

# *FILM INDUSTRY ANALYSIS*

*PRESENTATION  
BY GROUP 2.  
DSFT-H12*



# TABLE OF CONTENTS

- Overview
- Business Understanding
- Data Understanding
- Data Analysis
- Visualizations
- Recommendations
- Next Steps
- Thank You



0 0 0 0

# **BUSINESS UNDERSTANDING**



The movie industry is a dynamic and highly competitive space, where studios must carefully choose which films to produce to maximize profitability. This new movie studio aims to establish itself by leveraging data-driven insights to make informed decisions about film production.

## **PROJECT OVERVIEW**



This project aims to guide a new movie studio in making data-driven decisions on film production to maximize profitability. By analyzing historical box office trends, the project seeks to uncover patterns in successful films, helping the studio determine the most promising genres, themes, or production strategies.

# DATA UNDERSTANDING

To understand the Datasets, we do **Data Loading** which is importing the relevant CSV files and the Database then ensure they are structured correctly using Pandas Libraries and sqlite3.

Datasets Imported:

- Box Office Mojo Dataset (bom.movie\_gross.csv.gz): Contained information on movie titles, domestic and foreign gross revenue, studio, and year of release
- Budgets Dataset (tn.movie\_budgets.csv.gz): Contained movie titles, production budgets, worldwide gross revenue, and release dates.

# **DATA ANALYSIS**

Data analysis is the process of examining, cleaning, transforming, and interpreting data to extract valuable insights.

Steps taken in data analysis:

**1.Exploratory Data Analysis (EDA)** – Examine basic properties such as column types, summary statistics, and distributions.

Python functions used :

- df.head(): Displays the first rows of the Dataframe.
- df.info(): Provides an overview of the Dataframe, including column types and non-null values.
- df.describe(): Generates summary statistics for numerical columns.

# DATA ANALYSIS

**2. Data Cleaning:** Missing values and duplicates are handled, and outliers are removed using statistical methods like IQR.

Python functions used:

- df.duplicated().value\_counts() - checking duplicates
- df.isna().sum() - checking missing values
- df.drop\_duplicates(): Removes duplicate records.
- df.astype(type): Converts column data types.
- '.replace()' - Used to replace specific values

# DATA ANALYSIS

**3. Data Merging:** Datasets are merged to create a comprehensive final\_df. We applied the filling\_missing function to the merged dataframe (final\_df) to remove any remaining duplicates, drop sparse columns, and impute missing values in both numeric and categorical fields according to our defined strategy.

**4. Feature Engineering:** To enrich the dataset with actionable attributes, we created the following derived features: movie\_profit/loss, release\_month, and season.

# VISUALIZATIONS

- **Average Profit by Genre**

Bar chart showing mean profitability for each genre, highlighting top- and bottom-performing categories.

- **Average Profit by Season**

Bar chart comparing seasonal profitability (Spring, Summer, Fall, Winter) to identify optimal release windows.

- **Budget vs. Profit Relationship**

Scatter plot (with regression line) of production budget against movie profit to assess how investment levels influenced returns.

- **Top Directors and Actors**

Horizontal bar charts ranking the top 10 directors and top 10 actors by average profit to illustrate “star power” effects.

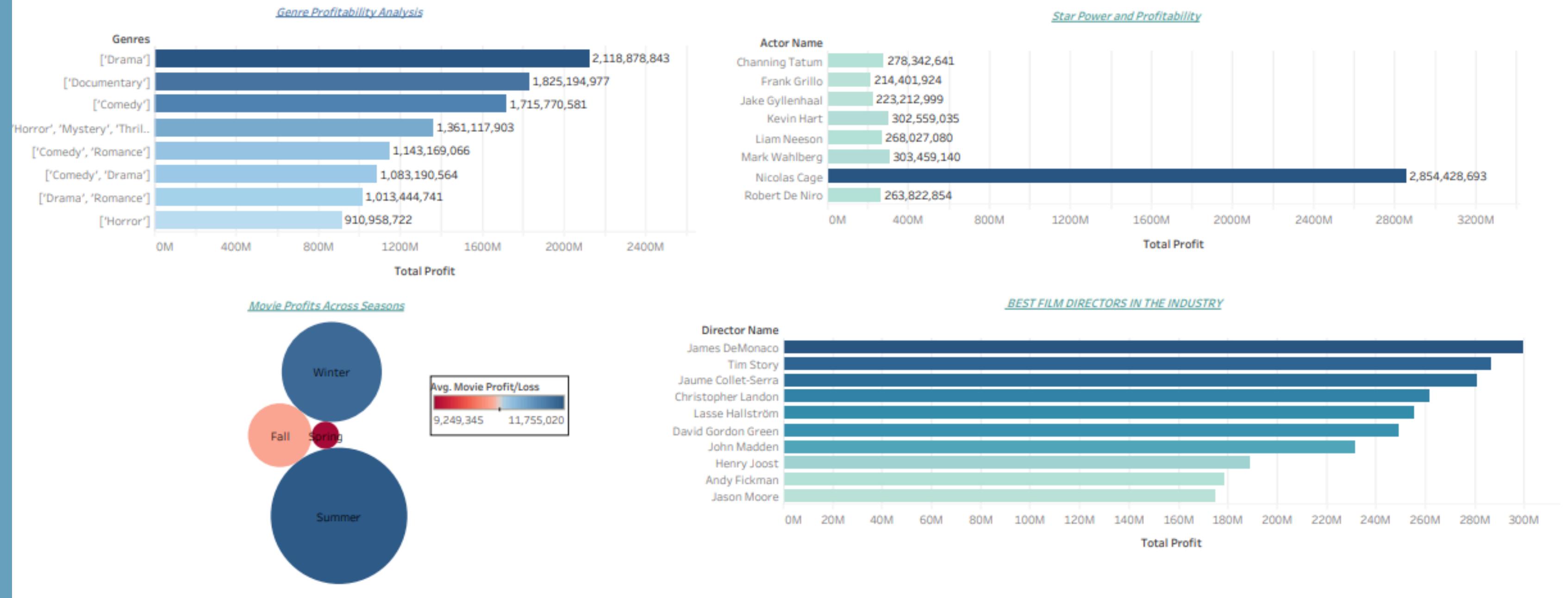
# **HYPOTHESIS TESTS**

We performed a hypothesis test to evaluate whether to accept or reject a null hypothesis based on the p-value and significance level (alpha) and the following were the findings:

- One-way ANOV - Genre has an effect on a movie's profit.
- One-way ANOVA - Season has no effect on movie profits.
- Chi-square statistic: Top actors and movie profit category are dependent on the movie profit.

# VISUALIZATIONS

## Factors Contributing to Film Success



# **INSIGHTS FROM THE ANALYSIS**

- **Genre Selection Drives Profitability:** Investing in genres with lower returns should be strategic, leveraging strong factors such as star power to boost success.
- **Seasonality Impacts Revenue:** Films released in Summer tend to perform significantly better financially.
- **Budget Optimization is Crucial:** The ideal budget range for maximizing return on investment is around \$80 million. While ultra-high budgets above \$150 million may generate large box office grosses, they also come with high financial risk due to unpredictable ROI.
- **Star Power Influences Success,** But Should Be Balanced Hiring top-performing actors and directors boosts profitability.

# **RECOMMENDATIONS**



## **Focus on High-Returning Genres**

- Prioritise the development of top-3 Genres ie. Action, Adventure & Comedy that have the highest average profit.
- Invest in genres with lower-than-average returns only if they are accompanied by other factors like top actors.

## **Optimize Release Timing**

- Schedule large releases over the summer, when average earnings were much greater.
- Summer releases perform better financially because of increased audience availability, school holidays, and high demand for blockbuster films.

# **RECOMMENDATIONS**

## **Budget Allocation Strategy**

- Set a production budget of around \$80 million to maximise return on investment.
- Be wary of ultra-high budgets (> \$150 million); while they can create enormous revenues, their ROI is more uncertain.

## **Leverage Star Power Selectively**

- Cast high-performing actors and directors who have consistently contributed to increased earnings.
- Balance iconic talent with rising talents to keep pay costs under check while still engaging audiences.

# **NEXT STEPS**

- **Invest in Profitable Genres** – Focus production on top-performing genres like Action, Adventure, and Comedy
- **Strategically Plan Release Dates** – Schedule major film releases in Summer, when profits are highest, to maximize audience turnout and revenue potential.
- **Optimize Budget Allocation** – Keep production budgets around \$80 million for best return on investment and avoid excessive spending beyond \$150 million unless backed by strong revenue projections.
- **Balance Star Power Wisely** – Cast high-performing actors and directors strategically, negotiating profit-sharing deals to reduce upfront

○ ○ ○ ○

**THANK  
YOU**