MOVIE INDUSTRY ANALYSIS

DSF-PT13 CLASS GROUP 6

17/12/2025

GROUP 6
- Abigael Kariuki
- Sylvester Nyachieo
- Ivy Cheptim
- Briege Meso
- Raphael Ndemo
- Lucy Muchiri

BUSINESS PROBLEM

 A new company is interested in creating original video content. They have been seeing big companies doing it and they want to join in on the fun. They have decided to create a new movie studio, but they don't know anything about creating movies. They have procured our services as Data Scientist and have charged us with exploring what types of films are currently doing the best at the box office. As Data Scientists, we are charged with translating our findings into actionable insights that the head of the new company movie studio will use to help them decide what type of films to create.

PROJECT OBJECTIVE

The objective of this project is to support the company's decision to launch a new movie studio by using data-driven insights. Specifically, this analysis aims to:
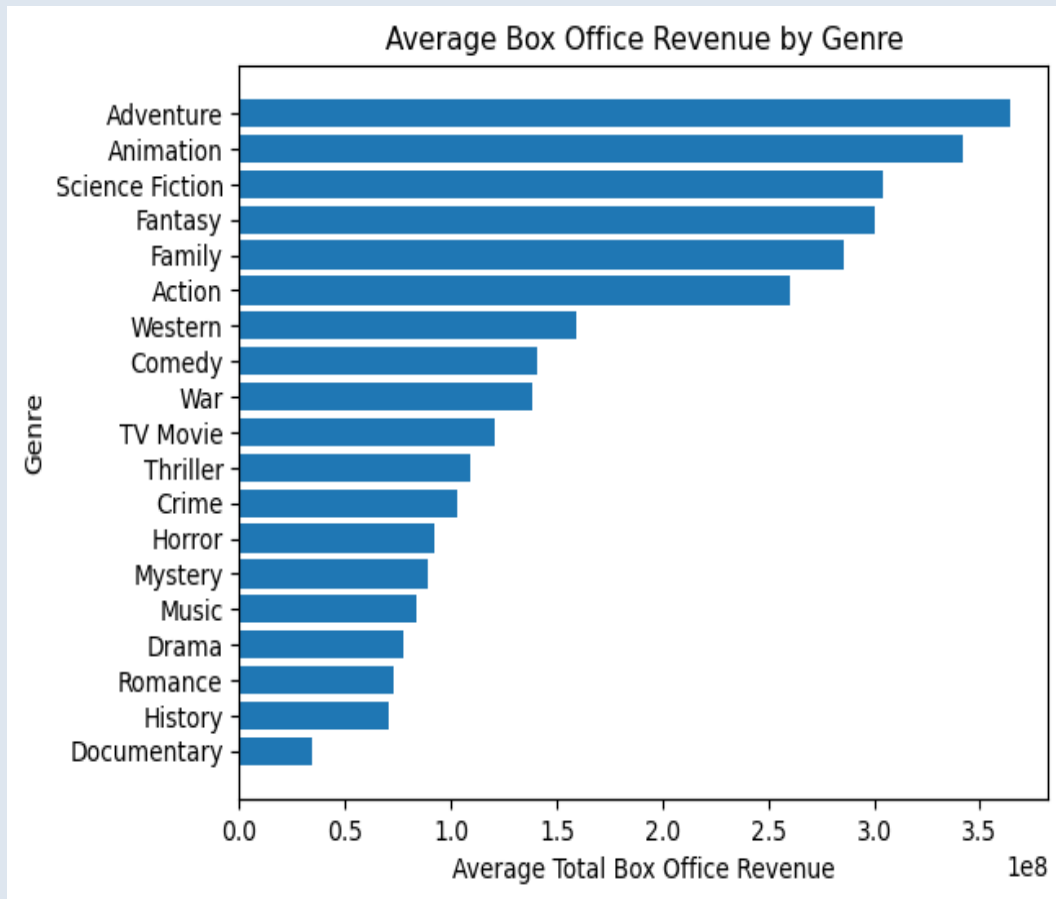
- To identify and prioritize genres based on revenue performance
- To determine audience ratings profiles to target
- To analyze the effect of production cost to revenue
- To evaluate revenue distribution but geography to determine top performing regions

Databases Analyzed:

- Box Office Mojo (revenue data)
- IMDB
- The Movie Database (TMDB - genres, ratings)
- The Numbers (production budgets)
- Rotten Tomatoes (critic reviews)

The project will use EDA and there after translate the findings into a clear and actionable recommendations for the head of the new movie company.

Objective 1- To identify and prioritize genres based on revenue performance.

Average Box Office Revenue by Genre

From the graph, we compare how genres perform based on average box office revenue

Top Performing Genres (Average Revenue):
Adventure - $364M average revenue (303 movies)
Animation - $341M average revenue (130 movies)
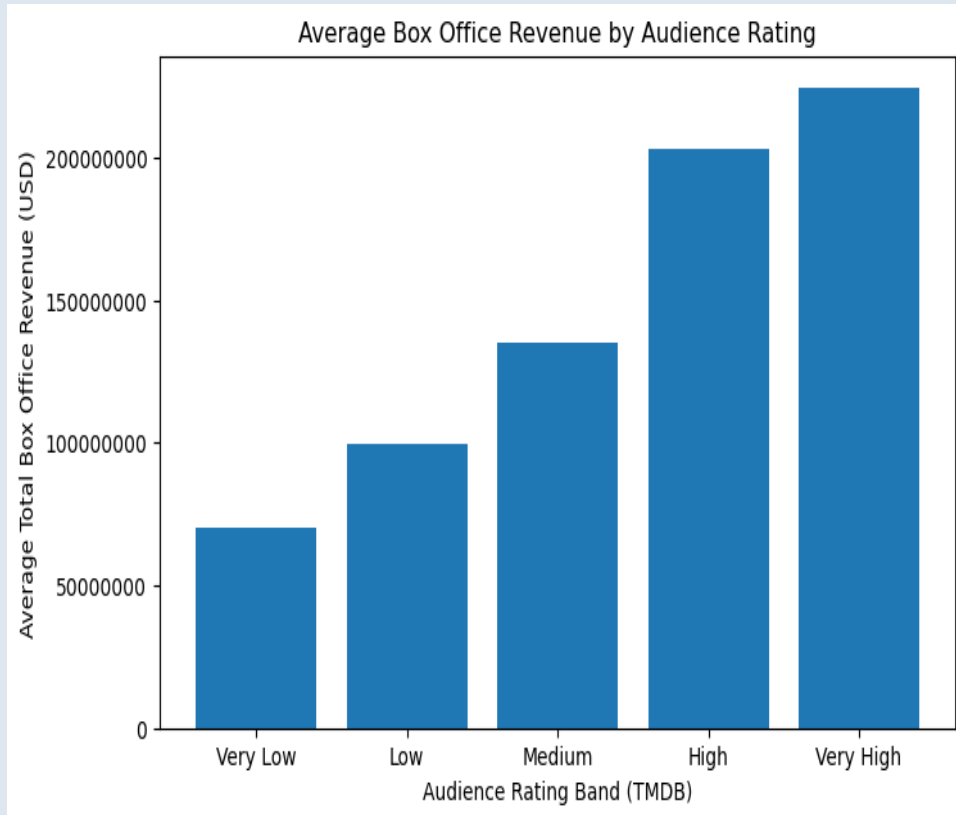**Science Fiction** - $304M average revenue (180 movies)
Fantasy - $300M average revenue (173 movies)
Family - $285M average revenue (175 movies)

These results suggest adventure and animation deliver stronger and more consistent box office returns than most other genres. In contrast, our analysis suggests lower revenue potential in genres such as horror, music, and documentary, so investing heavily in these may deliver weaker returns

Additionally, **Statistical Validation** using ANOVA test confirms significant revenue differences between genres (p-value < 0.05)
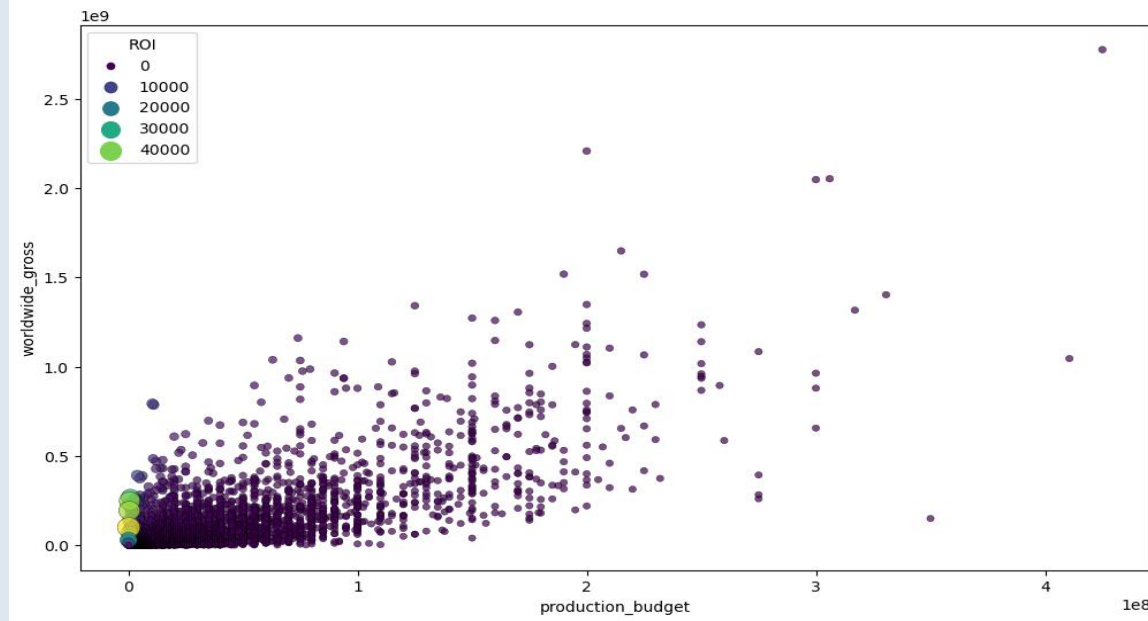
Revenue relationship with rating

Average Box Office Revenue by Audience Rating

- We grouped audience ratings into five bands: very low, low, medium, high, and very high. Average box office revenue rises as audience ratings increase.

- Low and very low rated films generate the lowest average revenues, while highly rated films perform better at the box office. High-rated films (7.0–8.0) average $203M in revenue.

- Very high-rated films (8.0–10.0) average $224M in revenue. An independent t-test confirms that high-rated films earn significantly higher average revenue than low-rated films.

Additionally, Statistical Validation using independent t-test confirms that high rated movies have higher average revenues than low rated films((p-value < 0.05)

Objective 3 - Analysis the effect of production budget to revenue(worldwide gross)
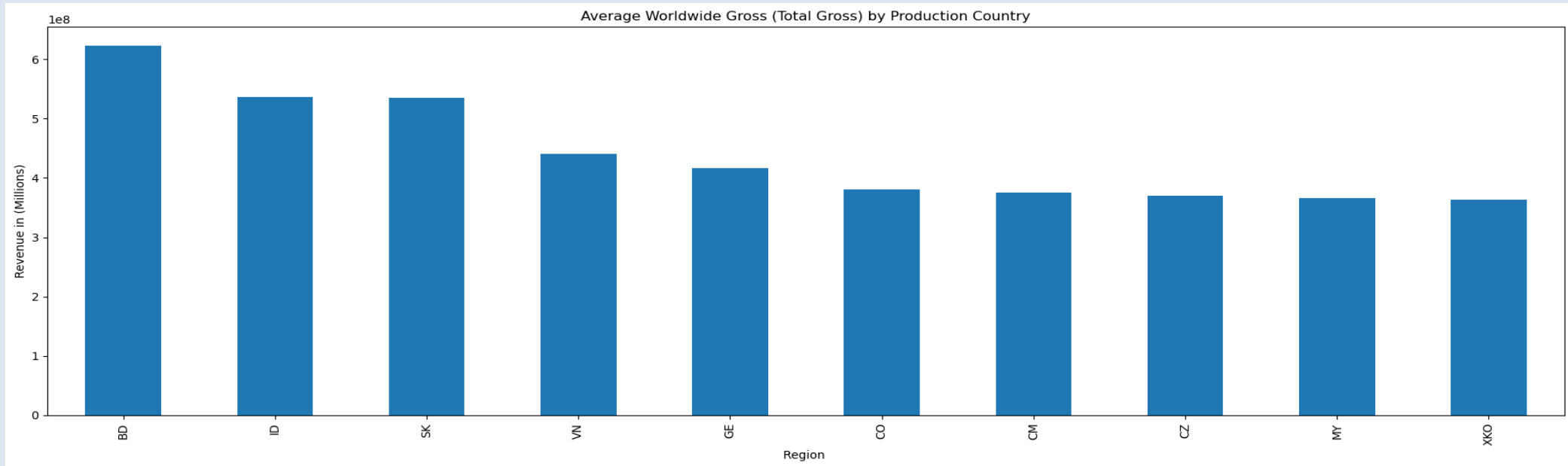

From our scatter plot, which is some what cone shaped, the graph shows an upward trend showing that higher production budgets can lead to higher worldwide gross. However the relationship is not strong or guaranteed- many high budget films do not achieve proportionally high returns.

There is also a huge variability at all production budgets. Low and mid budget films show a widespread in world gross, meaning some achieve strong box office success while other perform modestly.

Large bubble sizes indicating a higher ROI are mostly observed among low-to-mid budget films, indicating that smaller investments can sometimes lead to high profitability.
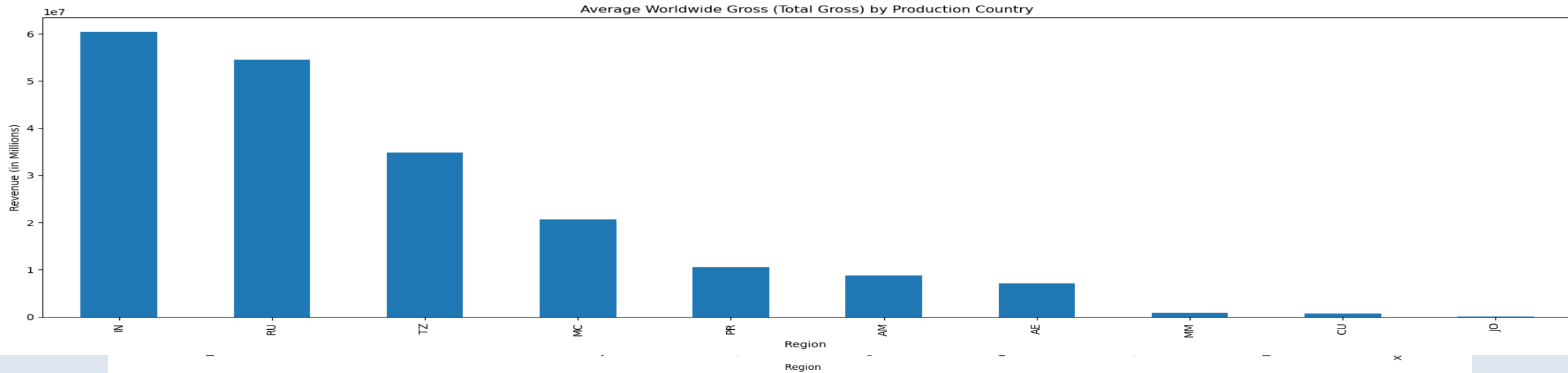
Therefore, investing in moderate-budget films may provide better risk-adjusted returns than consistently funding very high budget productions

Objective 4 - To evaluate revenue distribution by geography to determine top-performing regions

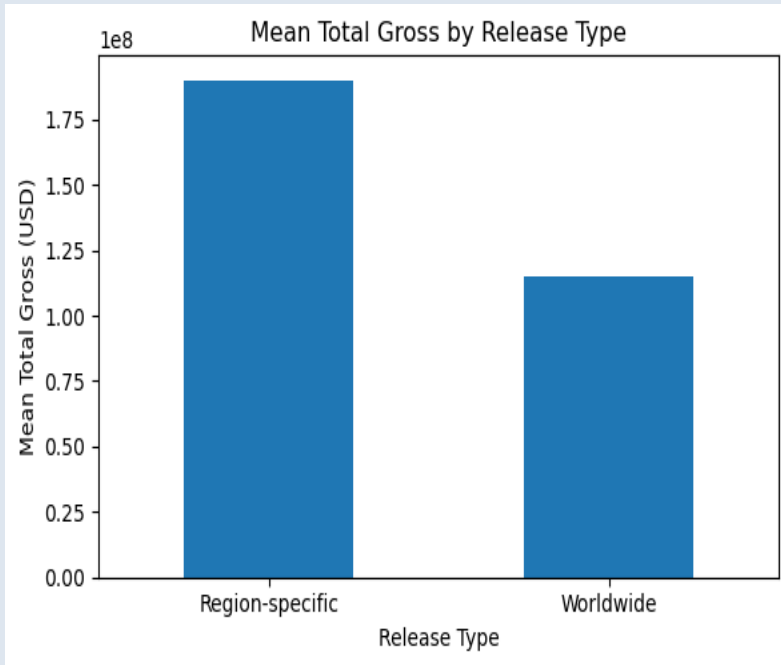Average Worldwide Gross (Total Gross) by Production Country

- Revenue is concentrated in a few regions. BD leads with $623M in total revenue, while ID and SK also perform strongly at $530M+ each.

- Overall, the chart suggests a small group of regions drive a large share of box office revenue in the IMDB sample.

...cont (Bottom 10 regions by revenue)

Average Worldwide Gross (Total Gross) by Production Country

- For the bottom 10 regions, IN , AM, IZ, and MC generate lower total revenue, indicating weaker box office performance in this dataset.

- This could reflect lower ticket sales, audience preferences, and language or distribution constraints. Smaller markets like Monaco and Armenia may also be limited by population size.

...Further analysis

We also went ahead to answer the question whether globally released films earn significantly more than region-specific films.
Two categories were created, region-specific and worldwide. From the distribution, region-specific movies were 5,167 and worldwide were 303.
One-way ANOVA was used to compare the mean total gross revenue among the different release types(genres)
Hypothesis testing:
Null Hypothesis - The mean total gross revenue is the same for globally released films and region-specific films.
- Alternative Hypothesis - The mean total gross revenue differs between globally released films and region-specific films.

The average comparison indicates that region-specific films have a higher mean total gross compared to worldwide, which implies effectiveness on targeted market.

Conclusion and Recommendation

- Prioritize genres that consistently deliver large box office revenues for commercial-scale investments such as Animation, Fantasy and Family films consistently generate the highest revenues.

- Audience reception and rating is an important signal for commercial performance of genres.

- For investments, the new company should focus producing their movies in high-revenue regions which show consistent box office returns . This will be done through prioritizing marketing, distribution and content tailored to these audiences. Strategically, avoid low-revenue regions.

- Additionally, adapt movie genres based on what resonates in top-performing regions and can also consider co-production or partnerships in strong markets to minimize risks and maximize reach

- Investing in moderate-budget films may provide better risk-adjusted returns than consistently funding very high budget productions. That high-budget films don't necessarily mean that the ROI will be high.