



Urban Happiness

Ayushi Sharma

Introduction

- + Relocating to another place is a very difficult task.
- + What if a family man wants to move and is unaware of the environment around his selected neighborhood? This is where the project Urban Happiness comes in.
- + Urban Happiness presents a map to the user with the crime ranges of the locality and the clustered neighborhood based on the venues. It also outputs the closest neighborhood from the venue given by the user.
- + The map marks two neighborhoods that are close to the venue given by the user.
- + The project clusters the neighborhoods based on all the venues and the crime rate in that neighborhood into three clusters and shows it to the user via a map.

Data

+ Crime Rate csv file

IncidentNur	Category	Descript	DayOfWee	Date	Time	PdDistrict	Resolution	Address	X	Y	Location	PdId	
1.2E+08	WEAPON I	POSS OF P	Friday	01/29/201	11:00	SOUTHERN	ARREST, B	800 Block	-122.403	37.77542	(37.775420	1.2E+13	
1.2E+08	WEAPON I	FIREARM,	Friday	01/29/201	11:00	SOUTHERN	ARREST, B	800 Block	-122.403	37.77542	(37.775420	1.2E+13	
1.41E+08	WARRANT	WARRANT	Monday	04/25/201	14:59	BAYVIEW	ARREST, B	KEITH ST /	-122.389	37.72998	(37.729980	1.41E+13	
1.6E+08	NON-CRIM	LOST PROF	Tuesday	#####	23:50	TENDERLO	NONE	JONES ST /	-122.413	37.78579	(37.785788	1.6E+13	
1.6E+08	NON-CRIM	LOST PROF	Friday	#####	00:30	MISSION	NONE	16TH ST /	-122.42	37.76505	(37.765050	1.6E+13	
1.6E+08	ASSAULT	BATTERY	Friday	#####	21:35	NORTHERN	NONE	1700 Block	-122.426	37.78802	(37.788018	1.6E+13	
1.6E+08	OTHER OF	PAROLE VI	Saturday	#####	00:04	SOUTHERN	ARREST, B	MARY ST /	-122.406	37.78088	(37.780878	1.6E+13	
1.6E+08	NON-CRIM	FIRE REPO	Saturday	#####	01:02	TENDERLO	NONE	200 Block	-122.412	37.78398	(37.783980	1.6E+13	
1.6E+08	WARRANT	WARRANT	Saturday	#####	12:21	SOUTHERN	ARREST, B	4TH ST / B	-122.393	37.77579	(37.775787	1.6E+13	
1.6E+08	MISSING P	FOUND PE	Friday	#####	10:06	BAYVIEW	NONE	100 Block	-122.387	37.72097	(37.720960	1.6E+13	
1.6E+08	LARCENY/	ATTEMPT	Friday	01/29/201	22:30	TARAVAL	NONE	1200 Block	-122.477	37.76448	(37.764478	1.6E+13	
1.6E+08	NON-CRIM	AIDED CAS	Saturday	#####	13:30	TARAVAL	NONE	2200 Block	-122.478	37.74574	(37.745738	1.6E+13	
1.6E+08	OTHER OF	RESISTING	Monday	01/25/201	23:20	BAYVIEW	ARREST, B	200 Block	-122.377	37.7357	(37.735697	1.6E+13	
1.41E+08	ASSAULT	AGGRAVAT	Thursday	09/15/201	07:40	INGLESIDE	ARREST, B	SILVER AV	-122.432	37.72927	(37.729270	1.41E+13	

+ San Francisco GeoJson file.

```
{
  "type": "FeatureCollection",
  "crs": {
    "type": "name",
    "properties": {
      "name": "urn:ogc:def:crs:OGC:1.3:CRS84"
    }
  },
  "features": [{
    "type": "Feature",
    "properties": {
      "OBJECTID": 1,
      "DISTRICT": "CENTRAL",
      "COMPANY": "A"
    },
    "geometry": {
      "type": "Polygon",
      "coordinates": [
        [
          [-122.40532134644249, 37.806867516866724],
          [-122.40440122046421, 37.80885380837723],
          [-122.40438743872008, 37.80886519707406],
          [-122.40436730880846, 37.808873066044306],
          [-122.40532134644249, 37.806867516866724]
        ]
      ]
    }
  }]
}
```


+ Postal Codes

```
df3['Pincode']=pincode
```

Pincodes of the neighbourhood added to the dataframe

```
df3
```

```
5]:
```

	Neighbourhood	Count	Pincode
0	BAYVIEW	14303	94124
1	CENTRAL	17666	94104
2	INGLESIDE	11594	94112
3	MISSION	19503	94114
4	NORTHERN	20100	94109
5	PARK	8699	94117
6	RICHMOND	8922	94121
7	SOUTHERN	28445	94105
8	TARAVAL	11325	94116
9	TENDERLOIN	9942	94102

+FourSquare API

- + The foursquare API is used to fetch the list of venues that are close to the given latitude and the longitude. The API uses the client ID and the Client Secret to fetch the details.
- + The response can then be converted to a pandas data frame and then be used for further operations.

```
: #venues dataframe for each location  
print(s_venues.shape)  
s_venues.head()
```

(177, 7)

27]:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	BAYVIEW	37.7309	-122.3886	Bayview Hunters Point YMCA	37.731851	-122.389733	Gym
1	BAYVIEW	37.7309	-122.3886	Foodway Liquors	37.730519	-122.388617	Liquor Store
2	BAYVIEW	37.7309	-122.3886	Palou And Lane 23 44 Bus Stop	37.732858	-122.388903	Bus Station
3	CENTRAL	37.7915	-122.4018	Pushkin	37.790943	-122.403877	Russian Restaurant
4	CENTRAL	37.7915	-122.4018	Blue Bottle Coffee	37.791320	-122.400983	Coffee Shop

Methodology

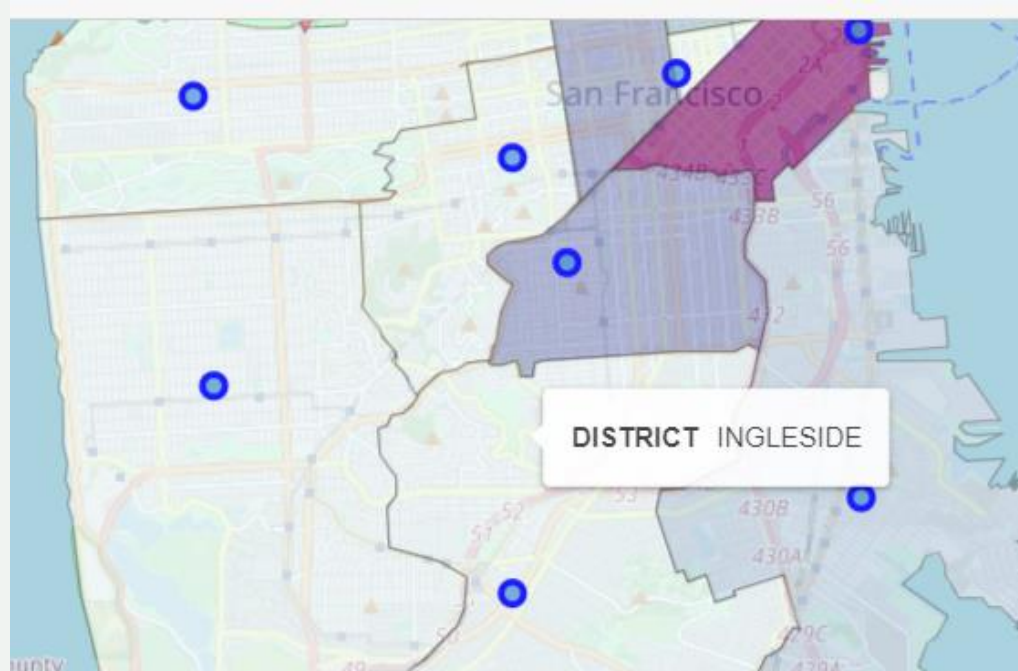
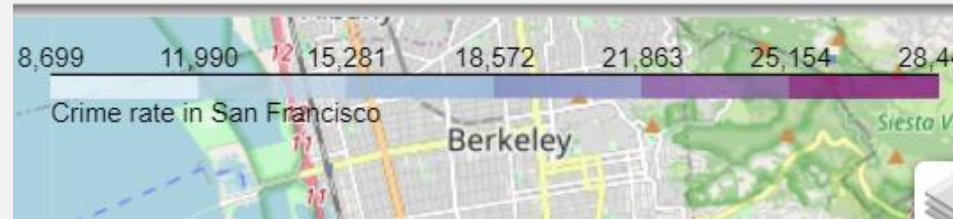
+ Data Pre-processing

- + The data from the csv file is processed so that the data is stored in the data frame with neighborhood, count as the columns.
- + One Hot Encoding is done on the foursquare response.
- + The data is then response is added to get the below data frame.
- + Each column is pre-processed and scaled down to a value between 0 and 1.

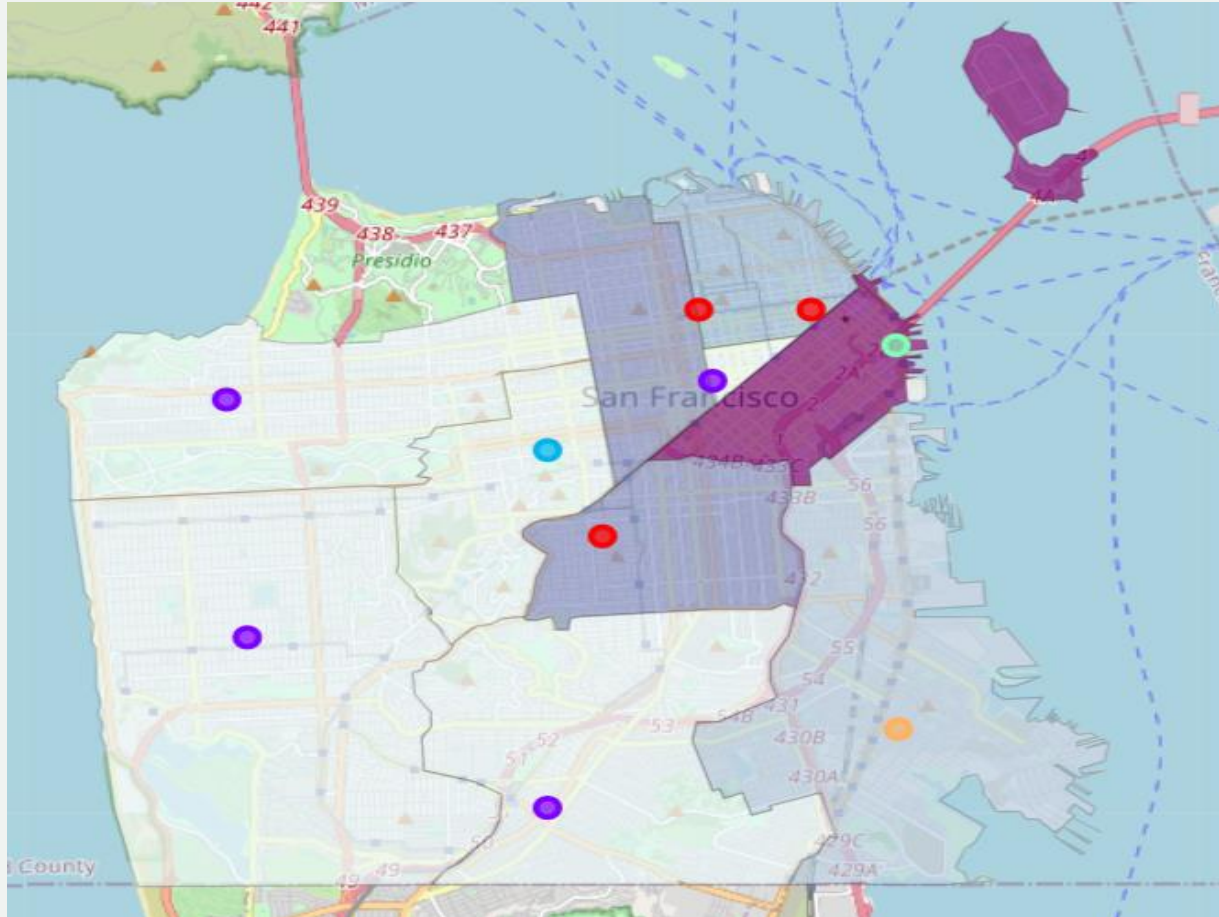
	Neighbourhood	Crime rate	Adult Boutique	American Restaurant	Art Gallery	Art Museum	Asian Restaurant	Bakery	Bar	Beer Bar
0	BAYVIEW	0.283804	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
1	CENTRAL	0.454117	0.000000	0.000000	0.00	0.000000	0.000000	0.021739	0.000000	0.000000
2	INGLESIDE	0.146612	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
3	MISSION	0.547149	0.000000	0.000000	0.05	0.000000	0.000000	0.000000	0.000000	0.000000
4	NORTHERN	0.577383	0.027778	0.027778	0.00	0.000000	0.000000	0.027778	0.111111	0.027778
5	PARK	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
6	RICHMOND	0.011293	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.200000	0.000000
7	SOUTHERN	1.000000	0.000000	0.071429	0.00	0.000000	0.000000	0.000000	0.000000	0.000000

+ Data Visualization

- + 1. Chloropleth map is shown to view crime rate



+ 2. Coloured markers for different clusters.

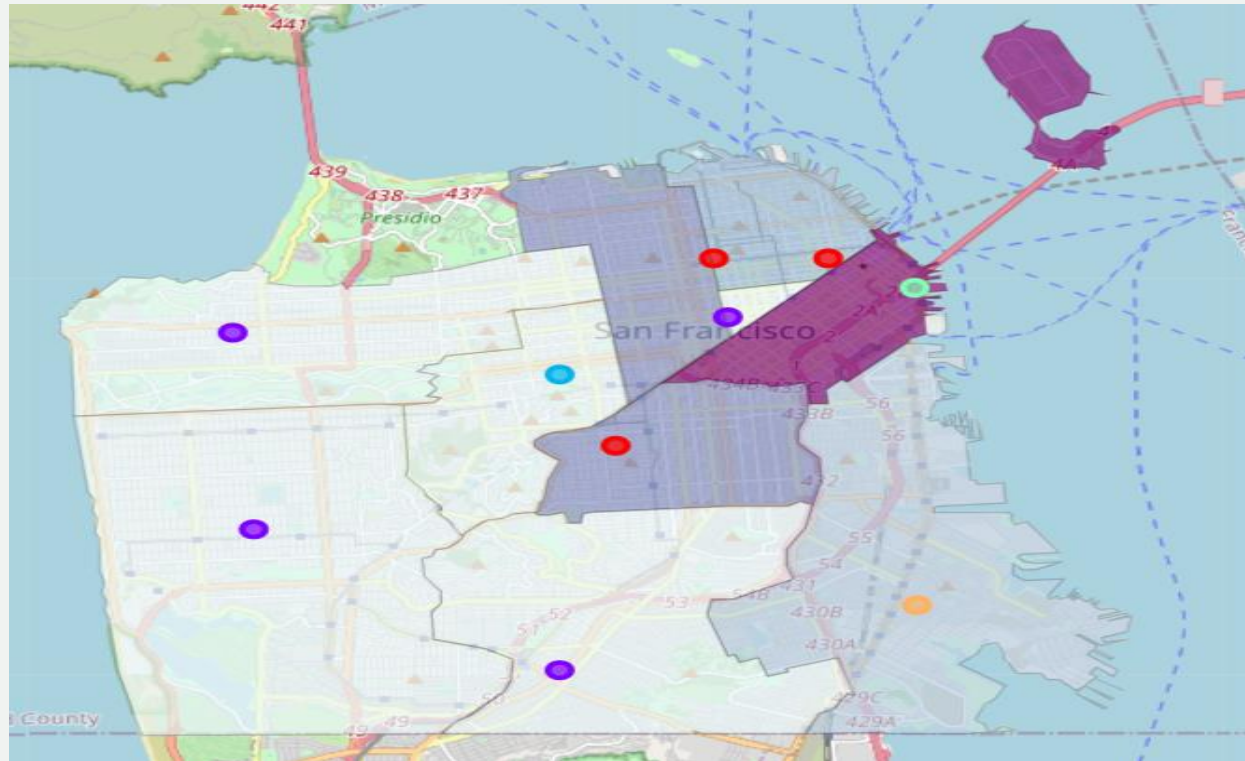


- + Marking the closest neighbourhood to the given venue.
- + Two nearest neighbourhood marked in the map are shown using the yellow line.



+Machine Learning

- + K-Means clustering is used to cluster the neighborhoods. Since the label in which the neighborhoods are not known classification algorithm cannot be used.
- + The output gives 5 clusters which are marked on the map.



Results

- + The user is shown the common venues of a location
- + The user is also shown the closest neighbourhood to the venue they selected.
- + The future scope of this project can also be to design a front end and show the users the most favourable neighbourhoods that the user can live in.



Thank You!