

# Sentiment Analysis of Yelp Reviews for Chipotle

STAT628 Module 3 Group 10  
Brian Tsai, Tinghui Xu, Lanxin Xiang



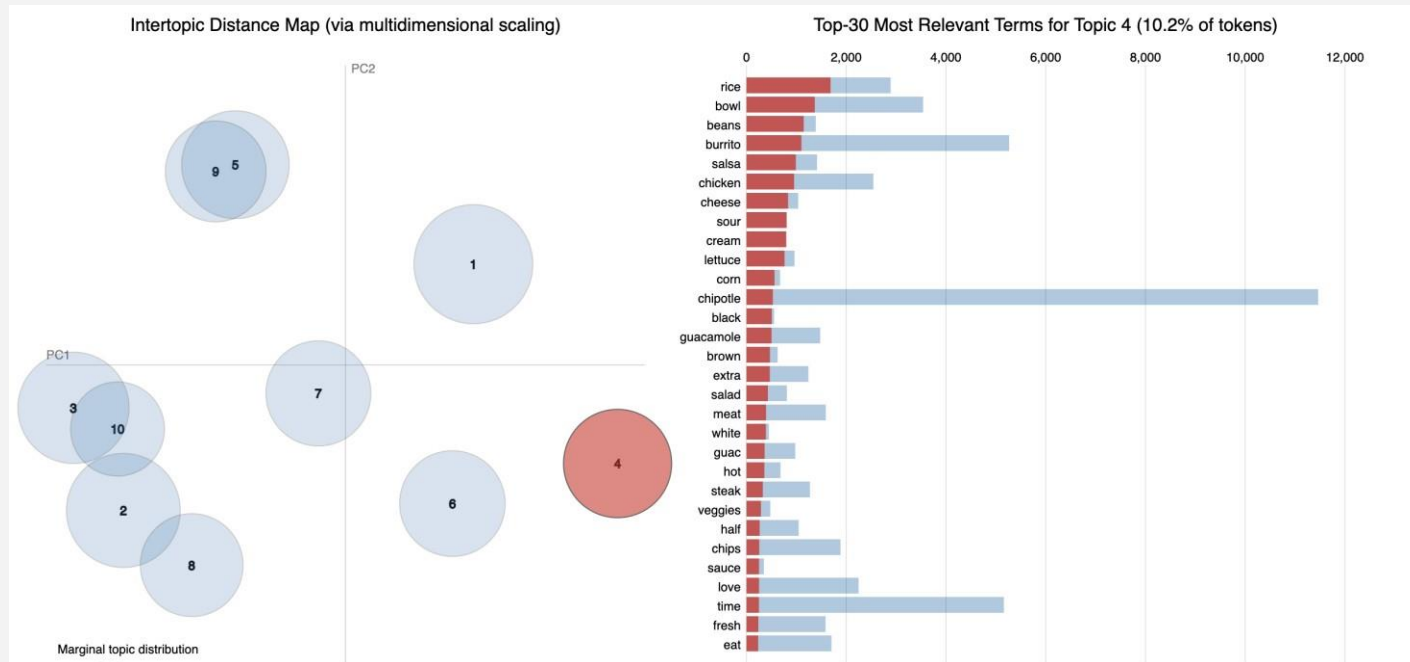


# Table of content

- ◆ Use LDA to filter out aspects
- ◆ Assign sentiment score to each aspect
- ◆ Compare scores between restaurants and give suggestions
- ◆ Discussion and Conclusion

# Latent Dirichlet Allocation (LDA)

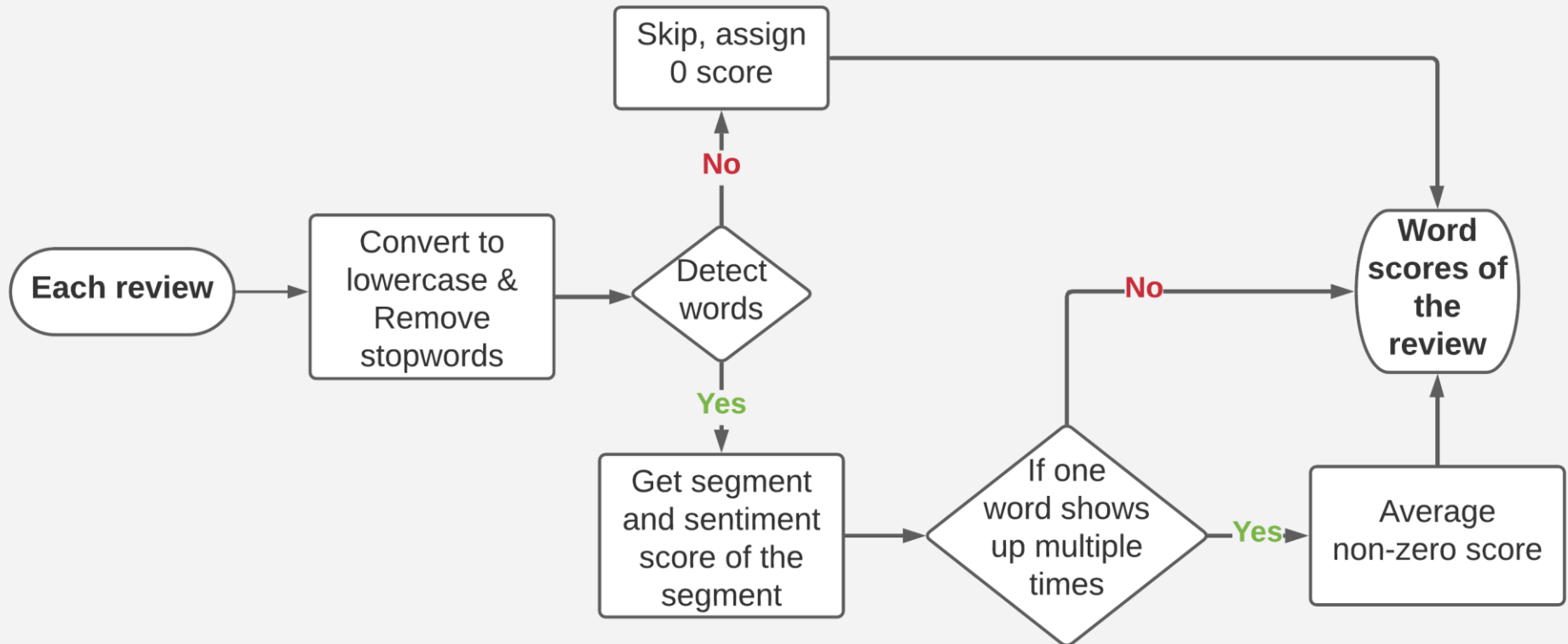
- ◆ Figure out important aspects from the top 30 words under each topic in the output of LDA.
- ◆ Select relative words under each aspects



# Four Aspects and Relevant Words

ASPECTS	Food	Service	Price	Ambience
Relevant words	burrito	time	price	ambience
	rice	attitude	cheap	atmosphere
	bowl	staff	affordable	table
	Sour cream	manager	expensive	seat
	ingredient	delivery	cost	decor
	...	...	...	...
	fresh	service	refund	parking
Number of words	37	44	8	19

# Calculate sentiment score for each word

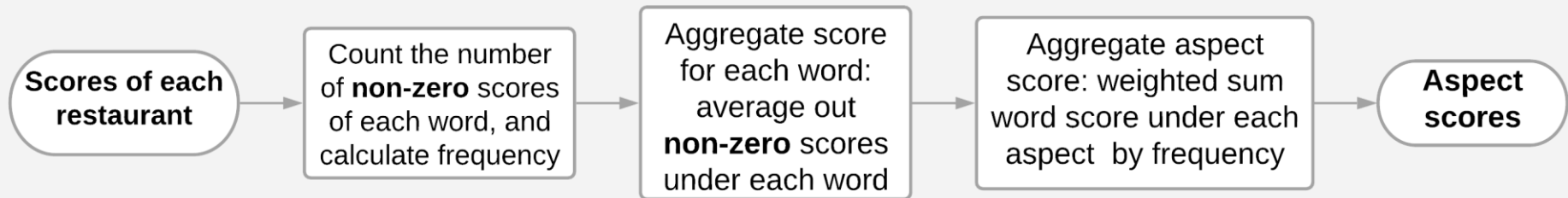


- ◆ Example:
- ◆ “I love Chipotle, *hands down my favorite place to go for a **burrito bowl**.*  
*The **food** is great, quality, quantity is all good. The only reason I won't give it 5 stars is because a couple of the **employees** give you unfriendly looks when you ask for a certain thing this way or that way but most are great! **Taste** wise it's the best Chipotle I've been to so far!”*

Food Aspect			Service Aspect
Burrito bowl	Food	Taste	Employee
0.459	0.625	0.807	-0.361



# Aggregate sentiment score for each restaurant



business_id	black	guac	burrito	door	table
1234567	0	0.1	0.3	-0.1	0
1234567	-0.1	0	-0.4	0	0.3
1234567	0.2	0	0	0	-0.1
1234567	0	0.1	0	0	0.1

business_id	black	guac	burrito	door	table
1234567	0.05	0.1	-0.05	-0.1	0.1

business_id	Food score	Service score
1234567	$0.05 \times \frac{2}{6} + 0.1 \times \frac{2}{6} - 0.05 \times \frac{2}{6} = \mathbf{0.033}$	$-0.1 \times \frac{1}{4} + 0.1 \times \frac{3}{4} = \mathbf{0.05}$



## Real example:

business_id	black	guac	amount	...	door	table	...	service	time	...	price	free	...
__a6cve2cp TdOP6uNXh ---	0	0.444	0.359	...	-0.024	-0.340	...	-0.089	0.137	...	0	0.462	...
business_id	Food score				Ambience score			Service score			Price score		
__a6cve2cp TdOP6uNX hRMw	0.174				-0.127			0.072			0.462		



# Suggestions Based on Sentiment Score

- ◆ Compare scores with **average** and **75% quantile**:
  - If the score is higher than 75% quantile score, we consider they are doing great in this part.
  - If the score is lower than average, we give suggestions on improving that.

	Food Score	Service Score	Price Score	Ambience Score
75% Quantile	0.307	0.189	0.463	0.117
Average	0.242	0.124	0.326	0.042



business_id	Food score	Service score	Price score	Ambience score
AtJFEVY58j8UU5K LRsA0UA	0.277	0.100	0.377	0.084

- **Your FOOD SCORE is slightly better than average store.** Customers enjoy guac, cilantro, salsa, cheese, sour cream, chicken, sauce. The followings are not as satisfying: water, cilantro, lettuce. As for taste, your restaurant gets less scores than the average.
- **Your SERVICE SCORE is lower than average.** The serving speed should be improved. Paying more attention on online services, including ordering with app and phone call, pickup, deliver services may help you attract more customers.
- **Your PRICE SCORE is slightly better than average score.** Customers think the food really worth its price.
- **Your AMBIENCE SCORE is slightly better than average score.** Sanitary condition needs improvement. Customers may find it inconvenient to park around your restaurant.



# Shiny app

Link: <https://tinghuixu1114.shinyapps.io/chipotleanalyzer/>



# Discussion and Conclusion

## Strength

- ◆ Explicitly calculate scores for four main aspects of reviews.
- ◆ Make comparisons between different Chipotle restaurants based on quantified performance.
- ◆ Being able to give specialized advice for each aspect for each restaurant.
- ◆ Being able to know how much a word contribute to overall score.

## Weakness

- ❑ The selection of words is somehow subjective, even if we applied LDA method.
- ❑ Same words may used to describe multiple aspects.
- ❑ Some times the segmentation is not accurately enough, especially when the object and its adjective are not in the same segment.

An aerial photograph of a city waterfront at sunset. The sun is low on the horizon, casting a warm, golden glow over the scene. The city skyline is visible on the left, with various buildings and a prominent tall skyscraper. The water is dark blue with many sailboats scattered across it. The text "THANK YOU." is overlaid in large, white, sans-serif capital letters.

THANK YOU.



2017