RETURNING USERS: OHBIBI CASE STUDY

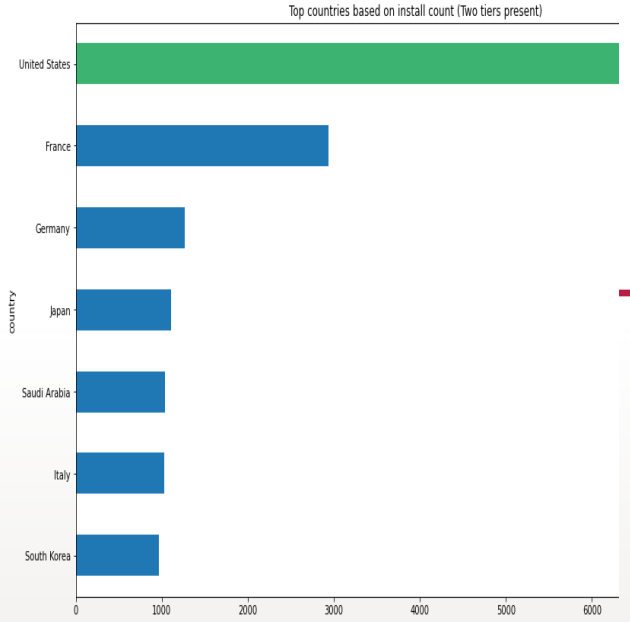
- GOKUL KRISHNAN R

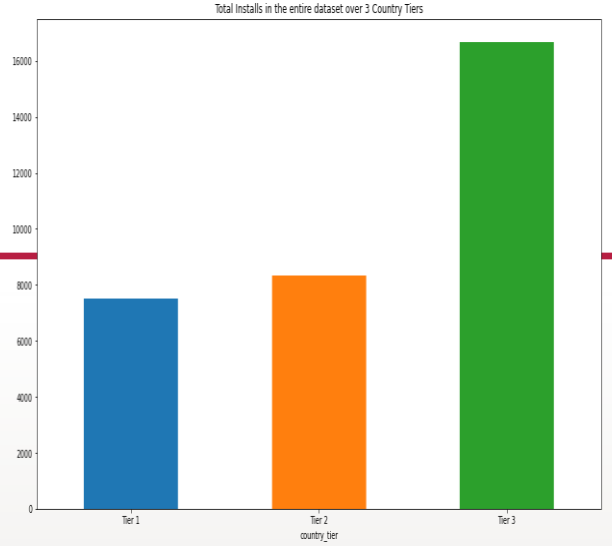
DATA OVERVIEW

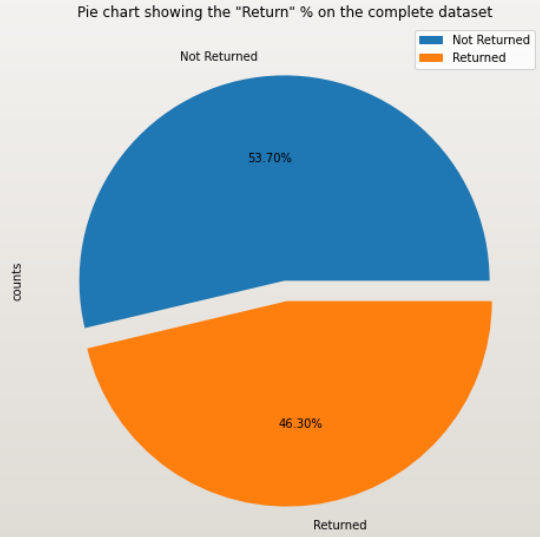
- Around 32.5k users installed the game from 194 countries.
- Installs happened on all seven days from both Facebook and Instagram.
- Platforms: iOS and Android
- Player performance is measured: battles, victories, scores, frag stats, etc.
- Player engagement is tracked: chests, diamonds spent, coins spent, club join, etc.
- **Return on the next day? Yes or No based on the 39 user attributes.**

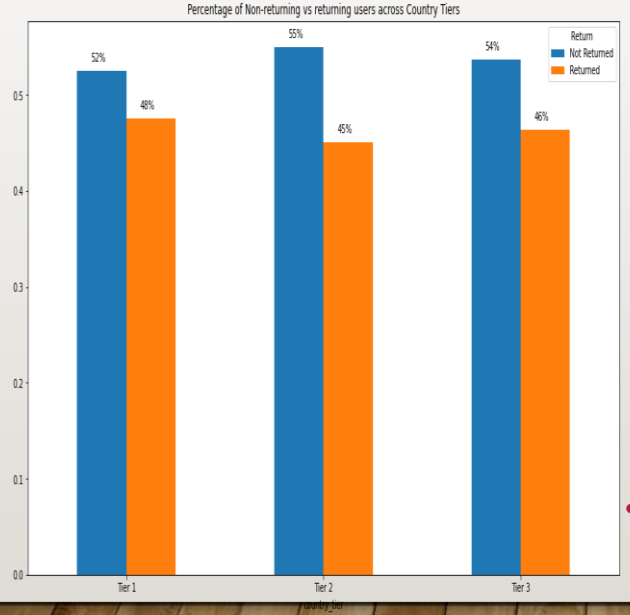
DATA: MISSING VALUES

- 'Best runs' and 'Worst runs' metric for users have 4k null values. We dropped the entire fields from the dataset as we cannot impute them reasonably without knowing the data generating process.
- 'Last score' of 11 users were missing; handled assigned value of 250.
- 1 user's country detail was missing; assigned the user to USA.
- 1 user's player performance metrics total_player_frags','max_player_frags',etc. was missing; dropped the user.







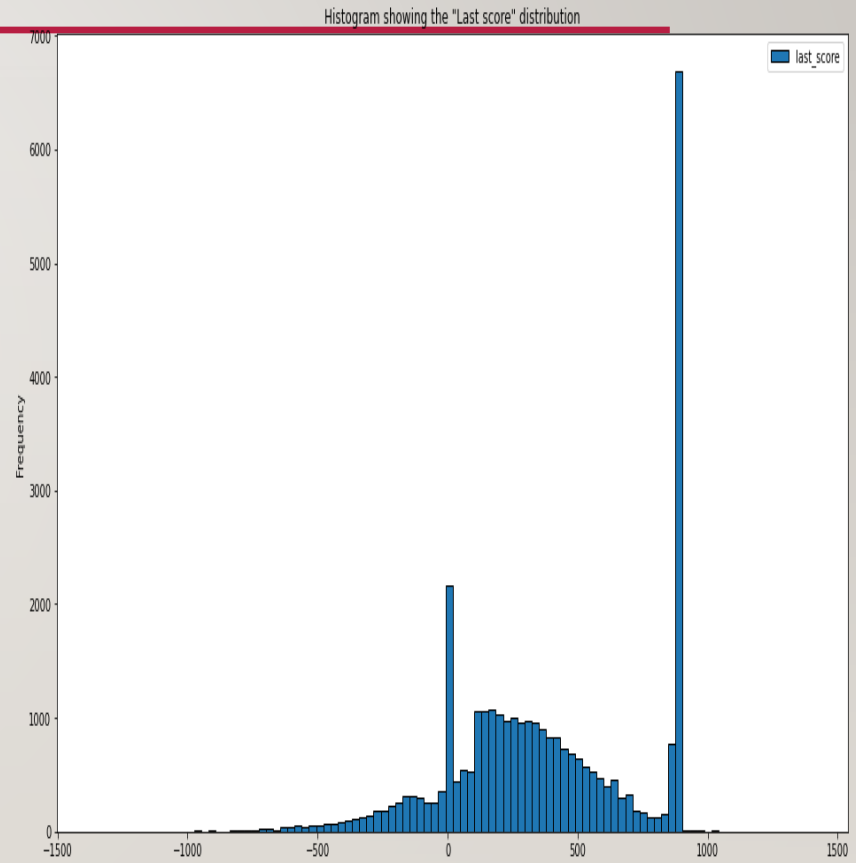


EXPLORATORY DATA ANALYSIS

- Around 46% of users returned the next day.
- Most users came from the following countries arranged in descending order:
  - Tier 1
    - United States
  - Tier 2
    - France
    - Germany
    - Japan
    - Saudi Arabia
    - Italy
    - South Korea
  - Tier 3 – Rest of the countries
- Users from Country Tier 1 have a better chance of returning back the next day.

EXPLORATORY DATA ANALYSIS (CONTD.)

- Facebook installs return more as compared to Instagram in USA
- Installs from Instagram tend to have better retention.
- Iphone users have a better chance of returning as compared to android phone users.
- Weekend install day users are less likely to return when compared to weekday-install day users.
- Last Score metric is heavily skewed with 2 modes.
- Users whose last score is either 0 or 900 have a higher chance of churning.



USER SEGMENTATION BASED ON PLAYER BATTLE PREFERENCE

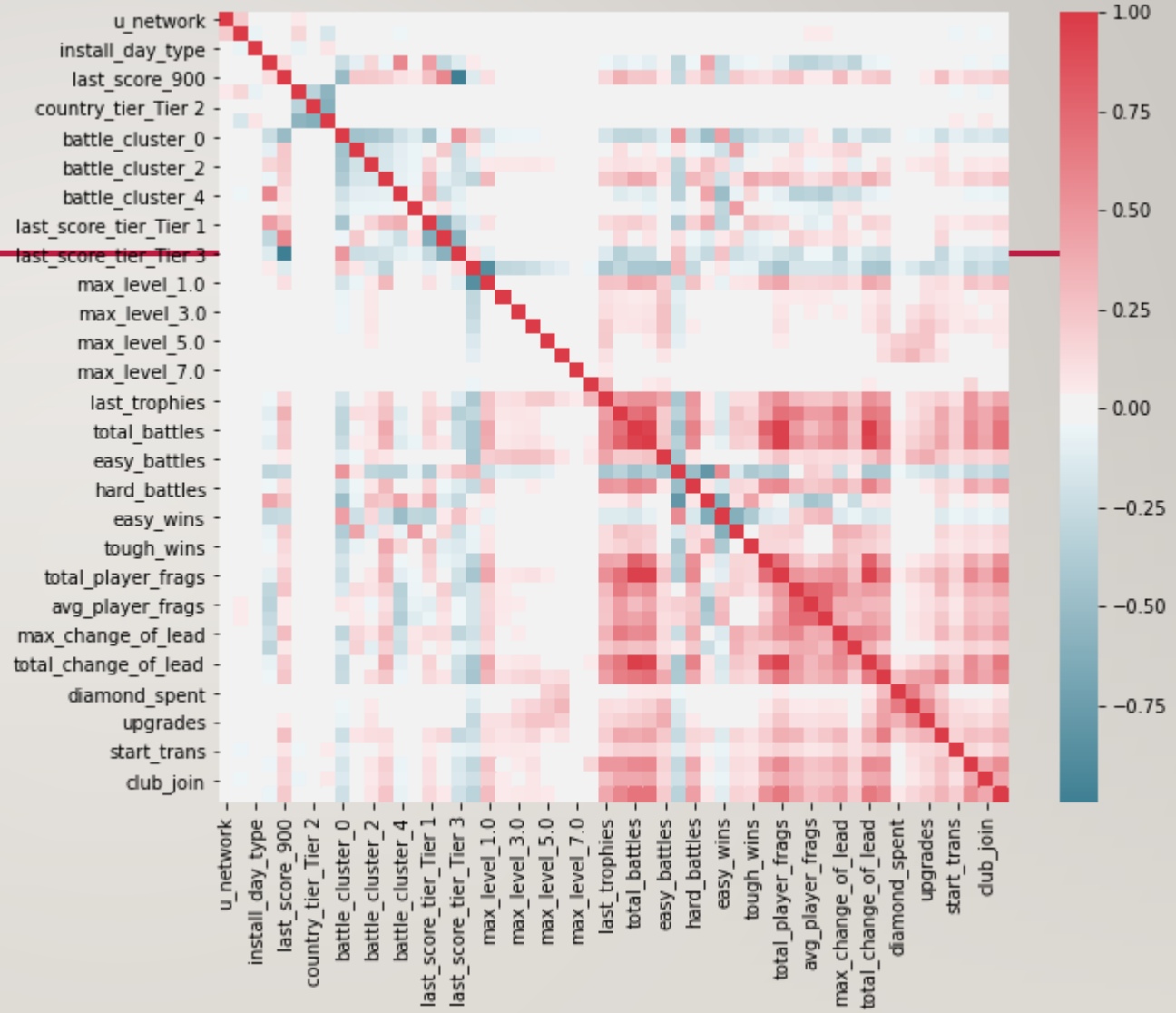
- Instead of using actual battle counts of the users ('easy_battles', 'fair_battles', 'hard_battles', 'tough_battles'), we use the proportion of games played at different difficulty levels.
- We use the computed proportions of victories and the corresponding difficultly of games played to segment the users using K-means clustering.
- Cluster 0 has a high tendency to play fair battles and are more likely to not return.

USER CLUSTERS AFTER SEGMENTATION

Cluster	easy_battles	fair_battles	hard_battles	tough_battles	easy_wins	hard_wins	tough_wins	Not Returned	Returned
0	0.000	0.993	0.002	0.004	0.987	0.012	0.001	10450	5684
1	0.002	0.918	0.016	0.064	0.679	0.307	0.015	2377	3116
2	0.036	0.723	0.038	0.203	0.946	0.024	0.030	2361	2719
3	0.030	0.676	0.101	0.193	0.654	0.157	0.189	950	3062
4	0.012	0.500	0.020	0.467	0.007	0.016	0.150	1028	271
5	0.011	0.751	0.015	0.223	0.067	0.914	0.019	304	208

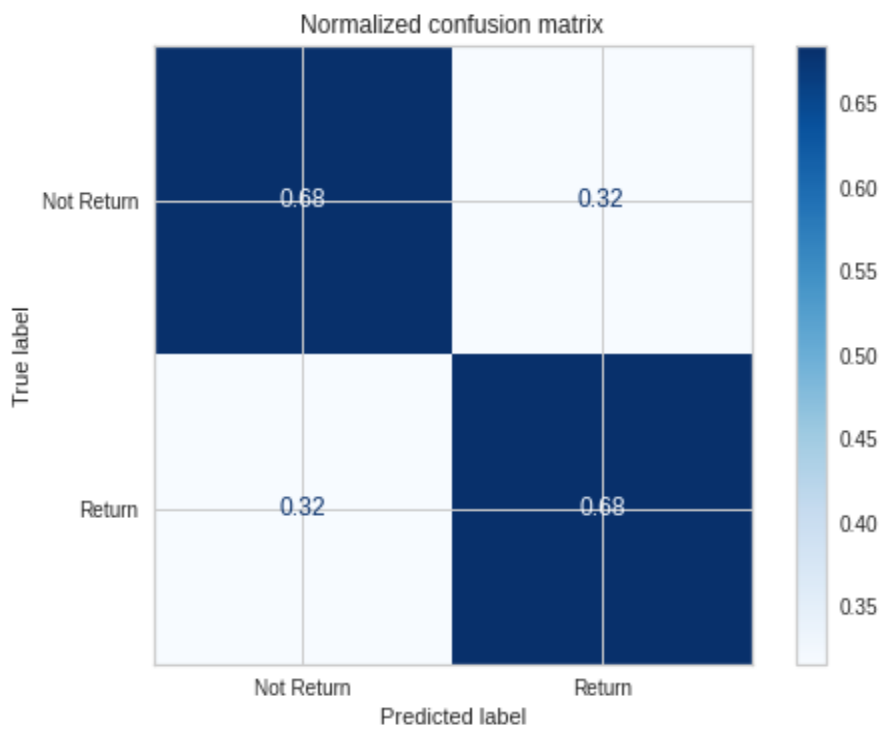
CORRELATION IN USER ATTRIBUTES

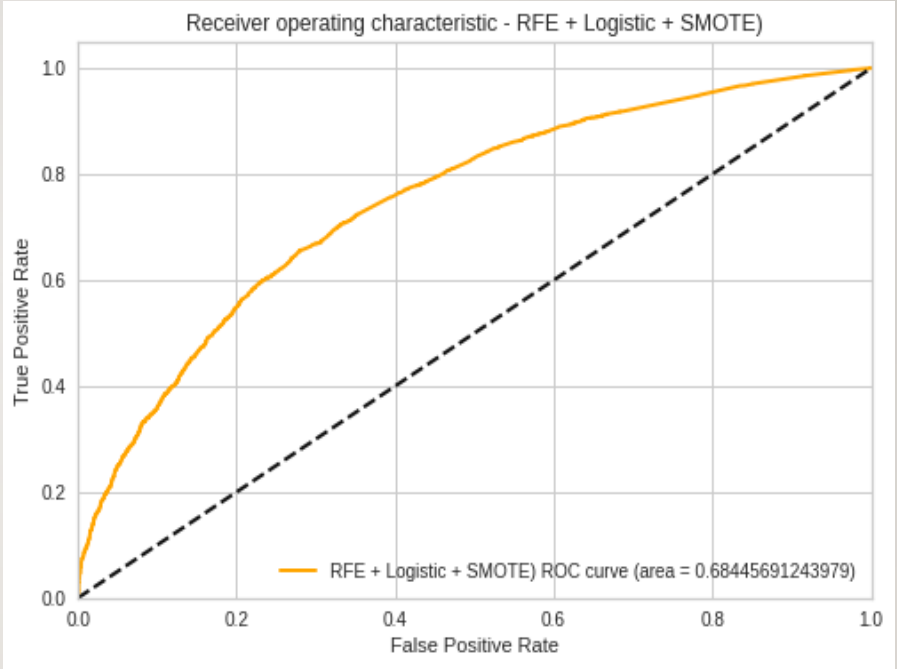
- As expected, there is a high degree of correlation among different user attributes.
- The users who generally make more progress in the game have a higher chance of having proportional user engagement like playing more games, making higher scores, etc.



MODELLING RETENTION – LOGISTIC REGRESSION

- Logistic Regression(Baseline):
  - Precision – 0.69, Recall – 0.69, F1 score – 0.69, AUC ROC – 0.682
- Logistic Regression + SMOTE:
  - Precision – 0.69, Recall – 0.69, F1 score – 0.69, AUC ROC – 0.683
- Logistic Regression + SMOTE + Recursive Feature Elimination:
  - Precision – 0.69, Recall – 0.68, F1 score – 0.68, AUC ROC – 0.684

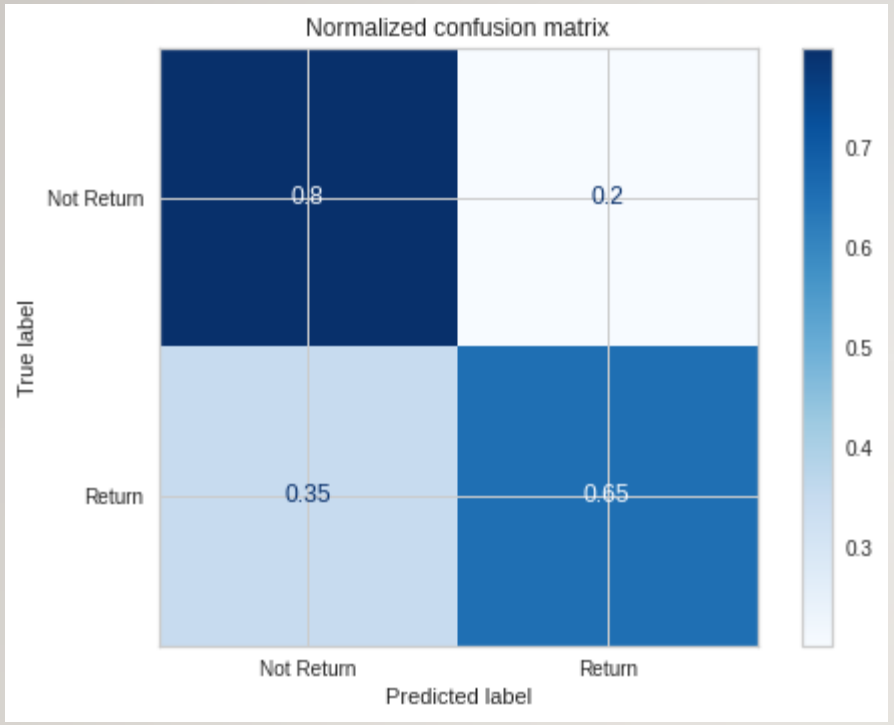


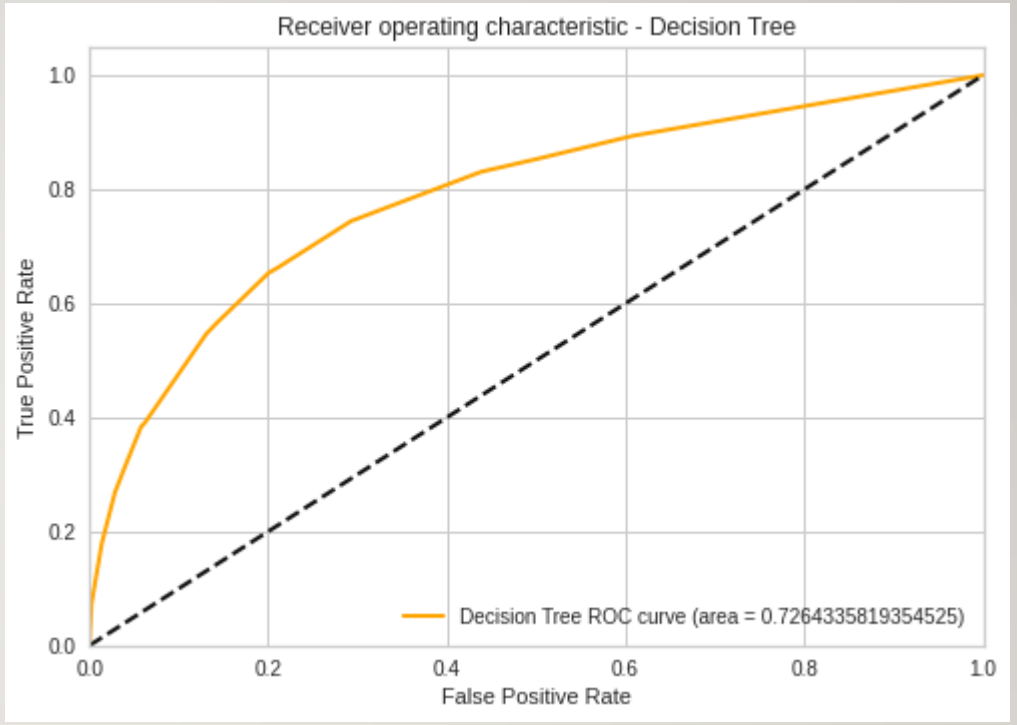


LOGISTIC - CONFUSION MATRIX, ROC CURVE

MODELLING RETENTION – DECISION TREE

- Since logistic regression is not able to improve the performance, we use decision trees to model non-linearity.
- We can use the unnormalized dataset as the functioning of decision tree is different from logistic regression. Also, this model is easy to interpret.
- Decision Tree classifier:
  - Precision – 0.73, Recall – 0.73, F1 score – 0.73 AUC ROC – 0.726 (Best model)





DECISION TREE - CONFUSION MATRIX, ROC CURVE

RECOMMENDATIONS

- Acquire more iphone users, it might be more expensive though!
- Provide more chests to users.
- Ensure that the user who have less than 7 victories, opens chest at least 3 times on the day of install. This can be programmed using push notification.
- Make users receive more than 5 diamonds on their install day.
- Try to encourage users to play more battles: cascade battles back to back with fair and hard players.
- Guide the user to a club close to the corresponding user segment.