What is SQL Query?

Imagine you have DB table with data in it and you got a requirement about your data. This requires data language which is SQL and which gives results.

SQL Clauses

SELECT Clause –

‘SELECT \*’ selects everything from the data.

‘FROM’ where to find the data.

SQL syntax - “SELECT \* FROM table”

Example of using SELECT statement (basic):

*SELECT \* FROM Data1*

*SELECT \*FROM Data2*

SQL Choose Columns -

‘SELECT Col1, Col2’ selects only columns which are mentioned.

‘FROM’ where to find the data.

SQL syntax – “SELECT Col1, Col2 FROM table”

Example:

***Retrieve only District, Growth and Literacy -*** *‘SELECT District, Growth, Literacy FROM Data1’*

WHERE Clause –

Filters your data based on the condition.

‘SELECT \*’ selects everything in the data.

‘FROM table’ where to find the data.

‘WHERE Condition’ conditions to be met in the query.

SQL syntax – “SELECT \* FROM table WHERE Condition”

Example:

*‘SELECT \* FROM Data1 WHERE Sex\_Ratio != 900’*

*‘SELECT \* FROM Data1 WHERE State = 'Maharashtra'*

*‘SELECT District, State FROM Data1 WHERE State = 'Maharashtra'*

Sort the data using ORDER BY -

You can sort the data in ascending (ASC) or descending (DSC).

SELECT \*

FROM table

ORDER BY column\_name ASC/DSC

Example:

*‘SELECT \* FROM Data1 ORDER BY Literacy DESC’*

*‘SELECT \* FROM Data1 ORDER BY Literacy ASC’*

Nested ORDER BY -

Only beneficial if the data has repetition value

*‘SELECT \* FROM Data1 ORDER BY District DESC, Literacy ASC’*

GROUP BY Clause -

Combines the rows with the same values. Aggregates the column by another column. If we add a new column name which is not been used for GROUP BY this will run into an error.

SELECT column\_name1, SUM(column\_name2)

FROM table

GROUP BY column\_name1

Example:

*‘SELECT State, SUM(Sex\_Ratio) AS Total\_Score FROM Data1 GROUP BY State’*

‘*SELECT State, SUM(Sex\_Ratio) AS Total\_Score, COUNT(State) AS Total\_appearance FROM Data1 GROUP BY State’*

HAVING Clause -

Filter data after aggregation. Used only after GROUP BY

SELECT column\_name1, SUM(column\_name2)

FROM table

GROUP BY column\_name1

HAVING SUM(column\_name2)>900

Example:

*‘SELECT State, AVG(Sex\_Ratio) FROM Data1 WHERE Sex\_Ratio != 0 GROUP BY State HAVING AVG(Sex\_Ratio) > 900’*

WHERE comes always before GROUP BY and HAVING comes after.

DISTINCT Clause –

Remove duplicates in your data.

SELECT DISTINCT

Column\_name

FROM table

Example:

*‘SELECT DISTINCT State FROM Data1’*

TOP –

Limit your data.

SELECT TOP 3 \*

FROM table

Example:

*‘SELECT TOP 3 \* FROM Data1’*

*‘SELECT TOP 3\* FROM Data1 ORDER BY Sex\_Ratio DESC’*

Executing order vs Coding order

Coding Order:

SELECT DISTINCT TOP 2

COL1,

COL2

FROM Table

WHERE COL=10

GROUP BY COL1

HAVING SUM(COL2)>30

ORDER BY COL1 ASC

Filter columns – SELECT

Filter Duplicates – DISTINCT

Filter Result Rows – TOP

Filter Rows Before Aggregations – WHERE

Filter Rows after aggregation – HAVING

Execution Order:

FROM

WHERE

GROUP BY

HAVING

SELECT DISTINCT

ORDER BY

TOP