

WGBD et BD01

*SQL comme langage d'interrogation :
Agrégation et partitionnement*

EPFC-ULB

Boris Verhaegen

boris@verhaegen.me

Fonctions d'agrégation

Les **fonctions** COUNT(), SUM(), MIN(), MAX() et AVG() permettent d'**agréger** des données.

C'est à dire **retourner un résultat pour un ensemble** de données (relation ou partition).

Employee

<u>SSN</u>	FName	LName	Salary
123	John	Doe	2000
456	John	Smith	1700
789	Paul	Smith	1900

Moyenne des salaires

```
SELECT AVG(e.Salary)
FROM Employee e
```

La requête SELECT SSN, MAX(Salary) FROM Employee est illégale. Pourquoi ?

Fonctions d'agrégation

Employee

<u>SSN</u>	FName	LName	Salary
123	John	Doe	2000
456	John	Smith	1700
789	Paul	Smith	1900

Le nombre de prénoms différents

```
SELECT COUNT(DISTINCT FName)
FROM Employee e
```

Le maximum des salaires

```
SELECT MAX(Salary)
FROM Employee e
```

Le nombre d'employés

```
SELECT COUNT(*)
FROM Employee e
```

Toutes les lignes

La clause GROUP BY

La clause **GROUP BY** <attributs> permet de **partitionner** une relation.
Il y aura une partition par combinaison de valeurs des attributs.

Syntaxe :

```
SELECT A1, ..., An, AGG(An+1), ..., AGG(An)
FROM ... WHERE ...
GROUP BY A1, ..., An
HAVING condition
```

Les attributs du SELECT (A1, ..., An) doivent être des clés de groupement (car uniques par groupe).

La condition du **HAVING** porte sur les partitions, donc sur les attributs A1, ..., An et sur des agrégations.

La clause GROUP BY

```
SELECT A1, ..., An, AGG(An+1), ..., AGG(An)
FROM ... WHERE ...
GROUP BY A1, ..., An
HAVING condition
```

Ordre intuitif d'évaluation :

Evaluation du FROM... WHERE...

Partitionnement selon les attributs du GROUP BY

Application de la condition du HAVING (sur les partitions)

Evaluation du SELECT

Moyenne des salaires pour les départements de moins de 3 employés

```
SELECT DNo, AVG(Salary)
```

```
FROM Employee
```

```
GROUP BY DNo
```

```
HAVING count(*) < 3
```

Employee

<u>SSN</u>	Salary	DNo
123	2000	1
456	1700	2
789	1900	1
888	2100	2
999	3000	1

Deux partitions

<u>SSN</u>	Salary	DNo
123	2000	1
789	1900	1
999	3000	1
456	1700	2
888	2100	2

Une seule a été retenue

<u>SSN</u>	Salary	DNo
456	1700	2
888	2100	2

DNo	AVG(Salary)
2	1900