



Técnicas y Métodos de Ciencia de Datos. Master Data Science

CURSO 2016/17

Propuesta análisis de datos

Notas:

- El trabajo se realizará de forma individual. Se entregará un informe en el que se recoja el problema planteado y el análisis de datos realizado, incluyendo el código R utilizado. Se aportará el fichero .Rmd, así como el fichero .pdf o .html con el texto y la salida del análisis.
- La entrega del informe se hará mediante **Aula Virtual**.

En el fichero *Deportistas.csv* aparecen datos sobre estatura, peso, índice de masa corporal, etc. en una muestra de deportistas profesionales australianos.

En concreto aparecen las siguientes variables:

Sex: Sexo del deportista (0=hombre, 1=mujer);

Ht: Altura (en cm.)

Wt: Peso (en Kg.);

BMI: Índice de masa corporal;

Bfat: Porcentaje de grasa corporal;

Sport: Deporte;

- Realiza un análisis descriptivo de los datos.
- Calcula un intervalo de confianza al 98% para el BMI medio de los deportistas masculinos. Obtén también el intervalo de confianza del mismo nivel para el BMI esperado en el caso de las mujeres deportistas.
- Plantea y realiza el contraste de hipótesis que permita afirmar si hay diferencias entre el BMI medio de los deportistas en base a su sexo.
- Y con respecto al porcentaje de grasa corporal, ¿qué se puede decir?