Chinese Restaurants Analysis on Yelp Data

STAT 628 Group 6

University of Wisconsin - Madison

Nov 21, 2019

Outline

- Background
- Analysis of Attributes
- Analysis of Reviews
- Rshiny

Background

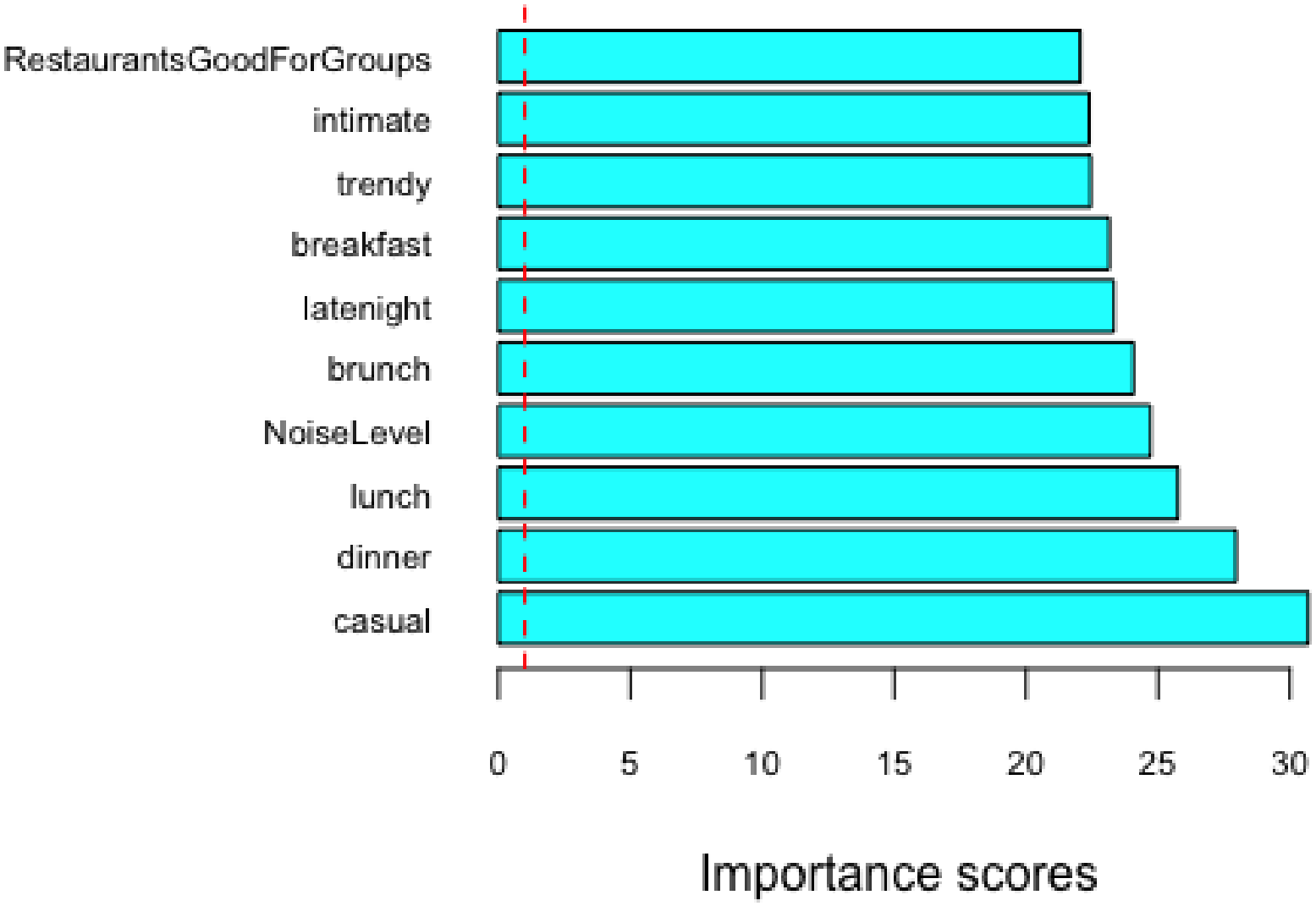
- Based on the reviews and restaurant information from yelp, we aim to provide useful and analytic insights to business owners.
- We are most interested in 4434 Chinese restaurants and want to help them to make data-driven decisions to improve their ratings on yelp.
- We build a R shiny Application to give a more vivid explanation.

Analysis of attributes

- Extract the main attributes and sub-attributes as new columns. Drop 11 attributes with one level. Then we have 60 attributes in total.
- Use GUIDE to construct a classification tree model and select the most important 10 attributes.
- Perform ANOVA and multiple comparisons for each important attribute.

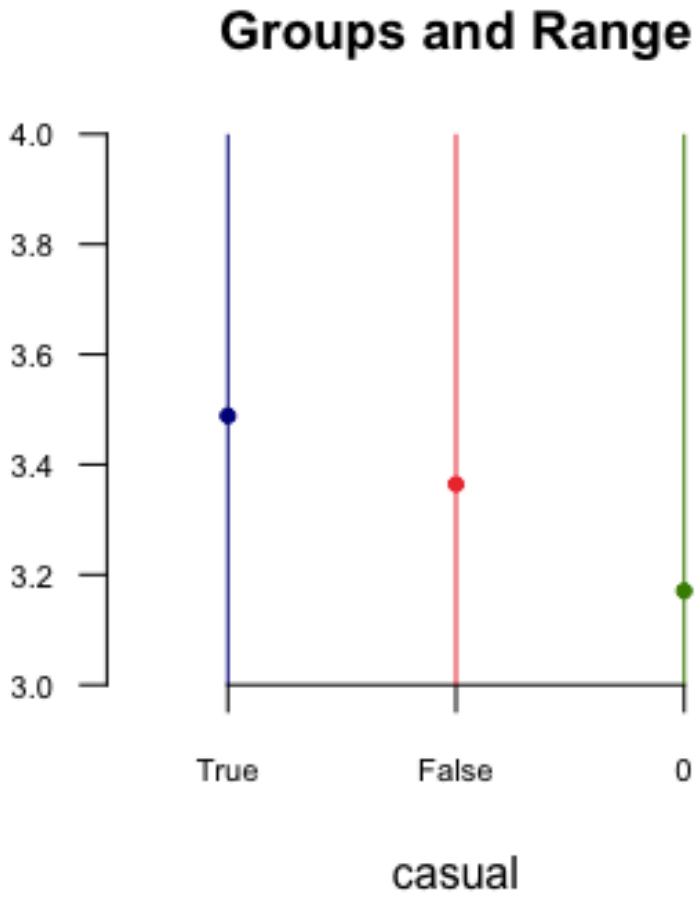
Analysis of attributes

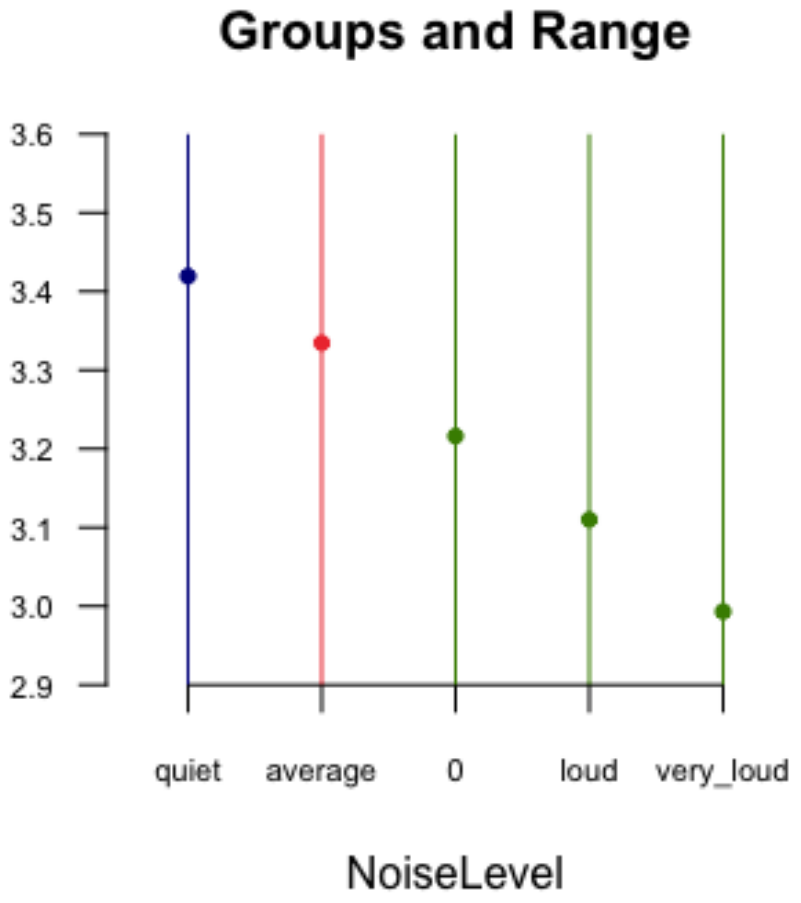
- Importance scores from GUIDE classification tree model.



Analysis of attributes

- Treat the missing data as a new level 0 in each attribute and perform one-way ANOVA.
- Perform multiple comparisons by Bonferroni method.





Analysis of Reviews

Focus on two aspects from the reviews:

- Overall Service
- Food

Analysis Of Reviews – Overall Service

- Get the Overall Service information from three sub aspects:

  1. "wait": wait, waiting
  2. "service": waiter, waitress, staff, serve, service, services, server, servers
  3. "price": paid, price, prices, pay, money, bill

- Convert three aspects to a vector for each review. Use it later in regression.
  e.g. A review does not include "wait", includes "service", "price" $\rightarrow (0, 1, 1)$

Analysis Of Reviews – Overall Service

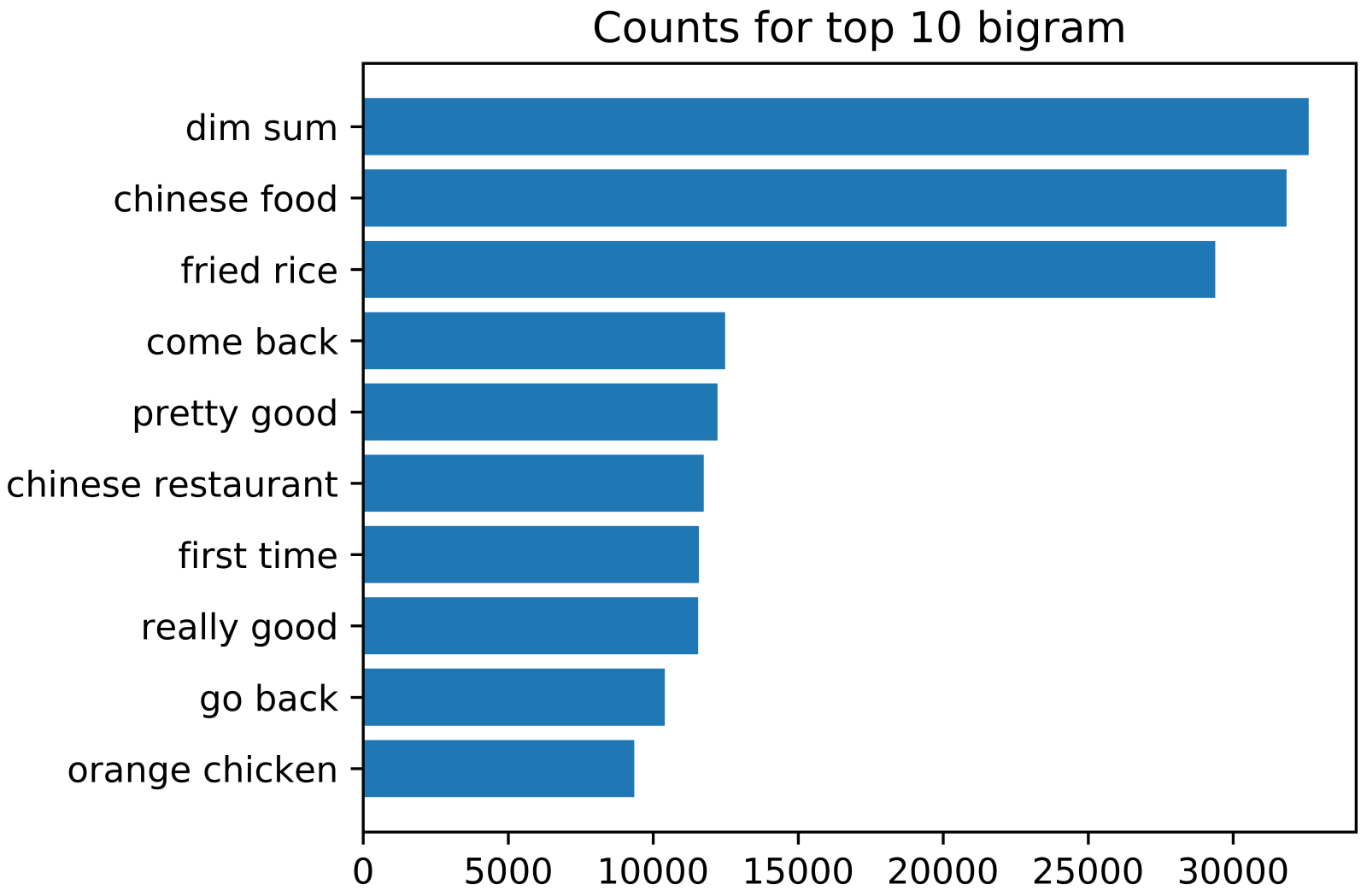
- For each review, find the adj/adv relate to three aspects. Display them on the Shiny App as an intuitive description.
- Ex. "Emerald Chinese Restaurant", stars: 2.5.

  wait: [('long', 4), ('short', 1)]],

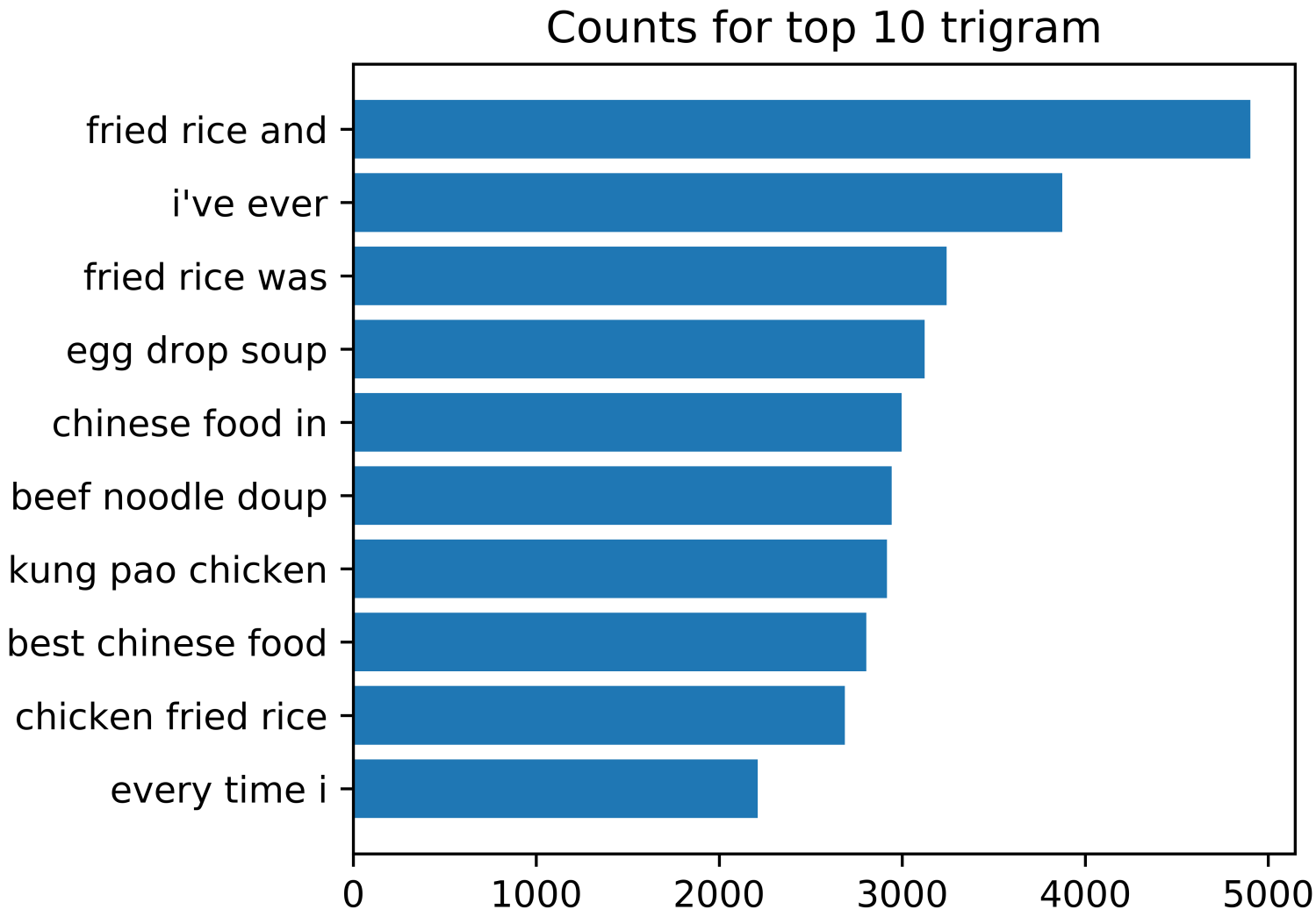
  service: [('good', 11), ('poor', 6), ('bad', 6), ('terrible', 3), ('friendly', 3), ('horrible', 3), ('pretty', 2), ('great', 2), ('attentive', 2), ('nice', 1)],

  price: [('higher', 3), ('reasonable', 3), ('high', 2), ('good', 1), ('expensive', 1)]

Analysis of Reviews - Food



Analysis of Reviews - Food



Analysis of Reviews - Food

Bigram: [('dim','sum'), ('fried','rice'), ('orange', 'chicken'), ('chow', 'mein'), ('pad', 'thai'), ('sour', 'soup'), ('hot', 'pot'), ('spring', 'rolls'), ('lo', 'mein'), ('bbq', 'pork'), ('mongolian', 'beef'), ('egg', 'roll'), ('wonton', 'soup'), ('soy', 'sauce'), ('ice', 'cream'), ('pork', 'belly'), ('milk', 'tea'), ('fried', 'chicken'), ('sour', 'chicken'), ('white', 'rice'), ('sesame', 'chicken'), ('chicken', 'wings'), ('peking', 'duck'), ('bubble', 'tea')]
Trigram: [('egg', 'drop', 'soup'), ('beef', 'noodle', 'soup'), ('kung', 'pao', 'chicken'), ('chicken', 'fried', 'rice'), ('pork', 'fried', 'rice'), ('black', 'bean', 'sauce'), ('shrimp', 'fried', 'rice'), ('xiao', 'long', 'bao'), ('egg', 'foo', 'young'), ('dan', 'dan', 'noodles')]

Analysis of Reviews

- We scan the top 100 Bigram and Trigram frequency distribution, and manually find 24 Bigram food phrases and 10 Trigram food phrases.
- Combine the "Overall service" and "Food" and get 37 dimensional vector for each review
- Use Linear Regression when reviews are more than 37, use Lasso Regression when reviews are less than 37

Analysis of Reviews

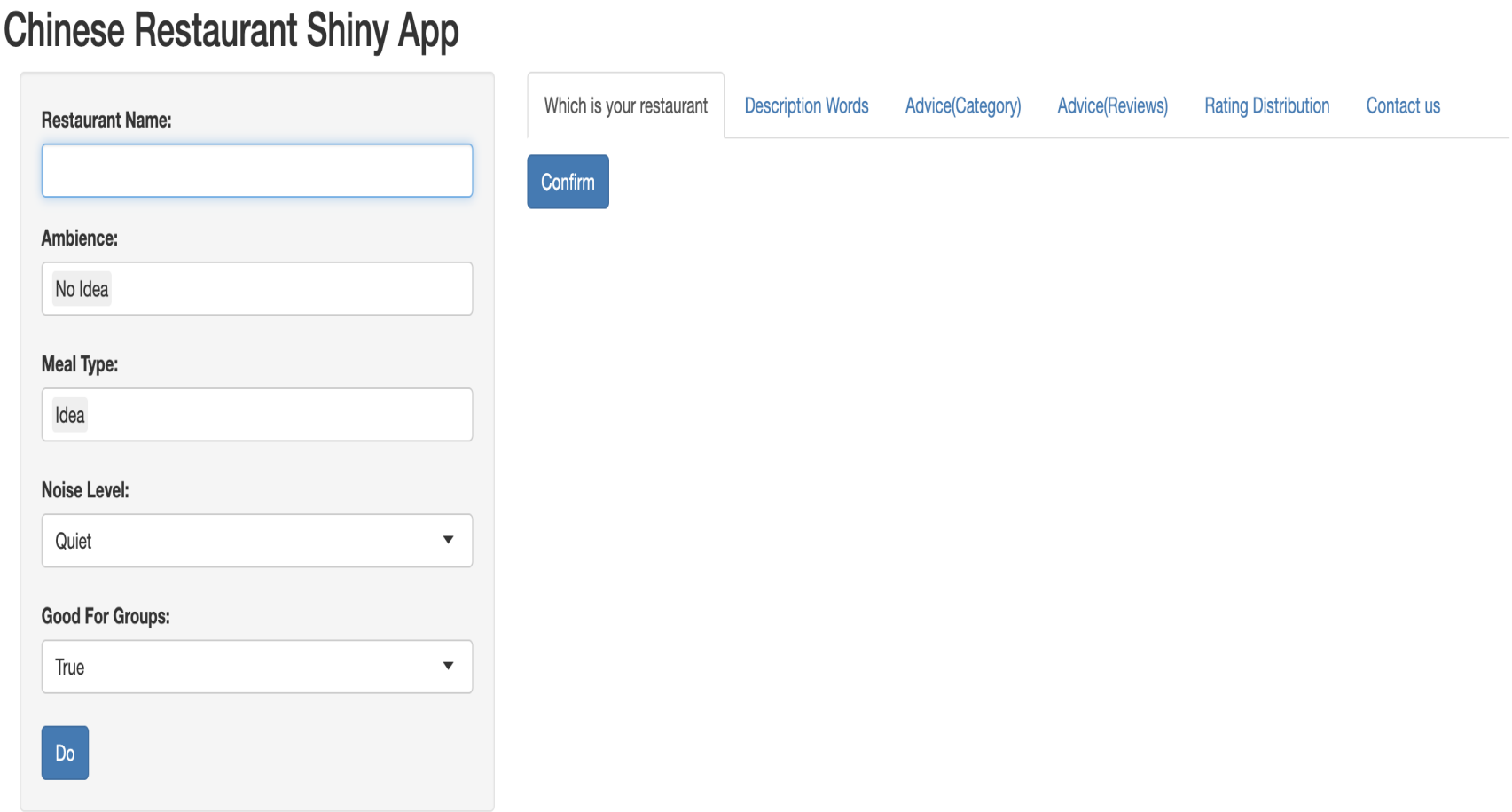
For example, for Emerald Chinese Restaurant. Using Lasso Regression to fit the data, we get the formula:

$$Y = -0.12 * service - 0.33 * price + 0.75 * \text{"}dim\ sum\text{"} - 0.36 * \text{"}fried\ rice\text{"}$$

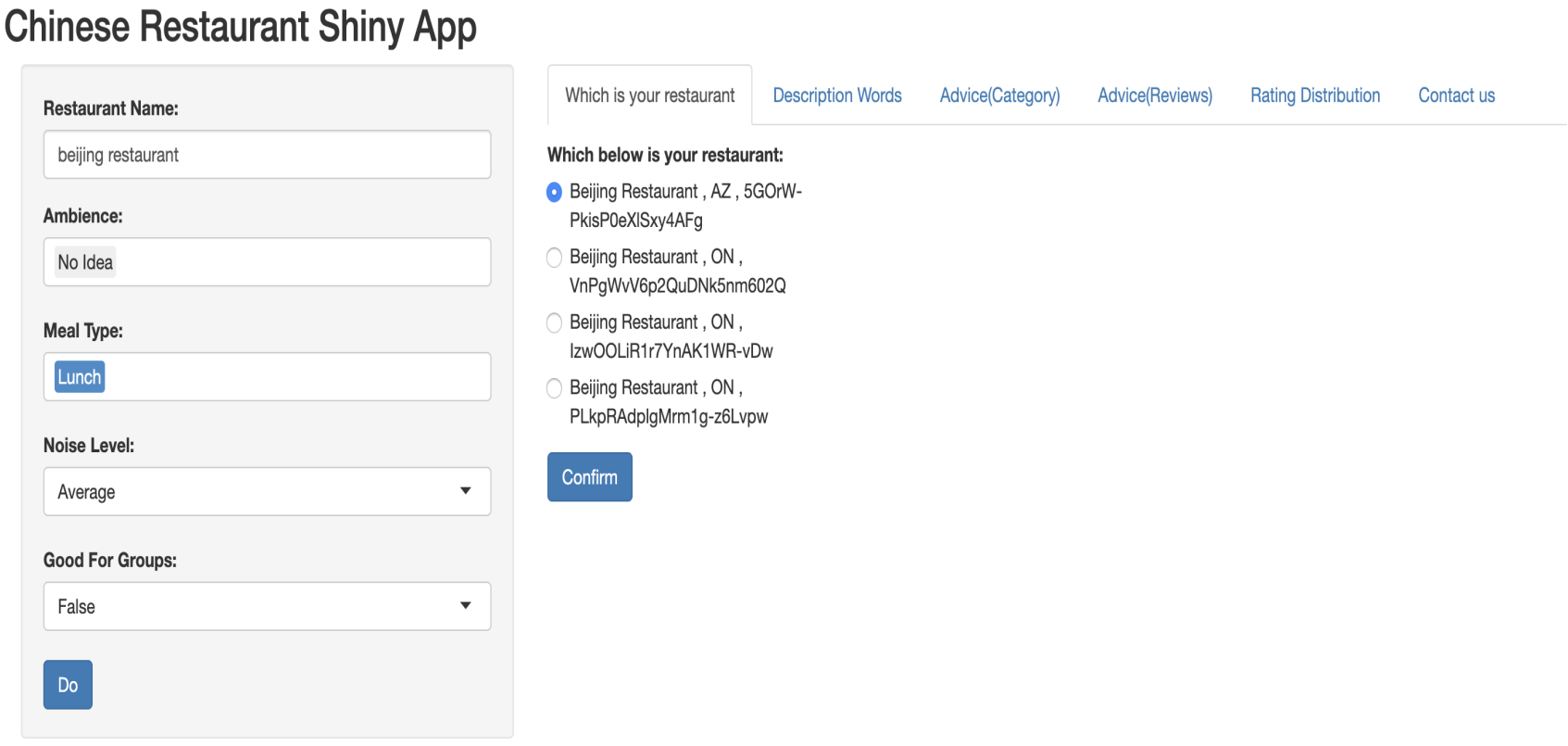
Thus, we can give the advice like this:

- Your ("dim", "sum") is fond of customers, so you should keep the flavor.
- Your ("fried","rice") more or less has a bad effect on your Yelp rating, so you should improve its recipe or remove ("fried","rice") from your menu.
- Your service lowers your Yelp rating, so you should improve your waiters' service level
- Your price lowers your Yelp rating, so you should cut your price.

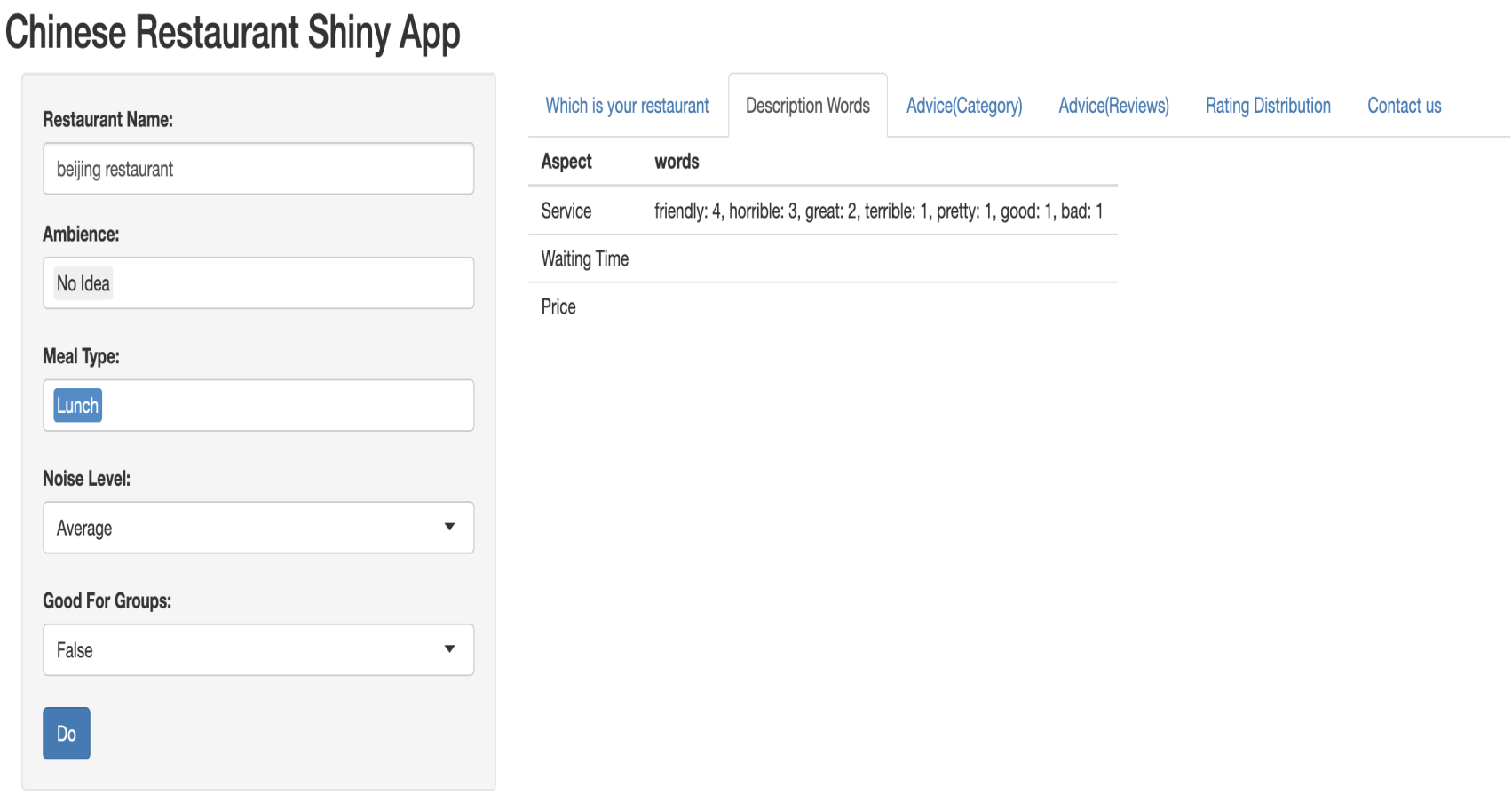
Rshiny


Chinese Restaurant Shiny App
Restaurant Name:
Ambience:
No Idea
Meal Type:
Idea
Noise Level:
Quiet
Good For Groups:
True
Do
Which is your restaurant
Description Words
Advice(Category)
Advice(Reviews)
Rating Distribution
Contact us
Confirm

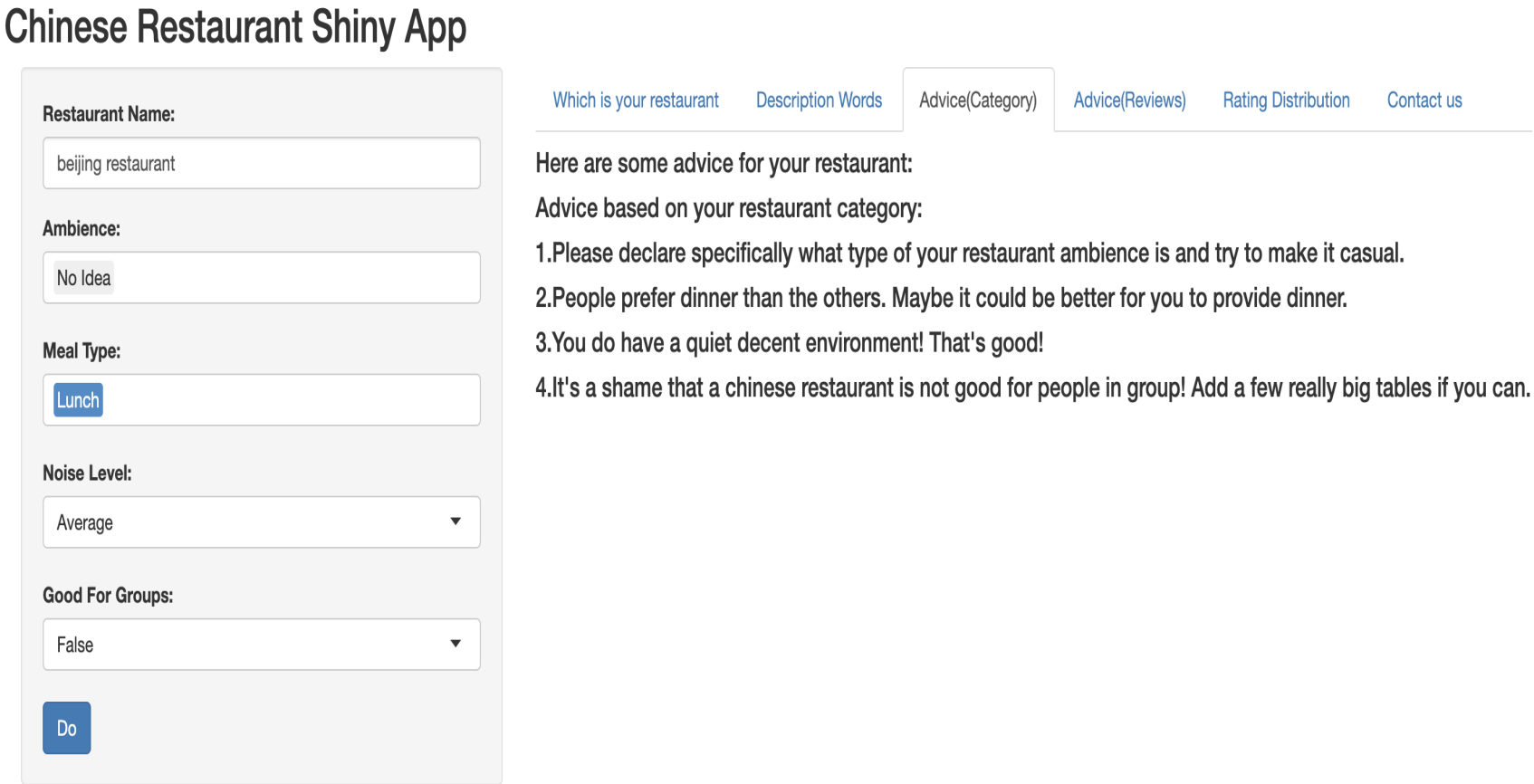
Rshiny


Chinese Restaurant Shiny App
Restaurant Name:
beijing restaurant
Ambience:
No Idea
Meal Type:
Lunch
Noise Level:
Average
Good For Groups:
False
Do
Which is your restaurant
Description Words
Advice(Category)
Advice(Reviews)
Rating Distribution
Contact us
Which below is your restaurant:
Beijing Restaurant , AZ , 5GOrW-PkisP0eXlSxy4AFg
Beijing Restaurant , ON , VnPgWvV6p2QuDNk5nm602Q
Beijing Restaurant , ON , lzwOOLiR1r7YnAK1WR-vDw
Beijing Restaurant , ON , PLkpRAdplgMrm1g-z6Lvpw
Confirm

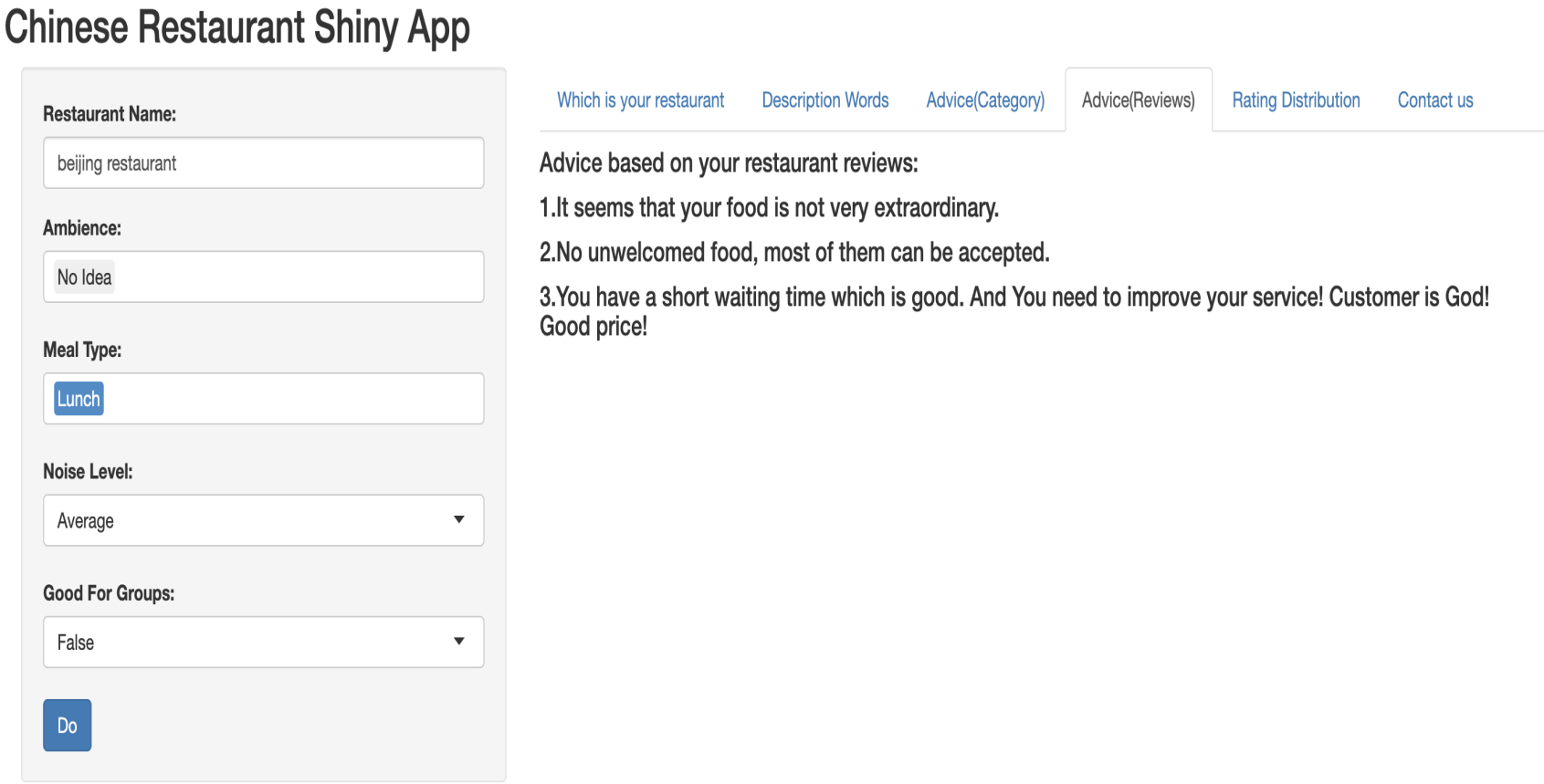
Rshiny


Chinese Restaurant Shiny App
Restaurant Name:
beijing restaurant
Ambience:
No Idea
Meal Type:
Lunch
Noise Level:
Average
Good For Groups:
False
Do
Which is your restaurant
Description Words
Advice(Category)
Advice(Reviews)
Rating Distribution
Contact us
Aspect
words
Service
friendly: 4, horrible: 3, great: 2, terrible: 1, pretty: 1, good: 1, bad: 1
Waiting Time
Price

Rshiny


Chinese Restaurant Shiny App
Restaurant Name:
beijing restaurant
Ambience:
No Idea
Meal Type:
Lunch
Noise Level:
Average
Good For Groups:
False
Do
Which is your restaurant
Description Words
Advice(Category)
Advice(Reviews)
Rating Distribution
Contact us
Here are some advice for your restaurant:
Advice based on your restaurant category:
1.Please declare specifically what type of your restaurant ambience is and try to make it casual.
2.People prefer dinner than the others. Maybe it could be better for you to provide dinner.
3.You do have a quiet decent environment! That's good!
4.It's a shame that a chinese restaurant is not good for people in group! Add a few really big tables if you can.

Rshiny


Chinese Restaurant Shiny App
Restaurant Name:
beijing restaurant
Ambience:
No Idea
Meal Type:
Lunch
Noise Level:
Average
Good For Groups:
False
Do
Which is your restaurant
Description Words
Advice(Category)
Advice(Reviews)
Rating Distribution
Contact us
Advice based on your restaurant reviews:
1.It seems that your food is not very extraordinary.
2.No unwelcomed food, most of them can be accepted.
3.You have a short waiting time which is good. And You need to improve your service! Customer is God! Good price!

Rshiny

Chinese Restaurant Shiny App

Restaurant Name:

beijing restaurant

Ambience:

No Idea

Meal Type:

Lunch

Noise Level:

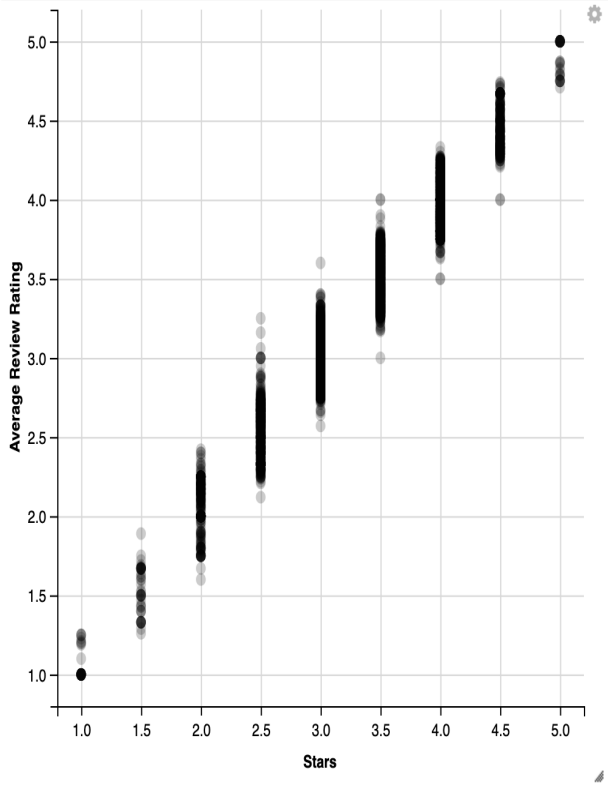
Average

Good For Groups:

False

Do

Which is your restaurant | Description Words | Advice(Category) | Advice(Reviews) | Rating Distribution | Contact us



Your restaurant star rating is 3.5 . And your average review rating is 3.42

Thank You!