**Team Name** :- Team phoenix

**Problem statement**

Analysing profiles of 2019 general election candidates using attributes like criminal cases, caste category etc.

**Problem statement analysis**

While analysing the problem statement we have considered what factors should be considered while examining a candidate's profile, how are the number of election candidates distributed between states, as well as the gender ratio in a particular constituency. Also, what is the distribution of criminal cases between states and how does it relate to their candidate profile and What is the age distribution of election candidates across each party, as well as the distribution of voters by category.

**Tools used**

To build the solution we have used the exploratory analysis. This technique has helped us to investigate the data for overall understanding and to take for the steps.

Also, for code building we have used google collab. The simple and distributed interface of google collab helped us to construct appropriate solution.

**Technical analysis**

To construct the solution using Python we have used several libraries and functions in Python. Some of the functions and libraries are as follows: -

* Scatter(): The scatter() method in the matplotlib library is used to draw a scatter plot. This scatter plot displays the gender, category, state, party, name and education of the candidate upon hovering the marker. The interactive scatter plot displays the age, education , state and party upon hovering the constituency in the scatter plot. At an instance one state with the constituencies will be displayed.
* Barplot() : The bar plot() method in the matplotlib library is used to draw the bar charts. We have used the same function to plot the education vs crime data , category wise representation of candidates , party wise split up of candidates, age vs crime data, state vs total votes.
* Pie() : We have used the pie() function to display the gender wise split of candidates appeared for the election and the gender wise split of winners also.

By using this libraries and functions we are able to construct a solution in a desired manner and able to represent it.

**Solution overview**

* We analyzed the profile of general election candidates based on their state, gender, age, education, and criminal history.
* We correlated the parameters of gender and age of the election candidates to understand the highest and lowest variables.
* After correlating all the parameters, we derived conclusions based on the resulting graphs.
* This correlation of parameters is useful to analyse the nature of election in that constituency.

**Conclusion**

After the exploratory analysis, solution building and Code implementation we can conclude following things: -

* Uttar Pradesh has the most candidates who have received the most votes, followed by West Bengal and Maharashtra. As a result, penetration is greater in these states. While Jharkhand has the lowest number.
* Male candidates outnumbered female candidates by a wide margin. While the proportion of female candidates winning elections is slightly greater.
* Correlation between criminal cases and education shows that criminal cases are highest among graduates. The Number of Criminal cases are highest in Kerala followed by U.P.

**Future scope**

This model has built for 2019 elections, but this model and its functionality are dynamic in nature. So, we can use this model for the analysis of further elections also.

This model illustrates the analysis of Lok Sabha elections but we can draw conclusions of state elections as well as ‘Sthanik Swaraj Sanstha’ Elections by using this model.