Coursera Capstone Project

Coursera IBM Data Science Certification

By: Josué Irizarry

1.0 Introduction

1.1 Scenario and Background

Here is my history and problem to resolve, currently I am living in the south of Puerto Rico island in a main municipality called Ponce, within 20 minutes to my work and another 20 far from my children's school. In these areas we have some venues and attractions, such as the local airport, cinema, fitness center, malls, international cuisine, hospitals, pharmacies and there are not a lot of options at all. I have the interest to move to work and live in another place like San Juan metropolitan area or Caguas municipally here in Puerto Rico, where I can have job options and schools not far from home and security. Also more options for entertainment, cuisine, shopping and hospital.

1.2 Problem to resolved:

How to find a new home in Puerto Rico metropolitan area or another place in Florida with following conditions:
- House or apartment with minimum of 3 bedrooms
- Monthly rend not to exceed US $2000 / month
- Available jobs around the area
- Schools around the new living area
- Centralized to venues and amenities like cinemas, fitness center, recreative parks malls, spa, international cuisines, hospital and pharmacies

1.3 Interested Audience

The audience is wide open to individuals or groups who want to take advantage of the model, approach and data analysis used in this data science capstone project. The use of FourSquare data and mapping techniques in conjunction with data analysis techniques will help to resolve the project's key questions.

2.0 Data Section

2.1 Data Requirements

- Geodata for current residence in Puerto Rico with venuess established thanks to FourSquare.
- List of Puerto Rico neighborhoods with clustered venus established via Foursquare: https://welcome.topuertorico.org/reference/zipcodes.shtml
- List of apartments or houses for rent in San Juan and Caguas municipalities with information on neighborhood location, address, number of beds, monthly rent price and complemented with geo data.
- List of schools K-12 in San Juan and Caguas municipalities
- Place to work in Caguas and San Juan for reference

2.2 Data Sources, Data Processing and Tools used

- Puerto Rico data map is to be created with help of tools: Nominatim, Foursquare and Folium mapping
- Neighborhoods are being obtained from Wikipedia and other available websites with geodata via Nominatim for mapping for Folium.
- List of Schools stations was obtained via Wikipedia and Google map.
- List of apartments or houses will be obtained from Internet information from Zillow.com or ClasificadosOnline. The geolocation (lat, long) will be obtained using Nominatim.
- Folium maps will be using the consolidated all the data in One map, where it is possible to visualize the insight to help with decisions of a new place to move.