

Aprendizaje Estadístico

Máster Interuniversitario en Tecnologías de Análisis de Datos Masivos: Big Data

Práctica 1. Evaluación Continua

Profesores: Jose Ameijeiras Alonso, Manuel F. Mucientes Molina

Análisis descriptivo de conjunto de datos

En grupos de una a tres personas deberéis obtener una base de datos y resolver las preguntas que aparecen a continuación. Deberéis entregar el *script* de R (fichero .R) con el que obtuvisteis las respuestas, así como un informe comentando los resultados obtenidos. El informe tendrá una extensión máxima de dos carillas sin contar imágenes o salidas de código. Dicho informe se deberá entregar en formato PDF. Para generar el PDF podéis usar, entre otros, RMarkdown (<https://rmarkdown.rstudio.com/>), LaTeX, MS Word, Libre Office, ...

Fecha máxima de entrega: 12 de octubre. Fecha recomendada: 2 de octubre.

Peso en la evaluación continua: 1 punto de los 5 que están asociados al primer bloque.

Para obtener la base de datos, tenéis varias opciones, hacer Web Scraping de una web (<https://www.datacamp.com/tutorial/r-web-scraping-rvest>), obtener datos de algún instituto de estadística (<https://www.ige.gal/> o <https://www.ine.es/>), usar las bases de datos disponibles en R (`library(help = 'datasets')`), ...

1. Describe el conjunto de datos y qué objetivos pretendes conseguir tras realizar el análisis estadístico.
2. Trata de obtener un subconjunto de tus datos que cumplan una cierta condición (si no eres capaz, coge el conjunto entero). ¿Cuántos datos tienes? ¿Cuál es el número de variables que tienes?
3. ¿Cómo son las variables que estás estudiando? Cuantitativas (discretas o continuas), cualitativas (nominales u ordinales).
4. Resume numéricamente tus datos. Puedes usar tablas de frecuencias y/o medidas características (posición, dispersión y forma). ¿Qué observas?
5. Resume de forma gráfica tus datos. ¿Qué observas?
6. Trata de hacer inferencia (intervalos de confianza o contrastes) con alguna de las características que consideres relevante. ¿Cuáles son tus conclusiones?