

# Analiza danych statystycznych dotyczących sprzedaży gier w latach 2005 - 2016

Marcin Mikula

## Wczytywanie danych

```
In [2]: import pandas as pd
import sqlite3
import seaborn as sns
pd.plotting.register_matplotlib_converters()
import matplotlib.pyplot as plt
```

Przypisanie danych do zmiennej.

```
In [4]: with sqlite3.connect('E:\Projekty\RPIS\projekt.db') as conn:
dg = pd.read_sql('select * from games ', conn)
dg.isnull().sum()
```

Czyszczenie danych zawierających null w sqlite i duplikatów w python.

```
In [12]: # DELETE
# FROM [games]
# WHERE Year LIKE '%N/A%'

len(dg)-len(dg.drop_duplicates())
dg.dropna(inplace=True)
dg['Year']=dg['Year'].astype(int)
```

```
In [25]: dg.head()
```

```
Out[25]:
```

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other
0	1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	
1	2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	
2	3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	
3	4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	
4	5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	

```
In [26]: dg.tail()
```

Out[26]:	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	
	16322	Woody Woodpecker in Crazy Castle 5	GBA	2002	Platform	Kemco	0.01	0.00	0.0	
	16323	Men in Black II: Alien Escape	GC	2003	Shooter	Infogrames	0.01	0.00	0.0	
	16324	SCORE International Baja 1000: The Official Game	PS2	2008	Racing	Activision	0.00	0.00	0.0	
	16325	Know How 2	DS	2010	Puzzle	7G//AMES	0.00	0.01	0.0	
	16326	Spirits & Spells	GBA	2003	Platform	Wanadoo	0.01	0.00	0.0	

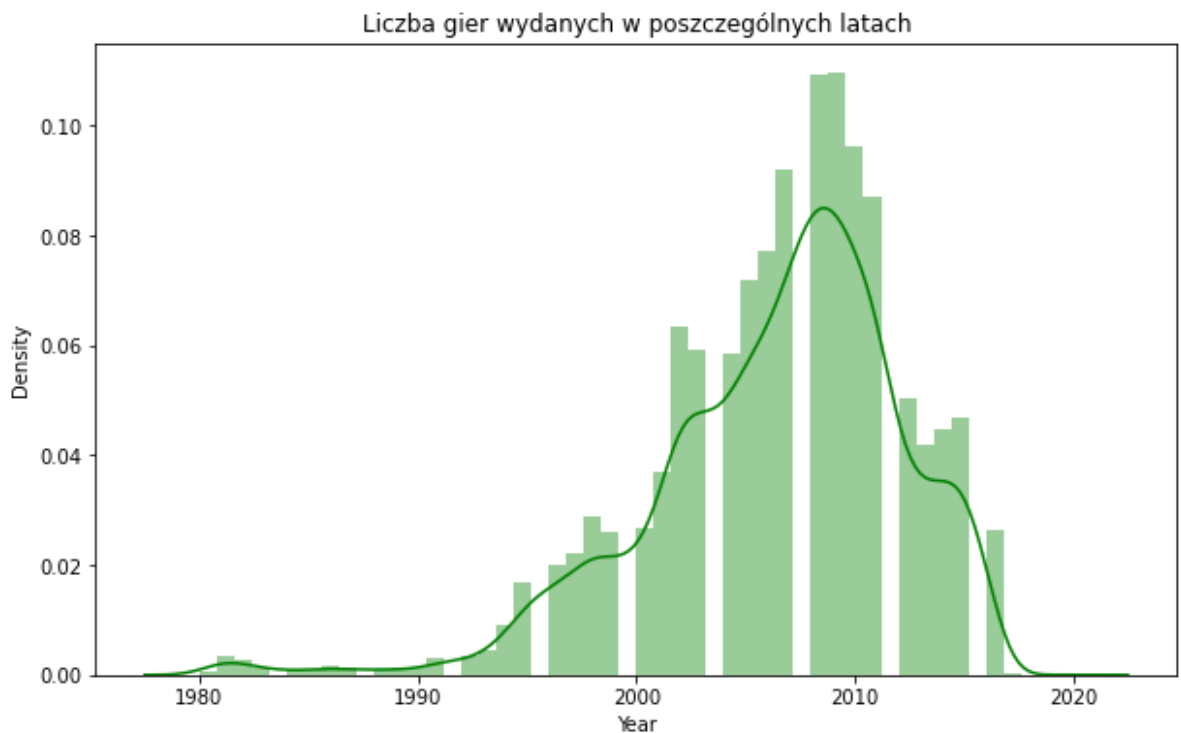
## Analiza danych dotyczących wydawania gier

Można zauważyć, że ilość gier wydawanych na przestrzeni lat osiągnęła największy przyrost w latach 2007-2010. Taki stan rzeczy bezpośrednio koreluje ze zwiększeniem konkurencyjności na rynku konsol (premiera PlayStation 3 oraz Xbox 360), oraz zwiększającym się zainteresowaniem E-sportem.

```
In [24]: ax = plt.figure(figsize=(10,6))
sns.distplot(dg['Year'],color='green')
plt.title('Liczba gier wydanych w poszczególnych latach')
avg = dg['Global_Sales'].mean()
median = dg['Global_Sales'].median()
plt.show()
print("Średnia:", round(avg * 10**5,0), " Mediana:", round(median * 10**5, 0))
```

C:\Python310\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

```
warnings.warn(msg, FutureWarning)
```



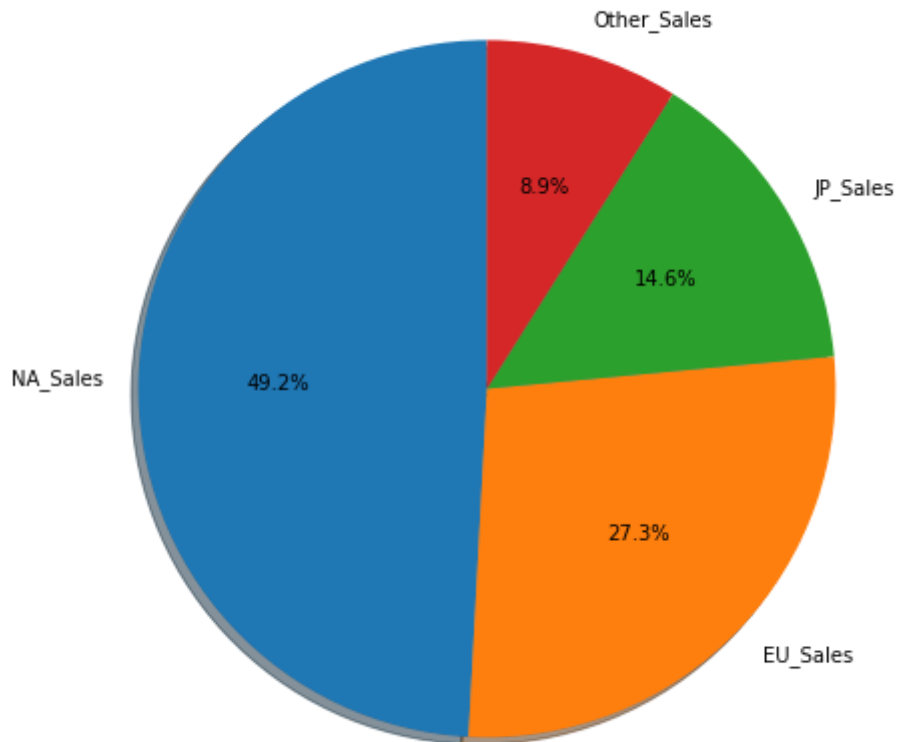
## Sprzedaż gier

Rynek gier komputerowych rozwija się bardzo dynamicznie, opracowane dane statystyczne wskazują, że zdecydowanym liderem sprzedaży gier na świecie są kraje Ameryki Północnej, pokrywając prawie 50% światowego zapotrzebowania.

```
In [37]: top_region = dg[['NA_Sales', 'EU_Sales', 'JP_Sales', 'Other_Sales']]
top_region = top_region.sum().reset_index()
top_region = top_region.rename(columns={"index": "region", 0: "sprzedaż"})

labels = top_region['region']
sizes = top_region['sprzedaż']
plt.figure(figsize=(10, 8))
plt.pie(sizes, labels=labels, autopct='%1.1f%%', shadow=True, startangle=90)
plt.title("Zestawienie sprzedaży gier w regionach świata")
plt.show()
```

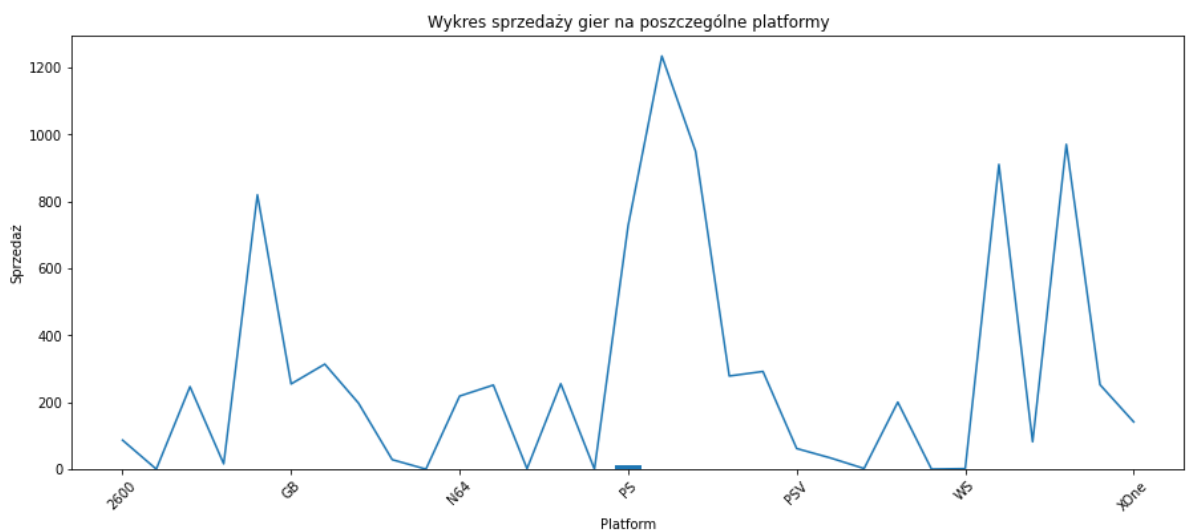
### Zestawienie sprzedaży gier w regionach świata



Dane dotyczące konsol na które tworzone są gry, ewidentnie wskazują na dużą przewagę PlayStation nad konkurencją. Taki stan rzeczy jest wynikiem bardzo długiej obecności tej konsoli na rynku, w tym czasie PlayStation zbudowało solidną renomę, dostarczając swoim odbiorcom ekskluzywne gry niedostępne na innych platformach w konkurencyjnych cenach.

In [8]:

```
ax = plt.figure(figsize=(15,6))
dg.groupby(['Platform'])['Global_Sales'].sum().plot().bar(15, 10)
plt.xticks(rotation=45)
plt.ylabel('Sprzedaż mln')
plt.title('Wykres sprzedaży gier na poszczególne platformy')
plt.show()
```

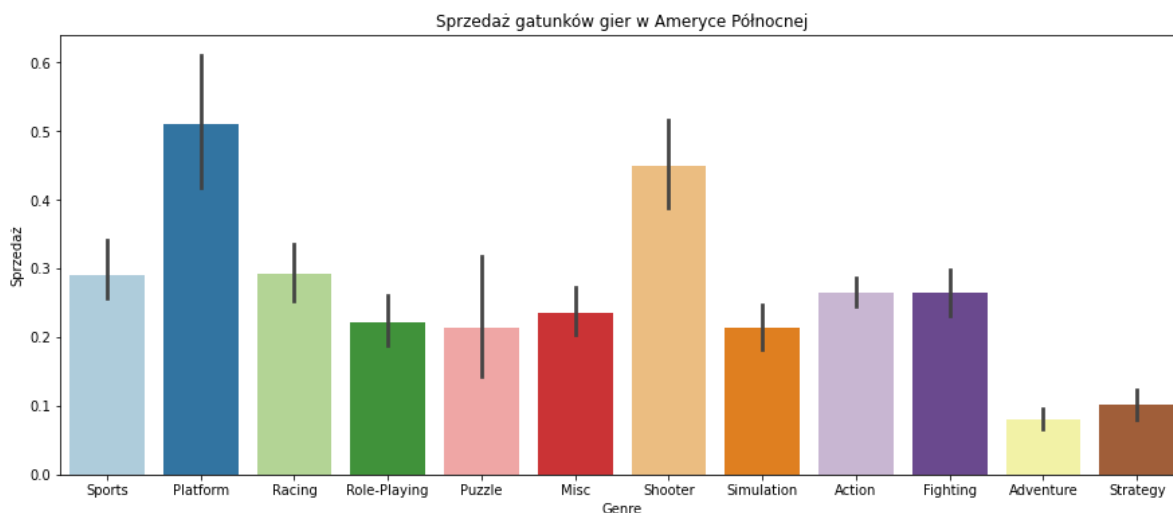


Poniżej opracowane zostały wyniki dotyczące sprzedaży poszczególnych gatunków gier w badanych regionach. Łatwo zauważyć, że w krajach Europy i Ameryki Północnej rozkład

wyglądają niemalże identycznie. Znaczna różnica pojawia się w zestawieniu tych regionów z Japonią. Bardzo popularne shootery na rynku Japońskim stanowią zaledwie mały ułamek w porównaniu z Europejskim i Amerykańskim, za to dużym zainteresowaniem wyróżniają się gry typu role-playing. Z całą pewnością można stwierdzić, że jest to wynikiem różnic kulturowych, oraz podejścia społeczności Azjatyckiej do gier w których pojawia się przemoc.

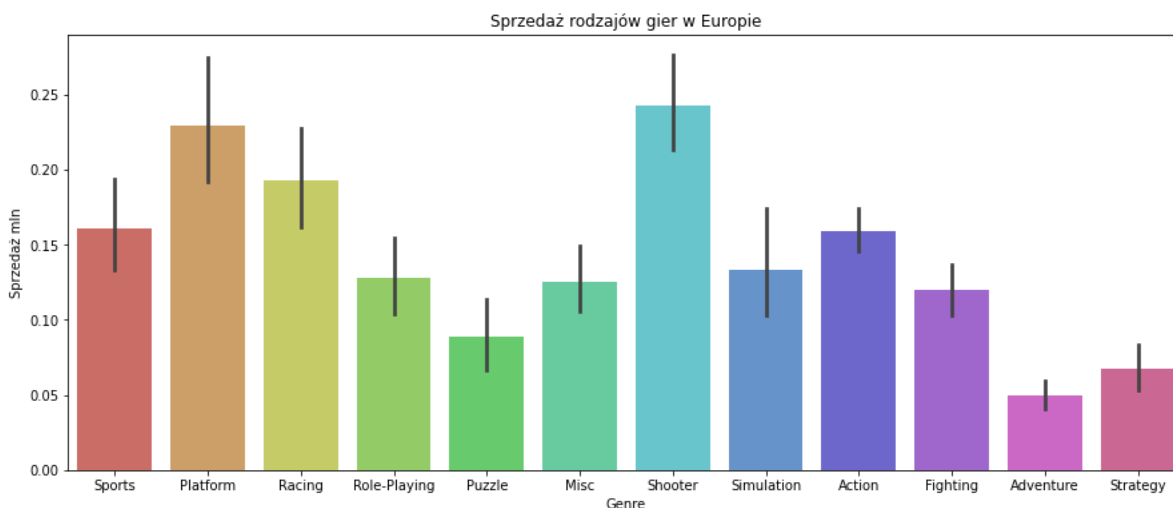
In [13]:

```
ax1=plt.figure(figsize=(15,6))
sns.barplot(x='Genre', y='NA_Sales', data=dg , palette='Paired')
plt.ylabel("Sprzedaż mln")
plt.title("Sprzedaż gatunków gier w Ameryce Północnej")
plt.show()
```



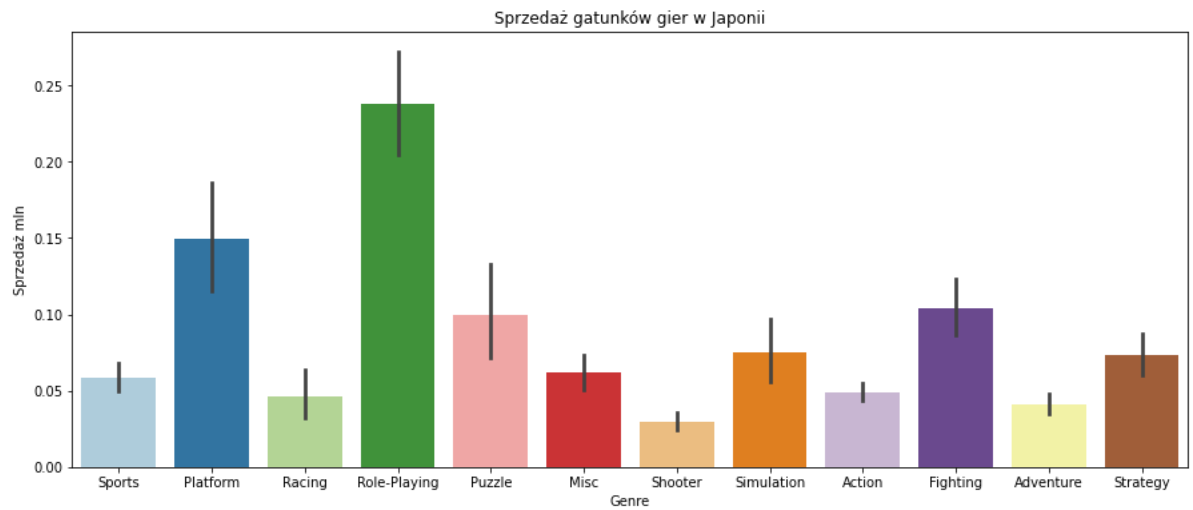
In [16]:

```
ax2=plt.figure(figsize=(15,6))
sns.barplot(x='Genre', y='EU_Sales', data=dg , palette='hls')
plt.ylabel("Sprzedaż mln")
plt.title("Sprzedaż rodzajów gier w Europie")
plt.show()
```



In [33]:

```
ax3=plt.figure(figsize=(15,6))
sns.barplot(x='Genre', y='JP_Sales', data=dg , palette='Paired')
plt.ylabel("Sprzedaż mln")
plt.title("Sprzedaż gatunków gier w Japonii")
plt.show()
```



## Podsumowanie analizy

Analizowane dane pokazują jak duży i zróżnicowany jest światowy rynek gier. Szczególny wpływ na zainteresowanie tą dziedziną ma rywalizacja w świecie konsol. Lata 2007-2010 były przełomowe w rozwoju rynku gier, odnotowany został wtedy największy przyrost zainteresowania konsolami oraz grami. Od tamtego czasu cała branża prężnie i stabilnie się rozwija, przewidywania na rok 2022 wskazują, że wartość rynku przekroczy 50 miliardów dolarów.