

18.1 - REQUÊTES : AGRÉGATIONS STATISTIQUES, ENSEMBLE ET SOUS-ENSEMBLE

https://www.sqlfacile.com/apprendre_bases_de_donnees/group_by_et_les_fonctions_d_agregation

SOMMAIRE :

- Les fonctions MIN et MAX
- La fonction Moyenne
- La fonction Somme
- La fonction de comptage
- La fonction de calcul de la variance
- La fonction écart type
- La commande GROUP BY
- La clause HAVING
- Mise en pratique sur la table « jeux_video »

```
1  USE SIMPLON ;
2  • SHOW TABLES ;
3  • select COLUMN_NAME from INFORMATION_SCHEMA.COLUMNS where
```

#	COLUMN_NAME
1	ID
2	nom
3	possesseur
4	console
5	prix
6	nbre_joueurs_max
7	commentaires

SYNTAXE :

SELECT fonction(colonne) FROM table ;

LES FONCTIONS MIN ET MAX :

Les fonctions MIN et MAX permettent de retrouver les valeurs minimum et maximum d'un champ.

```
select max(prix) as prix_max, min(prix) as prix_min
from jeux_video;
```

prix_min	prix_max
60	2

Attention :

```
select min(prix) as prix_min, nom, possesseur
from jeux_video;
```

prix_min	nom	possesseur
2	Super Mario Bros	Florent

prix_min	nom	possesseur
60	Super Mario Bros	Florent

Pas de correspondance entre le nom, le possesseur et le prix

LA FONCTION MOYENNE :

La fonction MOY ou AVG permet de retrouver la valeur moyenne d'une colonne.

Pour connaître le prix moyen des jeux vidéo :

```
select avg(prix) as prix_moyen  
from jeux_video;
```

prix_moyen
28.34

LA FONCTION SOMME :

La fonction SUM retour la somme des valeurs d'un champ.

Pour connaître le prix total de la valeur des jeux vidéo :

```
select sum(prix) as valeur_total  
from jeux_video;
```

valeur_total
1417

LA FONCTION DE COMPTAGE :

La fonction COUNT permet de retourner le nombre d'enregistrements d'une requête.

Pour connaître le nombre total de jeux :

```
select count(nom) as nb_jeux  
from jeux_video;
```

nb_jeux
50

LA FONCTION VAR OU VARIANCE :

Rappel :

En statistique et en théorie des probabilités, la variance est une mesure servant à caractériser la dispersion d'un échantillon ou d'une distribution.

Elle indique de quel manière la série statistique ou la variable aléatoire se disperse autour de sa moyenne.

- Une variance de 0 signifie que toutes les valeurs sont identiques.
- Une petite variance signifie que toutes les valeurs sont proches les unes des autres.
- Une variance élevée signifie que celles-ci sont très dispersées.

La fonction VAR ou VARIANCE permet de retourner la variance d'une colonne donnée et d'en déduire l'amplitude par rapport à la moyenne.

Pour connaître la variance des prix des jeux vidéos :

```
select variance(prix) as variance  
from jeux_video;
```

variance
317.42440000000005

LA FONCTION ÉCART TYPE STD OU STDDEV :

- En probabilité, l'écart type est une mesure de la dispersion d'une variable aléatoire.
- En statistique, il est une mesure de dispersion de données.
- Il est défini comme la racine carrée de la variance.

La fonction STD ou STDDEV retourne l'écart type d'une série de valeurs et permet d'en déduire leur dispersion ou leur homogénéité.

Pour connaître l'écart type du prix des jeux vidéo :

```
SELECT ROUND(STD(prix), 2) AS Ecart_type_prix FROM jeux_video ;
```

#	Ecart_type_prix
1	17.74

LA COMMANDE GROUP BY :

Les fonctions agrégatives peuvent être appliquées à l'intégralité ou à des sous-ensembles du résultat d'une requête.

Exemple :

Pour connaître le nombre de jeux par console :

```
select count(nom) as nb_jeux, console
from jeux_video;
```

nb_jeux	console
50	NES

Le résultat obtenu est faux.

Il n'y a qu'une seule valeur qui concerne toutes les consoles et non seulement la 'NES'.

Pour répondre à cette question il faut ajouter un critère de regroupement.

Pour ce faire on utilise la clause GROUP BY.

- Elle permet de constituer des sous-ensembles d'enregistrements l'or de la requête.
- Chaque attribut cité dans le GROUP BY doit être inclus dans le SELECT.
- Les résultats des clauses WHERE et FROM sont regroupés selon les critères du GROUP BY.

```
select count(nom) as nb_jeux, console
from jeux_video
group by console;
```

nb_jeux	console
3	NES
3	Megadrive
4	Nintendo 64
2	GameCube
10	Xbox
6	PC
3	SuperNES
9	PS2
2	GBA

Dans le résultat, chaque groupe est réduit à une occurrence et toutes les opérations qui suivent sont appliquées sur cette dernière.

Il est possible de spécifier plusieurs champs dans la commande GROUP BY.

```
select count(nom) as nb_jeux, console, possesseur
from jeux_video
group by console, possesseur;
```

nb_jeux	console	possesseur
1	NES	Florent
2	Megadrive	Patrick
2	Nintendo 64	Florent
1	GameCube	Michel
5	Xbox	Patrick

LA CLAUSE HAVING :

La clause HAVING est liée à la clause GROUP BY et ne peut être utilisée sans cette dernière. Elle permet de préciser un critère de restriction sur les résultats de l'agrégat.

```
select count(nom) as nb_jeux, possesseur  
from jeux_video  
group by possesseur;
```

nb_jeux	possesseur
17	Florent
13	Patrick
12	Michel
3	Mathieu
3	Sebastien
2	Corentin

```
select count(nom) as nb_jeux, possesseur  
from jeux_video  
group by possesseur  
having nb_jeux > 5;
```

nb_jeux	possesseur
17	Florent
13	Patrick
12	Michel

MISE EN PRATIQUE :

Afficher le nombre de films dans les quels à joué l'acteur « JOHNNY LOLLOBRIGIDA », regroupé par catégorie.

Ecrire la requête qui affiche les catégories dans les quels « JOHNNY LOLLOBRIGIDA » totalise plus de 3 films.

Afficher la durée moyenne d'emprunt des films par acteurs.

L'argent total dépensé au vidéos club par chaque clients, classé par ordre décroissant.