# MOVIE DATABASE ANALYSIS

ANNBELLAH NDUTA MBUNGU

# OVERVIEW

- Business Understanding
- Data Understanding
- Data Analysis
- Recommendations

# Business Understanding

The movie industry is complex and involves various stakeholders. Success depends on several factors, including script quality, star power, marketing, release date, and competition.

To analyze the industry, we need data from multiple sources and use data science techniques. A combination of quantitative and qualitative analysis techniques and domain expertise is necessary to make informed decisions for producing successful movies and improving the industry's profitability.

# The First Dataset

Name: tn.movie_budgets.csv

Type: csv

Time frame of the data: 1915 to 2020

Total data points : 5782 records

Data columns (total 6 columns):

| # | Column | Dtype |
| --- | ------ | ----- |
| 0 | id | int64 |
| 1 | release_date | object |
| 2 | movie | object |
| 3 | production_budget | object |
| 4 | domestic_gross | object |
| 5 | worldwide_gross | object |

dtypes: int64(1), object(5)

# Data Understanding

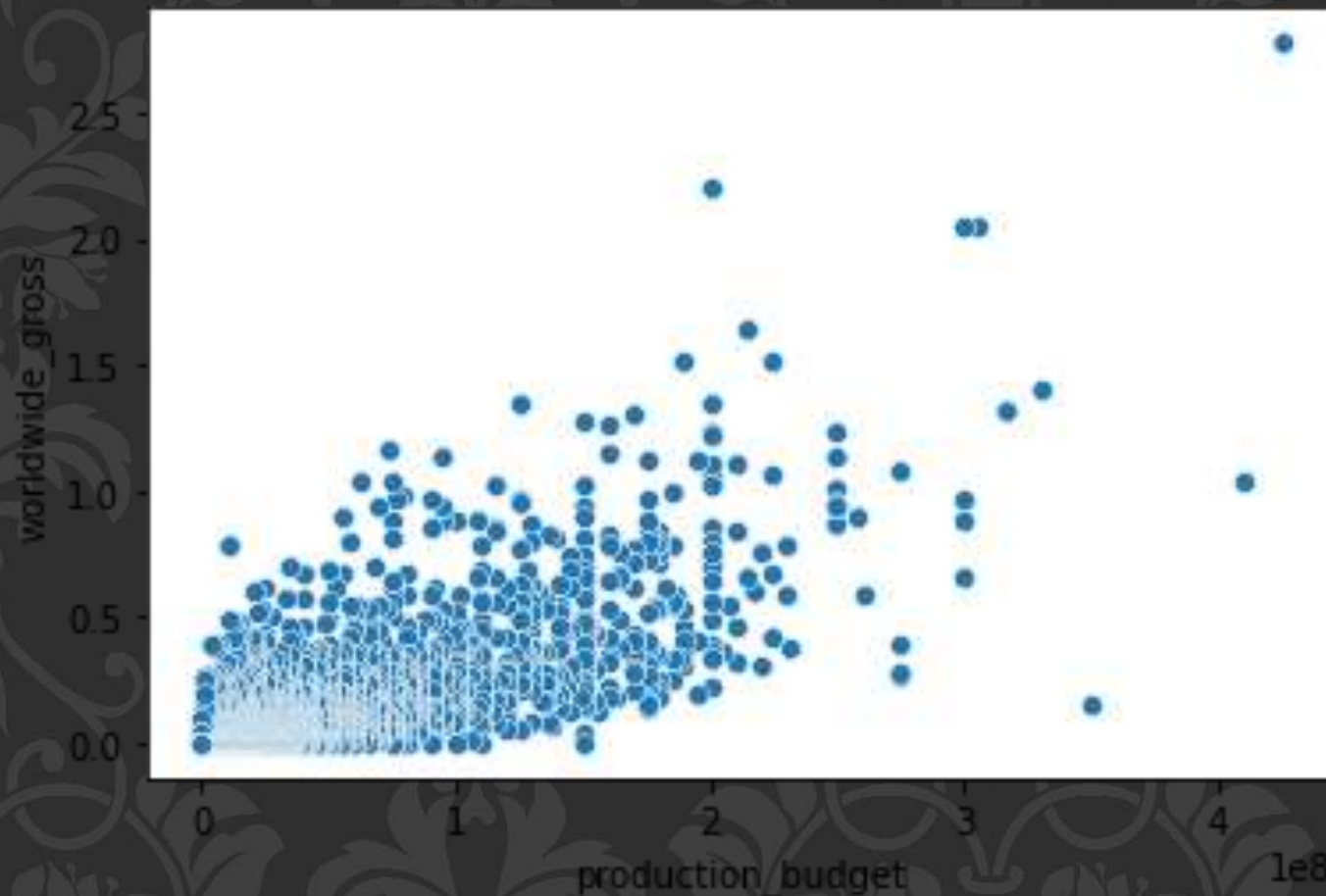Name: im.db

Type: sqlite database

| Table Names | columns |
|---|---|
| 0 movie_basics | (movie_id, primary_title, original_title, start_year, runtime_minutes, genres) |
| 1 directors | (movie_id, person_id) |
| 2 known_for | (person_id, movie_id) |
| 3 movie_akas | (movie_id, ordering, title, region, language, types, attributes, is_original_title) |
| 4 movie_ratings | (movie_id, averagerating, numvotes) |
| 5 persons | (person_id, primary_name, birth_year, death_year, primary_profession) |
| 6 principals | (movie_id, ordering, person_id, category, job, characters) |
| 7 writers | (movie_id, person_id) |

# WHAT IS THE RELATIONSHIP BETWEEN PRODUCTION BUDGET AND WORLDWIDE GROSS
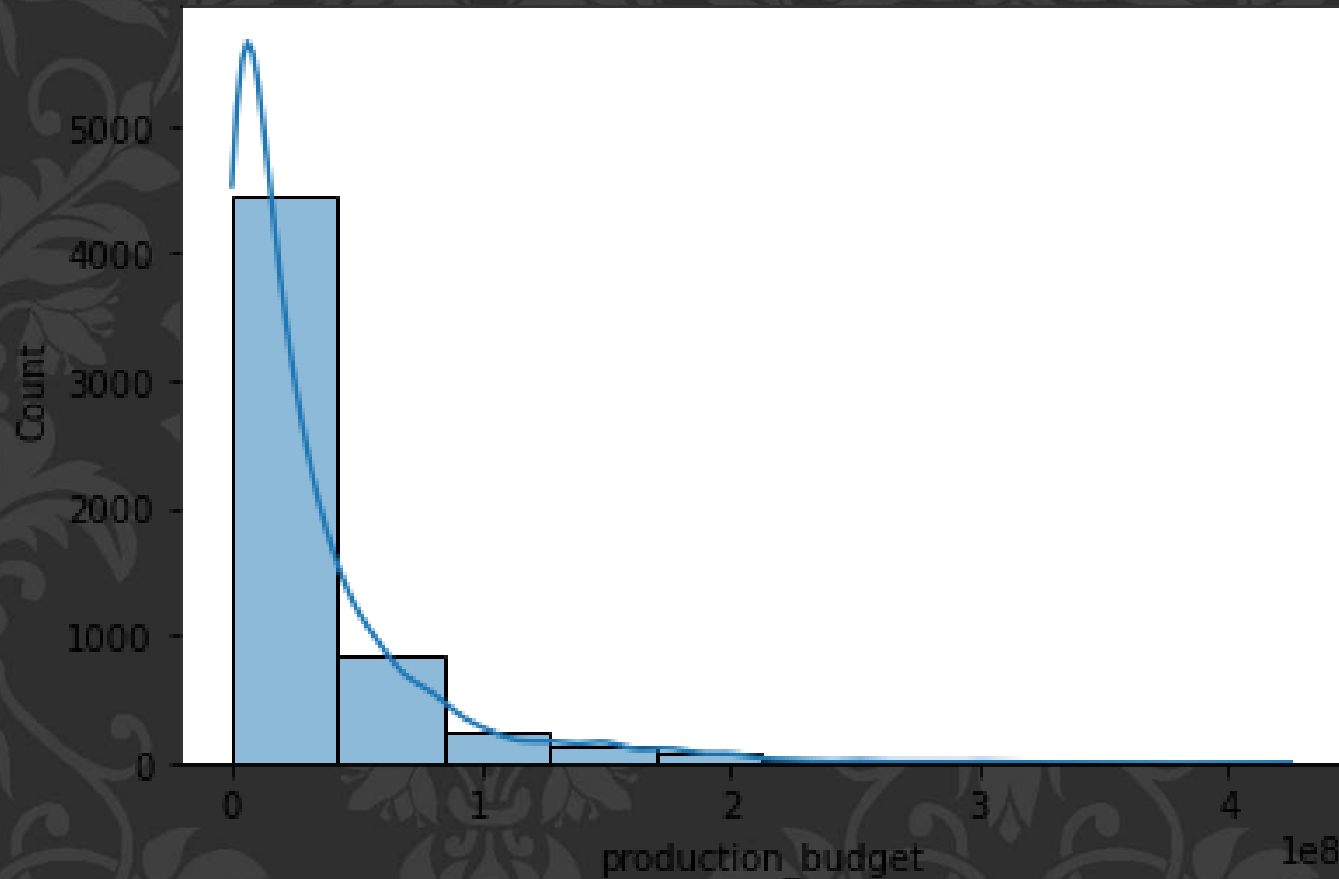


Correlation between production budget and worldwidegross

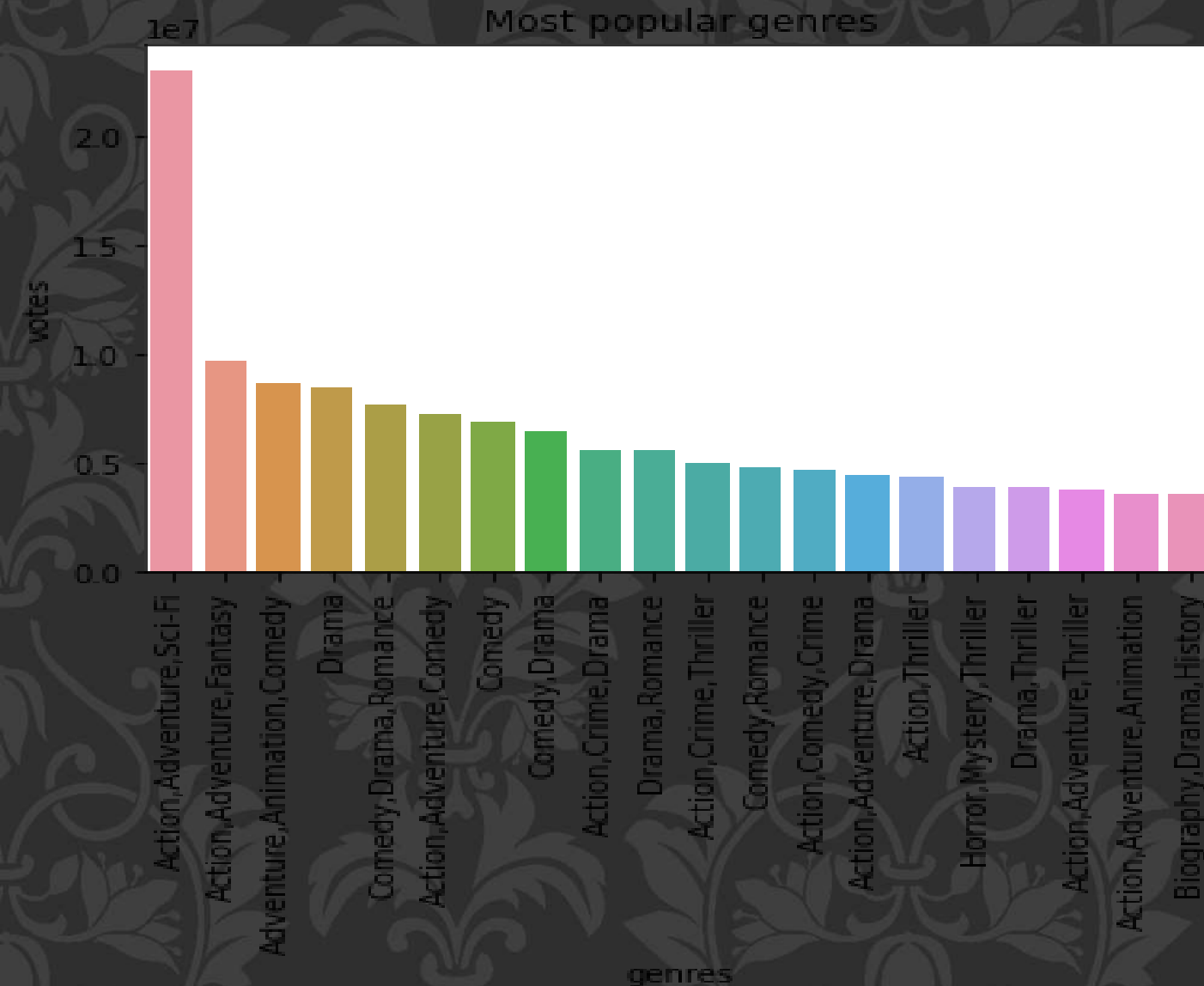There is a positive relationship between production budget and worldwide gross

Histogram of production_budget

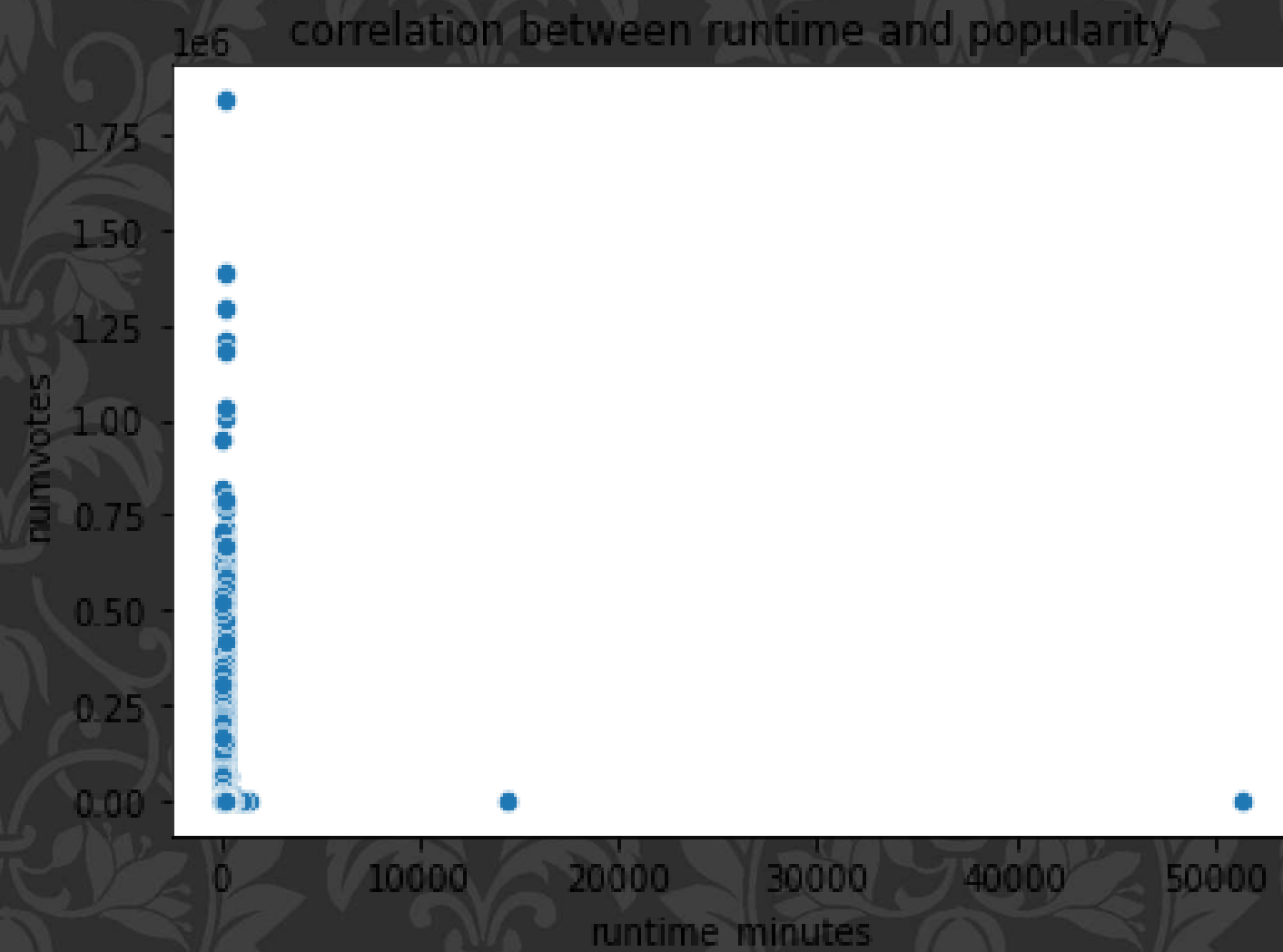The number of movies increases as the production budget increases.

# WHICH ARE THE MOST POPULAR GENRES



A combo of Action, Adventure and Sci-Fi are significantly more popular than other genres
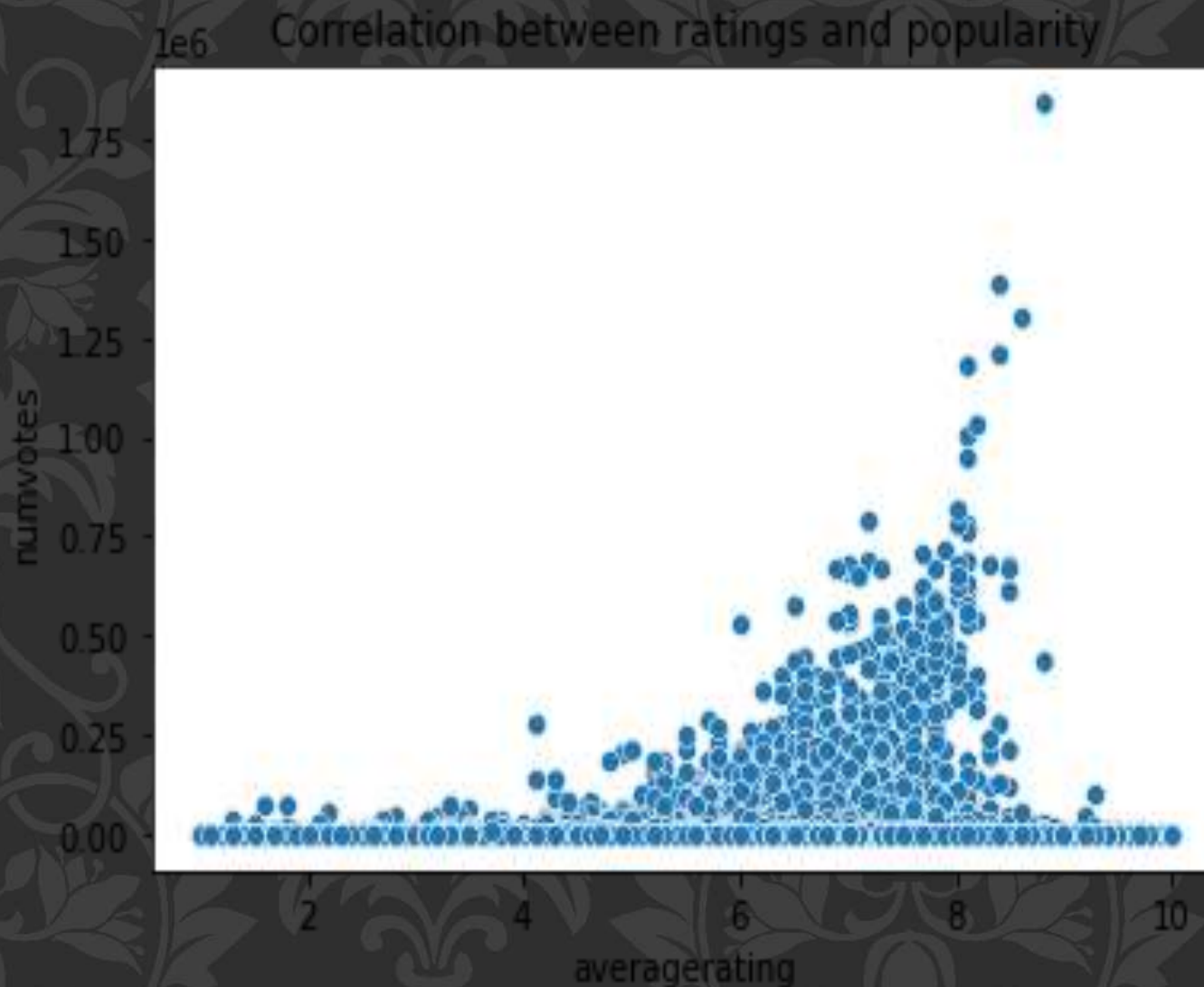
# What is the correlation between movie runtime and popularity


correlation between runtime and popularity

There is no correlation between runtime and popularity of movies

# What is the correlation between movie ratings and popularity?



There is a positive relationship between movie ratings and popularity.

1. Action as the most profitable genre, with Drama and Comedy following closely.

2. There was a  positive relationship between movie budgets and revenues, indicating that higher-budget films generally earned more. Interestingly, movie runtime showed minimal influence on revenue.

3. Microsoft should also consider scheduling movie releases for the months of May, June, and July. These months have consistently shown the highest profit margins. However, the underlying reasons behind the increased profits during these specific months require deeper investigation.

# THANK YOU

This presentation is prepared by Annbellah Nduta Mbungu.
If you have any enquiry, feel free to contact me at: annbellahn@gmail.com