



# CAHIER DE CHARGE DE PROJET FIL ROUGE

Projet : La Course Verte

*Analyse de la Transition  
Énergétique Mondiale*

Réalisé par :  
**TAHRI SOUAD**

Encadré par :  
**Yassine Ammami**

Février 2026

## Table des matières

<b>1</b>	<b>Introduction Générale</b>	<b>2</b>
1.1	Contexte du Projet .....	2
1.2	Problématique Métier .....	2
1.3	Objectifs du projet.....	2
1.4	Périmètre du projet .....	3
<b>2</b>	<b>Sources de données</b>	<b>3</b>
<b>3</b>	<b>Indicateurs &amp; KPI métiers</b>	<b>4</b>
<b>4</b>	<b>Architecture data &amp; pipeline</b>	<b>5</b>
<b>5</b>	<b>Préparation et nettoyage des données</b>	<b>6</b>
<b>6</b>	<b>Modélisation des données</b>	<b>6</b>
<b>7</b>	<b>Analyses prévues</b>	<b>7</b>
<b>8</b>	<b>Visualisation &amp; Dashboard</b>	<b>7</b>
<b>9</b>	<b>Automatisation &amp; outils techniques</b>	<b>7</b>
<b>10</b>	<b>Livrables</b>	<b>7</b>
<b>11</b>	<b>Risques &amp; limites</b>	<b>8</b>
<b>12</b>	<b>Résultats attendus &amp; valeur métier</b>	<b>8</b>

# 1 Introduction Générale

## 1.1 Contexte du Projet

Au cours des dernières décennies, le monde est confronté à des défis environnementaux majeurs liés au changement climatique. La croissance démographique, l'industrialisation et l'augmentation de la demande énergétique ont fortement accru l'utilisation des énergies fossiles (charbon, pétrole et gaz naturel).

Cette dépendance entraîne une hausse significative des émissions de CO<sub>2</sub> et contribue au réchauffement climatique. Face à cette situation, de nombreux pays s'orientent vers une transition énergétique basée sur le développement des énergies renouvelables telles que le solaire, l'éolien et l'hydroélectricité.

La transition vers un système énergétique plus propre est donc devenue un enjeu stratégique mondial pour concilier développement économique, accès à l'énergie et protection de l'environnement.

## 1.2 Problématique Métier

Malgré cette volonté de transition énergétique, le pilotage et l'évaluation des performances énergétiques des pays restent complexes.

Les données nécessaires à la prise de décision sont dispersées entre plusieurs sources : consommation d'énergie, mix énergétique, émissions de CO<sub>2</sub>, indicateurs économiques, démographie et accès à l'électricité. Cette dispersion rend difficile l'obtention d'une vision globale, cohérente et exploitable.

La problématique métier est donc :

Comment consolider, analyser et comparer ces données issues de différentes sources afin de mesurer efficacement la transition énergétique des pays et d'aider les décideurs à évaluer leur dépendance aux énergies fossiles et leurs progrès vers des solutions plus durables ?

## 1.3 Objectifs du projet

### Objectif général

Développer un système d'analyse et de visualisation des données énergétiques mondiales permettant d'évaluer la transition énergétique des pays et de fournir des indicateurs d'aide à la décision.

### Objectifs spécifiques

- Collecter les données énergétiques, économiques et démographiques depuis des sources ouvertes (OWID, World Bank)
- Nettoyer, préparer et fusionner les datasets dans un dataset final consolidé
- Construire des indicateurs clés (part renouvelables, part fossiles, CO<sub>2</sub>, énergie par habitant, etc.)
- Analyser l'évolution temporelle et comparer les pays

- Étudier la relation entre énergie émissions et croissance économique
- Fournir des recommandations basées sur les résultats.

## 1.4 Périmètre du projet

Afin d'assurer une gestion efficace du projet et de garantir la cohérence des livrables, le périmètre définit clairement les éléments pris en charge ainsi que ceux exclus de l'étude.

### Dans le périmètre (In scope)

Le projet comprend :

- La collecte de données ouvertes issues de sources publiques (Our World in Data et World Bank) sous format CSV
- L'automatisation de l'extraction, du traitement et de la consolidation des données à l'aide de scripts Python
- Le nettoyage, la préparation et la standardisation des données (gestion des valeurs manquantes, doublons et incohérences)
- La fusion des datasets pour créer un fichier analytique unique (CSV/Parquet)
- La construction d'indicateurs clés de performance (KPI) liés à la consommation énergétique, au mix énergétique (énergies renouvelables vs fossiles), aux émissions de CO<sub>2</sub>, à la population, au PIB et à l'accès à l'électricité
- L'analyse exploratoire et statistique des données afin d'identifier tendances, corrélations et comparaisons entre pays
- La conception de visualisations et d'un tableau de bord interactif pour faciliter l'aide à la décision
- La documentation technique, la restitution des résultats et la présentation finale du projet

### Hors périmètre (Out of scope)

Le projet ne couvre pas :

- L'utilisation de données propriétaires, payantes ou confidentielles
- La mise en place d'architectures Big Data complexes (Spark, Hadoop, etc.)
- Le développement de modèles avancés de Machine Learning ou Deep Learning
- Les analyses prédictives ou les projections futures à long terme
- Le développement d'une application web ou mobile dédiée
- Le traitement de flux de données en temps réel

## 2 Sources de données

Les données utilisées dans ce projet proviennent exclusivement de sources ouvertes et publiques reconnues pour leur fiabilité (Our World in Data et World Bank). Elles sont

fournies au format CSV afin de faciliter l'automatisation de l'extraction, du nettoyage et de l'intégration dans le pipeline data.

Le tableau suivant résume les principales caractéristiques de chaque dataset :

Dataset	Source	Type	Granularité	Colonnes principales
Energy Mix & Electricity	Our World in Data	CSV	Pays /Année	country, year, solar, wind, hydro, coal, gas, oil, renewables%, fossil %, electricity generation
CO2 emissions	Our World in Data	CSV	Pays \Année	country, year, co2, co2_per_capita
Population	World Bank	CSV	Pays /Année	Country, Year, Population
GDP	World Bank	CSV	Pays /Année	Country, Year, GDP, gdp_per_capita
Access to Electricity	World Bank	CSV	Pays /Année	country, year, access_to_electricity (%)

### 3 Indicateurs & KPI métiers

D'évaluer la transition énergétique et d'aider à la prise de décision, un ensemble d'indicateurs clés de performance (KPI) est défini. Ces indicateurs permettent de mesurer la dépendance aux énergies fossiles, l'adoption des énergies renouvelables, l'impact environnemental et le lien avec le développement économique :

#### KPI principaux

##### 1. Énergie

- Part des énergies renouvelables (%)
- Part des énergies fossiles (%)
- Production/consommation d'électricité.
- Consommation d'énergie par habitant

##### 2. Environnement

- Émissions totales de CO2
- CO2 par habitant
- Intensité carbone

### 3. Économie

- PIB total
- PIB par habitant (GDP per capita)

### 4. Social & démographie

- Population
- Accès à l'électricité (%)

## 4 Architecture data & pipeline

Afin d'assurer une gestion fiable, automatisée et reproductible des données, le projet repose sur une architecture data simple de type pipeline ETL (Extract – Transform – Load).

Cette architecture permet de collecter les données depuis plusieurs sources ouvertes, de les nettoyer et les transformer, puis de les centraliser dans un dataset unique (CSV/Parquet) avant leur exploitation analytique et leur visualisation.

Le flux de traitement suit les étapes suivantes :

### Étapes du pipeline

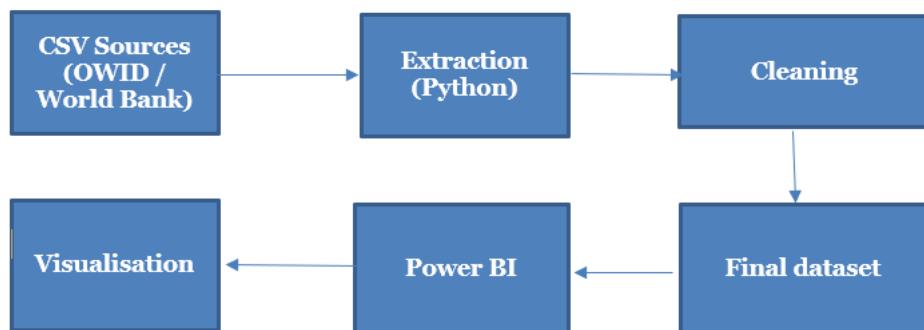


FIGURE 1 – Architecture globale.

#### 1. Extraction (E)

Les datasets sont téléchargés au format CSV depuis les plateformes Our World in Data et World Bank, puis importés automatiquement via des scripts Python.

#### 2. Cleaning (nettoyage)

Les données sont préparées afin d'améliorer leur qualité :

- suppression des valeurs manquantes
- gestion des doublons
- harmonisation des noms de pays
- correction des types de variables
- filtrage des années incohérentes

### **3. Transformation (T)**

Les données sont transformées pour créer des indicateurs métiers :

- calcul des pourcentages renouvelables et fossiles
- calcul du CO<sub>2</sub> par habitant
- calcul du GDP par habitant
- normalisation des unités

### **4. Merge (intégration)**

Tous les datasets sont fusionnés via les clés communes :

- country
- year

Afin de créer un fichier consolidé unique.

### **5. Load (L) – Stockage**

Les données finales sont stockées dans un fichier consolidé unique (CSV/Parquet) pour faciliter l'analyse et la visualisation.

### **6. Visualisation & analyse**

Les données sont exploitées dans un dashboard interactif (Power BI / Tableau / Python) permettant :

- comparaison entre pays
- suivi temporel
- analyse des corrélations
- aide à la décision

Cette architecture garantit :

- la reproductibilité
- l'automatisation
- la qualité des données
- la scalabilité du projet

## **5 Préparation et nettoyage des données**

Une phase de préparation est mise en place afin d'assurer la qualité et la fiabilité des données avant l'analyse.

- Traitement des valeurs manquantes (suppression ou interpolation)
- Suppression des doublons
- Harmonisation des noms de pays et des dates
- Normalisation des unités (% , TWh, tonnes de CO<sub>2</sub>)
- Conversion des types de données (numérique, date)
- Standardisation des clés communes (*country*, *year*) pour faciliter les jointures

Ces étapes garantissent des données propres, cohérentes et comparables.

## **6 Modélisation des données**

Les données nettoyées et fusionnées seront centralisées dans un dataset analytique unique (CSV/Parquet).

Chaque enregistrement sera identifié par les clés *country* et *year* et contiendra les indicateurs suivants :

- Indicateurs énergétiques (renewable %, fossile %, production électricité...)
- Indicateurs environnementaux (CO<sub>2</sub> total, CO<sub>2</sub>/habitant)
- Indicateurs socio-économiques (population, PIB, accès à l'électricité)

Cette structure simplifie l'analyse et permet une intégration directe avec Power BI pour le calcul des KPI et la création du dashboard.

## 7 Analyses prévues

Les analyses suivantes seront réalisées à partir du dataset consolidé :

- Tendances temporelles des énergies renouvelables, fossiles et émissions de CO<sub>2</sub>
- Comparaisons entre pays et régions
- Calcul d'indicateurs par habitant
- Corrélations entre consommation énergétique, émissions carbone et PIB
- Segmentation des pays selon leur niveau de transition énergétique
- Identification des pays les plus avancés et les plus dépendants aux énergies fossiles

Ces analyses permettront d'extraire des insights exploitables pour la prise de décision et l'orientation des politiques énergétiques.

## 8 Visualisation & Dashboard

Les résultats seront présentés via un **dashboard interactif** conçu pour faciliter l'interprétation et la prise de décision.

Le dashboard comprendra :

- Cartes KPI (renouvelables %, fossiles %, CO<sub>2</sub>, PIB/habitant)
- Graphiques d'évolution temporelle (line charts)
- Comparaisons entre pays (bar charts)
- Cartes géographiques (maps)
- Filtres dynamiques (pays, année, région)
- Un storytelling visuel guidant l'utilisateur vers les principaux insights

**Objectif :** transformer les données propres et consolidées en informations claires, visuelles et décisionnelles.

## 9 Automatisation & outils techniques

- Python pour l'extraction, le nettoyage et la transformation des données
- Bibliothèques : Pandas, NumPy, ...
- Stockage et consolidation dans un fichier CSV final
- Visualisation avec Power BI ou Tableau
- Scripts automatisés pour rejouer le pipeline (ETL)
- Versionnement du code avec Git et GitHub

## 10 Livrables

À l'issue du projet, les livrables suivants seront fournis afin de garantir la transparence, la réutilisabilité et la documentation complète du travail :

- **Dépôt GitHub** : code source, scripts, notebooks et documentation technique.
- **Scripts Python** : extraction, nettoyage et transformation des données.
- **Dataset final consolidé** : fichier CSV ou Parquet prêt à l'analyse.
- **Dashboard interactif** : Power BI ou Tableau avec KPI, graphiques et cartes.
- **Rapport analytique** : Data storytelling détaillant la démarche, les analyses et les recommandations.
- **Support de présentation** : slides et démonstration du dashboard.
- **Jira** : suivi des tâches et gestion de projet (planning, priorisation, jalons).
- **Confluence** : documentation centralisée du projet, procédures, méthodologie et résultats pour consultation par l'équipe ou les décideurs.

## 11 Risques& limites

- Valeurs manquantes dans certains pays
- Données hétérogènes entre sources
- Comparabilité limitée selon les méthodologies
- Données annuelles uniquement (pas temps réel)
- Contraintes de temps pour analyses avancées

## 12 Résultats attendus & valeur métier

- Suivi clair de la transition énergétique mondiale
- Comparaison renouvelables vs fossiles
- Analyse de l'impact sur les émissions de CO2
- Identification des pays performants ou en retard
- Aide à la décision pour politiques énergétiques durables
- Dashboard interactif facilitant l'exploration des données

