

AIB_18기_정하람

Section 1 : Project 발표

CONTENTS

1. 지역에 따른 선호하는 게임 장르의 차이
2. 연도별 게임의 트렌드
3. 인기가 많은 게임에 대한 분석
4. 다음 분기 게임 설계 기획



지역에 따른 선호하는 게임 장르



NA, EU, JP, Other
4개 지역으로 나뉘었다.



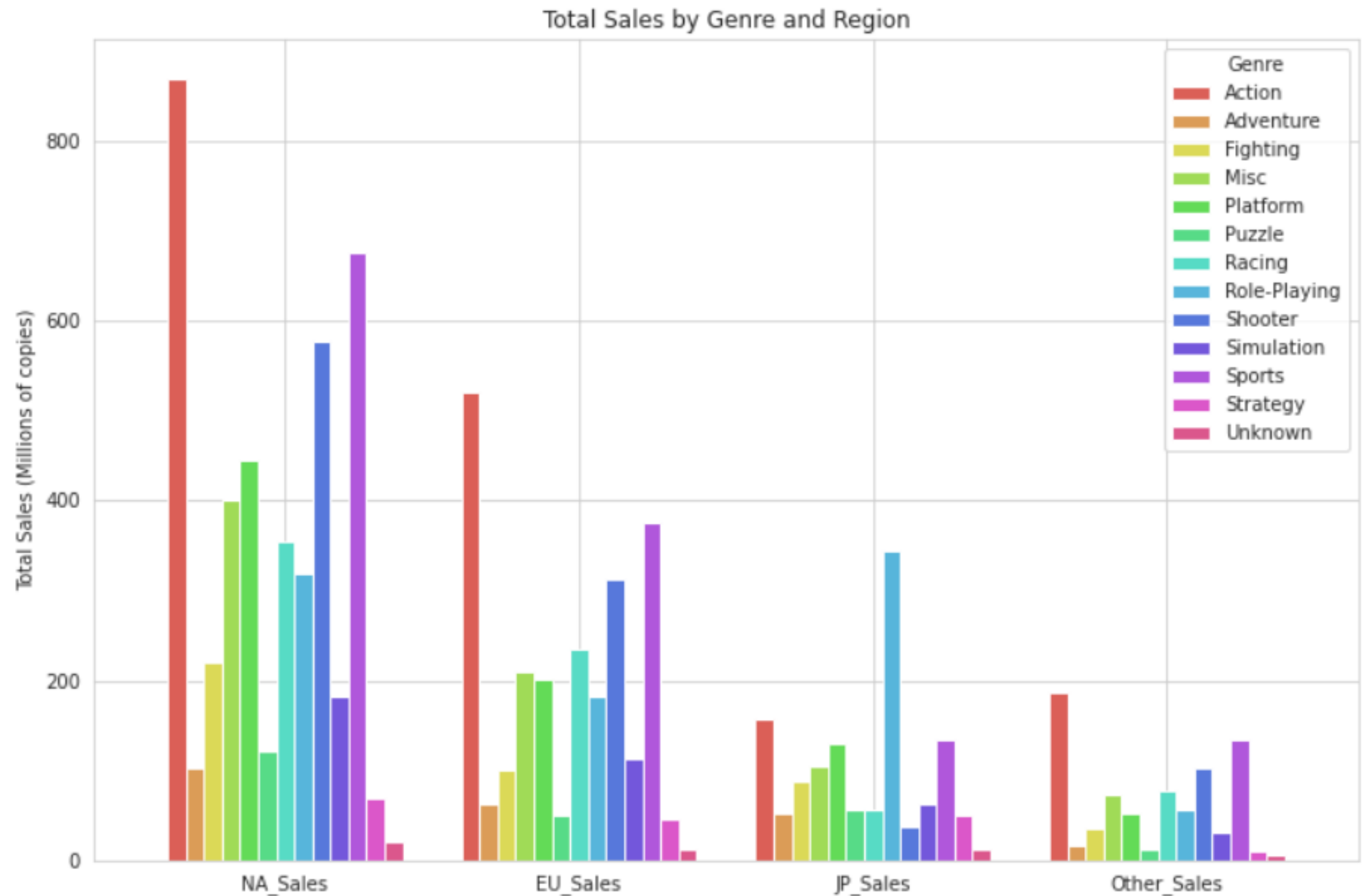
지역별 선호하는
게임 장르를 파악하고
시각화하였다.



지역별 선호하는
게임 장르의 시각화 정보를
토대로 그 차이를 파악하였다.

지역별, 장르별 게임 출고량

NA 지역에서 가장 많은 게임이 팔린 것을
알 수 있고, Action의 출고량이 가장 높았다.



코드 구현

```
#지역별, 장르별 출고량을 시각화
df_genre_sales = df.groupby('Genre')[['NA_Sales', 'EU_Sales', 'JP_Sales', 'Other_Sales']].sum()

sns.set_style('whitegrid')
colors = sns.color_palette("hls", len(df_genre_sales.index))
df_genre_sales.T.plot(kind='bar', figsize=(12, 8), title='Total Sales by Genre and Region', color=colors, width=0.8)
plt.xticks(rotation=0)
plt.ylabel('Total Sales (Millions of copies)')
plt.show()
```

df_genre_sales는 df데이터의 장르를 토대로 각 판매량의 총 합을 구하여 그룹화 시켰다.

whitegrid 스타일을 지정하여 하얀색의 격자무늬 배경을 그리고 그 위에 막대그래프를 그렸다.

기본 색상으로 하면 장르끼리 색상이 겹치기 때문에 팔레트 'hls'을 이용하여 색 구분을 해주었다.

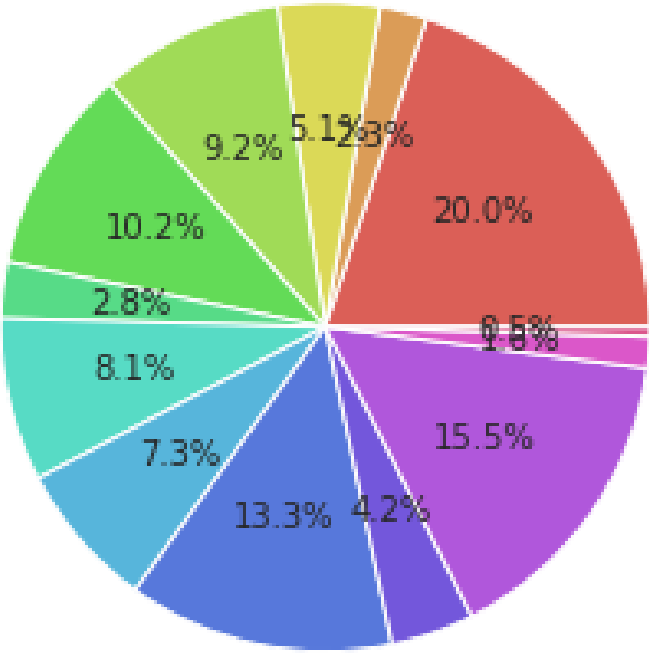
df_genre_sales의 역행렬로 그림을 그렸기 때문에 x축의 글씨도 0으로 돌려주었다.

ylabel의 이름을 지어주었다.

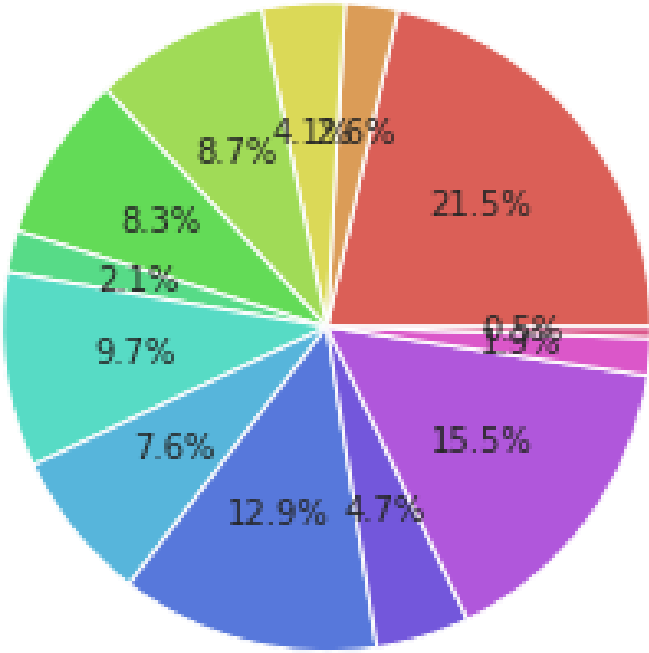
지역별, 장르별 게임 출고량의 비율

Japan을 제외한 모든 지역에서, Action 장르의 출고량이 가장 높았고, Japan에서만 유일하게 Role-Playing이 가장 높았다.

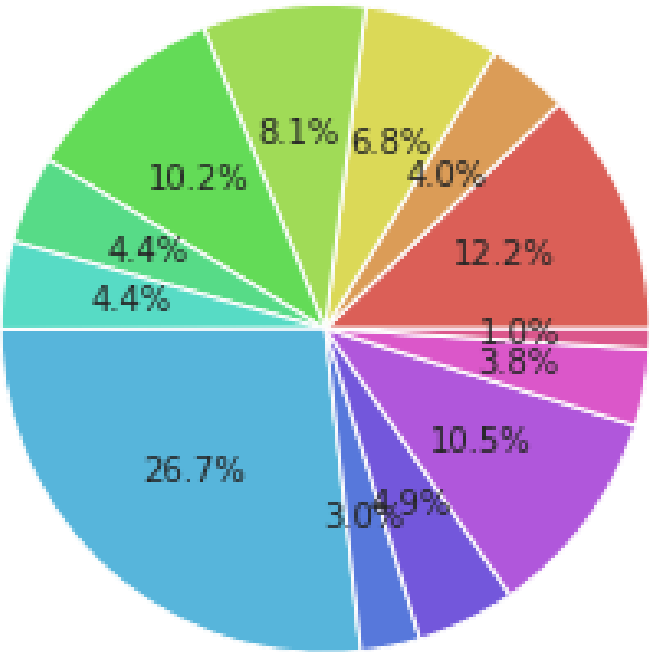
North America



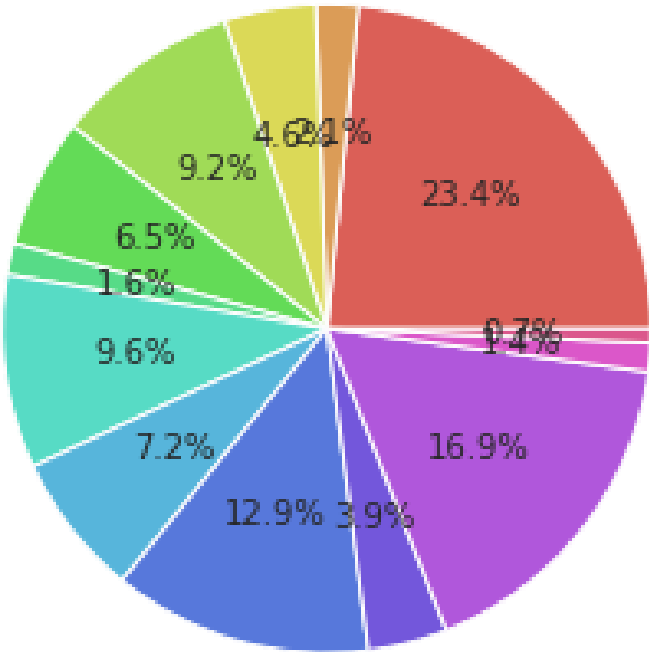
Europe



Japan



Other



코드 구현

```
#서브플롯 그리기
fig, axes = plt.subplots(2, 2, figsize=(10,10))

colors = sns.color_palette("hls", len(df_genre_sales.index))

# 각 서브플롯에 파이차트 그리기

axes[0][0].pie(df_genre_sales['NA_Sales'], autopct="%1.1f%%", colors=colors)
axes[0][0].set_title("North America")
axes[0][1].pie(df_genre_sales['EU_Sales'], autopct="%1.1f%%", colors=colors)
axes[0][1].set_title("Europe")
axes[1][0].pie(df_genre_sales['JP_Sales'], autopct="%1.1f%%", colors=colors)
axes[1][0].set_title("Japan")
axes[1][1].pie(df_genre_sales['Other_Sales'], autopct="%1.1f%%", colors=colors)
axes[1][1].set_title("Other")

fig.legend(df_genre_sales.index, loc="lower center", ncol=5, prop={"size": 12})

# 플롯 보여주기
plt.show()
```

그림을 2x2 총 4개로 분할하고 크기를 지정해주었다.

색상을 팔레트 'hls'으로 지정해주었다.

분할된 axes에 [0],[0]부터 [1],[1]까지
지역별, 장르별 출고량을 pie그림으로 그렸다.

장르의 종류가 많고 장르별 색상을 지정해주었기
때문에 범례를 따로 지정해주었다.

NORTH AMERICA

Action

20%

—

SPORTS : 15.5%

SHOOTER : 13.3%

EUROPE

Action

21.5%

—

SPORTS : 15.5%

SHOOTER : 12.9%

JAPAN

Role-playing

26.7%

—

ACTION : 12.2%

SPORTS : 10.5%

OTHER

Action

23.4%

—

SPORTS : 16.9%

SHOOTER : 12.9%

지역별로
선호하는 게임 장르의
—
차이가 있다.

연도별 트렌드의 변화

연도별 장르
트렌드의 변화

연도별 장르의 트렌드를
파악하고 시각화하였다.

연도별 플랫폼
트렌드의 변화

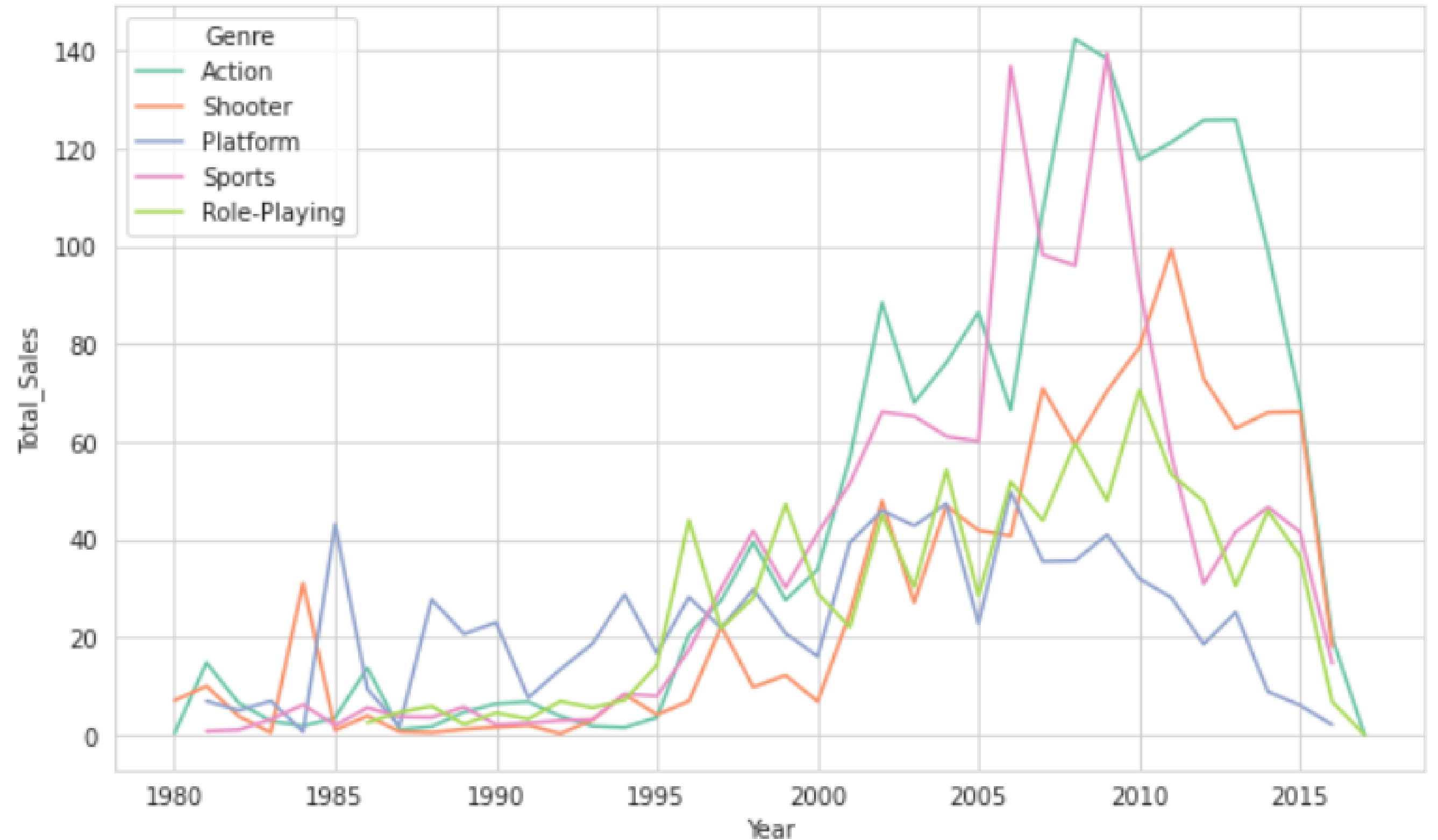
연도별 플랫폼의 트렌드를
파악하고 시각화하였다.

연도별 회사별
트렌드의 변화

연도별 회사의 트렌드를
파악하고 시각화하였다.

연도별 장르 트렌드

Top5 Genres in Total Sales by Year



Top 5. 장르에 대한 연도별 트렌드를 시각화 하였다. 2005년부터 2010년까지의 총 출고량이 가장 높았다.

코드 구현

```
# Top5 장르 선정
top5_genre = df.groupby('Genre')['Total_Sales'].sum().sort_values(ascending=False).head(5).index.tolist()

# 연도별 Top5 장르 선택
df_top5 = df[df['Genre'].isin(top5_genre)].groupby(['Year', 'Genre']).sum().reset_index()

# 시각화
plt.figure(figsize=(10, 6))
sns.lineplot(data=df_top5, x='Year', y='Total_Sales', hue='Genre')
plt.title('Top5 Genres in Total Sales by Year')
plt.show()
```

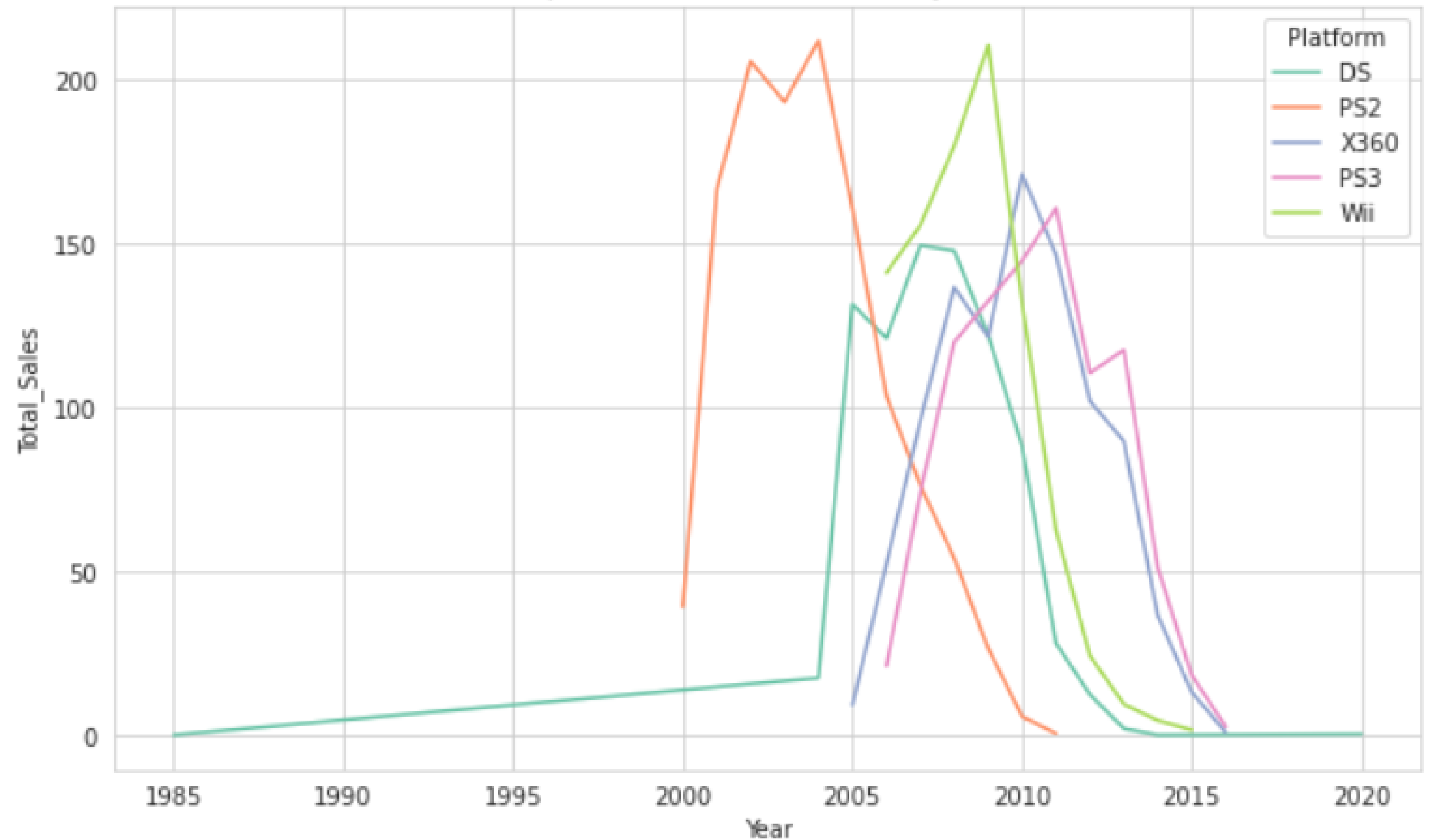
groupby를 이용해서 장르별 총 판매량을 구하고 내림차순하여 top 5를 구했다.

선택된 5개 장르에 대하여 연도별 총 판매량을 구한다.

line플롯을 활용하여 top 5 장르에 대한 판매량 추이를 확인하였다.

연도별 플랫폼 트렌드

Top5 Platform in Total Sales by Year



Top 5. 플랫폼에 대한 연도별 트렌드를 시각화 하였다. ps2, x360, ps3, wii, DS 순으로 높았다.

코드 구현

```
# Top5 플랫폼 선정
top5_platform = df.groupby('Platform')['Total_Sales'].sum().sort_values(ascending=False).head(5).index.tolist()

# 연도별 Top5 플랫폼 선택
df_top5_p = df[df['Platform'].isin(top5_platform)].groupby(['Year', 'Platform']).sum().reset_index()

# 시각화
plt.figure(figsize=(10, 6))
sns.lineplot(data=df_top5_p, x='Year', y='Total_Sales', hue='Platform')
plt.title('Top5 Platform in Total Sales by Year')
plt.show()
```

groupby를 이용해서 플랫폼별 총 판매량을 구하고 내림차순하여 top 5를 구했다.

선택된 5개 플랫폼에 대하여 연도별 총 판매량을 구한다.

line플롯을 활용하여 top 5 플랫폼에 대한 판매량 추이를 확인하였다.

연도별 회사별 트렌드

```
df['Platform'].value_counts()
```

```
# Sony: PS2, PS3, PS, PSP, PSV, PS4
```

```
# Nintendo: DS, Wii, GBA, GC, 3DS, N64, SNES, WiiU, NES, GB
```

```
# Microsoft: X360, XB, XOne
```

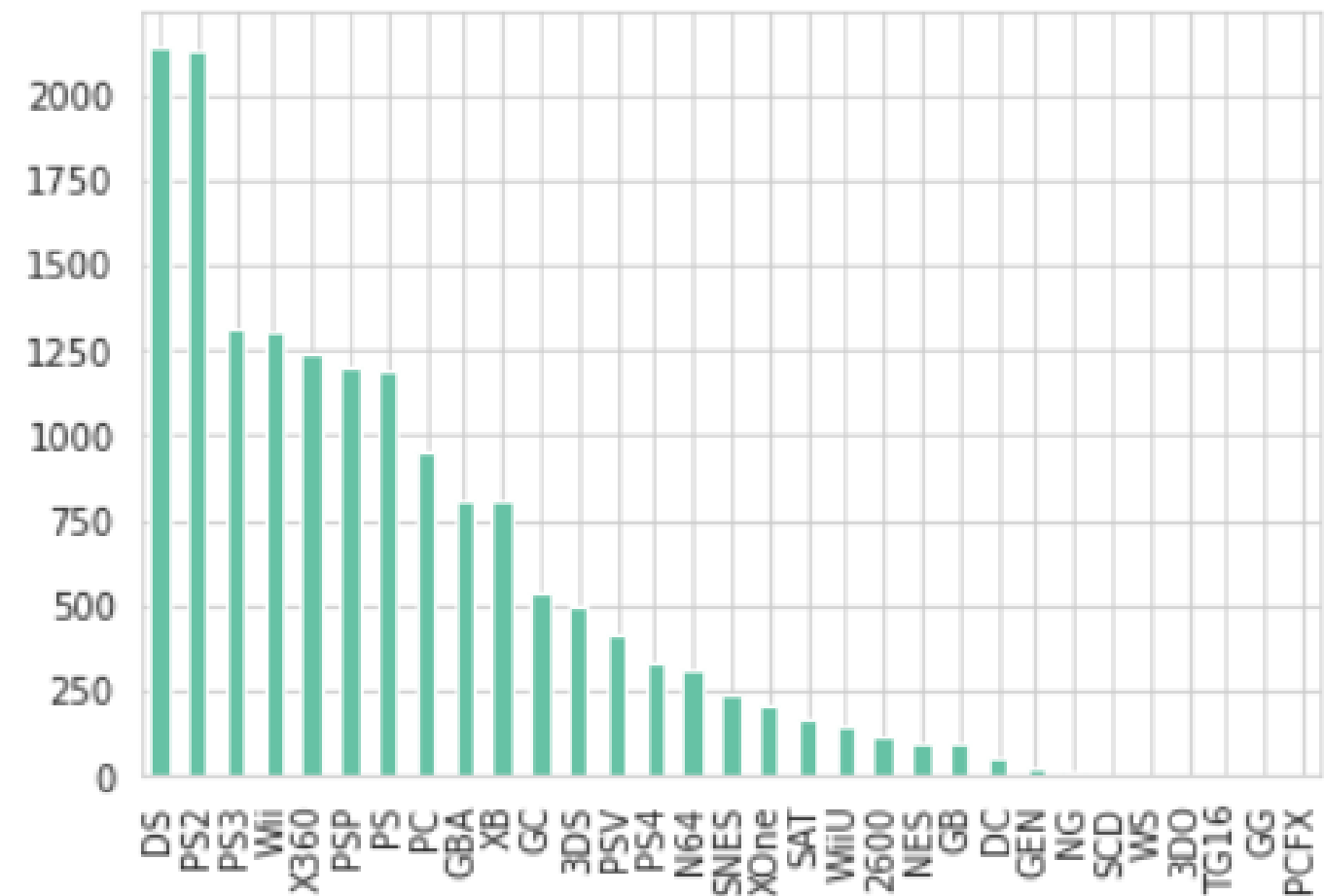
```
# PC: PC
```

```
# Sega: SAT, GEN, SC, DC
```

```
# 기타: 2600, NG, SCD, WS, 3DO, TG16, GG, PCFX
```

```
df['Platform'].value_counts().plot(kind='bar')
```

```
plt.show()
```



플랫폼이 너무 많기 때문에
회사별로 다시 분류하였다.

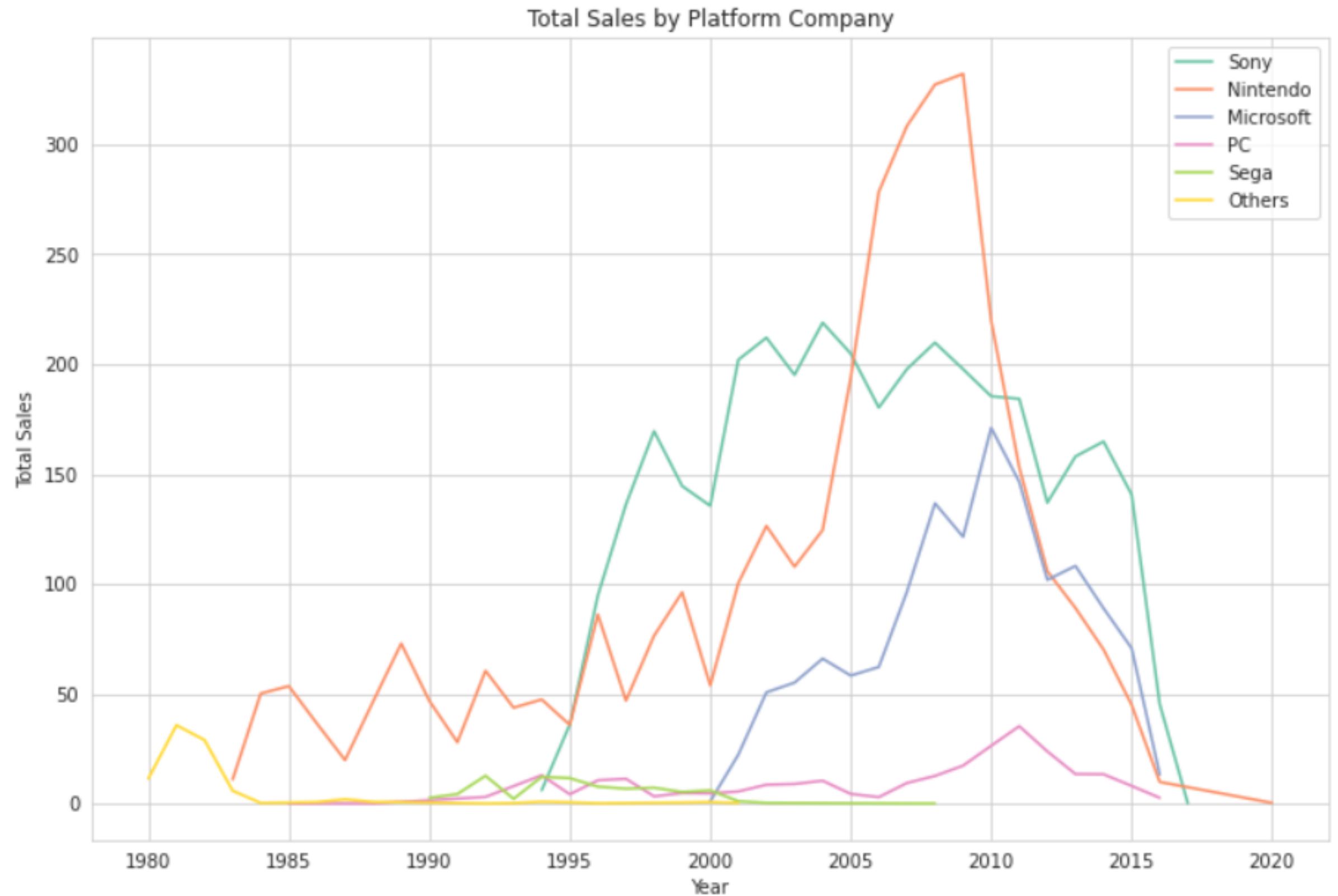
코드 구현

```
def categorize_company(platform):  
    if platform in ['PS2', 'PS3', 'PS', 'PSP', 'PSV', 'PS4']:  
        return 'Sony'  
    elif platform in ['DS', 'Wii', 'GBA', 'GC', '3DS', 'N64', 'SNES', 'WiiU', 'NES', 'GB']:  
        return 'Nintendo'  
    elif platform in ['X360', 'XB', 'XOne']:  
        return 'Microsoft'  
    elif platform == 'PC':  
        return 'PC'  
    elif platform in ['SAT', 'GEN', 'SCD', 'DC']:  
        return 'Sega'  
    else:  
        return 'Others'  
df['Company'] = df['Platform'].apply(categorize_company)
```

If 함수를 사용해서 각 플랫폼을 Company의 이름으로 다시 카테고리 분류 하였다.

연도별 회사별 트렌드

회사별 연도별, 출고량 추이를 확인하였다.
닌텐도가 가장 많은 것으로 파악되었다.



코드 구현

```
plt.figure(figsize=(12,8))

companies = ['Sony', 'Nintendo', 'Microsoft', 'PC', 'Sega', 'Others']

for company in companies:
    df_company = df[df['Company'] == company].groupby('Year').sum()
    plt.plot(df_company.index, df_company['Total_Sales'], label=company)

plt.xlabel('Year')
plt.ylabel('Total Sales')
plt.title('Total Sales by Platform Company')
plt.legend()
plt.show()
```

플랫폼을 회사별로 분류하여 시각화 하였다.

각 회사 이름을 리스트로 저장한 뒤 df[df['Company']] 에 따른 연도별 총 매출을 구했다.

인기가 많은 게임 분석

인기많은 게임
Top 20

인기많은 게임 Top20을 파악하고
시각화하였다.

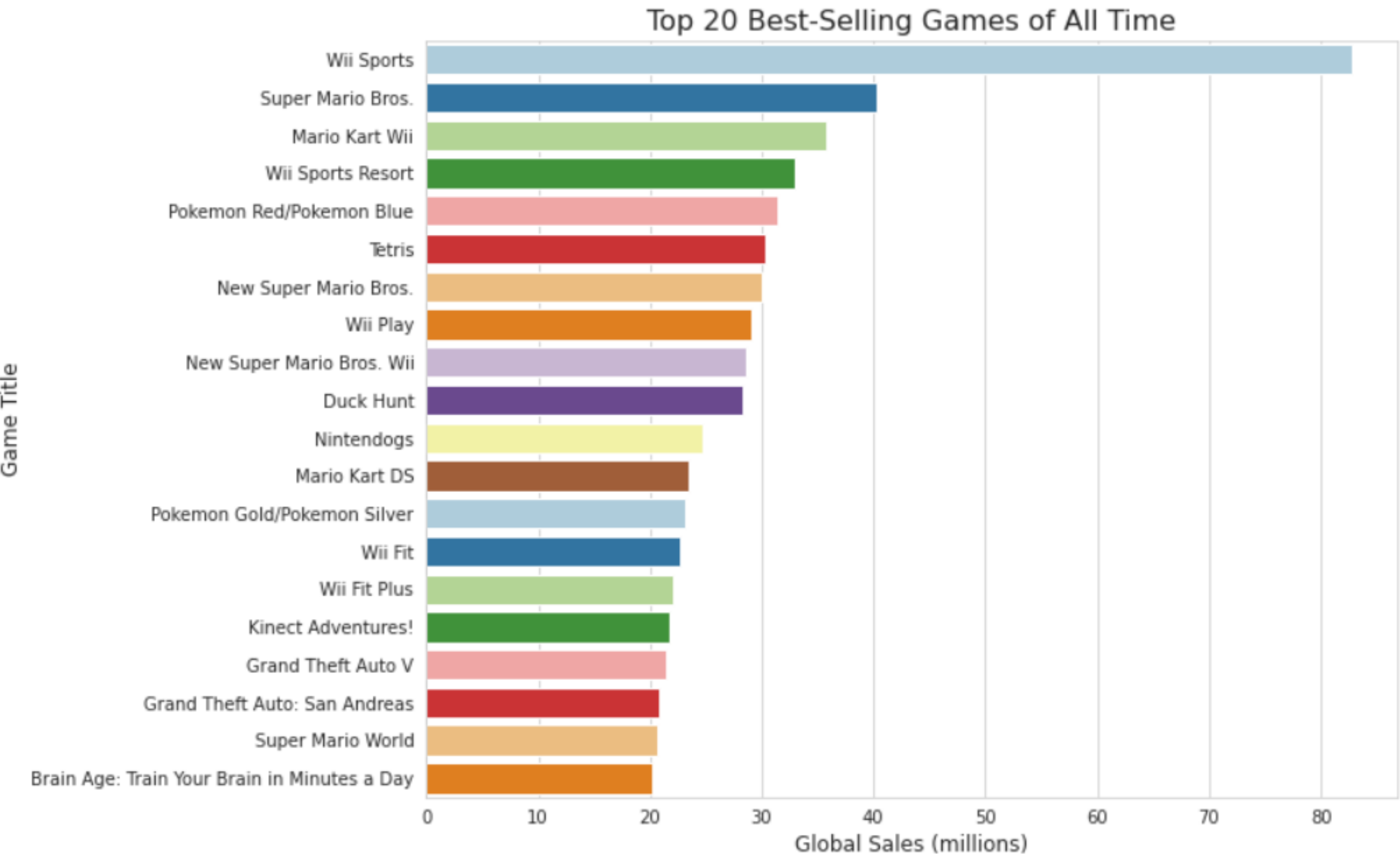
인기많은 장르

인기많은 장르를 파악하고
시각화하였다.

인기많은 플랫폼

인기많은 플랫폼을 파악하고
시각화하였다.

인기많은 게임 TOP.20



Wii Sports가 가장 많이 팔렸다

코드 구현

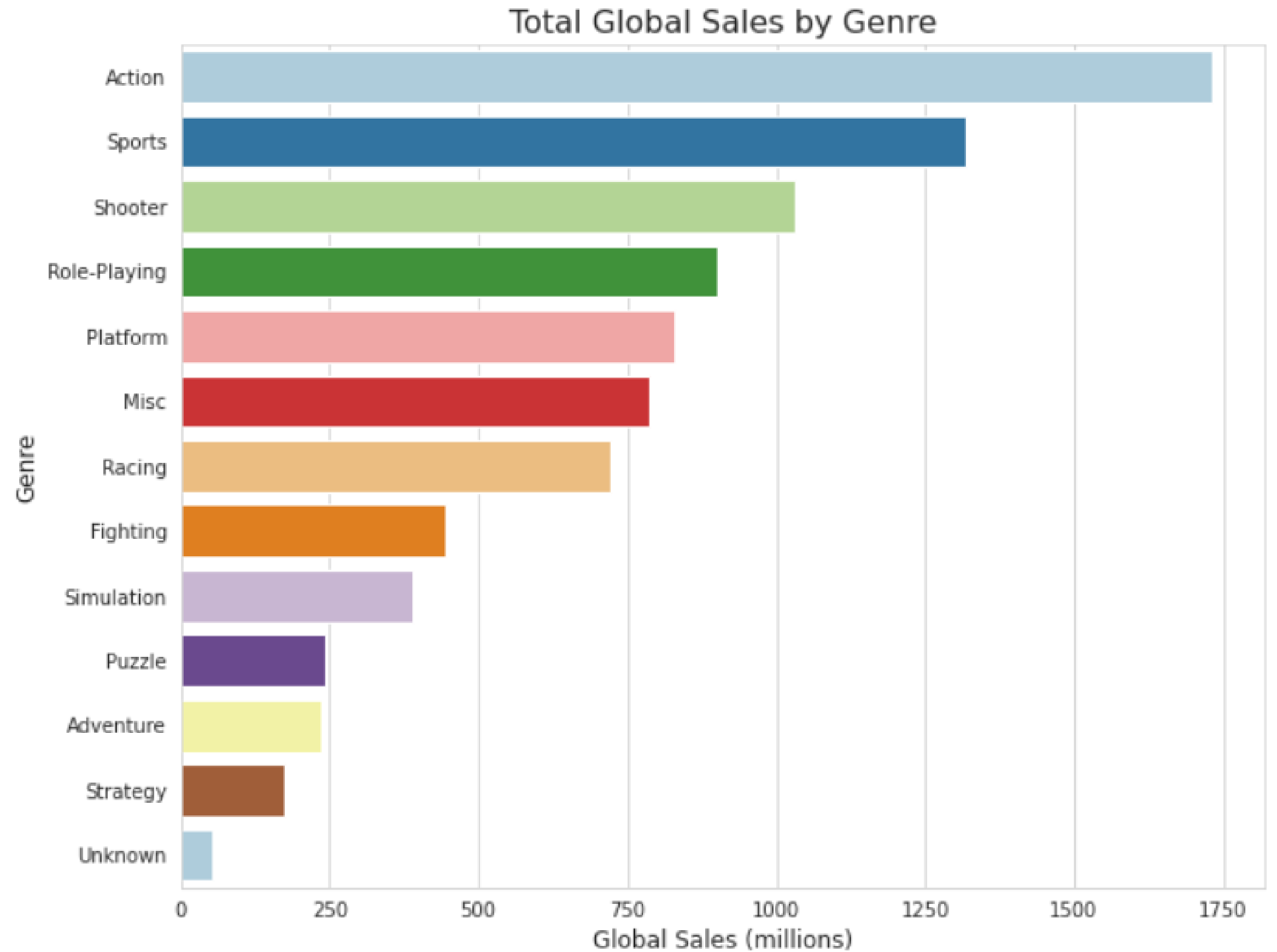
```
top20 = df.sort_values('Total_Sales', ascending=False).head(20)
genre_sales = df.groupby('Genre')['Total_Sales'].sum().sort_values(ascending=False)

plt.figure(figsize=(10,8))
sns.barplot(x='Total_Sales', y='Name', data=top20, palette='Paired')
plt.title('Top 20 Best-Selling Games of All Time', fontsize=16)
plt.xlabel('Global Sales (millions)', fontsize=12)
plt.ylabel('Game Title', fontsize=12)
plt.show()
```

총 출고량을 기준으로 Top 20을 구하여 Bar 플롯으로 시각화하였다.

인기많은 장르

전 세계적으로 Action 장르가 가장 인기가 많다.



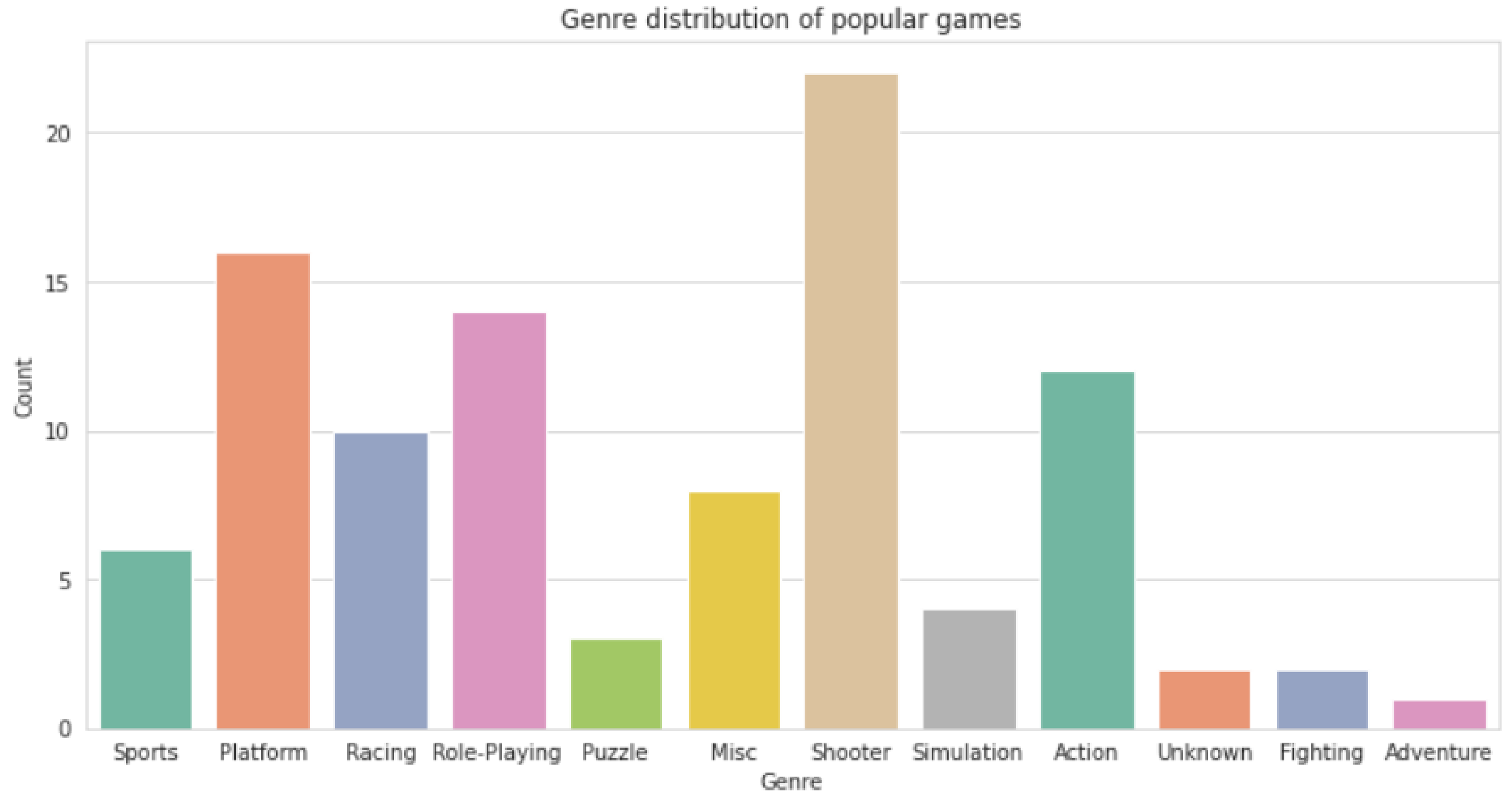
코드 구현

```
genre_sales = df.groupby('Genre')['Total_Sales'].sum().sort_values(ascending=False)

plt.figure(figsize=(10,8))
sns.barplot(x=genre_sales, y=genre_sales.index, palette='Paired')
plt.title('Total Global Sales by Genre', fontsize=16)
plt.xlabel('Global Sales (millions)', fontsize=12)
plt.ylabel('Genre', fontsize=12)
plt.show()
```

총 출고량을 기준으로 가장 많이 팔린 장르순으로 Bar플롯으로 시각화하였다.

인기많은 장르(TOP.20)



Top 20 게임의 포함된 장르 중에서는 Shooter가 가장 높았다.

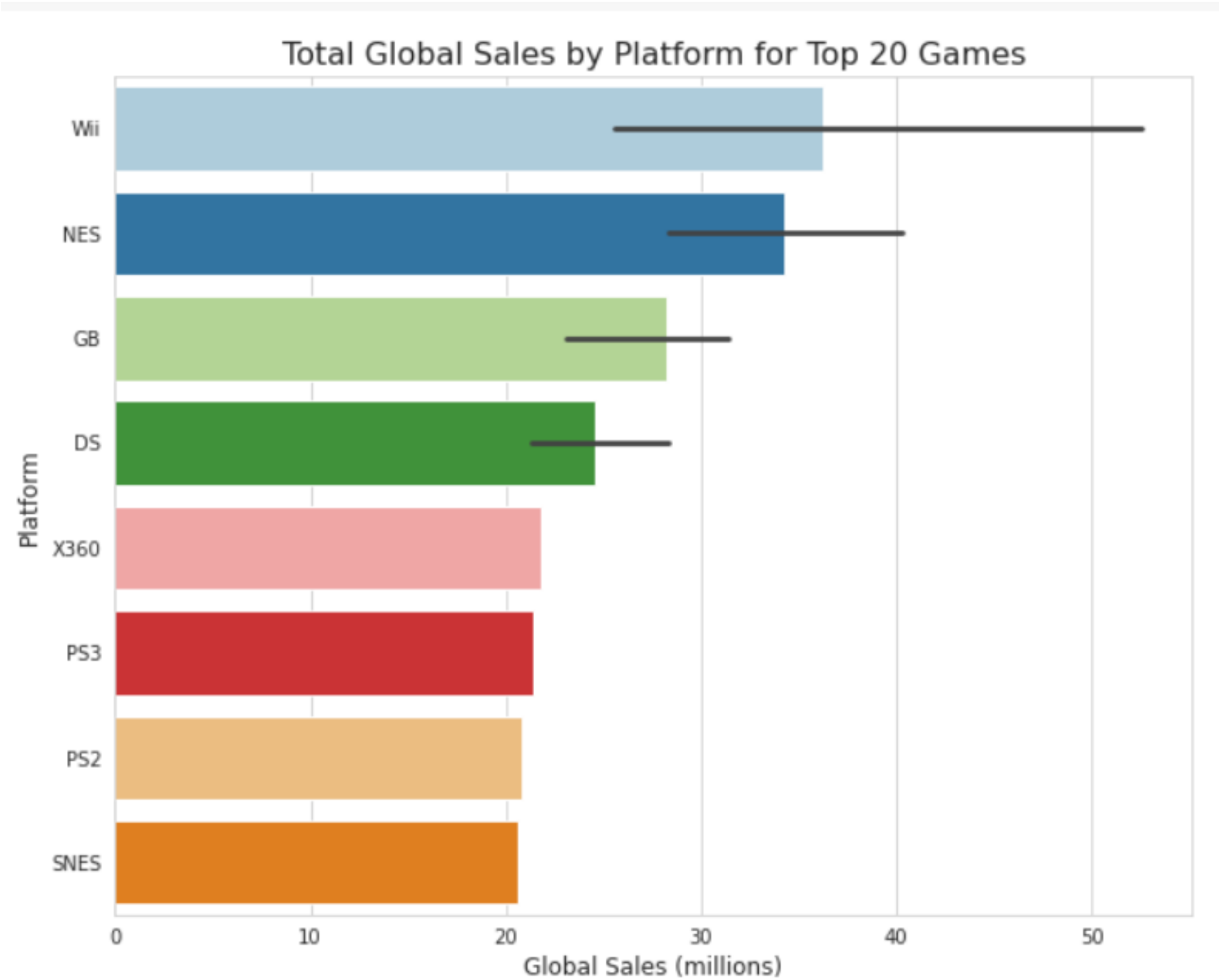
코드 구현

```
popular_games = df.sort_values('Total_Sales', ascending=False).head(100)

plt.figure(figsize=(12,6))
sns.countplot(x='Genre', data=popular_games, palette='Set2')
plt.title('Genre distribution of popular games')
plt.xlabel('Genre')
plt.ylabel('Count')
plt.show()
```

인기많은 게임 Top20을 기준으로 가장 많이 팔린 장르순으로 시각화하였다.

인기 많은 플랫폼



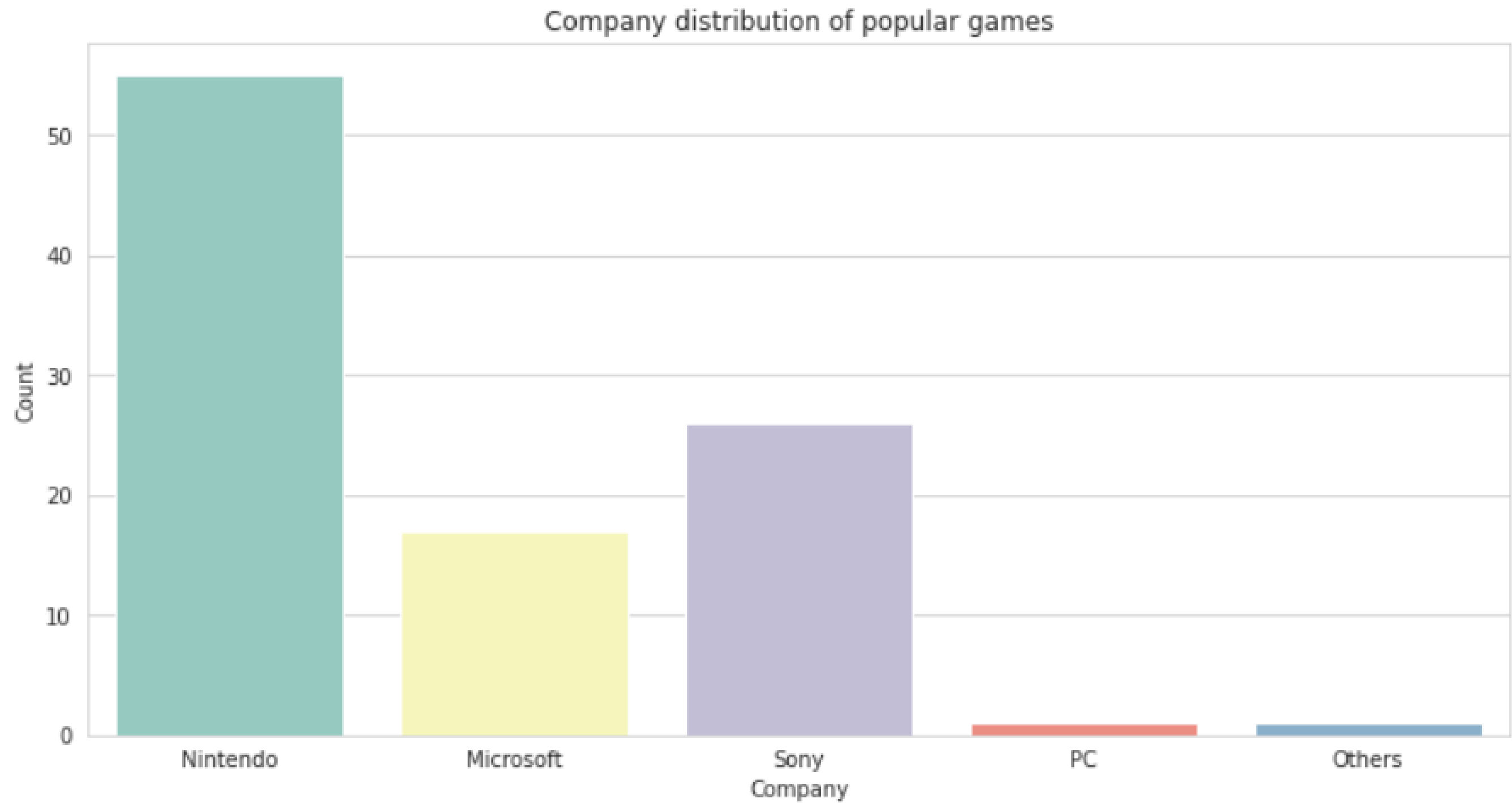
Top 20 게임에서 인기가 많은 플랫폼은 Wii이다.

코드 구현

```
plt.figure(figsize=(10,8))
sns.barplot(x='Total_Sales', y='Platform', data=top20, palette='Paired')
plt.title('Total Global Sales by Platform for Top 20 Games', fontsize=16)
plt.xlabel('Global Sales (millions)', fontsize=12)
plt.ylabel('Platform', fontsize=12)
plt.show()
```

인기많은 게임 Top20을 기준으로 가장 많이 팔린 플랫폼순으로 시각화하였다.

인기많은 회사



가장 인기가 많은 회사는 Nintendo이다.

코드 구현

```
plt.figure(figsize=(12,6))
sns.countplot(x='Company', data=popular_games, palette='Set3')
plt.title('Company distribution of popular games')
plt.xlabel('Company')
plt.ylabel('Count')
plt.show()
```

인기많은 게임 Top20을 기준으로 가장 많이 팔린 회사순으로 시각화하였다.

Nintendo



**ACTION, SHOOTER 장르의
게임이 인기가 많다.**

결론

가장 잘 팔리는 Nintendo의 최신 플랫폼에
Action, Shooter 장르의 게임을 출시한다.

BUT

최신 인사이트를 포함하면
연도별로 증가하는 Sony(PS), Microsoft(Xbox) 를 고려하여,
Action, Shooter, Adventure 장르의 AAA급 게임을
Cross Platform으로 출시하는 게 가장 좋다고 판단된다.



감사합니다

THANK YOU