

Battle of the Neighborhoods Applied Data Science Capstone by IBM

Exploring Downtown Eateries in Boston



1. Introduction

1.1 Background

The city of Boston was founded in 1630. The largest of the city's commercial districts, Downtown is the location of many corporate or regional headquarters; city, county, state and federal government facilities; and many of Boston's tourist attractions. Similar to other central business districts in the U.S., Downtown has recently undergone a transformation that included the construction of new condos and lofts, renovation of historic buildings, and arrival of new residents and businesses.

The term Downtown refers to main business, cultural and often the historical, political and geographic heart and commercial area of a town or city. Boston's downtown is majorly occupied by the Financial District.

Boston's downtown is extremely busy during the day as hundreds of thousands of office workers and tourists descend on the area. Though many of Boston's swankiest restaurants are located Downtown, budget restaurants tend to be geared towards the horde of office workers who descend on the city during the week so many of them are only open for weekday lunch.

1.2 Business Problem

I spent the first 8 months of 2019 working for a financial company in Boston, MA. My office was located right in the heart of the city, Financial District aka the Downtown. Every day, me and other interns would ask the same question – where to eat lunch? There are hundreds of eateries around the Financial District within a distance of 15-minute walk. Hence, as a result, I decided to explore the Financial District for food joints.

1.3 Interest

This project can be used by people working in the Financial District, Boston; specially the students who are interning or recent graduates who don't tend to cook and bring lunches to work-place. People can use this app to try different eateries according to cuisines and popularity; thus avoiding the same old routine places!

2. Data

2.1 Data Sources

Since I am exploring the city of Boston, I fetched the geo data from various sources as there was no dataset available online featuring all the neighborhoods along with their coordinates.

Files containing coordinates were download from following websites:

<https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/>

<https://simplemaps.com/data/us-zips>

2.2 Data Cleaning

One of the files contained Boston Neighborhoods along with their Zip codes. The other file was filtered and only contained Zip codes with their geographical coordinates for Boston. Both the files were merged in Microsoft Excel using the VLOOKUP function on the Zip Codes.

Duplicates were removed manually. Some neighborhoods were still missing coordinates which were entered manually fetching from www.google.com.

Foursquare API is used to fetch the numerous venues around various neighborhoods in Boston.

3. Data Analysis

3.1 Methodology

In this project, I explored all the food serving joints in Boston's Financial District which is a major part of Boston Downtown. For Step 1, I started the project with data collection. As stated earlier, data files from multiple websites were used. It was cleaned, merged and missing values were manually entered. As a result, a dataframe containing Boston Neighborhoods and their coordinates was prepared to get the project running.

I began Step 2 by initializing the Foursquare API to get the venues around the Financial District. To do so, I used my Foursquare API credentials, created URLs and requested venue-search calls to the API to get all the venues in downtown area. I set a radius to 1000m around its center so that all of the downtown venues are covered by the API while fetching the venues. I dataframe was prepared with all the food joints with their categories and coordinates.

Step 3 was to fetch the restaurants details like Ratings, Likes and Tips so that I could do deep analysis using these attributes. To fetch restaurant details, Premium API calls have to be made to the Foursquare API. Since just 50 calls are allowed in a day, it took me multiple days to fetch details about all 170 restaurants. I broke the dataframe into groups of 50, so that all restaurant details could be fetched over the span of several days. The final dataframe was created by concatenating the mini-dataframes and finally the restaurants details dataframe was built! Finally, the dataframes with geo details and restaurants details were joined on basis of 'restaurant id' and the final dataframe was created.

Step 4 was the analysis of the data (finally!). I started the analysis with exploring the venues based on ratings, tips and likes. This was followed by doing some geo-analysis of the restaurants using Folium maps. All the restaurants were plotted on a map to get a general sense of their locations. After this, I plotted the top restaurants and most favorite restaurants. Next type of analysis that I performed was Clustering. I clustered the restaurants based on their location and ratings.

4. Result and Conclusion

The analysis shows that there are about 170 food serving venues in and around Financial District having an average Rating of 8.47/10. Sporting more than 20 different cuisines with a minimum rating of 8.0, the most popular is Italian Cuisine with 27 restaurants, followed by Seafood

Restaurants. It is very interesting to observe that there are almost twice as many Italian restaurants as Seafood Restaurants in spite of Boston being a Coastal City.

Let's talk about a few restaurants which stand out in particular. Mike's Pastry is the most liked food joint in the Downtown with the most number of Tips as well! This place definitely becomes the must visit venue. Trillium Garden On The Greenway has got the best Ratings of all the restaurants in downtown area. Most of the highly rated restaurants are kind of evenly spread out throughout the region. So, if you are hungry, you can get one of the best foods in the city within a 10 minutes' walk. Some of the top liked places are mostly either in the northern half of the downtown i.e. near North End along with a few places near Chinatown and Seaport. 30% of the most liked restaurants serve Seafood.

Above analysis gives us a good view and trend of Downtown restaurants. The primary target audience for this project were employees working in the Financial District - which can include students as young as 21 years along with professional veterans with over 30 years of experience. This project is not restricted to just the target audience! Tourists can also make a good use of it. They can get to know the best and must visit places along with their locations so that they can plan their day even better.