**Household Energy Consumption**

**Suivi quotidien de la consommation d’énergie des ménages, de la température et de l’utilisation en heures de pointe**

**À propos du jeu de données**

🏡 **Household Energy Consumption – Avril 2025 (90 000 enregistrements)**

**Aperçu**

Ce jeu de données présente des enregistrements détaillés de consommation d’énergie de différents foyers durant le mois d’avril 2025.  
Avec **90 000 lignes** et plusieurs variables telles que la température, la taille du ménage, l’utilisation de la climatisation et la consommation en heures de pointe, il est idéal pour des analyses **de séries temporelles**, des projets **d’apprentissage automatique** et des recherches sur la **durabilité énergétique**.

**Description des colonnes**

| **Column Name** | **Data Type Category** | **Description** |
| --- | --- | --- |
| **Household\_ID** | Categorical (Nominal) | Identifiant unique de chaque ménage |
| **Date** | Datetime | Date de l’enregistrement de consommation |
| **Energy\_Consumption\_kWh** | Numerical (Continuous) | Consommation totale d’énergie du ménage (kWh) |
| **Household\_Size** | Numerical (Discrete) | Nombre d’individus dans le ménage |
| **Avg\_Temperature\_C** | Numerical (Continuous) | Température moyenne journalière (°C) |
| **Has\_AC** | Categorical (Binary) | Indique si le ménage possède une climatisation (Yes/No) |
| **Peak\_Hours\_Usage\_kWh** | Numerical (Continuous) | Consommation d’énergie durant les heures de pointe (kWh) |

**Résumé du dataset**

* **Rows** : 90 000
* **Plage temporelle** : 1er avril – 30 avril 2025
* **Granularité** : Journalière par ménage
* **Couverture** : Simulation mondiale
* **Format** : CSV (Comma-Separated Values)

**Bibliothèques utilisées**

**🔍 Manipulation et analyse de données**

* **pandas** : Lecture, nettoyage et transformation des données tabulaires
* **numpy** : Opérations numériques et manipulation de tableaux

**📊 Visualisation de données**

* **matplotlib** : Graphiques statiques (lignes, barres, histogrammes, etc.)
* **seaborn** : Visualisations statistiques (heatmaps, boxplots, etc.)
* **plotly** : Graphiques interactifs (séries temporelles, camemberts, barres, nuages de points, etc.)

**📈 Modélisation / Machine Learning**

* **scikit-learn** : Prétraitement, régression, classification, clustering
* **xgboost / lightgbm** : Modèles de gradient boosting pour une meilleure précision

**🧹 Prétraitement des données**

* **sklearn.preprocessing** : Encodage, mise à l’échelle, normalisation
* **datetime / pandas** : Conversion et manipulation de dates

**🧪 Évaluation des modèles**

* **sklearn.metrics** : Accuracy, MAE, RMSE, R², matrice de confusion, etc.

✅ Ces bibliothèques constituent une boîte à outils complète pour l’analyse, la modélisation et la visualisation des données.

**Cas d’utilisation potentiels**

🔮 **Prévision & séries temporelles**

* Prédire la consommation future selon les tendances passées et la météo
* Identifier les schémas saisonniers et journaliers

💡 **Analyse d’efficacité énergétique**

* Comparer la consommation des foyers avec et sans climatisation
* Étudier l’efficacité énergétique selon la taille des ménages

🌡️ **Études sur l’impact climatique**

* Analyser l’effet de la température sur la consommation électrique
* Modéliser l’impact du changement climatique sur la demande résidentielle

🔌 **Gestion de la charge de pointe**

* Prédire et gérer la demande durant les heures de pointe
* Contribuer à la recherche sur les réseaux intelligents (*smart grids*) et la tarification dynamique

🧠 **Projets Machine Learning**

* Régression / classification pour prédire la consommation
* Clustering des ménages selon les schémas de consommation
* Détection d’anomalies pour identifier des défauts

**Exemples de projets de départ**

* Prévision de séries temporelles avec **Facebook Prophet** ou **ARIMA**
* Régression avec **XGBoost** ou **LightGBM**
* Classification des comportements des foyers *AC vs non-AC*
* Systèmes de recommandation pour économiser l’énergie
* Heatmaps reliant température et consommation