

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 – ";" padrão.

dec=",": caractere que foi utilizado para separar as casas decimais na planilha excel.

header= TRUE : 1ª linha contém rótulos das variáveis

FALSE: dados se iniciam na primeira linha

Mais funções

- **cor(X, Y)**: retorna a correlação entre a variável X e Y.
- **choose(N,n)**: combinação de N tomado n a n;
- **factorial(x)**: fatorial;
- **sqrt(x)**: raiz quadrada.
- **text(coox, cooy, c("texto1", "texto2",...))**: insere um texto nas coordenadas do eixo x e y (coox, cooy).
- **text(locator(n), c("texto1", "texto2",...))**: insere *n* textos nas coordenadas do eixo x e y e permite que você manipule-os com o mouse.

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.

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")

outline=TRUE ou FALSE (plota e não plota outliers)

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",

1) 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**, colocando o nome da **Região Metropolitana** em seu respectivo ponto.

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

e) Identifique o ponto mais afastado da “reta ” do diagrama de dispersão do item c. Refaça o cálculo da correlação excluindo tal ponto (*excluir usando a ideia de extração de elementos []*).

2) Um biólogo interessado em saber se o número de aves está relacionado ao número de uma determinada espécie de árvore realizou amostras em 10 locais. Os valores obtidos foram:

Aves	22	27	37	34	13	24	39	5	33	32
Árvores	25	26	40	30	10	20	35	8	35	28

a) Faça um gráfico de pizza com a média das duas variáveis.

b) Faça um gráfico que mostra a relação entre o número de aves e o número de árvores e compare com a correlação entre tais variáveis.