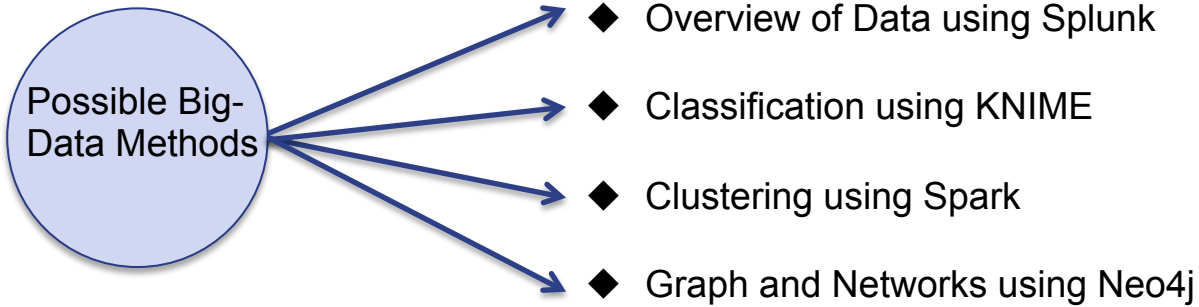
How can we increase revenue from Catch the Pink Flamingo?

By: Dibakar Sigdel

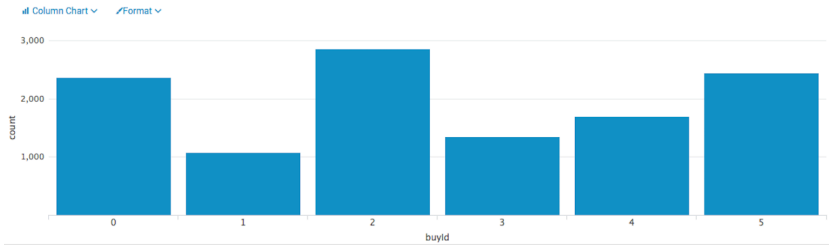
Problem Statement

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

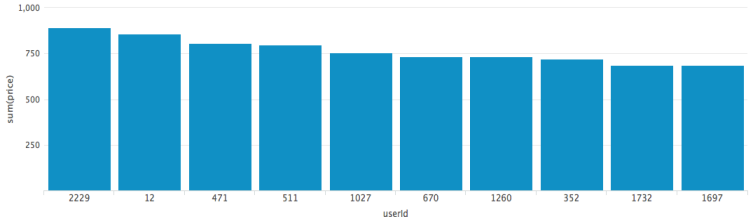
- Flamingo Data set
  - Add clicks info
  - Buy clicks info
  - Game clicks info
  - Users info
  - User Session info
  - Team info
  - Team assignment info
- Chat Data Set
  - Chat create team info
  - Chat item team info
  - Chat join team info
  - Chat leave team info
  - Chat mention team info
  - Chat respond team info



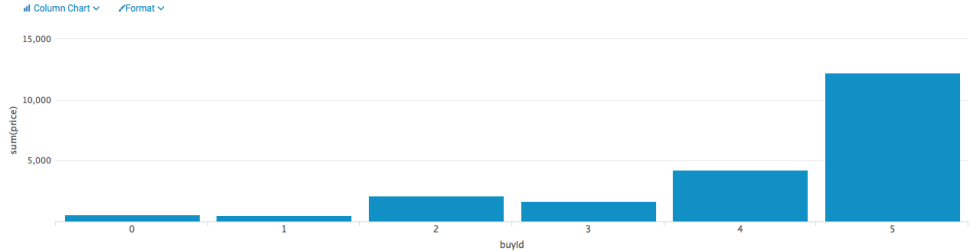
Data Exploration Overview



A histogram showing how many times each item is purchased:

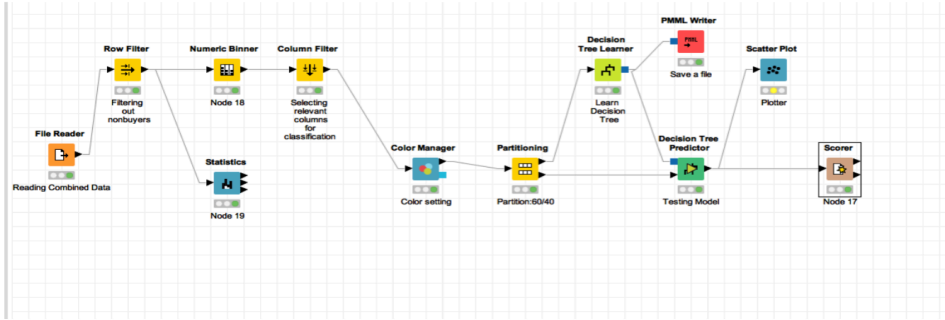


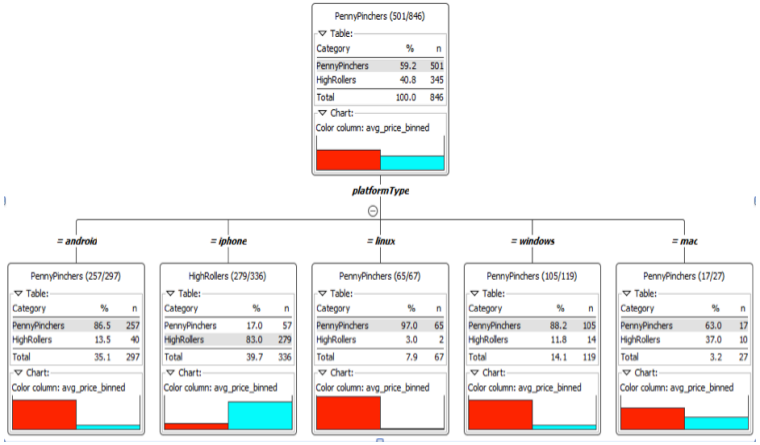
A histogram showing total amount of money spent by the top ten users



A histogram showing how much money was made from each item:

Classification-Analysis





Confusion Matrix

	PennyPitchers	HighRollers
PennyPitchers	True Positive TP(308)	False Negative (27)
HighRollers	False Positive (38)	True Negative (192)

Accuracy = (TP + TN) / (TP +TN + FP + FN)

= 308 + 192 / 308 + 192 + 38 + 27

= **0.88496 (which matches)**

Specific Recommendations to Increase Revenue
1. Since iphone users are HighRollers, the game offers should be high values if platformType is iphone, since that will be more appealing to them and are more likely to buy them.
2. Since other operating system device (i.e. android, mac, windows etc) and linux are generally PennyPinchers, they should be given low values. It is possible that they will purchase low value items but in more volume, so there should be some other strategy.

Clustering - Analysis

Selected attributes:

Price: characteristics of a user's purchase behavior. It is directly associated to profit made from user.

AdCounts: Ad clicks represents user's purchase frequency as well as add frequency. This represents dynamic relationship between users, add providers and host.

# of clusters	Cluster
2	(27.39, 23.86), (39.08, 115.26)
3	(34.62, 59.29), (40.87, 138.24), (25.33, 15.29)
4	(40.63, 23.53), (41.09, 146.39), (33.45, 69.41), (14.85, 12.65)

Three cluster center is the best approximation. 4-cluster shows over fitting!

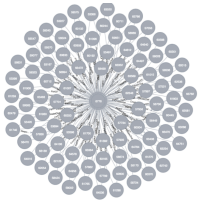
Recommended Action:

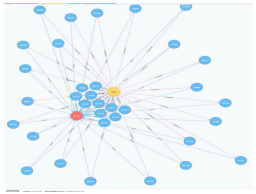
1. Increase ads quality by picking targeted ads for frequent clickers.

2. Place higher percentage of ads shown to the high revenue generating users.

3. Charge higher fees for hosting ads shown to higher revenue generating buyers.

Graph Analysis and exploration





Graph-nodes
1. User
2. Team
3. ChatItem
4. TeamChatSession

Graph Edges:
1. CreatChat
2. CreatSession
3. Joins
4. Leaves
5. OwnedBy
6. PartOf
7. Mentioned
8. RespondTo

Exploration :

1. **Finding the longest conversation chain and its participants**
2. **Analyzing the relationship between top chattiest users and top chattiest teams**
3. **How Active Are Groups of Users? : *Cluster Coefficient***

Most Active Users (based on Cluster Coefficients)

User ID	Coefficient
209	0.95
554	0.90
1087	0.8

Recommendation

1. Since iPhone users are High Rollers, the game offers should be high values if platform Type is iPhone.
2. Since other operating system device (i.e. android, mac, windows etc.) and Linux are generally Penny Pinchers, they should be given low values. It is possible that they will purchase low value items but in more volume, so there should be some other strategy.
3. Increase ads quality by picking targeted ads for frequent clickers
4. Place higher percentage of ads shown to the high revenue generating users
5. Charge higher fees for hosting ads shown to higher revenue generating buyers.
6. Chattiest group and chattiest users must be in focus. Individual users must be valued by their cluster coefficients.

THANK YOU!