Launching Italian Restaurant in Pune

1. Introduction

Every project starts with business understanding. I will provide analytic solution by defining the problem, project objectives and solution requirements from a business perspective.

In this project I will provide information about Pune City in India and then after I will provide information about specific areas [Kothrud and Hinjewadi] in Pune where someone can open an **Italian restaurant**.

I have chosen these two areas where Italian restaurant can be opened. Analysis will be provided about these areas with visualization graphs so that decision can be easy for the person opening restaurant in these areas.

I have lived in many cities over the years. However, PUNE holds a special place in my heart. No other place has given the level of thrill, fun, memories, holidays & opportunities as Pune. The city emerged drastically over the span of last 10 years and was awarded as the 2nd best city to live in India.

The top reasons why Pune is the best city to live in India.

Booming IT Sector Everyone knows that India produces the largest number of Software Experts in the world & IT Industry has become the backbone of the economy. With over 110 MNCs, Pune is home to some of the biggest and stunning IT parks of India. There are more than 5 official IT zones in Pune with Rajiv Gandhi Infotech Park in Hinjewadi being the biggest of all. This IT park is spread over an area of 13 KMs with the presence of all major

companies like TCS, Infosys, Wipro, Persistent, Cognizant, Accenture, etc. Such large landscape of companies gives great exposure to Job seekers in the Silicon Valley of India.

Educational Hub This is the sole thing Pune was famous for in the old days. Pune is known as the educational hub of India since years. There are world famous Universities and Colleges for each academic zone.

Pune University has an affiliation with 57 Engineering colleges which itself is a record. There are top management schools like Symbiosis. Pune has earned the reputation of "Oxford of the East." Being a top spot for Educational Institutes, it has also become a favorite place for Summits, Global Events, Concerts & Fests.

Food & Shopping These things are core elements of one's decision to open Restaurant in a city. Well, luckily you can get both at very reasonable price and with lots of variety in Pune. There are plenty of mega shopping malls in like Pheonix Market City, Amanora Town, Seasons Mall, Inorbit & Pune Central. As the city have plenty of young crowds, one can find many street shopping places on MG Road, Laxmi Road, FC Road, etc.

When it comes to Food, one can get all the variety of different lip-smacking cuisine in the city. From traditional street food to 7 courses dine at a five-star restaurant, you will get every option to try.
Everyone knows that you get to eat tasty Vadapav in Maharashtra. However, Puneri Misal beats it in all possible manner. There are more than 5000 restaurants in the city. You can get food even at 3 AM if you're planning to pull an all-nighter.

Comfortable Lifestyle Although Pune is next big city in Maharashtra after Mumbai, it is still pretty much cheaper when it comes to lifestyle**. Property rates are too much lower in comparison to Mumbai, and this really helps the budget for opening Restaurant in a city.**

2. Data Description

The analytic methods to be used require certain data content, formats and representations, guided by domain knowledge. I used **Foursquare API** to get the most common venues of Areas of Pune. Specifically, I have chosen Kothrud and Hinjewadi.

- In the initial data collection stage, As a data scientists I identified and gather the available data resources—structured, unstructured and semi-structured—relevant to the problem domain.

.

- After the original data collection, I typically used descriptive statistics and visualization techniques to understand the data content, assess data quality and discover initial insights about the data.
- For a given technique, I tried K-mean algorithm with their respective parameters to find the best model for the available variables.

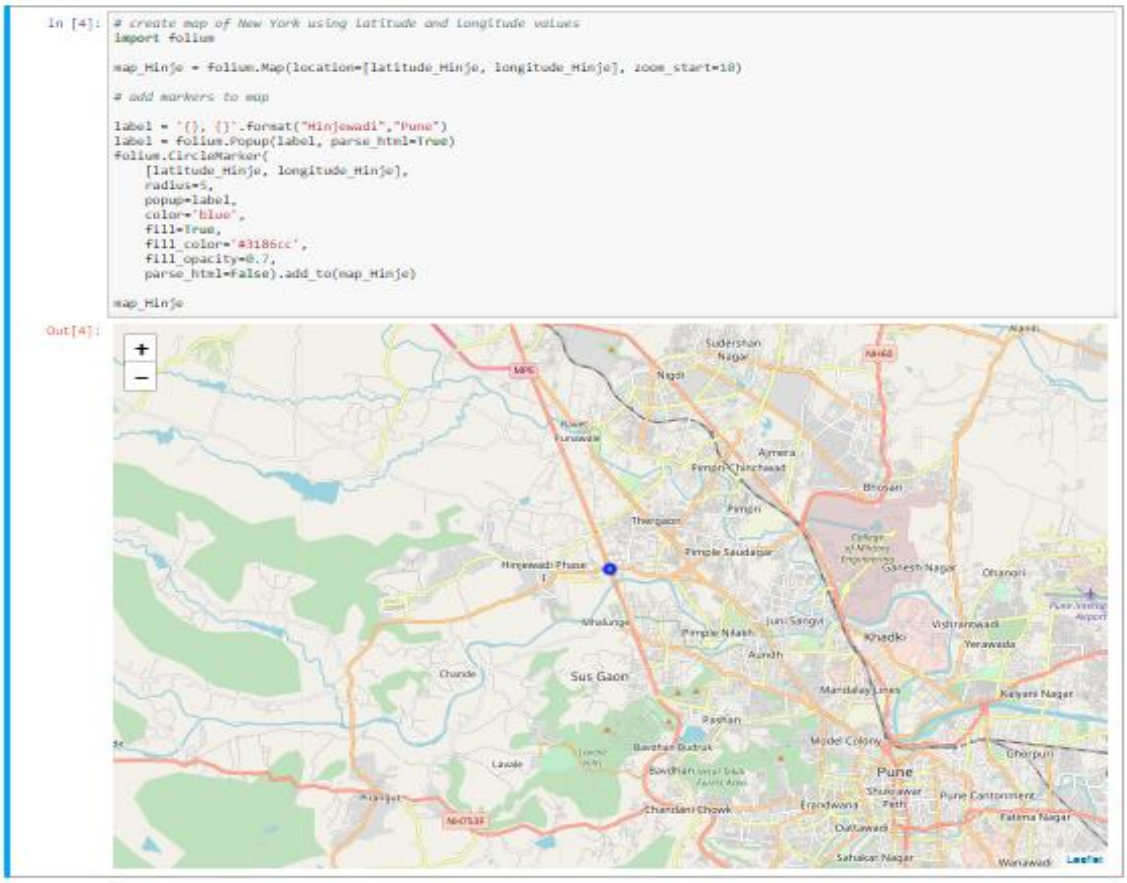
3. Methodology

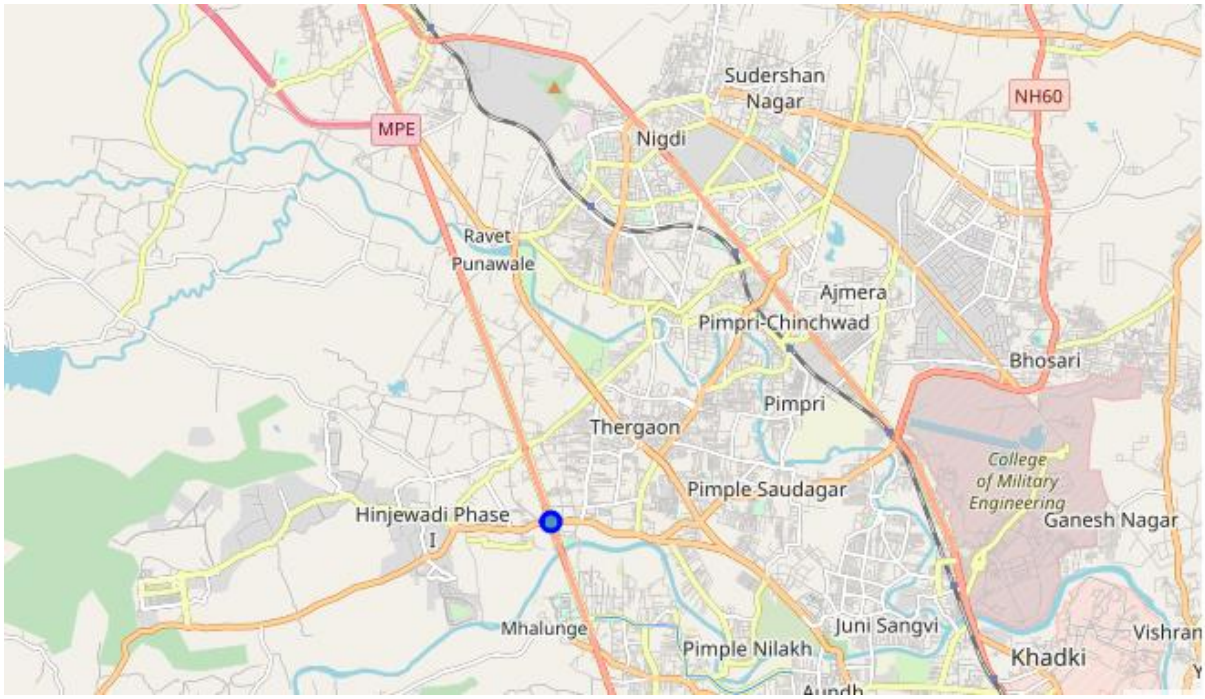
I used geopy library to get Logitude and latitude information of Hinjewadi location

```
: from geopy.geocoders import Nominatim
geolocator = Nominatim(user_agent="MyHinjewadi1")
location = geolocator.geocode("Hinjewadi,Pune")


latitude_Hinje = location.latitude
longitude_Hinje = location.longitude
print(latitude_Hinje)
print(longitude_Hinje)
```

```
18.5920937
73.7571423
```

I used python **folium** library to visualize geographic details of Hinjewadi Pune and used latitude and longitude values to get the visual as shown below.





I used foursquare API to get URL where I set Limit as 100 sothat 100 venues can be obtained and radius set as 2KM.

```
In [6]: LIMIT = 100 # limit of number of venues returned by Foursquare API
        radius = 2000 # define radius
        url_Hinje = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            latitude_Hinje,
            longitude_Hinje,
            radius,
            LIMIT)
        url_Hinje # display URL

Out[6]: 'https://api.foursquare.com/v2/venues/explore?&client_id=UUGKAHRK21E3FEBHEY1JBIDRXKTAA55M0USCTKV0R51MPFT0&client_secret=NDDHC4T
        WBDJIETCKMZAMV3PJ3DYWLDBOAMFZTY55BFTHJLO3&v=20180605&ll=18.5920937,73.7571423&radius=2000&limit=100'
```

Details of venues are obtained.

	name	categories	lat	lng
0	Natural Ice Cream	Ice Cream Shop	18.591192	73.752440
1	Sayaji	Hotel	18.599535	73.754995
2	Barbeque Nation	BBQ Joint	18.599390	73.755090
3	Courtyard by Marriott	Hotel	18.591527	73.746831
4	MoMo Cafe	Indian Restaurant	18.591650	73.747011
5	Little Italy	Italian Restaurant	18.591513	73.743668
6	121 Kitchen : Bar	Indian Restaurant	18.593894	73.763607
7	Portico - Sayaji	Indian Restaurant	18.599268	73.754935
8	The Gateway HOTEL	Hotel	18.591914	73.744877
9	Mainland China	Chinese Restaurant	18.591142	73.747642
10	D Mart	Department Store	18.591395	73.744324
11	Fortune Inn Exotica	Hotel	18.590988	73.748120
12	Ranjaai Garden	Indian Restaurant	18.593470	73.756332
13	Croma	Electronics Store	18.590243	73.770105
14	Saundarya Restaurant	Indian Restaurant	18.591475	73.751652
15	E Square Xion Hinjewadi	Multiplex	18.592393	73.744884
16	Cafe Coffee Day	Café	18.590903	73.747214
17	Chhatrapati Shivaji Sports Complex	Stadium	18.575559	73.761005
18	Cafe Coffee Day	Café	18.604069	73.753098
19	Sarovar	Diner	18.582646	73.761056

I identified Indian as well as Italian restaurant in the Hinjewadi Area.

In [21]:
```
df=nearby_venues_1[(nearby_venues_1['categories']=="Indian Restaurant")]
df
```

Out[21]:

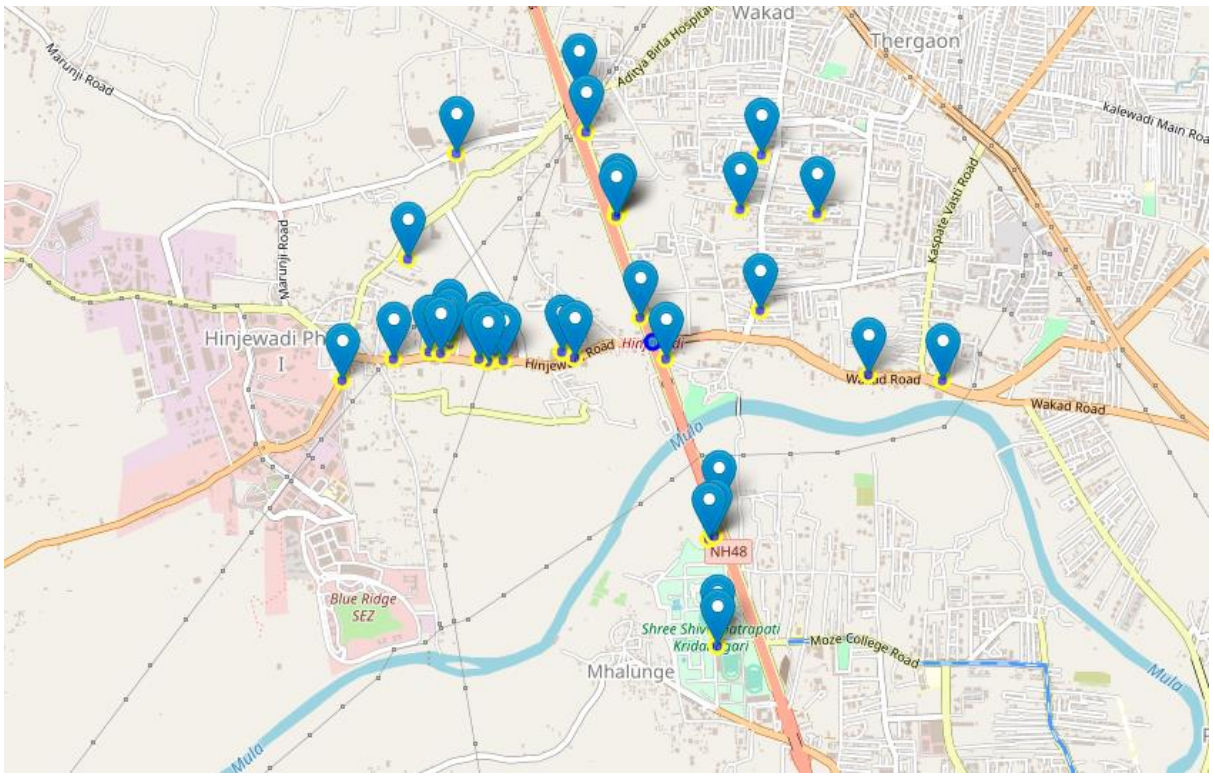
	name	categories	lat	lng
4	MoMo Cafe	Indian Restaurant	18.591650	73.747011
6	121 Kitchen : Bar	Indian Restaurant	18.593894	73.763607
7	Portico - Sayaji	Indian Restaurant	18.599268	73.754935
12	Ranjaai Garden	Indian Restaurant	18.593470	73.756332
14	Saundarya Restaurant	Indian Restaurant	18.591475	73.751652
23	Hotel Jalsa	Indian Restaurant	18.598985	73.763831
24	Hotel Shivratna	Indian Restaurant	18.581724	73.761343
25	ANCIENT HYDERBAD	Indian Restaurant	18.592443	73.744965
28	Vitthal Kamat	Indian Restaurant	18.590882	73.744416
31	Silver Spoon	Indian Restaurant	18.602798	73.745306

In [22]:
```
df1=nearby_venues_1[(nearby_venues_1['categories']=="Italian Restaurant")]
df1
```

Out[22]:

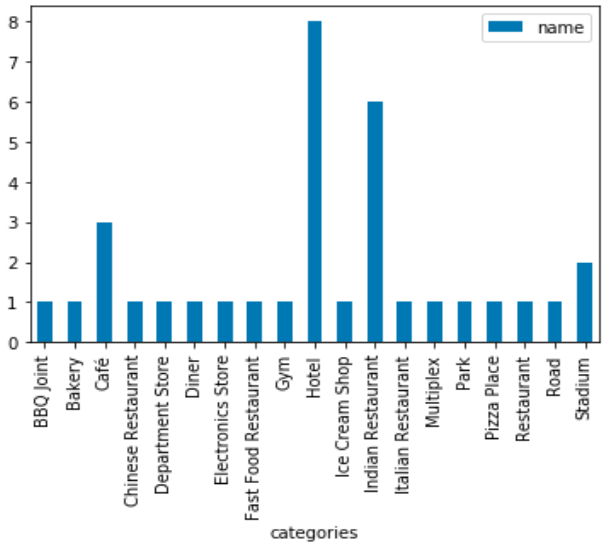
	name	categories	lat	lng
5	Little Italy	Italian Restaurant	18.591513	73.743668

It can be observed that there is only one Italian restaurant in the Hinjewadi Area named Little Italy.



All venues have been shown in bar chart as well.

```
In [19]: customer_group_1 = customers_1.groupby('categories').count()
         my_plot_1 = customer_group_1.plot(kind='bar')

         plt.show(my_plot_1)
```



Similarly, venues are identified for other location Kothrud in Pune.

In [32]:
```
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe

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

nearby_venues_Kothrud = json_normalize(venues) # flatten JSON
nearby_venues_Kothrud

# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.lat', 'venue.location.lng']
nearby_venues_2 =nearby_venues_Kothrud.loc[:, filtered_columns]

# filter the category for each row
nearby_venues_2['venue.categories'] = nearby_venues_2.apply(get_category_type, axis=1)

# clean columns
nearby_venues_2.columns = [col.split(".")[-1] for col in nearby_venues_2.columns]

nearby_venues_2
```

Out[32]:

	name	categories	lat	lng
0	Hidden Place - The Hangout	Pub	18.509107	73.812280
1	The Katta	Café	18.495362	73.812876
2	Cafe Coffee Day	Café	18.500140	73.814254
3	CCD	Café	18.507830	73.808498
4	Sudit Su's Chinese	Chinese Restaurant	18.495267	73.812488
5	Yashwantrao Chavan Natyagruha	Theater	18.502278	73.812180
6	Barometer	Café	18.498824	73.819240

It is observed that only one Italian Restaurant is present in the Kothrud Area.

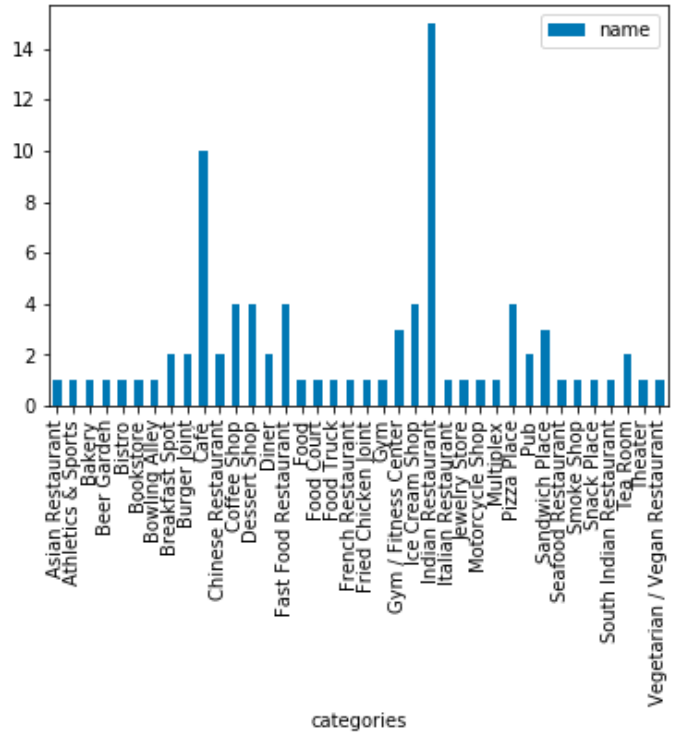
In [33]:
```
df2=nearby_venues_2[(nearby_venues_2['categories']=="Italian Restaurant")]
df2
```

Out[33]:

	name	categories	lat	lng
32	Feesta	Italian Restaurant	18.508699	73.812501

The venues are shown in the Bar Chart

```
In [35]: customer_group_2 = customers_2.groupby('categories').count()
         my_plot_2 = customer_group_2.plot(kind='bar')

         plt.show(my_plot_2)
```



So it can seen from analysis that Hinjewadi and Kothrud is having only one exclusive Italian Restaurant. Therefore it is great chance that some one can open Italian restaurant in these areas.

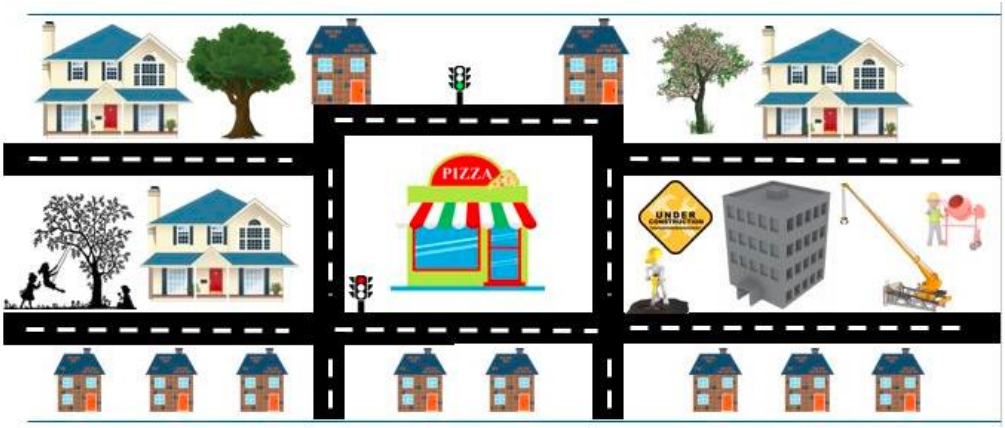
Now if I compare cost of these two areas then Hinjwadi seems to be cheaper than kothrud area. Hinjewadi costs around 5000/- 10000/-per sqft whereas Kothrud costs 10,000/ -20000/ per sqft.

If we compare the customer segments who regularly eat Italian food, then Hinjewadi would be more suitable place because a lot of young IT employees resides in Hinjewadi as compared to kothrud.

Hinjewadi is located towards the west of Pune and is 21 kilometres from the Railway station. Over a period of time, it has developed into one of the biggest industrial areas of Pune, transforming from a sleepy village to a bustling neighbourhood. Hinjewadi is also home to historical attractions such as Shaniwar Wada, Vishrambaug Wada and Aga Khan Palace. The pride of Hinjewadi is its state of the art 500-acre Rajiv Gandhi Infotech Park which hosts a cluster of IT/ITES and BPO units. It houses some of India's largest IT conglomerates like Wipro, Tech Mahindra, TCS, Infosys, IBM and Cognizant. Hinjewadi is being developed in multiple phases. Phase 1, 2 and 3 are completed and further development has been planned from phase 4 onwards.

Hinjewadi offers easy accessibility to Mumbai via the six-lane Pune- Expressway, and also to Tathawade (which houses several educational institutions) through the Mumbai-Bangalore bypass. To cater to the residential demand of the migrating IT population, Hinjewadi is witnessing several proposed and ongoing development of residential projects, serviced apartments and five

star hotels. With excellent means of road transport, Hinjewadi connects to Pune which is about 25 minutes to an hour's drive while the International Airport is about a three-hour drive away. Mumbai- Highway also connects Hinjewadi to the Airport which is about a 45-minute drive away.This location is highly attractive due to its proximity to the industrial belt of Pimpri and Chinchwad. This has also facilitated residential development of nearby areas like Wakad, Balewadi, Aundh and Baner. Further a lot of building construction is going on in Hinjewadi so chance of business is more as compared to Kothrud.



So looking at above analysis it can be finalised that restaurant can be open in Hinjewadi area

4. Challenges with the Project

Now next question would be where to open Italian shop in Hinjewadi. We need to take help for external marketing agency to do survey. major survey points are given below

1. How many customers are interested in Italian food in Hinjewadi Area.
2. What are their frequency of eating Italian food. [As Italian food is not a regular Indian food]
3. Customer's addresses need to be captured so that Longitude and Latitude information of Customer location can be obtained.

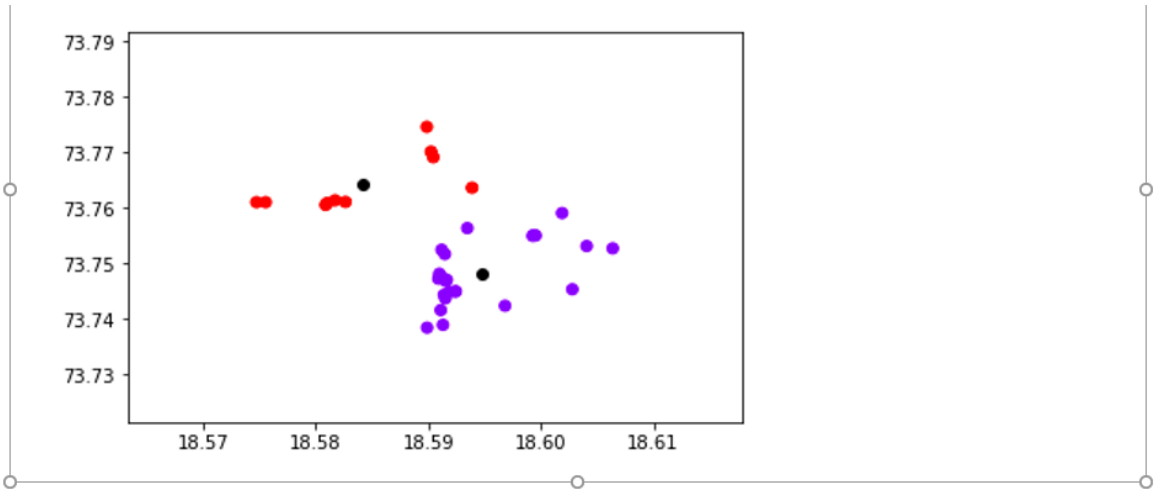
Lets say we could be able to get customer information of about 1000 customers who have shown interest in eating Italian food. now we can use K-mean Machine learning algorithm for finding the optimum location of Hotel so that all the customers can be served better way.

for simplicity I have taken Longitudes and latitudes of a few customers and use K-mean algorithm for getting clusters and centroid of the clusters would be Hotel Location.

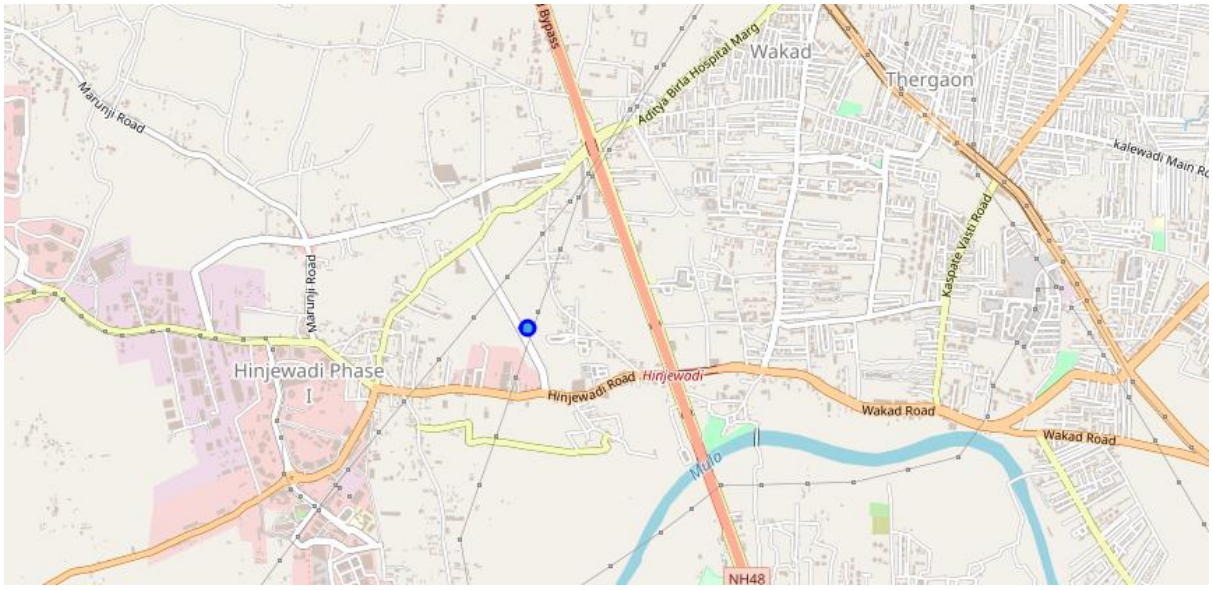
I got below result after applying K-Mean Algorithm

In [1]:
```
import matplotlib.pyplot as plt
import numpy as np
from sklearn.cluster import KMeans

X = np.array([[18.59119169262836, 73.75243955724582],
 [18.599535406189991, 73.7549952061101],
 [18.599389591869766, 73.75509038587762],
 [18.59152702926924, 73.74683097930983],
 [18.591649917117625, 73.74701112681768],
 [18.59151274236633, 73.74366780400008],
 [18.593893849415267, 73.76360660453777],
 [18.599267775472452, 73.75493532260735],
 [18.591914007855575, 73.7448765876003],
 [18.591142286156536, 73.74764190469449],
 [18.591394831323328, 73.7443236838029],
 [18.590888146970982, 73.7481201103472],
 [18.593470456931968, 73.75633183462378],
 [18.590242592662218, 73.77010487479674],
 [18.591474511375576, 73.75165224916849],
 [18.575558598527564, 73.76100540161133],
 [18.59239341022894, 73.74488449919531],
 [18.590902588918713, 73.74721396209324],
 [18.604069498485195, 73.75309781963126],
 [18.582646266961657, 73.76105613429112],
 [18.58088471470406, 73.76051975858331],
 [18.6063844887836, 73.75268411653326],
 [18.601887421437326, 73.7590447074031],
 [18.581723654525206, 73.76134297426894],
 [18.590448977020426, 73.76911092110393],
 [18.592443, 73.744965],
 [18.581002979550966, 73.76082952282827],
 [18.589897518554732, 73.7383957161389],
 [18.574745003725422, 73.76100540161133],
 [18.59681999373572, 73.74234298596807],
 [18.591111045087583, 73.74155403421064],
 [18.602797864403428, 73.74530626816272],
 [18.58987257713276, 73.77458047791653],
 [18.591308753644668, 73.73889927204515]])

kmeans = KMeans(n_clusters=2)
kmeans.fit(X)


print(kmeans.cluster_centers_)

print(kmeans.labels_)

plt.scatter(X[:,0],X[:,1], c=kmeans.labels_, cmap='rainbow')
plt.scatter(kmeans.cluster_centers_[:,0] ,kmeans.cluster_centers_[:,1], color='black')
plt.show()
```

```
[[18.59476975 73.74817938]
 [18.58410192 73.76431621]]
```



In [20]:
```
import folium
latitude_Hinje=18.59476975
longitude_Hinje=73.74817938
map_Hinje = folium.Map(location=[latitude_Hinje, longitude_Hinje], zoom_start=10)

# add markers to map

label = '{}, {}'.format("Hinjewadi","Pune")
label = folium.Popup(label, parse_html=True)
folium.CircleMarker(
    [latitude_Hinje, longitude_Hinje],
    radius=5,
    popup=label,
    color='blue',
    fill=True,
    fill_color='#3186cc',
    fill_opacity=0.7,
    parse_html=False).add_to(map_Hinje)

map_Hinje
```



Conclusion:

We obtained two centroid locations from K-Mean Algorithm. I put these values and tried to get the location of Hotel with latitude =18.59476975 longitude=73.74817938 and observed that first centroid location would be better initially, as Cost of opening Hotel is less as compared to other Centroid location [18.58410192 73.76431621].
This Hotel will be easily accessible by High profile customers from Blue Ridge, Godrej,Mega Polis and Vilas Jawdekar society. So, there are great chances of success of opening Italian Hotel at mentioned location.

Suhas Shirbavikar