

## **"Entrepôt de Données"**

### **Projet**

Le projet consiste à transformer des données climatiques brutes en un entrepôt de données :

- L'implémentation doit se faire exclusivement en *Python* et MySQL.
- Une fois le schéma en étoile modélisé, il doit être créé et stocker dans une base de données nommée « Weather\_DataWarehouse » qui représentera l'entrepôt de données.
- Les tables de l'entrepôt de données doivent être remplies en utilisant le processus ETL qui chargera les données à partir des fichiers plats vers l'entrepôt de données «Weather\_DataWarehouse ».
- Dans le but de faciliter l'analyse des données, il est demandé d'implémenter un tableau de bord à l'aide du framework [Dash](#) sous Python.
  - Des graphiques et des schémas dynamiques visualisant l'évolution des données climatiques doivent être inclus. À titre d'exemple :
    - Une carte géographique visualisant les différentes mesures avec des filtres sur l'année, la saison, le trimestre et le mois.
    - Une courbe traçant l'évolution de la température, la pluviométrie, la profondeur de la neige sur plusieurs années
    - Etc ...
- Rédiger un rapport (5 pages) qui doit bien expliquer les aspects suivants :
  - Le schéma en étoile de l'entrepôt de données,
  - Le processus traité par l'entrepôt de données,
  - La granularité de la table des faits,
  - Les mesures,
  - Les dimensions,
  - Le tableau de bord.

#### **Description du dataset :**

Le [dataset](#) fourni par *National Centers for Environmental Information*, comporte des données climatiques couvrant trois pays du Maghreb. On y trouve les dix attributs suivants :

- Précipitation (PRCP)
- Épaisseur de la neige (SNWD)
- Chute de neige (SNOW)
- Température moyenne (TAVG)
- Température maximale (TMAX)
- Température minimale (TMIN)
- Direction de la rafale de vent maximale (WDFG)
- Heure de pointe des rafales (PGTM)
- Vitesse maximale du vent en rafale (WSFG)
- Types de temps (WT\*\*)

En plus des données climatiques, on trouve des attributs qui fournissent des informations relatives au code de la station, son nom, sa position géographique et la date du prélèvement.

#### **Ressources :**

- Dataset : [Weather Data.zip](#)
- Dimension temporelle : [DateDim.csv](#)
- Documentation Dash : [dash.plotly.com](#)