Selection of a location for a Russian restaraunt in Tokyo

Andrey Slyadnev

July21, 2020

1. Introduction

1.1 Background

Tokyo metropolis is one of the most populous metrapolian areas in the word with more than 13.9 million residents as of 2019. Occupying the area of 2,194km^2, it consisits of 23 special wards, 26 cities, 1 district and 4 subprefectures. The metropolis is well known for its highly developed and very competitive food service industry which has been given a considerable boost in anticipation of the 2020 Summer Olympics that Tokyo was going to host. In 2020, however, the outbreak of COVID-19 pandemic confused all plans. The International Olympic Committee and local organizers decided to postpone the Olympic Games till 2021 leaving the food service and other businesses without long hoped-for revenues and customers. On the other hand, unlike other countries the Japanese authorities have never legally enforced any hard lockdowns so even during the pandemic the restaraunts and other food service companies continued their business operations.

1.2 Problem

A changing landscape may bring new opportunities. The purpose of this project is to acquire data on Tokyo restaurant venues, allocate venues to corresponding clusters with a focus on ethnic type venues. In particular, this research should help understand how many Russian food restaraunts are available throughout the area and if there is any room for a new Russian restaurant.

2. Data acquisition and wrangling

2.1 Data sources

The initial data will be obtained via Wikipedia, the Foursquare API and other internet resources as necessary.

2.2 Data wrangling and analysis

At the first step, the geographic coordinates of Tokyo districts will be obtained from Wikipedia or other sources. These data will be used to plot them and build a dataset of Tokyo Districts and their venues. Following that all food restaraunts will be sorted out, plotted and clusterized to get a general picture of restaraunts distribution. At the final step, Russian food restaraunts will be identified and plotted.