

TAREA CURSO BASICO DE R

Crear un archivo "tarea_curso.r" y colocar las preguntas a modo de comentario y las respuestas con el código. Debe enviar el archivo con las respuestas al correo danny.murillo@utp.ac.pa hasta el 11 de marzo de 2019.

Conceptos básicos

1. Crear 4 variables con los nombres "var_integer, var_double, _var_character, var_boolean", cuyo tipo: se integer, double, character, boolean.
2. Verificar el tipo de las 4 variables.
3. Duplicar las cuatro variables o colocar los nombres "var_integer2, var_double2, _var_character2, var_boolean2".
4. Borrar las variables "var_integer, var_double, _var_character, var_boolean".

Estructura de datos

5. Crear un vector atómico de 11 elementos llamado "suma_valores" con los siguientes valores (5 6 7 8 9 10 11 12 13 14 15) y multiplicarlo * 5, el resultado debe ser el siguiente vector (25 30 35 40 45 50 55 60 65 70 75).
6. Sume ambos vectores y con el resultado cree el vector "suma_valores"
7. Utilizando el vector "suma_valores" identifique a través de funciones: número de datos de un vector, suma de los datos del vector, la media de los datos del vector, la mediana de los datos del vector.
8. Importe la tabla "titanic.csv"
9. Haga un resumen de los datos y la estructura utilizando las funciones str() y summary().
10. Transformar las variables Survived, Sex, Pclass, Embarked a tipo factor.
11. Cree un nuevo dataframe de los datos importados de Titanic solo con las siguientes columnas (Survived, Pclass, Sex, Age, Cabin) usar el nombre de titanic_resumen.

Limpieza de datos

12. Utilice la función duplicated o unique para verificar si los datos de titanic están duplicados, el resultado puede arrojar 0 datos duplicados.
13. Verifique a través de la función **complete.cases**, cuales son los datos completos de titanic, asignar resultados a un nuevo dataframe, "resultados titanic".
14. Utilizando la función is.na() extraiga los datos que no están con celdas vacías.

Gráficos en R

15. Cree los siguientes gráficos en R, utilizando alguna de las variables categóricas: barplot, boxplot, histograma, colocar títulos a los gráficos.
16. En el diagrama de boxplot, muestre los datos atípicos de la variable utilizada (si existen)
17. Cree un gráfico de dispersión con las variables **titanic3\$Fare, titanic3\$Age**.
18. Calcule la media y la mediana de la variable FARE y gráfiquela en el gráfico de dispersión con la función abline().

19. Muestre 3 de los cuatro gráficos creados en una sola ventana utilizando la función **par(mfrow=c(1,3))** en caso de enviar error utilizar , utilizar función dev.off() para resetear área de trabajo previo a la función par.
20. Si se crea un segundo módulo de este curso, qué estaría interesado aprender o que incluyéramos en el contenido ?