

Ejercicios Trabajo Práctico 2

Fundamentos de Análisis de Datos

Realizar un informe con los ejercicios solicitados. Entregar el informe en formato electrónico PDF y los programas en R, ambos con nombre, enviarlos por mail a mgambini@itba.edu.ar.

1. Cargar los datos del archivo `data_multiple_regression_exercice.csv` que corresponden a datos de medidas de circunferencias corporales y pesos:
`data.body<-read.table(file="data_multiple_regression_exercice.csv")`. Se desea estudiar como depende la variable `weight` en función de las otras variables.
 - a) Separar el conjunto de registros aleatoriamente en dos conjuntos de datos `data.body1` y `data.body2`. El primer conjunto se utilizará para calcular los parámetros de cada modelo y el segundo se utilizará para realizar predicciones y evaluar los resultados.
 - b) Realizar el modelo de regresión lineal múltiple con todas las variables involucradas y la validación del modelo.
 - c) Realizar el modelo de regresión lineal múltiple utilizando el método de selección de variables *Mixed Selection*.
 - d) Estudiar el comando `pcr()` del paquete `pls`.
 - e) Realizar el modelo de regresión múltiple tomando las componentes principales del conjunto de variables explicativas.
 - f) Comparar los tres modelos utilizando el comando `predict()` sobre el segundo conjunto de datos `data.body2` y calculando el error cuadrático medio para cada caso.