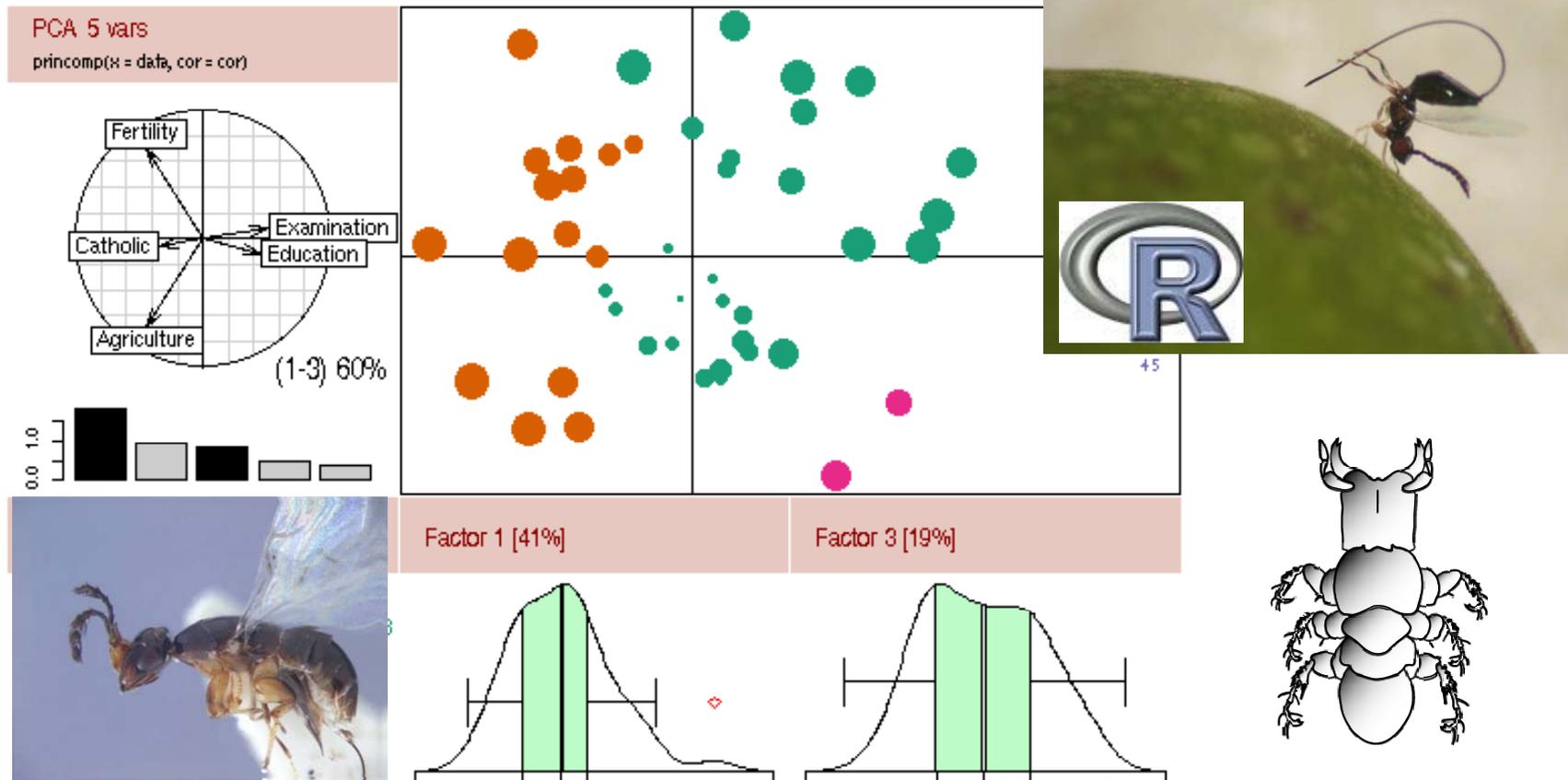


Uso da Linguagem R para Análise e Dados em Ecologia



Rodrigo A. S. Pereira
Depto de Biologia/USP-Ribeirão Preto

Objetivo da aula

**Apresentar os princípios básicos de
Análise Exploratória de Dados.**



Vocês me
conhecem?

Conhecer melhor os Dados



Análise Exploratória de Dados (AED)

- ✓ O que é AED?
- ✓ Quem utiliza AED?
- ✓ Objetivos da AED:
 - conhecer os dados
 - detectar padrões
 - orientar os testes de hipóteses
 - ajudar a interpretar os testes de hipóteses
- ✓ **Apresentação gráfica é muito útil para AED**

Explorando numericamente...

Algumas possibilidades no R

✓ Espiada básica:

```
summary(), str(), head(), tail()
```

✓ Procurar valores perdidos

```
is.na()
```

✓ Medidas de tendência central e dispersão

```
mean(trim=), mean(), median(), sd(), range()
```

Exemplos no R

Explorando visualmente...

Princípios básicos da apresentação gráfica:

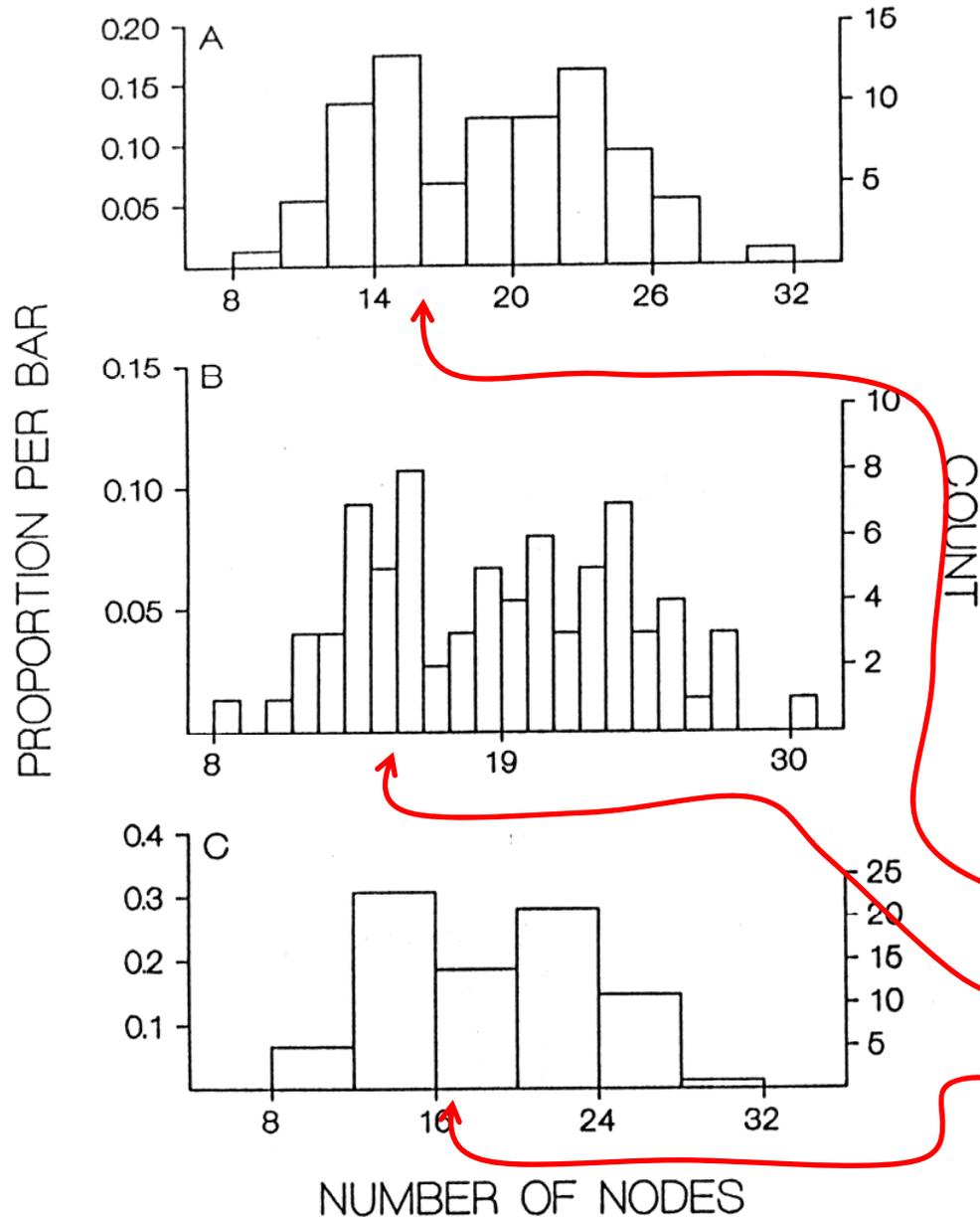
- ✓ Ressaltar os padrões de interesse;
- ✓ Manter a estrutura dos dados, de forma que o leitor possa reconstruir os dados a partir da figura;
- ✓ A figura deve ter uma razão **dado:tinta** alta;
- ✓ As figuras **não** devem distorcer, exagerar ou aparar os dados.

Tipos de gráficos para AED

Dados univariados

- ✓ Histograma: gráfico de densidade. Não confundir com gráfico de barras
 - Desvantagens: **esconde** os dados (**mostra apenas intervalos**) e as classes são **arbitrárias**.

Histograma



Número de nós por planta em 75 indivíduos de *Ailanthus altissima*.

A) Intervalo = 2

B) Intervalo = 1

C) Intervalo = 4

Tipos de gráficos para AED

Haste e folha

- ✓ Dados brutos são apresentados
- ✓ Mesmo problema visual que o histograma: classes são arbitrárias

0 8
1 0111
1 2223333333
1H 4444455555555
1 66777
1M 888889999
2 000000111
2H 2222233333333
2 4445555
2 6777
2
3 0

← Cada linha é uma haste

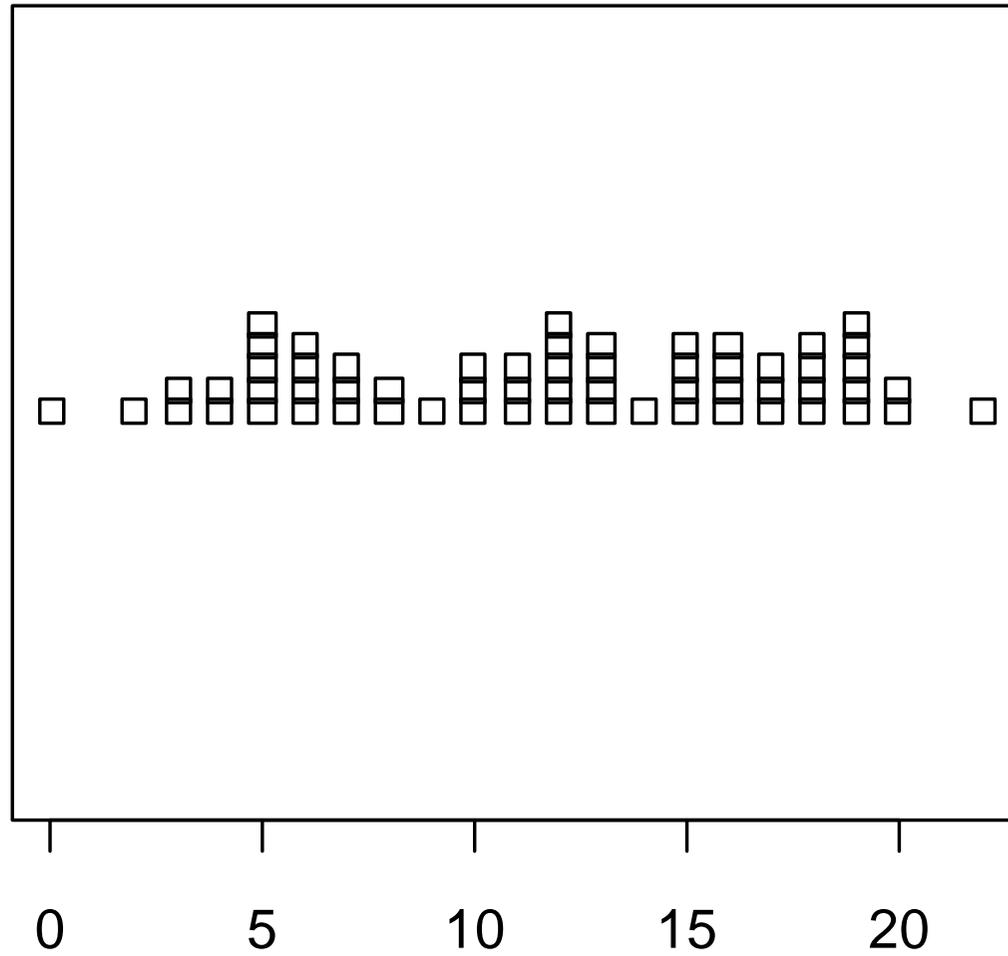
← Cada dado é uma folha

Exemplo:

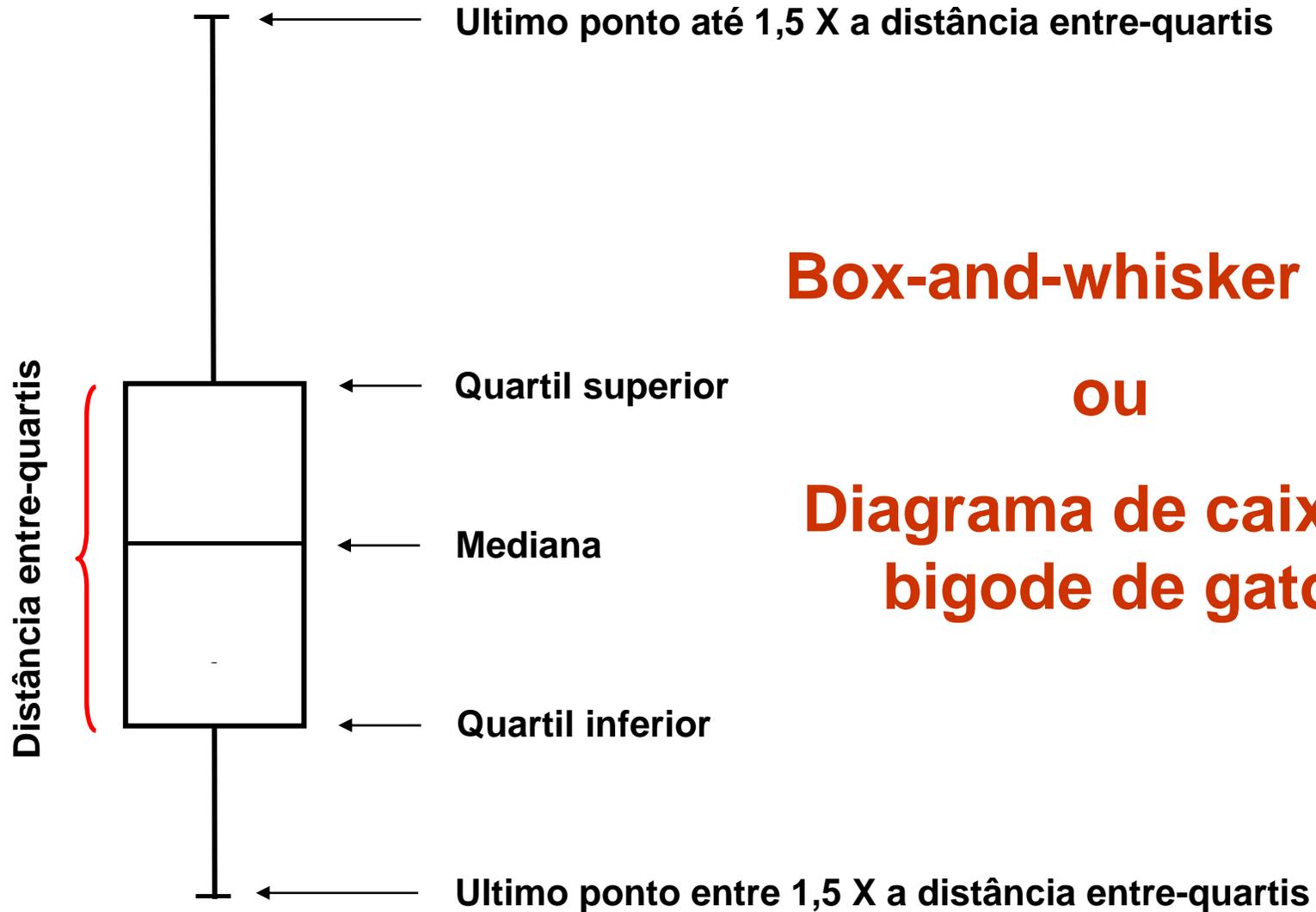
08
10, 11, 11, 11

M = mediana
H = quartil

Dit plot



* ← Valor extremo: + que 1,5 X a distância entre-quartis



Box-and-whisker plot

ou

Diagrama de caixa e bigode de gato

Tipos de gráficos para AED

Histogramas, haste e folha e dit plot podem representar possível bimodalidade nos dados

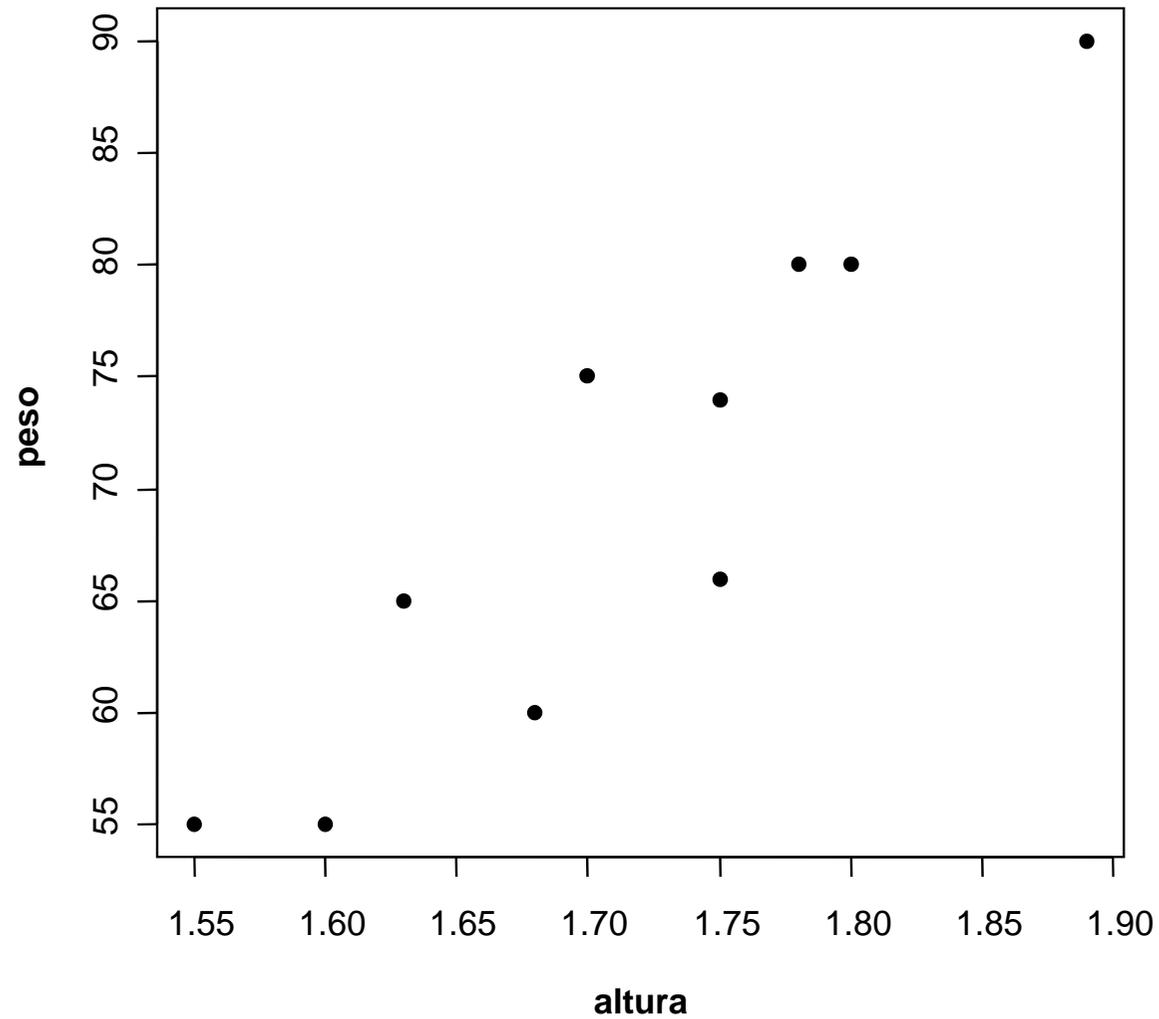
Box-plot esconde esse padrão

Exemplos no R

Dados bivariados

Dispersão

- ✓ Permite analisar a relação entre duas variáveis



Tipos de gráficos para AED

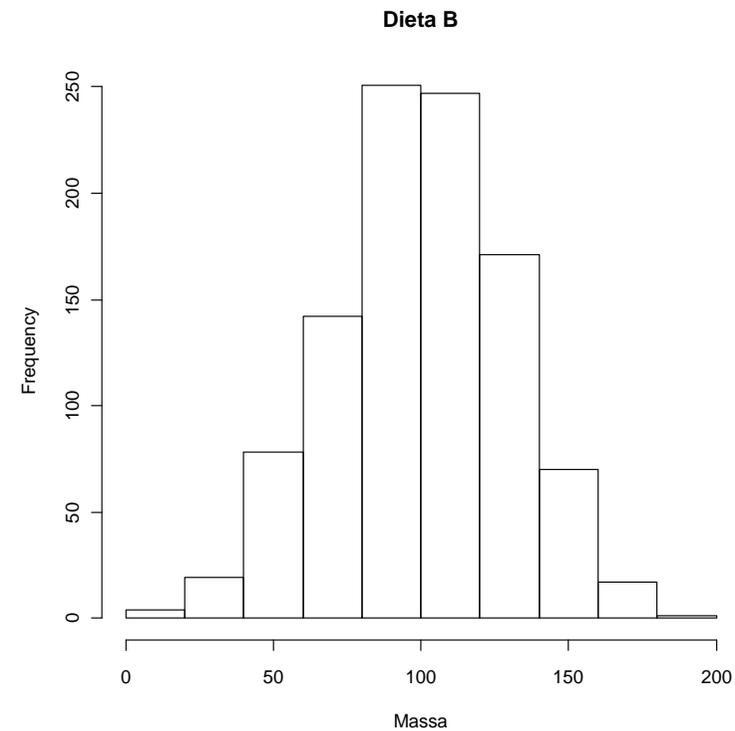
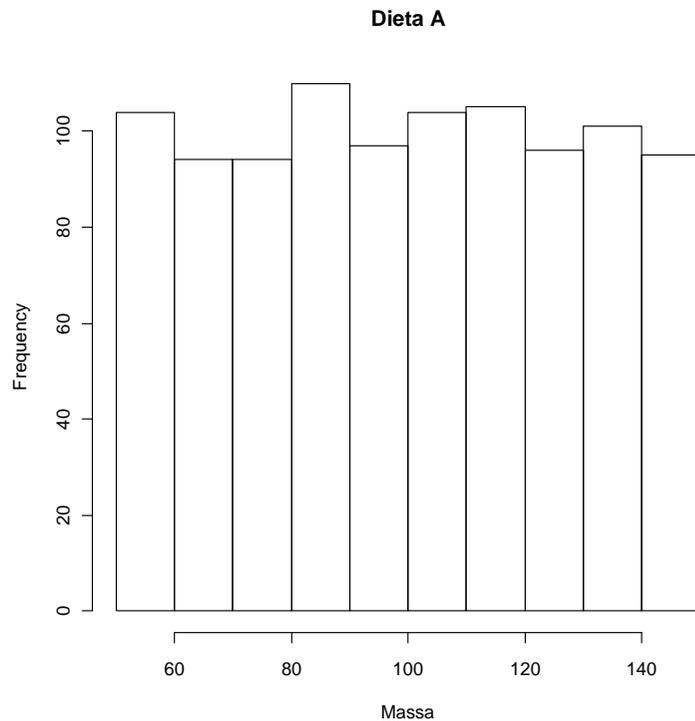
Quantil-Quantil

- ✓ Permite comparar a forma de duas distribuições
- ✓ Muito usado para **avaliar** se os dados têm **distribuição parecida** com alguma **distribuição teórica** (Normal, Poisson etc).

Tipos de gráficos para AED

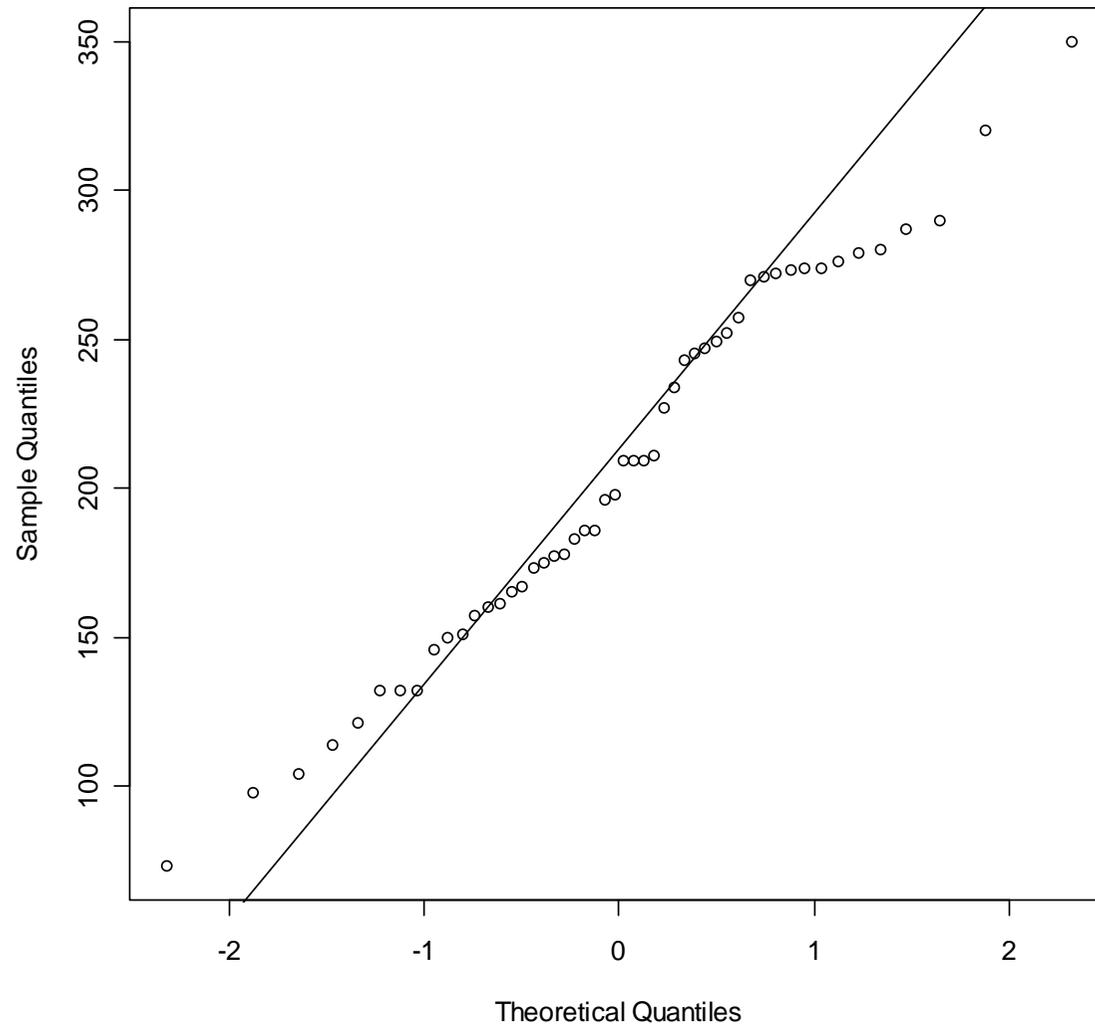
Distribuição é importante?

- ✓ **Massa larval de insetos criados em duas dietas ($X \pm SD$):**
 - **Dieta A**: $99,8 \pm 29,6$ mg (N = 1000 larvas)
 - **Dieta B**: $100,1 \pm 30,3$ mg (N = 1000 larvas)



Quantil-quantil

Normal Q-Q Plot

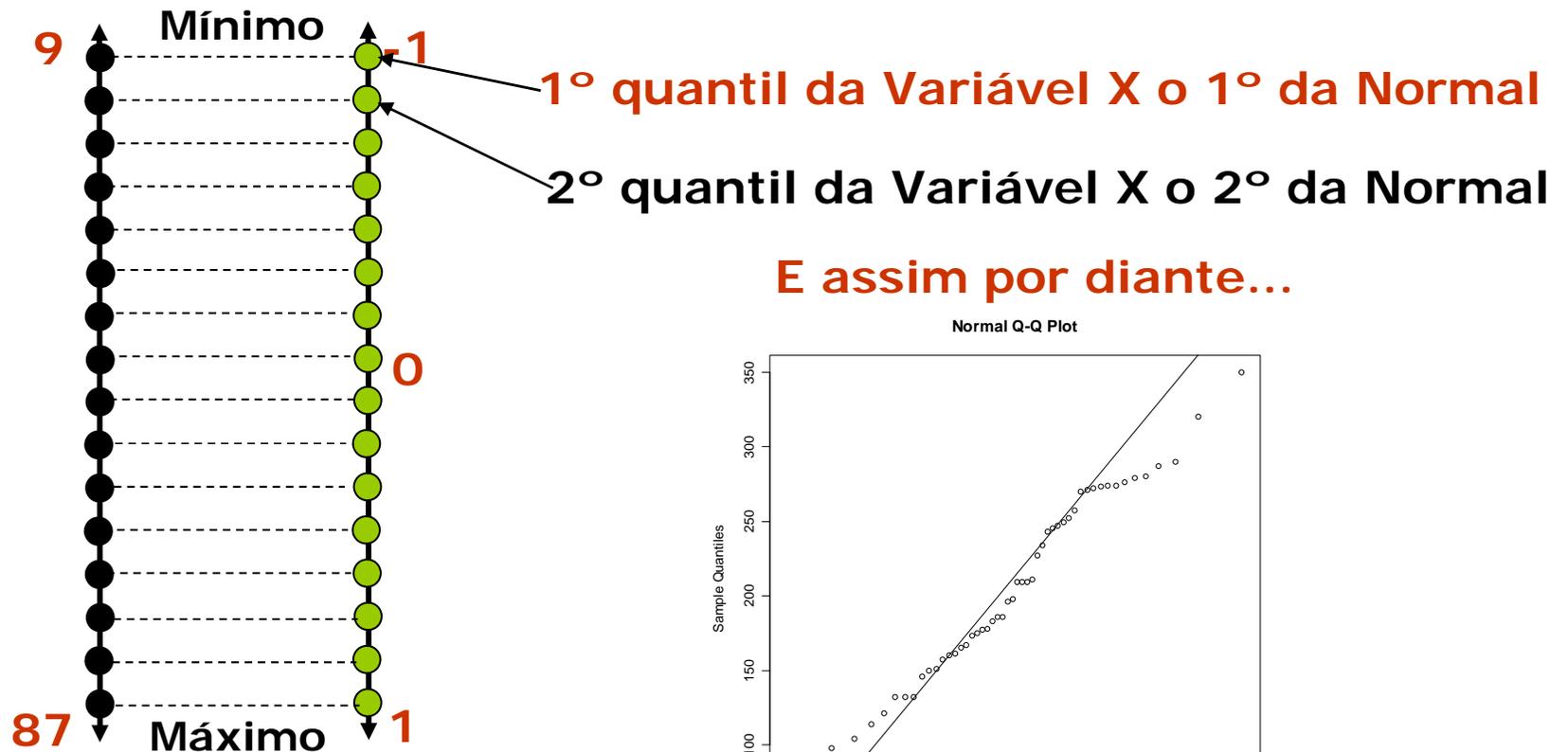


Quantil-quantil

1º) Ordenar as variáveis

2º) Plotar os quantis uns contra os outros

Variável Normal



E assim por diante...

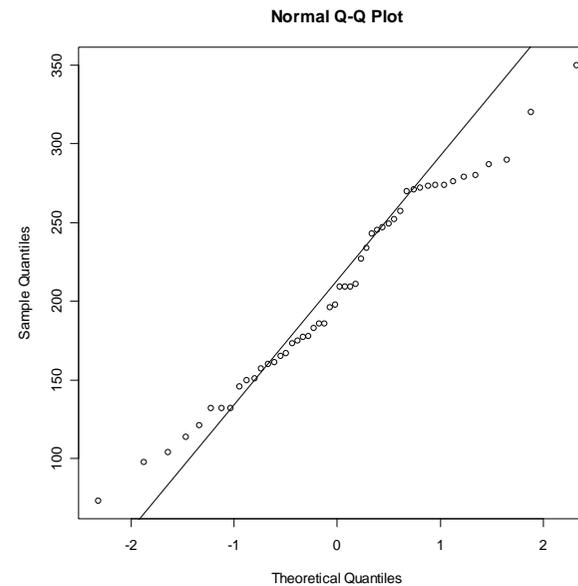
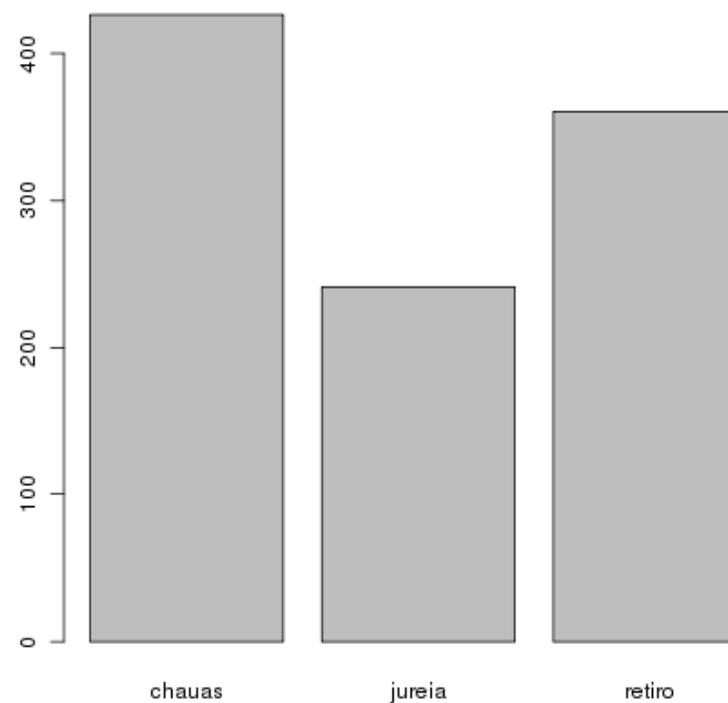


Gráfico de Barras

```
> table(caixeta$local)
```

```
chauas  jureia  retiro  
  426    241    360
```



Exemplos no R

Sugestão de leitura

Ellison, A. M. 1993. Exploratory data analysis and graphic display. In: Scheiner, S. M. (ed.), *Design and analysis of ecological experiments*. Chapman & Hall, pp. 14-45.

McGill, R., Tukey, J. W. and Larsen, W. A. 1978. Variations of Box Plots. *Am. Statist.* 32: 12-16.

