

Aggregate and Sorting Function

Relevel
by Unacademy



What is Aggregate Function?

An aggregate function in SQL returns one value after performing a calculation on multiple values of a column.

- SQL provides many aggregate functions that include avg, count, sum, min, max, etc.
- An aggregate function ignores NULL values when it performs the calculation, except for the count function.
- We often use aggregate functions with the GROUP BY and HAVING clauses of the SELECT statement



Aggregate functions in SQL

OPERATOR	DESCRIPTION
Min	returns the smallest value of the selected column in a set of non-Null values
Max	returns the largest value of the selected column in a set of non-Null values
Count	returns the number of rows in a column that matches specified criteria
Avg	returns the average value of a numeric column in a set of non-Null values
Sum	returns the total sum of a numeric column in a set of non-Null values

Min

Question. Write a query to find minimum daily typing pages.

employee_tbl

id	name	work_date	daily_typing_pages
1	John	2007-01-24	250
2	Ram	2007-05-27	220
3	Jack	2007-05-06	170
3	Jack	2007-04-06	100
4	Jill	2007-04-06	220
5	Zara	2007-06-06	300
5	Zara	2007-02-06	350

Query

```
SELECT MIN(daily_typing_pages)
FROM employee_tbl
```

Output

MIN(daily_typing_pages)
100

Instructions for practice questions



Log into <https://mode.com/>



Create a new report



Access database tutorial.sat_scores



Practice Question

Find the minimum marks in sat_verbal in the dataset

Instructions: Go to the coding console and use MIN function on sat_verbal.

Solution

Find the minimum marks in sat_verbal in the dataset

```
SELECT  
  MIN(sat_verbal) AS min_verbal  
FROM  
  tutorial.sat_scores
```

1 rows | 8B returned in 1s

min_verbal
200

Max

Question. Write a query to retrieve least and maximum daily typing page..

employee_tbl

id	name	work_date	daily_typing_pages
1	John	2007-01-24	250
2	Ram	2007-05-27	220
3	Jack	2007-05-06	170
3	Jack	2007-04-06	100
4	Jill	2007-04-06	220
5	Zara	2007-06-06	300
5	Zara	2007-02-06	350

Query

```
SELECT MIN(daily_typing_pages) least  
MAX(daily_typing_pages) max  
FROM employee_tbl
```

Output

least	max
100	350

Practice Question

Find the maximum marks in sat_writing in the dataset

Instructions: Go to the coding console and use MAX function on sat_writing.

Solution

Find the maximum marks in sat_writing in the dataset

```
1 SELECT
2   MAX(sat_writing) AS max_writing
3 FROM
4   tutorial.sat_scores
```

✓ 1 rows | 8B returned in 847ms

	max_writing	
1	799	

Count

Question. Write a query to find number of people with name 'Zara'.

employee_tbl

id	name	work_date	daily_typing_pages
1	John	2007-01-24	250
2	Ram	2007-05-27	220
3	Jack	2007-05-06	170
3	Jack	2007-04-06	100
4	Jill	2007-04-06	220
5	Zara	2007-06-06	300
5	Zara	2007-02-06	350

Query

```
SELECT COUNT(*)  
FROM employee_tbl WHERE  
name="Zara";
```

Output

COUNT(*)
2

Practice Question

Find the count of students in the dataset

Instructions: Go to the coding console and write code for counting student id.

Solution

Find the count of students in the dataset

```
SELECT  
  COUNT(student_id) AS total_students  
FROM  
  tutorial.sat_scores
```

1 rows | 8B returned in 911ms

total_students
135

Avg

Question. Write a query to find average typing papers..

employee_tbl

id	name	work_date	daily_typing_pages
1	John	2007-01-24	250
2	Ram	2007-05-27	220
3	Jack	2007-05-06	170
3	Jack	2007-04-06	100
4	Jill	2007-04-06	220
5	Zara	2007-06-06	300
5	Zara	2007-02-06	350

Query

```
SELECT AVG(daily_typing_pages)
FROM employee_tbl
```

Output

AVG(daily_typing_pages)
230.0000

Practice Question

Find the average marks in sat_math in the dataset

Instructions: Go to the coding console and use AVG function on sat_math.

Solution

Find the average marks in sat_math in the dataset

```
SELECT
  AVG(sat_math) AS avg_math
FROM
  tutorial.sat_scores
```

1 rows | 8B returned in 443ms

avg_math
517.4148

Sum

Question. Write a query to find total typing pages from 'employee_tbl'.

employee_tbl

id	name	work_date	daily_typing_pages
1	John	2007-01-24	250
2	Ram	2007-05-27	220
3	Jack	2007-05-06	170
3	Jack	2007-04-06	100
4	Jill	2007-04-06	220
5	Zara	2007-06-06	300
5	Zara	2007-02-06	350

Query

SELECT SUM(daily_typing_pages) FROM
employee_tbl

Output

SUM(daily_typing_pages)
1610

Practice Question

Find the sum of hrs_studied in the dataset

Instructions: Go to the coding console and write code to sum hrs_studied.

Solution

Find the sum of hrs_studied in the dataset

```
SELECT  
  SUM(hrs_studied) AS total_hrs_studied  
FROM  
  tutorial.sat_scores
```

1 rows | 8B returned in 412ms

total_hrs_studied
12870

GROUP BY CLAUSE

The GROUP BY clause is a SQL command that is used to **group rows that have the same values**.

- It is used to summarize data from the database.
- The queries that contain the GROUP BY clause are called grouped queries and only return a single row for every grouped item.
- GROUP BY clause applied on column(s) can be used to get unique records for those columns



GROUP BY IN A SINGLE COLUMN

Question. Write a query to classify gender.

gender
Female
Female
Male
Female
Male
Male
Male
Male
Male

Query

```
SELECT gender FROM members  
GROUP BY gender
```

gender
Female
Male

Practice Question

Find the list of unique schools in the data using Group by clause

Instructions: Go to the coding console and write code for grouping by column school

Solution

Find the list of unique school in the data using Group by clause

```
SELECT
  school
FROM
  tutorial.sat_scores
GROUP BY
  school
```

3 rows | 37B returned in 492ms

school	
Petersville HS	
St. John's	
Washington ...	

GROUP BY IN MULTIPLE COLUMN

Query

```
SELECT  
category_id, year_released  
FROM members  
GROUP BY  
category_id, year_released
```

category_id	year_released
1	2011
2	2008
NULL	2008
NULL	2010
8	2007
6	2007
6	2007
8	2005
NULL	2012
7	1920
8	NULL
8	1920

category_id	year_released
NULL	2008
NULL	2010
NULL	2012
1	2011
2	2008
6	2007
7	1920
8	1920
8	2005
8	2007

Practice Question

Find the list of unique school and teachers in the data using Group by clause

Instructions: Go to the coding console and write code for grouping by columns school and teacher

Solution

Find the list of unique school and teachers in the data using Group by clause

```
SELECT
  school,
  teacher
FROM
  tutorial.sat_scores
GROUP BY
  school,
  teacher
```

3 rows | 152B returned in 479ms

school	teacher	
Petersville HS	Perry	
Washington ...	Frederickson	
St. John's	Williams	
Petersville HS	Davis	
St. John's	Rajaram	
Washington ...	Spellman	
Petersville HS	Brown	
St. John's	Tran	

GROUP BY CLAUSE WITH AGGREGATE FUNCTION

SQL Aggregate functions aggregate data across the entire column/dataset. However, if we want to aggregate the data at a certain part of table or at certain unique values of a column, it is achieved by using aggregate function with Group By clause.

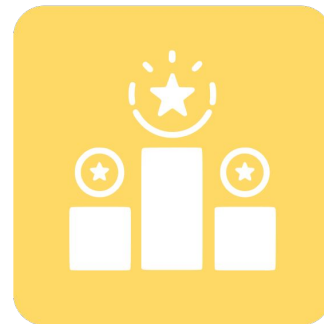
Typical Structure of a Aggregate function with Group by

SELECT <column, ...>, group/aggregate function(column)

FROM <table>

WHERE <condition>

[GROUP BY <column>]



GROUP BY CLAUSE WITH SUM FUNCTION

Question. Write a query to find unique names and total salary for each respectively from customers table.

CUSTOMERS

ID	NAME	AGE	ADDRESS	SALARY
1	Ramesh	32	Ahmedabad	2000.00
2	Ramesh	25	Delhi	1500.00
3	kaushik	23	Kota	2000.00
4	kaushik	25	Mumbai	6500.00
5	Hardik	27	Bhopal	8500.00
6	Komal	22	MP	4500.00
7	Muffy	24	Indore	10000.00

Query

```
SELECT NAME, SUM(SALARY) FROM  
CUSTOMERS  
GROUP BY NAME
```

Output

NAME	SUM(SALARY)
Hardik	8500.00
kaushik	8500.00
Komal	4500.00
Muffy	10000.00
Ramesh	3500.00

Practice Question

Find the total number of hours studied per school

Instructions: Go to the coding console and use sum function along with group by clause

Solution

Find the total number of hours studied per school

```
1 SELECT
2     school,
3     sum(hrs_studied) AS total_hrs_studied
4 FROM
5     tutorial.sat_scores
6 GROUP BY
7     school
```

✓ 3 rows | 61B returned in 971ms

	school	total_hrs_studied	
1	Petersville HS	4884	
2	St. John's	4791	
3	Washington ...	3195	

GROUP BY CLAUSE WITH COUNT FUNCTION

Question. Write a query to find number of unique locations from 'sales' table.

Sales

product	location	price	sold_at
Coffee	HQ	2	2020-09-01 09:42:33.085995
Coffee	HQ	2	2020-09-01 08:42:33.085995
Bagel	Downtown	3	2020-09-01 07:42:33.085995
Coffee	Downtown	2	2020-08-31 09:42:33.085995
Bagel	HQ	2	2020-08-30 09:42:33.085995
Bagel	1st Street	3	2020-08-30 08:42:33.085995
Coffee	1st Street	2	2020-08-29 09:42:33.085995
Bagel	HQ	3	2020-08-29 08:42:33.085995

Query

```
SELECT  
location, COUNT(*) AS number_of_sales  
FROM sales GROUP BY location
```

Output

location	number_of_sales
1st Street	2
HQ	4
Downtown	2

Practice Question

Find the number of students in each school

Instructions: Go to the coding console and use count function along with group by clause

Solution

Find the number of students in each school

```
1 SELECT
2   school,
3   COUNT(student_id) AS num_students
4 FROM
5   tutorial.sat_scores
6 GROUP BY
7   school
```

✓ 3 rows | 61B returned in 612ms

	school	num_students
1	Petersville HS	53
2	St. John's	51
3	Washington ...	31

FILTERING DATA WITH AGGREGATE FUNCTION

As we learnt about grouping data within rows, we might be tempted to filter certain groups from our data.

Sales

product	location	price	sold_at
Coffee	HQ	2	2020-09-01 09:42:33.085995
Coffee	HQ	2	2020-09-01 08:42:33.085995
Bagel	Downtown	3	2020-09-01 07:42:33.085995
Coffee	Downtown	2	2020-08-31 09:42:33.085995
Bagel	HQ	2	2020-08-30 09:42:33.085995
Bagel	1st Street	3	2020-08-30 08:42:33.085995
Coffee	1st Street	2	2020-08-29 09:42:33.085995
Bagel	HQ	3	2020-08-29 08:42:33.085995

Query

```
SELECT  
sold_at, COUNT(*) AS sales_per_day  
FROM sales  
WHERE COUNT(*) > 1 -- filter the groups?  
GROUP BY sold_at
```

Output

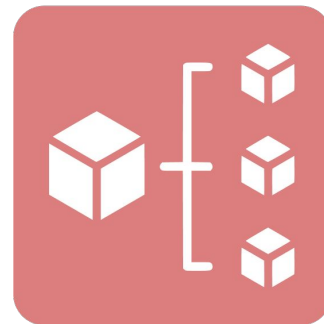
Error: aggregate functions are not allowed in group clause

Why is there an error?

Aggregate functions are not allowed in the WHERE clause because the WHERE clause is evaluated before the GROUP BY clause—there aren't any groups yet to perform calculations on.

But, there is a type of clause that allows us to filter, perform aggregations, and it is evaluated after the GROUP BY clause: **the HAVING clause**.

The HAVING clause is like a WHERE clause for your groups.



Practice Question

Find the average marks scored in sat_writing per teacher for the school 'Petersville HS'.

Instructions: Go to the coding console and use average function along with group by clause

Solution

Find the average marks scored in sat_writing per teacher for the school 'Petersville HS'.

```
1 SELECT
2   teacher,
3   AVG(sat_writing) AS avg_sat_writing
4 FROM
5   tutorial.sat_scores
6 WHERE
7   school = 'Petersville HS'
8 GROUP BY
9   teacher
```

3 rows | 39B returned in 501ms

	teacher	avg_sat_writing
1	Brown	480.3529
2	Davis	530.1538
3	Perry	572.6957

Practice Question

Find the maximum marks scored in sat_math per teacher for the school Washington HS'.

Instructions: Go to the coding console and use maximum function along with group by clause

Solution

Find the maximum marks scored in sat_math per teacher for the school 'Washington HS'.

```
1 SELECT teacher,  
2 MAX(sat_math) as max_sat_math  
3 from tutorial.sat_scores  
4 where school = 'Washington HS'  
5 group by teacher
```

✓ 2 rows | 36B returned in 1s

	teacher	max_sat_math
1	Frederickson	771
2	Spellman	796

Having Clause

As we learnt about grouping data within rows, we might be tempted to filter certain groups from our data.

Sales

product	location	price	sold_at
Coffee	HQ	2	2020-09-01 09:42:33.085995
Coffee	HQ	2	2020-09-01 08:42:33.085995
Bagel	Downtown	3	2020-09-01 07:42:33.085995
Coffee	Downtown	2	2020-08-31 09:42:33.085995
Bagel	HQ	2	2020-08-30 09:42:33.085995
Bagel	1st Street	3	2020-08-30 08:42:33.085995
Coffee	1st Street	2	2020-08-29 09:42:33.085995
Bagel	HQ	3	2020-08-29 08:42:33.085995

Query

```
SELECT
  sold_at AS date,
  COUNT(*) AS sales_per_day
FROM sales
GROUP BY sold_at
HAVING COUNT(*) > 1
```

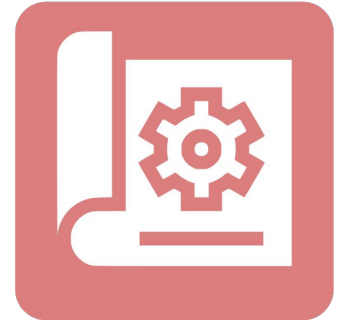
Output

date	sales_per_day
2020-09-01	3
2020-08-29	2
2020-08-30	2

ORDER BY

The ORDER BY statement in SQL is used to sort the fetched data in either ascending or descending according to one or more columns.

- By default ORDER BY sorts the data in ascending order.
- We can use the keyword DESC to sort the data in descending order and the keyword ASC to sort in ascending order.



ORDER BY

Sales

id	name	age	instrument	city
1	Dave	19	trumpet	New York City
2	Jess	46	flute	Los Angeles
3	Jenny	23	trombone	Chicago
4	Oscar	78	drums	London
5	Mike	38	violin	Paris
6	Sally	27	harp	Cape Town
7	Tom	42	tuba	Brisbane

Query

Select * FROM musicians
Order By name

Output

id	name	age	instrument	city
1	Dave	19	trumpet	New York City
3	Jenny	23	trombone	Chicago
2	Jess	46	flute	Los Angeles
5	Mike	38	violin	Paris
4	Oscar	78	drums	London
6	Sally	27	harp	Cape Town
7	Tom	42	tuba	Brisbane

Practice Question

Find the list of unique school and teachers in the data using Group by clause and sort it by ascending order for school and descending order for teacher name

Instructions: Go to the coding console and write code for grouping by columns school and teacher and use order by clause on school and teacher column

Solution

Find the list of unique school and teachers in the data using Group by clause and sort it by ascending order for school and descending order for teacher name

```
1 SELECT
2   school,
3   teacher
4 FROM
5   tutorial.sat_scores
6 GROUP BY
7   school,
8   teacher
9 ORDER BY
10  school,
11  teacher DESC
```

8 rows | 152B returned in 585ms

	school	teacher
1	Petersville HS	Perry
2	Petersville HS	Davis
3	Petersville HS	Brown
4	St. John's	Williams
5	St. John's	Tran
6	St. John's	Rajaram
7	Washington ...	Spellman
8	Washington ...	Frederickson

Practice Question

Find the list of all the teachers along with the minimum marks scored by their students in `sat_verbal`. Only consider those teachers where minimum marks is more than 220

Instructions: Go to the coding console and use `min` function along with `group by` and `having` clause

Solution

Find the list of all the teachers along with the minimum marks scored by their students in sat_verbal. Only consider those teachers where minimum marks is more than 220. Sort the output in ascending order

```
1 SELECT teacher,  
2 MIN(sat_verbal) as min_sat_verbal  
3 from tutorial.sat_scores  
4 group by teacher  
5 HAVING MIN(sat_verbal) >220  
6 order by min_sat_verbal
```

	teacher	min_sat_verbal
1	Frederickson	230
2	Tran	239
3	Spellman	286
4	Davis	330

TOP

It specifies the number of records to return from top and usually used to get a quick view of the schema (rows/columns) of the database when you start looking at a database for the first time.

Question. Write a query to find minimum daily typing pages.

CUSTOMERS

ID	NAME	AGE	ADDRESS	SALARY
1	Ramesh	32	Ahmedabad	2000.00
2	Ramesh	25	Delhi	1500.00
3	kaushik	23	Kota	2000.00
4	kaushik	25	Mumbai	6500.00
5	Hardik	27	Bhopal	8500.00
6	Komal	22	MP	4500.00
7	Muffy	24	Indore	10000.00

Query

SELECT TOP 3 * FROM CUSTOMERS

Output

ID	NAME	AGE	ADDRESS	SALARY
1	Ramesh	32	Ahmedabad	2000.00
2	Khilan	25	Delhi	1500.00
3	kaushik	23	Kota	2000.00

LAST

It returns the last value of the selected column(s).

Orders

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Query

```
SELECT LAST(OrderPrice) AS  
LastOrderPrice FROM orders
```

Output

LastOrderPrice
100

Limit

The LIMIT clause is used to set an upper limit on the number of tuples/records returned by SQL

CUSTOMERS

ID	NAME	AGE	ADDRESS	SALARY
1	Ramesh	32	Ahmedabad	2000.00
2	Khilan	25	Delhi	1500.00
3	kaushik	23	Kota	2000.00
4	Chaitali	25	Mumbai	6500.00
5	Hardik	27	Bhopal	8500.00
6	Komal	22	MP	4500.00
7	Muffy	24	Indore	10000.00

Query

SELECT * FROM CUSTOMERS LIMIT 3

Output

ID	NAME	AGE	ADDRESS	SALARY
1	Ramesh	32	Ahmedabad	2000.00
2	Khilan	25	Delhi	1500.00
3	kaushik	23	Kota	2000.00

Practice Question

Find the first 5 rows of the data

Instructions: Go to the coding console and use limit function or Top function

Solution

Find the first 5 rows of the data

```
1 SELECT
2   *
3 FROM
4   tutorial.sat_scores
5 LIMIT 5
```

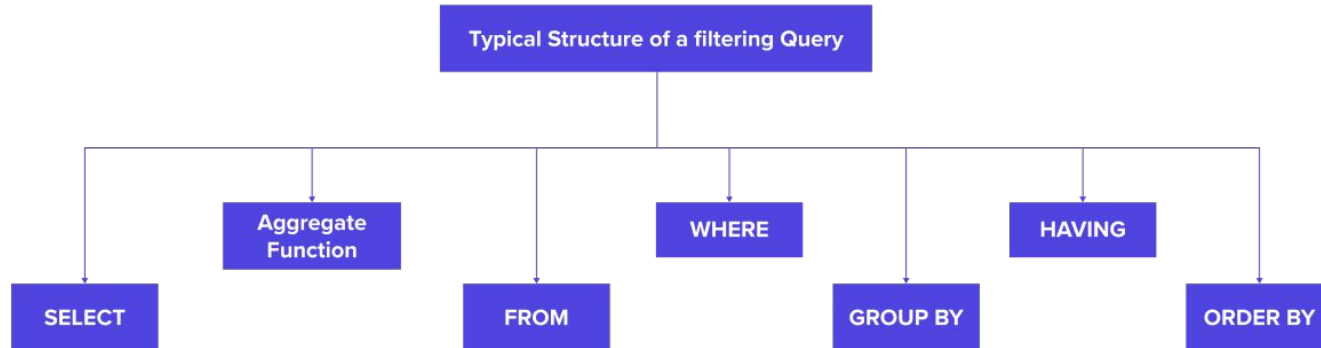
✓ 5 rows | 345B returned in 537ms



	school	teacher	student_id	sat_writing	sat_verbal	sat_math	hrs_studied	id
1	Washington ...	Frederickson	1	583	307	528	190	1
2	Washington ...	Frederickson	2	401	791	248	149	2
3	Washington ...	Frederickson	3	523	445	756	166	3
4	Washington ...	Frederickson	4	306	269	327	137	4
5	Washington ...	Frederickson	5	300	539	743	115	5

Standard Aggregate Query structure

Filtering data is a technique to extract the desired data from the database. It is achieved primarily via using 'Where clause' along with SQL logical operators.



Conclusion

In the next class we will study:



Problem solving in SQL Operators, filtering Data, aggregate and sorting functions