

TD 05 : Corrélation et régression

Exercice 01 : Les pluies mesurées pendant 25 ans aux stations suivantes (tableau 01), on remarque que la station P2 comporte des lacunes dans ses données.

1. Calculer le coefficient de corrélation R entre toutes les stations,
2. Trouver la relation de la droite de régression,
3. Combler les lacunes de la station P2, en utilisant la régression linéaire.

Tableau 01 : Les PJmax annuelles des trois stations

Année	P1	P2	P3
1980	50.9	42	51.8
1981	33.3	41.7	37.9
1982	39.3	46	22
1983	94.6	??	44.3
1984	72.1	58.7	23
1985	40.4	22.9	31.4
1986	43	48.4	37
1987	22.7	25	30
1988	54.6	39.5	40.6
1989	33.3	??	35.4
1990	37.3	44	33.6
1991	58.8	44	34
1992	108.6	82	49
1993	46.6	48	46.8
1994	54.5	35.5	31.8
1995	41.5	38	41.4
1996	24.7	15	52.5
1997	46	51	52.5
1998	69.1	56	55.7
1999	83	100	83
2000	45.7	32	38.2
2001	25.9	36	25.9
2002	88.2	68	62.5
2003	51.1	51	63.7
2004	91	72	48.4
2005	103.7	68	82.7
2006	80.8	??	100.3

Solution :

Tableau 02 : Calcul de coefficient de corrélation entre les deux stations (P3 et P2)

Année	P3 (X)	P2 (Y)	$(x - \bar{x})$	$(x - \bar{x})^2$	$(y - \bar{y})$	$(y - \bar{y})^2$	$(x - \bar{x}) * (y - \bar{y})$
1980							
1981							
1982							
1983							
1984							
1985							
1986							
1987							
1988							
1989							
1990							
1991							
1992							
1993							
1994							
1995							
1996							
1997							
1998							
1999							
2000							
2001							
2002							
2003							
2004							
2005							
2006							
Moyenne							

$$r_p = \frac{\sum_{i=1}^N (x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} * \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

Tableau 03 : Calcul de coefficient de corrélation entre les deux stations (P1 et P2)

Année	P1 (X)	P2 (Y)	$(x - \bar{x})$	$(x - \bar{x})^2$	$(y - \bar{y})$	$(y - \bar{y})^2$	$(x - \bar{x}) * (y - \bar{y})$
1980							
1981							
1982							
1983							
1984							
1985							
1986							
1987							
1988							
1989							
1990							
1991							
1992							
1993							
1994							
1995							
1996							
1997							
1998							
1999							
2000							
2001							
2002							
2003							
2004							
2005							
2006							
Moyenne							

$$r_p = \frac{\sum_{i=1}^N (x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} * \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$