Analyzing Restaurant Customer Review Using Sentiment Analysis

Piangkwan Jaikaew

17-12-2022

1. Introduction

Customer satisfaction in the restaurant industry has an outsized impact on success. Asking for feedback and review helps in improving the customer satisfaction. The reviews plays a vital role for restaurants to be successful and steadily grow in the future.

In this case study, I will conduct a sentiment analysis for the local restaurant. The original dataset was sourced from Kaggle datasets. It consists of 1000 reviews from customers.

2. Summary

The dataset demonstrates mainly positive reviews, approximately 50% from word counts.By using three different sentiment analysis methods, NRC, bing and loughran to extract and analyze in text mining, it shown the correlated result as a positive feedback.

3. Methodology

1. Clean and transform data using dplyr package.
2. Analyse sentiments using the syuzhet package based on the NRC sentiment dictionary.
3. Conduct sentiment analysis using the "bing" lexicon.
4. Conduct sentiment analysis using the "loughran" lexicon.
5. Visualization with ggplot2 package.

Import and transform data.

Sentiment Analysis

Syuzhet package is used for sentiment analysis

```
## # A tibble: 10 x 4
##    count emotion        perc labels
##    <dbl> <chr>         <dbl> <chr>
##  1   115 fear         0.0360 3.60%
##  2   149 disgust      0.0466 4.66%
##  3   153 anger        0.0479 4.79%
##  4   157 sadness      0.0492 4.92%
##  5   176 surprise     0.0551 5.51%
```

```
##  6   330 negative     0.103  10.33%
##  7   339 anticipation 0.106  10.61%
##  8   506 trust        0.158  15.84%
##  9   513 joy          0.161  16.06%
## 10   756 positive     0.237  23.67%
```

Chart 1: The table shows a number of count and percentage of ten different emotion categories.

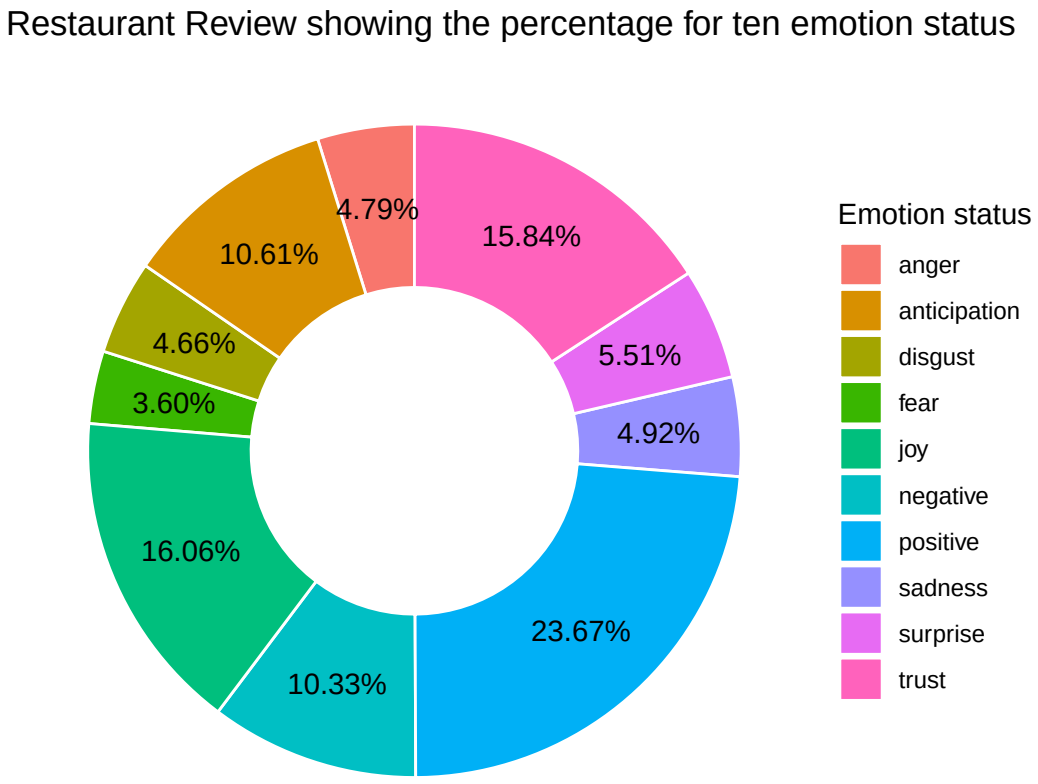

Chart 2: Approximately 50% of reviews expresses positive feeling toward the restaurant.

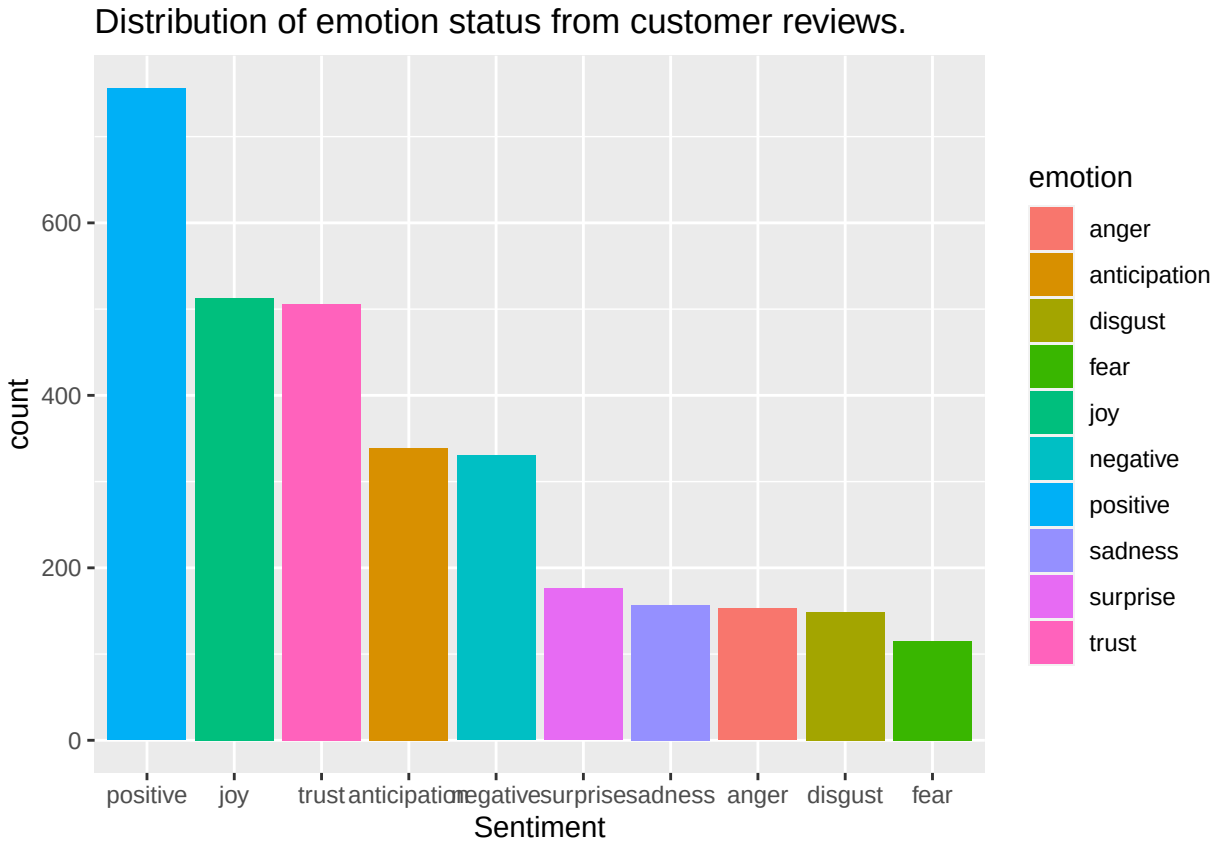

Chart 3: The word counts show mainly positive which is good news for restaurant.

Sentiment Analysis with Bing lexicon

```
## # A tibble: 23 x 3
##    word         sentiment     n
##    <fct>        <chr>     <int>
##  1 disappointed negative     19
##  2 bad          negative     18
##  3 worst        negative     15
##  4 bland        negative     11
##  5 slow         negative     11
##  6 terrible     negative     10
##  7 cold         negative      9
##  8 rude         negative      8
##  9 overpriced   negative      7
## 10 hard         negative      6
## # ... with 13 more rows
```

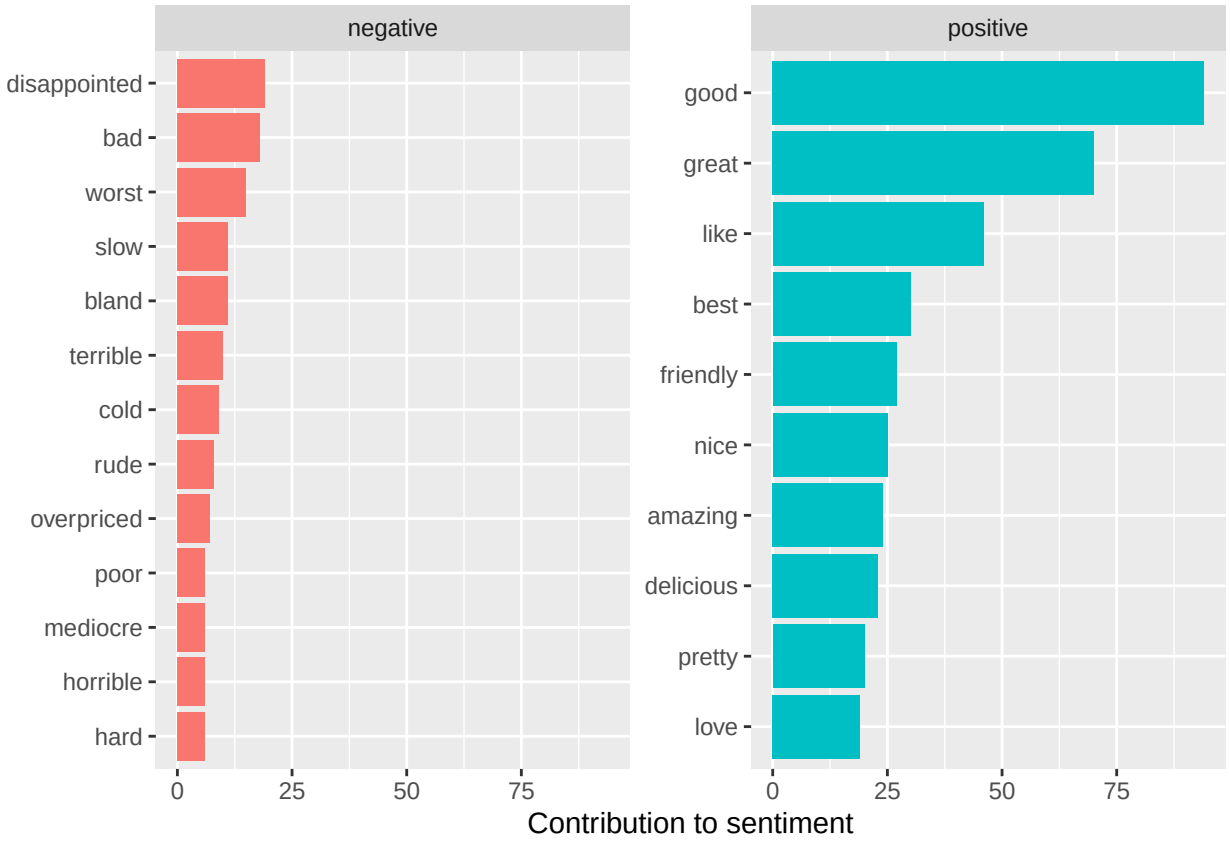

Chart 4: The positive sentiment is demonstrated in a greater proportion than negative sentiment from bing lexicon.

Sentiment Analysis with Loughran lexicon

```
## # A tibble: 44 x 3
##    word         sentiment     n
##    <fct>        <chr>     <int>
##  1 appealing    litigious     1
##  2 court        litigious     1
##  3 juries       litigious     1
##  4 law          litigious     1
##  5 lawyers      litigious     1
##  6 shall        litigious     1
##  7 whatsoever   litigious     1
##  8 disappointed negative     19
##  9 bad          negative     18
## 10 worst        negative     15
## # ... with 34 more rows
```

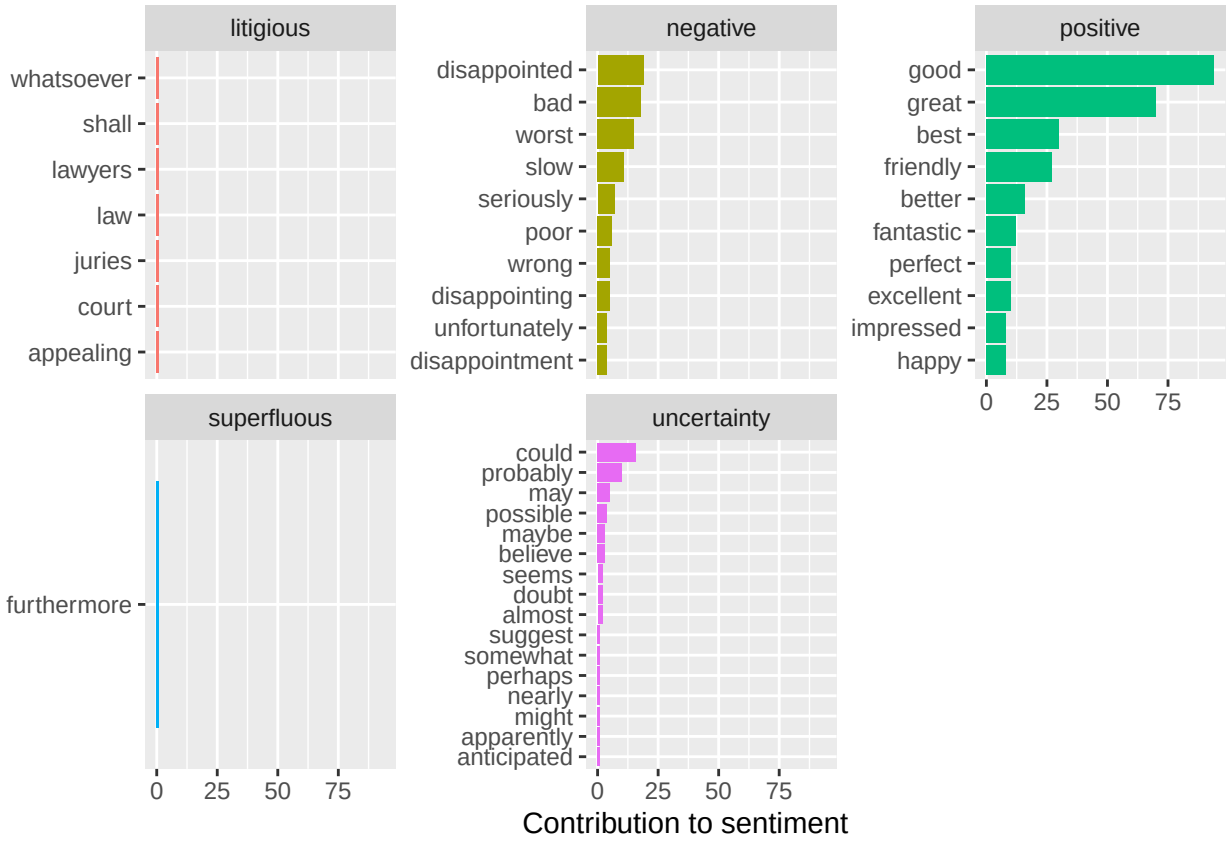

Chart 5: Still show primary positive! even more specific expressions such as uncertainty are shown by Loughran lexicon.

4. Conclusion

For limitation, the current data needs to provide a more oversized data frame and a specific period of time to measure on exact feedback in order to minimize a bias information to be as low as possible.