STAT463 Project: Sleep Health and Lifestyle

Lisa Lu 31088272

```
# Import libraries
library(dplyr)
library(ggplot2)
library(stringr)
library(gridExtra)
```

Data Exploration and Preprocessing

```
# Read in dataset
dataset <- read.table("Sleep_health_and_lifestyle_dataset.csv",
                      header = TRUE, sep = ',', na.strings = "na")

# Print the first few rows of data frame
head(dataset)
```

```
##   Person.ID Gender Age          Occupation Sleep.Duration Quality.of.Sleep
## 1         1   Male  27    Software Engineer            6.1                6
## 2         2   Male  28              Doctor            6.2                6
## 3         3   Male  28              Doctor            6.2                6
## 4         4   Male  28 Sales Representative            5.9                4
## 5         5   Male  28 Sales Representative            5.9                4
## 6         6   Male  28    Software Engineer            5.9                4
##   Physical.Activity.Level Stress.Level BMI.Category Blood.Pressure Heart.Rate
## 1                      42            6   Overweight         126/83         77
## 2                      60            8       Normal         125/80         75
## 3                      60            8       Normal         125/80         75
## 4                      30            8        Obese         140/90         85
## 5                      30            8        Obese         140/90         85
## 6                      30            8        Obese         140/90         85
##   Daily.Steps Sleep.Disorder
## 1        4200           None
## 2       10000           None
## 3       10000           None
## 4        3000    Sleep Apnea
## 5        3000    Sleep Apnea
## 6        3000       Insomnia
```

```
# Explore the structure of the dataset
str(dataset)
```

```
## 'data.frame':    374 obs. of  13 variables:
```

```
##  $ Person.ID              : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Gender                 : chr  "Male" "Male" "Male" "Male" ...
##  $ Age                    : int  27 28 28 28 28 28 29 29 29 29 ...
##  $ Occupation             : chr  "Software Engineer" "Doctor" "Doctor" "Sales Representative" ...
##  $ Sleep.Duration         : num  6.1 6.2 6.2 5.9 5.9 5.9 6.3 7.8 7.8 7.8 ...
##  $ Quality.of.Sleep       : int  6 6 6 4 4 4 6 7 7 7 ...
##  $ Physical.Activity.Level: int  42 60 60 30 30 30 40 75 75 75 ...
##  $ Stress.Level           : int  6 8 8 8 8 8 7 6 6 6 ...
##  $ BMI.Category           : chr  "Overweight" "Normal" "Normal" "Obese" ...
##  $ Blood.Pressure         : chr  "126/83" "125/80" "125/80" "140/90" ...
##  $ Heart.Rate             : int  77 75 75 85 85 85 82 70 70 70 ...
##  $ Daily.Steps            : int  4200 10000 10000 3000 3000 3000 3500 8000 8000 8000 ...
##  $ Sleep.Disorder         : chr  "None" "None" "None" "Sleep Apnea" ...
```

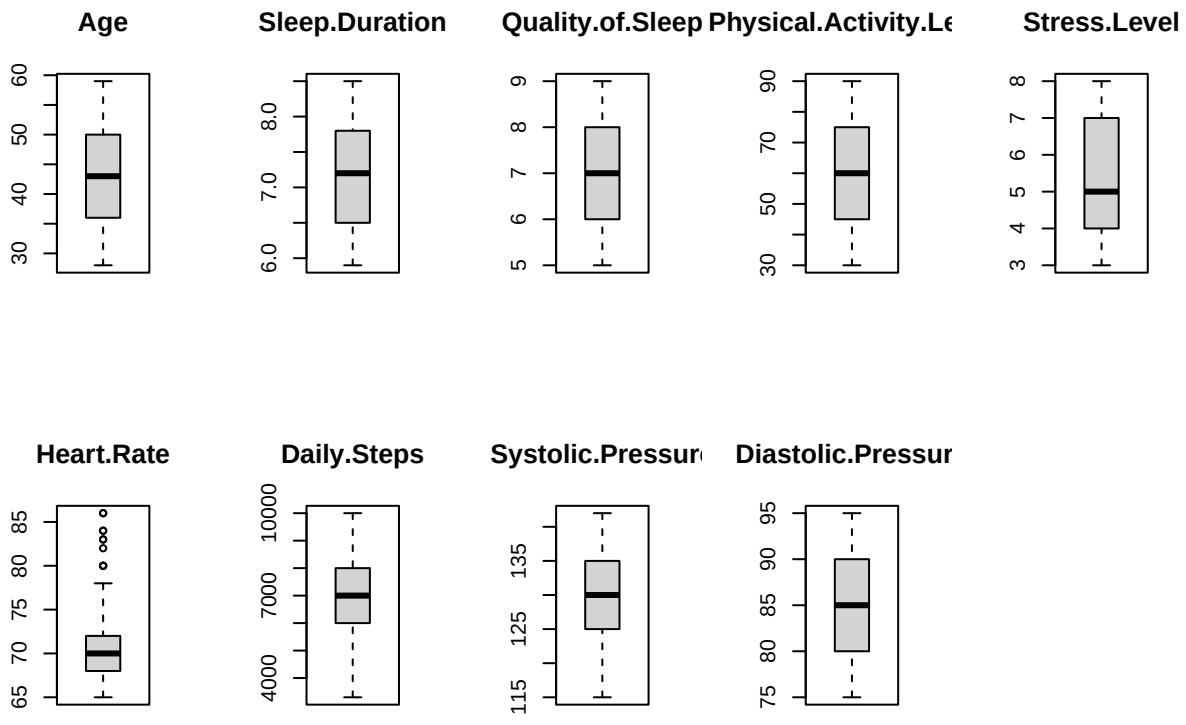
```r
# Get a descriptive statistics
summary(dataset)
```

```
##    Person.ID         Gender               Age         Occupation
##  Min.   :  1.00   Length:374         Min.   :27.00   Length:374
##  1st Qu.: 94.25   Class :character   1st Qu.:35.25   Class :character
##  Median :187.50   Mode  :character   Median :43.00   Mode  :character
##  Mean   :187.50                      Mean   :42.18
##  3rd Qu.:280.75                      3rd Qu.:50.00
##  Max.   :374.00                      Max.   :59.00
##  Sleep.Duration  Quality.of.Sleep Physical.Activity.Level  Stress.Level
##  Min.   :5.800   Min.   :4.000    Min.   :30.00           Min.   :3.000
##  1st Qu.:6.400   1st Qu.:6.000    1st Qu.:45.00           1st Qu.:4.000
##  Median :7.200   Median :7.000    Median :60.00           Median :5.000
##  Mean   :7.132   Mean   :7.313    Mean   :59.17           Mean   :5.385
##  3rd Qu.:7.800   3rd Qu.:8.000    3rd Qu.:75.00           3rd Qu.:7.000
##  Max.   :8.500   Max.   :9.000    Max.   :90.00           Max.   :8.000
##  BMI.Category       Blood.Pressure       Heart.Rate     Daily.Steps
##  Length:374         Length:374         Min.   :65.00   Min.   : 3000
##  Class :character   Class :character   1st Qu.:68.00   1st Qu.: 5600
##  Mode  :character   Mode  :character   Median :70.00   Median : 7000
##                                        Mean   :70.17   Mean   : 6817
##                                        3rd Qu.:72.00   3rd Qu.: 8000
##                                        Max.   :86.00   Max.   :10000
##  Sleep.Disorder
##  Length:374
##  Class :character
##  Mode  :character
##
##
##
```

```r
# Data preprocessing
# Split Blood Pressure column into systolic and diastolic pressure as numeric data
dataset[c('Systolic.Pressure', 'Diastolic.Pressure')] <- as.numeric(str_split_fixed(dataset$Blood.Pressu

# Combine "Normal" and "Normal Weight" values in BMI.Category
dataset$BMI.Category[dataset$BMI.Category == "Normal Weight"] <- "Normal"
# Change "Obese" into "Overweight"
```

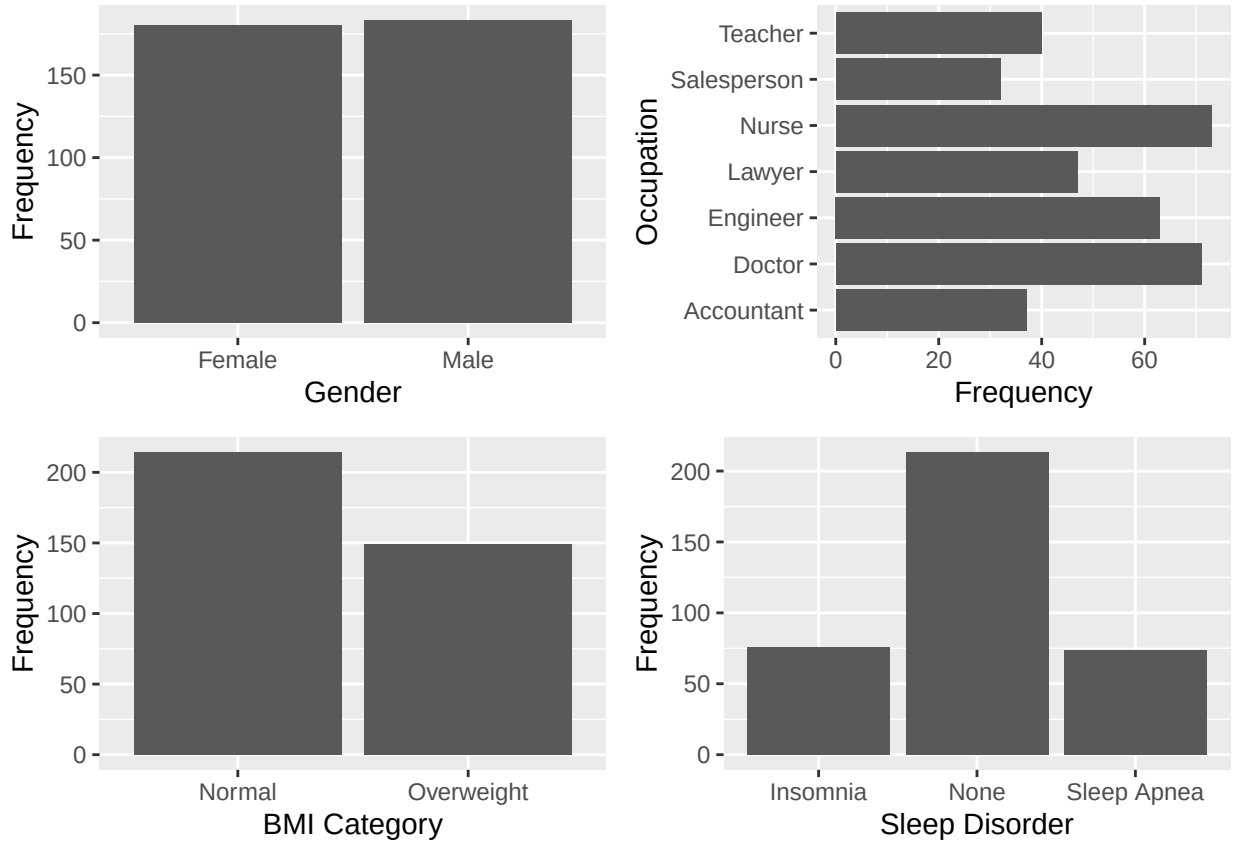
```r
dataset$BMI.Category[dataset$BMI.Category == "Obese"] <- "Overweight"

# Remove Occupations that has count under 10 (namely, Software Engineer, Scientist, Sales Representativ
dataset <- dataset[!(dataset$Occupation =="Software Engineer" | dataset$Occupation =="Scientist" |
          dataset$Occupation =="Sales Representative" |
          dataset$Occupation =="Manager"),]
```

```r
# Get the numeric data
numeric.data <- dataset[,c(3,5,6,7,8,11,12,14,15)]

# Plot the data -- Boxplot for numeric data and Histogram for categorical data
# Boxplots

par(mfrow=c(2,5))
for (i in 1:length(numeric.data)) {
        boxplot(numeric.data[,i], main=names(numeric.data[i]), type="l")

    }
```




```r
# Frequency charts of the categorical data
gender <- ggplot(data = dataset, aes(x = Gender)) +
geom_bar() +
labs(y = "Frequency", x = "Gender")

occupation <- ggplot(data = dataset, aes(y = Occupation)) +
```

```
geom_bar() +
labs(y = "Occupation", x = "Frequency")

bmi <- ggplot(data = dataset, aes(x = BMI.Category)) +
geom_bar() +
labs(y = "Frequency", x = "BMI Category")

sleep_disorder <- ggplot(data = dataset, aes(x = Sleep.Disorder)) +
geom_bar() +
labs(y = "Frequency", x = "Sleep Disorder")

grid.arrange(gender,occupation, bmi, sleep_disorder, ncol = 2, nrow = 2)
```



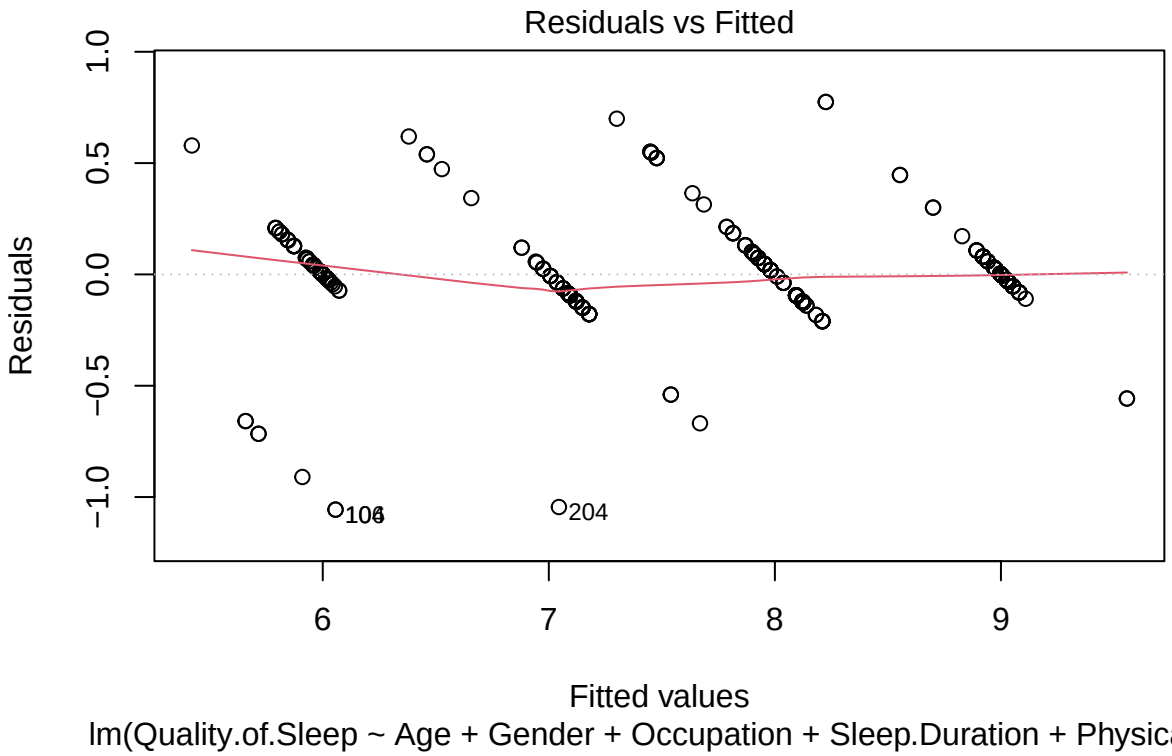
Explore the factors affecting quality of sleep

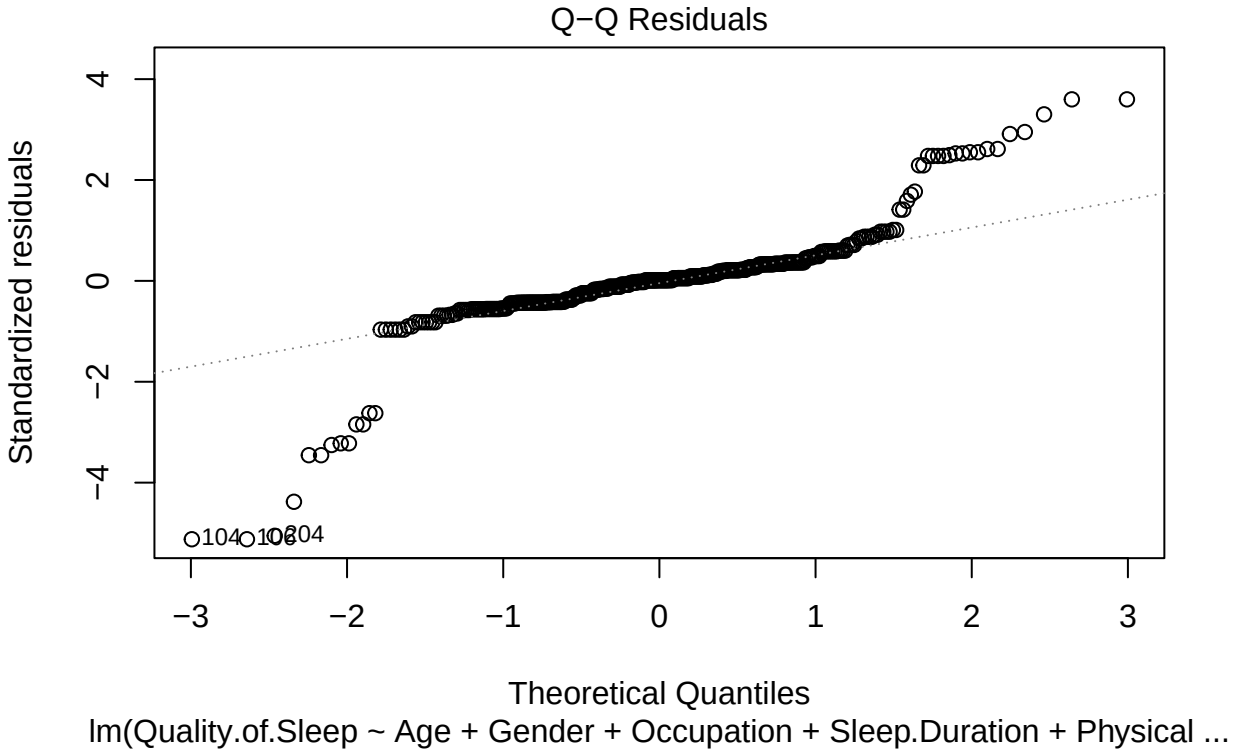
```
# Use linear regression model
sleep <- dataset[, -c(1, 10)]
# Get the full multiple linear regression model
lr_full <- lm(Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration + Physical.Activity.Level +

summary(lr_full)
```

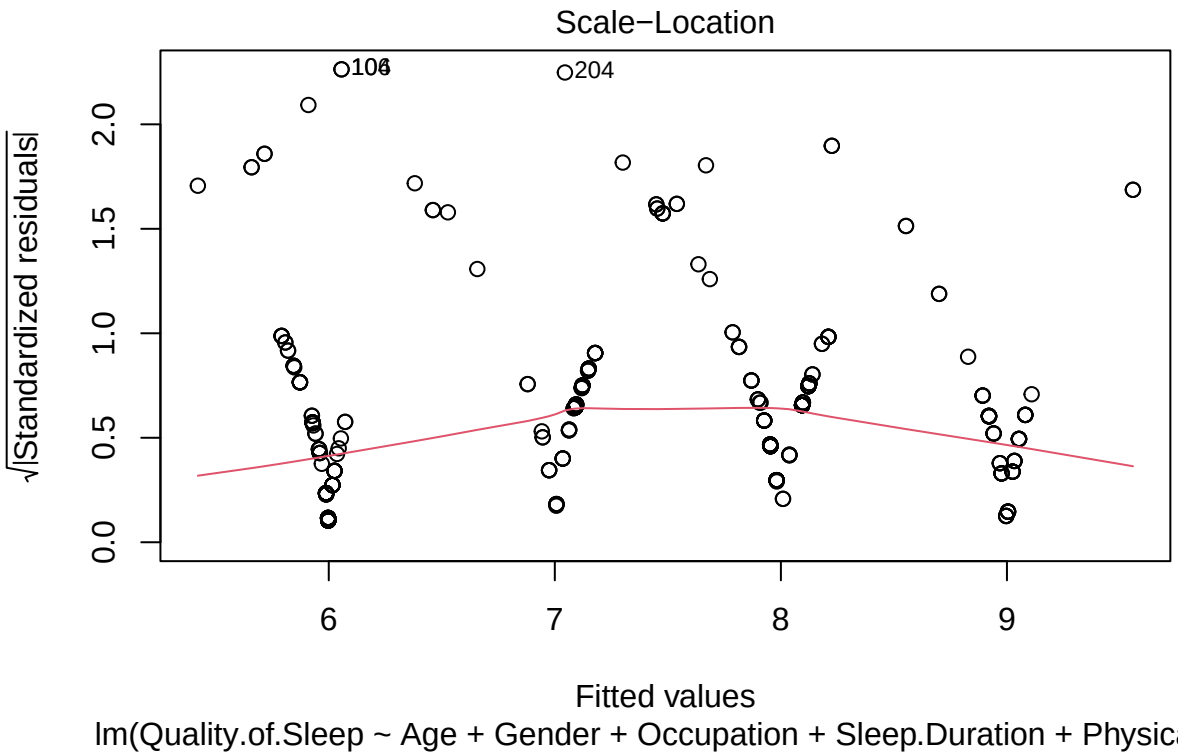
```
##
```

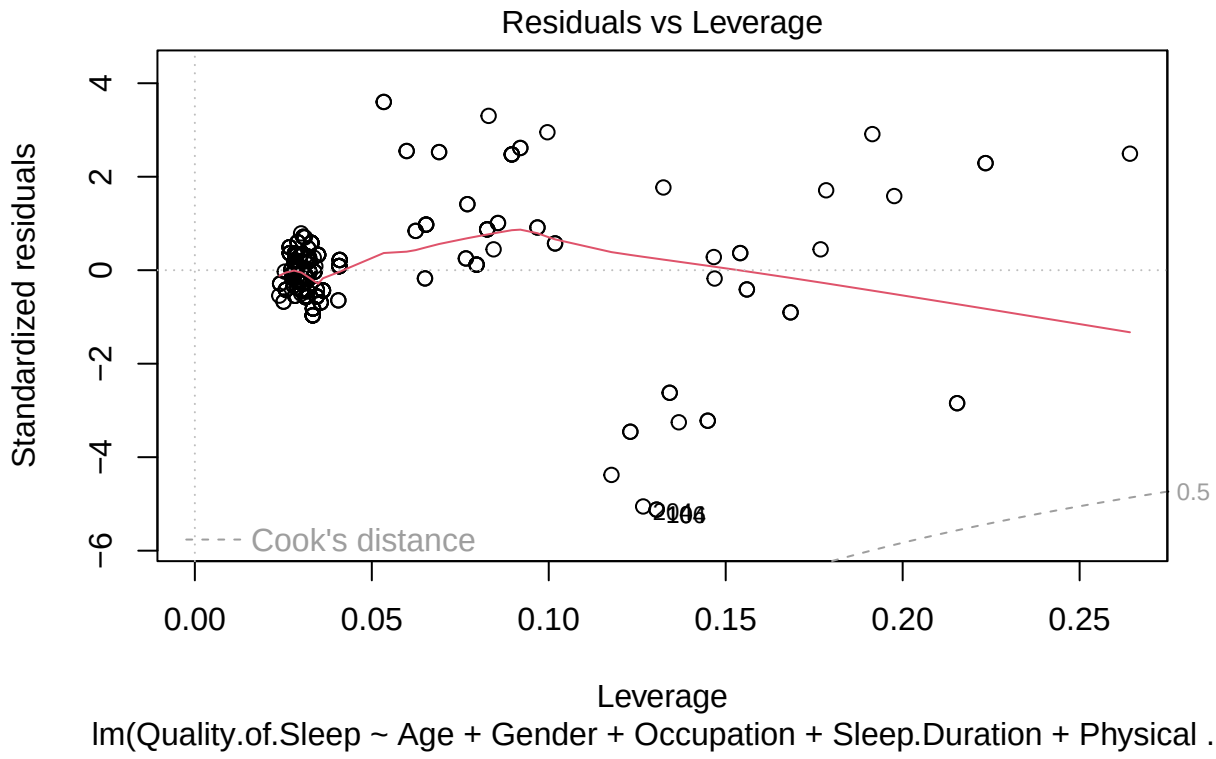
```
## Call:
## lm(formula = Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Heart.Rate + Daily.Steps +
##     Systolic.Pressure + Diastolic.Pressure + BMI.Category, data = dataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05646 -0.09108  0.00292  0.07118  0.77459
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               6.415e+00  8.176e-01   7.847 5.39e-14 ***
## Age                       5.608e-02  5.121e-03  10.950  < 2e-16 ***
## GenderMale                5.837e-01  6.983e-02   8.358 1.57e-15 ***
## OccupationDoctor         -4.938e-01  7.999e-02  -6.173 1.88e-09 ***
## OccupationEngineer       -6.836e-01  7.957e-02  -8.592 2.98e-16 ***
## OccupationLawyer         -3.983e-01  9.304e-02  -4.281 2.41e-05 ***
## OccupationNurse          -2.182e-01  8.422e-02  -2.591  0.00999 **
## OccupationSalesperson    -1.028e+00  8.924e-02 -11.516  < 2e-16 ***
## OccupationTeacher        -5.605e-01  7.511e-02  -7.463 6.92e-13 ***
## Sleep.Duration            2.833e-01  4.730e-02   5.989 5.29e-09 ***
## Physical.Activity.Level -2.281e-03  1.489e-03  -1.532  0.12641
## Stress.Level             -4.043e-01  2.103e-02 -19.220  < 2e-16 ***
## Heart.Rate               -6.741e-03  6.427e-03  -1.049  0.29503
## Daily.Steps               6.186e-05  2.045e-05   3.025  0.00267 **
## Systolic.Pressure         2.235e-02  1.530e-02   1.461  0.14486
## Diastolic.Pressure       -4.303e-02  2.046e-02  -2.103  0.03619 *
## BMI.CategoryOverweight   -4.278e-01  8.869e-02  -4.824 2.11e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2212 on 346 degrees of freedom
## Multiple R-squared:  0.9642, Adjusted R-squared:  0.9625
## F-statistic: 581.7 on 16 and 346 DF,  p-value: < 2.2e-16
```

```
plot(lr_full)
```


Residuals vs Fitted
Residuals
1.0
0.5
0.0
−0.5
−1.0
106
204
6
7
8
9
Fitted values
lm(Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration + Physical ...


Q−Q Residuals
Standardized residuals
4
2
0
−2
−4
104
106
204
−3
−2
−1
0
1
2
3
Theoretical Quantiles
lm(Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration + Physical ...


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
204
6
7
8
9
Fitted values
lm(Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration + Physical ...




```
lr_backward <- step(lr_full, direction = "backward")
```

```
## Start:  AIC=-1078.74
## Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Heart.Rate + Daily.Steps +
##     Systolic.Pressure + Diastolic.Pressure + BMI.Category
##
##                           Df Sum of Sq    RSS      AIC
## - Heart.Rate               1    0.0538 16.983 -1079.58
## <none>                                 16.929 -1078.74
## - Systolic.Pressure        1    0.1045 17.033 -1078.50
## - Physical.Activity.Level  1    0.1148 17.044 -1078.28
## - Diastolic.Pressure       1    0.2164 17.145 -1076.13
## - Daily.Steps              1    0.4478 17.377 -1071.26
## - BMI.Category             1    1.1385 18.067 -1057.11
## - Sleep.Duration           1    1.7550 18.684 -1044.93
## - Gender                   1    3.4180 20.347 -1013.98
## - Age                      1    5.8659 22.795  -972.74
## - Occupation               6   10.2728 27.202  -918.58
## - Stress.Level             1   18.0734 35.002  -817.05
##
## Step:  AIC=-1079.58
## Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Daily.Steps + Systolic.Pressure +
##     Diastolic.Pressure + BMI.Category
```

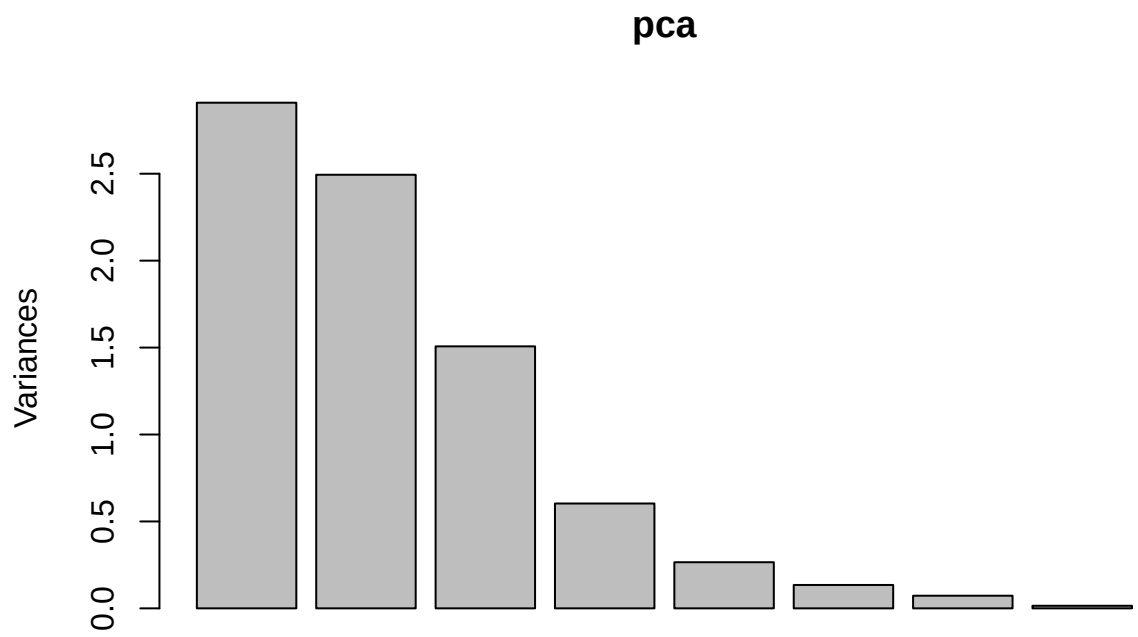
```
##
##                            Df Sum of Sq    RSS      AIC
## - Systolic.Pressure         1    0.0763 17.059 -1079.96
## <none>                                  16.983 -1079.58
## - Diastolic.Pressure        1    0.1779 17.161 -1077.80
## - Physical.Activity.Level   1    0.2759 17.258 -1075.73
## - Daily.Steps               1    0.6754 17.658 -1067.43
## - BMI.Category              1    1.5375 18.520 -1050.12
## - Sleep.Duration            1    1.7026 18.685 -1046.90
## - Gender                    1    3.9128 20.895 -1006.32
## - Age                       1    7.0850 24.068  -955.01
## - Occupation                6   10.4617 27.444  -917.36
## - Stress.Level              1   25.5119 42.494  -748.65
##
## Step:  AIC=-1079.96
## Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Daily.Steps + Diastolic.Pressure +
##     BMI.Category
##
##                            Df Sum of Sq    RSS      AIC
## <none>                                  17.059 -1079.96
## - Physical.Activity.Level   1    0.2274 17.286 -1077.15
## - Diastolic.Pressure        1    0.2383 17.297 -1076.92
## - Daily.Steps               1    0.7294 17.788 -1066.76
## - Sleep.Duration            1    1.7099 18.769 -1047.28
## - BMI.Category              1    2.1556 19.214 -1038.76
## - Gender                    1    4.2126 21.271 -1001.84
## - Occupation                6   10.4319 27.491  -918.74
## - Age                       1   12.2435 29.302  -885.58
## - Stress.Level              1   25.7039 42.763  -748.36
```

```
summary(lr_backward)
```

```
##
## Call:
## lm(formula = Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Daily.Steps + Diastolic.Pressure +
##     BMI.Category, data = dataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11013 -0.08980  0.00291  0.07526  0.75796
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               6.518e+00  6.127e-01  10.639  < 2e-16 ***
## Age                       6.146e-02  3.889e-03  15.804  < 2e-16 ***
## GenderMale                6.168e-01  6.653e-02   9.270  < 2e-16 ***
## OccupationDoctor         -4.883e-01  7.958e-02  -6.135 2.31e-09 ***
## OccupationEngineer       -6.658e-01  7.731e-02  -8.612 2.54e-16 ***
## OccupationLawyer         -3.726e-01  8.811e-02  -4.229 3.01e-05 ***
## OccupationNurse          -2.404e-01  8.319e-02  -2.890 0.004091 **
## OccupationSalesperson    -1.012e+00  8.881e-02 -11.398  < 2e-16 ***
## OccupationTeacher        -5.219e-01  7.084e-02  -7.368 1.27e-12 ***
```

```
## Sleep.Duration           2.518e-01  4.264e-02   5.906 8.33e-09 ***
## Physical.Activity.Level -2.714e-03  1.260e-03  -2.154 0.031957 *
## Stress.Level            -4.118e-01  1.798e-02 -22.899  < 2e-16 ***
## Daily.Steps              5.447e-05  1.412e-05   3.857 0.000137 ***
## Diastolic.Pressure      -1.447e-02  6.563e-03  -2.205 0.028104 *
## BMI.CategoryOverweight  -5.028e-01  7.582e-02  -6.631 1.27e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2214 on 348 degrees of freedom
## Multiple R-squared:  0.9639, Adjusted R-squared:  0.9624
## F-statistic: 663.3 on 14 and 348 DF,  p-value: < 2.2e-16
```

```
 # Use PCA analysis
# Standardise the Quality of sleep and stress level
Quality.of.Sleep <- data.frame(scale(dataset$Quality.of.Sleep))
Stress.Level <- data.frame(scale(dataset$Stress.Level))
Physical.Level <-data.frame(scale(dataset$Physical.Activity.Level))

pca_dataset <- select(dataset, Age, Physical.Activity.Level, Stress.Level, Sleep.Duration, Heart.Rate, D
pca <- prcomp(scale(pca_dataset))
screeplot(pca)
```




```
summary(pca)
```

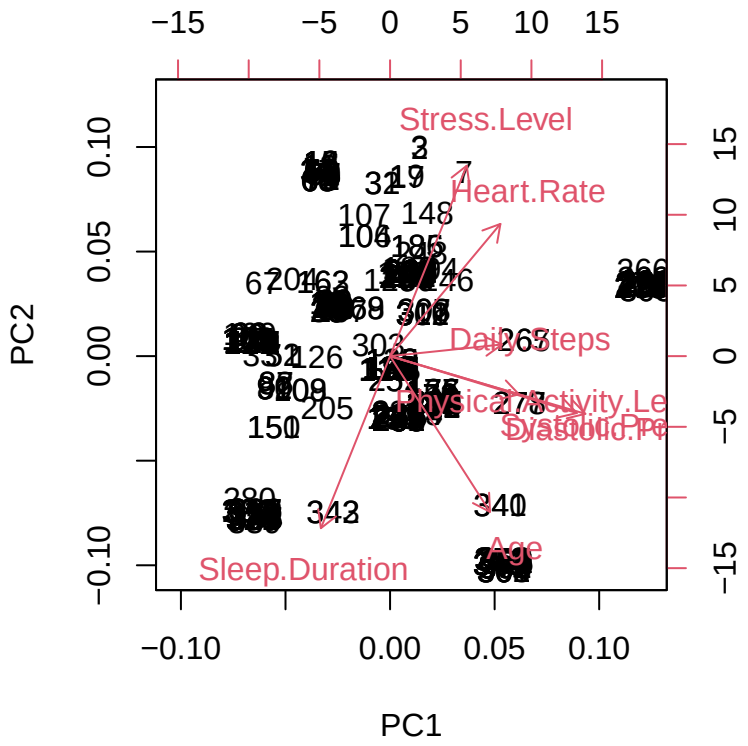
```
## Importance of components:
```

```
##                            PC1    PC2    PC3    PC4     PC5     PC6     PC7
## Standard deviation     1.7054 1.5792 1.2276 0.7766 0.51508 0.36686 0.26950
## Proportion of Variance 0.3636 0.3118 0.1884 0.0754 0.03316 0.01682 0.00908
## Cumulative Proportion  0.3636 0.6753 0.8637 0.9391 0.97223 0.98905 0.99813
##                            PC8
## Standard deviation     0.12220
## Proportion of Variance 0.00187
## Cumulative Proportion  1.00000
```

```
pca$rotation
```

```
##                                 PC1         PC2         PC3         PC4
## Age                       0.2719792 -0.45775940  0.25233251  0.06622055
## Physical.Activity.Level   0.3509205 -0.11243330 -0.59235213 -0.25620707
## Stress.Level              0.2086835  0.55957697  0.03128683  0.11152440
## Sleep.Duration           -0.1882257 -0.50575394 -0.21613685 -0.46131921
## Heart.Rate                0.3003581  0.38856803  0.07434625 -0.73786701
## Daily.Steps               0.3046078  0.03428766 -0.63132174  0.38965387
## Systolic.Pressure         0.5082469 -0.16728555  0.30556777  0.04173823
## Diastolic.Pressure        0.5315906 -0.16957723  0.20071641  0.08185187
##                                 PC5         PC6         PC7           PC8
## Age                       -0.77166106  0.1759377  0.14347212 -0.0377849586
## Physical.Activity.Level    0.14928427  0.2885336  0.58493765 -0.0009550409
## Stress.Level              -0.28999499 -0.5945634  0.43004786 -0.0858159758
## Sleep.Duration            -0.03918611 -0.6650008 -0.06725238 -0.0311275936
## Heart.Rate                -0.19971854  0.1562272 -0.37946130  0.0232273436
## Daily.Steps               -0.21673720 -0.1159727 -0.53210514  0.1099118223
## Systolic.Pressure          0.30491672 -0.1966373  0.01097240  0.6977292384
## Diastolic.Pressure         0.34173681 -0.1163604 -0.14336144 -0.7005653914
```

```
biplot(pca)
```




```
step.model <- step(lr_full, direction = "backward")
```

```
## Start:  AIC=-1078.74
## Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Heart.Rate + Daily.Steps +
##     Systolic.Pressure + Diastolic.Pressure + BMI.Category
##
##                           Df Sum of Sq    RSS      AIC
## - Heart.Rate               1    0.0538 16.983 -1079.58
## <none>                                 16.929 -1078.74
## - Systolic.Pressure        1    0.1045 17.033 -1078.50
## - Physical.Activity.Level  1    0.1148 17.044 -1078.28
## - Diastolic.Pressure       1    0.2164 17.145 -1076.13
## - Daily.Steps              1    0.4478 17.377 -1071.26
## - BMI.Category             1    1.1385 18.067 -1057.11
## - Sleep.Duration           1    1.7550 18.684 -1044.93
## - Gender                   1    3.4180 20.347 -1013.98
## - Age                      1    5.8659 22.795  -972.74
## - Occupation               6   10.2728 27.202  -918.58
## - Stress.Level             1   18.0734 35.002  -817.05
##
## Step:  AIC=-1079.58
## Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Daily.Steps + Systolic.Pressure +
##     Diastolic.Pressure + BMI.Category
```

```
##
##                           Df Sum of Sq    RSS      AIC
## - Systolic.Pressure        1    0.0763 17.059 -1079.96
## <none>                                 16.983 -1079.58
## - Diastolic.Pressure       1    0.1779 17.161 -1077.80
## - Physical.Activity.Level  1    0.2759 17.258 -1075.73
## - Daily.Steps              1    0.6754 17.658 -1067.43
## - BMI.Category             1    1.5375 18.520 -1050.12
## - Sleep.Duration           1    1.7026 18.685 -1046.90
## - Gender                   1    3.9128 20.895 -1006.32
## - Age                      1    7.0850 24.068  -955.01
## - Occupation               6   10.4617 27.444  -917.36
## - Stress.Level             1   25.5119 42.494  -748.65
##
## Step:  AIC=-1079.96
## Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Daily.Steps + Diastolic.Pressure +
##     BMI.Category
##
##                           Df Sum of Sq    RSS      AIC
## <none>                                 17.059 -1079.96
## - Physical.Activity.Level  1    0.2274 17.286 -1077.15
## - Diastolic.Pressure       1    0.2383 17.297 -1076.92
## - Daily.Steps              1    0.7294 17.788 -1066.76
## - Sleep.Duration           1    1.7099 18.769 -1047.28
## - BMI.Category             1    2.1556 19.214 -1038.76
## - Gender                   1    4.2126 21.271 -1001.84
## - Occupation               6   10.4319 27.491  -918.74
## - Age                      1   12.2435 29.302  -885.58
## - Stress.Level             1   25.7039 42.763  -748.36
```

```
summary(step.model)
```

```
##
## Call:
## lm(formula = Quality.of.Sleep ~ Age + Gender + Occupation + Sleep.Duration +
##     Physical.Activity.Level + Stress.Level + Daily.Steps + Diastolic.Pressure +
##     BMI.Category, data = dataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11013 -0.08980  0.00291  0.07526  0.75796
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               6.518e+00  6.127e-01  10.639  < 2e-16 ***
## Age                       6.146e-02  3.889e-03  15.804  < 2e-16 ***
## GenderMale                6.168e-01  6.653e-02   9.270  < 2e-16 ***
## OccupationDoctor         -4.883e-01  7.958e-02  -6.135 2.31e-09 ***
## OccupationEngineer       -6.658e-01  7.731e-02  -8.612 2.54e-16 ***
## OccupationLawyer         -3.726e-01  8.811e-02  -4.229 3.01e-05 ***
## OccupationNurse          -2.404e-01  8.319e-02  -2.890 0.004091 **
## OccupationSalesperson    -1.012e+00  8.881e-02 -11.398  < 2e-16 ***
## OccupationTeacher        -5.219e-01  7.084e-02  -7.368 1.27e-12 ***
```

```
## Sleep.Duration           2.518e-01  4.264e-02   5.906 8.33e-09 ***
## Physical.Activity.Level -2.714e-03  1.260e-03  -2.154 0.031957 *
## Stress.Level            -4.118e-01  1.798e-02 -22.899  < 2e-16 ***
## Daily.Steps              5.447e-05  1.412e-05   3.857 0.000137 ***
## Diastolic.Pressure      -1.447e-02  6.563e-03  -2.205 0.028104 *
## BMI.CategoryOverweight  -5.028e-01  7.582e-02  -6.631 1.27e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2214 on 348 degrees of freedom
## Multiple R-squared:  0.9639, Adjusted R-squared:  0.9624
## F-statistic: 663.3 on 14 and 348 DF,  p-value: < 2.2e-16
```

Explore the causes of sleep disorder

```
# Multinomial Logistic Regression: predict the probablities of categorically dependent variable
library(nnet)
mlr <- multinom(Sleep.Disorder ~ Gender + Occupation + Stress.Level + Physical.Activity.Level, data = sl
```

```
## # weights:  33 (20 variable)
## initial  value 398.796261
## iter  10 value 171.335140
## iter  20 value 161.253616
## iter  30 value 160.840628
## iter  40 value 160.828716
## iter  50 value 160.828480
## iter  50 value 160.828479
## iter  50 value 160.828479
## final  value 160.828479
## converged
```

```
summary(mlr)
```

```
## Call:
## multinom(formula = Sleep.Disorder ~ Gender + Occupation + Stress.Level +
##     Physical.Activity.Level, data = sleep)
##
## Coefficients:
##             (Intercept) GenderMale OccupationDoctor OccupationEngineer
## None           4.022528  0.9443797         2.416758         0.03370696
## Sleep Apnea  -11.113794  0.9860764        11.511691         8.00808294
##             OccupationLawyer OccupationNurse OccupationSalesperson
## None                0.752600       0.2801652             -3.522776
## Sleep Apnea         9.700794      13.7622114              8.234753
##             OccupationTeacher Stress.Level Physical.Activity.Level
## None                -2.842024   -0.6633934              0.01167303
## Sleep Apnea          9.106827   -0.5311657              0.04986580
##
## Std. Errors:
##             (Intercept) GenderMale OccupationDoctor OccupationEngineer
```

```
## None           1.599184  0.8684353         1.123280          0.8800371
## Sleep Apnea    1.549875  1.4183762         1.002882          1.1409667
##             OccupationLawyer OccupationNurse OccupationSalesperson
## None                1.094199       0.9962363              1.164531
## Sleep Apnea         1.076318       1.2843494              1.133791
##             OccupationTeacher Stress.Level Physical.Activity.Level
## None                0.6727419    0.2153846              0.01517467
## Sleep Apnea         0.8348767    0.2351505              0.01792520
##
## Residual Deviance: 321.657
## AIC: 361.657
```

```
exp(summary(mlr)$coefficients)
```

```
##              (Intercept) GenderMale OccupationDoctor OccupationEngineer
## None        5.584211e+01   2.571218         11.20946           1.034281
## Sleep Apnea 1.490529e-05   2.680696      99876.60227        3005.150532
##             OccupationLawyer OccupationNurse OccupationSalesperson
## None                2.122511    1.323348e+00           2.95174e-02
## Sleep Apnea     16330.564520    9.480964e+05           3.76971e+03
##             OccupationTeacher Stress.Level Physical.Activity.Level
## None             5.830756e-02    0.5151004                1.011741
## Sleep Apnea      9.016639e+03    0.5879193                1.051130
```

```
step(mlr, direction = 'backward')
```

```
## Start:  AIC=361.66
## Sleep.Disorder ~ Gender + Occupation + Stress.Level + Physical.Activity.Level
##
## trying - Gender
## # weights:  30 (18 variable)
## initial  value 398.796261
## iter  10 value 167.747995
## iter  20 value 161.675816
## iter  30 value 161.449674
## final  value 161.448415
## converged
## trying - Occupation
## # weights:  15 (8 variable)
## initial  value 398.796261
## iter  10 value 258.003099
## final  value 256.318488
## converged
## trying - Stress.Level
## # weights:  30 (18 variable)
## initial  value 398.796261
## iter  10 value 172.070021
## iter  20 value 166.464686
## iter  30 value 166.181825
## final  value 166.174564
## converged
## trying - Physical.Activity.Level
## # weights:  30 (18 variable)
```

```
## initial  value 398.796261
## iter  10 value 176.280707
## iter  20 value 165.925305
## iter  30 value 165.544055
## iter  40 value 165.537870
## iter  40 value 165.537870
## iter  40 value 165.537869
## final  value 165.537869
## converged
##                            Df      AIC
## - Gender                   18 358.8968
## <none>                     20 361.6570
## - Physical.Activity.Level 18 367.0757
## - Stress.Level             18 368.3491
## - Occupation                8 528.6370
## # weights:  30 (18 variable)
## initial  value 398.796261
## iter  10 value 167.747995
## iter  20 value 161.675816
## iter  30 value 161.449674
## final  value 161.448415
## converged
##
## Step:  AIC=358.9
## Sleep.Disorder ~ Occupation + Stress.Level + Physical.Activity.Level
##
## trying - Occupation
## # weights:  12 (6 variable)
## initial  value 398.796261
## iter  10 value 300.789075
## final  value 300.786852
## converged
## trying - Stress.Level
## # weights:  27 (16 variable)
## initial  value 398.796261
## iter  10 value 168.899272
## iter  20 value 166.495760
## iter  30 value 166.439242
## final  value 166.438730
## converged
## trying - Physical.Activity.Level
## # weights:  27 (16 variable)
## initial  value 398.796261
## iter  10 value 172.254164
## iter  20 value 166.856492
## iter  30 value 166.654338
## final  value 166.653685
## converged
##                            Df      AIC
## <none>                     18 358.8968
## - Stress.Level             16 364.8775
## - Physical.Activity.Level 16 365.3074
## - Occupation                6 613.5737
```

```
## Call:
## multinom(formula = Sleep.Disorder ~ Occupation + Stress.Level +
##     Physical.Activity.Level, data = sleep)
##
## Coefficients:
##             (Intercept) OccupationDoctor OccupationEngineer OccupationLawyer
## None           3.057756         3.140508          0.6260287         1.492406
## Sleep Apnea  -13.373748        13.850805         10.2223280        12.068248
##             OccupationNurse OccupationSalesperson OccupationTeacher
## None            -0.04068091             -2.765415         -2.627927
## Sleep Apnea     15.00246982             10.563566         10.806488
##             Stress.Level Physical.Activity.Level
## None          -0.5435703              0.01863057
## Sleep Apnea   -0.4278327              0.05415218
##
## Residual Deviance: 322.8968
## AIC: 358.8968
```

Generalised Linear Mixed Model

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
m1 <- lmer(Quality.of.Sleep ~ Sleep.Duration + Stress.Level + (1|Occupation), data = sleep)
summary(m1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Quality.of.Sleep ~ Sleep.Duration + Stress.Level + (1 | Occupation)
##    Data: sleep
##
## REML criterion at convergence: 199.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.0854 -0.4395  0.1433  0.5595  2.4792
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  Occupation (Intercept) 0.10052  0.3170
##  Residual               0.09112  0.3019
## Number of obs: 363, groups:  Occupation, 7
##
## Fixed effects:
##                Estimate Std. Error t value
## (Intercept)     6.76664    0.47721  14.179
## Sleep.Duration  0.39383    0.04984   7.903
## Stress.Level   -0.42085    0.02251 -18.699
##
## Correlation of Fixed Effects:
```

```
##             (Intr) Slp.Dr
## Sleep.Durtn -0.959
## Stress.Levl -0.895  0.869
```

```
m2 <- lmer(Quality.of.Sleep ~ Sleep.Duration + (1|Occupation), data=sleep)
summary(m2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Quality.of.Sleep ~ Sleep.Duration + (1 | Occupation)
##    Data: sleep
##
## REML criterion at convergence: 438.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3894 -0.4401  0.0525  0.6185  2.3162
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  Occupation (Intercept) 0.1418   0.3766
##  Residual               0.1800   0.4242
## Number of obs: 363, groups:  Occupation, 7
##
## Fixed effects:
##                Estimate Std. Error t value
## (Intercept)    -1.22745    0.28487  -4.309
## Sleep.Duration  1.20434    0.03463  34.776
##
## Correlation of Fixed Effects:
##             (Intr)
## Sleep.Durtn -0.862
```