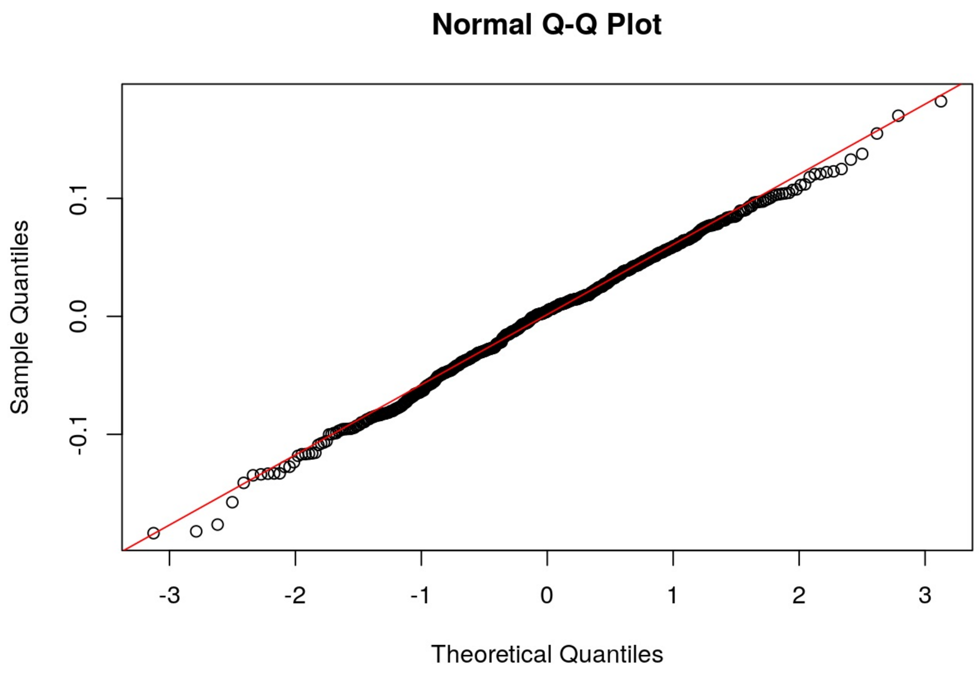
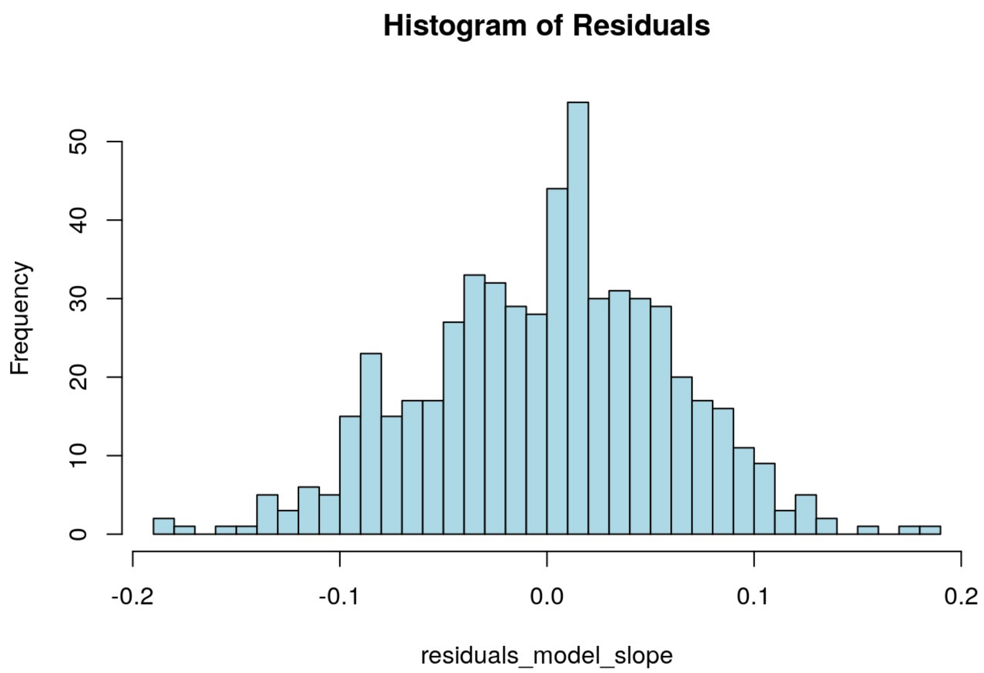
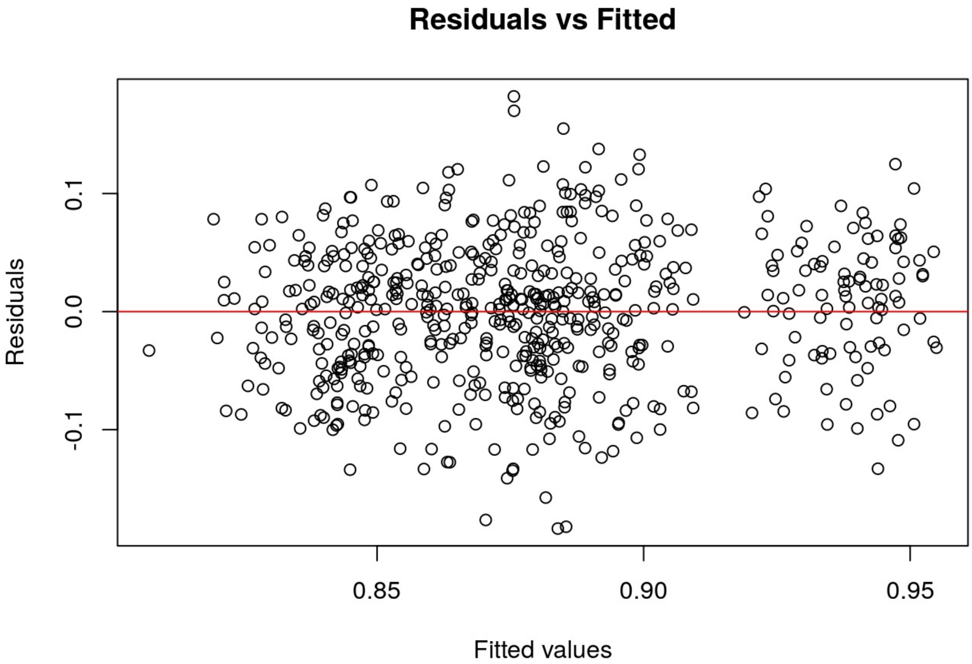


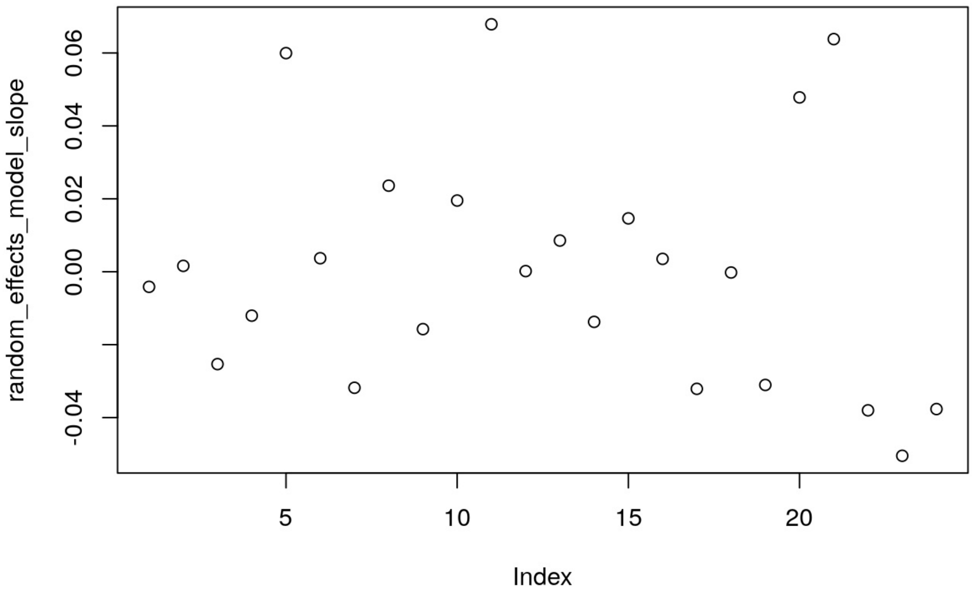

```
# I will create a histogram of residuals to visually inspect the residuals.

hist(residuals_model_slope, main = "Histogram of Residuals", breaks = 50, col = "lightblue", border = "black")
```



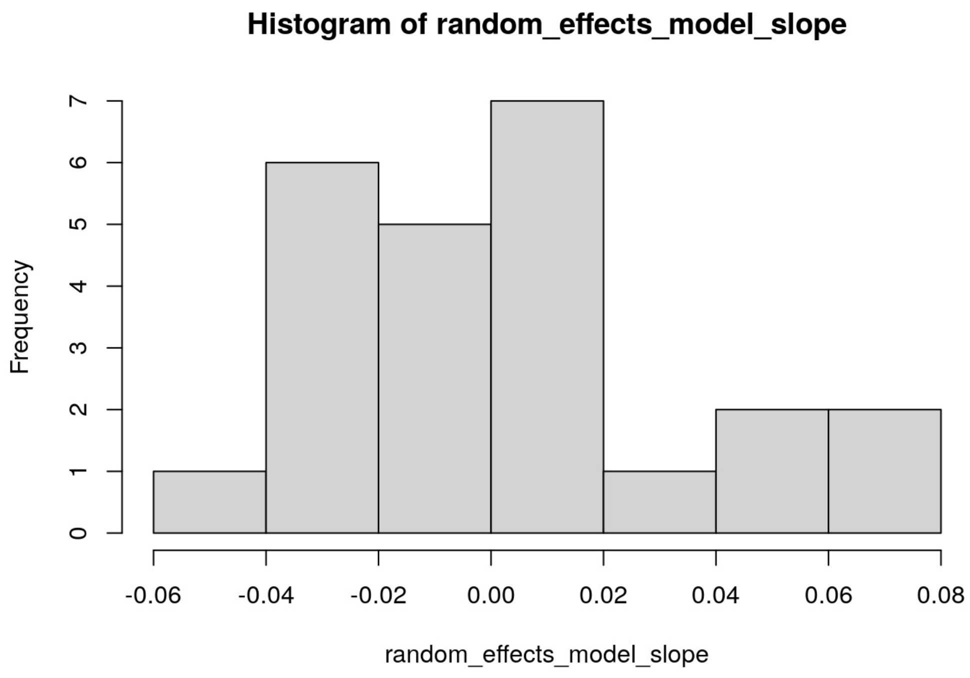

```
# Below is a residuals vs fitted values plot to check if the data is homoscedastic and the relationship between t
he dependent variables and the independent variables is linear. For these assumptions to be met, there must not b
e a specific shape in the plot (they should be scattered randomly).

plot(fitted(model_slope), resid(model_slope),
     main = "Residuals vs Fitted",
     xlab = "Fitted values",
     ylab = "Residuals")
abline(h = 0, col = "red")
```



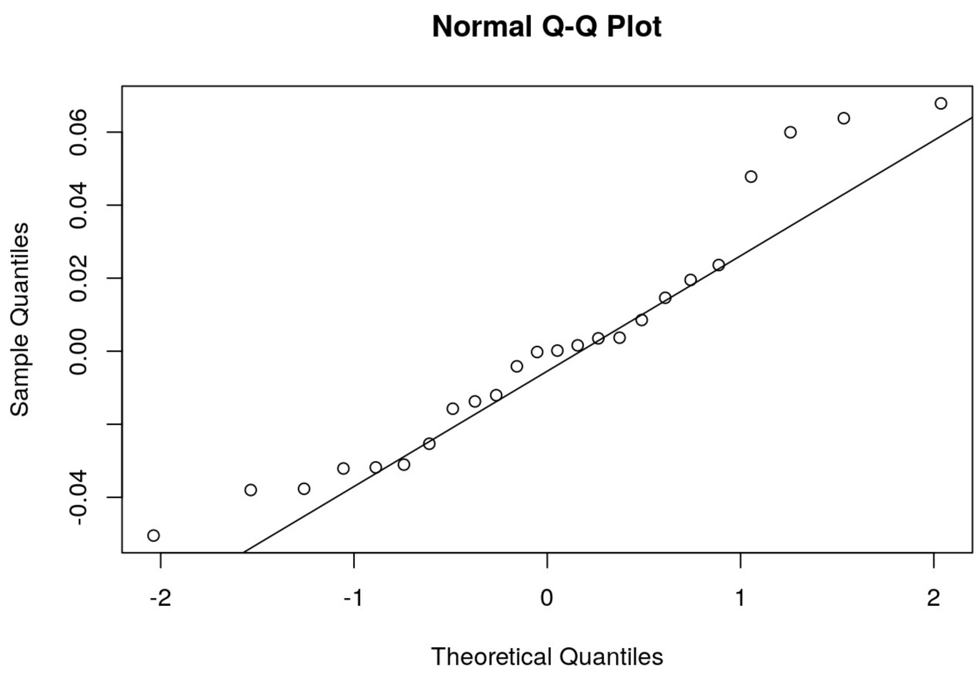

```
# One assumption is that the random intercepts have to be normal. Below I will get the random intercepts of my mo
del and plot them using a histogram and a QQ-plot to see if they are normally distributed visually. Then I will r
un a Shapiro-Wilk test to confirm the normality if the p-value is above the conventional threshold of 0.05. I wil
l also check that the mean of the random intercepts is 0 as this is another assumption.

random_effects_model_slope <- ranef(model_slope)$ID[, 1]
random_effects_model_slope <- as.numeric(random_effects_model_slope)

plot(random_effects_model_slope)
```




```
hist(random_effects_model_slope)
```




```
qqnorm(random_effects_model_slope)
qqline(random_effects_model_slope)
```




```
mean(random_effects_model_slope)
```

```
## [1] 0.0009317891
```

```
shapiro.test(random_effects_model_slope)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  random_effects_model_slope
## W = 0.93506, p-value = 0.1265
```

```
AIC(model_slope)
```

```
## [1] -1473.806
```

```
BIC(model_slope)
```

```
## [1] -1447.785
```

```
# This model also meets all the assumptions required for a mixed-effects model.

# This model fits the data better than the others, with an REML value of -1485.8, an AIC value of -1473.8 and a B
IC value of -1447.78.
```

```
# Above I have analysed the relationship between reaction time and the predictors, now I will attempt to do the s
ame with the accuracy of participants' responses as the outcome variable. I have a column that says whether the p
articipant recalled the headline correctly or not and this is stored as either "correct" or "incorrect". I want t
o transform this to 0 and 1, respectively, for my analysis.

df$accuracy_numeric <- ifelse(df$response == "correct", 1, 0)

# I will not check for normality as my accuracy column is binary.

# I am going to use Point-Biserial correlation, a specific type of Pearson correlation, to see if there is a corr
elation between accuracy and the predictor variables. This specific type of correlation is used when there is a b
inary categorical variable (accuracy) and continuous variables (predictors), as opposed to other forms of correla
tion.

cor.test(df$number_of_words, df$accuracy_numeric, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  df$number_of_words and df$accuracy_numeric
## t = -0.8289, df = 563, p-value = 0.4075
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.11706396  0.04771278
## sample estimates:
##         cor
## -0.03491286
```

```
cor.test(df$readability, df$accuracy_numeric, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  df$readability and df$accuracy_numeric
## t = 0.23868, df = 563, p-value = 0.8114
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.07248961  0.09247027
## sample estimates:
##        cor
## 0.01005877
```

```
cor.test(df$polarity, df$accuracy_numeric, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  df$polarity and df$accuracy_numeric
## t = -1.2222, df = 563, p-value = 0.2221
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.1333654  0.0311774
## sample estimates:
##         cor
## -0.05144313
```

```
cor.test(df$subjectivity, df$accuracy_numeric, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  df$subjectivity and df$accuracy_numeric
## t = -0.14839, df = 563, p-value = 0.8821
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.08869626  0.07627379
## sample estimates:
##          cor
## -0.006253788
```

```
# The results indicate that there is not a significant correlation between recall accuracy and word count, readab
ility, polarity and subjectivity, with p-values of 0.4075, 0.8114, 0.2221 and 0.8821, respectively. So, in this c
ase, there is not any correlation between accuracy and the predictor variables, although, in this case, ties may
affect the result. However, as my independent variables are not normally distributed, a vital assumption is viola
ted.
```

```
# I will resort to Kendall's correlation. However, it is important to mention that this correlation is not typica
lly used with binary variables. Kendall's correlation should manage to deal with the ties fairly well.

cor.test(df$number_of_words, df$accuracy_numeric, method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  df$number_of_words and df$accuracy_numeric
## z = 0.24208, p-value = 0.8087
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##         tau
## 0.008995349
```

```
cor.test(df$readability, df$accuracy_numeric, method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  df$readability and df$accuracy_numeric
## z = -0.519, p-value = 0.6038
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##        tau
## -0.0183596
```

```
cor.test(df$polarity, df$accuracy_numeric, method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  df$polarity and df$accuracy_numeric
## z = -1.3, p-value = 0.1936
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##         tau
## -0.04914103
```

```
cor.test(df$subjectivity, df$accuracy_numeric, method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  df$subjectivity and df$accuracy_numeric
## z = -0.22534, p-value = 0.8217
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##          tau
## -0.008389676
```