Analysis of films performances at the box office.

The analysis gives insights on success of films in the film industry by examinigng Data from films datasets.

Business Understanding

The company wants to embrace the film industry and is planning to build a new movie studio. Using data from IMDB and Movie Budget,insights have been made on how films produced in the studio can be succesful.

Data Understanding

Data Sources

- IMDB

- Movie Budgets

```python
import sqlite3
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df_1 = pd.read_csv('tn.movie_budgets.csv')
path = 'im.db'
conn = sqlite3.connect(path)
```

Data Cleaning

We clean the datasets before manipulating by :

-Checking for missing values

-Dropping duplicates

- Converting numeric values Datatype

```python
#Checking for missing values
df_1.isna().sum()
```

```
id                    0
release_date          0
movie                 0
production_budget     0
domestic_gross        0
worldwide_gross       0
dtype: int64
```

```python
#Dropping duplicates
df_1.drop_duplicates(inplace = True)
```
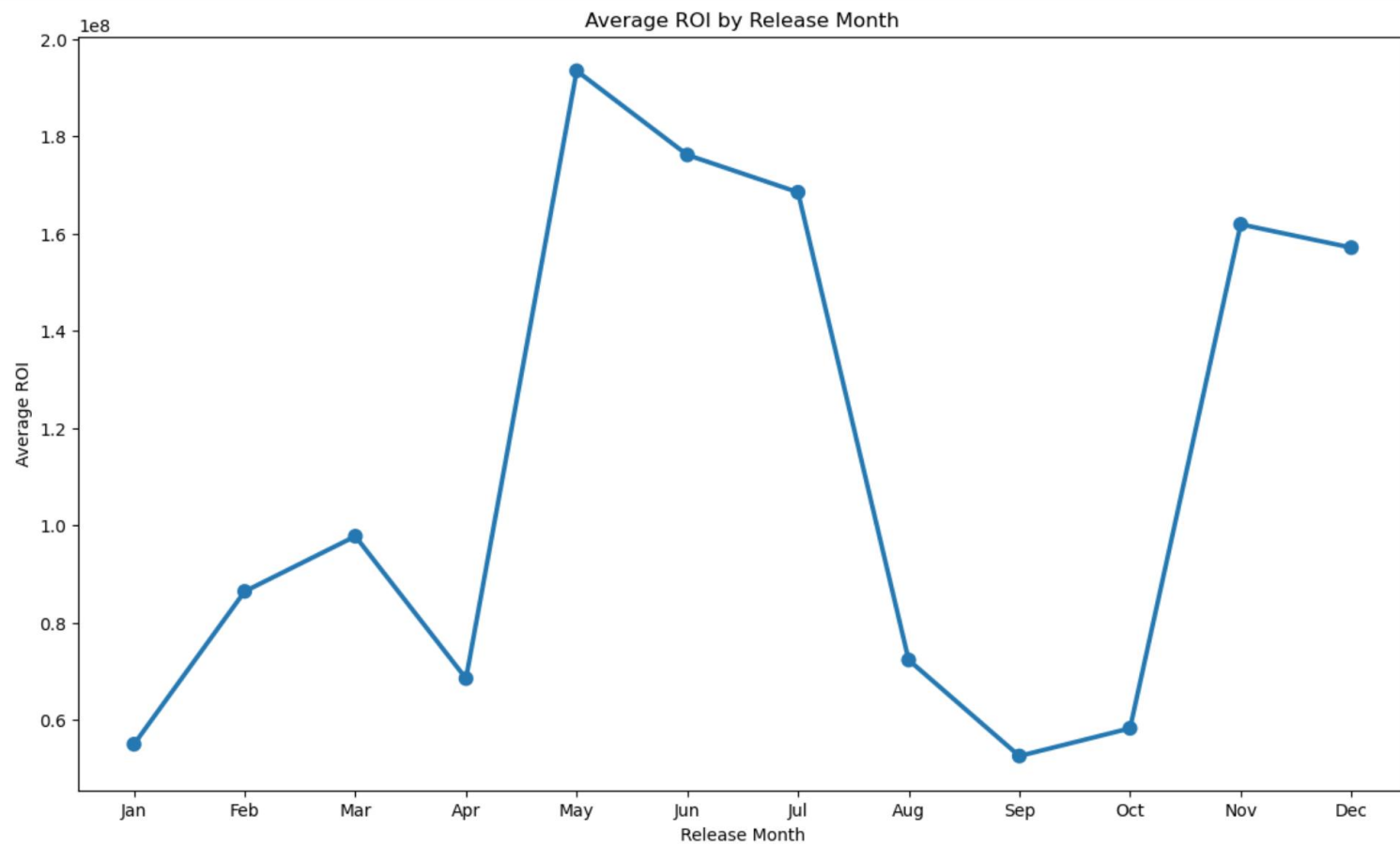
```python
#Removing non numeric signs
df_1['production_budget'] = df_1['production_budget'
df_1['domestic_gross'] = df_1['domestic_gross'].repl
df_1['worldwide_gross'] = df_1['worldwide_gross'].re
```

```python
#Changing the numeric values to float datatype.
df_1['production_budget'] = df_1['production_budget'
df_1['domestic_gross'] = df_1['domestic_gross'].asty
df_1['worldwide_gross'] = df_1['worldwide_gross'].as
```

Data Analysis

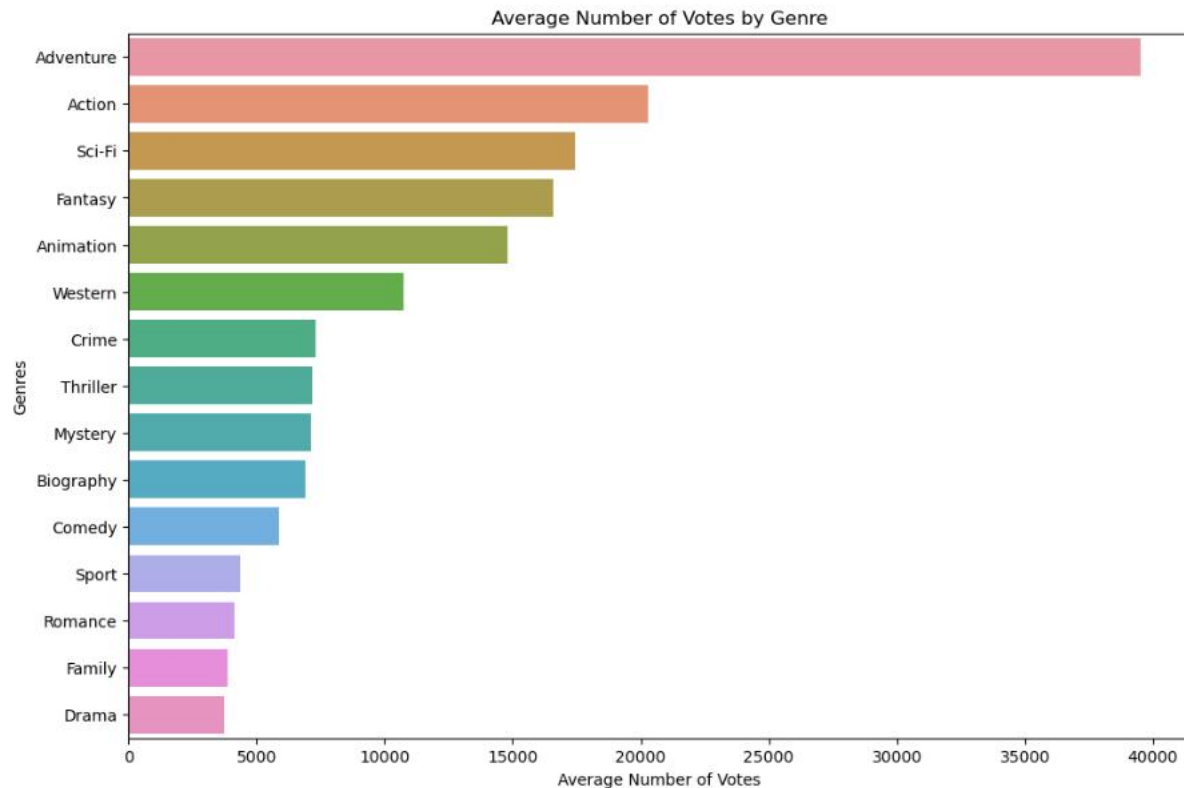- Analysis of Months and Return Of Investment

Certain months have high ROI while others have low.

It is important to take note of months with high ROI as the best release time for the movies. Months between April and August have a relatively high ROI.
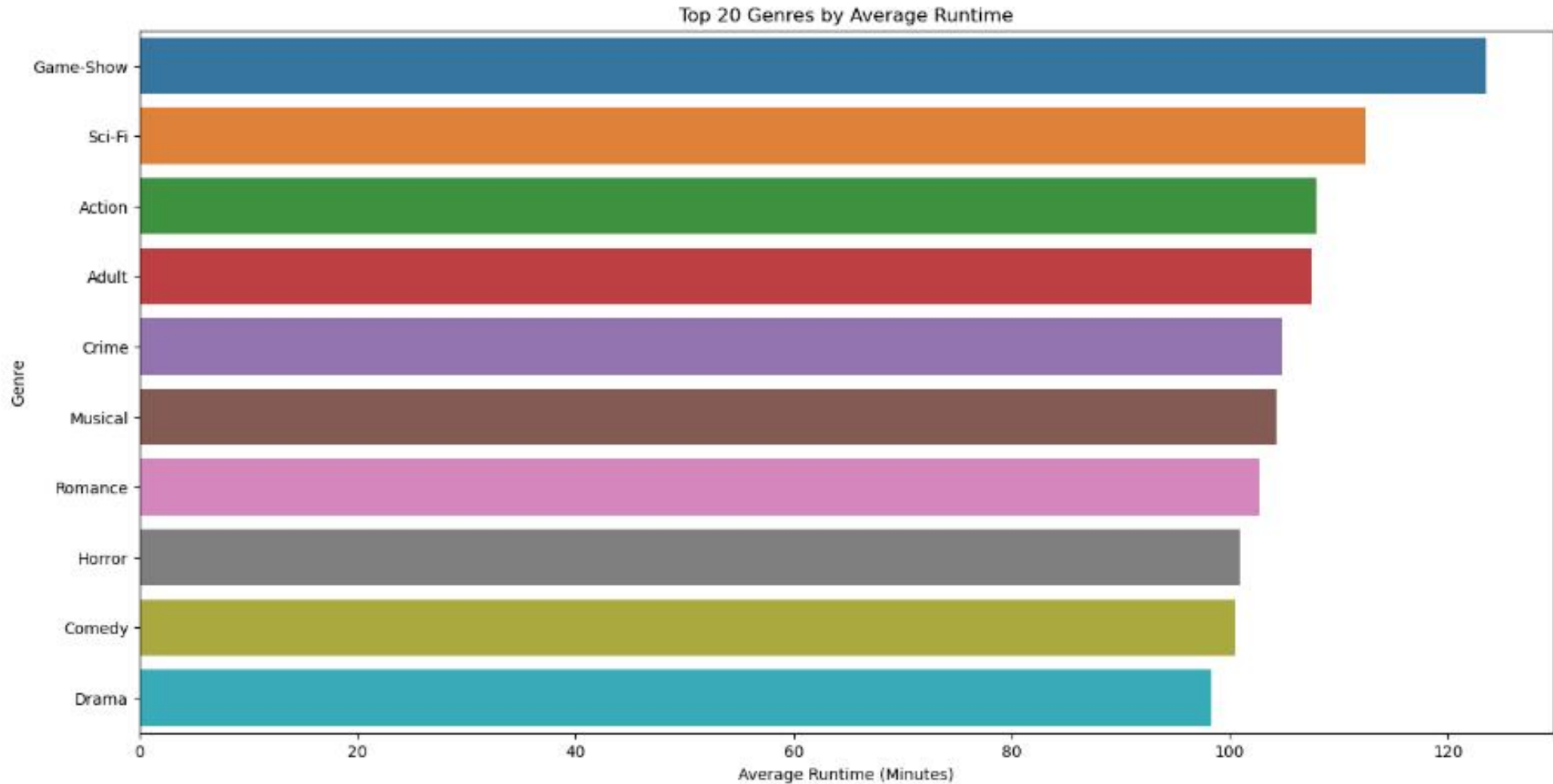
Average ROI by Release Month

.Analysis of popular genre based on number of votes

According to the insights ,certain genres are popular with the audience than others. Adventure,Action,Sci Fi and Fantasy films are among the top most popular genres in the box office.

Average Number of Votes by Genre

• Analysis of runtime of each genre

It is important to understand runtime of different genres.The runtime allocated to films matters a lot as shorter films are cheaper to produce and easy to edit.

Top 20 Genres by Average Runtime

Conclusion and Recommendation.

Months between April and August recorded the highest ROI.It is recommended that the studio should release films beteween this months to achieve high Return On Investment.

According to the distribution, the genres which had the most number of votes were :

Adventure,Action,Sci Fi and Fantasy.

The company should invest in producing movies with popular genres for good perfomance at the box office.

.The company should consider producing movies with short runtime as they are cheaper for production compared to longer running time which are expensive and take more time to edit.