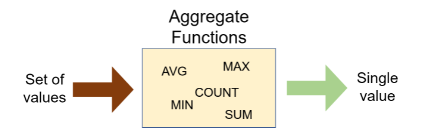
### AGGREGATE FUNCTIONs AND JOINs

**What is an aggregate function?**

In this section, we will introduce you to *aggregate functions*. Aggregate functions are functions that take a collection of values as input and return a single value. SQL has five built-in aggregate functions:

* COUNT (Count)
* MIN (Minimum)
* MAX (Maximum)
* SUM (Total)
* AVG (Average)



SUM and AVG functions must take only numeric values as input. However, the other functions (MIN, MAX, COUNT) may take non-numeric values (strings, date, etc.) as input other than numeric values. Before we dive into the aggregate functions, we should learn an important concept which is NULL.

NULL means no data and is a special value. It shows us that a piece of information is unknown or missing or not applicable. For instance, let's assume that we have a song database and we don't know the writers of some of them. We don't have that information. So, to store these unknown songwriters in a database table, we must use NULL. If a table has NULL value, then it will be displayed as NULL.

**💡Tips:**

* NULL value represents the unknown value or missing value or not applicable.
* NULL is not equal to zero or empty string.
* NULL is not equal to itself.

It's important to understand what nulls mean and to know how to work with them. They are often ignored in our aggregation functions.

Alright, let's start with COUNT function.

Q: What is an aggregate function?  
A: An aggregate function performs operations on a collection of values to return a single scalar value. Aggregate functions are often used with the GROUP BY and HAVING clauses of the SELECT statement.

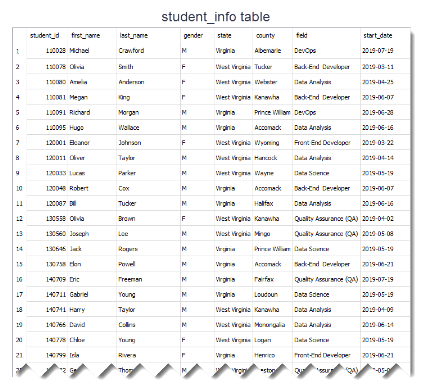
**COUNT**

### Introduction

We use COUNT function to count the numbers of records (a.k.a row) in a table. Here is the syntax of the COUNT function:

SELECT COUNT(column\_name)

FROM table\_name;

Use COUNT keyword immediately after SELECT keyword. Then specify the column on which you want to operate.   
Suppose that we want to count how many students enrolled in the courses.   
  


Let's write the query as follows:

query :

SELECT COUNT(first\_name)

FROM student\_info;

The output of the query:  
  
output :

COUNT(first\_name)

-----------------

32

32 students enrolled in the courses.   
This query counts all the first names in the student\_info table. You would reach the same result if you used COUNT(last\_name). There is another special character returning the number of rows in a table. That is \* character.  Use it inside the COUNT function as COUNT(\*).   
  
If you notice that the header of the output query is displayed as COUNT(first\_name). However, we can customize the header using AS keyword. AS is used to rename a column or table with an alias. This is the syntax:  
 SELECT column\_name AS alias\_name  
  
Let's rewrite the query above:  
  
query :

SELECT COUNT(first\_name) AS count\_of\_students

FROM student\_info;

output :

count\_of\_students

-----------------

32

As you see that what you wrote after AS keyword is displayed as a column header.   
  
An important point for COUNT(\*) function is that the result table includes NULL and duplicate values. If you want the number of non-null values, use the syntax COUNT(column\_name).

### COUNT DISTINCT

In some cases, we may want unique values. For instance, let's find how many different fields there are in the student\_info table.

### 

First, write this query:

query :

SELECT COUNT(field) AS count\_of\_field

FROM student\_info;

The output of the query:

output :

count\_of\_field

--------------

32

As you see that there are 32 fields. But, is that the correct number? Absolutely no. There are duplicate values. For instance, there is more than one student who enrolled in the course in the field of *Data Science*. COUNT(field) counts the duplicate rows as separate rows. So, we need to use COUNT(DISTINCT column\_name) function here.

query:

SELECT COUNT(DISTINCT field) AS count\_of\_field

FROM student\_info;

output:

count\_of\_field

--------------

6

Actually there are 6 unique or distinct fields in which students may enroll. COUNT(DISTINCT field) returns the number of unique and non-null values in column *field*.

We can also combine COUNT(DISTINCT) or COUNT() functions with WHERE clause.

For instance, display the number of students from Virginia.

query :

SELECT COUNT(\*) AS count\_of\_students

FROM student\_info

WHERE state = 'Virginia';

output :

count\_of\_students

-----------------

14

There are 14 students from Virginia state.