# RMproject2

Objective 1: Display the ability to perform EDA and build a logistic regression model for interpretation purposes.
Perform your multiple logistic regression analysis and provide interpretation of the regression coefficients including hypothesis testing, and confidence intervals. For simplicity sake, you do not need to include interactions with this model. Comment on the practical vs statistical significance of the deemed important factors.
note the difference between a model that is interpretable versus a model that is complex. (A model with a lot of predictors can still be interpretable) Interactions models shouldn't be used here as you should display your ability to interpret the regression coefficients. Effects plots may be used in addition too but not at the exclusions of coefficient interpretation.

Assessing risk factors and predicting if a woman with osteoperosis will have a bone fracture within the first year after joining the study. ?glow_bonemed for data description of variables.

```
#install.packages("aplore3")
library(aplore3)
## Warning: package 'aplore3' was built under R version 4.3.3
?glow_bonemed
## starting httpd help server ... done
library(ggplot2)
library(corrplot)
## corrplot 0.92 loaded
library(tidyverse)
## ── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
## ✓ dplyr     1.1.4     ✓ readr     2.1.5
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ lubridate 1.9.3     ✓ tibble    3.2.1
## ✓ purrr     1.0.2     ✓ tidyr     1.3.1
## ── Conflicts ────────────────────────────────────────
tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
library(GGally)
## Registered S3 method overwritten by 'GGally':
```

```
##   method from
##   +.gg   ggplot2
```

```
library(dplyr)


#EDA


data("glow_bonemed")
head(glow_bonemed)# see rows
```

```
##   sub_id site_id phy_id priorfrac age weight height      bmi premeno
momfrac
## 1      1       1     14        No  62   70.3    158 28.16055      No
No
## 2      2       4    284        No  65   87.1    160 34.02344      No
No
## 3      3       6    305       Yes  88   50.8    157 20.60936      No
Yes
## 4      4       6    309        No  82   62.1    160 24.25781      No
No
## 5      5       1     37        No  61   68.0    152 29.43213      No
No
## 6      6       5    299       Yes  67   68.0    161 26.23356      No
No
##   armassist smoke raterisk fracscore fracture bonemed bonemed_fu bonetreat
## 1        No    No     Same         1       No      No         No        No
## 2        No    No     Same         2       No      No         No        No
## 3       Yes    No     Less        11       No      No         No        No
## 4        No    No     Less         5       No      No         No        No
## 5        No    No     Same         1       No      No         No        No
## 6        No   Yes     Same         4       No      No         No        No
```

```
tail(glow_bonemed)
```

```
##     sub_id site_id phy_id priorfrac age weight height      bmi premeno
momfrac
## 495    495       2     70        No  87   64.4    158 25.79715      No
Yes
## 496    496       5    287       Yes  79   63.5    157 25.76169      No
Yes
## 497    497       5    296        No  64   48.1    149 21.66569      No
No
```

```
## 498    498       5    287       Yes  61   70.8    161 27.31376     Yes
No
## 499    499       3    181       Yes  81   77.6    153 33.14964      No
No
## 500    500       6    317        No  63   74.8    165 27.47475      No
No
##     armassist smoke raterisk fracscore fracture bonemed bonemed_fu
bonetreat
## 495       Yes    No     Same         9      Yes      No          No
No
## 496       Yes    No  Greater         8      Yes      No         Yes
No
## 497        No    No  Greater         2      Yes     Yes         Yes
Yes
## 498       Yes    No  Greater         4      Yes      No          No
No
## 499       Yes    No     Same         8      Yes      No          No
No
## 500        No    No     Less         1      Yes      No          No
No
```

```
View(glow_bonemed)
names(glow_bonemed)
```

```
##  [1] "sub_id"     "site_id"    "phy_id"     "priorfrac"  "age"
##  [6] "weight"     "height"     "bmi"        "premeno"    "momfrac"
## [11] "armassist"  "smoke"      "raterisk"   "fracscore"  "fracture"
## [16] "bonemed"    "bonemed_fu" "bonetreat"
```

```
summary(glow_bonemed) #summary stats
```

```
##      sub_id         site_id          phy_id       priorfrac      age
##  Min.   :  1.0   Min.   :1.000   Min.   :  1.00   No :374   Min.   :55.00
##  1st Qu.:125.8   1st Qu.:2.000   1st Qu.: 57.75   Yes:126   1st Qu.:61.00
##  Median :250.5   Median :3.000   Median :182.50             Median :67.00
##  Mean   :250.5   Mean   :3.436   Mean   :178.55             Mean   :68.56
##  3rd Qu.:375.2   3rd Qu.:5.000   3rd Qu.:298.00             3rd Qu.:76.00
##  Max.   :500.0   Max.   :6.000   Max.   :325.00             Max.   :90.00
##      weight           height           bmi        premeno   momfrac
armassist
##  Min.   : 39.90   Min.   :134.0   Min.   :14.88   No :403   No :435   No
:312
```

```
##  1st Qu.: 59.90   1st Qu.:157.0   1st Qu.:23.27   Yes: 97   Yes: 65
Yes:188
##  Median : 68.00   Median :161.5   Median :26.42
##  Mean   : 71.82   Mean   :161.4   Mean   :27.55
##  3rd Qu.: 81.30   3rd Qu.:165.0   3rd Qu.:30.79
##  Max.   :127.00   Max.   :199.0   Max.   :49.08
##  smoke        raterisk     fracscore      fracture  bonemed   bonemed_fu
##  No :465   Less   :167   Min.   : 0.000   No :375   No :371   No :361
##  Yes: 35   Same   :186   1st Qu.: 2.000   Yes:125   Yes:129   Yes:139
##            Greater:147   Median : 3.000
##                          Mean   : 3.698
##                          3rd Qu.: 5.000
##                          Max.   :11.000
##  bonetreat
##  No :382
##  Yes:118
##
##
##
##
```

```
sum(is.na(glow_bonemed))# missing values are 0 , if not could use to find
specifc na - sapply(glow_bonemed, function(x) sum(is.na(x)))
```

```
## [1] 0
```

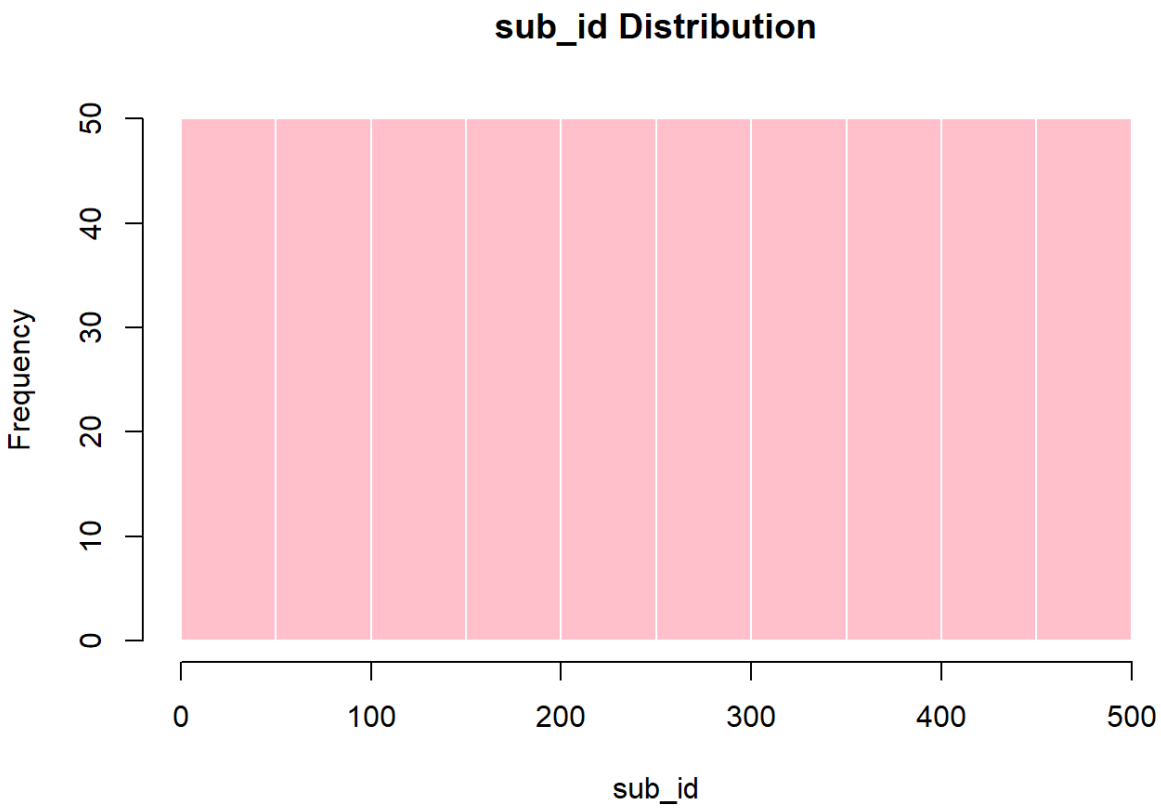
```
str(glow_bonemed) # see structure, integer, factor, levels etc.
```

```
## 'data.frame':    500 obs. of  18 variables:
##  $ sub_id    : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ site_id   : int  1 4 6 6 1 5 5 1 1 4 ...
##  $ phy_id    : int  14 284 305 309 37 299 302 36 8 282 ...
##  $ priorfrac : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 2 1 2 2 1 ...
##  $ age       : int  62 65 88 82 61 67 84 82 86 58 ...
##  $ weight    : num  70.3 87.1 50.8 62.1 68 68 50.8 40.8 62.6 63.5 ...
##  $ height    : int  158 160 157 160 152 161 150 153 156 166 ...
##  $ bmi       : num  28.2 34 20.6 24.3 29.4 ...
##  $ premeno   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ momfrac   : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
```

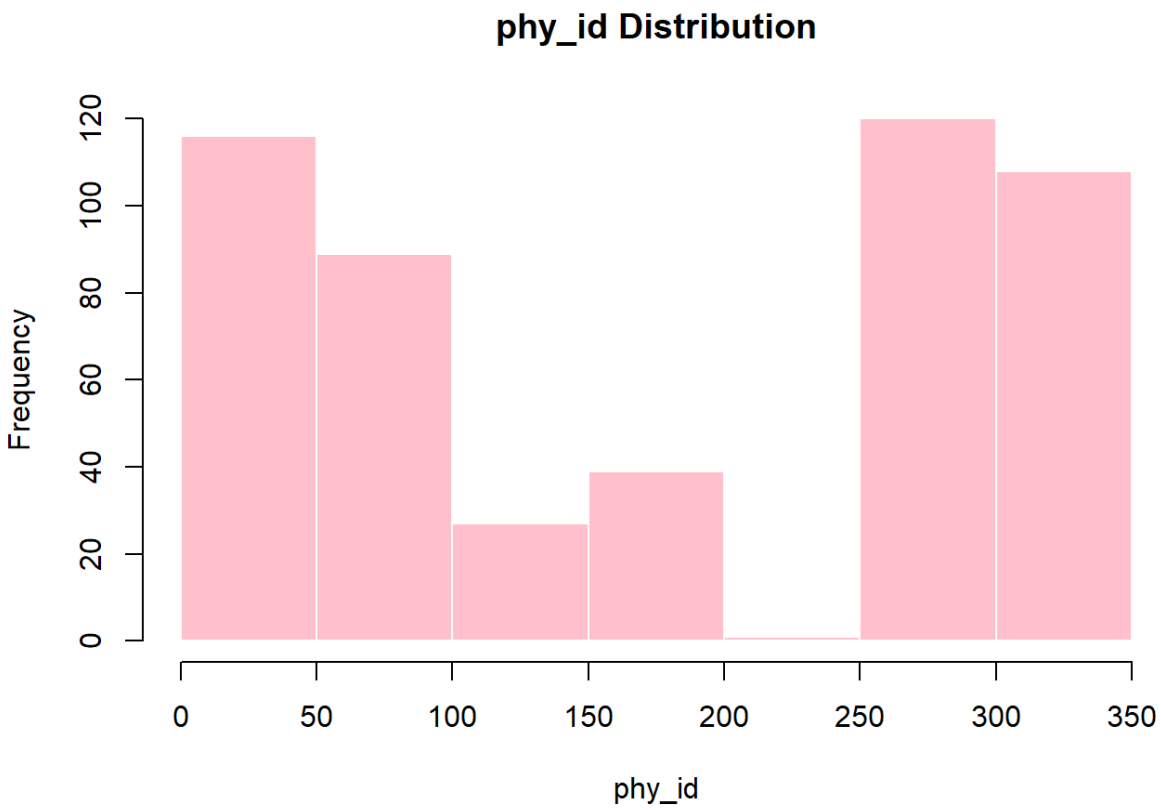
```
##  $ armassist : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
##  $ smoke     : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 2 1 1 1 1 ...
##  $ raterisk  : Factor w/ 3 levels "Less","Same",..: 2 2 1 1 2 2 1 2 2 1
...
##  $ fracscore : int  1 2 11 5 1 4 6 7 7 0 ...
##  $ fracture  : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ bonemed   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
##  $ bonemed_fu: Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
##  $ bonetreat : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
```

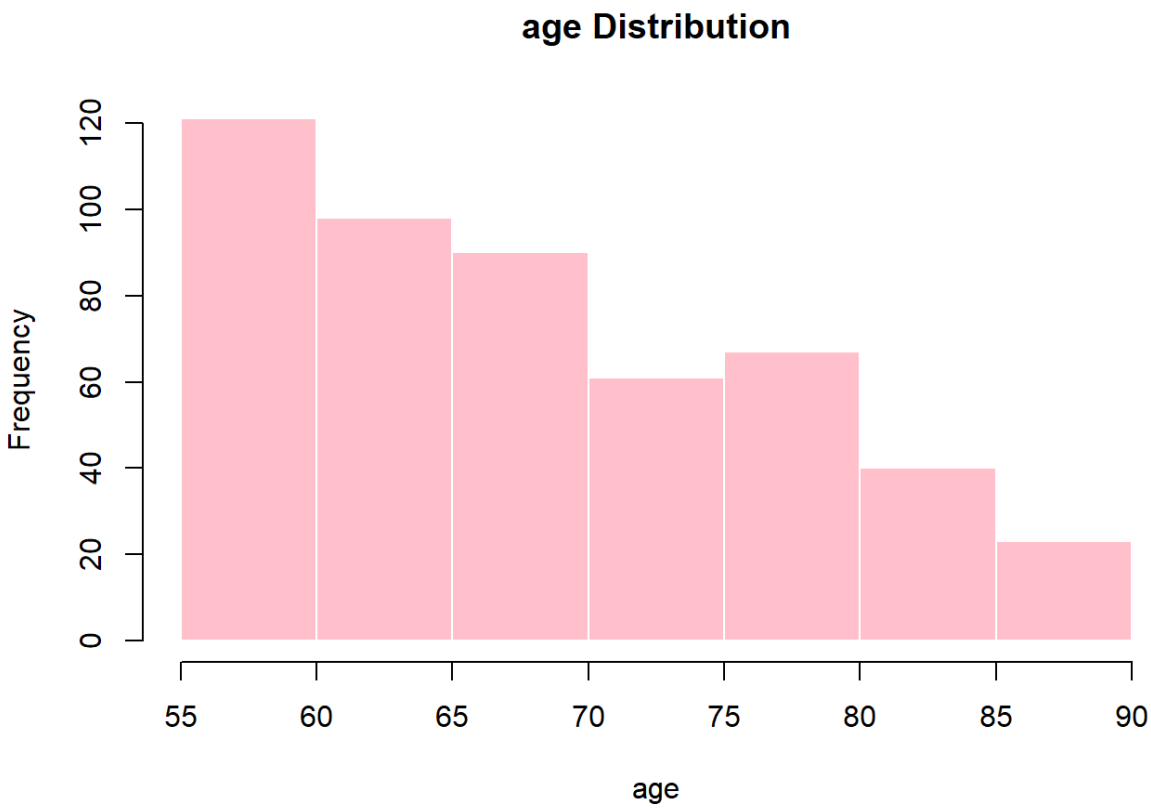
plots

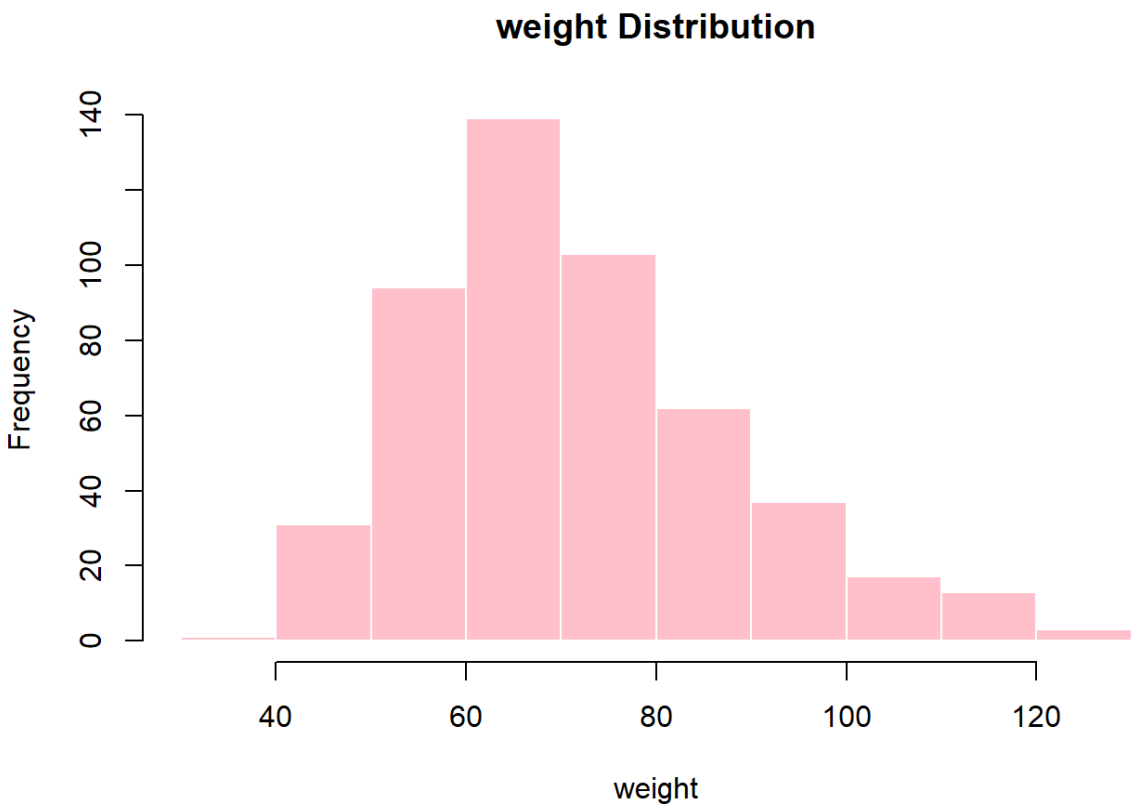
```
#simple histograms

#hist(glow_bonemed$age, main = "Age Distribution", xlab = "Age", col =
"pink", border = "white") # added color, choose a few simple plots for slide
show

#to loop though all numeric Identify numeric variables

numeric_variables <- sapply(glow_bonemed, is.numeric)

# Loop histograms

lapply(names(glow_bonemed)[numeric_variables], function(var_name) {
  hist(glow_bonemed[[var_name]],
       main = paste(var_name, "Distribution"),
       xlab = var_name,
       col = "pink",
       border = "white")
})
```

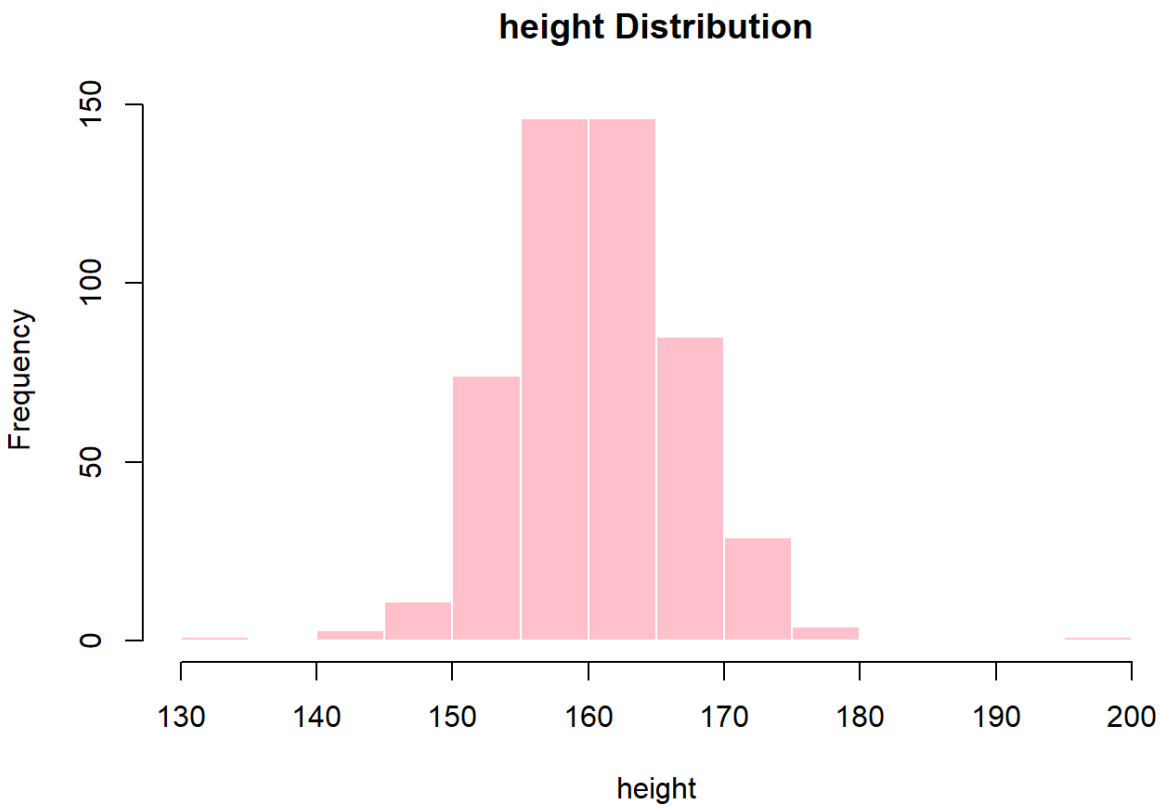

sub_id Distribution
Frequency
0
10
20
30
40
50
0
100
200
300
400
500
sub_id

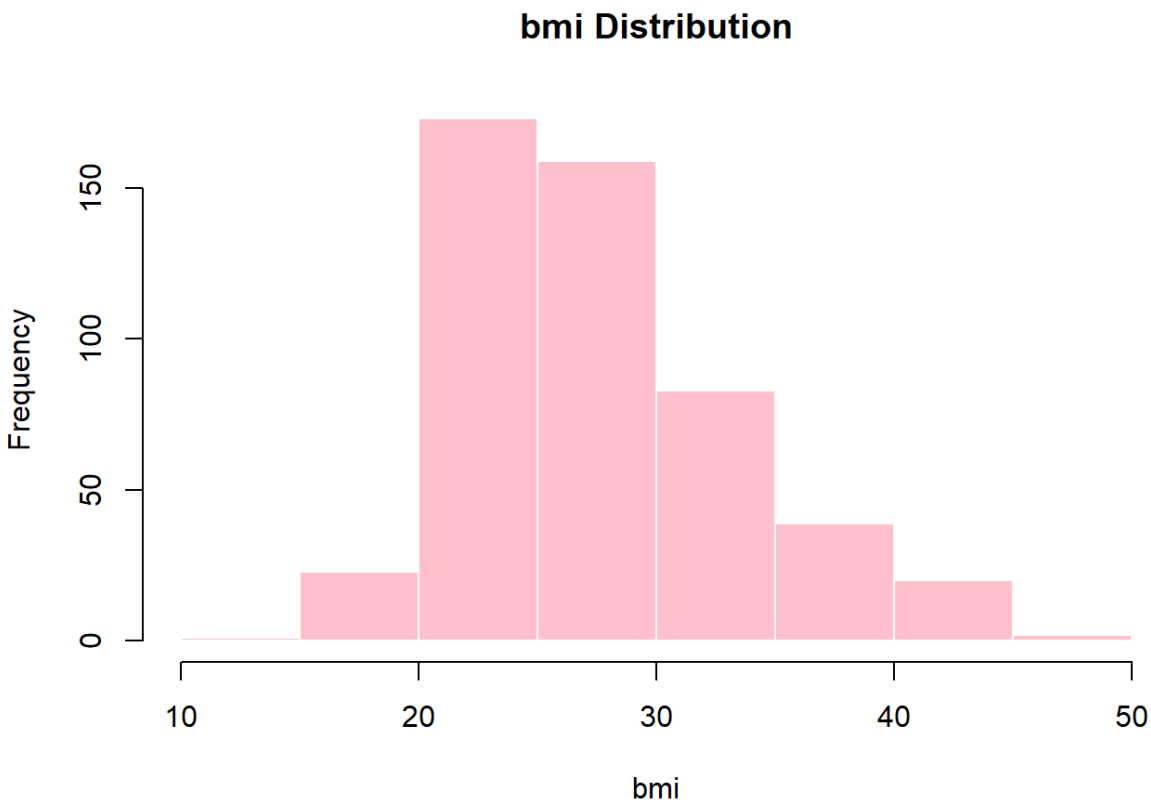

site_id Distribution
Frequency
0
20
40
60
80
100
120
1
2
3
4
5
6
site_id

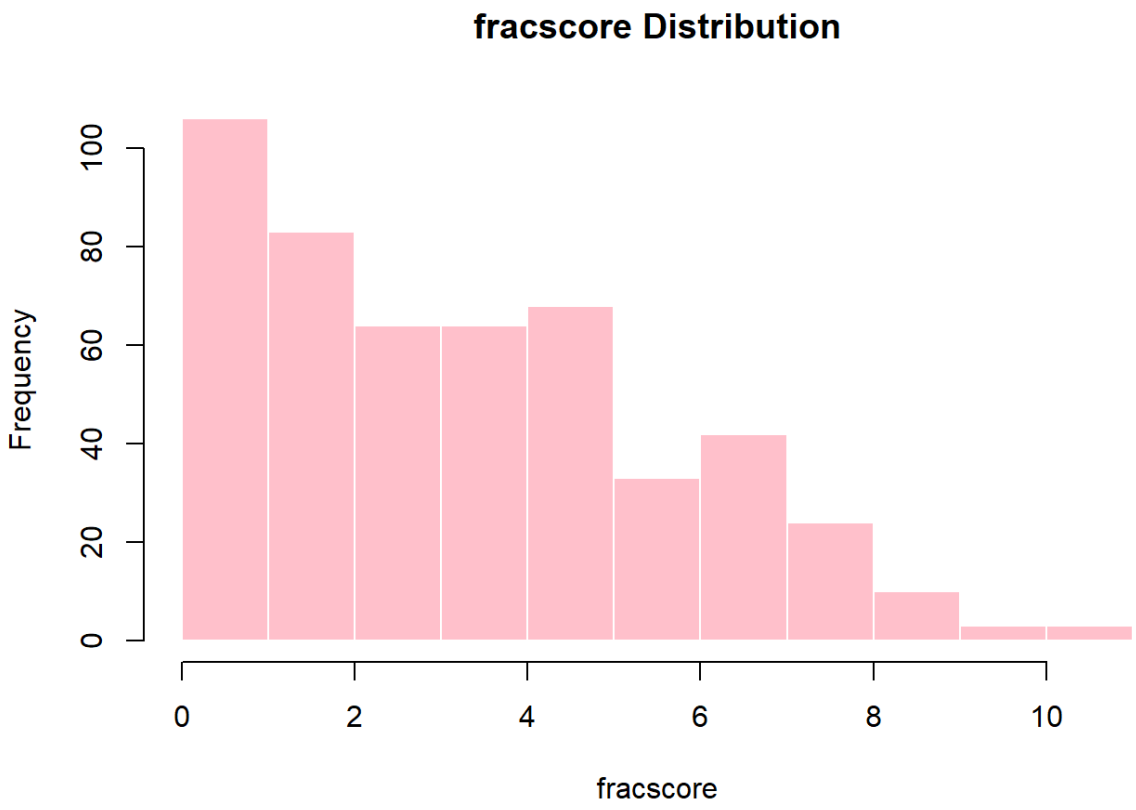

phy_id Distribution
Frequency
0
20
40
60
80
100
120
0
50
100
150
200
250
300
350
phy_id


age Distribution
Frequency
0
20
40
60
80
100
120
55
60
65
70
75
80
85
90
age


weight Distribution
Frequency
0
20
40
60
80
100
140
40
60
80
100
120
weight


height Distribution
Frequency
0
50
100
150
130
140
150
160
170
180
190
200
height


bmi Distribution
Frequency
0
50
100
150
10
20
30
40
50
bmi




```
## [[1]]
## $breaks
##  [1]   0  50 100 150 200 250 300 350 400 450 500
##
## $counts
##  [1] 50 50 50 50 50 50 50 50 50 50
##
## $density
##  [1] 0.002 0.002 0.002 0.002 0.002 0.002 0.002 0.002 0.002 0.002
##
## $mids
##  [1]  25  75 125 175 225 275 325 375 425 475
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
```

```
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
##
## [[2]]
## $breaks
##  [1] 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0
##
## $counts
##  [1] 107  90   0  65   0  36   0 120   0  82
##
## $density
##  [1] 0.428 0.360 0.000 0.260 0.000 0.144 0.000 0.480 0.000 0.328
##
## $mids
##  [1] 1.25 1.75 2.25 2.75 3.25 3.75 4.25 4.75 5.25 5.75
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
##
## [[3]]
## $breaks
## [1]   0  50 100 150 200 250 300 350
##
## $counts
## [1] 116  89  27  39   1 120 108
```

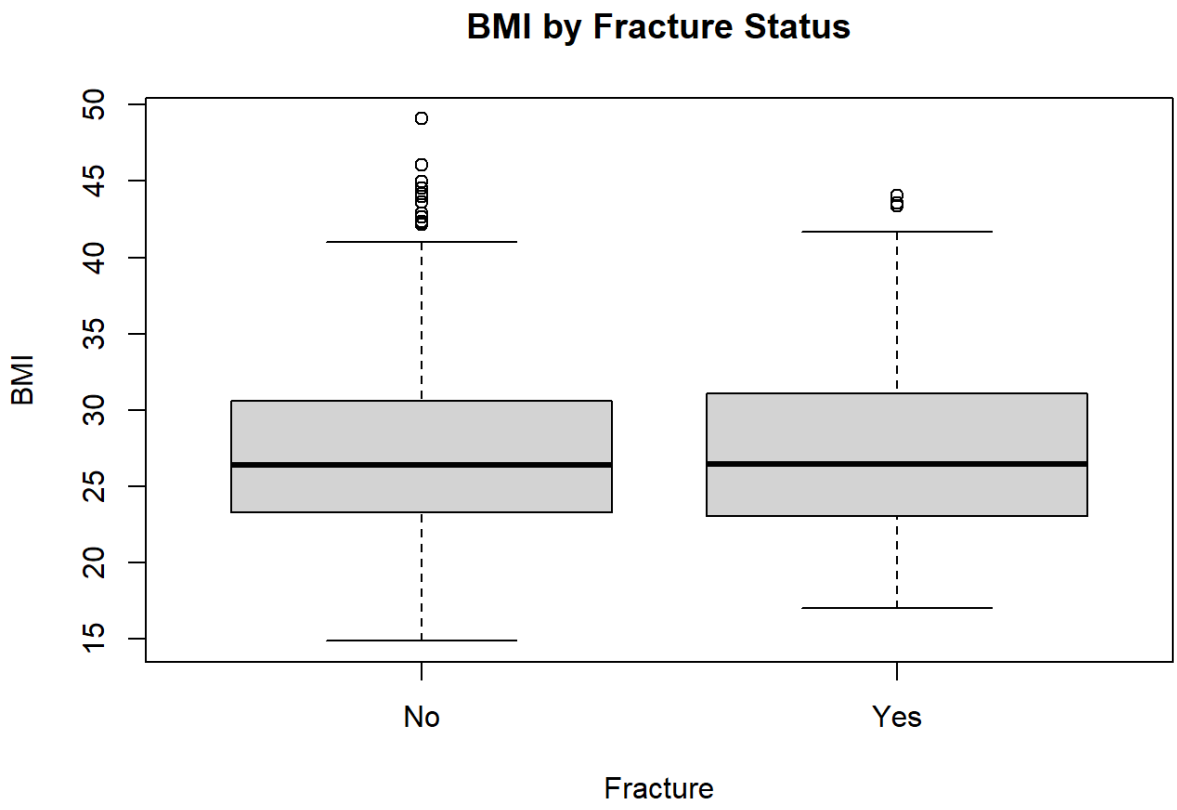
```
##
## $density
## [1] 0.00464 0.00356 0.00108 0.00156 0.00004 0.00480 0.00432
##
## $mids
## [1]  25  75 125 175 225 275 325
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
##
## [[4]]
## $breaks
## [1] 55 60 65 70 75 80 85 90
##
## $counts
## [1] 121  98  90  61  67  40  23
##
## $density
## [1] 0.0484 0.0392 0.0360 0.0244 0.0268 0.0160 0.0092
##
## $mids
## [1] 57.5 62.5 67.5 72.5 77.5 82.5 87.5
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
## $equidist
## [1] TRUE
```

```
##
## attr(,"class")
## [1] "histogram"
##
## [[5]]
## $breaks
##  [1]  30  40  50  60  70  80  90 100 110 120 130
##
## $counts
##  [1]   1  31  94 139 103  62  37  17  13   3
##
## $density
##  [1] 0.0002 0.0062 0.0188 0.0278 0.0206 0.0124 0.0074 0.0034 0.0026 0.0006
##
## $mids
##  [1]  35  45  55  65  75  85  95 105 115 125
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
##
## [[6]]
## $breaks
##  [1] 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200
##
## $counts
##  [1]   1   0   3  11  74 146 146  85  29   4   0   0   0   1
##
## $density
```

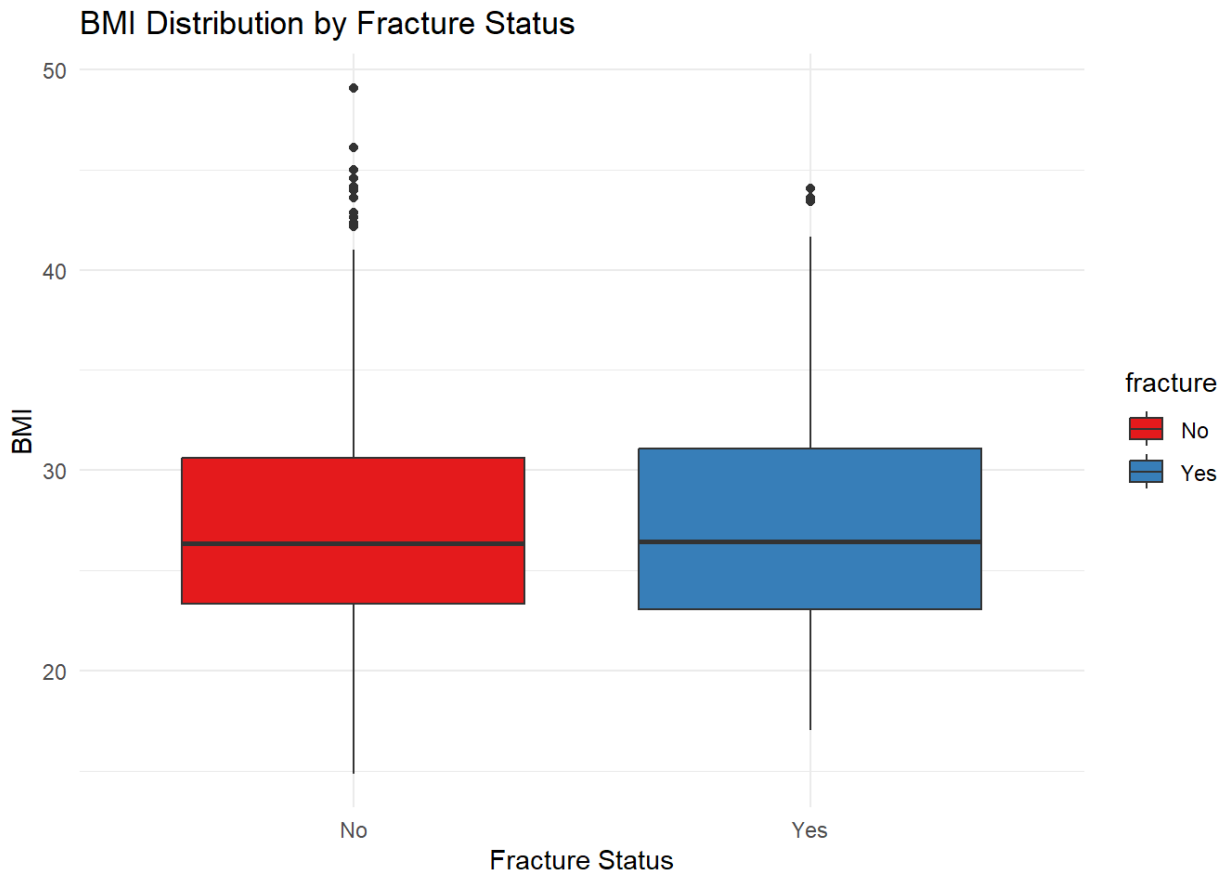
```
##  [1] 0.0004 0.0000 0.0012 0.0044 0.0296 0.0584 0.0584 0.0340 0.0116 0.0016
## [11] 0.0000 0.0000 0.0000 0.0004
##
## $mids
##  [1] 132.5 137.5 142.5 147.5 152.5 157.5 162.5 167.5 172.5 177.5 182.5
187.5
## [13] 192.5 197.5
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
##
## [[7]]
## $breaks
## [1] 10 15 20 25 30 35 40 45 50
##
## $counts
## [1]   1  23 173 159  83  39  20   2
##
## $density
## [1] 0.0004 0.0092 0.0692 0.0636 0.0332 0.0156 0.0080 0.0008
##
## $mids
## [1] 12.5 17.5 22.5 27.5 32.5 37.5 42.5 47.5
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
## $equidist
## [1] TRUE
```

```
##
## attr(,"class")
## [1] "histogram"
##
## [[8]]
## $breaks
##  [1]  0  1  2  3  4  5  6  7  8  9 10 11
##
## $counts
##  [1] 106  83  64  64  68  33  42  24  10   3   3
##
## $density
##  [1] 0.212 0.166 0.128 0.128 0.136 0.066 0.084 0.048 0.020 0.006 0.006
##
## $mids
##  [1]  0.5  1.5  2.5  3.5  4.5  5.5  6.5  7.5  8.5  9.5 10.5
##
## $xname
## [1] "glow_bonemed[[var_name]]"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
```

```
#simple boxplots
boxplot(bmi ~ fracture, data=glow_bonemed, main="BMI by Fracture Status",
xlab="Fracture", ylab="BMI")
```




```
# Creating boxplots for BMI by fracture status
ggplot(glow_bonemed, aes(x = fracture, y = bmi, fill = fracture)) +
  geom_boxplot() +
  scale_fill_brewer(palette = "Set1") +
  labs(y = "BMI", x = "Fracture Status", title = "BMI Distribution by
Fracture Status") +
  theme_minimal()
```

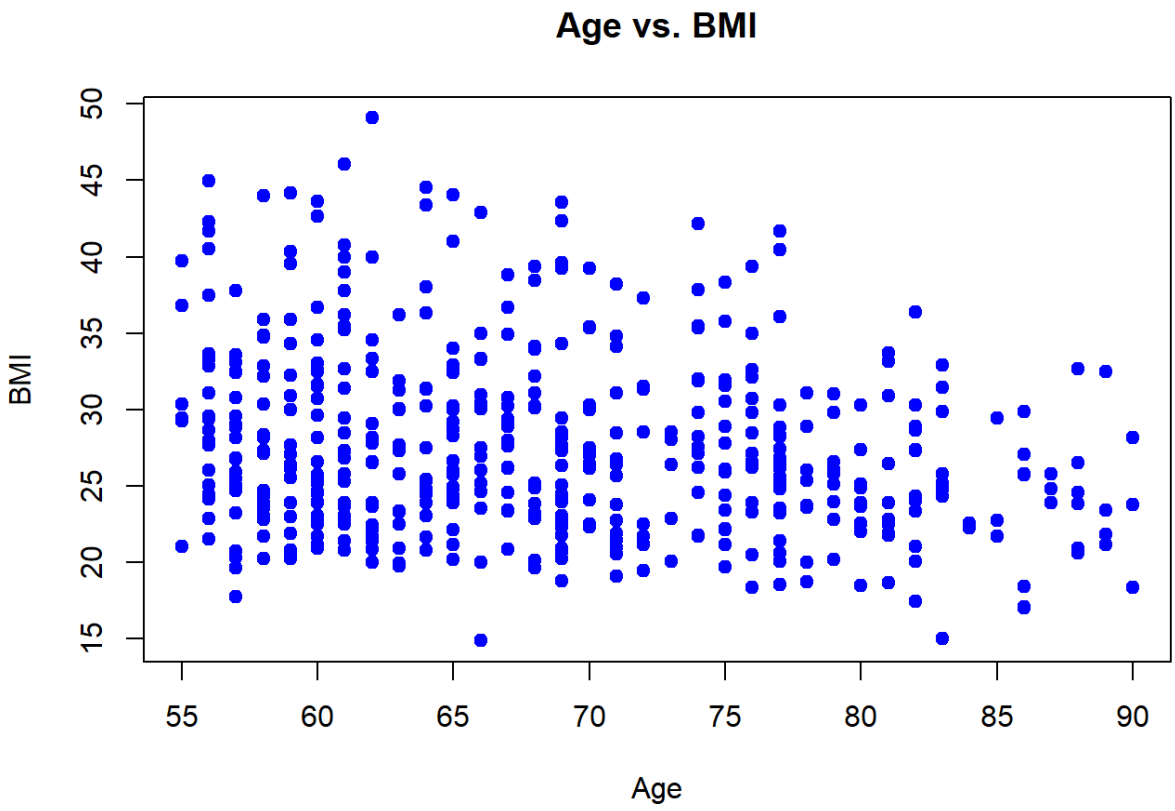



```
 # position="dodge" places the bars side by side instead of stacking them


#corr matrix
numeric_vars <- c("age", "weight", "height", "bmi", "fracscore")


cor_matrix <- cor(glow_bonemed[,numeric_vars])
corrplot(cor_matrix, method = "circle")
```

```
#scatterplot matrix
numeric_variables <- c("age", "weight", "height", "bmi", "fracscore")
ggpairs(glow_bonemed[, c(numeric_variables, "fracture")], aes(colour =
fracture))
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

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

```
# as age increases fracture score increases


# simple scatterplot
 plot(glow_bonemed$age, glow_bonemed$bmi, main = "Age vs. BMI", xlab = "Age",
ylab = "BMI", pch = 19, col = "blue")
```




```
head(glow_bonemed)
```

```
##   sub_id site_id phy_id priorfrac age weight height      bmi premeno
momfrac
## 1      1       1     14        No  62   70.3    158 28.16055      No
No
## 2      2       4    284        No  65   87.1    160 34.02344      No
No
```

```
## 3      3       6    305       Yes  88   50.8    157 20.60936      No
Yes
## 4      4       6    309        No  82   62.1    160 24.25781      No
No
## 5      5       1     37        No  61   68.0    152 29.43213      No
No
## 6      6       5    299       Yes  67   68.0    161 26.23356      No
No
##   armassist smoke raterisk fracscore fracture bonemed bonemed_fu bonetreat
## 1        No    No     Same         1       No      No         No        No
## 2        No    No     Same         2       No      No         No        No
## 3       Yes    No     Less        11       No      No         No        No
## 4        No    No     Less         5       No      No         No        No
## 5        No    No     Same         1       No      No         No        No
## 6        No   Yes     Same         4       No      No         No        No
```

```
str(glow_bonemed)
```

```
## 'data.frame':    500 obs. of  18 variables:
##  $ sub_id    : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ site_id   : int  1 4 6 6 1 5 5 1 1 4 ...
##  $ phy_id    : int  14 284 305 309 37 299 302 36 8 282 ...
##  $ priorfrac : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 2 1 2 2 1 ...
##  $ age       : int  62 65 88 82 61 67 84 82 86 58 ...
##  $ weight    : num  70.3 87.1 50.8 62.1 68 68 50.8 40.8 62.6 63.5 ...
##  $ height    : int  158 160 157 160 152 161 150 153 156 166 ...
##  $ bmi       : num  28.2 34 20.6 24.3 29.4 ...
##  $ premeno   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ momfrac   : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
##  $ armassist : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
##  $ smoke     : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 2 1 1 1 1 ...
##  $ raterisk  : Factor w/ 3 levels "Less","Same",..: 2 2 1 1 2 2 1 2 2 1
...
##  $ fracscore : int  1 2 11 5 1 4 6 7 7 0 ...
##  $ fracture  : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ bonemed   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
##  $ bonemed_fu: Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
##  $ bonetreat : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
```

```
#bonemed Bone medications at enrollment (1: No, 2: Yes)

#bonemed_fu Bone medications at follow-up (1: No, 2: Yes)

#bonetreat Bone medications both at enrollment and follow-up (1: No, 2: Yes)

#priorfrac History of Prior Fracture (1: No, 2: Yes)

# age Age at Enrollment (Years)

#fracscore Fracture Risk Score (Composite Risk Score)

#fractureAny fracture in first year (1: No, 2: Yes)



#if want to change, use mutate or pivot

#tidy data (Reshape wide to long), use pivot

#glow_bonemed_long <- glow_bonemed %>%

 # pivot_longer(cols = c(age, weight, height, bmi, fracscore), names_to =
"variable", values_to = "value")
```

logistic regression model #glm (Generalized Linear Model)/binomial family/ fitting a logistic regression model/ predict the log odds of the probability of the outcome occurring

```
str(glow_bonemed)# no codes 1, yes coded 2
## 'data.frame':    500 obs. of  18 variables:
##  $ sub_id    : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ site_id   : int  1 4 6 6 1 5 5 1 1 4 ...
##  $ phy_id    : int  14 284 305 309 37 299 302 36 8 282 ...
##  $ priorfrac : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 2 1 2 2 1 ...
##  $ age       : int  62 65 88 82 61 67 84 82 86 58 ...
##  $ weight    : num  70.3 87.1 50.8 62.1 68 68 50.8 40.8 62.6 63.5 ...
##  $ height    : int  158 160 157 160 152 161 150 153 156 166 ...
##  $ bmi       : num  28.2 34 20.6 24.3 29.4 ...
##  $ premeno   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ momfrac   : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
##  $ armassist : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
##  $ smoke     : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 2 1 1 1 1 ...
##  $ raterisk  : Factor w/ 3 levels "Less","Same",..: 2 2 1 1 2 2 1 2 2 1
...
##  $ fracscore : int  1 2 11 5 1 4 6 7 7 0 ...
##  $ fracture  : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ bonemed   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
```

```
##  $ bonemed_fu: Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
##  $ bonetreat : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
```

```
levels(glow_bonemed$fracture)
```

```
## [1] "No"  "Yes"
```

```
# Convert  factor to its underlying integer codes 1,2
as.numeric(glow_bonemed$fracture)
```

```
##   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [38] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [75] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [112] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [149] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [186] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [223] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [260] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [297] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [334] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [371] 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [408] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [445] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [482] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
```

```
str(glow_bonemed)
```

```
## 'data.frame':    500 obs. of  18 variables:
##  $ sub_id    : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ site_id   : int  1 4 6 6 1 5 5 1 1 4 ...
##  $ phy_id    : int  14 284 305 309 37 299 302 36 8 282 ...
##  $ priorfrac : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 2 1 2 2 1 ...
```

```
##  $ age       : int  62 65 88 82 61 67 84 82 86 58 ...
##  $ weight    : num  70.3 87.1 50.8 62.1 68 68 50.8 40.8 62.6 63.5 ...
##  $ height    : int  158 160 157 160 152 161 150 153 156 166 ...
##  $ bmi       : num  28.2 34 20.6 24.3 29.4 ...
##  $ premeno   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ momfrac   : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
##  $ armassist : Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...
##  $ smoke     : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 2 1 1 1 1 ...
##  $ raterisk  : Factor w/ 3 levels "Less","Same",..: 2 2 1 1 2 2 1 2 2 1
...
##  $ fracscore : int  1 2 11 5 1 4 6 7 7 0 ...
##  $ fracture  : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ bonemed   : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
##  $ bonemed_fu: Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
##  $ bonetreat : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 2 1 1 ...
```

```
#dependent variable is fracture : Any fracture in first year (1: No, 2: Yes)
#is it predicted by: "priorfrac"      "age"            "weight"
# "height"         "bmi"            "premeno"        "momfrac"
"armassist"      "smoke"
# "raterisk"       "fracscore"      "bonemed"        "bonemed_fu"
"bonetreat"
```

```
model1 <- glm(formula = fracture ~ priorfrac + age + weight + height + bmi +
premeno +
              momfrac + armassist + smoke + raterisk + fracscore + bonemed +
              bonemed_fu + bonetreat,
              family = binomial, data = glow_bonemed)
```

```
summary(model1)
```

```
##
## Call:
## glm(formula = fracture ~ priorfrac + age + weight + height +
##     bmi + premeno + momfrac + armassist + smoke + raterisk +
##     fracscore + bonemed + bonemed_fu + bonetreat, family = binomial,
##     data = glow_bonemed)
```

```
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -17.07822   12.80864  -1.333  0.18242
## priorfracYes      0.52585    0.39600   1.328  0.18420
## age               0.01618    0.05629   0.288  0.77371
## weight           -0.13254    0.08907  -1.488  0.13673
## height            0.08019    0.08120   0.988  0.32340
## bmi               0.36724    0.23143   1.587  0.11255
## premenoYes        0.19494    0.29255   0.666  0.50520
## momfracYes        0.61382    0.42858   1.432  0.15208
## armassistYes      0.16779    0.61066   0.275  0.78349
## smokeYes         -0.27693    0.54660  -0.507  0.61241
## rateriskSame      0.36390    0.28982   1.256  0.20926
## rateriskGreater   0.53726    0.31964   1.681  0.09280 .
## fracscore         0.09721    0.28046   0.347  0.72889
## bonemedYes        1.50963    0.65911   2.290  0.02200 *
## bonemed_fuYes     1.61953    0.50153   3.229  0.00124 **
## bonetreatYes     -2.66789    0.83779  -3.184  0.00145 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 562.34  on 499  degrees of freedom
## Residual deviance: 488.15  on 484  degrees of freedom
## AIC: 520.15
##
## Number of Fisher Scoring iterations: 4
confint(model1)
## Waiting for profiling to be done...
##                        2.5 %      97.5 %
## (Intercept)     -42.22434329  8.02050531
## priorfracYes     -0.24736648  1.30847429
```

```
## age              -0.09386000  0.12727158
## weight           -0.31025521  0.03961953
## height           -0.07963981  0.23885108
## bmi              -0.08069659  0.82854835
## premenoYes       -0.39043906  0.76026365
## momfracYes       -0.22673386  1.45795795
## armassistYes     -1.02811562  1.37123892
## smokeYes         -1.39807971  0.76400830
## rateriskSame     -0.20000696  0.93956144
## rateriskGreater  -0.08728069  1.16926797
## fracscore        -0.45491366  0.64695642
## bonemedYes        0.20460939  2.84614275
## bonemed_fuYes     0.64862664  2.63786064
## bonetreatYes     -4.34824962 -1.03068346
#AIC (Akaike Information Criterion) is 520.15


# interpretations


#"bonemed"  -  1.50963, p val 0.02200  CI[0.20460939 , 2.84614275]

# patient on bonemed at enrollment, log odds of having fracture expected to
increase by approx 1.51 compared to those not on meds, other factor beoing
consistent


#"bonemed_fu"   1.61953 , p val 0.0012 CI[0.64862664 , 2.63786064]

# patient on bonemed at follow up, log odds of having fracture expected to
increase by approx 1.62 compared to those not on meds, other factor being
consistent high p value, very stats sig.


#"bonetreat"  -2.66789 , p vale  0.00145 CI [-4.34824962 , -1.03068346]

# patient on bonemed at enrollment & follow up, log odds of having fracture
expected to decrease by approx 5.67 compared to those not on meds, other
factor being consistent high p value, very stats sig. CI is very wide so
uncertianty about this.


#model fit

# Predict probabilities of fracture

#predicted_probs <- predict(model1, type = "response")
```

```r
# Adding predicted prob back to the dataset
#glow_bonemed$predicted_prob = predicted_probs
#head(glow_bonemed)

#too
model2 <- glm(formula = fracture ~age + bonemed +
              bonemed_fu + bonetreat,
              family = binomial, data = glow_bonemed)

summary(model2)
```

```
##
## Call:
## glm(formula = fracture ~ age + bonemed + bonemed_fu + bonetreat,
##     family = binomial, data = glow_bonemed)
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -4.51622    0.85025  -5.312 1.09e-07 ***
## age            0.04472    0.01209   3.699 0.000216 ***
## bonemedYes     1.40723    0.64196   2.192 0.028374 *
## bonemed_fuYes  1.82394    0.47822   3.814 0.000137 ***
## bonetreatYes  -2.59557    0.81171  -3.198 0.001386 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 562.34  on 499  degrees of freedom
## Residual deviance: 519.29  on 495  degrees of freedom
## AIC: 529.29
##
## Number of Fisher Scoring iterations: 4
```

```r
confint(model2)
```

```
## Waiting for profiling to be done...
##                       2.5 %      97.5 %
## (Intercept)    -6.20871813 -2.86949335
## age             0.02112541  0.06860985
## bonemedYes      0.13661628  2.71394369
## bonemed_fuYes   0.90229270  2.80188014
## bonetreatYes   -4.22572102 -1.00798370
# AIC: 529.29 but now age is stat sig.
```