

# Nightlife in Las Vegas

Group 6 - Junxia Zhu, Yiqun Jiang, Yingjing Jiang

# Outline

- Goal & Data Filtering
- Business Analysis
- Review Analysis
- Suggestion
- Strengths & Weakness

# Goal & Data Filtering

- **Goal**

explore nightlife in Las Vegas

provide insights to help improve ratings

- **Data Filtering**

rules: categories: nightlife, bars

city: Las Vegas

size: business 1,201

reviews 265,847

# Business Analysis

- **Data cleaning**
  - extract all attributes
  - calculate average star for all business
  - missing value → “unknown”
- **Important attributes:** random forest importance score
  - NoiseLevel
  - RestaurantsPriceRange

# Business Analysis

- **Data cleaning**
  - extract all attributes
  - calculate average star for all business
  - missing value → “unknown”
- **Important attributes:** random forest importance score
  - NoiseLevel
  - RestaurantsPriceRange

# Business Analysis

- Important attributes

- One-way ANOVA

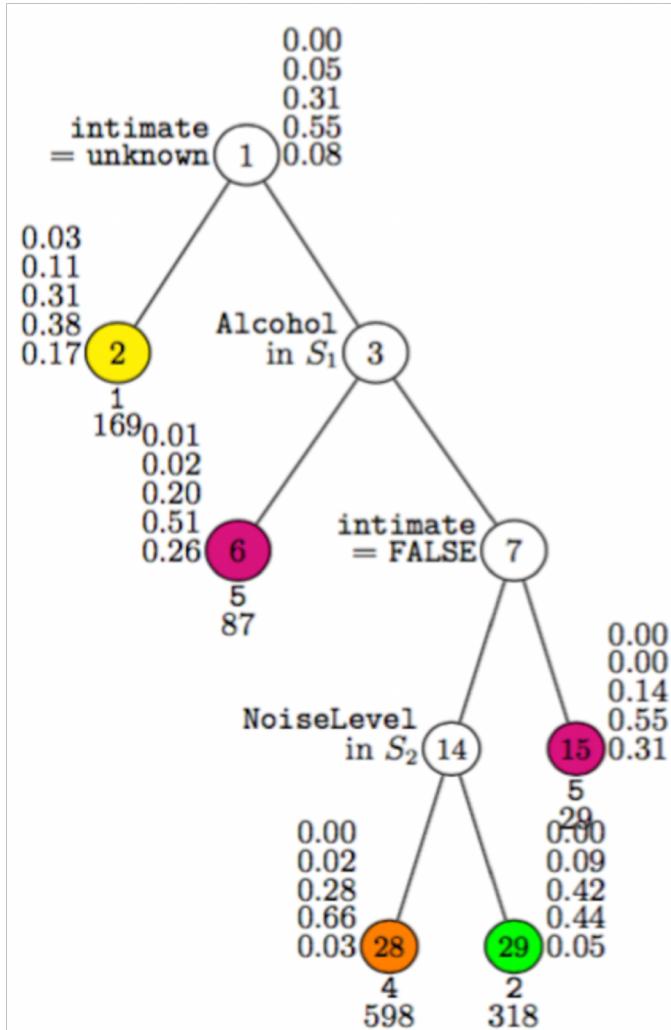
Term	P-value
NoiseLevel	4.36x10 <sup>-8</sup>
RestaurantPriceRange	2.44x10 <sup>-3</sup>

- Two-way ANOVA

Term	P-value
NoiseLevel	3.02x10 <sup>-8</sup>
RestaurantPriceRange	3.50x10 <sup>-3</sup>
NoiseLevel:RestaurantPriceRange	3.58x10 <sup>-2</sup>

# Business Analysis

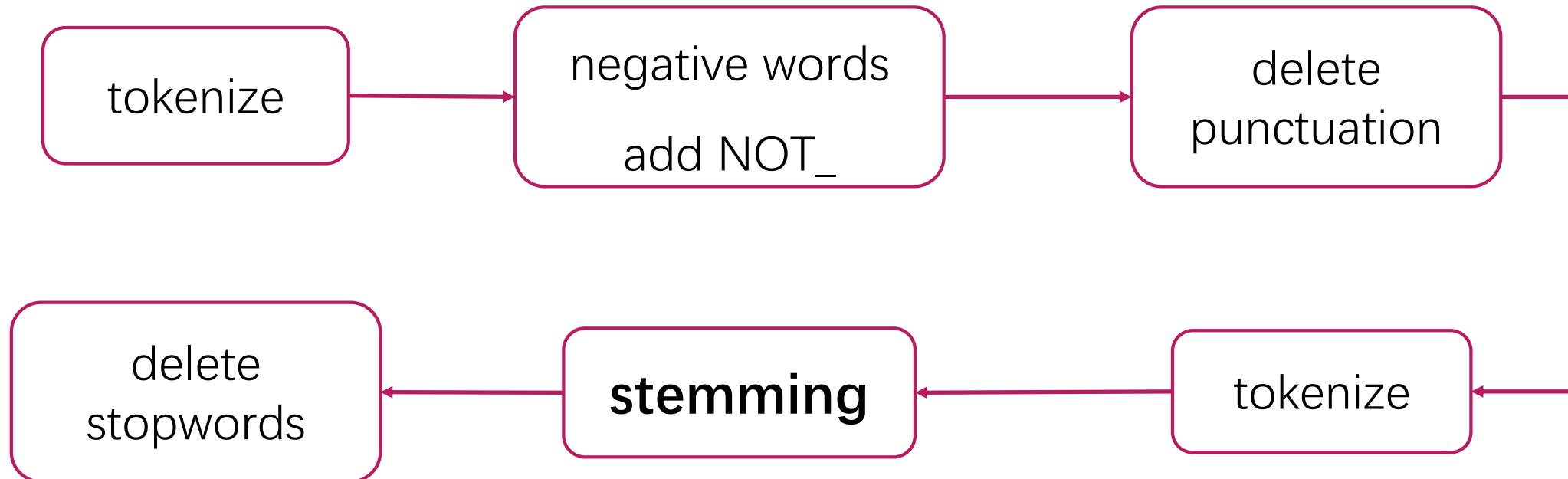
- Missing value analysis



GUIDE v.29.6 0.80-SE pruned classification tree for predicting stars using equal priors and unit misclassification costs. Maximum number of split levels is 11 and minimum node sample size is 12. At each split, an observation goes to the left branch if and only if the condition is satisfied. For splits on categorical variables, values not present in the training sample go to the right. Set  $S_1 = \{\text{'beer\_and\_wine'}, \text{'none'}\}$ . Set  $S_2 = \{\text{'average'}, \text{'quiet'}\}$ . Predicted classes and sample sizes printed below terminal nodes; class proportions for stars = 1, 2, 3, 4, and 5, respectively, beside nodes. Second best split variable at root node is Alcohol.

# Review Analysis

- Data cleaning



# Review Analysis

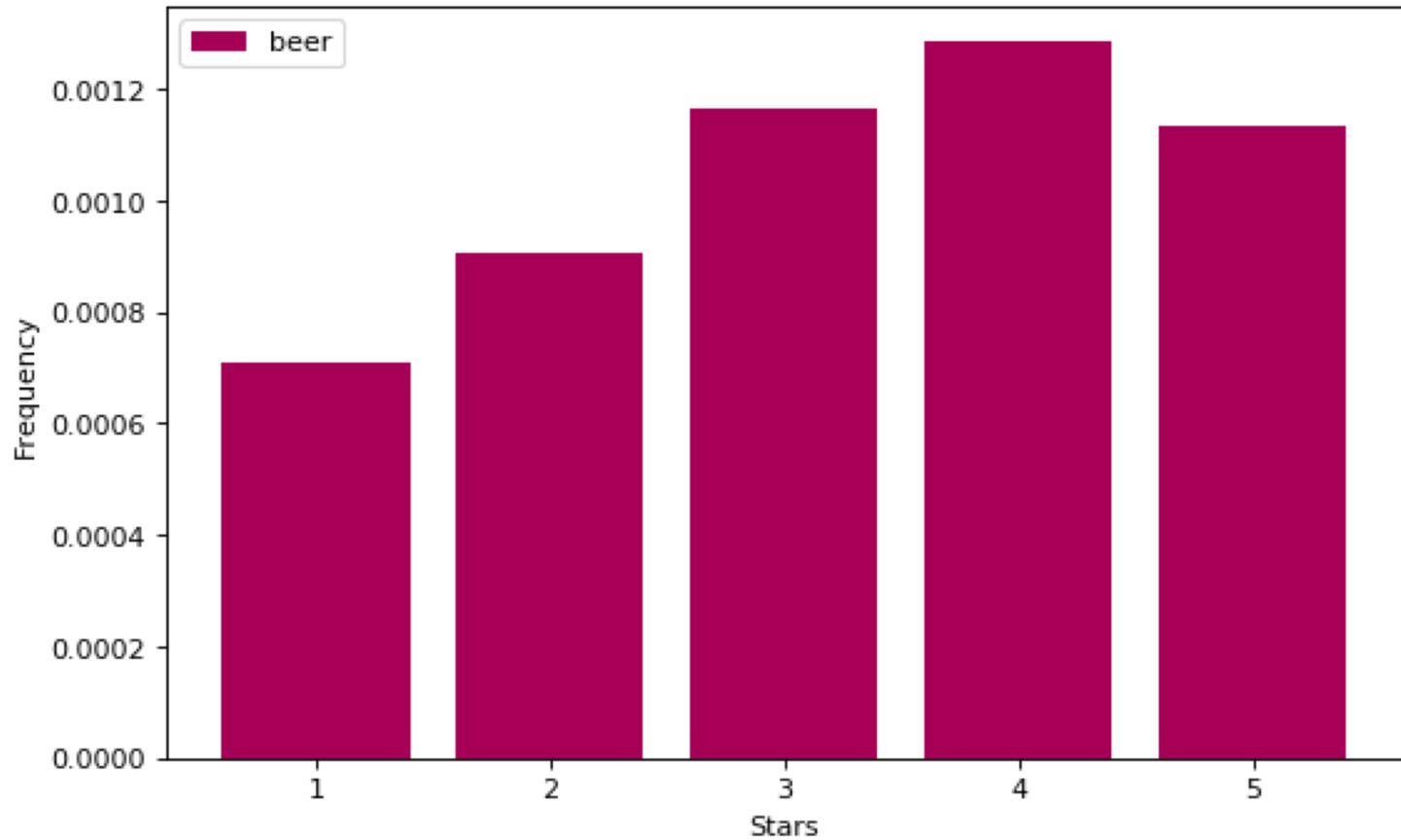
- **Unigram:** 4 classes

food

beverage

entertainment

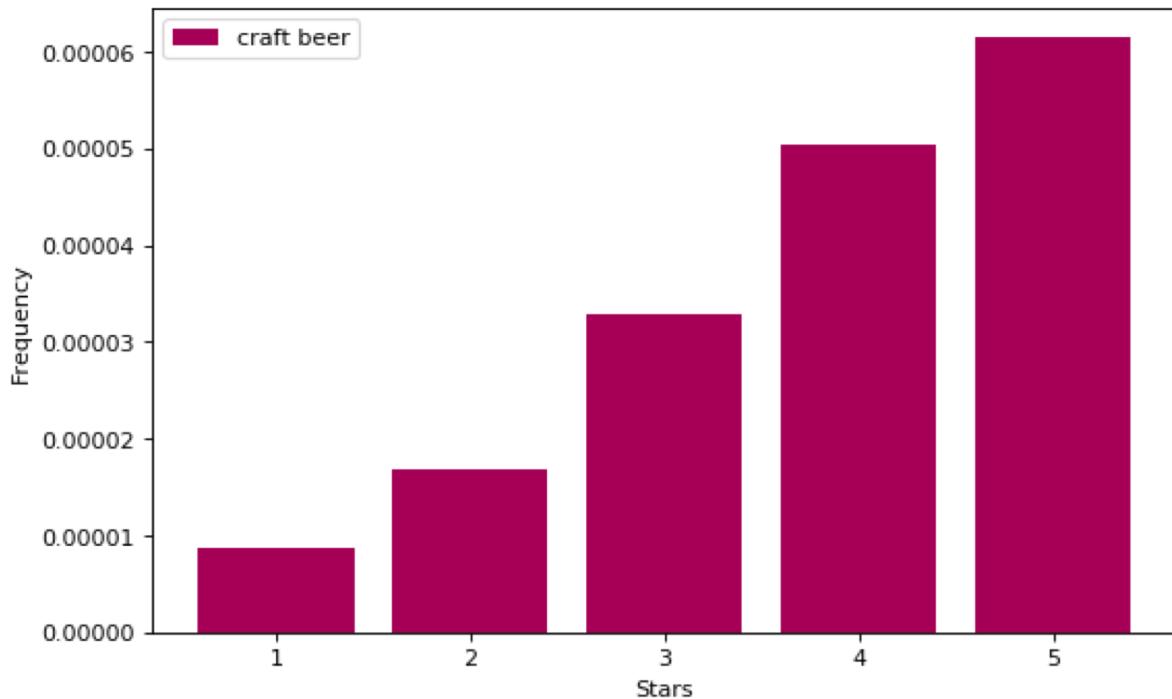
service



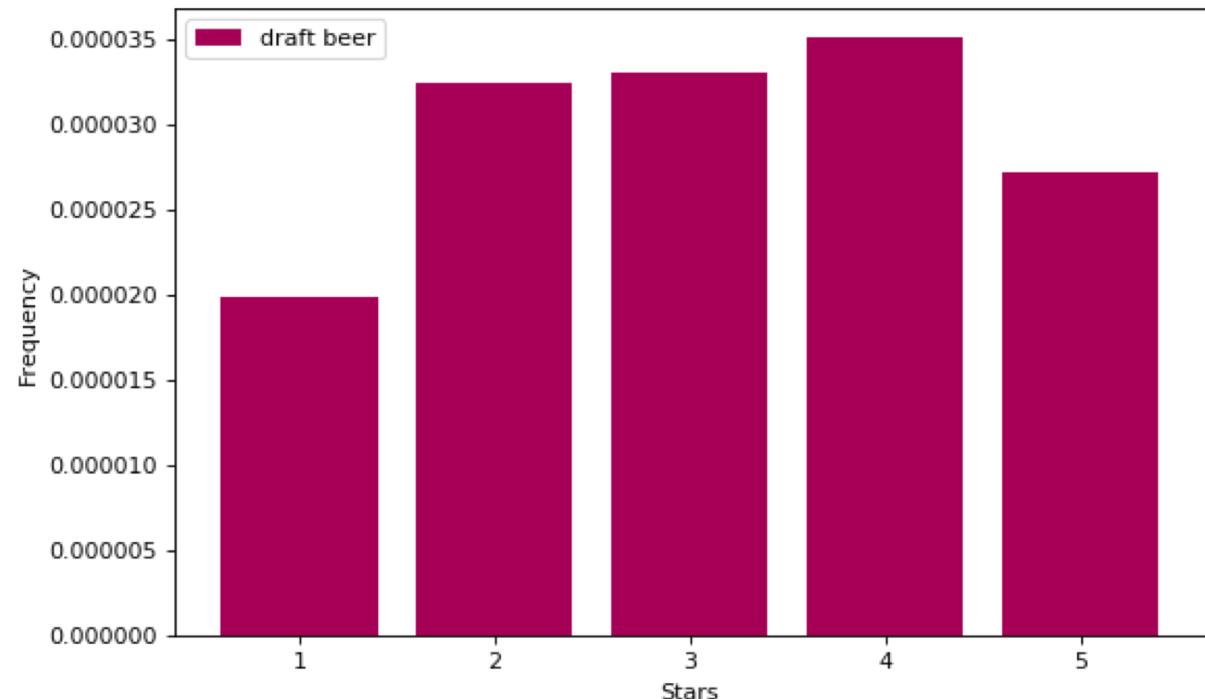
beer

# Review Analysis

- Bigram key words: beer → craft beer, draft beer



craft beer



draft beer

# Reviews Analysis

- Adjective words analysis

Great    location    so    if    you    want    iced    beer,    all    is    good.  
NNP            NN           RB          IN          PRP        VBP        JJ            NN        DT        VBZ        JJ

The diagram illustrates the parts of speech for the sentence "Great location so if you want iced beer, all is good." The words are underlined and grouped by part of speech: NNP (Great), NN (location), RB (so), IN (if), PRP (you). Verb groups are VBP (want) and JJ (iced). A noun group is NN (beer,). Articles are DT (all). A linking verb is VBZ (is). Adjectives are JJ (good). A checkmark icon is above want and iced, while a crossed-out X icon is above beer,.

- Example      key word: beer

Top 5 positive adjective: great, good, warm, special, free      total=313

Top 5 negative adjective: cold, small, flat, damn, poor      total=132

# Reviews Analysis

- Shiny app [https://yingjingjiang.shinyapps.io/shiny\\_app/](https://yingjingjiang.shinyapps.io/shiny_app/)

**Get suggestion for you business**

Enter business id and choose words to get the corresponding summary and suggestion.

business\_id  
130241

words  
cocktail  
beer  
beer select  
draft beer  
craft beer  
cocktail

Summary   Detail   Suggestion

average rating: 4.18

Percentage of positive and negative words

positive  
negative

The cocktail in this business is above average(90.1%).

# Reviews Analysis

- Shiny app [https://yingjingjiang.shinyapps.io/shiny\\_app/](https://yingjingjiang.shinyapps.io/shiny_app/)

## Get suggestion for you business

Enter business id and choose words to get the corresponding summary and suggestion.

**business\_id**

**words**

**Update**

**Summary**   **Detail**   **Suggestion**

This page shows the top 5 high frequency adjective words.  
**positive words:** classic, well, best, great, good  
**negative words:** old, usual, NOT\_sweet, poor, pricey

# Suggestion

- **attributes:** noise level, intimate
- **food:** salty level
- **service:** control attendances, waiters
- **entertainment:** casino, strip
- **beverage:** craft beer, draft beer, wide selection of beers

# Strengths & Weakness

- **Strength**
  - deal with missing value
  - use tree methods & ANOVA to confirm our results
  - adjective corpus obtained manually from data itself
- **Weakness**
  - fail to find patterns in time and hour analysis
  - just include several bigrams to give advice

**Thanks!**