



From Millionaire to Billionaire

Team: MonsterLearningSquared

Members: Hao Li, Louis Chu, Yiqin Chen, Tianai Guo

How to make a successful restaurant?

- ❖ To learn from **successful** ones
 - Highly rated
 - Popular (# review)
- ❖ To improve **relatively inexpensive** factors
 - Waiting time
 - Service

Do these improvements work?

Data Cleaning:

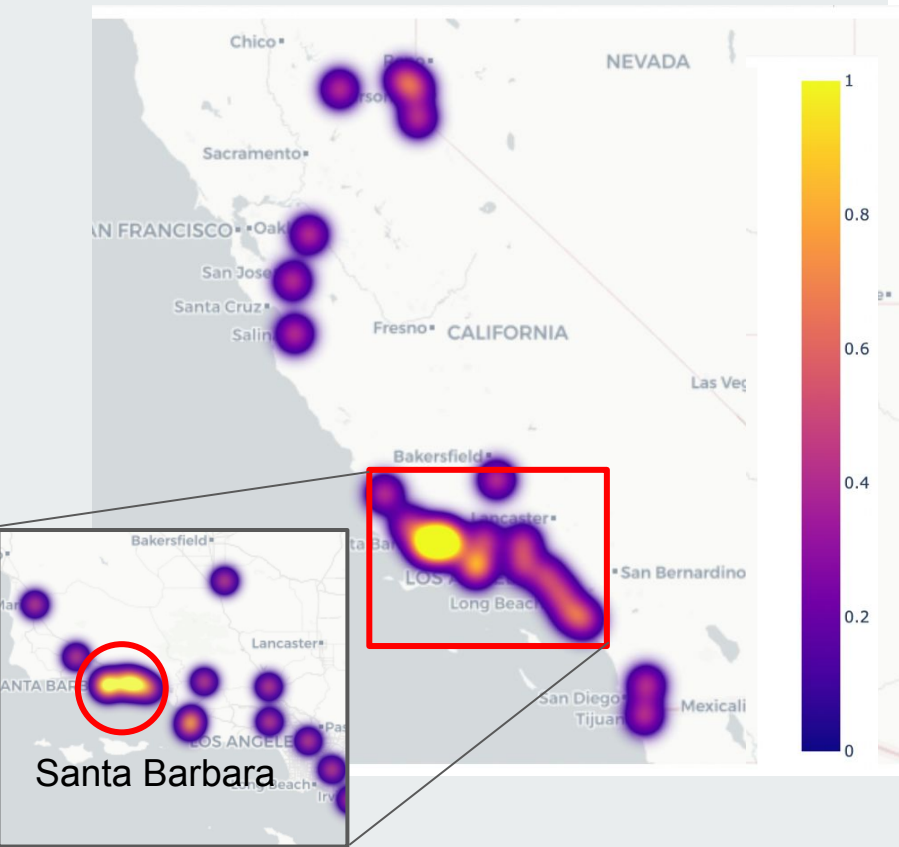
- Lemmatize Cities
- Extract Categories

Data Enrichment (on columns):

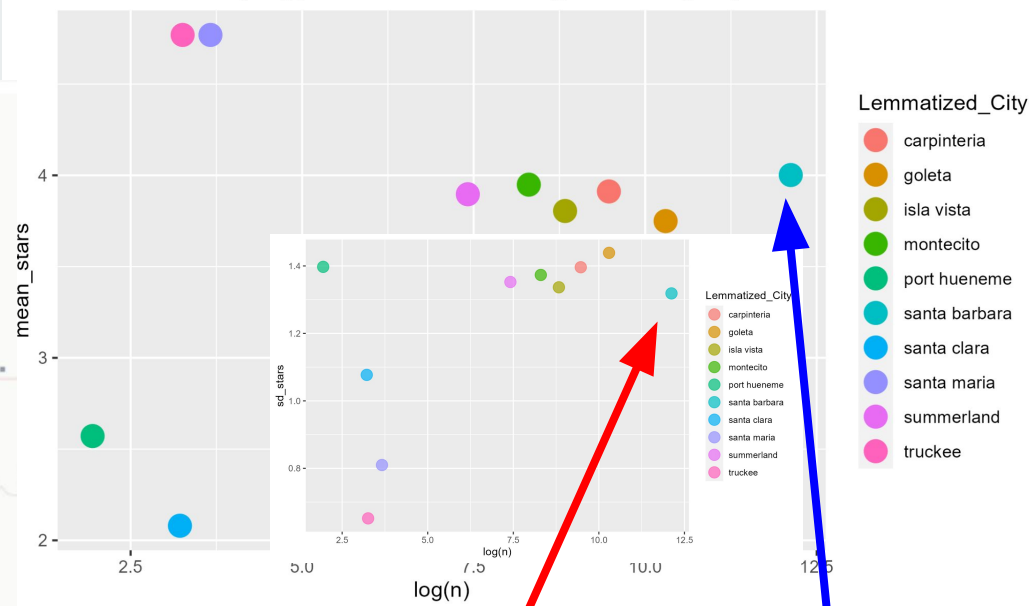
- Left-joined business and users on reviews
- Added encoder for the presence of top frequent words, sentiment score (from Vader)

Identify a cost-effective subgroup with sufficient reviews, reasonable star rating, impacted by waiting time and complaint handling

Recursively infer target subgroup by inference tests and correlations



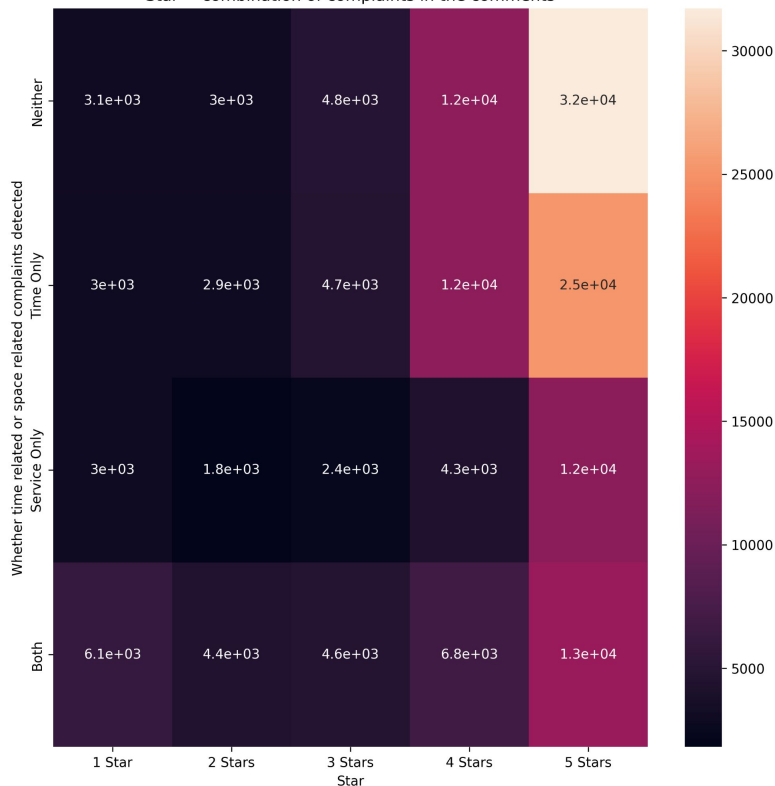
Mean of stars by log(number of comments), colored by city name



And... **risk averse** ... so Santa Barbara is chosen
More comments -> popularity -> more accurate estimators

Similar mean/ Lower s.d. In cities with >20 reviews -> may be more predictable
($MSE = Var + Bias^2$)
population variance is also lower

Star ~ combination of complaints in the comments

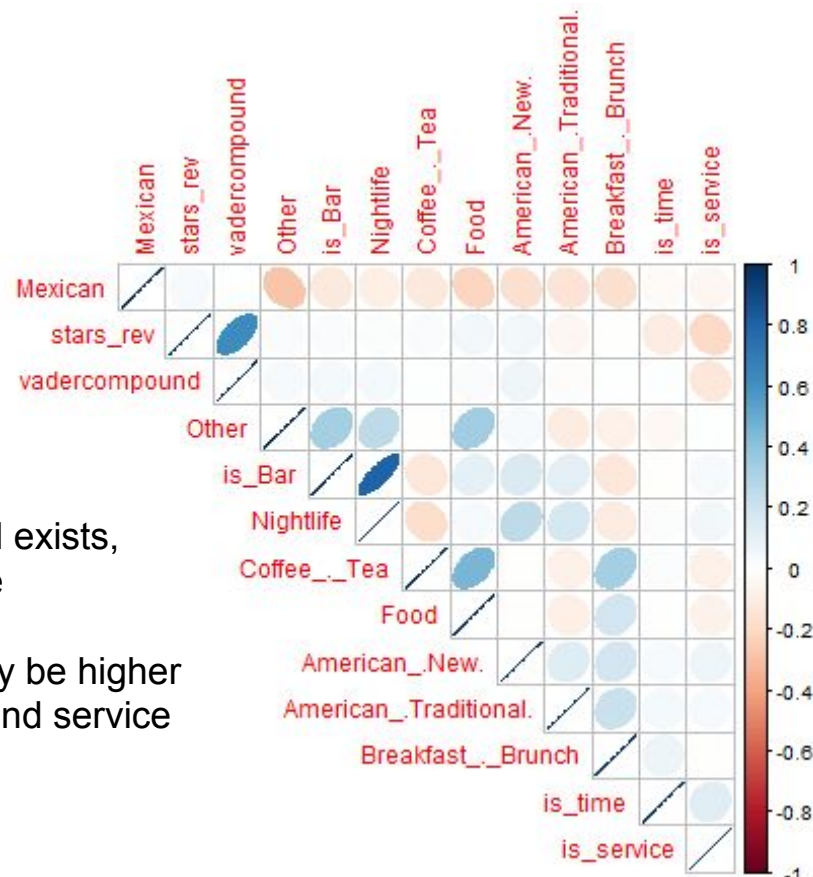


Outperform, or increase ratings by shortening waiting time and improving service?

X-squared 8365.7, df=12, **pval<2.2e-16**

What other aspects & keywords matters?

Correlation of potential predictors to stars given to restaurants in Santa Barbara, including features from frequency selection, encoding from presence of time/service complaints



Conclusion:

- Successful business model exists, i.e. Mesa Verde

- The rating may be higher if waiting time and service are better