QMB Assignment 8

Name - Manish Sanjay Talekar

U54500765

#Preprocessing

#q1

**> glow <- read\_xlsx("6304 Module 8 Assignment Data Set.xlsx")**

**> colnames(glow)=tolower(make.names(colnames(glow)))**

#q2

**> set.seed(54500765)**

**> main\_data <- sample\_n(glow, 150)**

#Analysis

#q1

**> all\_out <- glm(fracture~priorfrac + age + weight + height + bmi + menoby45 + momfrac + armassist, data = main\_data, family = "binomial")**

#q2

**> summary(all\_out)**

**Call:**

**glm(formula = fracture ~ priorfrac + age + weight + height +**

**bmi + menoby45 + momfrac + armassist, family = "binomial",**

**data = main\_data)**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-1.7190 -0.7679 -0.5632 1.0084 2.1266**

**Coefficients:**

**Estimate Std. Error z value Pr(>|z|)**

**(Intercept) 9.65056 21.73358 0.444 0.6570**

**priorfrac 1.01147 0.47025 2.151 0.0315 \***

**age 0.02721 0.02723 0.999 0.3178**

**weight 0.01157 0.06295 0.184 0.8542**

**height -0.19929 0.32398 -0.615 0.5385**

**bmi -0.05643 0.35032 -0.161 0.8720**

**menoby45 0.56553 0.53527 1.057 0.2907**

**momfrac 1.03785 0.52337 1.983 0.0474 \***

**armassist 0.14482 0.46287 0.313 0.7544**

**---**

**Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1**

**(Dispersion parameter for binomial family taken to be 1)**

**Null deviance: 175.96 on 149 degrees of freedom**

**Residual deviance: 156.32 on 141 degrees of freedom**

**AIC: 174.32**

**Number of Fisher Scoring iterations: 4**

#q3

INTERPRETATION – There is a difference of 19.64 between the null and residual deviance. Therefore, I believe there is markable difference between the two deviance factors.

#q4

Following variables will have the greatest influence in increasing the modeled probability that a subject will have recently experienced a bone fracture,

1. PRIORFRAC
2. AGE
3. WEIGHT
4. MENOBY45
5. MOMFRAC
6. ARMASSIST

#q5

Following variables will have the greatest influence in decreasing the modeled probability that a subject will have recently experienced a bone fracture,

1. HEIGHT
2. BMI

#q6

**> pred\_data <- expand.grid(priorfrac=unique(main\_data$priorfrac),**

**+ age=quantile(main\_data$age, c(0.25, 0.5, 0.75, 1)),**

**+ weight=quantile(main\_data$weight, c(0.25, 0.5, 0.75, 1)),**

**+ height=quantile(main\_data$height, c(0.25, 0.5, 0.75, 1)),**

**+ bmi=quantile(main\_data$bmi, c(0.25, 0.5, 0.75, 1)),**

**+ menoby45=unique(main\_data$menoby45),**

**+ momfrac=unique(main\_data$momfrac),**

**+ armassist=unique(main\_data$armassist)**

**+ )**

**> head\_data <- as.data.frame(head(pred\_data, 5))**

**>**

**> pred\_probs <- round(predict(all\_out, newdata = head\_data, type = "response"), 3)**

**> head\_data$probs <- pred\_probs**

**>**

**> head\_data**

**priorfrac age weight height bmi menoby45 momfrac armassist probs**

**1 0 60.00 135 66 23.28617 0 0 0 0.165**

**2 1 60.00 135 66 23.28617 0 0 0 0.352**

**3 0 67.00 135 66 23.28617 0 0 0 0.193**

**4 1 67.00 135 66 23.28617 0 0 0 0.396**

**5 0 75.75 135 66 23.28617 0 0 0 0.233**

#q7

**> head\_data[which.max(head\_data$probs), c("age", "weight", "probs")]**

**age weight probs**

**4 67 135 0.396**

**> head\_data[which.min(head\_data$probs), c("age", "weight", "probs")]**

**age weight probs**

**1 60 135 0.165**