

Data Analysis of Cases of Human Trafficking and Their Disposal

TEAM -15

SHIVAM AGARWALLA – 2017A7PS1589H

ADITYA ANIRVAN – 2018A7PS0367H

A. IMMANUEL BLESSY – 2017A7PS0255H

Abstract - In this report we try to analyse the cases of human trafficking and the way they were handled by the appropriate agencies related to dealing them. We used Python to analyse the data we obtained from the source.

Keywords— Human Trafficking, PCA, K-Means.

I. PROBLEM MOTIVATION

Human trafficking in India, till to this day remains a significant problem. People are frequently illegally trafficked through India for the purposes of commercial sexual exploitation and forced/bonded labour. There is an estimate this problem affects 20 to 65 million Indians.

Men, women and children are trafficked in India for diverse reasons. Women and girls are trafficked within the country for the purposes of commercial sexual exploitation and forced marriage. Men and boys are trafficked for the purposes of labour.

India is also a destination for women and girls from Nepal and Bangladesh trafficked for the purpose of commercial sexual exploitation.

Human trafficking in India results in women suffering from both mental (such as PTSD, depression and anxiety) and physical issues. The lack of control women have in trafficking increases their risk of suffering from mental disorders. Women who are forced into trafficking are at a higher risk for HIV, TB, and other STDs[2].

II. OBJECTIVES

Therefore, by analysing the data available to us, we thought of helping others to understand the way the cases are treated in order to bring general awareness, and hence helping the Govt. tackle the cases in a better way and hopefully come up with significant conclusions helping them to extend better care to the victims.

III. Background - Data Processing

We started off by taking the data available to us from data.gov.in^[1]

Of all the available data we did analyse the numbers for the field namely:

1. No. of cases withdrawn by the Govt. during investigation
2. Cases not investigated or in which investigation was refused
3. Cases declared false on account of mistake of fact or of law
4. Cases in which chargesheets were laid
5. Cases in which charge-sheets were not laid but final report submitted during year
6. Total cases (Chargesheeted + Final Report submitted)
7. Cases pending investigation at the end of the year
8. Cases pending trial from the previous year
9. Cases sent for trial during the year
10. No. of cases withdrawn by the Govt.
11. Cases compounded or withdrawn

As these fields captured most of the information we needed to analyse the patterns of treatment received by the victims.

Since the data for these was given under various categories, namely:

1. Buying of Girls for Prostitution
2. Immoral Traffic (P) Act, 1956
3. Importation of girls from Foreign Countries
4. Procuration of Minor Girls
5. Selling of Girls for Prostitution
6. Human Trafficking

Compounding this data, we performed Principal Component Analysis on the data we obtained from the above processes.

IV. Methodology - PRINCIPAL COMPONENT ANALYSIS (Pre - MidSem Work)

The results of PCA are as follows:

PCA of 2012 Data:

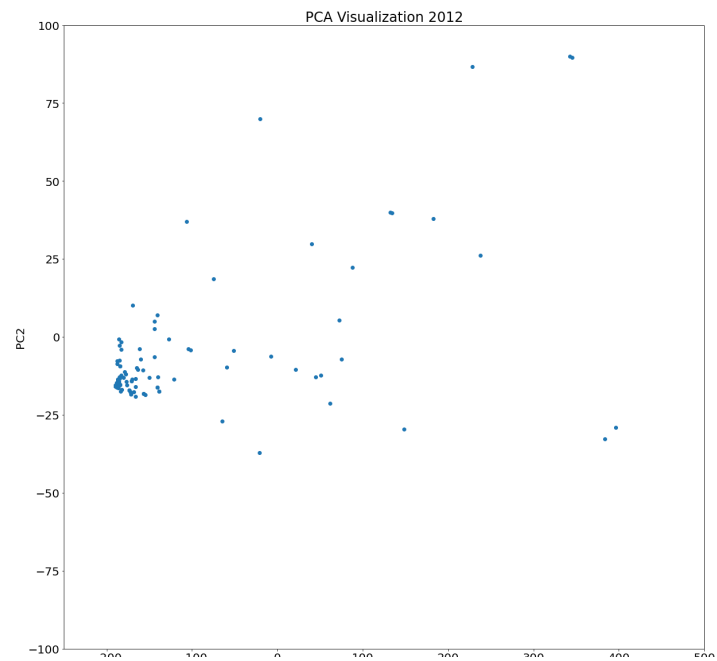


Fig.1 PCA Visualization of 2012 Data

PCA of 2013 Data:

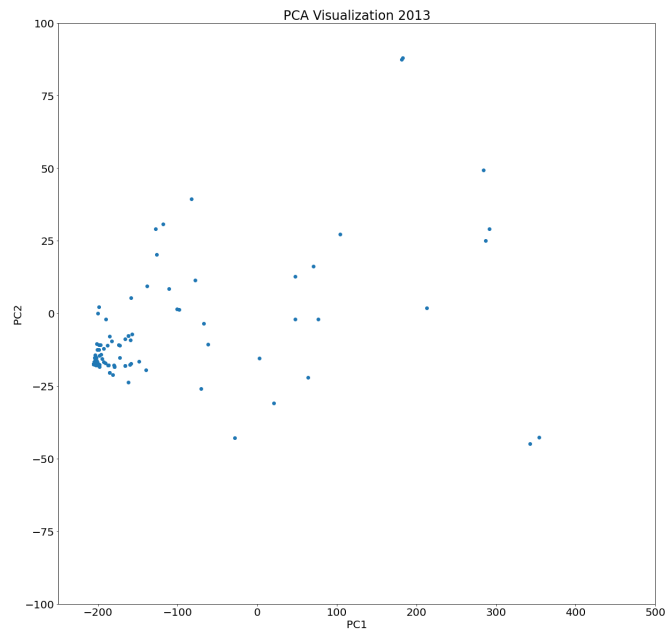
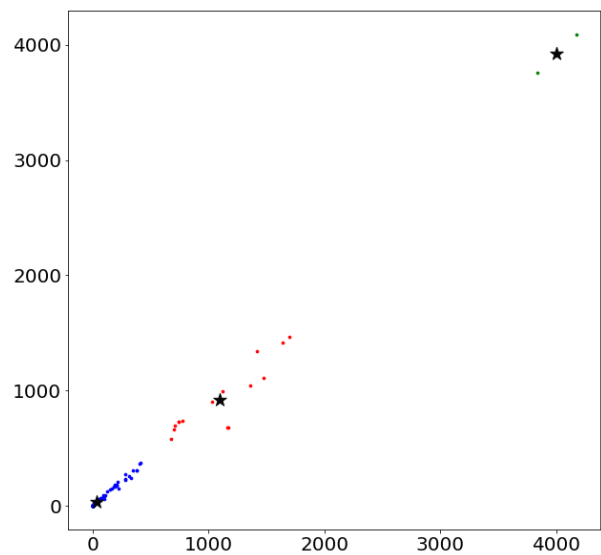
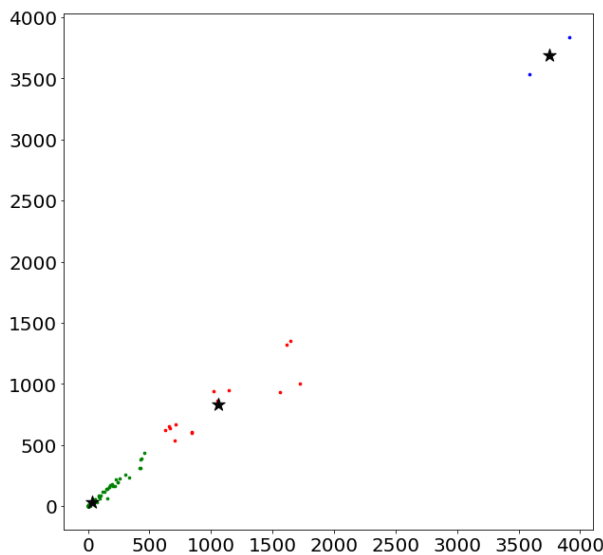
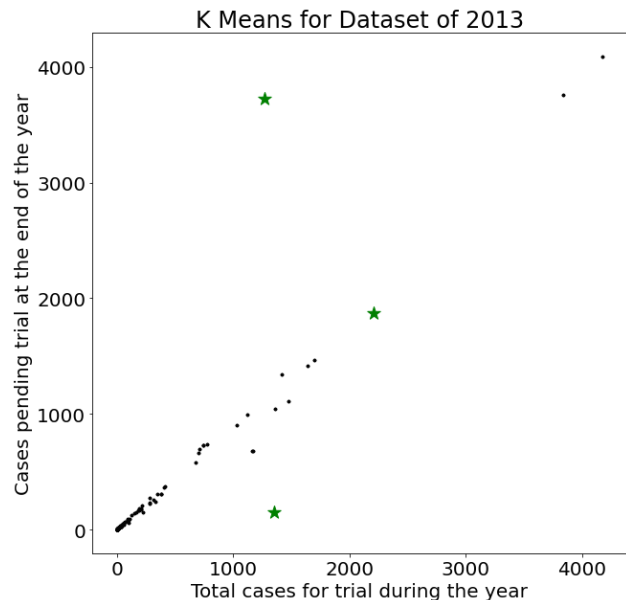
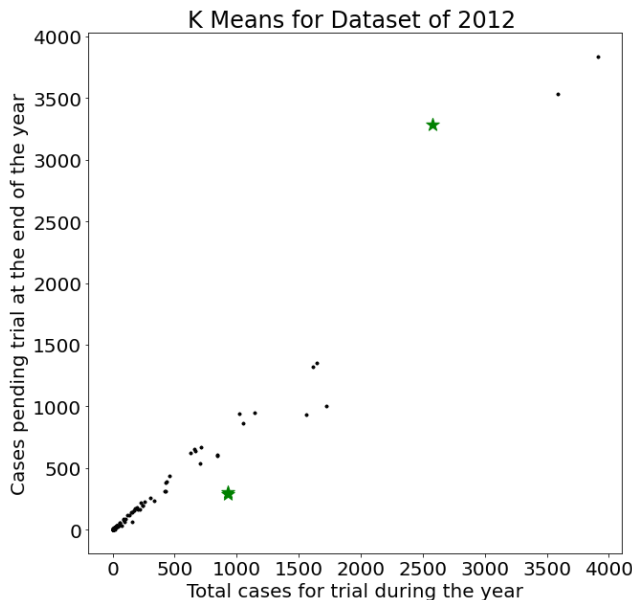


Fig.2 PCA Visualization of 2013 Data

IV. Methodology - K - MEANS (CLUSTERING ANALYSIS)

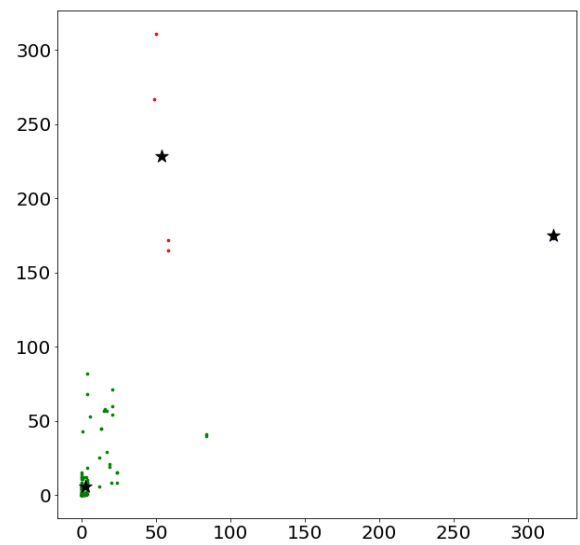
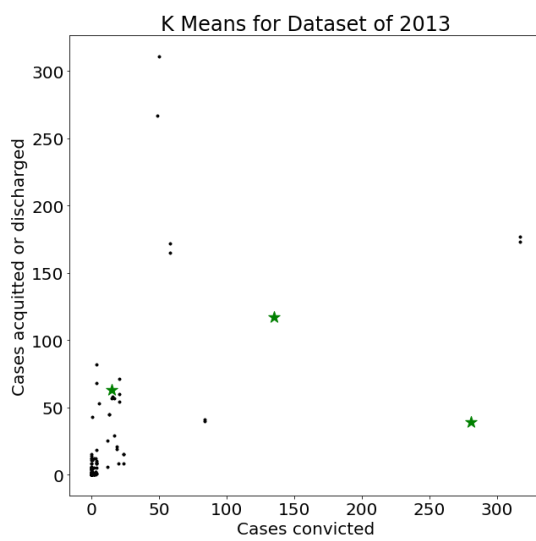
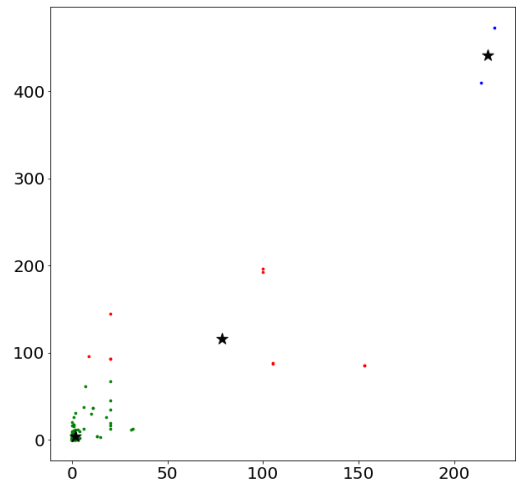
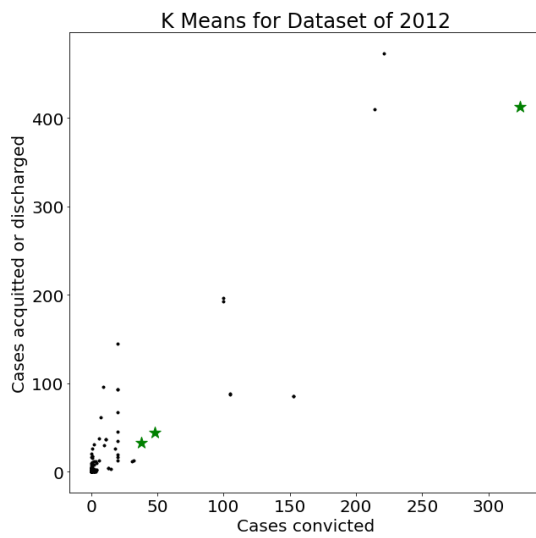
We applied K-Means as the method to process the data and find correlations in the data.

The data fields “**Total Cases at Trial During the Year**” were plotted against the “**Number of Cases Pending Trial at the End of the Year**” for estimating the way the judicial system works in its capacity to tackle these cases. The Initial results are as follows (Initial upon Final Plots):



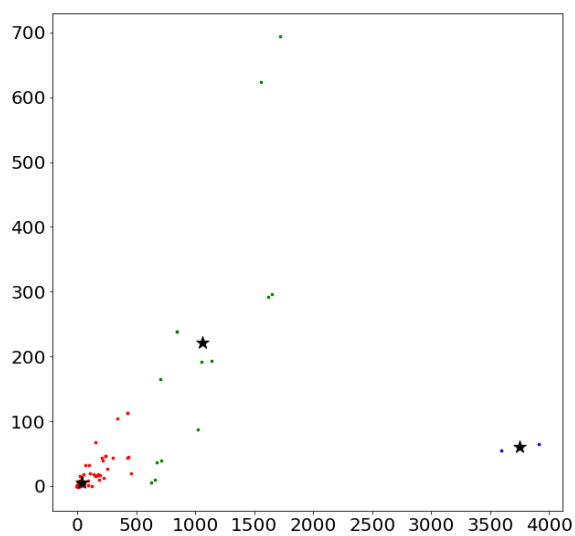
The Final clustering is as coloured as shown above.

Then we also went ahead and calculated the K-means for the data fields “**Cases convicted**” vs “**Cases acquitted or discharged**”. This gave us a better understanding of the workings of prosecution/ Government and their efficiency in working for the victims. We also wanted to know any anomalies, if exists. The related plots are as follows(Initial upon Final Plots):

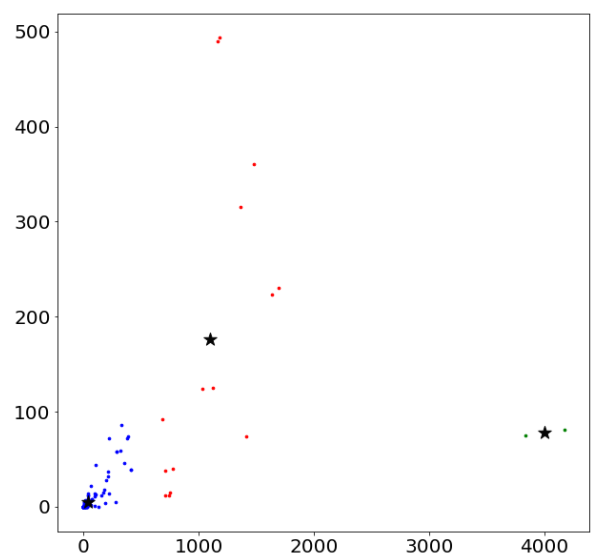


And we plotted the **No. Of Cases Completed Trial** vs **Total Cases Brought for Trial.**

2012



2013



V. Results & Insights

From the *first data set* (Total Cases at Trial During the Year vs. Number of Cases Pending Trial at the End of the Year), we observed that no. of pending cases high in those states with higher no. of cases highlighting the incapacity of the concerned prosecution team to go ahead with their jobs, due to various reasons they might come across.

From the *second data set*, we observe a higher ratio of acquitted/ discharged cases to convicted cases. Although we cannot come to any conclusive decision, but from the following(next data set) analysis, we came across the higher no. of pending cases, dragging out the cases that can be prosecuted out due to running out of Statue of Limitations, which is a real possibility considering the trial time or time spent in courts in our country and many states as well. Inability of procuring hard evidence, evidence manipulation owing to Political Influence etc., sometime false reports contribute to higher acquitted/discharged cases.

From the *third data set*, namely, No. Of trial completed cases vs. No of cases brought to trial, provides a much better introspect into the underlying problem that we thought to bring awareness. For example, 2 data points(**Green Cluster**) in 2013 plot, have 4000 cases, out of which fewer than 100 cases completed their trials. This highlights the need of judiciary wing to step their efforts to fast-track these sort of cases as most of the times victims are highly under-privileged or don't have enough resources to fight for the longer times, essentially putting their lives in further risk in addition to the ones already existing.

From this analysis, Government should work together with judiciary systems and allocate them enough resources to ensure the safety of all its citizens irrespective of their financial status or any other factor in that matter. Policy shaping should also ensure the time to process these cases is reduced and provide better protection to the victims.

REFERENCES

[1] Data Sources

<https://data.gov.in/search/site?query=human+trafficking>

[2] Information Source

https://en.wikipedia.org/wiki/Human_trafficking_in_India

[3] GitHub Repository

<https://github.com/Shivam-Agarwalla/Data-Mining-Project>