**Aggregate Functions**

SQL can not only retrieve data but it can calculate statistics about the data in the tables. For example:

SELECT COUNT(\*)

FROM mccMatches;

The aggregation function COUNT counts the number of rows, and (\*) signifies all rows. So this query will count the number of matches that have been played.

If you specify a column instead of \* it will count the number of data values in the column, ignoring any NULL values as these are not true values. For example:

SELECT COUNT(age)

FROM mccPlayer;

This will tell us how many players there are, whose ages are known.

Using DISTINCT removes duplicate values before the aggregation function is applied.

SELECT COUNT(DISTINCT g\_town)

FROM mccGround;

This will tell us how many different towns have cricket grounds.

Other aggregation functions are SUM (the sum), AVG (the average), MIN (the minimum value) and MAX (the maximum value). These functions must specify a numeric attribute in the aggregation clause. For example:

SELECT MAX(batting\_score)

FROM mccMatch\_performance;

This will tell us the highest score achieved.

However, the following query could produce misleading results, and in some SQL implementations will produce an error.

SELECT registration\_number, MAX(batting\_score)

FROM mccMatch\_performance;

This is because an aggregate function returns one value for the table as a whole, and you cannot also specify attributes that could have more than one value, for example two or more players, with different registration\_numbers, could have the same maximum score. SQL doesn't bother to test whether this is true or not, it just returns the first one (or an error).

You can restrict the rows that are considered by the aggregation function in the WHERE clause. For example, what is the highest batting score made since 2012?

SELECT MAX(batting\_score)

FROM mccMatch\_performance

WHERE match\_date > 120100;

The condition, match\_date > 120100, is applied first and the aggregate function is applied to the result.

Some queries for you to try.

1. What is the average batting score of all match performances?
2. What is the age of the youngest player?
3. How many matches have been played?
4. How many different teams have MCC played against?
5. What is the number and average age of people whose street name ends on “Street”?

A screen shot of a computer

Description automatically generated