ICT as a Predictor of Reading Scores_Factor

Nirmal Ghimire, Watson Colelge of EDucation, UNCW

2025-05-22

```
library(tidyverse)
library(ggplot2)
library(janitor)
library(stringr)
library(readr)
library(jtools)
library(lavaan)
library(tidySEM)
library(lavaanPlot)
library(GGally)
library(dplyr)
library(ggstatsplot)
```

```
ICT_data <- read.csv("ict1_data.csv")
# summary(ICT_data)

# Load the file reading_2018_pvs_by_student.csv that contains all 10 plausible values
pvs_data <- read.csv("reading_2018_pvs_by_student.csv")

# combine the two data sets by the common variables CNTSCHID and CNTSTUID
ICT_data <- ICT_data |>
  left_join(pvs_data, by = c("CNTSCHID", "CNTSTUID"))
# Checking the data
head(ICT_data)
```

```
  X CNTSCHID CNTSTUID COM_HOM INTERNET ICTHOME LOC_INFO UNDERSTD EVAL_REF
1 1 84000001 84000250       1        2      10 536.6629 516.4273 517.8831
2 2 84000001 84000304       0        2       5 428.4643 406.6853 422.0078
3 3 84000001 84000353       1        2       9 510.6508 500.6199 505.6186
4 4 84000001 84000536       1        2       6 432.2721 429.1067 436.2995
5 5 84000001 84001240       1        2      12 508.1031 536.2947 562.0626
6 6 84000001 84001624       1        2      11 397.6511 385.6474 393.5362
    SINGLE MULTIPLE READ_SCR W_FSTUWT PV1READ PV2READ PV3READ PV4READ PV5READ
1 559.1167 556.0618 544.2085 646.6246 590.330 516.809 556.074 519.524 575.069
2 428.9934 422.2820 432.2518 629.8343 394.507 388.612 424.154 473.115 480.034
3 507.8295 502.7360 503.9496 613.7567 501.424 491.149 516.066 498.120 504.878
4 469.6456 444.9006 437.7777 613.7567 463.757 458.004 423.998 430.991 435.542
5 517.7481 536.7814 535.9487 613.7567 575.470 566.900 512.585 518.828 512.627
6 428.5294 431.2862 449.0047 613.7567 454.268 361.564 420.643 456.744 479.670
  PV6READ PV7READ PV8READ PV9READ PV10READ
1 526.085 554.709 560.873 495.694  546.918
2 413.273 456.736 429.004 410.248  452.835
3 531.993 497.834 499.279 505.181  493.572
4 435.865 424.260 429.249 415.182  460.929
5 551.874 528.946 510.276 560.892  521.089
6 500.484 384.440 461.417 448.013  522.804
```

```r
# Get rid of the first three columns
ICT_data <- ICT_data[, -(1:3)]

# Save this data as Rdata file
save(ICT_data, file = "ICT_data.Rdata")
```

```r
# Indexing the columns that need to be changed
cols_to_change <- c("COM_HOM", "INTERNET")
# Passing the selected column through the `modify_at` function from `purrr` package.
ICT_data <- ICT_data |>
  purrr::modify_at(cols_to_change, factor)

# Getting rid of x variable
# ICT_data <- ICT_data[,-(1:3)]
# Checking for the Summary
summary(ICT_data)
```

```
 COM_HOM     INTERNET      ICTHOME          LOC_INFO        UNDERSTD   
 0   : 604   0   : 182   Min.   : 1.000   Min.   :155.6   Min.   :191.7  
 1   :4170   1   : 139   1st Qu.: 8.000   1st Qu.:425.2   1st Qu.:417.4  
 NA's:  64   2   :4261   Median :10.000   Median :503.5   Median :501.2  
             NA's: 256   Mean   : 9.413   Mean   :496.9   Mean   :495.7  
                         3rd Qu.:11.000   3rd Qu.:574.2   3rd Qu.:575.5  
                         Max.   :12.000   Max.   :784.6   Max.   :814.5  
                         NA's   :192                                     
    EVAL_REF         SINGLE         MULTIPLE        READ_SCR    
 Min.   :185.4   Min.   :185.4   Min.   :193.5   Min.   :157.3  
 1st Qu.:424.9   1st Qu.:420.7   1st Qu.:423.6   1st Qu.:425.9  
 Median :509.7   Median :502.6   Median :503.4   Median :504.8  
 Mean   :505.0   Mean   :497.3   Mean   :500.0   Mean   :500.6  
 3rd Qu.:587.2   3rd Qu.:578.0   3rd Qu.:579.3   3rd Qu.:578.4  
 Max.   :795.4   Max.   :783.4   Max.   :784.5   Max.   :810.5  
                                                                
    W_FSTUWT         PV1READ         PV2READ         PV3READ     
 Min.   : 262.8   Min.   :161.3   Min.   :176.5   Min.   :132.4  
 1st Qu.: 563.0   1st Qu.:423.5   1st Qu.:424.6   1st Qu.:423.2  
 Median : 661.7   Median :503.6   Median :505.2   Median :503.8  
 Mean   : 735.6   Mean   :500.2   Mean   :500.8   Mean   :500.3  
 3rd Qu.: 854.5   3rd Qu.:578.7   3rd Qu.:578.4   3rd Qu.:577.9  
 Max.   :2946.1   Max.   :868.9   Max.   :898.5   Max.   :858.4  
                                                                 
    PV4READ         PV5READ         PV6READ         PV7READ     
 Min.   :140.3   Min.   :137.7   Min.   :128.1   Min.   :148.7  
 1st Qu.:426.1   1st Qu.:423.2   1st Qu.:424.4   1st Qu.:424.4  
 Median :503.1   Median :504.5   Median :504.4   Median :502.6  
 Mean   :501.1   Mean   :500.5   Mean   :501.0   Mean   :499.9  
 3rd Qu.:579.3   3rd Qu.:579.0   3rd Qu.:579.3   3rd Qu.:578.6  
 Max.   :834.1   Max.   :853.5   Max.   :844.8   Max.   :815.3  
                                                                
    PV8READ         PV9READ         PV10READ    
 Min.   :170.9   Min.   :173.6   Min.   :167.8  
 1st Qu.:424.9   1st Qu.:426.0   1st Qu.:426.1  
 Median :504.0   Median :503.9   Median :503.6  
 Mean   :500.3   Mean   :501.1   Mean   :500.6  
 3rd Qu.:579.4   3rd Qu.:577.1   3rd Qu.:579.0  
 Max.   :823.4   Max.   :818.1   Max.   :834.1  
```

```
# Dimension
dim(ICT_data)
```

```
[1] 4838   20
```

Missing Values

Based on the above summary, we have some missing values in our data set. We have to make decisions based on whether or not they are missing in a patter or anything. First of all I want to see how many of the missing values in `COM_HOM [64 missing; ~1.3%]`, `INTERNET [256 missing; ~5.3%]`, and `ICTHOME [192 missing; ~4%]` are missing in the same rows.

```
na_variables <- c("COM_HOM", "INTERNET", "ICTHOME")

only_na_int <- ICT_data |>
  filter(is.na(INTERNET))
summary(only_na_int)
```

```
 COM_HOM     INTERNET      ICTHOME          LOC_INFO        UNDERSTD
 0   : 38   0   :  0   Min.   : 1.000   Min.   :155.6   Min.   :191.7
 1   :171   1   :  0   1st Qu.: 2.000   1st Qu.:343.4   1st Qu.:348.5
 NA's: 47   2   :  0   Median : 6.500   Median :421.1   Median :429.8
            NA's:256   Mean   : 6.109   Mean   :416.5   Mean   :428.9
                       3rd Qu.: 9.000   3rd Qu.:492.5   3rd Qu.:506.5
                       Max.   :11.000   Max.   :698.8   Max.   :683.7
                       NA's   :192
    EVAL_REF          SINGLE          MULTIPLE         READ_SCR
 Min.   :185.7   Min.   :214.6   Min.   :200.9   Min.   :200.6
 1st Qu.:358.2   1st Qu.:340.2   1st Qu.:344.7   1st Qu.:350.7
 Median :428.5   Median :427.2   Median :424.5   Median :427.6
 Mean   :432.1   Mean   :425.8   Mean   :427.2   Mean   :428.3
 3rd Qu.:508.6   3rd Qu.:501.9   3rd Qu.:505.0   3rd Qu.:504.2
 Max.   :703.4   Max.   :684.7   Max.   :683.9   Max.   :673.9

    W_FSTUWT          PV1READ          PV2READ          PV3READ
 Min.   : 262.8   Min.   :219.5   Min.   :189.8   Min.   :190.8
 1st Qu.: 541.5   1st Qu.:350.3   1st Qu.:348.1   1st Qu.:349.1
 Median : 684.1   Median :423.0   Median :425.0   Median :427.4
 Mean   : 777.4   Mean   :422.8   Mean   :426.5   Mean   :429.6
 3rd Qu.: 970.7   3rd Qu.:500.3   3rd Qu.:508.5   3rd Qu.:505.9
 Max.   :1943.1   Max.   :701.5   Max.   :663.6   Max.   :688.6

    PV4READ          PV5READ          PV6READ          PV7READ
 Min.   :180.0   Min.   :174.5   Min.   :162.6   Min.   :148.7
 1st Qu.:355.5   1st Qu.:346.9   1st Qu.:353.7   1st Qu.:343.9
 Median :429.1   Median :432.3   Median :436.9   Median :424.7
 Mean   :430.4   Mean   :431.5   Mean   :429.9   Mean   :426.5
 3rd Qu.:496.0   3rd Qu.:503.7   3rd Qu.:498.9   3rd Qu.:501.1
 Max.   :770.7   Max.   :675.9   Max.   :716.6   Max.   :679.3

    PV8READ          PV9READ          PV10READ
 Min.   :205.3   Min.   :173.7   Min.   :200.9
 1st Qu.:350.7   1st Qu.:349.4   1st Qu.:352.7
 Median :431.7   Median :432.3   Median :425.1
 Mean   :430.2   Mean   :428.5   Mean   :427.3
 3rd Qu.:506.1   3rd Qu.:500.3   3rd Qu.:505.9
 Max.   :703.1   Max.   :734.4   Max.   :671.7
```

```
only_na_com <- ICT_data |>
  filter(is.na(COM_HOM))
summary(only_na_com)
```

```
 COM_HOM   INTERNET     ICTHOME          LOC_INFO        UNDERSTD    
 0   : 0   0   : 4   Min.   : 2.000   Min.   :195.2   Min.   :191.7  
 1   : 0   1   : 1   1st Qu.: 2.000   1st Qu.:297.0   1st Qu.:311.4  
 NA's:64   2   :12   Median : 7.000   Median :375.2   Median :384.9  
           NA's:47   Mean   : 6.609   Mean   :383.0   Mean   :400.5  
                     3rd Qu.:10.000   3rd Qu.:468.4   3rd Qu.:491.1  
                     Max.   :12.000   Max.   :598.0   Max.   :613.8  
                     NA's   :41                                      
    EVAL_REF         SINGLE          MULTIPLE         READ_SCR    
 Min.   :185.4   Min.   :214.6   Min.   :200.9   Min.   :200.6  
 1st Qu.:341.9   1st Qu.:313.0   1st Qu.:302.0   1st Qu.:309.2  
 Median :426.3   Median :393.5   Median :392.4   Median :407.0  
 Mean   :427.2   Mean   :403.4   Mean   :399.6   Mean   :404.2  
 3rd Qu.:522.2   3rd Qu.:497.2   3rd Qu.:500.4   3rd Qu.:493.8  
 Max.   :647.3   Max.   :636.5   Max.   :640.4   Max.   :623.5  
                                                                 
    W_FSTUWT         PV1READ          PV2READ          PV3READ     
 Min.   : 436.5   Min.   :224.4   Min.   :189.8   Min.   :192.1  
 1st Qu.: 603.9   1st Qu.:313.8   1st Qu.:308.0   1st Qu.:311.1  
 Median : 748.1   Median :401.2   Median :388.5   Median :400.9  
 Mean   : 837.2   Mean   :406.1   Mean   :402.3   Mean   :402.6  
 3rd Qu.:1088.6   3rd Qu.:488.9   3rd Qu.:503.5   3rd Qu.:491.8  
 Max.   :1477.1   Max.   :633.0   Max.   :612.2   Max.   :638.3  
                                                                  
    PV4READ          PV5READ          PV6READ          PV7READ     
 Min.   :217.3   Min.   :174.5   Min.   :196.9   Min.   :148.7  
 1st Qu.:318.7   1st Qu.:324.0   1st Qu.:321.1   1st Qu.:319.9  
 Median :404.9   Median :408.6   Median :392.8   Median :403.5  
 Mean   :410.9   Mean   :406.0   Mean   :408.0   Mean   :405.5  
 3rd Qu.:487.9   3rd Qu.:492.4   3rd Qu.:484.5   3rd Qu.:486.3  
 Max.   :624.6   Max.   :659.6   Max.   :654.1   Max.   :627.5  
                                                                 
    PV8READ          PV9READ          PV10READ    
 Min.   :205.3   Min.   :173.7   Min.   :200.9  
 1st Qu.:303.6   1st Qu.:314.3   1st Qu.:312.9  
 Median :394.7   Median :406.4   Median :397.2  
 Mean   :401.3   Mean   :400.3   Mean   :398.7  
 3rd Qu.:498.3   3rd Qu.:490.0   3rd Qu.:486.2  
 Max.   :634.6   Max.   :640.8   Max.   :668.8  
```

```
only_na_icth <- ICT_data |>
  filter(is.na(ICTHOME))
summary(only_na_icth)
```

```
 COM_HOM    INTERNET      ICTHOME       LOC_INFO        UNDERSTD    
 0   : 25   0   :  0   Min.   : NA   Min.   :155.6   Min.   :214.3  
 1   :126   1   :  0   1st Qu.: NA   1st Qu.:348.5   1st Qu.:359.8  
 NA's: 41   2   :  0   Median : NA   Median :426.2   Median :436.9  
            NA's:192   Mean   :NaN   Mean   :421.8   Mean   :436.4  
                       3rd Qu.: NA   3rd Qu.:489.6   3rd Qu.:508.4  
                       Max.   : NA   Max.   :674.4   Max.   :677.2  
                       NA's   :192                                  
    EVAL_REF         SINGLE          MULTIPLE         READ_SCR    
 Min.   :224.9   Min.   :240.2   Min.   :230.0   Min.   :249.4  
 1st Qu.:363.7   1st Qu.:348.7   1st Qu.:362.6   1st Qu.:367.1  
 Median :434.6   Median :435.2   Median :432.6   Median :433.5  
 Mean   :439.9   Mean   :433.0   Mean   :434.7   Mean   :436.2  
 3rd Qu.:512.8   3rd Qu.:503.6   3rd Qu.:506.0   3rd Qu.:504.9  
 Max.   :678.5   Max.   :684.7   Max.   :679.2   Max.   :673.9  
                                                                
    W_FSTUWT          PV1READ          PV2READ          PV3READ     
 Min.   : 308.8   Min.   :219.5   Min.   :204.6   Min.   :192.1  
 1st Qu.: 560.0   1st Qu.:353.7   1st Qu.:359.8   1st Qu.:369.1  
 Median : 734.3   Median :424.6   Median :435.5   Median :438.5  
 Mean   : 804.5   Mean   :429.1   Mean   :432.7   Mean   :438.4  
 3rd Qu.:1021.3   3rd Qu.:500.7   3rd Qu.:508.9   3rd Qu.:505.9  
 Max.   :1943.1   Max.   :701.5   Max.   :663.6   Max.   :688.6  
                                                                 
    PV4READ         PV5READ          PV6READ          PV7READ     
 Min.   :216.5   Min.   :253.2   Min.   :233.2   Min.   :208.1  
 1st Qu.:369.3   1st Qu.:364.2   1st Qu.:369.2   1st Qu.:355.2  
 Median :432.0   Median :445.6   Median :442.4   Median :431.7  
 Mean   :438.8   Mean   :440.4   Mean   :437.6   Mean   :435.0  
 3rd Qu.:498.0   3rd Qu.:509.4   3rd Qu.:498.9   3rd Qu.:501.8  
 Max.   :770.7   Max.   :675.9   Max.   :716.6   Max.   :677.0  
                                                                
    PV8READ         PV9READ          PV10READ    
 Min.   :240.4   Min.   :215.8   Min.   :234.5  
 1st Qu.:362.8   1st Qu.:363.1   1st Qu.:366.7  
 Median :444.0   Median :441.9   Median :438.9  
 Mean   :439.1   Mean   :436.2   Mean   :434.5  
 3rd Qu.:508.1   3rd Qu.:500.3   3rd Qu.:505.9  
 Max.   :703.1   Max.   :674.3   Max.   :661.3  
```

```r
# Total Missing
sum(is.na(ICT_data))
```

```
[1] 512
```

There are at least 41 cases where all three variables have *NAs*, and 192 cases where `ICTHOME` and `INTERNET` share *NAs*.

Beaujean (2014) mentions that “missing data pattern is the configuration of observed and missing values in a dataset. Data are missing completely at random (MCAR) when the missing values on a given variable are unrelated to both that variable and any other variable in the dataset. Data are missing at random (MAR) if a given variable’s missingness is unrelated to the variable itself, but is related to other variables in the dataset. In addition, data are not missing at random (NMAR) if the missing values are not MCAR or MAR. Values that are NMAR are a problem because they yield biased parameter estimates with traditional techniques and could yield biased results with modern techniques could yield biased results with modern techniques as well”(p. 114-117).

The book further writes, “if there are missing responses in a dataset, the best method to deal with them is a function of: (a) **the type of missing data**; (b) **how much data are missing**; and (c) **the variables that have missing values**.

- if the data are MCAR and only make up a small percentage (i.e., <3 - 5%) of the entire dataset, and the sample size is relatively large (*n* > 200), then list wise deletion will likely not have a noticeable influence on the parameter estimates;
- Missing data on an endogenous variable pose different problems from missing data on an exogenous variable. When the data are MAR, then observations with missing values on the endogenous variable, but have values for all the exogenous variables, do not contribute any information to the outcome-predictor relationship. The information they provide is still useful, but they do not make a contribution to the path coefficient.

Dealing with missing data: Historically, missing data were addressed indirectly by one of the two methods, e.g., Traditional Methods, and Modern Methods.

Traditional Methods:

Historically, missing data were addressed indirectly by one of the two methods. First, observations with missing data were deleted either listwise or pairwise.

a. *Listwise Deletion*: with listwise deletion, observations with any missing values on variables used in a model are deleted before estimating any parameters. It provides unbiased parameter estimates only if data are MCAR. However, the price for removing entire observations is that the estimates’ standard errors increase and statistical power decreases.

b. *Pairwise Deletion*: With pairwise deletion, the maximum amount of bivariate data available is retained for a single parameter estimate. This method is usually used to estimate means and covariances, with are then used for other analyses (e.g., regression, latent variable model).

c. *Imputation*: Imputation replaces a missing value wit another plausible value, typically based on one of three strategies,

    i. the mean/median of all present values of the variable;
    ii. regression-based predictor scores; or
    iii. use a pattern-matching technique to find another observation that has similar responses across all the other variables in the dataset, and then replace the missing value with the matched observation’s value (e.g., cold-deck imputation). Mean imputation is never a good idea with any type of missing data. Regression and deck methods both sound reasonable, but because they only impute asingle value for each missing response, they tend to produce understimated variability and standard error estimates.

Modern Methods:

a. *Full Information Maximum Likelihood*: ML estimation involves an iterative procedure to find parameter values that are the most likely for a model, given the variables’ observed values. To use a ML estimator, there has to be an a priori assumption about the distribution of the variable, which is usually that they follow a multivariate normal distribution.

b. *Multiple Imputation*: Multiple imputation creates a multiple da1617 16tasets, each of which contain different plausible estimates of the missing values. It involves a three-step process. The first step is to create the *m* datasets with imputed data. It is the most complex step and differs by the computer program and the types of the data, e.g., categorical, continuous, etc. The second step in MI is the analysis of different datasets, which involves estimating the model parameters in all of th em complete datasets, separately. The third step in MI is to pool the parameter estimates form the m datasets to calculate the final parameter estimates and their standard errors. The parameter estimates are simply the average of the m parameter estimates.
c. *Auxiliary Variables*: An auxiliary variable is a varible that is not of interest in answering a research question, but is included in the model because it is either potential cause or correlate of missingness in the variables of interest. The AVs can be used with both FIML and MI, but for them to work well they should be storngly correlated (r >= 0.50) with the manifest varibles of interest and with each other.

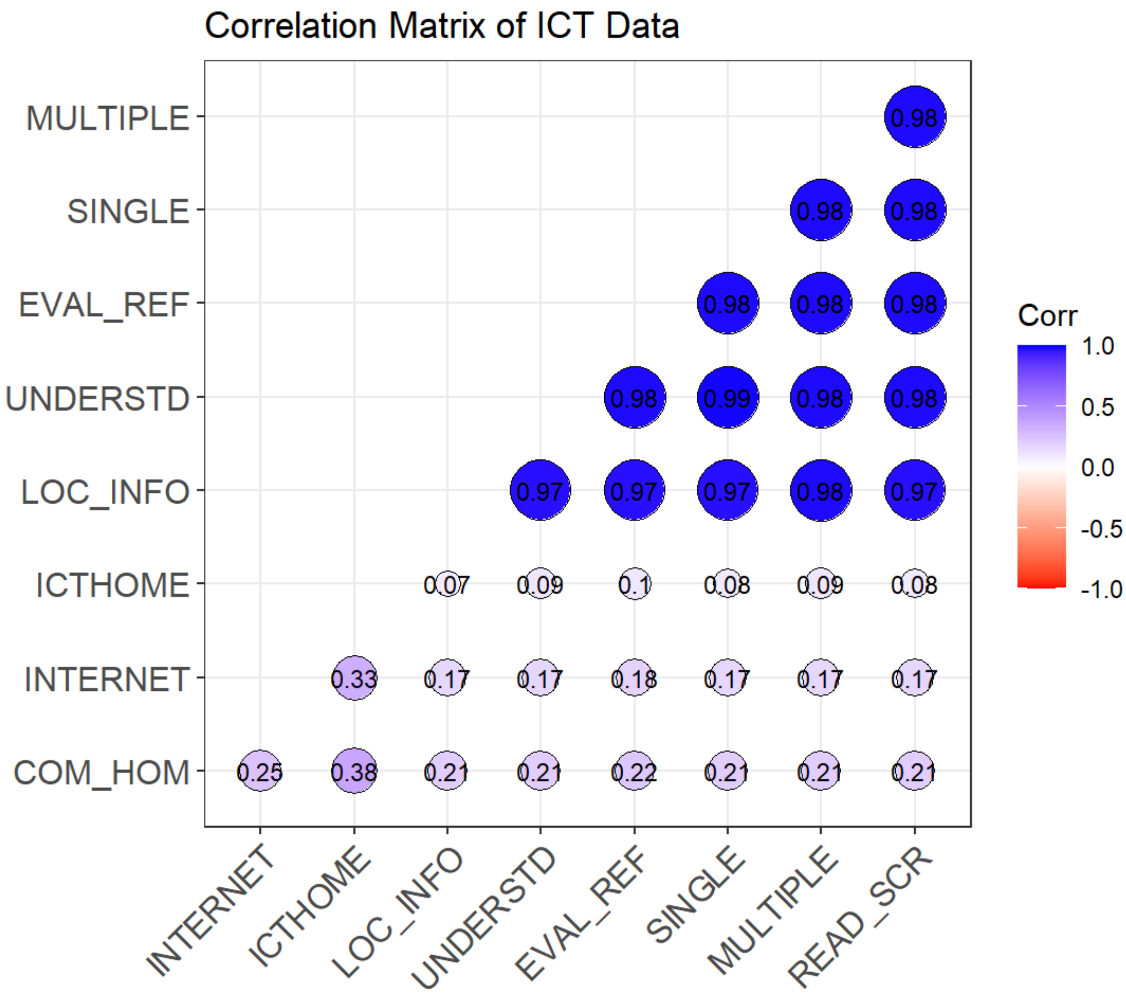
Before we take any action regarding deleting or imputing the missing data, I would like to check if the data are missing completely at random (MCAR)using multivariate test of MCAR proposed by Little (1988).

{r MCAR_test, out.width='100%'} misty::na.test(ICT_data)

library(mice) md.pattern(ICT_data) entire_data <- ICT_data
****

A missing value analysis indicated that Little's (1988) test of Missing Completely at Random (MCAR) was statistically significant, $\chi^2 = 662.75$, *DF* = 26, $p > .05$. When significant, Little's test suggests that the hypothesis the data are MCAR can be rejected. Whether a student has a computer at home is a key variable in the study. There are less than 1.5% of missing cases in this variable. It is not necessary for us to impute the missing data in this variable. Thus, we are going to get rid of missing data list wise.

```r
cor_data <- ICT_data

# Convert only factor or ordered variables to numeric (without changing original ICT_data)
cor_data$COM_HOM <- as.numeric(as.character(cor_data$COM_HOM))
cor_data$INTERNET <- as.numeric(as.character(cor_data$INTERNET))
cor_data$ICTHOME <- as.numeric(as.character(cor_data$ICTHOME))

# Correlation matrix lower triangle only
cor_matrix <- cor(cor_data[, c("COM_HOM", "INTERNET", "ICTHOME", "LOC_INFO", "UNDERSTD", "EVAL_R
EF", "SINGLE", "MULTIPLE", "READ_SCR")], use = "pairwise.complete.obs")
lower_tri <- cor_matrix
lower_tri[upper.tri(lower_tri)] <- NA
lower_tri[lower_tri == 1] <- NA
lower_tri[lower_tri == -1] <- NA
# display the lower triangle of the correlation matrix
library(ggcorrplot)
ggcorrplot(lower_tri,
  type = "lower",
  lab = TRUE,
  lab_size = 3,
  method = "circle",
  colors = c("red", "white", "blue"),
  title = "Correlation Matrix of ICT Data",
  ggtheme = theme_bw(),
  outline.col = "black",
  show.legend = TRUE,
  tl.srt = 45
)
```




```
## Variable Description
datawizard::describe_distribution(ICT_data)
```

```
Variable |   Mean |     SD |    IQR |             Range | Skewness | Kurtosis
-----------------------------------------------------------------------------
ICTHOME  |   9.41 |   2.16 |   3.00 |     [1.00, 12.00] |    -0.99 |     0.94
LOC_INFO | 496.89 | 102.77 | 149.03 |   [155.62, 784.56] |    -0.22 |    -0.40
UNDERSTD | 495.75 | 106.55 | 158.15 |   [191.67, 814.53] |    -0.13 |    -0.52
EVAL_REF | 504.97 | 110.50 | 162.43 |   [185.42, 795.43] |    -0.17 |    -0.53
SINGLE   | 497.33 | 108.28 | 157.39 |   [185.41, 783.42] |    -0.16 |    -0.51
MULTIPLE | 500.00 | 106.29 | 155.74 |   [193.49, 784.52] |    -0.14 |    -0.50
READ_SCR | 500.57 | 104.98 | 152.56 |   [157.30, 810.49] |    -0.14 |    -0.46
W_FSTUWT | 735.64 | 269.46 | 292.54 | [262.75, 2946.13] |     1.99 |     7.74
PV1READ  | 500.15 | 108.45 | 155.22 |   [161.34, 868.87] |    -0.10 |    -0.40
PV2READ  | 500.79 | 107.95 | 153.93 |   [176.46, 898.48] |    -0.10 |    -0.41
PV3READ  | 500.30 | 107.90 | 154.77 |   [132.42, 858.39] |    -0.10 |    -0.33
PV4READ  | 501.10 | 108.48 | 153.29 |   [140.29, 834.08] |    -0.10 |    -0.36
PV5READ  | 500.48 | 108.08 | 155.87 |   [137.74, 853.49] |    -0.11 |    -0.38
PV6READ  | 500.95 | 108.19 | 154.93 |   [128.11, 844.84] |    -0.10 |    -0.38
PV7READ  | 499.89 | 107.71 | 154.35 |   [148.74, 815.27] |    -0.11 |    -0.37
PV8READ  | 500.30 | 108.11 | 154.46 |   [170.91, 823.43] |    -0.11 |    -0.39
PV9READ  | 501.08 | 107.43 | 151.18 |   [173.64, 818.07] |    -0.09 |    -0.39
PV10READ | 500.63 | 107.93 | 153.02 |   [167.82, 834.09] |    -0.11 |    -0.38

Variable |    n | n_Missing
---------------------------
ICTHOME  | 4646 |       192
LOC_INFO | 4838 |         0
UNDERSTD | 4838 |         0
EVAL_REF | 4838 |         0
SINGLE   | 4838 |         0
MULTIPLE | 4838 |         0
READ_SCR | 4838 |         0
W_FSTUWT | 4838 |         0
PV1READ  | 4838 |         0
PV2READ  | 4838 |         0
PV3READ  | 4838 |         0
PV4READ  | 4838 |         0
PV5READ  | 4838 |         0
PV6READ  | 4838 |         0
PV7READ  | 4838 |         0
PV8READ  | 4838 |         0
PV9READ  | 4838 |         0
PV10READ | 4838 |         0
```

The variable `internet` has the Skeweness and Kurtosis values higher than the bearable limits. Thus, it's good to visualize the distribution and check. As this is a factor variable, the diagram shows that most of the 15-year-olds had access to internet at home compared to the and fairly small number of students noted to not have internet service at home, and not use it even when available at home.

```
# hist(ICT_data$INTERNET)
library(nlme)
null_m <- gls(READ_SCR ~ 1, data = ICT_data, method = "ML")
summary(null_m)
```

```
Generalized least squares fit by maximum likelihood
  Model: READ_SCR ~ 1
  Data: ICT_data
       AIC      BIC    logLik
  58762.09 58775.06 -29379.05

Coefficients:
               Value Std.Error  t-value p-value
(Intercept) 500.5677  1.509226 331.6718       0

Standardized residuals:
        Min          Q1         Med          Q3         Max
-3.27036287 -0.71117365  0.04073535  0.74167662  2.95261559

Residual standard error: 104.9645
Degrees of freedom: 4838 total; 4837 residual
```

```
fit1 <- lm(READ_SCR ~ COM_HOM, data = ICT_data)
summ(fit1)
```

Observations	4774 (64 missing obs. deleted)
Dependent variable	READ_SCR
Type	OLS linear regression

F(1,4772)	214.98
R²	0.04
Adj. R²	0.04

	Est.	S.E.	t val.	p
(Intercept)	444.96	4.15	107.17	0.00
COM_HOM1	65.14	4.44	14.66	0.00

Standard errors: OLS

```
# summ(fit1, robust = "HC1")
# summ(fit1, center = TRUE)
```

Model Fit 2

```
fit2 <- lm(READ_SCR ~ COM_HOM + INTERNET + COM_HOM * INTERNET, data = ICT_data)
summ(fit2)
```

Observations	4565 (273 missing obs. deleted)

Dependent variable	READ_SCR
Type	OLS linear regression

F(5,4559)	78.81
R^2	0.08
Adj. R^2	0.08

	Est.	S.E.	t val.	p
(Intercept)	442.05	10.31	42.88	0.00
COM_HOM1	20.67	14.92	1.39	0.17
INTERNET1	-30.52	19.72	-1.55	0.12
INTERNET2	10.04	11.35	0.88	0.38
COM_HOM1:INTERNET1	-37.61	24.51	-1.53	0.13
COM_HOM1:INTERNET2	44.65	15.74	2.84	0.00

Standard errors: OLS

```
# summ(fit2, center = TRUE)
car::Anova(fit2, type = "III")
```

```
Anova Table (Type III tests)

Response: READ_SCR
                   Sum Sq   Df   F value    Pr(>F)
(Intercept)      18172824    1 1838.5069 < 2.2e-16 ***
COM_HOM             18973    1    1.9195   0.16598
INTERNET            57165    2    2.8916   0.05559 .
COM_HOM:INTERNET   228460    2   11.5564 9.858e-06 ***
Residuals        45063690 4559
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(aov(fit2), ordered = TRUE)
```

```
  Tukey multiple comparisons of means
    95% family-wise confidence level
    factor levels have been ordered

Fit: aov(formula = fit2)

$COM_HOM
       diff      lwr      upr p adj
1-0 65.1498 56.39637 73.90324     0

$INTERNET
         diff      lwr       upr     p adj
0-1  70.55492 44.11763  96.99221 0.0000000
2-1 101.98253 81.82098 122.14407 0.0000000
2-0  31.42761 13.59461  49.26060 0.0001086

$`COM_HOM:INTERNET`
             diff        lwr       upr     p adj
0:1-1:1  16.94163 -38.514631  72.39790 0.9534272
0:0-1:1  47.46603   6.921556  88.01050 0.0110111
0:2-1:1  57.50176  26.462839  88.54069 0.0000020
1:0-1:1  68.13548  26.600445 109.67051 0.0000441
1:2-1:1 122.82020  94.516971 151.12344 0.0000000
0:0-0:1  30.52439 -25.682968  86.73176 0.6328555
0:2-0:1  40.56013  -9.227701  90.34796 0.1850953
1:0-0:1  51.19384  -5.732180 108.11987 0.1065013
1:2-0:1 105.87857  57.748680 154.00846 0.0000000
0:2-0:0  10.03574 -22.326048  42.39752 0.9504078
1:0-0:0  20.66945 -21.863234  63.20213 0.7358957
1:2-0:0  75.35417  45.606181 105.10217 0.0000000
1:0-0:2  10.63371 -22.960776  44.22820 0.9459925
1:2-0:2  65.31844  51.017961  79.61892 0.0000000
1:2-1:0  54.68472  23.600199  85.76925 0.0000082
```

Model Fit 3

```
fit3 <- lm(READ_SCR ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET * ICTHOME, data = ICT_da
ta)
# summ(fit3, robust = "HC1")
# summ(fit3, center = TRUE)
# summ(fit3)
car::Anova(fit3, type = "III")
```

```
Anova Table (Type III tests)

Response: READ_SCR
                          Sum Sq   Df  F value    Pr(>F)    
(Intercept)              2506725    1 254.5147 < 2.2e-16 ***
COM_HOM                    54613    1   5.5450 0.0185762 *  
INTERNET                  140952    2   7.1556 0.0007893 ***
ICTHOME                    79182    1   8.0395 0.0045969 ** 
COM_HOM:INTERNET           22830    2   1.1590 0.3138878    
COM_HOM:ICTHOME            49643    1   5.0404 0.0248112 *  
INTERNET:ICTHOME          130096    2   6.6045 0.0013673 ** 
COM_HOM:INTERNET:ICTHOME   61664    2   3.1305 0.0437920 *  
Residuals               44842682 4553                       
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# TukeyHSD(aov(fit3), ordered = TRUE)
```

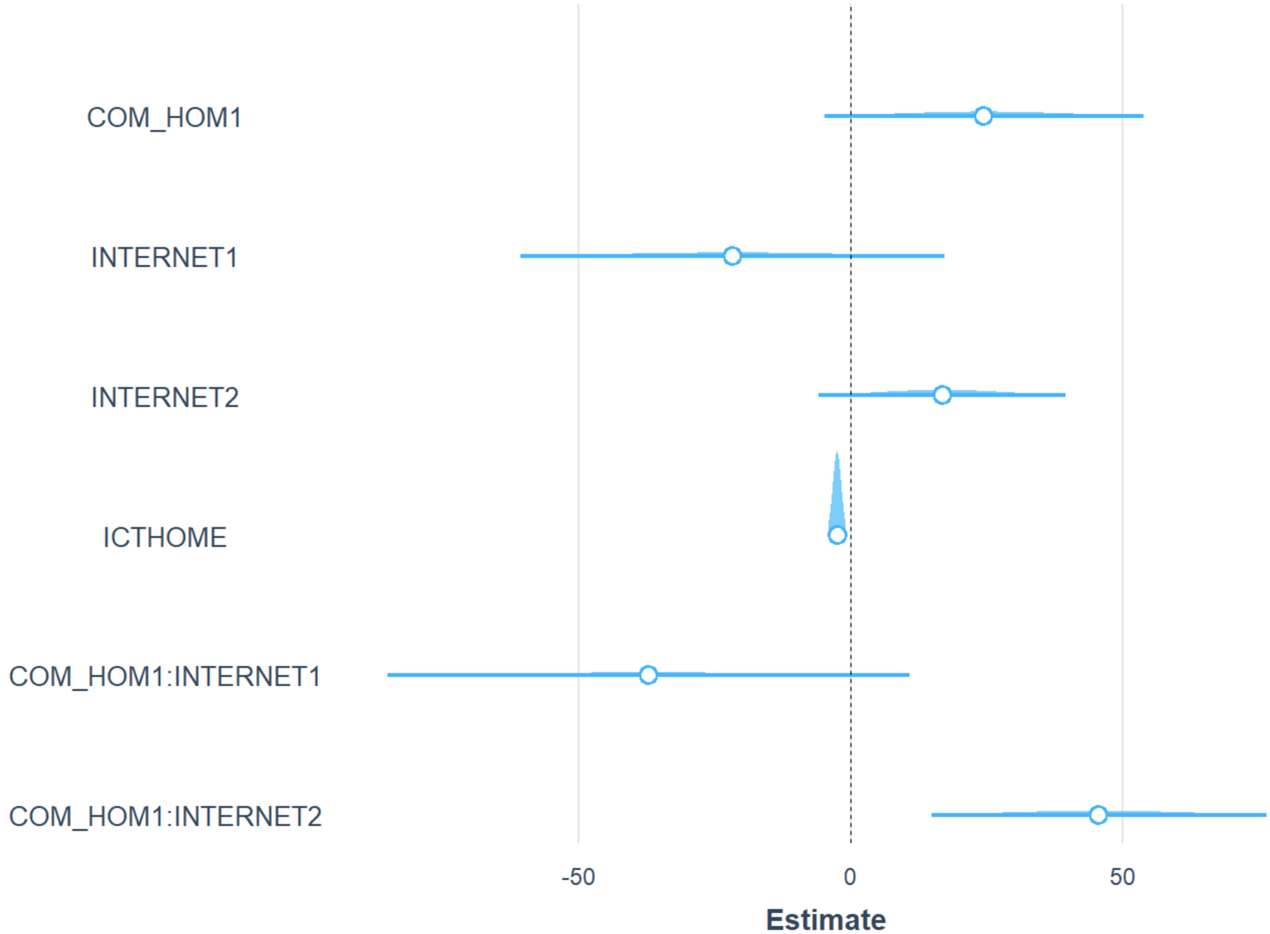
Model Fit 4

```
fit4 <- lm(READ_SCR ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET, data = ICT_data)
# summ(fit4)
car::Anova(fit4, type = "III")
```

```
Anova Table (Type III tests)

Response: READ_SCR
                   Sum Sq   Df   F value    Pr(>F)    
(Intercept)      16714871    1 1694.0230 < 2.2e-16 ***
COM_HOM             26631    1    2.6991  0.100477    
INTERNET            63281    2    3.2067  0.040581 *  
ICTHOME             90048    1    9.1262  0.002534 ** 
COM_HOM:INTERNET   233678    2   11.8414 7.424e-06 ***
Residuals        44973641 4558                        
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# TukeyHSD(aov(fit4), ordered = TRUE)
```

```
# effect_plot(fit3, pred = ICTHOME, interval = TRUE, plot.points = TRUE, jitter = 0.05)
# effect_plot(fit2, pred = INTERNET, interval = TRUE, plot.points = TRUE, jitter = 0.05)
# effect_plot(fit1, pred = COM_HOM, interval = TRUE, plot.points = TRUE, jitter = 0.05)
# plot_summs(fit3)
plot_summs(fit4, plot.distributions = TRUE, inner_ci_levels = .95)
```




```
# plot_summs(fit1, fit2, fit3)
export_summs(null_m, fit1, fit2, fit3,
  scale = FALSE,
  error_format = "[{conf.low}, {conf.high}]"
)
```

	Model 1	Model 2	Model 3	Model 4
(Intercept)	500.57 ***	444.96 ***	442.05 ***	380.95 ***
	[497.61, 503.53]	[436.82, 453.10]	[421.84, 462.26]	[334.14, 427.77]
COM_HOM1		65.14 ***	20.67	92.85 *
		[56.43, 73.85]	[-8.58, 49.92]	[15.55, 170.15]
INTERNET1			-30.52	69.30
			[-69.18, 8.13]	[-50.55, 189.16]

INTERNET2			10.04	111.09 ***
			[-12.22, 32.29]	[53.51, 168.68]
COM_HOM1:INTERNET1			-37.61	-124.20
			[-85.67, 10.45]	[-289.44, 41.03]
COM_HOM1:INTERNET2			44.65 **	-42.46
			[13.79, 75.51]	[-128.64, 43.73]
ICTHOME				12.41 **
				[3.83, 20.98]
COM_HOM1:ICTHOME				-14.10 *
				[-26.41, -1.79]
INTERNET1:ICTHOME				-16.91 *
				[-31.86, -1.96]
INTERNET2:ICTHOME				-17.55 ***
				[-27.07, -8.02]
COM_HOM1:INTERNET1:ICTHOME				16.27
				[-3.28, 35.83]
COM_HOM1:INTERNET2:ICTHOME				16.69 *
				[3.57, 29.81]
N	4838	4774	4565	4565
R2		0.04	0.08	0.08

*** p < 0.001; ** p < 0.01; * p < 0.05.

Effect Size (cohen’s d) for Fit Model 3

```
effectsize::omega_squared(fit2, alternative = "greater", verbose = TRUE, partial = TRUE, ci = 0.
95)
```

```
# Effect Size for ANOVA (Type I)

Parameter         | Omega2 (partial) |       95% CI
---------------------------------------------------
COM_HOM           |             0.04 | [0.04, 1.00]
INTERNET          |             0.03 | [0.02, 1.00]
COM_HOM:INTERNET  |         4.60e-03 | [0.00, 1.00]

- One-sided CIs: upper bound fixed at [1.00].
```

```
effectsize::omega_squared(fit3, alternative = "greater", verbose = TRUE, partial = TRUE, ci = 0.
95)
```

```
# Effect Size for ANOVA (Type I)

Parameter                 | Omega2 (partial) |       95% CI
-----------------------------------------------------------
COM_HOM                   |             0.04 | [0.04, 1.00]
INTERNET                  |             0.03 | [0.02, 1.00]
ICTHOME                   |         1.66e-03 | [0.00, 1.00]
COM_HOM:INTERNET          |         4.74e-03 | [0.00, 1.00]
COM_HOM:ICTHOME           |             0.00 | [0.00, 1.00]
INTERNET:ICTHOME          |         1.09e-03 | [0.00, 1.00]
COM_HOM:INTERNET:ICTHOME  |         9.33e-04 | [0.00, 1.00]

- One-sided CIs: upper bound fixed at [1.00].
```

```
effectsize::cohens_f(fit3, alternative = "greater", verbose = TRUE, partial = TRUE, ci = 0.95)
```

```
# Effect Size for ANOVA (Type I)

Parameter                 | Cohen's f (partial) |      95% CI
-------------------------------------------------------------
COM_HOM                   |                0.22 | [0.19, Inf]
INTERNET                  |                0.19 | [0.16, Inf]
ICTHOME                   |                0.04 | [0.02, Inf]
COM_HOM:INTERNET          |                0.07 | [0.05, Inf]
COM_HOM:ICTHOME           |            3.36e-03 | [0.00, Inf]
INTERNET:ICTHOME          |                0.04 | [0.01, Inf]
COM_HOM:INTERNET:ICTHOME  |                0.04 | [0.00, Inf]

- One-sided CIs: upper bound fixed at [Inf].
```

Cohen's-*f* rule of thumb (Cohen, 1988, p. 285-287) for multiple regression:

- f <= 0.14 : Small Effect
- f <= 0.39 : Medium Effect
- f >= 0.59 : Large Effect

Omega Squared rule of thumb:

- ω2 >= .01 : Small Effect
- ω2 >= .06 : Medium Effect
- ω2 >= .14 : Large Effect

```
names(ICT_data)
```

```
 [1] "COM_HOM"  "INTERNET" "ICTHOME"  "LOC_INFO" "UNDERSTD" "EVAL_REF"
 [7] "SINGLE"   "MULTIPLE" "READ_SCR" "W_FSTUWT" "PV1READ"  "PV2READ"
[13] "PV3READ"  "PV4READ"  "PV5READ"  "PV6READ"  "PV7READ"  "PV8READ"
[19] "PV9READ"  "PV10READ"
```

```
dim(ICT_data)
```

```
[1] 4838   20
```

```
entire_data <- ICT_data
```

```
entire_data[, c("COM_HOM", "INTERNET")] <- lapply(entire_data[, c("COM_HOM", "INTERNET")], order
ed)
str(entire_data)
```

```
'data.frame':	4838 obs. of  20 variables:
 $ COM_HOM : Ord.factor w/ 2 levels "0"<"1": 2 1 2 2 2 2 2 2 2 2 ...
 $ INTERNET: Ord.factor w/ 3 levels "0"<"1"<"2": 3 3 3 3 3 3 3 3 3 3 ...
 $ ICTHOME : int  10 5 9 6 12 11 9 10 6 9 ...
 $ LOC_INFO: num  537 428 511 432 508 ...
 $ UNDERSTD: num  516 407 501 429 536 ...
 $ EVAL_REF: num  518 422 506 436 562 ...
 $ SINGLE  : num  559 429 508 470 518 ...
 $ MULTIPLE: num  556 422 503 445 537 ...
 $ READ_SCR: num  544 432 504 438 536 ...
 $ W_FSTUWT: num  647 630 614 614 614 ...
 $ PV1READ : num  590 395 501 464 575 ...
 $ PV2READ : num  517 389 491 458 567 ...
 $ PV3READ : num  556 424 516 424 513 ...
 $ PV4READ : num  520 473 498 431 519 ...
 $ PV5READ : num  575 480 505 436 513 ...
 $ PV6READ : num  526 413 532 436 552 ...
 $ PV7READ : num  555 457 498 424 529 ...
 $ PV8READ : num  561 429 499 429 510 ...
 $ PV9READ : num  496 410 505 415 561 ...
 $ PV10READ: num  547 453 494 461 521 ...
```

```
colSums(is.na(entire_data))
```

```
 COM_HOM INTERNET  ICTHOME LOC_INFO UNDERSTD EVAL_REF   SINGLE MULTIPLE
      64      256      192        0        0        0        0        0
READ_SCR W_FSTUWT  PV1READ  PV2READ  PV3READ  PV4READ  PV5READ  PV6READ
       0        0        0        0        0        0        0        0
 PV7READ  PV8READ  PV9READ PV10READ
       0        0        0        0
```

```
library(reshape2)

long_data <- melt(entire_data, id.vars = c("COM_HOM", "INTERNET", "ICTHOME"))

long_data <- long_data |>
  rename(
    test_type = variable,
    scores = value
  )

str(long_data)
```

```
'data.frame':   82246 obs. of  5 variables:
 $ COM_HOM  : Ord.factor w/ 2 levels "0"<"1": 2 1 2 2 2 2 2 2 2 2 ...
 $ INTERNET : Ord.factor w/ 3 levels "0"<"1"<"2": 3 3 3 3 3 3 3 3 3 3 ...
 $ ICTHOME  : int  10 5 9 6 12 11 9 10 6 9 ...
 $ test_type: Factor w/ 17 levels "LOC_INFO","UNDERSTD",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ scores   : num  537 428 511 432 508 ...
```

```
head(long_data)
```

```
  COM_HOM INTERNET ICTHOME test_type   scores
1       1        2      10  LOC_INFO 536.6629
2       0        2       5  LOC_INFO 428.4643
3       1        2       9  LOC_INFO 510.6508
4       1        2       6  LOC_INFO 432.2721
5       1        2      12  LOC_INFO 508.1031
6       1        2      11  LOC_INFO 397.6511
```

```
summary(entire_data)
```

```
 COM_HOM     INTERNET      ICTHOME          LOC_INFO        UNDERSTD    
 0   : 604   0   : 182   Min.   : 1.000   Min.   :155.6   Min.   :191.7  
 1   :4170   1   : 139   1st Qu.: 8.000   1st Qu.:425.2   1st Qu.:417.4  
 NA's:  64   2   :4261   Median :10.000   Median :503.5   Median :501.2  
             NA's: 256   Mean   : 9.413   Mean   :496.9   Mean   :495.7  
                         3rd Qu.:11.000   3rd Qu.:574.2   3rd Qu.:575.5  
                         Max.   :12.000   Max.   :784.6   Max.   :814.5  
                         NA's   :192                                     
    EVAL_REF         SINGLE         MULTIPLE        READ_SCR    
 Min.   :185.4   Min.   :185.4   Min.   :193.5   Min.   :157.3  
 1st Qu.:424.9   1st Qu.:420.7   1st Qu.:423.6   1st Qu.:425.9  
 Median :509.7   Median :502.6   Median :503.4   Median :504.8  
 Mean   :505.0   Mean   :497.3   Mean   :500.0   Mean   :500.6  
 3rd Qu.:587.2   3rd Qu.:578.0   3rd Qu.:579.3   3rd Qu.:578.4  
 Max.   :795.4   Max.   :783.4   Max.   :784.5   Max.   :810.5  
                                                                
    W_FSTUWT         PV1READ         PV2READ         PV3READ     
 Min.   : 262.8   Min.   :161.3   Min.   :176.5   Min.   :132.4  
 1st Qu.: 563.0   1st Qu.:423.5   1st Qu.:424.6   1st Qu.:423.2  
 Median : 661.7   Median :503.6   Median :505.2   Median :503.8  
 Mean   : 735.6   Mean   :500.2   Mean   :500.8   Mean   :500.3  
 3rd Qu.: 854.5   3rd Qu.:578.7   3rd Qu.:578.4   3rd Qu.:577.9  
 Max.   :2946.1   Max.   :868.9   Max.   :898.5   Max.   :858.4  
                                                                 
    PV4READ         PV5READ         PV6READ         PV7READ     
 Min.   :140.3   Min.   :137.7   Min.   :128.1   Min.   :148.7  
 1st Qu.:426.1   1st Qu.:423.2   1st Qu.:424.4   1st Qu.:424.4  
 Median :503.1   Median :504.5   Median :504.4   Median :502.6  
 Mean   :501.1   Mean   :500.5   Mean   :501.0   Mean   :499.9  
 3rd Qu.:579.3   3rd Qu.:579.0   3rd Qu.:579.3   3rd Qu.:578.6  
 Max.   :834.1   Max.   :853.5   Max.   :844.8   Max.   :815.3  
                                                                
    PV8READ         PV9READ         PV10READ    
 Min.   :170.9   Min.   :173.6   Min.   :167.8  
 1st Qu.:424.9   1st Qu.:426.0   1st Qu.:426.1  
 Median :504.0   Median :503.9   Median :503.6  
 Mean   :500.3   Mean   :501.1   Mean   :500.6  
 3rd Qu.:579.4   3rd Qu.:577.1   3rd Qu.:579.0  
 Max.   :823.4   Max.   :818.1   Max.   :834.1  
```

```
xtabs(~ COM_HOM + INTERNET, data = entire_data)
```

```
       INTERNET
COM_HOM    0    1    2
      0   93   35  438
      1   85  103 3811
```

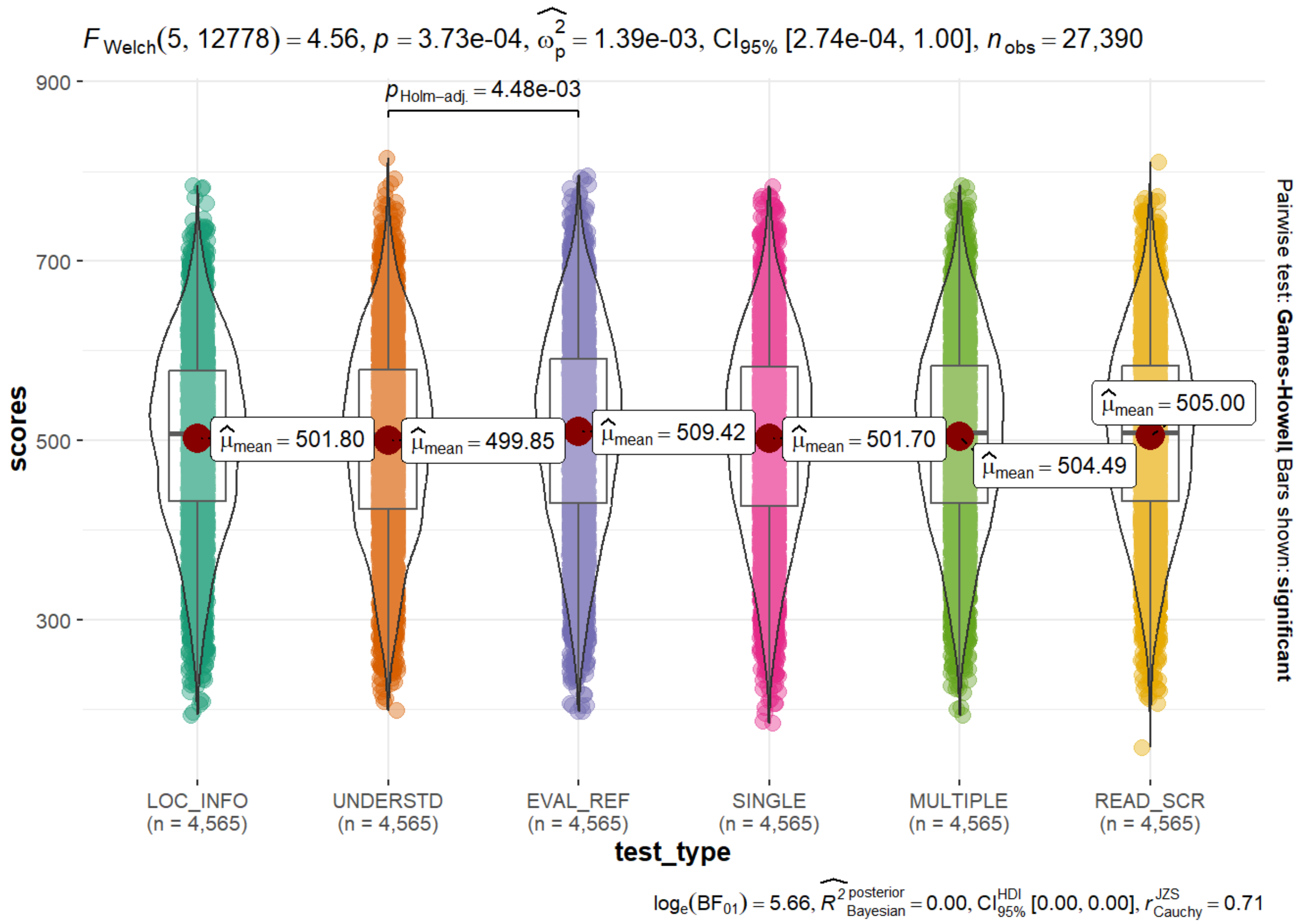
```
xtabs(~ COM_HOM + ICTHOME, data = entire_data)
```

```
       ICTHOME
COM_HOM   1   2   3   4   5   6   7   8   9  10  11  12
      0  12  10  18  35  73  77  87  95  46  47  44  35
      1   8   6   4  29  48 139 286 402 657 849 864 752
```

```
xtabs(~ INTERNET + ICTHOME, data = entire_data)
```

```
        ICTHOME
INTERNET   1   2   3   4   5   6   7   8   9  10  11  12
       0  14  10  14  20  29  24  27  22   7  13   2   0
       1   0   0   3   3   4   5   9   8  15  15  20  57
       2   0   1   3  38  89 181 332 464 674 866 882 731
```

```
options(digits = 3)
# Filter only first 6 levels of test_type
subset_data <- long_data %>%
  filter(test_type %in% levels(test_type)[1:6]) %>%
  na.omit()

# Plot
ggbetweenstats(
  data = subset_data,
  x = test_type,
  y = scores,
  type = "parametric", # use Welch's ANOVA
  pairwise.comparisons = TRUE,
  pairwise.display = "significant", # only show sig comparisons
  p.adjust.method = "holm",
  messages = FALSE
)
```




```
# Mean and SD of each test type
com_home_score <- entire_data |>
  na.omit() |>
  select(COM_HOM, ICTHOME, LOC_INFO, UNDERSTD, EVAL_REF, SINGLE, MULTIPLE, READ_SCR) |>
  group_by(COM_HOM) |>
  summarize(
    mean_ICTHOME = mean(ICTHOME),
    sd_ICTHOME = sd(ICTHOME),
    mean_LOC_INFO = mean(LOC_INFO),
    sd_LOC_INFO = sd(LOC_INFO),
    mean_UNDERSTD = mean(UNDERSTD),
    sd_UNDERSTD = sd(UNDERSTD),
    mean_EVAL_REF = mean(EVAL_REF),
    sd_EVAL_REF = sd(EVAL_REF),
    mean_SINGLE = mean(SINGLE),
    sd_SINGLE = sd(SINGLE),
    mean_MULTIPLE = mean(MULTIPLE),
    sd_MULTIPLE = sd(MULTIPLE),
    mean_READ_SCR = mean(READ_SCR),
    sd_READ_SCR = sd(READ_SCR)
  ) |>
  t()
com_home_score
```

```
              [,1]    [,2]
COM_HOM       "0"     "1"
mean_ICTHOME  "7.36"  "9.76"
sd_ICTHOME    "2.54"  "1.83"
mean_LOC_INFO "447"   "510"
sd_LOC_INFO   "92.6"  "99.5"
mean_UNDERSTD "441"   "508"
sd_UNDERSTD   " 95"   "104"
mean_EVAL_REF "445"   "519"
sd_EVAL_REF   "101"   "107"
mean_SINGLE   "443"   "510"
sd_SINGLE     " 97.5" "105.7"
mean_MULTIPLE "445"   "513"
sd_MULTIPLE   " 95.7" "103.3"
mean_READ_SCR "448"   "513"
sd_READ_SCR   " 94.8" "102.2"
```

```
# Mean and SD of each test type
internet_home_score <- entire_data |>
  na.omit() |>
  select(INTERNET, ICTHOME, LOC_INFO, UNDERSTD, EVAL_REF, SINGLE, MULTIPLE, READ_SCR) |>
  group_by(INTERNET) |>
  summarize(
    mean_ICTHOME = mean(ICTHOME),
    sd_ICTHOME = sd(ICTHOME),
    mean_LOC_INFO = mean(LOC_INFO),
    sd_LOC_INFO = sd(LOC_INFO),
    mean_UNDERSTD = mean(UNDERSTD),
    sd_UNDERSTD = sd(UNDERSTD),
    mean_EVAL_REF = mean(EVAL_REF),
    sd_EVAL_REF = sd(EVAL_REF),
    mean_SINGLE = mean(SINGLE),
    sd_SINGLE = sd(SINGLE),
    mean_MULTIPLE = mean(MULTIPLE),
    sd_MULTIPLE = sd(MULTIPLE),
    mean_READ_SCR = mean(READ_SCR),
    sd_READ_SCR = sd(READ_SCR)
  ) |>
  t()
internet_home_score
```

```
              [,1]    [,2]    [,3]   
INTERNET      "0"     "1"     "2"    
mean_ICTHOME  "5.70"  "9.97"  "9.61" 
sd_ICTHOME    "2.51"  "2.43"  "1.90" 
mean_LOC_INFO "451"   "398"   "507"  
sd_LOC_INFO   "93.9"  "93.3"  "98.9" 
mean_UNDERSTD "448"   "391"   "506"  
sd_UNDERSTD   " 98.9" " 95.5" "103.5"
mean_EVAL_REF "451"   "393"   "516"  
sd_EVAL_REF   "104.7" " 98.9" "106.9"
mean_SINGLE   "450"   "389"   "508"  
sd_SINGLE     "101.1" " 98.4" "104.9"
mean_MULTIPLE "450"   "398"   "510"  
sd_MULTIPLE   " 98.7" " 93.8" "102.8"
mean_READ_SCR "452"   "399"   "511"  
sd_READ_SCR   "100.0" " 92.3" "101.6"
```

```
# Mean and SD of each test type
ict_home_score <- entire_data |>
  na.omit() |>
  select(ICTHOME, LOC_INFO, UNDERSTD, EVAL_REF, SINGLE, MULTIPLE, READ_SCR) |>
  group_by(ICTHOME) |>
  summarize(
    mean_LOC_INFO = mean(LOC_INFO),
    sd_LOC_INFO = sd(LOC_INFO),
    mean_UNDERSTD = mean(UNDERSTD),
    sd_UNDERSTD = sd(UNDERSTD),
    mean_EVAL_REF = mean(EVAL_REF),
    sd_EVAL_REF = sd(EVAL_REF),
    mean_SINGLE = mean(SINGLE),
    sd_SINGLE = sd(SINGLE),
    mean_MULTIPLE = mean(MULTIPLE),
    sd_MULTIPLE = sd(MULTIPLE),
    mean_READ_SCR = mean(READ_SCR),
    sd_READ_SCR = sd(READ_SCR)
  ) |>
  t()
ict_home_score
```

```
               [,1] [,2]  [,3] [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11]
ICTHOME         1.0    2   3.0    4   5.0   6.0   7.0   8.0   9.0  10.0    11
mean_LOC_INFO 362.2  402 445.0  440 462.8 485.3 495.7 506.5 527.2 522.2   507
sd_LOC_INFO    73.9  110  68.8  102  91.5  89.1  91.6  94.4  95.0  98.8   100
mean_UNDERSTD 346.9  397 448.1  433 457.8 476.9 490.6 502.2 524.7 520.2   508
sd_UNDERSTD    82.3  122  68.4  103  93.7  93.0  95.6  97.6  98.6 104.3   107
mean_EVAL_REF 346.1  408 449.4  437 458.1 483.2 498.9 511.9 535.3 532.1   518
sd_EVAL_REF    86.3  119  70.4  108 100.1 100.0 102.3 102.4 101.5 106.9   110
mean_SINGLE   351.9  400 446.8  434 458.1 481.6 493.8 506.5 527.2 523.0   509
sd_SINGLE      86.3  130  67.8  105  95.1  95.1  97.6 100.5  99.8 105.3   108
mean_MULTIPLE 349.8  396 449.5  433 459.9 482.4 495.9 508.1 528.4 525.6   512
sd_MULTIPLE    87.6  120  68.0  101  92.2  92.7  96.2  98.0  97.9 103.0   106
mean_READ_SCR 345.4  400 446.1  439 463.7 484.5 497.0 509.0 530.4 525.7   511
sd_READ_SCR    96.8  121  74.9  102  91.7  92.6  95.5  97.6  96.4 101.7   105
              [,12]
ICTHOME          12
mean_LOC_INFO   471
sd_LOC_INFO     107
mean_UNDERSTD   471
sd_UNDERSTD     110
mean_EVAL_REF   480
sd_EVAL_REF     112
mean_SINGLE     470
sd_SINGLE       112
mean_MULTIPLE   476
sd_MULTIPLE     109
mean_READ_SCR   475
sd_READ_SCR     107
```

Locating Information Models

```
fit_loc_info1 <- lm(LOC_INFO ~ COM_HOM + INTERNET + sqrt(ICTHOME) + COM_HOM * INTERNET, data = I
CT_data)
# summ(fit_loc_info1)
# summary(fit_loc_info1)

fit_understand1 <- lm(UNDERSTD ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET, data = ICT_d
ata)
# summ(fit_understand1)
# summary(fit_understand1)

fit_eval_ref1 <- lm(EVAL_REF ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET, data = ICT_dat
a)
# summ(fit_understand1)
# summary(fit_eval_ref1)

fit_single1 <- lm(SINGLE ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET, data = ICT_data)
# summary(fit_single1)

fit_multiple1 <- lm(MULTIPLE ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET, data = ICT_dat
a)
# summary(fit_multiple1)

# Putting Results Together
export_summs(fit4, fit_loc_info1, fit_understand1, fit_eval_ref1, fit_single1, fit_multiple1,
  model.names = c("Reading Scores", "Locating Information", "Understanding Text", "Evaluating an
d Reflecting", "single Text", "Multiple Texts"),
  scale = TRUE,
  robust = TRUE,
  error_format = "[{conf.low}, {conf.high}]"
)
```

	Reading Scores	Locating Information	Understanding Text	Evaluating and Reflecting	single Text	Multiple Texts
(Intercept)	431.06 ***	430.41 ***	431.18 ***	434.22 ***	428.68 ***	433.17 ***
	[410.11, 452.01]	[410.71, 450.11]	[410.45, 451.91]	[412.44, 456.00]	[407.43, 449.93]	[412.60, 453.75]
COM_HOM1	24.58	23.53	19.00	19.90	22.72	18.94
	[-5.37, 54.53]	[-4.82, 51.87]	[-10.68, 48.68]	[-11.43, 51.23]	[-7.60, 53.04]	[-10.65, 48.54]
INTERNET1	-21.63	-21.97	-31.07	-31.05	-24.62	-26.69

	[-55.73, 12.47]	[-54.93, 10.99]	[-66.10, 3.96]	[-67.70, 5.60]	[-59.36, 10.12]	[-61.02, 7.64]
INTERNET2	16.92	16.48	9.51	11.74	13.06	12.56
	[-5.18, 39.03]	[-4.35, 37.31]	[-12.36, 31.38]	[-11.28, 34.77]	[-9.37, 35.49]	[-9.22, 34.34]
ICTHOME	-5.05 **		-4.23 *	-3.81 *	-5.68 **	-4.20 *
	[-8.44, -1.66]		[-7.63, -0.82]	[-7.37, -0.25]	[-9.17, -2.20]	[-7.61, -0.80]
COM_HOM1:INTERNET1	-37.08	-36.42	-30.08	-32.02	-40.53	-28.86
	[-81.69, 7.53]	[-79.81, 6.97]	[-75.57, 15.41]	[-79.63, 15.59]	[-86.32, 5.27]	[-73.71, 15.98]
COM_HOM1:INTERNET2	45.69 **	44.22 **	53.66 ***	58.10 ***	51.06 **	53.30 ***
	[14.34, 77.04]	[14.45, 74.00]	[22.57, 84.75]	[25.23, 90.96]	[19.24, 82.87]	[22.23, 84.37]
`sqrt(ICTHOME)`		-4.83 **				
		[-8.21, -1.45]				
N	4565	4565	4565	4565	4565	4565
R2	0.08	0.08	0.08	0.09	0.08	0.08

All continuous predictors are mean-centered and scaled by 1 standard deviation. The outcome variable is in its original units. Standard errors are heteroskedasticity robust. *** p < 0.001; ** p < 0.01; * p < 0.05.

Changing ICT home to a factor and running a regression and posthoc

```
fit_icthome <- lm(READ_SCR ~ factor(ICTHOME), data = ICT_data)
summ(fit_icthome)
```

Observations	4646 (192 missing obs. deleted)
Dependent variable	READ_SCR
Type	OLS linear regression

F(11,4634)	29.99
R²	0.07

Adj. R^2	0.06

	Est.	S.E.	t val.	p
(Intercept)	346.00	22.62	15.30	0.00
factor(ICTHOME)2	10.18	30.93	0.33	0.74
factor(ICTHOME)3	82.44	30.93	2.67	0.01
factor(ICTHOME)4	88.72	25.91	3.42	0.00
factor(ICTHOME)5	116.38	24.39	4.77	0.00
factor(ICTHOME)6	134.77	23.64	5.70	0.00
factor(ICTHOME)7	148.84	23.21	6.41	0.00
factor(ICTHOME)8	162.87	23.07	7.06	0.00
factor(ICTHOME)9	183.60	22.94	8.00	0.00
factor(ICTHOME)10	179.25	22.87	7.84	0.00
factor(ICTHOME)11	164.48	22.86	7.19	0.00
factor(ICTHOME)12	129.11	22.90	5.64	0.00

Standard errors: OLS

```
car::Anova(fit_icthome, type = "III")
```

```
Anova Table (Type III tests)

Response: READ_SCR
                  Sum Sq   Df F value Pr(>F)
(Intercept)      2394357    1     234 <2e-16 ***
factor(ICTHOME)  3374661   11      30 <2e-16 ***
Residuals       47409644 4634
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(aov(fit_icthome), ordered = TRUE)
```

```
  Tukey multiple comparisons of means
    95% family-wise confidence level
    factor levels have been ordered

Fit: aov(formula = fit_icthome)

$`factor(ICTHOME)`
        diff     lwr   upr p adj
2-1    10.18 -90.932 111.3 1.000
3-1    82.44 -18.673 183.6 0.244
4-1    88.72   3.999 173.4 0.031
5-1   116.38  36.638 196.1 0.000
12-1  129.11  54.225 204.0 0.000
6-1   134.77  57.486 212.1 0.000
7-1   148.84  72.940 224.7 0.000
8-1   162.87  87.453 238.3 0.000
11-1  164.48  89.722 239.2 0.000
10-1  179.25 104.477 254.0 0.000
9-1   183.60 108.603 258.6 0.000
3-2    72.26 -25.266 169.8 0.391
4-2    78.54  -1.864 158.9 0.063
5-2   106.19  31.060 181.3 0.000
12-2  118.93  48.966 188.9 0.000
6-2   124.59  52.064 197.1 0.000
7-2   138.66  67.610 209.7 0.000
8-2   152.69  82.157 223.2 0.000
11-2  154.30  84.472 224.1 0.000
10-2  169.06  99.227 238.9 0.000
9-2   173.41 103.336 243.5 0.000
4-3     6.28 -74.123  86.7 1.000
5-3    33.93 -41.198 109.1 0.947
12-3   46.67 -23.292 116.6 0.564
6-3    52.33 -20.194 124.9 0.434
7-3    66.40  -4.648 137.4 0.093
8-3    80.43   9.898 151.0 0.011
11-3   82.04  12.214 151.9 0.007
10-3   96.80  26.968 166.6 0.000
9-3   101.15  31.078 171.2 0.000
5-4    27.65 -23.319  78.6 0.832
12-4   40.39  -2.599  83.4 0.089
6-4    46.05  -0.994  93.1 0.062
7-4    60.12  15.384 104.9 0.001
8-4    74.15  30.237 118.1 0.000
11-4   75.76  32.993 118.5 0.000
10-4   90.53  47.739 133.3 0.000
9-4    94.87  51.698 138.1 0.000
12-5   12.73 -19.330  44.8 0.979
6-5    18.40 -18.932  55.7 0.905
7-5    32.47  -1.909  66.8 0.085
8-5    46.50  13.200  79.8 0.000
11-5   48.10  16.337  79.9 0.000
10-5   62.87  31.077  94.7 0.000
```

```
9-5     67.22  34.903  99.5 0.000
6-12     5.66 -19.693  31.0 1.000
7-12    19.73  -1.033  40.5 0.081
8-12    33.77  14.834  52.7 0.000
11-12   35.37  19.286  51.5 0.000
10-12   50.14  33.999  66.3 0.000
9-12    54.49  37.342  71.6 0.000
7-6     14.07 -14.150  42.3 0.898
8-6     28.10   1.203  55.0 0.031
11-6    29.71   4.730  54.7 0.006
10-6    44.48  19.462  69.5 0.000
9-6     48.83  23.151  74.5 0.000
8-7     14.03  -8.597  36.7 0.675
11-7    15.64  -4.670  35.9 0.329
10-7    30.40  10.054  50.8 0.000
9-7     34.75  13.597  55.9 0.000
11-8     1.61 -16.822  20.0 1.000
10-8    16.37  -2.102  34.8 0.142
9-8     20.72   1.362  40.1 0.024
10-11   14.77  -0.777  30.3 0.081
9-11    19.12   2.530  35.7 0.009
9-10     4.35 -12.289  21.0 0.999
```

Assumptions Testing

Assumption of Independence (Durbin Watson Test)

```
car::durbinWatsonTest(fit3)
```

```
 lag Autocorrelation D-W Statistic p-value
   1           0.167          1.67       0
 Alternative hypothesis: rho != 0
```

Assumption of No Multicollinearity

```
car::vif(fit3)
```

```
                            GVIF Df GVIF^(1/(2*Df))
COM_HOM                     78.3  1            8.85
INTERNET                   683.4  2            5.11
ICTHOME                     38.6  1            6.22
COM_HOM:INTERNET          6527.9  2            8.99
COM_HOM:ICTHOME            242.8  1           15.58
INTERNET:ICTHOME          4649.6  2            8.26
COM_HOM:INTERNET:ICTHOME 22320.1  2           12.22
```

```
mean(car::vif(fit4))
```

```
[1] 7.83
```

```
range(car::vif(fit4))
```

```
[1]  1.0 56.9
```

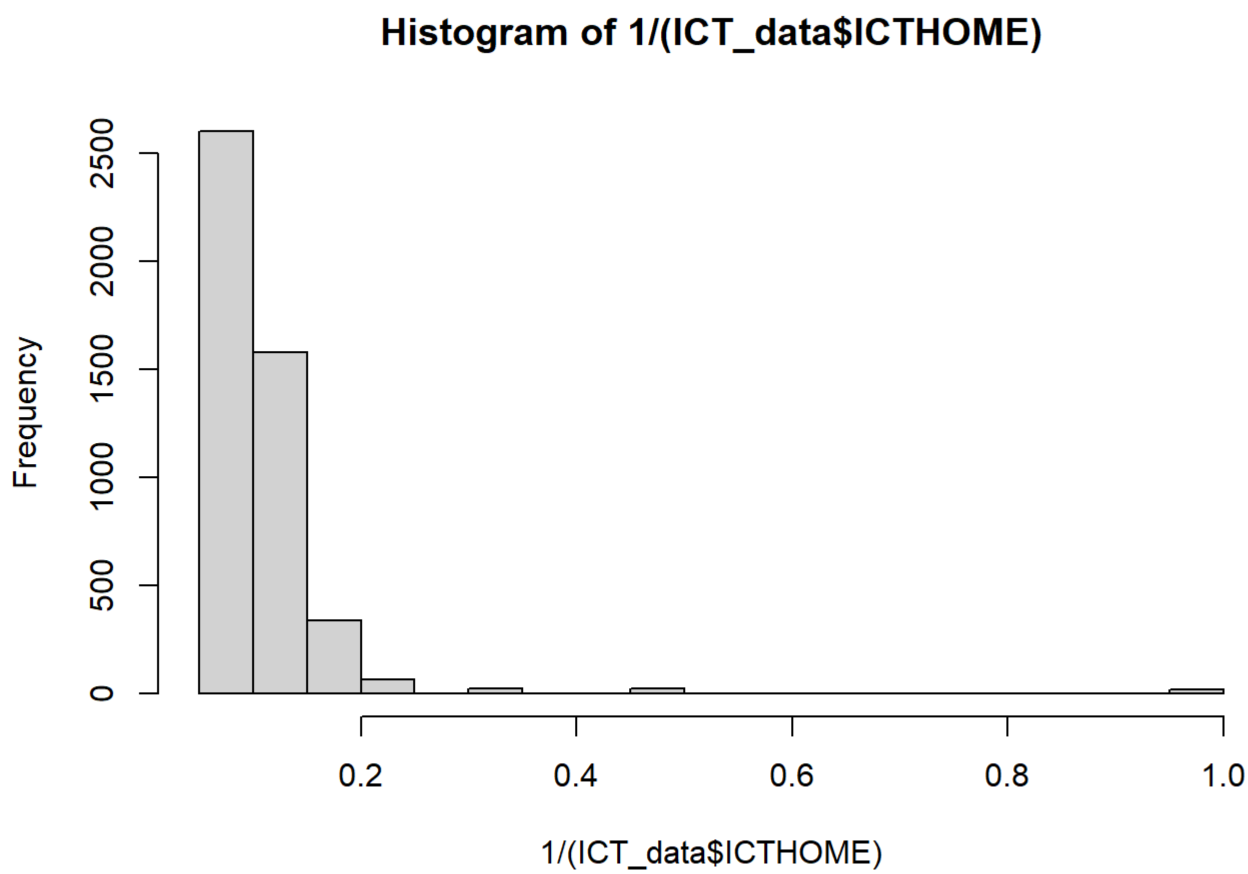
Test of Normality

```
options(scipen = 999)
pastecs::stat.desc(ICT_data[, c("COM_HOM", "INTERNET", "ICTHOME", "LOC_INFO", "UNDERSTD", "EVAL_
REF", "SINGLE", "MULTIPLE", "READ_SCR")], basic = FALSE, norm = TRUE)
```

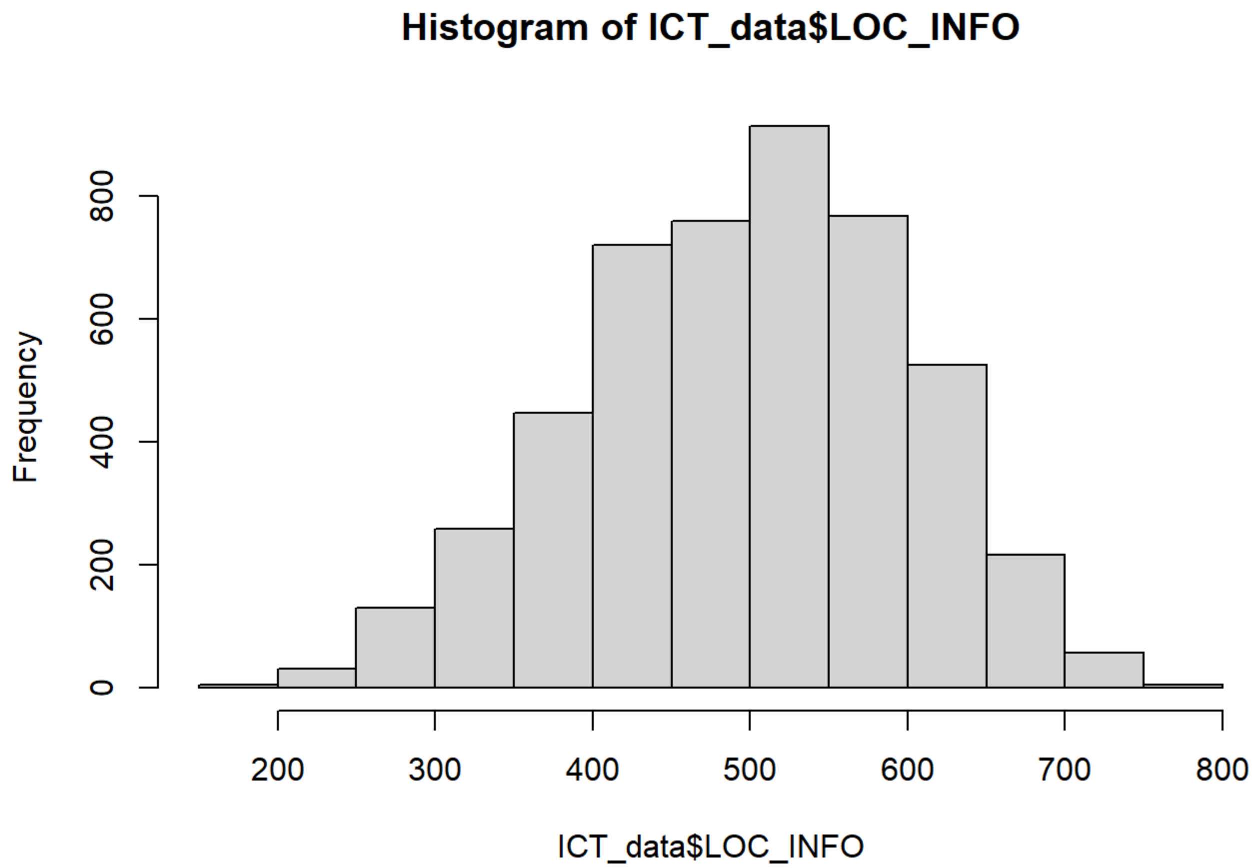
```
           COM_HOM INTERNET
median          NA       NA
mean            NA       NA
SE.mean         NA       NA
CI.mean         NA       NA
var             NA       NA
std.dev         NA       NA
coef.var        NA       NA
skewness        NA       NA
skew.2SE        NA       NA
kurtosis        NA       NA
kurt.2SE        NA       NA
normtest.W      NA       NA
normtest.p      NA       NA
                                                         ICTHOME
median       10.000000000000000000000000000000000000000000000000
mean          9.413043478260869179052860999945551156997680664062
SE.mean       0.031634277022760164121972792372616822831332683563
CI.mean       0.062018203880829271545938041754197911359369754791
var           4.649379884869190959761908743530511856079101562500
std.dev       2.156242074737711789111926918849349021911621093750
coef.var      0.229069596854352769410567702834669034928083419800
skewness     -0.986218178311966120475062780315056443214416503906
skew.2SE    -13.726099269098448019121860852465033531188964843750
kurtosis      0.937693858505030508609934258856810629367828369141
kurt.2SE      6.526773239428456996336080919718369841575622558594
normtest.W    0.907772389997938855543679892434738576412200927734
normtest.p    0.000000000000000000000000000000000000000000000142
                      LOC_INFO             UNDERSTD              EVAL_REF
median       503.4719000000000051   501.203050000000019   509.73424999999997453
mean         496.8873950806118387   495.747797126911962   504.96943985117815146
SE.mean        1.4775010244326423     1.531847595835663     1.58859180440906922
CI.mean        2.8965736028422979     3.003117585910887     3.11436202767431736
var        10561.3988830912185222 11352.643041124976662 12209.29252997053481522
std.dev      102.7686668352334891   106.548782447876789   110.49566747149199841
coef.var       0.2068248618352674     0.214925377511259     0.21881654363887185
skewness      -0.2155196847864908    -0.131073946560616    -0.16719533646856038
skew.2SE      -3.0608979118640707    -1.861565312350168    -2.37457593155363478
kurtosis      -0.3976695163469524    -0.519204722855075    -0.53369432280692042
kurt.2SE      -2.8245146153617195    -3.687738857985584    -3.79065367833863842
normtest.W     0.9935668094402292     0.993831883553952     0.99267812358104224
normtest.p     0.0000000000000608     0.000000000000139     0.00000000000000445
                        SINGLE             MULTIPLE          READ_SCR
median       502.5671999999999571   503.359600000000000   504.8434500000000
mean         497.3307972715998062   500.002211699049212   500.5676861306325
SE.mean        1.5567861253447643     1.528125416049594     1.5092258996877
CI.mean        3.0520084395043758     2.995820421621270     2.9587687788170
var        11725.2947478391361074 11297.539335361087069 11019.8165052023323
std.dev      108.2834001490493279   106.289883504316094   104.9753137894921
coef.var       0.2177291266559431     0.212578826687854     0.2097125257944
skewness      -0.1603654396513764    -0.139159240091768    -0.1447866948274
skew.2SE      -2.2775749688495726    -1.976395927988952    -2.0563193209098
```

```
kurtosis      -0.5070549455578677   -0.502867513308232   -0.4657524434249
kurt.2SE      -3.6014430215216926   -3.571701079774230   -3.3080850543402
normtest.W     0.9933676670488037    0.994120372352828    0.9947735864159
normtest.p     0.0000000000000332    0.000000000000351    0.0000000000032
```

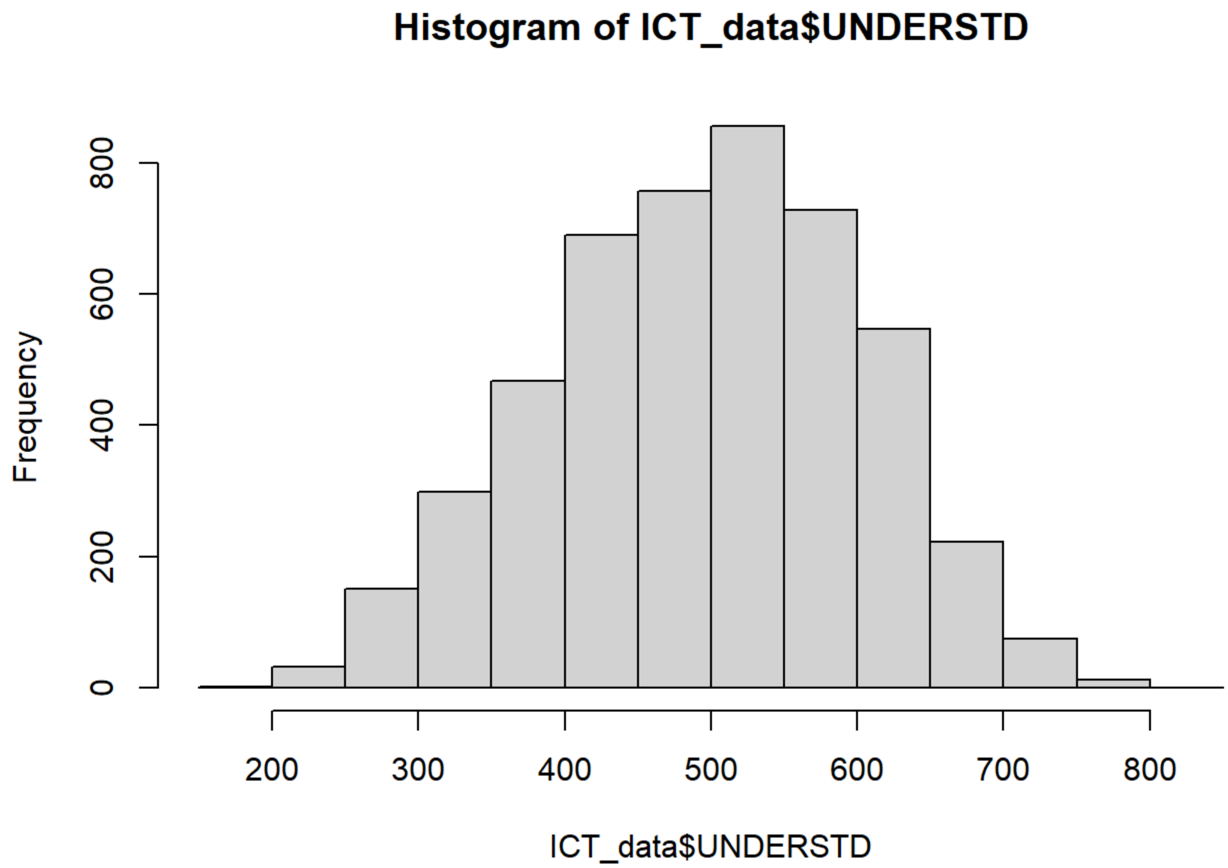
```
hist(1 / (ICT_data$ICTHOME))
```



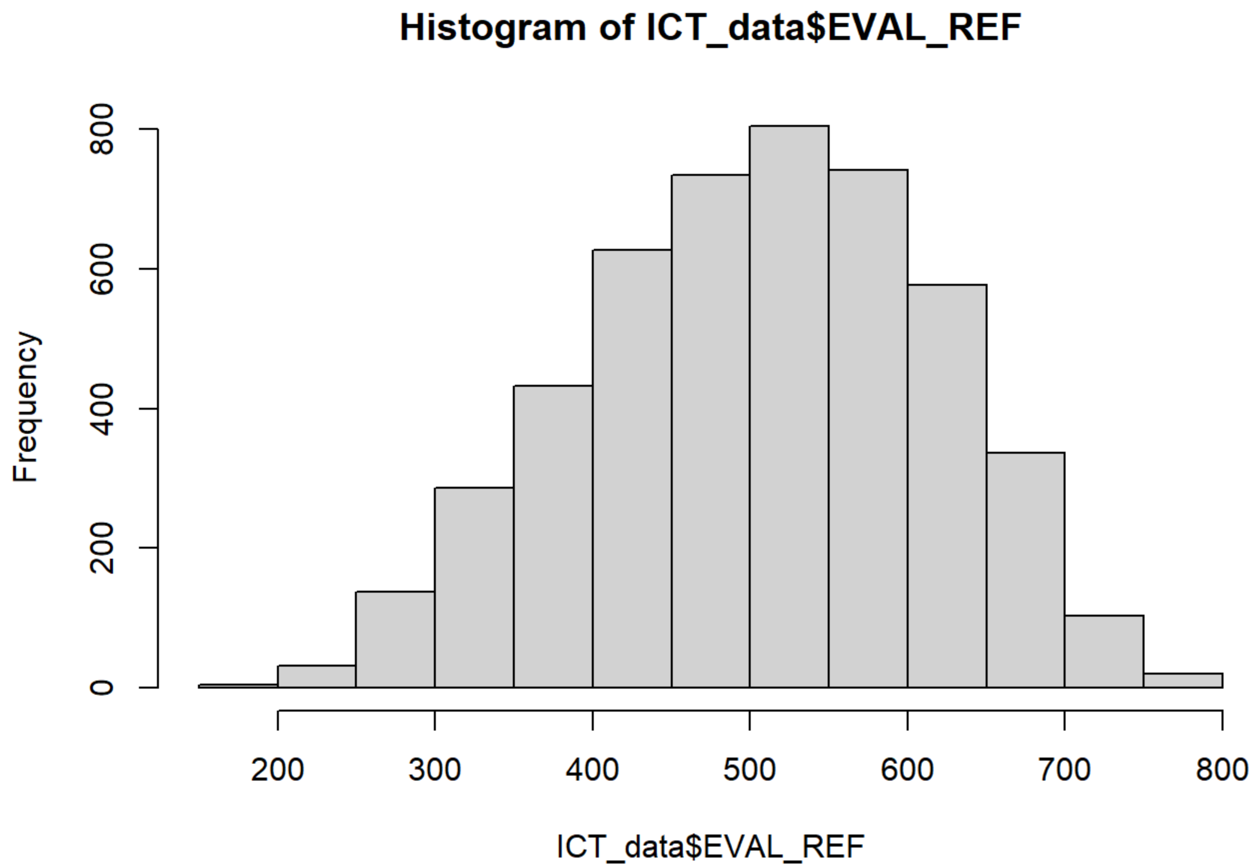

```
hist(ICT_data$LOC_INFO)
```



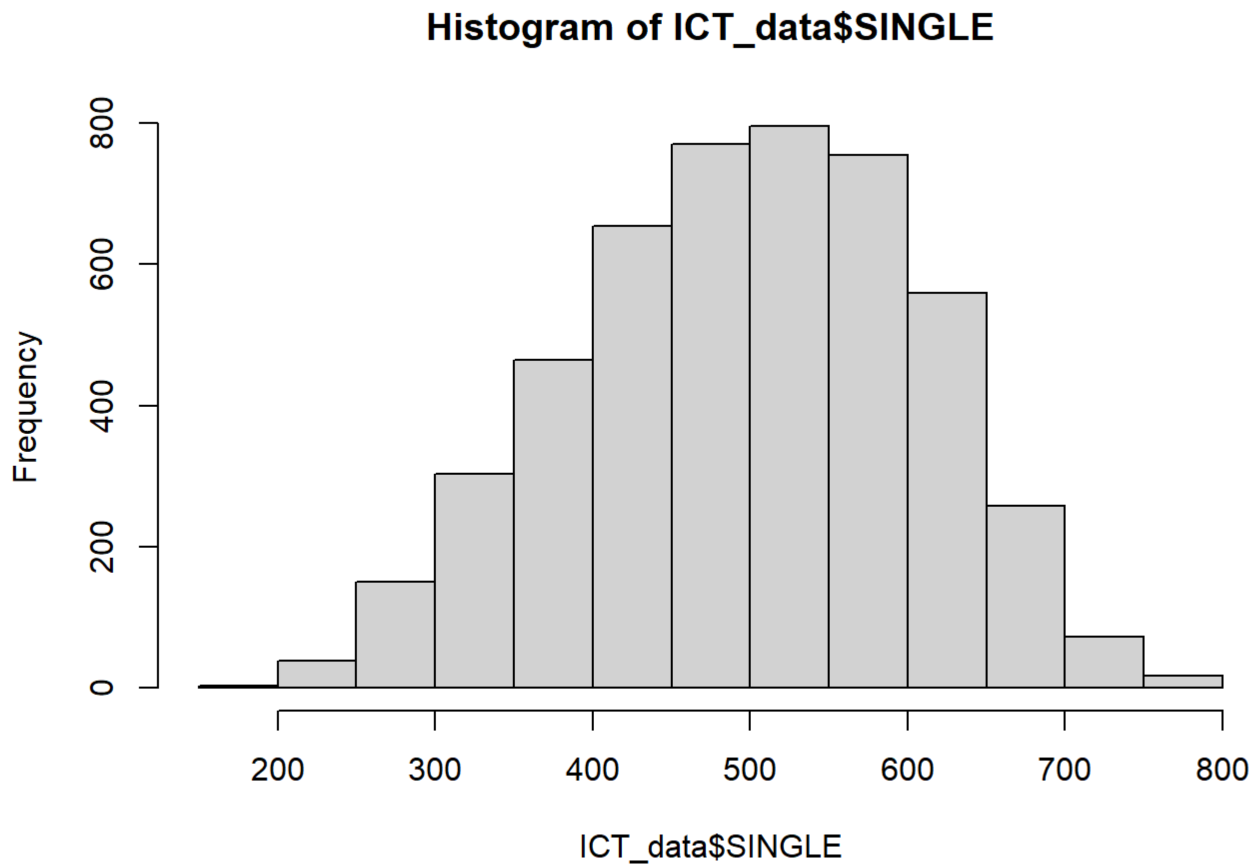

```
hist(ICT_data$UNDERSTD)
```



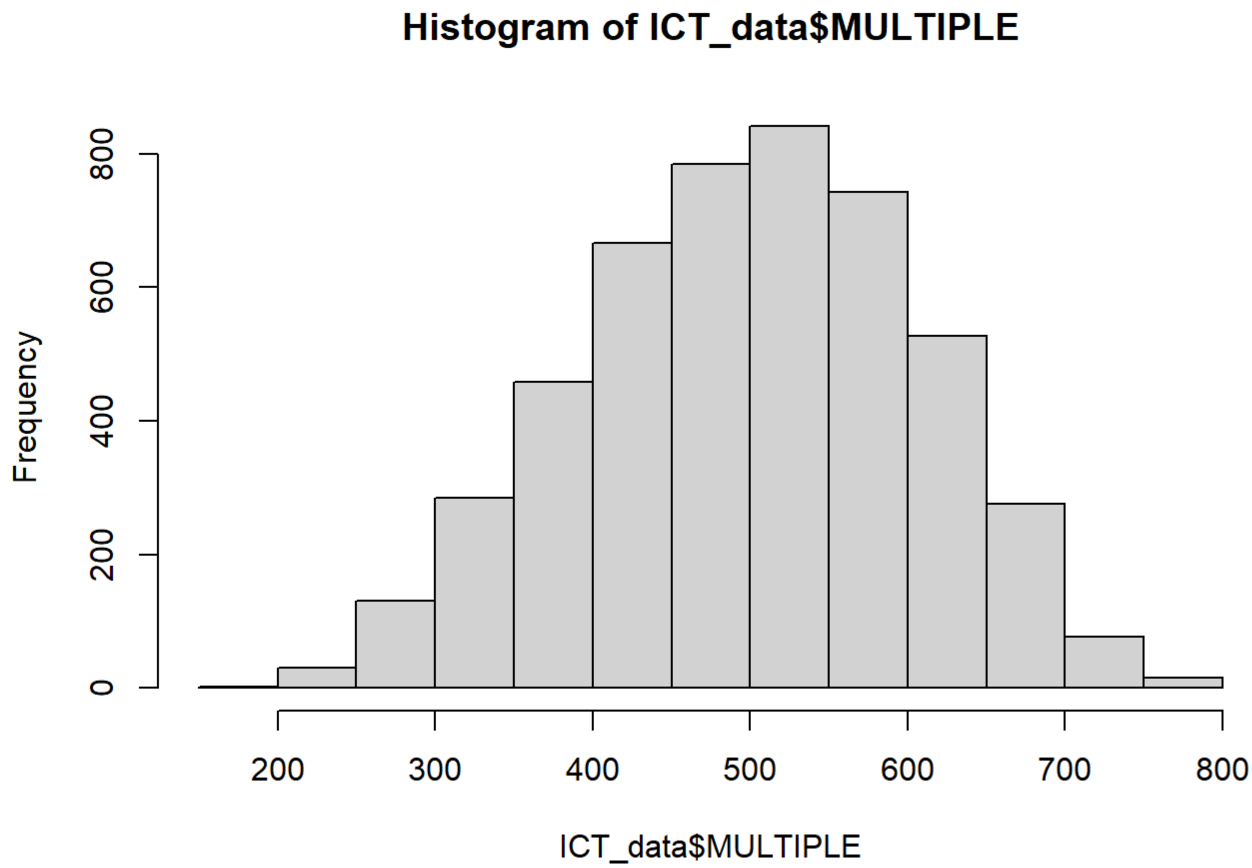

```
hist(ICT_data$EVAL_REF)
```



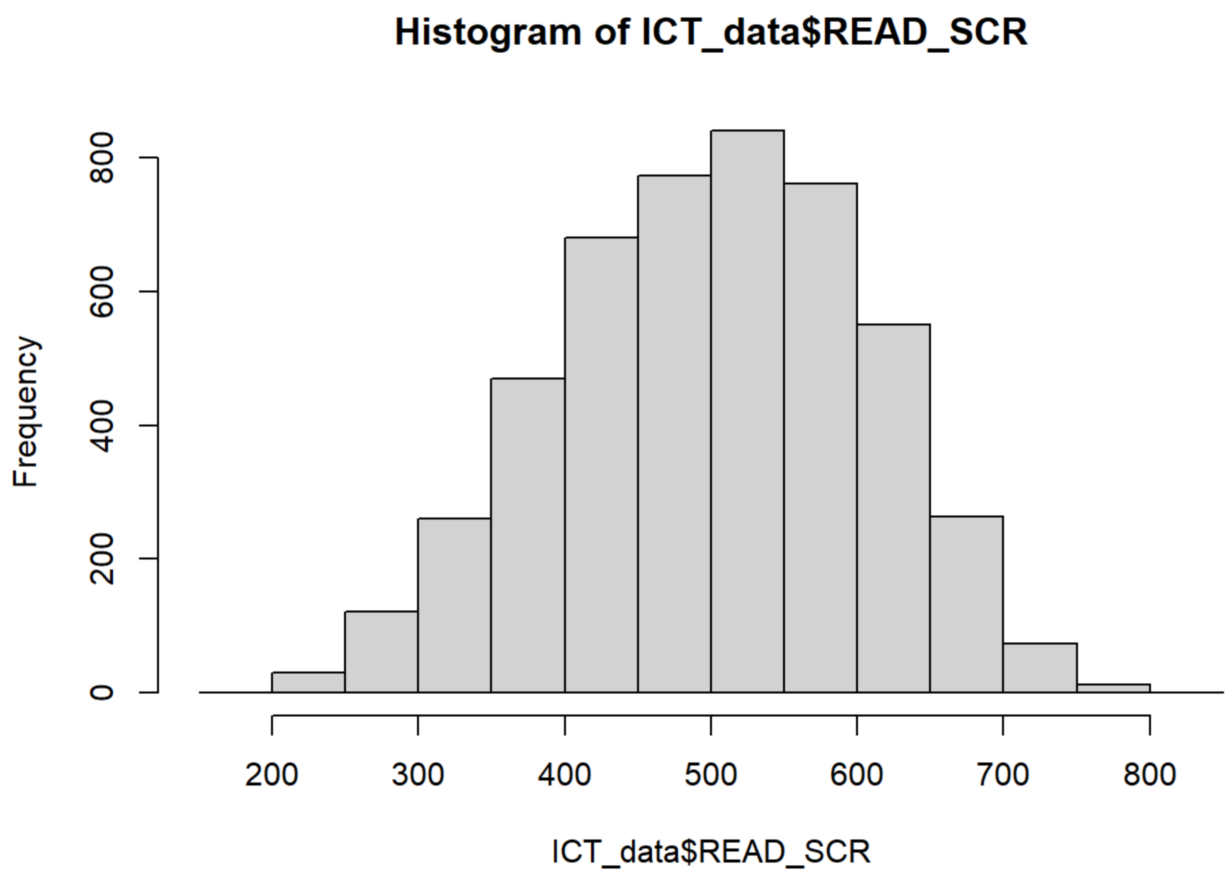

```
hist(ICT_data$SINGLE)
```




```
hist(ICT_data$MULTIPLE)
```

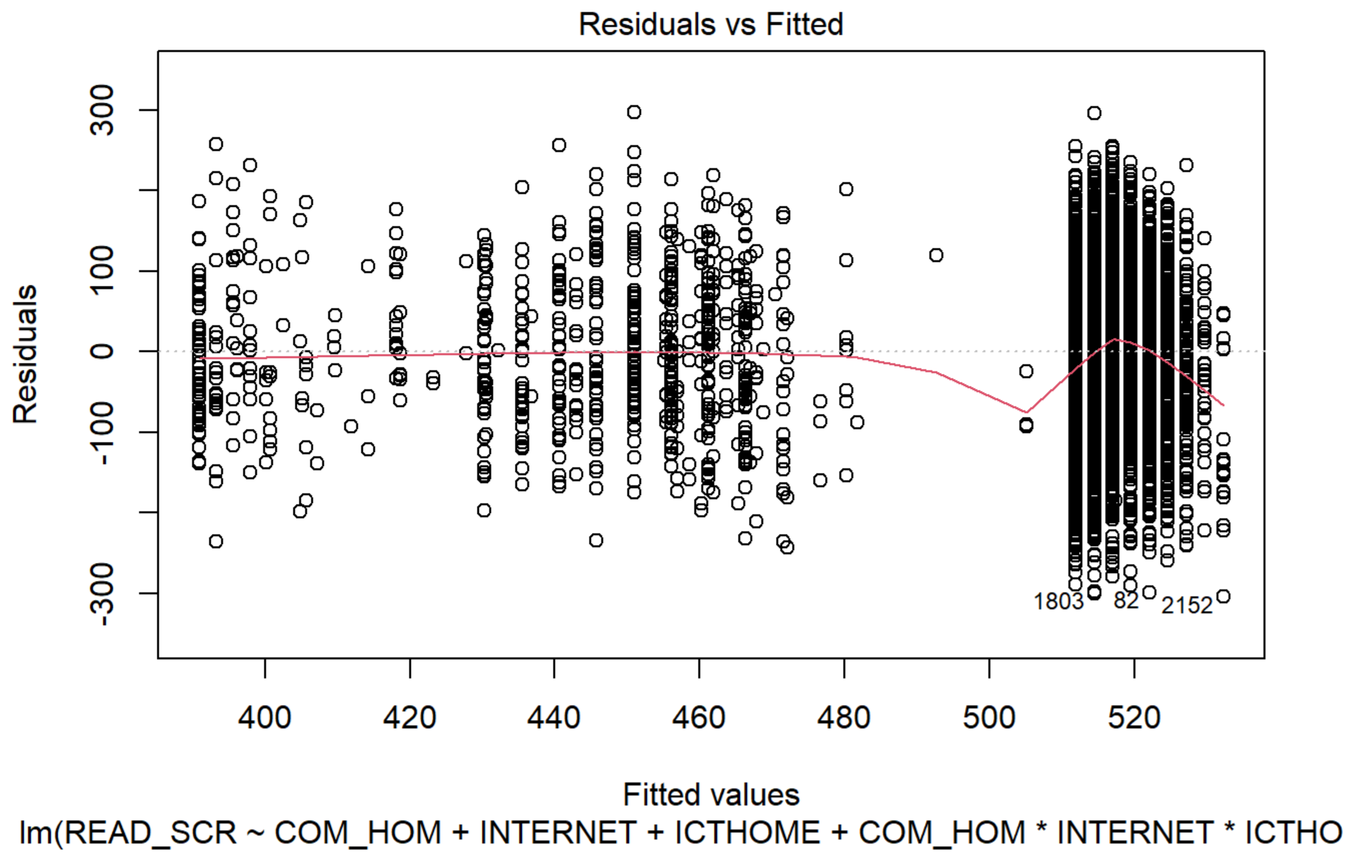



```
hist(ICT_data$READ_SCR)
```

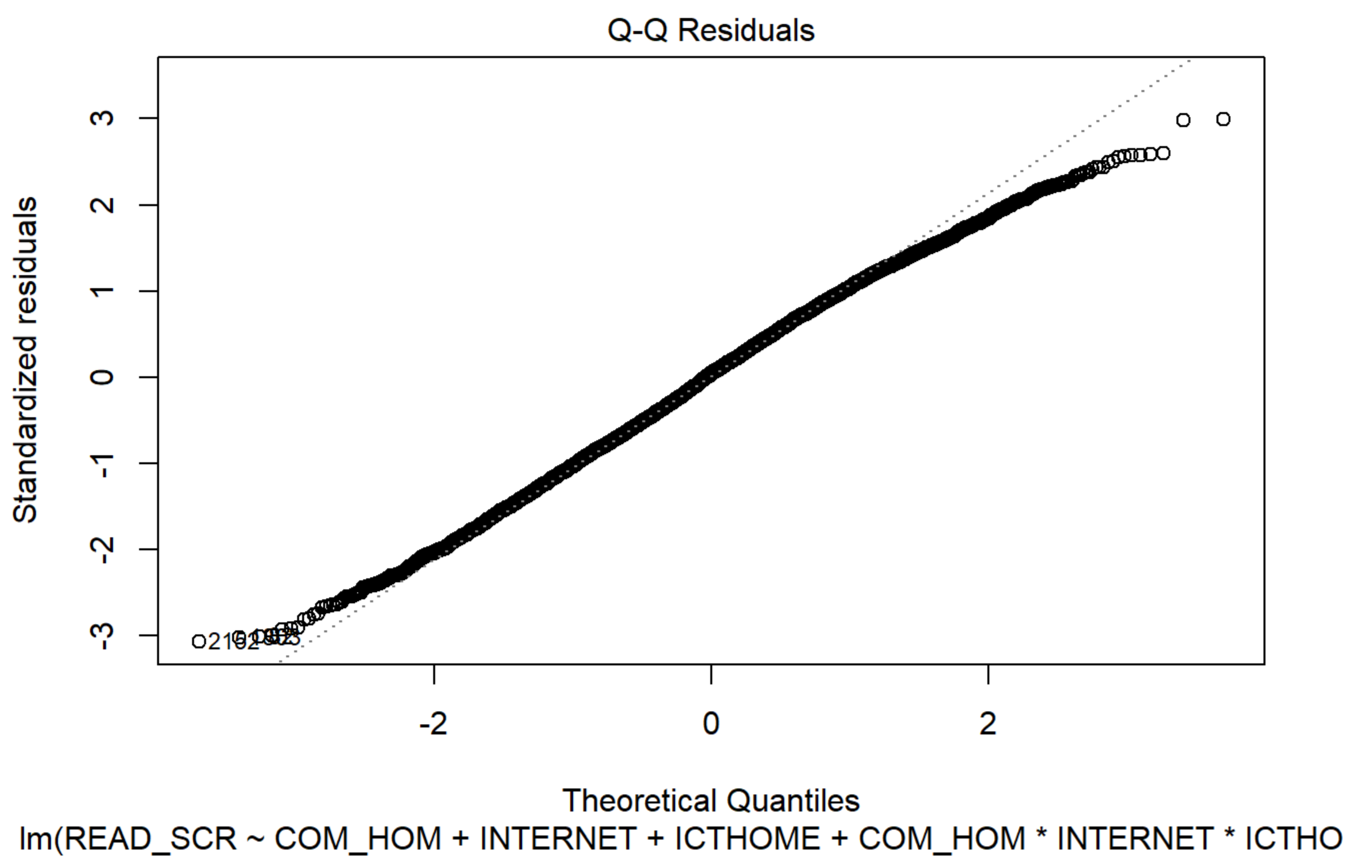


Residual Assumptions

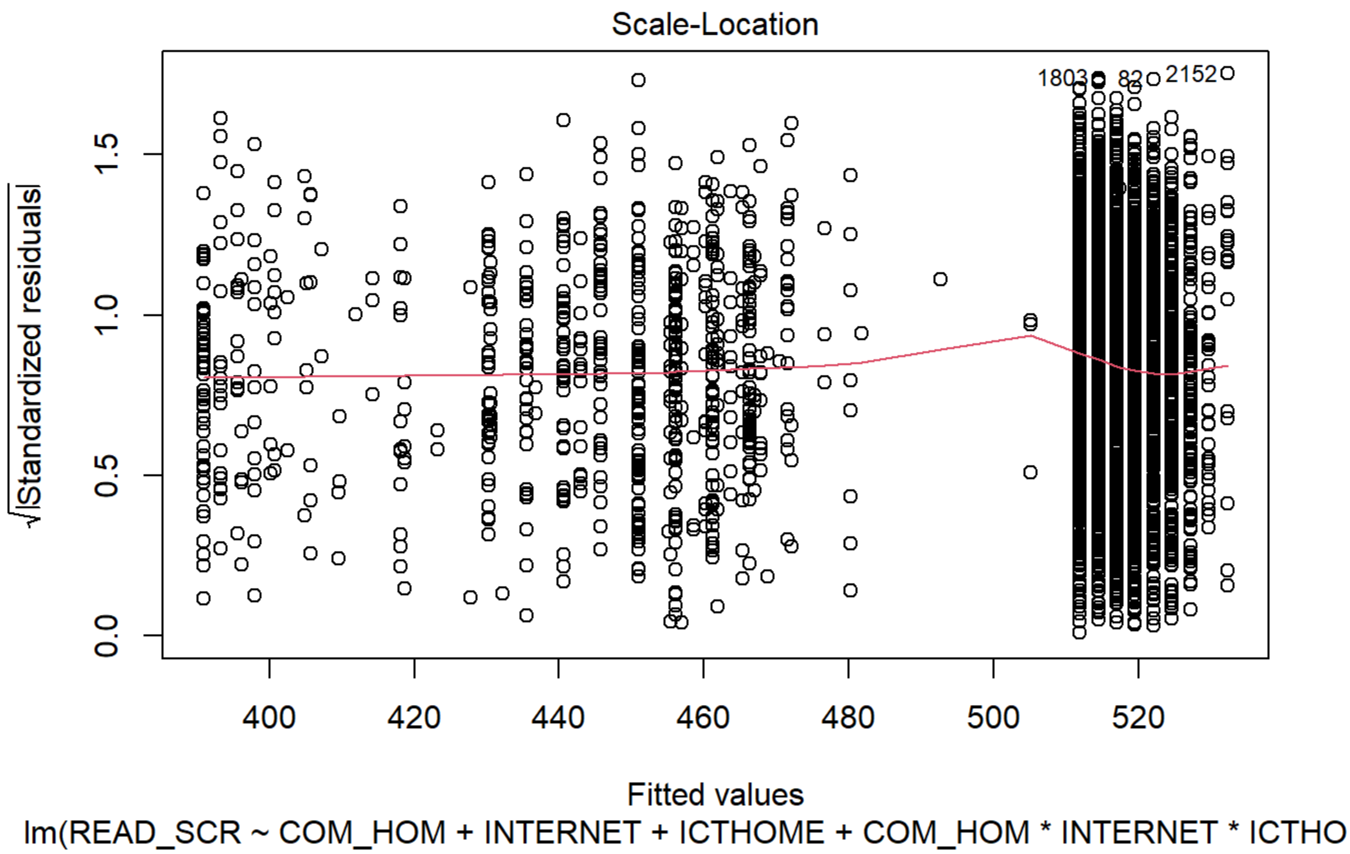
```
plot(fit3)
```

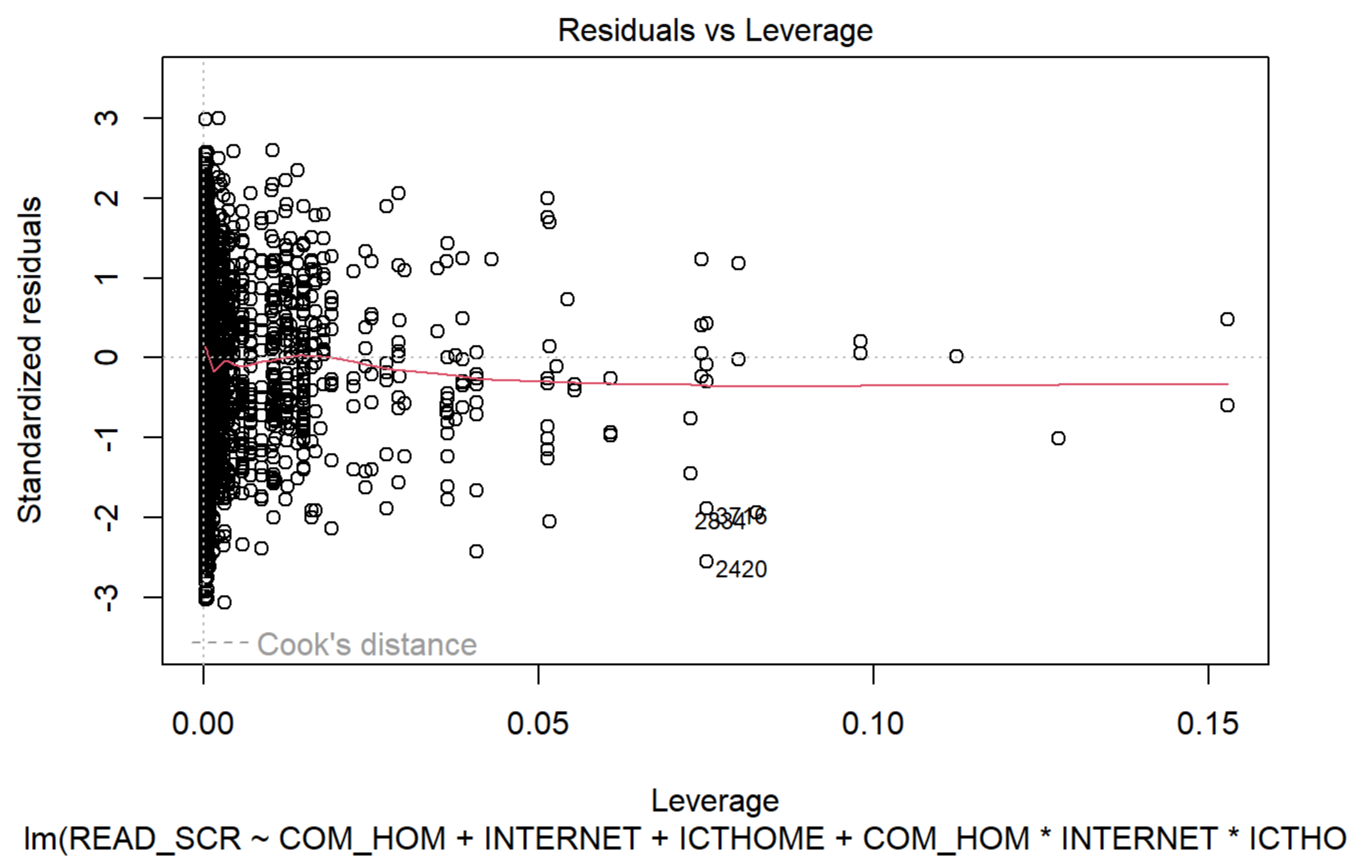


Residuals vs Fitted
Residuals
300
100
0
-100
-300
1803
82
2152
400
420
440
460
480
500
520
Fitted values
lm(READ_SCR ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET * ICTHO

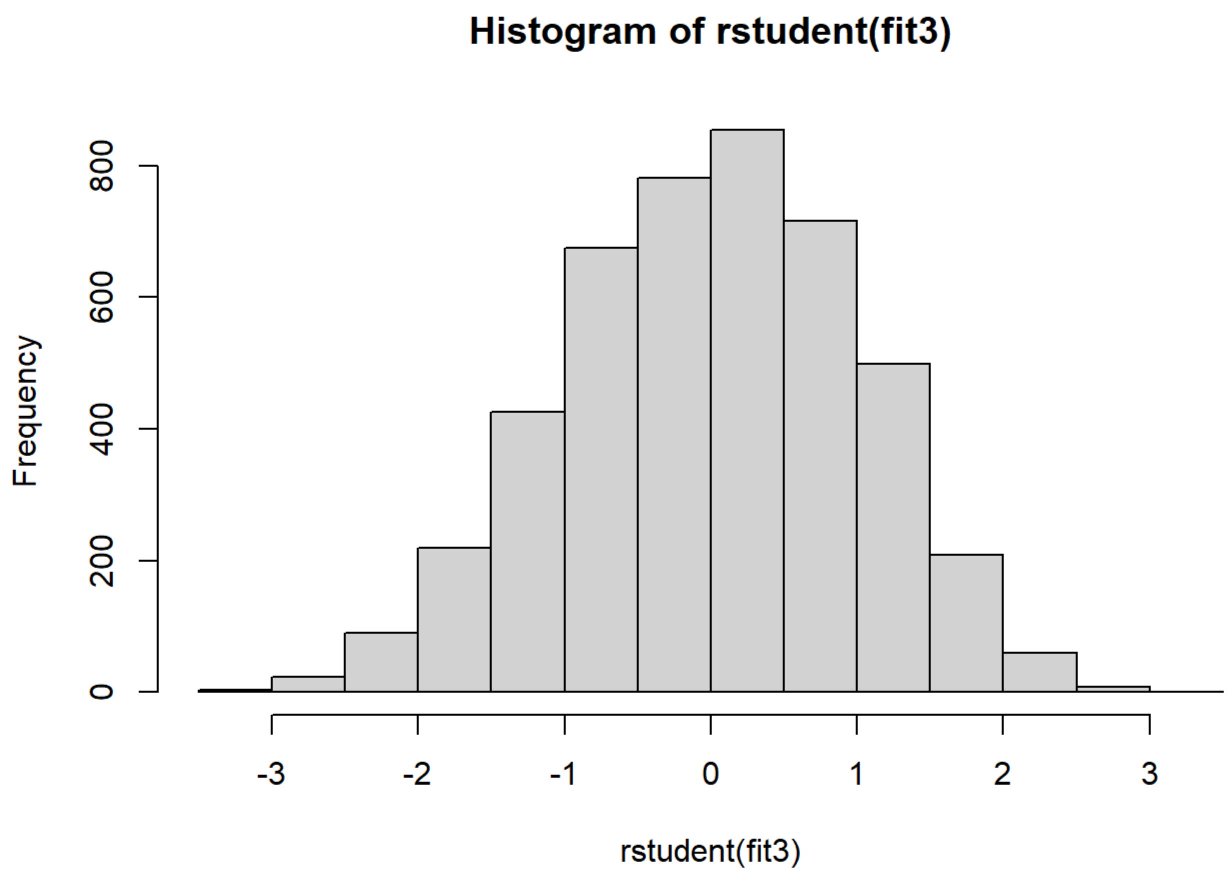


Q-Q Residuals
Standardized residuals
3
2
1
0
-1
-2
-3
-2
0
2
Theoretical Quantiles
lm(READ_SCR ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET * ICTHO


Scale-Location
√|Standardized residuals|
Fitted values
lm(READ_SCR ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET * ICTHO
1803
82
2152
0.0
0.5
1.0
1.5
400
420
440
460
480
500
520


Residuals vs Leverage
Standardized residuals
Leverage
lm(READ_SCR ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET * ICTHO
Cook's distance
2420
-3
-2
-1
0
1
2
3
0.00
0.05
0.10
0.15

```
hist(rstudent(fit3))
```




```
anova(fit2, fit3)
```

```
Analysis of Variance Table

Model 1: READ_SCR ~ COM_HOM + INTERNET + COM_HOM * INTERNET
Model 2: READ_SCR ~ COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET *
    ICTHOME
  Res.Df      RSS Df Sum of Sq    F Pr(>F)
1   4559 45063690
2   4553 44842682  6    221008 3.74  0.001 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Saving residuals with data and doing further assumptions test

```
ICT1_data <- na.omit(ICT_data)
ICT1_data$standardized_residual <- rstandard(fit3)
ICT1_data$studentized_residual <- rstudent(fit3)
ICT1_data$cooks_distance <- cooks.distance(fit3)
ICT1_data$dfbeta <- dfbeta(fit3)
ICT1_data$dffit <- dffits(fit3)
ICT1_data$leverage <- hatvalues(fit3)
ICT1_data$covariance_ratio <- covratio(fit3)

summary(ICT1_data)
```

```
 COM_HOM  INTERNET    ICTHOME         LOC_INFO      UNDERSTD      EVAL_REF  
 0: 566   0: 178   Min.   : 1.00   Min.   :193   Min.   :200   Min.   :198  
 1:3999   1: 138   1st Qu.: 8.00   1st Qu.:432   1st Qu.:423   1st Qu.:430  
          2:4249   Median :10.00   Median :508   Median :506   Median :514  
                   Mean   : 9.47   Mean   :502   Mean   :500   Mean   :509  
                   3rd Qu.:11.00   3rd Qu.:577   3rd Qu.:579   3rd Qu.:591  
                   Max.   :12.00   Max.   :785   Max.   :815   Max.   :795  
     SINGLE       MULTIPLE      READ_SCR      W_FSTUWT       PV1READ   
 Min.   :185   Min.   :193   Min.   :157   Min.   : 263   Min.   :161  
 1st Qu.:426   1st Qu.:430   1st Qu.:433   1st Qu.: 563   1st Qu.:429  
 Median :506   Median :509   Median :509   Median : 660   Median :509  
 Mean   :502   Mean   :504   Mean   :505   Mean   : 733   Mean   :505  
 3rd Qu.:582   3rd Qu.:583   3rd Qu.:582   3rd Qu.: 852   3rd Qu.:582  
 Max.   :783   Max.   :785   Max.   :810   Max.   :2946   Max.   :869  
    PV2READ       PV3READ       PV4READ       PV5READ       PV6READ   
 Min.   :176   Min.   :132   Min.   :140   Min.   :138   Min.   :128  
 1st Qu.:430   1st Qu.:428   1st Qu.:432   1st Qu.:428   1st Qu.:429  
 Median :509   Median :508   Median :507   Median :508   Median :509  
 Mean   :505   Mean   :505   Mean   :505   Mean   :505   Mean   :505  
 3rd Qu.:583   3rd Qu.:581   3rd Qu.:583   3rd Qu.:583   3rd Qu.:583  
 Max.   :898   Max.   :858   Max.   :834   Max.   :853   Max.   :845  
    PV7READ       PV8READ       PV9READ       PV10READ   standardized_residual
 Min.   :158   Min.   :171   Min.   :174   Min.   :168   Min.   :-3.072       
 1st Qu.:430   1st Qu.:429   1st Qu.:430   1st Qu.:432   1st Qu.:-0.702       
 Median :507   Median :508   Median :508   Median :509   Median : 0.045       
 Mean   :504   Mean   :505   Mean   :506   Mean   :505   Mean   : 0.000       
 3rd Qu.:581   3rd Qu.:583   3rd Qu.:580   3rd Qu.:582   3rd Qu.: 0.732       
 Max.   :815   Max.   :823   Max.   :818   Max.   :834   Max.   : 2.999       
 studentized_residual cooks_distance 
 Min.   :-3.075       Min.   :0.0000  
 1st Qu.:-0.702       1st Qu.:0.0000  
 Median : 0.045       Median :0.0000  
 Mean   : 0.000       Mean   :0.0002  
 3rd Qu.: 0.732       3rd Qu.:0.0001  
 Max.   : 3.001       Max.   :0.0440  
 dfbeta.(Intercept)  dfbeta.COM_HOM1   dfbeta.INTERNET1  dfbeta.INTERNET2   dfbeta.ICTHOME   dfb
eta.COM_HOM1:INTERNET1  dfbeta.COM_HOM1:INTERNET2  dfbeta.COM_HOM1:ICTHOME  dfbeta.INTERNET1:ICT
HOME  dfbeta.INTERNET2:ICTHOME  dfbeta.COM_HOM1:INTERNET1:ICTHOME  dfbeta.COM_HOM1:INTERNET2:ICT
HOME
 Min.   :-11.89    Min.   :-22.77    Min.   :-14.19    Min.   :-9.54     Min.   :-2.377    Min.   
:-22.66    Min.   :-11.89    Min.   :-1.877    Min.   :-1.877    Min.   :-1.877    Min.   :-3.00
7    Min.   :-3.007  
 1st Qu.:  0.00    1st Qu.:  0.00    1st Qu.:  0.00    1st Qu.: 0.00     1st Qu.: 0.000    1st Q
u.:  0.00    1st Qu.: -0.06    1st Qu.: 0.000    1st Qu.: 0.000    1st Qu.: 0.000    1st Qu.: 0.
000    1st Qu.:-0.008  
 Median :  0.00    Median :  0.00    Median :  0.00    Median : 0.00     Median : 0.000    Media
n :  0.00    Median :  0.00    Median : 0.000    Median : 0.000    Median : 0.000    Median : 0.
000    Median : 0.000  
 Mean   :  0.00    Mean   :  0.00    Mean   :  0.00    Mean   : 0.00     Mean   : 0.000    Mean   
:  0.00    Mean   :  0.00    Mean   : 0.000    Mean   : 0.000    Mean   : 0.000    Mean   : 0.00
0    Mean   : 0.000  
 3rd Qu.:  0.00    3rd Qu.:  0.00    3rd Qu.:  0.00    3rd Qu.: 0.00     3rd Qu.: 0.000    3rd Q
```

```
u.:  0.00    3rd Qu.:  0.08    3rd Qu.: 0.000    3rd Qu.: 0.000    3rd Qu.: 0.000    3rd Qu.: 0.
000    3rd Qu.: 0.007
 Max.   :  9.54    Max.   : 11.89    Max.   : 18.50    Max.   :11.89     Max.   : 1.877    Max.
: 22.77    Max.   : 22.77    Max.   : 3.007    Max.   : 2.377    Max.   : 2.377    Max.   : 1.97
7    Max.   : 1.877
     dffit            leverage       covariance_ratio
 Min.   :-0.727   Min.   :0.0003   Min.   :0.979
 1st Qu.:-0.018   1st Qu.:0.0003   1st Qu.:1.000
 Median : 0.001   Median :0.0004   Median :1.002
 Mean   :-0.002   Mean   :0.0026   Mean   :1.003
 3rd Qu.: 0.017   3rd Qu.:0.0010   3rd Qu.:1.003
 Max.   : 0.465   Max.   :0.1528   Max.   :1.183
```

```
options(scipen = 999)
assumption_values <- ICT1_data |>
  select(standardized_residual, studentized_residual, cooks_distance, dfbeta, dffit, leverage, c
ovariance_ratio) |>
  summarize(
    mean_standardized_residual = mean(standardized_residual),
    sd_standardized_residual = sd(standardized_residual),
    mean_studentized_residual = mean(studentized_residual),
    sd_studentized_residual = sd(studentized_residual),
    mean_cooks_distance = mean(cooks_distance),
    sd_cooks_distance = sd(cooks_distance),
    mean_dfbeta = mean(dfbeta),
    sd_dfbeta = sd(dfbeta),
    mean_dffit = mean(dffit),
    sd_dffit = sd(dffit),
    mean_leverage = mean(leverage),
    sd_leverage = sd(leverage),
    mean_covariance_ratio = mean(covariance_ratio),
    sd_covariance_ratio = sd(covariance_ratio)
  ) |>
  t()
assumption_values
```

```
                                 [,1]
mean_standardized_residual -0.0001573
sd_standardized_residual    1.0000602
mean_studentized_residual  -0.0001722
sd_studentized_residual     1.0002400
mean_cooks_distance         0.0002172
sd_cooks_distance           0.0012483
mean_dfbeta                -0.0000301
sd_dfbeta                   0.5145211
mean_dffit                 -0.0018402
sd_dffit                    0.0510372
mean_leverage               0.0026287
sd_leverage                 0.0085331
mean_covariance_ratio       1.0027203
sd_covariance_ratio         0.0098628
```

```r
library(boot)
bootregression <- function(formula, data, indices) {
  d <- data[indices, ]
  fit <- lm(formula, data = d)
  return(coef(fit))
}

boot_model3 <- boot(statistic = bootregression, formula = READ_SCR ~ COM_HOM + INTERNET + ICTHOM
E + COM_HOM * INTERNET * INTERNET, data = ICT_data, R = 5000)
boot_model3
```

```


ORDINARY NONPARAMETRIC BOOTSTRAP


Call:
boot(data = ICT_data, statistic = bootregression, R = 5000, formula = READ_SCR ~
    COM_HOM + INTERNET + ICTHOME + COM_HOM * INTERNET * INTERNET)


Bootstrap Statistics :
    original  bias    std. error
t1*   453.96  0.0968      10.943
t2*    24.58 -0.1368      15.179
t3*   -21.63  0.0585      17.502
t4*    16.92 -0.0222      11.299
t5*    -2.42 -0.0047       0.828
t6*   -37.08  0.3016      22.766
t7*    45.69  0.0494      15.857
```

```r
summary(boot_model3)
```

```
Number of bootstrap replications R = 5000
  original bootBias bootSE bootMed
1   453.96   0.0968 10.943  454.17
2    24.58  -0.1368 15.179   24.45
3   -21.63   0.0585 17.502  -21.59
4    16.92  -0.0222 11.299   16.85
5    -2.42  -0.0047  0.828   -2.43
6   -37.08   0.3016 22.766  -36.60
7    45.69   0.0494 15.857   45.77
```

```
boot.ci(boot_model3, type = "bca", index = 1)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL : 
boot.ci(boot.out = boot_model3, type = "bca", index = 1)

Intervals : 
Level       BCa          
95%   (432, 475 )  
Calculations and Intervals on Original Scale
```

```
boot.ci(boot_model3, type = "bca", index = 2)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL : 
boot.ci(boot.out = boot_model3, type = "bca", index = 2)

Intervals : 
Level       BCa          
95%   (-5.16, 53.55 )  
Calculations and Intervals on Original Scale
```

```
boot.ci(boot_model3, type = "bca", index = 3)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL : 
boot.ci(boot.out = boot_model3, type = "bca", index = 3)

Intervals : 
Level       BCa          
95%   (-55.6,  12.8 )  
Calculations and Intervals on Original Scale
```

```
boot.ci(boot_model3, type = "bca", index = 4)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL : 
boot.ci(boot.out = boot_model3, type = "bca", index = 4)

Intervals : 
Level       BCa          
95%   (-5.05, 39.38 )  
Calculations and Intervals on Original Scale
```

```
boot.ci(boot_model3, type = "bca", index = 5)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL : 
boot.ci(boot.out = boot_model3, type = "bca", index = 5)

Intervals : 
Level       BCa          
95%   (-4.019, -0.786 )  
Calculations and Intervals on Original Scale
```

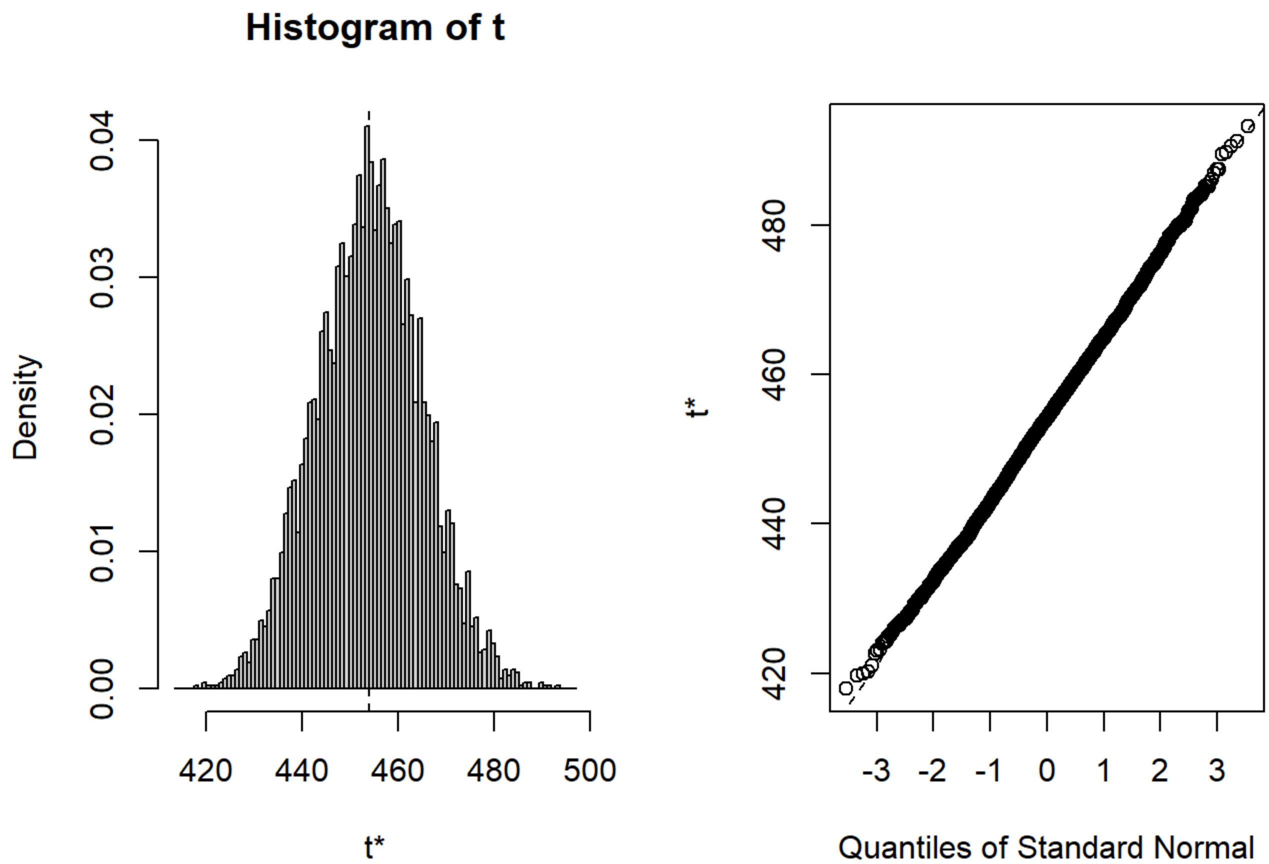
```
boot.ci(boot_model3, type = "bca", index = 6)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL :
boot.ci(boot.out = boot_model3, type = "bca", index = 6)

Intervals :
Level       BCa
95%   (-81.67,   7.12 )
Calculations and Intervals on Original Scale
```

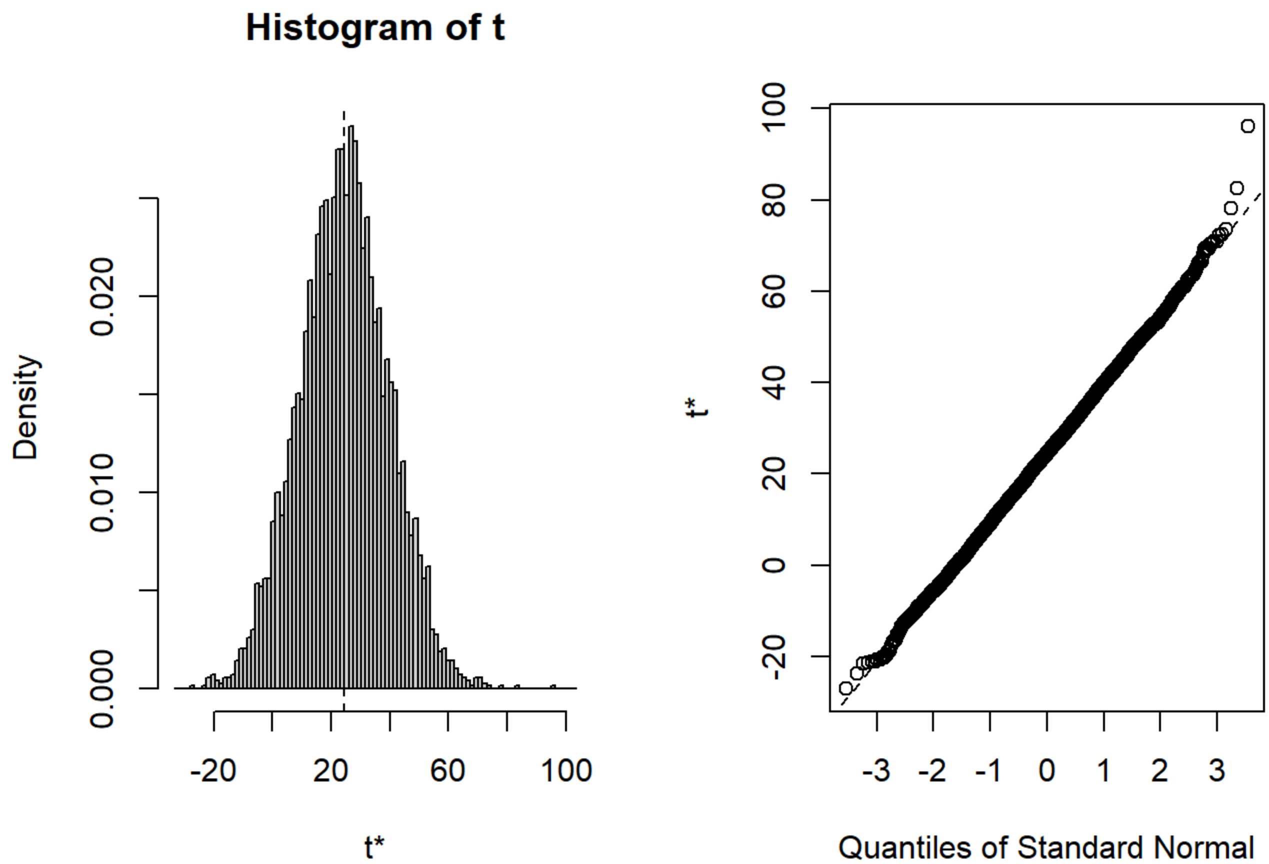
```
boot.ci(boot_model3, type = "bca", index = 7)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL :
boot.ci(boot.out = boot_model3, type = "bca", index = 7)

Intervals :
Level       BCa
95%   (14.8, 76.5 )
Calculations and Intervals on Original Scale
```

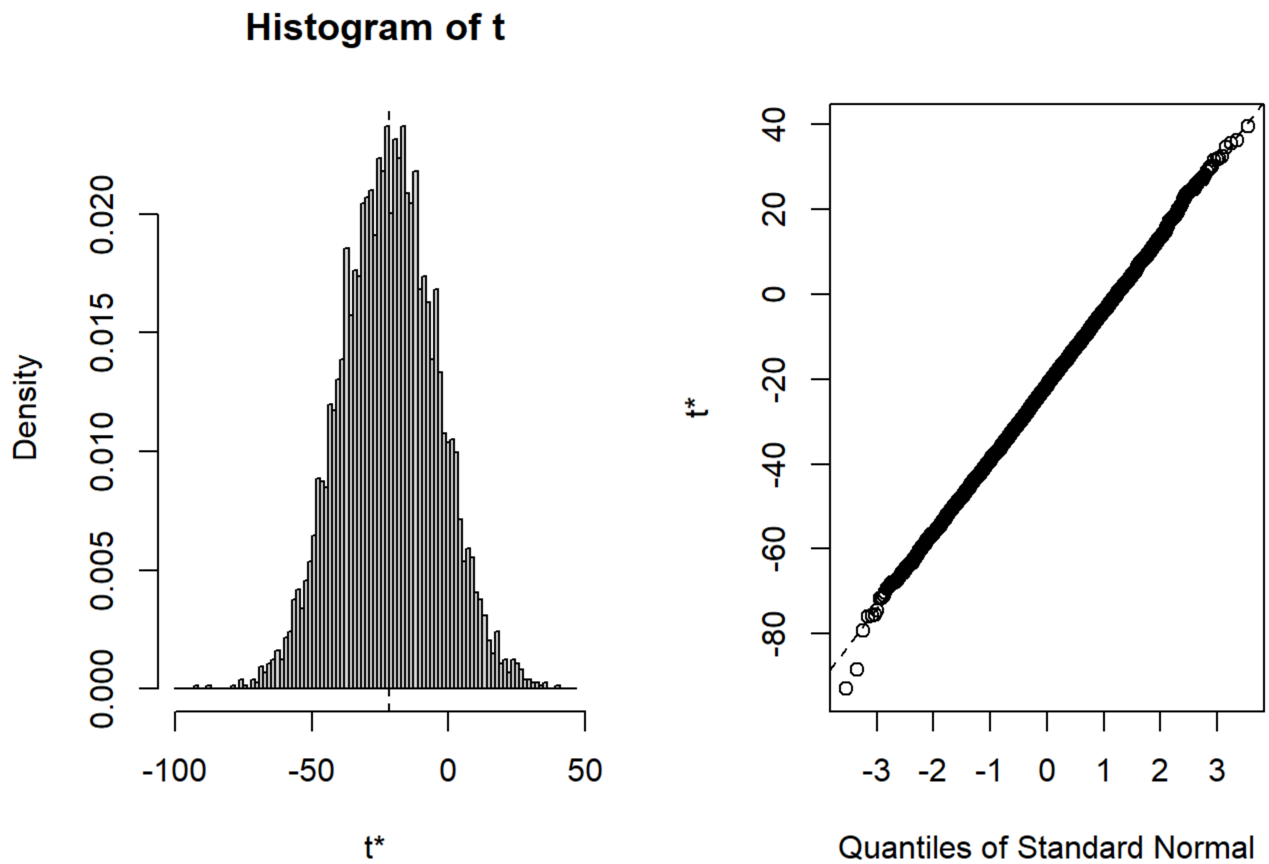
```
plot(boot_model3, index = 1)
```



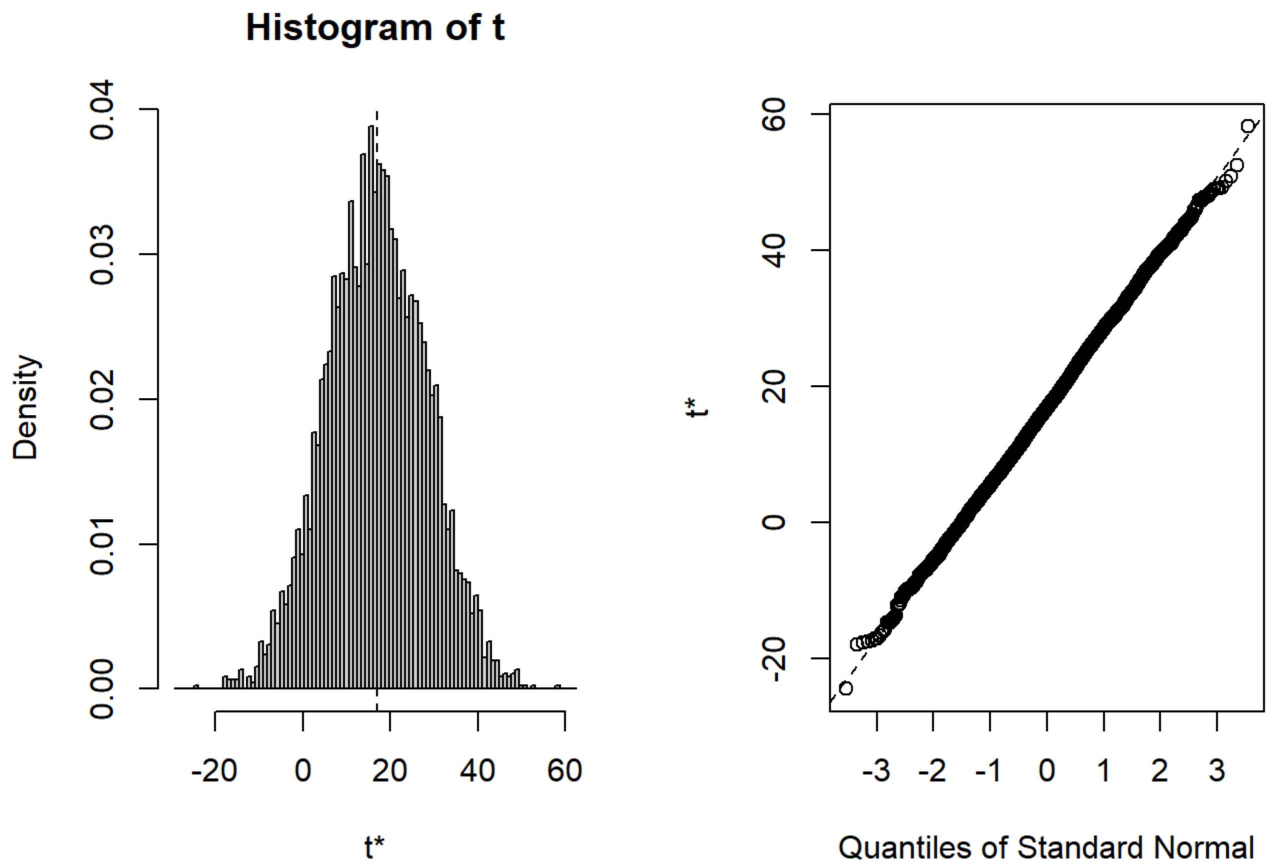

```
plot(boot_model3, index = 2)
```



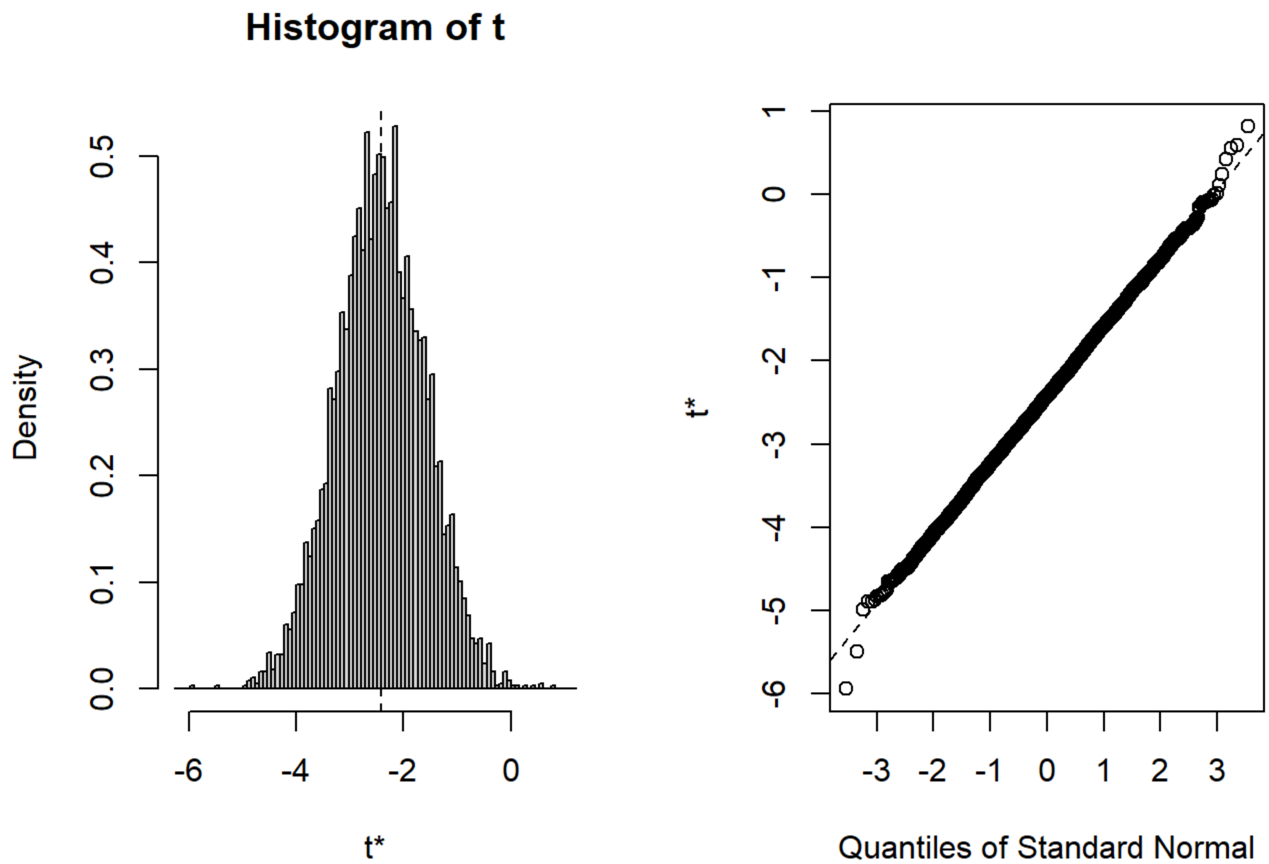

```
plot(boot_model3, index = 3)
```




```
plot(boot_model3, index = 4)
```




```
plot(boot_model3, index = 5)
```


Histogram of t
Density
0.0
0.1
0.2
0.3
0.4
0.5
-6
-4
-2
0
t*
t*
1
0
-1
-2
-3
-4
-5
-6
-3
-2
-1
0
1
2
3
Quantiles of Standard Normal