

Tarefa 3 | Ver antes de analisar

Objetivos

- Reconhecer a importância da representação dos dados no diagrama de dispersão.

O Quarteto de Anscombe é um grupo de quatro conjuntos de dados, construídos em 1973 pelo estatístico Francis Anscombe, com o objetivo de demonstrar a importância de se visualizar os dados antes de analisá-los. Os valores de x são comuns para os três primeiros conjuntos.

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89



Francis Anscombe
1918–2001

3.1) Insira os dados dos conjuntos I e II no menu  e obtenha as informações necessárias ao preenchimento da tabela seguinte:

Grupo	\bar{x}	\bar{y}	\tilde{x}	\tilde{y}	σ_x	σ_y
I						
II						

- 3.2) Obtenha o coeficiente de correlação e a equação da reta de regressão linear para cada um dos conjuntos I e II e compare os valores obtidos.
- 3.3) Obtenha a nuvem de pontos para cada uma dos conjuntos I e II. Visualize-as em simultâneo.
- 3.4) Em qual das situações acha que não deve utilizar a reta de regressão para prever o valor de y para $x = 13,5$?
- 3.5) Insira os dados do conjunto III em duas listas, obtenha a nuvem de pontos e a reta de regressão linear. Que dado poderá ser excluído do conjunto de modo a melhorar a regressão obtida? Confirme através da calculadora.
- 3.6) Compare os resultados obtidos nas alíneas anteriores. Que conclusões pode tirar sobre a importância da visualização dos dados relativamente à análise das medidas estatísticas por si só?