#Parte Andrei

# moda  
moda <- function(x)  
{  
 ux <- unique(x)  
 tab <- tabulate(match(x, ux))  
 ux[tab == max(tab)]  
}  
  
# formula sturges  
sturges <- function(x)  
{  
 floor((1 + log(length(x)) / log(2)))  
}  
  
# amplitude sturges  
amplitude <- function(x)  
{  
 max <- max(x)  
 min <- min(x)  
 nclass <- sturges(x)  
 (max - min) / nclass  
}  
  
# coeficiente de variacao  
coeficiente\_variacao <- function(x)  
{  
 s <- sd(x)  
 m <- mean(x)  
 s/m \* 100  
}  
  
# funcao de assimetria  
assimetria <- function(x)  
{  
 n <-length(x)  
 s <- sd(x)  
 m <- mean(x)  
 1/n \* sum((x - m)^3/s^3)  
}  
  
# funcao de achatamento ou curtose  
achatamento <- function(x)  
{  
 n <- length(x)  
 s <- sd(x)  
 m <- mean(x)  
 1/n \* sum((x - m)^4/s^4) - 3  
}  
  
# funcao de oulier quantiles  
outliers\_quantiles <- function(x, o)  
{  
 sx <- sort(x)  
 n <- length(x)  
 nps <- c(round(n \* .25), round(n \* .50), round(n \* .75))  
 qs <- c(sx[nps[1]],sx[nps[2]], sx[nps[3]])  
}  
  
# funcao de outlier moderado  
outlier\_moderado <- function(x, o)  
{  
 result <- FALSE  
 oqs <- outliers\_quantiles(x, o)  
 om <- c(oqs[1] - 1.5 \* (oqs[3] - oqs[1]), oqs[1] - 3 \* (oqs[3] - oqs[1]), oqs[3] + 1.5 \* (oqs[3] - oqs[1]) , oqs[3] + 3 \* (oqs[3] - oqs[1]))  
 result <- ((o >= om[1] && o <= om[2]) || (o >= om[3] && o <= om[4]))  
 result  
}  
  
# funcao de oulier severo  
outlier\_severo <- function(x, o)  
{  
 result <- FALSE  
 oqs <- outliers\_quantiles(x, o)  
 os <- c(oqs[1] - 3 \* (oqs[3] - oqs[1]), oqs[3] + 3 \* (oqs[3] - oqs[1]))  
 result <- (o < os[1]) || (o > os[2])  
 result  
}  
  
# funcao de outliers tipo moderado  
outliers\_type\_moderado <- function(x, os)  
{  
 result <- c()  
 if(!is.na(os[1]))  
 {  
 for(i in 1:length(os))  
 {  
 m <- outlier\_moderado(x, os[i])  
 if (m)  
 {  
 result <- c(result, os[i])  
 }  
 }  
 }  
 result  
}  
  
# funcao de outliers tipo severo  
outliers\_type\_severo <- function(x, os)  
{  
 result <- c()  
 if(!is.na(os[1]))  
 {  
 for(i in 1:length(os))  
 {  
 s <- outlier\_severo(x, os[i])  
 if (s)  
 {  
 result <- c(result, os[i])  
 }  
 }  
 }  
 result  
}

# importar dataset  
# alterar caminho para o vosso  
# importar dataset  
hour.dataset <- "C:/Users/andre/Desktop/ME/ME\_Trabalhos\_2122/hour.csv"  
  
# Todos os alugueres todas as horas  
hour <- read.csv(file = hour.dataset)  
  
# Todos os alugueres entre as 7 às 18  
hour.day <- subset(hour, 6<hr & hr<19)  
  
# Todos os alugueres entre as 19 às 6  
hour.night <- subset(hour, hr>18 | hr<7)  
  
# Ver tabelas importadas no RStudio  
View(hour)

# tabela frequencia  
local({  
 .Table <- with(hour, table(hr))  
 cat("\ncounts:\n")  
 print(.Table)  
 cat("\npercentages:\n")  
 print(round(100\*.Table/sum(.Table), 2))  
})

##   
## counts:  
## hr  
## 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19   
## 726 724 715 697 697 717 725 727 727 727 727 727 728 729 729 729 730 730 728 728   
## 20 21 22 23   
## 728 728 728 728   
##   
## percentages:  
## hr  
## 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15   
## 4.18 4.17 4.11 4.01 4.01 4.13 4.17 4.18 4.18 4.18 4.18 4.18 4.19 4.19 4.19 4.19   
## 16 17 18 19 20 21 22 23   
## 4.20 4.20 4.19 4.19 4.19 4.19 4.19 4.19

#quartis summary(hour.day)

## **1.1. temp**

### Classificação da variável

Variável Qualitativa nominal, e é tratada como Qualitativa nominal normal, visto que, tem grande variedade de variaveis distintas.

### Medidas de Tendência Central

# Contagem de valores  
table(hour.day$temp)

##   
## 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22 0.24 0.26 0.28 0.3 0.32   
## 5 4 6 8 17 31 52 103 65 169 194 198 221 133 272 295   
## 0.34 0.36 0.38 0.4 0.42 0.44 0.46 0.48 0.5 0.52 0.54 0.56 0.58 0.6 0.62 0.64   
## 318 333 200 334 289 219 248 152 249 271 265 274 145 315 305 322   
## 0.66 0.68 0.7 0.72 0.74 0.76 0.78 0.8 0.82 0.84 0.86 0.88 0.9 0.92 0.94 0.96   
## 298 144 343 325 332 285 113 265 171 118 117 47 86 47 17 16   
## 0.98 1   
## 1 1

#### Moda

hour.day.temp.sort <- sort(hour.day$temp)  
  
hour.day.temp.mode <- moda(hour.day.temp.sort)  
hour.day.temp.mode

## [1] 0.7

#### Média

hour.day.temperatura.media <- mean(hour.day$temp)  
hour.day.temperatura.media

## [1] 0.5236736

#### Mediana

hour.day.temperatura.mediana <- median(hour.day$temp)  
hour.day.temperatura.mediana

## [1] 0.52

#### Quartil

hour.day.temperatura.quantile <- quantile(hour.day$temp)  
hour.day.temperatura.quantile

## 0% 25% 50% 75% 100%   
## 0.02 0.36 0.52 0.70 1.00

### Medidas de Dispersão Central

#### Variância

hour.day.temperatura.variancia <- var(hour.day$temp)  
hour.day.temperatura.variancia

## [1] 0.0401759

#### Desvio Padrão

hour.day.temperatura.desvio\_padrao <- sd(hour.day$temp)  
hour.day.temperatura.desvio\_padrao

## [1] 0.2004393

#### Coeficiente de Variação

hour.day.temperatura.coeficiente\_variacao <- coeficiente\_variacao(hour.day$temp)  
hour.day.temperatura.coeficiente\_variacao

## [1] 38.27561

### Valores Extremos

#### Máximo

hour.day.temperatura.maxima <- max(hour.day$temp)  
hour.day.temperatura.maxima

## [1] 1

#### Mínimo

hour.day.temperatura.minima <- min(hour.day$temp)  
hour.day.temperatura.minima

## [1] 0.02

### Medidas de Assimetria

hour.day.temperatura.assimetria <- assimetria(hour.day$temp)  
hour.day.temperatura.assimetria

## [1] -0.03380892

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

### Medidas de Curtose ou Achatamento

hour.day.temperatura.achatamento <- achatamento(hour.day$temp)  
hour.day.temperatura.achatamento

## [1] -0.998132

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

### Histograma

# formula de sturges  
cat("Número de classes\n")

## Número de classes

hour.day.temp.sturges.nclass <- sturges(hour.day$temp)  
hour.day.temp.sturges.nclass

## [1] 14

# amplitude da classe  
cat("\nAmplitude da classe\n")

##   
## Amplitude da classe

hour.day.temp.sturges.amplitude <- amplitude(hour.day$temp)  
hour.day.temp.sturges.amplitude

## [1] 0.07

# tabela de frequencias  
cat("\nClasses\n")

##   
## Classes

hour.day.temp.sturges.breaks <- seq(hour.day.temperatura.minima, hour.day.temperatura.maxima, hour.day.temp.sturges.amplitude)  
hour.day.temp.sturges.breaks

## [1] 0.02 0.09 0.16 0.23 0.30 0.37 0.44 0.51 0.58 0.65 0.72 0.79 0.86 0.93 1.00

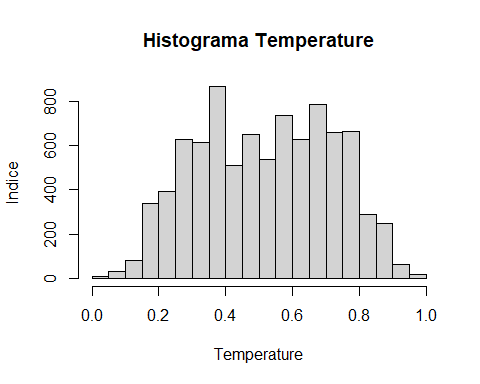
cat("\nTabela de frequências\n")

##   
## Tabela de frequências

table(cut(hour.day$temp, breaks=hour.day.temp.sturges.breaks))

##   
## (0.02,0.09] (0.09,0.16] (0.16,0.23] (0.23,0.3] (0.3,0.37] (0.37,0.44]   
## 18 100 531 824 946 823   
## (0.44,0.51] (0.51,0.58] (0.58,0.65] (0.65,0.72] (0.72,0.79] (0.79,0.86]   
## 868 955 942 1110 730 554   
## (0.86,0.93] (0.93,1]   
## 297 34

# histograma  
hist(hour.day$temp, main="Histograma Temperature", xlab="Temperature", ylab="Indice", nclass=hour.day.temp.sturges.nclass)



A partir do histograma, pode-se observar que o valor do índice aumenta no intervalo 0.02, a 0.72 e começa a diminuir de 0.72 a 1 tendo um comportamento crescente até atingir o maior índice no intervalo de 0.65 a 0.72.

### Diagrama de Extremos e Quartis

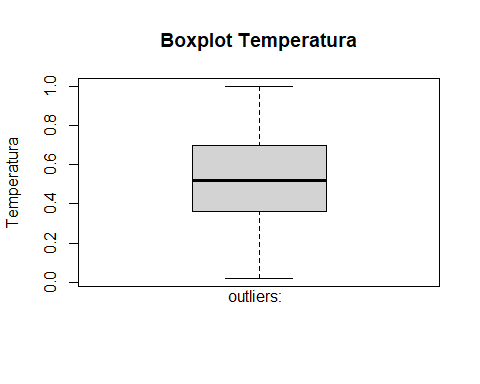
# Outliers  
cat("\nOutliers\n")

##   
## Outliers

hour.day.temperatura.outliers <- boxplot.stats(hour.day$temp)$out  
hour.day.temperatura.outliers

## numeric(0)

# Diagrama  
boxplot(hour.day$temp, ylab="Temperatura", main="Boxplot Temperatura")  
  
mtext(text = paste("outliers: ", paste(hour.day.temperatura.outliers, collapse = ", ")), side = 1)

 Verificou-se que não existem outliers presentes no gráfico.

## **2.1. weekday**

### Classificação da variável

Variável qualitativa ordinal, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variaveis distintas.

### Medidas de Tendência Central

#### Moda

hour.day.weekday.moda <- moda(hour.day$weekday)  
hour.day.weekday.moda

## [1] 0

#### Média

hour.day.weekday.media <- mean(hour.day$weekday)  
hour.day.weekday.media

## [1] 2.999886

#### Mediana

hour.day.weekday.mediana <- median(hour.day$weekday)  
hour.day.weekday.mediana

## [1] 3

#### Quartil

hour.day.weekday.quantile <- quantile(hour.day$weekday)  
hour.day.weekday.quantile

## 0% 25% 50% 75% 100%   
## 0 1 3 5 6

### Medidas de Dispersão Central

#### Variância

hour.day.weekday.variancia <- var(hour.day$weekday)  
hour.day.weekday.variancia

## [1] 4.020945

#### Desvio Padrão

hour.day.weekday.desvio\_padrao <- sd(hour.day$weekday)  
hour.day.weekday.desvio\_padrao

## [1] 2.00523

#### Coeficiente de Variação

hour.day.weekday.coeficiente\_variacao <- coeficiente\_variacao(hour.day$weekday)  
hour.day.weekday.coeficiente\_variacao

## [1] 66.84353

### Valores Extremos

#### Máximo

hour.day.weekday.maxima <- max(hour.day$weekday)  
hour.day.weekday.maxima

## [1] 6

#### Mínimo

hour.day.weekday.minima <- min(hour.day$weekday)  
hour.day.weekday.minima

## [1] 0

### Medidas de Assimetria

hour.day.weekday.assimetria <- assimetria(hour.day$weekday)  
hour.day.weekday.assimetria

## [1] -0.0001836458

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

### Medidas de Curtose ou Achatamento

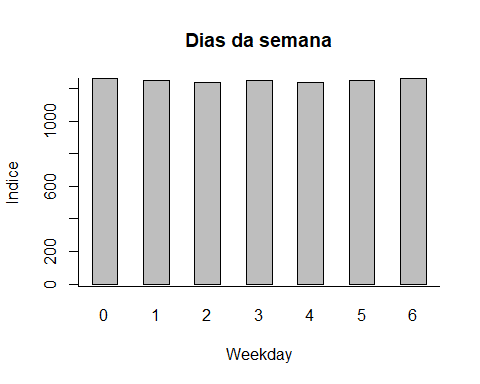
hour.day.weekday.achatamento <- achatamento(hour.day$weekday)  
hour.day.weekday.achatamento

## [1] -1.255535

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

### Grafico de barras

count <- table(hour.day$weekday)  
  
barplot(count, main="Dias da semana", xlab="Weekday", ylab="Indice", space =1.0)  
box(bty="L")

 A partir do Gráfico de barras, pode-se observar que o valor do índice mantém se no intervalo 0 a 6, tendo um comportamento crescente até atingir o maior índice que é 6.

### Diagrama de Extremos e Quartis

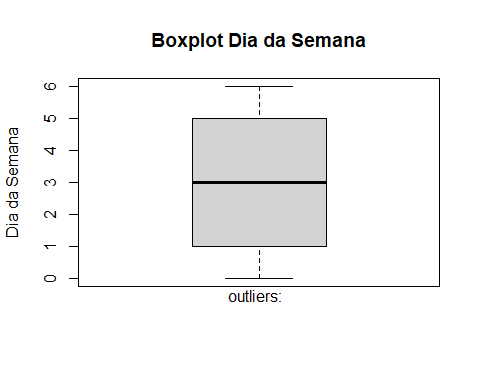
# Outliers  
cat("\nOutliers\n")

##   
## Outliers

hour.day.weekday.outliers <- boxplot.stats(hour.day$weekday)$out  
hour.day.weekday.outliers

## integer(0)

# Diagrama  
boxplot(hour.day$weekday, ylab="Dia da Semana", main="Boxplot Dia da Semana")  
  
mtext(text = paste("outliers: ", paste(hour.day.weekday.outliers, collapse = ", ")), side = 1)



Verificou-se que não existem outliers presentes no gráfico.

## **3.. casual**

### Classificação da variável

Variável quantitativa discreta, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variaveis distintas.

### Medidas de Tendência Central

#### Moda

hour.day.casual.moda <- moda(hour.day$casual)  
hour.day.casual.moda

## [1] 6

#### Média

hour.day.casual.media <- mean(hour.day$casual)  
hour.day.casual.media

## [1] 55.84882

#### Mediana

hour.day.casual.mediana <- median(hour.day$casual)  
hour.day.casual.mediana

## [1] 36

#### Quartil

hour.day.casual.quantile <- quantile(hour.day$casual)  
hour.day.casual.quantile

## 0% 25% 50% 75% 100%   
## 0 15 36 74 367

### Medidas de Dispersão Central

#### Variância

hour.day.casual.variancia <- var(hour.day$casual)  
hour.day.casual.variancia

## [1] 3436.565

#### Desvio Padrão

hour.day.casual.desvio\_padrao <- sd(hour.day$casual)  
hour.day.casual.desvio\_padrao

## [1] 58.62222

#### Coeficiente de Variação

hour.day.casual.coeficiente\_variacao <- coeficiente\_variacao(hour.day$casual)  
hour.day.casual.coeficiente\_variacao

## [1] 104.9659

### Valores Extremos

#### Máximo

hour.day.casual.maxima <- max(hour.day$casual)  
hour.day.casual.maxima

## [1] 367

#### Mínimo

hour.day.casual.minima <- min(hour.day$casual)  
hour.day.casual.minima

## [1] 0

### Medidas de Assimetria

hour.day.casual.assimetria <- assimetria(hour.day$casual)  
hour.day.casual.assimetria

## [1] 1.899737

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

### Medidas de Curtose ou Achatamento

hour.day.casual.achatamento <- achatamento(hour.day$casual)  
hour.day.casual.achatamento

## [1] 3.905951

Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

### Histograma

# formula de sturges  
cat("Número de classes\n")

## Número de classes

hour.day.casual.sturges.nclass <- sturges(hour.day$casual)  
hour.day.casual.sturges.nclass

## [1] 14

# amplitude da classe  
cat("\nAmplitude da classe\n")

##   
## Amplitude da classe

hour.day.casual.sturges.amplitude <- amplitude(hour.day$casual)  
hour.day.casual.sturges.amplitude

## [1] 26.21429

# tabela de frequencias  
cat("\nClasses\n")

##   
## Classes

hour.day.casual.sturges.breaks <- seq(hour.day.casual.minima, hour.day.casual.maxima, hour.day.casual.sturges.amplitude)  
hour.day.casual.sturges.breaks

## [1] 0.00000 26.21429 52.42857 78.64286 104.85714 131.07143 157.28571  
## [8] 183.50000 209.71429 235.92857 262.14286 288.35714 314.57143 340.78571  
## [15] 367.00000

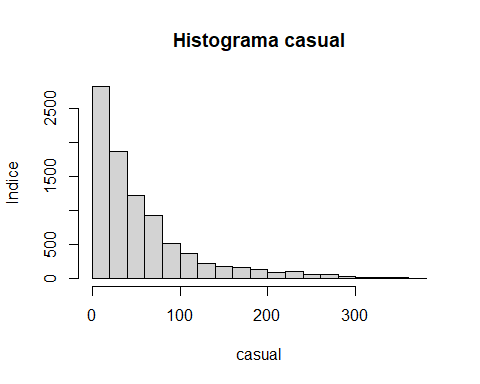
cat("\nTabela de frequências\n")

##   
## Tabela de frequências

table(cut(hour.day$casual, breaks=hour.day.casual.sturges.breaks))

##   
## (0,26.2] (26.2,52.4] (52.4,78.6] (78.6,105] (105,131] (131,157]   
## 3312 2043 1303 672 417 239   
## (157,184] (184,210] (210,236] (236,262] (262,288] (288,315]   
## 199 149 105 100 55 26   
## (315,341] (341,367]   
## 12 10

# histograma  
hist(hour.day$casual, main="Histograma casual", xlab="casual", ylab="Indice", nclass=hour.day.casual.sturges.nclass)



A partir do histograma, pode-se observar que o valor do índice diminui no intervalo 0, a 367, tendo um comportamento decrescente até atingir o maior índice no intervalo de 0 a 24.5.

### Diagrama de Extremos e Quartis

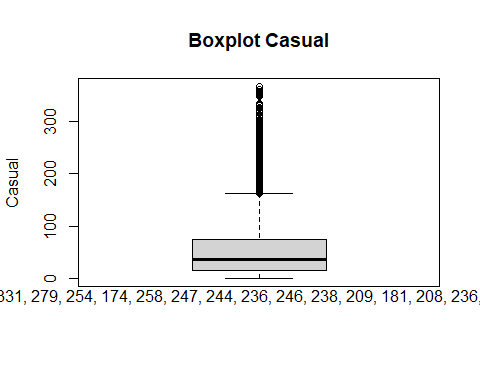
# Outliers  
cat("\nOutliers\n")

##   
## Outliers

hour.day.casual.outliers <- boxplot.stats(hour.day$casual)$out  
hour.day.casual.outliers

## [1] 174 168 170 175 172 166 219 240 174 167 181 170 179 182 171 180 168 168  
## [19] 205 197 167 178 185 184 217 191 196 166 195 183 206 173 222 191 187 232  
## [37] 204 191 191 164 180 171 168 172 187 201 180 173 216 237 183 221 194 214  
## [55] 179 188 193 180 168 169 176 193 210 175 175 197 197 174 177 178 178 177  
## [73] 168 206 179 193 215 198 248 225 194 195 181 166 177 237 242 235 224 236  
## [91] 240 222 167 182 171 171 187 221 201 205 183 234 185 177 200 168 176 176  
## [109] 175 175 180 169 178 196 204 187 186 170 174 214 245 205 218 196 204 187  
## [127] 217 170 185 180 170 164 168 167 180 176 191 256 251 262 221 174 189 212  
## [145] 225 272 198 223 208 186 164 195 171 242 166 179 177 167 169 204 182 177  
## [163] 169 164 192 171 188 191 165 165 177 165 183 229 186 167 176 212 201 208  
## [181] 199 226 286 286 352 357 367 291 221 222 198 218 240 229 233 165 199 172  
## [199] 208 191 219 264 237 202 184 263 265 240 275 243 238 180 187 190 283 295  
## [217] 236 232 173 240 218 184 172 187 251 320 355 326 321 354 299 227 195 198  
## [235] 229 254 260 232 185 207 274 308 288 311 253 251 197 186 227 275 298 282  
## [253] 266 286 262 184 183 260 262 291 278 267 267 259 223 226 281 279 240 188  
## [271] 164 221 220 217 187 241 230 204 207 197 226 229 181 195 206 293 257 269  
## [289] 254 233 164 189 255 228 190 225 191 177 235 276 332 361 356 331 279 254  
## [307] 174 258 247 244 236 246 238 209 181 208 236 265 307 274 261 235 199 209  
## [325] 268 301 270 317 290 258 275 266 189 254 233 272 238 181 223 179 250 279  
## [343] 259 297 275 248 171 199 276 265 267 236 226 192 168 218 196 236 239 218  
## [361] 219 238 175 184 173 204 186 173 184 164 213 254 293 264 295 238 244 256  
## [379] 196 260 267 255 203 240 232 203 177 167 203 226 238 217 196 214 194 215  
## [397] 227 237 253 197 181 203 186 164 164 177 190 163 181 219 269 243 226 199  
## [415] 170 171 182 191 178 238 256 254 245 188 173 173 186 206 208 199 183 181  
## [433] 231 195 191 164 222 195 183 205 197 179 169 208 164 203 214 228 248 246  
## [451] 227 220 167 163 178 213 235 213 186 164 166 180 266 289 242 250 287 256  
## [469] 225 217 186 174 225 166 219 187 237 240 257 209 205 184 226 248 263 249  
## [487] 179 222 219 227 187 191 209 163 168 165 181 173 220 233 260 175 172 194  
## [505] 247 238 232 284 228 213 168 228 287 327 325 312 350 295 232 183 188 240  
## [523] 225 215 194 177 189 228 273 250 307 253 195 171 186 250 257 266 265 254  
## [541] 227 217 191 233 302 271 275 256 174 222 206 180 208 230 202 195 198 258  
## [559] 362 310 269 279 317 268 183 192 202 235 243 251 294 193 174 200 218 249  
## [577] 213 203 193 204 267 273 335 308 325 347 165 167 191 236 243 235 240 182  
## [595] 165 197 264 310 325 310 257 248 258 268 280 216 227 183 180 195 262 292  
## [613] 304 260 179 175 175 186 224 240 233 164 167

# Diagrama  
boxplot(hour.day$casual, ylab="Casual", main="Boxplot Casual")  
mtext(text = paste("outliers: ", paste(hour.day.casual.outliers, collapse = ", ")), side = 1)



Verificou-se que existem outliers presentes no gráfico.

## **4.. atemp**

### Classificação da variável

Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variaveis distintas.

### Medidas de Tendência Central

#### Moda

hour.day.atemp.moda <- moda(hour.day$atemp)  
hour.day.atemp.moda

## [1] 0.6212

#### Média

hour.day.atemp.media <- mean(hour.day$atemp)  
hour.day.atemp.media

## [1] 0.4981943

#### Mediana

hour.day.atemp.mediana <- median(hour.day$atemp)  
hour.day.atemp.mediana

## [1] 0.5

#### Quartil

hour.day.atemp.quantile <- quantile(hour.day$atemp)  
hour.day.atemp.quantile

## 0% 25% 50% 75% 100%   
## 0.0000 0.3485 0.5000 0.6515 1.0000

### Medidas de Dispersão Central

#### Variância

hour.day.atemp.variancia <- var(hour.day$atemp)  
hour.day.atemp.variancia

## [1] 0.03187527

#### Desvio Padrão

hour.day.atemp.desvio\_padrao <- sd(hour.day$atemp)  
hour.day.atemp.desvio\_padrao

## [1] 0.1785365

#### Coeficiente de Variação

hour.day.atemp.coeficiente\_variacao <- coeficiente\_variacao(hour.day$atemp)  
hour.day.atemp.coeficiente\_variacao

## [1] 35.83671

### Valores Extremos

#### Máximo

hour.day.atemp.maxima <- max(hour.day$atemp)  
hour.day.atemp.maxima

## [1] 1

#### Mínimo

hour.day.atemp.minima <- min(hour.day$atemp)  
hour.day.atemp.minima

## [1] 0

### Medidas de Assimetria

hour.day.atemp.assimetria <- assimetria(hour.day$atemp)  
hour.day.atemp.assimetria

## [1] -0.1370249

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

### Medidas de Curtose ou Achatamento

hour.day.atemp.achatamento <- achatamento(hour.day$atemp)  
hour.day.atemp.achatamento

## [1] -0.8923441

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

### Histograma

# formula de sturges  
cat("Número de classes\n")

## Número de classes

hour.day.atemp.sturges.nclass <- sturges(hour.day$atemp)  
hour.day.atemp.sturges.nclass

## [1] 14

# amplitude da classe  
cat("\nAmplitude da classe\n")

##   
## Amplitude da classe

hour.day.atemp.sturges.amplitude <- amplitude(hour.day$atemp)  
hour.day.atemp.sturges.amplitude

## [1] 0.07142857

# tabela de frequencias  
cat("\nClasses\n")

##   
## Classes

hour.day.atemp.sturges.breaks <- seq(hour.day.atemp.minima, hour.day.atemp.maxima, hour.day.atemp.sturges.amplitude)  
hour.day.atemp.sturges.breaks

## [1] 0.00000000 0.07142857 0.14285714 0.21428571 0.28571429 0.35714286  
## [7] 0.42857143 0.50000000 0.57142857 0.64285714 0.71428571 0.78571429  
## [13] 0.85714286 0.92857143 1.00000000

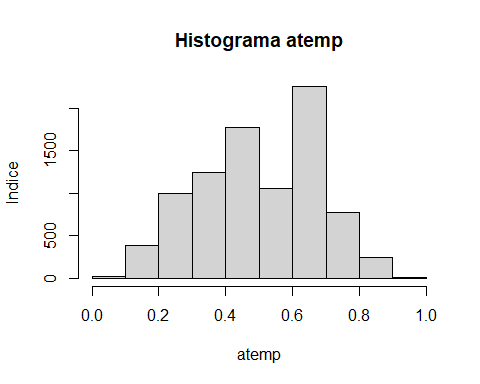
cat("\nTabela de frequências\n")

##   
## Tabela de frequências

table(cut(hour.day$atemp, breaks=hour.day.atemp.sturges.breaks))

##   
## (0,0.0714] (0.0714,0.143] (0.143,0.214] (0.214,0.286] (0.286,0.357]   
## 6 118 424 672 1120   
## (0.357,0.429] (0.429,0.5] (0.5,0.571] (0.571,0.643] (0.643,0.714]   
## 918 1145 799 1344 1333   
## (0.714,0.786] (0.786,0.857] (0.857,0.929] (0.929,1]   
## 503 294 57 4

# histograma  
hist(hour.day$atemp, main="Histograma atemp", xlab="atemp", ylab="Indice", nclass=hour.day.atemp.sturges.nclass)



A partir do histograma, pode-se observar que o valor do índice aumenta no intervalo 0 a 0.5337 e diminui no intervalo 0.5337 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.467 a 0.533.

### Diagrama de Extremos e Quartis

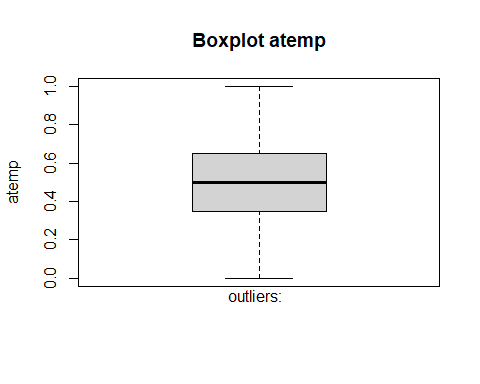
# Outliers  
cat("\nOutliers\n")

##   
## Outliers

hour.day.atemp.outliers <- boxplot.stats(hour.day$atemp)$out  
hour.day.atemp.outliers

## numeric(0)

# Diagrama  
boxplot(hour.day$atemp, ylab="atemp", main="Boxplot atemp")  
mtext(text = paste("outliers: ", paste(hour.day.atemp.outliers, collapse = ", ")), side = 1)



Verificou-se que não existem outliers presentes no gráfico.

## **5.. hum**

### Classificação da variável

Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, não tem grande variedade de variaveis distintas.

### Medidas de Tendência Central

#### Moda

hour.day.hum.moda <- moda(hour.day$hum)  
hour.day.hum.moda

## [1] 0.46

#### Média

hour.day.hum.media <- mean(hour.day$hum)  
hour.day.hum.media

## [1] 0.572013

#### Mediana

hour.day.hum.mediana <- median(hour.day$hum)  
hour.day.hum.mediana

## [1] 0.55

#### Quartil

hour.day.hum.quantile <- quantile(hour.day$hum)  
hour.day.hum.quantile

## 0% 25% 50% 75% 100%   
## 0.00 0.42 0.55 0.72 1.00

### Medidas de Dispersão Central

#### Variância

hour.day.hum.variancia <- var(hour.day$hum)  
hour.day.hum.variancia

## [1] 0.03764531

#### Desvio Padrão

hour.day.hum.desvio\_padrao <- sd(hour.day$hum)  
hour.day.hum.desvio\_padrao

## [1] 0.194024

#### Coeficiente de Variação

hour.day.hum.coeficiente\_variacao <- coeficiente\_variacao(hour.day$hum)  
hour.day.hum.coeficiente\_variacao

## [1] 33.91951

### Valores Extremos

#### Máximo

hour.day.hum.maxima <- max(hour.day$hum)  
hour.day.hum.maxima

## [1] 1

#### Mínimo

hour.day.hum.minima <- min(hour.day$hum)  
hour.day.hum.minima

## [1] 0

### Medidas de Assimetria

hour.day.hum.assimetria <- assimetria(hour.day$hum)  
hour.day.hum.assimetria

## [1] 0.2254248

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

### Medidas de Curtose ou Achatamento

hour.day.hum.achatamento <- achatamento(hour.day$hum)  
hour.day.hum.achatamento

## [1] -0.7835361

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

### Histograma

# formula de sturges  
cat("Número de classes\n")

## Número de classes

hour.day.hum.sturges.nclass <- sturges(hour.day$hum)  
hour.day.hum.sturges.nclass

## [1] 14

# amplitude da classe  
cat("\nAmplitude da classe\n")

##   
## Amplitude da classe

hour.day.hum.sturges.amplitude <- amplitude(hour.day$hum)  
hour.day.hum.sturges.amplitude

## [1] 0.07142857

# tabela de frequencias  
cat("\nClasses\n")

##   
## Classes

hour.day.hum.sturges.breaks <- seq(hour.day.hum.minima, hour.day.hum.maxima, hour.day.hum.sturges.amplitude)  
hour.day.hum.sturges.breaks

## [1] 0.00000000 0.07142857 0.14285714 0.21428571 0.28571429 0.35714286  
## [7] 0.42857143 0.50000000 0.57142857 0.64285714 0.71428571 0.78571429  
## [13] 0.85714286 0.92857143 1.00000000

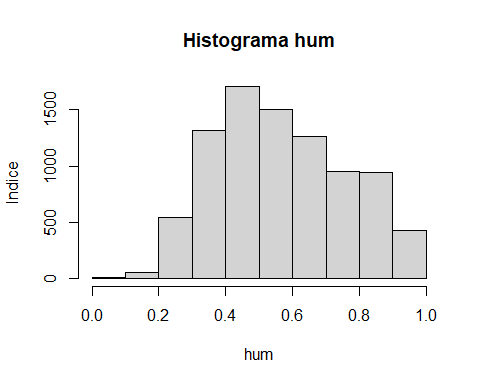
cat("\nTabela de frequências\n")

##   
## Tabela de frequências

table(cut(hour.day$hum, breaks=hour.day.hum.sturges.breaks))

##   
## (0,0.0714] (0.0714,0.143] (0.143,0.214] (0.214,0.286] (0.286,0.357]   
## 0 0 75 348 707   
## (0.357,0.429] (0.429,0.5] (0.5,0.571] (0.571,0.643] (0.643,0.714]   
## 1159 1339 1102 887 861   
## (0.714,0.786] (0.786,0.857] (0.857,0.929] (0.929,1]   
## 748 568 509 423

# histograma  
hist(hour.day$hum, main="Histograma hum", xlab="hum", ylab="Indice", nclass=hour.day.hum.sturges.nclass)



A partir do histograma, pode-se observar que o valor do índice aumenta no intervalo 0 a 0.6 e diminui no intervalo 0.6 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.533 a 0.6.

### Diagrama de Extremos e Quartis

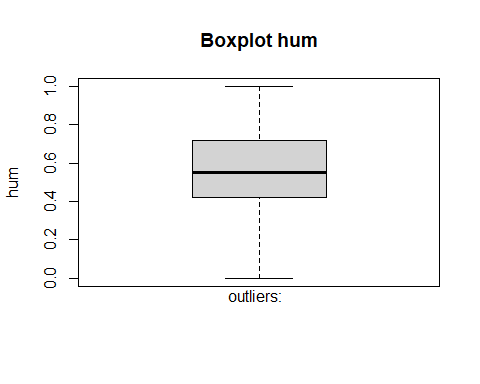
# Outliers  
cat("\nOutliers\n")

##   
## Outliers

hour.day.hum.outliers <- boxplot.stats(hour.day$hum)$out  
hour.day.hum.outliers

## numeric(0)

# Diagrama  
boxplot(hour.day$hum, ylab="hum", main="Boxplot hum")  
mtext(text = paste("outliers: ", paste(hour.day.hum.outliers, collapse = ", ")), side = 1)



Verificou-se que existem outliers presentes no gráfico no ponto 0.

## **6.. cnt**

### Classificação da variável

Variável qualitativa ordinal, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variaveis distintas.

### Medidas de Tendência Central

#### Moda

hour.day.cnt.moda <- moda(hour.day$cnt)  
hour.day.cnt.moda

## [1] 224

#### Média

hour.day.cnt.media <- mean(hour.day$cnt)  
hour.day.cnt.media

## [1] 280.9225

#### Mediana

hour.day.cnt.mediana <- median(hour.day$cnt)  
hour.day.cnt.mediana

## [1] 237

#### Quartil

hour.day.cnt.quantile <- quantile(hour.day$cnt)  
hour.day.cnt.quantile

## 0% 25% 50% 75% 100%   
## 1 142 237 382 977

### Medidas de Dispersão Central

#### Variância

hour.day.cnt.variancia <- var(hour.day$cnt)  
hour.day.cnt.variancia

## [1] 34893.5

#### Desvio Padrão

hour.day.cnt.desvio\_padrao <- sd(hour.day$cnt)  
hour.day.cnt.desvio\_padrao

## [1] 186.798

#### Coeficiente de Variação

hour.day.cnt.coeficiente\_variacao <- coeficiente\_variacao(hour.day$cnt)  
hour.day.cnt.coeficiente\_variacao

## [1] 66.4945

### Valores Extremos

#### Máximo

hour.day.cnt.maxima <- max(hour.day$cnt)  
hour.day.cnt.maxima

## [1] 977

#### Mínimo

hour.day.cnt.minima <- min(hour.day$cnt)  
hour.day.cnt.minima

## [1] 1

### Medidas de Assimetria

hour.day.cnt.assimetria <- assimetria(hour.day$cnt)  
hour.day.cnt.assimetria

## [1] 0.9866874

De acordo com o resultado, é assimétrica positiva (ou enviesada para a direita).

### Medidas de Curtose ou Achatamento

hour.day.cnt.achatamento <- achatamento(hour.day$cnt)  
hour.day.cnt.achatamento

## [1] 0.6186668

Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

### Histograma

# formula de sturges  
cat("Número de classes\n")

## Número de classes

hour.day.cnt.sturges.nclass <- sturges(hour.day$cnt)  
hour.day.cnt.sturges.nclass

## [1] 14

# amplitude da classe  
cat("\nAmplitude da classe\n")

##   
## Amplitude da classe

hour.day.cnt.sturges.amplitude <- amplitude(hour.day$cnt)  
hour.day.cnt.sturges.amplitude

## [1] 69.71429

# tabela de frequencias  
cat("\nClasses\n")

##   
## Classes

hour.day.cnt.sturges.breaks <- seq(hour.day.cnt.minima, hour.day.cnt.maxima, hour.day.cnt.sturges.amplitude)  
hour.day.cnt.sturges.breaks

## [1] 1.00000 70.71429 140.42857 210.14286 279.85714 349.57143 419.28571  
## [8] 489.00000 558.71429 628.42857 698.14286 767.85714 837.57143 907.28571  
## [15] 977.00000

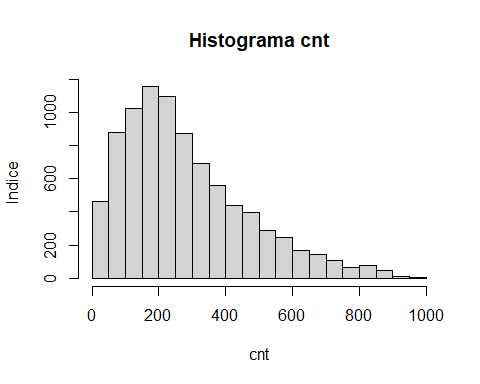
cat("\nTabela de frequências\n")

##   
## Tabela de frequências

table(cut(hour.day$cnt, breaks=hour.day.cnt.sturges.breaks))

##   
## (1,70.7] (70.7,140] (140,210] (210,280] (280,350] (350,419] (419,489]   
## 812 1329 1616 1370 1035 760 541   
## (489,559] (559,628] (628,698] (698,768] (768,838] (838,907] (907,977]   
## 434 312 204 130 108 69 15

# histograma  
hist(hour.day$cnt, main="Histograma cnt", xlab="cnt", ylab="Indice", nclass=hour.day.cnt.sturges.nclass)



A partir do histograma, pode-se observar que o valor do índice diminui no intervalo 0, a 977, tendo um comportamento decrescente até atingir o maior índice no intervalo de 1 a 66.1.

### Diagrama de Extremos e Quartis

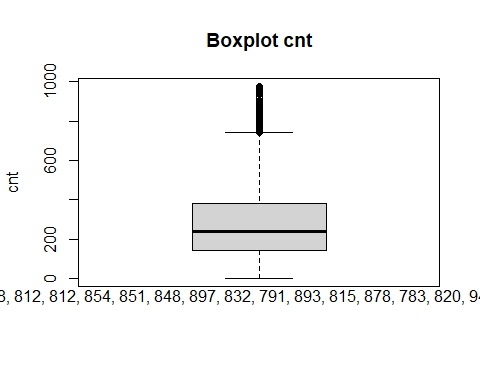
# Outliers  
cat("\nOutliers\n")

##   
## Outliers

hour.day.cnt.outliers <- boxplot.stats(hour.day$cnt)$out  
hour.day.cnt.outliers

## [1] 782 749 746 801 750 801 779 810 801 957 830 757 800 744 759 822 744 781  
## [19] 775 748 776 819 743 813 769 770 779 873 846 852 868 745 812 785 785 798  
## [37] 752 781 839 796 827 785 834 822 850 790 782 869 813 793 800 831 857 744  
## [55] 867 823 823 811 795 833 791 900 824 843 804 747 849 872 872 819 830 814  
## [73] 795 825 835 755 794 770 772 771 777 847 743 869 877 788 913 891 751 827  
## [91] 812 760 820 857 837 891 865 767 858 843 868 814 858 862 810 811 818 812  
## [109] 812 854 851 848 897 832 791 893 815 878 783 820 941 744 812 845 834 864  
## [127] 818 808 870 812 754 844 853 856 839 863 839 808 835 772 792 757 871 968  
## [145] 750 970 877 770 925 977 758 884 852 766 894 808 783 842 774 797 886 892  
## [163] 790 976 900 757 846 805 750 776 750 898 853 805 967 822 838 953 884 794  
## [181] 905 899 808 750 784 856 809 917 810 901 887 785 900 761 743 806 784 839  
## [199] 948 844 798 827 745 837 766 835 943 838 817 888 884 834 890 788 922 786  
## [217] 761 938 826 801 963 858 779 886 809 817 760 750 743 759

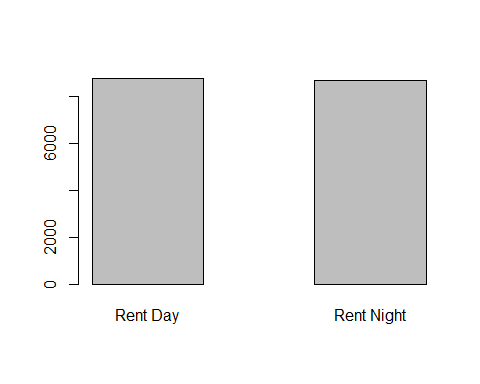
# Diagrama  
boxplot(hour.day$cnt, ylab="cnt", main="Boxplot cnt")  
mtext(text = paste("outliers: ", paste(hour.day.cnt.outliers, collapse = ", ")), side = 1)



Verificou-se que existem bastantes outliers presentes no gráfico.

### Gráfico de Barras

values <- c(nrow(hour.day), nrow(hour.night))   
labels <- c("Rent Day", "Rent Night")  
barplot(values, names.arg=labels, space =1.0)



### Gráfico Circular

x <- c(nrow(hour.day), nrow(hour.night))  
labels <- c("Rent Day", "Rent Night")  
piepercent<- round(100\*x/sum(x), 1)  
pie(x, labels = piepercent, main = "Rent hour.day pie chart",col = rainbow(length(x)))  
legend("topright", c("Rent Day", "Rent Night"), cex = 0.8,  
 fill = rainbow(length(x)))

