
Applied Data Science Capstone Project

Exploring and Analyzing Neighborhoods in Charlotte, North Carolina



IBM Data Science Professional Certificate

The IBM Data Science Professional program consists of 9 online courses that provide exposure to data tools and skills including open source tools and libraries, Python, databases, SQL, data visualization, data analysis, statistical analysis, predictive modeling, and machine learning algorithms.

This Professional Certificate has a strong emphasis on applied learning.

Data

A dataset with the required information could not be located, therefore a small dataset was manually created using Google Maps to identify neighborhood names and GPS coordinates.

This data was then imported from CSV format into a Pandas dataframe. After importing the data, Geopy was used to create a base map for Charlotte, North Carolina. Folium was used to place the neighborhoods into the base map.

Foursquare location data was used to identify the most popular venues in each area. K-Means clustering was then applied to group the neighborhoods according to their similarity to each other.