

Life Expectancy Analysis

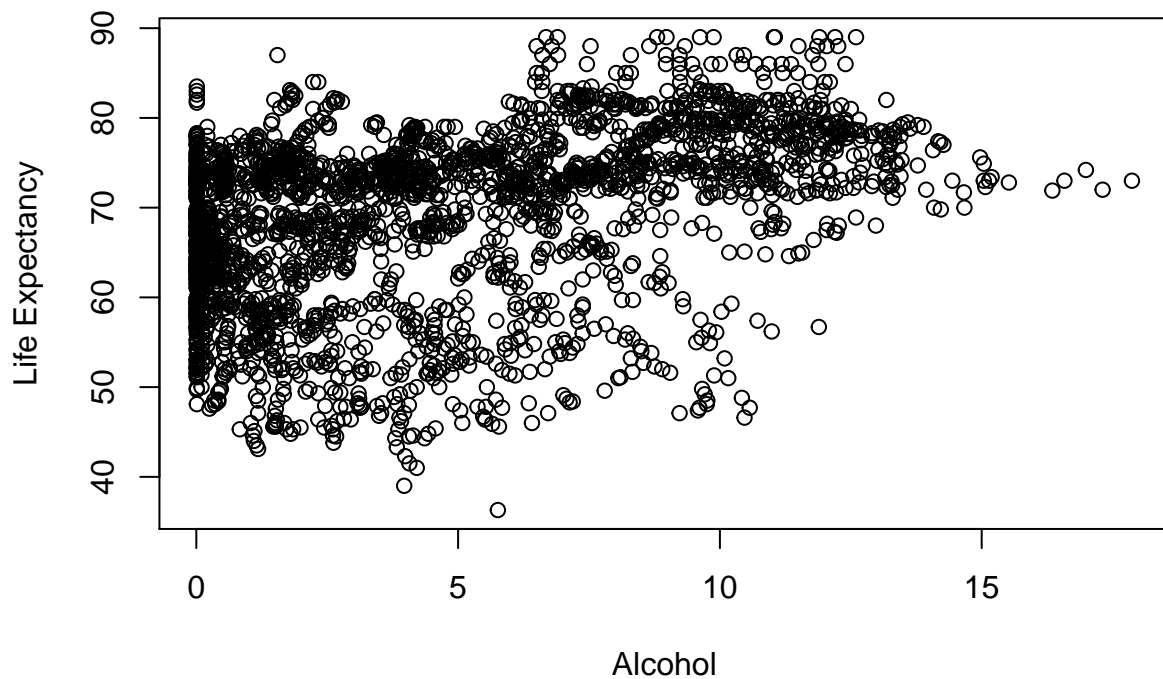
VMC

4/27/2021

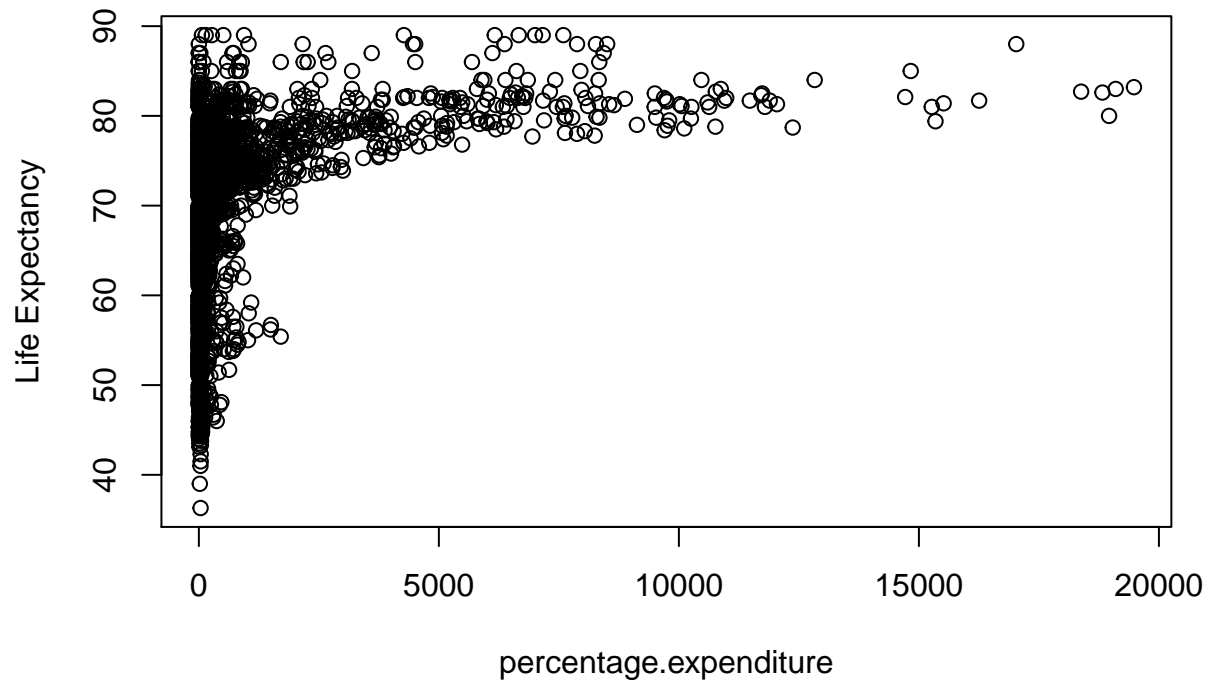
```
library(rmarkdown)
library(ggplot2)
life_expectancy_data <- read.csv("Life Expectancy Data.csv", header = TRUE)
```

Scatter Plots

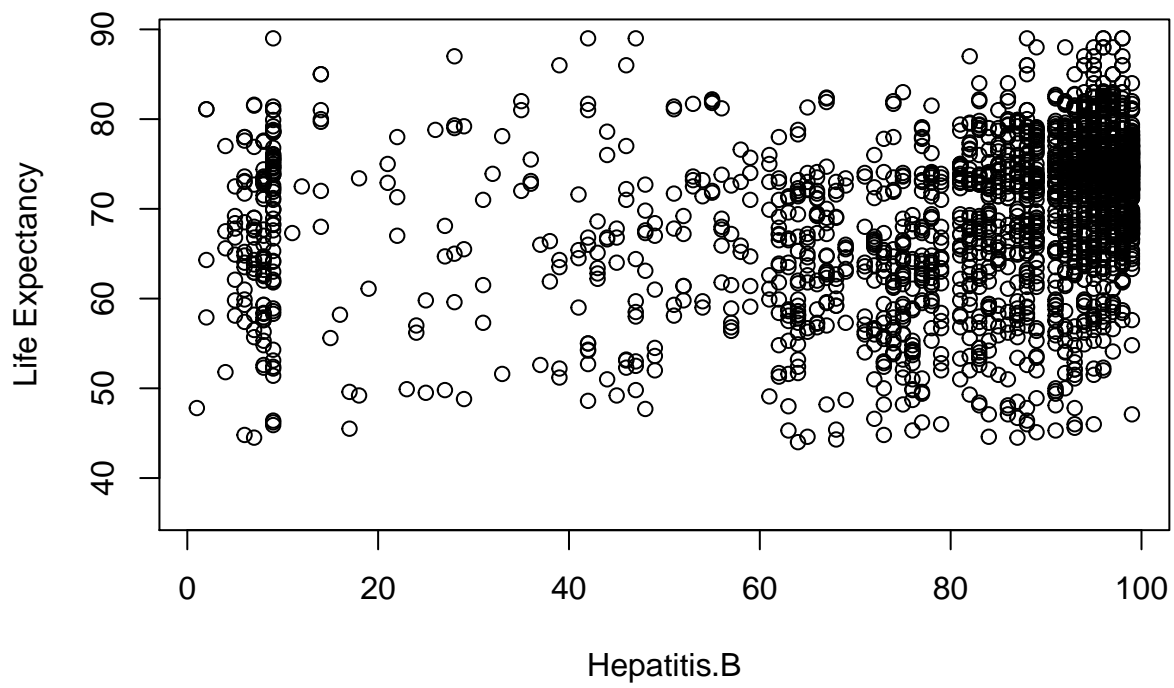
```
#ggplot(life_expectancy_data, aes(Year, Life.expectancy, fill = Life.expectancy)) + geom_point(shape=21)
life_expectancy_data$Developed <- as.integer(as.logical(life_expectancy_data$Status == "Developed"))
#ggplot(life_expectancy_data, Developed, Life.expectancy, ylab = "Life Expectancy", xlab = "0 = Developed")
plot(life_expectancy_data$Alcohol, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy", xlab = "Alcohol")
```



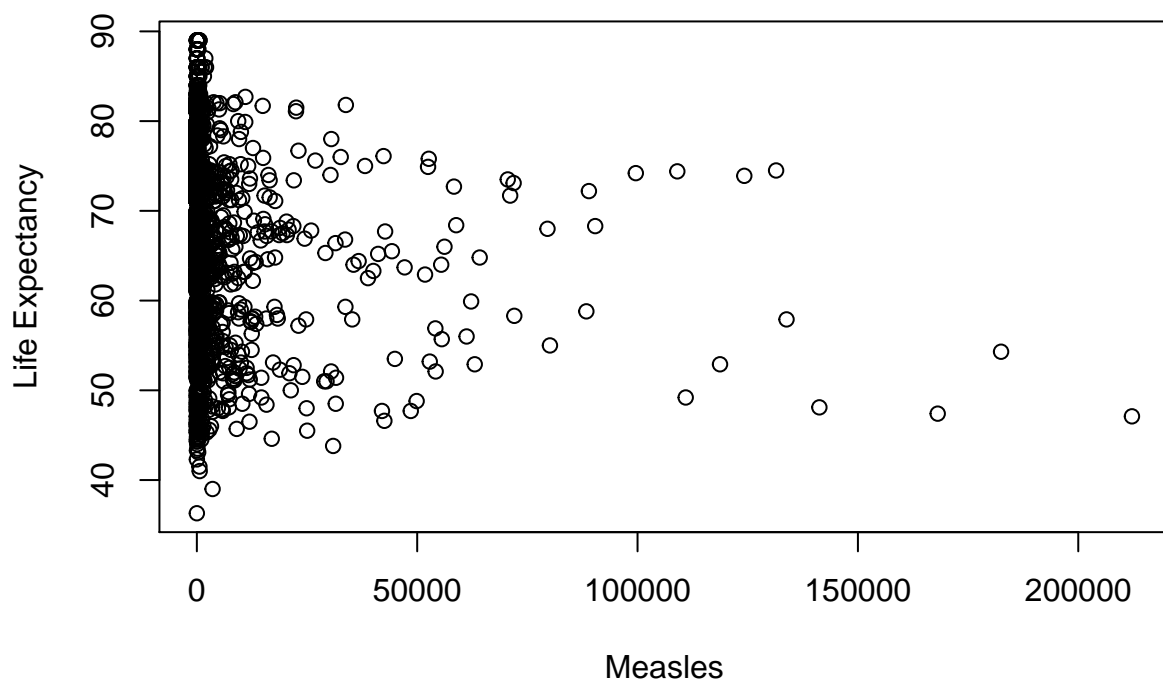
```
plot (life_expectancy_data$percentage.expenditure, life_expectancy_data$Life.expectancy, ylab = "Life E
```



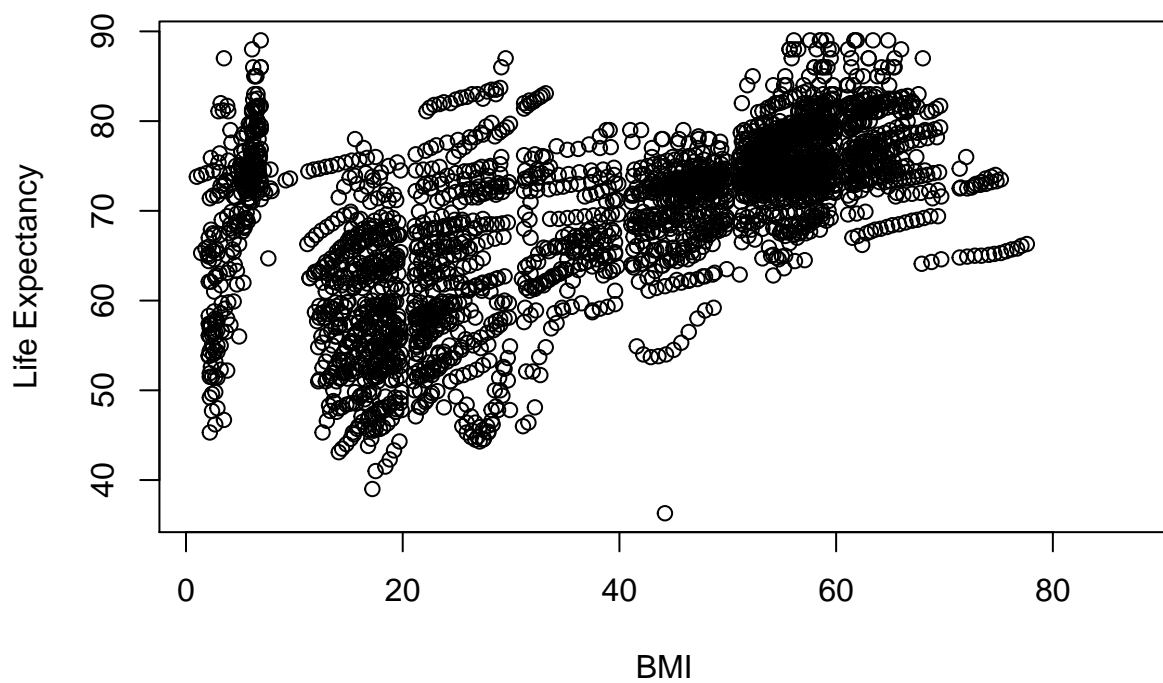
```
plot (life_expectancy_data$Hepatitis.B, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy",
```



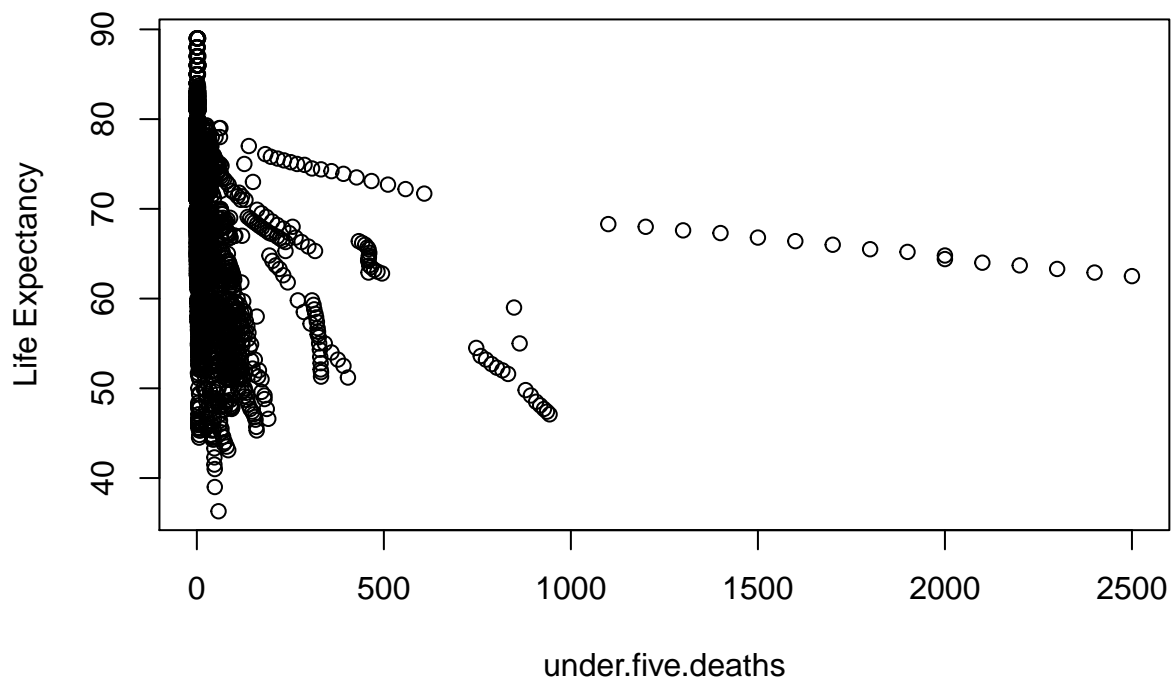
```
plot (life_expectancy_data$Measles, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy", xlab = "Hepatitis.B")
```



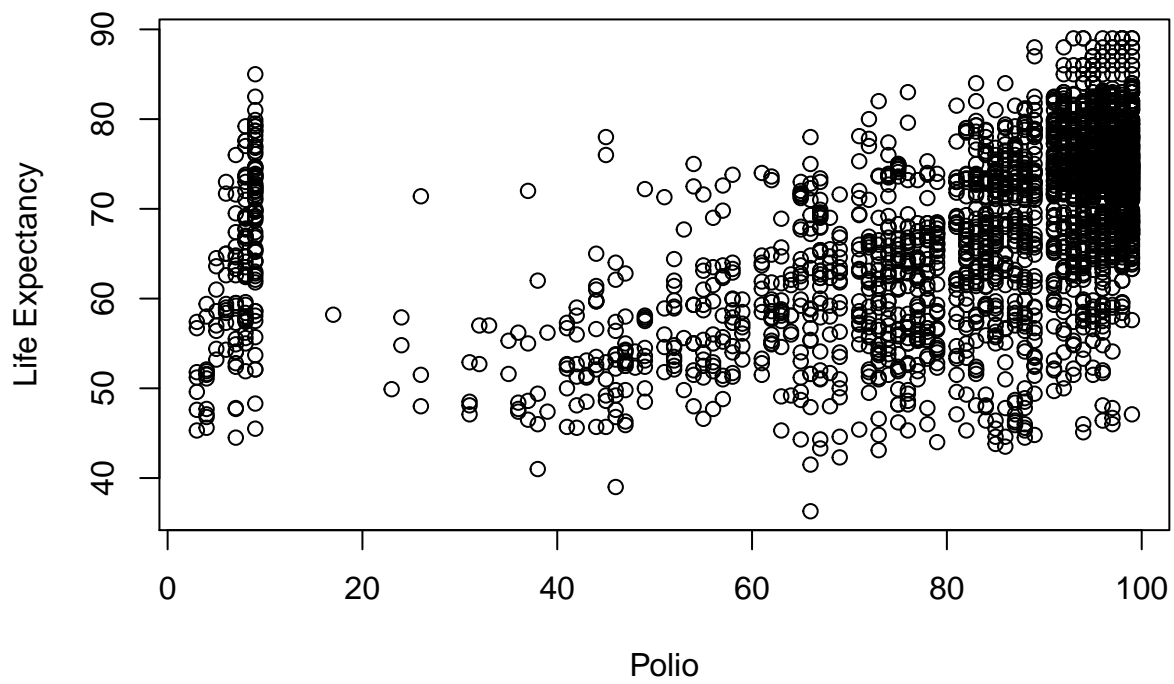
```
plot (life_expectancy_data$BMI, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy", xlab =
```



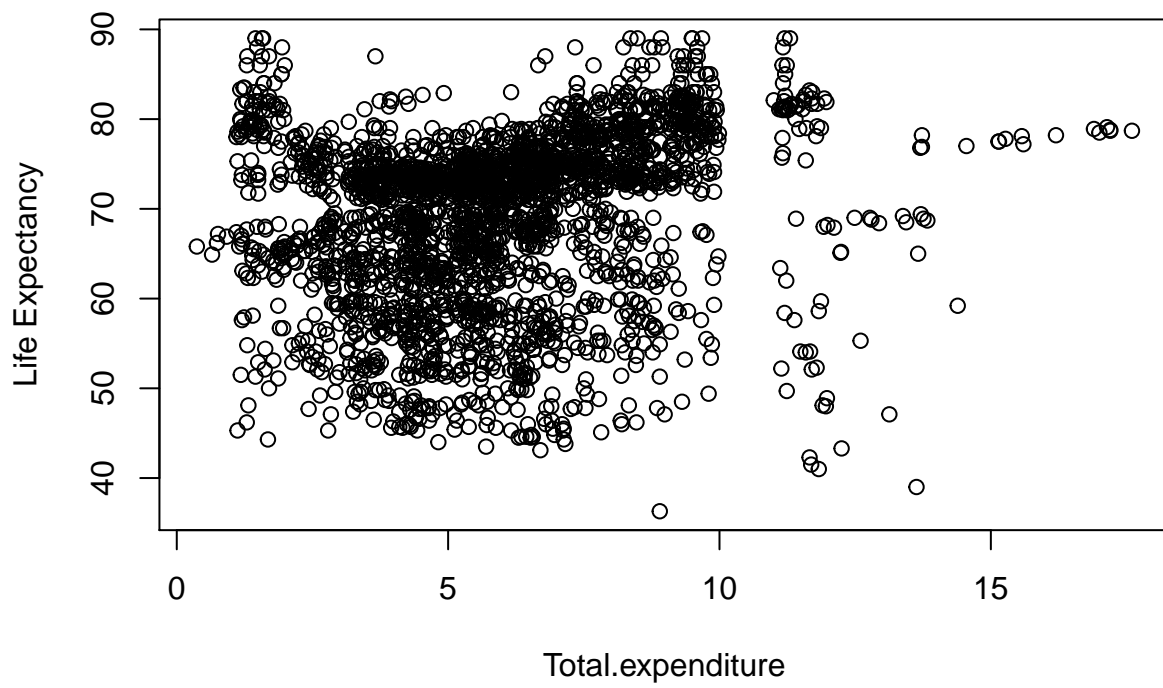
```
plot (life_expectancy_data$under.five.deaths, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy")
```



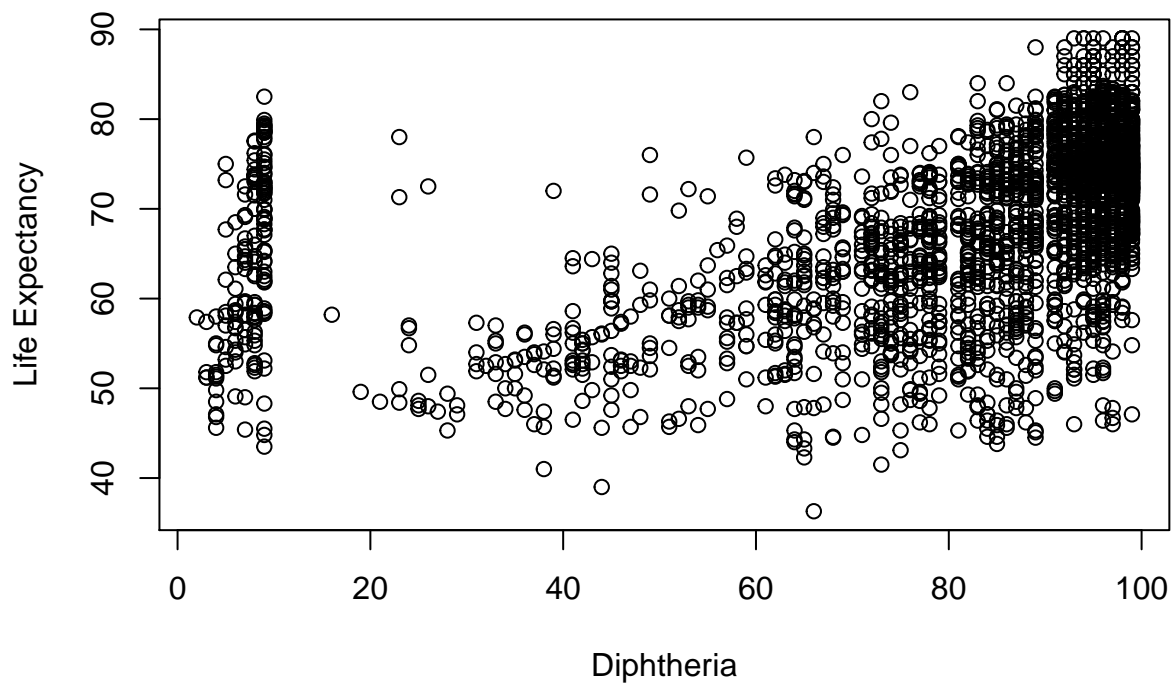
```
plot (life_expectancy_data$Polio, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy", xlab = "under.five.deaths")
```



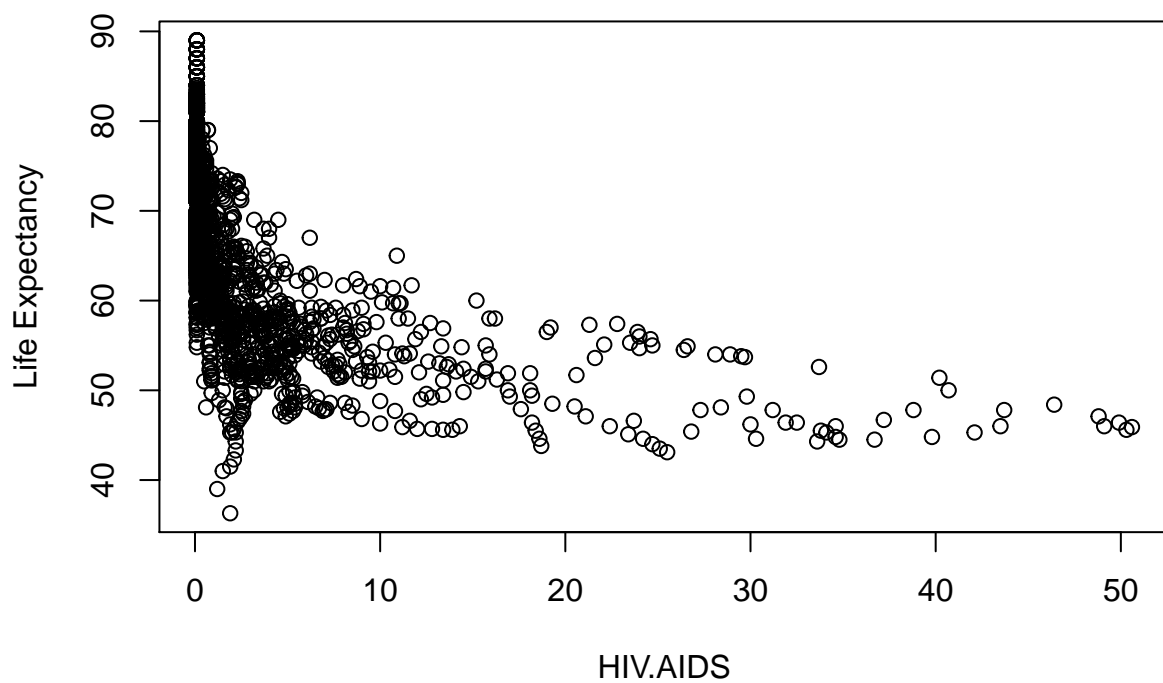
```
plot (life_expectancy_data$Total.expenditure, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy")
```



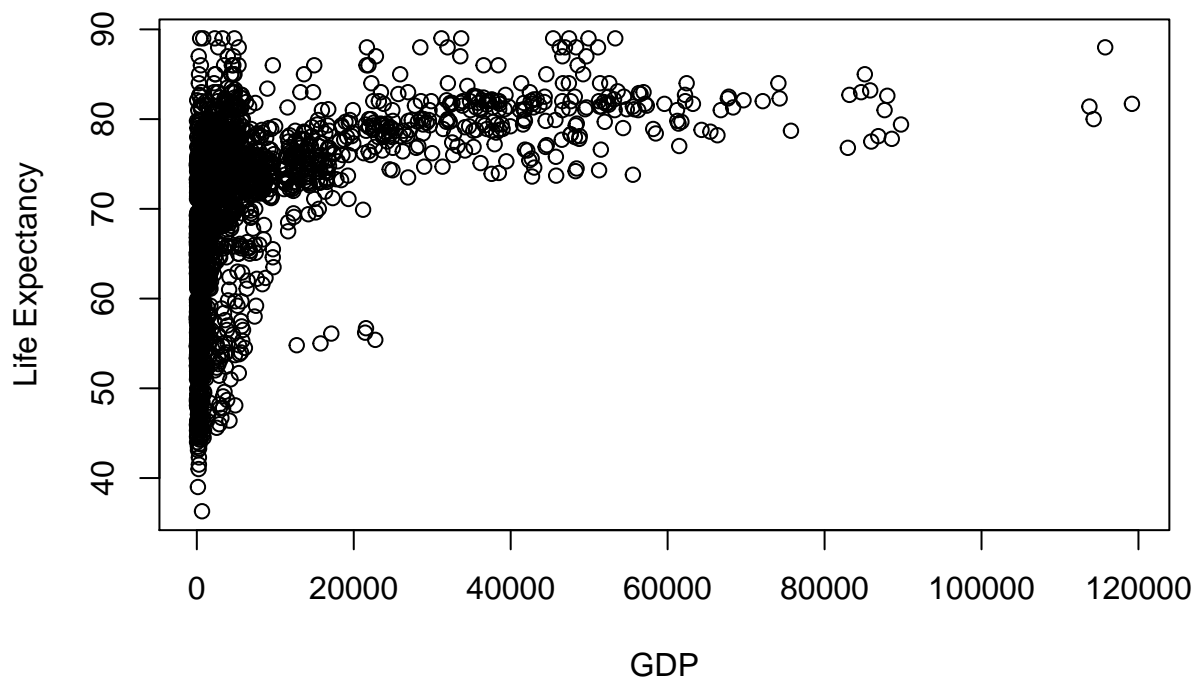
```
plot (life_expectancy_data$Diphtheria, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy",
```

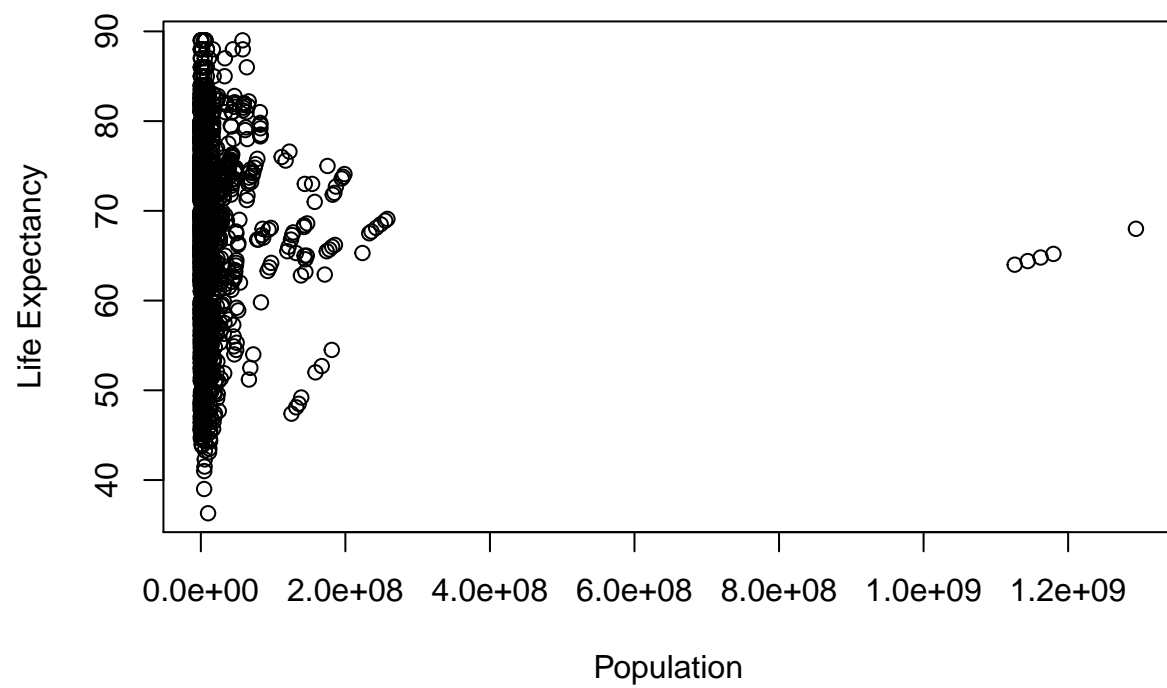
```
plot (life_expectancy_data$HIV.AIDS, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy", xlab = "Diphtheria")
```



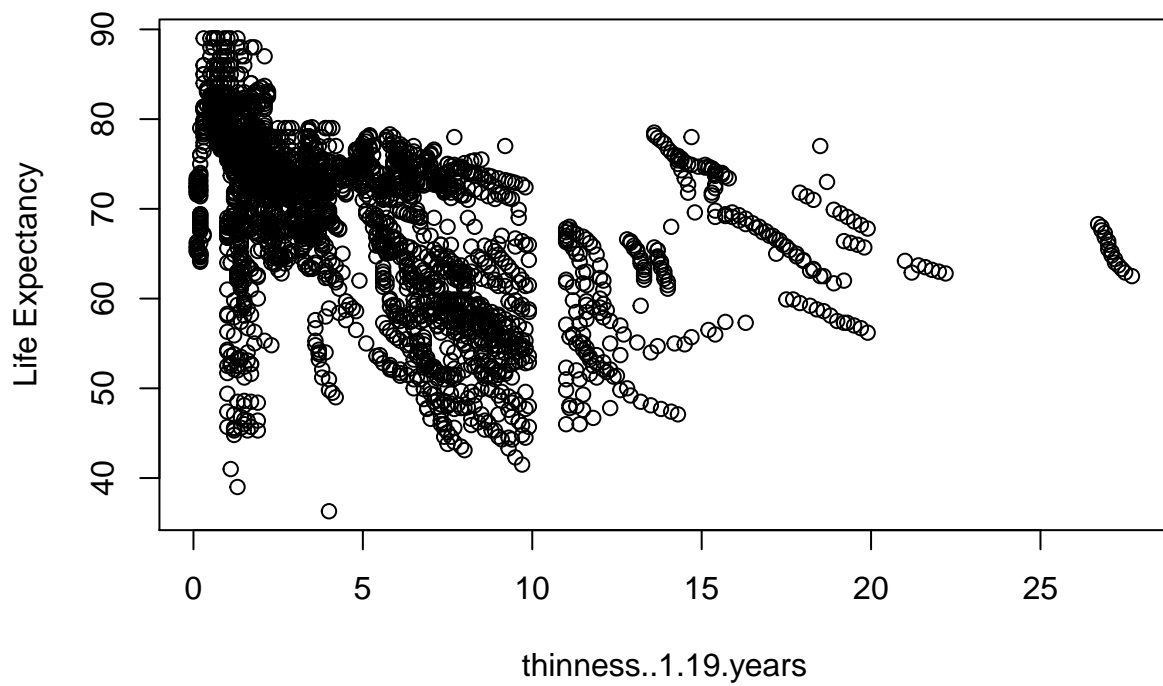
```
plot (life_expectancy_data$GDP, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy", xlab =
```



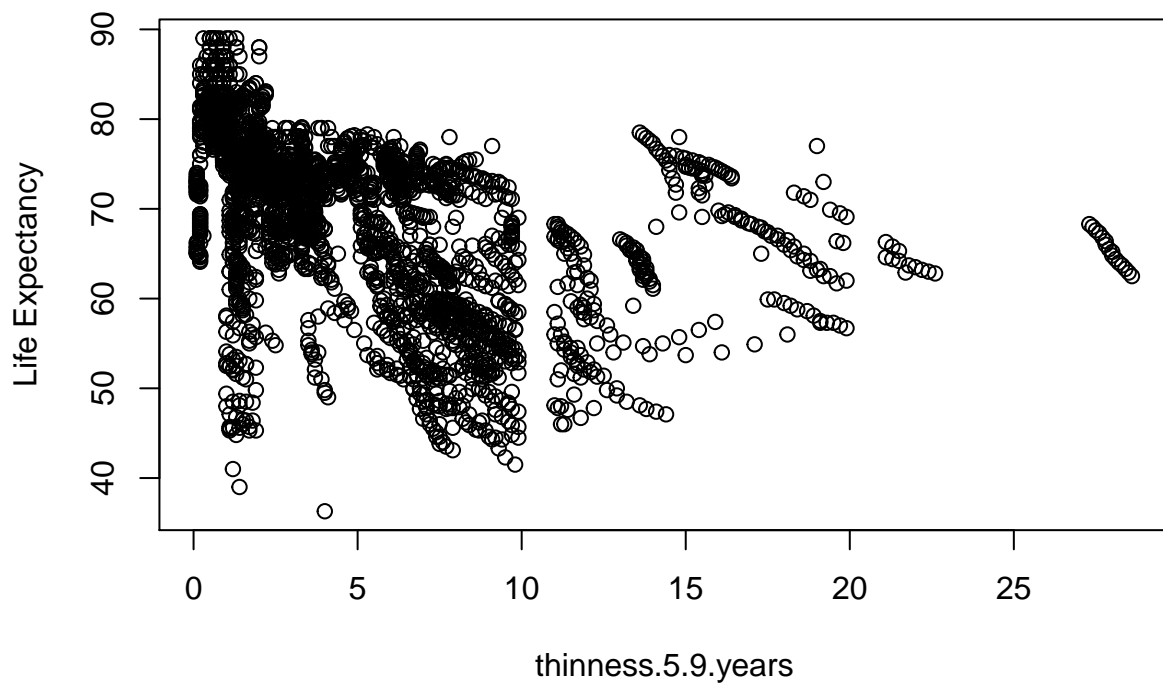
```
plot (life_expectancy_data$Population, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy",
```



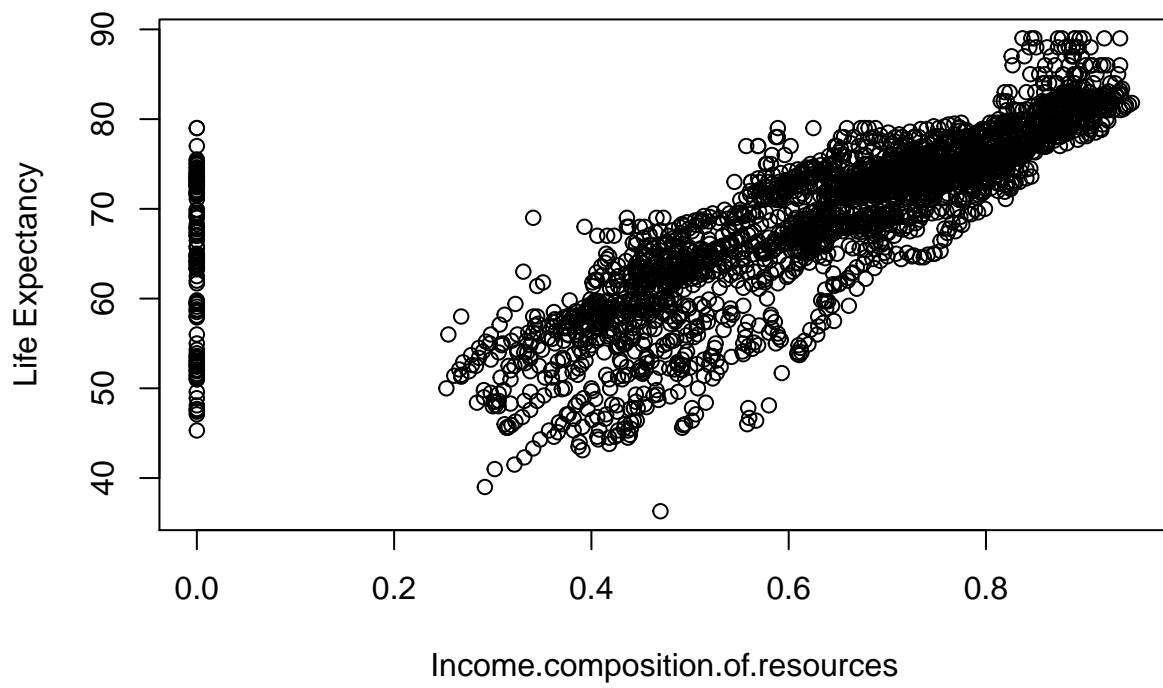
```
plot (life_expectancy_data$thinness..1.19.years, life_expectancy_data$Life.expectancy, ylab = "Life Exp
```



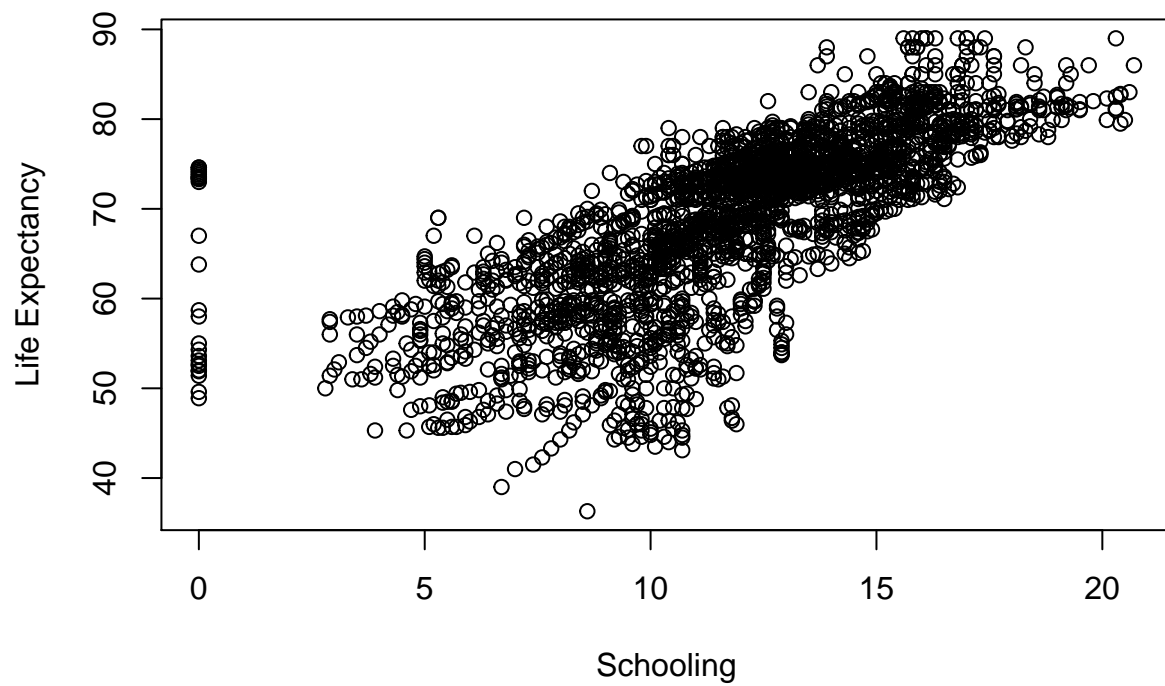
```
plot (life_expectancy_data$thinness.5.9.years, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy")
```



```
plot (life_expectancy_data$Income.composition.of.resources, life_expectancy_data$Life.expectancy, ylab =
```



```
plot (life_expectancy_data$Schooling, life_expectancy_data$Life.expectancy, ylab = "Life Expectancy", xlab = "Income.composition.of.resources")
```



Histograms

```
hist (life_expectancy_data$Life.expectancy, xlab="Life Expectancy", ylab="Frequency of Life Expectancy"
```




Fitting the model

In the following step, we will perform hypothesis testing to figure out variables that have a high linear correlation with life expectancy to arrive at a fitted multiple linear regression model

Residual Analysis

In the following step, we will perform residual analysis, to figure out outliers and variables that contribute for and against the accuracy of our model.

this is updated text