

# Data Exploration

## Data Set Overview

The table below lists each of the files available for analysis with a short description of what is found in each one.

File Name	Description	Fields
<b>ad-clicks.csv</b>	A line is added to this file when a player clicks on an advertisement in the Flamingo app.	timestamp: when the click occurred.  txID: a unique id (within ad-clicks.log) for the click  userSessionid: the id of the user session for the user who made the click  teamid: the current team id of the user who made the click  userid: the user id of the user who made the click  adID: the id of the ad clicked on  adCategory: the category/type of ad clicked on
<b>buy-clicks.csv</b>	A line is added to this file when a player makes an in-app purchase in the Flamingo app.	timestamp: when the purchase was made.  txID: a unique id (within buy-clicks.log) for the purchase  userSessionid: the id of the user session for the user who made the purchase  team: the current team id of the user who made the purchase  userid: the user id of the user who made the purchase  buyID: the id of the item purchased  price: the price of the item purchased
<b>users.csv</b>	This file contains a line for each	timestamp: when user first played

	<p>user playing the game.</p>	<p>the game.</p> <p>id: the user id assigned to the user.</p> <p>nick: the nickname chosen by the user.</p> <p>twitter: the twitter handle of the user.</p> <p>dob: the date of birth of the user.</p> <p>country: the two-letter country code where the user lives.</p>
<b>team.csv</b>	<p>This file contains a line for each team terminated in the game.</p>	<p>teamid: the id of the team</p> <p>name: the name of the team</p> <p>teamCreationTime: the timestamp when the team was created</p> <p>teamEndTime: the timestamp when the last member left the team</p> <p>strength: a measure of team strength, roughly corresponding to the success of a team</p> <p>currentLevel: the current level of the team</p>
<b>team-assignments.csv</b>	<p>A line is added to this file each time a user joins a team. A user can be in at most a single team at a time.</p>	<p>time: when the user joined the team.</p> <p>team: the id of the team</p> <p>userid: the id of the user</p> <p>assignmentid: a unique id for this assignment</p>
<b>level-events.csv</b>	<p>A line is added to this file each time a team starts or finishes a level in the game</p>	<p>time: when the event occurred.</p> <p>eventid: a unique id for the event</p> <p>teamid: the id of the team</p>

		<p>level: the level started or completed</p> <p>eventType: the type of event, either start or end</p>
<b>user-session.csv</b>	<p>Each line in this file describes a user session, which denotes when a user starts and stops playing the game. Additionally, when a team goes to the next level in the game, the session is ended for each user in the team and a new one started.</p>	<p>timeStamp: a timestamp denoting when the event occurred.</p> <p>userSessionId: a unique id for the session.</p> <p>userId: the current user's ID.</p> <p>teamId: the current user's team.</p> <p>assignmentId: the team assignment id for the user to the team.</p> <p>sessionType: whether the event is the start or end of a session.</p> <p>teamLevel: the level of the team during this session.</p> <p>platformType: the type of platform of the user during this session.</p>
<b>game-clicks.csv</b>	<p>A line is added to this file each time a user performs a click in the game.</p>	<p>time: when the click occurred.</p> <p>clickid: a unique id for the click.</p> <p>userid: the id of the user performing the click.</p> <p>usersessionid: the id of the session of the user when the click is performed.</p> <p>isHit: denotes if the click was on a flamingo (value is 1) or missed the flamingo (value is 0)</p> <p>teamId: the id of the team of the user</p> <p>teamLevel: the current level of the team of the user</p>

## Aggregation

Amount spent buying items	21407
# Unique items available to be purchased	6

Analysis on the # of unique items available:

The number of 6 items to be purchased is very limited. Users, especially gamers here, like to collect items. I would suggest to augment fiercely this number.

A histogram showing how many times each item is purchased:



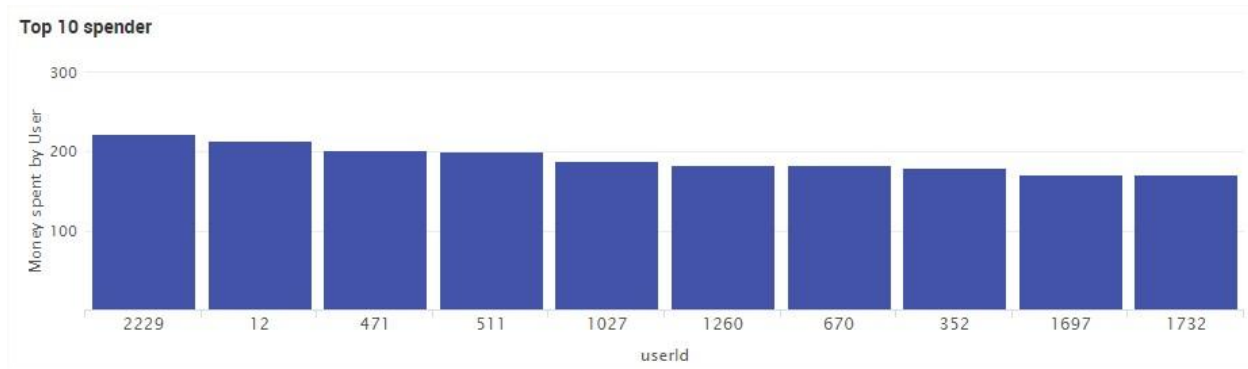
A histogram showing how much money was made from each item:



Analysis on the two above histograms: the repartition of the items bought and the amount spent by item are quite unequal. I would suggest to advertise more on the items less bought and average or quasi-equalize the price of the items.

## Filtering

A histogram showing total amount of money spent by the top ten users (ranked by how much money they spent).



The following table shows the user id, platform, and hit-ratio percentage for the top three buying users:

Rank	User Id	Platform	Hit-Ratio (%)
1	2229	iphone	0.115970
2	12	iphone	0.130682
3	471	iphone	0.145038

Analysis of the filtering tests:

The top 10 users by money spent or spenders have an average of 200 spent. The 3 top spenders seems to be iPhone owner and hit-ratio is going from 11% to 14%. I would suggest to expand the last analysis to more users to get a better sense of statistics. If we want, anyway, to conclude something of that last table: I would suggest

1/ to make more advertisement on Windows and Android users

and

2/ to lower the difficulty of the game as 11% or 16% of success hit shows a great difficulty to aim at the targets/flamingos. An easier game is played by more people, is more accessible.