

PREDIVIS: PRÉDICTION ET VISUALISATION

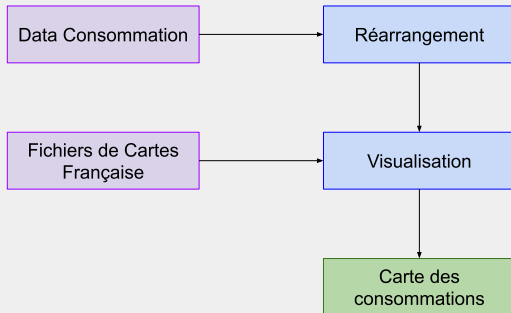
SUR LA CONSOMMATION ÉLECTRIQUE FRANÇAISE

LEROY NICOLAS
CHAKROUN MOHAMED MEHDI
DE MONTGOLFIER AXEL

UNIVERSITÉ DE MONTPELLIER

DECEMBER 9, 2022

VISUALISATION



Gestion de data :

Numpy : Calcul numérique.

Pandas : Gestion de dataframe.

Visualisation :

Dash : création d'applications web de tableaux de bord.

Plotly : Librairie de graphe.

DATAFRAME

	Année	Code INSEE de la commune	Nom de la commune	Consommation annuelle moyenne de la commune (MWh)
0	2019	76157	Canteleu	2.955
1	2019	76157	Canteleu	2.955
2	2019	76157	Canteleu	2.955
3	2019	76157	Canteleu	2.955
4	2019	76157	Canteleu	2.955
...
1603543	2021	94078	Villeneuve-Saint-Georges	4.135
1603544	2021	94078	Villeneuve-Saint-Georges	4.135
1603545	2021	94078	Villeneuve-Saint-Georges	4.135
1603546	2021	31588	Villeneuve-Tolosane	5.001
1603547	2021	31588	Villeneuve-Tolosane	5.001

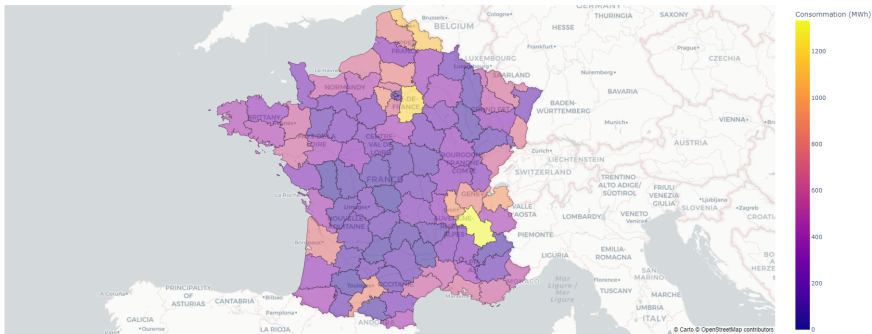
1603548 rows × 4 columns

DATAFRAME

Numero		Nom_dept	Consommation (MWh)
0	1	Ain	806.270
1	2	Aisne	354.479
2	3	Allier	148.999
3	4	Alpes de Haute-Provence	157.925
4	5	Hautes-Alpes	77.965
...
90	91	Essone	675.071
91	92	Hauts-de-Seine	125.132
92	93	Seine-Saint-Denis	148.686
93	94	Val-de-Marne	193.324
94	95	Val-d'Oise	655.549

95 rows x 3 columns

Average electricity consumption(MWh) in France by departments in 2018



```
Dash is running on http://127.0.0.1:8050/
```

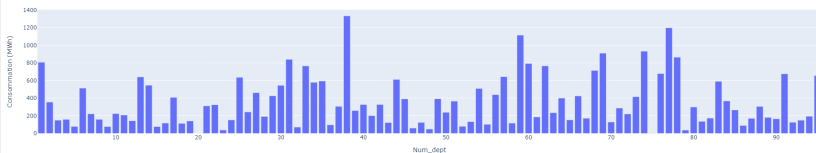
```
* Serving Flask app "visu_complet" (lazy loading)
* Environment: production
  WARNING: This is a development server. Do not use it in a production deployment.
  Use a production WSGI server instead.
* Debug mode: on
```

Average electricity consumption(MWh) in France by departments

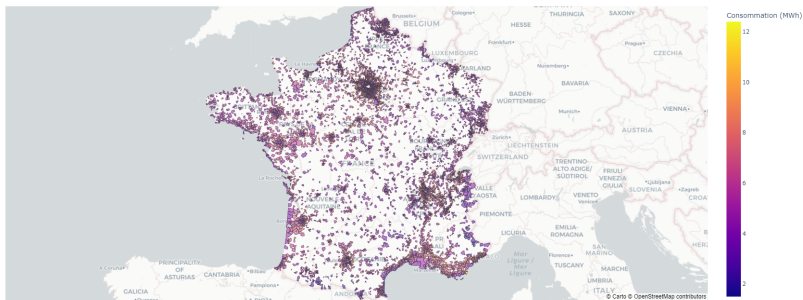
Select a Year:

☒ 2018 ☐ 2019 ☐ 2020 ☐ 2021

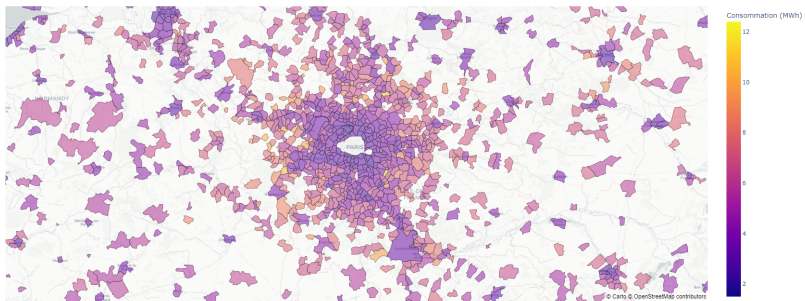
Histogramme de la consommation(MWh) moyenne par département en 2018



Carte de la consommation (MWh) par commune en 2018



Carte de la consommation (MWh) par commune en 2018



La consommation max du département 34 pour l'année 2018 est de 8.471 dans la commune de 301390	SUSSARGUES
La consommation min du département 34 pour l'année 2018 est de 2.024 dans la commune de 6762	GRANDE-MOTTE
La consommation max du département 34 pour l'année 2019 est de 8.405 dans la commune de 326100	Sussargues
La consommation min du département 34 pour l'année 2019 est de 1.986 dans la commune de 12375	La Grande-Motte
La consommation max du département 34 pour l'année 2020 est de 8.224 dans la commune de 253625	Sussargues
La consommation min du département 34 pour l'année 2020 est de 1.985 dans la commune de 1351	La Grande-Motte
La consommation max du département 34 pour l'année 2021 est de 8.447 dans la commune de 66689	Montferrier-sur-Lez
La consommation min du département 34 pour l'année 2021 est de 2.149 dans la commune de 31433	La Grande-Motte

Data consommation :

Site d'Enedis

(<https://data.enedis.fr/explore/dataset/consommation-annuelle-residentielle-par-adresse/information/>)

Fichier Carte :

Projet Github de gregoire david

(<https://france-geojson.gregoiredavid.fr/>)

PRÉDICTION

Traitement de données

- pandas : Bibliothèque d'analyse de données
- numpy : Bibliothèque d'outils mathématiques
- requests : Prise en charge des requêtes http

Prédiction

- prophet : Modèle de prédiction de séries temporelles
- statsmodels : Modèles statistiques divers

FONCTION : DATALOADING()

Input

filepath : L'endroit où est enregistré le fichier final.

Rôle

- Télécharger les deux jeux de données différents sur le site d'Enedis.
- Enlever les données inutiles et construire un index temporel unique adapté à pandas.
- Les mettre au même format et les concaténer. (interpoler linéairement)

FONCTION : PREPARE_DATA()

inputs

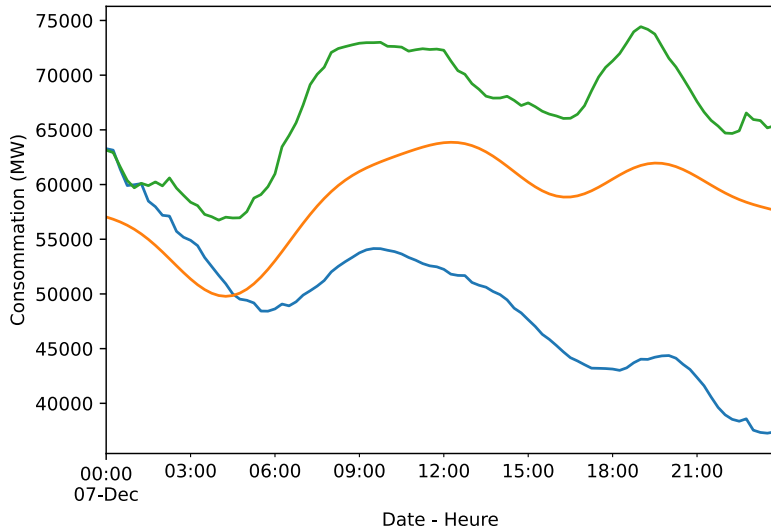
- `filepath(str)` : Chemin d'accès au tableau csv des données d'entraînements.
- `dateinitiale (str)` : Au format yyyy-mm-dd , date à partir de laquelle le modèle s'entraîne sur les données.
- `methode (str)` : Permet de choisir la méthode utilisée pour la prédiction ("prophet" ou "Holt-Winters").
- `sourceconso (str)` : La colonne du tableau à prédire.

FUNCTION : PREDICT_FOR_DAY()

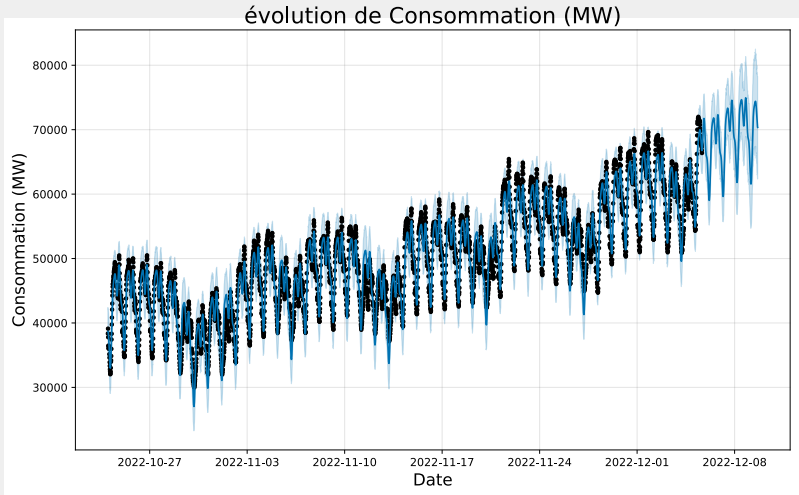
inputs

- `dfdata(pd.DataFrame)` : Tableau d'entraînement
- `dateprediction (str)` : Format yyyy-mm-dd , date à partir de laquelle le modèle s'entraîne sur les données.
- `methode (str)` : Permet de choisir la méthode utilisée.
- `sourceconso (str)` : La colonne du tableau à prédire.
- `savemodel (bool)` : Permet d'enregistrer l'entraînement avec la méthode prophet.
- `loadmodel (bool)` : Permet de charger un entraînement précédent avec la méthode prophet.

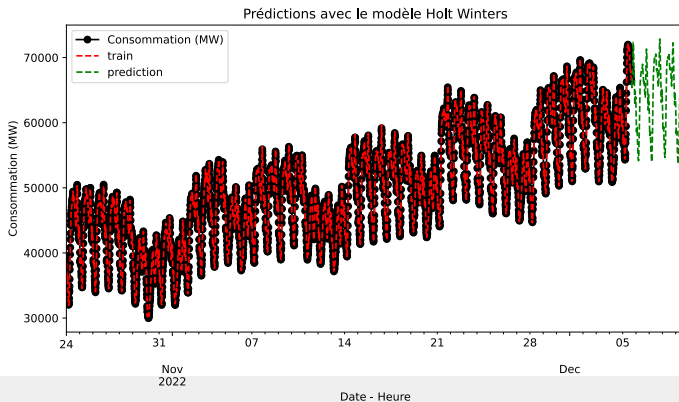
PREDICTION AU 7 DÉCEMBRE



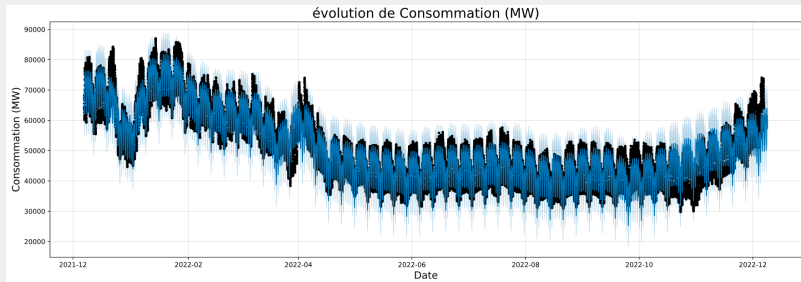
PROPHET COURTE PÉRIODE D'ENTRAÎNEMENT



HOLTWINTERS COURTE PÉRIODE D'ENTRAÎNEMENT



PROPHET AVEC ENTRAÎNEMENT LONG



Tendance

Caractéristique de long terme

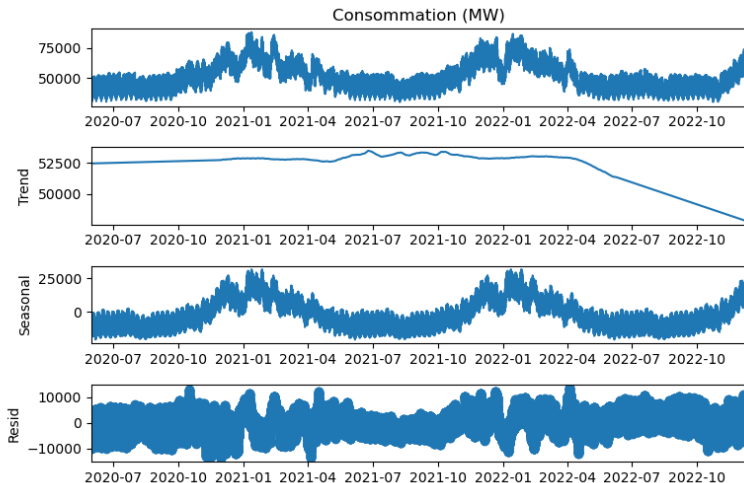
Saisonalité

Caractéristique Cyclique

Bruit

Reste

DÉCOMPOSITION CLASSIQUE



LISSAGE -> OBTENIR LA TENDANCE

Moyenne roulante

Associe à chaque instant t la moyenne sur l'intervalle $[t - p; t + p]$ avec p la période estimée du phénomène cyclique.

Moyenne exponentielle

Prend en compte toutes les observations mais les pondère par une exponentielle décroissante .

Inter-dépendance

- Est ce que le bruit dépend de la tendance ?
- Est ce que l'amplitude dépend de la tendance ?

Modèle

- Source : <https://peerj.com/preprints/3190.pdf>
- Utilise un modèle additif et une décomposition tendance-saisonalité..
- Considère la tendance comme un phénomène de croissance logistique
- Gère la saisonnalité par une décomposition en série de fourier
- Prend en considération particulière les "fêtes" et "vacances" connues.