

## exercicio2.R

Luiz Adriano, Gustavo Bereta, Andres Kindel

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

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

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")

# AMOSTRA SIMPLES AO ACASO
amostra400 <- base[sample(nrow(base), size=400), ]

#Analise battery_power
summary(amostra400$battery_power)

##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  502.0   818.2  1224.5  1228.5  1607.0  1998.0

quantile(amostra400$battery_power, probs=0.05)
```

```

##      5%
## 568.6

quantile(amostra400$battery_power, probs=0.95)

##      95%
## 1920.15

#Analise ram
summary(amostra400$ram)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      267    1203    2193    2114    3024    3990

quantile(amostra400$ram, probs=0.05)

##      5%
## 452.95

quantile(amostra400$ram, probs=0.95)

##      95%
## 3746

library(dplyr)
battery_power_price_range <- amostra400 %>%
  group_by(price_range) %>%
  summarise(
    media = mean(battery_power),
    mediana = median(battery_power),
    percentil_5 = quantile(battery_power, 0.05),
    percentil_25 = quantile(battery_power, 0.25),
    percentil_75 = quantile(battery_power, 0.75),
    percentil_95 = quantile(battery_power, 0.95),
    minimo = min(battery_power),
    maximo = max(battery_power)
  )

ram_price_range <- amostra400 %>%
  group_by(price_range) %>%
  summarise(
    media = mean(ram),
    mediana = median(ram),
    percentil_5 = quantile(ram, 0.05),
    percentil_25 = quantile(ram, 0.25),
    percentil_75 = quantile(ram, 0.75),
    percentil_95 = quantile(ram, 0.95),
    minimo = min(ram),
    maximo = max(ram)
  )

battery_power_price_range

```

```
## # A tibble: 4 × 9
##   price_range  media mediana percentil_5 percentil_25 percentil_75
percentil_95
##   <chr>      <dbl>   <int>      <dbl>      <dbl>      <dbl>
<dbl>
## 1 1: Baixo    1069.    984      551.      704      1425
1707.
## 2 2: Médio    1207.   1201      606.      826      1489
1925
## 3 3: Caro     1270.   1303      583       894      1652
1903.
## 4 4: Muito Caro 1419.   1617      570.     1086      1830
1945.
## # i 2 more variables: minimo <int>, maximo <int>

ram_price_range

## # A tibble: 4 × 9
##   price_range  media mediana percentil_5 percentil_25 percentil_75
percentil_95
##   <chr>      <dbl>   <int>      <dbl>      <dbl>      <dbl>
<dbl>
## 1 1: Baixo     735.    651      306.      454      1018
1268.
## 2 2: Médio    1455.   1412      520.     1077      1891
2326.
## 3 3: Caro     2907.   2890     2065.     2457      3377
3783.
## 4 4: Muito Caro 3422.   3451     2505.     3235      3707
3945
## # i 2 more variables: minimo <int>, maximo <int>
```

Conclusão:

1\2. Tanto a média quanto a mediana das variáveis "Battery power" e "RAM" apresentam valores próximos, e os percentis correspondentes também demonstram uma distribuição proporcional. É possível concluir que os valores estão bem distribuídos na amostra.

3. O aumento médio na capacidade da bateria tem uma relação pequena com base no intervalo de preço. Os valores mínimos e máximos não mostram uma dependência significativa em relação à variável analisada. Por outro lado, a

capacidade de RAM está tem uma relação grande com o preço, sendo mais elevada nos segmentos de preço mais elevados.