

# Rapport analyse de données

**1) Extraire de la base le nombre de ventes effectuées en 2022 par mois, vous obtenez un premier jeu de 12 valeurs noté V. Calculez :**

V est obtenu en résultat de la requête :

```
select MONTHNAME(STR_TO_DATE(MONTH(finVe), '%m')) as mois, count(idve)  
as ventes from VENTE natural join STATUT where nomst = "Validée" and  
YEAR(finVe) = 2022 group by MONTH(finVe);
```

Mois	Ventes
January	14
February	25
March	22
April	40
May	31
June	31
July	35
August	46
September	40
October	38
November	34
December	26

L'espérance de V = la moyenne de V donc

$$(14+25+22+40+31+31+35+46+40+38+34+26) / 12 = 31.83$$

La médiane de V = Il faut pour cela trier par ordre croissant et prendre la valeur "au milieu" donc [14,22,25,26,31,31,34,25,38,40,40,46] cela donne donc 31 et 34 donc la médiane est égale à  $(31+34)/2=32.5$

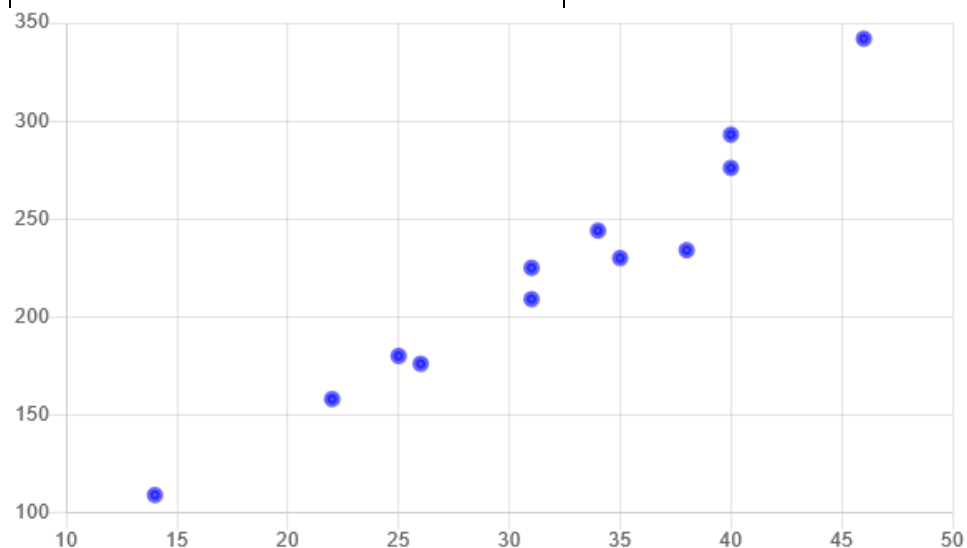
Le mode de V = 40 et 31 car ils se répètent un nombre égal de fois (2)

**2) Extraire le nombre d'enchères de 2022 par mois, ces valeurs constituent une seconde statistique notée E. Vous obtenez ainsi la statistique double (V, E). Tracez le nuage de points (V, E) :**

La requête pour obtenir cette série E est :

```
Select  MONTHNAME(STR_TO_DATE(MONTH(finVe), '%m')) as mois  
,count(montant) as nombreEncheres from VENTE natural join ENCHERIR  
where YEAR(finVe)=2022 group by MONTH(finVe);
```

Mois	nombreEncheres
January	109
February	180
March	158
April	276
May	225
June	209
July	230
August	342
September	293
October	234
November	244
December	176



**Voyez-vous apparaitre (visuellement) une corrélation linéaire entre V et E ?**

Le nuage de points montre une corrélation entre V et E, on remarque une droite qui se dessine, reliant les points.

**Calculez le coefficient de corrélation  $\rho_{V,E}$  du couple (V, E) :**

$$\text{Moyenne V} = 382/12 = 31.83$$

$$\text{Moyenne E} = 2676/12 = 223$$

Le calcul est donc

$$r = \frac{\sum [(x_i - \bar{x}) (y_i - \bar{y})]}{\sqrt{\sum (x_i - \bar{x})^2 * \sum (y_i - \bar{y})^2}}$$

Où r est le coefficient de corrélation, x correspond à V et y correspond à E.

Le calcul "brut" donne :

$$\begin{aligned} & ((14-31.83)(109-223)+(25-31.83)(180-223)+(22-31.83)(158-223)+(40- \\ & 31.83)(276-223)+(31-31.83)(225-223)+(31-31.83)(209-223)+(35- \\ & 31.83)(230-223)+(46-31.83)(342-223)+(40-31.83)(293-223)+(38- \\ & 31.83)(234-223)+(34-31.83)(244-223)+(26-31.83)(176-223))/\text{Sqrt}(((14- \\ & 31.83)^2+(25-31.83)^2+(22-31.83)^2+(40-31.83)^2+(31-31.83)^2+(31- \\ & 31.83)^2+(35-31.83)^2+(46-31.83)^2+(40-31.83)^2+(38-31.83)^2+(34- \\ & 31.83)^2+(26-31.83)^2)*((109-223)^2+(180-223)^2+(158-223)^2+(276- \\ & 223)^2+(225-223)^2+(209-223)^2+(230-223)^2+(342-223)^2+(293-223)^2+(234- \\ & 223)^2+(244-223)^2+(176-223)^2)) \\ & = 6076/\text{sqrt}(883.667*43960) \end{aligned}$$

$$\rho_{V,E} = 0.975$$

Le coefficient étant très proche de 1, on peut en déduire une corrélation entre ces 2 séries statistiques.

**3) En vous servant de la question précédente, quel nombre de ventes (sur un mois donné) anticipez-vous si le nombre d'enchères sur ce même mois est de 285.**

On peut anticiper un nombre d'enchères de  $40.8503 \sim 41$  ventes en s'il y a 285 enchères.