

Les outils pour gérer les bases de données que nous allons utiliser dans ce cours se trouvent sur un serveur distant. La première partie de ce TP est de vous montrer comment vous pouvez vous connecter au serveur (voir partie 1) et travailler dessus (voir partie 2) depuis les salles de TP, mais aussi depuis chez vous.

1 Se connecter au serveur

1.1 Depuis les salles TPs

1. Connectez vous sur un poste de travail (éventuellement en tant qu'invité).
2. Lancez une fenêtre terminal.
3. Dans cette fenêtre, tapez

```
ssh entid@etu-pgsql.math.univ-paris-diderot.fr
```

en remplaçant `entid` par votre identifiant ENT afin de vous connecter au serveur.

4. Entrez ensuite votre mot de passe ENT lorsque le serveur vous le demande et validez avec la touche Entrée. Vous êtes désormais connecté au serveur.

1.2 Depuis chez vous

Vous pouvez accéder à la base de données depuis chez vous en vous connectant en SSH. Sous Linux ou OSX (Mac), la démarche est identique à celle des salles de TP, il vous suffit de savoir lancer un terminal, référez-vous alors à la partie 1.1.

Sous Windows, il faut installer un logiciel capable de se connecter en SSH au serveur : PuTTY <http://www.putty.org/>. Après avoir téléchargé et exécuté PuTTY, voici les réglages qu'il faut faire :

1. Dans la case "Host Name (or IP address)", inscrivez :
`entid@etu-pgsql.math.univ-paris-diderot.fr`
en remplaçant `entid` par votre identifiant ENT.
2. Vérifiez que le champ "Port" contient bien 22 et que SSH est sélectionné dans le champ "Connection type".
3. Cliquez sur le bouton "Save" pour ne pas avoir à retaper ces informations à chaque fois.
4. Cliquez enfin sur "Open". Une fenêtre noire s'affiche et vous demande votre mot de passe. Entrez votre mot de passe ENT. Vous êtes désormais connecté au serveur.

2 Travailler sur le serveur

Sur la machine `etu-pgsql.math.univ-paris-diderot.fr` vous disposez de deux *clients* `psql` et `pgcli` pour interagir avec le *serveur* de la base de données. C'est en fait le même client mais `pgcli` ajoute par-dessus `psql` des fonctions un peu plus sympathiques : complétion automatique, coloration de la syntaxe etc. Cependant, `psql` étant le client officiel, il vous sera plus facile de trouver de l'aide sur internet concernant ce client.

Ici le serveur est *local*, il écoute sur port *officiel*, il suffit de préciser le catalogue sur lequel vous souhaitez travailler (`bd_2016`), en exécutant les commandes :

```
pgcli -d bd_2016
```

ou

```
psql -d bd_2016
```

Vous devez ensuite saisir (à nouveau) votre mot de passe ENT. Vous êtes désormais connecté à la base de données.

2.1 Schémas

Pour lister les schémas de votre environnement de travail :

```
bd_2016> SHOW search_path ;
```

Pour les modifier :

```
bd_2016> SET search_path TO world, entid, public ;
bd_2016> SHOW search_path ;
```

Lister les schémas Pour lister les tables des schemas inscrits sur votre search_path.

```
bd_2016> \d
```

Schema	Name	Type	Owner
world	city	table	stbouche
world	country	table	stbouche
world	countrylanguage	table	stbouche

2.2 Tables

Pour voir la définition (c'est-à-dire les différentes colonnes) d'une table :

```
bd_2016> \d world.country
```

Column	Type	Modifiers
code	character(3)	not null
name	text	not null
continent	text	not null
region	text	not null
surfacearea	real	not null
indepyear	smallint	
population	integer	not null
lifeexpectancy	real	
gnp	numeric(10,2)	
gnpold	numeric(10,2)	
localname	text	not null
governmentform	text	not null
headofstate	text	
capital	integer	
code2	character(2)	not null

2.3 Explorer les possibilités de psql ou de pgcli

```
bd_2016> \?
```

Dans la deuxième partie du TP, on s'intéresse aux requêtes, c'est-à-dire, les moyens d'extraire une information pertinente d'une base de données.

3 Écriture d'une requête

Pour extraire des informations d'une base de données, on utilise l'algèbre relationnelle (pour la théorie) et le langage SQL (pour la pratique).

L'algèbre relationnelle est un ensemble d'opération sur les tables. Chaque opération prend en argument une ou plusieurs tables et produit une nouvelle table. Nous commençons par introduire deux opérations importantes qui opèrent sur une table à la fois :

1. Projection : $\pi_{\text{liste de colonnes}}(\mathbf{R})$. Cette opération ne garde que les colonnes mentionnées de la table R . Par exemple $\pi_{\text{name, continent}}(\text{world.country})$ est la table avec deux colonnes `name`, `continent` et une ligne pour chaque ligne de la table `world.country`.
2. Sélection : $\sigma_{\text{condition}}(\mathbf{R})$. Cette opération ne garde que les lignes qui respectent la condition indiquée. Par exemple $\sigma_{\text{lifexpectancy} < 50}(\text{world.country})$ renvoie la table contenant les pays ayant une espérance de vie de moins de 50 ans.

Ces deux opérations peuvent être écrites en SQL ainsi :

```
SELECT colonne1, colonne2, ... FROM table WHERE condition;
```

Cette opération affiche les colonnes de `table` indiquées après le `SELECT` qui respectent la `condition`. Par exemple :

```
SELECT name, lifexpectancy  
FROM world.country  
WHERE lifexpectancy < 50 and continent = 'Asia';
```

affichera le nom et l'espérance de vie des pays d'Asie ayant une espérance de vie inférieure à 50 ans. Remarquez l'utilisation des apostrophes pour indiquer des chaînes de caractères ('Asia').

4 Requêtes monotables

Avec ce qui est expliqué à la partie 3, écrivez des requêtes en algèbre relationnelle et en SQL (dans `pgcli`) pour extraire les informations suivantes du schéma `world` :

1. Quels sont les noms des pays qui comptent plus de 100 000 000 d'habitants ?
2. Quelles sont les régions ?
3. Quelles sont les régions situées en Europe ?
4. Quels sont les pays situés en europe du sud ?
5. Quelles sont les capitales des pays situés en Europe de l'Ouest ? (quel est le type de la colonne `capital` ?)
6. A partir de la table `countrylanguage`, quels sont les langues qui sont officielles dans au moins un pays ?
7. Quels sont les codes des pays où le français est langue officielle ? (même question pour les langues de votre choix)
8. Quelle est la date d'indépendance de la France ?
9. Quelles sont les dates d'indépendance des pays d'Europe ?
10. Quelles sont les villes françaises de plus de 200 000 habitants ?
11. Pour chaque pays européens, calculer la densité, le gnp par habitant, et l'espérance de vie, ordonner par densité décroissante

12. Quels sont les pays où l'espérance de vie n'est pas inférieure à 77 ans et le pnb par habitant n'est pas supérieur à 0.010?
13. Quels sont les pays tels que la condition (espérance de vie supérieure ou égale à 77 ans ou PNB par habitant inférieur à 0.01) n'est pas vérifiée?
14. Quels sont les pays où une langue est officielle sans être parlée par au moins la moitié de la population?
15. Quels sont les pays qui ont au moins une langue officielle?

5 Requêtes multi-tables

On peut aussi combiner plusieurs tables. Pour ce TP, nous allons seulement présenter le produit cartésien de deux tables : $T \times S$ est la table dont les colonnes sont les colonnes de S et les colonnes de T . Ces lignes contiennent tous les couples (l_1, l_2) où l_1 est une ligne de T et l_2 est une ligne de S . En SQL, on écrira :

```
SELECT col1,col2 FROM table1,table2 WHERE condition;
```

Par exemple,

```
SELECT language FROM world.country as c,world.countrylanguage as l  
WHERE c.code = l.countrycode and c.continent = 'Europe';
```

affichera les langues parlées en Europe. Remarquez l'utilisation des `as` pour donner de nouveaux noms aux tables et l'utilisation de `c.code` pour lever l'ambiguïté sur des noms de colonnes qui seraient éventuellement les mêmes.

Avec ça, écrivez des requêtes pour les questions suivantes :

1. Quels sont les noms des capitales Sud-Américaines?
2. Quels sont les noms des pays où le français est langue officielle?
3. Quelles sont les pays où l'espagnol est langue officielle et la forme de gouvernement est `federal republic`?
4. Quels sont les pays qui ont plus de deux langues officielles?
5. Quels sont les pays qui n'ont pas de langue officielle?
6. Quels sont les pays qui comportent au moins deux villes de plus de 1 000 000 habitants?
7. Quelles sont les régions qui ne comportent qu'une seule forme de gouvernement?
8. Quelles sont les régions où on ne trouve pas de monarchie?