

# Project 423 - Kreslyn

2025-02-26

**Research Question 2:** What sleep-related variables mostly explain the variation in total work hours?

**Research Question 3:** Are there any significant interactions between the sleep-related variables? -> looked at work hour-related avaiables (i.e the predictors where total work hours are the response)

**Basic Correlation plot, can also just use Oliver's instead**

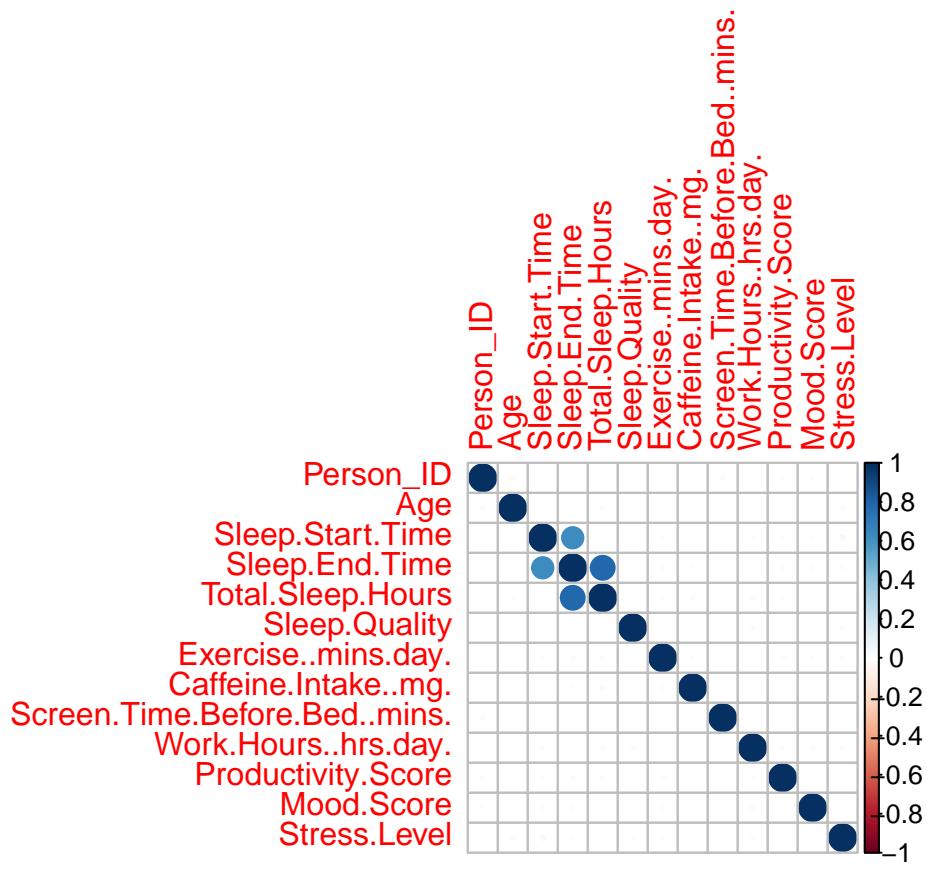
```
library(corrplot)

## Warning: package 'corrplot' was built under R version 4.3.3
## corrplot 0.95 loaded
library(tidyverse)

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr     1.1.4      v readr     2.1.5
## vforcats   1.0.0      v stringr   1.5.1
## v ggplot2   3.4.4      v tibble    3.2.1
## v lubridate 1.9.3      v tidyr    1.3.1
## v purrr    1.0.2

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
df <- read.csv("sleep_cycle_productivity.csv")

# Basic Correlation Between Variables:
# pairs(data_filtered, main = "Pairwise Scatter plots of Selected Variables")
cor_matrix <- cor(df[, sapply(df, is.numeric)])
corrplot(cor_matrix, method = "circle")
```



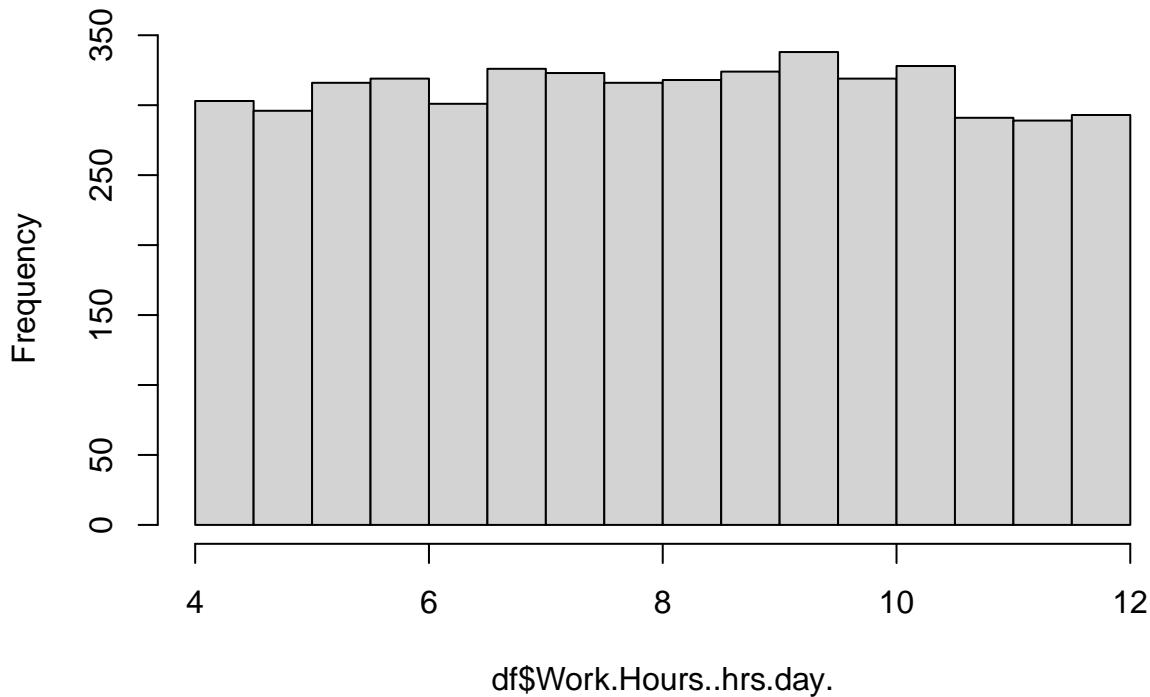
```
# Numeric variable scatter plot
```

```
numeric_vars <- c("Age", "Total.Sleep.Hours", "Exercise..mins.day.",
                 "Caffeine.Intake..mg.", "Screen.Time.Before.Bed..mins.", "Work.Hours..hrs.day.")

df_numeric <- df[, numeric_vars]

# No outliers in the response
hist(df$Work.Hours..hrs.day.)
```

### Histogram of df\$Work.Hours..hrs.day.



#### Findings (Expand in report):

Little to no correlation between variables, except start/end time and total hours no outliers.

```
df$Gender <- as.factor(df$Gender)
rating_vars <- c("Sleep.Quality", "Productivity.Score", "Mood.Score", "Stress.Level")
for (var in rating_vars) {
  new_var <- paste0(var, "Cat")
  df[[new_var]] <- cut(df[[var]], breaks = c(0, 3, 7, 10),
                        labels = c("Low", "Medium", "High"), right = TRUE)
  df[[new_var]] <- as.factor(df[[new_var]])
}

# Getting rid of unneeded columns
df_filtered <- df %>%
  select(-Date) %>%
  select(-Person_ID) %>%
  # These rows have been transformed into factors
  select(-Stress.Level) %>%
  select(-Productivity.Score) %>%
  select(-Mood.Score) %>%
  select(-Sleep.Quality) %>%
  # Will have co linearity for start, end and total hours for sleep (end-start = total)
  # Choosing to omit start time (see chunk for why we chose it...)
  select(-Sleep.Start.Time)
```

## How to choose end time or start time since we can't keep both

```
# Omit based on predictive power (P-value in isolation)
lm_start = lm(df$Work.Hours..hrs.day. ~ df$Sleep.Start.Time)
lm_end = lm(df$Work.Hours..hrs.day. ~ df$Sleep.End.Time)
lm_inter = lm(df$Work.Hours..hrs.day. ~ df$Sleep.Start.Time*df$Sleep.End.Time)

summary(lm_start)

##
## Call:
## lm(formula = df$Work.Hours..hrs.day. ~ df$Sleep.Start.Time)
##
## Residuals:
##     Min      1Q  Median      3Q      Max 
## -4.0299 -1.9539  0.0101  1.9231  4.0611 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 7.41866   0.61272 12.11   <2e-16 ***
## df$Sleep.Start.Time 0.02587   0.02781  0.93    0.352  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.276 on 4998 degrees of freedom
## Multiple R-squared:  0.0001731, Adjusted R-squared:  -2.692e-05 
## F-statistic: 0.8654 on 1 and 4998 DF,  p-value: 0.3523

summary(lm_end)

##
## Call:
## lm(formula = df$Work.Hours..hrs.day. ~ df$Sleep.End.Time)
##
## Residuals:
##     Min      1Q  Median      3Q      Max 
## -4.0202 -1.9649  0.0116  1.9194  4.0557 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 8.04724   0.09189 87.57   <2e-16 ***
## df$Sleep.End.Time -0.01192   0.01728  -0.69    0.49   
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.276 on 4998 degrees of freedom
## Multiple R-squared:  9.517e-05, Adjusted R-squared:  -0.0001049 
## F-statistic: 0.4757 on 1 and 4998 DF,  p-value: 0.4904

summary(lm_inter)

##
## Call:
## lm(formula = df$Work.Hours..hrs.day. ~ df$Sleep.Start.Time * 
##     df$Sleep.End.Time)
```

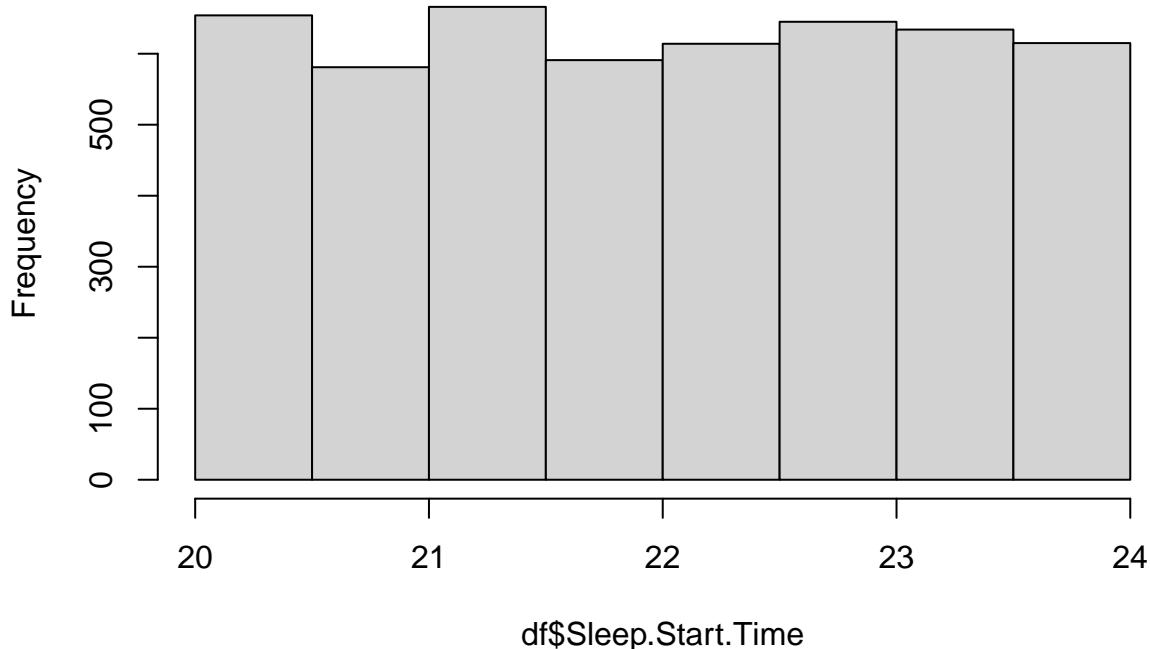
```

## 
## Residuals:
##   Min     1Q Median     3Q    Max 
## -4.0759 -1.9470  0.0161  1.9391  4.1013 
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                 8.52697   1.83432   4.649 3.43e-06 ***
## df$Sleep.Start.Time      -0.01723   0.08531  -0.202   0.840    
## df$Sleep.End.Time        -0.38799   0.34456  -1.126   0.260    
## df$Sleep.Start.Time:df$Sleep.End.Time  0.01598   0.01561   1.024   0.306  
## ---                        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 2.275 on 4996 degrees of freedom
## Multiple R-squared:  0.0009131, Adjusted R-squared:  0.0003131 
## F-statistic: 1.522 on 3 and 4996 DF,  p-value: 0.2066 

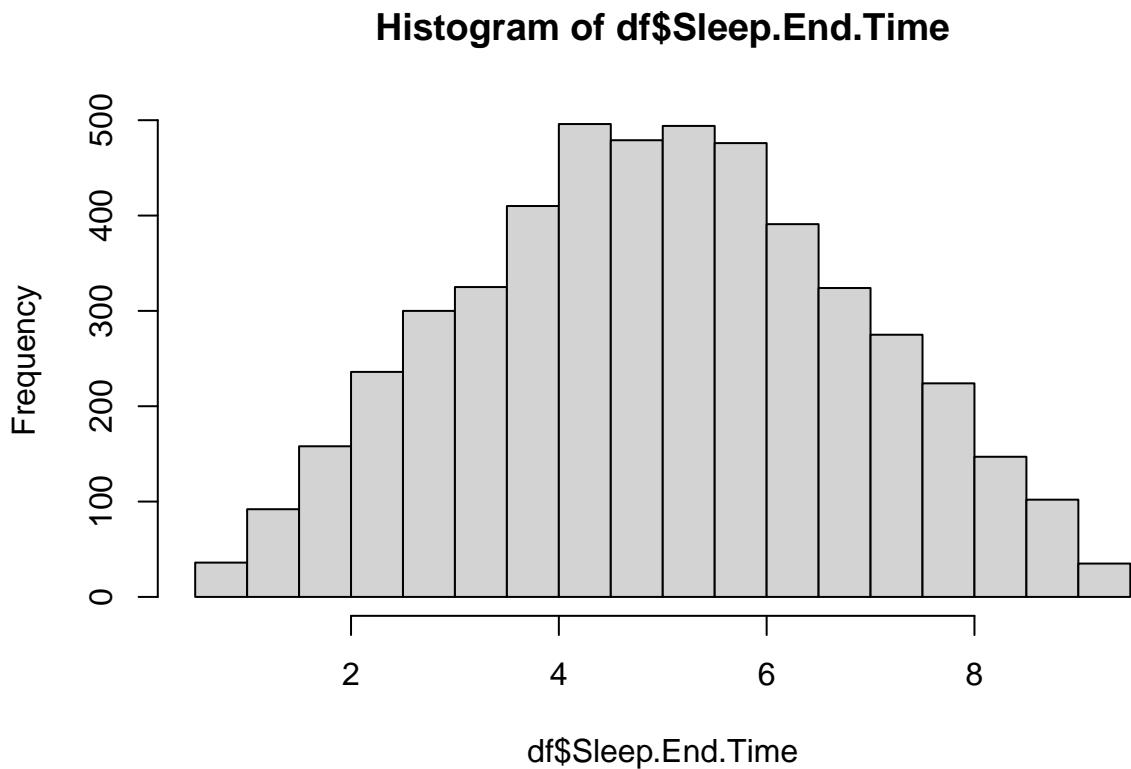
# Omit based on variation
hist(df$Sleep.Start.Time)

```

### Histogram of df\$Sleep.Start.Time



```
hist(df$Sleep.End.Time)
```



## Chose to omit start time

The distribution of end time is better, the prediction power is about the same

## Beginning of Total Work Hours Model

```

library(MASS)

##
## Attaching package: 'MASS'
## The following object is masked from 'package:dplyr':
##   select
library(car)

## Loading required package: carData
##
## Attaching package: 'car'
## The following object is masked from 'package:dplyr':
##   recode
## The following object is masked from 'package:purrr':
##   some

```

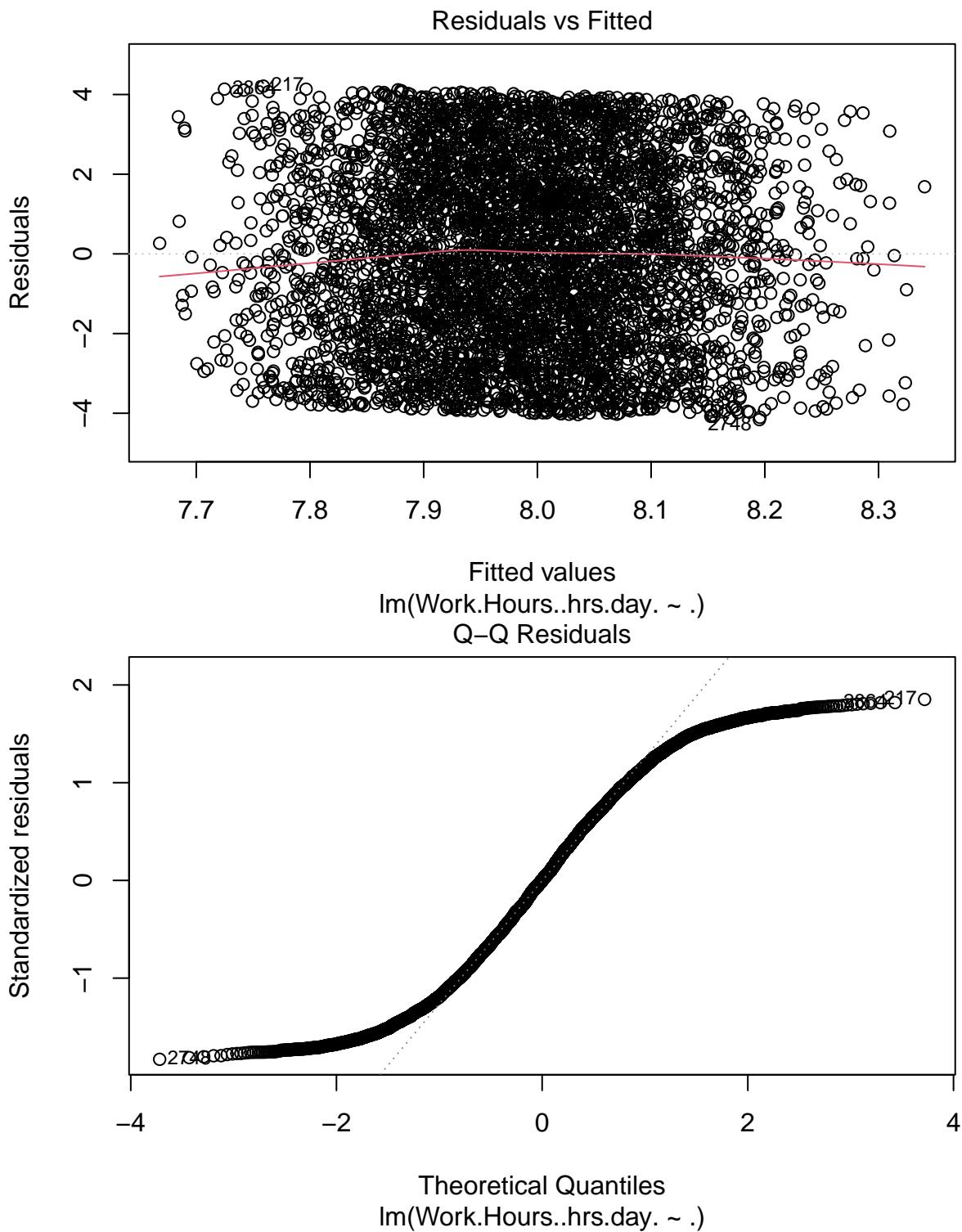
```

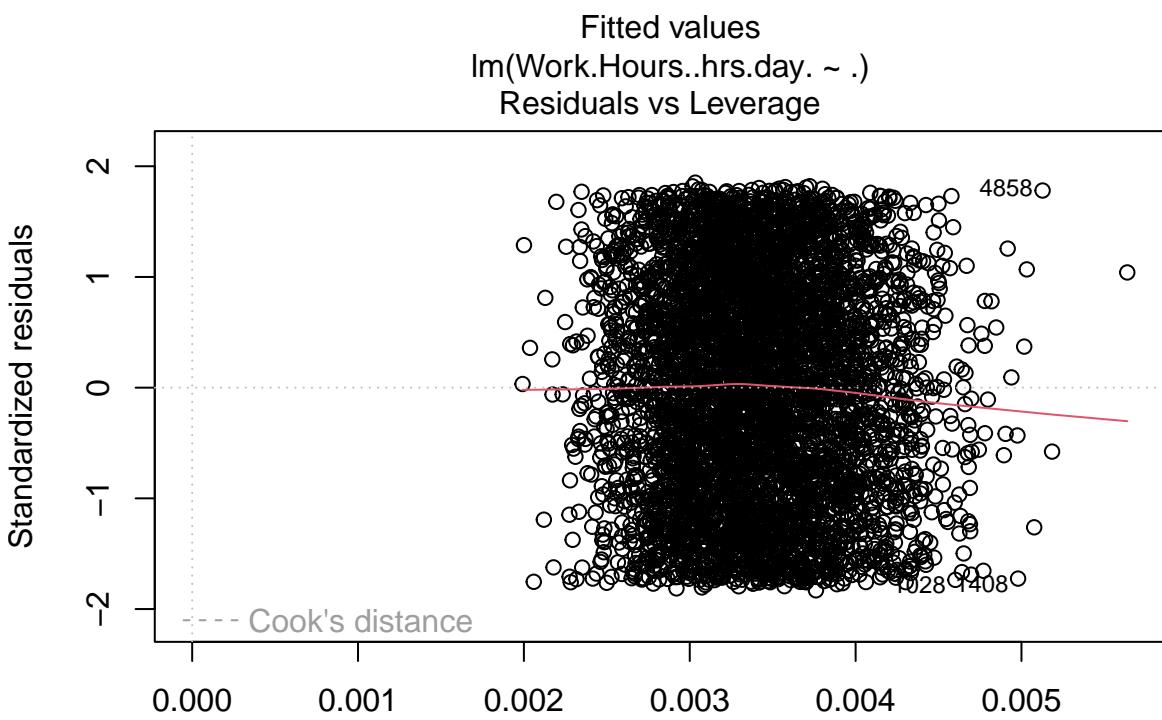
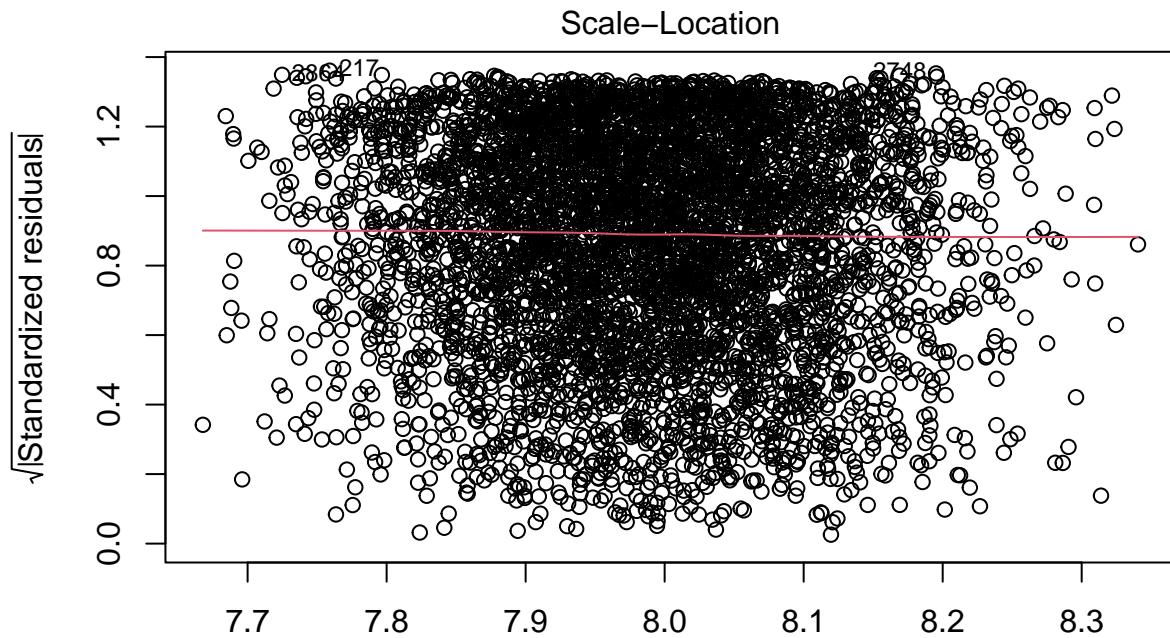
additive_model <- lm(Work.Hours..hrs.day. ~ ., data = df_filtered)
# basic variable only model
summary(additive_model)

## 
## Call:
## lm(formula = Work.Hours..hrs.day. ~ ., data = df_filtered)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -4.1605 -1.9384  0.0047  1.9224  4.2100 
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)               8.200e+00  2.437e-01 33.640 <2e-16 ***
## Age                     1.388e-03  2.620e-03  0.530  0.5963    
## GenderMale              -2.174e-02  7.826e-02 -0.278  0.7812    
## GenderOther              -9.012e-02  7.967e-02 -1.131  0.2580    
## Sleep.End.Time          2.547e-02  2.786e-02  0.914  0.3605    
## Total.Sleep.Hours       -6.075e-02 3.568e-02 -1.702  0.0887 .  
## Exercise..mins.day.     -3.288e-05 1.250e-03 -0.026  0.9790    
## Caffeine.Intake..mg.    3.299e-05  3.757e-04  0.088  0.9300    
## Screen.Time.Before.Bed..mins. 2.340e-04  6.190e-04  0.378  0.7054    
## Sleep.QualityCatMedium -4.426e-02  7.795e-02 -0.568  0.5702    
## Sleep.QualityCatHigh   -8.729e-02  8.367e-02 -1.043  0.2969    
## Productivity.ScoreCatMedium 6.745e-02  7.928e-02  0.851  0.3949    
## Productivity.ScoreCatHigh -3.319e-02  8.290e-02 -0.400  0.6889    
## Mood.ScoreCatMedium     3.748e-02  7.660e-02  0.489  0.6247    
## Mood.ScoreCatHigh       1.017e-01  8.398e-02  1.211  0.2258    
## Stress.LevelCatMedium   2.692e-03  7.836e-02  0.034  0.9726    
## Stress.LevelCatHigh     8.630e-02  8.280e-02  1.042  0.2973    
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 2.277 on 4983 degrees of freedom
## Multiple R-squared:  0.002257,  Adjusted R-squared:  -0.0009468 
## F-statistic: 0.7045 on 16 and 4983 DF,  p-value: 0.7922 

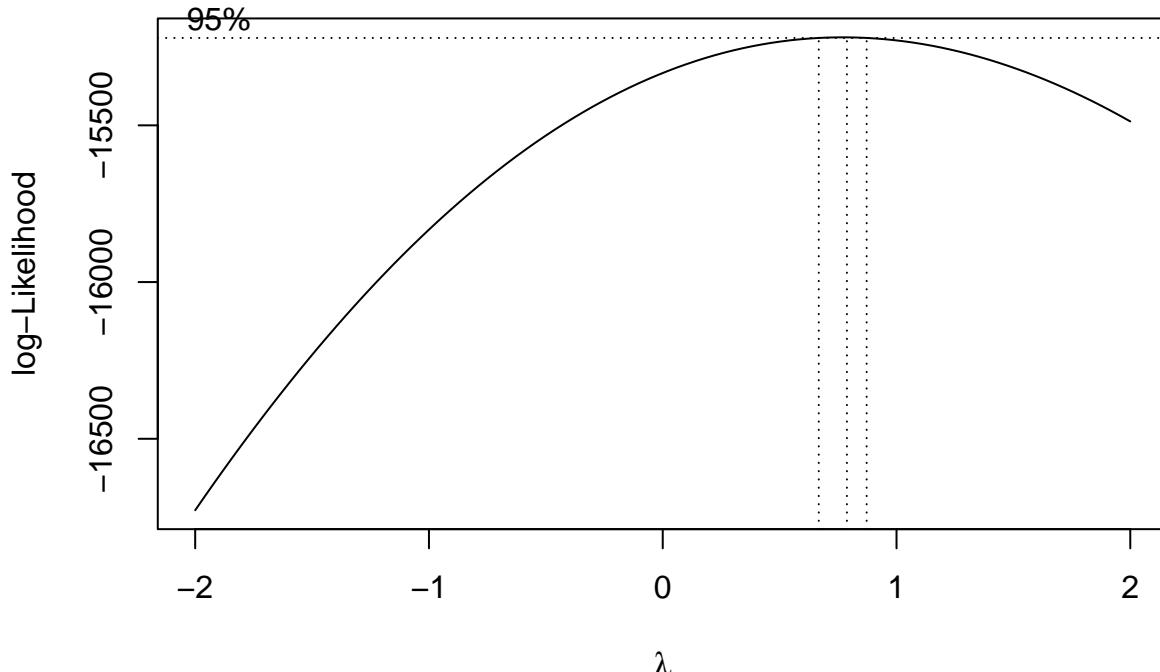
# residual plot
plot(additive_model)

```





```
# Transform the response for better results:
# response is strictly positive (box cox is appropriate - see notes on yeo-johnson in lecture)
# Box cox is theoretically used when we don't know what transformation is the most appropriate
bc <- boxcox(additive_model)
```



```

summary(bc)

##   Length Class  Mode
## x 100    -none- numeric
## y 100    -none- numeric
# Find lambda
lambda.hat <- bc$x[which.max(bc$y)]

## for transforming the data once you obtain lambda
df_bc <- df_filtered %>%
  mutate(Work.Hours..hrs.day. = (Work.Hours..hrs.day.^lambda.hat - 1) / lambda.hat)

# Basic model after box cox transformation
additive_bc <- lm(Work.Hours..hrs.day. ~ ., data = df_bc)
summary(additive_bc)

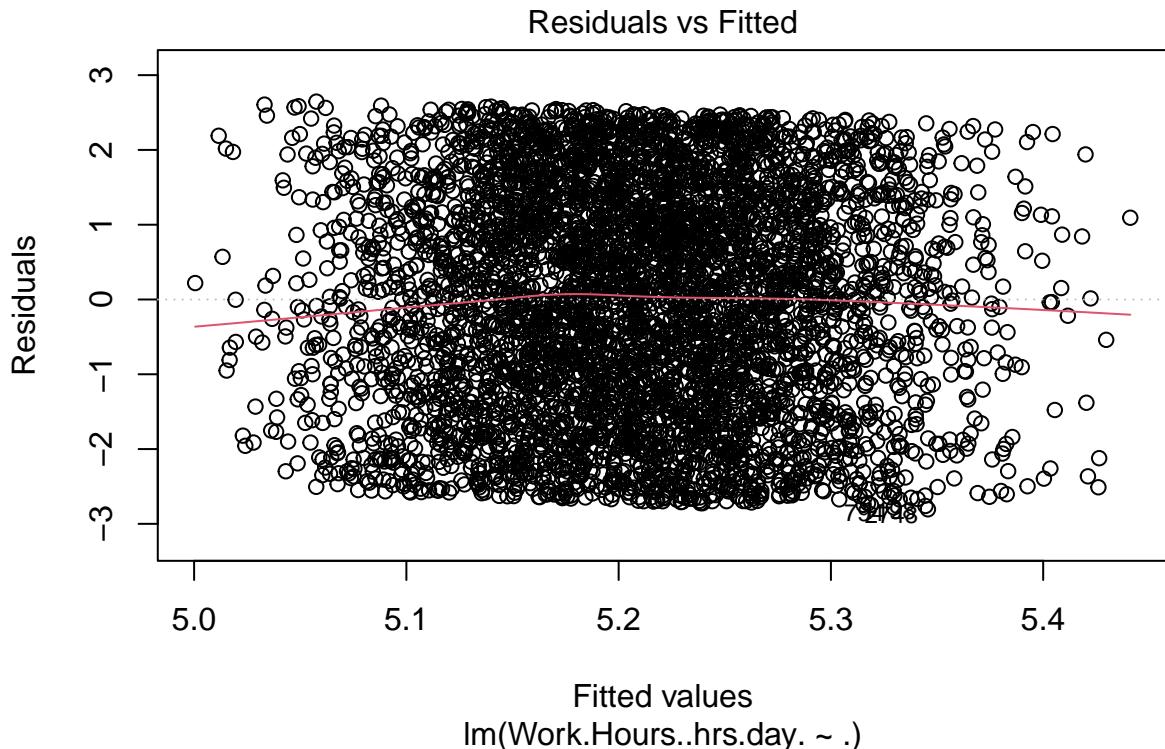
##
## Call:
## lm(formula = Work.Hours..hrs.day. ~ ., data = df_bc)
##
## Residuals:
##      Min       1Q     Median       3Q      Max 
## -2.80557 -1.24079  0.04865  1.25303  2.64576 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 5.347e+00  1.580e-01 33.838 <2e-16 ***
## Age         9.213e-04  1.698e-03  0.542  0.5875    
## GenderMale  -1.677e-02  5.074e-02 -0.330  0.7411    
## GenderOther -5.957e-02  5.165e-02 -1.153  0.2488    
## Sleep.End.Time 1.675e-02  1.806e-02  0.928  0.3537    
## Total.Sleep.Hours -3.937e-02  2.313e-02 -1.702  0.0888 .  

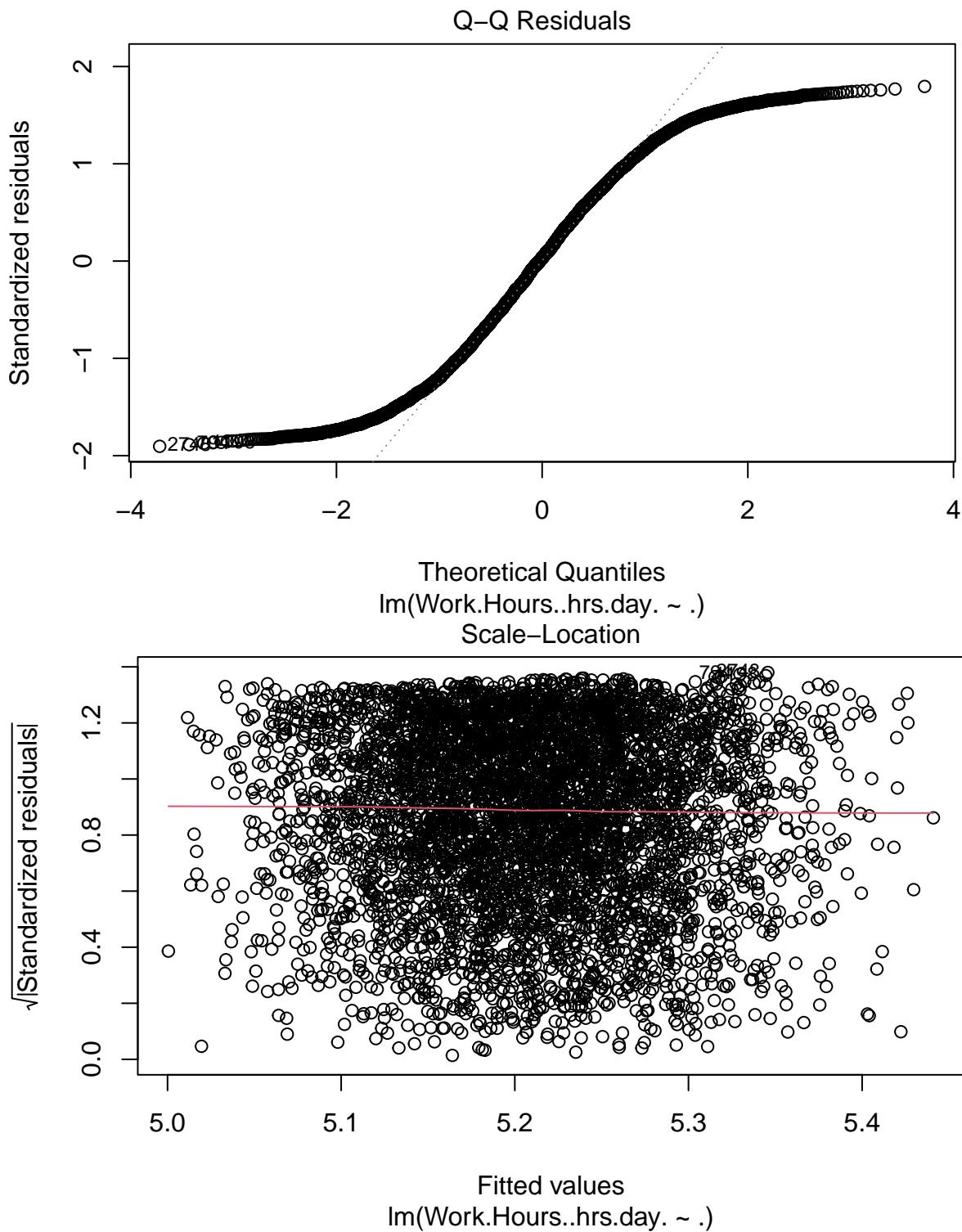
```

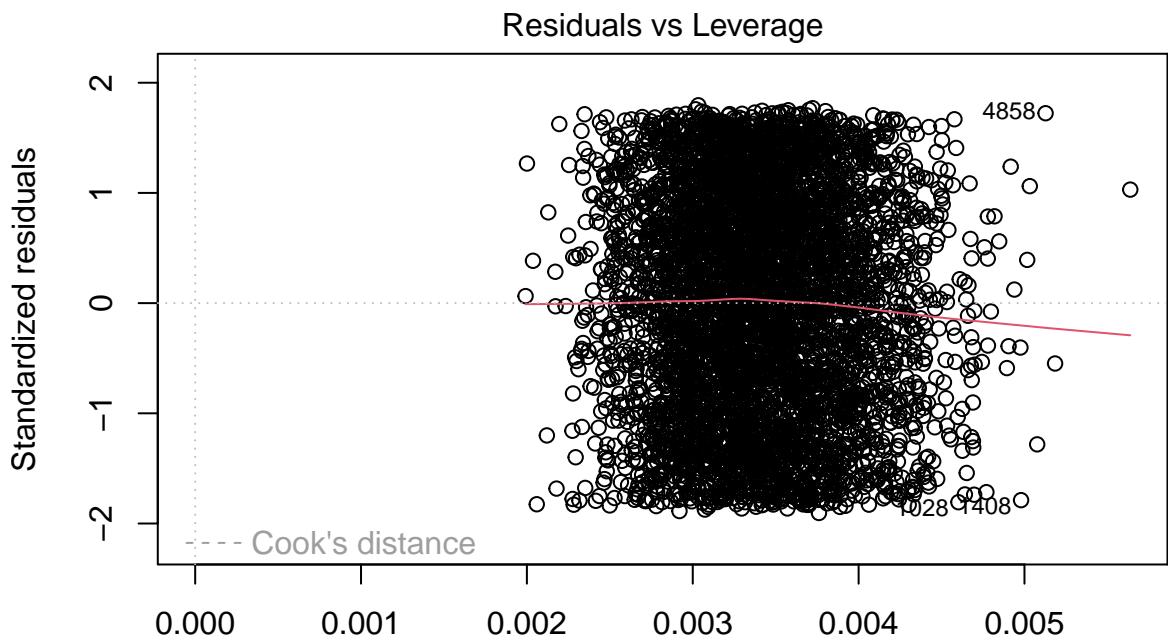
```

## Exercise..mins.day.          -9.722e-06 8.105e-04 -0.012  0.9904
## Caffeine.Intake..mg.        2.087e-05 2.435e-04  0.086  0.9317
## Screen.Time.Before.Bed..mins. 1.567e-04 4.013e-04  0.390  0.6963
## Sleep.QualityCatMedium     -3.164e-02 5.054e-02 -0.626  0.5313
## Sleep.QualityCatHigh       -5.768e-02 5.424e-02 -1.063  0.2877
## Productivity.ScoreCatMedium 4.164e-02 5.140e-02  0.810  0.4179
## Productivity.ScoreCatHigh   -2.469e-02 5.375e-02 -0.459  0.6460
## Mood.ScoreCatMedium        2.518e-02 4.966e-02  0.507  0.6121
## Mood.ScoreCatHigh          6.638e-02 5.445e-02  1.219  0.2228
## Stress.LevelCatMedium      2.571e-03 5.081e-02  0.051  0.9596
## Stress.LevelCatHigh         5.694e-02 5.368e-02  1.061  0.2888
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.476 on 4983 degrees of freedom
## Multiple R-squared:  0.002283,  Adjusted R-squared:  -0.0009203
## F-statistic: 0.7127 on 16 and 4983 DF,  p-value: 0.7837
plot(additive_bc)

```

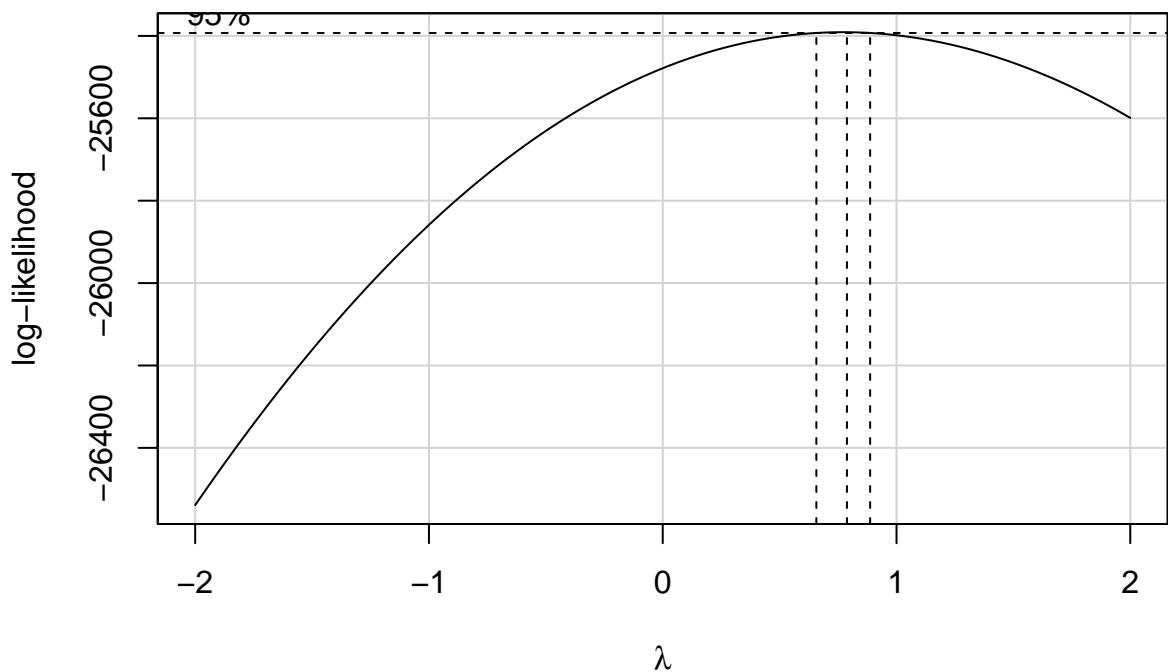






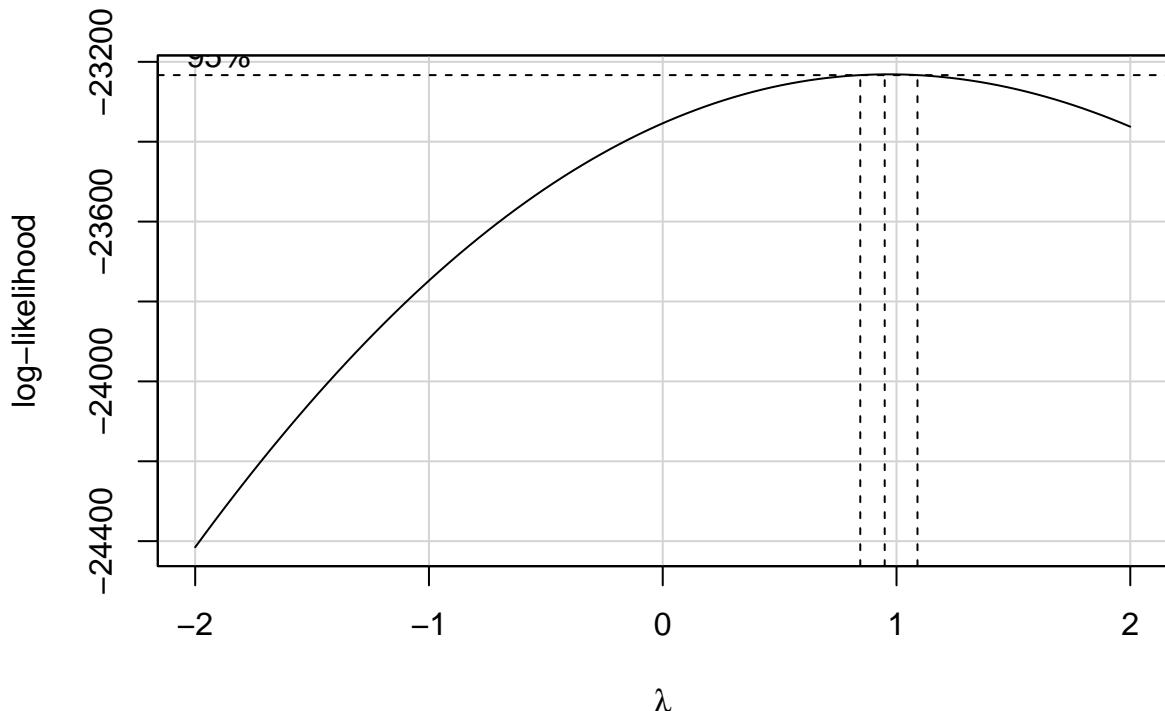
```
# Checking to see if lambda is now closer to 0
boxCox(additive_model,family="yjPower")
```

### Profile Log-likelihood



```
# To check if the log transformation worked (lambda should now be 0)
boxCox(additive_bc,family="yjPower")
```

## Profile Log-likelihood



```
# Check for colinearity issues
vif(additive_bc)
```

	GVIF	Df	GVIF^(1/(2*Df))
## Age	1.001600	1	1.000799
## Gender	1.005458	2	1.001362
## Sleep.End.Time	2.595692	1	1.611115
## Total.Sleep.Hours	2.595831	1	1.611158
## Exercise..mins.day.	1.003032	1	1.001515
## Caffeine.Intake..mg.	1.002563	1	1.001281
## Screen.Time.Before.Bed..mins.	1.002060	1	1.001030
## Sleep.QualityCat	1.003119	2	1.000779
## Productivity.ScoreCat	1.006021	2	1.001502
## Mood.ScoreCat	1.004532	2	1.001131
## Stress.LevelCat	1.006499	2	1.001621

Findings: The  $R^2$  actually got worse, do not move forward with the box cox transformed response

### Check for interactions in the model

```
# Basic full model
summary(additive_model)

##
## Call:
## lm(formula = Work.Hours..hrs.day. ~ ., data = df_filtered)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -10.0000  -2.5000   0.0000  10.0000  15.0000
```

```

## -4.1605 -1.9384  0.0047  1.9224  4.2100
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)               8.200e+00  2.437e-01 33.640 <2e-16 ***
## Age                      1.388e-03  2.620e-03  0.530  0.5963
## GenderMale                -2.174e-02 7.826e-02 -0.278  0.7812
## GenderOther                -9.012e-02 7.967e-02 -1.131  0.2580
## Sleep.End.Time             2.547e-02 2.786e-02  0.914  0.3605
## Total.Sleep.Hours          -6.075e-02 3.568e-02 -1.702  0.0887 .
## Exercise..mins.day.        -3.288e-05 1.250e-03 -0.026  0.9790
## Caffeine.Intake..mg.       3.299e-05 3.757e-04  0.088  0.9300
## Screen.Time.Before.Bed..mins. 2.340e-04 6.190e-04  0.378  0.7054
## Sleep.QualityCatMedium    -4.426e-02 7.795e-02 -0.568  0.5702
## Sleep.QualityCatHigh      -8.729e-02 8.367e-02 -1.043  0.2969
## Productivity.ScoreCatMedium 6.745e-02 7.928e-02  0.851  0.3949
## Productivity.ScoreCatHigh  -3.319e-02 8.290e-02 -0.400  0.6889
## Mood.ScoreCatMedium       3.748e-02 7.660e-02  0.489  0.6247
## Mood.ScoreCatHigh         1.017e-01 8.398e-02  1.211  0.2258
## Stress.LevelCatMedium     2.692e-03 7.836e-02  0.034  0.9726
## Stress.LevelCatHigh       8.630e-02 8.280e-02  1.042  0.2973
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.277 on 4983 degrees of freedom
## Multiple R-squared:  0.002257, Adjusted R-squared:  -0.0009468
## F-statistic: 0.7045 on 16 and 4983 DF,  p-value: 0.7922
full_model <- lm(Work.Hours..hrs.day. ~ .()^2, data = df_filtered)
summary(full_model)

##
## Call:
## lm(formula = Work.Hours..hrs.day. ~ .()^2, data = df_filtered)
##
## Residuals:
##      Min       1Q       Median      3Q      Max 
## -4.5916 -1.8845  0.0243  1.8236  4.6020 
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)               6.556e+00  1.238e+00
## Age                      3.700e-02  1.820e-02
## GenderMale                -2.474e-01 5.886e-01
## GenderOther                1.894e-01 6.026e-01
## Sleep.End.Time             -5.212e-02 1.821e-01
## Total.Sleep.Hours          9.892e-02 2.103e-01
## Exercise..mins.day.        2.146e-02 9.393e-03
## Caffeine.Intake..mg.       -4.121e-04 2.849e-03
## Screen.Time.Before.Bed..mins. -4.223e-03 4.567e-03
## Sleep.QualityCatMedium    1.152e+00 5.819e-01
## Sleep.QualityCatHigh      8.976e-01 6.273e-01
## Productivity.ScoreCatMedium 5.206e-01 5.923e-01
## Productivity.ScoreCatHigh  2.038e-01 6.217e-01
## Mood.ScoreCatMedium       -1.305e+00 5.771e-01

```

## Mood.ScoreCatHigh	-2.059e-01	6.340e-01
## Stress.LevelCatMedium	9.441e-01	5.800e-01
## Stress.LevelCatHigh	2.733e-03	6.174e-01
## Age:GenderMale	1.924e-03	6.452e-03
## Age:GenderOther	8.523e-03	6.559e-03
## Age:Sleep.End.Time	1.727e-03	2.265e-03
## Age:Total.Sleep.Hours	-3.621e-03	2.909e-03
## Age:Exercise..mins.day.	-3.151e-04	1.037e-04
## Age:Caffeine.Intake..mg.	2.872e-05	3.101e-05
## Age:Screen.Time.Before.Bed..mins.	2.047e-05	5.078e-05
## Age:Sleep.QualityCatMedium	-1.442e-02	6.403e-03
## Age:Sleep.QualityCatHigh	-1.215e-02	6.834e-03
## Age:Productivity.ScoreCatMedium	-7.570e-03	6.548e-03
## Age:Productivity.ScoreCatHigh	-1.338e-03	6.807e-03
## Age:Mood.ScoreCatMedium	3.280e-03	6.300e-03
## Age:Mood.ScoreCatHigh	-1.263e-02	6.878e-03
## Age:Stress.LevelCatMedium	-2.685e-03	6.473e-03
## Age:Stress.LevelCatHigh	4.744e-03	6.768e-03
## GenderMale:Sleep.End.Time	3.496e-04	6.870e-02
## GenderOther:Sleep.End.Time	4.004e-02	6.949e-02
## GenderMale:Total.Sleep.Hours	1.031e-02	8.836e-02
## GenderOther:Total.Sleep.Hours	-1.118e-01	8.915e-02
## GenderMale:Exercise..mins.day.	-2.267e-03	3.068e-03
## GenderOther:Exercise..mins.day.	-2.398e-03	3.106e-03
## GenderMale:Caffeine.Intake..mg.	6.678e-04	9.215e-04
## GenderOther:Caffeine.Intake..mg.	3.584e-04	9.338e-04
## GenderMale:Screen.Time.Before.Bed..mins.	-1.489e-03	1.509e-03
## GenderOther:Screen.Time.Before.Bed..mins.	1.262e-04	1.550e-03
## GenderMale:Sleep.QualityCatMedium	3.745e-01	1.915e-01
## GenderOther:Sleep.QualityCatMedium	1.644e-01	1.944e-01
## GenderMale:Sleep.QualityCatHigh	2.469e-01	2.052e-01
## GenderOther:Sleep.QualityCatHigh	1.631e-01	2.097e-01
## GenderMale:Productivity.ScoreCatMedium	3.889e-01	1.948e-01
## GenderOther:Productivity.ScoreCatMedium	2.432e-01	1.988e-01
## GenderMale:Productivity.ScoreCatHigh	9.892e-02	2.035e-01
## GenderOther:Productivity.ScoreCatHigh	3.143e-01	2.064e-01
## GenderMale:Mood.ScoreCatMedium	-2.064e-01	1.869e-01
## GenderOther:Mood.ScoreCatMedium	-2.023e-01	1.912e-01
## GenderMale:Mood.ScoreCatHigh	-1.429e-01	2.069e-01
## GenderOther:Mood.ScoreCatHigh	-3.138e-01	2.098e-01
## GenderMale:Stress.LevelCatMedium	-2.526e-01	1.926e-01
## GenderOther:Stress.LevelCatMedium	-3.486e-01	1.943e-01
## GenderMale:Stress.LevelCatHigh	1.034e-01	2.010e-01
## GenderOther:Stress.LevelCatHigh	9.570e-02	2.066e-01
## Sleep.End.Time:Total.Sleep.Hours	3.911e-03	1.282e-02
## Sleep.End.Time:Exercise..mins.day.	-3.712e-04	1.089e-03
## Sleep.End.Time:Caffeine.Intake..mg.	5.010e-04	3.316e-04
## Sleep.End.Time:Screen.Time.Before.Bed..mins.	-1.848e-04	5.417e-04
## Sleep.End.Time:Sleep.QualityCatMedium	6.717e-02	6.788e-02
## Sleep.End.Time:Sleep.QualityCatHigh	9.933e-02	7.271e-02
## Sleep.End.Time:Productivity.ScoreCatMedium	2.348e-02	6.899e-02
## Sleep.End.Time:Productivity.ScoreCatHigh	-5.158e-03	7.239e-02
## Sleep.End.Time:Mood.ScoreCatMedium	-1.666e-01	6.670e-02
## Sleep.End.Time:Mood.ScoreCatHigh	-5.951e-02	7.310e-02

## Sleep.End.Time:Stress.LevelCatMedium	-5.982e-02	6.852e-02
## Sleep.End.Time:Stress.LevelCatHigh	-7.940e-02	7.194e-02
## Total.Sleep.Hours:Exercise..mins.day.	-6.578e-04	1.399e-03
## Total.Sleep.Hours:Caffeine.Intake..mg.	-5.960e-04	4.231e-04
## Total.Sleep.Hours:Screen.Time.Before.Bed..mins.	6.452e-04	6.928e-04
## Total.Sleep.Hours:Sleep.QualityCatMedium	-1.306e-01	8.757e-02
## Total.Sleep.Hours:Sleep.QualityCatHigh	-1.686e-01	9.313e-02
## Total.Sleep.Hours:Productivity.ScoreCatMedium	-6.038e-02	8.890e-02
## Total.Sleep.Hours:Productivity.ScoreCatHigh	-2.419e-02	9.227e-02
## Total.Sleep.Hours:Mood.ScoreCatMedium	2.670e-01	8.556e-02
## Total.Sleep.Hours:Mood.ScoreCatHigh	1.317e-01	9.516e-02
## Total.Sleep.Hours:Stress.LevelCatMedium	1.862e-02	8.714e-02
## Total.Sleep.Hours:Stress.LevelCatHigh	8.212e-02	9.244e-02
## Exercise..mins.day.:Caffeine.Intake..mg.	3.402e-06	1.496e-05
## Exercise..mins.day.:Screen.Time.Before.Bed..mins.	-5.353e-06	2.434e-05
## Exercise..mins.day.:Sleep.QualityCatMedium	-4.325e-03	3.035e-03
## Exercise..mins.day.:Sleep.QualityCatHigh	1.481e-03	3.281e-03
## Exercise..mins.day.:Productivity.ScoreCatMedium	6.667e-04	3.093e-03
## Exercise..mins.day.:Productivity.ScoreCatHigh	-1.761e-04	3.247e-03
## Exercise..mins.day.:Mood.ScoreCatMedium	8.591e-04	2.992e-03
## Exercise..mins.day.:Mood.ScoreCatHigh	2.166e-03	3.265e-03
## Exercise..mins.day.:Stress.LevelCatMedium	-4.285e-03	3.085e-03
## Exercise..mins.day.:Stress.LevelCatHigh	1.652e-03	3.249e-03
## Caffeine.Intake..mg.:Screen.Time.Before.Bed..mins.	-4.627e-06	7.275e-06
## Caffeine.Intake..mg.:Sleep.QualityCatMedium	1.228e-03	9.188e-04
## Caffeine.Intake..mg.:Sleep.QualityCatHigh	8.203e-04	9.836e-04
## Caffeine.Intake..mg.:Productivity.ScoreCatMedium	3.245e-04	9.341e-04
## Caffeine.Intake..mg.:Productivity.ScoreCatHigh	7.880e-04	9.766e-04
## Caffeine.Intake..mg.:Mood.ScoreCatMedium	1.123e-04	9.123e-04
## Caffeine.Intake..mg.:Mood.ScoreCatHigh	1.254e-03	9.853e-04
## Caffeine.Intake..mg.:Stress.LevelCatMedium	-2.262e-04	9.248e-04
## Caffeine.Intake..mg.:Stress.LevelCatHigh	-1.625e-03	9.828e-04
## Screen.Time.Before.Bed..mins.:Sleep.QualityCatMedium	1.697e-04	1.507e-03
## Screen.Time.Before.Bed..mins.:Sleep.QualityCatHigh	-9.777e-04	1.626e-03
## Screen.Time.Before.Bed..mins.:Productivity.ScoreCatMedium	-7.558e-04	1.527e-03
## Screen.Time.Before.Bed..mins.:Productivity.ScoreCatHigh	1.370e-03	1.597e-03
## Screen.Time.Before.Bed..mins.:Mood.ScoreCatMedium	2.317e-03	1.487e-03
## Screen.Time.Before.Bed..mins.:Mood.ScoreCatHigh	1.367e-03	1.630e-03
## Screen.Time.Before.Bed..mins.:Stress.LevelCatMedium	3.231e-04	1.528e-03
## Screen.Time.Before.Bed..mins.:Stress.LevelCatHigh	-2.706e-04	1.617e-03
## Sleep.QualityCatMedium:Productivity.ScoreCatMedium	-2.754e-01	1.942e-01
## Sleep.QualityCatHigh:Productivity.ScoreCatMedium	-6.495e-02	2.082e-01
## Sleep.QualityCatMedium:Productivity.ScoreCatHigh	-3.971e-01	2.042e-01
## Sleep.QualityCatHigh:Productivity.ScoreCatHigh	-2.522e-01	2.170e-01
## Sleep.QualityCatMedium:Mood.ScoreCatMedium	2.215e-01	1.868e-01
## Sleep.QualityCatHigh:Mood.ScoreCatMedium	4.540e-02	2.015e-01
## Sleep.QualityCatMedium:Mood.ScoreCatHigh	4.269e-01	2.044e-01
## Sleep.QualityCatHigh:Mood.ScoreCatHigh	1.779e-01	2.194e-01
## Sleep.QualityCatMedium:Stress.LevelCatMedium	-4.120e-01	1.917e-01
## Sleep.QualityCatHigh:Stress.LevelCatMedium	-1.079e-01	2.069e-01
## Sleep.QualityCatMedium:Stress.LevelCatHigh	-1.846e-01	2.016e-01
## Sleep.QualityCatHigh:Stress.LevelCatHigh	2.732e-02	2.169e-01
## Productivity.ScoreCatMedium:Mood.ScoreCatMedium	1.945e-01	1.908e-01
## Productivity.ScoreCatHigh:Mood.ScoreCatMedium	-3.288e-01	1.999e-01

```

## Productivity.ScoreCatMedium:Mood.ScoreCatHigh -4.952e-02 2.109e-01
## Productivity.ScoreCatHigh:Mood.ScoreCatHigh -1.688e-01 2.190e-01
## Productivity.ScoreCatMedium:Stress.LevelCatMedium -5.247e-02 1.946e-01
## Productivity.ScoreCatHigh:Stress.LevelCatMedium -1.789e-02 2.041e-01
## Productivity.ScoreCatMedium:Stress.LevelCatHigh -5.005e-03 2.064e-01
## Productivity.ScoreCatHigh:Stress.LevelCatHigh 2.118e-01 2.136e-01
## Mood.ScoreCatMedium:Stress.LevelCatMedium -2.232e-02 1.881e-01
## Mood.ScoreCatHigh:Stress.LevelCatMedium -1.786e-01 2.053e-01
## Mood.ScoreCatMedium:Stress.LevelCatHigh -3.329e-02 1.984e-01
## Mood.ScoreCatHigh:Stress.LevelCatHigh -4.859e-01 2.185e-01
##
t value Pr(>|t|)
## (Intercept) 5.294 1.25e-07 ***
## Age 2.033 0.04207 *
## GenderMale -0.420 0.67424
## GenderOther 0.314 0.75328
## Sleep.End.Time -0.286 0.77467
## Total.Sleep.Hours 0.470 0.63819
## Exercise..mins.day. 2.285 0.02235 *
## Caffeine.Intake..mg. -0.145 0.88497
## Screen.Time.Before.Bed..mins. -0.925 0.35521
## Sleep.QualityCatMedium 1.979 0.04782 *
## Sleep.QualityCatHigh 1.431 0.15254
## Productivity.ScoreCatMedium 0.879 0.37949
## Productivity.ScoreCatHigh 0.328 0.74307
## Mood.ScoreCatMedium -2.262 0.02375 *
## Mood.ScoreCatHigh -0.325 0.74531
## Stress.LevelCatMedium 1.628 0.10363
## Stress.LevelCatHigh 0.004 0.99647
## Age:GenderMale 0.298 0.76552
## Age:GenderOther 1.299 0.19387
## Age:Sleep.End.Time 0.762 0.44582
## Age:Total.Sleep.Hours -1.245 0.21330
## Age:Exercise..mins.day. -3.040 0.00238 **
## Age:Caffeine.Intake..mg. 0.926 0.35446
## Age:Screen.Time.Before.Bed..mins. 0.403 0.68685
## Age:Sleep.QualityCatMedium -2.253 0.02431 *
## Age:Sleep.QualityCatHigh -1.778 0.07549 .
## Age:Productivity.ScoreCatMedium -1.156 0.24772
## Age:Productivity.ScoreCatHigh -0.197 0.84414
## Age:Mood.ScoreCatMedium 0.521 0.60264
## Age:Mood.ScoreCatHigh -1.836 0.06644 .
## Age:Stress.LevelCatMedium -0.415 0.67833
## Age:Stress.LevelCatHigh 0.701 0.48333
## GenderMale:Sleep.End.Time 0.005 0.99594
## GenderOther:Sleep.End.Time 0.576 0.56452
## GenderMale:Total.Sleep.Hours 0.117 0.90710
## GenderOther:Total.Sleep.Hours -1.254 0.20985
## GenderMale:Exercise..mins.day. -0.739 0.45998
## GenderOther:Exercise..mins.day. -0.772 0.44019
## GenderMale:Caffeine.Intake..mg. 0.725 0.46865
## GenderOther:Caffeine.Intake..mg. 0.384 0.70116
## GenderMale:Screen.Time.Before.Bed..mins. -0.987 0.32393
## GenderOther:Screen.Time.Before.Bed..mins. 0.081 0.93510
## GenderMale:Sleep.QualityCatMedium 1.956 0.05053 .

```

## GenderOther:Sleep.QualityCatMedium	0.845	0.39799
## GenderMale:Sleep.QualityCatHigh	1.203	0.22898
## GenderOther:Sleep.QualityCatHigh	0.778	0.43675
## GenderMale:Productivity.ScoreCatMedium	1.997	0.04592 *
## GenderOther:Productivity.ScoreCatMedium	1.223	0.22134
## GenderMale:Productivity.ScoreCatHigh	0.486	0.62697
## GenderOther:Productivity.ScoreCatHigh	1.522	0.12800
## GenderMale:Mood.ScoreCatMedium	-1.104	0.26943
## GenderOther:Mood.ScoreCatMedium	-1.058	0.29005
## GenderMale:Mood.ScoreCatHigh	-0.691	0.48990
## GenderOther:Mood.ScoreCatHigh	-1.496	0.13477
## GenderMale:Stress.LevelCatMedium	-1.312	0.18961
## GenderOther:Stress.LevelCatMedium	-1.794	0.07291 .
## GenderMale:Stress.LevelCatHigh	0.514	0.60713
## GenderOther:Stress.LevelCatHigh	0.463	0.64332
## Sleep.End.Time:Total.Sleep.Hours	0.305	0.76039
## Sleep.End.Time:Exercise..mins.day.	-0.341	0.73315
## Sleep.End.Time:Caffeine.Intake..mg.	1.511	0.13090
## Sleep.End.Time:Screen.Time.Before.Bed..mins.	-0.341	0.73301
## Sleep.End.Time:Sleep.QualityCatMedium	0.990	0.32243
## Sleep.End.Time:Sleep.QualityCatHigh	1.366	0.17196
## Sleep.End.Time:Productivity.ScoreCatMedium	0.340	0.73362
## Sleep.End.Time:Productivity.ScoreCatHigh	-0.071	0.94321
## Sleep.End.Time:Mood.ScoreCatMedium	-2.497	0.01255 *
## Sleep.End.Time:Mood.ScoreCatHigh	-0.814	0.41566
## Sleep.End.Time:Stress.LevelCatMedium	-0.873	0.38273
## Sleep.End.Time:Stress.LevelCatHigh	-1.104	0.26984
## Total.Sleep.Hours:Exercise..mins.day.	-0.470	0.63832
## Total.Sleep.Hours:Caffeine.Intake..mg.	-1.409	0.15898
## Total.Sleep.Hours:Screen.Time.Before.Bed..mins.	0.931	0.35176
## Total.Sleep.Hours:Sleep.QualityCatMedium	-1.492	0.13582
## Total.Sleep.Hours:Sleep.QualityCatHigh	-1.810	0.07030 .
## Total.Sleep.Hours:Productivity.ScoreCatMedium	-0.679	0.49706
## Total.Sleep.Hours:Productivity.ScoreCatHigh	-0.262	0.79324
## Total.Sleep.Hours:Mood.ScoreCatMedium	3.121	0.00182 **
## Total.Sleep.Hours:Mood.ScoreCatHigh	1.384	0.16632
## Total.Sleep.Hours:Stress.LevelCatMedium	0.214	0.83079
## Total.Sleep.Hours:Stress.LevelCatHigh	0.888	0.37441
## Exercise..mins.day.:Caffeine.Intake..mg.	0.227	0.82017
## Exercise..mins.day.:Screen.Time.Before.Bed..mins.	-0.220	0.82596
## Exercise..mins.day.:Sleep.QualityCatMedium	-1.425	0.15424
## Exercise..mins.day.:Sleep.QualityCatHigh	0.451	0.65170
## Exercise..mins.day.:Productivity.ScoreCatMedium	0.216	0.82932
## Exercise..mins.day.:Productivity.ScoreCatHigh	-0.054	0.95676
## Exercise..mins.day.:Mood.ScoreCatMedium	0.287	0.77404
## Exercise..mins.day.:Mood.ScoreCatHigh	0.663	0.50710
## Exercise..mins.day.:Stress.LevelCatMedium	-1.389	0.16492
## Exercise..mins.day.:Stress.LevelCatHigh	0.509	0.61107
## Caffeine.Intake..mg.:Screen.Time.Before.Bed..mins.	-0.636	0.52478
## Caffeine.Intake..mg.:Sleep.QualityCatMedium	1.336	0.18159
## Caffeine.Intake..mg.:Sleep.QualityCatHigh	0.834	0.40431
## Caffeine.Intake..mg.:Productivity.ScoreCatMedium	0.347	0.72834
## Caffeine.Intake..mg.:Productivity.ScoreCatHigh	0.807	0.41978
## Caffeine.Intake..mg.:Mood.ScoreCatMedium	0.123	0.90208

```

## Caffeine.Intake..mg.:Mood.ScoreCatHigh           1.273  0.20303
## Caffeine.Intake..mg.:Stress.LevelCatMedium      -0.245  0.80677
## Caffeine.Intake..mg.:Stress.LevelCatHigh         -1.653  0.09835 .
## Screen.Time.Before.Bed..mins.:Sleep.QualityCatMedium  0.113  0.91037
## Screen.Time.Before.Bed..mins.:Sleep.QualityCatHigh -0.601  0.54762
## Screen.Time.Before.Bed..mins.:Productivity.ScoreCatMedium -0.495  0.62057
## Screen.Time.Before.Bed..mins.:Productivity.ScoreCatHigh  0.858  0.39110
## Screen.Time.Before.Bed..mins.:Mood.ScoreCatMedium    1.558  0.11926
## Screen.Time.Before.Bed..mins.:Mood.ScoreCatHigh       0.839  0.40174
## Screen.Time.Before.Bed..mins.:Stress.LevelCatMedium   0.211  0.83256
## Screen.Time.Before.Bed..mins.:Stress.LevelCatHigh     -0.167  0.86710
## Sleep.QualityCatMedium:Productivity.ScoreCatMedium   -1.418  0.15631
## Sleep.QualityCatHigh:Productivity.ScoreCatMedium      -0.312  0.75504
## Sleep.QualityCatMedium:Productivity.ScoreCatHigh       -1.944  0.05192 .
## Sleep.QualityCatHigh:Productivity.ScoreCatHigh         -1.162  0.24510
## Sleep.QualityCatMedium:Mood.ScoreCatMedium            1.186  0.23580
## Sleep.QualityCatHigh:Mood.ScoreCatMedium              0.225  0.82177
## Sleep.QualityCatMedium:Mood.ScoreCatHigh               2.089  0.03674 *
## Sleep.QualityCatHigh:Mood.ScoreCatHigh                 0.811  0.41764
## Sleep.QualityCatMedium:Stress.LevelCatMedium          -2.150  0.03163 *
## Sleep.QualityCatHigh:Stress.LevelCatMedium            -0.522  0.60203
## Sleep.QualityCatMedium:Stress.LevelCatHigh             -0.916  0.35989
## Sleep.QualityCatHigh:Stress.LevelCatHigh              0.126  0.89976
## Productivity.ScoreCatMedium:Mood.ScoreCatMedium        1.019  0.30804
## Productivity.ScoreCatHigh:Mood.ScoreCatMedium          -1.644  0.10015
## Productivity.ScoreCatMedium:Mood.ScoreCatHigh          -0.235  0.81435
## Productivity.ScoreCatHigh:Mood.ScoreCatHigh            -0.771  0.44077
## Productivity.ScoreCatMedium:Stress.LevelCatMedium      -0.270  0.78746
## Productivity.ScoreCatHigh:Stress.LevelCatMedium        -0.088  0.93015
## Productivity.ScoreCatMedium:Stress.LevelCatHigh         -0.024  0.98066
## Productivity.ScoreCatHigh:Stress.LevelCatHigh           0.992  0.32136
## Mood.ScoreCatMedium:Stress.LevelCatMedium              -0.119  0.90555
## Mood.ScoreCatHigh:Stress.LevelCatMedium                -0.870  0.38426
## Mood.ScoreCatMedium:Stress.LevelCatHigh                 -0.168  0.86676
## Mood.ScoreCatHigh:Stress.LevelCatHigh                  -2.223  0.02625 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.274 on 4868 degrees of freedom
## Multiple R-squared:  0.02793,   Adjusted R-squared:  0.001767
## F-statistic: 1.068 on 131 and 4868 DF,  p-value: 0.2854
coefs <- summary(full_model)$coefficients
vars <- rownames(coefs)[which(coefs[, 4] < 0.05)]
print("Significant Interactions:")

## [1] "Significant Interactions:"
vars

## [1] "(Intercept)"
## [2] "Age"
## [3] "Exercise..mins.day."
## [4] "Sleep.QualityCatMedium"
## [5] "Mood.ScoreCatMedium"
## [6] "Age:Exercise..mins.day."

```

```

## [7] "Age:Sleep.QualityCatMedium"
## [8] "GenderMale:Productivity.ScoreCatMedium"
## [9] "Sleep.End.Time:Mood.ScoreCatMedium"
## [10] "Total.Sleep.Hours:Mood.ScoreCatMedium"
## [11] "Sleep.QualityCatMedium:Mood.ScoreCatHigh"
## [12] "Sleep.QualityCatMedium:Stress.LevelCatMedium"
## [13] "Mood.ScoreCatHigh:Stress.LevelCatHigh"

anova(full_model, additive_model)

## Analysis of Variance Table

## Model 1: Work.Hours..hrs.day. ~ (Age + Gender + Sleep.End.Time + Total.Sleep.Hours +
##     Exercise..mins.day. + Caffeine.Intake..mg. + Screen.Time.Before.Bed..mins. +
##     Sleep.QualityCat + Productivity.ScoreCat + Mood.ScoreCat +
##     Stress.LevelCat)^2

## Model 2: Work.Hours..hrs.day. ~ Age + Gender + Sleep.End.Time + Total.Sleep.Hours +
##     Exercise..mins.day. + Caffeine.Intake..mg. + Screen.Time.Before.Bed..mins. +
##     Sleep.QualityCat + Productivity.ScoreCat + Mood.ScoreCat +
##     Stress.LevelCat

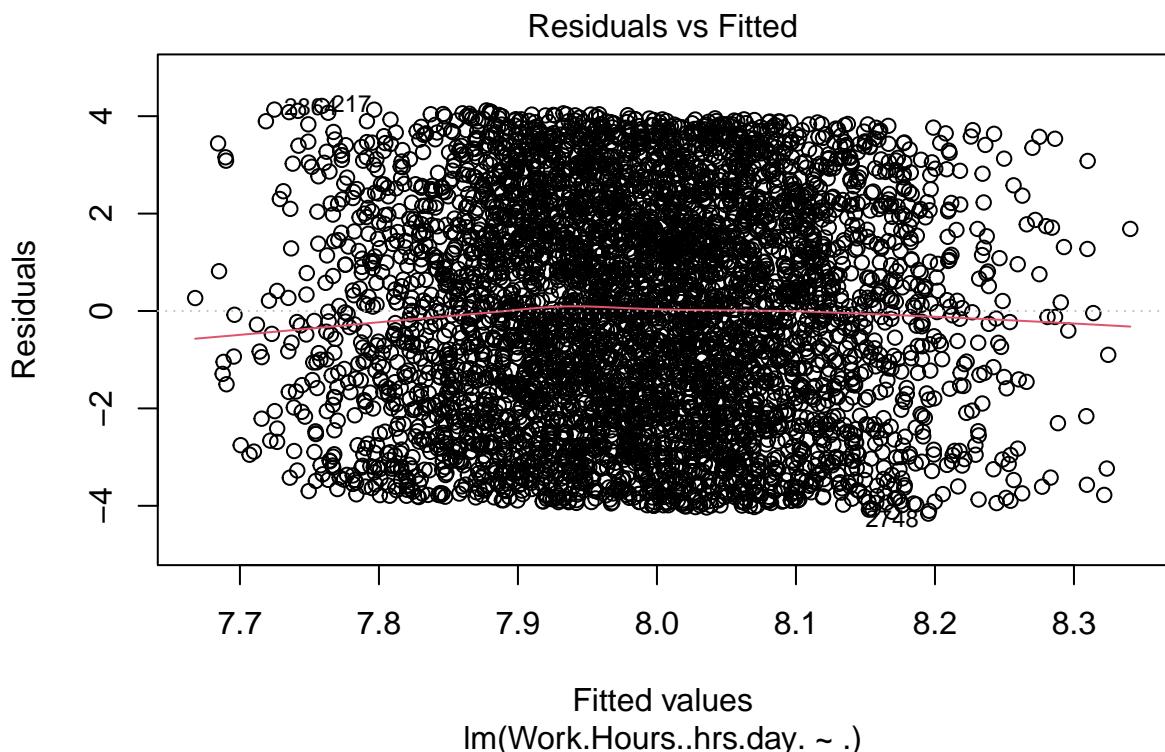
## Res.Df   RSS   Df Sum of Sq    F Pr(>F)
## 1   4868 25167
## 2   4983 25831 -115   -664.56 1.1178 0.1863

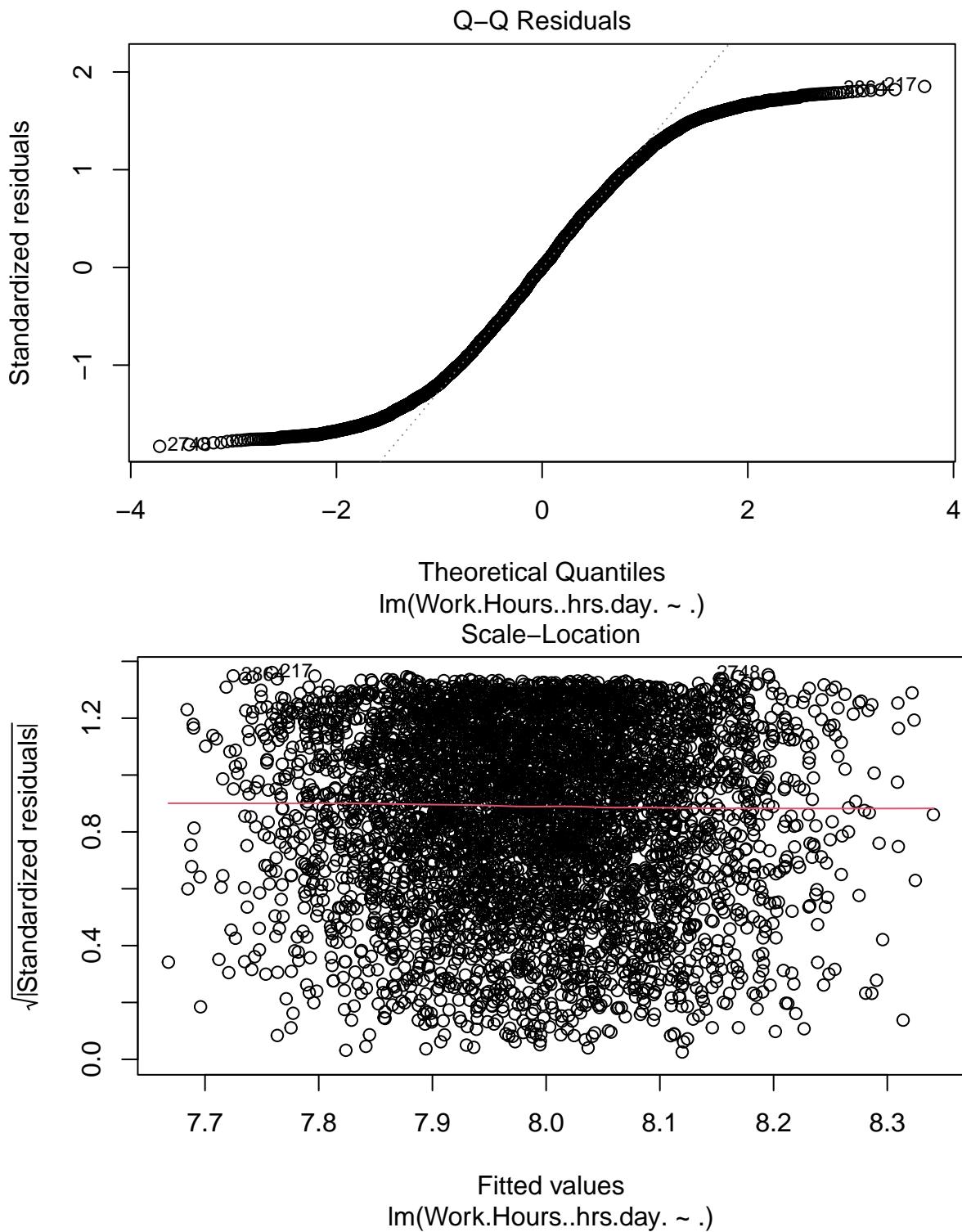
```

### Findings:

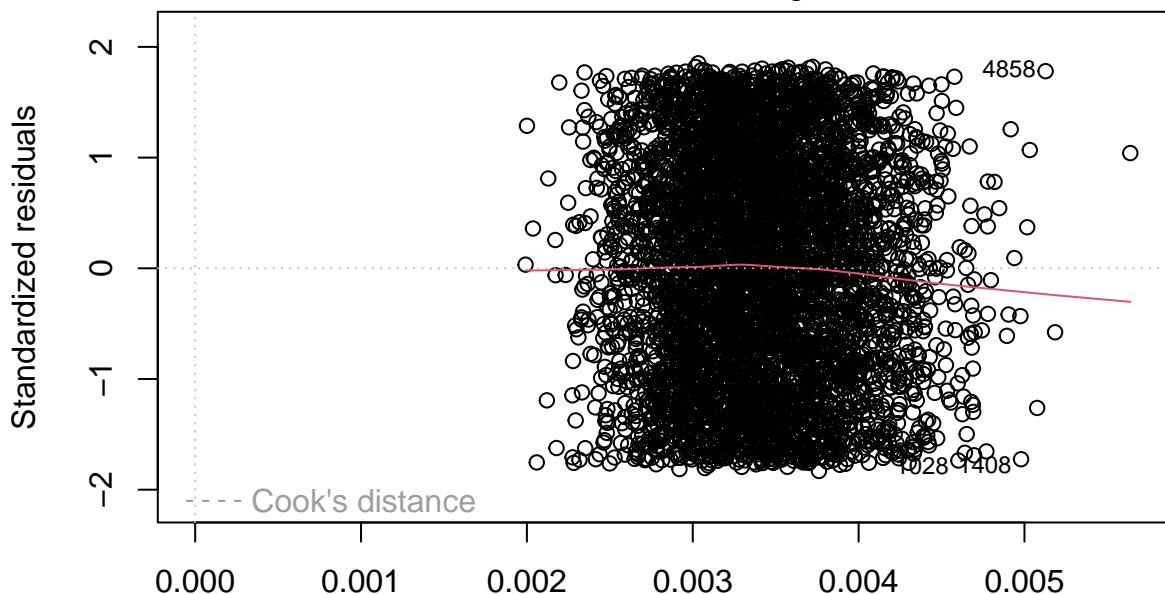
Using a significance level of 0.05 there are some significant interactions Using an anova test the p-value is high....

```
plot(additive_model)
```





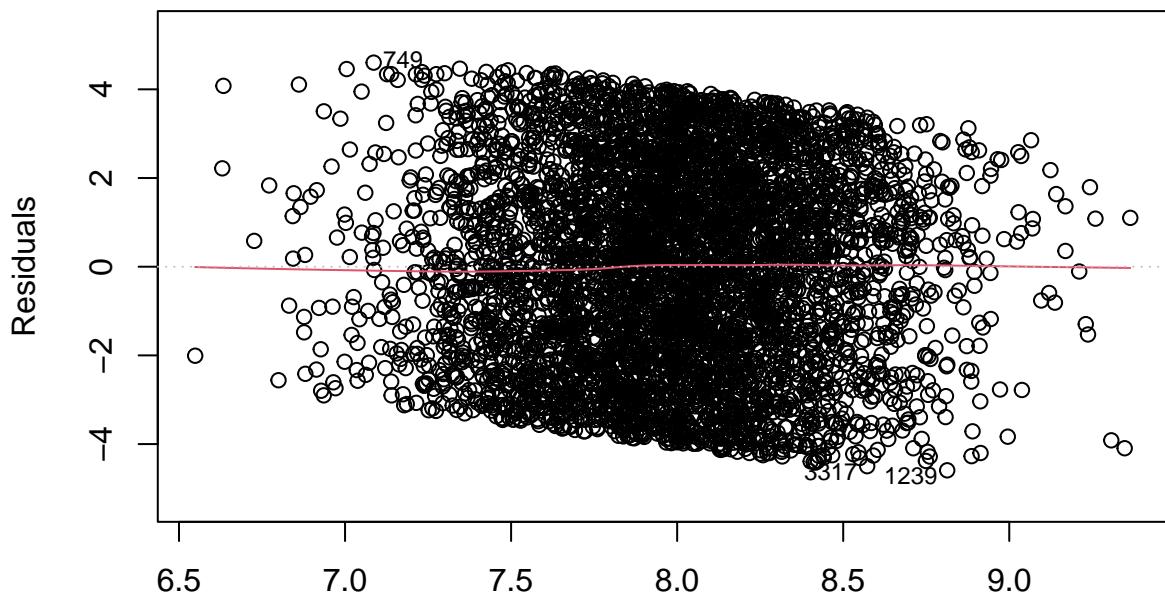
Residuals vs Leverage



Leverage  
lm(Work.Hours..hrs.day. ~ .)

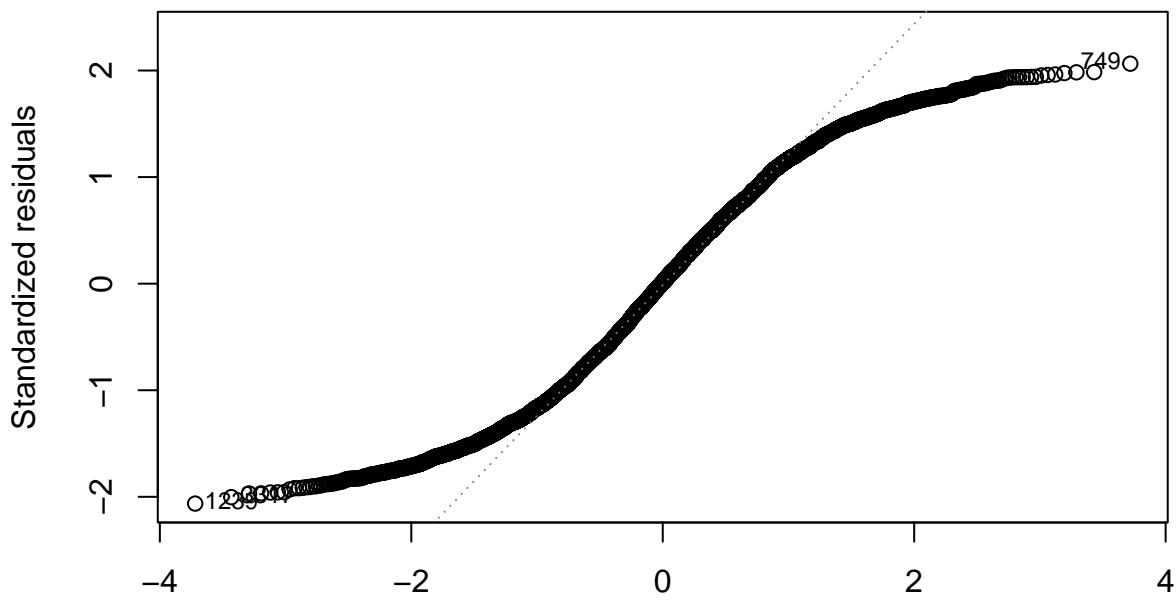
```
plot(full_model)
```

Residuals vs Fitted

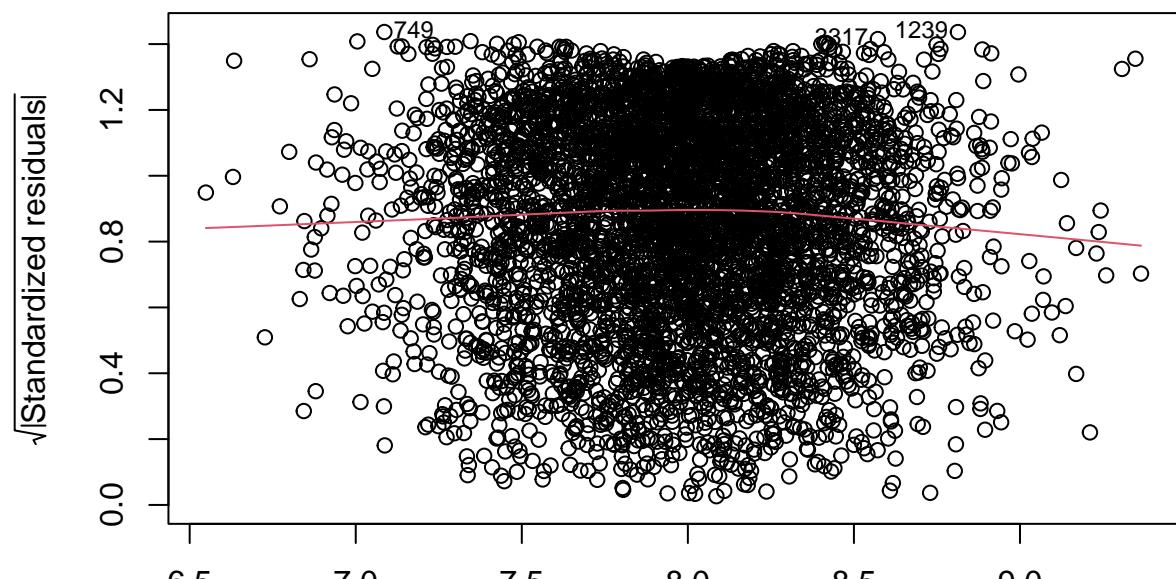


Fitted values  
lm(Work.Hours..hrs.day. ~ (.)^2)

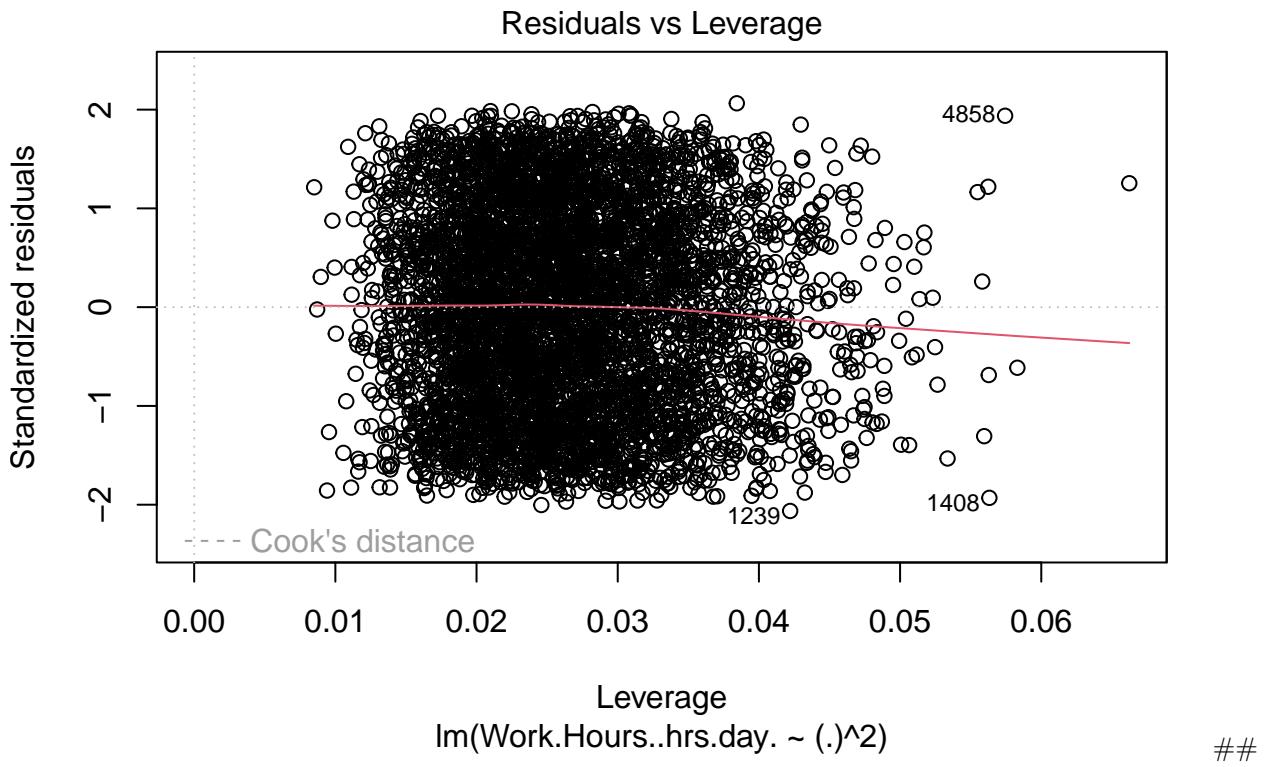
Q–Q Residuals



Theoretical Quantiles  
Im(Work.Hours..hrs.day. ~ (.)^2)  
Scale–Location



Fitted values  
Im(Work.Hours..hrs.day. ~ (.)^2)



## Last attempt at model selection to test for significant predictors

```
## Need to pick BIC or AIC (For BIC: k = log(nrow(df_bc))) before trace
# Using AIC, since the slides state it to be more accurate than BIC (also, overfitting is not a concern)
stepwise_model <- stepAIC(full_model, direction = "both", trace = FALSE)

# Display the summary of the selected model
summary(stepwise_model)

##
## Call:
## lm(formula = Work.Hours..hrs.day. ~ Age + Sleep.End.Time + Total.Sleep.Hours +
##     Exercise..mins.day. + Caffeine.Intake..mg. + Sleep.QualityCat +
##     Productivity.ScoreCat + Mood.ScoreCat + Age:Exercise..mins.day. +
##     Age:Sleep.QualityCat + Age:Mood.ScoreCat + Sleep.End.Time:Caffeine.Intake..mg. +
##     Sleep.End.Time:Mood.ScoreCat + Total.Sleep.Hours:Caffeine.Intake..mg. +
##     Total.Sleep.Hours:Mood.ScoreCat + Exercise..mins.day.:Sleep.QualityCat +
##     Productivity.ScoreCat:Mood.ScoreCat, data = df_filtered)
##
## Residuals:
##      Min      1Q Median      3Q      Max
## -4.5814 -1.9326  0.0211  1.8896  4.4367
##
## Coefficients:
## (Intercept)          Estimate Std. Error t value
## Age                  7.4434120  0.5384811 13.823
## Age                  0.0259032  0.0076666  3.379
```

```

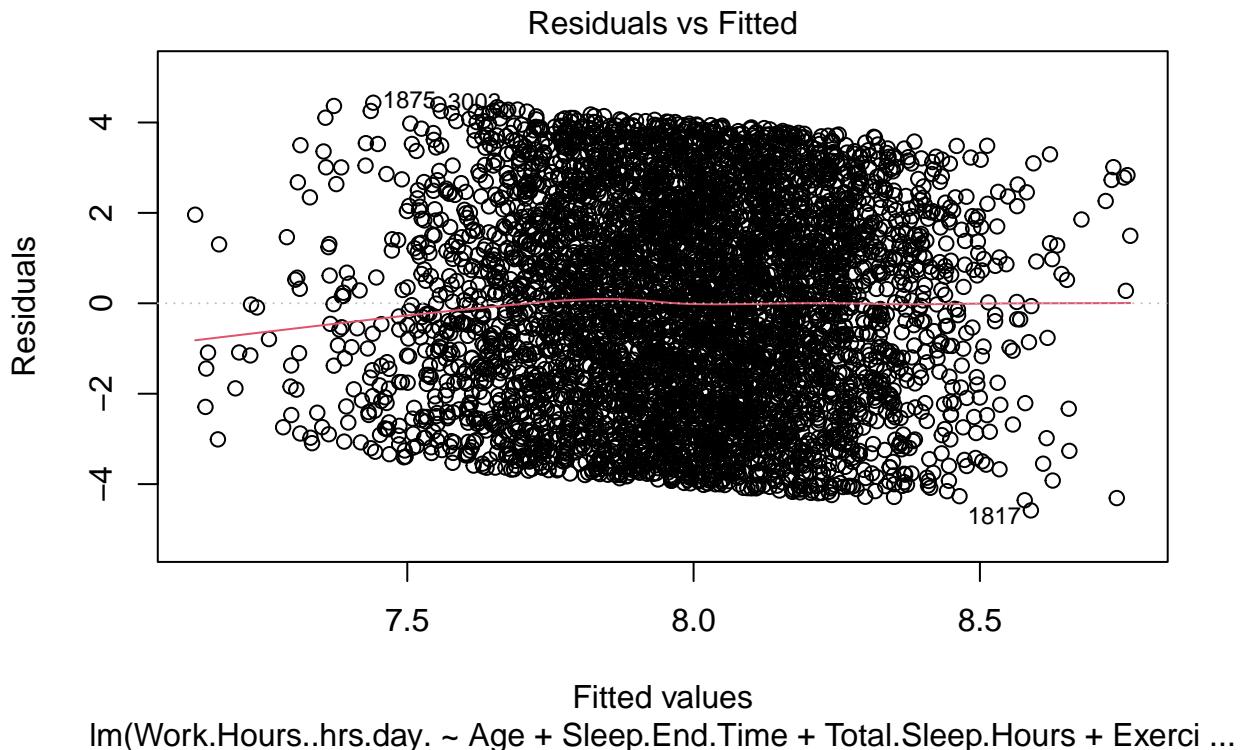
## Sleep.End.Time          0.0315214  0.0691137  0.456
## Total.Sleep.Hours      -0.1051131  0.0887827 -1.184
## Exercise..mins.day.     0.0136840  0.0045652  2.997
## Caffeine.Intake..mg.    0.0021680  0.0019567  1.108
## Sleep.QualityCatMedium 0.6937333  0.2877354  2.411
## Sleep.QualityCatHigh   0.2875609  0.3077839  0.934
## Productivity.ScoreCatMedium -0.0094452  0.1430943 -0.066
## Productivity.ScoreCatHigh 0.1469362  0.1487323  0.988
## Mood.ScoreCatMedium     -1.0898038  0.4813673 -2.264
## Mood.ScoreCatHigh       0.0629670  0.5320839  0.118
## Age:Exercise..mins.day. -0.0003144  0.0001020 -3.081
## Age:Sleep.QualityCatMedium -0.0136739  0.0063261 -2.161
## Age:Sleep.QualityCatHigh -0.0111778  0.0067511 -1.656
## Age:Mood.ScoreCatMedium  0.0039431  0.0062124  0.635
## Age:Mood.ScoreCatHigh   -0.0117715  0.0067821 -1.736
## Sleep.End.Time:Caffeine.Intake..mg. 0.0004803  0.0003269  1.469
## Sleep.End.Time:Mood.ScoreCatMedium -0.1572842  0.0659372 -2.385
## Sleep.End.Time:Mood.ScoreCatHigh   -0.0519702  0.0722716 -0.719
## Total.Sleep.Hours:Caffeine.Intake..mg. -0.0006490  0.0004155 -1.562
## Total.Sleep.Hours:Mood.ScoreCatMedium 0.2565001  0.0845115  3.035
## Total.Sleep.Hours:Mood.ScoreCatHigh   0.1201555  0.0938618  1.280
## Exercise..mins.day.:Sleep.QualityCatMedium -0.0047212  0.0029979 -1.575
## Exercise..mins.day.:Sleep.QualityCatHigh 0.0012498  0.0032437  0.385
## Productivity.ScoreCatMedium:Mood.ScoreCatMedium 0.1990204  0.1881790  1.058
## Productivity.ScoreCatHigh:Mood.ScoreCatMedium -0.3232027  0.1967645 -1.643
## Productivity.ScoreCatMedium:Mood.ScoreCatHigh -0.0484044  0.2081450 -0.233
## Productivity.ScoreCatHigh:Mood.ScoreCatHigh -0.1849050  0.2163547 -0.855
## Pr(>|t|)
< 2e-16 ***
## (Intercept) 0.000734 ***
## Age          0.648352
## Sleep.End.Time 0.236495
## Total.Sleep.Hours 0.002736 **
## Exercise..mins.day. 0.267916
## Caffeine.Intake..mg. 0.015944 *
## Sleep.QualityCatMedium 0.350197
## Sleep.QualityCatHigh 0.947375
## Productivity.ScoreCatMedium 0.323238
## Productivity.ScoreCatHigh 0.023619 *
## Mood.ScoreCatMedium 0.905803
## Mood.ScoreCatHigh 0.002073 **
## Age:Exercise..mins.day. 0.030705 *
## Age:Sleep.QualityCatMedium 0.097842 .
## Age:Sleep.QualityCatHigh 0.525649
## Age:Mood.ScoreCatMedium 0.082684 .
## Age:Mood.ScoreCatHigh 0.141817
## Sleep.End.Time:Caffeine.Intake..mg. 0.017099 *
## Sleep.End.Time:Mood.ScoreCatMedium 0.472116
## Sleep.End.Time:Mood.ScoreCatHigh 0.118416
## Total.Sleep.Hours:Caffeine.Intake..mg. 0.002417 **
## Total.Sleep.Hours:Mood.ScoreCatMedium 0.200558
## Exercise..mins.day.:Sleep.QualityCatMedium 0.115354
## Exercise..mins.day.:Sleep.QualityCatHigh 0.700039
## Productivity.ScoreCatMedium:Mood.ScoreCatMedium 0.290284

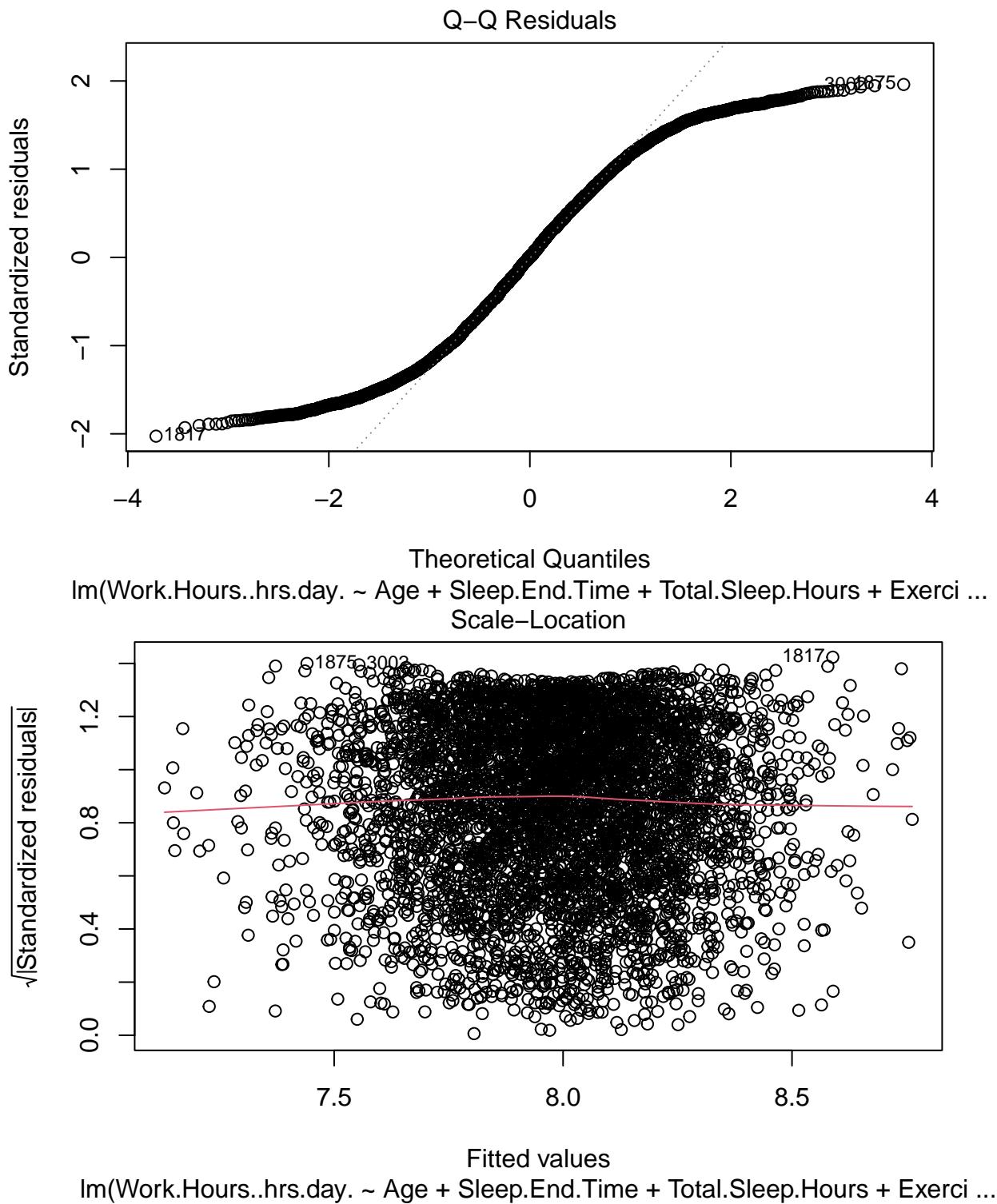
```

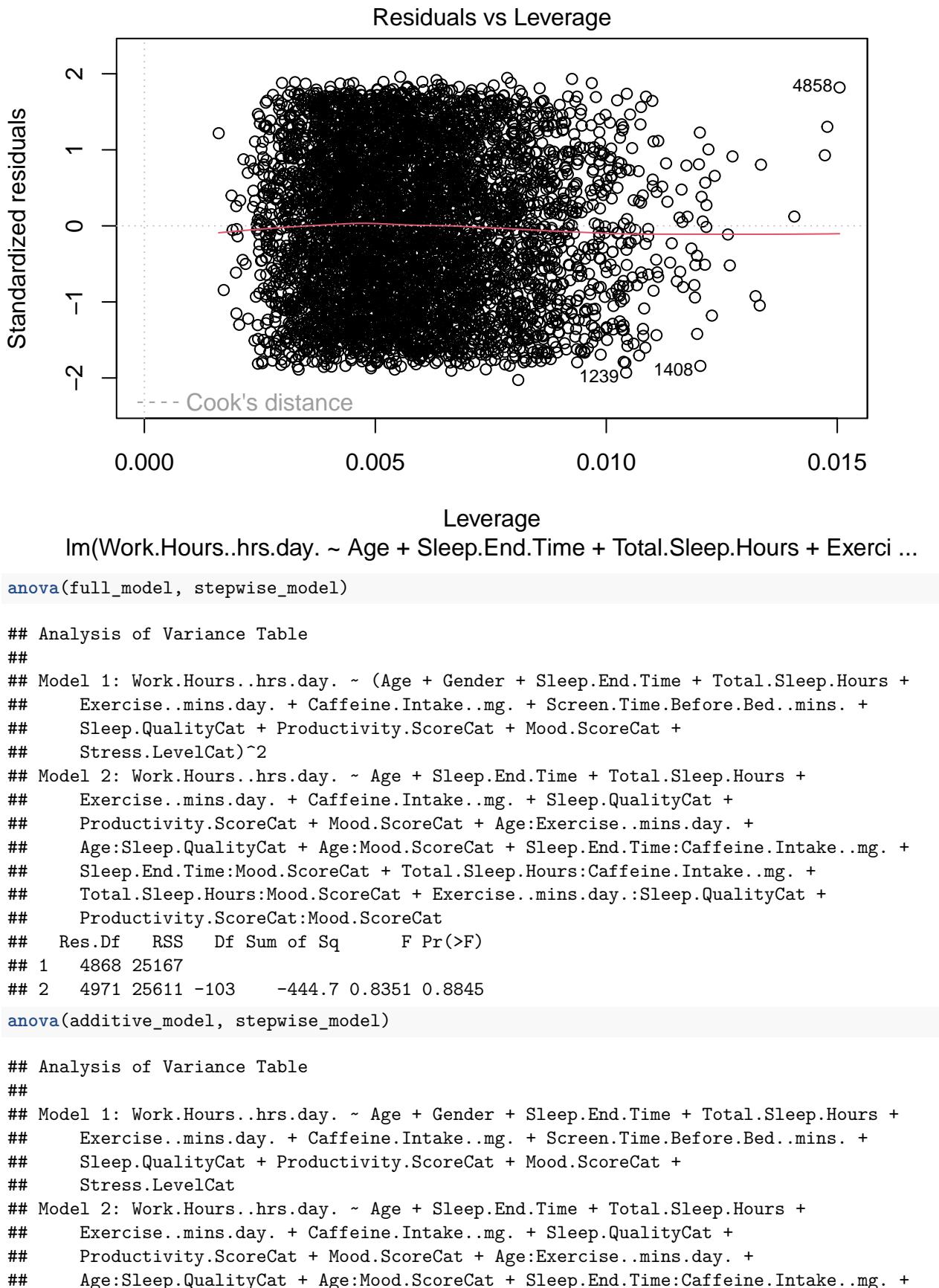
```

## Productivity.ScoreCatHigh:Mood.ScoreCatMedium  0.100532
## Productivity.ScoreCatMedium:Mood.ScoreCatHigh  0.816119
## Productivity.ScoreCatHigh:Mood.ScoreCatHigh    0.392793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.27 on 4971 degrees of freedom
## Multiple R-squared:  0.01075,   Adjusted R-squared:  0.005177
## F-statistic: 1.929 on 28 and 4971 DF,  p-value: 0.002323
plot(stepwise_model)

```







```

##      Sleep.End.Time:Mood.ScoreCat + Total.Sleep.Hours:Caffeine.Intake..mg. +
##      Total.Sleep.Hours:Mood.ScoreCat + Exercise..mins.day.:Sleep.QualityCat +
##      Productivity.ScoreCat:Mood.ScoreCat
##      Res.Df   RSS Df Sum of Sq      F    Pr(>F)
## 1    4983 25831
## 2    4971 25611 12     219.86 3.5561 2.707e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

LOOCV not useful (5,000 folds), k-fold cross validation - not useful since the response data is uniform and each fold will likely look the same

## Final Thoughts

No predictors are significant when looking at the anova tests, some predictors in the full model are significant

The predictors truly have no effect on work hours A different model might be better. More data may be needed for stronger statistical power.