



Presented by Raymond Lim

YELP RESTAURANTS BUSINESS REVIEWS

Business Analysis
Sentiment Analysis
Predict Restaurants Closure

PROBLEM STATEMENT

Singapore's F&B market is worth S\$8.3bn with expected 2.1% CAGR growth by 2023. Opportunities aside, F&B sector also operating in a competitive environment with the increase in numbers of restaurants.





WE OFFER

01. Restaurant Owners

Utilize user reviews data to help business owners understand why a business has good/bad ratings. One could learn, exploit and improve the weaknesses of a restaurant and thus get a business advantage.

02. Investors/Banks

Develop a classification model based on restaurants attributes to predict if a business will be closed. This enables investors to access the risk before investing in a restaurant.

DATA COLLECTED

Data Collected from Yelp business API, focusing on restaurants in Massachusetts and Oregon, United States.



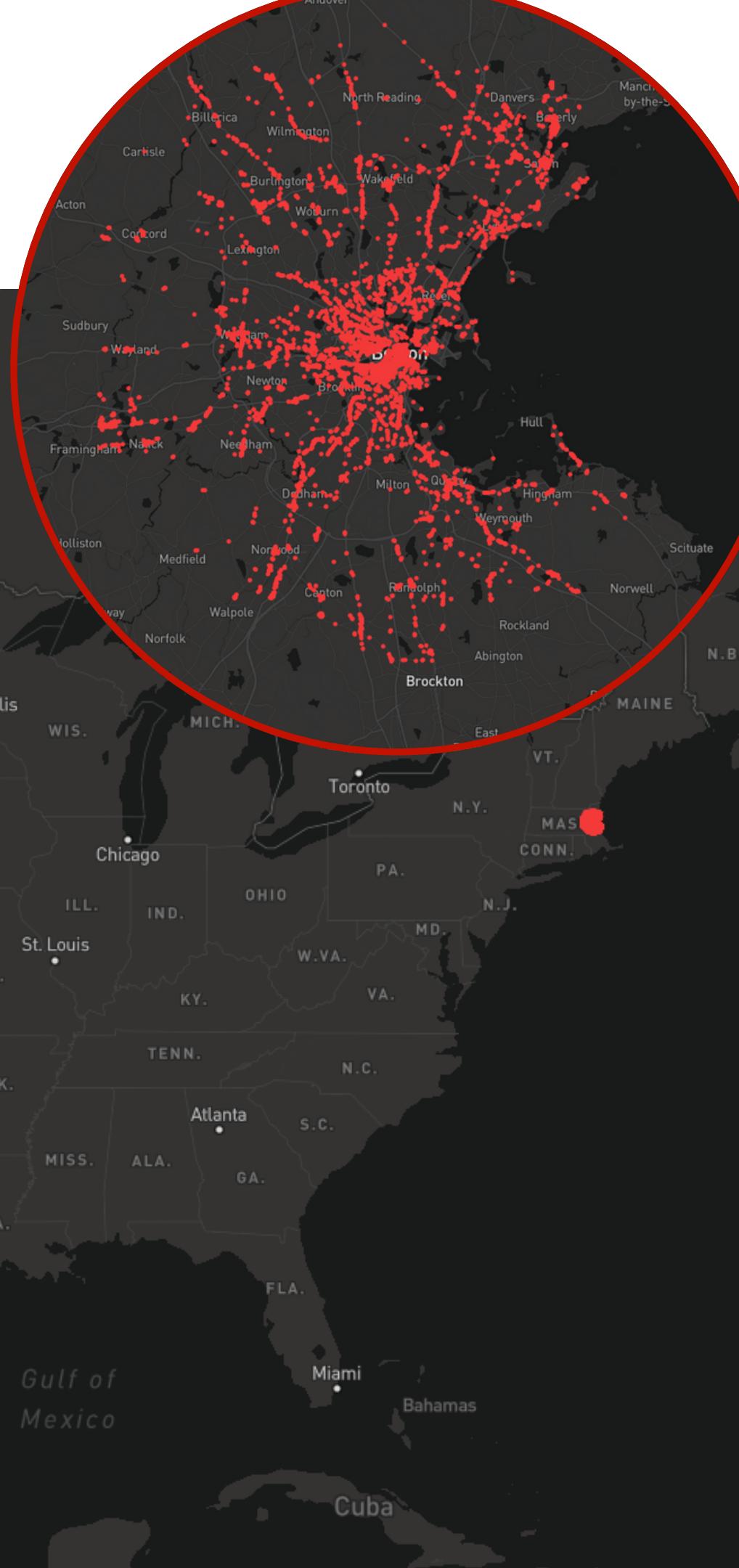
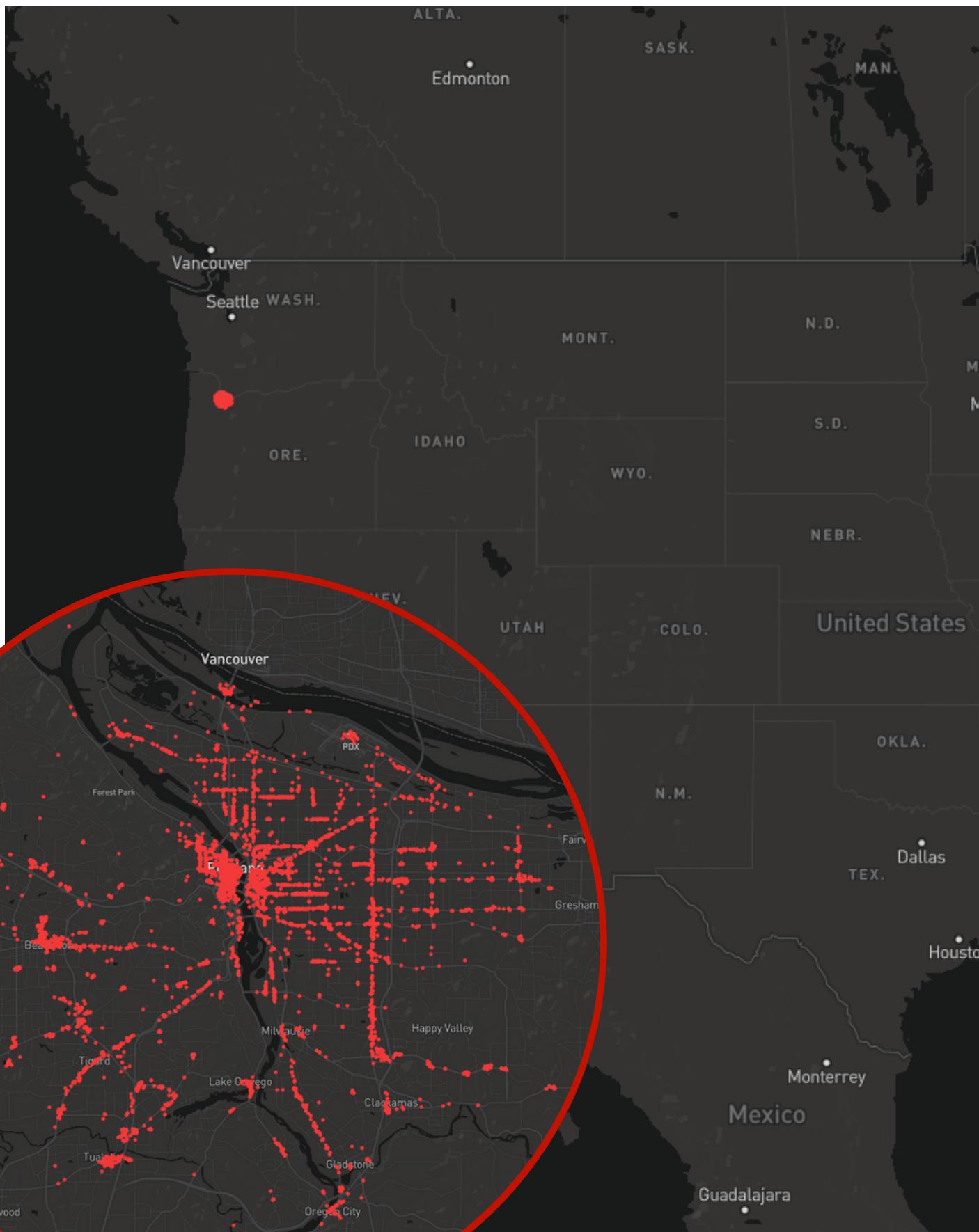
Initial: 16,256 Restaurants



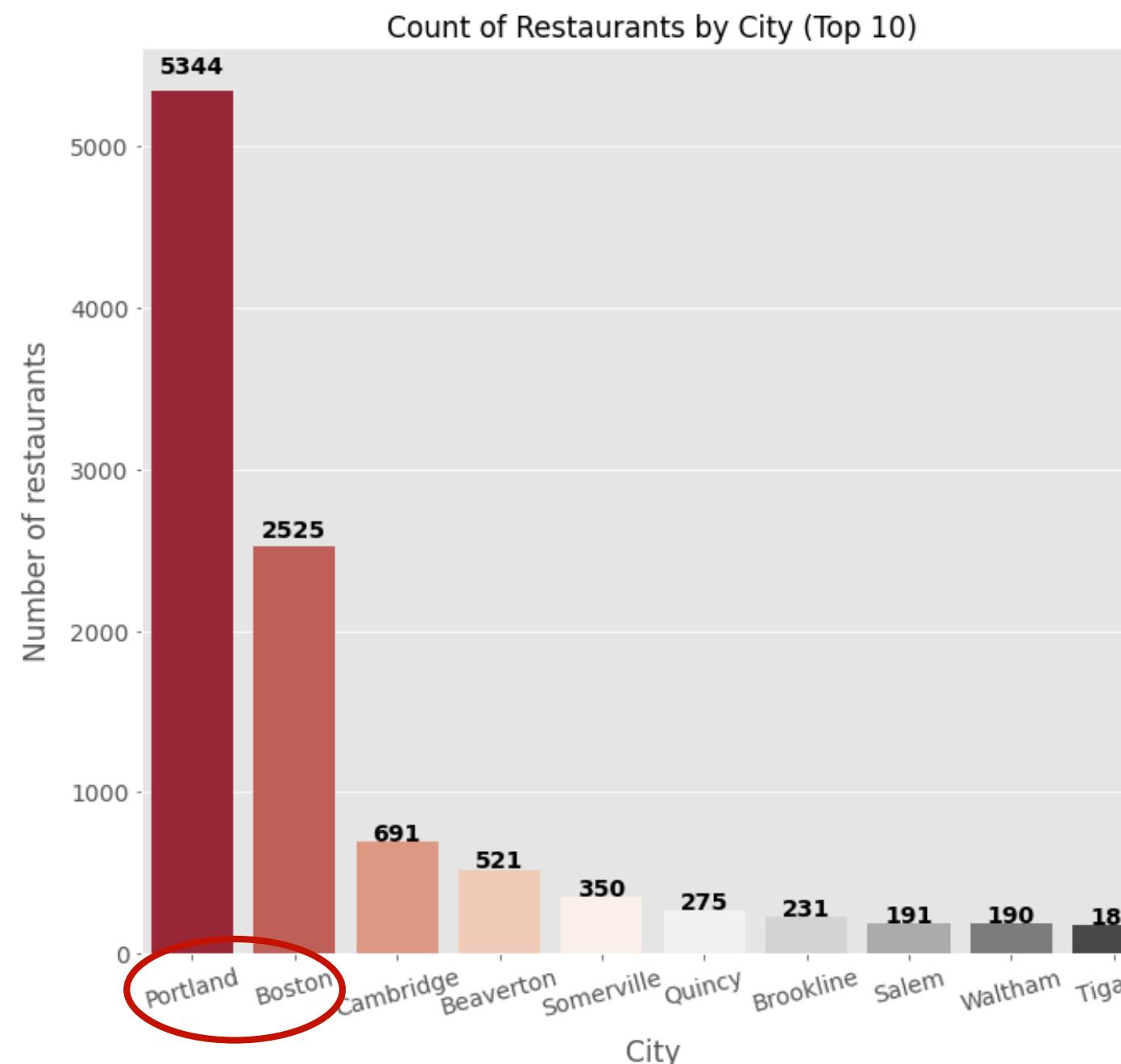
Initial: 2,282,296 Reviews



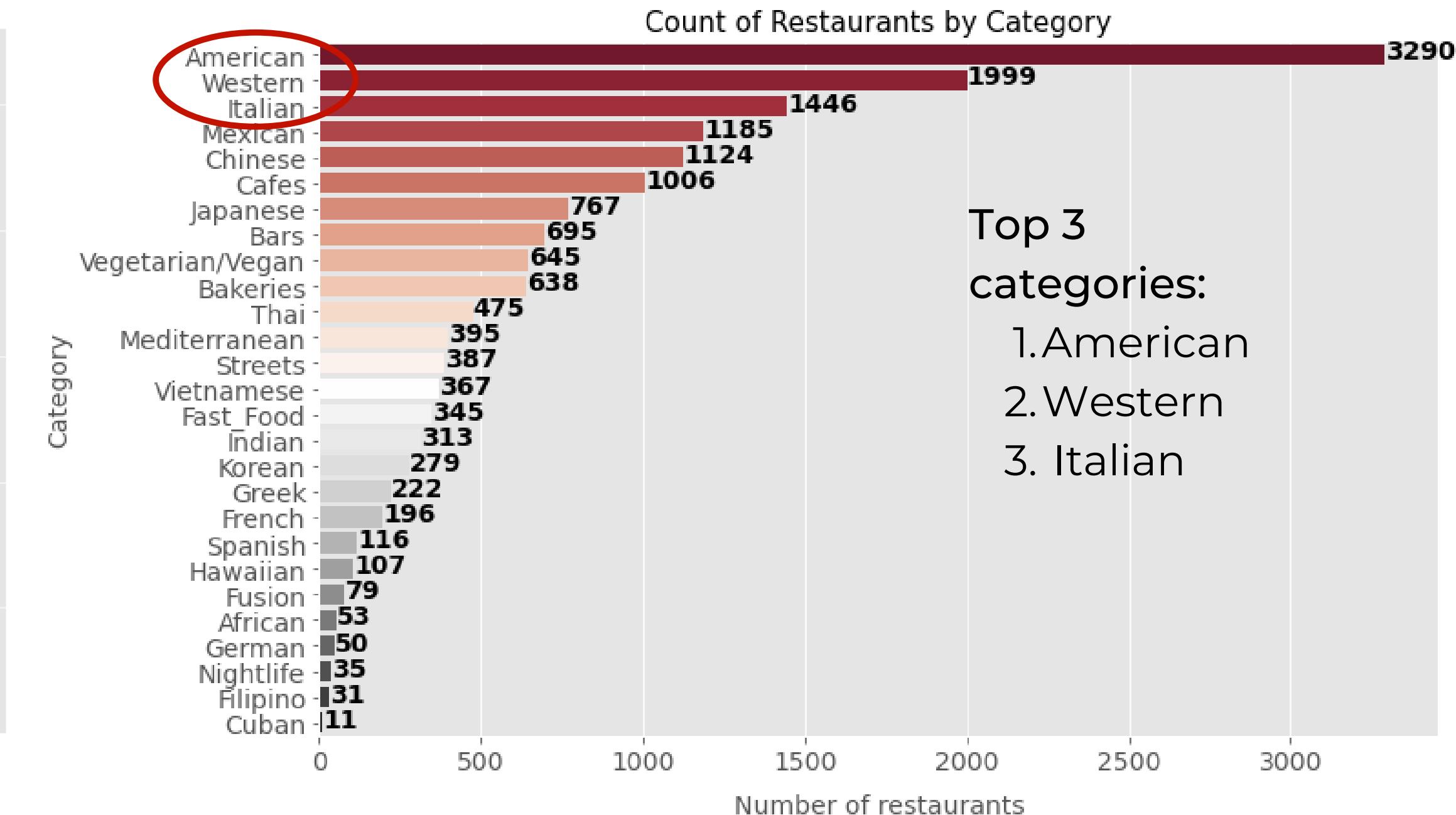
Initial: 636,249 Users



RESTAURANTS



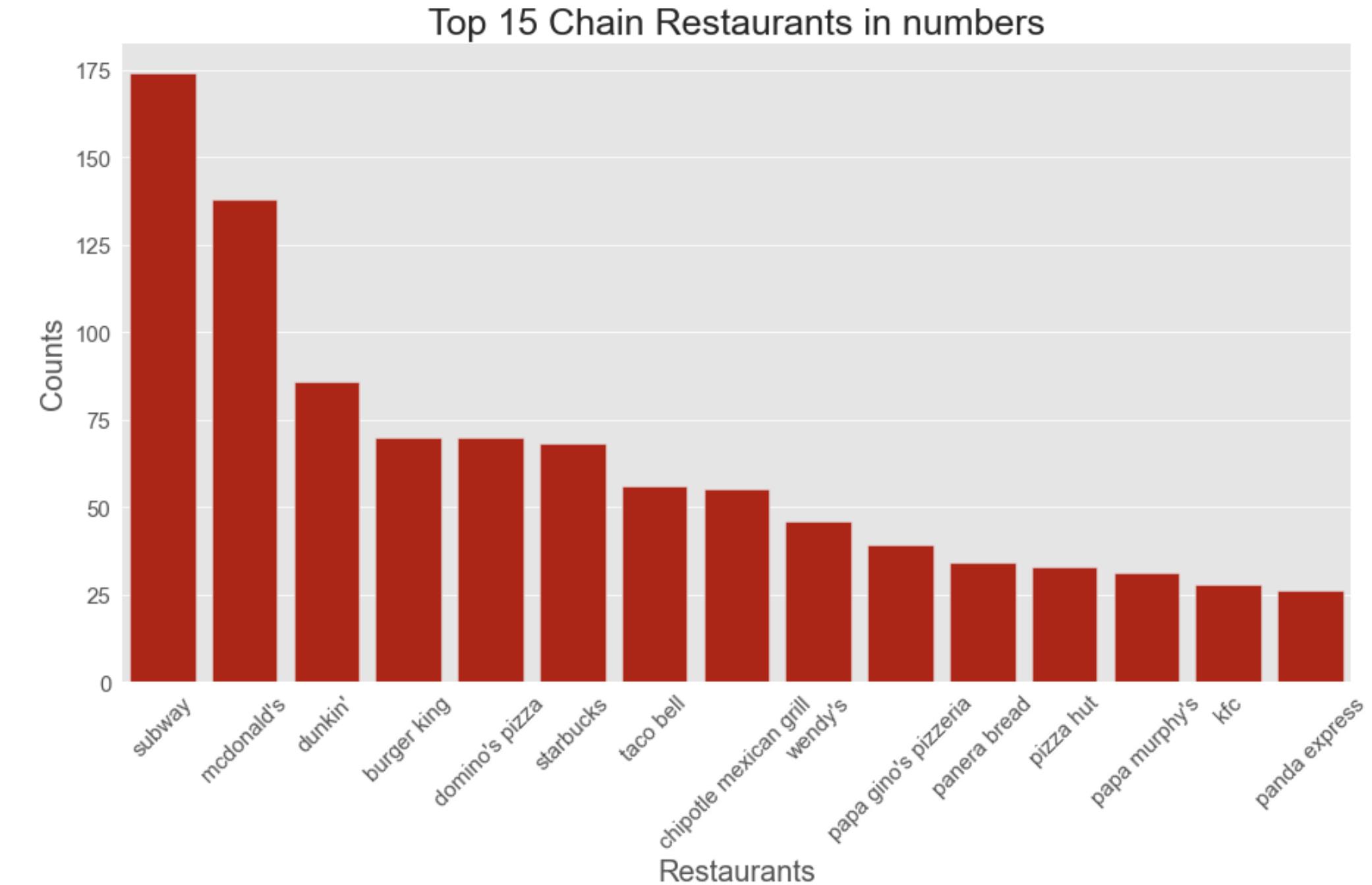
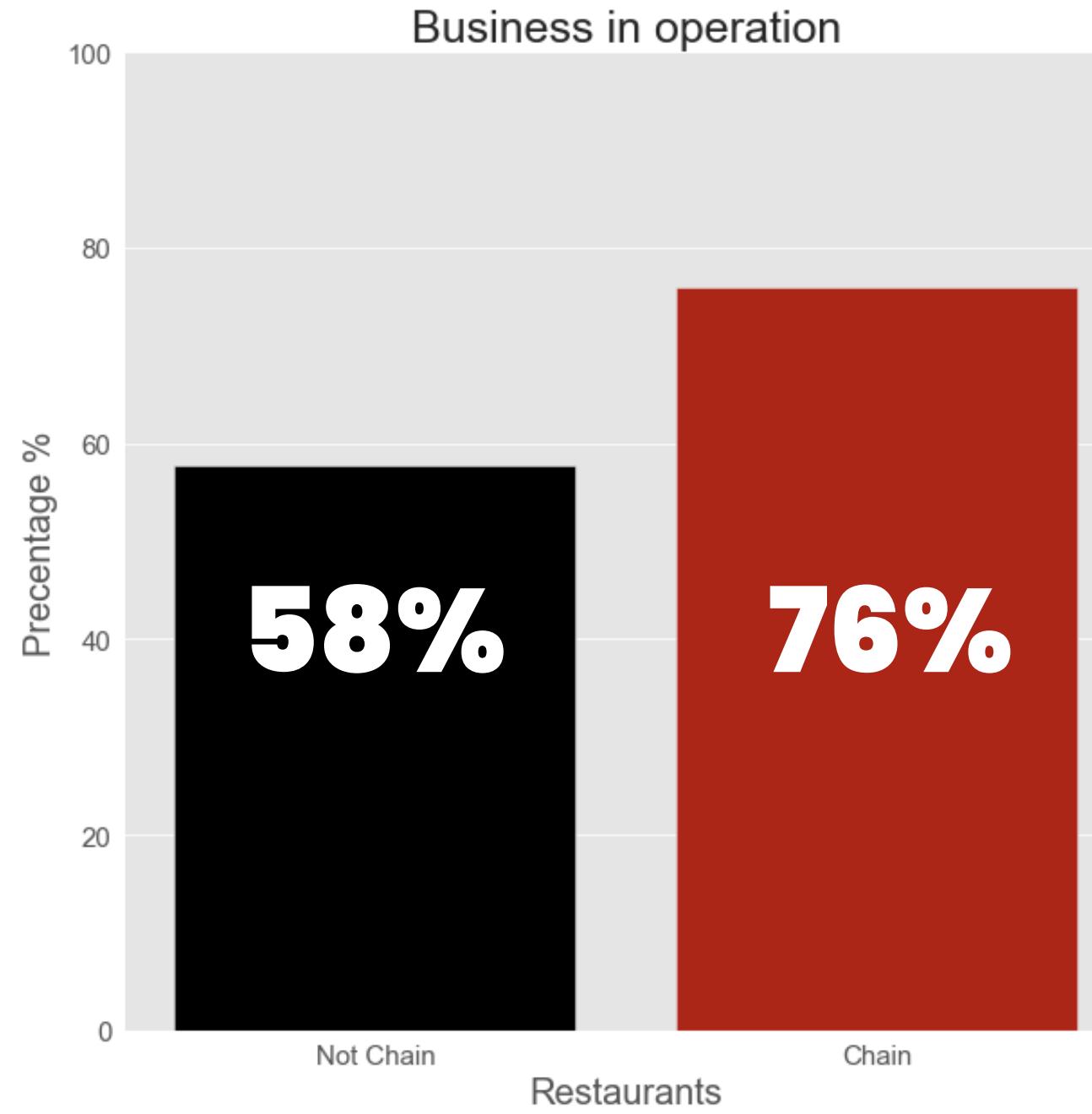
Portland & Boston are the cities with the higher number of restaurants count.



Restaurants are categorised into 27 categories.

CHAIN RESTAURANTS

Calculate the restaurants that has at least **5** identical entries.

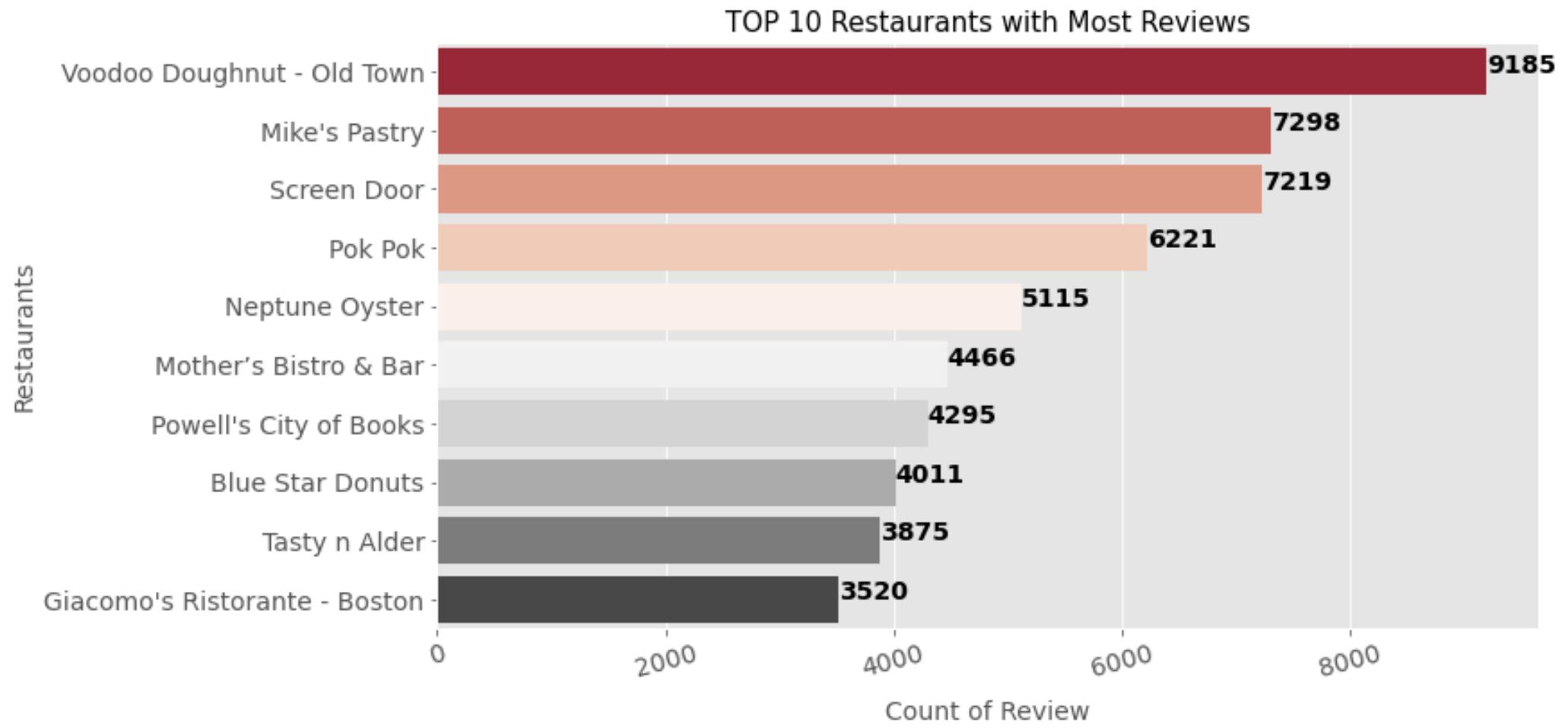


76% of Chain restaurants remain in operation, compared to non-chain restaurants at **58%** only.
Non-chain restaurants have a higher risk of failure compared to chain restaurants.

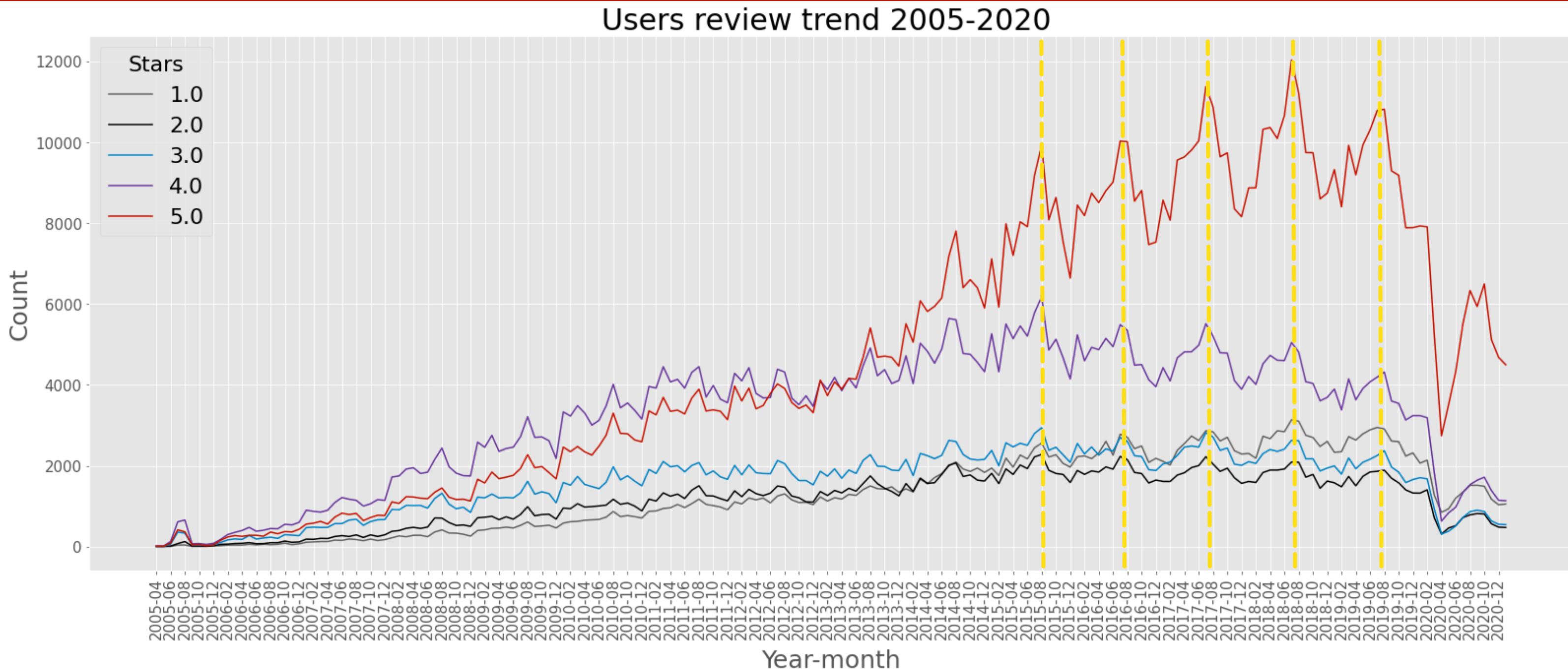
MOST REVIEWED

Top 3 most reviewed restaurants:

1. Voodoo Doughnut, Portland.
2. Mike's Pastry, Boston.
3. Screen Door, Portland.



REVIEWS TREND OVER YEARS



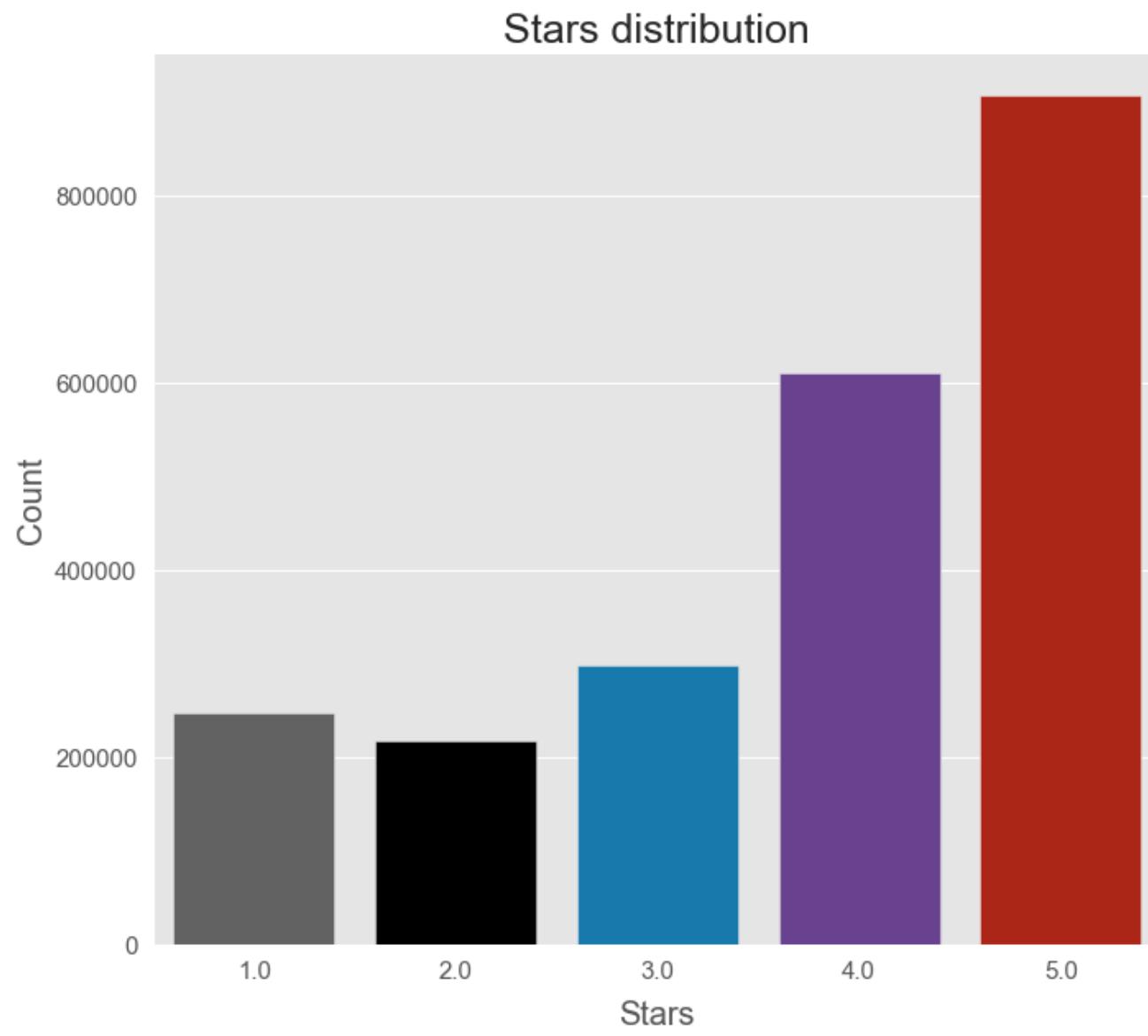
The number of reviews has increased gradually throughout the years.

Each year, the highest number of comments count recorded in August, which is the travel period for summer break.

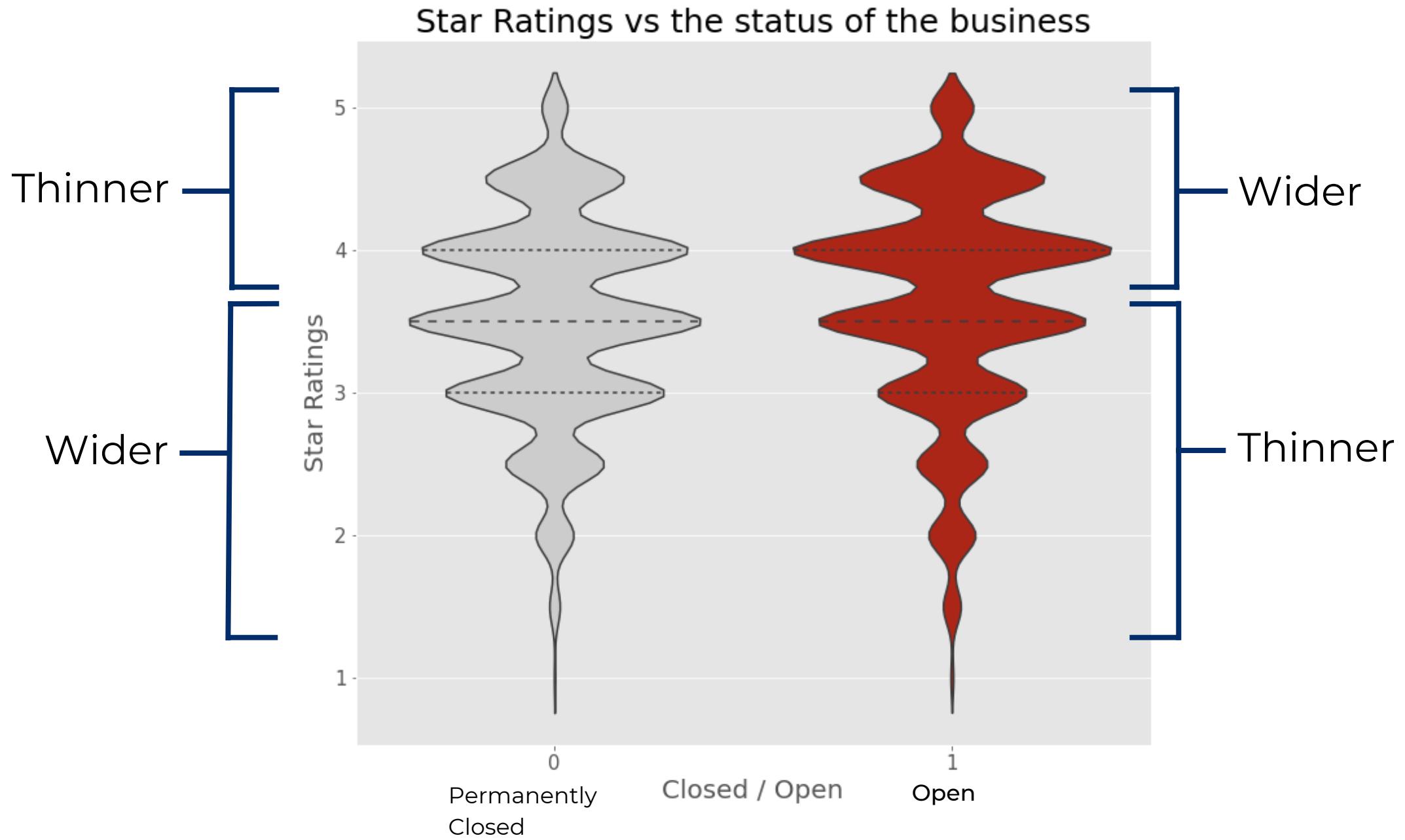
The number of reviews decreases drastically in end of 2019, 2020 due to COVID-19 pandemic lock down.

RATING DISTRIBUTIONS

For each restaurant review, each user left a rating on a scale of 1–5 that indicated their overall experience visiting a restaurant.



The distribution of review ratings was skewed to the left with more users giving higher ratings than lower ratings on average.

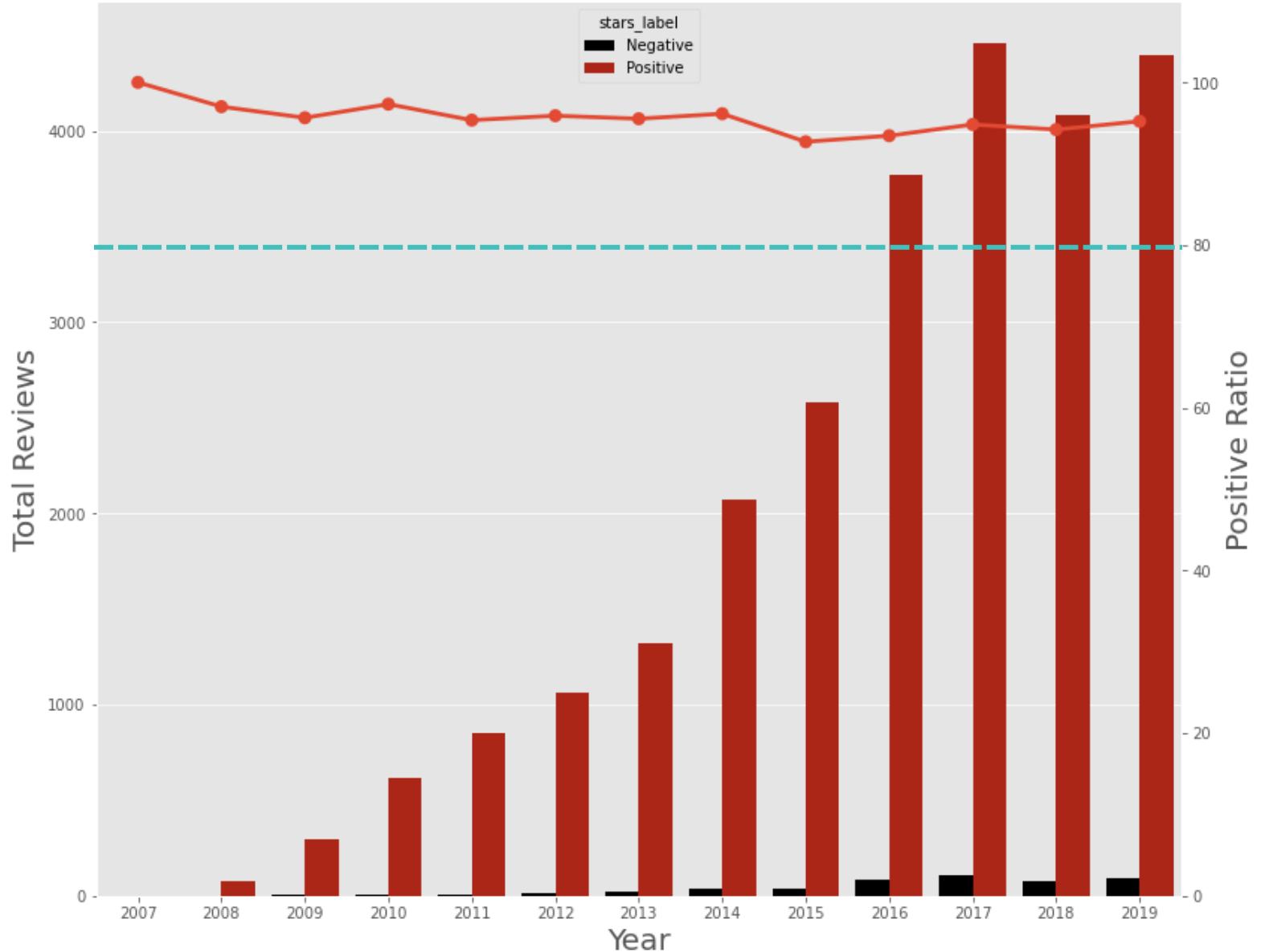


The violin plot suggests Permanently Closed restaurants have lower count of ratings above 4 compared to restaurants remains open. Those restaurants recorded higher numbers of ratings less than 3.5.

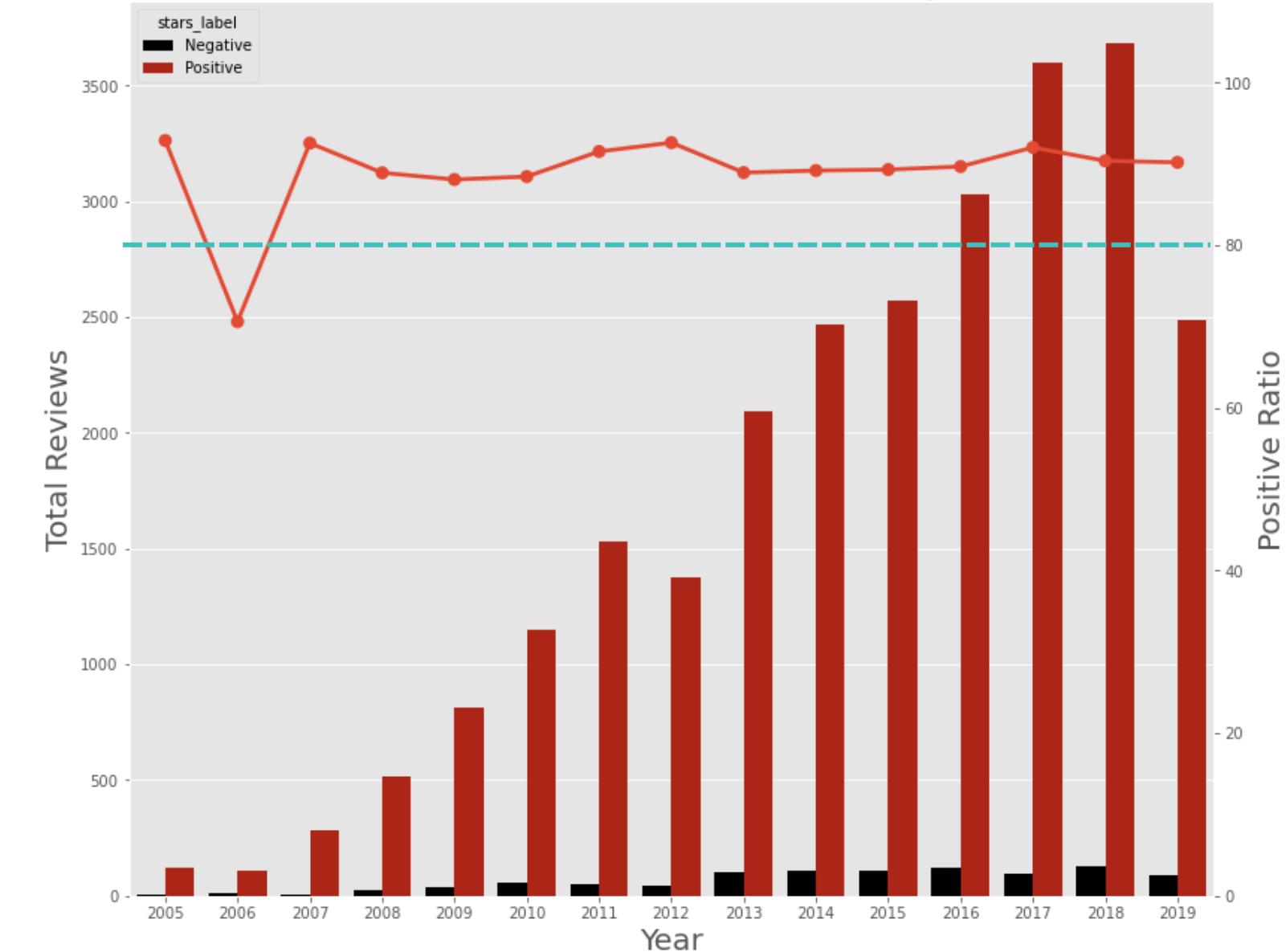
TOP BUSINESSES



Review trend of Screen Door



Review trend of Mike's Pastry

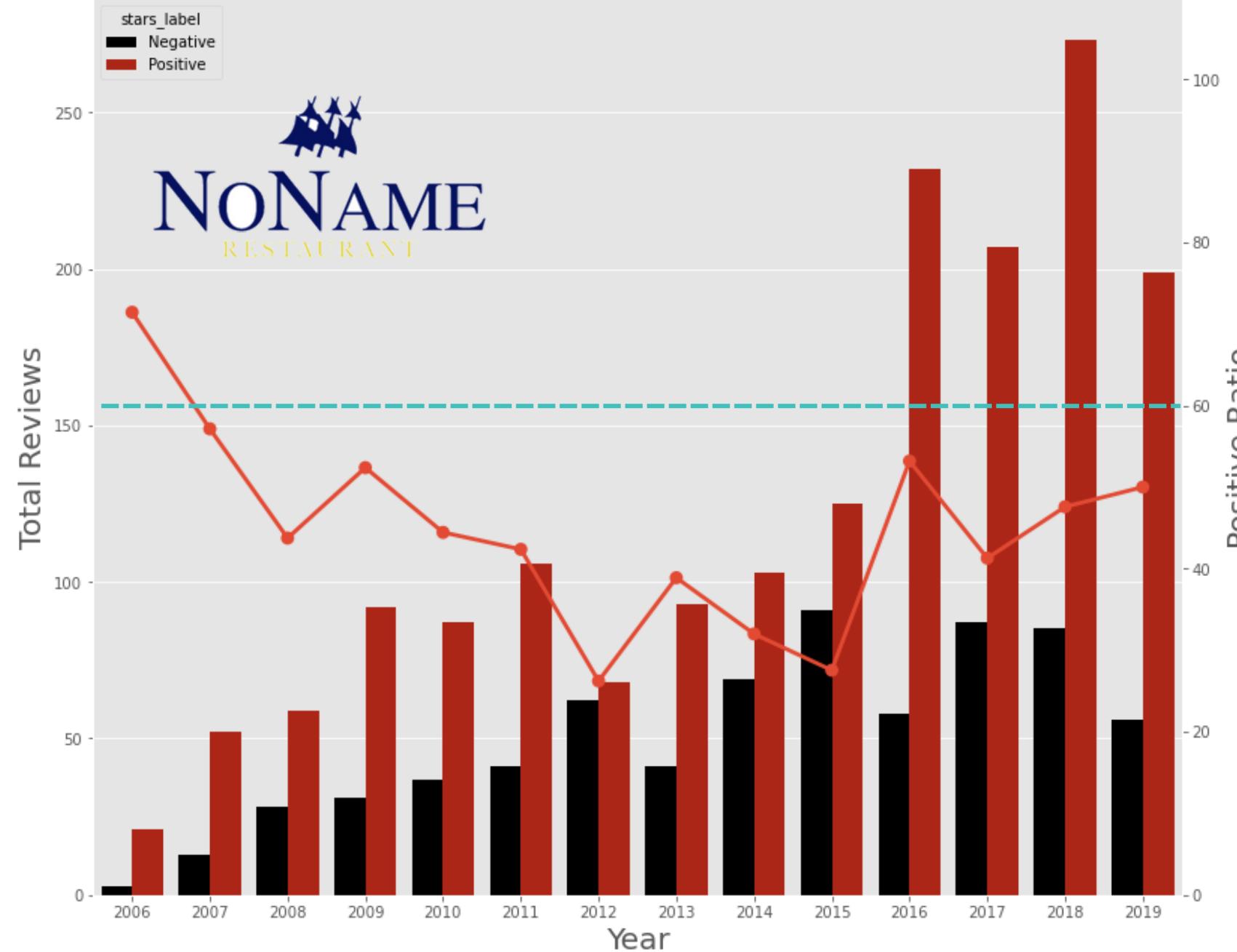


By looking at the top positively reviewed restaurants, we observed exponential growth of positive reviews over the years. At the same time, they are able to keep negative reviews low number.

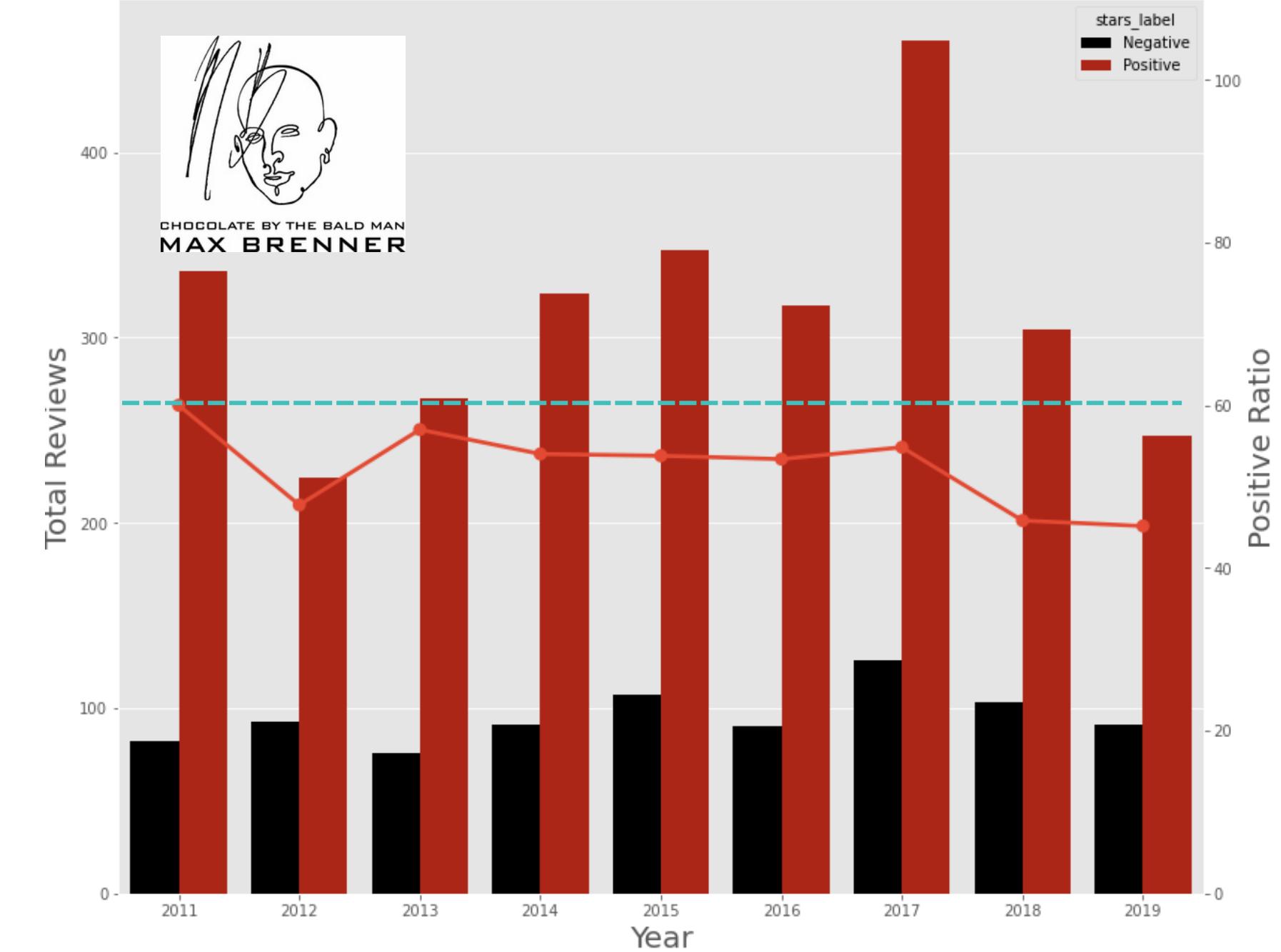
The positive reviews ratio is maintained at **above 80%** for top businesses.

LESS PERFORMING BUSINESSES

Review trend of No Name Restaurant



Review trend of Max Brenner - Boston



For the top negatively reviewed restaurants, the number of positive reviews increases with the increase in users. However, the number of negative reviews increases at the same time.

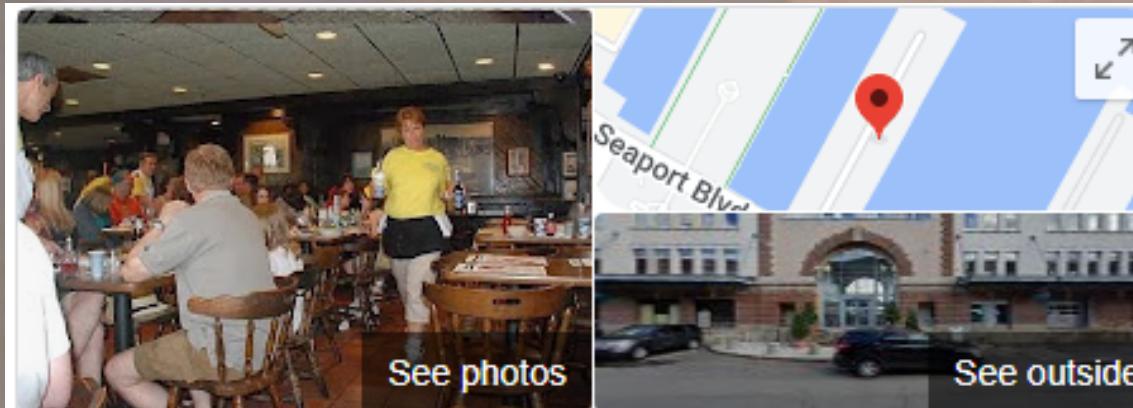
The positive reviews ratio is maintained at a **low 60%** for less-performing restaurants.

UNFORTUNATELY.. PERMANENTLY CLOSED..

Boston's landmark "No Name" restaurant closes after 102 years

BY RACHEL LAYNE
DECEMBER 31, 2019 / 11:46 AM / MONEYWATCH

f t p



Permanently closed

No Name Restaurant

4.0 ★★★★☆ 1,983 Google reviews

\$\$ · Seafood restaurant

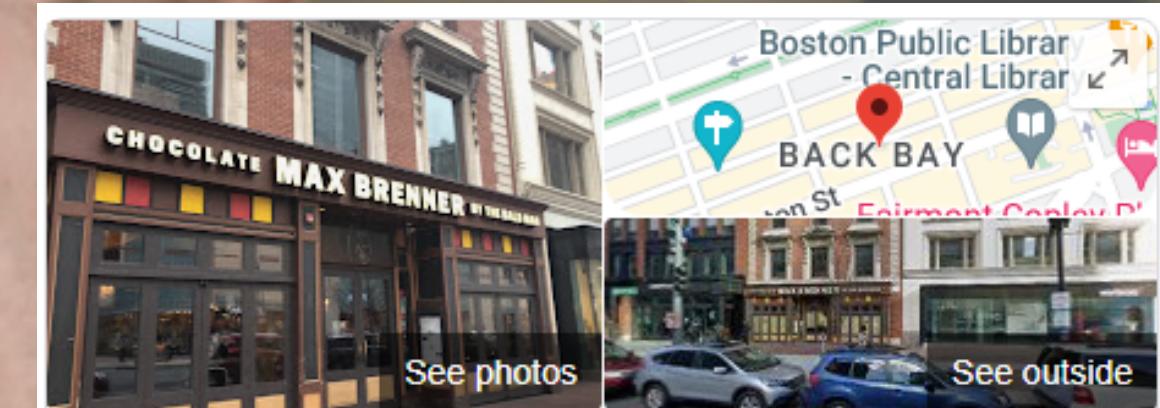
Address: 17 Fish Pier St E, Boston, MA 02210, United States

Menu: nonamerestaurant.com

BOSTON
Max Brenner in Boston's Back Bay Has Closed Permanently

By Marc Hurwitz • Published December 21, 2020 • Updated on December 21, 2020 at 3:03 pm

f t e



Permanently closed

Max Brenner

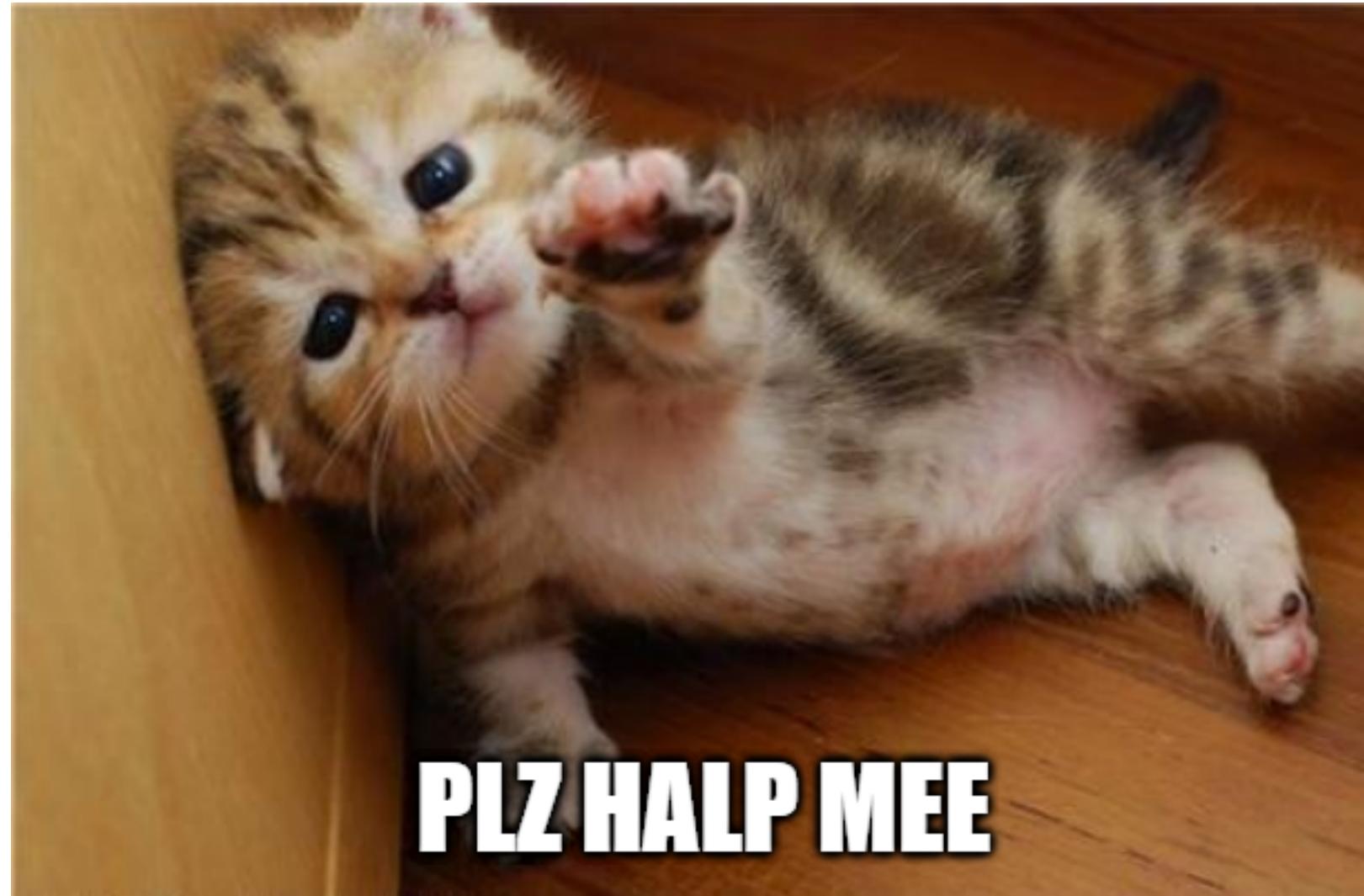
4.1 ★★★★☆ 1,601 Google reviews

\$\$ · Chocolate shop

Address: 745 Boylston St, Boston, MA 02116, United States

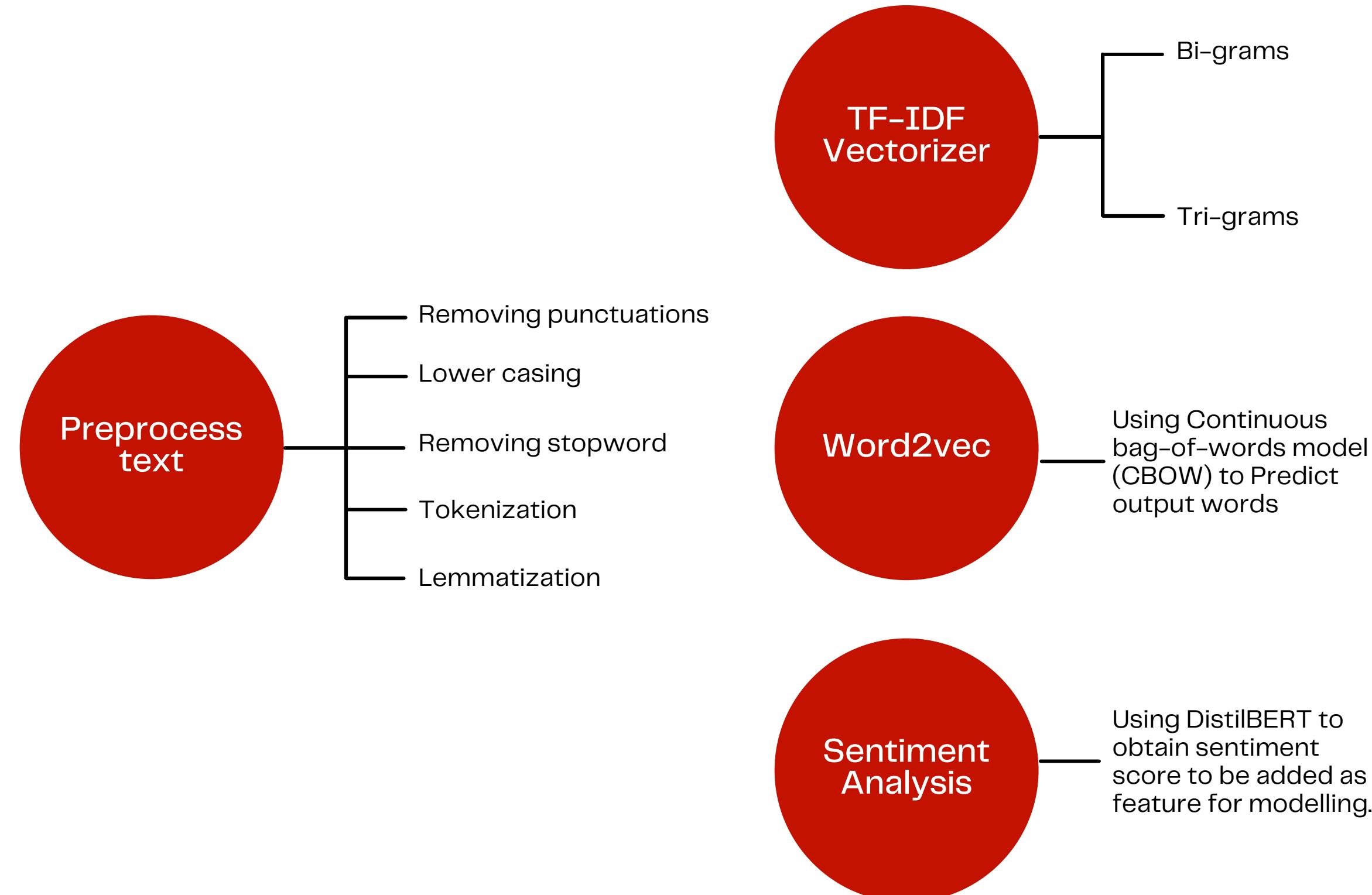
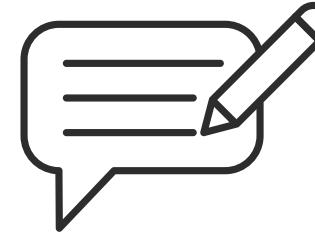
Menu: maxbrenner.com

SAVE THE BUSINESS!



PLZ HALP MEE

Customer Reviews analysis



1 star Reviews Trigram

Created a tree-map based on the 1 star reviews .



Most customers left 1-star reviews due to unsatisfied customer service.

Some customers were unhappy with the long waiting time, between 10 mins - 15 mins.

The customers feel wasted time and money if a situation is not handled properly in time by the staff.

They have also commented that they would not visit the place again, which would leave a negative impression on other users who saw the reviews.

5 stars Reviews Trigram

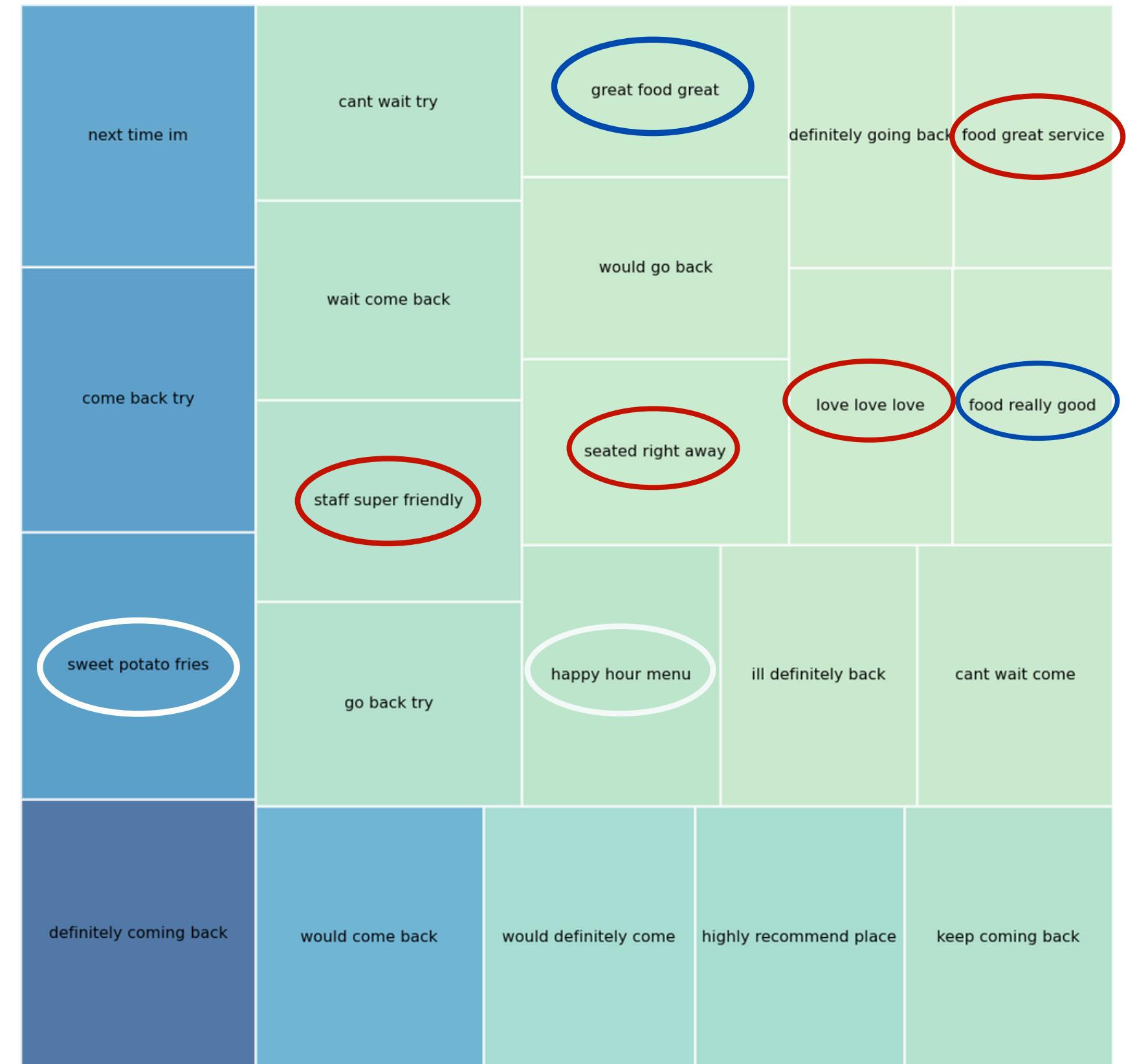
Created a tree-map based on the 5 star reviews .

Most customers left 5-star reviews due to satisfied customer service and food.

A small gesture like getting customers seated right away and friendly staff will make the customers feel loved.

Customers would love to look at the happy hour menu. Complement sides such as sweet potato fries is a plus point.

Happy customer commented they would definitely visit again, which would leave a positive impression on other users who saw the reviews.



Word2Vec - Why people love it

Predict output word from 5 stars reviews, sorted by probability

'service'



	word	probability
0	impeccable	0.419132
1	prompt	0.386828
2	spotty	0.077133
3	customer	0.075028
4	attentive	0.030484
5	friendly	0.002074

'food'



	word	probability
0	coma	0.303693
1	court	0.044934
2	trucks	0.030591
3	indian	0.030038
4	comfort	0.023506
5	allergies	0.022847

'staff'



	word	probability
0	friendly	0.729812
1	members	0.220202
2	member	0.027938
3	superfriendly	0.012504
4	friendliest	0.006315
5	courteous	0.001876
6	attentive	0.001018
7	knowledgeable	0.000123

'price'



	word	probability
0	tag	0.849519
1	reasonable	0.150423
2	moderate	0.000023
3	lower	0.000018
4	steep	0.000005
5	unbeatable	0.000005

Word2Vec - Why people dislike it

Predict output word from 1 star reviews, sorted by probability

'service'



	word	probability
0	customer	0.301797
1	poor	0.229803
2	atrocious	0.110427
3	slow	0.109642
4	deplorable	0.060949
5	abysmal	0.033704

'food'



	word	probability
0	poisoning	0.538474
1	court	0.059331
2	quality	0.019930
3	mediocre	0.014952
4	allergies	0.009028
5	truck	0.008439

'staff'



	word	probability
0	member	8.957837e-01
1	members	1.042160e-01
2	train	1.088223e-07
3	wait	5.236794e-08
4	friendly	5.085982e-08
5	trained	4.501265e-08
6	untrained	8.361844e-09
7	hire	6.647711e-09
8	incompetent	3.034652e-09
9	unfriendly	2.832359e-09

'price'



	word	probability
0	tag	0.906590
1	higher	0.082279
2	exorbitant	0.002630
3	outrageous	0.001566
4	reasonable	0.001303
5	steep	0.001096

Modeling

Use supervised ML model to predict if the restaurant is closed or open.

Metrics to be used to evaluate each models are:

- Accuracy
- F1 score
- Confusion Matrix
- ROC Curve
- Precision-Recall Curve

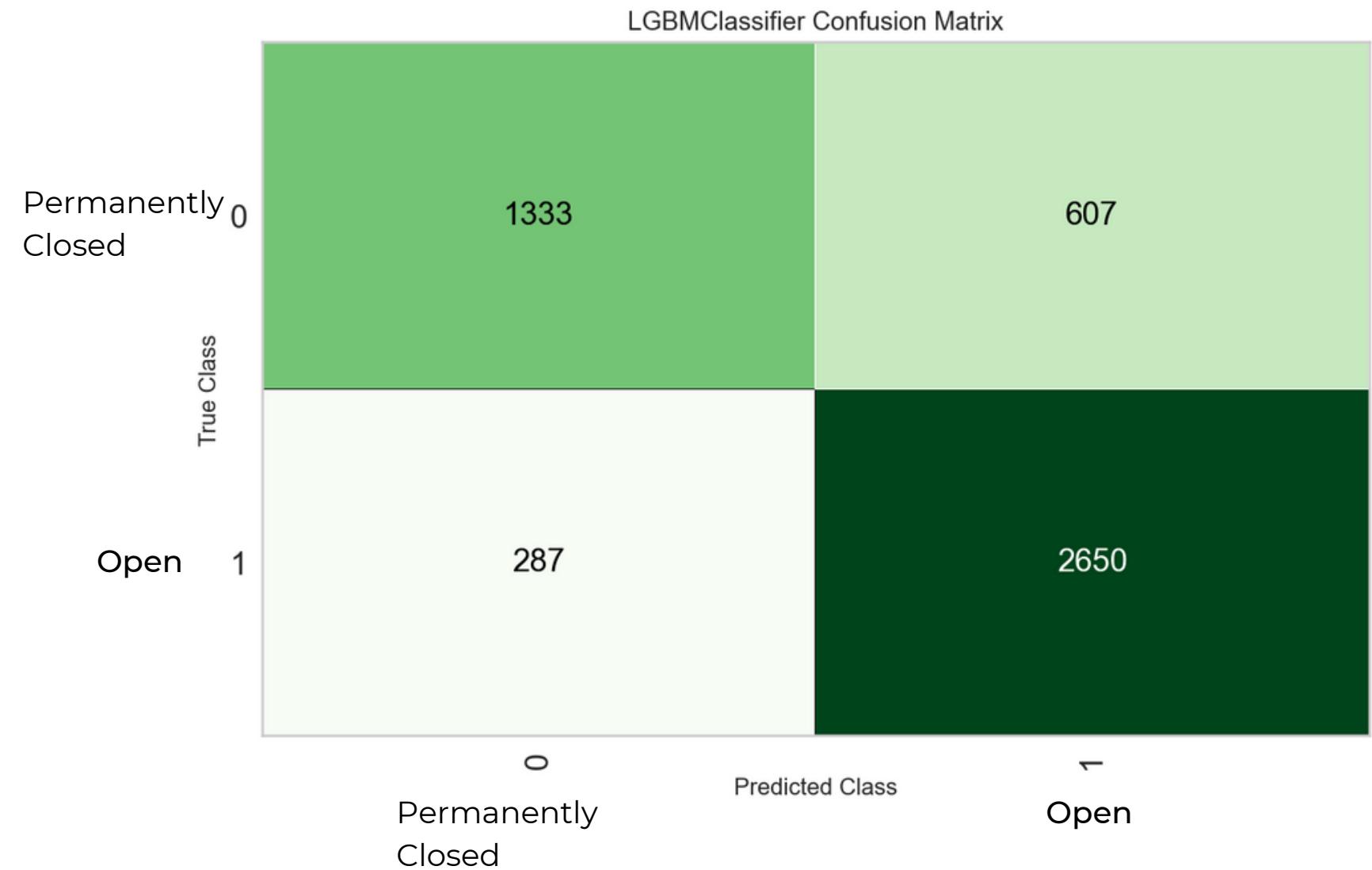
Model	Accuracy	F1
Light Gradient Boosting Machine	0.8172	0.8544
Random Forest Classifier	0.8140	0.8527
Gradient Boosting Classifier	0.8099	0.8501
Extra Trees Classifier	0.8018	0.8419
Logistic Regression	0.7996	0.8395

Light GBM model performed the best with 0.8172 accuracy score and 0.8544 F1 score.

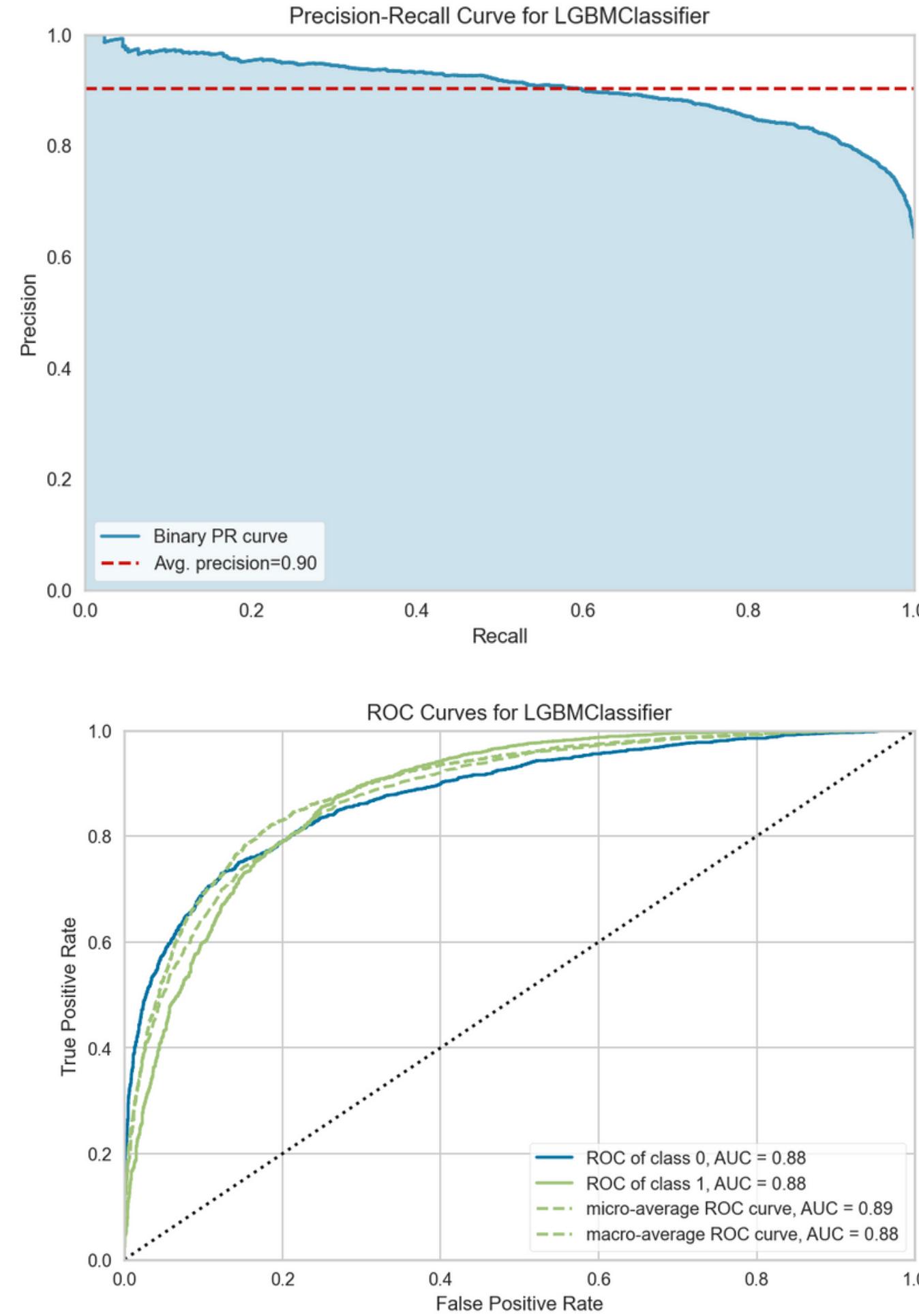
Confusion Martix

The precision of open restaurants is 90%. Only 10% remain as false positive. A bank that provides a loan based on this decision gives 10% fault rate, compared to 39% (Closed rate in the dataset)

The precision of the model is poor in case of closed restaurants. Out of the model prediction closed, Only 68% of the restaurants had truly closed down. It might mistakenly predict the restaurant remains open as closed.



	Open Restaurants	Closed Restaurants
Accuracy	0.8166	0.8166
Precision	0.901	0.687
Recall	0.8136	0.822

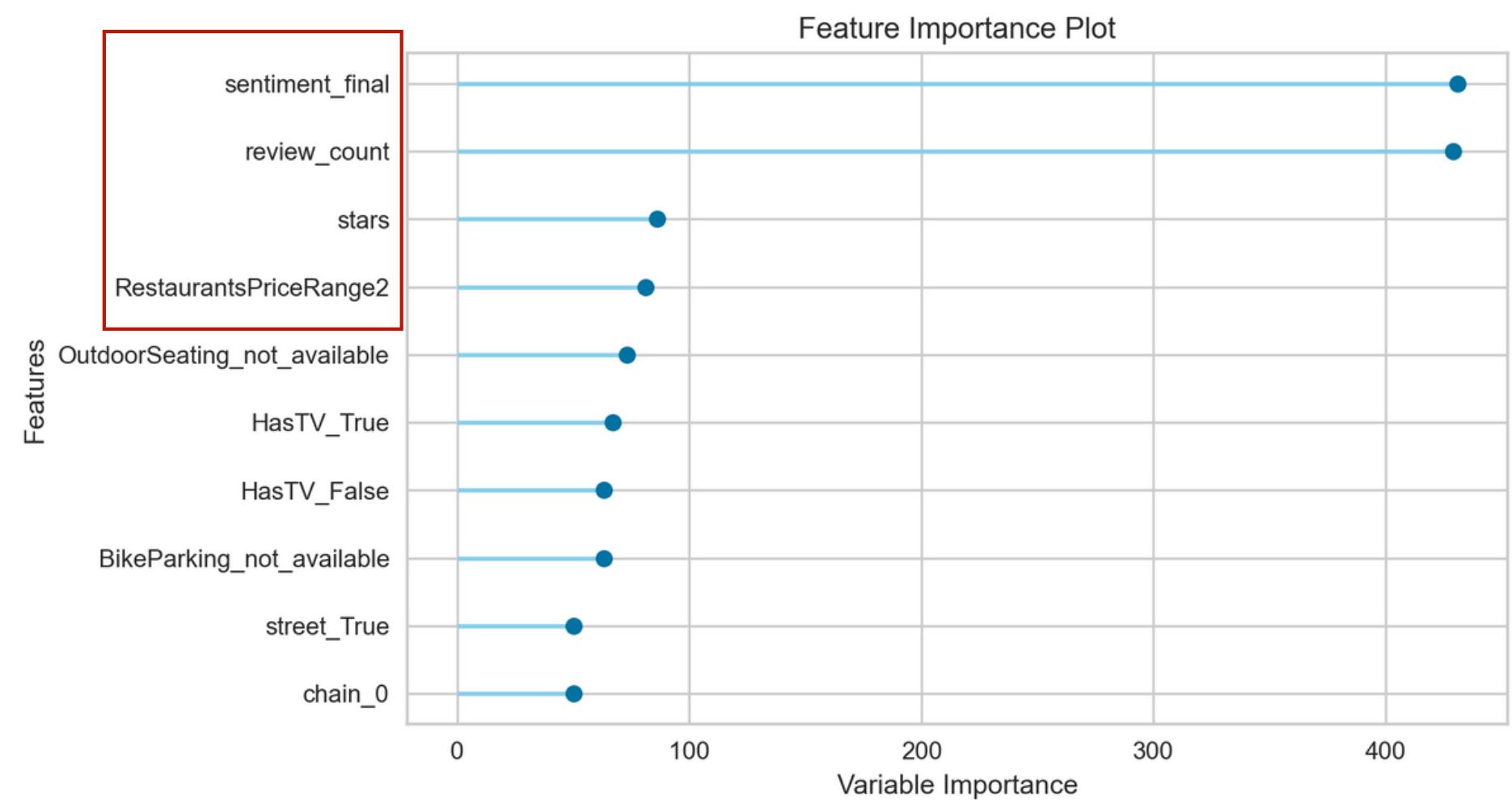


Model Evaluation

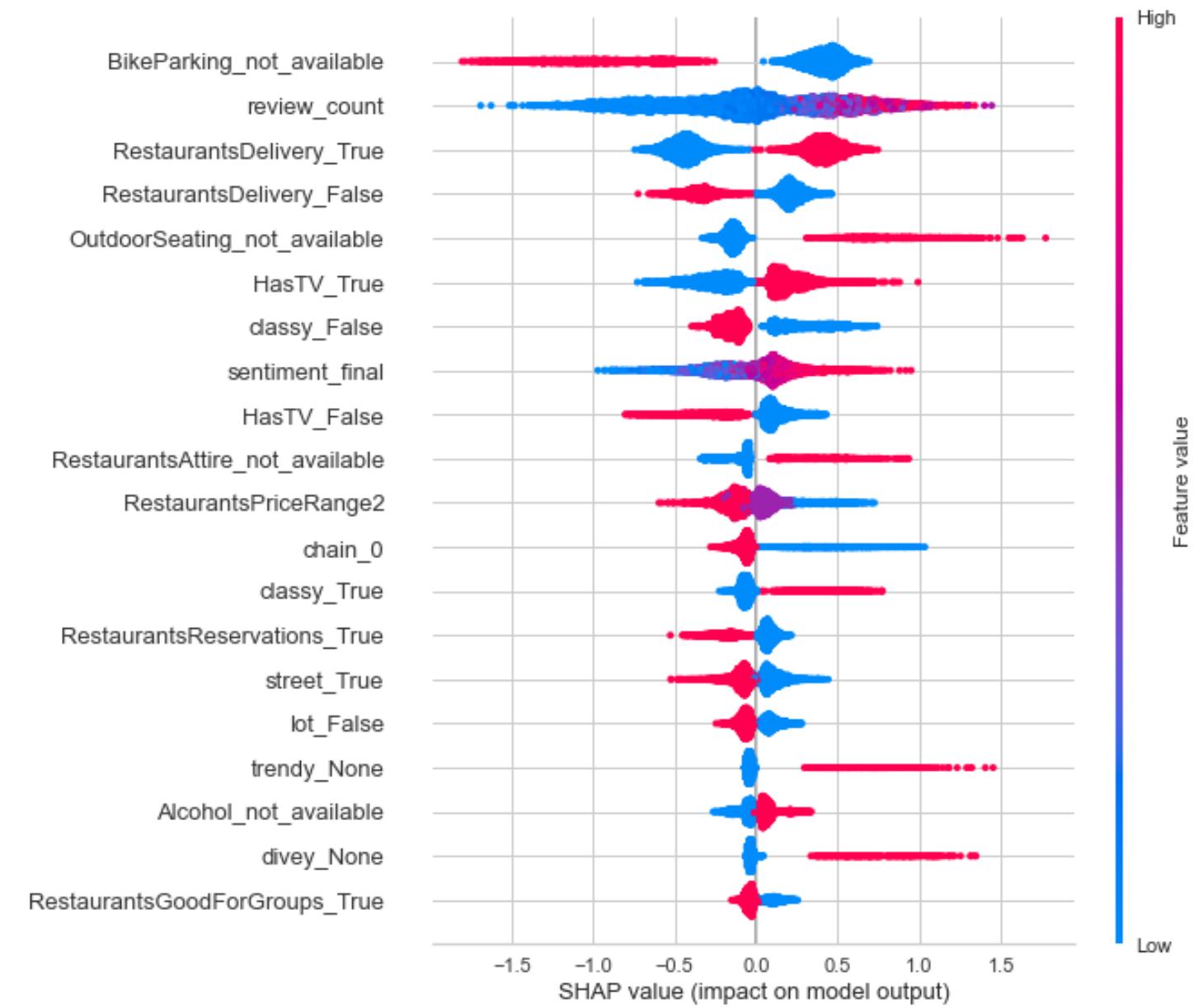
The average precision is around 90%. The rate of precision decreases as the recall rate increases.

The ROC plot shows a smooth curve with an AUC of 0.88. The model has an 88% chance can distinguish between an open or permanently closed restaurant.

Importance Feature



Sentiment score, Review counts, Ratings, and restaurant price range has a larger effect on the model to predict whether a restaurant is close or open.



SHAP value show the positive and negative relationships of the predictors with the target variable.
For example: Restaurants with delivery has positive impact on the model to predict the restaurant is open.

CONCLUSION



Key Elements of Successful Restaurants

- 01** Maintain High Positive Review Ratio
- 02** Well trained staffs
- 03** Quality Customer Service
- 04** Marketing Strategy
- 05** Franchise



We are not in the coffee business serving people,
we are in the people business serving coffee.

- Howard Schultz-
Starbucks CEO

RECOMMENDATION

01. Create the right culture

Cultivate an positive work environment where team members love what they do and where they work.

02. Invest in people

Provide benefits to the employee includes comprehensive training programs, strong benefits and opportunity for growth.
Happy employees = Happy customers