**Capstone Project Final Report**

**Introduction/Business Problem**

I own a fictional travel agency, and a client of mine is looking for a good destination for him and his wife. They want to travel to a city close to them and their two options are New York City and Toronto. Since both cities are big metropolises with storied histories and lots of things to see and do, it is a hard decision to make. However, their main issue is food, as they are vegan and don't always find good or unique vegan food on their travels. As a result, they want to decide based on which city is more vegan friendly, and therefore would be more welcoming for their dietary restrictions.

It also has the benefit of giving my clients a list of all the vegan/vegetarian restaurants in their city of choice, thus eliminating the headache of trying to find and choose vegan/vegetarian restaurants.

**Data**

Based on the definition of our problem, the factors that will influence my client's decision are:

1. The number of vegan/vegetarian restaurants in each city
2. Which city has the greater number of vegan/vegetarian restaurants.

For this analysis, I wanted to present both a list of all the restaurants in each city, as well as a map of all the restaurants in each city, to be able to properly visualize the number of vegan/vegetarian restaurants in each city and to provide a list that would make it easy for my clients to plan out their trip.

We will require the following for this analysis:

1. A list of all the neighborhoods in each city, along with the approximate latitude and longitudes.
2. A list of venues generated using the foursquare places API that will allow us to see what kinds of venue categories exist, and then narrow that down further to obtain a list of all the venues that are under the category 'Vegetarian / Vegan Restaurant'.
3. The coordinates of each city obtained using the Google Maps API geocoder package, which we will then use to create a base map.
4. A map of all the venues that we filtered out earlier, displayed in such a way that will allow my clients to see the relative position on the map to be able to use as a reference when planning stops to visit on their trip, to make sure that there is a place to eat for them near any sites they want to see.

**Below are snapshots of lists of each neighborhood, Borough, Latitude and Longitude in New York City and Toronto respectively.**

Table

Description automatically generated with low confidence

Graphical user interface, text, application

Description automatically generated

The data returned from the API call will contain the Neighborhood, Neighborhood Latitude, Neighborhood Longitude, Venue, Venue Latitude, Venue Longitude, Venue Category for each City.

A screenshot of a computer

Description automatically generated with medium confidence

Then I will reduce that down to a list of all the vegetarian/vegan restaurants in the area, and then map the locations for each venue in each city, to see not only which city has more vegan restaurants, but also which city has more variety.

**Methodology**

**Step 1: Obtaining the neighborhood datasets for each city, which include the Borough, Neighborhood, Latitude and Longitude of each neighborhood.**

For each city our approach will differ, primarily due to:

1. the location of each of the datasets

2. how much of the data is missing

3. how much preprocessing we will need to do

4. how much of the data will we need to clean and/or eliminate.

For New York City, we will obtain the data by grabbing the URL for the dataset that was given to us by the course admins for the previous lab. We will grab the raw JSON from the URL using the requests package. From the JSON, we will isolate the data that we need, which is under 'features' in the JSON dict. Once we have that sorted, we will create a data frame with the columns Borough, Neighborhood, Latitude and Longitude, and then load the data using a for loop.

For the Toronto Neighborhood data, our approach will differ because we are scraping data from two different sources and will require us to merge the data at the end to obtain a similar Neighborhood table with the same columns as the New York City Neighborhood Data.

Step 1 should result in the following tables.

Table

Description automatically generated with low confidence

Graphical user interface, text, application

Description automatically generated

**Step 2: Obtaining the Foursquare venue tables for each city**

After prepping the Foursquare Credentials to grab the venue data for each city, we use the getNearbyVenues function that we used in the lab to grab the list of venues for each city.

**NOTE: if there is a key error: ‘groups’ when running the code in the notebook, it might be because the number of calls I can make with my free foursquare account is limited and as a result the quota might have been exceeded. I couldn’t post photos of my venue data nor my map data as a result.**

A screenshot of a computer

Description automatically generated with medium confidence

This will result in two data frames containing the venues for each city.

**Step 3: Filtering the venue datasets to obtain only vegan/vegetarian restaurants in the city.**

Graphical user interface, website

Description automatically generatedThis is a lot of information to deal with, and to figure out which venues are vegan/vegetarian restaurants, we need to filter the data based on the venue category parameterto obtain a list of all the vegan/vegetarian restaurants in each city.

**Step 4: Mapping out the venues for each city.**

We need to map out the venues for each city, to provide a better visual analysis for my clients

We need to first use the geolocator package to obtain the geographical coordinates of each city, which we will then use to create a base map for each city.

Once we have the filtered data and the geographic coordinates, we then use the folium map package to first generate a map of each city, and then map the data points on each of the city maps.

Map

Description automatically generated**Map of New York City with Vegan Restaurants marked.**

Map

Description automatically generated

**Map of Toronto with Vegan Restaurants marked.**

Then we look at the marked points and based on those points see which city has more vegan/vegetarian restaurants.

Results

Based on the number of restaurants and the map visualization of this data, the clear winner in this case is New York City, and I would recommend New York City to my clients for travelling.

Discussion

When I was looking at the data, I realized that because I only focused on purely vegan/vegetarian restaurants and didn’t include restaurants which might have vegan/vegetarian options in addition to their non-vegetarian options, I might have eliminated some viable options for my clients. However, based on the sheer number of vegetarian options, I believe that even if they wanted a little more variety, there are a ton of options on the list and that NYC is very, very vegan friendly.

Conclusion

I can confidently recommend New York City as the better option for my clients to travel to. For the future and can use this approach to evaluate other cities for vegan friendliness and travel viability. I could also adjust the approach to check for restaurants that match the customers preferred cuisine and/or create food-based tours where the main focus is travelling and sampling famous local joints and eateries, thus making my package options very diverse, while offering a special theme that other places might not have.