Business Analysis of Pizza Hut based on Yelp data

STAT 628 Module 3 Group 8

1. Introduction and Motivation

Pizza is one of the most famous food all over the world, so it makes us interested in large pizza chains. From a report on the internet, we find top 10 pizza chains within the US in 2020. Firstly, we focus on top 10 pizza chains[1]: Domino's Pizza, Pizza Hut, Little Caesars, Papa John's Pizza, California Pizza Kitchen, Papa Murphy's, Sbarro, Marco's Pizza, Chuck E. Cheese's/Peter Piper Pizza, and Cici's Pizza. After the preliminary analysis, we find that Pizza Hut have the lowest user ratings. Hence, we decide to focus on Pizza Hut's business analysis based on Yelp data and try to give the owners of Pizza Hut recommendations.

2. Data Cleaning

2.1 Data Information

Due to the need of our analysis, we decided to use review.json, attribute.json and business.json files from Yelp data. Our original data includes about 8.64 million reviews and 160 thousand business.

2.2 Data Cleaning

Our goal is exploring businesses of Pizza Hut and gather insights about these businesses through reviews. To begin with, we filtered businesses with keywords which belong to Pizza Hut. Then we dropped some information we do not need. In the end, there are Pizza Hut stores' 3186 reviews and 145 businesses with useful columns: shops' average stars, review counts, attributes text, users' individual stars, and review text.

3. Exploratory Data Analysis

3.1 Features Extraction

3.1.1 Word Clouds

Firstly, we aimed to analyze word frequency from review texts to consider which factors customers care about. We ignore all punctuation marks and numbers, default stop words in package 'wordcloud' and some uninformative words like 'pizza', 'go'. We divided the data into two groups: high ratings (user's stars>2.5) and low ratings (user's stars<=2.5) and calculated the word frequency and drew world cloud figures.

High Ratings (stars>2.5)

Low Ratings (stars<=2.5)

Figure 1. Word Cloud

From figure 1, we can find that delivery is the most concerned aspect by users. Then, a notable point is that cold is a relative high frequency word in low ratings reviews while hot is

a high frequency word in high ratings, and as we all know these two words are antonyms, which tells us that perhaps the heat preservation of food is an aspect that users care about.

3.1.2 TF-IDF and Sentiment Analysis

Because sometimes word frequency and word clouds can not help us differentiate between specific restaurants clearly enough, we utilize TF-IDF, which emphasize words with high frequency in one document but not high in the whole documents set. This method can help us exclude some words frequently but not very informative and meaningful, which can be mathematically characterized by the following formula:

$$TFIDF = TF * IDF = \frac{a\ count\ in\ a\ review}{words\ count\ in\ a\ review} * log\frac{reviews\ count}{reviews\ containing\ a\ count} \quad (1)$$

where a is a specific word.

Then, to make the analysis clearer, we analyzed and got words with top TF-IDF index words from reviews with 5 stars and 1 star. We choose features common to both 5 stars and 1 star.

After features extraction, we calculated the scores of sentiment analysis between one and five stars for features that we considered valuable, and the characteristics with large differences in scores between the two became the characteristics of our concern. The result is as the following table:

features	*delivery*	*service*	*staff*	*driver*	*employee*	*pepperoni*	*sauce*
5-stars scores	*0.21*	*0.31*	*0.28*	*0.22*	*0.21*	*0.18*	*0.18*
1-star scores	*0.04*	*0.06*	*0.08*	*0.04*	*0.06*	*0.06*	*0.05*

Table 1

3.1.3 Word Tree[2]

A word tree displays textual data in a hierarchical manner: as a tree of elements, single words are linked by lines. The font size of words represents their frequency. We can use word tree to visualize the relationship between some particular words, a word tree in our analysis is as following figure:

Figure 2. word tree

3.2 Attribute Analysis

We analyze the attribute text, and process and plot the star ratings data of multiple features in the attribute text. We chose features that have large or obvious star ratings differences in different situations.

4. Key findings about Pizza Hut

After exploratory data analysis, we focused on 5 features:

From reviews text:	delivery, cold food, service
From attributes:	good for kids, and good for groups

Table 2. 5 features

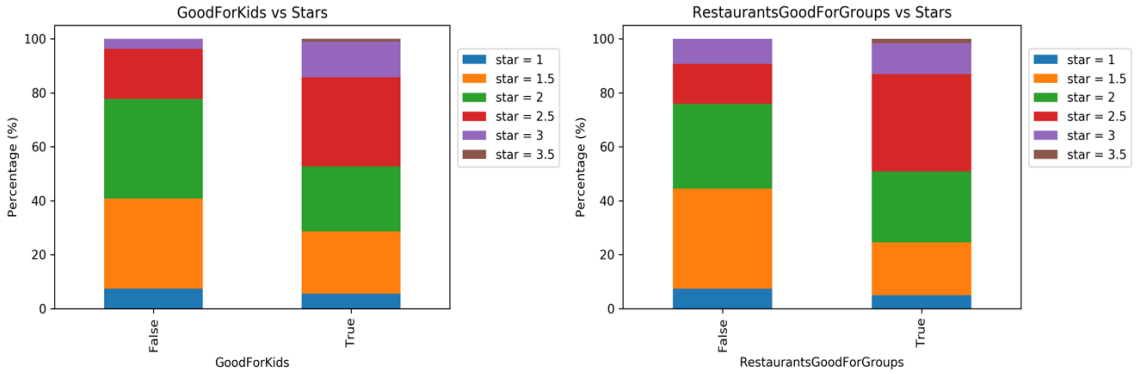

Figure 3. Star Ratings by good for kids/good for groups

Figure 3 indicates that frequency of each star ratings is correlated with these two attributes.

For an attribute, GoodForKids, we can see that restaurants that are more kids-friendly generally have a higher star rating. In specific statistical number, the average rating for GoodForKids is 2.14 while the average rating for NotGoodForKids is 1.89, which means a Pizza Hut restaurant with kids-friendly service had 0.25 star higher rating compared to a Pizza Hut restaurant without kids-friendly service. The result is convinced because the p-value of the t-test is 0.014 (<0.05).

And for GoodForGroups, like GoodForKids, the average rating for GoodForGroups is 2.17 while the average rating for NotGoodForGroups is 1.91, which means a Pizza Hut restaurant with groups-friendly service had 0.26 star higher rating compared to a Pizza Hut restaurant without groups-friendly service. The result is convinced because the p-value of the t-test is 0.006 (<0.05).

For significant features in reviews text, we obtain some insights.

For service, an amusing find is that the average rating of reviews associated with “manager” is 1.41 while the average rating of reviews including “friendly” within “manager” reviews is 3.04 (p-value=10^-4<0.05). It’s a huge different for Pizza Hut considering that the highest average store star of Pizza Hut is 3.5.

For delivery, we found that average rating associated with “deliver” is 1.75 and with

“wait|slow” in “deliver” reviews is 1.33. Hence, it is reasonable that Pizza Hut restaurants with faster delivery service tend to have higher ratings than Pizza Hut restaurants with slow delivery service.

For food, among all the 3186 reviews of Pizza Hut, 358 (about 11.2%) reviews mention the adjective “cold”. And the average sentiment score associated with “cold” is -0.096. A representative review is “There is no way I want a discount for cold pizza. It is unacceptable and really upsets me.”

5. Recommendations for Businesses

Based on our key findings about Pizza Hut, we have some general recommendations for Pizza Hut. Based on the characteristics of the Yelp data we use; the specific goal of our recommendations is to improve the restaurant’s user reviews.

The first thing put forward is that to make restaurants kids-friendly. For example, restaurants can provide the necessary child seating equipment, sell healthy children’s meals, and even build a small children’s playground.

Then, restaurants can make them more group-friendly by creating dedicated and larger spaces for group gatherings and assigning waiters to serve groups.

Next, for service, restaurants can provide managers with professional training to help them be kinder to their customers.

As for faster and on-time delivery, restaurants can optimize the meal process to speed up meals or use better delivery services for companies or people.

Last, restaurants can optimize food packaging to make it easier to keep warm, making it harder to get cold.

In our shiny app, we give unique and targeted advice to each of Pizza Hut's specific stores

6. Discussion

6.1 Weakness and Future work

In our analysis process, star ratings data of Pizza Hut violate normality assumption for hypothesis test because of discrete values and limit range. And we choose features which are significant and interesting to analyze, but perhaps there are other factors that can be analyzed to improve restaurants’ star ratings.

We concentrated on Pizza Hut's data in this project. In the future, we can also analyze and process data on other pizza restaurants, and it is likely that we will obtain some new and effective conclusions.

6.2 Summary

In this module, we pay most attention to one of the top 10 pizza chains in U.S., Pizza Hut. We combine different methodologies to help us extract significant and interesting feature from review text and attribute text data.

For these features, we try to use data analysis processes and ideas that are understandable to decision makers from non-statistical backgrounds, visually and concretely presenting our analysis results and recommendations.

7. Contributions

Yuyang He: wrote summary, revise presentation slides, statistical hypothesis test, data analysis.

Haoyue Shi: R shiny app, revise presentation slides, EDA code, data analysis

Shengwen Yang: wrote presentation slides, Data cleaning and merge, EDA code, data analysis

8. Reference

[1] 10 largest pizza chains in the US. Zippia. (n.d.). Retrieved November 18, 2021, from https://www.zippia.com/advice/largest-pizza-chains/

[2] Word Tree. (n.d.). Retrieved December 9, 2021, from https://www.jasondavies.com/wordtree/.