries. Examples of such factors would include bookstores and other entertainment facilities.

A further extension of the model showed how both external and internal factors of attractiveness should be considered (Jain and Mahajan, 1979). This was eventually used in a project in Germany which determined the agricultural mix of the farming sector i.e. whether farmers should go into cash crops, dairy production etc. (Neuenfelt, 2014). So other than external factors of attractiveness such as the socioeconomic factors, internal factors dealing with the farms’ capabilities were also considered (land, labour, capital etc.). In the context of our project, this would mean looking at the internal facilities of the library such as the availability of study spaces or a café.

All in all, based on our literature review, we are looking into including the following factors into our analysis at the moment. (They will be subject to changes depending on the progress of the project and the practicality of applying these factors):

1. Distance from Subzones/Planning Area to Library
2. Number of MRTs present within a Buffer
3. Number of Malls present within a Buffer
4. Number of Tuition Centres within a Buffer
5. Collection Size of the library
6. Gross Floor Area of the Library
7. Carpark accessibility
8. Branch Type (Mall, Stand-Alone, Regional)
9. MRT Centrality

## For Data Preparation & Model Calibration

### R Package - Multiplicative Competitive Interaction (MCI) Model

Packages are collections of R functions, data, and compiled code in a well-defined format. Market area models are used to analyze and predict store choices and market areas concerning retail and service locations. This package implements two market area models (Huff Model, Multiplicative Competitive Interaction Model) into R, while the emphases lie on 1.) fitting these models based on empirical data via OLS regression and nonlinear techniques and 2.) data preparation and processing (esp. interaction matrices and data preparation for the MCI Model).

### R Programming Language

R is a programming language and software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis. In our project, R programming language will be utilized to code back-end analytics models and front-end user interface.

### JMP Pro 13

JMP Pro 13 is a tool developed by the JMP division of SAS. As the data files are too large to be opened by conventional means such as Excel and Notepad, we will be using this tool to explore the data. Market Segment Analysis will also be done using the clustering function of this application.

## For Dashboard Visualization

### Tableau

Tableau is a visualization tool that will help us visualize and explore the data better so that we can use it to eventually develop a mock up for our dashboard.

### R Shiny & R Markdown

Shiny by RStudio is a web application framework for R and R Markdown from RStudio is a collection of interactive documents for R, both of which are planned to be used in our project for a more interactive and user-friendly interface.

### Javascript

JavaScript is a coding language for the web. We will be using JavaScript for most of the application’s user interfaces as it allows the implementation of various libraries to support user’s interactions and improve visualization.

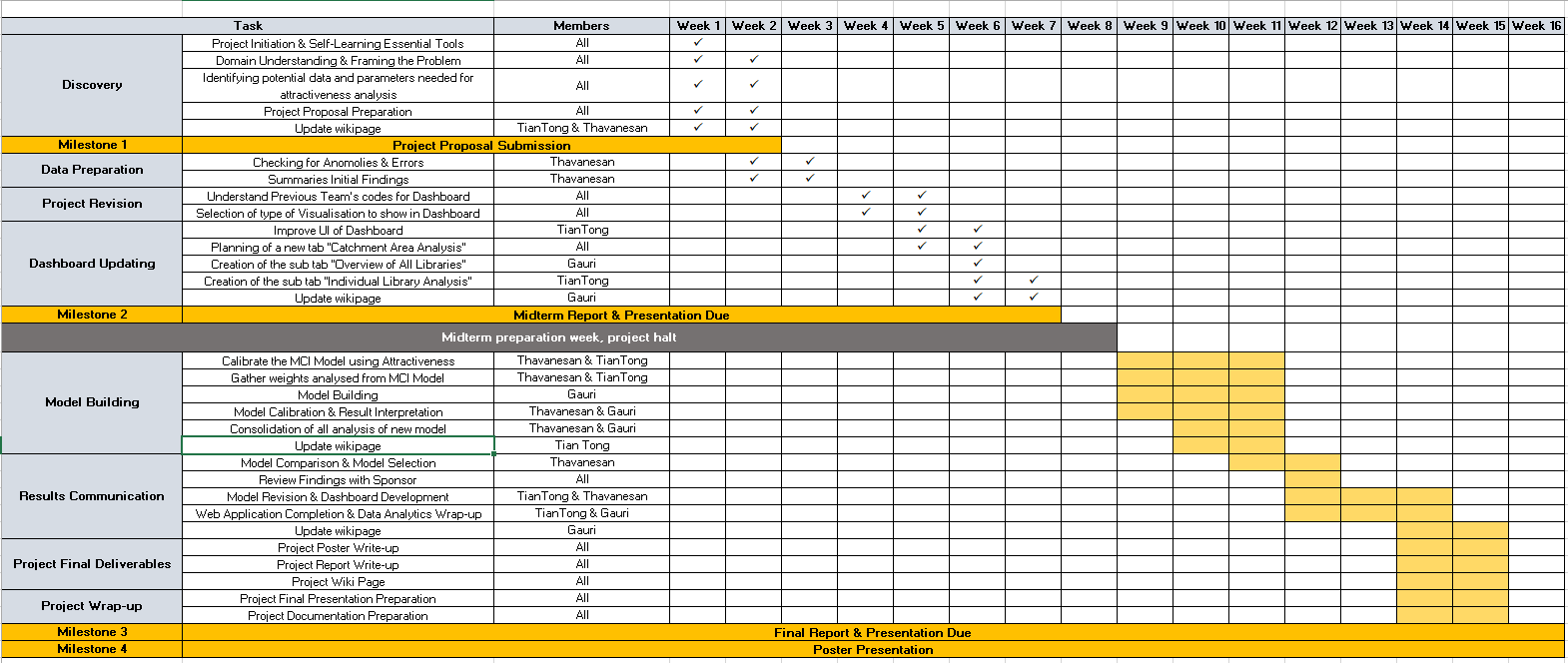
### Leaflet

Leaflet.js is an open source javascript library for interactive maps. This tool will be used to create a visualization page for the users where a map of Singapore, as well as point symbols representing various facilities will be displayed. The user can select the attribute to be considered for computing the attractiveness index by selecting or deselecting facility layers as well as varying buffer radius. This tool is selected as it provides a range of interactive maps and is easy to implement. It supports various plugins to extend its functionality.

# Revised Scope of Work

Our team intends to prioritize the model calibration first by ensuring that the “Huff’s Visualization” tab functions properly by using the MCI model. After which we would follow up with the improvement and completion of the “Catchment Area Analysis” tab with our desired mock-up of the dashboard which we showed above in the report.

# Revised Work Plan



# References

Brueckner, J. (2011). *Lectures on Urban Economics*. The MIT Press.

CRAN - Package MCI. (2017). Cran.r-project.org. Retrieved 6 January 2017, from <https://cran.r-project.org/web/packages/MCI/index.html>

Huff, D. L. (1962): “Determination of Intra-Urban Retail Trade Areas”. Los Angeles : University of California.

Huff, D. L. (1963): “A Probabilistic Analysis of Shopping Center Trade Areas”. In: Land Economics, 39, 1, p. 81-90.

Huff, D. L. (1964): “Defining and Estimating a Trading Area”. In: Journal of Marketing, 28, 4, p. 34-38.

Robert I., K. (2017). Quick-R: R Packages. Statmethods.net. Retrieved 6 January 2017, from <http://www.statmethods.net/interface/packages.html>

*R (programming language)*. (2017). *En.wikipedia.org*. Retrieved 6 January 2017, from <https://en.wikipedia.org/wiki/R_(programming_language)>

*Shiny*. (2017). *Shiny.rstudio.com*. Retrieved 6 January 2017, from <https://shiny.rstudio.com/>

Singapore Management University, (2016). *Geospatial Operational Insights for National Library Board (NLB) Project Proposal*. [online] Singapore Management University. Available at: https://wiki.smu.edu.sg/ANLY482/img\_auth.php/2/22/ANLY482\_NLB\_Proposal\_Qui\_Vivra\_Verra\_.pdf [Accessed 6 Jan. 2017].

Wieland, T. (2016, September 1). *Package ‘MCI’.* Retrieved January 6, 2017, from https://cran.r-project.org/web/packages/MCI/MCI.pdf