IBM Capstone Project

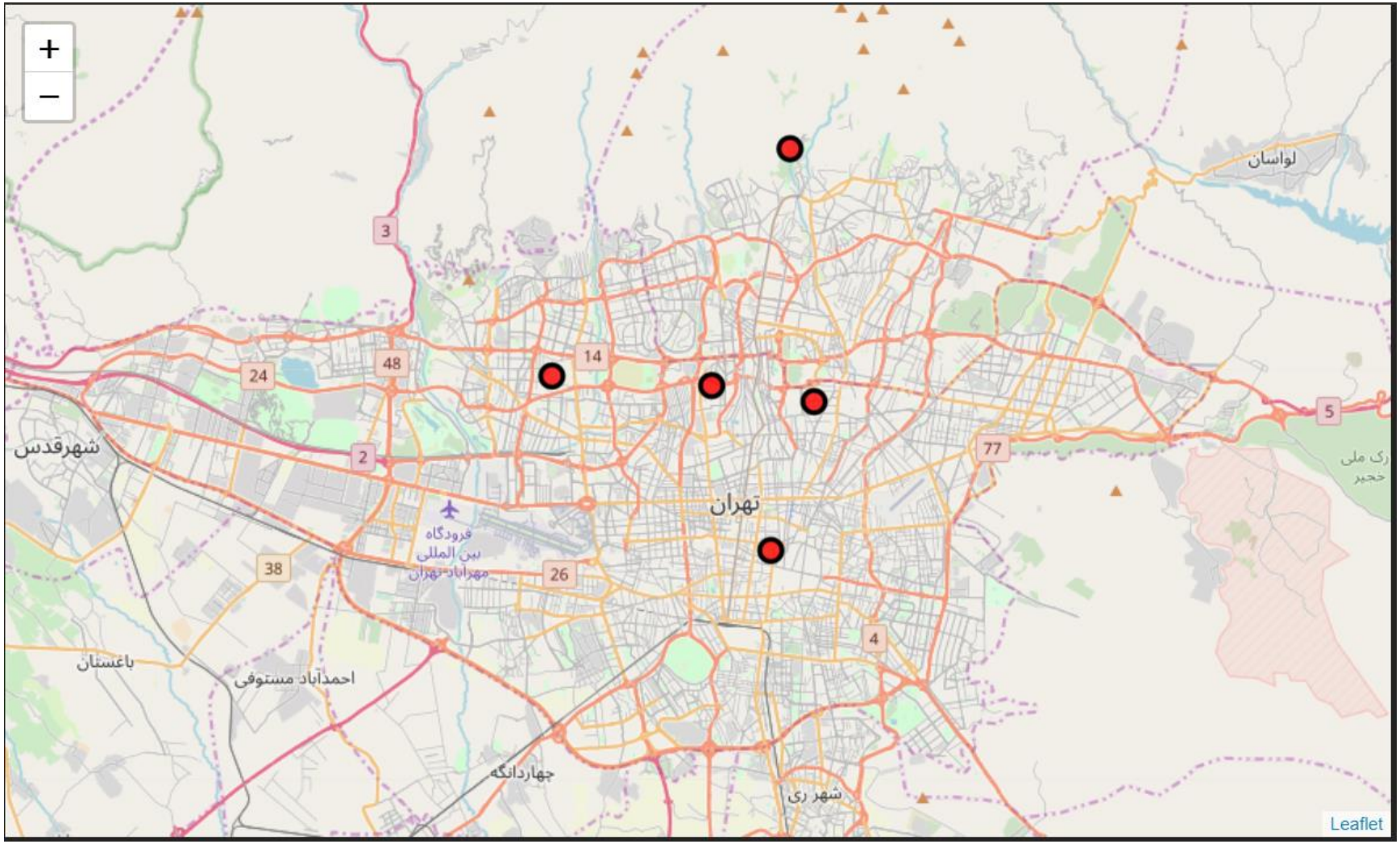
The battle of Neighborhoods: Tehran Venues analyze

Ashkan Yousefi Zadeh

Consider that you are coming to living in Tehran to start a business, you don't know much about this city, after a few searching about this city, you consider 6 neighborhoods for starting your business, but still, you don't know how to choose one of these neighborhoods which is the best location for your job. How you want to understand and solve the problem?! It′ s not 50 years ago when computers were not good enough so it's not proper for you to go walking in these neighborhoods and do your own research. So you need a consultant to tell you how you must choose the location?! And tells you why that location is the best for you? In this project, I chose 6 neighborhoods in Tehran and obtain their coordinates by Wikipedia and google maps. These neighborhoods are mentioned below with their coordinates. They are in central and north areas of Tehran.by analyzing these neighborhoods venues with assistance by Foursquare API, and visualizing various factors and features for the client, the client can find out which neighborhood is proper for start the business. Data locations are illustrated below:

	Neighborhood	Latitude	Longitude
0	Bagh Feiz	35.744200	51.322500
1	Abbas Abad	35.735556	51.435556
2	City Park	35.683333	51.416667
3	Amir Abad	35.740744	51.391598
4	Darakeh	34.448333	47.670833
5	Darband	35.823311	51.425028

Creating the Map:



Using Foursquare API:

In [6]:

```
def neighbor(names, latitudes, longitudes, radius=500):

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


        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&r
adius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)


        results = requests.get(url).json()["response"]['groups'][0]['items']


        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)
```

Data Analyzing:

In [7]:
```
venues_Tehran = neighbor(names=df['Neighborhood'],
                                    latitudes=df['Latitude'],
                                    longitudes=df['Longitude'],
                                   )
```

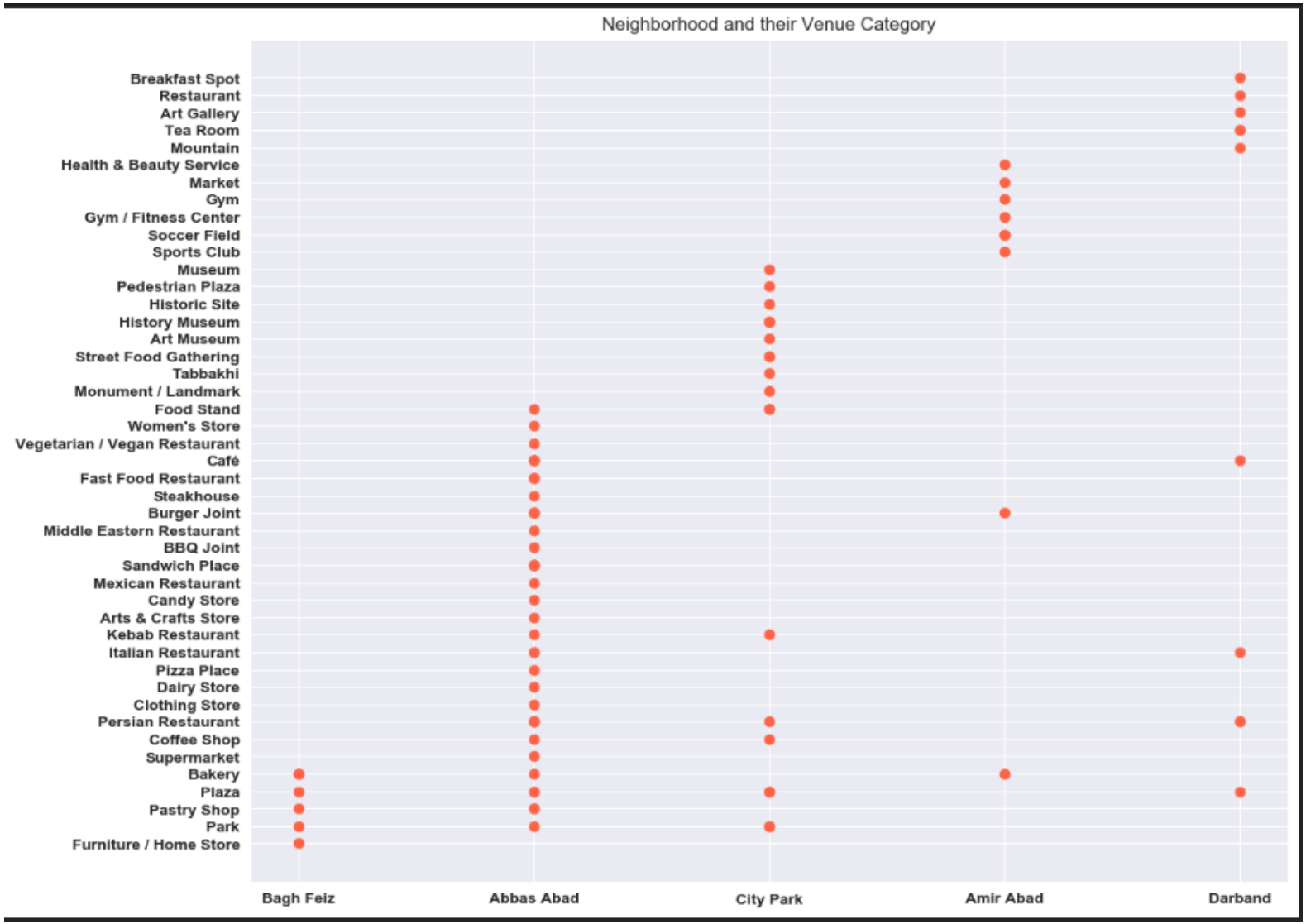
```
Bagh Feiz
Abbas Abad
City Park
Amir Abad
Darakeh
Darband
```

In [8]:
```
venues_Tehran.head()
```

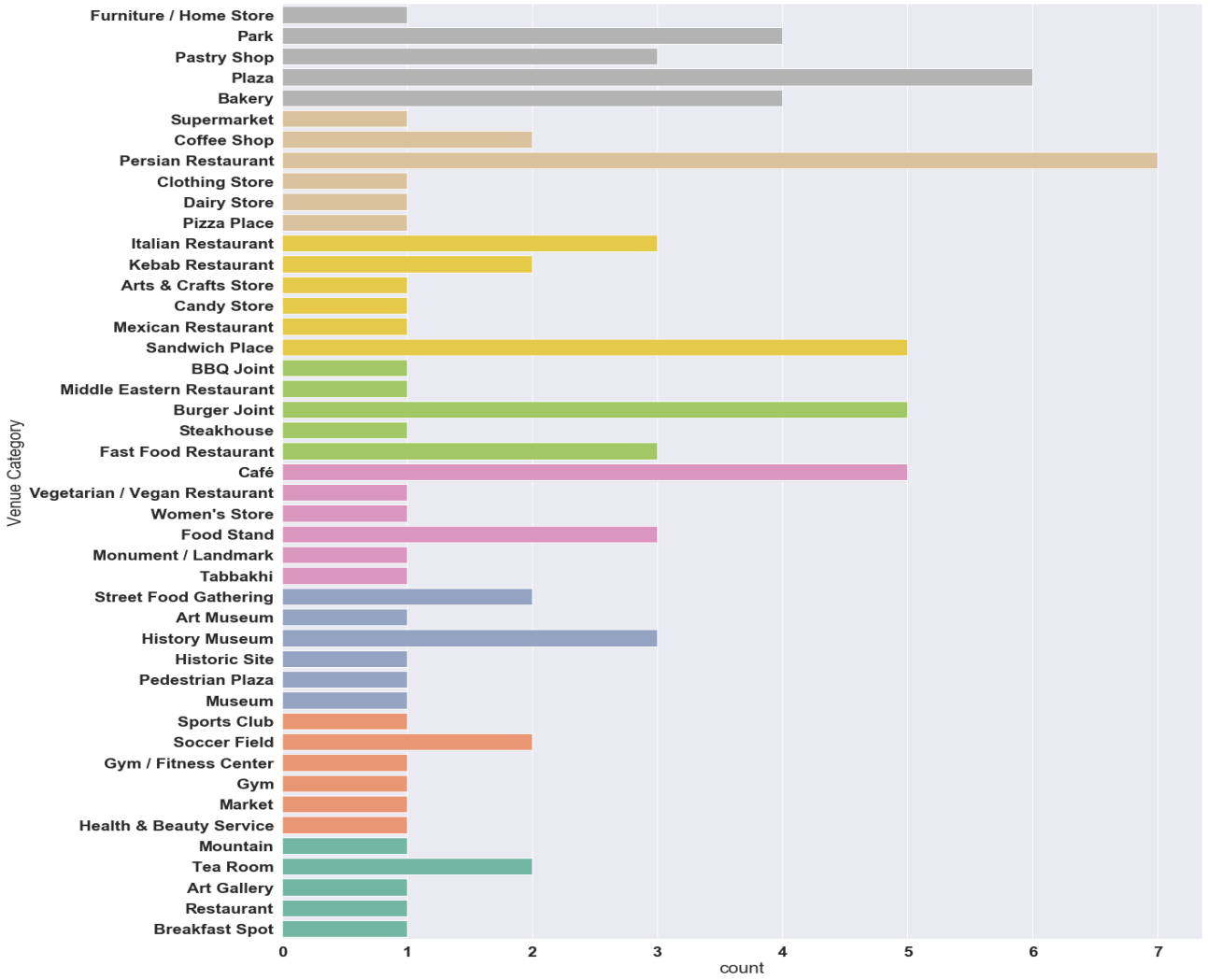
Out[8]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bagh Feiz	35.7442	51.3225	Eden Home Accessories	35.741793	51.322152	Furniture / Home Store
1	Bagh Feiz	35.7442	51.3225	Kaaj Park | بوستان کاج	35.746818	51.321806	Park
2	Bagh Feiz	35.7442	51.3225	West Diamond | شیرینی الماس غرب	35.740984	51.320589	Pastry Shop
3	Bagh Feiz	35.7442	51.3225	Bagh-e Feyz Square | میدان باغ فیض (میدان باغ ...	35.746028	51.325934	Plaza
4	Bagh Feiz	35.7442	51.3225	دانوایی سنگکی باغ فیض	35.746162	51.326198	Bakery

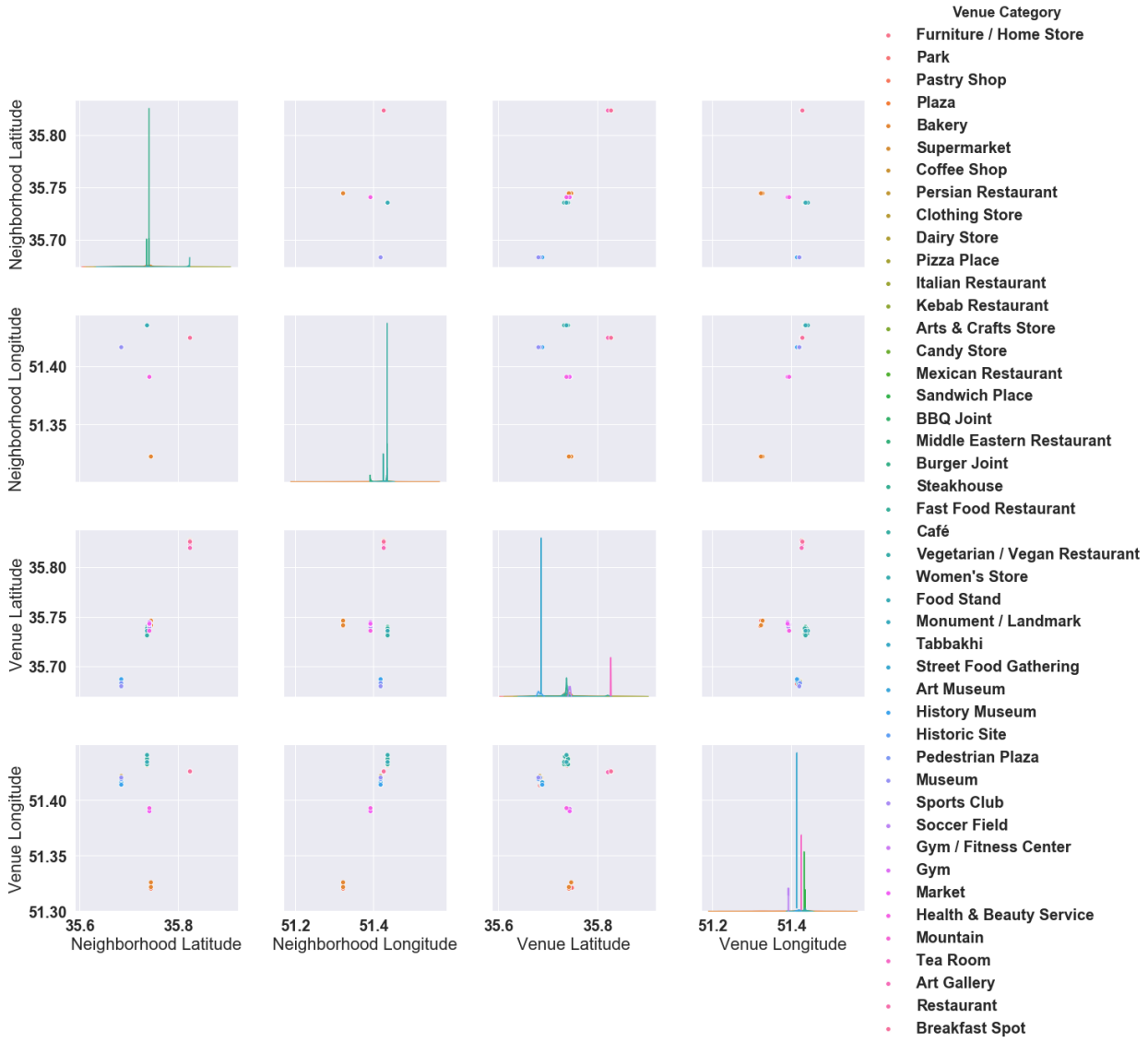
Neighborhood and their Venue Category:



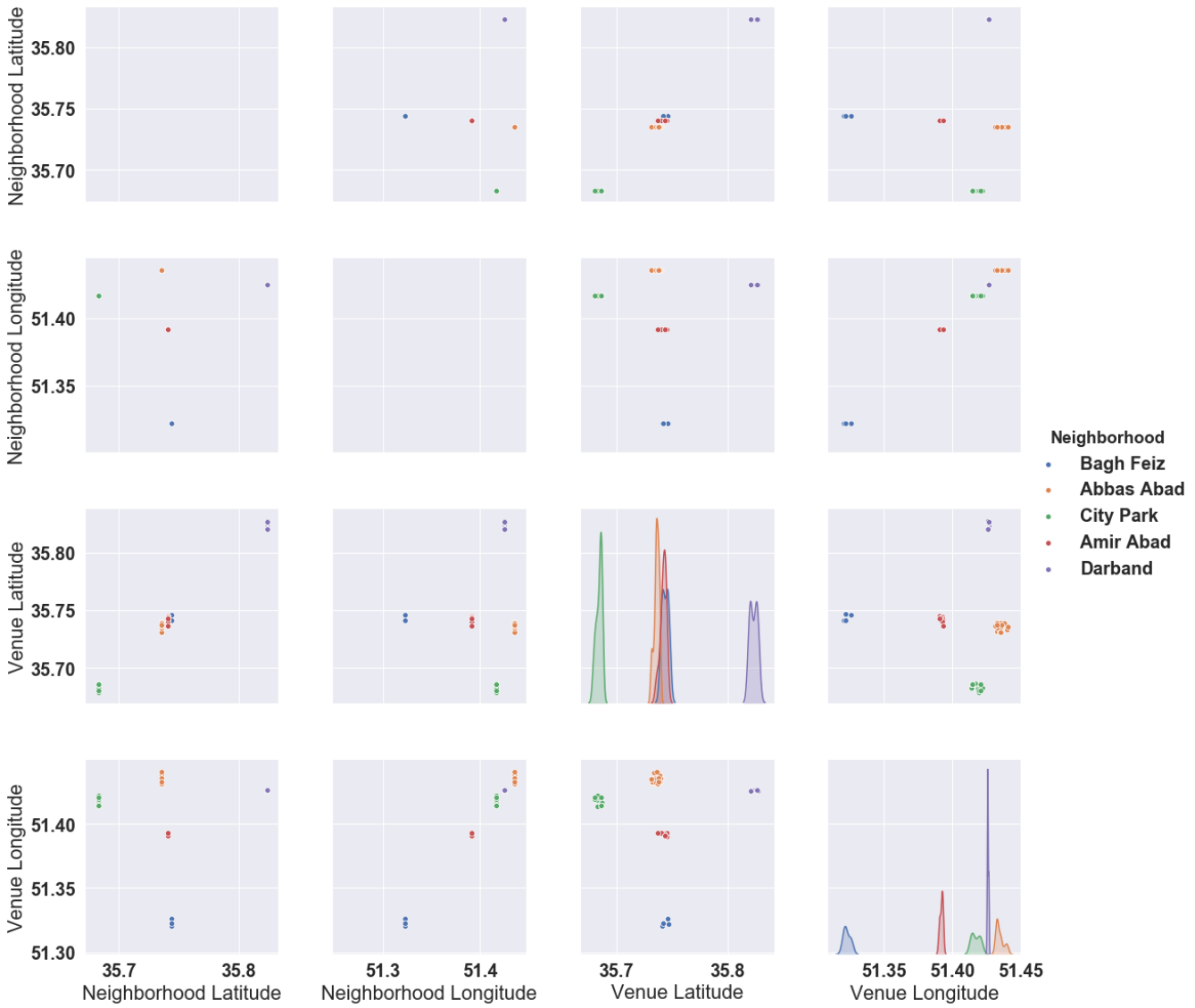
The below figure shows the most venues in these neighborhoods are Plaza and Pizza place:



Features analyze based on Venue Category:



Features analyze based on Venue Neighborhood:



Distinguishing venues values by dummies function:

In [16]:
```
tehran = pd.get_dummies(venues_Tehran[['Venue Category']], prefix="", prefix_sep="")
tehran['Neighborhood'] = venues_Tehran['Neighborhood']
fixed_columns = [tehran.columns[-1]] + list(tehran.columns[:-1])
tehran = tehran[fixed_columns]
tehran
```

Out[16]:

	Neighborhood	Art Gallery	Art Museum	Arts & Crafts Store	BBQ Joint	Bakery	Breakfast Spot	Burger Joint	Café	Candy Store	...	Sandwich Place	Soccer Field	Sports Club
0	Bagh Feiz	0	0	0	0	0	0	0	0	0	...	0	0	0
1	Bagh Feiz	0	0	0	0	0	0	0	0	0	...	0	0	0
2	Bagh Feiz	0	0	0	0	0	0	0	0	0	...	0	0	0
3	Bagh Feiz	0	0	0	0	0	0	0	0	0	...	0	0	0
4	Bagh Feiz	0	0	0	0	1	0	0	0	0	...	0	0	0
5	Bagh Feiz	0	0	0	0	1	0	0	0	0	...	0	0	0
6	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
7	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
8	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
9	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
10	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
11	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
12	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
13	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
14	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
15	Abbas Abad	0	0	0	0	0	0	0	0	0	...	0	0	0
16	Abbas Abad	0	0	1	0	0	0	0	0	0	...	0	0	0

```
tehran['Art Gallery'].value_counts()
```

```
0    88
1     1
Name: Art Gallery, dtype: int64
```

there is 1 Art Gallery in these Neighborhoods

```
tehran['Café'].value_counts()
```

```
0    84
1     5
Name: Café, dtype: int64
```

there is 5 Café in these Neighborhoods

Correlation between venues Types:

	Art Gallery	Art Museum	Arts & Crafts Store	BBQ Joint	Bakery	Breakfast Spot	Burger Joint	Café	Candy Store	Clothing Store	..
Art Gallery	1.000000	-0.250000	-0.250000	-0.250000	-0.369176	1.000000	-0.405603	0.600575	-0.250000	-0.250000	..
Art Museum	-0.250000	1.000000	-0.250000	-0.250000	-0.369176	-0.250000	-0.405603	-0.408203	-0.250000	-0.250000	..
Arts & Crafts Store	-0.250000	-0.250000	1.000000	1.000000	-0.277392	-0.250000	0.518556	0.624035	1.000000	1.000000	..
BBQ Joint	-0.250000	-0.250000	1.000000	1.000000	-0.277392	-0.250000	0.518556	0.624035	1.000000	1.000000	..
Bakery	-0.369176	-0.369176	-0.277392	-0.277392	1.000000	-0.369176	-0.143843	-0.527001	-0.277392	-0.277392	..
Breakfast Spot	1.000000	-0.250000	-0.250000	-0.250000	-0.369176	1.000000	-0.405603	0.600575	-0.250000	-0.250000	..
Burger Joint	-0.405603	-0.405603	0.518556	0.518556	-0.143843	-0.405603	1.000000	0.100888	0.518556	0.518556	..
Café	0.600575	-0.408203	0.624035	0.624035	-0.527001	0.600575	0.100888	1.000000	0.624035	0.624035	..
Candy Store	-0.250000	-0.250000	1.000000	1.000000	-0.277392	-0.250000	0.518556	0.624035	1.000000	1.000000	..
Clothing Store	-0.250000	-0.250000	1.000000	1.000000	-0.277392	-0.250000	0.518556	0.624035	1.000000	1.000000	..
Coffee Shop	-0.369287	0.890979	0.216883	0.216883	-0.502286	-0.369287	-0.165765	-0.118923	0.216883	0.216883	..
Dairy Store	-0.250000	-0.250000	1.000000	1.000000	-0.277392	-0.250000	0.518556	0.624035	1.000000	1.000000	..
Fast Food Restaurant	-0.250000	-0.250000	1.000000	1.000000	-0.277392	-0.250000	0.518556	0.624035	1.000000	1.000000	..
Food Stand	-0.318194	0.972592	-0.018011	-0.018011	-0.447836	-0.318194	-0.294308	-0.271663	-0.018011	-0.018011	..
Furniture / Home Store	-0.250000	-0.250000	-0.250000	-0.250000	0.946395	-0.250000	-0.405603	-0.408203	-0.250000	-0.250000	..
Gym	-0.250000	-0.250000	-0.250000	-0.250000	0.069348	-0.250000	0.698254	-0.408203	-0.250000	-0.250000	..

Below part shows some Venues density in Neighbors:

		Art Gallery
Neighborhood	Art Gallery	
Abbas Abad	0.000000	1
Amir Abad	0.000000	1
Bagh Feiz	0.000000	1
City Park	0.000000	1
Darband	0.090909	1

		Café
Neighborhood	Café	
Abbas Abad	0.093023	1.0
Amir Abad	0.000000	1.0
Bagh Feiz	0.000000	1.0
City Park	0.000000	1.0
Darband	0.090909	1.0

		Bakery
Neighborhood	Bakery	
Abbas Abad	0.023256	1.0
Amir Abad	0.111111	1.0
Bagh Feiz	0.333333	1.0
City Park	0.000000	1.0
Darband	0.000000	1.0

----Abbas Abad----

	venue	freq
0	Sandwich Place	0.12
1	Persian Restaurant	0.09
2	Burger Joint	0.09
3	Café	0.09
4	Fast Food Restaurant	0.07
5	Pastry Shop	0.05
6	Plaza	0.05

----Bagh Feiz----

	venue	freq
0	Bakery	0.33
1	Park	0.17
2	Pastry Shop	0.17
3	Furniture / Home Store	0.17
4	Plaza	0.17
5	Art Gallery	0.00
6	Restaurant	0.00

----Amir Abad----

	venue	freq
0	Soccer Field	0.22
1	Market	0.11
2	Gym	0.11
3	Bakery	0.11
4	Burger Joint	0.11
5	Sports Club	0.11
6	Health & Beauty Service	0.11

----City Park----

	venue	freq
0	History Museum	0.15
1	Food Stand	0.10
2	Plaza	0.10
3	Park	0.10
4	Street Food Gathering	0.10
5	Museum	0.05
6	Art Museum	0.05

----Darband----

	venue	freq
0	Tea Room	0.18
1	Persian Restaurant	0.18
2	Art Gallery	0.09
3	Plaza	0.09
4	Mountain	0.09
5	Breakfast Spot	0.09
6	Italian Restaurant	0.09

Most Common Venue:

	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
0	Abbas Abad	Sandwich Place	Café	Burger Joint	Persian Restaurant	Fast Food Restaurant	Pastry Shop	Italian Restaurant
1	Amir Abad	Soccer Field	Market	Health & Beauty Service	Gym / Fitness Center	Bakery	Gym	Sports Club
2	Bagh Feiz	Bakery	Furniture / Home Store	Park	Pastry Shop	Plaza	Women's Store	Dairy Store
3	City Park	History Museum	Park	Food Stand	Street Food Gathering	Plaza	Coffee Shop	Monument / Landmark
4	Darband	Tea Room	Persian Restaurant	Art Gallery	Café	Mountain	Plaza	Restaurant