

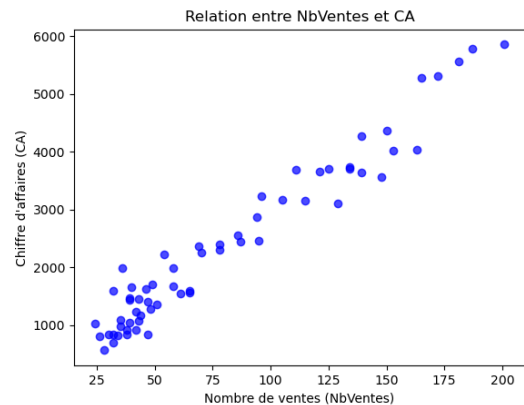
## SAE 2.04 – Exploitation d'une Base de Données

### Sujet n°1 : Livre Express

#### Rapport – Partie Analyse Statistique

#### Question 1

Pour tracer un nuage de points représentant le CA en fonction du nombre de ventes (pour nous, le nombre de ventes c'est le nombre de commandes effectuées et non le nombre de livres vendus voir la requête sql page 2), on utilise matplotlib.pyplot pour pouvoir faire le nuage de points.



#### Question 2

La distribution des points semble former une droite de la forme  $y=ax+b$ , pour affirmer cette hypothèse je vais calculer la corrélation:

$$\frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)} \cdot \sqrt{\text{var}(Y)}} \approx 0.97$$

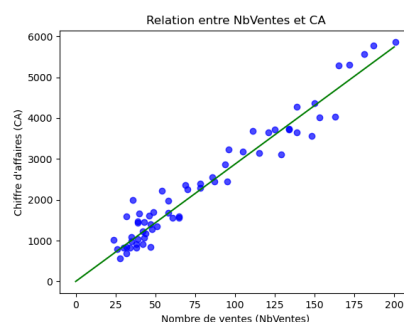
La corrélation entre le CA et le nombre de ventes est entre 0.8 et 1 ce qui veut dire qu'il y a une forte corrélation entre ces deux statistiques.

#### Question 3

Le modèle permettant d'expliquer la relation entre ces deux variables est une droite de régression linéaire donc j'ai calculer le a et le b de la fonction affine  $y=ax+b$  :

$$a = \frac{\text{cov}(X, Y)}{\text{var}(X)} \approx 28 \quad b = \bar{Y} - a \cdot \bar{X} \approx 82$$

la droite de régression linéaire qui modélise la relation est  $y \approx 28x+82$ , voici la fonction:



#### Question 4

Pour estimer le nombre de vente par rapport à notre modèle si notre chiffre d'affaires est de 1250€ nous pouvons utiliser la droite de régression linéaire qu'on a calculer avant.

Il faut retrouver l'antécédent de 1250 donc il faut résoudre cette équation :  $1250 = 28x+82$

Nous trouvons  $x \approx 41.7$ , donc il faut faire au moins 42 commandes pour pouvoir faire un chiffre d'affaires de 1250€.

### Question 5

Notre modèle est pertinent car il est construit sur les données fournies par la base de données mais des événements aléatoires peuvent affecter ce modèle comme par exemple le covid ou une réduction du pouvoir d'achat des citoyens ce qui va affecter directement les ventes et notre modèle ne prend pas en compte ces événements.

**Requete SQL pour le nuage de points :**

```
SELECT
    COUNT(DISTINCT numcom) AS NbVentes,
    SUM(prixvente * qte) AS CA
FROM COMMANDE
NATURAL JOIN DETAILCOMMANDE
GROUP BY MONTH(datecom), YEAR(datecom)
```