

Autoevaluación 1

[0010] Representación y tabulación de datos

Enunciado

Vamos a analizar los datos que se encuentran en el fichero `tiendas.csv` sobre tiendas de campaña.

1 Descarga el fichero de datos `tiendas.csv` situado en <http://gauss.inf.um.es/datos/> en local o localiza la url donde se encuentra.

Lee el fichero (desde local o desde la url) usando correctamente los argumentos de la función `read.table()`: `header`, `sep`, `dec`. Guarda los datos en una variable llamada `tiendCamp`.

2. Utilizando una sola función responde a las siguientes preguntas:

¿Cuántas tiendas de campaña componen el conjunto de datos? ¿Cuántas variables se han recogido?

XX tiendas de campaña y XX variables

3. Utilizando una sola función responde a las siguientes preguntas:

- ¿Qué variables del conjunto de datos son del tipo...

1. Numérico:
2. Entero:
3. Factor:

- ¿Cuántas marcas de tiendas hay?

1. Numérico: XX

2. Entero: XX

3. Factor: XX Hay XX marcas

4. Selecciona las variables cuantitativas, guárdalas en una variable de nombre `tiendCampC` y haz un resumen de las variables.

Con dicha información, responde:

¿Qué precio tienen de media las tiendas de campaña? ¿Cuál es el peso máximo que alcanzan?

5. Crea dos conjuntos de datos denominados `mochi` y `nomochi`, uno con las tiendas de uso mochilero y otro con las tiendas del resto de usos. ¿Son más caras en media las tiendas para mochileros o las de no mochileros?

6. Añade una variable categórica al conjunto de datos que se llame `pesocat` que provenga de la recodificación de la variable `peso` con los siguientes valores:

- “ligera”: para cuando pesan menos de 2000 kg
- “media”: cuando pesa entre 2000 y 4000 kg (ambos inclusive)
- “pesada”: cuando pesa más de 4000 kg

Ayuda: se vió en la sesión 2 de introR (el curso anterior) cómo hacer la recodificación de variables

7. Utiliza Quick-R <https://www.statmethods.net/graphs/creating.html> u otros recursos online para dibujar un gráfico de dispersión simple que enfrente las variables `peso` y `altura`.

8. Utiliza las ayudas y tu creatividad para mejorar dicho gráfico. Por si necesitas inspiración, puedes:

- Etiquetar los ejes con nombres adecuados.
- Poner un título representativo.
- Especificar el rango de la variable peso entre 0 y 10000.
- Especificar el rango de la variable altura entre 80 y 220.
- Pintar los puntos por colores según la variable `uso`.
- Utilizar distintos símbolos en función de la capacidad de la tienda.
- Buscar en la ayuda de la función `par()` cómo cambiar el tamaño y color del título.
- Crear un modelo de regresión lineal que se ajuste a la nube de puntos. Utiliza para ello la función `lm(y ~ x, data = tiendCamp)`. Guarda el modelo en una variable y añádelo al gráfico con la función `abline()`

¿Qué entregar?

Para realizar esta tarea se utilizará Rmarkdown, por lo tanto, se entregará un fichero .Rmd que contenga tanto el código necesario para responder a las preguntas como la respuesta a las mismas.