

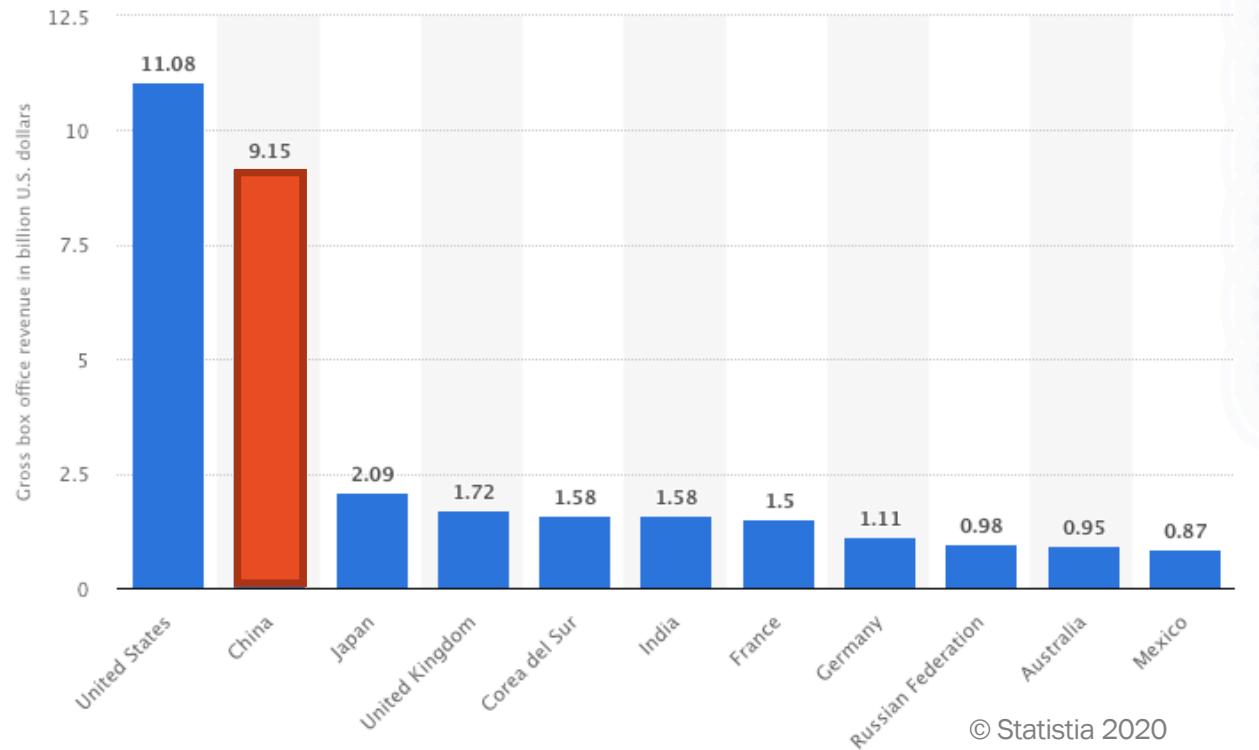
CHINESE BOX OFFICE ANALYSIS

Interpreting International
Film Performance in
Mainland China



What drives second largest box office in the world?

- Mainland Chinese box office revenue alone constitutes roughly 1/3 of the global market
- Greater understanding of this market can significantly boost worldwide film revenue

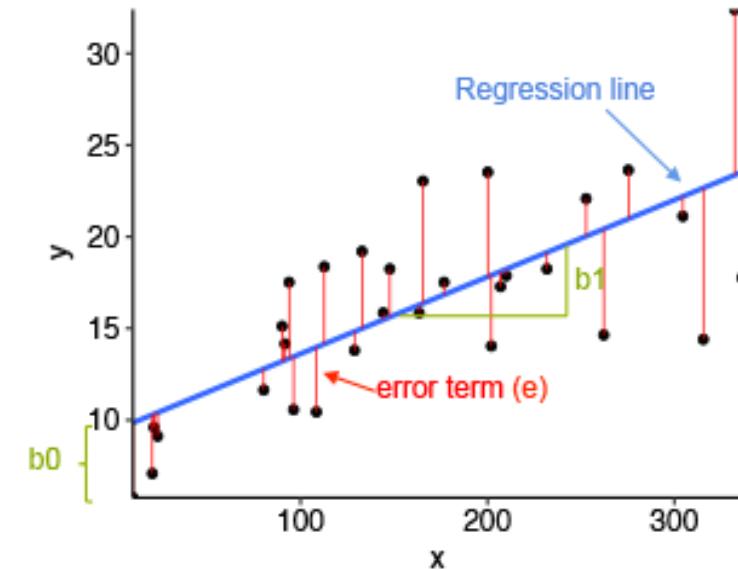


METHODOLOGY



Methodology

- GOAL:
Generate key insights to improve international film marketing strategy
- PROCESS:
use linear regression to interpret total gross revenue of internationally released films in the mainland Chinese market

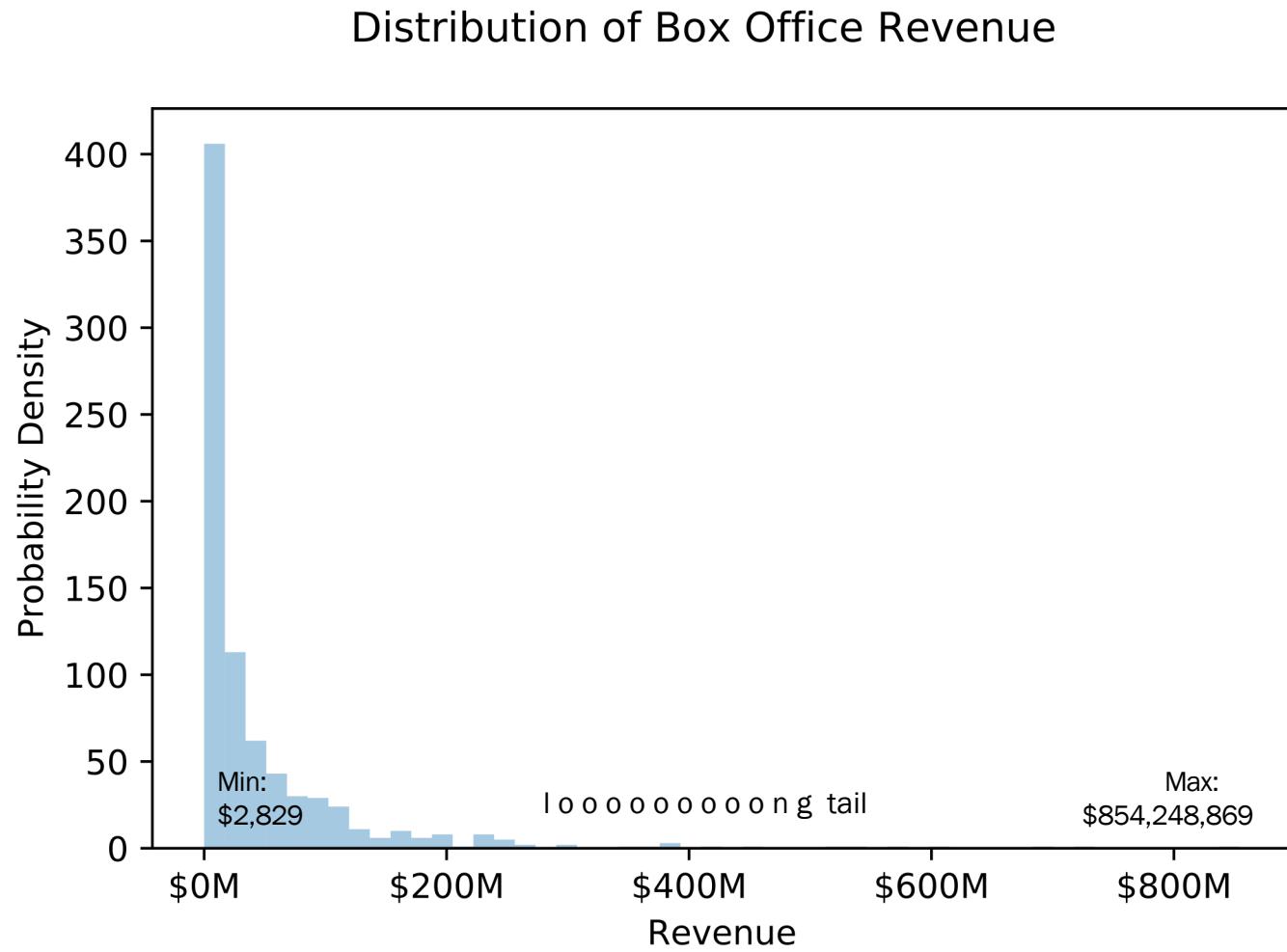


Methodology

- All data scraped from the following sources:
 - **Box Office Mojo** – *worldwide box office data 1994-2020*
 - out of ~16K films total, only ~800 released in China
 - HK excluded since it requires a unique marketing approach
 - **The Numbers** – *top 100 earning international actors*
 - **Wikipedia** – *top 10 highest-grossing film directors worldwide*



Target variable: box office revenue

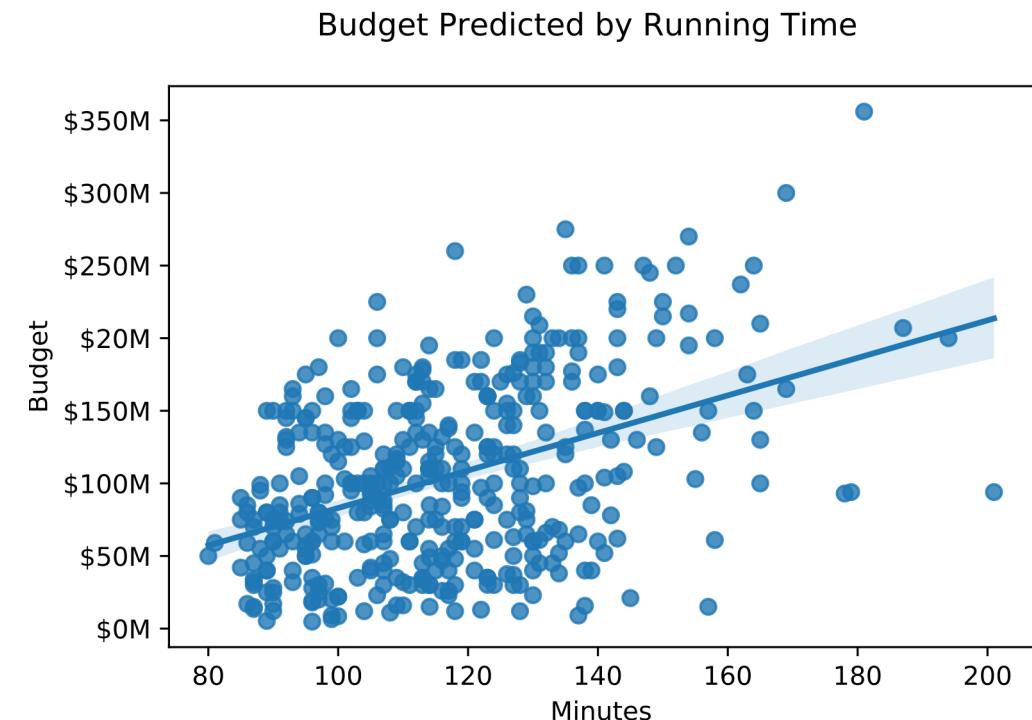


Challenges:

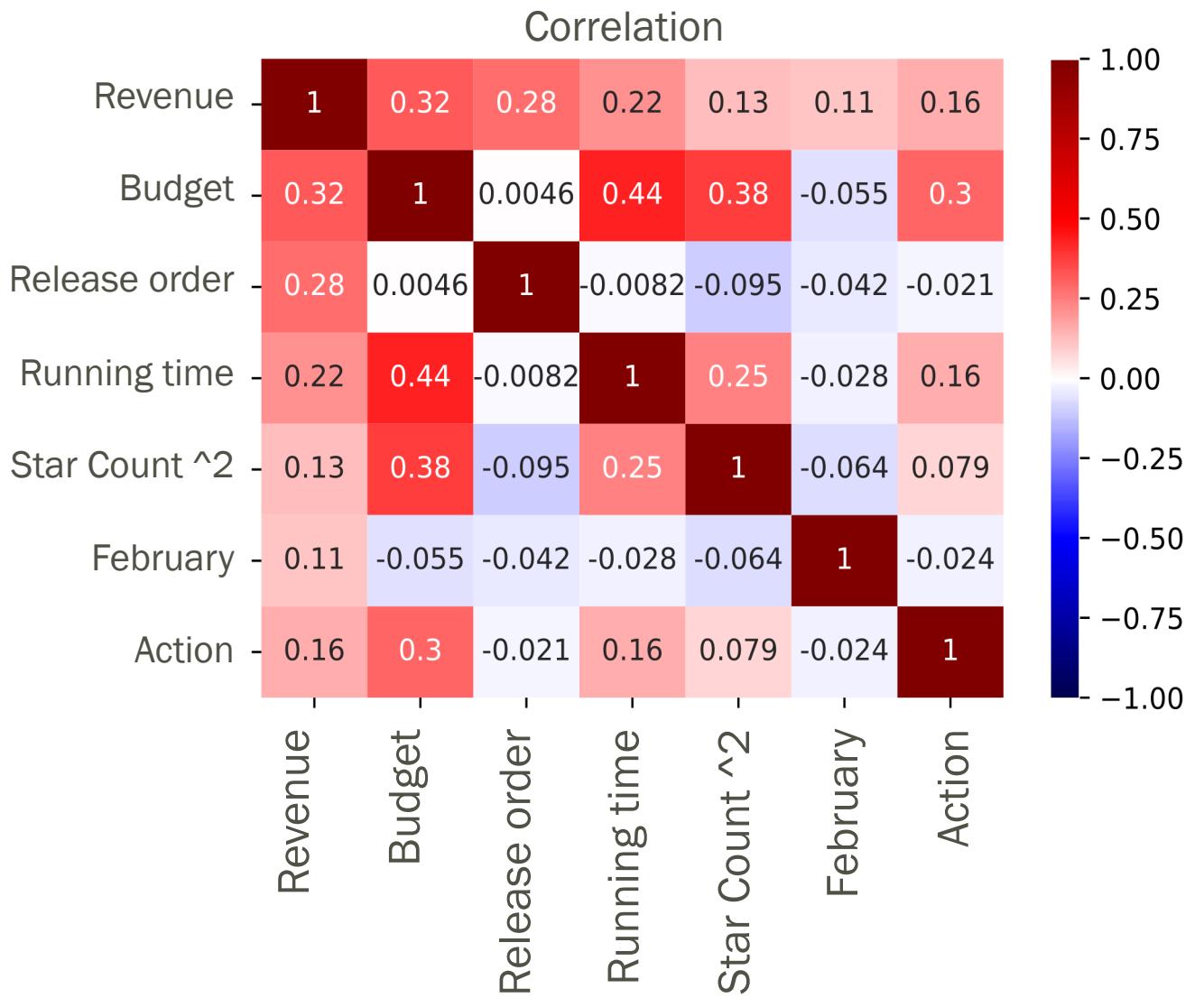
- Target variable distribution heavily skewed.
- Lack of strong correlation
- Imputing missing data

A model within a model...

- Problem:
 - *Budget influential to revenue prediction but nearly 50% of films missing budget info*
- Solution:
 - *Constructed linear regression model that could predict budget based using:*
 - Running time (see right), Genre, Worldwide gross revenue, Cast
- Results:
 - *2/3 of budget variance can be explained by model*
 - $R^2 = 0.676$
 - *Predictions accurate within margin of ~\$26M*
 - $MAE = 25,779,893$

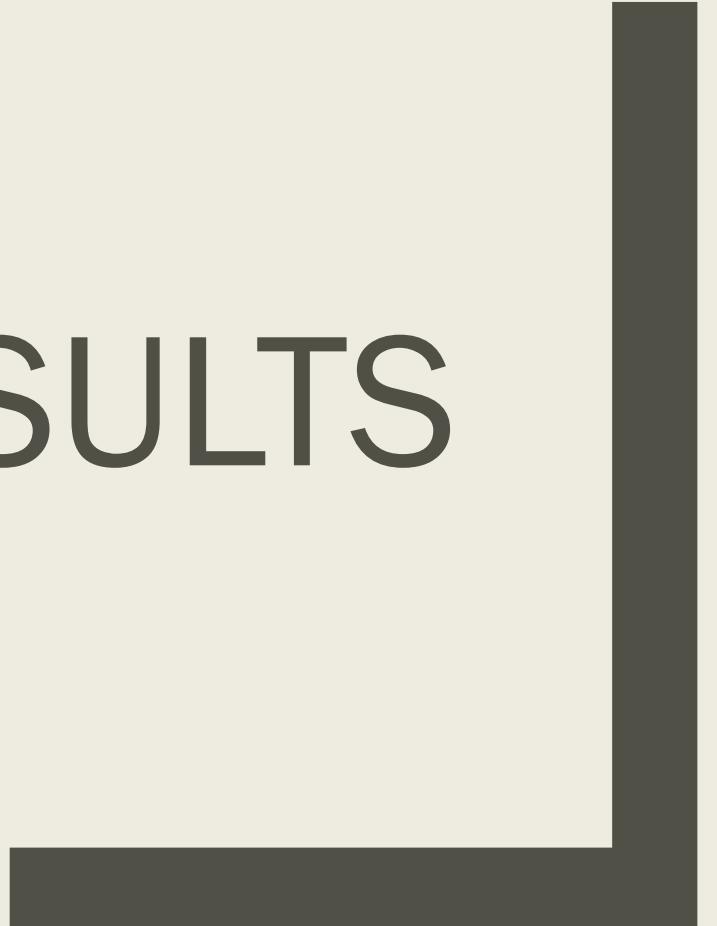


Key factors influencing revenue



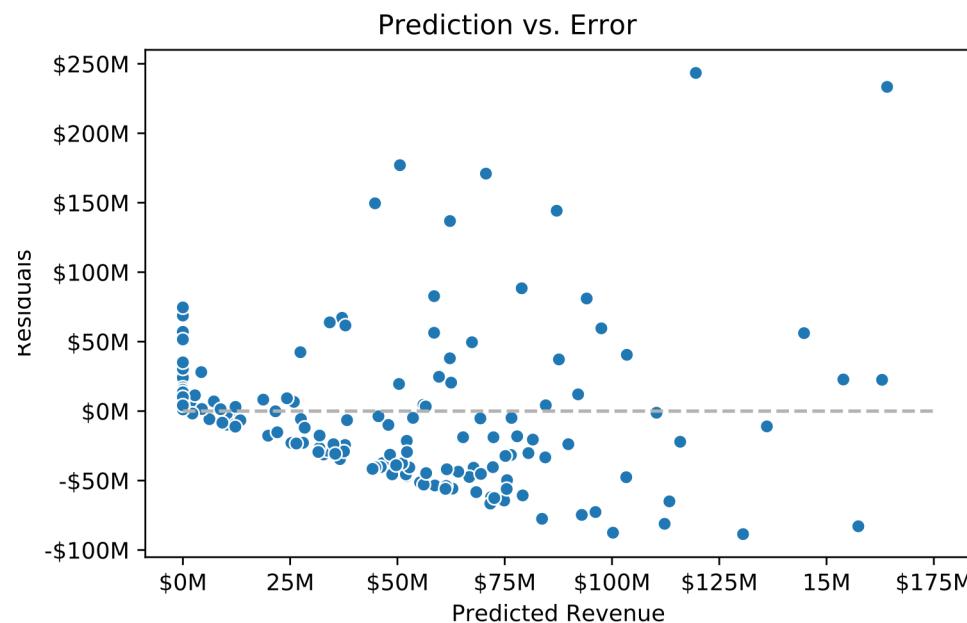
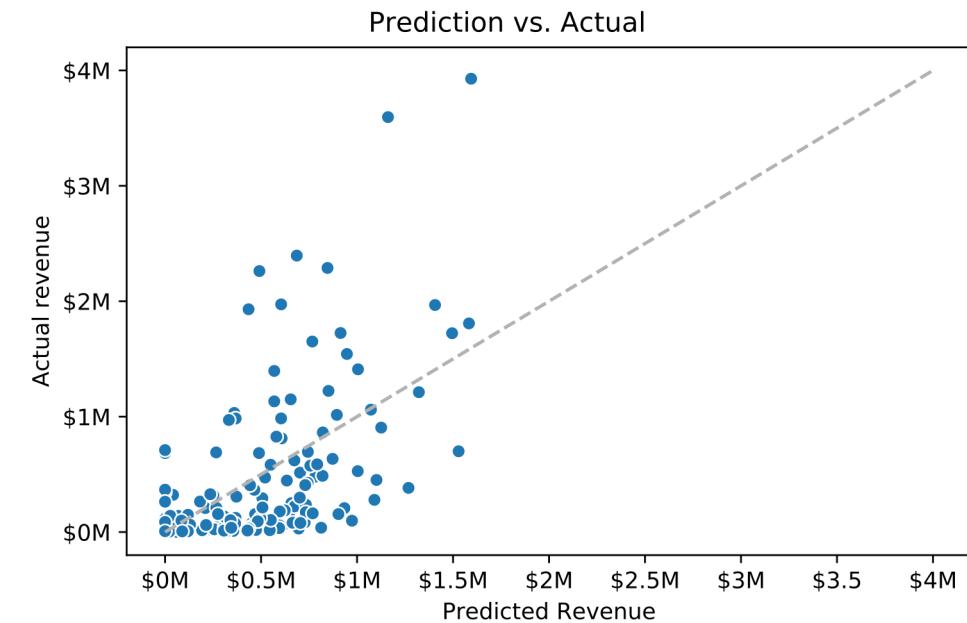
- Used domain knowledge and experimentation to engineer and select features with best correlation with revenue
 - A/B tested many combos using R^2 to judge result
 - Data set too small to benefit from regularization and/or widespread polynomial feature adoption
- Not pictured: many, many features that were left on the cutting room floor

RESULTS

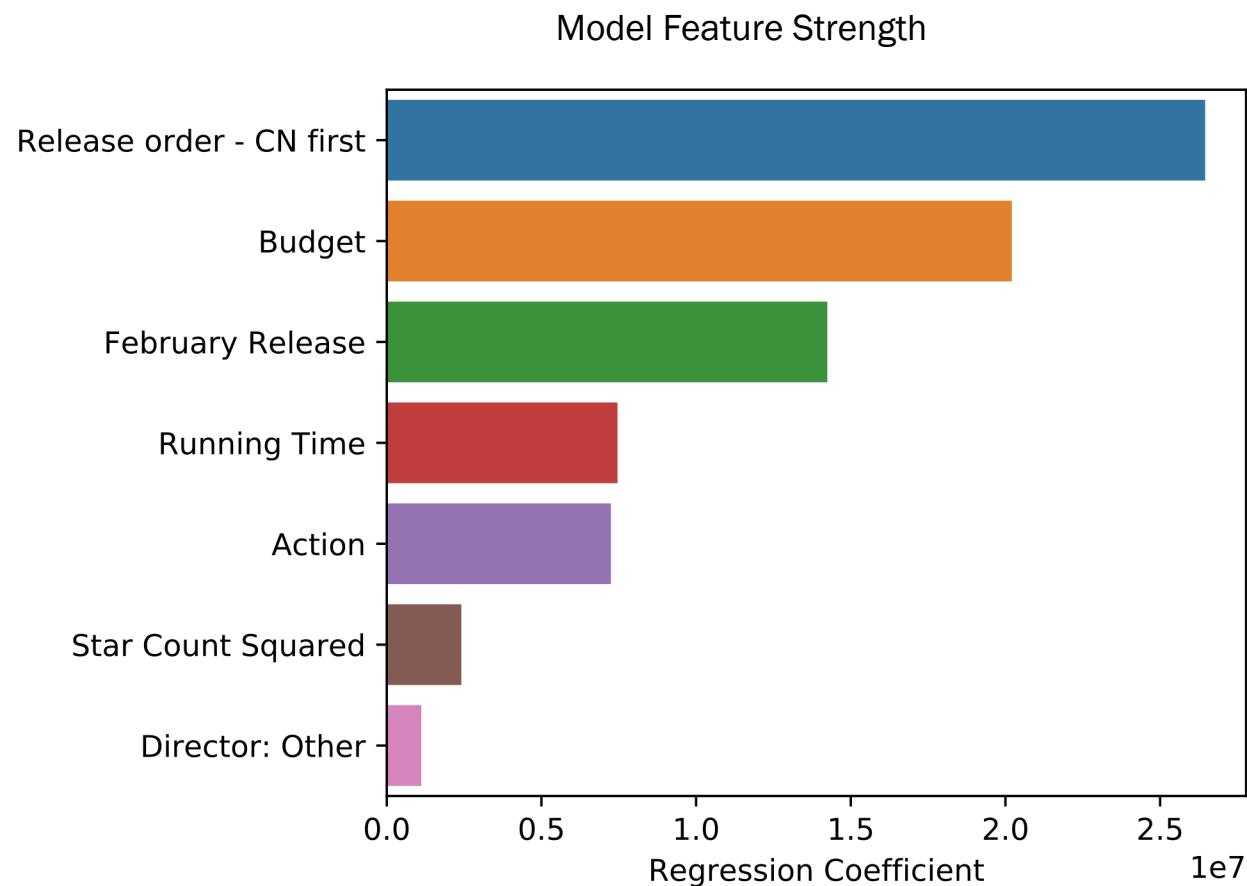


Revenue Predictions

- Results:
 - 30% of revenue variance can be explained by model
 - $R^2 = 0.306$
 - Predictions accurate within ~\$38 million
 - MAE = 38,373,286
- Analysis:
 - Overpredicted smaller films and underpredicted whales
 - Model's strength lies in interpretation of feature strength rather than revenue prediction



Strategic insights



KEY TAKEAWAYS:

1. **Timing matters**
 - *Release during Lunar New Year*
 - *Don't delay Chinese release*
2. **High budget yields high revenue**
3. **Action translates to foreign audiences easily**

CONCLUSION

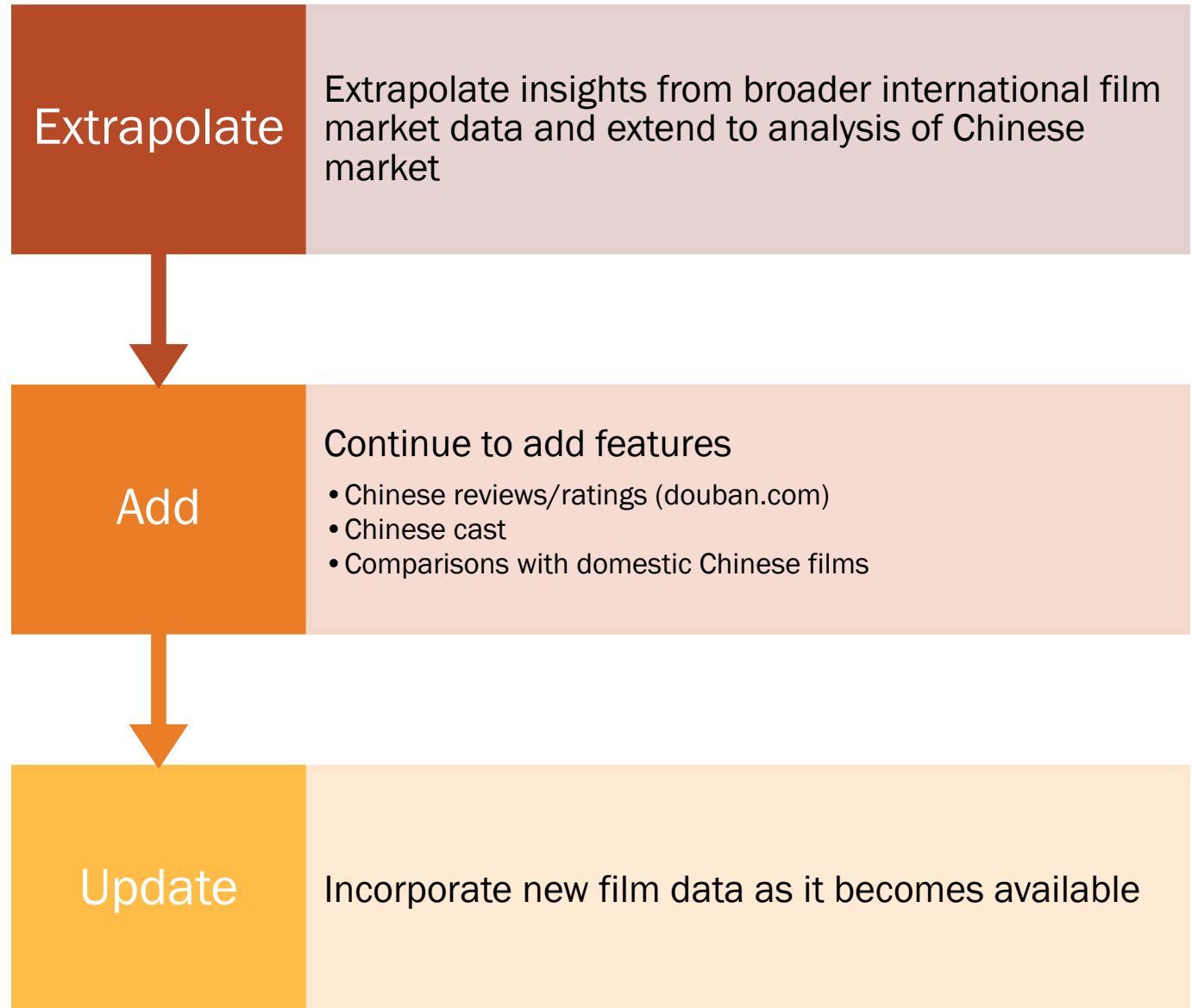


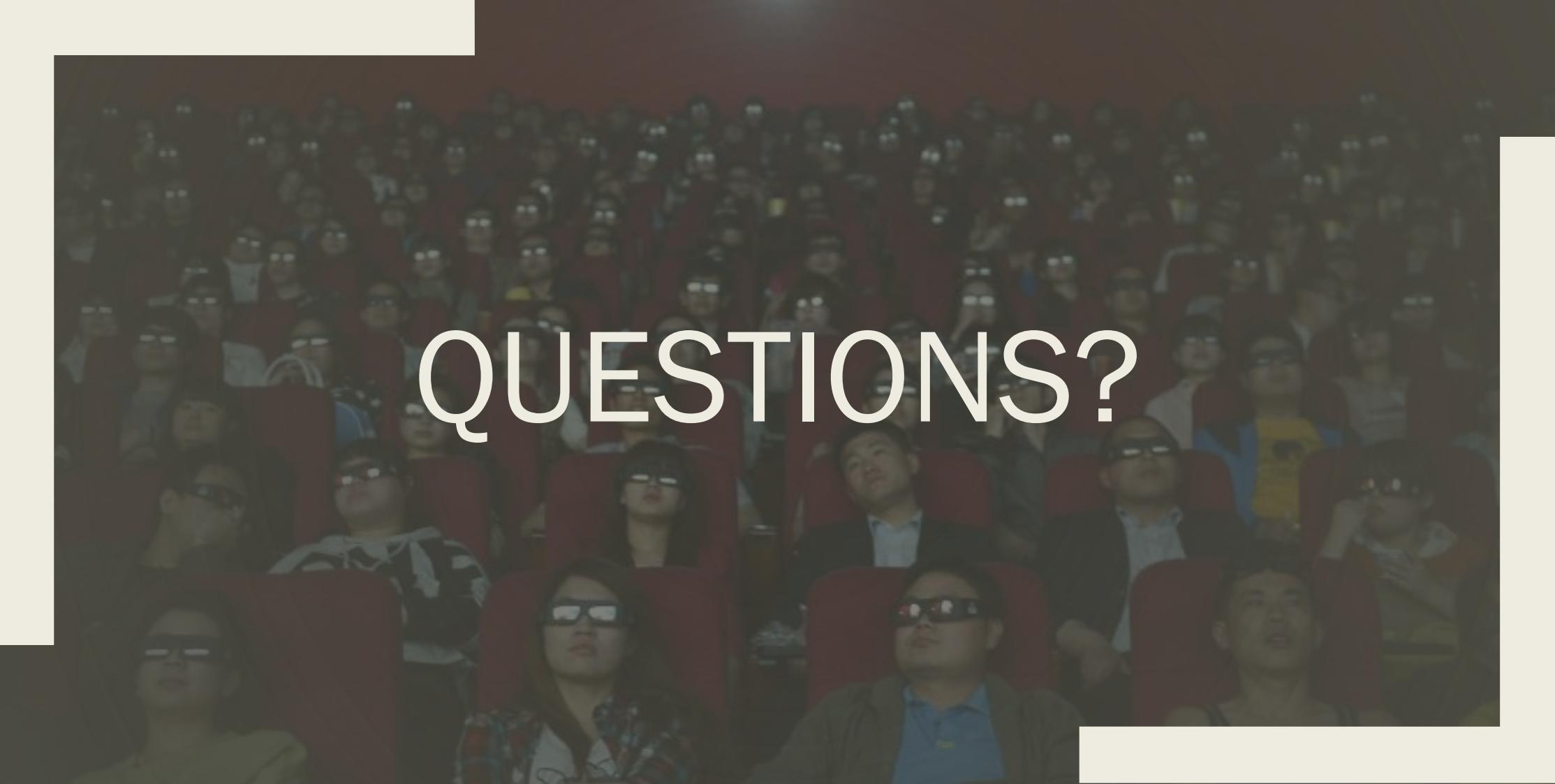
Conclusion

- Chinese film market requires specific strategic considerations, including
 - *Timing*
 - *Spend*
 - *Genre*
- Revenue is difficult to predict, but attempting to do so yields valuable insights for successfully marketing international films



Next steps





QUESTIONS?