Guide pour le TP SQL – M1 GLET

Version Windows

# Objectifs

Manipuler des données produites par le SIE Français avec Postgres/Postgis et un Jupyter Notebook.

En cas de problème n’hésitez pas à me contacter soit :

* Par mail : [n.dhuygelaere@oieau.fr](mailto:n.dhuygelaere@oieau.fr)
* Par Skype : n.dhuygelaere

L’entraide entre membre de la promotion est autorisée à condition que les rendus finaux ne soient identiques entre les étudiants.

Le rendu du Jupyter Notebook servira d’évaluation pour le module bases de données. Le rendu doit être envoyé sur ma boite mail avant le vendredi 15 mai 14H00.

La trame du Notebook est disponible sur Github à l’adresse : <https://github.com/ndhuygelaere/tp_sql>

# Prérequis

* Un accès Internet depuis la machine d’installation
* Windows 10 (le manuel utilise un exemple réalisé avec un Windows 10 64 bits, si vous utilisez une machine 32 bits vous devrez choisir les composants logiciels compatibles).

Les autres versions de Windows peuvent également fonctionner mais n’ont pas été testées, en cas de problème contactez moi.

Une installation sous un système Linux type Ubuntu fonctionne également très bien. Si vous choisissez cette option, contactez-moi pour avoir plus de détail sur les paquets nécessaires pour l’installation.

* RAM : 4 Go en plus du fonctionnement de base de la machine (un PC ou un portable récent fonctionne correctement)

# Installation de Postgres et des programmes annexes

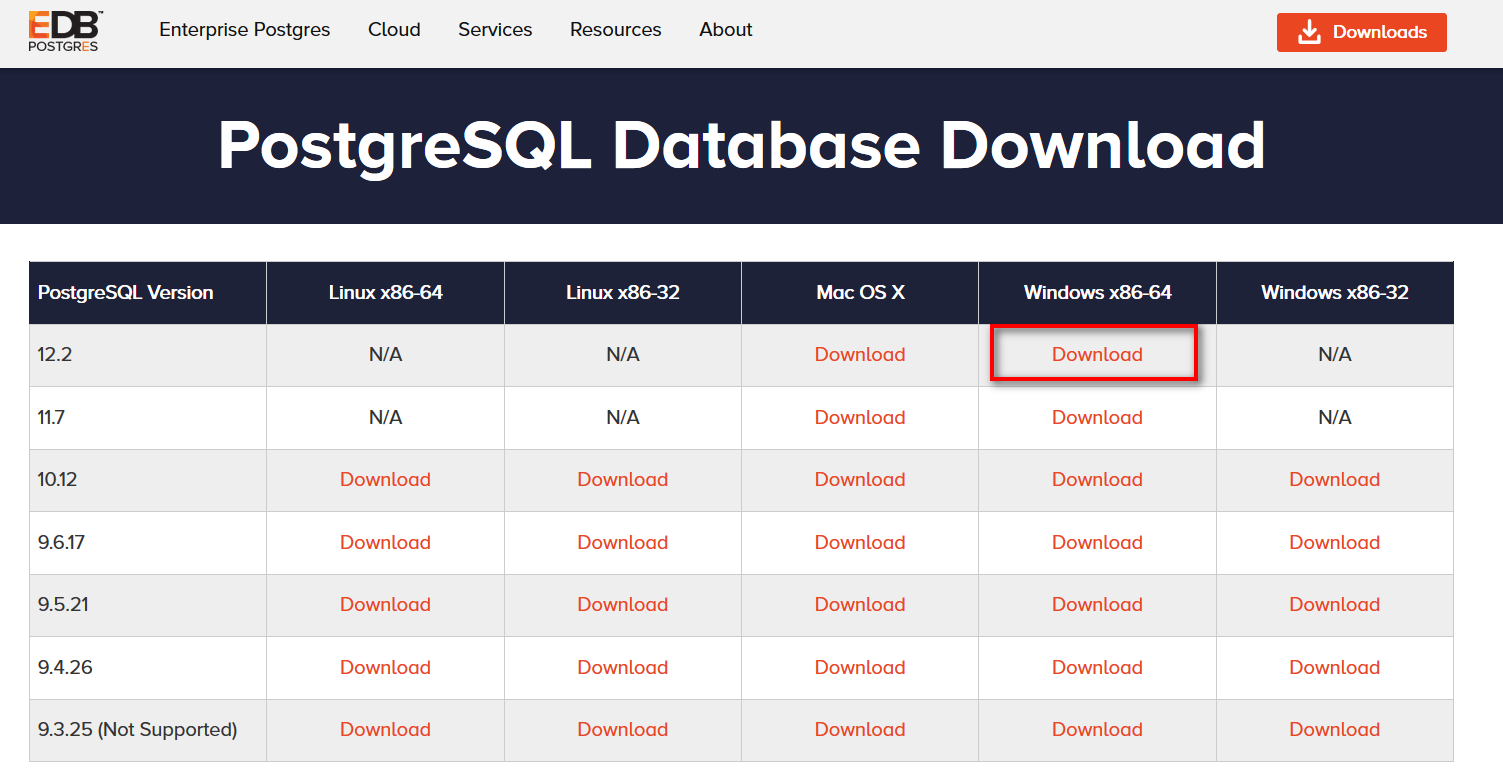
Postgres est le serveur de base de données qui sera utilisé pour ce dernier TP SQL du module base de données.

Le TP est réalisable avec les toutes les versions supérieures à la version 9.x de Postgres. Dans la mesure du possible, il est préférable de travailler avec la dernière version de Postgres c’est-à-dire la version 12.

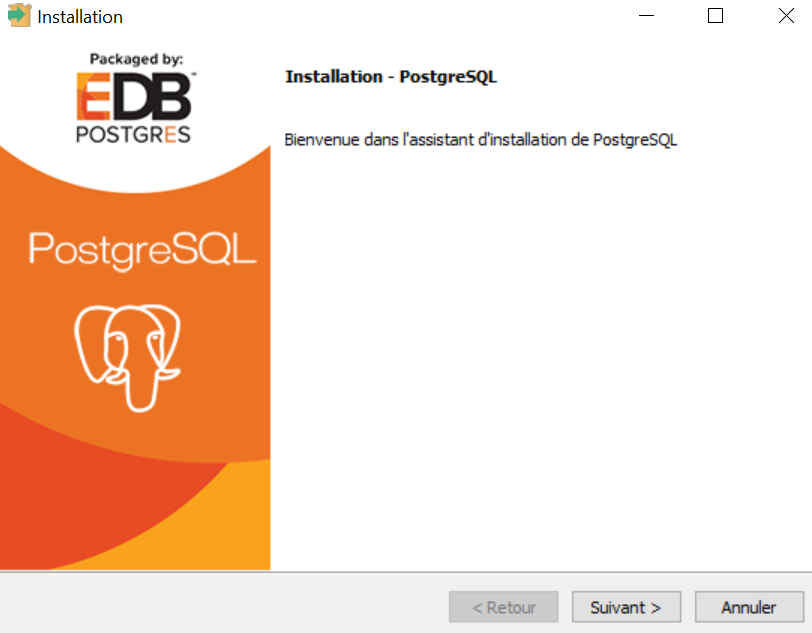
Le site officiel du projet Postgres propose différents exécutables d’installation et recommande en particulier d’utiliser ceux disponibles sur <https://www.enterprisedb.com>

L’installation suit les étapes suivantes :

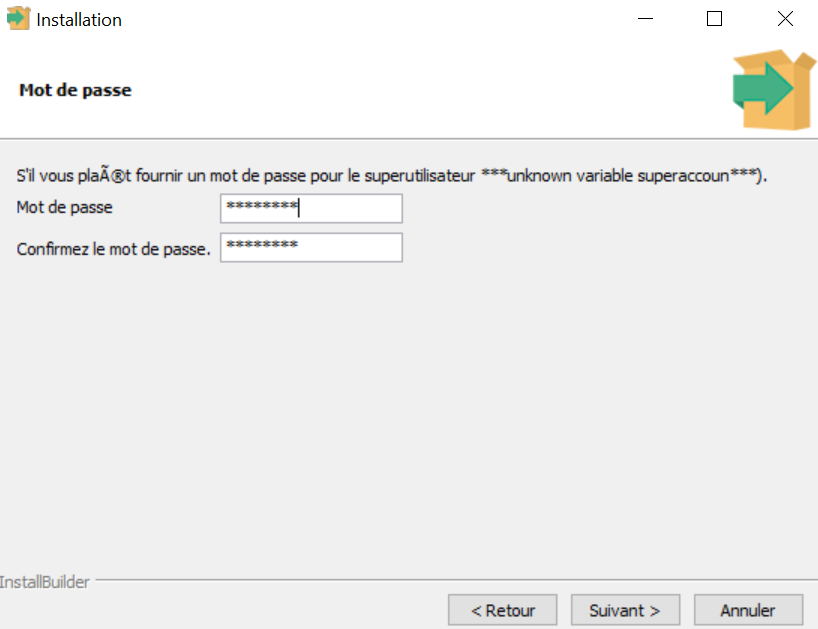
* Télécharger la version 12.x ou la version 11.x depuis l’adresse https://www.enterprisedb.com/downloads/postgres-postgresql-downloads (l’exemple utilise la version 12.2)



* Une fois télécharger, cliquez sur le fichier postgresql-12.2-2-windows-x64.exe et suivez le guide d’installation en gardant les valeurs par défaut et en cliquant sur « Suivant ». Il est impératif que PgAdmin et Postgis soient installés, vous en aurez besoin pour faire le TP.



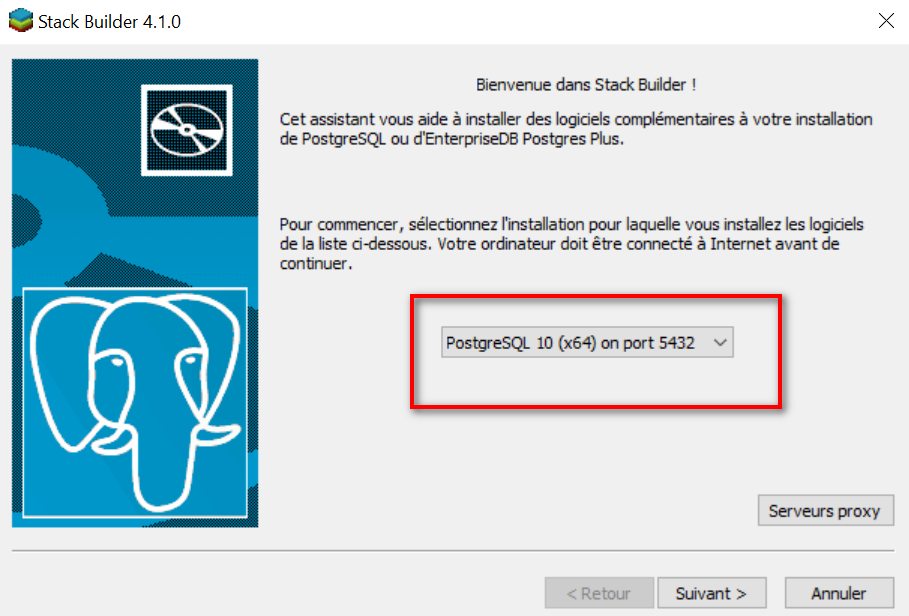
* Pour le login / mot de passe du super utilisateur mettre « **postgres » et « postres »**



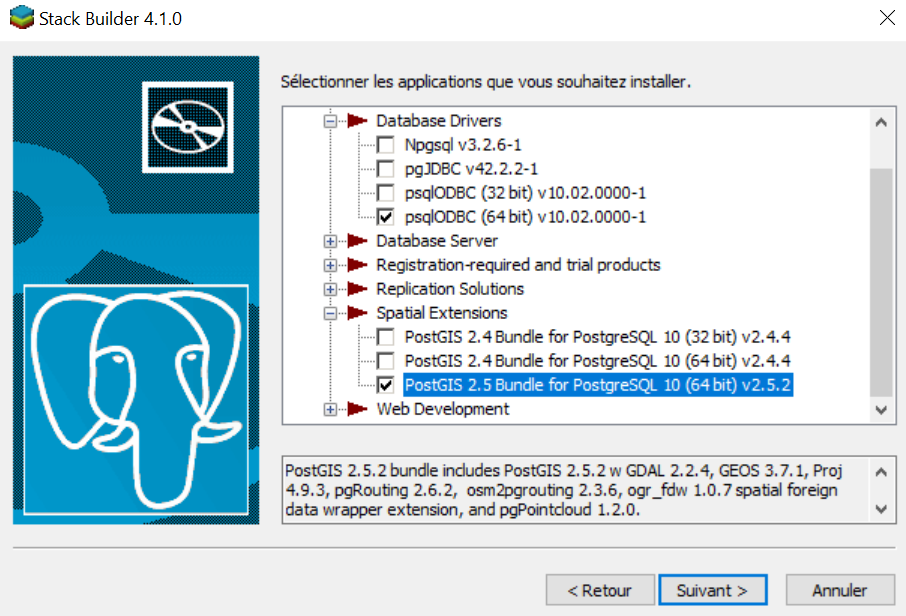
* L’exécution du script d’installation peut prendre plusieurs minutes (c’est le bon moment pour prendre un café),
* A la fin du script d’installation, cliquez sur « Terminer » en laissant coché la case concernant l’exécution de Stack Builder



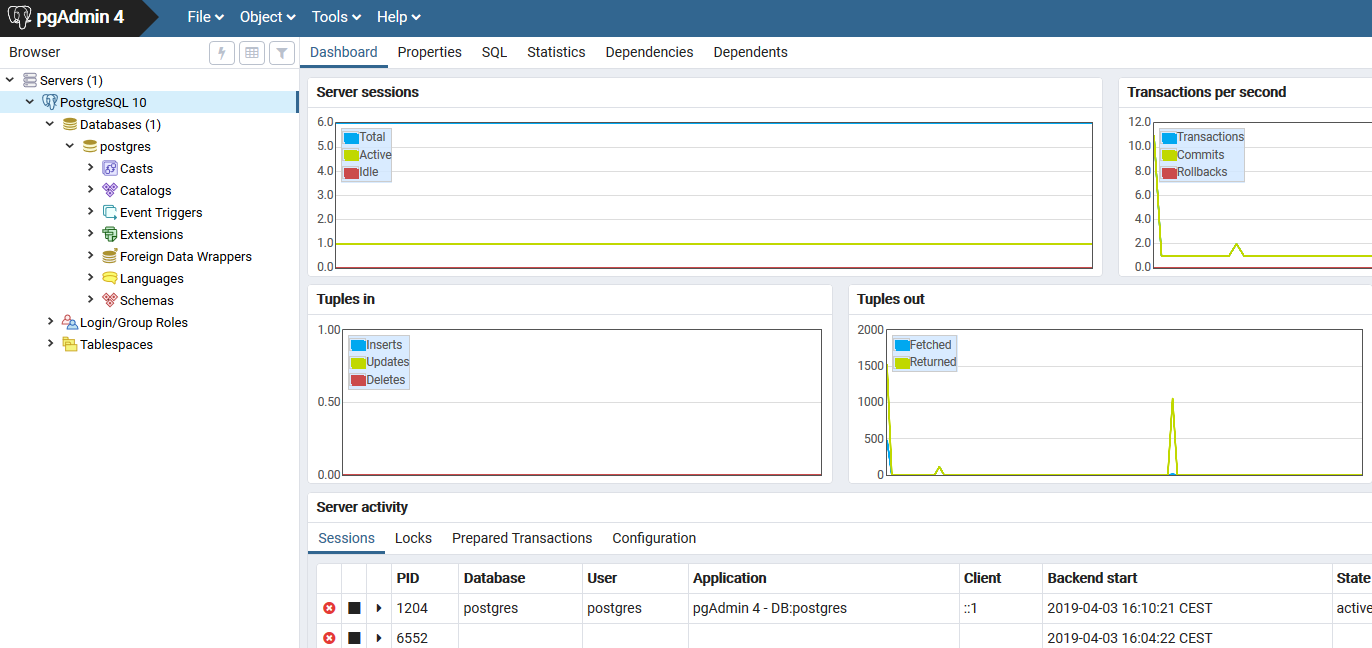
* Pour démarrer l’exécution de « Stack Builder » sélectionner, l’instance Postgres que vous venez d’installer sur votre machine : dans l’exemple : « PostgreSQL 12 (x64) on port 5432



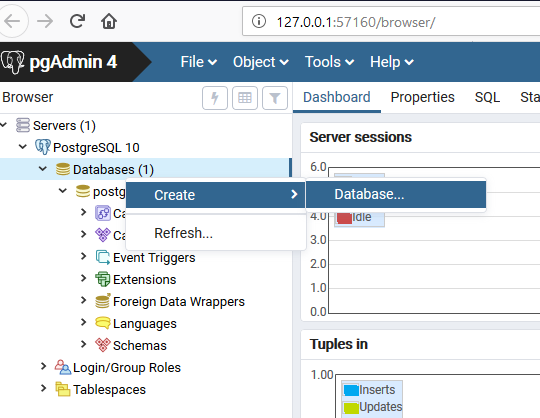
* Sélectionner les plugins qui vous seront utiles par exemple Postgis (2.4/2.5) pour les traitements géographiques et psqlODBC (optionnel) pour établir une connexion entre la base Postgres et une base Access, puis suivez le guide d’installation en cliquant sur « Suivant »,



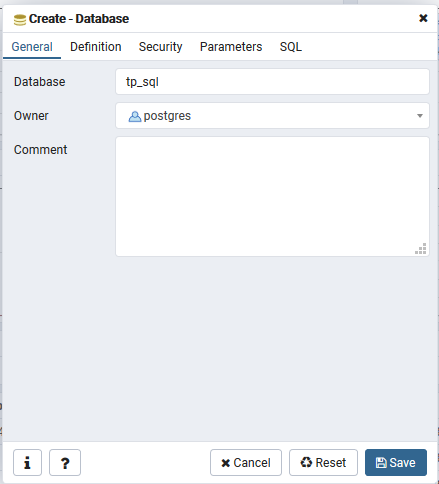
* Vérifiez l’installation de Postgres en démarrant l’application « PgAdmin » disponible dans le menu de Windows. Pour la connexion au serveur, utilisez le compte du « super utilisateur » défini lors de la procédure d’installation (par défaut : postgres/postgres).



* Avec PgAdmin, créez la base de données qui sera utilisée pour le TP SQL, pour cela faites un click droit sur l’item « Database » de votre serveur, puis « Create » et enfin « Database ».



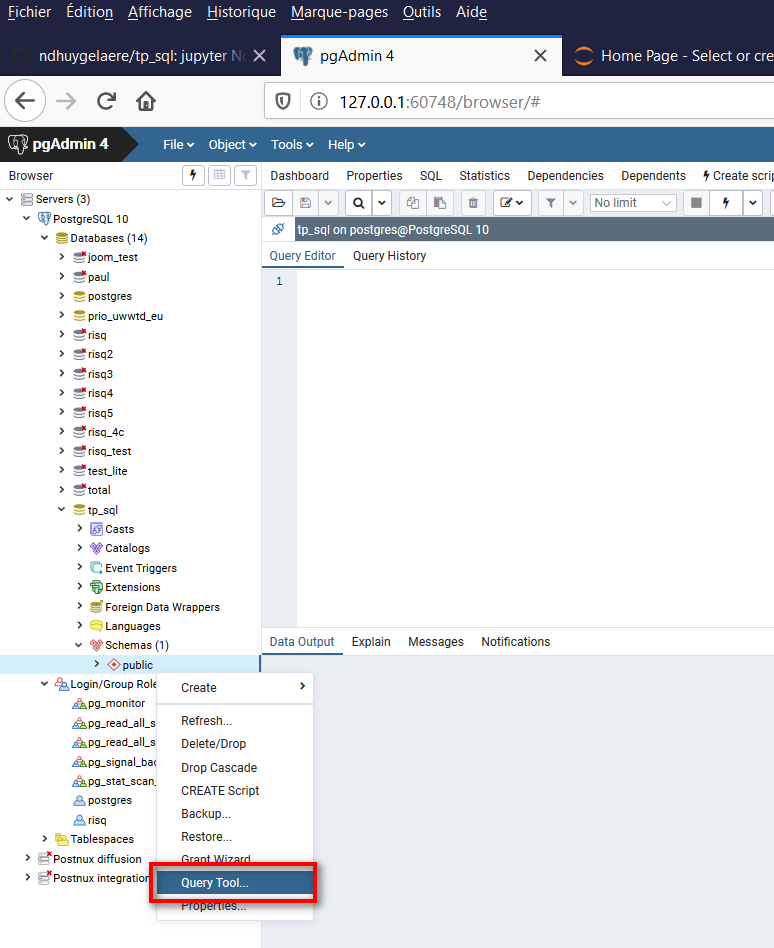
Dans le formulaire, entrez le nom de la base de données ex : « **tp\_sql** » et définissez l’utilisateur « **postgres** » précédemment créé comme propriétaire (champ « Owner »). Dans l’onglet « Definition » vérifiez que le champ « Encoding » soit bien défini avec la valeur « UTF8 » (à défaut changer la valeur)



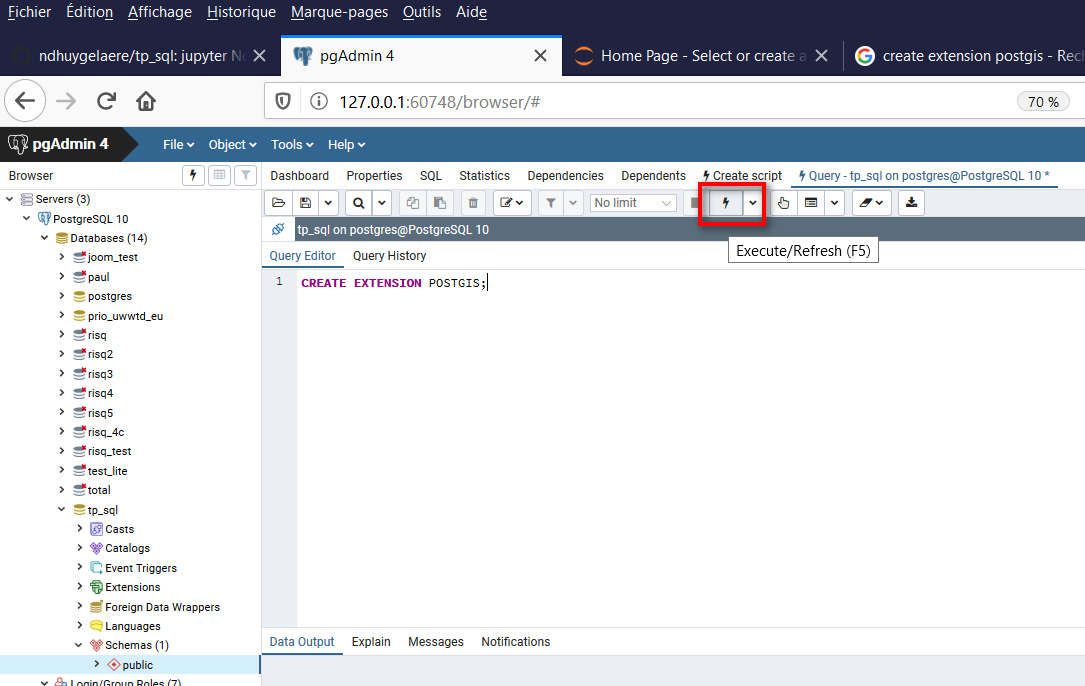
Après avoir créé votre base de données, celle-ci doit désormais apparaitre dans la liste des bases de votre serveur postgres.

Nous allons maintenant, activer l’extension **postgis** pour votre base de données. Pour cela, naviguer dans la liste des bases, puis sélectionnez la base « tp\_sql », puis le schéma « public » dans la liste des schémas. Dans postgres, le schéma « public » est le schéma par défaut. Vous avez la possibilité de créer autant de schéma que vous le souhaitez, ce qui est pratique pour conserver par exemple les données sources dans un stockage différent de celui de l’application de gestion ou des résultats issus de calculs.

Enfin fait un clic droit sur le schéma « **public** » et sélectionnez l’item **« Query tool …»**

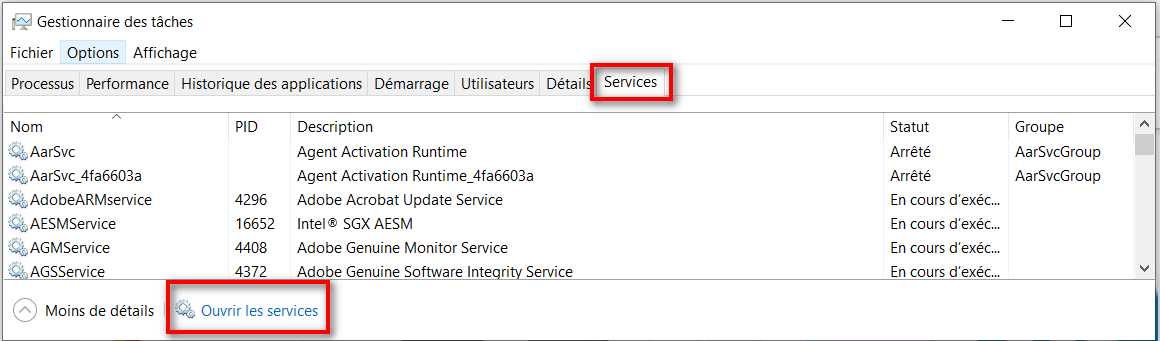


Dans l’onlet « Query editor », entrez la commande « **CREATE EXTENSION POSTGIS ;** » et exécuter la requête en cliquant sur « l’éclair » dans la barre d’outil :

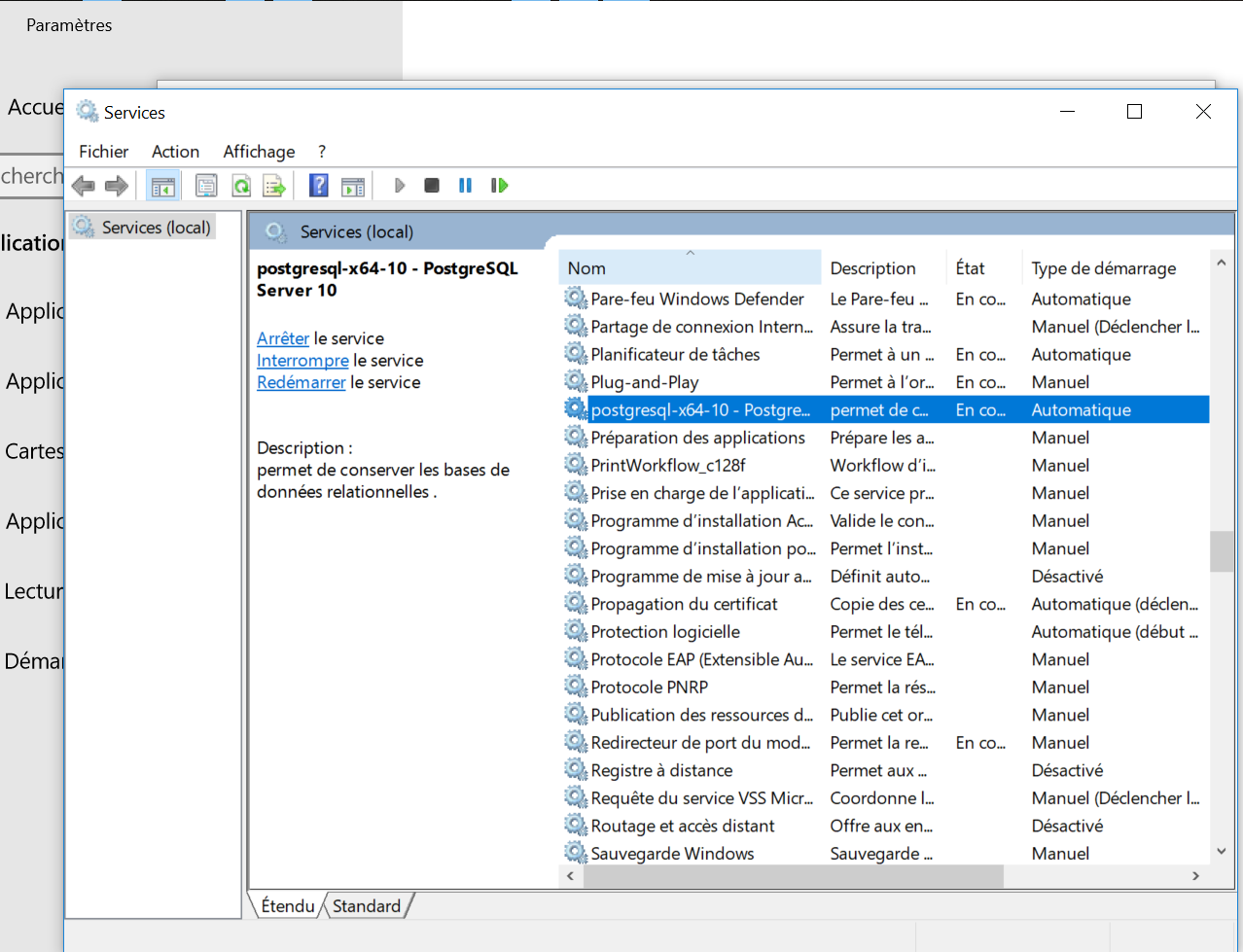


Votre base de donnée est prête !

Enfin, à l‘aide du « gestionnaire des tâches » Windows, aller dans l’onglet «Service», puis en bas de page cliqez sur le lien « Ouvrir les services »,



Dans la fenêtre qui s’ouvre chercher le service et vérifier que le service « Postgresql …» soit en type de démarrage « automatique ». Ainsi, lors du redémarrage de la machine le service tournera en tâche de fond. Sinon il faudra démarrer le service « postgresql » à chaque redémarrage de votre machine.



# Import des données dans postgres

Pour le TP, nous utiliserons 3 couches de données géographiques diffusées sur le site du Sandre. Je vous invite à lire les fiches de métadonnées et à télécharger la version shape de chaque couche d’information. Notez que les 3 couches sont en Lambert 93 (EPSG :2154) et qu’il s’agit donc d’une projection métrique.

|  |  |  |
| --- | --- | --- |
| **Nom de la couche** | **Description** | **URL** |
| Obstacles à l'écoulement – Métropole  (POINT) | Un obstacle à l’écoulement est un ouvrage lié à l’eau qui est à l’origine d’une modification de l’écoulement des eaux de surface (dans les talwegs, lits mineurs et majeurs de cours d'eau et zones de submersion marine). Seuls les obstacles artificiels (provenant de l’activité humaine) sont pris en compte. | <http://www.sandre.eaufrance.fr/atlas/srv/fre/catalog.search;jsessionid=node01w45y76i7ityrt8w55m5ahtre453803.node0#/metadata/070df464-73d3-4c00-be2f-93f2a97ef8f5> |
| Cours d'eau - Métropole 2017 - BD Carthage  (LINE) | Le concept principal du découpage hydrographique est l’entité hydrographique définie par la circulaire n°91-50.  L’entité hydrographique est un cours d’eau naturel ou aménagé, un bras naturel ou aménagé, une voie d’eau artificielle (canal,…), un plan d’eau ou une ligne littorale. | <http://www.sandre.eaufrance.fr/atlas/srv/fre/catalog.search;jsessionid=node01w45y76i7ityrt8w55m5ahtre453803.node0#/metadata/7381de46-42f7-42df-9abe-0ecd4b946034> |
| Sous-secteurs hydrographiques  - Métropole 2017 - BD Carthage  (POLYGON) | Le découpage hydrographique est l'ensemble des quatre partitions hiérarchisées du territoire français réalisé selon des aires hydrographiques décroissantes : - région hydrographique (1er ordre), - secteur hydrographique (2ème ordre), - sous-secteur hydrographique (3ème ordre), - zone hydrographique (4ème ordre). | <http://www.sandre.eaufrance.fr/atlas/srv/fre/catalog.search;jsessionid=node01w45y76i7ityrt8w55m5ahtre453803.node0#/metadata/7895748c-9991-4647-b03b-812347dde485> |

Une fois les couches géographiques téléchargées, dans le répertoire par défaut de votre utilisateur windows, créer le répertoire « **tp\_sql** ». Le chemin par défaut se trouve dans :

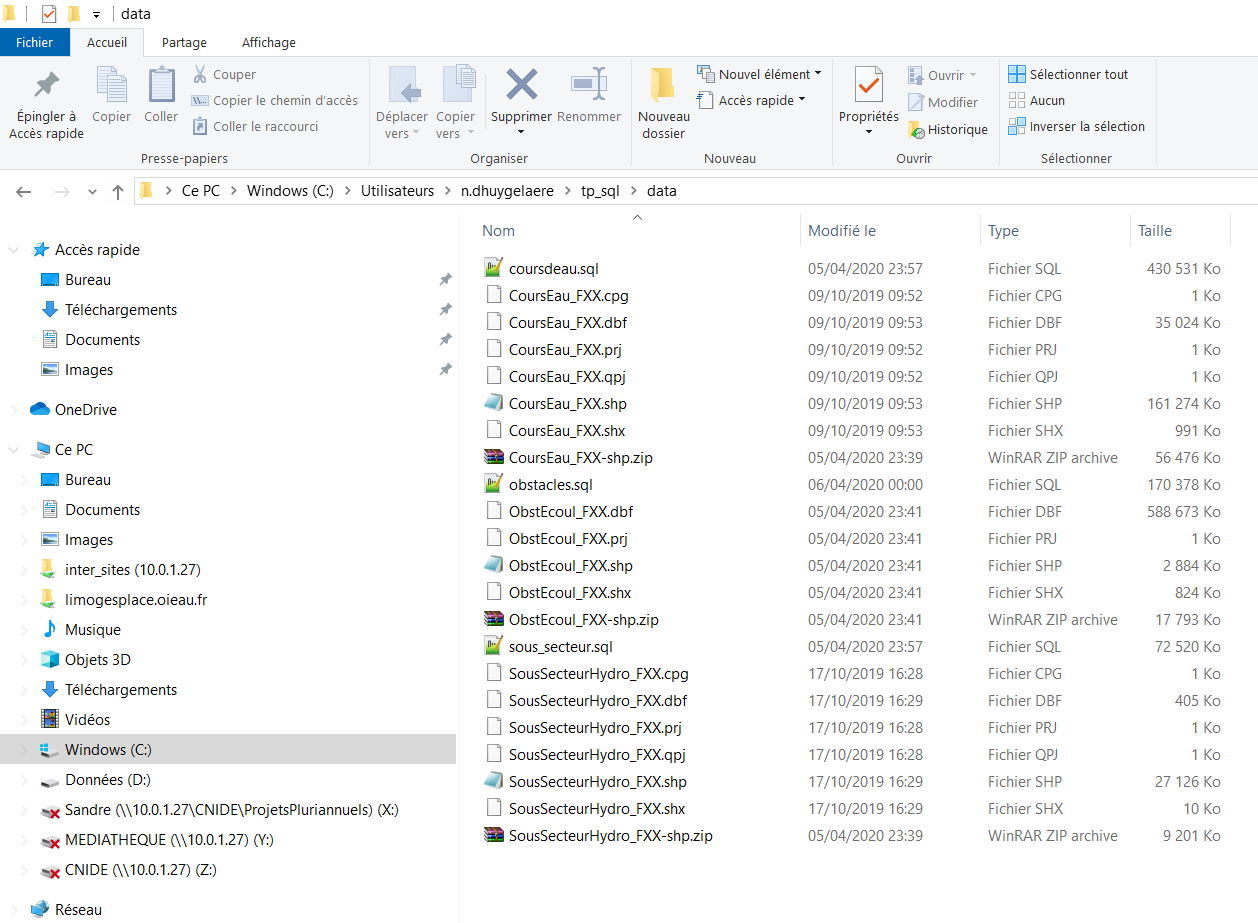
C:\Users\[*mon compte utilisateur*] \

Ou dans son alias :

C:\Utilisateurs\[*mon compte utilisateur*]

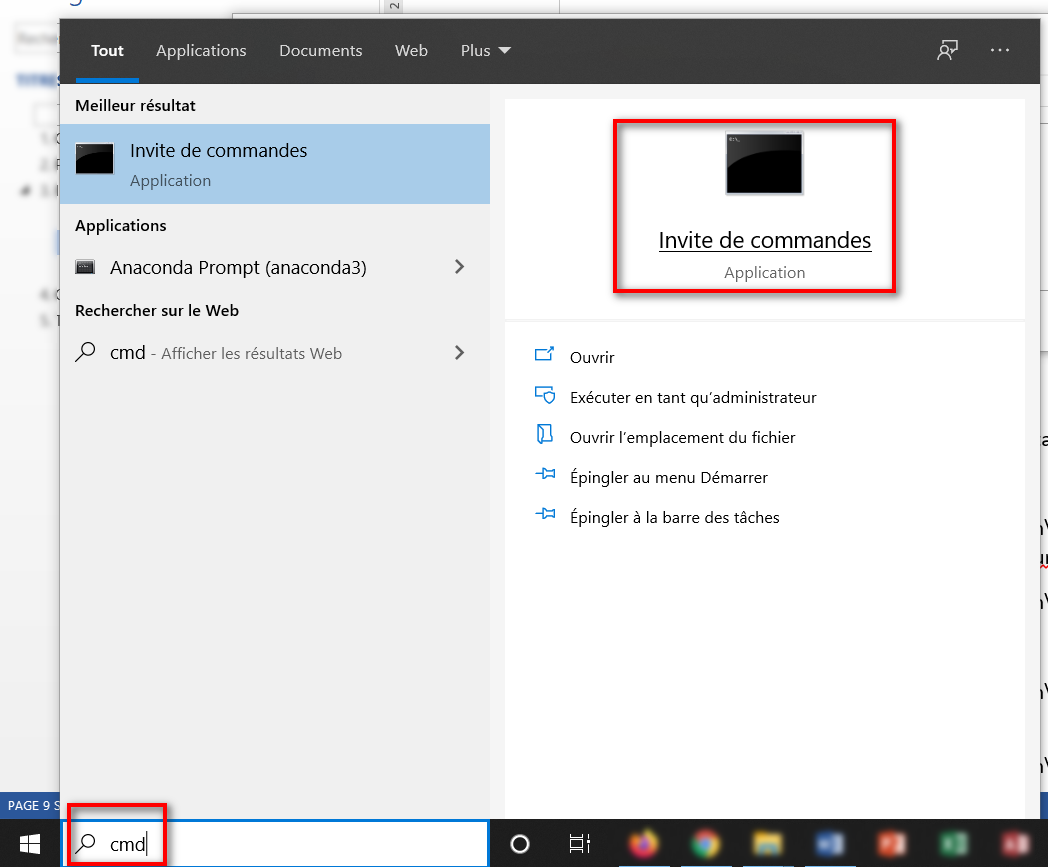
Avec *mon compte utilisateur =* ***n.dhuygelaere***

Puis créer le répertoire « **data** » dans le répertoire « **tp\_sql** » et décompresser toutes les archives ZIP contenant les couches géographiques. A la fin de cette opération, vous obtiendrez cet écran :

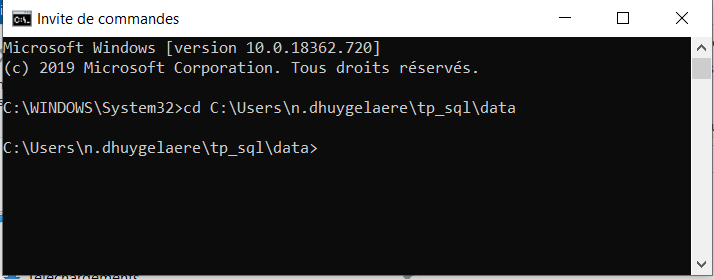


**A partir de maintenant, vous commercer votre vie de géomaticien développeur !**

Maintenant, nous allons charger les données géographiques contenues dans les fichiers shape dans la base postgres (sans passer par QGIS) en utilisant la commande shp2pgsql. Pour cela, démarrer votre console Windows également appelée « l’invite de commande ». Dans votre barre de menu Windows entrez « cmd », puis cliquez sur le programme « Invite de commandes »



Puis entrez les commandes suivantes dans la console :

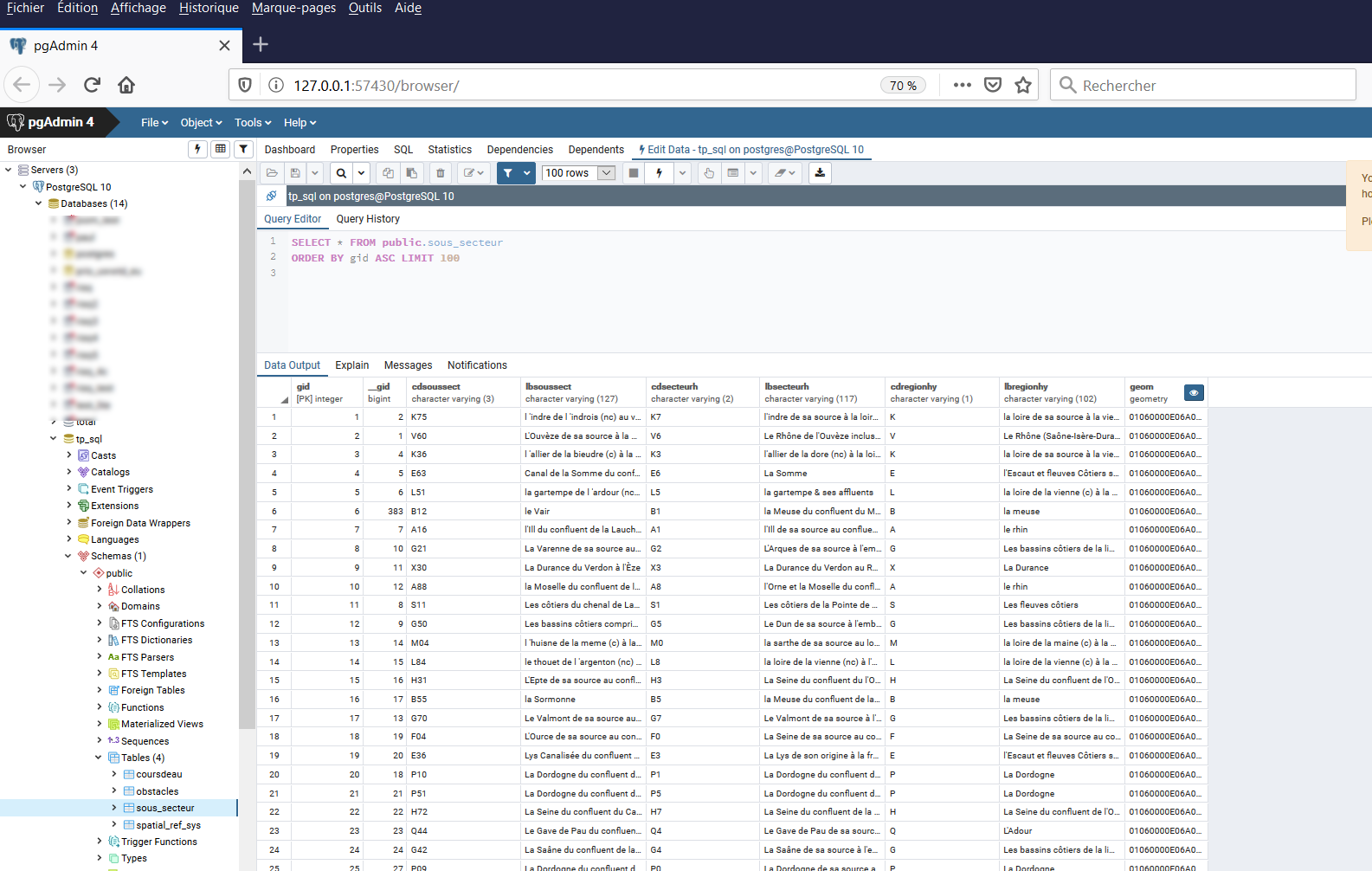


Remarque :

* Il est possible de copier/coller les commandes dans la console en faisant un CRTL+C pour copier le texte et de coller dans la console en positionnant le curseur sur la dernière ligne de console et en faisant un clic droit avec la souris.
* Pensez à changer le nom du compte utilisateur (en rouge) et la version de postgres (en orange) afin de coller à votre installation

|  |
| --- |
| cd C:\Users\**n.dhuygelaere**\tp\_sql\data  C:\"Program Files"\PostgreSQL\**12**\bin\shp2pgsql.exe -s 2154 -I -d SousSecteurHydro\_FXX.shp "public"."sous\_secteur" > sous\_secteur.sql  C:\"Program Files"\PostgreSQL\**12**\bin\psql.exe -U postgres -d tp\_sql < sous\_secteur.sql  C:\"Program Files"\PostgreSQL\**12**\bin\shp2pgsql.exe -s 2154 -I -d CoursEau\_FXX.shp "public"."coursdeau" > coursdeau.sql  C:\"Program Files"\PostgreSQL\**12**\bin\psql.exe -U postgres -d tp\_sql < coursdeau.sql  C:\"Program Files"\PostgreSQL\**12**\bin\shp2pgsql.exe -s 2154 -I -d ObstEcoul\_FXX.shp "public"."obstacles" > obstacles.sql  C:\"Program Files"\PostgreSQL\**12**\bin\psql.exe -U postgres -d tp\_sql < obstacles.sql |

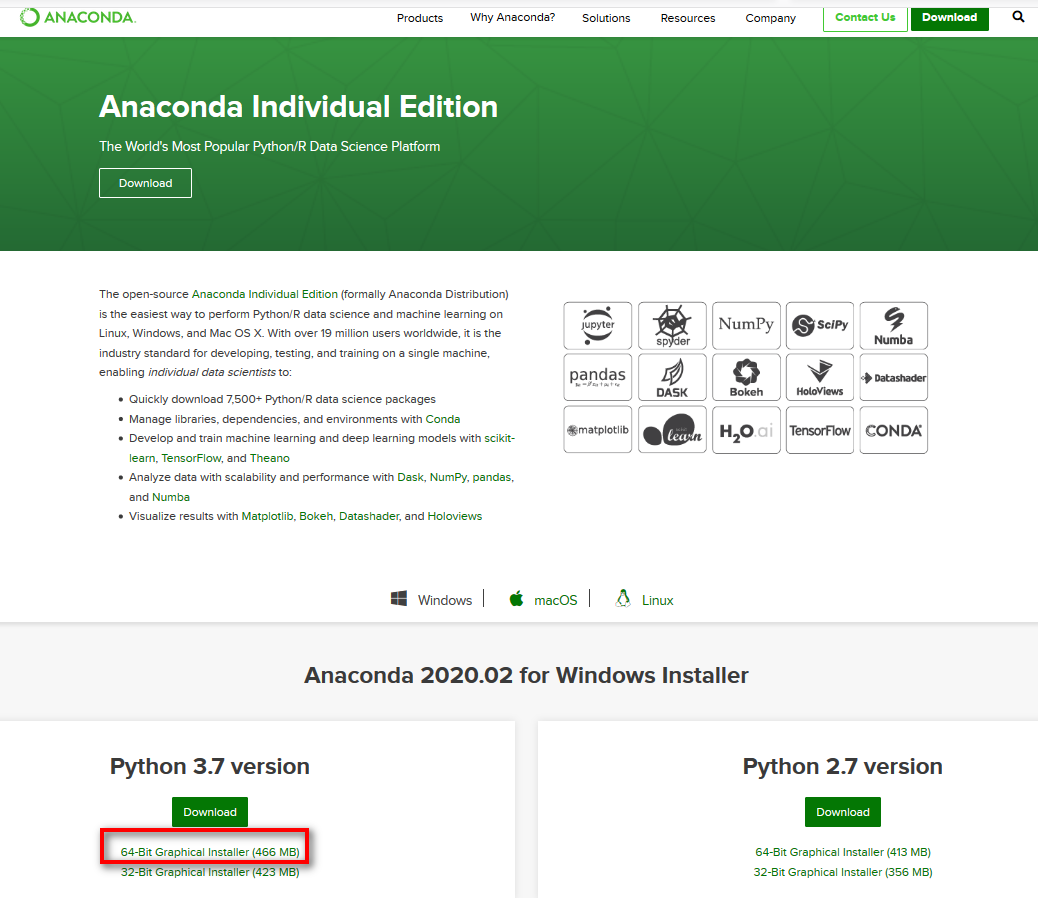
Vous pouvez désormais consulter vos données avec PgAdmin.



# Anaconda

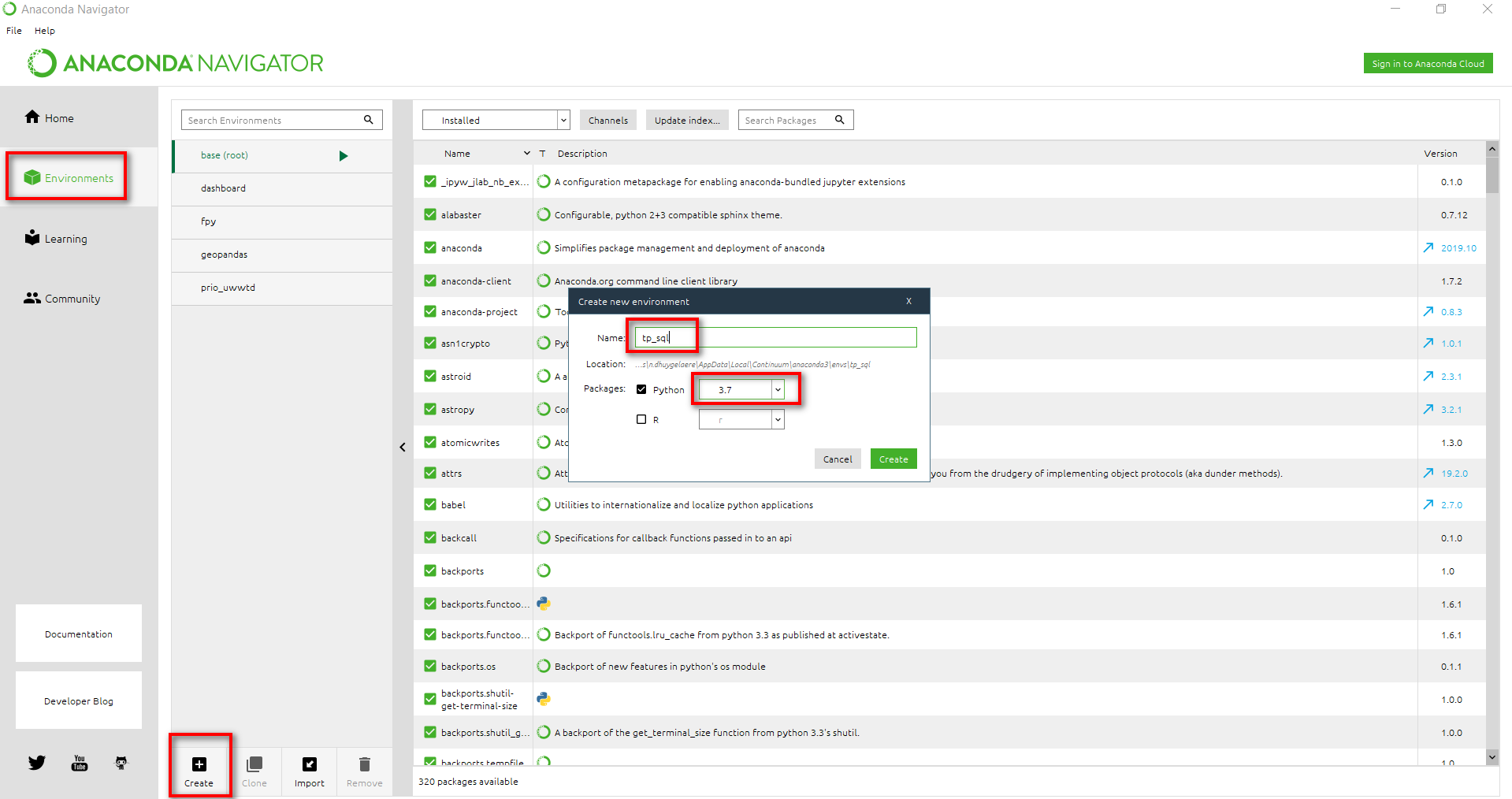
**Anaconda** est une distribution libre et open source des langages de programmation Python et R appliqué au développement d'applications dédiées à la science des données et à l'apprentissage automatique, qui vise à simplifier la gestion des paquets et de déploiement. Vous trouverez, le fichier d’installation à cette adresse : <https://www.anaconda.com/distribution/>

Choisissez la version **python 3.7** :

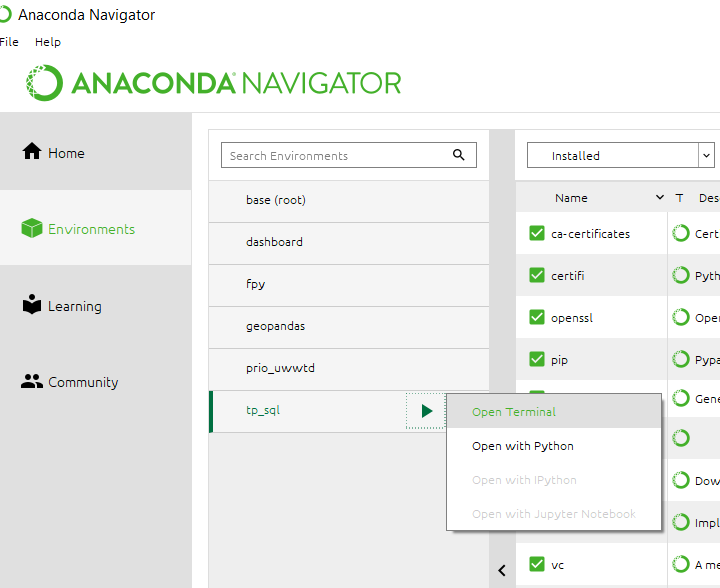


Après avoir téléchargé le fichier, lancer l’exécution de l’installation, celle-ci peut prendre pas mal de temps, donc si besoin, un deuxième café s’impose…

Une fois l’installation terminée, démarrez le « Anaconda navigator » depuis votre menu Windows afin de créer votre « environnement de développement ». Pour cela, comme sur l’image ci-dessous, allez sur l’onglet « **environnement** » sur le côté gauche de l’écran, plus cliquez sur le bouton « **create** » en bas de page. Enfin, entrez le nom de votre environnement ex : « **tp\_sql** » et activez le package **« python » « 3.7 ».**



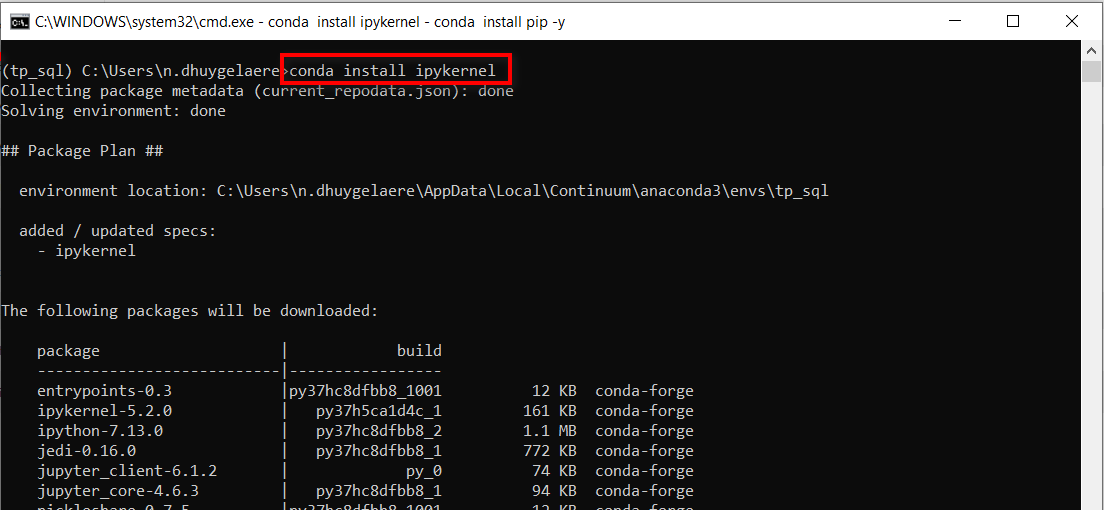
Après avoir créé votre environnement, cliquez sur l’item correspondant et démarrer la console python en cliquant sur l’item « **Open terminal** ».



Puis, copier coller dans la console les lignes ci-dessous afin d’installer les librairies python nécessaires au déroulement du TP :

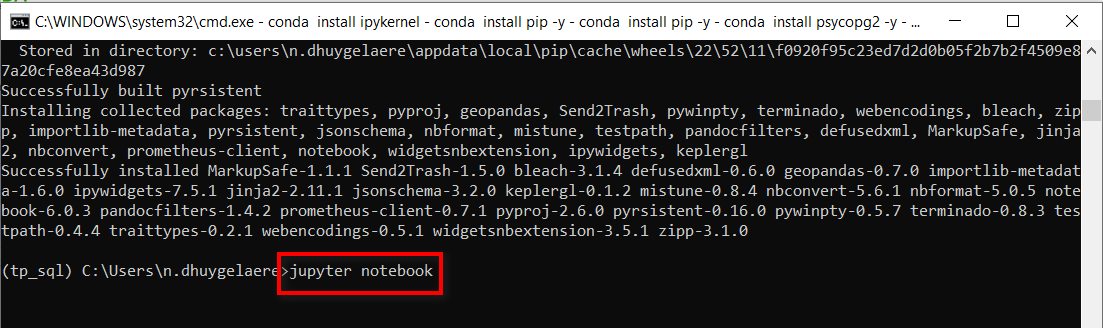
|  |
| --- |
| conda install ipykernel -y  python -m ipykernel install --user --name=tp\_sql --display-name='Python(tp\_sql)'  conda config --add channels conda-forge  conda install pip -y  conda install psycopg2 -y  conda install matplotlib -y  conda install numpy -y  conda install pandas -y  conda install fiona -y  conda install folium -y |

N’oubliez pas de taper sur « Entrer » à la fin de chaque ligne.

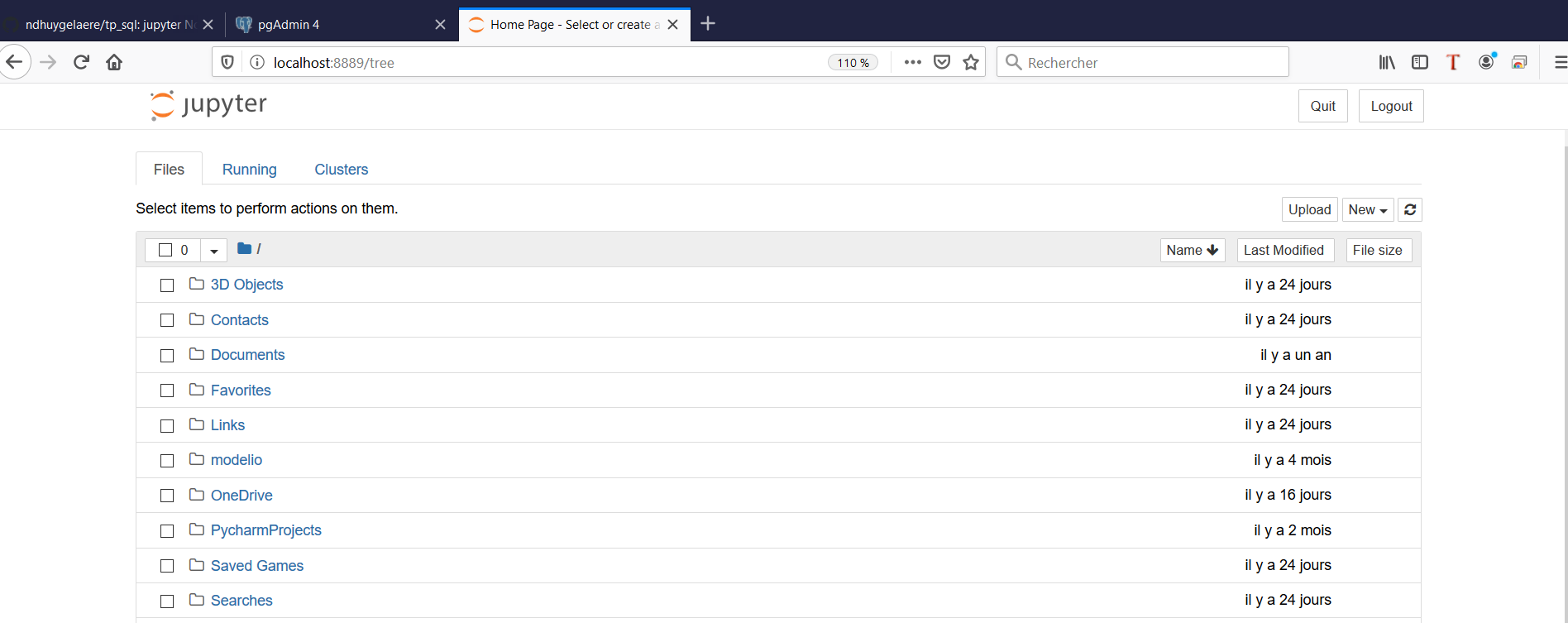


Enfin, afin de démarrer « **Jupyter** » entrez dans la console la commande suivante

|  |
| --- |
| jupyter notebook |



Si tout s’est bien passé, la fenêtre suivante devrait s’ouvrir dans votre navigateur web :

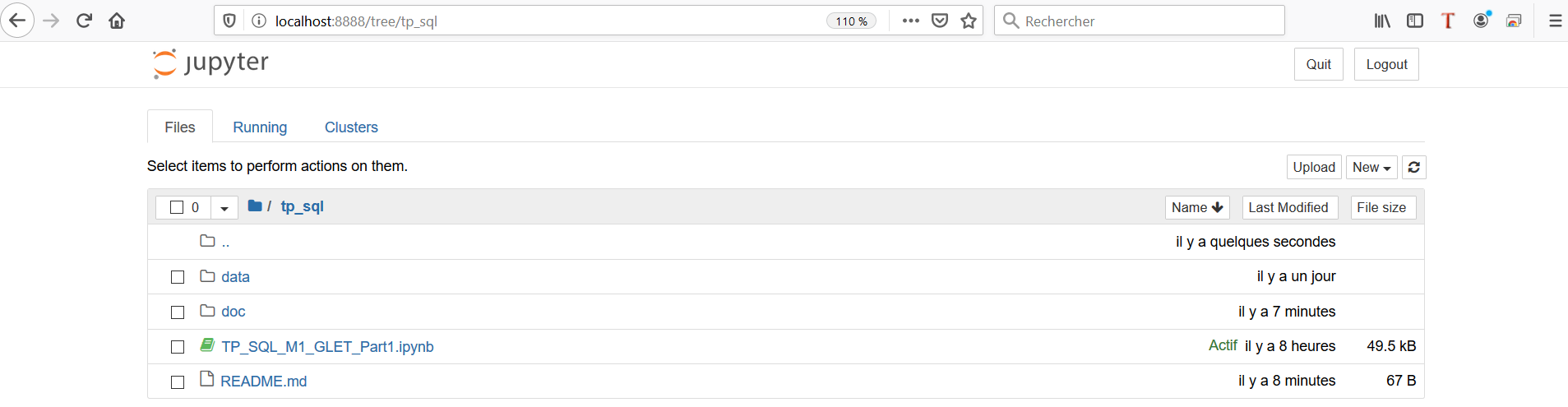


Nous y sommes presque, il ne reste plus qu’à récupérer le Notebook pour faire le TP.

# Installation du Notebook

Depuis l’adresse : <https://github.com/ndhuygelaere/tp_sql>, téléchargez les fichiers TP\_SQL\_M1\_GLET\_Part1.ipynb et TP\_SQL\_M1\_GLET\_Part2.ipynb et placez-les dans le répertoire C:\Users\n.dhuygelaere\tp\_sql

Depuis l’onglet Jupiter de votre navigateur web, rafraichissez votre page et aller dans le répertoire : tp\_sql. Ainsi vous devriez avoir une interface similaire à celle-ci :



Pour ouvrir, les notebooks cliquez simplement sur son nom.

Voilà, vous y êtes !