Project Presentation for Coursera Capstone

THE BATTLE OF NEIGHBORHOODS

MANHATTAN GROCERY STORE LOCATION SUGGESTION

Agenda:

- ✓ Objective and Problem Statement
- ✓ Resolution
- ✓ Automation
- ✓ Input/ Output
- ✓ Benefits
- ✓ Limitations

Objective and Problem Statement

- Objective

  Using dataset containing NYC Boroughs and Neighborhood features coupled with data science and Machine Learning techniques, enable potential clients to make informed decision to find the right Manhattan's neighborhoods for opening a Grocery store.

- Business Problem

  New York City (NYC) is the most populous city and most densely populated major city in the United States. A global power city, New York City has been described uniquely as the financial capital of the world and exerts a significant impact upon commerce. Needless to say, the city and especially its borough Manhattan provides immense opportunities for entrepreneurs and businessmen.

  NYC's food culture includes an array of international cuisines influenced by the city's immigrant history. This is evident from the fact that the city has a humongous number of restaurants, bars, cafes, joints et al. in its neighborhoods which creates a huge demand for raw materials to be supplied to these food outlets with quality and in a timely manner. In other words, a large scale Grocery store which would cater mainly in storing and supplying items required for restaurants' inventory. One of the challenge for any business to flourish is to carefully select the neighborhoods where it wants to target its customers.

Resolution:

With NYC Boroughs and Neighborhoods data coupled with data science and Machine Learnig techniques, one can derive useful information about setting up a new Grocery store based in a Manhattan neighborhood based on the venues and venue categories existing in the area.

The automated solution using Machine Learning techniques parses the necessary data/features from the json dataset which includes all of New York City's Boroughs and their neighborhoods and the features related to them viz. coordinates, venues, venue categories, et al. After cleansing, data is further condensed by selecting it only for the Borough Manhattan whose neighorhoods are the area of choice in this project. The data is further transformed by utilizing Foursquare API calls to extract only relevant information about the venues. Using the one-hot encoding and Machine Learning (K-Means clustering) techniques categorical variable (venue categories) are converted and categorized into 5 Clusters. These Clusters are then sorted based on the total population of the venue categories telling us which Cluster of the neighborhoods demands for a new Grocery store considering the population of already existing grocery stores. With NYC Boroughs and Neighborhoods data coupled with data science and Machine Learnig techniques, one can derive useful information about setting up a new Grocery store based in a Manhattan neighborhood based on the venues and venue categories existing in the area.

The automated solution using Machine Learning techniques parses the necessary data/features from the json dataset which includes all of New York City's Boroughs and their neighborhoods and the features related to them viz. coordinates, venues, venue categories, et al. After cleansing, data is further condensed by selecting it only for the Borough Manhattan whose neighorhoods are the area of choice in this project. The data is further transformed by utilizing Foursquare API calls to extract only relevant information about the venues. Using the one-hot encoding and Machine Learning (K-Means clustering) techniques categorical variable (venue categories) are converted and categorized into 5 Clusters. These Clusters are then sorted based on the total population of the venue categories telling us which Cluster of the neighborhoods demands for a new Grocery store considering the population of already existing grocery stores.

Automation Script:

Scripting language	Python 3.5
IDE	Notebook in IBM Watson
Input	JSON file with Coordinates of each NYC Neighborhood Foursquare location data
Output	1. List of suggested neighborhoods in Manhattan to open a new Grocery Store. 2. Suggested list of neighborhoods in Manhattan plotted on map. 3. Population of venue categories in the selected cluster with list neighborhoods.

Benefits:

- ✓ Dynamic recommendations
- ✓ Flexibility in choosing the list of venue categories
- ✓ Flexibility in choosing the Borough with its neighborhoods
- ✓ Suggestions plotted on the map for clarity

Thank You!