How can we increase revenue from Catch the Pink Flamingo?

Chakresh Kumar Singh

Problem Statement

How can we use the following data sets to understand options for increasing revenue from game players?

> Users Data -

- Summarising user behavior while playing the game
    - 8 tables connected via primary keys, can be processes as a relational database

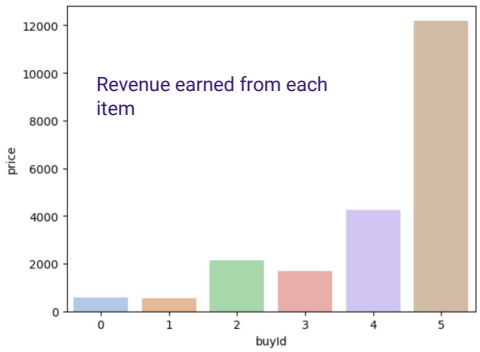
> Chat Data -

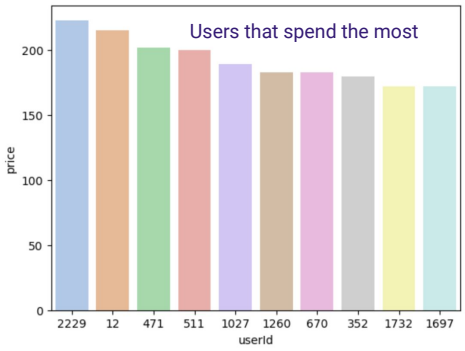
- Information on the interaction between users within teams as a graph dataset
    - 6 tables with unique Ids for users, teams and chat Items

With access to the in-depth game data we aim to identify ways to increase revenue for the pink-flamingo game. The Data is divided into two parts -

- User data - These are tables with data related to user behavior and their interactions while playing the game. These can be used in relational databases for analysis
- Chat data - These are tables containing data of user interaction within their teams. This can be used in graph-based databases to identify patterns in user interactions.

Data Exploration Overview

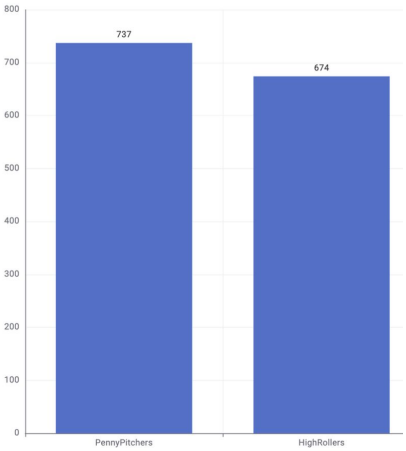




*price is in USD

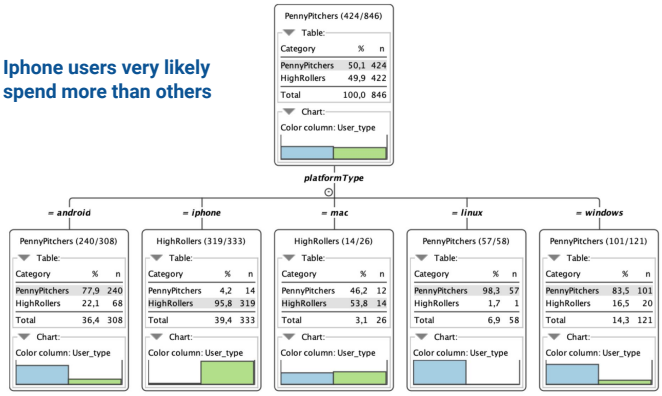
The data exploration reveals the patterns in the revenue earned. Product with id-5 is the most bought product whereas user with ids listed in the graph at right above are the top spenders.

What have we learned from classification?



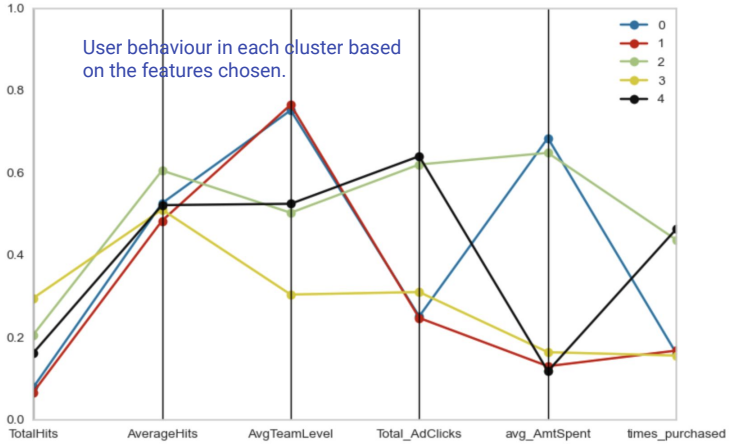
Users are divided between two categories based on their expenditure

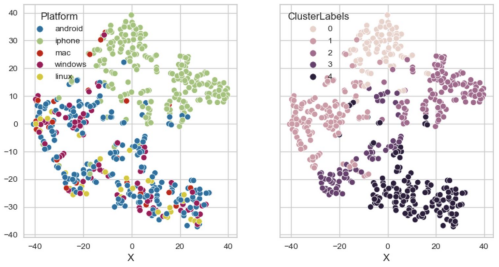
Iphone users very likely spend more than others



Users are divided between two categories based on their expenditure. To identify the dominating factors underlying user behaviors we run classification algorithms. We observe that iphone users are always spending the most. Whereas linux users are spending the least.

What have we learned from clustering?

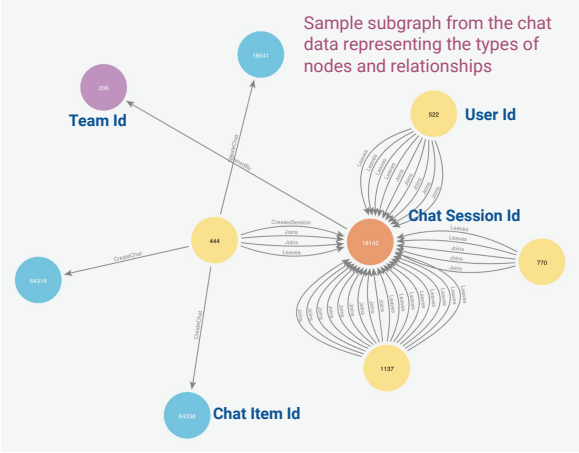




Similar users form clusters.

Graph on the left - Group 0 - clicks few ads but spends the most. Group 4 - clicks on most ads but spend the least. Group 3- probably new teams. Group2 - Ideal group. Graph on the right - Running dimensionality reduction on the standardised features shows clusters of users based on their patterns colored by their platform type (left) and the cluster they belong to (right)

From our chat graph analysis, what further exploration should we undertake?



Potential directions for further exploration

- Score and rank users, teams based on graph metrics
- Temporal analysis of the chat graph
- Correlation study between individual user behaviour and their interaction in the chat graph

- Interaction within teams should be rewarded by taking into consideration users influence score in the network for their rankings.
- Temporal analysis of the chat graph should reveal the moments when the teams are most interactive with most users online.
- Correlation between team interactions and user behaviors and help better identify factors governing user behavior while playing and the influence of their teams.

Recommendation

- Target iphone users for more ads and future products.
  - Offer discounts in general to promote more ad clicks and sales
- Users who are influential in the team should be targeted with promotions and more offers. They can help with reachability within teams.
- Temporal Graph analysis should be used to find moments when the teams have most users online and are interactive.
  - These time windows can be used to promote ads and offers.

1. Since we observed that iphone users spend the most, they should be targeted for future products and with more ads preferably with some rewards to promote sales. In general offer some discounts to promote users to click ads and buy.
2. Influential users can be effective information diffusers and should be targeted for offers and new products.
3. Finding time windows when most users are online and interactive within teams are ideal for promotion and can have a positive impact on revenues