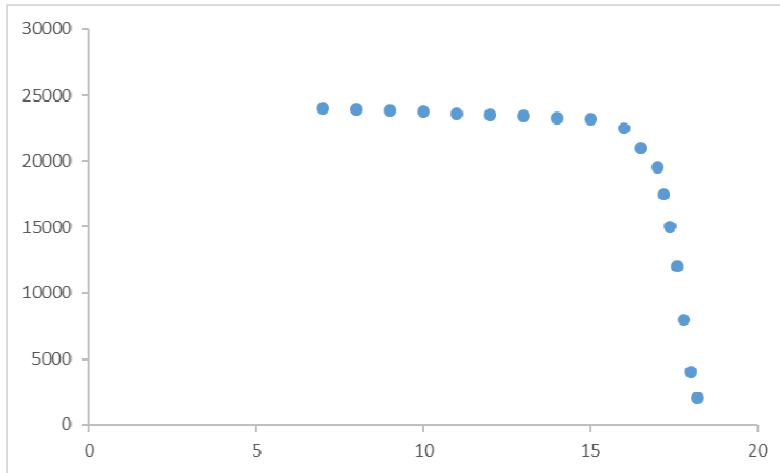


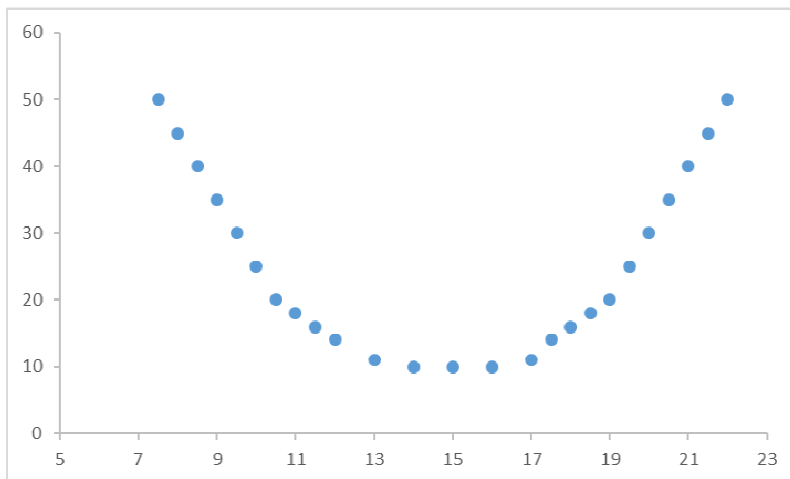
TD4 : Paramètres d'association (covariance)

**Questions de réflexion**

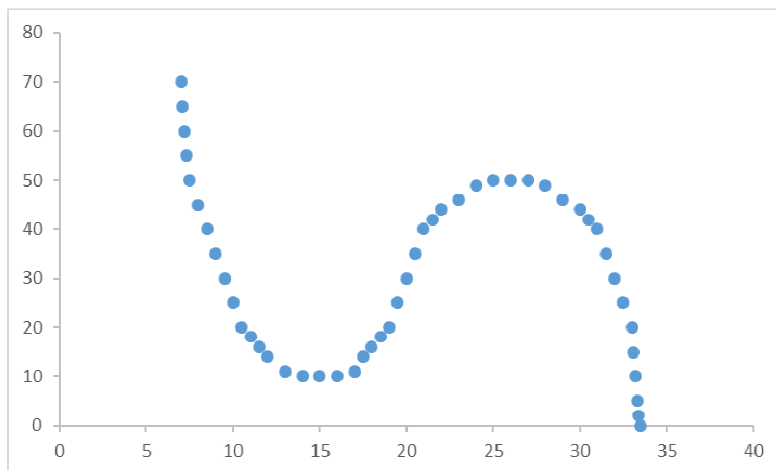
1/ Comment qualifieriez-vous ce type de relation :



Celui-ci :



Celui-ci :



**2/** Un ami vous indique qu'il a trouvé une covariance de 231 959 183 entre deux variables. Que pensez-vous de la valeur de cette covariance ?

**3/** Énoncez les étapes du calcul de la variance et de la covariance. Quelles étapes sont communes aux calculs de la variance et de la covariance ?

## Exercices

A) On s'intéresse à la relation existant entre la taille et le poids chez des patients de type boulimique. Voici nos données :

Numéro du patient	Taille (m)	Poids (kg)
1	1,65	66
2	1,70	82
3	1,68	67
4	1,74	90
5	1,54	57
6	1,58	63
7	1,49	55
8	1,62	72
9	1,71	81
10	1,56	58
11	1,59	67
12	1,63	64
13	1,69	93
14	1,61	68
15	1,57	59

NB : dans les calculs que nous ferons, nous nous intéresserons uniquement à notre groupe de patients, et n'aurons pas pour but de généraliser nos résultats à la population dont ce groupe est issu.

1/ Représentez le diagramme de dispersion de ces données.

2/ Calculez les variances et écarts-types des deux variables.

3/ Calculez la covariance entre les deux variables. Interprétez ce résultat.

4/ On décide de reprendre les données en convertissant la taille en cm plutôt qu'en m :

Numéro du patient	Taille (cm)	Poids (kg)
1	165	66
2	170	82
3	168	67
4	174	90
5	154	57
6	158	63
7	149	55
8	162	72
9	171	81
10	156	58
11	159	67
12	163	64
13	169	93
14	161	68
15	157	59

Recalculez la covariance à partir de ces données transformées. Comparez avec la covariance que vous avez obtenue à la question 2. Que pouvez-vous conclure sur la relation entre les variables taille et poids ?

**B)** On s'intéresse à la relation entre le taux de présence en cours de statistiques (de 0 à 14 puisqu'il y a 14 heures de cours de stat en tout) et la note aux examens de fin d'année (sur 20). Voici nos résultats sur un groupe de 13 étudiants :

Numéro de l'étudiant	Taux de présence	Note
1	12	8
2	10	7
3	13	5
4	14	16
5	9	7
6	12	10
7	13	11
8	14	9
9	14	17
10	12	16
11	14	6
12	13	4
13	14	9

NB : Nous aurons pour but de généraliser nos résultats à l'ensemble de la promo dont sont issus ces étudiants.

1/ Représentez les données sur un diagramme de dispersion.

2/ Calculez la covariance entre les deux variables.

3/ Que pensez-vous de la relation entre les deux variables au vu du diagramme de dispersion et de la valeur de la covariance ?

(Question subsidiaire) Quelqu'un qui a déjà suivi le TD5 vous souffle que 7% de la variance de la note sont expliqués par le taux de présence en cours. Que pensez-vous de ce résultat ? Quels facteurs pourraient l'influencer selon vous ?