Capstone Notebook

August 5, 2021

1. Introduction/Business Problem

1.1 Entrepreneurial Ecosystem:

An entrepreneurial ecosystem consists of stakeholders such as non profits, government, academia, private practitioners, financial institutions, among other support organizations. These stakeholders provide support in different areas such as education, access to capital, prototyping and even business plan. In order to maximize the resources of these support organizations they should know the business stage, size, and industry of the client they are helping. Specially, they should consider the location of the business since that data provides an additional layer of context.

1.2 Location:

In Puerto Rico, a territory of 3.1 million, there are more than 250 entrepreneurial support organizations. Sometimes, however, these organizations do not have the data in an accessible and digestible form that allows them to make data-driven decisions to improve their services.

1.3 Use Cases:

In this project we will be exploring geographic data with the idea of providing input to organizations that could help them to:

provide a better service in a personalized way

be able to design projects and develop grants based on the data

hire specialized personnel to serve the geographic area in which they are located

and develop public policy projects that help businesses in their area.

2. Data

2.1 Data Sources:

For this project we will be using Foursquare's Places API. The membership tier is the Sandbox. It includes:

950 Regular Calls/Day

50 Premium Calls/Day

1 Photo per Venue

1 Tip per Venue.

With this API it is possible to search, explore, get trends and recommendations regarding venues. In the same way it is possible to get data about users such as details, check-ins, venue history, among other information.

The zip codes were accessed through https://www.zipdatamaps.com/list-of-zip-codes-in-puerto-rico.php.

2.2 Feature Selection:

In this project the data will be limited to venues in Puerto Rico. The data that will be explored is categories. With this information entrepreneurial support organizations will be able to identify business clusters in different locations across Puerto Rico and areas of opportunity for new business owners or those that are looking to expand established businesses.

2.3 Data merged:

The data from ZipDataMaps will be merged with the Foursquare's Place API in order to access information of different venues in Puerto Rico.

3. Methodology

[1]:
```
!pip install beautifulsoup4
```

```
Requirement already satisfied: beautifulsoup4 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (4.9.3)
Requirement already satisfied: soupsieve>1.2; python_version >= "3.0" in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
beautifulsoup4) (2.2.1)
```

[2]:
```
import pandas as pd

contents = pd.read_csv("PRZipCodes.csv", dtype={'tablescraper-selected-row':␣
↪object})

zips_df = pd.DataFrame(contents)


zips_df = zips_df.drop(columns =['tablescraper-selected-row␣
↪href','tablescraper-selected-row 2','tablescraper-selected-row href 2',␣
↪'tablescraper-selected-row href 3'])

zips_df.rename(columns = {'tablescraper-selected-row':
↪'ZIP','tablescraper-selected-row 3':'Borough','tablescraper-selected-row 4':
↪'County'}, inplace = True)
zips_df['ZIP'] = zips_df['ZIP'].astype(str)

zips_df.head()
```

[2]:

	ZIP	Borough	County
0	00601	Adjuntas	Adjuntas
1	00602	Aguada	Aguada
2	00603	Aguadilla	Aguadilla
3	00604	Aguadilla	Aguadilla
4	00605	Aguadilla	Aguadilla

[3]:
```
LatLong_df = pd.read_csv("codes_latlong.csv",␣
 ↪dtype={'tablescraper-selected-row': object})

LatLong_df['ZIP'] = LatLong_df['ZIP'].apply(lambda x: '{0:0>5}'.format(x))
LatLong_df['ZIP'] = LatLong_df['ZIP'].astype(str)

LatLong_df.head()
```

[3]:

	ZIP	LAT	LNG
0	00601	18.180555	-66.749961
1	00602	18.361945	-67.175597
2	00603	18.455183	-67.119887
3	00606	18.158345	-66.932911
4	00610	18.295366	-67.125135

[4]:
```
zips_df.drop_duplicates()

LatLong_df.drop_duplicates()
```

[4]:

	ZIP	LAT	LNG
0	00601	18.180555	-66.749961
1	00602	18.361945	-67.175597
2	00603	18.455183	-67.119887
3	00606	18.158345	-66.932911
4	00610	18.295366	-67.125135
…	…	…	…
33139	99923	56.002315	-130.041026
33140	99925	55.550204	-132.945933
33141	99926	55.138352	-131.470424
33142	99927	56.239062	-133.457924
33143	99929	56.370751	-131.693301

[33144 rows x 3 columns]

[5]:
```
join_df = pd.merge(zips_df, LatLong_df, on='ZIP', how='inner')

join_df.head()
```

[5]:

	ZIP	Borough	County	LAT	LNG
0	00601	Adjuntas	Adjuntas	18.180555	-66.749961
1	00602	Aguada	Aguada	18.361945	-67.175597
2	00603	Aguadilla	Aguadilla	18.455183	-67.119887
3	00606	Maricao	Maricao	18.158345	-66.932911
4	00610	Anasco	Anasco	18.295366	-67.125135

[6]:
```
!pip install folium
```

```
Requirement already satisfied: folium in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (0.5.0)
Requirement already satisfied: requests in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from folium)
(2.25.1)
Requirement already satisfied: six in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from folium)
(1.15.0)
Requirement already satisfied: branca in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from folium)
(0.4.2)
Requirement already satisfied: jinja2 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from folium)
(3.0.1)
Requirement already satisfied: idna<3,>=2.5 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->folium) (2.10)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->folium) (1.26.6)
Requirement already satisfied: certifi>=2017.4.17 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->folium) (2021.5.30)
Requirement already satisfied: chardet<5,>=3.0.2 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->folium) (4.0.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
jinja2->folium) (2.0.1)
```

[7]:
```
import folium
# create map of Toronto using latitude and longitude values
map_pr = folium.Map(location=[18.303910000000002, -66.32617900000001],␣
 ↪zoom_start=9)

# add markers to map
for lat, lng, borough, county in zip(join_df['LAT'], join_df['LNG'],␣
 ↪join_df['Borough'], join_df['County']):
    label = '{}, {}'.format(borough, county)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
```

```
        fill_opacity=0.7,
        parse_html=False).add_to(map_pr)

map_pr
```

[7]:
```
<folium.folium.Map at 0x7fc293994860>
```

[8]:
```
CLIENT_ID = 'ONI3JSZV10EKQVW4DX4H3B3H454FPB1W4ZYVDVN5HYJKJXPI' # your␣
 ↪Foursquare ID
CLIENT_SECRET = 'PM3SHQOEVNHZXIO2JOFW1MCPGFRIVNIWRFTGNRGZROEFCZ1Q' # your␣
 ↪Foursquare Secret
VERSION = '20180605' # Foursquare API version
LIMIT = 100 # A default Foursquare API limit value

radius = 5000 # define radius

url = 'https://api.foursquare.com/v2/venues/explore?
 ↪&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    18.303910000000002,
    -66.32617900000001,
    radius,
    LIMIT)
url
```

[8]:
```
'https://api.foursquare.com/v2/venues/explore?&client_id=ONI3JSZV10EKQVW4DX4H3B3
H454FPB1W4ZYVDVN5HYJKJXPI&client_secret=PM3SHQOEVNHZXIO2JOFW1MCPGFRIVNIWRFTGNRGZ
ROEFCZ1Q&v=20180605&ll=18.303910000000002,-66.32617900000001&radius=5000&limit=1
00'
```

[9]:
```
import requests


results = requests.get(url).json()
results
```

[9]:
```
{'meta': {'code': 200, 'requestId': '610b2aa154f76f571f962930'},
 'response': {'headerLocation': 'Padilla',
  'headerFullLocation': 'Padilla',
  'headerLocationGranularity': 'city',
  'totalResults': 10,
  'suggestedBounds': {'ne': {'lat': 18.348910045000046,
    'lng': -66.2788692956976},
   'sw': {'lat': 18.258909954999957, 'lng': -66.37348870430242}},
  'groups': [{'type': 'Recommended Places',
```

```
    'name': 'recommended',
    'items': [{'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4c10e6eda6c19521746636dd',
       'name': 'El Yagrumo',
       'location': {'address': 'PR-159',
        'lat': 18.326534114706522,
        'lng': -66.34989312233385,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.326534114706522,
          'lng': -66.34989312233385}],
        'distance': 3552,
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['PR-159', 'Corozal', 'Puerto Rico']},
       'categories': [{'id': '4bf58dd8d48988d144941735',
         'name': 'Caribbean Restaurant',
         'pluralName': 'Caribbean Restaurants',
         'shortName': 'Caribbean',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/caribbean_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4c10e6eda6c19521746636dd-0'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4c0a92613c70b713ab05285b',
       'name': 'El Gran Cafe Restaurant',
       'location': {'address': 'Cervantes # 1',
        'lat': 18.340956533288097,
        'lng': -66.31698660688099,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.340956533288097,
          'lng': -66.31698660688099}],
        'distance': 4236,
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['Cervantes # 1', 'Corozal', 'Puerto Rico']},
```

```
       'categories': [{'id': '4bf58dd8d48988d143941735',
         'name': 'Breakfast Spot',
         'pluralName': 'Breakfast Spots',
         'shortName': 'Breakfast',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/breakfast_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4c0a92613c70b713ab05285b-1'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4c4791cf89a6c9b601cb9a88',
       'name': 'El Limo Viejo',
       'location': {'address': 'Pr- 159',
        'lat': 18.338022387928774,
        'lng': -66.3222914784637,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.338022387928774,
          'lng': -66.3222914784637}],
        'distance': 3819,
        'postalCode': '00783',
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['Pr- 159', 'Corozal 00783', 'Puerto Rico']},
       'categories': [{'id': '4bf58dd8d48988d116941735',
         'name': 'Bar',
         'pluralName': 'Bars',
         'shortName': 'Bar',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/nightlife/pub_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4c4791cf89a6c9b601cb9a88-2'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4cd5d944886cb60c709d8689',
       'name': 'Balalaika Restaurant',
       'location': {'address': 'Carr. 159 Km. 16.0',
        'crossStreet': 'Bo. Mavillas',
```

```
        'lat': 18.341657090311102,
        'lng': -66.30547214326874,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.341657090311102,
          'lng': -66.30547214326874}],
        'distance': 4737,
        'postalCode': '00783',
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['Carr. 159 Km. 16.0 (Bo. Mavillas)',
         'Corozal 00783',
         'Puerto Rico']},
       'categories': [{'id': '4bf58dd8d48988d16c941735',
         'name': 'Burger Joint',
         'pluralName': 'Burger Joints',
         'shortName': 'Burgers',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/burger_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4cd5d944886cb60c709d8689-3'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4c5d8a526ebe2d7fcc0dd32e',
       'name': 'El Rancho en Corozal',
       'location': {'lat': 18.3192965837838,
        'lng': -66.29264352517542,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.3192965837838,
          'lng': -66.29264352517542}],
        'distance': 3936,
        'cc': 'PR',
        'country': 'Puerto Rico',
        'formattedAddress': ['Puerto Rico']},
       'categories': [{'id': '4bf58dd8d48988d15e941735',
         'name': 'Pool',
         'pluralName': 'Pools',
         'shortName': 'Pool',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/parks_outdoors/pool_',
          'suffix': '.png'},
         'primary': True}],
```

```
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4c5d8a526ebe2d7fcc0dd32e-4'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4c44a91cf97fbe9a107dbb30',
       'name': 'Burger King',
       'location': {'address': 'Corozal',
        'lat': 18.342239278811025,
        'lng': -66.31022036075592,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.342239278811025,
          'lng': -66.31022036075592}],
        'distance': 4587,
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['Corozal', 'Corozal', 'Puerto Rico']},
       'categories': [{'id': '4bf58dd8d48988d16e941735',
         'name': 'Fast Food Restaurant',
         'pluralName': 'Fast Food Restaurants',
         'shortName': 'Fast Food',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/fastfood_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4c44a91cf97fbe9a107dbb30-5'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4d6937e99bd6a143b92c5cdf',
       'name': 'Walgreens',
       'location': {'address': 'Rd 159 Km 153 Barrio Pueblo',
        'lat': 18.341208,
        'lng': -66.310024,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.341208,
          'lng': -66.310024}],
        'distance': 4489,
        'postalCode': '00783',
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
```

```
        'country': 'Puerto Rico',
        'formattedAddress': ['Rd 159 Km 153 Barrio Pueblo',
         'COROZAL 00783',
         'Puerto Rico']},
       'categories': [{'id': '4bf58dd8d48988d10f951735',
         'name': 'Pharmacy',
         'pluralName': 'Pharmacies',
         'shortName': 'Pharmacy',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/shops/pharmacy_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4d6937e99bd6a143b92c5cdf-6'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4e551af262e1d184432675e7',
       'name': 'Supermercado Amigo Corozal',
       'location': {'lat': 18.340947898506673,
        'lng': -66.3114862132842,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.340947898506673,
          'lng': -66.3114862132842}],
        'distance': 4405,
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['Corozal', 'Puerto Rico']},
       'categories': [{'id': '52f2ab2ebcbc57f1066b8b46',
         'name': 'Supermarket',
         'pluralName': 'Supermarkets',
         'shortName': 'Supermarket',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/shops/food_grocery_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4e551af262e1d184432675e7-7'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4c96a583f244b1f7ed55361d',
       'name': 'La Plaza de Corozal',
```

```
       'location': {'lat': 18.340941895950376,
        'lng': -66.31720193438595,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.340941895950376,
          'lng': -66.31720193438595}],
        'distance': 4230,
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['Corozal', 'Puerto Rico']},
       'categories': [{'id': '4bf58dd8d48988d164941735',
         'name': 'Plaza',
         'pluralName': 'Plazas',
         'shortName': 'Plaza',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/parks_outdoors/plaza_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4c96a583f244b1f7ed55361d-8'},
     {'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4eda4abb722e1da302182dd2',
       'name': "Church's Chicken",
       'location': {'lat': 18.342600618564397,
        'lng': -66.31017600071418,
        'labeledLatLngs': [{'label': 'display',
          'lat': 18.342600618564397,
          'lng': -66.31017600071418}],
        'distance': 4627,
        'cc': 'PR',
        'city': 'Corozal',
        'state': 'Corozal Municipio',
        'country': 'Puerto Rico',
        'formattedAddress': ['Corozal', 'Puerto Rico']},
       'categories': [{'id': '4d4ae6fc7a7b7dea34424761',
         'name': 'Fried Chicken Joint',
         'pluralName': 'Fried Chicken Joints',
         'shortName': 'Fried Chicken',
         'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/friedchicken_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
```

```
'referralId': 'e-0-4eda4abb722e1da302182dd2-9'}]}]}}
```

[10]:
```
def get_category_type(row):
    try:
        categories_list = row['categories']
    except:
        categories_list = row['venue.categories']

    if len(categories_list) == 0:
        return None
    else:
        return categories_list[0]['name']
```

[11]:
```
from pandas.io.json import json_normalize

venues = results['response']['groups'][0]['items']

nearby_venues = json_normalize(venues) # flatten JSON

# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.lat',␣
 ↪'venue.location.lng']
nearby_venues =nearby_venues.loc[:, filtered_columns]

# filter the category for each row
nearby_venues['venue.categories'] = nearby_venues.apply(get_category_type,␣
 ↪axis=1)

# clean columns
nearby_venues.columns = [col.split(".")[-1] for col in nearby_venues.columns]

nearby_venues.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-
packages/ipykernel_launcher.py:5: FutureWarning: pandas.io.json.json_normalize
is deprecated, use pandas.json_normalize instead
  """
```

[11]:

	name	categories	lat	lng
0	El Yagrumo	Caribbean Restaurant	18.326534	-66.349893
1	El Gran Cafe Restaurant	Breakfast Spot	18.340957	-66.316987
2	El Limo Viejo	Bar	18.338022	-66.322291
3	Balalaika Restaurant	Burger Joint	18.341657	-66.305472
4	El Rancho en Corozal	Pool	18.319297	-66.292644

[12]:
```
print('{} venues were returned by Foursquare.'.format(nearby_venues.shape[0]))
```

```
10 venues were returned by Foursquare.
```

[13]:
```
def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?
↪&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item␣
↪in venue_list])
    nearby_venues.columns = ['Neighborhood',
                  'Neighborhood Latitude',
                  'Neighborhood Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

[14]:
```
PR_venues = getNearbyVenues(names=join_df['County'],
                                    latitudes=join_df['LAT'],
                                    longitudes=join_df['LNG']
```

```
)
```

```
Adjuntas
Aguada
Aguadilla
Maricao
Anasco
Arecibo
Arecibo
Barceloneta
Cabo Rojo
Cabo Rojo
Penuelas
Camuy
Lares
Sabana Grande
Ciales
Utuado
Dorado
Guanica
Florida
Arecibo
Guanica
Guayanilla
Hatillo
Hormigueros
Isabela
Jayuya
Lajas
Lares
Las Marias
Manati
Moca
Rincon
Quebradillas
Mayaguez
Mayaguez
San German
San Sebastian
Morovis
Arecibo
Aguadilla
Vega Alta
Vega Baja
Vega Baja
Yauco
Aguas Buenas
```

Salinas
Aibonito
Maunabo
Arroyo
Ponce
Ponce
Ponce
Naguabo
Naranjito
Orocovis
Patillas
Caguas
Caguas
Ponce
Canovanas
Ponce
Ponce
Ceiba
Cayey
Fajardo
Cidra
Fajardo
Humacao
Rio Grande
Salinas
San Lorenzo
Santa Isabel
Vieques
Villalba
Yabucoa
Coamo
Las Piedras
Loiza
Luquillo
Culebra
Juncos
Gurabo
Ponce
Comerio
Corozal
Guayama
Aibonito
Humacao
Barranquitas
Juana Diaz
San Juan
San Juan
San Juan

```
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
San Juan
Guaynabo
San Juan
Toa Baja
Dorado
Toa Baja
Toa Baja
Toa Alta
Bayamon
Bayamon
Bayamon
Bayamon
Bayamon
Catano
Guaynabo
Guaynabo
Guaynabo
Guaynabo
Guaynabo
Trujillo Alta
Carolina
Carolina
Carolina
Carolina
Carolina
```

[15]:
```
print(PR_venues.shape)
PR_venues.head()
```

```
(554, 7)
```

[15]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	\
0	Adjuntas	18.180555	-66.749961	
1	Aguada	18.361945	-67.175597	

```
2       Aguada              18.361945              -67.175597
3       Aguada              18.361945              -67.175597
4       Aguada              18.361945              -67.175597

                         Venue  Venue Latitude  Venue Longitude  \
0              Aeropuerto Ruyán       18.179777       -66.754370
1            El Sarten Criollo       18.360852       -67.172921
2                      Club 4x4       18.364132       -67.176337
3  Panaderia Reposteria Lorenzo       18.364824       -67.176945
4                   El Tal-Ivan       18.360521       -67.172333

         Venue Category
0               Airport
1  Caribbean Restaurant
2                   Bar
3                Bakery
4                   Bar
```

[16]:
```
PR_venues.groupby('Neighborhood').count()
```

[16]:

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue
Adjuntas	1	1	1
Aguada	5	5	5
Aguadilla	7	7	7
Aibonito	5	5	5
Barceloneta	3	3	3
Barranquitas	1	1	1
Bayamon	48	48	48
Cabo Rojo	6	6	6
Caguas	18	18	18
Carolina	34	34	34
Catano	1	1	1
Cidra	5	5	5
Comerio	2	2	2
Corozal	1	1	1
Fajardo	10	10	10
Guayanilla	6	6	6
Guaynabo	61	61	61
Gurabo	2	2	2
Hormigueros	6	6	6
Humacao	4	4	4
Isabela	2	2	2
Juncos	1	1	1
Las Piedras	2	2	2
Manati	2	2	2
Mayaguez	4	4	4

Moca	2	2	2
Morovis	2	2	2
Naguabo	1	1	1
Naranjito	4	4	4
Ponce	16	16	16
Quebradillas	1	1	1
Rincon	1	1	1
Rio Grande	1	1	1
Salinas	1	1	1
San Juan	276	276	276
San Lorenzo	3	3	3
Santa Isabel	3	3	3
Toa Baja	1	1	1
Utuado	1	1	1
Vega Alta	1	1	1
Villalba	2	2	2
Yauco	1	1	1

Neighborhood	Venue Latitude	Venue Longitude	Venue Category
Adjuntas	1	1	1
Aguada	5	5	5
Aguadilla	7	7	7
Aibonito	5	5	5
Barceloneta	3	3	3
Barranquitas	1	1	1
Bayamon	48	48	48
Cabo Rojo	6	6	6
Caguas	18	18	18
Carolina	34	34	34
Catano	1	1	1
Cidra	5	5	5
Comerio	2	2	2
Corozal	1	1	1
Fajardo	10	10	10
Guayanilla	6	6	6
Guaynabo	61	61	61
Gurabo	2	2	2
Hormigueros	6	6	6
Humacao	4	4	4
Isabela	2	2	2
Juncos	1	1	1
Las Piedras	2	2	2
Manati	2	2	2
Mayaguez	4	4	4
Moca	2	2	2
Morovis	2	2	2

```
Naguabo                    1                1                1
Naranjito                  4                4                4
Ponce                     16               16               16
Quebradillas               1                1                1
Rincon                     1                1                1
Rio Grande                 1                1                1
Salinas                    1                1                1
San Juan                 276              276              276
San Lorenzo                3                3                3
Santa Isabel               3                3                3
Toa Baja                   1                1                1
Utuado                     1                1                1
Vega Alta                  1                1                1
Villalba                   2                2                2
Yauco                      1                1                1
```

[17]:
```
print('There are {} uniques categories.'.format(len(PR_venues['Venue Category'].
 ↪unique())))
```

```
There are 156 uniques categories.
```

[18]:
```
# one hot encoding
PR_onehot = pd.get_dummies(PR_venues[['Venue Category']], prefix="",␣
 ↪prefix_sep="")

# add neighborhood column back to dataframe
PR_onehot['Neighborhood'] = PR_venues['Neighborhood']

# move neighborhood column to the first column
fixed_columns = [PR_onehot.columns[-1]] + list(PR_onehot.columns[:-1])
PR_onehot = PR_onehot[fixed_columns]

PR_onehot.head()
```

[18]:

	Neighborhood	Accessories Store	Airport	Airport Lounge	Airport Terminal	\
0	Adjuntas	0	1	0	0	
1	Aguada	0	0	0	0	
2	Aguada	0	0	0	0	
3	Aguada	0	0	0	0	
4	Aguada	0	0	0	0	

	American Restaurant	Argentinian Restaurant	Art Museum	\
0	0	0	0	
1	0	0	0	
2	0	0	0	
3	0	0	0	
4	0	0	0	

```
   Arts & Crafts Store  Asian Restaurant  …  Vegetarian / Vegan Restaurant  \
0                    0                 0  …                              0
1                    0                 0  …                              0
2                    0                 0  …                              0
3                    0                 0  …                              0
4                    0                 0  …                              0

   Veterinarian  Video Store  Volleyball Court  Warehouse Store  Whisky Bar  \
0             0            0                 0                0           0
1             0            0                 0                0           0
2             0            0                 0                0           0
3             0            0                 0                0           0
4             0            0                 0                0           0

   Wine Shop  Winery  Wings Joint  Women's Store
0          0       0            0              0
1          0       0            0              0
2          0       0            0              0
3          0       0            0              0
4          0       0            0              0

[5 rows x 157 columns]
```

[19]:
```
PR_grouped = PR_onehot.groupby('Neighborhood').mean().reset_index()
PR_grouped
```

[19]:
```
     Neighborhood  Accessories Store  Airport  Airport Lounge  \
0        Adjuntas               0.00   1.0000          0.0000
1          Aguada               0.00   0.0000          0.0000
2       Aguadilla               0.00   0.0000          0.0000
3        Aibonito               0.00   0.0000          0.0000
4     Barceloneta               0.00   0.0000          0.0000
5    Barranquitas               0.00   0.0000          0.0000
6         Bayamon               0.00   0.0000          0.0000
7       Cabo Rojo               0.00   0.0000          0.0000
8          Caguas               0.00   0.0000          0.0000
9        Carolina               0.00   0.0000          0.0000
10         Catano               0.00   0.0000          0.0000
11          Cidra               0.00   0.0000          0.0000
12        Comerio               0.00   0.0000          0.0000
13        Corozal               0.00   0.0000          0.0000
14        Fajardo               0.00   0.0000          0.0000
15     Guayanilla               0.00   0.0000          0.0000
16       Guaynabo               0.00   0.0000          0.0000
17         Gurabo               0.00   0.0000          0.0000
18    Hormigueros               0.00   0.0000          0.0000
```

19	Humacao	0.00	0.0000	0.0000
20	Isabela	0.00	0.0000	0.0000
21	Juncos	0.00	0.0000	0.0000
22	Las Piedras	0.00	0.0000	0.0000
23	Manati	0.00	0.0000	0.0000
24	Mayaguez	0.00	0.0000	0.0000
25	Moca	0.00	0.0000	0.0000
26	Morovis	0.00	0.0000	0.0000
27	Naguabo	0.00	0.0000	0.0000
28	Naranjito	0.25	0.0000	0.0000
29	Ponce	0.00	0.0625	0.0625
30	Quebradillas	0.00	0.0000	0.0000
31	Rincon	0.00	0.0000	0.0000
32	Rio Grande	0.00	0.0000	0.0000
33	Salinas	0.00	0.0000	0.0000
34	San Juan	0.00	0.0000	0.0000
35	San Lorenzo	0.00	0.0000	0.0000
36	Santa Isabel	0.00	0.0000	0.0000
37	Toa Baja	0.00	0.0000	0.0000
38	Utuado	0.00	0.0000	0.0000
39	Vega Alta	0.00	0.0000	0.0000
40	Villalba	0.00	0.0000	0.0000
41	Yauco	0.00	0.0000	0.0000

	Airport Terminal	American Restaurant	Argentinian Restaurant	Art Museum	\
0	0.0000	0.000000	0.000000	0.000000	
1	0.0000	0.000000	0.000000	0.000000	
2	0.0000	0.000000	0.000000	0.000000	
3	0.0000	0.000000	0.000000	0.000000	
4	0.0000	0.000000	0.000000	0.000000	
5	0.0000	0.000000	0.000000	0.000000	
6	0.0000	0.000000	0.000000	0.000000	
7	0.0000	0.000000	0.000000	0.000000	
8	0.0000	0.000000	0.000000	0.000000	
9	0.0000	0.000000	0.000000	0.000000	
10	0.0000	0.000000	0.000000	0.000000	
11	0.0000	0.000000	0.000000	0.000000	
12	0.0000	0.000000	0.000000	0.000000	
13	0.0000	0.000000	0.000000	0.000000	
14	0.0000	0.000000	0.000000	0.000000	
15	0.0000	0.000000	0.000000	0.000000	
16	0.0000	0.000000	0.081967	0.000000	
17	0.0000	0.000000	0.000000	0.000000	
18	0.0000	0.000000	0.000000	0.000000	
19	0.0000	0.000000	0.000000	0.000000	
20	0.0000	0.000000	0.000000	0.000000	
21	0.0000	0.000000	0.000000	0.000000	

22	0.0000	0.000000	0.000000	0.000000
23	0.0000	0.000000	0.000000	0.000000
24	0.0000	0.000000	0.000000	0.000000
25	0.0000	0.000000	0.000000	0.000000
26	0.0000	0.000000	0.000000	0.000000
27	0.0000	0.000000	0.000000	0.000000
28	0.0000	0.000000	0.000000	0.000000
29	0.0625	0.000000	0.000000	0.000000
30	0.0000	0.000000	0.000000	0.000000
31	0.0000	0.000000	0.000000	0.000000
32	0.0000	0.000000	0.000000	0.000000
33	0.0000	0.000000	0.000000	0.000000
34	0.0000	0.003623	0.007246	0.003623
35	0.0000	0.000000	0.000000	0.000000
36	0.0000	0.000000	0.000000	0.000000
37	0.0000	0.000000	0.000000	0.000000
38	0.0000	0.000000	0.000000	0.000000
39	0.0000	0.000000	0.000000	0.000000
40	0.0000	0.000000	0.000000	0.000000
41	0.0000	0.000000	0.000000	0.000000

	Arts & Crafts Store	Asian Restaurant	…	Vegetarian / Vegan Restaurant	\
0	0.000000	0.000000	…	0.000000	
1	0.000000	0.000000	…	0.000000	
2	0.000000	0.142857	…	0.000000	
3	0.000000	0.000000	…	0.000000	
4	0.000000	0.000000	…	0.000000	
5	0.000000	0.000000	…	0.000000	
6	0.020833	0.020833	…	0.000000	
7	0.000000	0.000000	…	0.000000	
8	0.000000	0.000000	…	0.000000	
9	0.000000	0.000000	…	0.000000	
10	0.000000	0.000000	…	0.000000	
11	0.000000	0.000000	…	0.000000	
12	0.000000	0.000000	…	0.000000	
13	0.000000	0.000000	…	0.000000	
14	0.000000	0.000000	…	0.000000	
15	0.000000	0.000000	…	0.000000	
16	0.016393	0.016393	…	0.000000	
17	0.000000	0.000000	…	0.000000	
18	0.000000	0.000000	…	0.000000	
19	0.000000	0.000000	…	0.000000	
20	0.000000	0.000000	…	0.000000	
21	0.000000	0.000000	…	0.000000	
22	0.000000	0.000000	…	0.000000	
23	0.000000	0.000000	…	0.000000	
24	0.000000	0.000000	…	0.000000	

```
25              0.000000          0.000000  …                     0.000000
26              0.000000          0.000000  …                     0.000000
27              0.000000          0.000000  …                     0.000000
28              0.000000          0.000000  …                     0.000000
29              0.000000          0.000000  …                     0.000000
30              0.000000          0.000000  …                     0.000000
31              0.000000          0.000000  …                     0.000000
32              0.000000          0.000000  …                     0.000000
33              0.000000          0.000000  …                     0.000000
34              0.003623          0.007246  …                     0.018116
35              0.000000          0.000000  …                     0.000000
36              0.000000          0.000000  …                     0.000000
37              0.000000          0.000000  …                     0.000000
38              0.000000          0.000000  …                     0.000000
39              0.000000          0.000000  …                     0.000000
40              0.000000          0.000000  …                     0.000000
41              0.000000          0.000000  …                     0.000000

    Veterinarian  Video Store  Volleyball Court  Warehouse Store  Whisky Bar  \
0       0.000000     0.000000          0.000000         0.000000    0.000000
1       0.000000     0.000000          0.000000         0.000000    0.000000
2       0.000000     0.000000          0.000000         0.000000    0.000000
3       0.000000     0.000000          0.000000         0.000000    0.000000
4       0.000000     0.000000          0.000000         0.000000    0.000000
5       0.000000     0.000000          0.000000         0.000000    0.000000
6       0.000000     0.020833          0.000000         0.020833    0.000000
7       0.000000     0.000000          0.000000         0.000000    0.000000
8       0.000000     0.000000          0.000000         0.000000    0.000000
9       0.000000     0.000000          0.000000         0.000000    0.000000
10      0.000000     0.000000          0.000000         0.000000    0.000000
11      0.000000     0.000000          0.000000         0.000000    0.000000
12      0.000000     0.000000          0.000000         0.000000    0.000000
13      0.000000     0.000000          0.000000         0.000000    0.000000
14      0.000000     0.000000          0.000000         0.000000    0.000000
15      0.000000     0.000000          0.000000         0.000000    0.000000
16      0.000000     0.000000          0.000000         0.000000    0.000000
17      0.000000     0.000000          0.000000         0.000000    0.000000
18      0.000000     0.000000          0.000000         0.000000    0.000000
19      0.000000     0.250000          0.000000         0.000000    0.000000
20      0.000000     0.000000          0.000000         0.000000    0.000000
21      0.000000     0.000000          0.000000         0.000000    0.000000
22      0.000000     0.000000          0.000000         0.000000    0.000000
23      0.000000     0.000000          0.000000         0.000000    0.000000
24      0.000000     0.000000          0.000000         0.000000    0.000000
25      0.000000     0.000000          0.000000         0.000000    0.000000
26      0.000000     0.000000          0.000000         0.000000    0.000000
27      0.000000     0.000000          0.000000         0.000000    0.000000
```

28	0.000000	0.000000	0.000000	0.000000	0.000000
29	0.000000	0.000000	0.000000	0.000000	0.000000
30	0.000000	0.000000	0.000000	0.000000	0.000000
31	0.000000	0.000000	0.000000	0.000000	0.000000
32	0.000000	0.000000	0.000000	0.000000	0.000000
33	0.000000	0.000000	0.000000	0.000000	0.000000
34	0.003623	0.000000	0.003623	0.000000	0.007246
35	0.000000	0.000000	0.000000	0.000000	0.000000
36	0.000000	0.000000	0.000000	0.000000	0.000000
37	0.000000	0.000000	0.000000	0.000000	0.000000
38	0.000000	0.000000	0.000000	0.000000	0.000000
39	0.000000	0.000000	0.000000	0.000000	0.000000
40	0.000000	0.000000	0.000000	0.000000	0.000000
41	0.000000	1.000000	0.000000	0.000000	0.000000

	Wine Shop	Winery	Wings Joint	Women's Store
0	0.000000	0.000000	0.000000	0.000000
1	0.000000	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.020833
7	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000	0.000000
14	0.000000	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000	0.000000
16	0.000000	0.016393	0.000000	0.000000
17	0.000000	0.000000	0.000000	0.000000
18	0.000000	0.000000	0.000000	0.000000
19	0.000000	0.000000	0.000000	0.000000
20	0.000000	0.000000	0.000000	0.000000
21	0.000000	0.000000	0.000000	0.000000
22	0.000000	0.000000	0.000000	0.000000
23	0.000000	0.000000	0.000000	0.000000
24	0.000000	0.000000	0.000000	0.000000
25	0.000000	0.000000	0.000000	0.000000
26	0.000000	0.000000	0.000000	0.000000
27	0.000000	0.000000	0.000000	0.000000
28	0.000000	0.000000	0.000000	0.000000
29	0.062500	0.000000	0.000000	0.000000
30	0.000000	0.000000	0.000000	0.000000

```
31   0.000000  0.000000     0.000000       0.000000
32   0.000000  0.000000     0.000000       0.000000
33   0.000000  0.000000     0.000000       0.000000
34   0.003623  0.000000     0.003623       0.000000
35   0.000000  0.000000     0.000000       0.000000
36   0.000000  0.000000     0.000000       0.000000
37   0.000000  0.000000     0.000000       0.000000
38   0.000000  0.000000     0.000000       0.000000
39   0.000000  0.000000     0.000000       0.000000
40   0.000000  0.000000     0.000000       0.000000
41   0.000000  0.000000     0.000000       0.000000

[42 rows x 157 columns]
```

[20]:
```
num_top_venues = 5

for hood in PR_grouped['Neighborhood']:
    print("----"+hood+"----")
    temp = PR_grouped[PR_grouped['Neighborhood'] == hood].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).
 ↪head(num_top_venues))
    print('\n')
```

```
----Adjuntas----
                  venue  freq
0               Airport   1.0
1             Nightclub   0.0
2          Noodle House   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Aguada----
                  venue  freq
0                   Bar   0.4
1                 Hotel   0.2
2  Caribbean Restaurant   0.2
3               Bakery   0.2
4     Accessories Store   0.0


----Aguadilla----
                venue  freq
```

```
0      Sandwich Place  0.29
1             Bakery  0.14
2  Italian Restaurant  0.14
3       Cocktail Bar  0.14
4    Asian Restaurant  0.14


----Aibonito----
                  venue  freq
0       Scenic Lookout   0.2
1                Diner   0.2
2  Caribbean Restaurant   0.2
3     Basketball Court   0.2
4                  Bar   0.2


----Barceloneta----
               venue  freq
0     Auto Workshop  0.33
1              Food  0.33
2     Grocery Store  0.33
3          Pharmacy  0.00
4  Paella Restaurant  0.00


----Barranquitas----
                  venue  freq
0     Basketball Court   1.0
1    Accessories Store   0.0
2            Nightclub   0.0
3  Other Great Outdoors   0.0
4    Paella Restaurant   0.0


----Bayamon----
                  venue  freq
0  Fast Food Restaurant  0.08
1           Food Truck  0.06
2        Shopping Mall  0.06
3                  Gym  0.04
4     Department Store  0.04


----Cabo Rojo----
                venue  freq
0          Restaurant  0.17
1   Electronics Store  0.17
2          Food Truck  0.17
```

```
3  Chinese Restaurant  0.17
4              Bakery  0.17


----Caguas----
                  venue  freq
0  Fast Food Restaurant  0.17
1            Steakhouse  0.11
2           Supermarket  0.11
3            Food Truck  0.06
4      Sushi Restaurant  0.06


----Carolina----
                  venue  freq
0          Pizza Place  0.09
1  Fast Food Restaurant  0.09
2  Caribbean Restaurant  0.09
3   Fried Chicken Joint  0.06
4              Pharmacy  0.06


----Catano----
                  venue  freq
0                   Bar   1.0
1     Accessories Store   0.0
2             Nightclub   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Cidra----
               venue  freq
0        Pizza Place   0.4
1           Pharmacy   0.2
2              Plaza   0.2
3               Café   0.2
4  Paella Restaurant   0.0


----Comerio----
                  venue  freq
0  Fast Food Restaurant   0.5
1          Post Office   0.5
2     Accessories Store   0.0
3              Pharmacy   0.0
4     Paella Restaurant   0.0
```

```
----Corozal----
                  venue  freq
0  Medical Supply Store   1.0
1     Accessories Store   0.0
2              Pharmacy   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Fajardo----
             venue  freq
0    Boat or Ferry   0.4
1             Café   0.2
2            Hotel   0.1
3           Hostel   0.1
4  Harbor / Marina   0.1


----Guayanilla----
                  venue  freq
0  Fast Food Restaurant  0.33
1        Sandwich Place  0.17
2                 Hotel  0.17
3                Casino  0.17
4          Dessert Shop  0.17


----Guaynabo----
                    venue  freq
0                Pharmacy  0.08
1  Argentinian Restaurant  0.08
2        Sushi Restaurant  0.05
3              Restaurant  0.05
4          Ice Cream Shop  0.05


----Gurabo----
               venue  freq
0         Campground   0.5
1         Playground   0.5
2  Accessories Store   0.0
3           Pharmacy   0.0
4  Paella Restaurant   0.0


----Hormigueros----
               venue  freq
```

```
0  Convenience Store  0.17
1      Shopping Mall  0.17
2     Sandwich Place  0.17
3  Electronics Store  0.17
4         Food Truck  0.17


----Humacao----
                 venue  freq
0               Resort  0.25
1                Beach  0.25
2   Seafood Restaurant  0.25
3          Video Store  0.25
4  Peruvian Restaurant  0.00


----Isabela----
                           venue  freq
0                    Coffee Shop   0.5
1             Mexican Restaurant   0.5
2                   Noodle House   0.0
3              Paella Restaurant   0.0
4  Paper / Office Supplies Store   0.0


----Juncos----
                  venue  freq
0        Baseball Field   1.0
1     Accessories Store   0.0
2             Nightclub   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Las Piedras----
                  venue  freq
0  Caribbean Restaurant   0.5
1               Bakery   0.5
2     Accessories Store   0.0
3              Pharmacy   0.0
4     Paella Restaurant   0.0


----Manati----
                venue  freq
0      Ice Cream Shop   0.5
1  Chinese Restaurant   0.5
2   Accessories Store   0.0
```

```
3                Pier   0.0
4   Paella Restaurant   0.0


----Mayaguez----
                  venue  freq
0               Resort  0.25
1  Caribbean Restaurant  0.25
2       Harbor / Marina  0.25
3            Food Truck  0.25
4              Pharmacy  0.00


----Moca----
                  venue  freq
0     Recreation Center   0.5
1                   Bar   0.5
2     Accessories Store   0.0
3   Peruvian Restaurant   0.0
4  Other Great Outdoors   0.0


----Morovis----
                  venue  freq
0                Market   0.5
1           Beer Garden   0.5
2     Accessories Store   0.0
3              Pharmacy   0.0
4  Other Great Outdoors   0.0


----Naguabo----
                  venue  freq
0      Basketball Court   1.0
1     Accessories Store   0.0
2             Nightclub   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Naranjito----
               venue  freq
0  Accessories Store  0.25
1        Pizza Place  0.25
2        Beer Garden  0.25
3         Restaurant  0.25
4                Pub  0.00
```

```
----Ponce----
                venue  freq
0              Bakery  0.12
1          Sports Bar  0.12
2  Italian Restaurant  0.06
3  Spanish Restaurant  0.06
4              Museum  0.06


----Quebradillas----
                           venue  freq
0              Convenience Store   1.0
1              Accessories Store   0.0
2                           Pier   0.0
3              Paella Restaurant   0.0
4  Paper / Office Supplies Store   0.0


----Rincon----
                       venue  freq
0  Latin American Restaurant   1.0
1          Accessories Store   0.0
2                   Pharmacy   0.0
3       Other Great Outdoors   0.0
4          Paella Restaurant   0.0


----Rio Grande----
                           venue  freq
0                         Resort   1.0
1                        Pharmacy   0.0
2           Other Great Outdoors   0.0
3              Paella Restaurant   0.0
4  Paper / Office Supplies Store   0.0


----Salinas----
                  venue  freq
0                 Beach   1.0
1     Accessories Store   0.0
2             Nightclub   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----San Juan----
                  venue  freq
```

```
0  Caribbean Restaurant  0.06
1                   Bar  0.05
2    Mexican Restaurant  0.04
3            Food Truck  0.03
4                 Café  0.03


----San Lorenzo----
                  venue  freq
0        Baseball Field  0.67
1                 River  0.33
2             Nightclub  0.00
3  Other Great Outdoors  0.00
4     Paella Restaurant  0.00


----Santa Isabel----
                  venue  freq
0                   Bar  0.67
1             Gastropub  0.33
2     Accessories Store  0.00
3   Peruvian Restaurant  0.00
4  Other Great Outdoors  0.00


----Toa Baja----
                  venue  freq
0    Seafood Restaurant   1.0
1     Accessories Store   0.0
2              Pharmacy   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Utuado----
                  venue  freq
0            Food Truck   1.0
1     Accessories Store   0.0
2              Pharmacy   0.0
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Vega Alta----
                  venue  freq
0           Music Store   1.0
1     Accessories Store   0.0
2             Nightclub   0.0
```

```
3  Other Great Outdoors   0.0
4     Paella Restaurant   0.0


----Villalba----
                  venue  freq
0                   Bar   0.5
1               Bakery   0.5
2     Accessories Store   0.0
3   Peruvian Restaurant   0.0
4  Other Great Outdoors   0.0


----Yauco----
                  venue  freq
0           Video Store   1.0
1     Accessories Store   0.0
2   Peruvian Restaurant   0.0
3          Noodle House   0.0
4  Other Great Outdoors   0.0
```

[21]:
```
def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

[22]:
```
import numpy as np

num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['Neighborhood']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = PR_grouped['Neighborhood']
```

```
for ind in np.arange(PR_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] =␣
↪return_most_common_venues(PR_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

[22]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adjuntas	Airport	Women's Store	Dive Bar	Food	Flower Shop	Fast Food Restaurant	Farmers Market	Fabric Shop	Electronics Store	Donut Shop
1	Aguada	Bar	Hotel	Caribbean Restaurant	Bakery	Women's Store	Donut Shop	Flower Shop	Fast Food Restaurant	Farmers Market	Fabric Shop
2	Aguadilla	Sandwich Place	Italian Restaurant	Asian Restaurant	Burger Joint	Bakery	Cocktail Bar	Argentinian Restaurant	Art Museum	Food	Flower Shop
3	Aibonito	Basketball Court	Caribbean Restaurant	Scenic Lookout	Bar	Diner	Women's Store	Dive Bar	Fast Food Restaurant	Farmers Market	Fabric Shop
4	Barceloneta	Food	Auto Workshop	Grocery Store	Flower Shop	Fast Food Restaurant	Farmers Market	Fabric Shop	Electronics Store	Donut Shop	Dive Bar

[23]:
```
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans

# set number of clusters
kclusters = 10

PR_grouped_clustering = PR_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(PR_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

[23]:
```
array([5, 0, 5, 5, 5, 6, 5, 5, 5, 5], dtype=int32)
```

[24]:
```
# add clustering labels
neighborhoods_venues_sorted.insert(0, 'Cluster Labels 2', kmeans.labels_)

PR_merged = join_df
neighborhoods_venues_sorted= neighborhoods_venues_sorted.rename(columns =␣
 ↪{"Neighborhood":"County"})

# merge manhattan_grouped with manhattan_data to add latitude/longitude for␣
 ↪each neighborhood
PR_merged = PR_merged.join(neighborhoods_venues_sorted.set_index('County'),␣
 ↪on='County')

PR_merged.head() # check the last columns!
```

[24]:

	ZIP	Borough	County	LAT	LNG	Cluster Labels 2	\
0	00601	Adjuntas	Adjuntas	18.180555	-66.749961	5.0	
1	00602	Aguada	Aguada	18.361945	-67.175597	0.0	
2	00603	Aguadilla	Aguadilla	18.455183	-67.119887	5.0	
3	00606	Maricao	Maricao	18.158345	-66.932911	NaN	
4	00610	Anasco	Anasco	18.295366	-67.125135	NaN	

	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	\
0	Airport	Women's Store	Dive Bar	
1	Bar	Hotel	Caribbean Restaurant	
2	Sandwich Place	Italian Restaurant	Asian Restaurant	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	

	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	\
0	Food	Flower Shop	Fast Food Restaurant	
1	Bakery	Women's Store	Donut Shop	
2	Burger Joint	Bakery	Cocktail Bar	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	

	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	\
0	Farmers Market	Fabric Shop	Electronics Store	
1	Flower Shop	Fast Food Restaurant	Farmers Market	
2	Argentinian Restaurant	Art Museum	Food	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	

	10th Most Common Venue
0	Donut Shop
1	Fabric Shop

```
2            Flower Shop
3                    NaN
4                    NaN
```

[25]:
```
from matplotlib import cm
import matplotlib.colors as colors

PR_merged = PR_merged.fillna(0)
PR_merged['Cluster Labels 2'] = PR_merged['Cluster Labels 2'].astype(int)

# create map
map_clusters = folium.Map(location=[18.303910000000002, -66.32617900000001],␣
 ↪zoom_start=9)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(PR_merged['LAT'], PR_merged['LNG'],␣
 ↪PR_merged['County'], PR_merged['Cluster Labels 2']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

[25]:
```
<folium.folium.Map at 0x7fc335890ba8>
```

4. Results

There are 554 venues in the Foursquare's Places API under zip codes from Puerto Rico. These venues belong to 156 unique categories in 42 counties.

The following counties had the greatest number of venues:

San Juan -276

Guaynabo - 61

Bayamon - 48

Carolina - 34

Caguas - 18

These counties are located in the metro area of Puerto Rico.

The following counties had the lowest number of venues (just one):

Adjuntas

Barranquitas

Catano

Corozal

Juncos

Naguabo

Quebradillas

Rincon

Rio Grande

Salinas

Toa Baja

Utuado

Vega Alta

Yauco

Most of these counties are located out of the metro area of Puerto Rico.

5. Discussion

There are some clear patterns in the clusters:

Cluster 1 is composed of San Lorenzo and Juncos. Making this cluster an area for sports such as Baseball Fields.

Cluster 8 is composed of Vega Alta. Making this cluster an area for Music Stores.

Cluster 9 is composed of Rincon. Making this cluster an are for Latin American restaurants.

6. Conclusion

Although we can say that this strategy for data analysis is a correct and useful approach for decision making in a business ecosystem, the project has some limitations. The main limitation is the low representation of Puerto Rican businesses in the API used. This makes the determination of clusters not so effective. In order to make decisions based on these results, the effort must be accompanied with a strategy for more businesses to complete their profile on the Foursquare platform.

[ ]: