

Analytics Practicum Team Meeting 01

MINUTES

AUGUST 19, 2016

1500 - 1700

SMU SIS BUILDING LEVEL 3

MEETING CALLED BY	Hui Min
TYPE OF MEETING	Project Proposal Preparation
FACILITATOR	-
NOTE TAKER	Chong Xin
TIMEKEEPER	Chong Xin
ATTENDEES	Chong Xin, Bowei, Hui Min

Agenda topics

1500 - 1545

APPLICATION OF THE HUFF'S MODEL

ALL MEMBERS

DISCUSSION	<ul style="list-style-type: none"> - Why is there a need to match the different subzones to the dataset? - We are using subzones as location parameter - Column headings <i>Locale Planning ADZID</i> – need to ask Prof about what does each value represent (about 1000 unique values) - How do we define the starting and ending points of the distance parameter of the Huff's model? - Centroid of subzone? Coordinates of library? - Distance plays a pivotal role in the Huff's Model, and any inaccuracies in measuring the distance between a subzone and a library may result in inaccuracies in the model/system 	
CONCLUSIONS	<ul style="list-style-type: none"> - To ask Prof about: (1) The column <i>Locale Planning ADZID</i> in the Patron dataset, and (2) How should we pinpoint the exact distance between a library and a subzone? E.g. coordinates of library, centroid of subzone, nearest point of subzone to the library? 	
ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
- To ask Prof the above-mentioned in Supervisor Meeting 02	All members	Supervisor Meeting 02

1545 - 1640

MARKET SEGMENTATION

ALL MEMBERS

DISCUSSION	<p>Define probability of a patron from subzone i visiting library j</p> <ul style="list-style-type: none"> - By dividing the (number of patron in each subzone i that visited a library j) by the (total number of patron in subzone i) <p>OR</p> <ul style="list-style-type: none"> - For each patron, find out the frequency of visits to each library, and integrate it into the probability of the patron visiting each library (decided to go ahead with this approach) - To cluster patrons for each library - To cluster patrons based on RFM model, and match the library to its dominant cluster (main bulk of patrons) - To find out each patron's RFM value before proceeding with clustering - To have market segmentation, either to use Sparks or SpatialLite - To refine project proposal <p>RFM in NLB context</p> <ul style="list-style-type: none"> - Recency – dropped due to data irrelevance - Frequency – number of visits in a year - Monetary – average number of books per transaction - Loyalty – number of unique libraries visited (unsure) 	
CONCLUSIONS	<ul style="list-style-type: none"> - To determine the RFM values for each patron & thereafter conduct market segmentation - To rethink whether there is a need for the "Loyalty" variable in the RFM model 	
ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
<ul style="list-style-type: none"> - All members to conduct the market segmentation to ensure correctness - Hui Min and Bowei to research on Apache Sparks and SpatialLite to check plausibility in using them in our project 	All members	Team Meeting 02

1640 - 1700

MEETING SCHEDULING

ALL MEMBERS

DISCUSSION	<ul style="list-style-type: none"> - To have the supervisor meeting in-between our regular team meetings, so to facilitate discussion before and after each supervisor meetings - Supervisor Meeting 02 to be scheduled on 24 August 2016 (Wed), to have Team Meeting 02 about an hour before and after Supervisor Meeting 02. 	
ACTION ITEMS	PERSON RESPONSIBLE	DEADLINE
-	-	-

OBSERVERS	-
SPECIAL NOTES	<p>Next Supervisor Meeting (02) will be tentatively scheduled on 24 August 2016 (Wed), subjected to Prof's availability. All members will present on their findings in the Team Meeting prior.</p>