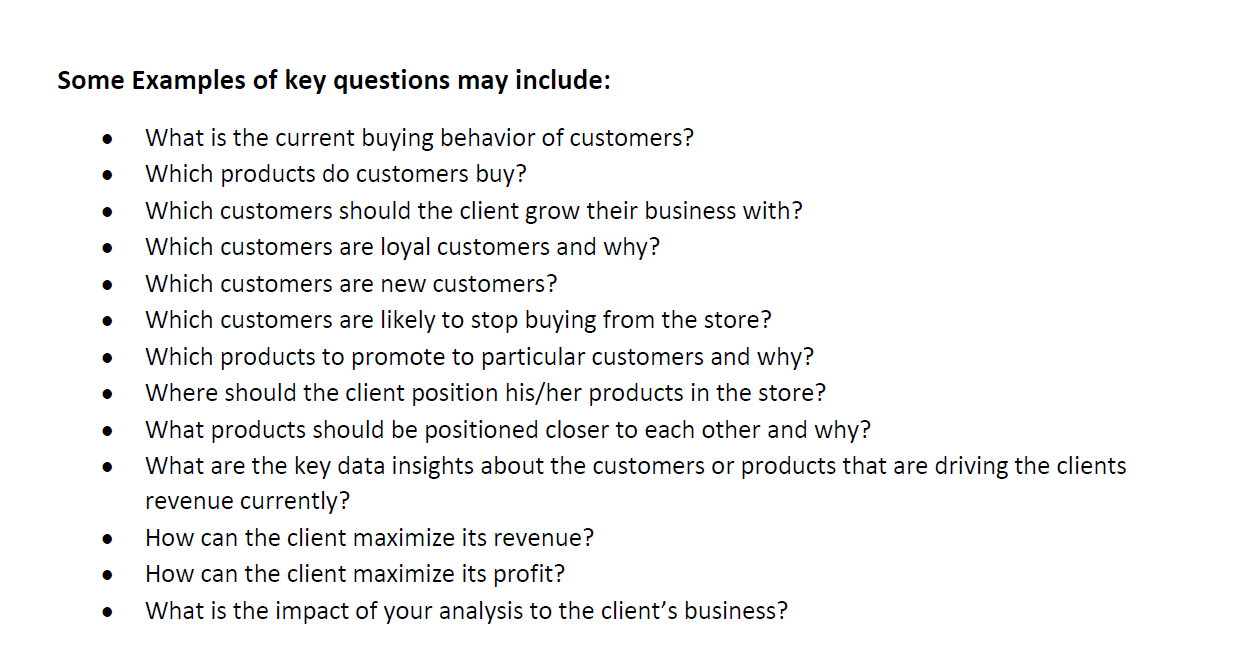
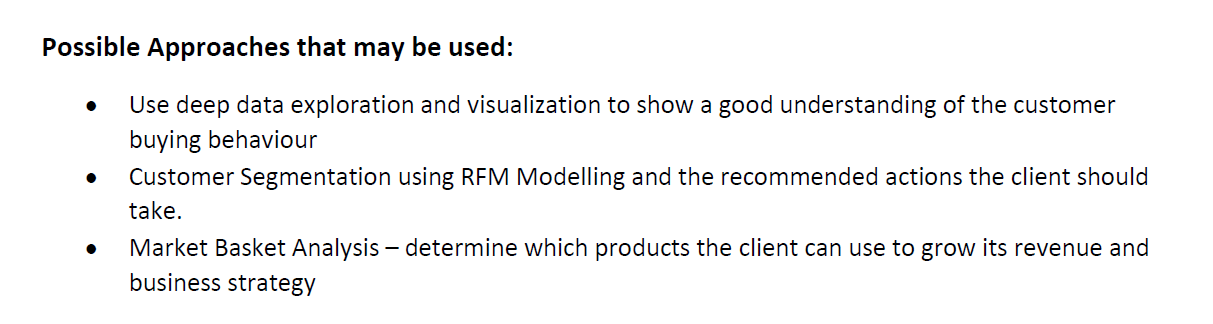
Meeting 23/8/2020 1500 -

Present: Shao Min, Mekala, Michael, Zhihao, Yifan





1. Which products do customers buy?
2. Customers profiling: loyal, new, leaving (Shao Min, by monday noon)
3. Time Series (shaomin)
4. Positioning closer to each other
5. Products to promote to particular customers (Mekala, by tues noon/night)
6. Pricing across clusters vs average pricing in Categories table (Yifan)
7. Maximisation its revenue (conclusion)

Shao Min : Data is skewed, the RFM may not be effective. **Find a better segmentation**. Define a new criteria for recency. Look at the 3 recent dates, if they are far apart, penalise the recency score.

Zhihao: Find intervals between purchases instead. Basket rules, a priori. Shopping rules per category. What products should we advertise?

Mekala: Hopkin’s Test. We can do k-means on all R,F,M → however this will give similar if not same results as RFM scoring if we follow the same weights so this can be alternative to RFM scoring

A specific basket of items for certain groups of people. Find the average basket price.

Spend/Volume does not equal to a constant, does not equal to the value of the category in the category csv, indicating a presence of different branded product in the data. If 0 volume, then spend/pack.

Assumptions: We have two hypothesis: either it's different brands or certain days may have promotions. After looking at the dataset, we assume it's different brands.

Investigate the expenditure on certain days? Maybe have some special day promotions?

Spend: may be 0

Coding: ShaoMin; YiFan, Mekala

Presentation: Zhihao, Michael

Timeline:

* Coding Team meet on Tuesday night,
* Thursday/Friday meet up with visualisation team
* Final meetup on sunday, wrap up with scripts or etc