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
ad-clicks.csv	A record of each ad that was clicked on.	timestamp: date / time the click occurred. txId: a unique id 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 of ad clicked on
buy-clicks.csv	A list of purchases.	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
game-clicks.csv	A record of each click a user performed during the game.	timestamp: when the click occurred.
		clickId: a unique id for the click.
		userId: the id of the user performing the click.
		userSessionId: the id of the session of the user when the click is performed.
		isHit: denotes if the click was on a flamingo (value is 1) or missed the flamingo (value is 0)
		teamId: the id of the team of the user
		teamLevel: the current level of the team of the user
level-events.csv	A record of each level event for a team. Level events are recorded when a team ends or begins a new level	timestamp: when the event occurred.
		eventId: a unique id for the event
		teamId: the id of the team
		teamLevel: the level started or completed

		eventType: the type of event, either start or end
team-assignments.csv	A record of each time a user joins a team.	timestamp: when the user joined the team. team: the id of the team userId: the id of the user assignmentId: a unique id for this assignment
team.csv	A record of each team in the game.	teamId: the id of the team name: the name of the team teamCreationTime: the timestamp when the team was created teamEndTime: the timestamp when the last member left the team strength: a measure of team strength, roughly corresponding to the success of a team currentLevel: the current level of the team
user-session.csv	A record of each session a user plays. When a team levels up, each current user session ends and a new session begins with the new level.	teamId: the id of the team name: the name of the team teamCreationTime: the timestamp when the team was created

|		teamEndTime: the timestamp when the last member left the team strength: a measure of team strength, roughly corresponding to the success of a team currentLevel: the current level of the team
users.csv	A record for each user playing the game	timestamp: when user first played the game. userId: the user id assigned to the user. nick: the nickname chosen by the user. twitter: the twitter handle of the user. dob: the date of birth of the user. country: the two-letter country code where the user lives.

Aggregation

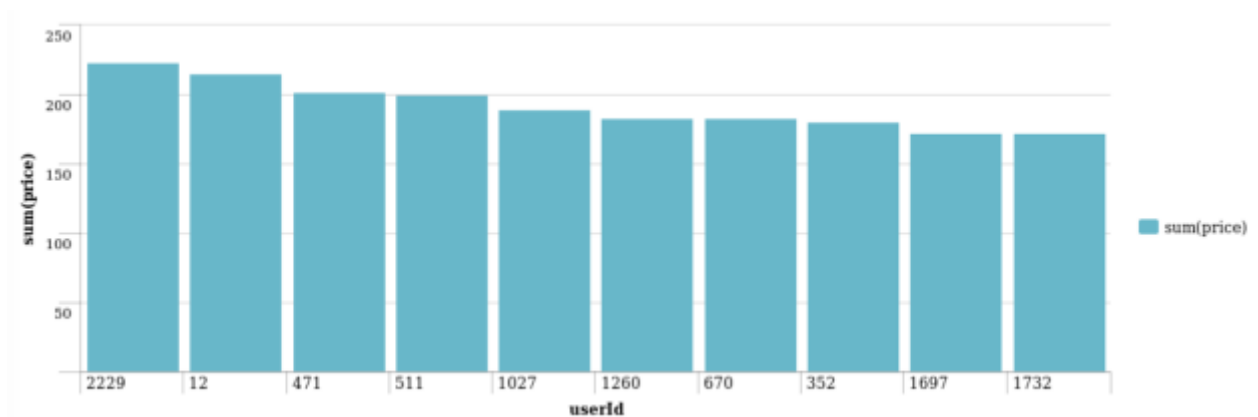|	
Amount spent buying items	$ 21407
Number of unique items available to be purchased	6

A histogram showing how many times each item is purchased:
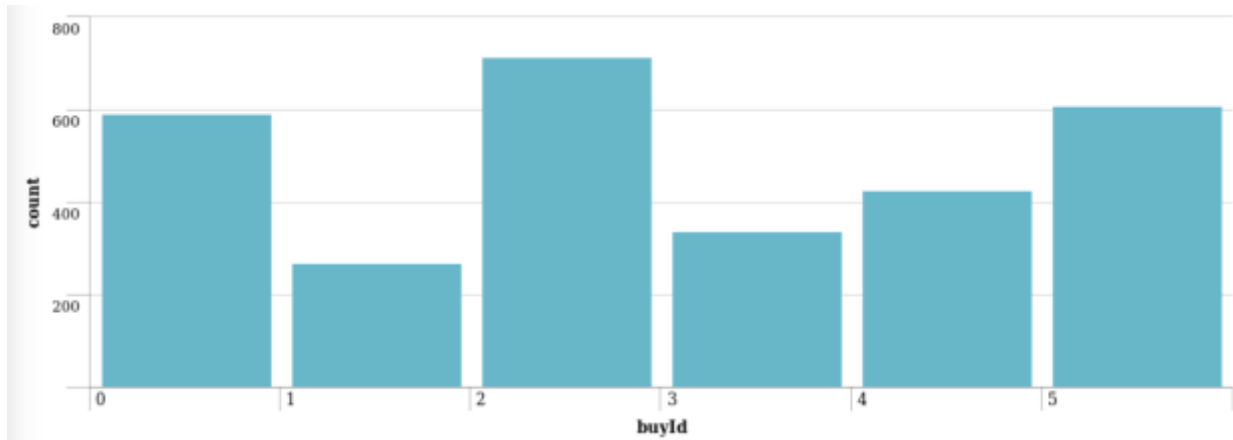
A histogram showing how much money was made from each item:
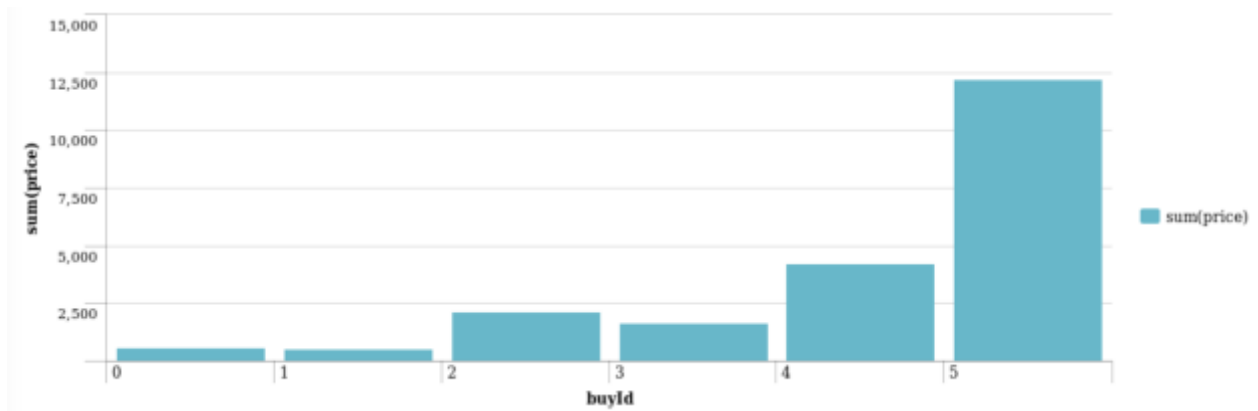
Filtering

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	11.5%
2	12	iPhone	13%
3	471	iPhone	14.5%