

Complément SQL/Illustrations

HAVING Vs. WHERE

id_patient	type_analyse	valeur	date_analyse
1	Glucose	1.1	2025-03-01
1	Cholestérol	2.3	2025-03-01
2	Glucose	1.5	2025-03-02
2	Cholestérol	3.1	2025-03-02
3	Glucose	1.2	2025-03-03
3	Cholestérol	2.0	2025-03-03
4	Glucose	1.8	2025-03-04
4	Cholestérol	2.7	2025-03-04

La liste des patients dont la moyenne de glucose est supérieure à 1.3

La liste des patients dont la moyenne de glucose est supérieure à 1.3

```
SELECT id_patient, AVG(valeur) AS moyenne_glucose
FROM Analyses_Sang
WHERE type_analyse = 'Glucose'
GROUP BY id_patient
HAVING AVG(valeur) > 1.3;
```

-- Filtrer les lignes : WHERE type_analyse = 'Glucose' garde uniquement les analyses de glucose.

-- Regrouper par patient : GROUP BY id_patient regroupe les résultats par patient.

-- Calculer la moyenne : AVG(valeur) AS moyenne_glucose calcule la moyenne de glucose pour chaque patient.

-- Filtrer les groupes : HAVING AVG(valeur) > 1.3 ne garde que les patients dont la moyenne de glucose dépasse 1.3.

id_patient	moyenne_glucose
2	1.5
4	1.8

La liste des patients dont la moyenne de glucose est supérieure à 1.3

Pourquoi on ne peut pas utiliser la clause WHERE ?

```
SELECT id_patient, AVG(valeur) AS moyenne_glucose  
FROM Analyses_Sang  
WHERE type_analyse = 'Glucose' AND AVG(valeur) > 1.3  
GROUP BY id_patient;
```

Erreur SQL : AVG(valeur) ne peut pas être utilisé dans WHERE
car il est calculé après le GROUP BY

- WHERE filtre les lignes avant l'agrégation (GROUP BY).
- HAVING filtre les résultats après l'agrégation.

La liste des patients dont la moyenne de glucose est supérieure à 1.3

Solution sans HAVING ⇒ Sous requête

```
SELECT id_patient, moyenne_glucose
FROM (
    SELECT id_patient, AVG(valeur) AS moyenne_glucose
    FROM Analyses_Sang
    WHERE type_analyse = 'Glucose'
    GROUP BY id_patient
)
WHERE moyenne_glucose > 1.3;
```

Division

- Utilisée pour trouver les éléments d'une relation R qui sont liés à tous les éléments d'une relation S.
- Ce type de requête est souvent utilisé pour répondre à des questions comme :
 - Quels étudiants ont validé toutes les matières ?
 - Quels fournisseurs fournissent toutes les pièces requises pour un produit ?"

Division

Etudiants

id_etudiant	Nom
1	Alice
2	Bob
3	Charlie

Matieres

id_matiere	nom_matiere
101	Mathématiques
102	Biologie
103	Chimie

Validation

id_etudiant	id_matiere
1	101
1	102
1	103
2	101
2	103
3	101
3	102

Trouver les étudiants qui ont validé toutes les matières obligatoires

Trouver les étudiants qui ont validé toutes les matières obligatoires

```
SELECT e.id_etudiant, e.nom
FROM Etudiants e
WHERE NOT EXISTS (
    SELECT m.id_matiere
    FROM Matieres m
    WHERE NOT EXISTS (
        SELECT v.id_matiere
        FROM Validations v
        WHERE v.id_etudiant = e.id_etudiant
        AND v.id_matiere = m.id_matiere
    )
);
```

-- Vérifie s'il existe une matière que l'étudiant n'a pas validée

Division

Soient $r(R)$ et $s(S)$ deux instances de relation,
avec $S \subset R$,

$r \div s$ est la relation de schéma $Q=R-S$ définie par:

$$r \div s = \{ t_q \in \pi_Q(r) \mid \forall t_s \in s, t_q \bowtie t_s \in r \}$$

$r \div s$ regroupe tous les éléments qui dans r sont reliés à tous les éléments de s :

$r \div s$ est le plus grand ensemble q de $\pi_Q(r)$ tel que $(q \bowtie s) \subset r$

Division

Quels sont les athlètes qui ont participé à toutes les épreuves?

$r \div s$ est la relation de schéma $Q=R-S$ définie par:

$$r \div s = \{ t_q \in \pi_Q(r) \mid \forall t_s \in s, t_q \bowtie t_s \in r \}$$

r

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	400 m
Jacques	200 m

$s = \prod_{\text{Epreuve}}(r)$

Epreuve
200 m
400 m
800 m

q

Athlète
Pierre

Division

Propriété

La division s'exprime en fonction des opérateurs précédents :

$$\begin{aligned}r \div s &= \pi_Q(r) - \pi_Q((\pi_Q(r) \bowtie s) - r) \\ &= \pi_Q(r) - \pi_Q((\pi_Q(r) \times s) - r)\end{aligned}$$

Division

r

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	400 m
Jacques	200 m

$S = \prod_{\text{Epreuve}(r)}$

Epreuve
200 m
400 m
800 m

$\prod_Q(r)$

Athlète
Pierre
Paul
Jacques

$S \times \prod_Q(r)$

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	200 m
Paul	400 m
Paul	800 m
Jacques	200 m
Jacques	400 m
Jacques	800 m

Les athlètes qui ont participé à toutes les épreuves?

Division

Quels sont les athlètes qui ont participé à toutes les épreuves?

$S \times \Pi_Q(r)$

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	200 m
Paul	400 m
Paul	800 m
Jacques	200 m
Jacques	400 m

r

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	400 m
Jacques	200 m

$temp$

Athlète	Epreuve
Paul	200 m
Paul	200 m
Jacques	400 m
Jacques	800 m

$\Pi_Q(temp)$

Athlète
Paul
Jacques

Division

Quels sont les athlètes qui ont participé à toutes les épreuves?

