

# 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.	<ul style="list-style-type: none"><li>● <b>timestamp</b>: when the click occurred.</li><li>● <b>txId</b>: a unique id (within ad-clicks.log) for the click</li><li>● <b>userSessionId</b>: the id of the user session for the user who made the click</li><li>● <b>teamid</b>: the current team id of the user who made the click</li><li>● <b>userid</b>: the user id of the user who made the click</li><li>● <b>adId</b>: the id of the ad clicked on</li><li>● <b>adCategory</b>: the category/type of ad clicked on</li></ul>
<b>buy-clicks.csv</b>	A line is added to this file when a player makes an in-app purchase in the Flamingo app.	<ul style="list-style-type: none"><li>● <b>timestamp</b>: when the purchase was made.</li><li>● <b>txId</b>: a unique id (within buy-clicks.log) for the purchase</li><li>● <b>userSessionId</b>: the id of the user session for the user who made the purchase</li><li>● <b>team</b>: the current team id of the user who made the purchase</li><li>● <b>userId</b>: the user id of the user who made the purchase</li><li>● <b>buyId</b>: the id of the item purchased</li><li>● <b>price</b>: the price of the item purchased</li></ul>

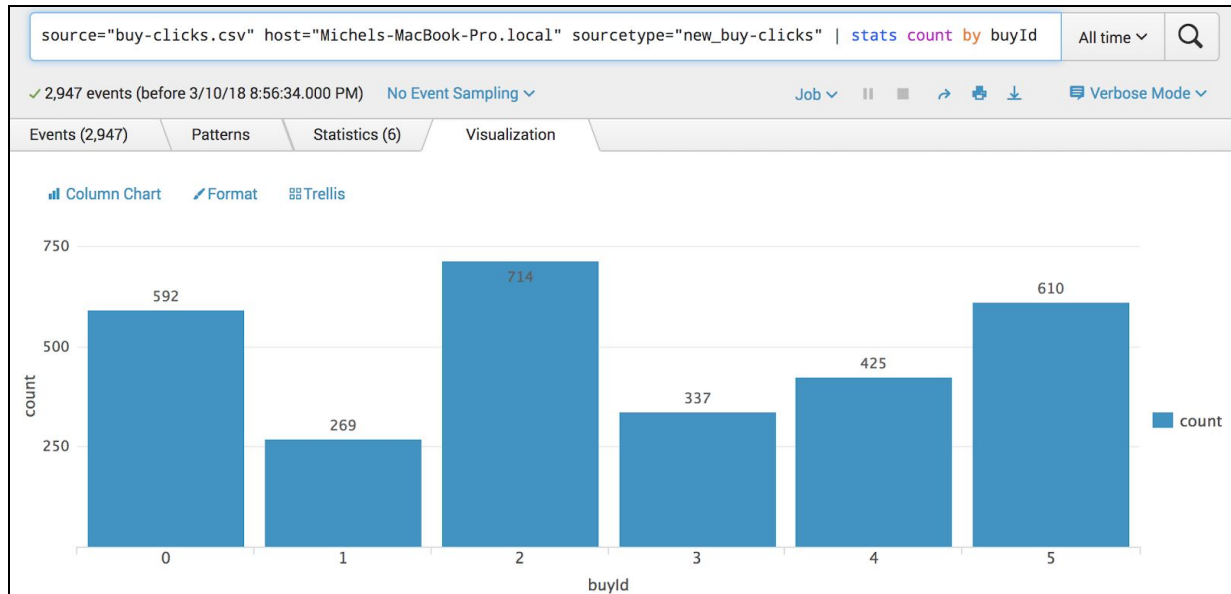
<b>users.csv</b>	This file contains a line for each user playing the game.	<ul style="list-style-type: none"> <li>• <b>timestamp:</b> when user first played the game.</li> <li>• <b>userId:</b> the user id assigned to the user.</li> <li>• <b>nick:</b> the nickname chosen by the user.</li> <li>• <b>twitter:</b> the twitter handle of the user.</li> <li>• <b>dob:</b> the date of birth of the user.</li> <li>• <b>country:</b> the two-letter country code where the user lives.</li> </ul>
<b>team.csv</b>	This file contains a line for each team terminated in the game.	<ul style="list-style-type: none"> <li>• <b>teamId:</b> the id of the team</li> <li>• <b>name:</b> the name of the team</li> <li>• <b>teamCreationTime:</b> the timestamp when the team was created</li> <li>• <b>teamEndTime:</b> the timestamp when the last member left the team</li> <li>• <b>strength:</b> a measure of team strength, roughly corresponding to the success of a team</li> <li>• <b>currentLevel:</b> the current level of the team</li> </ul>
<b>team-assignments.csv</b>	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.	<ul style="list-style-type: none"> <li>• <b>timestamp:</b> when the user joined the team.</li> <li>• <b>team:</b> the id of the team</li> <li>• <b>userId:</b> the id of the user</li> <li>• <b>assignmentId:</b> a unique id for this assignment.</li> </ul>
<b>level-events.csv</b>	A line is added to this file each time a team starts or finishes a level in the game	<ul style="list-style-type: none"> <li>• <b>timestamp:</b> when the event occurred.</li> <li>• <b>eventId:</b> a unique id for the event</li> <li>• <b>teamId:</b> the id of the team</li> <li>• <b>teamLevel:</b> the level started or completed</li> <li>• <b>eventType:</b> the type of event, either start or end</li> </ul>

<b>user-session.csv</b>	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.	<ul style="list-style-type: none"> <li>• <b>timestamp:</b> a timestamp denoting when the event occurred.</li> <li>• <b>userSessionId:</b> a unique id for the session.</li> <li>• <b>userId:</b> the current user's ID.</li> <li>• <b>teamId:</b> the current user's team.</li> <li>• <b>assignmentId:</b> the team assignment id for the user to the team.</li> <li>• <b>sessionType:</b> whether the event is the start or end of a session.</li> <li>• <b>teamLevel:</b> the level of the team during this session.</li> <li>• <b>platformType:</b> the type of platform of the user during this session.</li> </ul>
<b>game-clicks.csv</b>	A line is added to this file each time a user performs a click in the game.	<ul style="list-style-type: none"> <li>• <b>timestamp:</b> when the click occurred.</li> <li>• <b>clickId:</b> a unique id for the click.</li> <li>• <b>userId:</b> the id of the user performing the click.</li> <li>• <b>userSessionId:</b> the id of the session of the user when the click is performed.</li> <li>• <b>isHit:</b> denotes if the click was on a flamingo (value is 1) or missed the flamingo (value is 0)</li> <li>• <b>teamId:</b> the id of the team of the user</li> <li>• <b>teamLevel:</b> the current level of the team of the user</li> </ul>

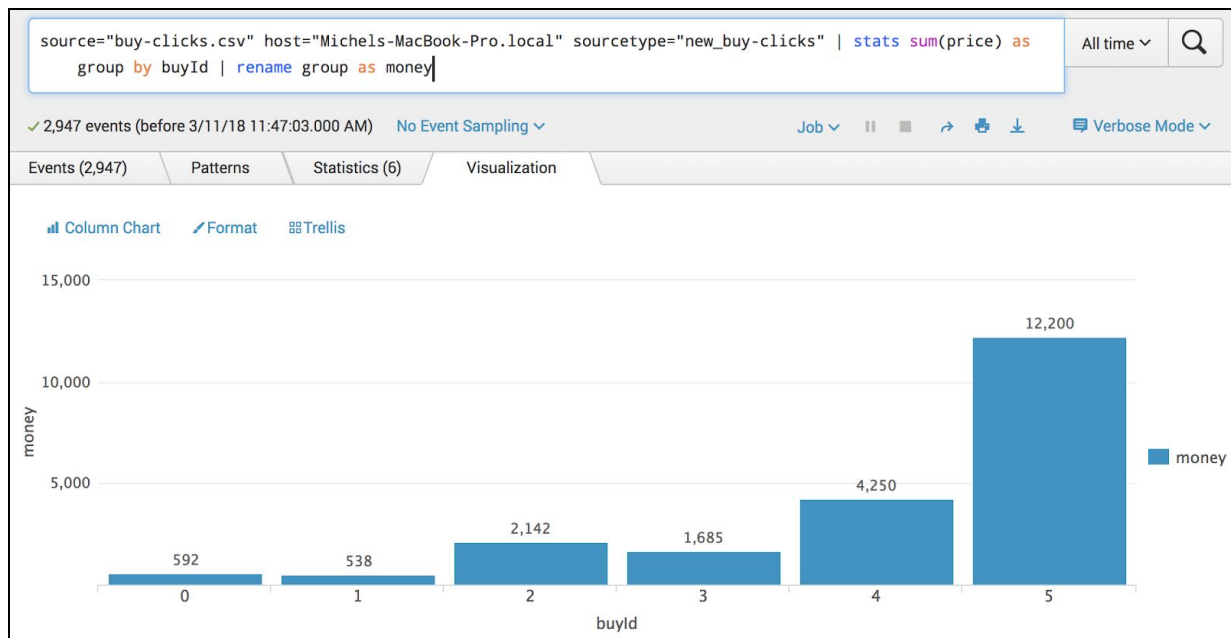
## Aggregation

Amount spent buying items	21407
Number of unique items available to be purchased	6

A histogram showing how many times each item is purchased:

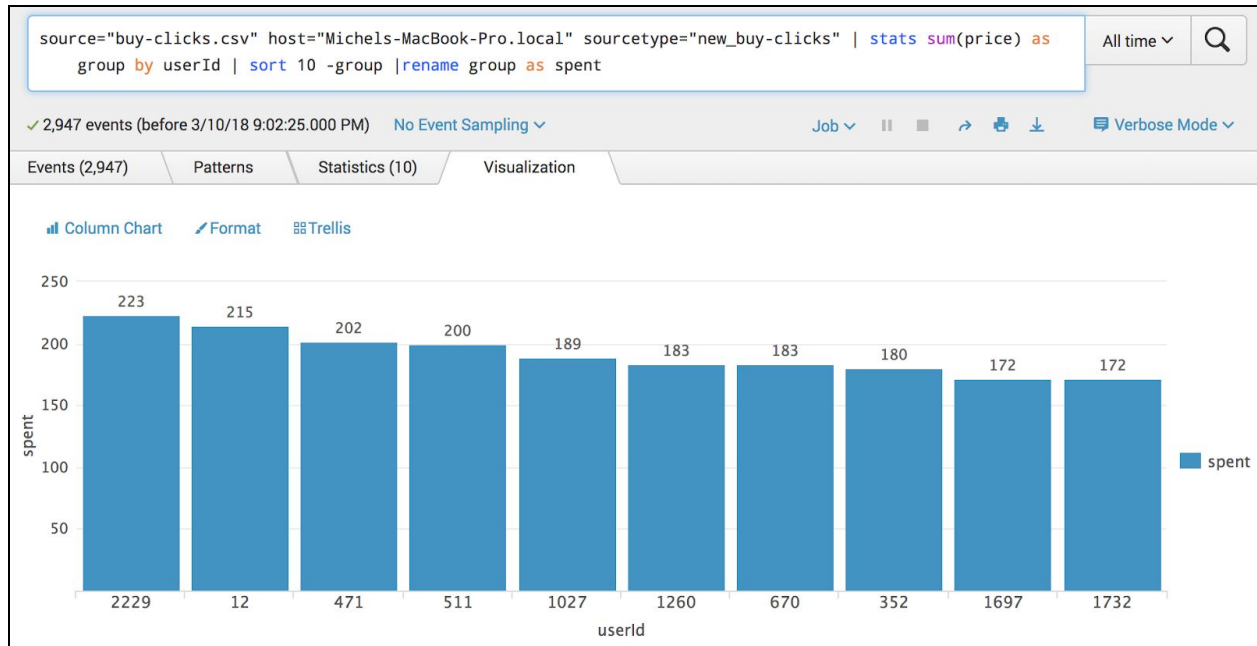


A histogram showing how much money was made from each item:



## 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	11.59
2	12	iphone	13.06
3	471	iphone	14.50