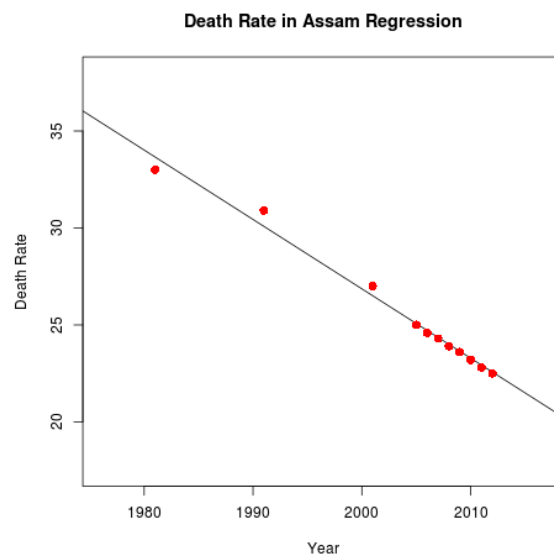


R Lab Assignment

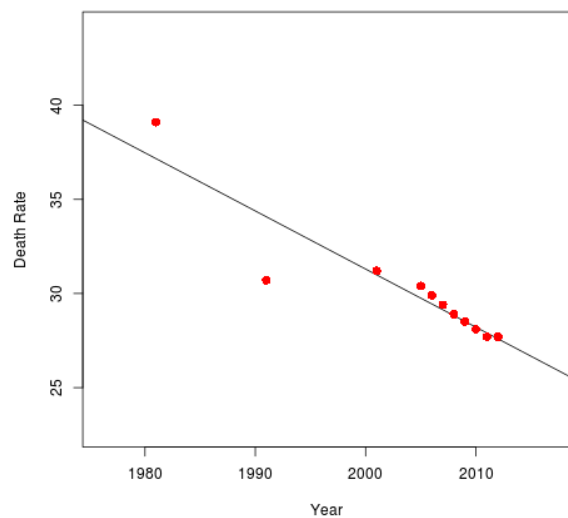
By: Vaibhav Jain (11634)

IVth Sem, B.Tech(IT & MI)

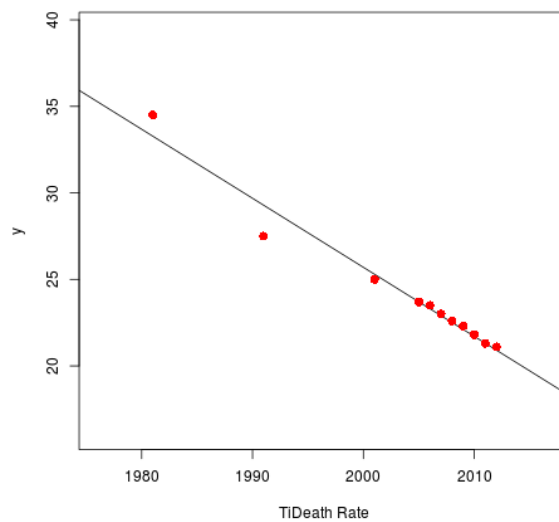
Problem#1: Choose any 10 states. Find the regression line between the death rate of total population and time.



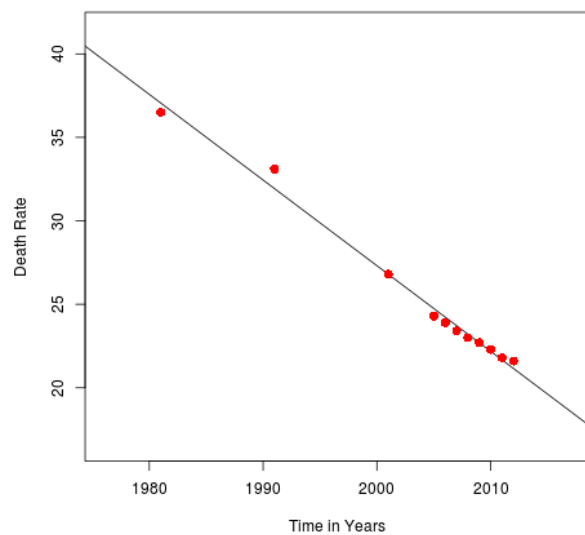
Death Rate in Bihar Regression



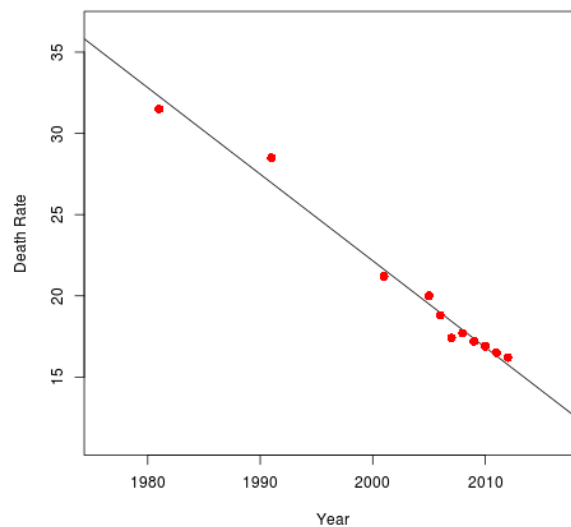
Death Rate in Gujarat



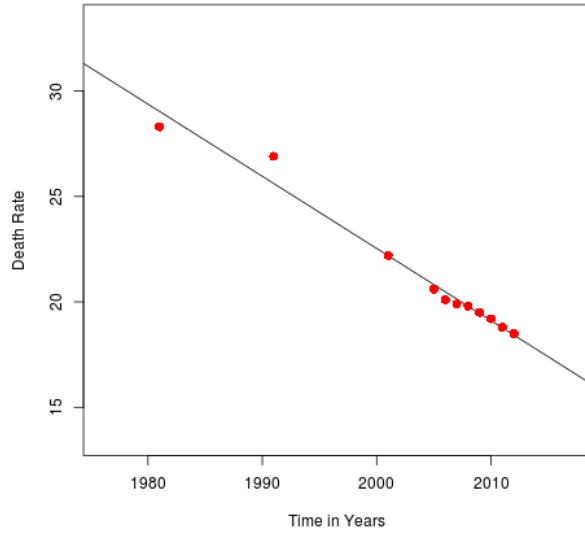
Death Rate in Haryana Regression



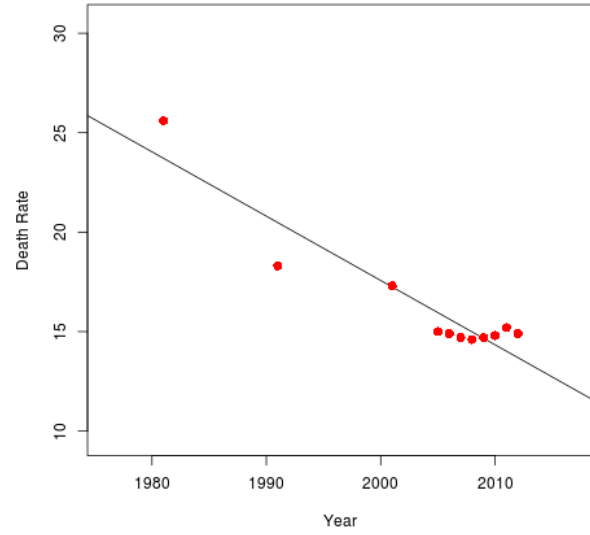
Death Rate in Himachal Regression



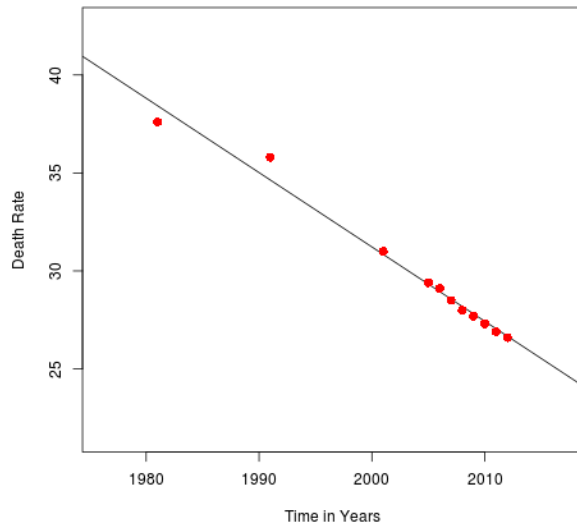
Death Rate in Karnataka



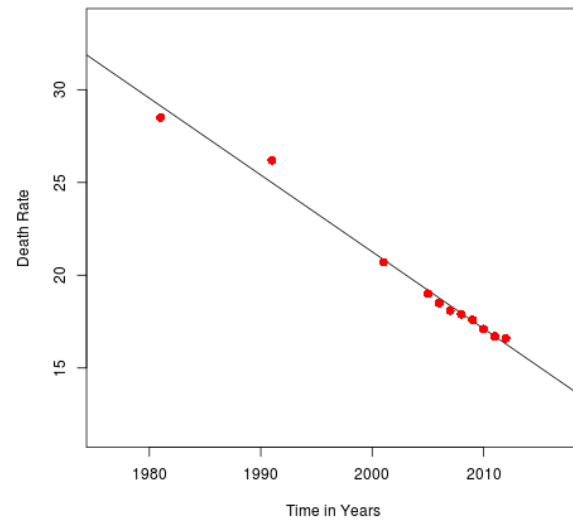
Death Rate in Kerala Regression



Death Rate in Madhya Pradesh Regression



Death Rate in Maharashtra Regression



10 states selected which were selected are:

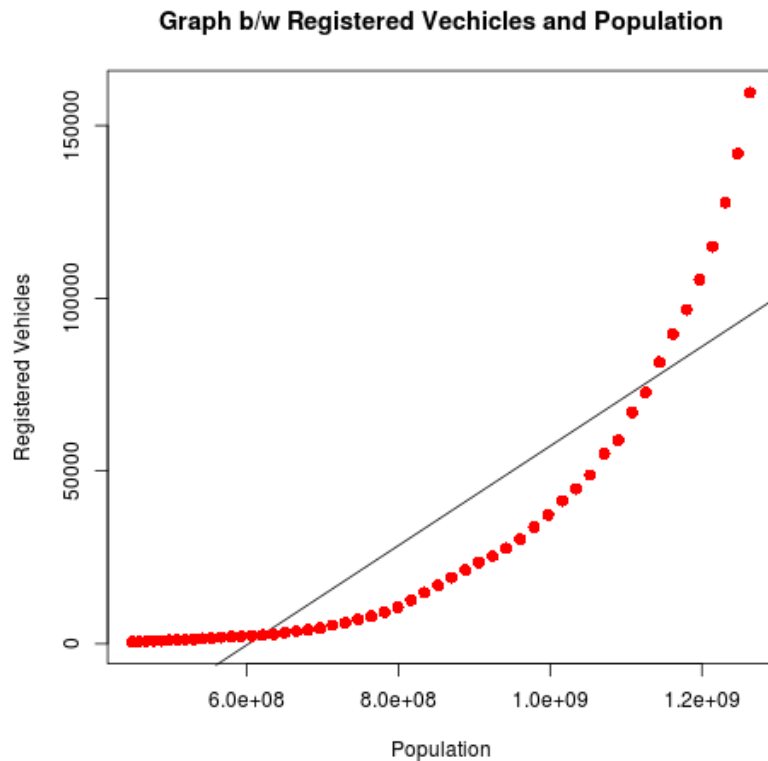
- Andhra Pradesh
- Assam
- Bihar
- Gujrat
- Haryana
- Himachal
- Karnataka
- Kerela
- Madhya Pradesh
- Maharashtra

In these graphs, death rate per 1000 people in the respected states is plotted against time in years. From the graphs, it is obvious that the death rate will decrease in the coming years since the linear regression has negative slope.

We also observe from above graph that the states - Madhya Pradesh, Maharashtra, Karnataka, Himachal Pradesh, Haryana, Assam, Andhra Pradesh have a good fitting of the linear model, which means that the regression line can predict the death rate of the states from the above graphs more accurately.

It can also be interpreted that the reason behind declining death rate can be better health care facilities in these states.

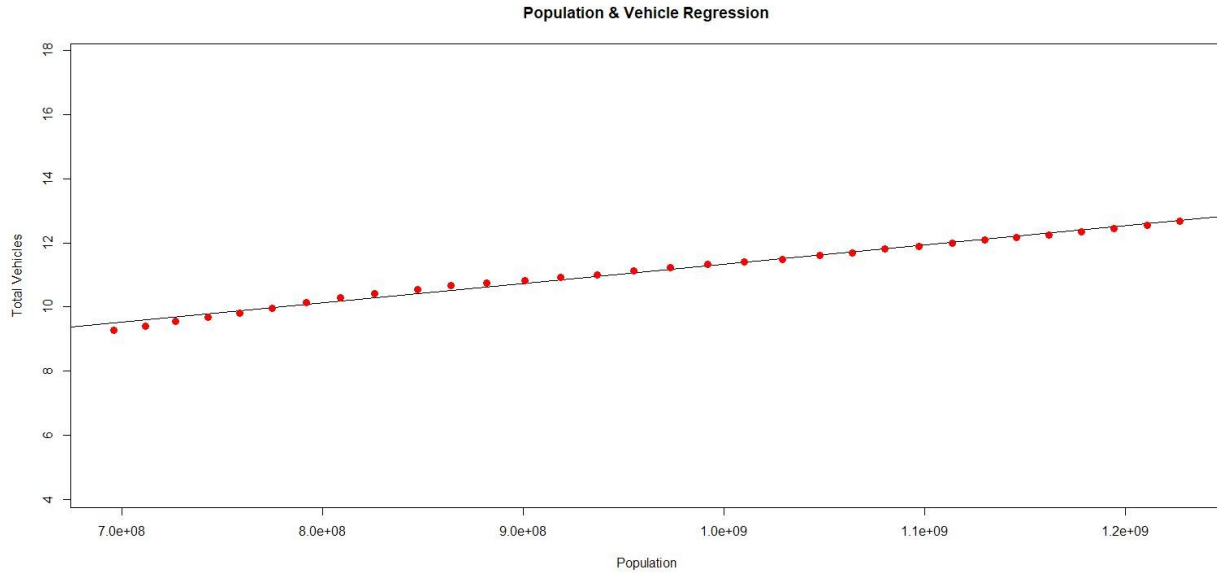
Problem#2: Linear Regression between registered vehicles and population



In this graph, number of registered vehicles are plotted against the population. This graph gives us the insights about the extent of dependence on vehicle among the population in a given area. It also help the government and automobile industries to plan their future policies.

From this regression plot, we can observe that:

1. The number of registered vehicles has increased with increase of population.
2. We can see that the linear model is not a very good fitting for the data. A more complex curve may generate a better fitting.
3. Although the fitting is not very accurate, it can still predict that in future more number of vehicles will be required.

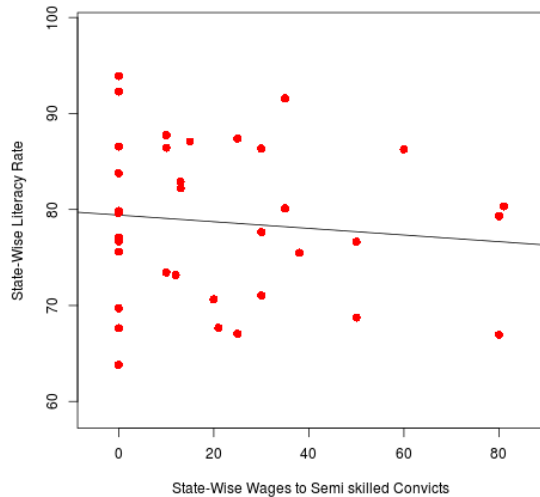


The above graph represents the logarithm of population plotted against the total number of vehicles. It can be observed that:

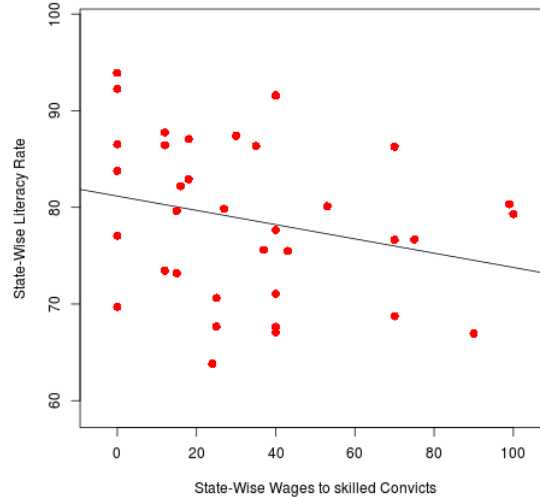
1. The linear model gives a good fit for the above data.
2. From this model, the prediction made should be very accurate. It implies that the required number of vehicles can be better predicted by taking logarithm of it and applying linear regression on it.

Problem#3: Linear Regression between wages of convicts and literacy rate

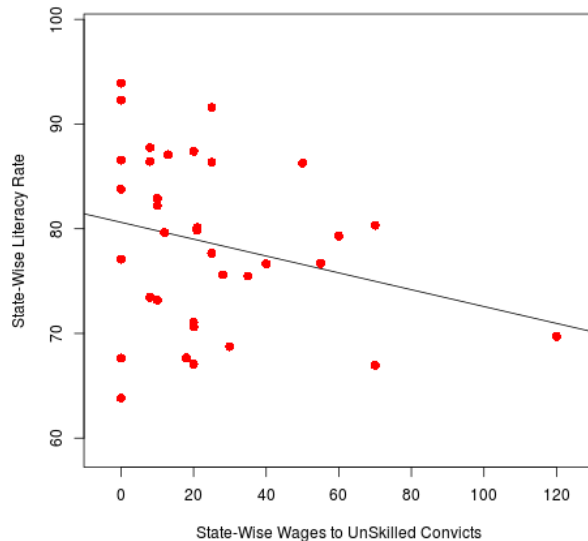
State-Wise Wages to the Semi Skilled Convicts and Literacy Rate in 2011



State-Wise Wages to the Skilled Convict and Literacy Rate in 2011



State-Wise Wages to the UnSkilled Convicts and Literacy Rate in 2011



The above three graphs give us a plot between the literacy rate of a state and wages of different convicts of those states.

Upon analysis, an absence of concrete relationship between the two parameters was found indication towards the fact that those two are not

related and are both independent factors that need to be taken into consideration separately as demonstrated clearly in the above graph. The following conclusions have been made –

1. There is no correlation between the two features given
2. There is no significance of fitting a linear model here as the features are not correlated. Hence, any predictions made on the data is not reliable.