

Predicting the Characteristics of Trending Vegetarian Restaurants

Katherine Ann Willyard

April 7, 2018

1. Introduction

Many restaurants fail within the few years of their opening. For example, according to [Business Insider](#), 80% of New York restaurants close within their first five years. However, potential restaurant owners can reduce the likelihood of their failure by using geographic data on successful restaurants in their city. The purpose of this project is to provide potential vegetarian City restaurant entrepreneurs recommendations on where they should open by looking at the relationship between trending restaurants and the characteristics of surrounding top vendors. This information is important for entrepreneurs to decide where they should locate a vegetarian restaurant in their city.

2. Data

This project uses data on New York City neighborhoods and the vendors located in these locations. A list of New York neighborhood names from 2014 were collected from New York University's Spatial Data Repository at https://geo.nyu.edu/catalog/nyu_2451_34572. This dataset involves all 5 New York City boroughs consisting of 306 neighborhoods. A list of neighborhood vendors were collected from Foursquare. For each neighborhood, the top vendors within a 500 meter radius from each neighborhood center (a total of 10,298 vendors), are explored.

3. Methods

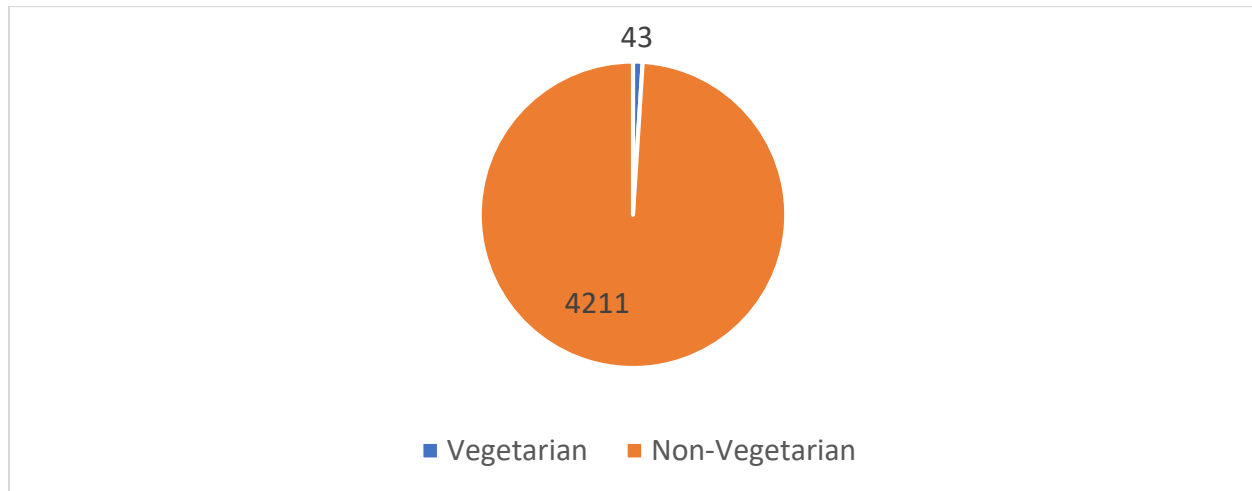
This project involved several exploratory analysis at the restaurant and neighborhood level. Several summary statistics were examined. Additionally, I created a correlation matrix to examine relationships between the number of vegetarian restaurants in a neighborhood and the number of other types of venues including: other restaurants, offices, transportation, retail, nightlife, entertainment, and total venues.

I also conducted inferential statistical testing, specifically the machine learning technique multiple regression. This technique was used because it best describes correlations between two variables. Analysis was conducted at the restaurant and neighborhood level. At the restaurant level, since the dependent variable is dichotomous, logit regression analysis was conducted. At the neighborhood level, since the dependent variable was not dichotomous, ordinary least squares regression analysis was used.

4. Results

As shown in Figure 1, few trending restaurants are vegetarian restaurants. Of the 4,254 trending venues that are restaurants, only 43 (1%) were vegetarian.

Figure 1: Portion of Trending Vegetarian Restaurants



Vegetarian restaurants, in comparison to non-vegetarian restaurants, tend to be located in neighborhoods with more venues, more nightlife, and fewer other nearby restaurants.

As shown in Table 1, on average, while non-vegetarian restaurants have 2.85 nightlife venues, vegetarian venues have 5.45 nightlife venues. The average number of nearby venues for non-vegetarian restaurants is 61.6 and the average number of nearby venues for vegetarian restaurants is 88.16. The average number of nearby other restaurants for non-vegetarian restaurants is 0.80 and the average number of other restaurants for vegetarian restaurants is 0.47.

Table 1: Restaurant Summary Statistics

| | Other Restaurants | Offices | Transportation | Retail | Residential | Nightlife | Entertainment | Total Vendors |
|----------------|-------------------|---------|----------------|--------|-------------|-----------|---------------|---------------|
| Non-Vegetarian | 26.50 | 2.36 | 0.80 | 10.85 | 0.83 | 2.86 | 4.97 | 61.60 |
| Vegetarian | 34.51 | 3.21 | 0.47 | 15.19 | 1.12 | 5.37 | 7.02 | 88.16 |

As shown in Table 2, there is a significant positive relationship between restaurants that are vegetarian and the number of nearby trending nightlife venues, and total trending venues. Holding all other variables constant, for each additional nearby trending nightlife venue, the log odds a restaurant is vegetarian (versus non-vegetarian) increases by 0.1791. For each additional nearby trending venue, the log odds a restaurant is vegetarian (versus non-vegetarian) increases by 0.0978. There is also a significant negative relationship between restaurants that are vegetarian and the number of nearby other restaurants, offices, transportation, retail, and entertainment. Holding all other variables constant, for each additional nearby trending other restaurant, the log odds a restaurant is vegetarian (versus non-vegetarian) decreases by 0.2837. For each additional nearby trending office, the log odds a restaurant is vegetarian (versus non-

vegetarian) decreases by 0.2348. For each additional nearby trending transportation venue, the log odds a restaurant is vegetarian (versus non-vegetarian) decreases by 1.6075. For each additional nearby trending retail space, the log odds a restaurant is vegetarian (versus non-vegetarian) decreases by 0.1664. For each additional nearby trending entertainment space, the log odds a restaurant is vegetarian (versus non-vegetarian) decreases by 0.2523.

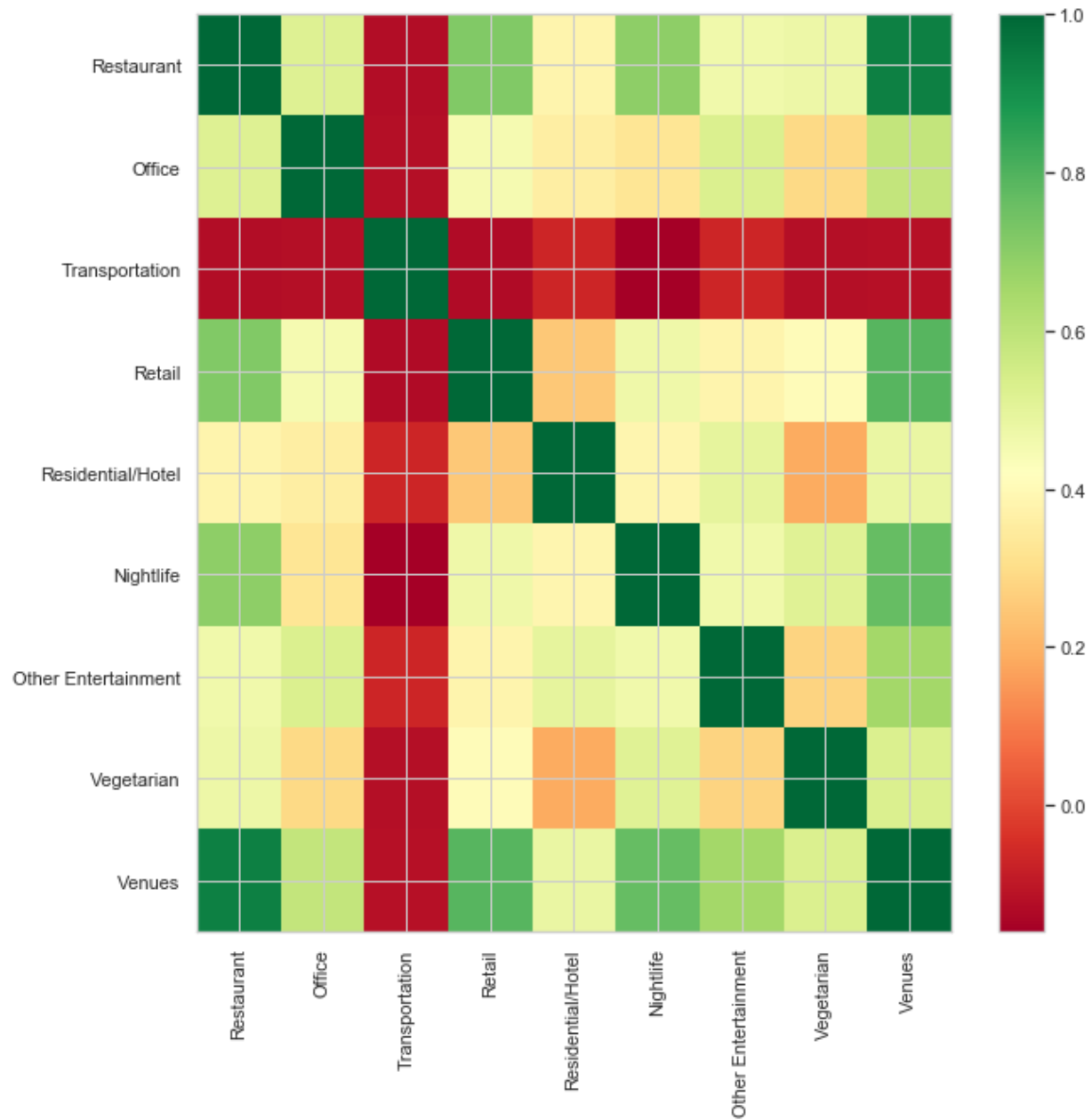
Table 2: Logit Regression Results

| | | | |
|---------------------|------------------|-------------------|----------|
| Model: | Logit | Pseudo R-squared: | -0.630 |
| Dependent Variable: | Vegetarian | AIC: | 799.7277 |
| Date: | 2019-04-07 19:38 | BIC: | 850.5726 |
| No. Observations: | 4254 | Log-Likelihood: | -391.86 |
| Df Model: | 7 | LL-Null: | -240.34 |
| Df Residuals: | 4246 | LLR p-value: | 1.0000 |
| Converged: | 1.0000 | Scale: | 1.0000 |
| No. Iterations: | 10.0000 | | |

| | Coef. | Std.Err. | z | P> z | [0.025 | 0.975] |
|---------------|---------|----------|---------|--------|---------|---------|
| Restaurants | -0.2837 | 0.0305 | -9.3115 | 0.0000 | -0.3435 | -0.2240 |
| Offices | -0.2348 | 0.0938 | -2.5043 | 0.0123 | -0.4186 | -0.0510 |
| Transport | -1.6075 | 0.2119 | -7.5870 | 0.0000 | -2.0228 | -1.1922 |
| Retail | -0.1664 | 0.0347 | -4.7910 | 0.0000 | -0.2345 | -0.0983 |
| Residential | 0.0652 | 0.1122 | 0.5813 | 0.5610 | -0.1547 | 0.2851 |
| Nightlife | 0.1791 | 0.0578 | 3.1011 | 0.0019 | 0.0659 | 0.2923 |
| Entertainment | -0.2523 | 0.0584 | -4.3174 | 0.0000 | -0.3668 | -0.1377 |
| Venues | 0.0978 | 0.0201 | 4.8565 | 0.0000 | 0.0583 | 0.1373 |

As shown in Figure 2, neighborhoods with more vegetarian restaurants also have more other restaurant types, more nightlife venues, and more total vendors. Also, neighborhoods with fewer vegetarian restaurants have more residential spaces, and other entertainment venues.

Figure 2: Neighborhood Venues Correlations Heat Map



As shown in Table 3, neighborhoods with more successful vegetarian restaurants have more venues, and fewer other restaurants. Holding all other variables in the model constant, for every additional other restaurant, the expected number of vegetarian restaurants decreases by 0.0009. Also, for every additional trending venue, the expected number of vegetarian restaurants increases by 0.0006, holding all other variables in the model constant.

Table 3: Ordinary Least Squares Regression Results

| | | | | | | |
|---------------------|------------------|---------------------|-------------|--------|---------|---------|
| Model: | OLS | Adj. R-squared: | 0.017 | | | |
| Dependent Variable: | Vegetarian | AIC: | -7539.3277 | | | |
| Date: | 2019-04-07 19:51 | BIC: | -7494.8384 | | | |
| No. Observations: | 4254 | Log-Likelihood: | 3776.7 | | | |
| Df Model: | 7 | F-statistic: | 11.67 | | | |
| Df Residuals: | 4247 | Prob (F-statistic): | 8.90e-15 | | | |
| R-squared: | 0.019 | Scale: | 0.0099338 | | | |
| ----- | | | | | | |
| | Coef. | Std.Err. | t | P> t | [0.025 | 0.975] |
| ----- | | | | | | |
| Restaurants | -0.0009 | 0.0003 | -2.7385 | 0.0062 | -0.0015 | -0.0003 |
| Offices | 0.0002 | 0.0009 | 0.1648 | 0.8691 | -0.0016 | 0.0019 |
| Transport | -0.0007 | 0.0012 | -0.5417 | 0.5881 | -0.0031 | 0.0018 |
| Retail | -0.0000 | 0.0003 | -0.0538 | 0.9571 | -0.0007 | 0.0007 |
| Nightlife | 0.0007 | 0.0008 | 0.8988 | 0.3688 | -0.0009 | 0.0023 |
| Entertainment | -0.0009 | 0.0005 | -1.7654 | 0.0776 | -0.0019 | 0.0001 |
| Venues | 0.0006 | 0.0002 | 2.4807 | 0.0132 | 0.0001 | 0.0011 |
| ----- | | | | | | |
| Omnibus: | 6429.367 | Durbin-Watson: | 1.989 | | | |
| Prob(Omnibus): | 0.000 | Jarque-Bera (JB): | 1575518.300 | | | |
| Skew: | 9.664 | Prob(JB): | 0.000 | | | |
| Kurtosis: | 95.277 | Condition No.: | 66 | | | |

5. Discussion

Results lead to several recommendations.

First, restaurant entrepreneurs should locate vegetarian restaurants in areas where there is a large number of trending venues. I make this recommendation because there is a consistent positive correlation between vegetarian restaurants and the number of total nearby trending venues.

Second, restaurant entrepreneurs should locate vegetarian restaurants in areas where there is more nightlife. I make this recommendation because trending restaurants that are vegetarian are in areas with more nightlife. There is also a positive correlation between the number of vegetarian restaurants in a neighborhood and the number of trending nightlife venues.

My final recommendation is that restaurant entrepreneurs locate vegetarian restaurants in areas where there are fewer other trending restaurants. I make this recommendation because, when examining both restaurant and neighborhood regression models, there is a significant negative relationship between trending vegetarian restaurants and other restaurants.

6. Conclusion

In this study, I analyzed the relationship between trending vegetarian restaurants, and the characteristics of surrounding vendors at the restaurant and neighborhood level. I found that trending restaurants are more likely to be vegetarian when they are located in neighborhoods with more venues, and more nightlife, and fewer other restaurants. I also found that neighborhoods with more trending vegetarian restaurants have more venues and fewer other restaurants. This led me to recommend that vegetarian restaurants be located in neighborhoods with a large number of venues, more nightlife, and fewer other restaurants. This information can be used by entrepreneurs to decide where they should locate a vegetarian restaurant in their city.