

How can we increase revenue from Catch the Pink Flamingo?

Li Xiaowei

Problem Statement

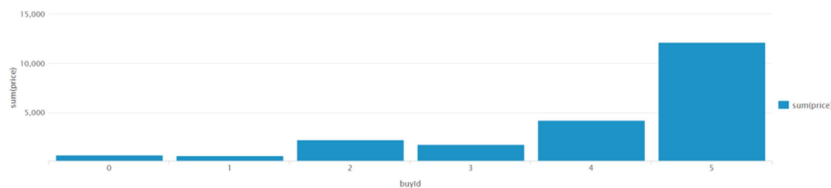
How can we use the following data sets to understand options for increasing revenue from game players?

File	Description
Chat-data.zip	Contains the user interactions data
Combined-data.zip	Contains the combined data set of user and user purchases
Flamingo-data.zip	Contains the user, team, game, advertisement and purchase data. The basic properties of the game revenue.

Basically, this study is find one or several ways to increase revenue from game players with the very riched data across game, revenue properties, user interaction records.

Data Exploration Overview

The 6 csv files provide very resourceful information on users' behaviors of clicking advertisements and in-app purchases. And below histogram is particularly interesting given that the buy clicks of all the 6 available items are around the same. But item 5 generated significant more revenue.



It could be a good idea to do something on item 5.

The purchase amount of the 6 items are ranging from 250 to 700, while the revenue generated from each item is from 600 to 12,200. Without any doubt, item 5 is the most expensive one but it's still very popular and demanding. Putting more resources in studying the impact and benefit of item 5 to understand the sentiment of why user purchasing this item can drive up the sales. Or having some promotions on this item to sell it to more users. Developing similar products is another option.

What have we learned from classification?

We could use one or many attributes to GUESS roughly whether a user is a high-value customer. And in this classification analysis, just by knowing user's mobile brands we would be able to determine whether the user is a "HighRollers" or not, with an error rate of only around 10%

Confusion Matrix - 4:5 - Scorer (Compute confusion...)

buyer \ Pr...	PennyPincher	HighRollers
PennyPincher	308	27
HighRollers	38	192

Correct classified: 500
Accuracy: 88.496 %
Cohen's kappa (κ) 0.76

Wrong classified: 65
Error: 11.504 %

= iPhone

HighRollers (279/336)

Category	%	n
PennyPincher	17.0	57
HighRollers	83.0	279
Total	39.7	336

Chart:
Color column: buyer

iPhone users produce significant more revenue and have much higher chance to spend more money.

It's quite surprising that just by knowing which platform the user is from we can determine whether this user is a "HighRoller" or "PennyPincher", with just 10% error rate. In most of the cases iPhone is the most expensive mobile phone/platform. And iPhone users are more willing to spend money in the game. We could develop some iPhone specific add-ons or in-app items to promote the sales revenue.

What have we learned from clustering?

We have done some research on finding whether there is a correlation between user hit ratio and the money they have spent. The 546 users are divided into two clusters with center information as below.

Cluster #	Cluster Center (hitratio, revenue)
1	Array([0.12591902, 117.86206897])
2	Array([0.11197414, 24.29847495])

Even though the revenue dimension varies a lot for the two cluster centers, there is really very little difference on the hit ratio dimension.

Not much valuable information found from this approach,
but at least it's worthy trying and conclude there is no correlation.



We want to explore the correlation between user hit ration and the money they have spent. Maybe one user with higher hit ration enjoys the game more and thus more willing to spend money. We do see two cluster centers far apart from each other, which is a good sign for us to cluster a given user. However, the hit ration dimension is really close to each other and we'd rather treat them as no difference. So not much suggestion can be given from this specific study but we do know that there is indeed very little correlation.

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

The chattiest user groups show a very high “dense” neighborhood. Usually a high cohesion is easier to build among them.

Further study on identifying these neighborhoods and pushing some promotions to one or some users from the neighborhoods should easily attract others attention. **The per user revenue may be dropped but overall will increase.**




Usually a high cohesion can be found from a high “dense” neighbourhood because they know better about each other and share more common language. And in this study we realized the most chattiest users usually form a “dense” neighbourhood. Some in-app purchases may be bought by few users from the neighbourhood. We can let the neighbourhood know there is someone already purchase something, this will stimulate the others to make the same purchases as they share more common language and interests.

Recommendation

To summarize on the studies performed.

1. The in-app purchase item 5 is quite demanding and generates almost half of the revenue. Having a round of promotion should easily attract more purchase and users.
2. iPhone users produce significant more revenue and have much higher chance to spend more money. Exclusive add-ons for iPhone users can be considered.
3. The chattiest users usually form very "dense" neighborhoods. Pushing some surprise promotions to the few users from the neighborhoods should easily attracts the community's attention and thus generate more revenue.



So by putting all the studies and findings together, we can focus on selling and developing item-5-like products. iPhone users are definitely our priority one for our new products and promotions, exclusive add-ons for iPhone can also be considered. For a more chatty user, we can put more resources in building and studying his neighbourhood and treat this neighbourhood as a unit to make advertisement and promotions. For example, let him/her know the status of playing game and item purchase from another one in the neighbourhood he/she is familiar with.