

Etapes de traitements	Sous-étapes	Nom fichier Python	Fichiers d'entrée	Fichier de sortie	Notes
<b>1- ANALYSE FICHIER BRUT</b>		Energie_TA_Pre_processing.ipynb	eco2mix-regional-cons-def.csv	eco2mix-regional-cons-def.csv	Eco2Mix 2nd version 2013-2023 -Aucune modification du fichier- Exploration des colonnes-Heatmaps-Boxplots des variables
<b>2-PREPARATION Eco2Mix_TCH_keep</b>		Energie_TA_1.py	eco2mix-regional-cons-def.csv	df_eco2_cleaned_tco_tch.csv	Modification des entêtes de colonnes et des noms de régions- Création colonnes 'export' et 'import' - AJOUT des colonnes TCO - Présence des TCH RTE de 2020-2022
<b>3-Capacites MAX production</b>	Calcul des capacités maximales	3-capacites Productions.py	registre_national_installation_production_stockage_electricite.csv	df_result_cap_prod_2013_2022	Calcul des capacités de production maximales par région et filière.Temps de traitement long
<b>4-Indisponibilités</b>	4-1-Preprocessing	4-1-preprocessing_indisponibilites.py	DonneesIndisponibilitesProduction_2013, DonneesIndisponibilitesProduction_2014, DonneesIndisponibilitesProduction_2015, DonneesIndisponibilitesProduction_2016, DonneesIndisponibilitesProduction_2017, DonneesIndisponibilitesProduction_2018, DonneesIndisponibilitesProduction_2019, DonneesIndisponibilitesProduction_2020, DonneesIndisponibilitesProduction_2021, DonneesIndisponibilitesProduction_2022	df_result_ind_prod.csv	Création d'un fichier des indisponibilités 2013-2022 par dates et régions identiques à celles d'Eco2Mix
	4-2-Prefusion_indispo	4-2-Prefusion_indisponibilites.py	df_result_ind_prod.csv	df_result_ind_prod_prefusion.csv	Regroupements par region, filiere_globale et date en cumulant les valeurs de 'capacite_indisponible' et 'capacite_disponible'
<b>5-Fusion Eco2_Cap_Ind</b>		5-Fusion_eco2_capM_Ind.py	df_eco2_cleaned_tco_tch.csv df_result_cap_prod_2013_2022.csv df_result_ind_prod_prefusion.csv	df_eco2_cap_ind_fusion.csv (de 2013 à 2022)	Fusion des dataframes Eco2Mix, Registre Installations et indisponibilités - Ajout des colonnes 'filieres_p_disp' et 'cap_prod-max_filiere'. Les calculs des TCH sur la base de nos calculs sont suspendus.
<b>6-Corrections TCH calculs-RTE</b>		6-Corrections_TCH_calculs_RTE.py	df_eco2_cap_ind_fusion.csv	df_eco2_tch_rte_extrapolé.csv	Création des colonnes 'cmax_[filieres]_rte' (capacités maximales de production sur base données TCH de RTE) - Brèves stats sur 2020-2023 pour extrapolations sur 2013-2019 - Assignation des cmax et TCH sur période 2013-2019 selon données 2020-2023
<b>7-Analyses TCO - TCH -capacités de production</b>	Analyses des TCO et TCH de la METROPOLE	7-1-Analyses_TCO_TCH_METROPOLE.py	df_eco2_tch_rte_extrapolé.csv	df_blackout_ML_pop.csv	Visualisations des TCO et TCH
	Capacites de production - GRAPHIQUES	7-2-capacites Productions_GRAPHIQUES.py	df_eco2_tch_rte_extrapolé.csv		Visualisations des capacités de production
<b>8-Machine Learning Echanges Physiques</b>	8-1-Choix modele	8-1- Machine Learning -choix modele.py	df_blackout_ML.csv	df_blackout_ML_pop.csv	Machine Learning Supervisé sur la variable 'echanges physiques' - Tests de modèles
	8-2-Optimisation XGBoost	8-2- Machine Learning -HyperP_random_XGB.py	df_blackout_ML_pop.csv		Recherche des hyperparamètres optimaux sans transformation variable cible
	8-3-Predictions	8-3- ML -Previsions_XGBoost	df_blackout_ML_pop.csv	resultats_prev_xgb_20_23_complet.csv	Prévision des données 2020 à 2023
		8-4_Previsions_XGBoost_graph Relatives.py	resultats_prev_xgb_20_23_complet.csv		Visualisations pour comparer valeurs réelles et valeurs prévues
<b>9-ML Ech_Phys_Log_Modulus</b>	9-1-Choix du modèle	9-1- Machine Learning -choix modele.py	df_blackout_ML_pop.csv		Machine Learning Supervisé sur la variable 'echanges physiques' transformée avec Log-Modulus - Tests de modèles
	9-2-Recherche des hyperparametres optimaux	9-2- ML -HyperP_random_XGB_LOG.py	df_blackout_ML_pop.csv		Recherche des hyperparametres optimaux
	9-3 -Prévisions	9-3- ML -Previsions_DecisionTree.py		results_dt_glob_20_23.csv	
	9 - 4 - Affichage des graphiques	9_4_Previsions_XGBoost_Log_graph.py	results_dt_glob_20_23.csv		