

MVP of Video Game Sales

The Goal of MVP make sure the data is clean and display without error .

In the first I import the library pandas to read the dataset, matplotlib to visualize data, and clean the text after that I read the data , Show first five data frame.

Explore the data:

The remaining are :

1) Name

2) Year

3) Genre

4) Publisher

5) North America Sales

6) European Union Sales

7) Japanese Sales

8) Other Sales

9) Total Global Sales

Setp 1: Load Dataset

The dataset I will be use [Kaggle](#)

Setp 2: Clean Dataset

To Call specified columns

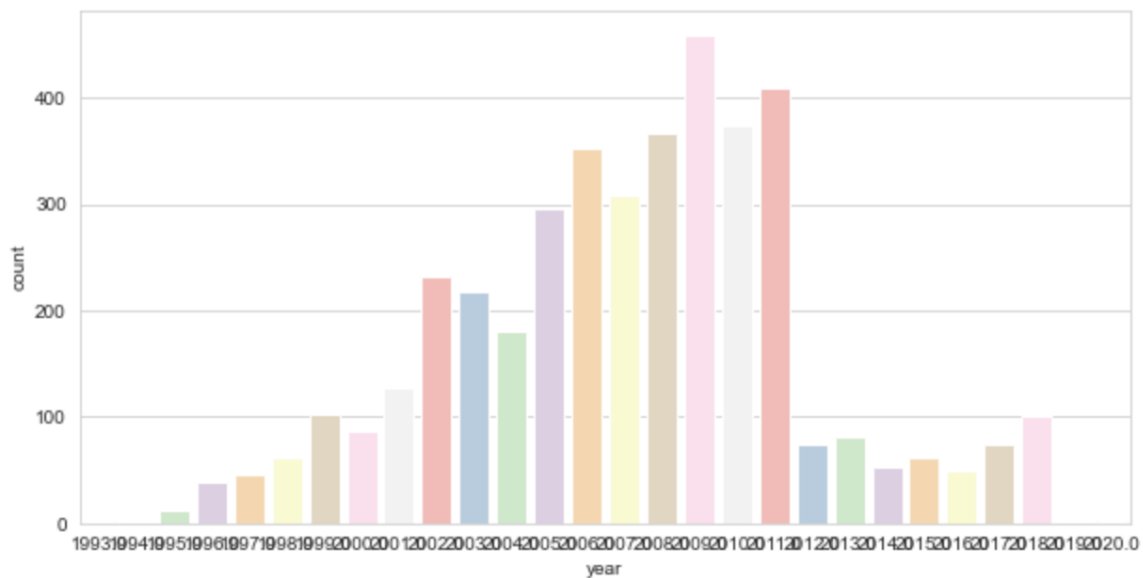
```
df.columns=['genre', 'esrb_rating', 'platform', 'publisher', 'developer', 'critic_score', 'global_sales', 'year']
```

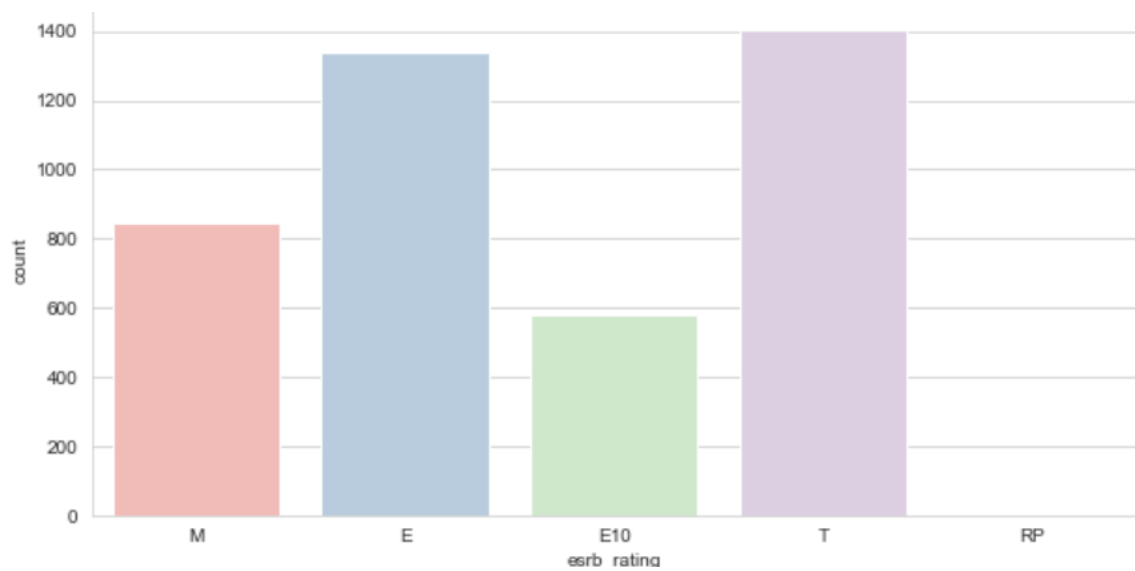
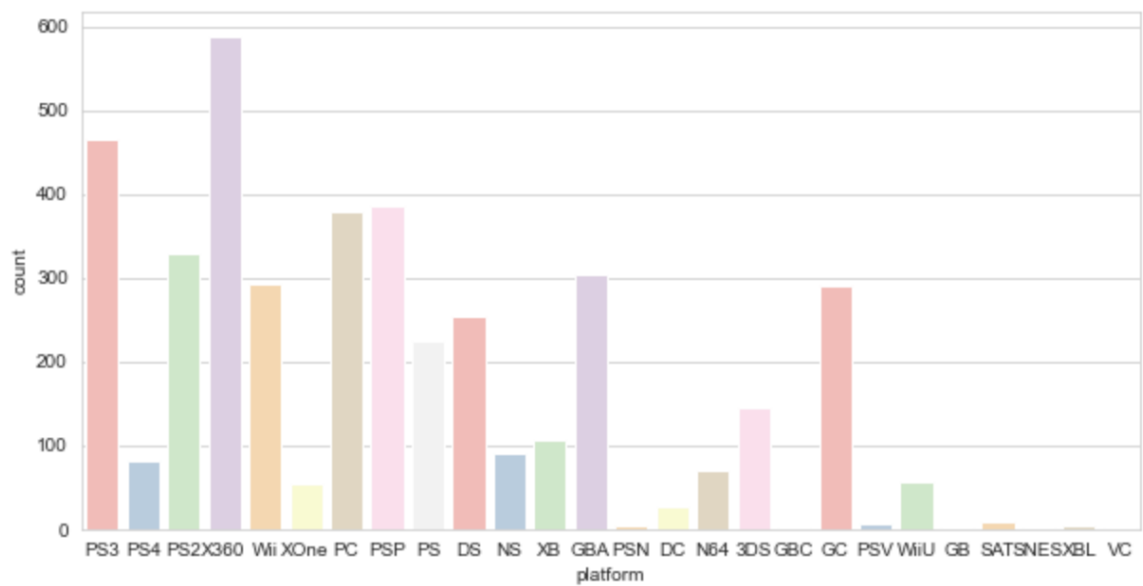
To Drop specified columns

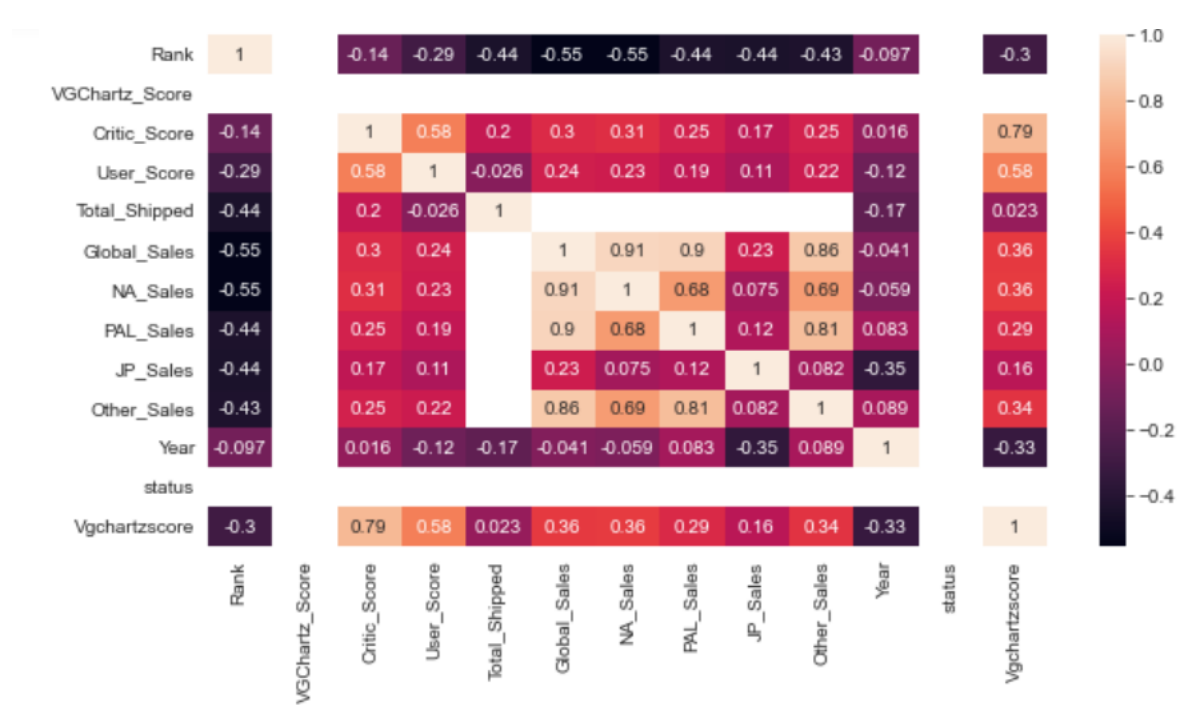
```
df2.drop( columns = ['img_url', 'NA_Sales', 'PAL_Sales', 'JP_Sales', ....], inplace=True )
```

Step 3 :EDA

The Target Column will be global_sales







Step 4 :Model

Frist I use sklearn model:

1.The linear model is LinearRegression()

2.The metrics is 35.137807122953326
4390685636713176.0

3. The ensemble is RandomForestRegressor()