Capstone Final Project

Finding the best city for starting Multiplex in Andhra Pradesh.

TADISETTI RAJU

JULY 2019

1. Introduction

1.1 Background

A multiplex is a movie theater complex with multiple screens within a single complex. They are usually housed in a specially designed building. Sometimes, an existing venue undergoes a renovation where the existing auditoriums are split into smaller ones, or more auditoriums are added in an extension or expansion of the building. The largest of these complexes can sit thousands of people and are sometimes referred to as a megaplex. With the urbanization the count of multiplexes also increasing and there is huge competition among themselves for attracting more audience to their multiplexes. Therefore, new multiplex organizers looking for the best suitable conditions to start their multiplex.

1.2 Problem

The problem is the customer wants to start new multiplex in Andhra Pradesh a state in India, which consist of some major and minor cities. He wants best city where he can get maximum audience and customers with less competition. The project aims to find best city for starting a multiplex to the customer. Targeted audience are businessmen those who want to construct a mall, multiplex and real estate people.

1.3 Interest

The project aims to find the best location for starting a new multiplex for the constructor or organizer. The targeted audience are businessmen those who want to construct a mall, multiplex. Like real estate people or even film producers or distributors etc.

2. Data

2.1 Data source

For solving problem, I need data related to cities in Andhra Pradesh and Foursquare data.
The cities in Andhra Pradesh, I got that data from Wikipedia by using this link https://en.wikipedia.org/wiki/List_of_cities_in_Andhra_Pradesh_by_population
Along with cities I need latitudes and longitudes of those cities

The foursquare venues data I need:

a) Multiplex data
b) School data
c) Universities data
d) Offices data

By using that data, I can find the best city for starting a mall to customer.
Wikipedia data consists of city names, its population.

Malls venues data consists of name of the mall and its location parameters School data consists of school names College data consists of college names

I need geo locations of venue in city where it is located, I fetched that data using geolocator.

2.2 Data cleaning:

I scraped Andhra Pradesh cities data from Wikipedia it contains city names district names and its population, I need only city names so, I need to remove the extra details like district name and its population. I stored the data into one data frame then I map the geo locations of cities like latitude and longitude. Data presents in the below figure:

Out[22]:

	cities	Latitude	Longitude
0	Visakhapatnam	17.723128	83.301284
1	Vijayawada	16.506759	80.618510
2	Guntur	16.291519	80.454159
3	Nellore	14.449372	79.987376
4	Kurnool	15.830925	78.042537
5	Rajahmundry	17.005045	81.780473
6	Kadapa	14.467149	78.822888
7	Kakinada	16.943738	82.235061
8	Tirupati	13.631637	79.423171
9	Eluru	16.710426	81.115382
10	Anantapur	14.550000	77.416667
11	Vizianagaram	18.115498	83.397881
12	Nandyal	15.473629	78.480659
13	Ongole	15.505871	80.049945
14	Adoni	15.625331	77.273089
15	Madanapalle	13.555798	78.500872
16	Machilipatnam	16.181939	81.135130
17	Tenali	16.237773	80.646422
18	Proddatur	14.755595	78.550153
19	Chittoor	13.303741	78.987701
20	Hindupur	13.826383	77.493772
21	Bhimavaram	16.542769	81.527344
22	Srikakulam	18.294940	83.893875
23	Guntakal	15.161098	77.376893
24	Dharmavaram, Anantapur	14.413840	77.720843
25	Gudivada	16.432998	80.993715

Fig 2.1 Cities data with latitudes and longitudes

Now I need multiplex, schools, colleges and offices data for each and every city, I get that data by using Foursquare API then I visualize those venues on maps using latitudes and longitudes of those venues.

3. Exploratory Data Analysis

I collected the multiplex data for each city within radius of 10km by using Foursquare API credentials. Total 67 multiplex are there in cities of Andhra Pradesh with in 10km radius.

Out[153]:

	cities	city Latitude	city Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Visakhapatnam	17.723128	83.301284	Inox Chitralaya Mall	17.712356	83.301176	Multiplex
1	Visakhapatnam	17.723128	83.301284	Sangam Sarat	17.725549	83.326355	Multiplex
2	Visakhapatnam	17.723128	83.301284	Inox	17.734076	83.318304	Multiplex
3	Visakhapatnam	17.723128	83.301284	screen 4 varun inox	17.709473	83.312653	Multiplex
4	Visakhapatnam	17.723128	83.301284	Inox	17.711077	83.315785	Multiplex

Fig 3.1 Multiplex data in Andhra Pradesh

Once I got the data I visualize that data in maps using latitudes and longitudes, the visualized maps are useful for identification of venues.

Out[127]:

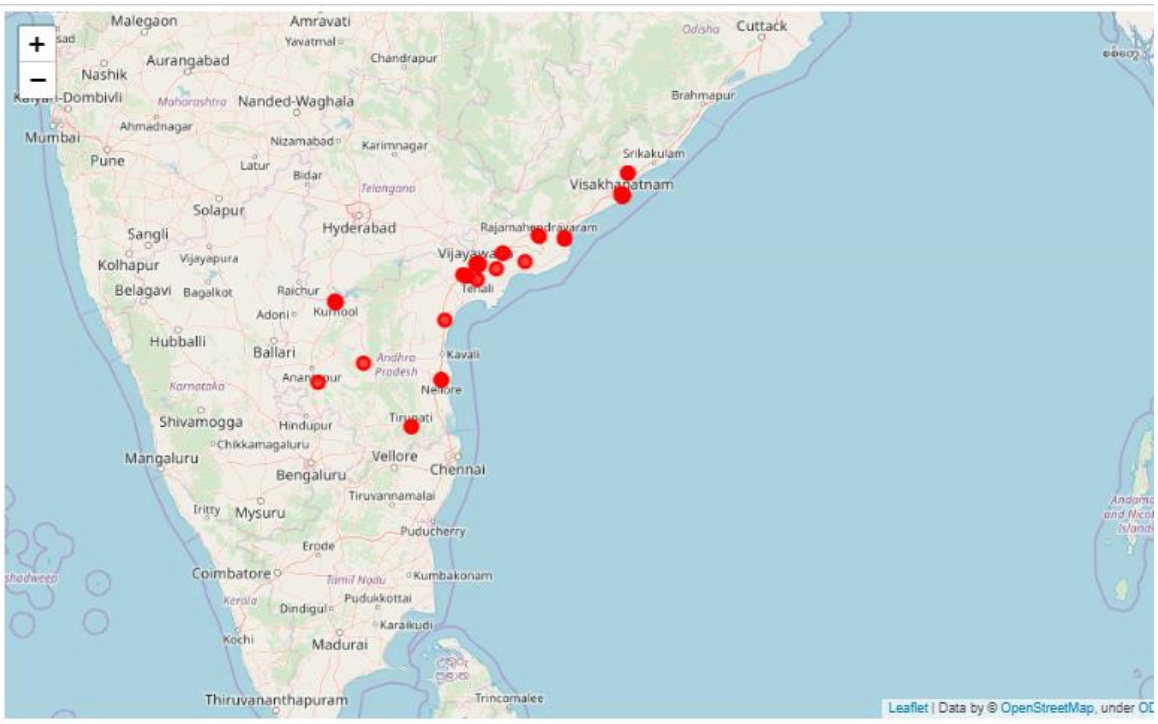

Fig 3.2 Multiplex in Andhra Pradesh

Once multiplex data completed now, I collected schools data in those cities by using foursquare API. Total 62 high schools are there in Andhra Pradesh

Out[128]:

	cities	city Latitude	city Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Visakhapatnam	17.723128	83.301284	Pollocks School	17.726652	83.305002	High School
1	Visakhapatnam	17.723128	83.301284	Silver Oaks International School	17.785889	83.378420	High School
2	Visakhapatnam	17.723128	83.301284	Mother Theresa English Medium School	17.765557	83.225806	High School
3	Visakhapatnam	17.723128	83.301284	Ramanath Secondary School	17.747763	83.234585	High School
4	Visakhapatnam	17.723128	83.301284	All saints convent	17.720388	83.309760	High School

Fig 3.3 Schools data in Andhra Pradesh

Schools locations visualizations is given below.

Out[130]:

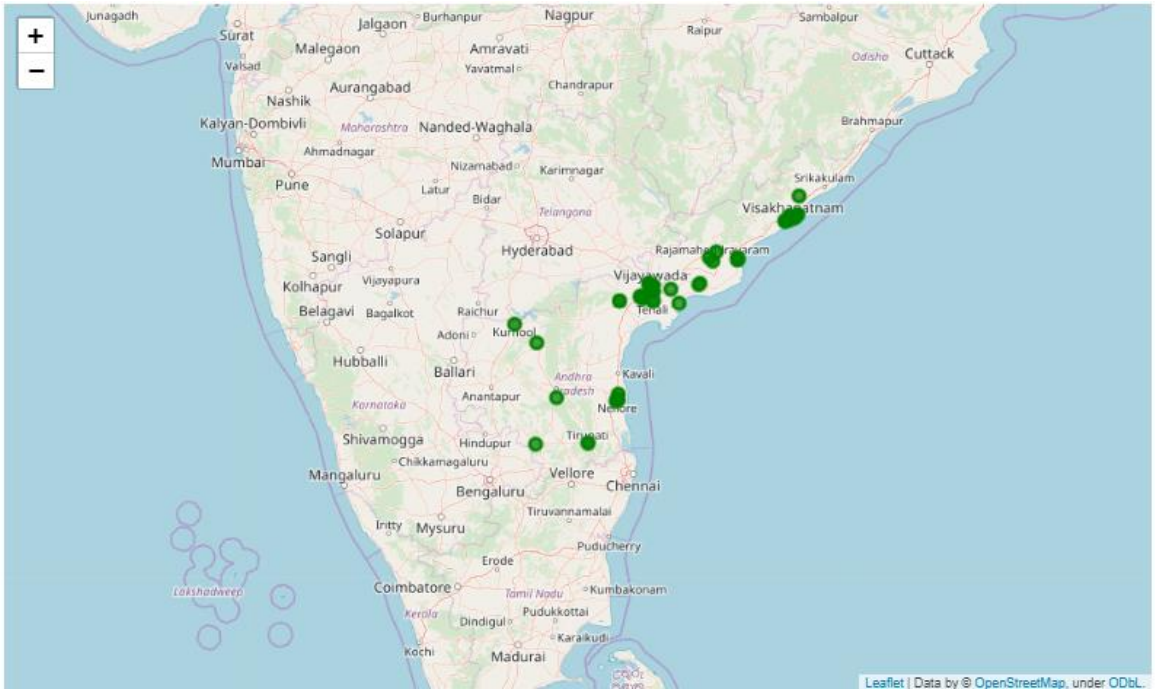

Fig 3.4 schools map in Andhra

Now I fetch universities data universities data using foursquare API. Total 40 universities and colleges in Andhra Pradesh.

Out[131]:

	cities	city Latitude	city Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Visakhapatnam	17.723128	83.301284	gitam	17.736346	83.309425	University
1	Visakhapatnam	17.723128	83.301284	GITAM University	17.781064	83.376782	University
2	Visakhapatnam	17.723128	83.301284	Dept. of Education	17.721780	83.320506	University
3	Visakhapatnam	17.723128	83.301284	Andhra University Delta Studies	17.737383	83.327916	University
4	Visakhapatnam	17.723128	83.301284	Andhra University Civil Department	17.687468	83.212322	University

Fig 3.5 Universities data in Andhra Pradesh

And visualizations of universities data

Out[149]:

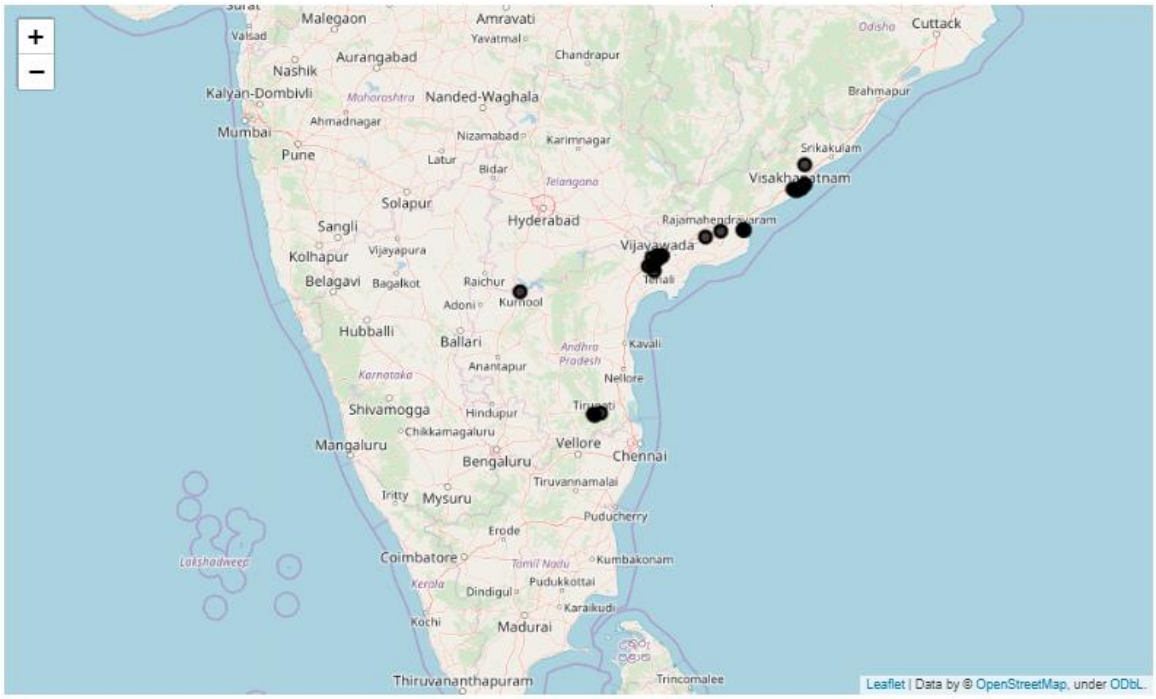

Fig 3.6 Universities in Andhra Pradesh

There are 715 offices in Andhra Pradesh and the data is

Out[134]:

	cities	city Latitude	city Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Visakhapatnam	17.723128	83.301284	Divisional Railway Manager Office	17.724348	83.299380	Government Building
1	Visakhapatnam	17.723128	83.301284	Dwarakanagar	17.725860	83.306840	Parking
2	Visakhapatnam	17.723128	83.301284	Reliance Web World	17.726236	83.303430	Tech Startup
3	Visakhapatnam	17.723128	83.301284	Narasimha swamy temple	17.766856	83.249161	Temple
4	Visakhapatnam	17.723128	83.301284	Indus Hospitals	17.711569	83.302108	Hospital

Fig 3.7 Offices Data in Andhra Pradesh

Visualizing that data in map

)ut[151]:

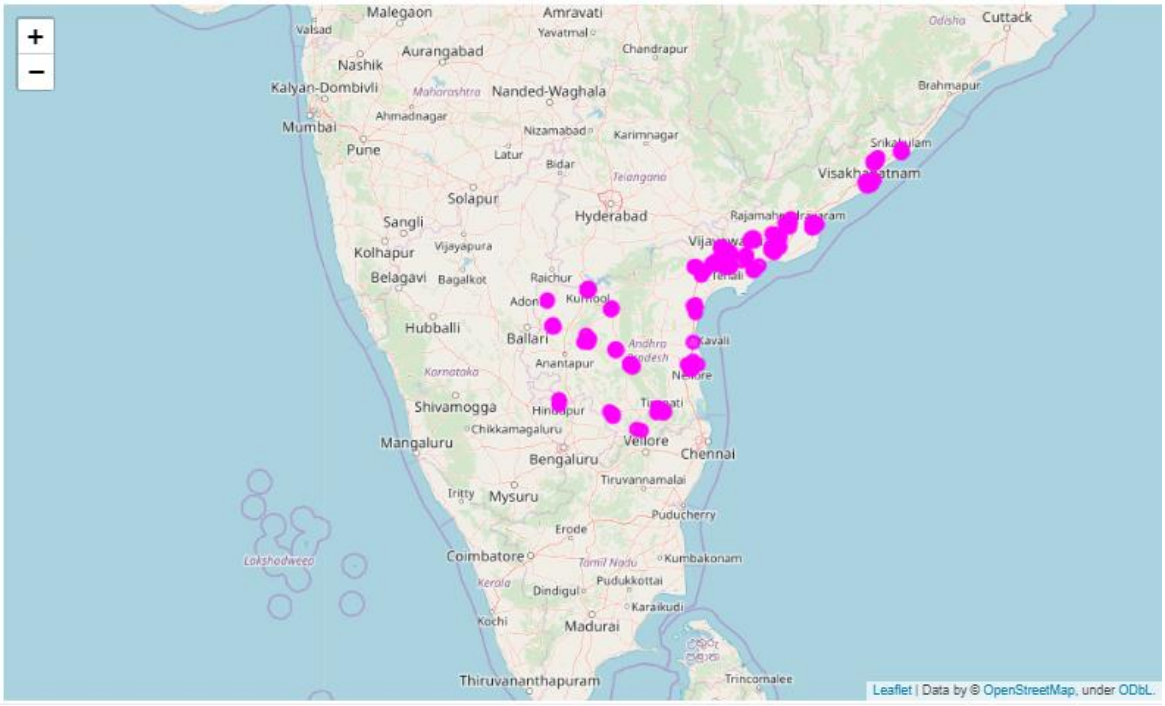

Fig 3.8 offices in Andhra Pradesh

Now I combined the whole venues data the data frame is

Out[138]:

	cities	Latitude	Longitude	malls	high schools	universities	offices
0	Visakhapatnam	17.723128	83.301284	11.0	14.0	12.0	49.0
1	Vijayawada	16.508759	80.618510	23.0	6.0	9.0	45.0
2	Guntur	16.291519	80.454159	6.0	6.0	3.0	46.0
3	Nellore	14.449372	79.987376	3.0	8.0	0.0	48.0
4	Kurnool	15.830925	78.042537	5.0	1.0	1.0	29.0
5	Rajahmundry	17.005045	81.780473	2.0	3.0	1.0	42.0
6	Kadapa	14.467149	78.822888	0.0	1.0	0.0	20.0
7	Kakinada	16.943738	82.235061	4.0	3.0	4.0	45.0
8	Tirupati	13.631637	79.423171	2.0	2.0	4.0	50.0
9	Eluru	16.710426	81.115382	2.0	0.0	0.0	44.0
10	Anantapur	14.550000	77.416667	0.0	0.0	0.0	0.0
11	Vizianagaram	18.115498	83.397881	3.0	1.0	1.0	48.0
12	Nandyal	15.473629	78.480659	0.0	1.0	0.0	8.0
13	Ongole	15.505871	80.049945	1.0	0.0	0.0	24.0
14	Adoni	15.625331	77.273089	0.0	0.0	0.0	3.0
15	Madanapalle	13.555798	78.500872	0.0	1.0	0.0	8.0
16	Machilipatnam	16.181939	81.135130	0.0	1.0	0.0	11.0
17	Tenali	16.237773	80.646422	1.0	2.0	1.0	20.0
18	Proddatur	14.755595	78.550153	1.0	0.0	0.0	9.0
19	Chittoor	13.303741	78.987701	0.0	0.0	0.0	2.0
20	Hindupur	13.826383	77.493772	0.0	0.0	0.0	11.0
21	Bhimavaram	16.542769	81.527344	1.0	2.0	0.0	47.0
22	Srikakulam	18.294940	83.893875	0.0	0.0	0.0	15.0
23	Guntakal	15.161098	77.376893	0.0	0.0	0.0	5.0
24	Dharmavaram, Anantapur	14.413840	77.720843	1.0	0.0	0.0	0.0

Fig 3.9 venues data of Andhra Pradesh

Visualizing the count of venues for some cities in bar chart

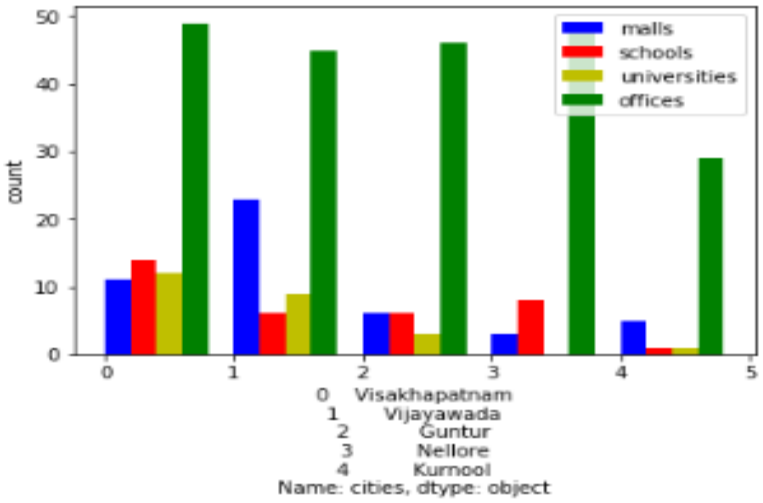

Fig 3.10 count of venues in some cities

4. Methodology

For each city, all multiplex, school, university and office venues have been collected form foursquare, then for each city the sum of multiplex, school, offices and universities were computed. For each of these 4 categories, a weight was defined to find best city.

- Multiplex have been weighted with -4, because multiplex will reduce the profit
- Schools have been weighted with 1, because school students are good customers.
- Universities have been weighted with 2, because students will easily attract to movies.
- Offices have been weighted with 3, because employees are independent and good customers.

Once weights assigning completed, score for each city computed and score is assigned to each city based on that score best city for multiplex finalized.

The city, which get maximum score is the best city for starting a new multiplex.

5. Results

After assigning weights and calculated the score for each city Tirupati got the maximum score among the remaining cities. The result data is in below table.

Out[143]:

	cities	score
8	Tirupati	152.0
30	Amaravati (state capital)	148.0
0	Visakhapatnam	141.0
3	Nellore	140.0
21	Bhimavaram	139.0
11	Vizianagaram	135.0
7	Kakinada	130.0
2	Guntur	126.0
9	Eluru	124.0
5	Rajahmundry	123.0
4	Kurnool	70.0
13	Ongole	68.0
1	Vijayawada	67.0
6	Kadapa	61.0
17	Tenali	60.0
22	Srikakulam	45.0
16	Machilipatnam	34.0
20	Hindupur	33.0
25	Gudivada	33.0
28	Tadepalligudem	29.0
15	Madanapalle	25.0
12	Nandyal	25.0
27	Tadipatri	24.0
18	Proddatur	23.0
26	Narasaraopet	22.0
23	Guntakal	15.0
29	Chilakaluripet	13.0
14	Adoni	9.0
19	Chittoor	6.0
31	Kavali	3.0
10	Anantapur	0.0
24	Dharmavaram, Anantapur	-4.0

Fig 5.1 Score for each city

Tirupati got maximum score then we conclude that Tirupati is suitable for starting new multiplex.

Venues in Tirupati

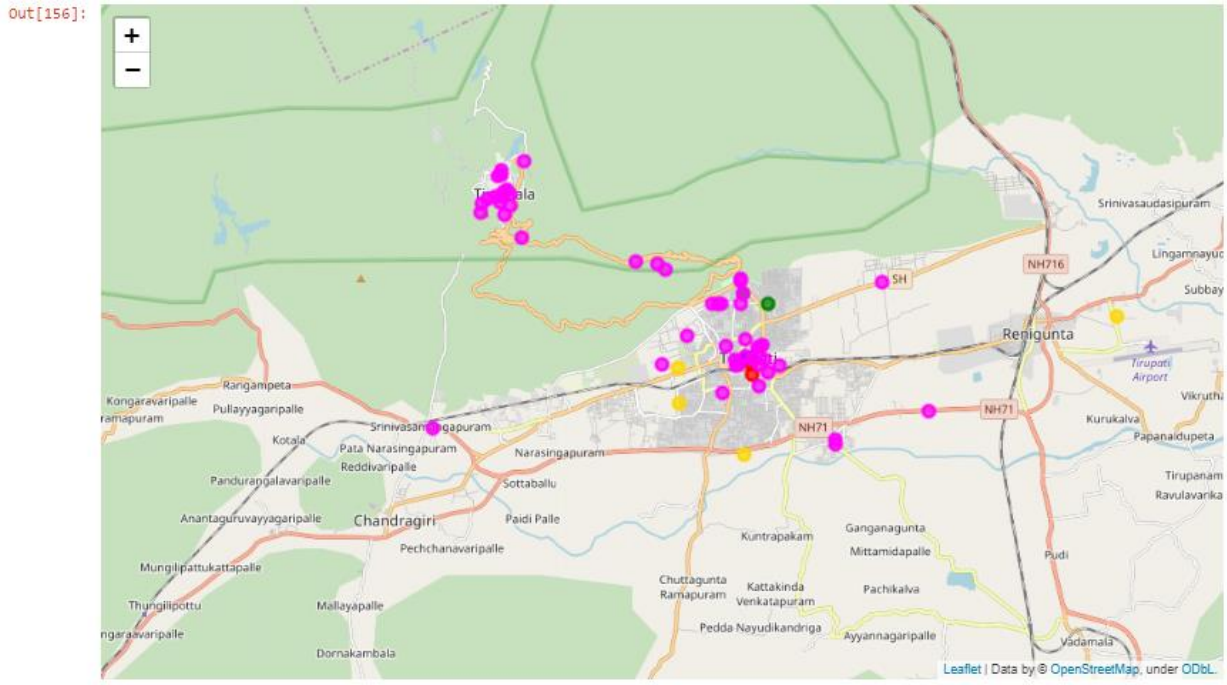

Fig 5.2 Venues in Tirupati

6. Conclusion

From the result of the above work, I conclude that Tirupati city has more benefits for starting new multiplex because in Tirupati there are less number of multiplexes and more schools, universities and offices, means new multiplex will have less competition and more audience.

Future Work:

We can also improve the result accuracy by considering life style of people in each city and their interest in movies and multiplex culture

We can also apply the clustering techniques like k-means clustering for grouping the venues by we can get best accuracy for finding the best locality for new multiplex.