**Projet : Prévision de la Consommation Énergétique des Foyers au Royaume Uni**

**Contexte Général :** Vous allez travailler sur le jeu de données " Foyers au Royaume Uni ". Votre mission est de développer des modèles capables de prévoir la consommation électrique de différents types de foyers (segments ACORN). Ce projet vous amènera à explorer les données, à comprendre les comportements de consommation, à modéliser l'impact de la météo, et enfin à construire et évaluer des modèles de prévision à court et moyen termes.

**Objectifs et Déroulement du projet (par groupe) :** Chaque groupe se concentrera sur **trois segments ACORN** (qui vous seront attribués, disponibles dans le dossier data/02\_processed[[1]](#footnote-1)). Pour ce(s) segment(s) :

1. **Phase 1 : Analyse et Caractérisation**
   * Analyser de façon exploratoire les profils de consommation électrique typiques (journaliers, hebdomadaires, saisonniers si possible).
   * Quantifier l'influence des conditions météorologiques (notamment la température) sur la consommation.
   * Identifier les dynamiques temporelles clés (autocorrélation, saisonnalités). Cette phase est cruciale pour bien comprendre les données avant de modéliser.
2. **Phase 2 : Prévision à Court Terme (Horizon 48h - pas de 30 minutes)**
   * Développer un modèle pour prévoir la consommation électrique au pas de 30 minutes pour les 48 prochaines heures : pour la période du **13 janvier 2014 à 00:00 au 14 janvier 2014 à 23:30**.
   * Utiliser les données historiques antérieures à cette période pour l'entraînement et la validation de votre modèle.
   * Évaluer rigoureusement la performance de votre modèle.
3. **Phase 3 : Prévision à Moyen Terme (Horizon 1 mois - pas journalier)**
   * Développer un modèle pour prévoir la consommation électrique journalière agrégée pour un mois complet : du **13** **janvier au 13 février 2014**).
   * Utiliser les données historiques antérieures à ce mois pour l'entraînement et la validation.
   * Évaluer rigoureusement la performance de votre modèle.
4. **(Bonus) Phase 4 : Restitution Interactive**
   * Présenter vos analyses, modèles et résultats de prévision de manière interactive à l'aide d'un tableau de bord développé avec [**Streamlit**](https://docs.streamlit.io/) **ou Dash** (Streamlit est plus simple à prendre en main pour des débutants).

**Données Fournies :**

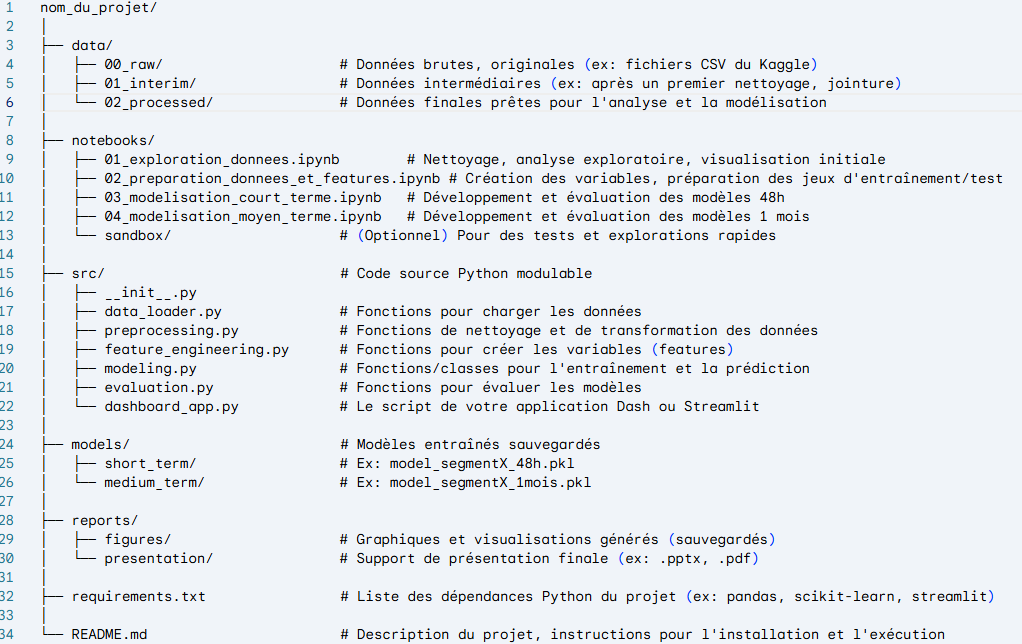
* Données de consommation électrique (prétraitées sous forme de profils moyens par segment ACORN). La période historique pour l'entraînement sera précisée (ex: données de 2012-2013).
* Données météorologiques historiques (température, humidité, vitesse du vent, etc.) synchronisées avec les données de consommation.
* Données des jours fériés
* Pour les périodes de prévision (13-14 janvier 2014 et du 13/01/2014 au 13/02/2014), vous utiliserez les données météorologiques réelles de ces périodes comme si elles étaient des prévisions parfaites.

**Critères d'Évaluation Principaux :**

1. **Qualité de l'Analyse Exploratoire (Phase 1) :** Profondeur de la compréhension des données, pertinence des visualisations et des interprétations.
2. **Rigueur de la Démarche de Modélisation (Phases 2 & 3) :** Choix **justifiés** des modèles, stratégie de validation correcte, feature engineering pertinent.
3. **Performance des Prévisions :** Qualité des prévisions mesurée par des métriques appropriées (RMSE).
4. **Clarté et Pertinence de la Restitution :** Qualité de la présentation orale, clarté des explications, et (pour le bonus) qualité, interactivité et utilité du tableau de bord Dash/Streamlit**.**
5. **Travail d'Équipe et Gestion de Projet :** Capacité à collaborer efficacement et à livrer un projet complet dans les temps.

**Quelques conseils :**

* **Commencez simple :** Implémentez toujours un modèle de référence simple (baseline) pour pouvoir juger de la plus-value des modèles plus complexes.
* **Itérez :** Ne cherchez pas la perfection du premier coup. Construisez une première version de vos modèles, évaluez-la, puis améliorez-la.
* **Communiquez :** Échangez régulièrement au sein de votre groupe, partagez vous les tâches et n'hésitez pas à solliciter les encadrants.
* **Soyez critiques :** Portez un regard critique sur vos propres résultats. Un modèle n'est jamais parfait.
* **Arborescence suggérée pour votre projet :**



**Bon projet à tous !**

1. Deux versions des fichiers sont disponibles en csv et en parquet. Utilisez de préférence « parquet » qui permet de conserver les types de données [↑](#footnote-ref-1)