Analytics Practicum Sponsor Meeting 01

MEETING CALLED BY	Prof
TYPE OF MEETING	Project Updates to Sponsor
FACILITATOR	-
NOTE TAKER	Hui Min, Bowei
TIMEKEEPER	Bowei
ATTENDEES	Chong Xin, Bowei, Hui Min

Agenda topics

1530 - 1545	PROJECT UPDATES TO SPONSOR (LIBRARY)	CHONG XIN

DISCUSSION	Demonstration - Go through initial visualizations - No. of Patrons may add up to more than 100% due to same person being able to visit multiple libraries - regional libraries serve a larger cluster, this is proven by the visualization - Central library has visitors from all over the country, and has a lot from Bedok - people at Bedok is rather localized - people at Bedok could be working at city areas, borrow books during work etc - other central libraries have the same pattern of having a large spread - Chinatown has a lot of people from Bukit Merah, perhaps due to the proximity of the two locations - Bedok residents have a difference of taste? High class libraries? - Patron distribution is due to Bedok being by far the most populous planning area. What about Tampines? - How about the patrons that visit multiple libraries? Double count Sponsor Feedback - Final design will be able to find out specific details of facilities. Which mall is 1 mall? Etc. - Should be able to show a listing of all malls, when you click the mall, it will zoom in to show mall location and details - the bubble size should be consistent across all visualizations - Could we drill down the map into the subzones?

ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
- As per mentioned above	All members	Sponsor Meeting 02

1545 - 1600	PROJECT UPDATES TO SPONSOR (SUBZONE)	CHONG XIN

DISCUSSION	Demonstration - People from Mandai has a direct bus to Yishun, explains the large amount of patron movement to Yishun - Queenstown is an inaccessible area Sponsor Feedback - Show a proportion of people from Queenstown over the entire population, to represent those that do not visit the libraries

ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
- As per mentioned above	All members	Sponsor Meeting 02

1600 - 1630	PROJECT UPDATES TO SPONSOR (RFM)	CHONG XIN

DISCUSSION	Demonstration - NLB Analysis team already has something similar to RFM done - Esplanade has AV materials, and a limit on what you can borrow - Plus, esplanade is out of the way - HHH could be children, but HLH could be interesting for Woodlands, because most people should go there due to tuition - HLH can be explained as it is a regional library, they can check out more materials, due to the variety etc.

	- Regional library can attract more people from further away, so they borrow more at once due to less frequent visits
	- Chinatown more dominated by HHL and HHH
	- Frequency can be caused by proximity to living distance from library
	- Boon Lay has LLL, but it has the Jurong libraries that are near, so that is strange.
	- Not going to acquire tableau, hence the app will be built for NLB using open source
	- An alternative could be using cliksense/clickview to build the dataset. But there are some challenges using cliksense. The background map will have problems, cliksense is not designed for it. Power BI could allow you to put your map in. Need to discuss with NLB in terms of deployment, options are Tableau, PowerBI, Open Source.
	Sponsor feedback
	- For RFM, show what each measure represents
	- What does low M mean? Lesser than 4?2? etc
	- What is the mode?
	- The focus should be on the low clusters, to reach out to them to visit more frequently
	- The high ones, we can engage them more to continue
	- The ones with high R and low others could be new members
	- Need for us to look at the products, to identify AV/Books etc.
	- Are there similarities across clusters if the frequency is low etc? can we find out what causes low frequency using this method?
	- LLL still means they borrow books, therefore we can look into why they are visiting so little
	- We can also look at those that don't borrow at all etc, that are currently not included in our visualization as of yet
	- There is interest in primary school students, that might cross over to 12 years old, so age groups of 5 may be less accurate
	- Might request customized aggregated age group
	- Financial year 2015 will probably still be included
	- Are there any further analysis done on the clusters?
	- Drill down into clusters
	- Penetration rate of patrons? Absolute values might not be very good
	- Perhaps another dashboard to focus on people not coming to the library?
	- Identify patrons, where are they coming from? Is the assumption of distance decay true for Singapore?
	- What other data do we put in to make sense of the data? What are correlated? Childcare? Kindergarten, Elderly care for huff's model. But these are demographic kinds of data
	- Can we get behavioral data? Other attendance or participation data we can use to find a relationship? No. Not in this exercise, but perhaps in the future.
	- The challenge of the project is that its very location specific, we want a predictive model to determine behavior/population patterns
	- Prof Kam: If we add a new library/ new infrastructure is built nearby, how will it affect the traffic? What if analysis of potential impacts

ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
- Shift application to alternate platform	All Members	Sponsor Meeting 02

1630 - 1700	ANOMALIES & COMMENTS	BOWEI

DISCUSSION	Sponsor comments
	- Which mall to go to if you want to relocate the library into the malls
	- Number of buses means connected?
	- 3 bus services to get to library does it mean is connected
	- Does shopping mall makes an impact? (it is proven)
	Anomalies
	- Prof will provide only households living in HDB and Condo to match the exact location or address of the patron
	- 32 books is a good indication/threshold to use
	- Anomaly 1 – NLB data integrity can be improved
	- Anomaly 2 – No comments.
	- Anomaly 3 – Branch code strives to be clean by sponsor
	- Anomaly 4 – DEAR not a patron, but are programs partnered with institutions. Sponsor suggests to remove all DEAR patron borrower category code from analysis. Birthyears '1990', '2015' and '2016' were set to the year that the institution was set up (to exclude)
	- Anomaly 5 – Exclude those branch codes not in the Collection Dataset.
	To-do
	- Send the pdf to prof then prof will send it to NLB
	- 4 weeks for now – the prototype plus model
	- Next meeting a month from now
	- First preview of model and prototype – complete everything by 26 November
	- Give a detailed timeline
	- Include our milestones in the next meeting

	- New batch of students will take over the project - NLB's main concern is to reach out to patrons who go to the library but does not borrow books		
ACTION ITEMS		**PERSON RESPONSIBLE**	**DEADLINE**
- As per mentioned above.		All members	Sponsor Meeting 02

OBSERVERS	NLB Team of 5
SPECIAL NOTES	Next Sponsor Meeting (02) will be tentatively scheduled next month. All members will present on their findings in the Team Meeting & Supervisor Meeting prior.