**Battle of Neighbourhood: Toronto VS Vizag**

**Searching the possibilities of replicating the real estate investment model of Toronto City in Vizag by comparing the neighbourhoods of both cities.**





**Contents**

1. **Introduction**
2. **Data collection and preprocessing**
3. **Methodology**
4. **Results**
5. **Discussion and Conclusion**

**1.1 Description and Discussion of the Background**

***Are the neighbourhood areas in Vizag and Toronto any similar? Do they share any common characteristics?***

I have decided to explore the neighborhoods of Vizag, AP, India with the neighbourhoods of Toronto, Canada understand the investment opportunities and the city overall growth and development at par with Toronto using Clustering & Segmentation techniques, ML (Machine Learning). Data Visualizations (using seaborn and matplotlib in Python) are created to explore GDP, Per Capita Income, climatic conditions, tourism and educational institutions of both the cities. All these data points will help us understand the rate of growth in Visakhapatnam and scope of development in different sectors.

The real estate investment patterns of Toronto can be observed by analysing the data of Neighbourhood of Toronto city. Then for exploring the possibility of replicating those investment model in real estate market of Vizag, we can compare the neighbourhood data of both cities.

**1.2.Problem**

Searching the possibilities of replicating the real estate investment model of Toronto City in Vizag by comparing the neighbourhoods of both cities.

**1.3.Interest**

This project will highlight the investor opportunities with increased scope of attracting real estate investors in investing in Vizag which can help Vizag to realize its ambitious economic growth goals while preserving and enhancing livability for the benefit of local citizens.

**2. Data**

**2.1 Data Requirements**

Following datasets have been used in the project:

* Postal Codes of Visakhapatnam. Data has been scraped and cleaned from Yo!Vizag — City’s Exclusive Magazine and Portal [1] using Beautiful Soup and pandas libraries and saved in .csv format.
* Foursquare API to get the most common venues of given boroughs of Visakhapatnam and Toronto respectively.[2]
* Visakhapatnam [3] and Toronto Wikipedia Pages [4] have been scraped and cleaned for creating Word clouds.
* Zip codes of Toronto. Data has been downloaded in .csv format from <https://datasf.org/>and cleaned using pandas.

**2.2.Data Analysis:**

2 Cities will be analyzed in this project: Visakhapatnam and Toronto.

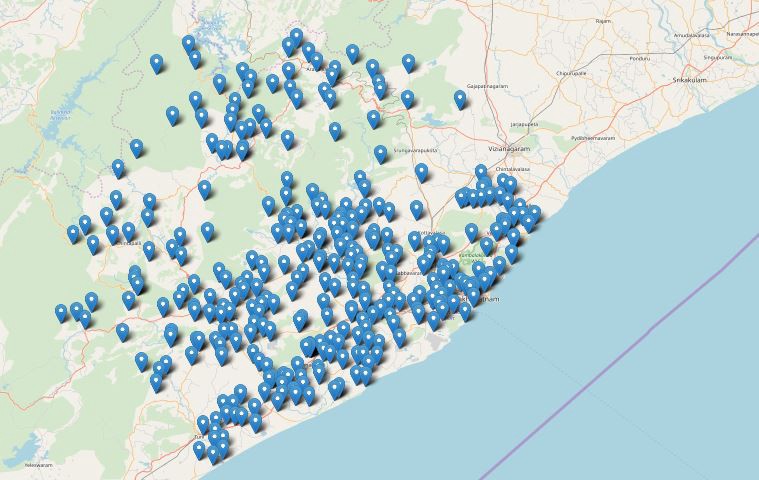
I will be using the below datasets for analyzing Visakhapatnam.

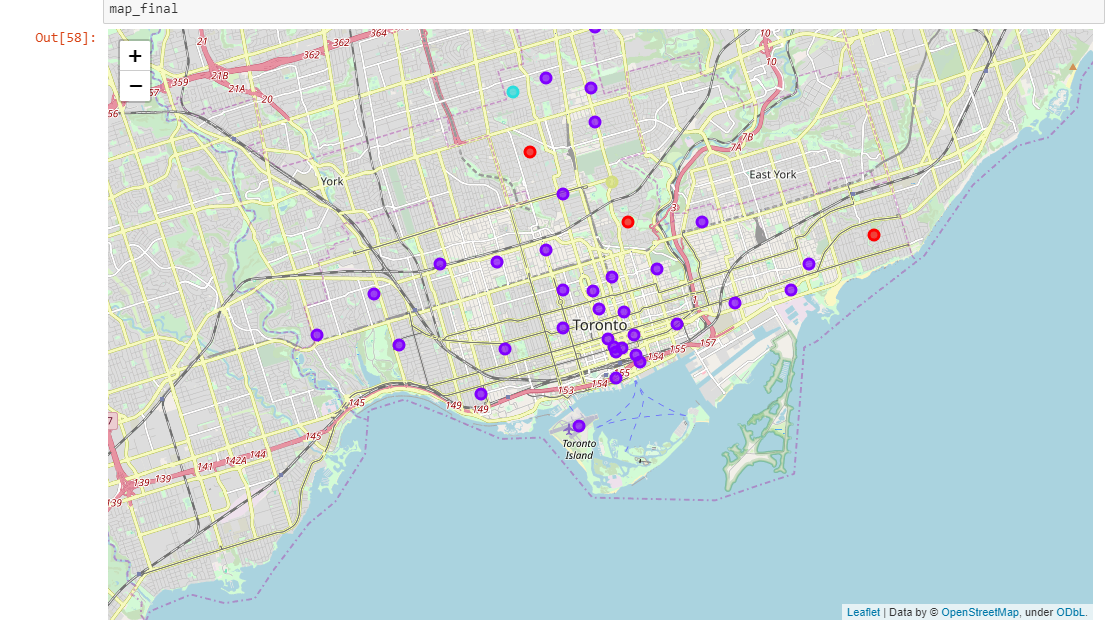
**Data 1:**Neighborhood has a total of 684 areas. Most notable areas of the city include urban areas like Dwaraka Nagar, Gajuwaka, Gopalapatnam, Jagadamba Centre, Maddilapalem, Madhurawada, Seethammadhara and semi-rural suburbs such as Simhachalam, Pendurthi, and Parwada.

Data has been scraped and cleaned from [Yo!Vizag — City’s Exclusive Magazine and Portal](https://www.yovizag.com/visakhapatnam-vizag-pin-code/)– using Beautiful Soup and pandas libraries and saved in .csv format.







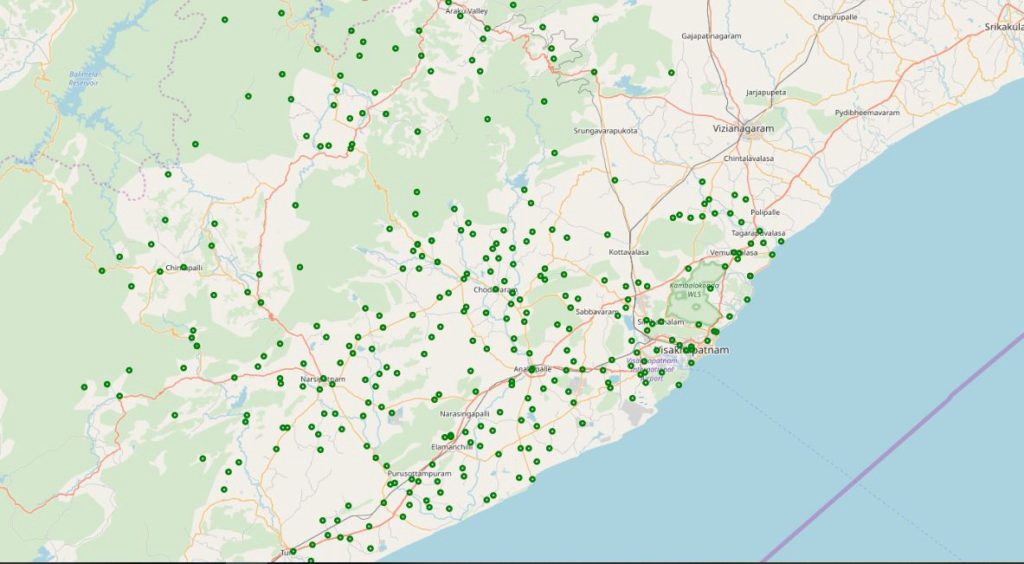


**3. Methodology: Exploratory Data Analysis:**

Data 1: Visakhapatnam Geographical Coordinates Data.

We use geopy and folium libraries to create a map of Visakhapatnam city with neighborhoods imposed on it. 326 areas are plotted using their latitude and longitude values to obtain a high-level visualization of the neighborhoods.

**Fig: Visakhapatnam Neighborhood Visualization**



Now let’s explore venues around Andhra University, one of the most prestigious and oldest university in Andhra Pradesh located in central Vizag. We selected this location as Andhra University is located on the uplands of Visakhapatnam, the university campus is scenic, with the [Bay of Bengal](https://en.wikipedia.org/wiki/Bay_of_Bengal) on one side of it and on the other, the green [Kailasagiri](https://en.wikipedia.org/wiki/Kailasagiri) hill range. This location is apt for our analysis as Toronto was also chosen because of the geographical similarity.

Longitude and Latitude values of Andhra University, Sivajipalem Road, Sector 4, Pedda Waltair, Visakhapatnam, Andhra Pradesh, 530001, India are 17.7376312,83.3300513027767.

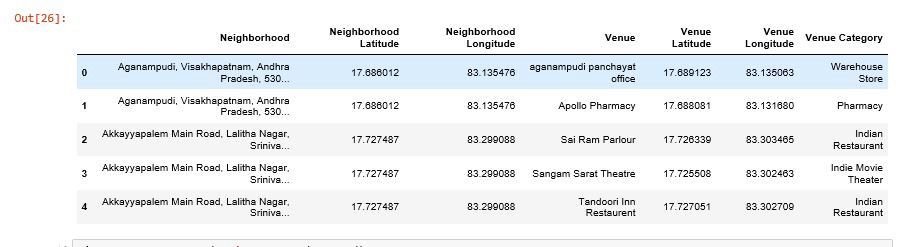
Now, let’s get the top 10 venues that are in Andhra university within a radius of 500 meters.

Foursquare API gave only 2 unique venues



.

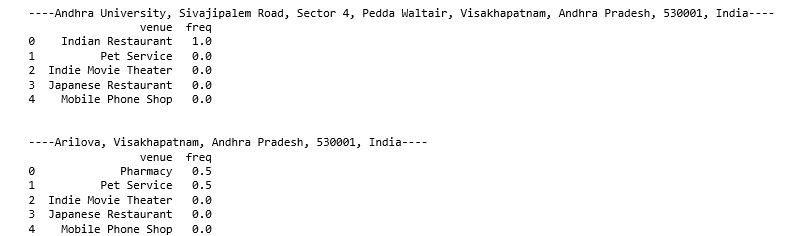
Now we repeat the same steps for all the neighborhoods around Andhra university to get the most common venue categories. Snap shot of first 5 neighborhoods and their venue categories.

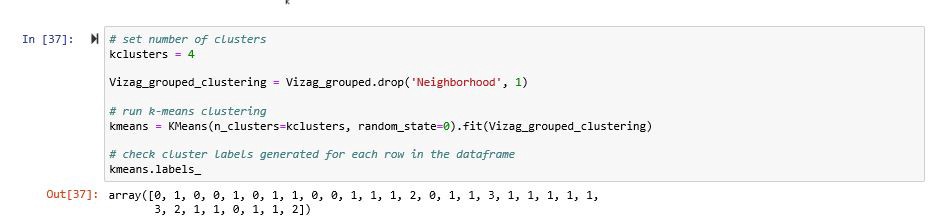


There are 39 unique categories of venues in the neighborhoods of Andhra University.

Now we repeat the same for all the neighborhoods in Visakhapatnam city. Let’s look at first 2 neighborhoods with the top 5 most common venues to get an idea. Refer here for [Code](https://github.com/SriramyaK/Machine-Learning/blob/master/IBM%20Final%20Applied%20Capstone%20Coursera%20Project.ipynb)

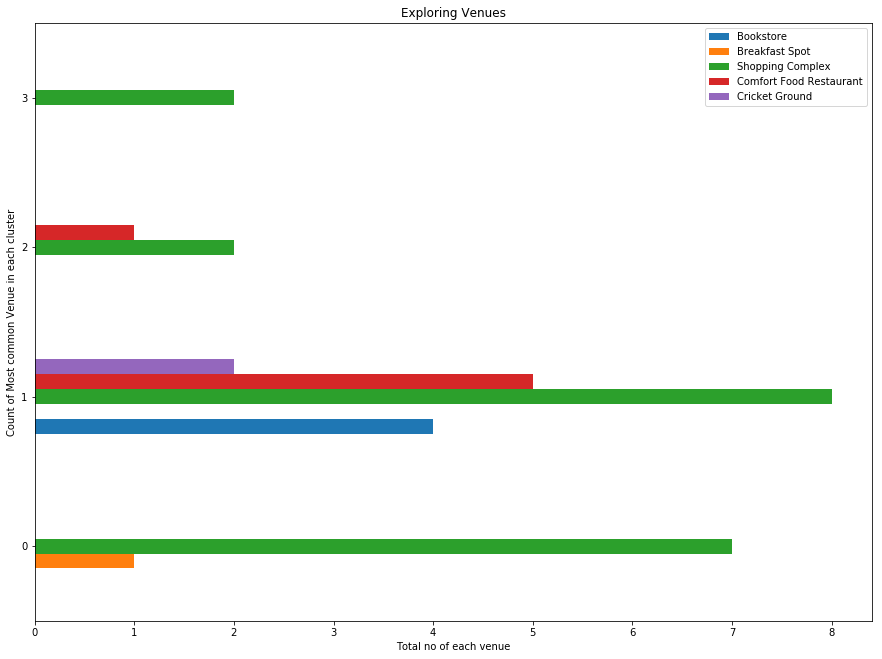




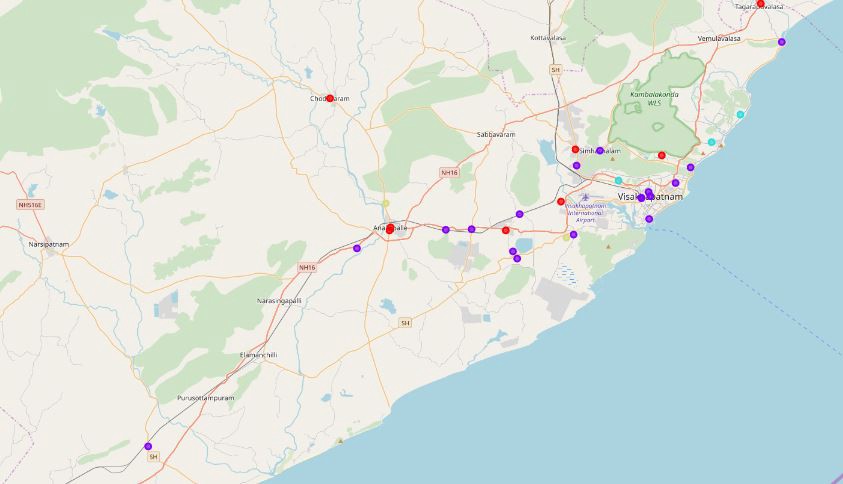


Now we run the k-means algorithm to cluster the neighborhoods into 4 clusters. The no. of clusters is decided by using Elbow method for optimal k. In our scenario the optimal no. of k Is 4.

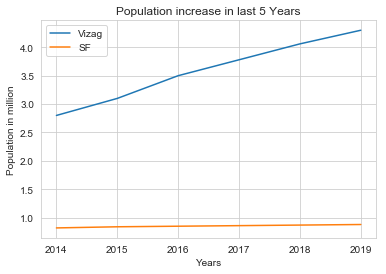
Below horizontal Bar Chart shows the count of most common venues in each cluster. Based on the analysis, we can clearly see the presence of clothing Store/Shopping complex in every cluster which shows the amount of urbanization and development throughout every neighborhood of Visakhapatnam. Breakfast spots, food restaurants are other common venues in cluster 1 and 2.



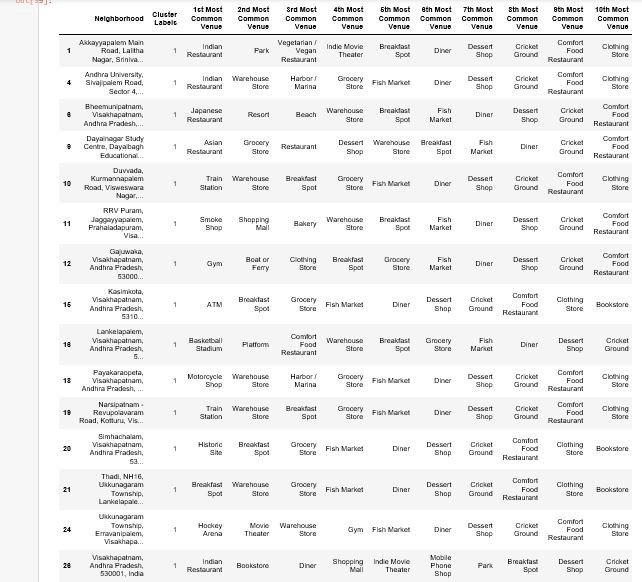
**Fig. Clustering and segmentation of Visakhapatnam using k means algorithm –**



Cluster 1 has the maximum no. of venues and development. There is a significant population increase in recent past. Below is the Bar chart depicting the population of both the cities in last 5 years.



Let’s explore the data further.



We can see the presence of Historic sites, harbor, fish markets and beach which gives us some idea on the geographical similarity between Vizag and SFO. Let’s Visualize this in word clouds with Tourism data of Vizag and Toronto scraped from the travel website TripAdvisor.

**4. Results:**

Though we could show limited results in demographic and geographical factors from the given data set in the clustering and segmentation of the two cities and word clouds of the Wikipedia pages of Visakhapatnam and Toronto, but we could bring out some business ideas on the new venue categories like dance studios, juice bars, coffee shops, event spaces and wide range of restaurants like sushi restaurant, Mediterranean restaurant etc. which can be tailored based on the priorities and interests of local population in Visakhapatnam.

The data can be used to compare the neighbourhood of both cities and can help investors to find appropriate opportunities of real estate investments in Vizag.

**5. Discussion and Conclusion:**

1. The neighbourhoods of both cities have shown multiple similarities in terms of nearby facilities availability.
2. Educational Institutions data can be explored further.
3. Business investor looking for real estate investment can further explore areas/neighborhoods in cluster 1 of Visakhapatnam as these are the areas having the highest development with restaurants, breakfast spots, shopping complex etc. as compared to the places in other clusters.
4. For people interested in coming up with startup ideas in the food sector of smart city — dance studios, juice bars, coffee shops, event spaces and wide range of restaurants like sushi restaurant, Mediterranean restaurant etc. are some of the new business ideas that can be experimented with based on further data analysis.
5. Individual investors looking for investment in residential plots can further explore areas in cluster 0 and cluster 2 of Visakhapatnam.

**References:**

* [1] [Yo!Vizag — City’s Exclusive Magazine and Portal](https://www.yovizag.com/visakhapatnam-vizag-pin-code/)
* [2] [Foursquare API](https://developer.foursquare.com/)
* [3] [Visakhapatnam Wikipedia page](https://en.wikipedia.org/wiki/Visakhapatnam)
* [4] [Toronto Wikipedia Page](https://en.wikipedia.org/wiki/San_Francisco)
* [5] <https://india.smartcitiescouncil.com/article/vizag-smart-city-model-itself-san-francisco>
* [6] <https://www.smartvizag.in/>
* [7] <http://apedb.gov.in/about-visakhapatnam-district.html>
* [8] [https://www.opendatanetwork.com](https://www.opendatanetwork.com/)