Chennai based Real Estate Investment Indicator using IT Park as Indices

Abstract

In metro cities people show interest in investing in lands for short term and long term profits. To increase the profit people should be aware of selecting the developed and developing regions. One factor that increases the land value is the urban development and in Indian cities like Chennai Mumbai, Banglore IT industries play a vital role in it. So the Project focuses on identifying the locations of IT companies and provides cluster based index to invest near IT industries. And the project focuses on the Chennai IT industries based on the construction status like completed, started, under construction and planning.

Introduction

Data analytics plays a major role in business and industries to reduce the investment and to increase the benefits. Data Science is playing a vital role in analytics to uncover the solutions in business. The machine learning algorithms automated the analytical process as the data grows. The big data repositories are analysed to bring solution to existing problems also data science has brought in new business environments for the online platforms. The individuals even can use the analytical reports published online. This way people are analysing the trends in stock market, weather conditions, people habits that enables a profit for business. AS data science provides cool solutions in various aspects people started depending on analytics unknowing from rating a zomato order to buying a top branded shoes. Take this as an opportunity, we explore the tech park location Chennai, Tamilnadu, India, to enable buyers to invest in proper location to multiple their investments in few years

Problem Description

The Chennai city is also one among the silicon cities, that accommodate above 4000 it companies and above 4 million it sector employees [Wikipedia]. Manu of the employees are from other states and cities. The next largest business is real estate both rental and selling. The regions near the IT parks are of highly sought after for the residential purpose to minimize the transportation to work place. Large investors can easily explore in the developed regions. The small investor who want to invest in the right place for residential or commercial lands to multiple their investment cannot afford places in developed regions, they can invest in the developing areas where the IT parks are planned or under construction in the city.

This project is to support those investors big or small to make a decision on where to buy a land in Chennai for huge rental income or high resale value. So the analytics produce results region wise density of the IT industry and status wise whether it is completed or under construction or stated for future investments.

2. A description of the data and how it will be used to solve the problem. (**15 marks)**

The required data for the project are taken from foursquare.com. The analytics need to explore two datasets. The first one is the IT industries location data in the city and the status of the company which includes categories like

1. Completed
2. Started
3. Planning
4. Under construction

The data is available in https://en.wikipedia.org/wiki/List_of_tech_parks_in_Chennai.

The another data set is the geographical location data of the city to plot the clusters in the city map. The dataset is available in https://en.wikipedia.org/wiki/List_of_neighbourhoods_of_Chennai.

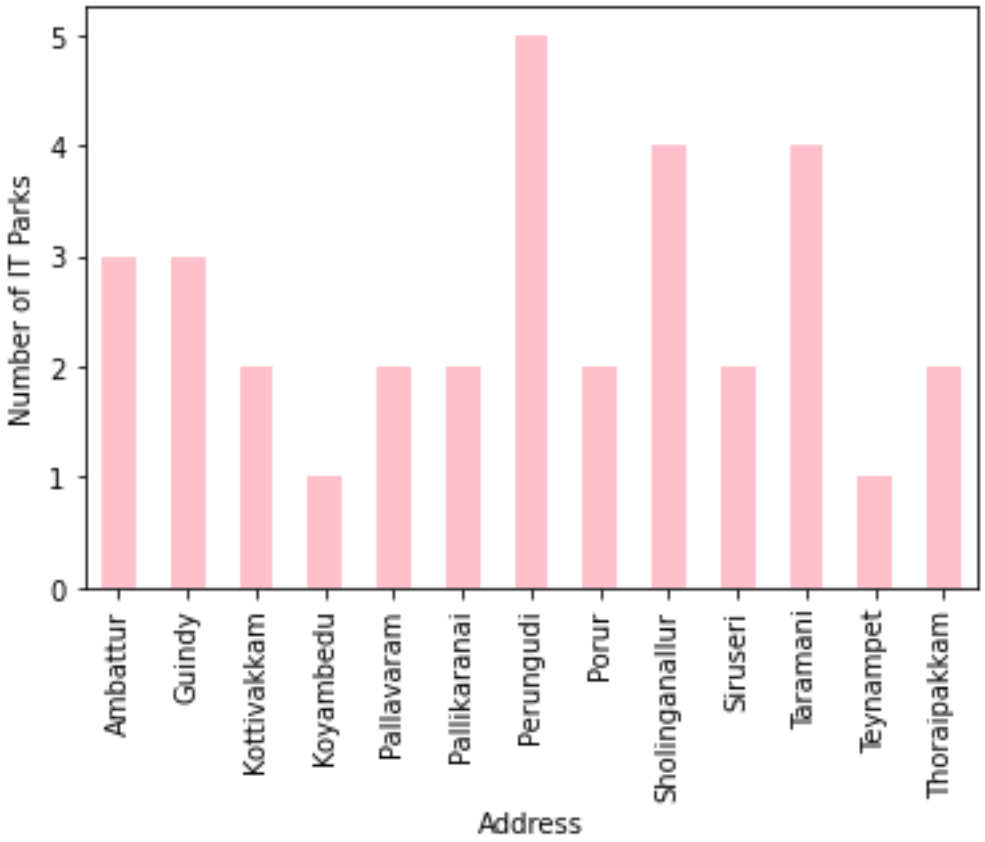
As the project focuses on the location of the IT industry and not on the IT companies working in the tech park those details are cleaned during the data cleaning and preparation step. In the same way the location also takes the postal address and latitude and longitude to get the geographical location.

Later the Status are converted to numeric data as status code to cluster it based on the status of the techno park. And two clusters are formed based on location and status.

Methodology

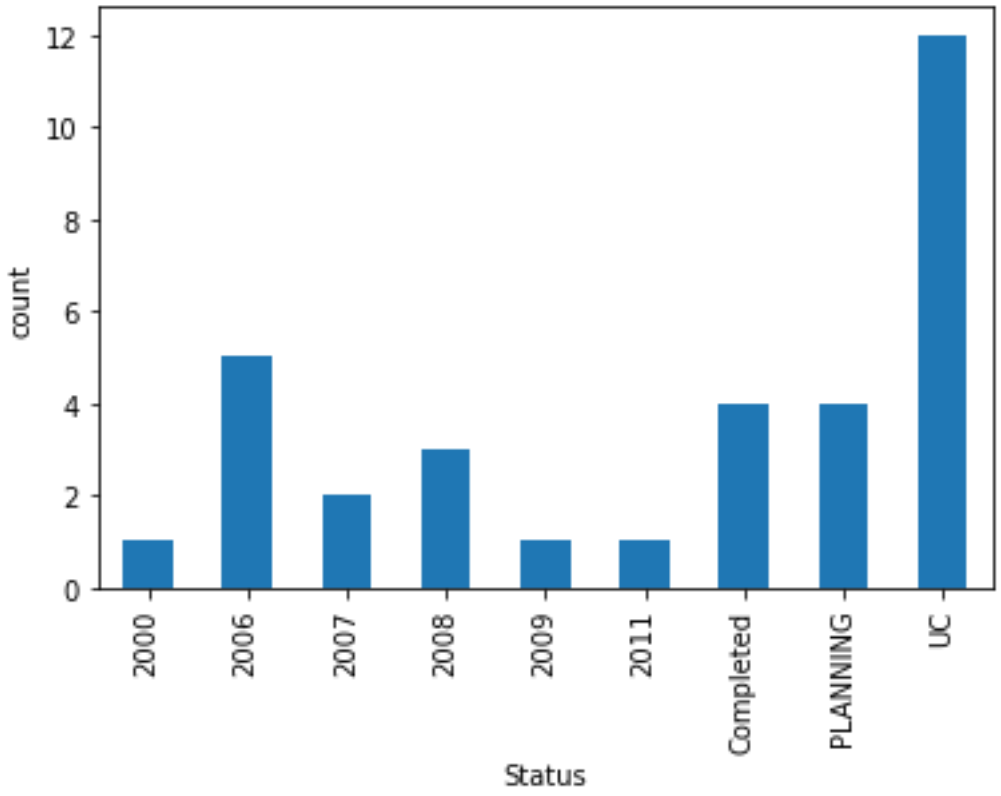
The data combines the it park data in Chennai city and the postal codes and categorize the data based on location and status. The data set is as follows:

The following graph categorize the data based on number of it companies based on location.



The following grapsh categorize the data based on the current status of the company.

The following graphs plots the location and status based maps.



Conclusion

The above maps indicates the best locations to buy for investors that yields better profit.