Data Science Capstone Project

Destination Cities - Vacation Travel Planning

June 7th, 2020

Tim Krompinger

Introduction/Business Problem:

When making decisions about vacation plans, individuals may not know specifically where they want to travel next or they may want to visit a new location but the location they desire is over their budget. In these cases, can we automatically provide suggestions for similar travel destinations based on venues and attributes by leveraging the Foursquare location data? By providing location recommendations the traveler who may not know where they want to visit next but knows the locations they enjoyed previously can narrow their destination options to those which are similar to locations they have already enjoyed. In the case of the individual who may have a location in mind that exceeds their budget, providing similar destinations can allow the individual to seek a similar experience in a different location that may fit in their budget. This application of the Foursquare data can be useful in search logic for travel or hotel sites.

Data:

In order to solve this problem and execute an example of this functionality a list of possible vacation locations has been predetermined by myself as a sample set. These locations include specific areas of possible destination cities that may be of interest to visit on a vacation. A total of 14 U.S. locations have been included in the sample list which is composed of the following:

"Savannah Historic District,Savannah,Georgia"

"Downtown Charleston,Charleston,South Carolina"

"Southwest Orlando,Orlando,Florida"

"Garden District,New Orleans,Louisiana"

"Fisherman's Wharf,San Francisco,California"

"Beacon Hill,Boston,Massachusetts"

"Bell Rock,Sedona,Arizona"

"Downtown,Key West,Florida"

"The Loop,Chicago,Illinois"

"Downtown,Houston,Texas"

"Downtown,Nashville,Tennessee"

"Union Station,Denver,Colorado"

"Downtown,Asheville,North Carolina"

"Downtown,St. Augustine,Florida"

Specific areas in each city are intentionally included to replicate a traveler who would be staying in a hotel or rental within a specific area of the city and therefore would tie their experience to the area they visited. The latitude and longitude coordinates will need to be determined for each of these locations in order to build proper data associations. A Nominatim geocoder tool will be used to establish these coordinates. Once destination cities and coordinates have been established, Foursquare location data from the Foursquare database will be used to find venue information for the indicated area in the destination cities above. This information includes such venues as:

"Concert Hall"

"Tea Room"

"Bar"

"Burger Joint"

"Historic Site"

ect.

The venue information will be analyzed across all locations and used to determine similarities between the possible destinations. For example, destinations with a large number of historic sites may be similar or locations with a high number of concert halls or music venues. Additional information regarding the complete analysis of this section will be expressed in the Methodology section of this report.