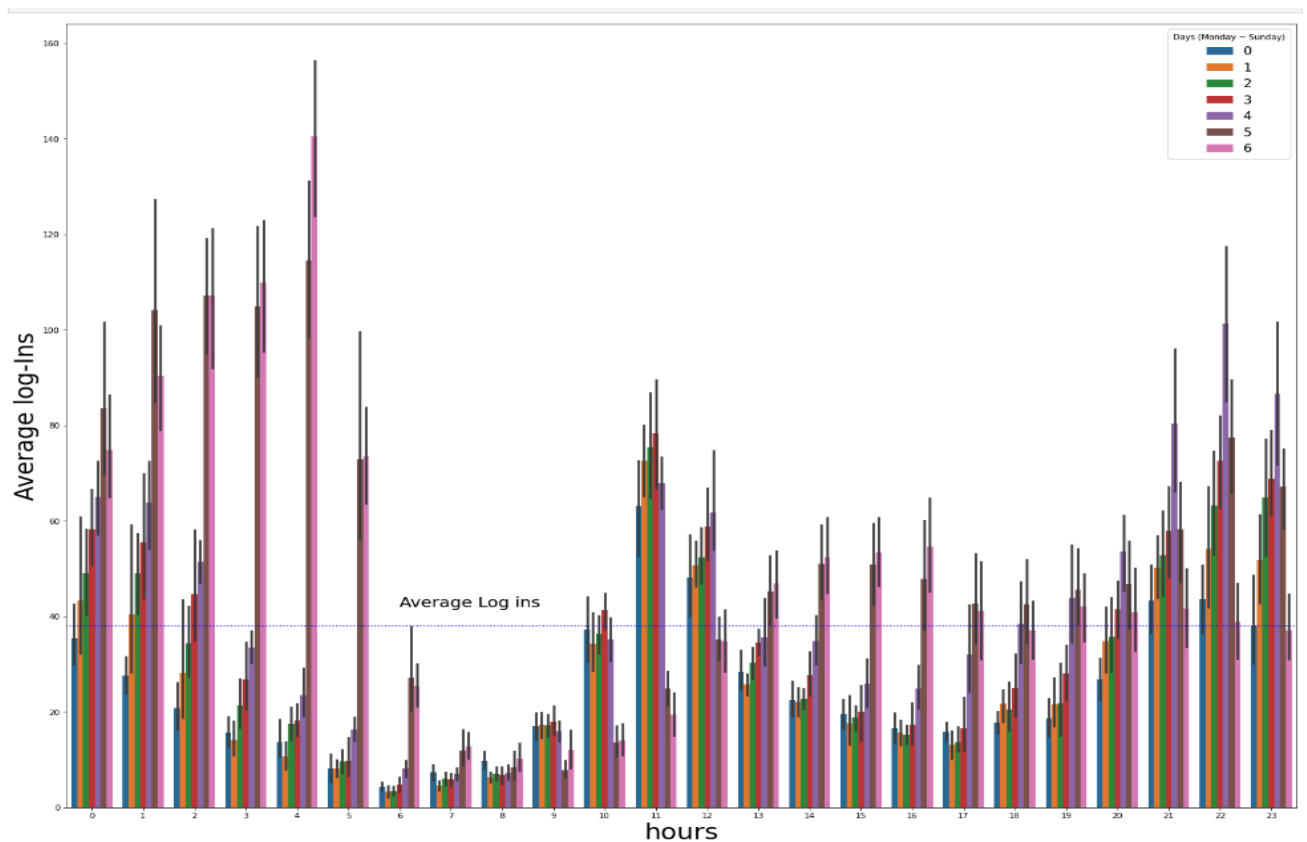
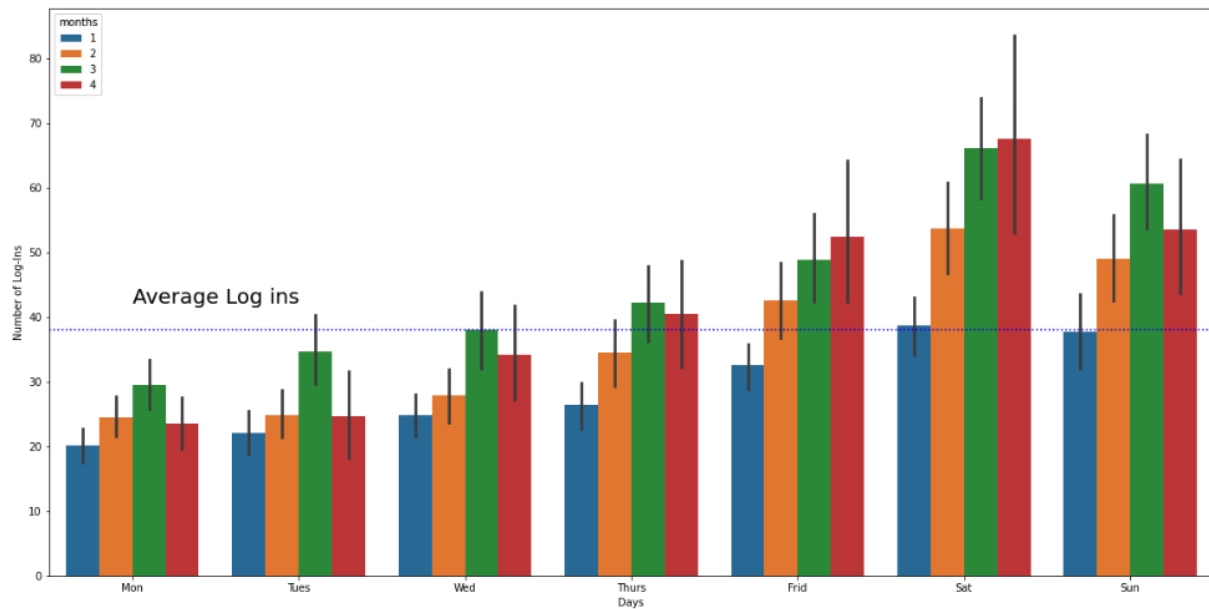


By Taehyun Yoon

Part 1:

There were 92265 clicks from 105 days in the dataset. Daily logins on weekends were significantly higher compared to the logins on weekdays.



In addition, around 3~6 am had the most click times. Compared to the click times in the afternoon.

Part 2: Experiment and Metric Design

• Days for rides • Weekday ride histories and Weekend ride histories before the toll bridge • Purchase for every ride before the toll • Weekday ride histories and Weekend ride histories after the toll bridge • Purchase for every ride after the toll

2.

Before we conduct the ANOVA test, we will choose two groups.

H0 - Null Hypothesis

H1 -

- Gotham
 - H0 - No difference in average purchase(rides) before and after a month from reimbursement during night time.
 - H1 - Increase in average purchase(rides) before and after a month from reimbursement during night time.
- Metropolis
 - H0 No difference in average purchase(rides) before and after a month from reimbursement during day time.
 - H1 - Increase in average purchase(rides) before and after a month from reimbursement during day time.
- One tailed ANOVA test
 - • To test if the average cost before the toll bridge and the average cost after the toll bridge to see if they are any statistical difference. • It gives you a P value.
 - 1. Null Hypothesis: There will be no difference in mean between groups before and after the toll bridge formation.
 - 2. Alternative Hypothesis: There will be increase in average price.
 - 2. Set the level of Significance(Criterion): $p < 0.05$ • If p values less than 0.05, we reject the null hypothesis.
 - 3. Compute test Statistics: • We create sample.
 - 4. Get results

For the Z test, we will need the following two groups. • Z test one tail • If we know the population standard deviation, then we can proceed with this test.

- Gotham

Ho - No difference in mean cars during night a month before reimbursement and a month after reimbursement.

H1 - Increase mean in cars during night a month before reimbursement and a month after reimbursement.

- Metropolis

Ho - No difference in mean cars during daytime a month before reimbursement and a month after reimbursement.

H1 - Increase in average cars during night a month before reimbursement and a month after reimbursement.

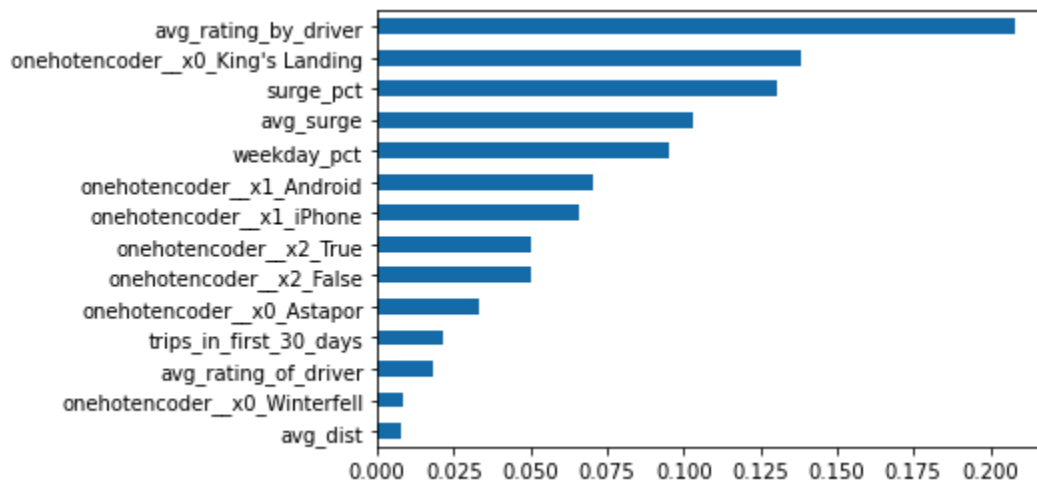
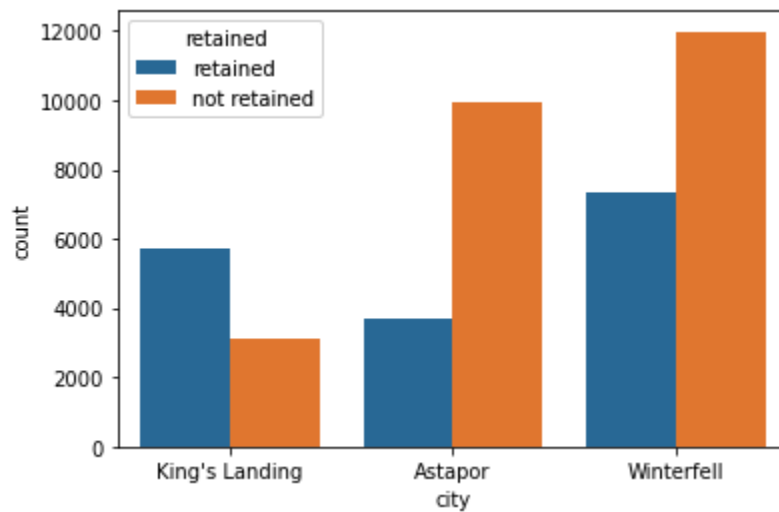
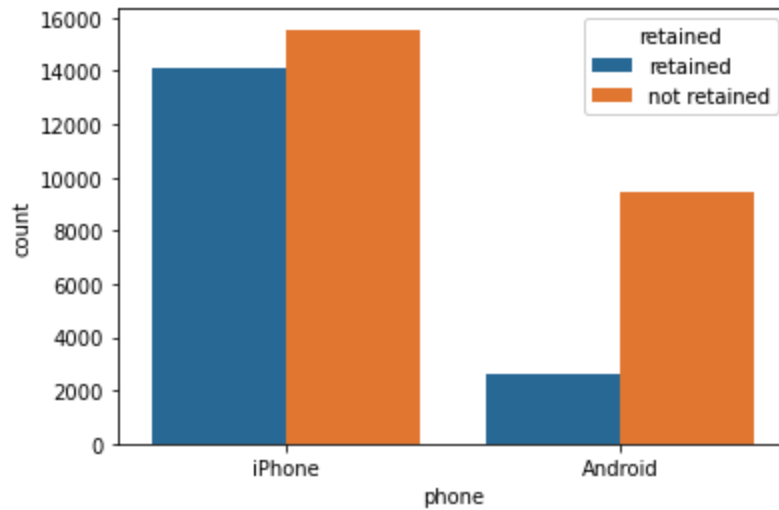
I will choose the alpha to be 0.05. If the p value is lower than alpha, then we will reject the null hypothesis and tell the manager to continue the reimbursement.

I will choose the alpha to be 0.05. If the p value is higher than alpha, we fail to reject the null hypothesis. Then we would have to decrease the toll fee or figure something else to increase the partner driving between the two cities.

If the p value is mixed, only one hypothesis is >0.05 , then the result is inconclusive.

3. Predictive Modeling

- User retention data were inspected and out of 41733 users and 25016 were not retained and the remaining users are retained which in percentage it would be 59% being not retained and 41% being retained. Through Exploratory Data Analysis, Winterfell had more people who were not retained and King's Landing had the most users with retained users.



- I built a Random Forest Classifier which has a 76% accuracy.
- Based on the feature importance chart,

- The rate by the driver had the most influence over the classification.
 - King's Landing and the surge_pct was also a important factors deciding the users retaining rate.
- Therefore, based on the results, I have 3 suggestions for the business team.
 - 1. Pick the better + rating drivers as much as possible. Also, create some sort of exam to become one of the drivers for the app for the best to customers.
 - 2. Investigate Ling's Landing. There must be something special about that area which attracts user to retain.
 - 3. Surge Pct - Possibly lower the Surge_pct. Adjust the price or time of this pct.

In addition, a further investigation about Iphone and Andriod should be done.
There might be a reason why iphone users has more retained users.