rm(list = ls())

Data = read.csv("Data.csv", stringsAsFactors = FALSE)

n = dim(Data)[1] # The number of rows of the data.frame

### ID is the 43th column in the dataset

# The code below generates a spread sheet with 101 predictors (X1 - X101)

# and the word corresponding to the names of predictors

iters = c(5:7, 13:42, 44:dim(Data)[2])

k = integer(1)

predictor.table = data.frame(predictor = character(101), idno = 1:101, stringsAsFactors = FALSE)

for (i in iters) {

k = k + 1

predictor.table$predictor[k] = names(Data)[i]

names(Data)[i] = paste("X", k, sep = "")

}

write.csv(predictor.table, "predictor\_table.csv", row.names = FALSE)

predictors = Data[,-c(1:4,8:12,43)]

############################################################

#Check the randomness of NA rows

data\_na = Data[is.na(Data$X5),]

boxplot(data\_na$stars)

hist(data\_na$stars)

freqs = table(data\_na$stars)

null.probs = table(Data$stars)

barplot(freqs)

barplot(null.probs)

chisq.test(freqs, null.probs) #Goodness of fit test

# if(!require(MissMech)) {

# install.packages("MissMech");

# require(MissMech)

# }

# MCAR\_test = TestMCARNormality(predictors, del.lesscases = 6, imputation.number = 1, method = "Auto", imputation.method = "Dist.Free", nrep = 100, n.min = 30, seed = 110, alpha = 0.05, imputed.data = NA)

#

# if(!require(BaylorEdPsych)) {

# install.packages("BaylorEdPsych");

# require(BaylorEdPsych)

# }

# if(!require(mvnmle)) {

# install.packages("mvnmle");

# require(mvnmle)

# }

# LittleMCAR(predictors)

#imputation

if(!require(mice)) {

install.packages("mice");

require(mice)

}

imputed\_mice\_ = mice(predictors, m = 3, maxit = 100)

imputed\_mice1 = complete(imputed\_mice\_,1)

imputed\_mice2 = complete(imputed\_mice\_,2)

imputed\_mice3 = complete(imputed\_mice\_,3)

sen\_score=rowMeans(cbind(imputed\_mice1$X5,imputed\_mice2$X5,imputed\_mice3$X5))

imputed\_mice1$X5 = sen\_score

imputed\_mice = imputed\_mice1

write.csv(imputed\_mice, "imputed\_mice.csv", row.names=F)

#imputed\_mice=read.csv("imputed\_mice.csv", stringsAsFactors = FALSE)

if(!require(missForest)) {

install.packages("missForest");

require(missForest)

}

imputed\_forest = missForest(predictors,maxiter = 10, ntree = 100, verbose = T)$ximp

write.csv(imputed\_forest, "imputed\_missForest.csv", row.names=F)

#imputed\_missForest=read.csv("imputed\_missForest.csv", stringsAsFactors = FALSE)

if(!require(mi)) {

install.packages("mi");

require(mi)

}

imputed\_mi = mi(predictors, verbose = T)

mi2stata(imputed\_mi, m=3, "imputed\_mi.csv")

imputed\_mi = read.csv("imputed\_mi.csv", stringsAsFactors = FALSE)