CAPSTONE PROJECT

—

Venue Recommendations to place a Mexican Restaurant

Introduction

Due in part to business and immigration, United States has hispanic presence across several States, on which there is a widespread likability of Mexican food. In Addition, more traditional Mexican dishes have also grown in popularity in the United States,like tortillas, salsa, chips, chili, burritos, and tacos. Therefore, the intention of this project is to collect information and provide recommendations on potential places to place a Mexican restaurant based on statistical data, open sources and FourSquare API.

PROBLEM DESCRIPTION

Many Mexican people migrating to various states of United States would like to have traditional Mexican dishes near their location. The project aims to create an analysis of features for a place comparative analysis. The features include influence of Hispanic people, location near important venues and shopping malls. This would help decision makers to become aware of places of potential interest.

Data Sources

- For this business analysis we need to acquire geo-locational information about specific cities with a high influence of latin-american people, which have more likability of Mexican food.

- Census data of the United States Government already provides this information.

- The main point of interest of the stakeholders is the State of Maryland, as in they would have strategic Vendors on this State.

- In order to perform a comparison of the potential locations with a high venue and also potential competence nearby, we will use the Foursquare API. This dataset comprises latitude and longitude of landmarks, as well as ZIP Codes, which will be used as key to be cross-referenced with the other datasets.

DATASOURCE 1
US Government Census Data

	SUMLEV	STATE	COUNTY	STNAME	CTYNAME	SEX	ORIGIN	AGEGRP	IMPRACE	RESPOP
0	50	1	1	Alabama	Autauga County	1	1	1	1	1337
1	50	1	1	Alabama	Autauga County	1	1	1	2	356
2	50	1	1	Alabama	Autauga County	1	1	1	3	2
3	50	1	1	Alabama	Autauga County	1	1	1	4	13
4	50	1	1	Alabama	Autauga County	1	1	1	6	33

DATASOURCE 2
Maryland Government Population Data

	OBJECTID_1	ZCTA5CE10	FIRST_STAT	FIRST_GEOI	FIRST_CLAS	FIRST_MTFC	FIRST_FUNC	ZCTA5N	STATE	AREALAND	AREAWATR	POP100	HU100	NHW	NH
0	1	20601	24	2420601	B5	G6350	S	20601	24	115635266	387684	24156	8722	9785	111
1	2	20602	24	2420602	B5	G6350	S	20602	24	35830723	352762	24955	9736	8466	130
2	3	20603	24	2420603	B5	G6350	S	20603	24	44239637	219356	28967	10317	9625	150
3	4	20606	24	2420606	B5	G6350	S	20606	24	7501011	1248760	431	230	377	
4	5	20607	24	2420607	B5	G6350	S	20607	24	54357590	448221	9802	3504	2165	63

DATASOURCE 3
Foursquare API Data

	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	postalCode	state	
0	Bank of America	Bank	9111 Riggs Rd	US	Adelphi	United States	NaN	128	[9111 Riggs Rd, Adelphi, MD 20783, United States]	[{'label': 'display', 'lat': 39.00382745145823...	39.003827	-76.970979	20783	MD	4bbb3
1	Basil's Pizza	Pizza Place	9113 Riggs Rd	US	Adelphi	United States	NaN	140	[9113 Riggs Rd, Adelphi, MD 20783, United States]	[{'label': 'display', 'lat': 39.00380609518559...	39.003806	-76.970784	20783	MD	4baa
2	7-Eleven	Convenience Store	9151 Riggs Rd	US	Adelphi	United States	Adelphi Rd	329	[9151 Riggs Rd (Adelphi Rd), Adelphi, MD 20783...	[{'label': 'display', 'lat': 39.0044208023493,...	39.004421	-76.968715	20783	MD	4b99
3	Cool Spring ES	Student Center	NaN	US	Hyattsville	United States	NaN	69	[Hyattsville, MD 20783, United States]	[{'label': 'display', 'lat': 39.00276294590971...	39.002763	-76.972798	20783	MD	4eca
4	Rite Aid	Pharmacy	9139 Riggs Rd	US	Adelphi	United States	NaN	196	[9139 Riggs Rd, Adelphi, MD 20783, United States]	[{'label': 'display', 'lat': 39.00369666813991...	39.003697	-76.969983	20783	MD	4ce

Methodology

PYTHON LIBRARIES USED FOR THE ANALYSIS:

- Pandas
- NumPy
- GeoPy
- Folium
- Matplotlib
- kMeans

DATA SOURCE LOADING

- Path to the datasources.
- Census data of the United States Government:

```
#US Government Census on Immigration
df=pd.read_csv('https://www2.census.gov/programs-surveys/popest/datasets/2010/modified-race-data-2010/stco-mr2010-1.csv')
```

```
#Maryland Census on Population across Counties
Maryland_Data=pd.read_csv('https://opendata.arcgis.com/datasets/eb706b48117b43d482c63d02017fc3ff_1.csv')
```

- As the main point of interest of the stakeholders is the State of Maryland, we have considered more details on this specific location.

DATA EXPLORATION

Filtering of the State of interest

]:

	SUMLEV	STATE	COUNTY	STNAME	CTYNAME	SEX	ORIGIN	AGEGRP	IMPRACE	RESPOP
486942	50	24	1	Maryland	Allegany County	1	2	7	1	96
486944	50	24	1	Maryland	Allegany County	1	2	8	1	65
486948	50	24	1	Maryland	Allegany County	1	2	9	1	60
487150	50	24	1	Maryland	Allegany County	2	2	7	1	15
487154	50	24	1	Maryland	Allegany County	2	2	8	1	28

Perform Grouping of Relevant Counties by number of habitants and therefore review ZIP Codes with greather percentage of presence of Hispanics.

CTYNAME	RESPOP
Montgomery County	37272
Prince George's County	27380
Anne Arundel County	7084
Baltimore County	6627
Baltimore city	5262

```
ZCTA5CE10
20783    59.4
20737    48.1
20903    47.9
21649    44.4
20722    41.7
Name: PHISP, dtype: float64
```

DATA ANALYSIS

We filter the relevant ZIP Codes during the exploration phase, on which potentially the Restaurant can be placed:

	OBJECTID_1	ZCTA5CE10	FIRST_STAT	FIRST_GEOI	FIRST_CLAS	FIRST_MTFC	FIRST_FUNC	ZCTA5N	STATE	AREALAND	AREAWATR	POP100	HU100	NHW
66	67	20712	24	2420712	B5	G6350	S	20712	24	1892811	12394	9031	3927	1112
72	73	20722	24	2420722	B5	G6350	S	20722	24	3810192	79461	5711	1945	740
79	80	20737	24	2420737	B5	G6350	S	20737	24	10724278	76543	20684	6227	2449
105	106	20781	24	2420781	B5	G6350	S	20781	24	6312229	205096	11440	4093	2786
106	107	20782	24	2420782	B5	G6350	S	20782	24	10973600	99086	30560	11552	4656

We evaluate the population number based on each ZIP Code:

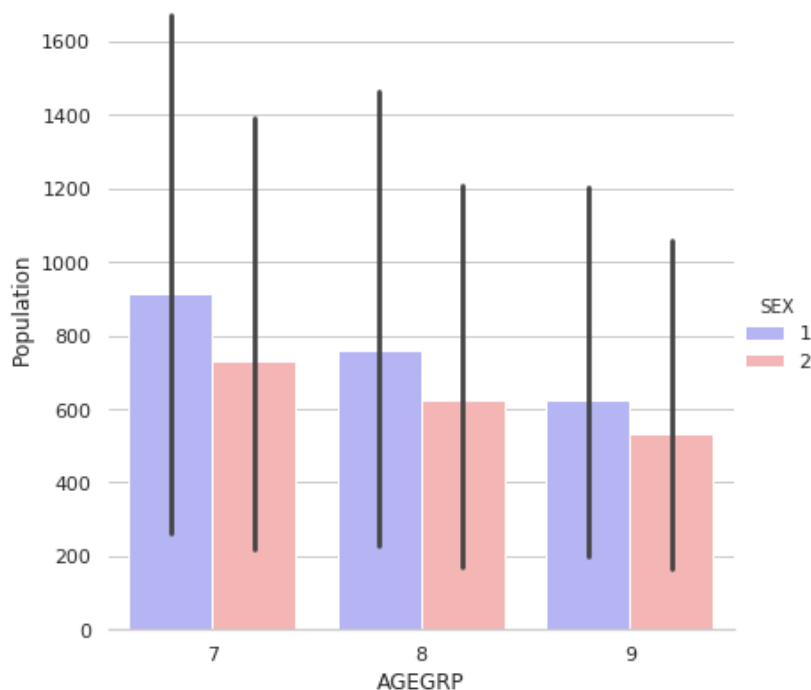| ZCTA5CE10	
20783	26424
20902	16532
20877	12872
20903	11323
20782	10294

Based on the analysis of the Census Data, we have determined that our main target audience is Hispanic people between 30 and 59 years:

Group 7: 30-39 years
Group 8: 40-49 years
Group 9: 50-59 years

FOURSQUARE

Connecting to the Foursquare API in order to retrieve relevant information about the places located nearby, which helps to evaluate the venue relevance. According to the Data Analysis on the Census Data and exploration of locations. The city of '**Adelphi, MD**' is of potential interest.

After finding the neighborhood of interest, we then connect to the Foursquare API to gather information about landmarks of interest on a radius of 500 meter.

One of the most relevant points of interest are the shopping malls nearby:

	Shopping_Center	Latitude	Longitude
0	Adelphi Shopping Center	38.984889	-76.967799
1	Adelphi Plaza	38.983813	-76.975459
2	Metzerott Plaza	39.004754	-76.969526

CLUSTERING

We prepare the data for clustering based on categories and closeness of relevant landmarks:

	name	lat	lng	American Restaurant	Bank	Chinese Restaurant	Church	Convenience Store	Deli / Bodega	Fast Food Restaurant	Financial or Legal Service	Fried Chicken Joint	Laundry Service	Liquor Store	Mobile Phone Shop	Pharmacy	Piz Pla
0	Bank of America	39.003827	-76.970979	0	1	0	0	0	0	0	0	0	0	0	0	0
1	Basil's Pizza	39.003806	-76.970784	0	0	0	0	0	0	0	0	0	0	0	0	0
2	7-Eleven	39.004421	-76.968715	0	0	0	0	1	0	0	0	0	0	0	0	0
3	Cool Spring ES	39.002763	-76.972798	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Rite Aid	39.003697	-76.969983	0	0	0	0	0	0	0	0	0	0	0	0	1

Then we apply the k-Means Clustering method, selecting 5 clusters and we append the result to the DataFrame:

```
##Clustering
cluster_LandmarksHispanicPlaces = LandmarksHispanicPlaces_firstcolumn.drop('name',1)
kclusters=5
kmeans_ = KMeans(n_clusters=kclusters, random_state=0).fit(cluster_LandmarksHispanicPlaces)
kmeans_.labels_
```
```
]: array([2, 1, 0, 0, 0, 0, 0, 0, 0, 4, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 1, 0,
          0], dtype=int32)
```

```
##Add the Cluster column to the DataFrame
LandmarksHispanicPlaces_firstcolumn['Cluster'] = pd.Series(kmeans_.labels_,index=LandmarksHispanicPlaces_firstcolumn.index)
LandmarksHispanicPlaces_firstcolumn.head()
```
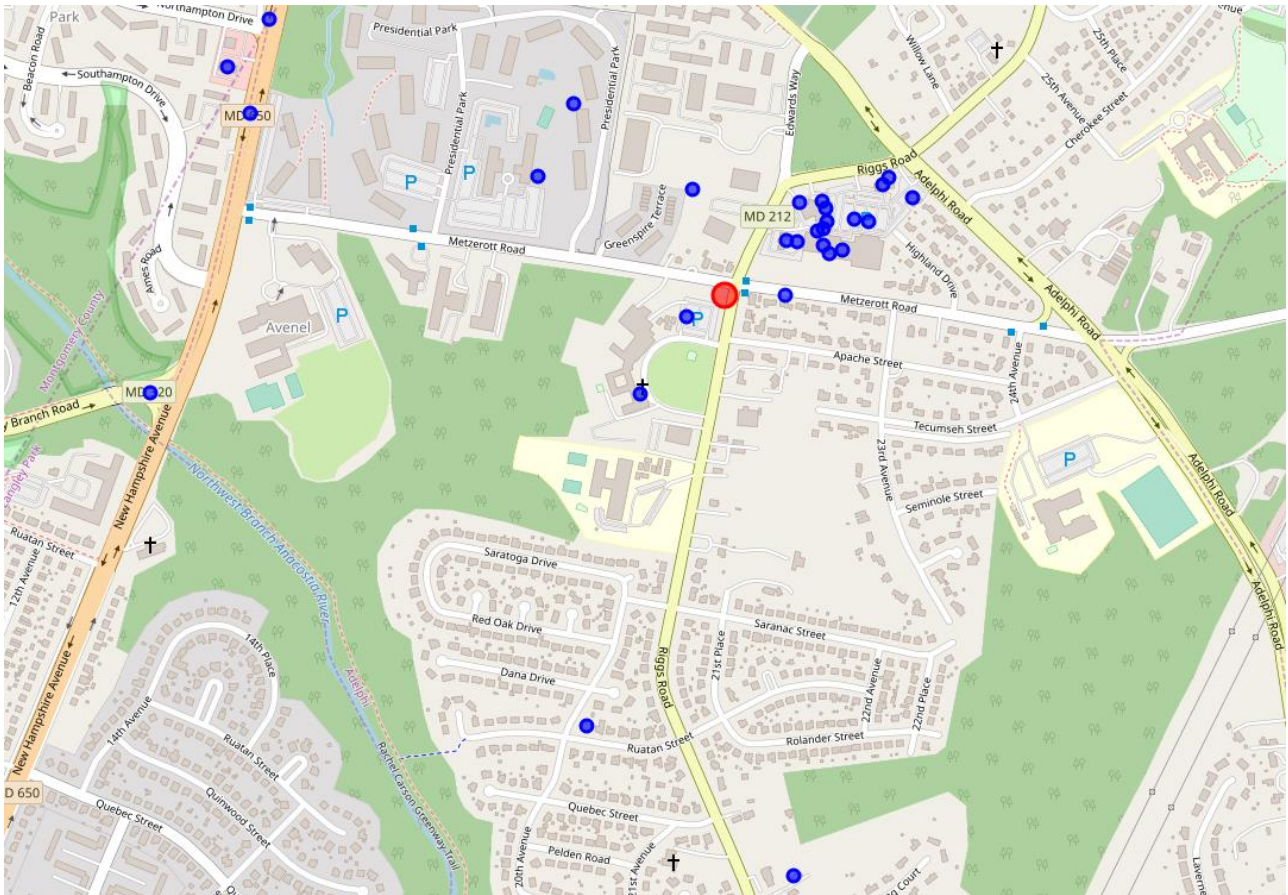
# Results

## MAPS GENERATED WITH FOLIUM BASED ON FOURQUARE API AND DATA SOURCES
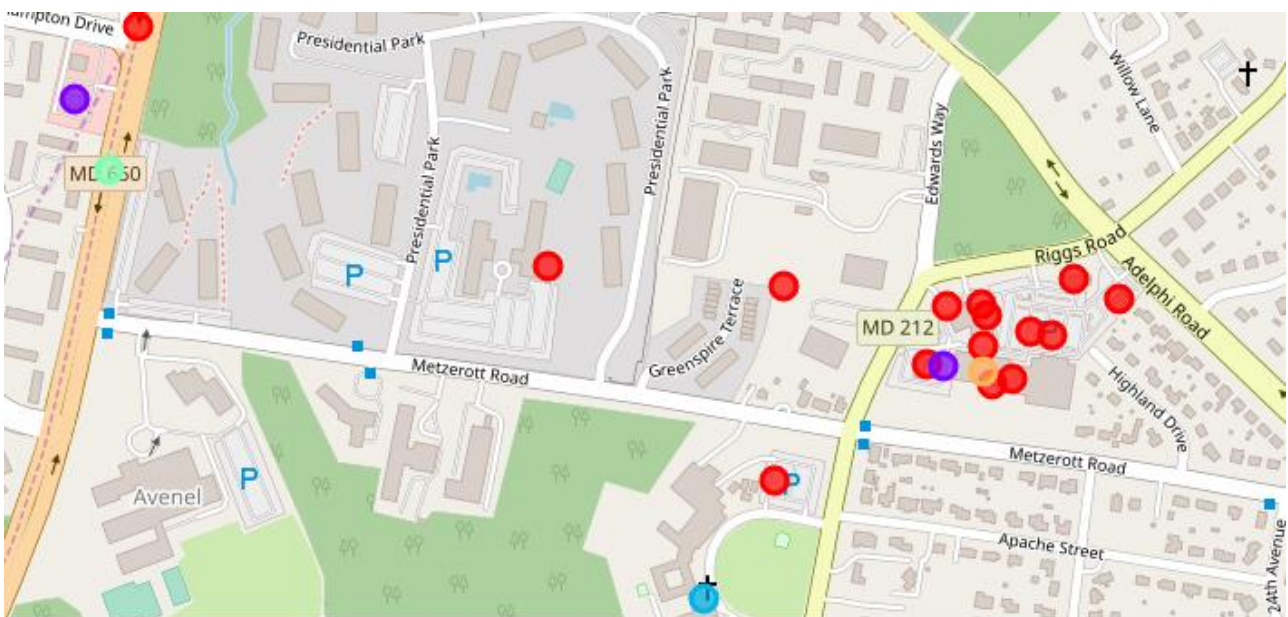
### Strategic Shopping Malls Nearby
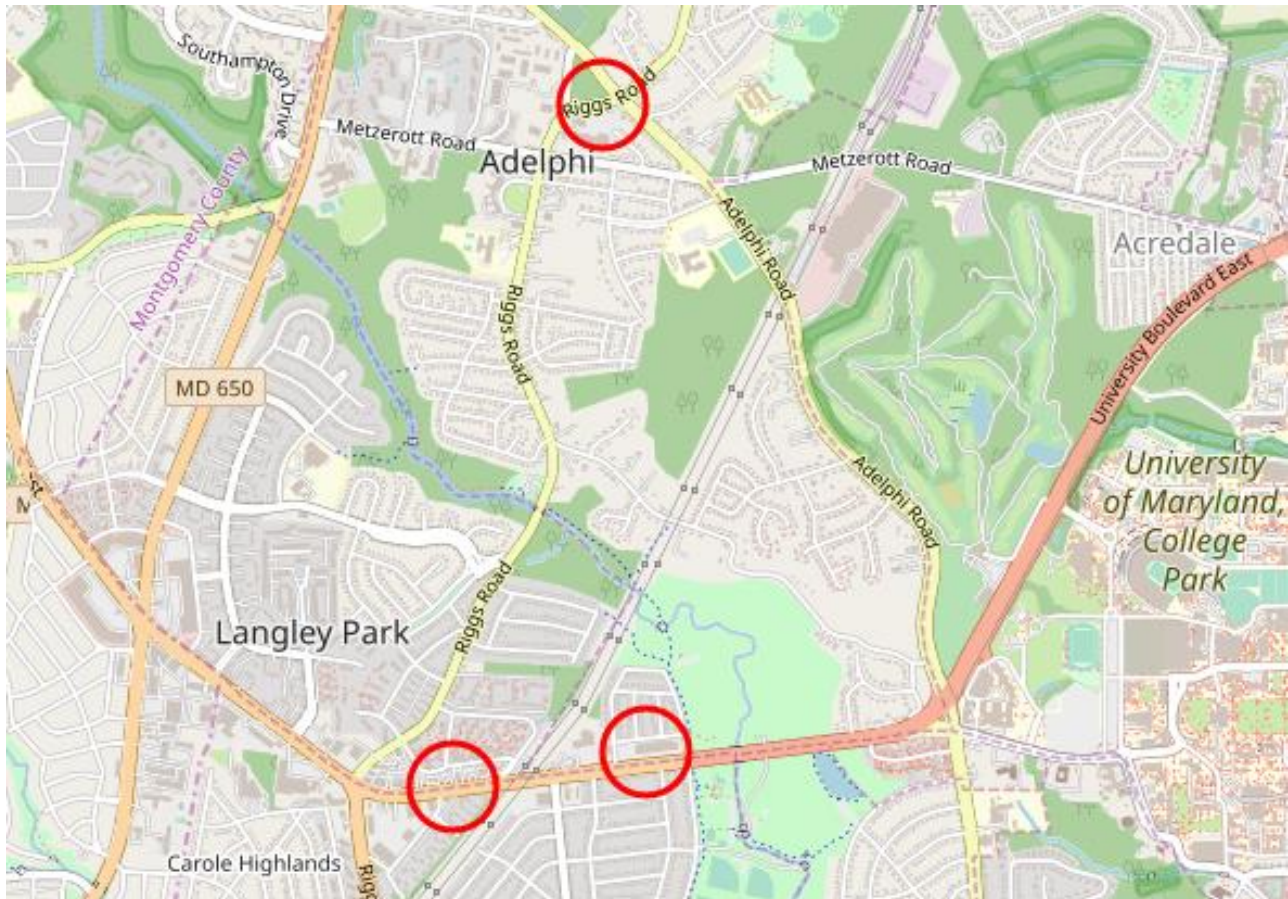


| | Shopping_Center | Latitude | Longitude |
|---|---|---|---|
| 0 | Adelphi Shopping Center | 38.984889 | -76.967799 |
| 1 | Adelphi Plaza | 38.983813 | -76.975459 |
| 2 | Metzerott Plaza | 39.004754 | -76.969526 |

## Relevant Landmarks near Adelphi City Center



## Clustering of relevant venues near Metzerott Plaza

# Conclusion

The city of Adelphi, MD, with a population of almost 60% of latinamerican people. Represents a good opportunity to place a Mexican Restaurant.

Based on Foursquare data, only one Mexican Restaurant was located called '*Los Antojitos*'. Therefore, a new Restaurant would represent an opportunity based on the potential demand for traditional dishes.

Based on clustering, we have identified two potential places near strategic Shopping Malls, which represent a higher concurrence of people.