TASK 1 Amazon prime

Dataset description : Amazon prime is a popular streaming service that offers a vast catalog of movies, TV shows, and original content. The data consist of contents added to Prime from 2008 to 2021. The oldest content is as old as 1925 and the newest is 2021

There are 12 variables and 8807 instance in the data set

Data type : Most of the data type is object and integer

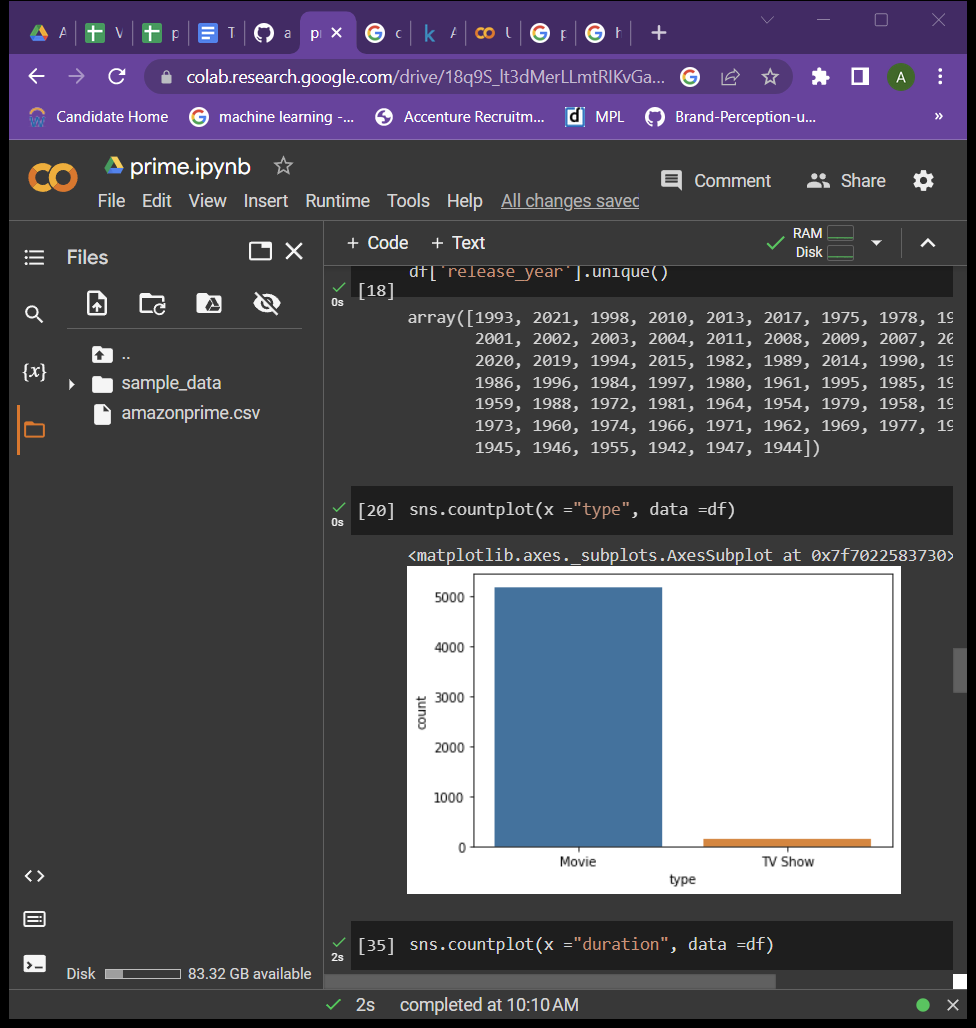
MISSING VALUES:

There are missing values in the data set (director- 2634, cast – 823, country – 831, date\_added-10)

#missing values has been removed from the dataset

Exploratory data Analysisi:

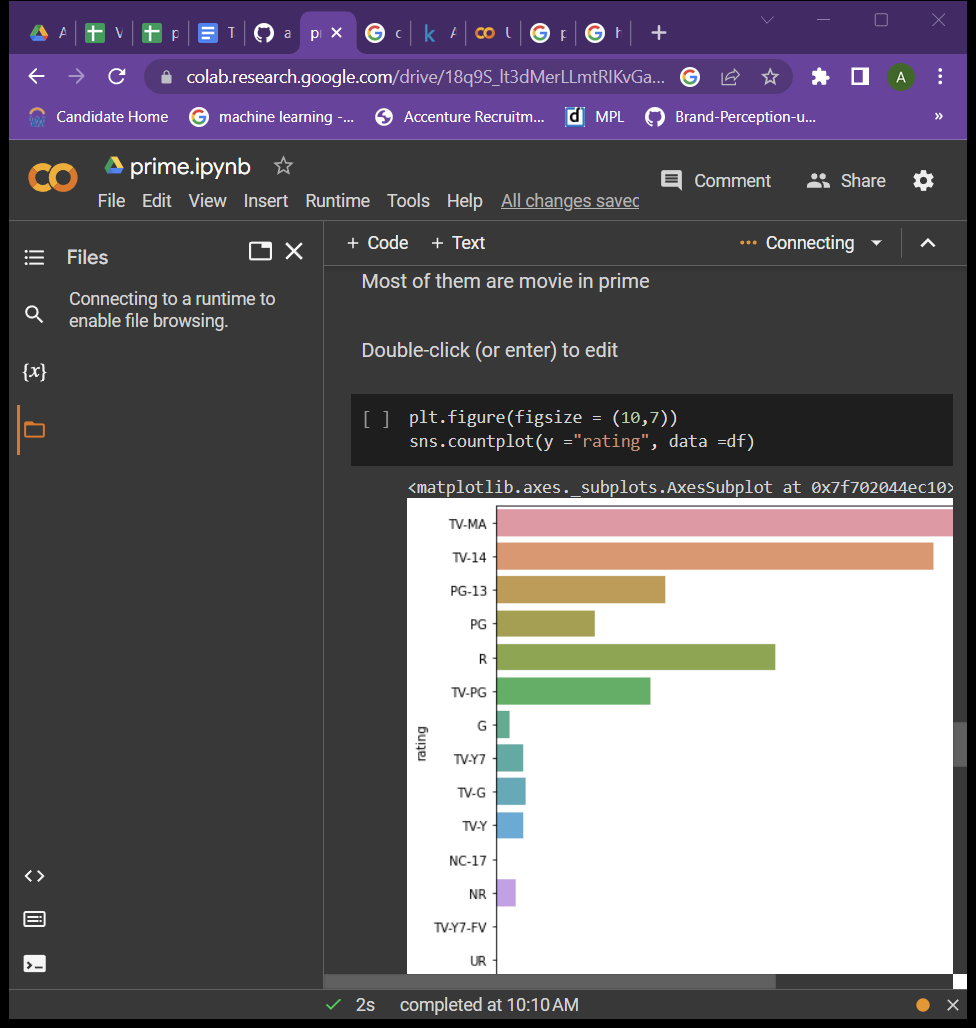
* Countplot:



In amazon prime most of the streaming shows are movies rather than tv shows

Duration : highest duration of the show is 140 minutes

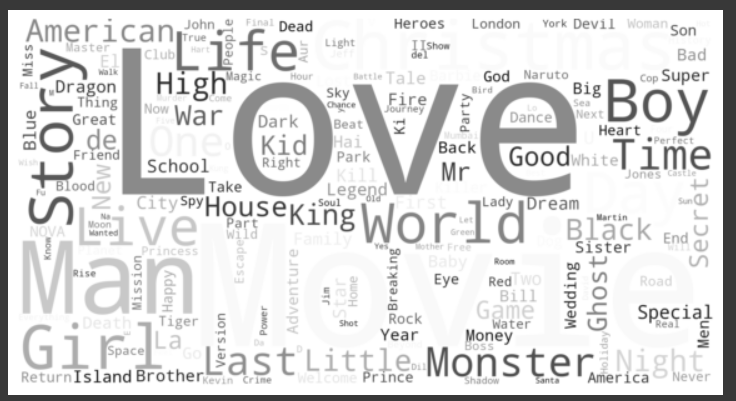
Rating :



Based on the rating we find that Most of the shows are for mature audience

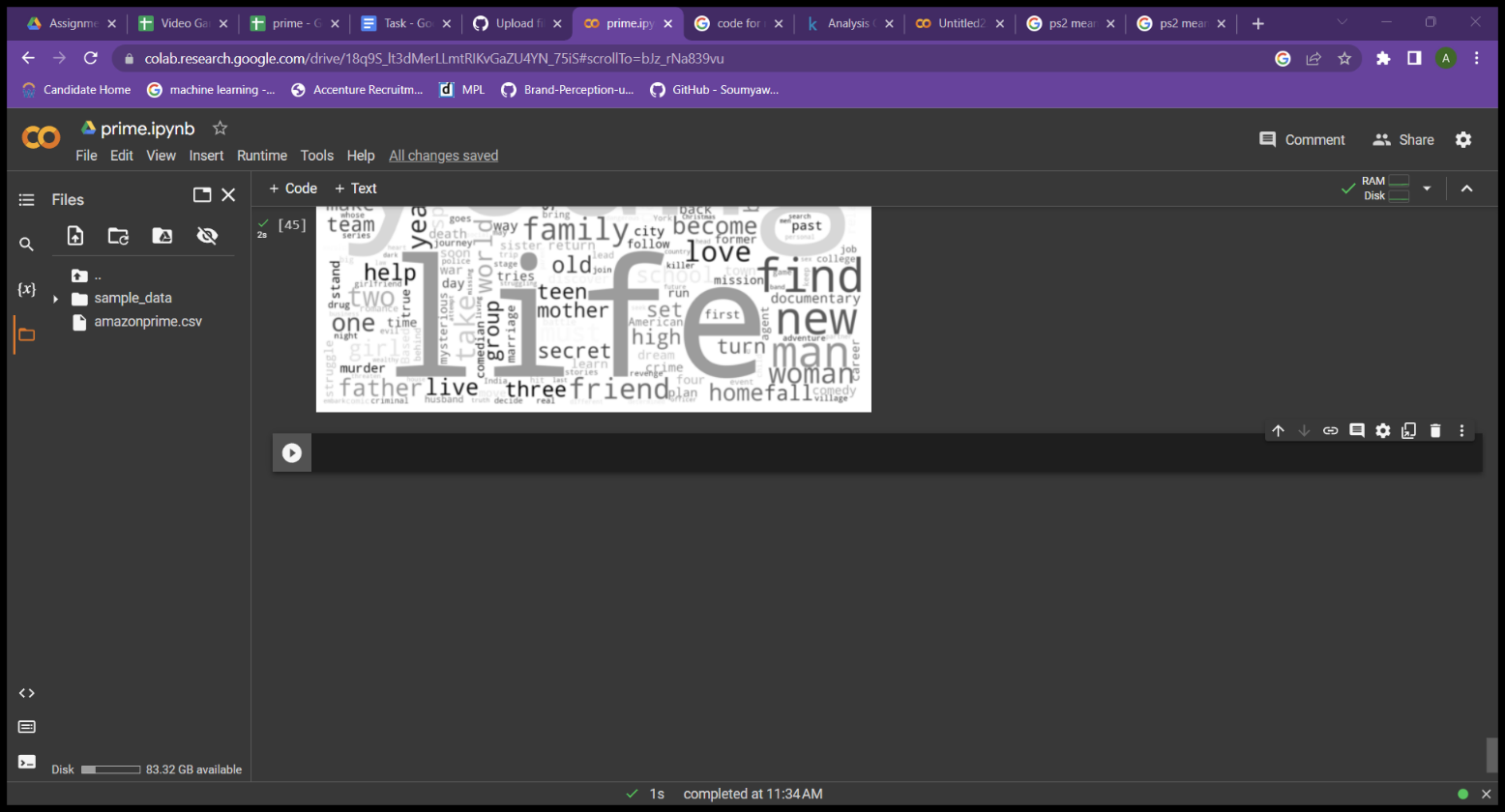
WORD COUNT:

Title:



Most repeated word in title is LOVE

WORD COUNT FOR DISCRIPTION:



Most repeated word in Discription is Life

Insights and Inference:

* In amazon prime most of the streaming shows are movies rather than tv shows
* Highest duration of the show is 140 minutes
* Based on the rating we find that Most of the shows are for mature audience

So kids safe browser should be used.

**TASK 2 THE VIDEO GAMING SALES**

DATASET DISCRIPTION :

The video games sales (see attached) dataset contains sales data for video games from all around the world, across different platforms, genres, and regions. Armed with this data and analysis, future developers can better understand what types of gameplay and mechanics resonate more with players to create a new gaming experience

There are 1907 variables and 13 instances

DATA TYPE : There are

* Float
* Object
* Int

MISSING VALUES:

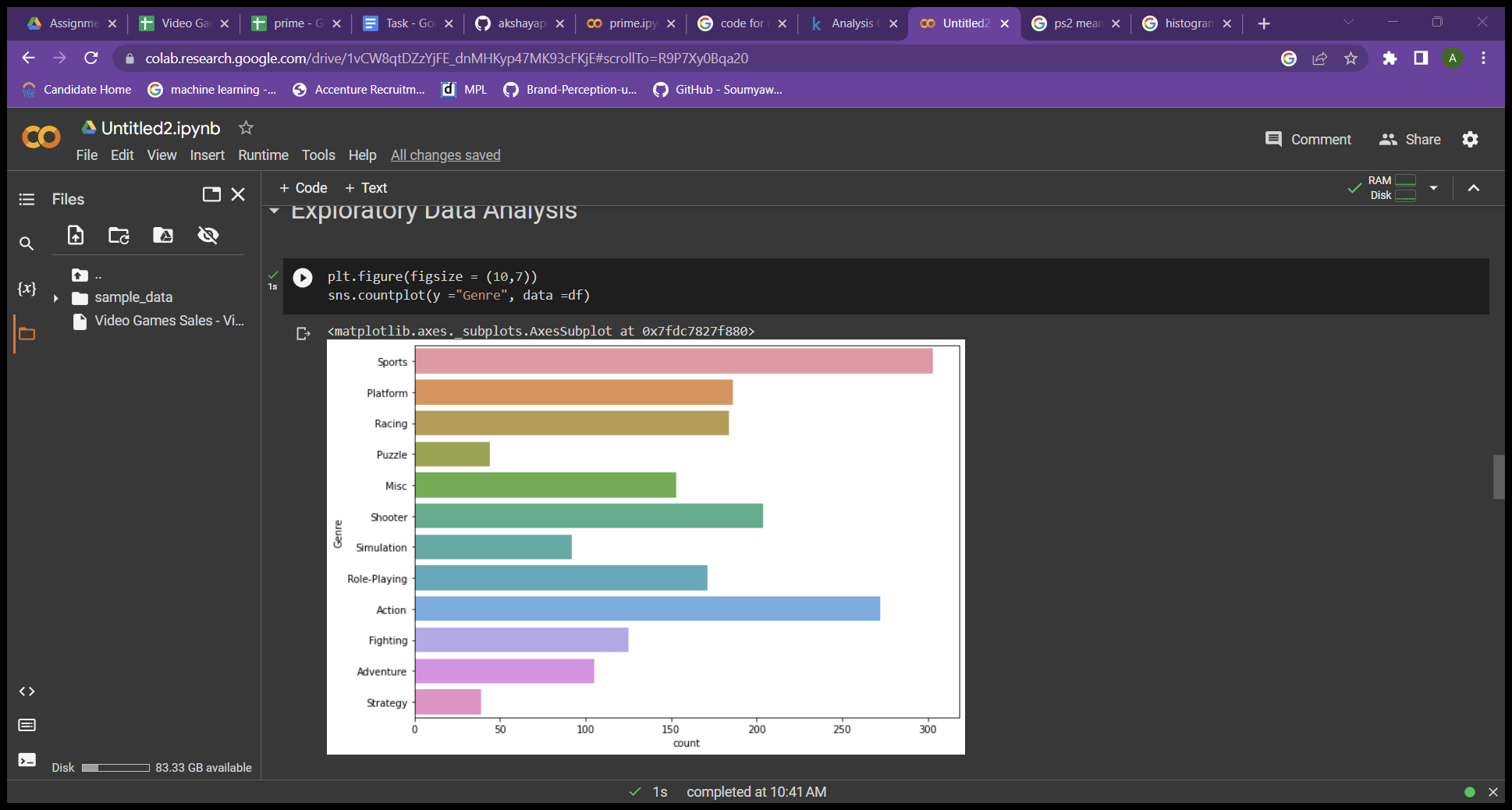
There is missing value in year attribute and publisher

#missing value has been removed

EXPLORATORY DATA ANALYSIS

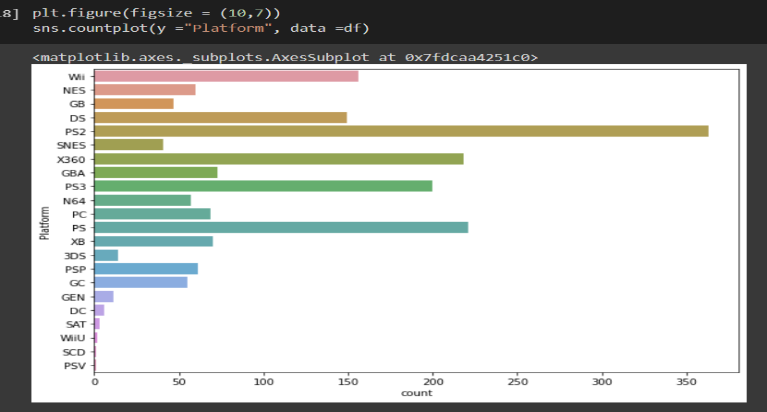
COUNT PLOT:

* GERNER:



From countplot we understand that most of the people like sport garner followed by Action

* PLATFORM :



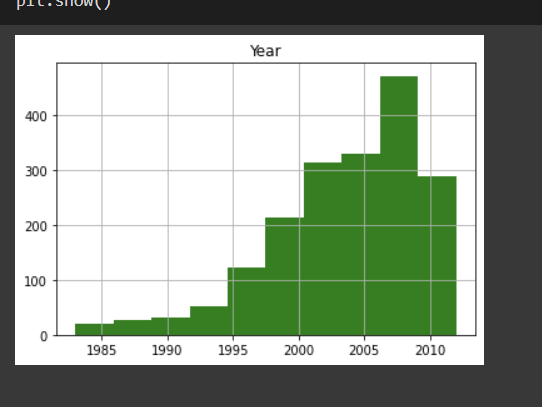
Most of the people like to play in (ps2) platform followed by ps

* PUBLISHER

Most of the games are published by Electronic arts

HISTOGRAM

* YEAR



Most of the games are published in the year between( 2005 to 2010)

INSIGHT AND INFERENCE:

* Most popular Genre is Sports And 2nd most popular gerner is Action
* Most of the people use play station to play games and publisher of most of the video games are Electronic arts
* Most of the games are lanched between 2005 to 2010 , Histogram is right skewed
* Most of the people like to play game in ps2 and ps

So new games can be created in Sports and Action so that people will like