

PHASE 2: PROJECT

DATA SCIENCE

DSF-PT11

Film Industry Analysis for Strategic Studio Launch

Business Recommendations Based on Industry Data

PRESENTED BY

Group 6



DATE

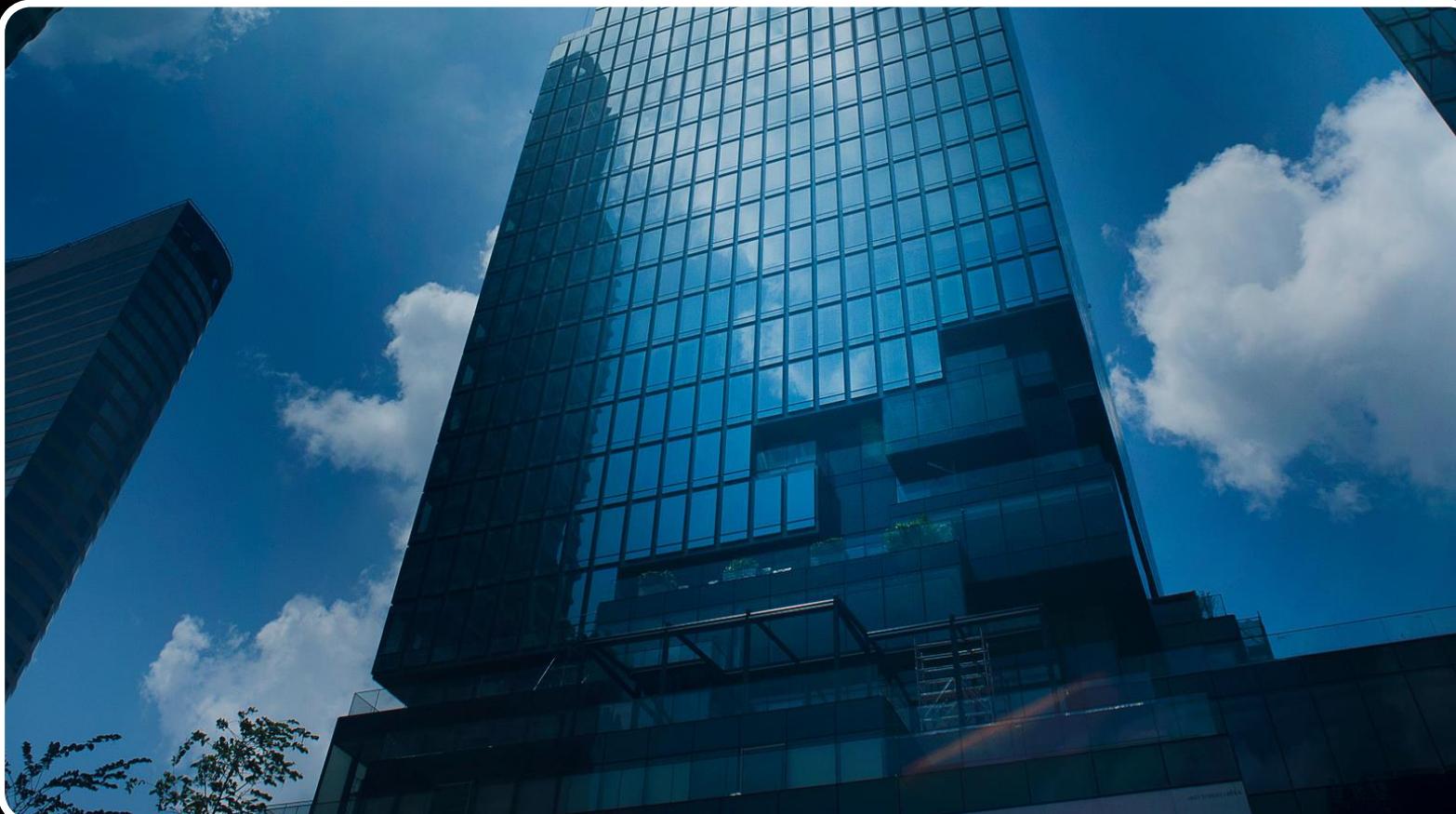
Jun. 12, 2025



PHASE 2: PROJECT

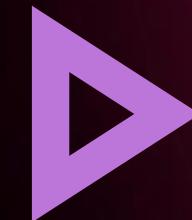
GROUP MEMBERS

INSTRUCTOR: GEORGE KAMUNDIA



- 01. EMMANUEL KIPLETING**
- 02. ERICK KIBUGI**
- 03. HILDA JEROTICH**
- 04. ALICE MUIA**
- 05. DAVID MURITHII**
- 06. BONIFACE NJERI**
- 07. BERNICE NYAWIRA**

SECTIONS



OVERVIEW AND BUSINESS UNDERSTANDING

This section introduces the goals of the project, the stakeholders' needs, and the context surrounding Golden Arc Pictures' entry into the movie industry.

DATA UNDERSTANDING

Here, we present the diverse sources of our data, including TheMovieDB, IMDb, Box Office Mojo, Rotten Tomatoes, and The Numbers.

DATA ANALYSIS

This section focuses on the insights derived from the data. We explore trends, relationships, and correlations using visualizations and statistical methods.

RECOMMENDATIONS AND NEXT STEPS

Based on our analysis, we offer actionable recommendations that address the core business questions.

OVERVIEW



Understanding What Makes A Movie Successful *

Golden Arc Pictures is preparing to enter the film production industry and requires a strategic foundation to guide its content development decisions. The objective of this project is to support that launch by identifying the characteristics that define a successful film. To achieve this, we conducted a data-driven analysis of historical movie data from multiple industry sources. Our focus was to uncover trends in genres, production budgets, runtimes, languages, and audience preferences that correlate with high profitability and popularity. These insights provide a fact-based framework for making informed investment decisions in the competitive movie market.

BUSINESS UNDERSTANDING

Business Context *



Golden Arc Pictures is new to film production and wants to minimize investment risk while maximizing box office success.

Business Needs *



- Identify genres with high profitability and popularity
- Determine budget strategies that optimize ROI
- Understand factors like language, genre, and marketing impact

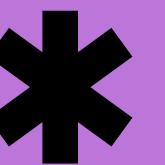
Key Questions *



- 1) How does budget relate to ROI?
- 2) What languages and genres perform best?
- 3) How do runtime and rating categories influence audience satisfaction?

PHASE 2: PROJECT

Data Sources



- ❑ TMDb & IMDb – Metadata on films, genres, popularity
- ❑ Box Office Mojo & The Numbers – Budgets, revenues
- ❑ Rotten Tomatoes – Audience and critic scores

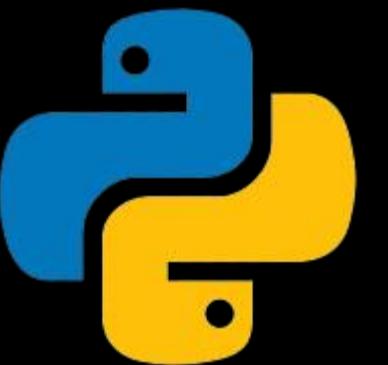
Preparation Steps

- ✓ Removed duplicates and missing values
- ✓ Standardized formats (e.g., dates)
- ✓ Created new variables like Revenue, Profit, and RO

Tools Used

- Python, Pandas, Numpy (data handling)
- Tableau, Seaborn, Matplotlib (visualization)
- SciPy, Statsmodels (hypothesis testing)

DATA UNDERSTANDING



+ a b | e a u

IMDb

matplotlib

**BOX
OFFICE
mojo**

**Rotten
Tomatoes®**

DATA ANALYSIS



Extracting Patterns to Drive Business Strategy *

- ✓ With our cleaned and integrated datasets, we conducted exploratory data analysis (EDA) to uncover meaningful trends and relationships. Our goal was to identify which film characteristics most strongly influence success – both in terms of audience reception and financial returns.
- ✓ To validate these visual insights, we used statistical testing:
 1. ANOVA tests confirmed significant differences in ratings across genres.
 2. T-tests helped compare ROI between different budget categories.
- ✓ Together, these insights form the basis of our final business recommendations – ensuring they're backed not only by patterns, but by statistical significance.

PHASE 2: PROJECT

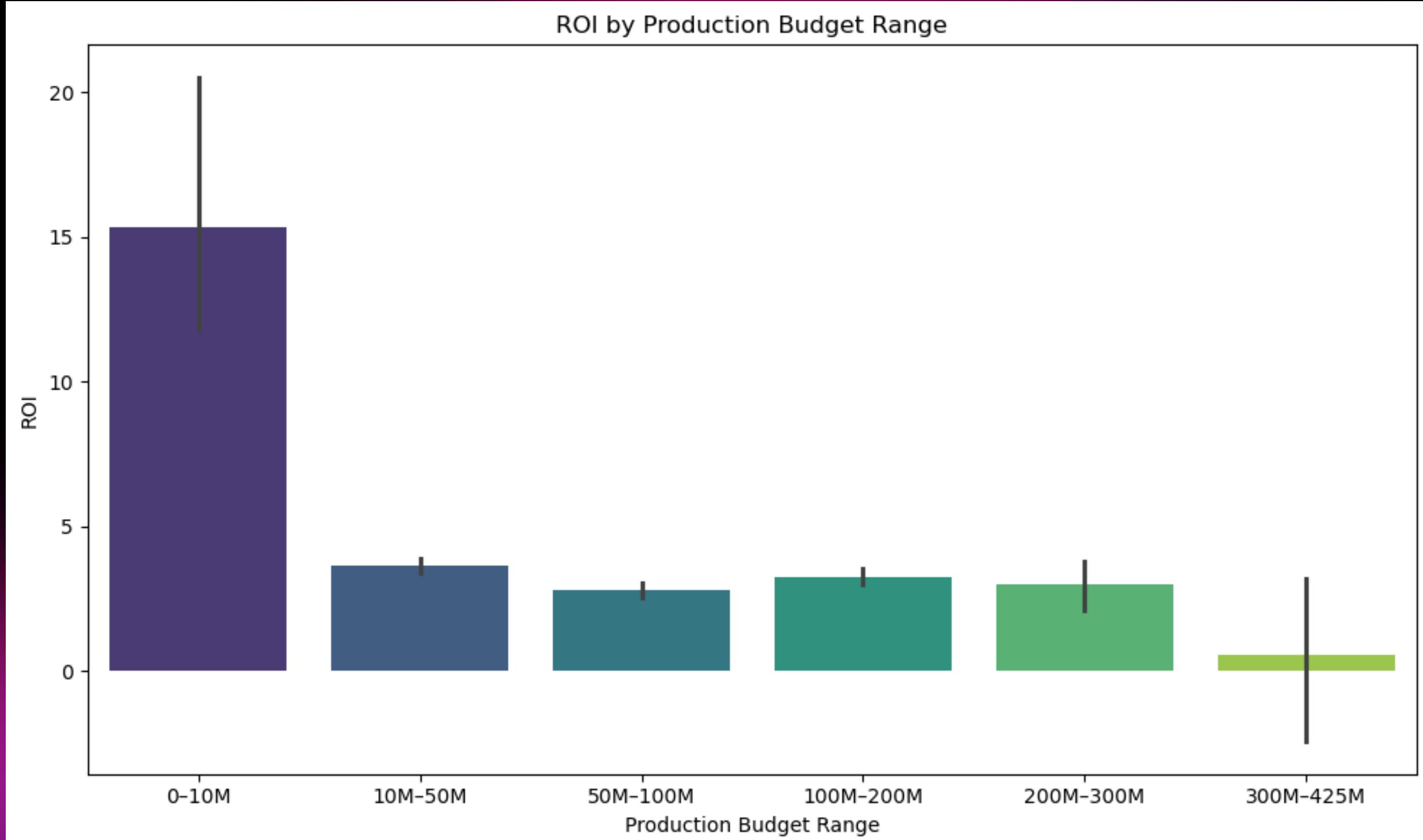
INSIGHT

Movies with low production budgets (< \$10M) have significantly higher Return on Investment (ROI) compared to higher-budget films.

INTERPRETATION

As production costs increase, ROI tends to decrease—possibly due to larger marketing and distribution expenses.

BUDGET VS ROI



PHASE 2: PROJECT

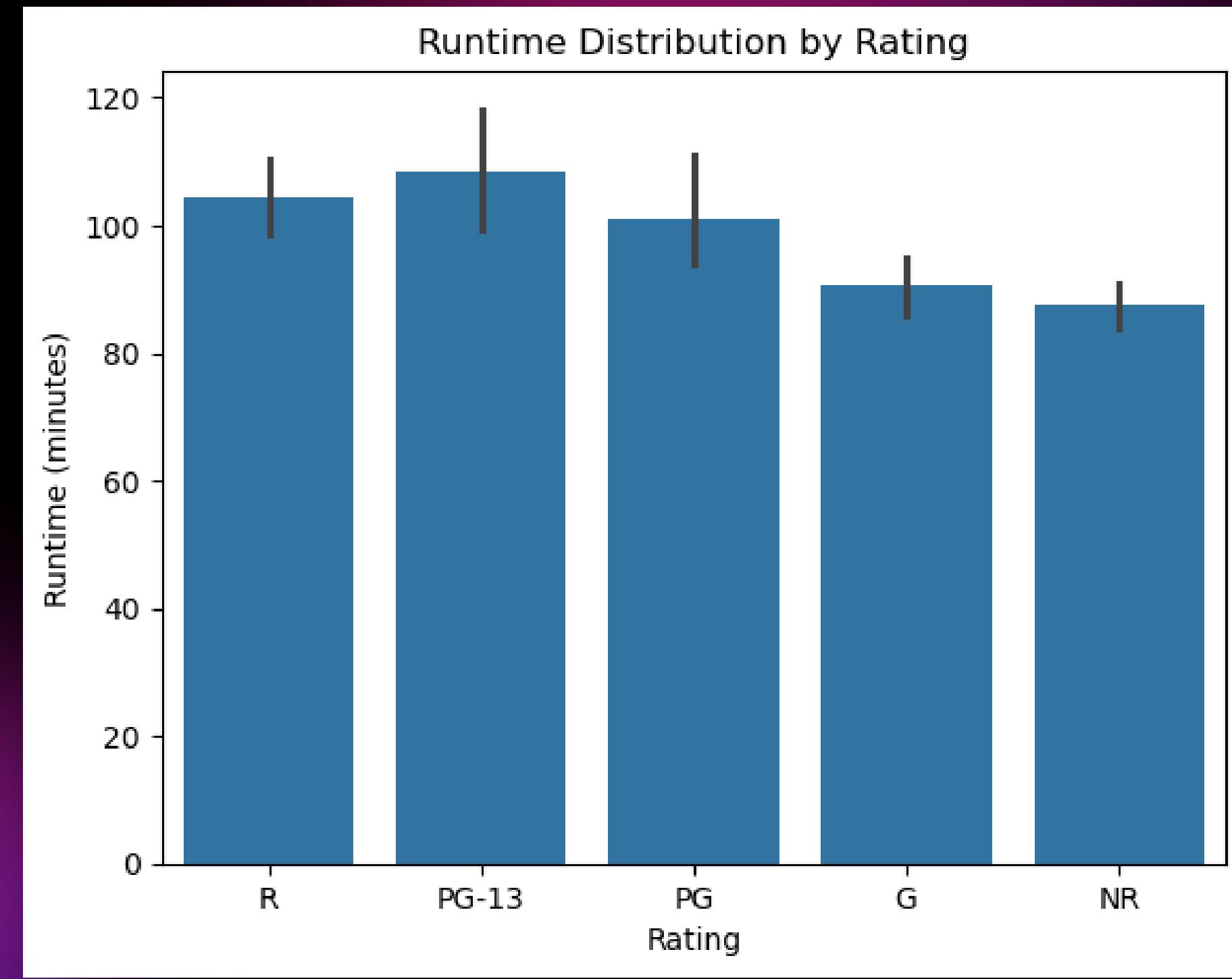
INSIGHT

PG-13 films with runtimes between 110–120 minutes earn the highest average audience ratings.

INTERPRETATION

Longer runtimes may allow for better storytelling in more mature-rated films, but shorter runtimes are better for younger audiences.

RUNTIME VS RATING



PHASE 2: PROJECT

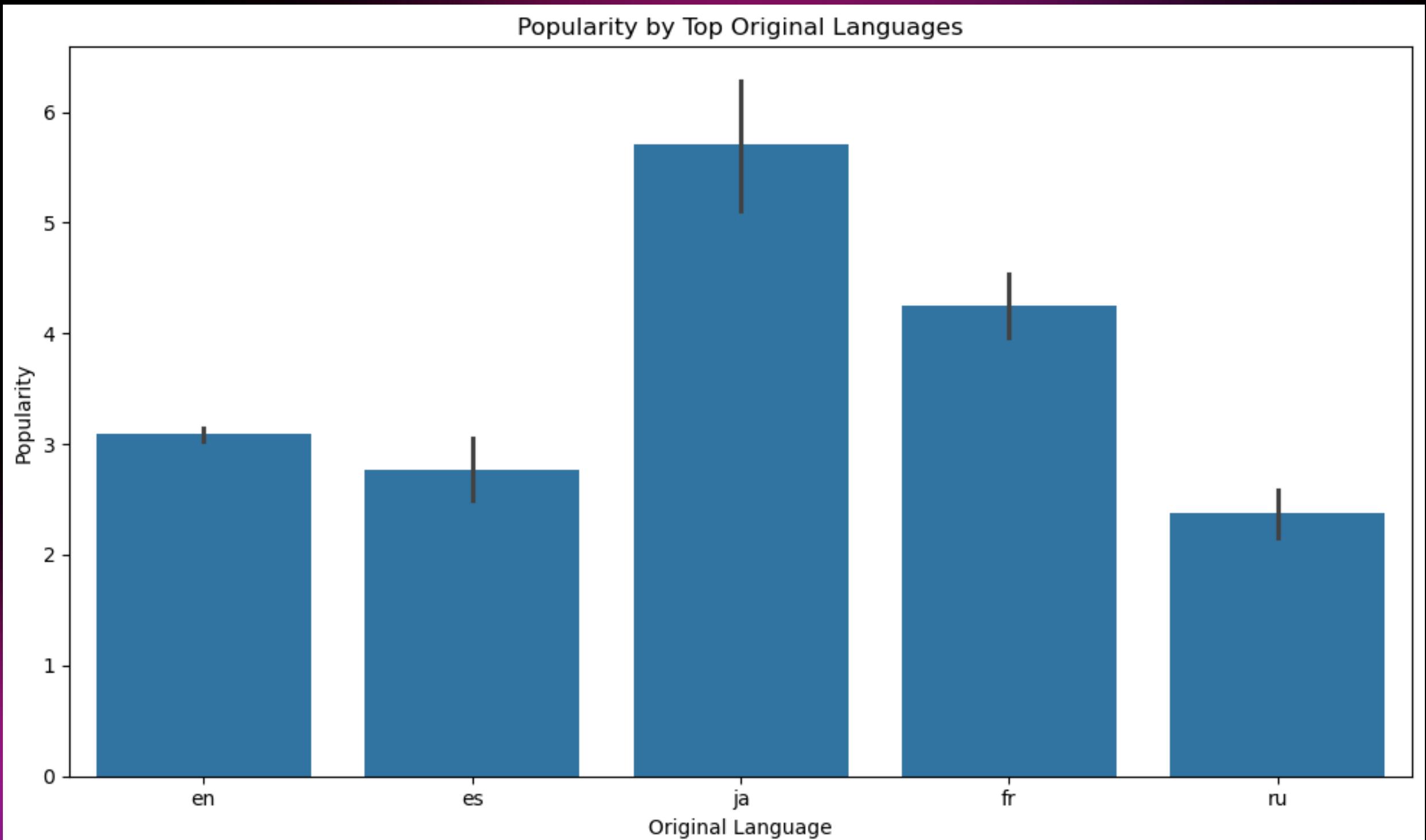
INSIGHT

Movies in Japanese, French, and English have the highest global popularity scores.

INTERPRETATION

Cross-cultural and multilingual content performs better globally, expanding audience reach.

LANGUAGE VS POPULARITY



RECOMMENDATIONS

1. Budget Strategy *



Golden Arc Pictures should invest in films with budgets under \$10M to achieve higher ROI.

2. Global Reach *



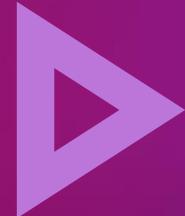
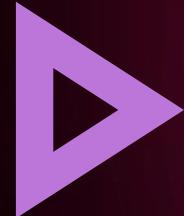
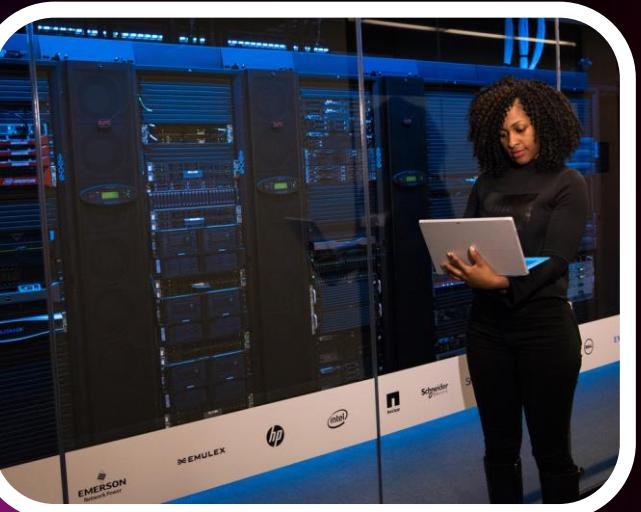
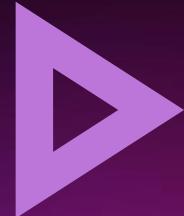
Studio should target a more global niche for the film to increase its popularity and reach. This will capture a large audience and should incorporate different cultures i.e. Japanese culture.

3. Genre to Focus *



Studio should focus on Comedy and Documentary genre as they are the most popular in the box office ratings are most likely successful in the market.

NEXT STEPS



LAUNCH A PILOT PROJECT

Begin with a low-budget film that aligns with the insights from our analysis – ideally in a high-performing genre like Comedy or Documentary.

PURSUE GLOBAL REACH

Explore partnerships for co-productions or invest in multilingual content strategies. Localizing films or incorporating cultural elements (e.g., Japanese or French language, international themes)

MONITOR PERFORMANCE METRICS

Implement a consistent KPI tracking system on key metrics such as audience ratings, revenue, and return on investment (ROI) to measure success and guide iterative improvements.

CONDUCT AUDIENCE TESTING

Before scaling production, run focus groups or early audience screenings to test preferences on runtime, genre, and storytelling style.

PHASE 2: PROJECT

THANK YOU

GITHUB REPOSITORY:

<https://github.com/hildajerotich/GROUP-6-PHASE-2-PROJECT>

