Project 3

Ana Saavedra | Analiza Asuncion | Avantika Singh | Shilpa Nagendra | Gatlin Rash

**Story –**

We are exploring the global popularity of VIDEO GAMES. We will be exploration the trends of the gaming industry for the last 40 years with popularity of consoles and specific games. With the data provided we can visualize the data with the different attributes we have learned in class with an interactive dashboard and visuals.

**Links –**

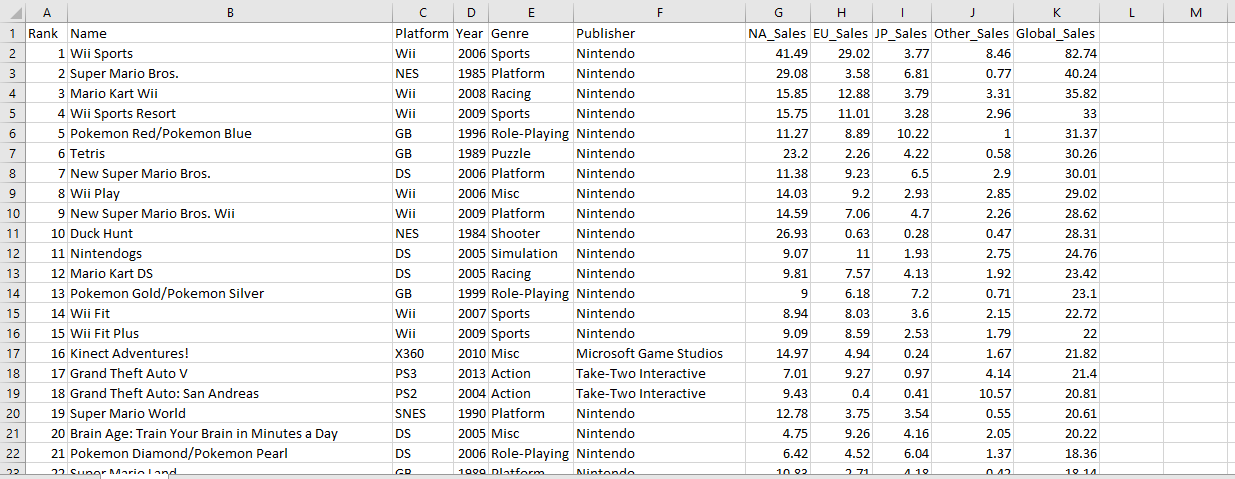
This is where we obtained our data.

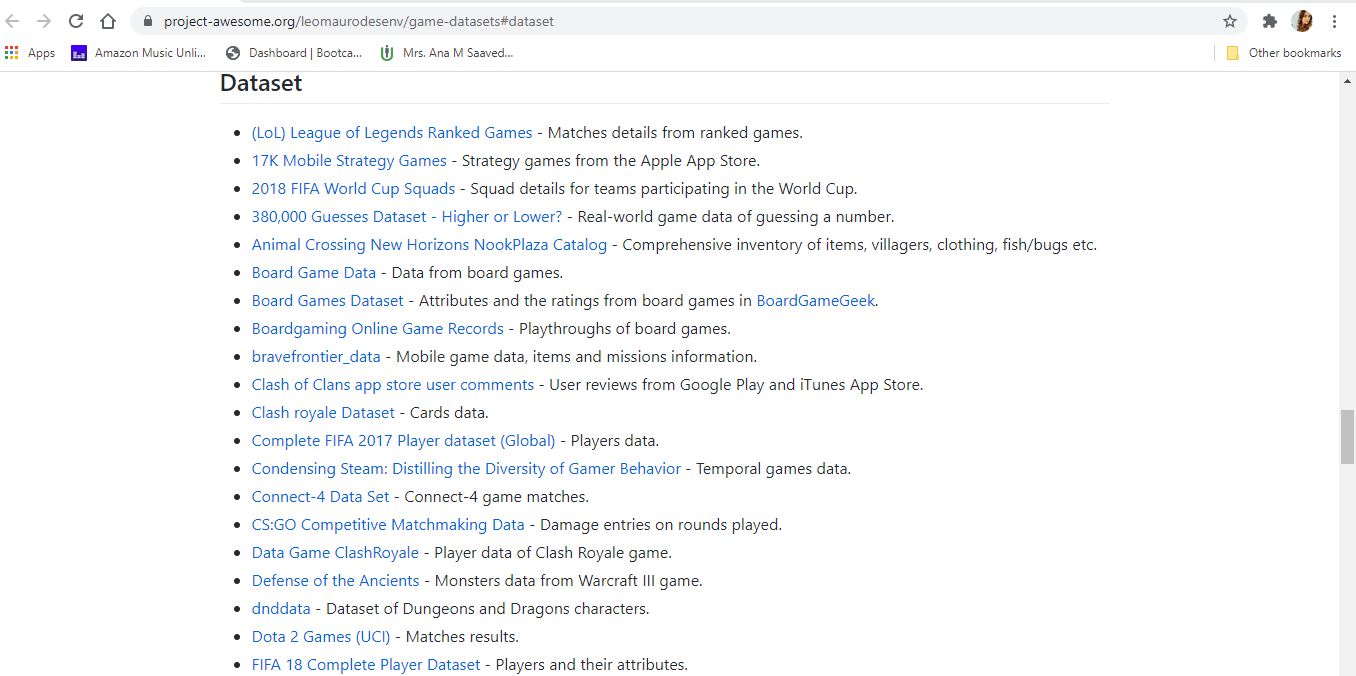
<https://www.kaggle.com/gregorut/videogamesales>

* Rank - Ranking of overall sales
* Name - The games name
* Platform - Platform of the games release (i.e. PC,PS4, etc.)
* Year - Year of the game's release
* Genre - Genre of the game
* Publisher - Publisher of the game
* NA\_Sales - Sales in North America (in millions)
* EU\_Sales - Sales in Europe (in millions)
* JP\_Sales - Sales in Japan (in millions)
* Other\_Sales - Sales in the rest of the world (in millions)
* Global\_Sales - Total worldwide sales.

<https://project-awesome.org/leomaurodesenv/game-datasets#dataset>

**Screenshots –**

The following is the snippet of the dataset we found to execute our project. ****

****

**AWS and Cloud Services –**

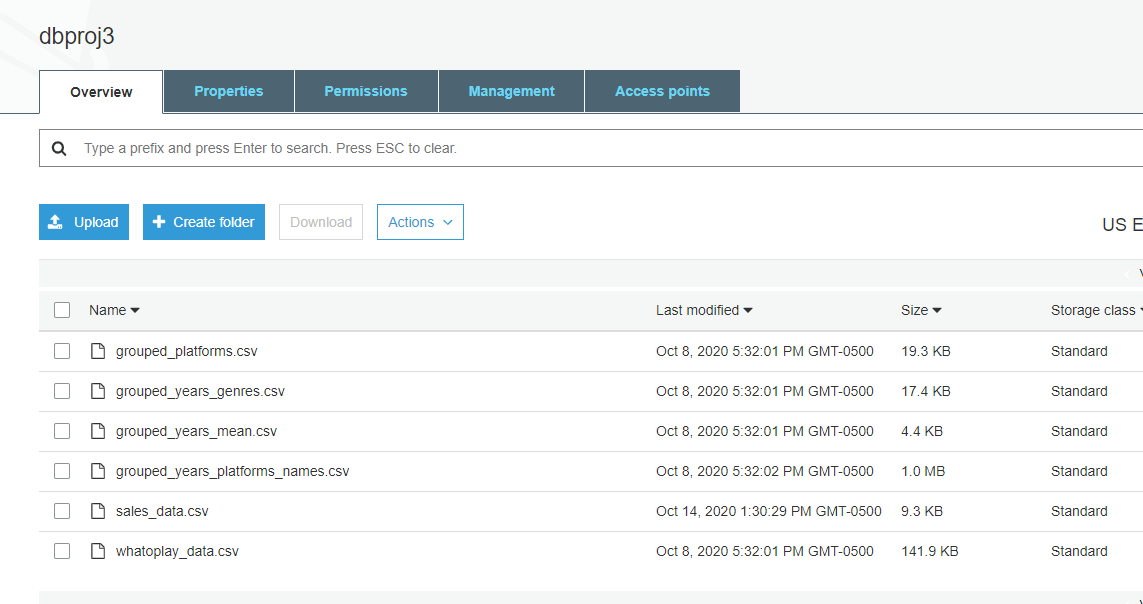
This section illustrates how we utilized a few services available in AWS:

* S3 Bucket
* RDS
* EC2
* Quicksight

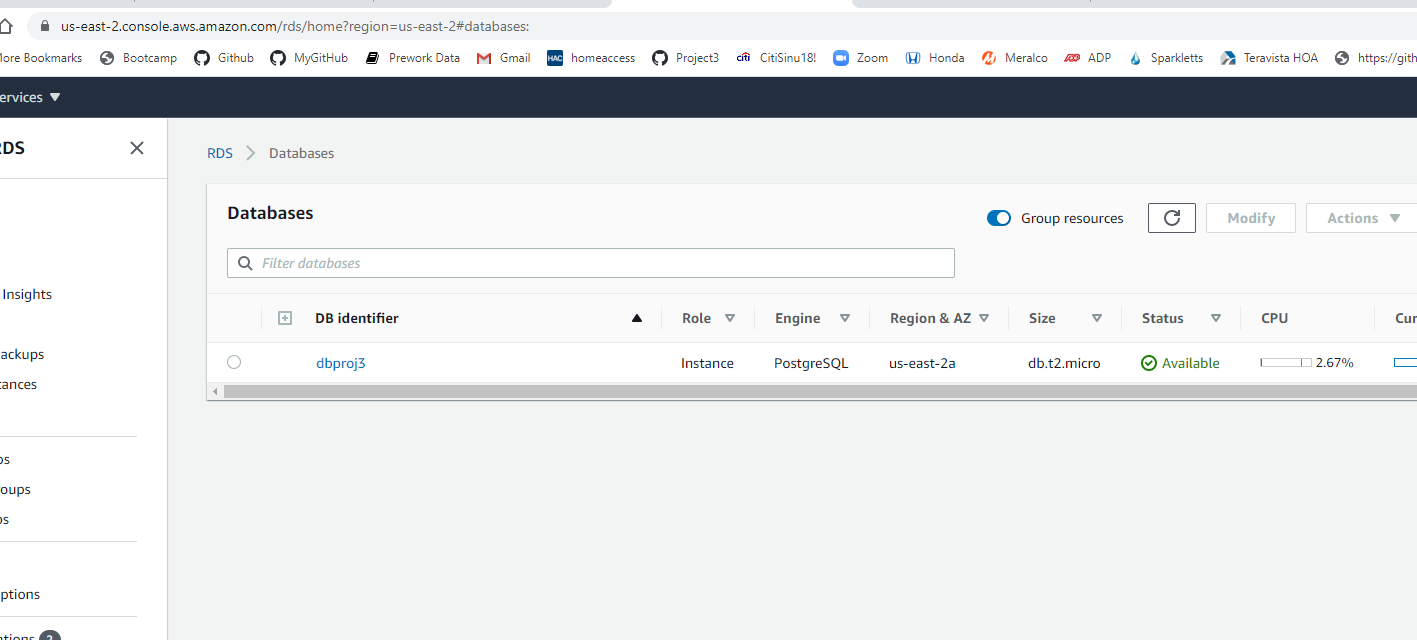
Other tools used:

* 5. Google Collaborator
* 6. Python
* 7. Postgres

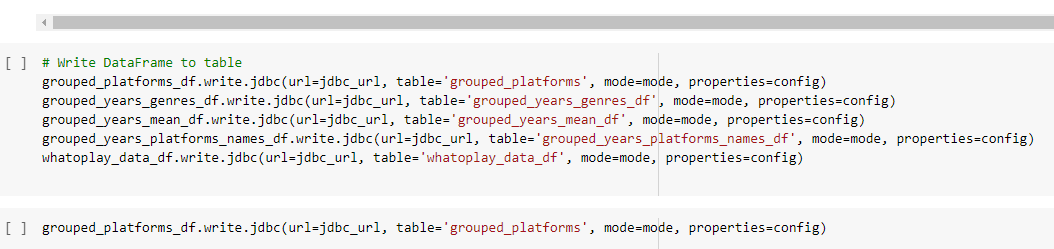
dbproj3 bucket with CSV files:



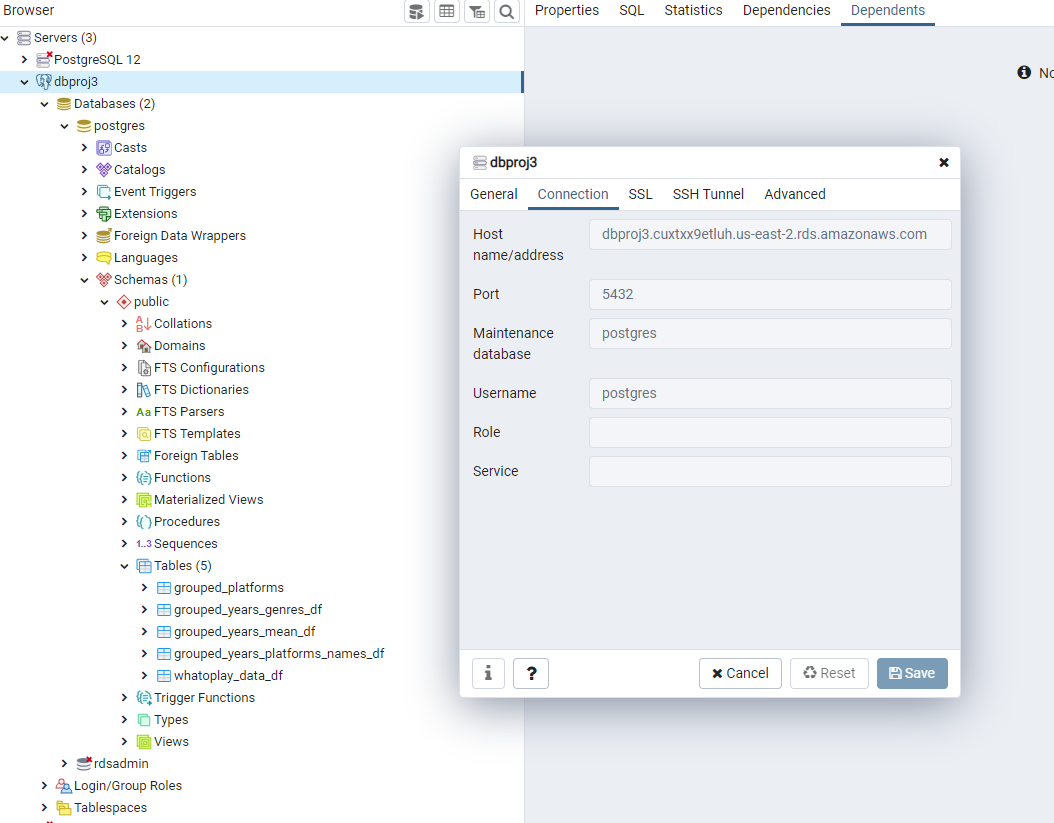
RDS database:



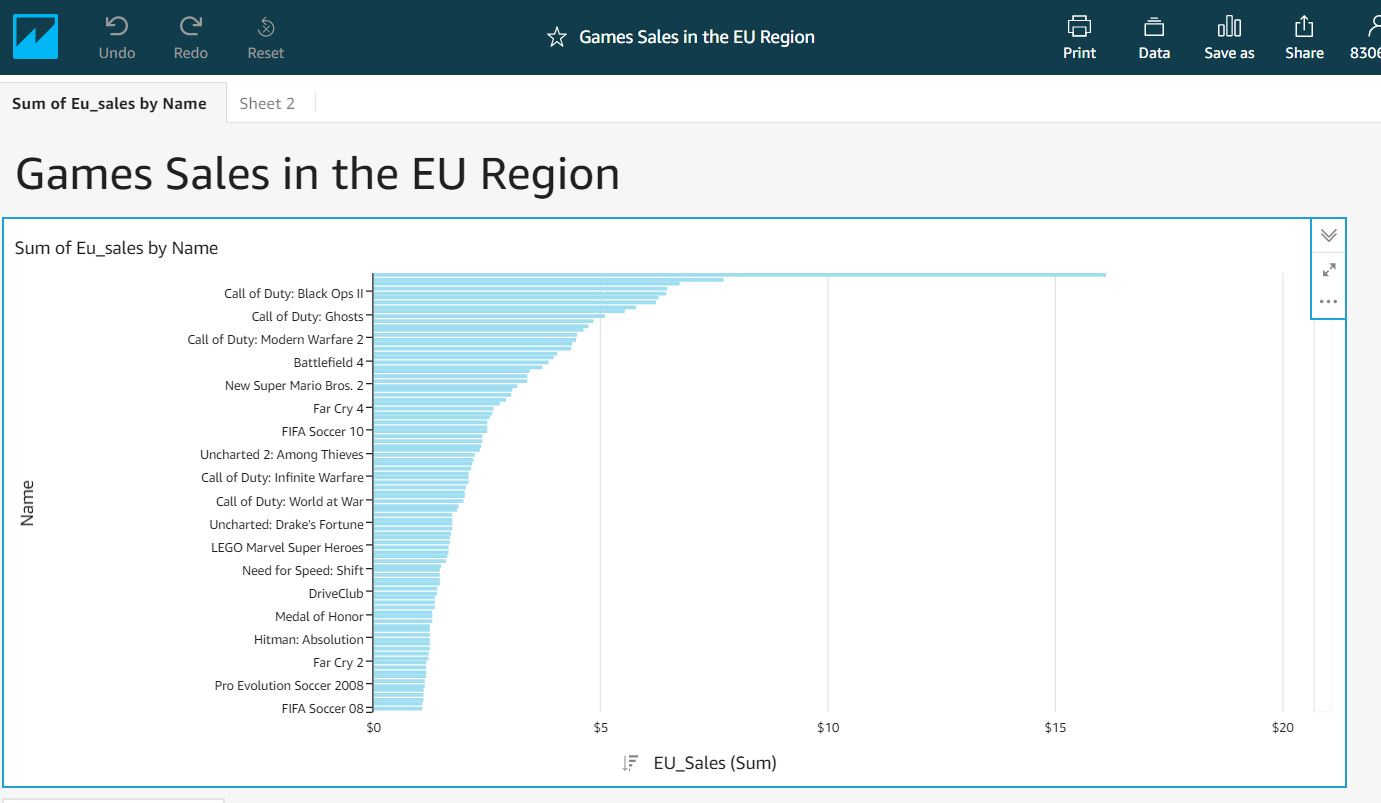
Python file used to write the data frames to Postgres using Google Collaborator:



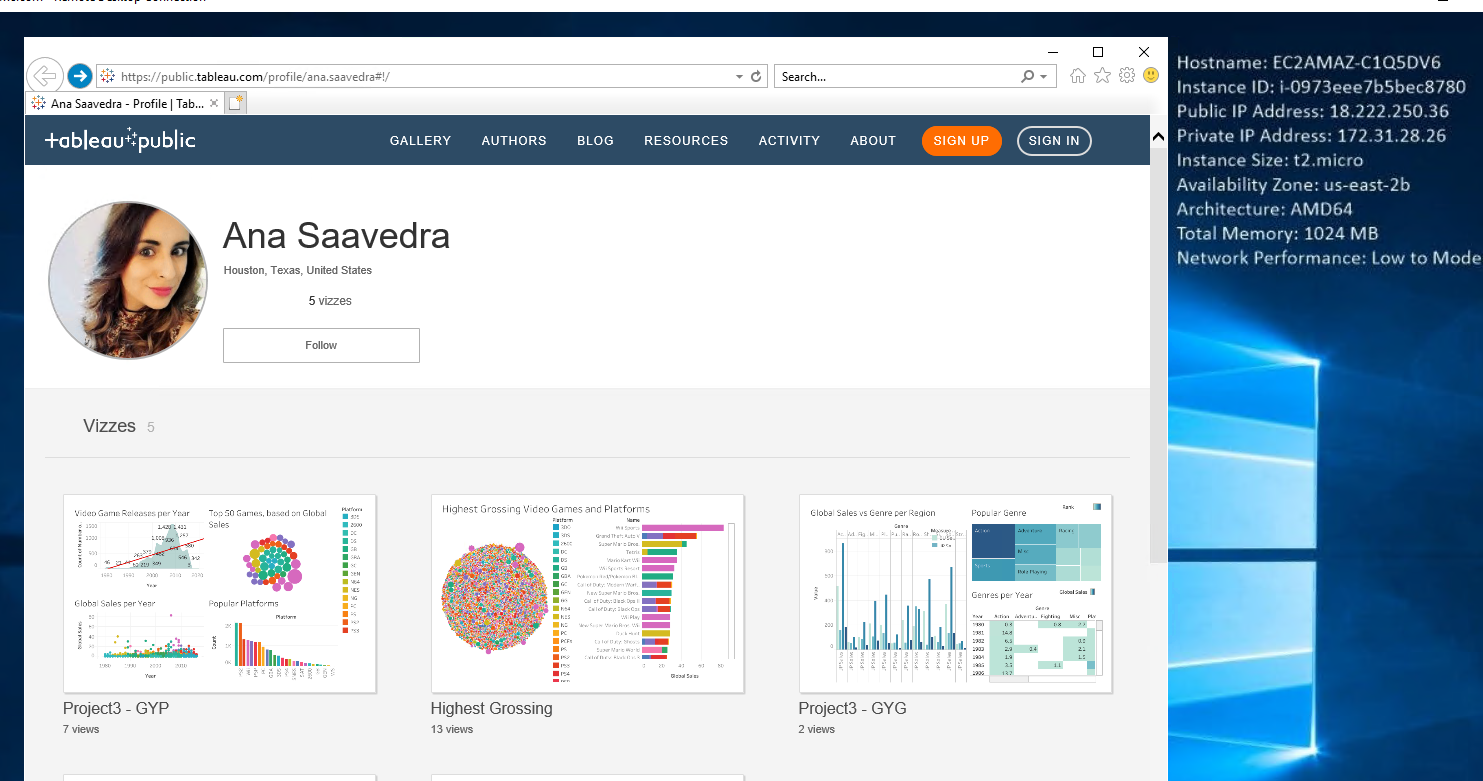
RDS dbproj3 and Postgres SQL Query Editor:



Quicksight Analytics using RDS as a data source:



EC2 showing Tableau graphs:



Challenge: No Admin rights to upgrade the IE browser or install Google Chrome. Unable to access the Quicksight website. Very slow.

**Tableau Visualizations and Dashboards –**

**Machine Learning –**

The Video Game data had three scores in the form of rating score which were Play Score, Game Score and Critic Score respectively. To describe each of them, the Play Score is the score obtained from the Player Ratings on various games, Game Score is the rating given to each game by the developer of the game and the Critic Score is obtained from the critic rating for the games. To study the relevance of these scores, we ran some Machine Learning Algorithms on our data (the analysis was done on North America Sales since it is the biggest market in terms of sales) and following are the conclusions that can be drawn.

**Final design –**

**GitHub Repository –**

This is the link to our GitHub repository.

https://github.com/Anabn357/dbproj3