

Can Vizag Smart City model itself as San Francisco?

Sriramya Kannepalli

May 7, 2019

1. Introduction

1.1 Background



This Project is a part of the Final IBM Coursera Data Science Capstone Course done based on the article published in Business-Standard.com on 15th June 2016 regarding Vizag Smart city to remodel itself as San Francisco.

Can Vizag remodel itself as San Francisco? If yes, what is the current growth rate of both the cities? How are they interrelated with respect to geography and demographics?

Hans Rosling was a young guest student in India when he first realized that Asia had all the capacities to reclaim its place as the world's dominant economic force. At TEDIndia, he graphs global economic growth since 1858 and predicts the exact date that India and China will outstrip the US. Hans Rowling's TEDIndia video:

https://www.ted.com/talks/hans_rosling_asia_s_rise_how_and_when?utm_source=whatsapp&utm_medium=social&utm_campaign=tedsread

Based on the report published by [SmartCitiesCouncil India](#), Fluentgrid Ltd (formerly Phoenix IT Solutions Ltd), in association with Greater Visakhapatnam Municipal Corporation (GVMC), has launched a state-of-the-art centralized City Command Center.

IBM, the Council's lead partner in India, partnering with two other US-based organizations — AECOM and KPMG — has prepared the entire master plan for Vizag Smart City, marking a strong American role in the city's ambitious plan.

Visakhapatnam (also known as Vizag) is the largest city and the financial capital of the Indian State of Andhra Pradesh. It is also the ninth-most populous metropolitan area in India with a population of 5,018,000. With an output of \$43.5 billion, Visakhapatnam is the ninth-largest contributor to India's overall gross domestic product as of 2016. (Wikipedia)

For the larger project where the US agencies are involved based on an event published in [USTDA Blog](#), the government authorities have been benchmarking Vizag with San Francisco when it comes to the targeted outcomes. San Francisco was also chosen because of the geographical similarity. Situated in northern California, the city is on the tip of a peninsula, surrounded by the Pacific Ocean and a bay. San Francisco is also known for its hilly landscape, among other picturesque places. Vizag, too, has a hilly terrain, with several big formations such as Kailasagiri and Rishikonda overlooking the Bay of Bengal. [5]

As a resident of Vizag, I decided to explore the neighborhoods of Vizag, AP, India with the neighborhoods of San Francisco, US to understand the investment opportunities and the city overall growth and development at par with San Francisco 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.

1.2 Problem

Data that shows current status of the 2 cities and identifies potential areas and different sectors of investment in Visakhapatnam. This is achieved by comparing the neighborhoods of Vizag and San Francisco and visualizing data for identifying patterns in their geographical and demographic similarities.

1.3 Interest

This project will highlight the investor opportunities with increased scope of attracting NRI (Non-Resident Indians) investments which can help Vizag to realize its ambitious economic growth goals while preserving and enhancing livability for the benefit of local citizens.[6]. Best

attempt has been done to bring all the resources at one place to highlight the potential of Vizag and showcase the efforts being spent by the government authorities in achieving the smart city goals.

For further reading and reference on current undergoing projects and opportunities –

1. https://smartnet.niua.org/sites/default/files/resources/Vizag%20Smart%20City_Draft%20Strategies%20%26%20Process_Nov%2010_AECOM_PMC.pdf
2. <https://www.smartvizag.in/index.php/smart-infrastructure-opportunities/>

2. Data

2.1 Data Requirements

Following datasets have been used in the project:

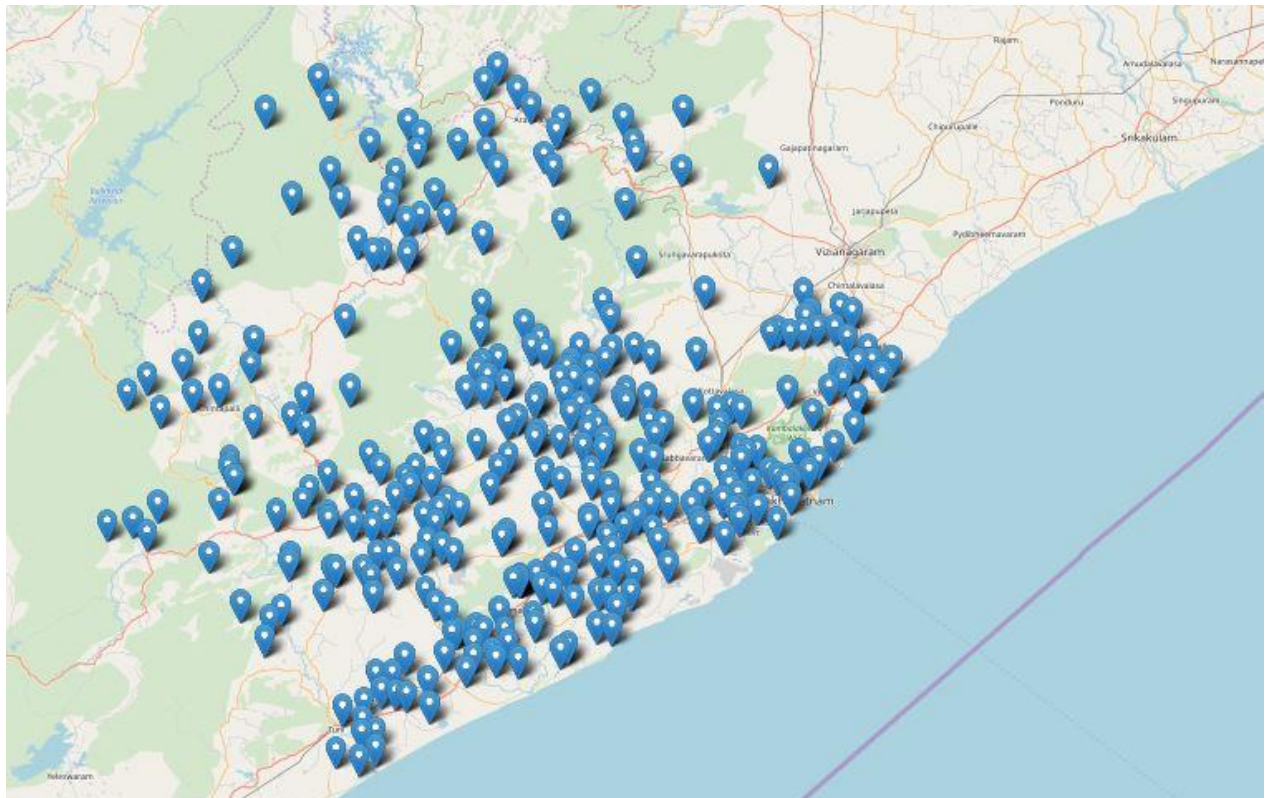
- Postal Codes of Visakhapatnam. Data has been scraped and cleaned from Yo!Vizag – City's Exclusive Magazine and Portal [1] using Beautiful Soup and pandas libraries and saved in .csv format.
- Foursquare API to get the most common venues of given boroughs of Visakhapatnam and San Francisco respectively.[2]
- Visakhapatnam [3] and San Francisco Wikipedia Pages [4] have been scraped and cleaned for creating Word clouds.
- Zip codes of San Francisco. Data has been downloaded in .csv format from <https://datasf.org/> and cleaned using pandas.
- GDP and Per Capita Income data of Visakhapatnam - https://en.wikipedia.org/wiki/Economy_of_Visakhapatnam
- GDP and Per Capita Income data of San Francisco - <https://fred.stlouisfed.org/series/NGMP41860>
- Population Data of Visakhapatnam - <https://indiapopulation2019.com/population-of-visakhapatnam-2019.html>
- Population data of San Francisco - https://www.opendatanetwork.com/entity/1600000US0667000/San_Francisco_CA/demographics.population.count?year=2017
- Tourism Data -

2.2 Data Analysis:

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

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](#)- using Beautiful Soup and pandas libraries and saved in .csv format. Below are the 1st five areas:

```
In [4]: df = pd.read_csv('Vizag_Pincodes.csv')
df.head()
```

Out[4]:

	Location	Pincode	City	State
0	A Kothapalle	531022	Visakhapatnam	Andhra Pradesh
1	A U engg college	530003	Visakhapatnam	Andhra Pradesh
2	A Veeranarayanam	531027	Visakhapatnam	Andhra Pradesh
3	Adakula	531115	Visakhapatnam	Andhra Pradesh
4	Addumanda	531077	Visakhapatnam	Andhra Pradesh

San Francisco Data:

Data 2: SFO has total of 36 neighborhoods [4]. But due to limited data available we could analyze only 26 neighborhoods. Data has been downloaded in .csv format from <https://datasf.org/> and cleaned using pandas. Below are the 1st five neighborhoods:

```
In [7]: df = pd.read_csv('SFO Pincodes.csv')
df['City'] = 'San Francisco'
df['State'] = 'California'
df.head()
```

Out[7]:

	Zip Codes	Neighborhood Name	City	State
0	94102	Hayes Valley	San Francisco	California
1	94103	SoMa	San Francisco	California
2	94104	Financial District	San Francisco	California
3	94105	Embarcadero South	San Francisco	California
4	94107	Portrero Hill	San Francisco	California

Data 3: For the below analysis we will get data from Wikipedia:

1. Visakhapatnam and San Francisco City Demographics.
2. Visakhapatnam Tourism and Attractions.
3. San Francisco Tourism and conventions.

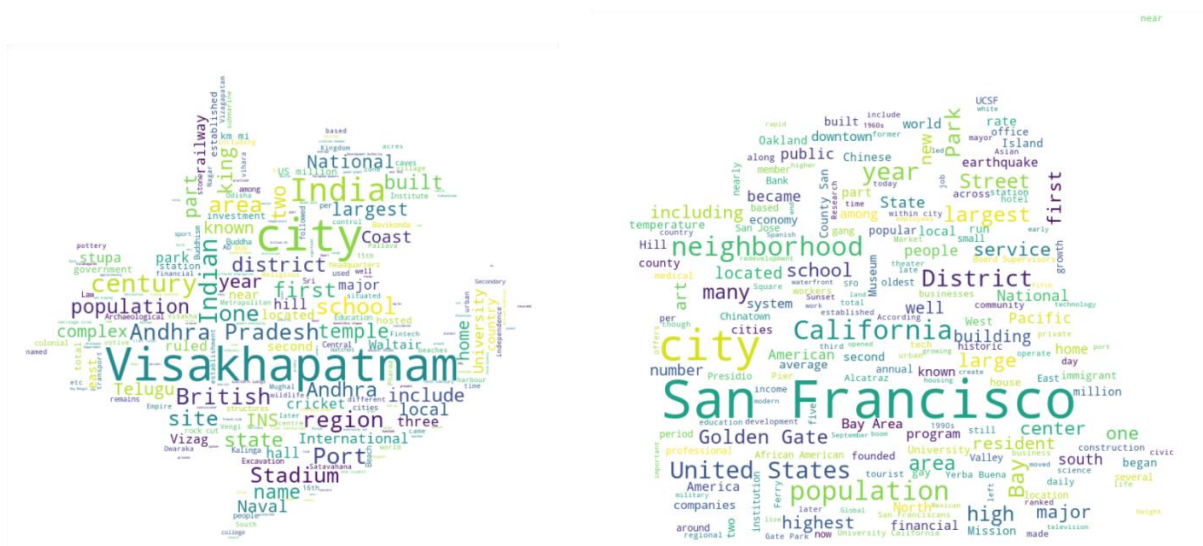


Fig. Word clouds of Vizag and SFO

Data 4: Visakhapatnam and SFO geographical coordinates will be utilized as input for Foursquare API, that will be leveraged to extract information for each neighborhood respectively. We will use the Foursquare

API to explore neighborhoods in Visakhapatnam and SFO City. The below is image of Foursquare API data for the 2 cities.

Visakhapatnam Foursquare API data:

Out[5]:

	Latitude	Longitude	Neighborhood
0	18.211488	82.535531	Adduru, Visakhapatnam, Andhra Pradesh, India
1	17.888012	83.135476	Aganampudi, Visakhapatnam, Andhra Pradesh, 530...
2	17.721842	83.235434	Visakhapatnam International Airport, NH16, Kak...
3	17.727487	83.299088	Akkayypalem Main Road, Lalitha Nagar, Sriniva...
4	17.931495	83.003157	Alamanda, Visakhapatnam, Andhra Pradesh, India

SFO Foursquare API data:

Out[6]:

	Latitude	Longitude	Neighborhood
0	37.776685	-122.422938	Hayes Valley, SF, California, 94102, USA
1	37.780893	-122.400952	South of Market, SF, California, 94104, USA
2	37.793647	-122.398938	Financial District, SF, California, 94111, USA
3	37.792884	-122.396912	Embarcadero, Market Street, Financial District...
4	37.794301	-122.406376	Chinatown, SF, California, 94107, USA

Data 5: Population, GDP, Per capita Income, Tourism, Educational Institutions and Weather data of Vizag and San Francisco.

3. Methodology:

Analytic Approach:

In this project, first part is clustering of Visakhapatnam using k means algorithm. Visakhapatnam has 648 pin codes/areas/postal codes, geocodes of only 326 locations have been included in the data analysis. We will explore the areas around central Visakhapatnam and compare it with the neighborhoods of San Francisco to understand the geographical similarities.

2nd part comprises of clustering of San Francisco. For San Francisco out of 36 neighborhoods venues of 27 neighborhoods have been explored in this project using Foursquare API.

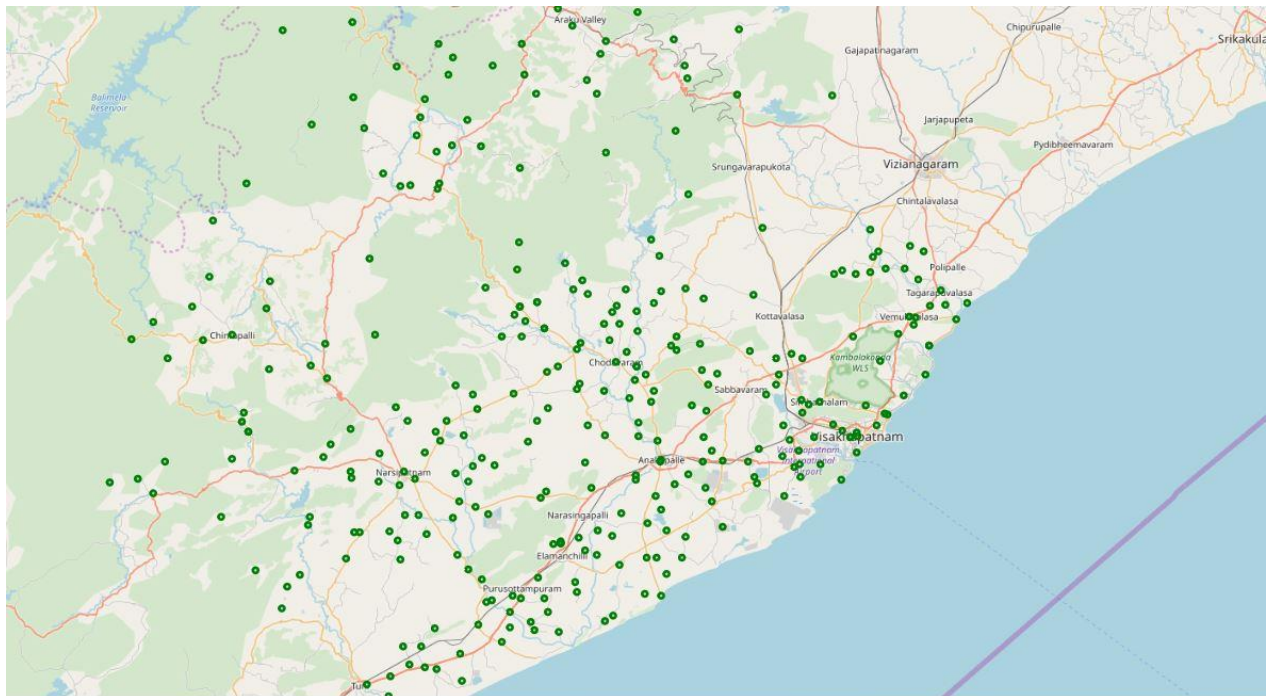
3rd part includes data visualizations and comparison of available data of both the cities for insights to take investment decisions in Vizag. Word clouds created from the wiki pages of Vizag and SFO further add value to our discussion.

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](#) on one side of it and on the other, the green [Kailasagiri](#) hill range. This location is apt for our analysis as San Francisco 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

Out[22]:

	name	categories	lat	lng
0	my restaurant	Indian Restaurant	17.738538	83.326882
1	Fring	Indian Restaurant	17.741645	83.331995

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.

Out[26]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Aganampudi, Visakhapatnam, Andhra Pradesh, 530...	17.688012	83.135476	aganampudi panchayat office	17.689123	83.135063	Warehouse Store
1	Aganampudi, Visakhapatnam, Andhra Pradesh, 530...	17.688012	83.135476	Apollo Pharmacy	17.688081	83.131680	Pharmacy
2	Akkayyapalem Main Road, Lalitha Nagar, Sriniva...	17.727487	83.299088	Sai Ram Parlour	17.726339	83.303465	Indian Restaurant
3	Akkayyapalem Main Road, Lalitha Nagar, Sriniva...	17.727487	83.299088	Sangam Sarat Theatre	17.725508	83.302463	Indie Movie Theater
4	Akkayyapalem Main Road, Lalitha Nagar, Sriniva...	17.727487	83.299088	Tandoori Inn Restaurant	17.727051	83.302709	Indian Restaurant

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.

```
----Andhra University, Sivajipalem Road, Sector 4, Pedda Waltair, Visakhapatnam, Andhra Pradesh, 530001, India----
  venue  freq
0  Indian Restaurant  1.0
1    Pet Service    0.0
2  Indie Movie Theater  0.0
3  Japanese Restaurant  0.0
4  Mobile Phone Shop  0.0

----Arlilova, Visakhapatnam, Andhra Pradesh, 530001, India----
  venue  freq
0   Pharmacy  0.5
1   Pet Service  0.5
2  Indie Movie Theater  0.0
3  Japanese Restaurant  0.0
4  Mobile Phone Shop  0.0
```

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.

```
In [37]: # set number of clusters
kclusters = 4

Vizag_grouped_clustering = Vizag_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Vizag_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_
```

Out[37]: array([0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 2, 0, 1, 1, 3, 1, 1, 1, 1, 1, 3, 2, 1, 1, 0, 1, 1, 2])

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.

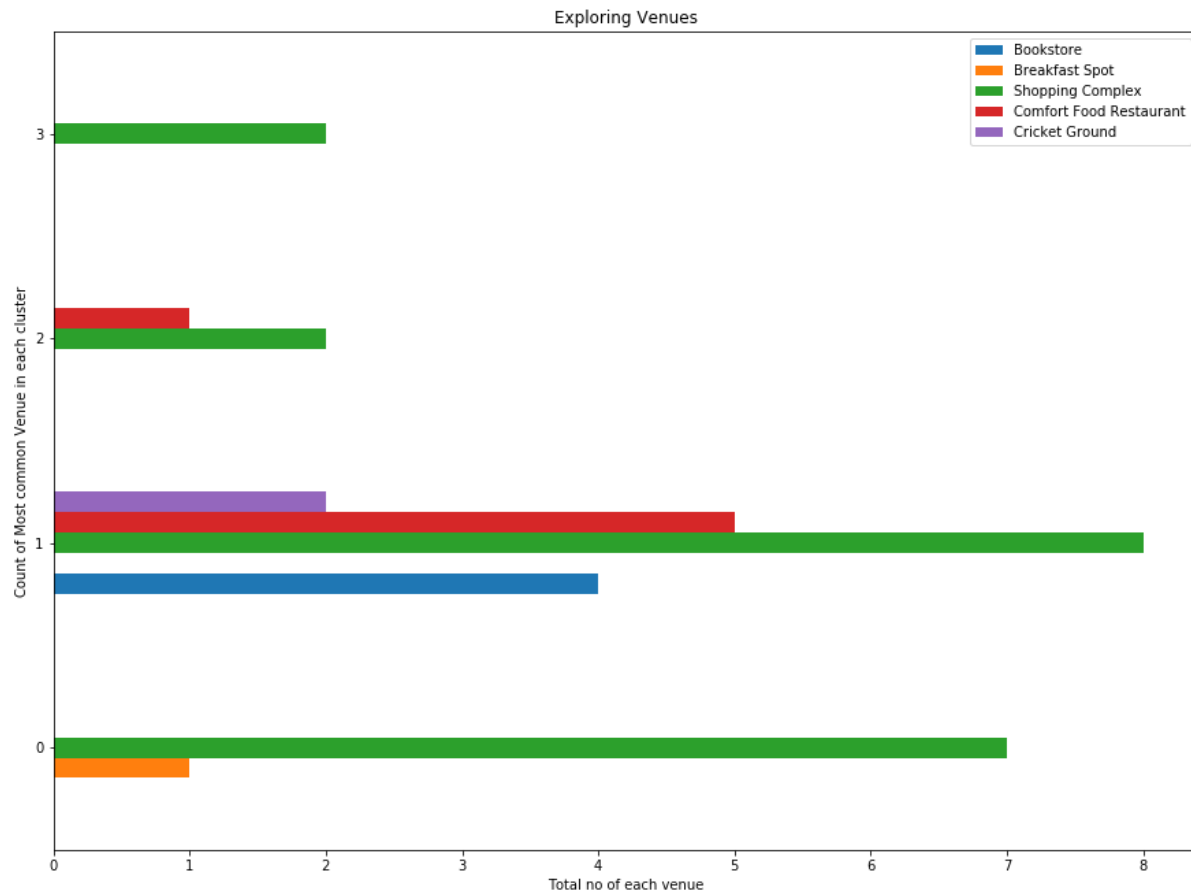
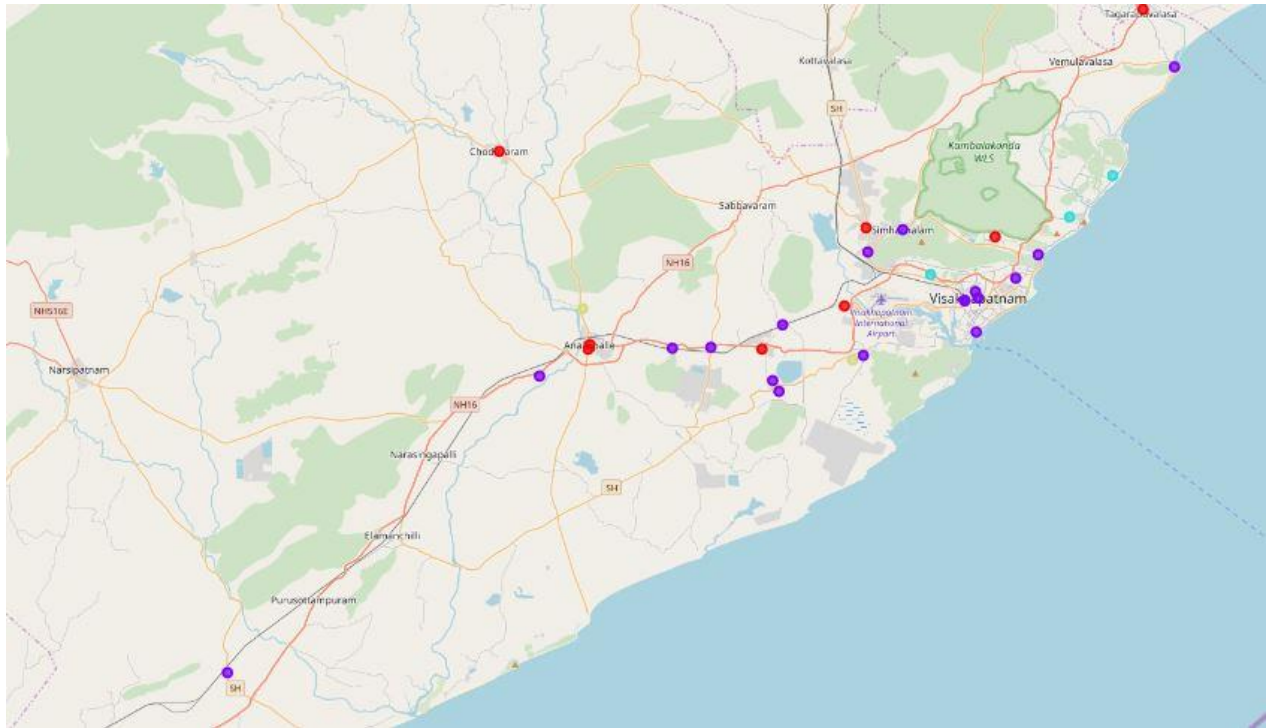
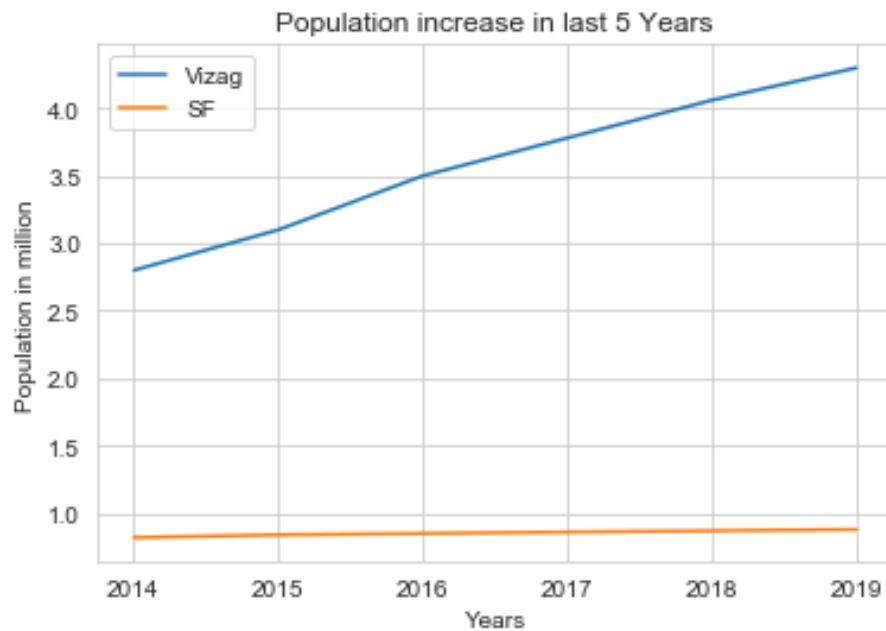


Fig. Clustering and segmentation of Visakhapatnam using k means algorithm -



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.

	Neighborhood	Cluster Labels	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Akkayyapalem Main Road, Lalitha Nagar, Sriniva...	1	Indian Restaurant	Park	Vegetarian / Vegan Restaurant	Indie Movie Theater	Breakfast Spot	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store
4	Andhra University, Shajjapalem Road, Sector 4,...	1	Indian Restaurant	Warehouse Store	Harbor / Marina	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store
6	Bheemunipatnam, Visakhapatnam, Andhra Pradesh,...	1	Japanese Restaurant	Resort	Beach	Warehouse Store	Breakfast Spot	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant
9	Dayanagar Study Centre, Dayabagh Educational...	1	Asian Restaurant	Grocery Store	Restaurant	Dessert Shop	Warehouse Store	Breakfast Spot	Fish Market	Diner	Cricket Ground	Comfort Food Restaurant
10	Duvvada, Kurmannapalem Road, Visweswara Nagar,...	1	Train Station	Warehouse Store	Breakfast Spot	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store
11	RRV Puram, Jaggayyapalem, Prahaladapuram, Visa...	1	Smoke Shop	Shopping Mall	Bakery	Warehouse Store	Breakfast Spot	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant
12	Gajuwaka, Visakhapatnam, Andhra Pradesh, 53000...	1	Gym	Boat or Ferry	Clothing Store	Breakfast Spot	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant
15	Kasimkota, Visakhapatnam, Andhra Pradesh, 5310...	1	ATM	Breakfast Spot	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store	Bookstore
16	Lankelapalem, Visakhapatnam, Andhra Pradesh, 5...	1	Basketball Stadium	Platform	Comfort Food Restaurant	Warehouse Store	Breakfast Spot	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground
18	Payakaraoopeta, Visakhapatnam, Andhra Pradesh, ...	1	Motorcycle Shop	Warehouse Store	Harbor / Marina	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store
19	Narsipatnam - Revupolavaram Road, Kotturu, Vis...	1	Train Station	Warehouse Store	Breakfast Spot	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store
20	Simhachalam, Visakhapatnam, Andhra Pradesh, 53...	1	Historic Site	Breakfast Spot	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store	Bookstore
21	Thadi, NH16, Ulkunagaram Township, Lankelapale...	1	Breakfast Spot	Warehouse Store	Grocery Store	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store	Bookstore
24	Ulkunagaram Township, Erravanipalem, Visakhapa...	1	Hockey Arena	Movie Theater	Warehouse Store	Gym	Fish Market	Diner	Dessert Shop	Cricket Ground	Comfort Food Restaurant	Clothing Store
26	Visakhapatnam, Andhra Pradesh, 530001, India	1	Indian Restaurant	Bookstore	Diner	Shopping Mall	Indie Movie Theater	Mobile Phone Shop	Park	Breakfast Spot	Dessert Shop	Cricket Ground

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 San Francisco scraped from the travel website TripAdvisor:

Fig. Word cloud of San Francisco list of tourist Attractions:



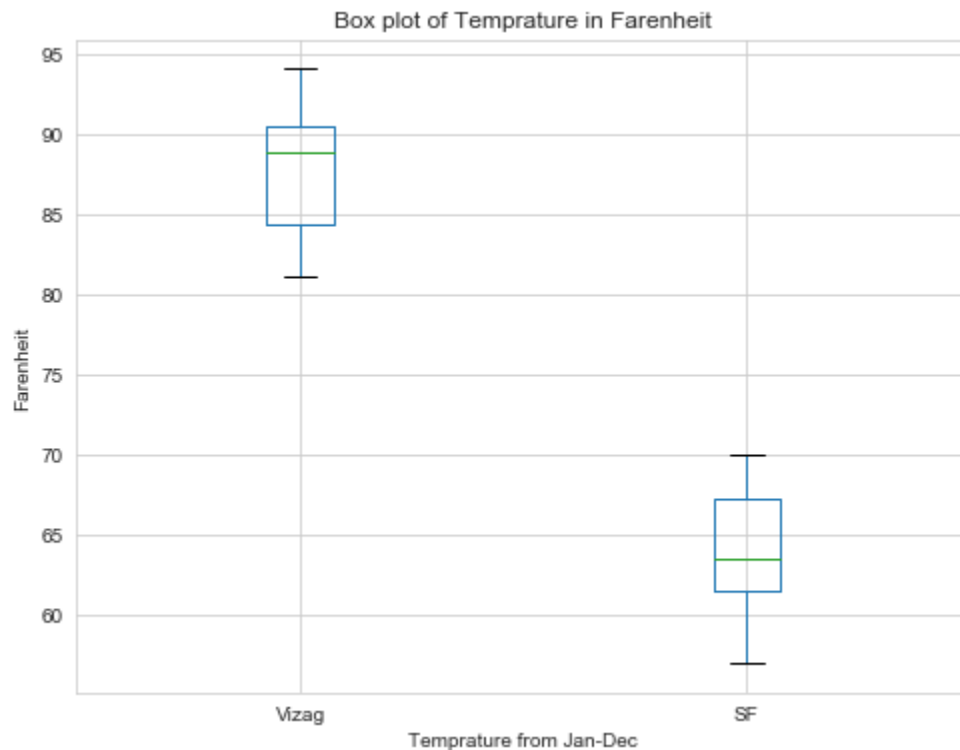
Word Cloud of list of Visakhapatnam tourist attractions:



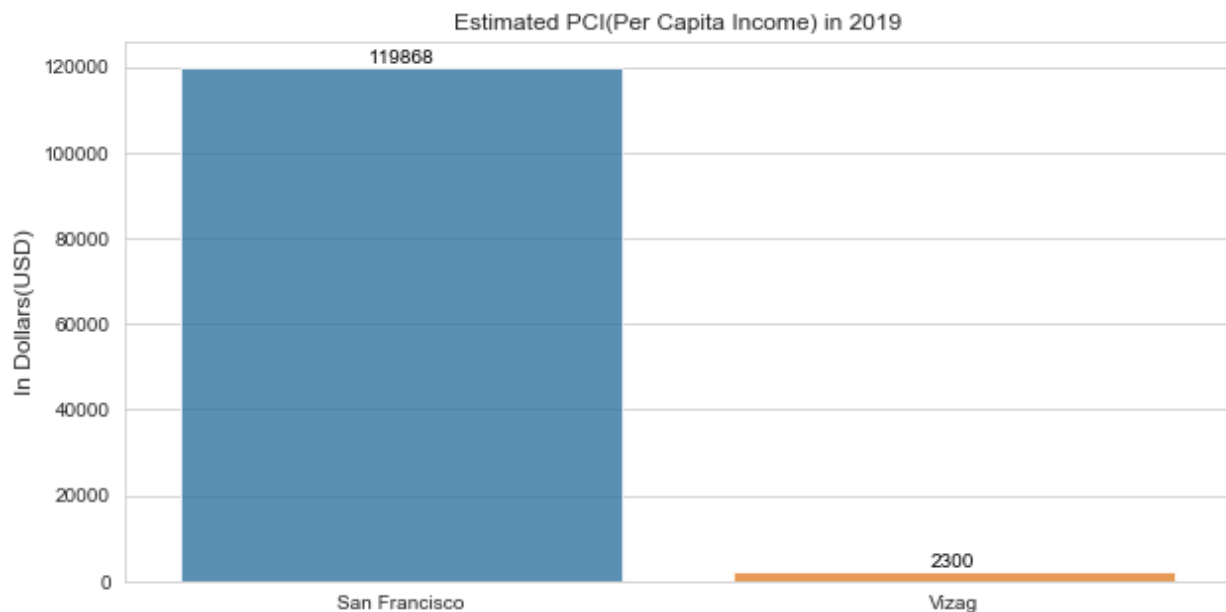
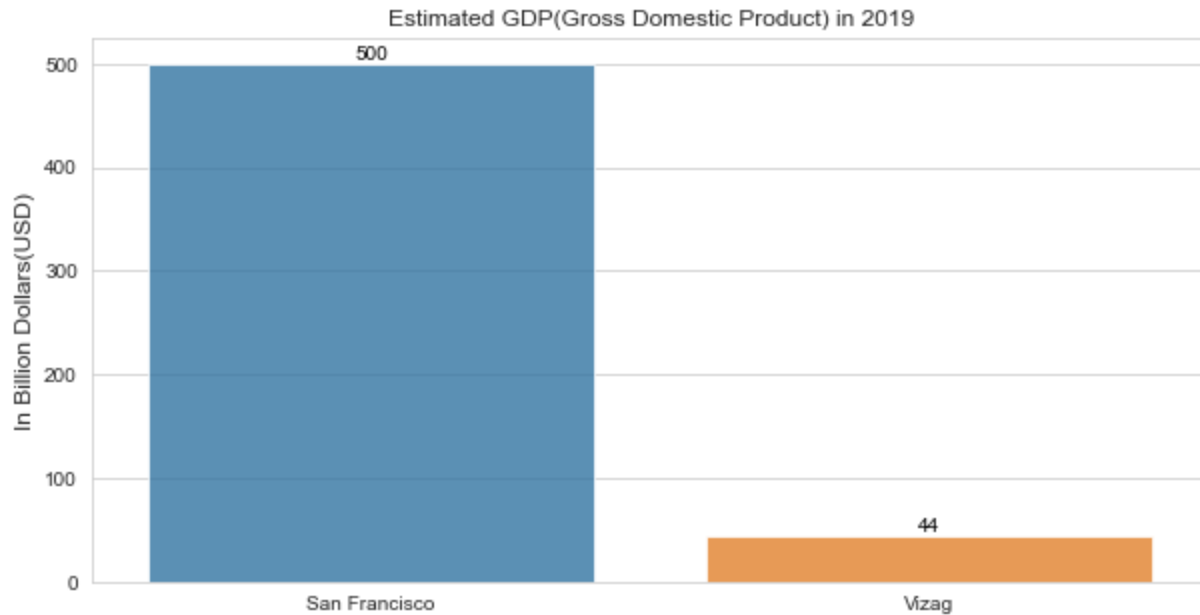
Above word clouds signify the similarity in the two cities Museum, Park and Beach/Bay being the most common among them. Some other already existing natural tourist spots adding to the beauty of the city are waterfalls, caves, hills, wildlife and temples in Visakhapatnam.

But when we closely observe the word cloud of tourism of San Francisco, there are several untapped opportunities like Fisherman's wharf, Pier 39, Twin Peaks, Big Bus Hop on Hop off tour etc. that can be implemented in Visakhapatnam due to similar geographical features and weather conditions.

Box plot of weather conditions of Visakhapatnam and San Francisco in a Year:



The hot and humid conditions of Visakhapatnam as compared to San Francisco clearly show huge scope for establishment of amusement water parks and recreational activities. **Cruises, Sailing, Hiking trails and Water tours** can create major spike in tourism and boost **GDP** of Visakhapatnam.

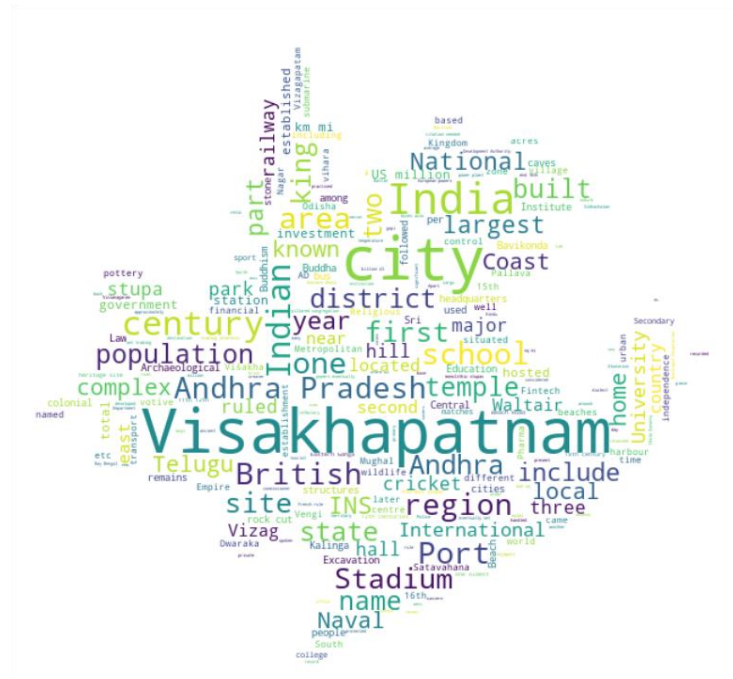


Though there is a significant difference in the GDP and Per Capita Income of Vizag and San Francisco, Visakhapatnam has managed to top the charts of urban population amongst all the 13 districts in Andhra Pradesh, India. According to data uploaded onto the CM's Dashboard, the 2011 Census of India states that Visakhapatnam stood first in the state with 47.45% of urban populace.

The difference in GDP and Per Capita Income of the two cities signify the importance of technology and investments required for the city to remodel itself as San Francisco in the next 10 years.[1] Achieving the vision will require a "Smart City" approach to regional development and infrastructure planning and delivery. For further information please refer the below link - <https://www.smartvizag.in/index.php/projects/>

To Summarize, we created word cloud using seaborn libraries and web scraping Wikipedia page using beautiful soup.

Fig. Word Cloud of Visakhapatnam Wikipedia Page



In this word cloud we can clearly see that Visakhapatnam has a coast, port, railway, naval base, university, stadium and is a metropolitan city with historic sites and international airport.

With this information we move on to the analysis of San Francisco and identify potential ideas for development.

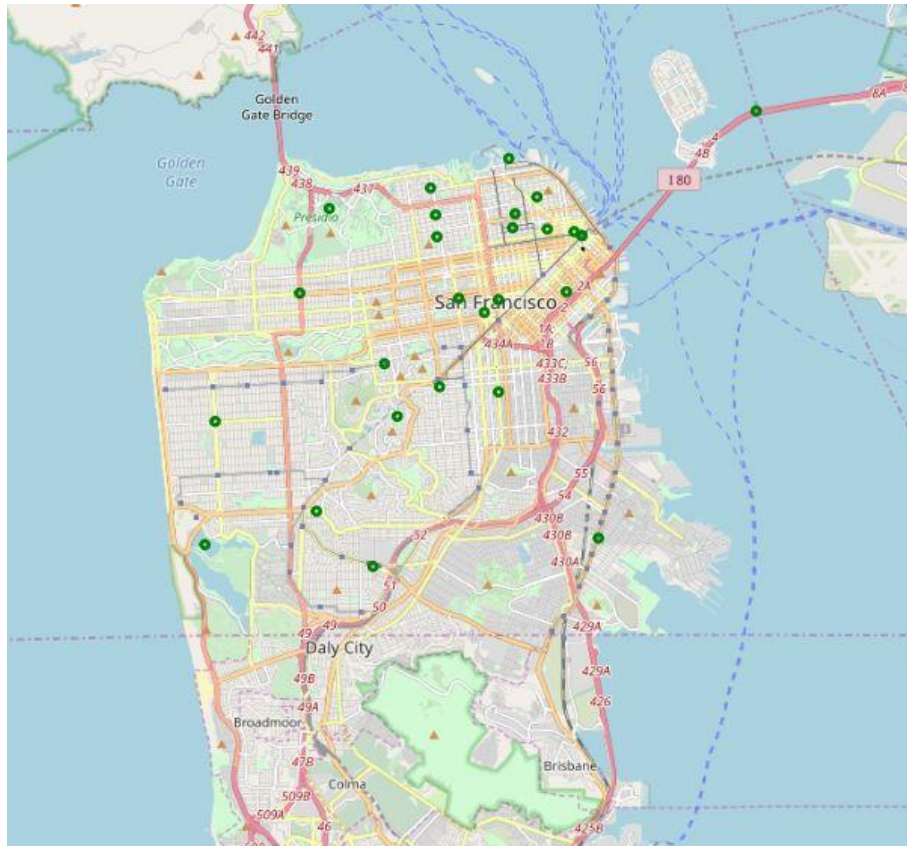
Data 2: SFO Geographical Coordinates Data is downloaded in .csv format from <https://datasf.org/> and cleaned using pandas. We explored 27 neighborhoods of San Francisco in our analysis.

Below are the first 5 neighborhoods.

Out[7]:

	Zip Codes	Neighborhood Name	City	State
0	94102	Hayes Valley	San Francisco	California
1	94103	SoMa	San Francisco	California
2	94104	Financial District	San Francisco	California
3	94105	Embarcadero South	San Francisco	California
4	94107	Portrero Hill	San Francisco	California

Fig. SFO Neighborhood Visualization using Folium and geopy libraries.



As we explore each neighborhood further for identifying similarities with Visakhapatnam, let's start with venues around the neighborhood surrounded with Beach in SFO.

```
In [20]: df_.loc[24, 'Neighborhood']
Out[20]: 'North Beach, SF, California, 94133, USA'

In [21]: neighborhood_latitude = df_.loc[24, 'Latitude']
neighborhood_longitude = df_.loc[24, 'Longitude']
neighborhood_name = df_.loc[24, 'Neighborhood']

print('Longitude and Latitude values of {} are {},{}'.format(neighborhood_name, neighborhood_latitude, neighborhood_longitude))

Longitude and Latitude values of North Beach, SF, California, 94133, USA are 37.8011749,-122.4090021.
```

Now, let's get the top 100 venues that are in North Beach, SF, California within a radius of 500 meters.

Foursquare API gave 100 unique venues. Let's explore the data –

```
Out[25]:
```

	name	categories	lat	lng
0	Tony's Pizza Napoletana	Pizza Place	37.800387	-122.409146
1	Golden Boy Pizza	Pizza Place	37.799840	-122.408070
2	Glow Yoga & Wellness	Yoga Studio	37.800125	-122.409114
3	Sotto Mare Oysteria and Seafood Restaurant	Seafood Restaurant	37.799625	-122.408357
4	Park Tavern	New American Restaurant	37.801097	-122.409301

We will do the same analysis for all the neighborhoods of North Beach, SF and explore the venues returned by Foursquare API to understand the most common venue categories.

```
In [25]: print(NorthBeaches_SF0_venues.shape)
NorthBeaches_SF0_venues.head()

(1670, 7)

Out[25]:
```

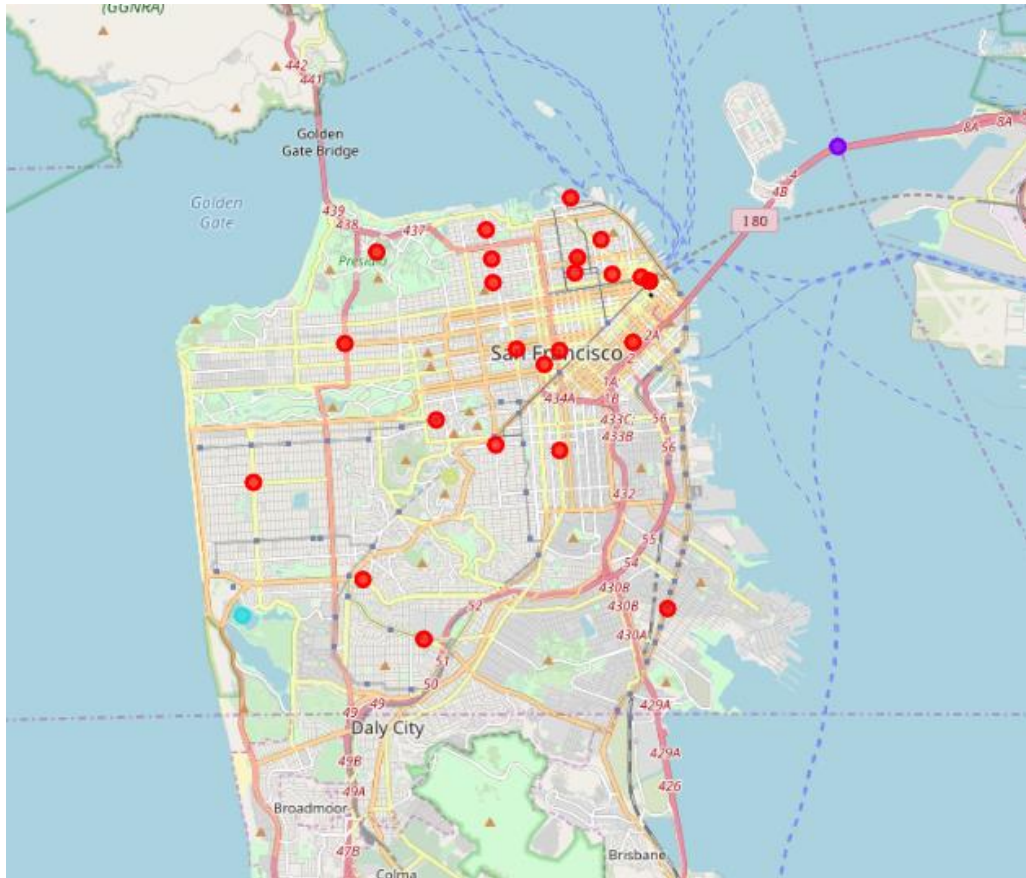
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Hayes Valley, SF, California, 94102, USA	37.776685	-122.422936	SFJazz Center	37.776350	-122.421539	Jazz Club
1	Hayes Valley, SF, California, 94102, USA	37.776685	-122.422936	Ritual Coffee Roasters	37.776476	-122.424281	Coffee Shop
2	Hayes Valley, SF, California, 94102, USA	37.776685	-122.422936	Blue Bottle Coffee	37.776430	-122.423224	Coffee Shop
3	Hayes Valley, SF, California, 94102, USA	37.776685	-122.422936	Birba	37.777750	-122.424159	Wine Bar
4	Hayes Valley, SF, California, 94102, USA	37.776685	-122.422936	Fatted Calf	37.775935	-122.423146	Butcher

```
In [26]: print('There are {} unique categories of venues in SF'.format(len(NorthBeaches_SF0_venues['Venue Category'].unique())))

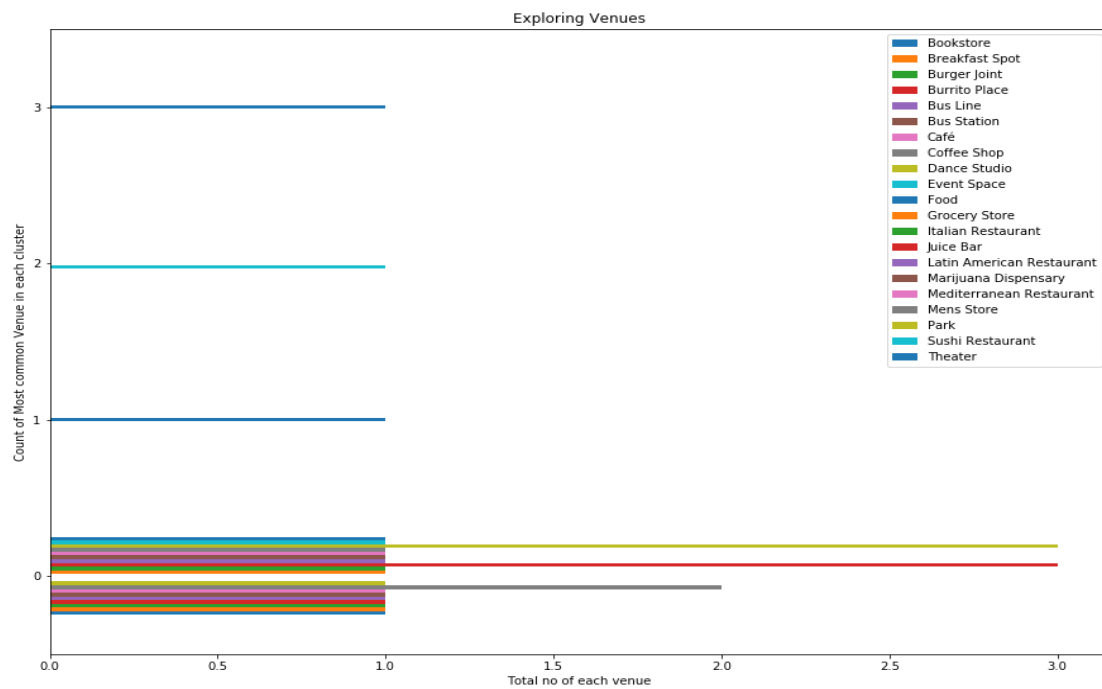
There are 261 unique categories of venues in SF
```

We repeat the same for all the neighborhoods of SF. There are 261 unique categories in SF. 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.

Fig. Clustering of neighborhoods of San Francisco using k means algorithm



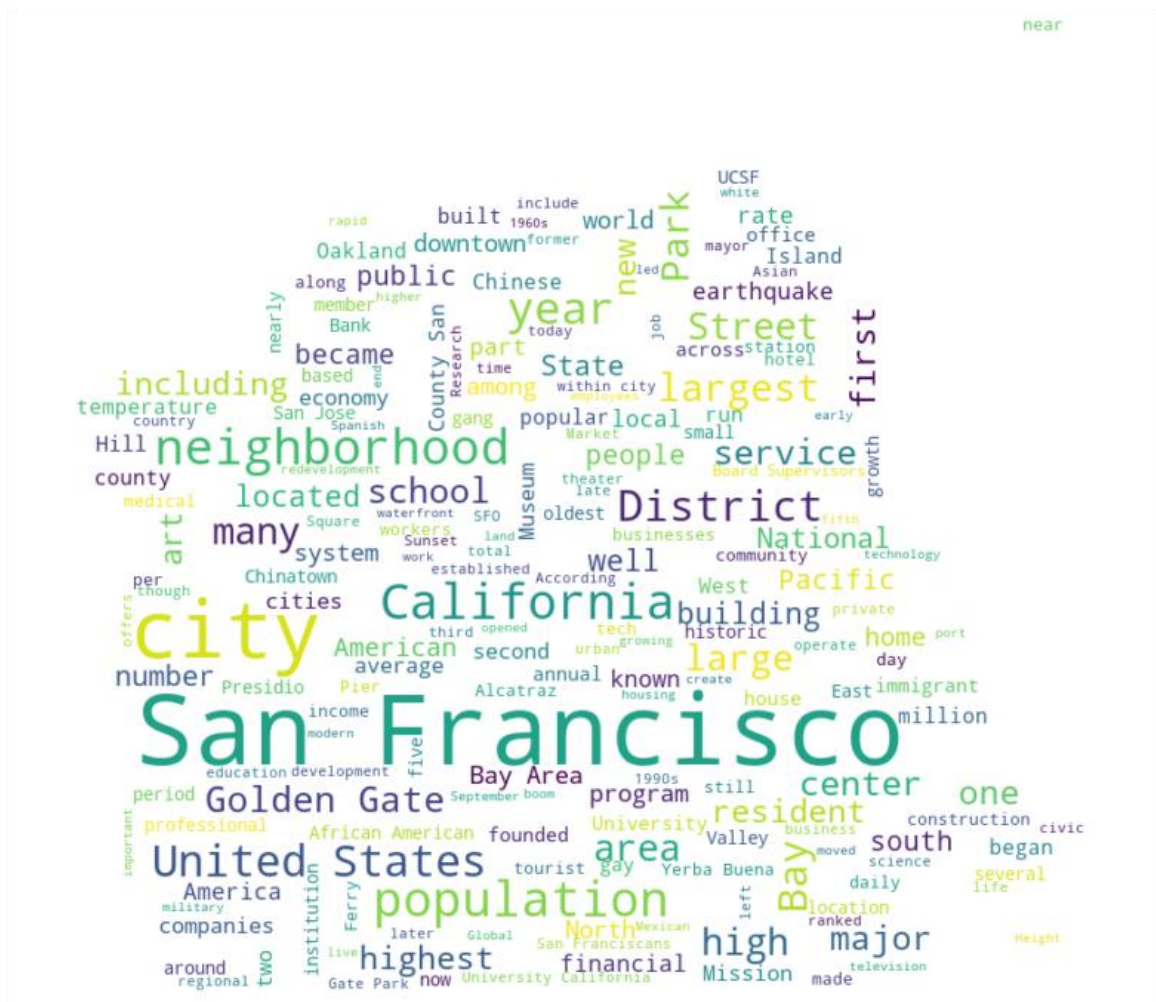
Below is the horizontal bar chart for most common venues in each cluster



As we can see from the above analysis, neighborhoods in cluster 0 are highly developed with wide range of restaurants, dance studios, juice bars, coffee shops, event spaces etc. The venues in Vizag and San Francisco are largely different and unique in nature due to different levels of development/urbanization rates in both the cities.

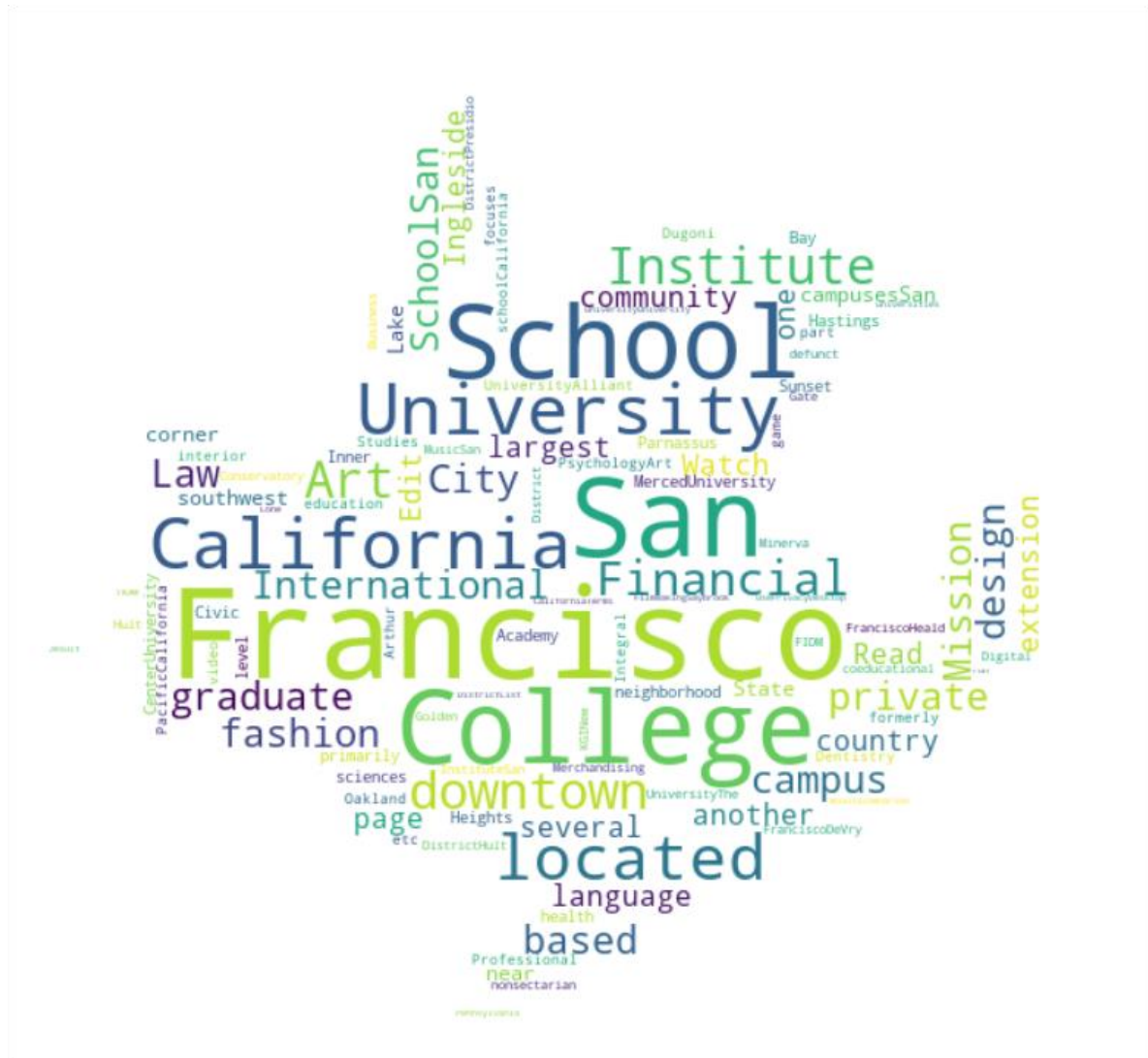
Finally, we will look at the word cloud of San Francisco created from Wikipedia to explore further.

Fig. Word cloud of San Francisco created from Wikipedia



We can see the words military, Bay area, hill, Pacific Ocean, Ferry, waterfront, historic building etc. which show some similarity in the geographic and demographic data of Visakhapatnam and San Francisco.

Word cloud of Educational institutions and universities in San Francisco:



Despite its limited geographical space, San Francisco, California is home to a multitude of colleges and universities. [San Francisco Conservatory of Music](#), [San Francisco School of Digital Filmmaking](#), [San Francisco Art Institute](#) and [Art Institute of California - San Francisco](#), a private campus which focuses on [video game](#) and design-based education (interior, fashion etc.) are some of the unique colleges and universities which can be further explored and established in Visakhapatnam.

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 San Francisco, 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. Tourism when developed in right way with advanced technologies and FDI can play major role boosting city's economy to remodel itself as San Francisco in near future. Educational Institutions form one more area of potential development.

5. Discussion and Conclusion:

1. Tourism has huge potential of development as a part of Smart city initiatives in Vizag. **Cruises, Sailing, Hiking trails and Water tours** can create major spike in tourism and boost **GDP** of Visakhapatnam.
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 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](#)
- [2] [Foursquare API](#)
- [3] [Visakhapatnam Wikipedia page](#)
- [4] [San Francisco Wikipedia Page](#)
- [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/entity/310M200US41860/San_Francisco_Metro_Area_CA/economy.gdp.per_capita_gdp?year=2017