Lars S. Madsen
Applied Data Science Capstone
May 2020

Culinary Comparison between the 10 largest cities in USA

What is the difference between the largest american cities when it comes to restaurants? The answer to this could be of interest, not only to the restaurant sector itself, but also to the tourism sector, healthcare, journalists and the government. Here the fictional Tourism Council of US Cities have requested basic research that can be used for a series of articles about the food scene in the largest cities. In this report I will look into the population size, the ethnic composition, income information, geographic location and the restaurant types in the 10 largest cities in USA.

The data that will be used in this report is information about the 10 largest cities sourced form wikipedia.org combined with detailed restaurant information from foursquare.com. For ease I will define a city as the location and a 20 mile radius (approximately the radius of New York City).

The data from Wikipedia will be scraped using Beautifulsoup while the Foursquare information will be sourced through their APIs. On the following two pages I will dive deeper into the data required.

The Foursquare API makes it easy to extract the needed information for a given location. For example:

https://api.foursquare.com/v2/venues/search?near=%27New%20York, %20USA%27&categoryId=4d4b7105d754a06374d81259

Which yields a JSON response beginning with this information:

```
{"meta":
{"code":200,"requestId":"5ebfe2e43907e7001bb7b735"},"resp
onse":{"venues":
[{"id":"4c5ef77bfff99c74eda954d3","name":"Eataly
Flatiron","location":{"address":"200 5th
Ave","crossStreet":"at W 23rd
St","lat":40.74198696405861,"lng":-73.98991319280194,"lab
eledLatLngs":
[{"label":"display","lat":40.74198696405861,"lng":-73.989
91319280194}],"postalCode":"10010","cc":"US","city":"New
York","state":"NY","country":"USA","formattedAddress":
["200 5th Ave (at W 23rd St)","New York, NY
10010","USA"]},"categories":
[{"id":"4bf58dd8d48988d1f5941735","name":"Gourmet
Shop","pluralName":"Gourmet
```

… which can be traversed through using any of the libraries for JSON interpretation.

From the Wikipedia page by using Beautifulsoup we can get population size, area size, income information and ethnographic composition data for the cities in the comparison.

The data will be processed using python in a Jupyter Notebook which will be provided along with this report. The notebook will also be used to generate informative charts and maps using the libraries Matplotlib and Folium.

In summary the data used will be:

Source	Content	Structure
https://en.wikipedia.org/wiki/ List_of_United_States_cities_by_ population	Top City List	Table
Wikipedia (for example https:// en.wikipedia.org/wiki/ New_York_City)	Population count	Table
Wikipedia (for example https:// en.wikipedia.org/wiki/ New_York_City)	Ethnographic information	Table

Source	Content	Structure
https://en.wikipedia.org/wiki/List_of_New_York_locations_by_per_capita_income	Income information	Table
Wikipedia (for example https://en.wikipedia.org/wiki/New_York_City)	City location	latitude and longitude
Foursquare API	Restaurant details	JSON structure

So for example for New York City by calling Foursquare API (https://api.foursquare.com/v2/venues/explore?ll=40.7128,-74.0060&radius=32000&categoryId=4bf58dd8d48988d10a941735) we'll discover that there are 61 Ethiopian restaurants listed on Foursquare within 20 miles from the centre of New York City. In comparison there are 40 in Los Angeles. New York City has a population of 8.4 million whereas Los Angeles has a population of roughly 4 million. That means that Los Angeles has 10 Ethiopian restaurants per million while New York City only has just over 7!