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

Responeu a les següent qüestions:

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

Els atributs que identifiquen una persona de manera única i inequívocament són el DNI i el número de la SS.

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

que has utilitzat.

dades\_subset = dades[,c(3,4,5)]. Amb aquesta comanda el que fem es que creem una nova variable “dades\_subset” on hi guardarem el contingut de les columnes 3, 4 i 5 que es corresponen a CP, Edat i salari.

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

Responeu a les següent qüestions:

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

Hem trobat que tant per Edat, CP i Salari hi ha persones que tenen valors únics i llavors qualsevol dels 3 atributs poden identificar de manera única a una persona concreta. A més a més també hem trobat que les combinacions CP-Edat, CP-Salari, Edat-Salari i CP-Edat-Salari identifiquen a una persona concretament.

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

dades\_CP\_freq = table(dades\_subset$CP)

Amb aquesta comanda creem una nova taula a partir del subset que hem treballat en l’apartat anterior per a la columna CP. Com a sortida d’aquesta comanda rebrem una taula amb els diferents valors que pren el CP i la freqüència amb la que aquests valors surten.

Si ara volem fer una combinació, la comanda resultant es molt similar:

dades\_CP\_Edat\_freq = table(dades\_subset$CP,dades\_subset$Edat)

Ara tenim una taula que ens mesura per a la variable1(CP) i per a la variable2(Edat) les combinacions possibles i quants cops apareixen.

Un cop hem probat totes les combinacions i tots els camps que considerem que poden ser quasi-identificadors utilitzarem:

'1' %in% TAULA

on TAULA és el nom de la taula generada anteriorment com per exemple dades\_CP\_freq i ens mostrarà TRUE si hi ha cap individu que es pot identificar de manera única.

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

Responeu a les següent qüestions:

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

b. Executa el següent codi adjunt i explica els resultats obtinguts, justificant la resposta.