

Deure 1

Considera les dades **bank.sav** (veure els fitxers de dades que hi ha a la web de l'assignatura) de salaris de treballadors del sector de banca.

Imagina que hi ha un plet en aquesta empresa sobre discriminació per raó de gènere. De fet, el salari actual dels homes és molt superior al de les dones.

Fes les anàlisis de regressió simple i múltiple, amb les interpretacions i arguments corresponents, que aporten evidència empírica en defensa de l'associació que ha iniciat el plet contra l'empresa.

1. Inicialment fes un breu descriptiu (freqüències, mitjanes,...) de les variables del fitxer segons siguin categòriques o contínues.
2. Visualitza un scatterplot (gràfic de dispersió) de cadascuna de les variables contínues respecte a la variable del salari actual. Comenta la linealitat o no, de les relacions observades.
3. Calcula la correlació de Pearson de cadascuna de les variables contínues i la variable salari actual.
4. Transforma la variable Sexo a una variable numèrica.
5. Utilitza una regressió lineal simple per tal de verificar que efectivament hi ha diferències salarials entre homes i dones.
6. Explora ara diferents models lineals múltiples usant com a variable depenent Salari Actual i com a regressors del model la variable sexe i altres factors de control (Nivell educatiu, Mesos des del contracte, experiència prèvia,...) .
7. Un cop escollit el model anota quins són els coeficients del model, el seu nivell de significació, quin nivell d'ajust ha obtingut el model, això com l'histograma dels residus i el gràfic de dispersió dels valors ajustats sobre els residus. Comenta breument els diferents resultats
8. Incorpora en el model la variable categòrica Categoria laboral i reporta els mateixos elements que en el punt anterior.
9. Escribeu quines són les conclusions principals que s'extreuen del model pel que fa a la qüestió de la discriminació del salari.