

Opening of new restaurants in Manhattan, NYC and Toronto: A comparative analysis

Felix Milan

Contents

1	Introduction	1
2	Data acquisition	1

1 Introduction

Restaurants in metropolitan cities are not just places to dine, but also an opportunity for guests to experience new cultures via a variety of cuisines. Thus, it would be convenient to explore the districts where particular restaurants are located, when opening a new restaurant in a city. Incidentally, several factors contribute to the success of a new restaurant in a specific location [1]. We limit ourselves to the environment factor of the location, i.e. we try to find suitable locations with a high number of restaurants in their vicinity, since this may increase the number of potential customers [1]. We choose Manhattan, NYC and Toronto as the cities to perform the location data analysis project on, because both resemble each other in a few aspects. The business problem in this project maybe stated as follows: Can we find a suitable data analysis algorithm to predict whether it is more profitable to open a restaurant in either Manhattan, NYC or Toronto, and which neighbourhoods provide the best location? For this question we will also consider the kind of the restaurant in question, e.g. gourmet or street food, and the potential customers. The target audience in question are both gastronomical entrepreneurs and large and low scale investors, as the location data is useful for both small and big businesses. Since a restaurant's location can be crucial to its initial and long term success, location data analysis studies are crucial for restaurant owners and their investors.

2 Data acquisition

The following data will be used in the analysis of this project:

1. a json file containing the geospatial data of New York City [2]
2. a list of postal codes for neighborhoods in Toronto taken from Wikipedia [3]
3. a csv file containing the longitude and latitude of the Toronto neighbourhoods [4]

The json file containing the geospatial data of New York City [2] is important in order to extract the Manhattan neighbourhoods alongside the latitude and longitude coordinates. In this way we can obtain the corresponding venues via a Foursquare API “explore” query. In order to obtain the geospatial data of Toronto, we scrape a Wikipedia page [3] to obtain a table of postal codes and neighbourhood names of Toronto, using the BeautifulSoup package of Python. Then, we merge this table by the geospatial location data of each Toronto neighbourhood given in a csv file [4]. Finally, we can use a Foursquare API query to obtain the venues for each Toronto neighbourhood. For both the venues in Manhattan and in Toronto we will limit ourselves to “restaurant” category, since we are only interested in location data regarding restaurants.

References

- [1] <https://www.entrepreneur.com/slideshow/299849>
- [2] https://coc1.us/new_york_dataset
- [3] https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- [4] https://coc1.us/Geospatial_data