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

METAN KONE

# Problem Statement

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

The datasets presented to us come from logs resulting from the **Catch the pink flamingos** game application. These logs are records of the players actions, and capture the players habits and behaviour, even on which platform (android, iphone) they use to play. These data also capture the events related to all the purchases done on the application (which players category do lot of purchases, …).

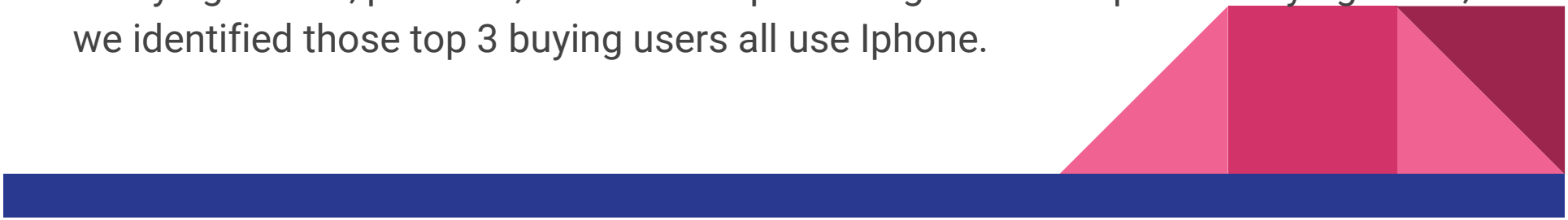
So by exploring and analyzing these big data with analytics technics, we can do recommandations in line with business strategy in order to increase the revenue of company.

# Data Exploration Overview

Our data data exploration analysis focused on the following csv files : **buy-clicks.csv, game-clicks.csv, user-session.csv, ad-clicks.csv, users.csv, team.csv, team-assignments.csv, level-events.csv.**

Using **Splunk**, We computed aggregations key indicators to identify for example the amount spent buying items or how many times each item is purchased.  We used histograms visualization for a better overview of these data.

Studying user id, platform, and hit-ratio percentage for the top three buying users, we identified those top 3 buying users all use Iphone.



# What have we learned from classification?

In the classification analysis, a new categorical attribute was created to enable analysis of players as broken into 2 categories (HighRollers and PennyPinchers). *HighRollers do more purchases while PennyPinchers spend very little money on the application.*

Using Knime, we predicted which users  falls  into HighRollers category, or in PennyPinchers category.
**We identified that players on iphone are high HighRollers while users on Linux are high PennyPinchers**.

# What have we learned from clustering?

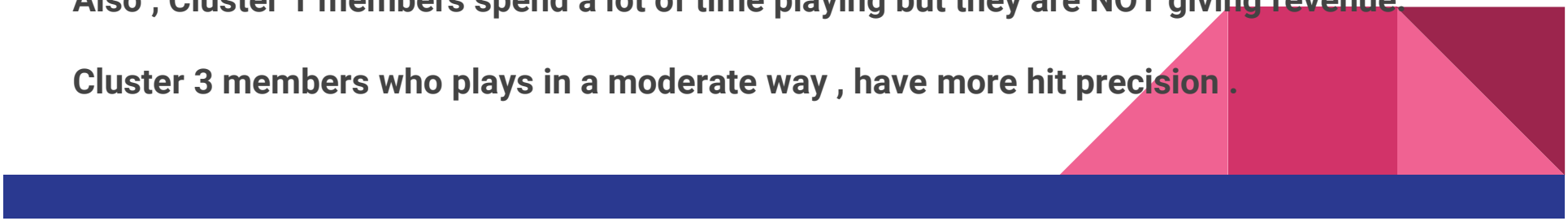
Using Apache Spark, we have done classification based on the 4 following features identified : **count_gameclicks, count_buyId, avg_price, hit_precision**.

Using elbow method, we saw optimum number of clusters was 5 :

**Cluster 0 has the highest number count_buyID, but we can notice the revenue is coming most from cluster 4.**

**Cluster 1 members are most playing . We can therefore notice they have the less precision. Also , Cluster 1 members spend a lot of time playing but they are NOT giving revenue.**

**Cluster 3 members who plays in a moderate way , have more hit precision .**



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

Chattier users and initiators of longer conversations can be more valuable, because of their potential to spread information to other communities members. Eglence, Inc. can target those members to increase its revenue by showing them more expensive items to such users, and showing them more adverts.

# Recommendation

**Show more advert to iphone users**.

**Increase advert price for iphone platform.**

**Chattier users and initiators of longer conversations can be more valuable, because of their potential to spread information to other communities members. Eglence, Inc. can target those members to increase its revenue by showing them more expensive items to such users, and showing them more adverts.**