**Régression linéaire sur les Données de Voitures**

1. Introduction :

La régression Linéaire est une méthode statistique utilisée pour prédire une variable dépendante y (par exemple, le prix de la vente d’une voiture) en fonction d’une ou plusieurs variables indépendantes x.

1. Exemple avec données réelles :

**Objectif :** Prédire le prix de vente en fonction de l’année

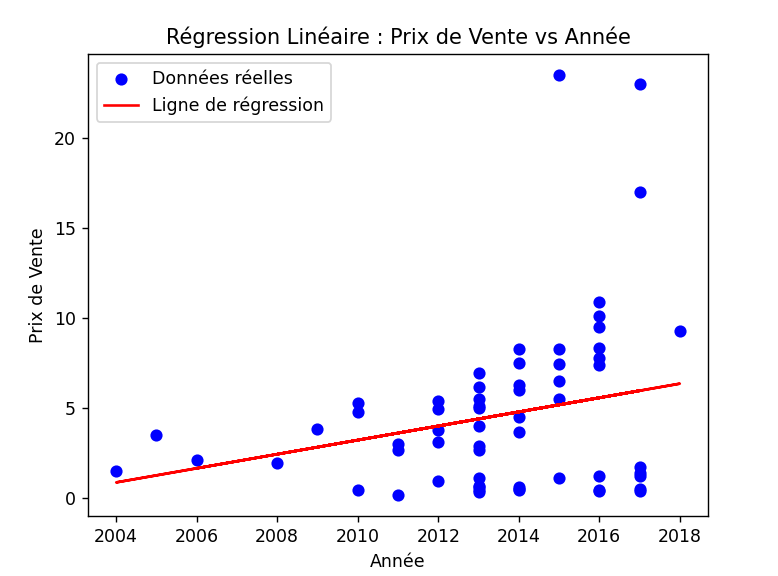
**Étapes :**

1. **Sélection des données** : On utilise uniquement les colonnes Année (variable indépendante) prix\_de\_vente et (variable dépendante).
2. **Entraînement du modèle** : Ajuster une droite pour minimiser l'erreur.
3. **Prédiction** : Utiliser le modèle pour prédire les prix de vente pour une année donnée.
4. Implémentation en Python :

A screen shot of a computer program

Description automatically generated

1. Résultats :



1. Analyse des données :

**Exemple des coefficients :**

* **Pente () : 0.35**  
  Cela signifie que pour chaque année plus récente, le prix de vente augmente en moyenne de **0.35** (dans l'unité monétaire utilisée).
* **Ordonnée à l'origine () : 708.5**  
  Cela représente une estimation théorique du prix de base.

**Points importants :**

1. **Prix récent :**  
   Une voiture fabriquée en **2016** a un prix estimé à :

**= 0.35 x 2016 + 708.5 (prix estimée en fonction de l’année)**

**= 1408.1 (prix estimé en unités monétaires)**

**Tendance générale :**

* + Les voitures récentes (après 2014) ont des prix plus élevés.
  + Les voitures anciennes (avant 2006) ont des prix beaucoup plus bas.

1. **Qualité du modèle :**
   * Les points éloignés de la ligne montrent que d'autres facteurs influencent les prix (kilométrage, carburant, etc.).
2. Exemple Mathématique :

Considérons des données fictives pour simplifier :

X = [2015,2016,2017]

Y = [5,4.5,5]

1. Formule de la droite :

y = m . x + b

1. Calcul de paramètre :

Moyenne de  :

Moyenne de  :

Pente () :

Ordonnée à l’origine () =

1. Prédictions :

Pour = 2018 : y = -0.5. 2018 + 1009.5 = 4.0