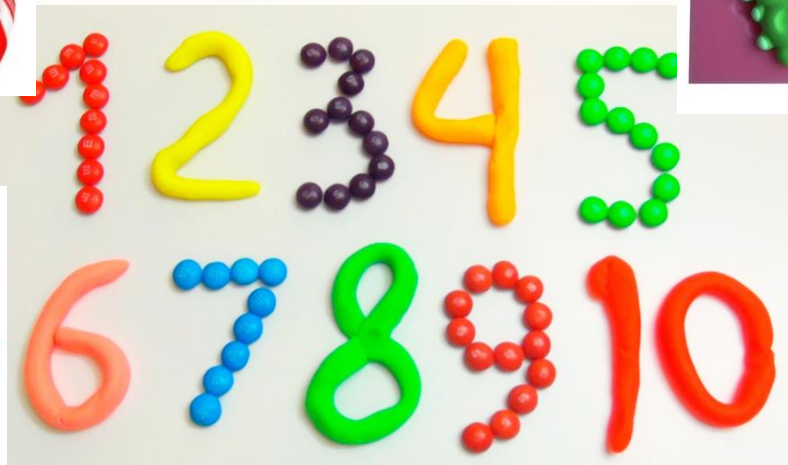


# Goodies\*



\* Goodies related to animals, plants and numbers...

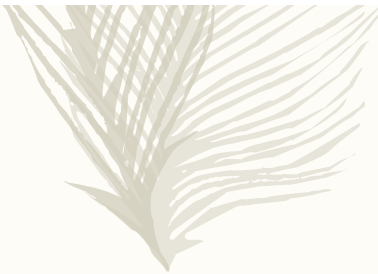
## Subjects

[Study options](#)[Non-degree courses](#)[Entry requirements](#)[Qualification indicator](#)[Academic entry explained](#)

[University](#) > [Subjects](#) > [Statistics](#) > Statistical Ecology MSc

# Statistical Ecology (MSc) 2020 entry

In this course, you will learn the modern statistical methods currently used by professionals in ecology. You will learn to formulate problems, conduct appropriate analyses and effectively communicate results to a variety of audiences. To balance theory and application, placement opportunities will be available with partner organisations within the UK and abroad.

[Apply now](#)

## Course type

Postgraduate, leading to a Master of Science (MSc)

## Course dates

- Start date: 7 September 2020
- End date: 30 September 2021

## Course duration

One year full time or two years part time

## Entry requirements

You should have one of the following qualifications:

- A good 2.1 undergraduate Honours degree in a relevant discipline (e.g. biological sciences, ecology, mathematics, statistics, environmental science or computer science).
- A 2.2 in a relevant discipline and equivalent work experience (for example, at least 12 months working in a relevant field). You should also have undergraduate training in mathematics and statistics at SCQF Level 8 or equivalent experience.

If you studied your first degree outside the UK, see the [international entry requirements](#).

## Tuition fees

**UK and EU:** £9,450

**Overseas:** £19,400

- [Find out about scholarships and funding.](#)
- [Further particulars regarding tuition fees.](#)

## Application deadline

Wednesday 12 August 2020. Applicants should apply as early as possible to be eligible for certain scholarships and for international visa purposes.

## Application requirements

- CV
- personal statement (optional)
- two original signed academic references
- academic transcripts and degree certificates
- [evidence of English language proficiency](#) (required if English is not your first language).

For more guidance, see [supporting documents and references](#) for postgraduate taught programmes.

## Subjects

[Study options](#) [Non-degree courses](#) [Entry requirements](#) [Qualification indicator](#) [Academic entry explained](#)

[University](#) > [Subjects](#) > [Statistics](#) > [Statistical Ecology MSc](#)

### Statistical Ecology (MSc) 2020 entry

In this course, you will learn the modern statistical methods currently used by professionals in ecology. You will learn to formulate problems, conduct appropriate analyses and effectively communicate results to a variety of audiences. To balance theory and application, placement opportunities will be available with partner organisations within the UK and abroad.

[Apply now](#)



**Tiago André Marques** @TiagoALOMarques · Oct 11  
This is the MSc I would have loved to take a few years ago!!!



**Andy Seaton** @ASeatonSpatial · Oct 10  
University of St Andrews has a new MSc in Statistical Ecology - starting in 2020. Come study in a lovely department with lovely people. Let your students know!

For more info and to apply:  
[st-andrews.ac.uk/subjects/stati...](http://st-andrews.ac.uk/subjects/stati...)

Also feel free to DM me if you are interested



1



2



16



I challenge the best of you to go there. Probably the best place on earth to study statistical ecology (ecological statistics!), with a team of world leading researchers supporting the program.

If you do want to go, feel free to ask me about St Andrews & you will continue to have my support !

Any way, you can ask your colleagues from last year, I will give you my support regardless of where you go – just drop me an email 😊 !



No sentido de melhorar o seu desempenho académico, promovemos um conjunto de iniciativas dirigidas a todos os estudantes da Universidade de Lisboa.

**Iniciativas de Apoio ao Estudante (2019-20):** candidaturas abertas entre 1 e 15 de outubro.

As candidaturas deverão ser formalizadas no sistema Fénix da Reitoria, não tendo qualquer custo associado. A posterior inscrição em cada ação tem um custo associado de 10€.

**Iniciativas de apoio ao estudante (2019-20):**

- | Ativa o teu semestre!
- | Comunicare – Formação em Soft Skills
- | Software de apoio à investigação
- | Métodos de Trabalho (3º ciclo)
- | Design Thinking (sessão de esclarecimentos)

**Mais informações**

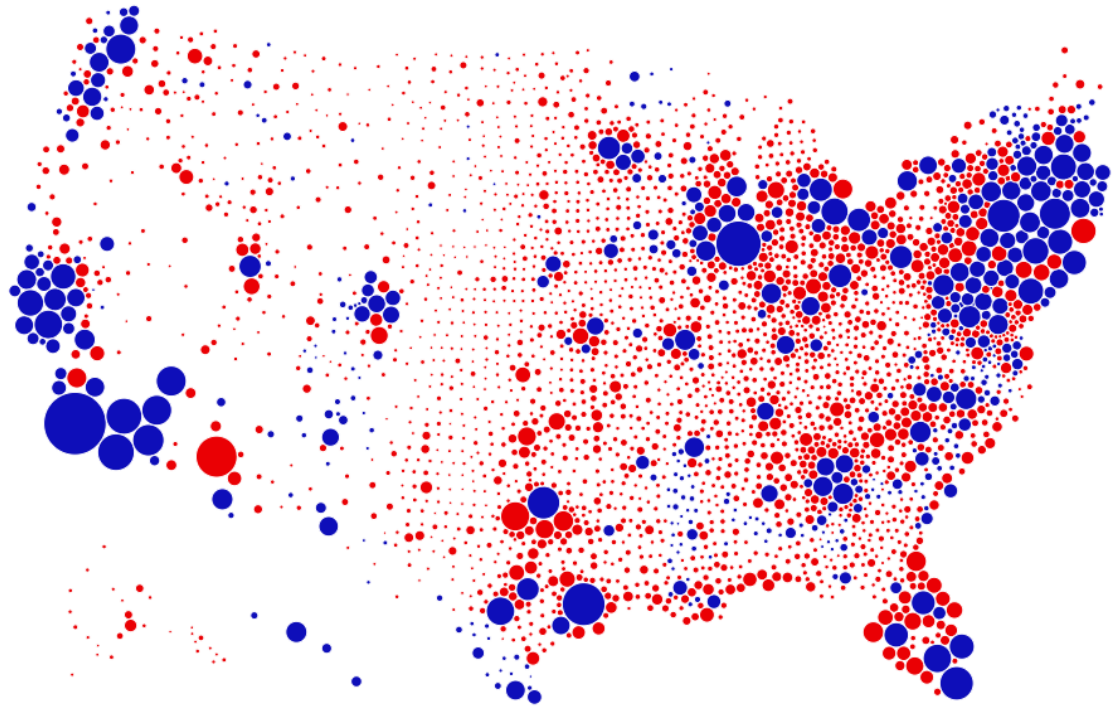
Lara Trump   
@LaraLeaTrump



83.6K 8:36 PM - Sep 28, 2019 - Briarcliff Manor, NY



Ilustração do poder da visualização  
(e manipulação) de dados



<https://observablehq.com/@karimdouieb/try-to-impeach-this-challenge-accepted>

**ILLEGAL**

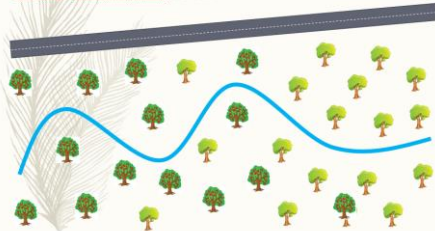
shutterstock.com • 545982400

TOP SECRET  
**CLASSIFIED**  
CONFIDENTIAL

## Protocolo de Amostragem

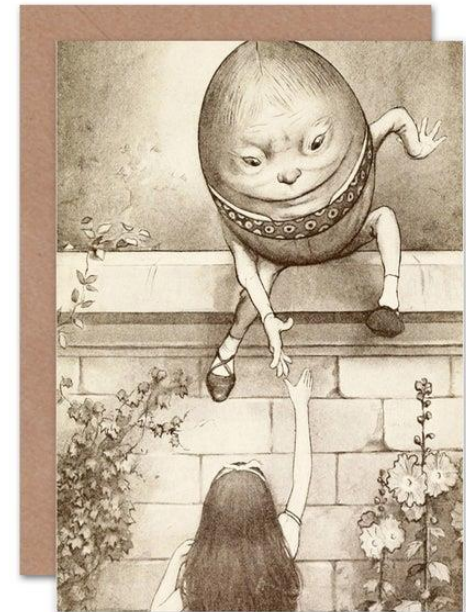
**WARNING: RGPD violation ....**

**Tarefa** (grupos de 3 ou 4):  
Defina um protocolo de amostragem para recolher um conjunto de dados que permita responder à seguinte pergunta: será que os insetos têm diferentes níveis de actividade durante o dia, e esses níveis de actividade dependem de qual próximo está de um rio? O esquema abaixo representa a área de estudo. Discuta as potenciais implicações e constrangimentos, justificando as escolhas que fizer. Enviar a resolução ao docente por e-mail.



Grupo no topo → + 0.5 no exame TP

os restantes nesta tabela, se a certa altura forem um ovo e estiverem em cima da cerca, provavelmente caem para o lado melhor...



Ecología Numérica - Aula Teórica 8 – 14-10-2019



The problem with assumptions is  
that we believe they are the truth.

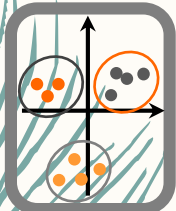
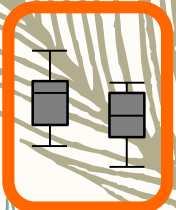
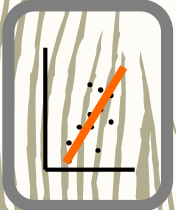
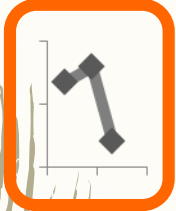
— Miguel Angel Ruiz —

AZ QUOTES

<https://www.azquotes.com/quote/520553>



# Ecologia Numérica



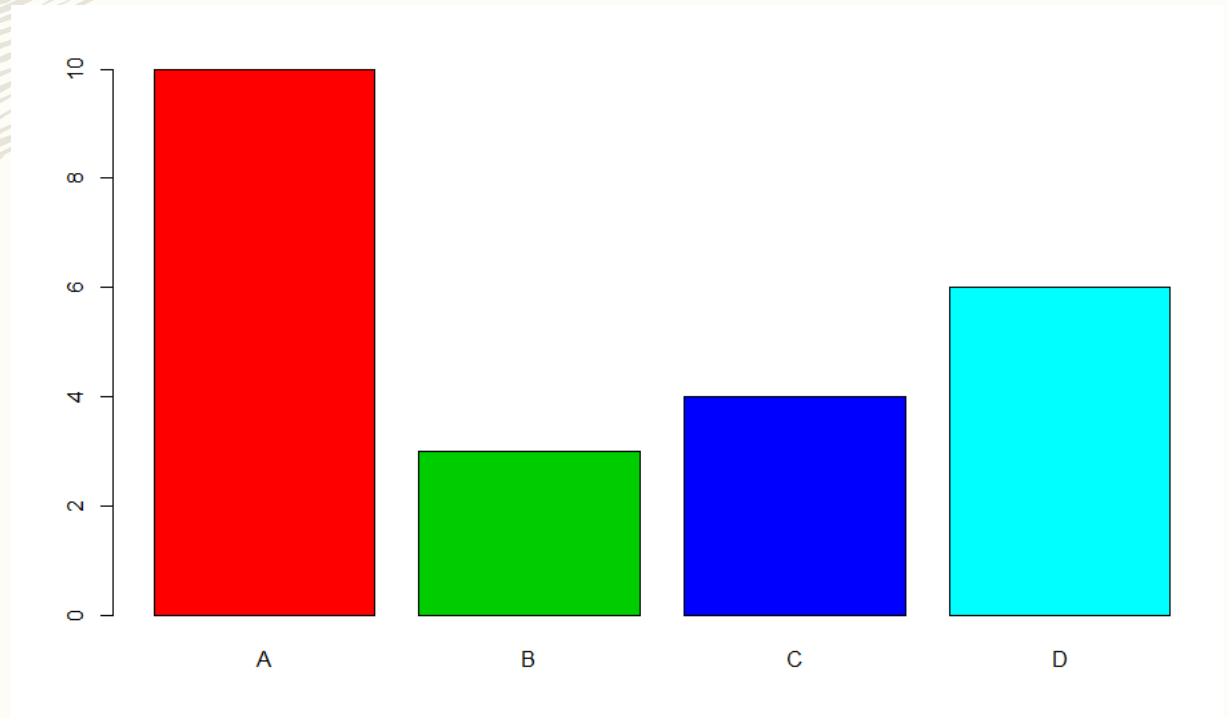
análise  
exploratória de  
dados

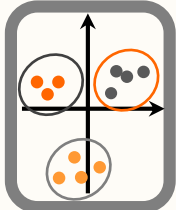
—  
Continuação...

# Gráficos de barras

Representar quantidades discriminadas por classes

```
barplot(c(10,3,4,6),names.arg=c("A","B","C","D"),col=2:5)
```



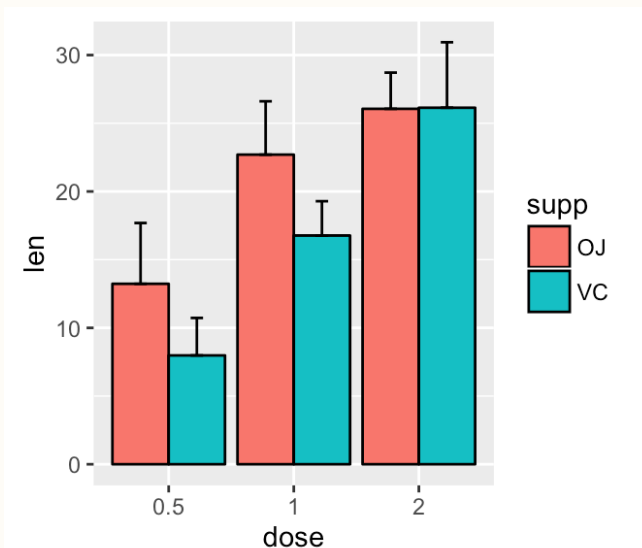


# análise exploratória de dados

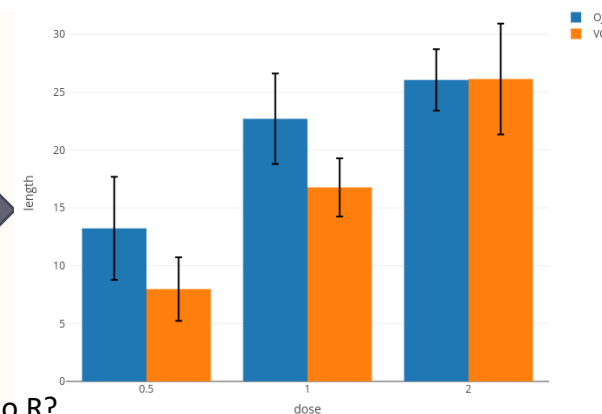
## Gráficos de barras

Não é a parte de cima de uma caixa com bigodes!

Representa a média e o erro padrão associado!!



This really is something like this!



Desafio: quem descobre como fazer barplots com barras de erro no R?

# Gráficos de barras – going beyond basics

```
par(mfrow=c(2,2),mar=c(4,4,2.5,0.5))
```

```
#default, with names
```

```
barplot(matrix(1:9,3,3),col=rep(2:4,times=3),names.arg =c("C1","C2","C3"))
```

```
#na horizontal
```

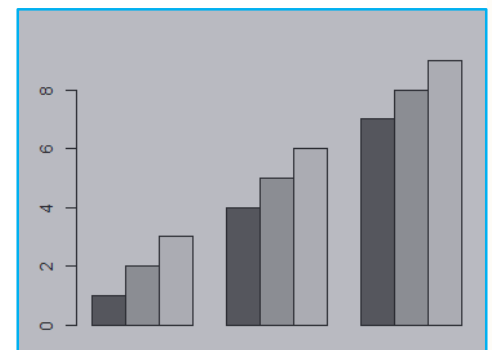
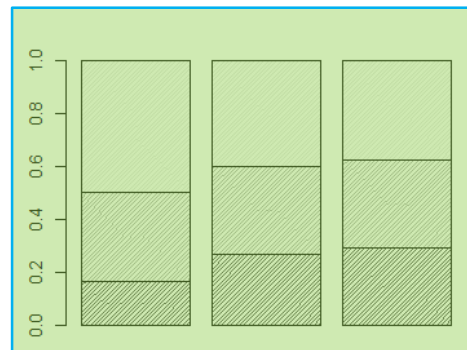
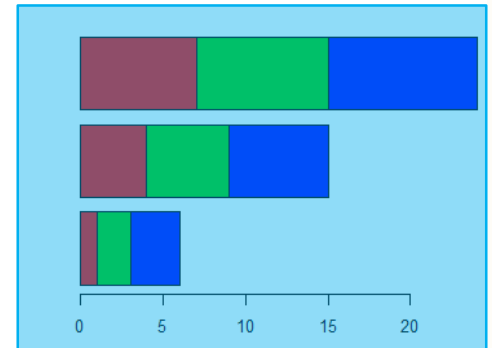
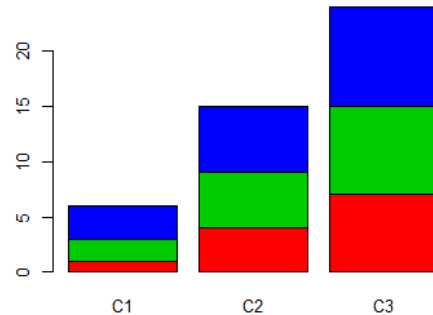
```
barplot(matrix(1:9,3,3),col=rep(2:4,times=3),horiz=TRUE)
```

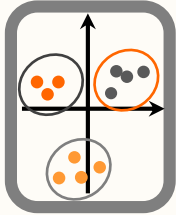
```
#com proporções
```

```
barplot(prop.table(matrix(1:9,3,3),2),density=40)
```

```
#lado a lado
```

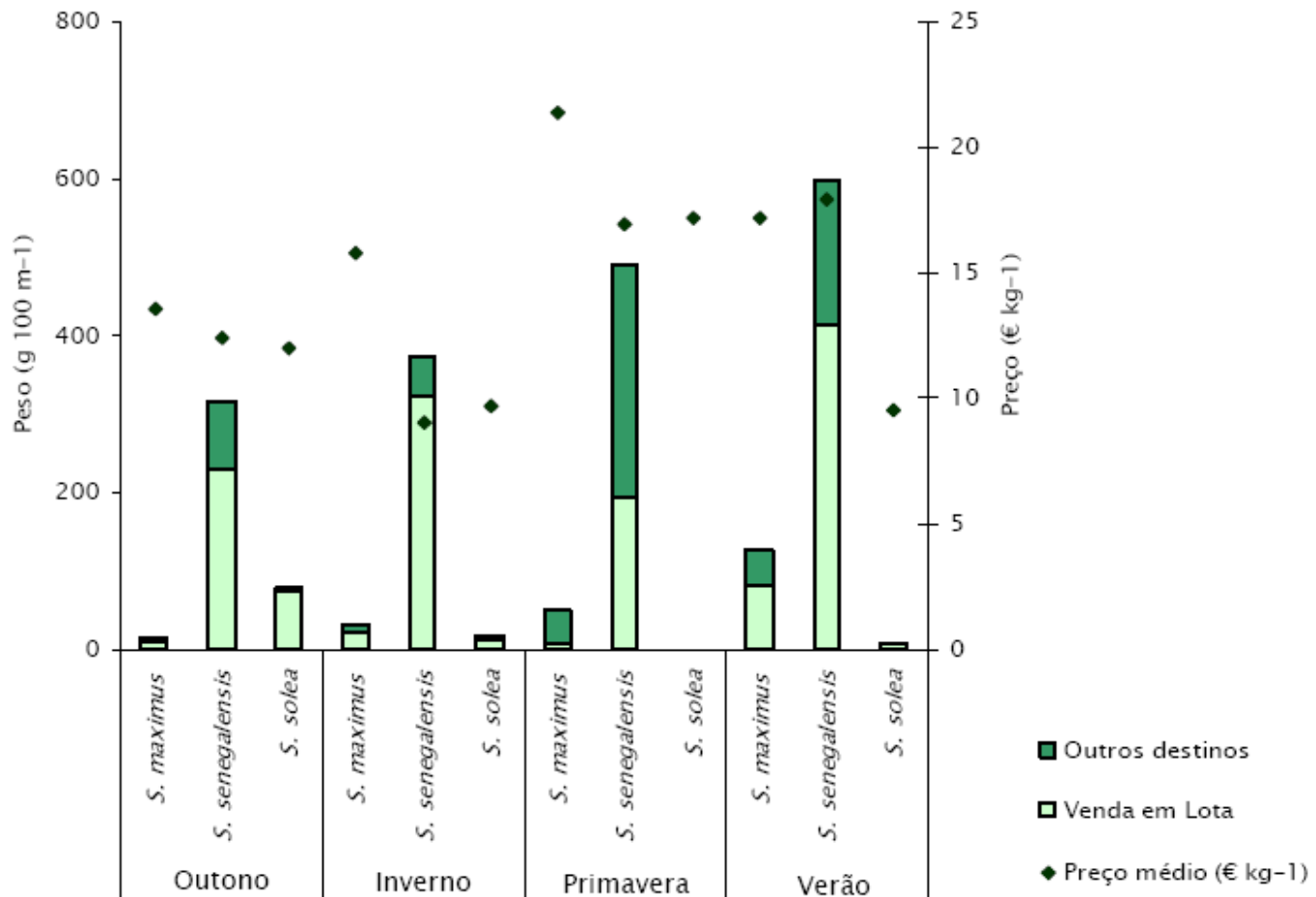
```
barplot(matrix(1:9,3,3),beside=TRUE)
```

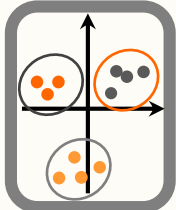




# análise exploratória de dados

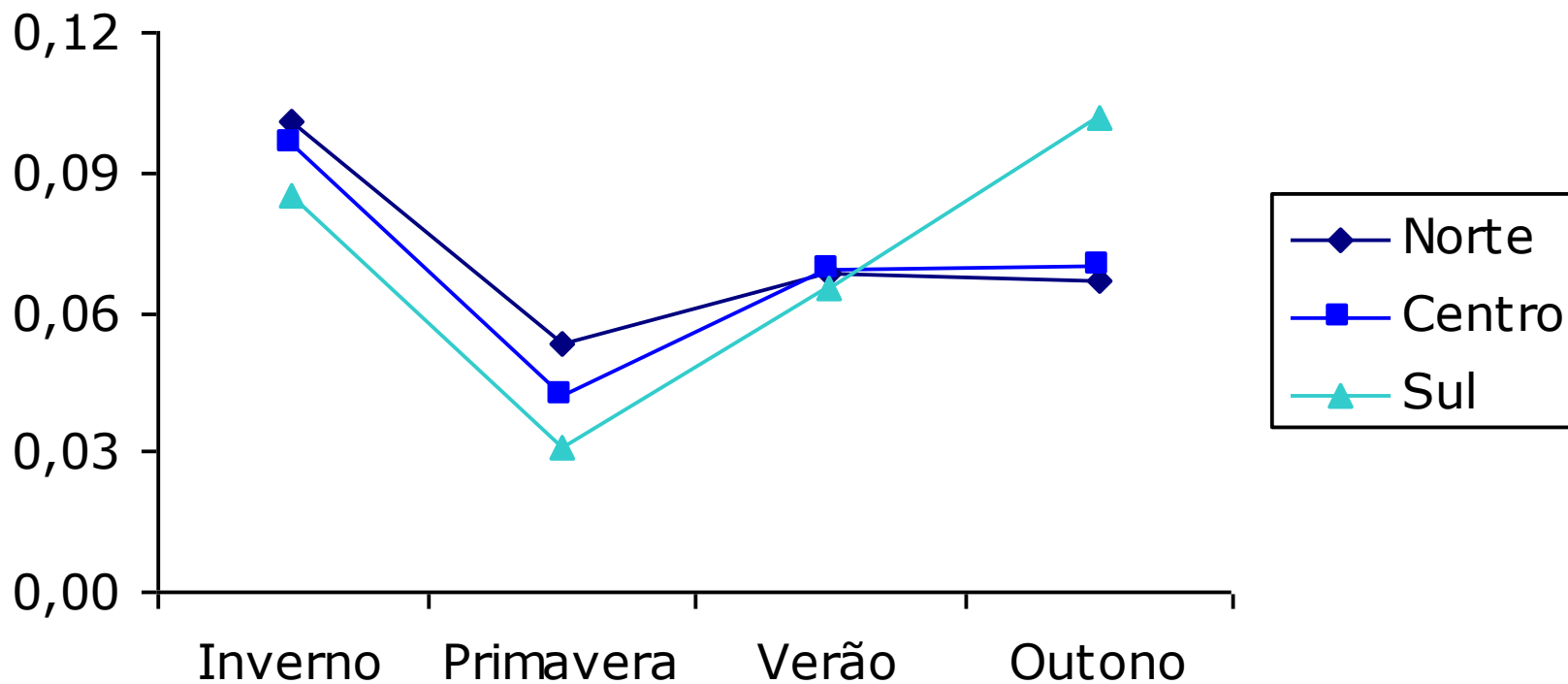
## Gráficos de barras





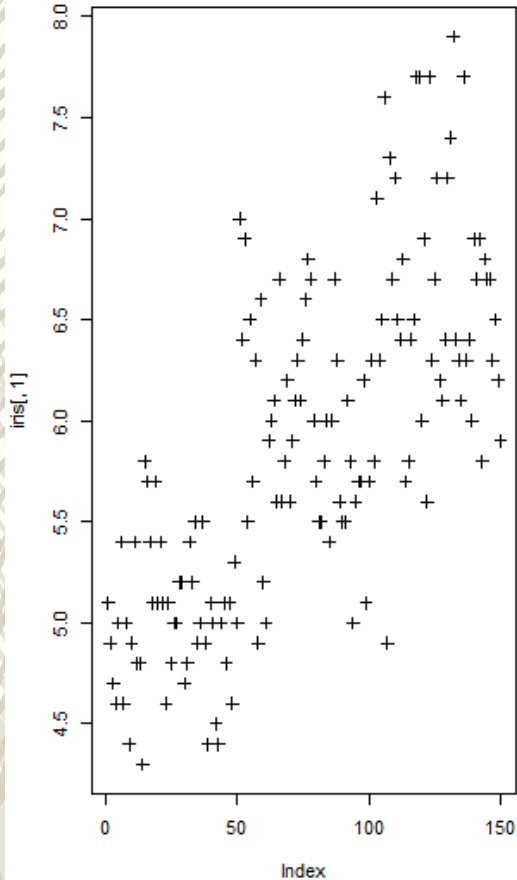
## Gráficos de linhas

### Machos

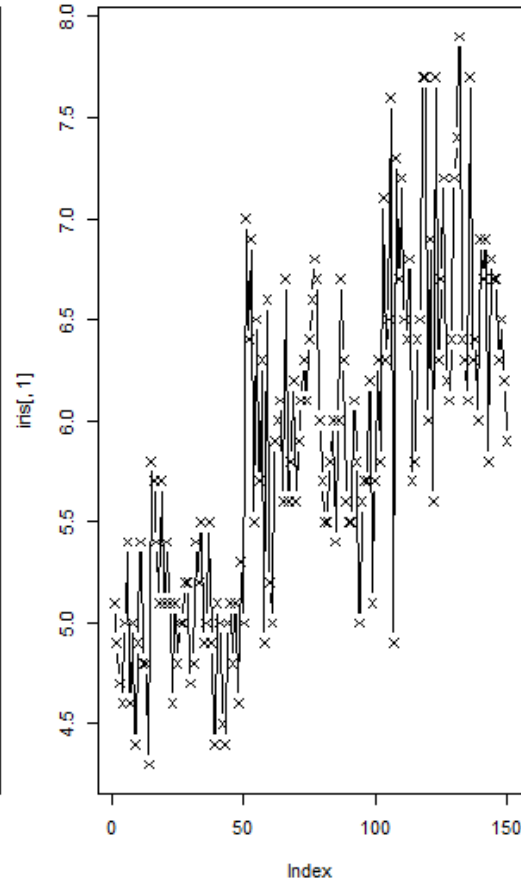


```
{r}  
par(mfrow=c(1,3),mar=c(4,4,2.5,0.5))  
plot(iris[,1],type="p",pch=3)  
plot(iris[,1],type="b",pch=4)|  
plot(iris[,1],type="l",col="yellow")  
}
```

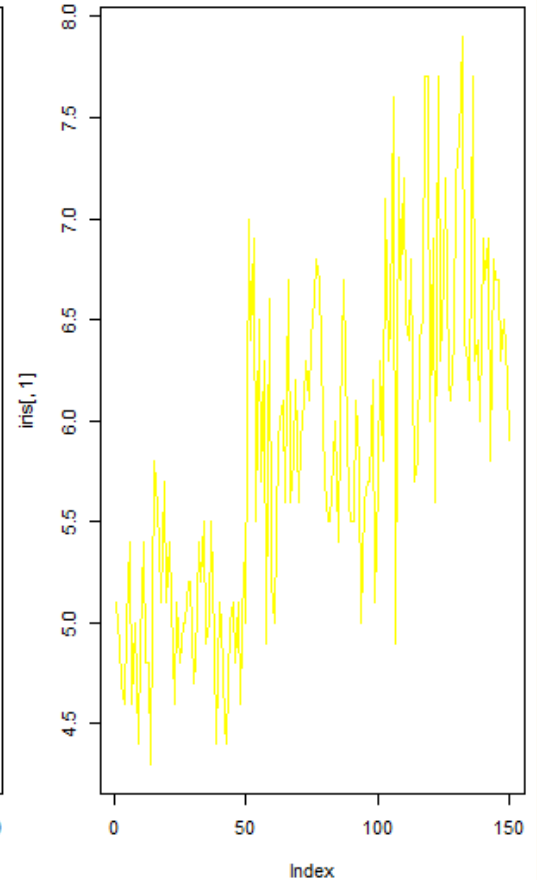
type="p" (points)



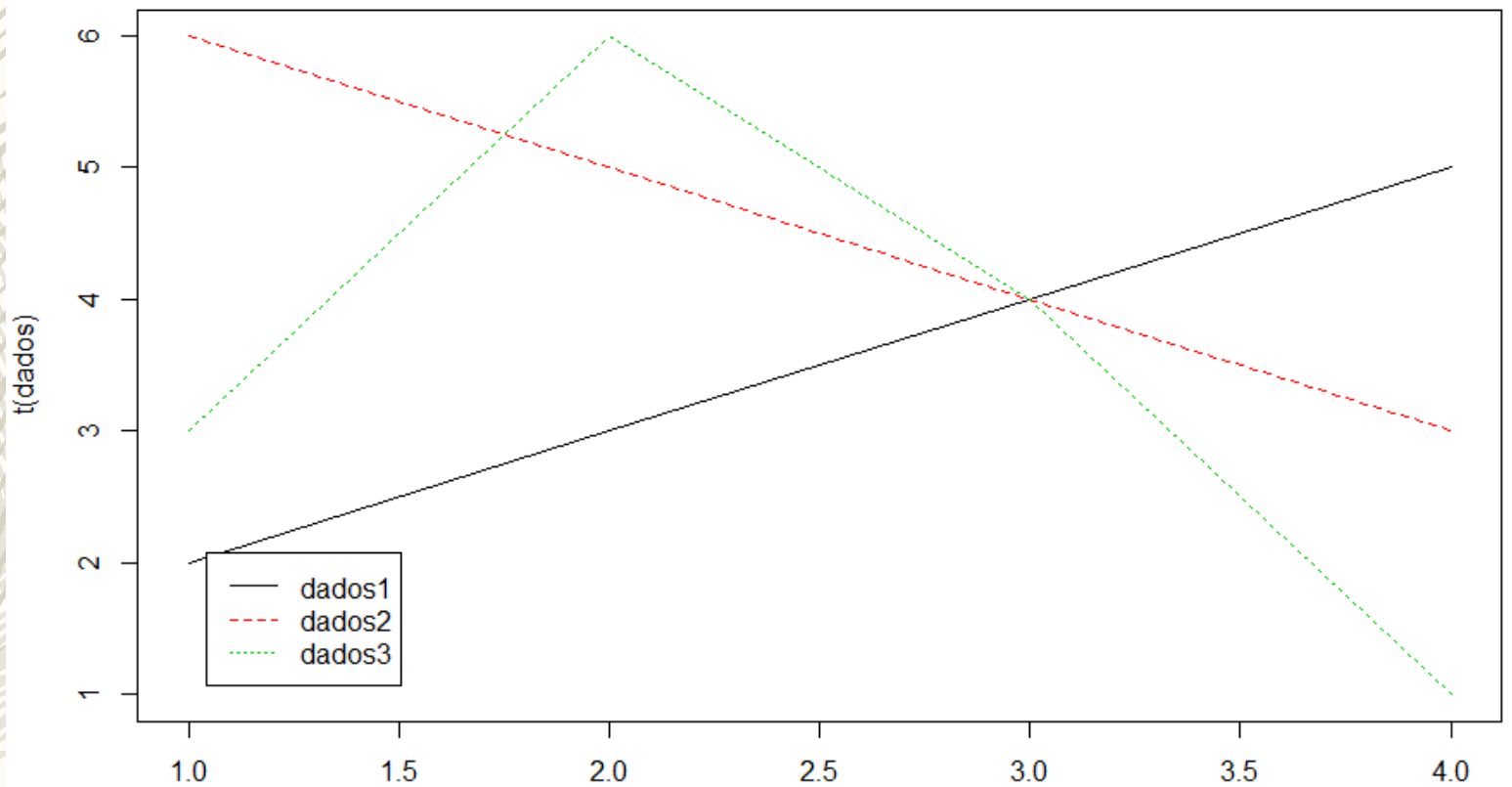
type="b" (both)



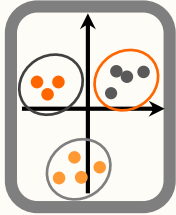
type="l" (lines)



```
{r}
dados1=c(2,3,4,5)
dados2=c(6,5,4,3)
dados3=c(3,6,4,1)
dados=rbind(dados1,dados2,dados3)
par(mfrow=c(1,1),mar=c(4,4,2.5,0.5))
matplot(t(dados),type="l")
legend("bottomleft",legend=c("dados1","dados2","dados3"),col=1:3,lty=1:3
,inset=0.05)
```

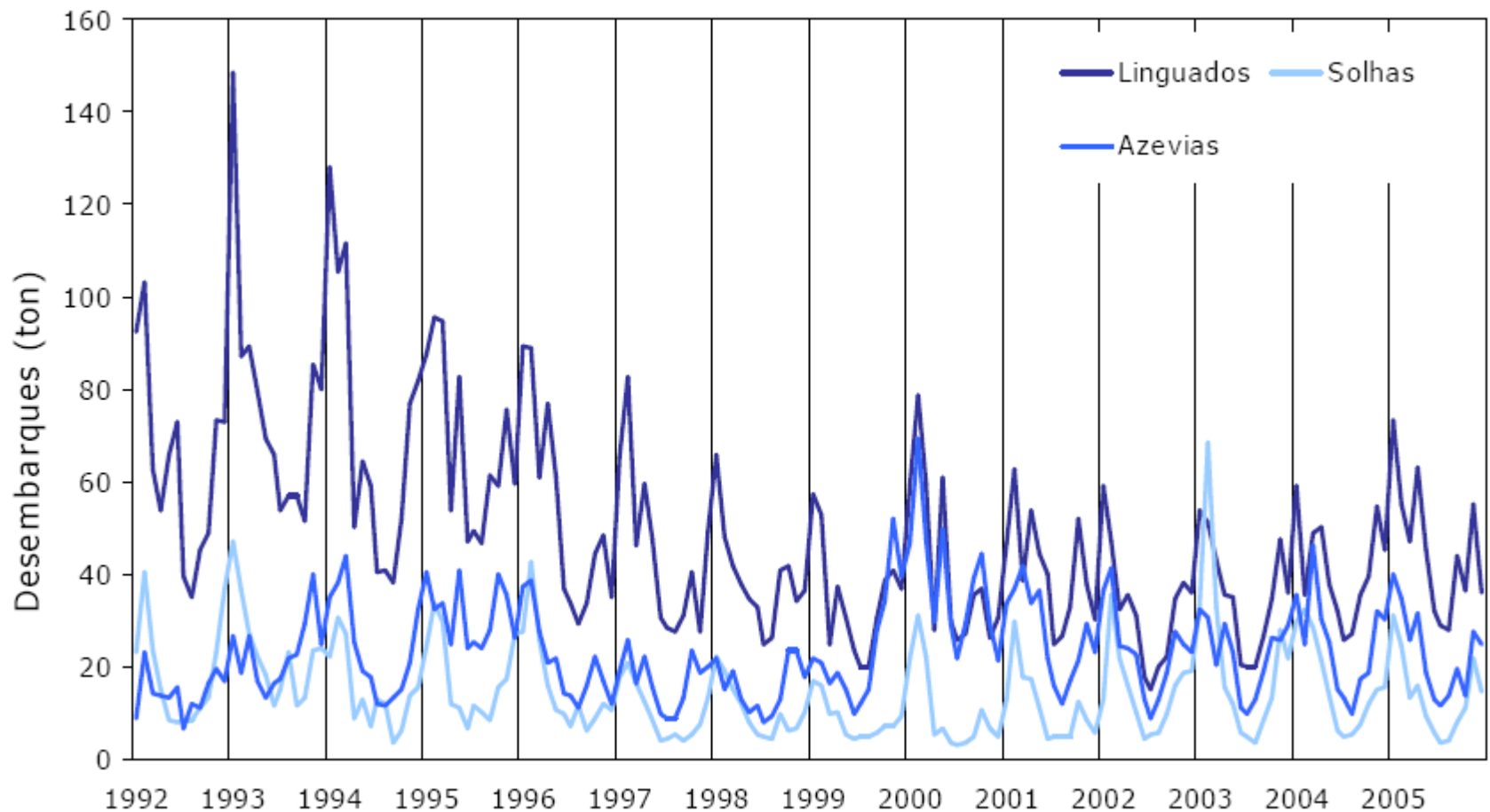


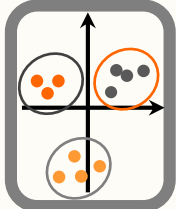




# análise exploratória de dados

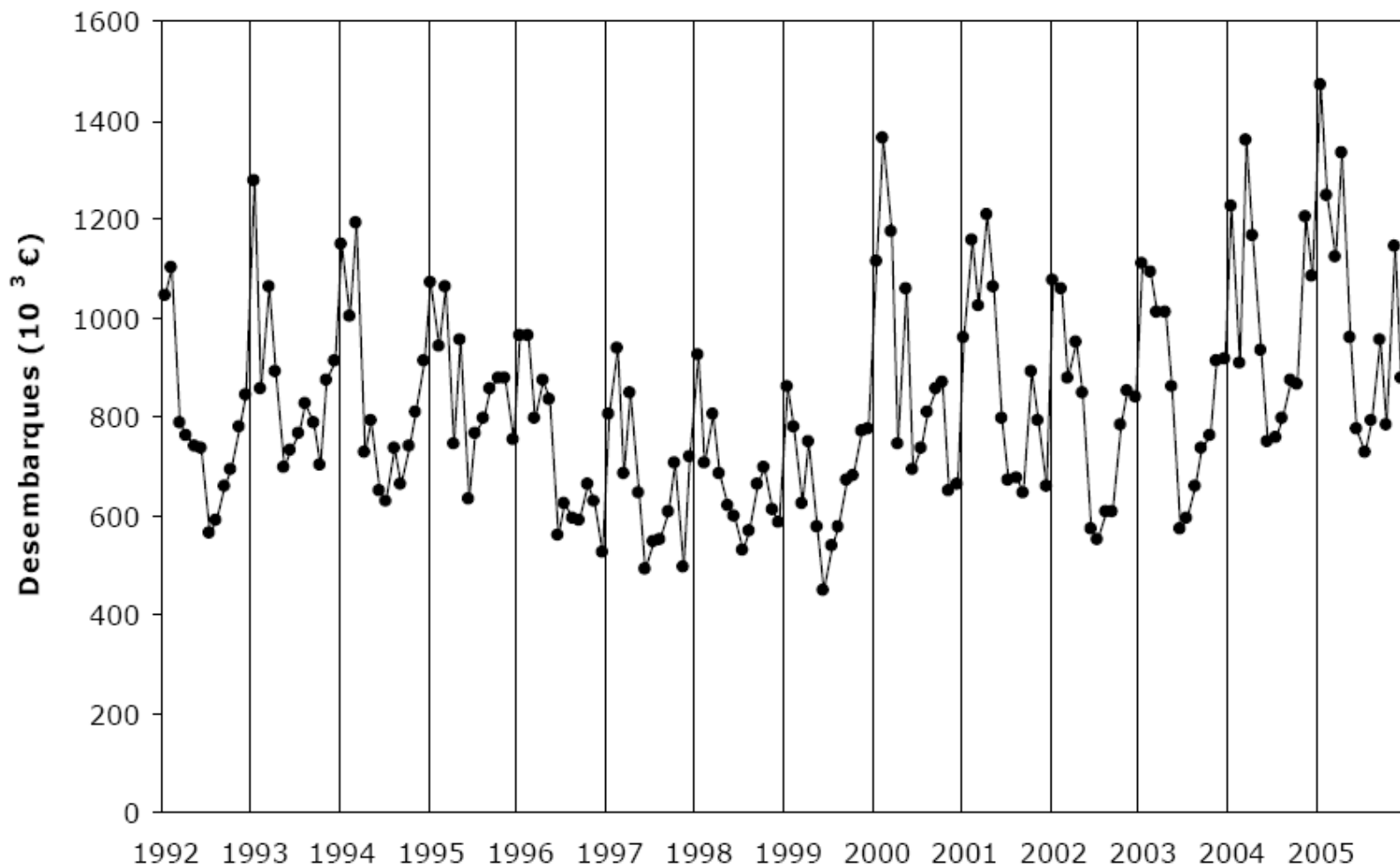
## Gráficos de linhas

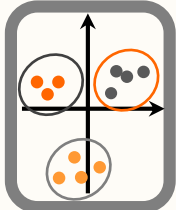




# análise exploratória de dados

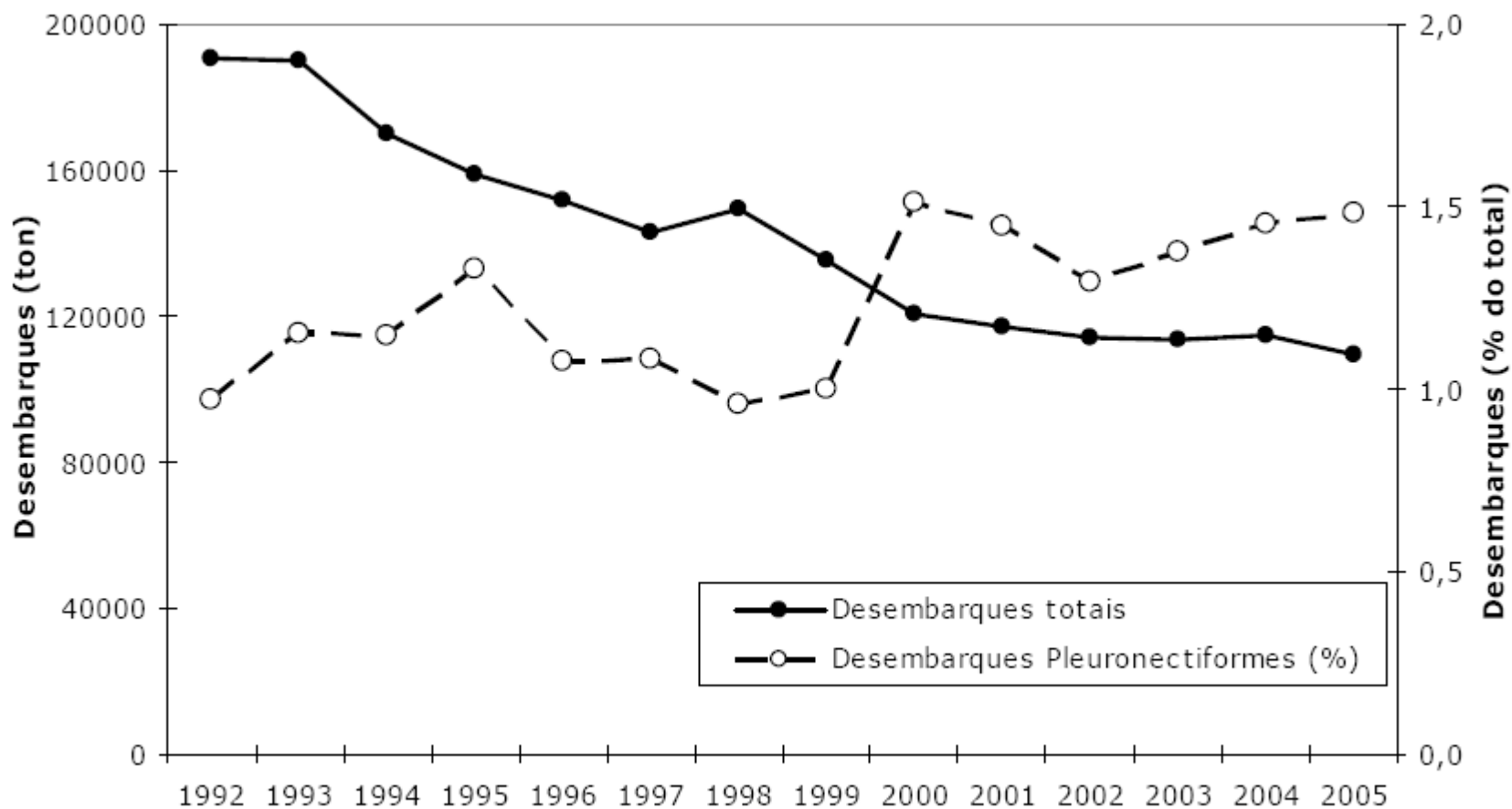
## Gráficos de linhas

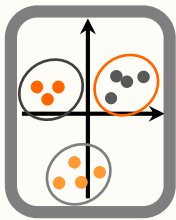




## análise exploratória de dados

# Gráficos de linhas

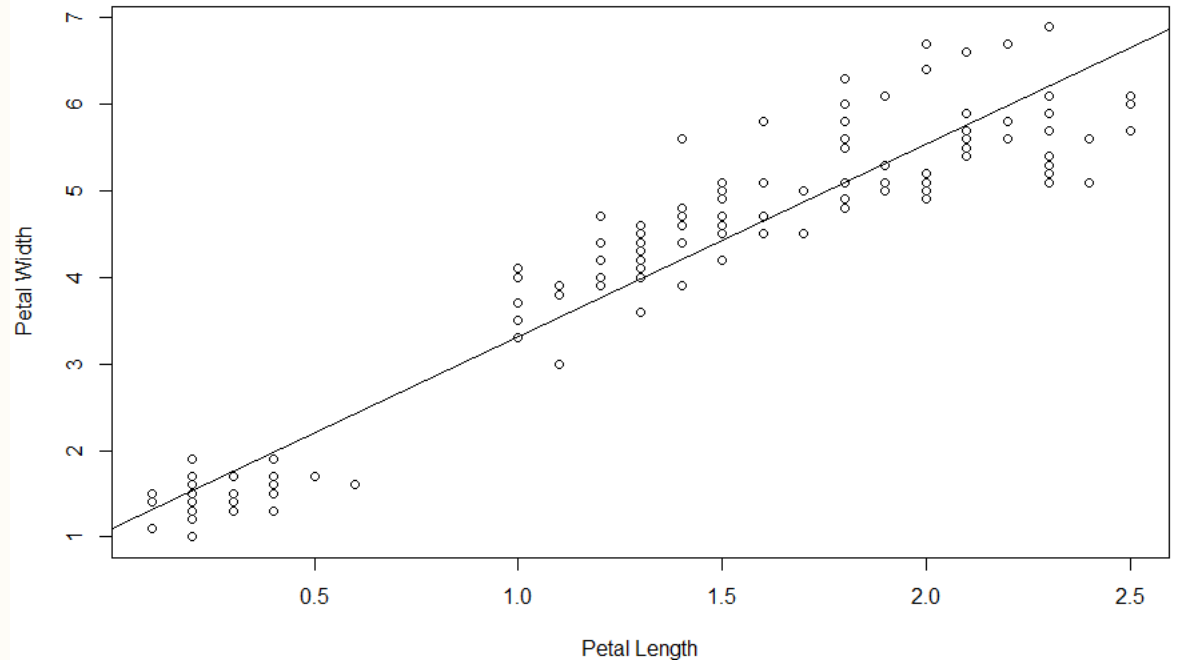


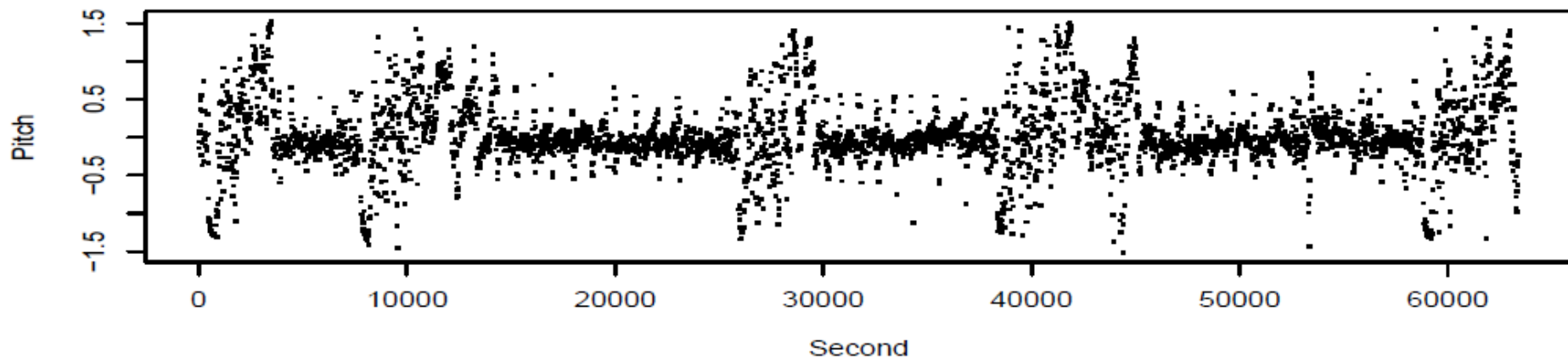
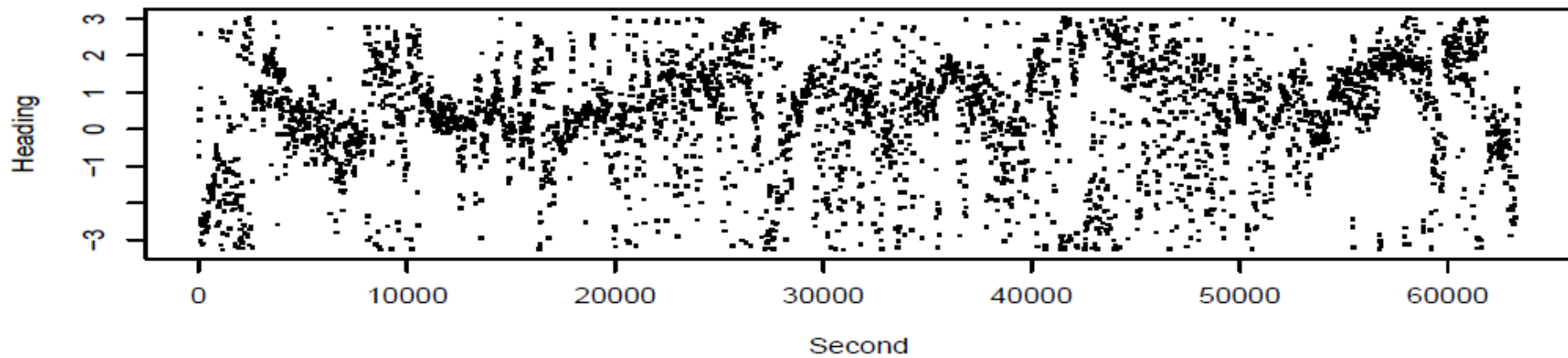
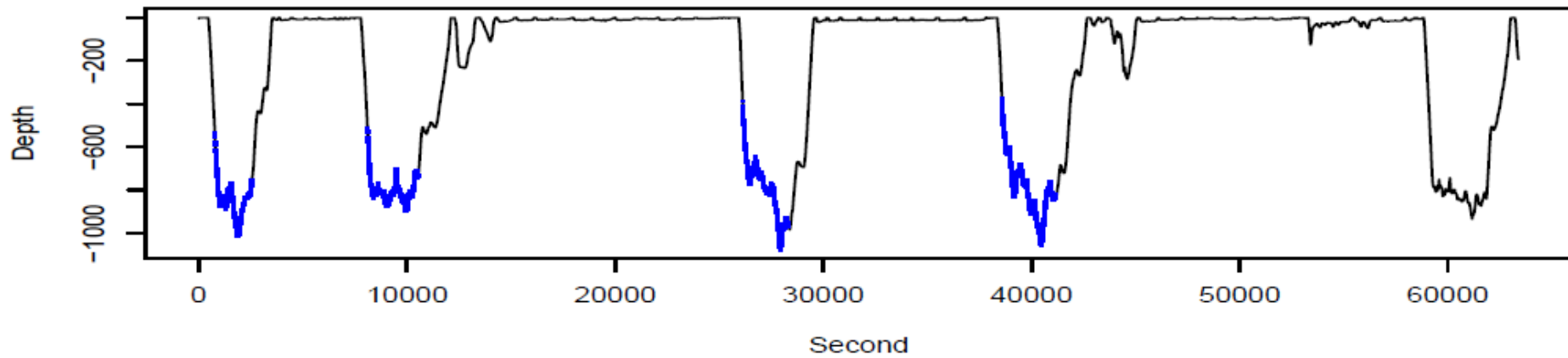


# análise exploratória de dados

## Gráficos de dispersão

```
... {r}  
data(iris)  
#plot the data |  
plot(iris[,4],iris[,3],xlab="Petal Length",ylab="Petal width")  
#add a linear model fit  
abline(lm(iris[,3]~iris[,4]))  
...
```





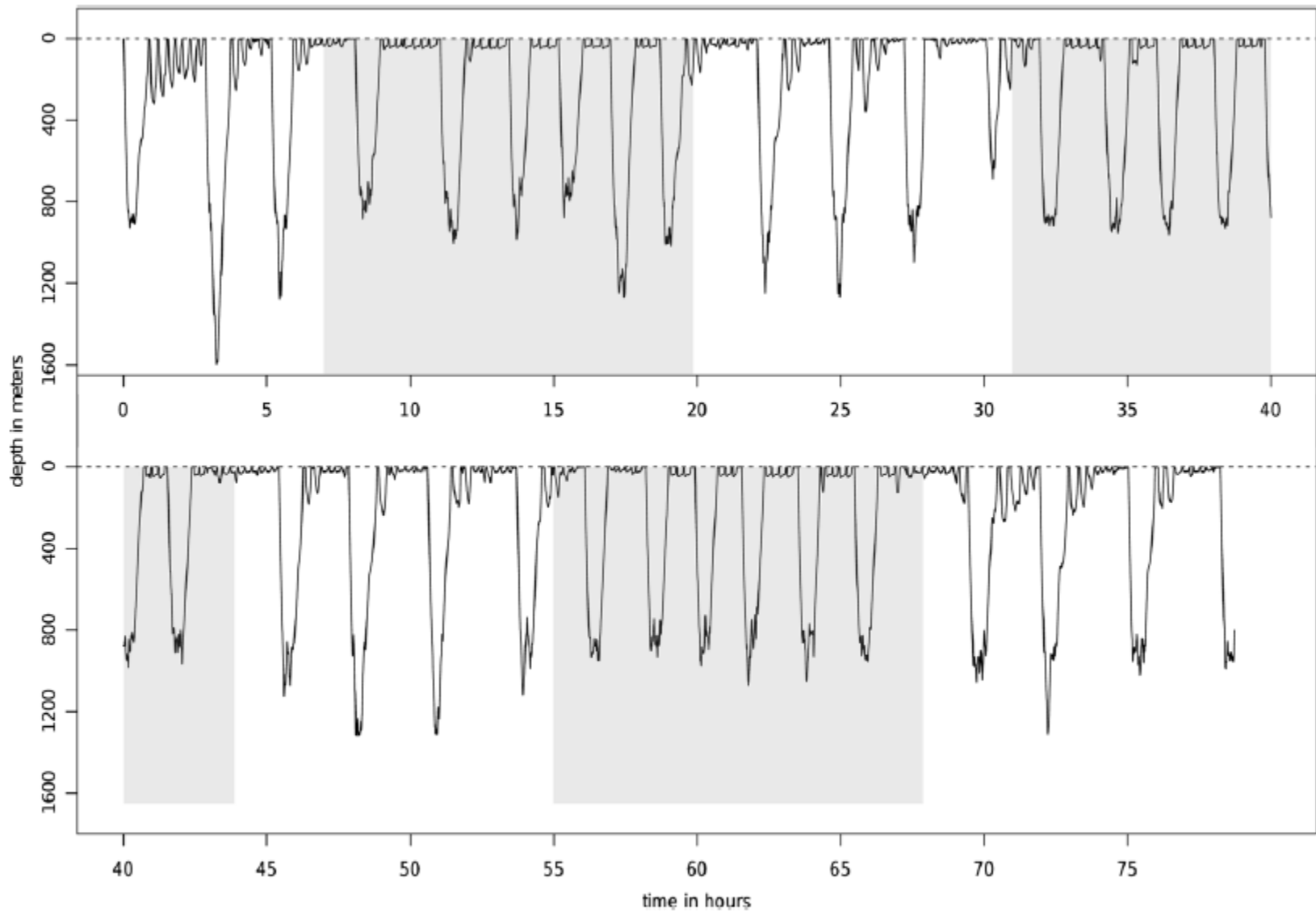
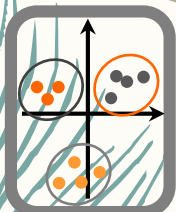
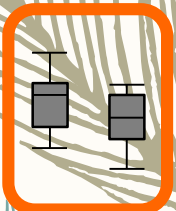
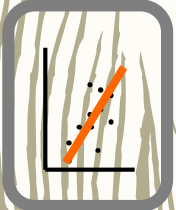
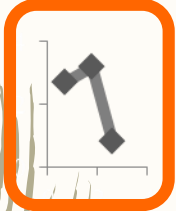
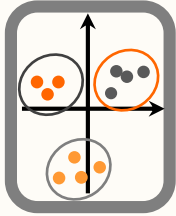


Figure 1. Beaked whale depth profile. The shaded periods correspond to “night” periods, between 17:44 h (sunset) and 06:37 h (sunrise).

# Ecologia Numérica



transformação  
de dados

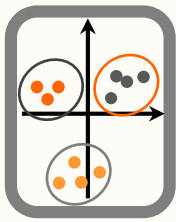


## transformação de dados

---

- Qual a necessidade de efectuar transformação de dados?
- Qual a legitimidade da transformação de dados?
- Serão as conclusões válidas quando obtidas com recurso a dados transformados?
- Que tipos de transformações são mais utilizadas e para que efeito?
- O que são pressupostos das análises estatísticas?
- Como avaliar os pressupostos?



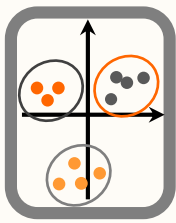


## transformação de dados

---

# Motivação para a transformação de dados

Para que se consiga atingir determinadas exigências que certas análises estatísticas têm (pressupostos).



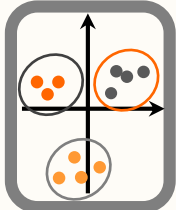
## transformação de dados

---

### **Será a transformação de dados legítima?**

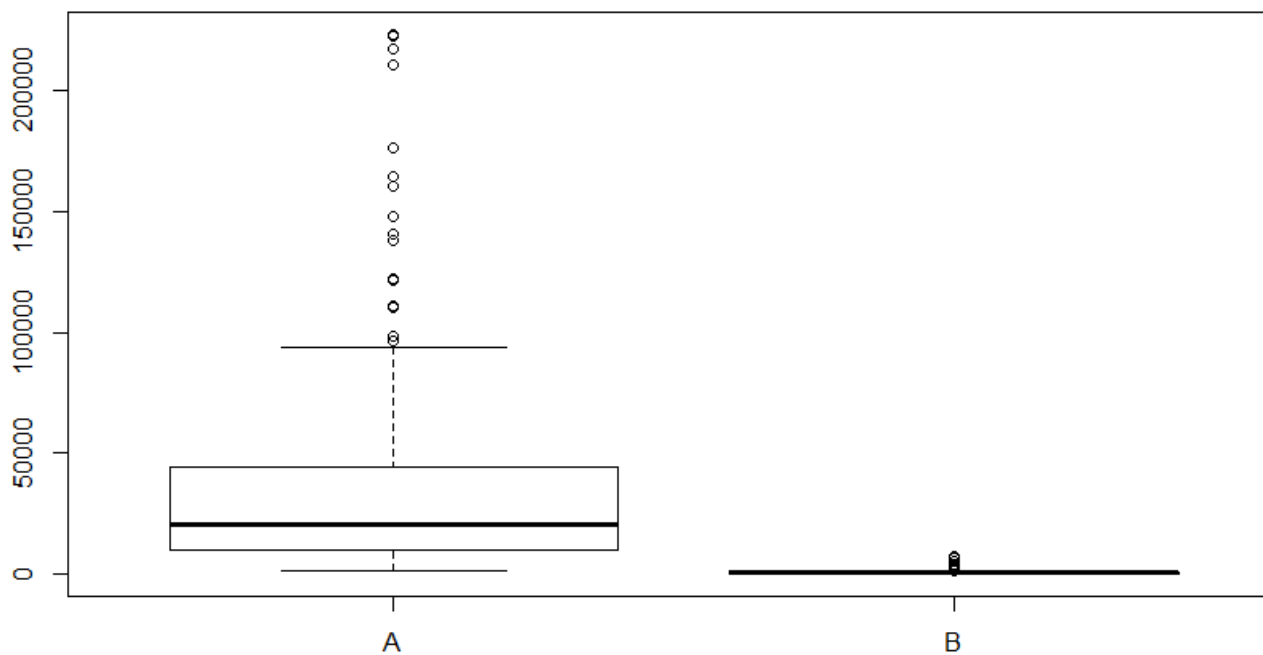
Na maioria das situações sim, mas é necessário ter algumas reservas dependendo do procedimento analítico.

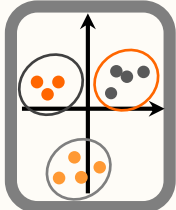
e.g. se eu quero dizer se A e B tem comprimentos diferentes, de que me interessa saber se o  $\log A$  é diferente do  $\log$  de B?



# transformação de dados

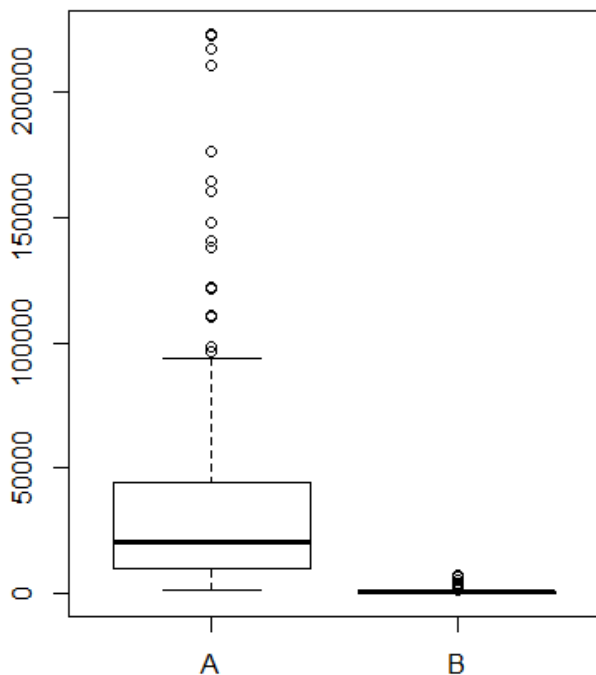
---



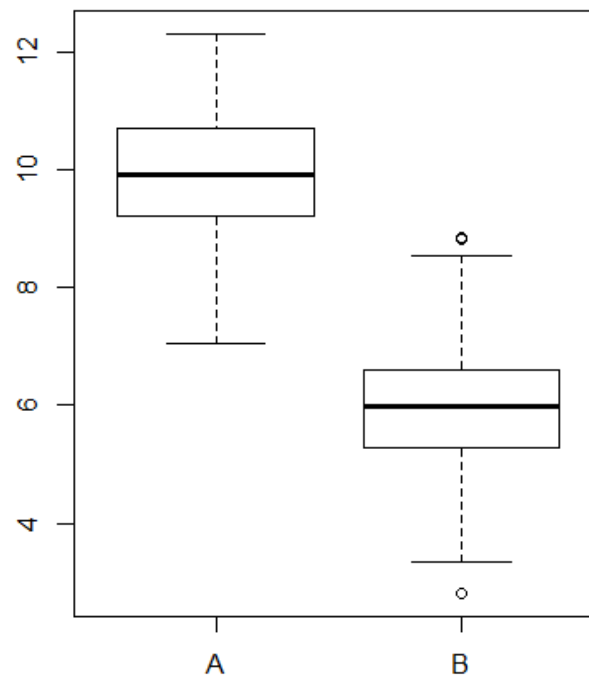


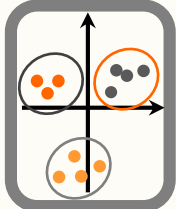
# transformação de dados

Dados Originais



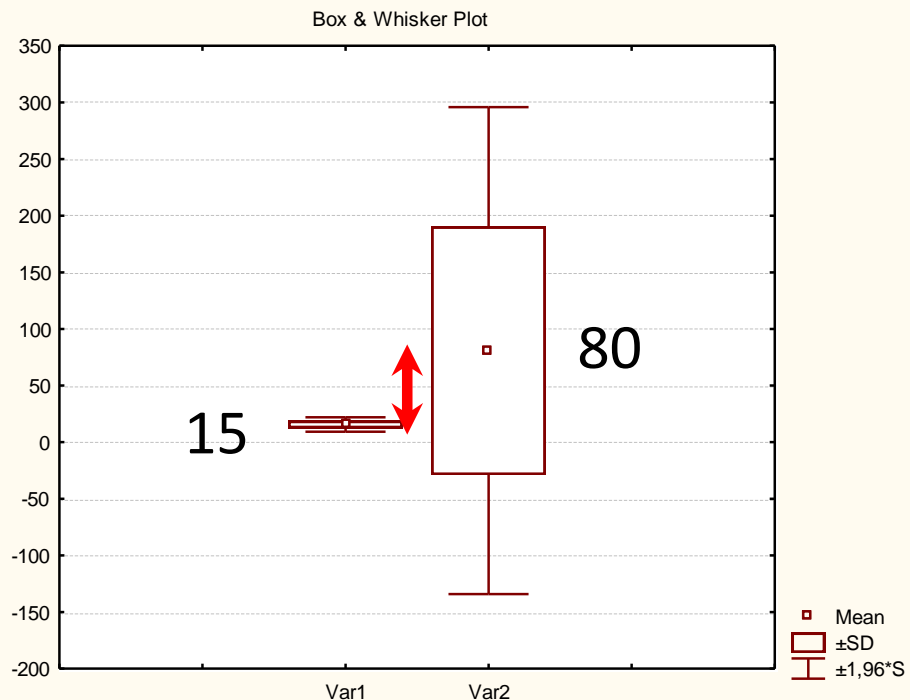
Log Dados Originais



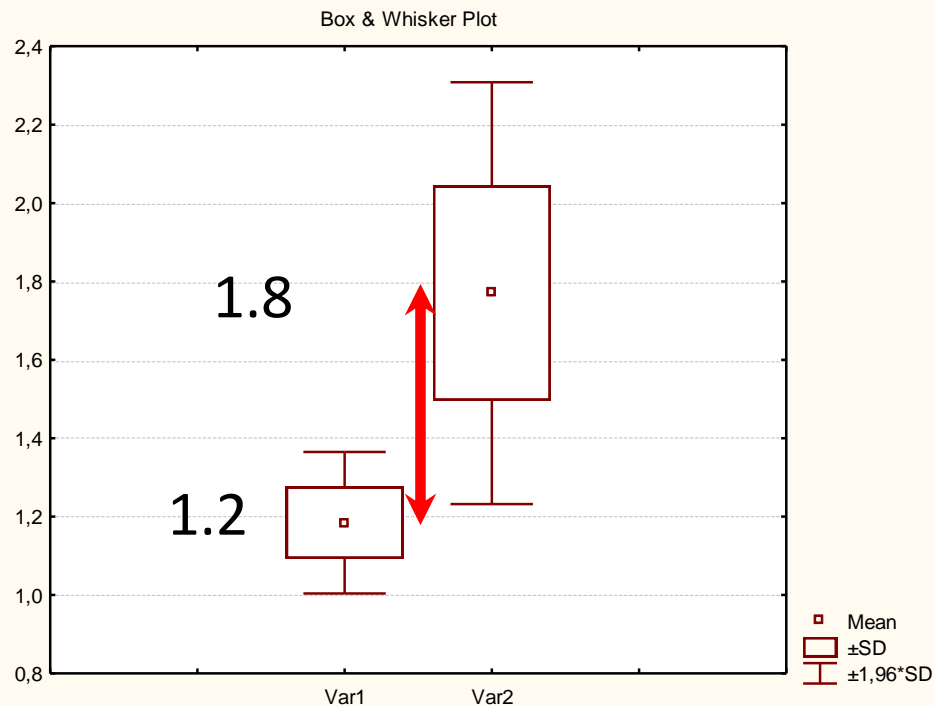


# transformação de dados

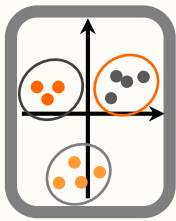
## Dados originais



## log



Uma imagem de SPSS...uma sobrevivente 😊



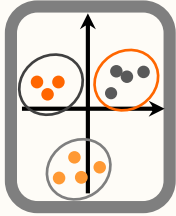
# transformação de dados

---

## **Tipos de transformação de dados mais frequentes**

Transformações lineares

Transformações não lineares

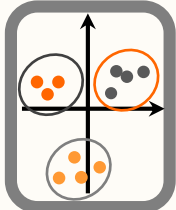


transformação de dados

---

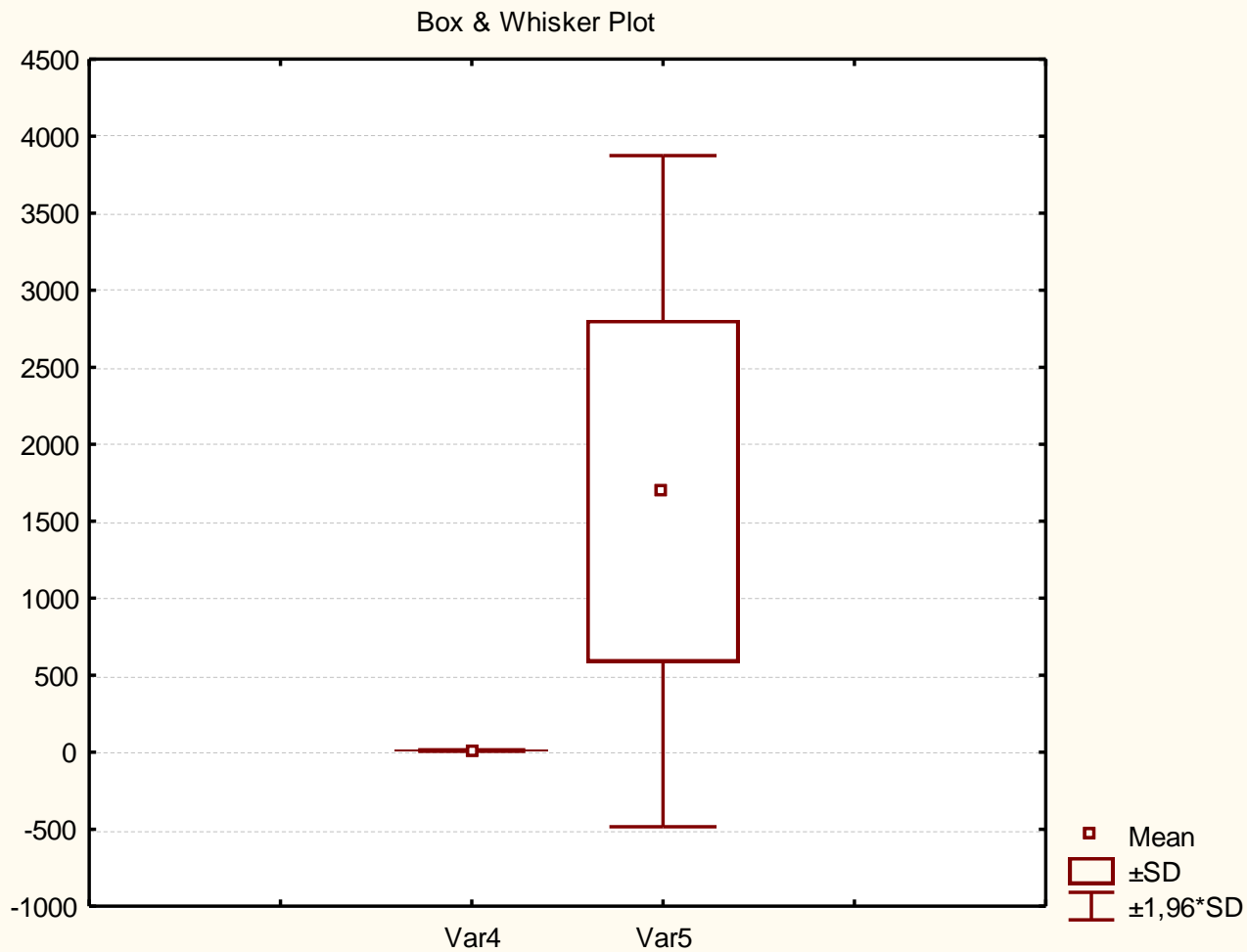
## Transformações lineares

Obtidas através de combinações lineares das variáveis originais (soma/subtração, produto/divisão).

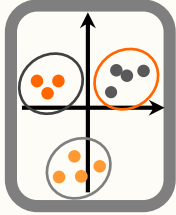


# transformação de dados

## Dados originais







transformação de dados

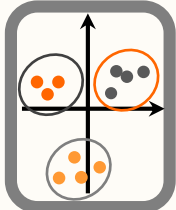
---

## Transformações lineares

Padronização pela variância (ou desvio-padrão) e pela média - Redução e centragem

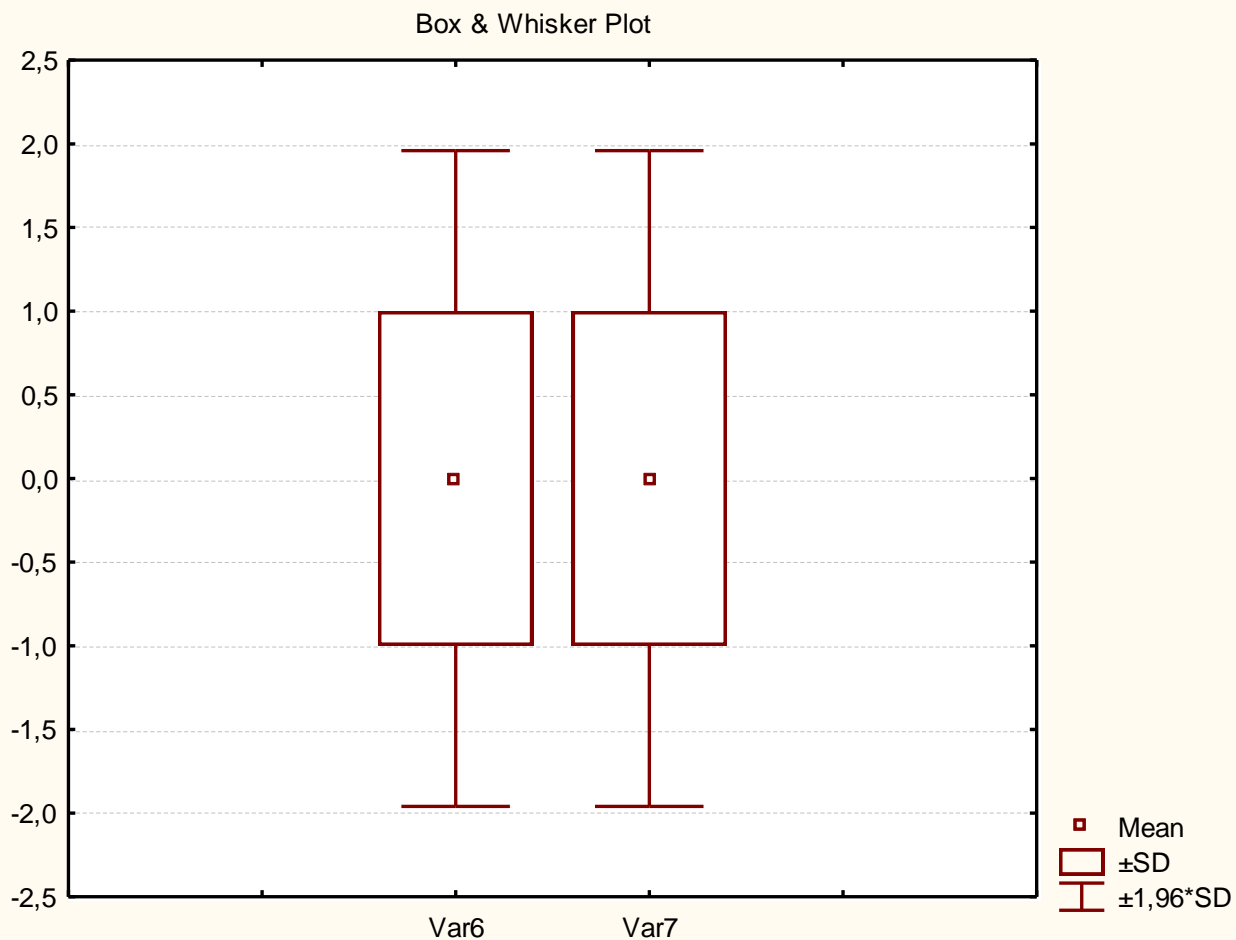
$$\frac{x_i - \bar{x}}{s}$$

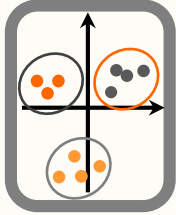
(mesma localização e dispersão: média=0; s=1)



# transformação de dados

## Dados padronizados: redução e centragem





transformação de dados

---

## Transformações lineares

Padronização por totais ou máximos

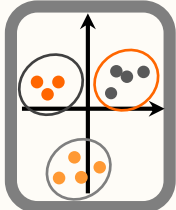
$$\frac{x_i}{T}$$

$$\frac{x_i}{max}$$

T=Soma de todos os  $x_i$

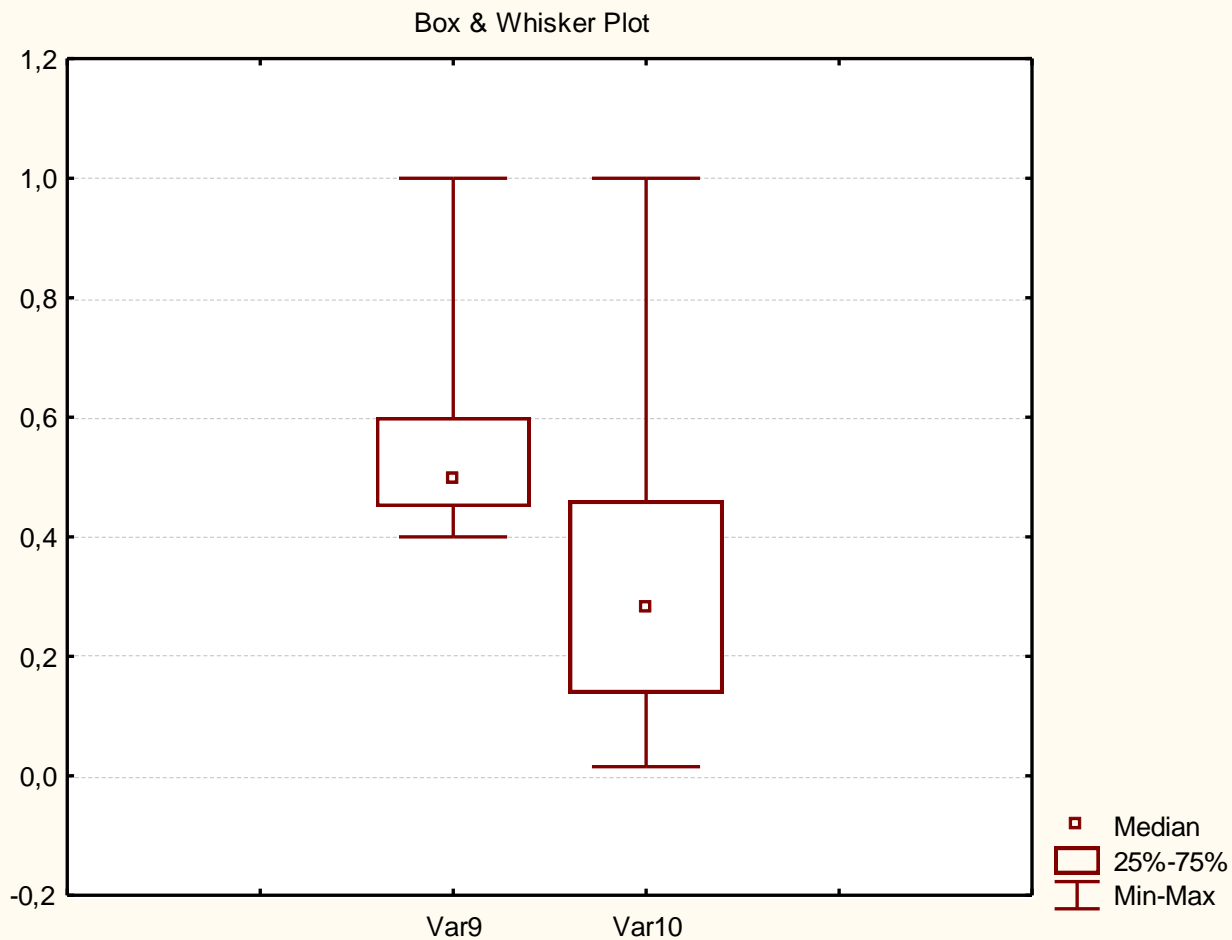
Max= máximos de todos os  $x_i$

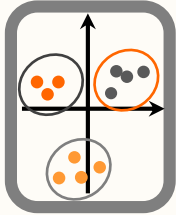
Todas as variáveis ficam com peso semelhante



# transformação de dados

## Dados padronizados: padronização pelo máximo





transformação de dados

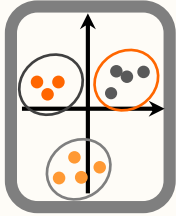
## Transformações não lineares

### Logaritmização

$$\log(x_i)$$

$$\log(x_i + 1)$$

Contribui para tornar as variâncias muito menores e desta forma mais facilmente obter homocedasticidade



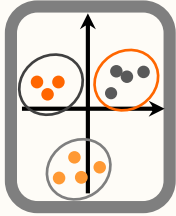
## Transformações não lineares

### Radiciação

$$\sqrt{x_i}$$

$$\sqrt{x_i + 0.5}$$

Contribui para tornar as variâncias muito menores e desta forma mais facilmente obter homocedasticidade



## Transformações não lineares

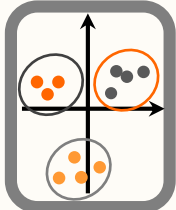
### Funções trigonométricas

Por exemplo:

$$\textit{arcsen}\sqrt{x_i}$$

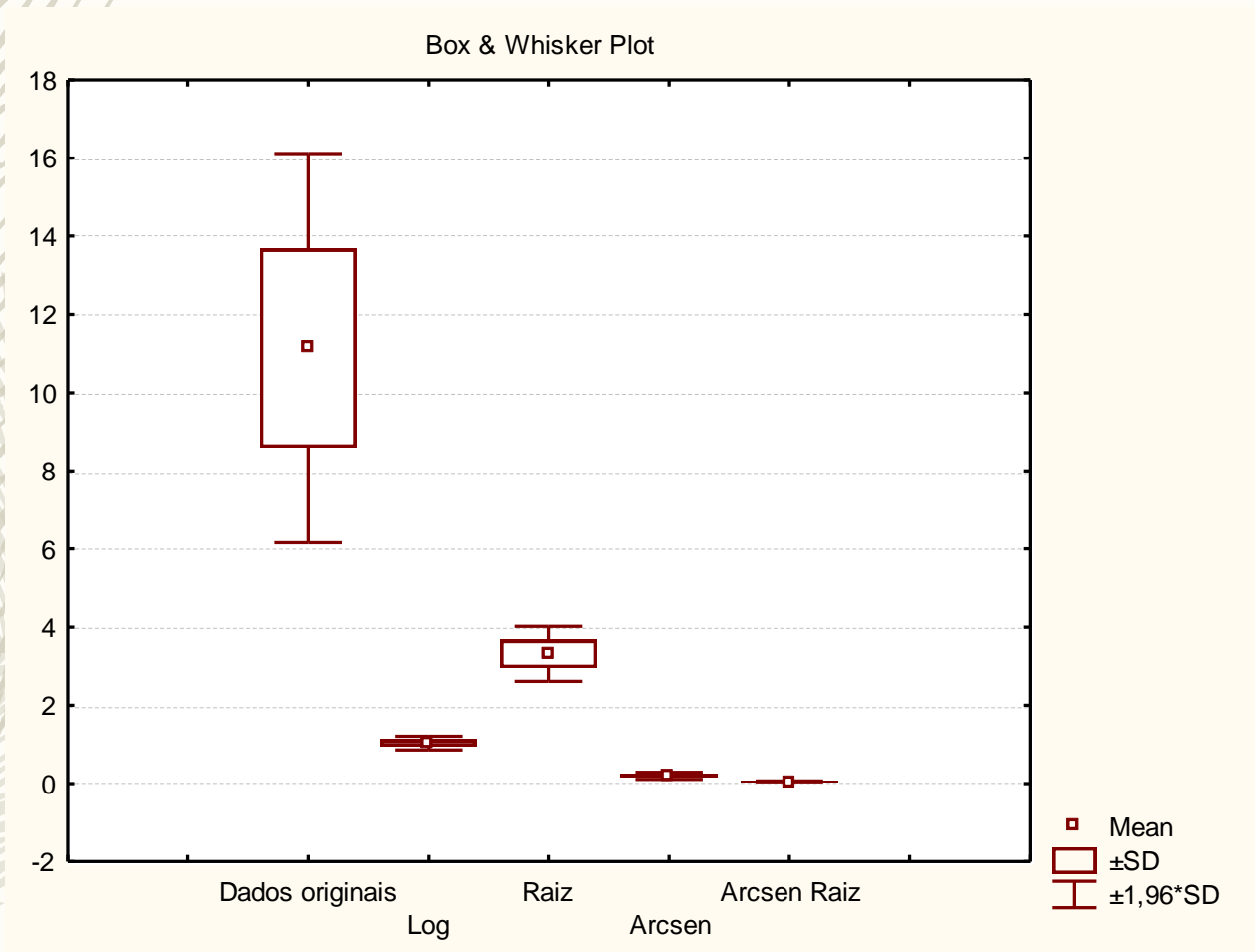
Contribuem para alterar as formas das distribuições.

Note-se que neste caso  $x_i$  deve estar entre 0 e 1, pelo que ou os dados são proporções ou temos antes que os transformar em proporções!

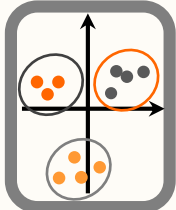


# transformação de dados

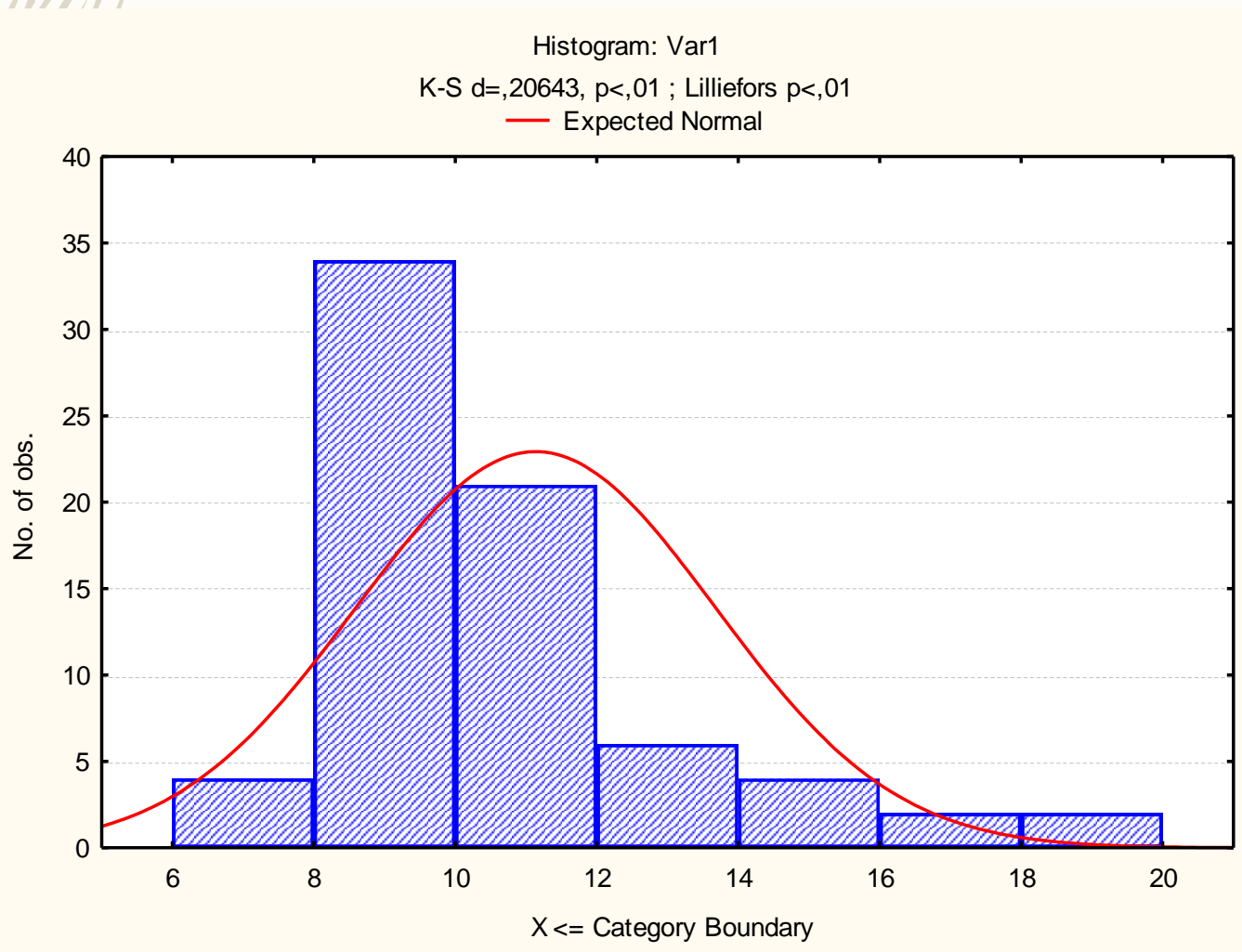
## Transformações não lineares

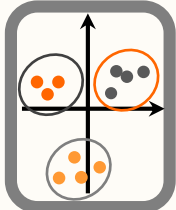






# transformação de dados

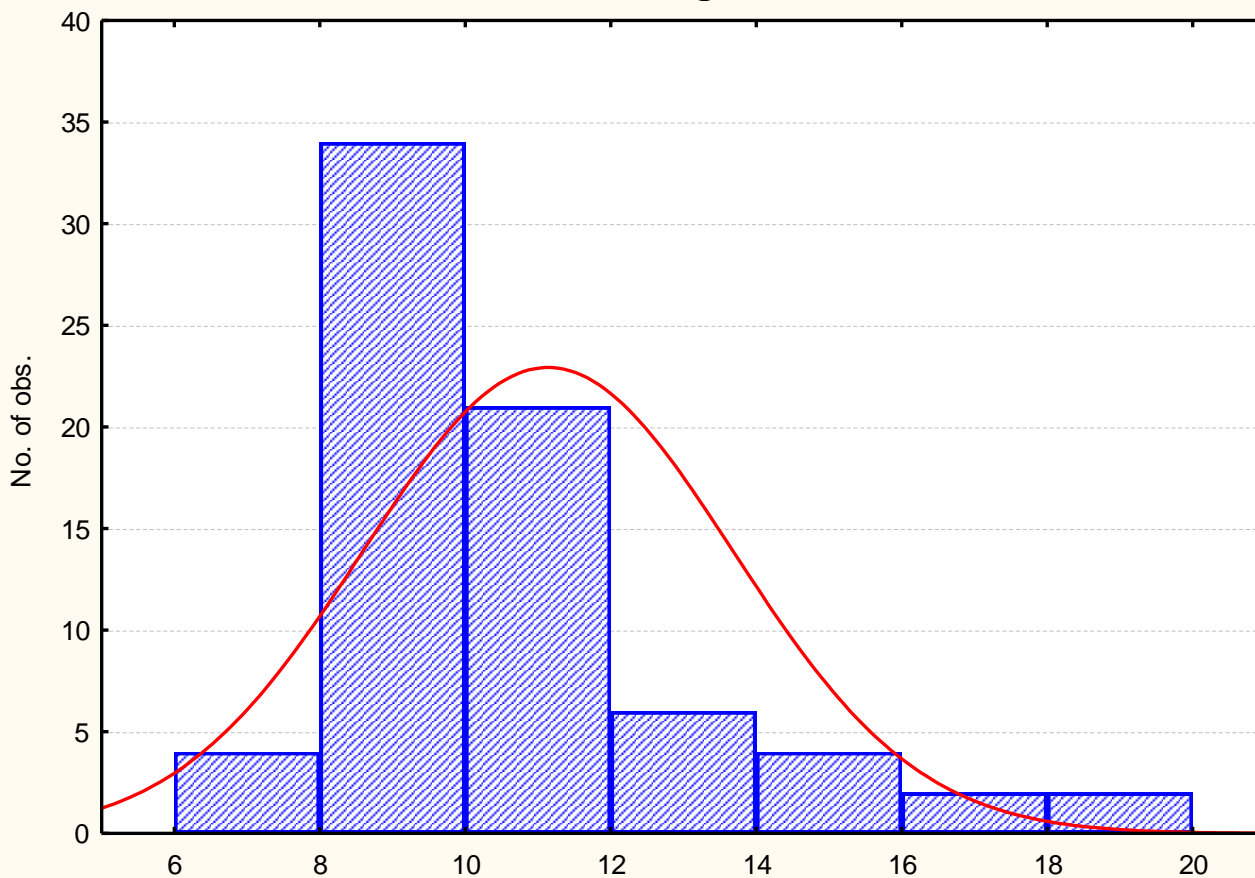


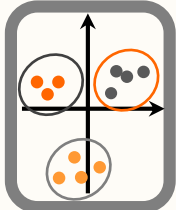


# transformação de dados

## Transformações não lineares

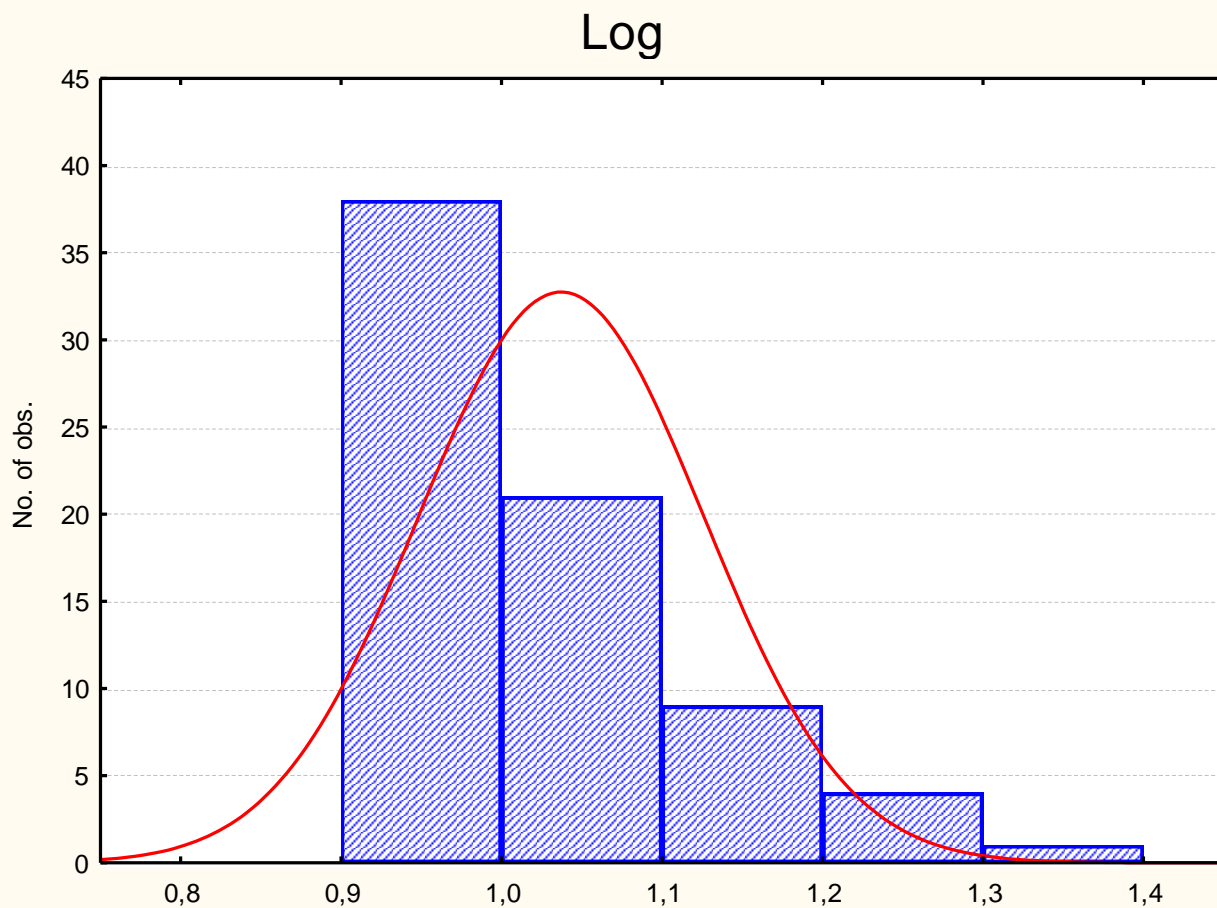
Dados originais

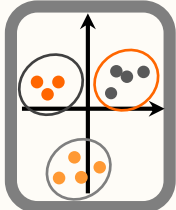




# transformação de dados

## Transformações não lineares



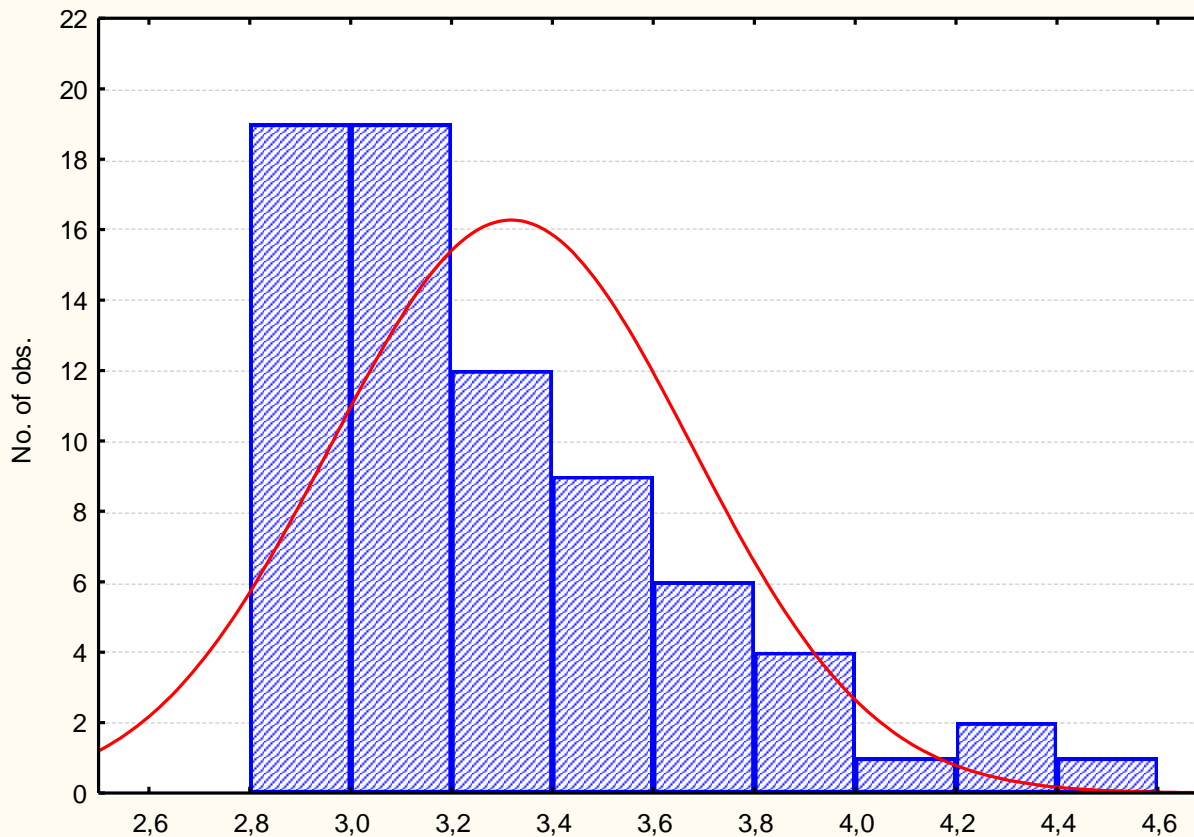


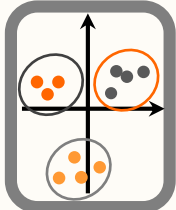
# transformação de dados

## Transformações não lineares



Raiz

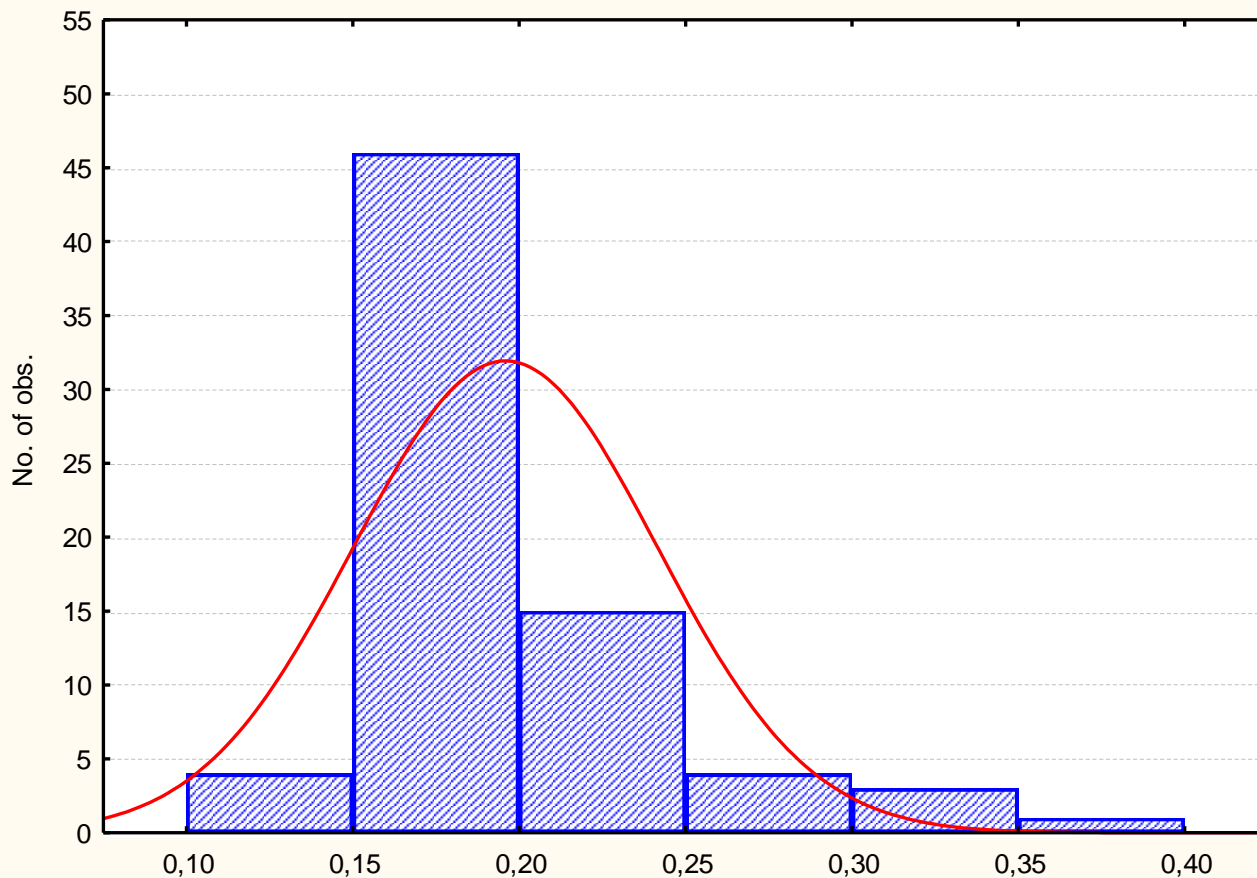


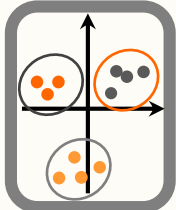


# transformação de dados

## Transformações não lineares

Arcsen

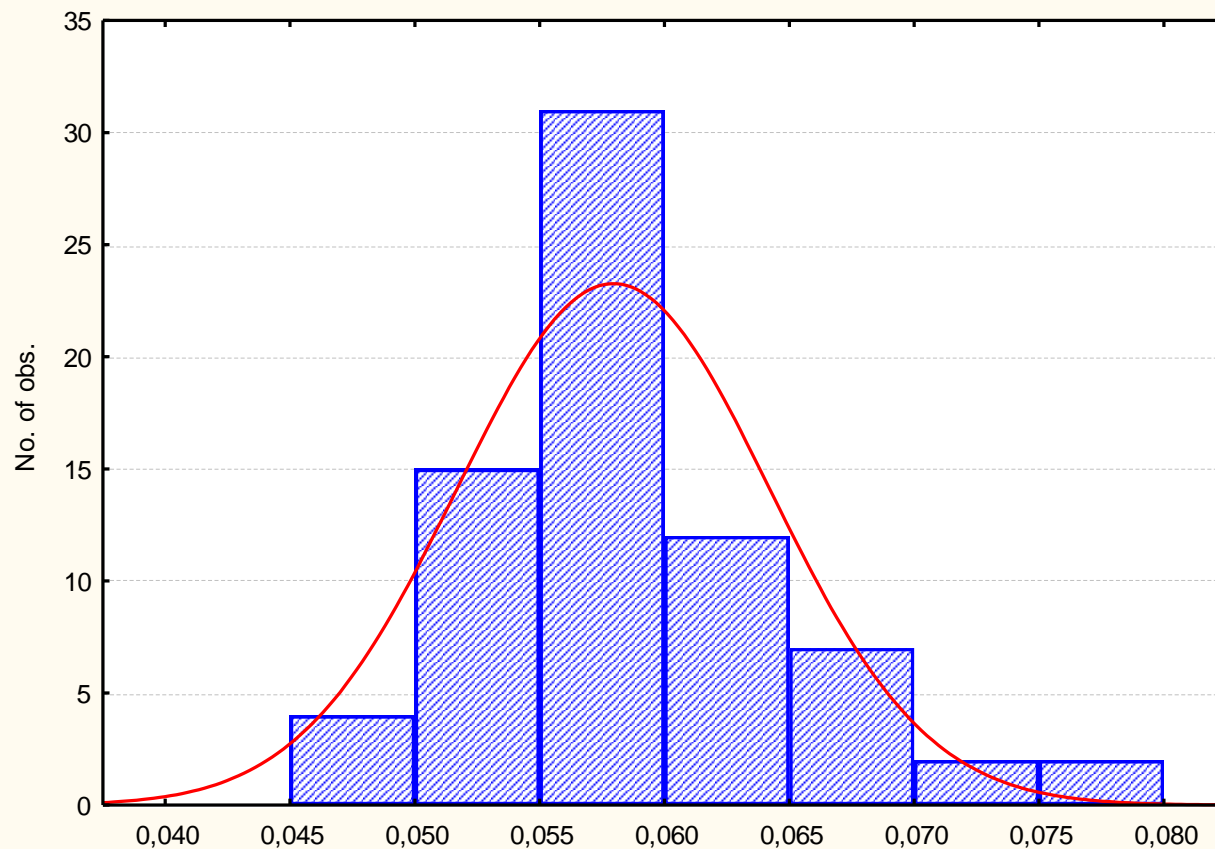


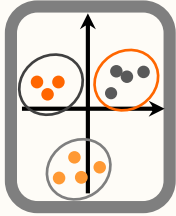


# transformação de dados

## Transformações não lineares

Arcsen Raiz





transformação de dados

---

## Transformações não lineares

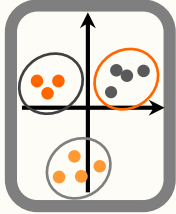
Funções trigonométricas

Por exemplo:

$$\textit{arcsen}\sqrt{x_i}$$

Contribuem para alterar as formas das distribuições.

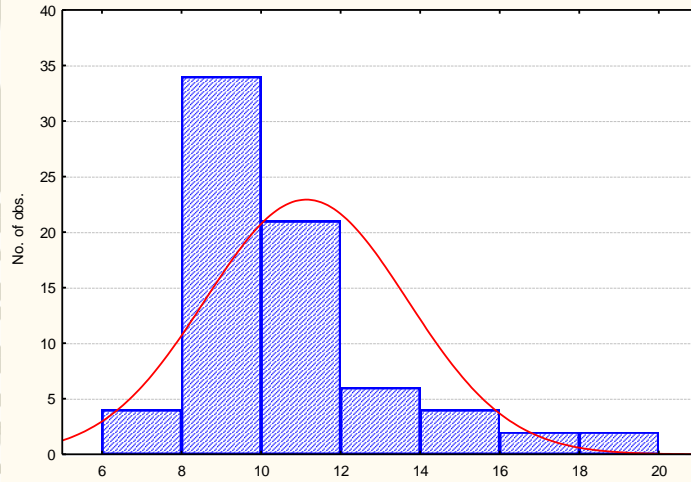
Note-se que o argumento do  $\textit{asin}$  deve estar entre 0 e 1, pelo que ou os dados são proporções ou temos antes que os transformar em proporções!



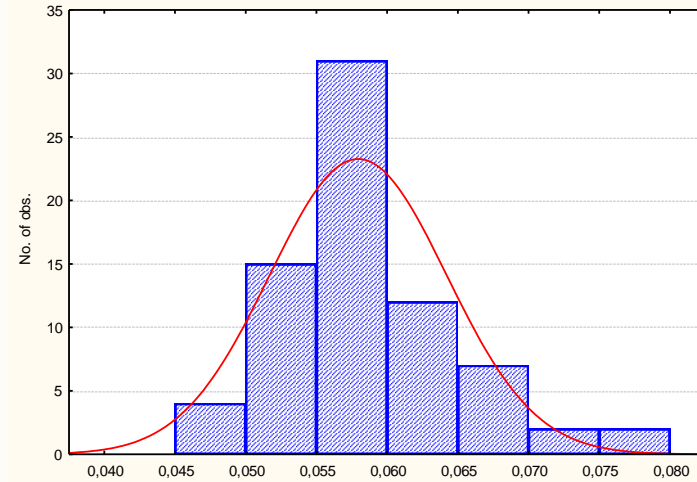
# transformação de dados

## Transformações não lineares

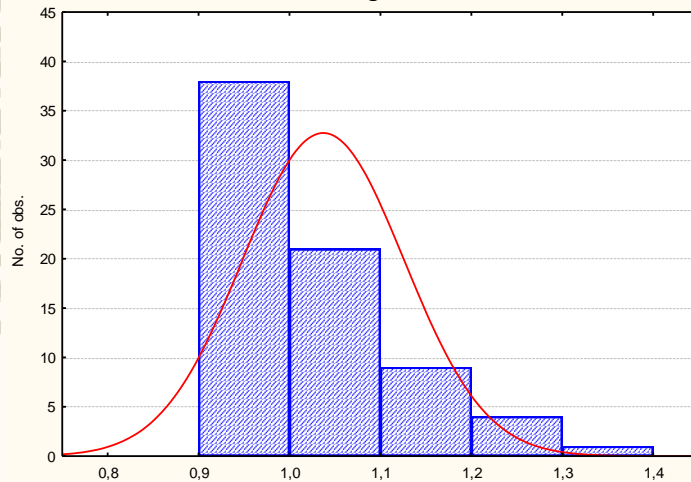
Dados originais



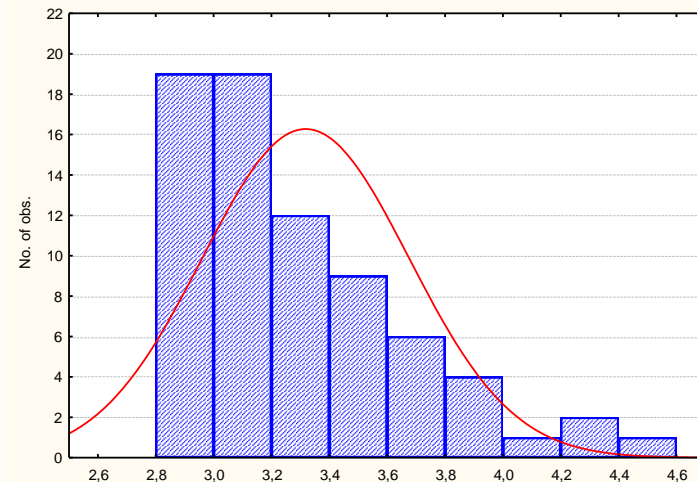
Arcsen Raiz



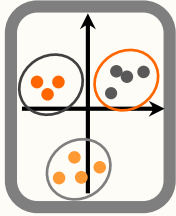
Log



Raiz

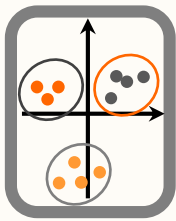






# Transformações dos dados

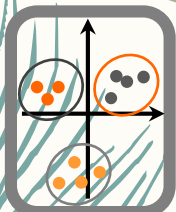
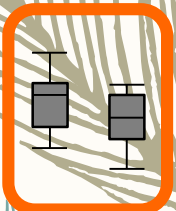
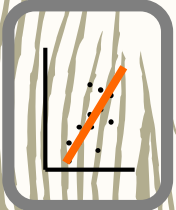
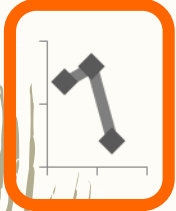
- Procedimento útil e necessário para que se consiga atingir o cumprimento de certos pressupostos distribucionais
- É mais fácil conter as variâncias do que alterar as formas das distribuições
- (mas, provavelmente, o melhor é mesmo encontrar metodologias que não exijam/requeiram os tais pressupostos!)



# Transformações dos dados

- É necessário uma análise cuidadosa para selecção da melhor e mais adequada transformação
- Cuidados e reservas na interpretação dos resultados!
- Muitas vezes é preferível seguir soluções alternativas
  - Métodos não paramétricos
  - Utilização de modelos que não tenham pressupostos tão estritos como Gaussianidade ou igualdade de variâncias!

# Ecologia Numérica



avaliação de  
pressupostos

—

# PRESSUPOSTOS

Exemplo de alguém que não percebeu os pressupostos envolvidos na utilização de um capacete branco na capa de um jornal

Como usar um capacete branco na capa de um jornal

**idealista** a maneira certa de encontrar casa

EDIÇÃO PORTO TER 15 SET 2015 **Maior produtora de queijo fresco aceita preço mínimo pelo leite** p19

**publico.pt** **Alunos portugueses são dos últimos da Europa a regressar às aulas**

Arranque oficial do ano lectivo acontece hoje, uma semana mais tarde do que o habitual, e em muitas escolas apenas na próxima semana haverá aulas. Cerca de 92 mil alunos entram no 1.º ano Portugal, 8 a 11

**LEGISLATIVAS 2015 PASSOS E PORTAS AGITAM FANTASMA DAS "ASNEIRAS" DO PS** Portugal, 6

**Decisão sobre quotas de refugiados adiada para 8 de Outubro**

Cimeira extraordinária de ontem termina em impasse. Controlo de fronteiras em três países p2 a 4

**ÁRTICO TIAGO MARQUES ANDOU A CONTAR URSOS POLARES** Ciência, 24/25

**Sanders e Carson baralham contas da campanha nos EUA**

Bernie Sanders, um "negro bem-educado", que combate a imagem negativa dos conservadores, e Ben Carson, um "socialista democrático" que revitaliza a esquerda, sobem nas sondagens p20

**Chuva e evento forte põem Norte do país em alerta vermelho**

Espera-se chuva forte e persistente, ventos de mais de 100 km/h e grande agitação marítima a partir de hoje. Metade do país, a norte do Mondego, em alerta laranja ou vermelho p13

**Crise em Angola não afecta turismo para Portugal**

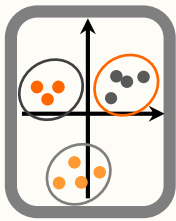
Receitas que entraram em Portugal subiram 2%, para 311,4 milhões de euros, no primeiro semestre do ano. Angolanos mantiveram o ritmo de compras, mas gastaram menos 4% p16

**CAMBRIDGE SCHOOL PORTUGAL**

Ano lectivo inscrições abertas

Educação: o seu melhor investimento.

www.cambridge.pt INGLÊS | FRANCÊS | ALEMÃO | PORTUGUÊS PARA ESTRANGEIROS

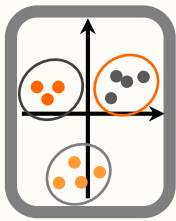


transformação de dados  
avaliação de pressupostos

---

## Avaliação de pressupostos

- A generalidade das inferências estatísticas tem por base testes (estatísticos) que tem pressupostos
- Os pressupostos são as condições básicas que deverão ser verificadas para se poder aplicar um determinado procedimento
- Existem diferentes formas de avaliar os pressupostos

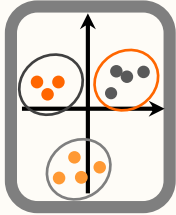


# Avaliação de pressupostos dos testes estatísticos

- Testes formais sobre os pressupostos
- Análises gráficas de dados / análise de resíduos (de um modelo)

# Os dois (ou três!) pressupostos mais comuns:

0. independência das observações (não se pode testar, pelo que ninguém fala disso!) – mas é possivelmente o pressuposto mais importante!
1. Resíduos (ou erros) Gaussianos (e.g. modelo linear, teste-t, ANOVA, etc)
2. Homogeneidade de variâncias = Igualdade de variâncias = Homocedasticidade (e.g. incluindo teste-t, ANOVA, etc)



## Normalidade

- Testes de normalidade
  - Teste de Kolmogorov-Smirnov  
*Estatística D*
  - Teste D'Agostino-Pearson  
*Estatística  $K^2$*
  - Teste de Shapiro-Wilk  
*Estatística W*
  - Outros: Anderson-Darling, Cramér–von Mises, etc
- Avaliação gráfica
- Análise de resíduos



A base de todos estes testes de ajustamento (*goodness-of-fit, gof*) é sempre a mesma:

Se os dados proviessem, de facto, de uma população Gaussiana\*, teriam um conjunto de características conhecidas.

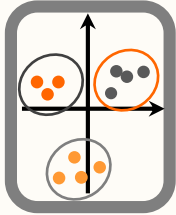
(que depende de teste para teste, daí eles não terem necessariamente de dar o mesmo resultado)

H0: a distribuição subjacente é Gaussiana\*

H1: a negação de H0 (a distribuição subjacente não é Gaussiana)

Observada a nossa amostra, é possível que os desvios em relação ao que seria expectável sejam tão grandes que a população parente dificilmente tenha sido a Gaussiana\*?

\*quem aqui diz Gaussiana diz outra qualquer para testes de ajustamento em geral!



# Testes de normalidade

## Teste de Kolmogorov-Smirnov

Estatística de teste:

$$D = \max_{1 \leq i \leq N} \left| F(Y_i) - \frac{i}{N} \right|$$

onde  $F$  é a distribuição teórica cumulativa da distribuição que está a ser testada.

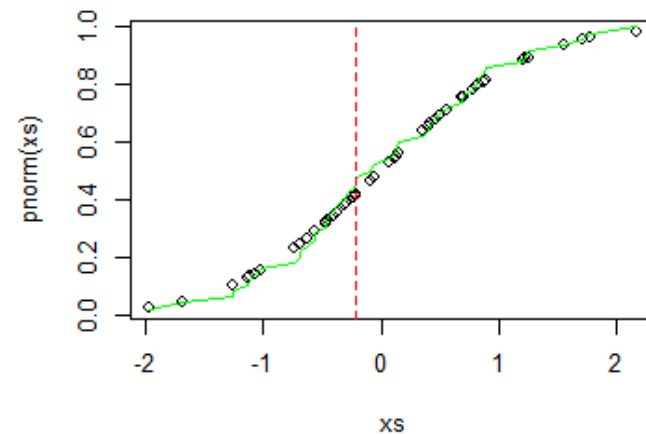
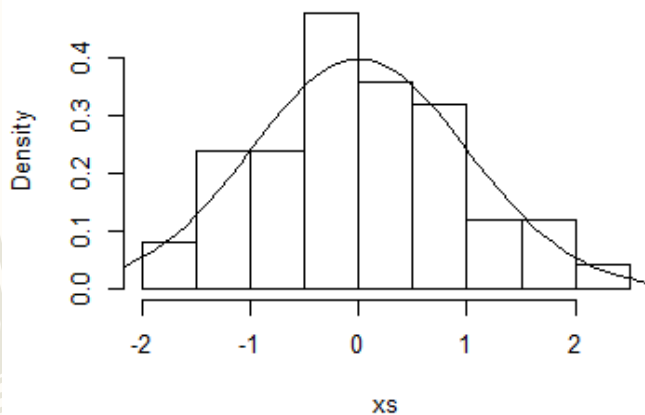
```
ks.test(rnorm(50), "pnorm")
```

```

set.seed(123)
#create a dataset for illustration; note sample size = 50 and
#the underlying distribution is truly Gaussian
n=50
xs=rnorm(n)
KStestEN(xs)

```

Adicionei o código desta função no FENIX, Outros Recursos



- Ecologia Numérica
  - Ecologia Numérica(Tecnologias de Informação)
  - Teóricas
  - Práticas
    - Week1
    - Week 2
    - Week 3
    - Week 4
  - PDFs
  - Outros Recursos
  - R Cheat Sheets

+ Criar

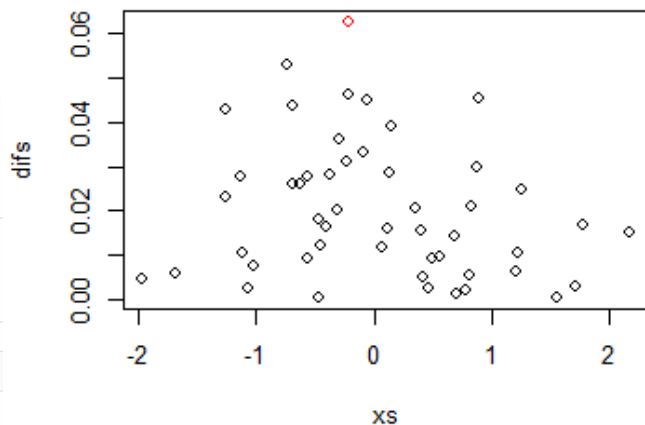
### Outros Recursos

Página Ficheiros 2 Permissões Link

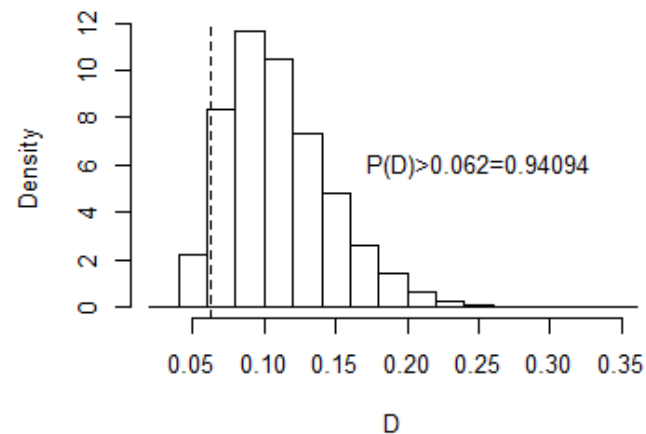
Adicionar Ficheiro

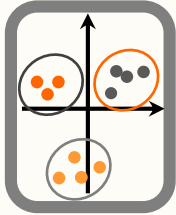
#	Nome
1	Elison2004.pdf
2	Código para fazer teste de KS - aula teórica 8 TA10xKStest.R

**D = 0.062**



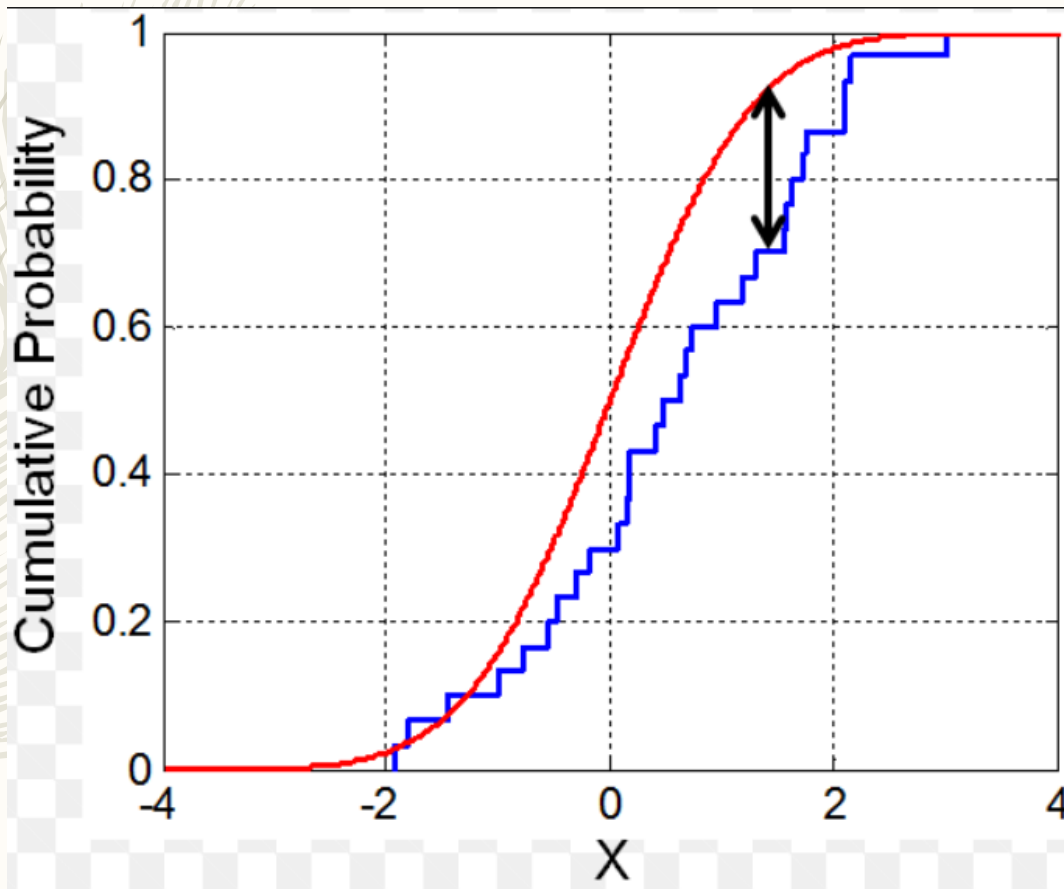
**Distribuição de D, sob H0**





# Testes de normalidade

## Teste de Kolmogorov-Smirnov



```
set.seed(123)
```

```
ks.test(rnorm(50), "pnorm")
```

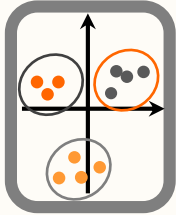
```
One-sample Kolmogorov-Smirnov test data: rnorm(50) D  
= 0.073034, p-value = 0.9347 alternative hypothesis:  
two-sided
```

```
ks.test(runif(50), "pnorm")
```

```
One-sample Kolmogorov-Smirnov test data: runif(50) D  
= 0.50421, p-value = 2.71e-12 alternative  
hypothesis: two-sided
```

```
ks.test(rgamma(50, 5), "pnorm")
```

```
One-sample Kolmogorov-Smirnov test data: rgamma(50,  
5) D = 0.99644, p-value = 8.882e-16 alternative  
hypothesis: two-sided
```



# Testes de normalidade

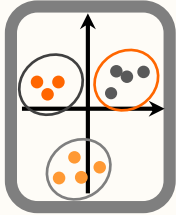
## Teste de Shapiro-Wilk

Estatística de teste:

```
shapiro.test(xs)
```

$$W = \frac{\left( \sum_{i=1}^n \alpha_i x_{(i)} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

onde  $x_{(i)}$  são os valores da amostra ordenados e  $\alpha_{(i)}$  são constantes obtidas a partir das médias, variâncias e covariâncias das estatísticas ordinais duma amostra de dimensão  $n$  duma população normal.



## Testes de normalidade

### Teste de D'Agostino-Pearson

**EXAMPLE 6.8** The D'Agostino-Pearson  $K^2$  test for normality applied to the data of Example 6.1.

$H_0$ : The sampled population is normally distributed.

$H_A$ : The sampled population is not normally distributed.

From Example 7.12:  $Z_{g_1} = -1.2291$

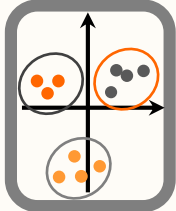
From Example 7.13:  $Z_{g_2} = 1.2763$

Therefore,

$$\begin{aligned} K^2 &= Z_{g_1}^2 + Z_{g_2}^2 = (-1.2291)^2 + (1.2763)^2 \\ &= 1.5107 + 1.6289 = 3.1396 \end{aligned}$$

Considering  $K^2$  to be from a  $\chi^2$  distribution with  $\nu = 2$ ,  $0.10 < P < 0.25$  [ $P = 0.21$ ]

Therefore, do not reject  $H_0$ .



## Testes de normalidade

### Teste de D'Agostino-Pearson

onde  $Z_{g_1}$  e  $Z_{g_2}$  são obtidos do seguinte modo

$$A = \sqrt{b_1} \sqrt{\frac{(n+1)(n+3)}{6(n-2)}}$$

$$B = \frac{3(n^2 + 27n - 70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)}$$

$$C = \sqrt{2(B-1)} - 1$$

$$D = \sqrt{C}$$

$$E = \frac{1}{\sqrt{\ln D}}$$

$$F = \frac{A}{\sqrt{\frac{2}{C-1}}}$$

$$Z_{g_1} = E \ln \left( F + \sqrt{F^2 + 1} \right)$$

Simetria

$$G = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

$$H = \frac{(n-2)(n-3)|g_2|}{(n+1)(n-1)\sqrt{G}}$$

$$J = \frac{6(n^2 - 5n + 2)}{(n+7)(n+9)} \sqrt{\frac{6(n+3)(n+5)}{n(n-2)(n-3)}}$$

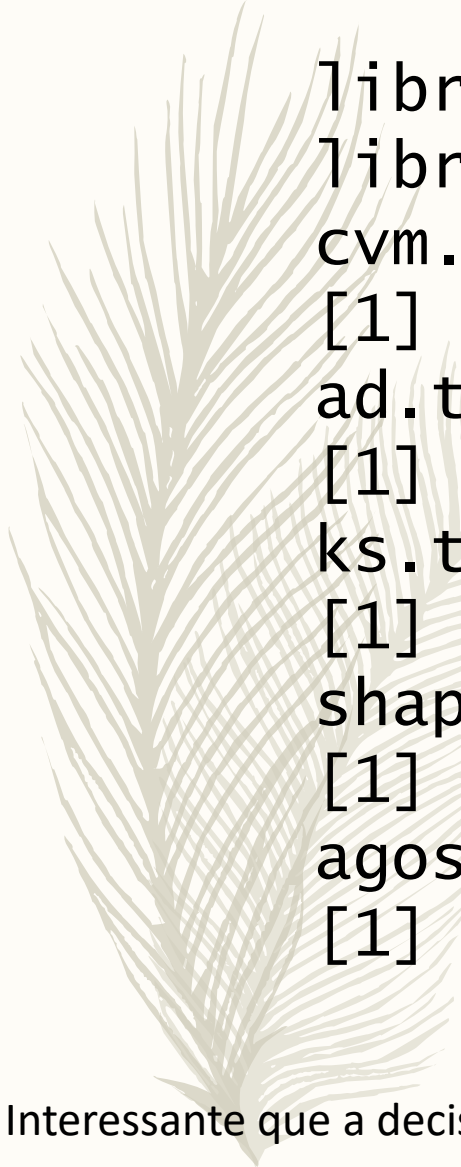
$$K = 6 + \frac{8}{J} \left[ \frac{2}{J} + \sqrt{1 + \frac{4}{J^2}} \right]$$

$$L = \frac{1 - \frac{2}{K}}{1 + H \sqrt{\frac{2}{K-4}}}$$

$$Z_{g_2} = \frac{1 - \frac{2}{9K} - \sqrt[3]{L}}{\sqrt{\frac{2}{9K}}}$$

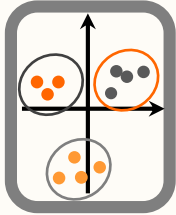
Curtose





```
library(moments)
library(goftests)
cvm.test(xs, "pnorm")$p.value
[1] 0.9627027
ad.test(xs, "pnorm")$p.value
[1] 0.9719898
ks.test(xs, "pnorm")$p.value
[1] 0.9346721
shapiro.test(xs)$p.value
[1] 0.9278568
agostino.test(xs)$p.value
[1] 0.5910889
```

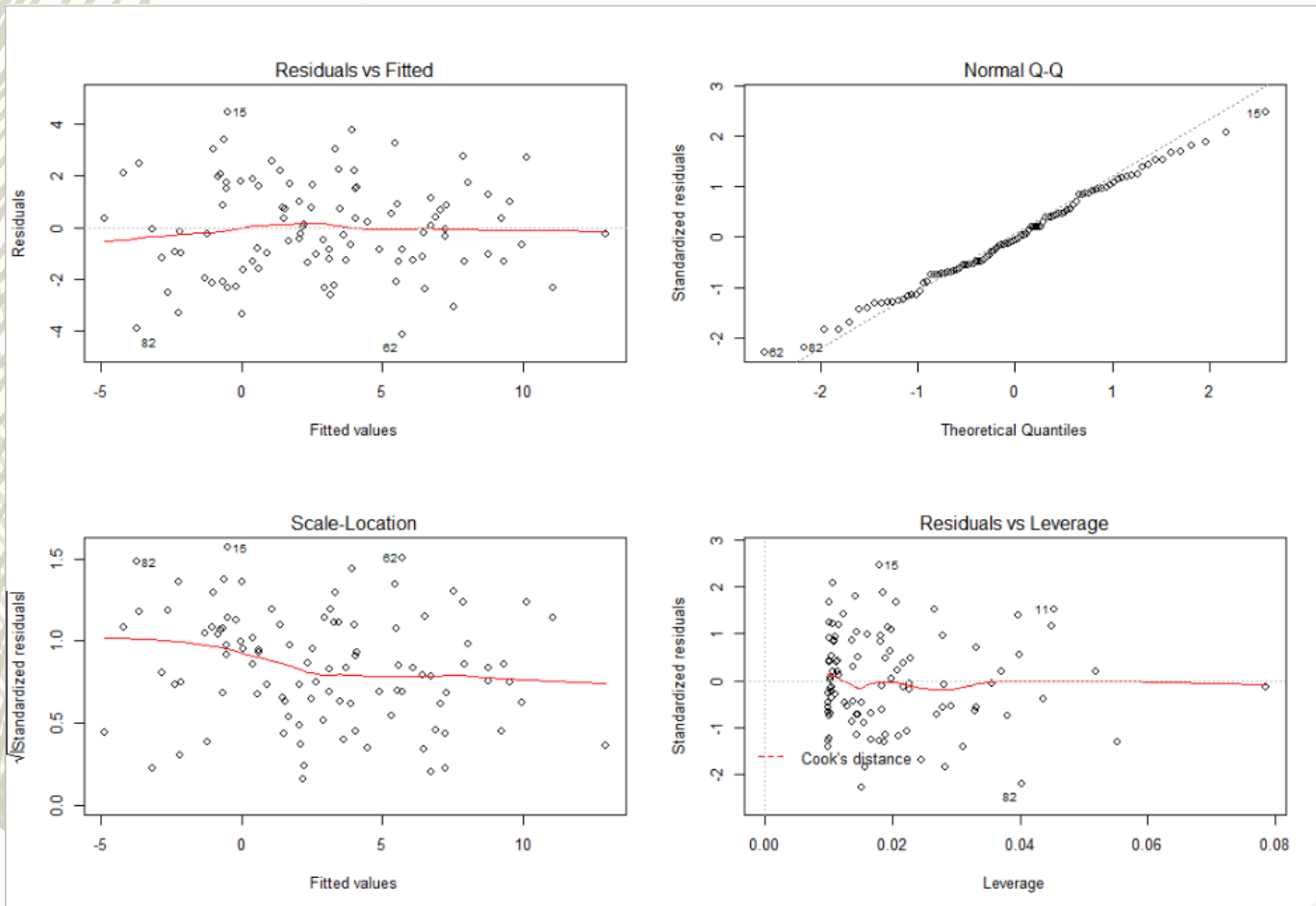
Interessante que a decisão seria sempre, neste caso, a de não rejeição da hipótese nula, mas nem sempre os resultados dos vários testes são consistentes.

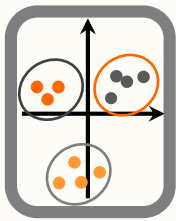


# avaliação de pressupostos

```
xs=rnorm(100)  
ys=3+4*xs+rnorm(100,0,2)  
par(mfrow=c(2,2))  
plot(lm(ys~xs))
```

## Análise de resíduos



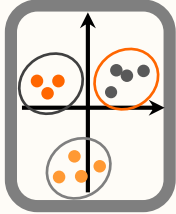


# Análise de resíduos

Veremos mais tarde que a diferença entre os valores observados e os valores ajustados (ou esperados) é denominada por resíduo.

**Importante:** os pressupostos de um procedimento estatístico são, geralmente, sobre os resíduos, não sobre os dados ou as observações propriamente ditas!

(E daí fazer muito pouco sentido transformar os dados para obter uma distribuição “mais Gaussiana” antes de ter ajustado um modelo!!!) - (erro comum, mesmo em analistas experientes!)



# avaliação de pressupostos

## Gráficos Q-Q

```
{r}  
par(mfrow=c(2,2),mar=c(4,4,2.5,0.5))  
set.seed(12345)  
qqnorm(rnorm(20),main="n=20")  
abline(0,1)  
qqnorm(rnorm(50),main="n=50")  
abline(0,1)  
qqnorm(rnorm(500),main="n=500")  
abline(0,1)  
qqnorm(rnorm(5000),main="n=5000")  
abline(0,1)
```

Permitem-nos avaliar a qualidade do ajustamento de um conjunto de dados a um modelo teórico

