

e Pratique : Acquisition et Traitement d'Informations sur les Plus Grandes Banques

Préliminaires : Installation des Bibliothèques et Téléchargement des Données

Avant de créer le code, vous devez installer les bibliothèques requises.

Les bibliothèques nécessaires au code sont :

- requests
- bs4
- pandas
- sqlite3
- numpy
- datetime

Installez les bibliothèques requises à partir de la fenêtre du terminal. La syntaxe de la commande est la suivante :

```
python3.11 -m pip install <library_name>
```

Téléchargez également le fichier de taux de change à l'aide de la commande suivante :

```
wget
```

```
https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMSkillsNetwork-PY0221EN-Coursera/labs/v2/exchange_rate.csv
```

Structure du Code

Créez le fichier `banks_project.py` dans le chemin `/home/project/`. Copiez et collez la structure de code suivante dans le fichier:

```
# Code for ETL operations on Country-GDP data

# Importing the required libraries

def log_progress(message):

    """ This function logs the mentioned message of a given stage of the

        code execution to a log file. Function returns nothing"""

def extract(url, table_attribs):

    """ This function aims to extract the required

        information from the website and save it to a data frame. The

        function returns the data frame for further processing. """

    return df

def transform(df, csv_path):

    """ This function accesses the CSV file for exchange rate

        information, and adds three columns to the data frame, each

        containing the transformed version of Market Cap column to

        respective currencies"""

    return df

def load_to_csv(df, output_path):

    """ This function saves the final data frame as a CSV file in

        the provided path. Function returns nothing."""

def load_to_db(df, sql_connection, table_name):

    """ This function saves the final data frame to a database

        table with the provided name. Function returns nothing."""

def run_query(query_statement, sql_connection):

    """ This function runs the query on the database table and

        prints the output on the terminal. Function returns nothing. """

""" Here, you define the required entities and call the relevant
```

functions in the correct order to complete the project. Note that this portion is not inside any function."

Tâches du projet

Tâche 1 : Fonction de Journalisation

Écrivez une fonction `log_progress()` pour enregistrer la progression du code à différentes étapes dans un fichier `code_log.txt`.

Tâche 2 : Extraction des Données

Extraire les informations tabulaires de l'URL donnée sous le titre « Par capitalisation boursière » et les enregistrer dans un dataframe.

Tâche 3 : Transformation des Données

Transformez le dataframe en ajoutant des colonnes pour la capitalisation boursière en GBP, EUR et INR.

Tâche 4 : Chargement au Format CSV

Chargez le dataframe transformé dans un fichier CSV de sortie.

Tâche 5 : Chargement dans la Base de Données

Chargez le dataframe transformé sur un serveur de base de données SQL sous forme de table.

Tâche 6 : Exécution des Requêtes sur la Base de Données

Exécutez des requêtes sur la table de base de données.

Tâche 7 : Vérification des Entrées du Journal

Vérifiez que les entrées du journal ont été complétées à toutes les étapes en vérifiant le contenu du fichier `code_log.txt`.

Conclusion

Le projet permet d'acquérir une expérience pratique en travaillant avec des données réelles. Vous serez en mesure de pratiquer l'extraction, la transformation, et le chargement (ETL) des données sur des bases de données, tout en loguant la progression de chaque étape du processus.