IBM®

DATA SCIENCE
PROFESSIONAL CERTIFICATE

WEEK 4 SUBMISSION

INTRODUCTION / BUSINESS PROBLEM

Section to discuss the business problem and who would be interested in the project

DATA SCIENCE
PROFESSIONAL CERTIFICATE

Week 4 Submission

INTRODUCTION

Original Business Problem & Proposal

- When traversing Long Island and interacting with the local residents, you will find that almost all of them share pride in – their food. This is especially true for both their pizza and bagels.
- If a local Long Islander travels somewhere and is served a pizza or bagel, they cannot resist saying, “This is good. But, not nearly as good as the ones that we have on Long Island. You see, our water has a lower concentration of calcium and magnesium which makes them better than everywhere else.”
- Having lived on Long Island for the majority of my life, I wanted to prove this element to be true. So, I would like to analyze all of New York State using FourSquare’s API. I would like to take a look into the Ratings of Bagel Stores, sort the venue ratings by County, and discover which County has the best rated Bagel Stores.

New Business Problem & Proposal

- (a necessity to purchase a Premium FourSquare Developer Account prevented me to accomplishing the Original Proposal)
- Friends of mine who have never been to Long Island are coming to visit. After bragging about the local area having the best bagels in the world, I need to ensure I take them to the highest rated Bagel Shops.
- I would like to take a look into the Ratings of Bagel Stores in New York City using FourSquare’s API and rank them from highest to lowest. Then, I would like to plot out all Bagel Store locations in New York City on an interactive map so that we are able to see if one is in close range to us.
- Audience: myself and friends touring New York for the first time. The results are important to the audience for we all wish to find Bagel Venues in NYC, as well as which ones are the highest rated.

COMPONENT DESCRIPTION

This section expands on the details of the data used in the project.

1. INPUT DATA

- This module requires geographical location data, specifically of New York State.
- Attributes such as Zip Code, State, City, Latitude, Longitude, County, etc. serve as the main driving force as the key features to answering the Business Problem

#	Name	Website	Description
1.1	US-CITY-STATE.csv	Open DataSoft	Table containing attributes 'Zip', 'City', 'Timezone', 'Daylight savings time flag', 'Latitude', and 'Longitude'. (data manipulated/divided in CSV to practice merging/appending in Python) Zip = 5 Digit Postal Code State = String Name of State City = String Name of City Timezone = Integer of UTC Timezone Daylight savings time flag = Flag with 0 for False, 1 for True in terms of participation in Daylight Savings Time
1.2	US-COORDINATES.csv	Open DataSoft	Table containing attributes 'Zip', 'City', 'Timezone', 'Daylight savings time flag', 'Latitude', and 'Longitude'. (data manipulated/divided in CSV to practice merging/appending in Python) Zip = 5 Digit Postal Code Latitude = Degree Latitude Coordinates with 6 decimal places Longitude = Degree Longitude Coordinates with 6 decimal places
1.3	NY-COUNTY.csv	Zip Codes to Go	Table containing attributes 'Zip Code', 'City', 'County', and 'Zip Code Map (Link)' Zip Code = 5 Digit Postal Code City = String Name of City County = String Name of County Zip Code Map (Link) = Hyperlink to Google Maps Postal Code Location

2. OUTPUT DATA

#	Name	Website	Description
2.1	BAGEL-RATINGS.csv	FourSquare API	Table containing attributes ‘Venue’, ‘Rating’, ‘Ratings Count’, and ‘Tip’. (data manipulated/divided in CSV to practice merging/appending in Python) Venue = String Name of Individual Venue Rating = Average Rating for a Store out of 10 Ratings Count = The number of ratings contributed to a particular location Tips = The number of tips that were commented on a particular location

- Initial data will be combined to make parameters for input into the FourSquare API.
- Venue names, ratings, tips, and other content will be pulled from FourSquare.
- This data, all-in-all, can be manipulated, altered, sorted, filtered, etc. to create many different types of results.
- By pulling venue name, venue category, and rating from FourSquare, we will be able to solve the initial Business Problem.

THANK YOU!