Tarea 1

Entrega 2 julio

Temas: vectores, dataframes, casos de estadística, dplyr, limpieza de datos.

1. Cree un vector que contenga las notas de 10 estudiantes, debe utilizar la función names(), además aplique un plot.

Estudiantes = c (59,62,69,95,100,26,85,89,47,98)

names (Estudiantes) = c ("Carlos", "Maria", "Oscar", "Jesus", "Fernanda","Jose", "Edwin","Orlando", "Tanya", "Natalia")

names (Estudiantes)

plot(Estudiantes)

1. Genere una base de datos en XLS, impórtela (trate que la base de datos contenga columnas cuantitativas, es decir campos como, por ejemplo, edad, año de nacimiento, factores, cantidad de ventas, cantidad de compras, boolean), genere un proceso exploratorio y de transformación de datos, donde utilice las siguientes funciones:

**#filter: Buscar los materiales del sector de Finanzas**

View(Finanzas)

Finanzas = data.frame(financials)

View (filter (Finanzas, Sector == "Materials"))

**#Group by: Saber el precio máximo de cada sector de finanzias**

dataAgrupada = Finanzas%>%group\_by(Sector)%>%summarise(max(Price))

**#Arrege: Ordenar los caracteres por su precio**

View(arrange(Finanzas, Price))

**#Mutate: Restar el precio de la venta de los prodctos para saber su precio**

#final

Finanzas = mutate(Finanzas, Final.Price = Price - Price.Sales)

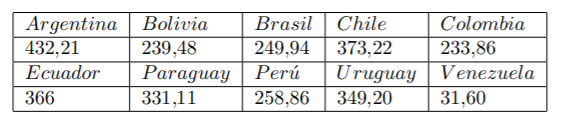
View(Finanzas)

**#Select: ver solo los datos mas relevantes de la tabla.**

View(select(Finanzas, Symbol, Name, Sector, Price, Final.Price))

1. Una investigación realizada por la Universidad Hispanoamérica detalla que los salarios mínimos para ciertos países se conforman de la siguiente forma:

**salarios=c(432.21,239.48,249.94,373.22,233.86,366,331.11,258.86,349.20,31.60)**



Salarios=c(432.21,239.48,249.94,373.22,233.86,366,331.11,258.86,349.20,31.60)

names(Salarios) = c("Argentina","Bolivia","Brasil","Chile","Colombia",

"Ecuador","Paraguay","Perú","Uruguay","Venezuela")

mean (Salarios)

sd (Salarios)

min(Salarios)

max(Salarios)

Texto

Descripción generada automáticamente

1. De la base de datos proporciona, genere las estructuras donde muestre la cantidad de facturas hechas por empleado, además muestre el nombre de los empleados. Ejemplo:

|  |  |  |
| --- | --- | --- |
| ID Empleado | Cantidad | Nombre |
| 1 | 2 | Brayner |
| 2 | 14 | Gerardo |
|  |  |  |

FacturadeEmpleados = data.frame(Facturas)

NombredeEmpleados = data.frame(Empleados)

View(merge(FacturadeEmpleados, NombredeEmpleados, by = "EmployeeID"))

Captura de pantalla de computadora

Descripción generada automáticamente

1. A continuación, se les muestra el ingreso de unidades de sacos de café por un periodo de tiempo son: 7, 6, 5, 8, 7, 7, 7, 5, 6, 8. Calcular la desviación estándar.

UnidadesCafe = c (7,6,5,8,7,7,7,5,6,8)

sd (UnidadesCafe)

Texto

Descripción generada automáticamente

1. Considere los datos en el ejercicio 5. Hallar los percentiles 0, 25, 50, 75 y 100.

quantile(UnidadesCafe, prob = c(0,0.25,0.50,0.75,1))

Un conjunto de letras blancas en un fondo blanco

Descripción generada automáticamente con confianza media

1. Realice una pequeña investigación sobre qué es stringr de la colección de paquetes Tidyvers

Strings no son componentes glamorosos y de alto perfil de R, pero desempeñan un papel importante en muchas tareas de limpieza y preparación de datos. El paquete de larguero proporciona un conjunto cohesivo de funciones diseñadas para hacer que trabajar con cadenas sea lo más fácil posible. Si no está familiarizado con las cadenas, el mejor lugar para comenzar es el capítulo sobre cadenas en R para ciencia de datos.

stringr se basa en stringi, que utiliza la biblioteca ICU C para proporcionar implementaciones rápidas y correctas de manipulaciones de cadenas comunes. stringr se centra en las funciones de manipulación de cadenas más importantes y comúnmente utilizadas, mientras que stringi proporciona un conjunto completo que cubre casi cualquier cosa que pueda imaginar. Si encuentra que al larguero le falta una función que necesita, intente buscar en stringi. Ambos paquetes comparten convenciones similares, por lo que una vez que haya dominado el larguero, debería encontrar stringi igualmente fácil de usar.