

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

file_path = '/content/IMDB-Movie-Data.xlsx'
df = pd.read_excel(file_path)

print(df.info())
print(df.describe())
print(df.head())

print(df.isnull().sum())

df['Metascore'] = df['Metascore'].fillna(df['Metascore'].mean())
df['Revenue (Millions)'] = df['Revenue (Millions)'].fillna(df['Revenue (Millions)'].mean())
df['Director'] = df['Director'].fillna('Unknown')
df['Actors'] = df['Actors'].fillna('Unknown')

df = df.dropna(subset=['Rating', 'Votes'])

df['Year'] = pd.to_datetime(df['Year'], format='%Y')

df['Genres'] = df['Genre'].str.split(',')

df_exploded = df.explode('Genres')

genres_over_time = df_exploded.groupby(['Year', 'Genres']).size().unstack().fillna(0)

plt.figure(figsize=(14,7))
genres_over_time.plot(kind='area', stacked=True, figsize=(14,7))
plt.title('Popular Genres Over Time')
plt.xlabel('Year')
plt.ylabel('Number of Movies')
plt.legend(loc='upper left')
plt.show()

plt.figure(figsize=(10,6))
sns.scatterplot(x='Rating', y='Votes', data=df)
plt.title('Correlation Between Ratings and Votes')
plt.xlabel('Rating')
plt.ylabel('Votes')
plt.show()

avg_ratings = df.groupby(df['Year'].dt.year)['Rating'].mean()

plt.figure(figsize=(12,6))
avg_ratings.plot(kind='line')
plt.title('Average Movie Ratings Over the Years')
plt.xlabel('Year')
plt.ylabel('Average Rating')
plt.grid(True)
plt.show()

top_directors = df.groupby('Director')['Rating'].mean().sort_values(ascending=False).head(10)
plt.figure(figsize=(12,6))
top_directors.plot(kind='bar')
plt.title('Top 10 Directors by Average Rating')
plt.xlabel('Director')
plt.ylabel('Average Rating')
plt.xticks(rotation=45)
plt.show()

movies_per_year = df['Year'].dt.year.value_counts().sort_index()

plt.figure(figsize=(12,6))
movies_per_year.plot(kind='bar')
```

```
plt.title('Number of Movies Released Each Year')
plt.xlabel('Year')
plt.ylabel('Number of Movies')
plt.grid(True)
plt.show()

actor_list = df['Actors'].str.split(',').explode('Actors')
top_actors = actor_list.value_counts().head(10)

plt.figure(figsize=(12,6))
top_actors.plot(kind='bar')
plt.title('Top 10 Popular Actors')
plt.xlabel('Actors')
plt.ylabel('Number of Movies')
plt.xticks(rotation=45)
plt.show()
```

```
<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
```

	Rank	Year	Runtime (Minutes)	Rating	Votes
count	1000.000000	1000.000000	1000.000000	1000.000000	1.000000e+03
mean	500.500000	2012.783000	113.172000	6.723200	1.698083e+05
std	288.819436	3.205962	18.810908	0.945429	1.887626e+05
min	1.000000	2006.000000	66.000000	1.900000	6.100000e+01
25%	250.750000	2010.000000	100.000000	6.200000	3.630900e+04
50%	500.500000	2014.000000	111.000000	6.800000	1.107990e+05
75%	750.250000	2016.000000	123.000000	7.400000	2.399098e+05
max	1000.000000	2016.000000	191.000000	9.000000	1.791916e+06

	Revenue (Millions)	Metascore
count	872.000000	936.000000
mean	82.956376	58.985043
std	103.253540	17.194757
min	0.000000	11.000000
25%	13.270000	47.000000
50%	47.985000	59.500000
75%	113.715000	72.000000
max	936.630000	100.000000

	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

Rank	0
Title	0
Genre	0
Description	0
Director	0
Actors	0
Year	0
Runtime (Minutes)	0
Rating	0
Votes	0
Revenue (Millions)	128
Metascore	64

```
dtype: int64
```

```
<Figure size 1400x700 with 0 Axes>
```

Popular Genres Over Time





