

SMART Trip Planner

IBM Data Science Professional certification-
Coursera Capstone Project

Doneria, Kriti

7/14/2019

Contents

1. Introduction	2
2. Data Description.....	2
3. Methodology	2
Creating a Foursquare developer account	2
Define the city and get its latitude & longitude	3
Search and fetching results from the API	3
Cleaning the data	4
4. Results.....	4
5. Discussion and future scope	4
Extending it to other cities	4
Integrating more data sources (Google openAPI, Microsoft Here etc).....	4
Deploying it as a web app	4
6. Conclusion.....	4

Introduction

Happyevents is a company in Delhi, India that manages events for business tourists. The company has to put a good program, including a hotel, a hall for meetings, places of landscape to visit, stores for shopping, restaurants and cafes. So this project is about taking input from a user in form of an Indian (metro) city and makes a list of places of landscape in the selected metro city, including the nearest restaurants, cafes, and shopping hangouts.

Data Description

The data used in this project is provided by Foursquare location data. The data are grouped by landscape area, and each area included the information about this area and all information about restaurants, cafes, and stores which in this area.

Methodology

Creating a Foursquare developer account

Creating an account here is simple, and takes less than 5 minutes here.

<https://foursquare.com/developers/apps>

Tech setup

Installing Anaconda

Installing libraries

```
import requests # to handle requests

import pandas as pd # for data analysis

import numpy as np # to handle data in a vectorized manner

# !conda install -c conda-forge geopy --yes # to get latitude and longitude

from geopy.geocoders import Nominatim # module to convert an address into latitude and longitude values

# libraries for displaying images

from IPython.display import Image

from IPython.core.display import HTML

#transforming json file into a pandas dataframe library

from pandas.io.json import json_normalize

# !pip install folium

import folium # plotting library
```

Define the city and get its latitude & longitude

Currently, users can pick from the four metro cities in India (Delhi, Mumbai, Chennai and Kolkata).

Search and fetching results from the API

The steps are as follows:

- Fetching the latitude and longitude of the city using geocode
- Sending a request to the foursquare API for the selected city
- Define the category of venue (Hotel, Park etc)
- Define the radius (by default in Miles)
- Firing the query combining all the above inputs
- Fetching the query results as json and converting it into dataframe
- Getting the relevant attribute from response result(venue)
- Data cleaning

Cleaning the data

- Retain only locational relevant columns
- Extracting the category of venue
- Retaining only the relevant category
- Delete unnecessary columns
- Printing the result

Results

The five dataframes having final results are:

- Df_hotels
- Df_park
- Df_cafeteria
- Df_Restaurants
- Df_shopping

It contains name and geographical data of the venues.

Discussion and future scope

Future extensions of this work can include:

- Extending it to other cities.
- Integrating more data sources (Google openAPI, Microsoft Here etc)
- Deploying it as a web app.

Limitations:

- Dataframe columns may vary based on the availability of data for certain cities.
- Data sanitization is a big challenge. Some hotels list themselves as office for instance.
- The API uses Euclidian distance which might not always be optimized for drivable distance.

Conclusion

New projects are all about looking at what other people have done and how beautifully can you combine this to make it do stuff differently.

This is a wonderful example I got to work on as a part of Final Capstone project for IBM Data science professional certification.

Combining different data sources should be done while taking care that the granularity is the same and the structure is fixed for each set of output produced.

This project can be a useful alternative to paid services and can save people lots of time.

Github link: