



## Fundamentos de la Ciencia de Datos. Prueba de Laboratorio 1 (PL1)

La Prueba de Laboratorio constará de dos partes:

1. Realización de un ejercicio en clase con ayuda del profesor en el que se va a realizar un análisis con R de descripción de Datos aplicando todos los conceptos vistos en el tema. Para realizar el ejercicio vamos a utilizar dos ficheros de datos:
  - 1.1. El primer fichero de datos será de tipo .txt, es decir, de texto plano, y estará formado por los datos de los satélites menores de Urano<sup>1</sup> que hemos utilizado en la descripción teórica del tema. Lo denominaremos satelites.txt. El objetivo es obtener utilizando R los valores de las mismas magnitudes cuyo valor hemos calculado de forma manual.
  - 1.2. El segundo fichero de datos será de tipo .sav, es decir, un fichero de datos procedente de SPSS, se denomina cardata.sav, y estará formado por datos de automóviles, como su consumo en mpg (millas por galón), cilindrada, aceleración, año de fabricación modelo, etc.

Para cada uno de los ficheros vamos a calcular todas las magnitudes vistas en el tema, es decir vamos a obtener las frecuencias, la media, que será la aritmética; las medidas de dispersión, desviación típica y varianza, y las medidas de ordenación, mediana y cuartiles, incluido el cuantil 54. Para los datos del fichero cardata.sav la variable analizada será mpg<sup>2</sup>.

2. Desarrollo por parte de cada grupo del enunciado y la solución de un ejercicio en el que se realice un análisis con R de descripción de Datos introduciendo modificaciones sobre el ejercicio hecho en clase (por ejemplo: los datos se leen desde un fichero generado con Excel o los ficheros que hay en un directorio se listan con la función dir()).

Se debe completar en el documento del grupo utilizando Sweave y Noweb, para realizar Literate-Programming y entregar los archivos .Rnw, .tex y .pdf generados por R con el resultado de las dos partes de la PL1. La PL1 se entregará como actividad en Blackboard. El nombre de los ficheros será: Grupo.NúmerodeGrupo-PL1

---

<sup>1</sup> (Nombre satélite, radio en km): Cordelia, 13; Ofelia, 16; Bianca, 22; Crésida, 33; Desdémona, 29; Julieta, 42; Rosalinda, 27; Belinda, 34; Luna-1986U10, 20; Calíbano, 30; Luna-999U1, 20; Luna 1999U2, 15.

<sup>2</sup> Para mpg(cardata) no hace falta obtener las frecuencias ni el cuantil 54