

Bases de données et langage SQL

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

Nayl Saifoudine BUT SD Galton

04/12/2023



Sommaire

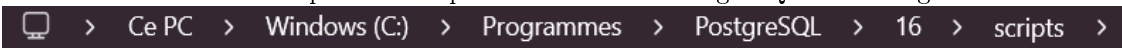
1	Mise en place de l'environnement	2
1.1	Installation de PostgreSQL	2
1.2	Exécution du script de création et peuplement de la base de données	2
1.3	Mise en place de Metabase et connexion à votre base PostgreSQL	2
1.4	Génération d'un premier graphique avec Metabase	2
2	Compréhension de la base de données	3
2.1	Explication détaillée de la base de données	3
2.2	Requête SQL permettant de calculer la moyenne des deux indicateurs de liberté en 2020 pour chaque région	3
2.3	Requêtes SQL permettant de produire deux diagrammes à barres des 2 indicateurs de liberté en 2020	4
2.3.1	Droit Politique	4
2.3.2	Liberté civile	4
3	Travail exploratoire	5
3.1	Première étude	5
3.2	Deuxième étude personnelle	5

1 Mise en place de l'environnement

1.1 Installation de PostgreSQL

- Se rendre sur le site :
<https://www.postgresql.org/download/>
- Cliquez sur l'onglet avec le système d'exploitation de son ordinateur (ex : Windows)
- Cliquez sur le lien "Download the installer"
- Téléchargez la version la plus récente compatible avec votre ordinateur
- Exécuter le fichier téléchargé en double-cliquant dessus
- Spécifiez le répertoire d'installation si besoin sinon cliquez sur "Suivant"
- Cliquez sur "Suivant" jusqu'à ce que l'installateur vous demande quels drivers installer
- Vous cochez les cases : **PostgreSQL (64 bit) v.16.1-1, psqlODBC (64 bit) v.13.02.0000-1, pgAgent (64 bit) for PostgreSQL 16 v4.2.2-1, SQL/Protect for PostgreSQL 16 (64 bit v16.0-2)**
- Cliquez sur "Suivant" et démarrez l'installation
- Remplissez ensuite les informations nécessaires à la fin de l'installation
- Bravo, vous avez installé PostgreSQL!

1.2 Exécution du script de création et peuplement de la base de données

- Se rendre sur le site :
http://gurau-audibert.hd.free.fr/josdblog/wp-content/uploads/2023/11/freedomFull.sql_.zip
Si le lien ne marche pas, remplacer à la main le symbole "underscore" et réessayez
- Extrayez le fichier téléchargé
- Placez ce fichier dans le répertoire scripts dans 16 dans PostgreSQL dans Programmes :

- Rendez vous sur votre "SQL Shell"
- Connectez-vous sur ce dernier
- Exécutez le fichier "freedomFull.sql" téléchargé avec la commande suivante :
`\i freedomFull.sql`
- Voilà votre base de données remplie!

1.3 Mise en place de Metabase et connexion à votre base PostgreSQL

- Installer Java (Minimum 11) sur votre ordinateur
- Aller sur le site :
<https://www.metabase.com/start/oss/>
- Télécharger le fichier JAR mis à disposition
- Tapez ensuite sur l'invite de commande Windows (ou équivalent sur Mac) l'instruction suivante :
`java -jar metabase.jar`
Remarque : Vous devez laisser votre invite de commande (ou terminal) ouvert avec la commande exécutée en arrière-plan pour vous connecter à Metabase
- Aller sur votre navigateur et copiez-collez ce lien :
<http://localhost:3000/>
- * Pour vous reconnecter ensuite à Metabase, vous devrez recommencer les 2 étapes précédentes*
- Renseignez ensuite les informations nécessaires à la connexion avec PostgreSQL grâce au SQL shell
- Bravo, votre base de données Metabase est maintenant relié à votre PostgreSQL!

1.4 Génération d'un premier graphique avec Metabase

A partir du menu de Metabase, vous pourrez explorer vos collections d'analyses ainsi que vos données de PostgreSQL. Vous pourrez écrire des requêtes SQL comme dans un terminal! Mais ce n'est pas tout! Vous pourrez également, après avoir écrit une requête SQL, la visualiser à travers un graphique!
Pour ce faire, voici comment procéder :

- Cliquez sur "Nouveau" en haut à droite de l'écran puis cliquez sur "Requête SQL"

- Sélectionnez ensuite votre base de données
- Ecrivez une requête en SQL
- Exécutez la en appuyant sur le triangle en bleu ou sur votre clavier avec **Ctrl + Enter**
- Cliquez ensuite sur "Visualisation" en bas à gauche de votre écran
- Choisissez le type de graphique que vous voulez créer
- Cliquez ensuite sur la roue crantée située à droite du bouton "Visualisation" ou sur le type de graphique
- Vous arriverez sur un menu composé de trois colonnes, cliquez sur "Données"
- Choisissez ensuite les données que vous voulez étudier sur l'abscisse et l'ordonnée
- Vous avez réussi à créer votre premier graphique sur Metabase! Félicitations!

2 Compréhension de la base de données

2.1 Explication détaillée de la base de données

Après avoir exécuté la commande `\i freedomFull.sql`, votre base de données devrait être composée de 4 tables différentes nommées :

country, freedom, region, status. Pour vérifier que votre base est belle et bien remplie, vous pouvez taper cette commande sur votre "SQL Shell" : `\d`

Si votre base contient bien les 4 tables alors vous pouvez continuer sinon vous avez oublié / mal effectué une étape précédente. Dans ce cas-là, vérifiez que vous avez bien suivi toutes les étapes précédentes correctement.

La première table de la base de données est la table **status**.

Elle est composée d'une colonne nommée *status*. Elle est de type *VARCHAR*, c'est à dire une chaîne de caractères de longueur variable. Elle est aussi une clé primaire.

Les modalités de cette colonne sont "NF", "PF" et "F" qui signifient "Not Free", "Partly Free" et "Free" et qui se traduisent par "Pas Libre", "Partiellement Libre" et "Libre".

La deuxième table est la table **region**.

Elle est composée de deux colonnes nommées *region_code* et *name*. La première est de type *INTEGER*, donc un entier, et est une clé primaire qui ne peut pas être nulle. Quant à la deuxième, elle est de type *CHARACTER VARYING*, donc une chaîne de caractères de longueur variable. Cette table représente les noms des 5 continents et leurs codes de continent uniques pour les identifier.

La troisième table est la table **country**.

Elle est composée de quatre colonnes nommées *id_country*, *name*, *region_code* et *is_ldc*. La première est de type *INTEGER*, donc un entier, et est une clé primaire. Quant à la deuxième, elle est de type *VARCHAR*, donc une chaîne de caractères de longueur variable, qui définit le nom du pays. La troisième colonne est de type *INTEGER* et c'est aussi une clé étrangère qui fait référence à la colonne *region_code* dans la table **region**. La dernière colonne est de type *BOOLEAN* qui définit si le pays en question est considéré comme étant l'un des pays les moins développés du globe ou pas.

La dernière table est la table **freedom**.

Elle est composée de cinq colonnes nommées *id_country*, *year*, *civil_liberties*, *political_liberties* et *status*. La première est de type *INTEGER*, donc un entier, qui est une clé étrangère faisant référence à la colonne *id_country* dans la table **country**. Quant à la deuxième, elle est aussi de type *INTEGER* définissant l'année de l'étude. La troisième table est de type *INTEGER* et elle définit le degré de liberté civile dans le pays. La quatrième table est également de type *INTEGER* et elle définit le degré de liberté politique dans le pays. Et la dernière table est de type *VARCHAR* et c'est une clé étrangère qui fait référence à la table *status*. La clé primaire de cette table est la combinaison de deux tables : les tables *id_country* et *year*. Cette table détermine le niveau de liberté de chaque pays.

2.2 Requête SQL permettant de calculer la moyenne des deux indicateurs de liberté en 2020 pour chaque région

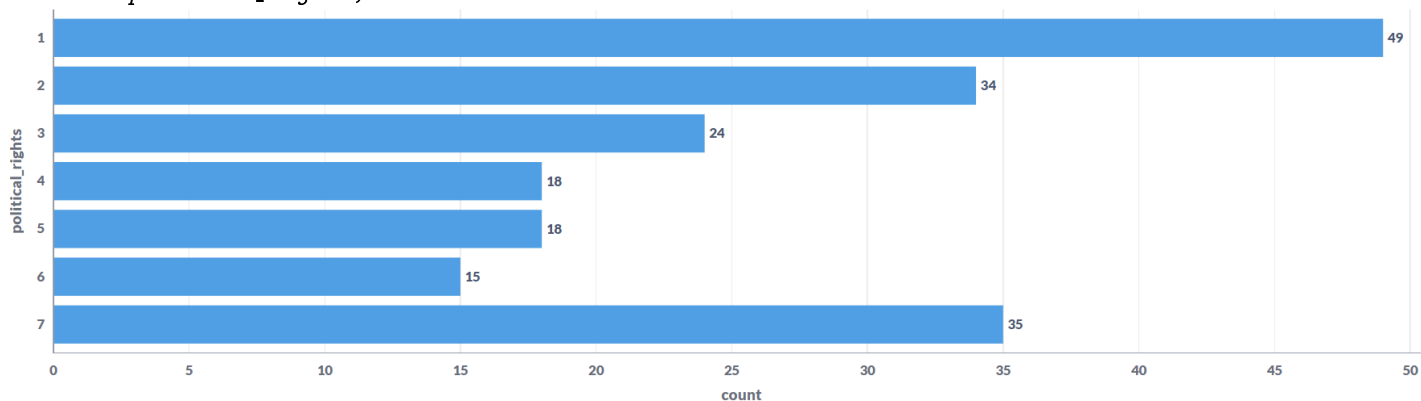
```
SELECT AVG(political_rights) AS politics, AVG(civil_liberties) AS civil, region.name FROM freedom
JOIN country USING(id_country) JOIN region USING(region_code) WHERE year = 2020 GROUP BY region.name;
```

▼ politics	▼ civil	name ▼
1.71	1.71	Oceania
2.51	2.57	Americas
4.81	4.44	Africa
4.98	4.83	Asia
1.86	1.84	Europe

2.3 Requêtes SQL permettant de produire deux diagrammes à barres des 2 indicateurs de liberté en 2020

2.3.1 Droit Politique

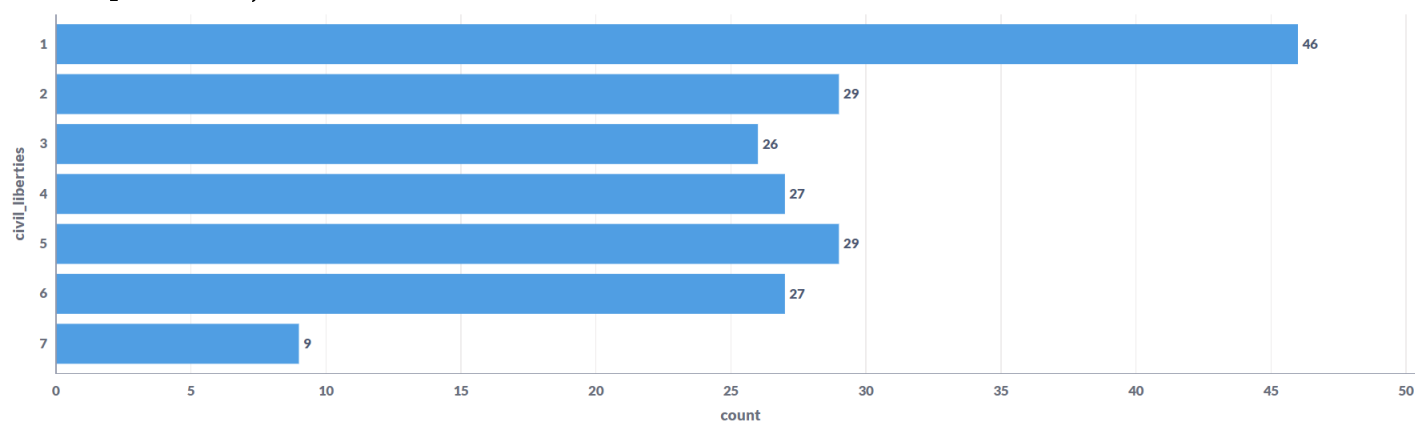
```
SELECT political_rights, COUNT(*) FROM freedom WHERE year = 2020 GROUP BY political_rights ORDER BY political_rights;
```



Avec cette requête SQL, on compte le nombre de pays en fonction des droits politiques et on tri les modalités par ordre croissant. On remarque, par exemple, qu'il y a 49 pays dont le degré des droits politiques est de 1, qui est le degré de droit politique le plus faible.

2.3.2 Liberté civile

```
SELECT civil_liberties, COUNT(*) FROM freedom WHERE year = 2020 GROUP BY civil_liberties ORDER BY civil_liberties;
```

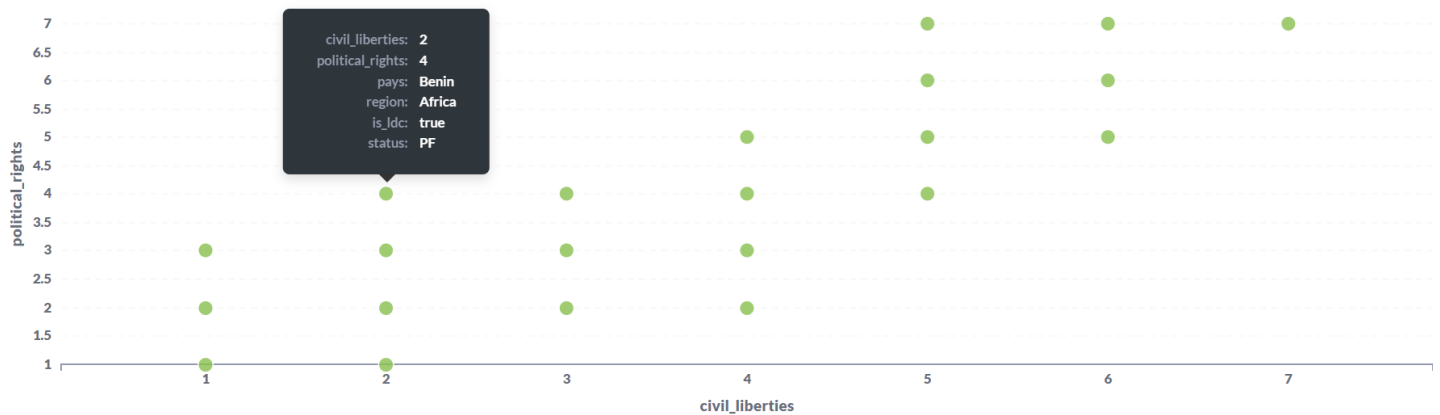


Avec cette requête SQL, on compte le nombre de pays en fonction de sa liberté civile et on tri les modalités par ordre croissant. On remarque, par exemple, qu'il y a seulement 9 pays dont le degré de liberté civile est de 7, qui est le degré de liberté civile le plus fort.

3 Travail exploratoire

3.1 Première étude

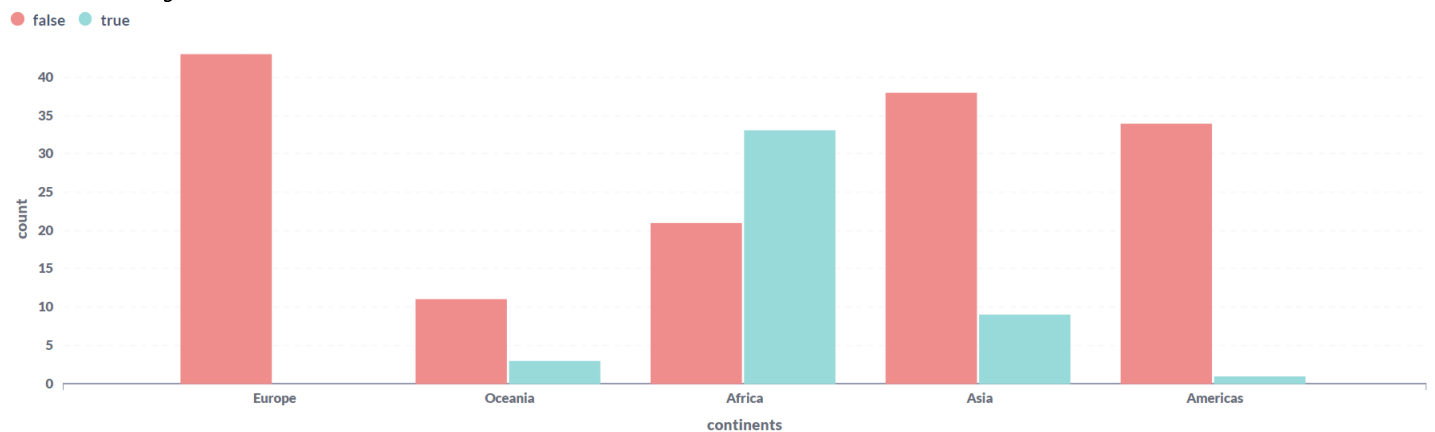
```
SELECT civil_liberties, political_rights, country.name AS Pays, region.name AS region, is_ldc,
status FROM freedom JOIN country USING (id_country) JOIN region USING(region_code) WHERE year
= 2020 ;
```



Cette requête permet de visualiser les pays en fonction de sa liberté civile et de ses droits politiques. En se positionnant sur l'un des points du graphique, on peut voir quels pays appartiennent à cette catégorie, le continent de ce pays, la condition du pays (si c'est l'un des pays les moins développés) et le statut de liberté du pays. Par exemple, on voit dans le graphique que le Benin, situé en Afrique, est l'un des pays les moins développés, a le statut de pays "Partiellment Libre", a un degré de liberté civile de 2 et un degré de droits politiques de 4. On remarque cependant que les pays ayant le même niveau de liberté civile et droits politiques sont superposés donc on ne voit pas tous les pays lorsque l'on se positionne sur un point du graphique.

3.2 Deuxième étude personnelle

```
SELECT is_ldc, region.name AS Continents, COUNT(*) FROM country JOIN region USING(region_code)
GROUP BY (region.name, is_ldc);
```



Cette requête permet de visualiser le nombre de pays considérés comme étant les moins développés en fonction des continents. Comme le montre la légende, les barres verticales bleues sont les pays les moins développés et les barres verticales rouges sont les pays qui ne sont pas considérés comme l'un des pays les moins développés.

On remarque donc que le continent ayant le plus de pays considérés comme étant les moins développés est l'Afrique. En effet, c'est le seul continent ayant plus de pays considérés comme les moins développés que de pays développés. A l'inverse, l'Europe est le seul continent n'ayant pas de pays considérés comme les moins développés (et l'Amérique se rapprochant grandement avec 1 seul pays considéré comme ceci). Ce graphique montre donc que la répartition des pays considérés comme étant les moins développés est inégale selon les régions du monde et que c'est l'Afrique qui est le continent le plus touché par cette inégalité.