

# Entrada de dados externos

**Microsoft Excel:** Salvar a planilha do excel em **.csv** (valores separados por vírgulas);

```
> dados = read.csv(  
  "caminho_e_nome_do_arquivo_.csv", ...)
```

**sep=";"** : caractere utilizado para a separação dos campos e valores.

**dec=","**: caractere utilizado para separar as casas decimais.

**header= TRUE** : 1ª linha contém rótulos das variáveis

**FALSE**: dados se iniciam na primeira linha

# Exemplo

- Abra uma planilha do Excel e crie os seguintes dados:

Pessoa	Idade	Peso
1	18	55,3
2	19	63,1
3	18	59,2
4	19	61,0
5	19	68,4

- Salve o arquivo na área de trabalho como **EXEMPLO.csv**

- **Importe o dados provenientes do arquivo EXEMPLO.csv:** identifique o caminho a ser digitado no R.

> Exemplo = **read.csv**("C:/Documents and Settings/Desktop/EXEMPLO.csv", sep=";", dec=".", header=TRUE)

> Exemplo

	Pessoa	Idade	Peso
1	1	18	55.3
2	2	19	63.1
3	3	18	59.2
4	4	19	61.0
5	5	19	68.4

# Mais funções

- **ls()**: apresenta os objetos criados na sessão atual;
- **summary(X)**: Min/Q1/Md/Me/Q3/Max;
- **subset(table(X),table(X) == max(table(X)))**: retorna a moda;
- **choose(N,n)**: combinação de N tomado n a n;
- **factorial(x)**: fatorial;
- **sqrt(x)**: raiz quadrada.

# PLOT

**> plot(X,Y, ...)**

**type=** “p” (pontos)

“l” (linhas)

“b” (pontos e linhas)

“h”, (linhas verticais)

**pch = 1...** Escolhe o tipo de caractere

**lwd = 1...** Espessura do caractere.

# Gráfico de pizza

> `pie(X, ...)`

`labels = c("Rótulo1", "Rótulo2")` (vetor contendo os rótulos de cada fatia)

`radius = 1` (raio da circunferência da pizza)

`col = c(col1, col2,...)` (vetor contendo as cores das fatias)

`main = "Título",`

# HISTOGRAMA

**> hist(X, ...)**

**breaks=c(x1, x2, ),** (vetor contendo os pontos de definição das larguras das barra do histograma)

**freq= TRUE ou FALSE,** (FALSE para  $f_r$ )

**main="TÍTULO",** (NULL para branco)

**xlab="TÍTULO EIXO x",**

**xlim=c(start, end),**

**ylim=c(start, end),**

**ylab=" TÍTULO EIXO y",**

**col="COR DAS COLUNAS (em inglês)"** colors()

# BOX PLOT

```
> boxplot(X,Y, ...)
```

```
xlab="TÍTULO EIXO x",
```

```
ylab=" TÍTULO EIXO y",
```

```
col="COR DE PREENCHIMENTO DA CAIXA"
```

```
names=c("nome caixa 1" ,"nome caixa 2")
```



# Configurações de gráficos

Configurações controladas pela função `par`.

- **mfrow=c(l,c)/mfcol=c(l,c):** (divide a janela onde os gráficos serão construídos),
- **ps=n** (configura o tamanho dos textos nos gráficos),
- **bg** (cor do fundo da janela do gráfico),
- **fg** (cor dos eixos dos gráficos) ,
- **col.main / col.lab / col.sub / col.axis** (cor)
- **cex.main/ cex.lab / cex.sub / cex.axis:** (tamanho)

- **fonte.main / font.lab / font.sub / font.axis**  
controla a fonte da seguinte forma:

→ 1: Normal

→ 2: Negrito

→ 3: Itálico

→ 4: Negrito + Itálico

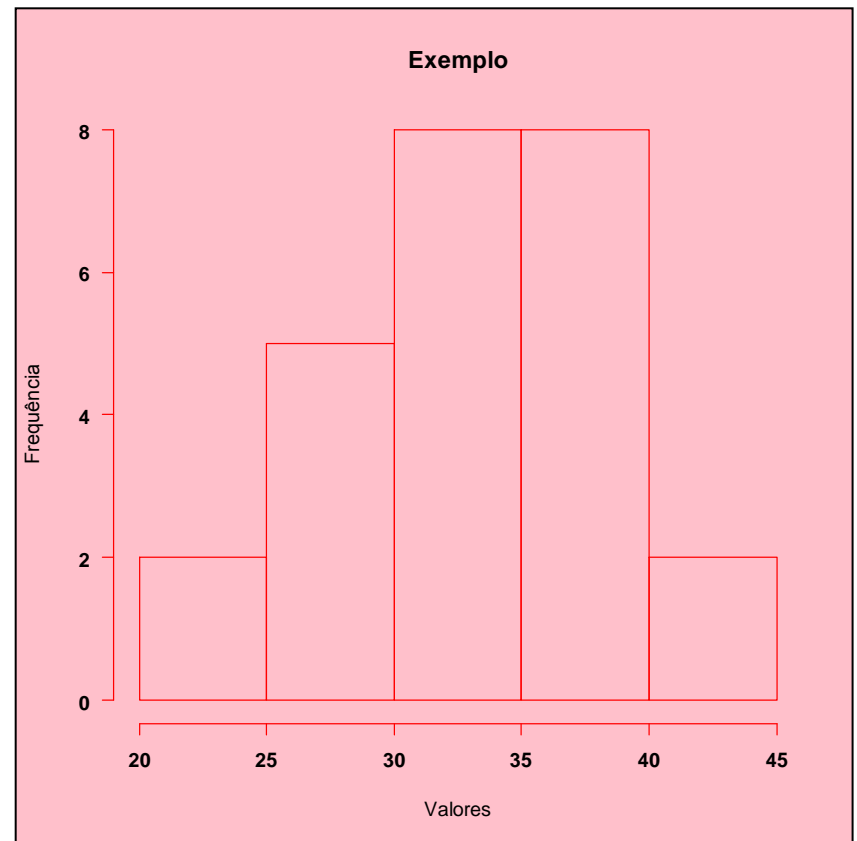
→ ... 20

- **las=1**: posiciona os valores do eixo X e Y na horizontal.

# Exemplos

```
par(bg="pink",fg="red",las=1,font.axis=2)
```

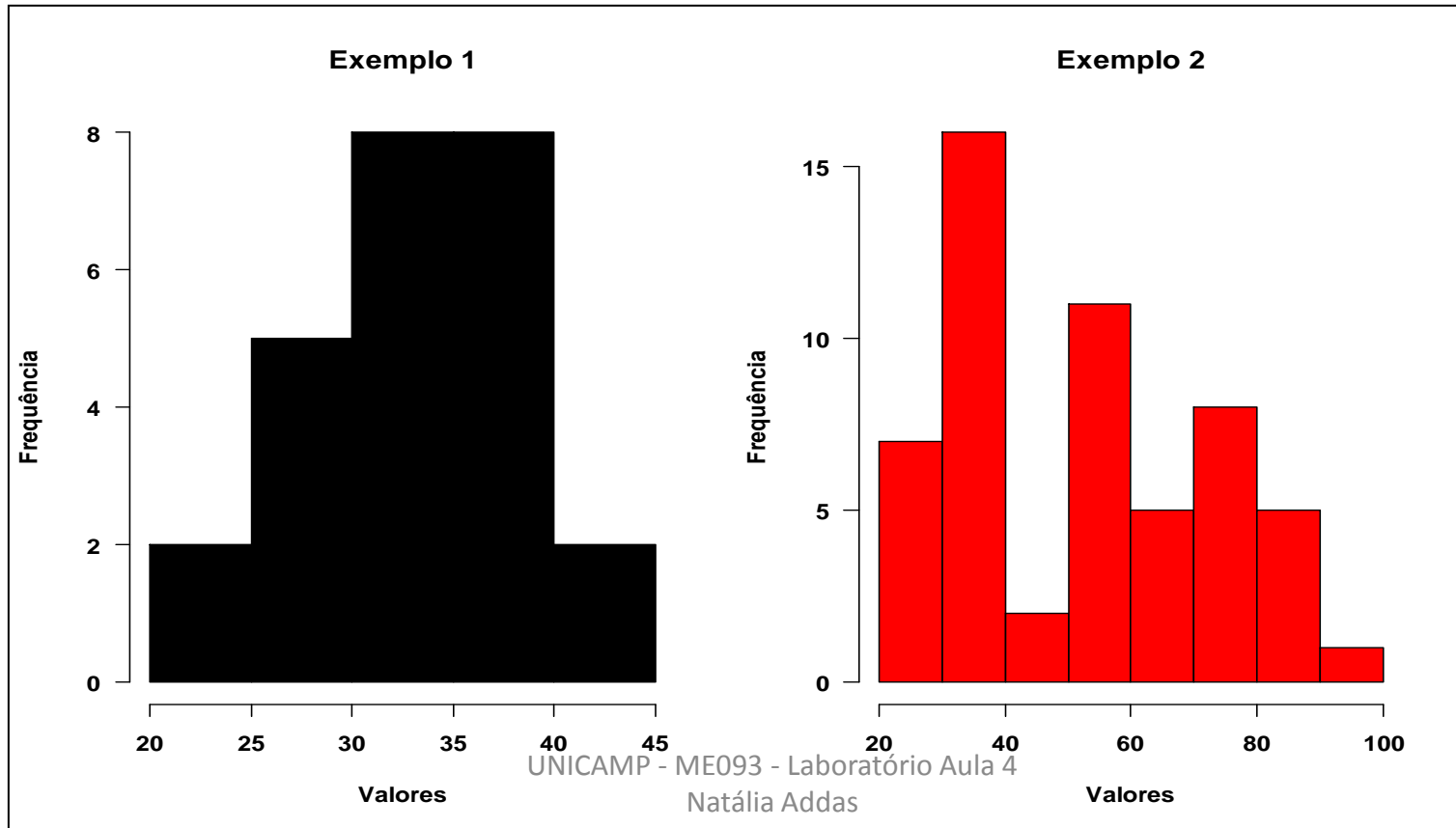
```
hist(A, main="Exemplo", xlab="Valores",  
ylab="Frequência")
```



```
par(mfcol=c(1,2),las=1,font.axis=2, font.lab=2)
```

```
hist(A, main="Exemplo 1", xlab="Valores",  
ylab="Frequência",col="black")
```

```
hist(S, main="Exemplo 2", xlab="Valores",  
ylab="Frequência",col="red")
```



1) Crie um arquivo **.csv** com as seguintes alturas de 20 indivíduos.

<b>Indivíduo</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Altura</b>	67.75	72.25	66.25	72.25	71.25
<b>Indivíduo</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<b>Altura</b>	74.75	69.75	72.5	74	73.5
<b>Indivíduo</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>
<b>Altura</b>	74.5	76	69.5	71.25	69.5
<b>Indivíduo</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
<b>Altura</b>	66	71	71	67.75	73.5

a) Importe os dados para o R.

b) Encontre o **Min/Q1/Md/Me/Q3/Max** e a **Moda** da variável **Altura**.

- c) Faça uma tabela de distribuição de frequência absoluta.
- d) Faça uma tabela de distribuição de frequências relativa.
- e) Apresente um histograma dos dados, com eixo y representando as frequências relativas.
- f) Considere os seguintes intervalos para as realizações da variável Altura.

Intervalo	1	2	3	4	5
	[66,68)	[68,70)	[70,72)	[72,74)	[74,76]

Faça um novo histograma estabelecendo os limites dos intervalos descritos acima.

2) Os dados a seguir são referentes à porcentagem da população economicamente ativa empregada no setor primário e o respectivo índice de analfabetismo para algumas regiões metropolitanas brasileiras.

Regiões Metropolitanas	Setor Primário	Índice de Analfabetismo
São Paulo	2,0	17,5
Rio de Janeiro	2,5	18,5
Belém	2,9	19,5
Belo Horizonte	3,3	22,2
Salvador	4,1	26,5
Porto Alegre	4,3	16,6
Recife	7,0	36,6
Fortaleza	13,0	38,4

*Fonte: Indicadores Sociais para Áreas Urbanas, IBGE, 1977, apud Morettin & Bussab, Estatística Básica 5ª edição, pag 90.*

Faça um **arquivo.csv** das 3 variáveis.

**b) Encontre o fator de correlação entre as variáveis **Setor Primário** e **Índice de Analfabetismo**.**

**c) Faça um diagrama de dispersão entre as variáveis **Setor Primário** e **Índice de Analfabetismo**.**

**d) Identifique o ponto mais afastado da “reta ” do diagrama anterior. Refaça o cálculo da correlação excluindo tal ponto.**

**e) Faça um **boxplot** contendo as variáveis **Setor Primário** e **Índice de Analfabetismo** em uma única janela gráfica.**