

MACS 30200 Final Paper

Spring 2018

Fangfang Wan

A Research on Yelp.com – What Affect Restaurants Ratings and an Investigation on Contagion Effect in Catering Industry

Abstract

In this research, I focus on the Yelp.com, a well-known online business rating and commenting platform, and investigate contagion effect in catering industry. I collected information of top 975 restaurants in Chicago (Yelp - The Best 10 Restaurants in Chicago, IL, 2018) using web scraping in Python. There are 2 research questions that I have answered in this research. The first question is what factors can affect a restaurant's Yelp rating. Linear regression model in Python shows that there are many factors associated in this process such as loudness, ambience and accessibility of wheelchair. K-fold cross validation with $K=4$ yields mean-squared-error of 0.1. In these factors, I focus on geographical location specifically. Therefore, the second research question to investigate whether there exists contagion effect, a theory in psychology and behavioral economics, in catering industry or not. Computing correlations between related variables shows that restaurants that are closer in geographical distance have closer Yelp ratings, and they also bear higher cosine similarity scores.

Introduction

Customer review has played a critical role in the operation of businesses, and it has

long been a topic behavioral economics, business and marketing researches. In the old days, customers' perceptions are spread from person to person through daily conversations, and previous literatures have provided evidence that traditional viva voce customer ratings affect business revenue. For example, Mahajan et al. (1984) discussed the behaviors of word-of-mouth spread based on a new "diffusion model" in movie markets, and provided evidence that better customer perceptions lead to higher movie revenue, while negative customer reviews can result in lower revenue. However, with the development social media and technology, the online business rating platforms have become increasingly important in business operation and management, and online review systems have substituted the role of traditional word-of-mouth to some extent. There are more abundant literatures on how online reviews affect sales and revenue of products, partly due to the fact that online review data are more easily collected and analyzed. For instance, Liu (2006) discussed the patterns and effects of movie reviews on Yahoo movies, claiming that the forum is most active just before the release of the movies, while the reviews has the strongest effect on box office revenue when movies are on.

Additionally, Hu et al. (2014) extend their studies to geographical effect – higher rating of a restaurant tends to mildly increase the ratings of other surrounding restaurants. This implies that geographically closer restaurants may exhibit more similar characteristics, and therefore they may have closer level of customer reviews. Such phenomenon is associated with "contagion effect", which is a theory that has been deeply investigated in

psychology (Hatfield et al., 1993). In behavioral economics academia, scholars have incorporated psychology theories into economics studies, and “contagion effect” is one of them. It has been especially deeply researched in behavioral finance literatures to study irrational behaviors of participants in financial markets (Hirshleifer and Teoh, 2003).

In this research, I will focus on the Yelp.com, a business rating and commenting platform, and investigate contagion effect in catering industry. There are 2 research questions that I will answer in this research. The first question is what factors can affect a restaurant’s Yelp rating. There are probably many factors associated in this process, and I will discuss them later. However, among all those factors, I will focus on geographical location specifically. Therefore, the second research question to investigate whether there is contagion effect in catering industry or not.

This paper is divided into seven parts. The first part is the introduction above, and then followed by literature review in related social science fields and a formal definition of theories and hypotheses associated with this paper. The fourth and fifth parts are data description and methods. Results are presented subsequently, and this paper will end by a conclusion.

Literature Review

In the old days, customers’ perceptions are spread from person to person through daily conversations, and previous literatures have provided evidence that traditional viva voce customer ratings affect business revenue. For example, Mahajan et al. (1984)

discussed the behaviors of word-of-mouth spread based on a new “diffusion model” in movie markets, and provided evidence that better customer perceptions lead to higher movie revenue, while negative customer reviews can result in lower revenue. Not only in commercial areas, similar patterns also exhibit in public sectors. Mackay and Crompton (2006) analyzed customer reviews in recreation service, and also provided evidence that better guest reviews increase revenue. Recently, with the development social media and technology, the online business rating platforms have become more and more important in business operation and management, and online review systems have substituted the role of traditional word-of-mouth to a great extent. Furthermore, online ratings have the same type of boosting effect on restaurant revenue as viva voce guest reviews. There are more abundant literatures on how online reviews affect sales and revenue of products, partly due to the fact that online review data are more easily collected and analyzed, and some authors even employed more recent computational methods as innovative tools to do analysis to a deeper level. Such effect also exists in various industries. For instance, Liu (2006) discussed the patterns and effect of movie reviews on Yahoo movies, claiming that the forum is most active just before the release of the movies, while the reviews has the strongest effect on box office revenue when movies are on. As for books market, Chevalier and Mayzlin (2006) analyzed customer reviews and sales of books on Amazon.com and Barnesandnoble.com, and found out that better reviews on one site lead to better sales in the same site, and 1-star reviews have larger effect on sales than 5-star

reviews. More closely related to my research topic on catering industry, Zhang et al. (2010) investigated reviews by consumers and found out that consumer reviews have positive effect on people's intention to visit a restaurant, while editor reviews exhibit negative effect. Park and Nicolau used dataset of around 5000 reviews of 45 restaurants in London and New York, and claimed the asymmetric effect of customer reviews – customers tend to react more to extremely good or bad reviews, while do not react much to others. As for Yelp.com specifically, Luca (2016) discussed the positive effect of higher ratings to restaurant revenue. Furthermore, the author also described the behaviors of customers and restaurant ratings – a review of one restaurant does not affect chained restaurants, and customers are affected by a review that conveys more information, while exhibiting an asymmetric and less attention to other reviews. Anderson and Magruder (2012) did a similar analysis and yield similar result that positive reviews lead to higher restaurant revenue, while claiming different behaviors that such effect tend to be larger when other information of this restaurant is not available. Due to such influence of rating on restaurants, some restaurants even generate fake reviews to attract more customers and therefore increase their revenue (Luca, 2016).

Literatures in the previous paragraph have provided strong evidence that guest reviews play a significant role in determining a restaurant's revenue, and the boosting effect of customer reviews on business revenue provides my research with the ultimate legitimacy and underlying theoretical basis. Information about what lead to higher

customer reviews is valuable for both restaurant owners and customers. For restaurant operators, they need information about what underlying factors lead to higher ratings so that they can improve their ratings and subsequently revenue accordingly. For customers, they can adjust their own expectation for a restaurant. For instance, suppose quietness of a restaurant is positively correlated with the rating of a restaurant. If, however, a customer does not mind loud environment or even prefer loudness, he or she may expect that the restaurant is better than what the rating suggests.

There are some literatures that investigate what factors lead to higher customer rating and satisfaction in different areas. The influencing factors on guest satisfaction vary according to industries. In business area, and specifically, hospitality industry, Ramanathan (2011) did research on hotels in United Kingdom, and found out that “value for money” is critical for returned stay of customers – it is crucial for increasing customers’ satisfaction, while inadequacy in “value for money” may lead to negative effect that cannot be compensated by other factors. “Customer service”, “Room quality” and “Quality of good” are dissatisfier factors, which means that lower performance in these areas can lead to negative result in the number of returned customers. In finance sector, Yang and Fang (2004) used content analysis methods to investigate 52 items in 16 major service quality dimensions from a dataset of 740 reviews in online securities brokerage services. They found out that the most significant factor contributing to higher customer service is the primary service quality, and lower customer satisfaction is mainly

resulted from information systems quality. Tripp et al. (2010) analyzed travelers at an information center, and investigated their behaviors. They concluded that the most significant factors affecting their choices are sanitation, food quality, and service quality. McAuley and Leskovec (2013) used innovative computational methods in analyzing review texts, found out the factors influencing customer ratings, and even performed predictions of user ratings. As for Yelp.com, Huang et al. (2014) used computational methods - Latent Dirichlet Allocation (LDA) algorithms, predicted hidden topics in customer reviews, and claimed that service, values, taking out and decorations are the four most crucial factors that influence Yelp rating of a restaurant, followed by healthiness of food, waiting time, background music, breakfast availability, dinner availability, and lunch availability. Hu et al. (2014) even extend their studies to geological effect – higher rating of a restaurant tends to mildly increase the ratings of other surrounding restaurants, which motivates my further investigation on “contagion effect” in catering industry.

Contagion effect has been long researched in psychology. Hatfield et al. (1993) is a notable research paper in psychology academia discussing emotional contagion. Behavioral economics use theories in psychology to explain irrational behaviors of human beings in economy, so that the contagion effect has also been a topic in behavioral economics, especially in behavioral finance. For example, Bikhchandani and Sharma (2000) provide an overview of herding behavior in financial market. Hirshleifer and Teoh

(2003) also investigated contagious behavior of financial market participants such as analyst, firms, and various kinds of investors, while Çelen and Kariv, S. (2004) shows that scholars in behavioral finance should distinguish cascading behavior from herding behavior.

Theories and Hypotheses

There are mainly two theories and 2 corresponding hypotheses associated with this paper. First, the boosting effect of customer reviews on business revenue provides my research with the ultimate legitimacy and underlying theoretical basis, and the corresponding hypothesis is that there are a lot of factors that can affect a restaurant's Yelp ratings, such as loudness of a restaurant, which is expected to exhibit the pattern that the louder a restaurant is, the lower the rating is likely to be. The second theory is the "contagion effect" theory in psychology and behavioral economics, and the corresponding hypothesis is that the closer two restaurants are in geographical distance, the more similar they are in terms of restaurant attributes, and the closer star ratings on Yelp they have.

Data

In this research, I used BeautifulSoup library in Python and scraped data of top 975 restaurants (Yelp - The Best 10 Restaurants in Chicago, IL, 2018). For each restaurant, I collected information about the name, rating stars, the number of reviews they have, latitude, longitude, operating hours every day throughout a week, neighborhood, price

level (\$, \$\$, \$\$\$, or \$\$\$\$), category (Asian, New American, Modern European, etc.), whether they take reservations or not, delivery availability, take-out permitted, acceptance of credit cards, Apply Pay and Android Pay, what it is good for (brunch, dinner, etc.), car and bike parking availability, wheelchair accessibility, whether it is good for kids or not, whether it is good for groups or not, attire requirement, ambience, noise level, alcohol availability, outdoor seating availability, Wi-Fi and TV availability, and whether it has waiter service and caters or not.

Since a part of data collected is presented as categorical variable, I have converted them to category dummy variables. For example, as for alcohol, there are only 3 kinds of values – Full Bar, No Alcohol, and Beer & Wine Only. Therefore, I have created 3 dummy variables for each of them. Missing values are recorded as 0.

Some special variables should be noticed. One is the neighborhood variable. I only included neighborhoods that have at least 30 restaurants in this dataset. The other one is category. Since there are more than 100 different categories, I have generated some new categories. For example, Chinese, Japanese and Korean are all categorized into “North Eastern Asian.

Methods

In the data collection part, I used web scraping in Python as described in the section above. As for computational methods to analyze my dataset, I first used sklearn in Python to run linear regressions, while the attributes of restaurants are independent variables, and

the star rating is the dependent variable. I used K-fold cross validation to compute the Mean-Squared-Error (MSE) of my model. To investigate contagion effect, I match each restaurant with all other possible restaurants one by one. Since there are nearly 1000 restaurants, there are nearly half a million restaurant pairs. I compute the correlation between star rating difference and geological distance between each pair (since I have data about their latitude and longitude, I can compute their distance in kilometer in Python). Then I used sklearn library in Python again to compute cosine similarity between the two restaurants in each pair, and compute the correlation between their cosine similarity and geological distance in kilometer.

Results

I first did a statistical description of star rating distribution in the dataset. We can see that more than half of the top restaurants earn a 4.0 star rating, which is consistent with my daily life experience. Some earn 4.5 starts, but only a few earn 5 stars. Since this dataset contains top restaurants in Chicago, it is reasonable that not a lot of restaurants have 3.5 and 3.0 stars.

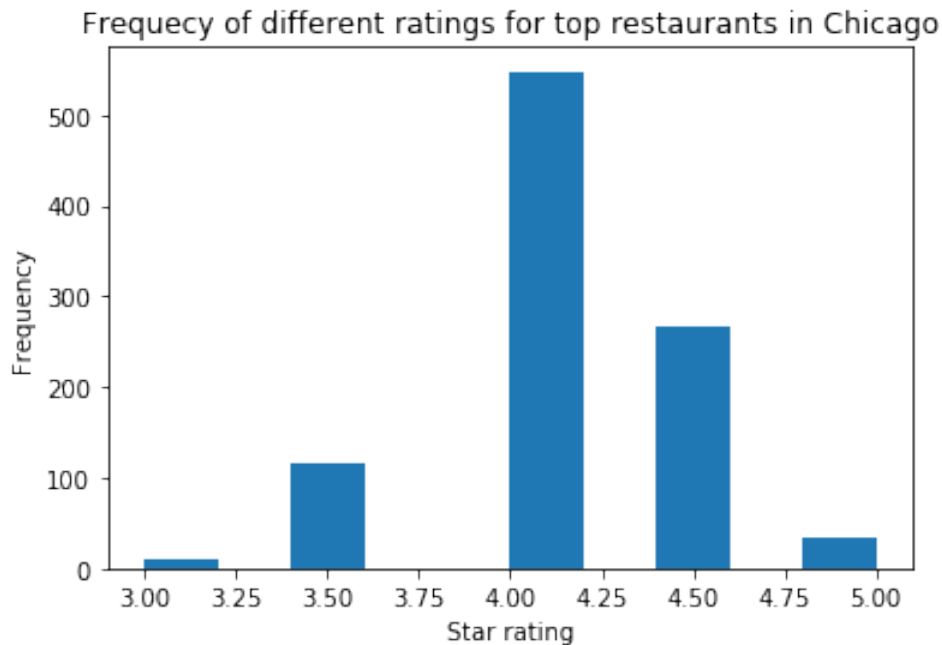


Figure 1: Frequency of different ratings for top restaurants in Chicago

Then I run the basic model, a linear regression, on the whole dataset. Below is the regression table.

| | | | |
|--------------------------|------------------|----------------------------|----------|
| Dep. Variable: | stars | R-squared: | 0.325 |
| Model: | OLS | Adj. R-squared: | 0.270 |
| Method: | Least Squares | F-statistic: | 5.865 |
| Date: | Sun, 03 Jun 2018 | Prob (F-statistic): | 1.53e-40 |
| Time: | 17:16:07 | Log-Likelihood: | -210.10 |
| No. Observations: | 975 | AIC: | 570.2 |
| Df Residuals: | 900 | BIC: | 936.4 |
| Df Model: | 74 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P> t | [0.025 | 0.975] |
|-----------------------|------------|----------|--------|-------|-----------|----------|
| const | 4.4810 | 0.101 | 44.254 | 0.000 | 4.282 | 4.680 |
| review_count | 1.816e-05 | 1.73e-05 | 1.052 | 0.293 | -1.57e-05 | 5.2e-05 |
| Monday hrs | -0.0114 | 0.003 | -3.497 | 0.000 | -0.018 | -0.005 |
| Tuesday hrs | -0.0034 | 0.005 | -0.666 | 0.506 | -0.013 | 0.007 |
| Wednesday hrs | 0.0168 | 0.008 | 2.088 | 0.037 | 0.001 | 0.033 |
| Thursday hrs | -0.0108 | 0.008 | -1.325 | 0.186 | -0.027 | 0.005 |
| Friday hrs | 0.0057 | 0.006 | 0.887 | 0.375 | -0.007 | 0.018 |
| Saturday hrs | -0.0150 | 0.004 | -3.545 | 0.000 | -0.023 | -0.007 |
| Sunday hrs | 0.0060 | 0.003 | 1.912 | 0.056 | -0.000 | 0.012 |
| West Loop | -0.0577 | 0.111 | -0.521 | 0.602 | -0.275 | 0.160 |
| Wicker Park | -0.2007 | 0.064 | -3.125 | 0.002 | -0.327 | -0.075 |
| Lincoln Park | -0.0219 | 0.054 | -0.405 | 0.685 | -0.128 | 0.084 |
| Near West Side | 0.0427 | 0.105 | 0.406 | 0.685 | -0.164 | 0.249 |
| DePaul | -0.0763 | 0.065 | -1.173 | 0.241 | -0.204 | 0.051 |
| Ukrainian Village | 0.0042 | 0.076 | 0.055 | 0.956 | -0.144 | 0.153 |
| Logan Square | 0.0564 | 0.049 | 1.147 | 0.252 | -0.040 | 0.153 |
| River North | -0.0186 | 0.063 | -0.294 | 0.769 | -0.143 | 0.106 |
| Near North Side | -0.0660 | 0.038 | -1.753 | 0.080 | -0.140 | 0.008 |
| Lakeview | -0.0096 | 0.041 | -0.235 | 0.814 | -0.090 | 0.071 |
| West Town | 0.0763 | 0.049 | 1.560 | 0.119 | -0.020 | 0.172 |
| The Loop | -0.1414 | 0.049 | -2.873 | 0.004 | -0.238 | -0.045 |
| Reservation2 | -0.0137 | 0.028 | -0.494 | 0.621 | -0.068 | 0.041 |
| Delivery | -0.0308 | 0.024 | -1.267 | 0.205 | -0.078 | 0.017 |
| Takeout | 0.0059 | 0.038 | 0.155 | 0.877 | -0.069 | 0.081 |
| credit card | 0.1258 | 0.061 | 2.074 | 0.038 | 0.007 | 0.245 |
| Apple Pay | -0.0105 | 0.037 | -0.284 | 0.776 | -0.083 | 0.062 |
| Android Pay | -0.0449 | 0.040 | -1.128 | 0.260 | -0.123 | 0.033 |
| outdoor2 | -0.0274 | 0.024 | -1.162 | 0.246 | -0.074 | 0.019 |
| Wi-Fi | -1.613e-16 | 2.8e-16 | -0.577 | 0.564 | -7.1e-16 | 3.88e-16 |
| TV | 0.0061 | 0.024 | 0.253 | 0.800 | -0.041 | 0.053 |
| waiter service | 0.0016 | 0.026 | 0.062 | 0.950 | -0.049 | 0.052 |
| catering | -0.0154 | 0.025 | -0.619 | 0.536 | -0.064 | 0.033 |
| Breakfast | 0.0267 | 0.047 | 0.563 | 0.574 | -0.066 | 0.120 |
| Brunch | 0.0240 | 0.029 | 0.824 | 0.410 | -0.033 | 0.081 |
| Lunch | 0.0315 | 0.029 | 1.096 | 0.273 | -0.025 | 0.088 |
| Dinner | -0.0282 | 0.031 | -0.896 | 0.371 | -0.090 | 0.034 |
| Dessert | 0.0889 | 0.050 | 1.775 | 0.076 | -0.009 | 0.187 |
| Late Night | -0.0460 | 0.053 | -0.870 | 0.385 | -0.150 | 0.058 |
| Valet | -0.0246 | 0.034 | -0.726 | 0.468 | -0.091 | 0.042 |
| Street | 0.0389 | 0.033 | 1.180 | 0.238 | -0.026 | 0.103 |
| Garage | -0.0593 | 0.045 | -1.304 | 0.192 | -0.149 | 0.030 |
| Private Lot | -0.0740 | 0.042 | -1.771 | 0.077 | -0.156 | 0.008 |
| Validated parking | 0.0102 | 0.055 | 0.183 | 0.855 | -0.099 | 0.119 |
| wheelchair accessible | 0.0584 | 0.025 | 2.350 | 0.019 | 0.010 | 0.107 |
| Good for Kids | 0.0772 | 0.032 | 2.447 | 0.015 | 0.015 | 0.139 |
| Good for groups | -0.1131 | 0.039 | -2.890 | 0.004 | -0.190 | -0.036 |
| Casual attire | -0.2134 | 0.079 | -2.708 | 0.007 | -0.368 | -0.059 |
| Dressy | -0.2607 | 0.097 | -2.680 | 0.008 | -0.452 | -0.070 |
| Formal | -0.2455 | 0.346 | -0.710 | 0.478 | -0.924 | 0.433 |
| Casual ambience | -0.1173 | 0.031 | -3.794 | 0.000 | -0.178 | -0.057 |
| Romantic | -5.184e-18 | 5.67e-17 | -0.091 | 0.927 | -1.16e-16 | 1.06e-16 |
| Classy | 0.0832 | 0.048 | 1.737 | 0.083 | -0.011 | 0.177 |
| Intimate | 0.1340 | 0.050 | 2.677 | 0.008 | 0.036 | 0.232 |
| Trendy | 0.0287 | 0.028 | 1.019 | 0.308 | -0.027 | 0.084 |
| Upscale | 0.1638 | 0.071 | 2.321 | 0.021 | 0.025 | 0.302 |
| Hipster | -0.0672 | 0.051 | -1.308 | 0.191 | -0.168 | 0.034 |
| Touristy | -0.2926 | 0.239 | -1.223 | 0.222 | -0.762 | 0.177 |
| Full Bar | 0.1384 | 0.086 | 1.616 | 0.106 | -0.030 | 0.306 |
| No Alcohol | 0.2820 | 0.082 | 3.435 | 0.001 | 0.121 | 0.443 |

| | | | | | | |
|-----------------------------|---------|-------|--------|-------|--------|--------|
| Beer & Wine Only | 0.1267 | 0.093 | 1.365 | 0.173 | -0.055 | 0.309 |
| Very Loud | -0.3114 | 0.129 | -2.406 | 0.016 | -0.565 | -0.057 |
| Loud | -0.2055 | 0.079 | -2.596 | 0.010 | -0.361 | -0.050 |
| Avg_loudness | -0.1809 | 0.066 | -2.757 | 0.006 | -0.310 | -0.052 |
| Quiet | -0.0725 | 0.074 | -0.974 | 0.331 | -0.219 | 0.074 |
| African | -0.0220 | 0.139 | -0.159 | 0.874 | -0.294 | 0.250 |
| NE Asian | 0.0073 | 0.034 | 0.211 | 0.833 | -0.060 | 0.075 |
| SE Asian | 0.0232 | 0.071 | 0.327 | 0.744 | -0.116 | 0.163 |
| Indian | 0.1258 | 0.109 | 1.151 | 0.250 | -0.089 | 0.340 |
| Middle Eastern | 0.0141 | 0.077 | 0.183 | 0.855 | -0.137 | 0.165 |
| N American | 0.0252 | 0.023 | 1.097 | 0.273 | -0.020 | 0.070 |
| S American | 0.0031 | 0.026 | 0.120 | 0.904 | -0.048 | 0.054 |
| European | 0.0249 | 0.024 | 1.017 | 0.309 | -0.023 | 0.073 |
| Bars | -0.0133 | 0.022 | -0.595 | 0.552 | -0.057 | 0.030 |
| Bakery | 0.0147 | 0.049 | 0.298 | 0.766 | -0.082 | 0.111 |
| BBQ | 0.0509 | 0.057 | 0.894 | 0.372 | -0.061 | 0.163 |
| W Asian E European | 0.0227 | 0.195 | 0.117 | 0.907 | -0.359 | 0.405 |
| Price | -0.0421 | 0.026 | -1.602 | 0.110 | -0.094 | 0.009 |

Table 1. Linear Regression on All Restaurants in the Dataset

From the result of linear regression, we can see that the R-squared is 0.325, which is reasonable, since there are many complicated factors that can define a restaurant's rating. For instance, if a group of people rate high or low on business purpose for a restaurant, then the rating will be higher or lower than the real customer review level, respectively. Moreover, there might also be "contagion effect" among customers, in which they may rate close to the average rating.

The regression result shows that not many coefficients are significant. However, we can still detect some patterns and associations between restaurant attributes and star ratings on Yelp. Restaurants in West Town are likely to be better rated. Better dessert is likely to bring better customer reviews. Restaurants that are compatible with wheelchair and that are good for kids are higher rated, and the coefficients are significant at 95% confidence level. Restaurants with casual ambience are lower rated, and the coefficient is

significant at 99.9% confidence level, while it is also statistically significant that restaurants that have intimate and that has upscale ambience are higher rated. It is significant at 99% confidence level that restaurants with no alcohol are rated 0.282 stars higher. The general trend is that the louder a restaurant, the lower the rating is. Furthermore, the more expensive, the lower ratings, holding other factors constant. K-fold cross validation with $K=4$ yields $MSE=0.108$.

Next, to investigate contagion effect, I compute the correlation of start difference and geological distance in kilometers, and the result is 0.051. This means that further geological distance between two restaurants is associated with more distinct Yelp ratings. To investigate whether such difference is associated with restaurant attribute distribution according to geographical location, I then compute the cosine similarity between 2 restaurants in each pair, and then the correlation between cosine similarity and geological distance in kilometer is -0.055, which shows that restaurants that are closer are indeed more similar in terms of restaurant characteristics. Such result indicates the existence of contagion effect in these restaurants. I have plotted a scatterplot of restaurant pairs with distance in kilometer as x-axis, and cosine similarity as y-axis. The scatterplot below does not exhibit an obvious pattern, but it is also roughly consistent with the trend – similar restaurants cluster in the left side.

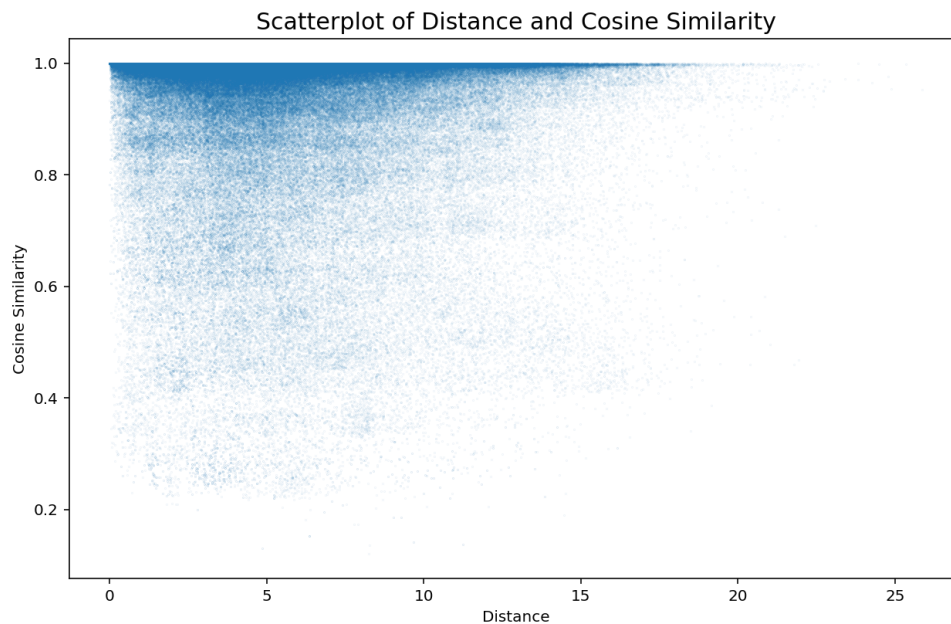


Figure 2. Scatterplot of Distance and Cosine Similarity

Conclusion

This research paper has 2 main research goals. The first is to investigate what factors can affect a restaurant's rating on Yelp.com. The R-squared is not exceptionally high, and some coefficients are not significant. However, we can still detect some patterns from linear regression, and found out that a restaurant's ambience, wheelchair accessibility, loudness, kids accessibility, and alcohol availability have relatively significant effect on a restaurant's rating. However, the result also exhibits that there are some limitations in this research. The R-squared is 0.325, which means that my linear regression model explains 32.5% of variations in Yelp ratings of top 975 restaurants in Chicago. More than half of the variations are not explained. This is mostly because that information on Yelp is limited, and there are still some other factors that we cannot obtain from Yelp, or they

are hard to collect and/or quantify. For example, if some people highly or lowly rate a restaurant on business purpose, then the restaurant will have higher or lower rating than the real customer review level, respectively. Another limitation is that my dataset only contains about 1000 restaurant in Chicago. Other restaurants in Chicago may exhibit different behaviors, and restaurants in other places in United States and over the world are possible to violate the model that I have established in this paper.

The second research question is to investigate the contagion effect in catering industry in Chicago. From my computation of cosine similarity and 2 types of correlations, it shows that restaurants that are closer in geographical locations have closer Yelp ratings, and they are also more similar in terms of restaurant attributes. A possible extension on this topic is to find out the reason behind this pattern, and future work on this topic will contribute to behavioral economics academia.

References

- Anderson, M., & Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563), 957-989.
- Bikhchandani, S., & Sharma, S. (2000). Herd behavior in financial markets. *IMF Staff papers*, 47(3), 279-310.
- Çelen, B., & Kariv, S. (2004). Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review*, 94(3), 484-498.
- Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3), 345-354.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current directions in psychological science*, 2(3), 96-100.
- Hirshleifer, D., & Hong Teoh, S. (2003). Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management*, 9(1), 25-66.
- Huang, J., Rogers, S., & Joo, E. (2014). Improving restaurants by extracting subtopics from yelp reviews. iConference 2014 (Social Media Expo).
- Hu, L., Sun, A., & Liu, Y. (2014, July). Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval (pp. 345-354). ACM.
- Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, 70(3), 74-89.
- Luca, M. (2016). Reviews, reputation, and revenue: The case of Yelp. com.
- Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62(12), 3412-3427.
- Mackay, K. J., & Crompton, J. L. (1988). A conceptual model of consumer evaluation of recreation service quality. *Leisure Studies*, 7(1), 40-49.
- Mahajan, V., Muller, E., & Kerin, R. A. (1984). Introduction strategy for new products with positive and negative word-of-mouth. *Management Science*, 30(12), 1389-1404.

- McAuley, J., & Leskovec, J. (2013, October). Hidden factors and hidden topics: understanding rating dimensions with review text. In Proceedings of the 7th ACM conference on Recommender systems (pp. 165-172). ACM.
- Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, 50, 67-83.
- Ramanathan, U., & Ramanathan, R. (2011). Guests' perceptions on factors influencing customer loyalty: An analysis for UK hotels. *International Journal of Contemporary Hospitality Management*, 23(1), 7-25.
- Tripp, C., Greathouse, K. R., Shanklin, C. W., & Gregoire, M. B. (1995). Factors influencing restaurant selection by travelers who stop at visitor information centers. *Journal of Travel & Tourism Marketing*, 4(2), 41-50.
- Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services. *International Journal of Service Industry Management*, 15(3), 302-326.
- Yelp.com (2018). *The Best 10 Restaurants in Chicago, IL*. Retrieved from https://www.yelp.com/search?find_loc=Chicago,+IL&cflt=restaurants.
- Zhang, Z., Ye, Q., Law, R., & Li, Y. (2010). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, 29(4), 694-700.