Sleep Efficiency Regression Analysis

Graeme Ko

2023-12-05

```
library(ggplot2)
library(leaps)
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
library(olsrr)
```

```
##
## Attaching package: 'olsrr'
```

```
## The following object is masked from 'package:datasets':
##
##     rivers
```

```
library(lmtest)
```

```
## Warning: package 'lmtest' was built under R version 4.3.2
```

```
## Loading required package: zoo
```

```
## Warning: package 'zoo' was built under R version 4.3.2
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(car)
```

```
## Loading required package: carData
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following object is masked from 'package:car':
##
##     recode

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(leaps)
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select

## The following object is masked from 'package:olsrr':
##
##     cement
```

```r
df = read.csv('Sleep_Efficiency.csv')
head(df)
```

```
##   ID Age Gender          Bedtime     Wakeup.time Sleep.duration
## 1  1  65 Female  2021-03-06 1:00 2021-03-06 7:00            6.0
## 2  2  69   Male  2021-12-05 2:00 2021-12-05 9:00            7.0
## 3  3  40 Female 2021-05-25 21:30 2021-05-25 5:30            8.0
## 4  4  40 Female  2021-11-03 2:30 2021-11-03 8:30            6.0
## 5  5  57   Male  2021-03-13 1:00 2021-03-13 9:00            8.0
## 6  6  36 Female 2021-07-01 21:00 2021-07-01 4:30            7.5
##   Sleep.efficiency REM.sleep.percentage Deep.sleep.percentage
## 1             0.88                   18                    70
## 2             0.66                   19                    28
## 3             0.89                   20                    70
## 4             0.51                   23                    25
## 5             0.76                   27                    55
## 6             0.90                   23                    60
##   Light.sleep.percentage Awakenings Caffeine.consumption Alcohol.consumption
## 1                     12          0                    0                    0
```

```
## 2                    53          3                    0                    3
## 3                    10          1                    0                    0
## 4                    52          3                   50                    5
## 5                    18          3                    0                    3
## 6                    17          0                   NA                    0
##   Smoking.status Exercise.frequency
## 1            Yes                  3
## 2            Yes                  3
## 3             No                  3
## 4            Yes                  1
## 5             No                  3
## 6             No                  1
```

Reformatting Bedtime and Wakeup.time, removing duplicates, removing NAs

```
df = na.omit(df)

# check for duplicates
df[duplicated(df)]
```

```
## data frame with 0 columns and 388 rows
```

```
# change format of columns
df$Bedtime <- as.POSIXct(df$Bedtime, format = "%Y-%m-%d %H:%M")
df$Wakeup.time <- as.POSIXct(df$Wakeup.time, format = "%Y-%m-%d %H:%M")

# Extract only the time part
df$Bedtime <- format(df$Bedtime, format = "%H:%M")
df$Wakeup.time <- format(df$Wakeup.time, format = "%H:%M")
df$Sleep.duration <- as.integer(df$Sleep.duration)

convert_time_to_numeric <- function(time_str) {
  if (is.na(time_str)) {
    return(NA)  # Return NA for missing values
  }

  time <- as.POSIXct(time_str, format = "%H:%M")
  hours <- as.numeric(format(time, "%H"))
  minutes <- as.numeric(format(time, "%M"))

  total_hours <- hours + minutes / 60  # Convert minutes to hours

  if (format(time, "%p") == "PM") {
    total_hours <- 12 - total_hours    # Subtract 12 hours for PM times
  }

  return(total_hours)
}

# Apply the function to Bedtime and Wakeup.time
df$Bedtime <- sapply(df$Bedtime, convert_time_to_numeric)
```

```r
df$Wakeup.time <- sapply(df$Wakeup.time, convert_time_to_numeric)
str(df)
```

```
## 'data.frame':    388 obs. of  15 variables:
##  $ ID                    : int  1 2 3 4 5 7 8 9 10 11 ...
##  $ Age                   : int  65 69 40 40 57 27 53 41 11 50 ...
##  $ Gender                : chr  "Female" "Male" "Female" "Female" ...
##  $ Bedtime               : num  1 2 -9.5 2.5 1 -9 0.5 2.5 1 0.5 ...
##  $ Wakeup.time           : num  7 9 5.5 8.5 9 3 10.5 8.5 10 8.5 ...
##  $ Sleep.duration        : int  6 7 8 6 8 6 10 6 9 8 ...
##  $ Sleep.efficiency      : num  0.88 0.66 0.89 0.51 0.76 0.54 0.9 0.79 0.55 0.92 ...
##  $ REM.sleep.percentage  : int  18 19 20 23 27 28 28 28 18 23 ...
##  $ Deep.sleep.percentage : int  70 28 70 25 55 25 52 55 37 57 ...
##  $ Light.sleep.percentage: int  12 53 10 52 18 47 20 17 45 20 ...
##  $ Awakenings            : int  0 3 1 3 3 2 0 3 4 1 ...
##  $ Caffeine.consumption  : int  0 0 0 50 0 50 50 50 0 50 ...
##  $ Alcohol.consumption   : int  0 3 0 5 3 0 0 0 0 0 ...
##  $ Smoking.status        : chr  "Yes" "Yes" "No" "Yes" ...
##  $ Exercise.frequency    : int  3 3 3 1 3 1 3 1 0 3 ...
##  - attr(*, "na.action")= 'omit' Named int [1:64] 6 20 21 25 27 34 38 58 61 63 ...
##   ..- attr(*, "names")= chr [1:64] "6" "20" "21" "25" ...
```

```r
head(df,100)
```

```
##      ID Age Gender Bedtime Wakeup.time Sleep.duration Sleep.efficiency
## 1     1  65 Female     1.0         7.0              6             0.88
## 2     2  69   Male     2.0         9.0              7             0.66
## 3     3  40 Female    -9.5         5.5              8             0.89
## 4     4  40 Female     2.5         8.5              6             0.51
## 5     5  57   Male     1.0         9.0              8             0.76
## 7     7  27 Female    -9.0         3.0              6             0.54
## 8     8  53   Male     0.5        10.5             10             0.90
## 9     9  41 Female     2.5         8.5              6             0.79
## 10   10  11 Female     1.0        10.0              9             0.55
## 11   11  50   Male     0.5         8.5              8             0.92
## 12   12  55   Male   -10.5         6.0              7             0.93
## 13   13  30 Female     2.5        11.5              9             0.93
## 14   14  28   Male     1.0         9.5              8             0.64
## 15   15  36 Female     1.5        10.0              8             0.54
## 16   16  32 Female   -10.0         5.5              7             0.92
## 17   17  21 Female     1.0         8.0              7             0.54
## 18   18  40 Female   -11.0         6.5              7             0.50
## 19   19  43 Female     0.0         9.0              9             0.98
## 22   22  32   Male     2.5        10.0              7             0.71
## 23   23  29 Female   -10.0         6.0              8             0.84
## 24   24  63 Female     2.5        10.5              8             0.98
## 26   26  52   Male     0.0         6.0              6             0.91
## 28   28  35   Male     2.5         9.5              7             0.84
## 29   29  23   Male    -9.5         6.5              9             0.65
## 30   30  47 Female   -10.5         5.5              7             0.91
## 31   31  24   Male     0.0         7.0              7             0.57
## 32   32  18   Male     0.5         8.5              8             0.68
## 33   33  26   Male     0.0         7.0              7             0.55
```

```
## 35   35  46 Female    -9.5         5.0              7            0.94
## 36   36  61 Female     1.0         8.0              7            0.64
## 37   37  38   Male   -10.0         5.0              7            0.87
## 39   39  28 Female     0.0         7.5              7            0.87
## 40   40  58 Female     2.0         9.0              7            0.63
## 41   41  46   Male   -10.0         4.0              6            0.83
## 42   42  52   Male     1.5         7.5              6            0.83
## 43   43  29 Female     0.0        10.0             10            0.59
## 44   44  31 Female   -10.5         6.0              7            0.87
## 45   45  34 Female   -11.0         7.0              8            0.77
## 46   46  40 Female     1.5         8.5              7            0.86
## 47   47  55   Male     0.0         8.0              8            0.91
## 48   48  27 Female     0.0         7.5              7            0.71
## 49   49  21 Female     0.0         8.0              8            0.81
## 50   50  37 Female    -9.5         4.5              7            0.81
## 51   51  29 Female   -10.0         5.0              7            0.71
## 52   52  65   Male     2.5        10.0              7            0.71
## 53   53   9 Female     1.0         9.5              8            0.52
## 54   54  16 Female     1.0         9.5              8            0.65
## 55   55  18   Male     1.5         9.0              7            0.84
## 56   56  37   Male   -10.0         7.0              9            0.84
## 57   57  54   Male     1.0         9.0              8            0.99
## 59   59  34 Female     1.5         8.5              7            0.71
## 60   60  34 Female    -9.0         4.5              7            0.80
## 62   62  56   Male   -10.5         5.5              7            0.91
## 66   66  21   Male     1.0         9.0              8            0.77
## 67   67  25   Male     0.0         7.0              7            0.84
## 68   68  52   Male     1.5         9.0              7            0.87
## 69   69  46   Male     0.0         7.0              7            0.80
## 70   70  40   Male   -10.0         4.0              6            0.52
## 71   71  30 Female     2.5        10.0              7            0.54
## 72   72  32 Female   -11.0         6.0              7            0.88
## 73   73  55   Male    -9.5         3.5              6            0.64
## 74   74  58   Male     2.5        10.0              7            0.87
## 75   75  47   Male   -10.0         5.0              7            0.77
## 77   77  47 Female    -9.0         4.0              7            0.94
## 78   78  46   Male     2.0        11.0              9            0.70
## 79   79  43   Male   -10.0         5.0              7            0.79
## 80   80  54   Male     2.5        10.0              7            0.77
## 81   81  55   Male     1.5        10.0              8            0.86
## 82   82  29   Male    -9.0         4.5              7            0.81
## 83   83  66   Male     2.0         9.5              7            0.88
## 84   84  40 Female   -10.5         6.5              8            0.78
## 85   85  50   Male   -10.0         4.0              6            0.64
## 87   87  50   Male     0.0         7.0              7            0.78
## 88   88  50   Male     0.0         7.0              7            0.71
## 89   89  48   Male     2.0         9.5              7            0.97
## 90   90  24 Female     0.0         7.5              7            0.77
## 91   91  45   Male     2.5         7.5              5            0.80
## 92   92  57   Male    -9.5         5.5              8            0.94
## 93   93  39   Male    -9.0         4.0              7            0.52
## 94   94  43   Male   -10.0         5.0              7            0.76
## 95   95  51   Male     0.0         7.5              7            0.73
## 96   96  61 Female     2.0         9.0              7            0.73
```

```
## 97   97  41   Male    -9.0         4.0              7             0.96
## 98   98  53   Male   -11.0         6.5              7             0.91
## 99   99  32 Female     2.0        11.0              9             0.80
## 100 100  65   Male     1.5         9.5              8             0.77
## 101 101  38 Female     0.5         8.5              8             0.81
## 102 102  53   Male     1.0         9.5              8             0.71
## 103 103  55 Female     0.0         8.0              8             0.95
## 104 104  44 Female   -11.0         5.0              6             0.64
## 105 105  38 Female   -11.0         5.0              6             0.94
## 106 106  40 Female   -10.0         6.0              8             0.87
## 107 107  53   Male    -9.0         4.0              7             0.63
## 108 108  24   Male     0.5         9.0              8             0.88
## 109 109  39 Female   -11.0         7.0              8             0.90
## 110 110  47   Male     0.5         8.0              7             0.94
## 111 111  51   Male     0.5         5.5              5             0.91
## 112 112  25   Male   -11.0         6.0              7             0.90
## 113 113  23   Male   -11.0         7.0              8             0.80
## 114 114  54   Male     1.5         7.5              6             0.85
##     REM.sleep.percentage Deep.sleep.percentage Light.sleep.percentage
## 1                     18                    70                     12
## 2                     19                    28                     53
## 3                     20                    70                     10
## 4                     23                    25                     52
## 5                     27                    55                     18
## 7                     28                    25                     47
## 8                     28                    52                     20
## 9                     28                    55                     17
## 10                    18                    37                     45
## 11                    23                    57                     20
## 12                    18                    60                     22
## 13                    24                    58                     18
## 14                    28                    25                     47
## 15                    20                    32                     48
## 16                    25                    55                     20
## 17                    28                    22                     50
## 18                    18                    20                     62
## 19                    20                    67                     13
## 22                    23                    58                     19
## 23                    23                    60                     17
## 24                    22                    65                     13
## 26                    18                    72                     10
## 28                    24                    60                     16
## 29                    27                    20                     53
## 30                    20                    60                     20
## 31                    27                    20                     53
## 32                    22                    30                     48
## 33                    23                    23                     54
## 35                    25                    60                     15
## 36                    23                    22                     55
## 37                    18                    72                     10
## 39                    28                    60                     12
## 40                    22                    23                     55
## 41                    22                    58                     20
## 42                    20                    67                     13
```

```
## 43  28  20  52
## 44  23  60  17
## 45  22  65  13
## 46  19  63  18
## 47  23  57  20
## 48  28  60  12
## 49  23  60  17
## 50  25  55  20
## 51  22  57  21
## 52  22  59  19
## 53  18  35  47
## 54  18  35  47
## 55  18  70  12
## 56  18  70  12
## 57  27  55  18
## 59  22  65  13
## 60  28  60  12
## 62  18  60  22
## 66  19  63  18
## 67  15  70  15
## 68  18  72  10
## 69  18  70  12
## 70  20  35  45
## 71  28  20  52
## 72  18  62  20
## 73  20  35  45
## 74  23  60  17
## 75  22  58  20
## 77  30  60  10
## 78  18  70  12
## 79  20  70  10
## 80  23  60  17
## 81  20  67  13
## 82  15  67  18
## 83  24  60  16
## 84  22  57  21
## 85  22  22  56
## 87  22  57  21
## 88  26  58  16
## 89  23  60  17
## 90  26  56  18
## 91  24  60  16
## 92  27  55  18
## 93  18  35  47
## 94  22  58  20
## 95  22  58  20
## 96  20  67  13
## 97  28  55  17
## 98  22  57  21
## 99  20  67  13
## 100  20  70  10
## 101  28  55  17
## 102  20  67  13
## 103  23  60  17
```

```
## 104                  20                  32                  48
## 105                  28                  60                  12
## 106                  22                  57                  21
## 107                  25                  20                  55
## 108                  19                  63                  18
## 109                  22                  63                  15
## 110                  23                  57                  20
## 111                  23                  57                  20
## 112                  20                  65                  15
## 113                  15                  65                  20
## 114                  20                  67                  13
##     Awakenings Caffeine.consumption Alcohol.consumption Smoking.status
## 1            0                   0                   0            Yes
## 2            3                   0                   3            Yes
## 3            1                   0                   0             No
## 4            3                  50                   5            Yes
## 5            3                   0                   3             No
## 7            2                  50                   0            Yes
## 8            0                  50                   0            Yes
## 9            3                  50                   0             No
## 10           4                   0                   0             No
## 11           1                  50                   0            Yes
## 12           0                   0                   0             No
## 13           0                  50                   0             No
## 14           4                   0                   0             No
## 15           2                  25                   1            Yes
## 16           0                  50                   2             No
## 17           4                   0                   2            Yes
## 18           3                  50                   2            Yes
## 19           0                  25                   1             No
## 22           3                  50                   4             No
## 23           2                  75                   0             No
## 24           0                  50                   0             No
## 26           1                   0                   0             No
## 28           2                   0                   0            Yes
## 29           3                   0                   0             No
## 30           1                   0                   3             No
## 31           3                   0                   1            Yes
## 32           3                  25                   0             No
## 33           3                   0                   4            Yes
## 35           0                   0                   0             No
## 36           2                  50                   3             No
## 37           1                   0                   0            Yes
## 39           1                  50                   0            Yes
## 40           3                  50                   3             No
## 41           3                   0                   0             No
## 42           1                  25                   0            Yes
## 43           2                  50                   0            Yes
## 44           1                  50                   0             No
## 45           3                  25                   1             No
## 46           1                  50                   0             No
## 47           1                   0                   0             No
## 48           1                  50                   1            Yes
## 49           3                   0                   0             No
```

```
## 50          1           0           5          No
## 51          1          25           0          No
## 52          2          50           2          No
## 53          2           0           0          No
## 54          2           0           0          No
## 55          2           0           0          No
## 56          1           0           0          No
## 57          1           0           3          No
## 59          3          50           0          No
## 60          4          50           0          No
## 62          0           0           0          No
## 66          4           0           3          No
## 67          4           0           4          No
## 68          0          25           2          No
## 69          2           0           0          No
## 70          2           0           2         Yes
## 71          1          50           3         Yes
## 72          1          50           2          No
## 73          1           0           4         Yes
## 74          0          25           0         Yes
## 75          4           0           0          No
## 77          1           0           0          No
## 78          1           0           5         Yes
## 79          4           0           0          No
## 80          3           0           0         Yes
## 81          1          25           0         Yes
## 82          2         200           0          No
## 83          0           0           0         Yes
## 84          4          25           0          No
## 85          4           0           3          No
## 87          2           0           0          No
## 88          3           0           3          No
## 89          0           0           0         Yes
## 90          4          25           0          No
## 91          4           0           0         Yes
## 92          0           0           3          No
## 93          3           0           3         Yes
## 94          4           0           0          No
## 95          4           0           3          No
## 96          1          25           2         Yes
## 97          0         200           0          No
## 98          1          25           0          No
## 99          2          25           0          No
## 100         4           0           0          No
## 101         1          25           1          No
## 102         4           0           1          No
## 103         1          25           0          No
## 104         4          25           5          No
## 105         0          50           0          No
## 106         0          25           0          No
## 107         3           0           5          No
## 108         0          50           0         Yes
## 109         1          50           2          No
## 110         1          50           0         Yes
```

```
## 111          1                  50                  0           Yes
## 112          0                 100                  0            No
## 113          4                  75                  0            No
## 114          1                  25                  0           Yes
##     Exercise.frequency
## 1                    3
## 2                    3
## 3                    3
## 4                    1
## 5                    3
## 7                    1
## 8                    3
## 9                    1
## 10                   0
## 11                   3
## 12                   3
## 13                   1
## 14                   3
## 15                   0
## 16                   5
## 17                   0
## 18                   3
## 19                   0
## 22                   1
## 23                   2
## 24                   1
## 26                   2
## 28                   2
## 29                   3
## 30                   4
## 31                   3
## 32                   0
## 33                   0
## 35                   3
## 36                   0
## 37                   3
## 39                   1
## 40                   0
## 41                   0
## 42                   1
## 43                   1
## 44                   1
## 45                   0
## 46                   1
## 47                   0
## 48                   1
## 49                   0
## 50                   4
## 51                   2
## 52                   3
## 53                   0
## 54                   0
## 55                   3
## 56                   3
```

```
## 57                   3
## 59                   1
## 60                   1
## 62                   4
## 66                   3
## 67                   2
## 68                   1
## 69                   0
## 70                   3
## 71                   1
## 72                   4
## 73                   3
## 74                   2
## 75                   0
## 77                   3
## 78                   3
## 79                   3
## 80                   2
## 81                   1
## 82                   2
## 83                   3
## 84                   3
## 85                   0
## 87                   0
## 88                   0
## 89                   3
## 90                   1
## 91                   2
## 92                   3
## 93                   3
## 94                   0
## 95                   0
## 96                   0
## 97                   3
## 98                   2
## 99                   0
## 100                  3
## 101                  0
## 102                  3
## 103                  0
## 104                  0
## 105                  1
## 106                  3
## 107                  2
## 108                  1
## 109                  4
## 110                  3
## 111                  3
## 112                  3
## 113                  2
## 114                  1
```

Initial model

```
mod1 = lm(Sleep.efficiency~Age+Gender+Bedtime+Wakeup.time+Sleep.duration+REM.sleep.percentage+Deep.sleep
summary(mod1)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + Gender + Bedtime + Wakeup.time +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     Smoking.status + Exercise.frequency, data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.172484 -0.042354  0.005555  0.039865  0.146482
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.3436989  0.0434220   7.915 2.81e-14 ***
## Age                     0.0009757  0.0002459   3.968 8.68e-05 ***
## GenderMale              0.0017585  0.0069883   0.252  0.80146
## Bedtime                -0.0004669  0.0009456  -0.494  0.62177
## Wakeup.time            -0.0000356  0.0027415  -0.013  0.98965
## Sleep.duration          0.0021003  0.0041473   0.506  0.61285
## REM.sleep.percentage    0.0066962  0.0009558   7.006 1.14e-11 ***
## Deep.sleep.percentage   0.0055714  0.0002387  23.343  < 2e-16 ***
## Awakenings             -0.0317903  0.0025326 -12.553  < 2e-16 ***
## Caffeine.consumption    0.0002451  0.0001142   2.147  0.03246 *
## Alcohol.consumption    -0.0061242  0.0021201  -2.889  0.00409 **
## Smoking.statusYes      -0.0448981  0.0069467  -6.463 3.19e-10 ***
## Exercise.frequency      0.0056564  0.0024931   2.269  0.02384 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06082 on 375 degrees of freedom
## Multiple R-squared:  0.8054, Adjusted R-squared:  0.7991
## F-statistic: 129.3 on 12 and 375 DF,  p-value: < 2.2e-16
```

Checking multicollinearity

```
vif(mod1)
```

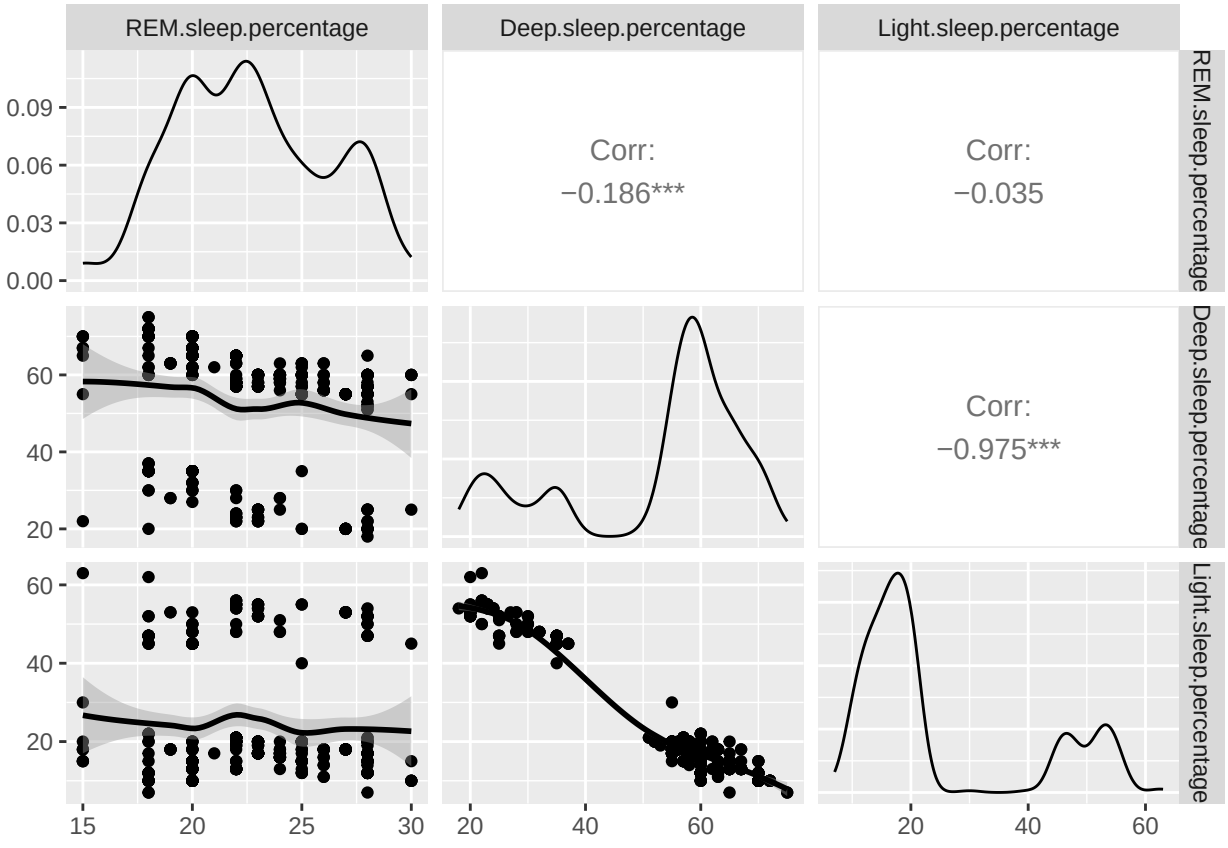
```
##                    Age                 Gender                Bedtime
##               1.136319               1.280531               2.923099
##            Wakeup.time         Sleep.duration   REM.sleep.percentage
##               3.245744               1.379663               1.124628
##  Deep.sleep.percentage             Awakenings   Caffeine.consumption
##               1.444977               1.233673               1.146675
##    Alcohol.consumption         Smoking.status     Exercise.frequency
##               1.223044               1.140210               1.362979
```

Removing Wakeup.time to due to multicollinearity - the data is accounted for by Bedtime and Sleep.duration

```
sleeppercdata <- subset(df, select=c("REM.sleep.percentage","Deep.sleep.percentage","Light.sleep.percen

sleeptimedata <- subset(df, select=c("Bedtime","Wakeup.time"))
```

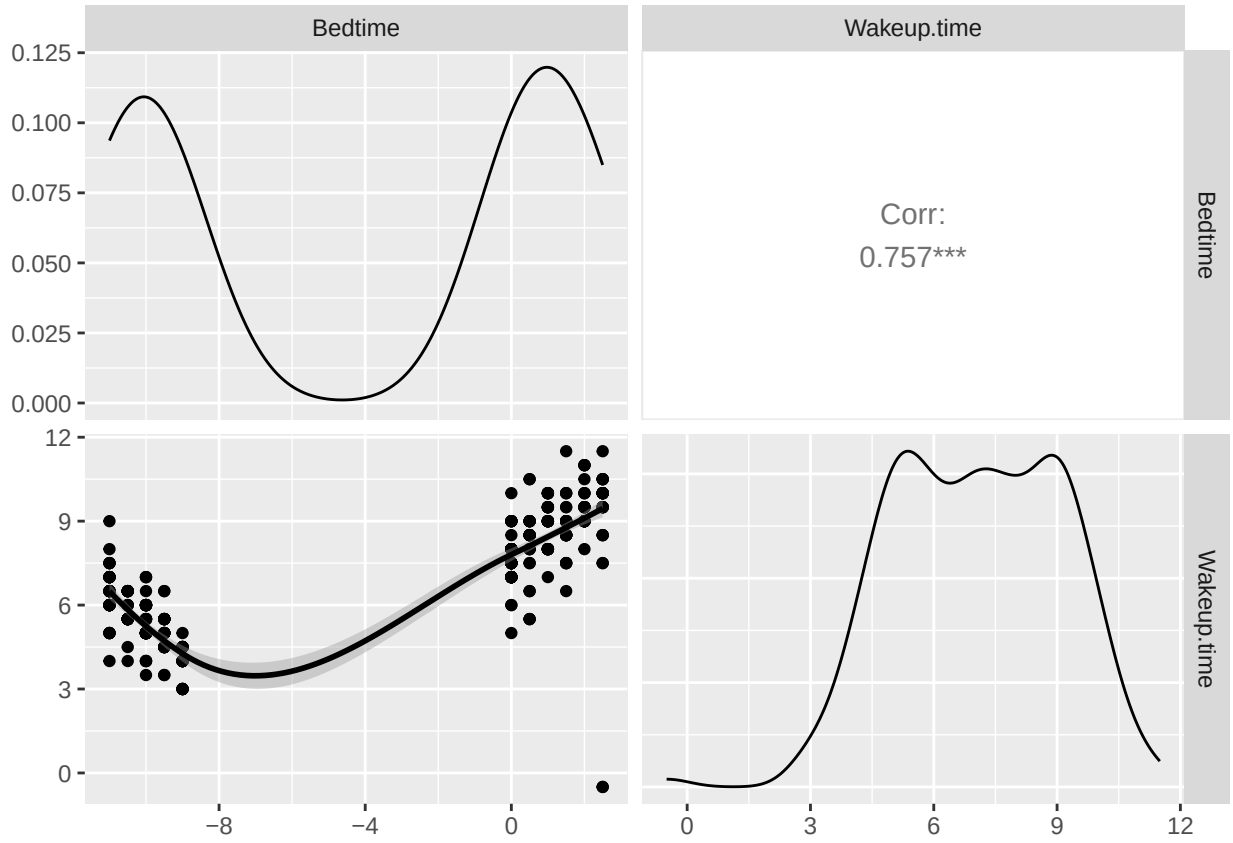
Checking correlation between columns which produced errors due to aliasing in the linear model

```
ggpairs(sleeppercdata,lower = list(continuous = "smooth_loess", combo =
"facethist", discrete = "facetbar", na = "na"))
```



Checking correlation between variables with multicollinearity

```
ggpairs(sleeptimedata,lower = list(continuous = "smooth_loess", combo =
"facethist", discrete = "facetbar", na = "na"))
```



Re-estimating model adjusted for multicollinearity

```
basemod = lm(Sleep.efficiency~Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage+ Deep.sleep

summary(basemod)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency, data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.172453 -0.042393  0.005563  0.039850  0.146470
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.3435592  0.0420132   8.177 4.51e-15 ***
## Age                         0.0009755  0.0002450   3.982 8.22e-05 ***
## factor(Gender)Male          0.0017635  0.0069683   0.253  0.80034
## Bedtime                    -0.0004761  0.0006268  -0.760  0.44803
## Sleep.duration              0.0020734  0.0035846   0.578  0.56334
```

```
## REM.sleep.percentage       0.0066984  0.0009398   7.128 5.25e-12 ***
## Deep.sleep.percentage      0.0055716  0.0002378  23.429  < 2e-16 ***
## Awakenings                -0.0317907  0.0025290 -12.570  < 2e-16 ***
## Caffeine.consumption       0.0002450  0.0001135   2.159  0.03150 *
## Alcohol.consumption       -0.0061233  0.0021163  -2.893  0.00403 **
## factor(Smoking.status)Yes -0.0449020  0.0069311  -6.478 2.91e-10 ***
## Exercise.frequency         0.0056560  0.0024895   2.272  0.02366 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06074 on 376 degrees of freedom
## Multiple R-squared:  0.8054, Adjusted R-squared:  0.7997
## F-statistic: 141.4 on 11 and 376 DF,  p-value: < 2.2e-16
```

Conducting best subset method for finding significant variables

```
best.subset<- regsubsets(Sleep.efficiency~Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage
```

Summary of best subset model

```
reg.summary <- summary(best.subset)
summary(best.subset)
```

```
## Subset selection object
## Call: regsubsets.formula(Sleep.efficiency ~ Age + factor(Gender) +
##     Bedtime + Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency, data = df, nv = 11)
## 11 Variables  (and intercept)
##                            Forced in Forced out
## Age                            FALSE      FALSE
## factor(Gender)Male             FALSE      FALSE
## Bedtime                        FALSE      FALSE
## Sleep.duration                 FALSE      FALSE
## REM.sleep.percentage           FALSE      FALSE
## Deep.sleep.percentage          FALSE      FALSE
## Awakenings                     FALSE      FALSE
## Caffeine.consumption           FALSE      FALSE
## Alcohol.consumption            FALSE      FALSE
## factor(Smoking.status)Yes      FALSE      FALSE
## Exercise.frequency             FALSE      FALSE
## 1 subsets of each size up to 11
## Selection Algorithm: exhaustive
##          Age factor(Gender)Male Bedtime Sleep.duration REM.sleep.percentage
## 1  ( 1 )  " " " "                " "     " "            " "
## 2  ( 1 )  " " " "                " "     " "            " "
## 3  ( 1 )  " " " "                " "     " "            "*"
## 4  ( 1 )  " " " "                " "     " "            "*"
## 5  ( 1 )  "*" " "                " "     " "            "*"
## 6  ( 1 )  "*" " "                " "     " "            "*"
```

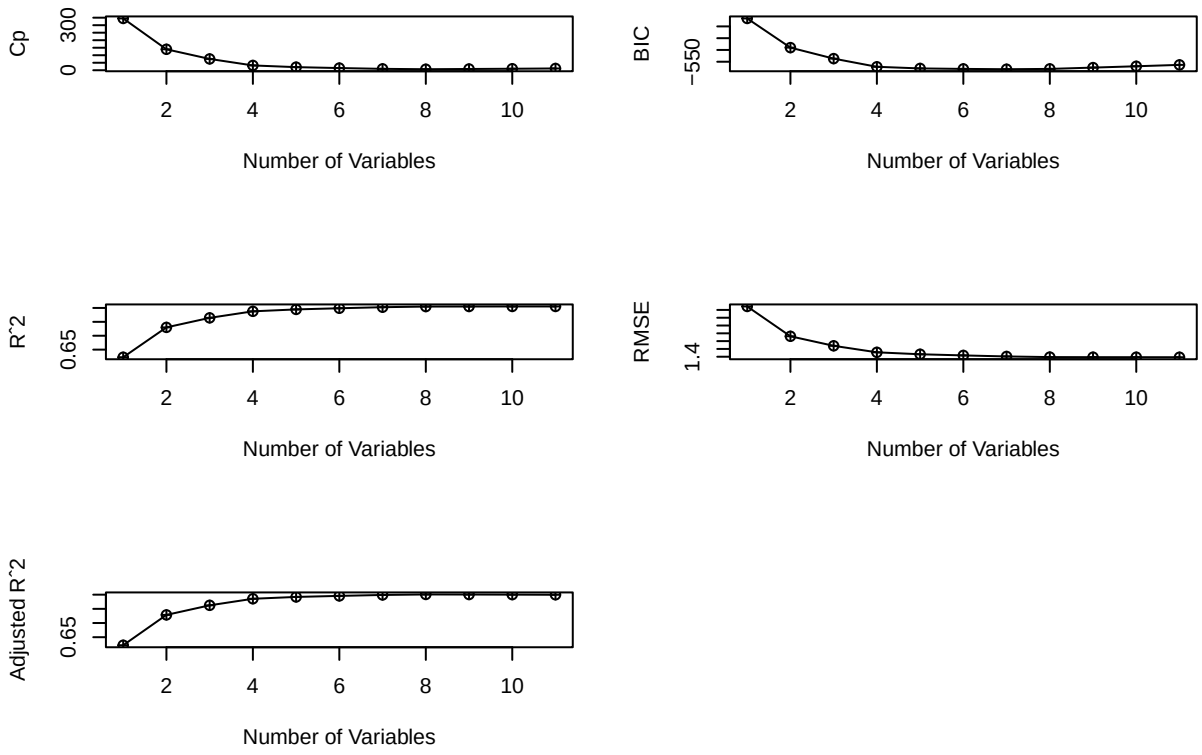
```
## 7  ( 1 )  "*" " "                 " "      " "             "*"
## 8  ( 1 )  "*" " "                 " "      " "             "*"
## 9  ( 1 )  "*" " "                 "*"      " "             "*"
## 10  ( 1 ) "*" " "                 "*"      "*"             "*"
## 11  ( 1 ) "*" "*"                 "*"      "*"             "*"
##           Deep.sleep.percentage Awakenings Caffeine.consumption
## 1  ( 1 )  "*"                   " "        " "
## 2  ( 1 )  "*"                   "*"        " "
## 3  ( 1 )  "*"                   "*"        " "
## 4  ( 1 )  "*"                   "*"        " "
## 5  ( 1 )  "*"                   "*"        " "
## 6  ( 1 )  "*"                   "*"        " "
## 7  ( 1 )  "*"                   "*"        " "
## 8  ( 1 )  "*"                   "*"        "*"
## 9  ( 1 )  "*"                   "*"        "*"
## 10  ( 1 ) "*"                   "*"        "*"
## 11  ( 1 ) "*"                   "*"        "*"
##           Alcohol.consumption factor(Smoking.status)Yes Exercise.frequency
## 1  ( 1 )  " "                 " "                       " "
## 2  ( 1 )  " "                 " "                       " "
## 3  ( 1 )  " "                 " "                       " "
## 4  ( 1 )  " "                 "*"                       " "
## 5  ( 1 )  " "                 "*"                       " "
## 6  ( 1 )  "*"                 "*"                       " "
## 7  ( 1 )  "*"                 "*"                       "*"
## 8  ( 1 )  "*"                 "*"                       "*"
## 9  ( 1 )  "*"                 "*"                       "*"
## 10  ( 1 ) "*"                 "*"                       "*"
## 11  ( 1 ) "*"                 "*"                       "*"
```

```
rsquare<-c(reg.summary$rsq)
cp<-c(reg.summary$cp)
AdjustedR<-c(reg.summary$adjr2)
RMSE<-c(reg.summary$rss)
BIC<-c(reg.summary$bic)
cbind(rsquare,cp,BIC,RMSE,AdjustedR)
```

```
##         rsquare         cp       BIC     RMSE AdjustedR
##  [1,] 0.6226586 344.914740 -366.2247 2.689344 0.6216810
##  [2,] 0.7300297 139.504685 -490.1810 1.924101 0.7286273
##  [3,] 0.7642764  75.349899 -536.8530 1.680021 0.7624349
##  [4,] 0.7877469  32.011795 -571.5856 1.512746 0.7855301
##  [5,] 0.7946762  20.626359 -578.5028 1.463360 0.7919887
##  [6,] 0.7988358  14.591269 -580.4828 1.433714 0.7956678
##  [7,] 0.8026031   9.313934 -581.8570 1.406864 0.7989668
##  [8,] 0.8049135   6.850856 -580.4641 1.390398 0.8007956
##  [9,] 0.8051562   8.381974 -574.9861 1.388668 0.8005171
## [10,] 0.8053208  10.064050 -569.3530 1.387495 0.8001569
## [11,] 0.8053540  12.000000 -563.4581 1.387259 0.7996595
```

```
par(mfrow=c(3,2)) # split the plotting panel into a 3 x 2 grid
plot(reg.summary$cp,type = "o",pch=10, xlab="Number of Variables",ylab= "Cp")
plot(reg.summary$bic,type = "o",pch=10, xlab="Number of Variables",ylab= "BIC")
```

```r
plot(reg.summary$rsq,type = "o",pch=10, xlab="Number of Variables",ylab= "R^2")
plot(reg.summary$rss,type = "o",pch=10, xlab="Number of Variables",ylab= "RMSE")
plot(reg.summary$adjr2,type = "o",pch=10, xlab="Number of Variables",ylab= "Adjusted R^2")
```



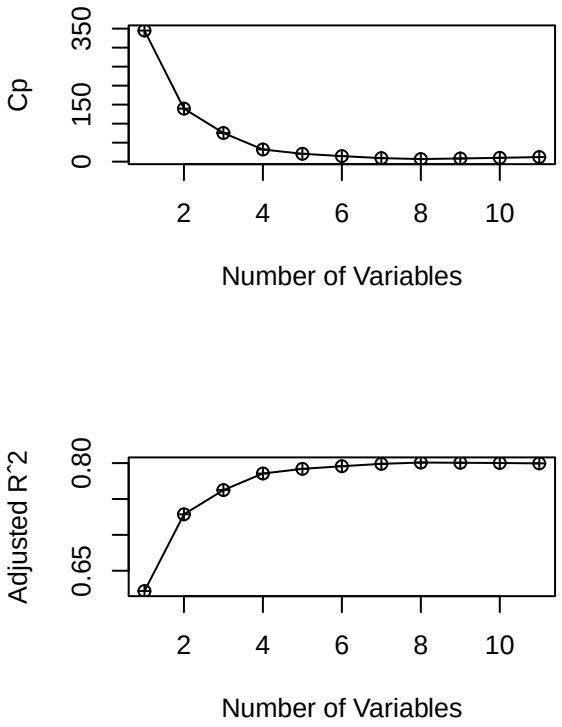
Conducting stepwise method to compare to best subset method

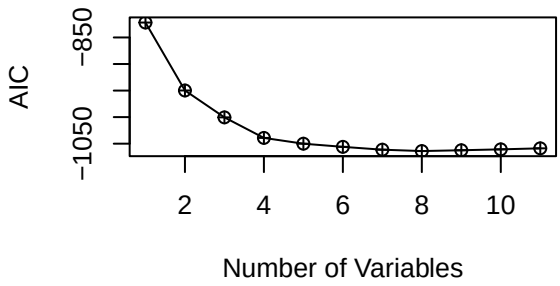
```r
ks=ols_step_best_subset(basemod, details=TRUE)
# for the output interpretation
AdjustedR<-c(ks$adjr)
cp<-c(ks$cp)
AIC<-c(ks$aic)
cbind(AdjustedR,cp,AIC)
```

```
##       AdjustedR         cp        AIC
##  [1,] 0.6216810 344.914740  -821.9265
##  [2,] 0.7286273 139.504685  -949.8438
##  [3,] 0.7624349  75.349899 -1000.4768
##  [4,] 0.7855301  32.011795 -1039.1704
##  [5,] 0.7919887  20.626359 -1050.0486
##  [6,] 0.7956678  14.591269 -1055.9896
##  [7,] 0.7989668   9.313934 -1061.3248
##  [8,] 0.8007956   6.850856 -1063.8929
##  [9,] 0.8005171   8.381974 -1062.3759
```

```
## [10,] 0.8001569  10.064050 -1060.7038
## [11,] 0.7996595  12.000000 -1058.7699
```

```
par(mfrow=c(2,2)) # split the plotting panel into a 2 x 2 grid
plot(ks$cp,type = "o",pch=10, xlab="Number of Variables",ylab= "Cp")
plot(ks$aic,type = "o",pch=10, xlab="Number of Variables",ylab= "AIC")
plot(ks$adjr,type = "o",pch=10, xlab="Number of Variables",ylab= "Adjusted R^2")
```





We will choose 8 and 11 variable models to compare for the final model.

The 8-variable model is chosen based on the high adjusted R-squared and low CP

The 11-variable model is chosen based on the CP being equal to p+1 where p is the number of predictors. This implies low bias even though the CP value is not as low.

```
vars <- c(ks$predictors)
vars
```

```
##  [1] "Deep.sleep.percentage"
##  [2] "Deep.sleep.percentage Awakenings"
##  [3] "REM.sleep.percentage Deep.sleep.percentage Awakenings"
##  [4] "REM.sleep.percentage Deep.sleep.percentage Awakenings factor(Smoking.status)"
##  [5] "Age REM.sleep.percentage Deep.sleep.percentage Awakenings factor(Smoking.status)"
```

```
##  [6] "Age REM.sleep.percentage Deep.sleep.percentage Awakenings Alcohol.consumption factor(Smoking.s
##  [7] "Age REM.sleep.percentage Deep.sleep.percentage Awakenings Alcohol.consumption factor(Smoking.s
##  [8] "Age REM.sleep.percentage Deep.sleep.percentage Awakenings Caffeine.consumption Alcohol.consump
##  [9] "Age Bedtime REM.sleep.percentage Deep.sleep.percentage Awakenings Caffeine.consumption Alcohol
## [10] "Age Bedtime Sleep.duration REM.sleep.percentage Deep.sleep.percentage Awakenings Caffeine.consu
## [11] "Age factor(Gender) Bedtime Sleep.duration REM.sleep.percentage Deep.sleep.percentage Awakenings
```

Evaluating the 11-variable model:

Testing interaction terms for the 11-variable model

```
intmod <- lm(Sleep.efficiency~(Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage+ Deep.slee

summary(intmod)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ (Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency)^2, data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.164211 -0.028762  0.004846  0.031347  0.132248
##
## Coefficients:
##                                                Estimate Std. Error t value
## (Intercept)                                  -1.326e-02  3.192e-01  -0.042
## Age                                           4.375e-03  3.697e-03   1.183
## factor(Gender)Male                            1.454e-01  1.130e-01   1.287
## Bedtime                                       6.130e-03  8.919e-03   0.687
## Sleep.duration                                9.945e-03  3.681e-02   0.270
## REM.sleep.percentage                          8.241e-03  1.101e-02   0.749
## Deep.sleep.percentage                         1.080e-02  2.634e-03   4.099
## Awakenings                                    5.006e-02  3.675e-02   1.362
## Caffeine.consumption                          1.708e-03  2.028e-03   0.843
## Alcohol.consumption                           6.804e-02  2.678e-02   2.541
## factor(Smoking.status)Yes                    -2.128e-01  9.606e-02  -2.216
## Exercise.frequency                           -3.265e-02  3.956e-02  -0.825
## Age:factor(Gender)Male                       -9.081e-05  6.128e-04  -0.148
## Age:Bedtime                                  -7.143e-05  5.771e-05  -1.238
## Age:Sleep.duration                            1.787e-04  3.270e-04   0.546
## Age:REM.sleep.percentage                     -6.786e-06  8.038e-05  -0.084
## Age:Deep.sleep.percentage                    -7.305e-05  1.695e-05  -4.309
## Age:Awakenings                               -5.291e-04  1.928e-04  -2.744
## Age:Caffeine.consumption                     -5.616e-06  1.111e-05  -0.506
## Age:Alcohol.consumption                      -3.588e-05  1.604e-04  -0.224
## Age:factor(Smoking.status)Yes                -1.229e-03  5.643e-04  -2.178
## Age:Exercise.frequency                        1.016e-04  2.314e-04   0.439
## factor(Gender)Male:Bedtime                    1.530e-03  1.702e-03   0.899
## factor(Gender)Male:Sleep.duration            -1.240e-02  1.017e-02  -1.220
```

```
## factor(Gender)Male:REM.sleep.percentage          1.466e-03  2.258e-03   0.649
## factor(Gender)Male:Deep.sleep.percentage        -1.154e-03  6.247e-04  -1.847
## factor(Gender)Male:Awakenings                    5.722e-03  5.992e-03   0.955
## factor(Gender)Male:Caffeine.consumption         -2.808e-04  2.696e-04  -1.041
## factor(Gender)Male:Alcohol.consumption          -2.110e-02  5.181e-03  -4.072
## factor(Gender)Male:factor(Smoking.status)Yes     1.528e-02  1.964e-02   0.778
## factor(Gender)Male:Exercise.frequency            8.191e-03  6.527e-03   1.255
## Bedtime:Sleep.duration                          -4.426e-04  6.888e-04  -0.643
## Bedtime:REM.sleep.percentage                    -9.283e-05  2.086e-04  -0.445
## Bedtime:Deep.sleep.percentage                    3.489e-05  4.748e-05   0.735
## Bedtime:Awakenings                               1.758e-04  4.710e-04   0.373
## Bedtime:Caffeine.consumption                    -2.011e-05  3.205e-05  -0.627
## Bedtime:Alcohol.consumption                      2.581e-04  4.375e-04   0.590
## Bedtime:factor(Smoking.status)Yes                1.105e-03  1.402e-03   0.788
## Bedtime:Exercise.frequency                      -8.336e-04  5.623e-04  -1.482
## Sleep.duration:REM.sleep.percentage             -1.117e-05  1.191e-03  -0.009
## Sleep.duration:Deep.sleep.percentage            -1.552e-04  2.580e-04  -0.601
## Sleep.duration:Awakenings                       -1.325e-03  3.150e-03  -0.420
## Sleep.duration:Caffeine.consumption              1.348e-04  1.893e-04   0.712
## Sleep.duration:Alcohol.consumption              -4.863e-03  2.425e-03  -2.006
## Sleep.duration:factor(Smoking.status)Yes        -6.357e-03  8.088e-03  -0.786
## Sleep.duration:Exercise.frequency                2.588e-03  3.208e-03   0.807
## REM.sleep.percentage:Deep.sleep.percentage       1.899e-05  7.173e-05   0.265
## REM.sleep.percentage:Awakenings                 -7.712e-04  7.765e-04  -0.993
## REM.sleep.percentage:Caffeine.consumption       -3.840e-05  3.233e-05  -1.188
## REM.sleep.percentage:Alcohol.consumption        -1.415e-03  6.548e-04  -2.160
## REM.sleep.percentage:factor(Smoking.status)Yes   2.806e-03  2.214e-03   1.268
## REM.sleep.percentage:Exercise.frequency          1.101e-04  9.377e-04   0.117
## Deep.sleep.percentage:Awakenings                -6.992e-04  1.893e-04  -3.694
## Deep.sleep.percentage:Caffeine.consumption      -2.591e-05  1.303e-05  -1.988
## Deep.sleep.percentage:Alcohol.consumption       -1.108e-04  1.456e-04  -0.761
## Deep.sleep.percentage:factor(Smoking.status)Yes  3.102e-03  5.272e-04   5.883
## Deep.sleep.percentage:Exercise.frequency         8.579e-05  2.130e-04   0.403
## Awakenings:Caffeine.consumption                  6.485e-05  1.085e-04   0.598
## Awakenings:Alcohol.consumption                   3.768e-03  1.678e-03   2.245
## Awakenings:factor(Smoking.status)Yes             1.132e-02  5.630e-03   2.011
## Awakenings:Exercise.frequency                   -3.027e-03  1.914e-03  -1.581
## Caffeine.consumption:Alcohol.consumption        -3.684e-04  1.249e-04  -2.948
## Caffeine.consumption:factor(Smoking.status)Yes   4.800e-04  3.627e-04   1.323
## Caffeine.consumption:Exercise.frequency          4.381e-05  1.373e-04   0.319
## Alcohol.consumption:factor(Smoking.status)Yes    7.870e-03  4.971e-03   1.583
## Alcohol.consumption:Exercise.frequency           4.696e-03  1.732e-03   2.712
## factor(Smoking.status)Yes:Exercise.frequency    -3.248e-03  6.781e-03  -0.479
##                                                 Pr(>|t|)
## (Intercept)                                     0.966898
## Age                                             0.237489
## factor(Gender)Male                              0.199093
## Bedtime                                         0.492345
## Sleep.duration                                  0.787197
## REM.sleep.percentage                            0.454633
## Deep.sleep.percentage                           5.26e-05 ***
## Awakenings                                      0.174086
## Caffeine.consumption                            0.400083
## Alcohol.consumption                             0.011533 *
```

```
## factor(Smoking.status)Yes                          0.027422 *
## Exercise.frequency                                 0.409746
## Age:factor(Gender)Male                             0.882289
## Age:Bedtime                                        0.216709
## Age:Sleep.duration                                 0.585228
## Age:REM.sleep.percentage                           0.932771
## Age:Deep.sleep.percentage                          2.19e-05 ***
## Age:Awakenings                                     0.006411 **
## Age:Caffeine.consumption                           0.613553
## Age:Alcohol.consumption                            0.823118
## Age:factor(Smoking.status)Yes                      0.030165 *
## Age:Exercise.frequency                             0.661053
## factor(Gender)Male:Bedtime                         0.369074
## factor(Gender)Male:Sleep.duration                  0.223374
## factor(Gender)Male:REM.sleep.percentage            0.516834
## factor(Gender)Male:Deep.sleep.percentage           0.065612 .
## factor(Gender)Male:Awakenings                      0.340392
## factor(Gender)Male:Caffeine.consumption            0.298456
## factor(Gender)Male:Alcohol.consumption             5.88e-05 ***
## factor(Gender)Male:factor(Smoking.status)Yes       0.437133
## factor(Gender)Male:Exercise.frequency              0.210446
## Bedtime:Sleep.duration                             0.520966
## Bedtime:REM.sleep.percentage                       0.656573
## Bedtime:Deep.sleep.percentage                      0.463047
## Bedtime:Awakenings                                 0.709125
## Bedtime:Caffeine.consumption                       0.530920
## Bedtime:Alcohol.consumption                        0.555619
## Bedtime:factor(Smoking.status)Yes                  0.431391
## Bedtime:Exercise.frequency                         0.139200
## Sleep.duration:REM.sleep.percentage                0.992524
## Sleep.duration:Deep.sleep.percentage               0.547976
## Sleep.duration:Awakenings                          0.674447
## Sleep.duration:Caffeine.consumption                0.476831
## Sleep.duration:Alcohol.consumption                 0.045702 *
## Sleep.duration:factor(Smoking.status)Yes           0.432429
## Sleep.duration:Exercise.frequency                  0.420435
## REM.sleep.percentage:Deep.sleep.percentage         0.791394
## REM.sleep.percentage:Awakenings                    0.321390
## REM.sleep.percentage:Caffeine.consumption          0.235800
## REM.sleep.percentage:Alcohol.consumption           0.031495 *
## REM.sleep.percentage:factor(Smoking.status)Yes     0.205862
## REM.sleep.percentage:Exercise.frequency            0.906610
## Deep.sleep.percentage:Awakenings                   0.000259 ***
## Deep.sleep.percentage:Caffeine.consumption         0.047615 *
## Deep.sleep.percentage:Alcohol.consumption          0.447209
## Deep.sleep.percentage:factor(Smoking.status)Yes    1.01e-08 ***
## Deep.sleep.percentage:Exercise.frequency           0.687368
## Awakenings:Caffeine.consumption                    0.550387
## Awakenings:Alcohol.consumption                     0.025431 *
## Awakenings:factor(Smoking.status)Yes               0.045167 *
## Awakenings:Exercise.frequency                      0.114813
## Caffeine.consumption:Alcohol.consumption           0.003431 **
## Caffeine.consumption:factor(Smoking.status)Yes     0.186684
## Caffeine.consumption:Exercise.frequency            0.749886
```

```
## Alcohol.consumption:factor(Smoking.status)Yes   0.114308
## Alcohol.consumption:Exercise.frequency          0.007051 **
## factor(Smoking.status)Yes:Exercise.frequency    0.632211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05288 on 321 degrees of freedom
## Multiple R-squared:  0.8741, Adjusted R-squared:  0.8482
## F-statistic: 33.76 on 66 and 321 DF,  p-value: < 2.2e-16
```

```
intmod <- lm(Sleep.efficiency~Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage+ Deep.sleep

summary(intmod)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency + Age * factor(Smoking.status) +
##     Age * Deep.sleep.percentage + Age * Awakenings + Age:factor(Smoking.status) +
##     factor(Gender):Alcohol.consumption + Sleep.duration:Alcohol.consumption +
##     REM.sleep.percentage:Alcohol.consumption + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:Caffeine.consumption + Deep.sleep.percentage:factor(Smoking.status) +
##     Awakenings:Alcohol.consumption + Awakenings:factor(Smoking.status) +
##     Awakenings:Exercise.frequency + Caffeine.consumption:Alcohol.consumption +
##     Alcohol.consumption:Exercise.frequency, data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.154006 -0.031054  0.003856  0.034966  0.138319
##
## Coefficients:
##                                                 Estimate Std. Error t value
## (Intercept)                                    7.268e-02  6.001e-02   1.211
## Age                                            5.474e-03  8.761e-04   6.248
## factor(Gender)Male                             2.175e-02  7.706e-03   2.822
## Bedtime                                       -7.493e-05  5.700e-04  -0.131
## Sleep.duration                                 5.516e-03  3.785e-03   1.457
## REM.sleep.percentage                           8.603e-03  1.047e-03   8.218
## Deep.sleep.percentage                          8.758e-03  7.382e-04  11.863
## Awakenings                                     1.986e-02  1.137e-02   1.746
## Caffeine.consumption                           5.237e-04  4.882e-04   1.073
## Alcohol.consumption                            3.969e-02  2.006e-02   1.979
## factor(Smoking.status)Yes                     -1.402e-01  3.017e-02  -4.648
## Exercise.frequency                             6.628e-03  3.245e-03   2.043
## Age:factor(Smoking.status)Yes                 -1.428e-03  4.634e-04  -3.082
## Age:Deep.sleep.percentage                     -6.855e-05  1.364e-05  -5.025
## Age:Awakenings                                -5.170e-04  1.626e-04  -3.180
## factor(Gender)Male:Alcohol.consumption        -1.463e-02  4.361e-03  -3.354
## Sleep.duration:Alcohol.consumption            -3.998e-03  1.997e-03  -2.002
## REM.sleep.percentage:Alcohol.consumption      -1.026e-03  5.022e-04  -2.042
## Deep.sleep.percentage:Awakenings              -6.131e-04  1.621e-04  -3.783
## Deep.sleep.percentage:Caffeine.consumption    -5.956e-06  8.176e-06  -0.728
```

```
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.531e-03  4.041e-04   6.263
## Awakenings:Alcohol.consumption                    4.333e-03  1.445e-03   2.999
## Awakenings:factor(Smoking.status)Yes              1.304e-02  4.923e-03   2.648
## Awakenings:Exercise.frequency                    -2.806e-03  1.532e-03  -1.831
## Caffeine.consumption:Alcohol.consumption         -1.961e-04  1.013e-04  -1.935
## Alcohol.consumption:Exercise.frequency            4.722e-03  1.400e-03   3.373
##                                                  Pr(>|t|)
## (Intercept)                                      0.226650
## Age                                              1.17e-09 ***
## factor(Gender)Male                               0.005035 **
## Bedtime                                          0.895482
## Sleep.duration                                   0.145941
## REM.sleep.percentage                             3.72e-15 ***
## Deep.sleep.percentage                             < 2e-16 ***
## Awakenings                                       0.081606 .
## Caffeine.consumption                             0.284189
## Alcohol.consumption                              0.048601 *
## factor(Smoking.status)Yes                        4.71e-06 ***
## Exercise.frequency                               0.041807 *
## Age:factor(Smoking.status)Yes                    0.002211 **
## Age:Deep.sleep.percentage                        7.92e-07 ***
## Age:Awakenings                                   0.001599 **
## factor(Gender)Male:Alcohol.consumption           0.000880 ***
## Sleep.duration:Alcohol.consumption               0.046060 *
## REM.sleep.percentage:Alcohol.consumption         0.041829 *
## Deep.sleep.percentage:Awakenings                 0.000182 ***
## Deep.sleep.percentage:Caffeine.consumption       0.466788
## Deep.sleep.percentage:factor(Smoking.status)Yes 1.07e-09 ***
## Awakenings:Alcohol.consumption                   0.002892 **
## Awakenings:factor(Smoking.status)Yes             0.008458 **
## Awakenings:Exercise.frequency                    0.067965 .
## Caffeine.consumption:Alcohol.consumption         0.053719 .
## Alcohol.consumption:Exercise.frequency           0.000824 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05237 on 362 degrees of freedom
## Multiple R-squared:  0.8607, Adjusted R-squared:  0.8511
## F-statistic: 89.46 on 25 and 362 DF,  p-value: < 2.2e-16
```

```
intmod <- lm(Sleep.efficiency~Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage+ Deep.sleep

summary(intmod)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency + Age * factor(Smoking.status) +
##     Age * Deep.sleep.percentage + Age * Awakenings + Age:factor(Smoking.status) +
##     factor(Gender):Alcohol.consumption + Sleep.duration:Alcohol.consumption +
##     REM.sleep.percentage:Alcohol.consumption + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
```

```
##     Awakenings:factor(Smoking.status) + Awakenings:Exercise.frequency +
##     Alcohol.consumption:Exercise.frequency, data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.158714 -0.030835  0.004594  0.036660  0.132823
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                     7.057e-02  5.912e-02   1.194
## Age                                             5.554e-03  8.772e-04   6.332
## factor(Gender)Male                              2.029e-02  7.660e-03   2.649
## Bedtime                                        -1.463e-04  5.618e-04  -0.260
## Sleep.duration                                  5.764e-03  3.792e-03   1.520
## REM.sleep.percentage                            8.768e-03  1.046e-03   8.384
## Deep.sleep.percentage                           8.702e-03  7.065e-04  12.318
## Awakenings                                      2.063e-02  1.137e-02   1.815
## Caffeine.consumption                            1.000e-04  1.028e-04   0.973
## Alcohol.consumption                             3.431e-02  1.980e-02   1.733
## factor(Smoking.status)Yes                      -1.398e-01  3.011e-02  -4.643
## Exercise.frequency                              6.123e-03  3.221e-03   1.901
## Age:factor(Smoking.status)Yes                  -1.479e-03  4.559e-04  -3.245
## Age:Deep.sleep.percentage                      -6.903e-05  1.365e-05  -5.057
## Age:Awakenings                                 -5.248e-04  1.628e-04  -3.224
## factor(Gender)Male:Alcohol.consumption          -1.018e-02  3.626e-03  -2.808
## Sleep.duration:Alcohol.consumption              -4.084e-03  1.999e-03  -2.043
## REM.sleep.percentage:Alcohol.consumption        -1.040e-03  5.029e-04  -2.067
## Deep.sleep.percentage:Awakenings                -6.121e-04  1.620e-04  -3.778
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.580e-03  4.041e-04   6.385
## Awakenings:Alcohol.consumption                   4.096e-03  1.443e-03   2.839
## Awakenings:factor(Smoking.status)Yes             1.294e-02  4.934e-03   2.622
## Awakenings:Exercise.frequency                   -3.031e-03  1.531e-03  -1.979
## Alcohol.consumption:Exercise.frequency           5.481e-03  1.338e-03   4.095
##                                                 Pr(>|t|)
## (Intercept)                                     0.233442
## Age                                             7.15e-10 ***
## factor(Gender)Male                              0.008429 **
## Bedtime                                         0.794771
## Sleep.duration                                  0.129338
## REM.sleep.percentage                            1.14e-15 ***
## Deep.sleep.percentage                            < 2e-16 ***
## Awakenings                                      0.070317 .
## Caffeine.consumption                            0.331172
## Alcohol.consumption                             0.083953 .
## factor(Smoking.status)Yes                       4.80e-06 ***
## Exercise.frequency                              0.058107 .
## Age:factor(Smoking.status)Yes                   0.001284 **
## Age:Deep.sleep.percentage                       6.77e-07 ***
## Age:Awakenings                                  0.001378 **
## factor(Gender)Male:Alcohol.consumption          0.005249 **
## Sleep.duration:Alcohol.consumption              0.041779 *
## REM.sleep.percentage:Alcohol.consumption        0.039418 *
## Deep.sleep.percentage:Awakenings                0.000185 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 5.22e-10 ***
```

```
## Awakenings:Alcohol.consumption                  0.004780 **
## Awakenings:factor(Smoking.status)Yes            0.009096 **
## Awakenings:Exercise.frequency                   0.048548 *
## Alcohol.consumption:Exercise.frequency          5.20e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0525 on 364 degrees of freedom
## Multiple R-squared:  0.8592, Adjusted R-squared:  0.8503
## F-statistic: 96.61 on 23 and 364 DF,  p-value: < 2.2e-16
```

Splitting data to prepare for pairs plots

```
half1data <- subset(df, select=c("Sleep.efficiency","Sleep.duration","REM.sleep.percentage","Deep.sleep
```

```
half2data <- subset(df, select=c("Sleep.efficiency","Age","Bedtime","Exercise.frequency"))
```

```
half3data <- subset(df, select=c("Sleep.efficiency","Caffeine.consumption","Alcohol.consumption","Awake
```

```
agedata <- subset(df, select=c("Sleep.efficiency","Age"))
```

```
ggpairs(half1data,lower = list(continuous = "smooth_loess", combo =
 "facethist", discrete = "facetbar", na = "na"))
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : pseudoinverse used at 7
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : neighborhood radius 1
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : reciprocal condition number 0
```

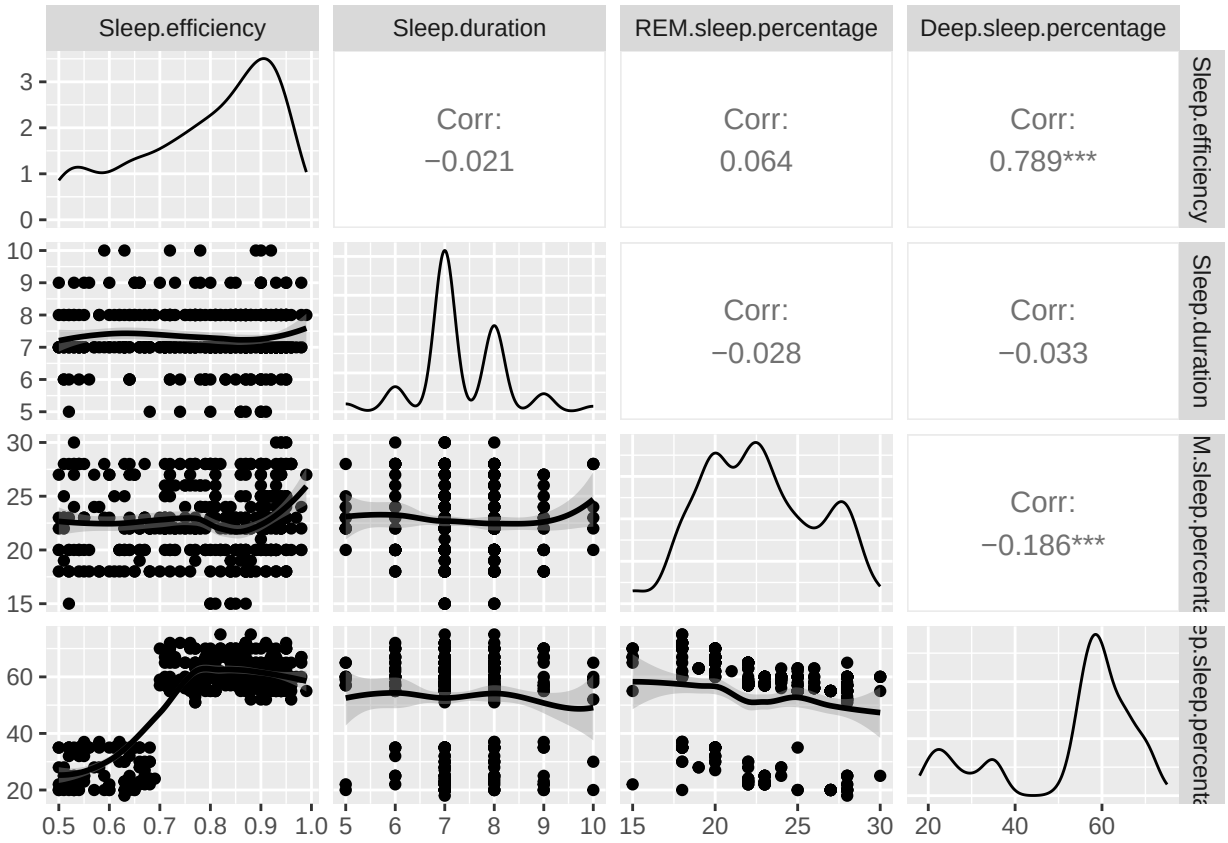
```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : There are other near singularities as well. 1
```

```
## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used at 7
```

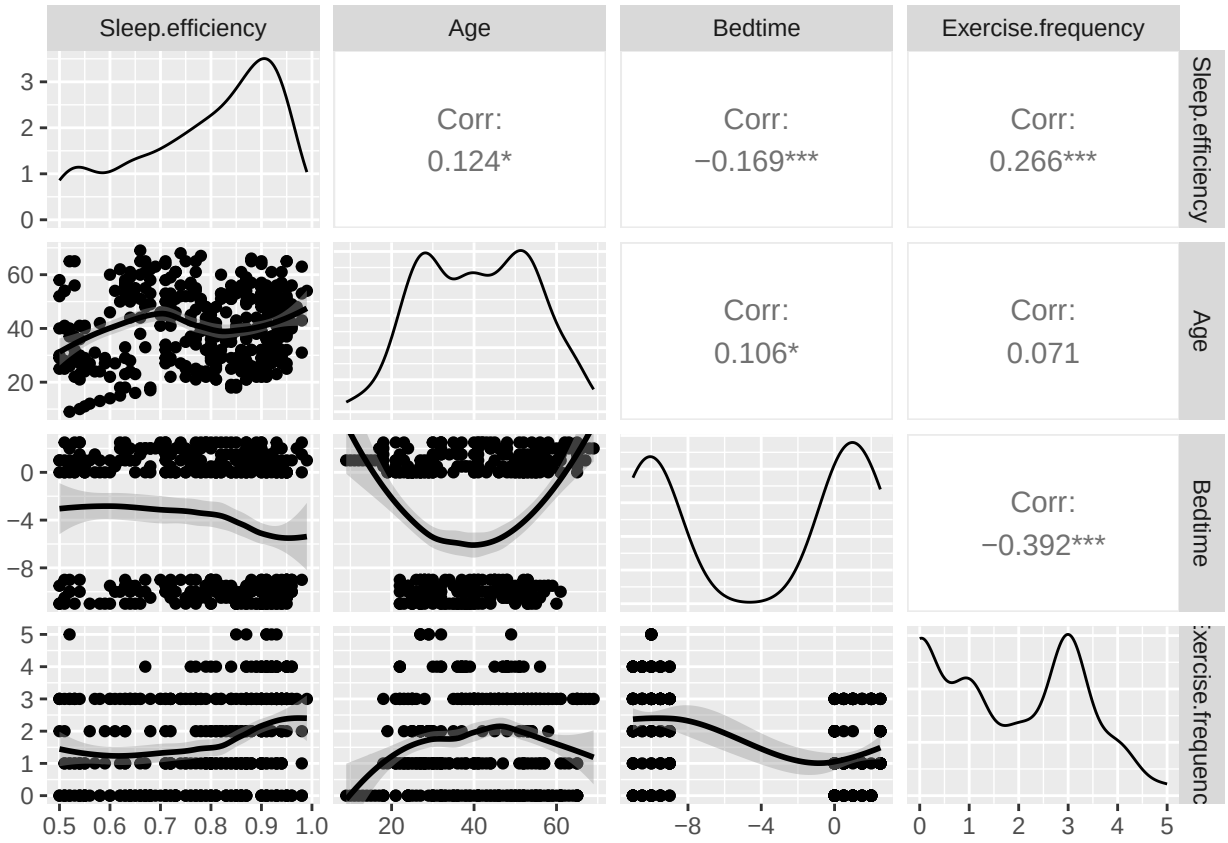
```
## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius 1
```

```
## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal condition
## number 0
```

```
## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other near
## singularities as well. 1

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : pseudoinverse used at 7

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : neighborhood radius 1

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : There are other near singularities as well. 1

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used at 7

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius 1

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal condition
## number 0

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other near
## singularities as well. 1
```




```
ggpairs(half2data,lower = list(continuous = "smooth_loess", combo =
 "facethist", discrete = "facetbar", na = "na"))
```




```
ggpairs(half3data,lower = list(continuous = "smooth_loess", combo =
 "facethist", discrete = "facetbar", na = "na"))
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : pseudoinverse used at -1
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : neighborhood radius 51
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : reciprocal condition number 1.5249e-16
```

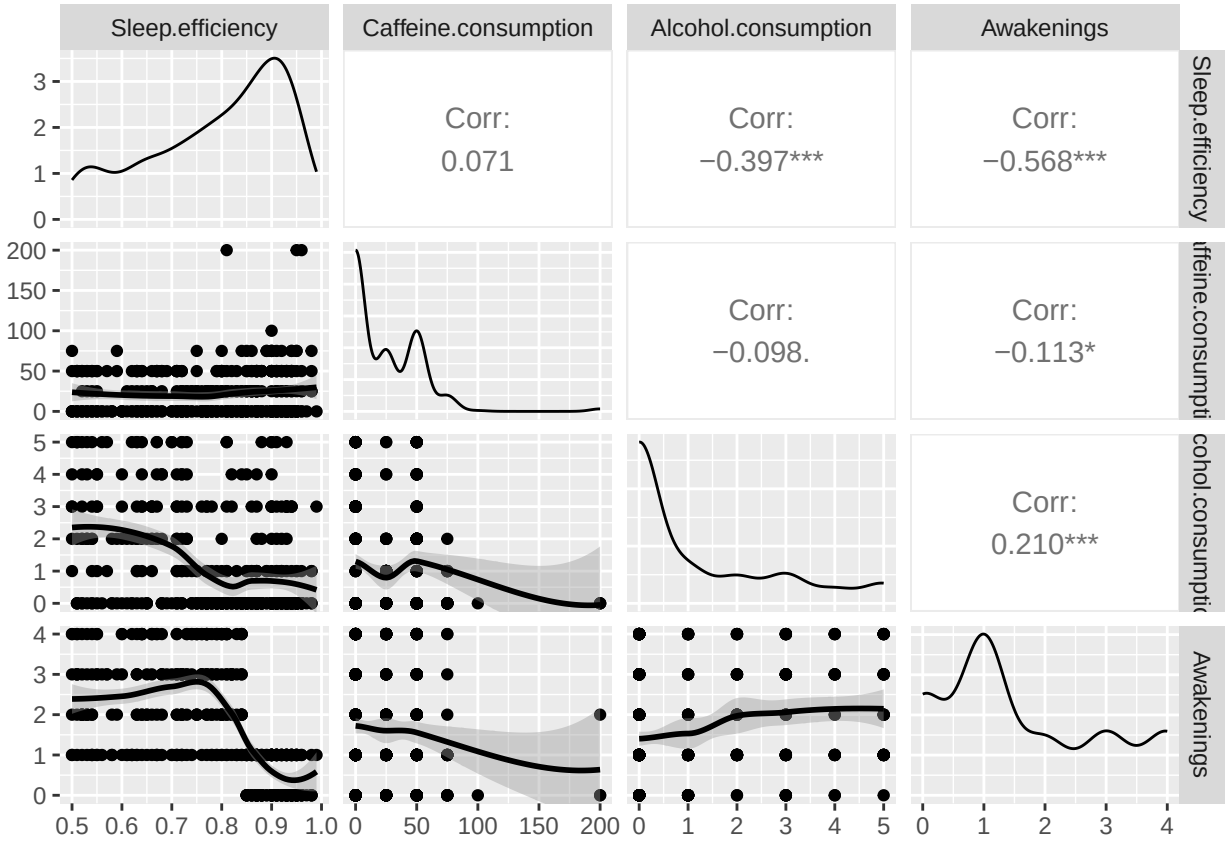
```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : There are other near singularities as well. 2500
```

```
## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used at
## -1
```

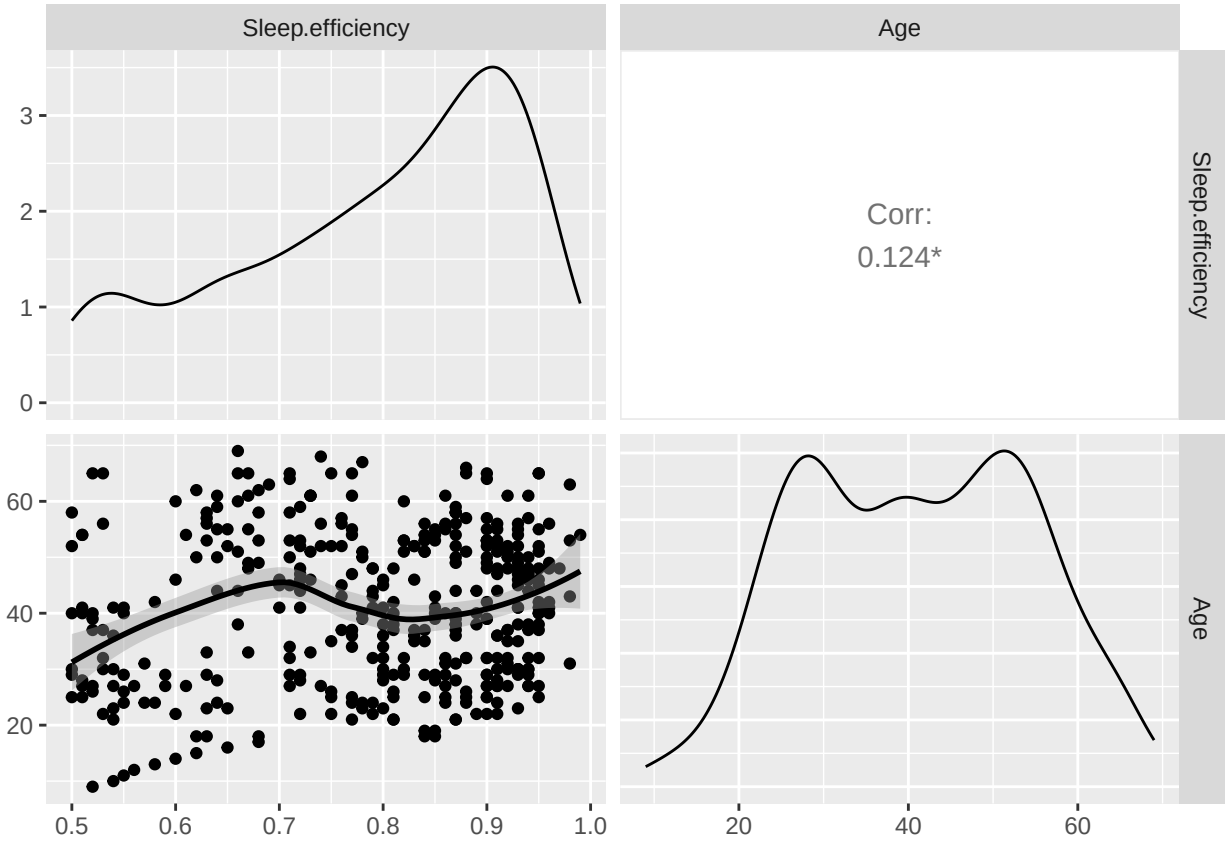
```
## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius 51
```

```
## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal condition
## number 1.5249e-16

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other near
## singularities as well. 2500

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : pseudoinverse used at -1

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : neighborhood radius 51

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : reciprocal condition number 1.5249e-16

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : There are other near singularities as well. 2500

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used at
## -1

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius 51

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal condition
## number 1.5249e-16

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other near
## singularities as well. 2500

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : pseudoinverse used at -0.025

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : neighborhood radius 2.025

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : reciprocal condition number 5.5142e-16
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,
## : There are other near singularities as well. 4

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used at
## -0.025

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.025

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal condition
## number 5.5142e-16

## Warning in predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x
## else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other near
## singularities as well. 4
```



```
ggpairs(agedata,lower = list(continuous = "smooth_loess", combo =
 "facethist", discrete = "facetbar", na = "na"))
```




```
elevenpowertry <- lm(Sleep.efficiency~Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage+ D

summary(elevenpowertry)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency + Age * factor(Smoking.status) +
##     Age * Deep.sleep.percentage + Age * Awakenings + Age:factor(Smoking.status) +
##     factor(Gender):Alcohol.consumption + Sleep.duration:Alcohol.consumption +
##     REM.sleep.percentage:Alcohol.consumption + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Awakenings:factor(Smoking.status) + Awakenings:Exercise.frequency +
##     Alcohol.consumption:Exercise.frequency + I(Age^2), data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.155735 -0.029843  0.003238  0.034601  0.136865
##
## Coefficients:
```

```
##                                                  Estimate Std. Error t value
## (Intercept)                                     5.078e-02  6.028e-02   0.842
## Age                                             7.689e-03  1.596e-03   4.819
## factor(Gender)Male                              1.908e-02  7.681e-03   2.484
## Bedtime                                         1.693e-04  5.943e-04   0.285
## Sleep.duration                                  5.877e-03  3.784e-03   1.553
## REM.sleep.percentage                            8.475e-03  1.059e-03   8.000
## Deep.sleep.percentage                           8.483e-03  7.181e-04  11.813
## Awakenings                                      2.109e-02  1.135e-02   1.858
## Caffeine.consumption                            9.659e-05  1.026e-04   0.941
## Alcohol.consumption                             3.266e-02  1.978e-02   1.651
## factor(Smoking.status)Yes                      -1.488e-01  3.057e-02  -4.868
## Exercise.frequency                              5.641e-03  3.228e-03   1.748
## I(Age^2)                                       -2.861e-05  1.788e-05  -1.600
## Age:factor(Smoking.status)Yes                  -1.337e-03  4.636e-04  -2.884
## Age:Deep.sleep.percentage                      -6.629e-05  1.373e-05  -4.828
## Age:Awakenings                                 -5.103e-04  1.627e-04  -3.136
## factor(Gender)Male:Alcohol.consumption          -9.194e-03  3.670e-03  -2.505
## Sleep.duration:Alcohol.consumption             -4.234e-03  1.997e-03  -2.120
## REM.sleep.percentage:Alcohol.consumption        -9.670e-04  5.038e-04  -1.919
## Deep.sleep.percentage:Awakenings                -6.285e-04  1.620e-04  -3.880
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.633e-03  4.046e-04   6.509
## Awakenings:Alcohol.consumption                  3.990e-03  1.441e-03   2.769
## Awakenings:factor(Smoking.status)Yes            1.219e-02  4.946e-03   2.466
## Awakenings:Exercise.frequency                  -2.943e-03  1.529e-03  -1.925
## Alcohol.consumption:Exercise.frequency          5.769e-03  1.348e-03   4.281
##                                                Pr(>|t|)
## (Intercept)                                    0.400121
## Age                                            2.13e-06 ***
## factor(Gender)Male                             0.013458 *
## Bedtime                                        0.775928
## Sleep.duration                                 0.121273
## REM.sleep.percentage                           1.69e-14 ***
## Deep.sleep.percentage                           < 2e-16 ***
## Awakenings                                     0.063916 .
## Caffeine.consumption                           0.347088
## Alcohol.consumption                            0.099665 .
## factor(Smoking.status)Yes                      1.69e-06 ***
## Exercise.frequency                             0.081388 .
## I(Age^2)                                       0.110439
## Age:factor(Smoking.status)Yes                  0.004160 **
## Age:Deep.sleep.percentage                      2.04e-06 ***
## Age:Awakenings                                 0.001850 **
## factor(Gender)Male:Alcohol.consumption         0.012683 *
## Sleep.duration:Alcohol.consumption             0.034679 *
## REM.sleep.percentage:Alcohol.consumption       0.055724 .
## Deep.sleep.percentage:Awakenings               0.000124 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 2.52e-10 ***
## Awakenings:Alcohol.consumption                 0.005915 **
## Awakenings:factor(Smoking.status)Yes           0.014134 *
## Awakenings:Exercise.frequency                  0.055010 .
## Alcohol.consumption:Exercise.frequency         2.38e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.05239 on 363 degrees of freedom
## Multiple R-squared:  0.8602, Adjusted R-squared:  0.851
## F-statistic: 93.09 on 24 and 363 DF,  p-value: < 2.2e-16
```

Not significant - We will not use higher order terms

Based on the patterns in the plot above, we added in a squared term for Deep.sleep.percentage.

Evaluating the 8-variable model

```
eightvar <- lm(Sleep.efficiency~Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.consu
summary(eightvar)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency,
##     data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.17106 -0.04091  0.00488  0.03992  0.14676
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.3634105  0.0298044  12.193  < 2e-16 ***
## Age                         0.0009562  0.0002349   4.071 5.70e-05 ***
## REM.sleep.percentage        0.0066438  0.0009314   7.133 5.02e-12 ***
## Deep.sleep.percentage       0.0055625  0.0002365  23.518  < 2e-16 ***
## Awakenings                 -0.0318493  0.0024973 -12.753  < 2e-16 ***
## Caffeine.consumption        0.0002333  0.0001101   2.119  0.03477 *
## Alcohol.consumption        -0.0062024  0.0021082  -2.942  0.00346 **
## factor(Smoking.status)Yes -0.0457656  0.0066573  -6.875 2.57e-11 ***
## Exercise.frequency          0.0064591  0.0022283   2.899  0.00396 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06057 on 379 degrees of freedom
## Multiple R-squared:  0.8049, Adjusted R-squared:  0.8008
## F-statistic: 195.5 on 8 and 379 DF,  p-value: < 2.2e-16
```

```
eightint <- lm(Sleep.efficiency~(Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.cons
summary(eightint)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ (Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency)^2,
```

```
##     data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.173444 -0.029579  0.004725  0.032983  0.132355
##
## Coefficients:
##                                                 Estimate Std. Error t value
## (Intercept)                                    9.345e-02  1.210e-01   0.773
## Age                                            6.807e-03  2.021e-03   3.368
## REM.sleep.percentage                           8.024e-03  5.455e-03   1.471
## Deep.sleep.percentage                          8.883e-03  1.575e-03   5.641
## Awakenings                                     4.769e-02  2.226e-02   2.142
## Caffeine.consumption                           1.824e-03  1.092e-03   1.671
## Alcohol.consumption                            1.430e-02  1.830e-02   0.782
## factor(Smoking.status)Yes                     -2.565e-01  6.256e-02  -4.100
## Exercise.frequency                            -6.534e-03  2.215e-02  -0.295
## Age:REM.sleep.percentage                      -3.782e-05  7.219e-05  -0.524
## Age:Deep.sleep.percentage                     -7.302e-05  1.577e-05  -4.631
## Age:Awakenings                                -4.954e-04  1.757e-04  -2.819
## Age:Caffeine.consumption                      -9.260e-06  9.575e-06  -0.967
## Age:Alcohol.consumption                       -1.250e-04  1.484e-04  -0.842
## Age:factor(Smoking.status)Yes                 -1.323e-03  4.952e-04  -2.671
## Age:Exercise.frequency                         8.952e-05  1.782e-04   0.502
## REM.sleep.percentage:Deep.sleep.percentage     1.947e-05  6.470e-05   0.301
## REM.sleep.percentage:Awakenings               -9.131e-04  7.235e-04  -1.262
## REM.sleep.percentage:Caffeine.consumption     -3.099e-05  2.934e-05  -1.056
## REM.sleep.percentage:Alcohol.consumption      -1.025e-03  5.736e-04  -1.786
## REM.sleep.percentage:factor(Smoking.status)Yes  3.930e-03  2.020e-03   1.946
## REM.sleep.percentage:Exercise.frequency        5.973e-04  7.380e-04   0.809
## Deep.sleep.percentage:Awakenings              -7.222e-04  1.809e-04  -3.992
## Deep.sleep.percentage:Caffeine.consumption    -1.301e-05  1.044e-05  -1.246
## Deep.sleep.percentage:Alcohol.consumption     -1.444e-04  1.377e-04  -1.048
## Deep.sleep.percentage:factor(Smoking.status)Yes 3.021e-03  5.035e-04   5.999
## Deep.sleep.percentage:Exercise.frequency      -5.589e-05  1.843e-04  -0.303
## Awakenings:Caffeine.consumption               -3.682e-05  9.914e-05  -0.371
## Awakenings:Alcohol.consumption                 3.056e-03  1.598e-03   1.912
## Awakenings:factor(Smoking.status)Yes           1.289e-02  5.261e-03   2.450
## Awakenings:Exercise.frequency                 -2.659e-03  1.659e-03  -1.602
## Caffeine.consumption:Alcohol.consumption      -9.645e-05  9.283e-05  -1.039
## Caffeine.consumption:factor(Smoking.status)Yes  9.918e-05  2.763e-04   0.359
## Caffeine.consumption:Exercise.frequency        5.851e-05  1.022e-04   0.572
## Alcohol.consumption:factor(Smoking.status)Yes  3.358e-03  4.516e-03   0.744
## Alcohol.consumption:Exercise.frequency         4.790e-03  1.616e-03   2.964
## factor(Smoking.status)Yes:Exercise.frequency  -3.537e-03  5.387e-03  -0.656
##                                               Pr(>|t|)
## (Intercept)                                   0.440297
## Age                                           0.000841 ***
## REM.sleep.percentage                          0.142233
## Deep.sleep.percentage                         3.49e-08 ***
## Awakenings                                    0.032897 *
## Caffeine.consumption                          0.095563 .
## Alcohol.consumption                           0.434966
## factor(Smoking.status)Yes                     5.13e-05 ***
```

```
## Exercise.frequency                                0.768157
## Age:REM.sleep.percentage                          0.600613
## Age:Deep.sleep.percentage                         5.14e-06 ***
## Age:Awakenings                                    0.005084 **
## Age:Caffeine.consumption                          0.334174
## Age:Alcohol.consumption                           0.400091
## Age:factor(Smoking.status)Yes                     0.007919 **
## Age:Exercise.frequency                            0.615749
## REM.sleep.percentage:Deep.sleep.percentage        0.763637
## REM.sleep.percentage:Awakenings                   0.207794
## REM.sleep.percentage:Caffeine.consumption         0.291503
## REM.sleep.percentage:Alcohol.consumption          0.074922 .
## REM.sleep.percentage:factor(Smoking.status)Yes    0.052496 .
## REM.sleep.percentage:Exercise.frequency           0.418910
## Deep.sleep.percentage:Awakenings                  7.99e-05 ***
## Deep.sleep.percentage:Caffeine.consumption        0.213545
## Deep.sleep.percentage:Alcohol.consumption         0.295233
## Deep.sleep.percentage:factor(Smoking.status)Yes 4.94e-09 ***
## Deep.sleep.percentage:Exercise.frequency          0.761859
## Awakenings:Caffeine.consumption                   0.710542
## Awakenings:Alcohol.consumption                    0.056628 .
## Awakenings:factor(Smoking.status)Yes              0.014764 *
## Awakenings:Exercise.frequency                     0.109988
## Caffeine.consumption:Alcohol.consumption          0.299553
## Caffeine.consumption:factor(Smoking.status)Yes    0.719784
## Caffeine.consumption:Exercise.frequency           0.567487
## Alcohol.consumption:factor(Smoking.status)Yes     0.457602
## Alcohol.consumption:Exercise.frequency            0.003240 **
## factor(Smoking.status)Yes:Exercise.frequency      0.511939
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05341 on 351 degrees of freedom
## Multiple R-squared:  0.8595, Adjusted R-squared:  0.8451
## F-statistic: 59.66 on 36 and 351 DF,  p-value: < 2.2e-16
```

```
eightint <- lm(Sleep.efficiency~(Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.cons
summary(eightint)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ (Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency +
##     Age:factor(Smoking.status) + REM.sleep.percentage:factor(Smoking.status) +
##     Deep.sleep.percentage:Awakenings + Deep.sleep.percentage:factor(Smoking.status) +
##     Awakenings:Alcohol.consumption + Awakenings:factor(Smoking.status) +
##     Alcohol.consumption:Exercise.frequency), data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.161682 -0.030809  0.004468  0.035789  0.146607
##
## Coefficients:
```

```
##                                                Estimate Std. Error t value
## (Intercept)                                   0.3401953  0.0410341   8.291
## Age                                           0.0012297  0.0002629   4.677
## REM.sleep.percentage                          0.0061863  0.0010570   5.853
## Deep.sleep.percentage                         0.0061274  0.0004780  12.820
## Awakenings                                    0.0029936  0.0103371   0.290
## Caffeine.consumption                          0.0002059  0.0001016   2.026
## Alcohol.consumption                          -0.0195224  0.0046602  -4.189
## factor(Smoking.status)Yes                    -0.1993890  0.0555239  -3.591
## Exercise.frequency                            0.0044177  0.0025260   1.749
## Age:factor(Smoking.status)Yes                -0.0010150  0.0004639  -2.188
## REM.sleep.percentage:factor(Smoking.status)Yes  0.0023610  0.0017795   1.327
## Deep.sleep.percentage:Awakenings             -0.0007349  0.0001674  -4.389
## Deep.sleep.percentage:factor(Smoking.status)Yes  0.0024569  0.0004265   5.761
## Awakenings:Alcohol.consumption                0.0034796  0.0014931   2.330
## Awakenings:factor(Smoking.status)Yes          0.0090595  0.0050798   1.783
## Alcohol.consumption:Exercise.frequency        0.0040663  0.0013496   3.013
##                                                Pr(>|t|)
## (Intercept)                                   2.08e-15 ***
## Age                                           4.08e-06 ***
## REM.sleep.percentage                          1.06e-08 ***
## Deep.sleep.percentage                          < 2e-16 ***
## Awakenings                                    0.772287
## Caffeine.consumption                          0.043473 *
## Alcohol.consumption                           3.50e-05 ***
## factor(Smoking.status)Yes                     0.000374 ***
## Exercise.frequency                            0.081128 .
## Age:factor(Smoking.status)Yes                 0.029281 *
## REM.sleep.percentage:factor(Smoking.status)Yes  0.185383
## Deep.sleep.percentage:Awakenings              1.48e-05 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 1.76e-08 ***
## Awakenings:Alcohol.consumption                0.020318 *
## Awakenings:factor(Smoking.status)Yes          0.075334 .
## Alcohol.consumption:Exercise.frequency        0.002763 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05507 on 372 degrees of freedom
## Multiple R-squared:  0.8417, Adjusted R-squared:  0.8353
## F-statistic: 131.9 on 15 and 372 DF,  p-value: < 2.2e-16
```

```
eightint <- lm(Sleep.efficiency~(Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.cons
summary(eightint)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ (Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency +
##     Age:factor(Smoking.status) + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Alcohol.consumption:Exercise.frequency), data = df)
##
## Residuals:
```

```
##       Min        1Q    Median        3Q       Max
## -0.162598 -0.033368  0.002092  0.036578  0.150256
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                     0.3047812  0.0374978   8.128
## Age                                             0.0012161  0.0002637   4.612
## REM.sleep.percentage                            0.0070711  0.0008619   8.204
## Deep.sleep.percentage                           0.0063402  0.0004637  13.673
## Awakenings                                      0.0082698  0.0097640   0.847
## Caffeine.consumption                            0.0002131  0.0001019   2.091
## Alcohol.consumption                            -0.0196203  0.0046754  -4.196
## factor(Smoking.status)Yes                      -0.1116384  0.0268750  -4.154
## Exercise.frequency                              0.0044790  0.0025335   1.768
## Age:factor(Smoking.status)Yes                  -0.0009879  0.0004642  -2.128
## Deep.sleep.percentage:Awakenings               -0.0007840  0.0001634  -4.799
## Deep.sleep.percentage:factor(Smoking.status)Yes  0.0020460  0.0003813   5.365
## Awakenings:Alcohol.consumption                  0.0034743  0.0014981   2.319
## Alcohol.consumption:Exercise.frequency          0.0039028  0.0013513   2.888
##                                                 Pr(>|t|)
## (Intercept)                                     6.46e-15 ***
## Age                                             5.50e-06 ***
## REM.sleep.percentage                            3.79e-15 ***
## Deep.sleep.percentage                            < 2e-16 ***
## Awakenings                                        0.3976
## Caffeine.consumption                              0.0372 *
## Alcohol.consumption                             3.39e-05 ***
## factor(Smoking.status)Yes                       4.05e-05 ***
## Exercise.frequency                                0.0779 .
## Age:factor(Smoking.status)Yes                     0.0340 *
## Deep.sleep.percentage:Awakenings                2.31e-06 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 1.42e-07 ***
## Awakenings:Alcohol.consumption                    0.0209 *
## Alcohol.consumption:Exercise.frequency            0.0041 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05527 on 374 degrees of freedom
## Multiple R-squared:  0.8397, Adjusted R-squared:  0.8341
## F-statistic: 150.7 on 13 and 374 DF,  p-value: < 2.2e-16
```

```
eightpow <- lm(Sleep.efficiency~Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.consu
summary(eightpow)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency +
##     Age:factor(Smoking.status) + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Alcohol.consumption:Exercise.frequency + I(Age^2), data = df)
##
## Residuals:
```

```
##       Min        1Q    Median        3Q       Max
## -0.171403 -0.033387  0.001426  0.036831  0.152237
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                     2.595e-01  4.023e-02   6.450
## Age                                             5.212e-03  1.392e-03   3.744
## REM.sleep.percentage                            6.781e-03  8.591e-04   7.893
## Deep.sleep.percentage                           6.067e-03  4.685e-04  12.950
## Awakenings                                      8.016e-03  9.667e-03   0.829
## Caffeine.consumption                            1.920e-04  1.012e-04   1.899
## Alcohol.consumption                            -2.152e-02  4.674e-03  -4.604
## factor(Smoking.status)Yes                      -1.327e-01  2.757e-02  -4.814
## Exercise.frequency                              2.696e-03  2.581e-03   1.045
## I(Age^2)                                       -4.978e-05  1.703e-05  -2.923
## Age:factor(Smoking.status)Yes                  -7.312e-04  4.679e-04  -1.563
## Deep.sleep.percentage:Awakenings               -7.797e-04  1.618e-04  -4.820
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.218e-03  3.821e-04   5.804
## Awakenings:Alcohol.consumption                  3.525e-03  1.483e-03   2.377
## Alcohol.consumption:Exercise.frequency          4.688e-03  1.365e-03   3.436
##                                                Pr(>|t|)
## (Intercept)                                    3.48e-10 ***
## Age                                            0.000209 ***
## REM.sleep.percentage                           3.32e-14 ***
## Deep.sleep.percentage                           < 2e-16 ***
## Awakenings                                     0.407555
## Caffeine.consumption                           0.058401 .
## Alcohol.consumption                            5.70e-06 ***
## factor(Smoking.status)Yes                      2.15e-06 ***
## Exercise.frequency                             0.296915
## I(Age^2)                                       0.003684 **
## Age:factor(Smoking.status)Yes                  0.118943
## Deep.sleep.percentage:Awakenings               2.09e-06 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 1.38e-08 ***
## Awakenings:Alcohol.consumption                 0.017982 *
## Alcohol.consumption:Exercise.frequency         0.000658 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05472 on 373 degrees of freedom
## Multiple R-squared:  0.8433, Adjusted R-squared:  0.8374
## F-statistic: 143.4 on 14 and 373 DF,  p-value: < 2.2e-16
```

```
eightpow <- lm(Sleep.efficiency~Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.consu
summary(eightpow)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency +
##     Age:factor(Smoking.status) + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Alcohol.consumption:Exercise.frequency + I(Age^2) + I(Age^3),
```

```
##     data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.170523 -0.034207  0.002006  0.035337  0.154360
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                     2.394e-01  5.791e-02   4.134
## Age                                             7.426e-03  4.798e-03   1.548
## REM.sleep.percentage                            6.702e-03  8.753e-04   7.657
## Deep.sleep.percentage                           6.017e-03  4.806e-04  12.518
## Awakenings                                      7.615e-03  9.713e-03   0.784
## Caffeine.consumption                            1.823e-04  1.032e-04   1.765
## Alcohol.consumption                            -2.158e-02  4.681e-03  -4.610
## factor(Smoking.status)Yes                      -1.357e-01  2.826e-02  -4.801
## Exercise.frequency                              2.660e-03  2.585e-03   1.029
## I(Age^2)                                       -1.083e-04  1.225e-04  -0.884
## I(Age^3)                                        4.758e-07  9.865e-07   0.482
## Age:factor(Smoking.status)Yes                  -7.138e-04  4.698e-04  -1.519
## Deep.sleep.percentage:Awakenings               -7.725e-04  1.626e-04  -4.751
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.251e-03  3.887e-04   5.792
## Awakenings:Alcohol.consumption                  3.501e-03  1.486e-03   2.356
## Alcohol.consumption:Exercise.frequency          4.686e-03  1.366e-03   3.430
##                                                Pr(>|t|)
## (Intercept)                                    4.41e-05 ***
## Age                                            0.122524
## REM.sleep.percentage                           1.66e-13 ***
## Deep.sleep.percentage                           < 2e-16 ***
## Awakenings                                     0.433555
## Caffeine.consumption                           0.078303 .
## Alcohol.consumption                            5.54e-06 ***
## factor(Smoking.status)Yes                      2.29e-06 ***
## Exercise.frequency                             0.304204
## I(Age^2)                                       0.377199
## I(Age^3)                                       0.629874
## Age:factor(Smoking.status)Yes                  0.129521
## Deep.sleep.percentage:Awakenings               2.90e-06 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 1.48e-08 ***
## Awakenings:Alcohol.consumption                 0.018986 *
## Alcohol.consumption:Exercise.frequency         0.000671 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05477 on 372 degrees of freedom
## Multiple R-squared:  0.8434, Adjusted R-squared:  0.8371
## F-statistic: 133.6 on 15 and 372 DF,  p-value: < 2.2e-16
```

In the end, we will only choose the Age^2 higher order term for this model as it appears possibly non-linear in the scatterplot and is significant in the model. Trying Age^3 in our model did not work

We will conduct tests on our final 8-variable model

```
finalmodel_8 <- lm(Sleep.efficiency~Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.
summary(finalmodel_8)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency +
##     Age:factor(Smoking.status) + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Alcohol.consumption:Exercise.frequency + I(Age^2), data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.171403 -0.033387  0.001426  0.036831  0.152237
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                     2.595e-01  4.023e-02   6.450
## Age                                             5.212e-03  1.392e-03   3.744
## REM.sleep.percentage                            6.781e-03  8.591e-04   7.893
## Deep.sleep.percentage                           6.067e-03  4.685e-04  12.950
## Awakenings                                      8.016e-03  9.667e-03   0.829
## Caffeine.consumption                            1.920e-04  1.012e-04   1.899
## Alcohol.consumption                            -2.152e-02  4.674e-03  -4.604
## factor(Smoking.status)Yes                      -1.327e-01  2.757e-02  -4.814
## Exercise.frequency                              2.696e-03  2.581e-03   1.045
## I(Age^2)                                       -4.978e-05  1.703e-05  -2.923
## Age:factor(Smoking.status)Yes                  -7.312e-04  4.679e-04  -1.563
## Deep.sleep.percentage:Awakenings               -7.797e-04  1.618e-04  -4.820
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.218e-03  3.821e-04   5.804
## Awakenings:Alcohol.consumption                  3.525e-03  1.483e-03   2.377
## Alcohol.consumption:Exercise.frequency          4.688e-03  1.365e-03   3.436
##                                                Pr(>|t|)
## (Intercept)                                    3.48e-10 ***
## Age                                            0.000209 ***
## REM.sleep.percentage                           3.32e-14 ***
## Deep.sleep.percentage                           < 2e-16 ***
## Awakenings                                     0.407555
## Caffeine.consumption                           0.058401 .
## Alcohol.consumption                            5.70e-06 ***
## factor(Smoking.status)Yes                      2.15e-06 ***
## Exercise.frequency                             0.296915
## I(Age^2)                                       0.003684 **
## Age:factor(Smoking.status)Yes                  0.118943
## Deep.sleep.percentage:Awakenings               2.09e-06 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 1.38e-08 ***
```

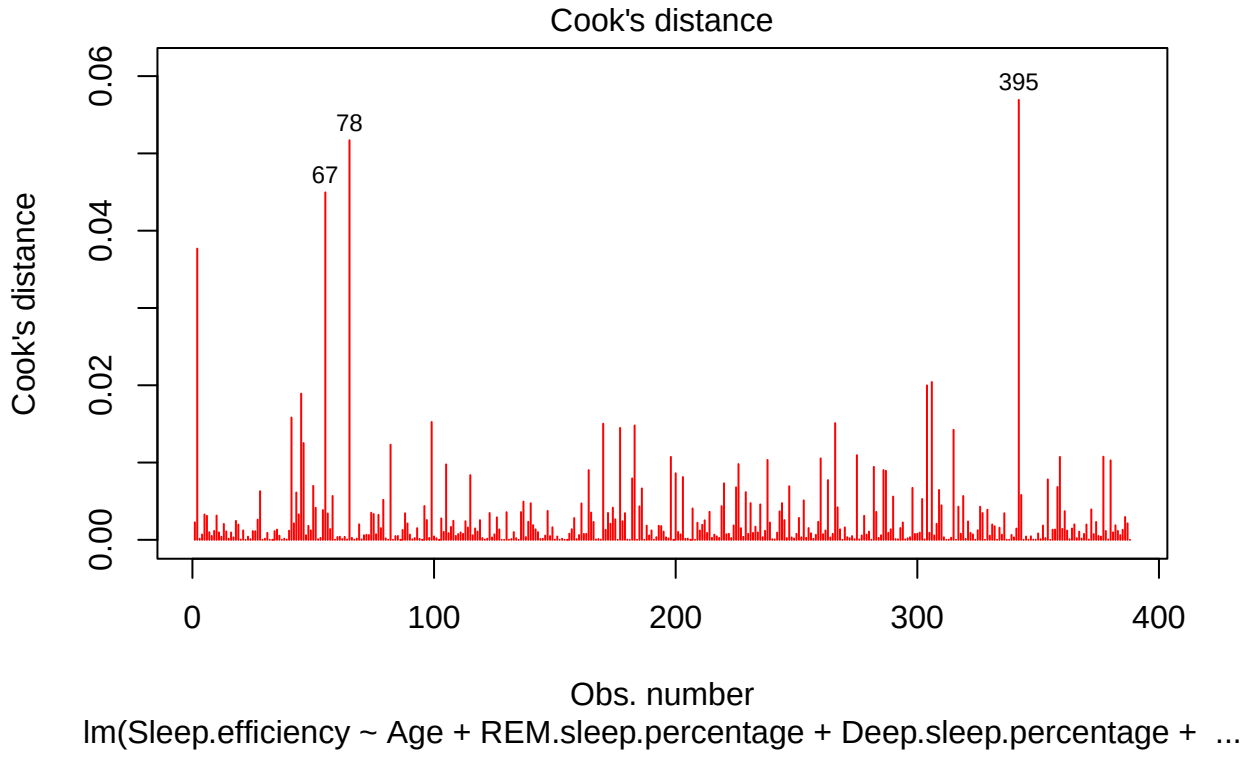
```
## Awakenings:Alcohol.consumption                  0.017982 *
## Alcohol.consumption:Exercise.frequency          0.000658 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05472 on 373 degrees of freedom
## Multiple R-squared:  0.8433, Adjusted R-squared:  0.8374
## F-statistic: 143.4 on 14 and 373 DF,  p-value: < 2.2e-16
```

```
df[cooks.distance(finalmodel_8)>1,]
```

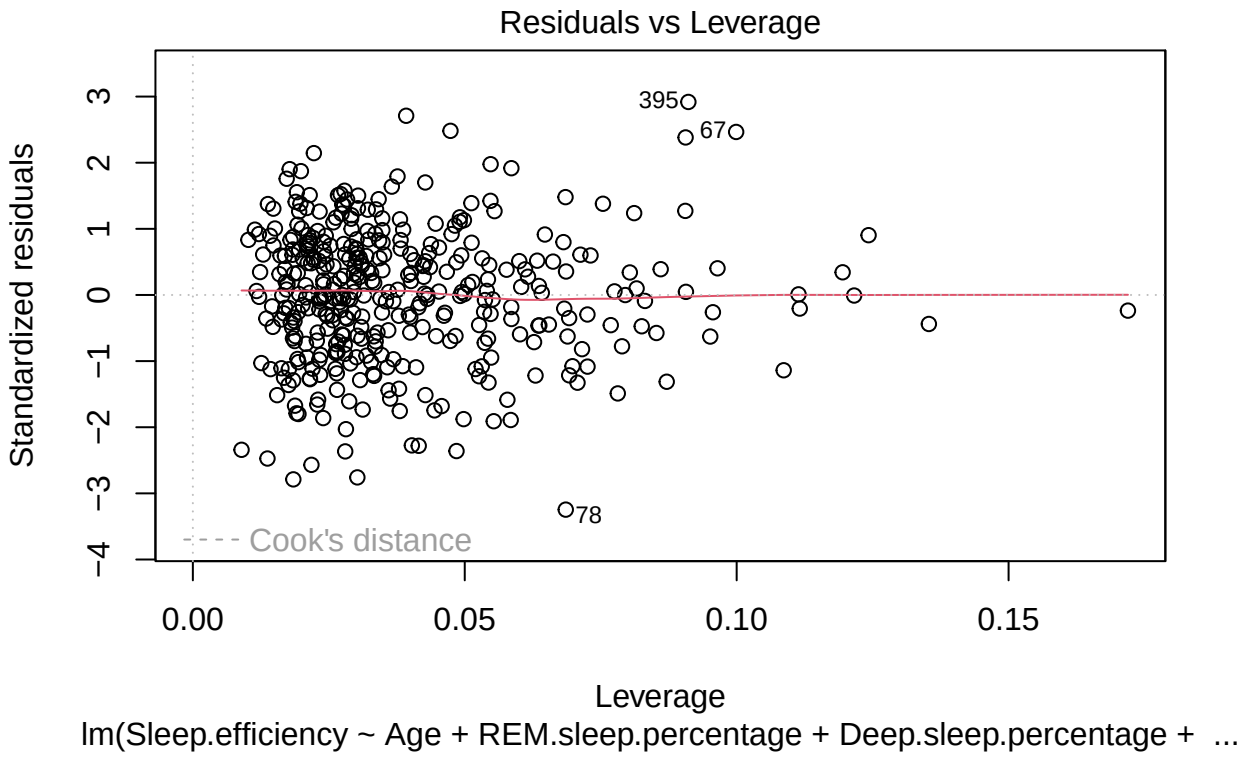
```
##  [1] ID                     Age                    Gender
##  [4] Bedtime                Wakeup.time            Sleep.duration
##  [7] Sleep.efficiency       REM.sleep.percentage   Deep.sleep.percentage
## [10] Light.sleep.percentage Awakenings             Caffeine.consumption
## [13] Alcohol.consumption    Smoking.status         Exercise.frequency
## <0 rows> (or 0-length row.names)
```

No points outside a cook's distance of 1 found.

```
plot(finalmodel_8,pch=18,col="red",which=c(4))
```




```
plot(finalmodel_8,which=5)
```




```
lev=hatvalues(finalmodel_8)
p = length(coef(finalmodel_8))
n = nrow(df)
outlier3p = lev[lev>(3*p/n)]
print(outlier3p)
```

```
##        82       258       303       379       425
## 0.1353442 0.1719497 0.1242562 0.1216035 0.1194863
```

```
outi <- c(82,258,303,379,425)
df_new = df[-outi,]
```

```
finalmodel_8 <- lm(Sleep.efficiency~Age+REM.sleep.percentage+Deep.sleep.percentage+Awakenings+Caffeine.
summary(finalmodel_8)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + REM.sleep.percentage +
##     Deep.sleep.percentage + Awakenings + Caffeine.consumption +
##     Alcohol.consumption + factor(Smoking.status) + Exercise.frequency +
##     Age:factor(Smoking.status) + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Alcohol.consumption:Exercise.frequency + I(Age^2), data = df_new)
##
```

```
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.173435 -0.034358  0.000541  0.037103  0.149176
##
## Coefficients:
##                                                    Estimate Std. Error t value
## (Intercept)                                       2.597e-01  4.040e-02   6.427
## Age                                               5.074e-03  1.395e-03   3.638
## REM.sleep.percentage                              6.782e-03  8.650e-04   7.840
## Deep.sleep.percentage                             6.104e-03  4.697e-04  12.995
## Awakenings                                        8.589e-03  9.847e-03   0.872
## Caffeine.consumption                              1.987e-04  1.013e-04   1.961
## Alcohol.consumption                              -2.071e-02  4.702e-03  -4.404
## factor(Smoking.status)Yes                        -1.379e-01  2.788e-02  -4.944
## Exercise.frequency                                2.640e-03  2.591e-03   1.019
## I(Age^2)                                         -4.818e-05  1.708e-05  -2.822
## Age:factor(Smoking.status)Yes                    -6.102e-04  4.730e-04  -1.290
## Deep.sleep.percentage:Awakenings                 -7.889e-04  1.644e-04  -4.799
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.233e-03  3.853e-04   5.797
## Awakenings:Alcohol.consumption                    3.371e-03  1.491e-03   2.261
## Alcohol.consumption:Exercise.frequency            4.532e-03  1.368e-03   3.312
##                                                  Pr(>|t|)
## (Intercept)                                      4.02e-10 ***
## Age                                              0.000314 ***
## REM.sleep.percentage                             4.87e-14 ***
## Deep.sleep.percentage                             < 2e-16 ***
## Awakenings                                       0.383667
## Caffeine.consumption                             0.050680 .
## Alcohol.consumption                              1.40e-05 ***
## factor(Smoking.status)Yes                        1.16e-06 ***
## Exercise.frequency                               0.309006
## I(Age^2)                                         0.005037 **
## Age:factor(Smoking.status)Yes                    0.197782
## Deep.sleep.percentage:Awakenings                 2.32e-06 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 1.45e-08 ***
## Awakenings:Alcohol.consumption                   0.024350 *
## Alcohol.consumption:Exercise.frequency           0.001016 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05473 on 369 degrees of freedom
## Multiple R-squared:  0.844,  Adjusted R-squared:  0.8381
## F-statistic: 142.6 on 14 and 369 DF,  p-value: < 2.2e-16
```

```
shapiro.test(residuals(finalmodel_8))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(finalmodel_8)
## W = 0.99472, p-value = 0.2121
```

We do not reject the null hypothesis that the data is normally distributed. The normality condition is satisfied.

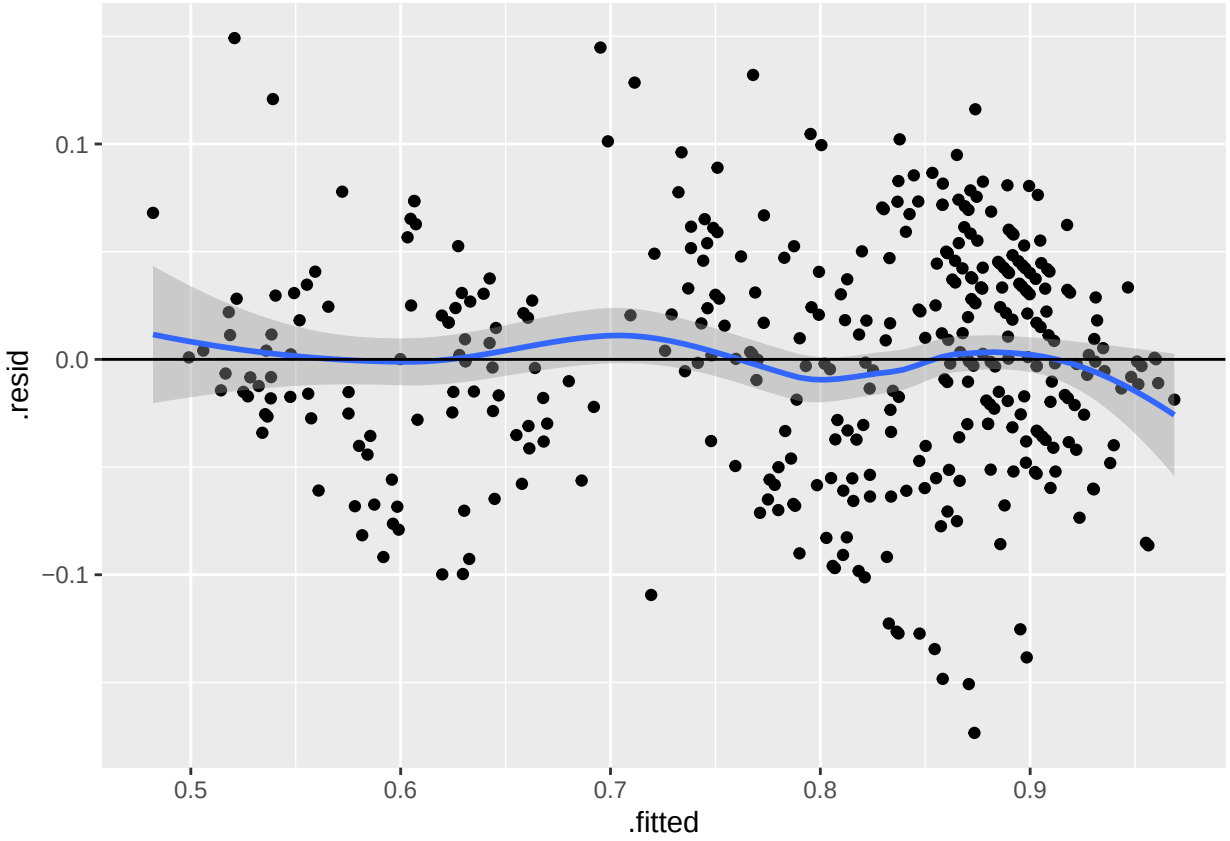
```
bptest(finalmodel_8)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  finalmodel_8
## BP = 29.855, df = 14, p-value = 0.00799
```

Reject the null hypothesis that the data is homoskedastistic. The homoskedasticity condition is not satisfied.

```
ggplot(finalmodel_8, aes(x=.fitted, y=.resid)) +
geom_point() + geom_smooth()+
geom_hline(yintercept = 0)
```

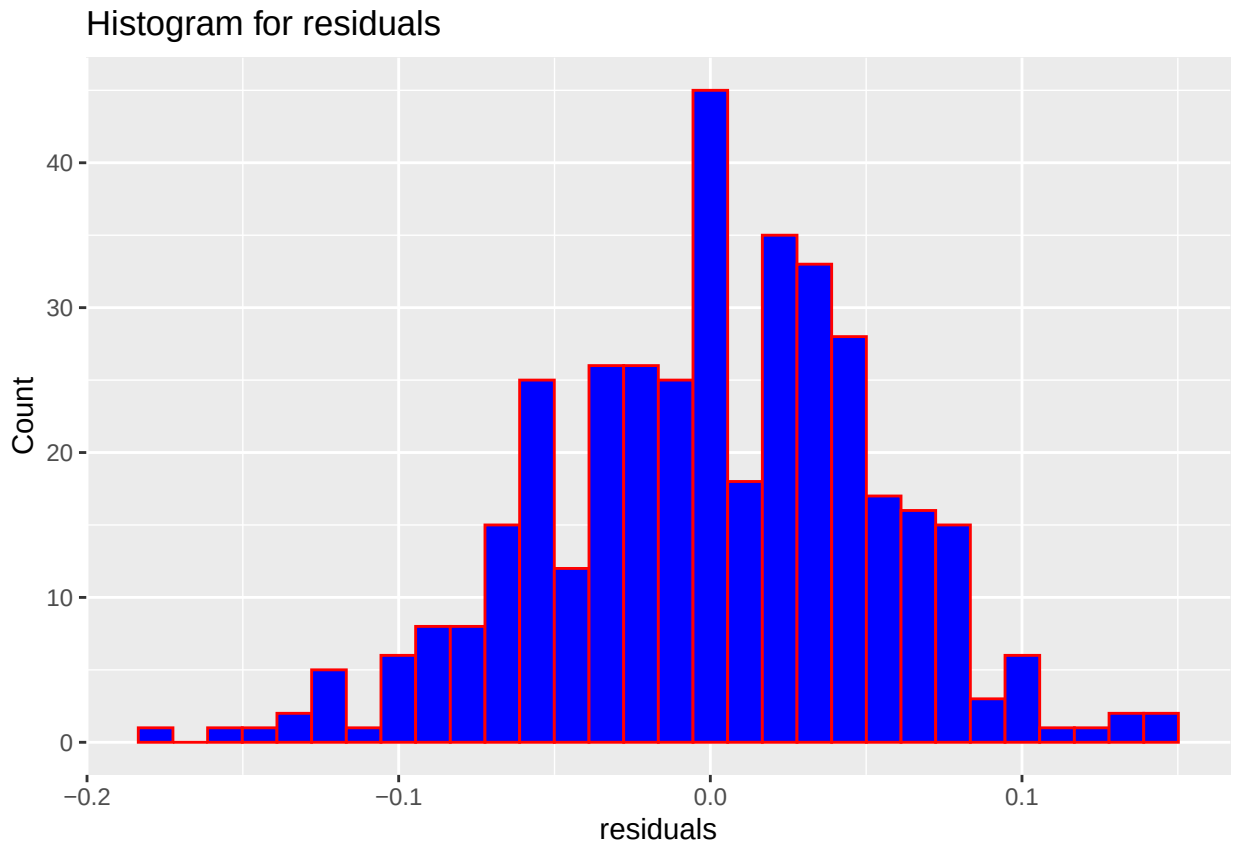
```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



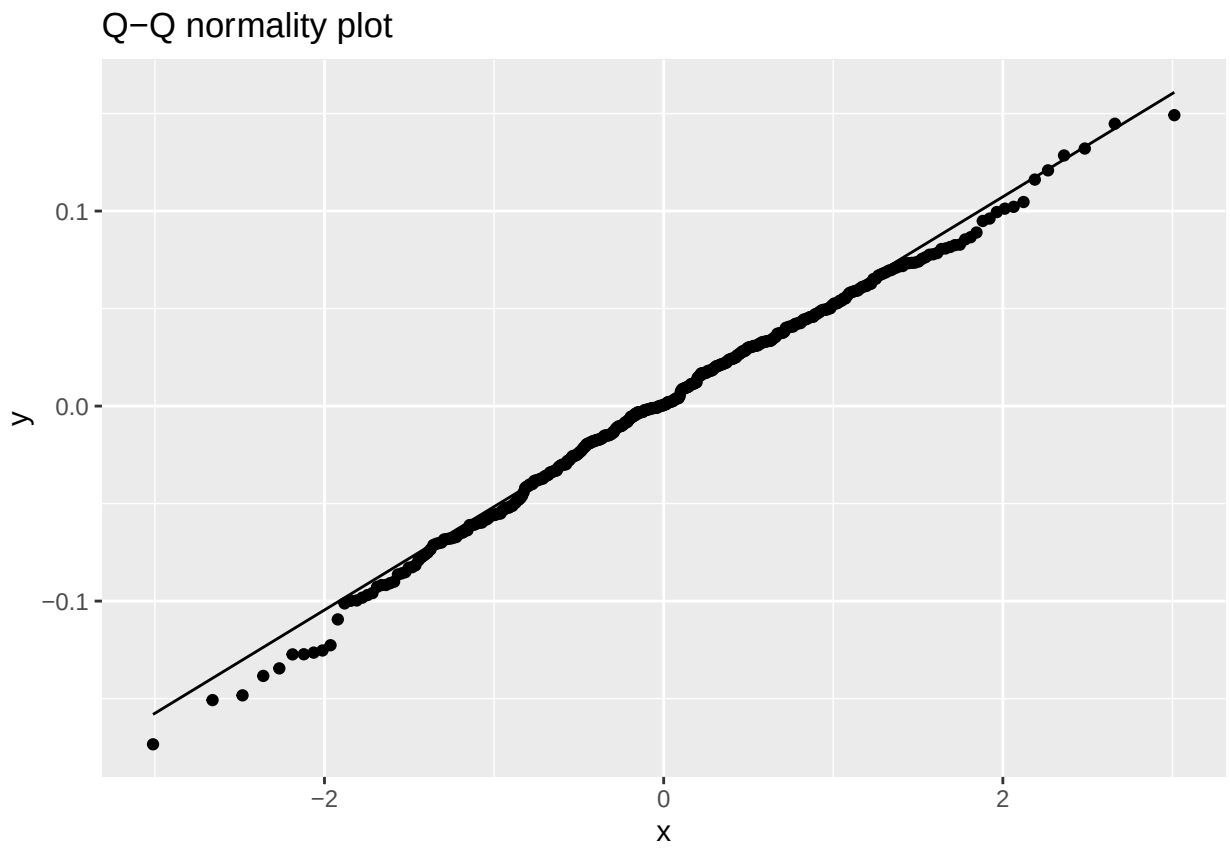
There does not appear to be a discernible pattern in the residual plot.

```
ggplot(data=df_new, aes(residuals(finalmodel_8))) +
geom_histogram(color='red',fill='blue') +
labs(title="Histogram for residuals") +
labs(x="residuals", y="Count")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

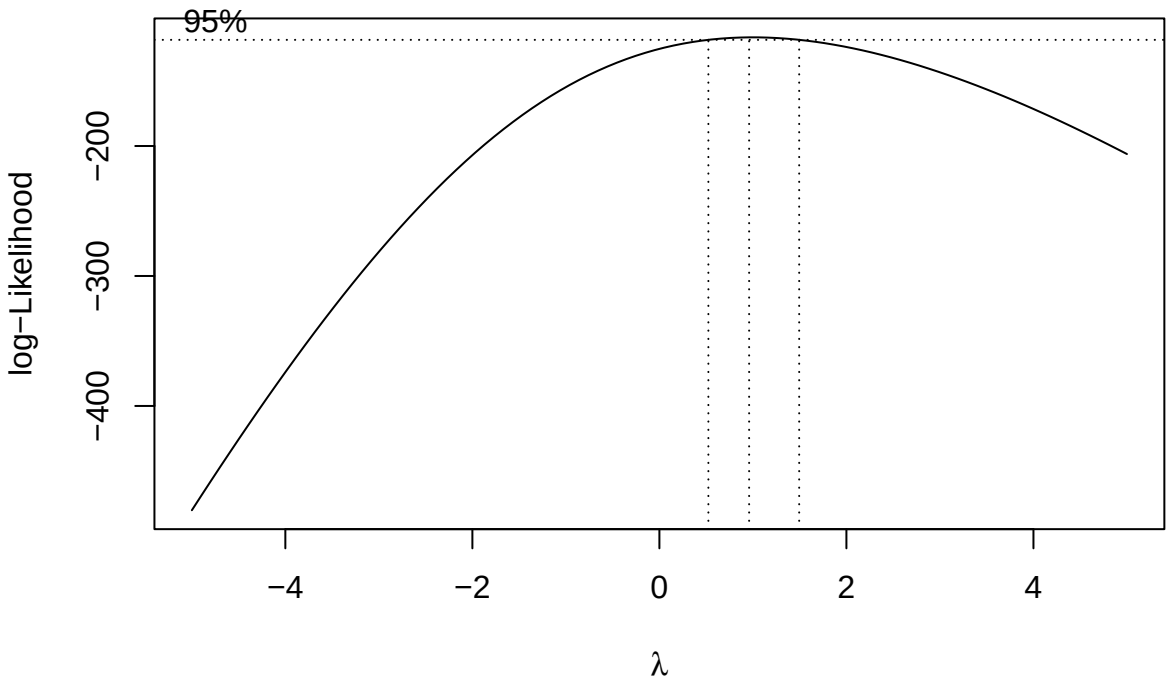



```
ggplot(df_new, aes(sample=finalmodel_8$residuals)) +
stat_qq() +
stat_qq_line() +
ggtitle("Q-Q normality plot")
```



To adjust for the heteroskedasticity and normality, we will perform a Box-Cox transformation on the 8-variable model

```
bc = boxcox(finalmodel_8, lambda=seq(-5,5))
```




```
bestlambda=bc$x[which(bc$y==max(bc$y))]
bestlambda
```

```
## [1] 0.959596
```

```
bcmodel=lm((((Sleep.efficiency^bestlambda)-1)/bestlambda)~Age+REM.sleep.percentage+Deep.sleep.percentage
summary(bcmodel)
```

```
##
## Call:
## lm(formula = (((Sleep.efficiency^bestlambda) - 1)/bestlambda) ~
##     Age + REM.sleep.percentage + Deep.sleep.percentage + Awakenings +
##         Caffeine.consumption + Alcohol.consumption + factor(Smoking.status) +
##         Exercise.frequency + Age:factor(Smoking.status) + Deep.sleep.percentage:Awakenings +
##         Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##         Alcohol.consumption:Exercise.frequency + I(Age^2), data = df_new)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.175378 -0.035472  0.000377  0.037411  0.151720
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                    -7.479e-01  4.082e-02 -18.323
## Age                                             5.138e-03  1.409e-03   3.646
```

```
## REM.sleep.percentage                                6.866e-03  8.739e-04   7.856
## Deep.sleep.percentage                               6.167e-03  4.745e-04  12.997
## Awakenings                                          8.554e-03  9.949e-03   0.860
## Caffeine.consumption                                2.013e-04  1.024e-04   1.966
## Alcohol.consumption                                -2.090e-02  4.750e-03  -4.400
## factor(Smoking.status)Yes                          -1.413e-01  2.817e-02  -5.016
## Exercise.frequency                                  2.656e-03  2.618e-03   1.015
## I(Age^2)                                           -4.879e-05  1.725e-05  -2.828
## Age:factor(Smoking.status)Yes                      -6.094e-04  4.778e-04  -1.275
## Deep.sleep.percentage:Awakenings                   -7.928e-04  1.661e-04  -4.774
## Deep.sleep.percentage:factor(Smoking.status)Yes     2.285e-03  3.892e-04   5.871
## Awakenings:Alcohol.consumption                      3.399e-03  1.506e-03   2.257
## Alcohol.consumption:Exercise.frequency              4.581e-03  1.382e-03   3.314
##                                                    Pr(>|t|)
## (Intercept)                                         < 2e-16 ***
## Age                                                0.000305 ***
## REM.sleep.percentage                               4.37e-14 ***
## Deep.sleep.percentage                               < 2e-16 ***
## Awakenings                                         0.390456
## Caffeine.consumption                               0.050075 .
## Alcohol.consumption                                1.42e-05 ***
## factor(Smoking.status)Yes                          8.22e-07 ***
## Exercise.frequency                                 0.310981
## I(Age^2)                                           0.004936 **
## Age:factor(Smoking.status)Yes                      0.202974
## Deep.sleep.percentage:Awakenings                   2.61e-06 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes     9.68e-09 ***
## Awakenings:Alcohol.consumption                     0.024598 *
## Alcohol.consumption:Exercise.frequency             0.001011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05529 on 369 degrees of freedom
## Multiple R-squared:  0.8446, Adjusted R-squared:  0.8387
## F-statistic: 143.3 on 14 and 369 DF,  p-value: < 2.2e-16
```

```
shapiro.test(residuals(bcmodel))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(bcmodel)
## W = 0.99471, p-value = 0.2107
```

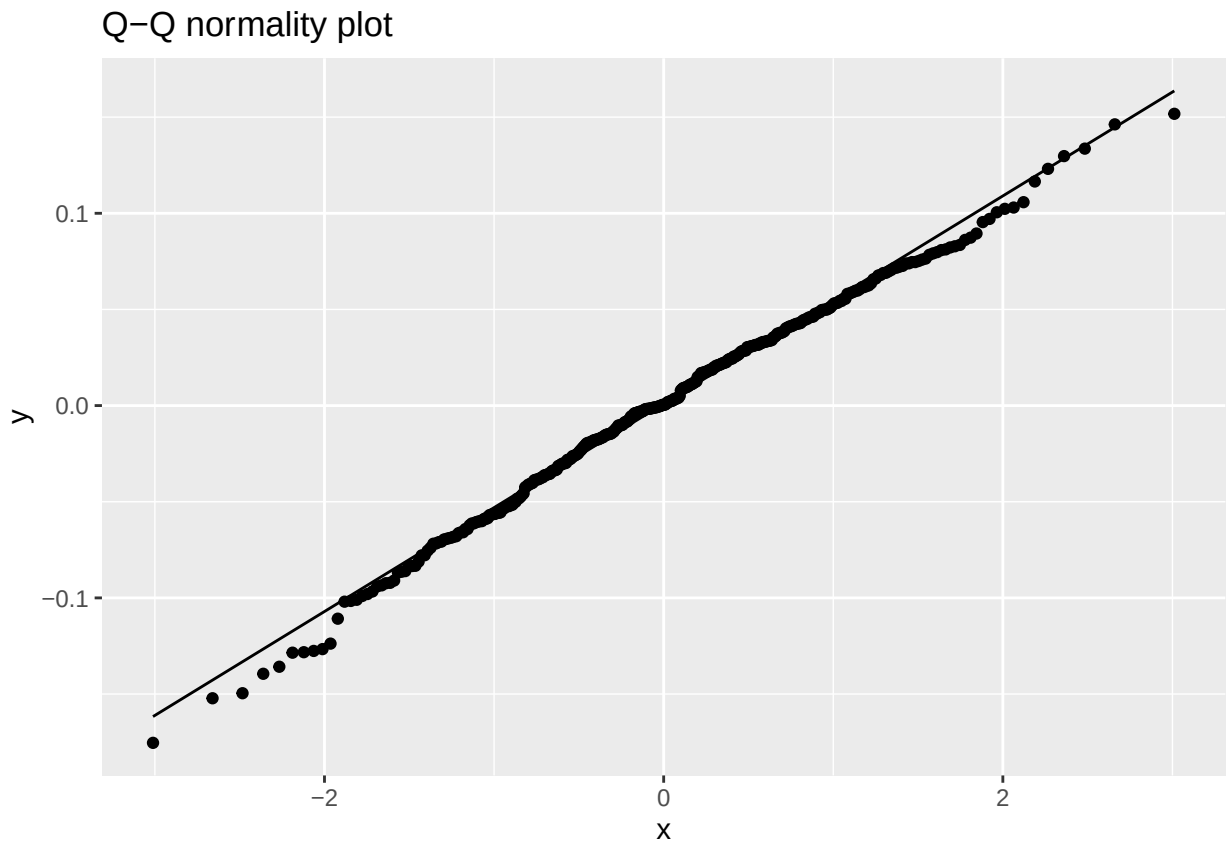
We still cannot reject the null hypothesis and the normality condition is satisfied # ADD HYPOTHESIS

```
bptest(bcmodel)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  bcmodel
## BP = 29.946, df = 14, p-value = 0.007764
```

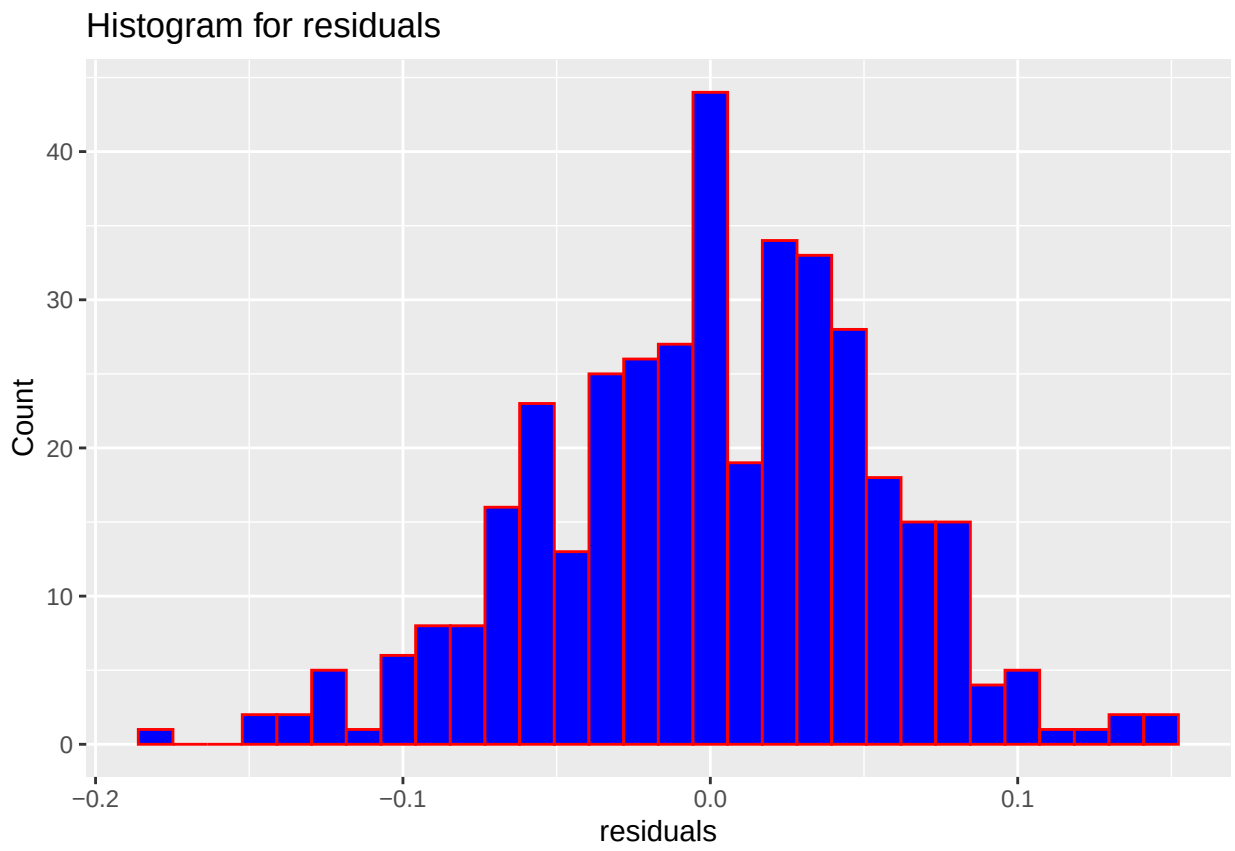
We still reject the null hypothesis and the homoskedasticity condition is still not satisfied.

```
ggplot(df_new, aes(sample=bcmodel$residuals)) +
stat_qq() +
stat_qq_line() +
ggtitle("Q-Q normality plot")
```



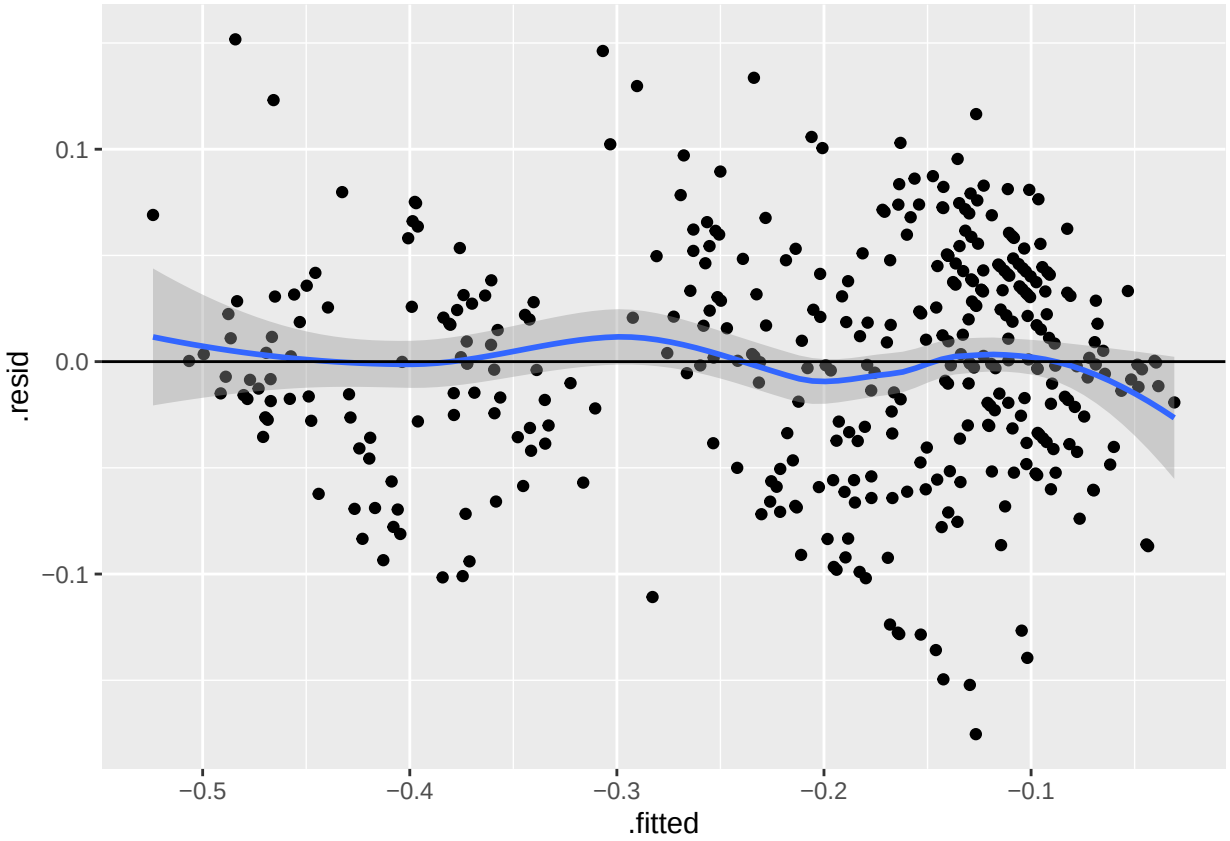

```
ggplot(data=df_new, aes(residuals(bcmodel))) +
geom_histogram(color='red',fill='blue') +
labs(title="Histogram for residuals") +
labs(x="residuals", y="Count")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
ggplot(bcmodel, aes(x=.fitted, y=.resid)) +
geom_point() + geom_smooth()+
geom_hline(yintercept = 0)
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



Evaluating the 11-variable model

```
finalelevenmod <- lm(Sleep.efficiency~Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage+ D

summary(finalelevenmod)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency + Age * factor(Smoking.status) +
##     Age * Deep.sleep.percentage + Age * Awakenings + Age:factor(Smoking.status) +
##     factor(Gender):Alcohol.consumption + Sleep.duration:Alcohol.consumption +
##     REM.sleep.percentage:Alcohol.consumption + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Awakenings:factor(Smoking.status) + Awakenings:Exercise.frequency +
##     Alcohol.consumption:Exercise.frequency, data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.158714 -0.030835  0.004594  0.036660  0.132823
##
```

```
## Coefficients:
##                                                   Estimate Std. Error t value
## (Intercept)                                       7.057e-02  5.912e-02   1.194
## Age                                               5.554e-03  8.772e-04   6.332
## factor(Gender)Male                                2.029e-02  7.660e-03   2.649
## Bedtime                                          -1.463e-04  5.618e-04  -0.260
## Sleep.duration                                    5.764e-03  3.792e-03   1.520
## REM.sleep.percentage                              8.768e-03  1.046e-03   8.384
## Deep.sleep.percentage                             8.702e-03  7.065e-04  12.318
## Awakenings                                        2.063e-02  1.137e-02   1.815
## Caffeine.consumption                              1.000e-04  1.028e-04   0.973
## Alcohol.consumption                               3.431e-02  1.980e-02   1.733
## factor(Smoking.status)Yes                        -1.398e-01  3.011e-02  -4.643
## Exercise.frequency                                6.123e-03  3.221e-03   1.901
## Age:factor(Smoking.status)Yes                    -1.479e-03  4.559e-04  -3.245
## Age:Deep.sleep.percentage                        -6.903e-05  1.365e-05  -5.057
## Age:Awakenings                                   -5.248e-04  1.628e-04  -3.224
## factor(Gender)Male:Alcohol.consumption           -1.018e-02  3.626e-03  -2.808
## Sleep.duration:Alcohol.consumption               -4.084e-03  1.999e-03  -2.043
## REM.sleep.percentage:Alcohol.consumption         -1.040e-03  5.029e-04  -2.067
## Deep.sleep.percentage:Awakenings                 -6.121e-04  1.620e-04  -3.778
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.580e-03  4.041e-04   6.385
## Awakenings:Alcohol.consumption                    4.096e-03  1.443e-03   2.839
## Awakenings:factor(Smoking.status)Yes              1.294e-02  4.934e-03   2.622
## Awakenings:Exercise.frequency                    -3.031e-03  1.531e-03  -1.979
## Alcohol.consumption:Exercise.frequency            5.481e-03  1.338e-03   4.095
##                                                  Pr(>|t|)
## (Intercept)                                      0.233442
## Age                                              7.15e-10 ***
## factor(Gender)Male                               0.008429 **
## Bedtime                                          0.794771
## Sleep.duration                                   0.129338
## REM.sleep.percentage                             1.14e-15 ***
## Deep.sleep.percentage                             < 2e-16 ***
## Awakenings                                       0.070317 .
## Caffeine.consumption                             0.331172
## Alcohol.consumption                              0.083953 .
## factor(Smoking.status)Yes                        4.80e-06 ***
## Exercise.frequency                               0.058107 .
## Age:factor(Smoking.status)Yes                    0.001284 **
## Age:Deep.sleep.percentage                        6.77e-07 ***
## Age:Awakenings                                   0.001378 **
## factor(Gender)Male:Alcohol.consumption           0.005249 **
## Sleep.duration:Alcohol.consumption               0.041779 *
## REM.sleep.percentage:Alcohol.consumption         0.039418 *
## Deep.sleep.percentage:Awakenings                 0.000185 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 5.22e-10 ***
## Awakenings:Alcohol.consumption                   0.004780 **
## Awakenings:factor(Smoking.status)Yes             0.009096 **
## Awakenings:Exercise.frequency                    0.048548 *
## Alcohol.consumption:Exercise.frequency           5.20e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

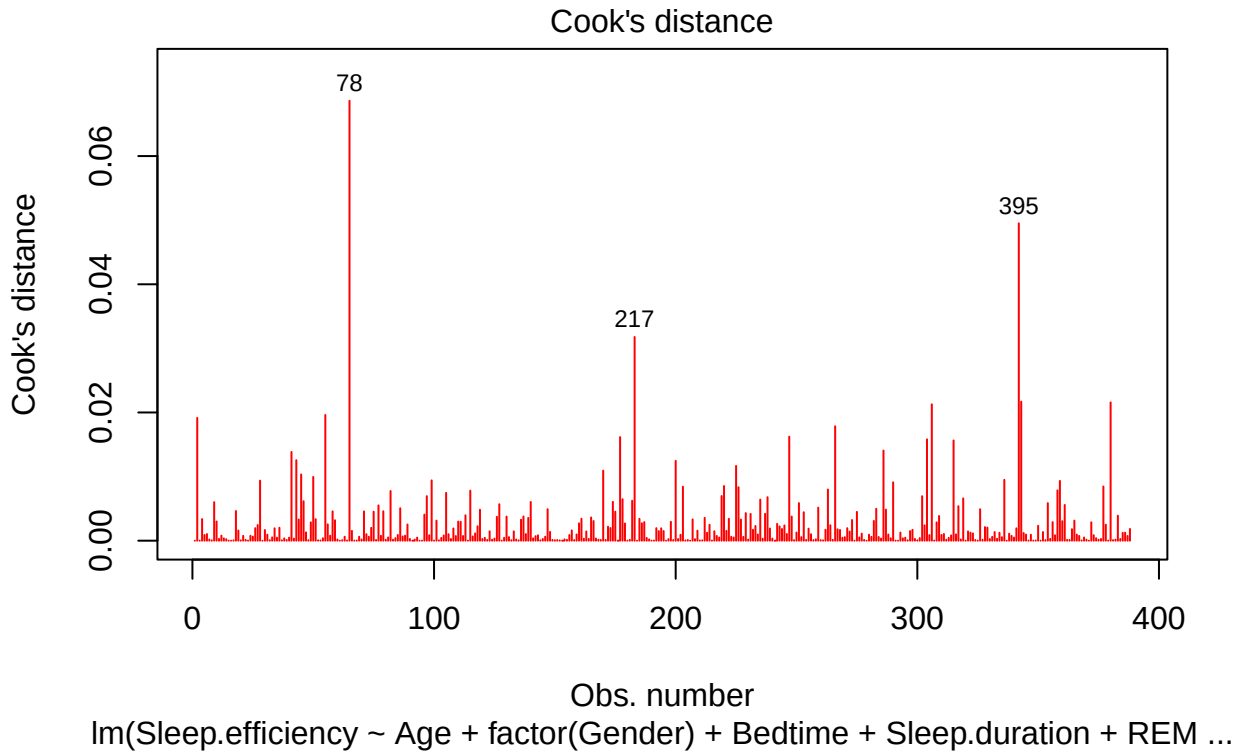
```
## Residual standard error: 0.0525 on 364 degrees of freedom
## Multiple R-squared:  0.8592, Adjusted R-squared:  0.8503
## F-statistic: 96.61 on 23 and 364 DF,  p-value: < 2.2e-16
```

```
df[cooks.distance(finalelevenmod)>1,]
```

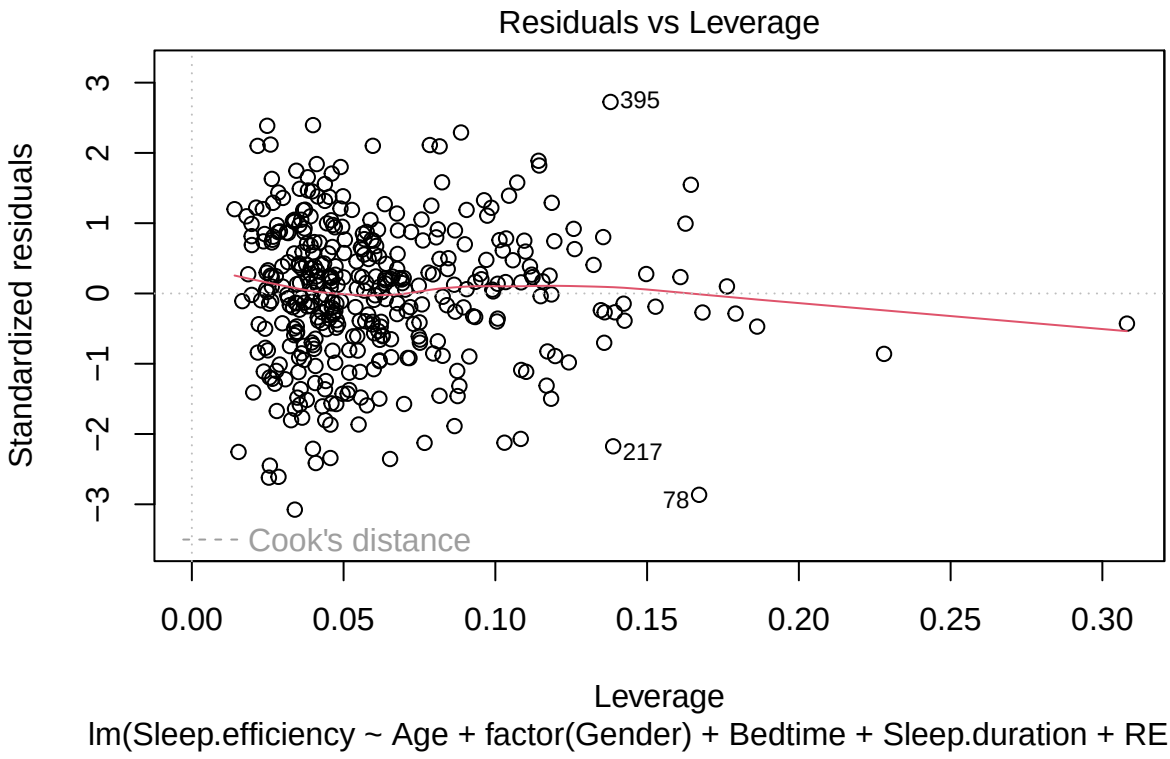
```
##  [1] ID                     Age                    Gender
##  [4] Bedtime                Wakeup.time            Sleep.duration
##  [7] Sleep.efficiency       REM.sleep.percentage   Deep.sleep.percentage
## [10] Light.sleep.percentage Awakenings             Caffeine.consumption
## [13] Alcohol.consumption    Smoking.status         Exercise.frequency
## <0 rows> (or 0-length row.names)
```

No points outside a cook's distance of 1 found.

```
plot(finalelevenmod,pch=18,col="red",which=c(4))
```




```
plot(finalelevenmod,which=5)
```




```
lev=hatvalues(finalelevenmod)
p = length(coef(finalelevenmod))
n = nrow(df)
outlier3p = lev[lev>(3*p/n)]
print(outlier3p)
```

```
##       258       336       379
## 0.3081001 0.2280643 0.1862574
```

```
outi <- c(258,336,379)
df_new = df[-outi,]
```

Building the model again without the outliers

```
finalelevenmod <- lm(Sleep.efficiency~Age+factor(Gender)+Bedtime+Sleep.duration+REM.sleep.percentage+ D

summary(finalelevenmod)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
```

```
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency + Age * factor(Smoking.status) +
##     Age * Deep.sleep.percentage + Age * Awakenings + Age:factor(Smoking.status) +
##     factor(Gender):Alcohol.consumption + Sleep.duration:Alcohol.consumption +
##     REM.sleep.percentage:Alcohol.consumption + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Awakenings:factor(Smoking.status) + Awakenings:Exercise.frequency +
##     Alcohol.consumption:Exercise.frequency, data = df_new)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.158461 -0.030808  0.003561  0.036317  0.127831
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                     6.986e-02  5.926e-02   1.179
## Age                                             5.531e-03  8.806e-04   6.281
## factor(Gender)Male                              2.057e-02  7.680e-03   2.679
## Bedtime                                        -1.234e-04  5.662e-04  -0.218
## Sleep.duration                                  5.793e-03  3.798e-03   1.525
## REM.sleep.percentage                            8.825e-03  1.049e-03   8.410
## Deep.sleep.percentage                           8.695e-03  7.090e-04  12.264
## Awakenings                                      2.009e-02  1.140e-02   1.762
## Caffeine.consumption                            1.019e-04  1.030e-04   0.989
## Alcohol.consumption                             3.621e-02  1.990e-02   1.820
## factor(Smoking.status)Yes                      -1.404e-01  3.019e-02  -4.651
## Exercise.frequency                              6.211e-03  3.231e-03   1.922
## Age:factor(Smoking.status)Yes                  -1.481e-03  4.594e-04  -3.224
## Age:Deep.sleep.percentage                      -6.889e-05  1.371e-05  -5.026
## Age:Awakenings                                 -5.036e-04  1.638e-04  -3.073
## factor(Gender)Male:Alcohol.consumption          -1.082e-02  3.666e-03  -2.951
## Sleep.duration:Alcohol.consumption              -4.175e-03  2.016e-03  -2.071
## REM.sleep.percentage:Alcohol.consumption        -1.101e-03  5.072e-04  -2.170
## Deep.sleep.percentage:Awakenings                -6.174e-04  1.624e-04  -3.802
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.594e-03  4.057e-04   6.395
## Awakenings:Alcohol.consumption                  4.405e-03  1.464e-03   3.008
## Awakenings:factor(Smoking.status)Yes            1.241e-02  4.959e-03   2.502
## Awakenings:Exercise.frequency                  -3.029e-03  1.535e-03  -1.973
## Alcohol.consumption:Exercise.frequency          5.533e-03  1.342e-03   4.123
##                                                Pr(>|t|)
## (Intercept)                                    0.239232
## Age                                            9.66e-10 ***
## factor(Gender)Male                             0.007723 **
## Bedtime                                        0.827543
## Sleep.duration                                 0.128139
## REM.sleep.percentage                           9.61e-16 ***
## Deep.sleep.percentage                           < 2e-16 ***
## Awakenings                                     0.078884 .
## Caffeine.consumption                           0.323322
## Alcohol.consumption                            0.069563 .
## factor(Smoking.status)Yes                      4.63e-06 ***
## Exercise.frequency                             0.055352 .
## Age:factor(Smoking.status)Yes                  0.001381 **
## Age:Deep.sleep.percentage                      7.88e-07 ***
```

```
## Age:Awakenings                                    0.002278 **
## factor(Gender)Male:Alcohol.consumption            0.003377 **
## Sleep.duration:Alcohol.consumption                0.039061 *
## REM.sleep.percentage:Alcohol.consumption          0.030668 *
## Deep.sleep.percentage:Awakenings                  0.000168 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 4.98e-10 ***
## Awakenings:Alcohol.consumption                    0.002814 **
## Awakenings:factor(Smoking.status)Yes              0.012776 *
## Awakenings:Exercise.frequency                     0.049236 *
## Alcohol.consumption:Exercise.frequency            4.64e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05258 on 361 degrees of freedom
## Multiple R-squared:  0.8591, Adjusted R-squared:  0.8501
## F-statistic: 95.66 on 23 and 361 DF,  p-value: < 2.2e-16
```

```
shapiro.test(residuals(finalelevenmod))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(finalelevenmod)
## W = 0.99315, p-value = 0.07705
```

Fail to reject. The normality condition is satisfied.

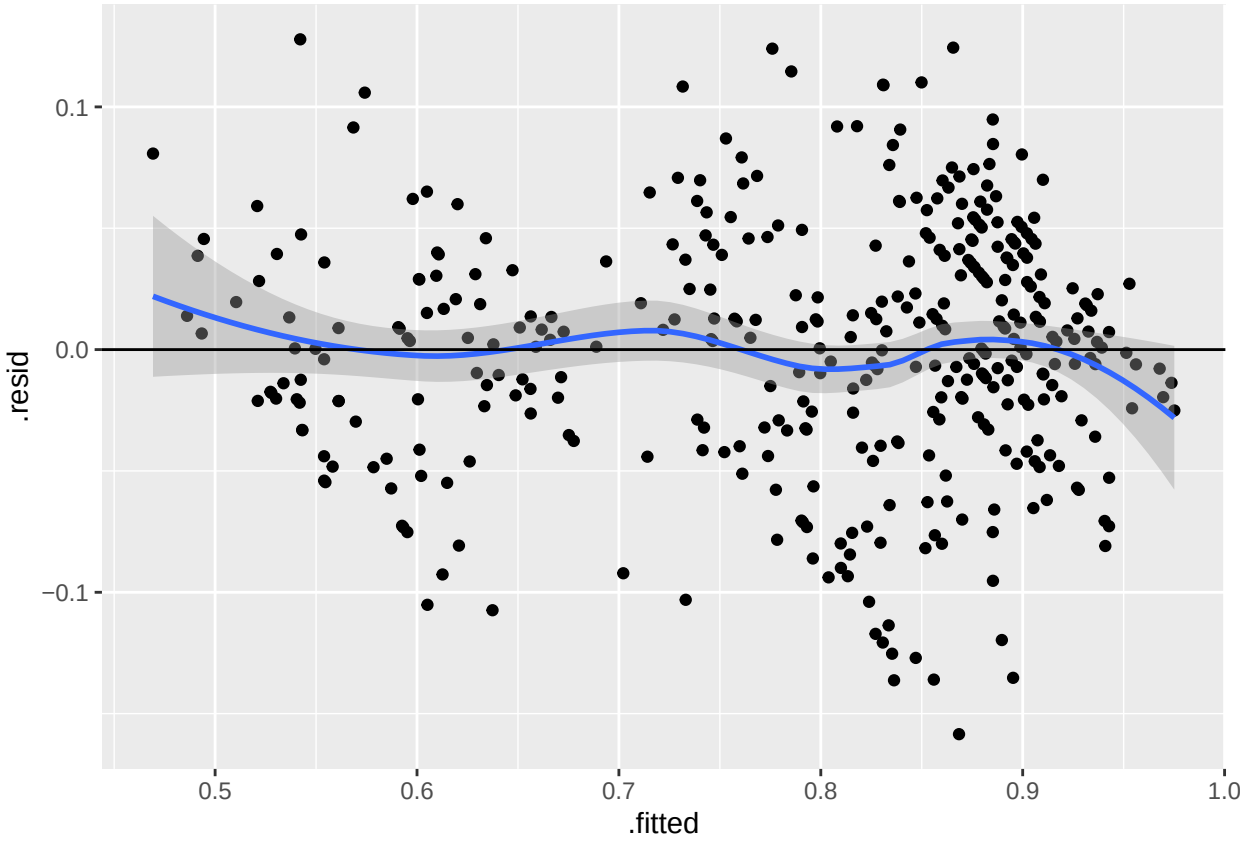
```
bptest(finalelevenmod)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  finalelevenmod
## BP = 23.071, df = 23, p-value = 0.4566
```

Fail to reject. The homoskedasticity condition is satisfied.

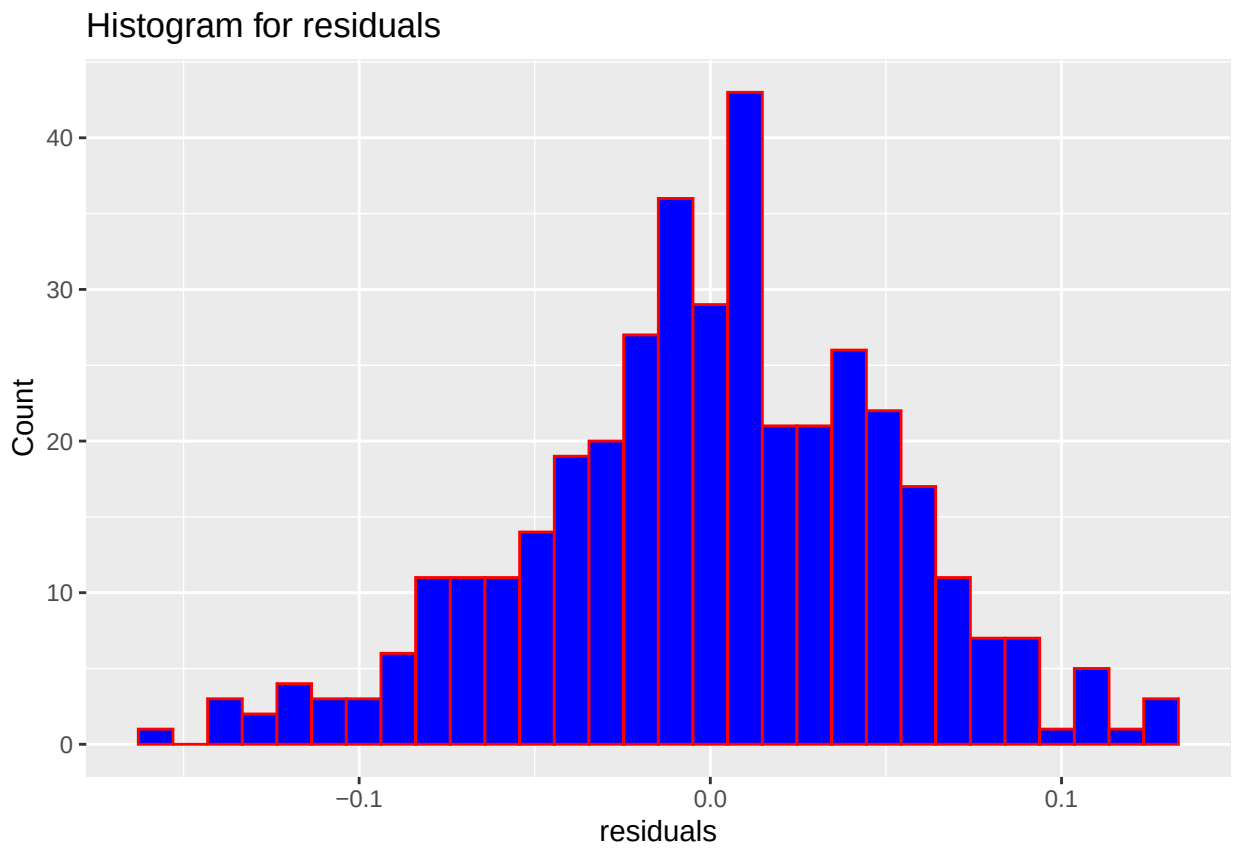
```
ggplot(finalelevenmod, aes(x=.fitted, y=.resid)) +
geom_point() + geom_smooth()+
geom_hline(yintercept = 0)
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

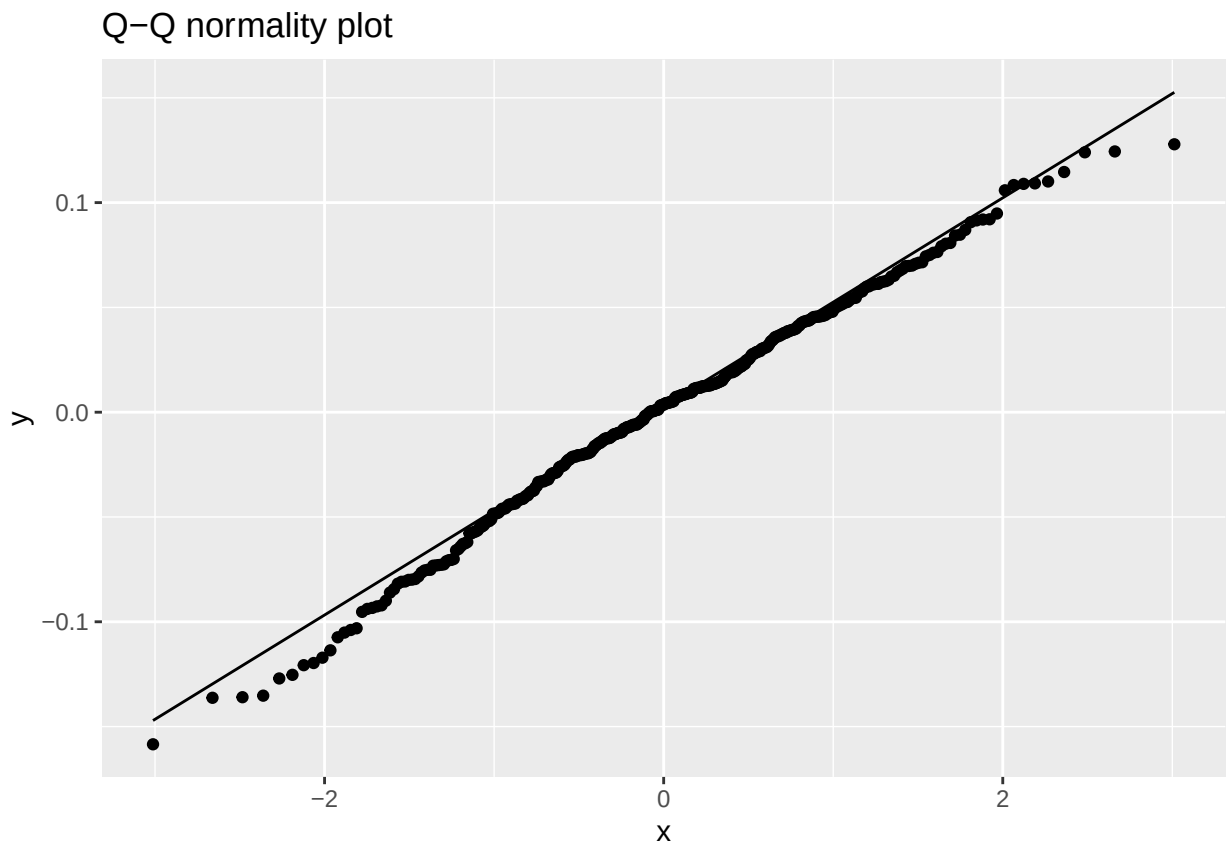



```
ggplot(data=df_new, aes(residuals(finalelevenmod))) +
geom_histogram(color='red',fill='blue') +
labs(title="Histogram for residuals") +
labs(x="residuals", y="Count")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
ggplot(df_new, aes(sample=finalelevenmod$residuals)) +
stat_qq() +
stat_qq_line() +
ggtitle("Q-Q normality plot")
```



Our final model is the 11-variable model.

```
summary(finalelevenmod)
```

```
##
## Call:
## lm(formula = Sleep.efficiency ~ Age + factor(Gender) + Bedtime +
##     Sleep.duration + REM.sleep.percentage + Deep.sleep.percentage +
##     Awakenings + Caffeine.consumption + Alcohol.consumption +
##     factor(Smoking.status) + Exercise.frequency + Age * factor(Smoking.status) +
##     Age * Deep.sleep.percentage + Age * Awakenings + Age:factor(Smoking.status) +
##     factor(Gender):Alcohol.consumption + Sleep.duration:Alcohol.consumption +
##     REM.sleep.percentage:Alcohol.consumption + Deep.sleep.percentage:Awakenings +
##     Deep.sleep.percentage:factor(Smoking.status) + Awakenings:Alcohol.consumption +
##     Awakenings:factor(Smoking.status) + Awakenings:Exercise.frequency +
##     Alcohol.consumption:Exercise.frequency, data = df_new)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.158461 -0.030808  0.003561  0.036317  0.127831
##
## Coefficients:
##                                                Estimate Std. Error t value
## (Intercept)                                   6.986e-02  5.926e-02   1.179
## Age                                           5.531e-03  8.806e-04   6.281
## factor(Gender)Male                            2.057e-02  7.680e-03   2.679
```

```
## Bedtime                                          -1.234e-04  5.662e-04  -0.218
## Sleep.duration                                    5.793e-03  3.798e-03   1.525
## REM.sleep.percentage                              8.825e-03  1.049e-03   8.410
## Deep.sleep.percentage                             8.695e-03  7.090e-04  12.264
## Awakenings                                        2.009e-02  1.140e-02   1.762
## Caffeine.consumption                              1.019e-04  1.030e-04   0.989
## Alcohol.consumption                               3.621e-02  1.990e-02   1.820
## factor(Smoking.status)Yes                        -1.404e-01  3.019e-02  -4.651
## Exercise.frequency                                6.211e-03  3.231e-03   1.922
## Age:factor(Smoking.status)Yes                    -1.481e-03  4.594e-04  -3.224
## Age:Deep.sleep.percentage                        -6.889e-05  1.371e-05  -5.026
## Age:Awakenings                                   -5.036e-04  1.638e-04  -3.073
## factor(Gender)Male:Alcohol.consumption           -1.082e-02  3.666e-03  -2.951
## Sleep.duration:Alcohol.consumption               -4.175e-03  2.016e-03  -2.071
## REM.sleep.percentage:Alcohol.consumption         -1.101e-03  5.072e-04  -2.170
## Deep.sleep.percentage:Awakenings                 -6.174e-04  1.624e-04  -3.802
## Deep.sleep.percentage:factor(Smoking.status)Yes  2.594e-03  4.057e-04   6.395
## Awakenings:Alcohol.consumption                    4.405e-03  1.464e-03   3.008
## Awakenings:factor(Smoking.status)Yes              1.241e-02  4.959e-03   2.502
## Awakenings:Exercise.frequency                    -3.029e-03  1.535e-03  -1.973
## Alcohol.consumption:Exercise.frequency            5.533e-03  1.342e-03   4.123
##                                                 Pr(>|t|)
## (Intercept)                                     0.239232
## Age                                             9.66e-10 ***
## factor(Gender)Male                              0.007723 **
## Bedtime                                         0.827543
## Sleep.duration                                  0.128139
## REM.sleep.percentage                            9.61e-16 ***
## Deep.sleep.percentage                            < 2e-16 ***
## Awakenings                                      0.078884 .
## Caffeine.consumption                            0.323322
## Alcohol.consumption                             0.069563 .
## factor(Smoking.status)Yes                       4.63e-06 ***
## Exercise.frequency                              0.055352 .
## Age:factor(Smoking.status)Yes                   0.001381 **
## Age:Deep.sleep.percentage                       7.88e-07 ***
## Age:Awakenings                                  0.002278 **
## factor(Gender)Male:Alcohol.consumption          0.003377 **
## Sleep.duration:Alcohol.consumption              0.039061 *
## REM.sleep.percentage:Alcohol.consumption        0.030668 *
## Deep.sleep.percentage:Awakenings                0.000168 ***
## Deep.sleep.percentage:factor(Smoking.status)Yes 4.98e-10 ***
## Awakenings:Alcohol.consumption                  0.002814 **
## Awakenings:factor(Smoking.status)Yes            0.012776 *
## Awakenings:Exercise.frequency                   0.049236 *
## Alcohol.consumption:Exercise.frequency          4.64e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05258 on 361 degrees of freedom
## Multiple R-squared:  0.8591, Adjusted R-squared:  0.8501
## F-statistic: 95.66 on 23 and 361 DF,  p-value: < 2.2e-16
```

Prediction

We thought it would be interesting to evaluate our model against one of our own group member's sleep data. Luckily, we had historical sleep data on Graeme's sleep which contained all the variables used in the model.

```
my_eff = (7+(10/60)) / (9+(50/60))
my_age = 21
my_gend = "Male"
my_btime = convert_time_to_numeric(as.POSIXct("2021-04-02 1:24", format = "%Y-%m-%d %H:%M"))
my_dur = 8.25
my_REM = ((12.5+17.5+13.5+30+5+10)/60) / my_dur * 100
my_deep = 40
my_awake = 4
my_caff = 0
my_alcohol = 0
my_smoke = "No"
my_exer = 1

my_df = data.frame(Sleep.efficiency=my_eff,Age=my_age,Gender=my_gend,Bedtime=my_btime,Sleep.duration=my_

head(my_df)
```

```
##   Sleep.efficiency Age Gender Bedtime Sleep.duration REM.sleep.percentage
## 1        0.7288136  21   Male     1.4           8.25              17.87879
##   Deep.sleep.percentage Awakenings Caffeine.consumption Alcohol.consumption
## 1                    40          4                    0                   0
##   Smoking.status Exercise.frequency
## 1             No                  1
```

```
my_eff_perc = my_eff*100
my_eff_output <- paste("Actual sleep efficiency for Graeme:",
                       sprintf("%.2f%%", my_eff_perc),"efficiency")
print(my_eff_output)
```

```
## [1] "Actual sleep efficiency for Graeme: 72.88% efficiency"
```

```
prediction <- predict(finalelevenmod,my_df,interval="predict")
eff_output <- paste("Model predicted sleep efficiency for Graeme:",
                      sprintf("%.2f%%", prediction[,1]*100),"efficiency")

inteff_output <- paste("With 95% prediction interval between","(",
                       sprintf("%.2f%%", prediction[,2]*100),",",
                       sprintf("%.2f%%", prediction[,3]*100),")")

print(eff_output)
```

```
## [1] "Model predicted sleep efficiency for Graeme: 63.53% efficiency"
```

```
print(inteff_output)
```

```
## [1] "With 95% prediction interval between ( 52.83% , 74.23% )"
```

We can see that Graeme's actual sleep efficiency of 72.88% is within the 95% prediction interval estimated by the model. Though much more data would be required for further model validation.

A more recent night of sleep was also recorded, displayed in the dataframe below.

```
my_eff = 0.68
my_age = 23
my_gend = "Male"
my_btime = convert_time_to_numeric(as.POSIXct("2023-12-08 1:48", format = "%Y-%m-%d %H:%M"))
my_dur = 4.417
my_REM = (38/60) / my_dur * 100
my_deep = 25
my_awake = 7
my_caff = 0
my_alcohol = 9
my_smoke = "No"
my_exer = 0

my_df = data.frame(Sleep.efficiency=my_eff,Age=my_age,Gender=my_gend,Bedtime=my_btime,Sleep.duration=my_

head(my_df)
```

```
##   Sleep.efficiency Age Gender Bedtime Sleep.duration REM.sleep.percentage
## 1             0.68  23   Male     1.8          4.417             14.33854
##   Deep.sleep.percentage Awakenings Caffeine.consumption Alcohol.consumption
## 1                    25          7                    0                   9
##   Smoking.status Exercise.frequency
## 1             No                  0
```

```
my_eff_perc = my_eff*100
my_eff_output <- paste("Actual sleep efficiency for Graeme:",
                       sprintf("%.2f%%", my_eff_perc),"efficiency")
print(my_eff_output)
```

```
## [1] "Actual sleep efficiency for Graeme: 68.00% efficiency"
```

```
prediction <- predict(finalelevenmod,my_df,interval="predict")
eff_output <- paste("Model predicted sleep efficiency for Graeme:",
                    sprintf("%.2f%%", prediction[,1]*100),"efficiency")

inteff_output <- paste("With 95% prediction interval between","(",
                       sprintf("%.2f%%", prediction[,2]*100),",",
                       sprintf("%.2f%%", prediction[,3]*100),")")

print(eff_output)
```

```
## [1] "Model predicted sleep efficiency for Graeme: 69.69% efficiency"
```

```
print(inteff_output)
```

```
## [1] "With 95% prediction interval between ( 50.69% , 88.69% )"
```

Here we tried a more recent night of sleep that was somewhat more irregular given the larger alcohol content, frequent awakenings, and short duration. The actual efficiency prediction from our model was less than 2% different from the measured value, though the 95% confidence interval for this test was fairly large as well.