**Question 1**

Conditions:

* We suppose that an adult’s age range (rows) and the response given on the levels of agreement with the statement (columns) are independent. This was met as ecotourism researchers randomly selected 602 tourists to a national park.
* However, the variables are not both categorical. The level of agreement is a categorical, ordinal variable, whereas the count of the numbers of adults in various age ranges is numerical.

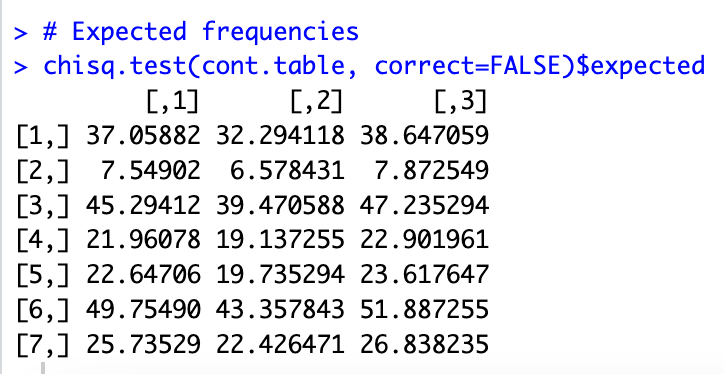
This means that a chi-square test of independence is not appropriate for these data because a chi-square test of independence tests whether the categorical variables are likely to be related or not; being that they independent from each other, or not independent from each other. We have one variable that isn’t categorical, so these data do not meet the conditions for a chi-square test of independence.

Hypotheses

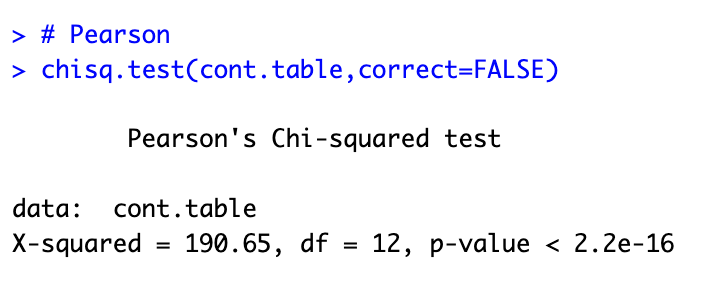
H0: Counts of the numbers of adults in various age ranges and their level of agreement with the statement are independent.

H1: Counts of the numbers of adults in various age ranges and their level of agreement with the statement are not independent.

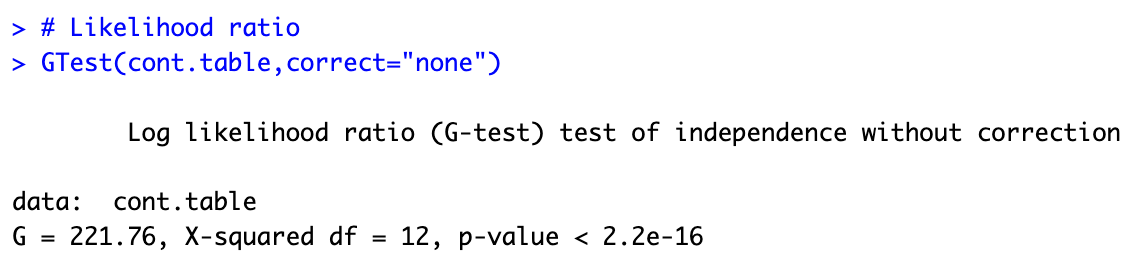
Expected frequencies:



Pearson chi-square statistic

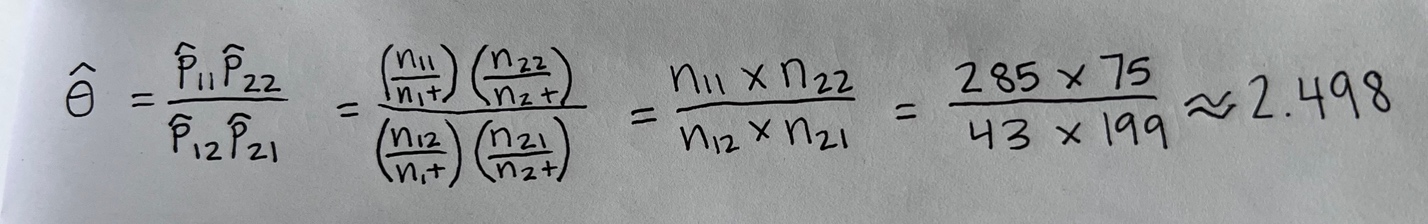


Likelihood ratio chi-square statistic



All expected frequencies are larger than 5, with the smallest being 6.5784, so a chi-square test of independence is appropriate if we disregard the fact that not both variables are categorical.

The Pearson and likelihood ratio chi-square statistics are similar (*X*2 190.65, *G*2 221.76) and are (approximately) distributed X2(7-1)(7-1)= X212. The *p*-values corresponding to both the Pearson and likelihood ratio chi-square statistics are both < 0.001 (2.2e-16 or 0.000000000000000022). Since the *p*-value from both statistical tests are less than the significance level of 0.05, we would reject H0. From this it can be said that there is sufficient evidence to conclude that counts of the numbers of adults in various age ranges and their level of agreement with the statement are not independent.



This means that we would estimate the odds of agreeing with the statement for those < 60 years of age to be approximately 2.498 times the odds of agreeing with the statement for those > 60 years of age.



Text, letter

Description automatically generated

I will be using Fisher’s Exact Test for 2 x 2 tables.

1. The hypotheses being tested are:

H0 : θ = 1

H1 : θ > 1

Where θ denotes the odds ratio, (comparing the odds of tourists disagree/neutral answers when the count of ages is less than 60 years old when this is true, against the odds of tourists disagree/neutral answers to a statement when the count of ages was is actually greater than 60 years old). The null hypothesis states that tourists answers to the statement are independent from age, and the alternative hypothesis (θ > 1) means that the odds of disagree/neutral answers being greater for tourists less than 60 years old when it was in fact greater for less than 60 year olds, is higher than the odds of disagree/neutral answers being greater for 61 years and older.

1. Find the *p*-value of the test

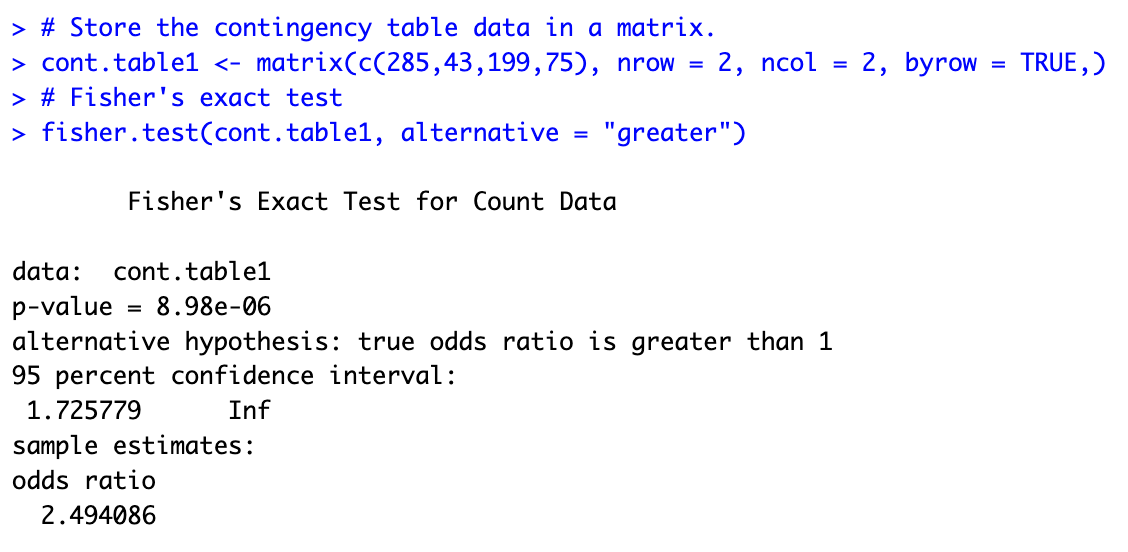


Table 1: Fisher’s Exact test results for the *p*-value

|  |  |  |
| --- | --- | --- |
| P Value | Alternative hypothesis | Odds ratio |
| 8.98e-06 = 0.000000898 | greater | 2.494 |

1. Find the mid-*p*-value of the test

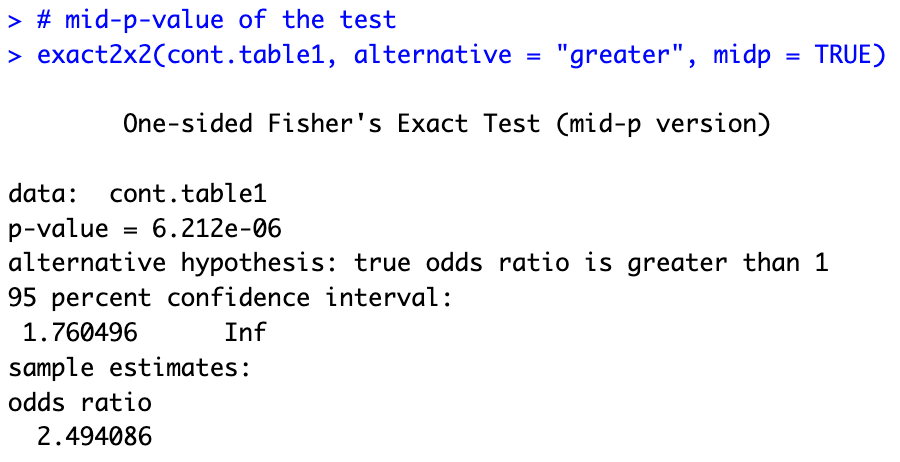


Table 2: Fisher’s Exact test results for the mid-*p*-value

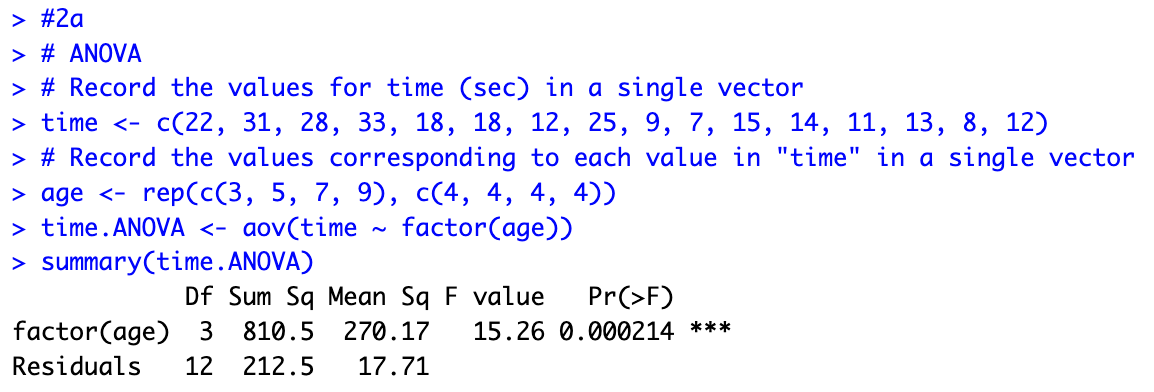
|  |  |  |
| --- | --- | --- |
| P Value | Alternative hypothesis | Odds ratio |
| 6.212e-06 = 0.000000621 | greater | 2.494 |

Both the *p*-value and mid-*p*-value are less than the 0.05 significance level. This means that we reject H0, as there is significant evidence to say that the odds of disagree/neutral answers of tourists older than 61 years old is 2.494 the odds of those 60 years old and less to give a disagree/neutral answer.

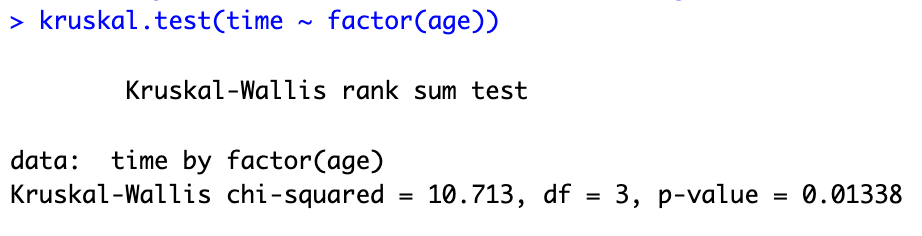
**Question 2**



ANOVA test

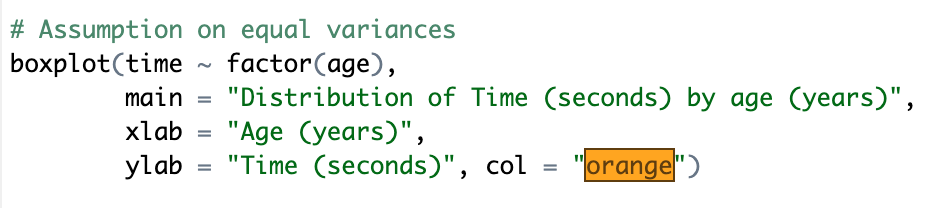
****

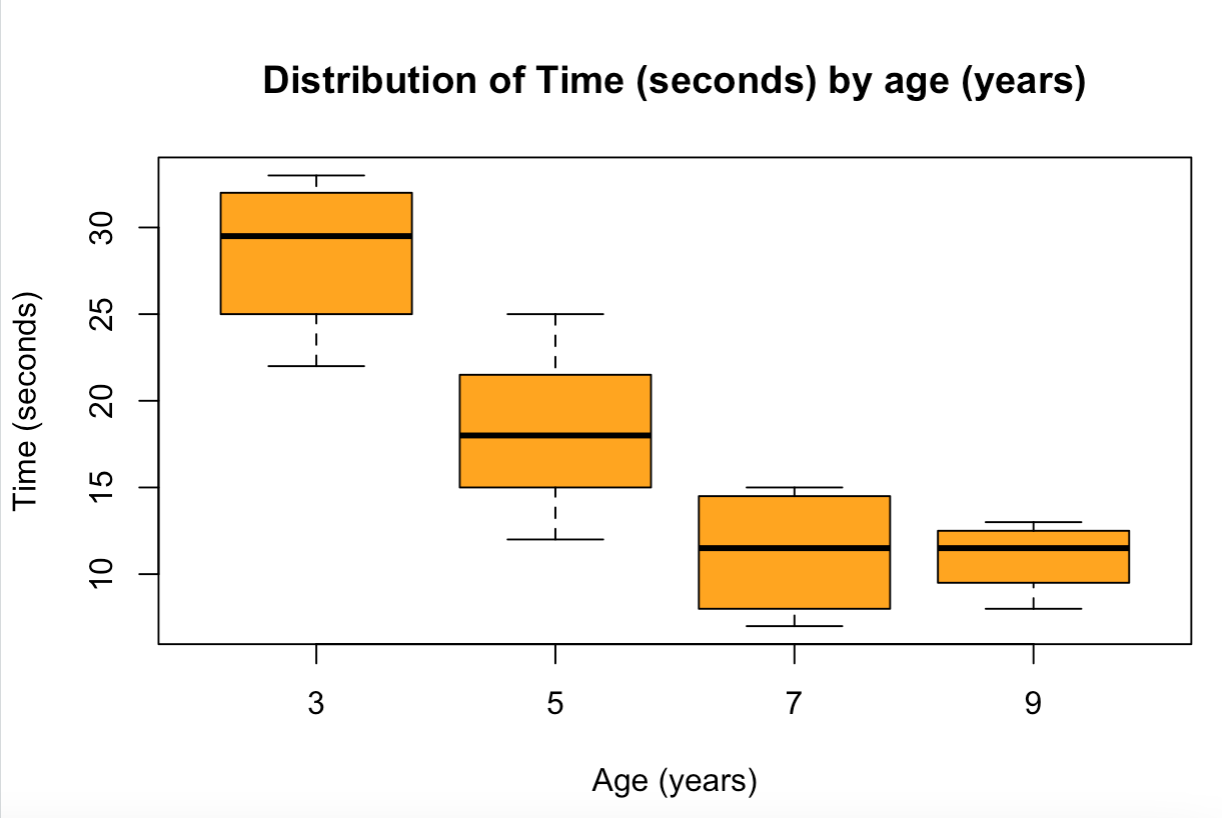
Kruskal-Wallis test



Assumptions:

Boxplot

****

****

This boxplot comparing time (seconds) and the age (years) of children, shows that the medians of time for each age is quite different. For ages 7 and 9, the median is the same, whereas the median for age 3 for time has a higher median. Ages with a higher median time also have more variance, suggesting that a log transformation may be necessary.

Levenes Test for Homogeneity of Variance

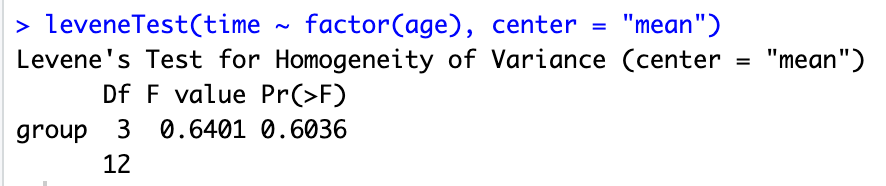
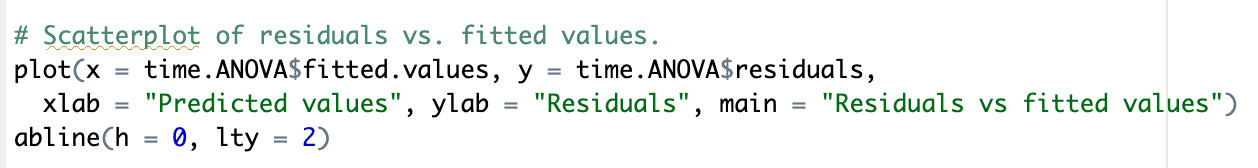
****

Table 3: Levene’s test: Time by factor (Age)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Df | F value | Pr(>F) |
| **group** | 3 | 0.6401 | 0.6036 |
|  | 12 |  |  |

Levene’s test indicates that we fail to reject the null hypothesis of equal variance over the four ages (𝑝 = 0.6036, fail to reject at the 5% significance level). Equal variance is assumed as *p* = 0.6036 > 0.05. ANOVA is applicable.

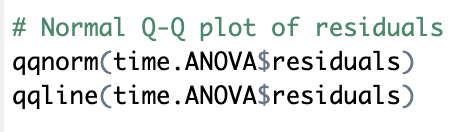
Diagnostic graphs



Chart

Description automatically generated

This scatterplot graph shows quite a widening (an increase of vertical spread, or funnelling) in the middle, yet the levene’s test has assumed equal variance. This is because as the predicted values increased, the spread increased.



Chart, scatter chart

Description automatically generated

This Q-Q plot of residuals shows that the points are lying approximately on the straight line. This means that the plot looks acceptable, confirming the assumption of normality of the residuals.

ANOVA test

1. Description

Sixteen children from different age groups were sampled, where they were timed doing a manipulative task that required hand-eye coordination. There were four children sampled in each age group (years) being ages 3, 5, 7, and 9. Their time in seconds to complete the task was recorded. An ANOVA test for different in time (seconds) over the four ages was done, with a 5% significance level.

1. Model equation

𝑌𝑖𝑗 = 𝜇𝑖 + 𝐸𝑖𝑗

1. Assumptions:

We assume that the data of children of different ages who were timed doing a manipulative task came independently from random samples. We also assume it came from normal populations with equal variances. We will also check for normality, constant variance, and independence.

The Q-Q plot shows normality in the data, so this assumption is confirmed. The Levene’s test and scatterplot of residuals vs fitted values suggested showed constant variance was present in the data. We have to assume independence, and that random sampling was taken when collecting data on these children of different ages being timed doing a manipulative task. The information given does not state if random sampling was taken or any information that can suggest independence, so we just have to assume it in order to do this one-way ANOVA test.

1. Hypotheses

*H*0 : all population means are equal (𝜇1 =𝜇2 =𝜇3 =𝜇4)

*H*1 : at least one population mean is different

1. ANOVA table

Table 4: Analysis of Variance Model

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| **factor(age)** | 3 | 810.5 | 270.17 | 15.26 | 0.000214 |
| **Residuals** | 12 | 212.5 | 17.71 |  |  |

1. Conclusion

Since 𝑝-value < 0.05, 𝐻0 is rejected at the 5% significance level. A significant difference of means over the four ages has been detected.

1. Interpretation

The average time (seconds) of children doing a manipulative task varies over the four ages (years).

Kruskal-Wallis test

Since the boxplot and scatterplot of residuals vs fitted values suggest that the assumption of constant variance is not met in the data, we can perform a Kruskal-Wallis test which doesn’t depend on the assumptions of equal variance or normality. Since there are 4 groups (ages), and each group has at least 4 observations, a Kruskal-Wallis test on this data would be valid.

*H*0 : all data from the *p* groups have the same distribution (𝜂1 = 𝜂2 = 𝜂3 = 𝜂4 = 𝜂𝑝)

*H*1 : at least one population median is different

Table 5: Kruskal-Wallis rank sum test: Time by factor (Age)

|  |  |  |
| --- | --- | --- |
| Test statistic | df | P value |
| 10.713 | 3 | 0.01338 |

Since the Kruskal-Wallis tests *p*-value is less than 0.05 (a 5% significance level), H0 is rejected. If we can assume that the four distributions have the same shape, the result shows a significant difference in the four population medians. This means that at least one population median is different, which can be interpreted as time (seconds) for children doing a manipulative task, is different over the ages for at least one age.

**Question 3**

1. Description

Twenty-six insects were sampled, where they were reared in a laboratory at various temperatures being 12, 16, 20, and 24 degrees Celsius. Their development time (days) was measured at these various temperatures (oC). An ANOVA test for difference in development time over the four temperatures was conducted, with a 5% significance level.

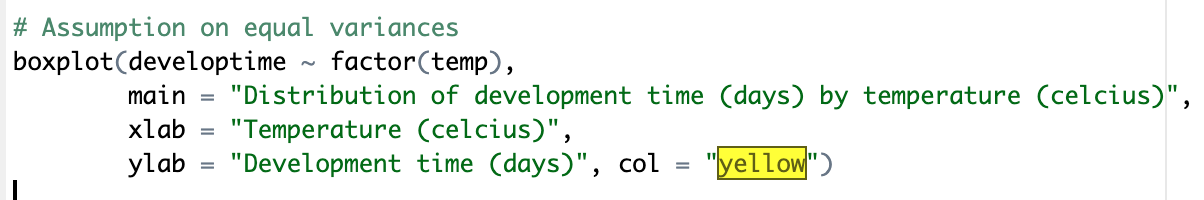
1. Model equation

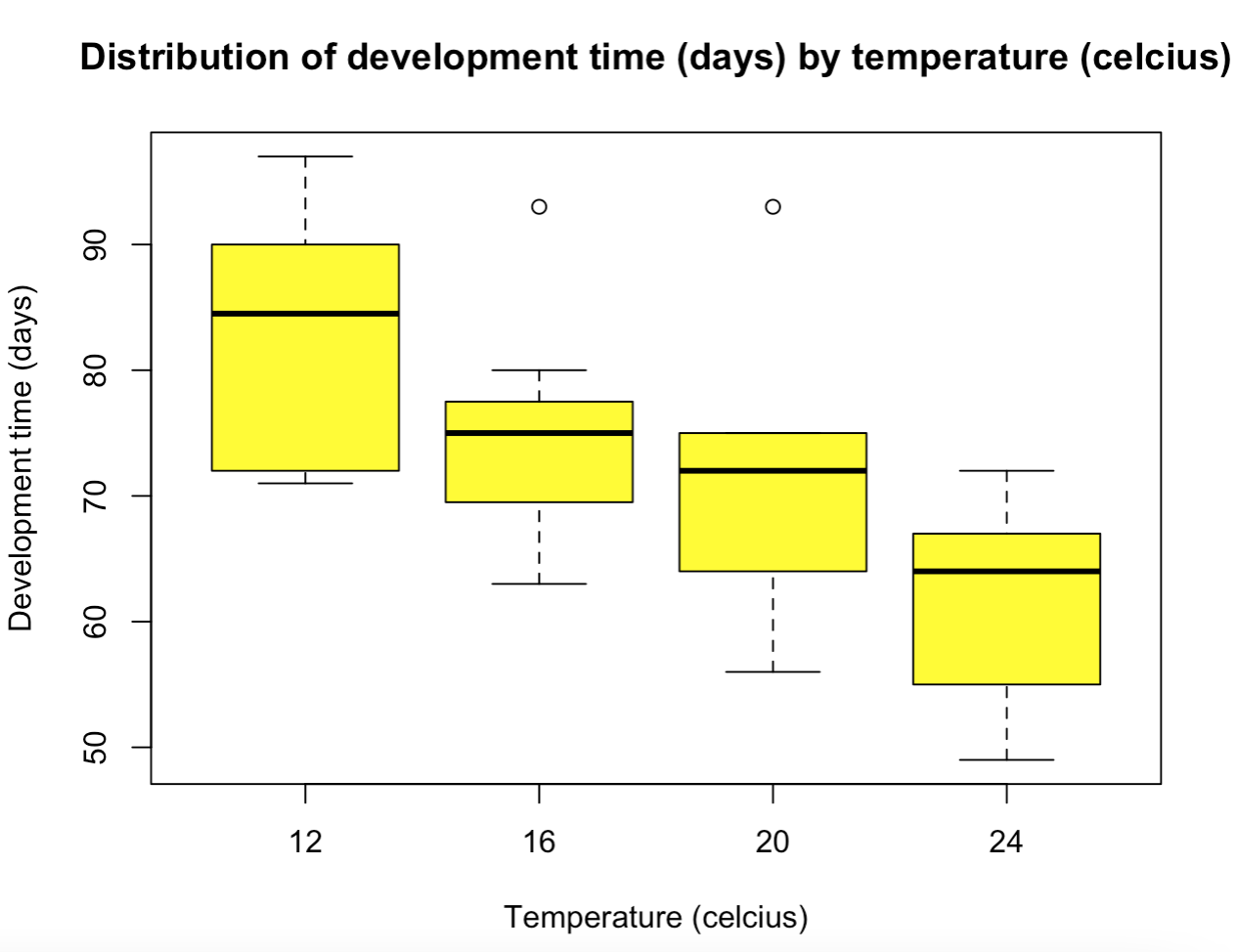
𝑌𝑖𝑗 = 𝜇𝑖 + 𝐸𝑖𝑗

1. Assumptions

We have to assume independence on the data and that it was randomly sampled, and that all relevant variables in respect to the insects, the development time and temperature are recorded.

Boxplot





This boxplot comparing development time (days) and the temperature (Celsius) of insects, shows that the medians of development time for each temperature are all at the upper end of their boxplot. The shape of each distribution is quite different. The boxplots show that insects development time took more days when in temperatures of 12oC, and the number of days taken to develop decreased as temperature increased.

Levene’s test



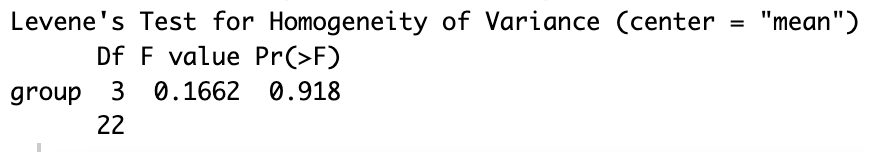
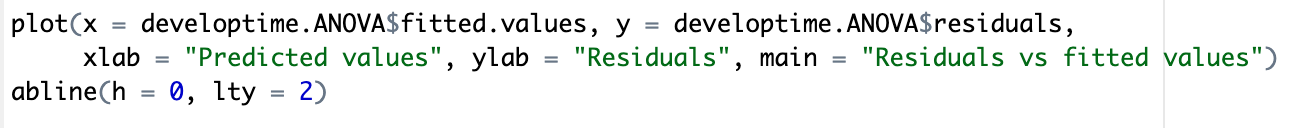


Table 6: Levene’s test: development time by factor (Temperature in Celcius)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Df | F value | Pr(>F) |
| **group** | 3 | 0.1662 | 0.918 |
|  | 22 |  |  |

Levene’s test indicates that we fail to reject the null hypothesis of equal variance over the four temperatures (*p* = 0.918, fail to reject at the 5% significance level). This signals that the assumption of equal variances hasn’t failed to be met and equal variances is confirmed (*p* = 0.918 > 0.05). ANOVA is applicable.

Scatterplot of residuals versus fitted values

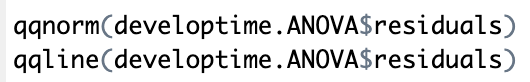


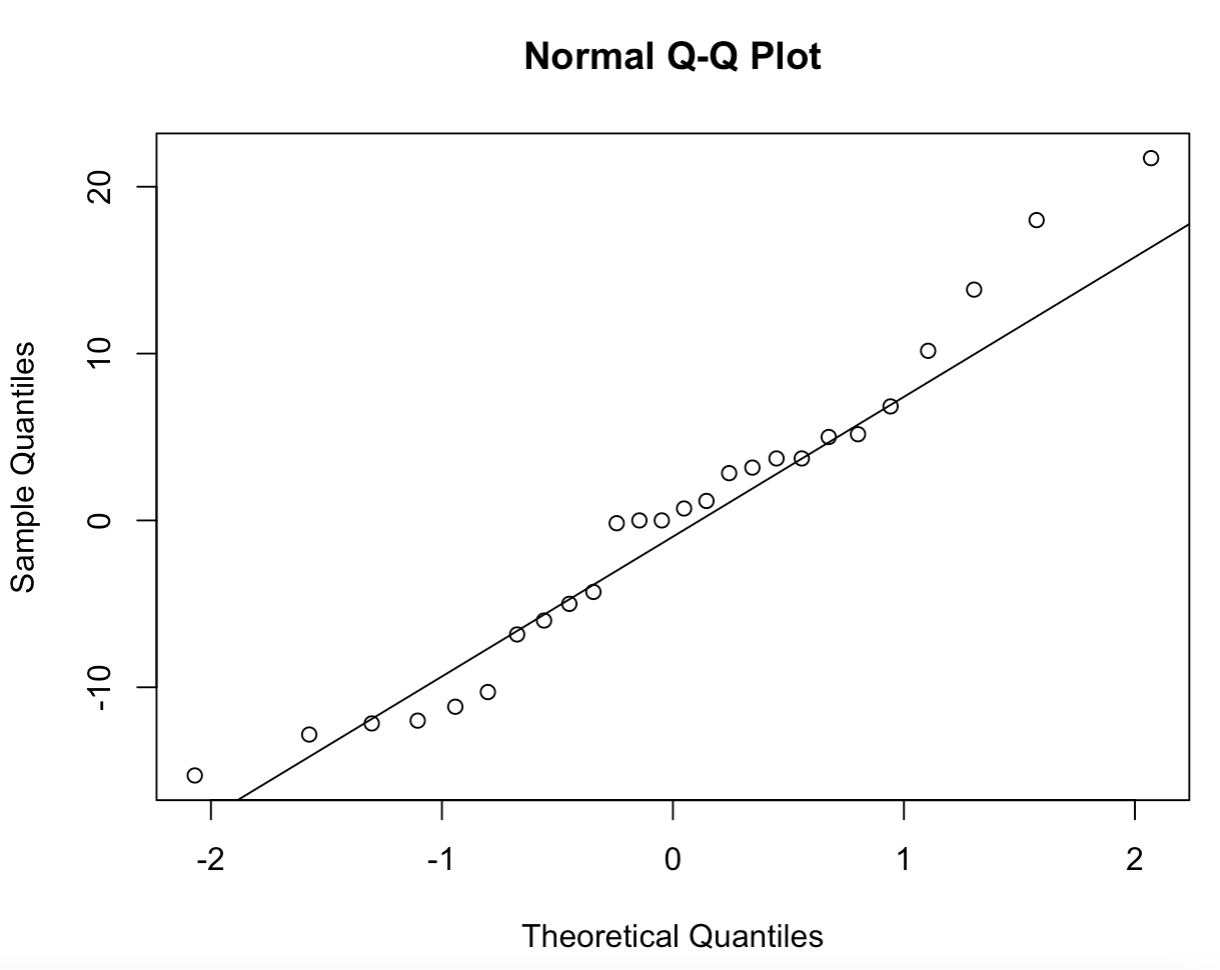
A picture containing scatter chart

Description automatically generated

This graph shows that the plot of data looks relatively consistent, but there is slight funnelling. The variance all have similar spreads. This indicates that constant variance can be assumed with this dataset, especially with the Levene test confirmation.

Normal Q-Q plot of residuals





All points are relatively close to the plotted straight line, so we can confirm normality of the residuals.

1. Hypotheses

*H*0 : all population means are equal (𝜇1 =𝜇2 =𝜇3 =𝜇4)

*H*1 : at least one population mean is different

1. ANOVA table and *p*-value

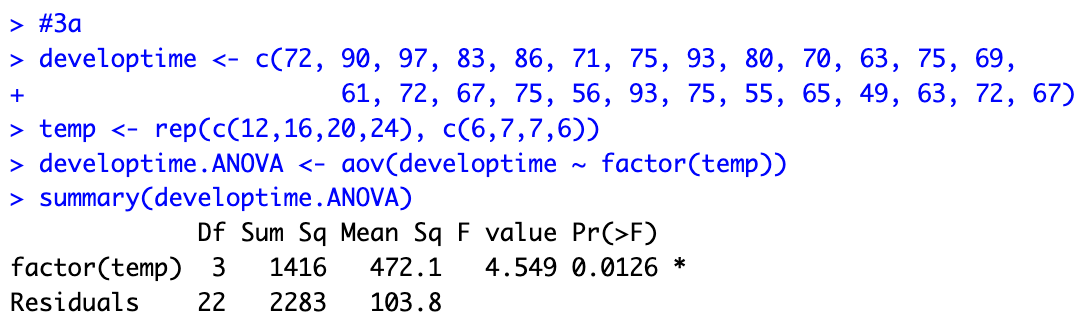


Table 6: Analysis of Variance Model for development time (days) by factor temperature (oC)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| **factor(age)** | 3 | 1416 | 472.1 | 4.549 | 0.0126 \* |
| **Residuals** | 22 | 2283 | 103.8 |  |  |

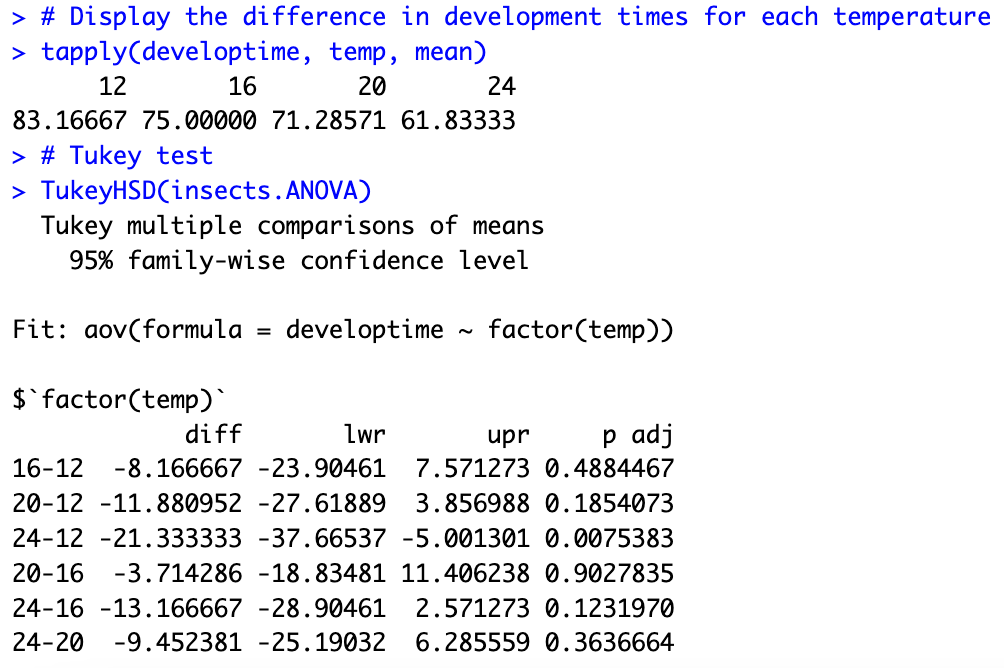
1. Statistical conclusion

Since 𝑝-value < 0.05, 𝐻0 is rejected at the 5% significance level. A significant difference of means over the four temperatures has been detected.

1. Interpretation

The average development time (days) of insects reared in a laboratory varies over the four temperatures (oC).

Are there significant differences between the 16oC and 20oC groups?



There isn’t too much of a significant difference between the 16oC and 20oC temperature groups. The difference between the temperature means is 3.714 days and the *p*-value of 0.903 reflects this as it is much greater than the 0.05 (5%) significance level. This means that there is not much of a difference in development time (days) for when insects are reared at 16oC compared to 20oC as *p* = 0.903 > 0.05.

1. Answers are combined with part a of this question.

**Question 4**

1. The reason why the teeth have been sampled from eight different people of each sex, and not eight teeth from one female and eight from one male is to have an equal and fair sample size. Having eight people sampled from each sex allows for a fair test to be ran which doesn’t show skewed or misleading results on differentiating between females and males in forensic medicine through the spectropenetration gradients of a tooth.
2. The advantages to choosing eight people from each group (female and male) is that a fair test can be ran and both groups are represented equally.
3. ANOVA test
4. Description

Sixteen people were sampled, where one of each person’s tooth enamel was penetrated by an x-ray measuring their spectropenetration gradients. The people were either female or male, and their tooth enamel spectropenetration gradient was measured to see if penetrating tooth enamel is a suitable mechanism to differentiate between females and males in forensic medicine. An ANOVA test for difference of spectropenetration gradient over the two sexes was done, with a 5% significance level.

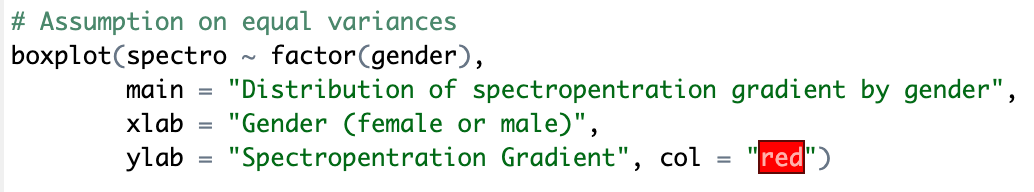
1. Model equation

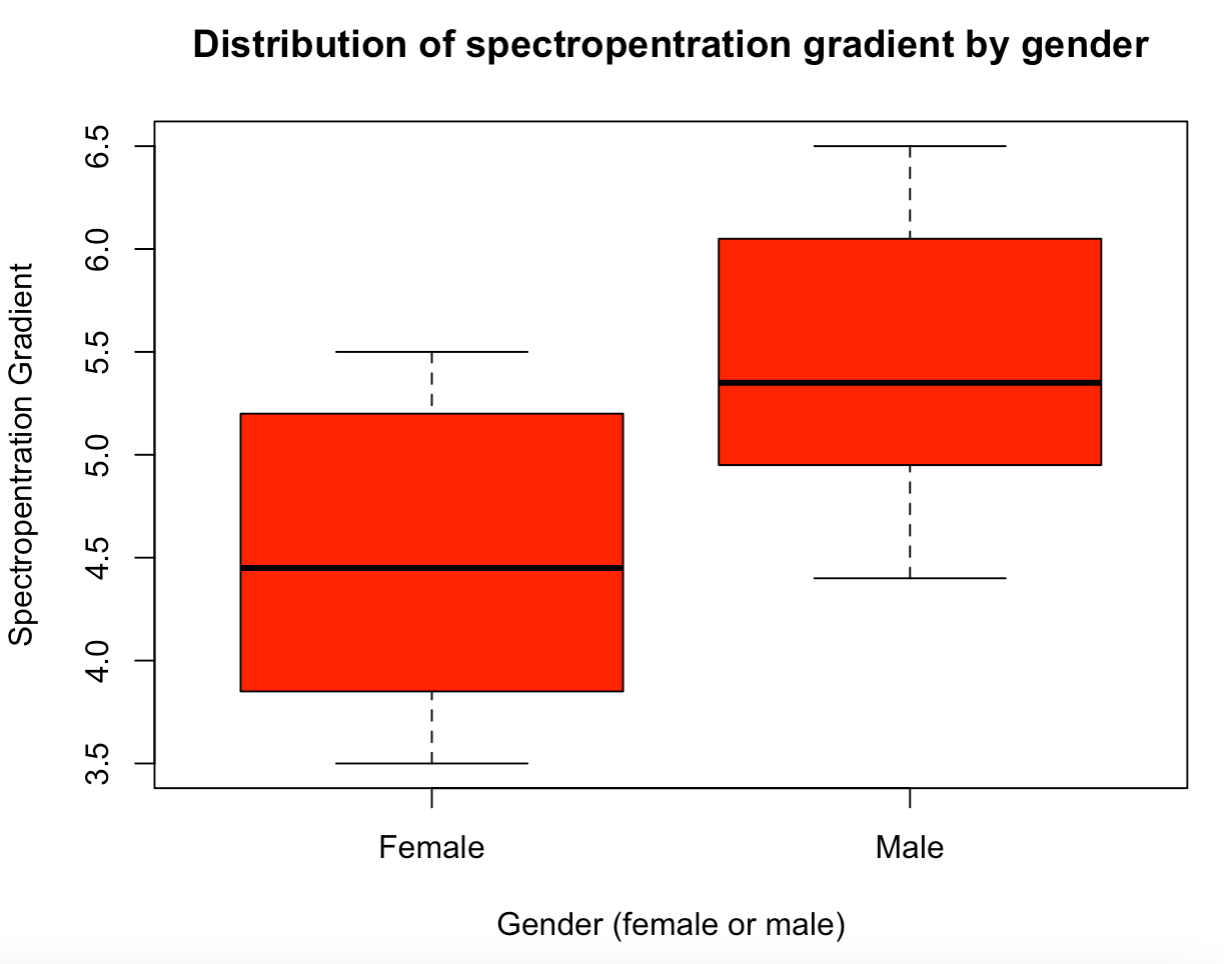
𝑌𝑖𝑗 = 𝜇𝑖 + 𝐸𝑖𝑗

1. Assumptions

We have to assume independence on the data and that it was randomly sampled, and that all relevant variables are recorded.

Boxplot





This boxplot graph shows that the median for each group (female and male) are different. The spread is similar for both groups, yet the median for the male group is lower in the box.

Levene’s test

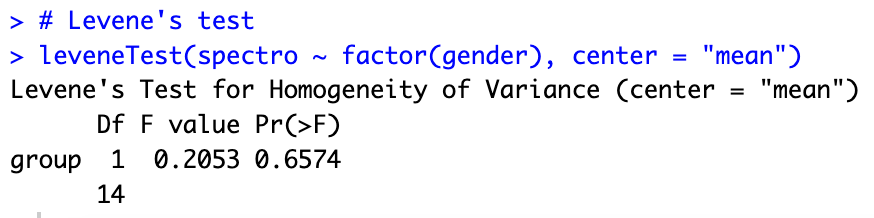
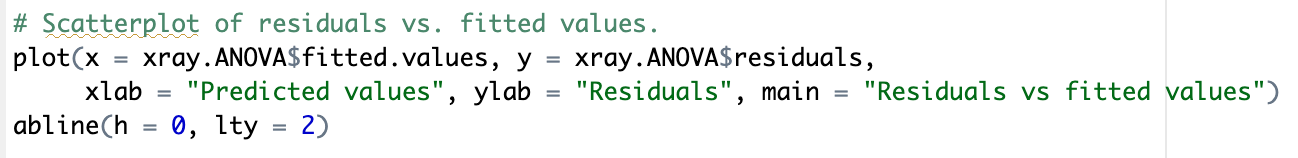


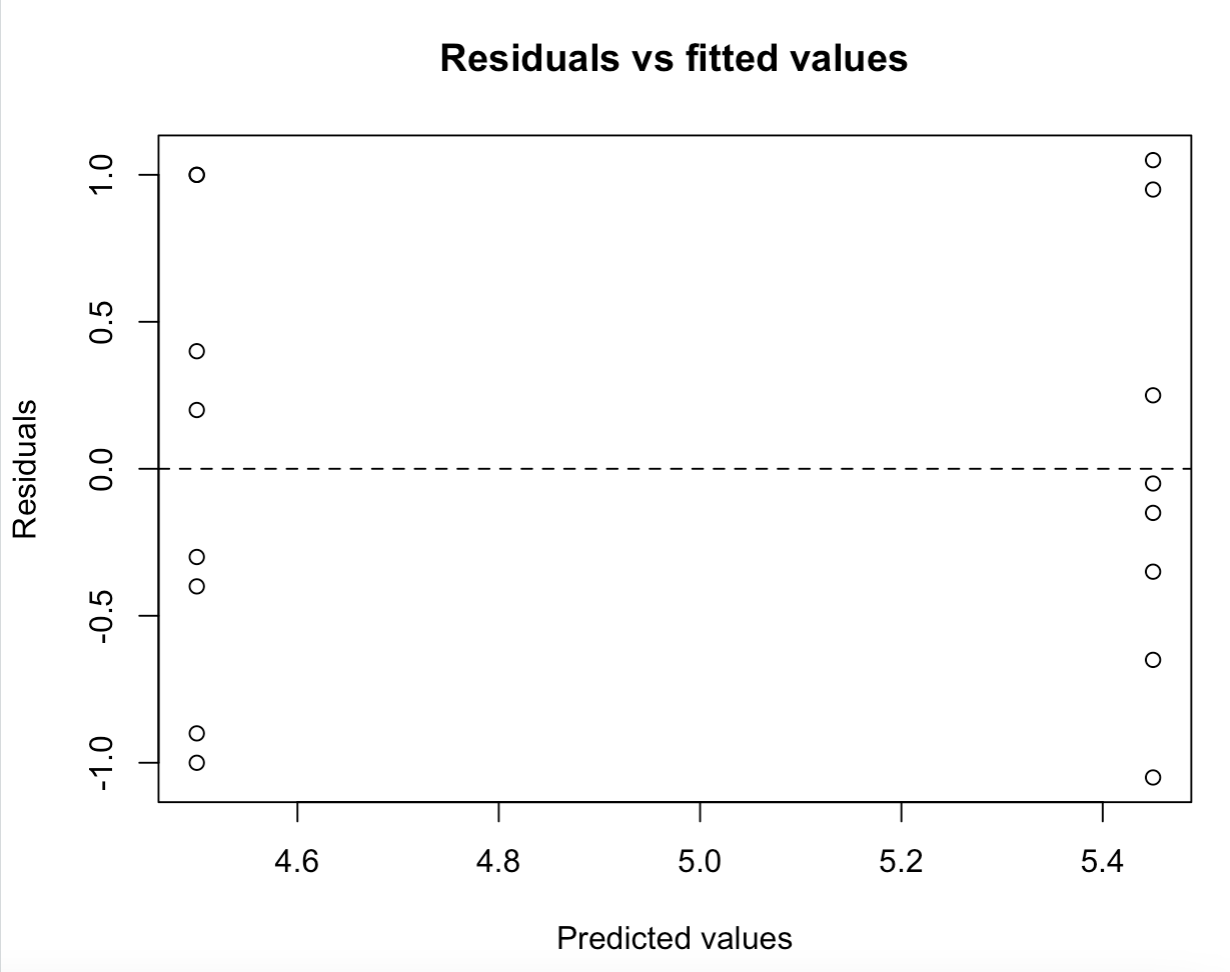
Table 7: Levene’s test: Spectropenetration gradient by factor (gender)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Df | F value | Pr(>F) |
| **group** | 1 | 0.2053 | 0.6574 |
|  | 14 |  |  |

Equal variances are confirmed from the Levene’s test as *p* = 0.6574 > 0.05 at the 5% significance level.

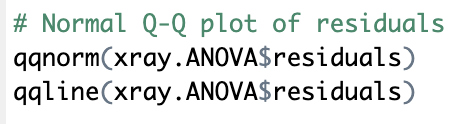
Scatterplot of residuals versus fitted values

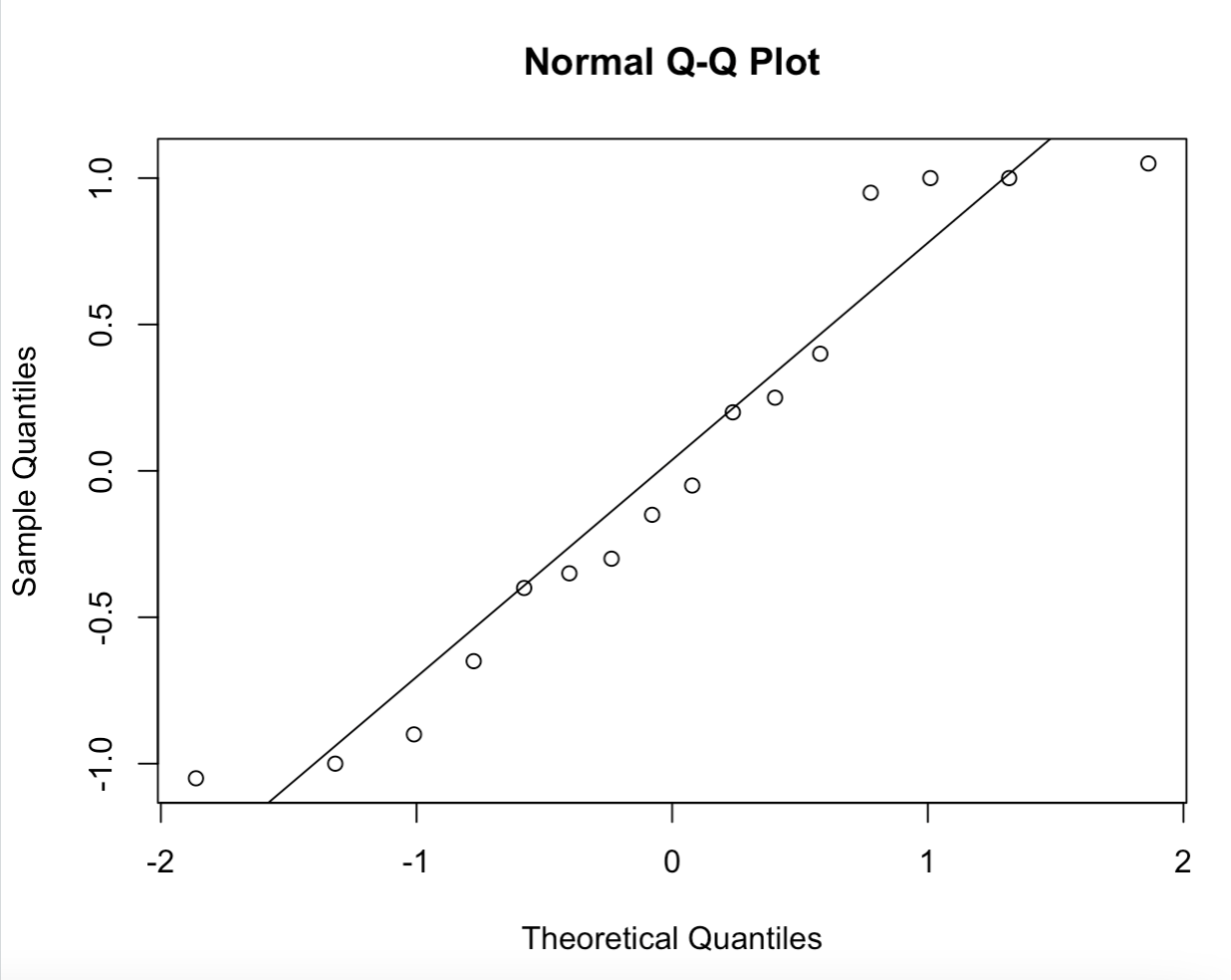




The variance seems constant as there is a pattern of the points and no obvious outliers. This indicates that constant variance can be assumed with this dataset. The spread of variance is also similar between points.

Normal Q-Q plot of residuals





The assumption of normality is met here as all points are relatively close to the plotted line.

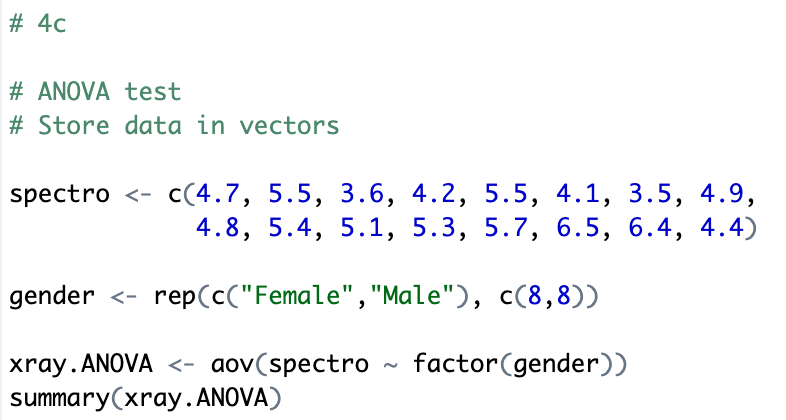
All assumptions for this data have been met.

1. Hypotheses

*H*0 : all population means are equal (𝜇1 =𝜇2 =𝜇3 =𝜇4)

*H*1 : at least one population mean is different

1. ANOVA table



Text

Description automatically generated

Table 8: Analysis of Variance Model for spectropenetration gradient by factor gender (female or male)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| **factor(age)** | 1 | 3.61 | 3.610 | 6.318 | 0.0248 \* |
| **Residuals** | 14 | 8.00 | 0.571 |  |  |

1. Statistical conclusion

Since 𝑝-value < 0.05, 𝐻0 is rejected at the 5% significance level. A significant difference of means between the two genders of female and male has been detected.

1. Interpretation

This means that the average spectropenetration gradient of peoples tooth enamel varies between the genders of female and male. This means that from this ANOVA test, it would be suitable to use X-rays to penetrate tooth enamel to differentiate between females and males in forensic medicine.

1. There is no point doing a Tukey test with these data because there are only two groups being female and male so we will already see if there’s a significant difference between groups and their spectropenetration gradient through the ANOVA test.