In [1]:
```
import numpy as np
import pandas as pd
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

In [2]:
```
import plotly
import plotly.express as px
import plotly.graph_objects as go
import seaborn as sns
import matplotlib.pyplot as plt
from plotly import tools
from plotly.subplots import make_subplots
from plotly.offline import iplot, init_notebook_mode
init_notebook_mode()
```

In [3]:
```
df=pd.read_csv("Video_Games.csv")
```

In [4]:
```
df.head(7)
```

Out[4]:

	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Score	User_Count	Developer	Rating
0	Wii Sports	Wii	2006.0	Sports	Nintendo	41.36	28.96	3.77	8.45	82.53	76.0	51.0	8	322.0	Nintendo	E
1	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24	NaN	NaN	NaN	NaN	NaN	NaN
2	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.68	12.76	3.79	3.29	35.52	82.0	73.0	8.3	709.0	Nintendo	E
3	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.61	10.93	3.28	2.95	32.77	80.0	73.0	8	192.0	Nintendo	E
4	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37	NaN	NaN	NaN	NaN	NaN	NaN
5	Tetris	GB	1989.0	Puzzle	Nintendo	23.20	2.26	4.22	0.58	30.26	NaN	NaN	NaN	NaN	NaN	NaN
6	New Super Mario Bros.	DS	2006.0	Platform	Nintendo	11.28	9.14	6.50	2.88	29.80	89.0	65.0	8.5	431.0	Nintendo	E

In [5]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 16719 entries, 0 to 16718
Data columns (total 16 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   Name             16717 non-null  object 
 1   Platform         16719 non-null  object 
 2   Year_of_Release  16450 non-null  float64
 3   Genre            16717 non-null  object 
 4   Publisher        16665 non-null  object 
 5   NA_Sales         16719 non-null  float64
 6   EU_Sales         16719 non-null  float64
 7   JP_Sales         16719 non-null  float64
 8   Other_Sales      16719 non-null  float64
 9   Global_Sales     16719 non-null  float64
 10  Critic_Score     8137 non-null   float64
 11  Critic_Count     8137 non-null   float64
 12  User_Score       10015 non-null  object 
 13  User_Count       7590 non-null   float64
 14  Developer        10096 non-null  object 
 15  Rating           9950 non-null   object 
dtypes: float64(9), object(7)
memory usage: 2.0+ MB
```

In [6]:
```
df.isna().sum()
```

Out[6]:
```
Name                 2
Platform             0
Year_of_Release    269
Genre                2
Publisher           54
NA_Sales             0
EU_Sales             0
JP_Sales             0
Other_Sales          0
Global_Sales         0
Critic_Score      8582
Critic_Count      8582
User_Score        6704
User_Count        9129
Developer         6623
Rating            6769
dtype: int64
```

In [7]:
```
pd.unique(df['Platform'])
```

Out[7]:
```
array(['Wii', 'NES', 'GB', 'DS', 'X360', 'PS3', 'PS2', 'SNES', 'GBA',
       'PS4', '3DS', 'N64', 'PS', 'XB', 'PC', '2600', 'PSP', 'XOne',
       'WiiU', 'GC', 'GEN', 'DC', 'PSV', 'SAT', 'SCD', 'WS', 'NG', 'TG16',
       '3DO', 'GG', 'PCFX'], dtype=object)
```

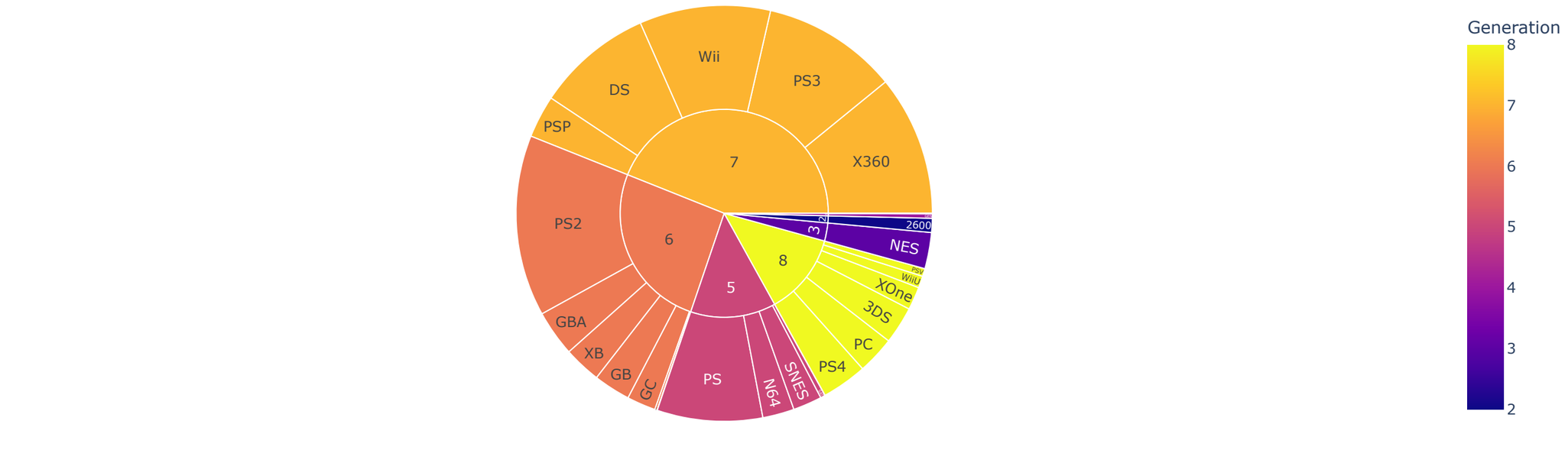
In [8]:
```
code={'Wii':7,'GEN':4,'NES':3,'GB':6,'DS':7,'X360':7,'PS3':7,'PS2':6,'SNES':5,'GBA':6,'PS4':8,'3DS':8,'N64':5,'PS':5,'XB':6,'PC':8,'2600':2,'PSP':7,'XOne':8,'WiiU':8,'GC':6,'DC':6,'PSV':8,'SAT':5,'SCD
df['Generation']=df['Platform'].map(code)
df
```

Out[8]:

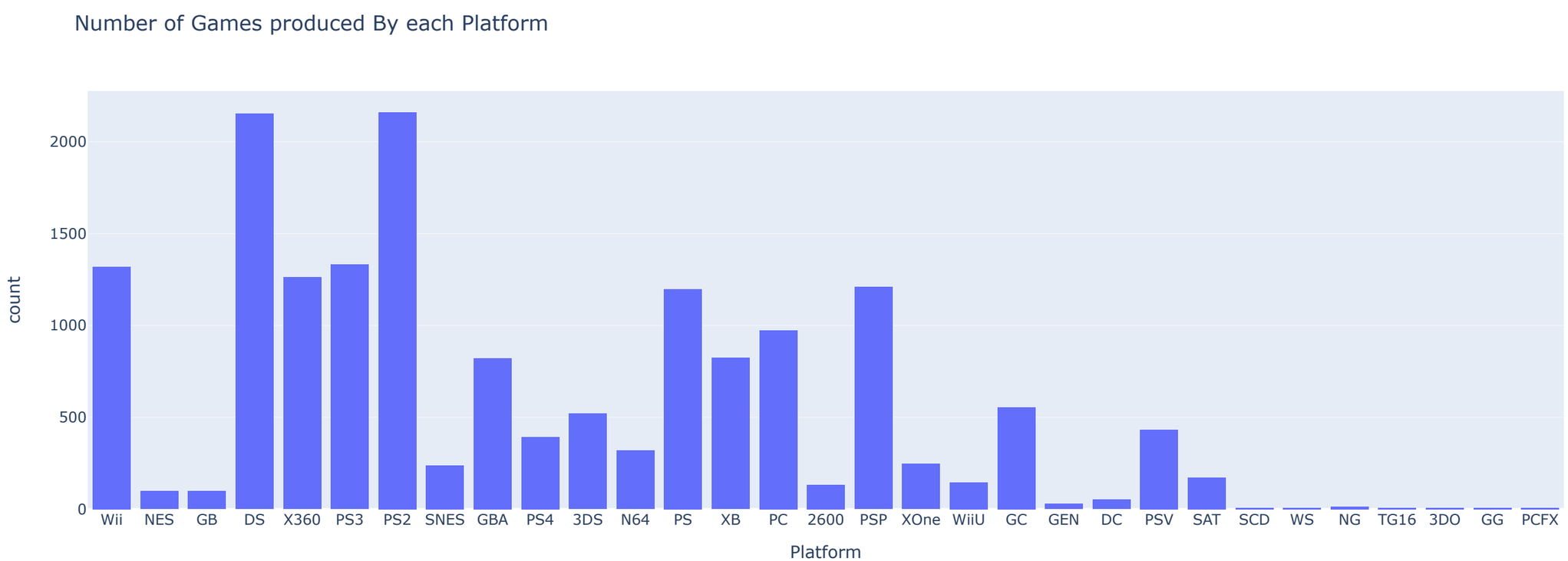
	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Score	User_Count	Developer	Rating	Generation
0	Wii Sports	Wii	2006.0	Sports	Nintendo	41.36	28.96	3.77	8.45	82.53	76.0	51.0	8	322.0	Nintendo	E	7
1	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24	NaN	NaN	NaN	NaN	NaN	NaN	3
2	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.68	12.76	3.79	3.29	35.52	82.0	73.0	8.3	709.0	Nintendo	E	7
3	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.61	10.93	3.28	2.95	32.77	80.0	73.0	8	192.0	Nintendo	E	7
4	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37	NaN	NaN	NaN	NaN	NaN	NaN	6
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
16714	Samurai Warriors: Sanada Maru	PS3	2016.0	Action	Tecmo Koei	0.00	0.00	0.01	0.00	0.01	NaN	NaN	NaN	NaN	NaN	NaN	7
16715	LMA Manager 2007	X360	2006.0	Sports	Codemasters	0.00	0.01	0.00	0.00	0.01	NaN	NaN	NaN	NaN	NaN	NaN	7
16716	Haitaka no Psychedelica	PSV	2016.0	Adventure	Idea Factory	0.00	0.00	0.01	0.00	0.01	NaN	NaN	NaN	NaN	NaN	NaN	8
16717	Spirits & Spells	GBA	2003.0	Platform	Wanadoo	0.01	0.00	0.00	0.00	0.01	NaN	NaN	NaN	NaN	NaN	NaN	6
16718	Winning Post 8 2016	PSV	2016.0	Simulation	Tecmo Koei	0.00	0.00	0.01	0.00	0.01	NaN	NaN	NaN	NaN	NaN	NaN	8

16719 rows × 17 columns

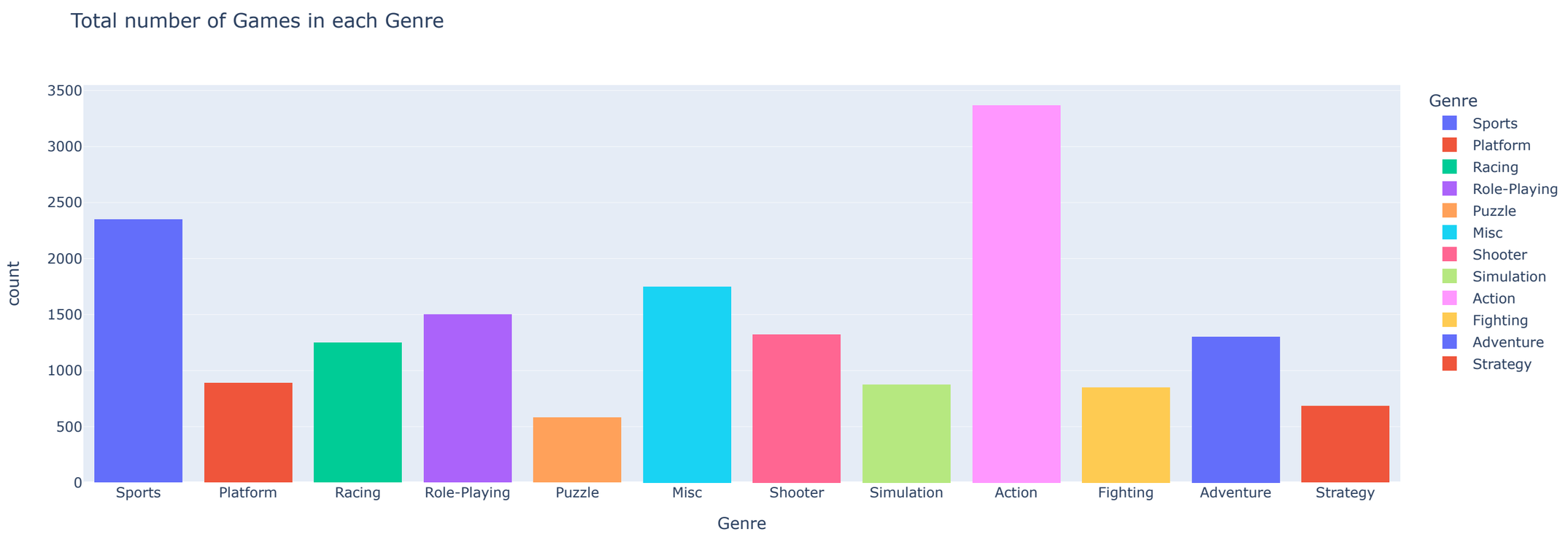
In [9]:
```
fig = px.sunburst(df, path=['Generation', 'Platform'], values='Global_Sales',
                  color='Generation', title = 'Global Sales By Console generations')
fig.show()
```



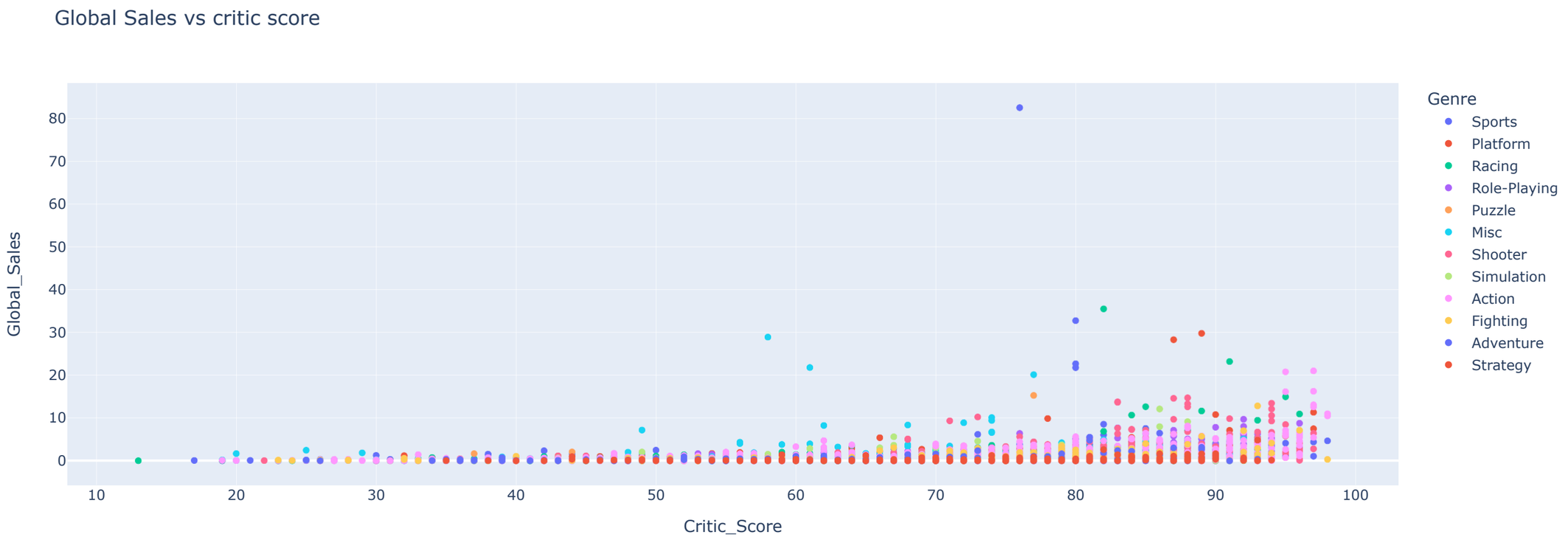
In [10]:
```
fig = px.histogram(df, x="Platform", title = 'Number of Games produced By each Platform')
fig.show()
```



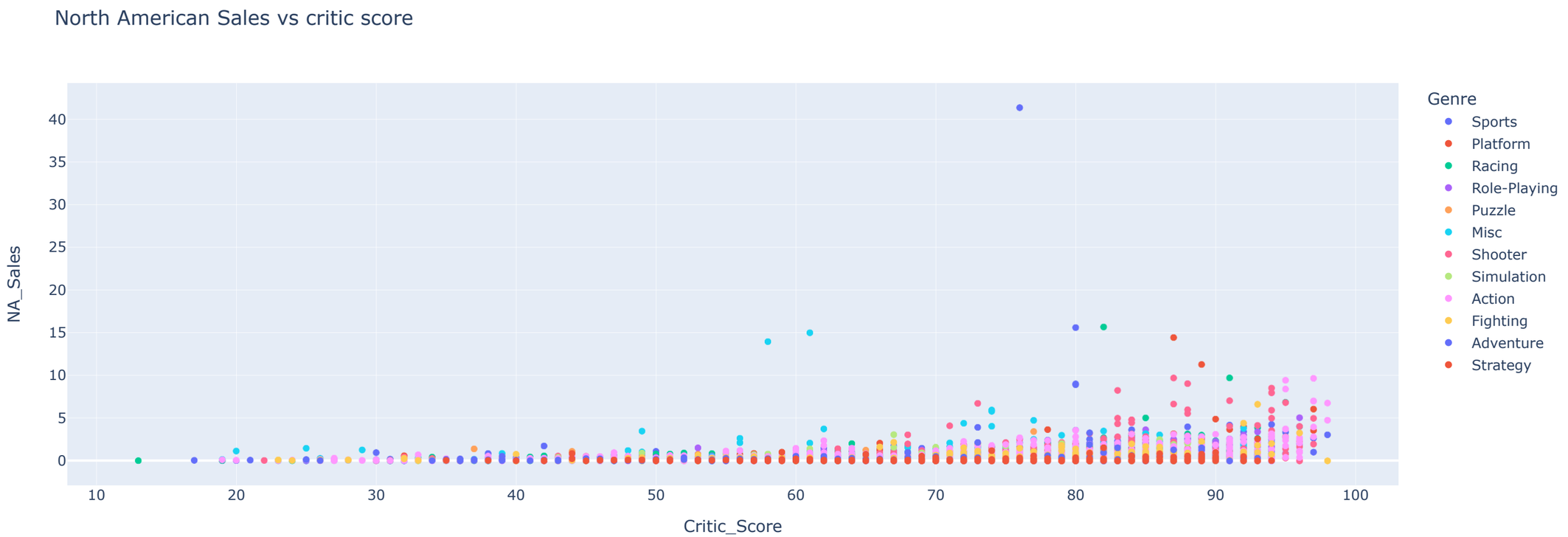
In [11]:
```
fig = px.histogram(df, x="Genre", color = 'Genre', title = 'Total number of Games in each Genre')
fig.show()
```



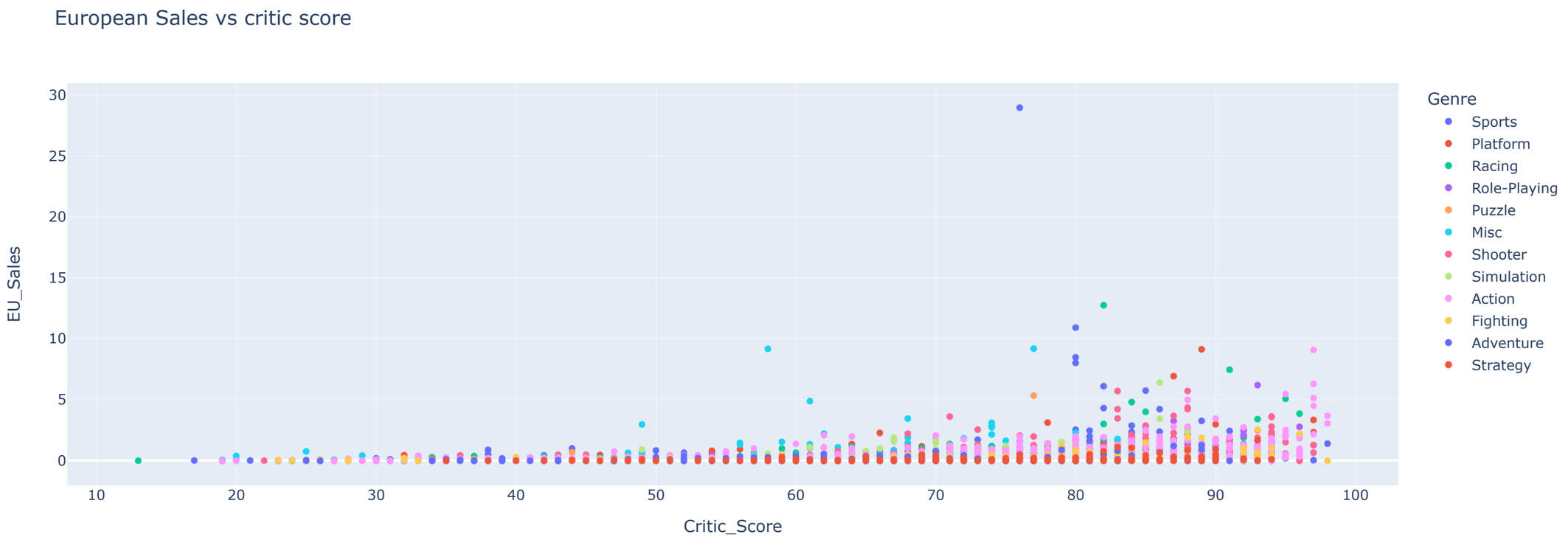
In [12]:
```
fig = px.scatter(df, x="Critic_Score", y="Global_Sales", color="Genre", hover_name = 'Name', title = 'Global Sales vs critic score')
fig.show()
```



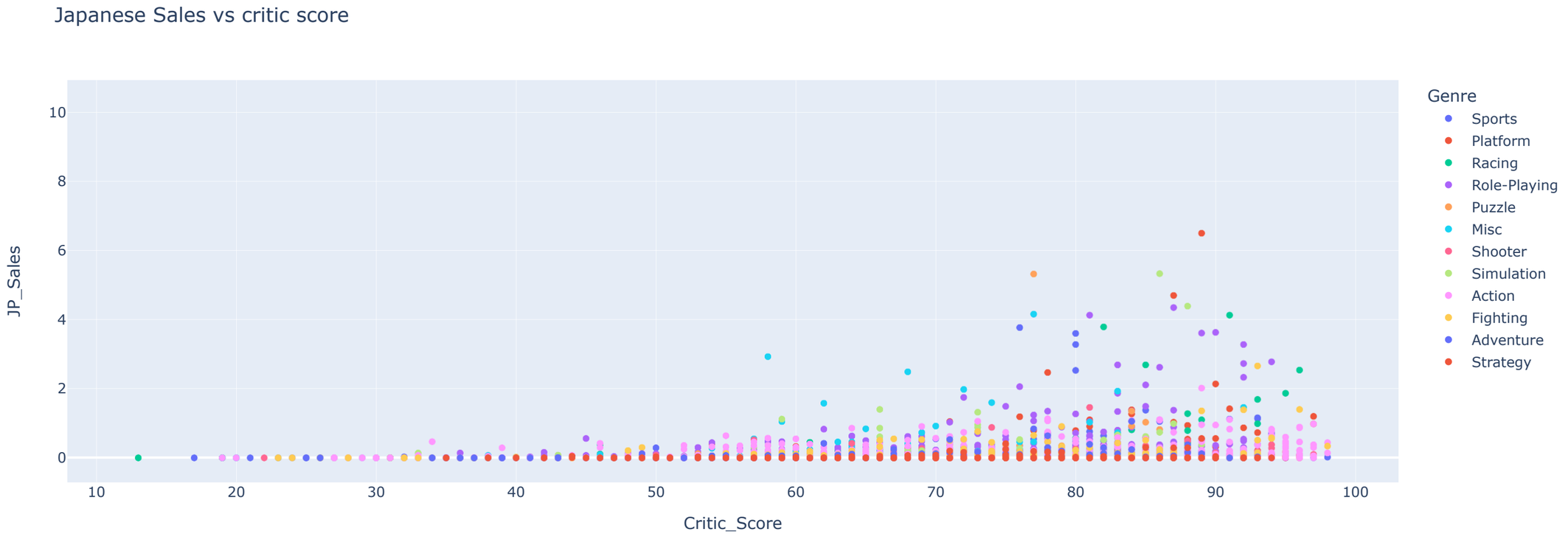
In [13]:
```
fig = px.scatter(df, x="Critic_Score", y="NA_Sales", color="Genre", hover_name = 'Name', title = 'North American Sales vs critic score')
fig.show()
```



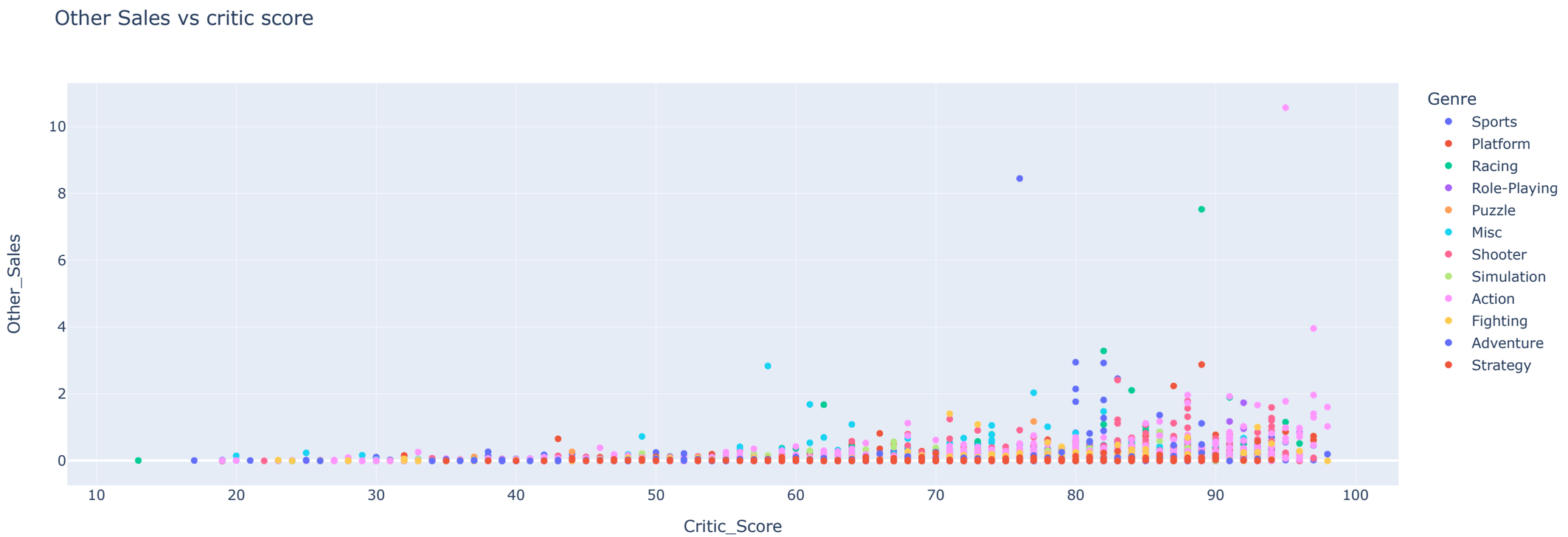
In [14]:
```
fig = px.scatter(df, x="Critic_Score", y="EU_Sales", color="Genre", hover_name = 'Name', title = 'European Sales vs critic score')
fig.show()
```



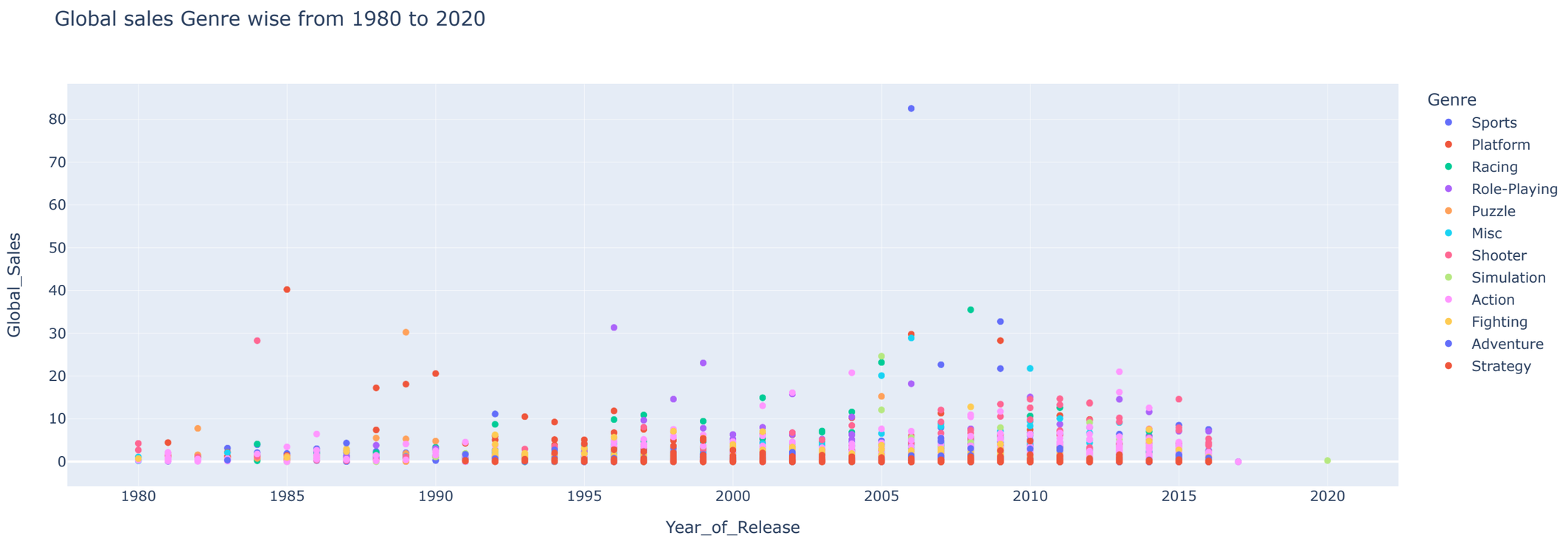
In [15]:
```
fig = px.scatter(df, x="Critic_Score", y="JP_Sales", color="Genre", hover_name = 'Name', title = 'Japanese Sales vs critic score')
fig.show()
```



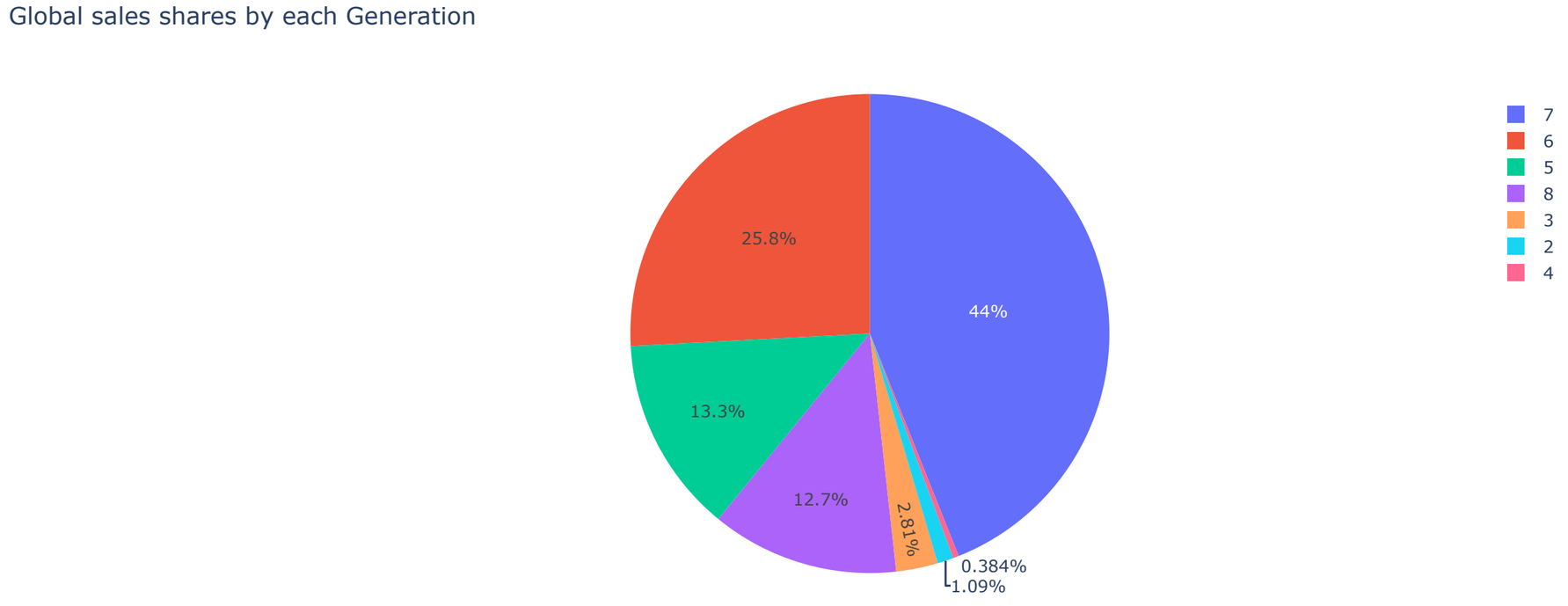
In [16]:
```
fig = px.scatter(df, x="Critic_Score", y="Other_Sales", color="Genre", hover_name = 'Name', title = 'Other Sales vs critic score')
fig.show()
```



In [17]:
```
fig = px.scatter(df, x="Year_of_Release", y="Global_Sales", color="Genre", hover_name = 'Name', title = 'Global sales Genre wise from 1980 to 2020')
fig.show()
```

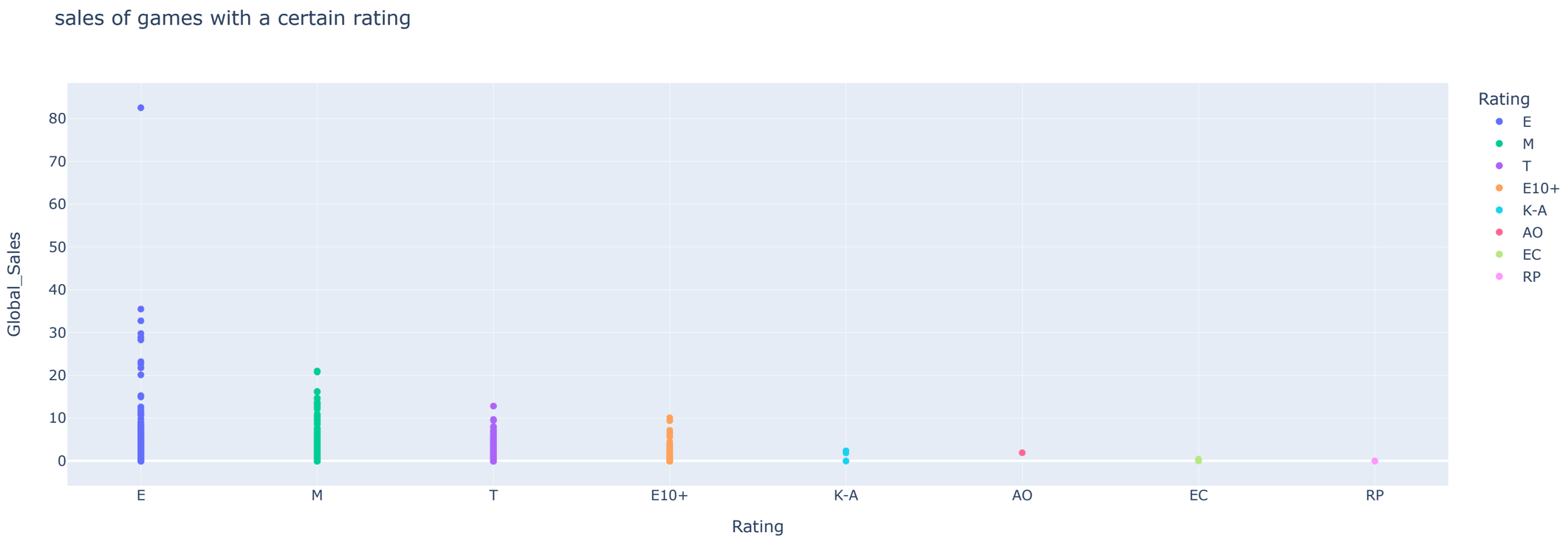


In [18]:
```
fig = px.pie(df, values='Global_Sales', names='Generation', title='Global sales shares by each Generation')
fig.show()
```



In [19]:
```
plt.figure(figsize=(12,10))
sns.heatmap(df.corr(), annot = True, fmt= '.2f')
plt.show()
```

In [20]:
```
fig = px.scatter(df, x="Rating", y="Global_Sales", color="Rating", hover_name = 'Name', title = 'sales of games with a certain rating')
fig.show()
```



In [26]:
```
pivot = df.pivot_table(index=['Genre'], values=['Global_Sales'], aggfunc='sum')
pivot
```

Out[26]:

Genre	Global_Sales
Action	1745.27
Adventure	237.69
Fighting	447.48
Misc	803.18
Platform	828.08
Puzzle	243.02
Racing	728.90
Role-Playing	934.40
Shooter	1052.94
Simulation	390.42
Sports	1332.00
Strategy	174.50

In [28]:
```
pivot = df.pivot_table(index=['Genre'], values=['Global_Sales'], aggfunc='max')
pivot
```

Out[28]:

Genre	Global_Sales
Action	21.04
Adventure	11.18
Fighting	12.84
Misc	28.92
Platform	40.24
Puzzle	30.26
Racing	35.52
Role-Playing	31.37
Shooter	28.31
Simulation	24.67
Sports	82.53
Strategy	5.45

In [31]:
```
pivot = df.pivot_table(index=['Genre'], values=['Global_Sales'], aggfunc='mean')
pivot
```

Out[31]:

Genre	Global_Sales
Action	0.517884
Adventure	0.182417
Fighting	0.527067
Misc	0.458960
Platform	0.932523
Puzzle	0.419000
Racing	0.583587
Role-Playing	0.622933
Shooter	0.795873
Simulation	0.446705
Sports	0.567291
Strategy	0.255490

In [35]:
```
pivot= df.pivot_table(values='Global_Sales',index='Genre',columns='Generation',fill_value=0, aggfunc=lambda x: round(sum(x)/sum(df['Global_Sales']) * 100,2))

pivot
```

Out[35]:

Generation Genre	2	3	4	5	6	7	8
Action	0.33	0.32	0.03	1.88	4.76	9.33	2.91
Adventure	0.02	0.05	0.00	0.30	0.72	1.25	0.31
Fighting	0.01	0.07	0.08	1.46	1.46	1.58	0.35
Misc	0.04	0.04	0.00	0.70	1.98	5.65	0.58
Platform	0.15	1.07	0.19	1.87	2.76	2.50	0.74
Puzzle	0.16	0.24	0.00	0.26	0.80	1.18	0.09
Racing	0.03	0.11	0.00	1.78	2.64	3.02	0.57
Role-Playing	0.00	0.21	0.00	1.37	3.08	3.72	2.09
Shooter	0.30	0.40	0.00	0.77	2.14	5.89	2.31
Simulation	0.01	0.00	0.00	0.47	0.77	2.21	0.92
Sports	0.04	0.30	0.04	1.96	4.29	7.11	1.21
Strategy	0.00	0.00	0.00	0.45	0.43	0.51	0.57

In [40]:
```
pivot = df.pivot_table(index=['Genre'], values=['Global_Sales'], aggfunc='min')
pivot
```

Out[40]:

Genre	Global_Sales
Action	0.01
Adventure	0.01
Fighting	0.01
Misc	0.01
Platform	0.01
Puzzle	0.01
Racing	0.01
Role-Playing	0.01
Shooter	0.01
Simulation	0.01
Sports	0.01
Strategy	0.01

In [ ]: