wambuiGicheha /
Project1

Code   Issues   Pull requests   Actions   Projects   Wiki   Security   Insights   Settings

0 stars   0 forks   1 watching   1 Branch   0 Tags   Activity

Public repository

1 Branch   0 Tags   Go to file   t   Go to file   Add file   Code   ...

wambuiGicheha final commits		87d0c0d · now
Visualizations	Read me additional content	18 hours ago
data	data commit	18 hours ago
.gitignore	adding a .git ignore file	18 hours ago
Presentation.pdf	pdf commit	11 minutes ago
Presentation.pptx	ppt content addition 2	14 minutes ago
README.md	final commits	now
index.ipynb	add bom movie dataset	15 hours ago

README

Project Overview

For this project, you will use exploratory data analysis to generate insights for a business stakeholder.

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.

Business Understanding

Initiative

Microsoft is seeking to venture into the film industry and wants to create a new movie studio and original video content.

Challenge

Lack of experience and knowledge in movie creation.

Goal:

Identify successful film types to guide Microsoft's new movie studio.

Personal Objective

1. To analyze current box office trends and generate actionable insights for Microsoft's new movie studio

Data Understanding and Analysis

Sources of data

In the folder zippedData are movie datasets from: Box Office Mojo - bom.movie_gross.csv IMDB - im.db Rotten Tomatoes TheMovieDB The Numbers

Description of the data

Two data sources were used for this data analysis namely the Box Office Mojo and IMDB database, due to time constraints for the project.

Data cleaning and preprocessing

It was established that both datasets had missing values

Movies basics

runtime_minutes column – 21% of that column had missing values(31,739/773,855 rows)

Original title column – 0.014% (21 /773,855)

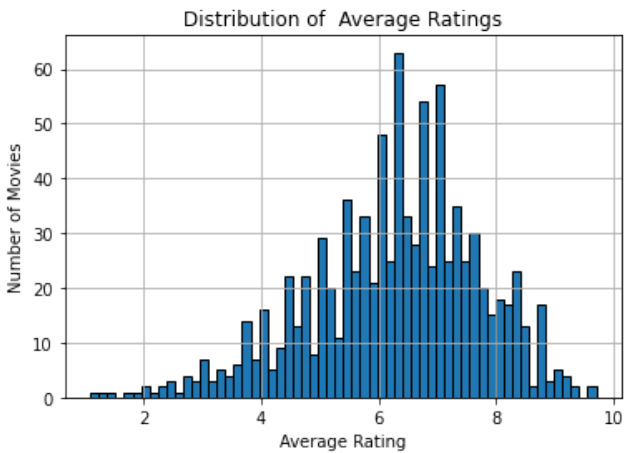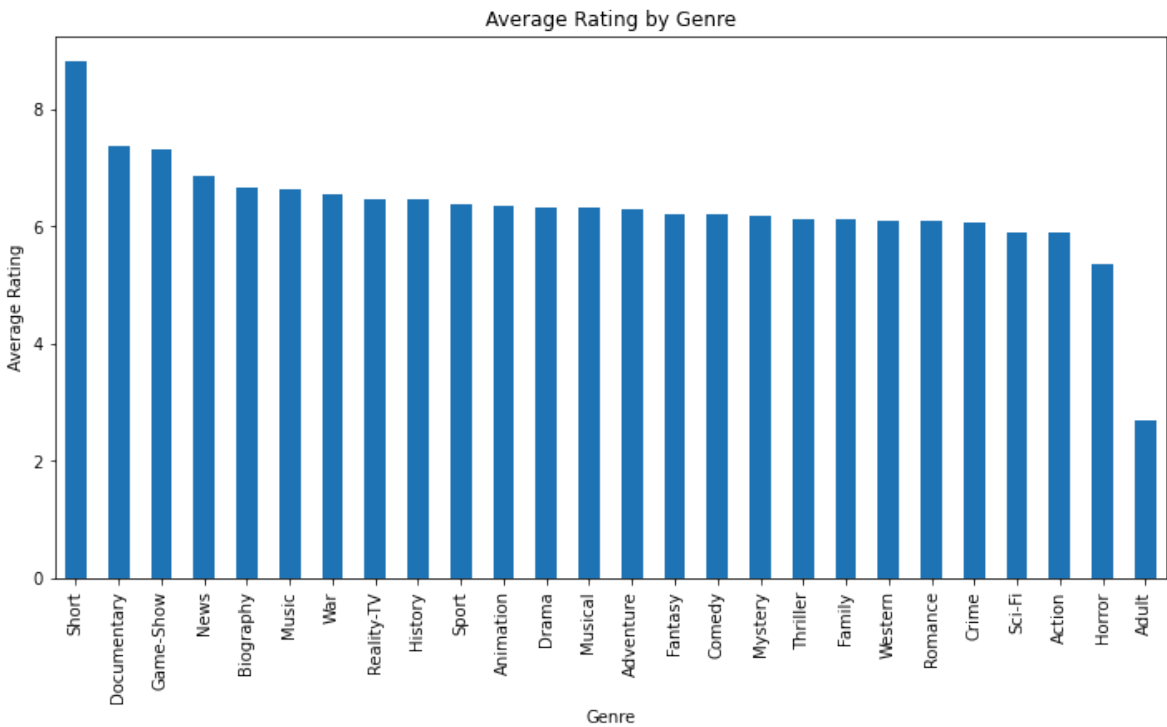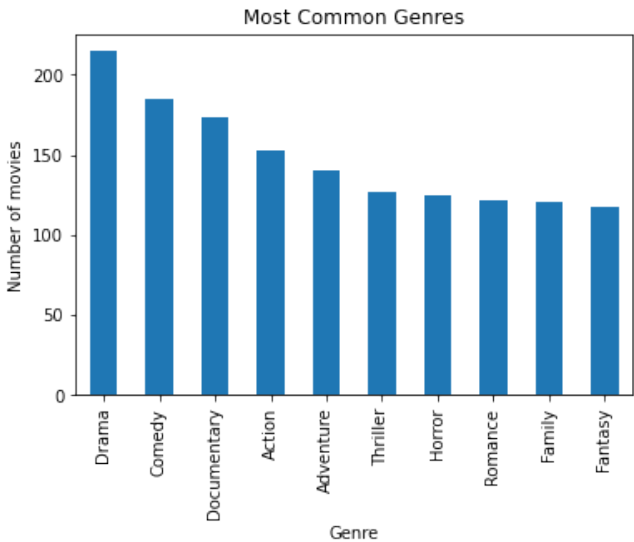Genres column – 3.7% (5408/773,855)
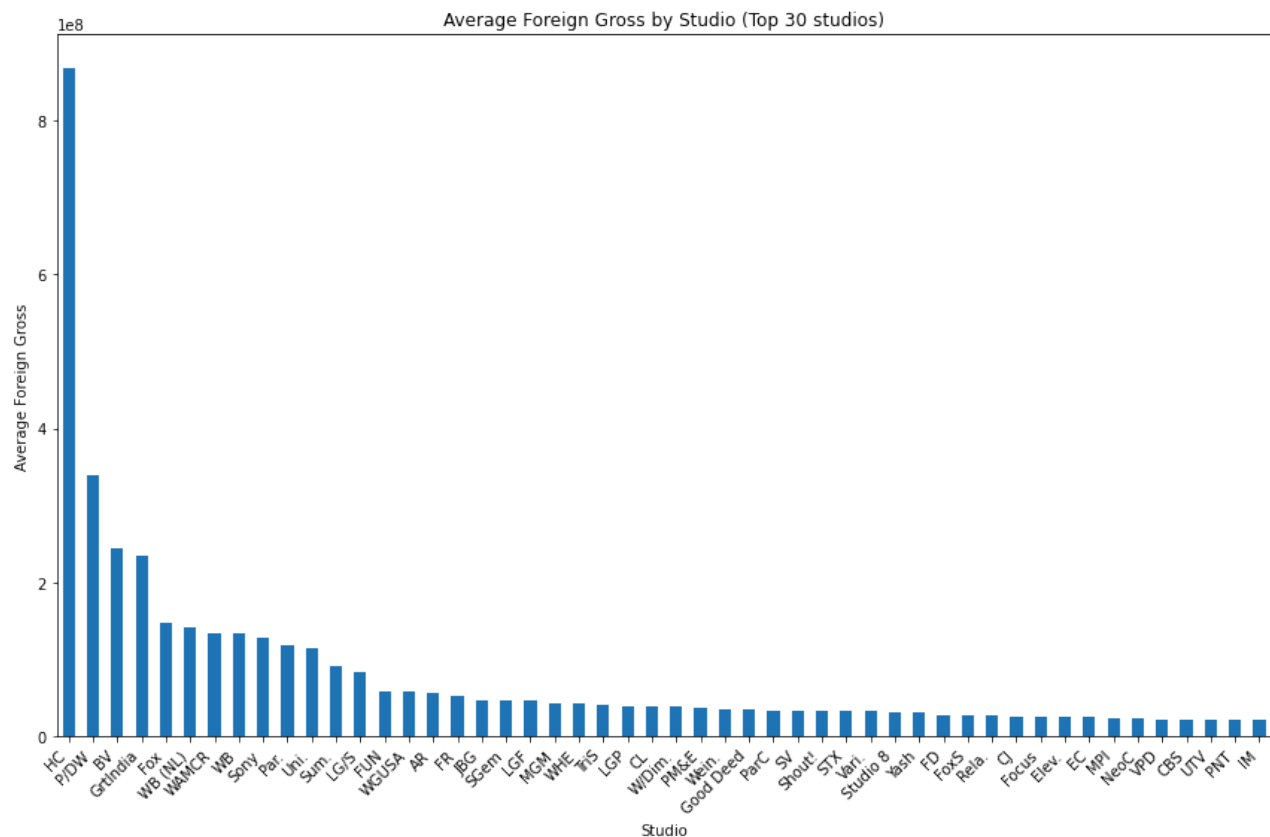
Box office movies

Studio 5/3386 (0.15%)

domestic_gross 28/3386 (0.8%)

foreign_gross 1350/3386 (40%)

Data Visualization

Most Common Genres

Average Rating by Genre

Distribution of Average Ratings

Average Foreign Gross by Studio (Top 30 studios)

Findings and Conclusion

Common movie genres are

Documentaries

Action

Drama

Adventure

Comedy

Their potential competitors will be

BV (Beirut visuals)

DW (DreamWorks)

P\DW (DreamWorks)

WB (Warner Bros)

Uni (Universal Studios)

Recommendations

1. Focus on popular movie genres with a high rating – Invest in genres that are consistently performing well at the box office(Documentaries, Action, Drama, Adventure, Comedy,)

2. Benchmark on studios that have a high domestic gross and foreign gross(BV, P\DW, WB, Uni,)- possibly collaborate in new movies with them.

3. Adapt to Trends: Monitor and adapt to changing genre trends to stay relevant.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● **Jupyter Notebook** 100.0%