DATA ANALYSIS OF VIDEO GAME SALES

DATA IMPORTATION

In [4]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [5]:

```
pd.read_csv?
vgsales_df = pd.read_csv("vgsales.csv")
```

In [6]:

```
vgsales_df
```

Out[6]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_S
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	
...	...	...	...	...	...	...	...	...	
16593	16596	Woody Woodpecker in Crazy Castle 5	GBA	2002.0	Platform	Kemco	0.01	0.00	
16594	16597	Men in Black II: Alien Escape	GC	2003.0	Shooter	Infogrames	0.01	0.00	
16595	16598	SCORE International Baja 1000: The Official Game	PS2	2008.0	Racing	Activision	0.00	0.00	
16596	16599	Know How 2	DS	2010.0	Puzzle	7G//AMES	0.00	0.01	
16597	16600	Spirits & Spells	GBA	2003.0	Platform	Wanadoo	0.01	0.00	

16598 rows × 11 columns

DATA PREPARATION FOR VISUALIZATION

UNDERSTANDING DATATYPE

In [28]:

```
vgsales_df.dtypes
```

Out[28]:

```
Name             object
Platform         object
Year             object
Genre            object
Publisher        object
NA_Sales        float64
EU_Sales        float64
JP_Sales        float64
Other_Sales     float64
Global_Sales    float64
dtype: object
```

SETTING UP INDEX

In [7]:

```
vgsales_df.set_index('Rank', inplace=True)
```

DROPPING NULL VALUES

In [7]:

```
vgsales_df.dropna(how='any')
```

Out[7]:

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	C
1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	
2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	
3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	
4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	
5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	
...	...	...	...	...	...	...	...	...	
16596	Woody Woodpecker in Crazy Castle 5	GBA	2002.0	Platform	Kemco	0.01	0.00	0.00	
16597	Men in Black II: Alien Escape	GC	2003.0	Shooter	Infogrames	0.01	0.00	0.00	
16598	SCORE International Baja 1000: The Official Game	PS2	2008.0	Racing	Activision	0.00	0.00	0.00	
16599	Know How 2	DS	2010.0	Puzzle	7G//AMES	0.00	0.01	0.00	
16600	Spirits & Spells	GBA	2003.0	Platform	Wanadoo	0.01	0.00	0.00	

16291 rows × 10 columns

CHANGING THE DATE FORMAT

In [8]:

```
vgsales_df['Year'] = pd.to_datetime(vgsales_df['Year'], format = '%Y')
```

In [9]:

```
vgsales_df['Year'] = vgsales_df['Year'].dt.strftime('%Y')
```

In [10]:

```
vgsales_df
```

Out[10]:

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Ot
1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	
2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	
3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	
4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	
5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	
...	...	...	...	...	...	...	...	...	
16596	Woody Woodpecker in Crazy Castle 5	GBA	2002	Platform	Kemco	0.01	0.00	0.00	
16597	Men in Black II: Alien Escape	GC	2003	Shooter	Infogrames	0.01	0.00	0.00	
16598	SCORE International Baja 1000: The Official Game	PS2	2008	Racing	Activision	0.00	0.00	0.00	
16599	Know How 2	DS	2010	Puzzle	7G//AMES	0.00	0.01	0.00	
16600	Spirits & Spells	GBA	2003	Platform	Wanadoo	0.01	0.00	0.00	

16598 rows × 10 columns

NOW DATA IS READY FOR VISUALIZATION

THE VISUALIZATION WAS DONE IN POWER BI CAUSE MY SEABORN AND MAPLOTLIB ARE MALFUNCTIONING

In [13]:

```
vgsales_df.to_csv('C:\\Users\\J6IX\\Desktop\\vgsale.csv')
```

In [ ]:

```

```