**Index:**

**Content Page no’s**

1.Abstract……………………………………………………………………..2

2.PreviousWorks……………………………………………………………...3

3.OurApproach……………………………………………………………….3

4.SoftwareRequirement………………………………………………………4

5.Design………………………………………………………………………4

6.Steps………………………………………………………………………...5

7.Code………………………………………………………………………..6-9

8.Output………………………………………………………………….......9-14

9.Chronic offender point system……………………………………..………15

10.FutureScope……………………………………………………………….16

11.Conclusion………………………………………………………………...16

**Abstract:**

There has been enormous increase in the crime rate last few years, especially in urban places of India . It is estimated that about **80.7%** of population in USA lives in cities and by the year 2030 about 60% of world population might reach the cities of their respective nations . So it will be a major task for Police departments and Government to keep a check on crimes happening in the state/country.

One effective way of reducing the crime rate is “**Predicting the Crimes**” by studying the link and similarities between them. As once said “*Prevention is better than Cure*”, predicting the crimes efficiently before happening will ensure maximum safety and security of citizens. It also makes **E-Governance** easy.

So here we use statistical prediction methods which identify the likely targets to prevent crime events. The techniques which we are going to use are done with the help of historic information,social media etc.

**Broken Window Theory** was the base for this Crime Prediction.

* **Previous Works:**

**LAPD**  has been implementing “**Predictive Policing**” past few years and has been successful in decreasing the crime rate in few situations.

States like Chicago , New York , Brisbane , LA etc. are using “***Operation Laser***” in their respective states and predicted few major crimes before happening and stopped them.

“**Chronic Offender Points System**” which is playing an important role in Criminal catching has its advantages and Disadvantages.It uses “**Point System**” which sometimes fail in catching the right person which is a big problem.

**Our Approach**

* We want to offer few amendents for “Chronic Offender System” with the help of proper analysis and algorithms.
* Our major task is “Predicting the crimes based on interesting Patterns obtained from previous crimes”
* Our work involves Clustering , Association and Classification techniques.

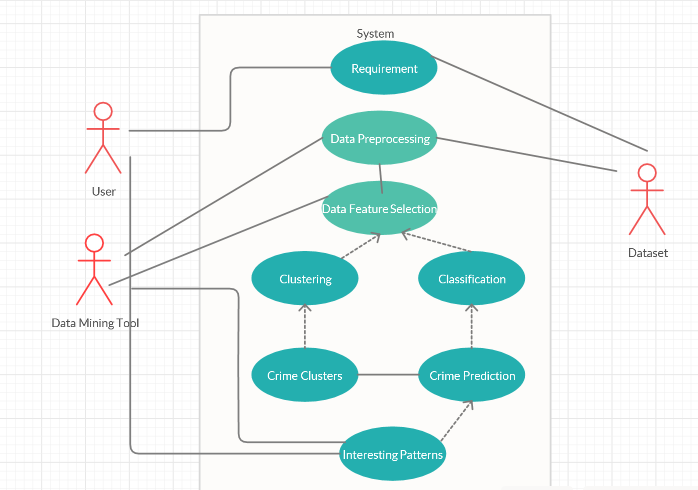
**Software Requirement:**

1.Spyder

2.Python3

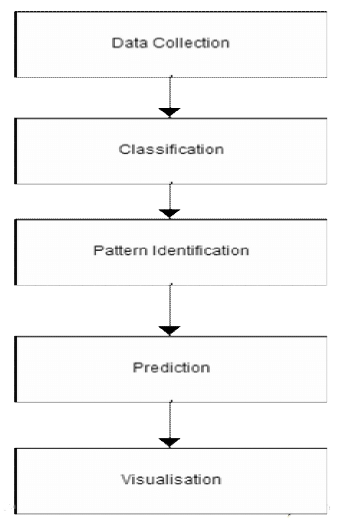
3.R(optional)

**Design**

****

Use case diagram

**Steps:**

****

**Code:**

# for some basic operations

import numpy as np

import pandas as pd

# for visualizations

import matplotlib.pyplot as plt

import seaborn as sns

import folium

import squarify

# for providing path

import os

print(os.listdir(r"C:\Users\Pawan\Downloads\sanfranciso-crime-dataset"))

# reading the dataset

data = pd.read\_csv(r'C:\Users\Pawan\Downloads\sanfranciso-crime-dataset\sf crime.csv')

# check the shape of the data

print(data.shape)

# checking the head of the data

print(data.head())

# describing the data

print(data.describe())

# checking if there are any null values

print(data.isnull().sum())

# filling the missing value in PdDistrict using the mode values

print(data['PdDistrict'].fillna(data['PdDistrict'].mode()[0], inplace = True))

print(data.isnull().any().any())

# different categories of crime

plt.rcParams['figure.figsize'] = (20, 9)

plt.style.use('dark\_background')

sns.countplot(data['Category'], palette = 'gnuplot')

plt.title('Major Crimes in Sanfrancisco', fontweight = 30, fontsize = 20)

plt.xticks(rotation = 90)

plt.show()

# plotting a tree map

y = data['Category'].value\_counts().head(25)

plt.rcParams['figure.figsize'] = (15, 15)

plt.style.use('fivethirtyeight')

color = plt.cm.magma(np.linspace(0, 1, 15))

squarify.plot(sizes = y.values, label = y.index, alpha=.8, color = color)

plt.title('Tree Map for Top 25 Crimes', fontsize = 20)

plt.axis('off')

plt.show()

# Regions with count of crimes

plt.rcParams['figure.figsize'] = (20, 9)

plt.style.use('seaborn')

color = plt.cm.spring(np.linspace(0, 1, 15))

data['PdDistrict'].value\_counts().plot.bar(color = color, figsize = (15, 10))

plt.title('District with Most Crime',fontsize = 30)

plt.xticks(rotation = 90)

plt.show()

# Regions with count of crimes

plt.rcParams['figure.figsize'] = (20, 9)

plt.style.use('seaborn')

color = plt.cm.ocean(np.linspace(0, 1, 15))

data['Address'].value\_counts().head(15).plot.bar(color = color, figsize = (15, 10))

plt.title('Top 15 Regions in Crime',fontsize = 20)

plt.xticks(rotation = 90)

plt.show()

# Regions with count of crimes

plt.style.use('seaborn')

'''data['DayOfWeek'].value\_counts().head(15).plot.pie(figsize = (15, 8), explode = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1))

plt.title('Crime count on each day',fontsize = 20)

plt.xticks(rotation = 90)

plt.show()'''

# Regions with count of crimes

plt.style.use('seaborn')

color = plt.cm.winter(np.linspace(0, 10, 20))

data['Resolution'].value\_counts().plot.bar(color = color, figsize = (15, 8))

plt.title('Resolutions for Crime',fontsize = 20)

plt.xticks(rotation = 90)

plt.show()

data['Date'] = pd.to\_datetime(data['Date'])

data['Month'] = data['Date'].dt.month

plt.style.use('fivethirtyeight')

plt.rcParams['figure.figsize'] = (15, 8)

sns.countplot(data['Month'], palette = 'autumn',)

plt.title('Crimes in each Months', fontsize = 20)

plt.show()

# checking the time at which crime occurs mostly

'''import warnings

warnings.filterwarnings('ignore')'''

color = plt.cm.twilight(np.linspace(0, 5, 100))

data['Time'].value\_counts().head(20).plot.bar(color = color, figsize = (15, 9))

plt.title('Distribution of crime over the day', fontsize = 20)

plt.show()

df = pd.crosstab(data['Category'], data['PdDistrict'])

color = plt.cm.Greys(np.linspace(0, 1, 10))

df.div(df.sum(1).astype(float), axis = 0).plot.bar(stacked = True, color = color, figsize = (18, 12))

plt.title('District vs Category of Crime', fontweight = 30, fontsize = 20)

plt.xticks(rotation = 90)

plt.show()

t = data.PdDistrict.value\_counts()

table = pd.DataFrame(data=t.values, index=t.index, columns=['Count'])

table = table.reindex(["CENTRAL", "NORTHERN", "PARK", "SOUTHERN", "MISSION", "TENDERLOIN", "RICHMOND", "TARAVAL", "INGLESIDE", "BAYVIEW"])

table = table.reset\_index()

table.rename({'index': 'Neighborhood'}, axis='columns', inplace=True)

table

gjson = r'https://cocl.us/sanfran\_geojson'

'''sf\_map = folium.Map(location = [37.77, -122.42], zoom\_start = 12)

#generate map

sf\_map.choropleth(

geo\_data=gjson,

data=table,

columns=['Neighborhood', 'Count'],

key\_on='feature.properties.DISTRICT',

fill\_color='YlOrRd',

fill\_opacity=0.7,

line\_opacity=0.2,

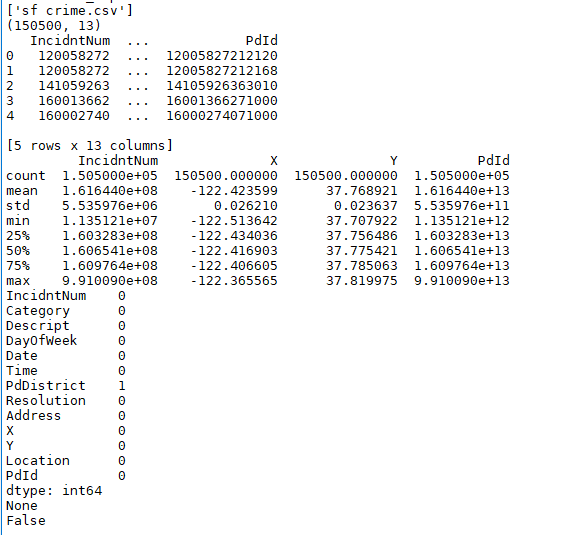
legend\_name='Crime Rate in San Francisco'

)

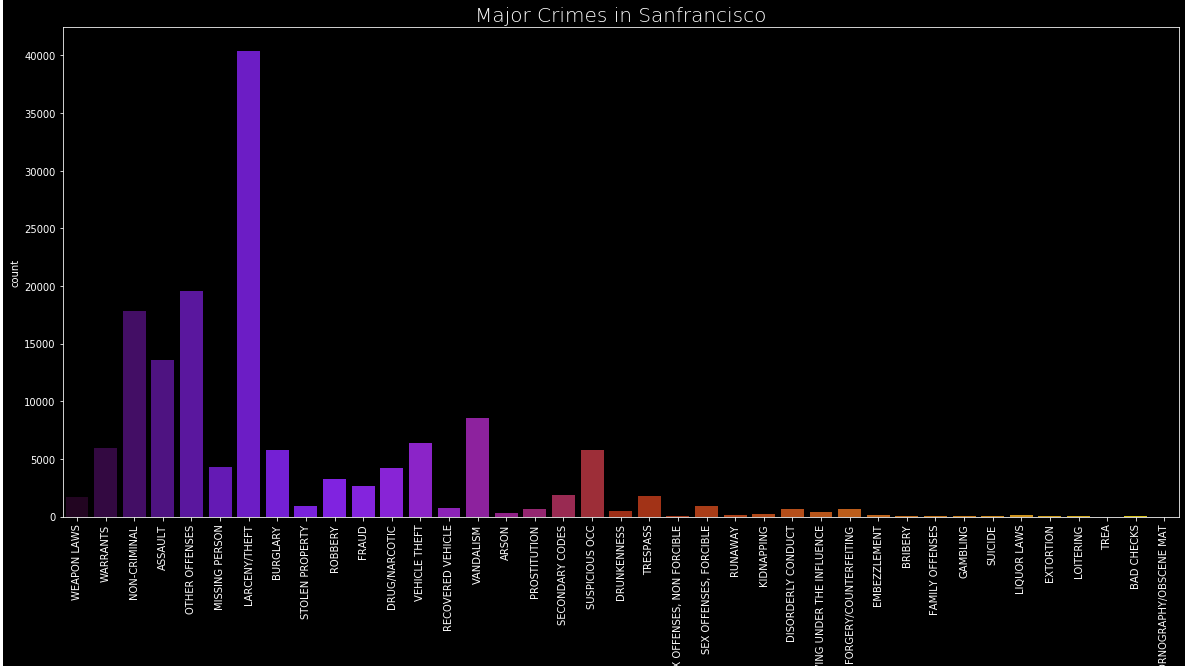
sf\_map'''

**Output**:

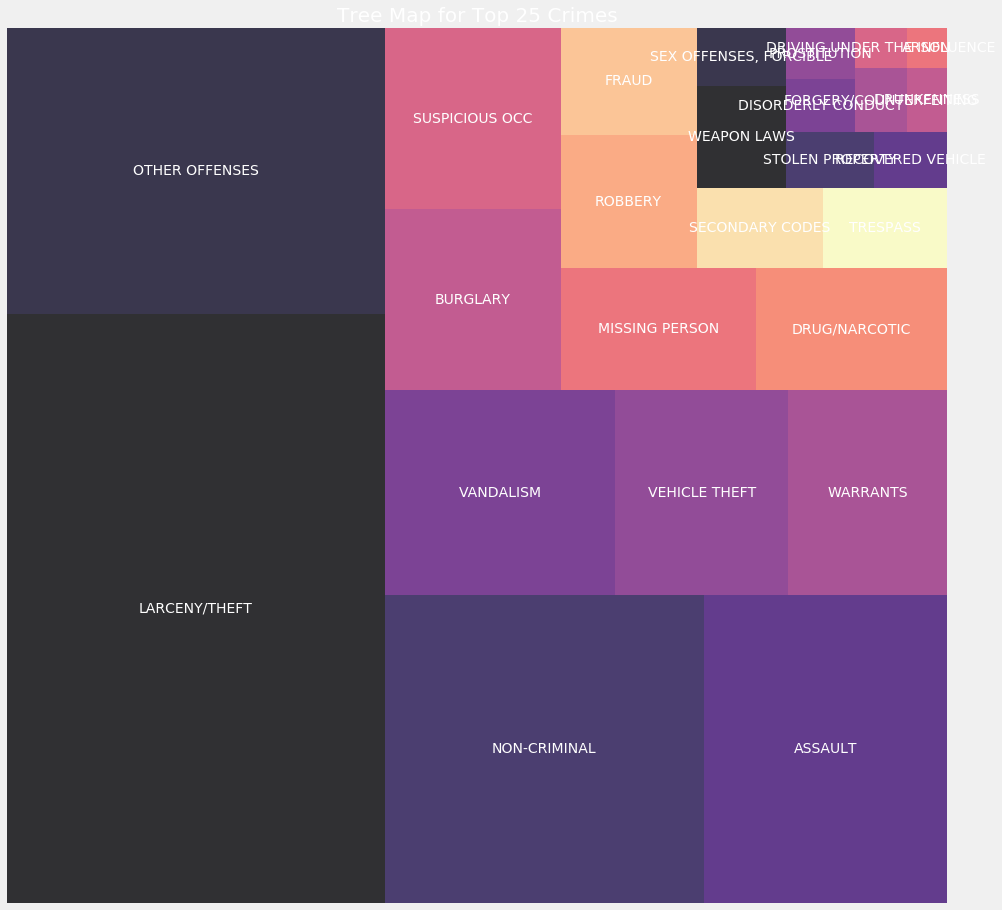
1.Data Description



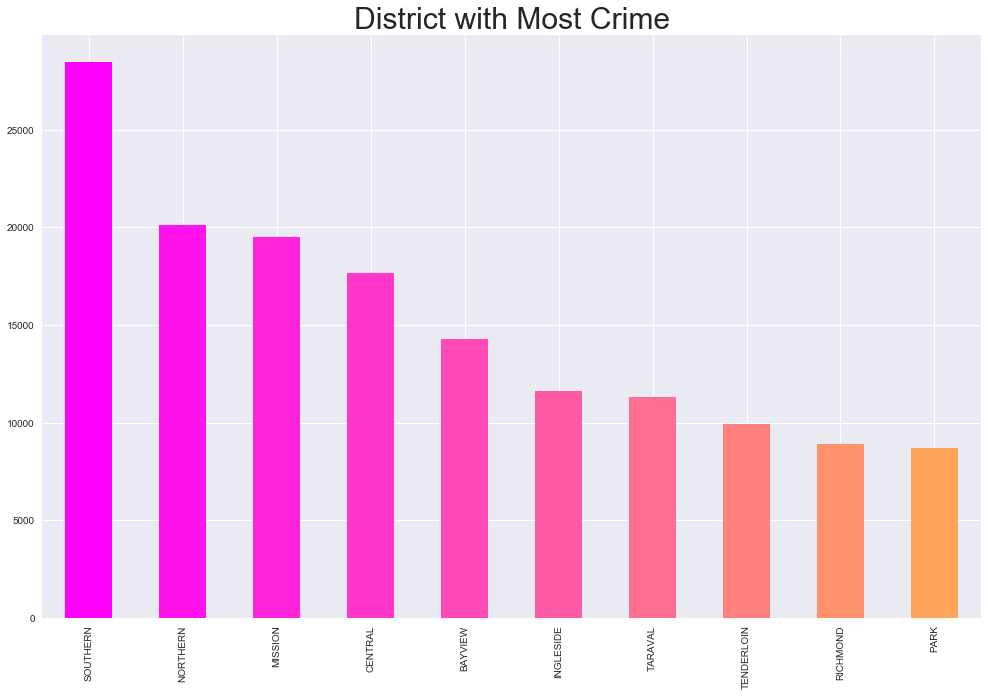
2.Major Type of crimes



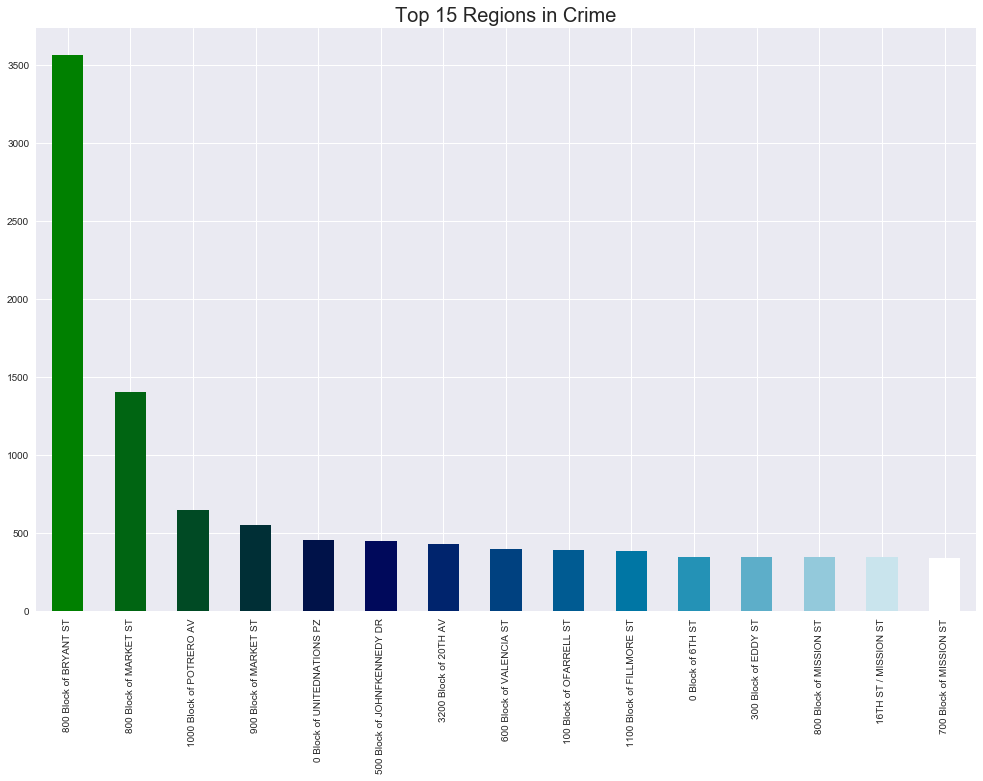
3.Tree map of Crimes

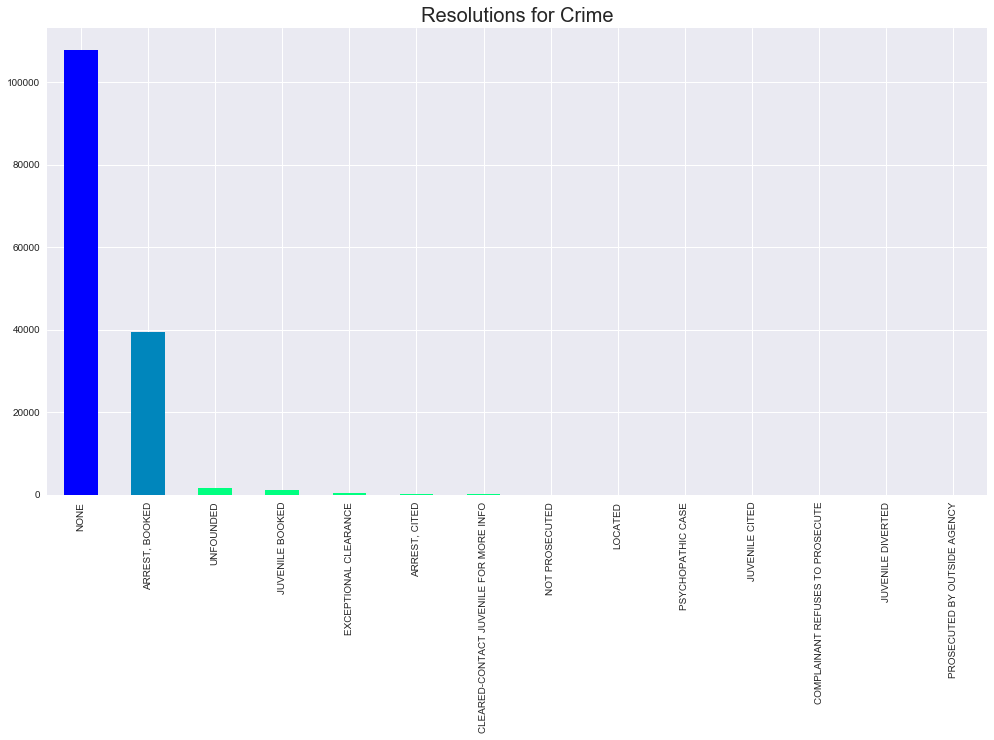


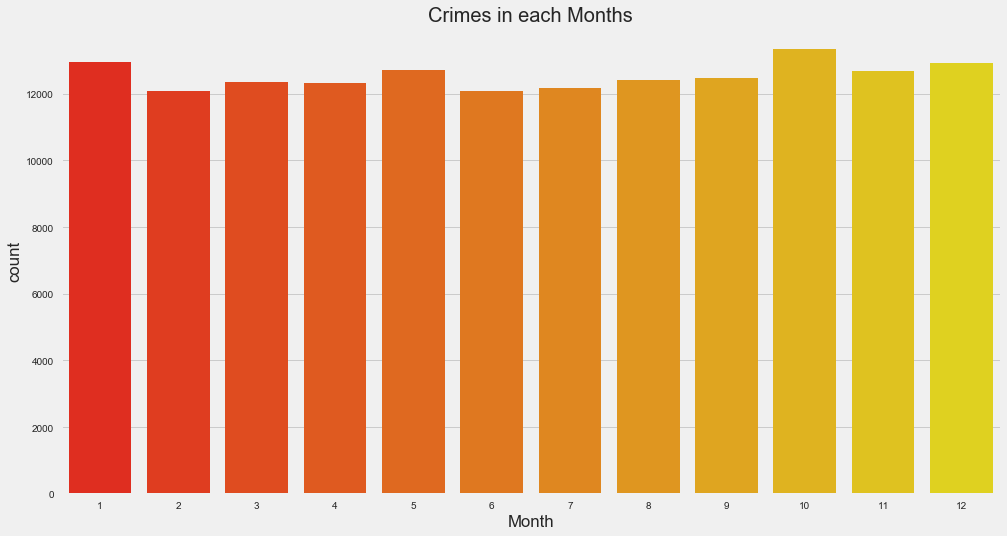
4.District with Most Crimes



5.Top 15 places with most crimes

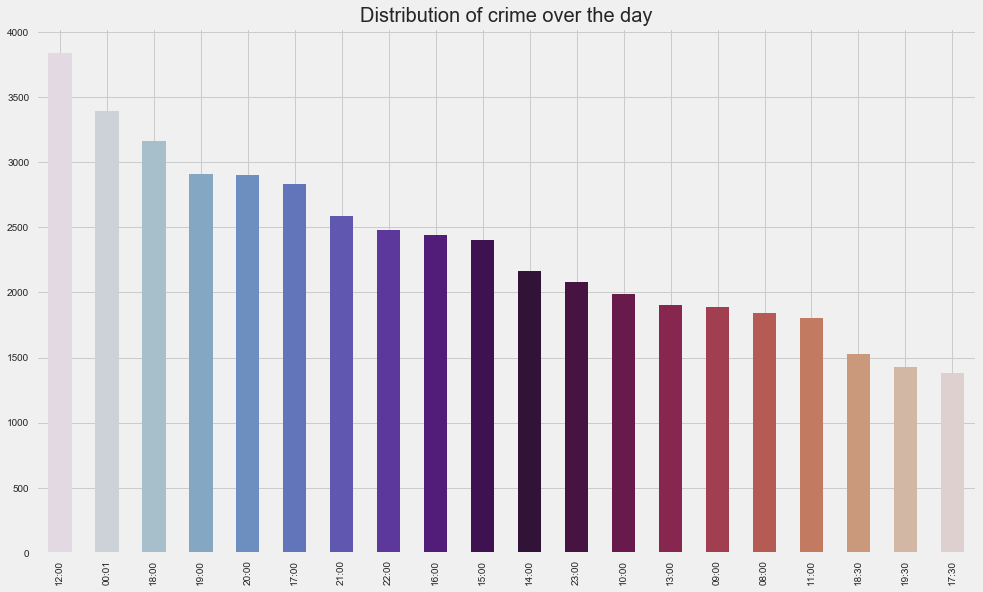


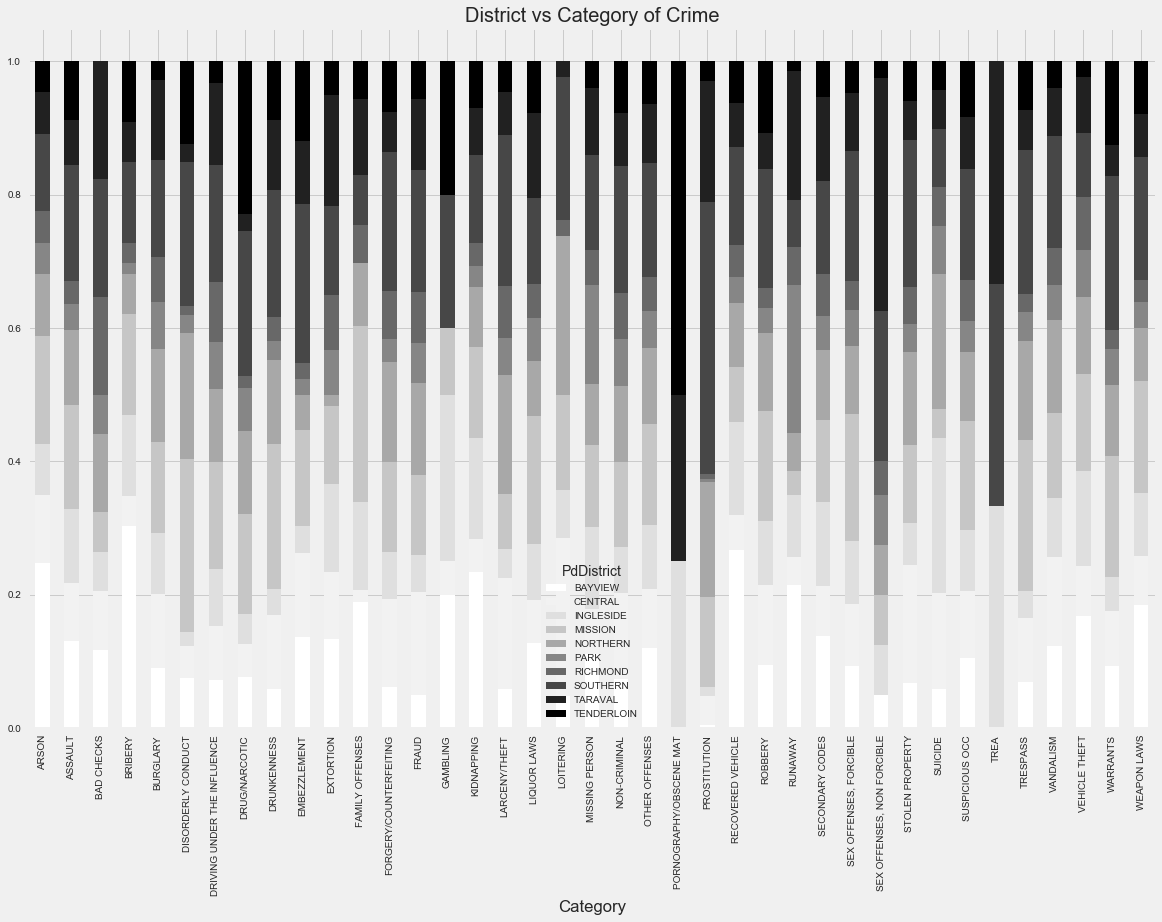
6.



7.Crimes in each Months

8.Time wise Crime occurences



9.District vs Category of Crime

***Part -II Theoretical***

***Our take on “Chronic Offender Point System”***

* **What is this “*Chronic Offender Point System”:***
  + ***One effort, known as Operation LASER, which began in 2011, crunches information about past offenders over a two-year period, using technology developed by the shadowy data analysis firm Palantir, and scores individuals based on their rap sheets.***
  + ***If you’ve ever been in a gang, that’s five points. If you’re on parole or probation? Another five. Every time you’re stopped by police, every time they come knocking on your door, that could land you more points. The higher the points, the more likely you are to end up on something called the Chronic Offender Bulletin, a list of people the data says are most at risk of reoffending and ought to be kept on close watch.***
  + ***The algorithm is always going to augment the system it’s in, and if the system is biased, is unjust, then the algorithm is going to replicate that.***
  + ***Instead of giving offenders points based on location or hotspots and previous records we would like to suggest that there should be the usage a much more sophisticated technology like IoT or AI which reduces the risk of arresting innocent people.***
  + ***Usually in Los Angeles many Blacks were targetted because of this.***
  + ***So it is better to record patterns of crimes happened and acquire much more interesting patterns to catch culprits.***

**Future Scope:**

* **To implement Chronic offender point system not based on location and person index but on previous crime records and similarities between crimes.**
* **To implement Map based interactive system so that we can geographically point crime hotspot.**

**Conclusion:**

**Crime prediction and Analysis can prevent loss before happening and we can achieve a safe and fruitful governance.**

**As we previously said “Prevention is Better Than Cure”**