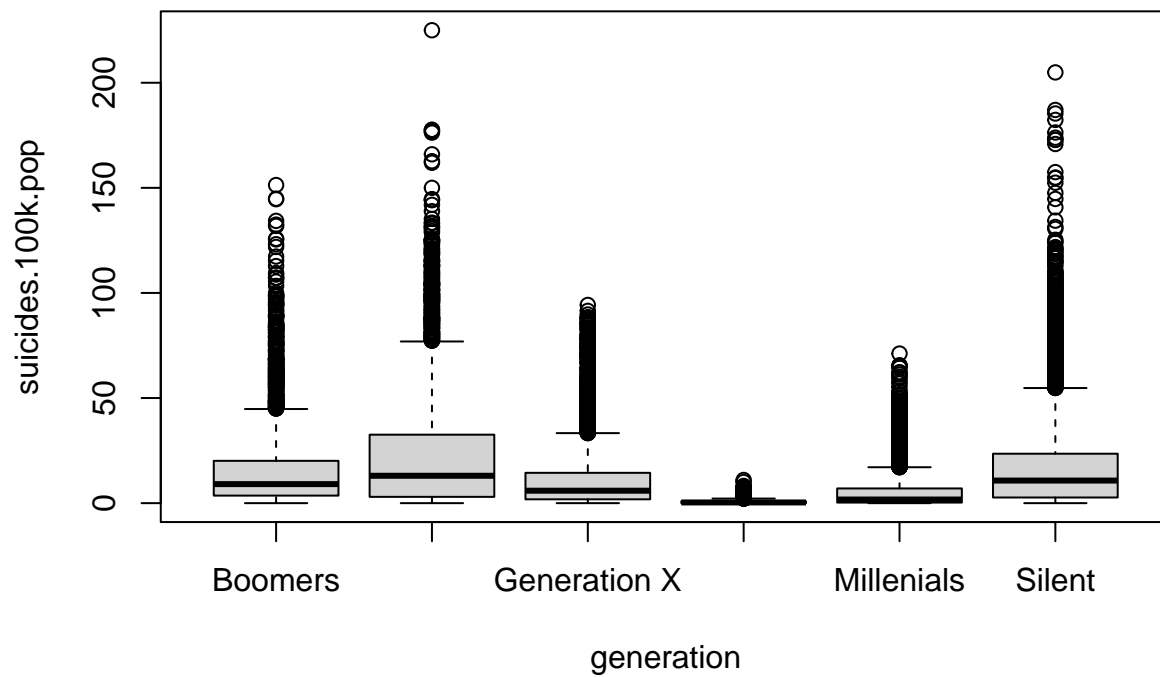


## 209 Project

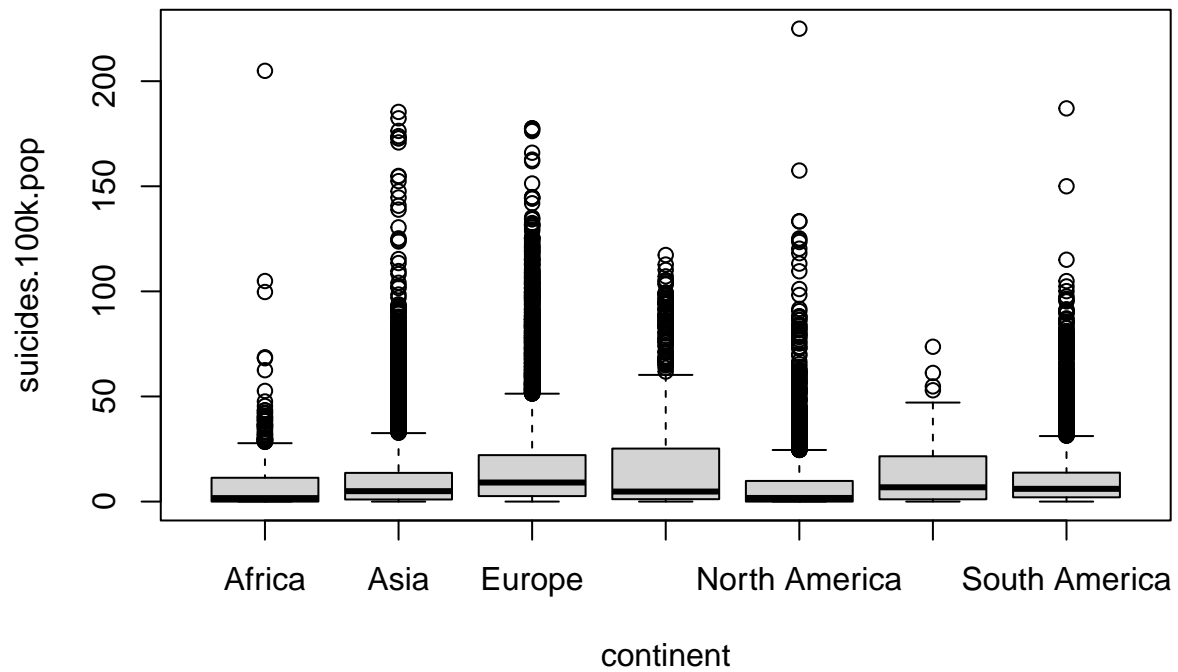
```
library(ggplot2)
```

```
dat <- read.csv("data_cleaned.csv")  
suicide <- read.csv("suicide_cleaned.csv")  
happy <- read.csv("Happiness_cleaned.csv")
```

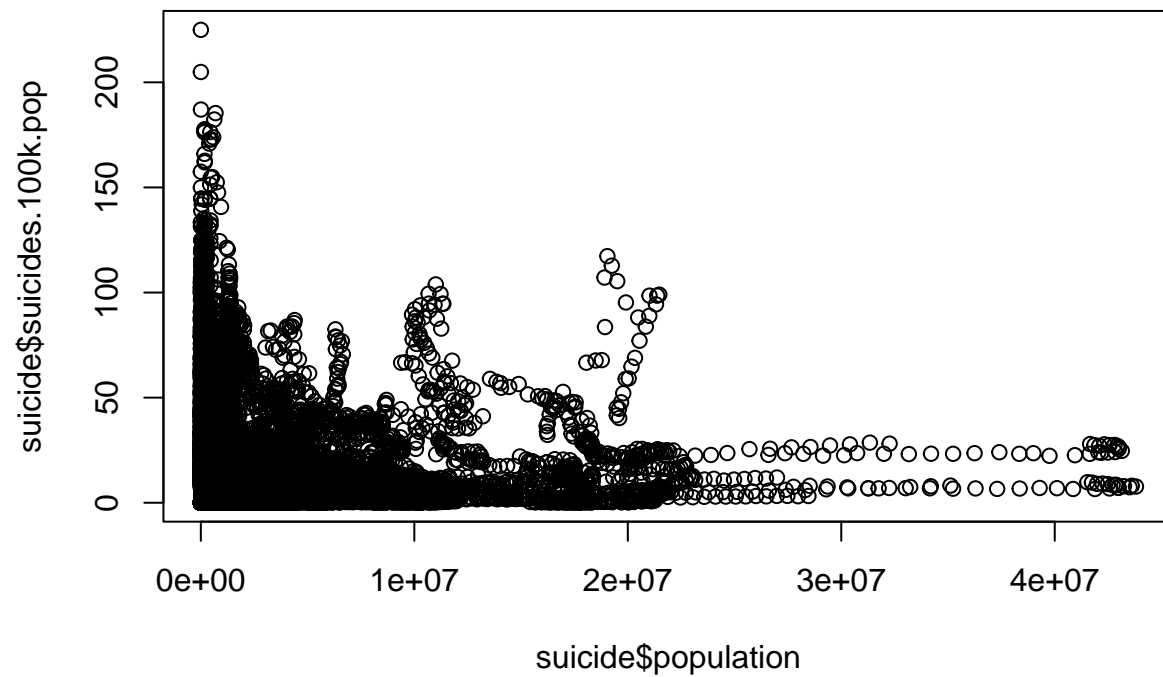
```
boxplot(suicides.100k.pop ~ generation, data = suicide)
```



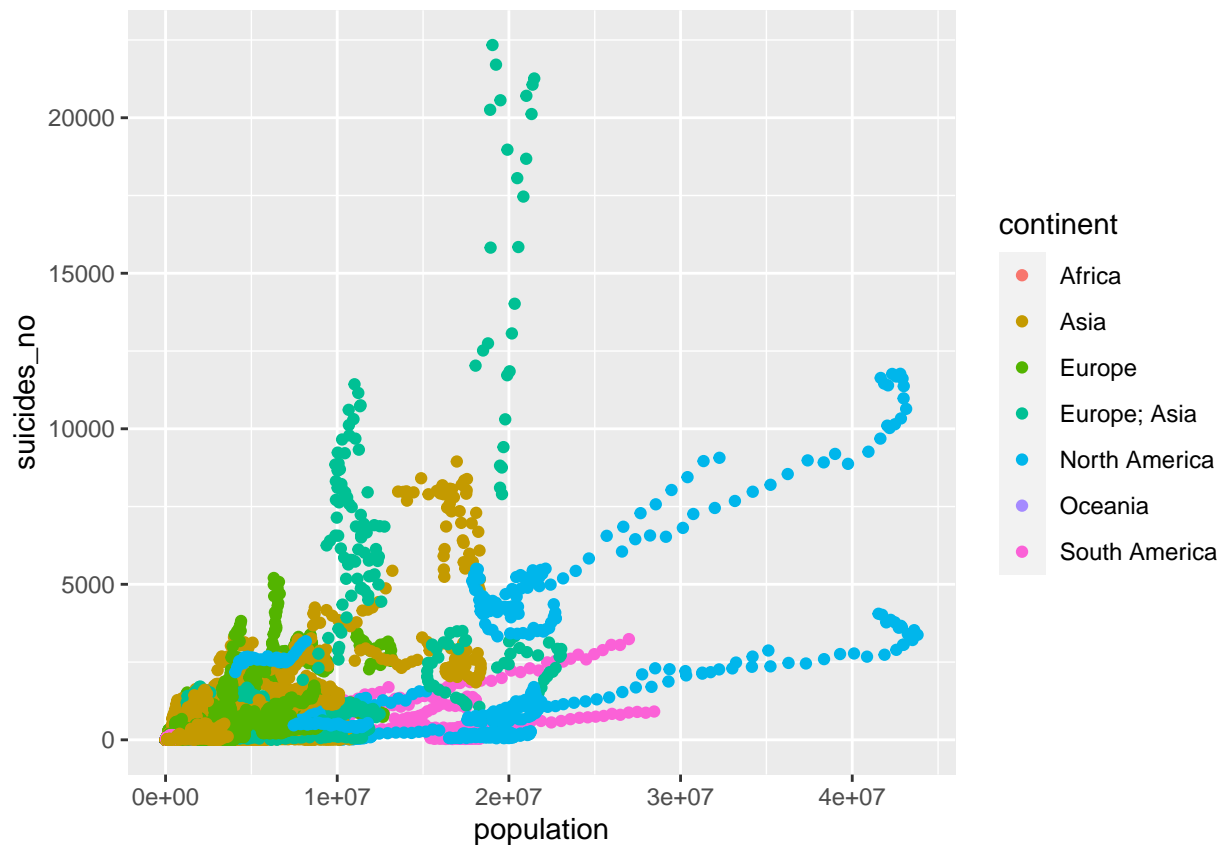
```
boxplot(suicides.100k.pop ~ continent, data = suicide)
```



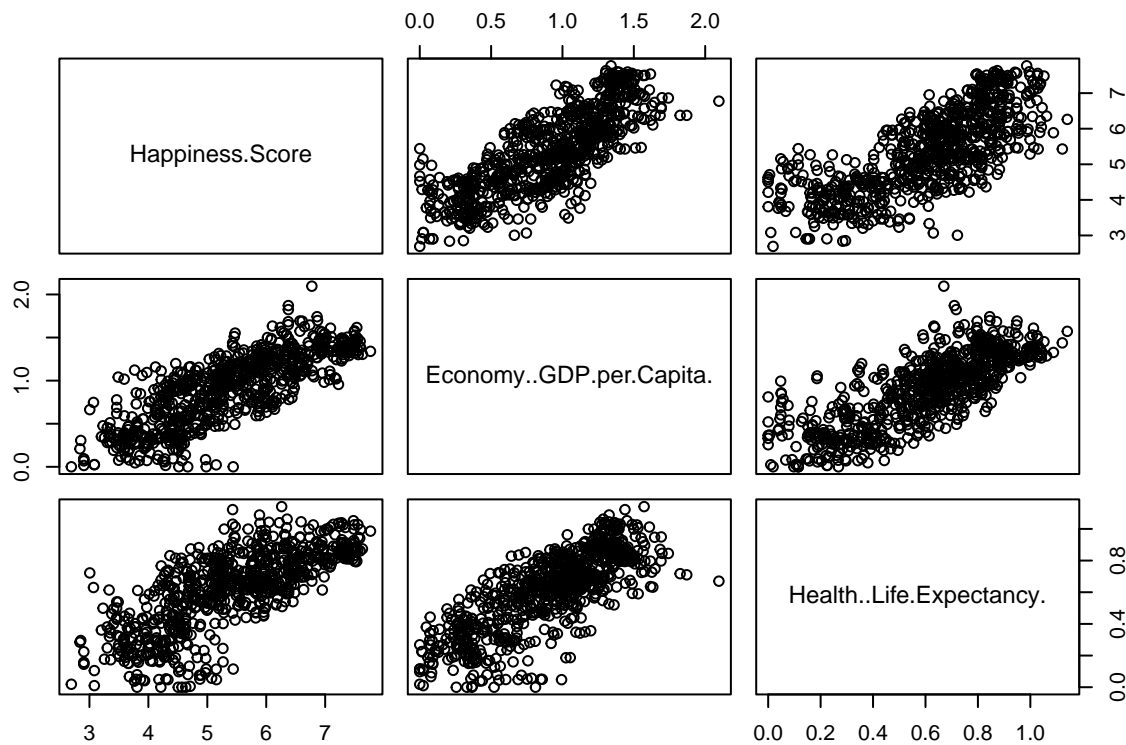
```
plot(suicide$population, suicide$suicides.100k.pop)
```



```
ggplot(data = suicide, aes(x = population, y = suicides_no, col = continent)) + geom_point()
```



```
plot(happy[c("Happiness.Score", "Economy..GDP.per.Capita.", "Health..Life.Expectancy.")])
```



```
Multiple.Model <- lm(Happiness.Score ~ Economy..GDP.per.Capita. + Health..Life.Expectancy., data = happy)
Cm <- Multiple.Model$coefficients
```

```

Y.Pred.Multiple <- Cm[1] + Cm[2]*happy$Economy..GDP.per.Capita. + Cm[2]*happy$Health..Life.Expectancy.
Y.True <- happy$Happiness.Score
Economy.Model <- lm(Happiness.Score ~ Economy..GDP.per.Capita., data = happy)
Ce <- Economy.Model$coefficients
Life.Model <- lm(Happiness.Score ~ Health..Life.Expectancy., data = happy)
Cl <- Life.Model$coefficients
Y.Pred.Economy <- Ce[1] + Ce[2] * happy$Economy..GDP.per.Capita.
Y.Pred.Life <- Cl[1] + Cl[2] * happy$Health..Life.Expectancy.
cor(Y.Pred.Multiple, Y.True)

```

```
## [1] 0.8164584
```

```
cor(Y.Pred.Life, Y.True)
```

```
## [1] 0.7440825
```

```
cor(Y.Pred.Economy, Y.True)
```

```
## [1] 0.7919431
```

```
Cm
```

```
##          (Intercept) Economy..GDP.per.Capita. Health..Life.Expectancy.
##          3.103927          1.506355          1.460409
```

We can see here that the Multiple Regression Model fits better than with just GDP per Capita or Life Expectancy. We get the formula of Happiness Index as

$HappinessIndex \approx 3.123 + 1.489 * GDP + 1.456 * LifeExpectancy$

```
boxplot(Happiness.Score ~ Continent, data = happy)
```

