

INDIA CRIMES CAPSTONE PROJECT

Sankul Pandey

sankul99@gmail.com

DS 0722

After a grueling yet exciting few months of just studying for our elective subjects and working for Flip robo technologies I was delighted to find out that I will be given the opportunity to work on a capstone project. I was excited because it was after a year or so I was working on a proper project, the last one being my FYP for engineering. The Capstone project excites me because it gives us the opportunity to finally show the skills that we have learned during our training or learning period. Given to us on the 26th of June we were given around 3 weeks to complete the project. My initial thoughts on seeing the project was of intrigue as the whole project was divided into 4 phases each targeting skills that we learned during our learning and training period. Phase 1 gave us the task to collect data. The data we had to collect was of population , literacy rate and area of each state from the year 2001 to 2012. We could collect the data any way using multiple techniques such as web scraping or could even do it using Google. I personally collected the data through the survey of India website and the data which i couldn't find I took the help of Google. After the data was collected I saved the data in Excel and as well as a CSV file.

We then move on to Phase 2 . In phase 2 we were given the task of making beautiful and easy to understand visual analysis of the data we have collected along with the data that was already given to us by the institution. I personally used two of the skills that I love using. Python for analysis and Tableau. Using python I first imported the required CSV files that I made in phase 1 and then imported the India Crimes files which were given to us.

In phase 2 we were asked to complete the analysis of Total crimes with parameters such as Literacy rate , population and area of the state. There was one complication that the data given to us showed the crime for each and every district of every single state for multiple years. This was a problem I had to work out through because I only needed a certain amount of data to complete my visual analysis. How I tackled this problem was by removing the columns I don't require and after that merging the crimes of each district using the state column. After doing this I got the total crimes committed for every state. Using this dataframe I just built it individually compared the state wise crimes with all the 3 parameters given to me. The results were very informative and told me a lot about the data set I was analyzing. It showed me how crimes committed in every state of the country is directly dependent on the area, literacy rate and population of that state. I further continued my analysis using the tool Tableau. Tableau is a very user-friendly tool and easy to understand. It has multiple analysis options and gives visually stunning graphics for your data. I made a complete tableau workbook of 9 pages in my analysis on the basis of the parameters given to us along with a few added analyses of my own. It was of great fun to use Tableau and Python to create these Visual analysis that were also easy to understand.

Then came Phase 3 which targeted our skills of using SQL. We were given a few queries that we had to solve. I approached this task by solving the simple queries using PostgreSQL. We later had to merge two queries that we had gotten and form a single data frame. I opened the database on a jupyter notebook and then merged the two required dataframes and finally completed the visual analysis of that data frame using Tableau. The screenshots I took I later uploaded on the same python notebook that thereby completed phase 3 of the project.

Phase 4 was a simple task of clustering using Machine learning. I created a separate Python notebook using Jupyter. I then uploaded the total crime dataset and then using Kmeans in minmaxscaler lib clustered the dataset on the basis of three area wise. Sensitive areas, peaceful areas and moderate areas. This gave us information such as which places are the best to live, places which have the most amount of crimes etc.

With this came an end to my Capstone project journey which was an exceptional learning journey. I learned and got to practice a lot of things. A capstone project is essential in the life of every person who has gotten into training and wants to prove to his superiors that whatever he has learned he knows how to put it into use. What I personally loved about this particular capstone project is that it made us focus on everything that we have learnt from the very beginning in August of 2022 till now and focused on every skill that we have learnt till now and harness the abilities to its optimum state and work through all the barriers that were given to us during this capstone. This project was definitely a great experience and I hope that whatever I learned and imputed in this particular project will help in getting me to the place that I always wanted to be at. I would lastly like to thank you data trained for giving me the opportunity of this particular project as it was not only a task for our own benefit but also gave us an insight into the crimes that our country is facing and not only that it makes us understand the ways that we can improve the society around us and by the time I have completed project I know that change comes from one person.

Thank you.