



# Module 3

Group 1 : Yunqing Shao  
Xu Zou  
Cuizhuo Lu





# Content

---



Data  
Processing



Exploratory  
Data Analysis



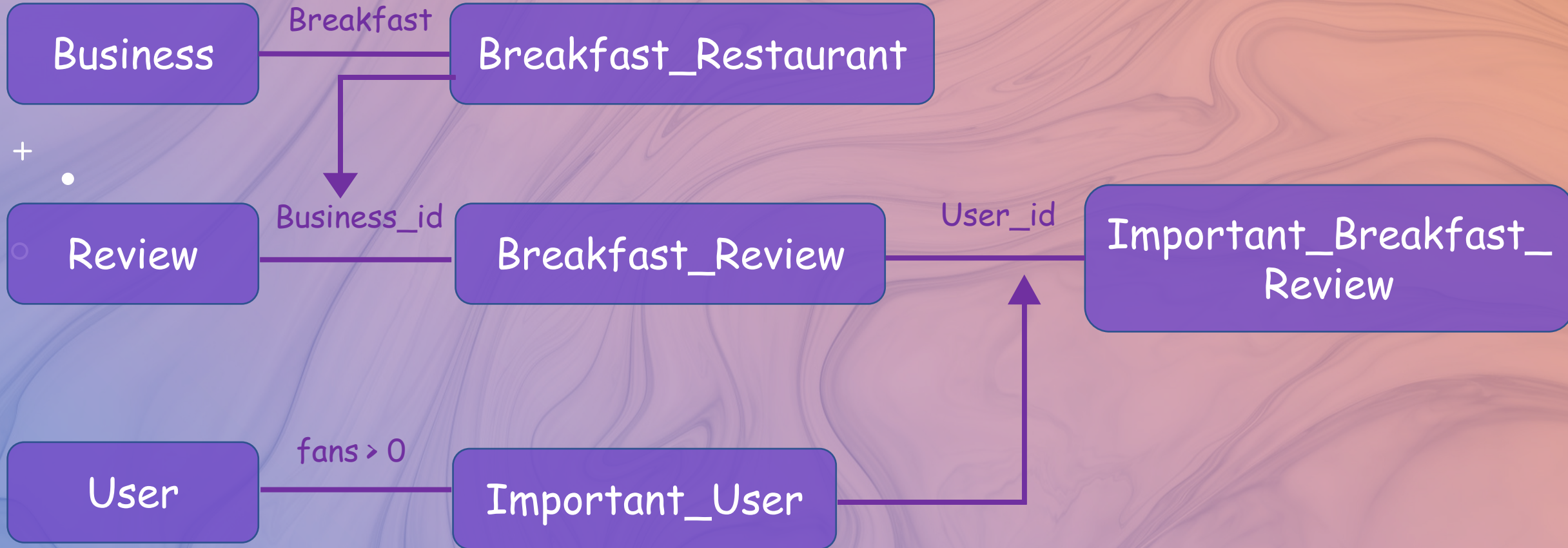
Model  
Building



Recommendation  
For Business

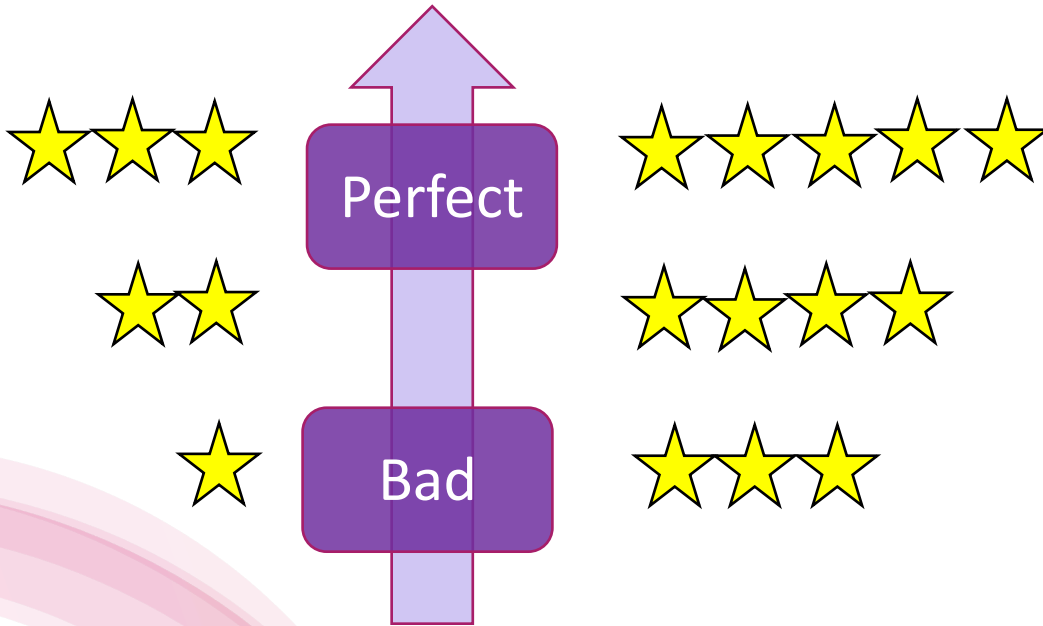


# Data Processing & EDA



# Standardize the Ratings

Why we need standardization ?



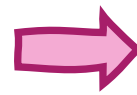
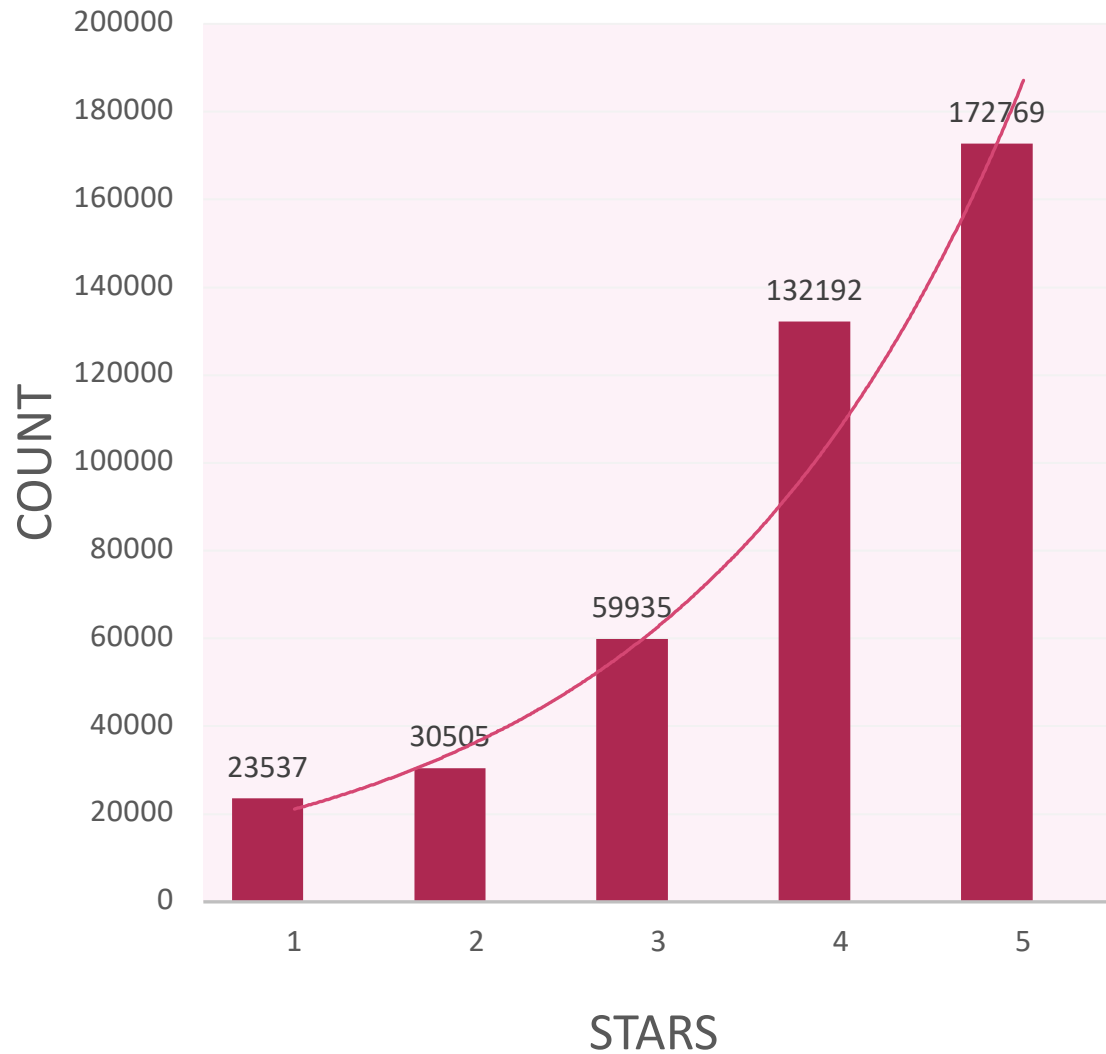
How to scale the rating ?

- Get the mean and variance of users' historical ratings

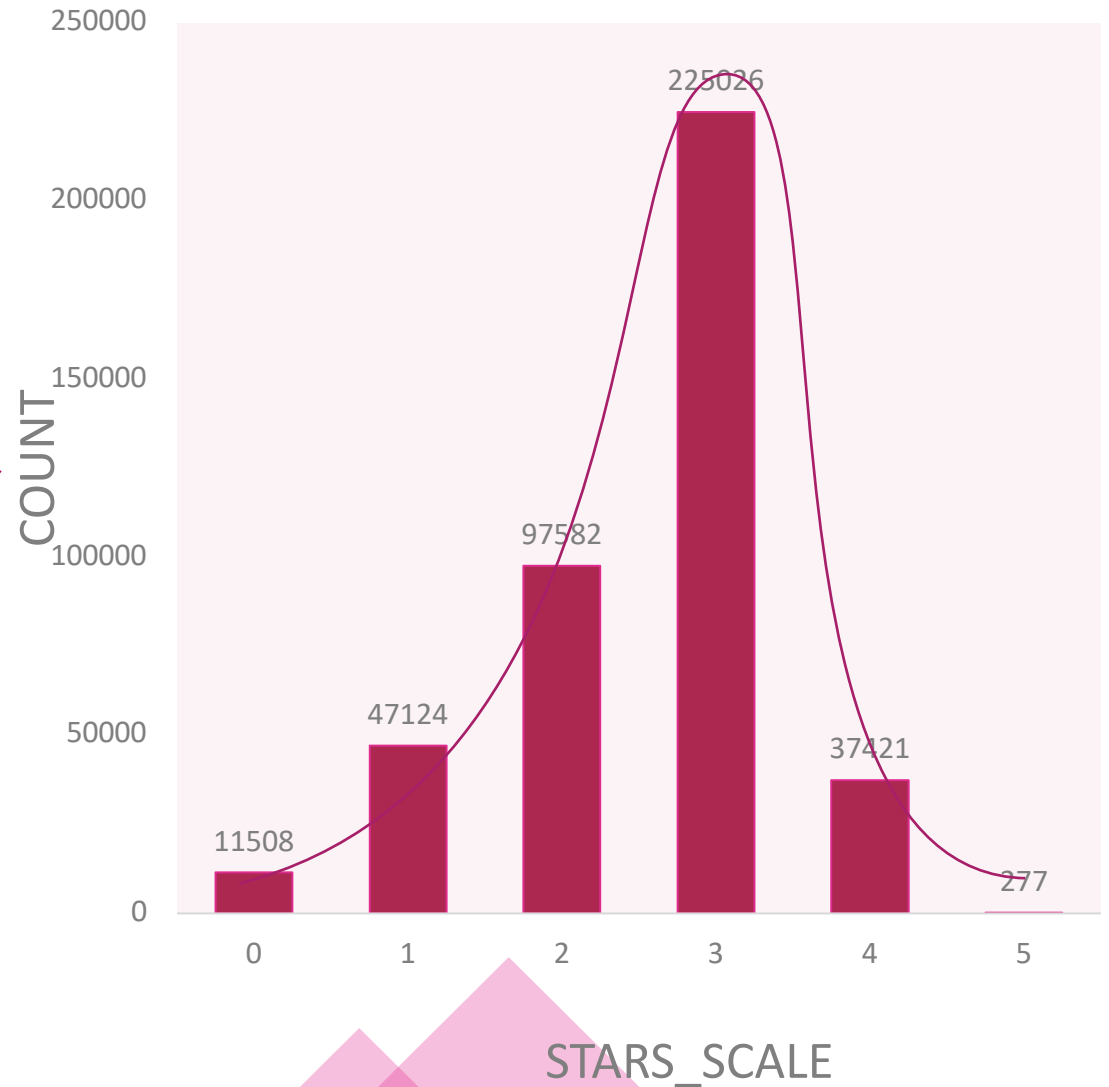
- $stars\ adjust_{user-i} = \frac{stars_{user-i} - mean_i}{sd_i}$

# Change of Rating distribution

Restaurant



Restaurant

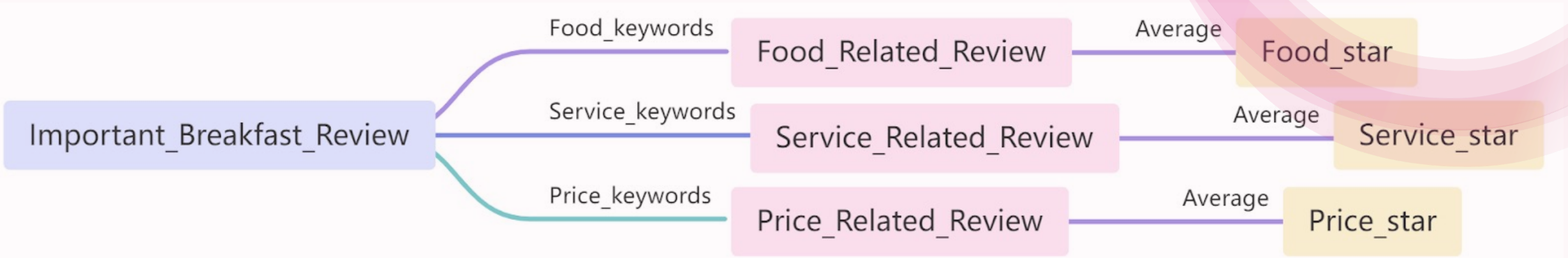


## Text Analysis Procedure:



## Model Building

### Form 3 kinds of key words:



word categories	key words
Food_keywords	food, delicious, coffee, cheese, fresh, meal, chicken, eggs, sandwich, tasty, drinks, sauce, salad...
Service_keywords	service, friendly, staff, parking...
Price_keywords	price, expensive, cheap, affordable...

Take Taco Bell as an example :



Select a state

IN

Select a city:

Carmel

Select a name:

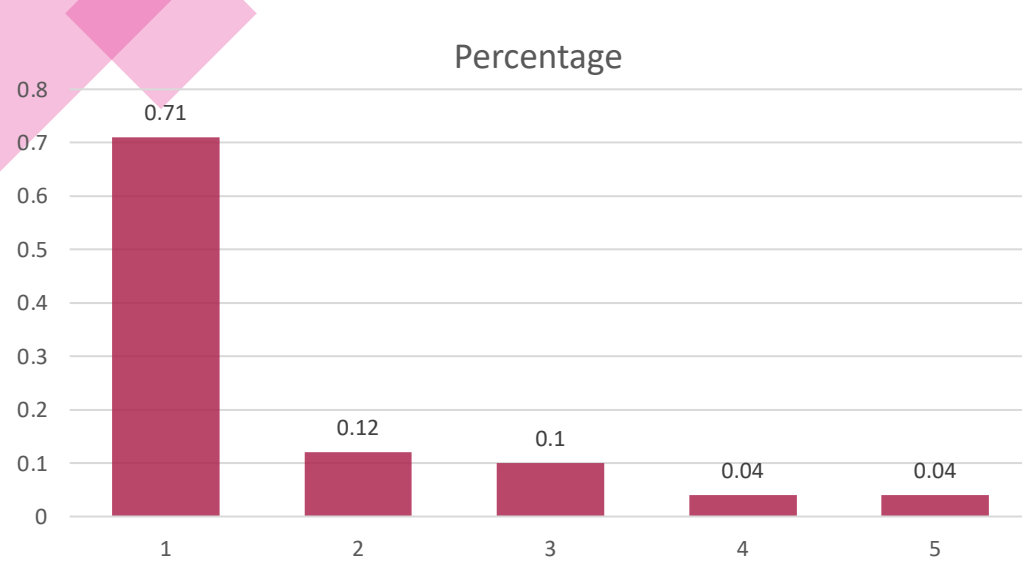
Taco Bell

Select a business ID:

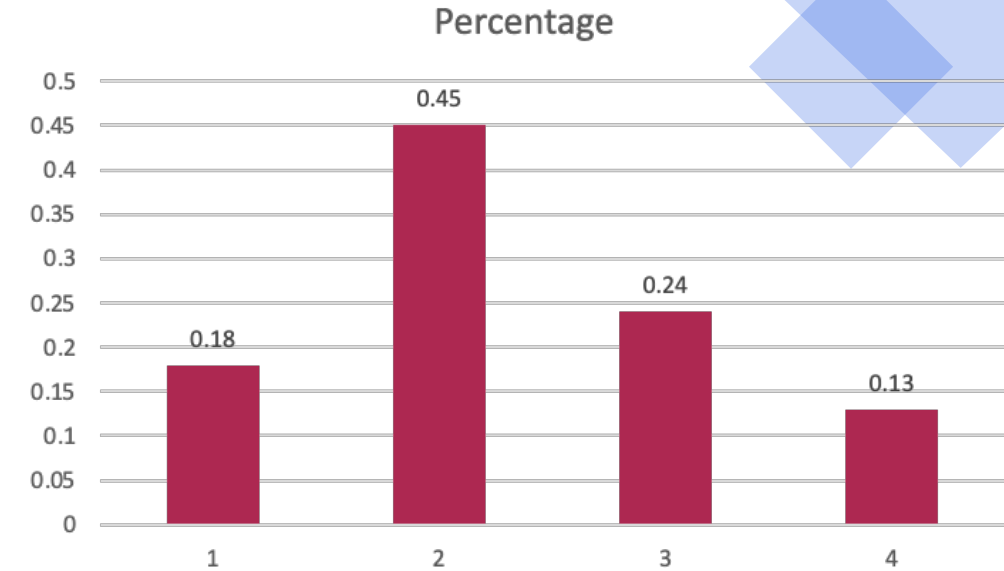
S5a0iHy8KDVmPFnZ1Sze-g

Recommendation  
For Business

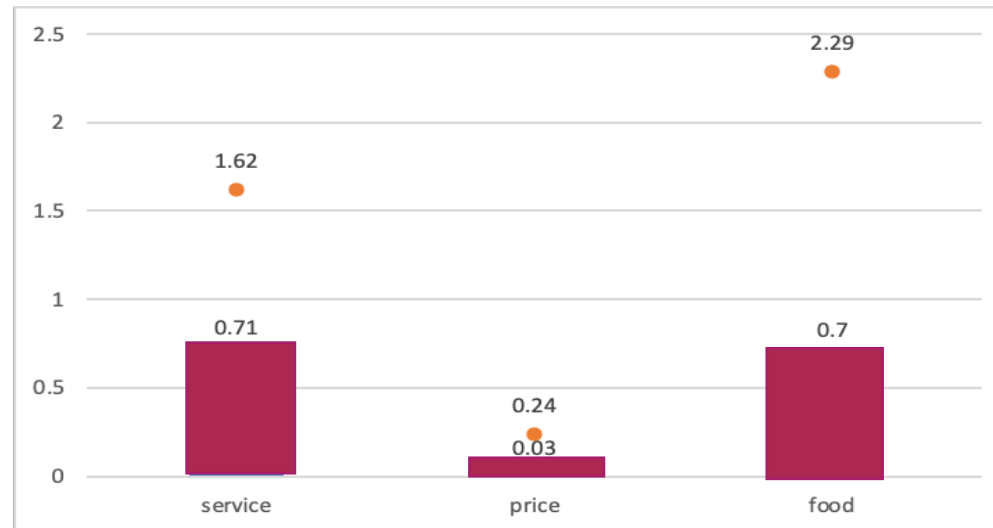
# Recommendation for business



Original\_stars count

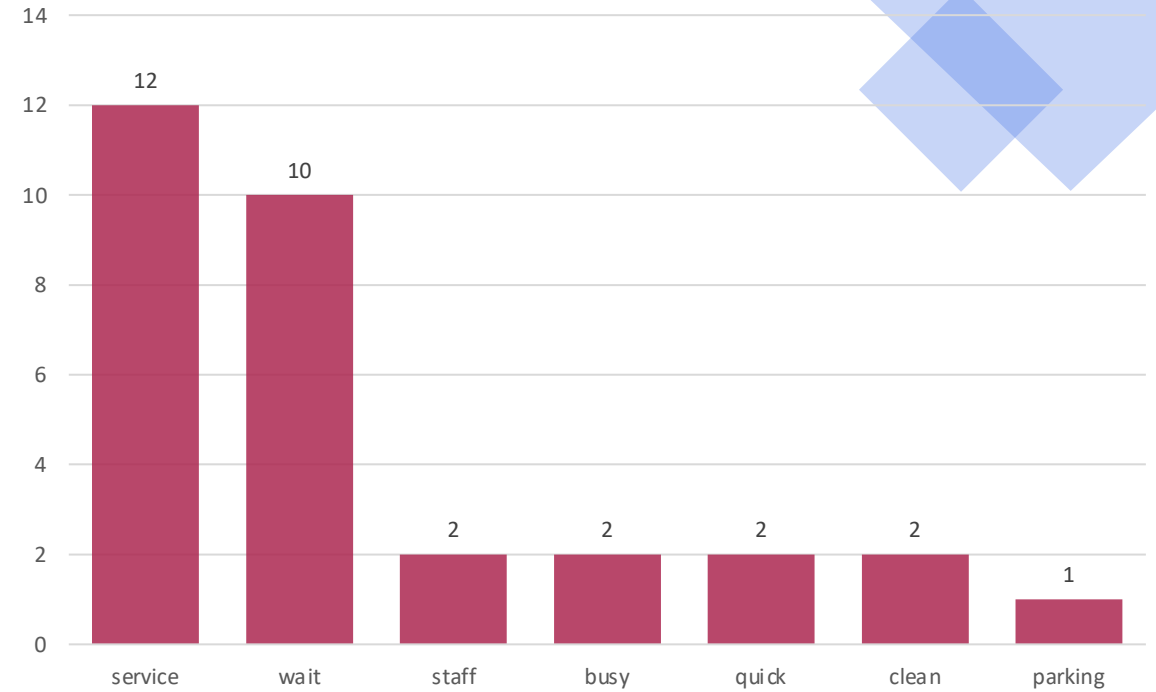
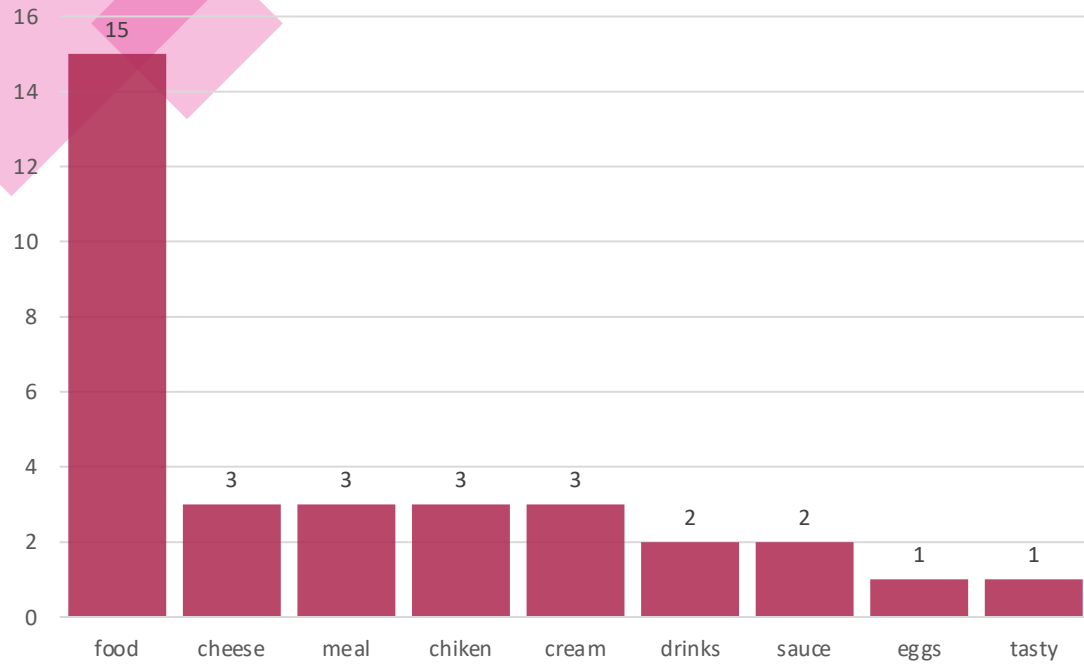


Scaled\_stars count



Stars on Three Aspects





### Further Analysis on Food and Service

- Taco Bell should care more about the quality of cheese and chicken.
- Taco Bell should cook faster to avoid users' long waiting.
- The staff of Taco Bell should be more friendly.
- Taco Bell should care more about the price.

+  
•  
○

Thank you !

+  
•  
○