

# **INTRODUCTION**

La question du climat est restée, jusque dans les années 80, du domaine des chercheurs et des experts. Ainsi, la Conférence de Stockholm1 (1972) se limite-t-elle à l'environnement. Il faut attendre le Sommet de la Terre de Rio (1992), pour voir le climat commencer à investir le champ de la politique internationale, avant de devenir le sujet de société que nous connaissons aujourd'hui. Il est vrai que les enjeux sont considérables : ils touchent la biodiversité et les écosystèmes, le cadre de vie en société, les questions sanitaires, l'économie et la finance, les relations entre les Etats (la géopolitique) pour aboutir à notre quotidien. Autant de sujets à aborder de façon globale, car la question du changement climatique met en cause nos modèles de production et de consommation, avec un développement axé largement sur des énergies carbonées. Le réchauffement global aujourd'hui n'est plus guère contesté; il ne s'agit pas d'une hypothèse lointaine avec un horizon 2100 comme on l'entend souvent. Le réchauffement a commencé; les scientifiques en dressent déjà le constat, tandis que ses effets sont désormais prévisibles.

Le réchauffement qui est en route n'est pas une question d'opinion, c'est une question de connaissance, pour ne pas dire de compétence chez les responsables: il s'agit en effet d'intégrer le processus du changement climatique, de voir en quoi nous sommes concernés et de déterminer comment nous pouvons y faire face.

Pour cela nous essayerons à répondre dans ce projet à ces questions essentielles qui se posent quand il s'agit du changement climatique et du réchauffement: Comment expliquer le changement? Comment approcher les causes du réchauffement? Quels sont les effets et conséquences possibles? Que pouvons-nous faire pour remédier ce phénomène?

Notre *objectif* est de mieux Comprendre l'évolution du réchauffement global aux États-Unis et avoir des insights sur les indicateurs qui concernent soit les causes, soit les effets de ce changement climatique. Et d'autres joueront un rôle clé dans la décarbonisation de nos systèmes énergétiques dans les décennies à venir .Collectivement, les tendances décrites dans ces indicateurs fournissent des preuves importantes de "ce à quoi ressemble le changement climatique" et les éventuelles relations entre ces indicateurs.

A la fin de ce projet nous allons pouvoir de savoir les causes, effets de réchauffement global à U.S. ainsi on va savoir si les indicateurs proposés sont des solutions efficaces pour remédier ce problème.

# CONCEPTION DU PROJET

# > Dataset

Les sources de notre ensemble de données sont les suivantes en les classant en trois catégories:

#### ✓ Causes:

- <a href="https://www.epa.gov/climate-indicators/climate-change-indicators-residential-energy-use">https://www.epa.gov/climate-indicators/climate-change-indicators-residential-energy-use</a>
- https://data.worldbank.org/indicator/SP.POP.TOTL?locations=US

#### ✓ Effets:

- <a href="https://www.epa.gov/climate-indicators/climate-change-indicators-high-and-low-temperatures">https://www.epa.gov/climate-indicators/climate-change-indicators-high-and-low-temperatures</a>
- https://www.epa.gov/climate-indicators/climate-change-indicators-seasonaltemperature
- https://www.epa.gov/climate-indicators/climate-change-indicators-sea-surfacetemperature
- https://www.epa.gov/climate-indicators/climate-change-indicators-drought
- <a href="https://www.kaggle.com/dataenergy/natural-disaster-data">https://www.kaggle.com/dataenergy/natural-disaster-data</a>
- <a href="https://www.epa.gov/climate-indicators/climate-change-indicators-us-greenhouse-gas-emissions">https://www.epa.gov/climate-indicators/climate-change-indicators-us-greenhouse-gas-emissions</a>

#### ✓ Solution:

https://ourworldindata.org/renewable-energy

# Indicateurs proposés pour le calcul

Nous pouvons décomposer nos indicateurs en quatre parties afin d'attaquer ce sujet sous différents angles et de l'analyser avec différentes visions.

# A. Comment détecter que notre climat est en train de changer?

- 1. Comment les émissions de gaz à effet de serre évoluent actuellement?
- 2. L'évolution de la température au cours du temps.
- 3. L'évolution de la température par saison.
- 4. L'évolution de la température de surface de la mer au fil du temps.

# B. Quels sont les causes de ce changement climatique ?

- 1. L'impact de l'évolution de nombre de la population de l'U.S. sur le taux de consommation des énergies non renouvelable.
- 2. Le taux d'émissions de gaz à effet de serre par secteurs.
- 3. Quel est le secteur d'activité qui impacte le plus sur le climat ?
- **4.** L'impact de la consommation des énergies non renouvelable sur l'évolution d'émission des gaz à effet de serre.

# C. Où l'on peut voir les effets du réchauffement climatique ?

- 1. Quel gaz est le plus concentré dans l'atmosphère et représente une cause directe de ce changement climatique ?
- 2. L'impact d'émission du gaz à effet de serre sur la température ?
- **3.** Impact de l'augmentation de la température sur l'apparition de la sécheresse & des catastrophes naturelles?
- 4. Taux de dégât économiques dus aux catastrophes naturelles ?

### D. Comment on peut remédier ce phénomène (solution)?

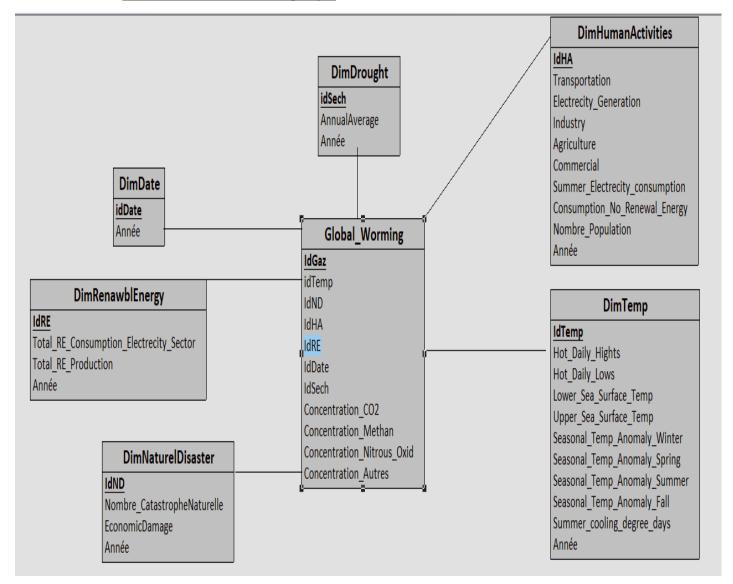
- 1. Taux d'évolution de la consommation en énergie d'origine décarbonée (énergie renouvelable) dans le secteur « électricité ».
- **2.** Taux de consommation des énergies renouvelable *VS* énergie non renouvelable dans le secteur d'activité au fil du temps.
- **3.** Quelle vitesse la production d'énergie renouvelable évolue-t-elle aux Etats Unis?
- **4.** Comparez la consommation d'énergie renouvelable et l'émission de gaz à effet de serre. Représente-t-elle une solution pour remédier à ce réchauffement climatique?

# **Matrice Dimensionnelle**

 Nous utiliserons dans cette matrice les indicateurs en fonction de leur catégorie, que j'ai déjà mentionnée en montrant la relation de chacun avec les axes d'analyse (dimension).

Axes d'analyse  Indicateurs	Température	Sécheresse	Natural Disaster	Human Activities	Renawble Energy	Date
Catégorie A	*					*
Catégorie B				*		*
Catégorie C	*	*	*			*
Catégorie C Catégorie D				*	*	*

# Modélisation du notre projet



# Outils utilisés :

- Modélisation: Merise, Access
- Pré-traitement des données : Open-Refine
- ETL: Microsoft SQL Server Management Studio
- Analyse et prédiction : Python

# REALISATION DU PROJET

Nettoyage

ETL

Analyse

Visualisation

Prédiction

- J'utilise le logiciel « OpenRefine » afin de réaliser cette étape de Nettoyage.
- OpenRefine?

C'est un logiciel libre de nettoyage et de mise en forme de données, qui utilise le langage GREL afin de faciliter la préparation des données.

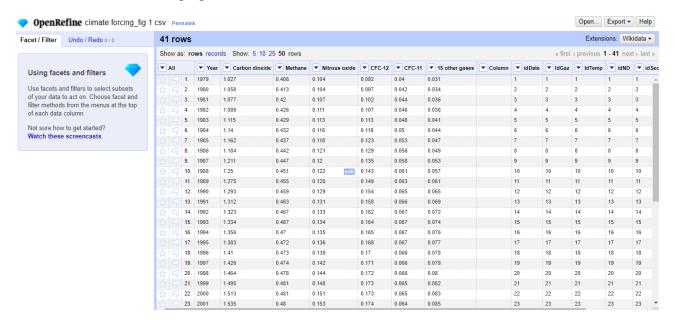


IMAGE1: « VISUALISATION DE L'OUTIL OPENREFINE »

- ✓ Au niveau de cette étape de nettoyage, on va identifier les valeurs manquantes , en le remplaçant par cette expérience : « N/A ».
- ✓ On va supprimer aussi les espaces initiaux et finaux.
- ✓ Supprimer colonnes vides

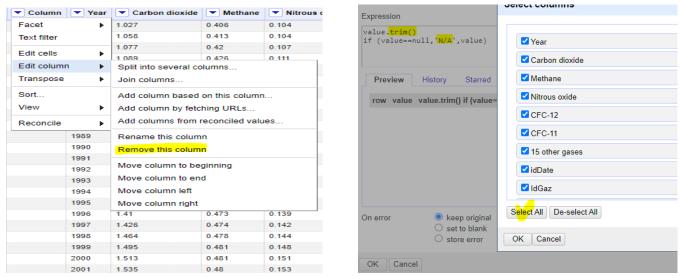


IMAGE2: « NETTOYAGE DES DONNEES »

✓ Après nettoyage, on exporte notre jeu donné en extension csv.

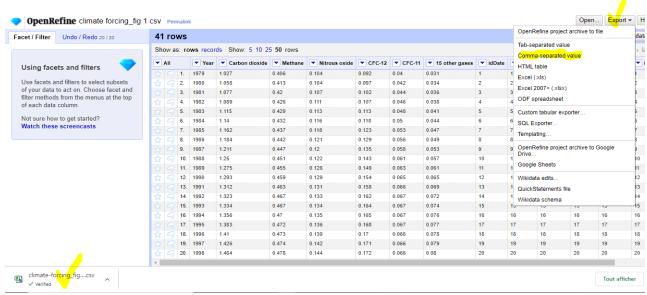


IMAGE3: « EXPORTATION DES DONNEES EN FORMAT CSV »

# Nettoyage

#### ETL

# Analyse

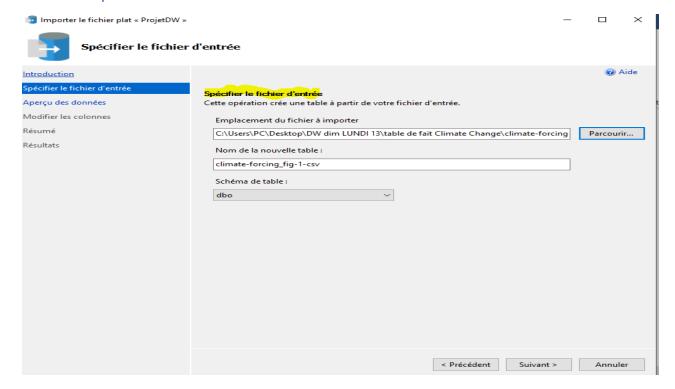
# Visualisation

Prédiction

- Avant de commencer la partie ETL, on doit :
- I. Créer une BD nommée « ProjectDW » ou nous téléchargerons nos dataset → Cette BD représente la source à partir de laquelle nous allons extraire nos données.
  - CREATION DE LA BD

#### CREATE DATABASE ProjetDW

- TELECHARGER NOS 13 DATASET:
  - ProjetDW >> Taches >> Importer le fichier plat
  - ♣ Spécifier le fichier d'entrée



# Aperçu des données

#### Apercu des données

Cette opération a analysé la structure du fichier d'entrée pour générer l'aperçu ci-dessous pour les 50 premières lignes.

Year	Carbon_dioxide	Methane	Nitrous_oxide	CFC_12	CFC_11
1979	1.027	0.406	0.104	0.092	0.04
1980	1.058	0.413	0.104	0.097	0.042
1981	1.077	0.42	0.107	0.102	0.044
1982	1.089	0.426	0.111	0.107	0.046
1983	1.115	0.429	0.113	0.113	0.048
1984	1.14	0.432	0.116	0.118	0.05
1985	1.162	0.437	0.118	0.123	0.053
1986	1.184	0.442	0.121	0.129	0.056
1987	1.211	0.447	0.12	0.135	0.058
1988	1.25	0.451	0.122	0.143	0.061
1989	1.275	0.455	0.126	0.149	0.063
1990	1.293	0.459	0.129	0.154	0.065
1991	1.312	0.463	0.131	0.158	0.066
<			·	1	>

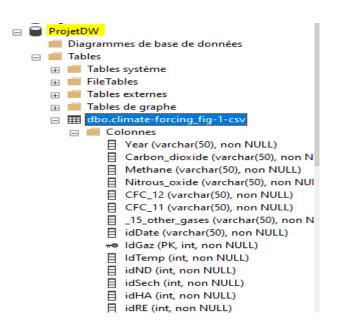
Utilisez la détection de type de données riche qui peut être mieux adaptée. Toutefois, les cellules avec des valeurs anormales peuvent être supprimées.

< Précédent	Suivant >	Annuler

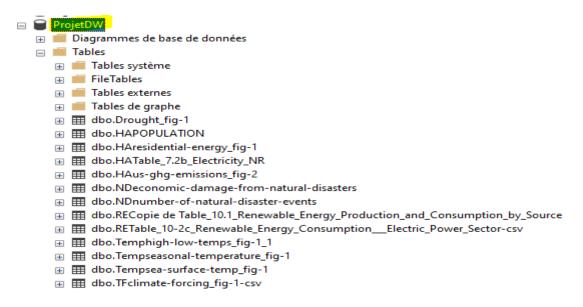
#### Modifier les colonnes

Modifier les colonn Cette opération a généré le schéma de table suivant. Vérifiez si le schéma est correct et, dans le cas contraire, apportez-y les modifications nécessaires. Nom de la colonne Type de données Year varchar(50) -Carbon\_dioxide varchar(50) Methane varchar(50) Nitrous\_oxide varchar(50) CFC\_12 varchar(50) CFC\_11 varchar(50) \_15\_other\_gases varchar(50) idDate int IdGaz int - $\overline{\phantom{a}}$ ldTemp int idND int idSech int idHA int idRE int

#### Résultat



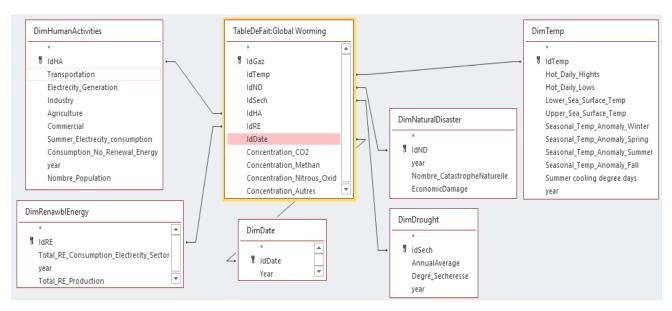
➡ Visualiser les tables créer de notre BD « ProjectDW »



- II. On va créer une deuxième BD nommée « Output » ou on va créer structure/schéma de nos tables → Cette BD représente la source de destination ou on va charger nos dataset.
  - CREATION DE LA BD « OUTPUT »

#### CREATE DATABASE Output

- CREATION DES TABLES (SCHEMA)
- On va suivre ce schéma ci-dessous créer avec SGBD « Access » dans la création de ces tables.



#### 1) DIM DATE:

```
CREATE TABLE dbo.Date
(
IdDate INT NOT NULL IDENTITY(1,1),
Year DATE NULL,
CONSTRAINT PK_Dates PRIMARY KEY (IdDate)
);
GO
```

#### 2) DIM NATURALDISASTER

```
CREATE TABLE dbo.NaturalDisatser
(
IdND INT NOT NULL IDENTITY(1,1),
Year DATE NULL,
Nombre_Catastrophe_Naturelle INT NULL,
Economic_Damage Float NULL
CONSTRAINT PK_NaturalDisaster PRIMARY KEY (IdND)
);
GO
```

#### 3) DIM REWABLENERGY

```
CREATE TABLE dbo.RenawblEnergy
(
IdRE INT NOT NULL IDENTITY(1,1),
Year DATE NULL,
Total_RE_Consumption_Electrecity_Sector INT NULL,
TOTal_RE_Production Float NULL,
CONSTRAINT PK_RenawblEnergy PRIMARY KEY (IdRE)
);
```

#### 4) DIM HUMANACTIVITIES

```
CREATE TABLE dbo.HumanActivities
(
IdHA INT NOT NULL IDENTITY(1,1),
Year DATE NULL,
Nombre_Population INT NULL,
Transportation INT NULL,
Electrecity_Generation Float NULL,
Industry Float NULL,
Agriculture Float NULL,
Commercial Float NULL,
Summer_Electrecity_consumption Float NULL,
Consumption_No_Renewal_Energy Float NULL,
CONSTRAINT PK_HumanActivities PRIMARY KEY (IdHA)
);
```

#### 5) DIM DROUGHT

```
CREATE TABLE dbo.Drought
(
IdSech INT NOT NULL IDENTITY(1,1),
Year DATE NULL,
Annual_Average Float NULL,
Degre_Secheresse varchar(50) NULL,
CONSTRAINT PK_Drought PRIMARY KEY (IdSech)
);
GO
```

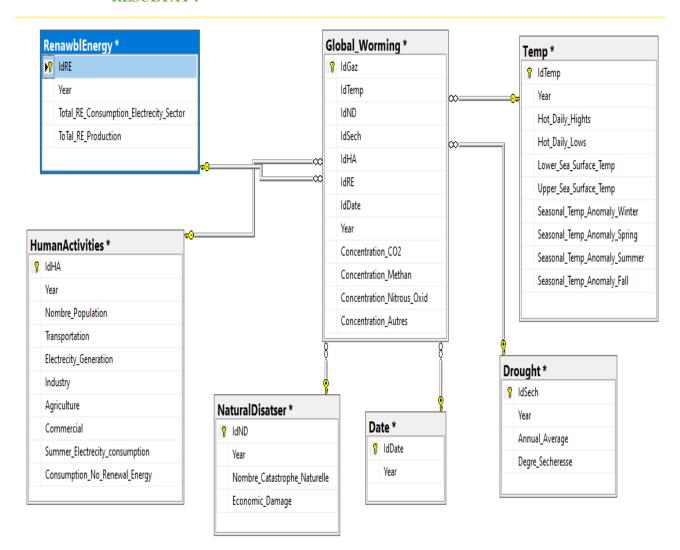
#### 6) DIM TEMP

```
CREATE TABLE dbo.Temp
(
IdTemp INT NOT NULL IDENTITY(1,1),
Year DATE NULL,
Hot_Daily_Hights Float NULL,
Lower_Sea_Surface_Temp Float NULL,
Upper_Sea_Surface_Temp Float NULL,
Seasonal_Temp_Anomaly_Winter Float NULL,
Seasonal_Temp_Anomaly_Spring Float NULL,
Seasonal_Temp_Anomaly_Summer Float NULL,
Seasonal_Temp_Anomaly_Fall Float NULL,
CONSTRAINT PK_Temp PRIMARY KEY (IdTemp)
);
GO
```

#### 7) TABLE DE FAIT: GOBAL WORMING

```
CREATE TABLE dbo.Global_Worming
IdGaz INT NOT NULL IDENTITY(1,1),
IdTemp INT NOT NULL ,
IdND INT NOT NULL,
IdSech INT NOT NULL,
IdHA INT NOT NULL ,
Idre Int Not Null,
IdDate INT NOT NULL ,
Year DATE NULL,
Concentration_CO2 Float NULL,
Concentration_Methan Float NULL,
Concentration_Nitrous_Oxid Float NULL,
Concentration_Autres Float NULL,
CONSTRAINT PK Global Worming PRIMARY KEY (IdGaz)
);
GO
```

### • **RESULTAT**:



→ MAINTENANT ON VA COMMENCER ETL AU NIVEAU DU VISUAL STUDIO:

• Au niveau du Visual studio, vous trouverez CJ les requêtes SQL utilisées pour chaque table :

**♣** DIM DATE

SELECT [Year], idDate

FROM [climate-forcing\_fig-1-csv]

#### **♣** DIM TEMP

SELECT [Tempsea-surface-temp\_fig-1].[Year], [Tempsea-surface-temp\_fig-

1].Lower\_95\_confidence\_interval, [Tempsea-surface-temp\_fig-

1]. Upper 95 confidence interval, [Temphigh-low-temps fig-1 1]. Hot daily highs,

[Temphigh-low-temps\_fig-1\_1].Hot\_daily\_lows, [Temphigh-low-temps\_fig-

1\_1].Winter, [Temphigh-low-temps\_fig-1\_1].Spring, [Temphigh-low-temps\_fig-

1\_1].Summer, [Temphigh-low-temps\_fig-1\_1].Fall,

[Tempsea-surface-temp\_fig-1].IdTemp

FROM [Temphigh-low-temps\_fig-1\_1] INNER JOIN

[Tempsea-surface-temp\_fig-1] ON [Temphigh-low-temps\_fig-1\_1].idTemp = [Tempsea-surface-temp\_fig-1].IdTemp

#### ♣ DIM HA

SELECT HAPOPULATION.[Year], HAPOPULATION.IdHA, HAPOPULATION.Population, HAPOPULATION.Transportation\_Million\_metroc\_tons\_CO2, HAPOPULATION.Electricity\_generation, HAPOPULATION.Industry,

HAPOPULATION.Agriculture, HAPOPULATION.Commercial,

HAPOPULATION.Electricity\_Net\_Generation\_From\_Coal\_Electric\_Power\_Sector\_Million Kilowatthours,

HAPOPULATION. Electricity Net Generation From Petroleum Electric Power Sector,

HAPOPULATION.Electricity\_Net\_Generation\_From\_Natural\_Gas\_Electric\_Power\_Sector

HAPOPULATION.Electricity\_Net\_Generation\_From\_Other\_Gases\_Electric\_Power\_Sector,

 ${\tt HAPOPULATION.Electricity\_Net\_Generation\_From\_Nuclear\_Electric\_Power\_Electric\_Power\_Sector\,,}$ 

[HAresidential-energy\_fig-1].Summer\_electricity\_use\_per\_capita\_KWh\_P

FROM HAPOPULATION INNER JOIN

[HAresidential-energy\_fig-1] ON HAPOPULATION.IdHA = [HAresidential-energy\_fig-1].IdHA

#### **↓** DIMRENAWBLENERGY

SELECT [RECopie-de-Table\_10-1 Renewable Energy Production and Consumption by Source-csv (1)]. Annual Total, [RECopie-de-Table\_10-1\_Renewable\_Energy\_Production\_and\_Consumption\_by\_Sourcecsv (1)].IdRE, [RETable\_10-2c Renewable Energy Consumption Electric Power Sectorcsv].Total\_Renewable\_Energy\_Consumed\_by\_the\_Electric\_Power\_Sector, [RECopie-de-Table\_10-1\_Renewable\_Energy\_Production\_and\_Consumption\_by\_Source-csv (1)].Total Renewable Energy Production FROM [RECopie-de-Table 10-1\_Renewable\_Energy\_Production\_and\_Consumption\_by\_Source-csv (1)] INNER JOIN [RETable 10-2c\_Renewable\_Energy\_Consumption\_\_\_Electric\_Power\_Sector-csv] ON [RECopie-de-Table\_10-1\_Renewable\_Energy\_Production\_and\_Consumption\_by\_Source-csv (1)].IdRE = [RETable 10-2c Renewable Energy Consumption | Electric Power Sector-csv].IdRE

#### **♣** DIM NATURELDISASTER

SELECT [NDnumber-of-natural-disaster-events].Number\_of\_reported\_natural\_disasters\_reported\_disasters, [NDnumber-of-natural-disaster-events].idND, [NDnumber-of-natural-disaster-events].[Year],

[economic-damage-from-natural-disasters\_US

FROM [economic-damage-from-natural-disasters] INNER JOIN

[NDnumber-of-natural-disaster-events] ON [economic-damage-from-natural-disasters].IdND = [NDnumber-of-natural-disaster-events].idND

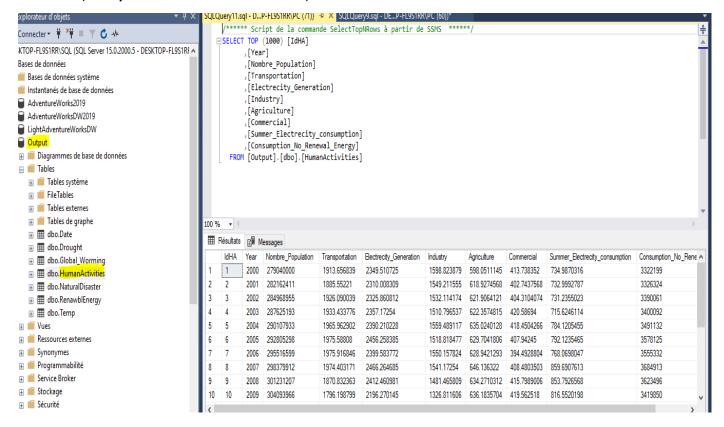
Where ([NDnumber-of-natural-disaster-events].[Year]>=2000 and [NDnumber-of-natural-disaster-events].[Year]<2019)

#### **♣** GLOBALWORMING

SELECT [Year], Carbon\_dioxide, Methane, Nitrous\_oxide, CFC\_12, CFC\_11, \_15\_other\_gases, idDate, IdGaz, idRE, idHA, idSech, idND, IdTemp

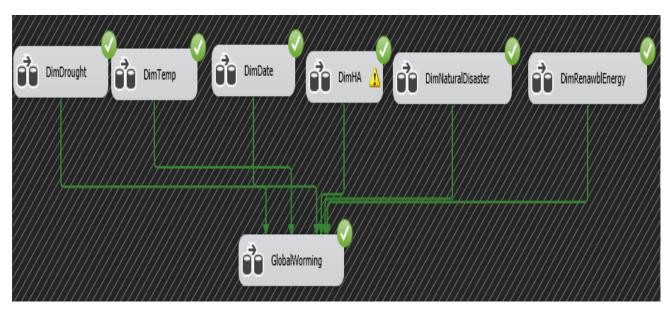
FROM [climate-forcing\_fig-1-csv]

Après l'exécution de ces requêtes, on consulte notre base de données au niveau de Management Studio (*Exemple de HumanActivities Table*) :



• Concernant la partie Transformation : j'ai travaillé avec : colonne dérivée (créer des nouvelles colonnes à partir des colonnes existants), conversion de données , trie de données.

#### **APRES DEBOGAGE:**



Nettoyage > ETL Analyse > Visualisation > Prédiction

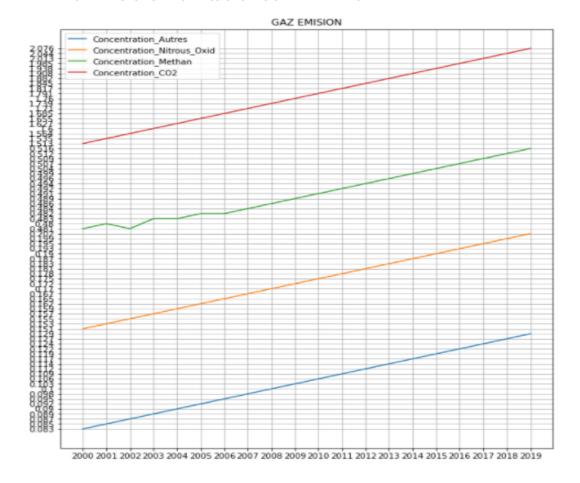
♣ J'ai utilisé Python comme outil pour faire l'analyse :

### → LIER MSSQL & PYTHON:

◆ On va essayer d'analyser ce sujet sous 4 angles differents :

# ? COMPRENDRE CE CHANGEMENT CLIMATIQUE ET COMMENT LE DETECTER :

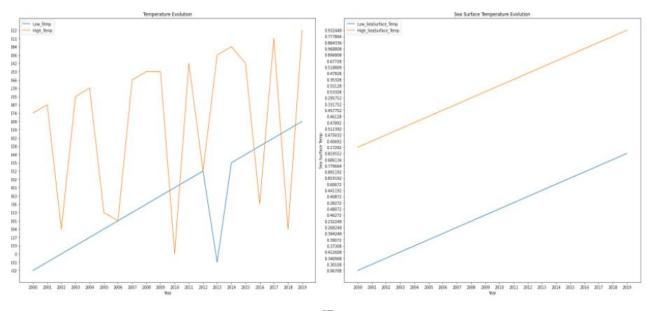
ANALYSER L'EVOLUTION LES EMISSIONS DES GAZ A EFFET DE SERRE

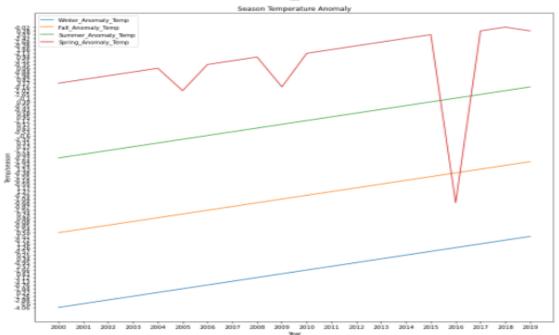




On remarque dans ce graphique de 2000 à 2019, l'émission des gaz à effets de serre (CO2, Nitrouis\_Oxid, Methan,Autre) à U.S. a augmenté de façon exponentielle. Et on trouve le CO2 en premier place.

#### • TEMPERATURE





Le premier graphique, nous voulons analyser l'évolution de la hight\_temperature & lowest\_temperature de 2000 à 2019 aux États-Unis. Et nous voyons qu'il y a des variations dans ces deux mesures mais les deux continuent à augmenter.



Concernant le deuxième graphique, il montre l'évolution de la température moyenne de surface des océans des États-Unis depuis 2000, et nous pouvons voir qu'il y a toujours une augmentation de la température de surface supérieure & inférieure de la mer.

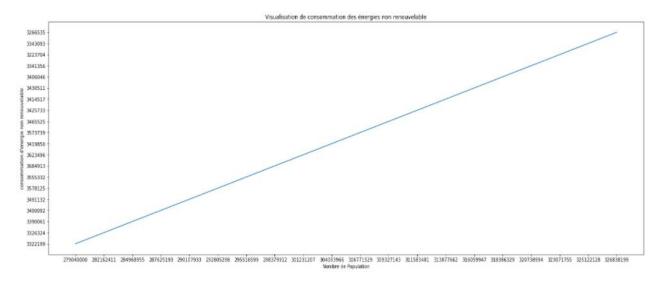
Le troisième graphique montre l'évolution de la température moyenne de chaque saison aux États-Unis entre 2000 et 2020. Les saisons sont définies comme étant l'hiver

(décembre, janvier, février), le printemps (mars, avril, mai), l'été (juin, juillet, août) et l'automne (septembre, octobre, novembre).

Tous ces changements climatiques qui n'a pas cessé de se dégrader, nous aide à comprendre ce qu'est un réchauffement climatique. Mais quels sont ces causes ?ces effets ?Comment on peut le réduire ? c'est ce qu'on va découvrir par la suite .

# ? DECOUVRIR LES CAUSES PRINCIPALES DERRIERE CE CHANGEMENT

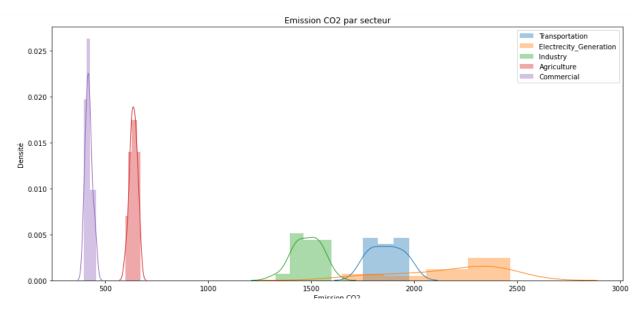
#### ANALYSER NOMBRE DE POPULATION AUX U.S. VS LE TAUX DE CONSOMMATION DES ENERGIES



Ţ

On constate que la consommation des énergies non renouvelable augmente avec l'accroissement de la population, ce qui a un impact sur notre climat.

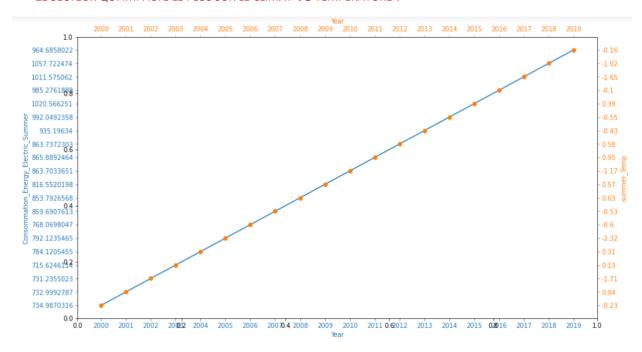
### • EMISSION DU CO2 PAR SECTEUR D'ACTIVITE





Nous pouvons voir sur ce graphique que les activités humaines sont responsables de l'émission de CO2, nous remarquons que le secteur de l'électricité arrive en première position, puis nous trouvons le transport, l'industrie, l'agriculture et enfin le secteur commercial.

#### • LE SECTEUR QUI IMPACTE LE PLUS SUR LE CLIMAT **VS** TEMPERATURE :



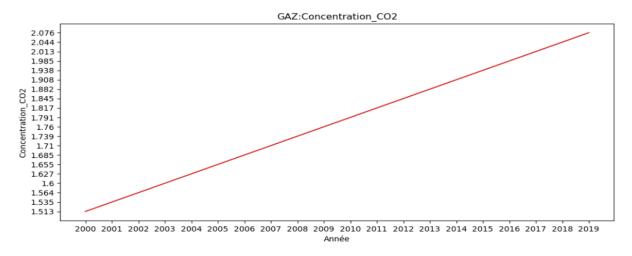


Dans ce graphique, nous essayons de voir comment la consommation d'énergie électrique pendant la saison d'été (dans notre cas) évolue en fonction des changements de température pendant la même période.

Nous avons remarqué que le secteur de l'électricité est le secteur qui émet le plus de CO2, et nous voyons dans ce graphique que la consommation de cette énergie augmente de plus en plus avec le changement de température, ce qui mettra notre planète dans une situation difficile à gérer.

# ? ANALYSER LES EFFETS DE CE CHANGEMENT

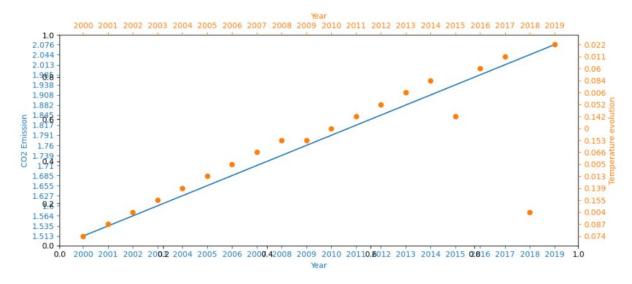
• EVOLUTION DE L'EMISSION DU CO2 PAR RAPPORT AU TEMPS.





Nous constatons une augmentation exponentielle du CO2 à partir de l'année 2000 et n'a pas cessé de diminuer.

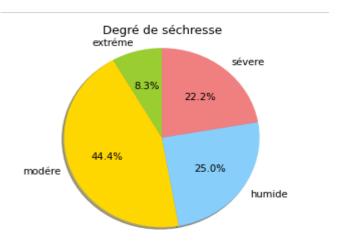
#### L'EFFET DE CE GAZ SUR LA TEMPERATURE

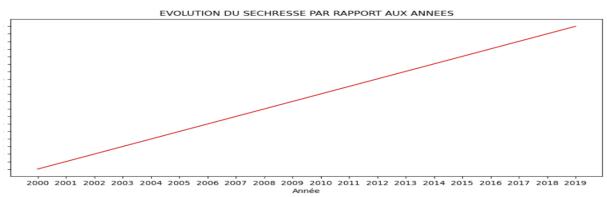




Nous constatons que la température augmente avec l'augmentation de l'émission du CO2 en air. Ce qui assure que le gaz CO2 est la cause principale de ce changement climatique et celui qui a dû à la naissance du réchauffement climatique.

#### • LE PHÉNOMÈNE DE LA SÉCHERESSE

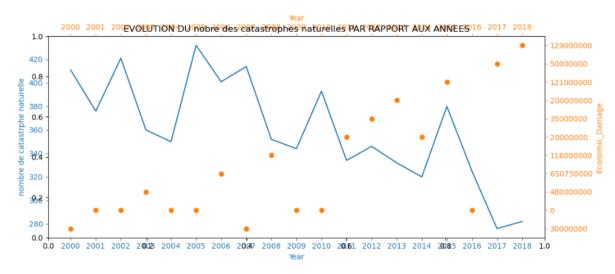






Nous remarquons que la sécheresse ne cesse pas de se dégrader année après l'autre avec un % de 22.2 de sécheresse sévère 8.3% de sécheresse extrême. La chose qui montre les effets dangereux de ce réchauffement climatique sur U.S. et sur la planète en général.

Nombre de catastrophes naturelles et leur damage economique (2000\_2019)



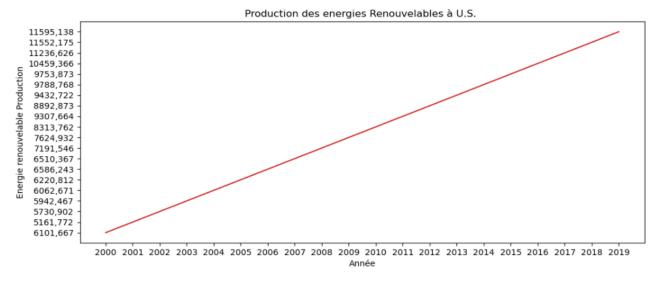


Nous remarquons dans ce graphique que le nombre de catastrophes naturelles a diminué de 2016 à 2018 mais les dommages économiques qu'elles ont causés sont les plus importants de toutes les années précédentes ce qui montre le degré de sévérité de ces catastrophes naturelles.

? VISUALISER EST CE QUE LES ENERGIES RENOUVELABLES REPRESENTENT UNE SOLUTION AU PROBLEME DU CLIMAT.

Pour réduire le CO2 la pollution atmosphère locale le monde doit rapidement évoluer vers des sources d'énergie à faible émission de carbone comme les technologies nucléaires et *renouvelables*.

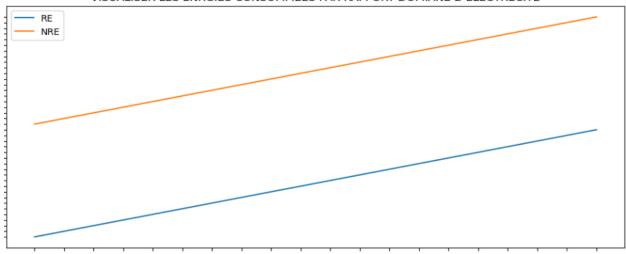
les énergies renouvelables joueront un rôle clé dans la décarbonisation de nos systèmes énergétiques dont les destinées à venir mais à quelle vitesse la production d'énergie renouvelable évolue-t-elle à U.S.?





Nous remarquons que la production des énergies renouvelable augmentent à partir 2000 à 2019 aux Etats-Unis d'une manière rapide.

#### VISUALISER LES ENRGIES CONSOMMEES PAR RAPPORT DOMIANE D'ELECTRECITE



2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019



On constate que les énergies non renouvelables sont toujours consommées par les Américains pour répondre à leurs besoins puisque les énergies renouvelables ne suffisent pas s'ils les utilisent seules.

→LES ÉNERGIES RENOUVELABLES SONT DONC UNE SOLUTION QUI NE PERMET QUE DE REMÉDIER AU PROBLÈME DU RÉCHAUFFEMENT CLIMATIQUE.

Nettoyage ETL Analyse Visualisation Prédiction

Cette phase est très intéressante car les scientifiques doivent utiliser leurs modèles (informations sur les climats passés) pour prévoir les climats futurs. Cela leur permettra de prévoir ce qui va se passer dans le futur et d'essayer de remédier à ce problème climatique.

Dans ce projet, je vais faire des prédictions du gaz CO2 (le plus émis dans l'air) en utilisant le modèle de régression linéaire (ML) et essayer de trouver de bonnes prédictions à la fin.

#### Création du modèle RL :

```
# Train set e Test set
X = df_fus[['Year','NO Renawbale Energy consumption']] #input or attributes
Y = df fus[['CO2']] # output or label
X train, X test, y train, y test = train test split(X, Y, random state=1 ,shuffle=False)
regression = LinearRegression()
regression.fit(X_train, y_train)
y_pred = regression.predict(X_test)
pred_y=pd.DataFrame(pred_y)
#prediction = pd.DataFrame({'NRE_Consumption': [i[0] for i in X_test], 'CO2_Emission': [k for k in pred_y]})
#prediction
#Creer nouveau Dataframe combine entre X_test[0] et y_predit
y_pred=pd.DataFrame(y_pred,columns=['pred'])
  # initialize list of lists
OUTPUT = [[X_test.Year.iloc[0],y_test.CO2.iloc[0], y_pred.pred.iloc[0]],[X_test.Year.iloc[1],y_test.CO2.iloc[1], y_pred.pred.iloc[0]
 # Create the pandas DataFrame
df = pd.DataFrame(OUTPUT , columns = ['Year','CO2', 'CO2_Pred'])
df
```

#### Visualiser les résultats trouvés

```
#Définir la colonne datetime comme indice de notre dataframe
df = df.set_index(['Year'])
df
```

$\sim$	~~	CO2	
			PIGO

Year		
2015	1.938	1.932374
2016	1.985	1.959659
2017	2.013	1.986315
2018	2.044	2.015790
2019	2.076	2.042935

# • Evaluer ces résultats :

	r2_score	min	max	mean	mse	mae	rmse	taux_erreur
CO2	0.710765	1 938	2 076	2 0112	0.000655	0.023785	0.025593	0.012328

```
# Visualisation des prédictions
df[['CO2','CO2_Pred']].plot()
plt.rcParams["figure.figsize"] = (20,5)

208
206
204
202
200
198
196
194
205.0 2015.5 2016.0 2016.5 2017.0 2017.5 2018.0 2018.5 2019.0
```

→On voit que notre modèle nous a donnée des bonnes prédictions (r2\_square de 70%) avec un RMSE de 2% et avec un taux d'erreur de 1%.