

# Estrutura

Marco Tulio Alves de Barros

2022-09-16

## FUNÇÕES UTEIS

```
fr = function(fi, N) {  
  round((fi / N) * 100, 2);  
}  
  
MEDIA = function(vetor) {  
  media = sum(vetor) / length(vetor)  
}  
  
MEDIA_POND = function(fi, yi) {  
  yf = 0  
  for (i in 1:length(fi)) {  
    yf = fi[i] * yi[i] + yf  
  }  
  yf = yf / sum(fi)  
}  
  
czuber = function(Li, a, d1, d2) {  
  # Li = limite inferior da classe  
  # a = amplitude da classe modal  
  # d1 = diferença entre a frequência da classe modal e a anterior  
  # d2 = diferença entre a frequência da classe modal e a posterior  
  Li + (d1 / (d1+d2)) * a  
}  
  
pearson = function(fi, yi, Md) {  
  # fi = vetor das frequências absolutas  
  # Md = mediana  
  # y = média amostral (vetor dos pontos médios das classes)  
  y = MEDIA_POND(fi, yi)  
  3 * Md - 2 * y  
}  
  
MEDIANA = function(Li, N, Fac, fmd, a) {  
  Emd = N / 2; # Mediana esperada  
  # Li = limite inferior da classe  
  # N = população total  
  # Fac = freq acumulada da classe anterior  
  # fmd = qntd de elementos na classe da mediana  
  # a = amplitude da classe
```

```

Md = Li + ((Emd - Fac) / fmd) * a
}

quantil = function(Li, Fq, Fac, a, i, n) {
  # i = por cento do quantil desejado
  # Li = limite inferior da classe quantilica
  # a = amplitude das classes
  # Fac = freq acumulada da classe anterior
  # Fq = freq da classe quantilica
  Li + (((i*n - Fac)) * a / Fq);
}

```

Ex 1:

```

glicemia = c(80, 85, 86, 90, 95, 96, 99, 100, 101, 103,
             103, 103, 104, 105, 108, 108, 109, 110, 110, 110)

# a) Agrupar os dados em classes e calcular a média, moda e a mediana;
k = round(sqrt(length(glicemia))); k

```

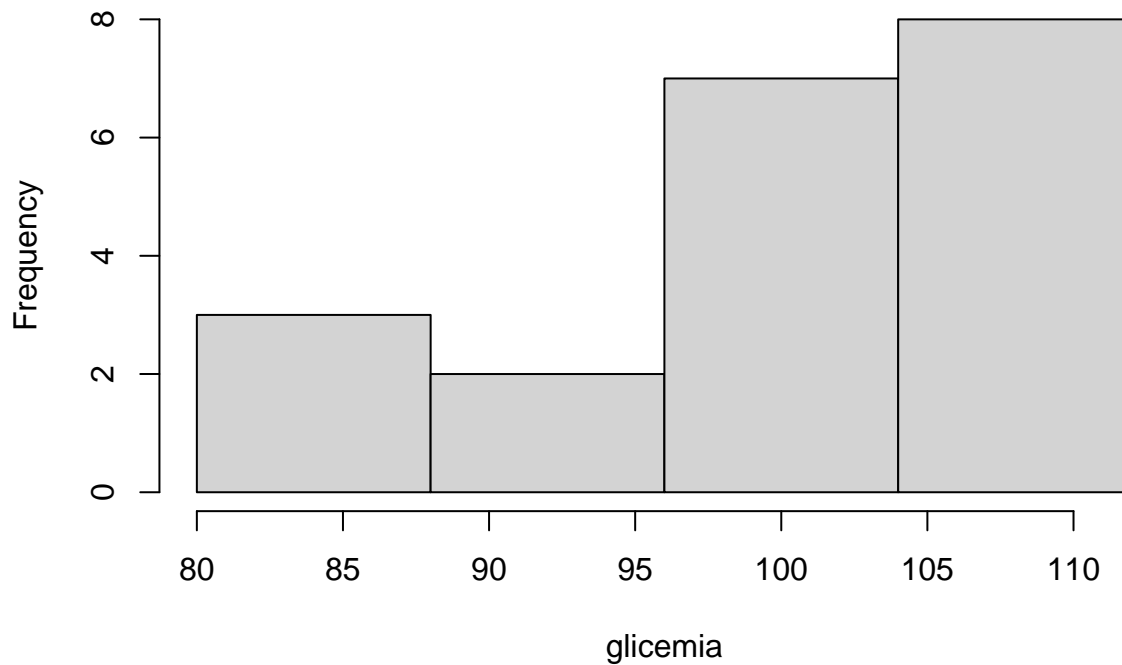
```
## [1] 4
```

```
a = round((max(glicemia) - min(glicemia)) / k); a
```

```
## [1] 8
```

```
tabela = hist(glicemia, plot=T, breaks=c(80, 88, 96, 104, 112), right=F)
```

## Histogram of glicemia



```
fi = round(tabela$count); fi
```

```
## [1] 3 2 7 8
```

```
frp = round(fi/sum(fi)*100, 2);
Fi = cumsum(fi);
Frp = cumsum(frp);
dados = cbind(fi, frp, Fi, Frp);
rownames(dados) = c("80|-88", "88|-96", "96|-104", "104|-112"); dados
```

```
##      fi frp Fi Frp
## 80|-88   3  15  3  15
## 88|-96   2  10  5  25
## 96|-104   7  35 12  60
## 104|-112  8  40 20 100
```

```
media = MEDIA(glicemia); media
```

```
## [1] 100.25
```

```
mediana = MEDIANA(96, 20, 5, 7, a); mediana
```

```
## [1] 101.7143
```

```
moda_czu = czuber(104, a, 1, 8); moda_czu
```

```
## [1] 104.8889
```

```
# b) Determine, sem agrupar em classes (dados brutos): a média, moda e mediana  
media_bruta = MEDIA(glicemia); media_bruta
```

```
## [1] 100.25
```

```
mediana_bruta = (glicemia[10] + glicemia[11]) / 2; mediana_bruta;
```

```
## [1] 103
```

```
moda_bruta = 110
```

## Ex 2:

```
l = c(48, 53, 58, 62, 64, 66, 69, 71, 77, 81,  
      49, 54, 58, 62, 64, 67, 69, 72, 77, 82,  
      50, 55, 59, 63, 65, 67, 70, 73, 78, 83,  
      52, 56, 60, 64, 65, 67, 70, 74, 78, 86,  
      52, 57, 61, 64, 66, 68, 71, 76, 80, 90)  
leitos = sort(l); leitos
```

```
## [1] 48 49 50 52 52 53 54 55 56 57 58 58 59 60 61 62 62 63 64 64 64 64 65 65 66  
## [26] 66 67 67 67 68 69 69 70 70 71 71 72 73 74 76 77 77 78 78 80 81 82 83 86 90
```

```
# a) Determine a média e mediana dos dados e, também, determine o terceiro  
# quartil e interprete-o;  
media = MEDIA(leitos); media
```

```
## [1] 66.46
```

```
mediana = leitos[25]; mediana
```

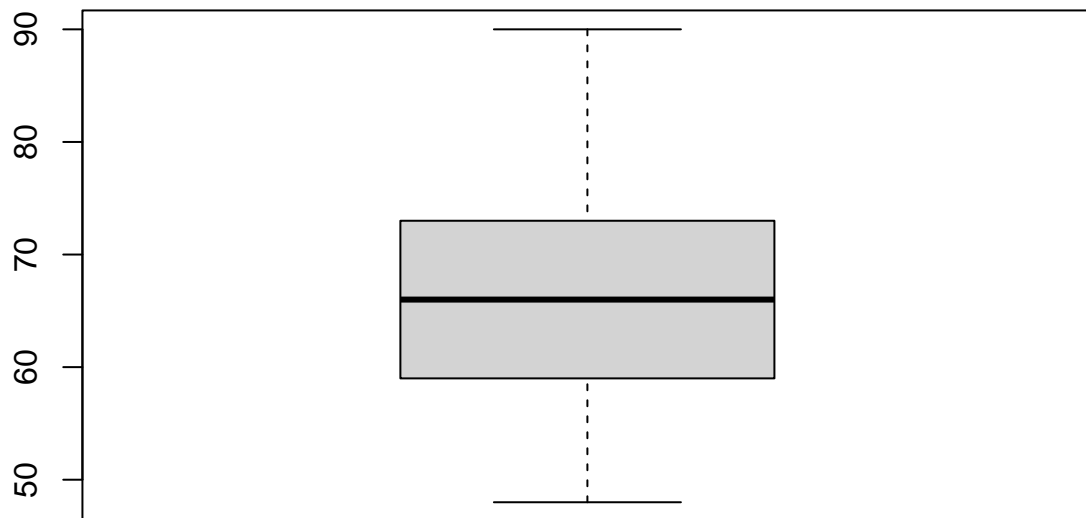
```
## [1] 66
```

```
moda = 64;  
q3 = quantil(72, 12, 24, 6, 0.75, 50); q3
```

```
## [1] 78.75
```

```
## Significa que das 50 instituições analisadas, 75% tem disponibilidade  
## de leitos inferior a 78.75; enquanto 25% dessas possuem disponibilidade  
## superior a 78.75
```

```
# b) Construa o gráfico de caixas (box plot).  
boxplot(leitos);
```



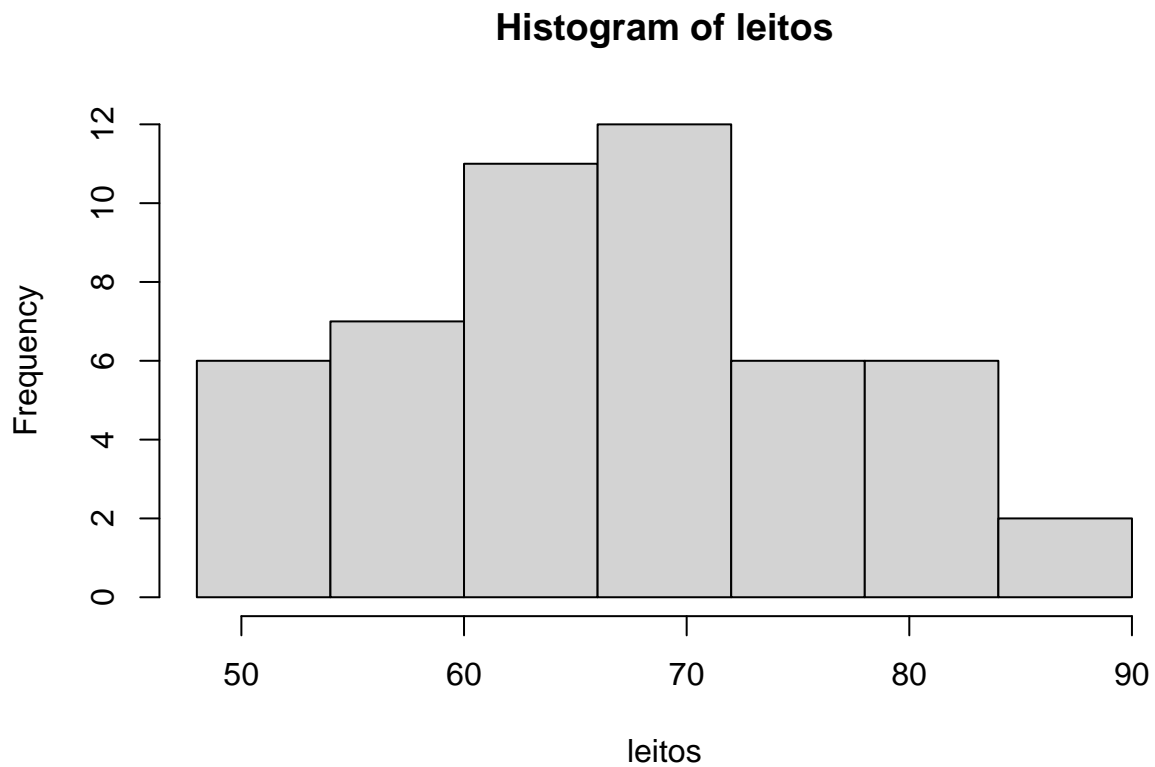
```
# c) Agrupar os dados em classes (sqrt(n)) e calcular a média, moda e a mediana.
k = round(sqrt(length(leitos))); k
```

```
## [1] 7
```

```
a = round((max(leitos) - min(leitos)) / k); a
```

```
## [1] 6
```

```
tabela = hist(leitos, plot=T, breaks=c(48, 54, 60, 66, 72, 78, 84, 90), right=F)
```



```
fi = round(tabela$count); fi
```

```
## [1] 6 7 11 12 6 6 2
```

```
frp = round(fi/sum(fi)*100, 2);
Fi = cumsum(fi);
Frp = cumsum(frp);
dados = cbind(fi, frp, Fi, Frp);
rownames(dados) = c("48|-54", "54|-60", "60|-66", "66|-72", "72|-78", "78|-84", "84|-90")
dados;
```

```
##      fi frp Fi Frp
## 48|-54 6 12 6 12
## 54|-60 7 14 13 26
## 60|-66 11 22 24 48
## 66|-72 12 24 36 72
## 72|-78 6 12 42 84
## 78|-84 6 12 48 96
## 84|-90 2 4 50 100
```

```
media = MEDIA(leitos); media
```

```
## [1] 66.46
```

```
mediana = MEDIANA(66, length(leitos), 24, 12, a); mediana
```

```
## [1] 66.5
```

```
moda = czuber(66, a, 1, 6); moda
```

```
## [1] 66.85714
```

### Ex 3:

```
cr = c(131, 119, 138, 125, 129, 126, 131, 132, 126, 128, 128, 131)
cranios = sort(cr); cranios
```

```
## [1] 119 125 126 126 128 128 129 131 131 131 132 138
```

```
len = length(cranios); len
```

```
## [1] 12
```

```
media = MEDIA(cranios);
mediana = (cranios[6] + cranios[7]) / 2; mediana
```

```
## [1] 128.5
```

```
moda_bruta = 131;
```

### Ex 4:

```
AT = 500 #  $AT = MAX(c) - MIN(c)$ 
k = 20 # qntd de classes
a = AT/k; a # amplitude das classes
```

```
## [1] 25
```

```
intervalos = c("35|-60", "60|-85", "85|-110", "110|-135", "135|-160", "...")
classe5_inf = 135;
classe5_medio = (135+12 + 135+13) / 2; classe5_medio # 24 elementos na classe
```

```
## [1] 147.5
```

### Ex 5:

```
fi = c(3, 6, 18, 35, 16, 2)
xi = c(0.1, 0.3, 0.5, 0.7, 0.9, 1.1) # ponto medio

media = MEDIA_POND(fi, xi); media
```

```
## [1] 0.6525
```

```
mediana = MEDIANA(0.6, 80, 27, 35, 0.2); mediana
```

```
## [1] 0.6742857
```

```
moda = czuber(0.6, 0.2, 17, 19); moda
```

```
## [1] 0.6944444
```

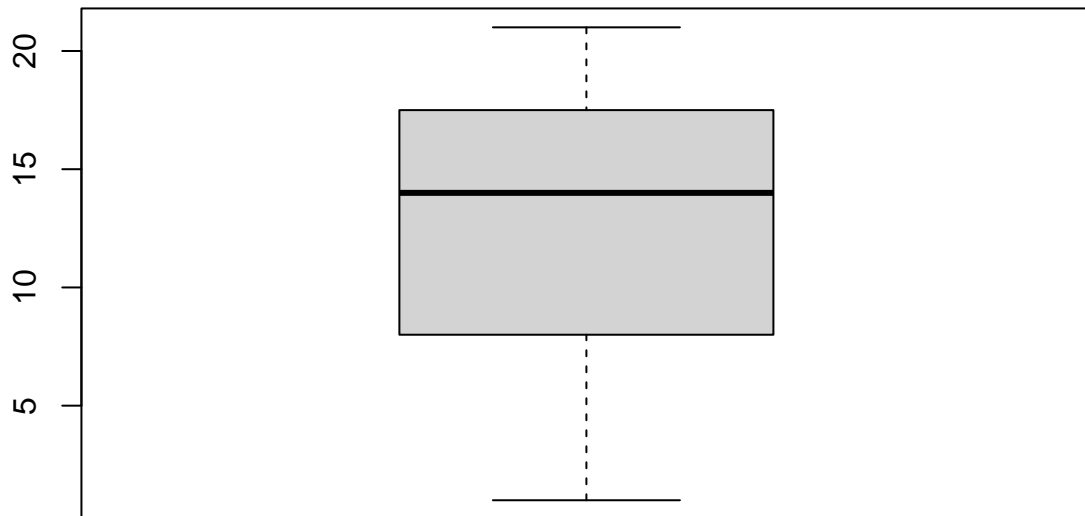
## Ex 6:

```
observacoes = c(1, 4, 7, 9, 10, 13, 15, 17, 17, 18, 19, 21)
quantile(observacoes); #q1 = 8.5; q2 = 14; q3 = 17.25
```

```
##      0%   25%   50%   75%  100%
##  1.00  8.50 14.00 17.25 21.00
```

```
boxplot(observacoes)
```





## Ex 7:

```
library(readxl)
require(readxl)
dados_rat0s = read_excel("C:/Users/marco/OneDrive/Área de Trabalho/AULAS/Estatística/dados_rat0s.xlsx")
with(dados_rat0s, boxplot(Peso ~ Idade));

# Peso médio para cada idade:
pesos = with(dados_rat0s, c(Peso));
idades = with(dados_rat0s, c(Idade));

i = 1
aux1 = aux2 = aux3 = aux4 = aux5 = 0
medias = c(0, 0, 0, 0, 0)
while(i <= length(idades)) {
  if (idades[i] == 30) {
    medias[1] = medias[1] + pesos[i]
    aux1 = aux1+1
  } else if (idades[i] == 34) {
    medias[2] = medias[2] + pesos[i]
    aux2 = aux2+2
  } else if (idades[i] == 38) {
    medias[3] = medias[3] + pesos[i]
    aux3 = aux3+1
  }
  i = i + 1
}
```

```

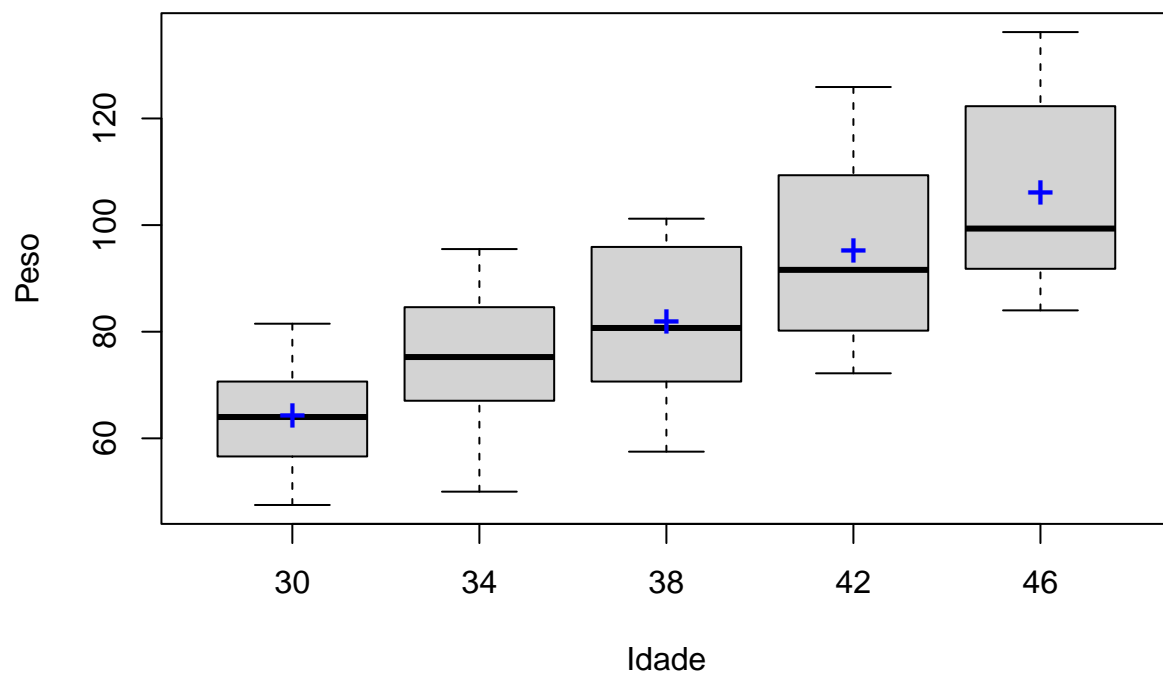
    } else if (idades[i] == 42) {
      medias[4] = medias[4] + pesos[i]
      aux4 = aux4+1
    } else if (idades[i] == 46) {
      medias[5] = medias[5] + pesos[i]
      aux5 = aux5+1
    }

    i = i+1
  }
Medias = c(medias[1]/aux1, medias[2]/aux2, medias[3]/aux3, medias[4]/aux4, medias[5]/aux5)
Medias

```

```
## [1] 63.93750 37.45625 81.65000 95.05000 105.88750
```

```
points(Medias, pch="+", cex=1.5, col="blue")
```



```
media_pesos = sum(pesos)/length(pesos); media_pesos
```

```
## [1] 84.2875
```

## Ex 8:

```
cores = c("vermelho", "azul", "verde", "amarelo", "preto", "marrom", "laranja", "roxo", "totais")
fi = c(4, 11, 4, 3, 5, 1, 1, 1, 30);
frp = c(fr(4, 30), fr(11, 30), fr(4, 30), fr(3, 30), fr(5, 30), fr(1, 30), fr(1, 30), fr(1, 30), 100)
data = data.frame(cores, fi, frp); data
```

```
##      cores fi    frp
## 1 vermelho  4  13.33
## 2   azul   11  36.67
## 3   verde  4  13.33
## 4 amarelo  3  10.00
## 5   preto  5  16.67
## 6  marrom  1   3.33
## 7 laranja  1   3.33
## 8   roxo   1   3.33
## 9  totais 30 100.00
```

```
# escolha_modal = azul, com 11 ocorrências
azul_fr = frp[2]; azul_fr
```

```
## [1] 36.67
```

## Ex 9:

```
xi = c(2, 3, 4, 5, 6); # peso das frequencias
fi = c(3, 6, 10, 6, 3);
frp = c(fr(3, 28), fr(6, 28), fr(10, 28), fr(6, 28), fr(3, 28))
Fi = c(3, 9, 19, 25, 28); # frequencias de cada valor
Frp = c(fr(3, 28), fr(9, 28), fr(19, 28), fr(25, 28), fr(28, 28))
data = data.frame(xi, fi, frp, Fi, Frp); data
```

```
##  xi fi  frp Fi   Frp
## 1  2  3 10.71  3  10.71
## 2  3  6 21.43  9  32.14
## 3  4 10 35.71 19  67.86
## 4  5  6 21.43 25  89.29
## 5  6  3 10.71 28 100.00
```

```
media_pond = MEDIA_POND(fi, xi); media_pond
```

```
## [1] 4
```

```
# mediana_bruta = 5;
mediana = MEDIANA(4, sum(fi), 9, 10, 1); mediana
```

```
## [1] 4.5
```

## Ex 10:

```
classes = c("4|-9", "9|-14", "14|-19", "19|-24");
fi = c(8, 12, 17, 3)
frp = c(fr(8, 40), fr(12, 40), fr(17, 30), fr(3, 40))
Fi = c(8, 20, 37, 40);
Frp = c(fr(8, 40), fr(20, 40), fr(37, 40), fr(40, 40))
data = data.frame(classes, fi, frp, Fi, Frp); data
```

```
##   classes fi   frp Fi   Frp
## 1    4|-9  8 20.00  8  20.0
## 2    9|-14 12 30.00 20  50.0
## 3   14|-19 17 56.67 37  92.5
## 4   19|-24  3  7.50 40 100.0
```

```
quarto_decil = quantil(9, 12, 8, 5, 0.4, sum(fi)); quarto_decil # 40%
```

```
## [1] 12.33333
```

```
primeiro_quartil = quantil(9, 12, 8, 5, 0.25, sum(fi)); primeiro_quartil # 25%
```

```
## [1] 9.833333
```

```
septuagesimo_segundo_percentil = quantil(14, 17, 20, 5, 0.72, sum(fi)); septuagesimo_segundo_percentil
```

```
## [1] 16.58824
```