 

**Data Science**

**“Analyzing Film Industry Rates”**

**Computer Science Department**

**University of Sulaimani**

**7th Semester**

**Prepared By: Azhin Nabaz, Omar Adil, Makwan Abdulla Supervised By: Dr. Miran Taha Abdullah**

**January 2024**

Table of Contents

[1. Introduction 3](#_Toc157495327)

[2. Methodology 4](#_Toc157495328)

[2.1 Data Collection: 4](#_Toc157495329)

[2.2 Exploratory Data Analysis (EDA): 4](#_Toc157495330)

[2.3 Statistical Analysis and Visualization: 5](#_Toc157495331)

[2.4 Impact of Movie Attributes: 5](#_Toc157495332)

[2.5 Implications and Recommendations: 6](#_Toc157495333)

[3. Implementation 7](#_Toc157495334)

[3.1 Loading and Displaying Data: 7](#_Toc157495335)

[3.2 Identifying Movies with High Reviews: 7](#_Toc157495336)

[3.3 Finding the Movie(s) with the Highest Rating: 8](#_Toc157495337)

[3.4 Determining the Movie(s) with the Highest Views: 8](#_Toc157495338)

[3.5 Analyzing Movies Released in a Specific Year: 9](#_Toc157495339)

[3.6 Visualizing the Average Review by Year: 9](#_Toc157495340)

[3.7 Plotting the Number of Movies Produced Each Year: 10](#_Toc157495341)

[3.8 Analyzing the Average Number of Views for Each Language: 10](#_Toc157495342)

[3.9 Plotting the Average Rating of Movies by Type: 11](#_Toc157495343)

[3.10 Visualizing the Total Views by Movie Type: 11](#_Toc157495344)

[3.11 Analyzing the Frequency of Each Movie Type: 12](#_Toc157495345)

[3.12 Determining the Preferred Time for a Good Movie: 12](#_Toc157495346)

[3.13 Analyzing the Correlation between Views and Time: 13](#_Toc157495347)

[4. Conclusion 14](#_Toc157495348)

[5. REFERENCE 15](#_Toc157495349)

# Introduction

In this report, we explore into the analysis of film industry rates using a dataset comprising information about various movies. Our exploration encompasses diverse facets such as movie ratings, views, types, languages, and production trends spanning different years.

By analyzing film industry rates, we aim to uncover the underlying patterns, trends, and dynamics that govern the reception and performance of movies. This project serves as a critical exploration into the mechanisms driving audience engagement, critical acclaim, and commercial success within the film industry. Through our analysis, we seek to provide actionable insights and recommendations for filmmakers, producers, distributors, and other industry players to enhance their decision-making processes and optimize their strategies for success.

# Methodology

Our analysis begins with the collection of a comprehensive dataset containing essential columns such as Movie Name, Type (categorized into genres like Drama, Crime, etc.), Views, Review ratings, Year of release, Duration, Accolades received (Prize Won), and Language. This dataset serves as the foundation for our exploration into the film industry's rate.

## 2.1 Data Collection:

-We start by gathering data on a diverse range of movies, ensuring our dataset represents various genres, languages, and release years.

- Each column in the dataset provides valuable information about the movie's characteristics, audience reception, and industry recognition.

## 2.2 Exploratory Data Analysis (EDA):

**-** We conduct exploratory data analysis to gain insights into the distribution and characteristics of the data.

- EDA helps us identify patterns, trends, and outliers in the dataset, guiding our subsequent analysis.

## 2.3 Statistical Analysis and Visualization:

- Using statistical techniques, we analyze the relationship between different variables in the dataset.

- We visualize the data using charts and graphs to present our findings in an easy-to-understand manner.

## 2.4 Impact of Movie Attributes:

- We explore the impact of movie attributes such as genre, Time, and language on audience engagement and critical reception.

- Understanding how these attributes influence a film's success provides valuable insights for filmmakers and producers.

## Implications and Recommendations:

- Based on our analysis, we draw conclusions about the factors driving success in the film industry.

- We provide recommendations for filmmakers, producers, and industry stakeholders to enhance their chances of success in the competitive landscape of cinema.

By following this methodology, we aim to provide a comprehensive analysis of the film industry's rate, offering valuable insights for industry professionals and enthusiasts alike.

# 

# 3. Implementation

## 3.1 Loading and Displaying Data:

We commence by scrutinizing the initial five rows of the dataset to grasp its structure and contents.

**Step 1:** Begin

**Step 2:** Load dataset from the specified file path

**Step 3:** Display the first five rows of the dataset

**Step 4:** End

## 3.2 Identifying Movies with High Reviews:

Identifying movies that have garnered exceptionally high reviews, characterized by ratings above 9.

**Step 1:** Begin

**Step 2:** Filter movies with review ratings greater than 9

**Step 3:** Display the filtered movies

**Step 4:** End

## 3.3 Finding the Movie(s) with the Highest Rating:

Pinpointing the movie(s) attaining the highest rating

**Step 1:** Begin

**Step 2:** Find the maximum review rating in the dataset

**Step 3:** Identify the movie(s) with the highest review rating

**Step 4:** Display the movie(s) with the highest rating

**Step 5:** End

## 3.4 Determining the Movie(s) with the Highest Views:

Identifying the movie(s) boasting the highest number of views.

**Step 1:** Begin

**Step 2:** Find the maximum number of views in the dataset

**Step 3:** Identify the movie(s) with the highest number ofviews

**Step 4**: Display the movie(s) with the highest views

**Step 5:** End

## 3.5 Analyzing Movies Released in a Specific Year:

Delving into movies released within a specific year, exemplified by the year

**Step 1:** Begin

**Step 2:** Filter movies released in the specified year (e.g., 2017)

**Step 3:** Display the filtered movies

**Step 4:** End

## 3.6 Visualizing the Average Review by Year:

Analyzing the average review score for movies annually and visually representing it through a bar plot.

**Step 1:** Begin

**Step 2:** Group movies by year and calculate the average review rating for each year

**Step 3**: Plot a bar chart showing the average review by year

**Step 4:** End

## 3.7 Plotting the Number of Movies Produced Each Year:

Examining the count of movies produced annually and illustrating the trend across different years.

**Step 1:** Begin

**Step 2:** Count the number of movies produced each year

**Step 3:** Plot a bar chart showing the number of movies produced each year

**Step 4:** End

## 3.8 Analyzing the Average Number of Views for Each Language:

Analyzing the average views garnered by movies in distinct languages and visualizing the outcomes.

**Step 1:** Begin

**Step 2:** Group movies by language and calculate the average number of views for each language

**Step 3:** Plot a bar chart showing the average number of views for each language

**Step 4:** End

## 3.9 Plotting the Average Rating of Movies by Type:

Determining the average rating across different movie types and illustrating it using a bar plot.

**Step 1:** Begin

**Step 2:** Group movies by type and calculate the average rating for each type

**Step 3:** Plot a bar chart showing the average rating of movies by type

**Step 4:** End

## 3.10 Visualizing the Total Views by Movie Type:

Analyzing the aggregate views for diverse movie types and visualizing the outcomes.

**Step 1:** Begin

**Step 2:** Group movies by type and calculate the total number of views for each type

**Step 3:** Plot a bar chart showing the total views by movie type

**Step 4:** End

## 3.11 Analyzing the Frequency of Each Movie Type:

Assessing the frequency of each movie type in the dataset and illustrating it through a bar plot.

**Step 1:** Begin

**Step 2:** Count the frequency of each movie type

**Step 3**: Plot a bar chart showing the frequency of each movie type

**Step 4:** End

## 3.12 Determining the Preferred Time for a Good Movie:

Determining the average duration of movies receiving high reviews (ratings above 5) and visualizing it.

**Step 1:** Begin

**Step 2:** Filter movies with review ratings above 5

**Step 3:** Calculate the average duration for these movies

**Step 4:** Display the average time for movies with reviews above 5

**Step 5:** End

## 

## 3.13 Analyzing the Correlation between Views and Time:

**Step 1:** Begin

**Step 2:** Calculate the correlation between views and duration for movies

**Step 3**: Plot a heatmap showing the correlation between views and time

**Step 4:** End

# 4. Conclusion

Through this comprehensive analysis, we unearth insights into various dimensions of the film industry, encompassing audience preferences, production dynamics, and the influence of diverse factors on movie ratings and views. This valuable information can empower filmmakers, producers, and industry stakeholders in making informed decisions and devising strategic initiatives for future endeavors.

# 5. REFERENCE

* <https://www.kurdcinama.com/>
* <https://www.imdb.com/?ref_=nv_home>
* <https://link.springer.com/chapter/10.1007/978-981-15-0135-7_31>
* https://dl.acm.org/doi/abs/10.1145/1460676.1460684