

Introdução à Estatística e Estatística Descritiva

ESTATÍSTICA?

Um conjunto de procedimentos e princípios para recolha, compilação, análise e interpretação de dados por forma a **ajudar na tomada de decisões quando na presença de incerteza.**

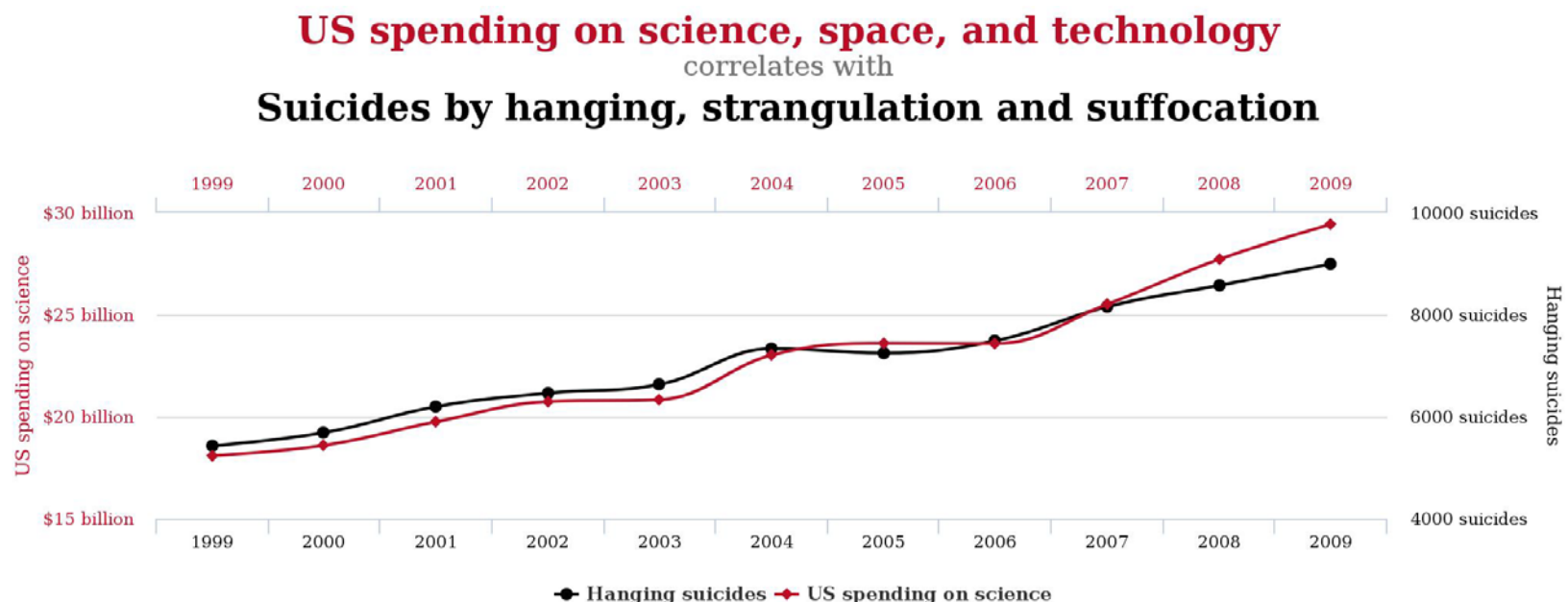
Herbert George Wells,
English author, said (circa 1940),

“Statistical thinking will one day be as necessary for
efficient citizenship as the ability to read and write”



Why it is important Statistical Literacy

- We are in the **age of information**
 - understanding statistics well enough to **be able** to **consume** the **information** we are inundated with on a daily basis, think **critically** about it, and **make good decisions** based on that information

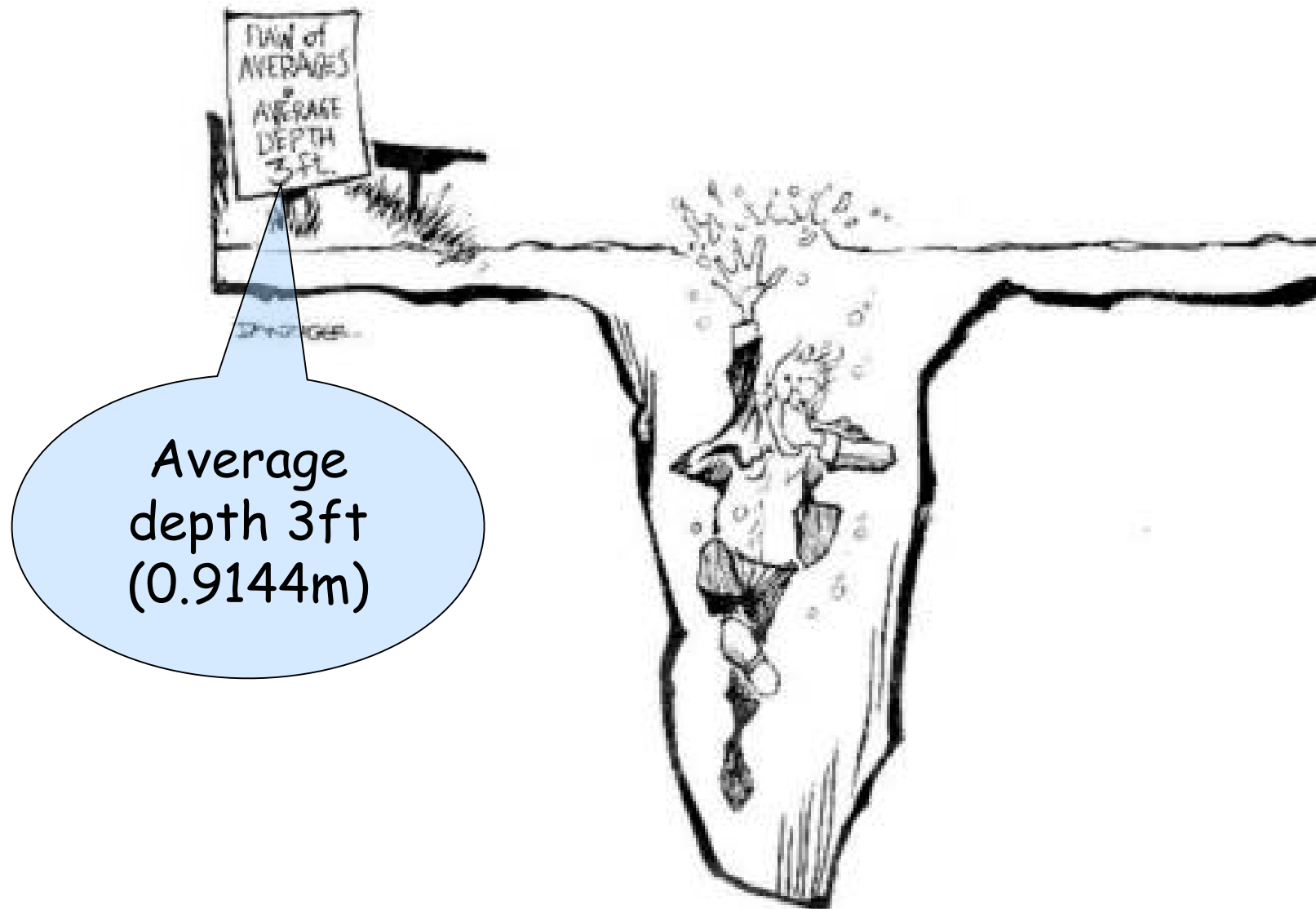


tylervigen.com

<http://tylervigen.com/spurious-correlations>

Why it is important Statistical Literacy

- Research scientist skills
 - the ability to identify questions, collect evidence (data), discover and apply tools to interpret the data, and communicate and exchange results
 - be able to **distinguish** the **proper** from the **improper** use of the scientific **method**
- The Challenge of Developing Statistical Reasoning , Joan Garfield , *Journal of Statistics Education* Volume 10, Number 3, 2002
- **students are (mainly) consumers and (at times) producers of information.**





Why do we need to understand statistics?

Reasoning with Uncertainty

- *from*
- Peter Donnelly: How juries are fooled by statistics
- <http://www.ted.com/index.php/talks/view/id/67>

Ex 1 - Coin Tossing

- Imagine tossing a coin successively, and waiting till the first time a particular pattern appears, say **HTT**
- For example, if the sequence of tosses was

HHTHH**TH**HTTTHTH

- The pattern HTT would first appear after the 10th toss

- Imagine that half of you toss a coin several times, each time till the sequence **HTT** occurs.
 - Record the average number of tosses till HTT occurs
- The other half of you prefer to count **HTH**
 - Record the average number of tosses till HTH occurs

- Which of the following is true:
 - A. The average number of tosses until HTH is larger than the average number of tosses until HTT
 - B. The average number of tosses until HTH is the same as the average number of tosses until HTT
 - C. The average number of tosses until HTH is smaller than the average number of tosses until HTT

Most people think that B is true but **A is true**.

The average number of tosses till HTH is 10 and the average number of tosses till HTT is 8.

- Intuitive explanation:
- Imagine that you win if HTH occurs
 - If the first toss gives a H you are excited and you get even more excited if the second is a T. If the third is H you win but if it is a T you have to start again and wait for the next H.
- If you win when HTT occurs
 - For the first two tosses the experience is the same. However, if the third toss is a H you loose but you already have the first H and are $1/3$ of the way to your pattern.

**It was an example of a simple
question on probabilities that
most people get wrong**



Cardozo falha um penalty em cada cinco

Por **Gonçalo Guimarães**

A- A A+

Ser ou não ser um especialista, eis a questão. Óscar Cardozo, 30 anos, falhou anteontem, em Barcelos, o 11.º penalty com a camisola do Benfica, em todas as provas oficiais, naquela que foi a sua 51.ª tentativa da marca dos onze metros. Isto significa que o ponta de lança paraguaio, na Luz desde 2007, tem uma percentagem de eficácia de 78,4 por cento. Dito de outra forma, falha praticamente um penalty em cada cinco vezes que é chamado a converter.

<http://www.abola.pt/nnh/ver.aspx?id=457643>

08:57 - 03-02-2014

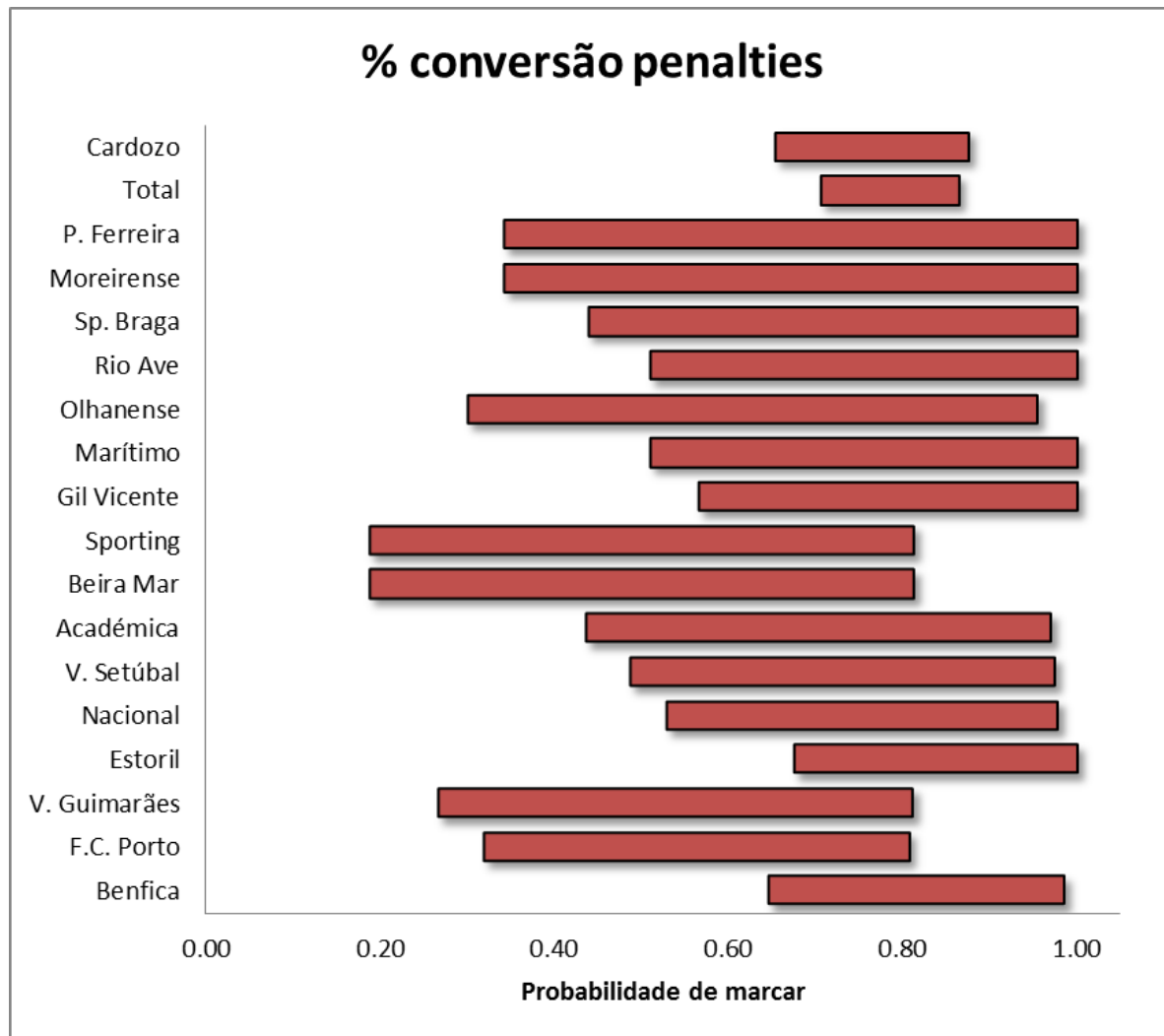
$$\frac{51 - 11}{51} = 0.784$$

O que significa?

Probabilidade de marcar

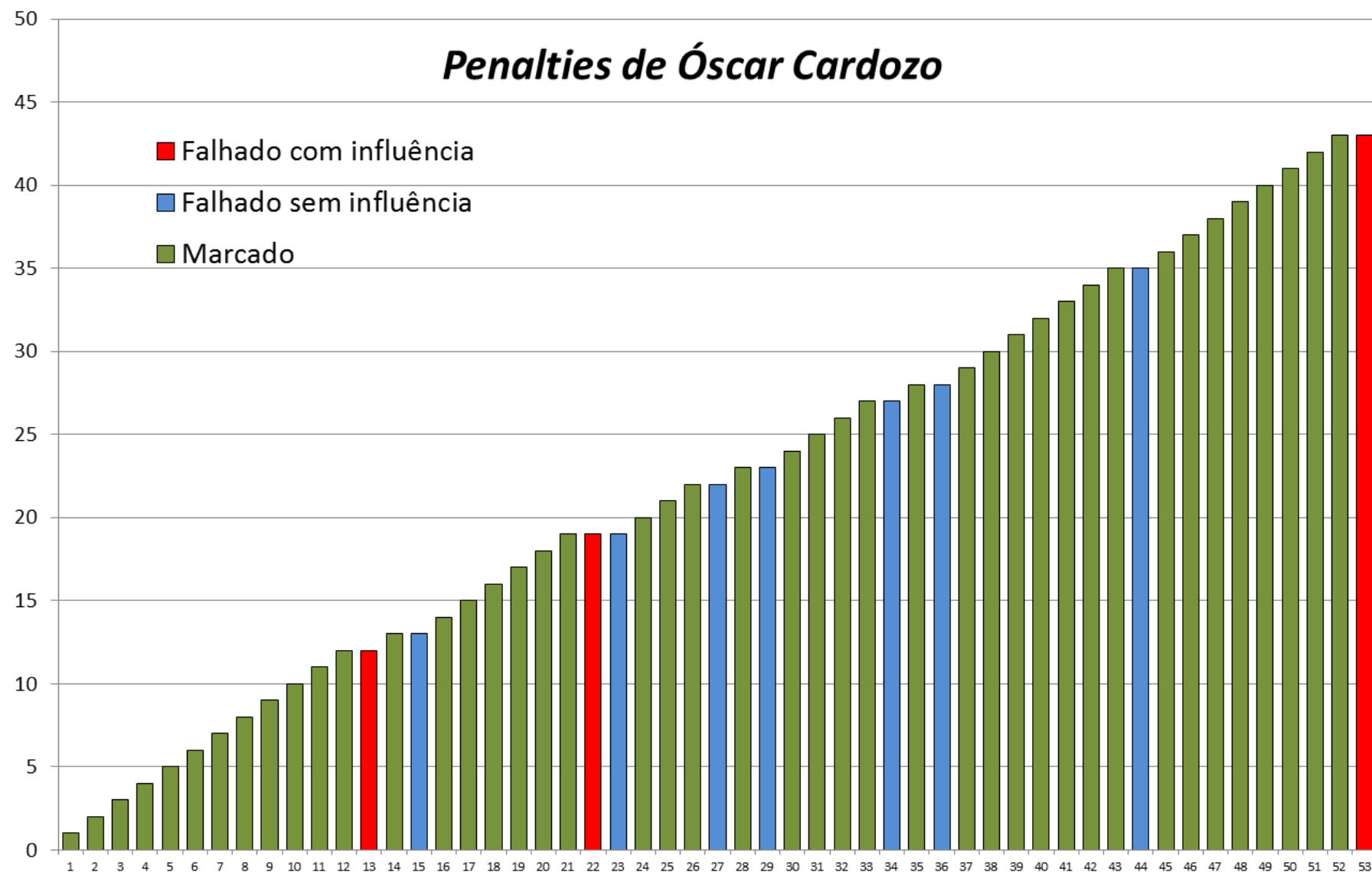
$$\hat{p} \in [0.65; 0.88]$$

http://en.wikipedia.org/wiki/Binomial_proportion_confidence_interval



Período	Quem	Total	Convertidos	Falhados	%	Wilson score interval	
						L.I.	L.S.
2012/2013	Benfica	12	11	1	0.92	0.65	0.99
2012/2013	F.C. Porto	12	7	5	0.58	0.32	0.81
2012/2013	V. Guimarães	9	5	4	0.56	0.27	0.81
2012/2013	Estoril	8	8		1.00	0.68	1.00
2012/2013	Nacional	8	7	1	0.88	0.53	0.98
2012/2013	V. Setúbal	7	6	1	0.86	0.49	0.97
2012/2013	Académica	6	5	1	0.83	0.44	0.97
2012/2013	Beira Mar	6	3	3	0.50	0.19	0.81
2012/2013	Sporting	6	3	3	0.50	0.19	0.81
2012/2013	Gil Vicente	5	5		1.00	0.57	1.00
2012/2013	Marítimo	4	4		1.00	0.51	1.00
2012/2013	Olhanense	4	3	1	0.75	0.30	0.95
2012/2013	Rio Ave	4	4		1.00	0.51	1.00
2012/2013	Sp. Braga	3	3		1.00	0.44	1.00
2012/2013	Moreirense	2	2		1.00	0.34	1.00
2012/2013	P. Ferreira	2	2		1.00	0.34	1.00
	Total	98	78	20	0.80	0.71	0.86
até Jan 2014	Cardozo	51	40	11	0.78	0.65	0.88

Penalties de Óscar Cardozo



Conclusions from the examples

- Randomness, uncertainty and chance are part of our life
- People make errors of logic when reasoning with uncertainty
- Errors in statistics may have serious consequences
 - (see [Sally Clark](#) example in the video)

It is very important to understand statistics!

e
**PORQUE É QUE UM
ENGENHEIRO INFORMÁTICO
PRECISA DE APRENDER
ESTATÍSTICA?!**

Desenvolvimento de Software apoiado em ferramentas

ex.

SVN

software versioning and revision control system
evolução do software

Bugzilla

gestão de bugs

Dr. Watson

dados de execução

SVN

Bugzilla

Dr. Watson

...

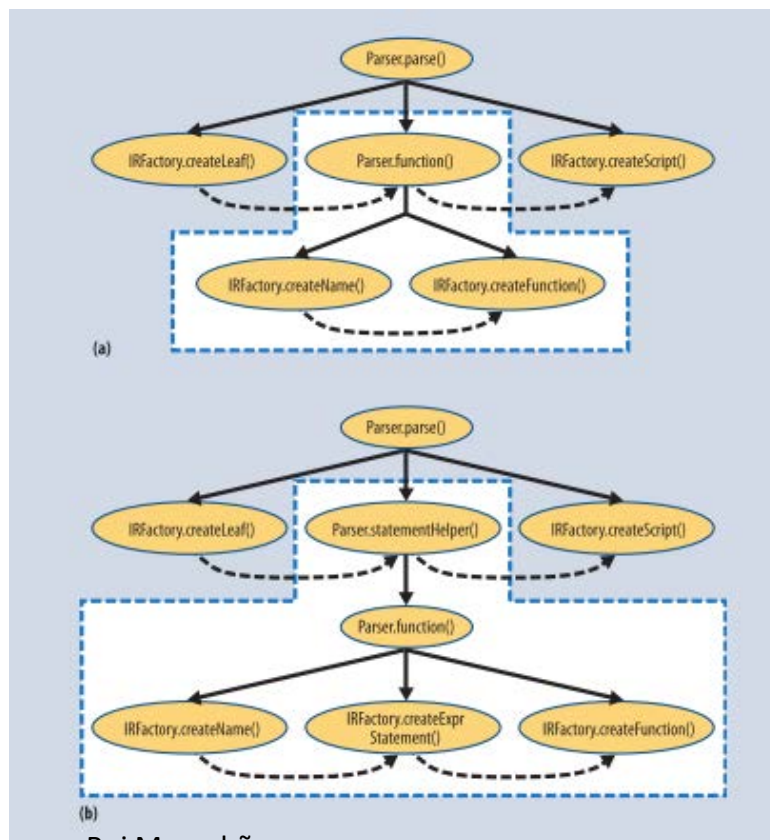
Engenharia de Software

gera muitos
dados

objetivo: extrair conhecimento útil dos dados

ex.

- qual a linha de código que está na origem do erro?
 - identificar padrões em erros ocorridos durante a execução de programas registadas com o Dr. Watson
 - aplicar esses padrões ao bug do programa que estou a desenvolver

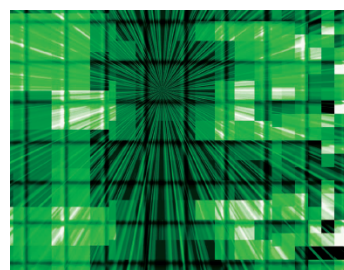


Xie, T., Thummalapenta, S., Lo, D., & Liu, C. (2009). Data Mining for Software Engineering. *Computer*, 42(8), 55–62. doi:10.1109/MC.2009.256

objetivo: extrair conhecimento útil dos dados
ex.
como?

Data Mining!

[e amigos]



DATA MINING FOR SOFTWARE ENGINEERING

Tao Xie and Suresh Thummalapenta, North Carolina State University
David Lo, Singapore Management University
Chao Liu, Microsoft Research

To improve software productivity and quality, software engineers are increasingly applying data mining algorithms to vari-

SE data concerns the 3Ps: people, processes, and products. People include software developers, testers, project managers and users. Processes include various develop-

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5197425>

Data Mining!

[e amigos]

para aprender (principalmente) em

- **Inteligência Artificial**
 - 3º ano
- **Extração de Conhecimento e Aprendizagem Computacional**
 - 4º ano

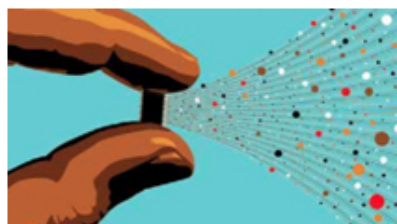
... e baseado em

ESTATÍSTICA!

O que deve ter um analista de dados para ser contratado?

19 de Dezembro de 2012 às 10:07:33 por computerworld

Com o aumento da competição pela contratação de profissionais da análise de dados, as empresas que conhecem as características exactas do que estão à procura têm vantagem.



Nenhuma outra área do mercado de trabalho de TI experimentou o crescimento de procura e a disparidade de oferta que o campo da análise de dados. À medida que o Big Data continua a crescer, e o campo da análise continua a amadurecer, eles estão tornando-se uma parte importante do negócio e do processo de tomada de decisão, levando a competição por grandes talentos a transformar-se numa verdadeira luta pelas contratações.


Num relatório sobre Big Data, o instituto McKinsey Global prevê que até 2018 os EUA passarão por uma escassez enorme de pessoal na área de análise.

"Ocorrerá uma escassez de talento necessário para que as organizações tirem proveito do Big Data. Até 2018, só os Estados Unidos podem encarar uma falta de 140 mil a 190 mil pessoas com profundas capacidades de análise, como também 1,5 milhões de gestores e analistas com o conhecimento para utilizarem a análise de grande volumes de dados a fim de tomarem decisões eficientes", diz o relatório.

O McKinsey Global Institute Report prevê também que as empresas que utilizam a análise no seu potencial completo podem aumentar as margens de operação em até 60%. As empresas estão a ter isso em conta, criando competição pelos melhores talentos.

Apenas reforçando o quão competitivo o mercado é, um relatório da Dice, do início deste ano, afirma que a actual taxa de desemprego da TI está em torno dos 3,6%.


"Os CIOs de hoje em dia precisam de alguns pensadores do tipo 'dados são dados' que não se intimidem pelo quão 'grande' os dados são ou se eles englobam ou não SKUs


Harvard Business Review

[THE MAGAZINE](#)
[BLOGS](#)
[VIDEO](#)
[BOOKS](#)
[CASES](#)
[WEBINARS](#)
[COURSES](#)

Guest
Subscribe today and get access to all current articles and HBR online archive







THE MAGAZINE
October 2012



ARTICLE PREVIEW To read the full article, [sign-in](#) or [register](#). HBR subscribers, click [here to register for FREE access](#) »

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

Comments (91)



RELATED

Executive Summary

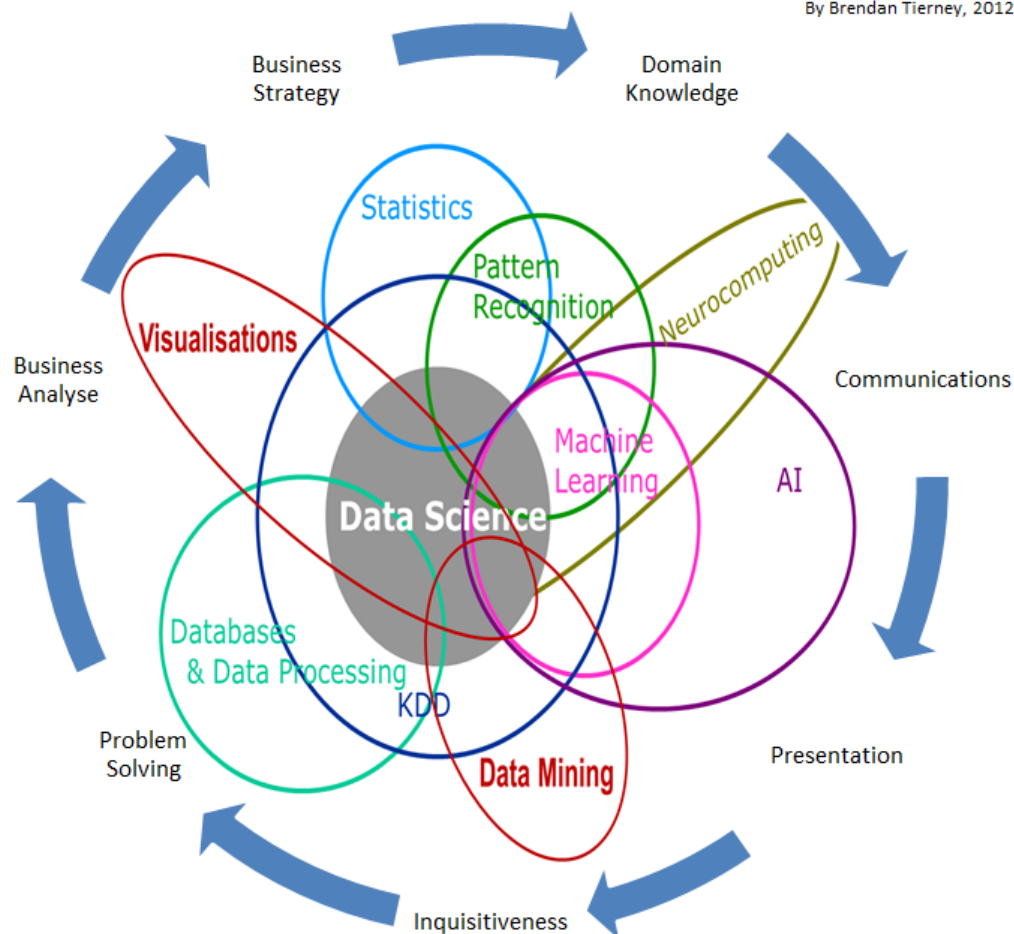
ALSO AVAILABLE

- Buy PDF

data science is a hot topic

Data Science Is Multidisciplinary

By Brendan Tierney, 2012





What is the problem here?

On average the
temperature is very
nice...

Estatística Descritiva

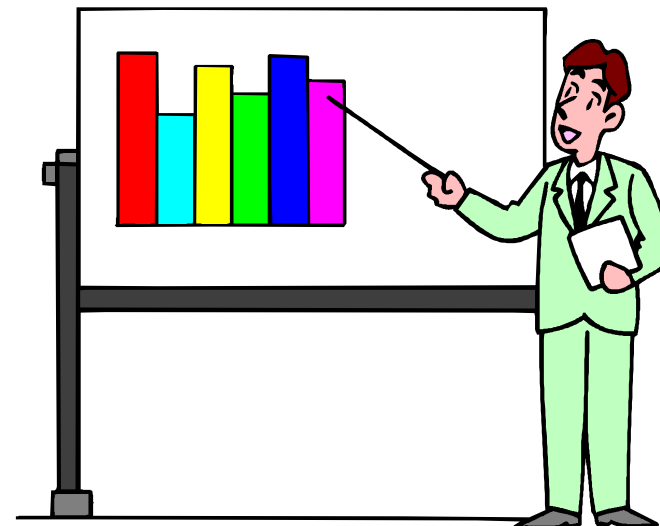
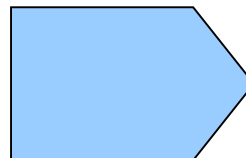
Na estatística descritiva procura-se sintetizar e representar de forma compreensível a informação contida num conjunto de dados

através da construção de

tabelas,

gráficos,

cálculo de medidas

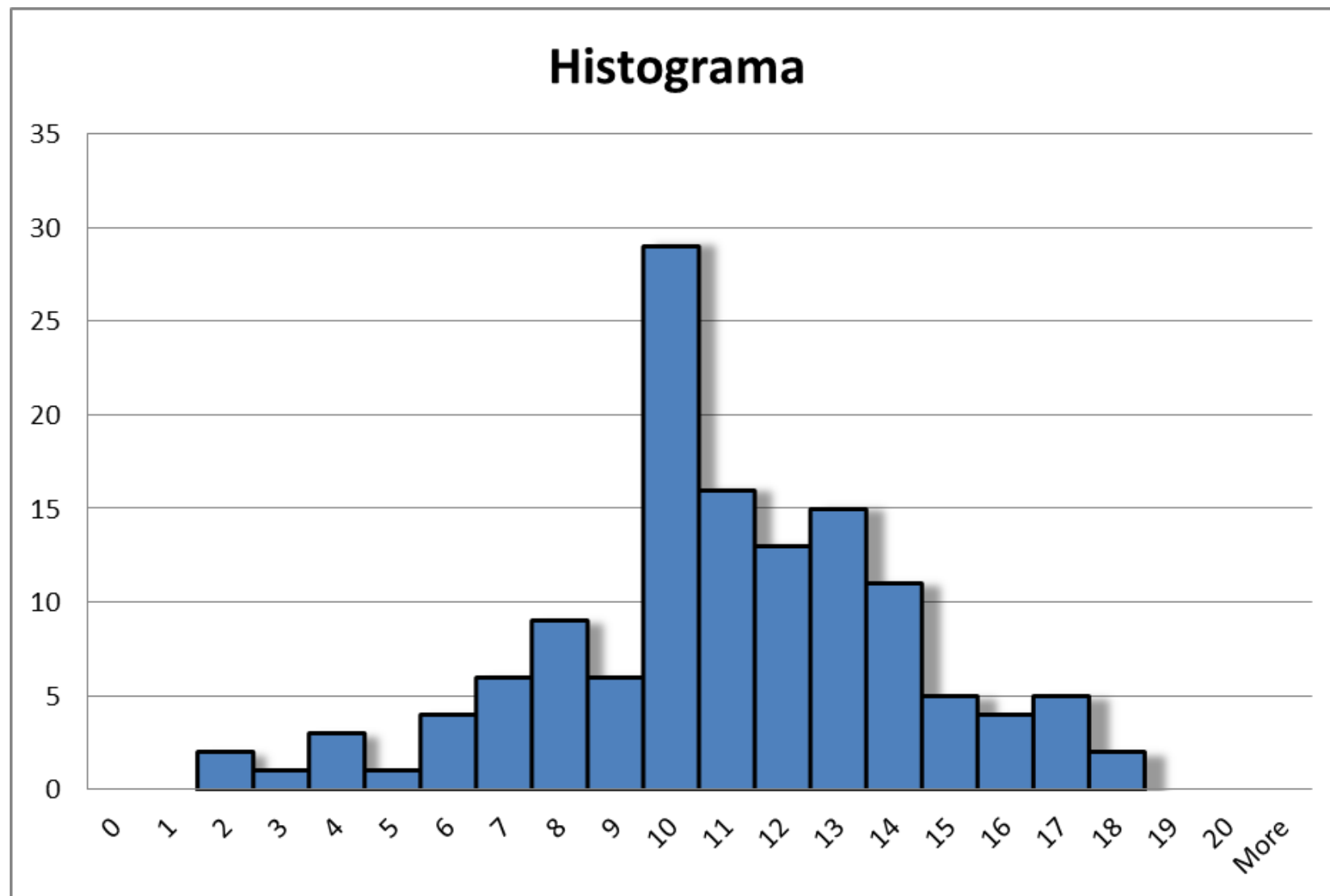


28

Exemplo: notas de Métodos Estatísticos 2004/2005

Aluno	Freq	Final	Aluno	Freq	Final	Aluno	Freq	Final	Aluno	Freq	Final	Aluno	Freq	Final	Aluno	Freq	Final
1			31	7.3	5.7425	61	5.1		91	8.7	9.50511	121	10.2	11.625	151	5.5	1.65
2			32	10.3	10.19511	62	7.4	8.17021	92	9.4	10.5201	122	11.5	13.7501	152	11.4	14.13011
3			33	11.9	10.15025	63	11.0	13.11761	93			123	8.1	10.1476	153	5.9	
4			34	9.8	12.92564	64	8.5	10.985	94	10.4	7.47764	124	9.6	9.60007	154		
5			35	10.0	11.87254	65			95	6.6	10	125	8.3	10.855	155	9.6	9.81
6	9.6	10.50661	36	16.0	16.5145	66	13.0	13.1225	96	8.5	13.275	126			156	9.3	11.50511
7	7.2	7.480175	37			67	12.3	9.67507	97	5.5		127	14.6	16.5075	157	6.3	4.690035
8	5.7	9.620105	38			68	7.3	7.7375	98	6.4		128			158	11.4	12.52011
9			39	13.3	14.875	69	7.0		99			129	7.7	3.64021	159	11.8	11.3275
10	7.7	9.52	40			70	9.1		100	10.0	10.665	130	9.5	8.06504	160	8.3	
11	10.2	11.075	41	9.9	11.0725	71			101	10.9	11.9851	131	7.0	6.61514	161		
12	5.8		42	12.3	14.61011	72			102			132	10.2	13.5075	162	7.3	10
13	15.6	17.665	43	8.7	9.5225	73	10.4	9.560105	103	17.3	16.84	133	12.8	13.0101	163	13.2	13.69
14	12.0	10.79254	44	7.4	15.55	74	8.8	9.51064	104	7.6		134			164	8.2	11.10025
15	13.3	15.75	45			75			105			135			165	6.1	7.01
16	7.8	11.51011	46	3.6	4.580175	76	11.3	13.52961	106	7.7	8.435	136	7.1	6.43507	166		
17	11.6	9.83271	47	7.8	7.065	77	8.7	10.66004	107	10.1	9.785	137	10.9	11.18	167		
18	7.5	7.447675	48	10.4	12.58761	78	10.2	10.5675	108	9.5	13.91	138	8.6	12.625	168		
19			49	12.3	12.28261	79	13.6	17.17	109	10.0	9.72011	139		6.02014	169	10.5	10.95511
20	12.3	9.5105	50	12.2	9.60005	80	6.7	9.517605	110			140	7.5	9.635	170	9.2	6.97764
21			51	14.9	15.915	81	7.5	10.23	111			141	9.1	10.2026	171	9.4	
22	11.2	11.37514	52	9.8	13.2002	82	12.9	16.19	112	7.1	10	142	9.9	12.5251	172	9.0	16.9102
23	4.4	6.325	53	6.1	8.025	83	5.6		113	10.2	12.545	143	6.1	8.13	173	10.1	11.5175
24	10.5	12.74	54	13.7	11.51261	84	9.8	10.78	114	8.0	9.61	144	6.3	8.34771	174	10.3	12.96011
25	7.5	11.175	55	15.0	14.93	85	11.3	9.508	115	14.4	12.6501	145	5.6		175	9.4	9.54
26			56	6.3	9.509605	86			116	12.5	15.8775	146	7.9	8.77507	176	5.4	
27	13.4	14.065	57	7.4		87	12.2	9.52271	117			147	13.9	13.5675	177	13.4	12.01757
28	9.4	8.105035	58	7.4	10	88	4.6		118	8.3	7.18	148			178	9.5	9.780105
29	8.8	11.56511	59	12.5	12.535	89	10.7	10.595	119	10.3	8.62021	149	10.5	9.78292	179	10.8	13.705
30	9.4	12.95261	60	7.3	3.74757	90	7.2	10	120			150	14.5	13.5376	180		
															181	8.8	9.535
															182	5.5	2.94514
															183	8.1	

N	132
Média	10.52
Mediana	10.51
Amplitude	16.29
Máximo	17.67
Mínimo	1.38
Quartil 25%	9.07
Quartil 75%	12.68
Desvio Padrão	3.21
Variância	10.29
Assimetria	-0.25

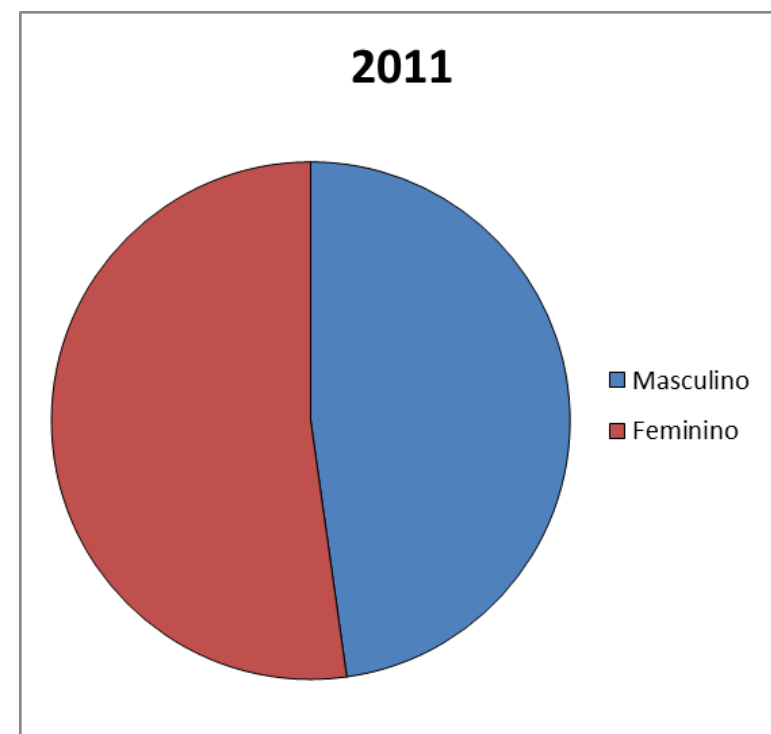
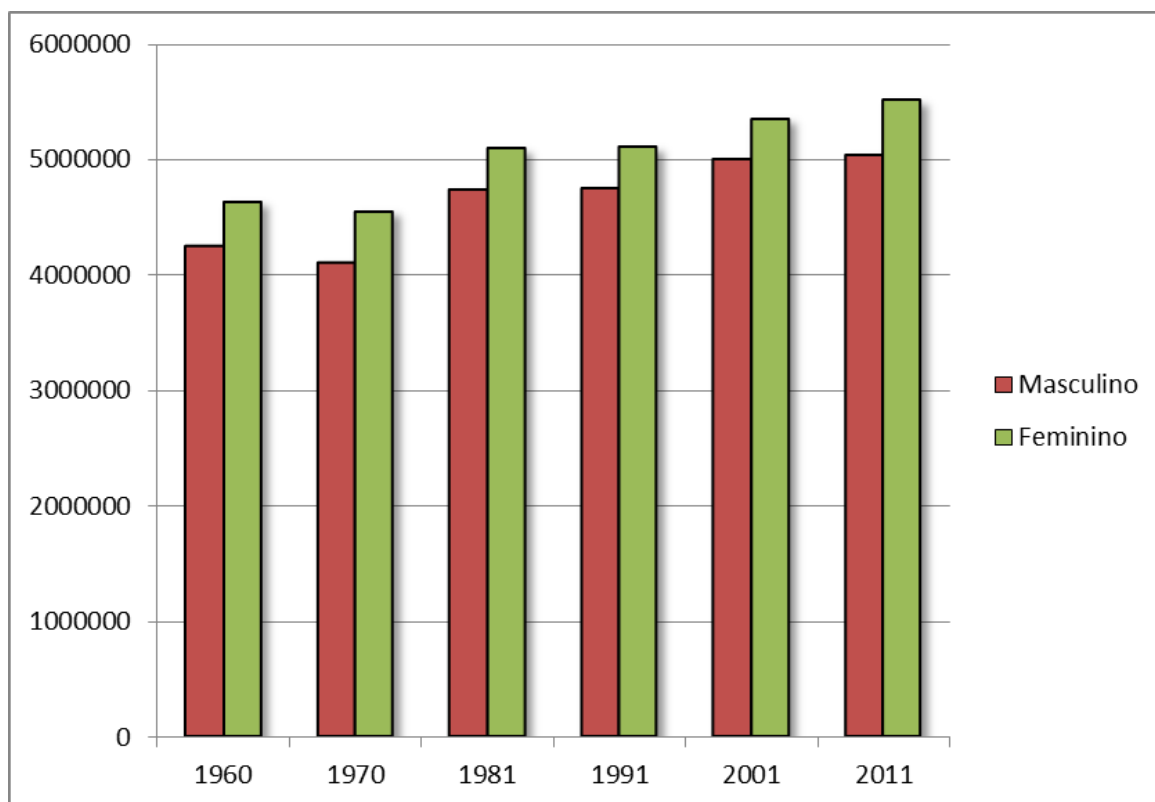


regra indicativa

<http://www.amstat.org/publications/jse/v6n3/applets/histogram.html>

Nº de células (K) $\approx \sqrt{\text{Nº de dados (N)}}$

População residente segundo os Censos: total e por sexo (dados do pordata.pt)

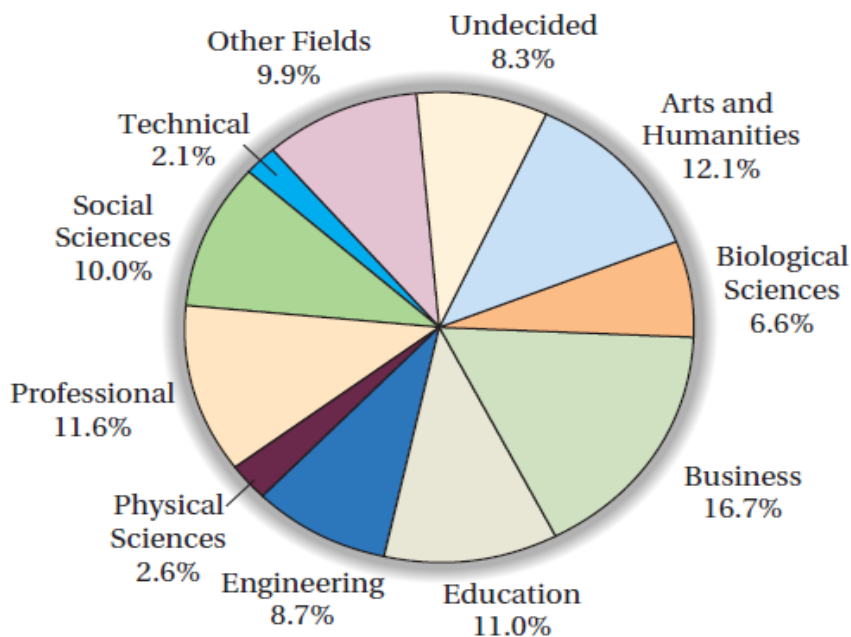


Em geral, é mais fácil comparar visualmente as alturas do que os ângulos.

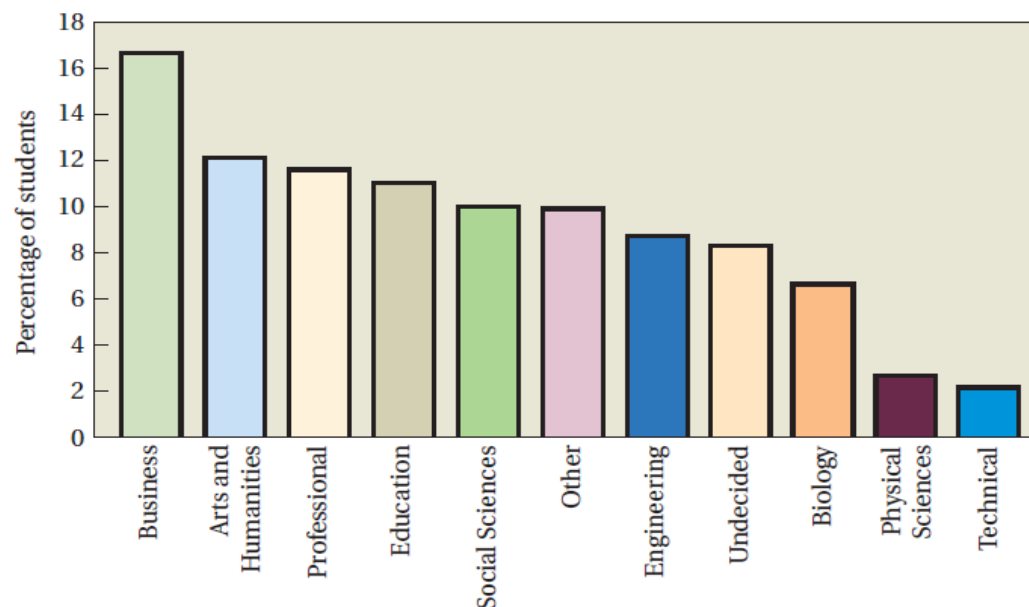
Carbon Dioxide Emissions

- The pie chart shows planned major areas for first-year college students.
- Example from: Using and Understanding Mathematics: A Quantitative Reasoning Approach, Jeffrey Bennett and ; 2004*

What Students Expect to Major In



What Students Expect to Major In



Qual dos dois é mais fácil de analisar?

Média e Mediana

Exemplo

x

10

12

14

11

7

14

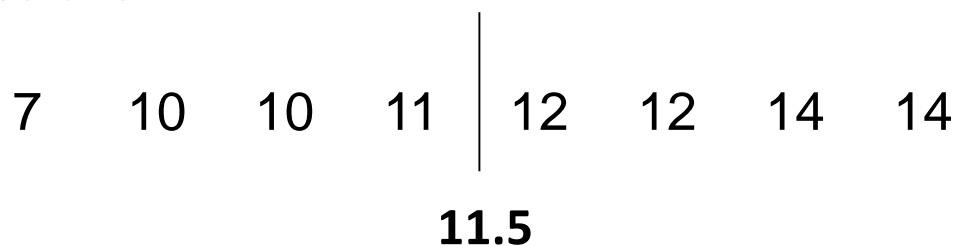
10

12

Média:

$$\bar{x} = (10 + 12 + 14 + 11 + 7 + 14 + 10 + 2) / 8 = 11.25$$

Mediana:



Média e Mediana

Exemplo

x

10

12

14

11

7

200

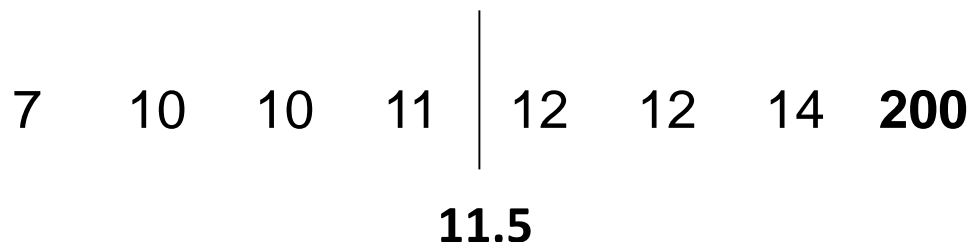
10

12

Média:

$$\bar{x} = (10 + 12 + 14 + 11 + 7 + 200 + 10 + 2) / 8 = 34.5$$

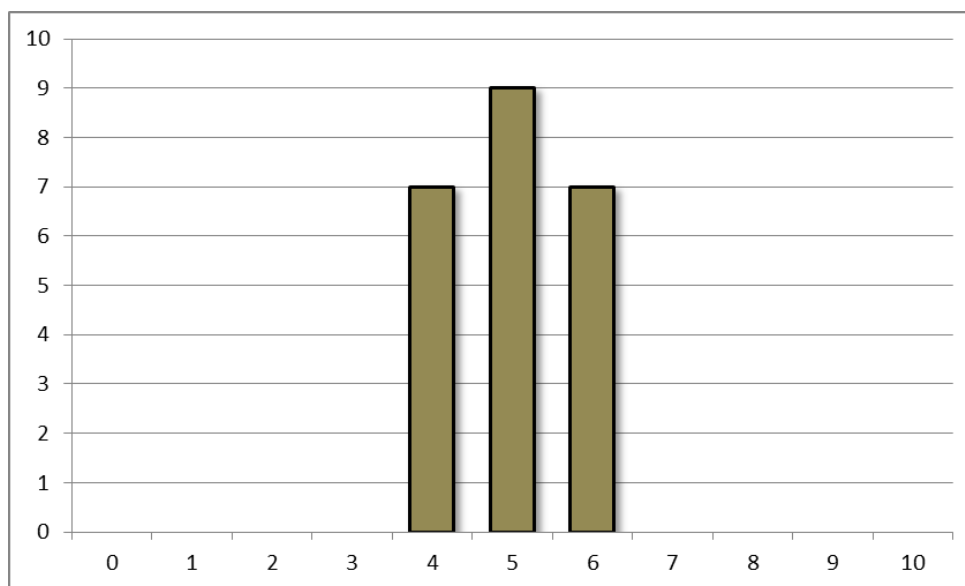
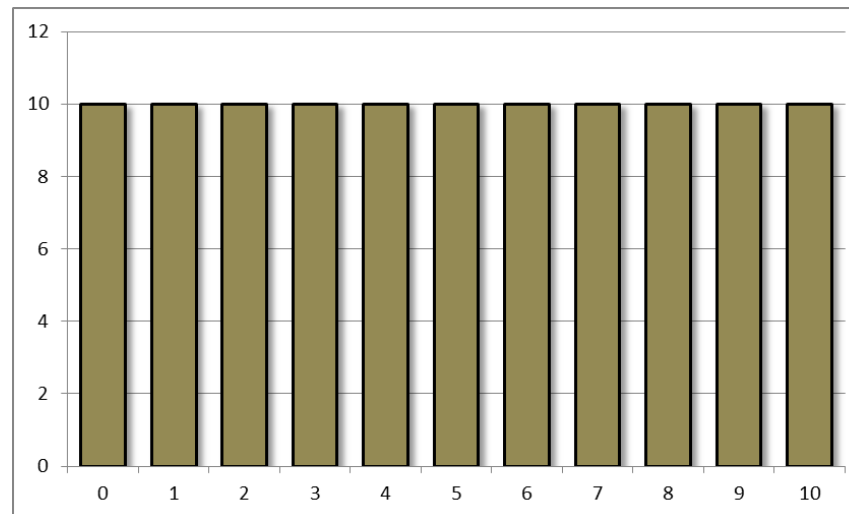
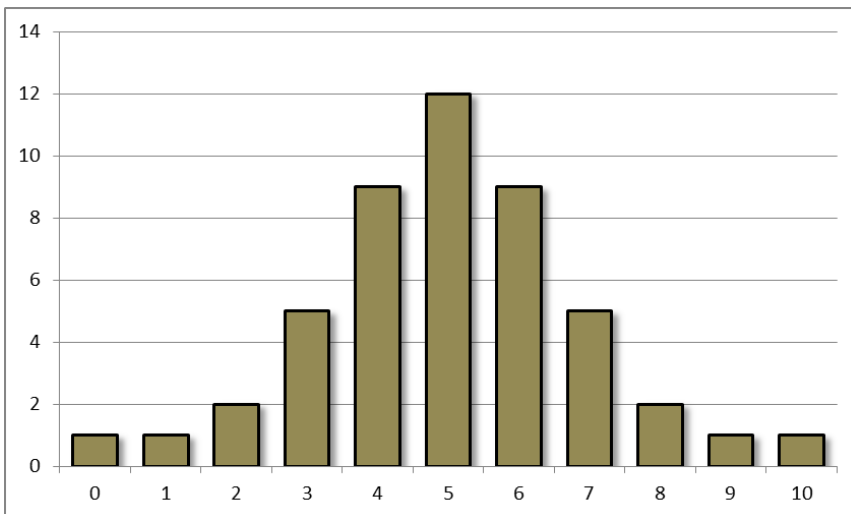
Mediana:



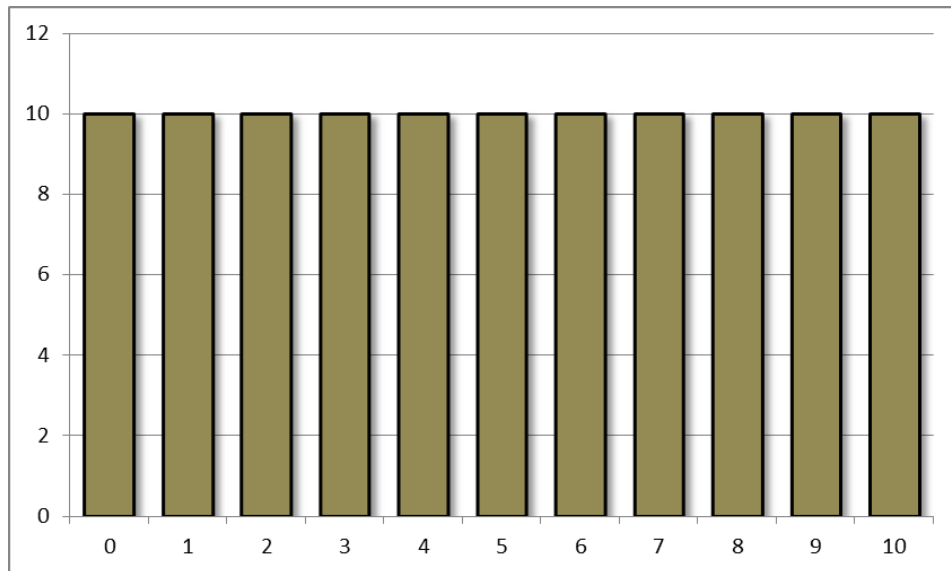
Média mais sensível a valores extremos!

ex. Salário médio vs. Salário mediano

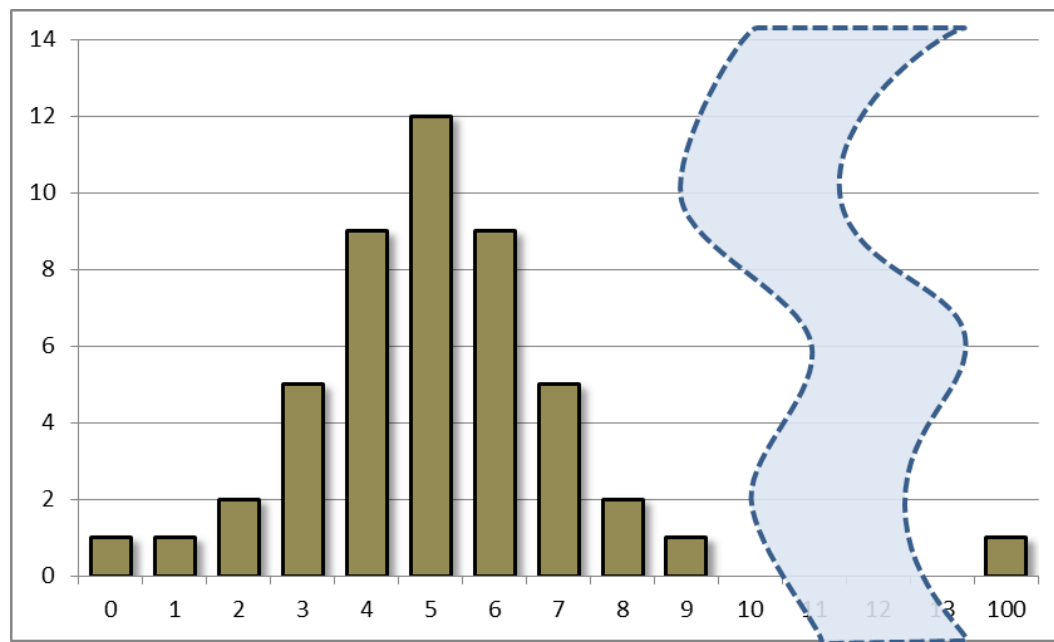
Which dataset has higher average?



How does the average compares?



and the median?



Variância e Desvio Padrão

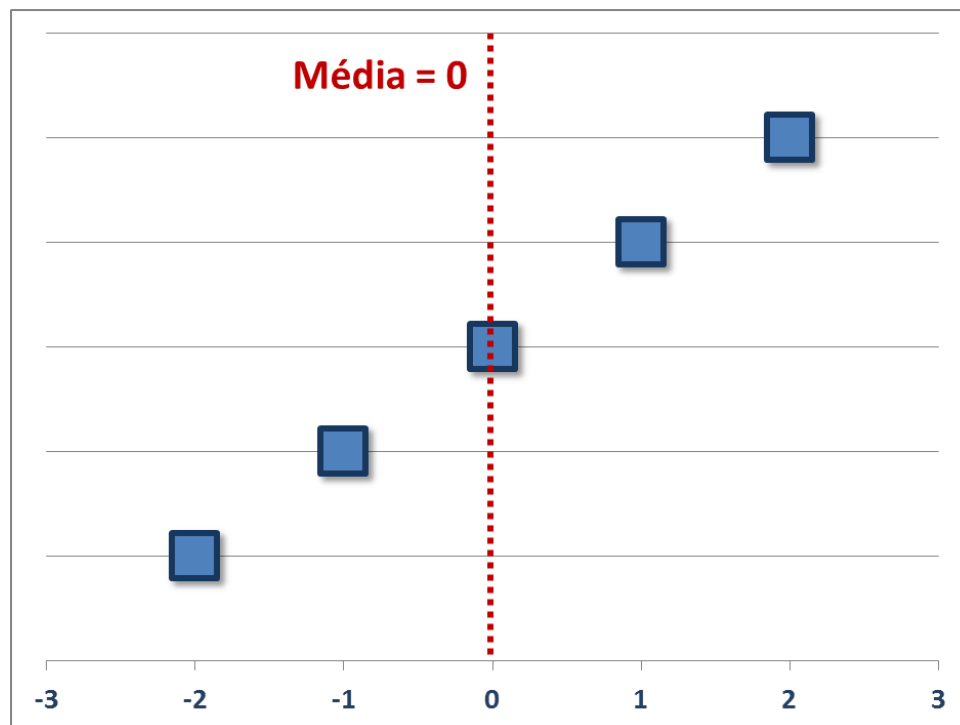
Para inferir da **variabilidade** de uma população a partir de uma amostra usa-se a **variância amostral** (s^2)

$$s^2 = \frac{1}{N-1} \cdot \sum_{n=1}^N (x_n - \bar{x})^2$$

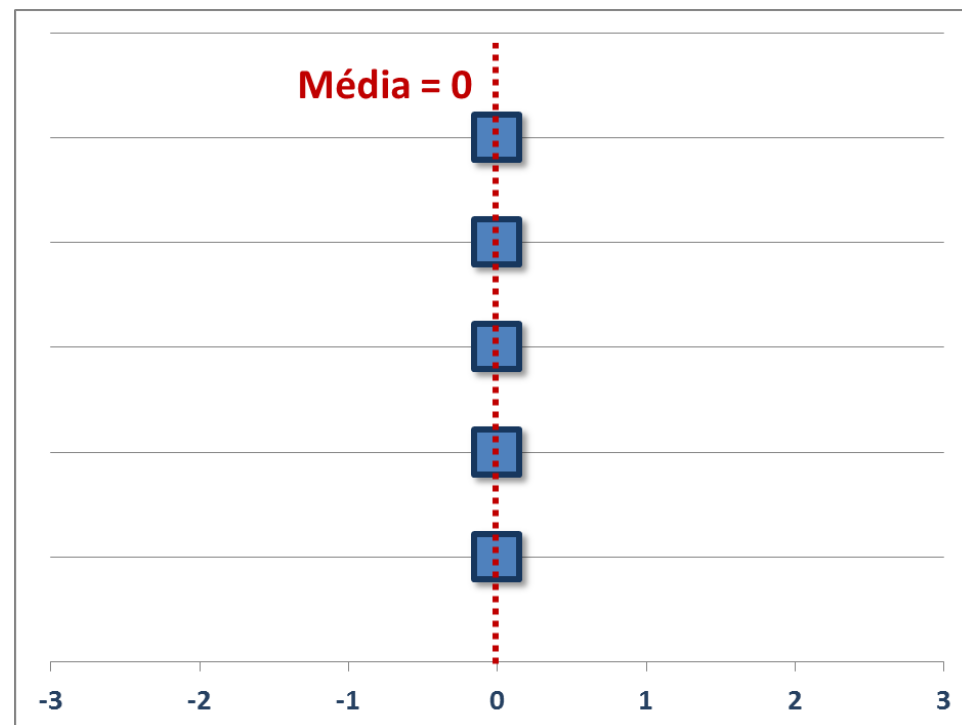
O **desvio padrão amostral** (s), raiz quadrada da variância amostral, tem a vantagem de ser **expresso nas mesmas unidades dos dados**

$$s = \sqrt{\frac{1}{N-1} \cdot \sum_{n=1}^N (x_n - \bar{x})^2}$$

dois conjuntos de valores com a mesma média.
como podemos descrever a diferença entre eles?

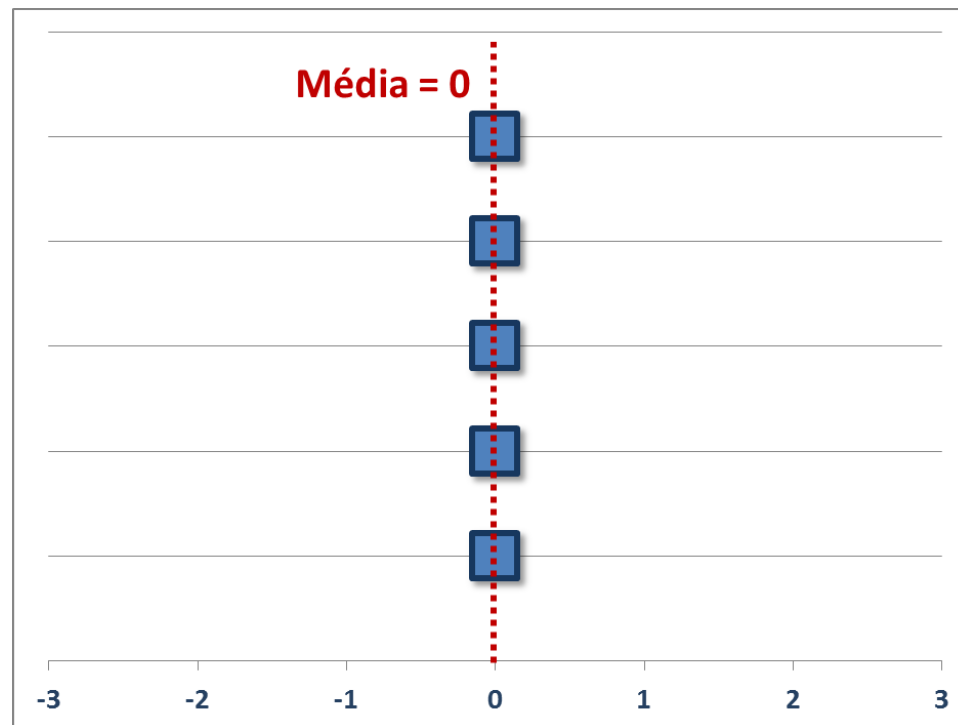
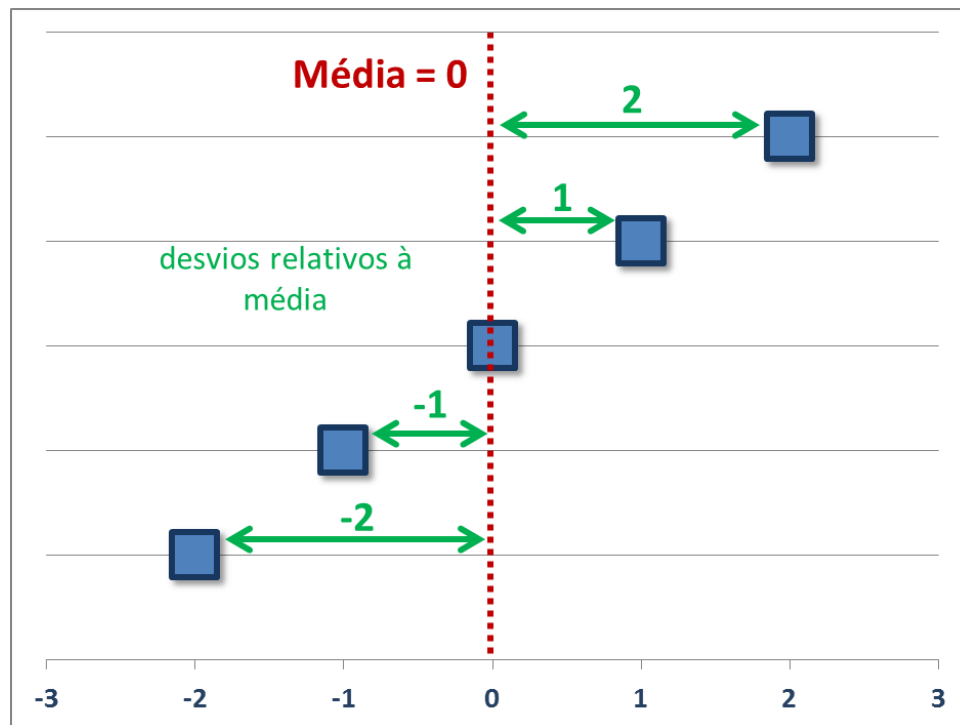


$(-2, -1, 0, 1, 2)$



$(0, 0, 0, 0, 0, 0)$

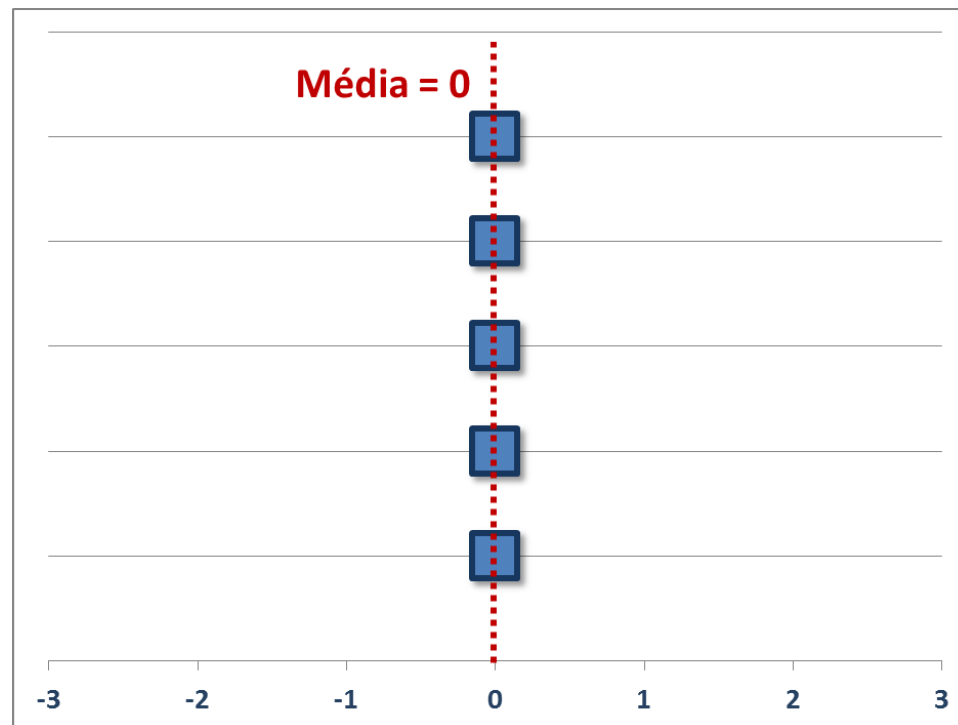
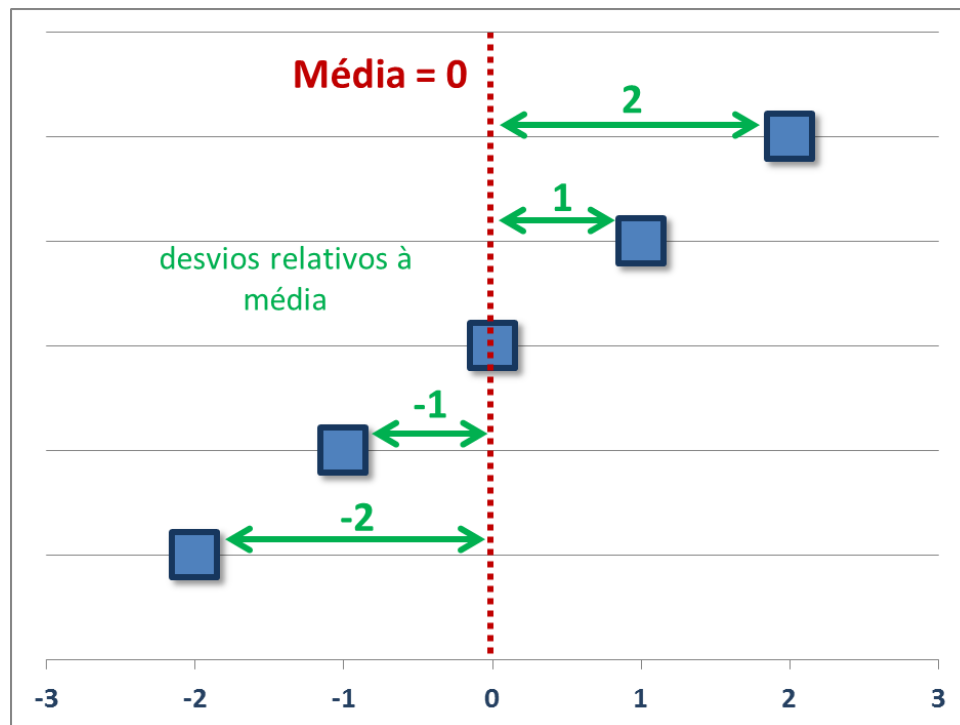
podemos calcular os desvios relativos à média.



O que podemos fazer com os desvios?

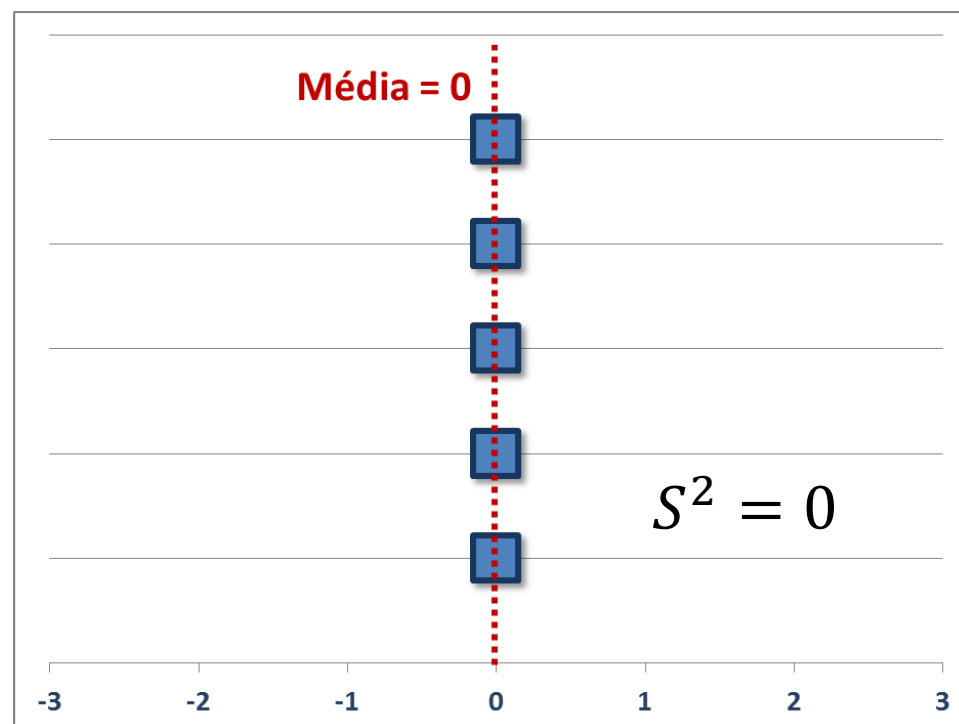
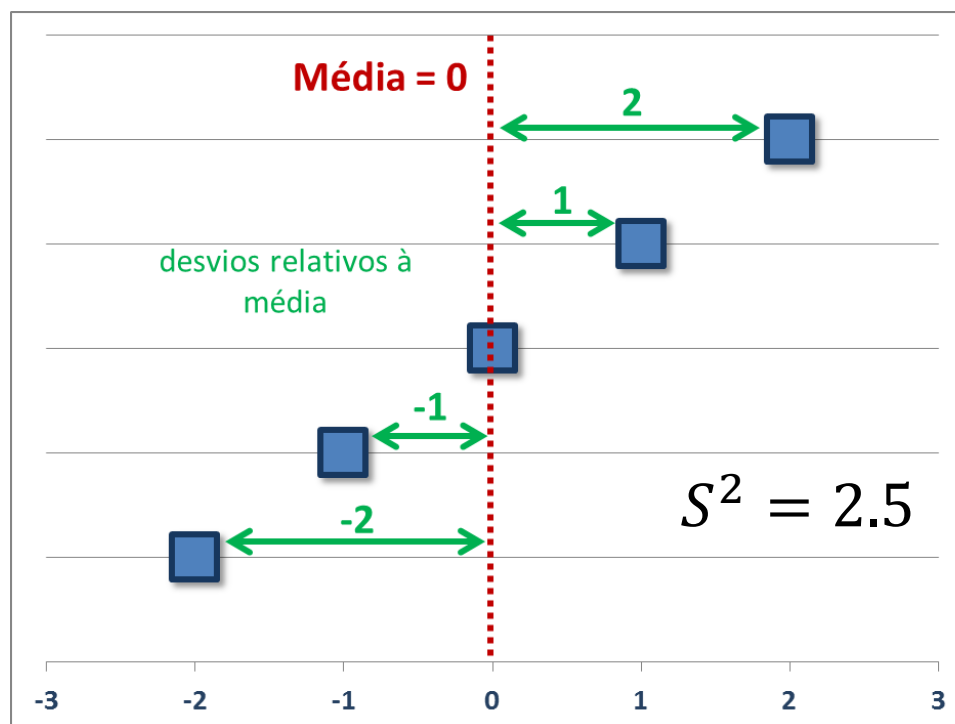
- Somar?
- Calcular a média?
- Calcular a média dos desvios absolutos (módulo)?
- Calcular a média dos desvios elevados ao quadrado?

podemos calcular os desvios relativos à média.



O que podemos fazer com os desvios?

- Somar? - **valor da soma ficaria a depender do número de elementos na amostra**
- Calcular a média? - **desvios negativos anulam os positivos**
- Calcular a média dos desvios absolutos (módulo)? - **ok. chama-se desvio absoluto médio**
- Calcular a média dos desvios elevados ao quadrado? - **ok. chama-se variância**



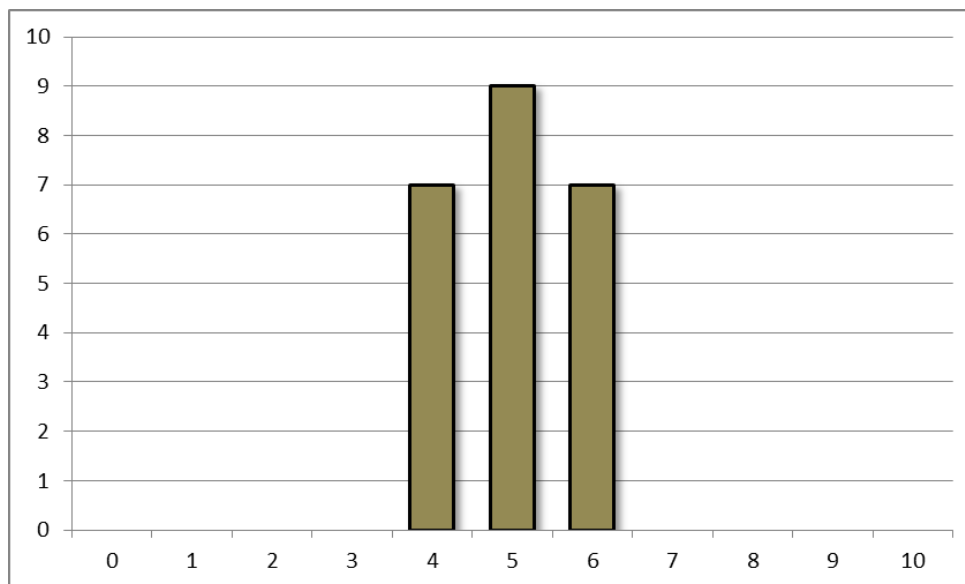
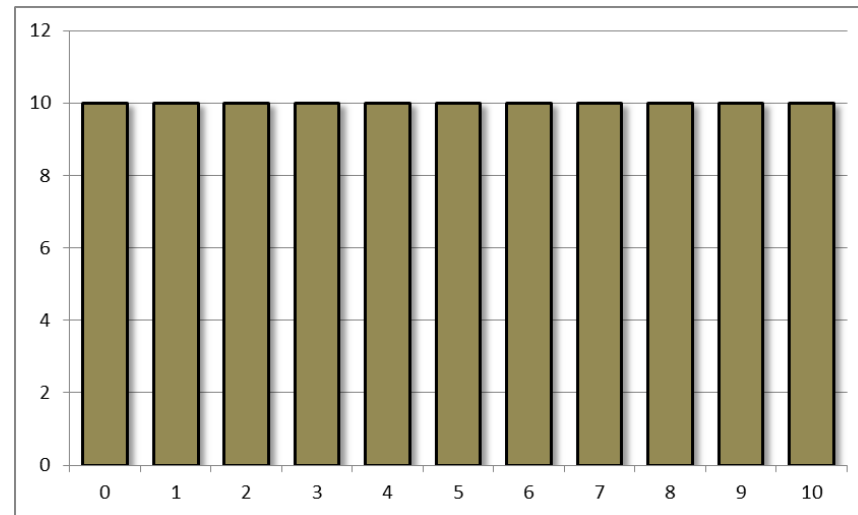
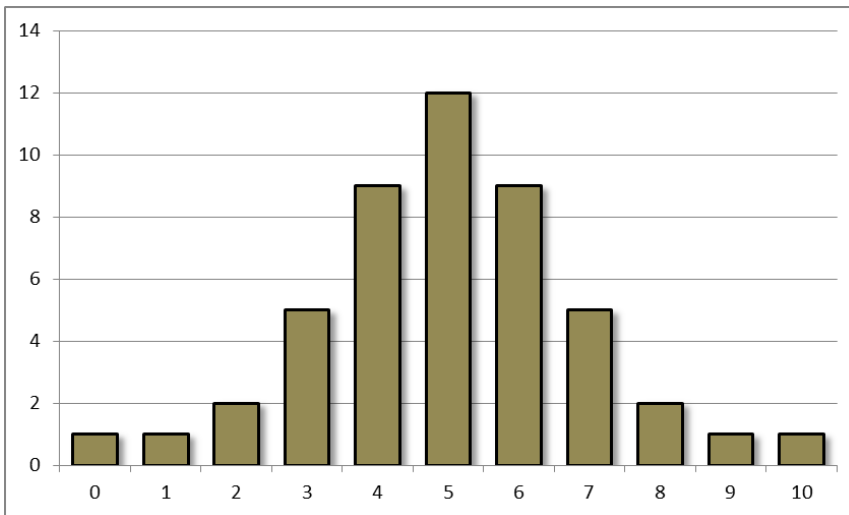
Variância $S^2 = (2^2 + 1^2 + (-1)^2 + (-2)^2) / (5 - 1) = 2.5$

se estivermos a medir metros a variância representa o desvio em m^2

para ter uma medida do desvio em m calculamos o desvio padrão

Desvio Padrão $S = \sqrt{S^2} = \sqrt{2.5} = 1.581$

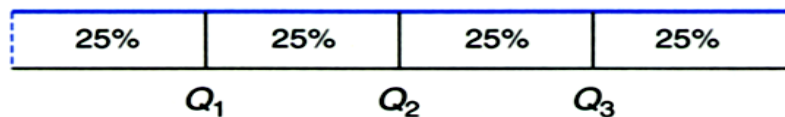
Which dataset has higher variability?



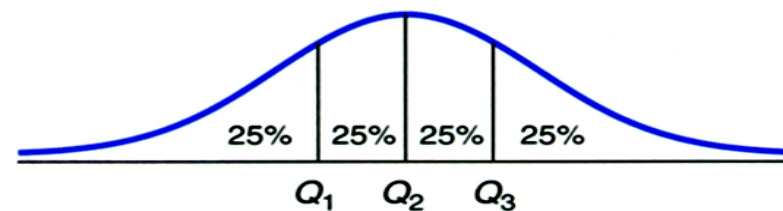
Coeficiente de assimetria (g_1)

$$g_1 = \frac{k_3}{s^3} \quad , \quad \text{COM} \quad k_3 = \frac{N^2}{(N-1) \cdot (N-2)} \cdot \left(\frac{1}{N} \cdot \sum_{n=1}^N (x_n - \bar{x})^3 \right)$$

mede desvios para a média ao cubo

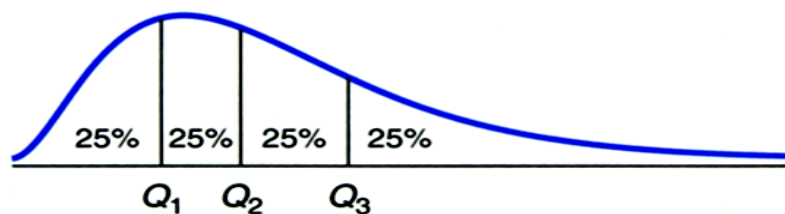


(a) Uniform



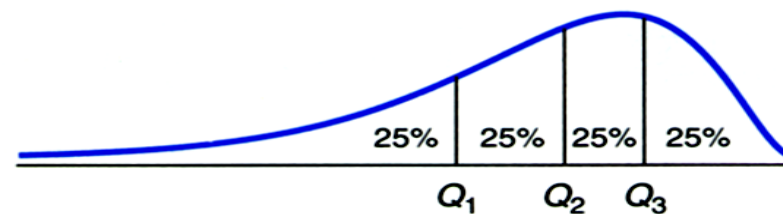
(b) Bell-shaped

$g_1 = 0$



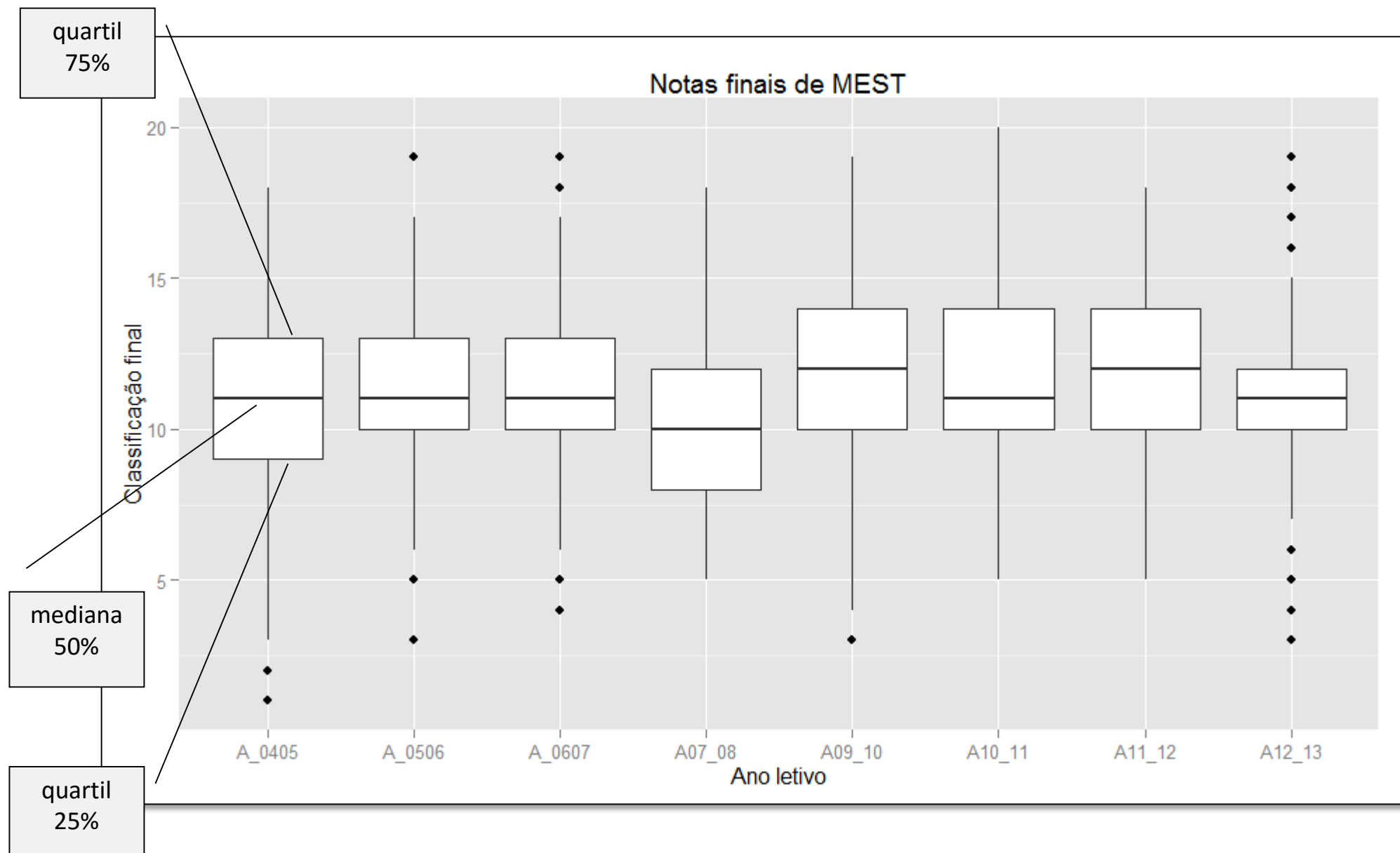
(c) Right skewed

$g_1 > 0$



(d) Left skewed

$g_1 < 0$



Análise detalhada das classificações de MEST em 04/05

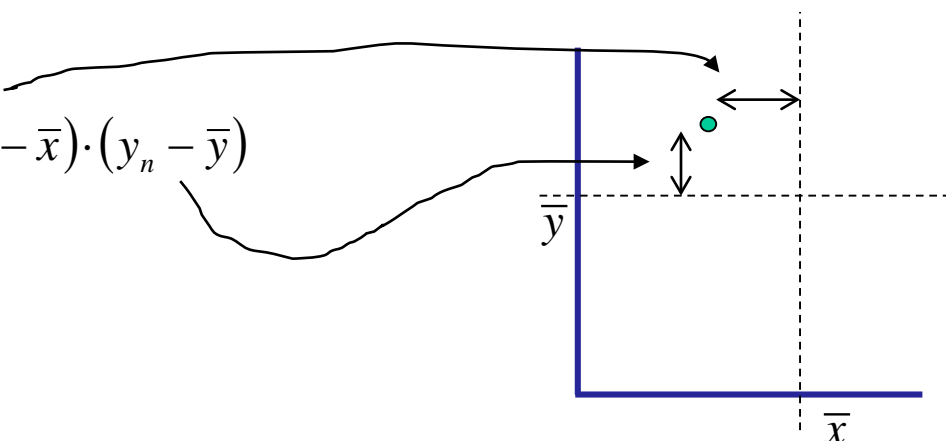
	2004/2005					
Nota Frequência	Alunos	Aprovados		Reprovados		S.E.
≥ 9.5	72	66	91.67%	6	8.33%	0 0.00%
≥ 8.5 e < 9.5	19	15	78.95%	2	10.53%	2 10.53%
≥ 7.5 e < 8.5	18	7	38.89%	8	44.44%	3 16.67%
< 7.5	36	4	11.11%	22	61.11%	10 27.78%
S.F.	38	0	0.00%	1	2.63%	37 97.37%
Total	183	92	50.27%	39	21.31%	52 28.42%

Conclusão:

estudar para os mini-testes é uma boa estratégia.

Measures the degree of adjustment of a linear relationship:

Sample covariance (infer about the population)

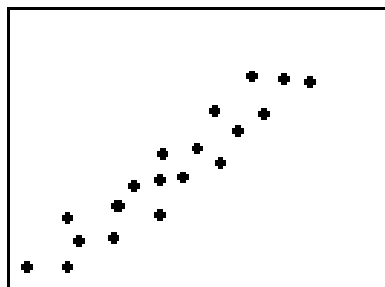
$$c_{XY} = \frac{1}{N-1} \cdot \sum_{n=1}^N (x_n - \bar{x}) \cdot (y_n - \bar{y})$$


Sample correlation coefficient (adimensional measure)

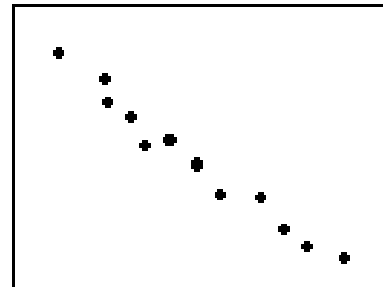
$$r_{XY} = \frac{c_{XY}}{s_X \cdot s_Y} = \frac{Cov\ xy}{DPx \cdot DPy} \quad (-1 \leq r_{XY} \leq 1)$$

$$r_{XY} = \frac{\frac{1}{N-1} \cdot \sum_{n=1}^N (x_n - \bar{x}) \cdot (y_n - \bar{y})}{\sqrt{\frac{1}{N-1} \cdot \sum_{n=1}^N (x_n - \bar{x})^2} \cdot \sqrt{\frac{1}{N-1} \cdot \sum_{n=1}^N (y_n - \bar{y})^2}} \quad (-1 \leq r_{XY} \leq 1)$$

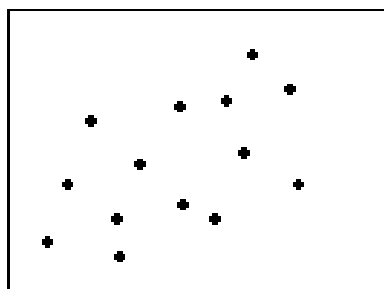
Degree of Correlation



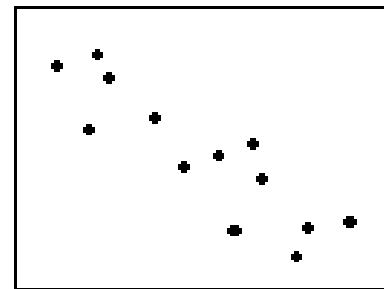
Strong Positive



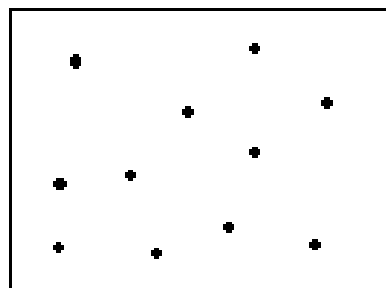
Strong Negative



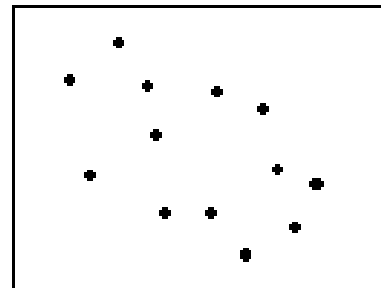
Weak Positive



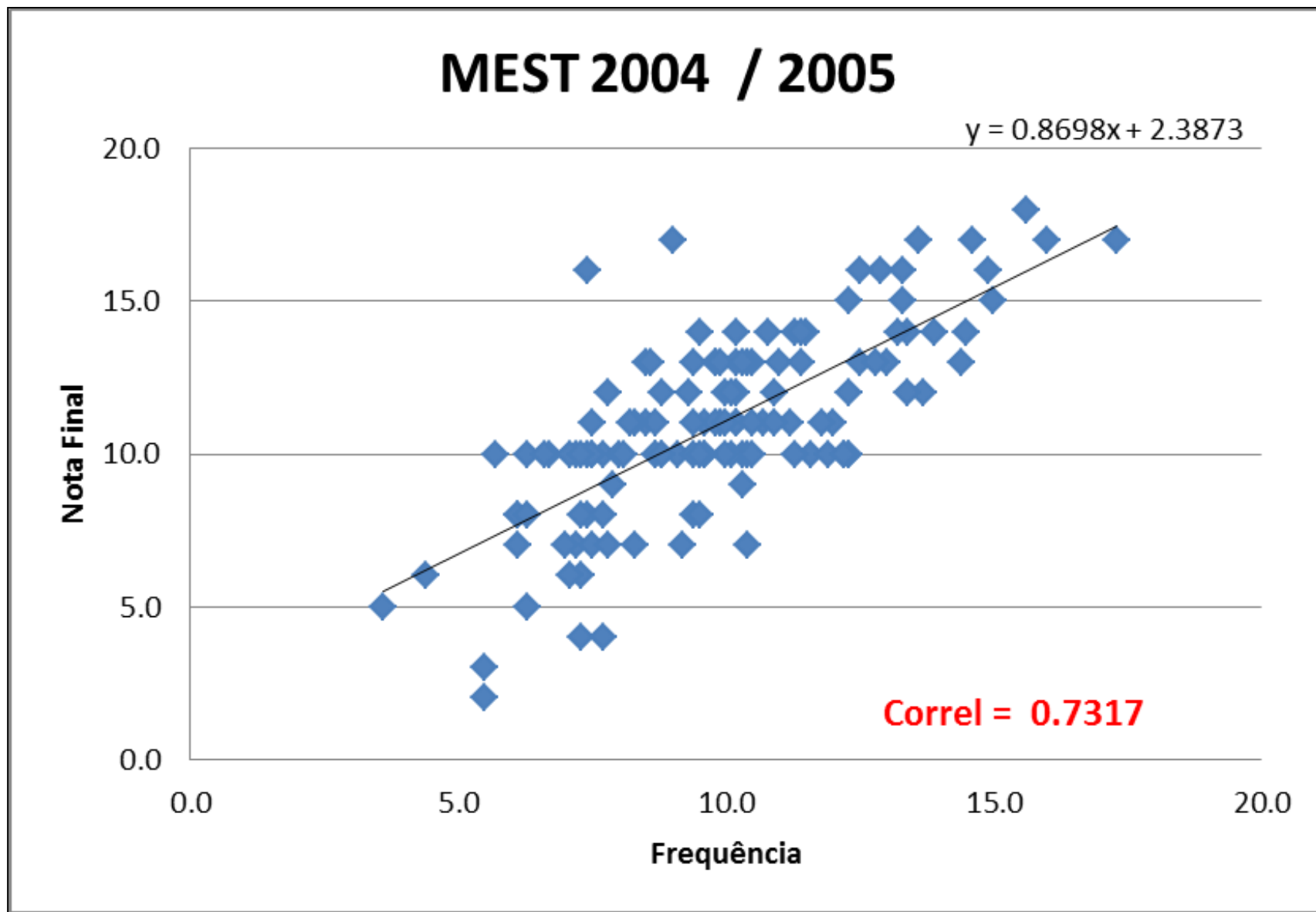
Moderate Negative



None



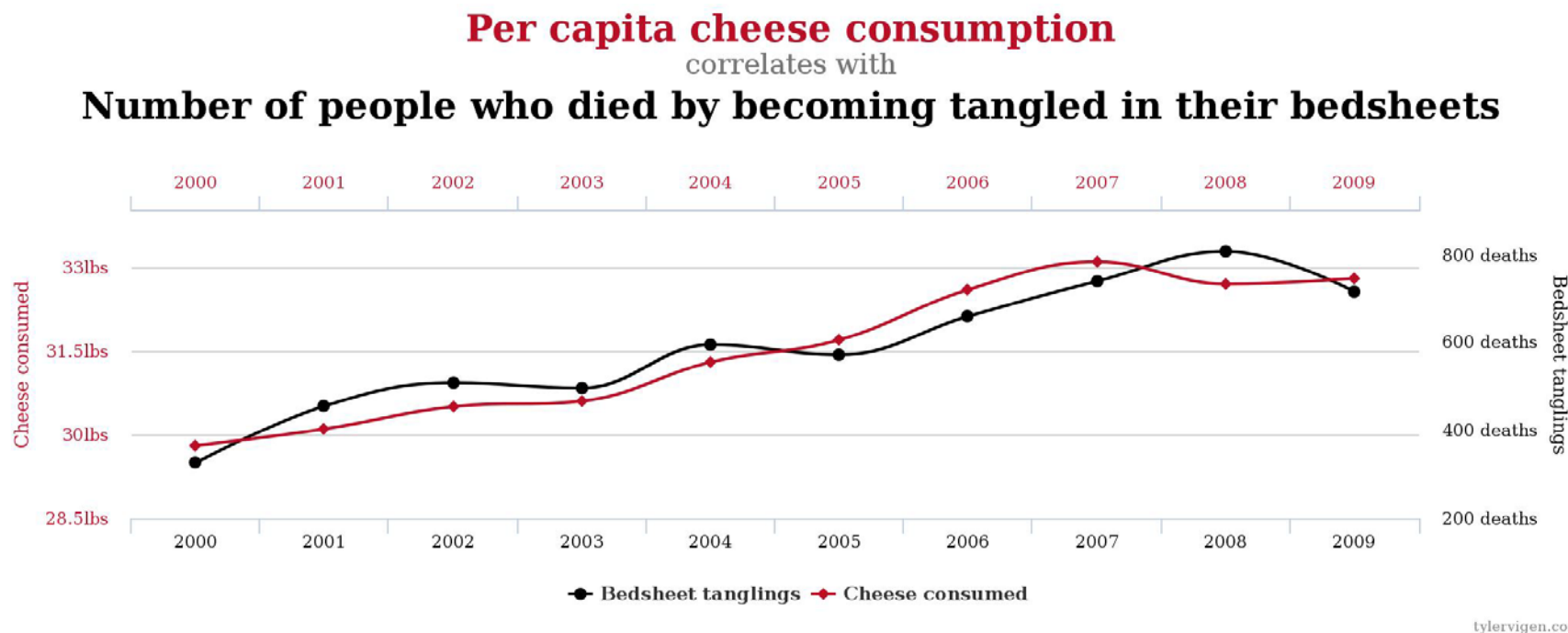
Weak Negative



Quem tem boa nota a frequência tende a ter boa nota no final.

Uma correlação alta não implica uma relação de causa-efeito entre as variáveis!!!

Spurious correlations



Golos campeonato 2012/2013

jornada	A	B	C	D
1	1	2	0	2
2	1	3	4	5
3	2	0	3	3
4	3	4	4	2
5	1	2	2	2
6	0	4	2	2
7	2	2	2	3
8	0	3	5	3
9	1	0	2	1
10	2	0	2	2
11	0	4	1	3
12	2	3	1	4
13	1	1	2	3
14	1	2	2	2
15	0	4	3	2
16	5	1	5	2
17	1	3	4	3
18	3	2	1	2
19	1	1	2	1
20	0	3	2	3
21	2	1	0	1
22	1	2	2	5
23	2	3	1	4
24	3	2	3	6
25	1	1	3	2
26	2	1	3	2
27	2	1	2	2
28	1	3	3	1
29	1	1	2	1
30	0	1	2	3

Class. Final	Equipa
1	FC Porto
2	Benfica
3	P. Ferreira
4	SC Braga

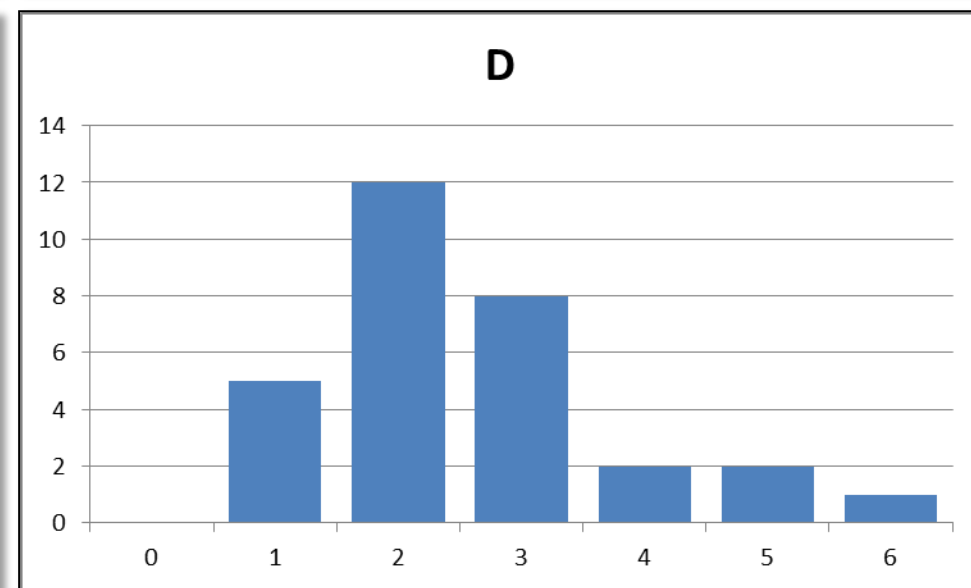
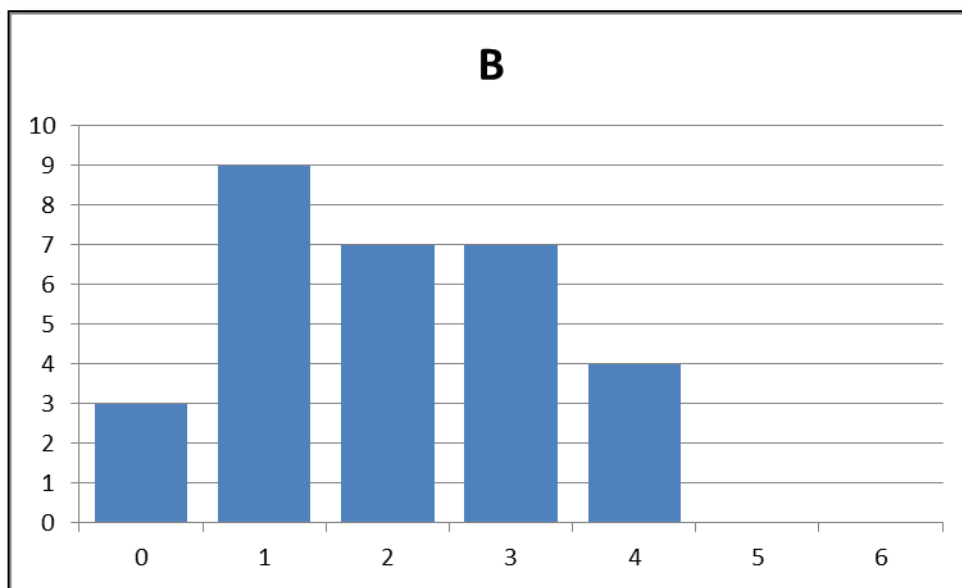
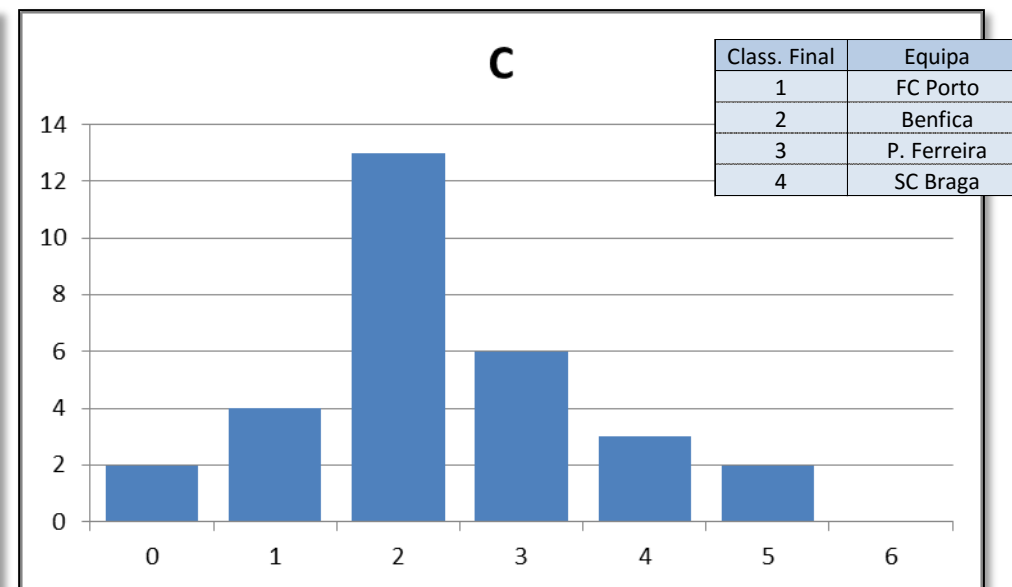
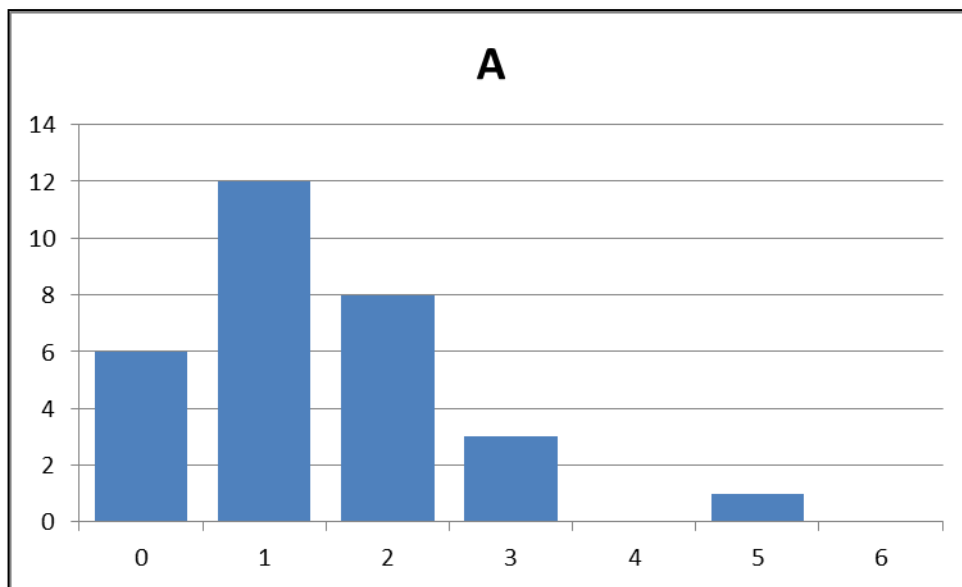
**Que equipa
corresponde ao
A, B, C e D?**

Golos marcados - Medidas de localização e de dispersão

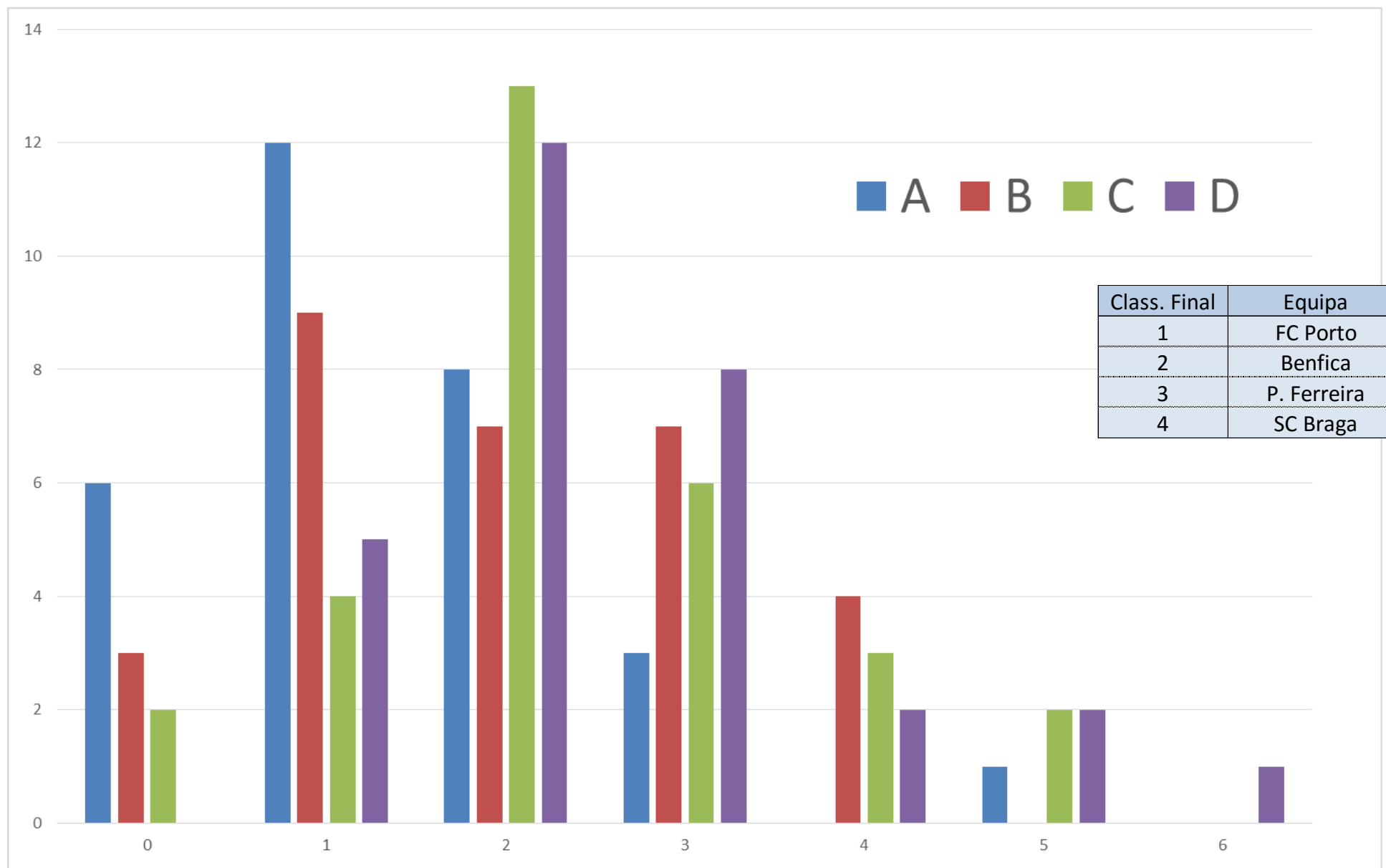
	A	B	C	D
n	30	30	30	30
total	42	60	70	77
mínimo	0	0	0	1
máximo	5	4	5	6
média	1,40	2,00	2,33	2,57
mediana	1	2	2	2
moda	1	1	2	2
amplitude	5	4	5	5
quartil 25%	1,00	1,00	2,00	2,00
quartil 75%	2,00	3,00	3,00	3,00
desvio padrão	1,13	1,23	1,24	1,25
variância	1,28	1,52	1,54	1,56
coef. assimetria	1,10	0,12	0,35	1,03

Class. Final	Equipa
1	FC Porto
2	Benfica
3	P. Ferreira
4	SC Braga

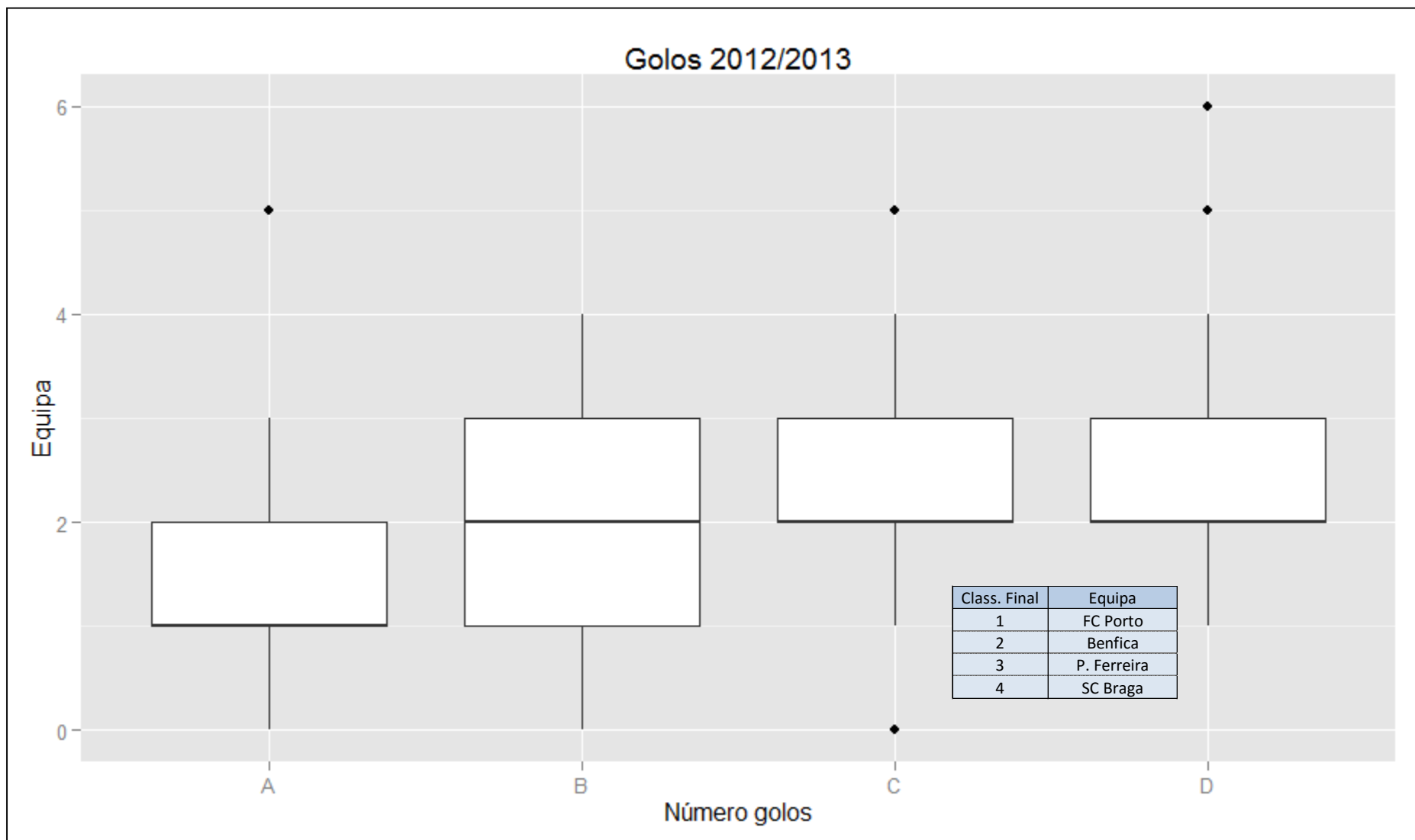
Golos marcados - gráficos de barras independentes



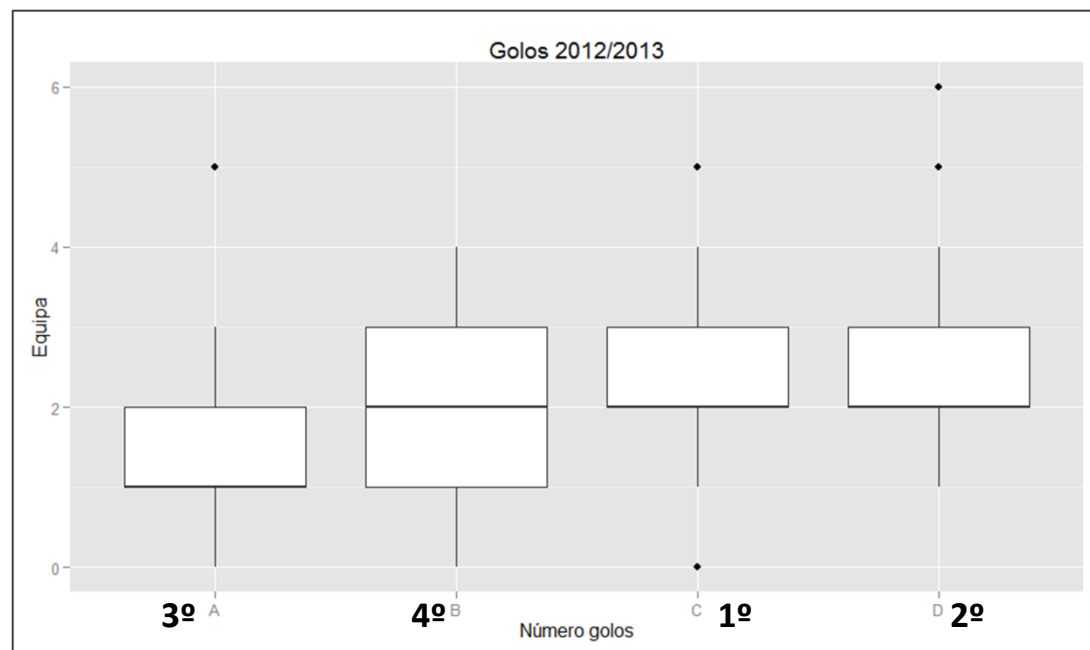
Golos marcados - um gráfico de barras



Golos marcados - box plot



CF	Equipa	P	J	V	E	D	GM	
1	FC Porto	78	30	24	6	0	70	C
2	Benfica	77	30	24	5	1	77	D
3	P. Ferreira	54	30	14	12	4	42	A
4	SC Braga	52	30	16	4	10	60	B

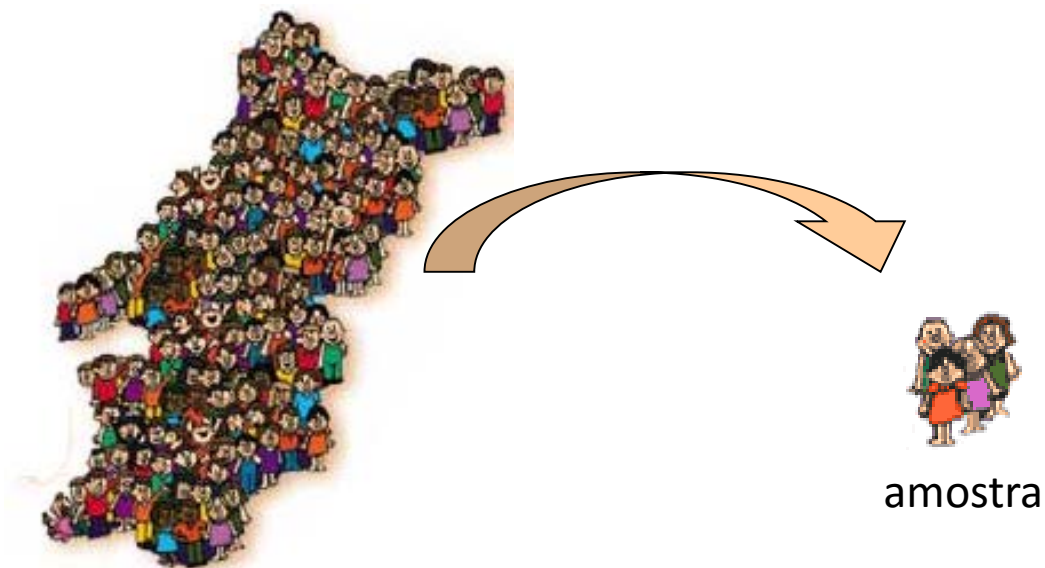


Atenção às conclusões que se tiram dos dados!
Número de golos não significa pontos!

População e Amostra

Designa-se por **população** (ou **universo**) o conjunto dos dados que expressam a característica em causa para todos os objetos sobre os quais a análise incide.

Uma **amostra** corresponde a um subconjunto de dados que pertencem à população.



Seleção de amostras

Quando todos os elementos da população têm igual probabilidade de ser incluídos na amostra evita-se qualquer enviesamento de seleção, designando-se então o processo por

amostragem aleatória

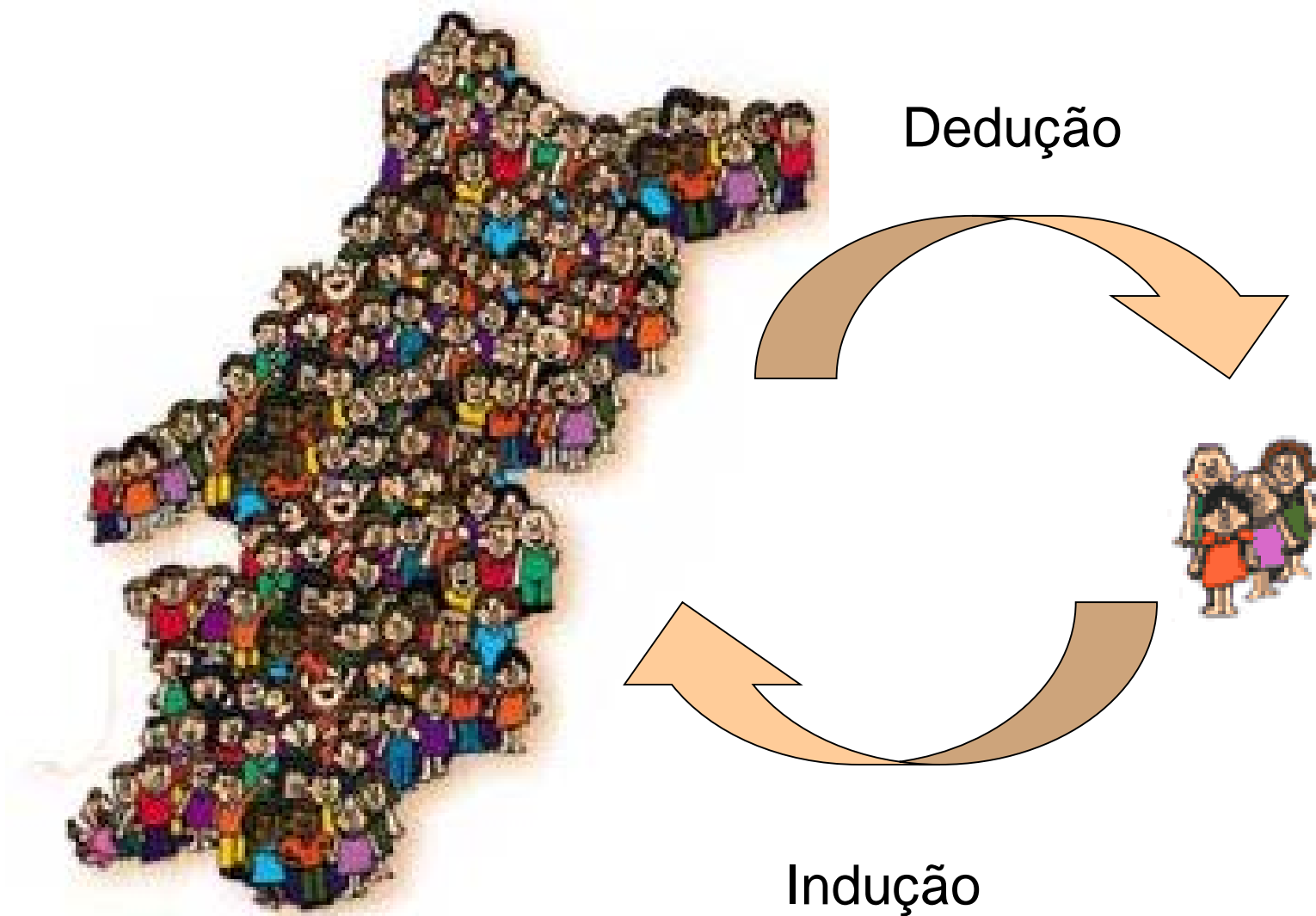
The 1936 election: the literary digest poll

- Candidates:
 - Democrat FD **Roosevelt** and
 - Republican Alfred **Landon**
- Prediction: Landon to win with 57% of the vote
- Outcome: Landon lost with only 38% of the vote
- Sample Size: 2.3 million people
- Literary Digest went bankrupt soon after

Why the Digest went wrong:

- **Bias in selection of sample**
 - 10,000,000 questionnaires sent out to
 - Magazine subscribers, car owners, telephone owners
- **Bias from non-response**
 - 20% bothered to reply
 - Presumably, those with strong views about the forthcoming election

Large sample size cannot compensate
for poor sample design!!!



DN e TSF -> Marktest



819 entrevistas e apresenta um erro de amostragem para um intervalo de confiança de 95 por cento, de mais ou menos 3,42 por cento.

Independente -> Instituto de Pesquisa de Opinião e Mercado (IPOM)



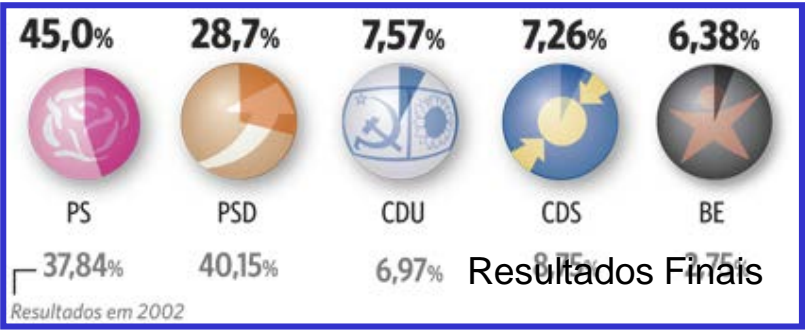
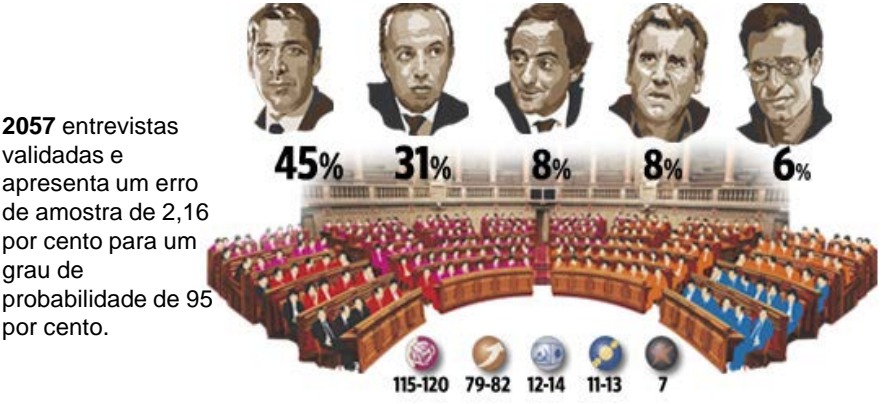
997 entrevistas validadas e apresenta um erro de amostragem, para um nível de confiança de 95,5 por cento, de mais ou menos 3,1 pontos percentuais.

JN -> Intercampus



1015 entrevistas, e apresenta um erro de amostragem, para um intervalo de confiança de 95 por cento, de mais ou menos 3,1 por cento.

EXPRESSO-SIC-Renascença -> Eurosondagem



- PS: 46% (118-124 deputados)
- PSD: 31% (80-84)
- CDU: 7% (8-12)
- BE: 7% (8-12)
- CDS-PP: 6% (6-10)
- Outros: 1% (0)
- Branco/nulos: 2%



5051 inquiridos, de 1,4 por cento com um nível de confiança de 95 por cento.

PÚBLICO, RTP e Antena 1 -> Universidade Católica

Resultados de Aprendizagem

- Porque tenho de aprender Estatística
- Calcular e interpretar estatísticas (Moda, Média, Mediana, Desvio padrão, Variância, coeficiente de assimetria e coeficiente de correlação)
- Construir e interpretar tabelas (frequências absolutas, frequências relativas e frequências acumuladas)
- Construir e interpretar gráficos (gráficos de barras, histogramas, gráficos circulares, gráficos de caixa 'box-plots' e scatter plots)