



Yelp Data Analyse - Part I

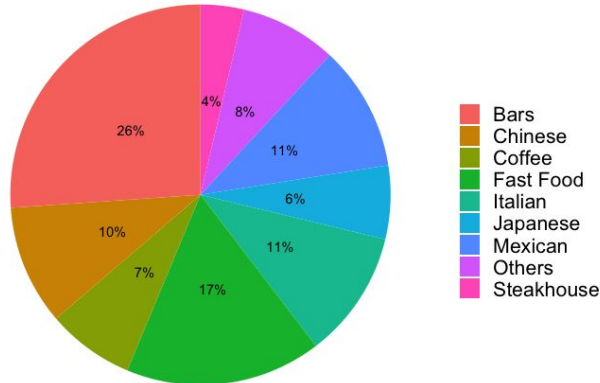
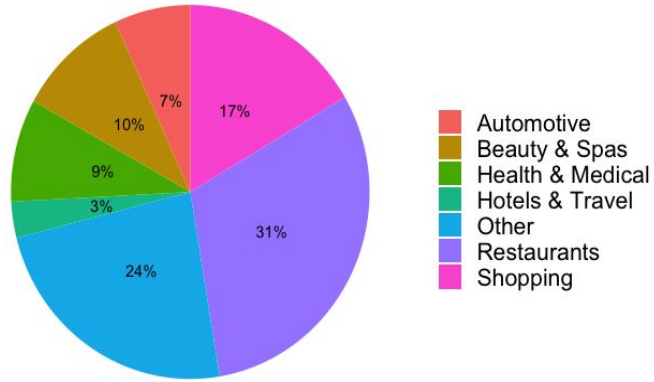
Oct 28, 2019

Xici Luan, Bi Qing Teng

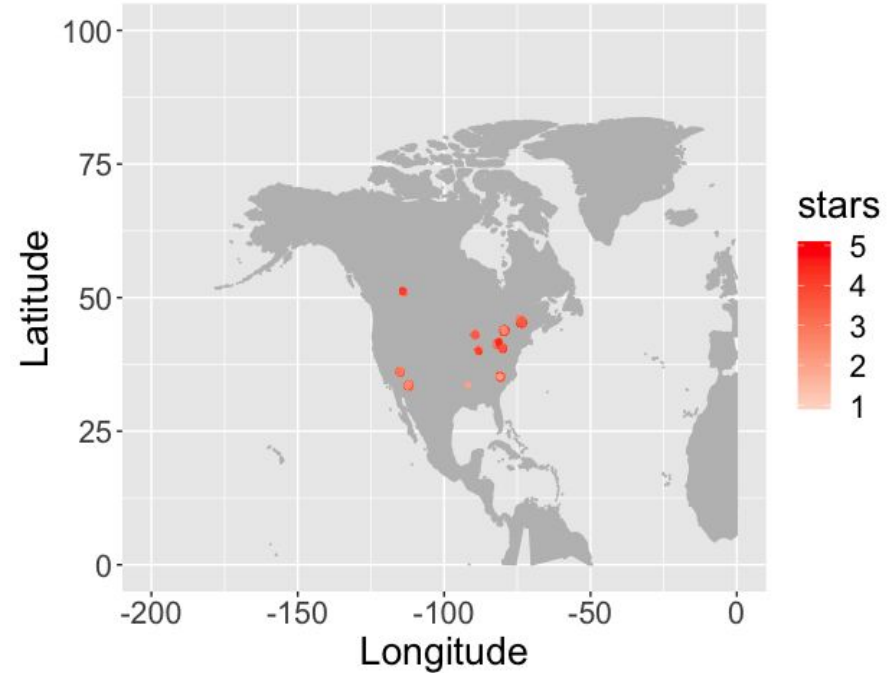
Yuhang Lan, Hongwei Pan



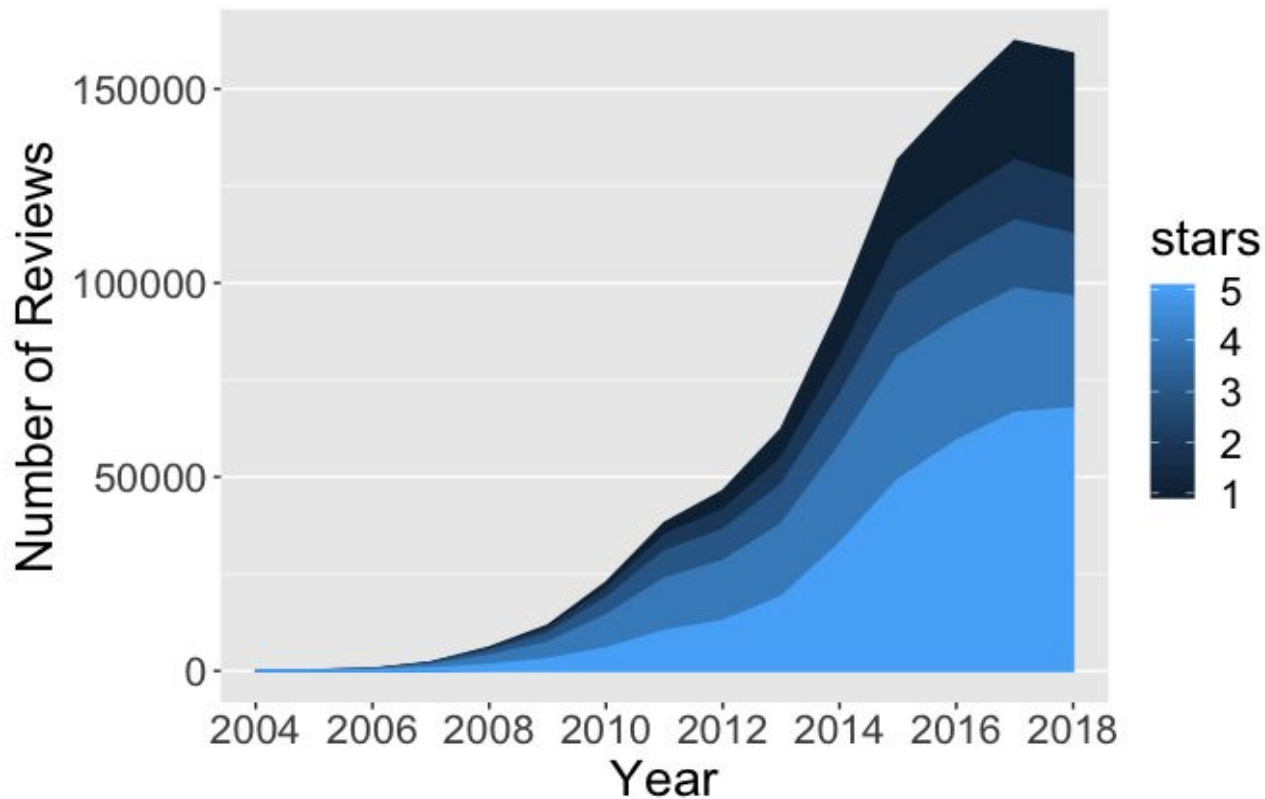
Businesses by Service



Fast Food Restaurants



Number of Review Submissions by Year



Part I

General Advice

1. Attributes
2. Reviews- Word Cloud
3. Reviews- Bar Plot

Part II

Business-Specific

1. Improve Rating Prediction
2. Tailor-made Suggestions
3. APP for better UX

TO BE CONTINUED...

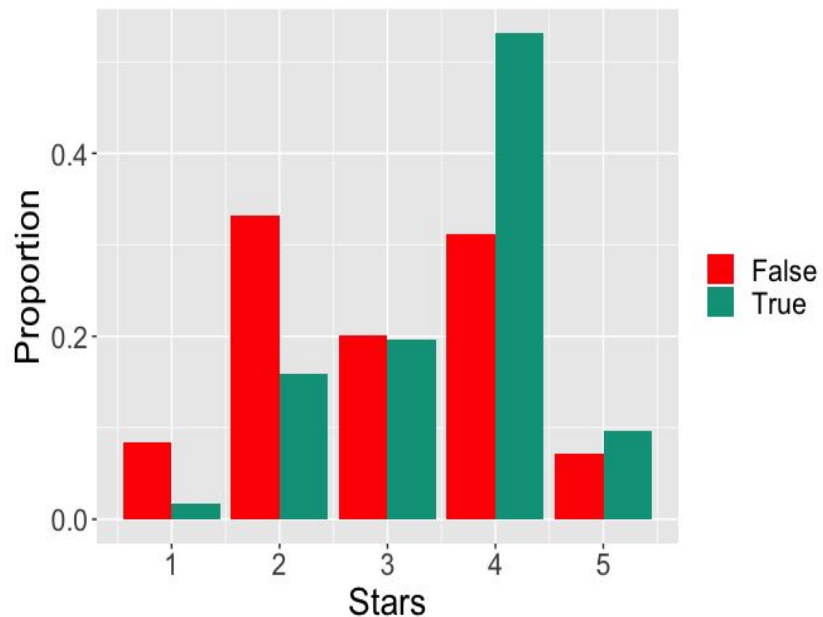
Business Attributes

- Business Parking
- Restaurants Delivery
- Restaurants Reservations
- Restaurants Take Out
- Noise Level
- Price Range
- WiFi

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
BusinessParking	1	450.21	450.21	837.43	$< 2.2 \times 10^{-16}$
RestaurantsDelivery	1	67.57	67.57	125.68	$< 2.2 \times 10^{-16}$
RestaurantsReservations	1	57.31	57.31	106.60	$< 2.2 \times 10^{-16}$
NoiseLevel	3	105.34	35.11	65.31	$< 2.2 \times 10^{-16}$
RestaurantsTakeOut	1	2.03	2.03	3.78	0.0519
RestaurantsPriceRange2	3	40.99	13.66	25.41	2.299×10^{-16}
WiFi	2	4.29	2.14	3.99	0.0186
Residuals	8827	4745.48	0.54		

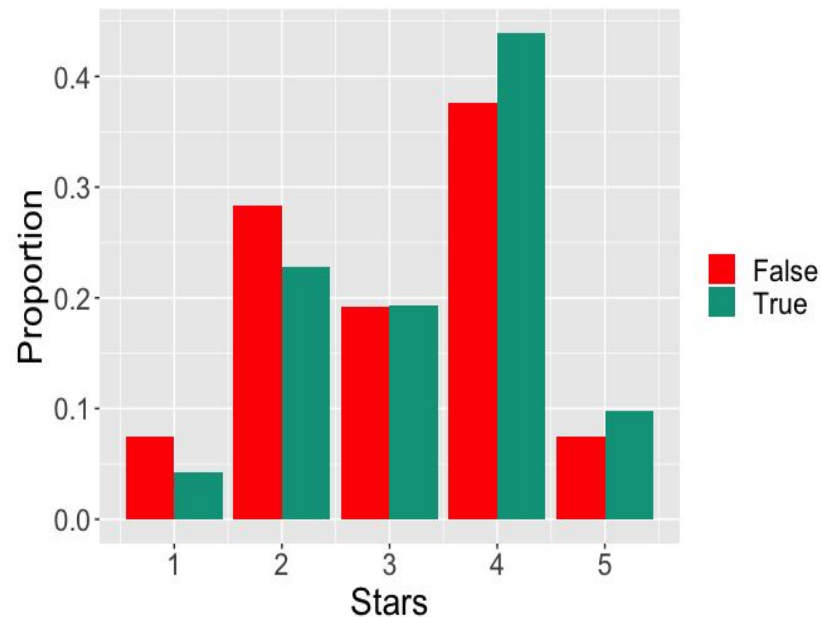
Table 1: ANOVA table obtained by regressing rating on business attributes.

Business Parking



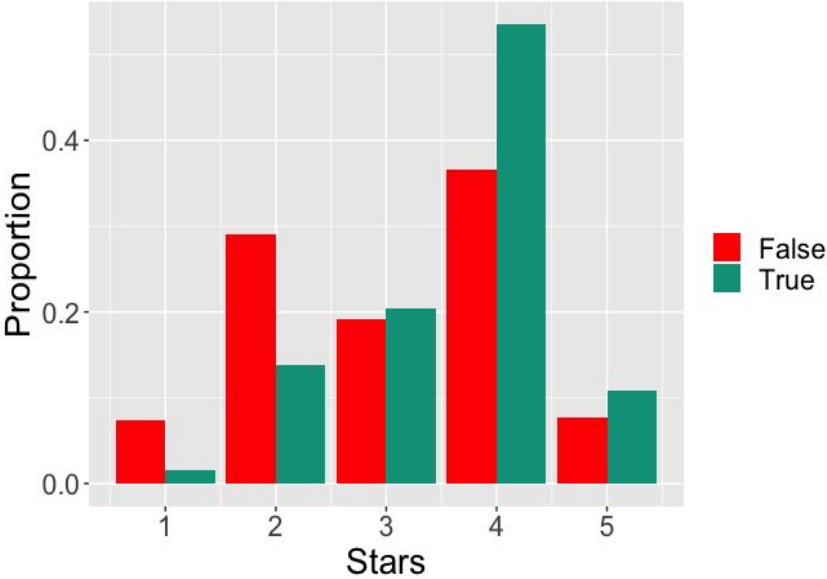
Business Parking	Average Rating
False	2.95
True	3.53

Restaurants Delivery



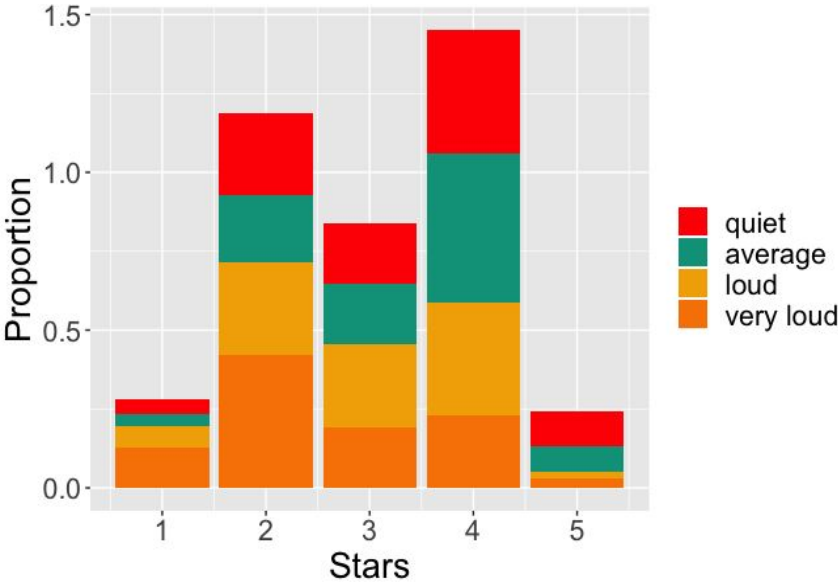
Restaurants Delivery	Average Rating
False	3.09
True	3.32

Restaurants Reservations



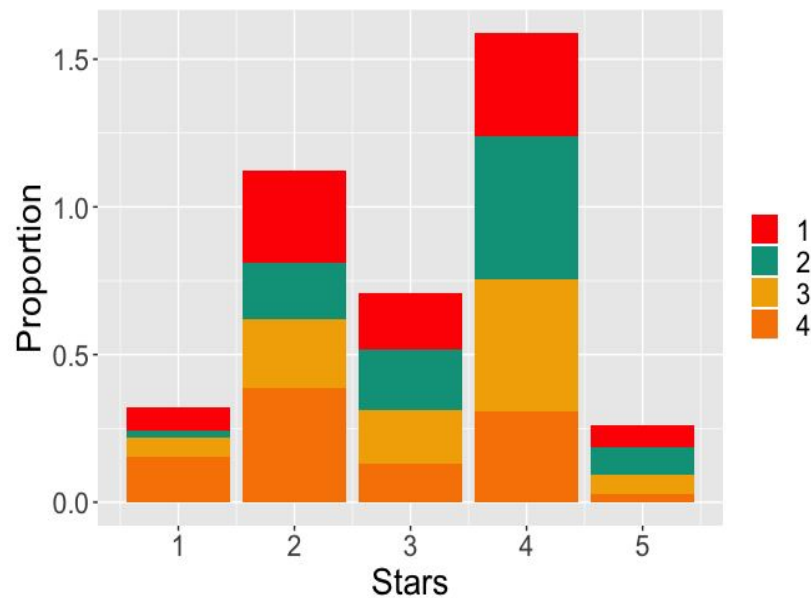
Restaurants Reservations	Average Rating
False	3.08
True	3.58

Noise Level



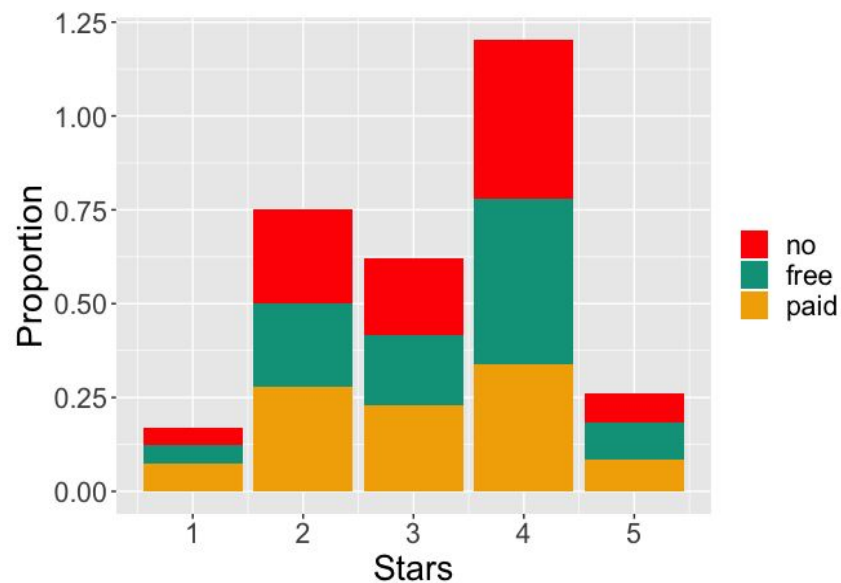
Noise Level	Average Rating
quiet	3.25
average	3.35
loud	2.97
very loud	2.62

Price Range



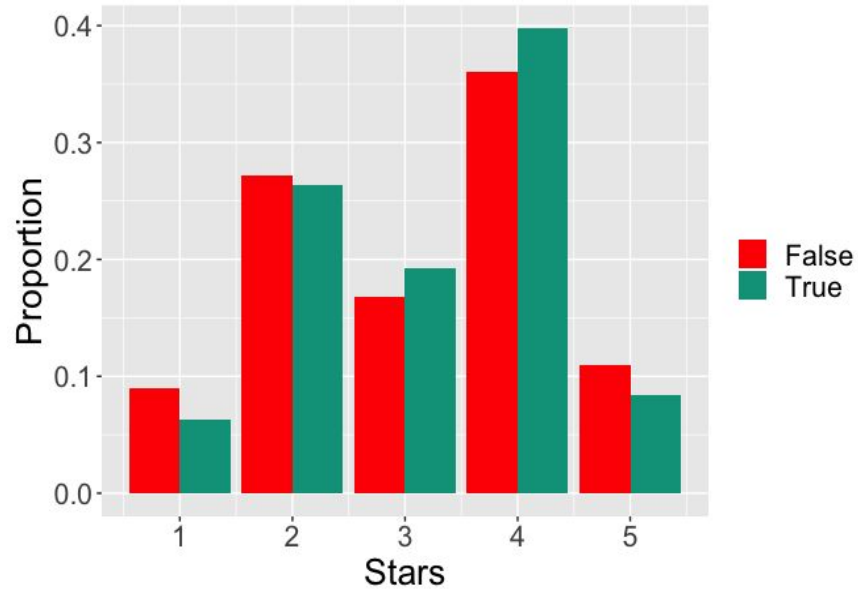
Price Range	Average Rating
1	3.02
2	3.42
3	3.22
4	2.67

WiFi



WiFi	Average Rating
no	3.23
free	3.32
paid	3.08

Restaurants Take Out



Restaurants Take Out	Average Rating
False	3.13
True	3.18

TIPS

- Business Parking
- Restaurants Delivery
- Restaurants Reservations
- Average Noise Level
- Price Range 2 -3
- Free WiFi

Reviews - Word Cloud

Feature Extraction from Reviews

From BoW (uni-gram)

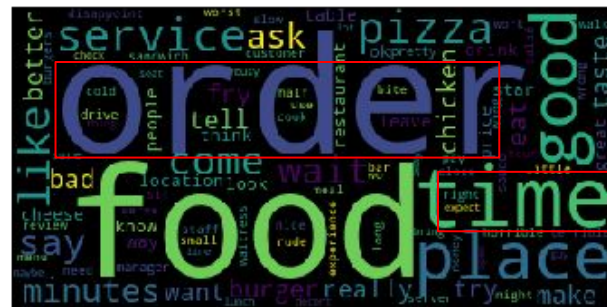


Positive

From TF-IDF (uni-gram)



Negative



Feature Extraction from Reviews

Word Cloud from Bigram TF-IDF:



Positive Comments



Negative Comments

Feature Extraction from Reviews

Word Cloud from Trigram TF-IDF:



Positive Comments

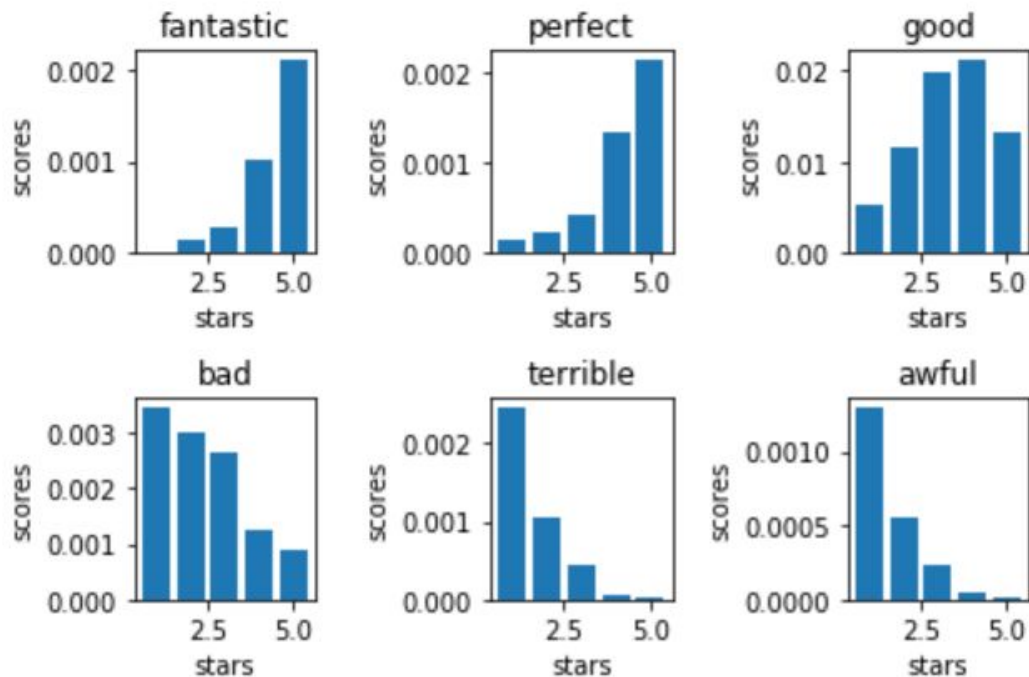


Negative Comments

Reviews - Bar Plot

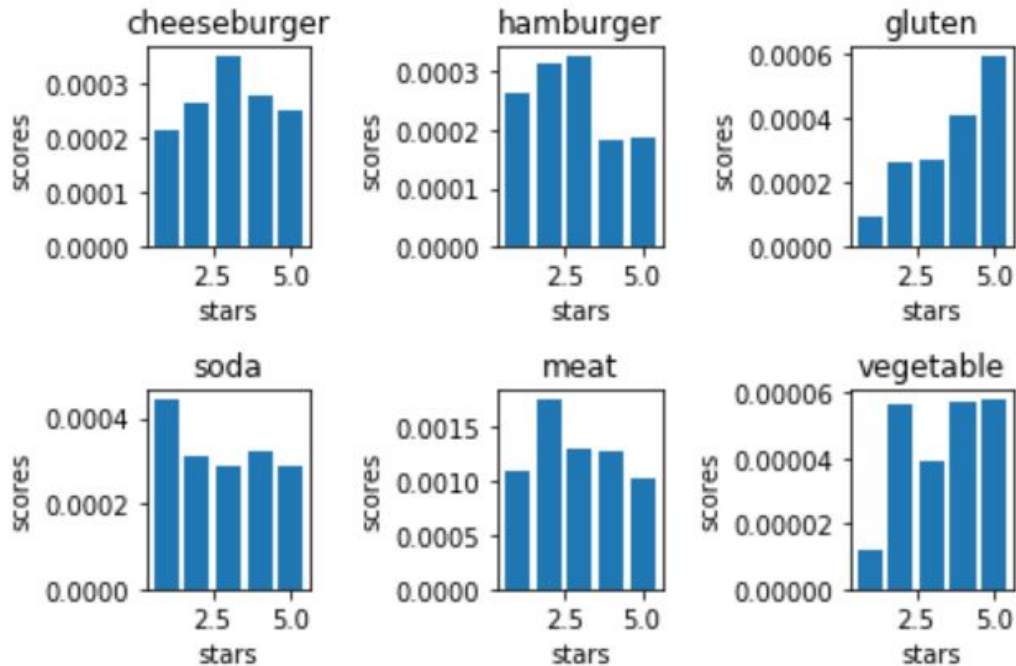
- Attitude
- Food
- Service
- Brand

Attitude



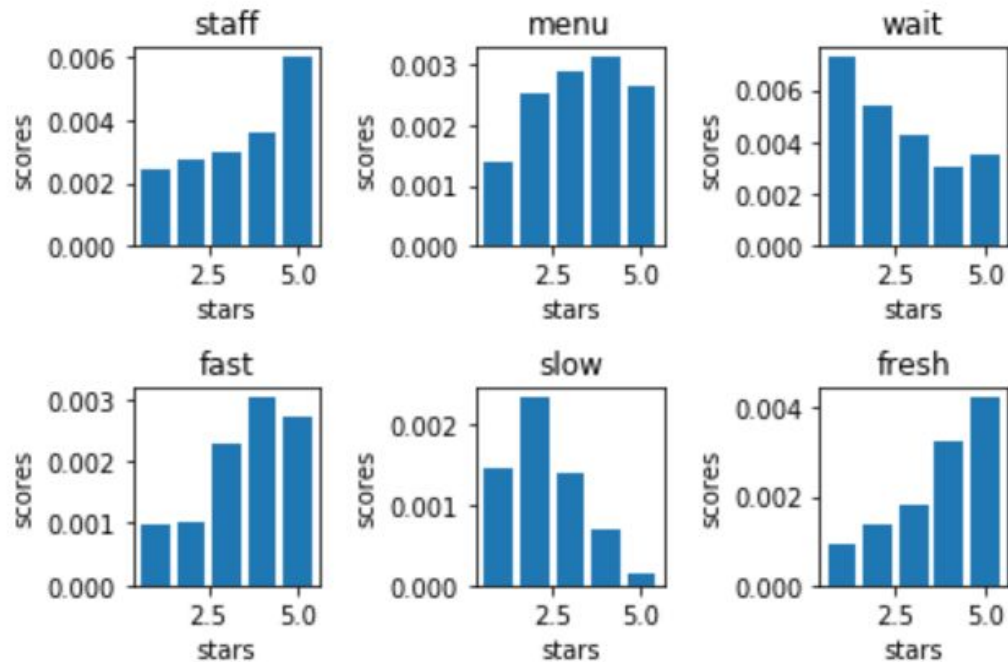
We adopted the right algorithm

Food



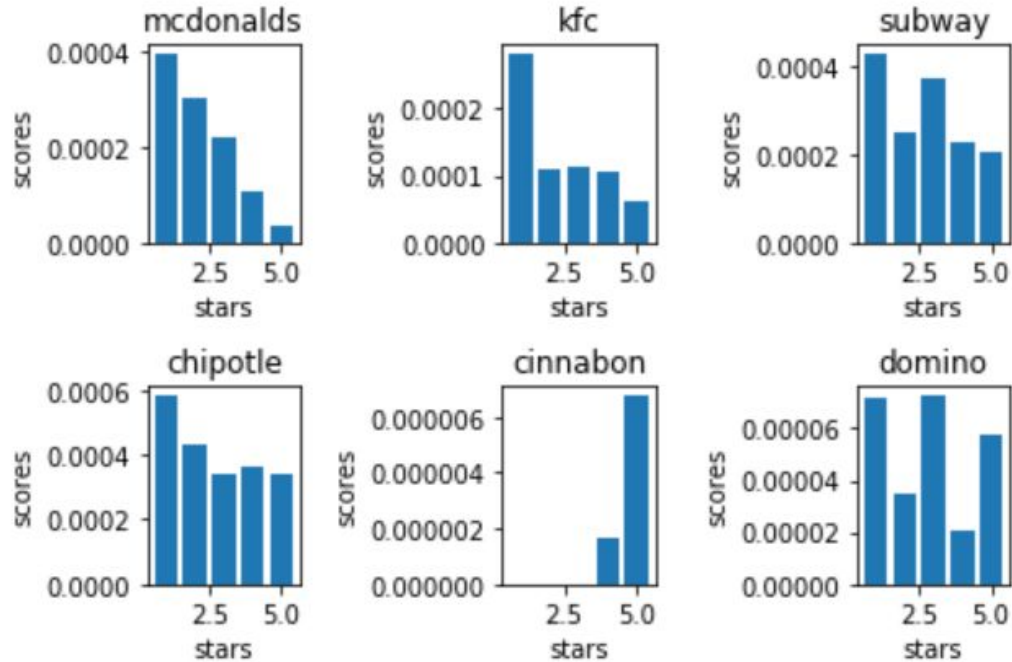
Remember to add gluten free food in your menu

Service



It might always be right to serve fast and fresh food

Brand



Be careful when choosing brand for franchise

not_neg today when_neg
chang toast back_neg
use
eat_neg
sweet
size eat_neg
lot
spot
hope
larg
long
absolut

do_neg
half
walk
order_neg
excel
were_neg run
instead again_neg
going_neg
fresh
fun
select
just_might
just_neg
front
● ● ●

me_neg
star
thank
still

talk
Day ask plu
felt us_neg
howev disappoint
mayb
leav manag
up_neg
highli

ok bad job
ok decent minut saw
seem
word okay
tell later
cold
noth
leave
our_neg
fantast

although
perfectli
will_neg
town
he
season
somemonth
sit line
least put
amaz
worth
arriv
bite late
computer
water
awesom
ice
service_neg
share

cool favorit care last
notder red
them_neg
may
perfect
their_neg
almost light
or_neg
total
how_neg
bbq
at_neg
away
stay