



Elina Rankova and Bekah McLaughlin

# Limelight Studios Launch

Phase 1 Data Analysis

# I

## **Founder Mission**

Business Problem and understanding

# II

## **Data Summary**

Brief dive into data exploration and analysis

# III

## **Recommendations**

Data driven insights for initial launch

# IV

## **Looking Ahead and Limitations**

Data gaps and a brief look at phase 2



# Founder Mission

- **PRIORITIZE PROFITABILITY FOR SUSTAINABILITY:** Identify the most profitable movie types & pinpoint those with lower production budgets
- **INCLUSIVITY AND GLOBAL REACH:** Aim to reach a global audience and invest in potential untapped markets
- **DEDICATION TO CRAFT:** Collaborate with top industry professionals & create high-quality, appealing films that resonate with audiences worldwide.

# Data Summary

- Assessing datasets for relevancy and usability
- Cleaned up missing, duplicate, and irrelevant entries
- Merged datasets for further analysis

Dataset	Contents	Size	Source
TN.MOVIE_BUDGETS.CSV	PRODUCTION BUDGET, WORLDWIDE GROSS, MOVIE TITLE, RELEASE DATE	4,045 MOVIES	THE-NUMBERS.COM
TMDB.MOVIES.CSV	MOVIE TITLE, RELEASE DATE, AVERAGE RATING, VOTE COUNT, POPULARITY	22930 MOVIES	THEMOVIEDB.ORG
IM.DB DATABASE	WRITERS, DIRECTORS, FILM PROFESSIONALS, MOVIE TITLES, RATINGS, VOTE COUNT, RELEASE YEAR	85232 MOVIES	IMDB



Final Data Analyzed	Size
TN.MOVIE_BUDGETS.CSV & TMDB.MOVIES.CSV	1297 MOVIES
IM.DB WRITERS AND MOVIE REVIEW TABLE	56714 MOVIES
IMDB DIRECTORS AND MOVIE REVIEW TABLES	64781 MOVIES

# Data Summary Cont.

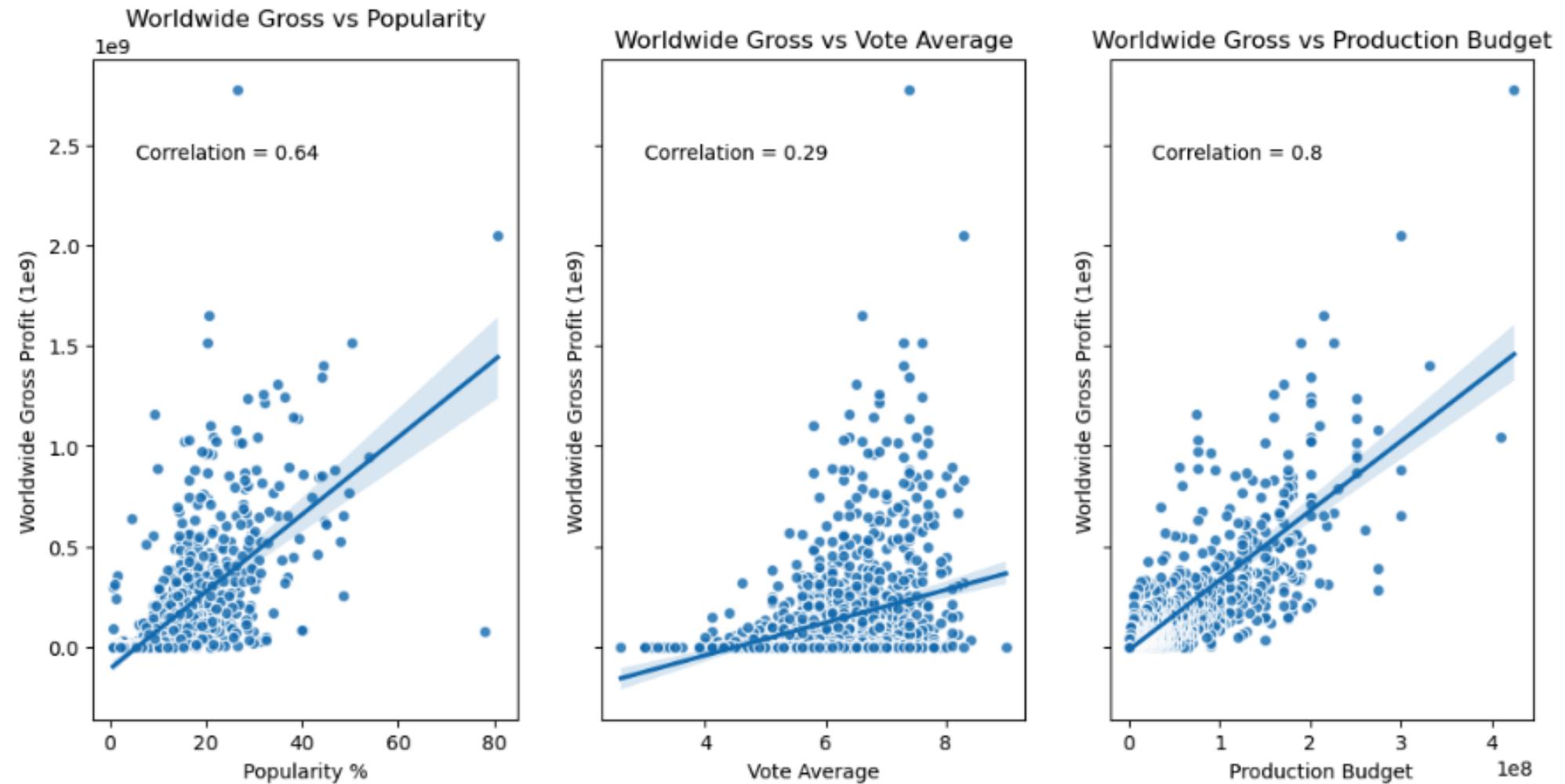
- Statistical analysis of data to confirm perceived patterns of initial exploration
- Correlation analysis movie gross factors
- Further filtering to find entertainment creator collaboration opportunities

Rule of Thumb for Interpreting the Size of a Correlation Coefficient<sup>4</sup>

Size of Correlation	Interpretation
.90 to 1.00 (−.90 to −1.00)	Very high positive (negative) correlation
.70 to .90 (−.70 to −.90)	High positive (negative) correlation
.50 to .70 (−.50 to −.70)	Moderate positive (negative) correlation
.30 to .50 (−.30 to −.50)	Low positive (negative) correlation
.00 to .30 (.00 to −.30)	negligible correlation

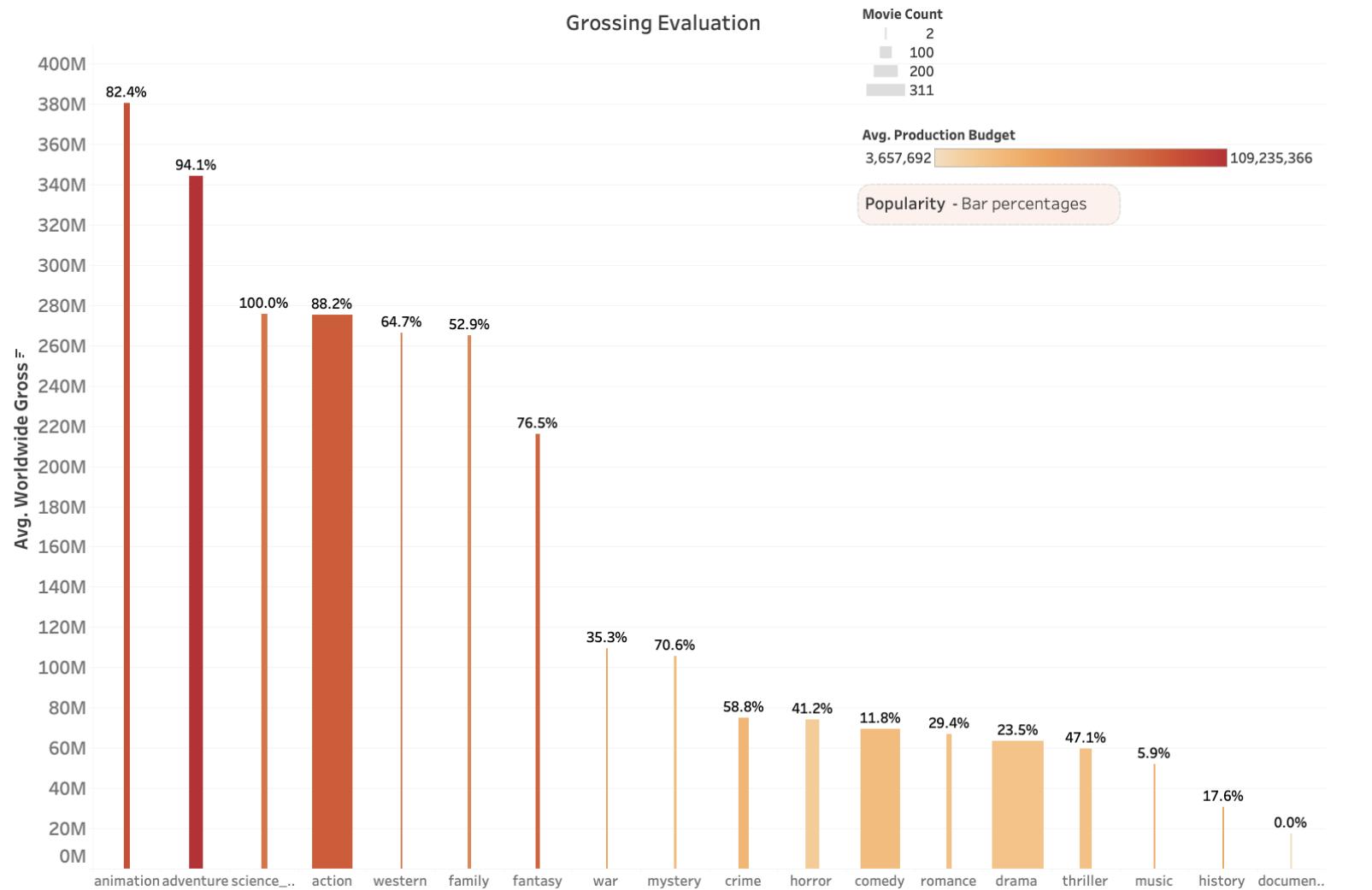
# Data Summary Cont.

We compared popularity, vote average, and production budget to worldwide gross

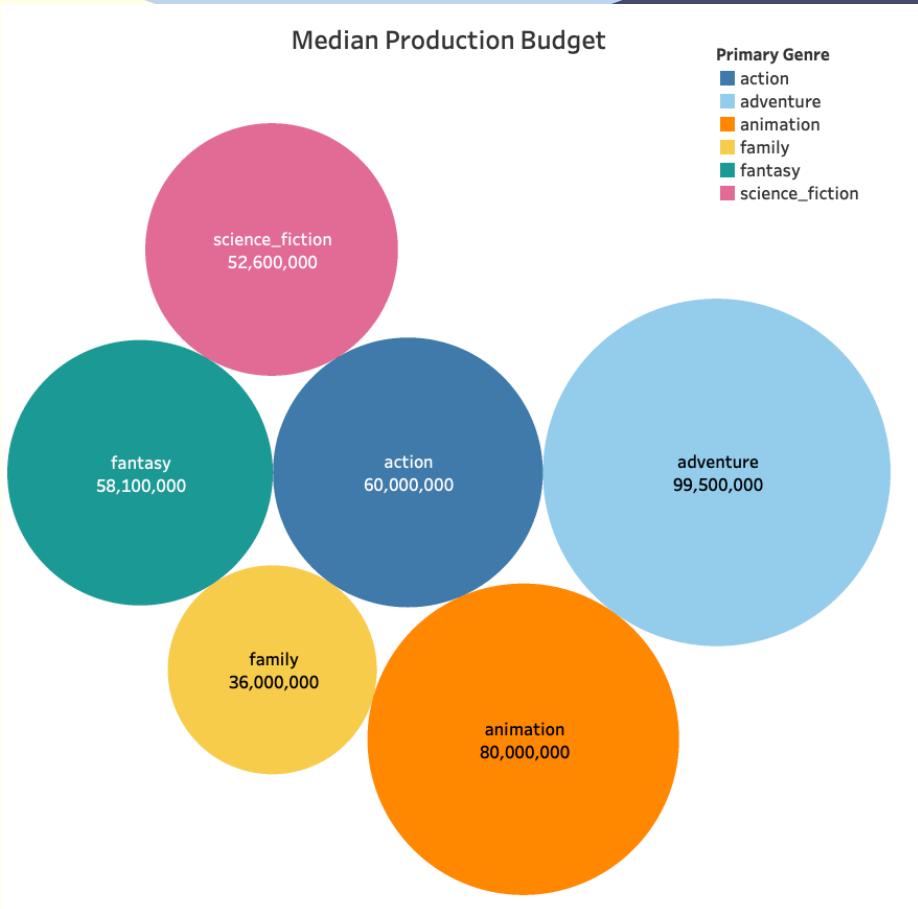


# Data Summary Cont.

- ANOVA test indicated worldwide gross varies depending on genre
- Thorough genre analysis, including mean worldwide gross, movie count, and mean production budget



# Recommendations and Limitations



1. TOP SIX GENRE RECOMMENDATIONS - Animation, Adventure, Sci-Fi, Action, Family and Fantasy
2. INITIAL BUDGET RECOMMENDATIONS - Budget aims per genre indicated in image
3. CREATOR OUTREACH RECOMMENDATIONS - Directors such as Erkut Altindag, Writers such as Usha Kaljai

## LIMITATIONS AND LOOKING AHEAD

- Missing data and data merge issues
- Data timeline limitations
- Phase 2 steps: recover missing data, deeper creator analysis, ROI analysis

# Questions for the Data Science Team?

Elina Rankova

<https://www.linkedin.com/in/elina-rankova-5b44b01bb/>



Bekah McLaughlin

<https://www.linkedin.com/in/bekah-mclaughlin-8a86b02b4/>

