

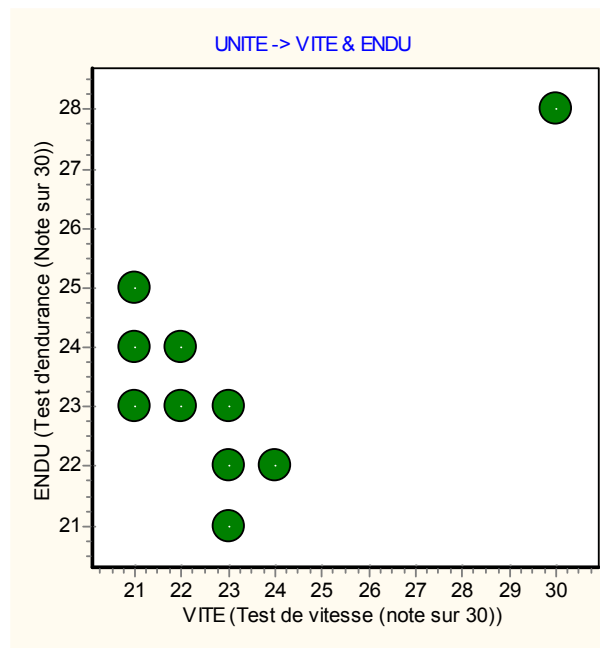
Dossier VITENDUR - Analyse

On cherche à connaître la relation entre la vitesse et l'endurance. En particulier, peut-on prédire le score d'endurance connaissant le score de vitesse ?

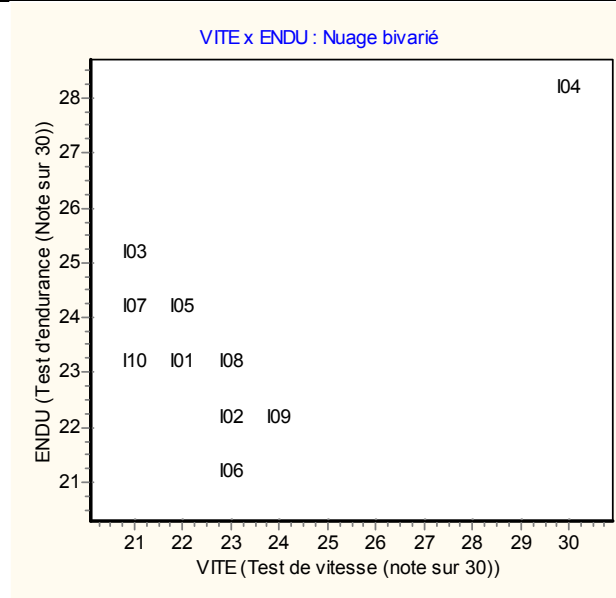
	Vitesse (X)	Endurance (Y)
i01	22	23
i02	23	22
i03	21	25
i04	30	28
i05	22	24
i06	23	21
i07	21	24
i08	23	23
i09	24	22
i10	21	23

Le coefficient de corrélation linéaire de Bravais-Pearson calculé sur ces 10 sportifs est égal à $+0.57$ et l'équation de la droite de régression permettant de prédire le score d'endurance en fonction du score de vitesse est : $Endurance = 0.42 Vitesse + 13.80$.

1) Avant de conclure, on représente le nuage de points associé au tableau ci-dessus. Que constate-t-on ?

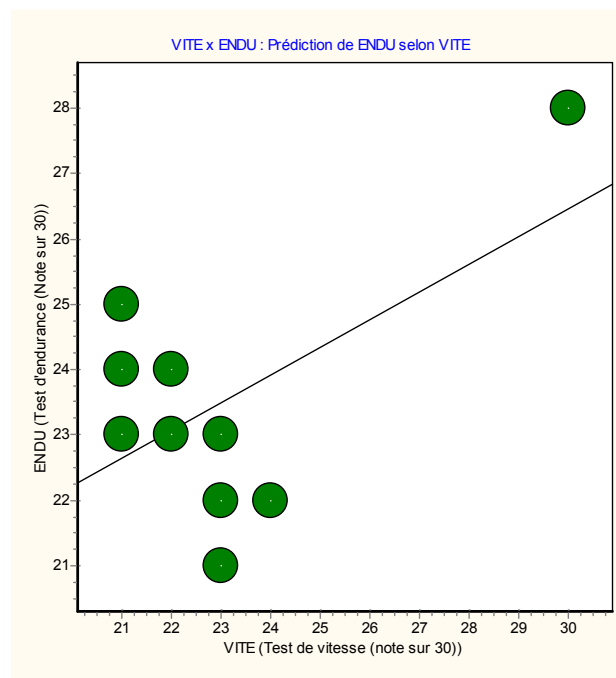


On constate l'existence d'un point atypique. Quel est-il ?



C'est le sportif i04. Il est à la fois très bon en vitesse et très bon en endurance.

La présence de ce sportif fait que le nuage à une pente positive (cf. graphe ci-dessous) :



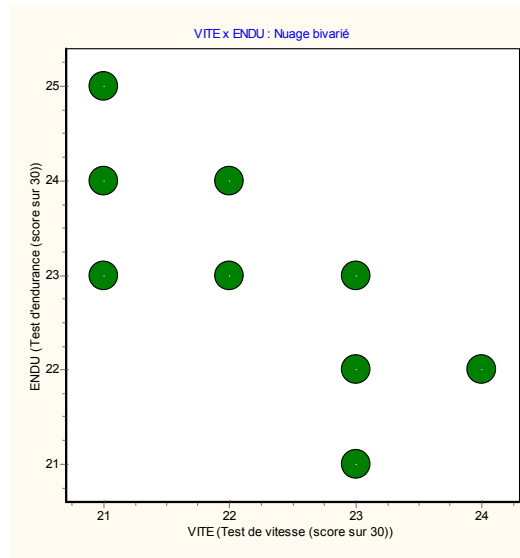
2) Si on retire un des sportifs de l'analyse on trouve :
 $r = -0.75$

$Endurance = -0.84 Vitesse + 41.60$

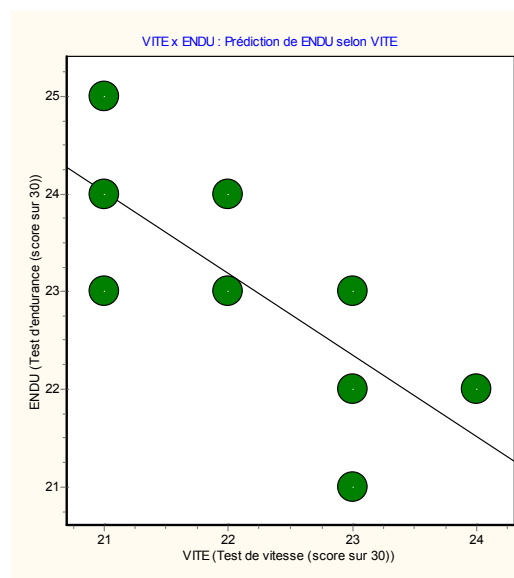
Rédiger une conclusion générale incluant les différents aspects ainsi mis en évidence.

On retire l'individu atypique i4 des calculs.

Le graphe a alors la forme suivante, indiquant cette fois une liaison négative entre les deux variables :



On peut représenter la nouvelle droite de régression sur ce graphe de corrélation. Elle a une pente négative.



En retirant le sujet i4, on constate que la corrélation est forte et négative et que la vitesse rend compte cette fois de 56% de la variance des scores en endurance.. C'était le sujet i4 qui donnait à lui seul une corrélation positive.

Conclusion :

Dans ce groupe de 10 sportifs, excepté un sportif (i4) qui a un score de vitesse et un score d'endurance importants (les plus élevés du groupe), on a une relation négative en le score de vitesse et celui d'endurance : les sportifs les plus rapides sont les moins endurants et réciproquement ($r = -0.75$).

Remarque finale :

Cet exemple souligne, une fois de plus, la nécessité de ne pas conclure l'analyse de la liaison entre deux variables numériques sans avoir représenté le graphe de corrélation. Ce graphe permet, entre autres, de détecter des valeurs atypiques que le coefficient de corrélation ne peut détecter.