

Current restaurants in New York

1. Introduction

Starting a new business, especially opening a new restaurant, requires a lot research on the current status of competing restaurants and the potential locations. New York is one of the largest cities in the world and to opening a new restaurant in the city of New York, we need to determine the appropriate location and the focus of the restaurant. Therefore, in this study, the distribution of the current restaurants in each borough is investigated.

The questions to be answered in this investigation include 1) the top 5 most popular restaurant types in each borough and 2) how the neighborhoods in each borough cluster based on the top 5 most popular types of restaurants.

2. Data

2.1 Data source

The data required in this project include the geological dataset that is available online ([link](#)). This dataset includes the name and coordinates (latitude and longitude) of the 5 boroughs and 306 neighborhoods of New York. The other dataset used in this project is obtained from Foursquare ([link](#)), containing the name, location, and category of the venues in New York.

2.2 Data preparation

First, the name and coordinates of the 5 boroughs and 306 neighborhoods are extracted from the dataset and stored in a Pandas data frame. Similarly, the name, location, and category of venues related to restaurants, such as American restaurant and Italian restaurants, of each neighborhood are exported and stored in another Pandas data frame.

Second, merge the two data frames together to a new data frame, and use the top 5 most common restaurant types to cluster the neighborhoods in each borough.