



ANLY482 - ANALYTICS PRACTICUM

PROJECT PROPOSAL

TEAM Enigma

Wei Xiaoxin

Wu Di

Zheng Xiye

Data Analytics for SMU Li Ka Shing Library

Overview

Situated at the heart of SMU campus, Li Ka Shing Library was officially opened on 24 January 2006. The library is named after Hong Kong businessman Dr. Li Ka Shing, chairman of Cheung Kong (Holdings) Limited and Hutchison Whampoa Limited. The Li Ka Shing Foundation donated an endowment to the library for collections.

With the vision to deliver exceptional services and build dynamic relationships within the SMU community and beyond, Li Ka Shing Library actively engages in gathering feedbacks from student community. One of the common feedback gathered from student community is that library should open earlier so as to fulfil the needs for students who prefer to study in the early morning. However, operational costs and the utilization rate associated with early opening hours is a big concern for library management team. Optimal opening hours is crucial in cutting back operational costs and yet, fulfilling students' academic needs across different time periods.

As such, our project will focus on analyse the traffic flow in between SMU buildings and library usage in the early mornings across different time periods so as to determine whether there is a great demand for earlier opening hours in SMU community.

Motivation

In recent years, digital disruption in every aspects of the commercial world has been prevailing. The ever-increasing trend of digital transformation can be attributed to the widespread adoption of business intelligence and data analytics approaches. Both of which are essential in facilitating organization's decision-making process and effectively bridging the gap between IT capabilities and business functions. In Singapore's context, Smart Nation initiative effectively drives adoption of data analytics among industry partners by harnessing the power of data technologies to create substantial business benefits.

In a sense, the rising industry needs drive employers to look out for individuals capable of drawing insights from numerous amount of data in seek of restructuring and optimizing ongoing business processes. As a group of graduating analytics students looking for hiring opportunities, our skills developed so far are mostly in line with market needs, which sparked off our desire to embark on a career in analytics. However, due to syllabus' constraint, skills developed in courses like, analytics foundation and data mining, are mostly in silos while in actual fact they constitutes an important component of the continuous Exploratory Data Analysis process. Meanwhile, most of our assignments and projects put too much emphasis on constructing the right model in delivering the right insight while neglecting essential steps like, data sourcing, pre-processing and iterative model refining.

Analytics Practicum project not only provides us with a hands-on experience on solving a real-life business problem with analytics methodologies. Most importantly, it draws linkage between skills developed through multiple courses so as to construct an integrated analytics solution. We hope to go beyond 'low-hanging fruits' like, simple model construction with numerous assumptions. Rather, we believe that the

'touch-and-feel' experience on analytics project lifecycle, will get us further down our career path in the near future.

Objectives

1. To identify and provide guidance for SMU Li Ka Shing library to make decision on re-scheduling open time in the morning
2. To identify the traffic volume trend from the other SMU buildings towards library in the morning changes on different days among the SMU academic year
3. To identify the traffic flow shifts from 24-hour study area to the main library areas in the morning around library opening time

Data

Through our initial discussion with our sponsor, three pieces of data may be of use in our scenario namely:

1. Library gantry access data - visitor data with demographics
2. Wi-Fi connection data - across different sections in all level of the library
3. Student traffic data - in between school buildings of different time periods.

Our sponsor has briefly walked us through the data-sets to give us an overview of its attributes and potential insights that can be drawn. He is still in progress of pre-processing the data to match student demographics with matric card data and block out unnecessary sensitive demographics information.

Nevertheless, from a high-level perspective, student traffic data may be of use in analysing traffic volume trend from other SMU buildings towards the library across different time periods. Meanwhile, since the number of Wi-Fi connections implies the number of students staying in different library sections, Wi-Fi connection data may be helpful in understanding traffic flow from library's 24-hour study area to main library areas around opening hours. On top of which, by combining the above-mentioned two pieces of data with library gantry access data, we may go further by analysing the demographics of students in both scenarios.

Methodology

1. Literature search

We gain understanding of the business domain through reading library website pages.

2. Trend analysis

Firstly, we will analyse traffic volume trend from other SMU buildings towards the library across different time periods.

Secondly, we will analyse the occupation of library in the early morning across different time periods. This is to understand whether there is a great variance for library occupation rate in the early morning across different time periods.

Scope of Work

| Task | Task Description | Job Allocation |
|--|---|-----------------------------|
| <i>Gather Requirements from Sponsor</i> | Learn about business background, operating model, and business requirements | WD(Wu Di), XY(Zheng Xiye) |
| <i>Research on SMU Li Ka Shing Library</i> | Conduct research to know more about the SMU Li Ka Shing Library | XX(Wei Xiaoxin) |
| <i>Research on Analytics Tools & Pick up Analytics Tools</i> | Explore different types of analytics tools and find out the suitable ones for our project | WD,XX |
| <i>Research on Data Models & Pick up Data Models</i> | Explore different types of data models and find out the suitable ones for our project | XY |
| <i>Brainstorm Ideas for Analysis</i> | Base on business requirement and collected data to brainstorm more analytic ideas | WD,XX,XY |
| <i>Proposal- Create Wiki Page</i> | Create and design project wiki page with related content | WD |
| <i>Proposal- Document meeting minutes</i> | Document sponsor meeting minutes and send follow-up email to sponsor after meeting, document internal meeting minutes | XX(Internal) XY(sponsor) |
| <i>Proposal- Proposal Report</i> | Combine information and write proposal report | WD,XX,XY |
| <i>Data Collection</i> | Collect data sets from sponsor, and ensure the data collected is correct | WD,XY |
| <i>Data Cleaning</i> | Detect and correct(or remove) dirty data (corrupt, incomplete or inaccurate data) for data set (1,2,3)* | WD(1),XX(2),XY(3) |
| <i>Data Transformation and Integration</i> | Transform data format and integrate related data for data set(1,2,3) | WD(1),XX(2),XY(3) |
| <i>Analyse and Design Data Presentation</i> | Explore insights(1,2)** and findings by analysing the relevant data; design proper data presentation (visualization) | WD(1),XX(2), XY(2) |
| <i>Present exploration</i> | Present exploration insights(1,2) to client | XY(presentation), |

| | | |
|--|--|--|
| <i>insights to client and gather feedback</i> | and gather client feedback about the insights and presentation(visualization) design | WD,XX(collect feedback) |
| <i>Further Exploratory Analysis base on client and supervisor feedback</i> | Further explore the analysis for insights(1,2) and data presentation based on the feedback from client and supervisor | WD(1),XX(2), XY(2) |
| <i>Analyse using chosen models</i> | Fit the data into chosen models and analyse | To be discussed based on chosen models |
| <i>Interim-Prepare Report</i> | Collaborate information and analyse results to write interim report | WD,XX,XY(finalize) |
| <i>Interim-Prepare presentation</i> | Collaborate analyse results to be ready for the interim presentation | WD,XX(finalize),XY |
| <i>Interim-Wiki page</i> | Ensure the wiki page is updated with current project progress | WD |
| <i>Meet with sponsor to get feedback and analyse results so far</i> | Meet with sponsor to get feedback and analyse results based on interim report and presentation | WD,XX,XY |
| <i>Refine model or perform analysis based on client feedback</i> | Refine model and analysis based on client feedback | To be discussed based on chosen models |
| <i>Integration of solutions & findings</i> | Collaborate and integrate current analyse results | To be discussed after Interim |
| <i>Application Building</i> | Build up final application based on client suggestion and team brainstorm | WD,XX,XY |
| <i>Finalize project results</i> | Collaborate all the analysis results and finalize | WD,XX,XY |
| <i>Update final wiki page</i> | Ensure the wiki page is updated to final stage with final outcomes | WD |
| <i>Update final report</i> | Integrate final analyse results and outcomes to write final report (need to collaborate the interim report feedback from client and supervisor) | WD,XX,XY(finalize) |
| <i>Prepare Poster</i> | Prepare poster for final presentation | WD |

| | | |
|-------------------------------------|--|----------|
| <i>Presentation to Client</i> | Present outcome, report and poster to client and gather feedback | WD,XX,XY |
| <i>Final Adjustment-Application</i> | Adjust final application based on presentation feedbacks | WD,XX,XY |
| <i>Final Adjustment-Report</i> | Adjust final report based on presentation feedbacks | WD,XX,XY |
| <i>Buffer</i> | For us to catch up or use due to any changes of our schedule. | |

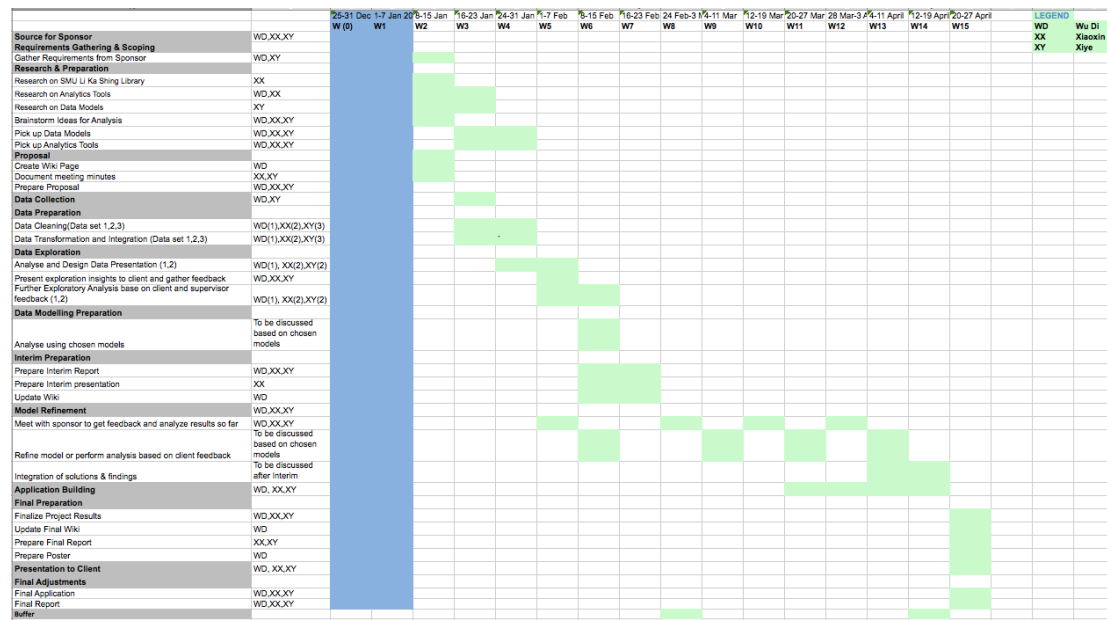
*data set(1,2,3) refers to

- 1 - Library visitor data - gantry
- 2 - Wi-Fi connection data - in different sections of the library
- 3 - Student traffic data - in between the buildings

** insights(1,2) refers to

- 1 - The traffic volume trend from the other SMU buildings towards library in the morning changes on different days among the SMU academic year.
- 2 - The traffic flow shifts from 24-hour study area to the main library areas in the morning around library opening time

Work Plan



* For weekly sponsor meeting, supervisor consultation, internal meetings:

| Weekly Meeting | Task | Job Allocation |
|-----------------|--------------------------------------|--|
| Sponsor meeting | Meeting minutes, Presentation slides | XY(minutes), WD,XX,XY(take turn to prepare slides) |

| | | |
|-------------------------|-----------------|----|
| Supervisor consultation | Meeting minutes | WD |
| Internal meetings | Meeting minutes | XX |

References

https://smux.smu.edu.sg/sites/default/files/smu/statistical_highlights_20160912.pdf

<https://library.smu.edu.sg/about-us-overview>

<https://library.smu.edu.sg/about-us/overview/about-us-li-ka-shing-library>

<https://library.smu.edu.sg/about-us-vision-and-mission>