Formation PostGIS

SLQ et clients PgSQL / PostGIS - Octobre 2020

Fabien Guerreiro

Supports sous Licence Ouverte Etalab

Crédits : détails des crédits dans les supports.

INSTITUT NATIONAL SUPÉRIEUR DES SCIENCES AGRONOMIQUES, DE L'ALIMENTATION ET DE L'ENVIRONNEMENT

MEMBRE DE agreenium l'institut agronomique, vétérinaire & forestier de France

UBFC UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

Général

Trois ensembles

- Langage de Définition de Données (LDD) : créer et supprimer des objets.
- Langage de Contrôle de Données (LCD) : gérer les droits sur les objets.
- Langage de Manipulation de Données (LMD) : recherche, insertion, mise à jour et suppression de données.

Syntaxe

SELECT (liste des attributs)
FROM (liste des tables)
WHERE (Conditions)
;

Opérateurs

Comparaison

A = B
A <> B (différent)
A < B
A > B
A <= B (inférieur ou égal)
A >= B (supérieur ou égal)
A BETWEEN B AND C (compris entre B et C)
A IN (B1, B2,...) : liste de valeurs

logiques

OR : pour séparer deux conditions dont au moins une doit être vérifiée.
AND : pour séparer deux conditions qui doivent être vérifiées simultanément.
NOT : permet d'inverser une condition.

Types de données

Types

- **CHARACTER** (ou **CHAR**) : caractère de texte ou chaîne de caractères, de longueur fixe ;
- **CHARACTER VARYING** (ou **VARCHAR**) : chaîne de caractères de longueur variable avec longueur maximale fixée;
- **TEXT** : chaîne de caractères sans limite de taille ;
- **NUMERIC** (ou DECIMAL ou DEC) : nombre décimal exact de précision arbitraire ;
- **INTEGER** (ou **INT**) : nombre entier ;
- **REAL** : nombre réel à virgule flottante (de précision limitée, donc inexact) ;
- **BOOLEAN** (ou LOGICAL) : valeur booléenne (vrai ou faux) ;
- **DATE** : date du calendrier grégorien.

Fonctions de transcriptage

- **cast** (expr as type) : convertir un type en un autre (PgSQL : **expr::type**).
- **LIMIT** indiquer le maximum d'enregistrements en retour.
- **OFFSET** décaler le nombre de lignes à obtenir

Fonctions classiques

Fonctions de chaînes de caractères

- **LENGTH** : longueur d'une chaîne
- **CHR** : caractère correspondant au code ASCII (exemple CHR(184) renvoi ©)
- **SUBSTR** : extraction d'une sous-chaîne de caractères substr(chaîne, position , longueur) ; le premier caractère a la position 1 et non pas 0.
- **UPPER** : convertit en majuscule
- **LOWER** : convertit en minuscule

Fonctions mathématiques

- **POW** : pour élever à une puissance. ex : POW(champ, 2) pour élever au carré.
- **SQRT** : pour obtenir la racine carrée.
- **ROUND** : qui permet d'arrondir un résultat

Tri et agrégation

Tri

- **ORDER BY** : classer le résultat
- **DESC** : tri décroissant
- **SUBSTR** : extraction d'une sous-chaîne de caractères substr(chaîne, position , longueur) ; le premier caractère a la position 1 et non pas 0.
- **UPPER** : convertit en majuscule
- **LOWER** : convertit en minuscule

Agrégation

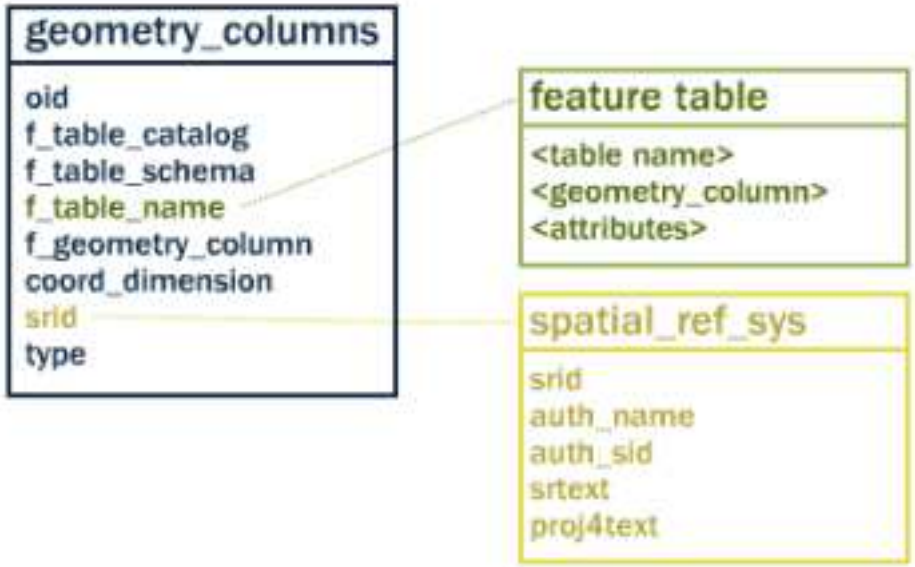
- **count()** : nombre d'enregistrements
- **sum()** : somme
- **max()** : maximum
- **min()** : minimum
- **avg()** : moyenne

- **GROUP BY** : groupement des autres colonnes
- **HAVING** : critère supplémentaire de sélection sur l'agrégation

Données géométriques et fonctions spatiales

Géométrie

- geometry, geom ou the_geom
- deux autres tables internes
  - geometry_columns
  - spatial_ref_sys



Agrégation

ST_SRID() : le code du système de projection de l'objet
ST_IsValid() : vérifie la géométrie des objets (pas d'erreur topologique)
ST_X() : la coordonnée X d'un point (et uniquement d'un point).
ST_Y() : coordonnée Y d'un point
ST_Centroid() : le centroide d'un polygone
ST_Area() : la surface d'un objet
ST_Buffer() : un nouvel objet tampon construit autour d'un objet
ST_Length() : la longueur ou le périmètre d'un objet

Exercice pratique

NOM_COMM	INSEE_COMM	STATUT	X_COMMUNE	Y_COMMUNE	SUPERFICIE	POPULATION	INSEE_CANT	INSEE_ARR	NOM_DEPT	INSEE_DEPT
SAINT-JEAN-DE-LA-MOTTE	72291	Commune simple	478935	6744018	3203	900	26	1	SARTHE	72
ARTHEZE	72009	Commune simple	466877	6748256	865	400	17	1	SARTHE	72
VAULANDRY	49380	Commune simple	472055	6726373	2765	300	04	3	MAINE-ET-LOIRE	49
CLEFS	49101	Commune simple	470066	6730106	2592	900	04	3	MAINE-ET-LOIRE	49
MAREIL-SUR-LOIR	72185	Commune simple	475371	6739051	1183	600	14	1	SARTHE	72
BOUSSE	72044	Commune simple	470515	6745247	1202	400	17	1	SARTHE	72
LE BAILLEUL	72022	Commune simple	462145	6746131	2746	1200	17	1	SARTHE	72
CLERMONT-CREANS	72084	Commune simple	473148	6741278	1782	1200	14	1	SARTHE	72
MALICORNE-SUR-SARTHE	72179	Chef-lieu de canton	469673	6750652	1513	2000	17	1	SARTHE	72
THOREE-LES-PINS	72357	Commune simple	477876	6733984	2818	700	16	1	SARTHE	72
LA FONTAINE-SAINT-MARTIN	72135	Commune simple	479050	6747256	1372	600	26	1	SARTHE	72
LA FLECHE	72154	Sous-préfecture	470872	6737445	7421	15400	14	1	SARTHE	72
VILLAINES-SOUS-MALICORNE	72377	Commune simple	467557	6744178	1916	1000	17	1	SARTHE	72
CRE	72108	Commune simple	464444	6733839	1719	800	14	1	SARTHE	72
CROSMIERES	72110	Commune simple	463343	6741281	2045	900	14	1	SARTHE	72
SAINT-QUENTIN-LES-BEAUREPAIRE	49315	Commune simple	467128	6731077	751	300	04	3	MAINE-ET-LOIRE	49
BAZOUGES-SUR-LE-LOIR	72025	Commune simple	461769	6736584	2990	1200	14	1	SARTHE	72
COURCELLES-LA-FORET	72100	Commune simple	473803	6748536	1960	400	17	1	SARTHE	72
LIGRON	72163	Commune simple	474237	6745574	1348	500	17	1	SARTHE	72

- Sélectionner les communes du département de la Sarthe de plus de 1500 habitants en affichant un tableau avec les noms de communes et leur population.
- Sélectionner les communes de la table COMMUNE dont le statut n'est pas chef-lieu de canton et afficher les colonnes NOM_COMM en lui donnant comme alias NOM et les colonnes, STATUT, POPULATION et SUPERFICIE
- Calculer pour chaque département ; la population totale, la densité moyenne de population des communes = moyenne(population commune /superficie commune) arrondie à deux décimales, la population de la commune la plus peuplée et celle de la moins peuplée, la superficie moyenne des communes.
- Quels sont les surfaces (en km^2) et périmètres (en km), arrondis à deux chiffres après la virgule, des communes du département de la Sarthe ? On considérera que la géométrie s'appelle geometry, et que le SRID est le 2154 (RGF93/Lambert93), ayant pour unité le mètre.

PgAdmin : présentation

pgAdmin 4

pgAdmin Fichier Objet Outils Aide

Navigateur | Tableau de bord | Statistiques | Dépendances | Dépendants | Propriétés

- formation
  - Catalogues
  - Extensions
  - Langages
  - Schémas (1)
    - public
      - Analyseurs de recherche plein texte
      - Collationnements
      - Configurations de recherche plein texte
      - Dictionnaires de recherche plein texte
      - Domaines
      - Fonctions
      - Fonctions déclencheurs
      - Modèle de recherche plein texte
      - Procédures
      - Séquences (5)
      - Tables (7)
        - commune
        - etablissement
        - iris_extrait72
        - ponctuel_hydrographique
        - route_xy
        - spatial_ref_sys
        - troncon_hydrographique
      - Tables distantes
      - Types
      - Vues (2)
        - geography_columns
        - geometry_columns
      - Vues matérialisées
  - Transtypages
  - Triggers sur évènements
  - Wrapper de données distantes

SQL

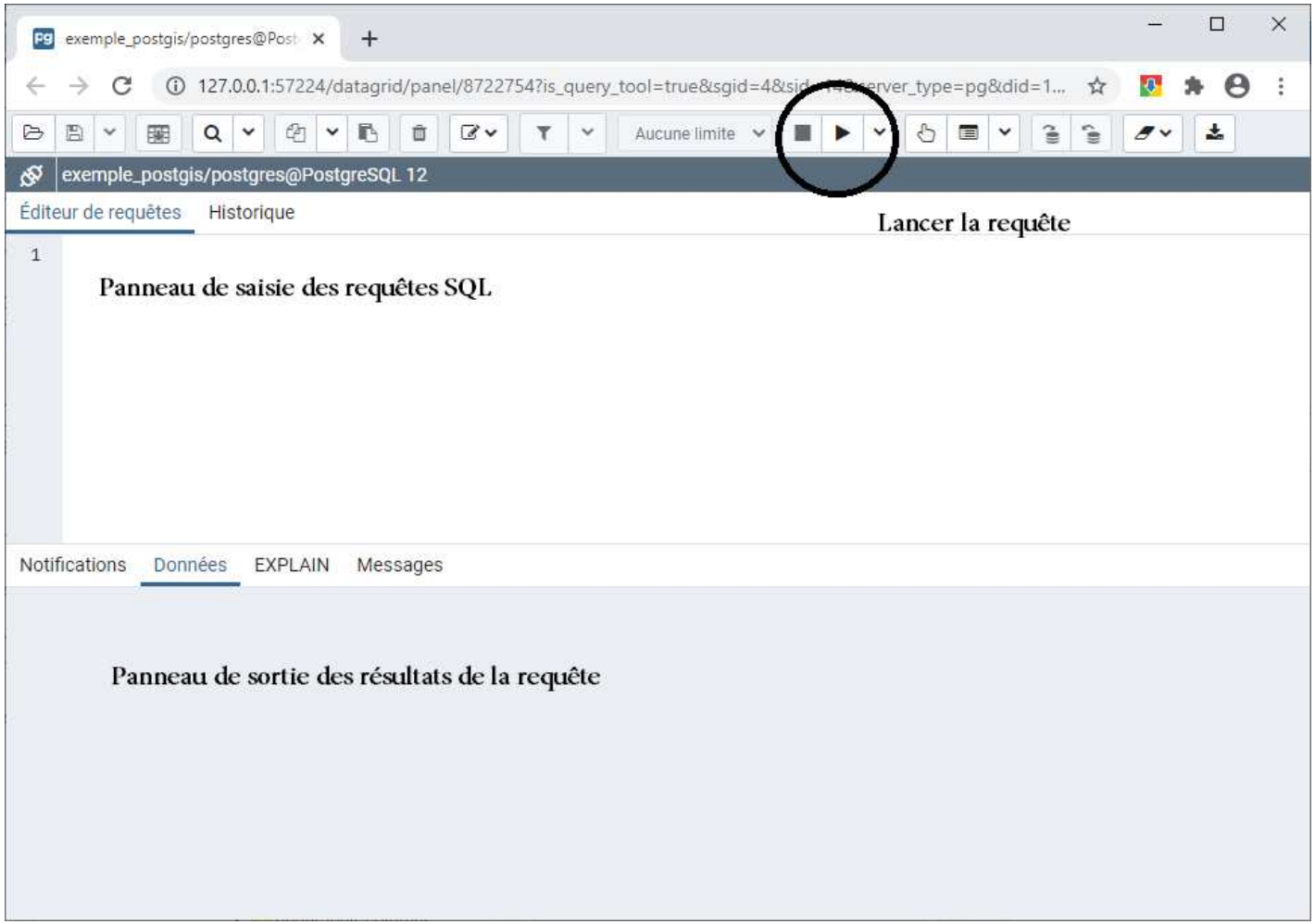
```sql
-- Table: public.commune

-- DROP TABLE public.commune;

CREATE TABLE public.commune
(
    id integer NOT NULL DEFAULT nextval('commune_id_seq'::regclass),
    geom geometry(MultiPolygon,2154),
    id_bdcarto bigint,
    nom_comm character varying(50) COLLATE pg_catalog."default",
    insee_comm character varying(5) COLLATE pg_catalog."default",
    statut character varying(20) COLLATE pg_catalog."default",
    x_commune integer,
    y_commune integer,
    superficie bigint,
    population integer,
    insee_cant character varying(2) COLLATE pg_catalog."default",
    insee_arr character varying(1) COLLATE pg_catalog."default",
    nom_dept character varying(30) COLLATE pg_catalog."default",
    insee_dept character varying(2) COLLATE pg_catalog."default",
    nom_region character varying(30) COLLATE pg_catalog."default",
    insee_reg character varying(2) COLLATE pg_catalog."default",
    CONSTRAINT commune_pkey PRIMARY KEY (id)
)

TABLESPACE pg_default;

ALTER TABLE public.commune
    OWNER to postgres;
-- Index: sidx_commune_geom

-- DROP INDEX public.sidx_commune_geom;

CREATE INDEX sidx_commune_geom
    ON public.commune USING gist
    (geom)
```

PgAdmin : base de données PostGIS

- postgres
  - Catalogues (2)
    - ANSI (information_schema)
    - PostgreSQL Catalog (pg_catalog)
  - Extensions (2)
    - adminpack
    - plpgsql
  - Langages (1)
    - plpgsql
  - Schémas (1)
    - public
      - Analyseurs de recherche plein texte
      - Collationnements
      - Configurations de recherche plein texte
      - Dictionnaires de recherche plein texte
      - Domaines
      - Fonctions
      - Fonctions déclencheurs
      - Modèle de recherche plein texte
      - Procédures
      - Séquences
      - Tables
      - Tables distantes
      - Types
      - Vues
      - Vues matérialisées
  - Transtypages
  - Triggers sur évènements
  - Wrapper de données distantes

- Schémas (1)
  - public
    - Analyseurs de recherche plein texte
    - Collationnements
    - Configurations de recherche plein texte
    - Dictionnaires de recherche plein texte
    - Domaines
    - Fonctions
    - Fonctions déclencheurs
    - Modèle de recherche plein texte
    - Procédures
    - Séquences
    - Tables
    - Tables distantes
    - Types
    - Vues
    - Vues matérialisées

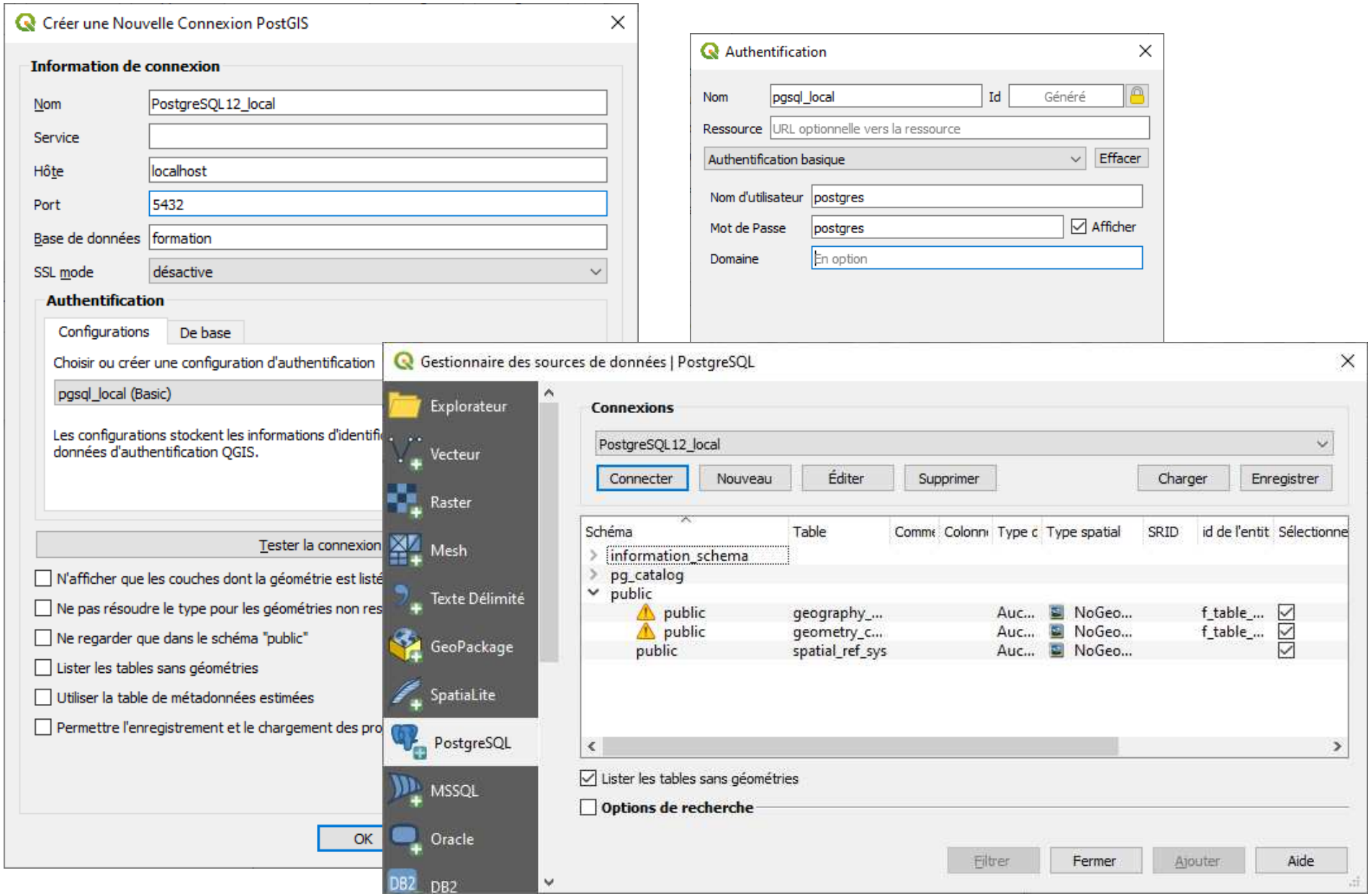
PgAdmin : Exercice pratique



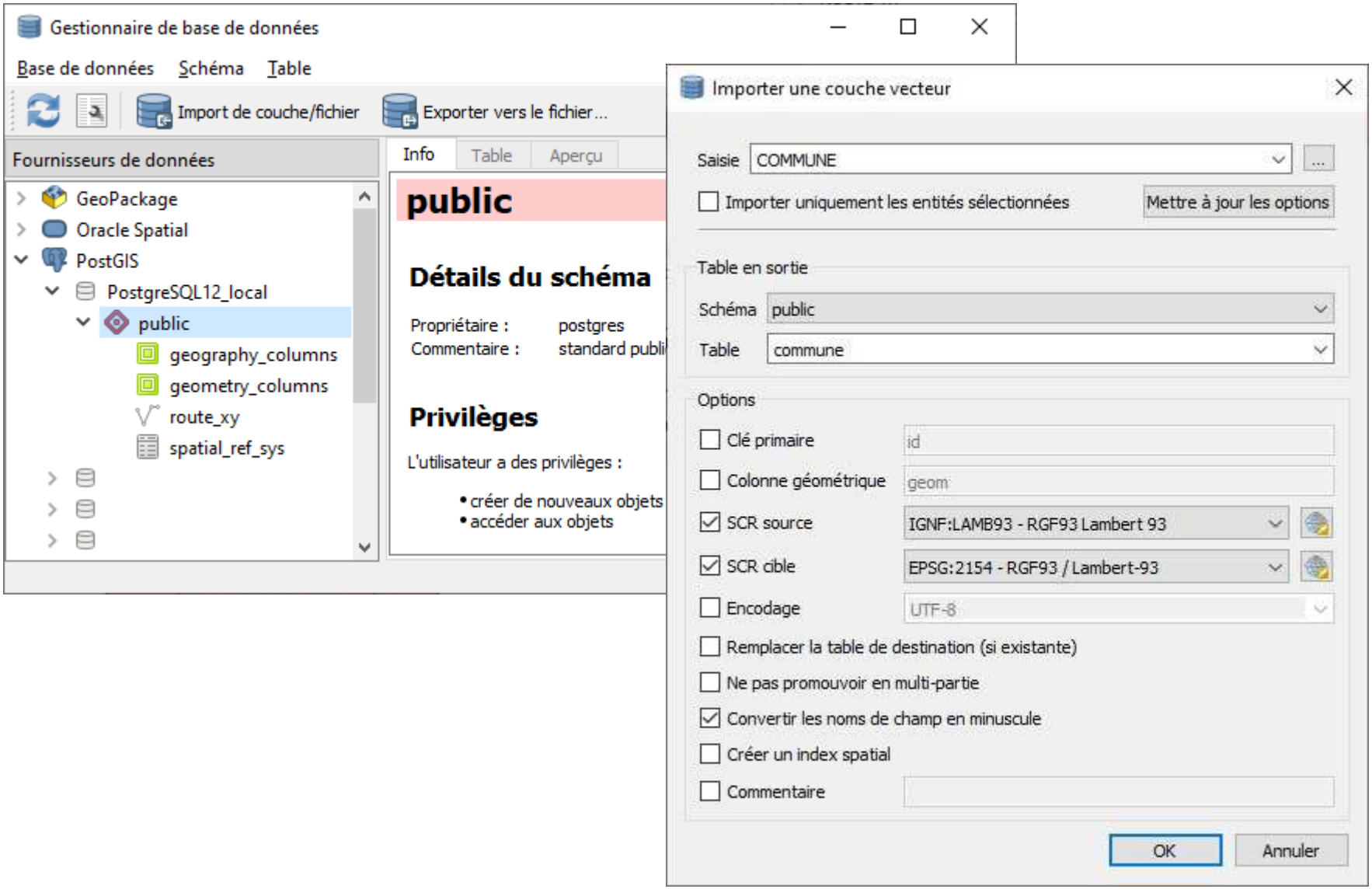
Créer une base de donnée formation avec l'extension PostGIS

- CREATE EXTENSION "postgis";
- SELECT version();
- SELECT PostGIS_full_version();

QGIS : connecter une base PostGIS / ajouter des données



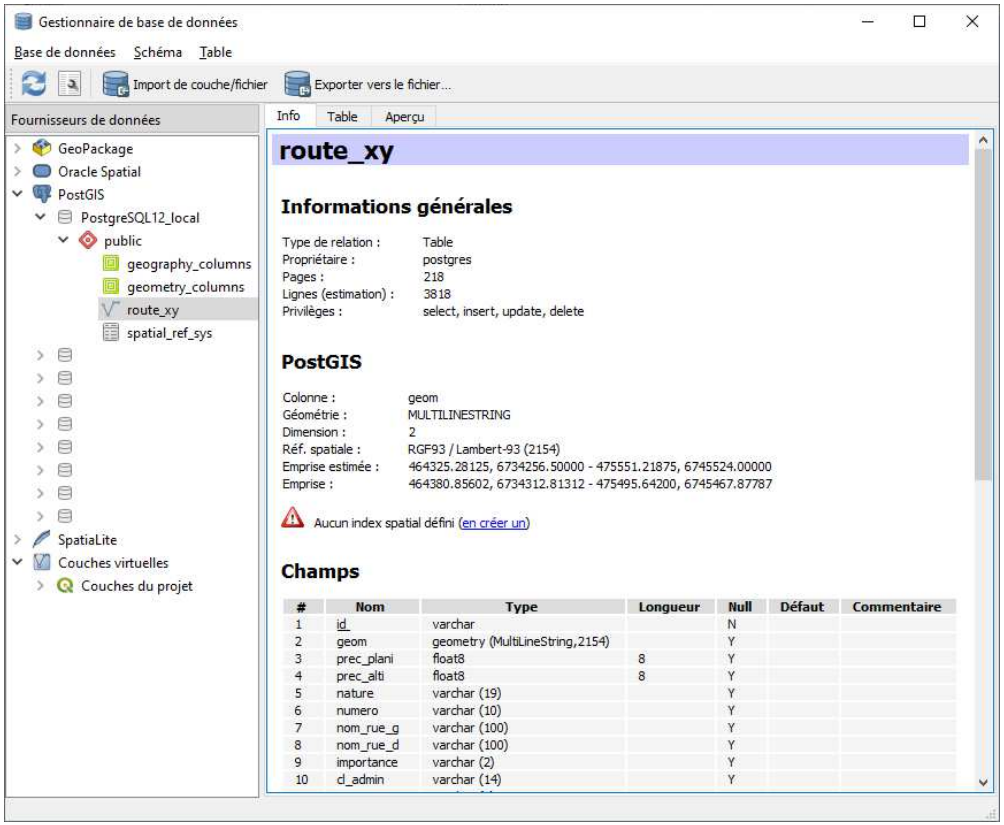
QGIS : importer des données dans PostGIS



DBmanager

QGIS : DB Manager

- Importer ROUTE_XY (répertoire Divers) dans PostGIS (extension DBManager)
- Contours_Iris\carto\IRIS_extrait72
- BD_CARTO\ADMINISTRATIF\COMMUNE
- BD_CARTO\HYDROGRAPHIE\PONCTUEL_HYDROGRAPHIQUE
- BD_CARTO\HYDROGRAPHIE\TRONCON_HYDROGRAPHIQUE
- BD_CARTO\TOPONYMIE\ETABLISSEMENT



- Attention au codage
- Attention à la casse

Indexation spatiale

GiST (Generalized Search Tree)

- L'indexation accélère les recherches
- Les index ajoutent aussi une surcharge au système

```
CREATE INDEX table_the_geom_gist
ON table USING GIST (the_geom)
```

L'optimiseur de requêtes sous PostGIS

- Un ordre SQL de type EXPLAIN ANALYZE suivi de la requête permet de récupérer la stratégie retenue par l'optimiseur.

```
EXPLAIN ANALYZE SELECT * from table
```

Exercice pratique

1. Sélectionner tous les IRIS (table IRIS_extrait72) de la commune de la FLECHE (colonne Nom_Com)

2. Sélectionner les communes du département de la Sarthe de plus de 1500 habitants en affichant un tableau avec les noms de communes et leur population.

3. Sélectionner les communes de la table COMMUNE dont le statut n'est pas chef-lieu de canton et afficher les colonnes NOM_COMM en lui donnant comme alias NOM et les colonnes, STATUT, POPULATION et SUPERFICIE

4. Sélectionner les différents noms des tronçons comportant le nom 'ruisseau' dans la colonne TOPONYME de la table TRONCON_HYDROGRAPHIQUE

5. A partir de la table COMMUNE, calculer pour chaque département ; la population totale, la densité moyenne de population des communes = moyenne(population commune /superficie commune) arrondie à deux décimales, la population de la commune la plus peuplée et celle de la moins peuplée, la superficie moyenne des communes.

6. Quels sont les surfaces (en km2) et périmètres (en km), arrondis à deux chiffres après la virgule, des communes du département de la Sarthe ?

7. Sélectionner le nombre de tronçons de la 'rivière le loir', par classe de largeur (colonne LARGEUR)

8. Quelle est la longueur de la 'rivière le loir' par type de largeur sur ce jeu de données ?