2020

Arora, Ankit

Capgemini

10/24/2020

Battle of Neighborhood



**Contents**

[**NEW FAST FOOD STORE PREFERENCES** 2](#_Toc54519791)

[**Introduction** 2](#_Toc54519792)

[**Data Collection** 2](#_Toc54519793)

[**Methodology** 3](#_Toc54519794)

[**References** 4](#_Toc54519795)

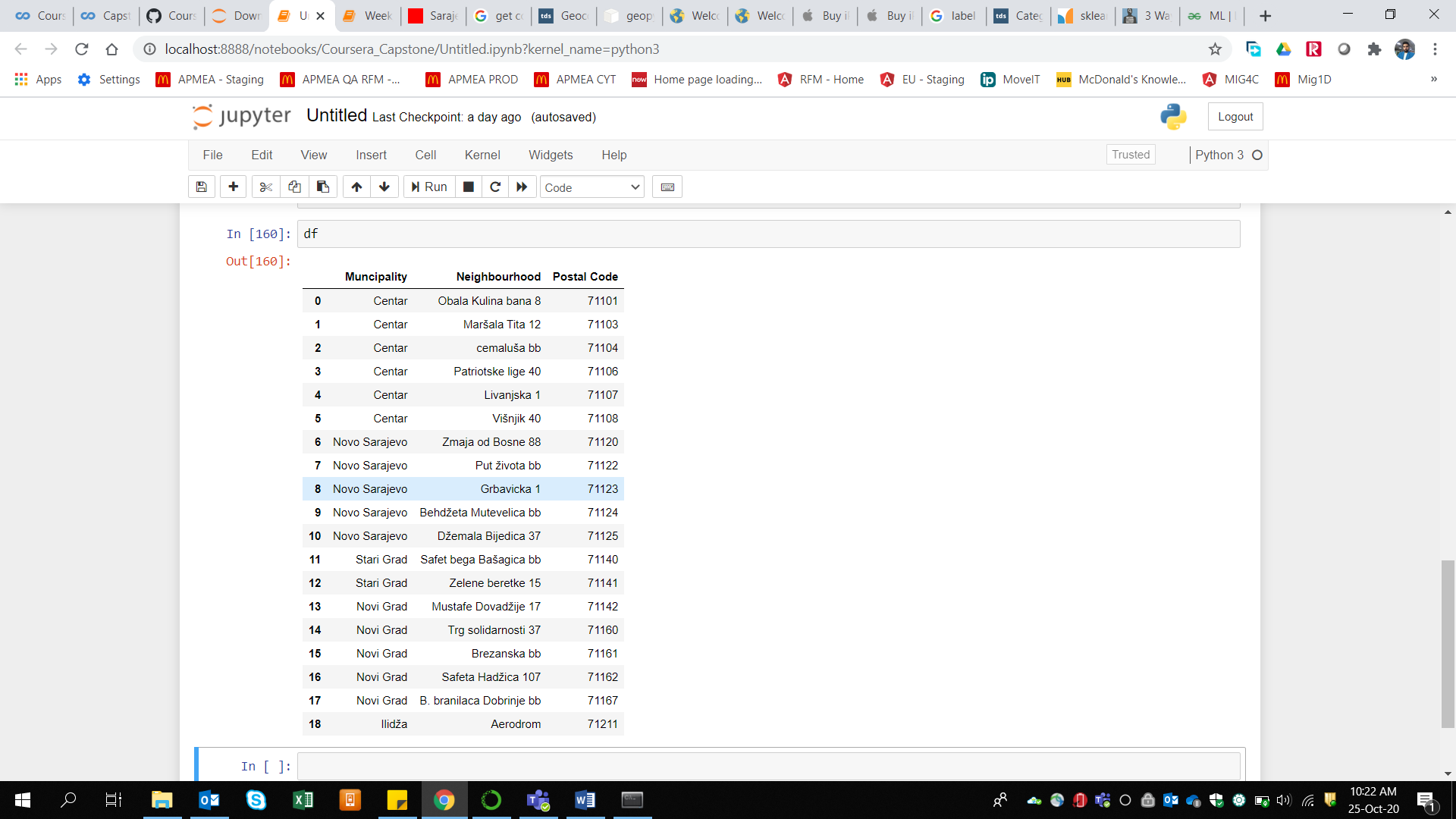
# **NEW FAST FOOD STORE PREFERENCES**

## **Introduction**

Whenever one have to start a business we need to first do market research in order to know the current trends of market as well as to find the premium location to set up an office where we can grab the attention of customers the most. Thus, we have to keep check on demographical, geospatial and environmental factors. Other major factor is to find potential competitors and the surroundings. In this proposal we are determined to evaluate above said features in order to find the best location in a city for opening of fast-food store. The proposal is intended to be for business management and thus the intention from the plan is to establish a new store and to get the business out of it. Initially, this setup will be helpful to determine the best location for a fast food store, but we can also build similar kind of models for other businesses and commodities.

## **Data Collection**

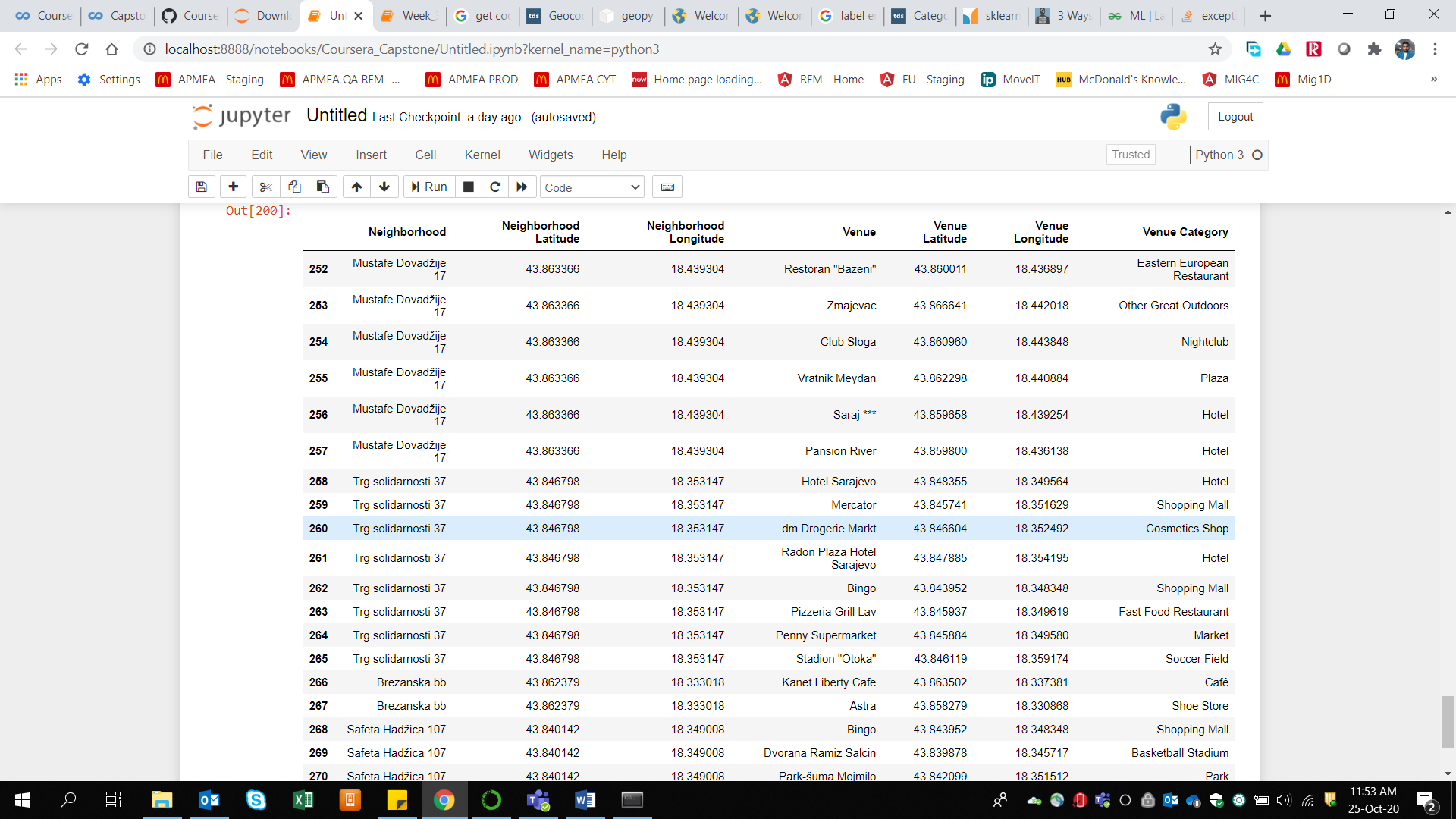
In order to accomplish the analytics task, we need to gather and process huge amount of data related to city locality, census, demographics, meteorological data and current businesses in the area. We have targeted Sarajevo for our analysis which is capital city of Bosnia in European continent and have collected the data for Municipalities and their neighborhoods from PostalCodeCountry.com. The data consist of Postal Codes, Municipality, and Neighborhood. Similarly, we have used geopy.geocoder module of Python to get longitude and latitude data of the localities.



*Fig. 1: Count of accidents with total no. of vehicles*

## **Methodology**

Methodology is the process that shows the approach followed to define and solve a business problem. In the introduction section, we have defined the business problem, then in data collection we have provided the reference of data we have taken for Sarajivo, Bosnia. We will be creating Jupyter notebook for interactive data analysis using Python programming language. We will be using Classification algorithms for explanatory data analysis and to find relation between independent and dependent variable. Confusion matrix will help determine the accuracy of our model. But before that we need to focus on Data Pre-processing and Feature Engineering which includes finding and removing outliers, filling up with NULL or missing values, Feature generation and data merging from various sources and datasets.



## **References**

Postal Codes: http://www.postalcodecountry.com/zipcode/Sarajevo-1-Bosnia%20Herzegovina

Venues: (Foursquare API) <https://api.foursquare.com/v2/venues/explore>

Geocoder reference: https://pypi.org/project/geopy/