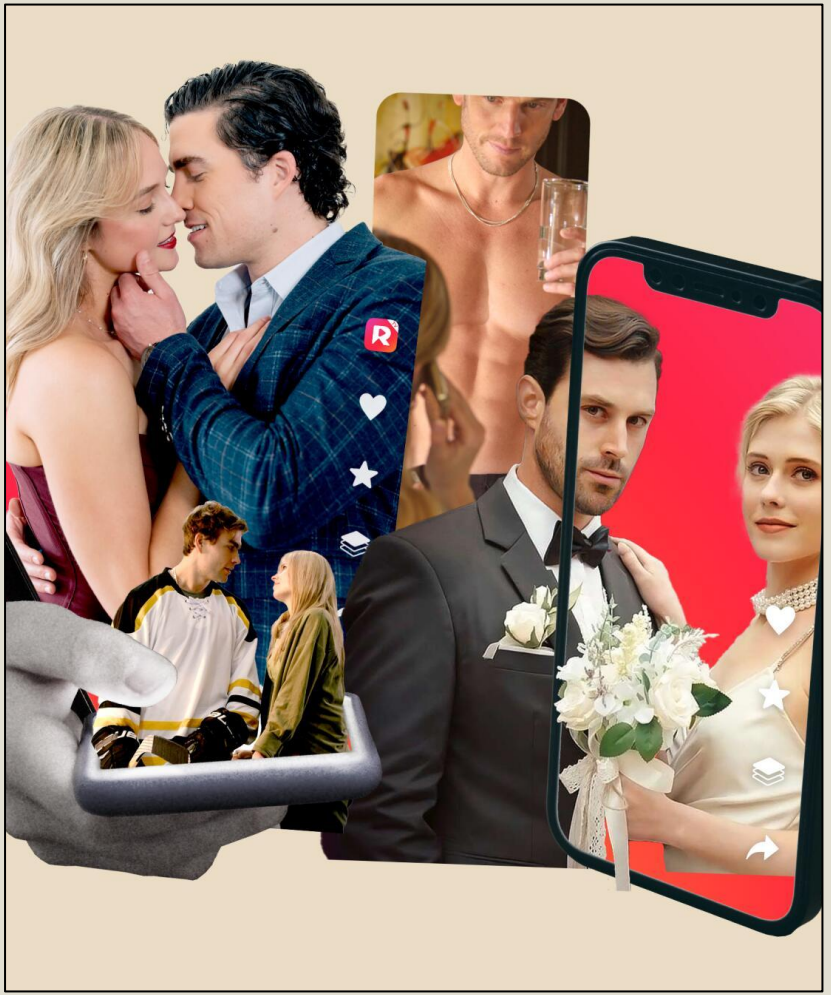
11/25/25
STUDENT: TITI LEE

USC DSCI510_FINAL PROJECT
PROFESSOR: ALEXEY TREGUBOV

VERTICAL MINIDRAMA DATABASE

/the first verticals only database



- ➔ What is the correlation between the social media following of leading actors on a show and the show’s overall popularity by views?
- ➔ Does an actor starring in a ‘hit’ predict a relevant increase of ‘hit’ shows?
- ➔ Does audience engagement correlate to overall views of a show?

/introduction

The ‘vertical drama’ (also known as ‘minidrama’ or ‘microdrama’) is a fast-growing form of video entertainment designed explicitly for mobile viewing and distributed through proprietary mobile apps.

KEY CHARACTERISTICS

- Introduced in the US in 2023 from China
- Shot in 9:16 aspect ratio
- Titles contain ~60-80 episodes, each 1-1.5 minutes long
- Typically fall in the romance genre
- Industry growth of 41% YoY as of October 2025

/data sources

Data Source 01: IMDB Show/Actor Credits

SOURCE URL: imdb.com
TYPE: Web Scraping (Semi-structured)
FORMAT: HTML → JSON → SQL DB

DATA SIZE: 321 Titles

Data Source 02: YouTube YouTube Comments

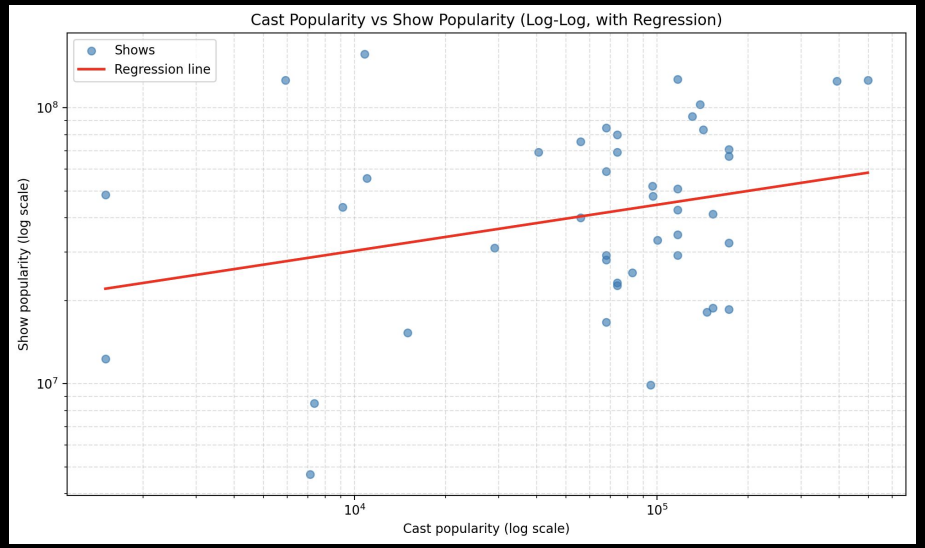
SOURCE URL: YouTube.com
TYPE: YouTube API
FORMAT: HTML response → CSV

DATA SIZE: 879 comments

Data Source 03: Instagram Social Media Following

SOURCE URL: Instagram.com
TYPE: Web Page
FORMAT: HTML response → SQL DB

DATA SIZE: 30 Actors



"WATCH OUT I'M THE LADY BOSS" (Figure 3d)

/results

SUMMARY OF RESULTS

- → Overall, there is a **weak correlation** between the social media following of the cast and the show's overall popularity (Figure 1a, 1b)
- → The average "hit rate" for actors is **20.42%** (Figure 2)
- → As show popularity increases, audience engagement grows, but **sentiment** becomes **mixed**. (Figure 3a, 3b, 3c)

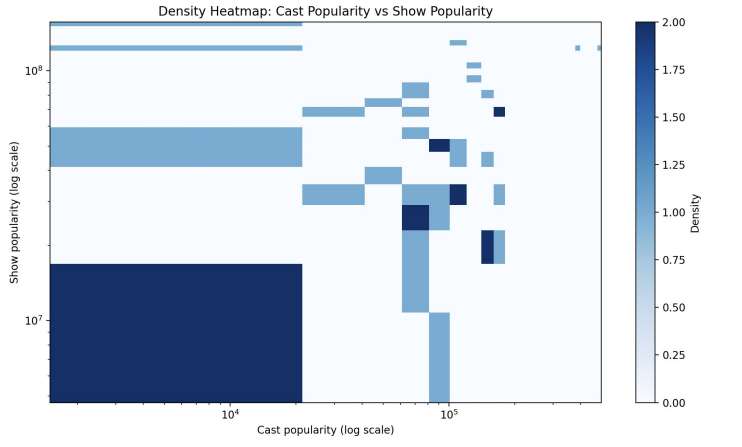
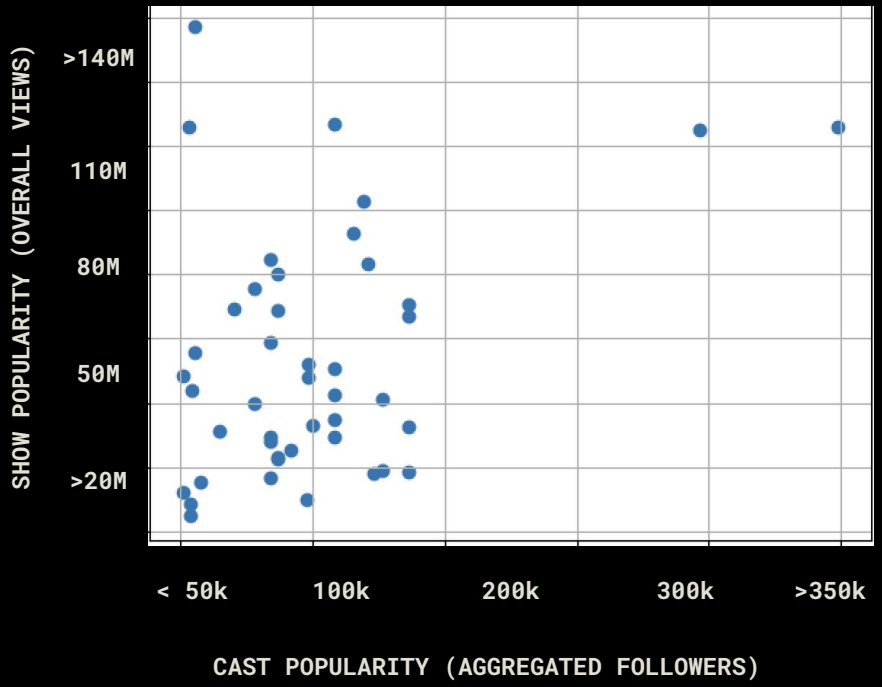


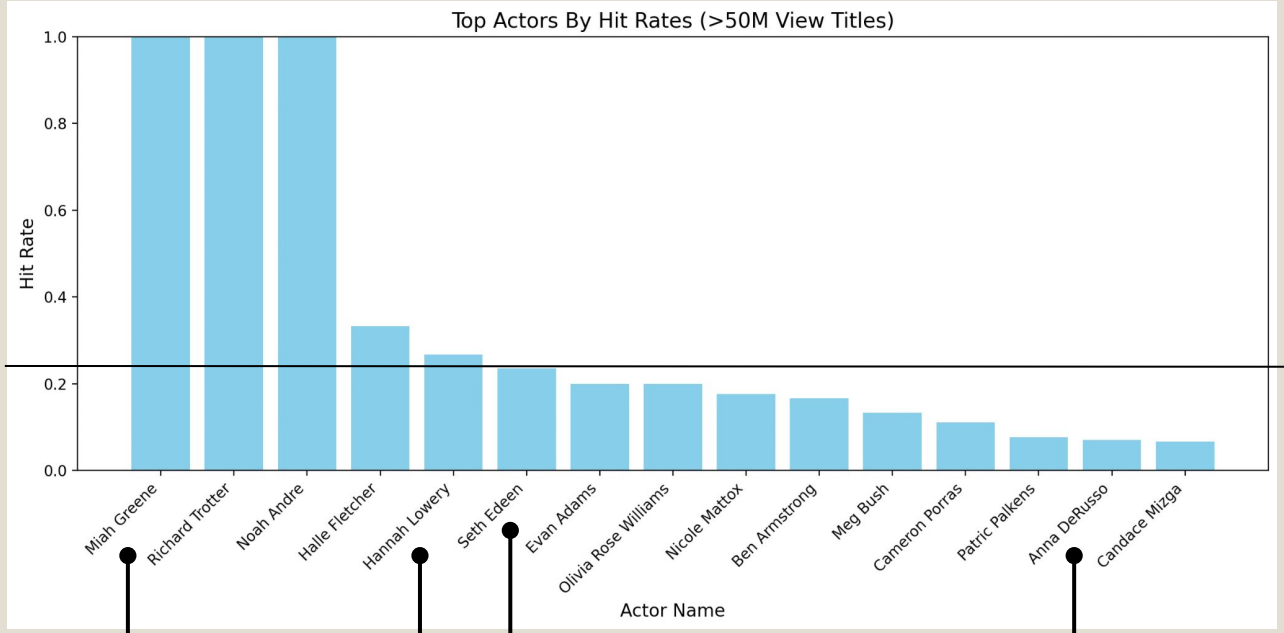
Figure 1b.
HEAT MAP (CAST VS SHOW POPULARITY)

Figure 1a.
SCATTERPLOT
CAST SOCIAL MEDIA VS SHOW POPULARITY



Correlation Coefficient = .3481

Figure 2.
TOP 15 ACTORS BY HIT RATES



AVERAGE HIT RATE: 20.42%

Calculating a "hit"

- **Over 50 million views**
- **Revenue > $3M**
- **Other internal markers**

Miah and Richard starred in 'How to Break a DILF' (500M+) only title listed

Hannah L: Social media following <100k

Seth E: social media following >100k

Anna D. stars in 'FYI My Ship Has Sailed Without You' (hit)

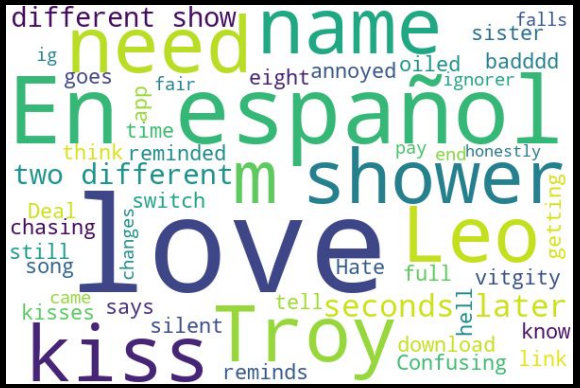
Figure 3a.
WORD CLOUD TO VISUALIZE AUDIENCE SENTIMENT

“30 YEARS FROZEN” (NETSHORT)

Audience engagement vs show popularity:

3a. **126.1M** views, **495** comments,
3b. **58.9M** views, **49** comments
3c. **100M** views, **75** comments
3d. **130M** views, **260** comments

“SUMMER SITUATIONSHIP” (REELSHORT)



“A DEAL WITH THE HOCKEY CAPTAIN” (DRAMABOX)

CONCLUSION

Based on the current data analysis, there is no evidence of strong correlation between actors with “heat” (aka recently starring in ‘hits’ or gaining large social media followings) and their ability to consistently star in future “hit” shows. This may be likely due to the rapidly changing tastes and interests of the audience, along with other variables that affect the show’s overall quality including script, production value, and marketing.

This is a growing industry and the “formula” for a “hit” is constantly changing. Actors with “hits” may command higher rates but this does not necessarily equate to higher revenue.

LOOKING AHEAD

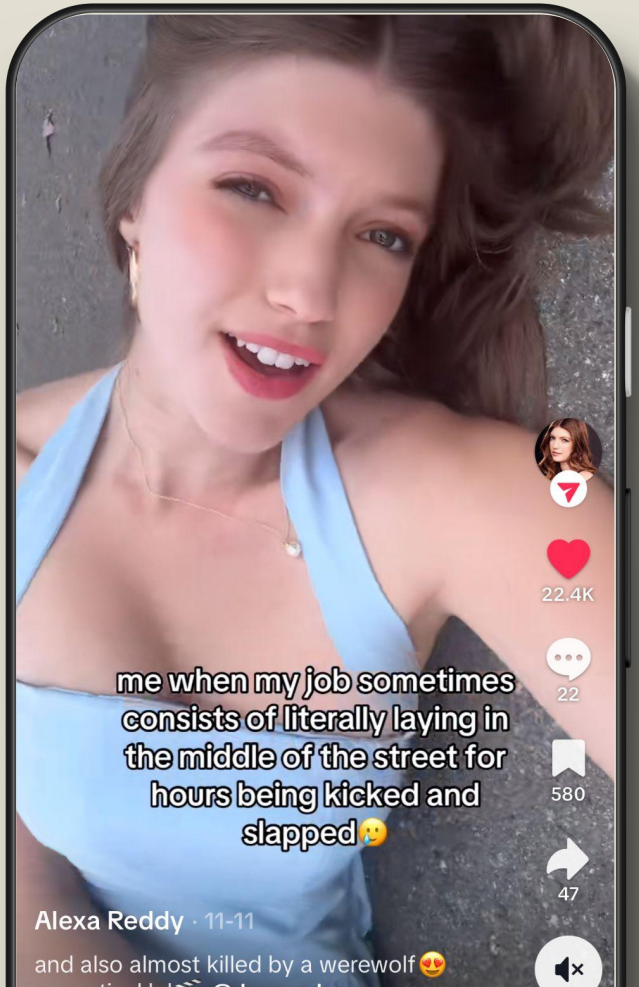
- Tracking views by release date vs audience sentiment over time (example: tracking first week vs first month)
- Do certain genres do better on certain platforms?
- Mapping co-star chemistry (hit shows when two actors play the lead together) across titles

CHALLENGES

Getting consistent data across competing platforms such as views and release dates proved harder than expected. While IMDB provides a starting point for previous titles, the data is messy and requires a lot of cleaning time to sort for vertical titles only. (Example, some shows would be categorized as “TV Movie” and others as “Mini-series” on IMDB).

I also faced roadblocks trying to extract follower counts from social media, as Instagram & Tik Tok hide their follower count under Javascript.

While the word cloud was interesting, it could be much improved to include aggregated comments across multiple sources such as Reddit, Tik Tok, Instagram Reels, and be modified to take key word arguments such as a show title or actor’s name rather than just a single video ID.



QUESTIONS?

➔ Email: tlee6409@usc.edu

➔ IG: @teresaleebot

THANK YOU

If you’re seeing this slide, it means my baby did not come early. Thank you, baby.