Prática

2022-10-01

# Carga de las bibliotecas

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggplot2)

# Carga de los datos

df\_facturas = read.csv('facturas.csv', sep=',')  
df\_productos = read.csv('productos.csv', sep=',')  
df\_productos\_vendidos = read.csv('productos\_vendidos.csv', sep=',')  
df\_productos\_vendidos = df\_productos\_vendidos %>%   
 mutate(idfactura = as.factor(idfactura),  
 idproducto = as.factor(idproducto))

# 1 a) Por cada cliente, halle la cantidad de facturas que generaron y ubique a quienes tengan exactamente 2 facturas. ¿Cuántos clientes son? ¿Cuáles son sus códigos de cliente?

df\_facturas %>%   
 group\_by(idcliente) %>%   
 summarise(Frecuencia = n()) %>%   
 filter(Frecuencia == 2)

## # A tibble: 2 × 2  
## idcliente Frecuencia  
## <int> <int>  
## 1 2 2  
## 2 4 2

Se encontraron a dos clientes que generaron dos facturas, los códigos de estos clientes son **2** y **4**.

# 1 b) Por cada vendedor, halle la cantidad de facturas y ordénelos de manera ascendente según su cantidad de ventas. Reporte el resultado e indique quiénes son los 3 vendedores con mayor cantidad de facturas generadas.

df\_facturas %>%   
 group\_by(idvendedor) %>%   
 summarise(Frecuencia = n()) %>%   
 arrange(Frecuencia)

## # A tibble: 7 × 2  
## idvendedor Frecuencia  
## <int> <int>  
## 1 1 1  
## 2 3 3  
## 3 7 3  
## 4 6 4  
## 5 5 5  
## 6 2 7  
## 7 4 7

Los 3 vendedores con mayores ventas son los vendedores 5, 2 y 4, cada uno de ellos con 5, 7 y 7 ventas respectivamente.

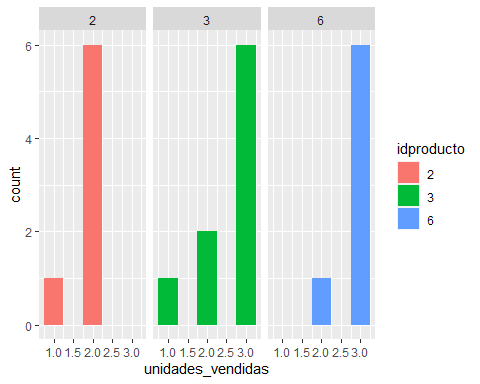
# 2 a) Mantenga todos los registros cuyo id de producto sea 2, 3 o 6. Grabe el resultado en un nuevo objeto y muestre su contenido.

productos\_vendidos <-   
 df\_productos\_vendidos %>%   
 filter(idproducto %in% c(2, 3, 6))  
  
productos\_vendidos

## idfactura idproducto unidades\_vendidas  
## 1 2 2 2  
## 2 3 3 3  
## 3 6 6 3  
## 4 8 2 2  
## 5 9 3 3  
## 6 12 6 3  
## 7 14 2 2  
## 8 15 3 3  
## 9 18 6 3  
## 10 20 2 2  
## 11 21 3 3  
## 12 24 6 3  
## 13 26 2 2  
## 14 27 3 3  
## 15 30 6 3  
## 16 1 2 1  
## 17 2 3 2  
## 18 5 6 2  
## 19 13 6 3  
## 20 15 2 2  
## 21 16 3 3  
## 22 12 3 1  
## 23 13 3 2

# 2 b) Considerando el resultado anterior, cree uno o más histogramas que permitan conocer la distribución de la variable “unidades\_vendidas”, por id de producto. Muestre la(s) gráfica(s) y comente qué presentación de “unidades\_vendidas” se vende con mayor frecuencia por producto.

productos\_vendidos %>%   
 ggplot(mapping = aes(x = unidades\_vendidas, fill = idproducto)) +  
 geom\_histogram(bins = 5) +  
 facet\_wrap(~idproducto)



# 3 a) Halle el monto total de ingresos en unidades monetarias por producto (considere unidades vendidas y luego el precio) e indique cuáles son los 3 productos que más ingresos trajeron a la compañía.

left\_join(  
 x = df\_productos\_vendidos %>% mutate(idproducto = as.integer(idproducto)),   
 y = df\_productos,   
 by = "idproducto") %>%   
 mutate(ingresos = precio \* unidades\_vendidas) %>%   
 group\_by(idproducto, nombre, descripcion) %>%   
 summarise(Monto\_Total = sum(ingresos)) %>%   
 arrange(Monto\_Total)

## `summarise()` has grouped output by 'idproducto', 'nombre'. You can override  
## using the `.groups` argument.

## # A tibble: 6 × 4  
## # Groups: idproducto, nombre [6]  
## idproducto nombre descripcion Monto\_Total  
## <int> <chr> <chr> <dbl>  
## 1 1 IndVT350 ml Botella de vino tinto de 350ml 115.  
## 2 5 IndVT2000 ml Botella de vino tinto de 2000ml 343.  
## 3 3 IndVT1000 ml Botella de vino tinto de 1000ml 428.  
## 4 4 PackVT1000 ml Pack de vino tinto de 1000ml 814   
## 5 2 PackVT350 ml Pack de vino tinto de 350ml 894.  
## 6 6 PackVT2000 ml Pack de vino tinto de 2000ml 3410

Los 3 productos que más ingresos trajeron a la compañía son Pack de vino tinto de 1000ml (814.00), Pack de vino tinto de 350ml (893.75) y Pack de vino tinto de 2000ml (3410.00).