

# STA 160 Project

Gabriel Jones

2024-11-04

## Libraries

```
##  
## Attaching package: 'dplyr'  
  
## The following object is masked from 'package:MASS':  
##  
##     select  
  
## The following objects are masked from 'package:stats':  
##  
##     filter, lag  
  
## The following objects are masked from 'package:base':  
##  
##     intersect, setdiff, setequal, union  
  
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --  
## vforcats    1.0.0    vreadr      2.1.4  
## vggplot2    3.4.3    vstringr    1.5.0  
## vlubridate  1.9.2    vtibble     3.2.1  
## vpurrr      1.0.2    vtidyr     1.3.0  
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()   masks stats::lag()  
## x dplyr::select() masks MASS::select()  
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors  
  
## Warning: package 'leaps' was built under R version 4.3.2  
  
## NOTE: Either Arial Narrow or Roboto Condensed fonts are required to use these themes.  
##       Please use hrbrthemes::import_roboto_condensed() to install Roboto Condensed and  
##       if Arial Narrow is not on your system, please see https://bit.ly/arialnarrow  
  
## Warning: package 'glmnet' was built under R version 4.3.3  
  
## Loading required package: Matrix  
##  
## Attaching package: 'Matrix'
```

```

## 
## The following objects are masked from 'package:tidyverse':
##   expand, pack, unpack
## 
## Loaded glmnet 4.1-8
## corrplot 0.92 loaded

```

## Load Data

```

data <- read.csv("Life Expectancy Data.csv")
data <- na.omit(data)

```

```
glimpse(data)
```

```

## Rows: 1,649
## Columns: 22
## $ Country <chr> "Afghanistan", "Afghanistan", "Afghani-
## $ Year <int> 2015, 2014, 2013, 2012, 2011, 2010, 20-
## $ Status <chr> "Developing", "Developing", "Developin-
## $ Life.expectancy <dbl> 65.0, 59.9, 59.9, 59.5, 59.2, 58.8, 58-
## $ Adult.Mortality <int> 263, 271, 268, 272, 275, 279, 281, 287-
## $ infant.deaths <int> 62, 64, 66, 69, 71, 74, 77, 80, 82, 84-
## $ Alcohol <dbl> 0.01, 0.01, 0.01, 0.01, 0.01, 0.01, 0.-
## $ percentage.expenditure <dbl> 71.279624, 73.523582, 73.219243, 78.18-
## $ Hepatitis.B <int> 65, 62, 64, 67, 68, 66, 63, 64, 63, 64-
## $ Measles <int> 1154, 492, 430, 2787, 3013, 1989, 2861-
## $ BMI <dbl> 19.1, 18.6, 18.1, 17.6, 17.2, 16.7, 16-
## $ under.five.deaths <int> 83, 86, 89, 93, 97, 102, 106, 110, 113-
## $ Polio <int> 6, 58, 62, 67, 68, 66, 63, 64, 63, 58, ~
## $ Total.expenditure <dbl> 8.16, 8.18, 8.13, 8.52, 7.87, 9.20, 9.-
## $ Diphtheria <int> 65, 62, 64, 67, 68, 66, 63, 64, 63, 58-
## $ HIV.AIDS <dbl> 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1-
## $ GDP <dbl> 584.25921, 612.69651, 631.74498, 669.9-
## $ Population <dbl> 33736494, 327582, 31731688, 3696958, 2-
## $ thinness..1.19.years <dbl> 17.2, 17.5, 17.7, 17.9, 18.2, 18.4, 18-
## $ thinness.5.9.years <dbl> 17.3, 17.5, 17.7, 18.0, 18.2, 18.4, 18-
## $ Income.composition.of.resources <dbl> 0.479, 0.476, 0.470, 0.463, 0.454, 0.4-
## $ Schooling <dbl> 10.1, 10.0, 9.9, 9.8, 9.5, 9.2, 8.9, 8-

```

```
count(data, Country)
```

	Country	n
## 1	Afghanistan	16
## 2	Albania	16
## 3	Algeria	11
## 4	Angola	8
## 5	Argentina	13
## 6	Armenia	15
## 7	Australia	14
## 8	Austria	15
## 9	Azerbaijan	13

```
## 10          Bangladesh 12
## 11          Belarus 15
## 12          Belgium 15
## 13          Belize 15
## 14          Benin 13
## 15          Bhutan 15
## 16          Bosnia and Herzegovina 11
## 17          Botswana 15
## 18          Brazil 15
## 19          Bulgaria 15
## 20          Burkina Faso 9
## 21          Burundi 11
## 22          Cabo Verde 13
## 23          Cambodia 9
## 24          Cameroon 10
## 25          Canada 12
## 26          Central African Republic 6
## 27          Chad 7
## 28          Chile 9
## 29          China 15
## 30          Colombia 15
## 31          Comoros 12
## 32          Costa Rica 15
## 33          Croatia 8
## 34          Cyprus 15
## 35          Djibouti 8
## 36          Dominican Republic 15
## 37          Ecuador 15
## 38          El Salvador 15
## 39          Equatorial Guinea 1
## 40          Eritrea 10
## 41          Estonia 11
## 42          Ethiopia 8
## 43          Fiji 15
## 44          France 15
## 45          Gabon 10
## 46          Georgia 15
## 47          Germany 15
## 48          Ghana 13
## 49          Greece 15
## 50          Guatemala 10
## 51          Guinea 8
## 52          Guinea-Bissau 6
## 53          Guyana 14
## 54          Haiti 2
## 55          Honduras 15
## 56          India 11
## 57          Indonesia 15
## 58          Iraq 11
## 59          Ireland 5
## 60          Israel 15
## 61          Italy 15
## 62          Jamaica 12
## 63          Jordan 15
```

## 64 Kazakhstan 15  
## 65 Kenya 13  
## 66 Kiribati 15  
## 67 Latvia 15  
## 68 Lebanon 15  
## 69 Lesotho 12  
## 70 Liberia 7  
## 71 Lithuania 15  
## 72 Luxembourg 15  
## 73 Madagascar 13  
## 74 Malawi 13  
## 75 Malaysia 15  
## 76 Maldives 15  
## 77 Mali 12  
## 78 Malta 12  
## 79 Mauritania 10  
## 80 Mauritius 15  
## 81 Mexico 15  
## 82 Mongolia 15  
## 83 Montenegro 9  
## 84 Morocco 15  
## 85 Mozambique 14  
## 86 Myanmar 12  
## 87 Namibia 5  
## 88 Nepal 12  
## 89 Netherlands 4  
## 90 Nicaragua 15  
## 91 Niger 6  
## 92 Nigeria 10  
## 93 Pakistan 12  
## 94 Panama 14  
## 95 Papua New Guinea 15  
## 96 Paraguay 12  
## 97 Peru 12  
## 98 Philippines 15  
## 99 Poland 15  
## 100 Portugal 15  
## 101 Romania 15  
## 102 Russian Federation 14  
## 103 Rwanda 13  
## 104 Samoa 15  
## 105 Sao Tome and Principe 12  
## 106 Senegal 11  
## 107 Serbia 11  
## 108 Seychelles 15  
## 109 Sierra Leone 8  
## 110 Solomon Islands 15  
## 111 South Africa 15  
## 112 Spain 15  
## 113 Sri Lanka 12  
## 114 Suriname 10  
## 115 Swaziland 15  
## 116 Sweden 4  
## 117 Syrian Arab Republic 8

```

## 118 Tajikistan 13
## 119 Thailand 15
## 120 Timor-Leste 7
## 121 Togo 7
## 122 Tonga 15
## 123 Trinidad and Tobago 13
## 124 Tunisia 15
## 125 Turkey 15
## 126 Turkmenistan 13
## 127 Uganda 13
## 128 Ukraine 15
## 129 Uruguay 15
## 130 Uzbekistan 13
## 131 Vanuatu 15
## 132 Zambia 10
## 133 Zimbabwe 15

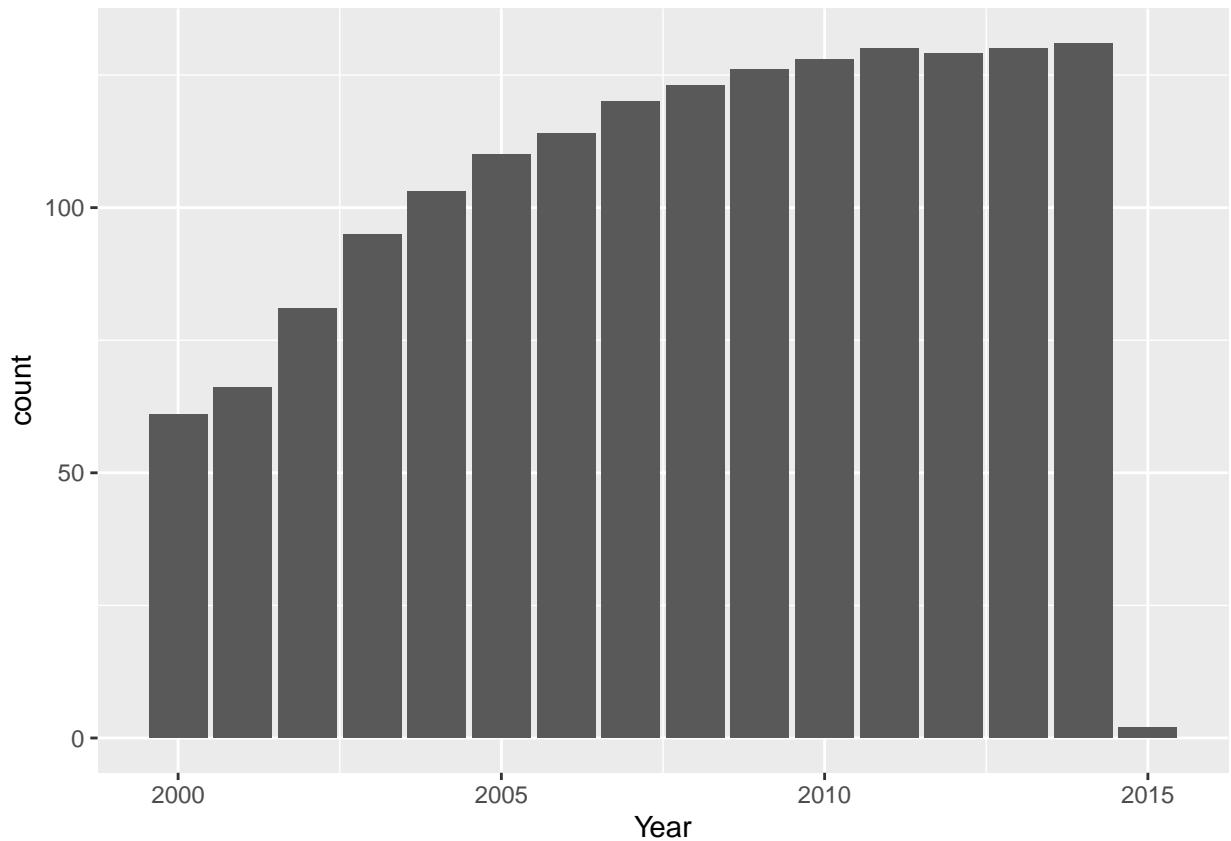
```

### Yearly data points

```

data%>%
  ggplot(aes(Year))+
  geom_bar()

```

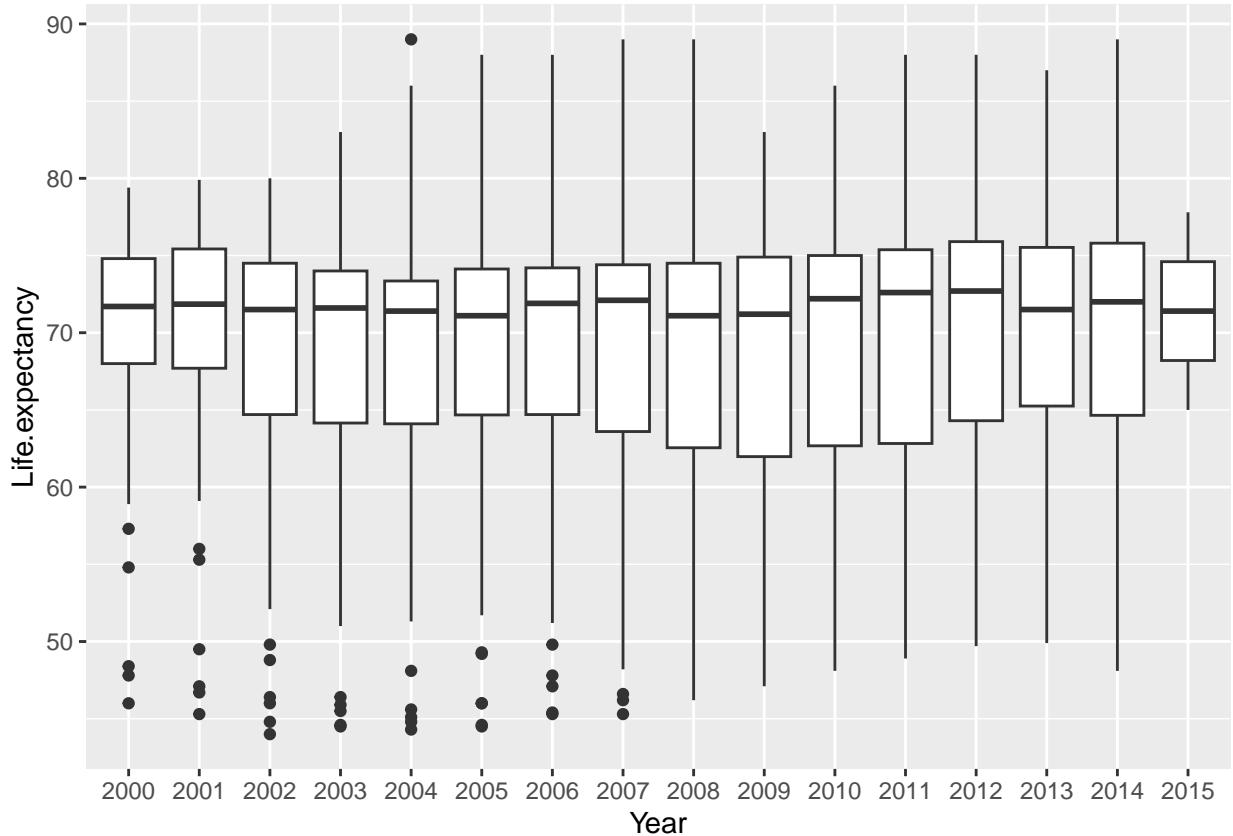


Overall Life Expectancy by Year

```

data %>%
  mutate(Year=as.factor(Year)) %>%
  ggplot(aes(Year, Life.expectancy)) +
  geom_boxplot()

```



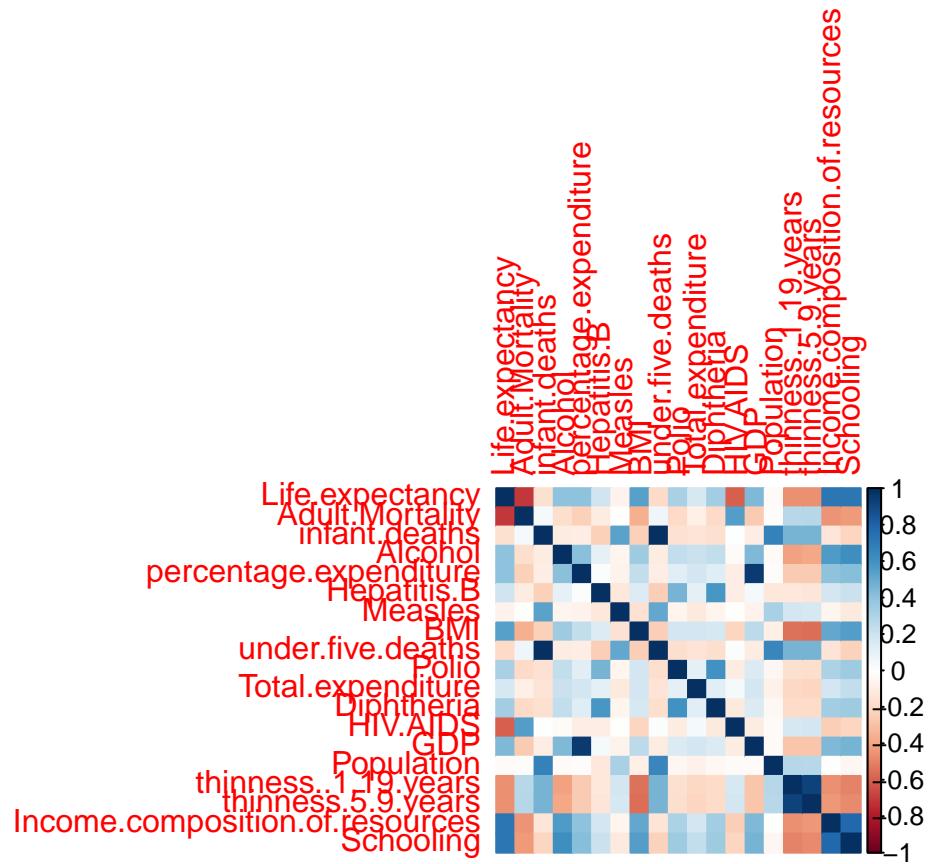
### Correlation Matrix

```

dataNumeric <- data %>%
  select(-"Country", -"Year", -"Status") %>%
  na.omit()

corrplot(cor(dataNumeric), method="color")

```



High Correlations include:

- Life expectancy & Adult Mortality (-0.70)
- Life expectancy & income (.72)
- Life expectancy & schooling (.72)
- GDP & percentage expenditure (.95)
- Income & schooling (.78) - thinness vars (.92)
- Diphtheria & Polio (0.61)
- Under five deaths & population (.65)
- infant deaths & population (0.67)
- Under five deaths & infant deaths (0.99)

Vars of interest may include: *Adult Mortality, HIV AIDS, Income, and Schooling.*

If the above are used, consider dropping either income or schooling

### Life Expectancy by Country Summary Statistics

```
# Numerical
sumstat <- data.frame("Country"=c(),
                      "Min."=c(),
```

```

    "X1st.Qu."=c(),
    "Median"=c(),
    "Mean"=c(),
    "X3rd.Qu."=c(),
    "Max."=c())
for (c in unique(data$Country)){
  country.dat <- data$Life.expectancy[data$Country==c]
  r <- cbind(data.frame("Country"=c(c)), data.frame(as.list(summary(country.dat))))
  sumstat <- rbind(sumstat, r)
}
kable(sumstat)

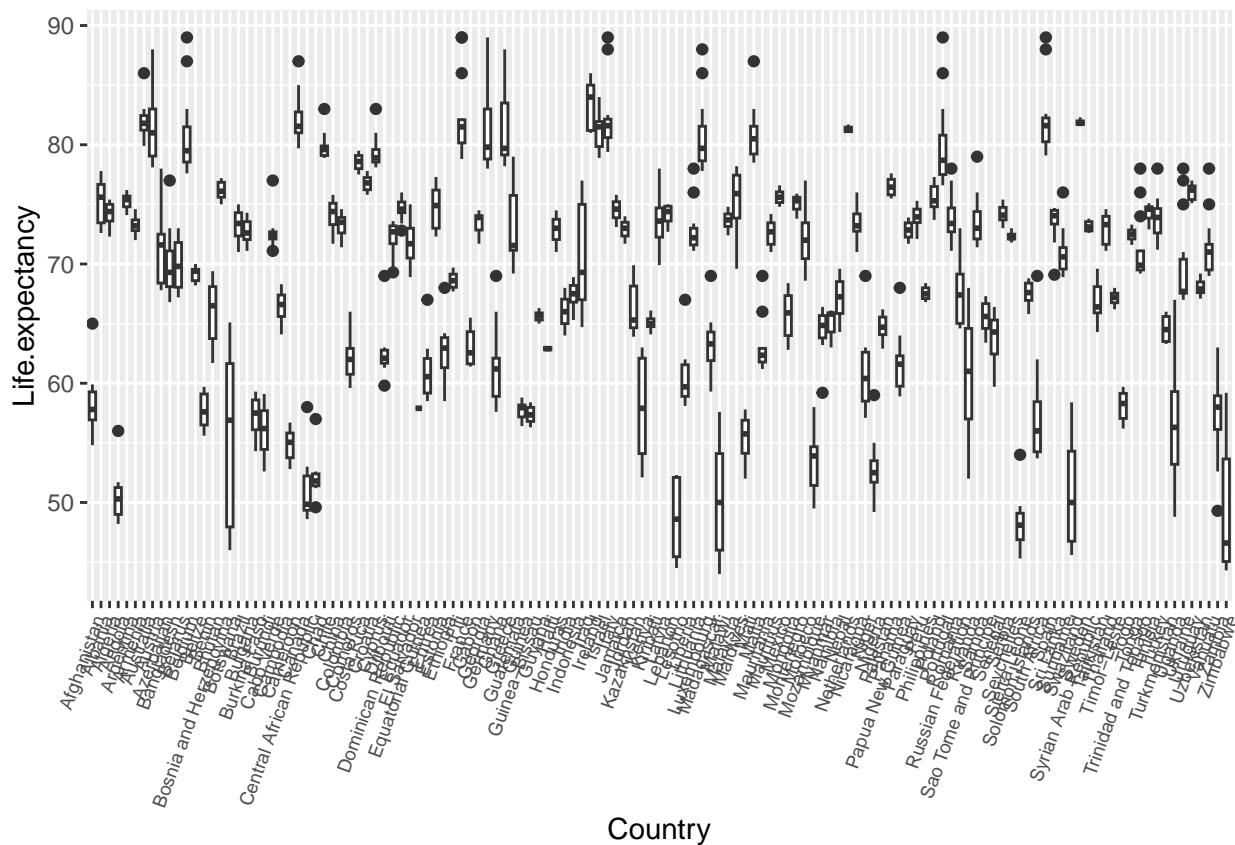
```

Country	Min.	X1st.Qu.	Median	Mean	X3rd.Qu.	Max.
Afghanistan	54.8	56.925	57.80	58.19375	59.275	65.0
Albania	72.6	73.450	75.60	75.15625	76.675	77.8
Algeria	72.3	73.600	74.40	74.20909	75.000	75.4
Angola	48.2	49.000	50.30	50.67500	51.250	56.0
Argentina	74.1	74.800	75.40	75.23846	75.700	76.2
Armenia	72.0	72.800	73.20	73.30667	73.700	74.6
Australia	79.9	81.225	81.80	81.90714	82.450	86.0
Austria	78.1	79.050	81.00	81.48000	83.000	88.0
Azerbaijan	67.8	68.400	71.60	71.14615	72.500	78.0
Bangladesh	66.8	68.100	69.30	69.96667	71.100	77.0
Belarus	67.2	68.050	69.80	69.74667	71.800	73.0
Belgium	77.6	78.550	79.50	80.65333	81.500	89.0
Belize	68.2	68.600	69.40	69.15333	69.550	70.0
Benin	55.6	56.500	57.60	57.70769	59.100	59.7
Bhutan	61.7	63.750	66.50	65.92000	68.100	69.4
Bosnia and Herzegovina	75.0	75.600	76.10	76.18182	76.850	77.2
Botswana	46.0	47.950	56.90	55.40667	61.650	65.1
Brazil	71.0	72.350	73.40	73.27333	74.300	75.0
Bulgaria	71.1	72.050	72.60	72.74000	73.550	74.3
Burkina Faso	54.3	56.100	57.50	57.23333	58.600	59.3
Burundi	52.6	54.450	56.20	56.02727	57.700	59.1
Cabo Verde	71.1	72.100	72.40	72.62308	72.700	77.0
Cambodia	64.1	65.600	66.60	66.43333	67.400	68.3
Cameroon	52.8	53.750	55.05	54.86000	55.825	56.7
Canada	79.7	81.000	81.55	82.23333	82.750	87.0
Central African Republic	48.6	49.350	49.85	51.41667	52.225	58.0
Chad	49.6	51.400	51.80	52.28571	52.400	57.0
Chile	78.9	79.100	79.60	79.94444	79.900	83.0
China	71.7	73.300	74.40	74.14000	75.100	75.8
Colombia	71.4	72.600	73.50	73.18667	73.900	74.6
Comoros	59.6	60.750	62.00	61.94167	62.925	66.0
Costa Rica	77.5	78.000	78.60	78.52667	79.100	79.5
Croatia	75.8	76.225	76.80	76.78750	77.250	77.8
Cyprus	78.1	78.550	78.90	79.32000	79.600	83.0
Djibouti	59.8	61.675	62.10	62.72500	62.775	69.0
Dominican Republic	69.3	71.700	72.70	72.24000	73.200	73.6
Ecuador	72.8	74.300	74.60	74.62667	75.200	76.0
El Salvador	68.9	70.500	71.70	71.62667	73.000	75.0
Equatorial Guinea	57.9	57.900	57.90	57.90000	57.900	57.9

Country	Min.	X1st.Qu.	Median	Mean	X3rd.Qu.	Max.
Eritrea	58.5	59.175	60.55	61.09000	62.075	67.0
Estonia	72.3	73.000	74.90	74.76364	76.200	77.3
Ethiopia	58.5	61.300	62.95	62.73750	63.825	68.0
Fiji	67.7	68.050	68.60	68.63333	69.150	69.7
France	78.8	80.150	81.50	82.20667	82.100	89.0
Gabon	61.4	61.625	62.55	63.00000	64.325	65.5
Georgia	71.7	72.850	73.90	73.44667	74.050	74.5
Germany	78.0	78.800	79.80	81.18667	83.000	89.0
Ghana	57.6	58.900	61.20	61.29231	62.100	69.0
Greece	78.2	79.150	79.70	81.23333	83.500	88.0
Guatemala	69.2	71.150	71.55	73.14000	75.750	79.0
Guinea	56.4	57.175	57.95	57.71250	58.175	58.8
Guinea-Bissau	56.3	56.800	57.35	57.36667	57.975	58.4
Guyana	65.0	65.300	65.65	65.61429	65.900	66.3
Haiti	62.7	62.800	62.90	62.90000	63.000	63.1
Honduras	71.0	72.050	73.00	72.88667	73.750	74.5
India	64.0	65.000	66.00	66.00000	67.050	68.0
Indonesia	65.3	66.800	67.50	67.45333	68.200	68.9
Iraq	64.7	67.000	69.30	70.39091	75.000	77.0
Ireland	81.0	81.200	84.00	83.44000	85.000	86.0
Israel	78.9	79.850	81.50	81.22000	81.950	84.0
Italy	79.4	80.600	81.60	82.15333	82.150	89.0
Jamaica	73.1	73.875	74.60	74.51667	75.225	75.8
Jordan	71.7	72.350	73.00	72.91333	73.500	74.0
Kazakhstan	63.9	64.650	65.30	66.41333	68.150	69.9
Kenya	52.1	54.100	57.90	57.88462	62.100	63.0
Kiribati	64.1	64.750	65.00	65.07333	65.400	66.1
Latvia	69.9	72.250	73.60	73.67333	74.700	78.0
Lebanon	72.7	73.600	74.40	74.15333	74.850	75.0
Lesotho	44.5	45.450	48.60	48.60833	52.100	52.3
Liberia	58.1	58.900	59.70	60.81429	61.550	67.0
Lithuania	71.1	71.600	72.20	72.75333	73.000	78.0
Luxembourg	77.8	78.650	79.70	80.70000	81.550	88.0
Madagascar	59.3	61.900	63.30	63.21538	64.300	69.0
Malawi	44.0	46.000	50.00	50.26154	54.100	57.6
Malaysia	72.4	73.150	73.70	73.67333	74.200	74.8
Maldives	69.6	73.850	75.90	75.34000	77.450	78.2
Mali	52.0	54.125	55.75	55.40000	56.900	57.8
Malta	78.5	79.225	80.50	80.88333	81.550	87.0
Mauritania	61.2	61.775	62.35	63.17000	62.925	69.0
Mauritius	71.0	71.650	72.70	72.58667	73.450	74.2
Mexico	74.8	75.150	75.60	75.65333	76.050	76.6
Mongolia	62.8	64.000	65.90	65.69333	67.350	68.4
Montenegro	73.8	74.600	75.30	75.06667	75.600	75.9
Morocco	68.6	70.450	72.00	72.01333	73.450	77.0
Mozambique	49.5	51.425	53.90	53.40714	54.675	58.0
Myanmar	59.2	63.800	64.85	64.43333	65.675	66.4
Namibia	63.0	64.300	65.80	65.02000	65.900	66.1
Nepal	64.3	65.850	67.25	67.14167	68.525	69.6
Netherlands	81.1	81.100	81.25	81.32500	81.475	81.7
Nicaragua	71.0	72.750	73.20	73.36000	74.200	76.0
Niger	57.1	58.500	60.40	61.35000	62.600	69.0

Country	Min.	X1st.Qu.	Median	Mean	X3rd.Qu.	Max.
Nigeria	49.2	51.700	52.50	52.84000	53.500	59.0
Pakistan	62.9	64.075	64.70	64.71667	65.550	66.2
Panama	75.5	75.800	76.45	76.45000	77.100	77.6
Papua New Guinea	58.9	59.750	61.60	61.60000	62.300	68.0
Paraguay	71.7	72.250	72.85	72.84167	73.450	73.9
Peru	72.1	73.475	73.95	73.89167	74.600	75.3
Philippines	66.8	67.100	67.50	67.51333	68.000	68.4
Poland	73.7	74.800	75.30	75.52667	76.500	77.3
Portugal	76.6	77.500	78.70	79.92000	80.800	89.0
Romania	71.1	72.700	73.40	73.98667	74.700	78.0
Russian Federation	64.6	65.025	67.40	67.44286	69.150	73.0
Rwanda	52.0	57.000	61.00	60.46154	64.600	68.0
Samoa	71.4	72.100	73.00	73.59333	74.400	79.0
Sao Tome and Principe	63.4	64.600	65.60	65.55000	66.675	67.3
Senegal	59.7	62.450	64.30	63.78182	65.300	66.4
Serbia	73.0	73.700	74.10	74.19091	74.750	75.4
Seychelles	71.8	72.100	72.20	72.32000	72.500	73.0
Sierra Leone	45.3	46.875	48.10	48.42500	49.100	54.0
Solomon Islands	65.8	66.950	67.60	67.61333	68.400	68.8
South Africa	53.7	54.250	56.00	57.14000	58.450	69.0
Spain	79.1	80.250	81.60	82.02000	82.250	89.0
Sri Lanka	69.1	73.350	74.05	73.46667	74.500	74.7
Suriname	68.9	69.575	70.60	71.04000	71.375	76.0
Swaziland	45.6	46.750	50.00	50.82000	54.300	58.4
Sweden	81.7	81.700	81.80	81.90000	82.000	82.3
Syrian Arab Republic	72.6	72.775	73.10	73.16250	73.550	73.8
Tajikistan	64.3	65.900	66.40	66.85385	68.100	69.6
Thailand	71.1	71.650	73.30	72.96000	74.000	74.6
Timor-Leste	66.2	66.750	67.20	67.14286	67.550	68.0
Togo	56.2	57.050	58.30	58.08571	59.150	59.7
Tonga	71.6	72.100	72.50	72.46667	72.850	73.3
Trinidad and Tobago	69.2	69.500	69.90	71.36154	71.100	78.0
Tunisia	72.9	73.850	74.60	74.29333	74.800	75.1
Turkey	71.2	72.600	73.90	73.78667	74.650	78.0
Turkmenistan	63.3	63.500	64.50	64.60000	65.600	66.0
Uganda	48.8	53.200	56.30	56.51538	59.300	67.0
Ukraine	67.0	67.550	67.70	69.84667	70.400	78.0
Uruguay	75.1	75.400	76.20	76.01333	76.550	77.0
Uzbekistan	67.1	67.600	67.90	68.04615	68.500	69.2
Vanuatu	69.0	69.500	71.00	71.34667	71.650	78.0
Zambia	49.3	56.125	58.00	57.25000	58.950	63.0
Zimbabwe	44.3	45.050	46.60	49.38667	53.650	59.2

```
# Graphical
data%>%
  ggplot(aes(x=Country, y=Life.expectancy)) +
  geom_boxplot() +
  scale_x_discrete(guide = guide_axis(angle = 70)) +
  theme(axis.text.x=element_text(size=7))
```



### Find Life Expectancy Outliers

```

out.dat <- data[FALSE,]

for (c in unique(data$Country)){
  country.dat <- data$Life.expectancy [data$Country==c]
  lowerq = quantile(data$Life.expectancy [data$Country==c]) [2]
  upperq = quantile(data$Life.expectancy [data$Country==c]) [4]
  iqr <- IQR(data$Life.expectancy [data$Country==c])
  upper.out = (iqr * 1.5) + upperq
  lower.out = lowerq - (iqr * 1.5)
  for (v in country.dat){
    if (v <= lower.out | v >= upper.out){
      out.dat <- rbind(out.dat, filter(data, Country==c, Life.expectancy==v))
    }
  }
}
kable(head(out.dat, 10))

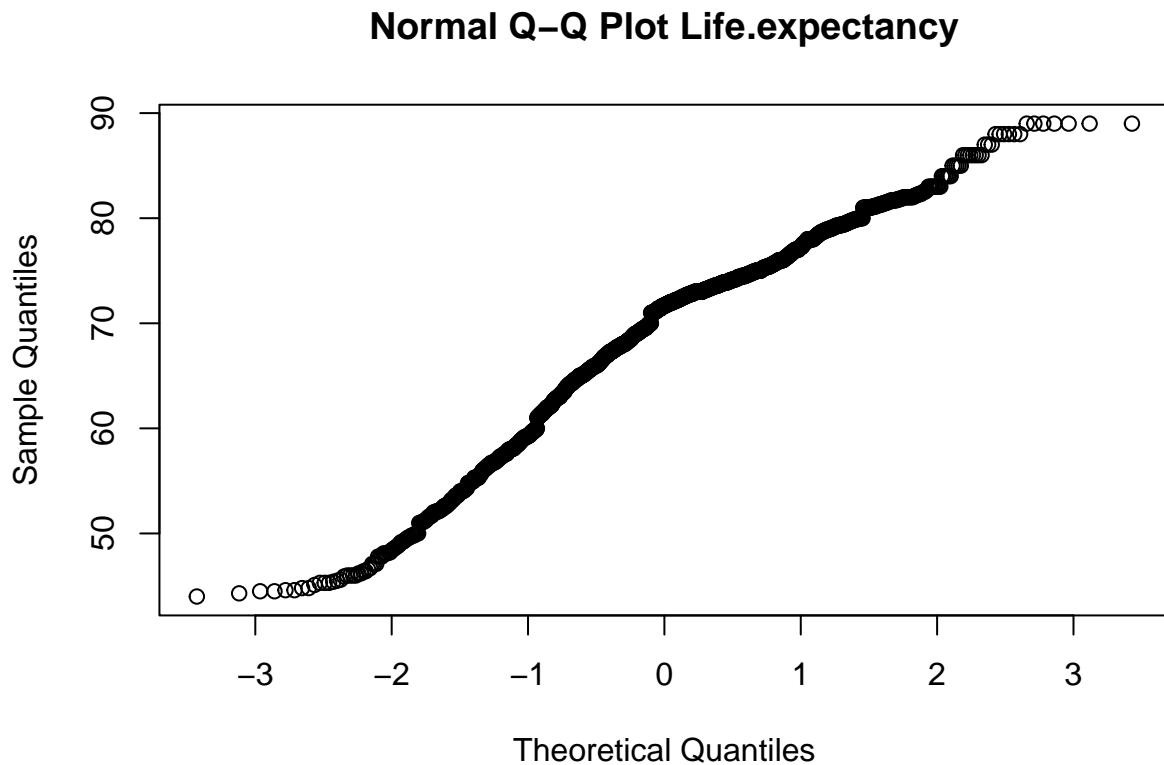
```

Country	Year	Status	Life.expectancy	Aptitude	Alcohol	HDI	GDP	Min.Poverty	MDG	MDG1	MDG2	MDG3	MDG4	MDG5	MDG6	MDG7	MDG8	Optimism	Spirits	Relationships	Innovation	Memory	Scholarship	Composition.of.res.
Afghanistan	2011	Developed	65.0	9263	62	0.0171.27962	115419.183	6	8.16	65	0.1	584.23972367424	17.3	0.479	10.1									
Angola	2011	Developed	56.0	9358	72	8.24256.12252	445821.5110	75	3.30	75	2.6	459825900158	8.6	0.508	10.3									
Australia	2010	Developed	86.0	9669	1	9.84588.56937	70	6.9	1	92	8.57	92	0.1	3472212780.7	0.6	0.908	20.7							
Bangladesh	2011	Developed	70.0	9137	111	0.0159.25893	198616.4139	94	3.80	94	0.1	856.342972853	19.0	0.557	9.9									
Belgium	2011	Developed	80.0	9676	0	12.67163.3489270	63.4	1	99	1.59	99	0.1	474391296570	1.0	0.890	16.3								

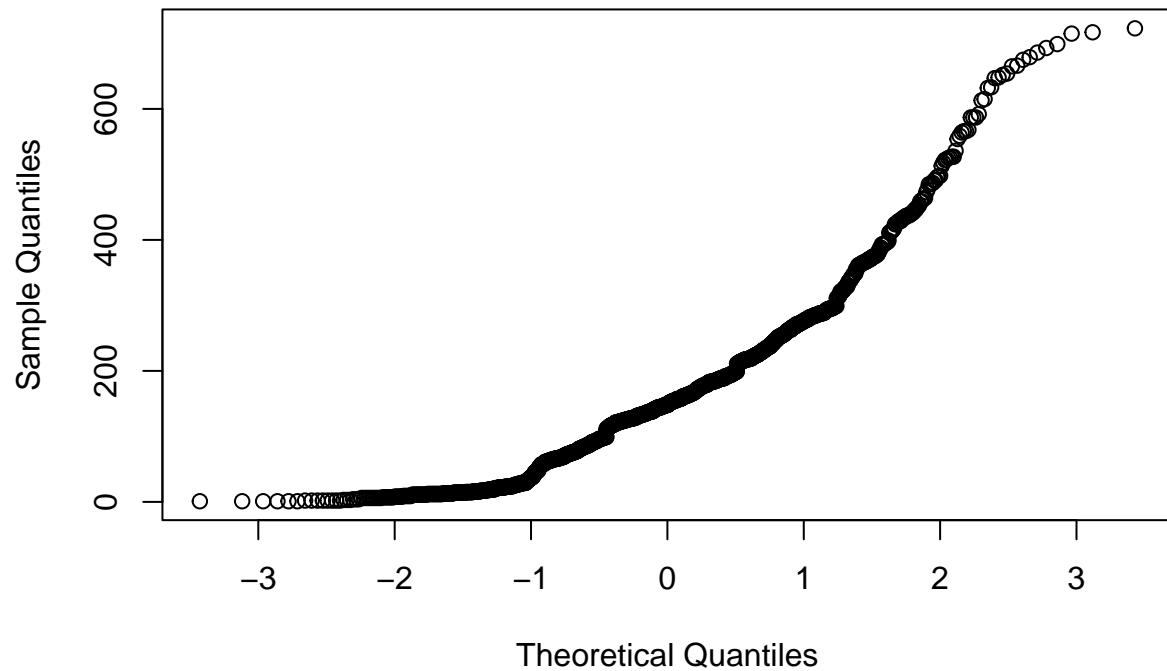
Country	Year	Stat	Life.expectancy	Agricultural land (% of land area)	Mining, quarrying and construction (% of land area)	Arable land (% of land area)	Net primary productivity (GDP\$)	Population density (people per square kilometer)	Urban population (% of total population)	GDP (\$Bn)	HDI	GDP per capita (\$)	Optimal life expectancy (years)	Optimal life expectancy (years) in 1950	Optimal life expectancy (years) in 1950/90s	Composition of resources			
Belgium	2013	Develped	77	0	11.77	0.35986	39	63.0	1	99	1.57	99	0.1	46511	13862	1807	0.9	0.889	16.3
Cabo Verde	2000	Develping	44	0	4.16	209.08650	0	23.2	0	93	5.00	92	0.9	17684	49213	38.9	8.9	0.572	11.3
Cabo Verde	2007	Develping	48	0	3.82	155.20727	0	22.6	0	92	5.17	91	0.8	13734	45216	6.2	9.1	0.569	11.3
Canada	2000	Develping	74	2	8.30	8433.9268162	63.0	2	89	1.30	92	0.1	46598	33240	73	0.4	0.897	15.8	
Central African Republic	2018	Develping	37	15	0.0153	43947	210	22.122	47	4.20	47	4.5	377.14523	35392	8.3	8.3	0.345	7.1	

## QQ plots

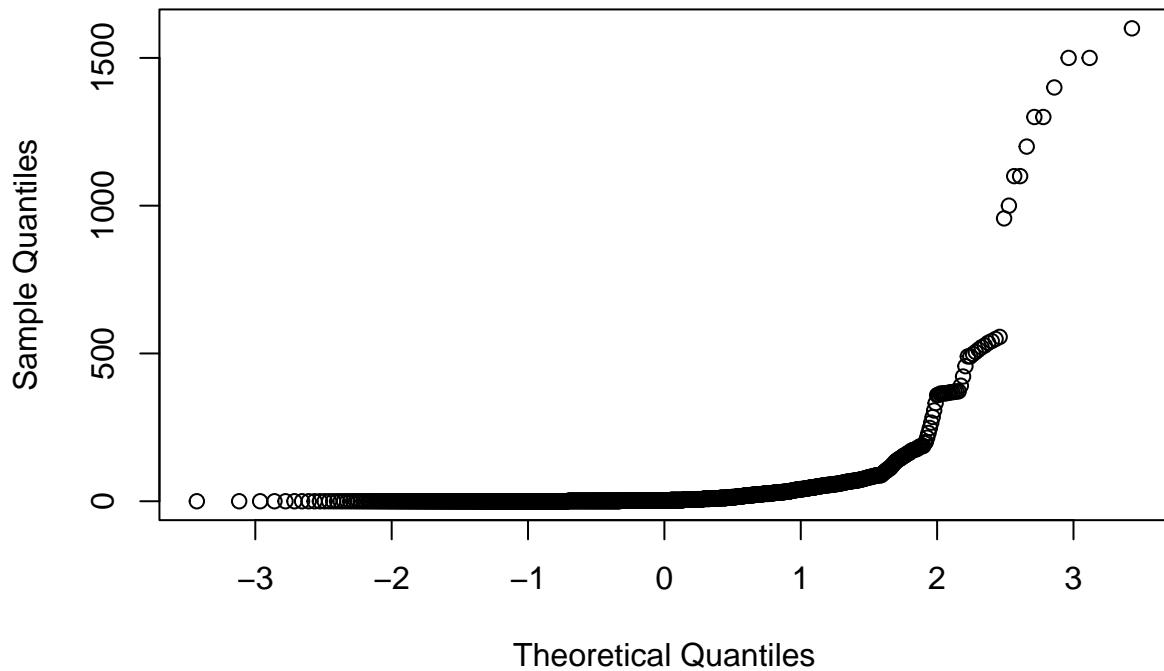
```
for (i in 1:length(colnames(dataNumeric))){
  qqnorm(dataNumeric[,i], main = paste0("Normal Q-Q Plot ", colnames(dataNumeric)[i]))
}
```



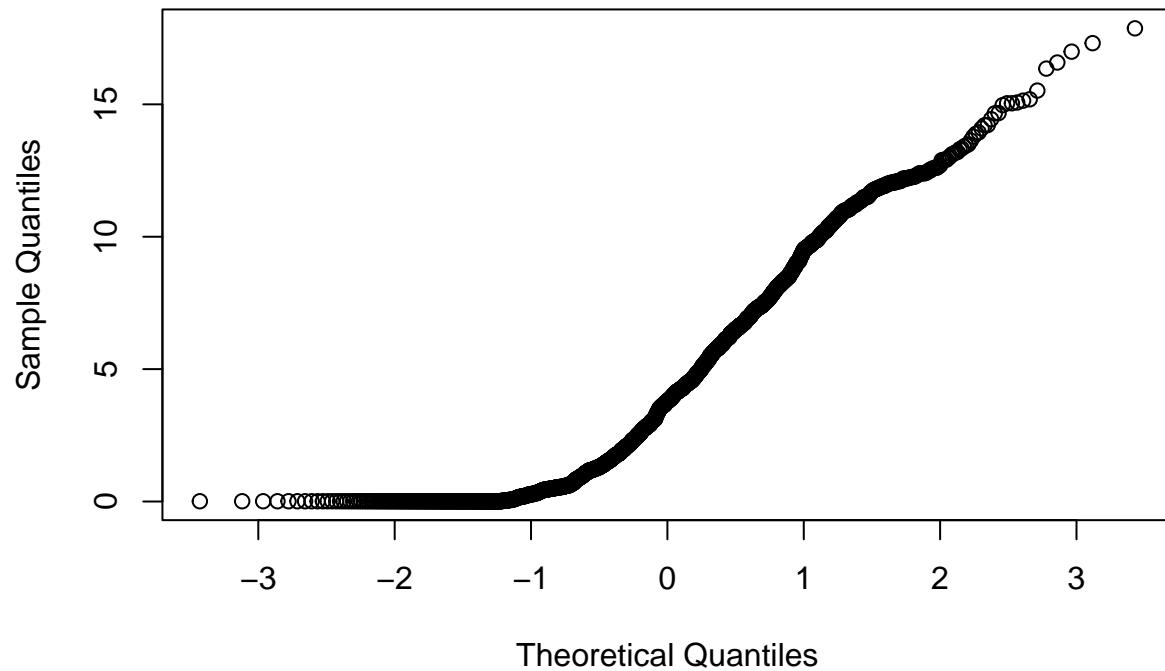
**Normal Q–Q Plot Adult.Mortality**



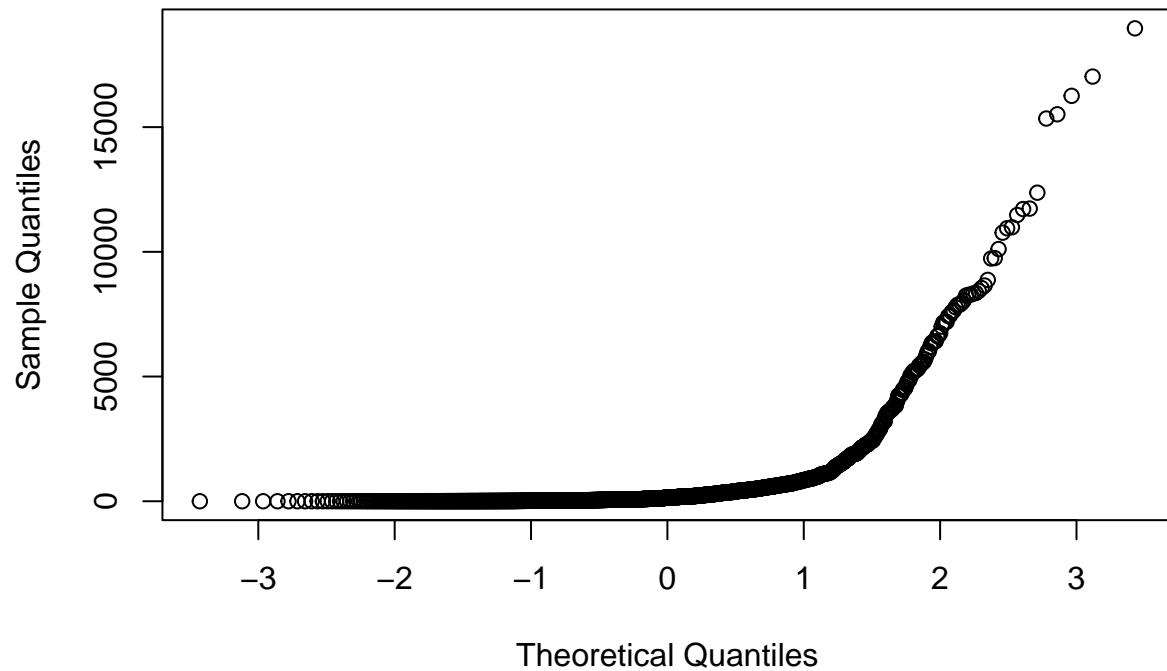
### Normal Q-Q Plot infant.deaths



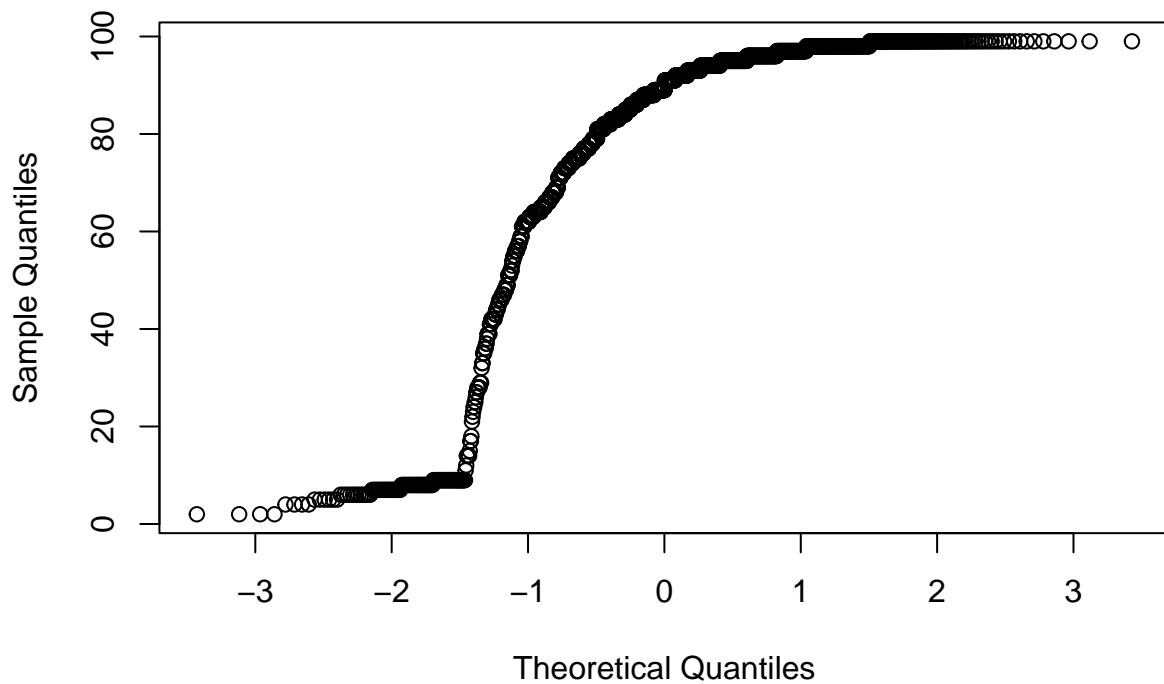
### Normal Q-Q Plot Alcohol



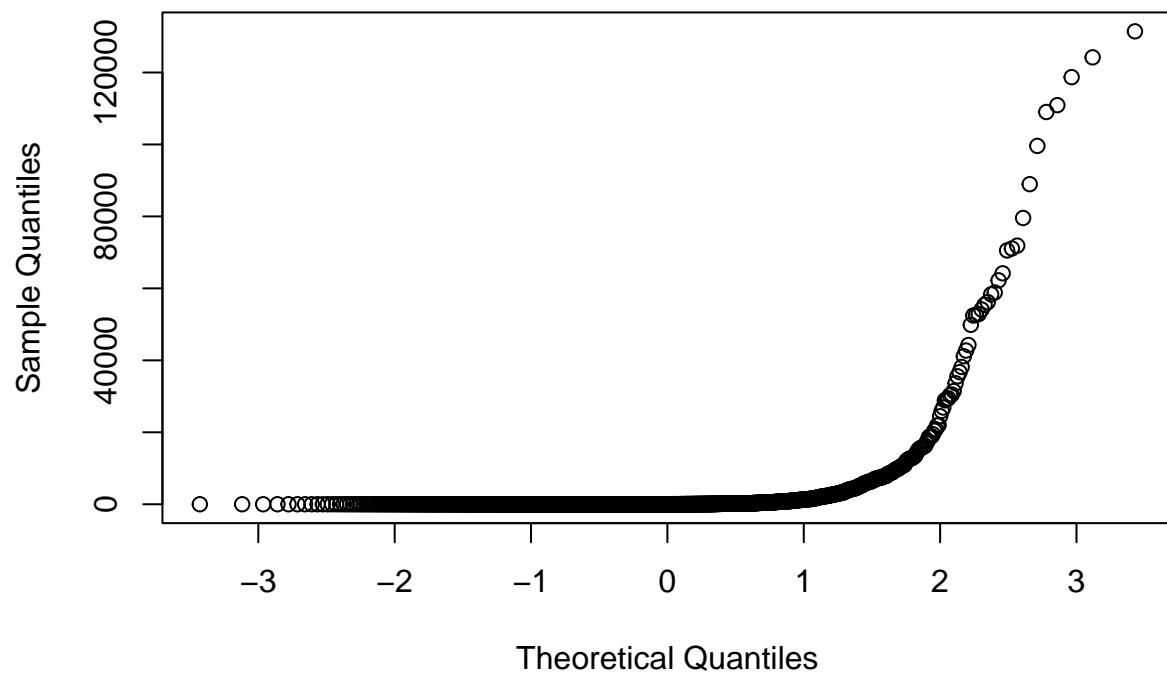
**Normal Q–Q Plot percentage.expenditure**



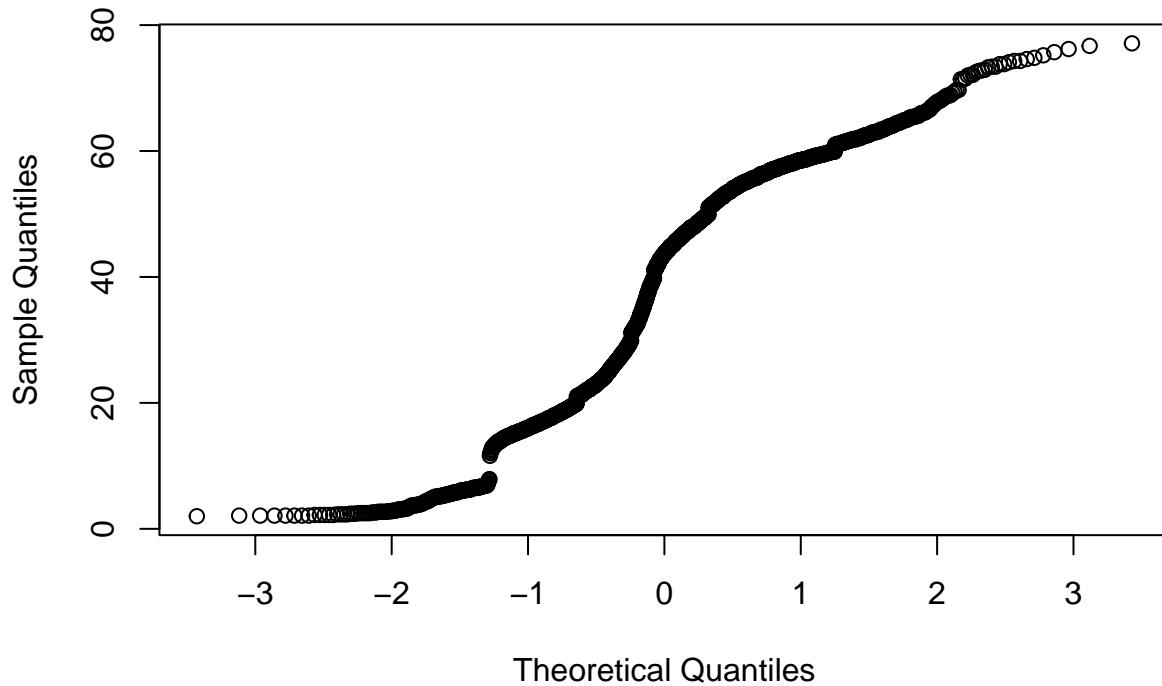
### Normal Q–Q Plot Hepatitis.B



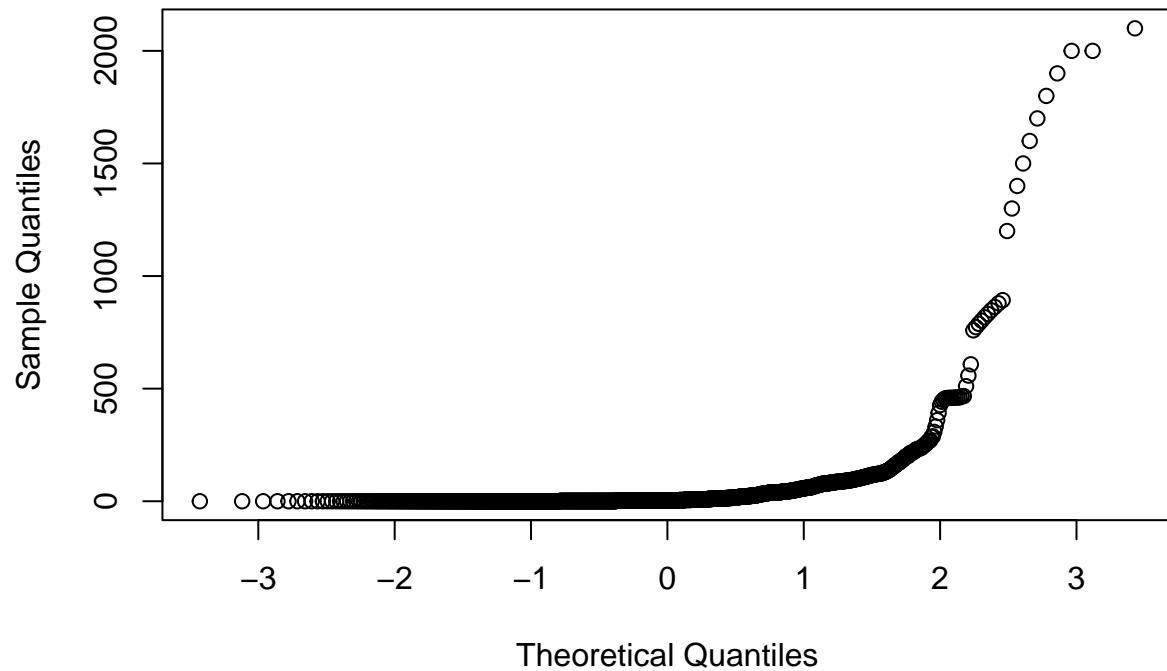
### Normal Q–Q Plot Measles



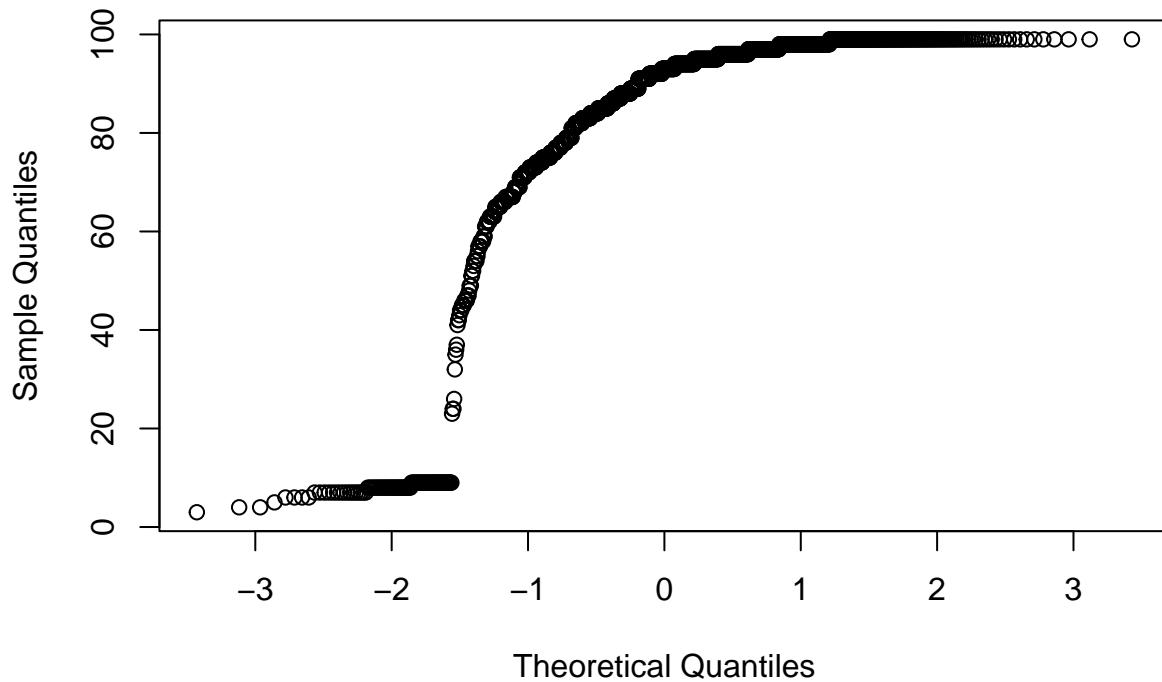
### Normal Q-Q Plot BMI



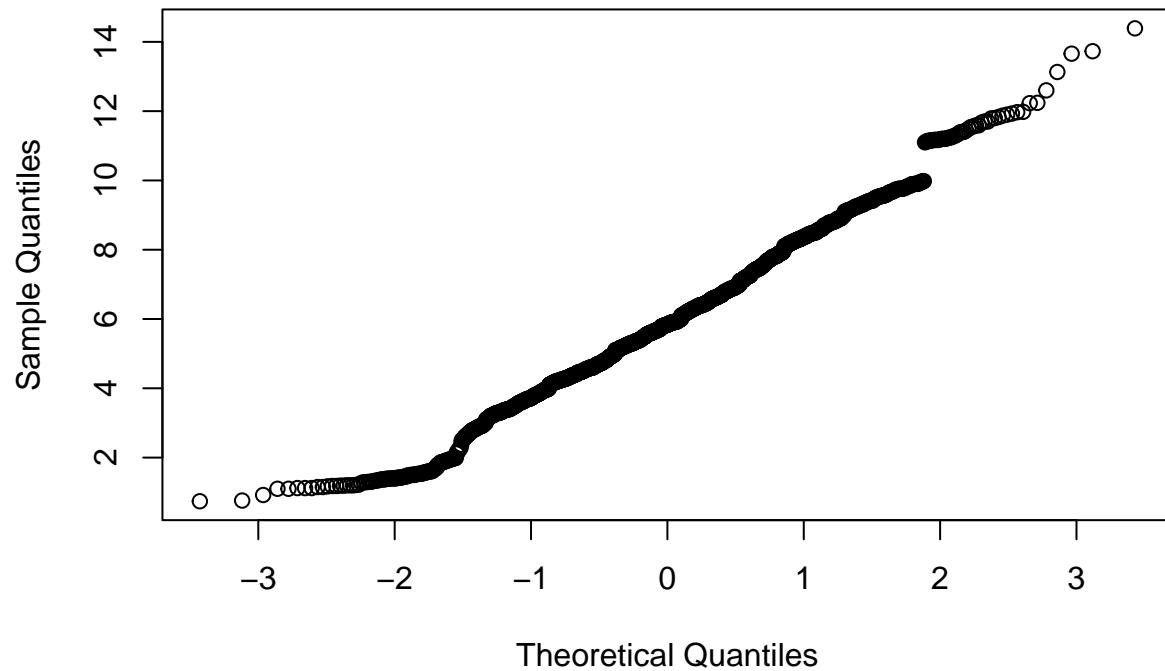
**Normal Q-Q Plot under.five.deaths**



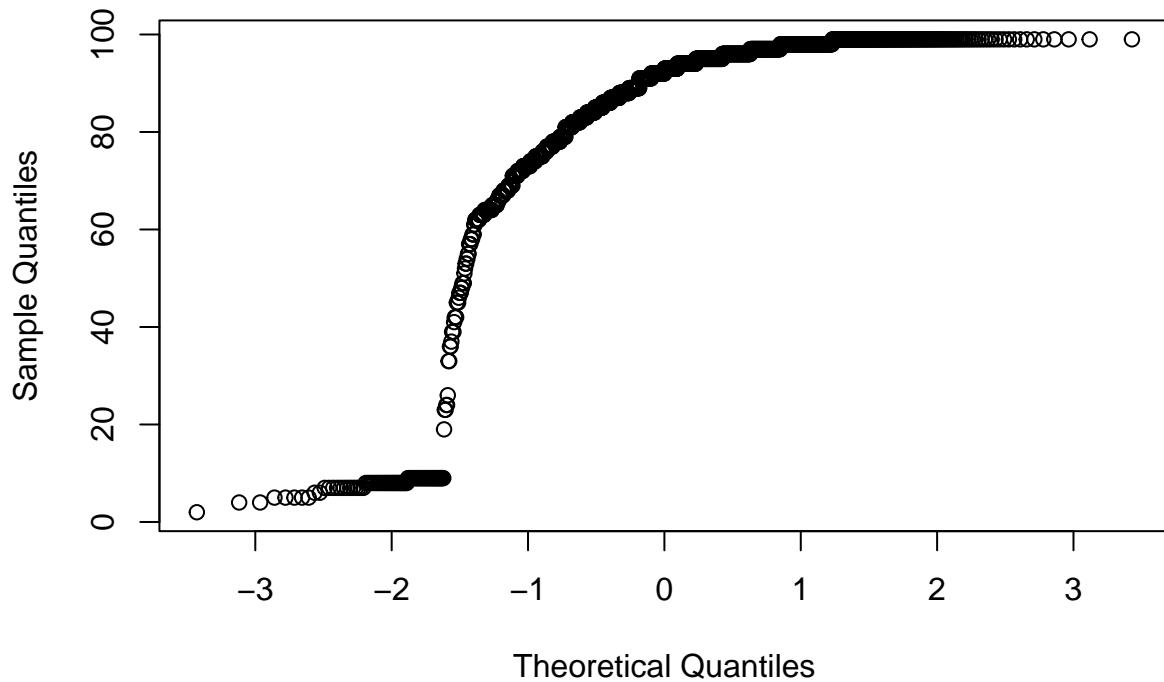
### Normal Q–Q Plot Polio



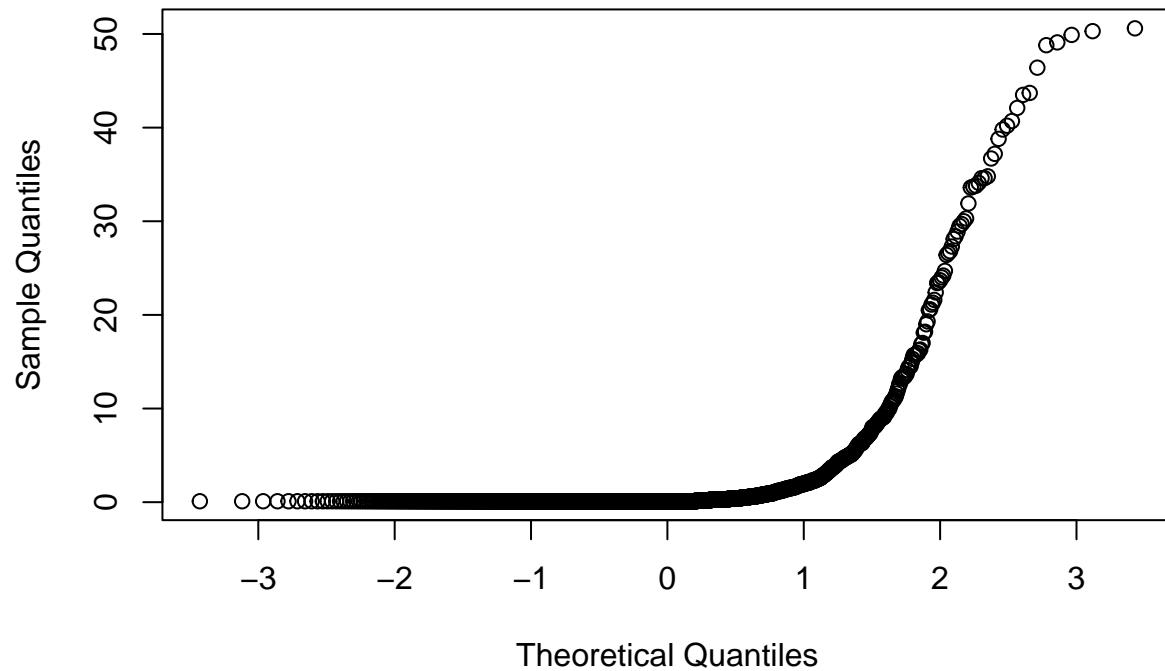
**Normal Q-Q Plot Total.expenditure**



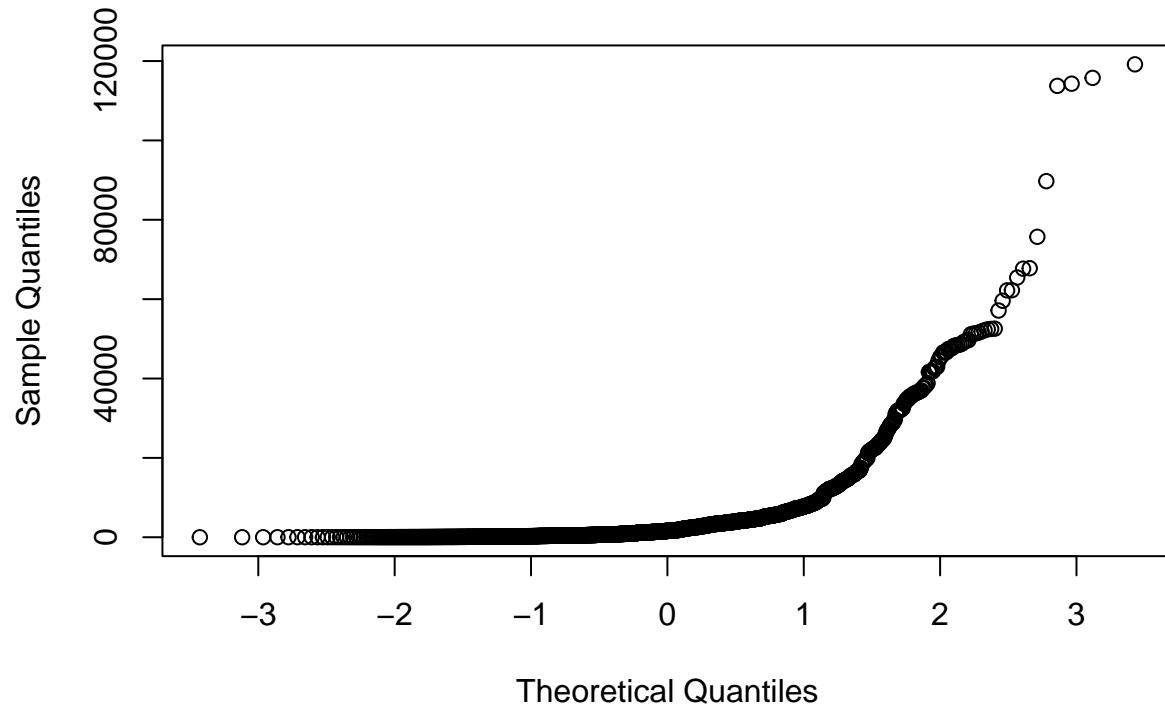
### Normal Q-Q Plot Diphtheria



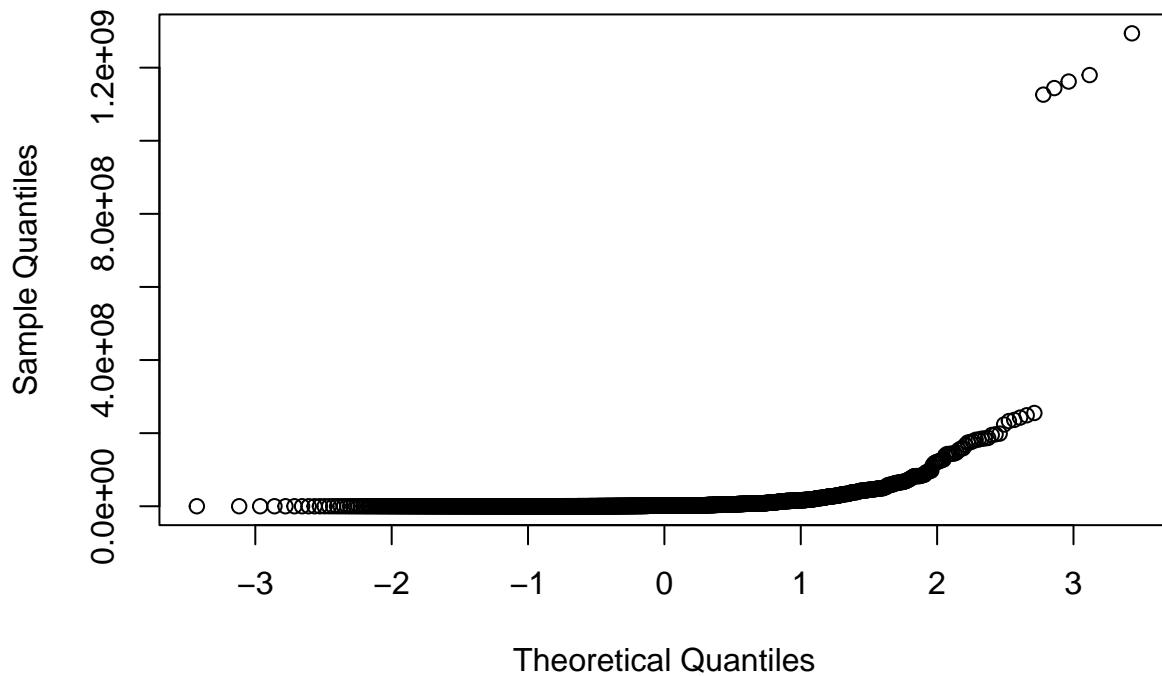
### Normal Q-Q Plot HIV.AIDS



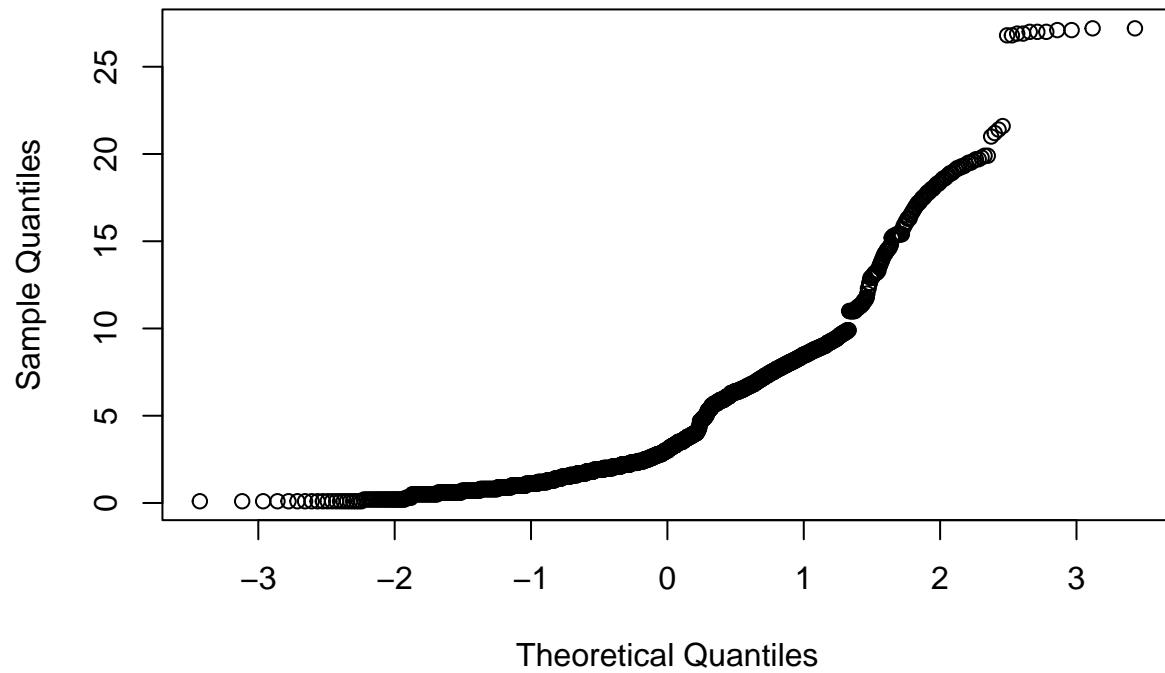
**Normal Q-Q Plot GDP**



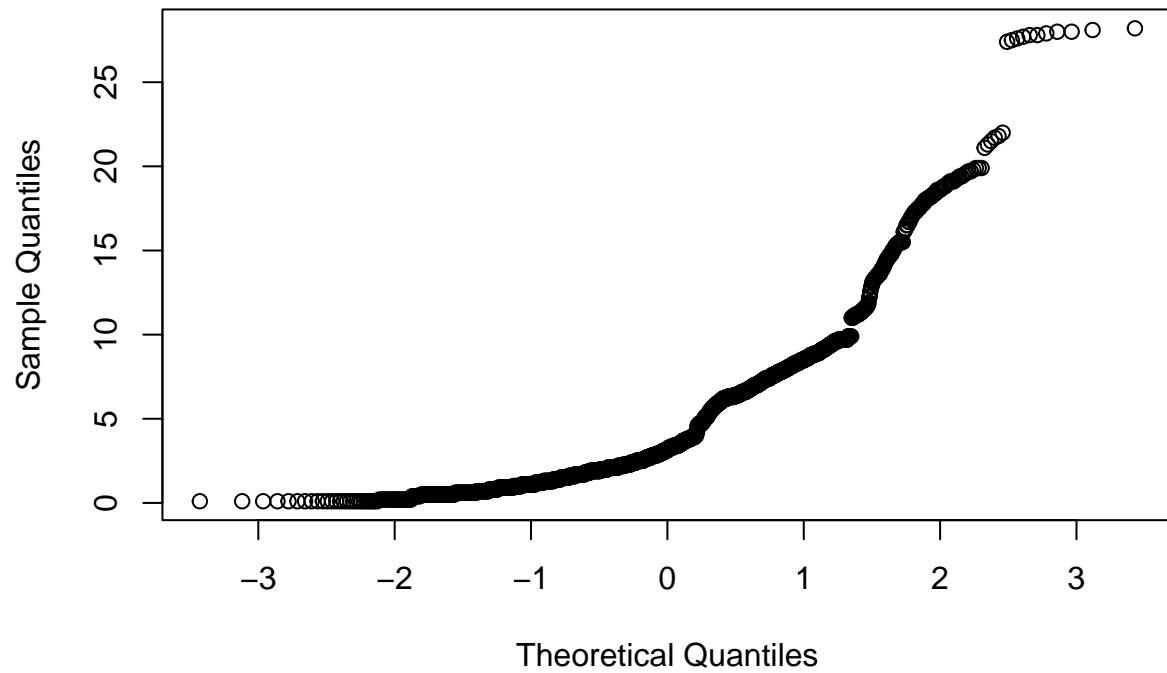
### Normal Q–Q Plot Population



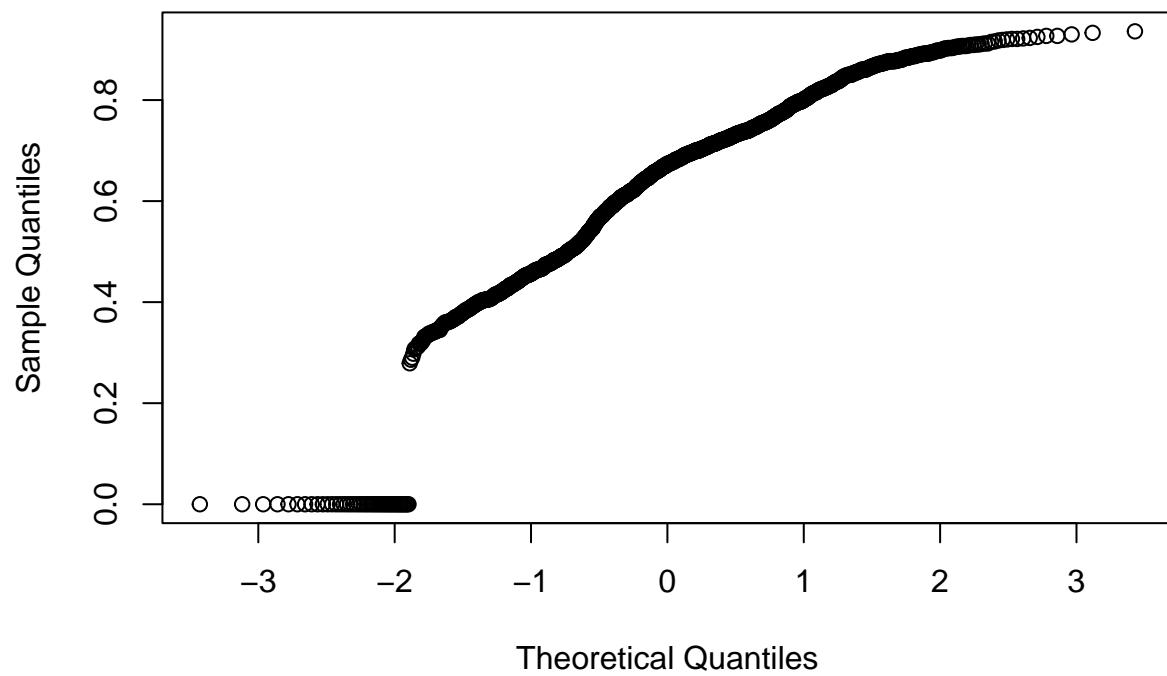
**Normal Q–Q Plot thinness..1.19.years**



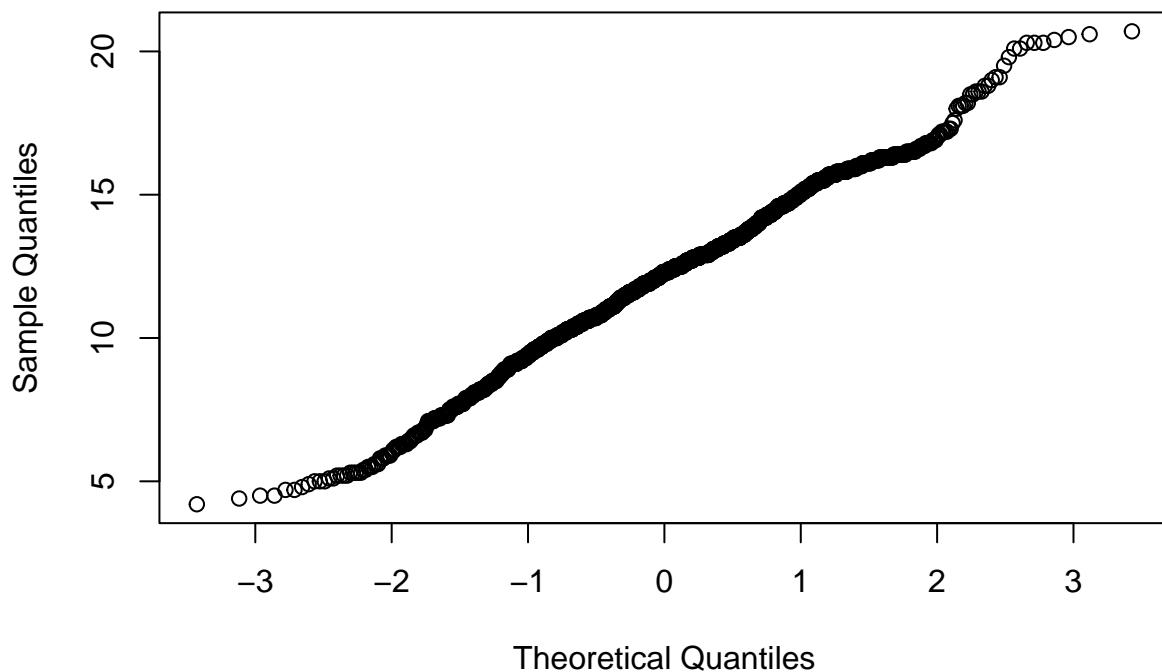
**Normal Q-Q Plot thinness.5.9.years**



**Normal Q–Q Plot Income.composition.of.resources**



## Normal Q-Q Plot Schooling

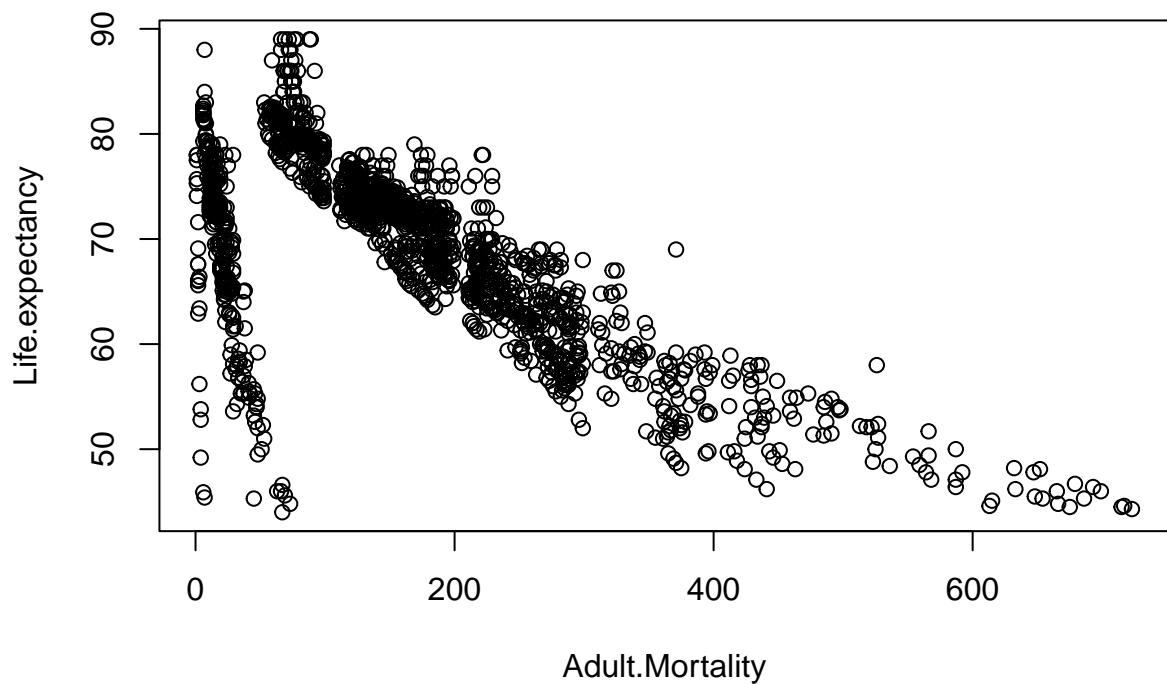


- *Adult Mortality* - Slightly Right Skewed
- *HIV Aids* - Very Right Skewed, many zeros
- *Income* - Slightly Left Skewed, many Zeros
- *Schooling* - Normal

### Scatterplots

```
for (i in 2:length(colnames(dataNumeric))){  
  plot(dataNumeric[,i], dataNumeric[,1],  
    main = paste0(colnames(dataNumeric)[i], "Scatterplot"),  
    xlab = colnames(dataNumeric)[i],  
    ylab = colnames(dataNumeric)[1])  
}
```

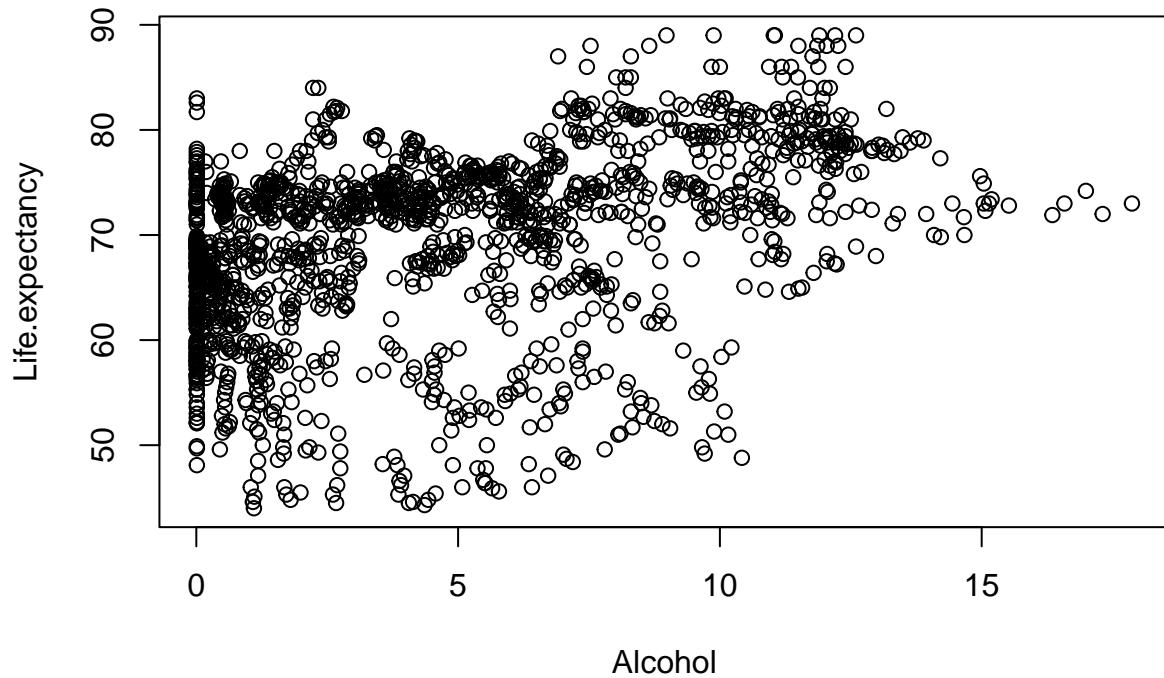
**Adult.MortalityScatterplot**



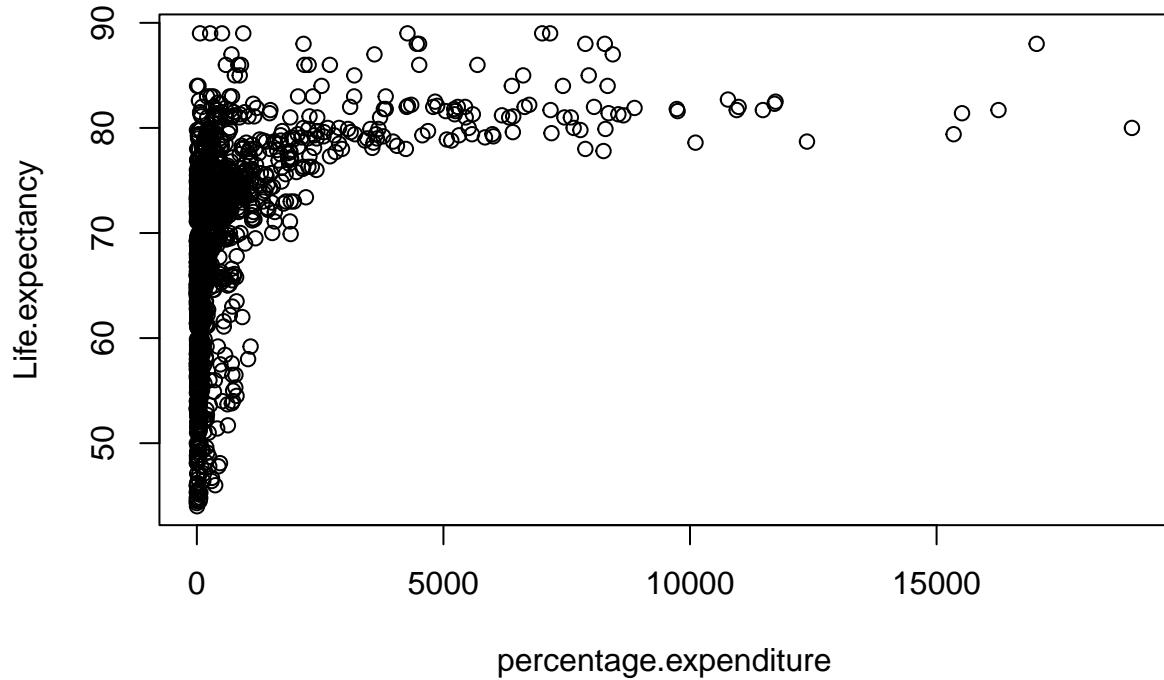
**infant.deathsScatterplot**



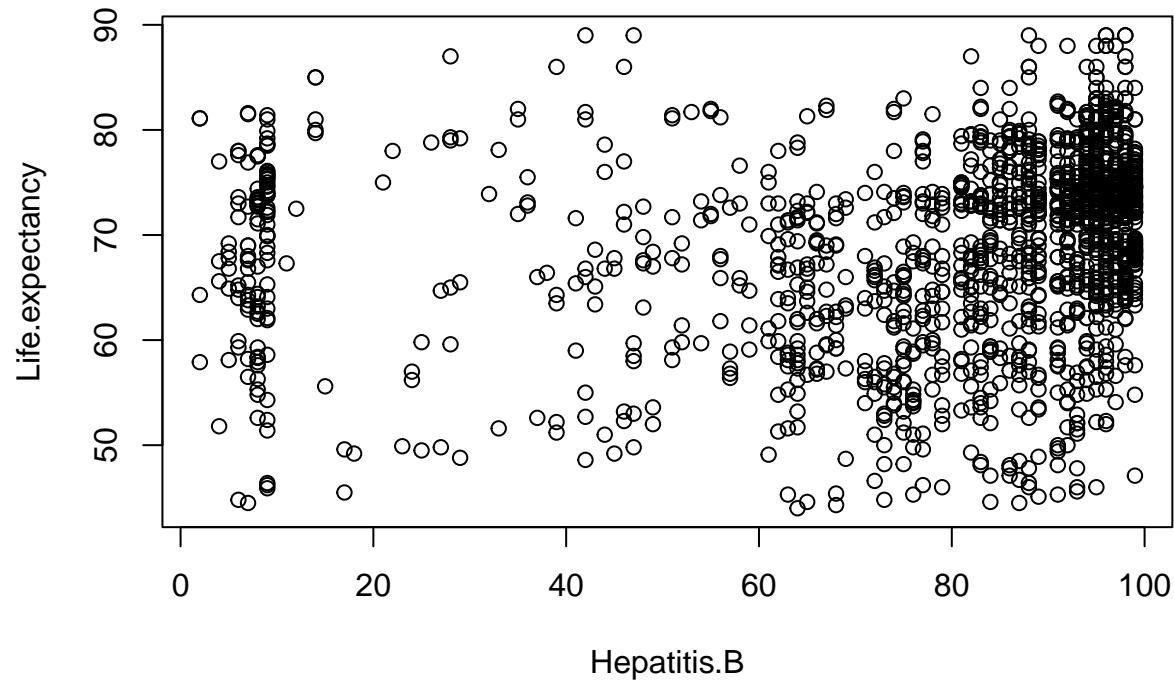
### AlcoholScatterplot



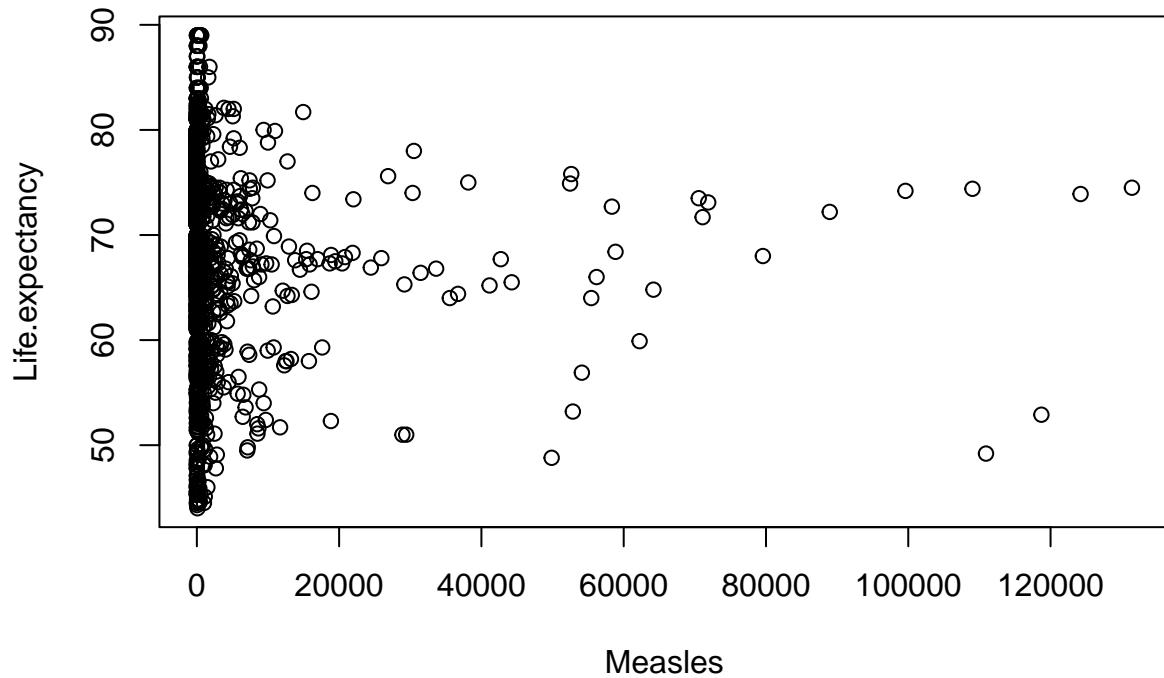
**percentage.expenditureScatterplot**



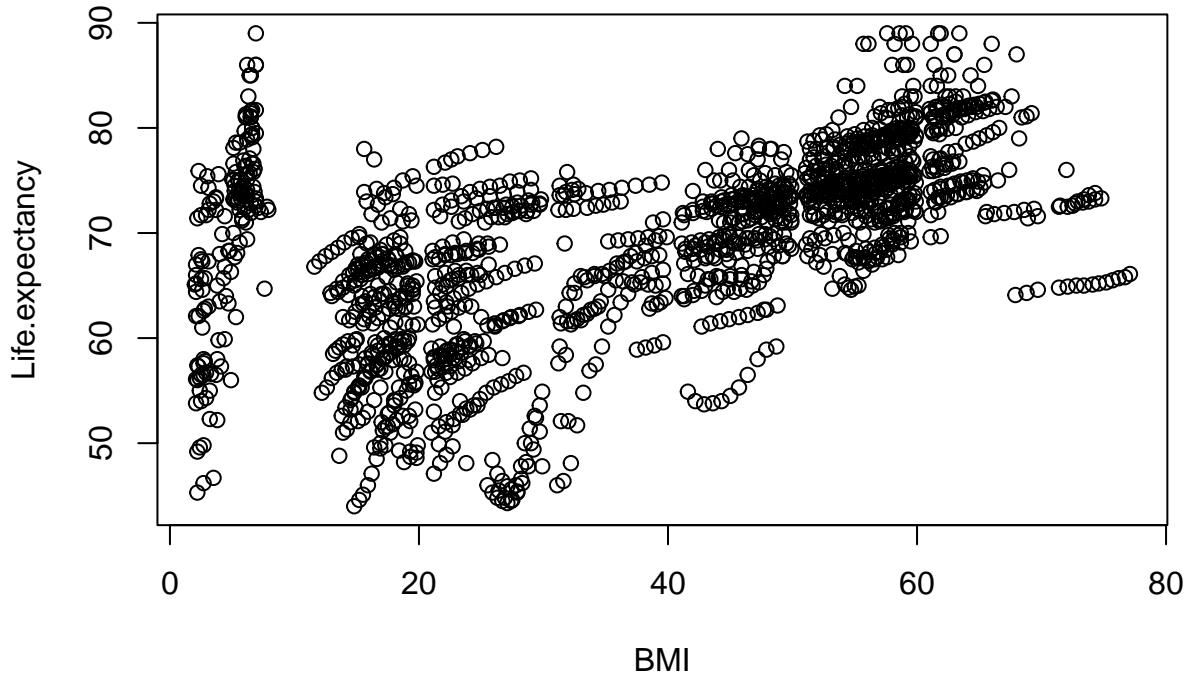
## Hepatitis.BScatterplot



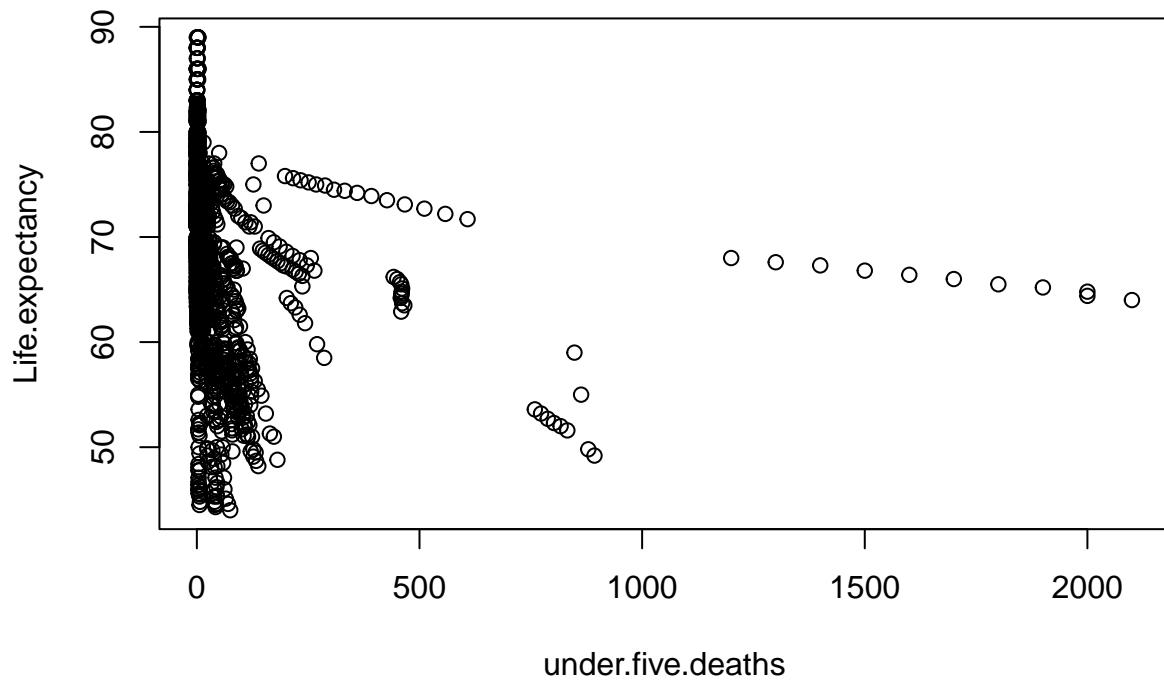
### MeaslesScatterplot



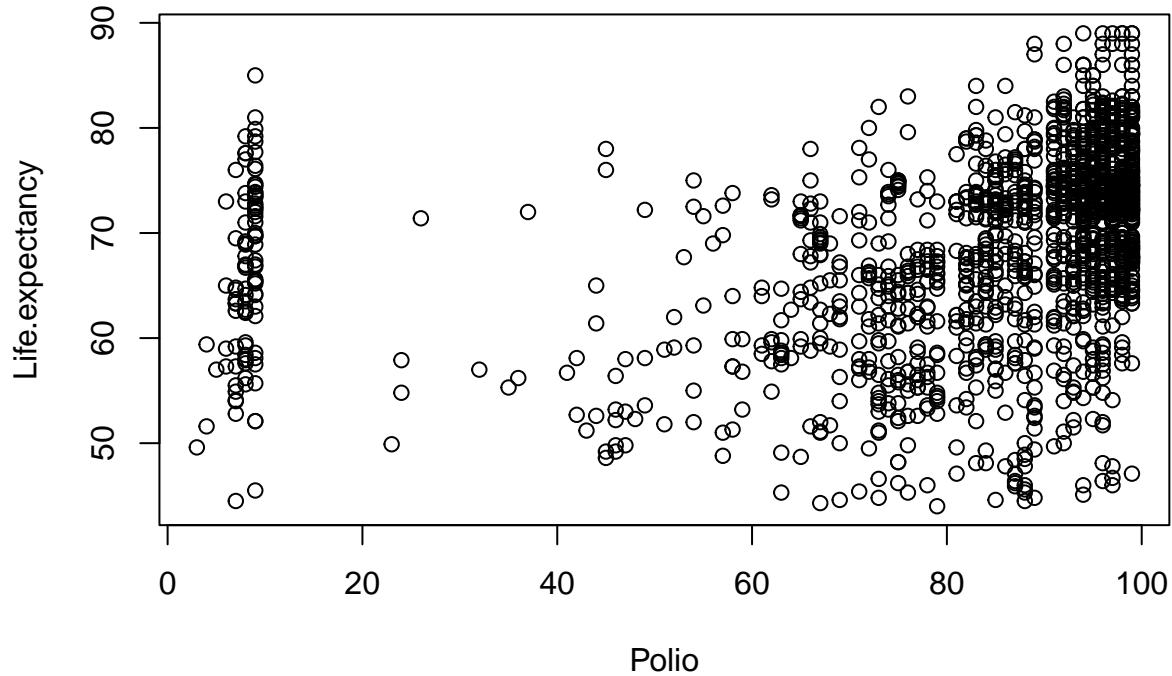
### BMI Scatterplot



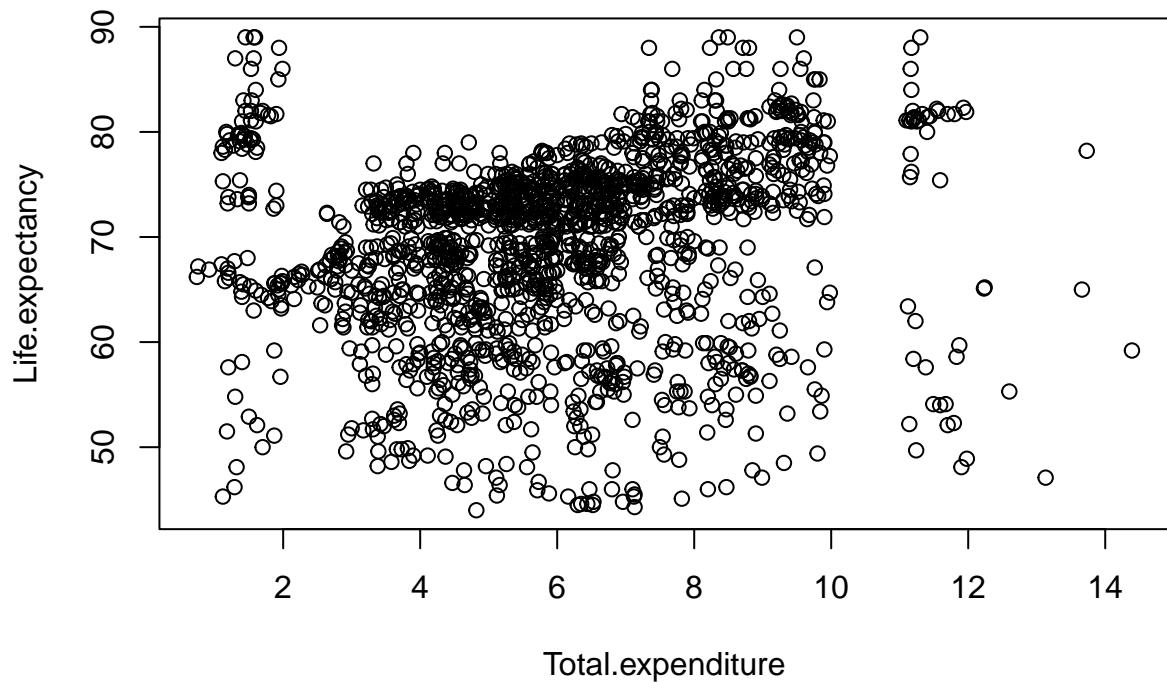
**under.five.deathsScatterplot**



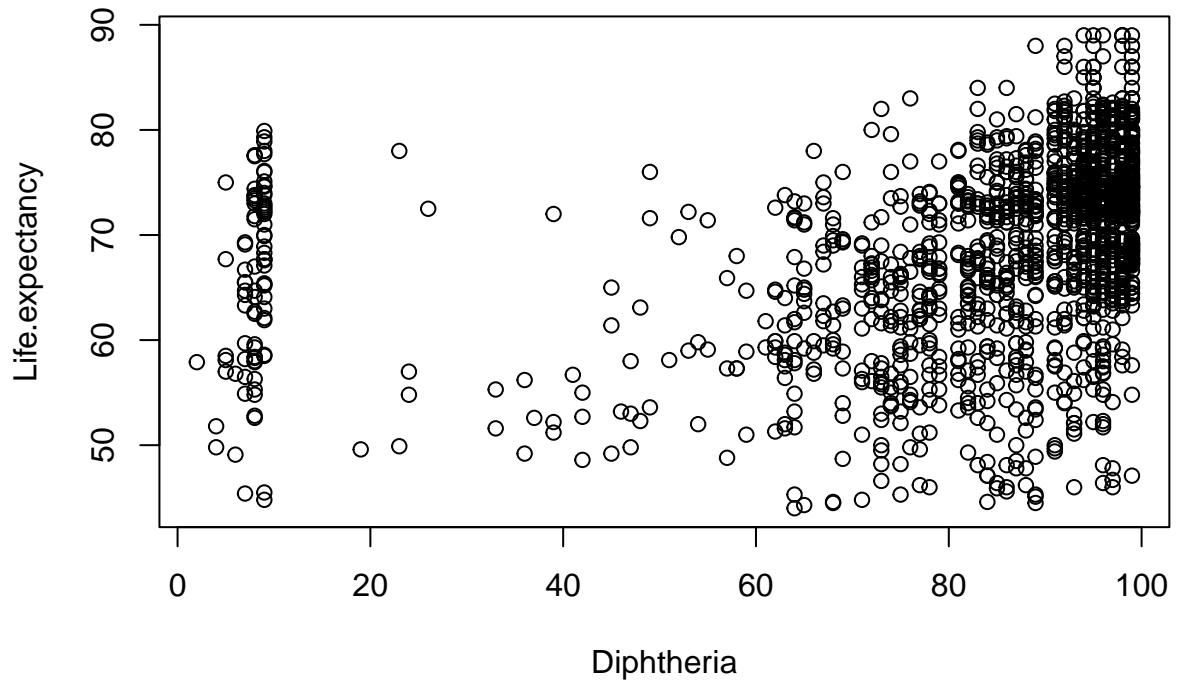
### PolioScatterplot



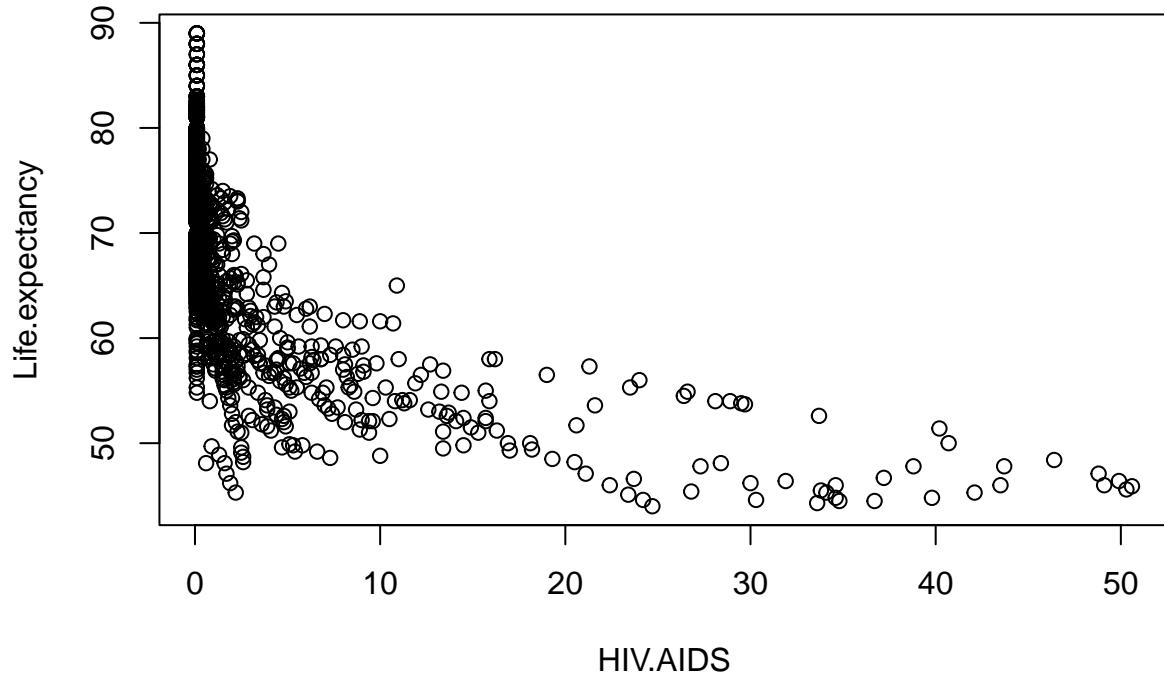
**Total.expenditureScatterplot**



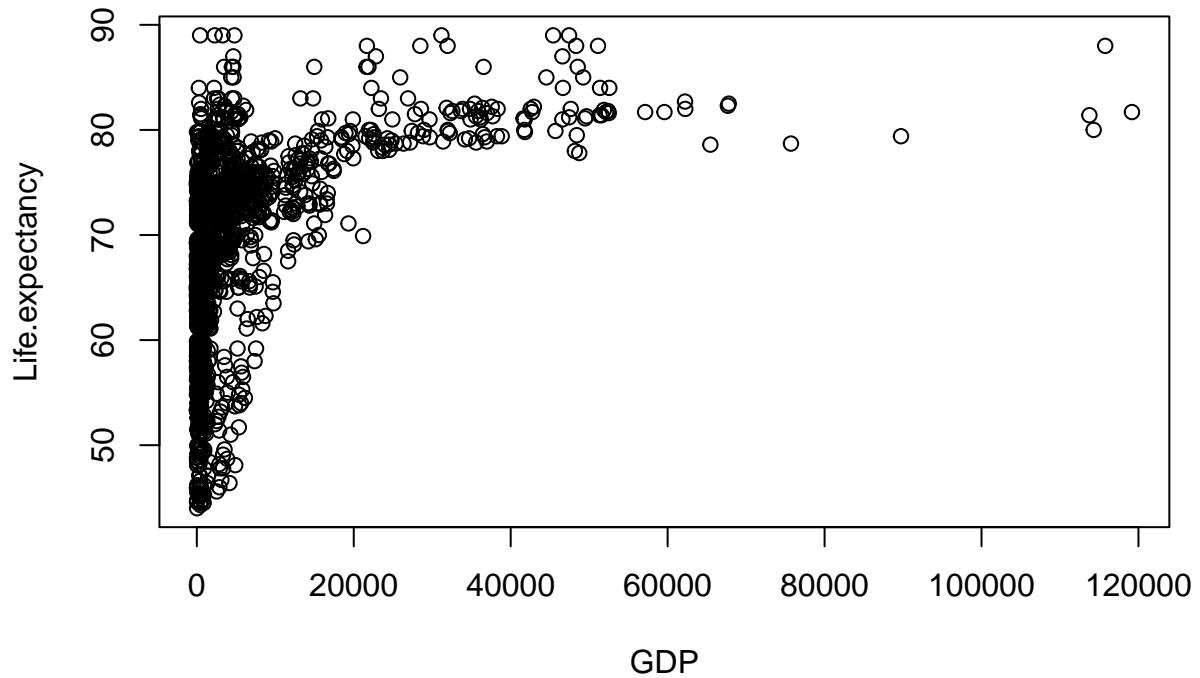
### DiphtheriaScatterplot



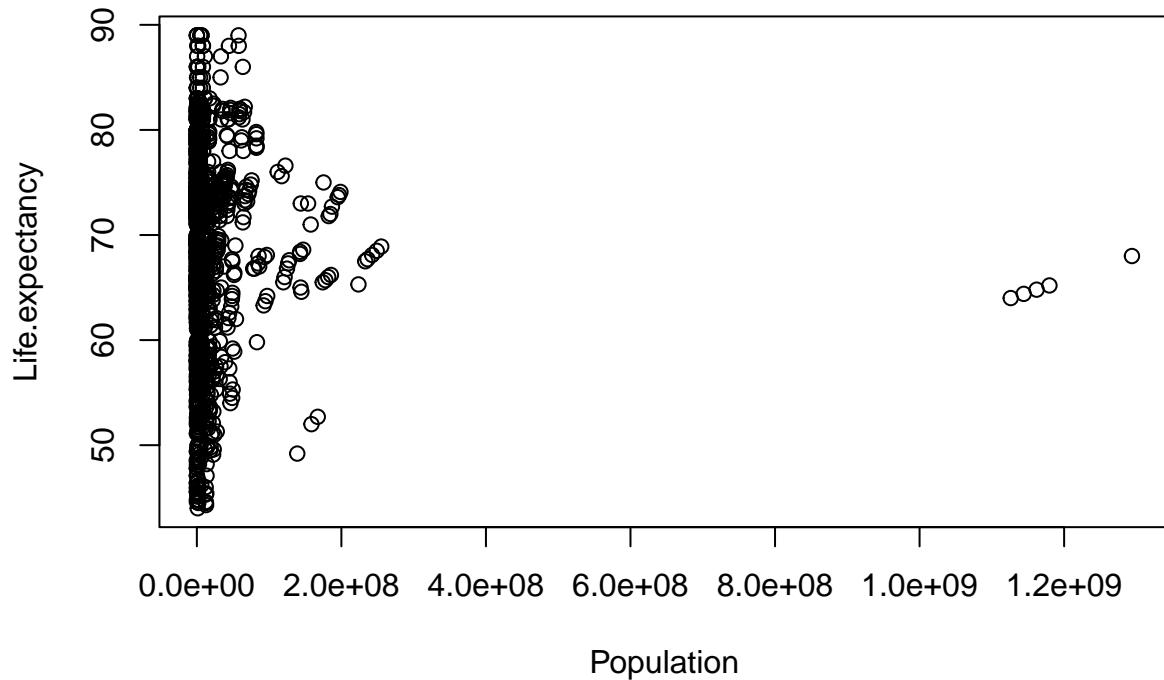
**HIV.AIDSScatterplot**



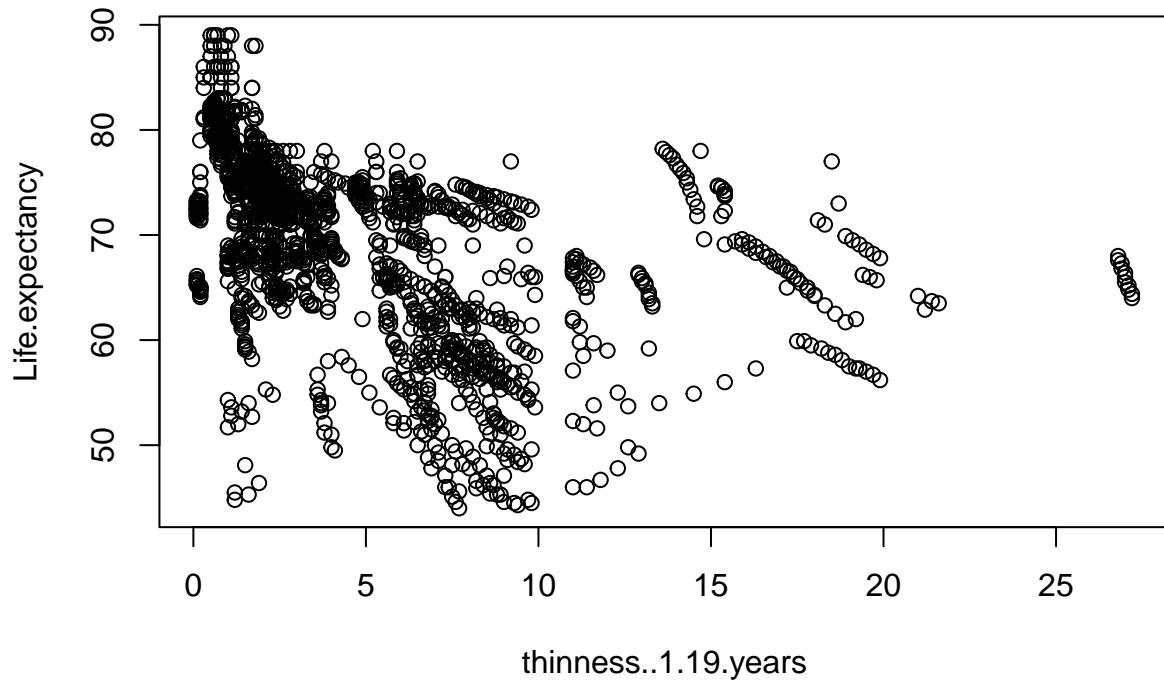
### GDPSscatterplot



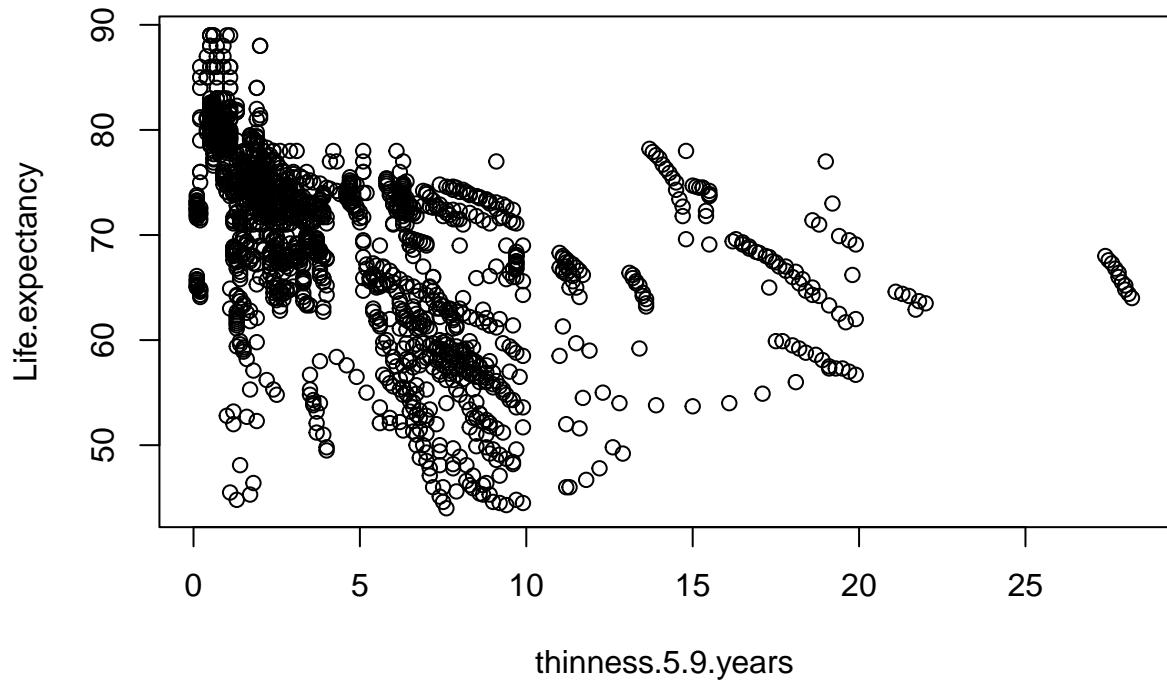
### PopulationScatterplot



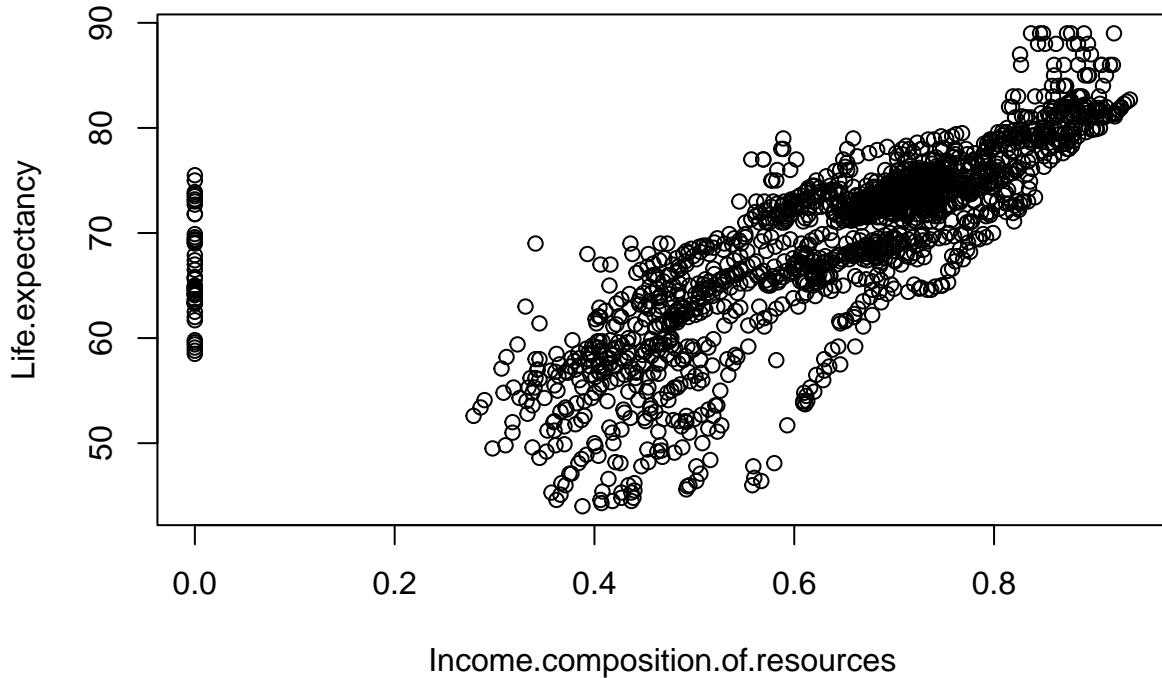
**thinness..1.19.yearsScatterplot**



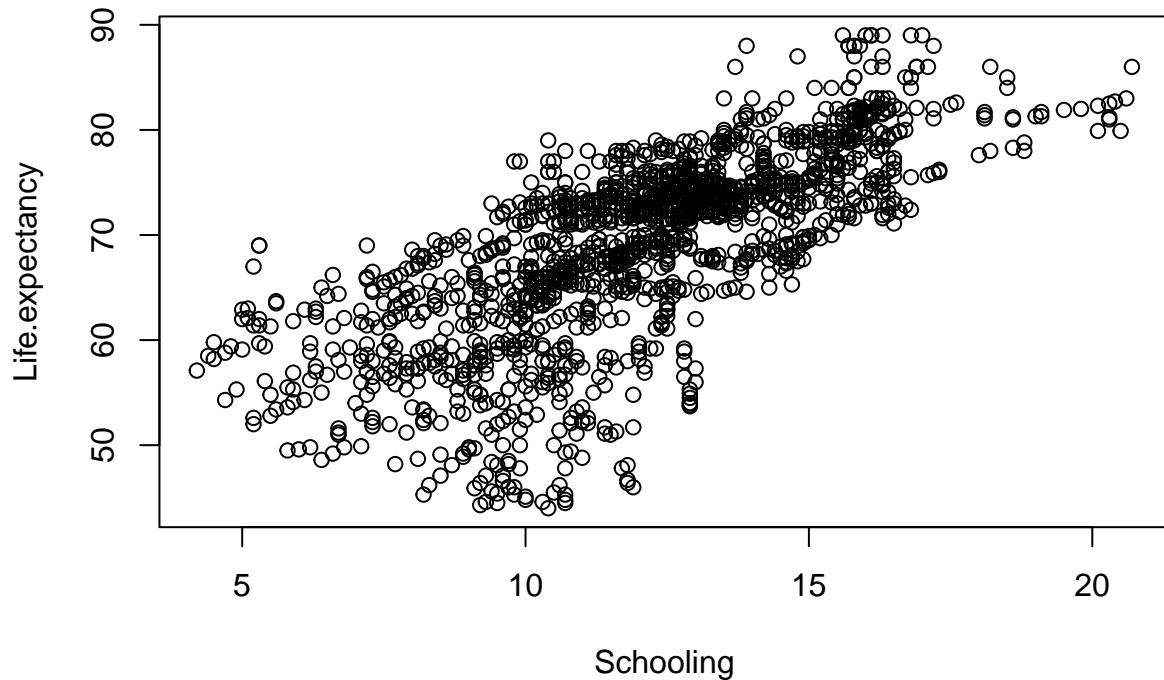
**thinness.5.9.yearsScatterplot**



**Income.composition.of.resourcesScatterplot**

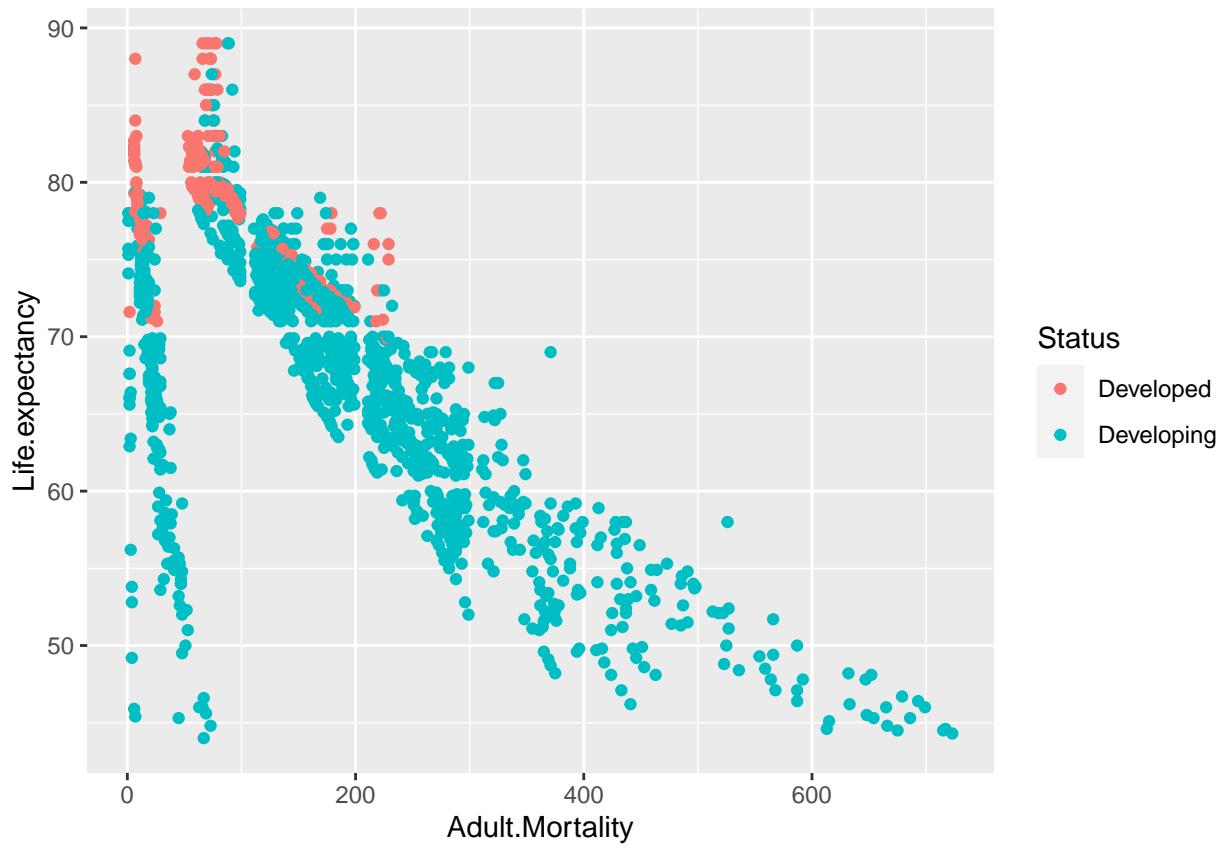


## SchoolingScatterplot



Trying to figure out cause of Adult Mortality Grouping

```
data%>%
  ggplot(aes(Adult.Mortality, Life.expectancy, color=Status))+
  geom_point()
```

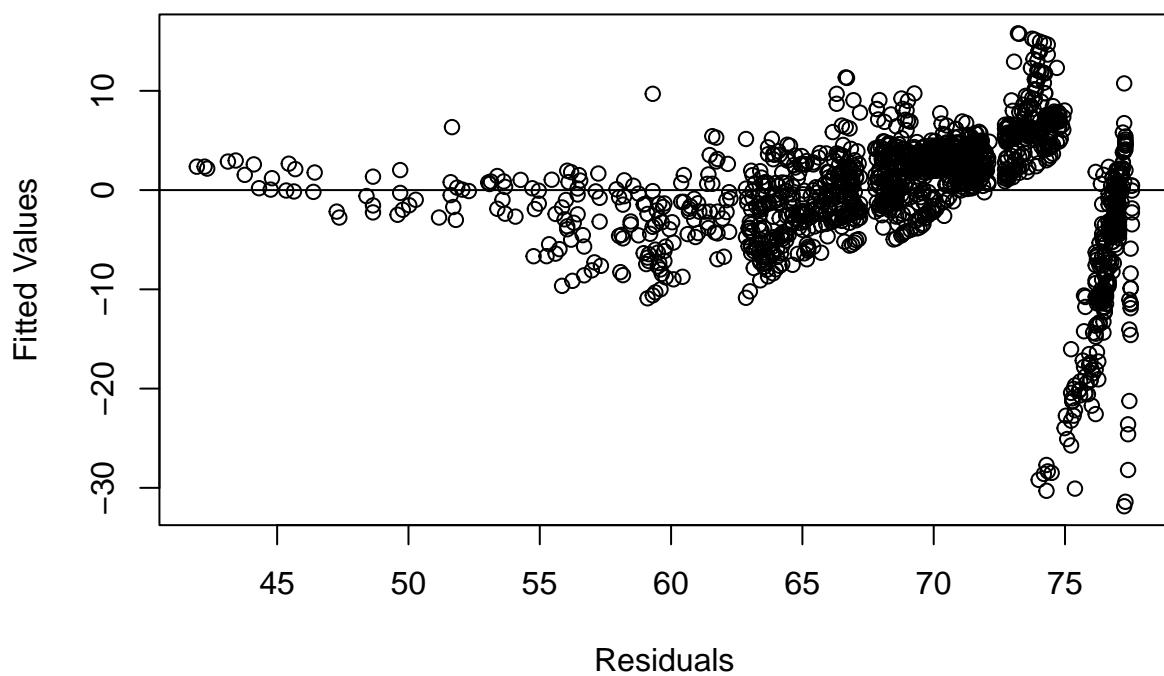


### Residual Plots

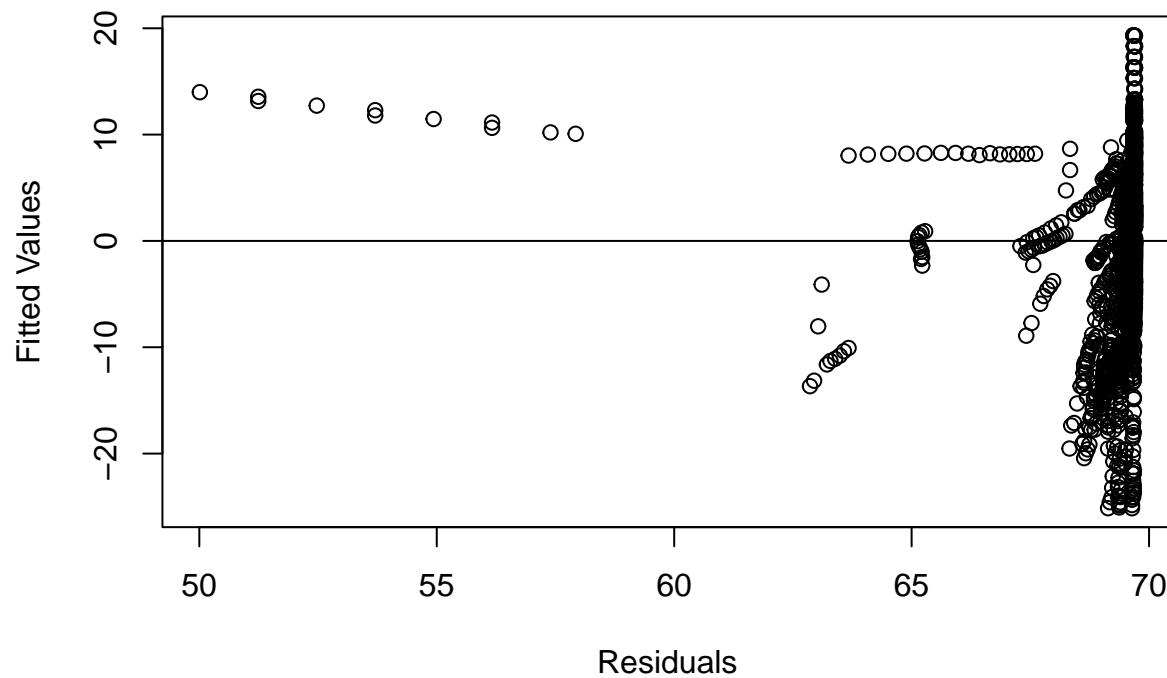
```

for (i in 2:length(colnames(dataNumeric))){
  model <- lm(dataNumeric[,1] ~ dataNumeric[,i])
  res <- resid(model)
  plot(fitted(model), res,
    main = paste0(colnames(dataNumeric)[i], " Residual Plot"),
    xlab = "Residuals",
    ylab = "Fitted Values")
  abline(0,0)
}
  
```

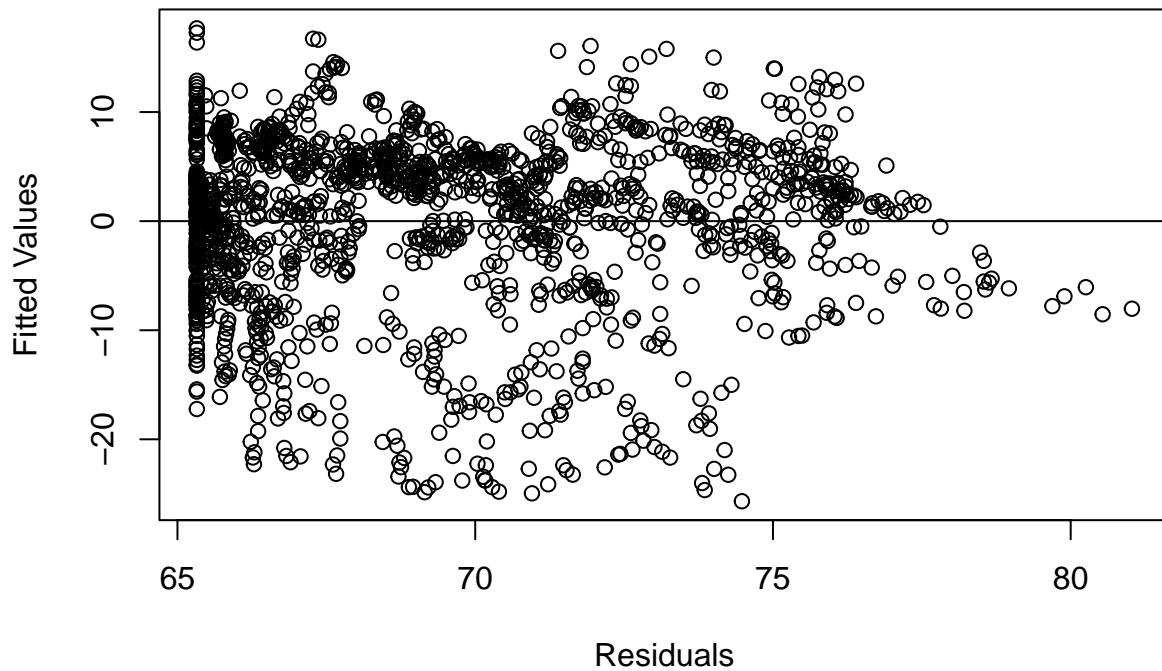
### Adult.Mortality Residual Plot



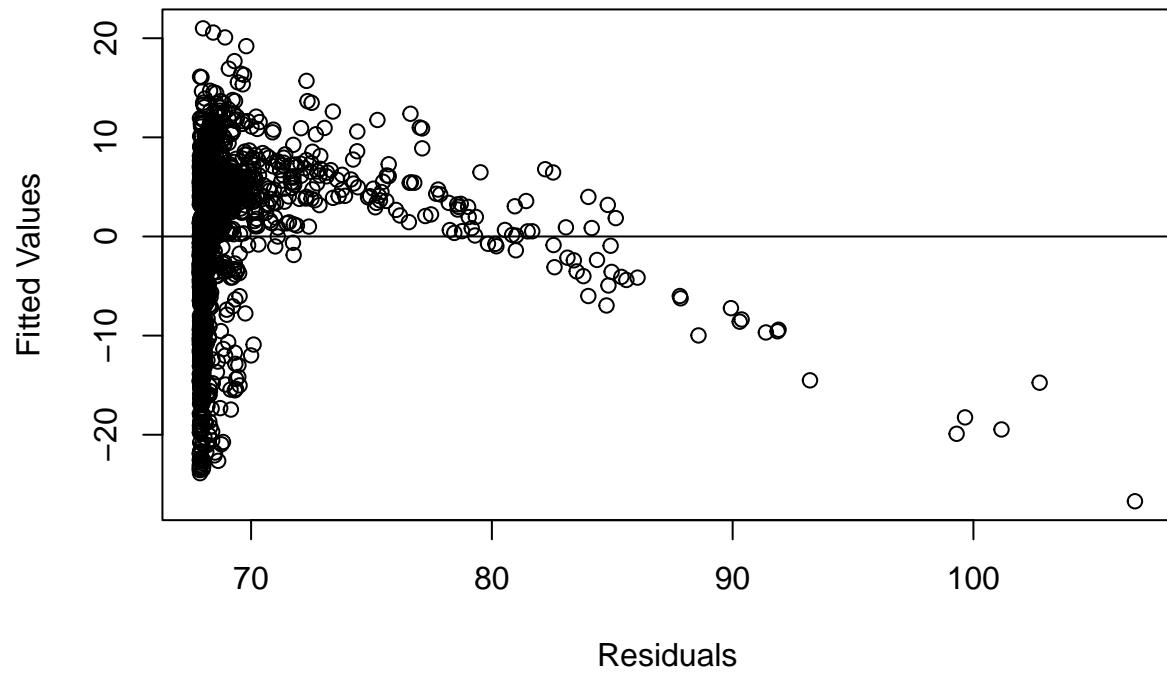
### infant.deaths Residual Plot



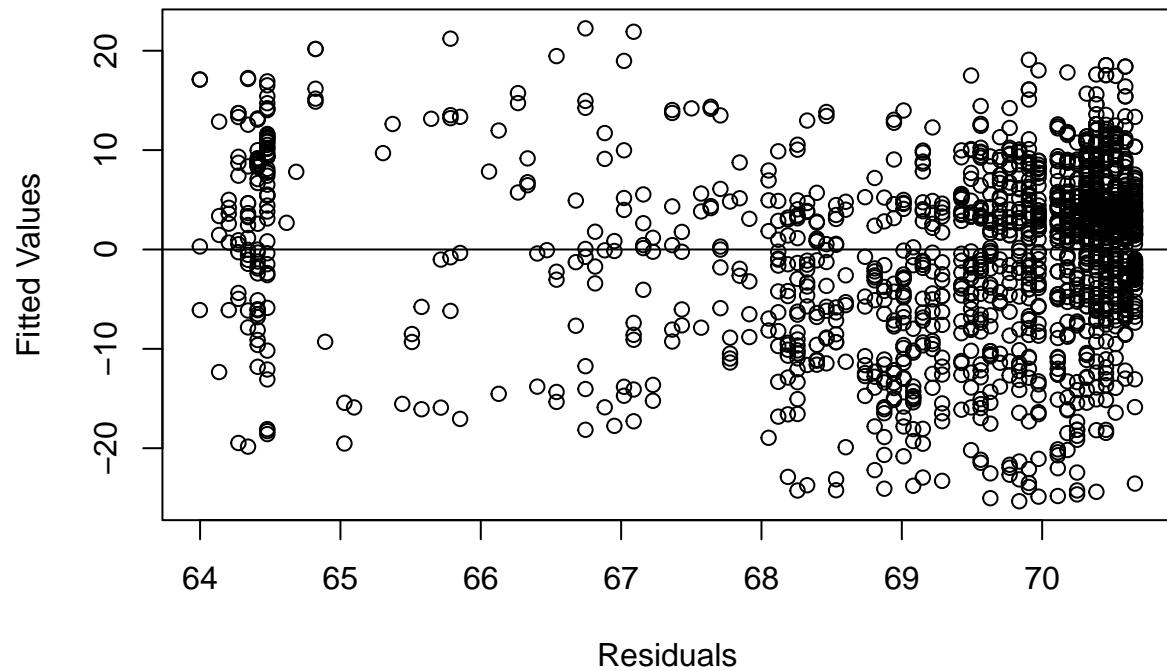
### Alcohol Residual Plot



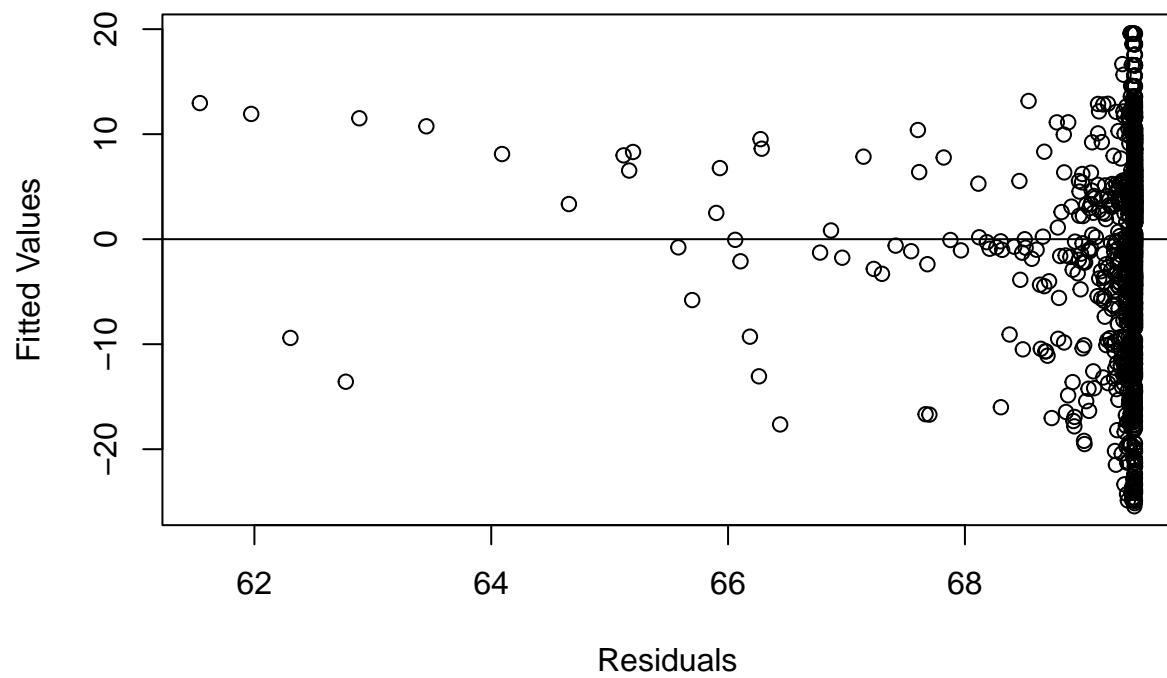
### **percentage.expenditure Residual Plot**



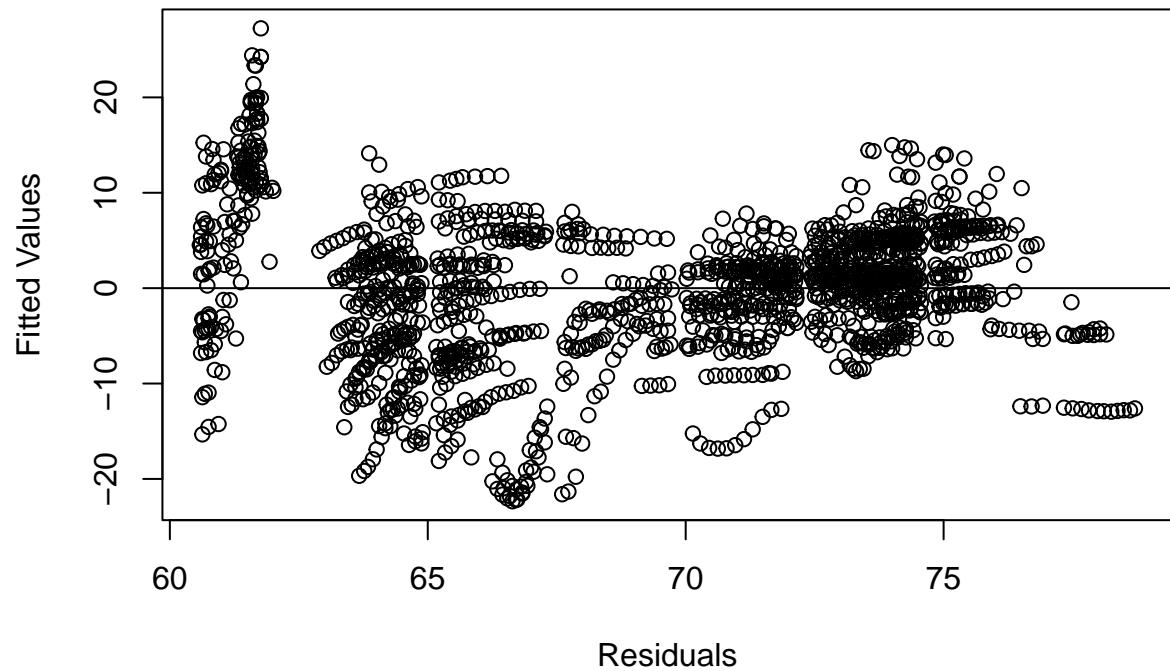
### Hepatitis.B Residual Plot



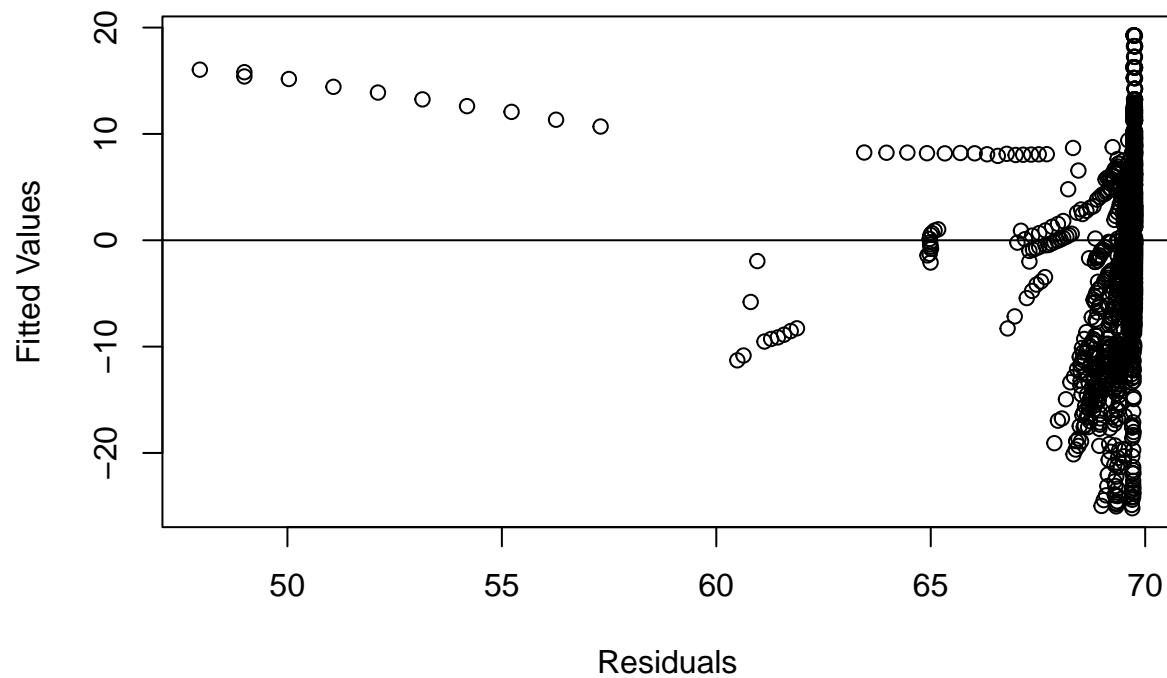
### Measles Residual Plot



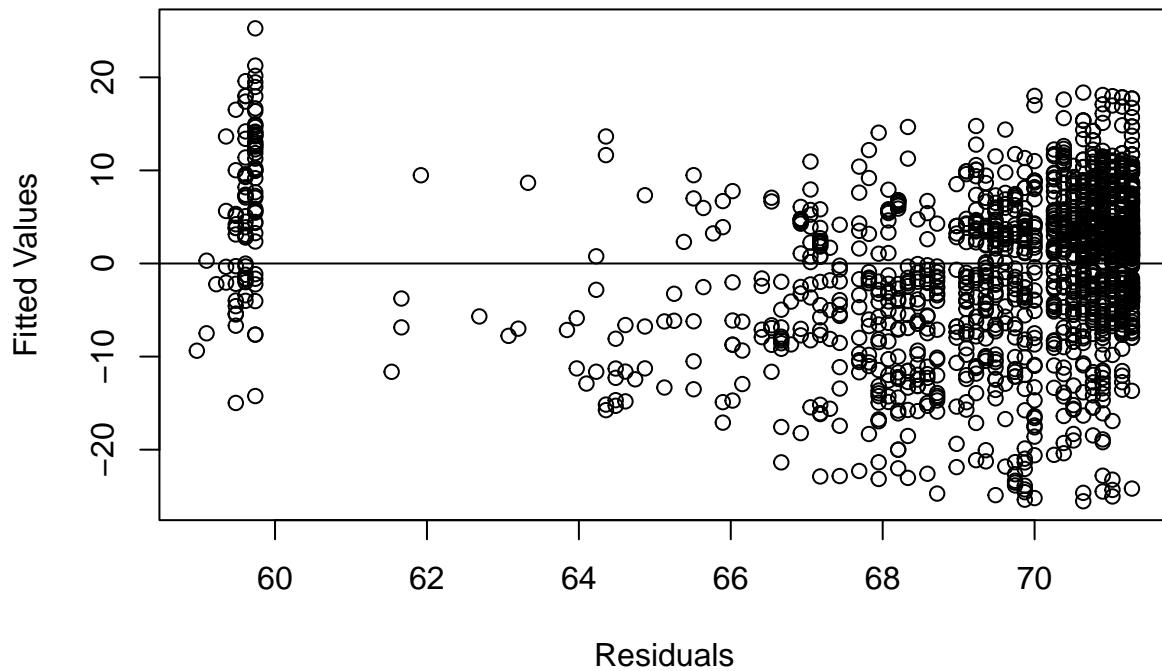
### BMI Residual Plot



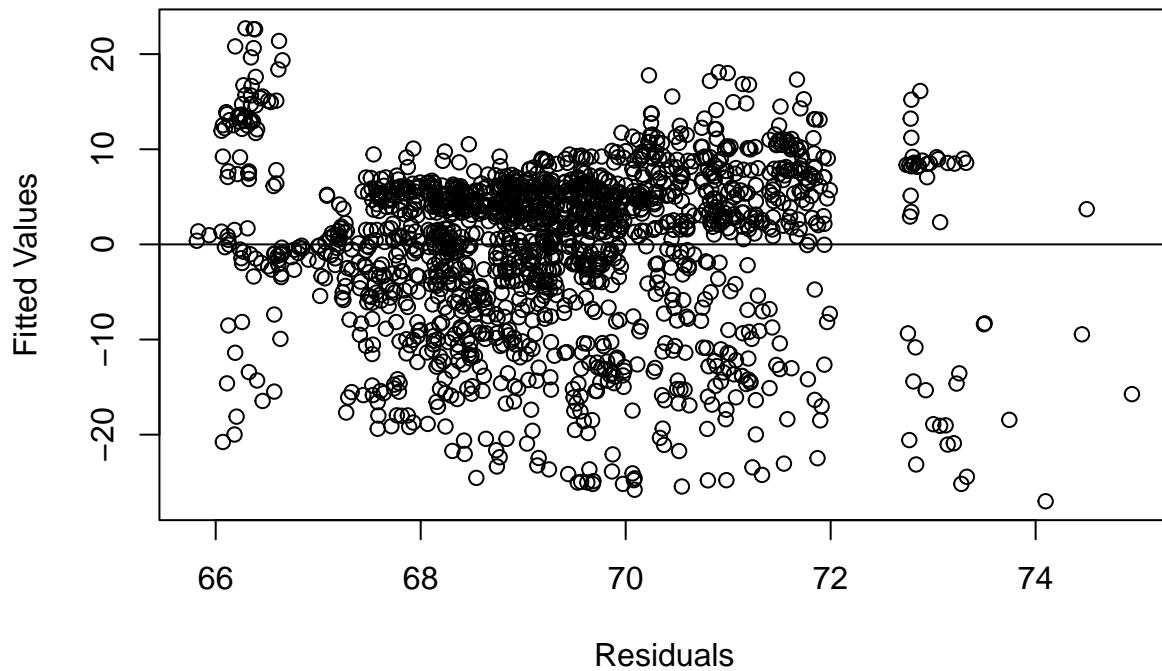
### **under.five.deaths Residual Plot**



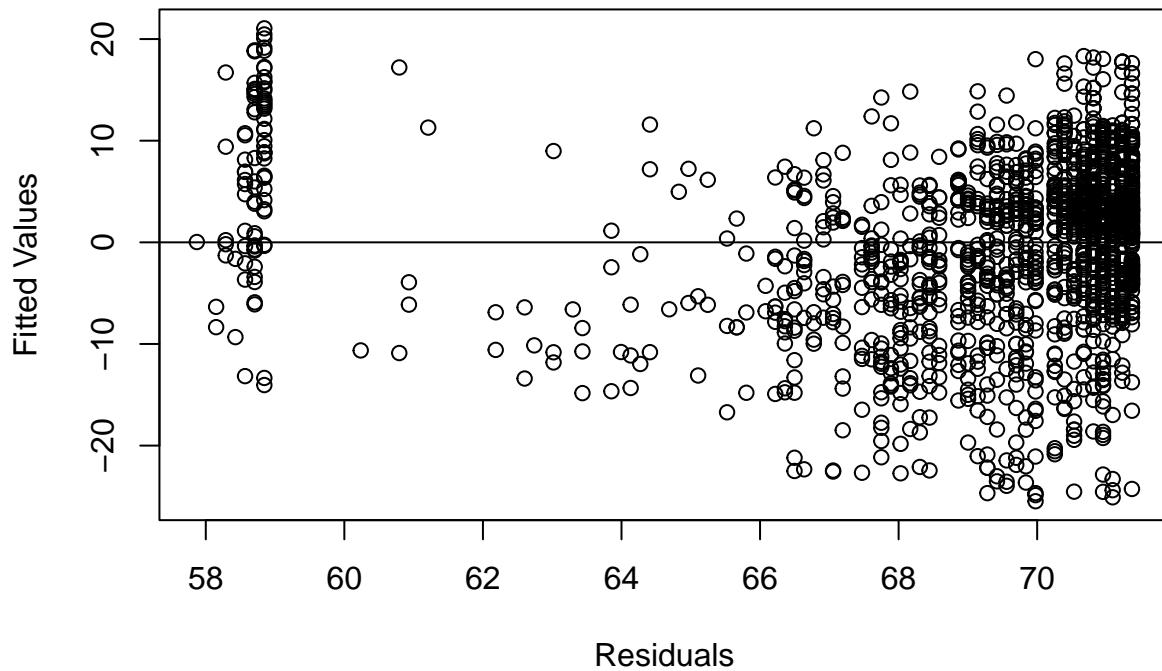
### Polio Residual Plot



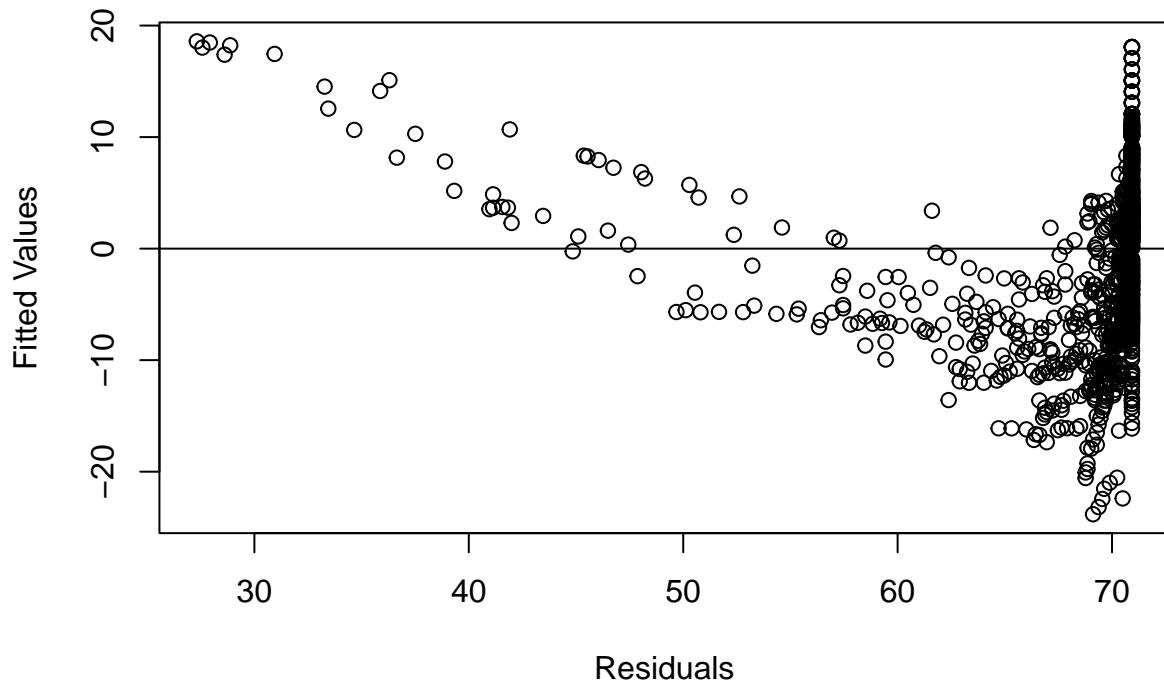
### Total.expenditure Residual Plot



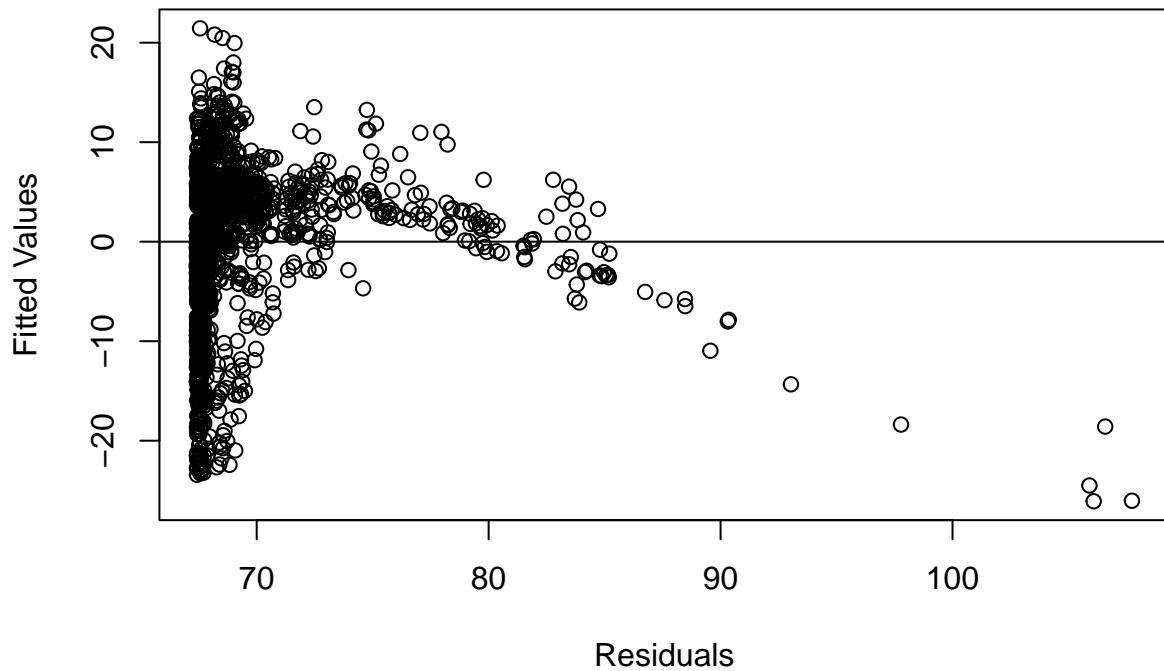
### Diphtheria Residual Plot



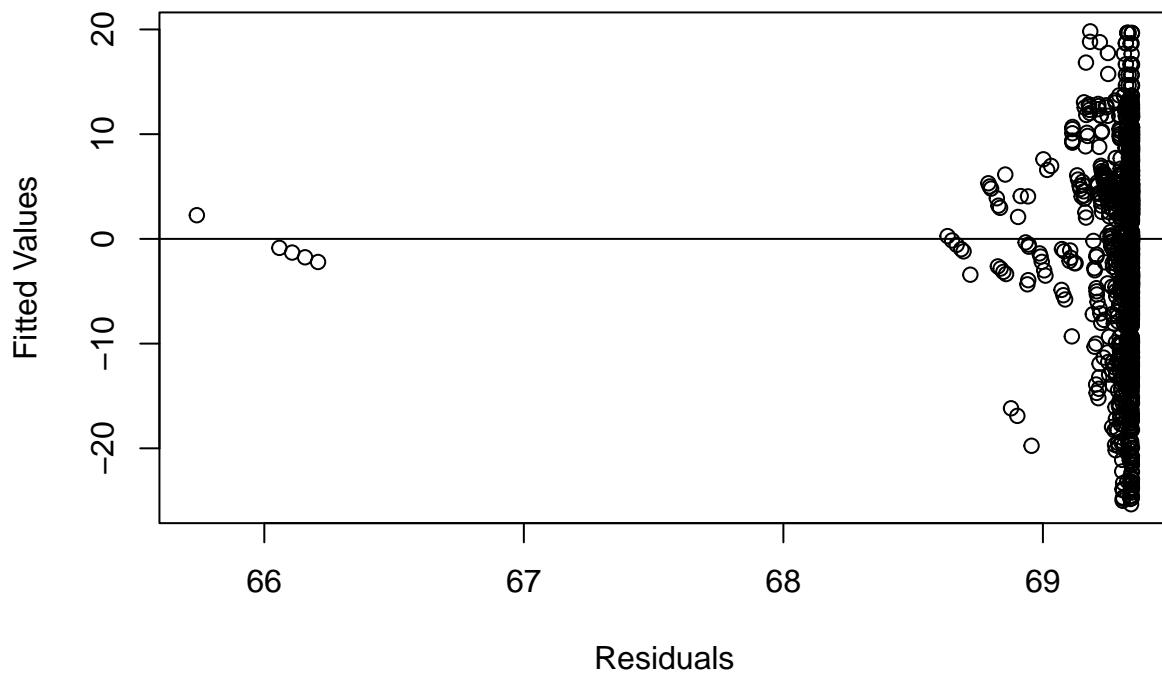
### HIV.AIDS Residual Plot



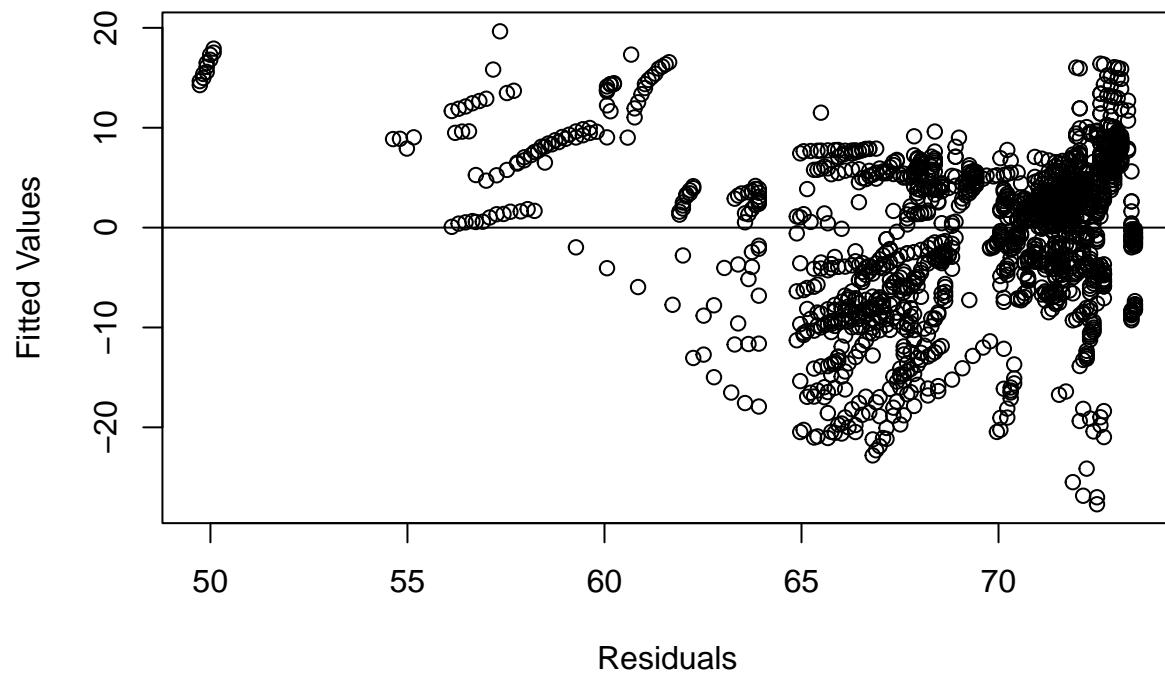
### GDP Residual Plot



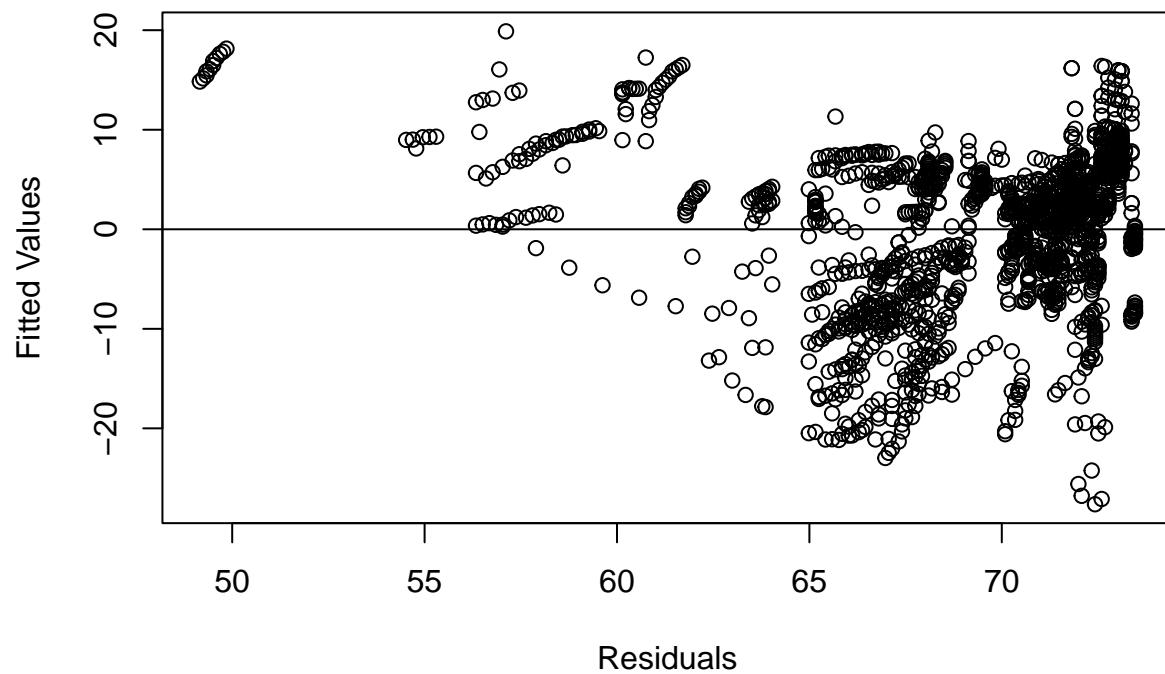
### Population Residual Plot



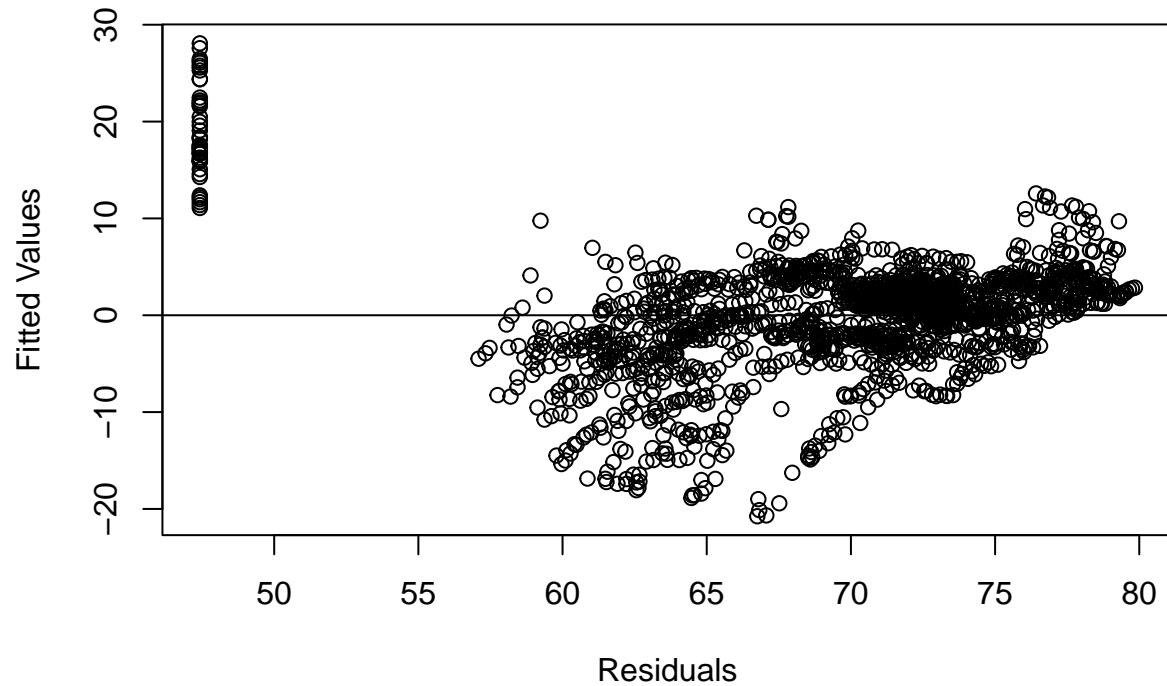
### **thinness..1.19.years Residual Plot**



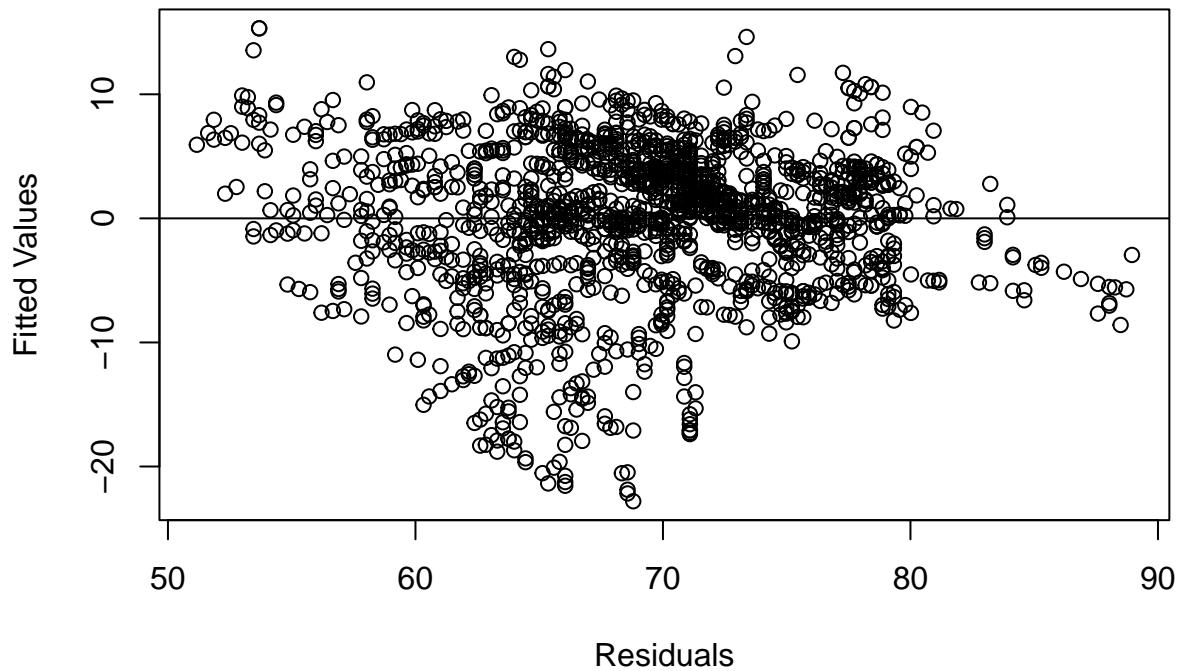
**thinness.5.9.years Residual Plot**



### Income.composition.of.resources Residual Plot

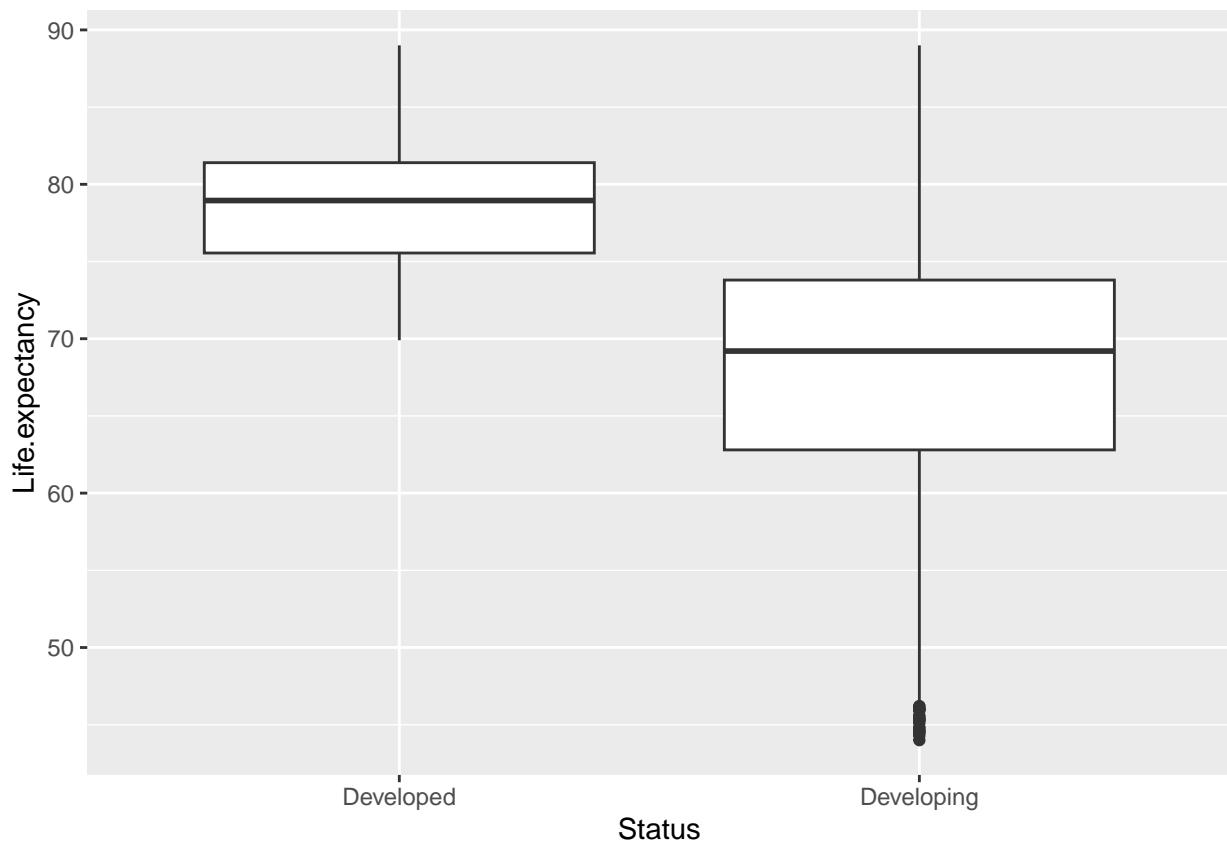


## Schooling Residual Plot

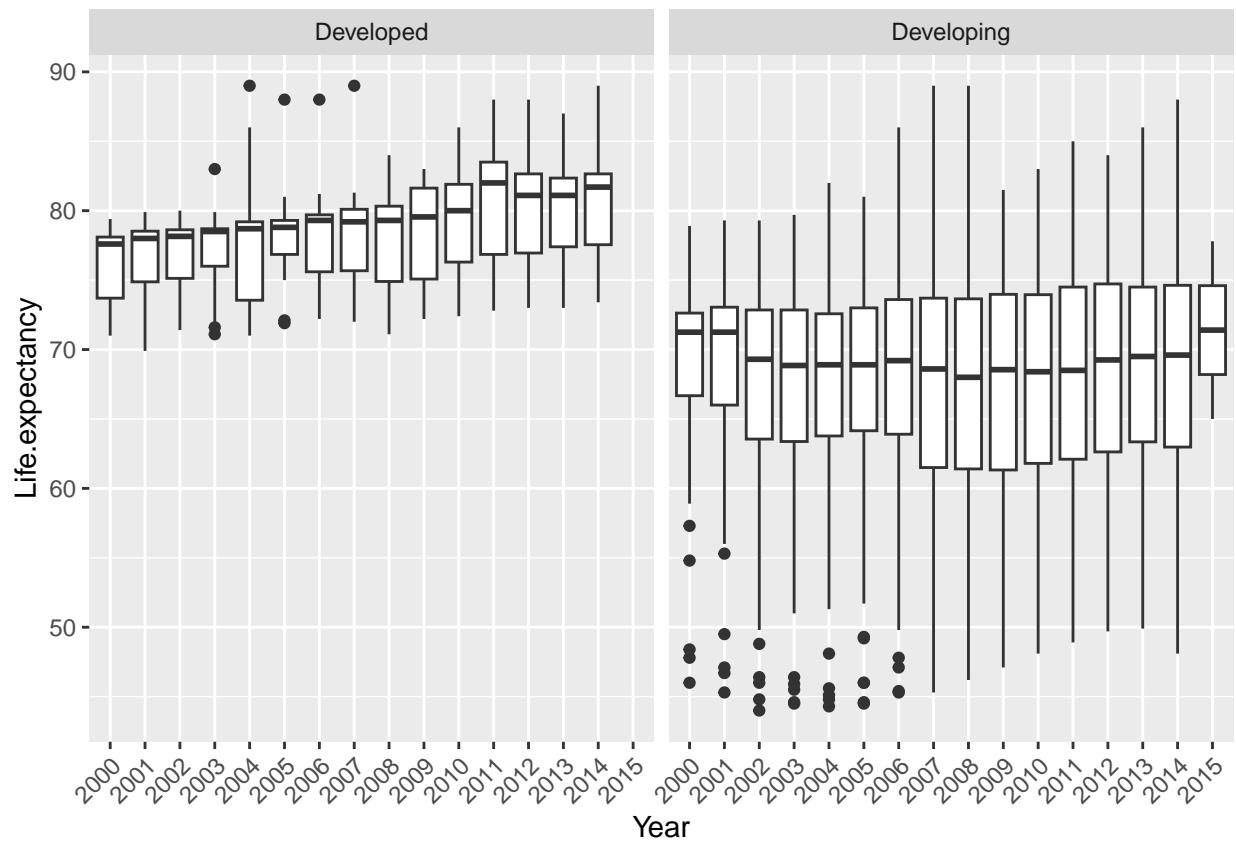


### Life Expectancy grouped by Status

```
data %>%
  ggplot(aes(Status, Life.expectancy)) +
  geom_boxplot()
```

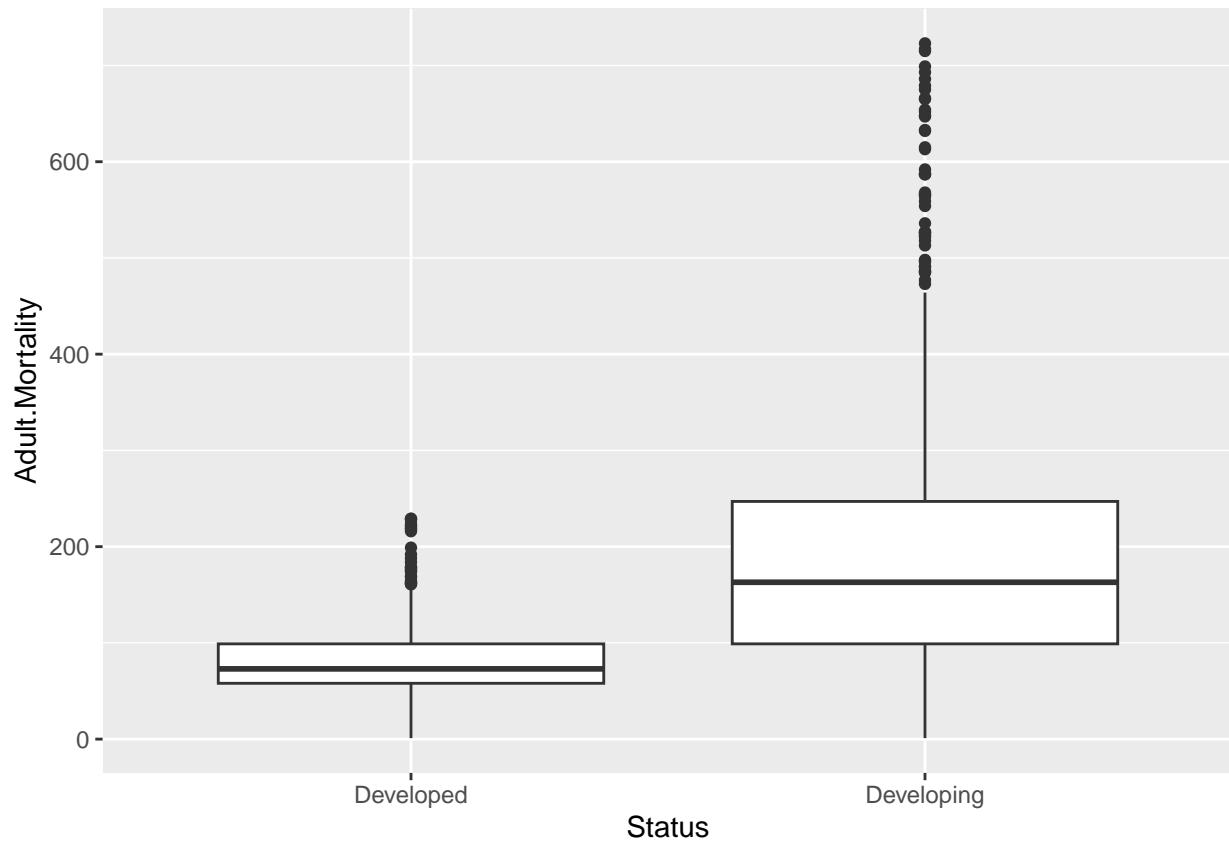


```
data%>%
  mutate(Year=as.factor(Year))%>%
  ggplot(aes(Year, Life.expectancy))+ 
  geom_boxplot()+
  facet_wrap(vars(Status))+ 
  scale_x_discrete(guide = guide_axis(angle = 45))
```

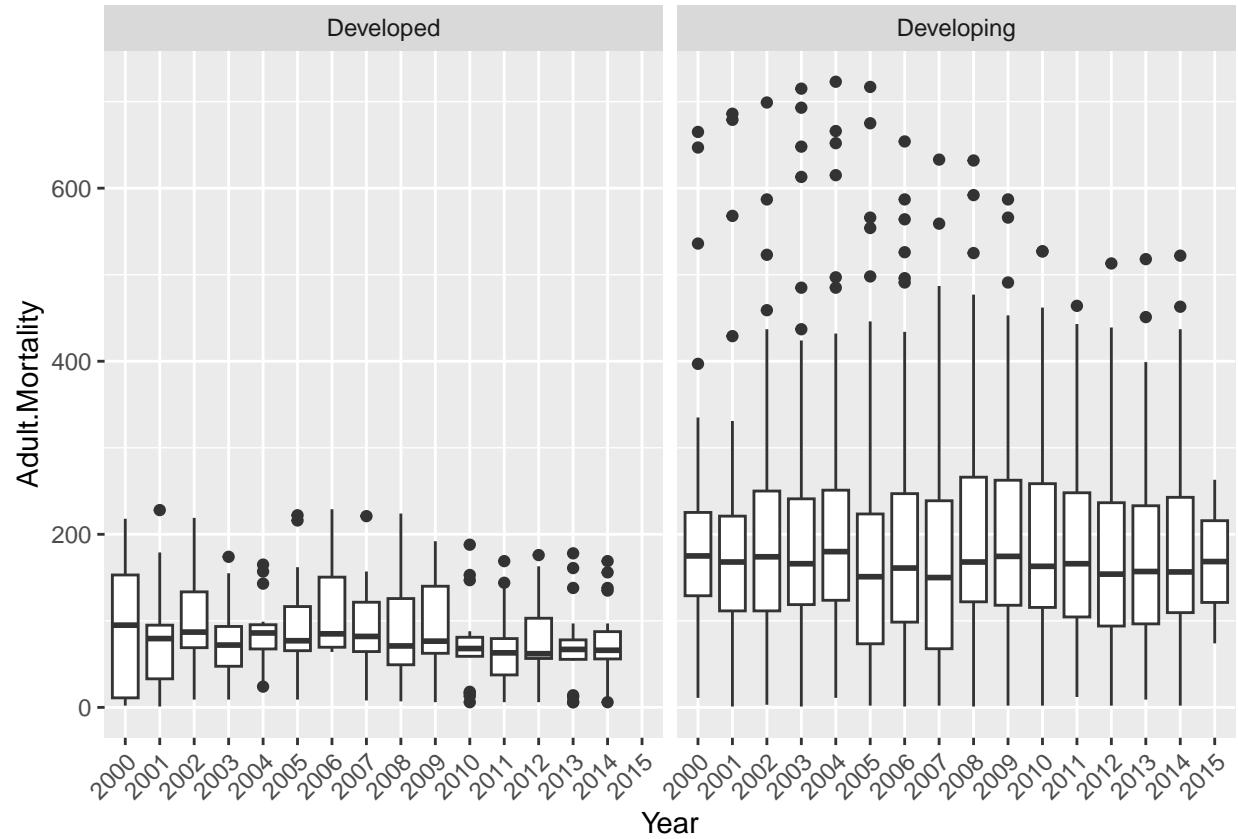


### Adult Mortality grouped by Status

```
data%>%
  ggplot(aes(Status, Adult.Mortality))+
  geom_boxplot()
```

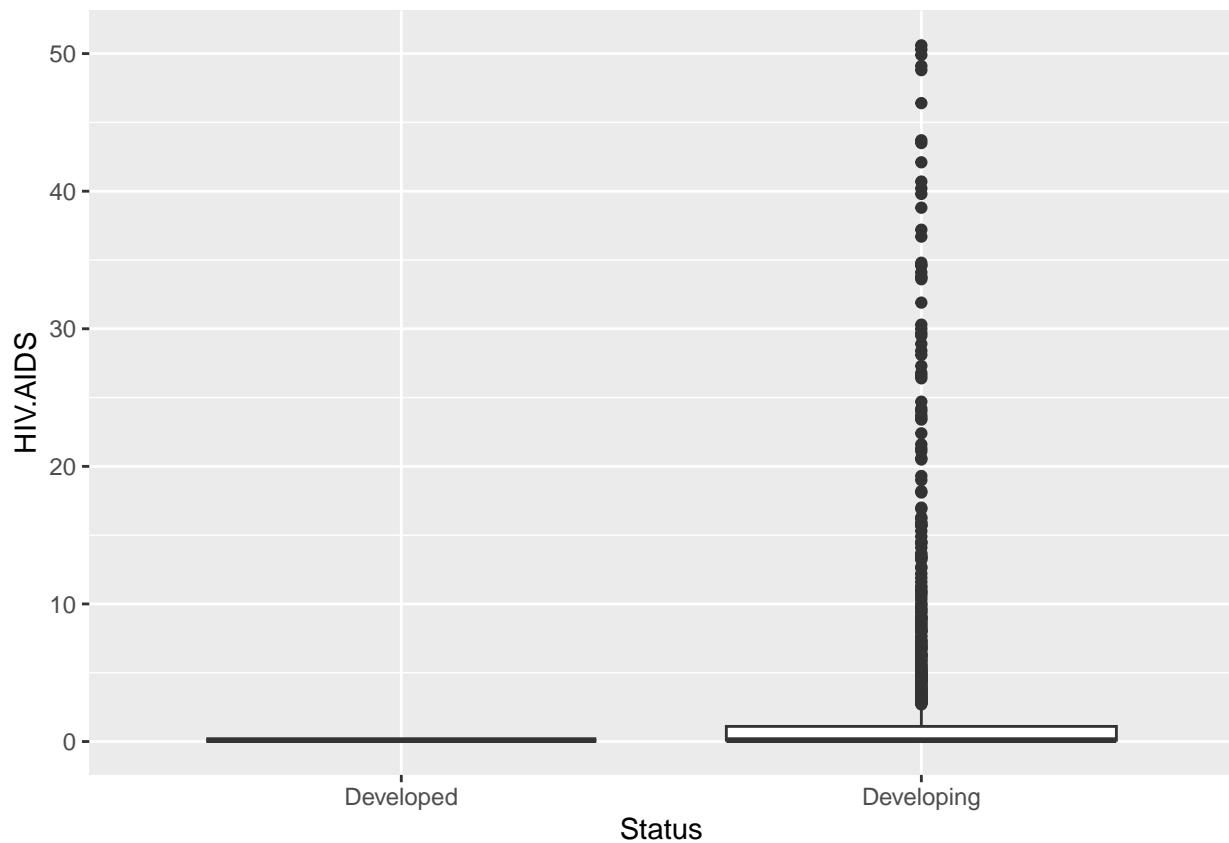


```
data%>%
  mutate(Year=as.factor(Year))%>%
  ggplot(aes(Year, Adult.Mortality))+  
  geom_boxplot()+
  facet_wrap(vars(Status))+  
  scale_x_discrete(guide = guide_axis(angle = 45))
```

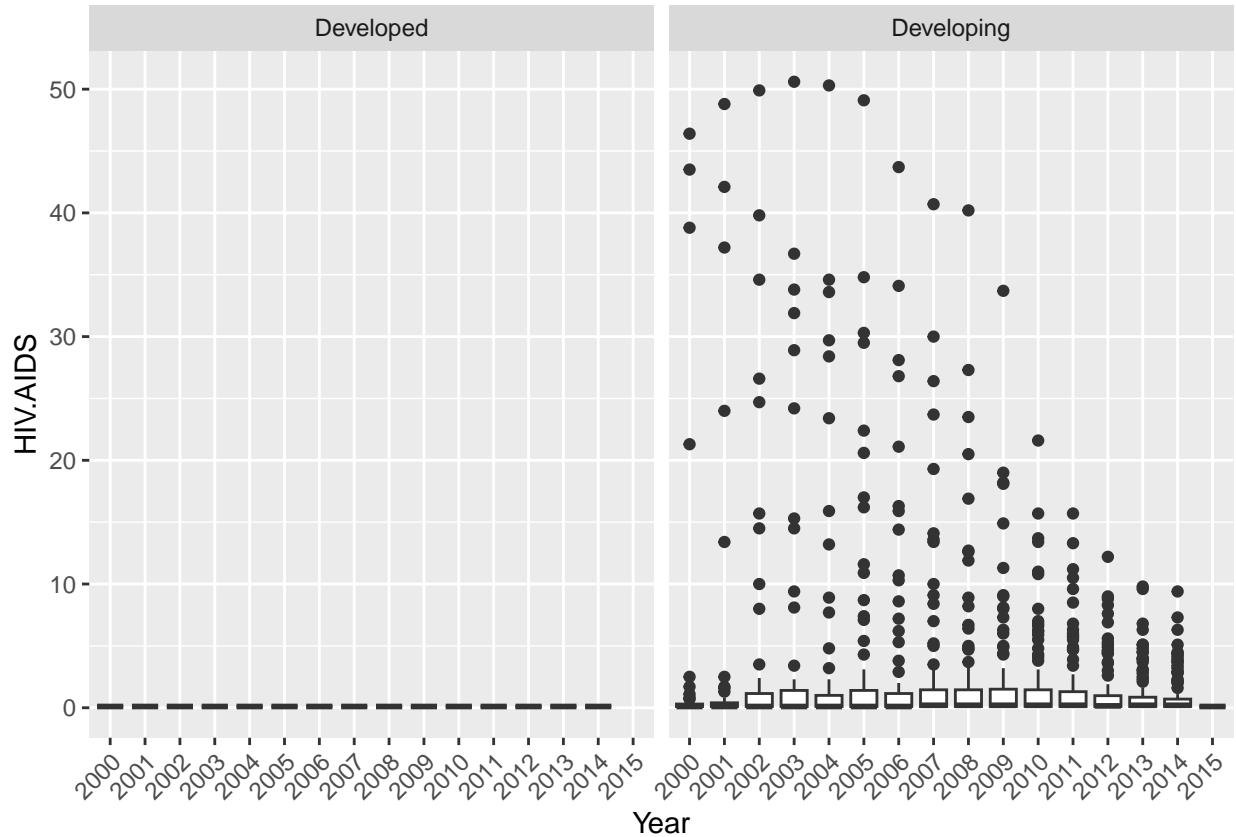


#### HIV/AIDS grouped by Status

```
data%>%
  ggplot(aes(Status, HIV.AIDS))+
  geom_boxplot()
```

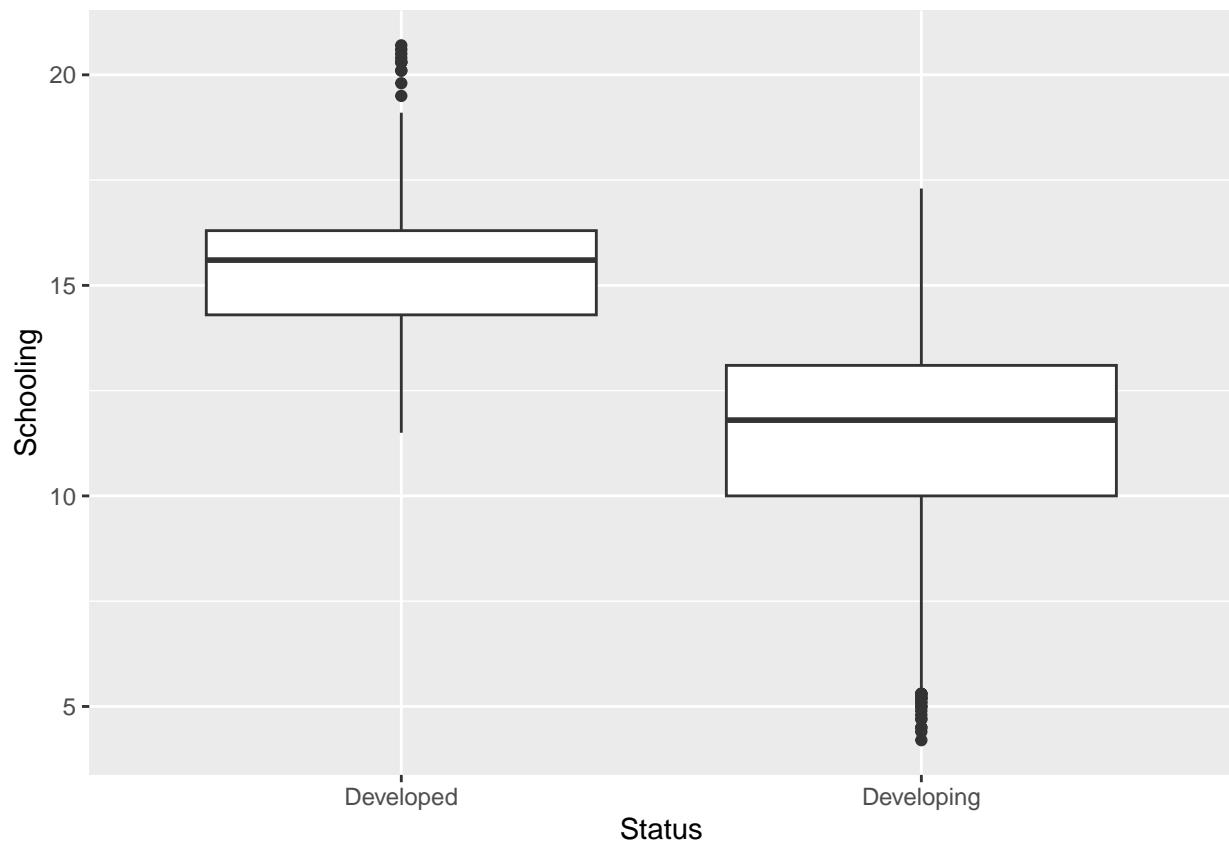


```
data%>%
  mutate(Year=as.factor(Year))%>%
  ggplot(aes(Year, HIV.AIDS))+
  geom_boxplot()+
  facet_wrap(vars(Status))+
  scale_x_discrete(guide = guide_axis(angle = 45))
```



#### Schooling grouped by Status

```
data%>%
  ggplot(aes(Status, Schooling))+
  geom_boxplot()
```



```
data%>%
  mutate(Year=as.factor(Year))%>%
  ggplot(aes(Year, Schooling))+
  geom_boxplot()+
  facet_wrap(vars(Status))+
  scale_x_discrete(guide = guide_axis(angle = 45))
```

