**PART 2 ANSWER**

install.packages("polycor")

library("polycor")

set1<-read.csv("Bank\_branch\_transactions.csv")

set2<-read.csv("Bank\_customer\_db.csv" )

df <- data.frame(corr=double(),i=integer(),j=integer())

a=0

q=0

for(i in 1:19)

{

for(j in 1:12)

{

setA<-as.integer(set1[,i])

setB<-as.integer(set2[,j])

if(mean(setA)!= setA[1] && mean(setB)!= setB[1] ){

r<-hetcor(setA,setB)

a<-c(a,r$correlations[2,1])

q<-q+1

df[q,]<-c(r$correlations[2,1],i,j)

}

}

}

p<-df[which.max(df[,1]),]

HighestCorrelation<-c(names(set1[p[1,2]]) , names(set1[p[1,3]]))

**SO THE HIGHEST CORRELATION AMONG 2 DATASETS IS BETWEEN:-**

**"Bank\_branch\_transactions.account\_number"**

**"Bank\_customer\_info.dependents"**