

# Base de Données et langage SQL

## Situation d'Apprentissage et d'Evaluation (SAE) 1-01

*Création et reporting à partir de données stockées dans un  
SGBD relationnel : Freedom in the world*

**Nom : Diallo**

**Prénom : Thierno**

**Mail : thiernosaliou0223@gmail.com**

**Groupe : Galton**

**Promotion : BUT SD année 1**

**Enseignant : Laurent Audibert**

### Table des matières

I.	Mise en place de l'environnement .....	2
1.	Installation de PostGreSQL. ....	2
2.	Exécution du script de peuplement de la base de données. ....	2
3.	Mise en place de Metabase.....	2
4.	Génération du premier graphique .....	3
II.	Compréhension de la base de données.....	3
1.	Explication détaillé de la base de données .....	3
2.	Requête SQL permettant de calculer la moyenne de deux indicateurs de liberté en 2020 pour chacune des régions.....	4
3.	Diagrammes en barre pour observer la répartition des pays en fonction des indicateurs.....	4
a)	Civil liberties .....	4
b)	Political rights. ....	4
III.	Travail exploratoire.....	5
1.	Création de nuage de point (requête, graphique, commentaire).....	5
2.	Représentation d'une observation personnelle.....	5

## I. Mise en place de l'environnement

Mon système d'exploitation est Ubuntu.

### 1. Installation de PostgreSQL.

Etant sur linux, nous utilisons la commande « `sudo apt-get install postgresql` » pour installer PostgreSQL. L'exécution donne :

```
thierno0223@thierno0223-X540NA:~$ sudo apt-get install postgresql
[sudo] Mot de passe de thierno0223 :
Lecture des listes de paquets... Fait
Préparation du dépaquetage de .../postgresql_14+238_all.deb ...
Dépaquetage de postgresql (14+238) ...
Paramétrage de postgresql (14+238) ...
thierno0223@thierno0223-X540NA:~$
```

Cela indique que l'installation a réussi

### 2. Exécution du script de peuplement de la base de données.

Afin de lancer le script de peuplement de la base de donnée, nous nous connectons à PostgreSQL avec la commande : « `sudo su postgres` » puis « `psql` ». Ensuite nous créons une nouvelle base de données avec la commande SQL « `CREATE DATABASE sae_sql` » puis basculons sur elle avec « `\c sae_sql` ». Pour donner l'accès au fichier à l'utilisateur Postgres, il faut utiliser la commande « `chmod 777 freedomFull.sql` » ensuite copier le fichier dans le dossier /tmp/ avec la commande « `cp freedomFull.sql /tmp` ». Enfin nous lançons la commande suivante pour exécuté le script SQL : « `\i /tmp/freedomFull.sql` »

```
thierno0223@thierno0223-X540NA:~/project/data_base/sae_sql$ cp freedomFull.sql /tmp/
thierno0223@thierno0223-X540NA:~/project/data_base/sae_sql$ sudo su postgres
[sudo] Mot de passe de thierno0223 :
postgres@thierno0223-X540NA:/home/thierno0223/project/data_base/sae_sql$ psql

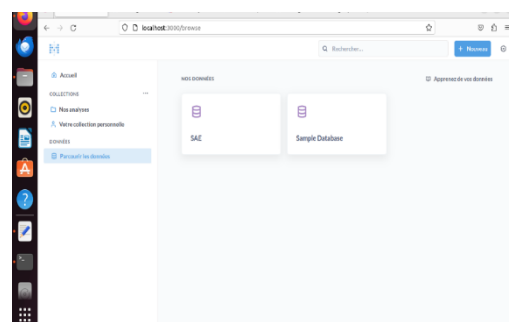
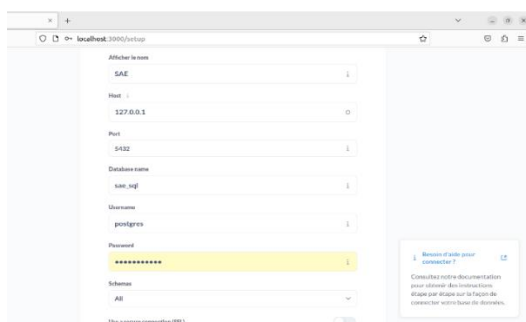
postgres=# create database sae_sql;
CREATE DATABASE
postgres=# \c sae_sql
You are now connected to database "sae_sql" as user "postgres".
sae_sql=# \i /tmp/freedomFull.sql
psql:/tmp/freedomFull.sql:1: NOTICE:  table "status" does not exist, skipping
DROP TABLE
psql:/tmp/freedomFull.sql:2: NOTICE:  table "country" does not exist, skipping
DROP TABLE
psql:/tmp/freedomFull.sql:3: NOTICE:  table "region" does not exist, skipping
DROP TABLE
psql:/tmp/freedomFull.sql:4: NOTICE:  table "freedom" does not exist, skipping
DROP TABLE
CREATE TABLE
CREATE TABLE
CREATE TABLE
CREATE TABLE
INSERT 0 1
INSERT 0 1
```

### 3. Mise en place de Metabase

Pour utiliser Metabase, nous avons besoin de java. Etant sur linux, java est installer par défaut et nous pouvons le vérifier avec la commande « `java --version` »

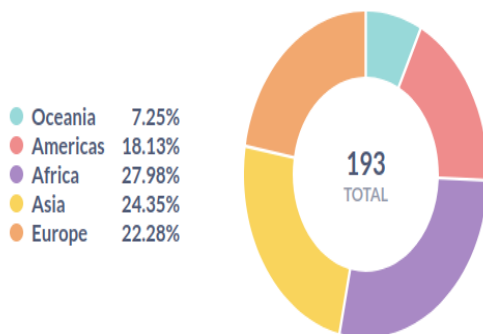
```
thierno0223@thierno0223-X540NA:~$ java --version
openjdk 11.0.20.1 2023-08-24
```

Ensuite, nous téléchargeons Metabase sur le site officiel, puis le rangeons dans un dossier. Pour lancer Metabase, nous tapons dans le terminal « `java -jar metabase.jar` », et complétons le formulaire de connexion.



## 4. Génération du premier graphique

nombre pays par region (question 1-4)



Pour générer notre premier graphique avec Metabase, nous devons cliquer sur 'nouveau' puis sur 'requête SQL' et nous écrivons notre requête : ici nous voulons récupérer le nombre de pays par région.

Cela s'écrit : « `SELECT COUNT (country.name), region.name FROM country JOIN region USING (region_code) GROUP BY region.region_code;` » puis nous

appuyons "CTRL + entrer".

Ensuite nous cliquons sur visualisation et choisissons diagramme circulaire et paramétrons le graphique ce qui donne ce graphique ci-dessus

## II. Compréhension de la base de données

### 1. Explication détaillé de la base de données

- Table status : elle contient l'ensemble des status possible pour un pays, c'est-à-dire F, PF, N
  - status : il est au format VARCHAR et est utiliser comme clé primaire car il y'a pas de répétition.
- Table region : Cette table contient le nom de toutes les régions du monde (continent)
  - Region\_code : cet attribut est de type INTEGER et contient le code attribué à chaque région. Ce code étant unique pour chaque région, cet attribut est aussi utiliser comme PRIMARY KEY.
  - Name : cet attribut est du type CHARACTER VARYING (une chaine de caractère variable) et contient le nom de la région
- Table country : cette table contient la liste de tous les pays et leurs différentes caractéristiques.
  - id\_country : c'est un INTEGER qui est unique à chaque pays et permet de l'identifier (PRIMARY KEY)
  - name : ce champ est du type VARCHAR (chaine de caractère variable) et contient le nom du pays
  - region\_code : ce champ de type INTEGER est une clé étrangère qui fait référence au champ region\_code de la table region et contient code de la region dans la quel il se situe.
  - is\_ldc : c'est un champ de type BOOLEAN qui contient l'état du pays. C'est à dire si un pays est un pays en voie de développement ou non.
- Table freedom : comprend l'ensemble des rapports émis par Freedom in the World pour chaque pays depuis 1972 (indice, année, etc...)
  - id\_country : de type INTEGER, id\_country est une clé étrangère qui fait référence à l'attribut id\_country de la table country, il contient identifiant du pays en question.
  - year : c'est un champ de type INTEGER qui contient l'année de publication du rapport.
  - civil\_liberties : de type INTEGER, ce champ contient un nombre entier entre 1 et 7 qui exprime le degré de liberté civil d'un pays l'année du rapport.

- political\_rights : de type INTEGER, ce champ contient un nombre entier entre 1 et 7 qui exprime le degré de liberté politique (droit politique) d'un pays l'année du rapport.
- status : ce champ de type VARCHAR est une clé étrangère qui fait référence à la colonne status de la table status, contient le status d'un pays l'année du rapport.
- (id\_country et year) : ces champs à eux deux, forment une clé primaire pour identifier chaque ligne de la table freedom.

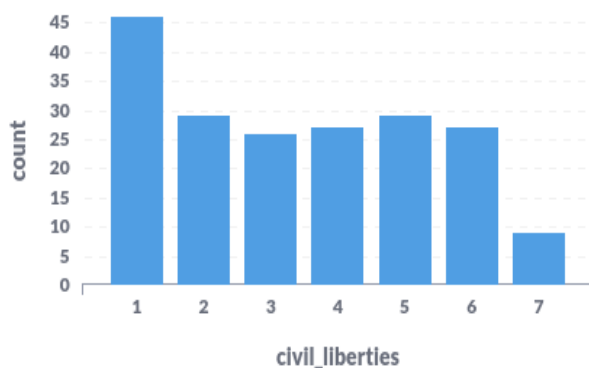
## 2. Requête SQL permettant de calculer la moyenne de deux indicateurs de liberté en 2020 pour chacune des régions.

**Requête :** « SELECT AVG(civil\_liberties) AS moy\_civil\_liberties, AVG(political\_rights) AS moy\_political\_rights, region.name FROM freedom JOIN country USING(id\_country) JOIN region USING(region\_code) WHERE freedom.year = 2020 GROUP BY region.name; »

## 3. Diagrammes en barre pour observer la répartition des pays en fonction des indicateurs.

### a) Civil liberties

**Requête :** « SELECT COUNT(civil\_liberties), civil\_liberties FROM freedom WHERE year = 2020 GROUP BY civil\_liberties; ». Ce qui nous donne :

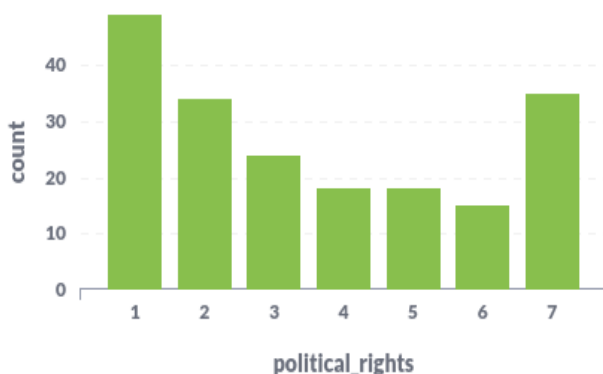


#### Commentaire :

Nous constatons que environ 25% des pays ont reçu une note de 1. Ensuite, le nombre de pays ayant obtenu une note de 2 à 6 est environ le même, soit 25 pays par modalité. Et enfin, nous avons environ 10 pays qui ont obtenu une note de 7. Nous pouvons dire ainsi que qu'il y'a plus de la moitié des pays qui ont un degré de liberté civil acceptable (soit moins de 3,5).

### b) Political rights.

**Requête :** « SELECT COUNT(political\_rights), political\_rights FROM freedom WHERE year = 2020 GROUP BY political\_rights; ». Ce qui nous donne :



#### Commentaire :

Nous voyons que le degré 1 (droit politique le plus respecté) compte le plus grand nombre de pays, soit 50 pays. Suivie par le 2 et le 3. Ensuite nous avons les degrés 4,5 et 6 qui comptent le moins de pays (environ 15pays/degré).enfin nous avons le degré 7 (pays où le droit politique est le moins respecté) qui compte de nombreux pays (35 pays).

Ainsi nous pouvons dire que moins le droit politique est respecté, moins il y'a de pays, excepter pour le 7 qui compte de nombreux pays.

### III. Travail exploratoire

#### 1. Création de nuage de point (requête, graphique, commentaire)

```
« SELECT STRING_AGG(country.name,';'), STRING_AGG(region.name,';'),
COUNT(*), civil_liberties, political_rights FROM freedom JOIN country
USING(id_country) JOIN region USING(region_code) WHERE year = 2020
GROUP BY ((civil_liberties, political_rights)); »
```



**Remarque :** Nous optons pour cette requête car il existe plusieurs pays qui ont les mêmes notes (ex: 1 et 1). Ainsi nous pouvons visualiser la densité de pays par couple de note

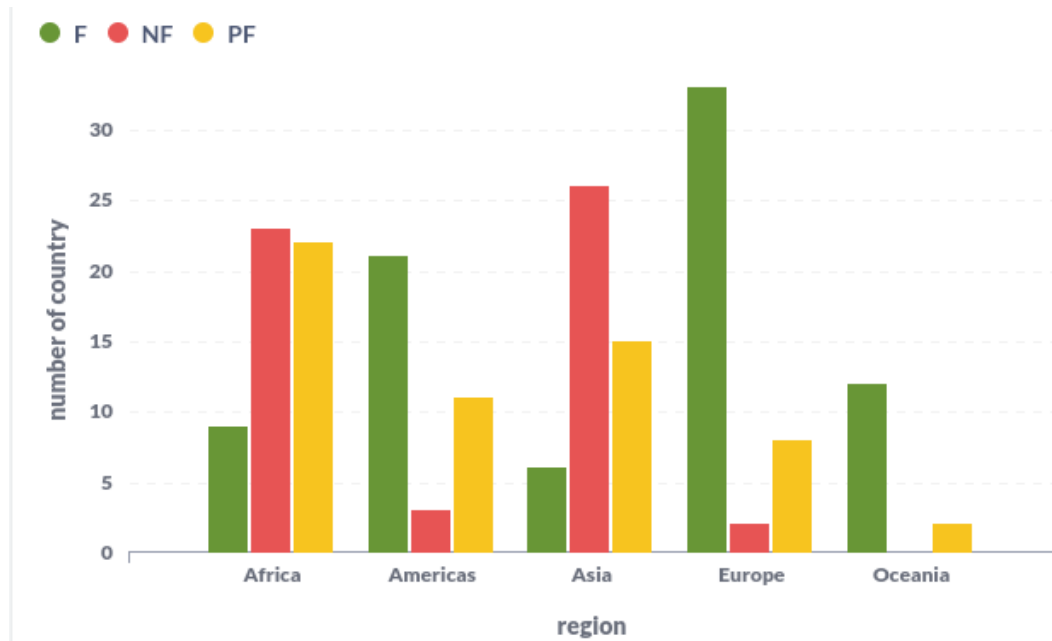
#### **Commentaire :**

Ce nuage de point nous présente la répartition des pays en fonction de la liberté civile et du droit politique en 2020. En abscisse, nous avons la liberté civile, en ordonnée le droit politique et la taille du cercle représente le nombre de pays. D'une part, pour la majorité des pays ayant un degré de liberté inférieur ou égal à 4 (les plus libres) ont également un degré de droit politique (droit politique le plus respecté) inférieur ou égal à 4. Et cette zone compte également le plus de cercles mais aussi de grands cercles, donc le plus de pays. D'autre part, pour la majorité des pays ayant un degré de liberté supérieur ou égal à 4 (les moins libres) ont également un degré de droit politique (droit politique le moins respecté) supérieur ou égal à 4. Et cette zone compte le plus petit nombre de cercles. Ainsi d'une manière générale, nous voyons que plus les pays respectent la liberté civile, plus ils respectent le droit politique et vice-versa.

#### 2. Représentation d'une observation personnelle

Comme observation personnelle, nous choisissons d'afficher le nombre de pays par statut et par régions afin de les comparer. Cette requête présente un intérêt car elle utilise de nombreuses connaissances (groupement, jointure, filtre, fonction d'agrégation, ordre) mais aussi elle permet de voir la répartition de tous les statuts pour chaque région.

```
Requête : « SELECT r.name AS région, f.status, COUNT(*) AS "number of_countries"
FROM region r JOIN country c USING (region_code) JOIN freedom f USING
(id_country) WHERE year = 2020 GROUP BY r.name, f.status ORDER BY r.name,
f.status; »
```



### Commentaire :

Ce diagramme en barre nous présente le nombre de pays par statut dans chaque région.

Ainsi pour la région « Afrique », nous avons la majorité des pays qui sont partiellement libre (22) ou pas libre (23). Sur le continent américain, nous avons 21 pays qui sont totalement libre et environ 10 qui le sont partiellement. Puis sur le continent européen, la quasi-totalité des pays sont totalement libre et en Océanie également. Cependant sur le continent asiatique la majorité des pays sont considéré comme non libre (26) et partiellement libre (15).

Nous voyons donc le nombre de pays libre varie beaucoup en fonction de la région. Mais aussi que la région européenne est la plus libre et la région asiatique la moins libre.