Contoso

THE BATTLE OF NEIGHBORHOOD - "SAN FRANCISCO"

JAYRAJSINH SISODIYA

INTRODUCTION

- **Background:** Best location is a huge concern for business owner when starting a new hotel business. Competitive analysis is an essential tactic for finding out what your competitors are doing and what kind of threat they present to your financial well-being.

- **Problem:** This project aims to find the best neighborhood for a hotel in San Francisco based on the competition with Airbnb properties. Exploring the neighborhoods of the region to find out 10 most common venues in every neighborhoods and cluster the neighborhoods based on K-means clustering algorithm.

- **Interest:** This project is for all those business owners who want to start their new hotel business in San Francisco and exploring the neighborhoods of San Francisco with common venues around.

Contoso

DATA ACQUISITION

- **Data Acquisition:** **The data acquired for the project is a combination of 2 data source:**

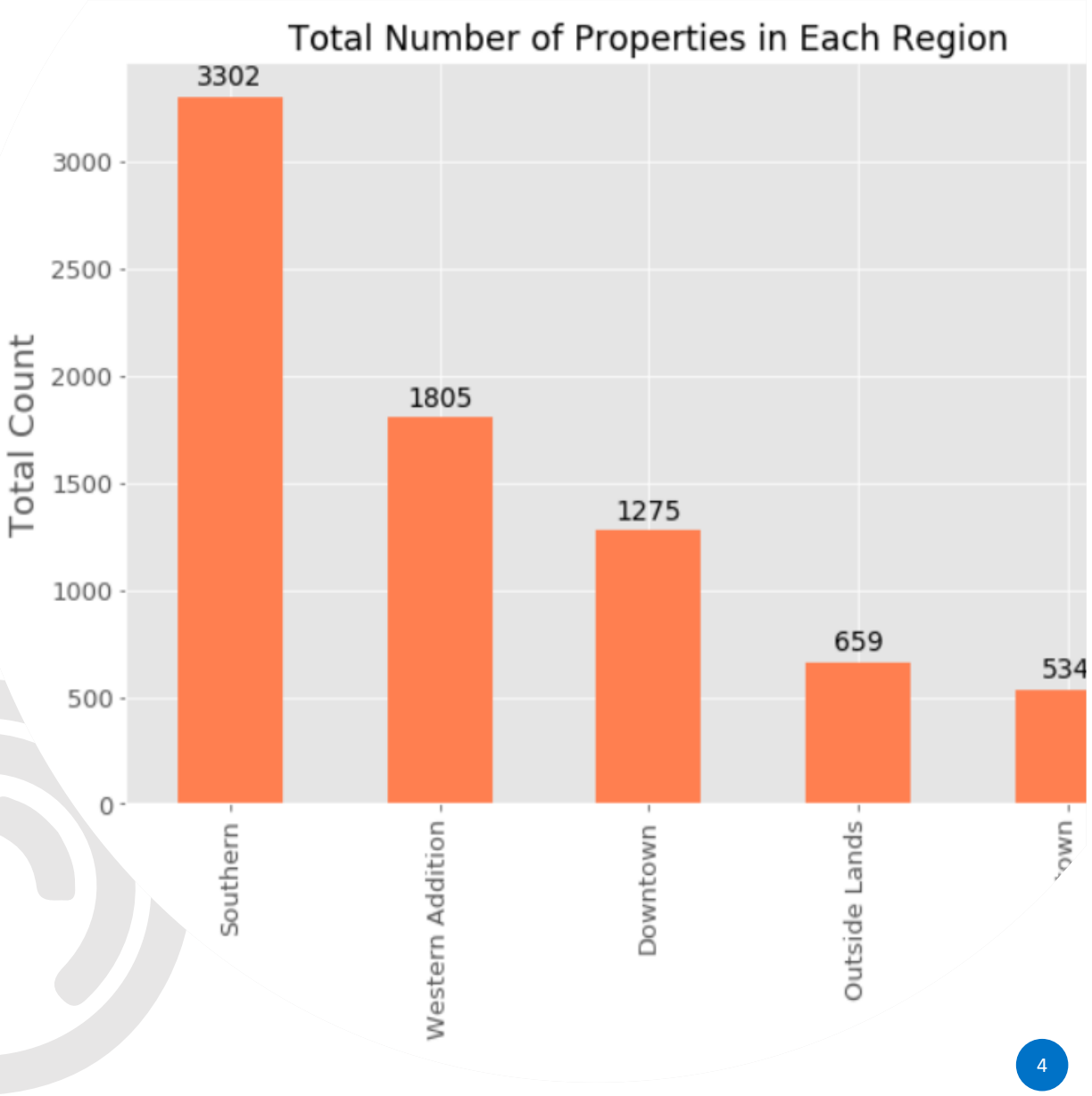
- **The first data source of the project uses a listing of San Francisco Airbnb properties that shows the Airbnb properties information in each region and neighborhood.**
- **The 2nd Dataset is the geographical co-ordinates data of San Francisco will be used for input from the Foursquare API which will be leveraged to provide venue information for the neighborhood.**

Contoso

METHODOLOGY

Exploratory Data Analysis

- As per June 2019, there are 7575 Airbnb properties in San Francisco that are in 5 different region and 55 neighborhoods.
- Highest number of properties are in “Southern” region followed by “Western Addition”, “Downtown”, “Outside Lands” & “North of Downtown”.



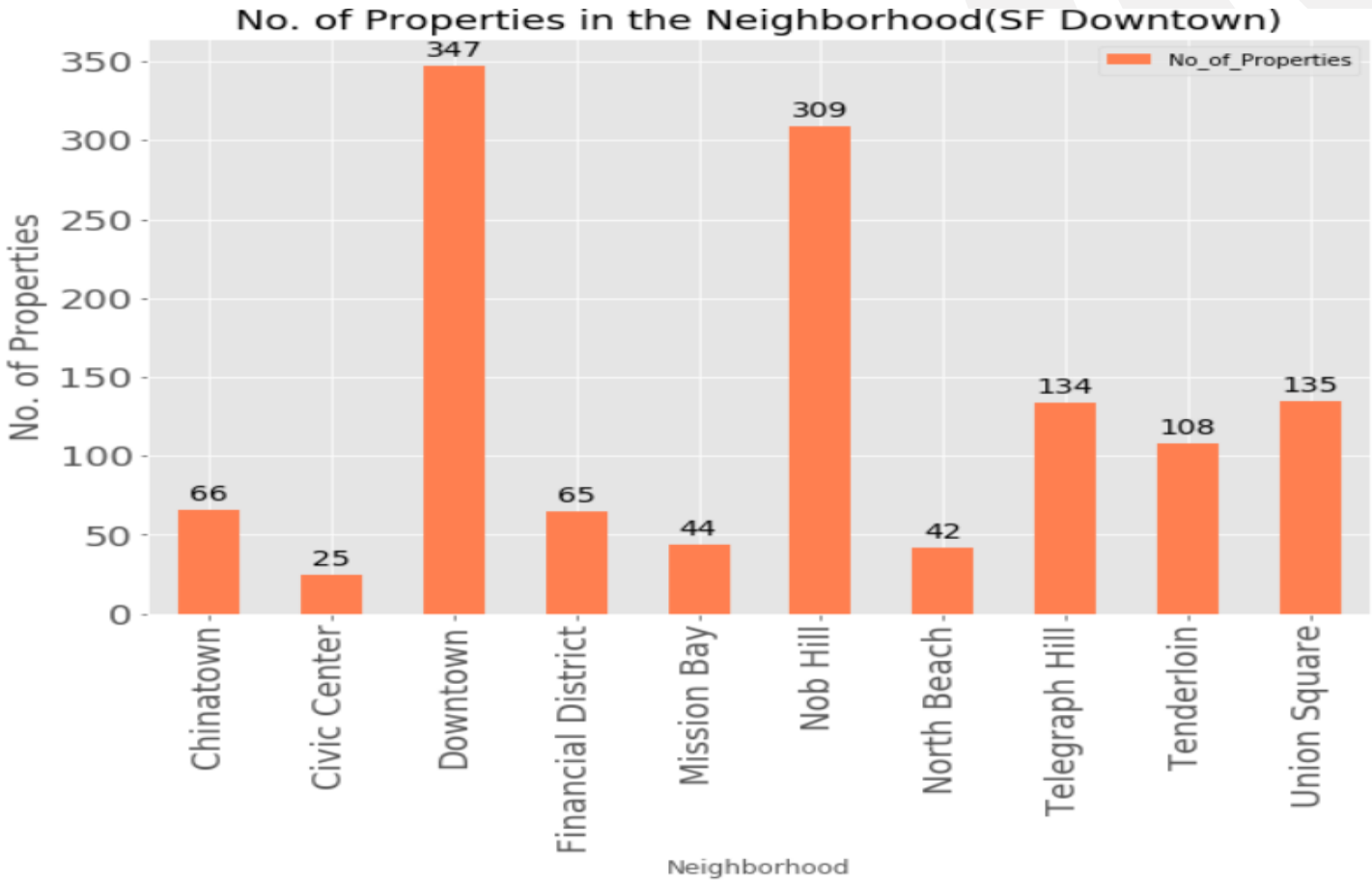
Contoso

PRIMARY FOCUS OF THE CLIENT

“As per our client's primary focus is to find a neighborhood that has a moderate number of Airbnb properties. The reason behind the condition is that the client wants a hotel location that neither has high competition from Airbnb nor he wants a hotel location with a smaller number of Airbnb properties where there is a low return of investment. Therefore, we will select “Downtown” region for our client where there is a moderate competition with 1275 Airbnb properties.”

Contoso

NO. OF PROPERTIES IN THE NEIGHBORHOOD (SF DOWNTOWN)



NEIGHBORHOODS IN SAN FRANCISCO DOWNTOWN

There are 9 neighborhoods in the downtown of San Francisco, and they are visualized using Folium Library through python.

MODELLING

- Using the final dataset that contains neighborhoods in San Francisco downtown with respective longitude and latitude. We can observe all the venues with in 500-meter radius of each neighborhood by using connecting dataset to Foursquare API.

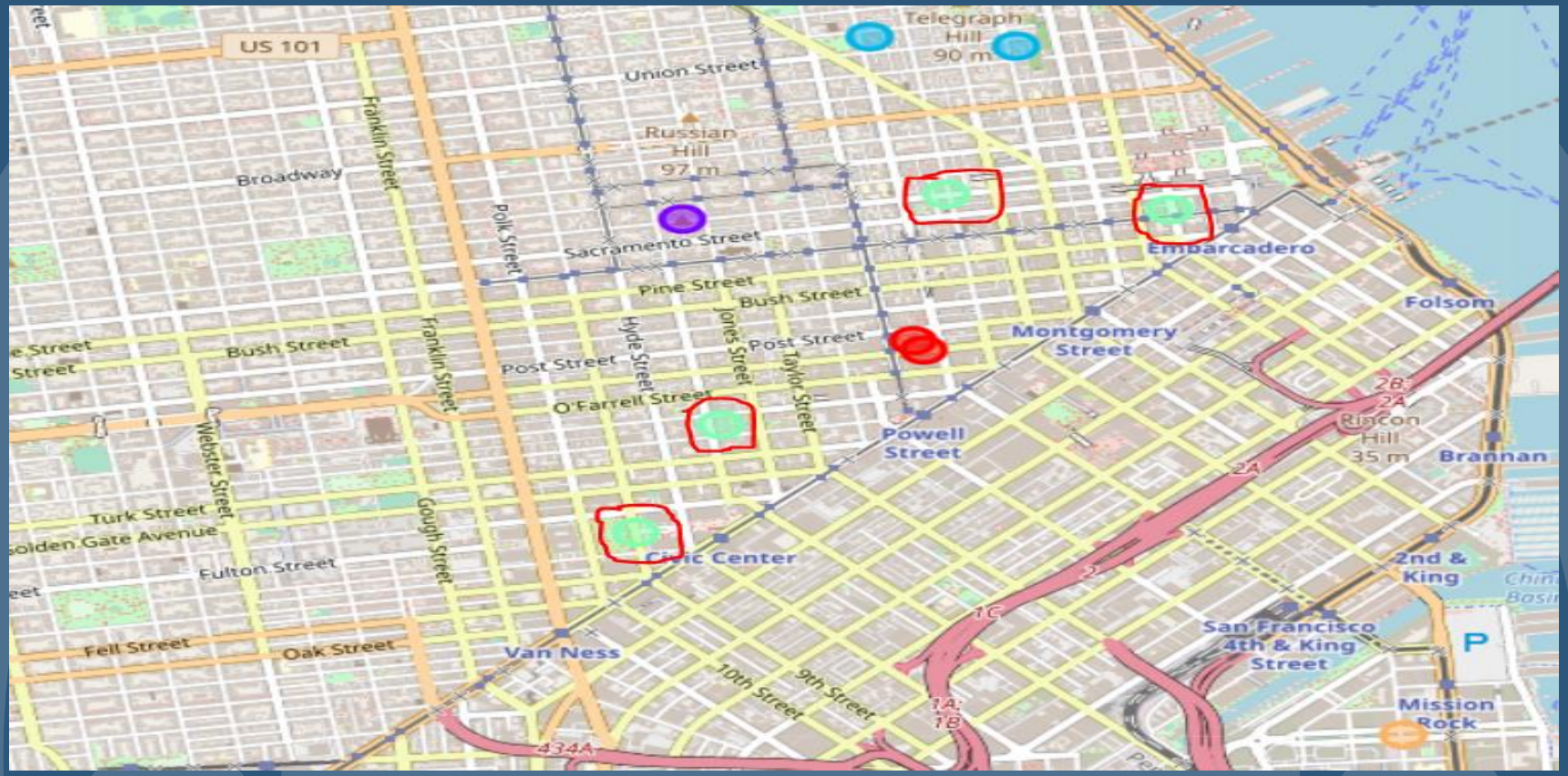
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Chinatown	37.794301	-122.406376	Blue Bottle Coffee	37.792771	-122.404833	Coffee Shop
1	Chinatown	37.794301	-122.406376	Hinodeya	37.794656	-122.404544	Ramen Restaurant
2	Chinatown	37.794301	-122.406376	Red Blossom Tea Company	37.794643	-122.406379	Tea Room
3	Chinatown	37.794301	-122.406376	Chapel Hill Coffee Co.	37.794041	-122.404247	Coffee Shop
4	Chinatown	37.794301	-122.406376	Mister Jiu's	37.793790	-122.406615	Chinese Restaurant

- To help business owners find similar neighborhood in the moderate competition, we will utilize K-means clustering algorithm which is a unsupervised machine learning algorithm that clusters data based on the predefined cluster size.
- We will cluster 9 neighborhoods into 5 cluster. The reason for K-means clustering is to cluster neighborhoods with similar venues together so that business owners can find location for the hotel based on the preferences.

Contoso

RESULTS

- After running K-means clustering, we can observe cluster created to see which neighborhoods are assigned to the clusters.



Cluster 1: Neighborhood (Downtown, Union Square)

	Neighborhood	Region	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	Downtown	Downtown	37.787514	-122.407159	0	Boutique	Jewelry Store	Clothing Store	Hotel	Theater
9	Union Square	Downtown	37.787936	-122.407517	0	Boutique	Hotel	Jewelry Store	Clothing Store	Theater

The cluster 1 has 2 neighborhoods out of 9. The cluster consist common venue such as Boutique, Jewelry Store, Hotel, Clothing Store, Jewelry Store, Theater.

Cluster 2: Neighborhood (Nob Hill)

	Neighborhood	Region	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
5	Nob Hill	Downtown	37.793262	-122.415249	1	Italian Restaurant	Hotel	Bar	Café	Coffee Shop

The cluster 2 has one neighborhood that consist venues such as Italian Restaurant, Hotel, Bar, Cafe, Coffee Shop.

Cluster 3: Neighborhood (North Beach, Telegraph Hill)

	Neighborhood	Region	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
6	North Beach	Downtown	37.801175	-122.409002	2	Italian Restaurant	Pizza Place	Bakery	Café	Cocktail Bar
7	Telegraph Hill	Downtown	37.800785	-122.404091	2	Italian Restaurant	Pizza Place	Cocktail Bar	Café	Coffee Shop

The cluster 3 has 2 neighborhoods out of 9 and cluster consist of venues such as Italian restaurant, Pizza place, bakery, cocktail bar and Cafe

Contoso

Cluster 4: Neighborhood (Chinatown, Civic Center, Financial District, Tenderloin)

	Neighborhood	Region	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Chinatown	Downtown	37.794301	-122.406376	3	Coffee Shop	Chinese Restaurant	Bakery	New American Restaurant	Men's Store
1	Civic Center	Downtown	37.779594	-122.416794	3	Café	Coffee Shop	Cocktail Bar	Vietnamese Restaurant	Beer Bar
3	Financial District	Downtown	37.793647	-122.398938	3	Coffee Shop	Food Truck	Café	Gym	Salad Place
8	Tenderloin	Downtown	37.784249	-122.413993	3	Vietnamese Restaurant	Coffee Shop	Thai Restaurant	Theater	Cocktail Bar

The cluster 4 has the highest number of neighborhoods in the cluster. It has 4 neighborhoods out of 9 and cluster consist of venues such as coffee shop, Cafe, Vietnamese restaurant, Chinese Restaurant, Food Truck, Bakery, gym, Beer Bar.

Cluster 4: Neighborhood (Mission Bay)

	Neighborhood	Region	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
4	Mission Bay	Downtown	37.770774	-122.391171	4	Food Truck	Gym	Coffee Shop	Pharmacy	Park

Cluster 5 has one neighborhood which has venues such as food truck, gym, coffee shop, pharmacy and park.

Contoso

DISCUSSION

- The aim of the project is to help business owners to find the best place in San Francisco downtown. The business owners can choose the location based on the most common venues around.
- As an example, if a business owner wants a hotel that has shopping venues around such as boutique, jewelry store, clothing store then they can choose a neighborhood from cluster 1.
- In cluster 2,3 & 4 most of the places are food places. Such as if business owner wants a hotel near Italian Restaurant, Hotel, Bar, Cafe, Coffee Shop then he/she would choose cluster 2.
- The neighborhood in cluster 3 has proximity to Italian restaurant, Pizza place, bakery, cocktail bar and Café.
- Cluster 4 that has the highest neighborhoods and scattered , the business owner will choose such neighborhood if the preference is in favor of coffee shop, Cafe, Vietnamese restaurant, Chinese Restaurant, Food Truck, Bakery, gym, Beer Bar.
- Business owner will choose cluster 5 where there is only one neighborhood if he/she wants a hotel near places such as food truck, gym, coffee shop, pharmacy and park.

Contoso

CONCLUSION

- This project aims to help business owners to have a better understanding of neighborhoods in comparison with most common places around neighborhood. It is essential to use the technology to gain advantage in business such as knowing more about location before starting the business in the region.
- In this project, the competition with Airbnb properties has been considered. The future scope of this project can include the competition with other hotel, price offered by other business owner, safety in the neighborhood.

Contoso

THANK YOU!

Contoso