

```
In [4]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

```
In [5]: data = pd.read_csv('vgsales.csv')
```

```
In [4]: data.head()
```

```
Out[4]:
```

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37

```
In [5]: data.shape
```

```
Out[5]: (16598, 11)
```

```
In [6]: data.isnull().sum()
```

```
Out[6]: Rank      0
Name      0
Platform    0
Year      271
Genre      0
Publisher   58
NA_Sales    0
EU_Sales    0
JP_Sales    0
Other_Sales 0
Global_Sales 0
dtype: int64
```

```
In [7]: data.dropna(inplace = True)
```

```
In [8]: data.shape
```

```
Out[8]: (16291, 11)
```

```
In [9]: data_decade = data[data.Year > 2010]
```

```
In [10]: data.corr()
```

```
Out[10]:
```

	Rank	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
Rank	1.000000	-0.178027	-0.400315	-0.370137	-0.360233	-0.332735	-0.436075

```
In [7]: data.dropna(inplace = True)
```

```
In [8]: data.shape
```

```
Out[8]: (16291, 11)
```

```
In [9]: data_decade = data[data.Year > 2010]
```

```
In [10]: data.corr()
```

```
Out[10]:
```

	Rank	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
Rank	1.000000	0.178027	-0.400315	-0.379137	-0.269323	-0.332735	-0.426975
Year	0.178027	1.000000	-0.091285	0.006108	-0.169387	0.041128	-0.074647
NA_Sales	-0.400315	-0.091285	1.000000	0.768923	0.451283	0.634518	0.941269
EU_Sales	-0.379137	0.006108	0.768923	1.000000	0.436379	0.726256	0.903264
JP_Sales	-0.269323	-0.169387	0.451283	0.436379	1.000000	0.290559	0.612774
Other_Sales	-0.332735	0.041128	0.634518	0.726256	0.290559	1.000000	0.747964
Global_Sales	-0.426975	-0.074647	0.941269	0.903264	0.612774	0.747964	1.000000

```
In [12]: data_decade.head()
```

```
Out[12]:
```

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
16	17	Grand Theft Auto V	PS3	2013.0	Action	Take-Two Interactive	7.01	9.27	0.97	4.14	21.40
23	24	Grand Theft Auto V	X360	2013.0	Action	Take-Two Interactive	9.63	5.31	0.06	1.38	16.38
29	30	Call of Duty: Modern Warfare 3	X360	2011.0	Shooter	Activision	9.03	4.28	0.13	1.32	14.76
32	33	Pokemon X/Pokemon Y	3DS	2013.0	Role-Playing	Nintendo	5.17	4.05	4.34	0.79	14.35
33	34	Call of Duty: Black Ops 3	PS4	2015.0	Shooter	Activision	5.77	5.81	0.35	2.31	14.24

```
In [13]: np.round(data.corr(),2)
```

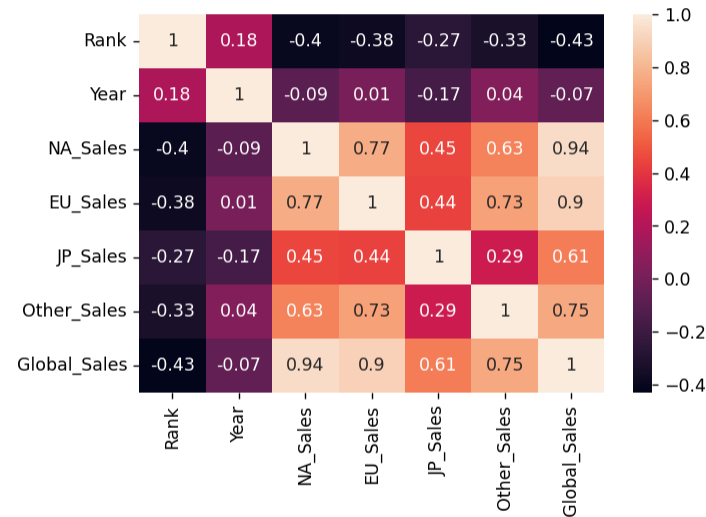
```
Out[13]:
```

	Rank	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
Rank	1.00	0.18	-0.40	-0.38	-0.27	-0.33	-0.43
Year	0.18	1.00	-0.09	0.01	-0.17	0.04	-0.07
NA_Sales	-0.40	-0.09	1.00	0.77	0.45	0.63	0.94
EU_Sales	-0.38	0.01	0.77	1.00	0.44	0.73	0.90
JP_Sales	-0.27	-0.17	0.45	0.44	1.00	0.29	0.61
Other_Sales	-0.33	0.04	0.63	0.73	0.29	1.00	0.75
Global_Sales	-0.43	-0.07	0.94	0.90	0.61	0.75	1.00

```
In [16]: plt.figure(dpi=125)
sns.heatmap(np.round(data.corr(),2),annot=True)
plt.show()
```

JP_Sales	-0.27	-0.17	0.45	0.44	1.00	0.29	0.61
Other_Sales	-0.33	0.04	0.63	0.73	0.29	1.00	0.75
Global_Sales	-0.43	-0.07	0.94	0.90	0.61	0.75	1.00

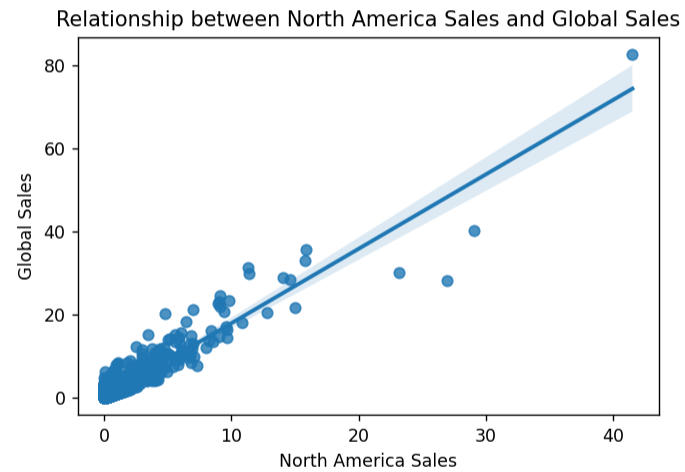
```
In [16]: plt.figure(dpi=125)
sns.heatmap(np.round(data.corr(),2),annot=True)
plt.show()
```



```
In [17]: plt.figure(dpi=125)
sns.regplot(x=data['NA_Sales'],y=data['Global_Sales'])
plt.xlabel('North America Sales')
plt.ylabel('Global Sales')
plt.title('Relationship between North America Sales and Global Sales')
plt.show()
```



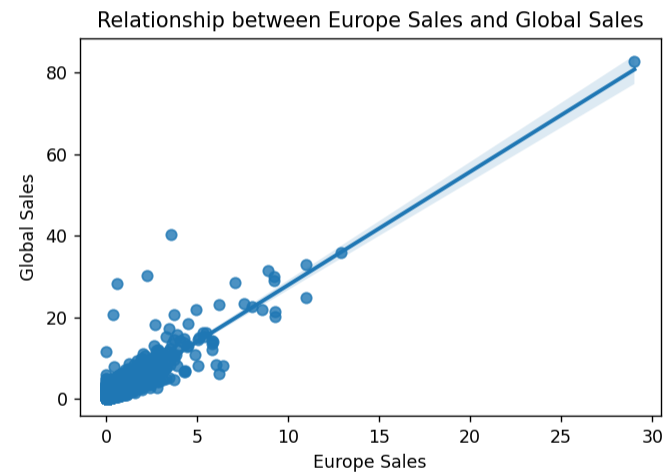
```
In [17]: plt.figure(dpi=125)
sns.regplot(x=data['NA_Sales'],y=data['Global_Sales'])
plt.xlabel('North America Sales')
plt.ylabel('Global Sales')
plt.title('Relationship between North America Sales and Global Sales')
plt.show()
```



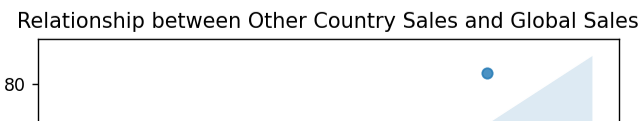
```
In [18]: plt.figure(dpi=125)
sns.regplot(x=data['EU_Sales'],y=data['Global_Sales'])
plt.xlabel('Europe Sales')
plt.ylabel('Global Sales')
plt.title('Relationship between Europe Sales and Global Sales')
plt.show()
```



```
In [18]: plt.figure(dpi=125)
sns.regplot(x=data['EU_Sales'],y=data['Global_Sales'])
plt.xlabel('Europe Sales')
plt.ylabel('Global Sales')
plt.title('Relationship between Europe Sales and Global Sales')
plt.show()
```

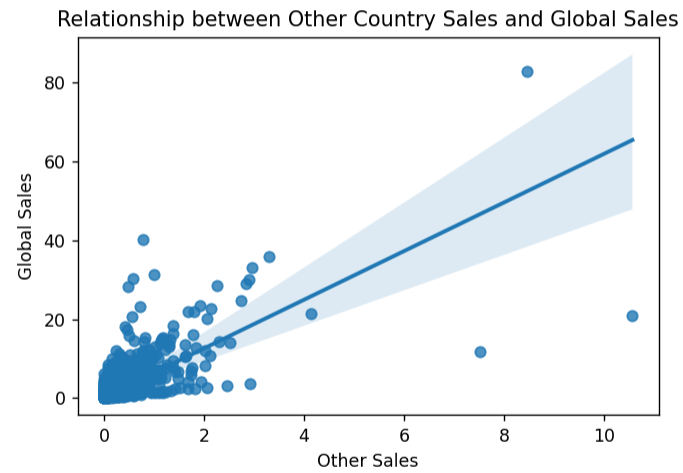


```
In [19]: plt.figure(dpi=125)
sns.regplot(x=data['Other_Sales'],y=data['Global_Sales'])
plt.xlabel('Other Sales')
plt.ylabel('Global Sales')
plt.title('Relationship between Other Country Sales and Global Sales')
plt.show()
```





```
In [19]: plt.figure(dpi=125)
sns.regplot(x=data['Other_Sales'],y=data['Global_Sales'])
plt.xlabel('Other Sales')
plt.ylabel('Global Sales')
plt.title('Relationship between Other Country Sales and Global Sales')
plt.show()
```

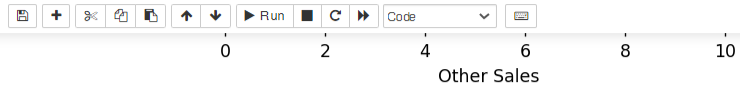


```
In [21]: plt.figure(dpi=125)
sns.countplot(data.head(100)['Publisher'])
plt.xlabel('Publisher Name')
plt.xticks(rotation=90)
plt.ylabel('Count')
plt.title('Best Publisher in Top 100 Video Games')
plt.show()
```

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\decorators.py:36: FutureWarning: Pass the following variable as a keyword argument: x. From version 0.12, the only valid positional argument will be 'data', and passing other arguments without an explicit keyword will result in an error or misinterpretation.

warnings.warn(

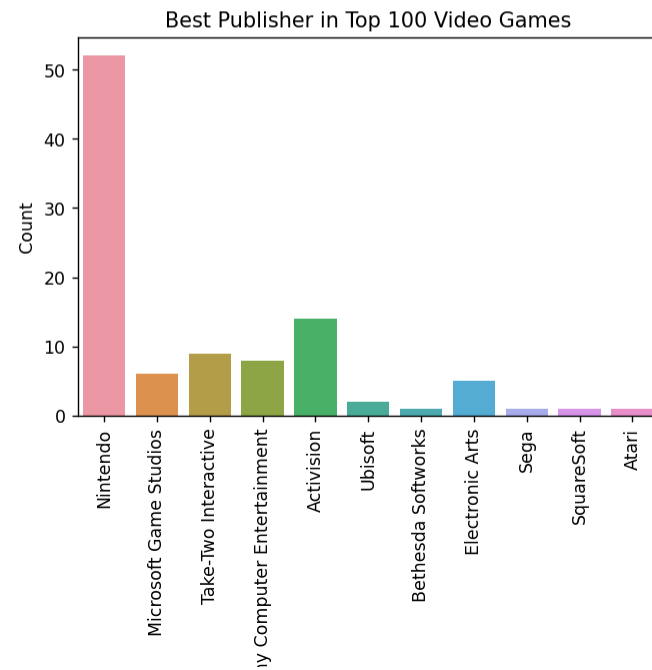
Best Publisher in Top 100 Video Games

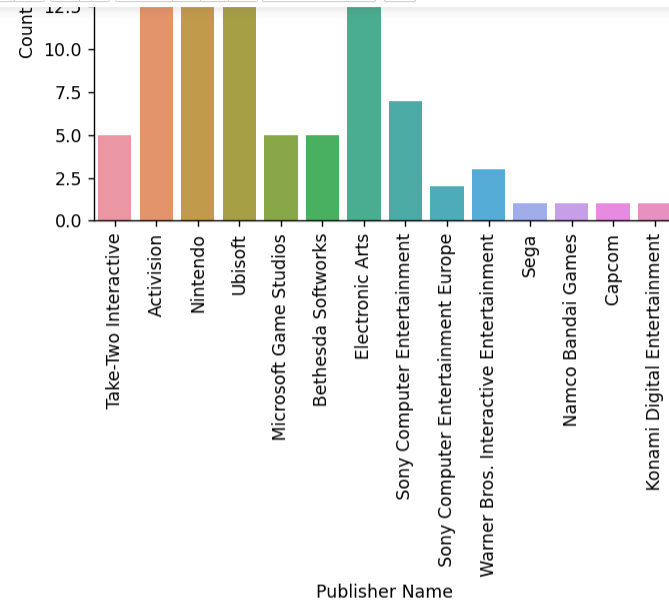


```
In [21]: plt.figure(dpi=125)
sns.countplot(data.head(100)['Publisher'])
plt.xlabel('Publisher Name')
plt.xticks(rotation=90)
plt.ylabel('Count')
plt.title('Best Publisher in Top 100 Video Games')
plt.show()
```

C:\ProgramData\Anaconda3\lib\site-packages\seaborn_decorators.py:36: FutureWarning: Pass the following variable as a keyword argument: x. From version 0.12, the only valid positional argument will be 'data', and passing other arguments without an explicit keyword will result in an error or misinterpretation.

warnings.warn(





```
In [6]: from wordcloud import WordCloud
```

```
-----  
ModuleNotFoundError: Traceback (most recent call last)  
Input In [6], in <cell line: 1>()  
----> 1 from wordcloud import WordCloud  
  
ModuleNotFoundError: No module named 'wordcloud'
```

```
In [ ]:
```