Vivian-Maiyo /
dsc-phase-1-project

<> **Code**   Pull requests   ▷ Actions   ⊞ Projects   📖 Wiki   ⊘ Security   📈 Insights   ⚙ Set

⚖ View license

☆ **0** stars   ⑂ **424** forks   ⊙ **0** watching   ∿ Activity

🌐 Public repository · Forked from learn-co-curriculum/dsc-phase-1-project

⑂ **master** ▾

⑂ Branches   🏷 Tags

This branch is **1 commit ahead** of learn-co-curriculum:master.   ⑂ Contribute ▾   ↻ Sync fork ▾

Vivian-Maiyo Project Phase 1   ...   4 minutes ago   ⟲ 9

View code

≣ **README.md**   ✎

Movies Data Analysis 🔗

Author : Vivian Maiyo

Project Overview 🔗

This project analyzes movies data to come up with a recommendation on the best film genre
for Microsoft's new movie studio. It takes into account film genres, ratings, runtime, cost of
production and revenue to come up with the most ideal film genres for production. Our
Analysis dataset contains movie data between the years 2010 and 2019 from IMDb movies
and The Numbers.

Business Problem 🔗

Microsoft sees all the big companies creating original video content and they want to get in
on the fun. They have decided to create a new movie studio, but they don't know anything
about creating movies. You are charged with exploring what types of films are currently doing
the best at the box office. You must then translate those findings into actionable insights that
the head of Microsoft's new movie studio can use to help decide what type of films to create.

The Data 🔗

For this analysis, we use the data sets; * imdb.title.basics - contains data on production information on movies * imdb.title.ratings - contains data on movie ratings * tn.movie_budgets - contains the cost of production and sales revenue from the below movie sites; * IMDB * The Numbers

Methodology 🔗

The project uses descriptive statistics and data visualization tools to draw conclusions from the datasets.

Results 🔗

We established that the most popular film genre is Documentary. We further stablished that the cost of production was higher in genres rated good than those rated excellent and average. This means it is possible to produced highly rated films at a low cost. There is no strong correlation between rating and runtime, hence runtime does not determine the film rating. The revenue increases with the cost of production.

Conclusion 🔗

The best movies genre for Microsoft movie studio is Documentaries. There is no gain in increasing the cost of production as high cost do not necessarily yield high ratings nor profits. The movie studio should focus more on film genres with high rating.

Limitations 🔗

The data set tn.movie_budgets had very low values compared to the rest of the data sets, hence there was high loss of data.

Further Research 🔗

Explore the relationship between rating and return on investment. Explore the preferences of film genres by directors with high rating and profitability. Analyze the effect of release date on ratings and Return on Investment.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● **Jupyter Notebook** 100.0%   3/3