

BIOESTATÍSTICA

M.I. Eng. Biomédica

2015-2016

Aula Teórica 2



Introdução

- População
 - Conjunto dos elementos cujos atributos são objecto de um determinado estudo.



Introdução

- População
 - Exemplos:
 - O conjunto das empresas transformadoras em Portugal em Setembro de 2011;
 - O conjunto dos cidadãos eleitores nas eleições legislativas de Junho de 2011;
 - O conjunto de valores de pressão arterial ao longo de um dia;



Introdução

- População
 - Para se conhecer de forma completa é necessário estudar todos os seus elementos efectuando um **censo**



Introdução

- Amostra
 - Subconjunto finito da população



Introdução

- Amostras
 - Exemplos
 - 25 empresas transformadoras escolhidas aleatoriamente a partir de dados publicados pelo MESS;
 - Um conjunto de 1000 cidadãos eleitores seleccionados à boca das urnas;
 - Medição da pressão arterial em intervalos sucessivos de 30 min durante um dia;



Introdução

- Atributos (variáveis)
 - Quantitativos (ex. altura)
 - Qualitativos
 - Nominais (ex. género)
 - Ordinais (ex. grau de escolaridade)



Introdução

- Representação numérica – qualquer que seja o tipo de variável a sua representação pode ser feita numericamente
 - Quantitativa – os valores numéricos são as intensidades;
 - Qualitativa – pode estabelecer-se uma correspondência entre o atributo e números inteiros;



Representação de dados

- Tabular;
- Gráfica;
- Indicadores numéricos;

Representação de dados

- Tabular

	Controlos (23)	Diabéticos (23)	p
Género			
M	8 (34.8%)	10 (43.5%)	0.763
Idade			
Média	50.39	64.43	0.001
Desvio padrão	12.409	12.978	
Min	32	28	
Max	74	85	
Fumador			
Sim		4 (17.4)	
Medicação			
Sim	3 (13.0%)	20 (87.0%)	<0.001
Nº dentes perdidos			
Média	7.57	10.35	0.123
Min	0	1	
Max	22	23	

Ana Reis *et. al.*

Representação de dados

- Tabular

	Amido	Lactose	Maltodextrina	Sacarose
Leite em pó	10 (18.9%)	29 (54.7%)	11 (20.8%)	3 (5.7%)
Papa	22 (16.8%)	33 (25.2%)	33 (25.2%)	43 (32.8%)
Papa líquida	3 (14.3%)	7 (33.3%)	4 (19.0%)	7 (33.3%)
Papa preparada	7 (26.9%)	11 (42.3%)	5 (19.2%)	3 (11.5%)
Sobremesa	17 (34.0%)	7 (14.0%)	7 (14.0%)	19 (38.0%)
Iogurte	8 (33.3%)	8 (33.3%)	0 (0.0%)	8 (33.3%)
Refeição	9 (69.2%)	1 (7.7%)	3 (23.1%)	0 (0.0%)

Joana Toscano *et. al.*

Representação de dados

- Tabular

Classificações do teste prático de biofísica

2.80	3.00	3.00	3.00	2.10	2.90	2.80
3.00	3.00	2.25	2.95	2.85		2.05
3.00	2.75	3.00	2.90	2.70	2.75	
3.00	2.95	3.00	2.60	2.80	2.90	2.90
2.95	2.60	3.00	2.90		2.70	2.35
2.70	2.85	3.00	3.00	1.70	2.75	2.90
2.90	2.80	2.85	2.95	2.90		2.75
2.30		1.65	2.70	2.80	1.00	2.20
2.90	2.90	2.50	2.70	1.70	0.60	
3.00	2.40	2.90	3.00	2.70	1.35	

Representação de dados

- Tabular

Classificações do teste prático de biofísica

0.60	1.00	1.35	1.65	1.70	1.70	2.05
2.10	2.20	2.25	2.30	2.35	2.40	2.50
2.60	2.60	2.70	2.70	2.70	2.70	2.70
2.70	2.75	2.75	2.75	2.75	2.80	2.80
2.80	2.80	2.80	2.85	2.85	2.85	2.90
2.90	2.90	2.90	2.90	2.90	2.90	2.90
2.90	2.90	2.90	2.95	2.95	2.95	2.95
3.00	3.00	3.00	3.00	3.00	3.00	3.00
3.00	3.00	3.00	3.00	3.00	3.00	3.00

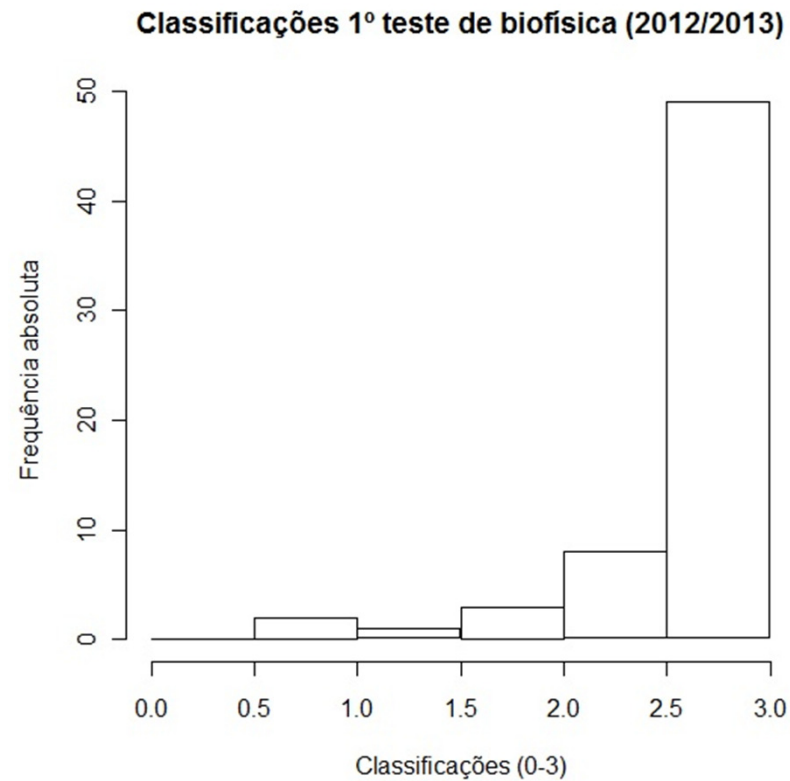
Representação de dados

- Gráfico de caule e folhas

Caule	Folhas
5	47
6	2356668889
7	0011233445557889
8	122344557
9	1558

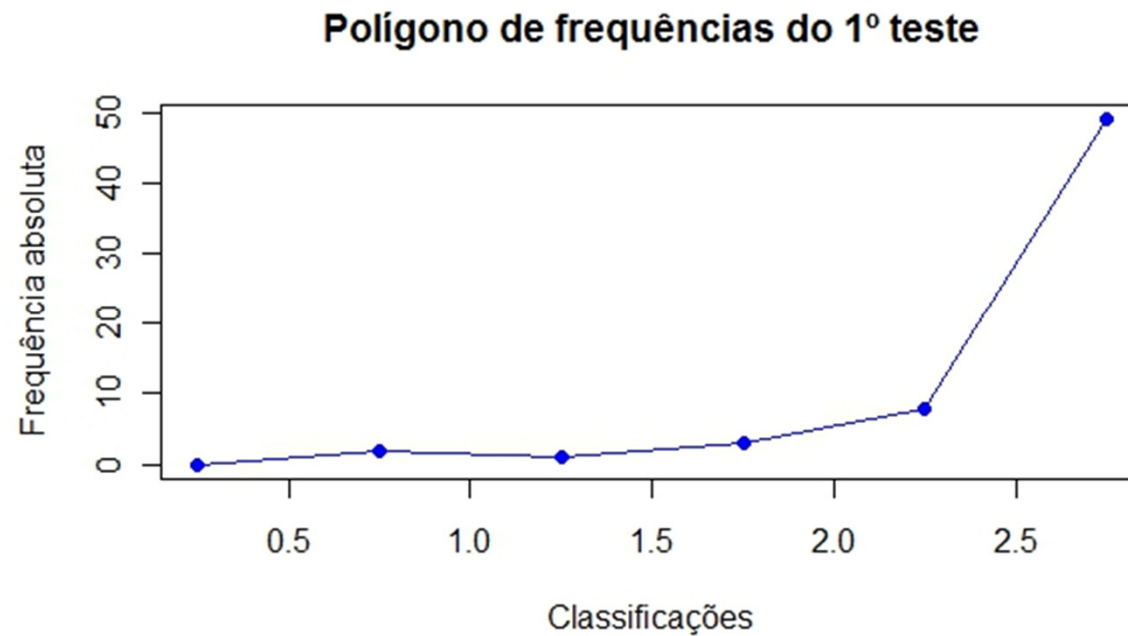
Representação de dados

- Gráfica: histograma



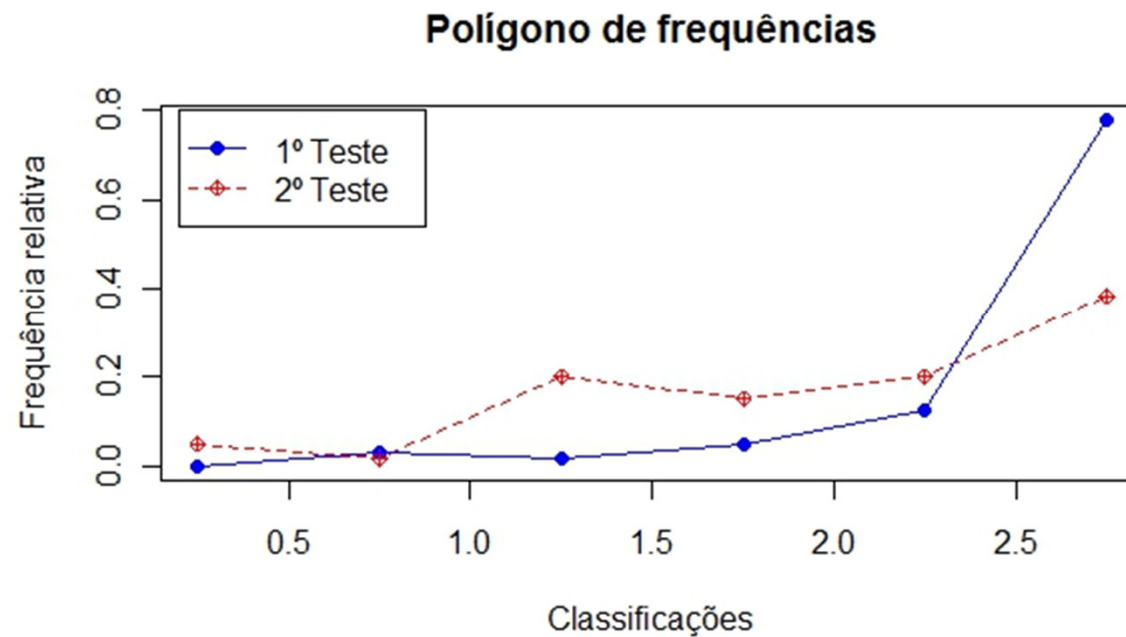
Representação de dados

- Gráfica: polígono de frequências



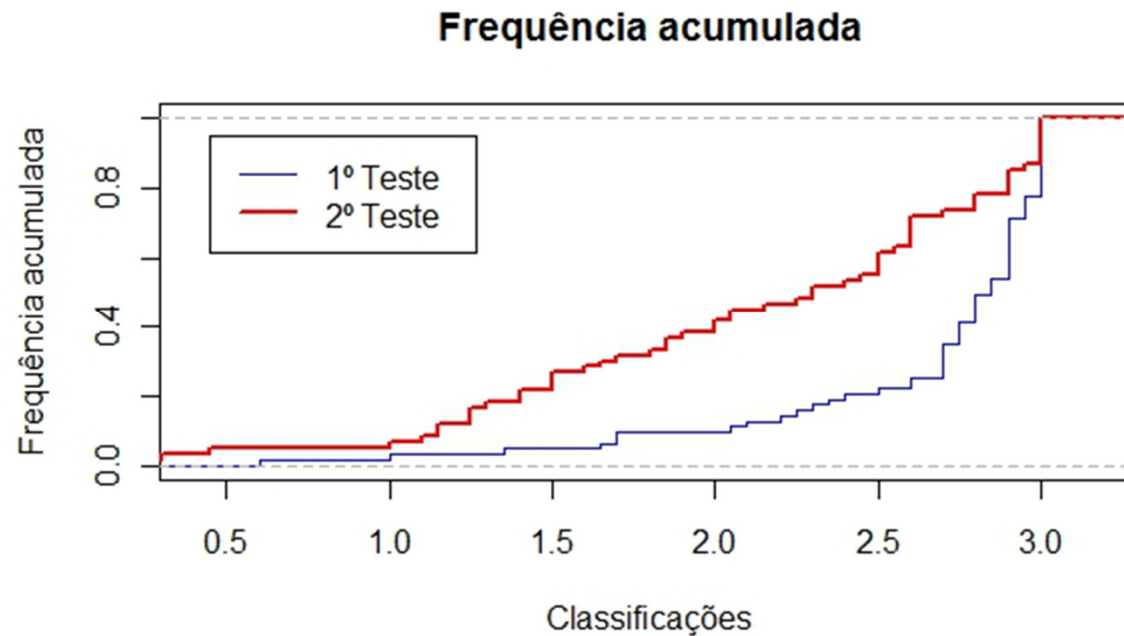
Representação de dados

- Gráfica: polígono de frequências



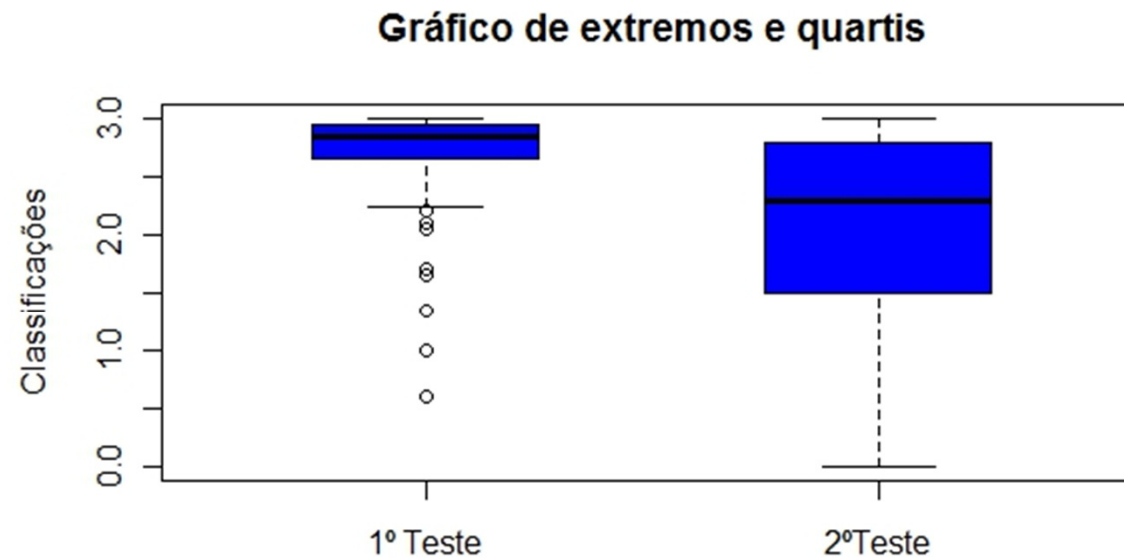
Representação de dados

- Gráfica: frequência acumulada



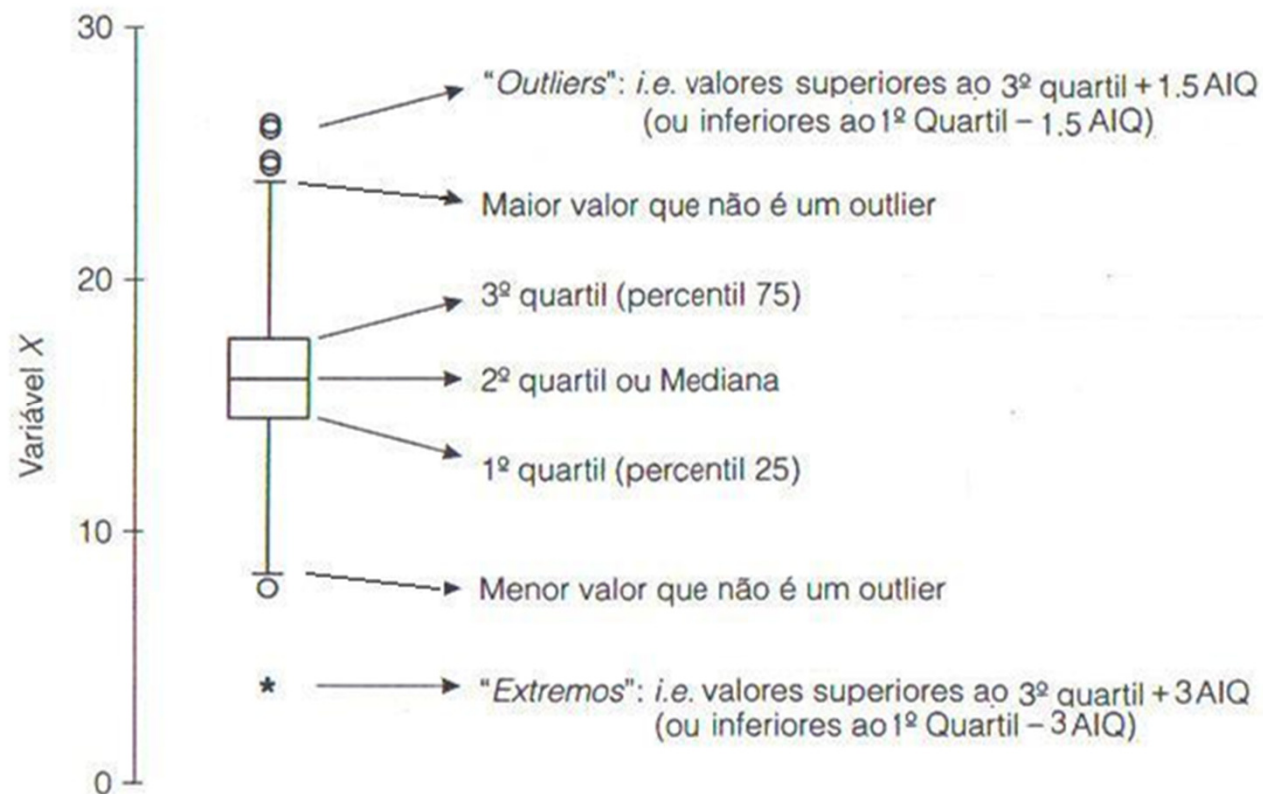
Representação de dados

- Gráfica: extremos e quartis



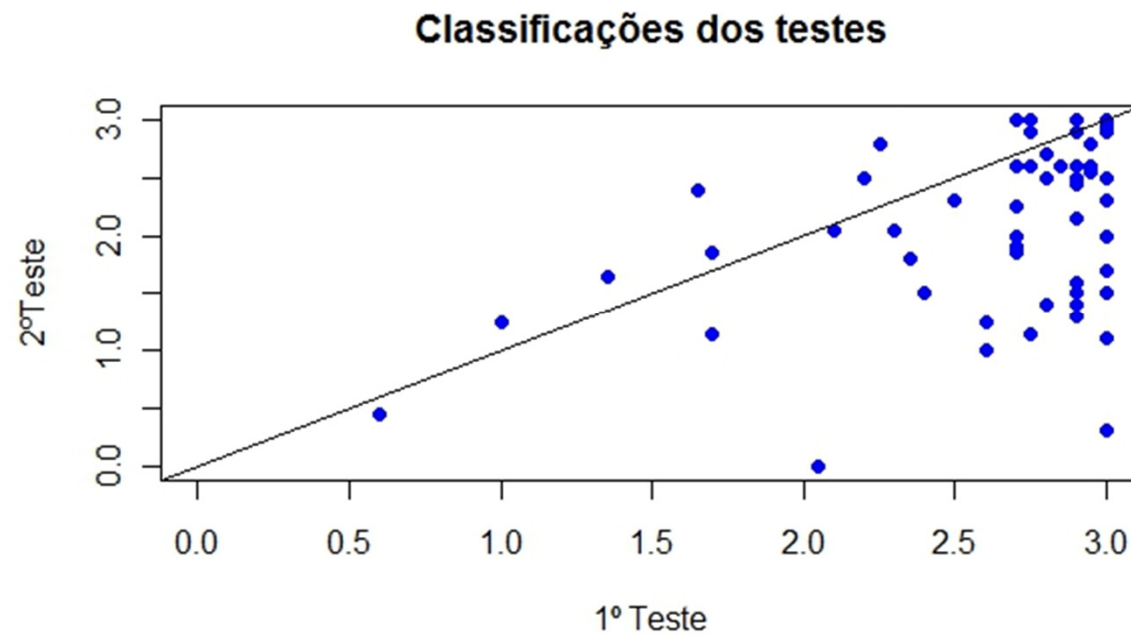
Representação de dados

- Caixa de bigodes (box plot)



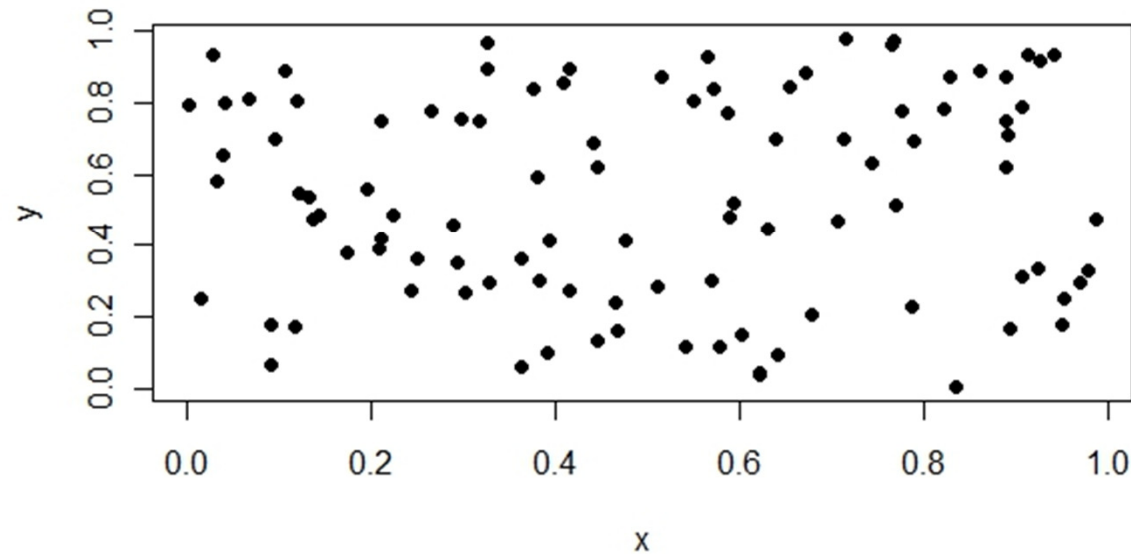
Representação de dados

- Gráfica: dispersão



Representação de dados

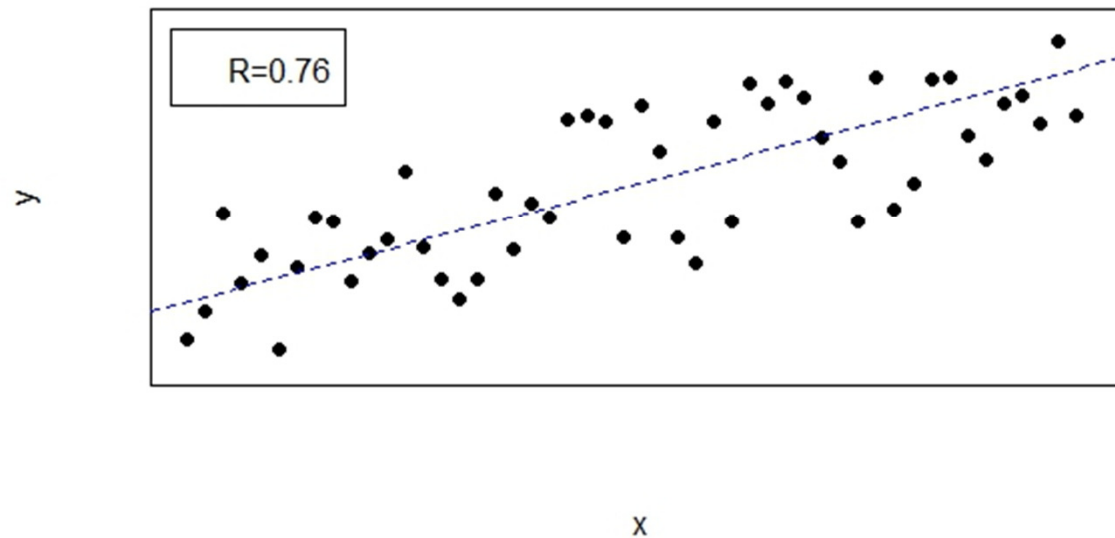
- Gráfica: dispersão (observação de correlação)



ausência de correlação

Representação de dados

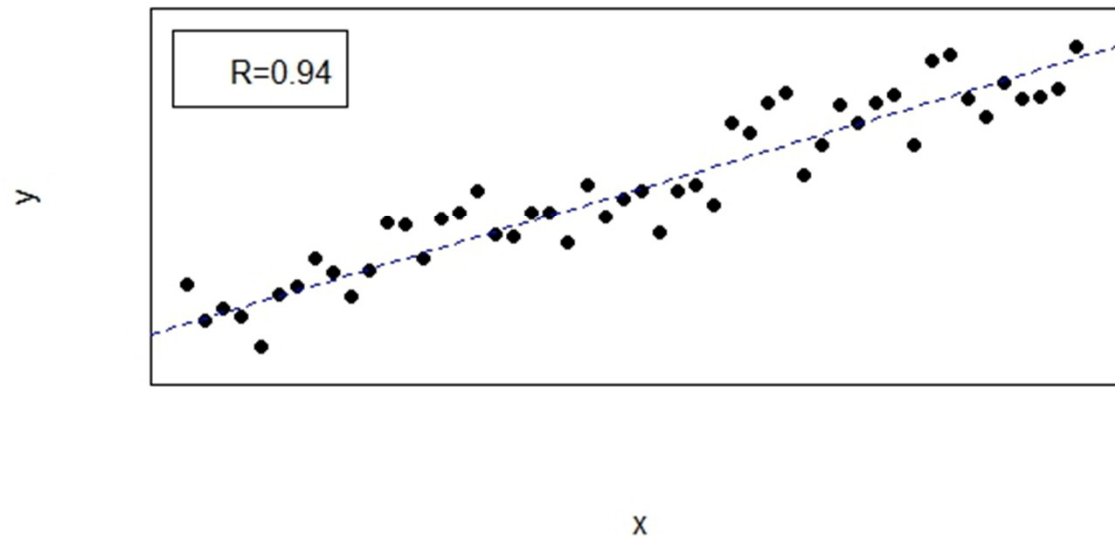
- Gráfica: dispersão (observação de correlação)



correlação positiva

Representação de dados

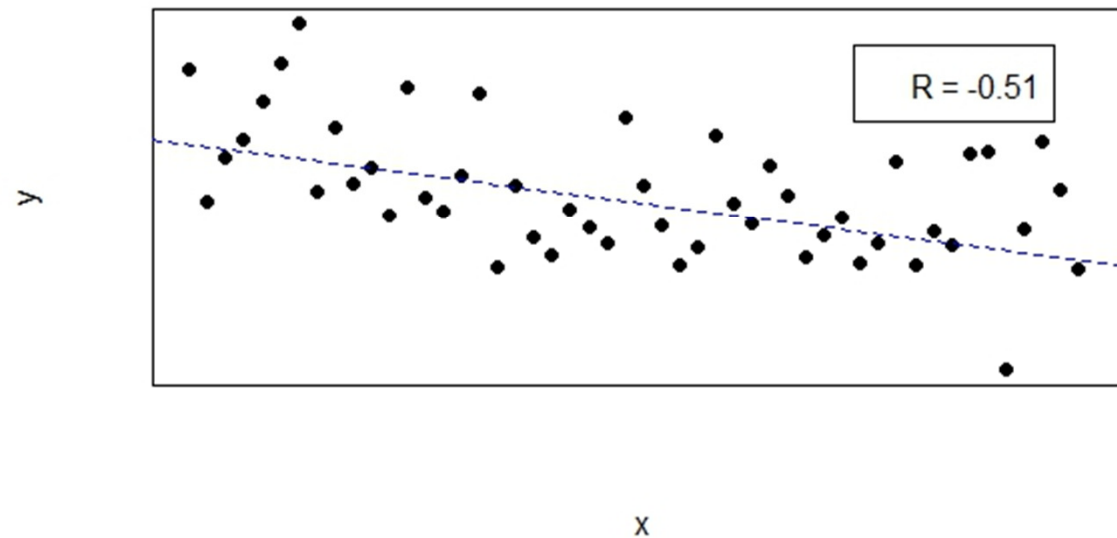
- Gráfica: dispersão (observação de correlação)



correlação positiva

Representação de dados

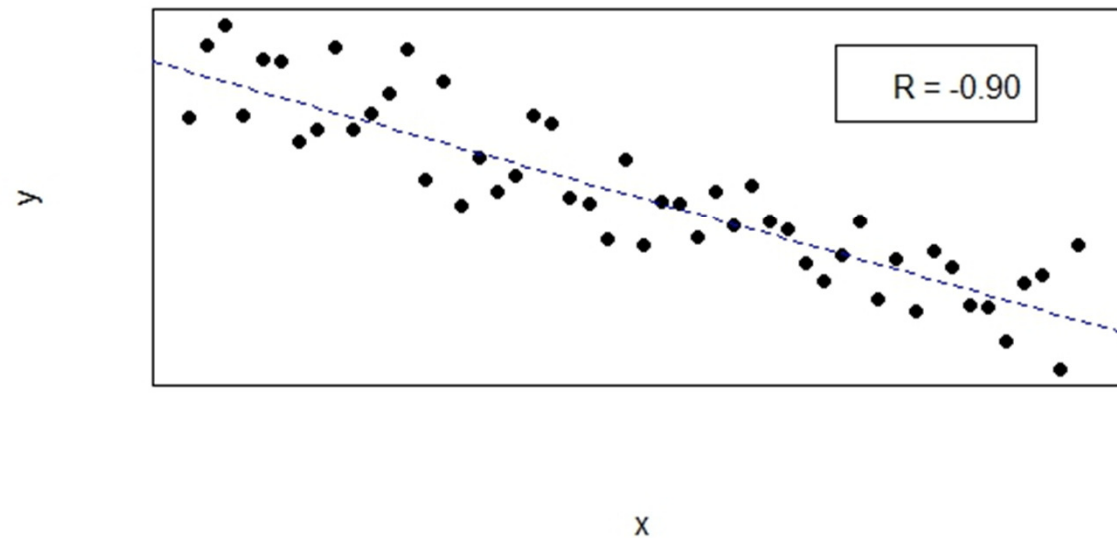
- Gráfica: dispersão (observação de correlação)



correlação negativa

Representação de dados

- Gráfica: dispersão (observação de correlação)



correlação negativa



Visualização de dados

- “Data visualization is the display of information in a graphic or tabular form.”
- “Successful visualization requires that the data (information) be converted into a visual format that the characteristics of the data and the relationships among data items or attributes can be analyzed or reported”



Visualização de dados

- “The goal of visualization is the interpretation of the visualized information by a person and the formation of a mental model of the information”

“Introduction to Data Mining “, Pang-Ning Tan et. al.

- “The main goal of data visualization is to communicate information clearly and effectively through graphical means.”

Friedman (2008)

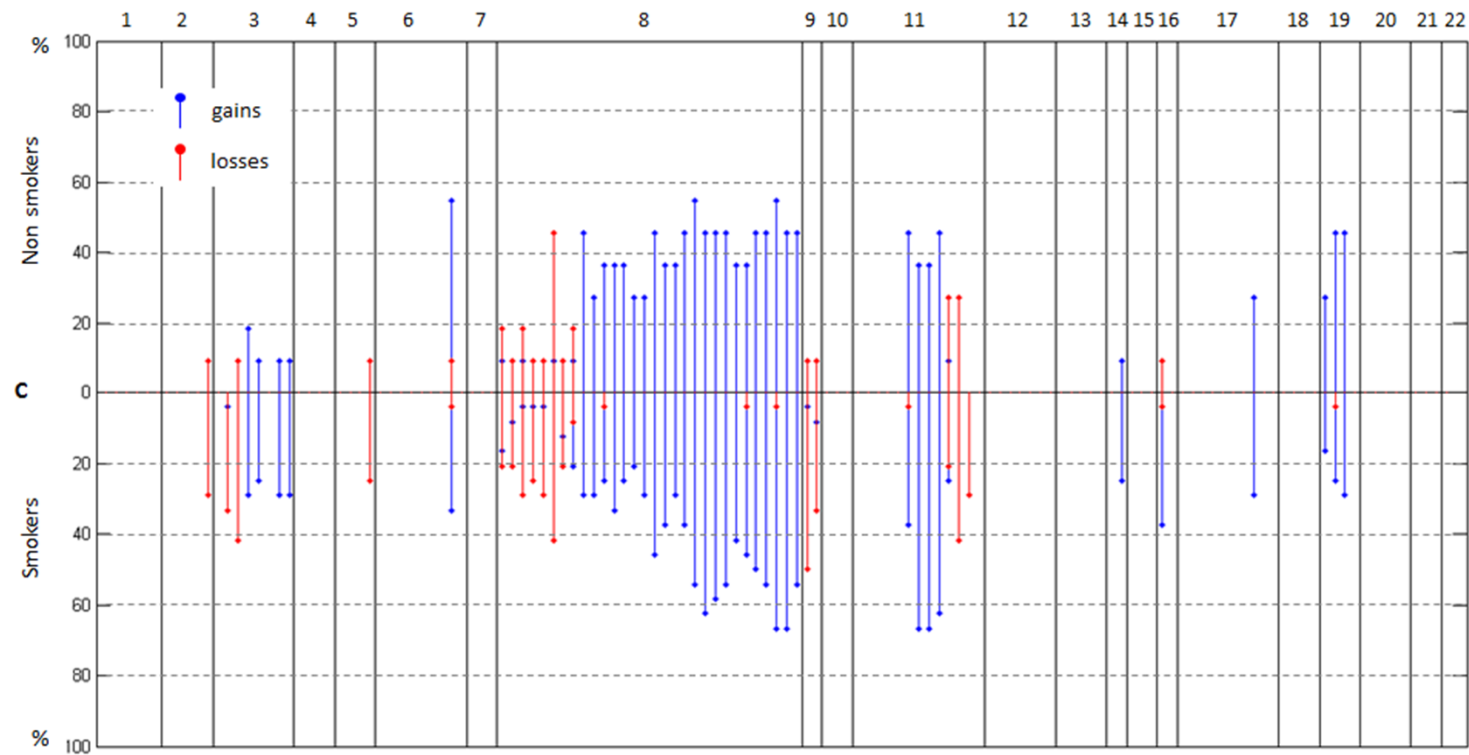


Visualização de dados

- Organizar os dados de forma:
 - a visualizar os padrões e as relações pertinentes;
 - a contar uma história
 - a fazer sentido
 - a focar apenas na informação pertinente
- Compressão de conhecimento

Visualização de dados

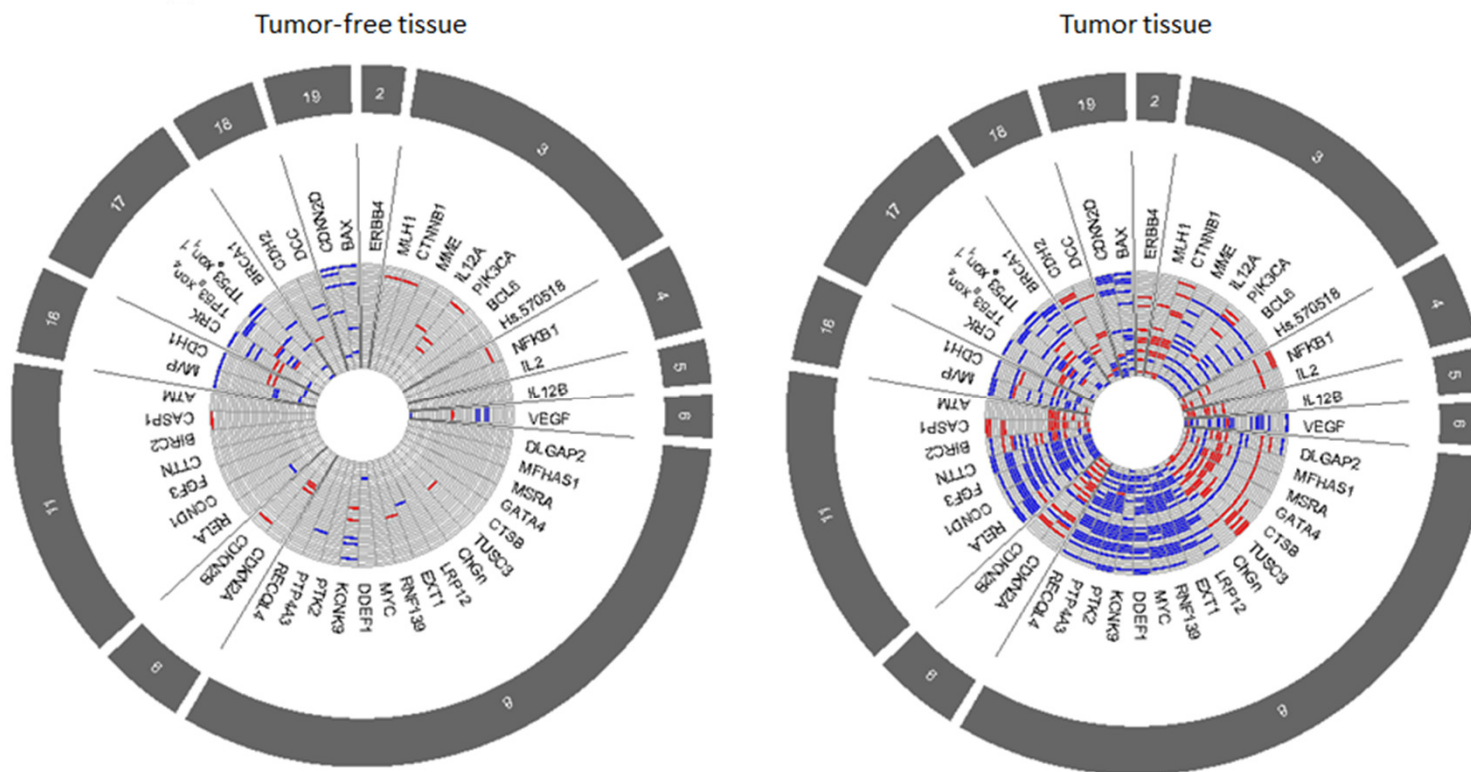
- Exemplo



Ilda Ribeiro *et al.*

Visualização de dados

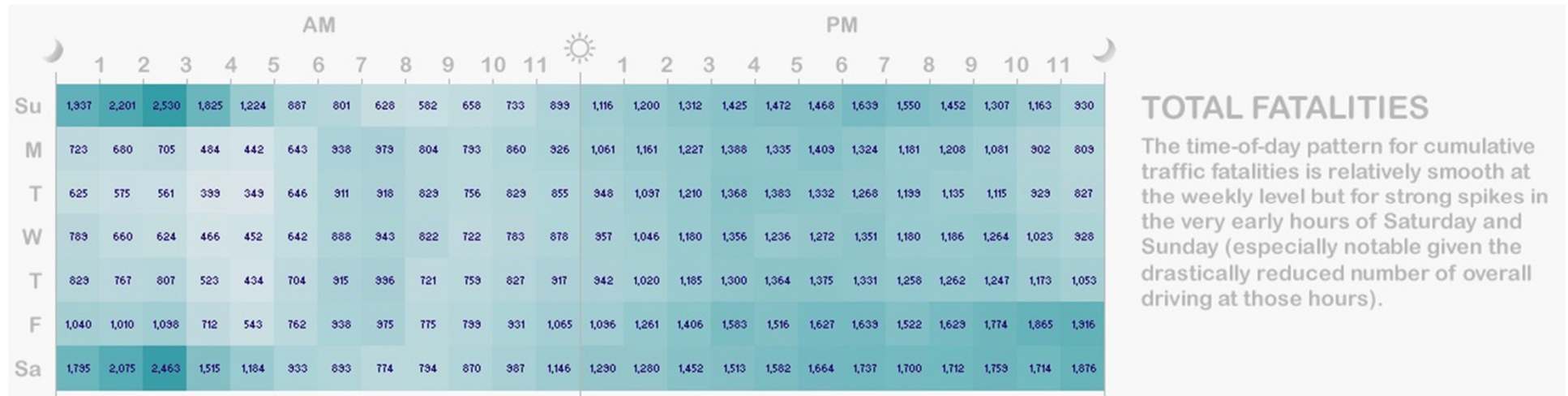
- Exemplo



Ilda Ribeiro *et al.*

Visualização de dados

- Exemplo

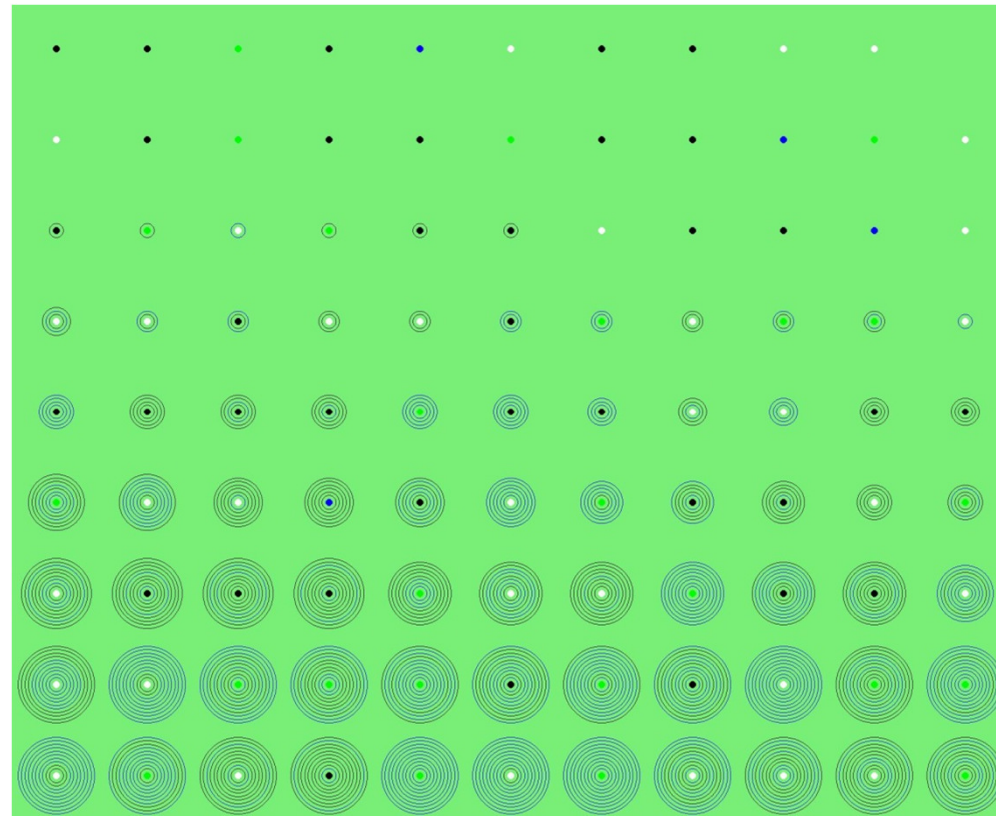


<http://www.flickr.com/photos/idvsolutions/8265288798/sizes/o/in/photostream/>

Visualização de dados

- Exemplo

Laboratório de Bioestatística e Informática Médica



Visualização de dados

- Exemplo

