

Felix-Awino / Phase\_1\_project Public

0 stars 0 forks

Star

Watch

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

main

...



Felix-Awino Update README.md ...

3 minutes ago 12

[View code](#)

README.md



Please fill out:

- Student name: Felix Awino
- Student pace: part time
- Scheduled project review date/time: N/A
- Instructor name: Everlyne Asiko
- Blog post URL:N/A

## PROBLEM STATEMENT

The project requires that we use exploratory data to review a set of movie data from a selected studio and use the findings as an insight into advising the head of Microsoft new movie studio on what films to create.

### The Project will attempt to answer the Questions below.

Having reviewed the dataset provided, I decided to work with the IM.DB SQL database which contained a good amount of data for consideration. The following key pointers could be retrived from the data through data cleaning and analysis as shall be demonstarted hereafter.

The following key issues will be relevant in advising the M/S movie production director appropriately.

- Which are the top ten genre of movies in production by volume.
- Which are the highest rated movies in production.

- Is there any correlation between the production volume and the movie ratings.

## DATA ANALYSIS PROCESS

---

In order to answer the subject question, we shall use Data Analysis with Pandas in order to work through the provided dataset. We shall go through the below key steps in our data analysis process.

1. Data exploration in Pandas
2. Data cleaning
3. Data analysis in Pandas

## Data

---

The data for this analysis is drawn from different movie studio databases including im.db sql database and tn.movie\_budgets.csv database. The data provides details of the movie names and ratings and goes ahead to provide the grossing amount made in the movie sales.

## Methods

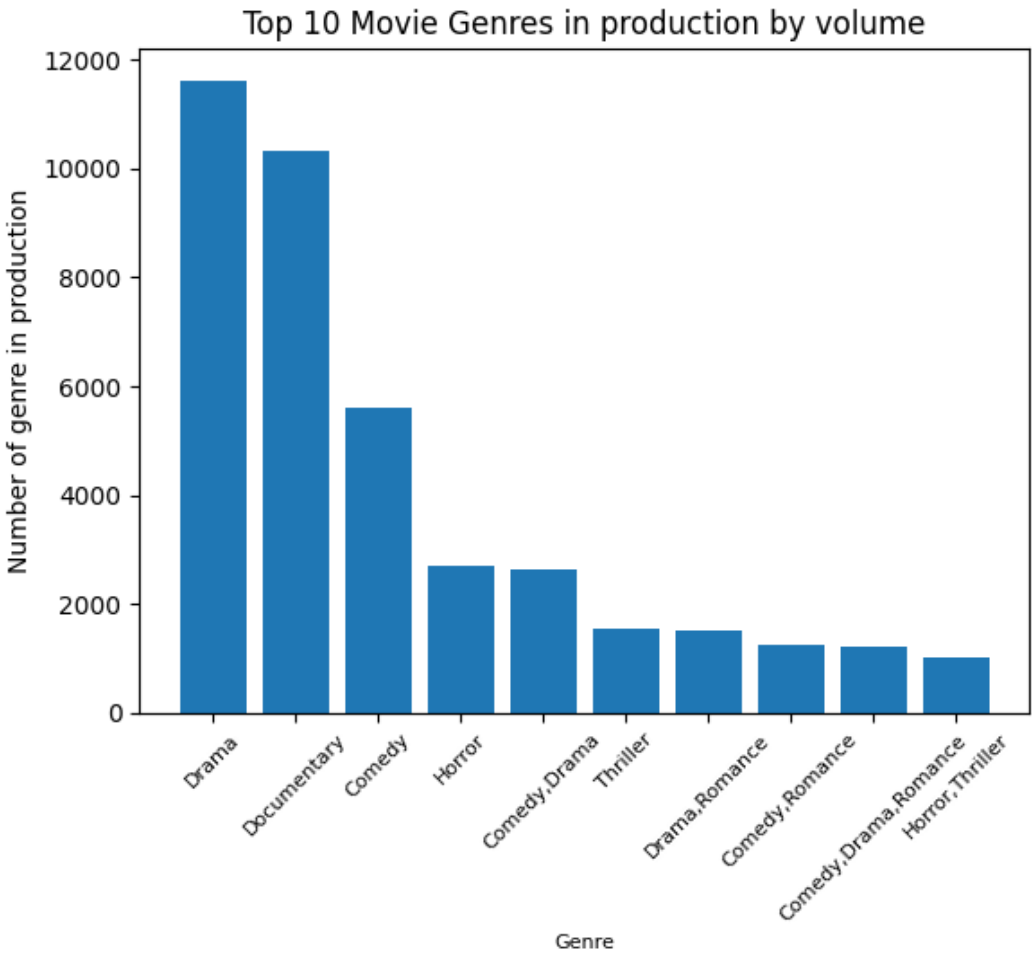
---

This project uses descriptive analysis, including description of the performance of the movies by the genre groupings. We were able to analyse the highest genres in active production, the highest rates genres and the highest grossing genres in terms of profits. This gave us a good overview of the kind of genres the production studio needs to focus on.

## Results

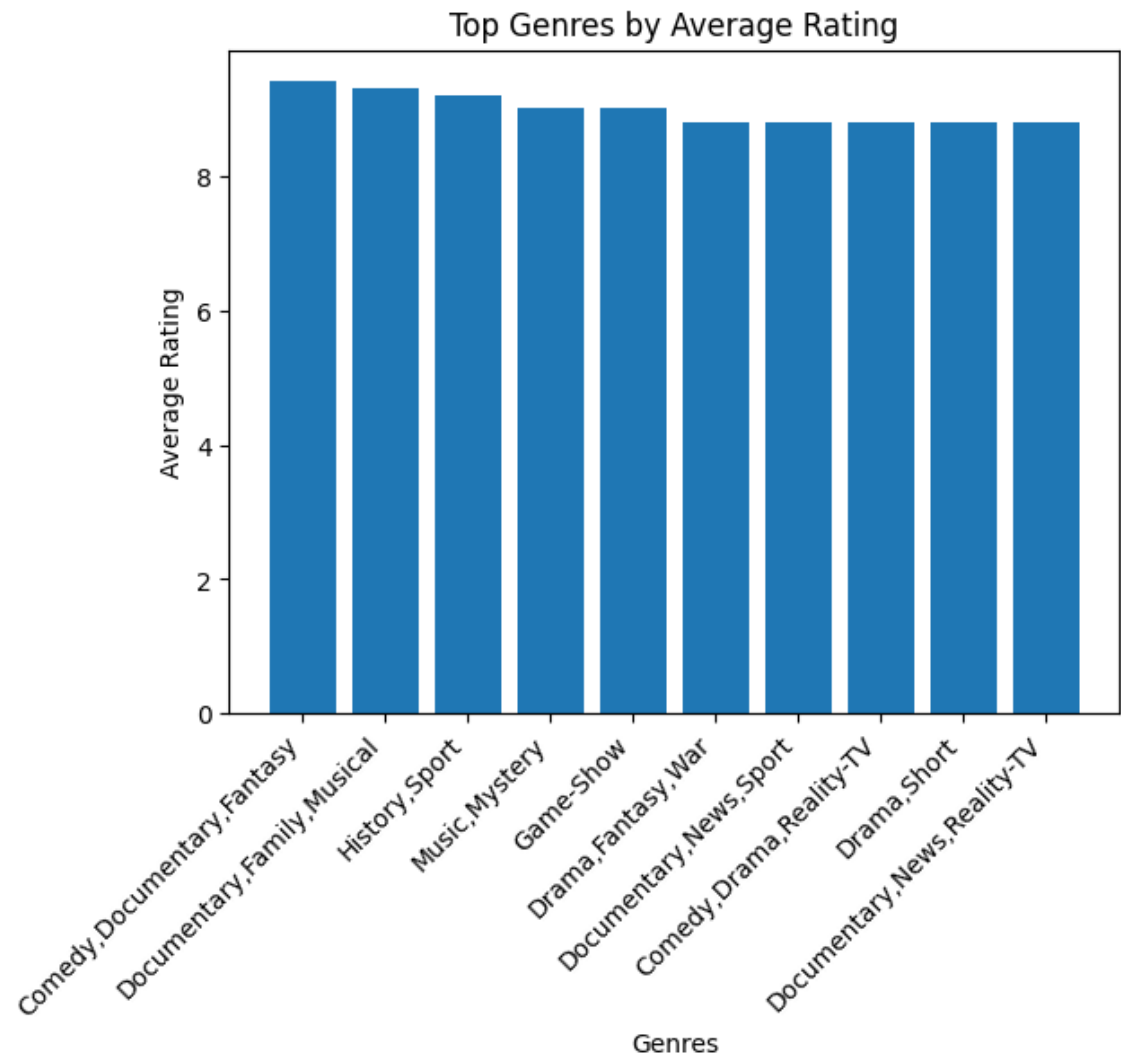
---

We were able to pick out the top ten movies in production by volume. This could point to the market share in terms of movie production and consumption numbers.



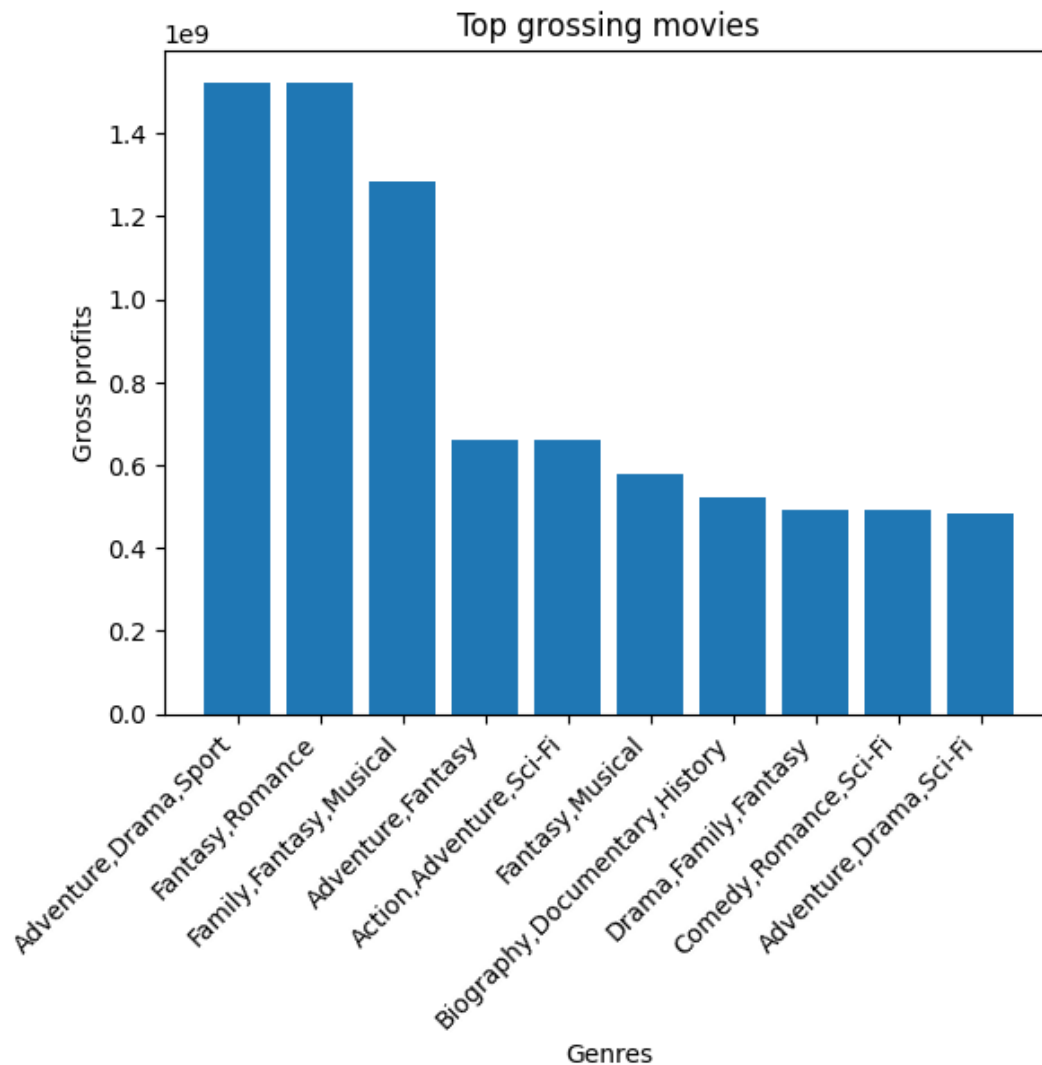
## Top Ten Rated Movies

We were also able to single out the top ten rated moviers by their genres.



### Top Ten grossing movies.

Next we analysed the average cost of production vs the gross sales per movie to determine which movies grossed the highest amounts to guide the area of production to focus on.



## SUMMARY AND CONCLUSIONS

From the analysis we were able to pick out the following facts around the movies in the given dataset.

- The three top genres in production are drama, documentary and comedy. In the event that the movie house intends to reach to the largest audience, these three genres would be a good place to get started.
- On the highest rated movies, the clear outcome from that analysis will be debunked in the next conclusion that the highest rating movies are not necessarily the ones that rack in the most profits or the highest viewership.
- We were also able to pick out the highest grossing movies. The top three being
  1. Adventure, drama, sport
  2. Fantasy, romance
  3. Family, fantasy, musical

These would be the top three recommendations to the production house were we to realise the highest profit margins.

## Next Steps

---

Further analyses could yield additional insights on movie consumptions by demographics, we could also look into the regional distribution of movie consumption by the market share.

We could also model a predictin for year by year trend on ovies consumed and predict the areas to focus in for future productions.

## For More Information

---

See the full analysis in the Jupyter Notebook or review this presentation.

For additional info, contact Felix Awino @[felix.awino@student.moringaschool.com](mailto:felix.awino@student.moringaschool.com)



### Releases

No releases published

[Create a new release](#)

### Packages

No packages published

[Publish your first package](#)

### Languages

● Jupyter Notebook 100.0%