

Compte rendu projet final

Certification Simplon 2021

Développer data

Apprenant :

Camille Lautrédou

Années :

2020-2021

SOMMAIRE

SOMMAIRE	1
INTRODUCTION	2
I. L'analyse de la demande	3
1. L'état de l'art.....	3
a. Nécessité et objectifs de la transition énergétique	3
b. Les enjeux	3
2. La problématique et les attentes du client	4
3. La traduction métier du besoin client	4
II. La mise en œuvre du projet.....	5
1. La gestion de projet	5
a. De l'initiation du projet à sa réalisation	6
b. Les outils d'organisation et collaboratifs	6
2. Les aspects techniques du projet.....	7
a. La collecte de la donnée	7
b. La création du schéma de base de données relationnelle	11
c. Le nettoyage de la donnée.....	14
d. Stockage de la donnée	15
III. Présentation de l'analyse et des résultats	16
1. Classement des 5 pays producteurs d'électricité par rapport à la population et par type de combustible.....	16
2. Classement des 5 pays producteurs d'électricité par type de combustible.....	17
3. Cartographie de la production d'électricité en 2018	18
4. Evolution de la production d'électricité de 2009 à 2018	19
5. La part des énergies dans la production d'électricité des pays européens (2018)	21
6. La part des importations par type de combustible (2018).....	22
CONCLUSION	23

INTRODUCTION

Nous allons aborder dans ce compte rendu la thématique de la transition énergétique et environnementale appliquée à un travail technique lié à la data dans le but de concevoir et développer une base de données afin de mettre en évidence des analyses et résultats liés à ce sujet. Ce projet a été réalisé collectivement dans la conception et le développement de la base de données et individuellement sur la partie visualisation des données.

Depuis la révolution industrielle, le mode de vie des sociétés occidentales est lié à la présence d'énergie abordable et facilement accessible. Ce modèle de civilisation « moderne » et énergivore s'est depuis lors largement répandu à l'échelle de la planète. Ceci a été rendu possible par un apport énergétique principalement assuré par les combustibles fossiles. Aujourd'hui, une transition énergétique de grande envergure s'impose. Celle-ci doit s'inscrire dans la transition vers une économie décarbonisée qu'une stabilisation du climat rendra nécessaire.

Tout d'abord, rappelons la définition et les enjeux de la transition énergétique. La transition énergétique est le terme utilisé pour désigner la transformation du système énergétique français. Plus concrètement, l'expression sert à désigner l'ensemble des changements engagés pour réduire l'impact environnemental de la production, de la distribution et de la consommation d'énergie (électricité, gaz...). Ces changements répondent à des engagements européens et sont inscrits dans une loi : la loi de transition énergétique pour la croissance verte, promulguée en 2015. Les enjeux sont multiples notamment d'un point de vue environnemental, sanitaire, économique et politique. La production d'électricité est un point important de la transition énergétique car certaines énergies utilisées ont un impact négatif sur notre environnement. Certaines énergies fossiles ou énergies nucléaires sont encore massivement utilisées et reste difficile à remplacer.

Il est donc légitime de se demander quel est l'état des lieux de la transition énergétique et environnementale en Europe ? Et plus particulièrement de répondre à la question suivante : Quelle est l'évolution de la production d'électricité selon le type de combustible en Europe ? Pour répondre à cette question, il est tout d'abord nécessaire de comprendre le besoin qui émane de ce projet, comment mettre en œuvre ce projet et enfin la présentation des analyses liées à cette problématique.

I. L'analyse de la demande

1. L'état de l'art

Dans l'introduction de ce compte rendu nous avons pu évoquer succinctement la définition et les enjeux de la transition énergétique et environnementale. Nous allons donc dans cette partie aborder plus précisément cette thématique pour en définir les contours et les enjeux et plus particulièrement sur la production d'électricité en Europe.

a. Nécessité et objectifs de la transition énergétique

Le réchauffement climatique, affectant déjà les écosystèmes, la météo ou les cultures, est majoritairement dû à l'émission de gaz à effets de serre générée par les activités humaines. Or, l'énergie en est en grande partie responsable, qu'il s'agisse de produire ou de consommer, de gaz ou d'électricité, d'essence ou de déchets. La transition énergétique doit permettre de lutter contre le changement ou de limiter le réchauffement climatique. La transition énergétique répond à la fois à un besoin et à une volonté, des gouvernements comme des populations. Si l'enjeu principal est la réduction de l'impact sur l'environnement, les conséquences de ce changement sont bien plus vastes : les enjeux de la transition énergétique touchent à la fois à l'économie, au confort, à la géopolitique et, bien sûr, à l'écologie.

La transition énergétique intègre plusieurs objectifs clairs et chiffrés inscrits dans la loi (loi de transition énergétique pour la croissance verte). Tous répondent à un même enjeu majeur : diminuer l'impact environnemental du système énergétique français. Une diminution qui bénéficiera à la planète comme à la nation et engagera à la fois les citoyens, les entreprises et les territoires (cf. www.geo.fr, « Transition énergétique : définition et enjeux »)

b. Les enjeux

▪ Les enjeux sanitaires et environnementaux

L'enjeu environnemental : en allégeant l'impact du système énergétique sur l'environnement, la transition énergétique doit préserver l'écosystème et limiter le réchauffement climatique dû aux émissions de gaz à effet de serre. Les énergies vertes quant à elles polluent peu ou pas, utilisent moins de ressources voire aucune ressource épuisable.

L'enjeu sanitaire : la pollution de l'air est responsable de nombreuses infections respiratoires. Le réchauffement climatique, en transformant le climat et les écosystèmes, engendre également des risques dus à la prolifération de certaines bactéries et autres insectes.

▪ Les enjeux économiques et politiques

Réduire la consommation globale d'énergie ou la production de déchets implique une réduction des coûts liée à toute la chaîne de valeur. Autrement dit, une réduction des dépenses

pour les collectivités ou les entreprises, favorisant la compétitivité et la rentabilité. La modification du système a également pour but une meilleure indépendance énergétique. Celle-ci limiterait à la fois les variations de prix dues à la situation politique dans les pays producteurs, et les tensions entre les pays vendant ou achetant de l'énergie. L'enjeu de la transition énergétique est également sécuritaire. En France, le système énergétique repose largement sur l'énergie nucléaire. Coûteuse en eau et productrice de déchets nucléaires, elle peut également donner lieu à des accidents potentiellement catastrophiques. La remplacer progressivement par des énergies renouvelables réduirait ce risque (cf. www.geo.fr, « Transition énergétique : définition et enjeux »).

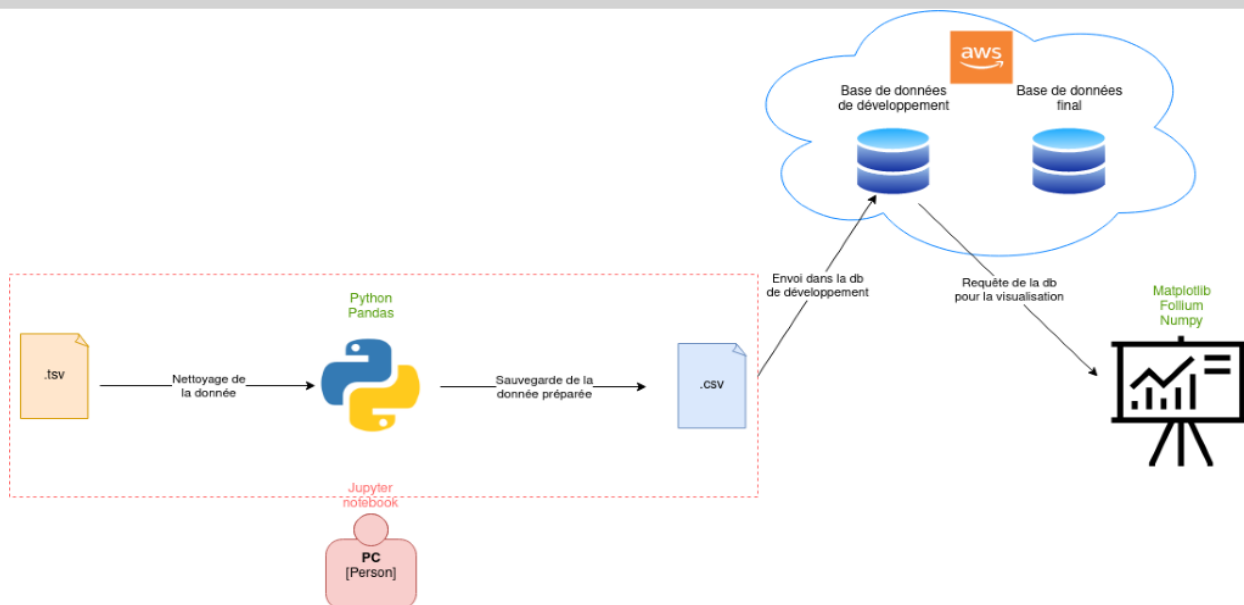
2. La problématique et les attentes du client

Nous avons débuté ce projet avec une thématique imposée : la transition énergétique et environnementale. Les attentes sont les suivantes : collecter de la donnée sur cette thématique, concevoir et développer une base de données et enfin visualiser cette donnée dans le but de mettre en évidence des résultats émanant d'une problématique préalablement définie. La base de données doit être composée d'un minimum de 30 000 lignes. Dans cette partie nous allons expliquer comment nous avons appréhendé la problématique et les attentes du client, quelles difficultés nous avons pu rencontrer.

La première étape de ce projet est la collecte de la donnée. Cette étape est primordiale et est un moment clé dans le démarrage d'un projet puisque c'est avec ces données que nous allons travailler pendant plusieurs semaines. Il faut prendre en compte le thème, le nombre de lignes imposé, récolter des données relativement récentes, gratuites et à partir de sources fiables (principalement via des sites officiels). Il faut également se poser les bonnes questions : qu'allons-nous mettre en avant à travers ces données ? Sont-elles suffisamment intéressantes pour en faire de la data visualisation et mettre en évidence des points spécifiques ? Il est souvent difficile de répondre à toutes ces questions en survolant le jeu de données car c'est souvent en allant manipuler la donnée, lire la documentation que l'on va vite se rendre compte des difficultés que l'on pourra rencontrer. Dans notre groupe nous avons chacun cherché des données sur notre thématique et avons choisi le jeu de données qui répondait le mieux aux attentes du client notamment avoir assez de matière pour travailler à trois.

3. La traduction métier du besoin client

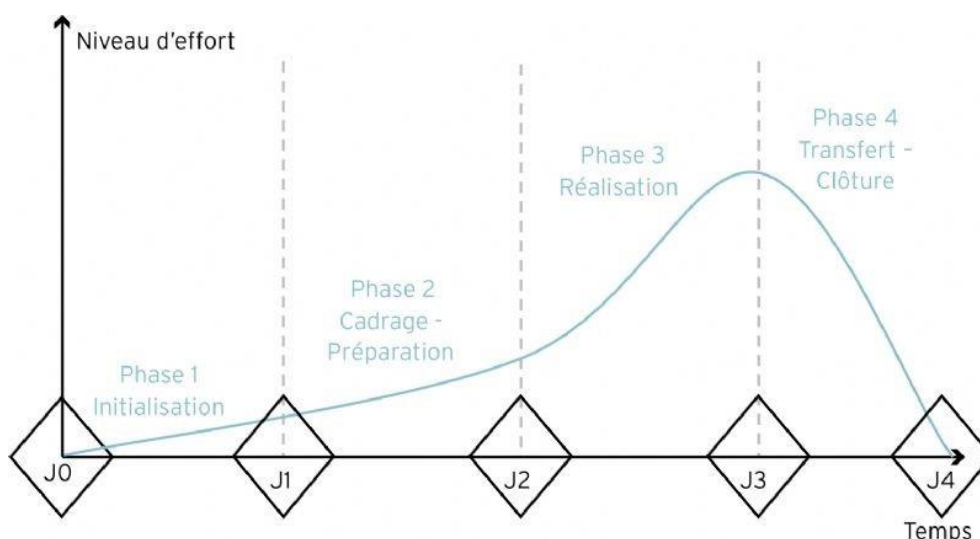
Afin de visualiser le déroulement d'un projet, j'ai mis en place un schéma fonctionnel sur l'architecture de ce projet. On peut y suivre les différentes étapes avec notamment la collecte de la donnée, la préparation, le stockage et la visualisation. Ces étapes d'un point de vue technique seront détaillées davantage dans la seconde partie appelée la mise en œuvre du projet. Ce schéma a pu être alimenté au fur et à mesure de son déroulement.



II. La mise en œuvre du projet

1. La gestion de projet

Nous avons débuté notre projet par la mise en commun des données collectées individuellement dans le but de choisir un jeu de données qui correspondait aux attentes finales. Nous avons sélectionné les données les plus pertinentes dans le cadre d'un projet collectif mais également individuel. Ce sont donc deux paramètres très importants à prendre en compte dans l'organisation et la gestion du projet afin de créer un suivi régulier de notre travail. Dans cette partie nous allons nous appuyer sur le cycle de vie d'un projet pour structurer sa mise en œuvre. (Voir ci-dessous).



Ce schéma permet de décrire l'organisation générale d'un projet en définissant la logique de son déroulement. L'objectif étant de fixer le premier niveau de planification du projet :

- En le découpant en grandes étapes, que l'on appelle phases pour les projets séquentiels.
- En définissant dès le départ la durée du projet et les étapes à réaliser.

a. De l'initiation du projet à sa réalisation

Phase 1 : l'initialisation

Dans cette première phase nous avons la prise de connaissance des consignes.

Phase 2 : cadrage et préparation

Préparation principalement orientée gestion de projet. Avec la mise en commun des outils, choisir la donnée (avec une source fiable et des données de qualité), le langage SQL utilisé, le choix de l'hébergeur de la base de données, orienter notre projet afin de réaliser un projet en groupe mais dont une partie sera réalisée individuellement.

Phase 3 : réalisation

Chacun travaille et évolue sur les tâches qui lui sont attribuées, des points réguliers sont effectués pour suivre l'évolution et les difficultés de chacun. Ces points permettent également de mettre à jour la base de données si besoin.

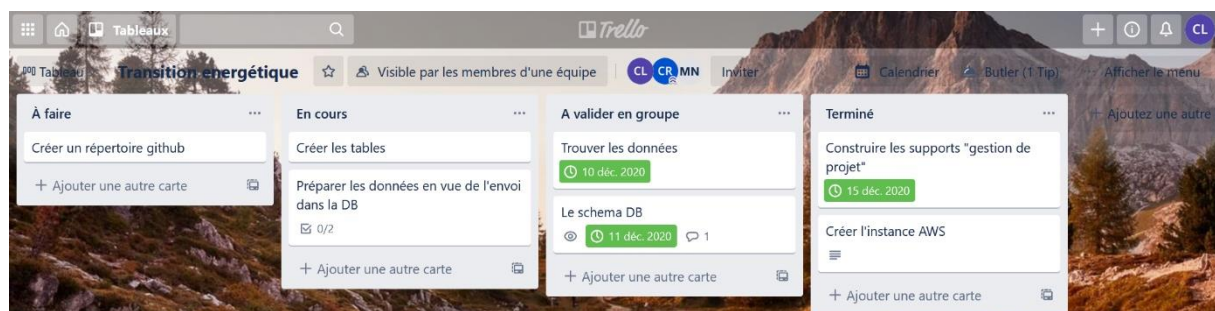
Phase 4 : clôture

Date limite du projet, soit le 21 janvier 2021.

b. Les outils d'organisation et collaboratifs

Pour l'organisation et le suivi de projet nous avons choisi de collaborer dans un premier temps via un google doc. Le but étant de mettre à plat nos idées et de les organiser au fur et à mesure. Après avoir délimité les étapes du projet, nous avons choisi comme outil « Trello ». Trello est un outil de gestion de projet en ligne, inspiré par la méthode Kanban de Toyota. Nous avons choisi d'utiliser la méthode de visualisation de Kanban qui consiste à visualiser le flux de travail en utilisant des colonnes. Ces colonnes représentent l'état d'avancement du projet avec entre autres :

- « A faire » : définition des étapes et sous étapes du projet
- « En cours » : pour que chaque personne du groupe puisse savoir qui avance sur quoi
- « A valider en groupe » : création d'une étape intermédiaire afin de savoir quelle étape doit être validée collectivement dans le but de passer définitivement à l'étape suivante
- « Terminé » : visualisation des étapes franchies au cours du projet



2. Les aspects techniques du projet

Voici un tableau reprenant l'ensemble des langages informatiques utilisé, les différentes bibliothèques et les outils pour la mise en œuvre du projet.

Langages	Bibliothèques	Outils
Python	Pandas	Jupyter notebook
Postgrès	Matplotlib	Dbeaver
	Folium	Vscode
	SqlAlchemy	AWS

a. La collecte de la donnée

Notre thématique imposée pour ce projet final est donc la transition énergétique et environnementale. Nous avons choisi d'aborder cette thématique à partir du site Eurostat (<https://ec.europa.eu/eurostat/fr/data/database>). Eurostat est l'Office statistique de l'Union européenne. Il a pour mission de fournir des statistiques de qualité sur l'Europe. Les statistiques sont harmonisées à l'aide du système statistique européen (SSE) afin d'être comparables entre les pays. On y trouve plusieurs thématiques comme l'économie, la population, l'agriculture, l'industrie, les transports, l'environnement, ou encore la science et les technologies.

Nous nous sommes donc principalement concentrés sur la partie environnementale dans la sous-catégorie « énergie ». Après analyse des différents jeux de données proposés nous avons orienté notre choix sur un premier fichier nommé « Bilan énergétique complets ». Afin de compléter notre projet de groupe nous avons choisi d'aborder 3 sous parties de la transition énergétique que chacun étudiera individuellement :

- Les déchets générés et le recyclage (Chloé Rontex)
- La production d'électricité par type de combustible (Camille Lautrédou)
- Et enfin la surface agricole biologique (Maxim Novikov)

Pour comprendre l'objectif et récapituler l'ensemble des jeux de données utilisé dans ce projet, voici un descriptif comprenant l'intitulé des fichiers, la source, le type de format, une présentation succincte, la série chronologique et la fréquence des données, l'unité de mesure sélectionnée et enfin le nombre de lignes.

Pour les deux fichiers présentés ci-dessous, "bilan énergétiques complets" et "production d'électricité et de chaleur par type de combustible", nous pouvions sélectionner les informations suivantes :

- **Le code GEO** : code et libellé des noms des pays de l'UE (28 pays de 2013 à 2020, 27 pays à partir de 2020), de l'AELE (Association européenne de libre-échange) et des candidats à l'adhésion à l'UE.

- **Le code NRG_BAL** : code et libellé du type de bilan énergétique (énergie disponible brute, consommation d'énergie primaire, production brute de chaleur, ...) .
- **Le code SIEC** : code et libellé des types de combustible (combustibles fossiles, charbon, éolien, biocarburant, ...).
- **Le code UNIT** : code et libellé des unités de mesure (térajoules, en kilotonnes d'équivalent pétrole et en gigawattheures).

Voici la description des fichiers sélectionnés pour ce projet :

1. Bilan énergétiques complets

https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_bal_c&lang=fr

Le bilan énergétique permet aux utilisateurs de voir la quantité totale d'énergie extraite de l'environnement, échangée, transformée et utilisée par différents types d'utilisateurs finaux. Il permet également de voir la contribution relative de chaque vecteur d'énergie (carburant, produit). Le bilan énergétique permet d'étudier l'ensemble du marché intérieur de l'énergie et de surveiller les impacts des politiques énergétiques. Le bilan énergétique offre une vision complète de la situation énergétique d'un pays dans un format compact, telle que la consommation d'énergie de l'ensemble de l'économie et des différents secteurs. Le bilan énergétique présente tous les produits énergétiques statistiquement significatifs (combustibles) d'un pays et leur production, transformation et consommation par différents types d'acteurs économiques (industrie, transport, etc.). Par conséquent, un bilan énergétique est le point de départ naturel pour étudier le secteur de l'énergie.

La collecte annuelle de données couvre les États membres de l'UE, les pays membres de l'AELE, les pays candidats à l'adhésion à l'UE et les pays candidats potentiels et récolte des données de 2009 à 2019.

Type de code	Définition	Valeurs sélectionnées	Libellé
CODE FREQ	Relative au temps	1/1	Annuelle
CODE GEO	Entité géopolitique	43/43	États membres de l'UE, AELE, pays candidats à l'adhésion à l'UE et les pays candidats potentiels
CODE NRG_BAL	Bilan énergétique	6/139	Production primaire, production brute d'électricité, production brute de chaleur, consommation finale d'énergie (2020-2030), importations et exportations
CODE SIEC	Classification internationale standard des produits de l'énergie	72/72	Intégralité des combustibles
CODE UNIT	Unité de mesure	1/3	Tonnes équivalent pétrole
CODE TIME	Temps	10/10	2009 à 2018

Nous arrivons donc à un total de lignes de :

43 pays x 6 bilans énergétiques x 72 types de combustibles x 10 années soit 185 760 lignes.

2. Production d'électricité et de chaleur par type de combustible

https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_bal_peh&lang=fr

Ce jeu de données présente la production d'électricité et de chaleur par type de combustible. J'ai choisi de sélectionner deux types de combustible ainsi que la totalité d'entre eux afin de rendre mon analyse plus pertinente et plus ciblée. Dans ce fichier cela correspond au code SIEC suivant :

- Les énergies fossiles (Combustibles fossiles solides)
- Les énergies renouvelables (Renouvelables et biocarburants)
- La totalité des combustibles (Total)

Dans ces données on pourra mesurer l'étendue de l'utilisation de l'énergie renouvelable et, par implication, le degré auquel les carburants renouvelables ont remplacé les carburants fossiles et par conséquent, contribué à la décarbonisation de l'économie de l'UE.

Ce jeu de données comprend les États membres de l'UE, les pays membres de l'AELE, pays candidats à l'adhésion à l'UE et les pays candidats potentiels et récolte des données de 2009 à 2018.

Type de code	Définition	Valeurs sélectionnées	Libellé
CODE FREQ	Relative au temps	1/1	Annuelle
CODE GEO	Entité géopolitique	43/43	États membres de l'UE, AELE, pays candidats à l'adhésion à l'UE et les pays candidats potentiels
CODE NRG_BAL	Bilan énergétique	1/139	Production brute d'électricité
CODE SIEC	Classification internationale standard des produits de l'énergie	3/72	Combustibles fossiles solides, Renouvelables et biocarburants et le total
CODE UNIT	Unité de mesure	1/3	Milliers de tonnes équivalent pétrole
CODE TIME	Temps	10/10	2009 à 2018

Chaque catégorie de combustible (fossile, renouvelable et le total) sera regroupée dans des fichiers distincts soit 3 fichiers au total. Ainsi, nous arrivons donc à un total de lignes de :
43 pays x 1 bilan énergétique x 1 types de combustibles x 10 années soit 430 lignes par fichier.

A présent, nous allons pouvoir décrire les fichiers secondaires (fichiers qui viennent compléter les fichiers principaux).

Ces fichiers s'ajouteront à notre base de données dans le but de venir faire des clés étrangères ou tables de jointure. Ces données ont été extraites du site Eurostats, il n'existe pas de fichiers de base créé dans le but de télécharger cette donnée. Nous avons donc du copier-coller pour les mettre en fichier texte. Il n'a pas été possible de "scraper" cette donnée en python. L'ensemble de ces données sont réparties en 2 colonnes, le code et son libellé.

<https://appsso.eurostat.ec.europa.eu/nui/submitModifiedQuery.do>

1. Code GEO

Ces données présentent :

- Les États membres de l'UE : l'Union Européenne à 27 pays (à partir de 2020) et l'Union Européenne à 28 pays (2013-2020),
- L'AELE (l'Association européenne de libre-échange) : ici Norvège et Islande
- Les pays candidats à l'adhésion à l'UE et les pays candidats potentiels : ici Turquie, Serbie

On trouve donc un total de 43 lignes.

2. Code NRG_BAL

Ces données présentent le type de bilan énergétique que l'on peut choisir pour spécifier une analyse, avec par exemple : La production primaire, les énergies disponibles brutes, disponible pour la consommation finale. Il existe 139 types de catégorie.

3. Code SIEC

Ensemble des combustibles d'après la classification internationale standard des produits de l'énergie. On trouve au total 72 catégories de combustible.

4. Evolution de la démographie

https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_gind&lang=fr

Ce jeu de données présente les États membres de l'UE, les pays membres de l'AELE, les pays candidats à l'adhésion à l'UE et les pays candidats potentiels et l'évolution de leur population de 1960 à 2020. Les données sont établies au 1er janvier de chaque année.

Type de code	Définition	Valeurs sélectionnées	Libellé
CODE FREQ	Relative au temps	1/1	Annuelle
CODE GEO	Entité géopolitique	43/43	États membres de l'UE, AELE, pays candidats à l'adhésion à l'UE et les pays candidats potentiels
CODE INDIC_DE	Indicateur démographique	1/28	Population au 1er janvier - totale
CODE TIME	Temps	61/61	1960 à 2020

Les autres fichiers sont ceux des deux autres membres du groupe comprenant :

Déchets générés par activité économique

<https://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=fr&pcode=ten00106&plugin=1>

Taux de recyclage des déchets municipaux

https://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=fr&pcode=t2020_rt120&plugin=1

Taux de recyclage des e-déchets

https://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=fr&pcode=t2020_rt130&plugin=1

Superficies cultivées de manière biologique par méthodes de production agricole et cultures végétales (à partir de 2012)

https://ec.europa.eu/eurostat/databrowser/view/org_cropar/default/table?lang=fr

b. La création du schéma de base de données relationnelle

Nous avons dans un premier temps utilisé le site de création de base de données dbdiagrame.io puis nous nous sommes rabattus vers dbdesigner. Un outil plus complet et plus simple d'utilisation.

Modélisation conceptuelle

Les éléments de cette base de données sont construits de la manière suivante : les entités, les attributs, les types de relation et l'identifiant. Comme vu dans la partie précédente, nous avons déterminé le contenu de notre base de données avec la sélection de nos fichiers. Ces fichiers seront donc nos entités :

- Bilan énergétique

- Production d'électricité fossile
- Production d'électricité renouvelable
- Production d'électricité totale
- Déchets
- Recyclage
- Surface biologique
- Évolution démographique

Chacun de ces fichiers comprend des codes sur les pays, sur les types de combustibles, sur le type de bilan énergétique, etc... Nous avons donc ajouté ces entités comme table car elles nous seront utiles par la suite pour créer des clés étrangères :

- Code pays
- Code siec
- Code bilan énergétique
- Code activité économique
- Code statut biologique
- Code type de production

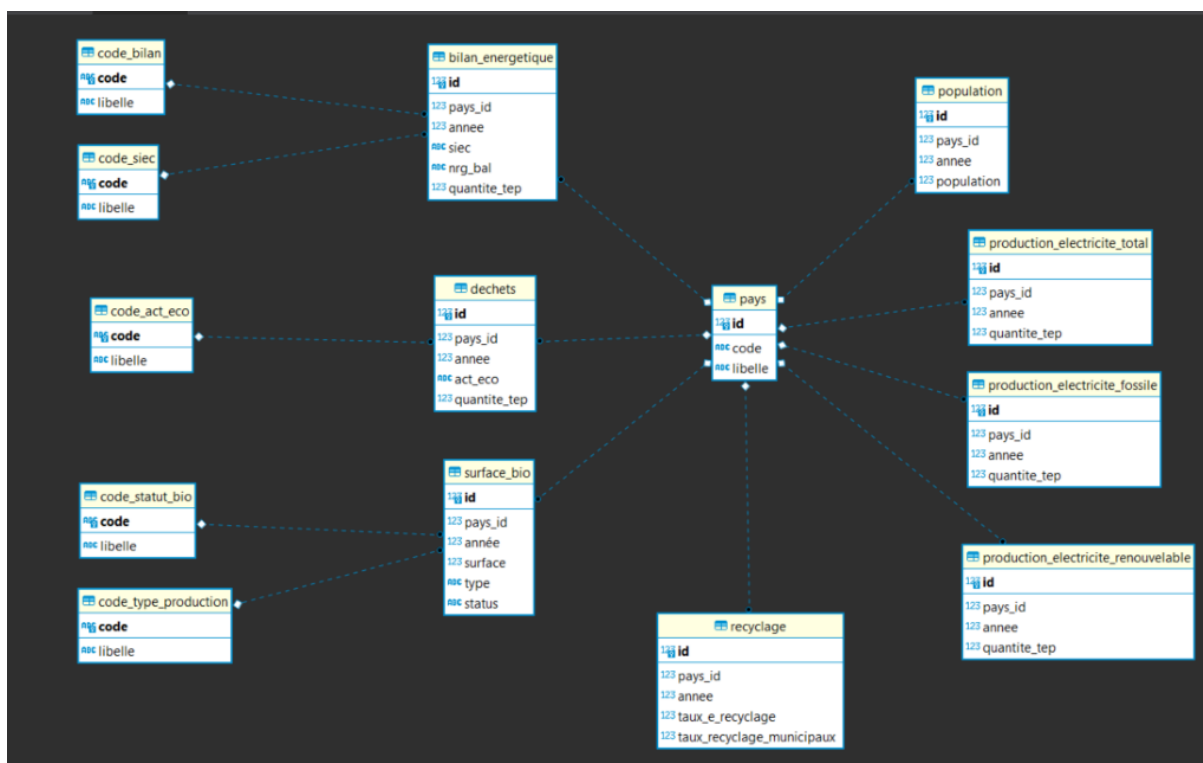
Nous avons ensuite déterminé les attributs (les colonnes) de nos tables qui correspondent aux données que nous voulons mettre en évidence. Chaque table sera identifiée par un "id".

Entités	Bilan énergétique	Production d'électricité totale	Production d'électricité fossile	Production d'électricité renouvelable	Déchets	Recyclage	Surface biologique	Population
Attributs	Code pays	Année	Année	Année	Année	Année	Année	Année
	Code siec	Quantité	Quantité	Quantité	Quantité	Taux de e-recyclage	Surface	Population
	Code nrg_bal					Taux de recyclage municipaux	Type de production	
	Année						Statut biologique	
	Quantité							

Entités	Code pays	Code siec	Code bilan	Code activité économique	Code statut biologique	Code type de production
Attributs	Code	Code	Code	Code	Code	Code
	Libellé	Libellé	Libellé	Libellé	Libellé	Libellé

A présent nous avons donc identifié nos tables et leurs colonnes. Il faut maintenant déterminer les relations entre ces tables. Les relations de nos tables sont toutes en « one to many » ou « une à plusieurs » en français. Par exemple, un bilan énergétique pour une année donnée ne concerne qu'un seul pays mais un pays à plusieurs bilans énergétiques (en fonction de l'année et du type de bilan). Ainsi, l'id de la table « [pays](#) » devient une clé étrangère dans la table « [bilan_énergétique](#) ». La création d'une table « [pays](#) » permet également d'éviter la redondance de la donnée. Nous aurions pu mettre dans chaque table (bilan énergétique, production d'électricité fossile, ...) le nom des pays mais ceci n'est pas une bonne pratique. Pour la table « [population](#) » nous avons fait la même chose. Nous aurions pu avoir une table « [pays](#) » regroupant le code, le libellé, l'année et l'évolution de la population du pays. Afin d'avoir un historique de cette évolution nous avons choisi de dédier une table nommée « [population](#) » et avoir les codes des pays uniquement sur une table.

Voici le schéma de notre base de données finale :



c. Le nettoyage de la donnée

Le nettoyage de la donnée a été fait à partir d'un jupyter notebook avec l'utilisation de la bibliothèque Pandas. Dans cette partie les étapes de la préparation de la donnée seront principalement énumérées par ordre chronologique de l'avancement du projet.

Transformation des années (en colonnes) en ligne

La première partie la plus importante dans ce travail de nettoyage et pour l'ensemble des fichiers, soit 5 au total, est de rendre notre fichier de base au format que l'on souhaite intégrer dans nos tables. Par exemple, pour la table « production d'électricité fossile », les années sont énumérées en ligne (voir table à droite), alors que dans notre fichier de base ce n'est pas le cas. Une boucle « for » en python a permis de résoudre ce problème.

production_electricite_fossile			
id	integer		
pays_id	integer		
annee	integer		
quantite_tep	decimal		
Add field			

APRES

de résoudre ce problème.

AVANT

GEO,NRG_BAL,SIEC,UNIT:TIME	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
0 Union européenne - 27 pays (à partir de 2020), Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	60.496,545	60.294,939	62.359,396	63.861,407	62.675,652	59.566,150	60.618,471	56.678,555	54.930,576	51.213,293
1 Union européenne - 28 pays (2013-2020), Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	69.356,132	69.546,358	71.683,730	76.139,395	73.875,738	68.185,153	67.142,824	59.315,580	56.867,846	52.660,533
2 Zone euro - 19 pays (à partir de 2015), Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	39.139,194	38.510,932	39.791,898	42.887,030	42.140,432	39.865,979	41.022,267	37.340,874	35.745,174	32.374,392
3 Belgique, Production brute d'électricité, Combustibles fossiles	445,228	360,877	298,882	291,402	258,641	179,725	178,392	33,491	7,816	7,850

id	pays_id	annee	quantite_tep	
0	1	1	2009	51552.228
1	2	2	2009	54039.770
2	3	3	2009	39937.608
3	4	4	2009	590.598
4	5	5	2009	369.767
5	6	6	2009	447.811
6	7	7	2009	865.262
7	8	8	2009	8735.082
8	9	9	2009	46.459
9	10	10	2009	383.855
10	11	11	2009	727.088

Utilisation de la méthode «split » sur une colonne

Lors de l'importation de nos fichiers une colonne commune qui décrit les codes GEO, SIEC, NRG_BAL étaient dans une seule colonne, il a donc fallu « splitter » cette colonne afin de récupérer les données qui nous intéressaient dans le but de faire nos clés étrangères. La méthode « split » a permis de séparer les données sur les virgules. Voici une illustration pour mieux comprendre.

	GEO,NRG_BAL,SIEC,UNIT:TIME	GEO	NRG_BAL	SIEC	UNIT:TIME
0	Union européenne - 27 pays (à partir de 2020), Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	Union européenne - 27 pays (à partir de 2020)	Production brute d'électricité	Combustibles fossiles solides	Milliers de tonnes équivalent pétrole (TEP)
1	Union européenne - 28 pays (2013-2020), Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	Union européenne - 28 pays (2013-2020)	Production brute d'électricité	Combustibles fossiles solides	Milliers de tonnes équivalent pétrole (TEP)
2	Zone euro - 19 pays (à partir de 2015), Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	Zone euro - 19 pays (à partir de 2015)	Production brute d'électricité	Combustibles fossiles solides	Milliers de tonnes équivalent pétrole (TEP)
3	Belgique, Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	Belgique	Production brute d'électricité	Combustibles fossiles solides	Milliers de tonnes équivalent pétrole (TEP)
4	Bulgarie, Production brute d'électricité, Combustibles fossiles solides, Milliers de tonnes équivalent pétrole (TEP)	Bulgarie	Production brute d'électricité	Combustibles fossiles solides	Milliers de tonnes équivalent pétrole (TEP)

Utilisation de “regex” pour modifier de la donnée

Après avoir utilisé la méthode « split » en séparant à partir des virgules, une autre problématique est apparue. Dans l’intitulé de l’Allemagne il y a une virgule (« Allemagne (jusqu'en 1990, ancien territoire de la RFA)») et également dans un intitulé de combustible « Énergie hydrocinétique, houlomotrice, marémotrice ». Ainsi, en faisant le split, ces données ont été séparées, cela a causé des problèmes lors des jointures, comme ces données ne correspondaient plus à l’intitulé exact, elles n’ont pas été prises en compte. C’est grâce à la vérification du nombre de lignes que je me suis aperçue de ce manque de données. Pour résoudre ce problème j’ai donc utilisé des « Regex » pour modifier l’intitulé et ne pas mettre de virgules.

Remplacement des caractères spéciaux

Plusieurs caractères spéciaux ont dû être remplacés afin de terminer le nettoyage de la donnée. Après le remplacement des caractères spéciaux il a fallu convertir la donnée en « integer » ou « float ». Enfin pour terminer, une vérification de la cohérence de la donnée et du nombre de lignes a été minutieusement appliqués.

d. Stockage de la donnée

Nous avons décidé d’héberger notre base de données sur AWS avec son service Relational Database Service (RDS). Pour ce projet nous avons mis en place un environnement de développement dans le but :

- D’alimenter la db (CREATE)
- Requête pour tester et construire notre analyse (READ)
- Mettre à jour les données (UPDATE/DELETE)

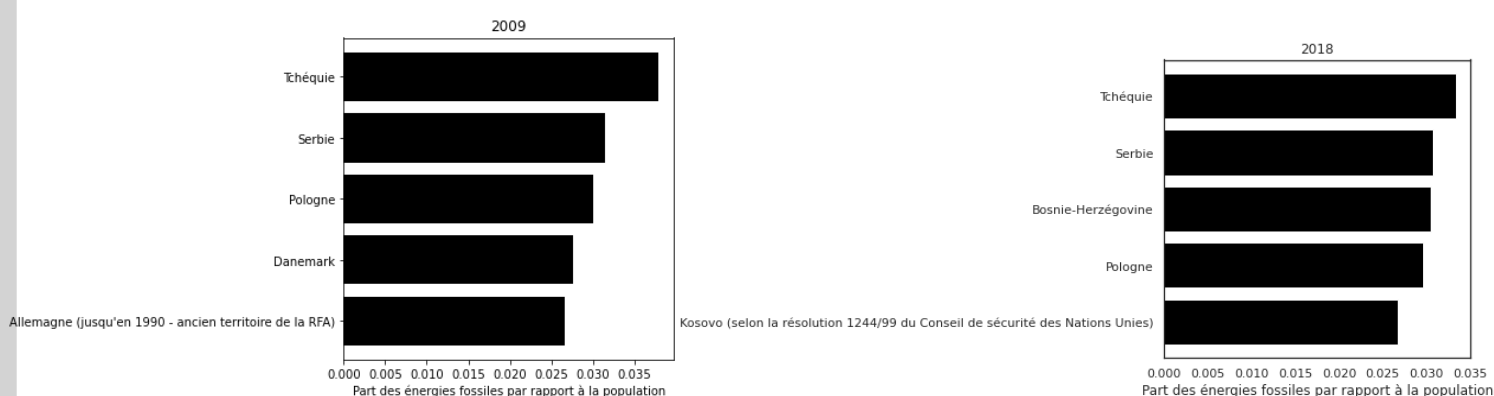
Et un environnement pour la mise en production a également été mis en place. Nous avons donc alimenté et testé notre base de données de développement. Au fur et à mesure de la réalisation de ce projet nous avons pu apporter quelques modifications. C’est pour cela qu’il est important de mettre en place un environnement de développement car à la fin de ce projet nous aurons une version aboutie de notre base de données et pourrons transférer ces données sur la base de données finale pour conserver un rendu définitif.

III. Présentation de l'analyse et des résultats

L'analyse de la donnée a été construite à partir d'un jupyter notebook. Pour rappel, la problématique énoncée dans l'introduction de ce rapport, est d'analyser l'évolution de la production d'électricité selon les types de combustible en Europe. Ainsi, pour constater cette évolution sur une période de 10 ans (entre 2009 et 2018) des graphiques et statistiques provenant de notre base de données permettront de faire un bilan et de mettre en lumière des points particuliers. Dans cette partie la production d'électricité sera analysée sous 3 angles : les énergies fossiles, les énergies renouvelables et enfin la production totale. Deux axes seront principalement abordés : la production d'électricité rapportée à la population et la production d'électricité du pays. Grâce à ces données nous pourrions mesurer l'étendue de l'utilisation de l'énergie renouvelable et par implication, le degré auquel les carburants renouvelables ont remplacé les carburants fossiles. Chaque sous partie représentera un point de l'analyse.

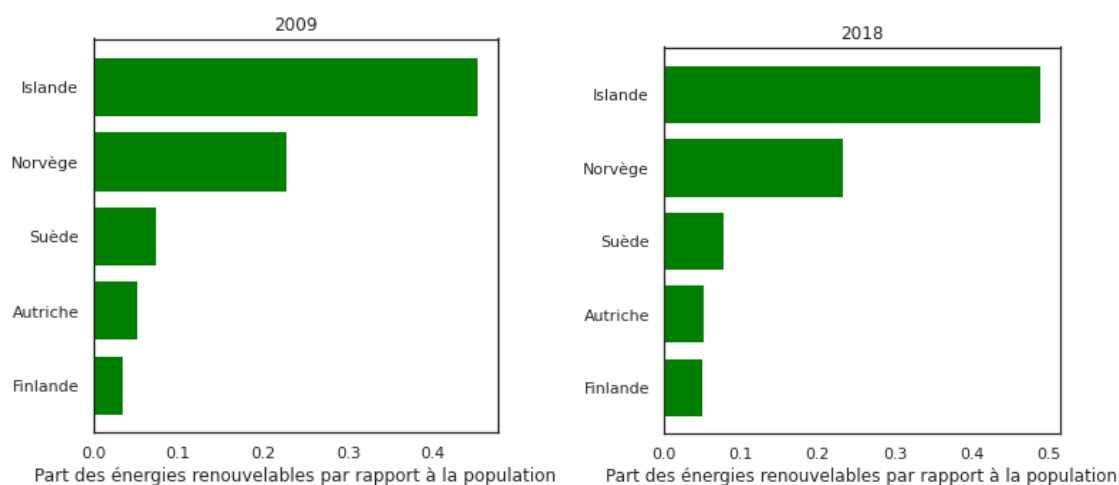
1. Classement des 5 pays producteurs d'électricité par rapport à la population et par type de combustible

Afin de comparer au mieux les pays, j'ai choisi de prendre comme année de référence 2009 et 2018 pour visualiser l'évolution du classement avec les énergies fossiles et renouvelables. Ces diagrammes représentent la part des énergies fossiles par rapport à la population. Le calcul est le suivant pour chaque pays : (quantité en milliers de tonnes équivalent pétrole/la population) * 100. Dans la requête j'ai trié ce calcul par ordre décroissant et sélectionné les 5 premiers pays. On peut donc constater que la République Tchèque et la Serbie par rapport au nombre d'habitant produisent beaucoup d'électricité fossile.



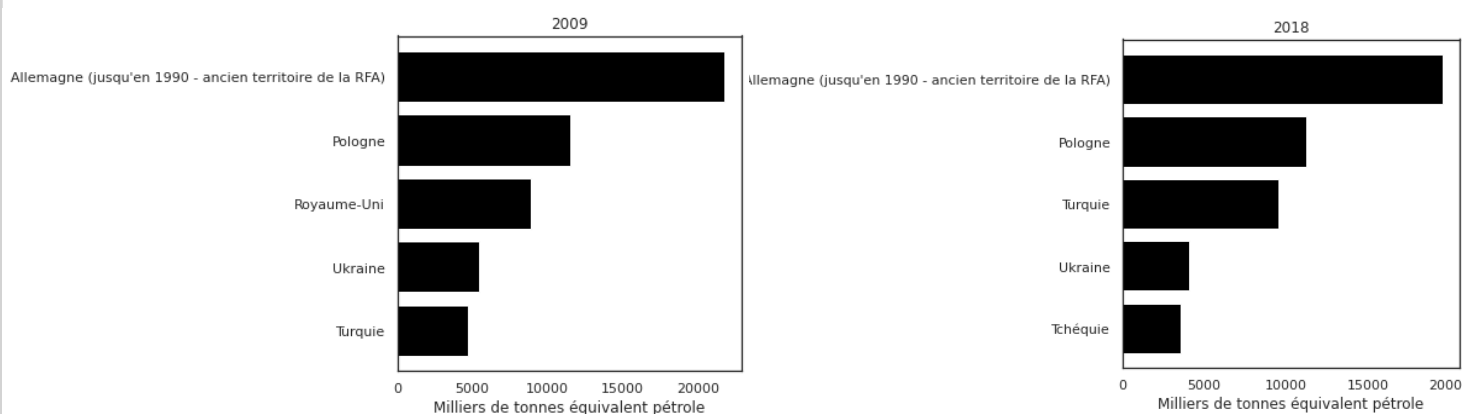
Ces diagrammes quant à eux représentent la part des énergies renouvelables par rapport à la population. Les pays nordiques reviennent constamment sur le podium. Des raisons principalement politiques et géographiques permettent d'expliquer ce classement. Les pays scandinaves ont su prendre un virage plus soutenu dans leur transition énergétique avec une politique énergétique développée. L'hydroélectricité est particulièrement utilisée dans les pays scandinaves comme la Norvège, la Suède, la Finlande ou encore l'Islande grâce à une

géographie qui s'adapte particulièrement à l'installation de Centrales Hydroélectriques (cf. hydroproyectos, les pays scandinaves et les énergies renouvelables).

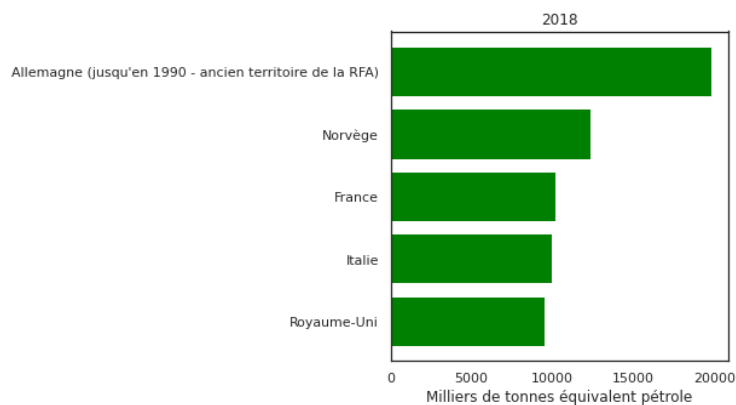
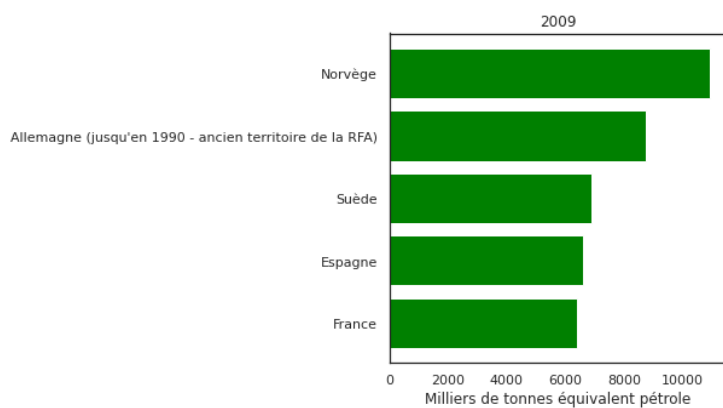


2. Classement des 5 producteurs d'électricité par type de combustible

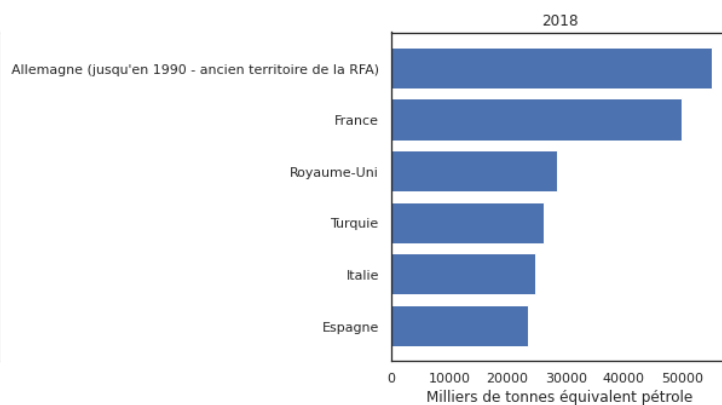
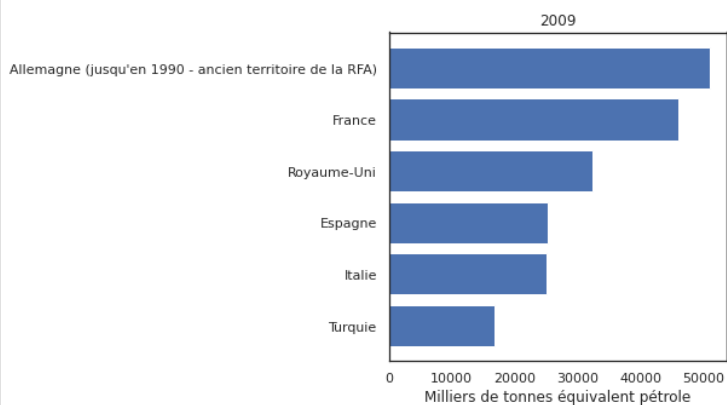
Ces diagrammes cette fois-ci ne représentent pas la production rapportée à la population mais la production du pays dans sa globalité. On constate des changements dans le classement, l'Allemagne est le 1^{er} producteur d'énergie fossile d'Europe et conserve sa position de leader depuis 10 ans. Même si nous constaterons, dans les diagrammes suivant que l'Allemagne a engagé un réel changement dans sa production en accomplissant un progrès dans l'électricité verte, elle poursuit un recours important au charbon.



Ces diagrammes représentent cette fois-ci les énergies renouvelables, on constate effectivement que l'Allemagne a su développer et conserver depuis 10 ans une place importante dans la production d'électricité renouvelable. La France apparaît 3^{ème} au podium en 2018, cette croissance résulte principalement du fort développement de l'éolien, des pompes à chaleur et des biocarburants (cf. actu-environnement, Chiffres clés des énergies renouvelables, Édition 2020).

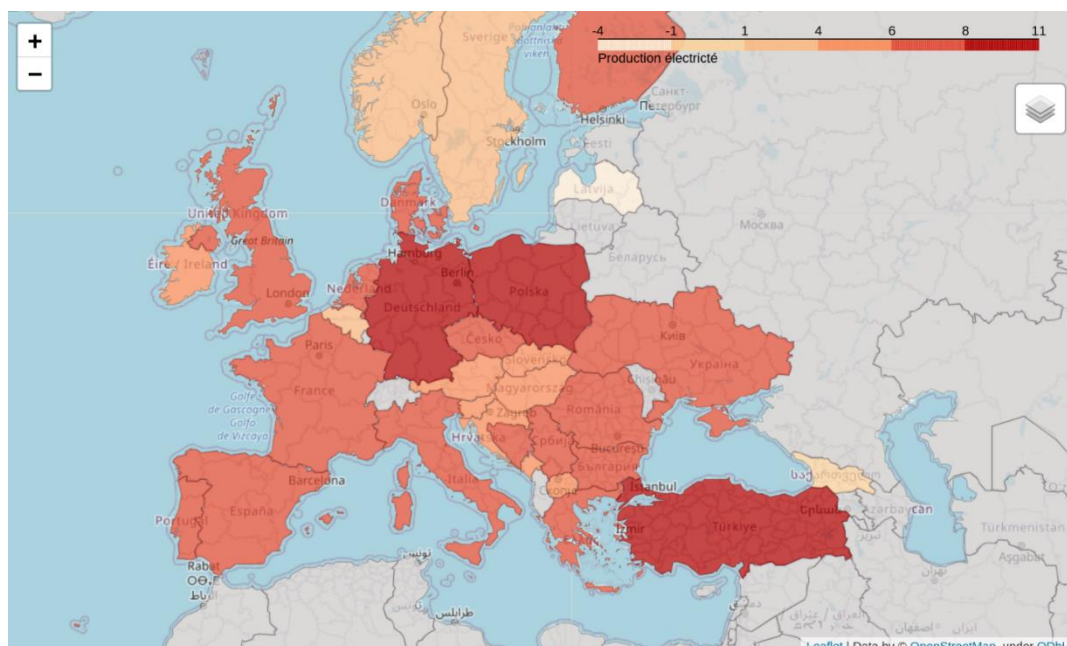


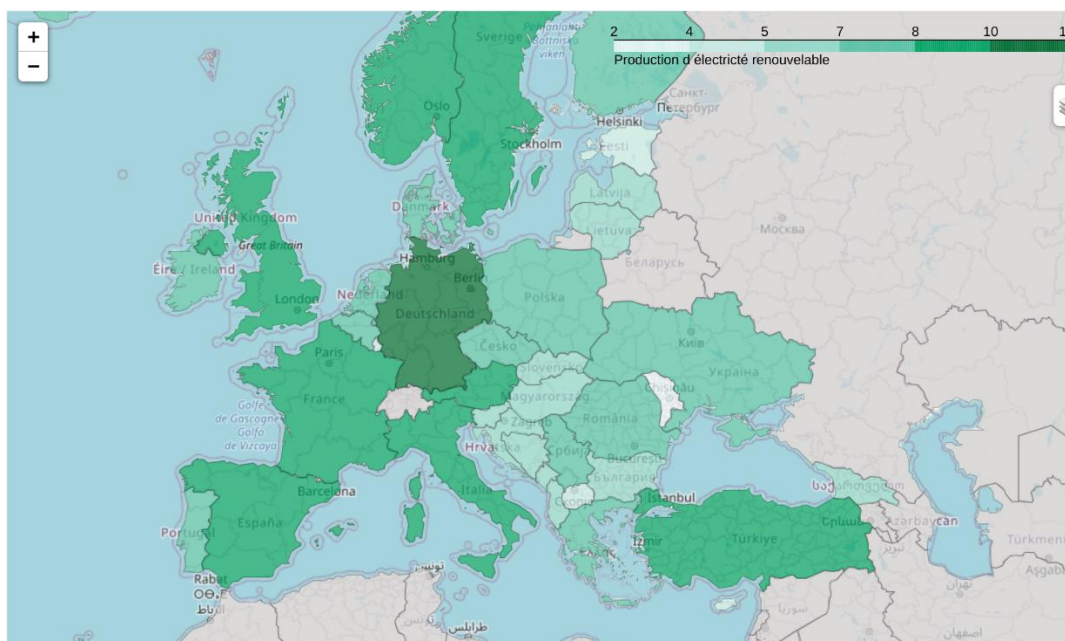
Concernant la production d'électricité totale, l'Allemagne, la France et le Royaume-Uni sont les 3 pays qui produisent le plus d'électricité en Europe. Ce sont les 3 pays ayant les populations les plus importantes d'Europe. Le besoin en électricité est donc plus important.



3. Cartographie de la production d'électricité en 2018

Pour visualiser plus facilement la production d'électricité des énergies fossiles et renouvelables, ces deux cartes permettent de constater les différences entre pays.

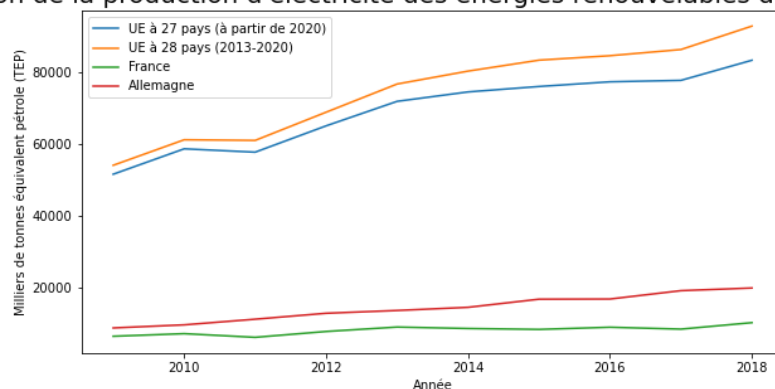




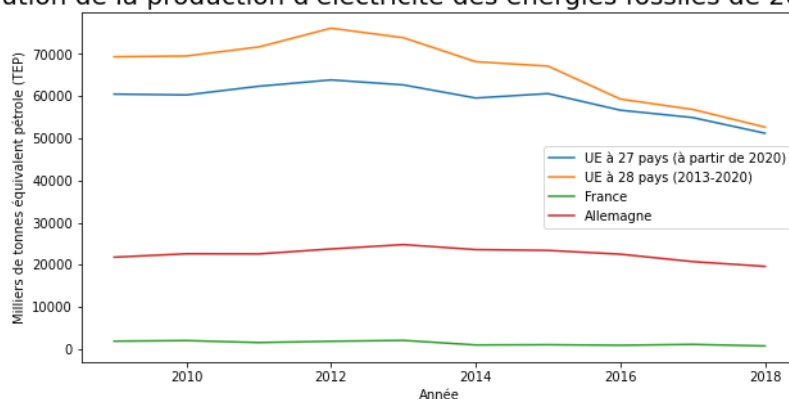
4. Evolution de la production d'électricité de 2009 à 2018

Ces graphiques représentent l'évolution de la production par type d'énergie pour la période 2009 à 2018. L'Europe à 27, l'Europe à 28, la France et l'Allemagne ont été sélectionné à titre de comparaison.

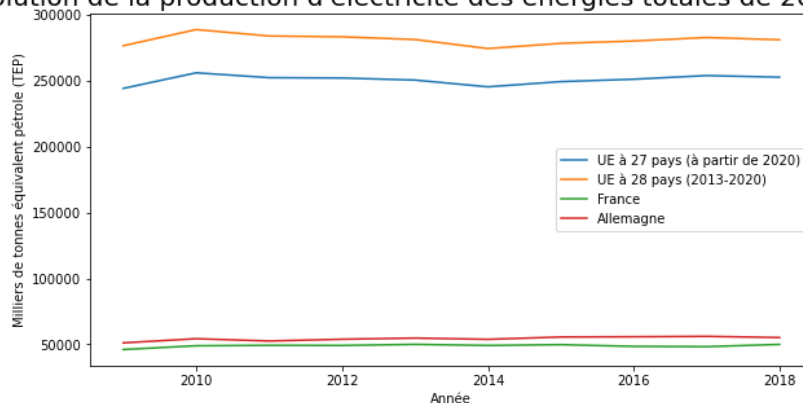
Evolution de la production d'électricité des énergies renouvelables de 2009 à 2018



Evolution de la production d'électricité des énergies fossiles de 2009 à 2018



Evolution de la production d'électricité des énergies totales de 2009 à 2018



Un bilan plutôt positif.

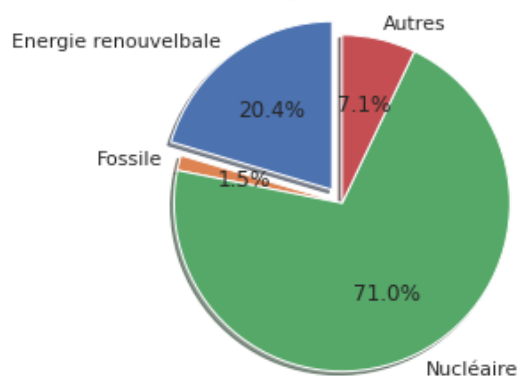
Concernant l'Europe on constate une augmentation de l'utilisation des énergies renouvelables dans la production d'électricité. Pour visualiser davantage cette augmentation on peut calculer le taux d'évolution entre 2009 et 2018 pour l'Europe à 28. En 2009, la quantité était 54 039 tep contre 92 814 tep en 2018 soit une augmentation de 72% en 10 ans.

Mais plus complexe qu'il n'y paraît.

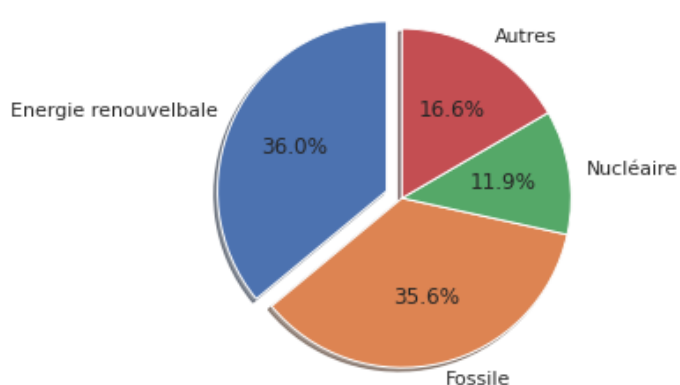
La France a vu sa part des énergies vertes augmenter de 59% en 10 ans (6 417 tep en 2009 contre 10206 tep en 2018) mais l'énergie nucléaire reste encore l'énergie la plus importante (72%).

L'Allemagne quant à elle réussit progressivement à sortir du nucléaire et voit la part de ses énergies vertes augmenter de 127% en 10 ans. Mais la part du charbon reste encore importante (30% en 2018).

Part des énergies en France en 2018



Part des énergies en Allemagne en 2018



5. La part des énergies dans la production d'électricité des pays européens (2018)

	libelle	annee	energie renouvelable	fossile	nucléaire	total
0	Albanie	2018	100.00	NaN	0.00	735.467
1	Islande	2018	99.99	NaN	0.00	1705.052
2	Norvège	2018	97.98	0.03	0.00	12601.582
3	Luxembourg	2018	87.62	NaN	0.00	189.245
4	Lituanie	2018	83.34	NaN	0.00	281.978
5	Géorgie	2018	82.59	0.10	0.00	1044.592
6	Autriche	2018	78.17	2.63	0.00	5897.163
7	Croatie	2018	72.37	10.66	0.00	1172.115
8	Danemark	2018	68.36	21.63	0.00	2611.916
9	Monténégro	2018	59.20	40.80	0.00	327.661
10	Suède	2018	55.81	0.21	41.95	14049.871
11	Lettonie	2018	52.02	0.01	0.00	578.235
12	Portugal	2018	51.37	20.13	0.00	5127.780
13	Finlande	2018	45.93	8.27	32.57	6017.380
14	Roumanie	2018	41.04	24.12	17.54	5578.372
15	Italie	2018	40.17	9.85	0.00	24858.748
16	Espagne	2018	38.76	13.61	20.33	23590.542
17	Allemagne (jusqu'en 1990 - ancien territoire d...	2018	35.99	35.56	11.85	55167.068
18	Macédoine du Nord	2018	35.05	50.80	0.00	482.130
19	Bosnie-Herzégovine	2018	34.71	64.91	0.00	1647.463
20	Irlande	2018	33.51	6.91	0.00	2677.024
21	Royaume-Uni	2018	33.45	5.09	19.67	28438.746
22	Slovénie	2018	33.22	28.30	35.37	1404.191
23	Zone euro - 19 pays (à partir de 2015)	2018	33.20	16.14	26.59	200574.261
24	Union européenne - 28 pays (2013-2020)	2018	32.99	18.72	25.27	281359.242
32	France	2018	20.42	1.45	71.04	49980.732

Dans ce tableau j'ai voulu montrer la part des énergies renouvelables, des énergies fossiles et ajouter également l'énergie nucléaire sur le total de la production d'électricité de chaque pays. J'ai utilisé mes tables des énergies renouvelables et fossiles et joint la table bilan énergétique pour récupérer la production d'électricité nucléaire. J'ai mis en avant les pays qui produisent le plus d'énergie renouvelable, ils sont donc triés par ordre décroissant sur cette colonne.

On constate que l'Albanie à une production 100% verte, l'Islande, la Norvège et le Luxembourg sont apriori les bons élèves de l'Europe. L'Europe à 28 génère un total de 33% d'électricité verte produite. L'Allemagne est 17ème avec 36% et la France 33ème avec 20%.

Grâce à ce classement nous venons de visualiser les pays qui produisent le plus d'énergie verte, il serait maintenant intéressant d'analyser les importations et le type de combustibles les plus représentés pour ces pays.

6. La part des importations par type de combustible (2018)

Certains pays ont donc misé sur la production d'énergie verte et investi dans des infrastructures pour exploiter les ressources naturelles comme les pays nordiques ou encore l'Albanie notamment avec des barrages hydrauliques. Mais regardons de plus près leur importation. Prenons pour exemple l'Albanie, l'Islande, la France et l'Europe à 28 pays afin d'analyser leurs importations et comparer avec leur production.

	libelle	annee	nrg_bal	siec	quantite_tep	Part des énergies
0	Albanie	2018	IMP	Total	1397.504	100.00
1	Albanie	2018	IMP	Énergie fossile	1198.456	85.76

	libelle	annee	nrg_bal	siec	quantite_tep	Part des énergies
0	Islande	2018	IMP	Total	1280.578	100.00
1	Islande	2018	IMP	Énergie fossile	1263.023	98.63

	libelle	annee	nrg_bal	siec	quantite_tep	Part des énergies
0	Norvège	2018	IMP	Total	11026.010	100.00
1	Norvège	2018	IMP	Énergie fossile	9902.617	89.81

	libelle	annee	nrg_bal	siec	quantite_tep	Part des énergies
0	Union européenne - 28 pays (2013-2020)	2018	IMP	Total	1493542.429	100.00
1	Union européenne - 28 pays (2013-2020)	2018	IMP	Énergie fossile	1451530.059	97.19
17	Union européenne - 28 pays (2013-2020)	2018	IMP	Renouvelables et biocarburants	22070.183	1.48

On constate que pour :

- **L'Albanie** : 86% de ses importations sont des énergies fossiles. Elle produit au total environ 735 tep mais importe 1 198 tep d'énergie fossile
- **L'Islande** : 99% des énergies fossiles sont importées. Elle produit au total environ 1 700 tep d'électricité verte mais importe 1 263 tep d'énergie fossile.
- **La France** : nous importons à 98% les énergies fossiles. Nous produisons au total environ 49 980 tep et importons environ 150 187 tep d'énergie fossile.
- **L'Europe à 28** : pour l'ensemble de l'Europe, les énergies fossiles représentent 97% des importations. 281 359 tep d'électricité sont produites en Europe avec un total de 1 493 542 tep d'électricité importé, principalement des énergies non renouvelables.

Ainsi, au travers de ces chiffres nous constatons que beaucoup de pays sont dépendants des importations et principalement pour les énergies fossiles. La totalité des importations en Europe représente 5 fois la production d'électricité. En France, nous importons 3 fois plus que nous produisons.

CONCLUSION

Au sein de l'Europe, on observe une grande hétérogénéité des énergies produites. Par exemple, la Pologne utilise encore massivement du charbon alors que la France l'a remplacé par le nucléaire ; l'Italie et le Royaume-Uni produisent leur électricité majoritairement avec du gaz naturel, tandis que la Norvège et la Suède sont presque parvenues à se détourner des hydrocarbures au profit des énergies renouvelables (notamment de l'hydroélectricité) ; et du nucléaire dans le cas de la Suède. Dans le cas de la France et de l'Allemagne, les deux États produisent à peu de chose près le même volume d'électricité, les allemands sont parvenus à produire 36% d'énergie renouvelable contre 20% pour la France en 2018. Mais l'utilisation importante du charbon en Allemagne et du nucléaire pour la France rend le tableau moins glorieux.

Ainsi, ces analyses entre 2009 et 2018 mettent en évidence les différentes politiques de chaque pays. L'exemple des pays scandinaves permettent de mieux comprendre les conditions et les moyens utilisés pour réussir leur transition énergétique. Ces bons résultats des pays scandinaves s'expliquent par la présence de ressources naturelles abondantes mais aussi par un volontarisme politique et civil, un cadre législatif souple. La réussite énergétique de ces pays tient aussi à une bonne coopération nouée au sein du Conseil des pays en cas d'intermittence (par exemple quand le vent se met à fléchir sur son territoire, le Danemark importe l'électricité hydraulique norvégienne) parallèlement à l'harmonie de la nature qui occupe une place particulière dans les mentalités et mythes scandinaves.

Il convient toutefois de nuancer ce constat : en analysant les importations nous constatons que les pays produisant beaucoup d'énergie verte importent des énergies fossiles. Indispensables, à l'utilisation dans certains domaines. Mais c'est la quantité importée et la part de certaines énergies dans la production d'un pays qui montrent que rien n'est encore joué pour remplacer en partie les énergies fossiles et nucléaires par des énergies plus vertes.

Pour aller plus loin, il serait intéressant d'ajouter à cette analyse des données sur les émissions de gaz à effet de serre pour se rendre compte réellement de l'impact sur le réchauffement climatique. Comme nous l'avons vu certains pays produisent en grande quantité des énergies renouvelables mais importent énormément d'énergie fossiles. Ainsi, en analysant les émissions de gaz à effet de serre générées par la production et les importations nous pourrions ajouter une analyse plus poussée sur qui sont réellement les pays qui polluent le moins en Europe.