KOVAI.CO

DATE: 18/10/2023

NAME : PRAGADEESHWARAN K J

ROLLNO: 20BIS031

COLLEGE: KUMARAGURU COLLEGE OF TECHNOLOGY (KCT)

TASK-2

Video Game Sales Data set

Analyse and Insights

GithubLink:

[Chat-Analysis-ML/Untitled1.ipynb at main · Prag-deesh/Chat-Analysis-ML (github.com)](https://github.com/Prag-deesh/Chat-Analysis-ML/blob/main/Untitled1.ipynb)

Used libraries

pandas

matplotlib.pyplot

seaborn

dataframes describe and info

df.head()

df.info()

df.describe

It tells about the dataframes

Plotting global hit games

plt.figure(figsize=(8, 6))

sns.histplot(data=df, x='Game Title', y='Global', bins=20,

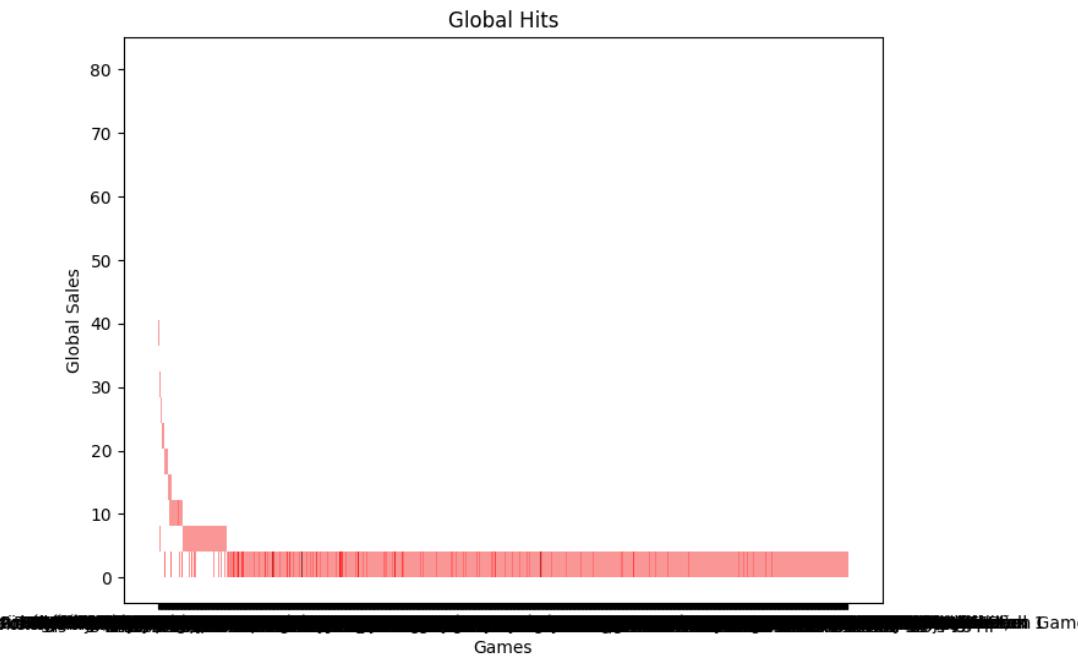
color='red',kde=True)

plt.title('Global Hits')

plt.xlabel('Games')

plt.ylabel('Global Sales')

plt.show()



Filtering and printing successful games based on global sales vales

success\_threshold = 1.0

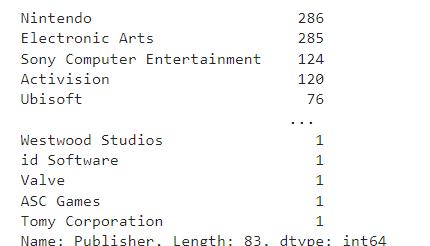
successful\_games = df[df['Global'] >= success\_threshold] successful\_games



Filtering and printing successful genres

succussful\_genres = successful\_games['Genre'].value\_counts()

succussful\_genres



Filtering and printing successful publishers

successful\_publisher = successful\_games['Publisher'].value\_counts() successful\_publisher



Correlation Matrix

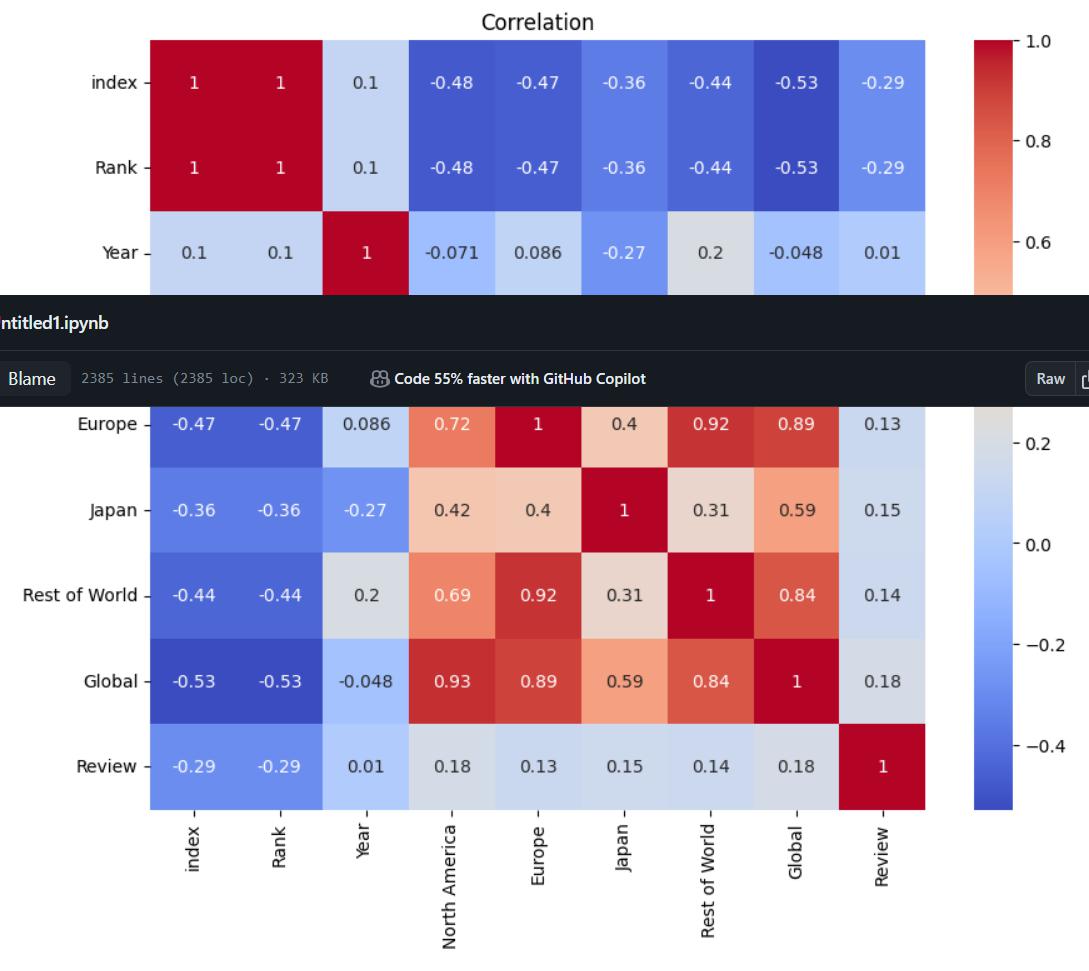
corr\_matrix = df.corr()

plt.figure(figsize=(10, 8))

sns.heatmap(corr\_matrix, annot=True, cmap='coolwarm')

plt.title('Correlation')

plt.show()



Categorising sales based on regions

total\_na = df['North America'].sum()

total\_eu = df['Europe'].sum()

total\_jpn = df['Japan'].sum()

total\_other = df['Rest of World'].sum()

In [124]:

regions = ['North America','Europe','Japan','Rest of World'] sales = [total\_na,total\_eu,total\_jpn,total\_other]

In [125]:

plt.figure(figsize=(10,6))

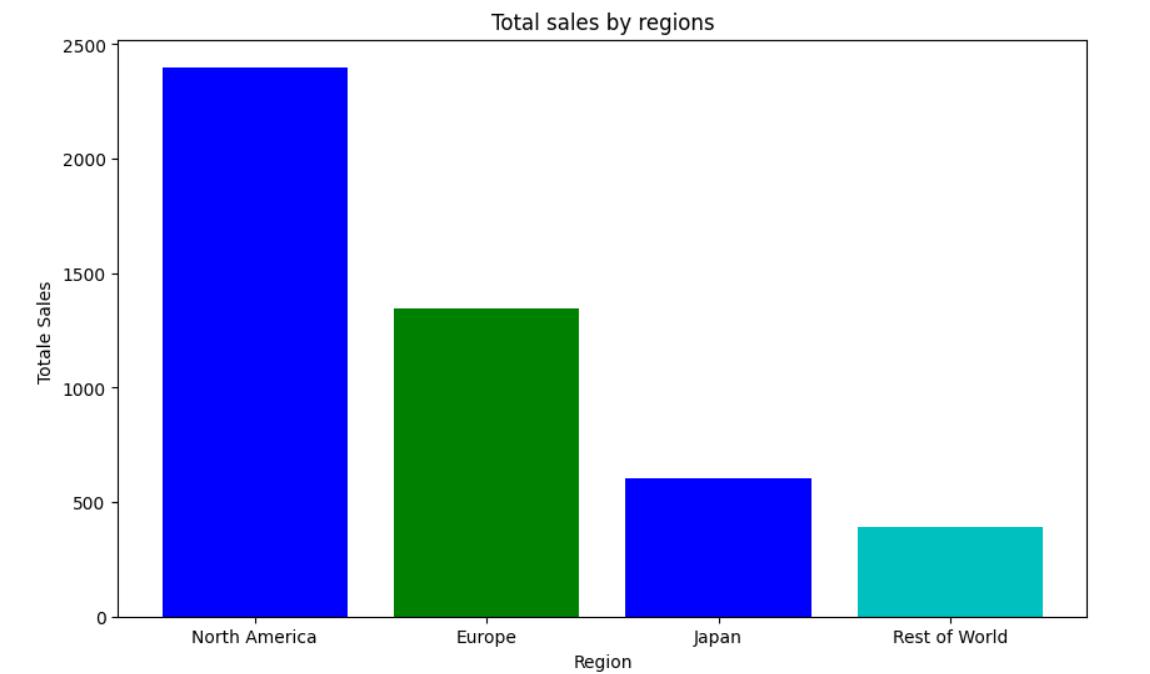
plt.bar(regions,sales,color=['b','g','b','c'])

plt.title("Total sales by regions")

plt.xlabel('Region')

plt.ylabel('Totale Sales ')

plt.show()



Categorising global sales over years

yearly\_sales = df.groupby('Year')

df['Global'].sum()

plt.figure(figsize=(10,6))

plt.plot(yearly\_sales.index,yearly\_sales.values,marker='o',linestyle='-')

plt.title('Global trend over the Years')

plt.xlabel('Year')

plt.ylabel('Global in years')

plt.grid=True

plt.xticks(rotation=45)

plt.show()

Distribution of games through genres

genre\_dist = df['Genre'].value\_counts()

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

plt.pie(genre\_dist, labels=genre\_dist.index, startangle=140)

plt.title('Distribution of Game Genres')

plt.show()

