Agrégations

Ludovic Deneuville

Données utilisées

id_joueuse	nom	prenom	elo	ville_club
1	Guichard	Pauline	2357	Asnieres
2	Daulyte	Deimante	2371	Chartres
3	Navrotescu	Andreea	2338	Asnieres
4	Sebag	Marie	2448	Bischwiller
5	Millet	Sophie	2366	Asnieres
6	Skrichempko	Almira	2361	Bischwiller

Questions

Quelques questions que l'on peut se poser :

- Quel est le id_joueuse maximum ?
- Quel est le nombre de joueuses ?
- Quelle est la moyenne elo par club ?

Compter les lignes

```
SELECT COUNT(*)
  FROM echecs.joueuse
 WHERE elo < 2400;
```

count
5

```
SELECT COUNT(DISTINCT prenom)
  FROM echecs.joueuse;
```

count
6

Compter avec un filtre

L'opérateur `FILTER` permet de calculer une agrégation conditionnelle.

```sql
SELECT COUNT(*) FILTER (WHERE elo BETWEEN 2300 AND 2399) AS niv_fm,
       COUNT(*) FILTER (WHERE elo BETWEEN 2400 AND 2499) AS niv_im
  FROM echecs.joueuse;
```

niv_fm	niv_im
5	1

Somme

```sql
SELECT SUM(elo)
  FROM echecs.joueuse;
```

sum
14223

Stats descriptives

```sql
SELECT MAX(elo),
       MIN(elo),
       AVG(elo)
  FROM echecs.joueuse;
```

max	min	avg
2448	2338	2373.5

> Tip
>
> Vous pouvez également calculer : `VARIANCE`, `STDDEV`.

Comment afficher la joueuse avec le plus grand elo ? 😵

Sous-requête

En utilisant une **sous-requête** :

- `SELECT MAX(elo) FROM echecs.joueuse` ➡️ renvoie un nombre
- Nous pouvons ensuite tester si *elo* est égal à ce nombre

```sql
SELECT nom,
       prenom,
       elo
  FROM echecs.joueuse
 WHERE elo = (SELECT MAX(elo) FROM echecs.joueuse);
```

Regroupement

Comment obtenir la moyenne elo par club ?

```
SELECT ville_club,    -- Supposons que la table joueuse a un attribut ville_club
       AVG(elo)
  FROM echecs.joueuse
 GROUP BY ville_club;
```

ville_club	avg_elo
Asnieres	2372.4
Chartres	2371.0
Bischwiller	2404.5

Colonnes du GROUP BY

Vous pouvez grouper en utilisant plusieurs colonnes.

> ! Important
>
> Les colonnes présentes dans le *SELECT* doivent impérativement être présentes dans le *GROUP BY*.
>
> L'inverse n'est pas vrai. Vous pouvez avoir des colonnes du *GROUP BY* que vous n'affichez pas dans le *SELECT*.

Filtrer suite au GROUP BY

Comment faire pour afficher uniquement les clubs où la moyenne élo est inférieure à 2400 ?

```
SELECT ville_club,
       AVG(elo)
  FROM echecs.joueuse
 GROUP BY ville_club
HAVING AVG(elo) < 2400;
```

club	avg_elo
Asnieres	2372.4
Chartres	2371.0

Having Vs Where

- **WHERE** - Filtre Avant l'Agrégation
- **HAVING** - Filtre Après l'Agrégation