Exercise 3

Chris Leonard

September 1, 2016

```
ddata <-read.csv("C:/Users/Chris/Desktop/Diabetiesresearch/diabetic_data.csv", header=TRUE)
##Data downloaded 8/30/16
##ddata is the list of events in the hopsital
```

```
mcodes <-read.csv("C:/Users/Chris/Desktop/Diabetiesresearch/IDs_mapping.csv")
type(ddata) ## Codes in case needed

##mcodes = coding of information in reference to ddata

summary(ddata)


addmission <- table(ddata[,c(9)])
addmissionhistogram <-ddata[,c(9)]
paste(addmission)
##part 2
paste("The data points in admissions that is of no use is identified in the accompanying document.  Ent
paste("Not available occurences:",addmission[4])
paste("Null or empty data occurences:", addmission[5])
paste("Not available occurences:", addmission[8])







class(ddata)

colnames(ddata)
##here is a list of ages for patients from data.  This is a check to see if there
##are any erronious data points.  This column looks ok
admitsource <- table(ddata[,c(9)])


addmission <- table(ddata[,c(7)])
addmissionhistogram <-ddata[,c(7)]
paste(admitsource)
eradd <-admitsource[7]
total <- as.numeric("101766")
paste(admitsource)
pERadmit <- ((eradd / total)* (as.numeric("100")))

paste("The precentage of admittence to the ER is:")
paste(pERadmit, "%")
```

```
##Here is admissions information, This must be compared with the coding on 2nd ##document
##For example, There have been 53,990 emergency admissions over the period of data collection

#invalid <-scan("C:/Users/Chris/Desktop/Diabetiesresearch/diabetic_data.csv",sep="\n",what="?")

paste(invalid)

addmission <- table(ddata[,c(7)])
addmissionhistogram <-ddata[,c(7)]
paste(addmission)

eradmit <-(addmission[1])##number of admissions to ER

events <-sum(addmission) #number of total events in data

dischargestat <- as.data.frame(ddata[,c(8)])
admitsource2 <- as.data.frame(ddata[,c(7)])


problem3 <- dim(ddata[ddata$discharge_disposition_id == 11,])[1]

problem3a <- (problem3/ total)* (as.numeric("100"))

paste("The probability of a patient who entered the ER that is registered as dececed is:", problem3a)



#9,8
#compare1 <-data.matrix(ddata[,c(7,8)])
#compare2 <-ddata[,c(7,8)]

compare3 <-subset(compare2,1,select =)
print(compare3)

##Question 4
paste(addmission [1]) ##Highest hospital admmit (ER)

dischage <- table(ddata[8])
paste("The highest discharge, to home, is", dischage[1])


paste("The most frequent discharge rating is discharge to home which occured",dischage [1])

problem4 <- (ddata$discharge_disposition_id[ddata$admission_type_id == 1])
problem4a <-(table(problem4))[1]

paste("The most frequent discharge status for people coming to the ER is",problem4a)

hist(addmissionhistogram)
```