

Slide 1

Hi, my name is Yeow Long. I'll now proceed to discuss on how we can increase revenue from Catch the Pink Flamingo.

Slide 2

It is important to obtain different kinds and sources of data so that we extract customers buying habits and patterns and determine the top ten spenders, how much they spend and what they spend on, to obtain understand this trends. After performing data exploration, it is possible to determine the options/possibilities/ways to increase revenue from the game players. It is also possible to uncover previously unseen options.

Slide 3

We see that out of a total of 4619 samples, only 1411 transactions were made. A new categorical feature attribute was created to enable analysis of players. As this is supervised learning, we need labels for our datasets. We break them down into HighRollers and PennyPinchers. The creation of this categorical attribute is also necessary as this is a classification problem. After completing the data exploration and preparation step, we decided to remove features such as avg_price, user_id and user_session_id as they were deemed unnecessary.

Slide 4

The overall accuracy of the model is 207 HighRollers predicted correctly, 10 HighRollers predicted incorrectly, 285 PennyPinchers predicted correctly, 63 PennyPinchers predicted incorrectly.

From analysis results, it was determined that iPhone users are HighRollers with a percentage rating of 93.2% and other platform type users are PennyPinchers 6.8%.

Slide 5

From cluster analysis, 3 cluster centers were identified. These clusters can be differentiated from each other as follows:

Cluster 1 is different from the others in that users produce average number of hits but made the most purchases.

Cluster 2 is different from the others in that the users who produce less hits and also made less purchases.

Cluster 3 is different from the others in that the users who produce the most hits but made less purchases.

There are two recommended actions.

Firstly, we should increase the number of adds to users who play alot and the rational is that users who play alot are also users who spend less and click less on adds, this adds increase will promote users to spend more and hence, increase the revenue.

Secondly, we should show higher price ads to users who spend more. One behaviour noticed was that they do not play often and when they do, they always spend and hence, by showing them the more valuable adds we can increase revenue.

Slide 6

From chat graph analysis, we have determined the longest conversation chains and its participants. There are 9 such chats and 5 users participated in this chain. We have also analyzed the relationship between the top 10 chattiest users and top 10 chattiest teams and noticed that the most chattiest users are not in the chattiest teams. Finally we determined how active groups of users are.

Slide 7

Recommendation. I would recommend Eglence to build a ad delivery system / recommendation engine based on data collected, processed and analyzed. For example, from the results of clustering, one behaviour noticed was that they do not play often and when they do, they always spend and hence, by showing them the more valuable adds we can increase revenue.