

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

La Prueba de Laboratorio constará de dos partes:

1. Realización de dos ejercicios con ayuda del profesor en los que se van a realizar, utilizando el entorno R, un análisis de descripción de datos y un análisis de asociación, aplicando todos los conceptos teóricos vistos en cada lección. Para realizar el ejercicio se van a utilizar dos ficheros de datos en formato de texto plano .txt:
 - 1.1. El primer fichero de datos, que se empleará para realizar el análisis de descripción de datos, estará formado por datos de una característica cualitativa, nombre, y otra cuantitativa, radio, de los satélites menores de Urano, es decir, aquellos que tienen un radio menor de 50 Km, dichos datos, los primeros cualitativos nominales, y los segundos cuantitativos continuos, son: (Nombre, 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.
 - 1.2. El segundo fichero de datos, que se empleará para realizar el análisis de asociación, estará formado por las siguientes 6 cestas de la compra: {Pan, Agua, Leche, Naranjas}, {Pan, Agua, Café, Leche}, {Pan, Agua, Leche}, {Pan, Café, Leche}, {Pan, Agua}, {Leche}.
2. Realización de dos ejercicios de forma autónoma por cada grupo de estudiantes en los que se van a realizar, utilizando el entorno R, un análisis de descripción de datos y un análisis de asociación, aplicando todos los conceptos teóricos vistos en cada lección. Para realizar el ejercicio se van a utilizar dos ficheros de datos en formato de texto plano .txt:
 - 2.1. El primer fichero de datos, que se empleará para realizar el análisis de descripción de datos, estará formado por datos de una característica cuantitativa, distancia, desde el domicilio de cada estudiantes hasta la Universidad, dichos datos, cuantitativos continuos, son: 16.5, 34.8, 20.7, 6.2, 4.4, 3.4, 24, 24, 32, 30, 33, 27, 15, 9.4, 2.1, 34, 24, 12, 4.4, 28, 31.4, 21.6, 3.1, 4.5, 5.1, 4, 3.2, 25, 4.5, 20, 34, 12, 12, 12, 12, 5, 19, 30, 5.5, 38, 25, 3.7, 9, 30, 13, 30, 30, 26, 30, 30, 1, 26, 22, 10, 9.7, 11, 24.1, 33, 17.2, 27, 24, 27, 21, 28, 30, 4, 46, 29, 3.7, 2.7, 8.1, 19, 16.
 - 2.2. El segundo fichero de datos, que se empleará para realizar el análisis de asociación, estará formado por las siguientes conjuntos de extras incluidos en 8 ventas de coches: {X, C, N, B}, {X, T, B, C}, {N, C, X}, {N, T, X, B}, {X, C, B}, {N}, {X, B, C}, {T, A}. Donde: {X: Faros de Xenon, A: Alarma, T: Techo Solar, N: Navegador, B: Bluetooth, C: Control de Velocidad}, son los extras que se pueden incluir en cada coche.

Se debe completar en el documento del grupo utilizando Microsoft Word, para realizar Literate-Programming y entregar los archivos .docx y .pdf 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.