

TD 2 : SQL

SQL Interrogation de données avec ChatGPT

2024-09-27



Avec solutions

Au delà des solutions calculées à la main, nous présentons des solutions proposées par un *chatbot*, en l'occurrence, *chatgpt* 3.5 de <https://openai.com>

- L3 MIASHS/Ingémath
- [Université Paris Cité](#)
- Année 2024-2025
- [Course Homepage](#)
- [Moodle](#)



Schéma world (rappel)

Utiliser les commandes suivantes pour retrouver les schémas de tables `country`, `countrylanguage` et `city`.

```
bd_2023> \d world."nom de table"
```

[Vue d'ensemble du schéma world](#)

Informier le chatbot du schéma utilisé

Ici nous utilisons *chatgpt* 3.5 (voir <https://openai.com>).

Nous partons d'une description du schéma `world` construite par le client [dbSchema](#). Le client `dbSchema` est connecté au schéma `world` du catalogue `bd_2023-24` sur le serveur `etu-pgsql`. Le modèle (menu Model) est exporté (sous menu Export) au format HTML/PDF (en markdown) et sauvé dans un fichier `world_schema.md`.

Dans la conversation avec le *chatbot*, nous attachons le fichier `world_schema.md` et nous précisons son usage.

```
The attachment contains the definition of the world database.
```

```
Write an sql query for determining the capitals of countries from region "Western Europe"
```

Requêtes

- Quels sont les noms des capitales d'Europe de l'ouest ? (14 lignes)

💡 Solution

Question posée au chatbot :

Write an sql query for determining the capitals of countries from region "Western Europe".

La réponse est :

Based on the structure provided from the world database, it appears that:

The country table contains the region, name_country, and capital fields. The city table contains city-related information, including the name of the city and its id, which is linked to the capital field in the country table through a foreign key. We can join the country and city tables using the capital (in country) and id (in city) fields to find the capital cities of countries in the "Western Europe" region.

Here is the SQL query:

```
SELECT
    city.name AS capital_name,
    country.name_country AS country_name
FROM
    world.country JOIN
    world.city ON
    country.capital = city.id
WHERE
    country.region = 'Western Europe';
```

This query retrieves the capital names and their corresponding country names for countries in the "Western Europe" region by joining the country and city tables.

La réponse à la main était :

```
SELECT city.name
FROM world.city JOIN
    world.country ON capital=id
WHERE continent = 'Western Europe';
```

- Quels sont les noms des pays où le français est langue officielle? (18 lignes)

Solution

Question au chatbot :

In which countries is French an official language?

La réponse est :

To find the countries where French is an official language, we need to reference the `countrylanguage` table. This table contains information about languages spoken in countries, including whether a language is official. The table can be joined with the `country` table to get the country names.

et la requête SQL est

```
SELECT
  country.name_country AS country_name
FROM
  world.country
JOIN
  world.countrylanguage ON
  country.countrycode = countrylanguage.countrycode
WHERE
  countrylanguage.language = 'French' AND
  countrylanguage.official = TRUE;
```

à comparer avec a requête construite à la main

```
SELECT name_country
FROM world.country NATURAL JOIN world.countrylanguage
WHERE official AND language='French';
```

- Quels sont les pays dont le gouvernement est monarchique (constitutionnelle ou pas) ? (43 lignes)

Solution

Question posée au chatbot :

Which countries are monarchies?

Réponse :

To identify countries that are monarchies, we need to check the governmentform column in the country table, as it likely contains information about the type of government each country has.

Et la requête SQL est :

```
SELECT name_country AS country_name
FROM world.country
WHERE governmentform LIKE '%Monarchy%';
```

This query retrieves the names of countries where the form of government includes the word "Monarchy" in the governmentform field.

- Quelles sont les monarchies d'Europe dont la capitale compte moins d'un million d'habitants ? (8 lignes)

💡 Solution

Question posée au chatbot :

In which European monarchies has the capital less than 1 million inhabitants?

Réponse du chatbot :

To find European monarchies where the capital city has less than 1 million inhabitants, we need to:

Filter for countries in Europe that are monarchies.

Join the country and city tables to get the population of the capital city.

Ensure the population of the capital city is less than 1 million.

Requête SQL :

```
SELECT
  country.name_country AS country_name,
  city.name AS capital_name,
  city.population AS capital_population
FROM
  world.country JOIN
  world.city ON
  country.capital = city.id
WHERE
  country.governmentform LIKE '%Monarchy%' AND
  country.continent = 'Europe' AND
  city.population < 1000000;
```

À comparer avec :

```
WITH T1 AS (
  SELECT *
  FROM world.country
  WHERE governmentform LIKE '%Monarchy%' AND continent='Europe' )

SELECT name_country
FROM T1 JOIN world.city ON capital=id
WHERE population < 1000000;
```

- Quelles sont les régions où on ne trouve pas de monarchie ? (9 lignes)

💡 Solution

Question posée au chatbot :

In which regions are there no monarchies?

Réponse :

To find the regions where there are no monarchies, we can:

Identify regions that do not have any countries with a monarchy.

Use the governmentform field in the country table to filter out monarchies.

Requête SQL :

```
SELECT DISTINCT region
FROM world.country
WHERE region NOT IN (
    SELECT region
    FROM world.country
    WHERE governmentform LIKE '%Monarchy%'
);
```

À comparer avec

```
WITH monarchy AS (
    SELECT region
    FROM world.country
    WHERE governmentform LIKE '%Monarchy%')

(SELECT region
 FROM world.country)
EXCEPT
(SELECT *
 FROM monarchy);
```

- Quelles sont les langues qui ne sont langues officielles dans aucun pays ? (355 lignes)



Solution

Question posée au chatbot :

Which languages are nowhere official?

Réponse :

To find languages that are not official in any country, we need to:

Query the countrylanguage table for languages where isofficial is FALSE across all entries.
Use GROUP BY and HAVING to ensure the language is not marked as official in any country.

Requête SQL :

```
SELECT
  language
FROM
  world.countrylanguage
GROUP BY
  language
HAVING
  SUM(CASE WHEN isofficial = TRUE THEN 1 ELSE 0 END) = 0;
```

À comparer avec :

```
WITH R AS (
  SELECT language
  FROM world.countrylanguage
  WHERE isofficial
)

(SELECT DISTINCT language
 FROM world.countrylanguage
)
EXCEPT
(SELECT *
 FROM R
);
```

On peut demander au chatbot de revoir sa copie :

Can you write an equivalent query without any aggregation?

Réponse :

Yes! To write an equivalent query without aggregation, we can look for languages that do not appear with `isofficial = TRUE` in the `countrylanguage` table. This can be done by using a `NOT EXISTS` subquery to exclude any language that is official in any country.

Here is the SQL query:

```
SELECT DISTINCT language
FROM world.countrylanguage c1
WHERE NOT EXISTS (
  SELECT 1
  FROM world.countrylanguage c2
  WHERE c1.language = c2.language
    AND c2.isofficial = TRUE
);
```

This query finds distinct languages from the `countrylanguage` table where no entry for that language exists with `isofficial = TRUE`.

- Quels sont les pays qui n'ont pas de langue officielle ? (49 lignes)

Solution

Question posée au chatbot :

Find the countries with no official language

Réponse :

To find the countries with no official language, we need to:

1. Identify countries that do not have any language marked as official in the `countrylanguage` table.
2. Use a `NOT EXISTS` subquery to check for countries without any official language.

Here is the SQL query:

```
SELECT name_country AS country_name
FROM world.country c
WHERE NOT EXISTS (
    SELECT 1
    FROM world.countrylanguage cl
    WHERE c.countrycode = cl.countrycode
    AND cl.official = TRUE
);
```

This query selects the names of countries where there is no corresponding entry in the `countrylanguage` table with `official = TRUE`.

À comparer avec :

```
WITH R AS (
    SELECT countrycode
    FROM world.countrylanguage
    WHERE official
),
S AS (
    SELECT countrycode
    FROM world.country
    EXCEPT
    SELECT * FROM R
)

SELECT name_country
FROM world.country NATURAL JOIN S;
```

- Quels sont les pays dont la seule langue officielle est le français ? (9 lignes)



💡 Solution

Question posée au chatbot :

In which countries is French the only official language?

Réponse :

To find countries where French is the only official language, we need to:

1. Identify countries where French is an official language.
2. Ensure no other languages are marked as official in those countries.

Here is the SQL query:

```
SELECT c.name_country AS country_name
FROM world.country c
JOIN world.countrylanguage cl ON c.countrycode = cl.countrycode
WHERE cl.language = 'French'
  AND cl.official = TRUE
  AND NOT EXISTS (
    SELECT 1
    FROM world.countrylanguage cl2
    WHERE cl2.countrycode = c.countrycode
      AND cl2.language != 'French'
      AND cl2.official = TRUE
  );
```

This query retrieves the countries where French is the only official language by checking that no other language is marked as official for those countries.

À comparer avec :

```
WITH R AS(
  SELECT DISTINCT T1.countrycode
  FROM world.countrylanguage as T1 JOIN
    world.countrylanguage as T2 ON
      T1.countrycode = T2.countrycode
  WHERE T1.language <> T2.language AND
    T1.official AND
      T2.official),
S AS (
  SELECT countrycode
  FROM world.countrylanguage
  WHERE language='French' AND official
  EXCEPT
    (SELECT * FROM R)
)

SELECT name_country
FROM world.country NATURAL JOIN S;
```

Ou encore, en utilisant EXCEPT :

```
WITH OnlyFrench AS (
  (SELECT countrycode
   FROM countrylanguage
   WHERE official AND language = 'French')
  EXCEPT
  (SELECT countrycode
   FROM countrylanguage
   WHERE official AND language <> 'French')
)
```

- Les noms des pays où le français n'est pas la seule langue officielle. (9 lignes)

Solution

Question posée au chatbot :

Find the countries where French is not the only official language.

Réponse :

To find countries where French is an official language but not the only official language, we need to:

1. Identify countries where French is an official language.
2. Ensure that there is at least one other language marked as official in those countries.

Here is the SQL query:

```
SELECT c.name_country AS country_name
FROM world.country c
JOIN world.countrylanguage cl ON c.countrycode = cl.countrycode
WHERE cl.language = 'French'
  AND cl.official = TRUE
  AND EXISTS (
    SELECT 1
    FROM world.countrylanguage cl2
    WHERE cl2.countrycode = c.countrycode
      AND cl2.language != 'French'
      AND cl2.official = TRUE
  );
```

This query retrieves the countries where French is one of the official languages, but there is at least one other official language in the same country.

À comparer avec :

```
WITH R AS (
  SELECT DISTINCT T1.countrycode FROM world.countrylanguage as T1
    JOIN world.countrylanguage as T2
      ON T1.countrycode = T2.countrycode
  WHERE T1.language <> T2.language AND
    T1.language = 'French' AND
    T1.official AND
    T2.official)

SELECT name_country FROM R NATURAL JOIN world.country;
```

En utilisant INTERSECT :

```
WITH AlsoFrench AS (
  (SELECT countrycode
   FROM world.countrylanguage
   WHERE official AND language = 'French')
  INTERSECT
  (SELECT countrycode
   FROM world.countrylanguage
   WHERE official AND language <> 'French'))
)

SELECT name_country
FROM world.country NATURAL JOIN
  AlsoFrench;
```

- Quelles sont les régions qui ne comportent qu'une seule forme de gouvernement ? (3 lignes)



💡 Solution

Question posée au chatbot :

Name the regions where there is only one form of government

Réponse :

To find regions where there is only one form of government across all countries in that region, we need to:

1. Group countries by region and government form.
2. Ensure that each region has only one unique government form.

Here is the SQL query:

```
SELECT region
FROM world.country
GROUP BY region
HAVING COUNT(DISTINCT governmentform) = 1;
```

This query groups the countries by region and checks that each region has only one distinct `governmentform`. The `HAVING` clause ensures that only regions with exactly one form of government are selected.

À comparer avec :

```
WITH R AS (
  SELECT T1.region
  FROM world.country AS T1 JOIN
    world.country AS T2 ON
    T1.region = T2.region
  WHERE T1.governmentform <> T2.governmentform)

SELECT region
FROM world.country
EXCEPT
SELECT *
FROM R;
```

👉 Le chatbot est capable de formuler une réponse qui n'utilise pas d'aggrégation.

Could you write another solution without using aggregation?

conduit à

Certainly! To solve this without aggregation, we can approach it by:

1. Finding regions where every country has the same government form.
2. Using a `NOT EXISTS` subquery to ensure there is no country in the region with a different government form.

Here's the SQL query:

```
SELECT DISTINCT c1.region
FROM world.country c1
WHERE NOT EXISTS (
  SELECT 1
  FROM world.country c2
  WHERE c1.region = c2.region
    AND c1.governmentform != c2.governmentform
);
```

This query selects distinct regions where no country in the same region has a different form of government, ensuring that only one government form exists within the region.

- Quelles sont les langues officielles des pays dont la capitale compte plus de 5 000 000 d'habitants ? (12 lignes)

Solution

Question posée au chatbot :

What are the official languages of countries
which capital has more than 5 000 000 inhabitants ?

Réponse :

To find the official languages of countries where the capital city
has more than 5 million inhabitants, we need to:

1. Join the `country` and `city` tables to filter countries
where the capital's population exceeds 5 million.
2. Then, join the `countrylanguage` table to get
the official languages of these countries.

Here is the SQL query:

```
SELECT DISTINCT cl.language
FROM
  world.country c JOIN
  world.city ci ON c.capital = ci.id JOIN
  world.countrylanguage cl ON c.countrycode = cl.countrycode
WHERE
  ci.population > 5000000 AND
  cl.isofficial = TRUE;
```

This query retrieves the distinct official languages
from the `countrylanguage` table, for countries
where the capital city has a population greater than 5 million.

À comparer avec :

```
SELECT DISTINCT language
FROM
  world.countrylanguage NATURAL JOIN
  world.city JOIN
  world.country ON capital=id
WHERE
  population > 5000000 and isofficial;
```

- Quels sont les pays où au moins trois langues sont parlées chacune par strictement plus de 10% de la population ? (35 lignes)



💡 Solution

Question posée au chatbot :

In which countries, can we find three languages which are spoken by at least 10% of the population?

Réponse :

To find countries where at least three languages are spoken by at least 10% of the population, we can:

1. Filter the `countrylanguage` table for languages spoken by at least 10% of the population.
2. Group the results by country and count how many languages meet the 10% threshold.
3. Filter countries where the count is 3 or more.

Here is the SQL query:

```
SELECT
  c.name_country AS country_name
FROM
  world.country c JOIN
  world.countrylanguage cl ON c.countrycode = cl.countrycode
WHERE
  cl.percentage >= 10
GROUP BY c.name_country
HAVING COUNT(cl.language) >= 3;
```

This query finds countries where at least three languages are spoken by 10% or more of the population, based on the `percentage` column in the `countrylanguage` table.

À comparer avec :

```
WITH R AS (
  SELECT T1.countrycode
  FROM world.countrylanguage AS T1 JOIN
  world.countrylanguage AS T2 ON T1.countrycode = T2.countrycode JOIN
  world.countrylanguage AS T3 ON T1.countrycode = T3.countrycode
  WHERE T1.language <> T2.language AND
  T2.language <> T3.language AND
  T1.language <> T3.language AND
  T1.percentage > 10 AND
  T2.percentage > 10 AND
  T3.percentage > 10
)

SELECT DISTINCT name_country
FROM world.country NATURAL JOIN R;
```

ou

```
WITH tpc AS (
  SELECT *
  FROM world.countrylanguage
  WHERE percentage > 10),
three AS (
  SELECT DISTINCT t1.countrycode
  FROM tpc AS t1 JOIN
  tpc AS t2 ON (t1.countrycode=t2.countrycode AND t1.language < t2.language) JOIN
  tpc AS t3 ON (t1.countrycode=t3.countrycode AND t3.language > t2.language)
)
```

```
SELECT co.*
```

- Quelles sont les régions où il existe deux pays dont les espérances de vie diffèrent par au moins 10 ans ? (16 lignes)

💡 Solution

Question posée au chatbot :

In which regions do the largest and the smallest life expectancy differ by more than 10 years ?

Réponse :

To find regions where the difference between the largest and smallest life expectancy among countries is more than 10 years, we can:

1. Group countries by region.
2. Calculate the maximum and minimum life expectancy for each region.
3. Filter regions where the difference between the maximum and minimum life expectancy is greater than 10 years.

Here is the SQL query:

```
SELECT region
FROM world.country
GROUP BY region
HAVING MAX(lifeexpectancy) - MIN(lifeexpectancy) > 10;
```

This query groups countries by region, calculates the difference between the maximum and minimum life expectancy within each region, and selects only those regions where the difference is greater than 10 years.

À comparer avec :

```
SELECT DISTINCT T1.region
FROM world.country AS T1 JOIN
      world.country AS T2 ON T1.region = T2.region
WHERE T1.countrycode <> T2.countrycode AND
      (T1.lifeexpectancy - T2.lifeexpectancy >= 10 OR
       T2.lifeexpectancy - T1.lifeexpectancy >= 10);
```

- Quels sont les pays où l'anglais et le français sont des *langues officielles*? (3 lignes)

💡 Solution

Question posée au chatbot :

In which countries are both English and French official languages?

Réponse :

To find the countries where both English and French are official languages, we need to:

1. Filter for countries where English is an official language.
2. Ensure that French is also an official language in the same country.

Here's the SQL query:

```
SELECT c.name_country AS country_name
FROM world.country c
JOIN world.countrylanguage cl1 ON c.countrycode = cl1.countrycode
JOIN world.countrylanguage cl2 ON c.countrycode = cl2.countrycode
WHERE cl1.language = 'English' AND cl1.official = TRUE
      AND cl2.language = 'French' AND cl2.official = TRUE;
```

This query selects countries where both English and French are marked as official languages by joining the `countrylanguage` table twice, once for each language.

À comparer avec :

```
WITH R AS (
  SELECT T1.countrycode
  FROM world.countrylanguage AS T1 JOIN
       world.countrylanguage AS T2 USING (countrycode)
  WHERE T1.language='French' AND
        T2.language='English' AND
        T1.official AND
        T2.official)

SELECT name_country
FROM R NATURAL JOIN
     world.country;
```

- Montrer comment calculer l'*intersection* de deux tables avec une *jointure* (sous certaines conditions).

💡 Solution

Si R et S sont deux tables de même schéma, une jointure naturelle des deux tables devrait (en première approximation) calculer l'intersection des deux tables.
Pourtant

```
(SELECT * FROM R)
INTERSECT
(SELECT * FROM S) ;
```

ne donne pas toujours le même résultat que

```
SELECT *
FROM R NATURAL JOIN S ;
```

Essayez par exemple pour l'intersection de la table `country` avec elle-même :

```
(SELECT * FROM country)
INTERSECT
(SELECT * FROM country);
```

retourne la table `country`, soit 239 lignes.
La jointure naturelle

```
SELECT *
FROM country AS c1 NATURAL JOIN
     country AS c2;
```

ne donne que 167 lignes. On peut réconcilier les deux résultats en ne conservant que les lignes pour lesquelles aucune colonne ne prend la valeur `NULL`.
En vérifiant quels attributs peuvent prendre la valeur `NULL` dans le schéma, on peut tester cela par la requête

```
SELECT *
FROM world.country
WHERE indepyear IS NOT NULL AND
      lifeexpectancy IS NOT NUL AND
      gnp IS NOT NULL AND
      gnpold IS NOT NULL AND
      headofstate IS NOT NULL AND
      capital IS NOT NULL;
```

qui retourne bien 167 lignes.
On peut aussi demander l'avis du chatbot :

```
Is it possible to compute the intersection of two tables using joins?
```