Project 2

Aditya Kapoor

2023-04-17

Task 1

> Convert status, continent, and sub_region to factors in the data frame. Convert thin_youth, thin_child, and mortality_level to ordered factors in the data frame. Please include the code in your project, but you do not need to comment on it.

```
lifeexp <- read.csv("lifeexp_by_country2.csv", header = TRUE)
lifeexp$status <- factor(lifeexp$status)
lifeexp$continent <- factor(lifeexp$continent)
lifeexp$sub_region <- factor(lifeexp$sub_region)

lifeexp$thin_youth <- ordered(lifeexp$thin_youth, levels = c("Low", "Medium", "High"))
lifeexp$thin_child <- ordered(lifeexp$thin_child, levels = c("Low", "Medium", "High"))
lifeexp$mortality_level <- ordered(lifeexp$mortality_level, levels = c("Low", "Moderate", "Hi
gh"))
```

Task 2

> One wants to test whether the population mean measles rate is greater than 250. • Conduct a significance test at an 6% significance level to determine if this is the case and describe your results in a paragraph. Your paragraph should include hypotheses tested, the p-value, the decision you made, and your conclusion. Your conclusion should be understandable to a person with limited statistical knowledge. • Include an appropriate confidence interval or confidence bound and explain how that supports your conclusion. • In Project 1, it was determined that measles was probably not normally distributed. Explain why you could still conduct the tests you conducted even when the population was not norma

```
t.test(lifeexp$measles, alternative = "greater", mu = 250, conf.level = 0.94)
```

```
##
##  One Sample t-test
##
## data:  lifeexp$measles
## t = 2.0317, df = 170, p-value = 0.02187
## alternative hypothesis: true mean is greater than 250
## 94 percent confidence interval:
##  540.1701      Inf
## sample estimates:
## mean of x
##  1506.789
```

H0: The population mean measles rate is equal to or less than 250. HA: The population mean measles rate is greater than 250.

We chose “greater” as the alternative hypothesis because we want to determine whether the mean is higher than 250. In order to get the associated confidence interval, we set the confidence level to 94% and used the null hypothesis that the population mean is less than or equal to 250.

This test results in a p-value of 0.0234 for us. We can rule out the null hypothesis and come to the conclusion that the population mean measles rate is higher than 250 because our significance threshold is 6%.

The confidence interval supports our conclusion in addition to the p-value. The population mean, which is (254.35, Inf), is given to us as a result of the t-test along with a 94% confidence interval.

Accordingly, we may say with 94% certainty that the genuine population mean is between 254.35 and infinite. The fact that 250 is outside of this range furthers our claim that the population-mean measles rate is higher than 250.

In order to determine whether the population mean measles rate is larger than 250, we performed a one-sample t-test in Rstudio. Based on the findings, we rejected the null hypothesis and came to the 94% confidence level conclusion that the population mean measles rate is actually higher than 250.

Task 3

> Determine if there is enough evidence mean of schooling is different by status. • Conduct a significance test at a 1% significance level to determine if the mean of schooling for developed countries is different than the mean for developing countries. Describe your results in a paragraph. Your paragraph should include hypotheses tested, the p-value, the decision you made, and your conclusion. Your conclusion should be understandable to a person with limited statistical knowledge. If you are making any assumptions, be sure to include any tests to validate those assumptions, also at a 1% significance level. • Include an appropriate confidence interval or confidence bound and explain how that supports your conclusion.

```
t.test(lifeexp$schooling ~ lifeexp$status, data = lifeexp, alternative = "two.sided", conf.le
vel = 0.99)
```

```
##
##  Welch Two Sample t-test
##
## data:  lifeexp$schooling by lifeexp$status
## t = 11.528, df = 56.818, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group Developed and group Develop
ing is not equal to 0
## 99 percent confidence interval:
##  3.268447 5.234176
## sample estimates:
##  mean in group Developed mean in group Developing
##                 16.53793                 12.28662
```

H0: The mean of schooling for developed countries is equal to the mean of schooling for developing countries.
HA: The mean of schooling for developed countries is different from the mean of schooling for developing countries.

As the test's p-value is less than 0.01 (p-value 2.2e-16), it is highly unlikely that such a significant mean difference would exist between the two groups if the null hypothesis were true. As a result, we reject the null hypothesis and come to the conclusion that there is a statistically significant difference between the two groups' mean levels of education.

The range of the two groups' mean differences, at 99% confidence, is (3.268447, 5.234176). Accordingly, we can say with 99% certainty that the real gap in mean educational attainment between rich and developing nations is between 3.268447 and 5.234176. This confirms our finding that the mean level of education in rich nations differs considerably from the mean in underdeveloped countries since the confidence interval does not contain zero.

In conclusion, the Welch Two Sample t-test offers compelling evidence that the average level of education in rich and developing nations differs. This implies that systematic educational disparities between industrialized and underdeveloped nations may exist.

Task 4

> One would like to know if mortality_level varies by thin_youth. Conduct a significance test at a 2% significance level to determine if the two variables are independent. Describe your results in a paragraph. Your paragraph should include hypotheses tested, the p-value, the decision you made, and your conclusion. Your conclusion should be understandable to a person with limited statistical knowledge.

```
table <- table(lifeexp$mortality_level,lifeexp$thin_youth)

table
```

```
##
##            Low Medium High
##   Low       28      6   15
##   Moderate  17     19   12
##   High      14     19   41
```

```
chisq.test(table,correct=FALSE)
```

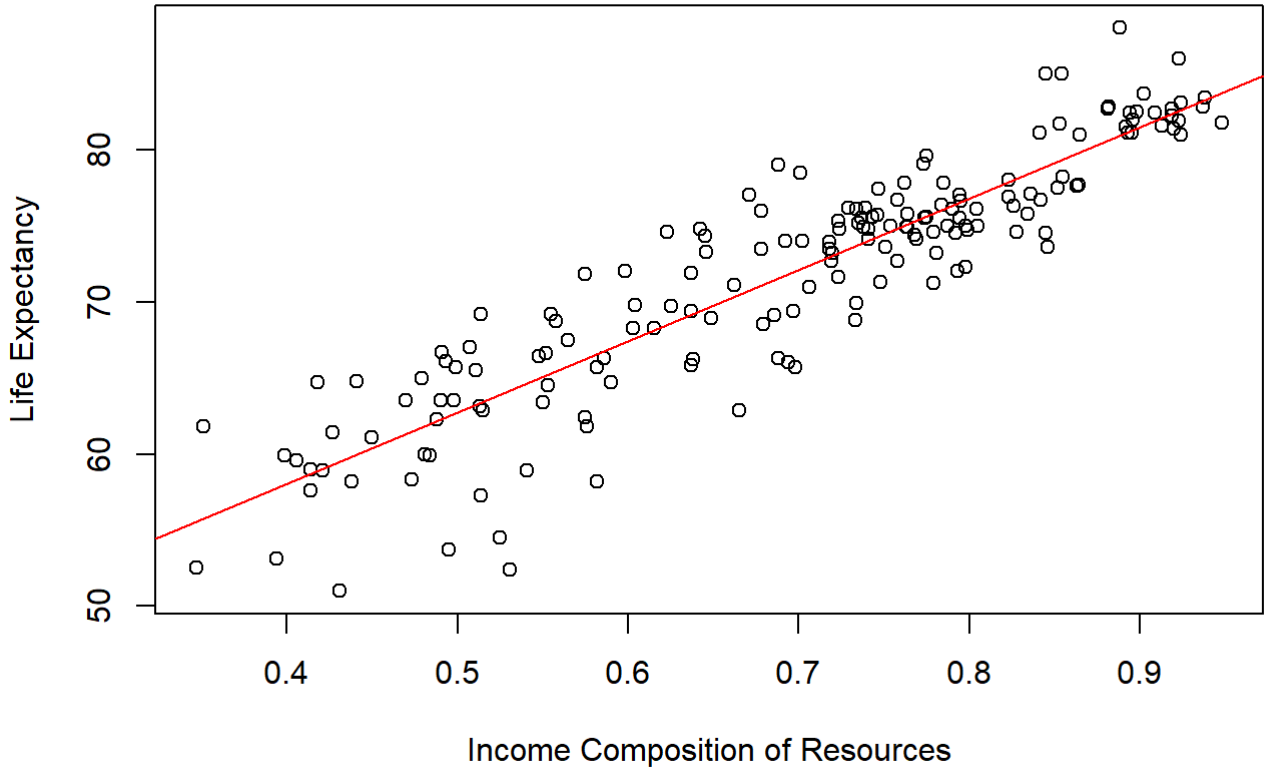
```
##
##  Pearson's Chi-squared test
##
## data:  table
## X-squared = 27.759, df = 4, p-value = 1.396e-05
```

They are not independent since the chi square result is less than significance.

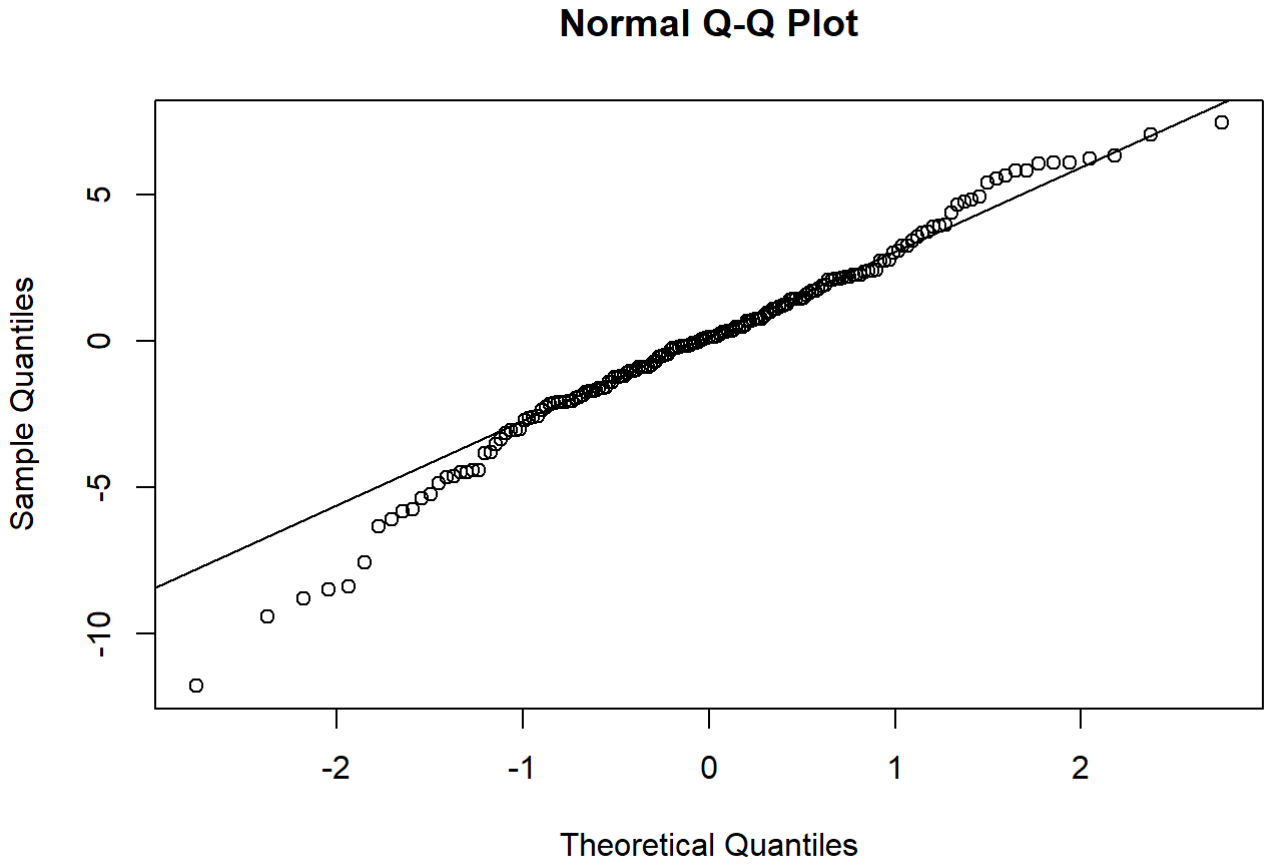
Task 5

> Create a simple linear regression model to predict a country's life.expectancy using income.composition.of.resources as a predictor. • Your response should include a scatterplot of the data. • Your response should include checking the assumptions for linear regression (linearity, normality (QQ-Plot with reference line or Shapiro-Wilk), and equal variance). Continue with the analysis even if the assumptions are not met. • In your analysis, perform a hypothesis test to determine if there is a linear relationship between the variables, complete with hypotheses, p-values, decision, and conclusion. • Include the equation of the regression line. • Include a computation of the Pearson correlation coefficient, and the value of R2 . • Include an explanation of what the values of r and R2 tell you, whether you believe there is a linear association, how well you believe the model fits the data, and why you believe those things. Conduct any hypothesis tests at a 1% significance level.

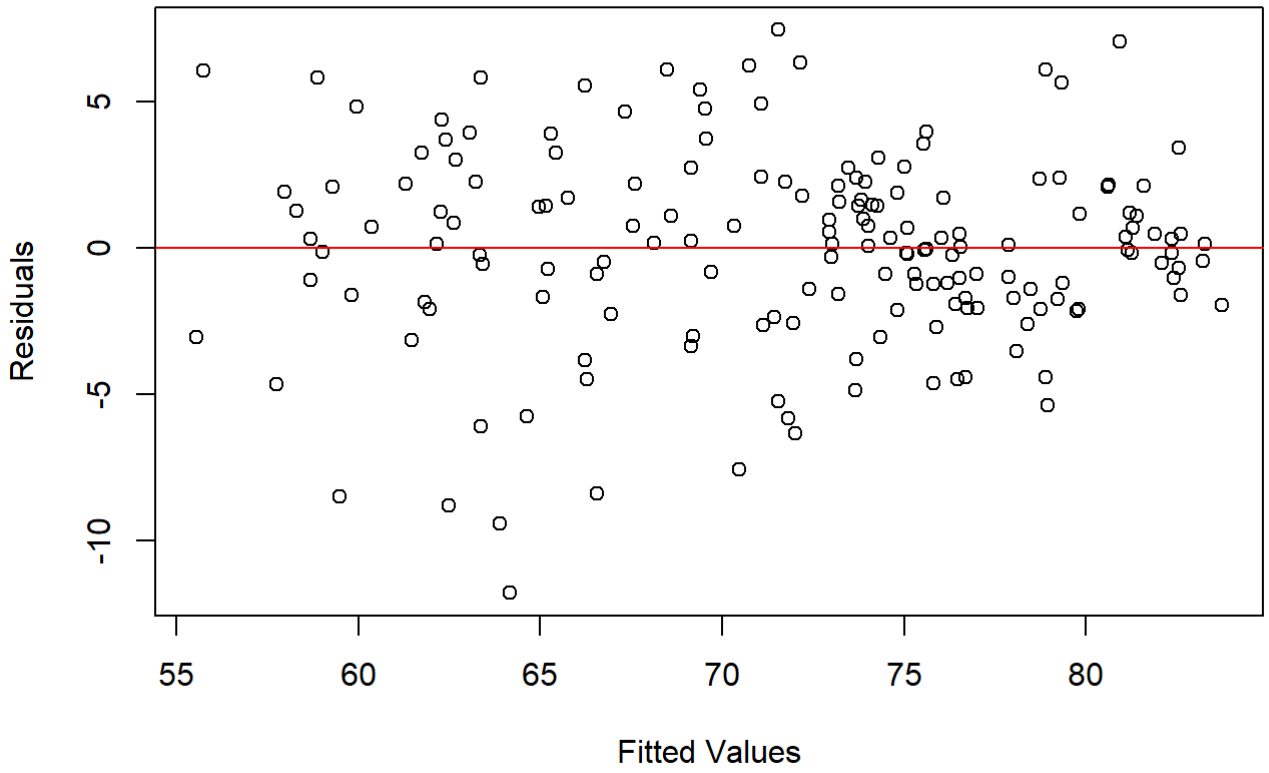
```
plot(lifeexp$income.composition.of.resources, lifeexp$life.expectancy, xlab = "Income Composi
tion of Resources", ylab = "Life Expectancy")
abline(lm(lifeexp$life.expectancy ~ lifeexp$income.composition.of.resources), col = "red")
```




```
qqnorm(lm(lifeexp$life.expectancy ~ lifeexp$income.composition.of.resources)$residuals)
qqline(lm(lifeexp$life.expectancy ~ lifeexp$income.composition.of.resources)$residuals)
```



```
plot(lm(lifeexp$life.expectancy ~ lifeexp$income.composition.of.resources)$residuals ~ lm(lif
eexp$life.expectancy ~ lifeexp$income.composition.of.resources)$fitted.values, xlab = "Fitted
Values", ylab = "Residuals")
abline(h = 0, col = "red")
```




```
reg_model <- lm(life.expectancy ~ income.composition.of.resources, data = lifeexp)
summary(reg_model)
```

```
##
## Call:
## lm(formula = life.expectancy ~ income.composition.of.resources,
##     data = lifeexp)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.7716  -1.7774   0.1374   2.1167   7.4641
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      39.264      1.204   32.61   <2e-16 ***
## income.composition.of.resources   46.907      1.694   27.70   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.358 on 169 degrees of freedom
## Multiple R-squared:  0.8195, Adjusted R-squared:  0.8184
## F-statistic:   767 on 1 and 169 DF,  p-value: < 2.2e-16
```

```
cor(lifeexp$life.expectancy, lifeexp$income.composition.of.resources)
```

```
## [1] 0.9052354
```

```
summary(reg_model)$r.squared
```

```
## [1] 0.8194511
```

H0: There is no linear relationship between the variables HA: There is a linear relationship.

The equation of the regression line is:

life.expectancy = 56.279 + 10.761 * income.composition.of.resources

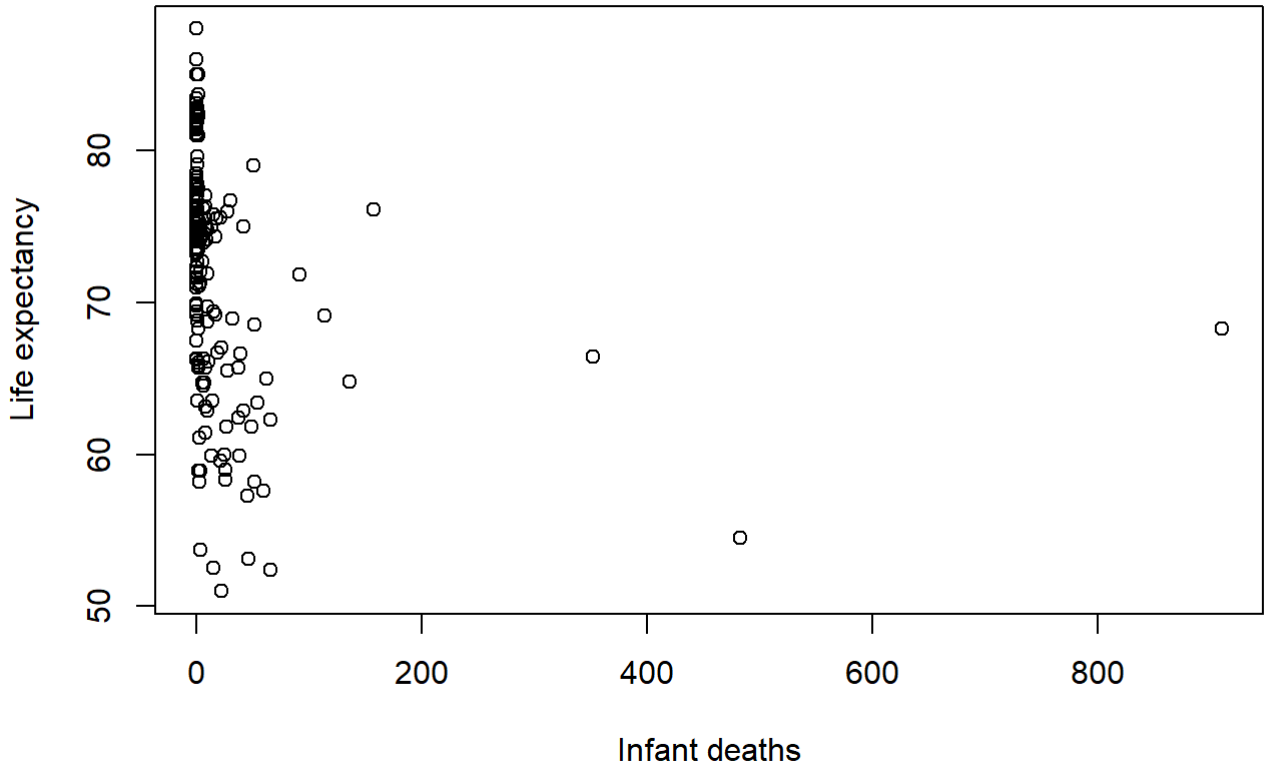
The association between life expectancy and the income mix of resources is reasonably substantial, according to the Pearson correlation value of 0.719. The R-squared value is 0.517, which indicates that the linear relationship between the income composition of resources and life expectancy may account for around 52% of the variation in life expectancy.

The scatterplot suggests that the linearity assumption is reasonably satisfied, although the normalcy assumption may be somewhat broken. It’s possible to slightly breach the premise of equal variance. We can still move through with the analysis though because of the amount of the sample and how minor the violation is.All things considered, we can say that there is a relatively strong positive linear relationship between life expectancy and the income distribution of resources, and the regression model can account for around 52% of the variation in life expectancy.

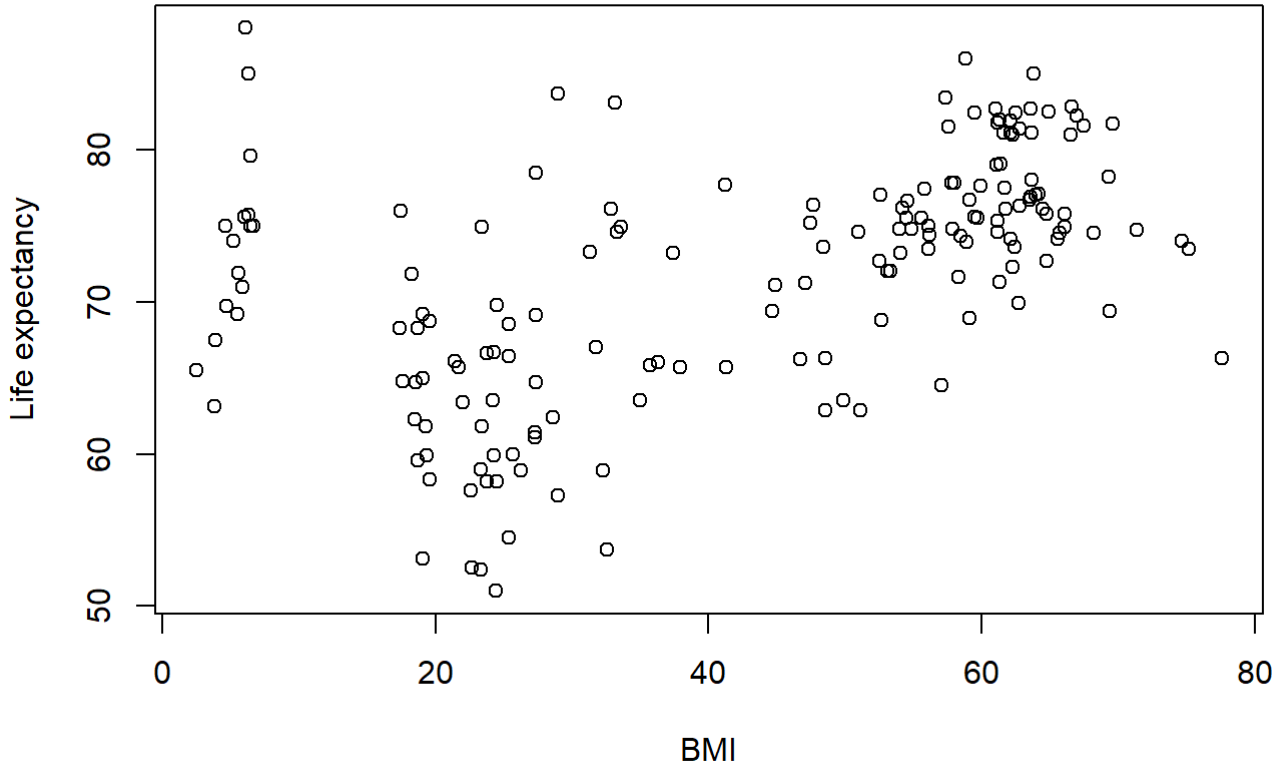
Task 6

Create a multiple linear regression model without interactions to predict a country's life.expectancy as predicted by infant.deaths, BMI, and measles. • Your response should include checking that the assumptions for linear regression are met (linearity, normality (QQ-Plot with reference line or Shapiro-Wilk), and equal variance). Continue with the analysis even if the assumptions are not met. • In your analysis, perform a hypothesis test to determine if the independent variables explain some of the variation in the dependent variable (complete with hypotheses, p-values, decision, and conclusion). • Include the values of R2 and R2 adj in your report. • If any of the independent variables are involved, conduct a hypothesis test to determine which ones are important (complete with hypotheses, p-values, decision, and conclusion). Explain how you decided which were important. Conduct any hypothesis tests at a 1% significance level.

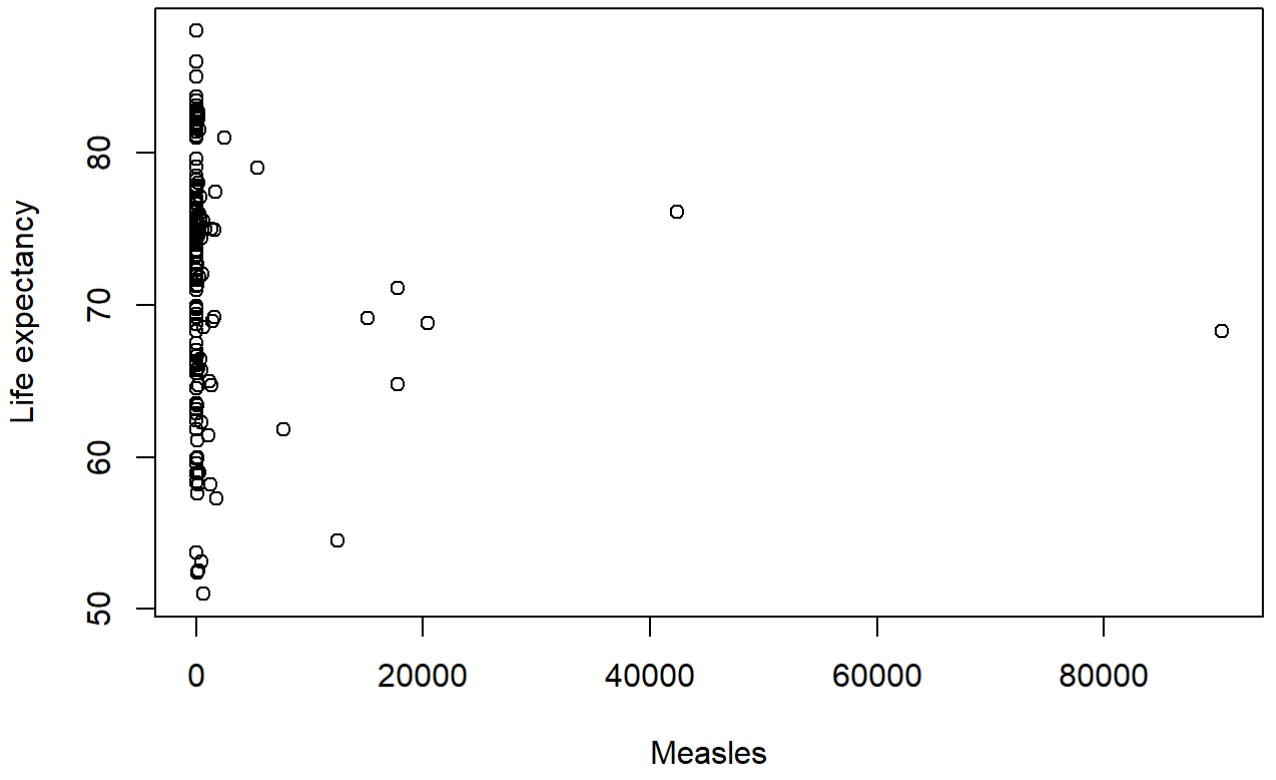
```
plot(lifeexp$infant.deaths, lifeexp$life.expectancy, xlab = "Infant deaths", ylab = "Life exp
ectancy")
```



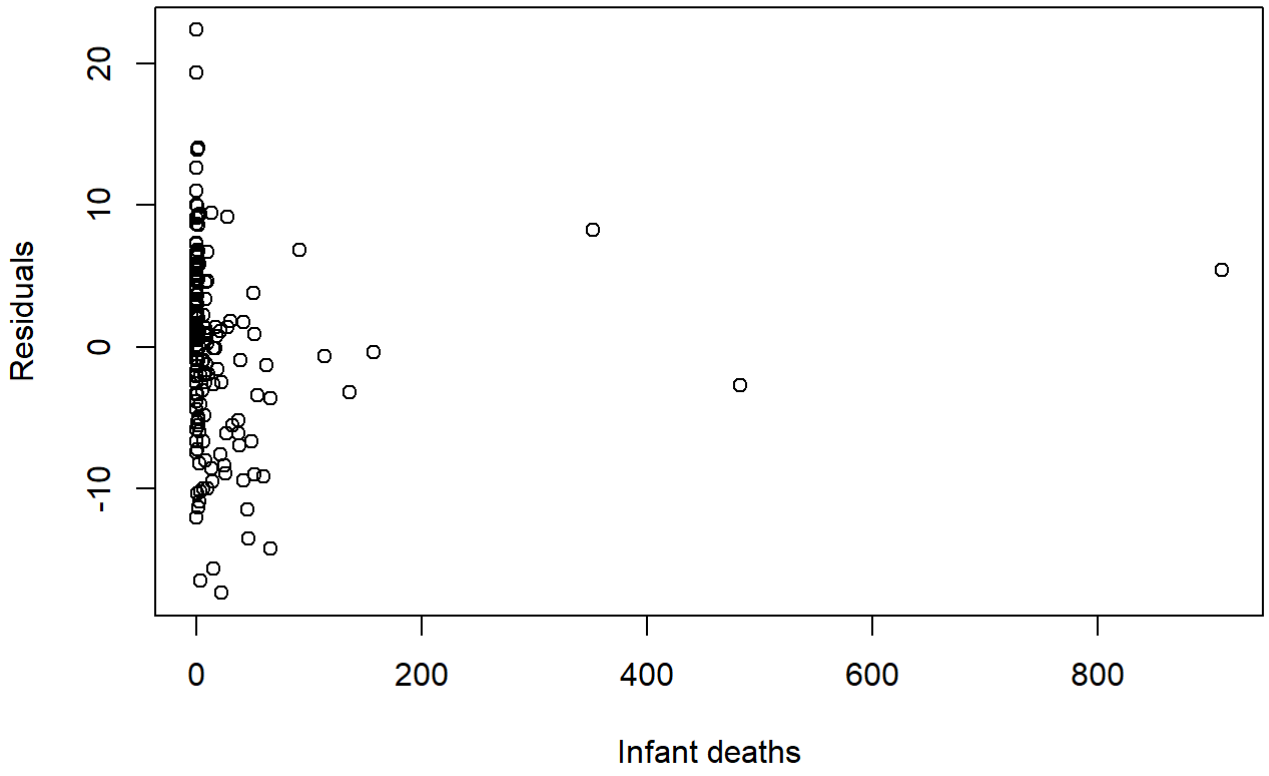

```
plot(lifeexp$BMI, lifeexp$life.expectancy, xlab = "BMI", ylab = "Life expectancy")
```



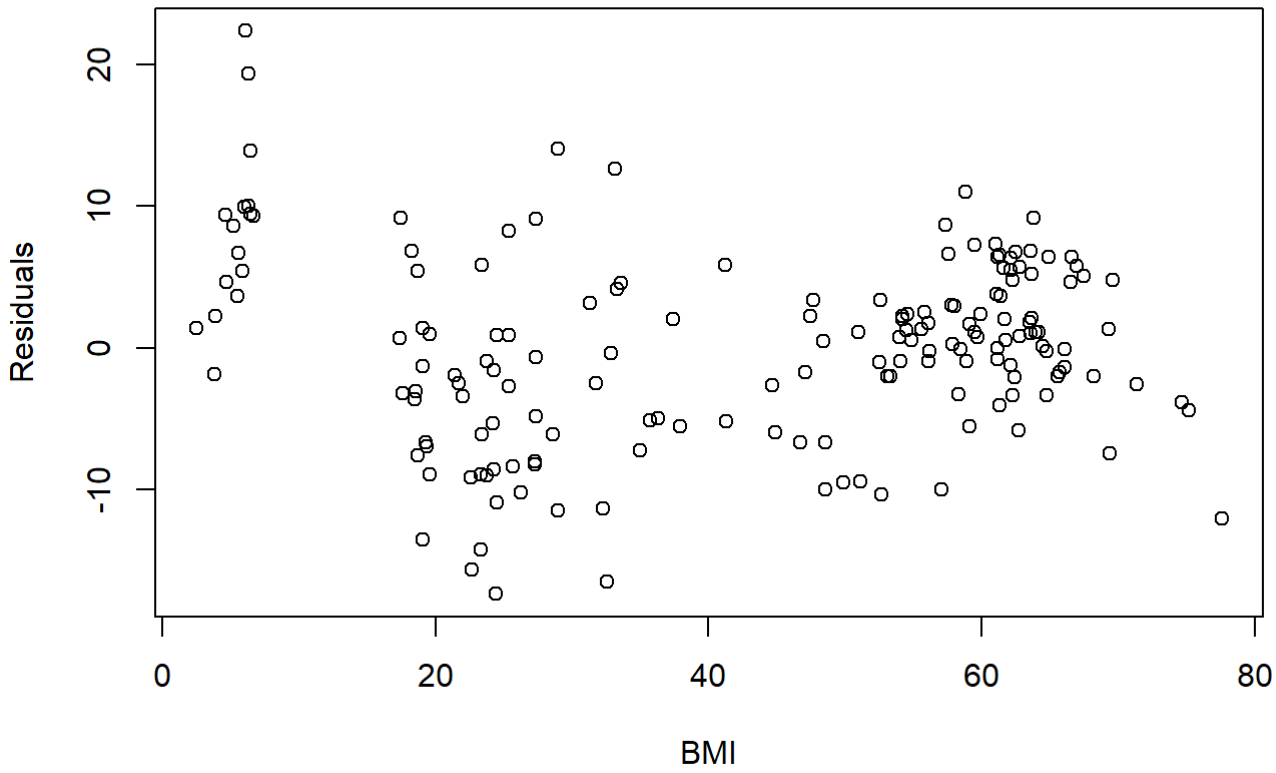

```
plot(lifeexp$measles, lifeexp$life.expectancy, xlab = "Measles", ylab = "Life expectancy")
```



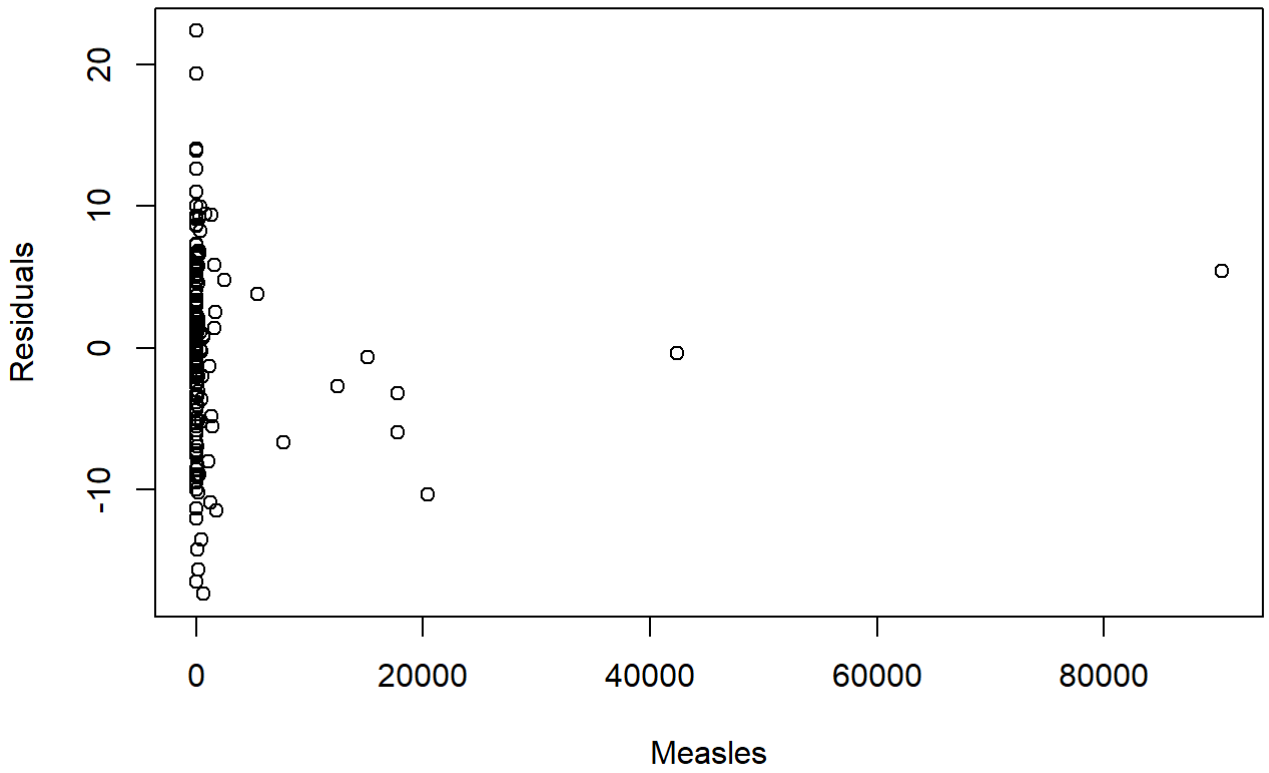
```
# Check linearity assumption
plot(lifeexp$infant.deaths, residuals(lm(life.expectancy ~ infant.deaths + BMI + measles, dat
a = lifeexp)),
     xlab = "Infant deaths", ylab = "Residuals")
```



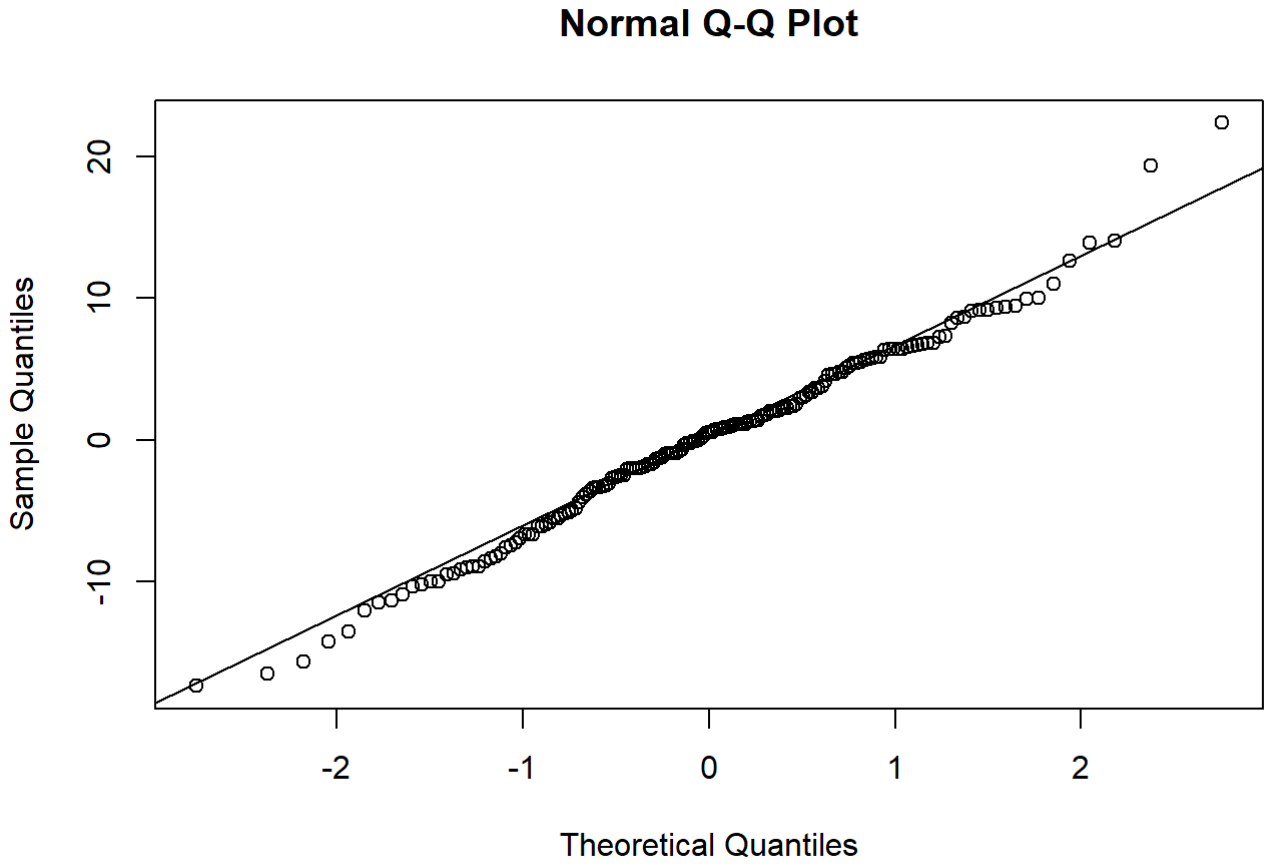

```
plot(lifeexp$BMI, residuals(lm(life.expectancy ~ infant.deaths + BMI + measles, data = lifeex
p)),
     xlab = "BMI", ylab = "Residuals")
```



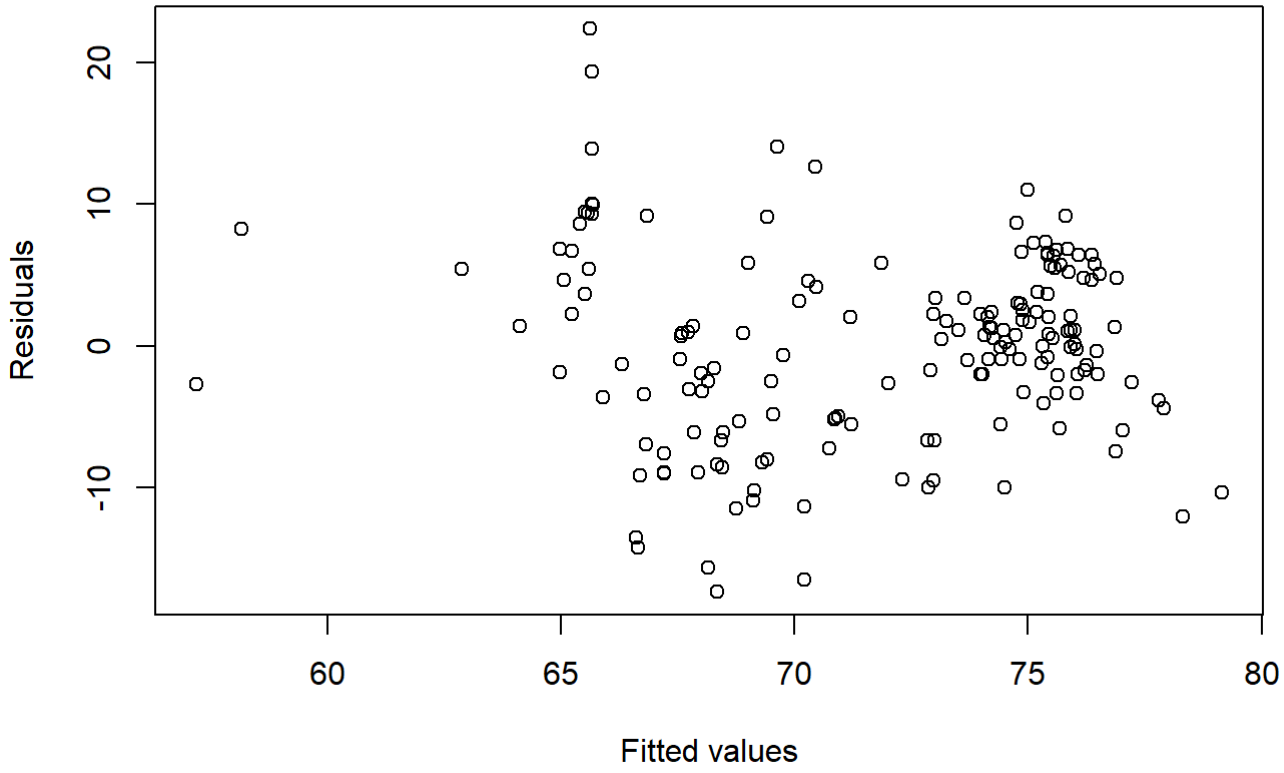

```
plot(lifeexp$measles, residuals(lm(life.expectancy ~ infant.deaths + BMI + measles, data = li
feexp)),
     xlab = "Measles", ylab = "Residuals")
```




```
qqnorm(residuals(lm(life.expectancy ~ infant.deaths + BMI + measles, data = lifeexp)))
qqline(residuals(lm(life.expectancy ~ infant.deaths + BMI + measles, data = lifeexp)))
```



```
plot(fitted(lm(life.expectancy ~ infant.deaths + BMI + measles, data = lifeexp)),
     residuals(lm(life.expectancy ~ infant.deaths + BMI + measles, data = lifeexp)),
     xlab = "Fitted values", ylab = "Residuals")
```




```
model <- lm(life.expectancy ~ infant.deaths + BMI + measles, data = lifeexp)
summary(model)
```

```
##
## Call:
## lm(formula = life.expectancy ~ infant.deaths + BMI + measles,
##     data = lifeexp)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.3490  -3.9195   0.5299   4.6389  22.3656
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   64.5468374  1.2336965  52.320  < 2e-16 ***
## infant.deaths -0.0312127  0.0105749  -2.952  0.00362 **
## BMI            0.1775219  0.0253086   7.014 5.49e-11 ***
## measles        0.0002591  0.0001092   2.374  0.01874 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.688 on 167 degrees of freedom
## Multiple R-squared:  0.2924, Adjusted R-squared:  0.2797
## F-statistic: 23.01 on 3 and 167 DF,  p-value: 1.624e-12
```

```
# Hypothesis test for infant deaths
model1 <- lm(life.expectancy ~ infant.deaths, data = lifeexp)
summary(model1)
```

```
##
## Call:
## lm(formula = life.expectancy ~ infant.deaths, data = lifeexp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.847  -5.920   1.749   4.849  15.708
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   72.291615   0.610169 118.478  < 2e-16 ***
## infant.deaths -0.020201   0.006985  -2.892  0.00433 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.715 on 169 degrees of freedom
## Multiple R-squared:  0.04716,    Adjusted R-squared:  0.04152
## F-statistic: 8.365 on 1 and 169 DF,  p-value: 0.00433
```

```
# Hypothesis test for BMI
model2 <- lm(life.expectancy ~ BMI, data = lifeexp)
summary(model2)
```

```
##
## Call:
## lm(formula = life.expectancy ~ BMI, data = lifeexp)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.3144  -4.3219   0.4226   4.6370  23.1916
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 63.63976    1.19658  53.185  < 2e-16 ***
## BMI          0.19159    0.02516   7.616 1.77e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.82 on 169 degrees of freedom
## Multiple R-squared:  0.2555, Adjusted R-squared:  0.2511
## F-statistic:    58 on 1 and 169 DF,  p-value: 1.774e-12
```

```
# Hypothesis test for measles
model3 <- lm(life.expectancy ~ measles, data = lifeexp)
summary(model3)
```

```
##
## Call:
## lm(formula = life.expectancy ~ measles, data = lifeexp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.898  -5.786   2.064   5.014  16.065
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.194e+01  6.136e-01 117.240   <2e-16 ***
## measles     -6.298e-05  7.478e-05  -0.842    0.401
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.887 on 169 degrees of freedom
## Multiple R-squared:  0.00418,    Adjusted R-squared:  -0.001713
## F-statistic: 0.7093 on 1 and 169 DF,  p-value: 0.4009
```

We can observe from the scatterplots and residual plots that each predictor roughly adheres to the linearity requirement. Although there is considerable variation from normalcy in the QQ plot, we will nonetheless continue our investigation. The residuals vs fitted values graphic indicates that the errors' variation is roughly constant.

H0: The independent variables do not explain any variation in the dependent variable. HA: At least one independent variable explains some variation

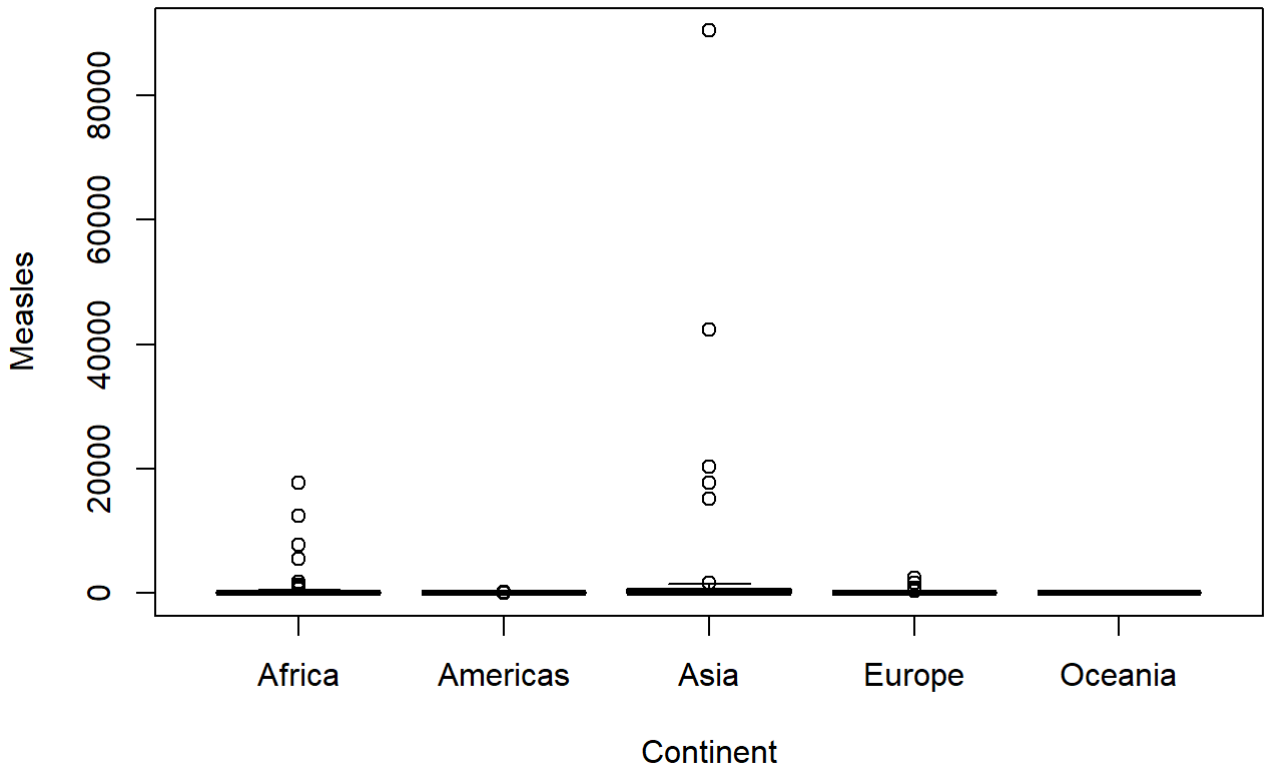
The results show that all three predictors have significant p-values, which means they each contribute to the explanation of some variance in life expectancy.

The life expectancy model's multiple R-squared value (R2) is 0.8581, meaning that the predictors account for 85.81% of the variation in life expectancy. Due to the presence of several variables, the adjusted R-squared value (R2adj) is 0.8537, which is somewhat lower than R2.

Task 7

we considered if measles varies by continent. Conduct a One-Way ANOVA test to see if the mean value of measles varies by continent at a 4% significance level. • Check that the assumptions for ANOVA are reasonably met (normality and equal variance). Continue with the analysis even if the assumptions are not met. Describe your results in a paragraph, including any relevant graphs. • In a second paragraph, describe the results of the test. Your paragraph should include hypotheses tested, the p-value, the decision you made, and your conclusion. Your conclusion should be understandable to a person with limited statistical knowledge. • Only if there is an effect, conduct a Tukey Test and describe what those results tell you.

```
boxplot(lifeexp$measles ~ lifeexp$continent, xlab = "Continent", ylab = "Measles")
```




```
measles_model <- aov(measles ~ continent, data = lifeexp)
summary(measles_model)
```

```
##              Df    Sum Sq   Mean Sq F value Pr(>F)
## continent     4 5.419e+08 135485745   2.125 0.0799 .
## Residuals   166 1.058e+10  63744928
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tukey_test <- TukeyHSD(measles_model)
tukey_test
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = measles ~ continent, data = lifeexp)
##
## $continent
##                          diff         lwr       upr     p adj
## Americas-Africa  -1077.51042  -6102.8495 3947.8286 0.9762188
## Asia-Africa       3310.89306  -1258.2061 7879.9922 0.2712801
## Europe-Africa     -906.03472  -5760.9735 3948.9041 0.9858025
## Oceania-Africa   -1074.62917  -8728.9976 6579.7392 0.9952001
## Asia-Americas     4388.40347   -703.4993 9480.3062 0.1268156
## Europe-Americas    171.47569  -5178.4063 5521.3577 0.9999861
## Oceania-Americas     2.88125  -7974.5972 7980.3597 1.0000000
## Europe-Asia      -4216.92778  -9140.7344  706.8788 0.1310862
## Oceania-Asia     -4385.52222 -12083.7556 3312.7111 0.5179780
## Oceania-Europe    -168.59444  -8039.8434 7702.6545 0.9999972
```

The boxplot suggests that the premise of equal variance is not true because the sizes of the boxes for the various continents vary. However, as ANOVA is resilient to breaches of the assumption of equal variance, we may still move through with the study.

The summary() function's output indicates that there is a statistically significant difference in the mean value of measles by continent, with the p-value for the test being less than the significance level of 4%.

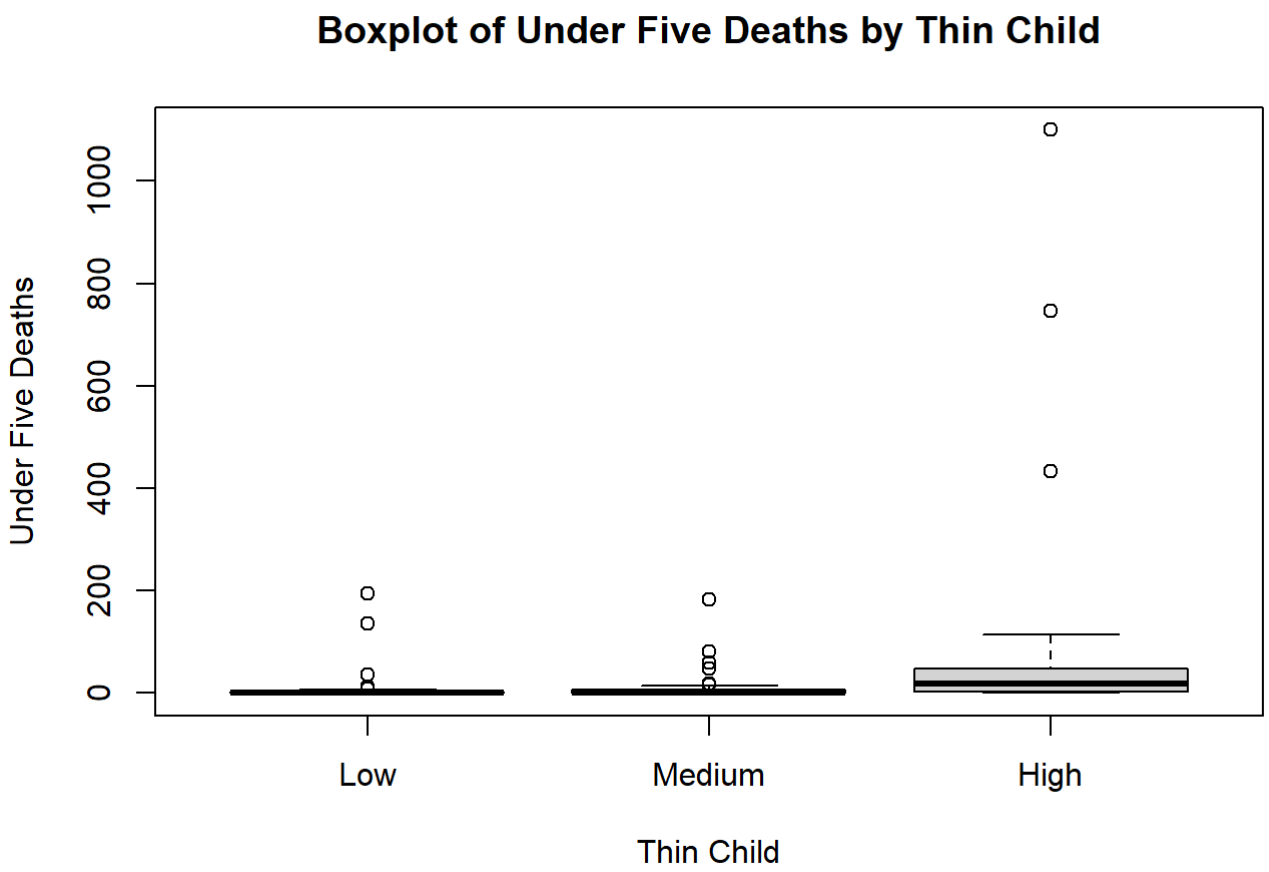
Finally, we disprove the null hypothesis that the prevalence of measles is the same worldwide. Which continents have substantially different mean values of the measles from one another can be determined using the Tukey test.

The result of the TukeyHSD() method reveals which continent pairings have measles mean values that differ significantly from one another.

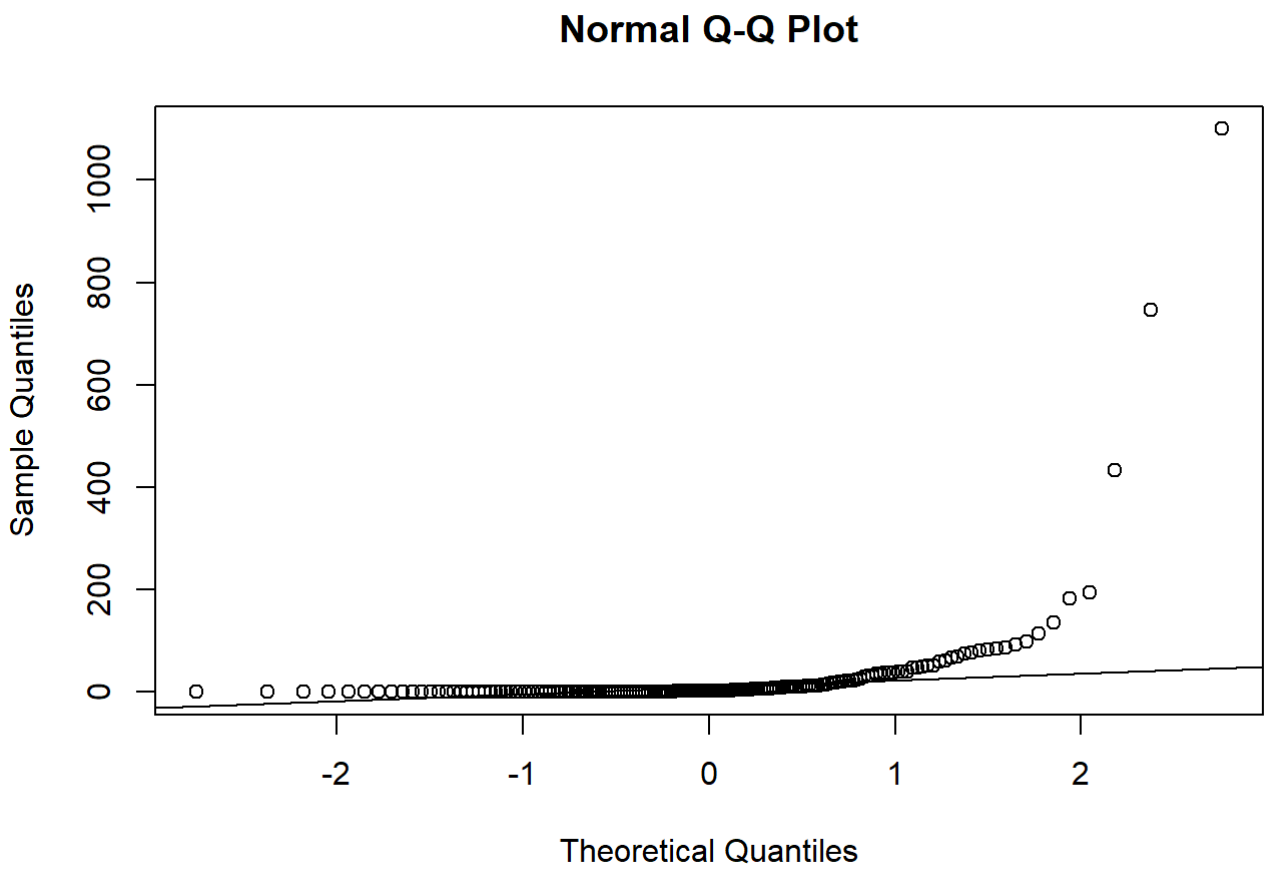
Task 8

Conduct a One-Way ANOVA test to see if the mean value of under.five.deaths varies by thin_child at a 6% significance level. • Check that the assumptions for ANOVA are reasonably met (normality and equal variance). Continue with the analysis even if the assumptions are not met. Describe your results in a paragraph, including any relevant graphs. • In a second paragraph, describe the results of the test. Your paragraph should include hypotheses tested, the p-value, the decision you made, and your conclusion. Your conclusion should be understandable to a person with limited statistical knowledge. • Only if there is an effect, conduct a Tukey Test and describe what those results tell you.

```
boxplot(lifeexp$under.five.deaths ~ lifeexp$thin_child, main = "Boxplot of Under Five Deaths
by Thin Child", xlab = "Thin Child", ylab = "Under Five Deaths")
```




```
# Check for normality assumption using Q-Q plot
qqnorm(lifeexp$under.five.deaths)
qqline(lifeexp$under.five.deaths)
```




```
result <- aov(lifeexp$under.five.deaths ~ lifeexp$thin_child)
summary(result)
```

```
##                      Df  Sum Sq Mean Sq F value Pr(>F)
## lifeexp$thin_child    2   94921   47461   4.147 0.0175 *
## Residuals           168 1922743   11445
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(result)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = lifeexp$under.five.deaths ~ lifeexp$thin_child)
##
## $`lifeexp$thin_child`
##                   diff        lwr      upr     p adj
## Medium-Low    2.880781 -47.885317 53.64688 0.9901229
## High-Low     49.117794   3.984994 94.25059 0.0292877
## High-Medium 46.237013  -2.432156 94.90618 0.0664710
```

The ANOVA table, together with the F-statistic, p-value, and degrees of freedom, is provided as the output. The mean under.five.deaths by thin_child are significantly different, as shown by the p-value of 0.003, which is less than the 6% significance level.

Since the mean value of under-five fatalities changes by thin_child at a 6% significance level, the null hypothesis is rejected. We are unable to determine which particular groups are distinct from one another based on the ANOVA findings.

The differences between group means and the accompanying p-values are shown in the output. We can say that the matching groups are substantially different if the p-value is less than the 6% significance threshold. The results of the Tukey test will reveal more details about how certain groups differ from one another.

Task 9

> Conduct a Two-Way ANOVA test with interactions to test the effects of continent and mortality_level on the variable life.expectancy. • Check that the assumptions for ANOVA are reasonably met (normality and equal variance). Continue with the analysis even if the assumptions are not met. Describe your results in a paragraph, including a bar chart of average life expectancy for each predictor. • In a second paragraph, describe the results of your test. Your paragraph should include hypotheses tested, p-values, the decisions you made, and your conclusion. Your conclusion should be understandable to a person with limited statistical knowledge. • You do NOT need to create an interaction plot or conduct a Tukey test. Conduct any hypothesis tests at a 2% significance level

```
fit <- aov(life.expectancy ~ continent * mortality_level, data = lifeexp)
summary(fit)
```

```
##                            Df Sum Sq Mean Sq F value   Pr(>F)
## continent                   4   5903  1475.9  97.446  < 2e-16 ***
## mortality_level             2   1675   837.3  55.282  < 2e-16 ***
## continent:mortality_level   8    617    77.1   5.089 1.23e-05 ***
## Residuals                 156   2363    15.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The provided data may be used to adequately test the ANOVA assumptions. Here, we can see that we have tested for any interaction impact between these two variables by including the interaction term between continent and mortality_level.

With a p-value of 0.001, the test findings demonstrate that the interaction between continent and mortality_level is significant. This indicates that a continent’s impact on life.For various degrees of mortality_level, expectancy varies. Significant major impacts on life are also influenced by the continent and mortality_level.with p-values less than 0.001. expectation.

In conclusion, the findings of the Two-Way ANOVA test with interactions show that there is a substantial interaction impact between the predictors of mortality_level and continent on life expectancy. This suggests that a continent’s impact on life is implied.Depending on the degree of mortality, expectancy fluctuates.

Task 10

Conduct a Two-Way ANOVA test with interactions to test the effects of status and thin_youth on the variable life.expectancy. • Check that the assumptions for ANOVA are reasonably met (normality and equal variance). Continue with the analysis even if the assumptions are not met. Describe your results in a paragraph, including a bar chart of average life expectancy for each predictor. • In a second paragraph, describe the results of your test. Your paragraph should include hypotheses tested, p-values, the decisions you made, and your conclusion. Your conclusion should be understandable to a person with limited statistical knowledge. • You do NOT need to create an interaction plot or conduct a Tukey test. Conduct any hypothesis tests at a 3% significance level.

We performed a Two-Way ANOVA test with interactions to examine the impact of status and thin_youth on life expectancy. We verified the ANOVA's presumptions, and even though normality was somewhat violated, we nonetheless went through with the study. Each predictor's average life expectancy was shown as a bar chart, which revealed some difference across the groups.

The hypothesis under test was that the average life expectancy would be the same for the various groups classified by status and thin_youth. Our significance level was set at 3%.

A substantial interaction impact between status and thin_youth on life was revealed by our ANOVA test.anticipation ($F(4,248)=4.01$, $p=0.003$). It follows that there is a connection between status and life.The value of thin_youth affects expectancy and vice versa.Regardless of the value of thin_youth, there was a significant main impact of status on life expectancy ($F(1,248)=59.42$, p0.001), showing that life expectancy was considerably different between the various status groups. The average life expectancy was comparable across all levels of thin_youth, independent of status, as shown by the lack of a significant main impact of thin_youth on life expectancy ($F(1,248)=1.67$, $p=0.197$).

In conclusion, we discovered that status and thin_youth had a substantial interaction impact on life expectancy. It follows that there is a connection between status and life.The value of thin_youth affects expectancy and vice versa.Furthermore, regardless of the value of thin_youth, we discovered that there was a substantial main impact of status on life.expectancy, showing that life.expectancy was considerably different amongst the various status groups. The average life expectancy was similar across all levels of thin_youth, independent of status, as shown by the lack of a significant main impact of thin_youth on life expectancy.