Clustering the different Neighborhoods in Metro Manila Using K-Means Clustering Algorithm
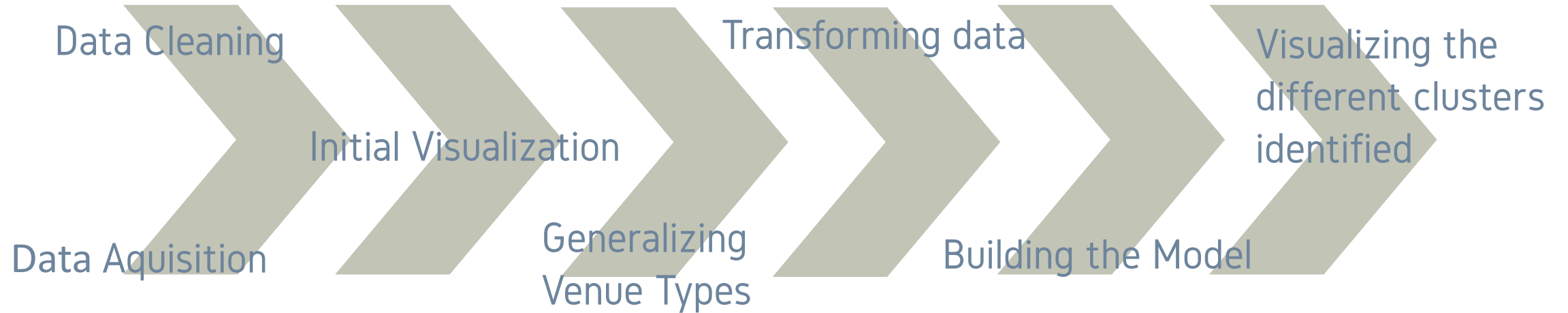
Businesses

Aim & Advantages

People

Customer targeting

Finding areas for new business locations

Knowledge to identify what neighborhoods match the lifestyle they desire.

Methodology

Data Cleaning

Data Aquisition

Initial Visualization

Generalizing Venue Types

Transforming data

Building the Model

Visualizing the different clusters identified

Data Acquisition and Cleaning

List of Neighborhoods in Metro Manila

Details and Types of Venues in Metro Manila

Geospatial Data of the Neighborhoods in Metro Manila

Wikipedia

FourSquare API

Nominatim API

Data Stored in Pandas Data Frame

Data frame with geospatial
data of all the Neighborhoods

Data frame details of the venues
found in all the neighborhoods

	ZIP code	Neighborhood	City	Latitude	Longitude
0	401	Asian Development Bank	San Juan	14.588076	121.058301
1	550	Febias College of Bible	Valenzuela	14.687899	120.981408
2	702	Citibank	Makati	14.607314	121.078924
3	704	Producers Bank	Makati	14.529699	121.041247
4	707	Canadian Embassy	Makati	14.560645	121.016578

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Asian Development Bank	14.588076	121.058301	Craft Coffee Revolution	14.585859	121.059212	Coffee Shop
1	Asian Development Bank	14.588076	121.058301	The Nostalgia Dining Lounge	14.587519	121.059753	Restaurant
2	Asian Development Bank	14.588076	121.058301	Gino's Brick Oven Pizza	14.585791	121.059640	Pizza Place
3	Asian Development Bank	14.588076	121.058301	The Café Mediterranean	14.585967	121.057122	Mediterranean Restaurant
4	Asian Development Bank	14.588076	121.058301	Wildflour Café + Bakery	14.585866	121.059573	Café

Initial Mapping of the Neighborhoods in Metro Manila
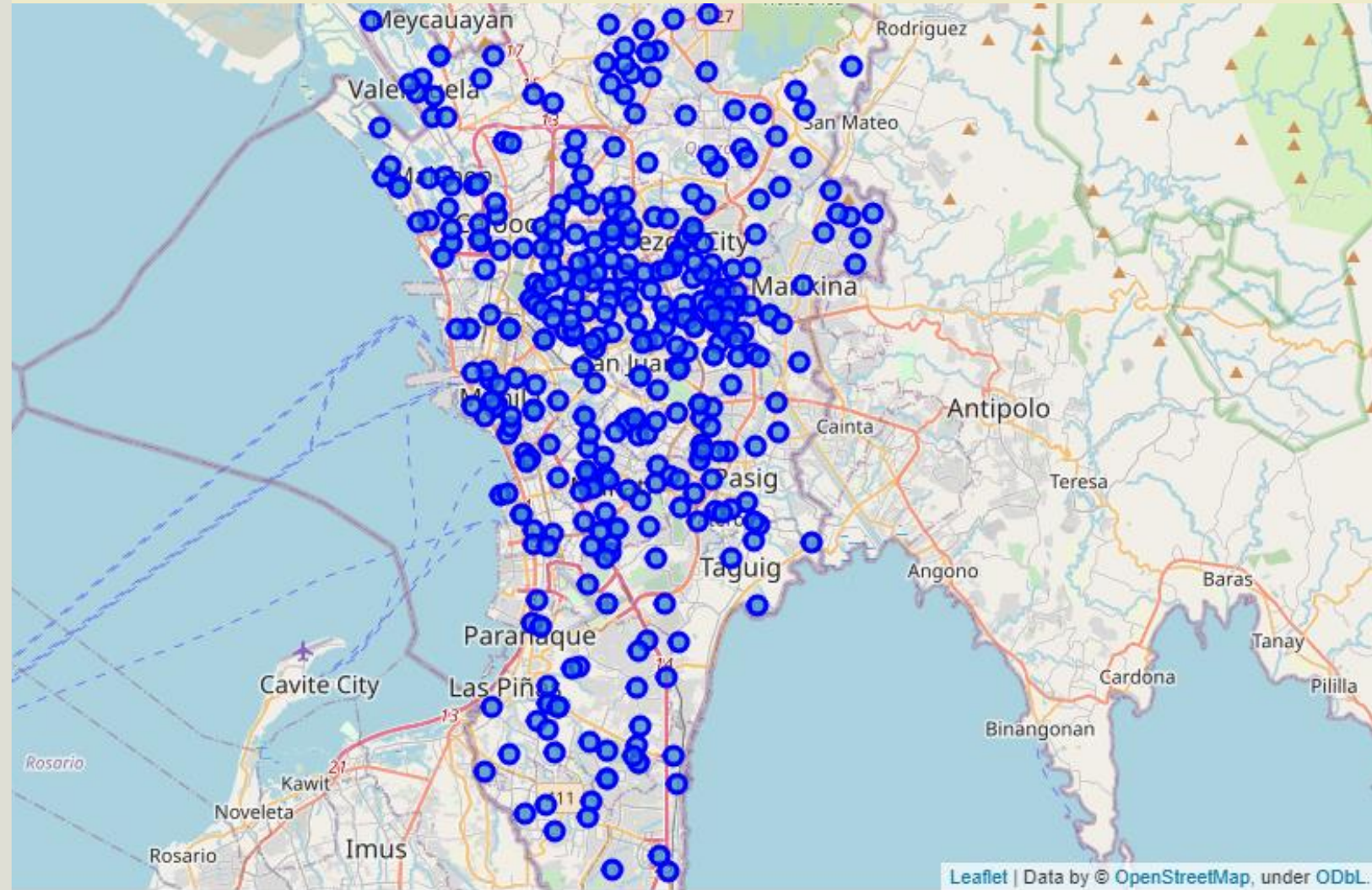
Generalizing the specific Venue Types

337 different venues types from the FourSquare Dataset

Consolidating similar venue types

Venue Type	Venue Category
Residential	Boarding House
	Residential Building (Apartment/Condo)

Venue Type	Venue Category
Accomodation	Hostel
	Hotel
	Hotel Pool
	Lounge
	Motel
	Rest Area
	Roof Deck
	Travel Lounge

	General Venue Types
1	Accomodation
2	Art/Culture/Museum
3	Business Hub
4	Café/Tea/Dessert
5	City Essentials
6	Convenience stores and Food Essentials
7	Cosmetics, Clothes, Beauty and Health Services
8	Entertainment
9	Fast Food
10	Gadgets, Hardware, Toys, etc.
11	Night Life
12	Residential
13	Restaurants
14	Sports & Fitness
15	Tourist Destination
16	Vehicle/Transport Related

Visualization of the Different Venue Types

Show a large bias towards restaurants which is expected of a Metropolitan Region

Skews the model
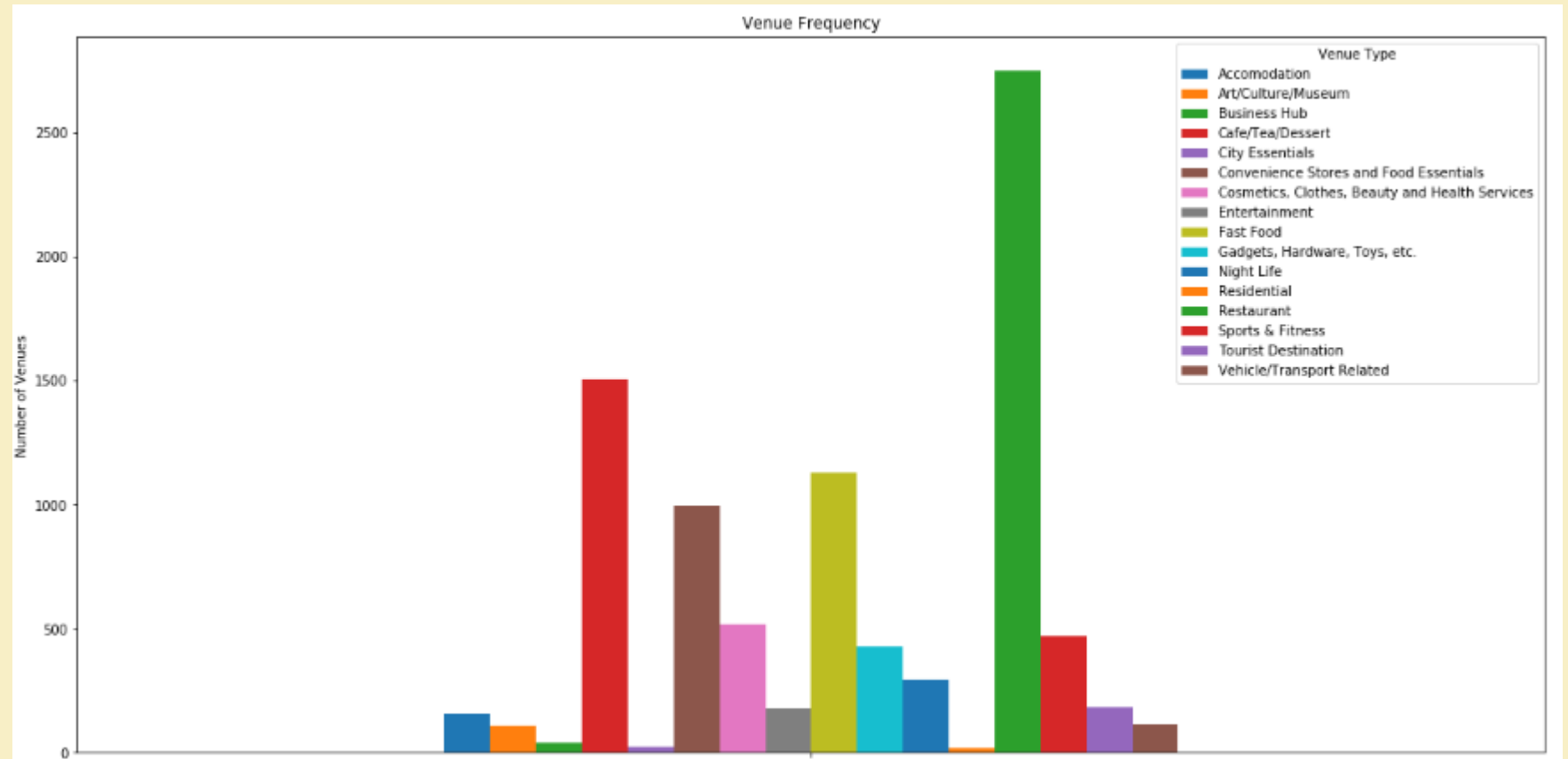
Applies an Importance Factor

Emphasizes lower count venue types

Shows a large prominence on neighborhoods that have high density of the low count venue types

Deemphasize high venue count types

	Importance factor
Venue Type	
Accomodation	0.107595
Art/Culture/Museum	0.155963
Business Hub	0.414634
Cafe/Tea/Dessert	0.011296
City Essentials	0.653846
Convenience Stores and Food Essentials	0.017068
Cosmetics, Clothes, Beauty and Health Services	0.032692
Entertainment	0.096591
Fast Food	0.015084
Gadgets, Hardware, Toys, etc.	0.040000
Night Life	0.057627
Residential	1.000000
Restaurant	0.006182
Sports & Fitness	0.036247
Tourist Destination	0.093923
Vehicle/Transport Related	0.150442

Transform Data using One-hot Encoding

Finding the best K for K-Means Clustering

Using Elbow Method
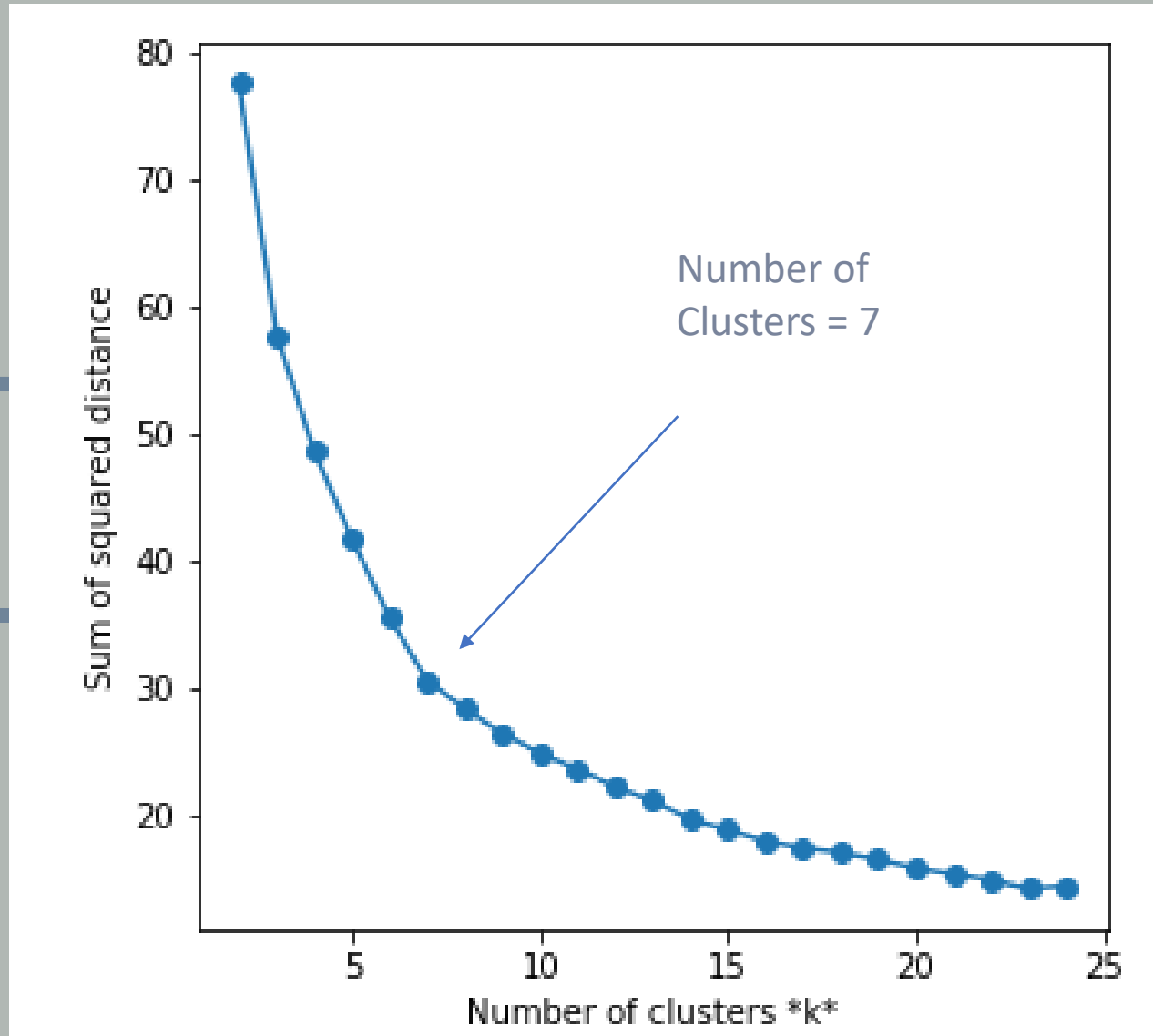
Most Optimal number of Cluster is 7

Visualizing the Clusters

Cluster 1: Arts, Culture and Science District

Notice the high value for Arts, Culture and Museum
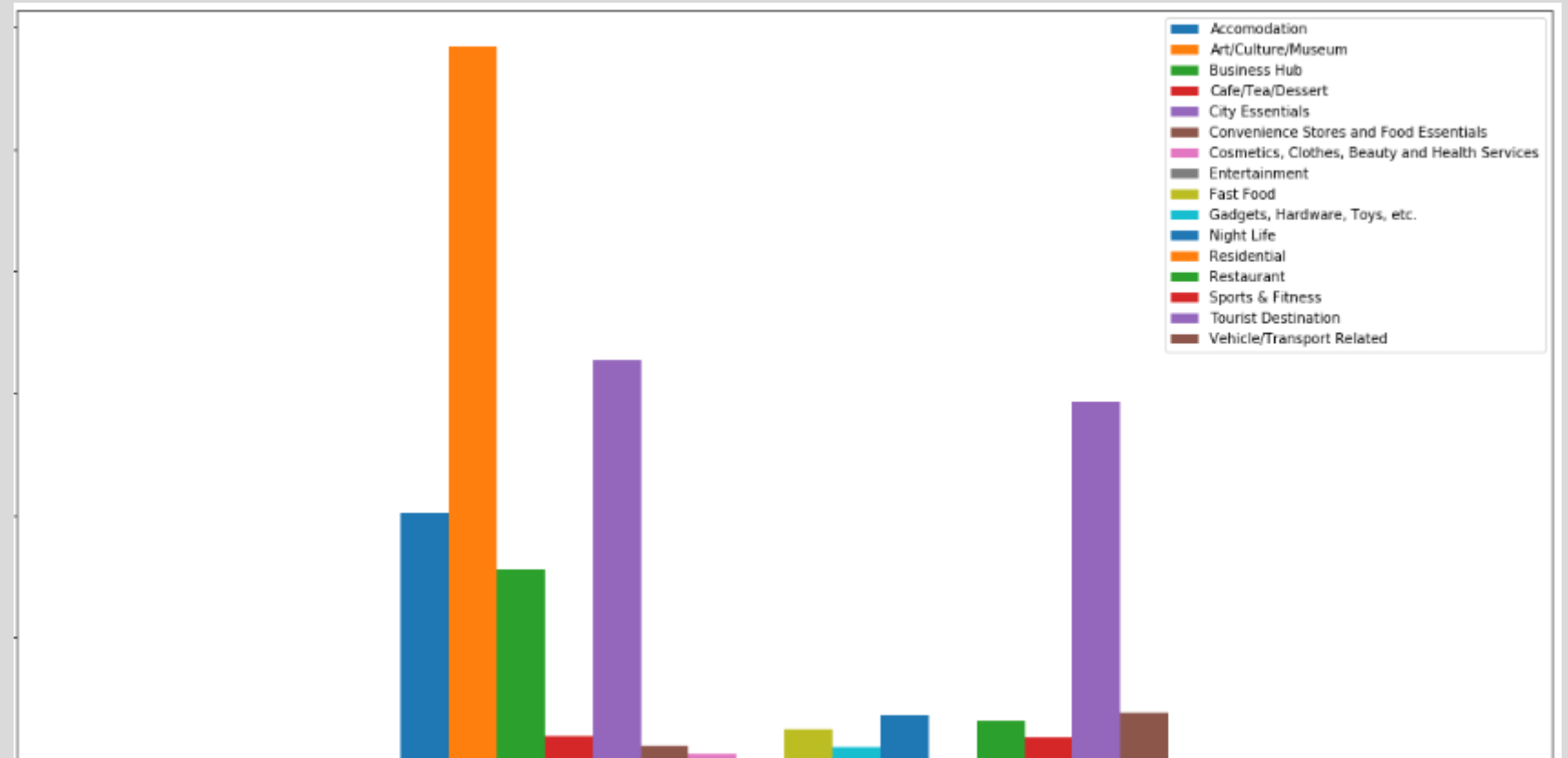
Cluster 2: Downtown Church, Hospital or School Zone

Notice the high value for the city essentials bar

Compare this graph to Cluster 7 which will be discussed later

Legend:
- Accomodation
- Art/Culture/Museum
- Business Hub
- Cafe/Tea/Dessert
- City Essentials
- Convenience Stores and Food Essentials
- Cosmetics, Clothes, Beauty and Health Services
- Entertainment
- Fast Food
- Gadgets, Hardware, Toys, etc.
- Night Life
- Residential
- Restaurant
- Sports & Fitness
- Tourist Destination
- Vehicle/Transport Related
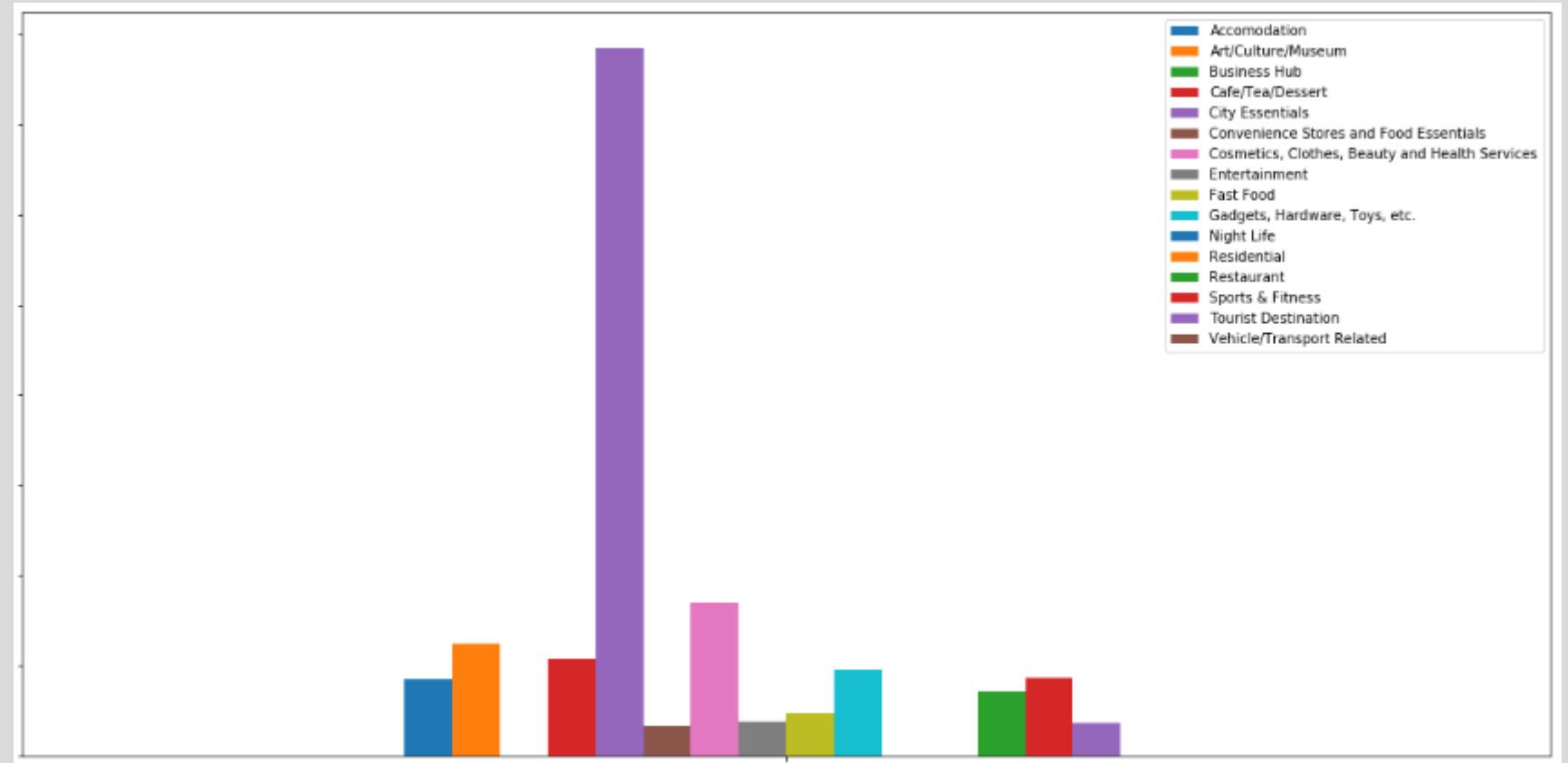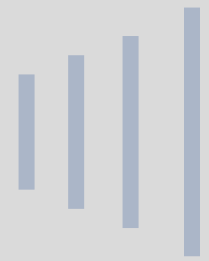
Cluster 3: City Center

Notice the high value for most of the bars especially the Vehicle/Transport related and Convenience Stores and Food essential venue types
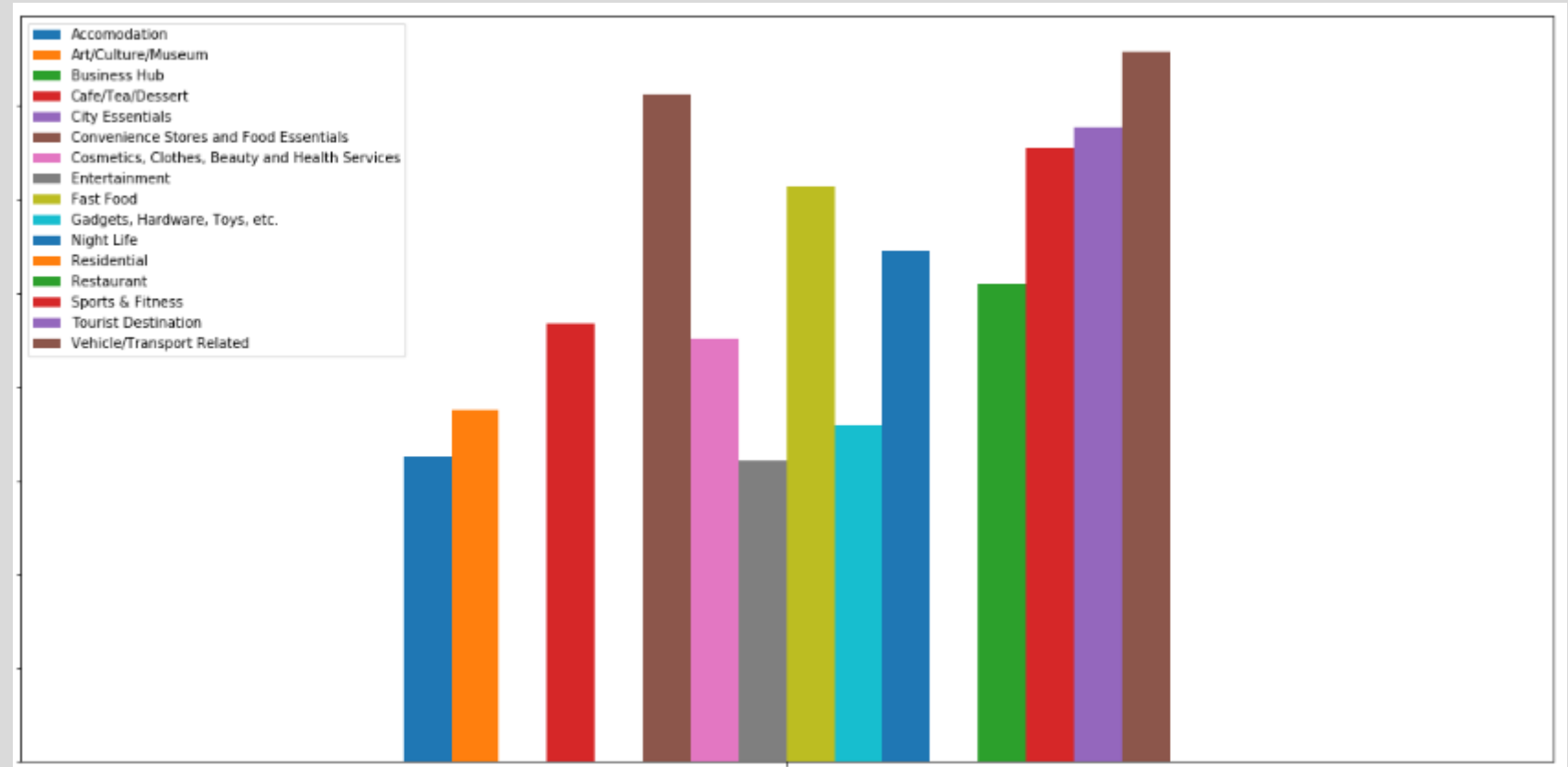
Cluster 4: Entertainment and Shopping District

Notice the high value for Entertainment and Cosmetics, Clothes, Beauty and Health services venue types

Legend:
- Accomodation
- Art/Culture/Museum
- Business Hub
- Cafe/Tea/Dessert
- City Essentials
- Convenience Stores and Food Essentials
- Cosmetics, Clothes, Beauty and Health Services
- Entertainment
- Fast Food
- Gadgets, Hardware, Toys, etc.
- Night Life
- Residential
- Restaurant
- Sports & Fitness
- Tourist Destination
- Vehicle/Transport Related
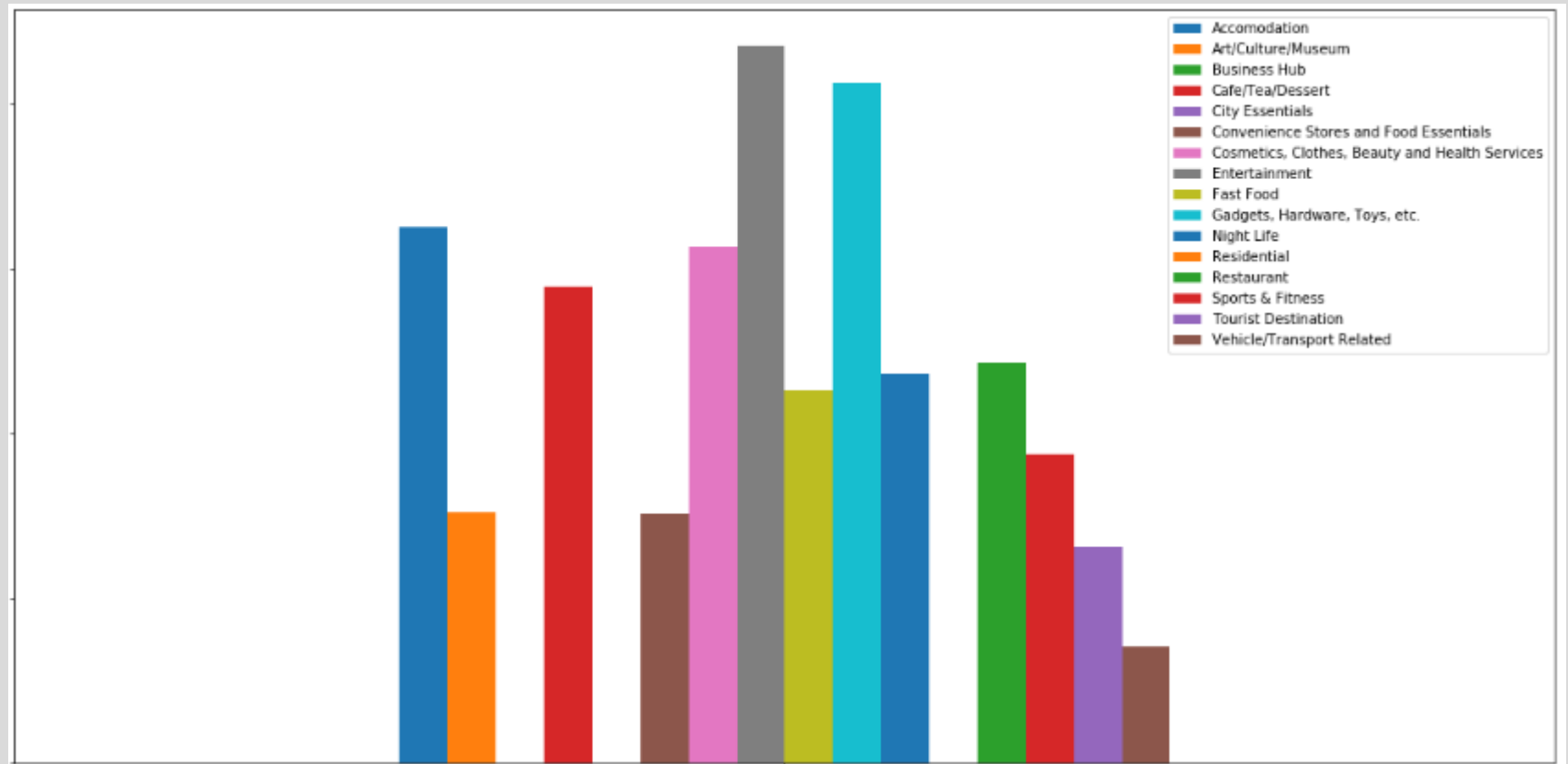
Cluster 5: Business District

Notice the high value for Business Hub Venue types

Legend:
- Accomodation
- Art/Culture/Museum
- Business Hub
- Cafe/Tea/Dessert
- City Essentials
- Convenience Stores and Food Essentials
- Cosmetics, Clothes, Beauty and Health Services
- Entertainment
- Fast Food
- Gadgets, Hardware, Toys, etc.
- Night Life
- Residential
- Restaurant
- Sports & Fitness
- Tourist Destination
- Vehicle/Transport Related
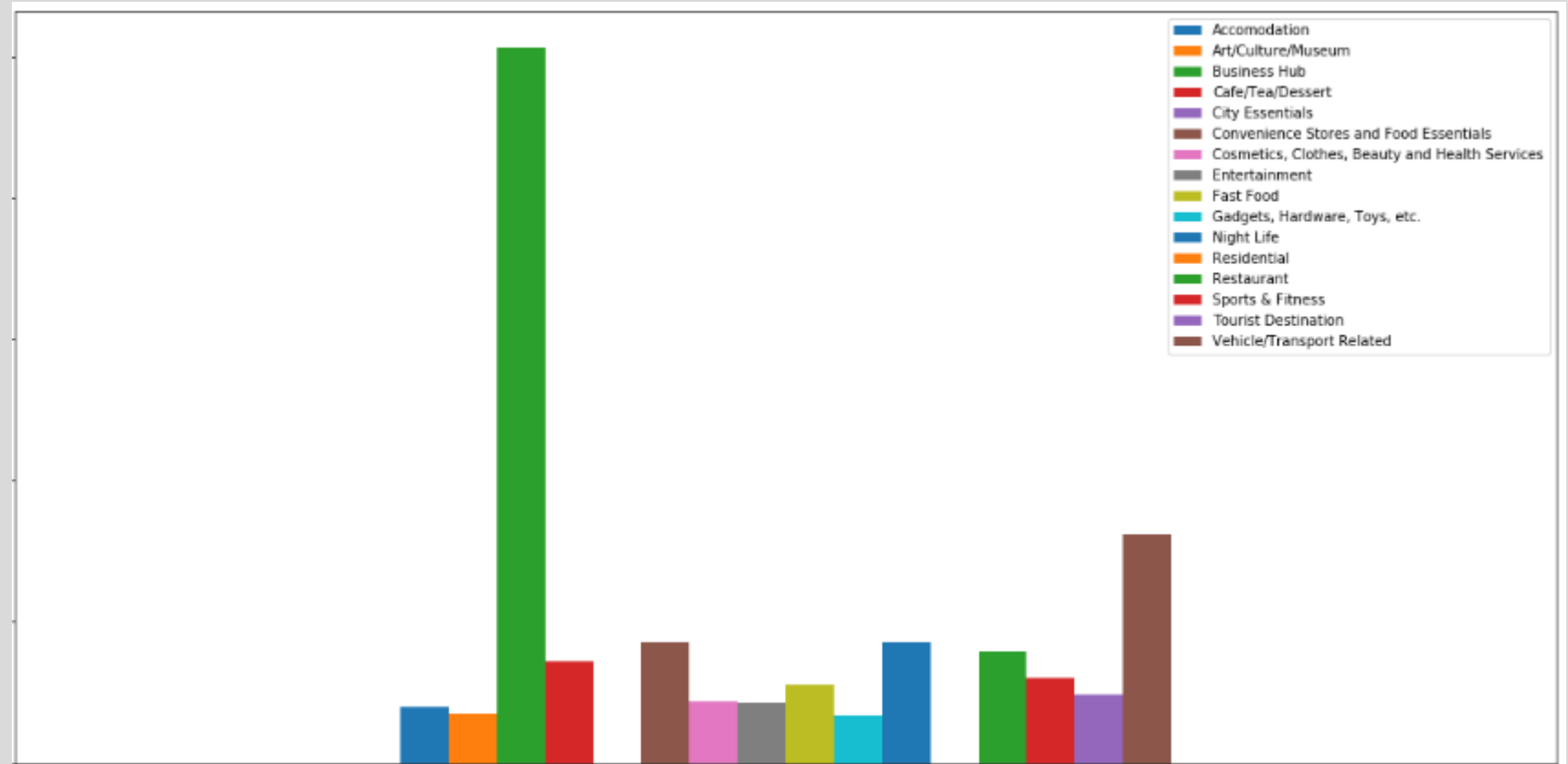
Cluster 6: Residential District

Notice the high value for residential venue types

Legend:
- Accomodation
- Art/Culture/Museum
- Business Hub
- Cafe/Tea/Dessert
- City Essentials
- Convenience Stores and Food Essentials
- Cosmetics, Clothes, Beauty and Health Services
- Entertainment
- Fast Food
- Gadgets, Hardware, Toys, etc.
- Night Life
- Residential
- Restaurant
- Sports & Fitness
- Tourist Destination
- Vehicle/Transport Related
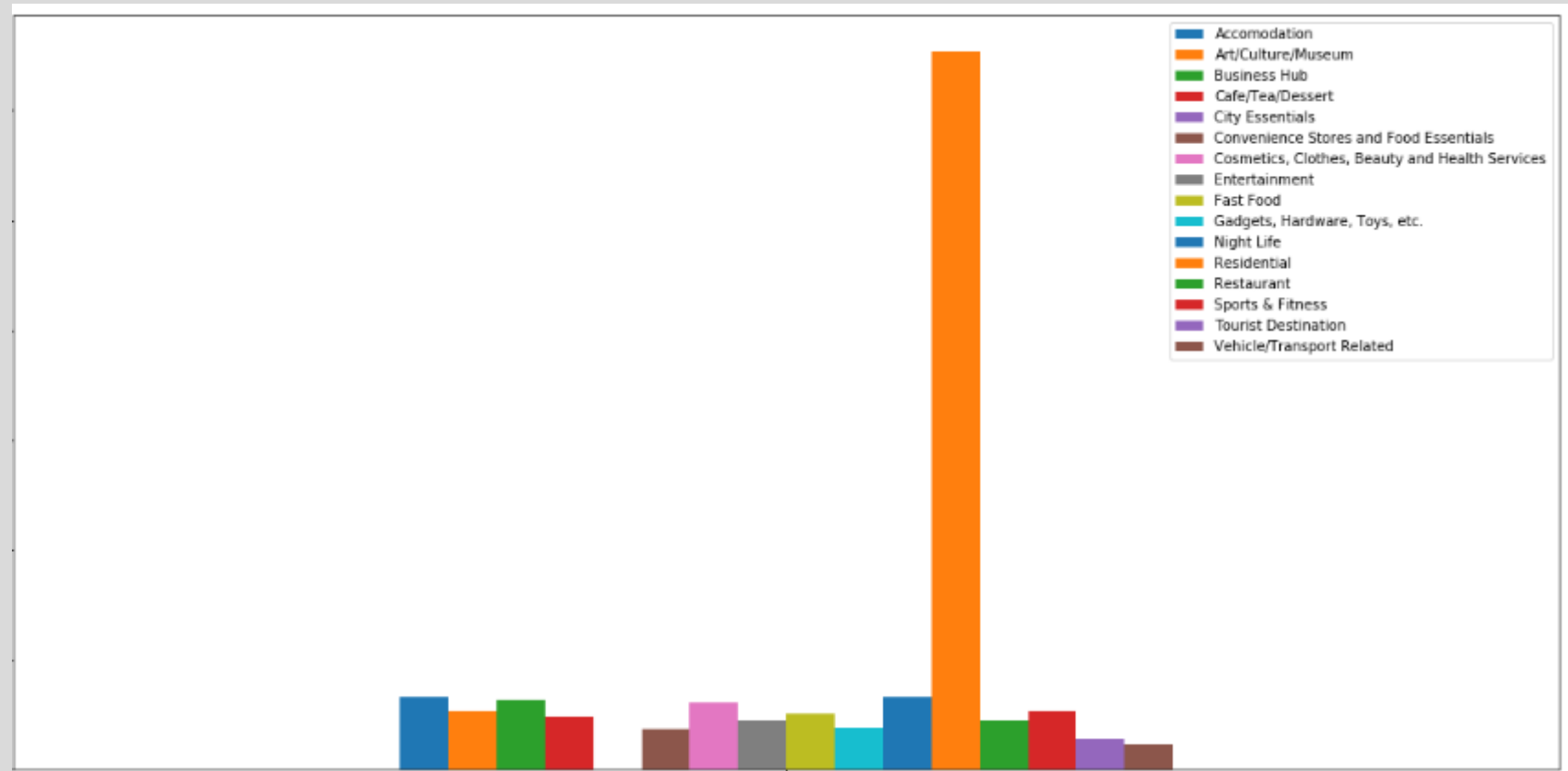
Cluster 7: Low Profile Church, Hospital or School Zone

Notice the high value for city essential venue types

Smaller value of the other venue types especially compared to Cluster 2

Legend:
- Accomodation
- Art/Culture/Museum
- Business Hub
- Cafe/Tea/Dessert
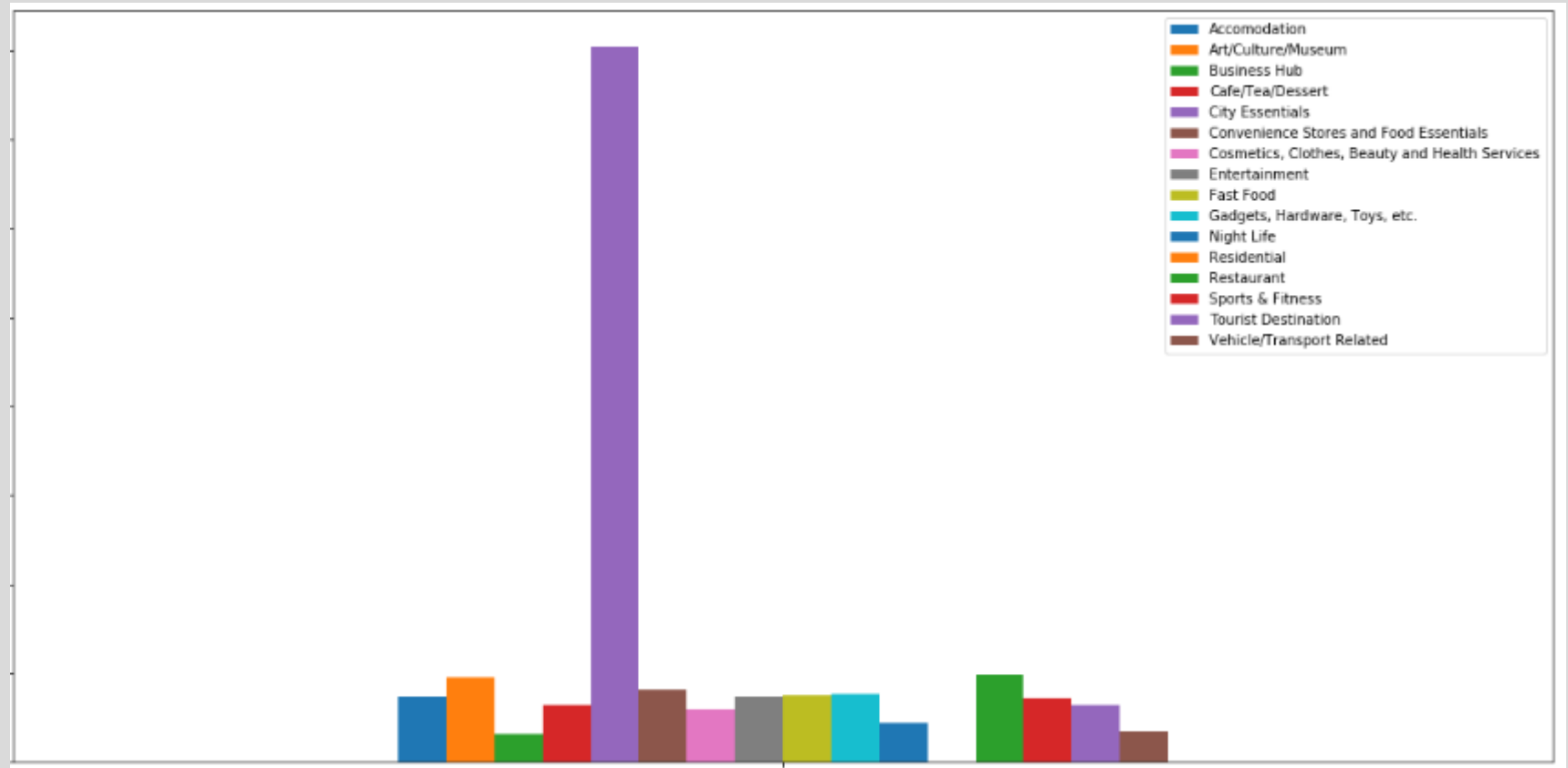- City Essentials
- Convenience Stores and Food Essentials
- Cosmetics, Clothes, Beauty and Health Services
- Entertainment
- Fast Food
- Gadgets, Hardware, Toys, etc.
- Night Life
- Residential
- Restaurant
- Sports & Fitness
- Tourist Destination
- Vehicle/Transport Related

Bar graph of all the Clusters

Better comparison between Cluster 2 (labeled 1) and Cluster 7 (labeled 6)

The magnitude matters in this case

Legend:
- Accomodation
- Art/Culture/Museum
- Business Hub
- Cafe/Tea/Dessert
- City Essentials
- Convenience Stores and Food Essentials
- Cosmetics, Clothes, Beauty and Health Services
- Entertainment
- Fast Food
- Gadgets, Hardware, Toys, etc.
- Night Life
- Residential
- Restaurant
- Sports & Fitness
- Tourist Destination
- Vehicle/Transport Related

Cluster Labels

Conclusion

Neighborhoods in Metro Manila is divided into 7 clusters

The information can be used by business owners and people who plans to migrate to the different neighborhoods

| Cluster	Description
1	Arts, Culture and Science District
2	Downtown Church, Hospital or School Zone
3	City Center
4	Entertainment and Shopping District
5	Business District
6	Residential District
7	Low Profile Church, Hospital or School Zone

Recommendations

- Add more datasets that would visualize the neighborhoods in Metro Manila better.
  - E.g. Real Estate Prices
  - Population Density and Demographics

- Use more official dataset other than scrapping it from websites
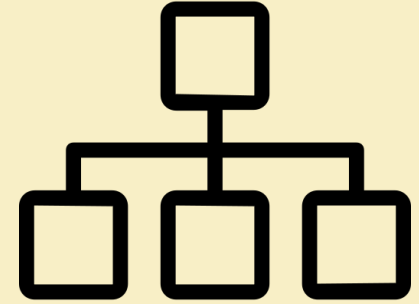
- Use other Clustering Models such as Hierarchical Clustering or DBSCAN Models