# Movie Studio

Industry Analysis

**Cayke Felipe dos Anjos**
**Dolgor Purbueva**

# Business Problem

The  company aspires to enter the competitive field of original video content production by establishing a new movie studio. To succeed in this endeavor, it's crucial to understand the current market trends. This project involves conducting exploratory data analysis to provide actionable insights for our business stakeholders.

Our  goal is to explore and analyze film industry data to identify:

- How should the movie be budgeted?
- What genres are most profitable?
- What are the best directors for each genre?
- When should a movie be released?
- What should be the approximate runtime?

# Data used

Used data with movies financial results such revenue and profit, genres, movie ratings, popularity, release dates, directors.

## TheNumbers

6 columns, 5782 rows

## TheMovieDb

10 columns, 26517 rows

## IMDB

8 tables

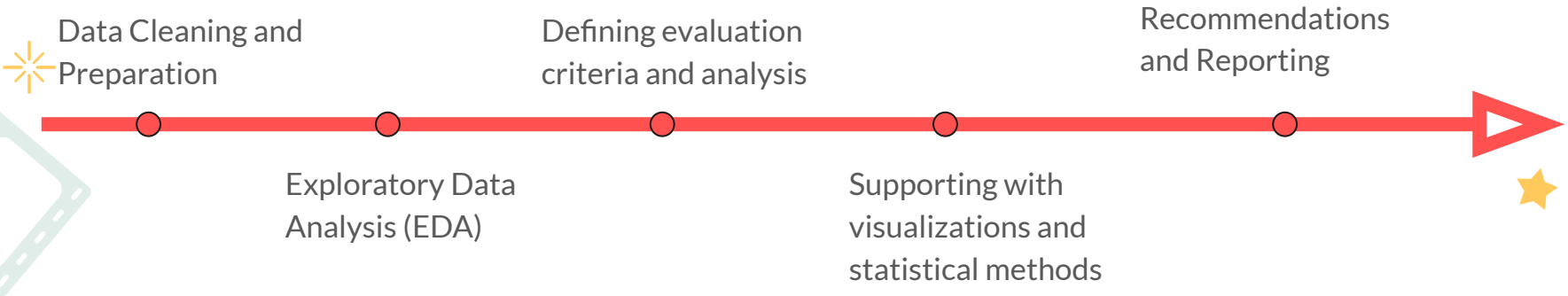
## Bom Office Mojo

5 columns, 3387 rows

## Rotten Tomatoes

12 columns, 1560 rows
8 columns, 54432 rows

# Data Analysis Steps

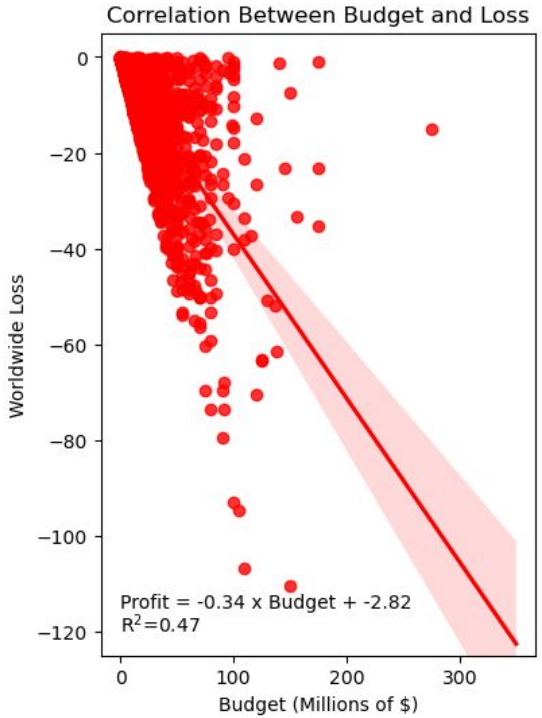


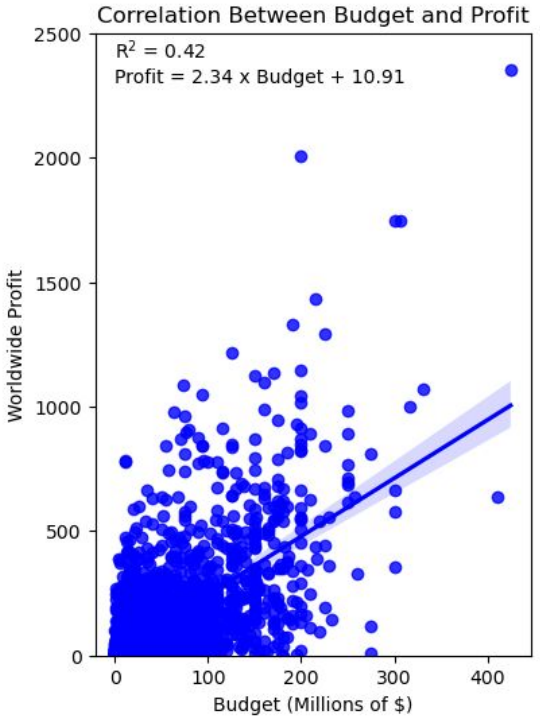

- EDA Tools: Python Pandas, SQL Lite, Numpy
- Visualization Tools: Seaborn and Matplotlib
- Statistical Tools: T-test, Linear regression modelling

# Investment Risk and Profit

## A higher investment may yield higher profits or higher losses

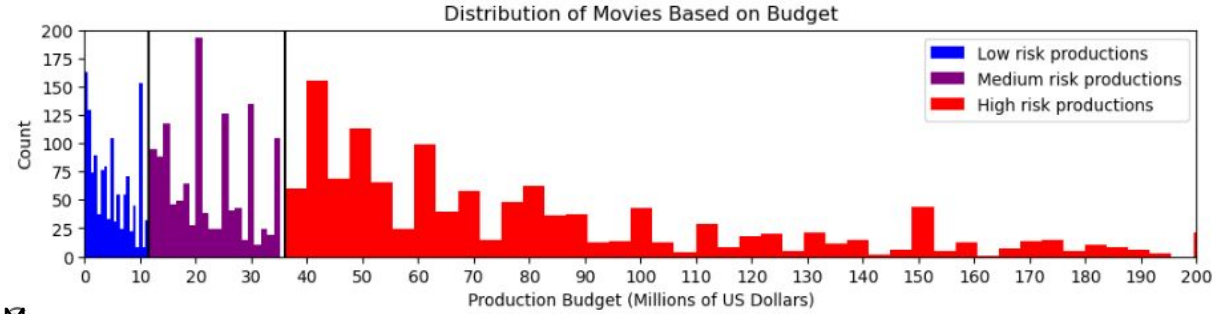
Correlation between higher budget and higher losses is positive.

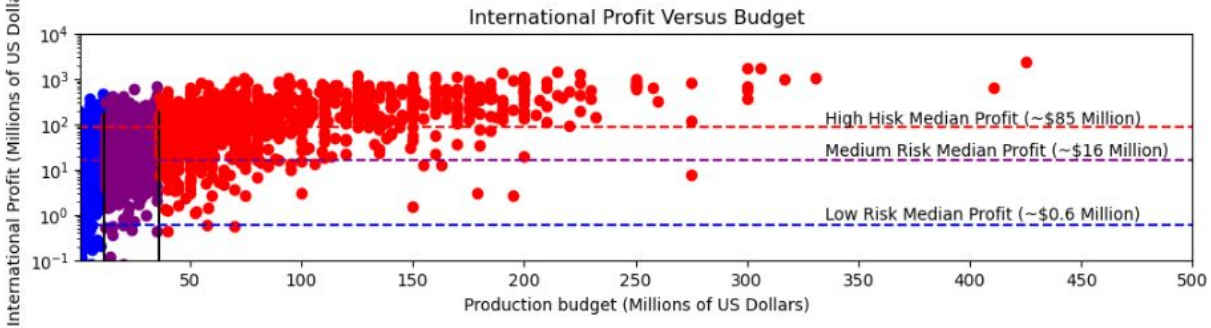


# Investment Risk and Profit

- Low Risk: Budget<$11.5M
- Medium Risk: $11.5M< Budget <$36M
- High Risk: Budget>$36M




**We split recommendations in low, medium and high risk based on profit/losses!**

# Fluxogram on risk analysis

Join The Numbers data with IMDB and Movie DB

# Fluxogram on risk analysis

Join The Numbers data with IMDB and Movie DB

Obtain genres with top profit

| | genres | worldwide_net |
|---|---|---|
| 0 | Horror | 15434588.0 |
| 1 | Mystery | 14371155.5 |
| 2 | Romance | 2681937.0 |
| 3 | Fantasy | 2438210.0 |
| 4 | Thriller | 1889796.0 |

# Fluxogram on risk analysis


## Horror

```
T scores for profit of:
$20 million = -3.6017645678829213
$35 million = -0.6797071532264725
$50 million = 2.242350261429976

Probability for profit of:
$20 million = 0.9997733937949502
$35 million = 0.7510316219789762
$50 million = 0.013343571801237686
Average cost of horror movies below $11.50 million are $4.51 millions
```

## Mystery

```
T scores for profit of:
$20 million = -2.769851906203676
$35 million = 0.009069978403655641
$50 million = 2.7879918630109874

Probability for profit f:
$20 million = 0.9966272916473534
$35 million = 0.49639116153943286
$50 million = 0.003203179748977769
Average cost of mystery movies below $11.50 million are $4.97 millions
```

# Fluxogram on risk analysis

Join The Numbers data with IMDB and Movie DB

Obtain genres with top profit

T-Test each genre distribution and select most profitable

Find best directors from genre

| | director | worldwide_net |
|---|---|---|
| 0 | Jordan Peele | 250367951.0 |
| 1 | John R. Leonetti | 250362920.0 |
| 2 | Tod Williams | 174512032.0 |
| 3 | Henry Joost | 169928918.0 |
| 4 | Adam Robitel | 157885588.0 |

# Fluxogram on risk analysis

Join The Numbers data with IMDB and Movie DB

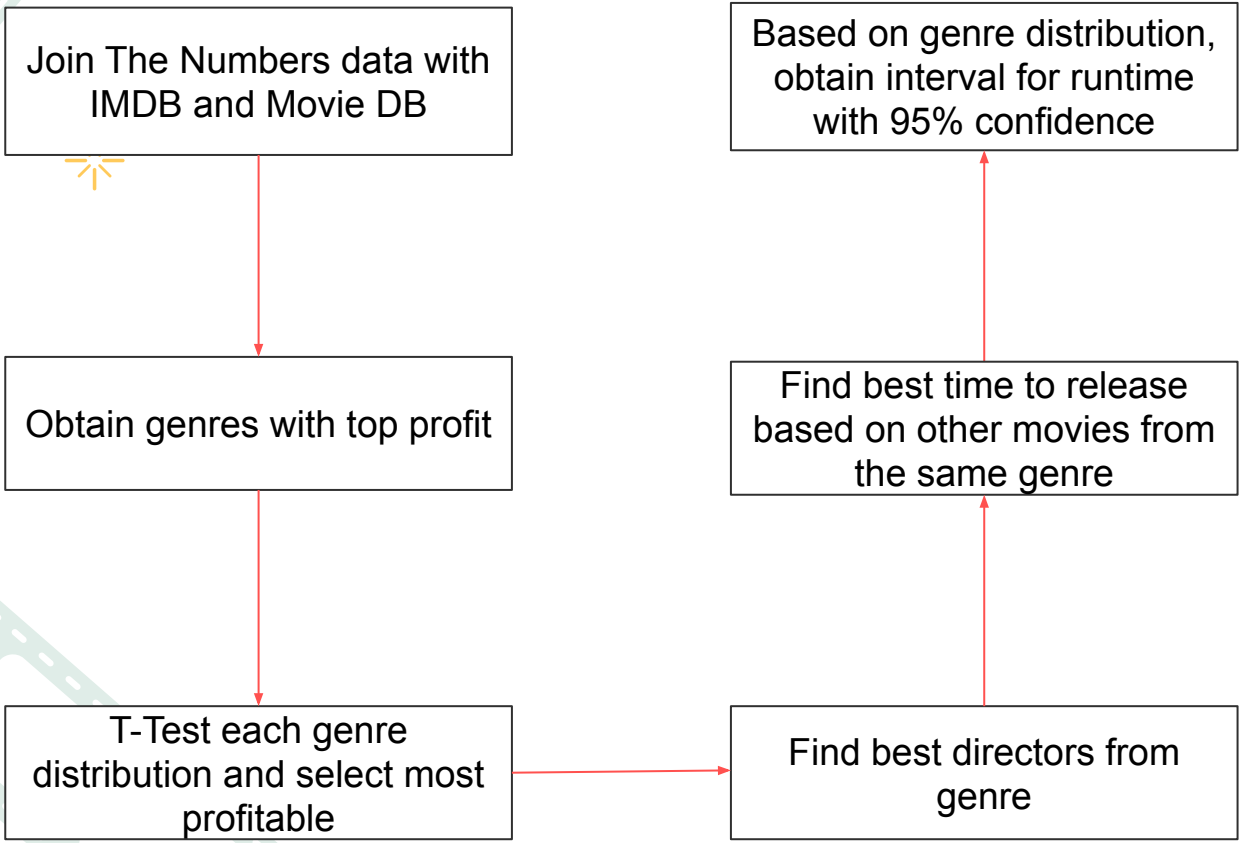
Obtain genres with top profit

T-Test each genre distribution and select most profitable

Find best directors from genre

Find best time to release based on other movies from the same genre

# Fluxogram on risk analysis



# Recommendations

## Low Risk:

- Genre: Horror
- Director: Jordan Peele or John R. Leonetti
- Best release date: Halloween (October) or Easter (April)
- Runtime: 96 to 100 minutes
- Average cost: $7.53 millions

## Medium Risk:

- Genre: Music
- Director: Damien Chazelle (La La Land)
- Best release date: August
- Runtime: 96 to 113 minutes
- Average cost: $21.93 millions

## High Risk:

- Genre: Sci-Fi
- Director: Michael Bay (Transformers), Anthony Russo (Marvel), Francis Lawrence (Hunger Games)
- Best release date: March (Spring Break) or June (Summer)
- Runtime: 112 to 120 minutes
- Average cost: $155.73 millions


# Thank you


**Cayke Felipe dos Anjos**

caykefelipe01@gmail.com

@cayke-fda

cayke-fda

**Dolgor Purbueva**

pdolgora@gmail.com

@dolgorp

purbuyeva