Applied Data Science Capstone Project

The Battle of Tokyo Neighborhoods Restaurants

Introduction/Business Problem

The objective of this capstone project is to help the travelers of Tokyo city to choose the best restaurant that fits their needs since Tokyo is well-known as the restaurant capital of the world with over 160,000 places , using data science methodology and machine learning techniques especially clustering, this project aims to provide solutions for this problem

Foursquare's slogan is "Foursquare helps you find the places you'll love, anywhere in the world". Since the launch of the Foursquare mobile app in 2009, Foursquare has helped 60 million users discover new exciting places worldwide [1]. The app provides personalized recommendations of places to visit in the vicinity of a user's current location based on "previous browsing history, purchases, or check-in history". [2] As a result, the Foursquare app has gained popularity for helping users to discover brand new places that match their interests. Using Foursquare API will help us collect the data that we need to resolve our Business Problem

Data section

For this project we need following data:
Tokyo data that contains list districts (Wards) along with their latitude and longitude.
We will Scrap Tokyo districts (Wards) Table from Wikipedia and get the coordinates of these 23 major districts using geocoder class of Geopy client.
Restaurants in each neighborhood of Tokyo:
Datasource : https://en.wikipedia.org/wiki/Special_wards_of_Tokyo#List_of_special_wards
Description : By using this API we will get all the venues in each neighborhood. We can filter these venues to get only restaurants.

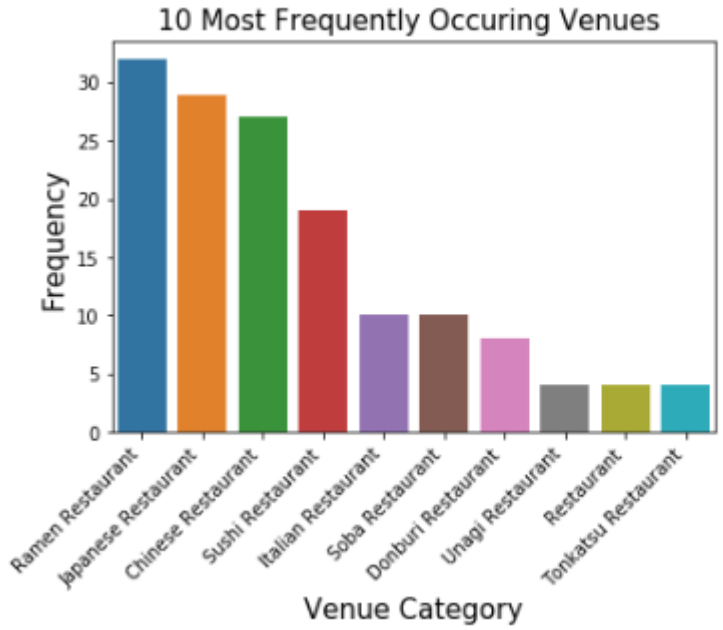
Data Preparation

We Scrap Data from Wikipedia Special Wards of Tokyo page to create my initial Dataframe using Pandas and make transformation on the dataset

Next, we get the coordinates of the 23 Tokyo major districts that we have in the dataset using geocoder class of Geopy client to add them to our initial dataset

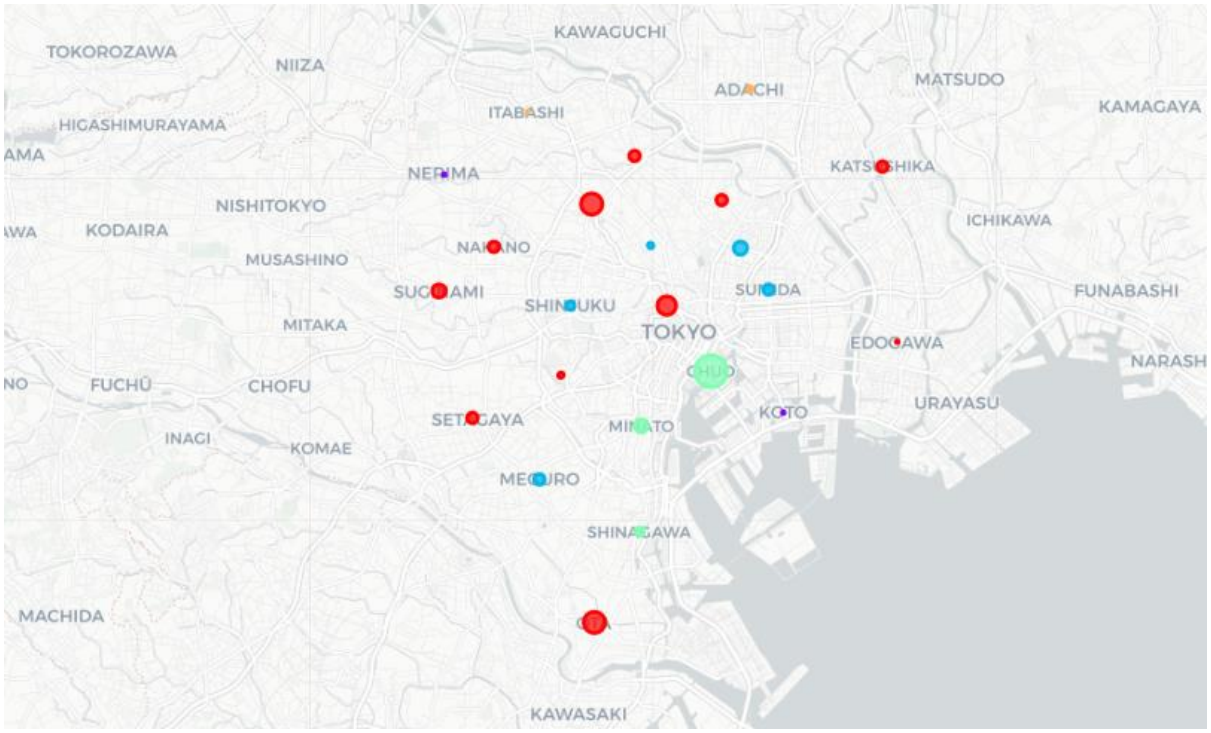
EDA

In our project we will concentrate in Restaurant Category only and explore all the 23 districts , let's now see the top 10 types of restaurant that our districts has as follow



So as we can see the most commun type of restaurants in Tokyo is the 'Ramen Restaurant' , so Ramen Restaurants would be your first choice if you visit Tokyo 😉 .

Finally , we use *clustering* (KMeans) to help a traveler decide a location to go for a restaurant , and in order to do that we try to cluster our 23 districts based on the restaurant categories and our expectation would be based on the similarities of venue categories, these districts will be clustered.



So as we can see Ramen restaurants has the most common venues in the 23 districts and Chuo ward and Chiyoda ward has maximum number of restaurants.

In our analysis, we have ignored other factors like distance of the venues from closest stations, range of prices of restaurants, Michelin Restaurants and so on, since we don't have such data and it would be difficult to farm it for a small exploratory study like ours. Hence, our analysis only helps travelers to get an overview of Restaurants distribution by categories in the 23 major districts of Tokyo.