**Exercici 1.- Eliminar els identificadors del conjunt de dades (0.5p)**

1. Quins són els atributs identificadors en el conjunt de dades?

DNI i número de la seguretat social “SS”.

1. Descriu les comandes que has emprat, indicant breument la funció de cadascun dels paràmetres que has utilitzat.

dades\_sense\_identificador<-dades[2:5]

Bàsicament dades\_sense\_identificar rep la variable dades però traient la primera y segona columna.

**Exercici 2.- Re-identificació (0.5p)**

1. Quins d’aquests atributs o combinacions poden conduir a la identificació única d’un registre del conjunt de dades?

Es poden donar diferents combinacions per a identificar, a mes a mes al ser unes dades amb tant poques persones es molt més fàcil. Identificar a una persona nomes el CP, perquè existeix una persona amb aquets codi únicament. També la identificació mitjançant la combinació CP i edat, com que les dades que tenim, i per últim, la combinació CP i salari.

1. Indica quines comandes has fet servir per obtenir aquesta informació, explicitant el significat de les comandes i dels seus paràmetres.

x<-dades2[1]

y<-dades2[2]

c<- dades2[3]

which(x==25560) #Resultat posició 18, identificat nomes amb CP.

which(x==17800 & y==32) #Resultat posició 1, identificat amb CP i edat.

which(x==43870 & c==78) #Resultat posició 94, identificat amb CP i salari.

Les primeres instruccions emmagatzemen les dades de CP, edat i salari. Which es una instrucció que troba una/es comparacions, retornant la posició.

**Exercici 3.- Aplicació de soroll additiu (0.5p)**

1. Indica la comanda que has fet servir, junt amb els paràmetres emprats i el significat de cadascun d’ells?

dades.an=addNoise(dades2,variables=’Edat’,noise=20,method=’additive’)

1. Executa el següent codi adjunt i explica els resultats obtinguts, justifica la resposta.

Es mostra un gràfic, aquest ens mostra dues dades combinades, com si fos un punt amb dos valors (x,y). Es a dir, al Original tenim el valor de dades columna edat i al Masked el valor de dades.an columna edat però amb el soroll ja afegit.

També tenim una línia vermella que aporta poca informació, per a referenciar bé el valor dels punts.

**Exercici 4.- Aplicació del mètode d’intercanvi de rang additiu (rang swap) (0.5p)**

1. Indica la comanda que has fet servir, junt amb els paràmetres emprats i el significat de cadascun d’ells.

dades.rs=rankSwap(dades2,variables=’Edat’,TopPercent=10, BottomPercent=10, K0=NULL,R0=NULL, missing=NA, seed=NULL)

Guardem el resultat en la variable dades.rs, la comanda rankSwap s’encarrega de intercanviar els valors dins d’un rang especificat. Aquesta comanda rep un paràmetres, aquells que estan en NULL no fens interessen, el paràmetre dades2 es d’on extraiem les dades, el paràmetre variables en dius que nomes mirem el la columna edat. Per aplicar el mètode pertorbatiu necessitem un percentatge del 10%, aquest especificat als dos paràmetres TopPercent i BottomPercent,

1. Genera una gràfica similar a la que hem emprat en l’exercici anterior apartat b), on es pugui veure de forma visual i ràpid la dispersió o alteració de l’atribut Edat després d’aplicar el mètode d’intercanvi del rang.

plot(cbind(dades$Edat,dades.rs$Edat),ylim=c(min(dades$Edat),max(dades$Edat)),xlim=c(min(dades$Edat),max(dades$Edat)),xlab="Original",ylab

="Masked", main="Metode Pertorbatiu - Edat (P=0.10) ")

abline(a=0, b=1, col=”red”)