STA108_TermProjectEdits

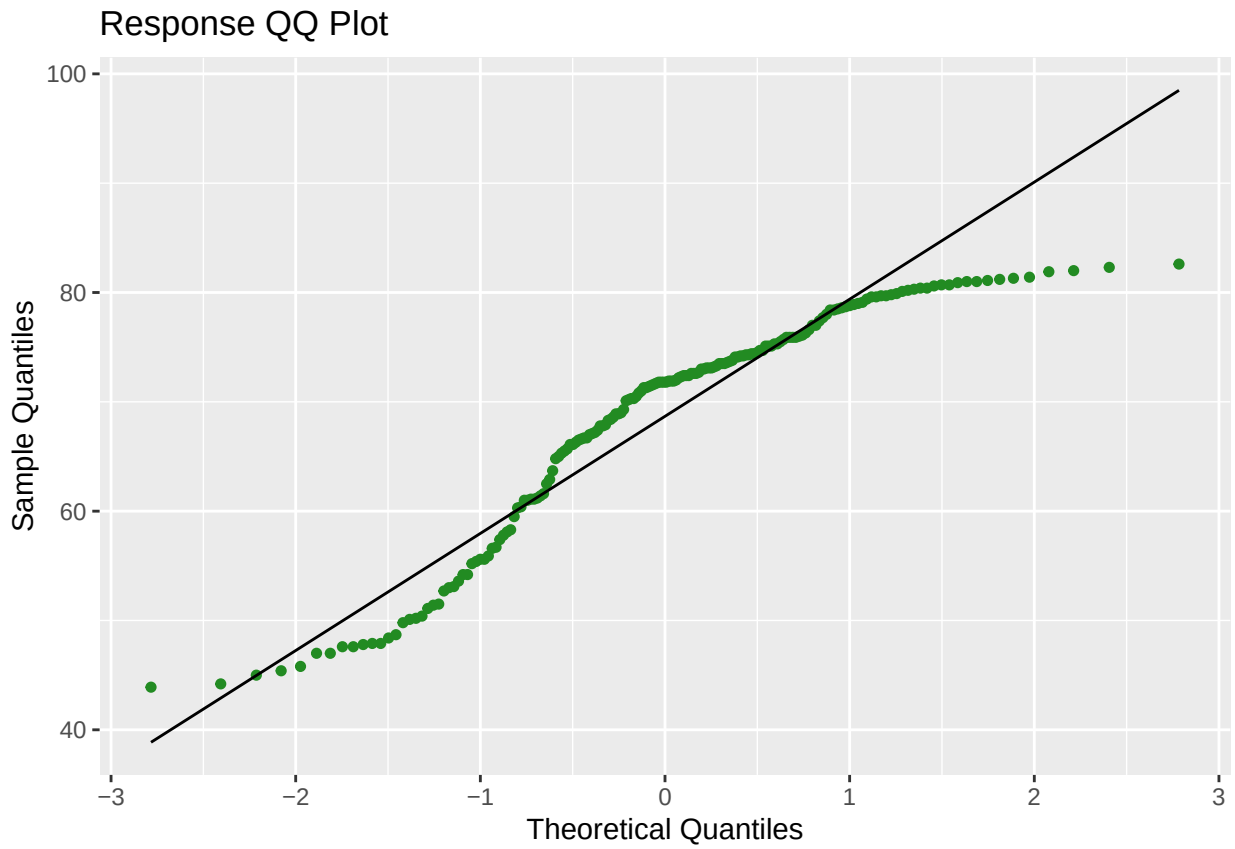
Gabriel Jones

2023-11-23

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select

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

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##     recode
```

Single Linear Regression Models & Analysis

Life Expectancy Analysis

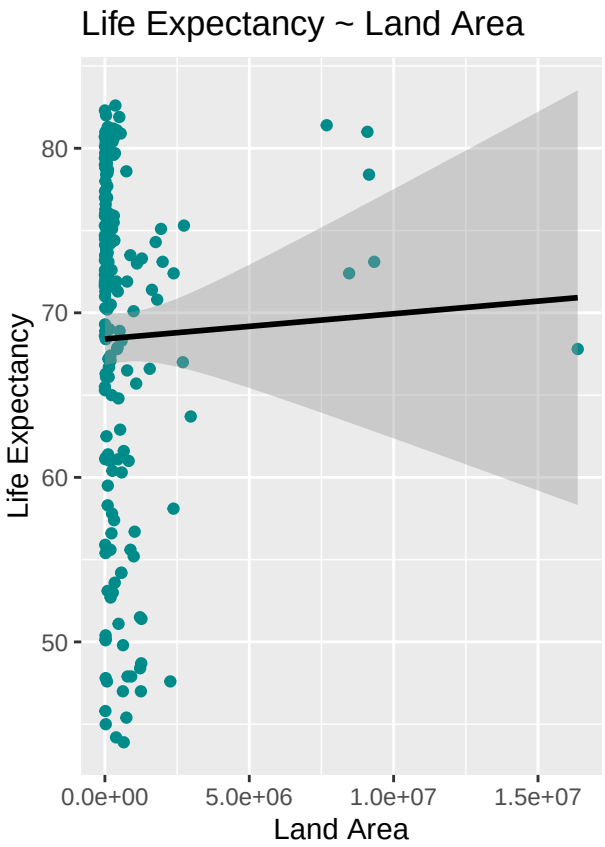


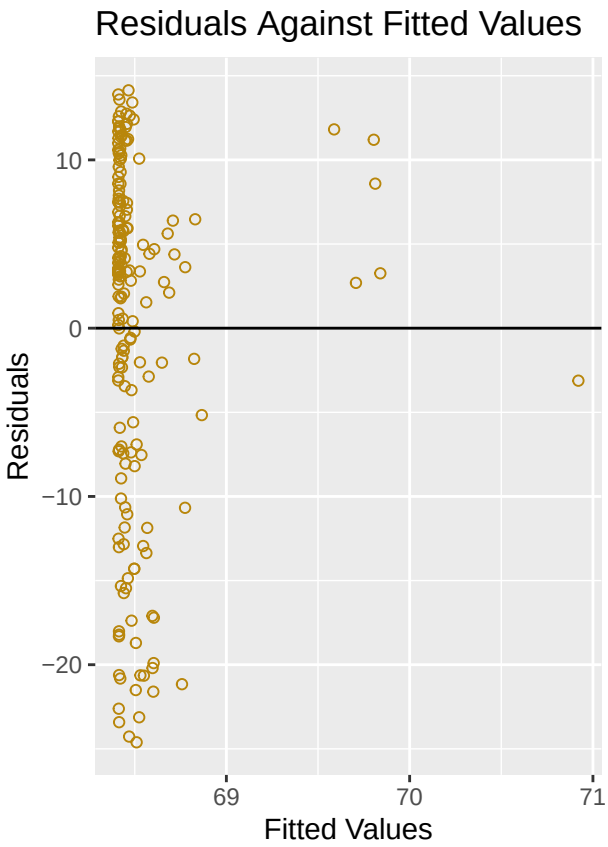
Life Expectancy ~ Land Area

Table 1: Life Expectancy ~ Land Area

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	68.4102866	0.8068269	84.789303	0.0000000
Land Area	0.0000002	0.0000004	0.379057	0.7050826

```
## 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 ~ Land Area
Life Expectancy
80
70
60
50
0.0e+00
5.0e+06
1.0e+07
1.5e+07
Land Area

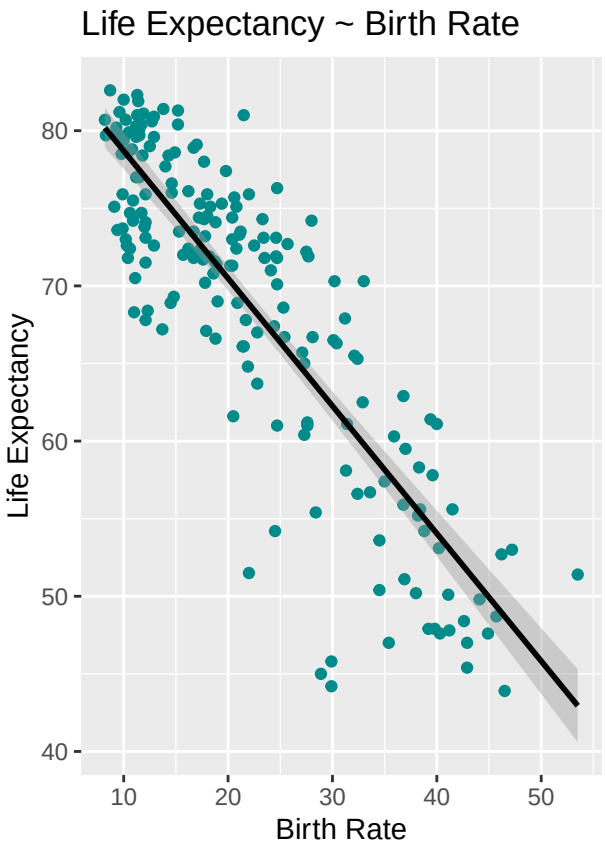

Residuals Against Fitted Values
Residuals
10
0
−10
−20
69
70
71
Fitted Values

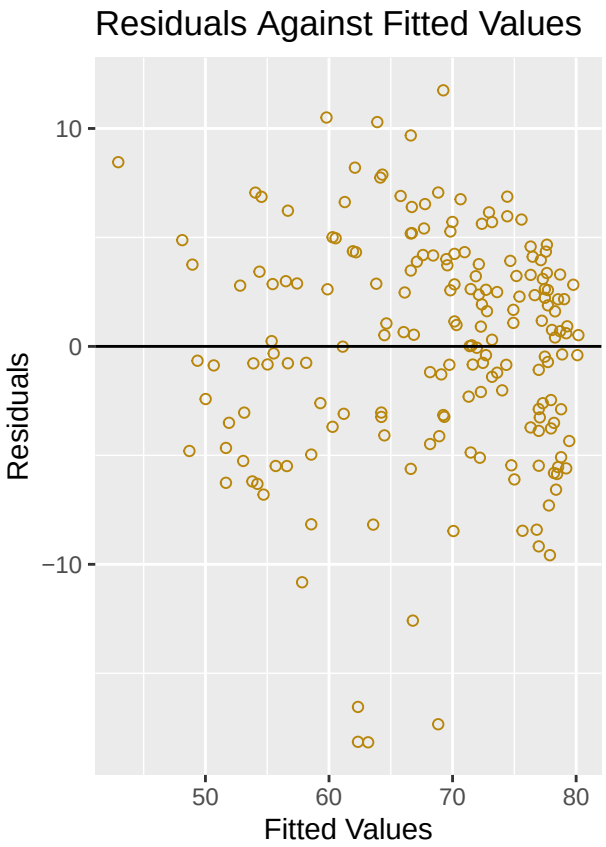
Life Expectancy ~ Birth Rate

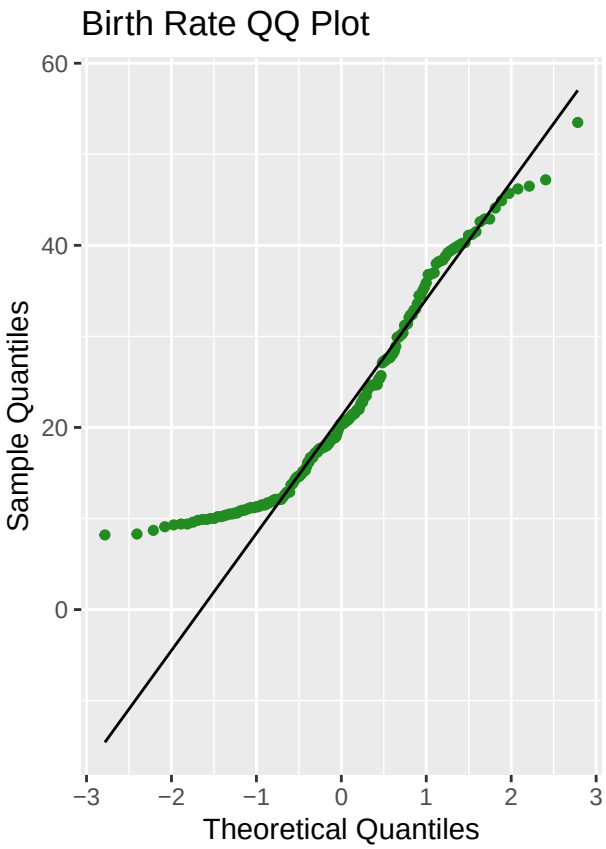
Table 2: Life Expectancy ~ Birth Rate

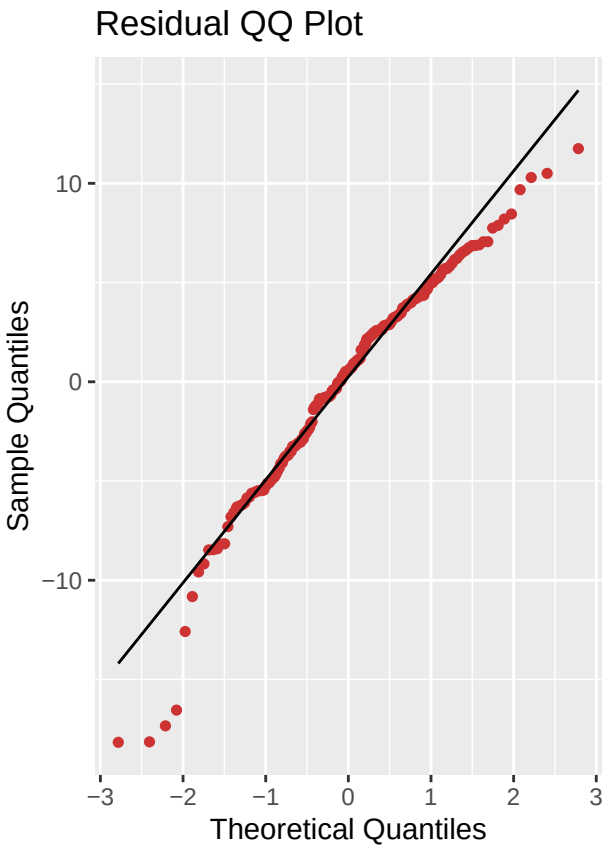
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	86.9231381	0.9012399	96.44839	0
Birth Rate	-0.8219937	0.0363068	-22.64021	0

```
## 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 ~ Birth Rate
Life Expectancy
Birth Rate
80
70
60
50
40
10
20
30
40
50


Residuals Against Fitted Values
Residuals
Fitted Values
10
0
−10
50
60
70
80



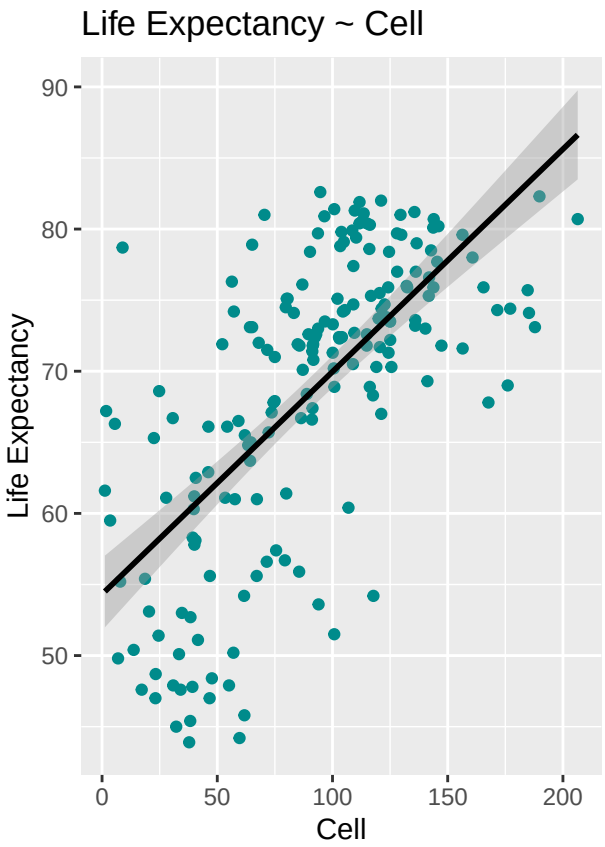


Life Expectancy ~ Cell

Table 3: Life Expectancy ~ Cell

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	54.3145708	1.2926162	42.01910	0
Cell	0.1565293	0.0128266	12.20351	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
Life Expectancy
Cell
90
80
70
60
50
0
50
100
150
200

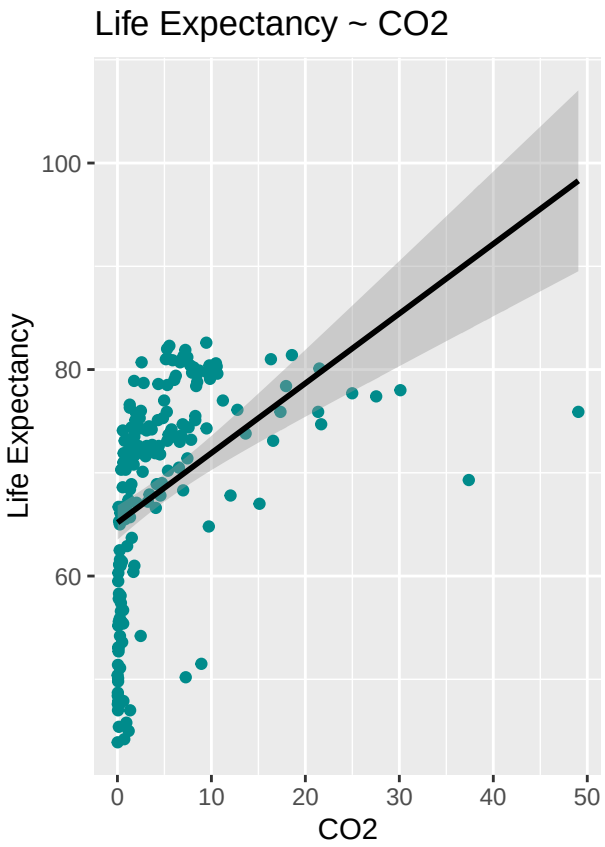

Residuals Against Fitted Values
Residuals
Fitted Values
20
10
0
−10
−20
60
70
80

Life Expectancy ~ CO2

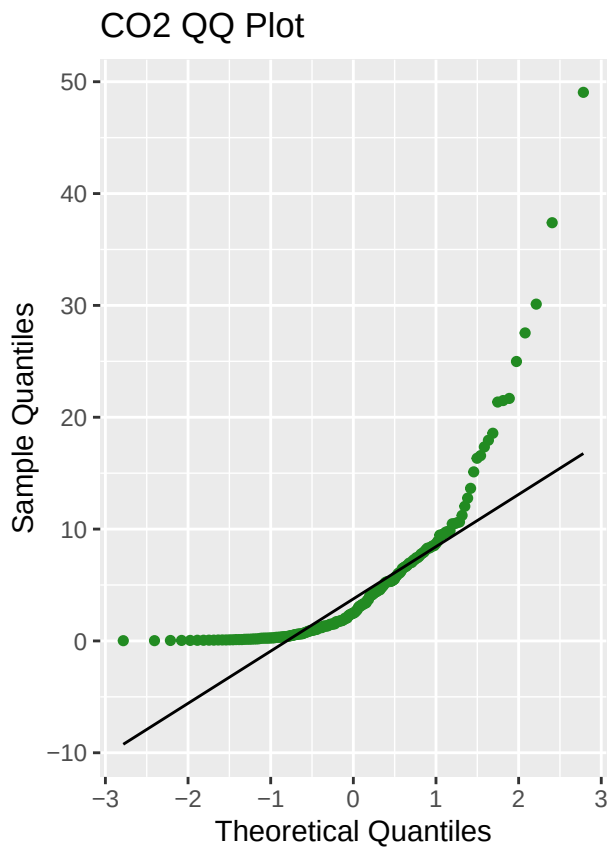
Table 4: Life Expectancy ~ CO2

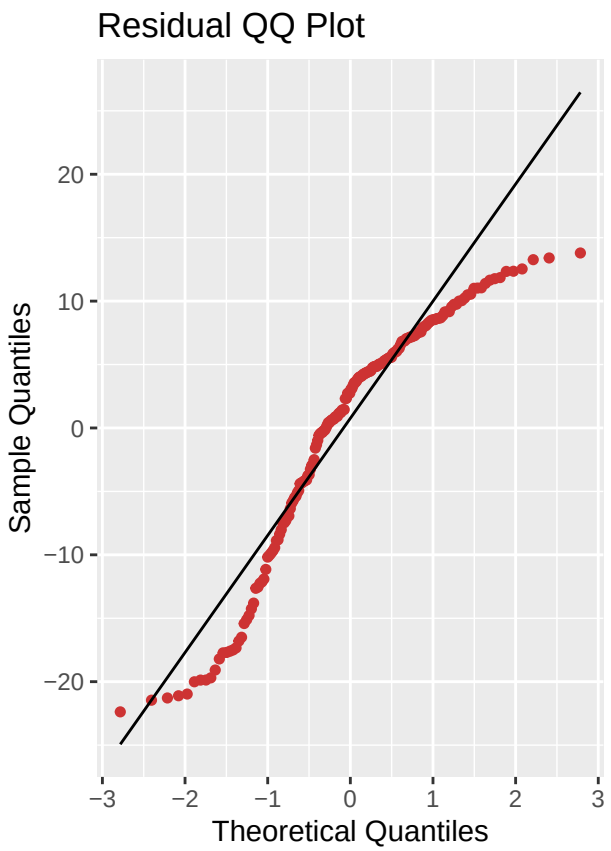
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	65.1673619	0.8378840	77.776114	0
CO2	0.6750591	0.0995408	6.781732	0

```
## 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 ~ CO2
Life Expectancy
100
80
60
0
10
20
30
40
50
CO2


Residuals Against Fitted Values
Residuals
10
0
−10
−20
70
80
90
Fitted Values



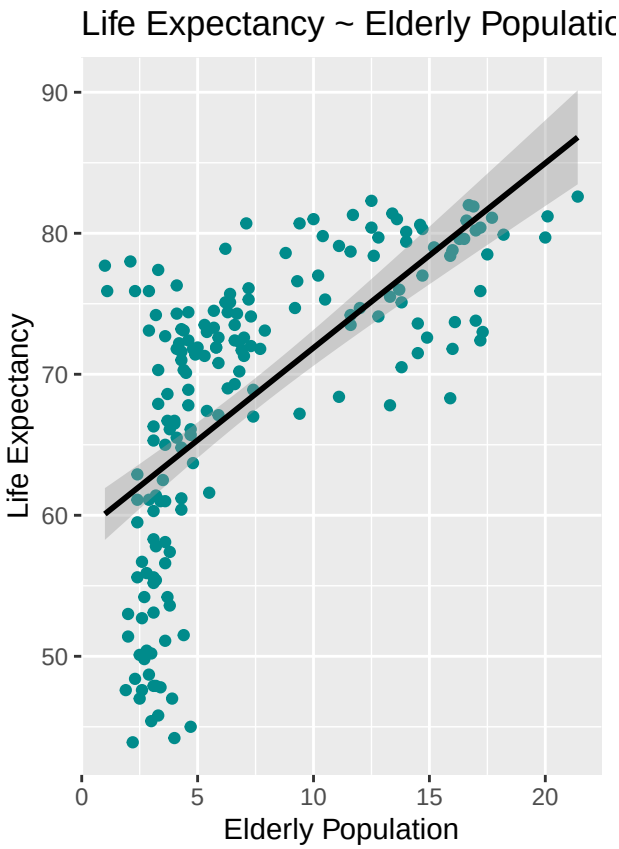


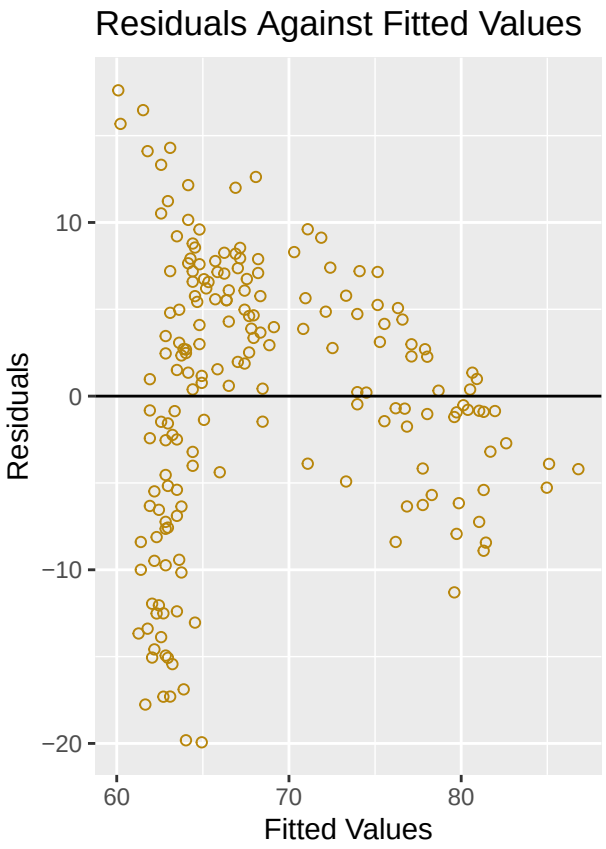
Life Expectancy ~ Elderly Population

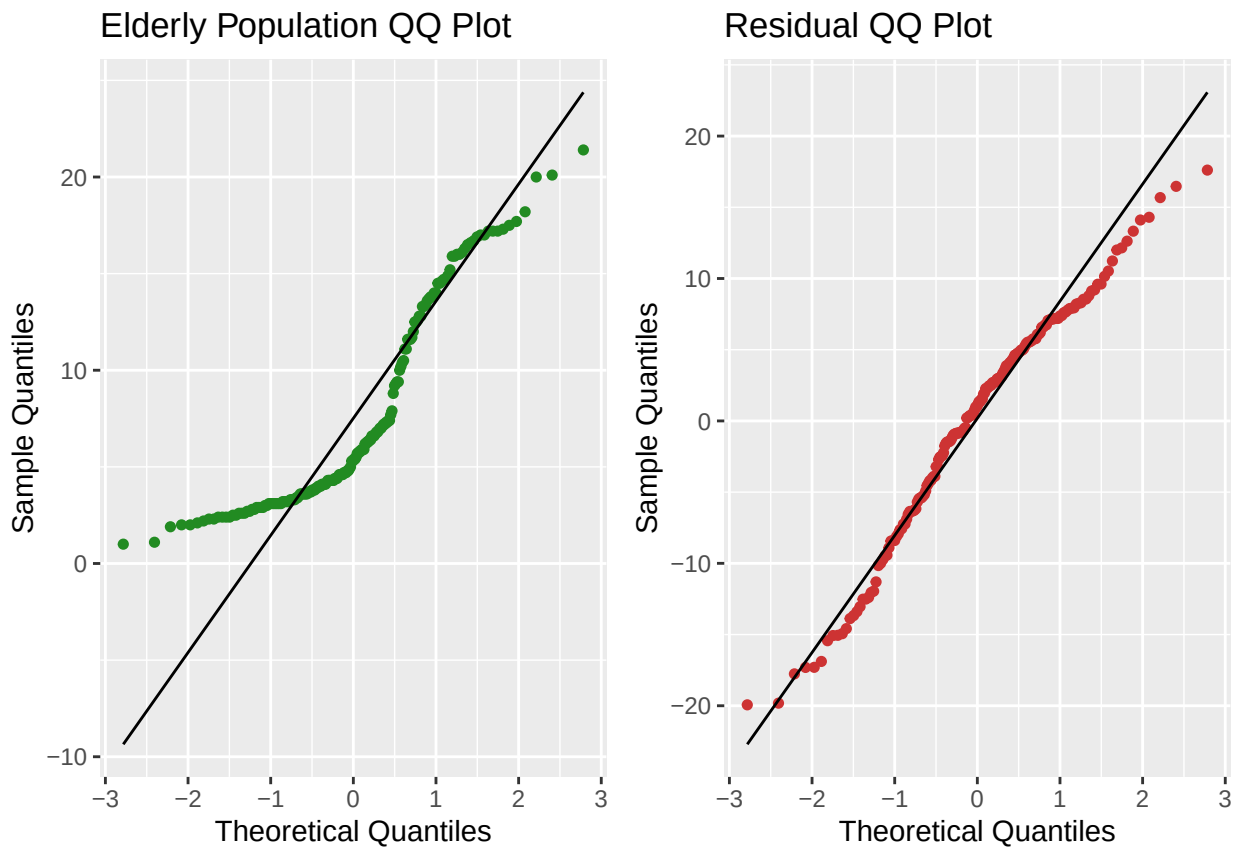
Table 5: Life Expectancy ~ Elderly Population

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	58.780154	1.0230021	57.45849	0
Elderly Population	1.309561	0.1136439	11.52337	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 ~ Elderly Populatic
Life Expectancy
Elderly Population
90
80
70
60
50
0
5
10
15
20


Residuals Against Fitted Values
Residuals
Fitted Values
10
0
−10
−20
60
70
80

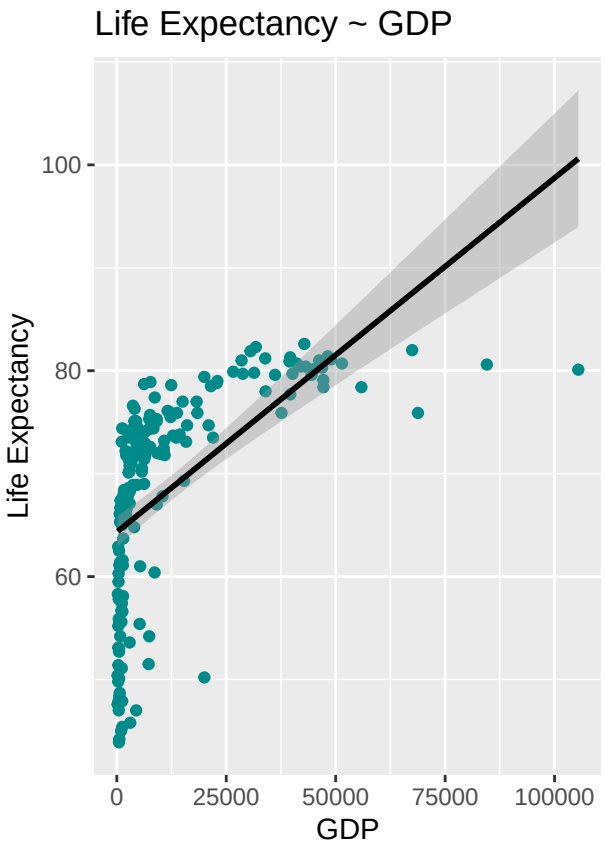


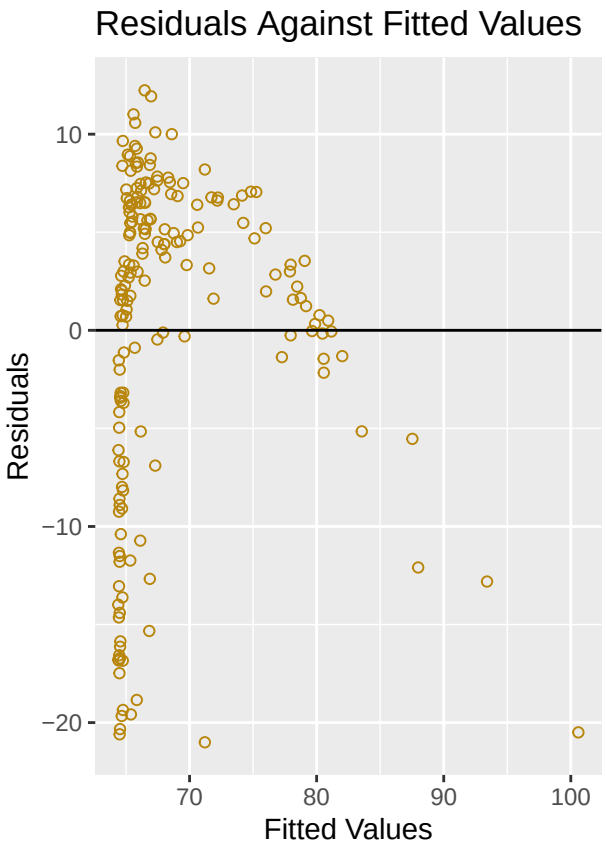
Life Expectancy ~ GDP

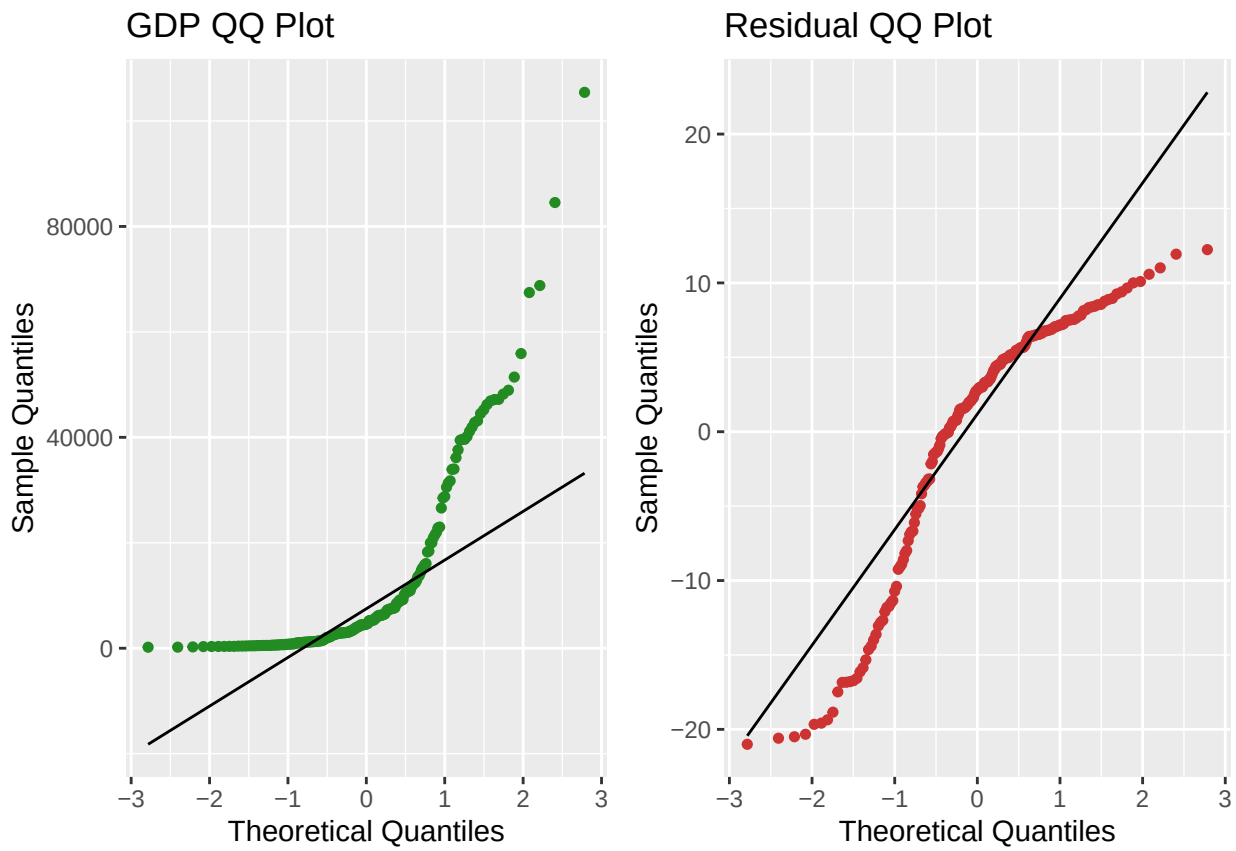
Table 6: Life Expectancy ~ GDP

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	64.323624	0.7525609	85.472984	0
GDP	0.000344	0.0000355	9.690967	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
Life Expectancy
100
80
60
0
25000
50000
75000
100000
GDP


Residuals Against Fitted Values
Residuals
10
0
−10
−20
70
80
90
100
Fitted Values

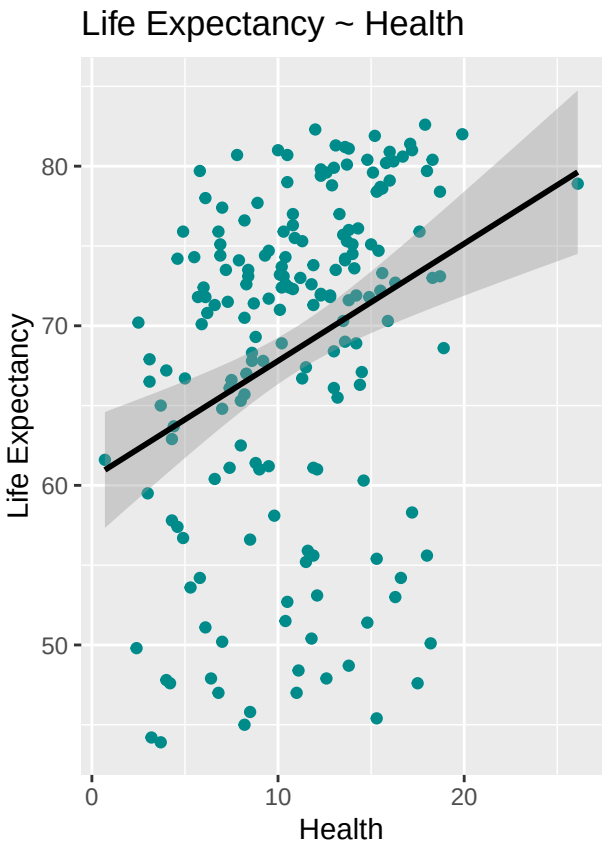


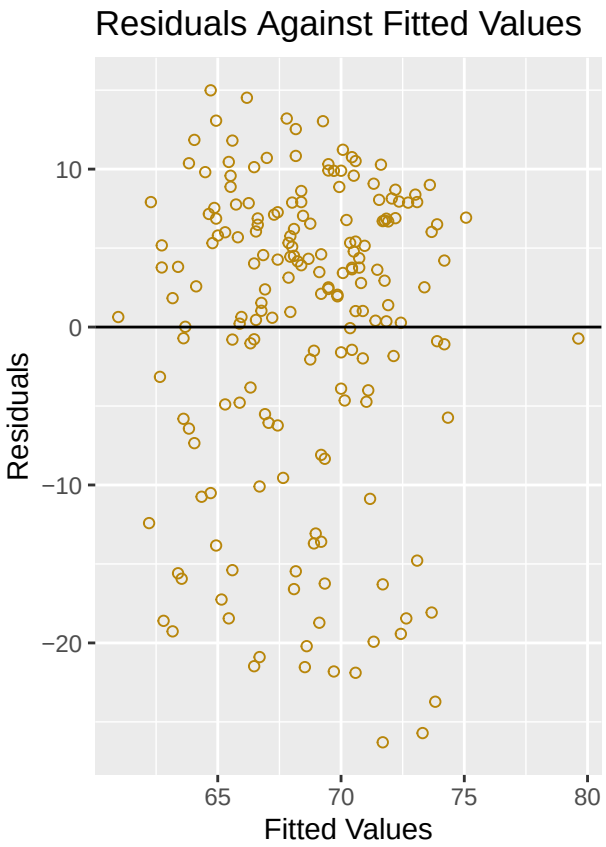
Life Expectancy ~ Health

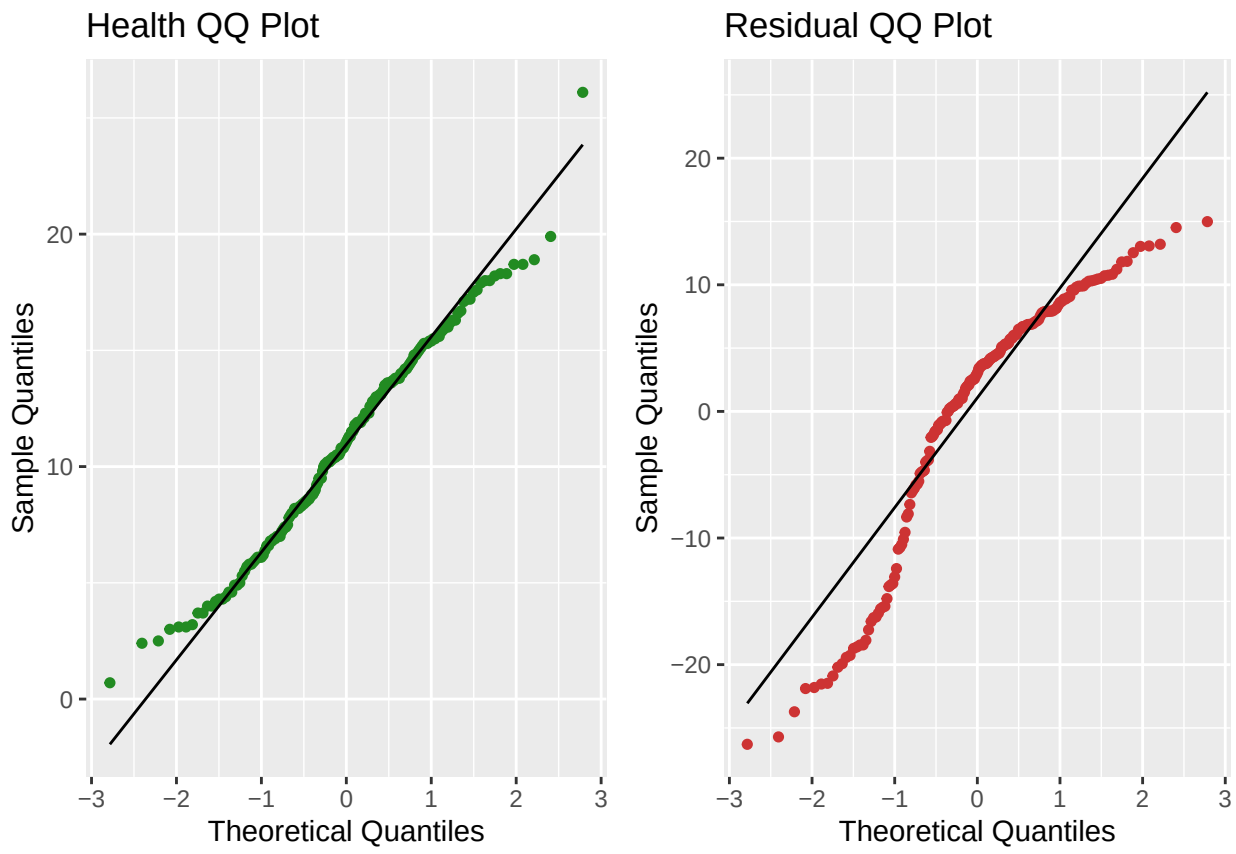
Table 7: Life Expectancy ~ Health

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	60.4502780	1.9492731	31.011703	0.00e+00
Health	0.7347805	0.1650418	4.452086	1.47e-05

```
## 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
Life Expectancy
80
70
60
50
0
10
20
Health


Residuals Against Fitted Values
Residuals
10
0
−10
−20
65
70
75
80
Fitted Values

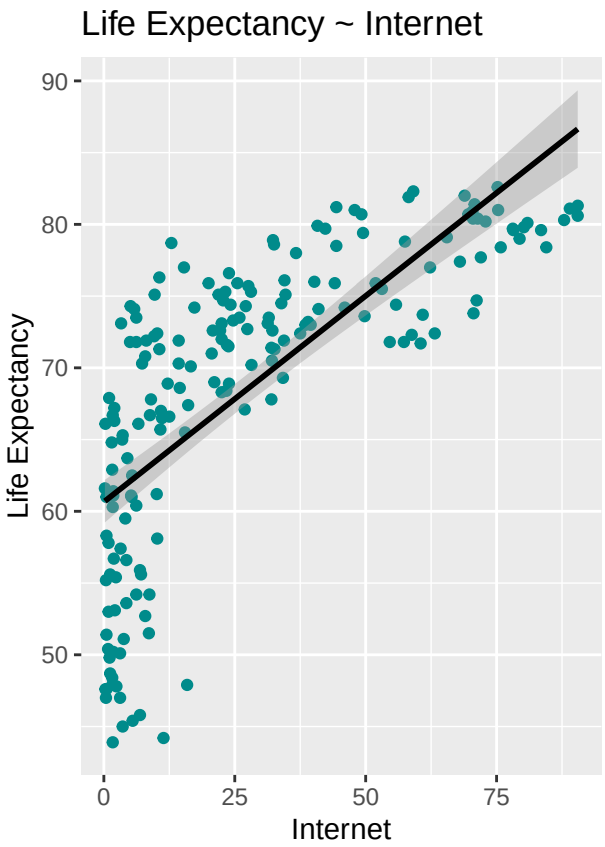


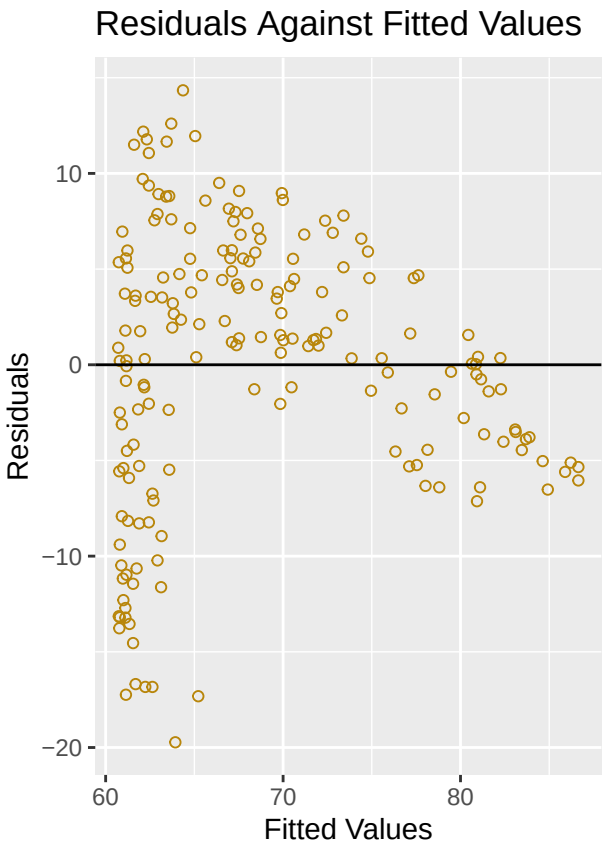
Life Expectancy ~ Internet

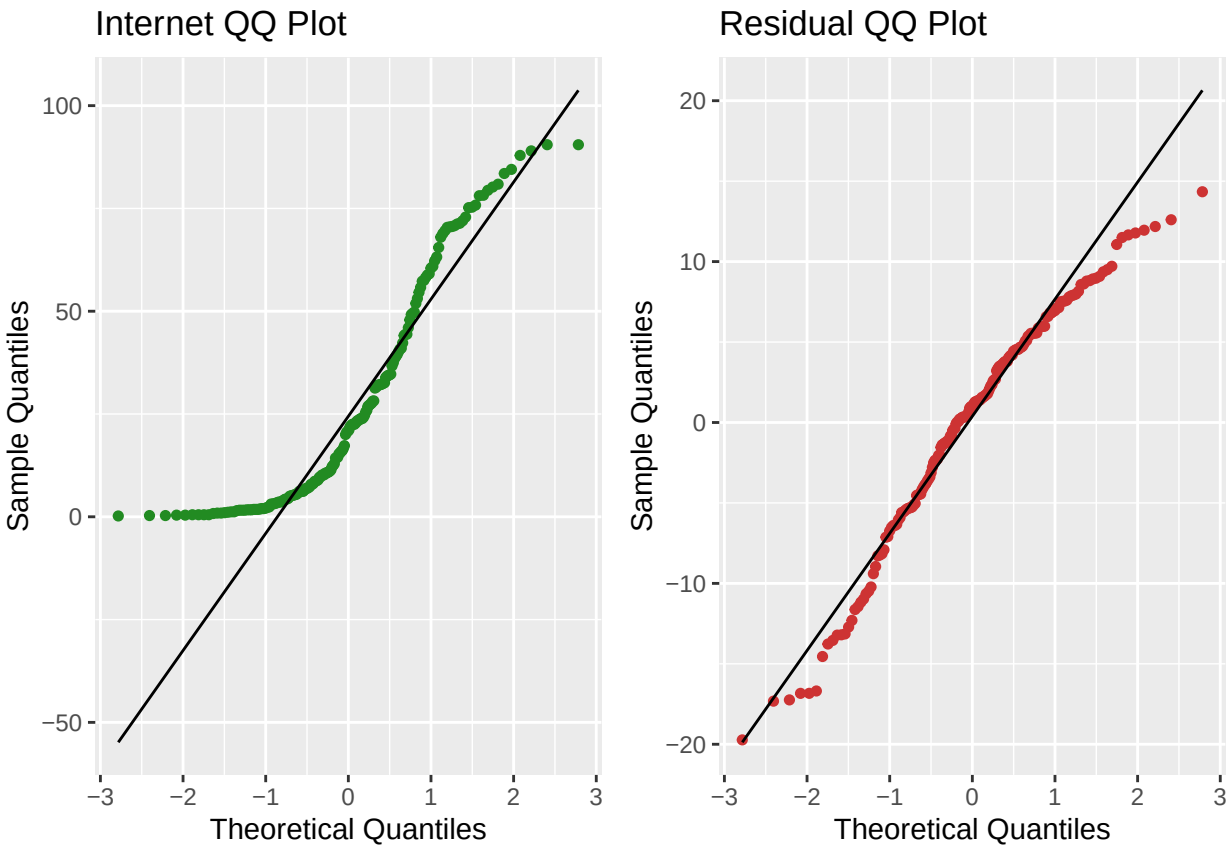
Table 8: Life Expectancy ~ Internet

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	60.6542290	0.7560181	80.22854	0
Internet	0.2871813	0.0200382	14.33166	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
Life Expectancy
Internet
90
80
70
60
50
0
25
50
75


Residuals Against Fitted Values
Residuals
Fitted Values
10
0
−10
−20
60
70
80

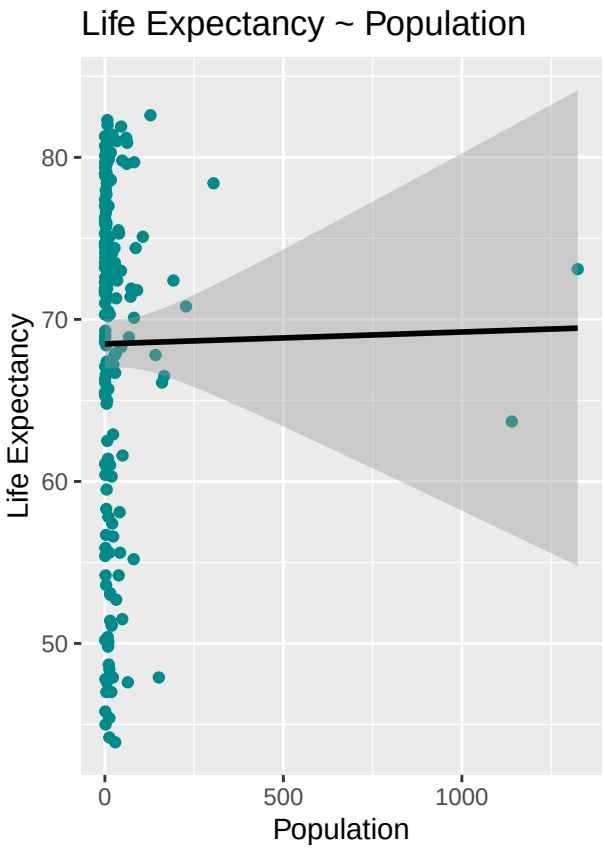


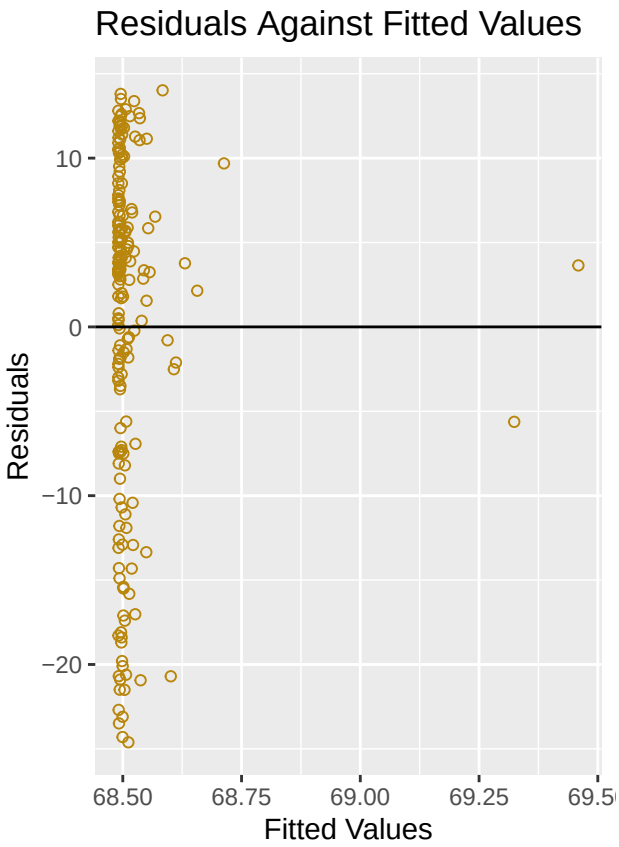
Life Expectancy ~ Population

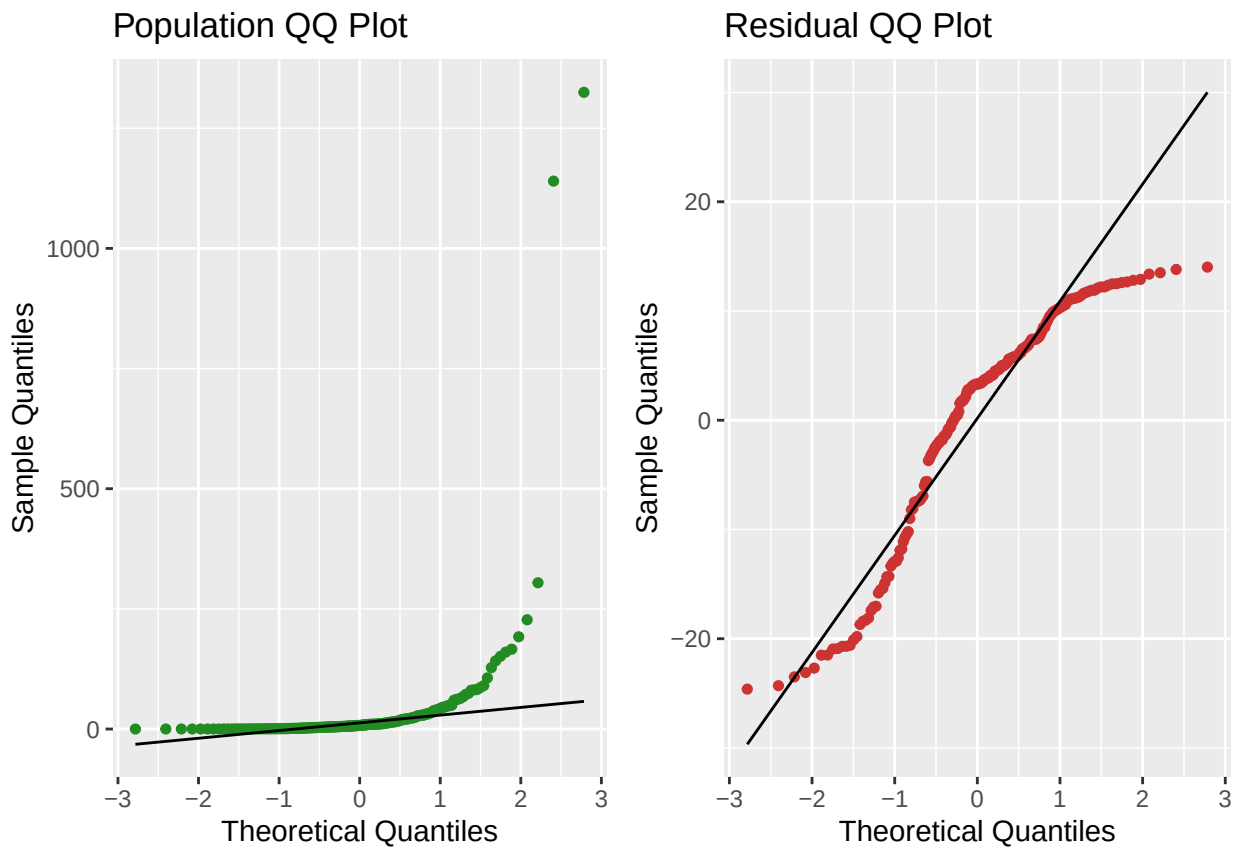
Table 9: Life Expectancy ~ Population

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	68.4904710	0.7841700	87.3413582	0.0000000
Population	0.0007312	0.0057423	0.1273414	0.8988093

```
## 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
80
70
60
50
Life Expectancy
0
500
1000
Population


Residuals Against Fitted Values
10
0
−10
−20
Residuals
68.50
68.75
69.00
69.25
Fitted Values

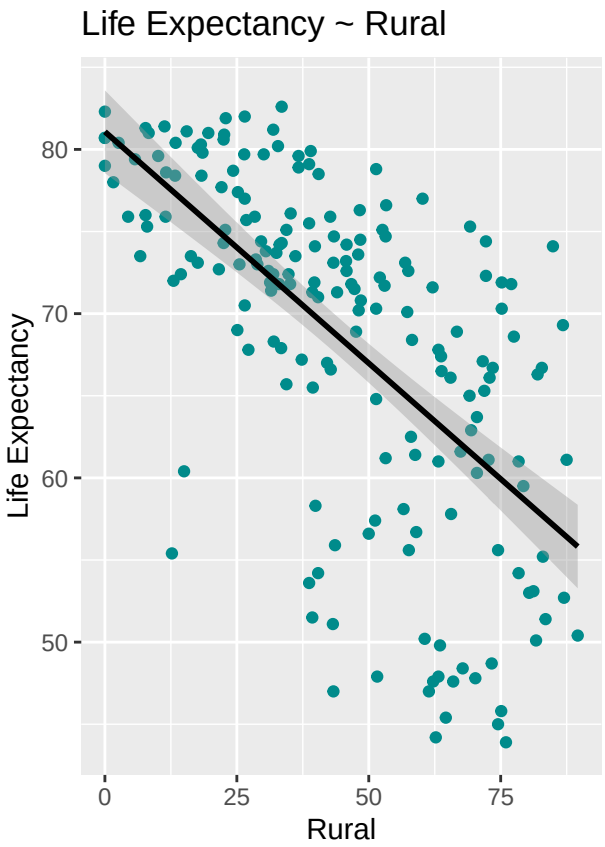


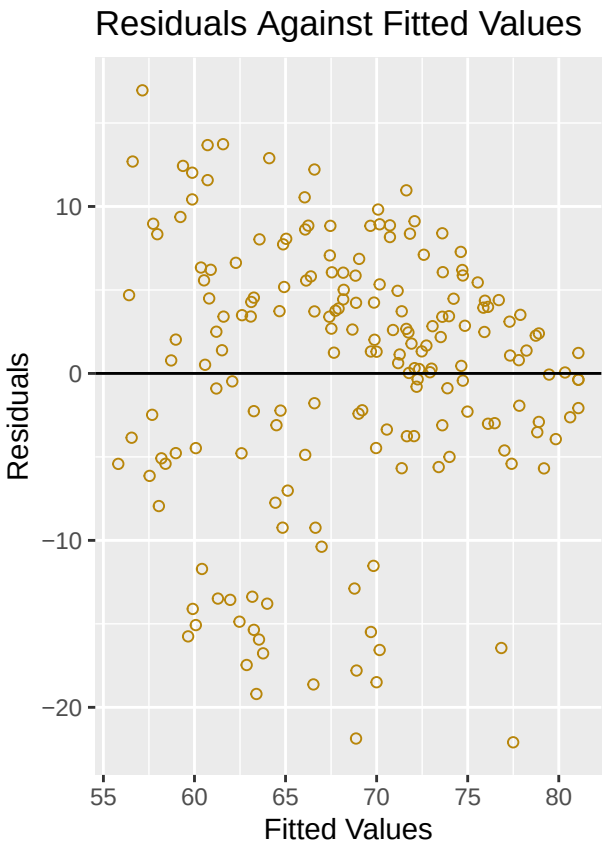
Life Expectancy ~ Rural

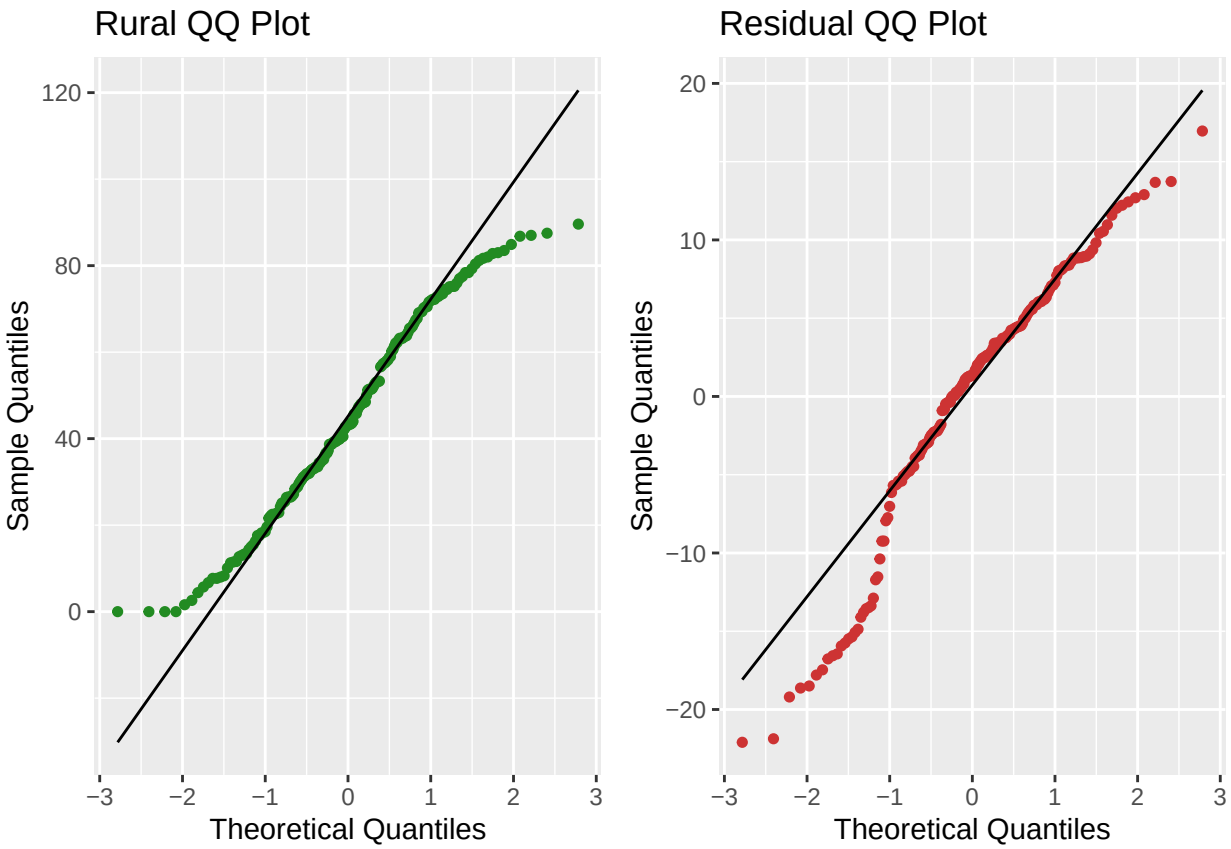
Table 10: Life Expectancy ~ Rural

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	81.0758804	1.2775025	63.46436	0
Rural	-0.2819026	0.0254756	-11.06561	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 ~ Rural
Life Expectancy
80
70
60
50
0
25
50
75
Rural


Residuals Against Fitted Values
Residuals
10
0
−10
−20
55
60
65
70
75
80
Fitted Values



Transformations

New Models

State All Rsq

Table 11: R Squared Values

XVar	Rsq	Adj.Rsq	Trans.Rsq
Birth Rate	0.7358520	0.7344164	NA
Cell	0.4473241	0.4443204	NA
CO2	0.1999718	0.1956238	0.5840630
Elderly Population	0.4191698	0.4160131	0.4694471
GDP	0.3379267	0.3343284	0.6272979
Health	0.0972474	0.0923412	NA
Internet	0.5274739	0.5249058	0.6156395
Land Area	0.0007803	-0.0046503	NA
Population	0.0000881	-0.0053462	NA
Rural	0.3995712	0.3963080	NA

Multifactor Models

Table 12: Forward Selection Predictiors

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

Table 13: Forward Selection Algorithm | nbest=8

	land_area	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 14: Backward Elimination Predictiors

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

Table 15: Backward Elimination Algorithm | nbest=8

	land_area	population	rural	health	internet	birth_rate	elderly_pop	co2	gdp	cell
1 (1)						*				
2 (1)					*	*				
3 (1)			*		*	*				

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

Best Model

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 16: VIF Values For Two Factor Models

	LandArea	Population	Rural	Health	Internet	BirthRate	ElderlyPop	CO2	GDP	Cell
LandArea	Inf	1.265	1.017	1.001	1.005	1.005	1.009	1.018	1.009	1.001
Population	1.265	Inf	1.004	1.008	1.000	1.004	1.001	1.001	1.001	1.006
Rural	1.017	1.004	Inf	1.038	1.724	1.576	1.314	1.890	2.334	1.568
Health	1.001	1.008	1.038	Inf	1.109	1.057	1.150	1.017	1.106	1.019
Internet	1.005	1.000	1.724	1.109	Inf	2.682	1.900	2.662	3.127	2.087
BirthRate	1.005	1.004	1.576	1.057	2.682	Inf	2.894	3.103	2.797	1.812
ElderlyPop	1.009	1.001	1.314	1.150	1.900	2.894	Inf	1.524	1.762	1.339
CO2	1.018	1.001	1.890	1.017	2.662	3.103	1.524	Inf	4.368	2.064
GDP	1.009	1.001	2.334	1.106	3.127	2.797	1.762	4.368	Inf	2.014
Cell	1.001	1.006	1.568	1.019	2.087	1.812	1.339	2.064	2.014	Inf