

Exploring the most Popularity genres and Revenue Trends in Cinema: Insights from TMDB

An Analytical Study Using Python
Presented By Mohamed Amr

A large, dark blue, abstract shape that starts as a thin line on the left and expands into a wide, solid block on the right, positioned below the subtitle text.

Agenda

- Introduction
- Data Overview
- Data Analysis Process
- Conclusion

Introduction

Introduction: How did Movies–Industry Start?

- Photography and Motion Study & Invention of Motion Picture Cameras
- The Birth of Cinema by Making First Films and Public Screenings
- The Rise of Hollywood & Formation of Studios & Star System
- The Sound Era & Golden Age of Hollywood
- Global Expansion
- Modern Era

Data Overview

Data Source: tmdb_movies.csv

Time Period: January 1960-
December 2015

Key Variables: imdb_id, popularity,
runtime, genres, release_date,
budget_adj and revenue_adj.

Data Analysis Process

Data Analysis Process

1-Problem Defining: In this project, the primary objective is to explore and analyze trends within the movie industry using the TMDB Movies dataset. We aim to address the following key questions:

- What are the top 5 most popular genres across different time intervals?
- How do movies genres preferences change across different seasons and time intervals?
- What is the monthly distribution of adjusted revenue (`revenue_adj`) for movies?
- How has the average runtime of movies changed over the years?
- What is the relationship between movie runtime and its popularity in the last interval?

Data Analysis Process

2- Data Collecting: The data for this project is sourced from the TMDB Movies dataset (tmdb_movies.csv). This dataset includes various attributes of movies such as title, genre, release date, budget, revenue, popularity, and runtime. The dataset spans several decades, offering a comprehensive overview of the movie industry.

3- Data Cleaning:

- Removing Duplicates
- Handling Missing Values: (`budget_adj`, `revenue_adj` and `homepage`)
- Creating Derived Features: Extract the month from the `release_date` column to analyze trends over time.
- Creating calculating field: (`Season` and `interval`)

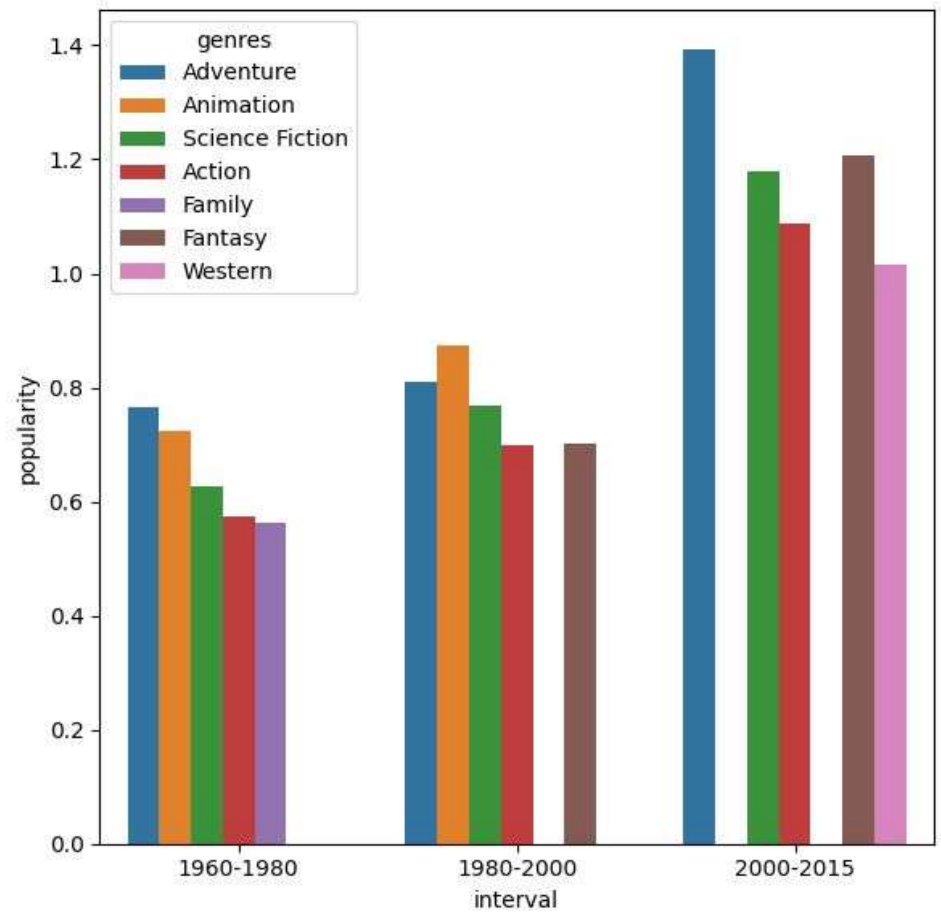
Data Analysis Process

4- Analyzing Data:

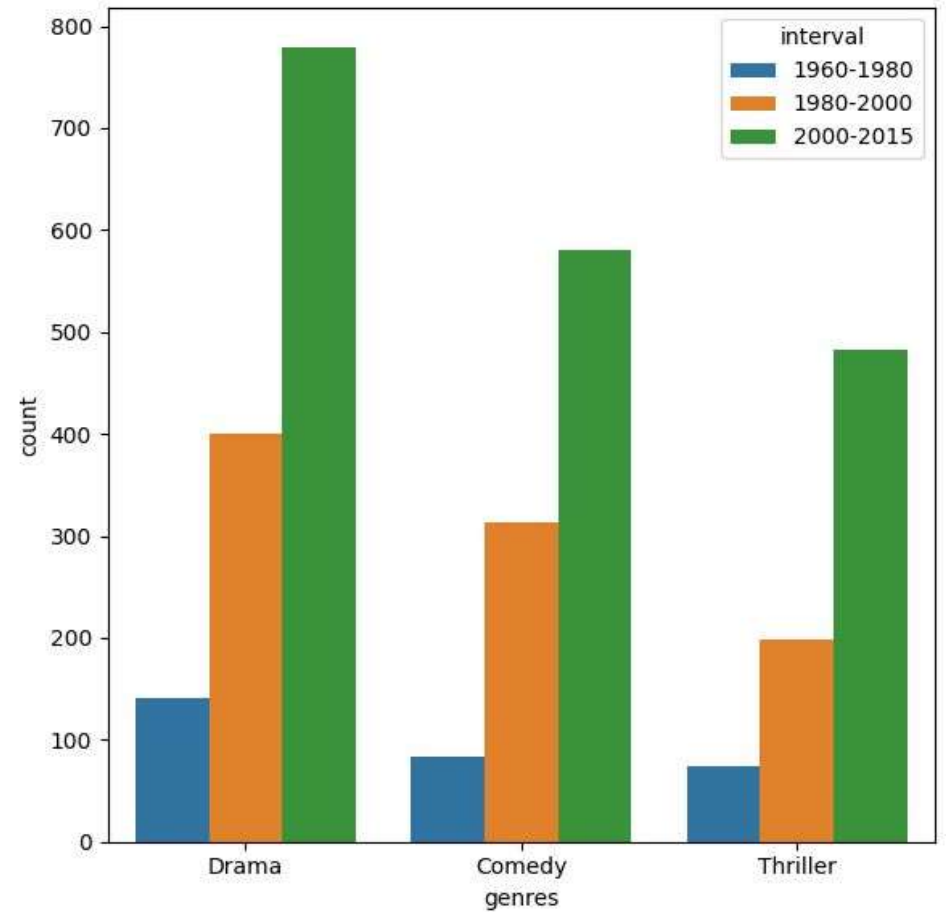
- Descriptive Statistics
- Grouping and Aggregating Data
- Trend Analysis
- Correlation Analysis

5-Data Visualization & Data Presentation:

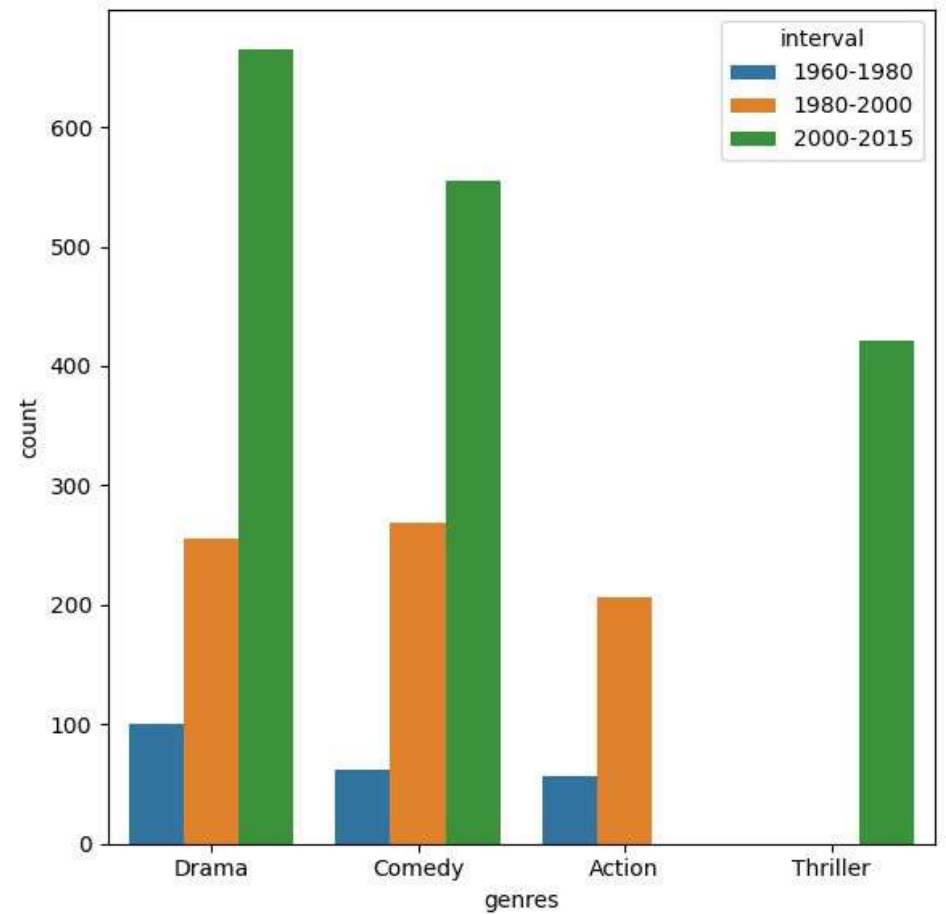
Top 5 most Genres by Interval



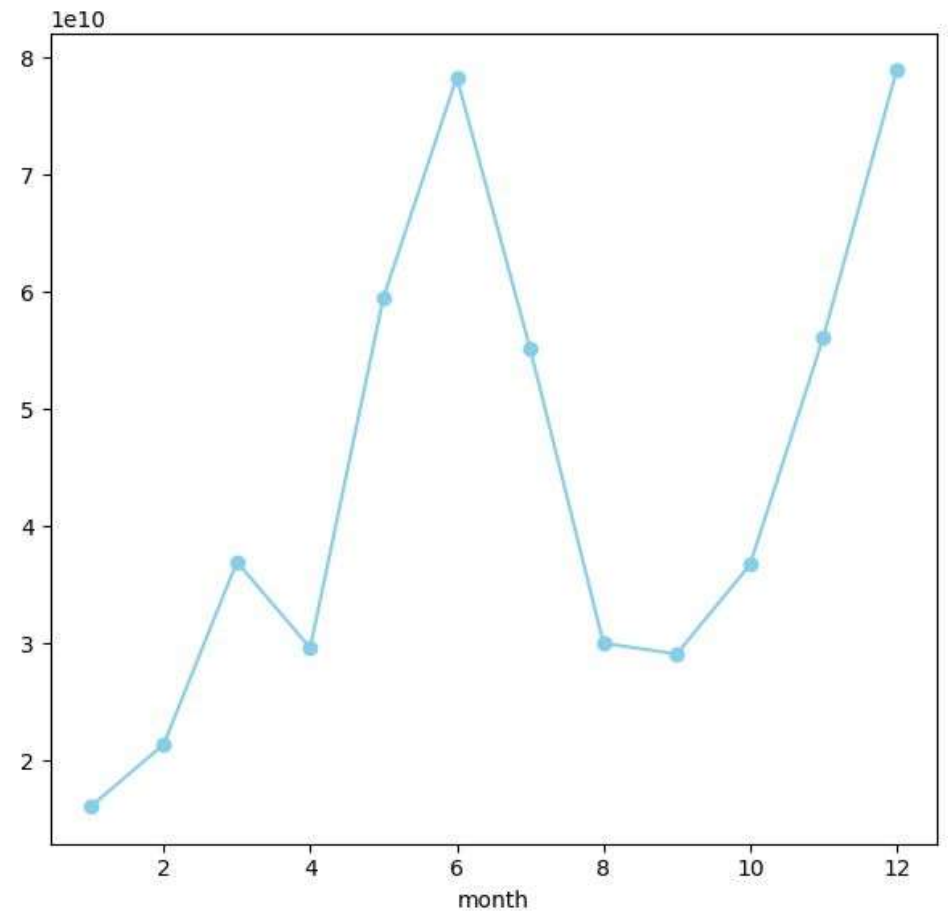
Top 3 Movie Genres by Fall Across Time Intervals



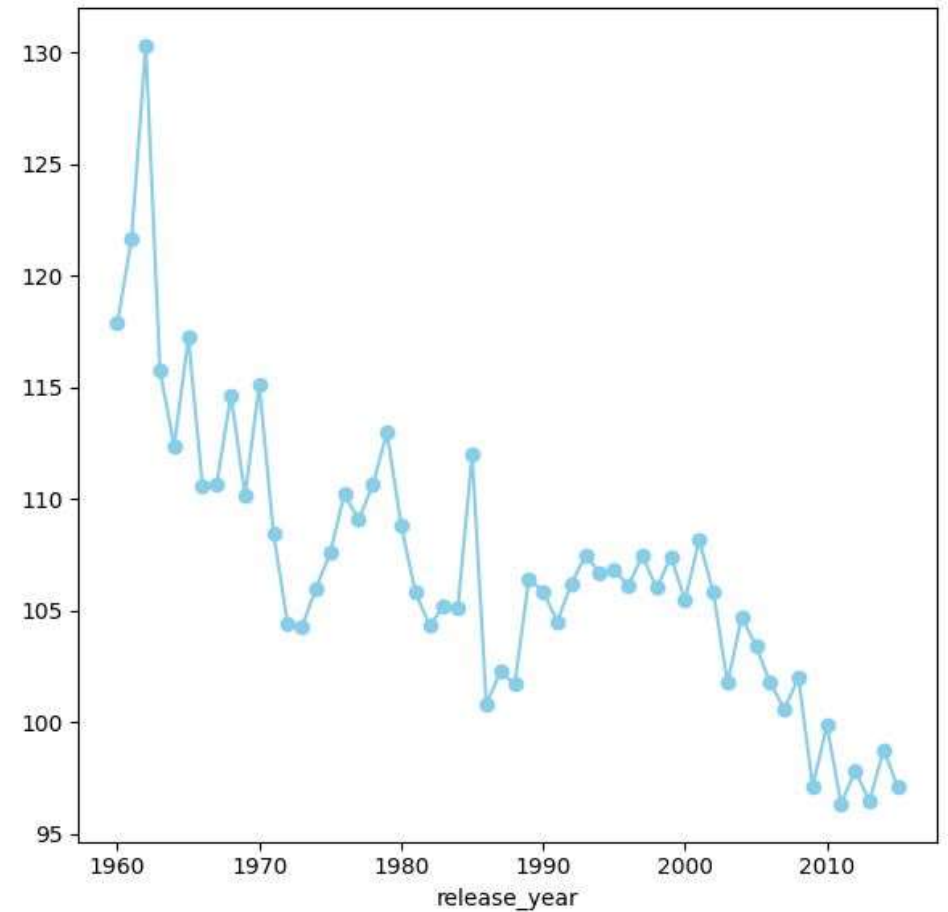
Top 3 Movie Genres by Spring Across Time Intervals



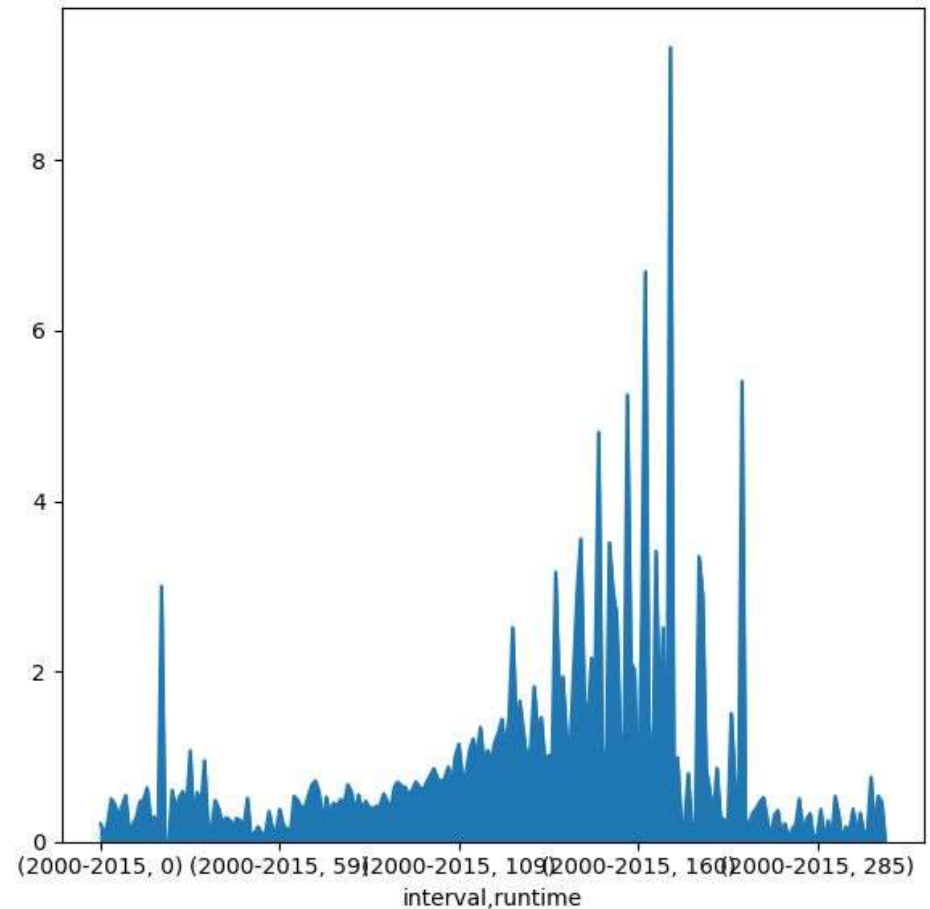
Revenue_adj for
each month



The Change in Average of Movies Runtime for each Year



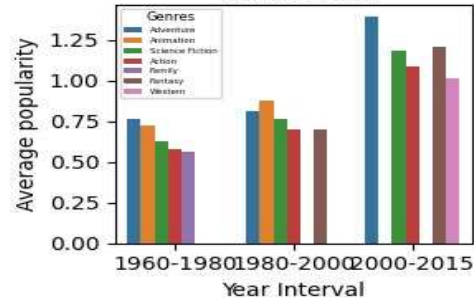
Exploring the Relationship Between Movie Runtime and ROI in Last interval



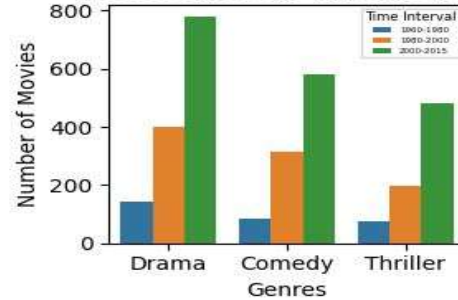
Conclusion

Dashboard

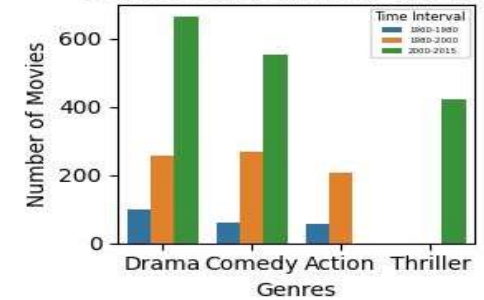
Top 5 Most Popular Genres by Interval



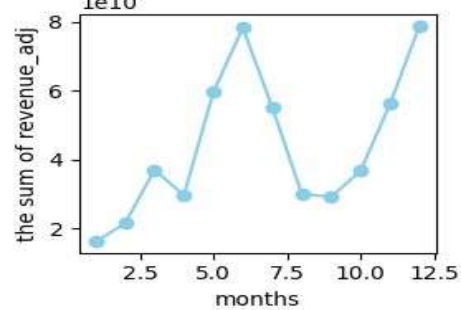
Top 3 Popular Genres by Fall Season Across Different Time Intervals



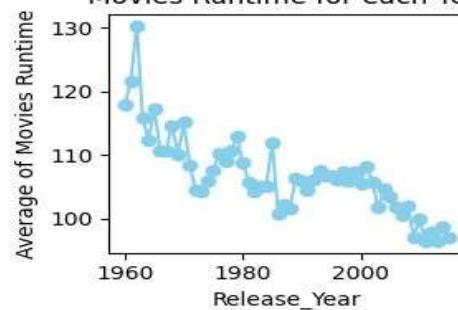
Top 3 Popular Genres by Spring Season Across Different Time Intervals



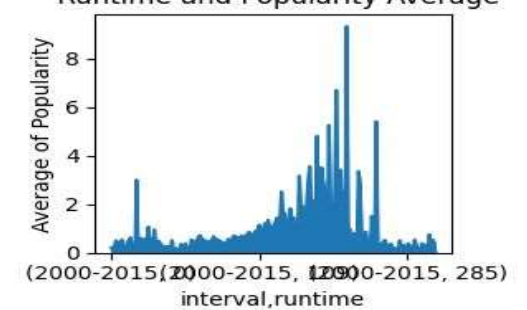
revenue_adj for each month



The Change in Average of Movies Runtime for each Year



The Relation between Movies Runtime and Popularity Average



Thank
You
Any Question?