

Projet 3
Requêtez une base de données SQL
-
Méthodologie

Parcours Data Analyst
OpenClassrooms

Eliane CAMUS

Requête 1 :

Lister les numéros de contrats (contrat_ID) avec leur surface pour la commune de Caen.

```
SELECT contrat_ID AS 'N° de contrat', surface AS 'Surface (m2) à Caen'  
FROM contrat c  
JOIN region r  
ON c.Code_dep_code_commune = r.Code_dep_code_commune  
WHERE r.com_nom_maj_court = upper ('Caen')
```

Méthodologie :

Nous avons besoin de faire apparaître à la fois les numéros de contrat (colonne “Contrat_ID”) et la surface des biens (colonne “Surface”), variables issues de la table “CONTRAT”) : on appelle donc dans SELECT “contrat_ID” et “surface” (en renommant l’intitulé des colonnes du résultat de la requête via la commande “AS”), tout en spécifiant que les colonnes sont présentes dans la table dénommée CONTRAT (via la commande “FROM”), que l’on appellera “c” par simplification.

Ces données doivent avoir pour condition la commune de Caen (colonne “com_nom_maj_court”), or cette variable est présente dans une autre table que “CONTRAT, qui est la table “REGION”. La table “REGION” et la table “CONTRAT” sont liées entre elles par le champ “Code_dep_code_commune” (qui en constitue la clé primaire dans la table “REGION” et la clé secondaire dans la table “CONTRAT”). Il est donc nécessaire pour que la requête puisse faire le lien entre les 2 tables, d’ajouter la commande “JOIN ON” (avec indication de la clé primaire de la 2^{de} table et de la clé secondaire de la 1^{ere} table).

Une fois que le lien entre les 2 tables est réalisé, nous pouvons finir avec la commande “WHERE” qui permet de spécifier la condition d’un nom de commune égal à “Caen”. Ceci permet de limiter les résultats de la requête aux surfaces uniquement des biens situés dans la commune de Caen, sans quoi nous aurions les surfaces de chaque commune listée dans la table “REGION”. Comme dans la base de données “REGION” toutes les communes sont listées en lettres capitales, j’ai ajouté le “UPPER” pour convertir “Caen” en “CAEN” (j’aurais tout aussi bien pu directement indiquer le code suivant à la place : **WHERE** r.com_nom_maj_court = 'CAEN').

Requête 2 :

Lister les numéros de contrats avec le type de contrat et leur formule pour les maisons du département 71.

```
SELECT contrat_ID AS 'N° de contrat', Type_contrat AS 'Type de contrat',  
Formule  
FROM contrat c  
JOIN region r  
ON c.Code_dep_code_commune = r.Code_dep_code_commune  
WHERE c.type_local = 'Maison'  
AND r.dep_code = 71
```

Méthodologie :

Nous avons besoin de faire apparaître à la fois les numéros de contrat (colonne “Contrat_ID”), le type de contrat (colonne “Type_contrat”) et la formule (colonne “Formule”), qui sont 3 variables issues de la table “CONTRAT” : on appelle donc dans SELECT “contrat_ID”, “type_contrat” et “formule” (en renommant l’intitulé des colonnes du résultat de la requête via la commande “AS”), tout en spécifiant que les colonnes sont présentes dans la table dénommée CONTRAT (via la commande “FROM”), que l’on appellera “c” par simplification.

Les résultats de la requête doivent répondre à 2 conditions :

- Que le type de bien soit une “Maison” : colonne “type_local” => table “CONTRAT”, table déjà appelée dans le “FROM”.
- Que le département dans lequel la maison se situe soit le 71 : colonne “dep_code” => table “REGION”, on fait donc le lien entre les 2 tables via la commande “JOIN ON”, comme expliqué dans la requête N°1.

Pour finir nous ajoutons les 2 conditions dans la commande “WHERE” (on ajoute la commande “AND” pour lister plusieurs conditions).

En ajoutant le “c.” dans le code “c.type_local = 'Maison'”, cela signifie que la colonne “type_local” se situe dans la table “CONTRAT” => “FROM contrat c”.
Idem avec le “r.” qui précède “r.dep_code = 71” : on indique que la colonne “dep_code” est issue de la table “REGION” => “JOIN region r”.

Requête 3 :

Lister le nom des régions de France.

```
SELECT reg_nom AS 'Régions de France'  
FROM region  
GROUP BY reg_nom
```

Méthodologie :

Nous avons besoin de faire apparaître uniquement les noms de régions de France (colonne “reg_nom”), issue de la table “CONTRAT” : on appelle donc cette unique colonne dans SELECT (en renommant l’intitulé de la colonne du résultat de la requête via la commande “AS”), tout en spécifiant que la colonne vient de la table dénommée CONTRAT (via la commande “FROM”), désignée comme “c” par simplification.

Comme je ne souhaite pas que le résultat de la requête me renvoie X fois chaque région de France dans ma liste (qui correspondrait au nombre d'occurrences d'apparition de chaque région dans la base de données “CONTRAT”), je dois spécifier la commande “GROUP BY reg_nom”.

Requête 4 :

Quels sont les 5 contrats qui ont les surfaces les plus élevées ?

```
SELECT contrat_ID AS 'N° de contrat', surface AS 'Surface'  
FROM contrat c  
ORDER BY surface DESC  
LIMIT 5
```

Méthodologie :

Nous avons besoin de faire apparaître à la fois les numéros de contrat (colonne “Contrat_ID”) et la surface des biens (colonne “Surface”), variables issues de la table “CONTRAT”) : on appelle donc dans SELECT “contrat_ID” et “surface” (en renommant l’intitulé des colonnes du résultat de la requête via la commande “AS”), tout en

spécifiant que les colonnes sont présentes dans la table dénommée CONTRAT (via la commande "FROM"), que l'on appellera "c" par simplification.

Ici il nous est demandé de limiter les résultats de la requête aux 5 résultats les plus élevés uniquement.

J'aimerais donc que les résultats de la requête apparaissent dans l'ordre décroissant, et que les résultats soient limités à 5 contrats/surfaces uniquement.

Pour afficher les données dans l'ordre décroissant, on ajoute la commande "ORDER BY" mais nous devons également spécifier le "DESC" sinon les résultats s'afficheront par défaut dans l'ordre croissant. Pour que la requête soit limitée à 5 résultats, j'ajoute la dernière commande "LIMIT 5".

Requête 5 :

Quel est le prix moyen de la cotisation mensuelle ?

```
SELECT ROUND(AVG(Prix_cotisation_mensuel), 2)
AS 'Prix moyen de la cotisation mensuelle'
FROM contrat c
```

Méthodologie :

Nous avons besoin de faire apparaître uniquement le prix de la cotisation mensuelle, variable issue de la table "CONTRAT" : on appelle donc dans SELECT "Prix_cotisation_mensuel" tout en spécifiant que la colonne est présente dans la table dénommée CONTRAT (via la commande "FROM"), que l'on appellera "c" par simplification.

Ici il est nécessaire de faire appel à la fonction AVG qui permet de calculer la moyenne de l'ensemble des surfaces listées dans la table "CONTRAT".

Comme nous souhaitons que la cotisation moyenne de notre requête s'affiche uniquement avec 2 décimales, nous imbriquons la fonction "AVG(Prix_cotisation_mensuel)" dans la commande "ROUND", avec la précision du nombre de décimales (ici 2) dans le 2nd paramètre.

Requête 6 :

Quel est le nombre de contrats pour chaque catégorie de prix de la valeur déclarée des biens ?

```
SELECT Valeur_declaree_biens AS 'Catégorie de prix de la valeur déclarée des
biens', COUNT(Contrat_ID) AS 'Nombre de contrats'
FROM contrat c
GROUP BY Valeur_declaree_biens
ORDER BY COUNT(Contrat_ID) DESC
```

Méthodologie :

Nous avons besoin de faire apparaître à la fois le nombre de contrats (colonne "Contrat_ID") et la catégorie de prix "Valeur_declaree_biens", variables issues de la table "CONTRAT" : on appelle donc dans SELECT ces 2 colonnes tout en spécifiant qu'elles sont présentes dans la table CONTRAT (via la commande "FROM"), appelée "c" par simplification. On renomme les intitulés de colonnes via "AS".

Ici il s'agit de faire apparaître le nombre de contrats et non pas le numéro de chaque contrat : on utilise donc la fonction COUNT sur la colonne "Contrat_ID" pour faire apparaître que le nombre de contrats par catégorie de prix.

Cela suppose également d'opérer une agrégation par catégorie de prix via la commande "GROUP BY Valeur_declaree_biens".

Ensuite je souhaiterais faire apparaître les résultats de la requête dans l'ordre décroissant du nombre de contrats par catégorie. Pour ce faire, j'utilise la commande "ORDER BY" sur le nombre de contrats, en spécifiant "DESC" pour obtenir un classement dans l'ordre décroissant.

Requête 7 :

Quel est le nombre de formule “integral” sur la région Pays de la Loire ?

```
SELECT COUNT(Formule) AS 'Nombre de formules "Integral"'
FROM contrat c
JOIN region r
ON c.Code_dep_code_commune = r.Code_dep_code_commune
WHERE r.reg_nom = 'Pays de la Loire'
```

Méthodologie :

Nous avons besoin de faire apparaître le nombre de formules “integral”, variable issue de la table “CONTRAT” : on appelle donc dans SELECT la colonne “Formule” tout en spécifiant qu’elle est issue de la table CONTRAT (via la commande “FROM”), que l’on appellera “c” par simplification. La colonne “Formule” est imbriquée dans le “COUNT” car on veut avoir le nombre de formule, pas l’ensemble des valeurs que peut prendre la colonne “Formule”.

Ici il est nécessaire de faire le lien entre la table “CONTRAT” et la table “REGION” car il y a une condition sur le nom de la région (colonne “reg_nom” issue de la table “REGION”). Le lien est fait grâce à la commande “JOIN ON” comme expliqué dans les requêtes précédentes.

Comme nous ne sommes intéressés que par la région “Pays de la Loire”, nous venons ajouter la commande “WHERE r.reg_nom = 'Pays de la Loire’”.

Requête 8 :

Lister les numéros de contrats avec le type de contrat et leur formule pour les maisons du département 71.

(=> même demande que requête N°2 : voir la requête N°2 pour le détail de la méthodologie)

```
SELECT contrat_ID AS 'N° de contrat', Type_contrat AS 'Type de contrat',  
Formule  
FROM contrat c  
JOIN region r  
ON c.Code_dep_code_commune = r.Code_dep_code_commune  
WHERE c.type_local = 'Maison'  
AND r.dep_code = 71
```

Requête 9 :

Quelle est la surface moyenne des contrats à Paris ?

```
SELECT ROUND(AVG(surface)) AS 'Surface moyenne des contrats à Paris (m2)'  
FROM contrat c  
JOIN region r  
ON c.Code_dep_code_commune = r.Code_dep_code_commune  
WHERE r.com_nom_maj_court LIKE '%PARIS%'
```

Méthodologie :

Nous avons besoin uniquement de faire apparaître la surface : on appelle donc dans SELECT la colonne “surface” tout en spécifiant qu’elle est issue de la table CONTRAT via la commande “FROM”.

Comme on veut la surface moyenne, on fait appel à la fonction AVG. Il est aussi préférable de faire apparaître la surface moyenne en nombre entier (cela a plus de sens), on imbrique la commande “AVG(surface)” dans “ROUND”.

La condition imposée ici est celle de la commune de Paris (colonne "com_nom_maj_court" de la table "REGION"), on fait donc le lien entre les tables "CONTRAT" et "REGION" via la commande "JOIN ON".

Pour spécifier la condition, on utilise la commande WHERE. Ici, la commune peut prendre plusieurs valeurs différentes à cause des différents arrondissements : "PARIS", "PARIS 1", "PARIS 2" etc.

Pour que la requête retienne toutes les valeurs présentes dans la colonne "com_nom_maj_court" comportant la mention "PARIS", on ajoute un "%" avant et après la chaîne de caractères "PARIS" dans la commande "LIKE" après le "WHERE".

Requête 10 :

Classements des 10 départements où le prix moyen de la cotisation est le plus élevé.

```
SELECT r.dep_nom AS 'Département', ROUND(AVG(c.Prix_cotisation_mensuel) ,2) AS  
'Prix moyen de la cotisation'  
FROM contrat c  
JOIN region r  
ON c.Code_dep_code_commune = r.Code_dep_code_commune  
GROUP BY r.dep_nom  
ORDER BY ROUND(AVG(c.Prix_cotisation_mensuel) ,2)  
DESC  
LIMIT 10
```

Méthodologie :

Nous avons besoin de faire apparaître le nom des départements et le prix de la cotisation mensuelle : on appelle donc dans SELECT les colonnes "dep_nom" et "Prix_cotisation_mensuel". Comme le nom du département fait appel à la table "REGION" contrairement aux prix de la cotisation qui fait appel à la table "CONTRAT", on fait le lien entre les 2 tables via "JOIN ON".

Comme on veut la cotisation moyenne, on fait appel à la fonction AVG. J'imbrique la commande "AVG(surface)" dans "ROUND" spécifiant le paramètre "2" : cela permet à la requête de me lister le prix moyen arrondi à 2 décimales.

On utilise ensuite la commande “GROUP BY” pour pouvoir afficher la cotisation moyenne par département, ce qui va permettre de grouper les résultats par nom de département sinon la requête exécuterait une moyenne de surface par numéro de contrat ce qui n’aurait aucun sens.

Il est ici demandé de faire apparaître uniquement les 10 moyennes les plus élevées. Pour cela nous allons demander à lister les tarifs moyens de cotisations par ordre décroissant via la commande “ORDER BY...DESC”. Ensuite nous limitons les résultats à 10 lignes uniquement via la commande “LIMIT 10”.

Requête 11 :

Liste des communes ayant eu au moins 150 contrats.

```
SELECT r.com_nom_maj_court AS 'Commune', COUNT(c.Contrat_ID) AS 'Nombre de  
contrats > 150'  
  FROM contrat c  
  JOIN region r  
  ON c.Code_dep_code_commune = r.Code_dep_code_commune  
  GROUP BY r.com_nom_maj_court  
  HAVING COUNT(c.Contrat_ID) >=150  
  ORDER BY COUNT(c.Contrat_ID) DESC
```

Méthodologie :

Nous avons besoin de faire apparaître le nom des communes et le nombre de contrats : on appelle donc dans SELECT les colonnes “com_nom_maj_court” et “Contrat_ID”. Comme le nom de la commune fait appel à la table “REGION” contrairement aux contrats qui font appel à la table “CONTRAT”, on fait le lien entre les 2 tables via la commande “JOIN ON”.

Comme on veut le nombre de contrats et non pas la liste des N° de contrats, on vient utiliser “COUNT”.

Ensuite il faut opérer une agrégation par nom de commune via “GROUP BY”.

Comme la condition est d’afficher uniquement les communes ayant au moins 150 contrats, on vient utiliser cette fois-ci la commande “HAVING” (et non pas “WHERE”

car la condition s'opère sur une agrégation qui a déjà eu lieu). On vient préciser ">=150" pour afficher les communes ayant au moins 150 contrats.

J'ai ensuite souhaité faire apparaître les communes en fonction de l'ordre décroissant du nombre de contrats via la commande "ORDER BY...DESC".

Requête 12 :

Quel est le nombre de contrats pour chaque région ?

```
SELECT reg_nom AS 'Région', COUNT(contrat_ID) AS 'Nombre de contrats par région'
FROM contrat c
JOIN region r
ON c.Code_dep_code_commune = r.Code_dep_code_commune
GROUP BY r.reg_nom
ORDER BY COUNT(contrat_ID)
DESC
```

Méthodologie :

Nous avons besoin de faire apparaître le nom des régions et le nombre de contrats : on appelle donc dans SELECT les colonnes "reg_nom" et "Contrat_ID". Comme le nom de la région fait appel à la table "REGION" contrairement aux contrats qui font appel à la table "CONTRAT", on fait le lien entre les 2 tables via la commande "JOIN ON".

Comme on veut le nombre de contrats et non pas la liste des N° de contrats, on vient utiliser "COUNT".

Ensuite il faut opérer une agrégation par région via "GROUP BY".

J'ai ensuite souhaité faire apparaître le nombre de contrats dans l'ordre décroissant via la commande "ORDER BY...DESC".