**Vegan Restaurant Venture**

1. **Introduction**

In this project, I evaluate the potential of three different cities – Philadelphia, New York, and Boston – for the purpose of opening a vegan restaurant venture. I utilize Foursquare, extract restaurant data, categorize that data, and determine which market is most suitable for such a business. To do so, I use restaurant likes as a proxy for business success and performance. I build linear and logistic regressions to evaluate the strength of the model and make the best-informed and wisest business decision based on the generated results. The project could be helpful to sustainable business enthusiasts.

1. **Data**

I use Foursquare data to retrieve the coordinates of the three cities mentioned above - Philadelphia, New York, and Boston. Then, I take advantage of the Foursquare API to get information on restaurants' 'name', 'categories', 'latitude', 'longitude', and 'id' for each city. I have decided to focus only on restaurants that are within 1,000 km of the three major cities I explore. Since the data also results in non-restaurant venues, I manually clean the data and leave only restaurant information needed.

1. **Methodology**

I have used both linear and logistic regression to train and test the data. Linear regression is most appropriate for predicting the number of “likes” a new restaurant in the above mentioned cities will acquire, while logistic regression is most appropriate for classifying the method. To classify by number of “likes”, I bin the data and then multinomial logistic regression to perform the analysis. For both methods, I have used the Ski-Kit Learn package.

1. **Results**

The first step has been to test a linear regression model on a random subsample of 80% and use the other 20% for testing purposes. To evaluate if the model is reasonable, I look at both the sum of squares and the variance calculated as 43503.22 and -0.04. The variance is negative which indicates that is not the best manner in which we can monitor the data. Thus, I have also utilized logistic regression.

The logistic regression model is trained on the same subsample and then is tested on the remaining 20%. The Jaccard score and log loss are 36.61% and 0.9494. The Jaccard score is at a reasonable enough level.

I further run the model on the complete dataset. The coefficients I get show that opening a restaurant that is Asian, casual, European, Latin, and located in Boston is negatively associated with likes. Therefore, I conclude that an American restaurant serving vegan food options and located in either Philadelphia or New York will generate the highest performance.

Here are some pictures of the more successful vegan restaurants in Philadelphia and New York.

****

****

1. **Discussion**

As already mentioned above logistic regression represents a better fir for the data over linear regression. Using logistic regression, we have found that the Jaccard score is 36.61%, which is more reasonable than the variance computed. I have used number of “likes” as a proxy for a restaurant’s performance in terms of brand and image. With the rising demand and interest in vegan products, I believe that the number of “likes” is an appropriate proxy for the potential of a vegan restaurant. However, it is worth noting that the number of “likes” can bring some bias to the results since the project itself has limitations due to the amount of data that can be fetched from the FourSquare API.

I proceed to break down the results of the logistic regression model. Based on the precision scores, I can conclude that this model is good at predicting if a restaurant will fall in to the best or worst percentile of likes. The three-bin method used has been able to yield the best Jaccard score.

In addition to predicting the performance of the restaurant, I have also attempted to get insights on the business strategy to be used. The coefficients generated suggest that the business to be opened is an American restaurant serving vegan options.

1. **Conclusion**

In conclusion, after analyzing 300 restaurants in some of the biggest cities in the U.S. – Philadelphia, New York, and Boston – I find that the best approach to maximize performance is to open a restaurant with American vegan cuisine in either Philadelphia or New York. Given the rise in vegan options and the awakening of people, I believe that a vegan restaurant will be a profitable venture if executed with precision.