Where to open a vegetarian restaurant in New York City?

Introduction

The purpose of this project is to determine where a new restaurant should be opening in New York City. According to a study by Ohio State University, 60% of new restaurants do not make it past their first year. (https://journals.sagepub.com/doi/abs/10.1177/0010880405275598).

Consequently entrepreneurs need to use existing data to give themselves a geographical advantage as the industry is competitive. Additionally the characteristics of successful restaurants will be examined. The result will be a recommendation of where and what type of restaurant new entrepreneurs should open in New York City.

Data

New York University's Spacial Data Repository (https://geo.nyu.edu/catalog/nyu_2451_34572.) contained data on 5 boroughs and 306 neighbourhoods.

Foursquare was utilised to provide neighbourhood vendors. A circle of influence of 500 meters was explored for each neighbourhood.

Methodology

At the neighbourhood and restaurant level exploratory analysis was conducted and summary level statistics were examined.

Multiple Linear regression was performed (machine learning technique) in order to determine the correlation between variables.

At the neighbourhood level, Ordinary Least Squares (OLS) regression analysis was performed as the dependent variable was not dichotomous.

At the restaurant level, Logit regression analysis was performed as the dependent variable was dichotomous.

Results

Of the 4,017 trending venue that are restaurants, only 34 (1%) were vegetarian. Vegetarian restaurants in comparison to non-vegetarian restaurants, tend to be located in neighbourhoods with more venues, more nightlife, and fewer other nearby restaurants.

As shown in table 1, on average, while non-vegetarian restaurants have 2.32 nightlife venues, vegetarian venues have 3.91 nightclub venues.

The average number of nearby venues for non-vegetarian restaurants is 57.20 and the average number of nearby venues for vegetarian restaurants is 76.01.

The average number of nearby other restaurants for non-vegetarian restaurants is 0.867 and the average number of other restaurants for vegetarian restaurants is 0.67.

Table 1: Restaurant summary statistics

	Restaurants	Offices	Transport	Retail	Residential	Nightlife	Entertainment	Venues
Vegetarian								
0	24.688677	2.390911	0.866934	9.820989	0.748431	2.320362	4.628170	57.197088
1	29.352941	3.117647	0.676471	12.176471	1.029412	3.911765	6.735294	76.088235

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 0.2467.

For each additional nearby trending venue, the log odds a restaurant is vegetarian (versus non-vegetarian) increases by 0.0938. 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 0.3087.

For each additional nearby trending office, the log odds a restaurant is vegetarian (versus non-vegetarian decreases by 0.3324.

For each additional nearby trending transportation venue, the log odds a restaurant is vegetarian (versus non-vegetarian) decreases by 1.3836.

For each additional nearby trending retail space, the log odds a restaurant is vegetarian (versus0 non-vegetarian) decreases by 0.1792.

For each additional nearby trending entertainment space, the log odds a restaurant is vegetarian (versus non-vegetarian) decreases by 0.1767.

Table 2" Logit Regression Results

```
Optimization terminated successfully.
        Current function value: 0.079988
        Iterations 10
                        Results: Logit
=================================================================
Model:                Logit          Pseudo R-squared: -0.638
Dependent Variable: Vegetarian       AIC:              656.6206
Date:                 2020-04-22 20:05 BIC:            700.7086
No. Observations:     4017           Log-Likelihood:   -321.31
Df Model:             6              LL-Null:          -196.10
Df Residuals:         4010           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    -0.3087   0.0400  -7.7243 0.0000  -0.3870 -0.2304
Offices        -0.3324   0.1040  -3.1955 0.0014  -0.5363 -0.1285
Transport      -1.3836   0.1967  -7.0346 0.0000  -1.7691 -0.9981
Retail         -0.1792   0.0538  -3.3332 0.0009  -0.2845 -0.0738
Nightlife       0.2467   0.1304   1.8922 0.0585  -0.0088  0.5022
Entertainment  -0.1767   0.0635  -2.7815 0.0054  -0.3012 -0.0522
Venues          0.0938   0.0308   3.0464 0.0023   0.0335  0.1542
=================================================================
```

As shown in Table 3, neighbourhoods with more successful vegetarian restaurants have more successful vegetarian restaurants have more venue, and fewer other restaurants. Holding all other variables in the model constant, for every additional restaurant, the expected number of vegetarian restaurants decreases by 0.0008. Also for every additional trending venue, the

expected number of vegetarian restaurants increases by 0.0006, holding all other variables in the model constant.

```
              Results: Ordinary least squares
=================================================================
Model:                 OLS              Adj. R-squared:      0.012
Dependent Variable: Vegetarian          AIC:                 -7809.5573
Date:                  2020-04-22 20:56 BIC:                 -7765.4692
No. Observations:      4017             Log-Likelihood:      3911.8
Df Model:              7                F-statistic:         7.820
Df Residuals:          4010             Prob (F-statistic):  1.96e-09
R-squared:             0.013            Scale:               0.0083646
-----------------------------------------------------------------
              Coef.    Std.Err.     t      P>|t|     [0.025    0.975]
-----------------------------------------------------------------
Restaurants   -0.0008   0.0003  -2.5093   0.0121   -0.0015   -0.0002
Offices        0.0001   0.0009   0.1688   0.8660   -0.0016    0.0019
Transport     -0.0003   0.0012  -0.2377   0.8122   -0.0026    0.0021
Retail        -0.0002   0.0004  -0.5279   0.5976   -0.0010    0.0006
Nightlife      0.0002   0.0009   0.2196   0.8262   -0.0016    0.0020
Entertainment -0.0007   0.0005  -1.3556   0.1753   -0.0016    0.0003
Venues         0.0006   0.0002   2.3512   0.0188    0.0001    0.0011
-----------------------------------------------------------------
Omnibus:              6412.289    Durbin-Watson:           1.967
Prob(Omnibus):        0.000       Jarque-Bera (JB):        2167401.848
Skew:                 10.633      Prob(JB):                0.000
Kurtosis:             114.790     Condition No.:           61
=================================================================
```

Conclusion

The results lead to several recommendations;

1. 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 trending nearby.

2. 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 neighbourhood and the number of trending nightlife venues.

3. Restaurant entrepreneurs locate vegetarian restaurants in areas where there are fewer other trending restaurants, I make this recommendation because when assessing both restaurants and neighbourhood regression models, there is a significant negative relationship between trending vegetarian restaurants and other restaurants.

In this study, I analysed the relationship between trending vegetarian restaurants, and the characteristics of surrounding vendors at the restaurant and neighbourhood level.

I discovered that trending restaurants are more likely to be vegetarian when they are located in neighbourhoods with more venues, more nightlife and fewer other restaurants, I also found that neighbourhoods with more trending vegetarian restaurants have more venues and fewer other restaurants,

This resulted in my recommendation that vegetarian restaurants be located in neighbourhoods with a large number of venues, more nightlife and fewer restaurants. The information can be used by entrepreneurs to decide where they should locate a vegetarian in their city.