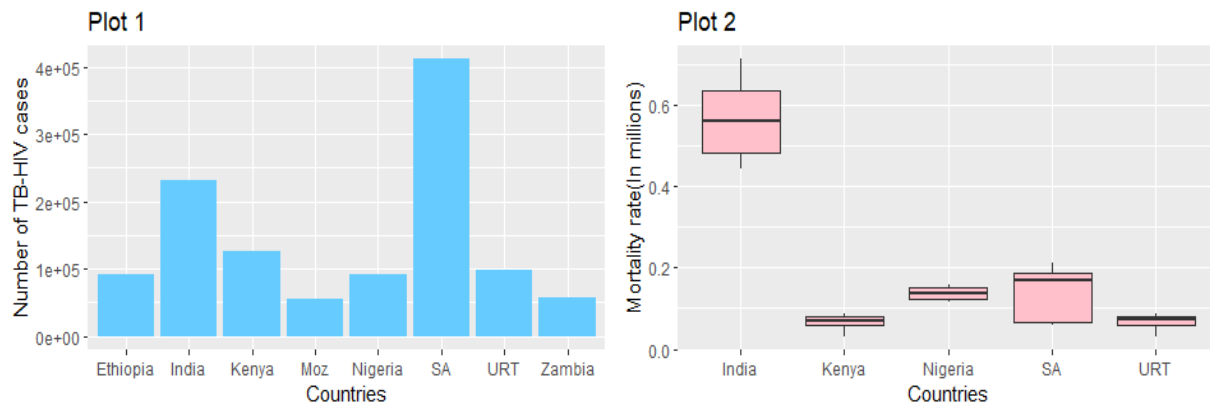


## Pair Assignment 2

**"TB Burden Estimates"** dataset has been used. It has **4272 observations** and **50 variables**, from which 7 variables are selected to test the hypothesis. The variables in the dataset show the estimates of number of cases and mortality rate of TB cases over the years 2000 to 2019, for various countries.

```
## Rows: 3,640
## Columns: 7
## $ country      <chr> "Afghanistan", "Afghanistan", "Afghanistan",
##               "...
## $ year         <int> 2000, 2001, 2002, 2003, 2004, 2005, 2006,
##               2007...
## $ e_inc_num    <int> 39000, 41000, 43000, 45000, 47000, 48000,
##               5000...
## $ e_inc_tbhiv_num <int> 110, 110, 110, 120, 140, 170, 200, 230, 250,
##               2...
## $ e_mort_num   <int> 14000, 14000, 13000, 14000, 13000, 12000,
##               1100...
## $ e_mort_tbhiv_num <int> 35, 65, 62, 69, 71, 79, 85, 89, 100, 120,
##               120,...
## $ e_mort_exc_tbhiv_num <int> 14000, 14000, 13000, 14000, 13000, 12000,
##               1100...
```

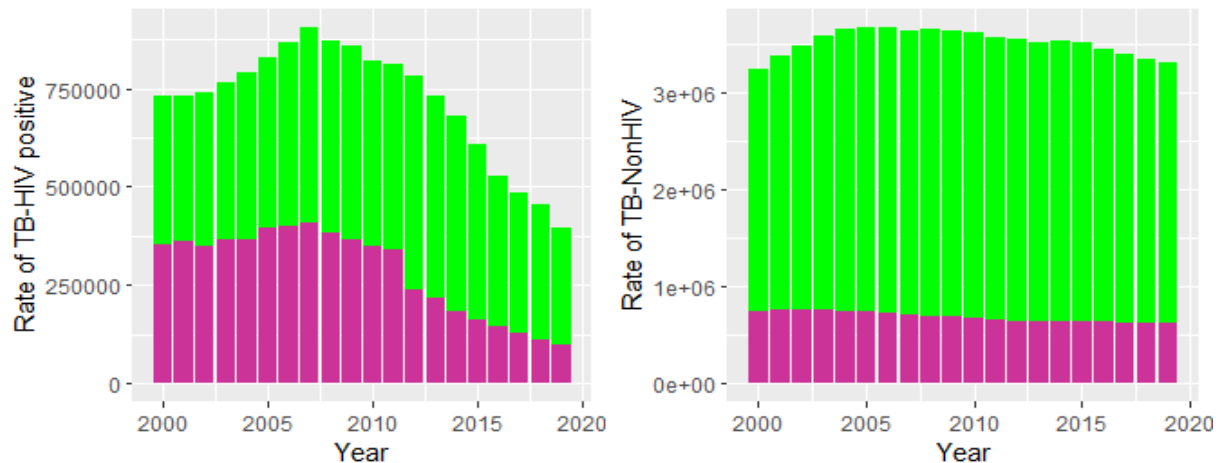
We have selected feature of TB cases with documented HIV against NON-HIV TB cases because the data explicitly differentiate between HIV positive and HIV negative TB cases.



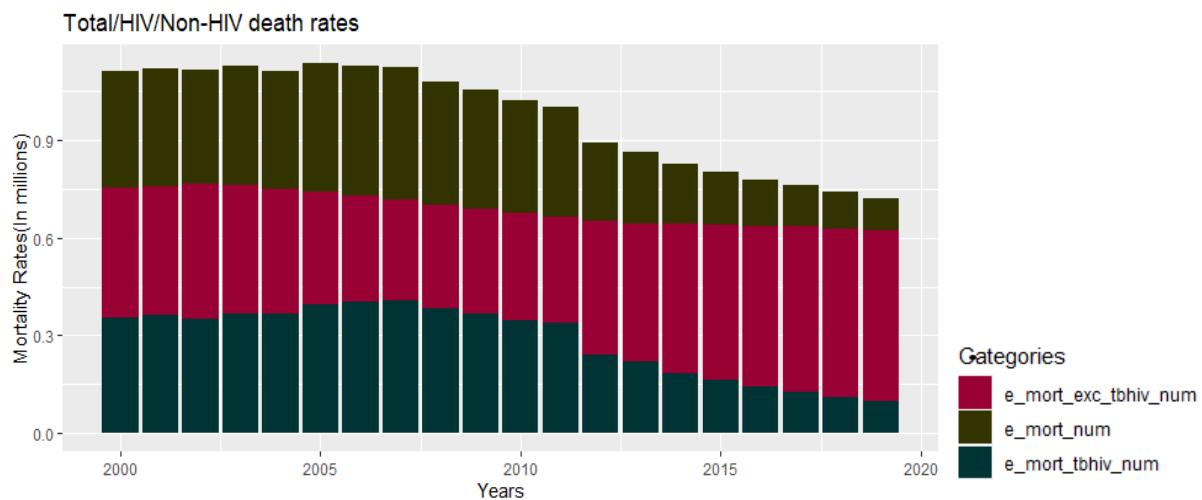
We have chosen countries with highest TB HIV cases so that we can better demonstrate the relationship between number of cases and mortality rate of TB HIV cases. (Plot 1)

We will analyze mortality rate of HIV positive and HIV negative TB cases over the years in the 5 countries with highest number of TB HIV cases. From given boxplot graph, we can see range and median values of mortality rates in selected countries.(Plot 2)

From the below graph, It is evident that there has been a reduction in number of cases of TB with documented HIV, so has the mortality rate of HIV positive TB cases. The same trend is seen in the graph of TB cases excluding HIV cases. Green shows the number of HIV and Non-HIV cases respectively on both plots, and Pink shows the mortality rates.



Let's see the relationship between total mortality rate of TB cases with mortality HIV and NON-HIV TB cases.



Mortality rate for TB documented HIV cases has decreased from **31% to 13%** with respect to total mortality rate from 2000 to 2019 respectively. Whereas, in last 2 decades, mortality for TB excluding HIV cases shows increase from **68% to 86%** with respect to total mortality cases.

It can be concluded that, Over the years, Overall mortality rate as well as TB with documented HIV mortality rate have been decreased. Whereas, Non-HIV TB mortality rate has been increased.

**Note:** Reason for Decrease in HIV mortality rates w.r.t total mortality rate could be either due to effective treatments or better preventive measures.