exercicio\_1.R

davi-softplan

2023-08-27

# Importação das bibliotecas  
library(data.table)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':  
##   
## between, first, last

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

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

library(ggplot2)  
library(psych)

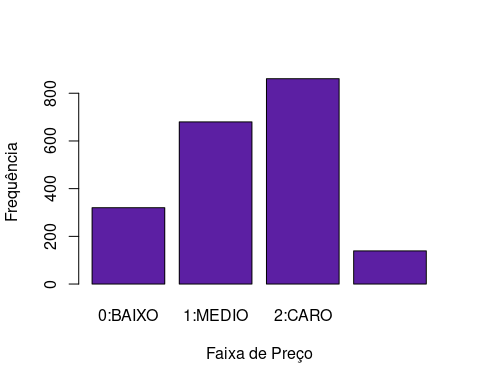
##   
## Attaching package: 'psych'

## The following objects are masked from 'package:ggplot2':  
##   
## %+%, alpha

# Base de dados  
dados <- fread(  
 input = paste0('mobile.csv'),  
 header = T,  
 na.strings = 'NA',  
 data.table = FALSE,  
 dec = ','  
)  
dados$price\_range <- recode(dados$price\_range, `0`='0:BAIXO', `1`='1:MEDIO', `2`='2:CARO', `3`='3:MUITO-CARO')  
dados$wifi <- recode(dados$wifi, `0`='0:NAO', `1`='1:SIM')  
  
# Tabela de faixa de preço  
price.tabela <- table(dados$price\_range, useNA = 'ifany')  
porc.price.tabela <- round(prop.table(price.tabela) \* 100, 1)  
price.tabela <- data.frame(price.tabela, porc.price.tabela)  
price.tabela <- price.tabela[, -3]  
colnames(price.tabela) <- c('Faixa\_Preco', 'Frequencia', 'Porcentagem')  
price.tabela

## Faixa\_Preco Frequencia Porcentagem  
## 1 0:BAIXO 320 16  
## 2 1:MEDIO 680 34  
## 3 2:CARO 861 43  
## 4 3:MUITO-CARO 139 7

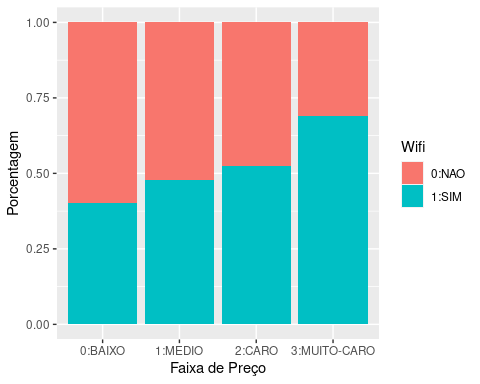
# Gráfico da tabela de faixa de preço  
barplot(  
 height = price.tabela$Frequencia,  
 names = price.tabela$Faixa\_Preco,  
 col=rgb(0.36, 0.12, 0.64, 1),  
 xlab = 'Faixa de Preço',  
 ylab = 'Frequência',  
 main = ''  
)



# Tabela de wifi x faixa de preço  
wifi.price.tabela <- table(dados$wifi, dados$price\_range)  
porc.wifi.price.tabela <- round(prop.table(wifi.price.tabela, 1) \* 100, 1)  
wifi.price.tabela <- data.frame(wifi.price.tabela, porc.wifi.price.tabela)  
wifi.price.tabela <- wifi.price.tabela[, c(-4, -5)]  
colnames(wifi.price.tabela) <- c('Wifi', 'Faixa\_Preco', 'Frequencia', 'Porcentagem')  
wifi.price.tabela

## Wifi Faixa\_Preco Frequencia Porcentagem  
## 1 0:NAO 0:BAIXO 191 19.2  
## 2 1:SIM 0:BAIXO 129 12.9  
## 3 0:NAO 1:MEDIO 354 35.5  
## 4 1:SIM 1:MEDIO 326 32.5  
## 5 0:NAO 2:CARO 409 41.0  
## 6 1:SIM 2:CARO 452 45.1  
## 7 0:NAO 3:MUITO-CARO 43 4.3  
## 8 1:SIM 3:MUITO-CARO 96 9.6

# Gráfico de relação Wifi x Faixa de preço  
ggplot(wifi.price.tabela, aes(fill=Wifi, y=Frequencia, x=Faixa\_Preco)) +  
 geom\_bar(position='fill', stat='identity') +  
 xlab('Faixa de Preço') +  
 ylab('Porcentagem')



# Tabela de faixa de preço x ram  
price.ram.tabela <- aggregate(dados$ram, by=list(dados$price\_range), FUN='mean')  
colnames(price.ram.tabela) <- c('Faixa de preço', 'RAM')  
price.ram.tabela

## Faixa de preço RAM  
## 1 0:BAIXO 771.7625  
## 2 1:MEDIO 1449.1735  
## 3 2:CARO 2947.6028  
## 4 3:MUITO-CARO 3439.8417

# Classificação baseada na RAM  
dados$ram\_interval[dados$ram <= 800] <- 'RAM1'  
dados$ram\_interval[dados$ram > 800 & dados$ram <= 1600] <- 'RAM2'  
dados$ram\_interval[dados$ram > 1600 & dados$ram <= 2400] <- 'RAM3'  
dados$ram\_interval[dados$ram > 1600 & dados$ram <= 2400] <- 'RAM3'  
dados$ram\_interval[dados$ram > 2400 & dados$ram <= 3200] <- 'RAM4'  
dados$ram\_interval[dados$ram > 3200] <- 'RAM5'  
  
# Tabela intervalo de RAM  
ram\_interval.tabela <- table(dados$ram\_interval)  
porc.ram\_interval.tabela <- round(prop.table(ram\_interval.tabela) \* 100, 1)  
ram\_interval.tabela <- data.frame(ram\_interval.tabela, porc.ram\_interval.tabela)  
ram\_interval.tabela <- ram\_interval.tabela[, -3]  
colnames(ram\_interval.tabela) <- c('Intervalo\_Ram', 'Frequencia', 'Porcentagem')  
ram\_interval.tabela

## Intervalo\_Ram Frequencia Porcentagem  
## 1 RAM1 304 15.2  
## 2 RAM2 414 20.7  
## 3 RAM3 441 22.0  
## 4 RAM4 406 20.3  
## 5 RAM5 435 21.8