

**Centre de formations professionnelles aux métiers du numérique**

**Projet fil rouge Data Analyst**

# **ANALYSE DES DONNEES CLIMATIQUES**

Sous l'encadrement de :

**M. DONTSA Cyprien**

Réalisé par l'apprenant :

**TAGNE FOSSI BRICE GAUTIER**

*Année académique : 2023 – 2024*

## Table des matières

<b>I. Contexte et objectifs.....</b>	<b>2</b>
<b>II. Exigences .....</b>	<b>Erreur ! Signet non défini.</b>
<b>III. Livrables .....</b>	<b>Erreur ! Signet non défini.</b>
<b>IV. Calendrier .....</b>	<b>Erreur ! Signet non défini.</b>

## I. Contexte et objectifs

### Contexte :

- **Zone d'étude** : Cameroun.
- **Période** : 1958-2023
- **Données disponibles** : données climatiques mensuelles provenant de nombreux stations météorologiques (température maximale, température minimale, précipitation, vitesse du vent, rayonnement descendant et le ruissellement)
- **Granularité** : Analyse et prédiction par région

### Objectifs :

- **Objectif principal** : Analyser les tendances climatiques historiques et prédire la variable température pour chaque région du Cameroun.
- **Objectifs spécifiques** :
  - Visualiser les tendances historiques et les variables saisonnières des variables climatiques à l'aide de power BI.
  - Identifier les changements significatifs dans les variables climatiques et quantifier leur ampleur.
  - Développer un modèle de machine Learning en python pour prédire les valeurs futures de la variable température.
  - Evaluer la performance du modèle de prédiction et identifier les variables les plus difficiles à prédire.

## II. Exigences

- **Données** :
  - Accès à des données climatiques historiques de qualité pour chaque région du Cameroun (température, précipitation, vitesse du vent, rayonnement descendant, ruissellement).
  - Les données doivent être cohérente, complètes et couvrent une période de temps suffisamment longue pour identifier les tendances climatiques et entraîner un modèle de machine Learning.
- **Outils et logiciel** :
  - Power BI pour la visualisation des données.

- Python et ses bibliothèques (Numpy, Pandas, scikit-learn, etc.) pour le développement du modèle de machine Learning.
- Connaissance des techniques de prétraitement des données, de sélection de caractéristiques, de modélisation et d'évaluation de modèles de machine Learning.
- **Expertise :**
  - Expertise en climatologie, changement climatiques, et analyse de données et de machine Learning.
  - Compréhension des algorithmes de machine Learning et des techniques de prédictions

### III. Livrables

- **Rapport final :**
  - Un rapport complet documentant la méthodologie, les résultats de l'analyse visuelle et les performances du modèle de machine Learning.
  - Le rapport doit inclure des captures d'écran des visualisations power BI, des tableaux de résultats et des indicateurs d'évaluation du modèle.
  - Un résumé en français.
- **Code source :**
  - Le code source du modèle de machine Learning en Python, documenté et organisé de manière claire.
- **Présentation :**
  - Une présentation PowerPoint résumant les résultats clés de l'analyse et les performances, les résultats du modèle de prédiction pour une communication orale.

### IV. Calendrier

- **Phase 1 :** Collecte et préparation des données (2 jour)
  - Rassembler les données climatiques de toutes les sources pertinentes.
  - Nettoyer, prétraiter et homogénéiser les données.
- **Phase 2 :** Analyse visuelle avec Power BI (4 jours) :

- Créer des visualisations interactives pour explorer les tendances climatiques et identifier les modèles.
- Analyser les variations saisonnières et les anomalies climatiques.
- **Phase 3** : Développement du modèle de machine Learning (4 jours) :
  - Sélectionner un algorithme de machine Learning adapté à la tâche de prédiction.
  - Entraîner et optimiser le modèle de machine Learning en utilisant les données historiques.
  - Évaluer la performance du modèle sur un ensemble de données de test indépendant.
- **Phase 4** : Rédaction du rapport et préparation des livrables (4 jours) :
  - Rédiger le rapport final documentant la méthodologie, les résultats de l'analyse visuelle et les performances du modèle de machine Learning.
  - Préparer le code source et la présentation PowerPoint.
- **Phase 5** : Validation et communication des résultats (2 jours) :
  - Soumettre le rapport final, le code source et la présentation pour validation.
  - Présenter les résultats aux parties prenantes et répondre aux questions.