**Introduction**

**Background**

Dallas city is one of the most popoulous city in U.S. and is home to many immigrant population in Texas after San Antonio and Houston. Furthermore, it is the fourth-largest metropolitan area in the U.S. at 7.5 million people as of 2018 with an estimated population of 7,846,293 residents. Being an metropolitan city, Dallas is also home to many restaurants which serves wide variety of cuisines. Owing to significant number of Indian expatriate population, Dallas City and its nearby Suburbs have handful of Indian restaurant.

So, as a part of this project Indian restaurants in Dallas City will be listed, visualized and rated

**Problem Description**

By utilizing the Dallas City restaurants inspection data, Indian Restaurants in Dallas City and their risk category will be Analyzed.

Secondly, a classsifier model will be built to predict the risk categories of resturants.

Furthermore, using the foursqure API we will get the ratings of Indian Restaurants in Dallas City.

## Data

For this project we will use the following data :

Dallas City restaurants inspection data from 2016-2019

**Data\_source**: <https://www.dallasopendata.com/api/views/dri5-wcct/rows.csv?accessType=DOWNLOAD>

**Description :** This data set contains 37876 rows and 114 coulmns with Restaurant Name, Street Name, violation descriptions along with their latitude and longitude.

This data will be downloaded and used.

Ratings of Indian resturants for selected locality in Dallas City

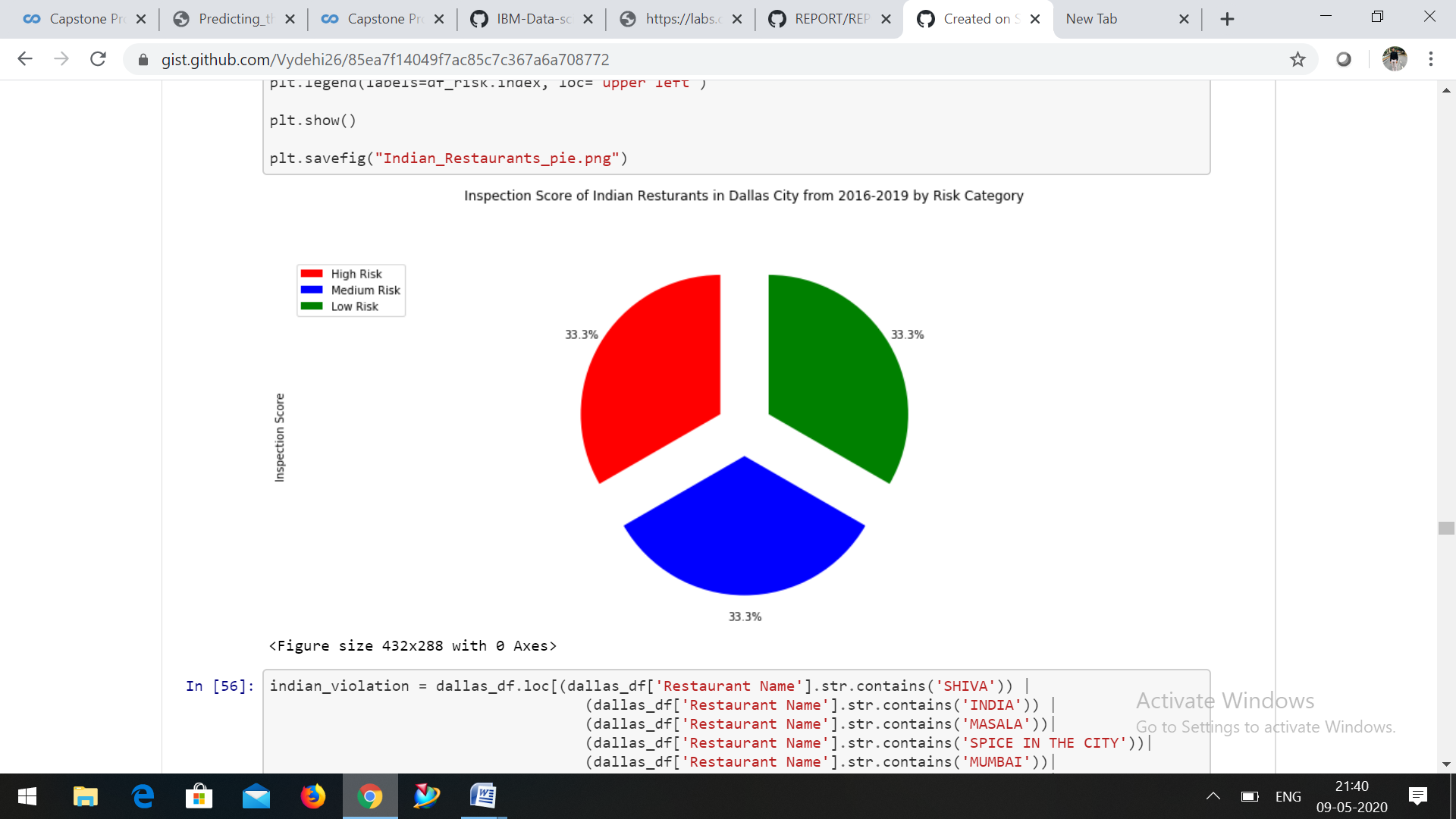
**Data source :** Fousquare API

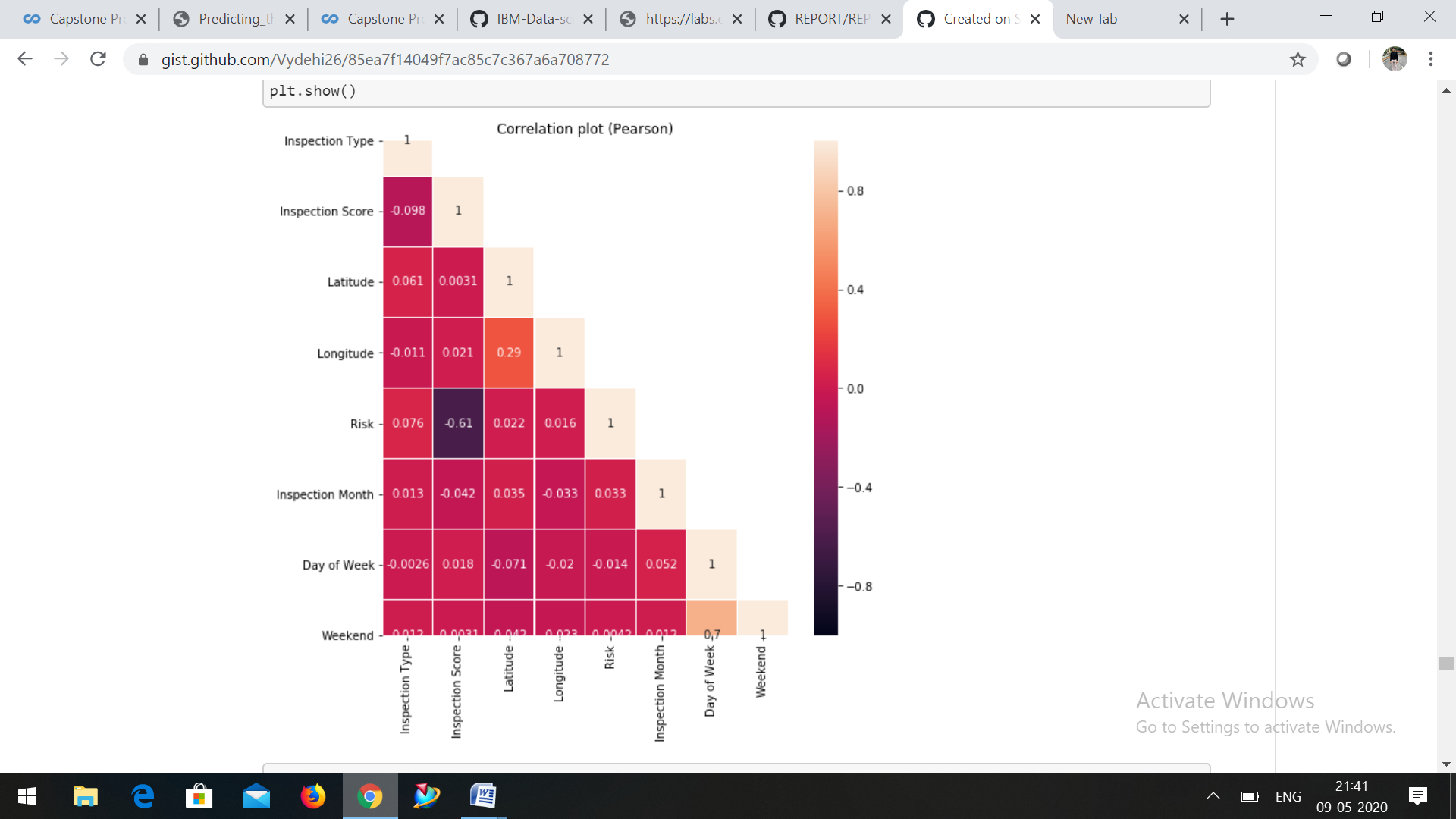
**Description** : By using this api we will get all the ratings for Indian restaurants in selected neighbourhood

## Methodology

* Collect the Dallas City Restaurants Inspection data from <https://www.dallasopendata.com/api/views/dri5-wcct/rows.csv?accessType=DOWNLOAD> Analyze the restaurants based on their risk category
* Built machine learning classifier model for predicting the risk category the resturants are placed
* Filter out all venues from the inspection data that are Indian Resturants.
* Visualize the Indian Restaurants using folium library(python)
* Find rating , tips and like count for each Indian Resturants using FourSquare API.

RESULTS:





**CONCLUSION:**

This project successfully completes my IBM Data Science Professional Certification Trianing. I am quite new to the data science and I had a steep learnig curve during the course. I have really enjoyed doing all the lab excersies and the courses were really informative. The following are the conclusions that I derive from this project:

Dallas City have only very few Indian restaurants. Hence, it has a potential market for opening a new Indian restaurant Roughly 80% of Indian restaurants that are cuurrently present in Dallas City are placed in low risk category based on the inspection data from 2016-2019 A decision tree classifier model is built for classifying the restaurants into various risk categories and the model performs well for the given data set. This will help the restaurants in predicting their risk category for a given year. The Indian restaurants in the Dallas City were visulaized using the folium map rendering library Using FourSquare API, the venue details for the Indian restaurants were analyzed and found that among all the restaurants in Dallas City India Palace is the best place to dine.