**HiveQL Manipulation**

Hive achieves this task by converting queries written in HiveQL into MapReduce tasks that are then run across the Hadoop cluster to fetch the desired results. As such, Hive is best suited for batch processing large amounts of data (such as in data warehousing) but is not ideally suitable as a routine transactional database because of its slow response times (it needs to fetch data from across a cluster).

**HiveQL - Select-Where**

The Hive Query Language (HiveQL) is a query language for Hive to process and analyze structured data in a Metastore.

SELECT statement is used to retrieve the data from a table. WHERE clause works similar to a condition. It filters the data using the condition and gives you a finite result. The built-in operators and functions generate an expression, which fulfils the condition.

**Syntax: -**

SELECT [ALL | DISTINCT] select\_expr, select\_expr, ...

FROM table\_reference

[WHERE where\_condition]

[GROUP BY col\_list]

[HAVING having\_condition]

[CLUSTER BY col\_list | [DISTRIBUTE BY col\_list] [SORT BY col\_list]]

[LIMIT number];

**HiveQL –ORDER BY Clause**

The ORDER BY clause is used to retrieve the details based on one column and sort the result set by ascending or descending order.

**Syntax: -**

his chapter explains how to use the ORDER BY clause in a SELECT statement. The ORDER BY clause is used to retrieve the details based on one column and sort the result set by ascending or descending order.

Syntax

Given below is the syntax of the ORDER BY clause:

SELECT [ALL | DISTINCT] select\_expr, select\_expr, ...

FROM table\_reference

[WHERE where\_condition]

[GROUP BY col\_list]

[HAVING having\_condition]

[ORDER BY col\_list]]

[LIMIT number];

**HiveQL – GROUP BY Clause**

The GROUP BY clause is used to group all the records in a result set using a particular collection column. It is used to query a group of records.

The GROUP BY clause is used to group all the records in a result set using a particular collection column. It is used to query a group of records.

**Syntax: -**

SELECT [ALL | DISTINCT] select\_expr, select\_expr, ...

FROM table\_reference

[WHERE where\_condition]

[GROUP BY col\_list]

[HAVING having\_condition]

[ORDER BY col\_list]]

[LIMIT number];

**HiveQL – Select – Join**

JOIN is a clause that is used for combining specific fields from two tables by using values common to each one. It is used to combine records from two or more tables in the database. It is more or less similar to SQL JOIN.

**Syntax**

join\_table:

table\_reference JOIN table\_factor [join\_condition]

| table\_reference {LEFT|RIGHT|FULL} [OUTER] JOIN table\_reference

join\_condition

| table\_reference LEFT SEMI JOIN table\_reference join\_condition

| table\_reference CROSS JOIN table\_reference [join\_condition]