

# **Final Project Report for Peer-graded Assignment: Capstone Project - The Battle of Neighborhoods (Week 1)**

## **Project Title:**

*Crime Data Analysis for states of India with top venues around*

## **Introduction/Business Problem:**

There are many travel sites on the Internet, including FourSquare, that will tell you all about places to visit, things to do, good restaurants, bars, nightclubs, breakfast places coffee shops, market places etc.

But, these sites do not give information about the law and order of a place. There may be some interesting area to visit but full of criminal activities. So, visiting those places without prior information about the crime is risky for travellers.

So, I have plan for this project.

In this project I have analyzed important venues of different states of India and also crime rate of those states.

## **Stakeholders/Audience for this project:**

This solution is targeted at the cautious traveller. The want to see all the main sites of a state that they have never visited before but at the same time, for whatever reasons unknown , they want to be able to do all that they can to make sure that they stay clear of trouble i.e. is it safe to visit this venue and this restaurant at night.

## Data section

I have done this project on the data available for the country **India**.

So, I have used the website <https://data.gov.in> to get crime rate data of different states of India.

Dataset URL: <https://data.gov.in/catalog/crime-india-2016>

The dataset has following structure:

State/UT (Col.2)	Total Cognizable IPC Crimes - R (Col.118)
Andhra Pradesh	206.4
Arunachal Pradesh	192
Assam	313.8
Bihar	157.4
Chhattisgarh	211.7
Goa	135.3
Gujarat	233.2
Haryana	320.6
Himachal Pradesh	188
Jammu & Kashmir	196.6
Jharkhand	120.4
Karnataka	237.2
Kerala	727.5
Madhya Pradesh	337.9
Maharashtra	217.1
Manipur	121.9

The website <http://www.foursquare.com> has been used to get the top venue information.

The **geopy** library of python has been used to get the latitude and longitude values of any location.

After analyzing, cleaning and filtering of data, the final data is stored in a data frame whose sample structure is as follows:

	State	NoOfCrime	Latitude	Longitude
0	Andhra Pradesh	206.4	15.924091	80.186381
1	Arunachal Pradesh	192.0	28.364199	94.758627
2	Assam	313.8	26.407384	93.255130
3	Bihar	157.4	25.644085	85.906508
4	Chhattisgarh	211.7	21.663736	81.840635