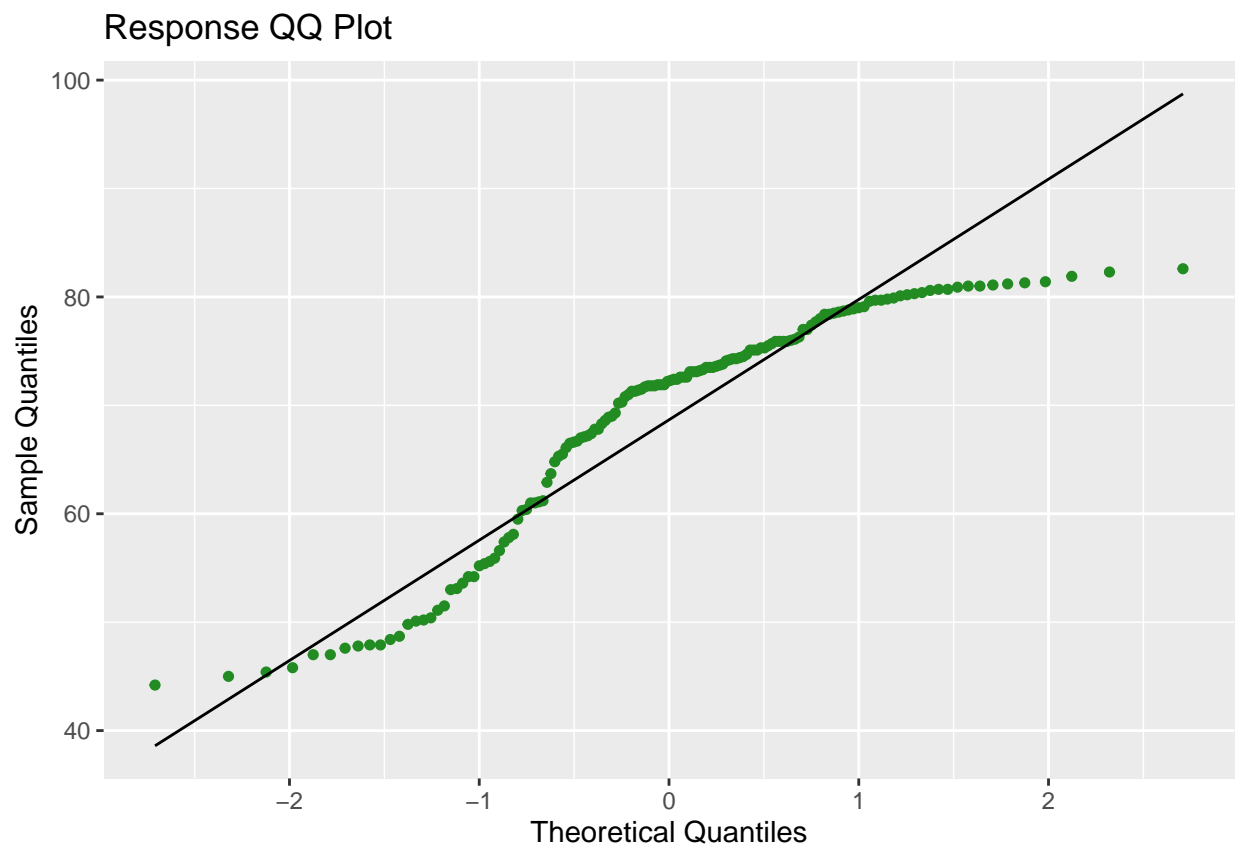# STA108_TermProjectEdits

Gabriel Jones

2023-11-23

```
## Warning: package 'leaps' was built under R version 4.3.2
```

Single Linear Regression Models & Analysis

Life Expectancy Analysis

Life Expectancy ~ Land Area

Table 1: Life Expectancy ~ Land Area

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	68.6437149	0.9344772	73.4568137	0.0000000
Land Area	0.0000002	0.0000004	0.4899179	0.6249274

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

Land Area QQ Plot

Residual QQ Plot
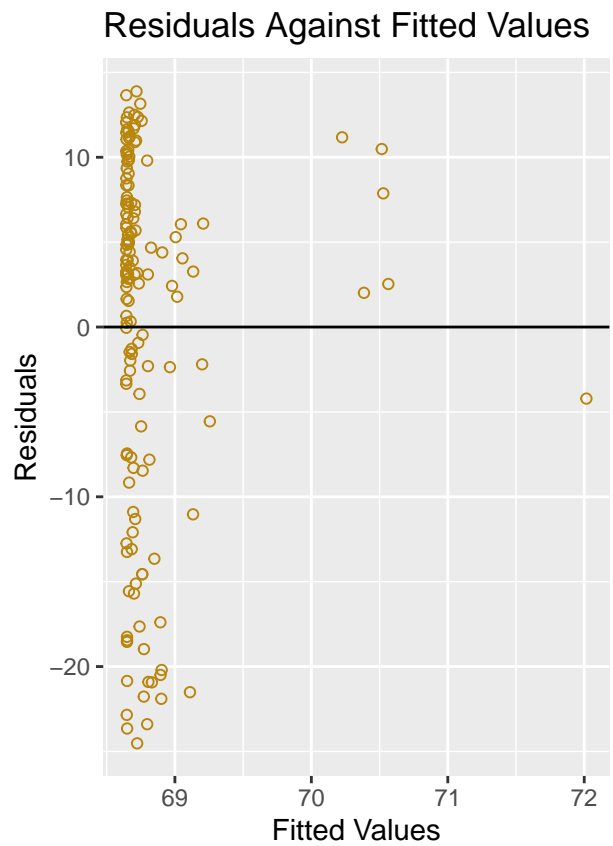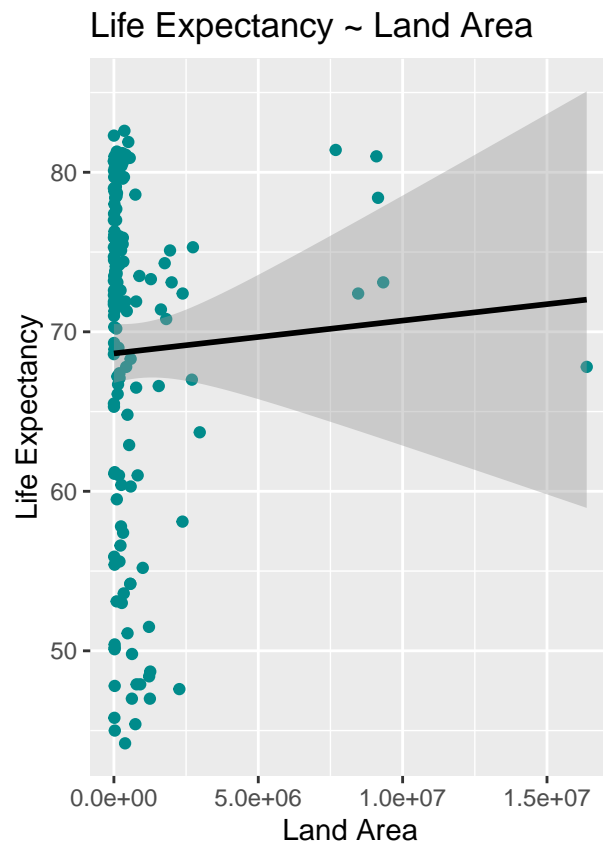
Life Expectancy ~ Birth Rate

Table 2: Life Expectancy ~ Birth Rate

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	88.2252805	1.0469920	84.26548	0
Birth Rate	-0.8837966	0.0431595	-20.47747	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```
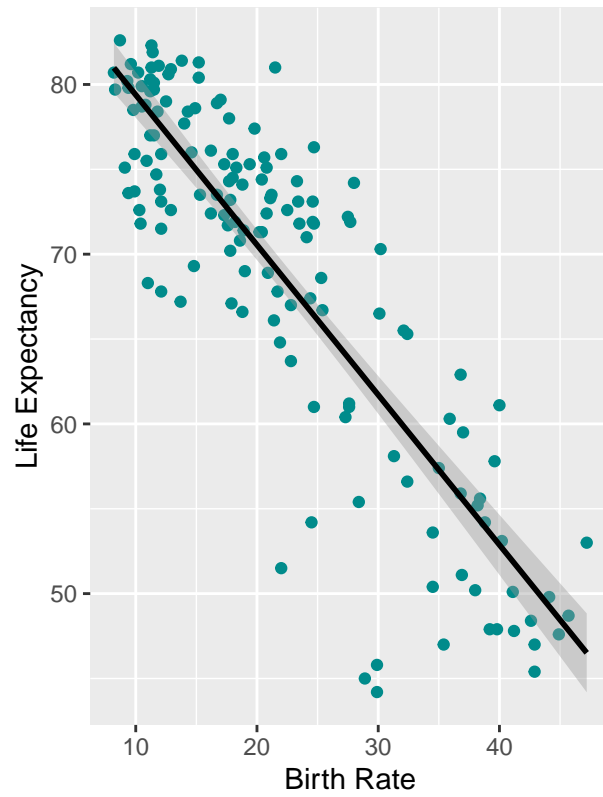
Birth Rate QQ Plot

Residual QQ Plot

Life Expectancy ~ Cell

Table 3: Life Expectancy ~ Cell

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	54.262300	1.5330388	35.39526	0
Cell	0.158516	0.0150848	10.50833	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

Life Expectancy ~ Cell

Residuals Against Fitted Values

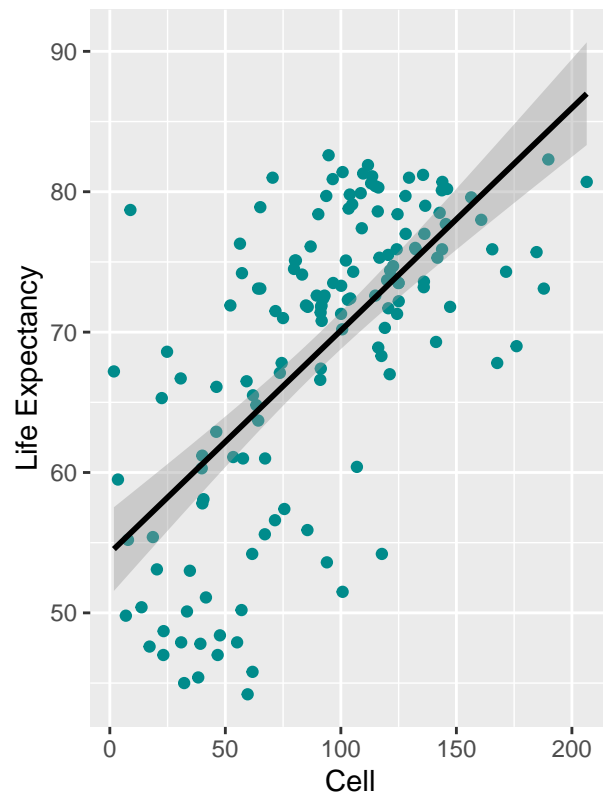## Cell QQ Plot

Residual QQ Plot
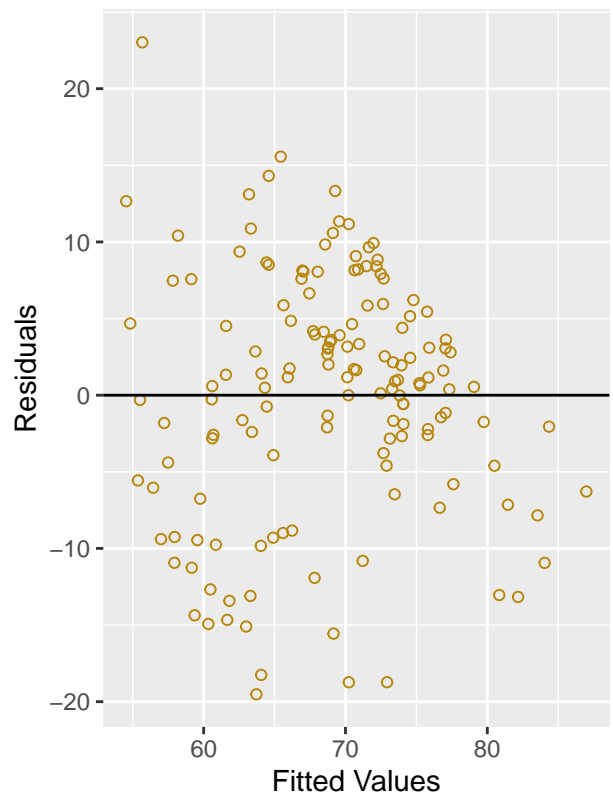
Life Expectancy ~ CO2

Table 4: Life Expectancy ~ CO2

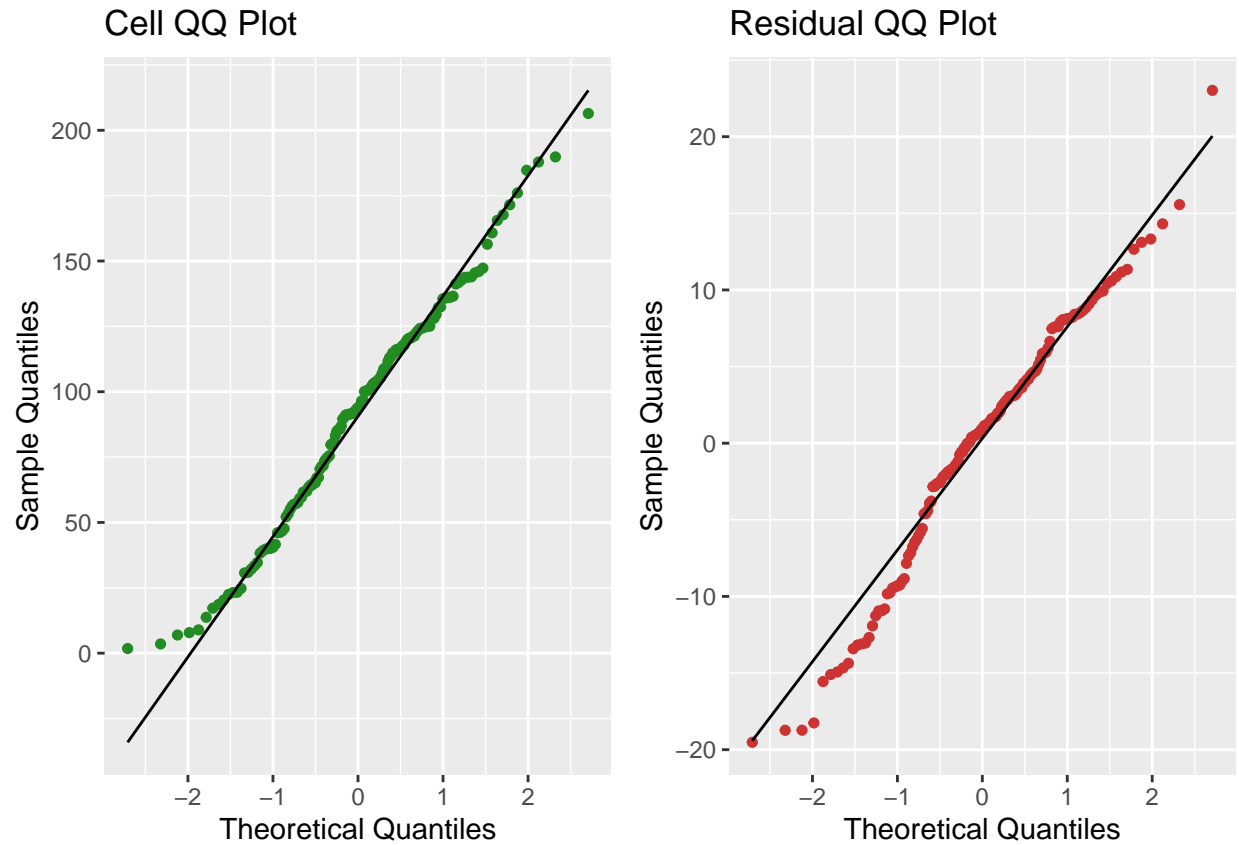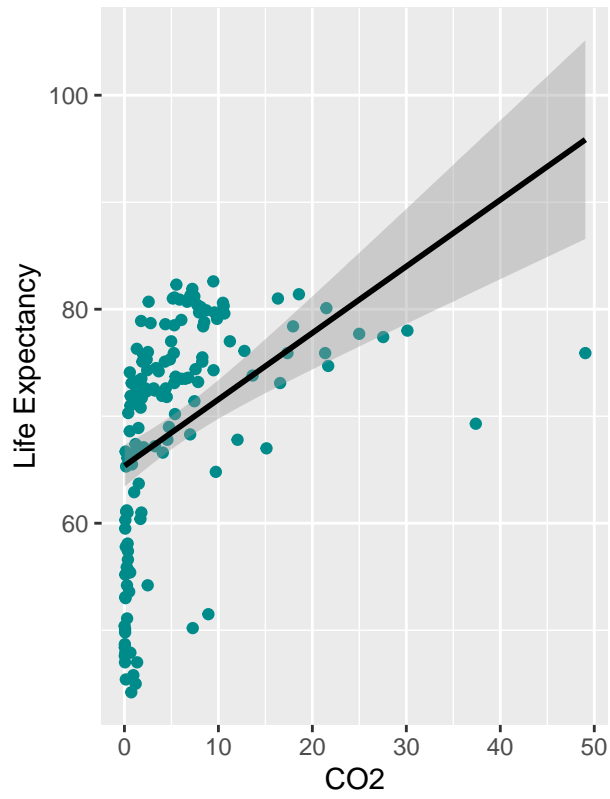|	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	65.3529718	0.9840266	66.413824	0
CO2	0.6216755	0.1064559	5.839747	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

CO2 QQ Plot

Residual QQ Plot
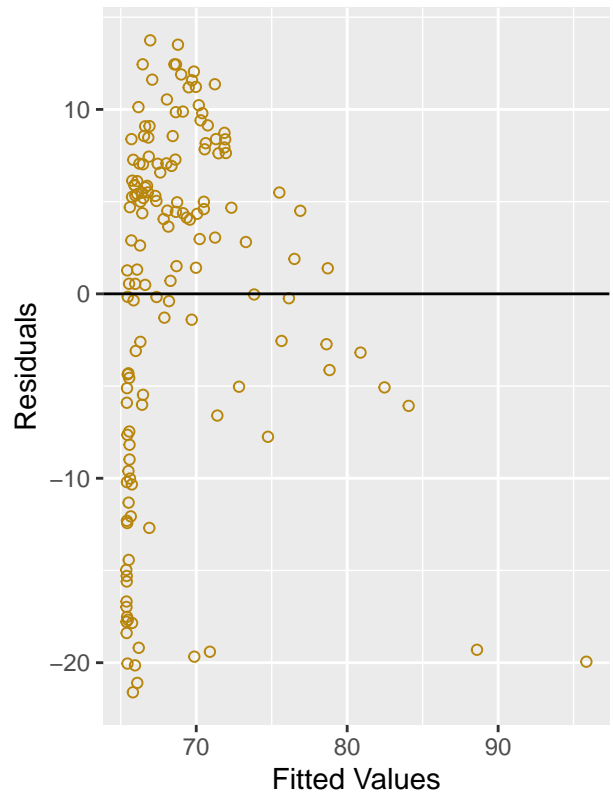
Life Expectancy ~ Elderly Population

Table 5: Life Expectancy ~ Elderly Population

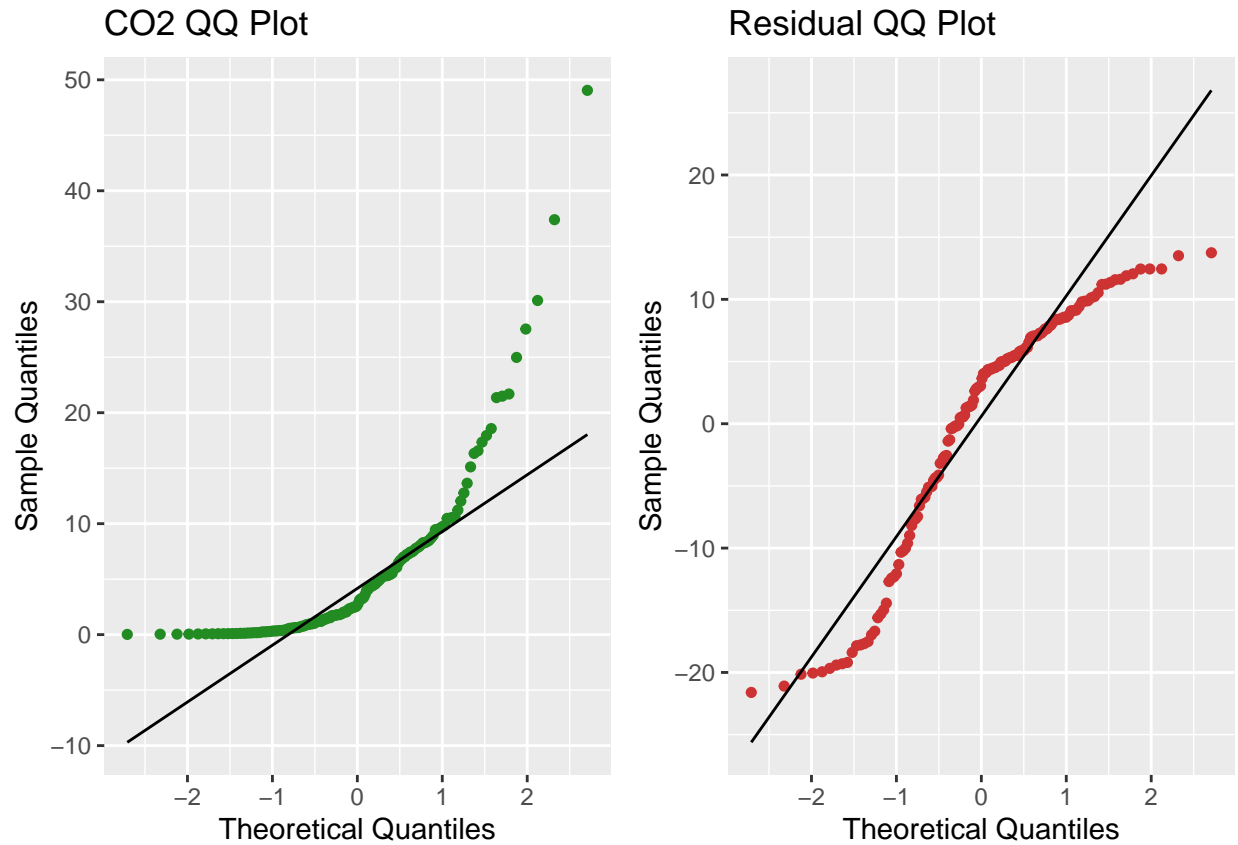|	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	58.987283	1.2078486	48.836652	0
Elderly Population	1.312309	0.1336064	9.822199	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

Elderly Population QQ Plot

Residual QQ Plot
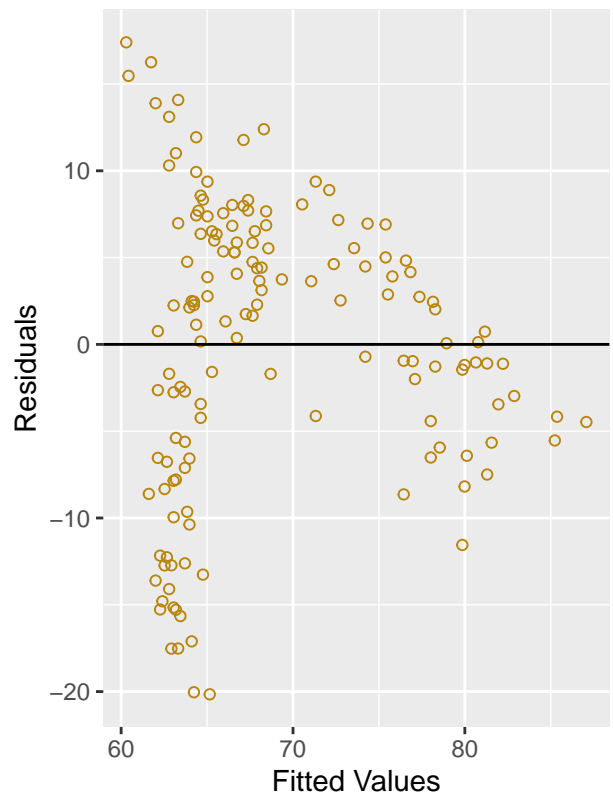
Life Expectancy ~ GDP

Table 6: Life Expectancy ~ GDP

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	64.2487516	0.8867131	72.457200	0
GDP	0.0003399	0.0000395	8.610075	0

```
## Saving 6.5 x 4.5 in image
## `geom_smooth()` using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## `geom_smooth()` using formula = 'y ~ x'
```

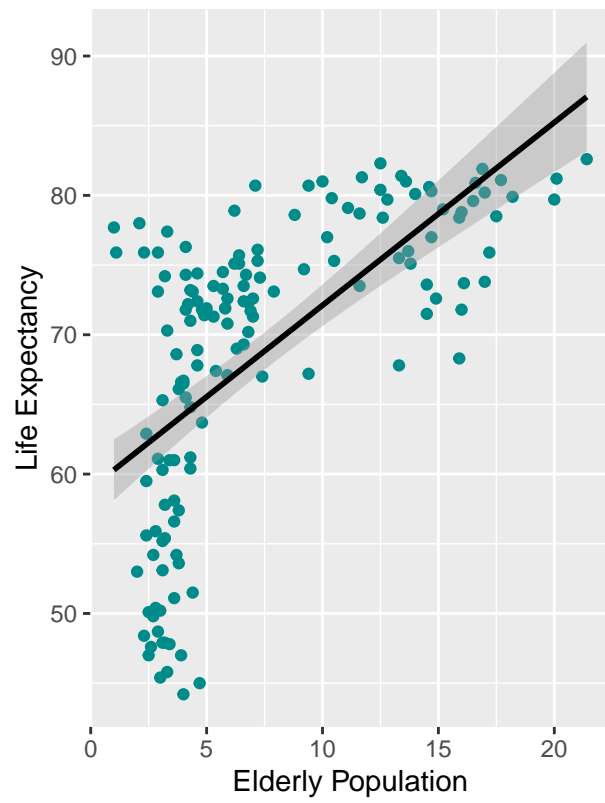## Life Expectancy ~ GDP

Residuals Against Fitted Values

GDP QQ Plot

Residual QQ Plot
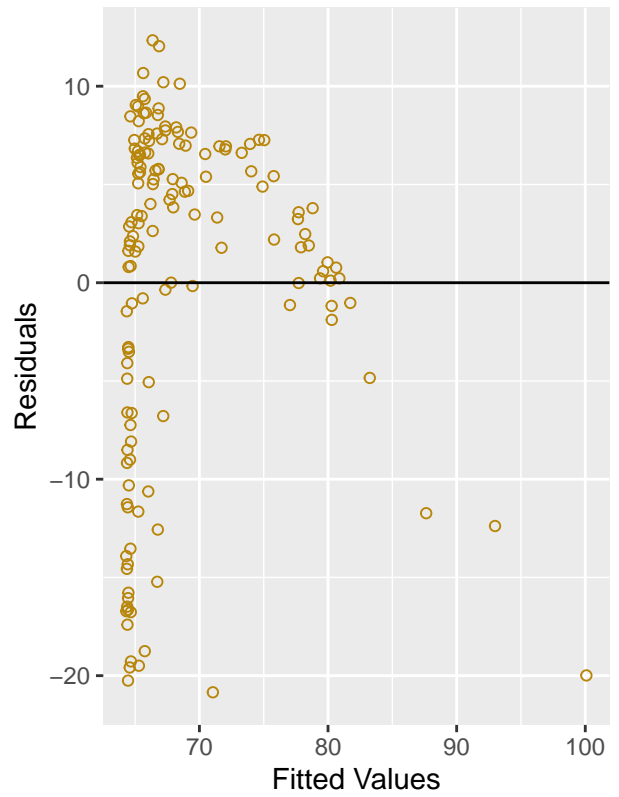
Life Expectancy ~ Health

Table 7: Life Expectancy ~ Health

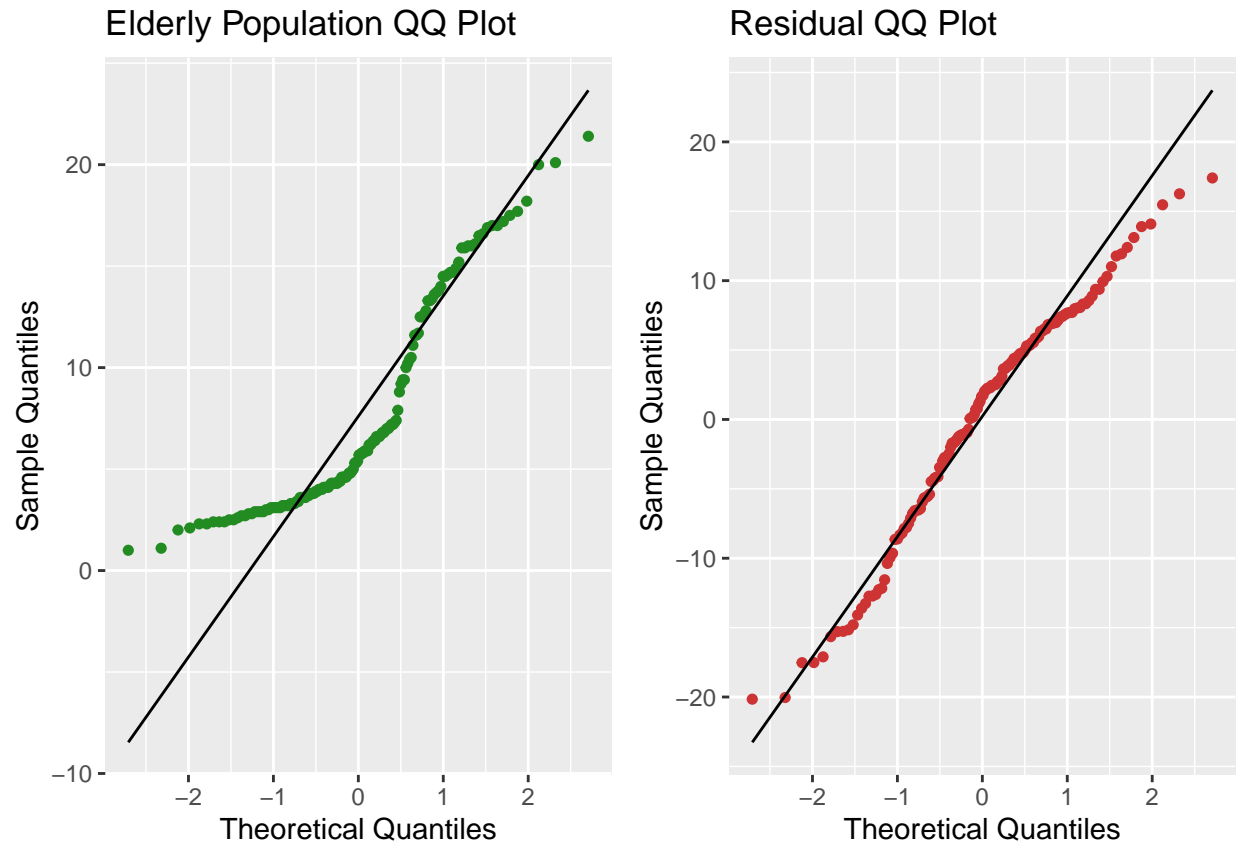|	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	60.6696446	2.3307451	26.030150	0.0000000
Health	0.7392994	0.1977105	3.739303	0.0002644

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

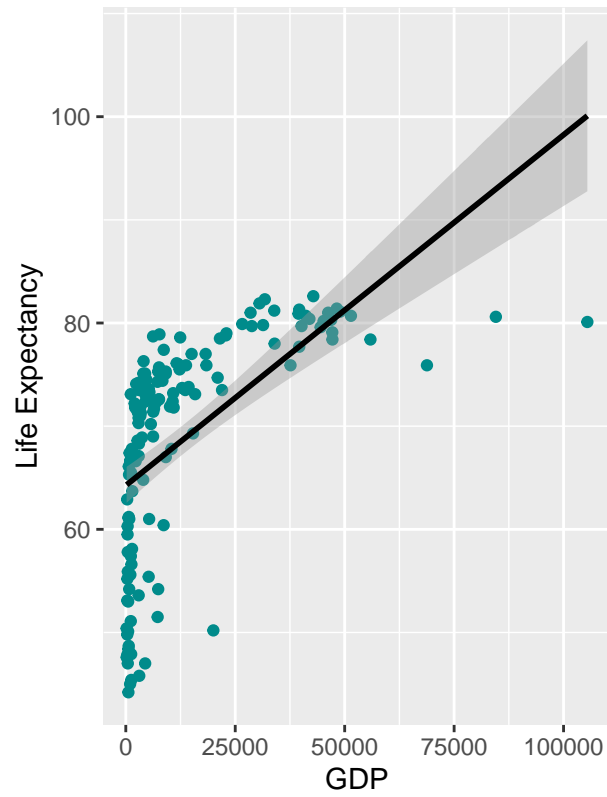## Life Expectancy ~ Health

Residuals Against Fitted Values

Health QQ Plot

Residual QQ Plot
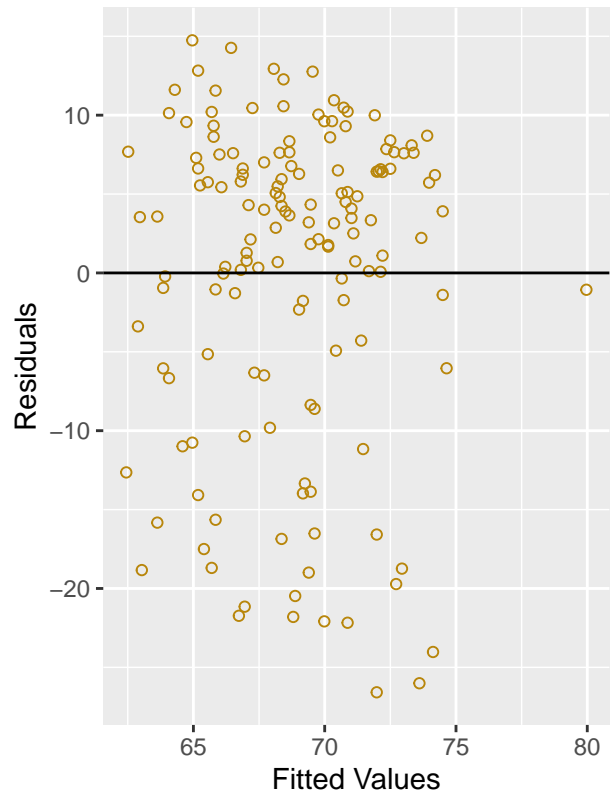
Life Expectancy ~ Internet

Table 8: Life Expectancy ~ Internet

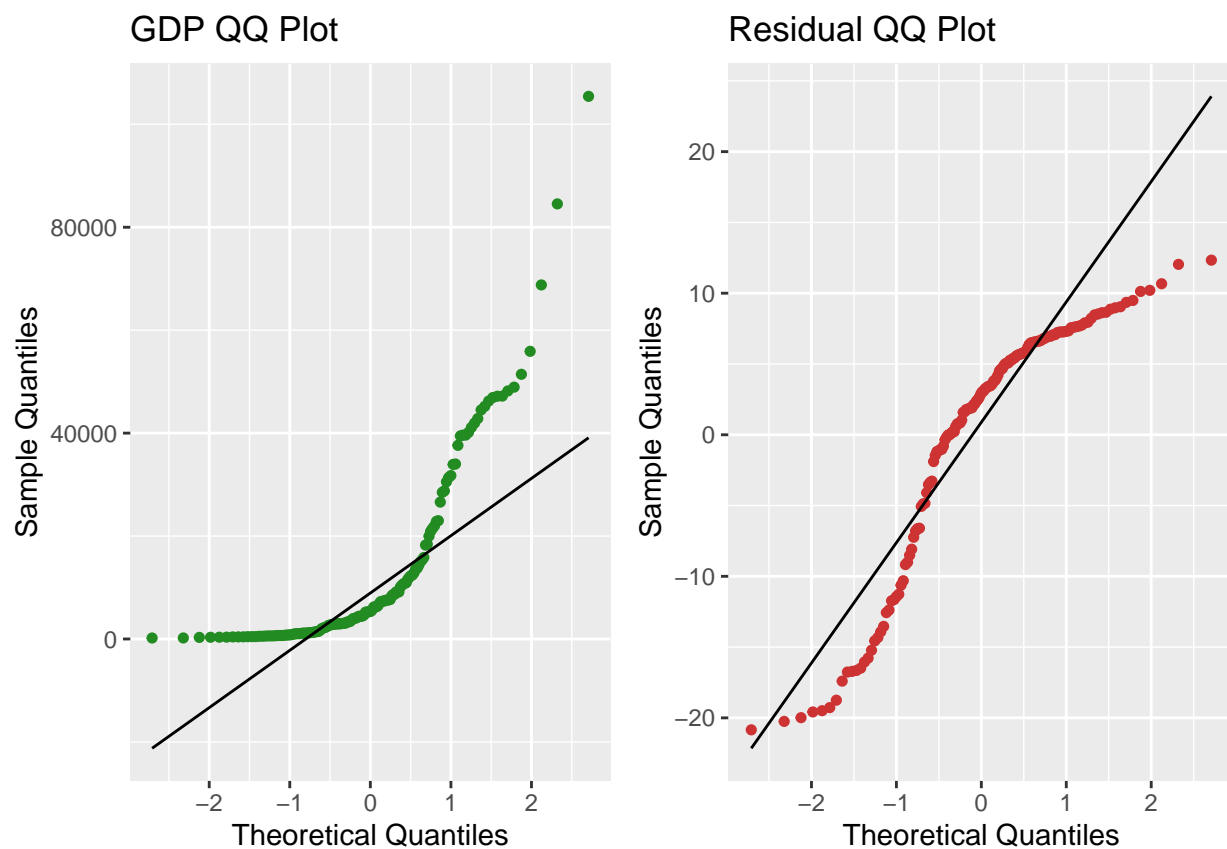|	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	60.725099	0.8927376	68.02122	0
Internet	0.286114	0.0230849	12.39400	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

Life Expectancy ~ Internet
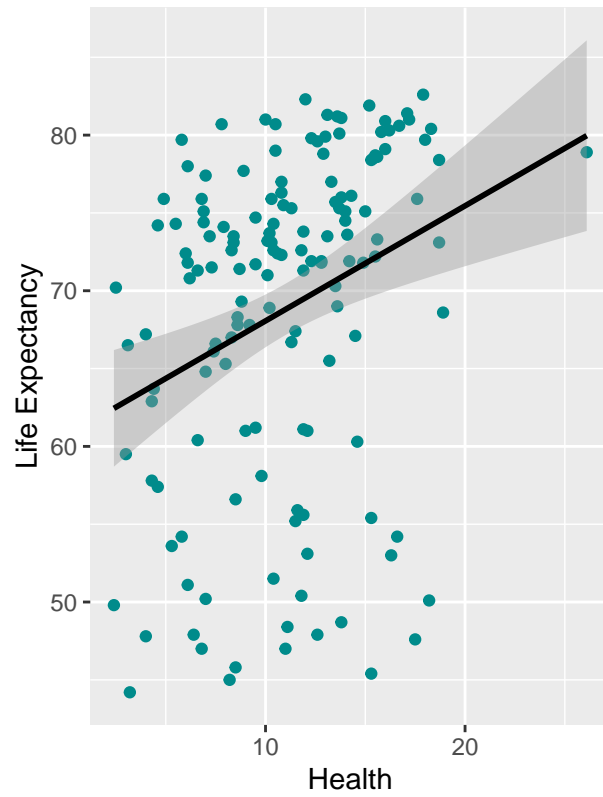
Residuals Against Fitted Values

Internet QQ Plot

Residual QQ Plot
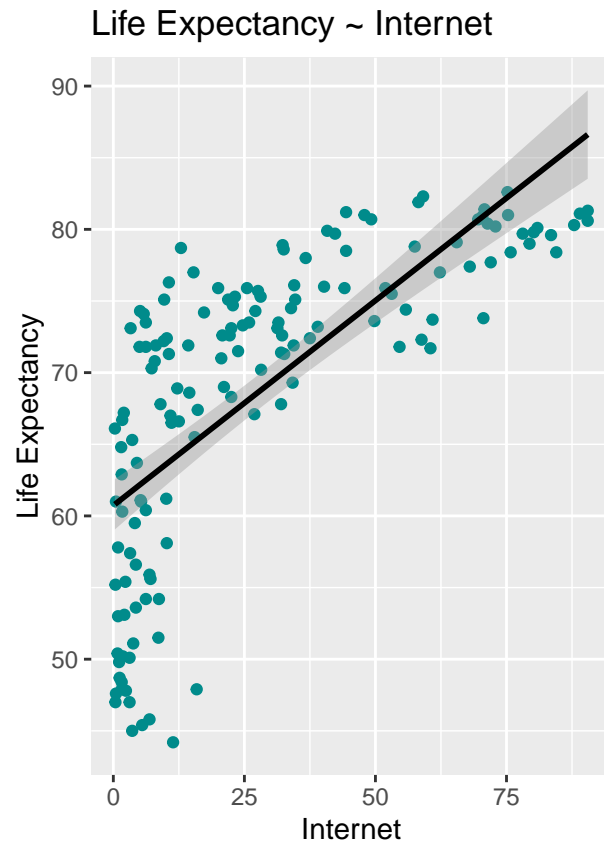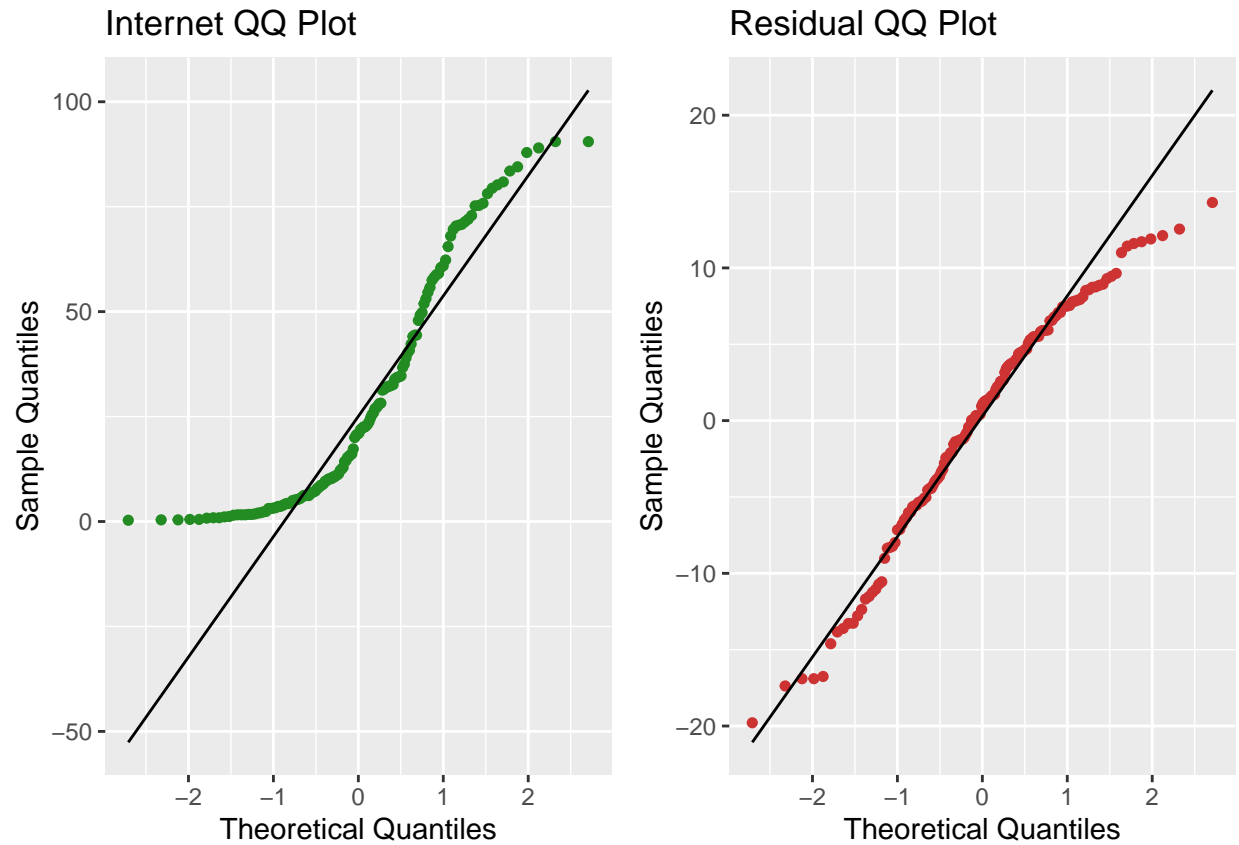
Life Expectancy ~ Population

Table 9: Life Expectancy ~ Population

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	68.7894859	0.9058768	75.9369106	0.0000000
Population	0.0004438	0.0059454	0.0746386	0.9406045

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

Life Expectancy ~ Population

Residuals Against Fitted Values

Population QQ Plot

Residual QQ Plot
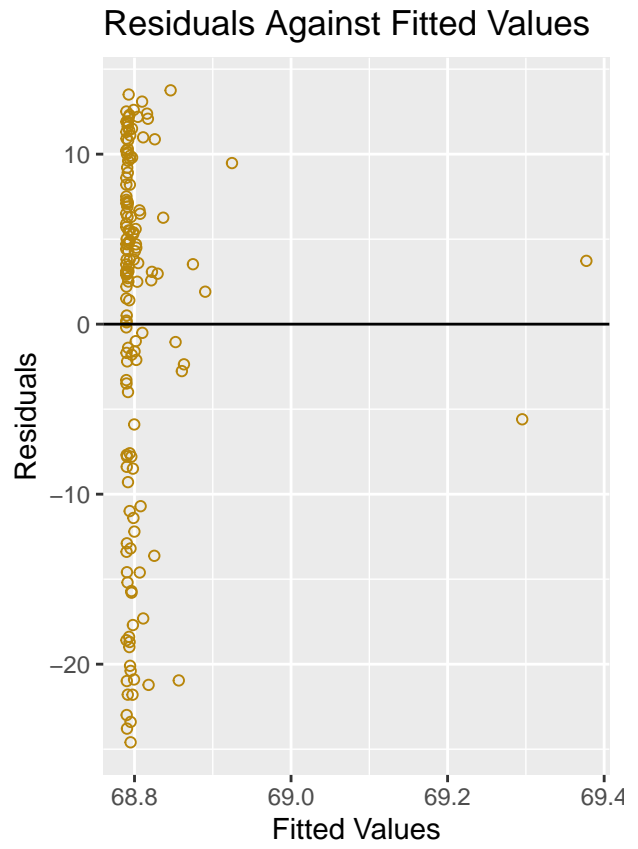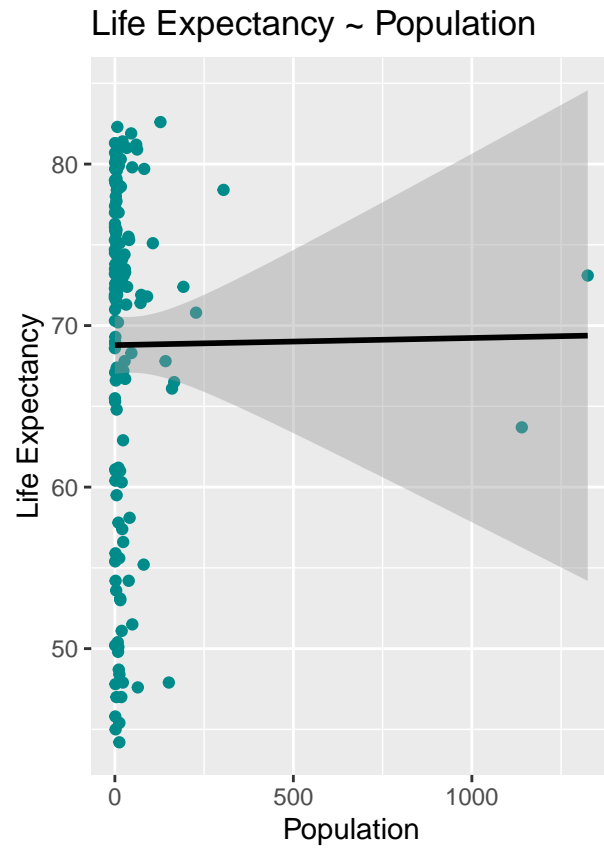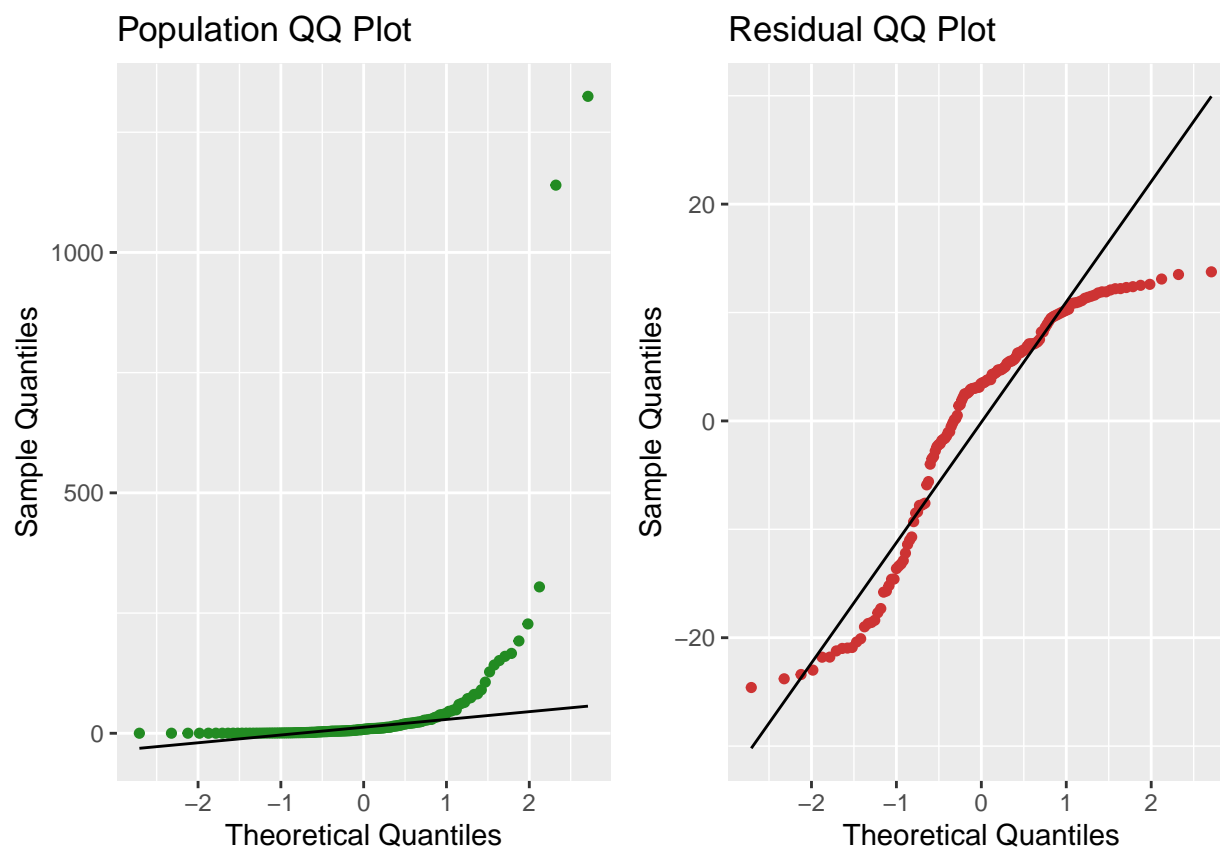
Life Expectancy ~ Rural

Table 10: Life Expectancy ~ Rural

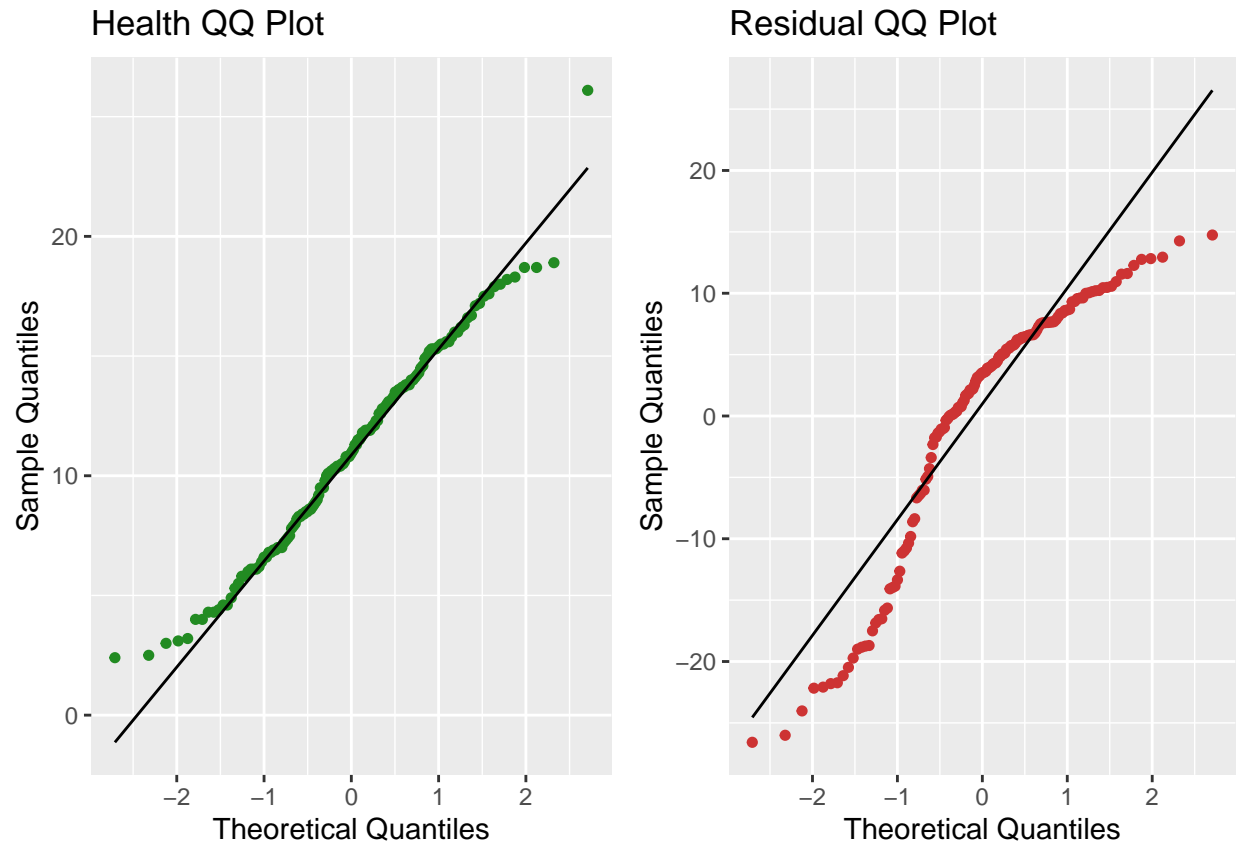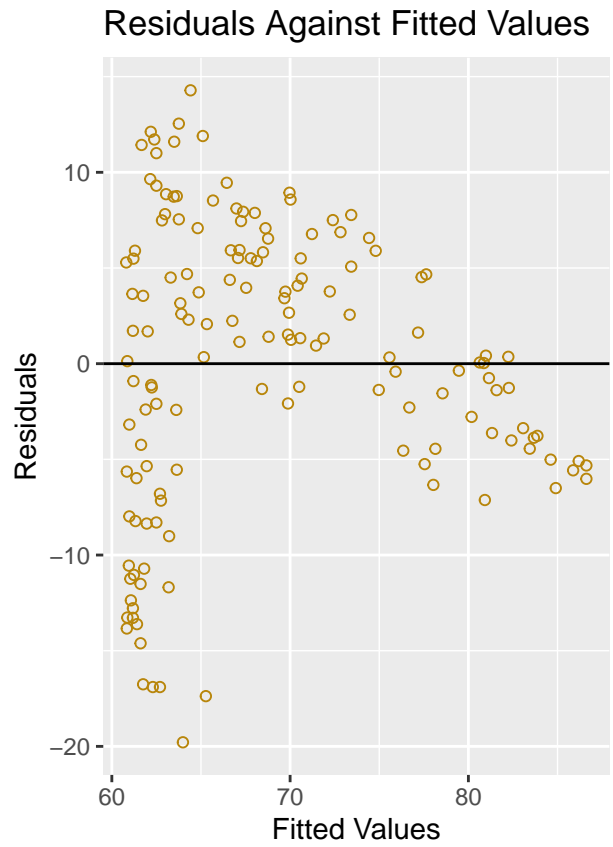|	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	81.2538126	1.4694641	55.294859	0
Rural	-0.2878945	0.0300757	-9.572321	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```
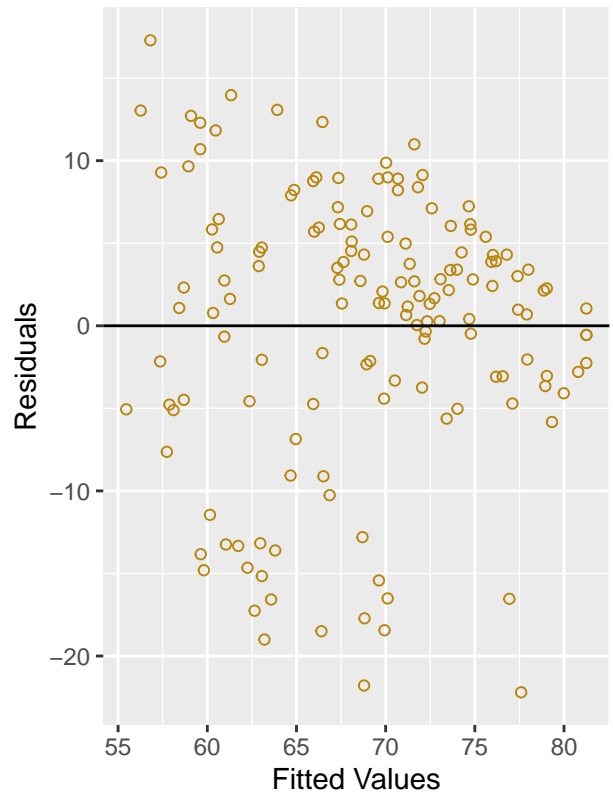
Transformations

New Models

Life Expectancy ~ log(CO2)

Table 11: Life Expectancy ~ log(CO2)

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	65.423520	0.6397774	102.25981	0
log(CO2)	4.622899	0.3473506	13.30903	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

log(CO2) QQ Plot

Residual QQ Plot
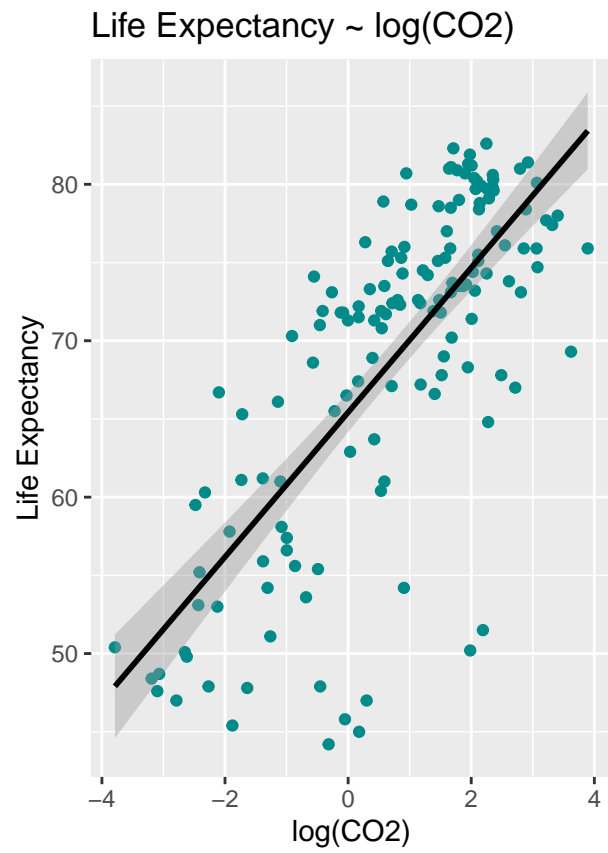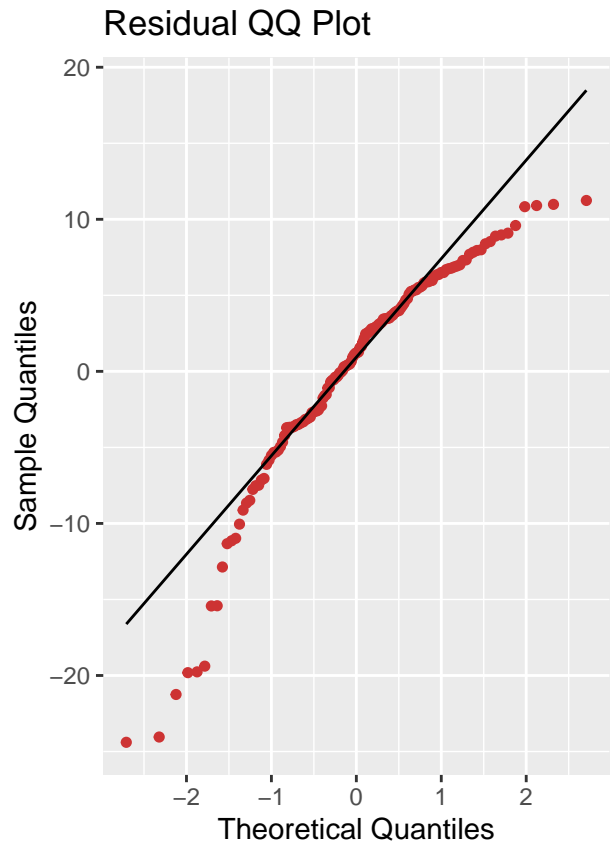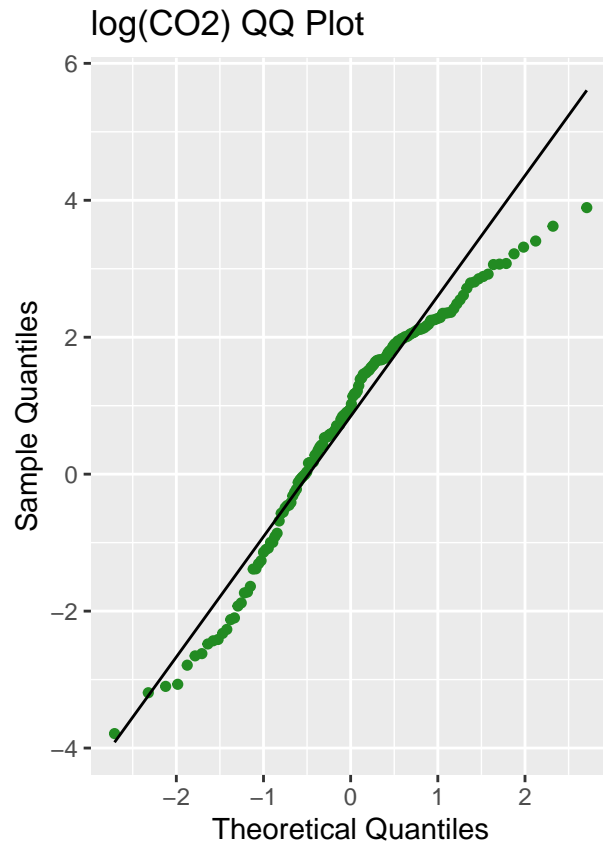
Life Expectancy ~ log(Elderly Population)

Table 12: Life Expectancy ~ log(Elderly Population)

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	50.37800	1.8762682	26.85011	0
log(Elderly Population)	10.29956	0.9816733	10.49184	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

Life Expectancy ~ log(Elderly Popu

Residuals Against Fitted Values

log(Elderly Population) QQ Plot

Residual QQ Plot
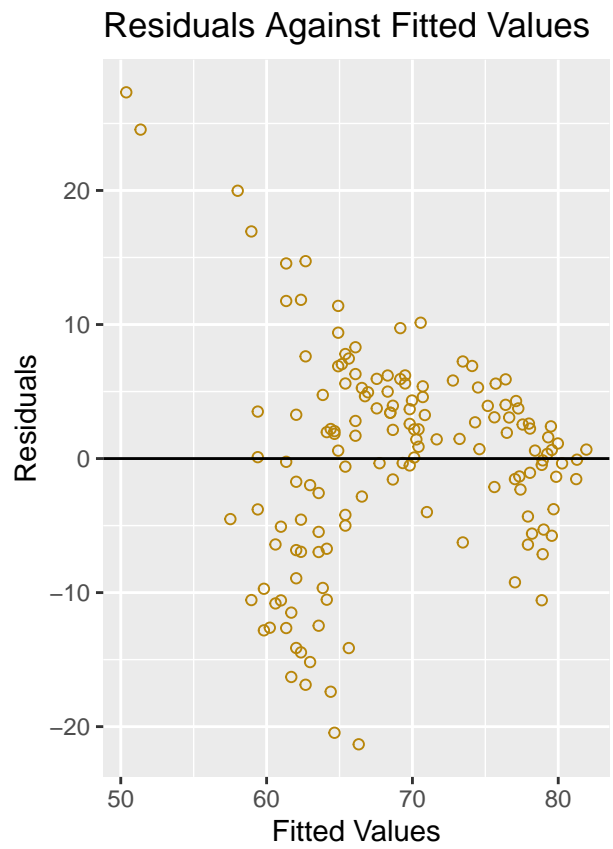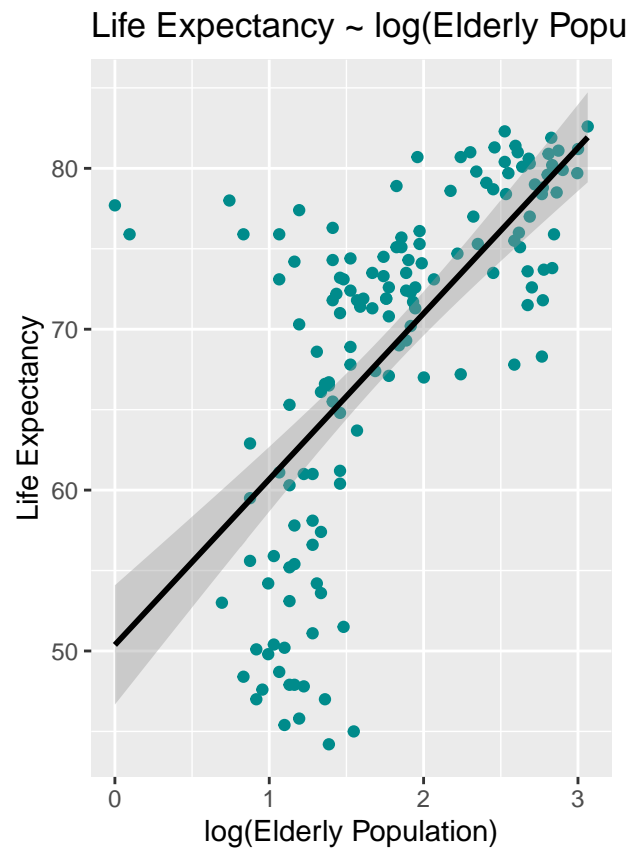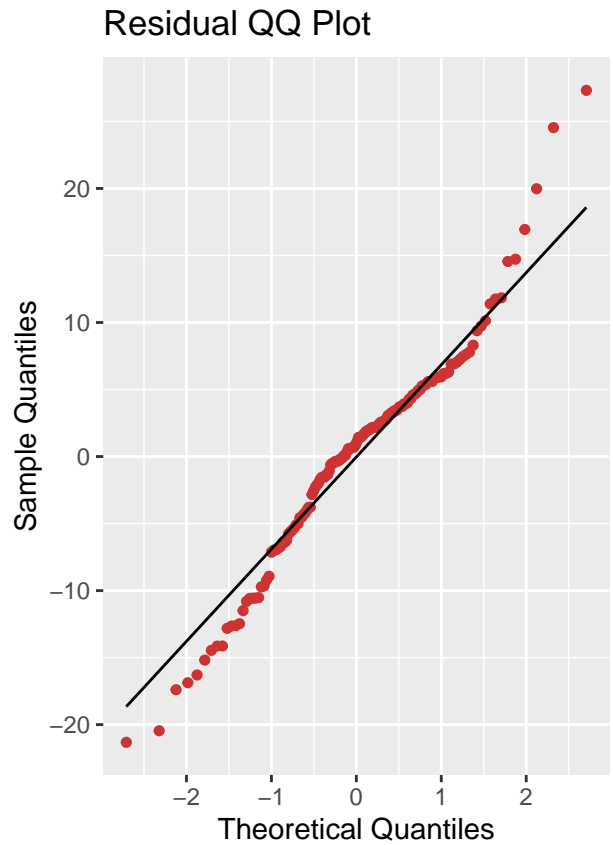
Life Expectancy ~ log(GDP)

Table 13: Life Expectancy ~ log(GDP)

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	22.885285	3.0930132	7.399026	0
log(GDP)	5.374861	0.3563291	15.083980	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```
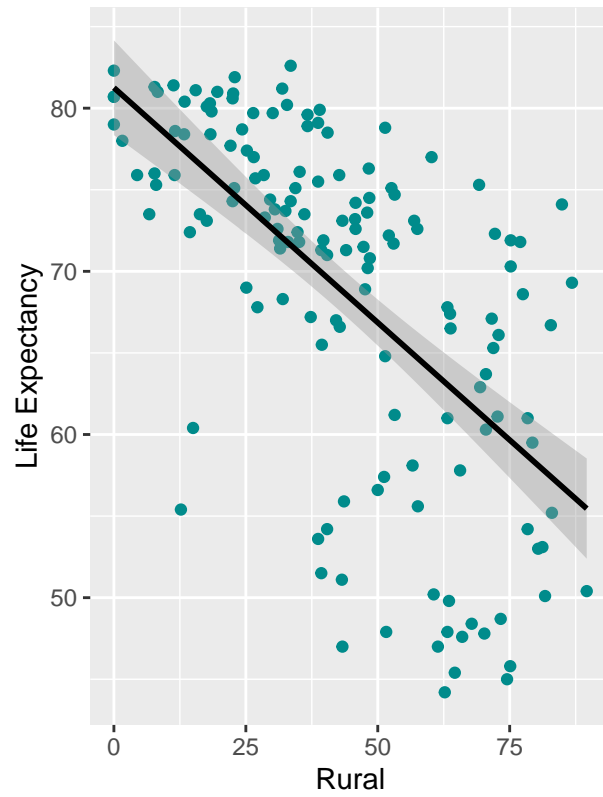
Life Expectancy ~ log(GDP)

Residuals Against Fitted Values

log(GDP) QQ Plot   Residual QQ Plot
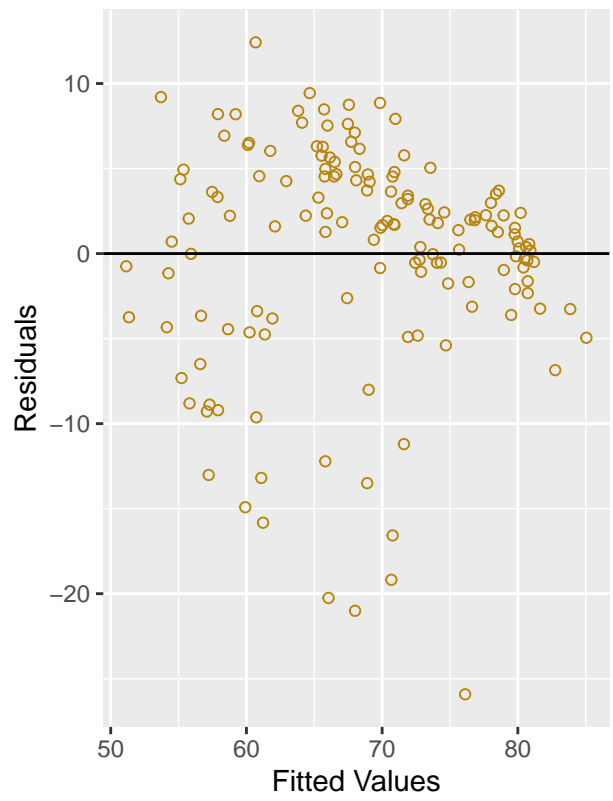
Life Expectancy ~ log(Internet)

Table 14: Life Expectancy ~ log(Internet)

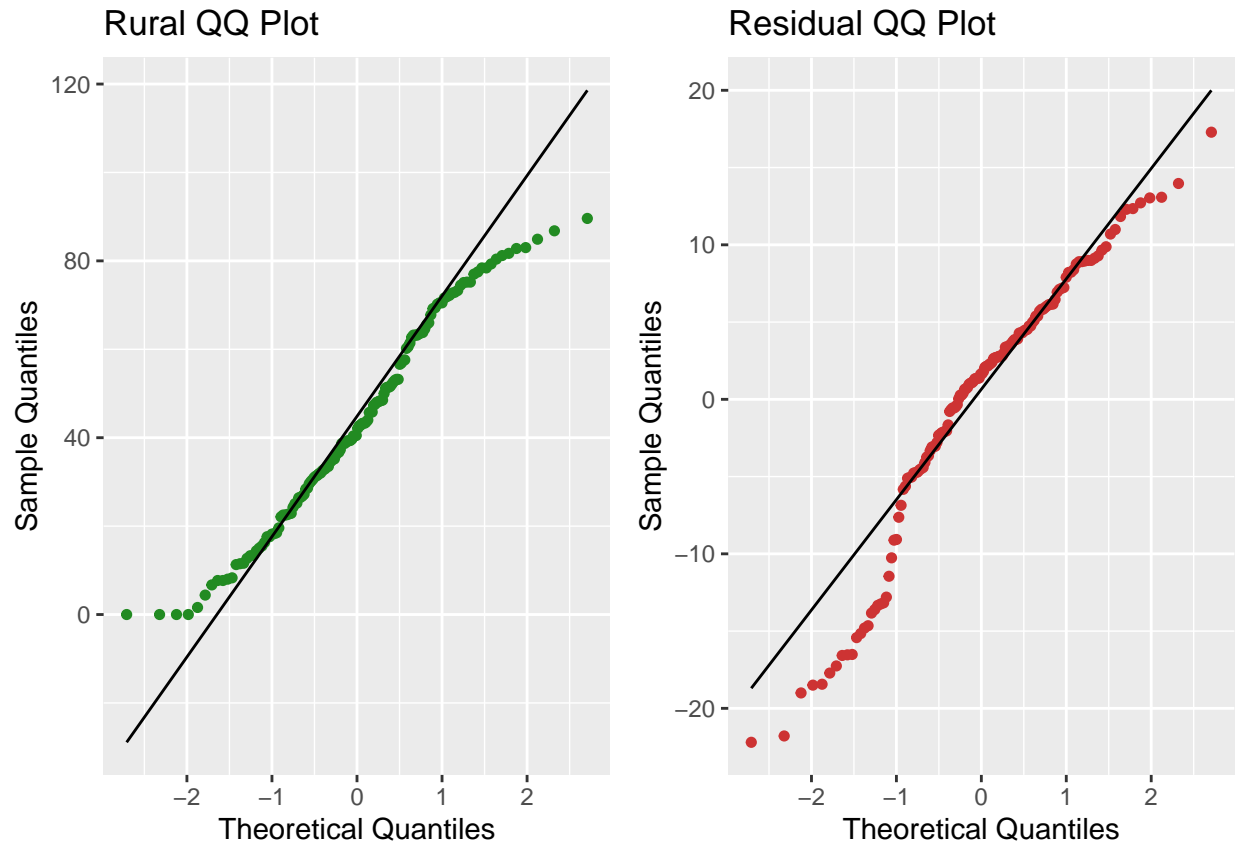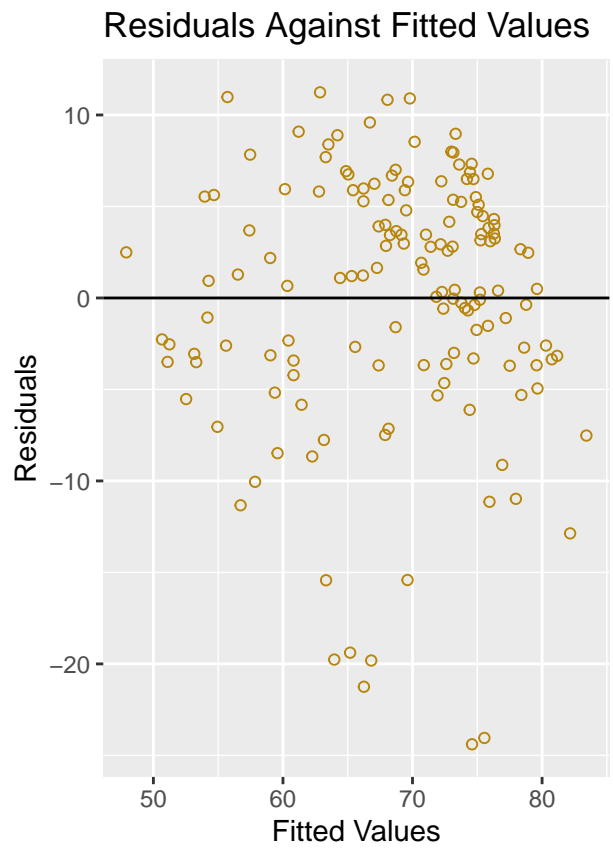|	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	53.175799	1.1599814	45.84194	0
log(Internet)	5.882667	0.3858553	15.24579	0

```
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## 'geom_smooth()' using formula = 'y ~ x'
```

log(Internet) QQ Plot

Residual QQ Plot
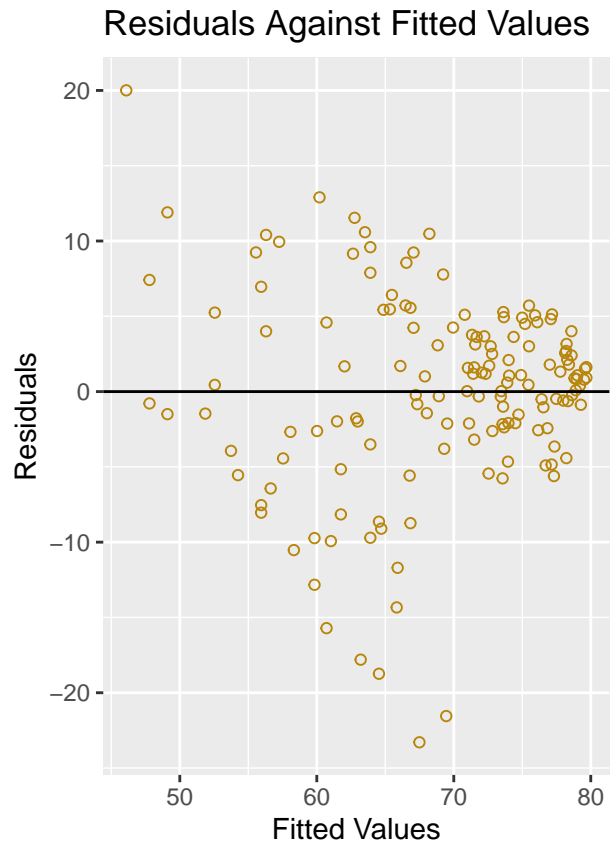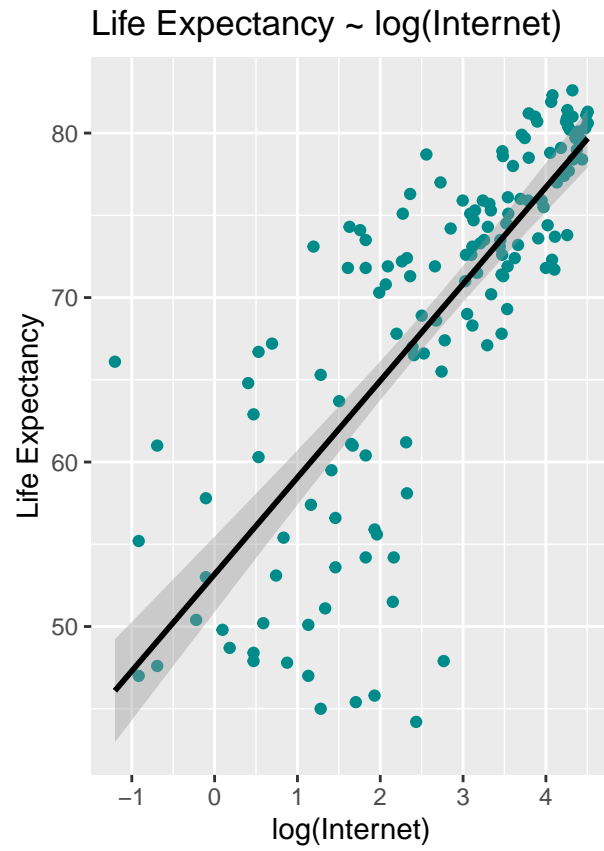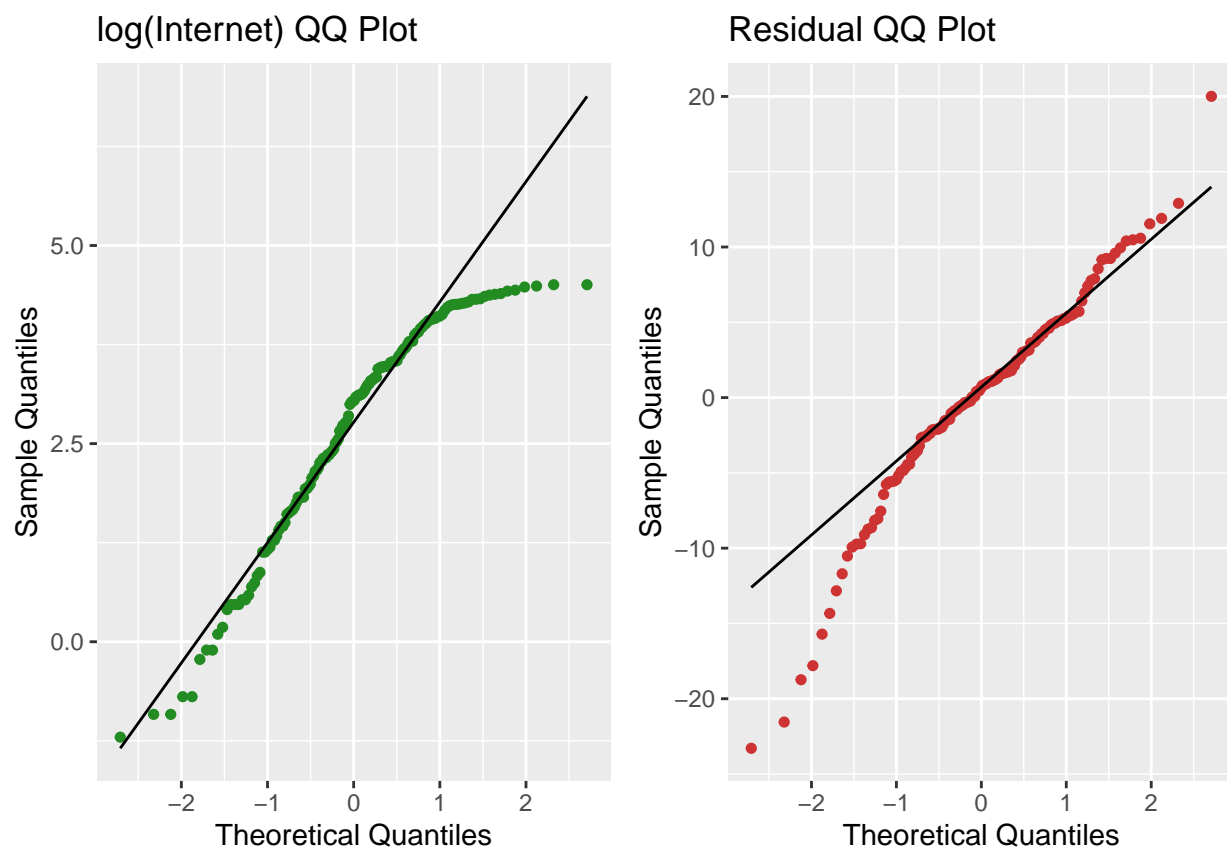
State All Rsq

Table 15: R Squared Values

XVar	Rsq	Adj.Rsq	Trans.Rsq
Birth Rate	0.7417423	0.7399734	NA
Cell	0.4306327	0.4267329	NA
CO2	0.1893512	0.1837988	0.5481698
Elderly Population	0.3978775	0.3937534	0.4298627
GDP	0.3367658	0.3322231	0.6091308
Health	0.0873995	0.0811488	NA
Internet	0.5127019	0.5093643	0.6141996
Land Area	0.0016413	-0.0051968	NA
Population	0.0000382	-0.0068109	NA
Rural	0.3855977	0.3813895	NA

Multifactor Models

Table 16: Forward Selection Predictiors

	Include
(Intercept)	TRUE
population	FALSE
rural	TRUE

	Include
health	TRUE
internet	TRUE
birth_rate	TRUE
elderly_pop	TRUE
co2	FALSE
gdp	FALSE
cell	FALSE

Table 17: Forward Selection Algorithm | nbest=5

	population	rural	health	internet	birth_rate	elderly_pop	co2	gdp	cell
1 (1)					*				
2 (1)				*	*				
3 (1)		*		*	*				
4 (1)		*	*	*	*				
5 (1)		*	*	*	*	*			
6 (1)		*	*	*	*	*			*
7 (1)		*	*	*	*	*	*		*
8 (1)		*	*	*	*	*	*	*	*

Table 18: Backward Elimination Predictiors

	Include
(Intercept)	TRUE
population	FALSE
rural	TRUE
health	TRUE
internet	TRUE
birth_rate	TRUE
elderly_pop	TRUE
co2	FALSE
gdp	FALSE
cell	FALSE

Table 19: Backward Elimination Algorithm | nbest=5

	population	rural	health	internet	birth_rate	elderly_pop	co2	gdp	cell
1 (1)					*				
2 (1)				*	*				
3 (1)		*		*	*				
4 (1)		*	*	*	*				
5 (1)		*	*	*	*	*			
6 (1)		*	*	*	*	*		*	
7 (1)		*	*	*	*	*	*	*	
8 (1)		*	*	*	*	*	*	*	*

Assess Multicollinearity

```
## Warning: 'select_()' was deprecated in dplyr 0.7.0.
## i Please use 'select()' instead.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

Table 20: VIF Values

	LandArea	Population	Rural	Health	Internet	BirthRate	ElderlyPop	CO2	GDP	Cell
LandArea	Inf	1.260	1.022	1.002	1.009	1.010	1.016	1.023	1.011	1.002
Population	1.260	Inf	1.006	1.010	1.000	1.004	1.001	1.001	1.002	1.009
Rural	1.022	1.006	Inf	1.023	1.775	1.550	1.231	1.791	2.306	1.675
Health	1.002	1.010	1.023	Inf	1.084	1.062	1.161	1.008	1.093	1.009
Internet	1.009	1.000	1.775	1.084	Inf	2.672	1.750	2.659	3.283	1.999
BirthRate	1.010	1.004	1.550	1.062	2.672	Inf	2.549	2.973	2.799	1.823
ElderlyPop	1.016	1.001	1.231	1.161	1.750	2.549	Inf	1.401	1.627	1.270
CO2	1.023	1.001	1.791	1.008	2.659	2.973	1.401	Inf	4.304	2.087
GDP	1.011	1.002	2.306	1.093	3.283	2.799	1.627	4.304	Inf	2.175
Cell	1.002	1.009	1.675	1.009	1.999	1.823	1.270	2.087	2.175	Inf

Best Model

Table 21: Best Model Summary

	Estimate	Std. Error	t value	Pr(>|t|)	RSq
(Intercept)	82.2440751	3.8475946	21.375452	0.0000000	0.7837944
rural	-0.0460924	0.0245643	-1.876394	0.0626535	NA
health	0.2307352	0.1063916	2.168734	0.0317680	NA
internet	1.5883390	0.5275125	3.010998	0.0030829	NA
birth_rate	-0.7054330	0.0816440	-8.640350	0.0000000	NA
elderly_pop	-1.5119081	1.0341320	-1.462007	0.1459489	NA

Best Model Residual Plot