

```
In [2]: import pandas as pd
import sqlite3 as sql
```

```
In [3]: df = pd.read_csv('IMDB_Movie_Data.csv')
```

```
In [4]: df.head()
```

Out[4]:

| | Rank | Title | Genre | Description | Director | Actors | Year | Runtime (Minutes) | Rating | Votes | Revenue (Millions) | Metascore |
|---|------|-------------------------|--------------------------|---|----------------------|---|------|-------------------|--------|--------|--------------------|-----------|
| 0 | 1 | Guardians of the Galaxy | Action,Adventure,Sci-Fi | A group of intergalactic criminals are forced ... | James Gunn | Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S... | 2014 | 121 | 8.1 | 757074 | 333.13 | 76.0 |
| 1 | 2 | Prometheus | Adventure,Mystery,Sci-Fi | Following clues to the origin of mankind, a te... | Ridley Scott | Noomi Rapace, Logan Marshall-Green, Michael Fa... | 2012 | 124 | 7.0 | 485820 | 126.46 | 65.0 |
| 2 | 3 | Split | Horror,Thriller | Three girls are kidnapped by a man with a diag... | M. Night Shyamalan | James McAvoy, Anya Taylor-Joy, Haley Lu Richar... | 2016 | 117 | 7.3 | 157606 | 138.12 | 62.0 |
| 3 | 4 | Sing | Animation,Comedy,Family | In a city of humanoid animals, a hustling thea... | Christophe Lourdelet | Matthew McConaughey,Reese Witherspoon, Seth Ma... | 2016 | 108 | 7.2 | 60545 | 270.32 | 59.0 |
| 4 | 5 | Suicide Squad | Action,Adventure,Fantasy | A secret government agency recruits some of th... | David Ayer | Will Smith, Jared Leto, Margot Robbie, Viola D... | 2016 | 123 | 6.2 | 393727 | 325.02 | 40.0 |

```
In [5]: df.isnull().sum()
```

Out[5]:

| | |
|--------------------|-------|
| 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 |

```
In [6]: df.info()
```

```
<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
```

```
In [7]: df['Revenue (Millions)'] = df['Revenue (Millions)'].fillna(round(df['Revenue (Millions)'].mean(), 2))
df['Metascore'] = df['Metascore'].fillna(round(df['Metascore'].mean(), 2))
```

```
In [8]: df.isnull().sum()
```

```
Out[8]: Rank                0
Title                    0
Genre                    0
Description              0
Director                 0
Actors                   0
Year                     0
Runtime (Minutes)       0
Rating                   0
Votes                    0
Revenue (Millions)      0
Metascore                0
dtype: int64
```

```
In [16]: conn = sql.connect('Project 12 (SQL).db')
```

```
In [17]: df.to_sql('IMDB_Movie_Data', conn, if_exists='replace')
conn.close()
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

File

Owner DB

Run

Export

Import

Client

Project 12 (SQL).db

0.1.3 beta

Table

IMDB_Movie_Data

Project 12 (Sql).db

Project 12 (Sql).db

SQLite

MariaDB

PostgreSQL

MS SQL

Project 12 (SQL).db

SQL

1 --SELECT * FROM IMDB_Movie_Data;

2 -- Top 5 directors having ratings more than average?

3 SELECT DISTINCT director, rating

4 FROM IMDB_Movie_Data

5 WHERE rating > (SELECT avg(rating) FROM IMDB_Movie_Data)

6 ORDER BY rating DESC

7 LIMIT 5;

Director

Rating

Christopher Nolan

9

Christopher Nolan

8.8

Nitesh Tiwari

8.8

Christopher Nolan

8.6

Makoto Shinkai

8.6

https://sqliteonline.com

1/1