

## Dureza de los árboles

1. La dureza de los árboles es difícil de medir directamente, sin embargo la densidad si es relativamente fácil de medir. Por ello es de gran interés disponer de un modelo que permita predecir la dureza de un árbol (medida en libras de fuerza (lbf)) a partir de su densidad (la masa por unidad de volumen de su madera, medida en  $kgm/m^3$ ). Por este motivo se ha tomado una muestra de 36 eucaliptos australianos y se les midió su densidad ( $X$ ) y su dureza ( $Y$ ).

Considere el conjunto de datos *dureza.csv*

- a) Plantee un modelo de regresión lineal que describa la relación entre la dureza de los eucaliptos en función de su densidad.
- b) Realice el scatter plot de estas variables y superponga la recta que mejor ajusta al modelo. ¿Describe adecuadamente la relación entre ambas variables?
- c) Analice la validez de los supuestos del modelo. ¿Hay observaciones atípicas o influyentes? Investigue la función *text* de R para marcar estas observaciones en los gráficos respectivos.
- d) Ajuste un nuevo modelo sin considerar las observaciones atípicas y de alto leverage y llame a este modelo *modelo2*. ¿Observa alguna diferencia entre ambos modelos? ¿Cuáles? ¿Con qué modelo se queda?
- e) A partir del model elegido:
  - i) interprete los parámetros del mismo.
  - ii) prediga el valor de la dureza que tendrá un eucalipt para una densidad de  $47\text{ }kgm/m^3$ .
  - iii) ¿Cuál es el porcentaje de la variabilidad explicada por el modelo?
    - 1) De una estimación de la varianza del error.
  - iv) ¿Son significativos los parámetros del modelo? Plantee las hipótesis a testear, el estadístico de contraste junto con su función de distribución, el pvalor y la decisión que toma a un nivel de significación del 5 %.
  - v) De los intervalos de confianza de nivel 90 % y 95 % para los parámetros del modelo.