

Internal Meeting 6

Date & Time: 04 Mar 1:00pm - 2:00pm

Venue: Student Study Lounge B1-4

Attendees: Ren Mengxi, Wang Sijia, Wang Tianjing

Absentees: Null

Agenda:

1. Update progress of statistical reconfirmation
2. Explore two new data columns (dean's list or not and Singaporean or not)
3. Update R shiny progress and split coding work

Details:

1. We joined our table with the two new columns and found some problems:
 - a. There are 1650 students without a dean list record. They are considered as non dean's listers for now.
 - b. There is one student with two nationality records (one 0 one 1). We guess the reason may be that he/she changed nationality during the undergraduate four years. We will treat him/her as Singaporean.

The above issues have been addressed to our sponsor.

2. We have added in Contingency analysis and one way analysis for our user level analysis.
 - a. We choose contingency analysis to compare the number of undergraduate visitors from different schools and batches. As the probability $>$ Chi Square is too low (lower than 50%), our null hypothesis could not be rejected, meaning that there is no distinct difference among different schools and batches.
 - b. We choose one way analysis to compare the difference in number of days among different year of study. Based on the p-value, we can conclude that freshman and sophomore students share the similar pattern, and junior and senior students share the similar pattern. The difference between these two groups (freshman and sophomore, junior and senior) is highly significant (p-Value $<$ 0.001).
 - c. We choose one way analysis to compare the difference in number of days among different schools. Based on p-value = 0.9001, Students from SOB and SOA have similar patterns of behaviors in terms of the number of the days they go to the library and the median of the number of days students study in library is 17 days in these six months; As the probability $>$ F is lower than 0.0001, we can reject our null hypothesis and conclude there is different patterns among different schools.

- d. We choose one way analysis to compare the difference in number of days among different user groups. As Prob > F is lower than 0.0001 and p-Value < 0.0001, we can conclude that different user groups have significant difference regarding to number of days they come to library.

Action Plan:

Item	Person in charge	Deadline
Explore data frame	Mengxi	Mar 8th
Explore how to do data preparation using Shiny	Tianjing	Mar 8th
Add in statistical insights for data exploration	Sijia	Mar 8th