



SME0803

ANÁLISE EXPLORATÓRIA DE DADOS

1º/2011

**Cibele Russo**

cibele@icmc.usp.br

<http://www.icmc.usp.br/~cibele>

Sala 3-160, ramal 8174

**CoteiaWIKI:**

[http://wiki.icmc.usp.br/index.php/SME-803\(cibele\)](http://wiki.icmc.usp.br/index.php/SME-803(cibele))



# PRINCIPAIS OBJETIVOS DO CURSO

- Apresentar as principais técnicas estatísticas para descrição e exploração de dados
- Conhecer os principais pacotes estatísticos
- Interpretar resultados de análises descritivas e exploratórias



# PROGRAMA RESUMIDO

1. Tipos de variáveis e organização de banco de dados.
2. Representação de variáveis qualitativas: tabelas de frequências, gráficos de barras e de setores.
3. Medidas descritivas de dados quantitativos: posição, dispersão, assimetria e concentração.
4. Representação de variáveis quantitativas: tabelas de frequências, histograma, gráfico de caixas, ramo-e-folhas e gráfico de pontos.
5. Tabelas de contingência e medidas de associação entre variáveis qualitativas.
6. Associação entre variáveis quantitativas e qualitativas.
7. Representação gráfica de dados multidimensionais.
8. Análise de agrupamentos.



# CRITÉRIO DE AVALIAÇÃO

$$\text{Média Final MF} = \frac{4 P1 + 5 P2 + NL}{10}$$

P1 Prova 1: 12/5/2011

P2 Prova 2: 30/6/2011

NL Nota de Listas de Exercícios

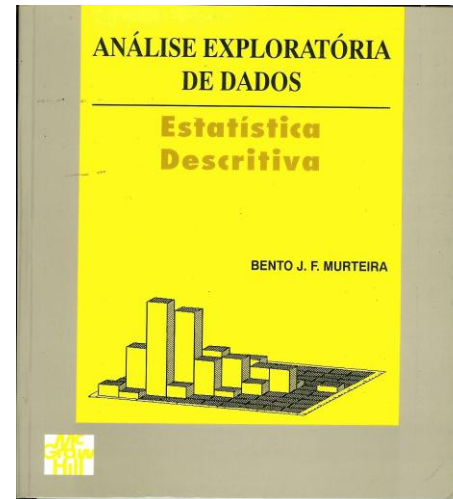
**Não haverá prova substitutiva.**

Critério de Recuperação: ver JúpiterWeb.

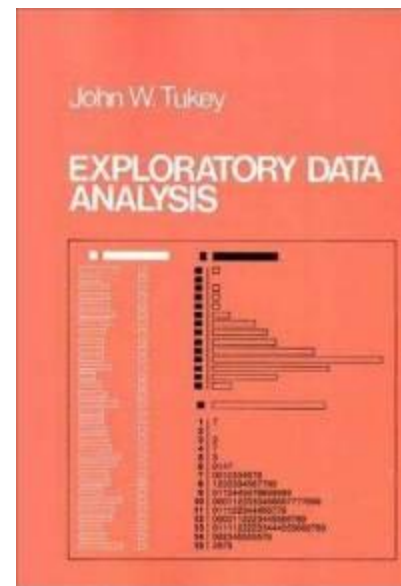


# BIBLIOGRAFIA PRINCIPAL

1. Murteira, B. J. F., Análise Exploratória de Dados. McGraw-Hill, Lisboa, 1993.

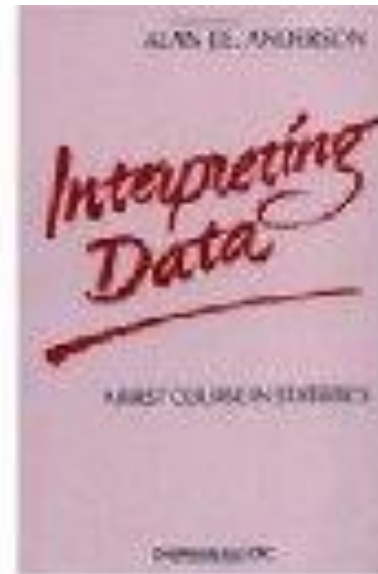


2. Tukey, J. W., Exploratory Data Analysis. Addison-Wesley, Reading-MA, 1977.  
(Disponível na biblioteca)



# BIBLIOGRAFIA COMPLEMENTAR

1. Anderson, A. J. B.,  
Interpreting Data: a First  
Course in Statistics.  
London:  
Chapman&Hall,1989.  
(Disponível na biblioteca)



2. Levine, D. M., Berenson, M. L., Stephan, D. Estatística: Teoria e Aplicações usando Microsoft Excel em Português. Rio de Janeiro:LTC, 1998.  
(Disponível na biblioteca)

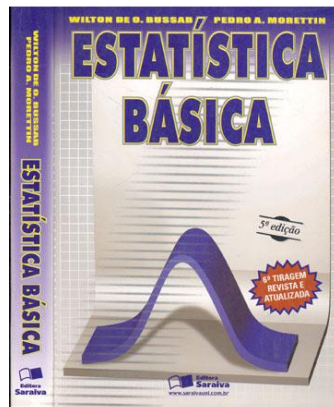


# BIBLIOGRAFIA COMPLEMENTAR

3. Magalhães, M.N.; Lima, A.C.P. Noções de Probabilidade e Estatística, 6ª ed., EDUSP, São Paulo, 2004. (Disponível na biblioteca)



4. Bussab, W.O.; Morettin, P.A. Estatística Básica, 6ª ed., Saraiva, São Paulo, 2009. (Disponível na biblioteca)

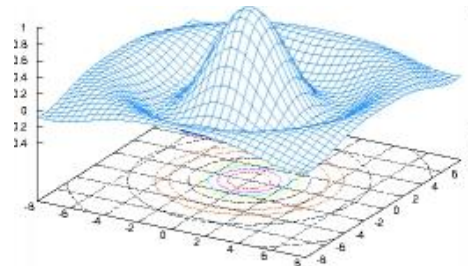
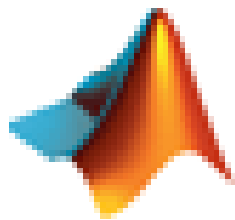
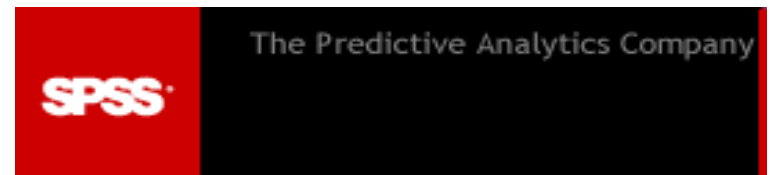


# APOIO COMPUTACIONAL

Excel.

R ([www.R-project.org](http://www.R-project.org)).

Minitab, S-PLUS, SAS, SPSS, Statistica, MatLab, Octave, Scilab, etc.





## O QUE É ESTATÍSTICA

Para muitos, Estatística não passa de conjuntos de tabelas. Os estatísticos são pessoas que coletam esses dados.

A Estatística originou-se com a coleta de dados e a construção de tabelas para os governos.

A situação evoluiu e a coleta de dados representa somente um dos aspectos da Estatística.

Estatística é um conjunto de técnicas utilizadas para a coleta, organização, processamento e análise de dados.

A Estatística é uma ciência baseada na Teoria da Probabilidade, cujo objetivo principal é nos auxiliar a tomar decisões ou tirar conclusões em situações de incerteza, a partir de dados.

## EM OUTRAS PALAVRAS...

**Estatística** é um conjunto de técnicas e métodos de pesquisa que envolve o planejamento do experimento a ser realizado, a coleta qualificada dos dados, a inferência, o processamento, a análise e a disseminação das informações  
(ENCE - Escola Nacional de Ciências Estatísticas)

# HISTÓRIA DA ESTATÍSTICA

Desenvolvimento da Estatística, principalmente no Brasil:

<http://www.redeabe.org.br/historia.htm>

Memória, J. M. P., Breve História da Estatística. EMBRAPA, Brasília, 2004.

[http://www.embrapa.br/publicacoes/tecnico/folderTextoDiscussao/arquivos-pdf/Miolo\\_21.pdf/view](http://www.embrapa.br/publicacoes/tecnico/folderTextoDiscussao/arquivos-pdf/Miolo_21.pdf/view)

Materials for the History of Statistics:

<http://www.york.ac.uk/depts/maths/histstat/>

História da Estatística (por Prof. Gauss Cordeiro):

[http://www.des.uem.br/uploads/arquivos\\_professor/1125192908.pdf](http://www.des.uem.br/uploads/arquivos_professor/1125192908.pdf)

# ANÁLISE EXPLORATÓRIA DE DADOS

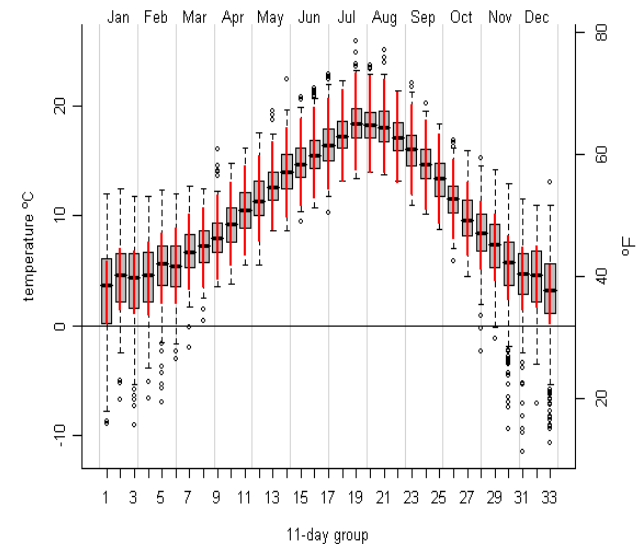
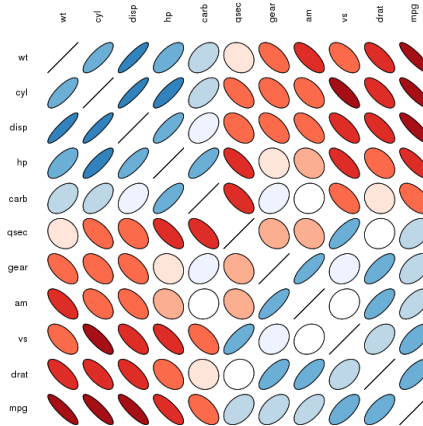
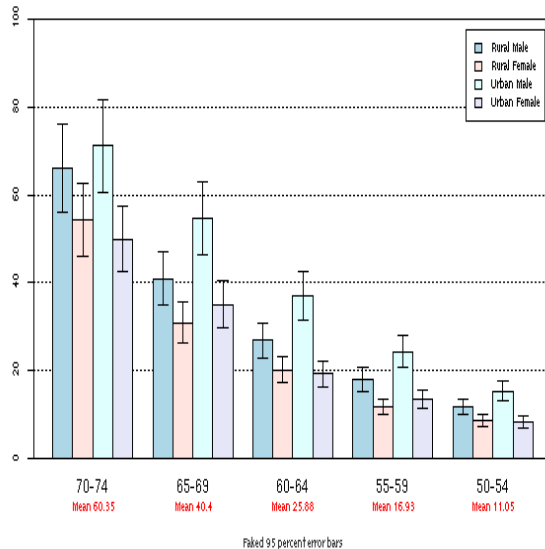
## Estatística Descritiva e Análise Exploratória

Etapas iniciais da **maioria** dos trabalhos estatísticos. Utilizadas para descrever, resumir e os dados.

A disponibilidade de uma grande quantidade de dados e de métodos computacionais muito eficientes revigorou estas áreas da Estatística.

A análise exploratória de dados utiliza técnicas gráficas.

Death Rates in Virginia



# ANÁLISE EXPLORATÓRIA DE DADOS



John W. Tukey (1915 – 2000)

<http://www.york.ac.uk/depts/maths/histstat/lifework.htm#t>

Exploratory Data Analysis is detective work – numerical detective work – or counting detective work – or graphical detective work ...

It can never be the whole story, but nothing else can serve as the foundation stone - as the first step.

## ANÁLISE CONFIRMATÓRIA DE DADOS

Apesar de ser uma etapa muito importante, o trabalho estatístico vai além da exploração e descrição de dados. É necessário **estender** as conclusões para além dos dados coletados.

Análise confirmatória é parte importante de muitos trabalhos. Técnicas de **inferência estatística** são utilizadas.

“Exploratory and confirmatory can – and should – proceed side by side”.

# VÍDEO: 200 PAÍSES, 200 ANOS, 4 MINUTOS

<http://www.gapminder.org/videos/200-years-that-changed-the-world-bbc/>

