

# Nightlife in Las Vegas

Group 6 - Junxia Zhu, Yiqun Jiang, Yingjing Jiang

# Outline

- Goal
- Data filtering
- Data analysis

Internal factors: attributes from business

External factors: factors from reviews

- Conclusion
- Future work

# Goal

- Explore nightlife in Las Vegas
- Provide insights to help improve ratings

# Data Filtering

- **Original train data:** business 154,606 reviews 5,364,626
- **Filter rules:** categories: nightlife, bars city: Las Vegas
- **Final train data:** business 1,201 reviews 265,847

# Data Analysis Outline

- **Internal factors:** attributes from business  
data cleaning  
important attributes  
attributes analysis
- **External factors:** factors from reviews  
data cleaning  
high-frequency words  
words analysis

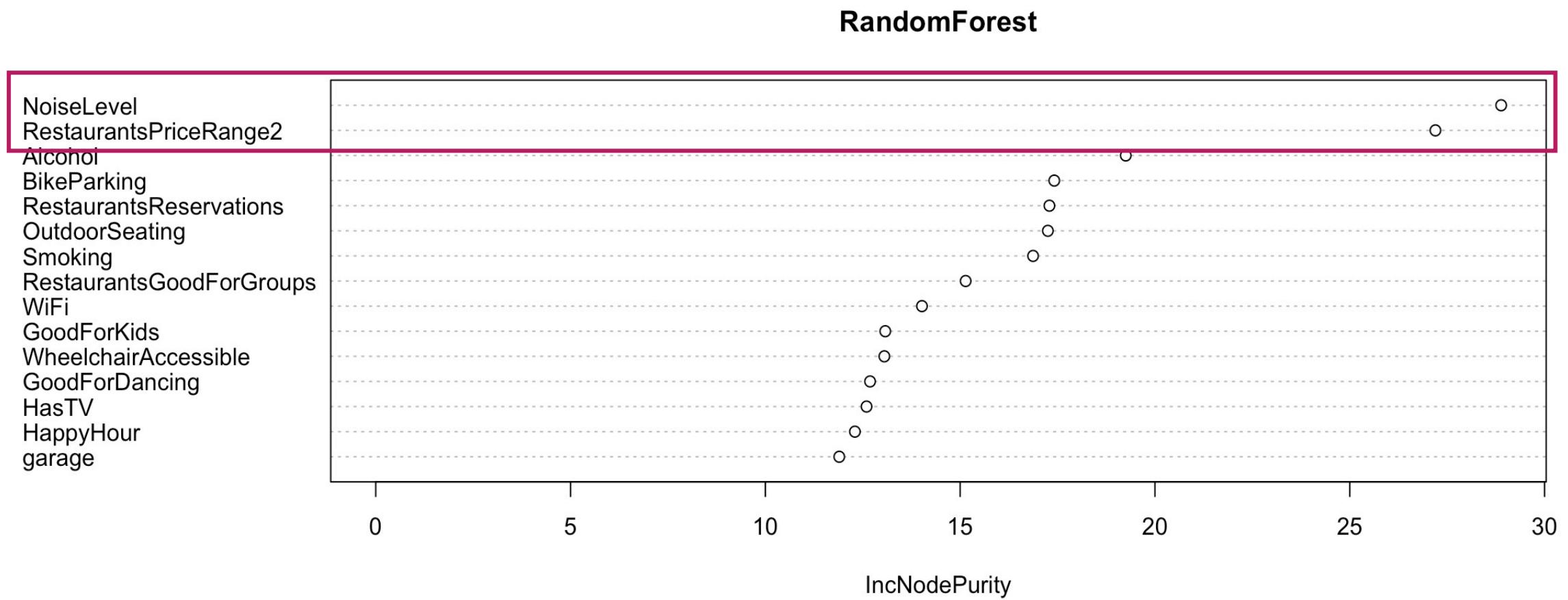
# Internal Factors

- Data cleaning
  - all attributes
  - average star for all business

business_id	attributes_1	...	attributes_n	average_star
1				1.5
...	unknown			...
m				4.3

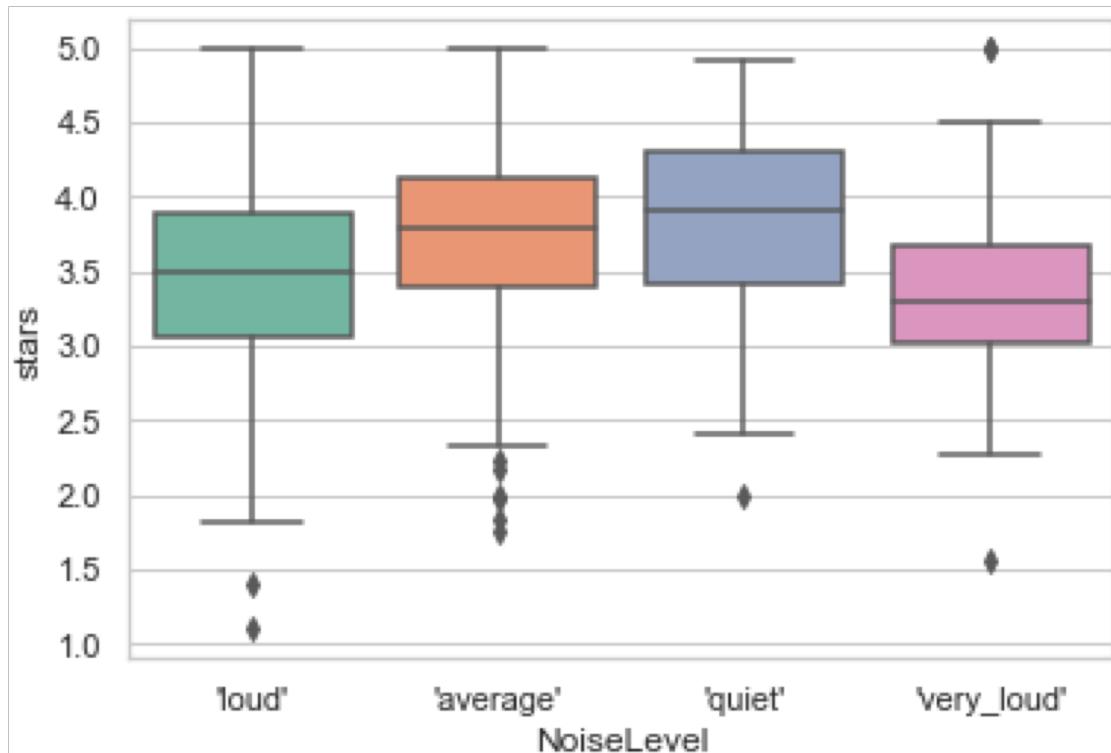
# Internal Factors

- Important attributes: random forest importance score



# Attributes Analysis

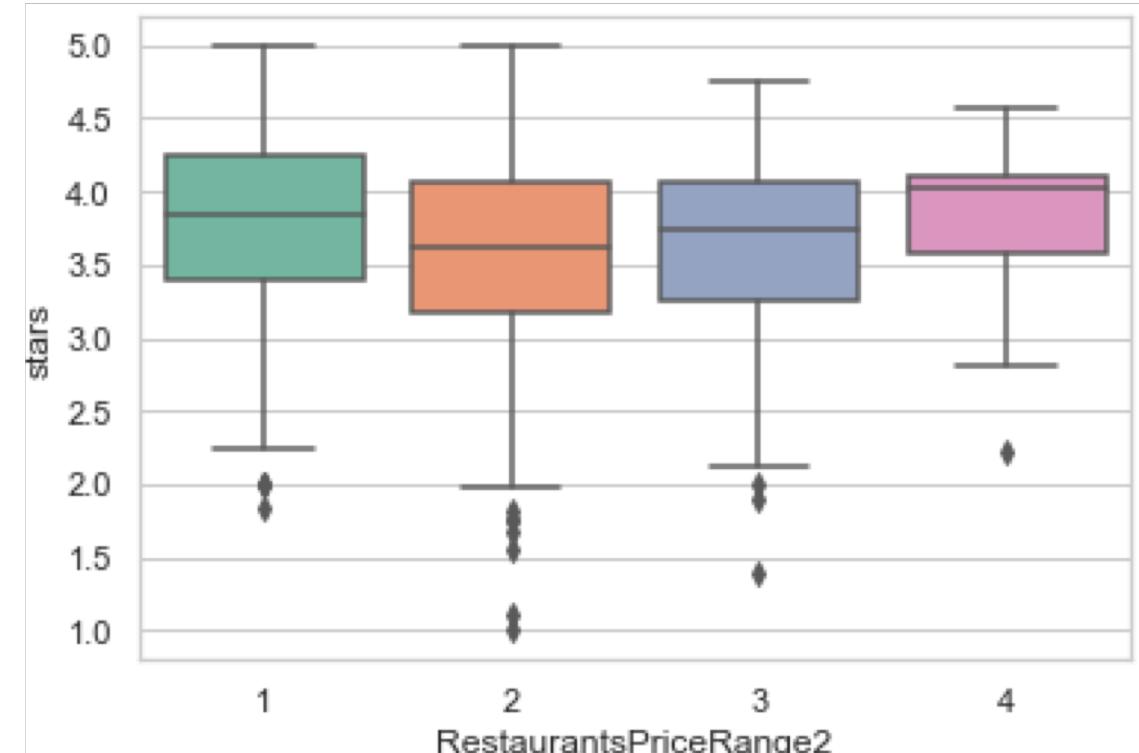
## NoiseLevel



One-way  
ANOVA

p-value:  $5 \times 10^{-10}$

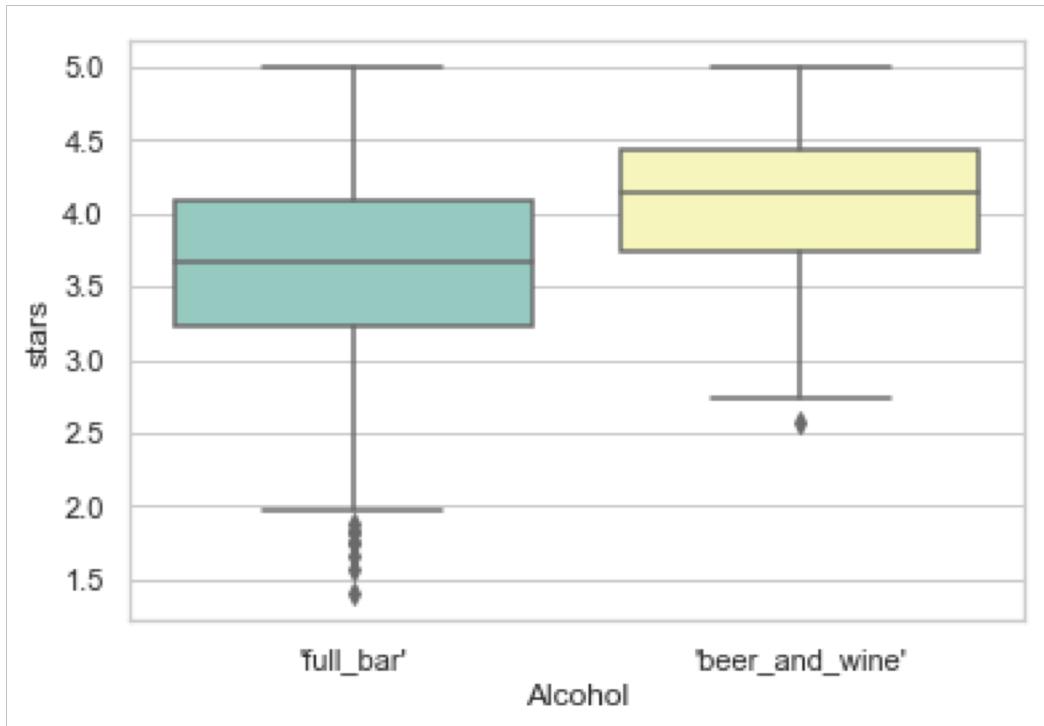
## RestaurantsPriceRange



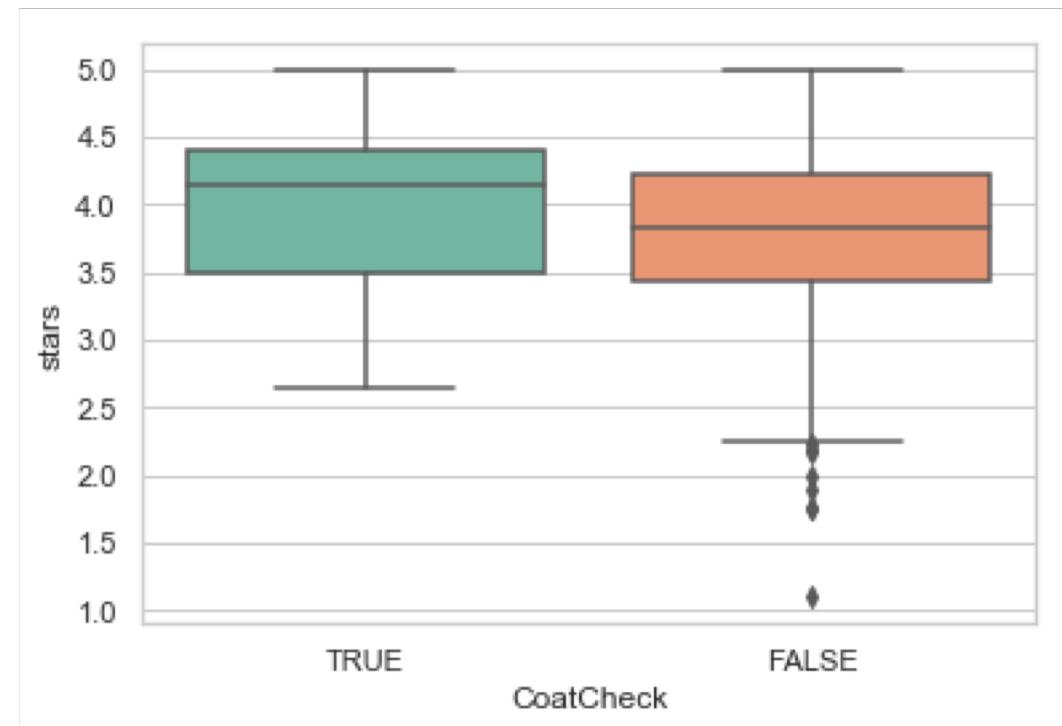
p-value:  $1 \times 10^{-3}$

# Attributes Analysis

## Alcohol



## CoatCheck



# External Factors

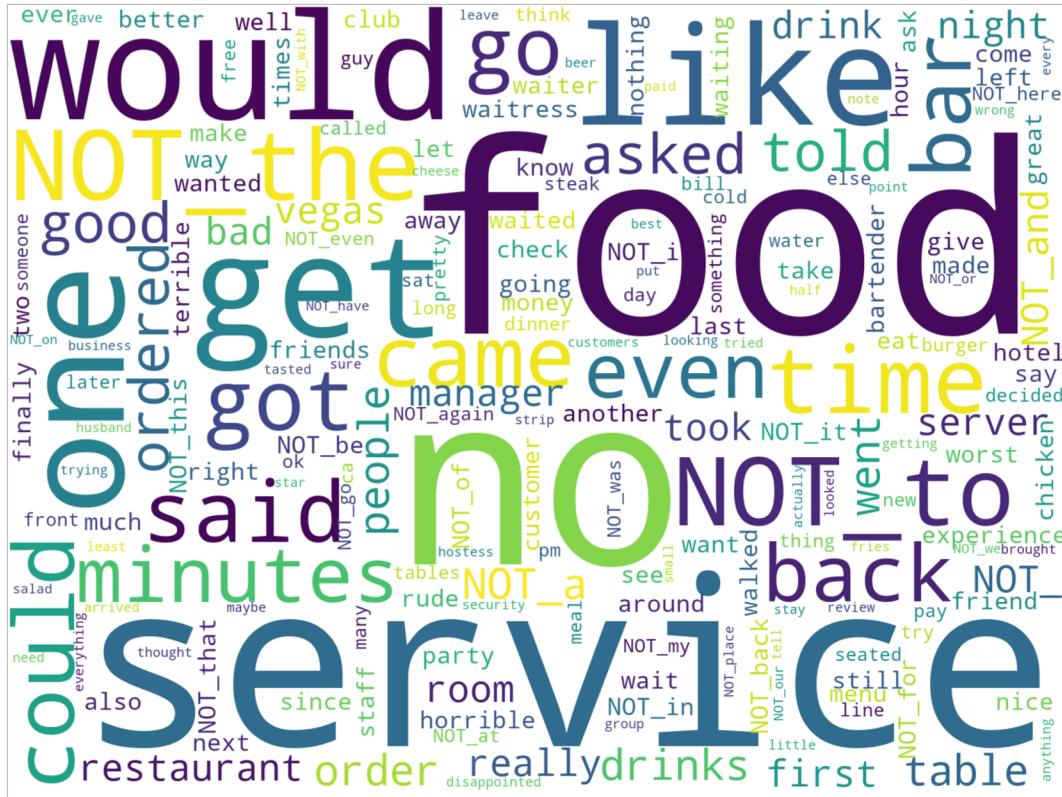
- Data cleaning
  - tokenize:
    1. paragraph → sentences
    2. negative words: n't → not → add NOT\_ to following words until next punctuation.  
eg: I don't think I will go back again.  
→ I do NOT\_think NOT\_I NOT\_will NOT\_go NOT\_back NOT\_again.
    3. sentences → words
  - delete punctuation
  - delete stopwords

high-frequency words  
for each star level

select words for  
different aspects

# External Factors

- ## • star 1



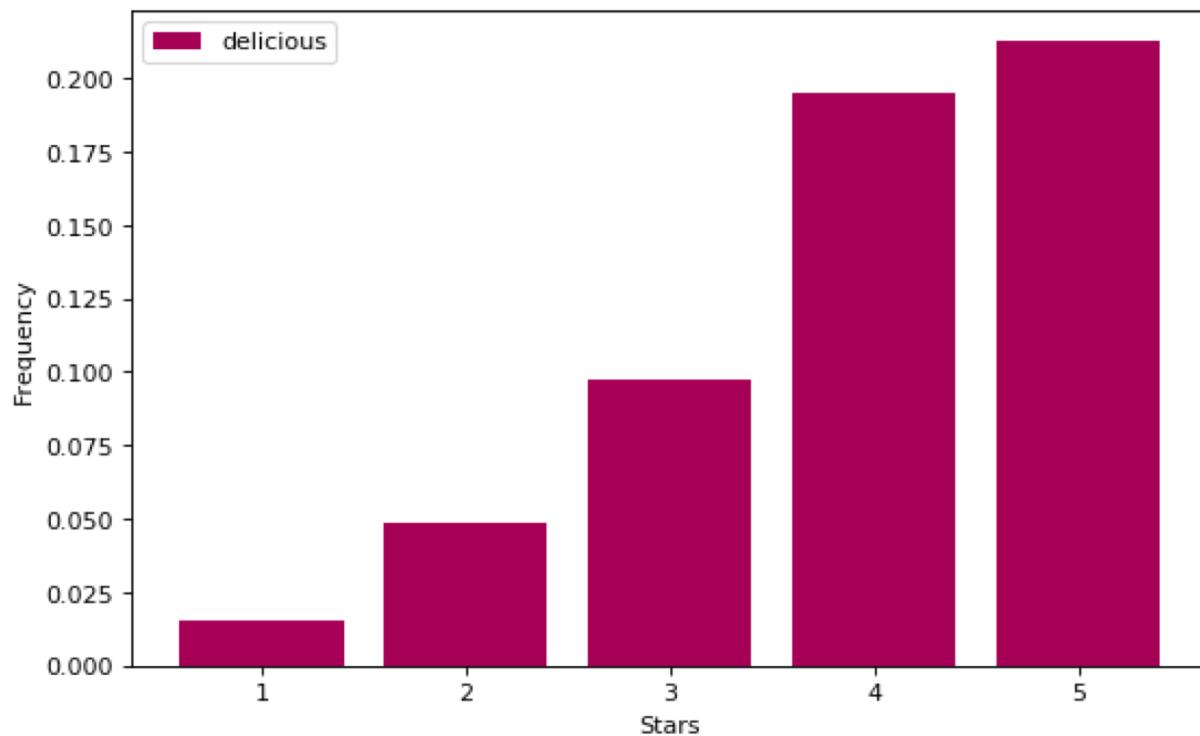
- star 5



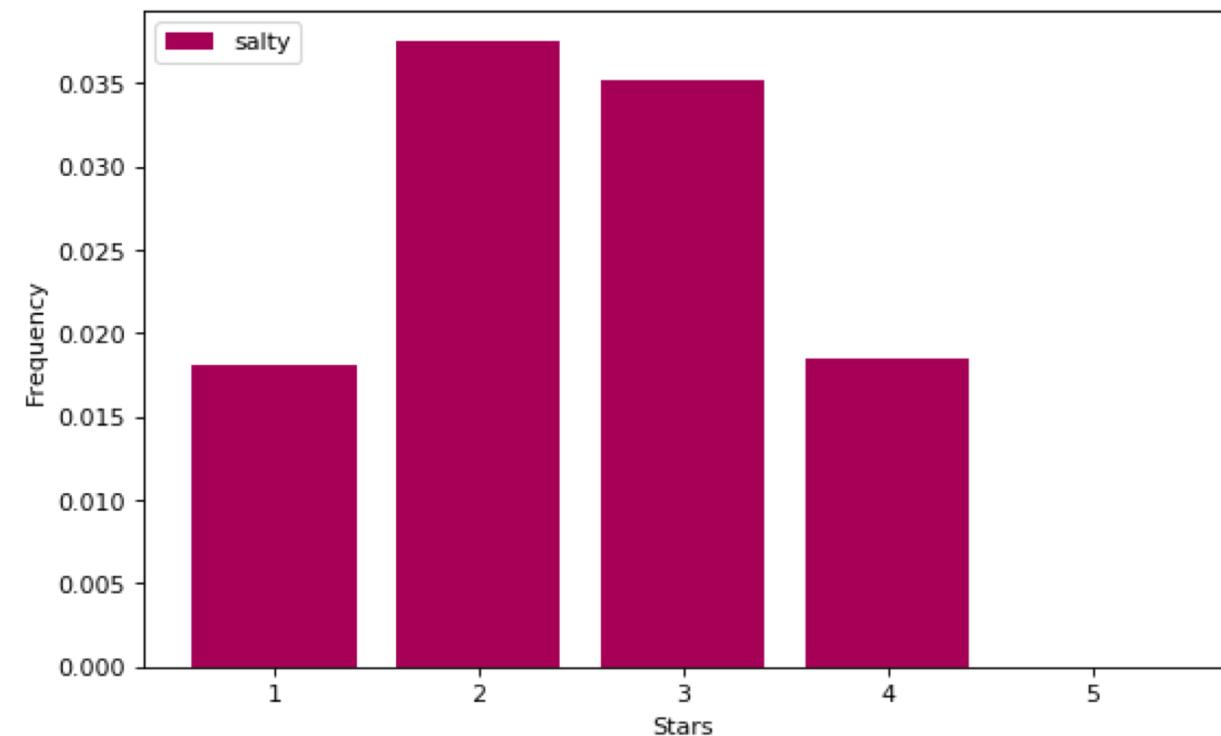
# Words Analysis

- Food

delicious



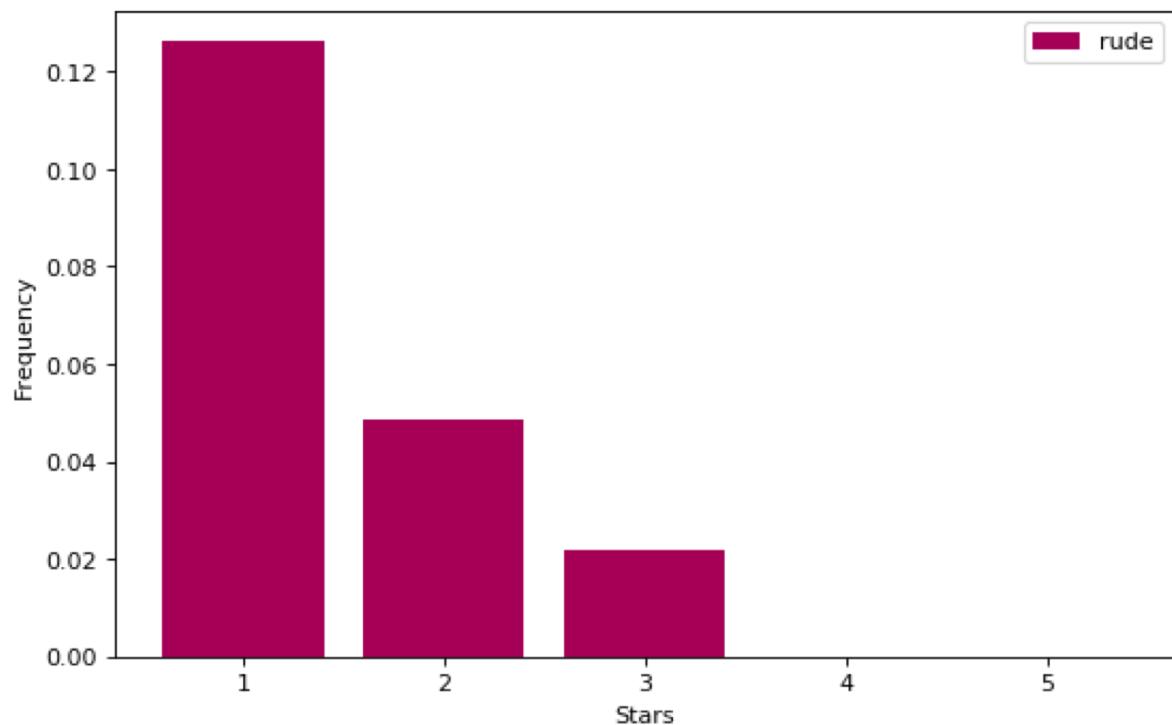
salty



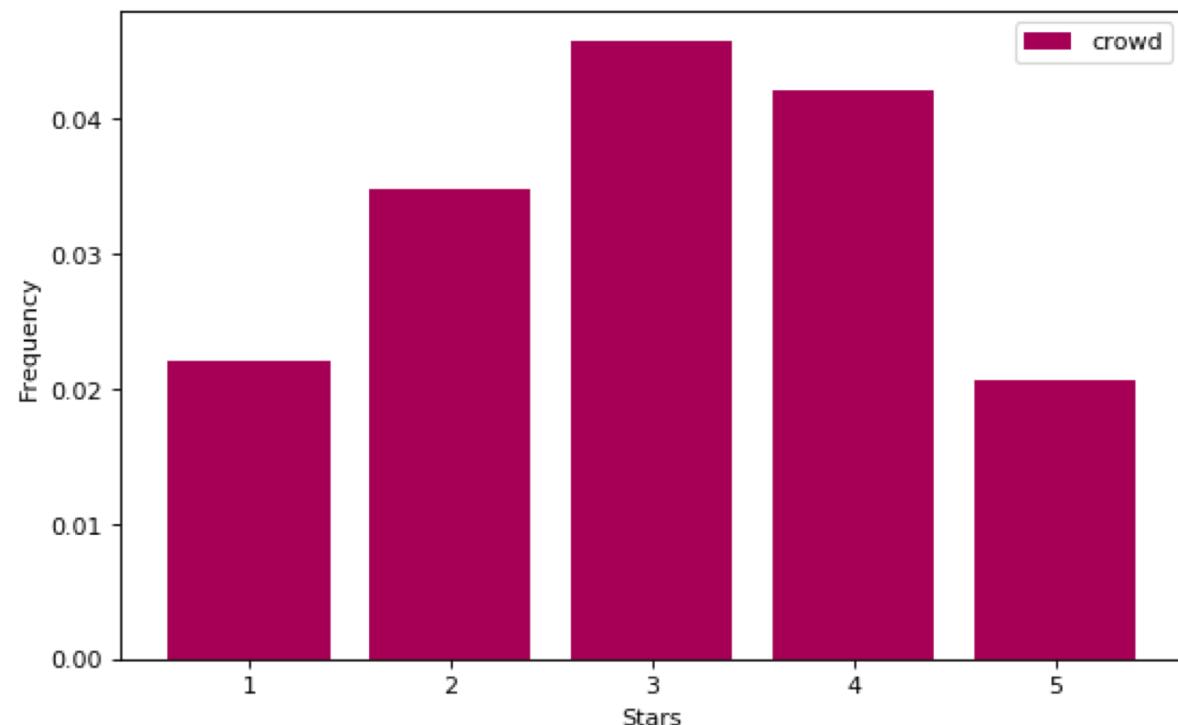
# Words Analysis

- Service

rude



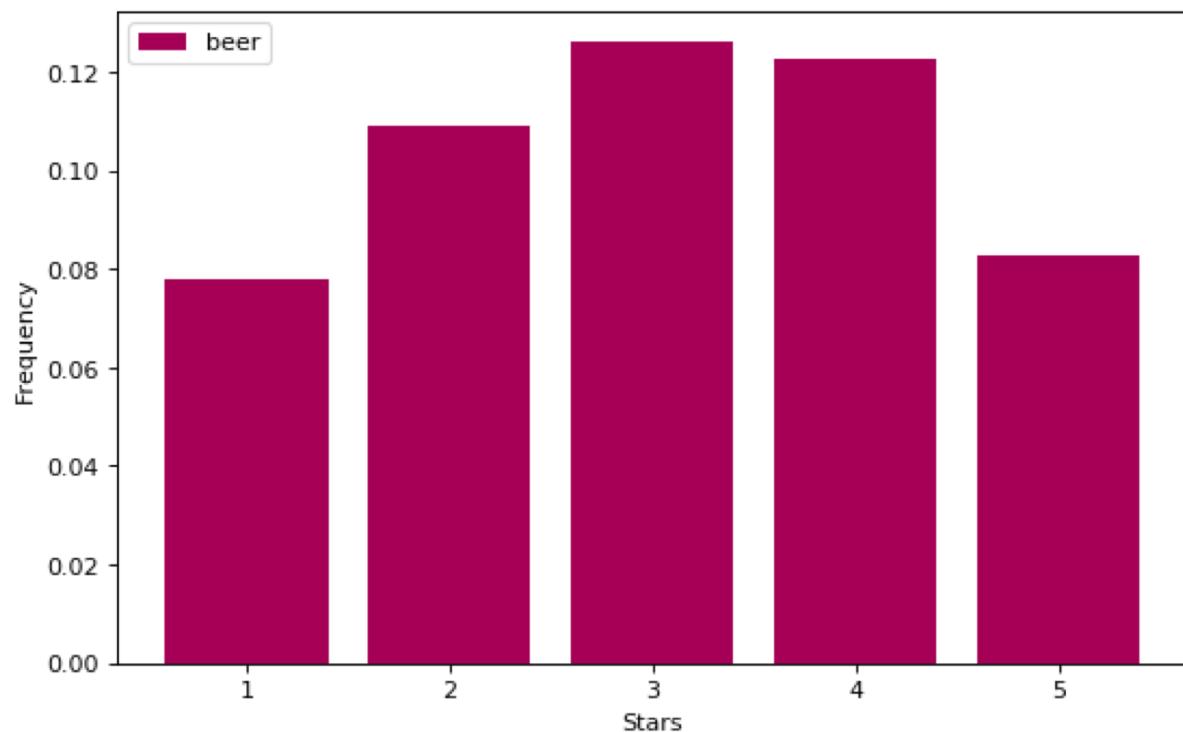
crowd



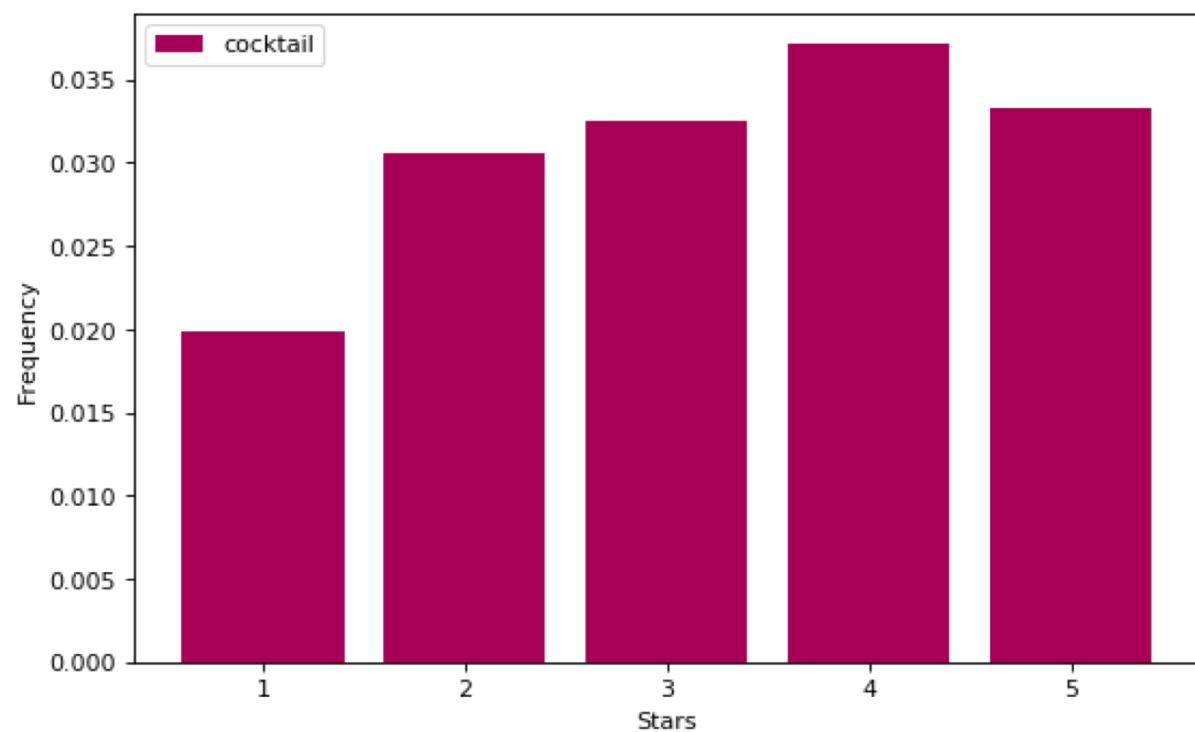
# Words Analysis

- Beverage

beer



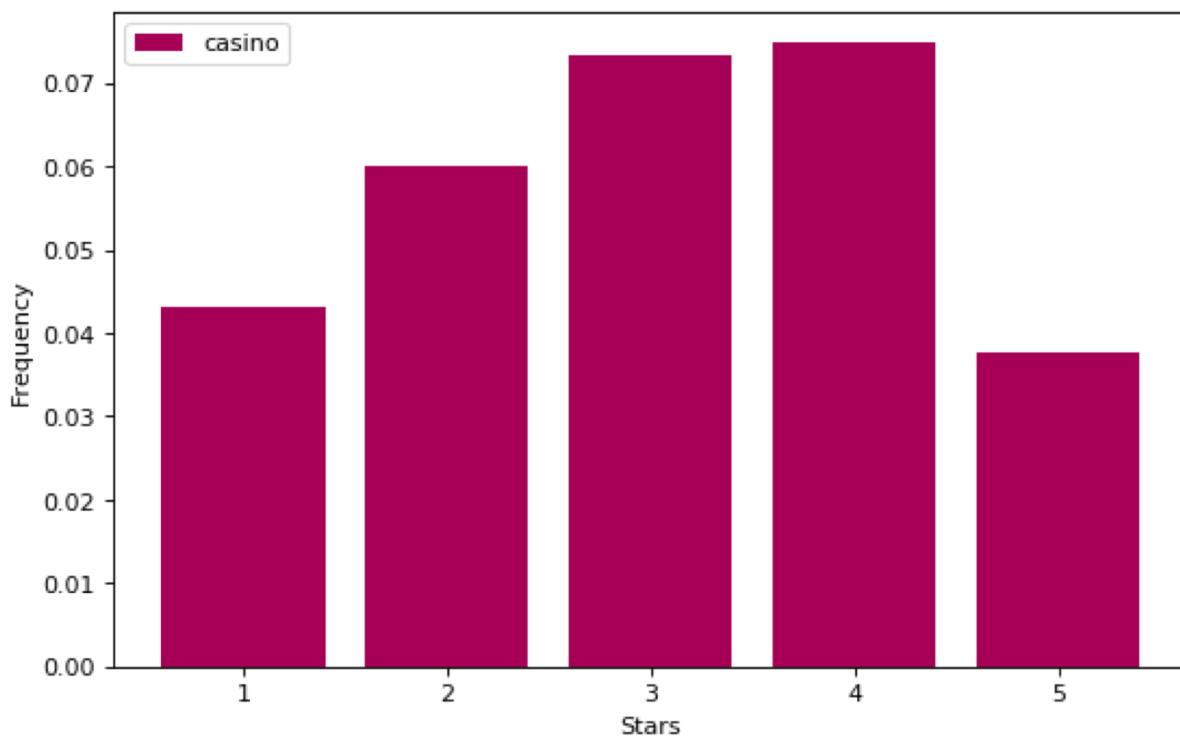
cocktail



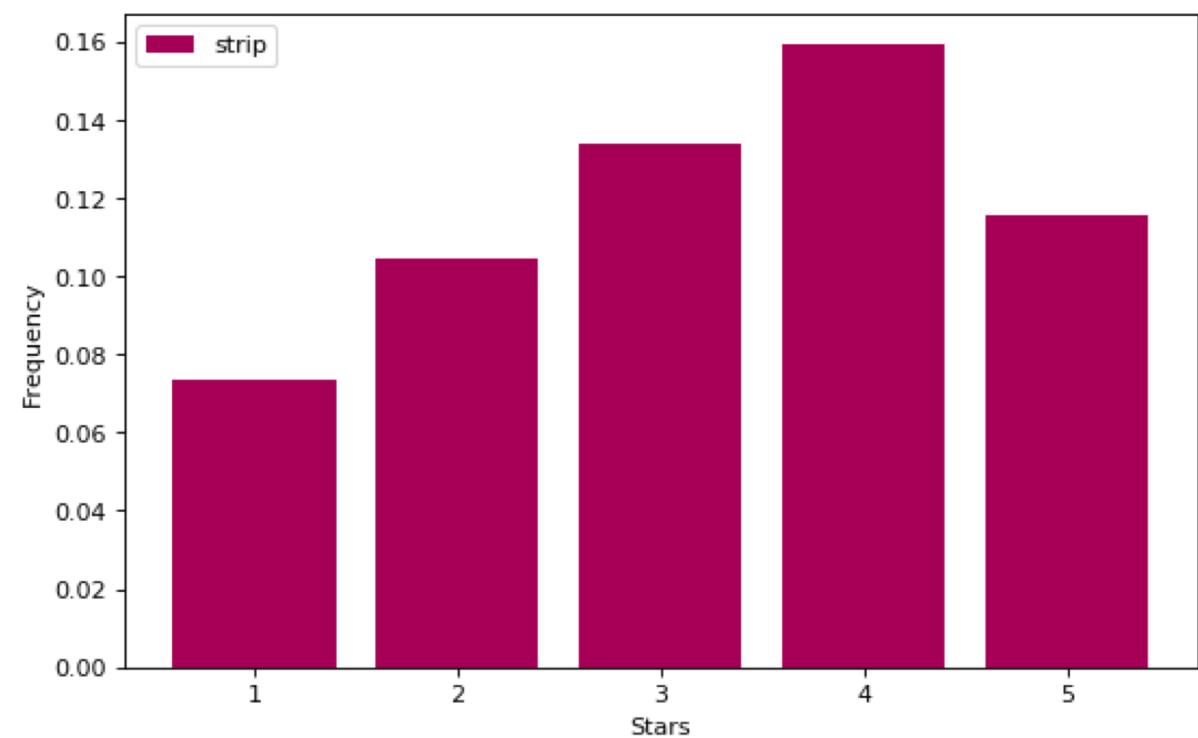
# Words Analysis

- Entertainment

casino



strip



# Future Work

- **Data cleaning**  
verb - tense  
nouns - singular or plural  
ngrams
- **Factors**  
open hours  
...

**Thanks!**