EXPLORATORY DATA ANALYSIS REPORT

(VIDEO GAME SALES DATA)

**CONTENTS:**

* **Introduction**: Our understanding of the video game sales industry
* Initial hypotheses
* Relevant questions for the analysis
* Detailed process description, using visualization

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

* **THE VIDEO GAME INDUSTRY**

The video game industry according to Wikipedia.org is the economic sector involved in the development, marketing and monetization of video games. It encompasses dozens of job discipline and its component parts employ thousands of people worldwide. As of 2018, the video games industry generated sales of US$134.90 billion annually worldwide. The video games industry has since its inception seen a significant growth in sales as the years goes by as its popularity grows bigger.

it has numerous genres namely sports, racing, action games, adventure etc and many publishers in the industry, some of which include Nintendo, Ubisoft, Sony, sega etc.

* **HYPOTHESIS**

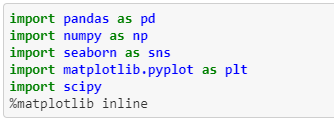
There has been an increase in video game global sales in recent years with the Play Station Platform and North American region taking the highest share.

* **QUESTIONS**

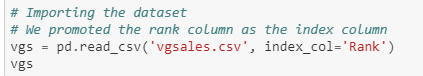
1. What is the trend of top 5 global video game sales over the years?
2. Which years had the highest and lowest video game sales?
3. Which regions have the highest share of global video game sales?
4. What is the trend of video game sales in North America over the years?
5. Which platforms have the highest share of global video game sales?

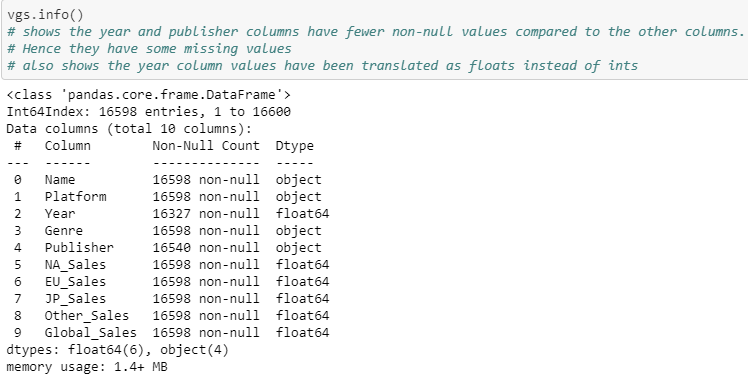
* **DETAILED PROCESS DESCRIPTION**
* First of all, we imported all of the libraries that we would need for the analysis and conducted a preliminary analysis as shown below:

**UNDERSTANDING THE DATASET**

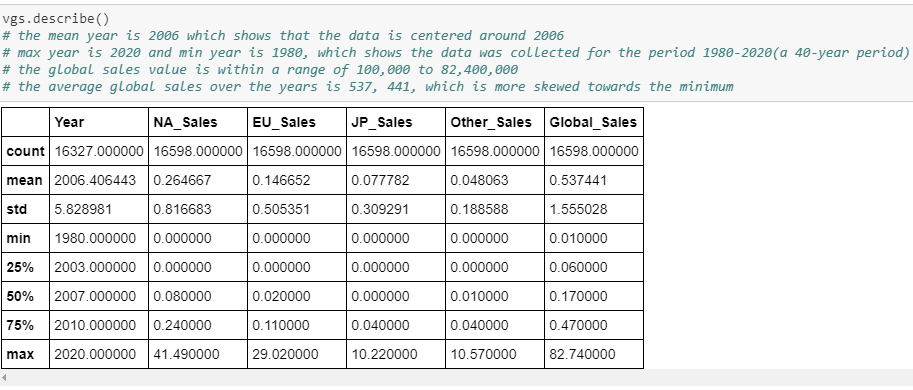


* Next, the dataset was loaded into python and inspection of all columns (variables) done:



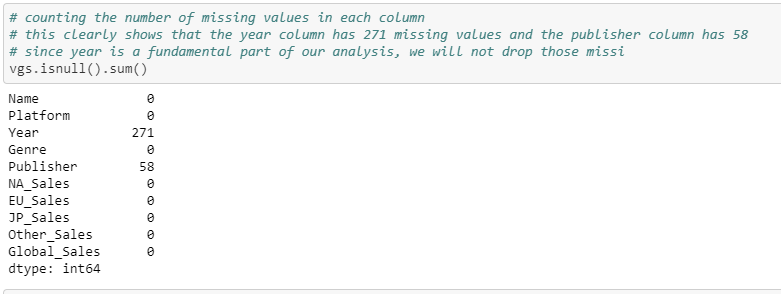


* And then a basic description of the whole datasets was carried out.



**CLEANING THE DATASET**

* After thorough inspection, and further exploration missing values were identified;



* Upon identifying these missing values, steps were taken to refill these missing values for the **year** column and the **publisher** column as indicated above as they play a major role in our analysis. (Details can be seen in the jupyter notebook attached)



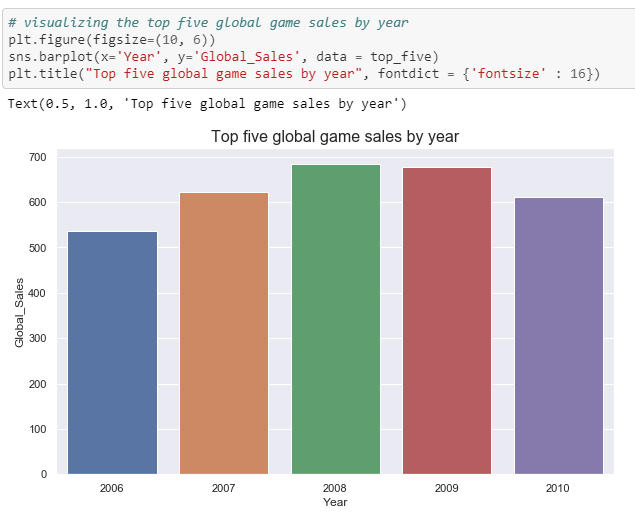
**ANALYZING RELATIONSHIP BETWEEN DATASET**

**(DATA VISUALIZATION)**

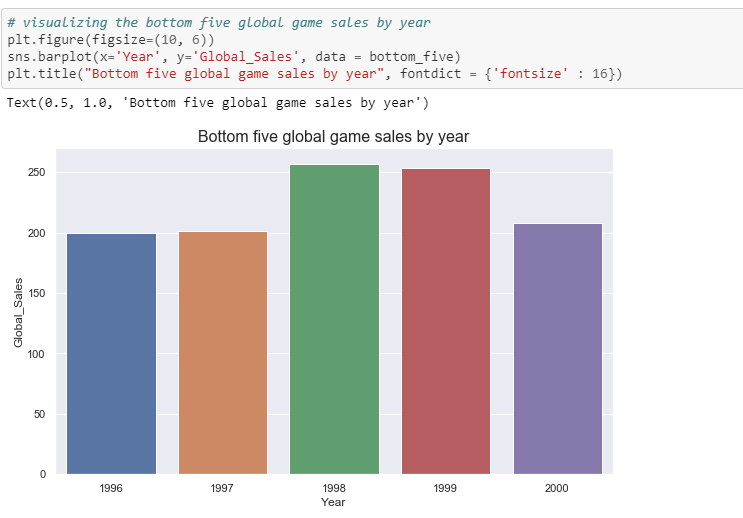
* Data Visualization with the help of the Matplotlib imported in python was used to help answer the relevant research questions raised for analysis. Diagrams such as the histogram, pie chart, scatter plot, bar plots etc were utilized.

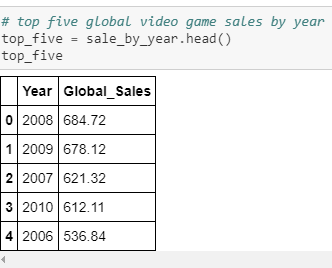
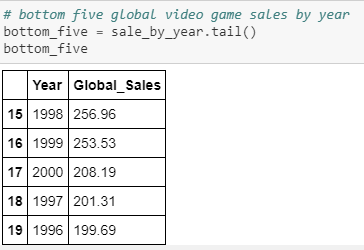
Below are diagrams/charts generated in python to answer questions that were raised:

1. **What is the trend of top 5 global video game sales over the years?**



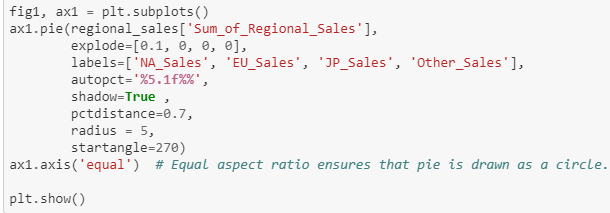
1. **Which years had the highest and lowest video game sales?**

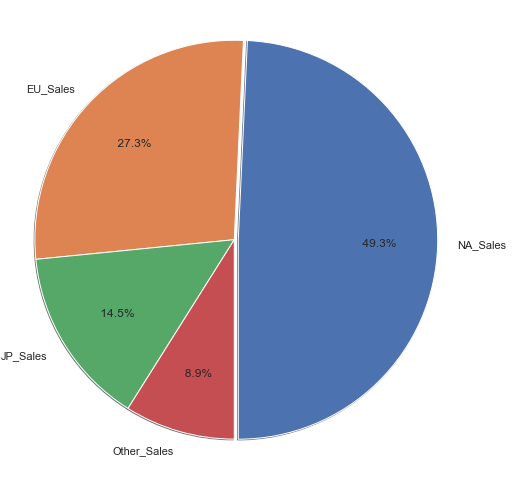


From the diagrams and tables above, it can be inferred that, the year **2008** recorded the **highest** global sales whiles the 1996 recorded the lowest global sales.

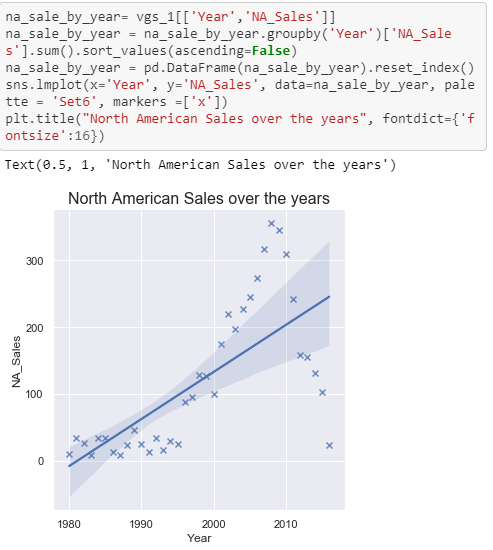
1. **Which regions have the highest share of global video game sales?**





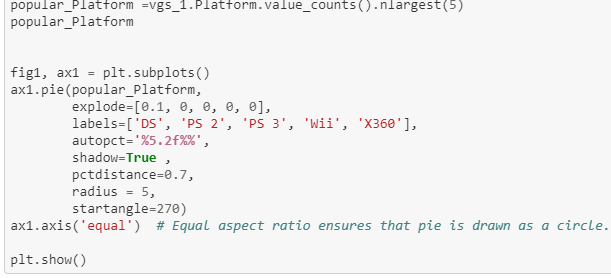
The pie chart above shows that North American sales have the highest share of global sales. This covers almost half (49.3%) of video games sold since 1980. Europe sales covers the second largest share, covering about 27.3% and the rest shared between Japan and other sales.

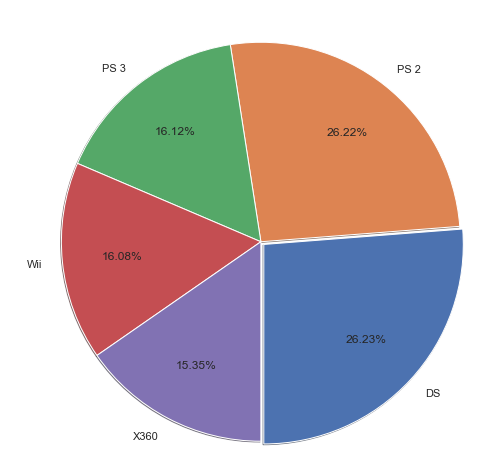
1. **What is the trend of video game sales in North America over the years?**



It can be inferred from the diagram above that North America sales increases slowly in the 1980s through to the middle of the 1990s. it then increases rapidly, reaches an apex in 2008 and begins to decline.

1. **Which platforms have the highest share of global video game sales?**





An initial look at the pie chart reveals that DS seem to have the highest share of global game sales (26.23%). However, the PlayStation platforms PS 2 and PS 3 have 16.12% and 26.22% of the global game sales shares. When combined, PlayStation has about 42.3% share of global video game sales. This is almost half of the entire sales worldwide. This is therefore a clear indication that **PlayStation platforms** are very popular among gamers.

* HYPOTHESIS RECHECK (CONCLUSION)

From our exploratory analysis, we can confidently accept the hypothesis raised from the onset of the analysis. This can be inferred from diagrams displayed in answering questions 3 and questions 5 where we have **North America** recording the highest of the global sales (Question 3) and **PlayStation** **platform** also having the highest share of the global video game sales.

Hence, our hypothesis raised that there has been an increase in video game global sales in recent years with the Play Station Platform and North American region taking the highest share can be accepted.

**GROUP MEMBERS**

1. Priscilla Baah
2. Raymond Bentum
3. Raymond Normeshie
4. Prince Adorkor