

# Week\_1

February 11, 2019

## 1 1- INTRODUCTION: Gourmet Bussines

A chef has contacted me, since he needs to do a study of the businesses that have been successful in the culinary field he develops.

With respect to his experience, he has been based on Gourmet businesses in certain areas of Mexico City where they have been well received by the population.

His intention is to open a gourmet-type business in a potential place for his success and where the concept has not yet arrived.

The proposal that he makes to me is the following:

### 1.1 Customer's proposal

Dear Minerva: I am in contact with you to formally make my intention to require your artificial intelligence / machine learning service, for which I will detail below the requirements of the project you will develop.

Approach

As part of a professional growth, I intend to open a gourmet store, which will have as main selling axes:

Butchery: National and imported cuts of beef (prime quality) will be sold, wagyu meat (imported from Japan), aged meat (inside the store by us in aging chambers), pork cuts, piglets, exotic meats such as wild boar, crocodile. Fish and seafood: Daily selection attached to sustainable fishing, fish and shellfish national and imported. Spices, fruits, vegetables and other items related to gourmet preparations. Workshops, courses.

Location

As part of the support I require from you, I need to determine the ideal location of the store, which could be in the Cuahutemoc or Benito Juarez delegation, for which I need to consider the following points in order to make a decision:

Purchase incidence: Take into account the purchasing trends of customers, according to nearby businesses. Opening hours and closing of the store. Market segmentation.

Taking these points into account, I ask for your expertise to be able to create a model and suitable with the needs that I have.

Without further ado, I remain attentive to your response.

## 2 2- DATA

We will rely on Foursquare to have the necessary information of the places, as well as their popularity and the business line that is of interest.

In order to locate each of the colonies in the determined areas of Mexico City, we will access the INEGI portal, which is the National Institute of Statistics and Geography of Mexico, through a web query, we can download a csv file with the geographic coordinates of all the neighborhoods of the City and to be able to work on these data.

In the question of the procedure, we can highlight the places of Gourmet turnaround related to the interest of my client and associate the types of businesses and sites that exist around it, in order to associate a set of behavior related to the gourmet sector and thus be able to extrapolate this set to a new town in the city

## 2.1 csv File filter and Downloaded from <https://www.inegi.org.mx/app/geo2/cng/>

Import dependencies that we will need

```
In [3]: import pandas as pd
import numpy as np

In [8]: pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json # library to handle JSON files

#!conda install -c conda-forge geopy --yes # uncomment this line if you haven't complete
from geopy.geocoders import Nominatim # convert an address into latitude and longitude v

import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe

# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors

# import k-means from clustering stage
from sklearn.cluster import KMeans

#!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if you haven't c
import folium # map rendering library

print('Libraries imported.')
```

Libraries imported.

Preview of the Data obtained

```
In [9]: geo=pd.read_csv('borough_data.csv')
geo.columns = [c.replace(' ', '_') for c in geo.columns]
geo.rename(index=str, columns={'Nombre_de_municipio': 'Borough', 'Nombre_geografico': 'Nei
geo.head()
```

```

Out[9]:
      Neighborhood  Nombre_conocido  Clase Termino_generico \
0      10 de Abril                NaN  Localidad          Colonia
1      10 de Mayo                NaN  Localidad          Colonia
2     12 de Diciembre                NaN  Localidad          Colonia
3      15 de Agosto                NaN  Localidad          Colonia
4     16 de Septiembre                NaN  Localidad          Colonia

      Clave_de_entidad  Nombre_de_entidad  Clave_de_municipio \
0                9  Ciudad de Mexico                16
1                9  Ciudad de Mexico                17
2                9  Ciudad de Mexico                 7
3                9  Ciudad de Mexico                 5
4                9  Ciudad de Mexico                16

      Borough  Clave_de_localidad  Nombre_de_localidad  Latitud \
0    Miguel Hidalgo                1    Miguel Hidalgo 19ř26'55'' N
1  Venustiano Carranza                1  Venustiano Carranza 19ř26'01'' N
2      Iztapalapa                1      Iztapalapa 19ř19'45'' N
3  Gustavo A. Madero                1  Gustavo A. Madero 19ř29'05'' N
4    Miguel Hidalgo                1    Miguel Hidalgo 19ř24'07'' N

      Longitud  Latitude  Longitude
0  99ř12'30'' W  19.448611  99.208333
1  99ř07'05'' W  19.433611  99.118056
2  99ř06'02'' W  19.329167  99.100556
3  99ř06'23'' W  19.484722  99.106389
4  99ř12'25'' W  19.401944  99.206944

```

### 2.1.1 Filter the DataFrame with the data of Borough,Neighborhood,Latitude and Longitude

```

In [10]: df = pd.DataFrame()
          df['Borough']=geo.Borough
          df['Neighborhood']=geo.Neighborhood
          df['Latitude']=geo.Latitude
          ###multiply longitud by -1 in order to use in Foursquare
          df['Longitude']=(geo.Longitude)*-1
          df.head()

```

```

Out[10]:
      Borough  Neighborhood  Latitude  Longitude
0    Miguel Hidalgo      10 de Abril  19.448611 -99.208333
1  Venustiano Carranza      10 de Mayo  19.433611 -99.118056
2      Iztapalapa     12 de Diciembre  19.329167 -99.100556
3  Gustavo A. Madero     15 de Agosto  19.484722 -99.106389
4    Miguel Hidalgo     16 de Septiembre  19.401944 -99.206944

```

Use geopy library to get the latitude and longitude values of Mexico City.

```

In [11]: address = 'Mexico City, MX'

```

```

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Mexico City are {}, {}'.format(latitude, longitude))

```

The geograpical coordinate of Mexico City are 19.4326009, -99.1333416.

### Create a map of Mexico City with neighborhoods superimposed on top.

```

In [13]: # create map of Mexico using latitude and longitude values
map_mx = folium.Map(location=[latitude, longitude], zoom_start=11.48)

# add markers to map
for lat, lng, borough, Neighborhood in zip(df['Latitude'], df['Longitude'], df['Borough'], df['Neighborhood']):
    label = '{} {}'.format(Neighborhood, borough)
    popup = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=popup,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_mx)

```

map\_mx

Out[13]: <folium.folium.Map at 0x7f1900953160>

In [ ]: