

Murder Victims in India

Venkatayashwanth Damera | Advisor Dr. Charles Hoot
Applied Computer Science | North West Missouri State University

- The crime rate in India is rapidly increasing day-by-day. Which motivated me to analyze the murder rates with respect to age and place.
- Predicting more number of victims by gender based and which place consists of more number of victims

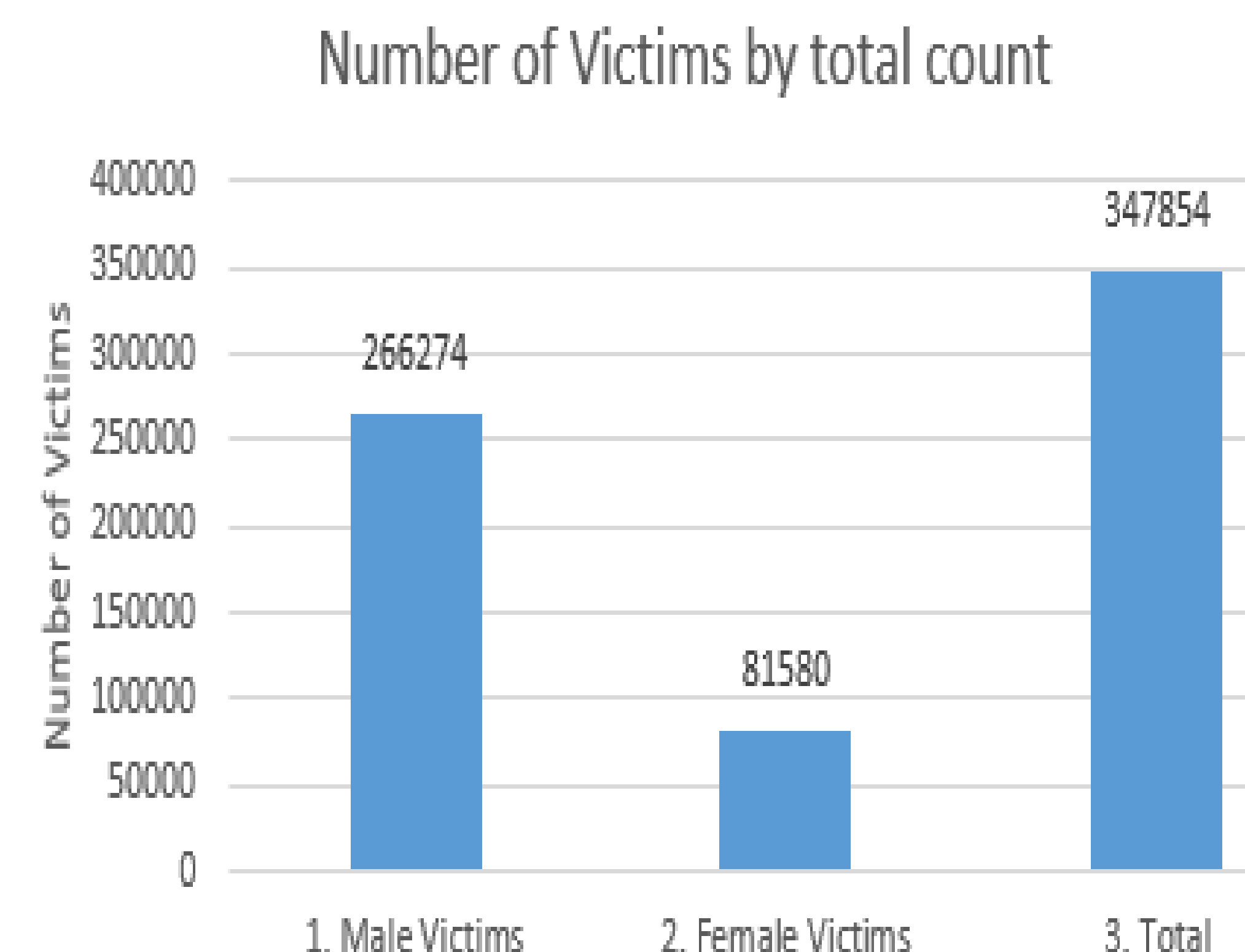
- Data Preprocessing: `dropna()` – clears out all null values.
- Prediction: `LinearRegression()` and for evaluating this, RMS Error.
- Classification: `DecisionTreeClassifier()` and `RandomForestClassifier()`.
- Clustering: `Kmeans(n_clusters=2)` and PCA for dimensional analysis.

IDE and Tool used - Jupyter Notebook

Data Source – Kaggle - https://www.kaggle.com/rajanand/crime-in-india?select=32_Murder_victim_age_sex.csv

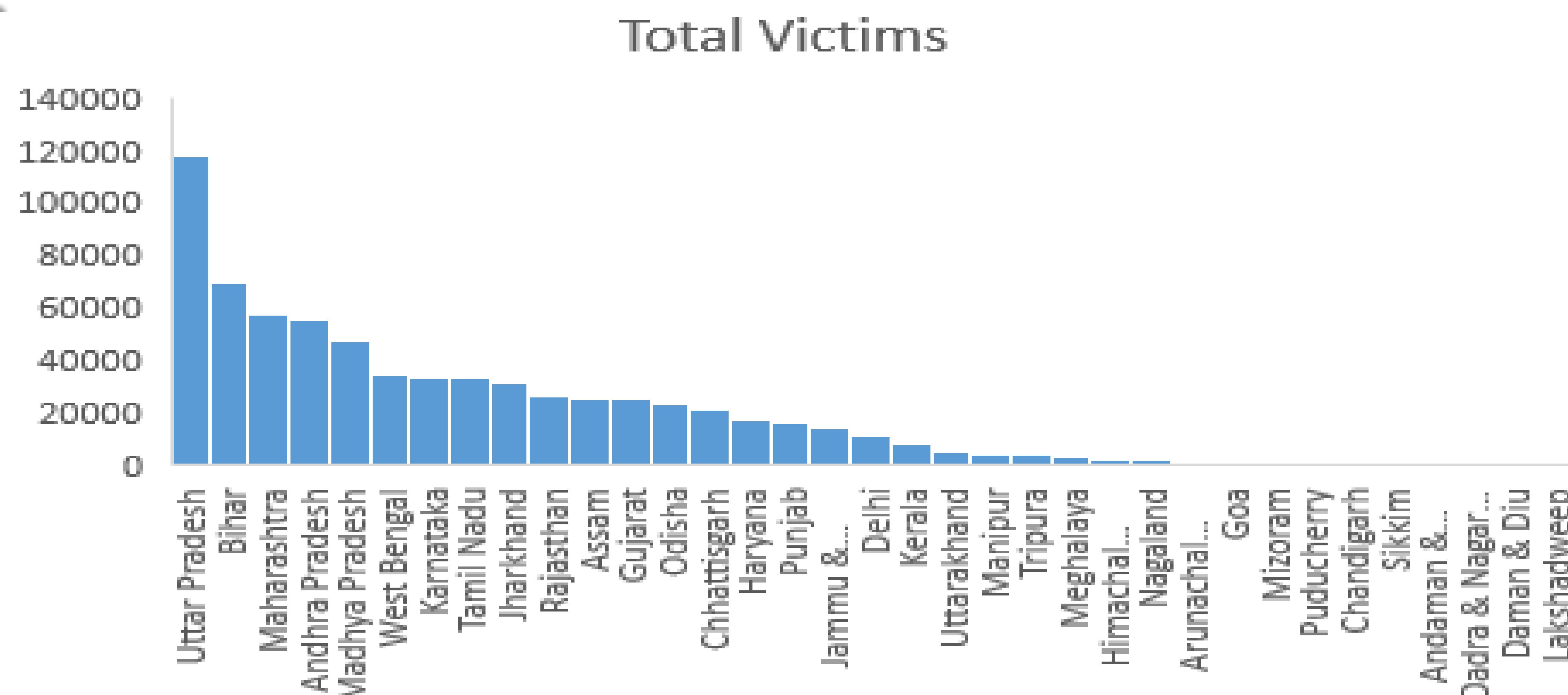
Python Model References - <https://scikit-learn.org/stable/>

Bar Graph 1 shows the number
Number of victims to gender.
While Bar graph 2 shows the
Number of total victims to the
Area.



Bar Graph 1

Bar Graph 2



- Evaluate the minimum number victims.
- The Bar Graph 1 representation show that as there are more number of male victims than female victims
- Using Random Forest Classifier accuracy of 98% is obtained.
- Accuracy of 92% is achieved with decision tree classifier which indicates good model.

I thank Dr. Hoot for giving me proper guidance through out this project.

<https://github.com/44-599-MachineLearning-S21/project-machine-learning-s21-Yashwanth-Damera>