

‘Prueba de Evaluación R’

Autor: _____

Actualizado septiembre 2022

Pregunta 1 (20 puntos): Un boleto del sorteo de la ONCE consta de dos partes, la primera es un número de 4 dígitos y la segunda es un número de tres dígitos que forman la serie del boleto.

Aquí consideramos sólo el número, por ejemplo,

0	2	0	9
---	---	---	---

Se pide:

- Genera todos los números que entran en el sorteo de la ONCE y mostrarlos con los cuatro dígitos.
- ¿Cuál es la suma de los números de un boleto que más se repite?

Pregunta 2 (20 puntos): En la carpeta covid_19 hay una serie de archivos sobre el covid-19 en España (Fuente:

<https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/situacionActual.htm>).

Se pide:

- Leer los archivos “datos_provincias.csv”, “CodProv.txt*” y “CodCCAA.dat “. Añade el código de la comunidad autónoma al fichero “datos_provincias.csv” (no manualmente).
- Selecciona los datos de la comunidad autónoma que te corresponda. Para saber cuál es tu comunidad autónoma realiza la siguiente operación

DNI o Pasaporte mod 17 por ejemplo (12345678 %% 17 = 6 → Castilla y León)
hay que seleccionar las provincias de Castilla y León

Gráficos

- Realizar un gráfico que muestre adecuadamente la evolución de los casos nuevos. Justifica el gráfico elegido.
- Presenta en único gráfico la evolución de las distintas variables (columnas) por medio de un gráfico de líneas múltiples. Utiliza diferentes colores y añade una leyenda muestre el origen de cada línea.

Pregunta 3 (30 puntos): Una tienda vende una serie de artículos cuyo precio y volumen de ventas están recogidos en el archivo “articulo.xlsx”. Al llegar el periodo de rebajas se plantea aplicar un descuento para lo cual decide clasificar los artículos en tres tipos A, B y C. Los artículos se clasifican en función del ingreso bruto que proporcionan ($\text{ingresobruto} = \text{PVP} \times \text{cantidad}$), de forma que el 20% de los que más aportan se clasifican de tipo A, el 30% siguiente de Clase B y el resto de clase C.

El porcentaje de rebaja que aplicará en función de esta clasificación se encuentra en el archivo “descuento_aplicar.txt”.

Se pide:

- Crear dos data frame de nombres, uno de nombre **articulo** leyendo la información de “articulo.xlsx” y el otro de nombre **descuento** que guarde la información de “descuento_aplicar.txt”. (5 ptos)
- Crear una variable llamada tipo que clasifique los artículos en los tipos A, B y C en la forma indicada y añadirlo al data frame articulo y calcula, **en una única sentencia**, el total de ingresos brutos por cada tipo de artículo.
- Unir los dos data frame, articulo y descuento, de forma adecuada, para crear el data frame clientes. Calcular la variable $\text{nuevo_pvp} = \text{pvp} - \text{cantidad a descontar}$. Suponiendo que el volumen de ventas se mantiene constante dar una estimación del porcentaje de decremento en los ingresos brutos al aplicar los descuentos.

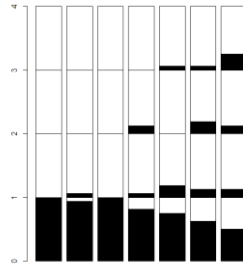
Pregunta 4 (30 puntos): La siguiente tabla representa puntuaciones de sensación de ardor para 16 sujetos en un estudio para probar un nuevo hidrogel. La primera columna da el número del sujeto. Las siguientes columnas dan la puntuación de sensación de ardor (en una escala de 1 a 4) para semanas 1 (S1) a 7 (S7). (La matriz de datos se encuentra en “matriz.R” , no se puede abrir el fichero y copiar su contenido en el script)

Nr	S1	S2	S3	S4	S5	S6	S7
1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	2
3	1	1	1	1	1	2	3
4	1	1	1	1	1	3	4
5	1	1	1	1	2	3	3
6	1	1	1	1	1	1	1
7	1	1	1	3	4	2	2
8	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1
10	1	1	1	1	1	1	4
11	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1
13	1	2	1	3	2	3	4
14	1	1	1	2	2	4	4
15	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1

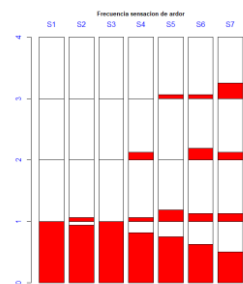
Propondremos una forma interesante de mostrar estos datos.

1.- Para la semana S_7 , calcule el vector $(f_1, 1 - f_1, f_2, 1 - f_2, f_3, 1 - f_3, f_4, 1 - f_4)$ donde f_i es la frecuencia de la modalidad $i \in \{1, 2, 3, 4\}$ observada en la semana S_7 sobre los 16 sujetos. (Sugerencia: use las funciones `tabulate()`, `cbind()`, `t()` y `as.vector()`). Ahora, use la función `apply()` para hacer el mismo cálculo para todas las demás semanas. Almacene el resultado en una matriz.

2.- Utilice la función `barplot()` y el argumento `col = c("black", "white")` en esta matriz. El gráfico que se obtiene ofrece una descripción general de la evolución de la Sensación de ardor con el tiempo.



3.- Cambie el gráfico anterior para que las barras que representan las frecuencias estén en rojo. Los números de las semanas deben estar en azul y en la parte superior del gráfico en lugar del fondo. Los números de modalidad deben estar a la izquierda, en azul. Agrega un título al gráfico



Instrucciones generales:

1. Es obligatorio enviar el código R que utilices en un fichero ASCII (texto plano).
2. Si el ejercicio lo requiere, se adjuntará un documento en pdf que explique lo que se ha hecho y lo que se ha obtenido.
3. Se puede sustituir los documentos (1) y (2) por un único documento en markdown.
4. Es obligatorio utilizar la codificación UTF-8 en los archivos *.R y *.Rmd.
5. Se valorará positivamente el uso de código vectorizado y negativamente el uso de bucles cuando estos puedan ser sustituidos por un código vectorizado.
6. Si se utiliza alguna librería debe decirse cual es y las ventajas respecto al código base.
7. Siempre que se pueda utilizar las funciones de R Base.
8. Si la función que necesita no existe se recomienda crear funciones para responder a las preguntas (si es posible y necesario).
9. Las funciones deben ser los más generales posibles.