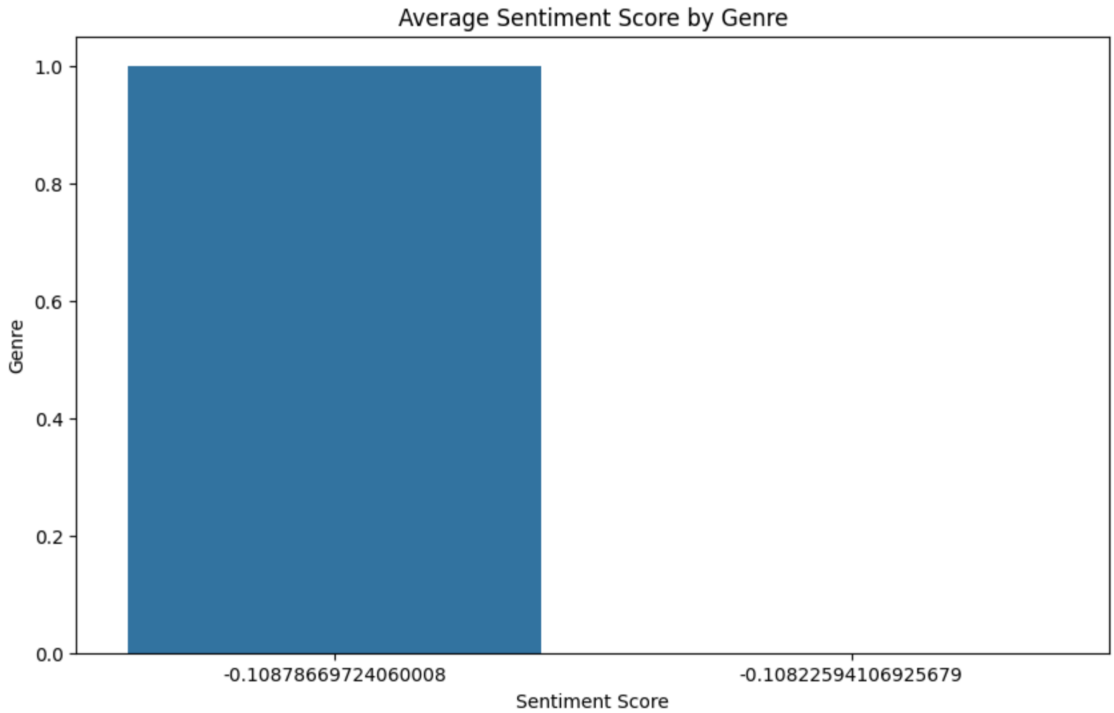
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

data=pd.read_csv('IMDB-Movie-Data.csv')

# Remove rows with missing revenue (target variable)
data=data.dropna(subset=['Revenue (Millions)'])

# Optional: Fill or drop other NaNs
# Instead of filling with empty strings, fill with a numeric value (e.g., 0)
data.fillna(0, inplace=True)  # Fill all NaNs with 0

import nltk
from nltk.sentiment.vader import SentimentIntensityAnalyzer

nltk.download('vader_lexicon')
sid = SentimentIntensityAnalyzer()
data['Sentiment_Score'] = data['Description'].apply(lambda x: sid.polarity_scores(x)['compound'])

if 'Genre' in data.columns:
    data['Genre'] = data['Genre'].apply(lambda x: x.split('|')[0])  # Or one-hot encode all genres
    data = pd.get_dummies(data, columns=['Genre'], drop_first=True)
else:
    print("Column 'Genre' not found. Available columns:", data.columns)

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

# Select features and target
features = ['Rating', 'Metascore', 'Votes', 'Sentiment_Score'] + [col for col in data.columns if col.startswith('Genre_')]
X = data[features]
y = data['Revenue (Millions)']

# Convert selected features to numeric, coerce errors to NaN, and fill NaN with 0
# This is done before the train-test split to ensure consistency
for col in features:
    X[col] = pd.to_numeric(X[col], errors='coerce').fillna(0)

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Model
model = LinearRegression()
model.fit(X_train, y_train)

# Predictions and evaluation
y_pred = model.predict(X_test)
print(f"R²: {r2_score(y_test, y_pred)}")
print(f"MSE: {mean_squared_error(y_test, y_pred)}")


# Remove rows with missing revenue (target variable)
data=data.dropna(subset=['Revenue (Millions)'])
for col in ['Rating', 'Metascore', 'Votes']:
    data[col] = pd.to_numeric(data[col], errors='coerce').fillna(0)
data.fillna("Unknown", inplace=True)

import seaborn as sns
import matplotlib.pyplot as plt
genre_columns = [col for col in data.columns if col.startswith('Genre_')]

# Calculate average sentiment for each genre
genre_sentiment = data.groupby(genre_columns)['Sentiment_Score'].mean()
# Reshape and sort for plotting
genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
genre_sentiment = genre_sentiment.groupby('value')['Sentiment_Score'].mean().sort_values(ascending=False)


# Plot
plt.figure(figsize=(10, 6))
sns.barplot(x=genre_sentiment.values, y=genre_sentiment.index)
plt.title('Average Sentiment Score by Genre')
plt.xlabel('Sentiment Score')
plt.ylabel('Genre')
plt.show()
```

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from nltk.sentiment.vader import SentimentIntensityAnalyzer
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.preprocessing import OneHotEncoder
import nltk
nltk.download('vader_lexicon')

data = pd.read_csv('IMDB-Movie-Data.csv')
print(data.head())
print(data.info())

# Drop missing target (revenue)
data = data.dropna(subset=['Revenue (Millions)'])

# Fill others as needed
data.fillna({'Metascore': data['Metascore'].median()}, inplace=True)
data.fillna("", inplace=True)

# Simplify genre (first listed genre only)
data['Main_Genre'] = data['Genre'].apply(lambda x: x.split(',')[0])

# Simulate basic review sentiment (real review scraping can be added later)
data['Reviews'] = data['Description']  # Using Description as proxy for reviews

# VADER Sentiment
sid = SentimentIntensityAnalyzer()
data['Sentiment'] = data['Reviews'].apply(lambda x: sid.polarity_scores(x)['compound'])

# One-hot encode Main Genre
data = pd.get_dummies(data, columns=['Main_Genre'], drop_first=True)

# Feature selection
features = ['Rating', 'Metascore', 'Votes', 'Sentiment'] + [col for col in data.columns if col.startswith('Main_Genre_')]
X = data[features]
y = data['Revenue (Millions)']

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Model
model = LinearRegression()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

# Evaluation
print("R^2 Score:", r2_score(y_test, y_pred))
print("MSE:", mean_squared_error(y_test, y_pred))

# Group and visualize
# Use the original 'Genre' column before one-hot encoding
genre_sentiment = data.groupby(data['Genre'].str.split(',').str[0])['Sentiment'].mean().sort_values()

plt.figure(figsize=(10, 6))
sns.barplot(x=genre_sentiment.values, y=genre_sentiment.index, palette='coolwarm')
plt.title("Average Sentiment by Genre")
plt.xlabel("Sentiment Score")
plt.ylabel("Genre")
plt.show()

data.to_excel("movie_analysis_output.xlsx", index=False)
```

```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data]   Package vader_lexicon is already up-to-date!
<ipython-input-2-e0e3c219fde2>:39: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-ver
  X[col] = pd.to_numeric(X[col], errors='coerce').fillna(0)
R²: 0.4520200184546934
MSE: 7151.215402650742
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
```

```
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `f
```

```
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
<ipython-input-2-e0e3c219fde2>:67: PerformanceWarning: DataFrame is highly fragmented.  This is usually the result of calling `i
  genre_sentiment = genre_sentiment.reset_index().melt(id_vars=['Sentiment_Score'], value_vars=genre_columns)
```




```
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data]   Package vader_lexicon is already up-to-date!
   Rank                    Title                     Genre  \
0     1  Guardians of the Galaxy   Action,Adventure,Sci-Fi
1     2               Prometheus  Adventure,Mystery,Sci-Fi
2     3                    Split           Horror,Thriller
3     4                     Sing   Animation,Comedy,Family
4     5            Suicide Squad  Action,Adventure,Fantasy

                                         Description             Director  \
0  A group of intergalactic criminals are forced ...           James Gunn
1  Following clues to the origin of mankind, a te...         Ridley Scott
2  Three girls are kidnapped by a man with a diag...   M. Night Shyamalan
3  In a city of humanoid animals, a hustling thea...  Christophe Lourdelet
4  A secret government agency recruits some of th...           David Ayer

                                              Actors  Year  Runtime (Minutes)  \
0  Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S...  2014                121
1  Noomi Rapace, Logan Marshall-Green, Michael Fa...  2012                124
2  James McAvoy, Anya Taylor-Joy, Haley Lu Richar...  2016                117
3  Matthew McConaughey,Reese Witherspoon, Seth Ma...  2016                108
4  Will Smith, Jared Leto, Margot Robbie, Viola D...  2016                123

   Rating   Votes  Revenue (Millions)  Metascore
0     8.1  757074              333.13       76.0
1     7.0  485820              126.46       65.0
2     7.3  157606              138.12       62.0
3     7.2   60545              270.32       59.0
4     6.2  393727              325.02       40.0
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 12 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Rank                1000 non-null   int64
 1   Title               1000 non-null   object
 2   Genre               1000 non-null   object
 3   Description         1000 non-null   object
 4   Director            1000 non-null   object
 5   Actors              1000 non-null   object
```

```
 6   Year                1000 non-null   int64
 7   Runtime (Minutes)   1000 non-null   int64
 8   Rating              1000 non-null   float64
 9   Votes               1000 non-null   int64
 10  Revenue (Millions)  872 non-null    float64
 11  Metascore           936 non-null    float64
dtypes: float64(3), int64(4), object(5)
memory usage: 93.9+ KB
None
R^2 Score: 0.5324709274020911
MSE: 6101.319788582008
<ipython-input-2-e0e3c219fde2>:138: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set

  sns.barplot(x=genre_sentiment.values, y=genre_sentiment.index, palette='coolwarm')
```