

1 Ejercicio Práctica 3 y 4

Fecha límite de entrega: 30 de abril de 2018 23:59.

Consideremos el Data Set de Galton (en la página de la materia, el archivo **GaltonMod.csv**). Dicho conjunto de datos representa los estudios llevados durante el siglo XIX por el polifacético investigador Francis Galton a fin analizar la relación entre la altura de un niño y su padre.

1. Abra el archivo con R.
2. Proceda a graficar en un *scatter plot* las alturas de los padres y los hijos.
3. Determine si se detectan anomalías o cuestiones raras en el conjunto de datos y lleve a cabo la limpieza que considere prudente. Una vez hecho esto, nuevamente realice el *scatter plot*.
4. Realice estudios de regresión lineal para determinar la relación entre la altura de un padre y su hijo. Ensaye con diversos modelos (ejemplo, lineal, cuadrático, entre otros), considere esquemas de training/testing para seleccionar y utilice un criterio para poder concluir o aventurar el tipo de relación entre ambas variables.
5. Grafique el *scatter plot* de los datos junto con una curva representando el modelo escogido en el inciso anterior. Haga el gráfico de la curva en una grilla fina, de resolución menor o igual a 0.01.

Se puede utilizar la función **lm** o bien una implementación propia.