

A large, thin black circle frames the central text. Two solid red dots are positioned on the circle's circumference, one in the upper-left quadrant and one in the lower-right quadrant.

# American Traditional Restaurant

——Based on Yelp Review Data



# CONTENT



Introduction



Review Text



Attributes



Suggestions



# 1 | Introduction

## □ Primary Goal:

- Develop suggestions for 6307 American traditional restaurants in our dataset

## □ Methods:

- Review Text: NLP, LASSO regression
- Attributes: Decision tree, ANOVA

## □ Suggestions:

- Miscellaneous suggestions for all the restaurants
- Specific suggestions for each business on Shiny App



Review Text

# Utilize NLP to process review text

- Lowercase
- Noun Phrases, replace space with underline
- Lemmatization
- Add negative tag before words
- Remove punctuation



# Lasso Regression

- Calculate Tf-idf for each word
- Select words of high frequency, they are basically nouns and adjectives
- Data set contains: 483009 Reviews, 125 Words

# Lambdas and Coefficients

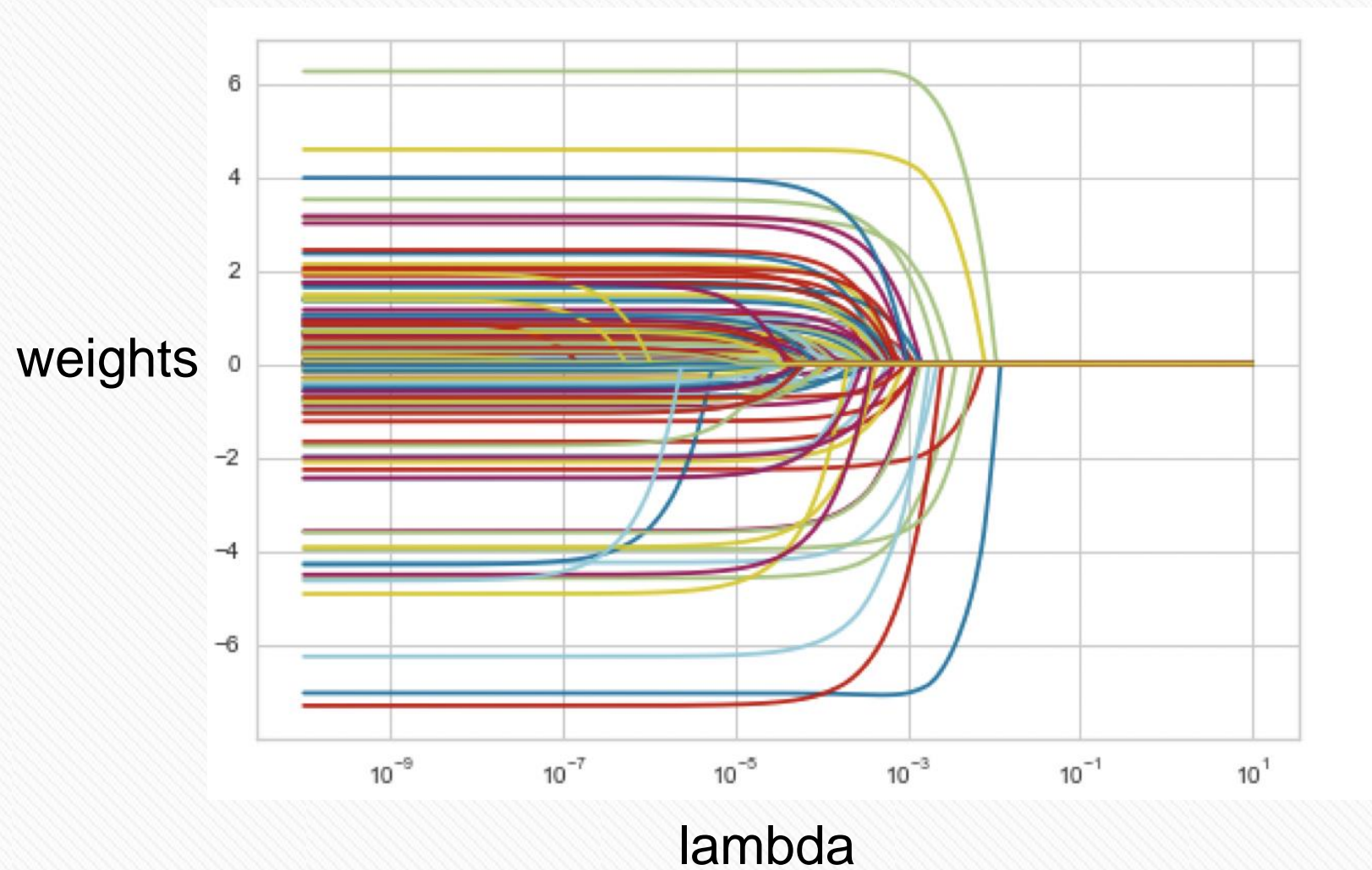


Figure 1: Variable selection plot



# Results

Use **Cross validation** to select **Best Lambda**:  $3.54e-06$

<b>slow</b>	<b>delicious</b>	<b>kitchen</b>	<b>reasonable price</b>	<b>fresh</b>
-4.59	1.00	-2.10	2.36	3.12
<b>minute</b>	<b>bland</b>	<b>location</b>	<b>courteous</b>	<b>loud</b>
-7.04	-7.32	-1.23	2.01	-1.97

Table 1: Part of Results



# Attributes

# Attributes

## □ Raw Data:

- More than 70 attributes and a lot of missing data in them

## □ Select Features:

- Attributes with at least 3000 non-missing entries

## □ Assign “None” to missing values:

- Treat “None” as a level when fit decision trees

# Decision Tree

12

Attributes of Importance:

- Caters
- Business Parking
- Good for meal
- Alcohol
- Good for kids
- ...

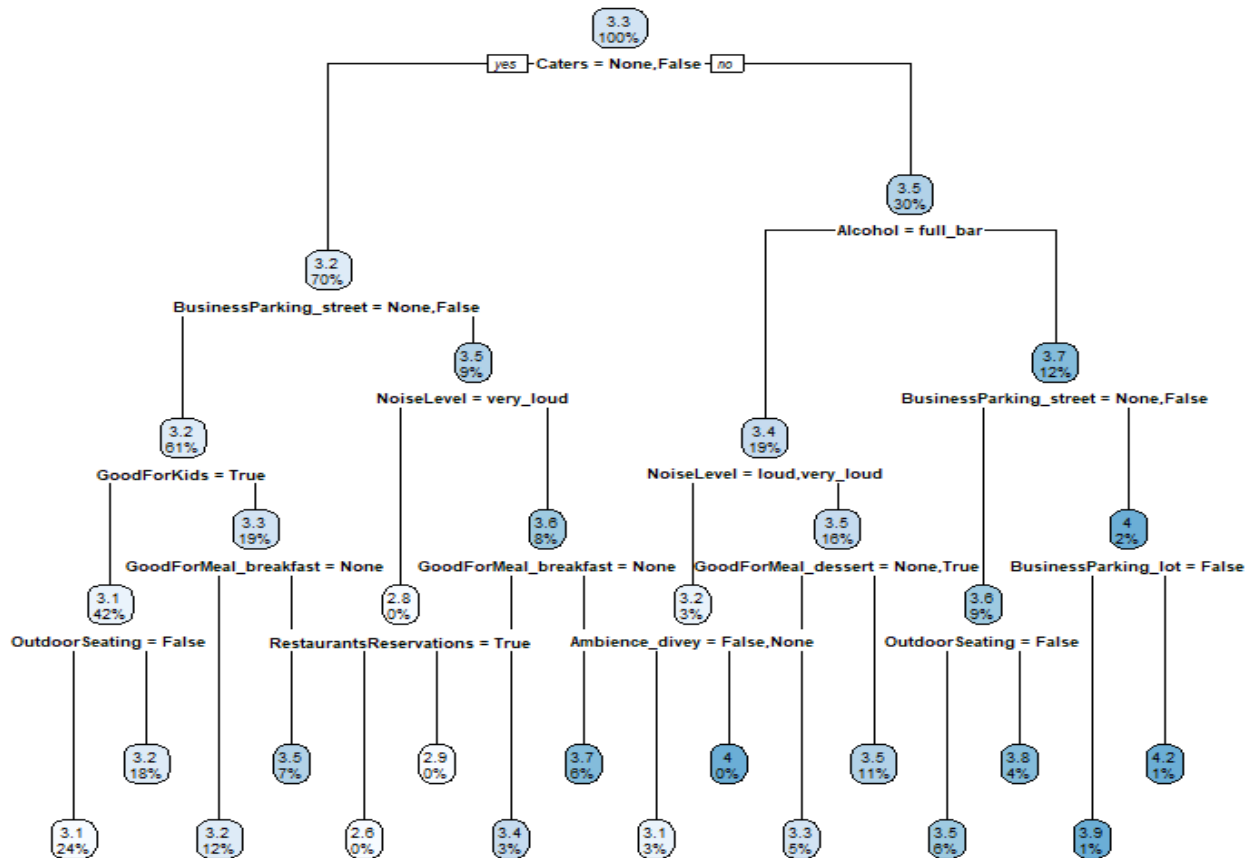


Figure 2: Decision Tree

# Regression Model for Categorical Variables

Table 2 & 3: Part of Results:

Coef (Significance)	Caters True	Business Parking	Good for kids
True	-0.045(.)	0.241(***)	0.003
False	0.207(***)	0.492(***)	-0.212(***)

. < 0.1; \* < 0.05; \*\* < 0.01 \*\*\* < 0.001

Coef (Significance)	Noise Level
Quiet	-0.07
Average	-0.08(*)
Loud	-0.35(***)
Very Loud	-0.58(***)





# Suggestions

Based on two parts of analysis in  
previous presentation

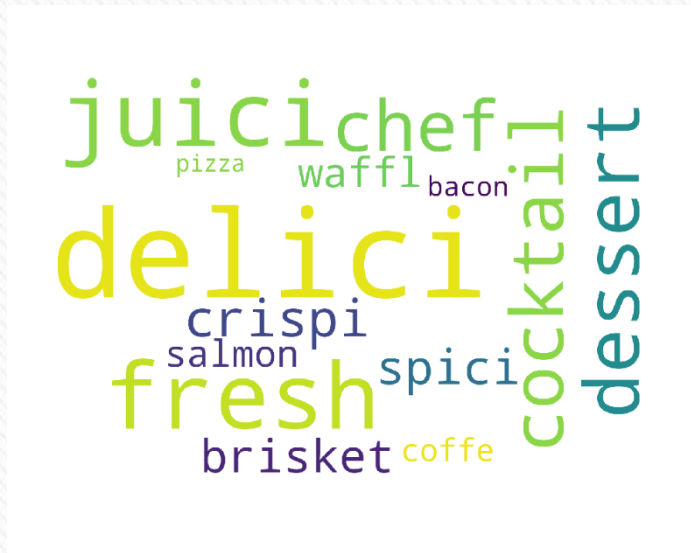


Figure 3a: positive words

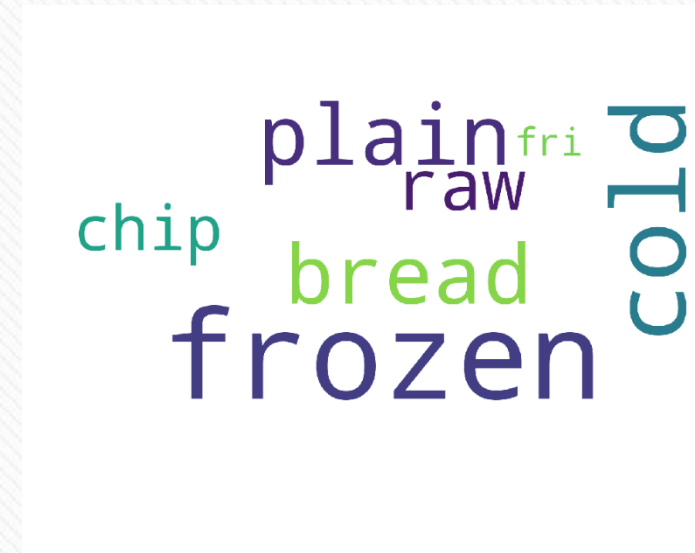


Figure 3b: negative words



- Fresh food is essential for your business. Keep the materials fresh!
- Delicious waffle, salmon and desserts attract customers.
- Crisp flavor and juicy style are favored.
- People would love tea and cocktails for their drinks.

# Service



Figure 4a: positive words



Figure 4b: positive words



- Train your waiters and bartenders well to be courteous and friendly.
- Make your customers feel welcomed.
- Work efficiently. Don't let your customers wait for a long time.

# Environment



Figure 5a: positive words



Figure 5b: positive words



- Having great background music may make your restaurant popular.
- Control the noise level.
- Keep your table and bathroom clean all the time.
-

# Facilities

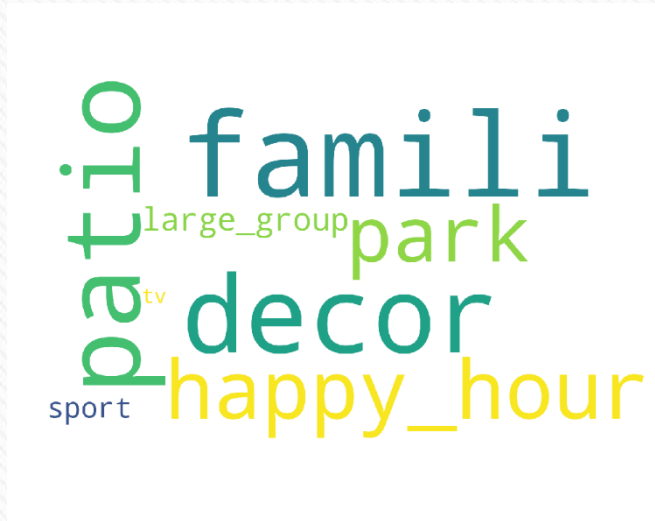


Figure 6a: positive words



Figure 6b: positive words



- Playing sports games on TV makes the customers happy.
- Try to provide parking places and patios in your restaurants.
- If your location is difficult to find, a marked sign may help a lot.



# Shiny App

## Yelp suggestion for Restaurant

Information of a traditional American Restaurant

Please input business\_id

Submit

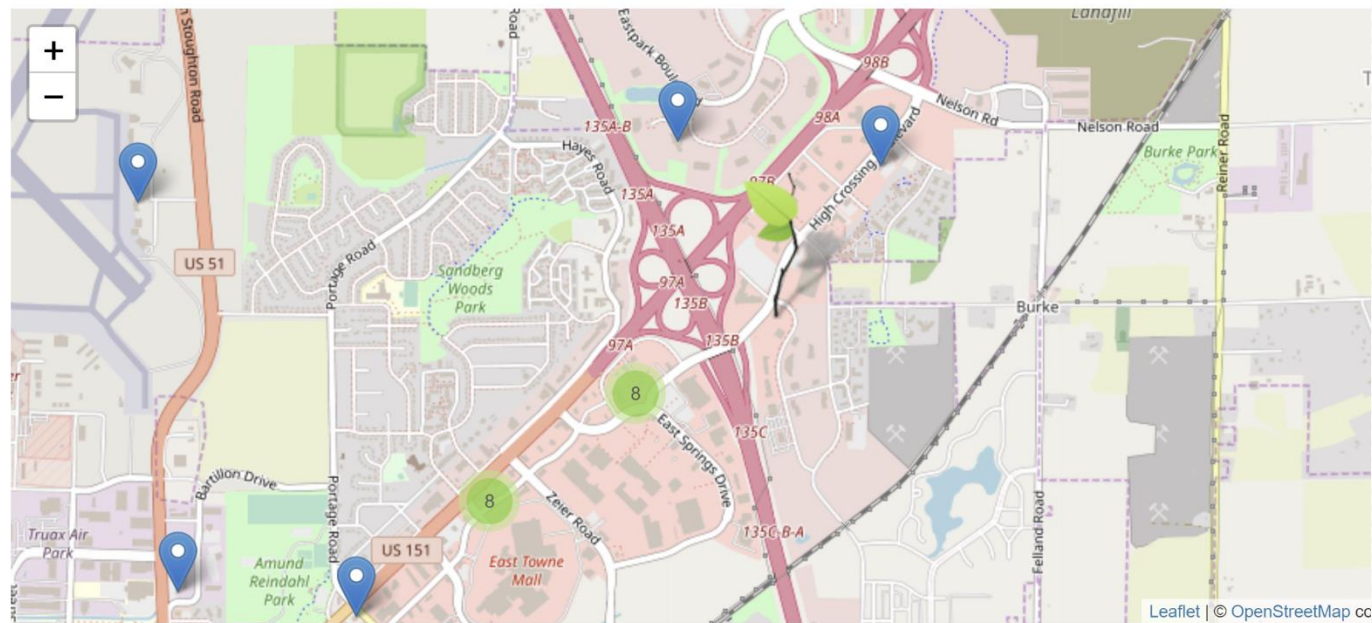
[Review Star for the Restaurant](#)

[Recommendation for Restaurant](#)

[Restaurant Map](#)

[business\\_id for the restaurant name you inp](#)

The map's center is: Uno Pizzeria & Grill



if you have any question, feel free to contact me: [chao.chang@wisc.edu](mailto:chao.chang@wisc.edu)

Link: [https://clarefrost.shinyapps.io/shiny\\_yelp/](https://clarefrost.shinyapps.io/shiny_yelp/)



# Thanks for your listening

谢 谢 聆 听