Analytics Practicum Supervisor Meeting 03

MEETING CALLED BY	Prof Kam
TYPE OF MEETING	Project Briefing
FACILITATOR	-
NOTE TAKER	Chong Xin
TIMEKEEPER	Chong Xin
ATTENDEES	Chong Xin, Bowei, Hui Min

Agenda topics

1600 - 1620	INITIAL VISUALISATION	ALL MEMBERS

DISCUSSION	Showed Prof the initial visualization on Tableau - Add in transparency of the circles and edge effects of the circles that are used to represent the proportion - The circle size should be based on absolute, and the colour of the circle should reflect the differences in percentages - In addition to show them the visualization, should also add in some observations which provides insights for the sponsor - E.g. The library@esplanade has a larger catchment area than the community ones - The Woodlands Regional Library also has a larger catchment area than the community ones - The sphere of influence of each library can come from a larger area if the library of a regional type Accuracy checking - Need to check through the no. of malls within 1KM to Jurong Regional Library. Currently it shows as 1 - Prof recommends adding in a "Year" filter to increase accuracy in reading the visualisations - Need to avoid the confusion – percentage colour should be more consistent Different Methods of Presenting the Data - Prof says need to show the sponsor the composition of patrons per subzone that visit each library - Another way to present it is to show the frequency: e.g. People who stay in Jurong West, how frequent do they visit the different libraries - In Tableau there is a "Parameter" function, and users will be able to choose between the "Absolute Unique Patrons" and "Frequency of Unique Patrons" - Line up all the variables (e.g. MRT stations, malls) - To go through the visualisations and show sponsor only the significant ones. Can summarise patterns which are the same (e.g. community libraries have very localized catchment area) - Include the no. of subzones that has significant amount of patronage (e.g. threshold = 5%) - Include the "patronage share" of each library e.g. Jurong Regional Library has 20% of all library patrons - To use Tableau to display the visualization, and also show the insights via Tableau instead of using PPT slides

ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
- Update the visualization by adding more specifics	All members	Sponsor Meeting 01

1620 - 1650	PROJECT PROPOSAL SLIDES	ALL MEMBERS

DISCUSSION	Assumptions - Bowei updates the Prof on the assumptions and anomalies - For Anomaly 1, we should exclude those MOLLY - For Anomalies 2 and 3, the percentage of anomalies is very small, hence it is okay to remove them from analysis - For Anomalies 4, we should clarify with the sponsor about "DEAR" category - Hui Min mentions that there are more branch codes than the 28 branch codes in the collection dataset. Prof mentions to include them in the slides as well RFM Analysis - We should do the interpretations more carefully. - Map the clusters back to the libraries, to show the cluster composition of each library - Prof showed the group k-means with Johnson Transform, which is similar to a log transformation - Prof showed the graph how to check the optimal transformation for a set of data (i.e. Data > Distribution > Continuous Fit > All)

	- Prof recommends checking through the different values of k, and not just rely on the optimal CCC generated by the software. If we use only k=3, some small clusters may be hidden and not shown in the interpretation

ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
-	All members	Sponsor Meeting 01

1650 - 1715   PROJECT PROPOSAL SUBMISSION   ALL MEMBERS

DISCUSSION	- Prof mentions not to include any initial findings and visualizations in the project proposal and wiki

ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
-	All members	2nd September (Fri)

OBSERVERS	-
SPECIAL NOTES	Sponsor Meeting 01 will be tentatively scheduled this week or next week. All members will present on their findings in the Team Meeting prior, and complete the PPT slides for the Sponsor Meeting.