#Loading the creditcardfraud csv file into a data frame

credit\_card=read.csv("creditcardfraud.csv")

credit\_card

#Checking if there are any NA values in the dataset

any(is.na(credit\_card))

# So from the output it is understood that there are NA values in the dataset

#Let us extract the count of NA values in the dataset

sum(is.na(credit\_card))

#Replacing the NA values in each column with the mean value of the values in

#the column

credit\_card$Time[is.na(credit\_card$Time)]=mean(credit\_card$Time,na.rm=TRUE)

credit\_card$V1[is.na(credit\_card$V1)]=mean(credit\_card$V1,na.rm=TRUE)

credit\_card$V2[is.na(credit\_card$V2)]=mean(credit\_card$V2,na.rm=TRUE)

credit\_card$V3[is.na(credit\_card$V3)]=mean(credit\_card$V3,na.rm=TRUE)

credit\_card$V4[is.na(credit\_card$V4)]=mean(credit\_card$V4,na.rm=TRUE)

credit\_card$V5[is.na(credit\_card$V5)]=mean(credit\_card$V5,na.rm=TRUE)

credit\_card$V6[is.na(credit\_card$V6)]=mean(credit\_card$V6,na.rm=TRUE)

credit\_card$V7[is.na(credit\_card$V7)]=mean(credit\_card$V7,na.rm=TRUE)

credit\_card$V8[is.na(credit\_card$V8)]=mean(credit\_card$V8,na.rm=TRUE)

credit\_card$V9[is.na(credit\_card$V9)]=mean(credit\_card$V9,na.rm=TRUE)

credit\_card$V10[is.na(credit\_card$V10)]=mean(credit\_card$V10,na.rm=TRUE)

credit\_card$V11[is.na(credit\_card$V11)]=mean(credit\_card$V11,na.rm=TRUE)

credit\_card$V12[is.na(credit\_card$V12)]=mean(credit\_card$V12,na.rm=TRUE)

credit\_card$V13[is.na(credit\_card$V13)]=mean(credit\_card$V13,na.rm=TRUE)

credit\_card$V14[is.na(credit\_card$V14)]=mean(credit\_card$V14,na.rm=TRUE)

credit\_card$V15[is.na(credit\_card$V15)]=mean(credit\_card$V15,na.rm=TRUE)

credit\_card$V16[is.na(credit\_card$V16)]=mean(credit\_card$V16,na.rm=TRUE)

credit\_card$V17[is.na(credit\_card$V17)]=mean(credit\_card$V17,na.rm=TRUE)

credit\_card$V18[is.na(credit\_card$V18)]=mean(credit\_card$V18,na.rm=TRUE)

credit\_card$V19[is.na(credit\_card$V19)]=mean(credit\_card$V19,na.rm=TRUE)

credit\_card$V20[is.na(credit\_card$V20)]=mean(credit\_card$V20,na.rm=TRUE)

credit\_card$V21[is.na(credit\_card$V21)]=mean(credit\_card$V21,na.rm=TRUE)

credit\_card$V22[is.na(credit\_card$V22)]=mean(credit\_card$V22,na.rm=TRUE)

credit\_card$V23[is.na(credit\_card$V23)]=mean(credit\_card$V23,na.rm=TRUE)

credit\_card$V24[is.na(credit\_card$V24)]=mean(credit\_card$V24,na.rm=TRUE)

credit\_card$V25[is.na(credit\_card$V25)]=mean(credit\_card$V25,na.rm=TRUE)

credit\_card$V26[is.na(credit\_card$V26)]=mean(credit\_card$V26,na.rm=TRUE)

credit\_card$V27[is.na(credit\_card$V27)]=mean(credit\_card$V27,na.rm=TRUE)

credit\_card$V28[is.na(credit\_card$V28)]=mean(credit\_card$V28,na.rm=TRUE)

credit\_card$Amount[is.na(credit\_card$Amount)]=mean(credit\_card$Amount,na.rm=TRUE)

credit\_card$Class[is.na(credit\_card$Class)]=mean(credit\_card$Class,na.rm=TRUE)

options(scipen = 5)

credit\_card

any(is.na(credit\_card))

#From the result it is understood that there are NA values in the dataset

#So the data is now cleaned

#Class of data object

class(credit\_card)

#Display Internal structure of data

str(credit\_card)

#Summary of data

summary(credit\_card)

#Column names

names(credit\_card)

#Dimensions of data

dim(credit\_card)

# Data of the top

head(credit\_card)

#Data from the top

tail(credit\_card)





















































