

NOTICE D'UTILISATION

SCRIPT DE RÉCUPÉRATION DE DONNÉES MÉTÉOROLOGIQUES

Open-Meteo Historical Weather Data Extractor

Description

Ce script permet de récupérer des données météorologiques historiques à partir de l'API Archive d'Open-Meteo et de les enregistrer au format Parquet. Il est particulièrement utile pour obtenir des séries temporelles de données météorologiques pour plusieurs emplacements géographiques sur une période donnée.

Prérequis

Le script nécessite les bibliothèques Python suivantes :

- openmeteo_requests
- requests_cache
- retry_requests
- pyarrow
- datetime
- os
- time

Installation des dépendances :

```
pip install openmeteo-requests requests-cache retry-requests pyarrow
```

Utilisation principale

python

```
from meteo_extractor import get_weather

# Définir Les emplacements (nom, Latitude, Longitude)
locations = [
    ("Paris", 48.85, 2.35),
    ("Lyon", 45.75, 4.85)
]

# Récupérer Les données
result = get_weather(
    locations,
    date_start="2023-01-01",
    date_end="2023-12-31",
    output_path="data",
    output_filename="meteo_france.parquet"
)

print(f"Récupération terminée : {result['num_rows']} enregistrements")
```

Paramètres de la fonction get_weather()

Paramètres obligatoires

- **locations** : Liste de tuples (nom_lieu, latitude, longitude) Liste des emplacements pour lesquels récupérer les données météorologiques. Exemple : `[("Paris", 48.85, 2.35), ("Lyon", 45.75, 4.85)]`
- **date_start** : Chaîne de caractères au format "YYYY-MM-DD" Date de début de la période pour laquelle récupérer les données. Exemple : `"2023-01-01"`
- **date_end** : Chaîne de caractères au format "YYYY-MM-DD" Date de fin de la période pour laquelle récupérer les données. Exemple : `"2023-12-31"`

Paramètres optionnels

- **hourly_variables** : Liste de chaînes de caractères (défaut: None) Variables météorologiques à récupérer. Si None, une liste par défaut complète est utilisée. Exemple : `["temperature_2m", "precipitation", "wind_speed_10m"]`
- **output_path** : Chaîne de caractères (défaut: None) Chemin du dossier où enregistrer le fichier Parquet. Si None, aucun fichier n'est créé. Exemple : `"data/meteo"`
- **output_filename** : Chaîne de caractères (défaut: None) Nom du fichier Parquet. Si None, un nom est généré automatiquement. Exemple : `"meteo_france_2023.parquet"`

- **partition_col** : Chaîne de caractères (défaut: None) Colonne de partitionnement pour filtrer les données. Options: "année", "mois", "jour", "heure". Exemple : `"mois"`

Variables météorologiques disponibles

Si aucune liste de variables n'est fournie (hourly_variables=None), les variables suivantes sont récupérées :

- **Vent** : wind_speed_10m, wind_speed_80m, wind_speed_120m, wind_speed_180m, wind_direction_10m, wind_direction_80m, wind_direction_120m, wind_direction_180m, wind_gusts_10m
- **Température** : temperature_80m, temperature_120m, temperature_180m, apparent_temperature, soil_temperature_0cm, soil_temperature_6cm, soil_temperature_18cm, soil_temperature_54cm
- **Humidité et précipitations** : rain, snowfall, snow_depth, precipitation, soil_moisture_0_to_1cm, soil_moisture_1_to_3cm, soil_moisture_3_to_9cm, soil_moisture_9_to_27cm, soil_moisture_27_to_81cm
- **Nuages et ensoleillement** : cloud_cover, cloud_cover_low, cloud_cover_mid, cloud_cover_high, sunshine_duration
- **Autres** : visibility, is_day, uv_index, surface_pressure, pressure_msl, weather_code

Format des données de sortie

Les données sont enregistrées au format Parquet avec les colonnes suivantes :

- datetime : Horodatage (timestamp Unix)
- location_name : Nom de l'emplacement
- latitude : Latitude de l'emplacement
- longitude : Longitude de l'emplacement
- [Variables météorologiques] : Une colonne par variable demandée

Remarques importantes

1. Le script inclut une pause de 60 secondes entre chaque emplacement pour éviter de surcharger l'API.
2. Les données sont récupérées à une résolution horaire.
3. L'option de partitionnement (partition_col) filtre les données selon la date actuelle.
4. L'API Open-Meteo a des limites d'utilisation, vérifiez leur documentation pour plus d'informations.

Exemple d'utilisation pour la surveillance météorologique des routes aériennes

L'exemple par défaut du script récupère les données météorologiques pour :

- 4 grands aéroports américains : Atlanta (ATL), Denver (DEN), Los Angeles (LAX), Chicago O'Hare (ORD)

- 12 points intermédiaires sur les routes aériennes entre ces aéroports
- Sur une période de 3 ans (2018-2020)
- Avec toutes les variables météorologiques disponibles

Ces données peuvent être utilisées pour analyser les conditions météorologiques le long des couloirs aériens et leur impact potentiel sur les opérations de vol.

Commentaire by Claude