

## Exercício 1

Andres, Gustavo Beretta e Luiz Adriano.

```
#Leitura da base de dados
library(data.table)
base <- fread(input = paste0("mobile.csv"), header = T, na.strings =
"NA", data.table = FALSE, dec = ",")

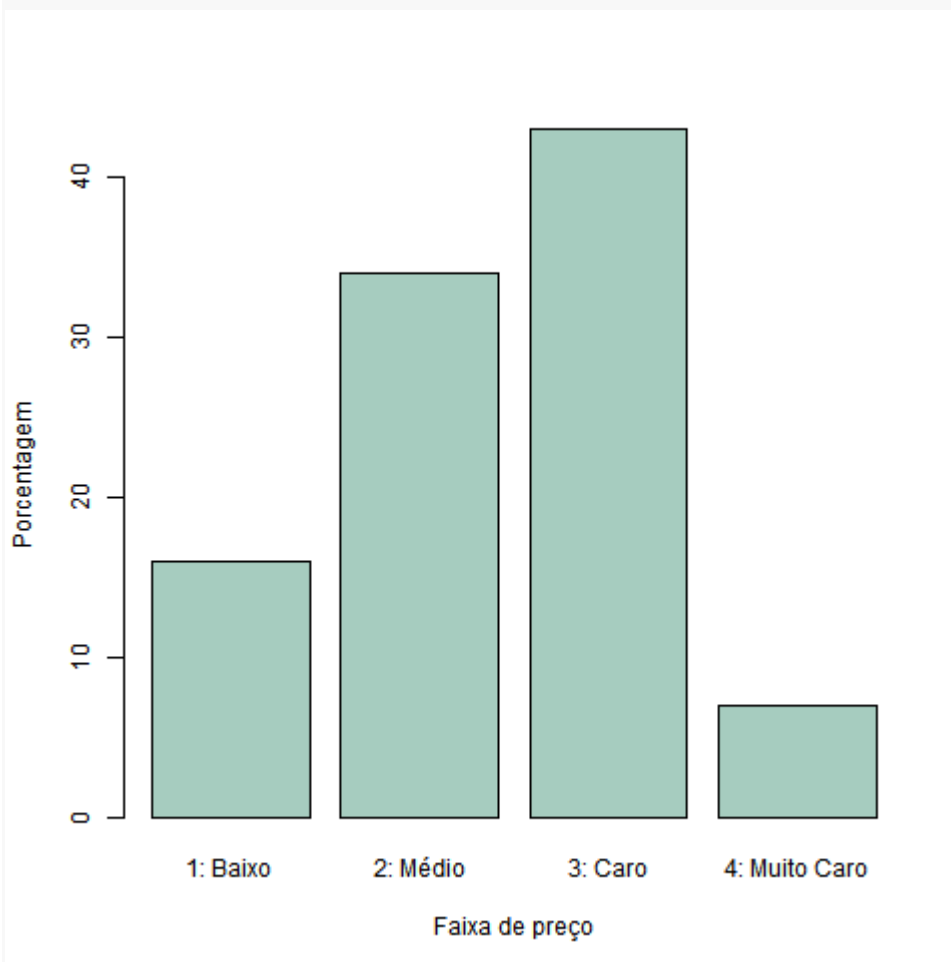
#ajuste de variaveis quantitativas e qualitativas
library(dplyr)

base$blue <- recode(base$blue, `0`="1: Não", `1`="2: Sim")
base$clock_speed <- as.numeric(base$clock_speed)
base$dual_sim <- recode(base$dual_sim, `0`="1: Não", `1`="2: Sim")
base$four_g <- recode(base$four_g, `0`="1: Não", `1`="2: Sim")
base$m_dep <- as.numeric(base$m_dep)
base$touch_screen <- recode(base$touch_screen, `0`="1: Não", `1`="2: Sim")
base$wifi <- recode(base$wifi, `0`="1: Não", `1`="2: Sim")
base$price_range <- recode(base$price_range, `0`="1: Baixo", `1`="2:
Médio", `2`="3: Caro", `3`="4: Muito Caro")

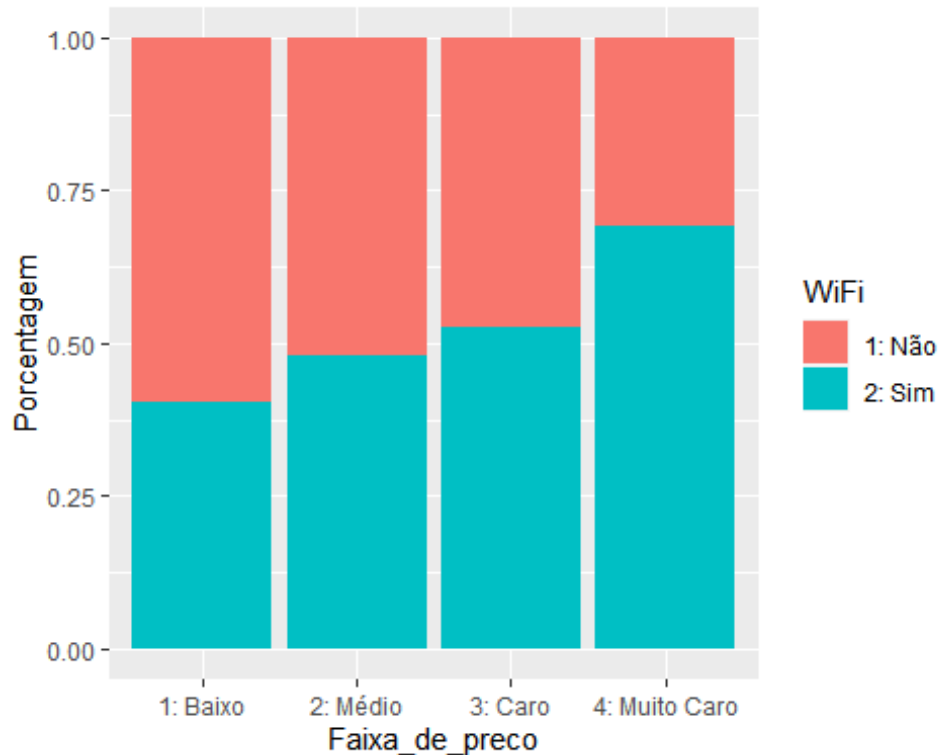
#tabela de faixa de preço
precos.freq <- table(base$price_range, useNA = "ifany")
precos.porc <- round(prop.table(precos.freq)*100,1)
precos.tabela <- data.frame(precos.porc)
colnames(precos.tabela) <- c("Faixa de preço", "Porcentagem")
precos.tabela

##   Faixa de preço Porcentagem
## 1      1: Baixo           16
## 2      2: Médio           34
## 3      3: Caro            43
## 4      4: Muito Caro        7
```

```
#grafico de faixa de preço  
png(file = "grafico_precos.png")  
grafico_precos <- barplot(height=precos.tabela$Porcentagem,  
names=precos.tabela$`Faixa de preço`,  
col=rgb(0.3,0.6,0.5,0.5),  
xlab="Faixa de preço",  
ylab="Porcentagem",  
main="")  
dev.off()
```



```
#gráfico de relação entre wifi e preço
wifi_x_preco.tabela <- data.frame(table(base$price_range,base$wifi))
colnames(wifi_x_preco.tabela) <- c("Faixa_de_preco","WiFi","Frequencia")
library(ggplot2)
ggplot(wifi_x_preco.tabela, aes(fill=WiFi, y=Frequencia,
x=Faixa_de_preco)) +
  geom_bar(position="fill", stat="identity") +
  ylab("Porcentagem")
```



```
#tabela de relação entre memória RAM e preço
ram_x_preco.tabela <- aggregate(base$ram, by=list(base$price_range),
FUN="mean")
colnames(ram_x_preco.tabela) <- c("Faixa de preço","Memória RAM")
ram_x_preco.tabela
```

```
## Faixa de preço Memória RAM
## 1 1: Baixo 771.7625
## 2 2: Médio 1449.1735
## 3 3: Caro 2947.6028
## 4 4: Muito Caro 3439.8417
```

```

#classificação por memória RAM
library(psych)

base$Grupo_Ram[base$ram <800] = "G1"
base$Grupo_Ram[base$ram >=800 & base$ram <1600] = "G2"
base$Grupo_Ram[base$ram >=1600 & base$ram <2400] = "G3"
base$Grupo_Ram[base$ram >=2400 & base$ram <3200] = "G4"
base$Grupo_Ram[base$ram >=3200] = "G5"
grupos_ram <- table(base$Grupo_Ram, useNA = "ifany")
grupos_ram

##
##  G1  G2  G3  G4  G5
## 304 414 441 406 435

```

#### 1) Classificação das variáveis:

1. battery\_power - quantitativa contínua
2. blue - qualitativa nominal
3. clock\_speed - quantitativa contínua
4. dual\_sim - qualitativa nominal
5. fc - quantitativa contínua
6. four\_g - qualitativa nominal
7. int\_memory - quantitativa contínua
8. m\_dep - quantitativa contínua
9. mobile\_wt - quantitativa contínua
10. n\_cores - quantitativa discreta
11. pc - quantitativa contínua
12. ram - quantitativa contínua
13. sc\_h - quantitativa contínua
14. sc\_w - quantitativa contínua
15. touch\_screen - qualitativa nominal
16. wifi - qualitativa nominal
17. price\_range - qualitativa ordinal

- 2) O segmento com mais celulares é o “caro”, conhecido também como intermediários premium ou custo-benefício. O segmento de topo de linha (“muito caro”) é o menor de todos.
- 3) Quanto mais caro o celular, mais comum é a presença de rede WiFi. Pelo menos 25% dos celulares ainda não têm WiFi e pelo menos 37,5% já têm.
- 4) Quanto mais caro o celular, mais memória RAM ele tem. A diferença de memória entre celulares de preços diferentes é muito grande.
- 5) Com exceção do grupo com menos de 800MB de memória RAM, todos os outros tem uma frequência muito parecida.