



Projet de spécialisation Data Business Intelligence

21 Novembre 2025
Ines Krissaane

Bordeaux - Lyon - Paris

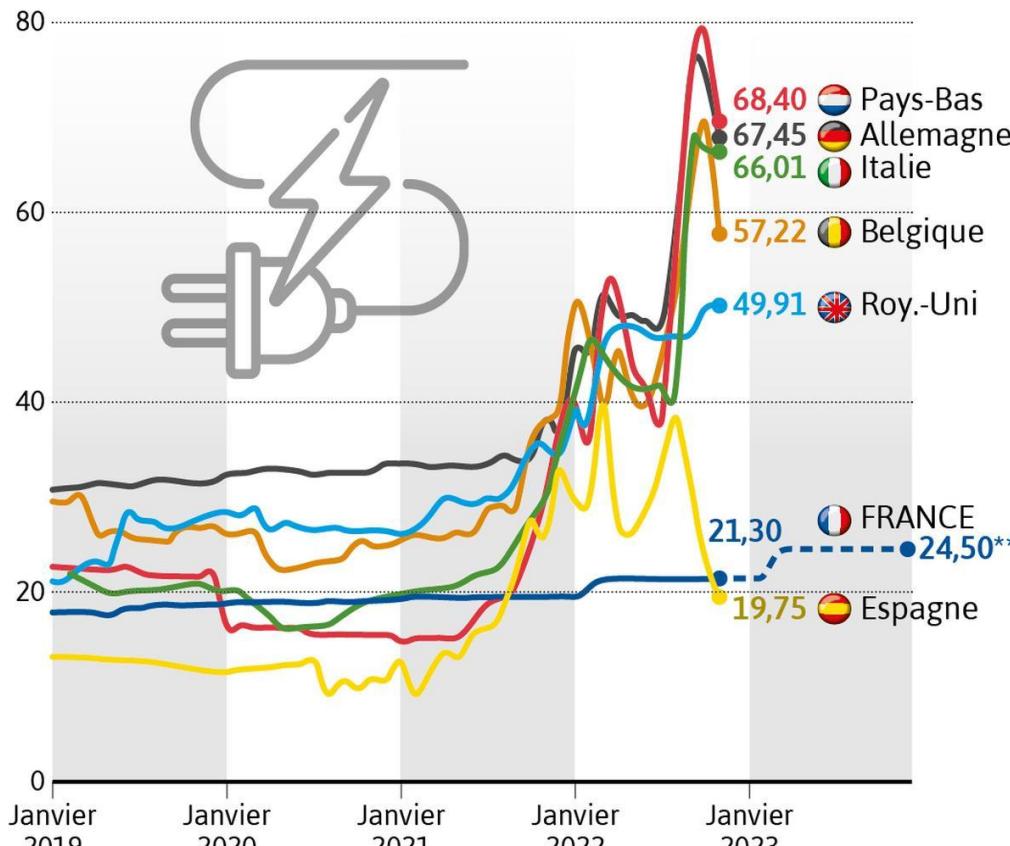
Data & Business Intelligence

Type d'intervention	Description	Date
Module 1	Module de spécialisation	29/09/2025
Module 2	Module de spécialisation	30/09/2025
Module 3	Module de spécialisation	01/10/2025
Module 4	Module de spécialisation	19/01/2025
Projet 1	Kick-Off + Workshop	21/11/2025
Projet 1	Workshop & Follow-up	12/12/2025
Projet 1	Workshop & soutenance	19/12/2025
Projet 2	Kick-Off + Workshop	20/01/2026
Projet 2	Workshop & Follow-up	26/01/2026
Projet 2	Workshop & soutenance	29/01/2026
Projet 3	Kick-Off + Workshop	20/03/2026
Projet 3	Workshop & Follow-up	03/04/2026
Projet 3	Workshop & soutenance	17/04/2026

Projet 1 - Prédiction du prix de l'électricité en Europe à partir de données multi-sources

Le prix de l'électricité en Europe pour les particuliers

Évolution en centimes d'€ par kWh*



*Kilowatt-heure. **prévisions. Source: Hello watt.

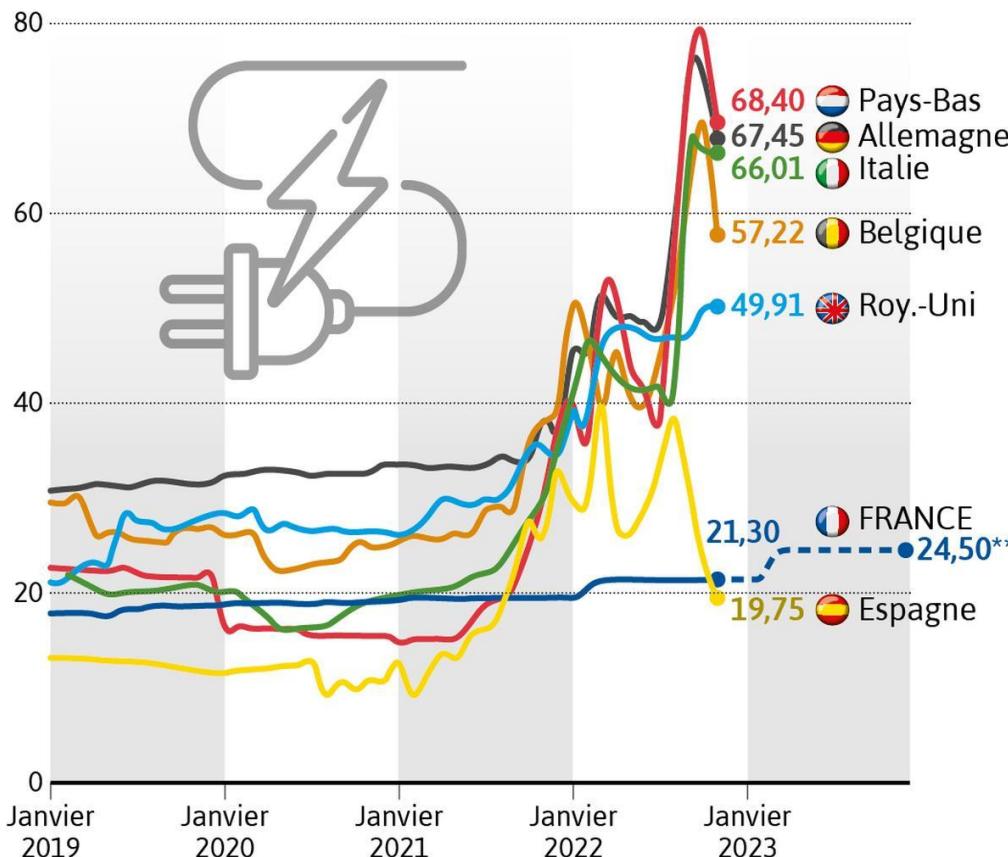
VISACTU

<https://www.sudouest.fr/economie/energie/infographie-quel-est-le-prix-de-l-electricite-en-europe-pourquoi-la-france-n-est-pas-a-plaindre-26064854.php>

Projet 1 - Prédiction du prix de l'électricité en Europe à partir de données multi-sources

Le prix de l'électricité en Europe pour les particuliers

Évolution en centimes d'€ par kWh*



- Renouvelables** – La production éolienne et hydraulique a été faible.
- Demande** – L'hiver froid et la reprise économique après le COVID-19 ont augmenté la consommation d'électricité et de chauffage.
- Nucléaire** – Certaines centrales nucléaires étaient en maintenance ou à puissance réduite.
- Géopolitique** – Les tensions Russie-Ukraine ont créé des incertitudes sur l'approvisionnement en gaz.

<https://www.sudouest.fr/economie/energie/infographie-quel-est-le-prix-de-l-electricite-en-europe-pourquoi-la-france-n-est-pas-a-plaindre-26064854.php>

Dans le contexte actuel de transition énergétique, la volatilité des prix de l'électricité constitue un enjeu économique et stratégique majeur pour les opérateurs, les gouvernements et les consommateurs.

Les fluctuations de ces prix dépendent de nombreux facteurs : production d'énergie renouvelable, conditions météorologiques, niveau de consommation, disponibilité du réseau ou encore coût des matières premières (gaz, charbon, etc.).

Projet 1 - Objectifs

L'objectif de ce projet est de prédire le **prix journalier de l'électricité day-ahead (€/MWh)** pour plusieurs pays européens à partir de variables explicatives issues de la production, de la consommation et des conditions météorologiques.

“Day-ahead” signifie « la veille pour le lendemain ».

Dans le marché de l'électricité, le prix day-ahead est le prix auquel l'électricité est achetée/vendue pour le jour suivant, sur la base d'enchères réalisées la veille.

Vous devez entraîner un modèle supervisé de séries temporelles tabulaires capable d'estimer, pour une date future t , le prix prévu de l'électricité, en utilisant comme entrées les informations disponibles jusqu'à $t-1$.

Projet 1 - Jeu de données

Le projet s'appuie sur les données publiques https://data.open-power-system-data.org/time_series/.

Ce jeu de données contient :

- les prix de l'électricité day-ahead (€/MWh) -> variable cible
- les volumes de production (totale, renouvelable, nucléaire, fossile),
- les données de consommation par pays,
- et des indicateurs météorologiques indirects (production solaire liée à l'ensoleillement, production eolienne liée à la vitesse du vent).
- Les fichiers sont disponibles en format CSV, avec une granularité horaire ou journalière, et couvrent plusieurs années de données (2015–2023 selon les pays).

Particularités du jeu de données

- Résolutions temporelles multiples : 15 min, 30 min, 60 min → nécessité de resampler les données pour obtenir des moyennes journalières.
- Données parfois manquantes (NaN) → traitement par ffill() ou bfill().
- Les prix day-ahead ne sont pas disponibles pour tous les pays.
- Plusieurs zones par pays (ex. Allemagne : DE_50hertz, DE_amprion, DE_tennet, DE_transnetbw ; Allemagne-Luxembourg : DE_LU).
- Les colonnes disponibles diffèrent selon les pays.
- Les prix ne sont pas disponibles pour toutes les années.

Projet 1 - Objectifs pédagogiques

- **Explorer et intégrer des données hétérogènes (données externes?)**
- **Mettre en œuvre des techniques avancées de feature engineering**
 - création de variables temporelles (lags, moyennes mobiles, tendances),
 - agrégations par jour et pays,
 - interactions entre production et conditions météo.
- **Concevoir, entraîner et optimiser un modèle de machine learning**
- **Évaluer et interpréter les résultats**
 - à l'aide de métriques adaptées (RMSE, MAE)
- **Communiquer et visualiser les résultats avec un dashboard interactif** permettant de :
 - visualiser les prédictions du modèle comparées aux valeurs réelles,
 - filtrer les résultats par pays ou période,
 - mettre en évidence les périodes d'erreur ou de forte volatilité,
 - analyser les variables les plus influentes dans les variations du prix.

1. Code et pipeline reproductible
2. Rapport écrit (max 10 pages) : Data canvas, présentation du jeu des données et des variables, Méthodologie, résultats (Dashboard), interprétation, discussion.
3. Dashboard interactif
4. Présentation orale : 15 minutes

Vous êtes autorisés à utiliser des jeux de données externes (documenter dans le rapport, donner les sources).

Utilisation de l'intelligence artificielle

L'utilisation d'outils d'intelligence artificielle est autorisée. Toutefois :

1. Chaque membre doit contribuer activement au projet.
2. Chaque étudiant du groupe doit maîtriser l'intégralité des aspects techniques, scientifiques et méthodologiques du projet (code, données, interprétations, résultats, visualisations, etc.).
3. Lors de la soutenance orale, tout membre doit être capable d'expliquer clairement le travail réalisé, les choix effectués, et de reproduire le raisonnement menant aux résultats.
4. Si un seul membre du groupe ne maîtrise pas les principales méthodes utilisées dans le projet, la note du groupe sera 0.

Accent sur la compréhension et l'interprétation

L'évaluation met l'accent sur :

- La qualité de l'analyse, la rigueur méthodologique et la clarté des explications ;
- L'interprétation des résultats, leur cohérence et leur mise en perspective ;
- La justification des choix techniques et scientifiques.

Les valeurs numériques finales (scores, métriques, performances) ne constituent pas le critère principal d'évaluation.

La compréhension, la pertinence des approches et la capacité à expliquer les résultats priment sur l'obtention d'un “bon score”.

Délais et rendus

Tout rendu (rapport, code, présentation, dashboard, etc.) doit être transmis avant la date limite indiquée. Aucun rendu en retard ne sera accepté.

Tout rendu remis après la deadline recevra la note de 0, sans exception.

Attentes pour les livrables

Le travail doit être original, documenté et reproductible.

Les sources de données, de code et d'outils externes doivent être clairement citées.

Toute forme de plagiat, d'utilisation non transparente d'IA, ou de reproduction non justifiée de code externe entraînera une note de 0.