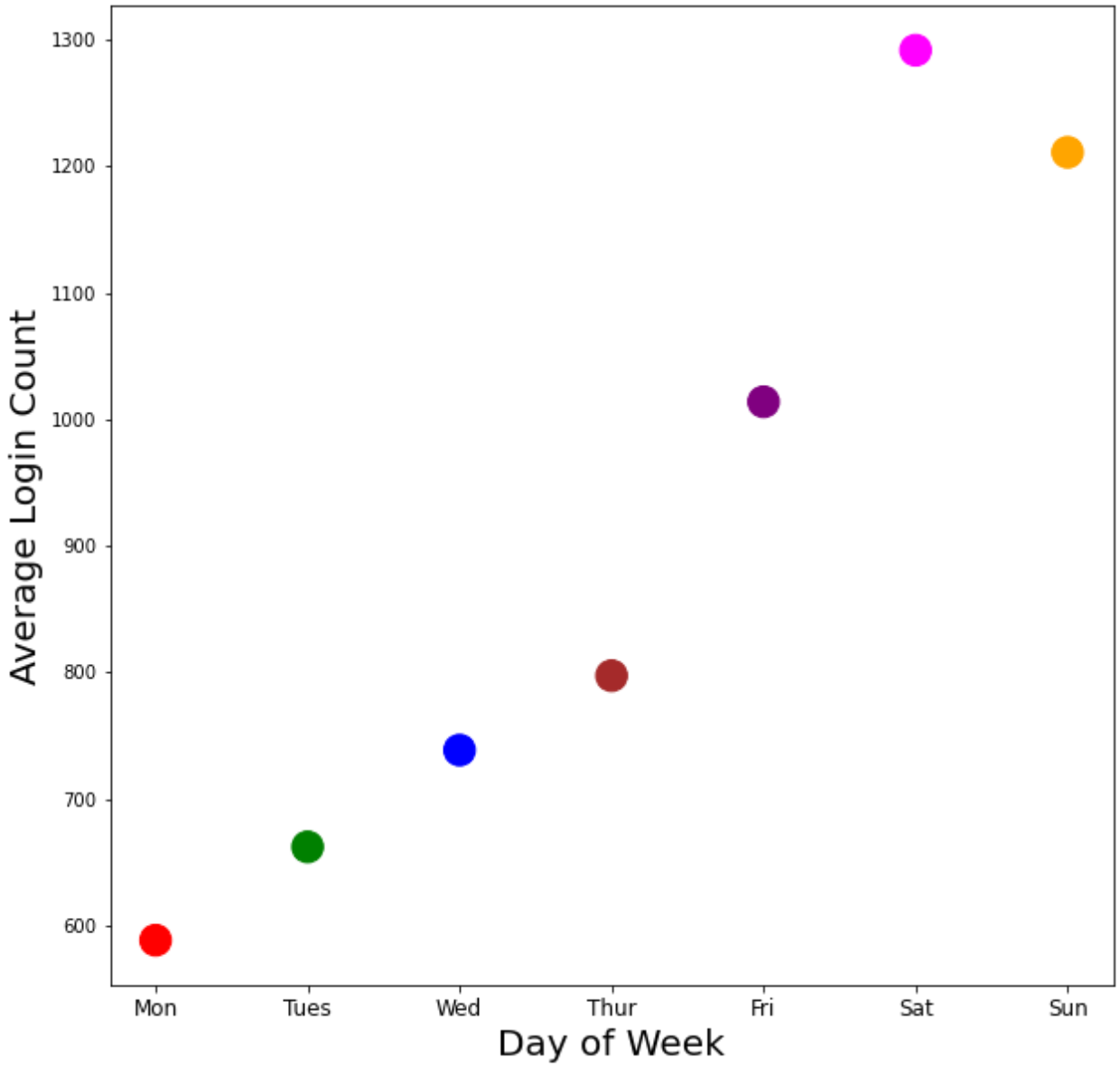
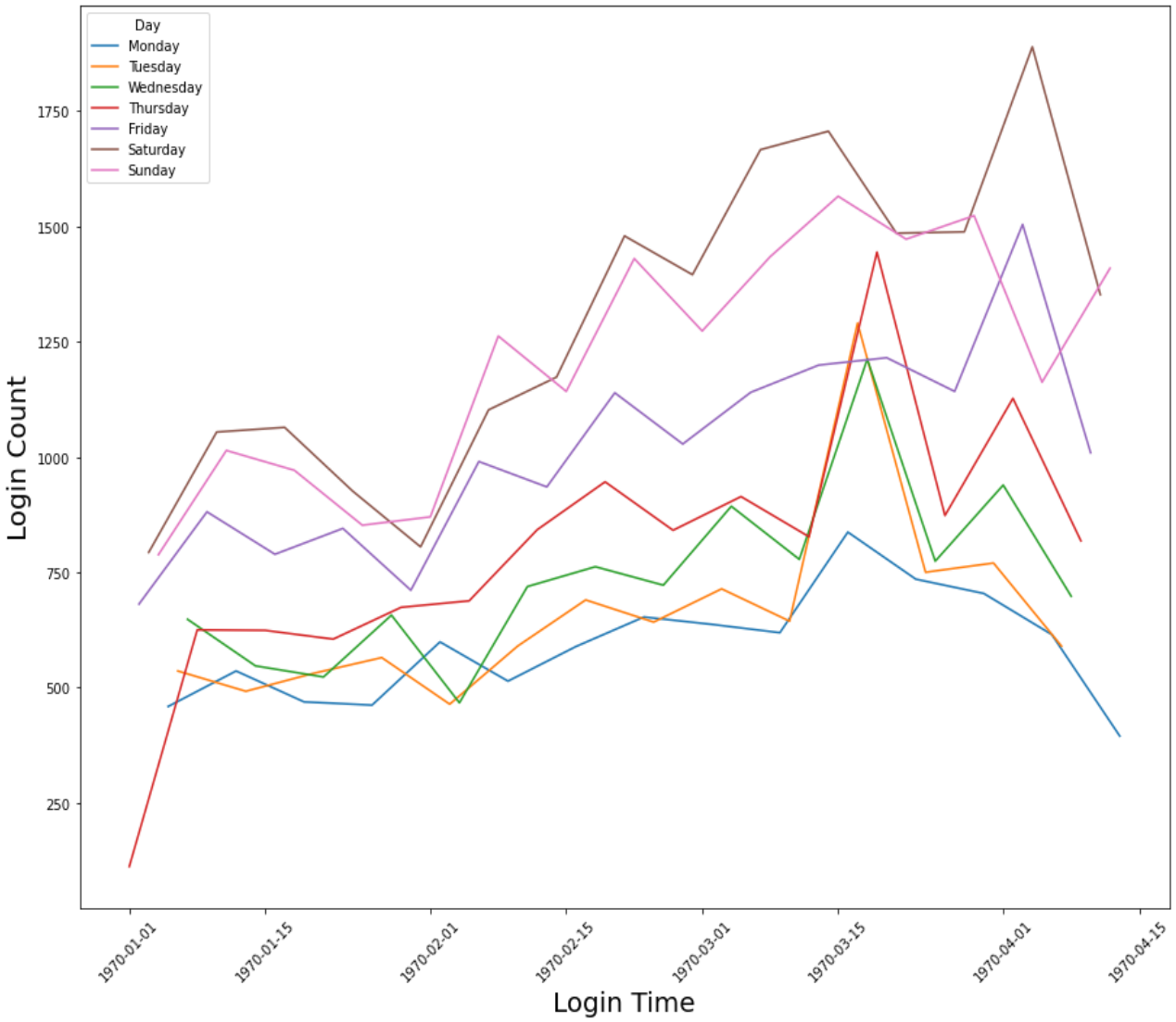
Ultimate Technologies Inc. Challenge

Part 1 – Exploratory Data Analysis

There were 93,142 timestamps of user logins over the 103 days in the dataset. Through explanatory analysis, I found that Saturday had the highest average daily logins, Sunday the second, and Friday the third.





Part 2 – Experiment and metrics design

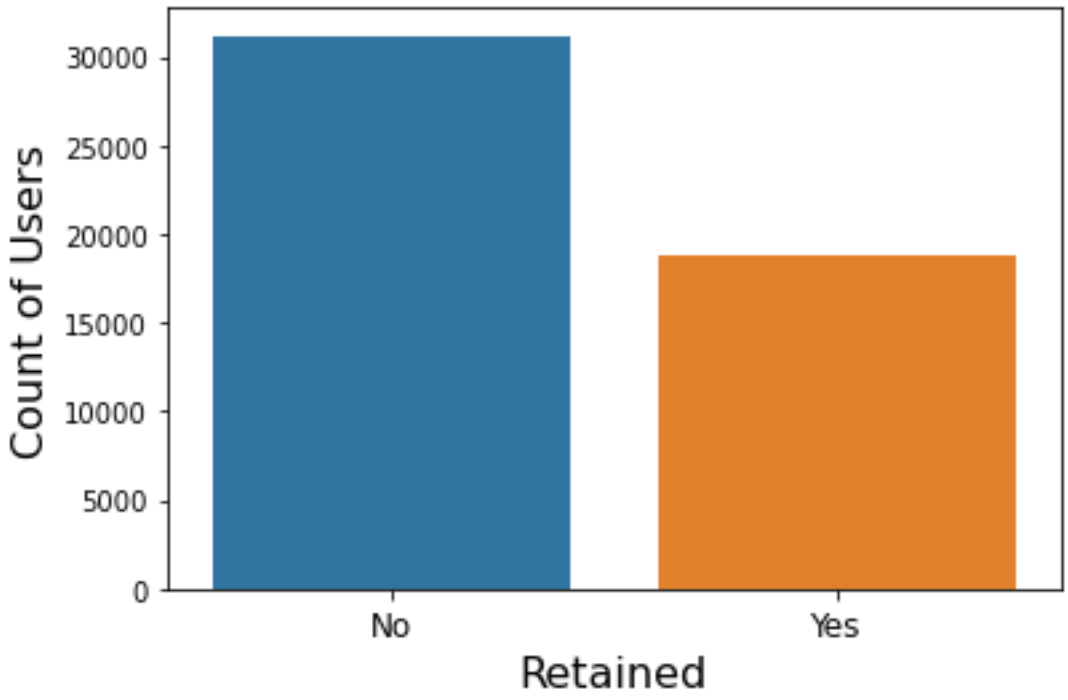
1. For the goal of increasing inter-city rides between Gotham and Metropolis, I would measure the total number of times drivers crossed the toll bridge in the testing period.
2. Randomly select 200 drivers and divide them into two groups of 100 each. Offer the experimental group reimbursement for all toll costs between Gotham and Metropolis. Meanwhile, the control group receives no toll cost reimbursement offer. Run the experiment for 4 weeks, collecting data on

the number of trips each driver makes across the toll bridge during that time. Conduct A/B testing between these two randomly selected groups. Follow through with a z-test to verify the significance of the observation. Based on the results, I could find the difference in number of inter-city trips between the control group (drivers with no toll reimbursement offer) and the experimental group (drivers offered toll reimbursement). I could provide this information to the Ultimate managers with analysis by day and by week. I could also consider whether any significant events during the testing period might have impacted the test results.

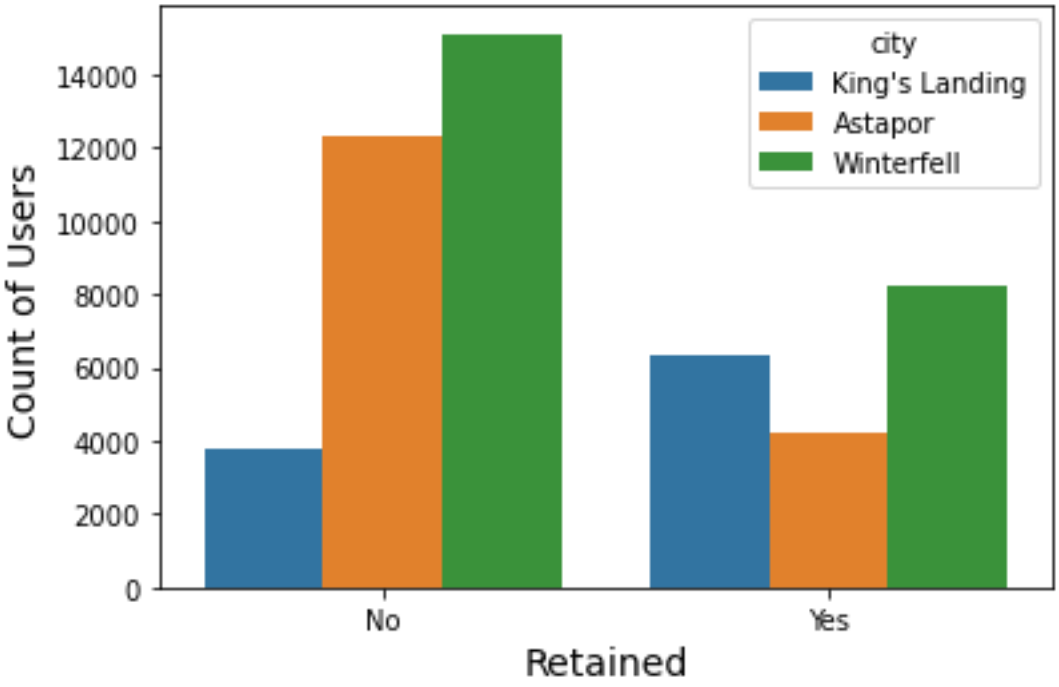
Part 3 – Predictive Modeling

After imputing the null values in avg_rating_of_driver and avg_rating_by_driver with their means respectively, I kept all the observed data.

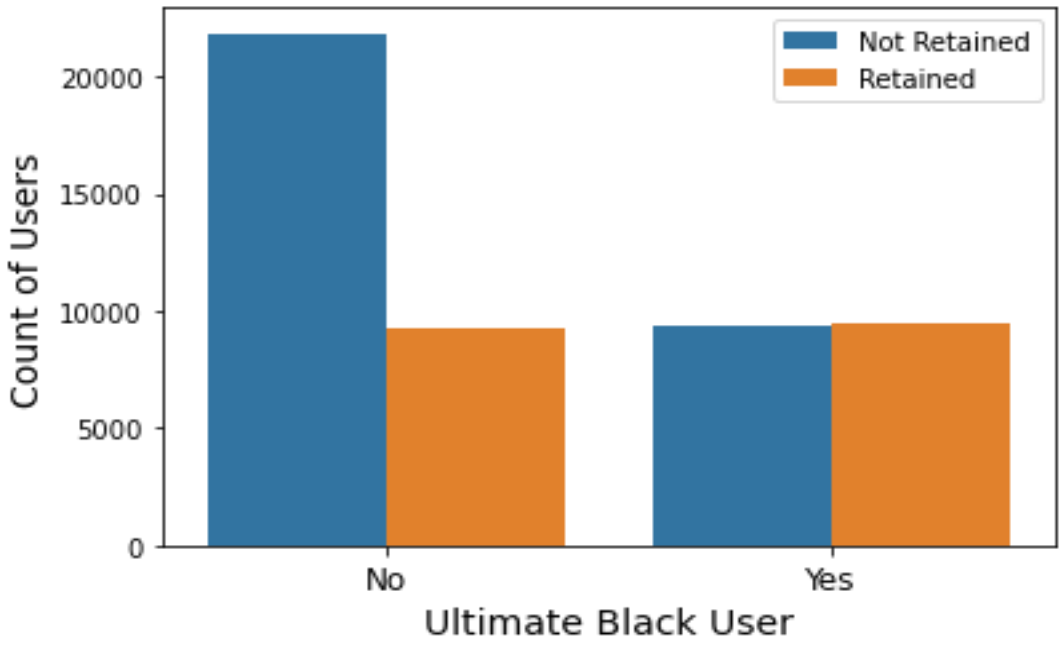
After analyzing the data provided, I found that 18,804 out of 50,000 users were retained in their 6th month on the system.



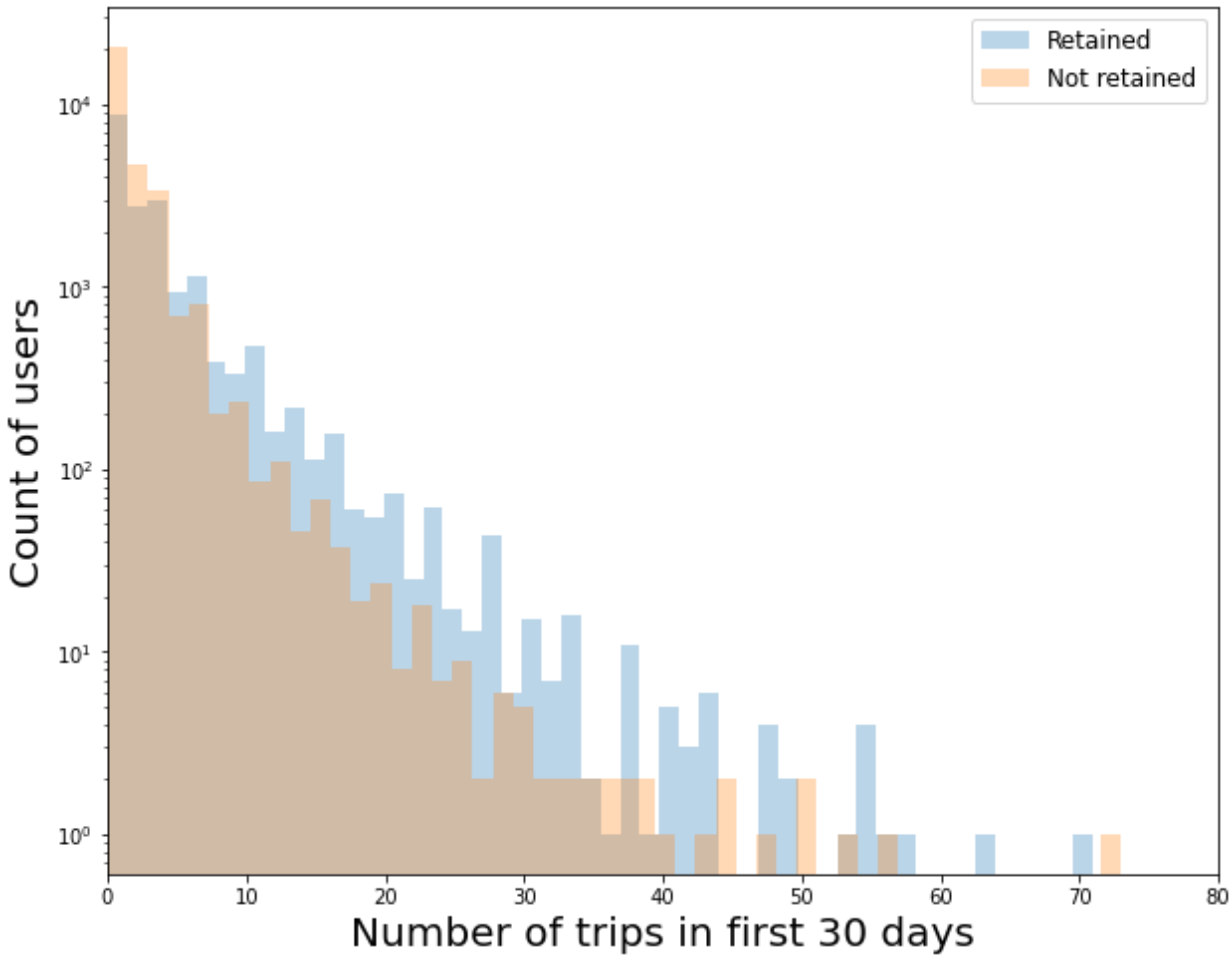
Through explanatory data analysis, I found that users who signed up in King's Landing have a 62.81% retention rate, which is much higher than 35.19% and 25.57% in Winterfell and Astapor respectively.



Ultimate Black users have a 50.37% retention rate, which is much higher than the Non-Ultimate Black users' retention rate of 29.88%



Also, the more trips that a user took in the first 30 days, the more likely it is that the user will be retained.



I built a logistic regression model that achieved about 70% accuracy and a Random Forest Classifier that achieved above 71% accuracy in predicting whether or not a user will be retained in their 6th month on system. The three features mentioned above are the most important predictors for retention.

Suggestions for the business management:

1. Promote the use of Ultimate Black.
2. Find out why retention rate is higher in King's Landing, and how other cities can learn from it.
3. Promote the use of Ultimate in a user's first 30 days.