Data Description

Data used for this project came from two sources. Data on destination counties were retrieved from Table1. Population Density and Non-Driving Commuting in Counties That Disproportionately Attract New Jersey Millennials, from njfuture.org website. Foursquare.com was used to gather data on list of common venue in the counties as well as their coordinates.

The coordinate for each destination country will be retrieved using geocoder. The latitude and longitude will be used to get the most common venues from the foursquare.com. List of 10 most common venues will be created. KMeans clustering method will be used to cluster the counties based on the common venues. Subsequently, the clusters will be compared with the population density, commuting percentage and number of migrants retrieved from destination counties table.