Machine Learning

Exploring the Taste of Chicago Neighborhoods

Introduction

Background

Chicago is the most popular city in the United States, home to the headquarters of the United Nations and an important center for international diplomacy. It just might be the most diverse city on the planet, as it is home to over 8.6 million people and over 800 languages. As quoted in an article - What Food Tells Us About Culture "Traditional cuisine is passed down from one generation to the next. It also operates as an expression of cultural identity. Immigrants bring the food of their countries with them wherever they go and cooking traditional food is a way of preserving their culture when they move to new places."

Problem

Undoubtedly, Food Diversity is an important part of an ethnically diverse metropolis. The idea of this project is to categorically segment the neighborhoods of Chicago into major clusters and examine their cuisines. A desirable intention is to examine the neighborhood cluster's food habits and taste. Further examination might reveal if food has any relationship with the diversity of a neighborhood. This project will help to understand the diversity of a neighborhood by leveraging venue data from Foursquare's 'Places API' and 'k-means clustering' unsupervised machine learning algorithm. Exploratory Data Analysis (EDA) will help to discover further about the culture and diversity of the neighborhood.

Stakeholders

This quantifiable analysis can be used to understand the distribution of different cultures and cuisines over 'the most diverse city on the planet – Chicago'. Also, it can be utilized by a new food vendor who is willing to open his or her restaurant. Or by a government authority to examine and study their city's culture diversity better.

Data

To examine the above said, following data sources will be used:

Chicago Dataset

Link: https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago

There are sometimes said to be more than 200 neighborhoods in Chicago, though few residents would agree on their names and boundaries. A city ordinance prescribing and mapping 178 neighborhoods is almost unknown and ignored even by municipal departments. Neighborhood names and identities have evolved over time due to real estate development and changing demographics. The City of Chicago is also divided into 77 community areas which were drawn by University of Chicago researchers in the late 1920s. Chicago's community areas are well-defined, generally contain multiple neighborhoods, and are less commonly used by city residents. More historical images of Chicago neighborhoods can be found in Explore Chicago Collections, a digital repository made available by Chicago Collections archives, libraries and other cultural institutions in the city. Here is an example of json data :

```
{'type': 'Feature',
 'id': 'nyu_2451_34572.306',
 'geometry': {'type': 'Point',
 'coordinates': [-74.08173992211962, 40.61731079252983]},
 'geometry_name': 'geom',
 'properties': {'name': 'Fox Hills',
 'stacked': 2,
 'annoline1': 'Fox',
 'annoline2': 'Hills',
 'annoline3': None,
 'annoangle': 0.0,
 'borough': 'Staten Island',
 'bbox': [-74.08173992211962,
```

Foursquare API

Link: https://developer.foursquare.com/docs

Foursquare API, a location data provider, will be used to make RESTful API calls to retrieve data about venues in different neighborhoods. This is the link to Foursquare Venue

Category Hierarchy. Venues retrieved from all the neighborhoods are categorized broadly into 'Arts & Entertainment', 'College &University', 'Event', 'Food', 'Nightlife Spot', 'Outdoors & Recreation', etc. An extract of an API call is as follows:

```
'categories': [{'id': '4bf58dd8d48988d110941735',
'name': 'Italian Restaurant',
'pluralName': 'Italian Restaurants',
'shortName': 'Italian',
'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/food/italian_',
'suffix': '.png'},
'primary': True}],
'verified': False,
'stats': {'tipCount': 17},
'url': 'http://eccorestaurantny.com',
'price': {'tier': 4, 'message': 'Very Expensive', 'currency'
```