Data Exploration

Data Set Overview

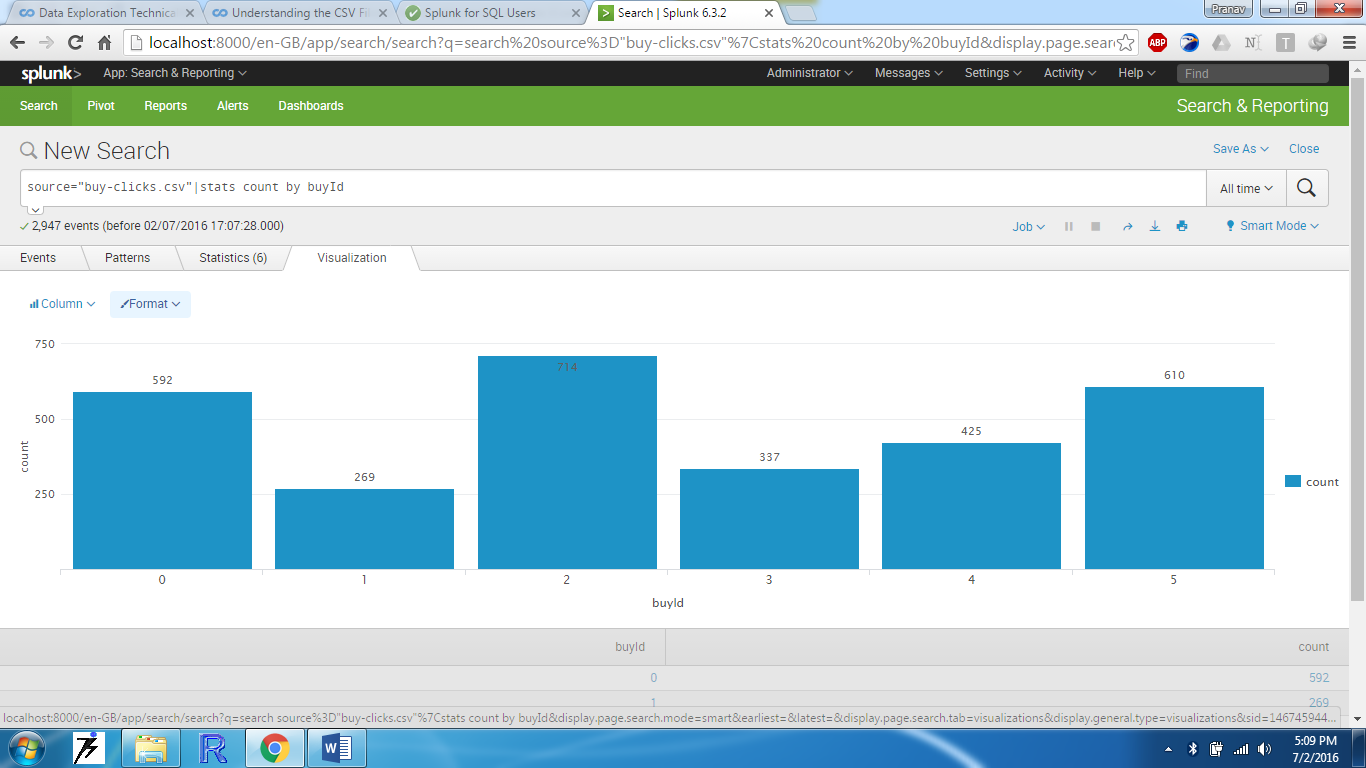
The table below lists each of the files available for analysis with a short description of what is found in each one.

|  |  |  |
| --- | --- | --- |
| **File Name** | **Description** | **Fields** |
| ad-clicks.csv | A line added to this file whenever an ad is clicked. | adId: Id of clicked ad  adCategory: Type of ad clicked on  teamId: current team id of user  timestamp:when the click occurred  txId: Unique id for each click  userId: Id of user  userSessionId: Id of user session in which user clicked |
| buy-clicks.csv | A line is added to this file on each app-purchase in Flamingo app. | buyId: Id of purchased item  price: Cost of item purchased  timestamp: time of click  team: current team id of user who made purchase  txId: Unique id for each click  userId: Id of user  userSessionId: Id of user session in which user clicked |
| users.csv | This file contains a line for each user in the game. | timestamp: when user first played the game.  id: 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. |
| team.csv | This file contains a line for each terminated 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 |
| team-assignments.csv | A line is added to the file each time a use joins the team. | time: 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 | A line is added to this file whenever a team starts or ends a new level. | time: when the event occurred.  eventid: a unique id for the event  teamid: the id of the team  level: the level started or completed  eventType: the type of event, either start or end |
| user-session.csv | Each line describes the user session of a player i.e. when session starts and ends. | 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 | A line is added to this file each time a user performs a click in the game. | time: 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 |

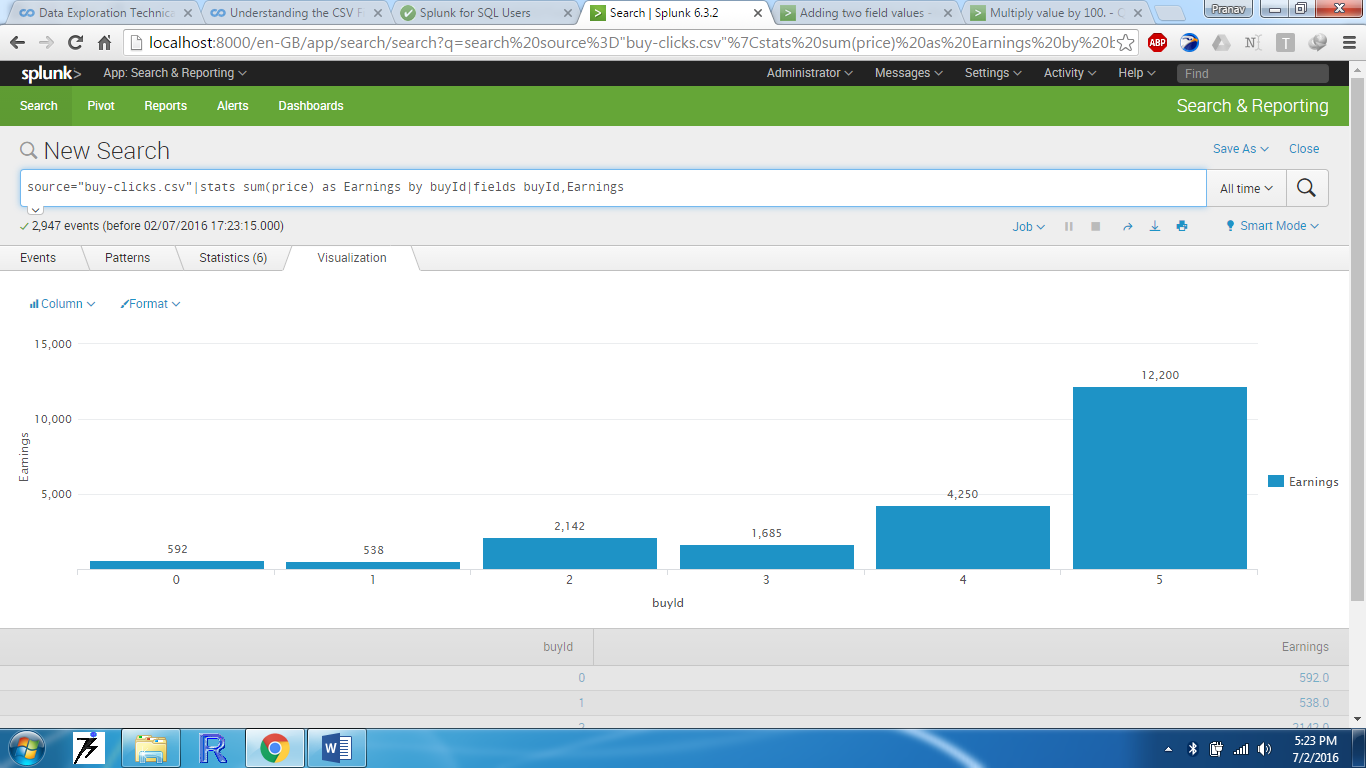
Aggregation

|  |  |
| --- | --- |
| Amount spent buying items | 21407.0 |
| # Unique items available to be purchased | 6 |

A histogram showing how many times each item is purchased:

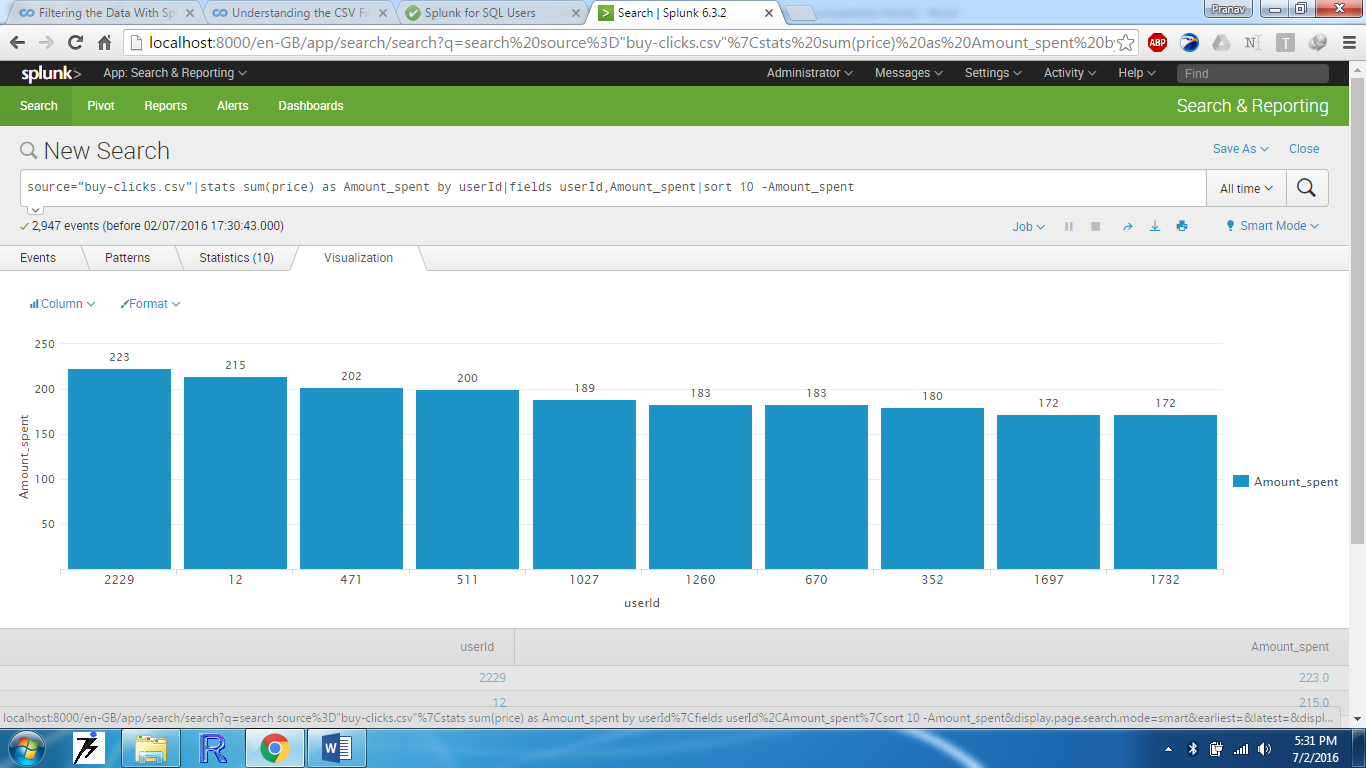


A histogram showing how much money was made from each item:



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 | 0.115970 |
| 2 | 12 | iphone | 0.130682 |
| 3 | 471 | iphone | 0.145038 |