

LE QUARTET D'ANSCOMBE

Preuve par la démonstration de l'utilité de la visualisation dans l'analyse de données

I		II		III		IV	
x	y	x	y	x	y	x	y
10,00	8,04	10,00	9,14	10,00	7,46	8,00	6,58
8,00	6,95	8,00	8,14	8,00	6,77	8,00	5,76
13,00	7,58	13,00	8,74	13,00	12,74	8,00	7,71
9,00	8,81	9,00	8,77	9,00	7,11	8,00	8,84
11,00	8,33	11,00	9,26	11,00	7,81	8,00	8,47
14,00	9,96	14,00	8,10	14,00	8,84	8,00	7,04
6,00	7,24	6,00	6,13	6,00	6,08	8,00	5,25
4,00	4,26	4,00	3,10	4,00	5,39	19,00	12,50
12,00	10,84	12,00	9,13	12,00	8,15	8,00	5,56
7,00	4,82	7,00	7,26	7,00	6,42	8,00	7,91
5,00	5,68	5,00	4,74	5,00	5,73	8,00	6,89

Ces données ont des propriétés statistiques identiques :

- La moyenne
- La variance
- le coefficient de corrélation
- L'équation de la droite de régression linéaire

Quartet d'Anscombe, créé par Francis Anscombe, 1973

REPRÉSENTATION GRAPHIQUE DU QUARTET D'ANSCOMBE

Pourtant, quand on compare leurs représentations graphiques, on s'aperçoit de deux choses :

- Elles sont très différentes les unes des autres
- Certaines valeurs sont incohérentes avec le reste de leur groupe

