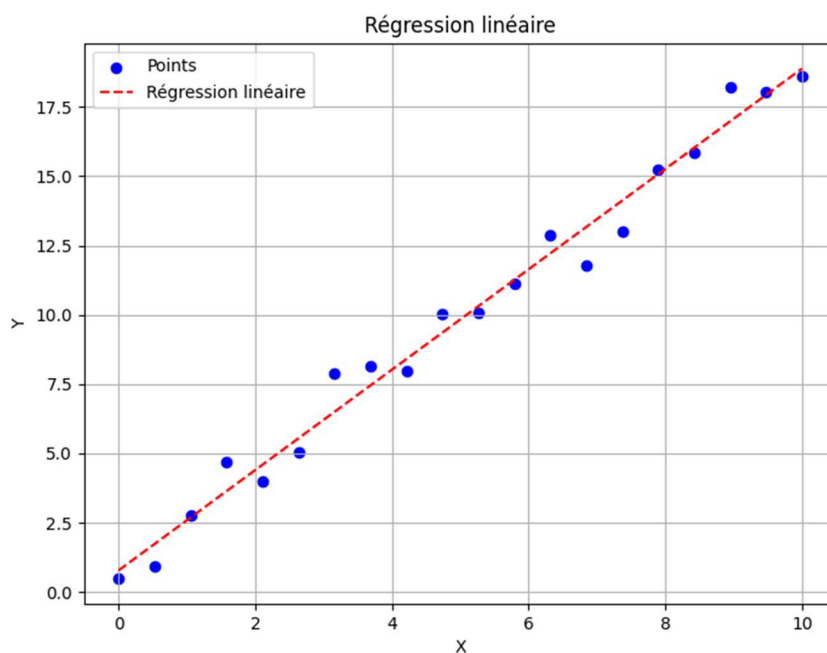


Réalisé par :
ASMAE KARMOUCHI
KENDI MOHAMMED AMINE

2A GL

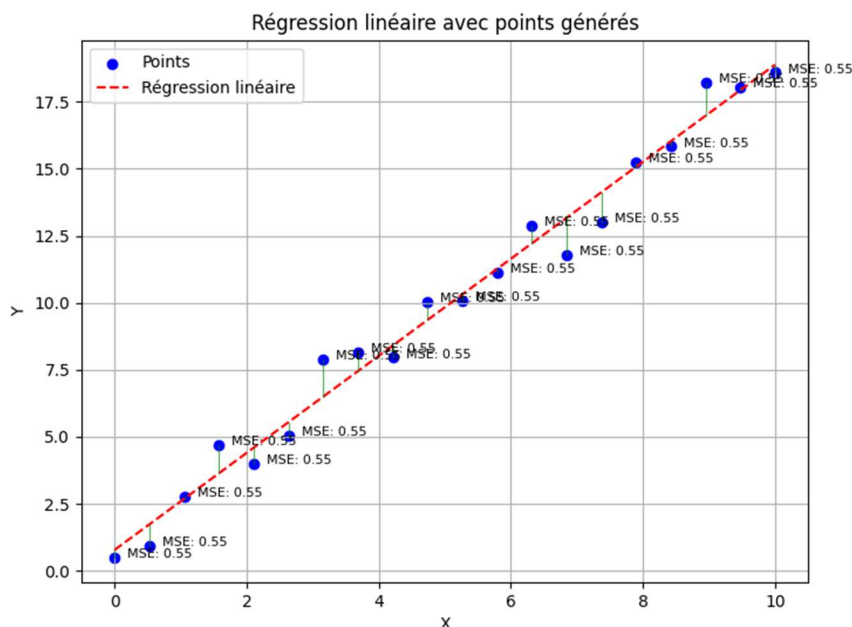
Régression linéaire simple :

- La régression linéaire simple est une méthode statistique qui vise à modéliser la relation entre une variable dépendante Y et une variable indépendante X à l'aide d'une relation linéaire de la forme $Y = \beta_0 + \beta_1 X + \epsilon$.
- Ici, β_0 est l'ordonnée à l'origine (intercept), β_1 est le coefficient de pente (slope) qui mesure le changement moyen de Y pour un changement unitaire de X , et ϵ est le terme d'erreur qui capture l'incertitude et les facteurs non observés.
- L'objectif est de trouver les coefficients β_0 et β_1 qui minimisent l'erreur quadratique moyenne entre les valeurs prédites et les valeurs observées.



Fonction de perte (loss function) :

- La fonction de perte est une mesure de l'erreur entre les valeurs prédites par le modèle et les valeurs réelles observées dans les données.
- En régression linéaire, la fonction de perte la plus couramment utilisée est la somme des erreurs au carré, également connue sous le nom de moindres carrés. Elle est définie comme $\sum (y_i - \hat{y}_i)^2$, où y_i sont les valeurs réelles et \hat{y}_i sont les valeurs prédites par le modèle pour chaque observation.
- La fonction de perte des moindres carrés (ou MSE pour "Mean Squared Error" en anglais), et elle est définie comme la moyenne des carrés des écarts entre les valeurs observées y_i et les valeurs prédites \hat{y}_i , soit : $MSE = 1/n \sum (y_i - \hat{y}_i)^2$.
- L'objectif est de minimiser cette fonction de perte, ce qui revient à trouver les coefficients du modèle qui produisent les meilleures prédictions possibles.

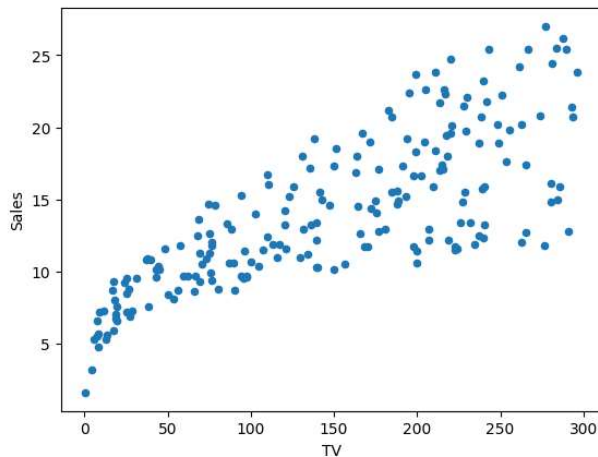


Tv Marketing Prediction :

Dans cet exemple, nous étudions la relation entre les ventes et le budget de marketing pour la télévision ('TV').

- Nous utilisons la régression linéaire simple pour prédire les ventes en fonction du budget alloué à la publicité télévisée:
 1. **Variables** : Nous avons deux variables : les ventes ('Sales') et le budget de marketing télévisé ('TV').
 2. **Modèle de Régression Linéaire Simple** :
- **Observation de la Visualisation des Données** : En visualisant les données à l'aide de graphiques tels que des nuages de points (scatter plots), nous avons remarqué une tendance linéaire apparente entre les ventes et le budget de publicité TV. Les points de données

semblaient être alignés de manière générale le long d'une ligne droite, ce qui suggère une relation linéaire entre les deux variables.



- **Calcul du Coefficient de Corrélation :** Nous avons calculé le coefficient de corrélation de Pearson entre le budget de publicité TV et les ventes. Le coefficient de corrélation obtenu était de 0.7822. Ce chiffre étant proche de 1, cela indique une corrélation positive forte entre les deux variables. En d'autres termes, lorsque le budget de publicité TV augmente, les ventes ont tendance à augmenter également, et vice versa.
- 3. **Prédiction :** Une fois le modèle entraîné, nous utilisons le budget de publicité TV pour faire des prédictions sur les ventes. En appliquant les coefficients obtenus à l'équation de la droite de régression, nous obtenons les valeurs prédites pour les ventes.
- 4. **Évaluation du Modèle :** Nous évaluons la performance de notre modèle en examinant des mesures telles que la perte (MSE) sur les données d'entraînement et de test.

En résumé, la régression linéaire simple est une méthode statistique puissante pour modéliser la relation entre une variable dépendante et une seule variable indépendante. Dans notre exemple d'application, nous utilisons cette méthode pour comprendre comment les variations dans le budget de publicité TV peuvent influencer les ventes.