The City Of San Diego CA

Introduction and Business problem:

Relocation is always a difficult task when we have to select between the cities. It is always a good idea to gather some information about the place you are planning to move and check if it is similar to your liking. Some initial survey greatly helps in taking the decision of whether should consider moving to that place or not. Everyone has a different criteria to define the best place they want to live in. for example- Best schools nearby, low crime rate, housing prices etc.

The city under consideration for analysis is San Diego CA. San Diego is a city in the U.S. state of California. It is in San Diego County, on the coast of the Pacific Ocean in Southern California, approximately 120 miles south of Los Angeles and immediately adjacent to the border with Mexico. San Diego has been called "the birthplace of California". San Diego is the eighth-largest city in the United States and second-largest in California.

The objective of this project is to help the target audience who are considering moving to San Diego and help them in deciding whether the city is similar to their liking or not and if it interests them enough to settle in .

Data:
Dataset 1: We are using the open data set for US cities from below website

1. https://simplemaps.com/data/us-cities

   The link contains data of all the cities of United States But for our project we have selected the data for San Diego city. It contains all the neighborhoods of San Diego and their latitude and longitude along with other details.

.0]:

	index	Neighborhood	County	Latitude	Longitude	population	density
0	15449	Jacumba Hot Springs	San Diego	32.6278	-116.1827	NaN	631.9
1	15513	Del Mar	San Diego	32.9633	-117.2627	4363.0	986.0
2	15557	Rainbow	San Diego	33.4098	-117.1394	NaN	189.6
3	15586	Winter Gardens	San Diego	32.8376	-116.9269	NaN	1428.0
4	15590	Rancho Santa Fe	San Diego	33.0240	-117.1992	NaN	357.9
5	15608	Fallbrook	San Diego	33.3693	-117.2258	NaN	2417.9
6	15619	Castle Park	San Diego	32.6103	-117.0675	NaN	6016.4
7	15637	Encinitas	San Diego	33.0492	-117.2613	63184.0	1296.0

This data will help us to plot the neighborhoods of the San Diego.

Dataset 2: The below data is for analyzing the demographics of San Diego

2. https://en.wikipedia.org/wiki/San_Diego

   Beautiful soup is used for scraping the below data from Wikipedia page.

	Racial composition	2010	1990	1970	1940
0	White	58.9%	67.1%	88.9%	96.9%
1	—Non-Hispanic	45.1%	58.7%	78.9%[99]	n/a
2	Black or African American	6.7%	9.4%	7.6%	2.0%
3	Hispanic or Latino (of any race)	28.8%	20.7%	10.7%[99]	n/a
4	Asian	15.9%	11.8%	2.2%	1.0%

Dataset 3: For crime analysis the data file is downloaded from below website:

3. https://www.sandag.org/index.asp?classid=14&subclassid=21&projectid=446&fuseaction=projects.detail

It contains the most recent 180 days of San Diego regional crime data as displayed on crimemapping.com. The data are in a comma separated value.

t[53]:

	Crime	Agency	Description	Date	Address	ZipCode	community
0	DRUGS/ALCOHOL VIOLATIONS	SHERIFF	SELL LIQUOR TO MINOR (M)	3/9/2019 16:40:00	300 S BLOCK SANTA FE AVENUE	92083.0	VISTA
1	THEFT/LARCENY	NATIONAL CITY	BURGLARY (SHOPLIFTING) (F)	3/22/2019 10:42:00	3000 BLOCK PLAZA BONITA ROAD	91950.0	NATIONAL CITY
2	THEFT/LARCENY	NATIONAL CITY	PETTY THEFT(All Other Larceny) (M)	2/6/2019 22:00:00	1300 E BLOCK 04TH STREET	91950.0	NATIONAL CITY
3	VEHICLE BREAK-IN/THEFT	NATIONAL CITY	PETTY THEFT(from Veh) (M)	2/27/2019 9:00:00	3000 BLOCK PLAZA BONITA ROAD	91950.0	NATIONAL CITY
4	ASSAULT	NATIONAL CITY	CAUSE HARM/DEATH OF ELDER /DEPENDENT ADULT (F)	3/3/2019 17:41:00	2000 BLOCK HIGHLAND AVENUE	91950.0	NATIONAL CITY

Methodology:

Below is the data analysis process followed:

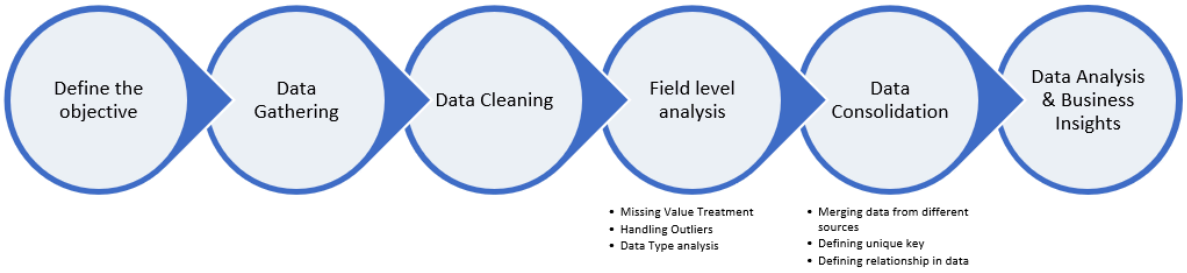


Image Ref: https://medium.com/datadriveninvestor/data-analytics-step-by-step-approach-757c6a0bd8a2

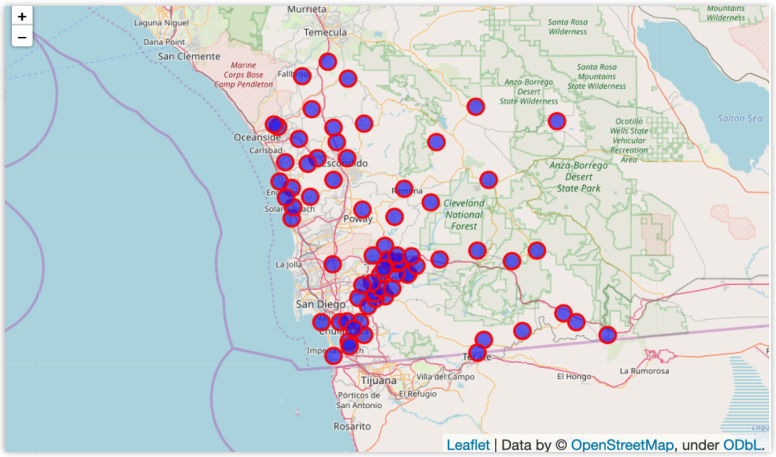
Data Cleaning : Removal of Nan values and selecting the required columns for analysis.

Exploratory Data Analysis: Once the data was collected, some preprocessing was required to get it in correct format. The data chosen was mainly related to the geography, demographics, crime rate and neighborhood of San Diego.

K-means clustering was used to explore the neighborhoods of San Diego which gave some interesting results about the popular venues in the area.

Results and Discussion:

Location of neighborhoods of San Diego:



Most Popular Venues of San Diego:

1. Restaurants: The restaurants that are most popular are mostly fast food joints and American restaurants.

Restaurents:

```
]: SD_merged.loc[SD_merged['Cluster Labels'] == 0, SD_merged.columns[[1] + list(range(5, 
]:
```

	County	density	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
15513	San Diego	986.0	0.0	American Restaurant	Italian Restaurant	Coffee Shop	Hotel Bar	New American Restaurant
15619	San Diego	6016.4	0.0	Fast Food Restaurant	Japanese Restaurant	Pizza Place	Mexican Restaurant	Sushi Restaurant
15637	San Diego	1296.0	0.0	Fast Food Restaurant	Mexican Restaurant	Sandwich Place	Pizza Place	Japanese Restaurant
15715	San Diego	1188.0	0.0	Park	Yoga Studio	Discount Store	Farm	Falafel Restaurant
15775	San Diego	1523.0	0.0	Burger Joint	Fast Food Restaurant	Bakery	Sandwich Place	Shopping Mall
15784	San Diego	3254.0	0.0	Taco Place	Auto Dealership	Mexican Restaurant	Convenience Store	Pub
15792	San Diego	2615.1	0.0	Pool	American Restaurant	Disc Golf	Farm	Falafel Restaurant
15828	San Diego	2097.0	0.0	Italian Restaurant	Coffee Shop	Yoga Studio	Farm	Falafel Restaurant

The area includes some Italian as well as Mexican restaurants. So, we can see there are many r

2. The area also includes some services and markets:

Services:

```
SD_merged.loc[SD_merged['Cluster Labels'] == 2, SD_merged.columns[[1] + lis
```

	County	density	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
15608	San Diego	2417.9	2.0	Home Service	Business Service	Yoga Studio	Discount Store	Farm
16932	San Diego	494.0	2.0	Business Service	Yoga Studio	Comic Shop	Farm	Falafel Restaurant

Markets:

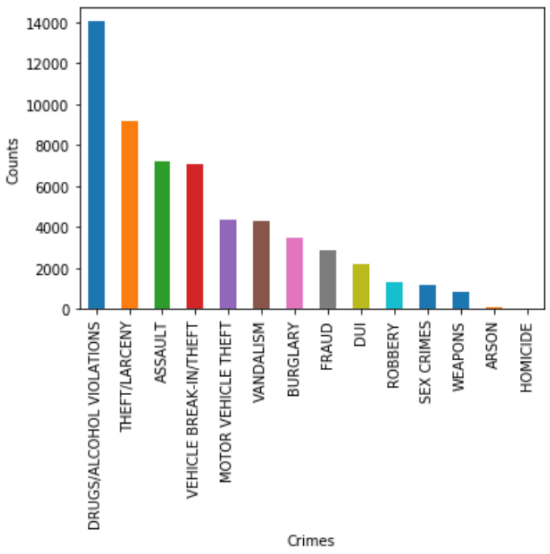
```
SD_merged.loc[SD_merged['Cluster Labels'] == 3, SD_merged.columns[[1] + lis
```

	County	density	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
16638	San Diego	1667.3	3.0	Flea Market	Yoga Studio	Farmers Market	Falafel Restaurant	Fabric Shop

Demographics:

	Racial composition	2010	1990	1970	1940
0	White	58.9%	67.1%	88.9%	96.9%
1	—Non-Hispanic	45.1%	58.7%	78.9%[99]	n/a
2	Black or African American	6.7%	9.4%	7.6%	2.0%
3	Hispanic or Latino (of any race)	28.8%	20.7%	10.7%[99]	n/a
4	Asian	15.9%	11.8%	2.2%	1.0%

Crime Statistics:



Conclusion:

San Diego has been called "the birthplace of California". San Diego is the eighth-largest city in the United States and second-largest in California. Over all the neighborhoods of san Diego has many interesting places including different types of restaurants, markets, gym and bars including stables for the people who enjoy riding. The demographics show that there is a good mix of racial composition since 2010. The crime data analysis can be little better but has shown decrease in total number of crimes over the years.

So, Overall San Diego is a pretty great city if you are considering a change of location.