

## Activité 2 : Regression Linéaire

---

On souhaite effectuer de la régression linéaire simple.

Dans cette activité, on se basera sur la base "weather.csv" contenant des données météorologiques.

- 1- Importer les données à partir du fichier.
- 2- Afficher la dimension de l'ensemble de données et en déduire le nombre des variables utilisées.
- 3- Afficher les noms de ces variables.

Pour faire de la régression linéaire simple, on se restreindra à deux variables : la valeur minimale de température (MinTemp) et sa valeur maximale (MaxTemp). On considère que "MinTemp" est la variable explicatif et "MaxTemp" et la variable expliquée. Par conséquent, on souhaite prédire la valeur de "MaxTemp" en fonction de la valeur enregistrée de "MinTemp".

- 4- Afficher toutes les observations de la base en considérant la valeur minimale de la température sur l'axe des abscisses et sa valeur maximale sur l'axe des ordonnées.
- 5- Construire le nouvel ensemble de données, composé seulement des deux variables "MinTemp" et "MaxTemp".
- 6- Diviser l'ensemble de données en un ensemble d'apprentissage qui contient 80% des observations et un ensemble de test (20%).
- 7- Déterminer les paramètres "a" et "b" de la droite de régression ( $y=ax+b$ ) en utilisant l'ensemble d'apprentissage.
- 8- Afficher les paramètres du modèle : la pente "a" et l'ordonné à l'origine "b".
- 9- Déterminer et afficher le coefficient de détermination  $R^2$ . Conclure.
- 10- En utilisant les températures minimales de l'ensemble de test et le modèle de régression calculé dans la question précédente, prédire les températures maximales.
- 11- Afficher, sur la même figure, les valeurs réelles observées et celles prédites des températures maximales en fonction des températures minimales, et ce à partir de l'ensemble de test.