The Relationship Between Exercise Time and Weight Loss

Bilal Mozaffar, Risheng Li, Thomas Janes

2/6/2021

```r
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.6.2
```

```r
library(ggplot2)
library(dplyr)
```

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

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

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

```r
library(reshape2)
```

```
##
## Attaching package: 'reshape2'

## The following object is masked from 'package:tidyr':
##
##     smiths
```

Abstract

The purpose of this data analysis is to apply statistical models to health data for a call center. While there were many data points missing for weight gain and the total metabolic minutes, many of these could be found using the other metrics available. The health data was processed to provide as many viable data points as possible without compromising the integrity of the analysis. Two models were constructed: a linear model to examine the relationship between total metabolic minutes and weight gain and a logistic regression model to examine the relationship between shift time and weight gain.

Introduction

We have been given several health measurements of employees working at a call center. Over eight months, the metrics have tracked information including work shift time, exercise time, weight gained (binary yes/no measurement), amount of weight gained, total metabolic minutes, and more. We hope to figure out whether certain factors, specifically total metabolic minutes and shift time, play a role in employees' weight gain.

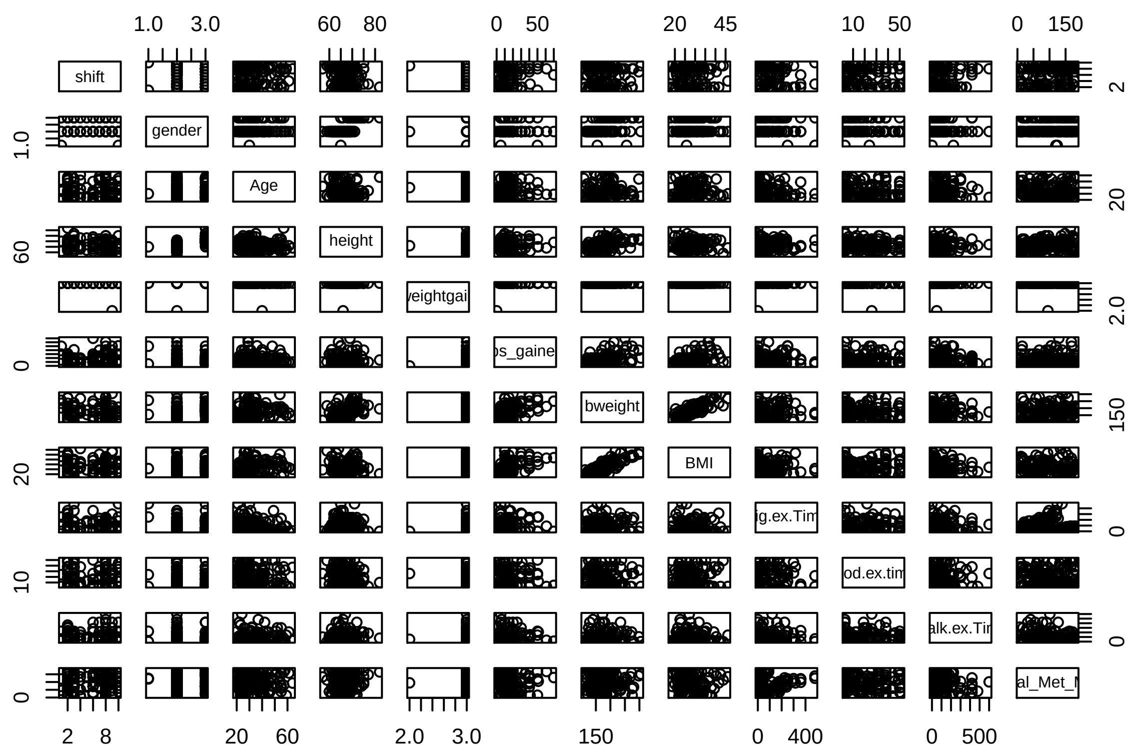
The Data

The data used in this analysis was provided by the call center. The data consists of 392 observations and 83 features, with variables such as member information, body measurements, exercise time, and weight situation. There are thirteen variables that we will use for this analysis, including shift time, gender, age, height, weight gain, pounds gained, body weight, body mass index, vigorous exercise time, moderate exercise time, walking exercise time, and total metabolic minutes. It is important to note that there were many missing values for total metabolic minutes, but these could be filled in using the other variables available. There were also missing values for pounds gained, but no features were available to account for this so these observations were removed.

We load and subset the raw data into a new dataset that contains the columns that we will use for our analysis. After the initial examination of the data, we renamed several variables by eliminating white spaces in order to improve the easiness of analysis. Using the formula provided, Total_met_min = 8 * Vig_ex_time + 4 * Mod_ex_time + 3.3 * Walk_ex_time, we filled in the missing values of variable Total_Met_Min.

```
### Already a numeric variable? No need to change
dat1$Total_Met_Min = as.numeric(dat1$Total_Met_Min)
```

```
pairs(dat1)
```

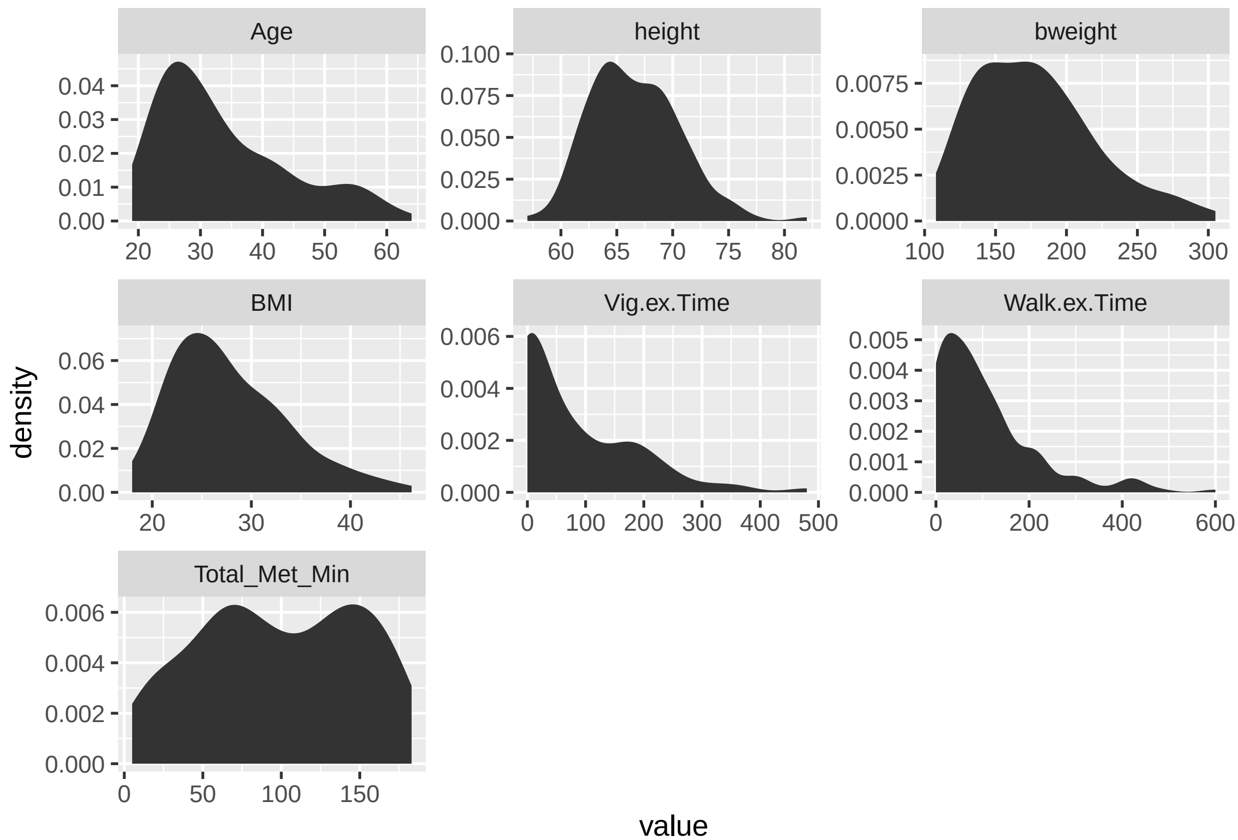



```
dat1.try <- dat1[c(3:4, 7:12)]
dat1.try <- melt(dat1.try)
```

```
## Using Mod.ex.time as id variables
```

```r
ggplot(data = dat1.try, aes(x = value)) + stat_density() + facet_wrap(~variable, scales = 'free')
```

```
## Warning: Removed 82 rows containing non-finite values (stat_density).
```



We see that Age and BMI are right-skewed with nonzero values, so we will do a log transformation on these variables.

```r
dat1$Age <- log(dat1$Age)
dat1$BMI <- log(dat1$BMI)
```

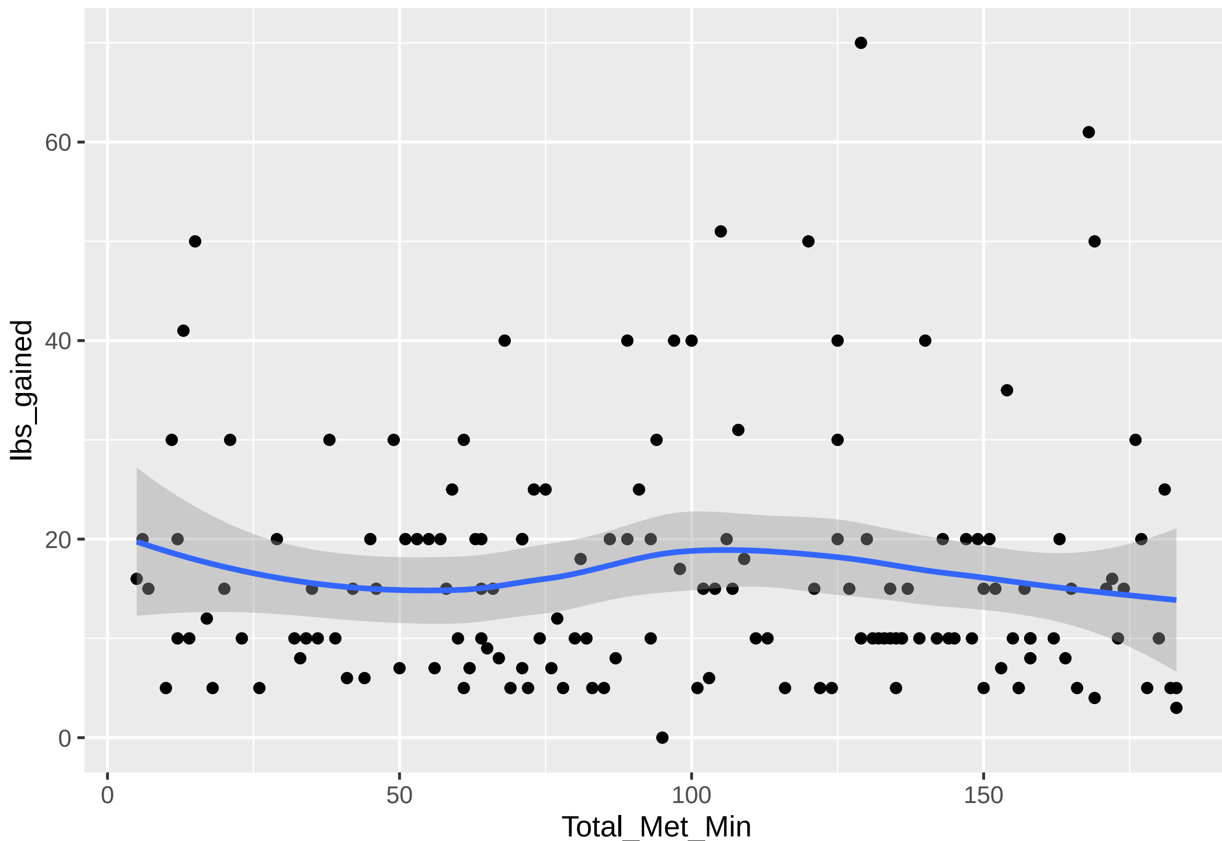
```r
lm = lm(lbs_gained ~ Total_Met_Min, data = dat1)

summary(lm)
```

```
##
## Call:
## lm(formula = lbs_gained ~ Total_Met_Min, data = dat1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -16.406  -7.288  -1.913   3.543  53.824
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)   17.048084   2.115141   8.060 2.03e-13 ***
## Total_Met_Min -0.006761   0.019050  -0.355    0.723
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.05 on 153 degrees of freedom
## Multiple R-squared:  0.0008225,  Adjusted R-squared:  -0.005708
## F-statistic: 0.126 on 1 and 153 DF,  p-value: 0.7232
```

```r
ggplot(data = dat1, aes(x = Total_Met_Min, y = lbs_gained)) + geom_point() + geom_smooth()
```

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




```r
i = 1
for (obs in dat1$lbs_gained) {
  if (obs > 45) {
    dat1 = dat1[-i,]
  } else {
    i = i + 1
  }
}
```

```r
i = 1
for (obs in dat1$Total_Met_Min) {
```

```r
  if (obs >= 5000) {
    dat1 = dat1[-i,]
  } else {
    i = i + 1
  }
}
```

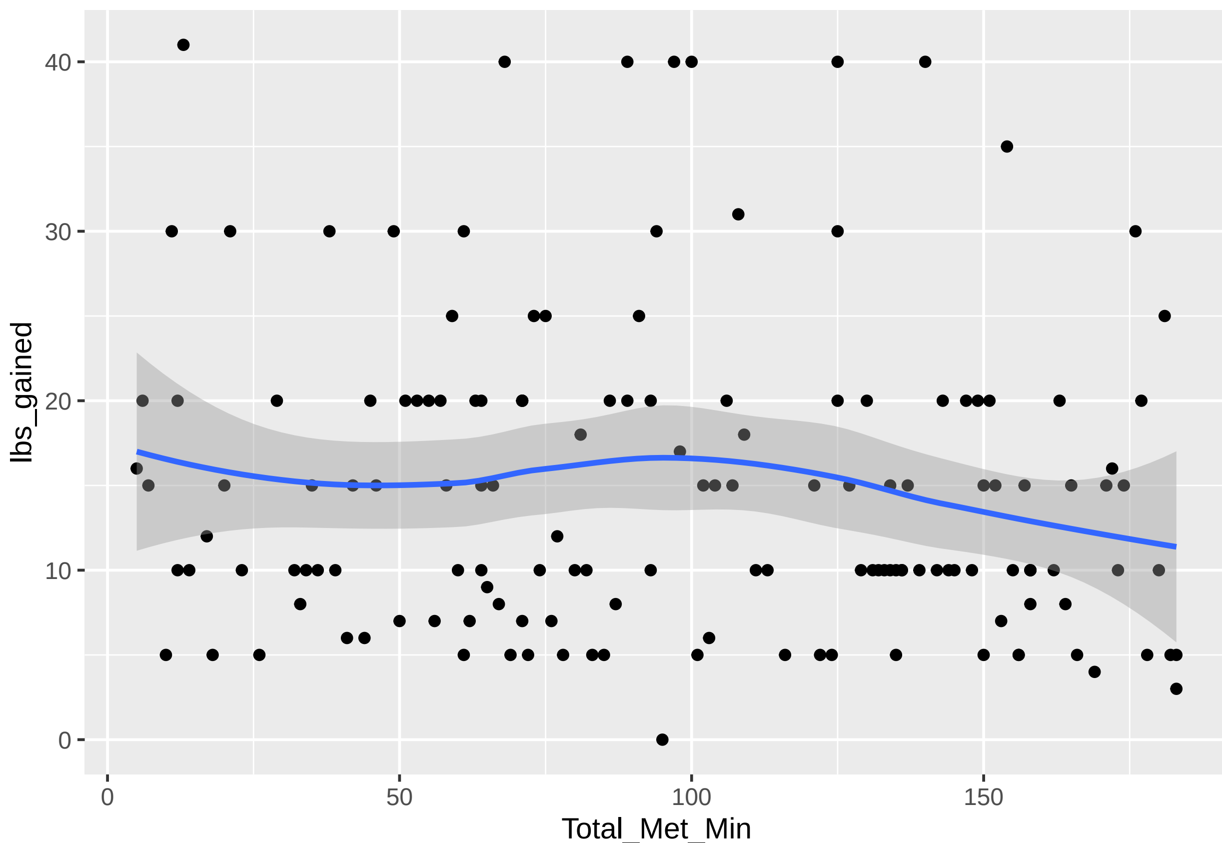
```r
lm = lm(lbs_gained ~ Total_Met_Min, data = dat1)

summary(lm)
```

```
##
## Call:
## lm(formula = lbs_gained ~ Total_Met_Min, data = dat1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -14.875  -6.588  -3.225   4.470  26.077
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   16.88414    1.62851  10.368   <2e-16 ***
## Total_Met_Min -0.02115    0.01477  -1.432    0.154
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.127 on 147 degrees of freedom
## Multiple R-squared:  0.01376,    Adjusted R-squared:  0.007054
## F-statistic: 2.051 on 1 and 147 DF,  p-value: 0.1542
```

P-value got bigger

```r
ggplot(data = dat1, aes(x = Total_Met_Min, y = lbs_gained)) + geom_point() + geom_smooth()
```

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



Risheng: replace lbs_gained value with 0 if weightgain = 0 and drop the rows that both weightgain and lbs_gained are missing. Also calculate the beginning weight by subtracting lbs_gain from body weight. Calculate the change in BMI using begeinning weight and height.

```
dat1$lbs_gained[dat1$weightgain == "No"] <- 0

dat1$beg_weight <- dat1$bweight - dat1$lbs_gained

dat1$Beg_BMI <- (dat1$beg_weight/(dat1$height^2))*703
dat1$BMIChange <- dat1$BMI - dat1$Beg_BMI

dat1 <- dat1 %>%
  filter_at(vars(weightgain, lbs_gained), any_vars(complete.cases(.)))
```

By turning weightgain into a dummy variable and running a logistic regression, we can measure if other variables have influence on weightgain. First, we factored the two categorical variables and built a model with variables include shift, gender, Age, height and Beg_BMI and Total_Met_Min. Because the change in BMI is very minimal, we decided to use the beginning BMI as one of the regressors. The half normal plot shows that there is no obvious outliers in the model:

```
dat1$weightgain.b <- ifelse(dat1$weightgain == 'Yes', 1, 0)

dat1
```

```
##     shift gender      Age height weightgain lbs_gained bweight      BMI
## 1     8am Female 3.258097     62        Yes         18     124 3.121483
## 2     9am Female 3.295837     65        Yes          6     151 3.223664
## 3     9am   Male 3.688879     69        Yes         10     180 3.280159
## 4     7am   Male 3.951244     72        Yes         20     190 3.249211
## 5     7am Female 4.060443     62        Yes         20      NA       NA
## 6   other Female 3.135494     NA        Yes          5      NA       NA
## 7     8am Female 3.496508     62        Yes         15     155 3.344627
## 8     8am   Male 3.401197     75        Yes          8     160 2.995732
## 9     8am Female 4.077537     63        Yes         10     150 3.279783
## 10    8am Female 3.713572     68        Yes         10     205 3.439456
## 11    8am Female 3.135494     64        Yes         10      NA       NA
## 12    8am   Male 3.091042     72        Yes         20     235 3.461665
## 13    7am   Male 3.988984     72        Yes         40     298 3.699077
## 14    8am   Male 3.988984     69        Yes         20     185 3.307619
## 15   11am Female 3.737670     66        Yes         20     209 3.518388
## 16    7am   Male 3.713572     70        Yes         10     204 3.376563
## 17  other Female 3.737670     63        Yes          5     129 3.128951
## 18    8am Female 3.737670     57        Yes         10     113 3.196630
## 19    8am Female       NA     NA        Yes          3      NA       NA
## 20    8am Female 3.828641     64        Yes         15     140 3.179303
## 21    9am Female 3.688879     66         No          0      NA       NA
## 22    8am   Male 3.401197     75        Yes         40     225 3.336481
## 23    9am Female 3.218876     60        Yes         30      NA       NA
## 24    9am Female 3.401197     61        Yes         15     135 3.239071
## 25    9am Female 3.828641     62        Yes         20     120 3.088767
## 26    8am Female 3.295837     64        Yes          7      NA       NA
## 27   10am Female 3.258097     62        Yes         31     170 3.436886
## 28   10am   Male 3.135494     69        Yes          5     170 3.222868
## 29   11am   Male 4.007333     65        Yes          5     108 2.888704
## 30    9am Female 3.091042     64        Yes          5     120 3.025291
## 31    8am   Male 3.218876     71        Yes         20      NA       NA
## 32    8am Female 3.091042     64        Yes          7     132 3.120601
## 33   10am   Male 3.850148     68        Yes         25     200 3.414772
## 34   10am Female 3.044522     64        Yes          5     130 3.105035
## 35    8am Female 3.178054     67        Yes          5     173 3.299165
## 36        Female 3.526361     61        Yes         30     202 3.641788
## 37  other Female 3.332205     65        Yes          5     135 3.111736
## 38    9am Female 3.178054     63        Yes         10      NA       NA
## 39   12pm Female 3.401197     64        Yes         10     269 3.832330
## 40   11am Female 3.258097     63        Yes          8     132 3.151881
## 41   11am   Male 3.178054     69        Yes          8     150 3.097837
## 42    2pm Female 3.044522     63        Yes         15     118 3.039749
## 43    1pm Female 3.135494     67        Yes          5     155 3.189241
## 44   11am Female 3.688879     61        Yes         41      NA       NA
## 45    2pm Female 3.401197     66        Yes         10     140 3.117507
## 46    2pm Female 3.465736     66        Yes         10      NA       NA
## 47  other Female 3.332205     66        Yes         10      NA       NA
## 48    2pm   Male       NA     NA        Yes         10     173       NA
## 49    9am   Male       NA     NA        Yes         25     180       NA
## 50    8am Female 3.332205     69        Yes         15     180 3.280159
## 51    2pm   Male       NA     72        Yes         30     250 3.523415
## 52    8am Female 3.806662     62        Yes         18     167 3.419037
## 53   11am Female 3.218876     67        Yes         20     155 3.189241
```

```
## 54    9am Female 3.295837     59          Yes           10      217 3.780090
## 55  other Female 4.007333     60          Yes           30      170 3.502550
## 56    8am Female 3.433987     63          Yes            7       NA       NA
## 57    8am Female 3.496508     64          Yes           10       NA       NA
## 58    8am Female 3.295837     67          Yes           20      145 3.122805
## 59    2pm Female 3.178054     65          Yes           17       NA       NA
## 60    8am Female 3.295837     69          Yes           15       NA       NA
## 61    1pm Female 3.688879     62          Yes            5       NA       NA
## 62   11am Female 3.583519     61          Yes           10       NA       NA
## 63    9am Female 3.401197     61          Yes            5      139 3.268047
## 64   10am Female 3.367296     63          Yes           30      235 3.728581
## 65    7am Female 3.218876     66          Yes           10      128 3.028199
## 66    8am   Male 3.258097     69          Yes           20      220 3.480625
## 67    8am   Male 3.433987     72          Yes           40      250 3.523415
## 68   12pm   Male 3.401197     73          Yes            5      200 3.272606
## 69    8am   Male       NA     NA          Yes            8      194       NA
## 70    8am Female 3.806662     68          Yes           20      200 3.414772
## 71    2pm Female       NA     61          Yes           20      135 3.239071
## 72   10am Female 3.367296     69          Yes            8      125 2.915606
## 73   10am   Male 3.526361     70          Yes           15      275 3.675034
## 74  other Female 3.931826     68          Yes           10      207 3.449035
## 75   11am   Male 3.258097     69          Yes            5      150 3.097837
## 76    1pm   Male 3.218876     73          Yes           10      185 3.194993
## 77   12pm Female 4.110874     66          Yes           10      165 3.282038
## 78   11am Female       NA     71          Yes           20       NA       NA
## 79   12pm Female 3.178054     66          Yes           10      170 3.312002
## 80    8am Female       NA     68          Yes           20       NA       NA
## 81          Male 3.610918     70          Yes           15      235 3.518091
## 82   10am   Male 4.043051     82          Yes           15      195 3.015045
## 83    8am Female 4.043051     67          Yes           10      188 3.382354
## 84   10am Female 4.007333     62          Yes           12      185 3.521348
## 85    9am Female 3.806662     68          Yes           10      165 3.222469
## 86    9am Female 3.610918     64          Yes            4      150 3.248046
## 87   12pm Female 3.637586     64          Yes           20      180 3.430433
## 88   10am Female 3.218876     67          Yes           10      178 3.327910
## 89    9am Female 3.465736     NA          Yes            5      135       NA
## 90    8am   Male 3.465736     71          Yes           15      187 3.261169
## 91    7am   Male 3.496508     68          Yes           20      180 3.309448
## 92   11am Female 3.761200     69          Yes           10      238 3.559340
## 93    8am Female 3.555348     64          Yes           30       NA       NA
## 94   10am Female 3.295837     62          Yes           20       NA       NA
## 95    8am   Male 3.367296     70          Yes           20      188 3.294725
## 96   10am Female 3.218876     65          Yes            5      115 2.951258
## 97    8am Female 4.007333     67          Yes           30       NA       NA
## 98   11am Female 3.178054     68          Yes           10      140 3.057768
## 99    7am Female 3.688879     68          Yes           25       NA       NA
## 100   7am Female 3.091042     64          Yes           15      162 3.325036
## 101   7am Female 3.258097     62          Yes           15      138 3.228430
## 102   2pm Female 2.995732     67          Yes            7      120 2.933325
## 103   2pm Female 3.912023     68          Yes           20       NA       NA
## 104  10am Female 3.951244     64          Yes           12      208 3.575151
## 105  10am Female 3.091042     62          Yes           20      150 3.311637
## 106  10am Female 3.091042     64          Yes           20      170 3.373484
## 107   8am Female 3.610918     66          Yes            6      140 3.117507
```

```
## 108  11am   Male 3.465736     75         Yes         16     185 3.140698
## 109   8am   Male 3.218876     70         Yes         15     170 3.194173
## 110  12pm Female 3.951244     63         Yes         16     185 3.489513
## 111  10am   Male 4.007333     72         Yes         10     210 3.349202
## 112  10am Female 3.332205     67         Yes          5      NA       NA
## 113   8am Female       NA     64         Yes          7     145 3.214466
## 114   8am Female 3.555348     70         Yes         20     240 3.538928
## 115  11am   Male 3.688879     70         Yes          5     215 3.429137
## 116   8am Female 3.465736     65         Yes         10     157 3.262701
## 117   7am Female 3.737670     67         Yes         15     201 3.449352
## 118   8am Female 4.158883     64         Yes         10     145 3.214466
## 119   8am Female 3.091042     66         Yes          9     142 3.132010
## 120   8am Female 3.178054     68         Yes         40     170 3.252311
## 121  11am   Male 3.891820     66         Yes         15     138 3.103240
## 122   8am Female 3.713572     62         Yes         15     187 3.532226
## 123   8am Female 3.688879     64         Yes         35     130 3.105035
## 124  10am Female 3.465736     69         Yes         15     195 3.360028
## 125  10am Female 3.135494     65         Yes          6      NA       NA
## 126  10am   Male 3.555348     70         Yes         10     273 3.667911
## 127  11am Female 3.135494     68         Yes          5     145 3.092859
## 128   7am   Male 3.737670     69         Yes         10     210 3.434310
## 129   8am Female 3.218876     64         Yes         20     133 3.128075
## 130 other   Male 3.218876     75         Yes         15     175 3.085116
## 131   9am Female 3.433987     70         Yes         40     215 3.429137
## 132   8am Female 2.944439     69         Yes         15     145 3.063858
## 133  10am Female 3.433987     NA         Yes          5     185       NA
## 134   9am Female 3.496508     65         Yes         10     165 3.312366
## 135   8am Female 3.295837     63         Yes         30     183 3.478467
## 136  11am Female 3.178054     66         Yes         10     140 3.117507
## 137   9am Female 3.258097     64         Yes          7     125 3.065725
## 138   2pm   Male 3.555348     74         Yes         25     230 3.385407
## 139  10am Female 3.135494     70         Yes         20     158 3.121042
## 140   8am Female 3.526361     61         Yes          5     123 3.145875
## 141   9am Female 3.258097     71         Yes         20     305 3.750210
## 142   8am   Male 3.258097     72         Yes         40     280 3.636796
## 143  10am   Male 3.091042     71         Yes         25     215 3.400530
## 144  11am   Male 3.044522     71         Yes         10     168 3.154017
## 145   8am Female 3.332205     65         Yes         10      NA       NA
## 146  10am Female 3.465736     67         Yes         15     165 3.251924
## 147  11am Female 3.295837     66         Yes         15      NA       NA
## 148 other        3.401197     65         Yes          5     155 3.249987
## 149  11am   Male 3.555348     77         Yes         10     220 3.261552
##     Vig.ex.Time Mod.ex.time Walk.ex.Time Total_Met_Min beg_weight  Beg_BMI
## 1           180         160          100            81        106 19.38554
## 2            40           0            0           103        145 24.12663
## 3            40          20           10           129        170 25.10187
## 4           180         180            0            71        170 23.05363
## 5            90          40            0           177         NA       NA
## 6             0         150           20           156         NA       NA
## 7            60          30           80           174        140 25.60354
## 8             0          40           30            87        152 18.99662
## 9            90           0          140            14        140 24.79718
## 10           50         160          125            36        195 29.64641
## 11            0          75           45           132         NA       NA
```

```
## 12          45         0        420          55        215 29.15606
## 13           0        40        105         140        258 34.98727
## 14           0         0         10         106        165 24.36358
## 15          90        60         60          12        189 30.50207
## 16         225       360         70         111        194 27.83306
## 17          90        60        420          78        124 21.96321
## 18           0        30        100         133        103 22.28655
## 19           0         0         30         183         NA       NA
## 20         180         0          0          35        125 21.45386
## 21           0        30         50          95         NA       NA
## 22         210       210        105          97        185 23.12089
## 23           0       120        120         176         NA       NA
## 24           0         0         60          64        120 22.67132
## 25           0         0         49          45        100 18.28824
## 26          20         0         20          76         NA       NA
## 27         300        80        189         108        139 25.42066
## 28         100        60         90          26        165 24.36358
## 29          75         0         20         156        103 17.13822
## 30         300       300         60         122        115 19.73755
## 31          60         0         90         163         NA       NA
## 32         160        60         80          56        125 21.45386
## 33           0         0         68          75        175 26.60575
## 34         180        30        100          61        125 21.45386
## 35         120        30        600         101        168 26.30965
## 36           0         0         35          11        172 32.49557
## 37          30        60        100         166        130 21.63077
## 38           0        90        150         173         NA       NA
## 39           0         0        140         134        259 44.45239
## 40         180        91         70          67        124 21.96321
## 41          60       180         60          33        142 20.96744
## 42           0      1680          0         150        103 18.24364
## 43         150       150         90          69        150 23.49076
## 44          60        50        150          13         NA       NA
## 45           0       120          0         136        130 20.98026
## 46           0         0         20         155         NA       NA
## 47         180        75         35          60         NA       NA
## 48           0         0        420          32        163       NA
## 49         160         0        420          91        155       NA
## 50          80         0         60         171        165 24.36358
## 51         160         0          0          21        220 29.83410
## 52           0        20         80         109        149 27.24948
## 53           0        60         60         130        135 21.14168
## 54           0         9          0         113        207 41.80437
## 55          60       210         60          38        140 27.33889
## 56          80         0          0         153         NA       NA
## 57           0       135          0         144         NA       NA
## 58           0         0        120         125        125 19.57563
## 59         360         0          0          98         NA       NA
## 60          15        10          0          42         NA       NA
## 61          16        16         90         135         NA       NA
## 62           0         0        135         131         NA       NA
## 63          90        40         30         182        134 25.31631
## 64         315        60         16          94        205 36.31015
## 65           0         0        210         158        118 19.04362
```

```
## 66          180        180          0          71        200 29.53161
## 67          200        240        140         100        210 28.47801
## 68           60         60         60         178        195 25.72434
## 69            0          0        210         158        186       NA
## 70            0        100         60         149        180 27.36592
## 71           40         30        280          29        115 21.72669
## 72           90          9          9         164        117 17.27599
## 73          200         60          0          58        260 37.30204
## 74            0          0        225         162        197 29.95048
## 75            0         30         30          72        145 21.41042
## 76          180        150         50          74        175 23.08594
## 77            0          0        210         158        155 25.01492
## 78            0         20         60          93         NA       NA
## 79            0        240         60          12        160 25.82185
## 80           90         60         20           6         NA       NA
## 81          120        120         60          46        220 31.56327
## 82            0        140         75         165        180 18.81916
## 83            0         60          0          80        178 27.87570
## 84          225          0        140          77        173 31.63866
## 85           60         30        210          23        155 23.56510
## 86            0         90        140         169        146 25.05811
## 87          120        120         80          51        160 27.46094
## 88            0         45         90         135        168 26.30965
## 89            0          0         30         183        130       NA
## 90          135         30         24          20        172 23.98651
## 91          180        180          0          71        160 24.32526
## 92          165         90        225          82        228 33.66604
## 93            0          0        120         125         NA       NA
## 94            0          0         60          64         NA       NA
## 95            0         60         90         143        168 24.10286
## 96            0         15          0         150        110 18.30296
## 97            0        225        300          61         NA       NA
## 98            0          0         60          64        130 19.76427
## 99          120        105        140          59         NA       NA
## 100         225        210        420         127        147 25.22974
## 101          60          0         60         157        123 22.49454
## 102         180        180          0          71        113 17.69637
## 103           0         90         60         147         NA       NA
## 104           0         60        300          17        196 33.63965
## 105           0        360        120          57        130 23.77471
## 106           0          0         80          89        150 25.74463
## 107           0          0        480          41        134 21.62580
## 108          12         16        210         172        169 21.12124
## 109           0        135         20         152        155 22.23776
## 110           0        140        140           5        169 29.93374
## 111          60        120          0         180        200 27.12191
## 112         240         40        120          83         NA       NA
## 113         120        100        100          50        138 23.68506
## 114         240        120         40          86        220 31.56327
## 115           0         30         40          85        210 30.12857
## 116          60          0          0         136        147 24.45941
## 117           0          0        100         107        186 29.12854
## 118           8         56         20          93        135 23.17017
## 119         200         40         75          65        133 21.46442
```

```
## 120          0          0          120          125          130 19.76427
## 121        360        120          120          121          123 19.85055
## 122         45        240          210           66          172 31.45578
## 123          0         40          150          154           95 16.30493
## 124        180        210          300          104          180 26.57845
## 125         30         45          360           44           NA       NA
## 126          0          0          210          158          263 37.73245
## 127         50         75          135           10          140 21.28460
## 128         30         60           20          145          200 29.53161
## 129        135         90          160           63          113 19.39429
## 130         30         45           20          137          160 19.99644
## 131          0          0           80           89          175 25.10714
## 132        240        120          210          102          130 19.19555
## 133        360          0          240          116          180       NA
## 134          0          0          160          142          155 25.79053
## 135          0         30           15           49          153 27.09977
## 136        180         20            0           39          130 20.98026
## 137          0        240          300           62          118 20.25244
## 138        135        105          210           73          205 26.31757
## 139        180         20           60           53          138 19.79878
## 140        120         40           40           18          118 22.29347
## 141         75          0            0          151          285 39.74509
## 142        240         30            0           68          240 32.54630
## 143         60         90           40          181          190 26.49673
## 144        480        240           60          139          158 22.03412
## 145         90         90          105           34           NA       NA
## 146          0          0          315            7          150 23.49076
## 147          0          0          140          134           NA       NA
## 148        480          0            0          124          150 24.95858
## 149          0          0          180          148          210 24.89965
##     BMIChange weightgain.b
## 1   -16.26405            1
## 2   -20.90296            1
## 3   -21.82171            1
## 4   -19.80442            1
## 5          NA            1
## 6          NA            1
## 7   -22.25891            1
## 8   -16.00089            1
## 9   -21.51740            1
## 10  -26.20695            1
## 11         NA            1
## 12  -25.69439            1
## 13  -31.28819            1
## 14  -21.05596            1
## 15  -26.98368            1
## 16  -24.45650            1
## 17  -18.83426            1
## 18  -19.08992            1
## 19         NA            1
## 20  -18.27455            1
## 21         NA            0
## 22  -19.78441            1
## 23         NA            1
```

```
## 24  -19.43225          1
## 25  -15.19947          1
## 26         NA          1
## 27  -21.98377          1
## 28  -21.14071          1
## 29  -14.24952          1
## 30  -16.71226          1
## 31         NA          1
## 32  -18.33326          1
## 33  -23.19098          1
## 34  -18.34882          1
## 35  -23.01048          1
## 36  -28.85378          1
## 37  -18.51903          1
## 38         NA          1
## 39  -40.62006          1
## 40  -18.81133          1
## 41  -17.86961          1
## 42  -15.20389          1
## 43  -20.30151          1
## 44         NA          1
## 45  -17.86275          1
## 46         NA          1
## 47         NA          1
## 48         NA          1
## 49         NA          1
## 50  -21.08342          1
## 51  -26.31069          1
## 52  -23.83044          1
## 53  -17.95244          1
## 54  -38.02428          1
## 55  -23.83634          1
## 56         NA          1
## 57         NA          1
## 58  -16.45282          1
## 59         NA          1
## 60         NA          1
## 61         NA          1
## 62         NA          1
## 63  -22.04827          1
## 64  -32.58157          1
## 65  -16.01542          1
## 66  -26.05099          1
## 67  -24.95459          1
## 68  -22.45173          1
## 69         NA          1
## 70  -23.95115          1
## 71  -18.48762          1
## 72  -14.36039          1
## 73  -33.62701          1
## 74  -26.50144          1
## 75  -18.31258          1
## 76  -19.89095          1
## 77  -21.73288          1
```

```
## 78         NA            1
## 79  -22.50985            1
## 80         NA            1
## 81  -28.04517            1
## 82  -15.80411            1
## 83  -24.49334            1
## 84  -28.11731            1
## 85  -20.34263            1
## 86  -21.81006            1
## 87  -24.03050            1
## 88  -22.98174            1
## 89         NA            1
## 90  -20.72534            1
## 91  -21.01581            1
## 92  -30.10670            1
## 93         NA            1
## 94         NA            1
## 95  -20.80813            1
## 96  -15.35170            1
## 97         NA            1
## 98  -16.70651            1
## 99         NA            1
## 100 -21.90470            1
## 101 -19.26611            1
## 102 -14.76304            1
## 103        NA            1
## 104 -30.06450            1
## 105 -20.46308            1
## 106 -22.37115            1
## 107 -18.50830            1
## 108 -17.98055            1
## 109 -19.04358            1
## 110 -26.44422            1
## 111 -23.77271            1
## 112        NA            1
## 113 -20.47059            1
## 114 -28.02434            1
## 115 -26.69943            1
## 116 -21.19671            1
## 117 -25.67918            1
## 118 -19.95570            1
## 119 -18.33241            1
## 120 -16.51196            1
## 121 -16.74731            1
## 122 -27.92355            1
## 123 -13.19990            1
## 124 -23.21842            1
## 125        NA            1
## 126 -34.06454            1
## 127 -18.19174            1
## 128 -26.09730            1
## 129 -16.26621            1
## 130 -16.91133            1
## 131 -21.67801            1
```

```
## 132 -16.13169           1
## 133        NA           1
## 134 -22.47817           1
## 135 -23.62131           1
## 136 -17.86275           1
## 137 -17.18672           1
## 138 -22.93216           1
## 139 -16.67773           1
## 140 -19.14759           1
## 141 -35.99488           1
## 142 -28.90950           1
## 143 -23.09620           1
## 144 -18.88010           1
## 145        NA           1
## 146 -20.23883           1
## 147        NA           1
## 148 -21.70859           1
## 149 -21.63809           1
```

```r
head(dat)
```

```
##   Snumber Date_Started     dept         Job shift gender Age height head
## 1       1     4/6/2005 Training       Other   8am Female  30     66    3
## 2       2    4/10/2005 Training       Other   7am Female  31     65    0
## 3       3    4/11/2005 Training       Other   8am Female  26     62    0
## 4       4    4/11/2005 Training       Other   7am Female  34     62    5
## 5       5    4/11/2005       CS       Other   9am Female  27     65    0
## 6       6    4/14/2005      CFS Collections  11am   Male  21     74    0
##   neck rshoul relbow rwrist lback rleg rknee rfoot eyes uback lshould
## 1    5      5      0      0     5    0     0     0    3     5       5
## 2    5      0      0      0     4    0     4     0    0     0       0
## 3    3      3      0      3     1    0     0     0    4     3       3
## 4    5      4      2      3     0    0     2     1    4     0       0
## 5    0      0      0      0     1    0     0     0    2     0       0
## 6    0      0      0      1     1    0     0     0    1     0       0
##   lelbow lwrist butt lleg lknee lfoot ddis ddiscomfort     oftendis
## 1      0      0    0    0     0     0   NA     #VALUE! Occasionally
## 2      0      0    0    0     0     0   NA     #VALUE!  When Active
## 3      0      3    0    0     0     0  730         730  When Active
## 4      0      1    0    0     2     1  365         365 Occasionally
## 5      0      0    0    0     0     0   NA     #VALUE! Occasionally
## 6      1      0    0    0     0     0  182         182 Occasionally
##            discomorig             workplace1       Job1       tool
## 1 Gradually over time                Seating No Concern No Concern
## 2 Gradually over time                Reaches No Concern No Concern
## 3 Gradually over time Adjustability; Seating     Breaks
## 4 Gradually over time             No Concern No Concern No Concern
## 5 Gradually over time             No Concern    Methods No Concern
## 6 Gradually over time             No Concern     Breaks No Concern
##                     handling                              environment
## 1                    Lifting                              Temperature
## 2 Carrying; Lifting; Push/Pull        Temperature; Lighting; Ventilation
## 3        Lifting; Lift Assists        Ventilation; Lighting; Temperature
## 4            Carrying; Lifting                      Temperature; Noise
```

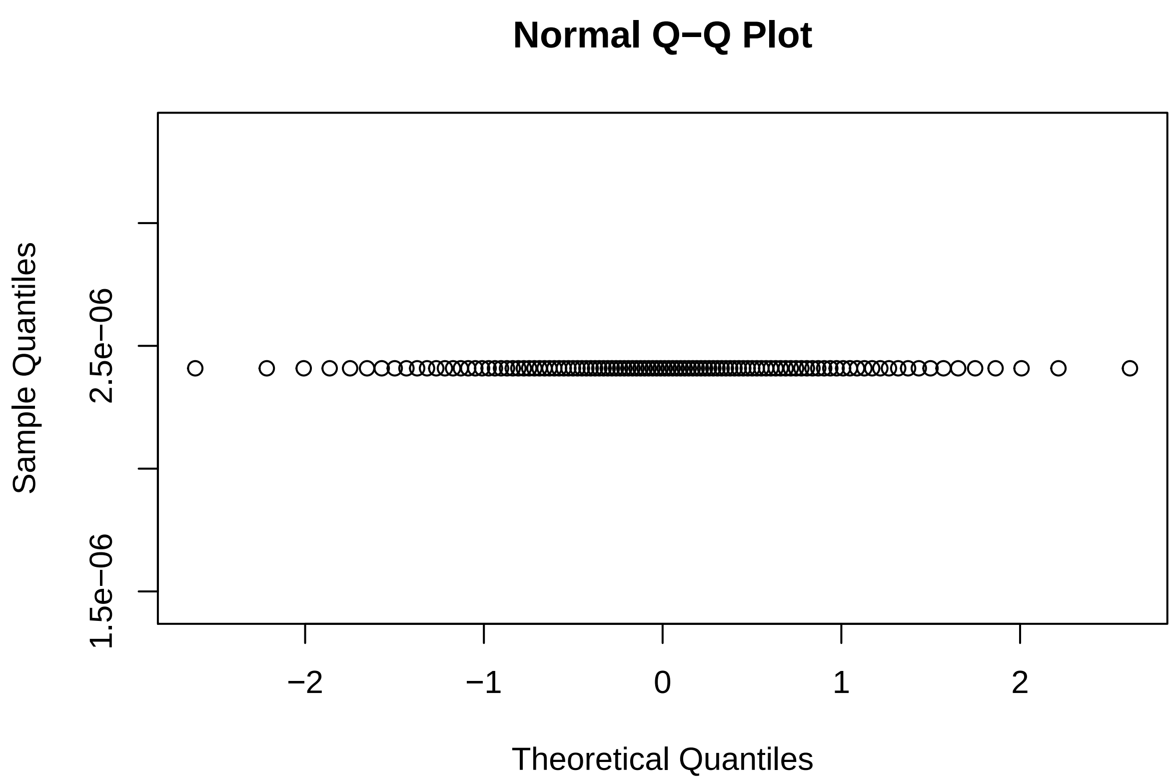
```
## 5                  No Concern                                  Lighting
## 6                  No Concern Noise; Temperature; Lighting; Ventilation
##   fatigue5min fatigue2 fatigue4 fatigue6 fatigue8 fatigue10 fatigue12
## 1           5        3        3        3        3         3         4
## 2           0        1        2        5        5        NA        NA
## 3           1        2        3        3        4         0         0
## 4           4        2        0        2        2        NA        NA
## 5           0        0        0        3        1         2         4
## 6           0        0        0        0        0         0         0
##   Days_to_less_fatigue injury                         injurypart
## 1                 14.0     No
## 2                 30.0    Yes
## 3                 60.0    Yes Neck; Hand/wrist; Lower Back; Foot
## 4                  1.5    Yes                               Neck
## 5                   NA     No
## 6                   NA     No
##        plabor prevposition weightgain lbs_gained pounds_gained
## 1 Very little           No         No         NA       #VALUE!
## 2    Moderate           No         No         NA       #VALUE!
## 3    Moderate           No        Yes         18            18
## 4    Moderate           No        Yes         NA       #VALUE!
## 5 Very little           No        Yes          6             6
## 6 Very little           No         No         NA       #VALUE!
##             weight_situation bweight exbreak expartic numstretch member
## 1    I am AT my ideal weight     135     Yes       No        0.0    Yes
## 2 I am ABOVE my ideal weight     135      No       No        0.0     No
## 3 I am ABOVE my ideal weight     124     Yes      Yes        0.0     No
## 4 I am ABOVE my ideal weight      NA     Yes      Yes        0.0     No
## 5 I am ABOVE my ideal weight     151     Yes      Yes        1.5     No
## 6 I am ABOVE my ideal weight     280     Yes      Yes        2.0     No
##   daysVHW Veg_Serv fruit_servings fried_servings vigorous7 vigur vigoroust
## 1     3.5       NA             NA             NA       5.5    NA        60
## 2     0.0        2              0              0       3.0    NA        30
## 3     0.0        3              5              2       3.0    NA        60
## 4      NA        2              2              0       3.0    NA        NA
## 5     0.0        2              3              2       2.0    NA        20
## 6     0.0        0              0              0       3.0    NA       120
##   vigtur moderate7 modur moderatet modtur walk7 walkur walkt walktur
## 1     NA         0    NA         0     NA     0     NA     0      NA
## 2     NA         0    NA         0     NA     3     NA    60      NA
## 3     NA         4    NA        40     NA     5     NA    20      NA
## 4      8         7    NA        NA      8     3     NA    NA       8
## 5     NA         0    NA         0     NA     0     NA     0      NA
## 6     NA         3    NA        60     NA     7     NA    90      NA
##   sitt7t situr Field66   BMI Total_ex_time Total_Ex_Time_Exclude
## 1     NA     8       1 21.79           330                    NA
## 2    300    NA       2 22.46           270                    NA
## 3    600    NA       3 22.68           440                    NA
## 4    240    NA       4    NA           104                    NA
## 5    540    NA       5 25.12            40                    NA
## 6    600    NA       6 35.95            NA                  1170
##   Vig.ex.Time Mod.ex.time Walk.ex.Time Total_Met_Min Total_Exclude_Met
## 1         330           0            0          2640                NA
## 2          90           0          180          1314                NA
```

```
## 3        180        160        100        2410                NA
## 4         24         56         24                            NA
## 5         40          0          0         320                NA
## 6        360        180        630                          5679
```

```
dat1$gender <- factor(dat1$gender)
dat1$shift <- factor(dat1$shift)


mod1 <- glm(weightgain.b ~ shift + gender + Age + height + Beg_BMI + Total_Met_Min, family = binomial,na
```

```
## Warning: glm.fit: algorithm did not converge
```

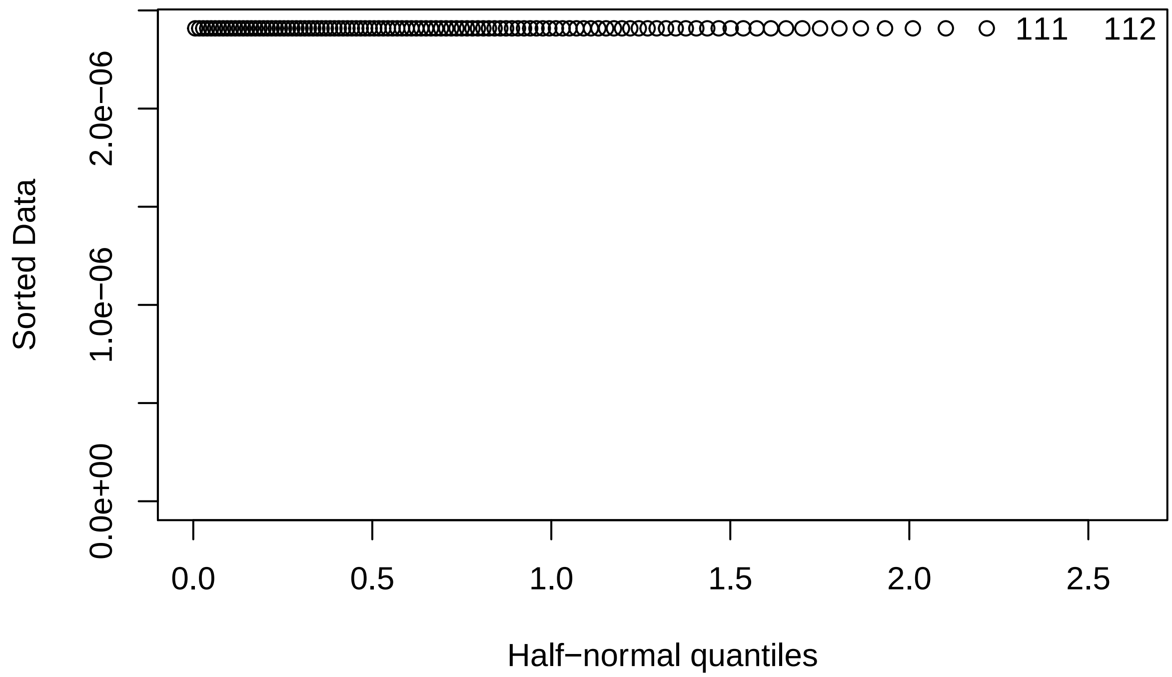
```
##qq plot
qqnorm(residuals(mod1))
```




```
##Outliers: no obvious outliers
library(faraway)
```

```
## Warning: package 'faraway' was built under R version 3.6.3
```

```
halfnorm(residuals(mod1))
```




```r
## test if shift should be removed from the model
mod0 <- glm(weightgain.b ~ gender + Age + height + Beg_BMI + Total_Met_Min, family = binomial,na.omit(da
```

```
## Warning: glm.fit: algorithm did not converge
```

```r
summary(mod1)
```

```
##
## Call:
## glm(formula = weightgain.b ~ shift + gender + Age + height +
##     Beg_BMI + Total_Met_Min, family = binomial, data = na.omit(dat1))
##
## Deviance Residuals:
##       Min         1Q     Median         3Q        Max
## 2.409e-06  2.409e-06  2.409e-06  2.409e-06  2.409e-06
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.657e+01  9.919e+05       0        1
## shift10am     -4.695e-06  2.728e+05       0        1
## shift11am     -2.580e-06  2.789e+05       0        1
## shift12pm      1.584e-07  2.974e+05       0        1
## shift1pm      -8.091e-07  3.662e+05       0        1
## shift2pm      -2.699e-06  3.245e+05       0        1
## shift7am      -5.463e-06  2.889e+05       0        1
```

```
## shift8am      -1.577e-06  2.691e+05       0        1
## shift9am      -1.728e-06  2.820e+05       0        1
## shiftother    -5.322e-07  3.117e+05       0        1
## genderFemale  -9.812e-08  3.916e+05       0        1
## genderMale     9.053e-07  4.036e+05       0        1
## Age           -1.096e-06  1.174e+05       0        1
## height        -1.187e-07  1.186e+04       0        1
## Beg_BMI       -1.285e-07  6.968e+03       0        1
## Total_Met_Min -1.341e-10  7.348e+02       0        1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 0.0000e+00  on 111  degrees of freedom
## Residual deviance: 6.4978e-10  on  96  degrees of freedom
## AIC: 32
##
## Number of Fisher Scoring iterations: 25
```

```
summary(mod0)
```

```
##
## Call:
## glm(formula = weightgain.b ~ gender + Age + height + Beg_BMI +
##     Total_Met_Min, family = binomial, data = na.omit(dat1))
##
## Deviance Residuals:
##       Min         1Q     Median         3Q        Max
## 2.409e-06  2.409e-06  2.409e-06  2.409e-06  2.409e-06
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.657e+01  9.572e+05       0        1
## genderFemale  -1.074e-08  3.593e+05       0        1
## genderMale    -8.714e-09  3.693e+05       0        1
## Age            3.807e-09  1.138e+05       0        1
## height        -3.932e-11  1.167e+04       0        1
## Beg_BMI        1.322e-10  6.518e+03       0        1
## Total_Met_Min -2.871e-11  6.991e+02       0        1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 0.0000e+00  on 111  degrees of freedom
## Residual deviance: 6.4978e-10  on 105  degrees of freedom
## AIC: 14
##
## Number of Fisher Scoring iterations: 25
```

```
anova(mod0, mod1, test = "Chi")
```

```
## Analysis of Deviance Table
##
## Model 1: weightgain.b ~ gender + Age + height + Beg_BMI + Total_Met_Min
## Model 2: weightgain.b ~ shift + gender + Age + height + Beg_BMI + Total_Met_Min
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1       105 6.4978e-10
## 2        96 6.4978e-10  9        0        1
```

```
##do not reject the null hypothesis that the reduced model is better(the model includes shift is less d
```

Solely from the model, we can see that only three variables have significant impact on the predictor: gender, Beg_BMI and Total_Met_Min. We perform a Chi-sqaure test to test if the model contains shift is better than the model without shifts. The null hypothesis is that the model without shift is a better model, and the alternative hypothesis is that the model without shift is not a better model. In order to compare two models, we omit the NA values in the dataset to make sure the number of cases used in each model is the same. The test statistic is is very large given a 95% confidence interval, therefore we fail to reject the null hypothesis that the reduced model is better, which means that the model contains shift is less favorable than the reduced model. The test result is shown below:

Analysis of Deviance Table

Model 1: weightgain.b ~ gender + Age + height + Beg_BMI + Total_Met_Min Model 2: weightgain.b ~ shift + gender + Age + height + Beg_BMI + Total_Met_Min Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1 202 236.06
2 194 228.44 8 7.6207 0.4714

Since the initial model shows that Total_Met_Min is a significant variable, we performed another analysis, using a forward stepwise selection to select the model with the most appropriate variables that produces the lowest AIC. The selected variables are exactly the same as our previous analysis, which are gender, Beg_BMI and Total_Met_Min, which reassures that total metabolic minutes do have an effect on weight gain and shift does not have an effect on weightgain. The results are shown below:

Call: glm(formula = weightgain.b ~ gender + Beg_BMI + Total_Met_Min, family = binomial, data = na.omit(dat1))

Coefficients: (Intercept) genderMale Beg_BMI Total_Met_Min
3.0281421 -0.7079166 -0.0615582 -0.0001601

Degrees of Freedom: 207 Total (i.e. Null); 204 Residual Null Deviance: 251.7 Residual Deviance: 237.4 AIC: 245.4

```
## Using stepwise to select the best model
mod00 <- glm(weightgain.b ~ 1, family = binomial,na.omit(dat1))
```

```
## Warning: glm.fit: algorithm did not converge
```

```
m.sw1 = step(mod00,
             scope=list(lower=mod00, upper=mod1),
             direction = "both", trace = F)
```

```
## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge
```

```
## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: algorithm did not converge
```

```
m.sw1$anova
```

```
##   Step Df Deviance Resid. Df   Resid. Dev AIC
## 1      NA       NA       111 6.497771e-10   2
```

```
m.sw1
```

```
##
## Call:  glm(formula = weightgain.b ~ 1, family = binomial, data = na.omit(dat1))
##
## Coefficients:
## (Intercept)
##       26.57
##
## Degrees of Freedom: 111 Total (i.e. Null);  111 Residual
## Null Deviance:       0
## Residual Deviance: 6.498e-10     AIC: 2
```

```
## The selected model with the lowest AIC contains only three variables: gender, BMI and Total_Met_Min,
```

The Relationship Between Total Metabolic Minutes and Weight Gain

Using various methods of analysis, we found that total metabolic minutes is a significant variable. The initial model demonstrated that there is a negative relationship between total metabolic minutes and weight gain, and this was corroborated by the subsequent forward stepwise model.

The Relationship Between Shift Time and Weight Gain

It does not appear that shift has significant impact on weight gain. From the Chi-square test, we found that the model without shift performs better than the model with shift. As mentioned, we failed to reject the null hypothesis that the reduced model is better, so we do not find it beneficial to include shift in our final model to predict weight gain.

Conclusion

Our initial goal was to see whether shift and total metabolic minutes have an impact on weight gain. Multiple forms of analysis indicate that shift is not a relevant factor in determining weight gain, while total metabolic minutes is. Total metabolic minutes has a slightly adverse relationship to weight gain. We conducted further analysis to go beyond the requested covariate relationships. For a final model selection, gender, Beg_BMI, and Total_Met_Min appear to be the most signficant and useful predictors of weight gain.