

Aula 6 - Construção de gráficos

Profa. Yana Borges

Março de 2022

Sempre que analisamos dados, a primeira coisa que devemos fazer é olhar para eles. Para cada variável, quais são os valores mais comuns? Quanta variabilidade está presente? Existem observações incomuns?

Uma forma de responder a essas perguntas é através de análise gráfica

Gráfico de setores

Usaremos os dados `Arthritis` do pacote `vcd`. Este pacote não está incluído na instalação padrão do R, portanto, instale-o antes do primeiro uso (`install.packages("vcd")`).

Observe que o pacote `vcd` não é necessário para criar gráficos de barras. Você está carregando para obter acesso ao conjunto de dados `Arthritis`.

```
require("vcd")
```

Antes de se criar o gráfico é preciso criar um objeto na forma de tabela para então representá-lo graficamente.

Gráfico de setores

O comando para se fazer gráficos de setores é dado por `pie(sexo)`, por exemplo

```
sexo <- table(Arthritis$Sex)
par(mar = c(0.5,.5,.5,.5))
pie(sexo)
```

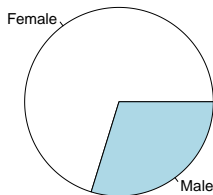


Figure 1: Gráficos de setores

Gráfico de Barras

Um gráfico de barras exibe a distribuição (frequência) de uma variável categórica por meio de barras verticais ou horizontais.

Assim como o Gráfico de setores, é preciso criar uma tabela

```
count <- table(Arthritis$Improved)
```

Gráfico de Barras

Você pode representar graficamente as contagens de variáveis usando um gráfico de barra vertical ou horizontal usando `barplot`.

```
par(mfrow = c(1,2), mar = c(2,2,1.5,1))  
barplot(count)  
barplot(count)
```

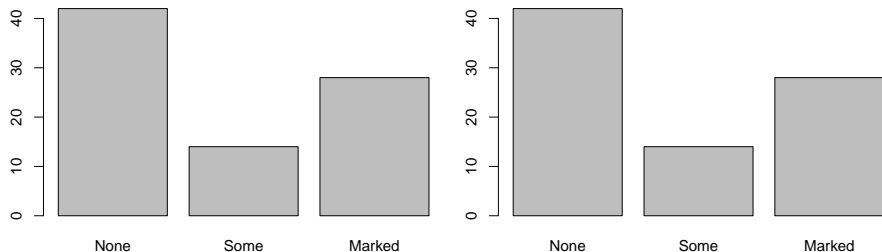


Figure 2: Gráficos de barras verticais e horizontais simples

Gráfico de barras

Se a variável a ser plotada for um fator, como "Arthritis\$Improved", você pode criar um gráfico de barra vertical rapidamente com a função `plot()`, sem precisar da função `table`.

```
par(mfrow = c(1,2), mar = c(2,2,1.5,1))  
plot(Arthritis$Improved);  
plot(Arthritis$Improved, horiz=TRUE)
```

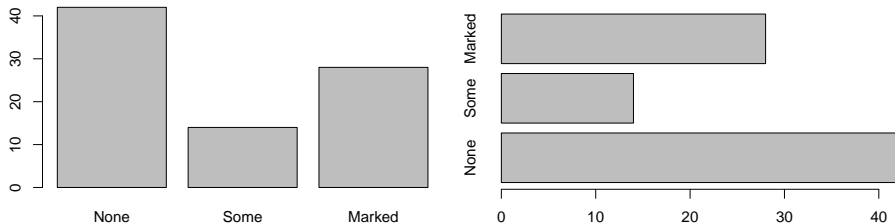


Figure 3: Gráficos de barras verticais e horizontais simples

Histograma

A forma mais utilizada para se representar variáveis quantitativas contínuas é através do histograma.

```
hist(Arthritis$Age)
```

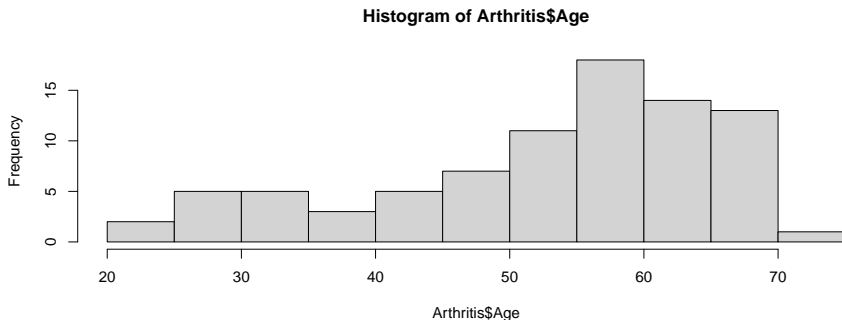


Figure 4: Histograma de frequência

Gráfico de linhas

O gráfico de linhas é o mais apropriado para variáveis observadas através do tempo e pode ser construído usando a função `plot` com o argumento `type`.

Para este, usaremos o conjunto de dados `mtcars` e a variável `wt` (*Peso (1000 lbs)*).

```
par(mar = c(3,2,1,1))  
plot(mtcars$wt, type = "l")
```

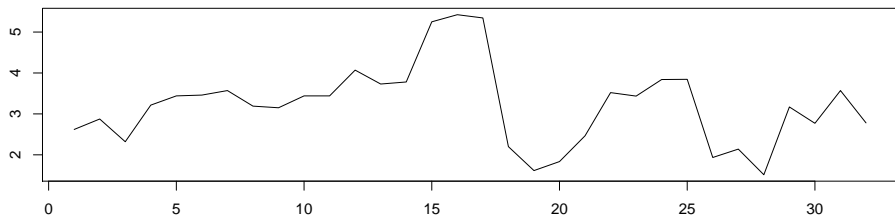


Figure 5: Gráfico de linhas

Gráfico de dispersão

Quando tratamos de duas variáveis quantitativas é interessante observar o comportamento conjunto entre elas. A maneira mais usual para se verificar esse comportamento é o gráfico de dispersão. Para construir esse gráfico basta utilizar a função `plot()` e inserir as duas variáveis analisadas.

Gráfico de dispersão

Ainda usando o conjunto de dados `mtcars`, iremos verificar as variáveis `wt` e `mpg` (*Milhas/(EUA) galão*).

```
plot(mtcars$mpg, mtcars$wt)
```

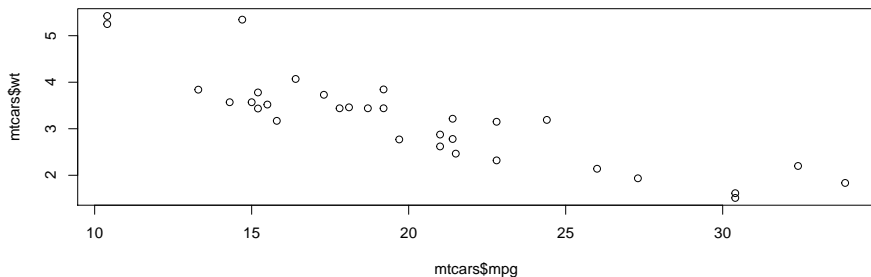


Figure 6: Gráfico de dispersão

Box Plot

Quando tratamos de variáveis quantitativas é interessante observar a distribuição dos dados de um modo geral. Uma forma de se verificar isso é utilizar Box Plot em que conseguimos visualizar os quartis, mediana, limites superior e inferior e outliers.

```
par(mar=c(.1,2,.1,1))  
boxplot(mtcars$mpg)
```

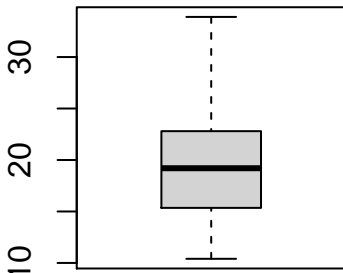


Figure 7: Box Plot

Box Plot

Voltando ao conjunto de dados Arthritis, podemos fazer o boxplot da variável idade, segundo o sexo

```
par(mar=c(2,2,.1,1))  
boxplot(Arthritis$Age ~ Arthritis$Sex)
```

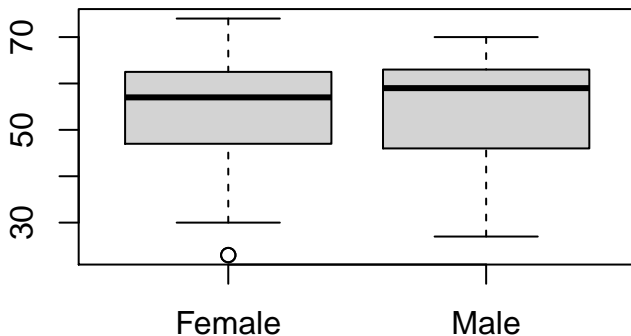


Figure 8: Box Plot

Ajustes gráficos

Em praticamente qualquer gráfico, teremos as seguintes opções:

- `main` Informa o título do gráfico
- `sub` Cria um subtítulo do gráfico
- `xlab` e `ylab`= Trocam as legendas dos eixos
- `xlim` e `ylim`= Trocam os limites dos eixos
- `col` Informa a cor do gráfico

Ajustes gráficos

```
par( mar = c(2,2,1.5,1))  
barplot(count,main = "Título do gráfico",  
        sub = "Subtítulo do gráfico",  
        xlab = "Altera legenda do eixo x",  
        ylab = "Altera legenda do eixo y",  
        ylim = c(1,50), col = 1:3  
)
```

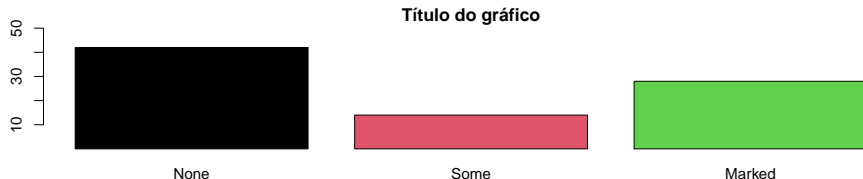


Figure 9: Gráficos de barras verticais e horizontais simples

Ajustes gráficos

```
sexo <- table(Arthritis$Sex)
par(mar = c(0.5,.5,.5,.5))
pie(sexo,
     main = "Título do gráfico",
     sub = "Subtítulo", labels = c("c", "d"))
```

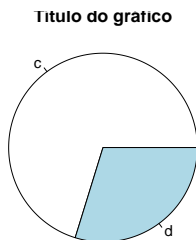
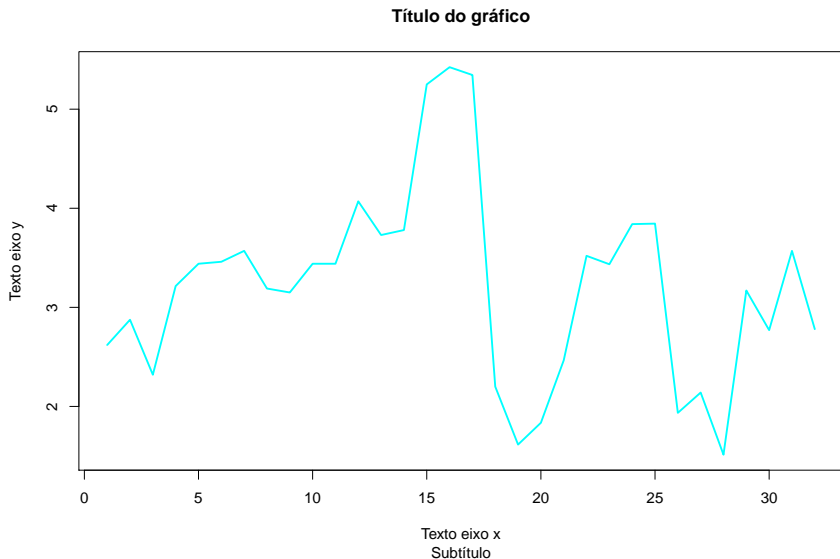


Figure 10: Gráficos de setores

Ajustes gráficos

```
plot(mtcars$wt, type = "l", col = "cyan", lwd=2,  
     main = "Título do gráfico", sub = "Subtítulo",  
     xlab = "Texto eixo x", ylab = "Texto eixo y")
```

Ajustes gráficos



Ajustes gráficos

`cex.axis` - Altera o tamanho da fonte para os valores dos eixos. Argumento numérico com padrão = 1. Pode ser usado junto com `font.axis` para alterar o tipo de fonte e `col.axis` para alterar a cor.

```
plot(mtcars$mpg, mtcars$wt,  
     xlab = "Variável preditora",  
     ylab = "Variável resposta", main = "Título",  
     xlim = c(10, 25), ylim = c(0, 10),  
     cex.axis = 3,  
     col.axis = "darkblue",  
     font.axis = 3)
```

Ajustes gráficos

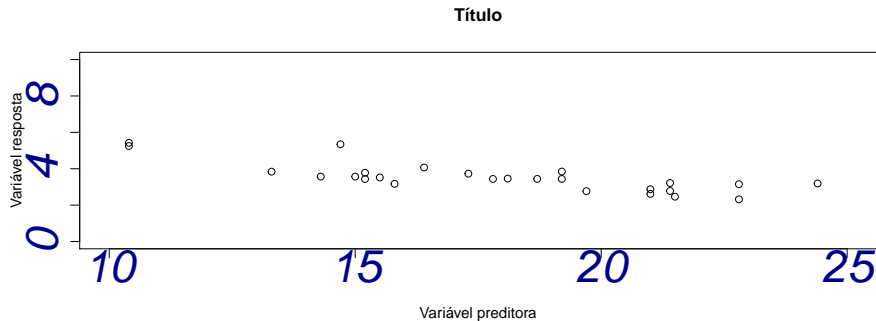


Figure 12: Gráfico de dispersão

Ajustes gráficos

`type`- Alterar o tipo de gráfico que será construído. Vários valores são aceitos, os principais são "p" (gráfico de pontos), "l" (gráfico de linhas), "b" (gráficos com pontos e linhas) e "n" (para não mostrar nada referente aos dados).

`xaxt` e `yaxt` - Especificar se o eixo x e y, respectivamente, deve ser mostrado. Quando o valor for "n" o eixo não é mostrado.

`bty` - Alterar o tipo de caixa (box) usado envolta do gráfico. Valores aceito "o", "l", "7", "c", "u" e "]" que resultam em contorno igual a forma do símbolo ou "n" para não mostrar o contorno.

Ajustes gráficos

```
plot(mtcars$mpg,type = "b",  
     xlab = "Variável preditora",  
     ylab = "Variável resposta", main = "Título",  
     cex.axis = 3,  
     col.axis = "darkblue",  
     font.axis = 3,  
     xaxt = "n",  
     bty = "]" )
```

Ajustes gráficos

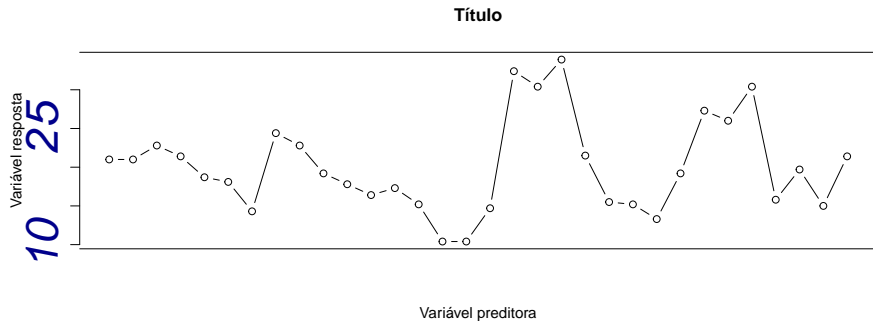


Figure 13: Gráfico de dispersão

Ajustes gráficos

A função auxiliar `points` adiciona pontos ou linhas ao gráfico.

Argumentos:

`x` e `y` - Valor único ou vetores com as coordenadas dos pontos para os eixos `x` e `y` respectivamente.

`pch` - Alterar o tipo de símbolo dos pontos. Valores aceitos entre 0 e 25. Se for um único valor todos os símbolos serão idênticos, se for um vetor cada ponto receberá um símbolo conforme especificado.

`cex` - Alterar o tamanho do símbolos. Argumento com valor numérico com padrão = 1. Se for um único valor todos os símbolos serão do mesmo tamanho, se for um vetor cada ponto será do tamanho especificado.

Ajustes gráficos



Figure 14: Gráfico de dispersão

Ajustes gráficos

Pode-se acrescentar nova variável e legenda. `lines` acrescenta nova variável; `lwd` controla espessura das linhas

```
plot(rnorm(100,5,3), type = "l", lwd=2, lty = 1)
lines(rnorm(100), col = 2, lwd = 2, lty = 2)
abline(h = 7.5, # acrescenta linha simples
       lty = 3,
       col = "darkgreen")

legend(x = "topleft",
      legend = c("N(5,3)", "N(0,1)"),
      lty = c(1,2),
      lwd = 2,
      bty = "n")
```

Ajustes gráficos

