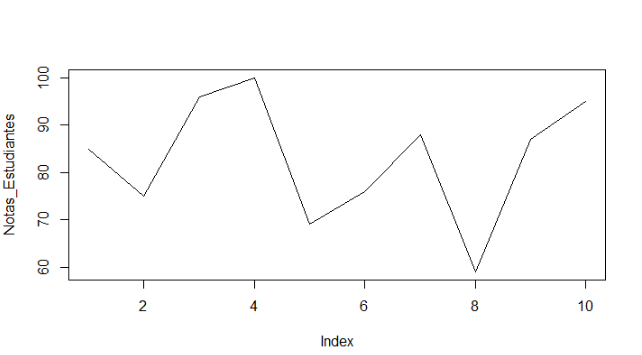
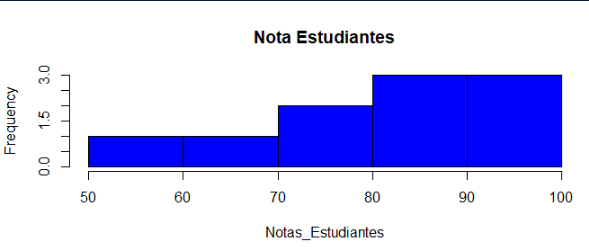
Tarea 1 . Jensy Orozco

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.

* Notas\_Estudiantes=c(85,75,96,100,69,76,88,59,87,95)
* names(Notas\_Estudiantes)=c("Eduardo","Marcela","Valeria","Victor","Emerson","Daniela", "Nicole","Carlos","Sandra","Issac")
* plot(Notas\_Estudiantes,type = "l")
* hist(Notas\_Estudiantes,main = "Nota Estudiantes",col="blue")





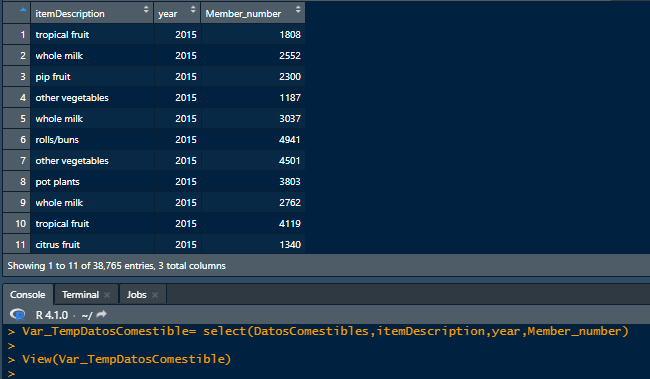
1. Genere una base de datos en XLS, impórtela (trate que la base de datos contenga columnas cuantitativas, es decir campos como, por ejemplo, edad, año de nacimiento, factores, cantidad de ventas, cantidad de compras, boolean), genere un proceso exploratorio y de transformación de datos, donde utilice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

NOTA: Archivo que tome es el *"Groceries data*"

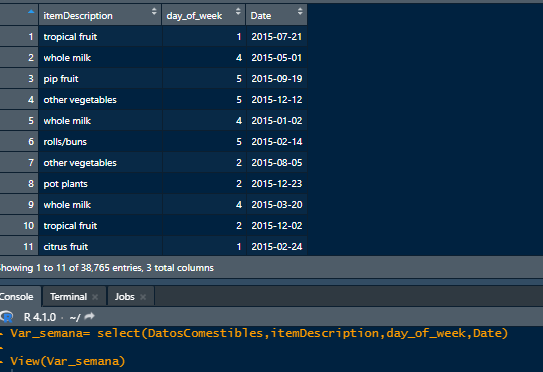
DatosComestibles= data.frame(DatosComestibles)

1. Guardar datos en variables temporales, es decir hijas de la variable de importación:

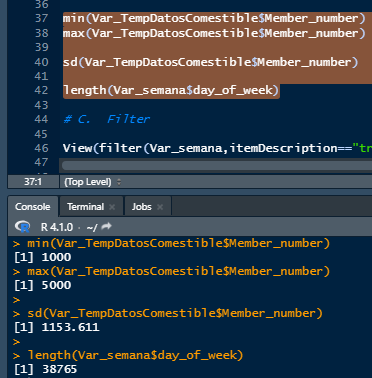
* Var\_TempDatosComestible= select(DatosComestibles,itemDescription,year,Member\_number)
* View(Var\_TempDatosComestible)



* Var\_semana= select(DatosComestibles,itemDescription,day\_of\_week,Date)
* View(Var\_semana)

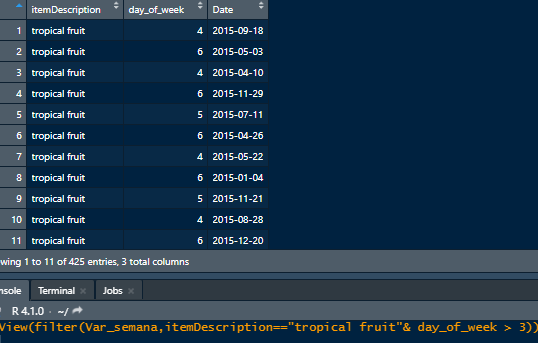


1. Al menos 2 operaciones aritméticas entre vectores :



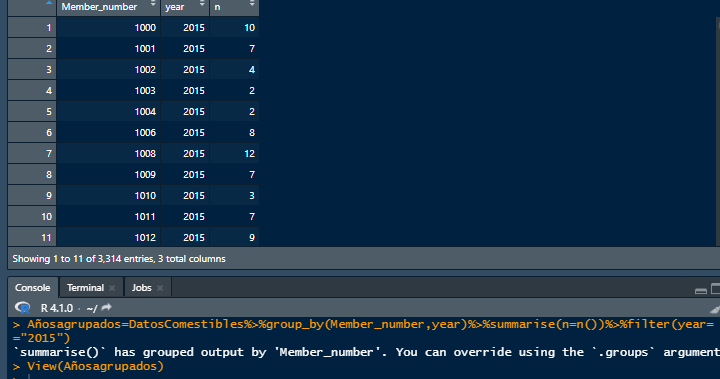
1. Filter :

View(filter(Var\_semana,itemDescription=="tropical fruit"& day\_of\_week > 3))



1. Group by :

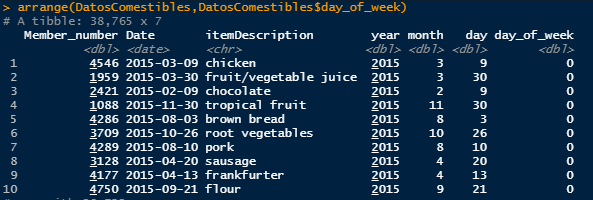
Añosagrupados=DatosComestibles%>%group\_by(Member\_number,year)%>%summarise(n=n())%>%filter(year=="2015")

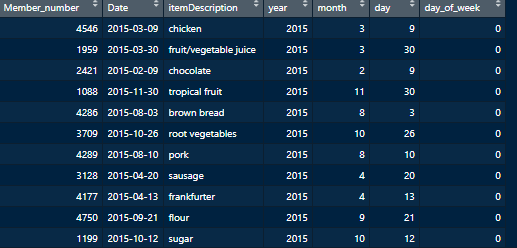


1. Arrege :

arrange(DatosComestibles,DatosComestibles$day\_of\_week)

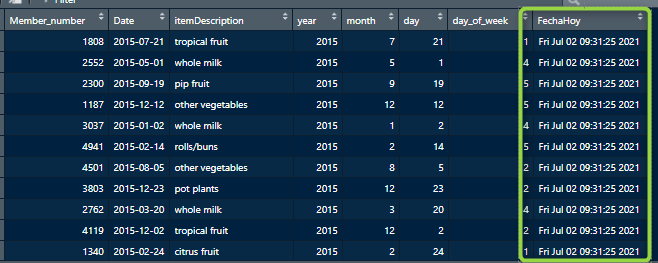
view(arrange(DatosComestible,DatosComestible$day\_of\_week))



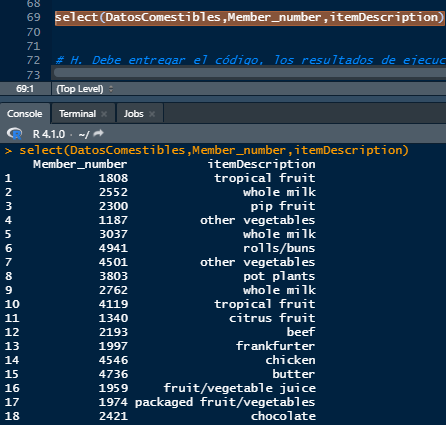


1. Mutate:

DatosComestibles = mutate ( DatosComestibles , FechaHoy = date ())



1. Select :



1. Debe entregar el código, los resultados de ejecución y una pequeña descripción de la idea que expuso en el modelo exploratorio

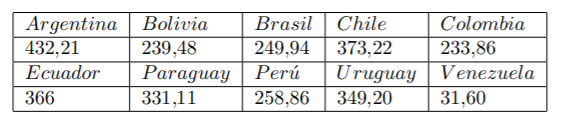
Con las ejecuciones anteriores lo que se denota es cuales productos

se consumieron por cada número de miembro y la cantidad de veces al año que lo

hicieron. A la vez los productos menos consumidos por semana.

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

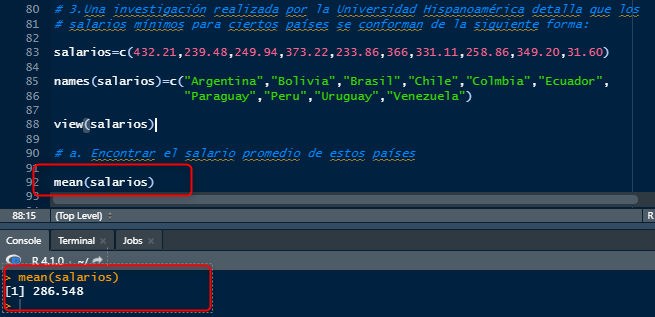
**salarios=c(432.21,239.48,249.94,373.22,233.86,366,331.11,258.86,349.20,31.60)**

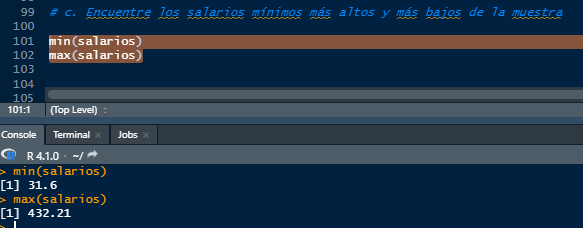


a. Encontrar el salario promedio de estos países.

b. Encontrar la desviación estándar de la muestra.

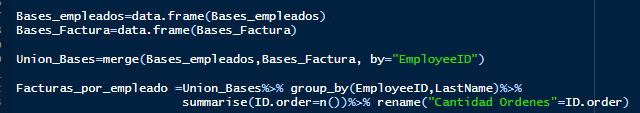
c. Encuentre los salarios mínimos más altos y más bajos de la muestra

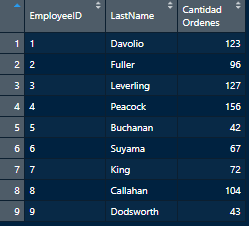




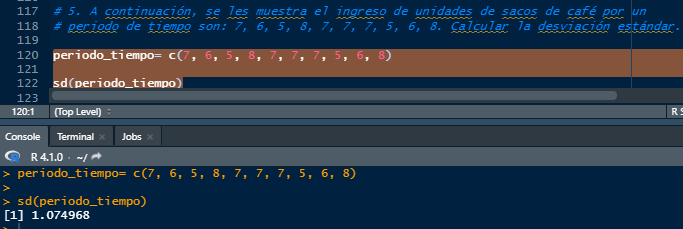
1. De la base de datos proporciona, genere las estructuras donde muestre la cantidad de facturas hechas por empleado, además muestre el nombre de los empleados. Ejemplo:

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

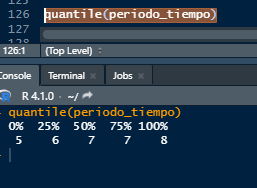
R/ 



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.



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



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

El modificar variables de texto para hacer que las variables tengan sentido humano es algo muy importante. Stringr (Cadena de caracteres) proporciona un conjunto cohesivo de funciones diseñadas para hacer que trabajar con cadenas sea lo más fácil posible.

Se puede crear una cadena utilizando comillas simples o dobles. A diferencia de otros lenguajes

Hay cuatro familias principales de funciones en stringr:

1. Manipulación de caracteres: estas funciones le permiten manipular caracteres individuales dentro de las cadenas en vectores de caracteres.
2. Herramientas de espacios en blanco para agregar, eliminar y manipular espacios en blanco.
3. Operaciones sensibles a la configuración regional cuyas operaciones variarán de una ubicación a otra.
4. Funciones de coincidencia de patrones. Estos reconocen cuatro motores de descripción de patrones. La más común son las expresiones regulares, pero hay otras tres herramientas.

El prefijo común str\_ es particularmente útil si utilizas RStudio, ya que al escribir str\_ se activa el autocompletado, lo que te permite ver todas las funciones.

Ejemplo: str\_length():