

Explication des Jeux de Données et Leurs Caractéristiques

Dans cette section, nous fournirons des explications détaillées pour chaque jeu de données, y compris leurs sources, les principales caractéristiques et leur pertinence pour la détection des anomalies dans la consommation d'énergie. De plus, les liens vers les API et les jeux de données sont fournis pour référence.

Sources de Données et Liens API sur la Consommation d'Énergie

Les liens suivants mènent aux API et aux jeux de données utilisés dans ce projet :

- [Courbe de Charge ELD/GRD Régional GRTgaz Teréga](#)
- [ecomix](#)
- [Consommation Nationale Horaire de Gaz - Données Provisoires GRTgaz Teréga](#)
- [Consommation Brute Corrigée - Client Direct](#)
- [Consommation Quotidienne Brute Régionale](#)
- [Consommation Annuelle d'Électricité et de Gaz par Département et par Secteur d'Activité](#)
- [Mouvements sociaux depuis 2002 : motif et participation](#)
- [Niveaux de Prix par Commune pour les Tarifs Réglementés de Vente de Gaz Naturel d'Engie](#)
- [Data Connect – Enedis DataHub](#)
- [Prix des Énergies en France](#)
- [Prix des Énergies depuis 2009](#)
- [Traces des Réseaux de Chaleur et de Froid](#)
- [Croissance Démographique, Réchauffement Climatique, Besoins Énergétiques - Comment vont évoluer les besoins en eau dans le monde ?](#)
- [Production Régionale Annuelle des Énergies Renouvelables 2008 à 2021](#)
- [Visualisation des Données sur les Prix des Carburants](#)
- [Production d'Électricité par Filière et Coûts de Production au Pas Horaire](#)
- [DPE Logements Existants depuis Juillet 2021](#)
- [DPE Logements Neufs depuis Juillet 2021](#)
- [ElecDom - Données de Consommation Annuelle](#)
- [Énergie et Patrimoine Communal - 5 Bâtiments](#)
- [Données Locales d'Énergie](#)
- [Consommation horaire de gaz](#)

Jeux de Données choisir & Explication Détaillée et Analyse de chaque donnée

1. Prévisions Météo France à 36h

Utilité : Les données météorologiques sont essentielles pour comprendre les variations de consommation d'énergie, car elles peuvent fortement influencer l'utilisation de chauffage ou de climatisation.

Lien : [Donnée prévisions météo](#)

Colonnes Pertinentes :

- **date_prevision** : Date et heure de la prévision, utile pour aligner avec les données de consommation.
- **maximum_temperature_at_2_metres** : Température maximale prévue, impacte la demande énergétique.
- **minimum_temperature_at_2_metres** : Température minimale prévue, impacte la demande énergétique.
- **relative_humidity_2m** : Humidité relative, peut affecter le confort thermique et la consommation.
- **wind_speed** : Vitesse du vent, influence la sensation thermique et la consommation.

Applications :

- **Détection des Anomalies** : Identifier les anomalies de consommation dues à des variations météorologiques extrêmes.
- **Prédiction des Incidents** : Prévoir les pics de consommation lors de conditions météorologiques extrêmes.

Après fusions de dataset nous avons décidé de garder ces colonnes :

TempMax_Deg
TempMin_Deg
Wind_kmh
Wet_percent
Visibility_km
CloudCoverage_percent
Dayduration_hour
date
region
opérateur

3. Consommation Quotidienne de Gaz Régionale

Utilité : Permet d'analyser les tendances régionales de la consommation de gaz.

Lien : [consommation quotidienne de gaz par region](#)

Colonnes Pertinentes :

- **date** : Date de consommation.
- **consommation_journaliere_mwh_pcs** : Consommation journalière en MWh, variable cible pour la détection des anomalies.
- **region** : Région, pour l'analyse géographique.
- **statut** : Statut de la consommation (définitive ou provisoire).

Applications :

- **Détection des Anomalies** : Identifier des anomalies de consommation à l'échelle régionale.
- **Détection des Anomalies** : Identifier des anomalies dans la consommation horaire.
- **Optimisation des Ressources** : Aider à planifier la distribution de gaz.
- **Prédiction des Incidents** : Prévoir des problèmes de distribution dans certaines régions.

4. Prix des Énergies en France

Utilité : Suivre les variations des prix de l'énergie pour comprendre leur impact sur la consommation.

Lien : [Prix des énergies en france](#)

Colonnes Pertinentes :

- **periode** : Période de mesure.
- **elec_ttc** : Prix de l'électricité TTC.
- **gaz_ttc** : Prix du gaz TTC.

Applications :

- **Détection des Anomalies** : Identifier des changements de consommation en réponse à des variations de prix.
- **Optimisation des Ressources** : Ajuster les stratégies de consommation en fonction des prix.

5. Mouvements Sociaux

Utilité : Comprendre l'impact des grèves et mouvements sociaux sur la consommation d'énergie.

Lien : [Mouvement sociaux](#)

Colonnes Pertinentes :

- **date_de_debut** : Date de début de l'événement.
- **date_de_fin** : Date de fin de l'événement.
- **motif** : Motif de la grève, pour comprendre le contexte.
- **nombre_grevistes** : Nombre de grévistes, pour quantifier l'impact.
- **taux_grevistes** : Taux de grévistes, pour évaluer l'ampleur.

Applications :

- **Détection des Anomalies** : Identifier des baisses ou hausses de consommation dues aux grèves.
- **Prédiction des Incidents** : Prévoir les impacts futurs des mouvements sociaux.

6. les Prix des Carburants en France - Flux Quotidien

Les données sur les prix des carburants fournissent des informations détaillées sur les prix des différents carburants à travers la France, avec des mises à jour quotidiennes.

Lien : [prix carburants](#)

Colonnes Pertinentes

- **date** : La date de mise à jour du prix est essentielle pour analyser les tendances temporelles des prix des carburants.
- **prix_valeur** : La valeur du prix est la principale variable quantitative à analyser.
- **prix_nom** : Le type de carburant (Gazole, SP95, etc.) permet de segmenter les analyses par type de carburant.
- **ville** : La ville fournit des informations géographiques pour analyser les variations locales des prix.
- **dep_name** et **reg_name** : Le département et la région permettent des analyses à des niveaux géographiques plus larges.
- **services_service** : Les services offerts à la station peuvent influencer la fréquentation et potentiellement les prix.

- **rupture_nom** et **rupture_debut** : Les informations sur les ruptures de stock sont cruciales pour détecter des anomalies et comprendre les variations de prix.
- **horaires_automate_24_24** : Indique si la station est ouverte 24h/24, ce qui peut affecter la disponibilité des carburants et les prix.

Applications pour la Détection d'Anomalies et la Prédiction

1. Détection des Anomalies :

- **Fluctuations de Prix** : Détecter des augmentations ou diminutions soudaines des prix du carburant.
- **Ruptures de Stock** : Identifier les périodes de pénurie de carburant et leur impact sur la consommation énergétique alternative (gaz, électricité).

2. Prédiction des Incidents :

- **Prix et Consommation** : Analyser comment les variations de prix influencent la consommation de carburant et prévoir les pics de demande.
- **Horaires et Accessibilité** : Utiliser les horaires d'ouverture pour prévoir les périodes de forte et faible demande.

Conclusion

Les jeux de données choisis sont pertinents pour le projet de détection d'anomalies dans la consommation énergétique. Voici un résumé des colonnes à intégrer dans votre dataset et les applications potentielles :

1. Prévisions Météo France :

- **Colonnes** : date_prevision, maximum_temperature_at_2_metres, minimum_temperature_at_2_metres, relative_humidity_2m, wind_speed.
- **Applications** : Analyser l'impact des conditions météorologiques sur la consommation.

2. Consommation Horaire Définitive de Gaz :

- **Colonnes** : date, consommation_journaliere_mwh_pcs, secteur_d_activite.
- **Applications** : Détection des anomalies horaires, optimisation de la distribution.

3. Consommation Quotidienne de Gaz Régionale :

- **Colonnes** : date, consommation_journaliere_mwh_pcs, region, statut.
 - **Applications** : Analyser les tendances régionales et détecter des anomalies.
4. **Prix des Énergies en France :**
- **Colonnes** : periode, elec_ttc, gaz_ttc.
 - **Applications** : Comprendre l'impact des variations de prix sur la consommation.
5. **Mouvements Sociaux :**
- **Colonnes** : date_de_debut, date_de_fin, motif, nombre_grevistes, taux_grevistes.
 - **Applications** : Analyser l'impact des grèves sur la consommation énergétique.
6. **Prix des Carburants en France - Flux Quotidien :**
- **Colonnes** : id, cp, adresse, ville, horaires, rupture, prix_maj, prix_nom, prix_valeur, com_arm_code, com_arm_name, epci_code,, epci_name, dep_code, dep_name, reg_code, reg_name, services_service
 - **Applications** : Analyser l'impact des grèves sur la consommation énergétique.

En combinant ces jeux de données, vous pourrez non seulement détecter les anomalies de consommation mais aussi comprendre les facteurs contextuels influençant ces anomalies, permettant ainsi de prédire et d'optimiser la consommation énergétique.

7. Consommation Quotidienne Brute Régionale

Ce jeu de données contient des données de consommation énergétique quotidienne pour diverses régions. Il offre une vue complète de l'utilisation de l'électricité et du gaz, segmentée par différents secteurs tels que l'agriculture, l'industrie et le résidentiel.

Référence d'Image :

+ Créer une base de données

Rechercher des espaces de noms

DBÉnergie

Consommation quotidi...

Consommation-horair...

Consommation-quotidi...

Données de consomm...

Niveaux-de-prix-par-c...

donnée-local-energie

écomix

prix-des-carburants

prod-region-annuelle...

échantillon_mflix

DBEnergy.Consommation brute quotidienne régionale

TAILLE DE STOCKAGE : 20 KO

TAILLE DES DONNÉES LOGIQUES : 7,53 KO

TOTAL DES DOCUMENTS : 20

TAILLE TOTALE DES INDEX : 20 KO

Trouver

Index

Anti-modèles de schéma

Agrégation

Index de recherche

Générer des requêtes à partir du langage naturel dans Compass

Filtre

Type a query: { field: 'value' }

RÉSULTATS DE LA REQUÊTE : 1-20 SUR 20

_identifiant: ObjectId(' 667fc8de65b11852abad0128 ')

date_heure: " 2021-06-28T11:30:00+00:00 "

date: " 2021-06-28 "

heure: " 13:30 "

code_insee_region: " 52 "

région: " Pays de la Loire "

consommation_brute_gaz_grtgaz: nul

statut_grtgaz: nul

consommation_brute_gaz_terega: nul

statut_terega: nul

consommation_brute_gaz_totale: nul

consommation_brute_electricite_rte: 2958

statut_rte: " Consolidé "

consommation_brute_totale: nul

_identifiant: ObjectId(' 667fc8de65b11852abad0129 ')

date_heure: " 2021-06-28T11:30:00+00:00 "

date: " 2021-06-28 "

heure: " 13:30 "

code_insee_region: " 53 "

région: " Bretagne "

consommation_brute_gaz_grtgaz: nul

statut_grtgaz: nul

consommation_brute_gaz_terega: nul

Caractéristiques :

- date_heure: Date et heure de la saisie des données.
- code_insee_region: Code INSEE de la région.
- region: Nom de la région.
- consommation_brute_electricite_rte: Valeur brute de la consommation d'électricité.
- statut_grtgaz: Statut de GRTgaz.
- consommation_brute_gaz_terega: Valeur brute de la consommation de gaz.

8. Consommation Quotidienne de Gaz

Ce jeu de données fournit des données horaires de consommation de gaz, y compris des informations sur les opérateurs et les secteurs d'activité. Il aide à comprendre les schémas d'utilisation du gaz tout au long de la journée.

Référence d'Image :

+ Créer une base de données

Q

Rechercher des espaces de noms

DBÉnergie

Consommation quotidi...

Consommation-horair...

Consommation-quoti...

Données de consomm...

Niveaux-de-prix-par-c...

donnée-local-energie

écomix

prix-des-carburants

prod-region-annuelle-...

échantillon_mflix

DBEnergy.Consommation-quotidienne-gaz

TAILLE DE STOCKAGE : 24 KO

TAILLE DES DONNÉES LOGIQUES : 11,41 KO

TOTAL DES DOCUMENTS : 20

TAILLE TOTALE DES INDEX : 20 KO

Trouver

Index

Anti-modèles de schéma

Agrégation

Index de recherche

Générer des requêtes à partir du langage naturel dans Compass

Filtre

Type a query: { field: 'value' }

RÉSULTATS DE LA REQUÊTE : 1-20 SUR 20

```

_identifiant: IDObjet(' 667fc738f6c9d7a6f2b91363 ')
date : " 2022-05-29 "
opérateur : " GRTgaz "
secteur_d_activité : « GRD/ELD »
région : " Grand Est "
00_00_00 : 915.8687
01_00 : 920.67043
02_00 : 918.714844
03_00 : 1015.063388
04_00 : 1211.500894
05_00_00 : 1595.616774
06_00_00 : 1158.458995
07_00 : 1157.757154
08_00 : 1215.434537
09_00 : 1303.719463
10_00 : 1363.096929
11_00_00 : 1340.05922
12_00_00 : 1213.782191
13_00 : 1107.573008
14_00 : 1010.063629
15_00 : 944.55784
16_00 : 922.341228
17_00 : 974.442598
18_00 : 1151.059333
19_00_00 : 1297.508782

```

Caractéristiques :

- date: Date de la saisie des données.
- operateur: Nom de l'opérateur.
- secteur_d_activite: Secteur d'activité.
- region: Nom de la région.
- Valeurs de consommation horaire (00_00, 01_00, etc.).

9. Donnée Locale Energie

Ce jeu de données comprend des données énergétiques locales, couvrant diverses mesures liées à la consommation et à la production au niveau local. Il fournit une ventilation détaillée de l'utilisation de l'énergie par secteur et par région.

Référence d'Image :

+ Créer une base de données

Q

Rechercher des espaces de noms

DBÉnergie

Consommation quotidi...

Consommation-horair...

Consommation-quotidi...

Données de consomm...

Niveaux-de-prix-par-c...

donnée-local-energie

écomix

prix-des-carburants

prod-region-annuelle-...

échantillon_mflix

DBEnergy.donnée-locale-energie

TAILLE DE STOCKAGE : 4,80 MO

TAILLE DES DONNÉES LOGIQUES : 19,79 MO

TOTAL DES DOCUMENTS : 51905

TAILLE TOTALE DES INDEX : 1,44 MO

Trouver

Index

Anti-modèles de schéma

Agrégation

Index de recherche

Générer des requêtes à partir du langage naturel dans Compass

Filtre

Type a query: { field: 'value' }

```

_identifiant: ObjectId(' 664742d5eae6c7672f67cca0 ')
OPÉRATEUR : " SRD "
ANNEE : 2017
FILIERE : " E "
TAP_ : " Région "
CODE : " 24 "
CONSOA : " 0,00 "
PDLA : 0
INDQUALA : " 0,900000 "
CONSOI : " 0,00 "
PDLI : 0
INDQUALI : " 0,900000 "
CONST_ : " 0,00 "
PDLT : 0
INDEQUAL_ : " 0,900000 "
CONSOR : " 62,37 "
PDLR : 11
INDEPEN_ : " 0,500000 "
THERMOR : " 0,00 "
PARTR : " 0,00 "
CONSONA : " 0,00 "
PDLNA : 0
INDQUALNA : " 0,500000 "

```

PRÉCÉDENT

1 à 20 parmi de nombreux résultats

Caractéristiques :

- OPERATEUR: Nom de l'opérateur.
- ANNEE: Année de saisie des données.
- FILIERE: Secteur énergétique.
- CONSOA: Consommation dans l'agriculture (en MWh).
- PDLA: Nombre de points de livraison dans l'agriculture.
- CONSOI: Consommation dans l'industrie (en MWh).
- PDLI: Nombre de points de livraison dans l'industrie.

10.Ecomix (voir avec alimou si donnée charger dans mongo)

Ecomix combine différents types de données énergétiques, fournissant une vue holistique de la production et de la consommation d'énergie. Ce jeu de données est crucial pour analyser l'écosystème énergétique global et identifier les tendances.

Référence d'Image :

Caractéristiques :

11.Prix des Carburants

Ce jeu de données contient des statistiques mensuelles sur les prix des carburants, essentielles pour comprendre les tendances des coûts et leur impact sur les schémas de consommation d'énergie.

Référence d'Image :

The screenshot displays the DBEnergy.prix-des-carburants interface. On the left, a sidebar lists various data categories under 'DBÉnergie', with 'prix-des-carburants' highlighted. The main panel shows search statistics (TAILLE DE STOCKAGE: 20 KO, TAILLE DES DONNÉES LOGIQUES: 2,21 KO, TOTAL DES DOCUMENTS: 20, TAILLE TOTALE DES INDEX: 20 KO) and a search bar. Below the search bar, a filter query is shown: 'Type a query: { field: 'value' }'. The results section, titled 'RÉSULTATS DE LA REQUÊTE : 1-20 SUR 20', displays five entries with their unique identifiers and associated metrics.

Identifiant	__id, reuse_id, metric_month, monthly_visit
IDObjet(' 664738a4bdcdb25a7d5e6e77 ')	" 1126410,635a9e6b04c1b9b37fb75a15,2022-10,700 "
IDObjet(' 664738a4bdcdb25a7d5e6e78 ')	" 1103413,635a9e6b04c1b9b37fb75a15,2022-11,2871 "
IDObjet(' 664738a4bdcdb25a7d5e6e79 ')	" 1100015,635a9e6b04c1b9b37fb75a15,2022-12,1494 "
IDObjet(' 664738a4bdcdb25a7d5e6e7a ')	" 217844,635a9e6b04c1b9b37fb75a15,2023-01,1974 "
ObjectId(' 664738a4bdcdb25a7d5e6e7b ')	" 229843,635a9e6b04c1b9b37fb75a15,2023-02,1589 "

Caractéristiques :

- metric_month: Mois de la saisie des données.
- monthly_visit: Statistiques des visites mensuelles.

12.Production Région Annuelle d'Énergie

Ce jeu de données fournit des données annuelles sur la production énergétique régionale, en mettant l'accent sur les sources renouvelables et non renouvelables. Il met en lumière les capacités de production et les contributions des différents types d'énergie dans chaque région.

Référence d'Image :

+ Créer une base de données

Q

Rechercher des espaces de noms

DBÉnergie

Consommation quotidi...

Consommation-horair...

Consommation-quotidi...

Données-de-consomm...

Niveaux-de-prix-par-c...

donnée-local-energie

ecomix

prix-des-carburants

prod-region-annuelle-...

sample_mflix

DBEnergy.prod-region-annuelle-enr

TAILLE DE STOCKAGE : 146 MO

TAILLE DES DONNÉES LOGIQUES : 2,35 MO

TOTAL DES DOCUMENTS : 208

TAILLE TOTALE DES INDEX : 24 KO

Trouver

Index

Anti-modèles de schéma

Agrégation

Index de recherche

Générer des requêtes à partir du langage naturel dans Compass

Filtre

Type a query: { field: 'value' }

```

_identifiant: IDObjet(' 667d2010cf1b75b6a9efbc48 ')
année: 2008
nom_insee_region: " Centre Val de Loire "
code_insee_region: 24
production_hydraulique_renouvelable: 189
production_bioénergies_renouvelable: 78.1
production_eolienne_renouvelable: 758
production_solaire_renouvelable: 0,4
production_electrique_renouvelable: 1025,5
production_gaz_renouvelable: 0
production_totale_renouvelable: 1025,5
région_géo_forme: " {"coordonnées": [[[2.936623684, 48.163455797], [2.963697564, 48.153076_ "
geo_point_region: " 47.48477569, 1.684391194 "

```

```

_identifiant: ObjectId(' 667d2010cf1b75b6a9efbc49 ')
année: 2008
nom_insee_region: " Nouvelle-Aquitaine "
code_insee_region: 75
production_hydraulique_renouvelable: 4318
production_bioénergies_renouvelable: 172,9
production_eolienne_renouvelable: 69
production_solaire_renouvelable: 17

```

PRÉCÉDENT

1 à 20 parmi de nombreux résultats

Caractéristiques :

- annee: Année de saisie des données.
- nom_insee_region: Nom INSEE de la région.
- production_hydraulique_renouvelable: Production hydraulique renouvelable.
- production_totale_renouvelable: Production totale renouvelable.

Les jeux de données décrits ci-dessus fournissent une base solide pour analyser la consommation d'énergie et détecter les anomalies. En exploitant ces jeux de données, il est possible d'obtenir des informations sur les schémas d'utilisation, d'identifier les tendances inhabituelles et d'améliorer les pratiques de gestion de l'énergie.