Data Exploration

Data Set Overview

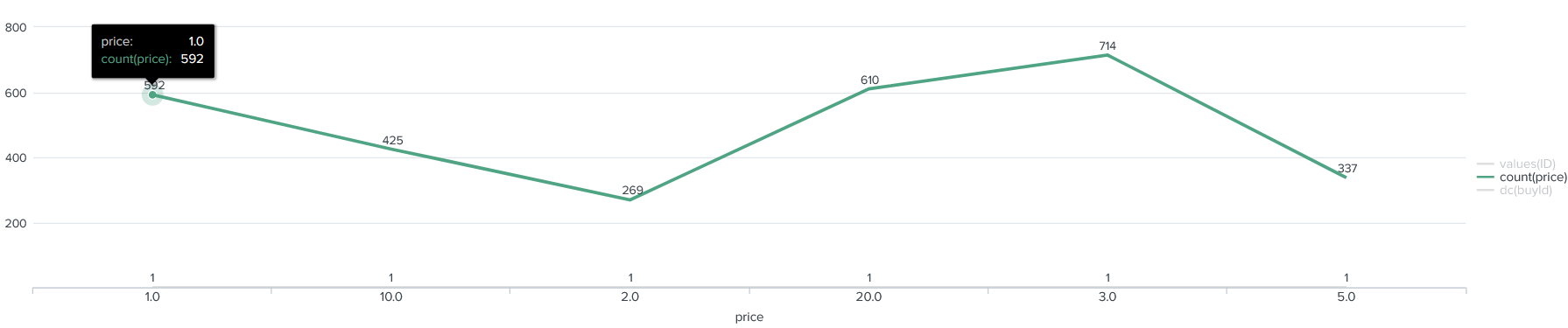
The table below lists each of the files available for analysis with a short description of what is found in each one.

|  |  |  |
| --- | --- | --- |
| **File Name** | **Description** | **Fields** |
| **ad-clicks.csv** | Each line in the file represents when a player clicks on an ad in a Flamingo game. | -timestamp: when 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/type of ad clicked on |
| **buy-clicks.csv** | Each line in this file represents when a player makes an in-app purchase in a Flamingo game. | -timestamp: when the purchase was made.  -txId: a unique id 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 |
| **users.csv** | This file contains a line for each user who plays 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 handles of the user.  -dob: the date of birth of the user.  -country: the two-letter country code where the user lives. |
| **team.csv** | This file contains a line for each team terminated 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 |
| **team-assignments.csv** | Each line in this file represents each time a user joins a team. The user can be in at most one team at a time. | -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 |
| **level-events.csv** | Each line in this file represents each time a team starts or ends with a level in the game | -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 |
| **user-session.csv** | 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. | -timestamp: a timestamp denoting when the event occurred.  -userSessionId: a unique id for the session.  -userId: the current user's ID.  -teamId: the current user's team.  -assignmentId: the team assignment id for the user to the team.  -sessionType: whether the event is the start or end of a session.  -teamLevel: the level of the team during this session.  -platformType: the type of platform of the user during this session. |
| **game-clicks.csv** | Each line in this file represents each time the user makes a click in 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. |

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:

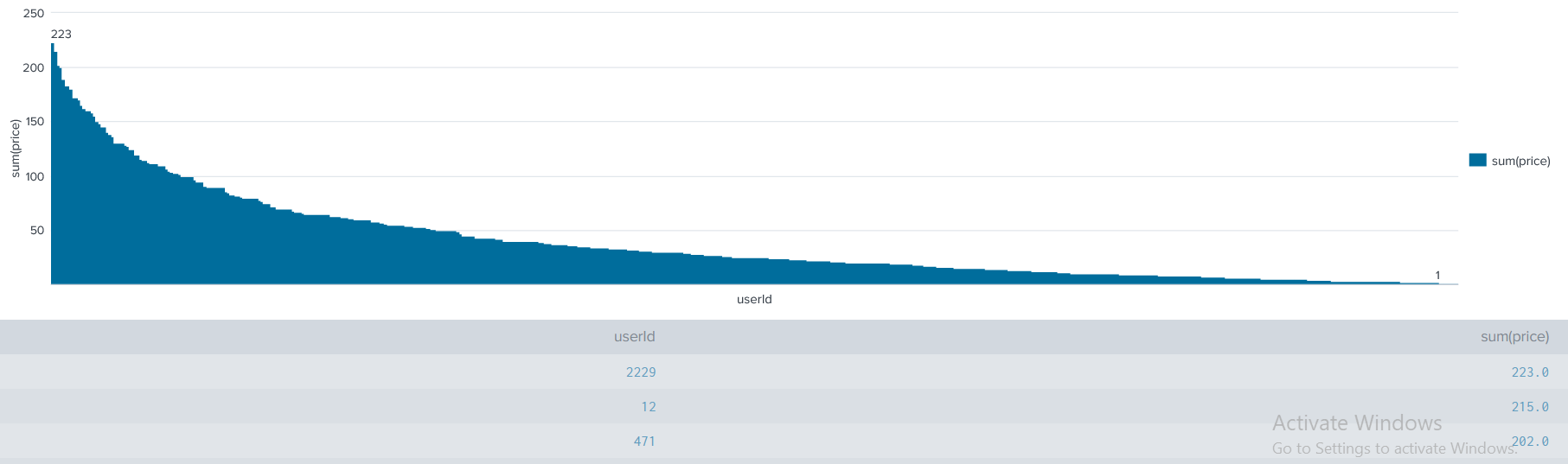


A histogram showing how much money was made from each item:



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 | 11.6% |
| 2 | 12 | iPhone | 13.07% |
| 3 | 471 | iPhone | 14.5% |