



Assess COVID-19's Impact on US Restaurants



Through Yelp Data

Team018

Yanlin Chen(ychen3402), Fangzhou He(fhe48), Xinyue Li(xli901),
Jianan Li(jli3059), Yaozhen Li(yli3307)



Summary

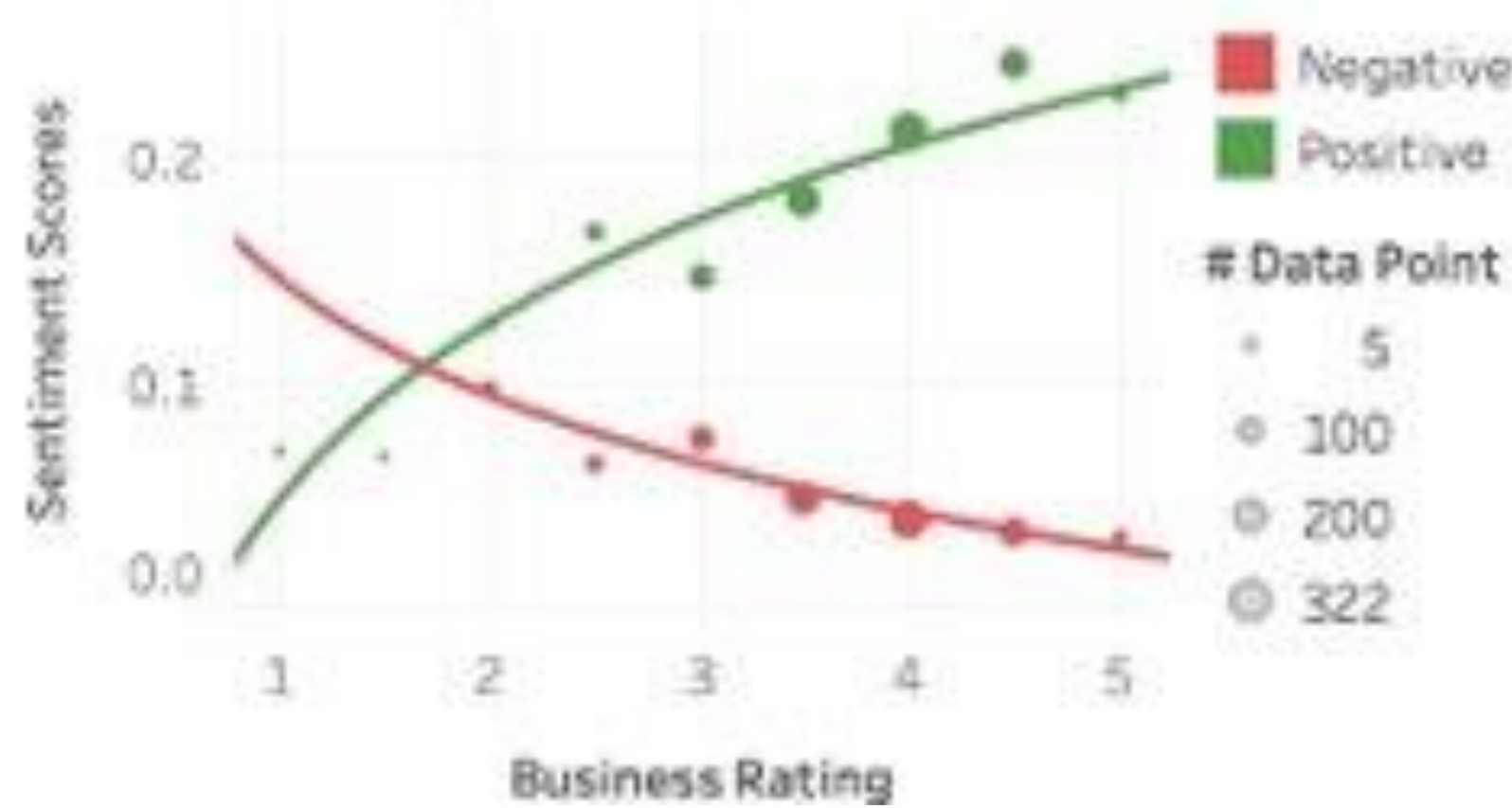
COVID-19 has tremendously impact to thousands of business. How should restaurant owners make changes to survive through? We performed a mixed approach, combining **interactive visualization** and **NLP analysis**, to guide restaurant owner to incrementally explore their neighborhood latest trends through reviews.

Sentiment Analysis by Review Topics

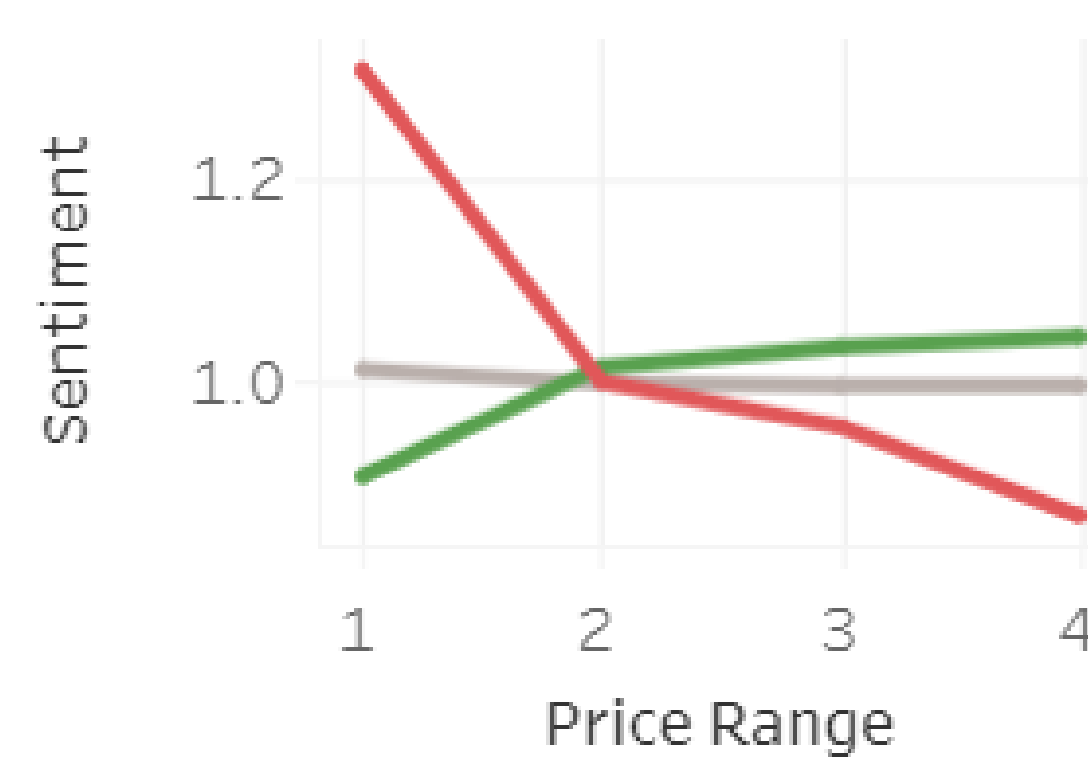
LDA review topics	Covid Status		
Delivery	After Covid	21.00%	76.00%
Order/Waiting time	Before Covid	65.80%	24.95%

Below is an example of Wyoming's closed restaurants' review topic, from which business owner could infer that they should enhance delivery service and order awaiting time to minimize risks of closing.

Business Rating VS. Sentiment Scores



Price Range VS. Sentiment

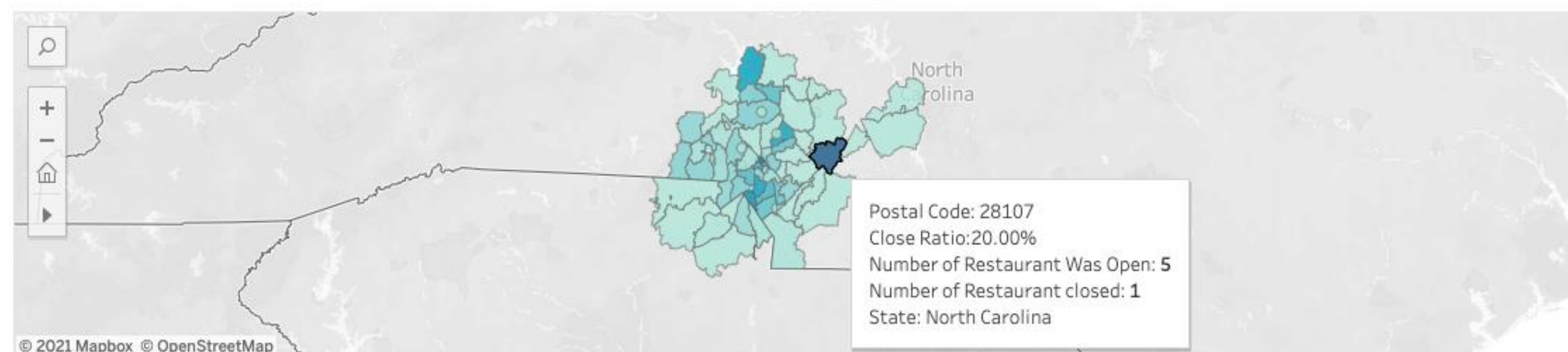


Observations

Sentiment analysis through review topic reveals that customers are more positive towards higher priced restaurants and are more negative to lower priced ones. Also, those remained opened during pandemic's have fewer negative comments on food quality. Meanwhile, there are more reviews on Delivery, Location, Order/Waiting time since these are the major ways in which restaurants engaged with customer during pandemic.

Postal Code
(All)

Closed Ratio Per Zipcode



There will be closed restaurant among different start range and price range based on selected state and zip code area as right histogram shows. This result got proved with NLP model result : Restaurants priced at lower range more likely to close, and restaurants with higher star ratings are more likely to be closed during the pandemic.

Sentiment Analysis by Review Topics

LDA review topics	Covid Status		
Atmosphere/Staff	Before Covid	28.21%	68.86%
	After Covid	31.10%	66.59%
Fish	Before Covid	27.62%	68.93%
	After Covid	28.75%	68.28%
Dessert/Coffee	Before Covid	23.51%	73.03%
	After Covid	25.10%	71.78%
Menu/Service	Before Covid	23.14%	73.51%
	After Covid	25.29%	71.76%
Delivery	Before Covid	21.77%	73.84%
	After Covid	22.18%	73.40%
Meat	Before Covid	21.28%	74.71%
	After Covid	21.98%	74.08%
Staple Food	Before Covid	19.03%	76.93%
	After Covid	19.97%	76.07%
Fast Food	Before Covid	18.50%	76.33%
	After Covid	18.97%	75.83%
Location	Before Covid	17.37%	78.16%
	After Covid	18.16%	77.30%
Order/Waiting time	Before Covid	14.64%	78.57%
	After Covid	15.07%	77.97%

Evaluation & Results

The interactive tool was designed to allow users to view the **results in different geographic levels** to aide in gaining an **overall understanding** of the local needs. Utilizing unstructured review data, it allows users to draw **personalized insights** of the COVID-19 situation and **customer preferences** of the restaurants in local neighborhoods. We had some restaurant owners **test** our public Tableau interface and received very **positive** feedback.

Finally, our tool & approach can be extended to assess user review topic changes before / after any public event similar COVID-19.

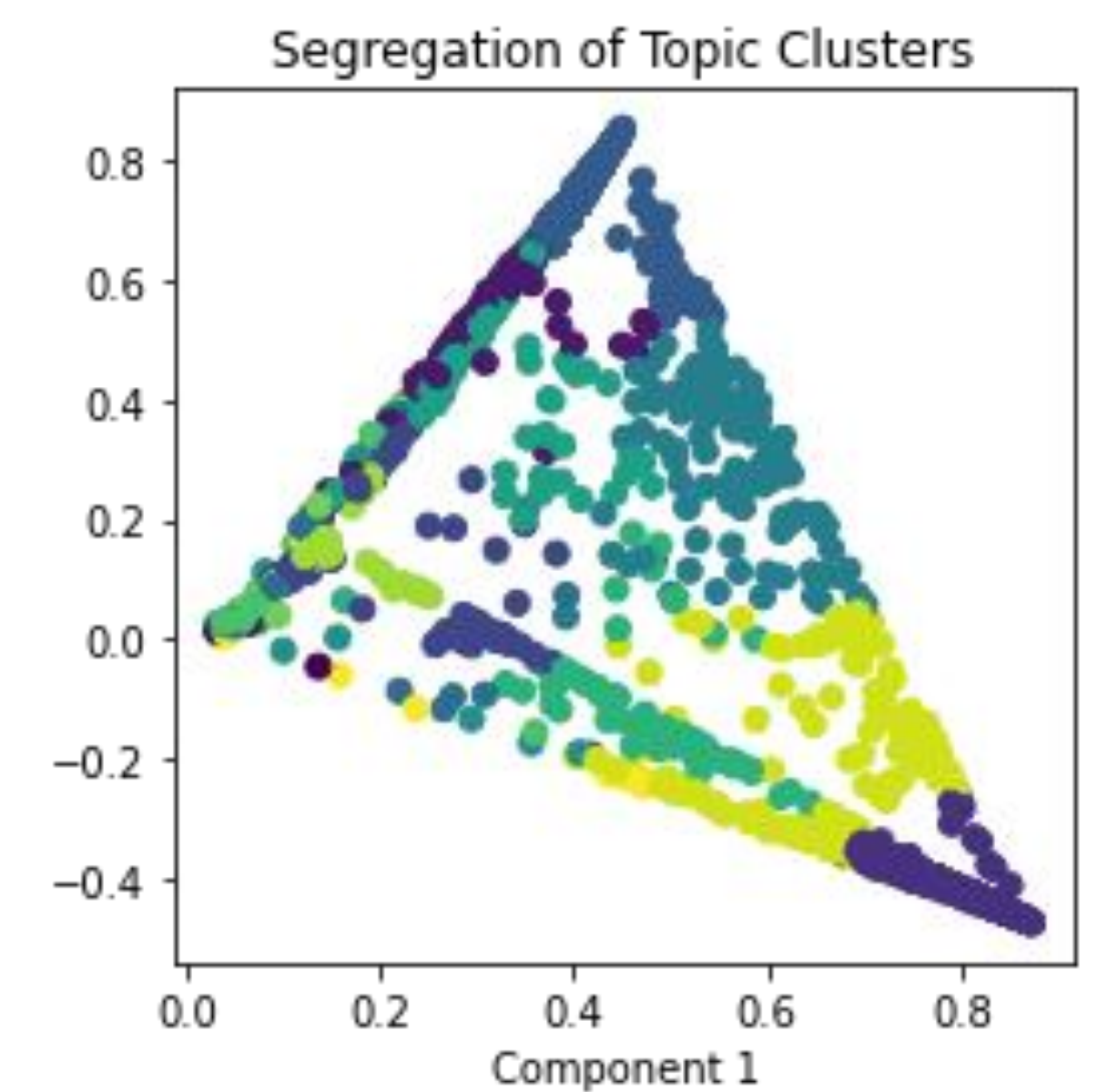
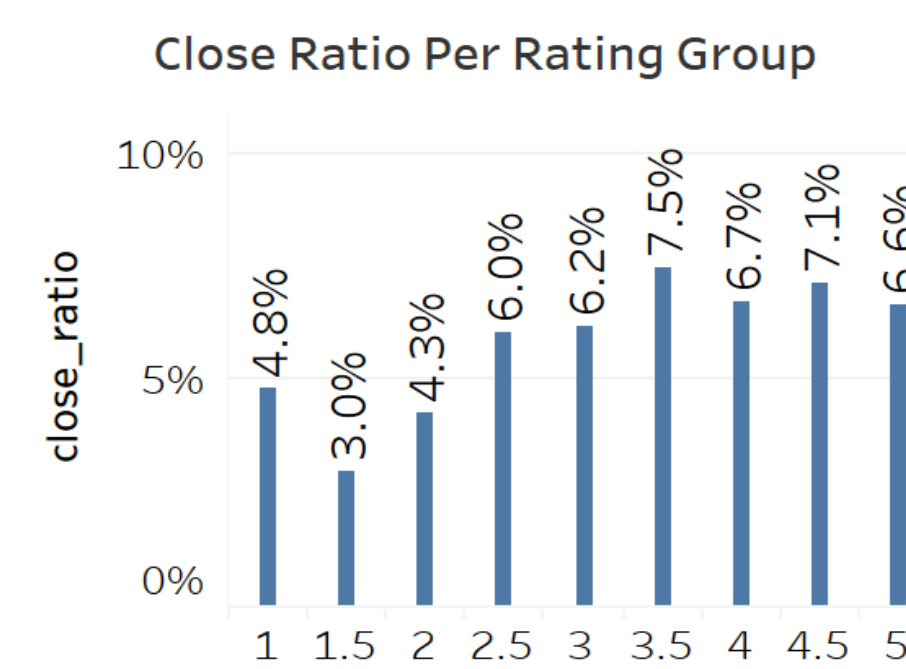
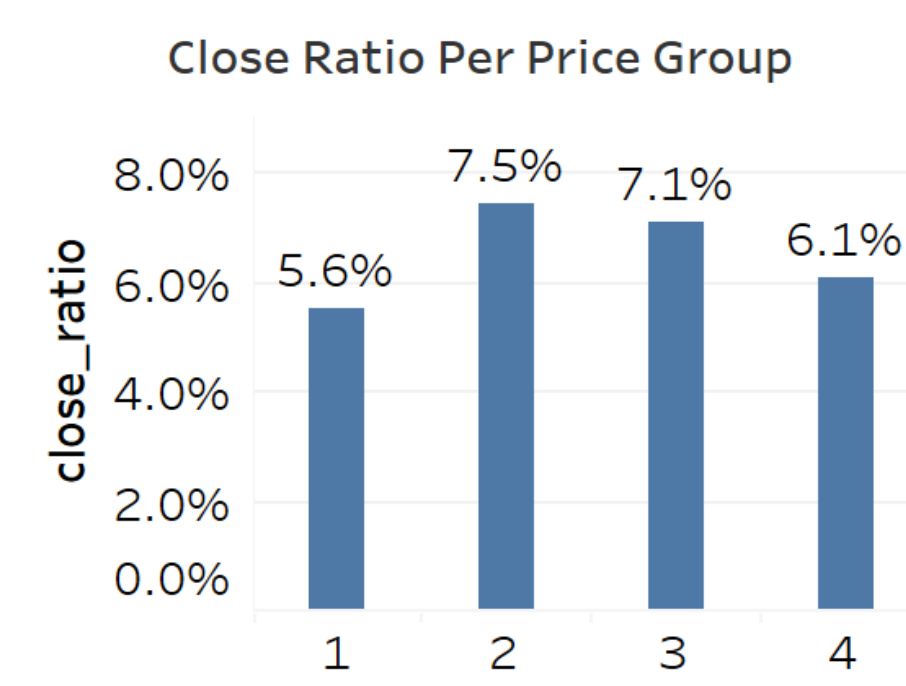
Data From Yelps'API

Type	Size	Period Covered
Business detail	200k records	Until April 2021
Customer review	8000k records	Until April 2021
COVID-19 policies	200k records	Until June 2020

Joined the data set through unique Business ID
Used 5% for sample test during proposal
Used 10% for optimize algorithm during progress
Used 20% for final delivery of results

Approach

We used Python for mining the text review and identify the geography close rate. For comparison purposes, we calculated the **close rate** in both state and zip code level. We grouped the closed business ID and used the **NLTK features and classifiers** to detects polarity with a score. We used **LDA** to generate topics for each category in reviews. Below plot shows our model generated 20 topics.



Interactive Visualization Tool

We provided a map visualization of the closing ratio for each state. For each mouse over region, there is detail close ratio information will be display in tooltips and closed group in bar in the left figure. Users can select a detail zip code area for more specific area. Over the left is an example of selecting all zip code areas available for North Carolina.

Topics% across Topics by Covid Status

LDA review topics	Covid Stat..	
Menu/Service	Before Covid	45.95%
	After Covid	42.35%
Order/Waiting time	Before Covid	27.58%
	After Covid	31.03%
Dessert/Coffee	Before Covid	9.31%
	After Covid	8.97%
Atmosphere/Staff	Before Covid	6.13%
	After Covid	4.45%
Staple Food	Before Covid	4.00%
	After Covid	5.12%
Meat	Before Covid	2.45%
	After Covid	2.28%
Fast Food	Before Covid	1.40%
	After Covid	2.34%
Delivery	Before Covid	1.36%
	After Covid	1.78%
Location	Before Covid	0.96%
	After Covid	0.87%
Fish	Before Covid	0.87%
	After Covid	0.82%