

SWEET!!
OR
TOO SWEET?

STAT 628 - M3 - Project

Niharika Chunduru

Siyan Wang

Yiyuan Li



THE PROBLEM WE WISH TO ADDRESS



Target Sample: Bakery Business Owners in Santa Barbara



Since 2005, more than a third of the bakeries got shut down over the years



A comparative analysis study of open and closed bakeries in Santa Barbara



Helping current and new business owners identify factors impacting their business star ratings

DATA HANDLING

DATA SOURCES:

Business & Reviews Data from Yelp, and Average Income Data from US Census Bureau.

DATA PRE-PROCESSING:

1. Word-cloud and Aspect-Based Sentiment analysis:

- Separated the reviews to closed and open bakeries reviews
- Classified reviews as positive reviews (≥ 3.5 stars) and negative reviews (< 3.5 stars).
- Tokenized word reviews and identified top 5 Aspect dimensions: Food, Drinks, Service, Ambiance, and Price

2. Zip-Code Analysis:

- Calculated the duration of operation of a bakery from the last posted review
- Extracted and processed income datasets for each zip-code in Santa Barbara from 2011 till 2021

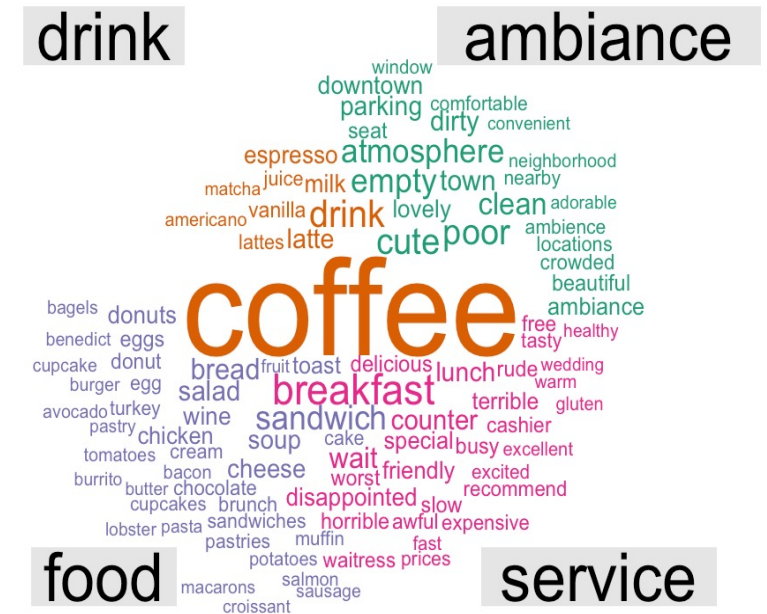
WORD-CLOUD ANALYSIS

Aspect	Example Words	
Food	croissant	cake
Drink	coffee	latte
Ambiance	beautiful	dirty
Service	warm	awful

Word analysis based on top four aspects dictionaries.

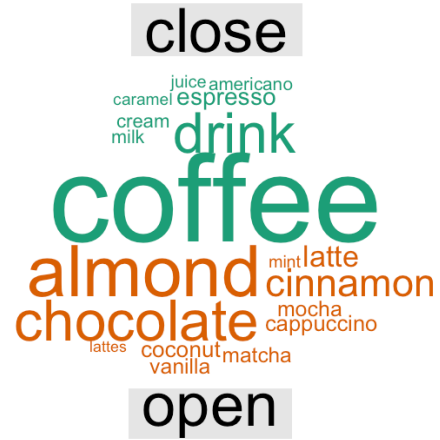


Negative Reviews from Closed Bakeries



Positive Reviews from Open Bakeries

Drinks

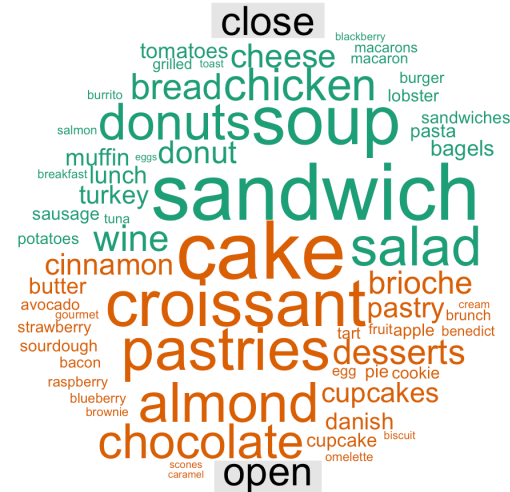


close



Ambiance

Food



close



open

Service

SUGGESTIONS TO BAKERY OWNERS:

- Providing more choice on beverages for customers, including more flavors, coffee types and milk options (Drinks)
- Try to spend more time making regular pastry food, rather than brunch items (Food)
- A comfortable environment and the parking space also help for the survival of bakeries (Ambiance)
- Be friendly to customers and don't let them wait for too long (Service)

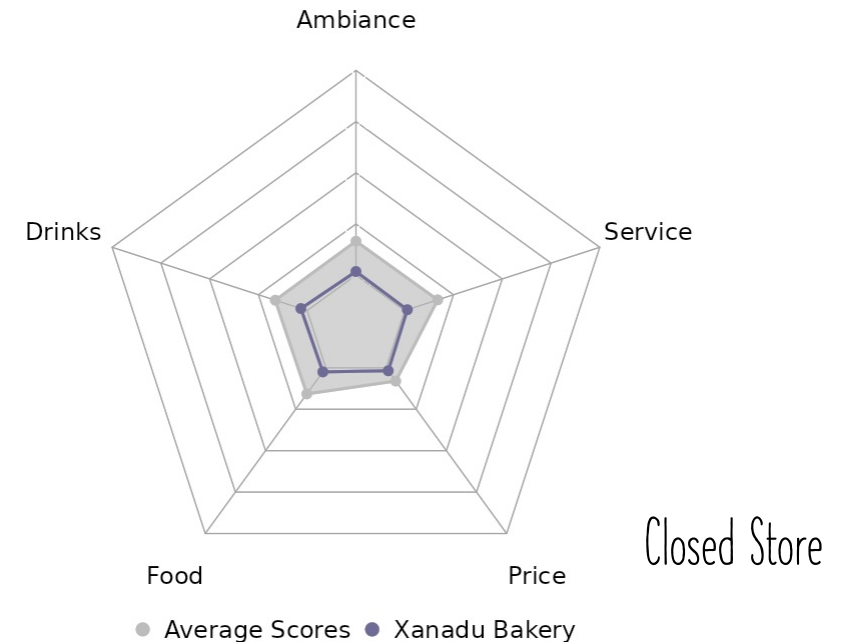
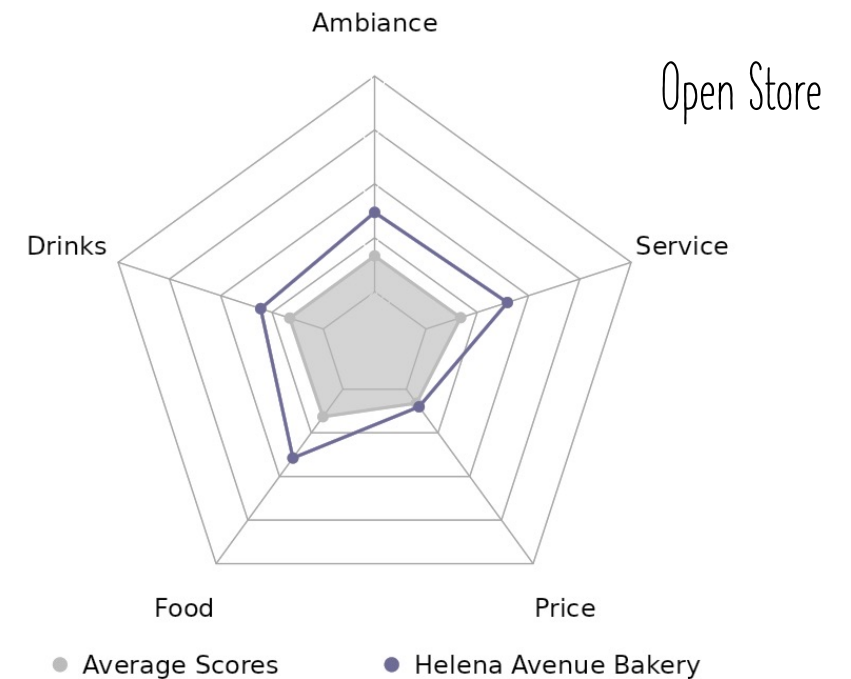
SENTIMENT SCORE ANALYSIS

- Sentiment lexicon "afinn"
- Calculated sentiment score for each store by their reviews
- Labeled each store by their sentiment score

Scores	Label
$\{score\} \geq 1.5$	The Best!
$0.7 < \{score\} < 1.5$	Worth a visit
$0 < \{score\} < 0.7$	Not bad
$\{score\} < 0$	Many faults

ASPECT SCORE ANALYSIS

- Scored the reviews based on top 5 aspect dictionaries
- Compared the bakery store scores with the average performance in the county
- Separated open and closed bakeries to provide easy comparison



STAR ~ {ASPECTS}

- Aspects are independent
- Aspect score per star-rating for each bakery
- Naïve Bayes Classifier with Aspects as features, and Star Rating as the variable
- Tells the contribution of each aspect in predicting star rating

1 Unit change in...	Open Bakeries	Closed Bakeries
Ambiance ↑	0.32 ↑	0.72 ↑
Drinks ↑	0.22 ↑	0.70 ↑
Food ↑	0.10 ↓	0.50 ↑
Price ↑	0.10 ↓	1.46 ↑
Service ↑	0.32 ↑	0.08 ↓



THANK YOU FOR LISTENING!