

Alyssa DERENY
Alexy GABORIAUD
Marylin FONTA
Solène ISSANES
Kenzo SALAUN

Documentation du Projet :

Visualisation des Données Météorologiques

Introduction

Ce projet a pour objectif de créer un pipeline de données météorologiques afin de visualiser des informations sur les températures de différentes villes françaises. L'application utilisera des technologies telles que MongoDB, Python, Dash, Pandas et Plotly afin d'offrir à l'utilisateur une interface interactive et le choix de réaliser des analyses pertinentes et approfondies.

Technologies utilisées

- **Python** : Langage de programmation principal utilisé pour gérer l'API, le traitement des données et la création de l'application
- **Dash** : Framework pour le développement d'applications web interactives
- **Pandas** : Bibliothèque pour la manipulation et l'analyse de données sous la forme de tableaux
- **Plotly** : Bibliothèque pour la création de graphiques interactifs par l'utilisateur
- **MongoDB** : Base de données NoSQL utilisée pour le stockage des données météorologiques
- **JSON** : Format de fichier utilisé pour les informations des villes, facilitant l'intégration et le traitement des données.

Rôles et tâches par les membres du groupe

- Chef de projet : Alyssa DERENY
- Développeur Backend : Kenzo SALAUN
- Intégrateur API : Solène ISSANES
- Spécialiste DevOps : Alexy GABORIAUD
- Analyste de données : Maryline FONTA

Description des données

Nous avons récupéré les données météorologiques depuis l'API REST du site *openweathermap.org*. Ce site fournit des informations sur les températures actuelles, la vitesse et la direction du vent, ainsi que les coordonnées géographiques (latitude et longitude) des villes.

Nous avons sélectionné les données **des 60 villes les plus peuplées de France**, qui ont été utilisées comme base pour nos analyses. Ces villes sont identifiées grâce à un fichier JSON contenant notamment l'identifiant de la ville, le nom de la ville, la latitude et la longitude. Ce fichier est intégré dans le processus de collecte et de visualisation des données.

Structure de l'application

L'application Dash est composée de quatre pages principales et accessibles via une interface utilisateur interactive :

1. **Page d'accueil** : Elle présente une introduction au projet avec les informations générales, les membres de l'équipe projet et l'école. Puis dessous, il y a des liens qui redirige vers chacune des pages créées.
2. **Page pour le dashboard global** : Elle affiche 3 indicateurs : le nombre de villes françaises analysées, la température moyenne du jour et la vitesse moyenne du vent. Puis dessous, il y a deux graphiques : à gauche, un premier diagramme en barre représentant la distribution des températures des villes françaises, puis sur le côté droit une carte géographique zoomée sur la France avec les températures et la description du temps de chacune des villes françaises.
3. **Page 1** : Elle affiche le top 10 des villes les plus chaudes et les plus froides en fonction de la température en France dans deux tableaux. En dessous des deux tableaux, il y a un box plot pour montrer la dispersion entre les valeurs de la variable "Températures" entre les différentes villes de France.
4. **Page 2** : L'utilisateur peut choisir de sélectionner l'ensemble des données ou uniquement celles de son choix à l'aide d'un déroulant. Le graphique est un nuage de points interactif permettant de visualiser les températures des villes. Puis dessous, il y a la liste des villes françaises avec leur température.
5. **Page 3** : Affiche deux nuages de point illustrant la corrélation entre la température et la vitesse du vent, puis entre la température et la latitude.
6. **Page 4** : Tableau récapitulatif des données météorologiques.

Flux de données

Le projet utilise **MongoDB** pour stocker et gérer les données météorologiques. Les étapes suivantes résument le flux de données :

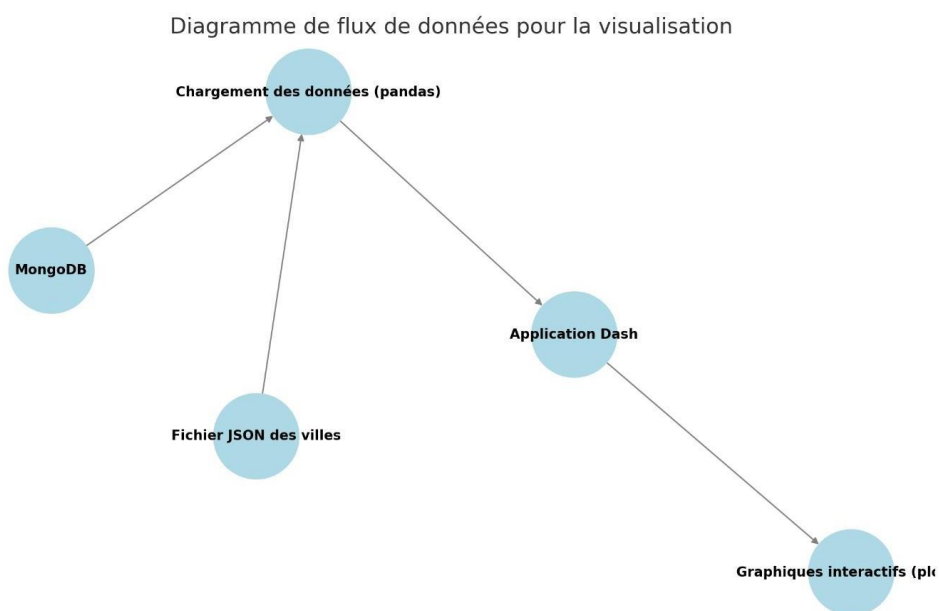
1. **Collecte des données** : à partir d'une source externe (comme OpenWeatherMap) et stockées dans un fichier JSON.
2. **Insertion dans MongoDB** : Les données JSON sont importées dans une base de données MongoDB pour un stockage structuré.
3. **Analyse avec Pandas** : Les données sont traitées avec Pandas pour effectuer des calculs et préparer les données nécessaires à la visualisation.
4. **Visualisation avec Dash et Plotly** : L'application Dash utilise les bibliothèques Plotly et Dash pour créer des graphiques interactifs et des tableaux.

Description du projet

1. **Source de données :**
 - **MongoDB** : Base de données où sont stockées les informations météorologiques.
 - **Fichier JSON des villes** : Contient les données des villes nécessaires à l'affichage des graphiques et à la sélection par l'utilisateur.
2. **Traitement des données :**
 - **Chargement et prétraitement** : Utilisation de Pandas pour charger et préparer les données issues de MongoDB et du fichier JSON.
3. **Application Dash :**
 - **Pages d'analyse** (Page 1, Page 2, Page 3) : Ces pages utilisent les données traitées pour afficher des graphiques interactifs et des tableaux.
4. **Visualisation des données :**
 - **Graphiques interactifs** : Utilisation de Plotly pour créer des graphiques basés sur les données de température, la corrélation entre la température et la vitesse du vent, ainsi que la température par ville.

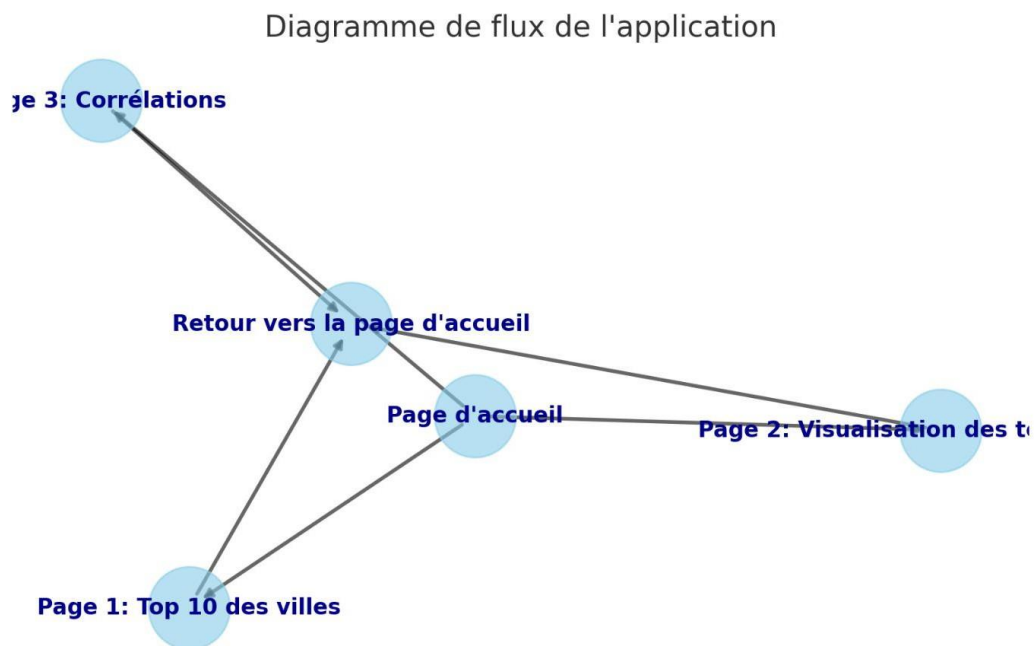
Diagramme

1. Fichier JSON des villes ↔ Chargement des données (Pandas) ↔ Application Dash
2. MongoDB ↔ Chargement des données (Pandas) ↔ Application Dash
3. Application Dash ↔ Graphiques interactifs (Plotly)



Structure du schéma (par étapes) :

- **Accueil -> Page 1 / Page 2 / Page 3** : L'utilisateur peut choisir la page à afficher.
- Chaque page d'analyse peut rediriger vers l'accueil.



Voici un diagramme de flux illustrant la structure du projet. Il montre comment les pages sont connectées entre elles, avec des flèches indiquant la direction de navigation entre la page d'accueil et les différentes pages d'analyse :

1. **Page d'accueil** :
 - L'utilisateur peut naviguer vers **Page 1**, **Page 2**, ou **Page 3** via les liens disponibles.
2. **Page 1, Page 2, et Page 3** :
 - Chaque page d'analyse permet à l'utilisateur de revenir à la **Page d'accueil**.

Conclusion

Ce projet offre une vue interactive des conditions météorologiques des villes les plus peuplées de France, permettant aux utilisateurs d'explorer des visualisations des températures, des tendances et des corrélations importantes pour l'analyse des données météorologiques.