

Introduction and Business Problem

Background

The Battle of Neighborhoods in this project I am going to compare the neighborhoods of the two cities (Rotterdam and The Hague) in South Holland province of the Netherlands. By the end of the project I am going to determine how similar or dissimilar they are. Rotterdam is a major port city and it is now known for bold and modern architecture. Rotterdam city was almost completely reconstructed after WWII. On the North Sea coast of the western Netherlands, the Hague city is located. The Hague city is home to the International Criminal Court (ICC) and the U.N.'s International Court of Justice. For people that are new to the Netherlands particularly to the South Holland province, both cities offer many interesting things to do such as visiting the beautiful buildings architecture in Rotterdam city and the beach in the Hague. At the same time, it would be fun and exciting to visit and eat in the restaurants which are customer's recommended and preferred choice. The locals are very outgoing and take every opportunity to taste different cuisines and food which is offered by the diverse and multinational community in the Netherlands. Similarly, the international students are trying their best to taste various food and find familiar taste of food offered by their country cooks. Importantly, **the main target audience of this project are the travelers** who have limited time to visit amazing places and will be provided a list of restaurants, marked on the map, in clusters based on the customers' (users) venue likes from Foursquare.

Business Problem

In this project, I am going to compare Rotterdam and the Hague city that how similar or dissimilar they are. In addition, I will provide two separate lists of the most liked restaurants marked on the map based on the users who have like the venues (restaurants) on Foursquare. One might think that travelers could google and/or use different websites to find restaurants which have the option of ratings and reviews. Similar to the machine learning and data science skills that google and other applications use to offer such choices, here I am going to replicate and offer that

service with the skills that I have learned joining IBM Data Science course. **Thus, the business problem that will be tackled in this project is to assist the visitors/travelers (to the Netherlands) and anyone interested to find customer's preferred choice restaurants.** For two main reasons it is important to solve this business problem. First, it is time saving and takes away the unnecessary anxiety of missing the opportunity to eat in the best restaurants in a single visit to the Netherlands. Second, for the businesses (restaurants) this project provides in clusters, categorization of restaurants, based the customer's experience, which could be used to improve their services.

Data Requirements

From Wikipedia, list of cities, towns and villages in South Holland are available. I am going to use BeautifulSoup, one of the available libraries in Python, to extract (Name, Municipality, Coordinates) data. Subsequently, after cleaning and preparing the data frame for the two cities through Foursquare API the following data on venues (restaurants) in two cities Rotterdam and the Hague of the Netherlands are compiled and stored for further analysis:

- Venue Name
- Venue ID
- Venue Location (lat and lng)
- Venue Category
- Venue User Likes

Data Acquisition Approach

Latitude and longitude coordinates for South Holland province focusing on Rotterdam and the Hague cities.

Through Foursquare API to get a list of all venues in Rotterdam and the Hague cities consist of venue name, venue id, location, category, and likes.

Methodology

Following similar process that, Christopher Jameson, has carried out that venue users “likes are a proxy for quality.” The number of “likes” for each of the restaurants are counted and stored in column (satisfactory category) categorizing the restaurant in four categories (Poor, Fair, Good, and Excellent). This will be done using k-means clustering algorithm grouping restaurants into 4 clusters. Finally, I will provide two separate lists of the most liked restaurants marked on the map, based on the users who have like the venues (restaurants) on Foursquare, to be able to choose among the restaurants that customers (users) have liked the venues (restaurants). Depending of the type of membership on Foursquare the regular and premium calls limits calls such as venue tips and photos.

References

Christopher Jameson: <https://medium.com/@chriswjameson/ibm-data-science-professional-certificate-capstone-project-blog-df3044cffe24>

Foursquare: <https://developer.foursquare.com/docs/api/endpoints>

Google search Rotterdam:

<https://www.google.com/search?q=rotterdam&oq=ro&aqs=chrome.69i59j69i60l3j69i57j69i59.1193j0j4&sourceid=chrome&ie=UTF-8>

Google search the Hague:

https://www.google.com/search?ei=AMUDXbLDDLDjkgXz5Lkl&q=the+hague&oq=the+&gs_l=psy-ab.3.0.0i67l6j0l4.75086.77585..78796...0.0..0.179.1104.8j4.....0....1..gws-wiz.....0..0i71.chRLr4Ni8GA

Wikipedia: https://en.wikipedia.org/wiki/List_of_cities,_towns_and_villages_in_South_Holland