→ Summarizing Data

↳ Aggregate Functions

Aggregate functions in SQL are special functions that allow us to perform calculations on groups of data.
These functions take a set values as input and return a single value as output.

Common aggregate function include

SUM(): Calculates the sum of a set of values.
AVG(): Calculates the average value of a set of values
MIN(): Finds the smallest value in a set of values
MAX() : Finds the largest value in a set of values
COUNT(): Count the nb of raws that meet a certain condition

Syntaxe

```
SELECT  AGGREGATE FUNCTION (attributename)
FROM  table name ;
```

Examples

```
SELECT MIN (price)
FROM stocks ;
(the minimum price in the stock)
SELECT  COUNT (*)
FROM  customers ;
(the nb of customers )
```

Notes !!

- Aggregate functions only operate on non-null values

  For example, the "COUNT()" function counts the number of non-null value in a column, so it will not count null values

- To include NULL values in the calculations we can use the "COUNT(*)" function, wich count all rows in the table regardless of their values

- By default aggregate functions take duplicate values, so if we want to exclude duplicate we have to use the "distinct" keyword.

  COUNT (DISTINCT client_id)

- we can operate these functions on expressions

  SUM (invoice_total * 1.1)

↳ GROUP BY Clause

- Aggregate functions can be used with the GROUP BY clause to group data by one or more columns.

- The GROUP BY clause divides the data into groups based on the values in the specified columns.

  The aggregate functions then operate on each group separately, returning a single result for each group

Example :

```
SELECT department, COUNT(*) as count
FROM employees
GROUP BY department;
```

In this example, the query will return the nb of employees in each department

NOTE !!

When using aggregate functions with GROUP BY, columns that appear in the SELECT clause and that are not used in an aggregate function, must appear in the Group By Clause.

✗
```
SELECT A, B, SUM(C)
FROM table
GROUP BY B;
```

✓
```
SELECT A,B SUM(C)
FROM table
GROUP BY A, B;
```

↳ HAVING Clause

- The HAVING Clause is used to filter the result of an aggregate function based on a specified condition.

- The HAVING Clause is different from the WHERE Clause, wich is used to filter rows before they are grouped. The HAVING Clause is applied after the rows have been grouped and the aggregate function have been calculated

Example:

```
SELECT department AVG (salary)
FROM employees
GROUP BY department
HAVING AVG (salary) > 50000;
```

In this query, we're selecting the department and the average salary for each department and find department in wich the average (salary) is greater than 50000

NOTE!!

- When using having, we must include at least one aggregate function in the select statment, the column that we use in the HAVING clause have to be part of the select statment.

- The Order of Actions matter

```
SELECT seled_list
FROM table
WHERE search_condition
GROUP BY group-by-expression
HAVING search_condition
ORDERBY order in ASC/DESC.
```