


# How can we increase revenue from Catch the Pink Flamingo?

PRANAV BHARGAVA

# Problem Statement

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

These datasets provide a very important V of big data i.e. **variety**. The datasets help us determine the features which play a major role in determining how the revenue increases and what influences it. Revenue is directly dependent on people buying in app purchases. But it is also important to look at not so dependent but extremely important factors which affect revenue.



A variety in datasets ensure that ad clicks, buy clicks , game clicks and even chat content is provided to search for ample features which affect the revenue of Eglence Inc.

This data has been generated for this project but an actual game data has velocity in it i.e. the real time data coming in from users active at this particular instant is also received. Earlier only advertisements or in app purchases were the independent variables on which revenue depended. But as is clear with this dataset hit ratio's , game clicks, time spent playing etc also provide new ways to increase revenue.

# Data Exploration Overview

The exploration phase helps one acquaint with the dataset and allows one to identify different variables in different files. Also we perform statistical measurements like mean or max etc. to understand the variables in-depth.



This phase showed us that the even though there are only 6 items to buy with the most expensive product costing only 20, the overall amount spent is 21407. We observed the top 10 users who spent maximum and how spending more isn't related to the hits in game. Users who have spent less need not have less hits.

## What have we learned from classification?

This phase led to creating new features which are in the data but not obvious. We made new variables from existing ones and made a decision tree to help us classify unclassified users into highRollers category or PennyPinchers category. According to the decision tree, the platform used to play this game is the main independent variable which affects our depending variable. Thus a user using **iphone** has greater chances of being a high roller.



Assume you have 2 minutes to present what you perceive to be the most important or remarkable points from your classification analysis.

We learnt how to make new variables by classifying a numeric variable into factors(based on limits). The conversion of a numeric variable to a new variable with just 2 categories led to the resulting conclusion of who tends to spend more and how. Certain variables were filtered. A training model was created based on given data and predictions were made on the new data which wasn't used to model.

## What have we learned from clustering?

Users in the 3rd cluster should be shown more ads and given more app purchases. This is because these users have the maximum adClicks and generate most revenue along with the highest hit-percentage. So these users can be offered more expensive products to buy just after they achieve a certain number of hits and be shown more ads.

Users in 2nd cluster should be given cheaper products and more ads. Users with lowest adClicks and least revenue should be shown cheaper products as they would be motivated to buy whenever they hit a flamingo.



Assume you have 2 minutes to present what you perceive to be the most important or remarkable points from your classification analysis.

Clustering made us create groups based on variables. It helped us to differentiate among users who spend less, spend medium and those who spend the most. The variables used to cluster were adCount, price and hit-percentage. I created 3 clusters and thus did my analysis.

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

Chat graph analysis determines longest chats, chattiest users and the people who chat more means they are more active. It helps segregate active and passive users. It also helps understand users more whether they are a Type A person or Type B.



Personality of users can be determined from the type of chat they do. How long they chat for and how negative or positive their chat messages are. Chat analysis also helps understand how groups or teams are formed, how like minded people come together to form teams.

## Recommendation

- For users spending more advertise items with high cost and vice versa.
- If a player plays continuously for supposedly 3 hours then the user should purchase any item (of any value) to continue playing or close game to return after 3 hours.
- If the hit-ratio of a user decreases then he can pay to increase pink flamingo sightings around him for say half an hour. But there is no guarantee than he hits the pink flamingo.
- Items that only a team as a whole can purchase should be introduced.

The rationale for the 1<sup>st</sup> idea is that users who spend more don't tend to consider low or high price and thus can be sold items with high prices only.

If a player played 3 0 hours continuously means he is addicted and thus cannot take a break for more than an hour and thus should purchase something to continue playing immediately.

If a player's hit ratio decreases they feel demotivated and thus to increase hit ratio can buy the feature to get a chance to increase their or their team's hit ratio.

Items purchased by a team would help to maintain team loyalty as everyone invests as whole and thus would make the team more united and make them strive to hit more pink flamingos.