

## Supervisor Meeting 4

Date & Time: 01 Feb 10:30am - 11:30am

Venue: SIS MR 4.1

Attendees: Prof. Kam Tin Seong (Supervisor), Wang Sijia, Ren Mengxi, Wang Tianjing

Absentees: Null

### Agenda:

1. Update supervisor on data exploratory progress and gather feedback
2. Consult on bin width difference problem

### Details:

1. Inform supervisor that we further bin the time based on the rule below:

Before 9AM	Early morning
9-11	morning
12-17	afternoon
18-23	night

However, as the time period is not evenly divided, the total number of visits for each basket may be bias. Same problem applies for binning the month into terms as well. Prof Kam suggested us to divide number of units for each basket and to use average for further analysis.

2. Try to explore frequency of visit (per day) for both school level and user level
3. Any graph about ranking should be sorted to give a better view (visit rate by school)
4. Apply the same rule for undergraduates and postgraduates and compare whether their usage is of the same pattern
5. Time period should not just apply at unique user level because it only catches earliest and latest entry time (great data loss). Apply same rule at per visit level.
6. Use mosaic plot to show school vs. time period
7. Always be careful to compare total number as there may be different base which creates bias. Use percentage.
8. Some ideas to analyse unique user data: Can get an idea about the particular person who went to the library most frequent. Monday Sunday? Who are the student who visit the library everyday. Are they morning persons or afternoon persons. How many times they visit by day of week. Every monday are they going to be there? (count visits by day of week -> 7 cols) what time do they prefer to visit? Are the pattern same for per visit level and per student level?

### Action Plan:

Item	Person in charge	Deadline
Update sponsor about our progress	Sijia	Feb.3
Refine and finish analysis	Mengxi, Tianjing	Feb.5
Start to look into R	All	Feb.8