Project : Capstone Battle of Neighorhoods

Objective

The main objective of this project is to apply the different skills learned in the course IBM Analytics by Coursera and does not pretend to be a Business Case. The target audience could be people who are interested and giving their first steps in Business Analytics with Python.

We will use Python for this project together with some libraries like pandas, Seaborn, Matplotlib, Choropleth, Folium, Foursquare and others that will help us to do the analysis.

Introduction.

México city is one of the most populated cities on the world.

Our interest is to find out if there are similarities on the type of cousins between Mexico and a country with similar population.

Business Understanding

There are several steps that we will go thru this project, included but not limited to:

Finding the data

Search a country with a similar population to México. There are a lot of resources to get this information, but we have decided to explore them thru the World Bank, an organization that provide high-quality international statistics and maintains a number of macro, financial and sector databases. We will explore https://datos.bancomundial.org to find out which country has a similar population to Mexico.

Understanding the data

Understanding the structure of the country in terms of administrative divisions (states, municipalities, etc.). Some of this information can be obtained from Wikipedia, and in other cases can be obtained from government resources of each country. For both countries we will show the population information segregated by administrative divisions.

Analizing the data

Based on the above results, we will decide which administrative division we will explore in more detail for each country.

We will use Foursquare to explore venues (in our case cousins types) around the areas selected.

At the end we will use K-means to find the similarities (if any) among the two cities selected and proceed to the conclusion of the project.

Conclusion

Provide conclusions based on the findings made in the analysis.