**HiveQL**

HiveQL (Hive Query Language) is a query language similar to SQL for querying and managing data in the Apache Hive environment, which runs on top of Hadoop.

**HiveQL DDL (Data Definition Language)**

DDL commands are used to define, alter, and drop tables, databases, and other objects in Hive.

**Common DDL Commands:**

1. **CREATE**
2. **DROP**
3. **ALTER**
4. **TRUNCATE**

**Examples of HiveQL DDL**

**1. CREATE DATABASE**

Used to create a new database.

CREATE DATABASE IF NOT EXISTS example\_db;

**2. USE DATABASE**

Switch to a specific database.

USE example\_db;

**3. CREATE TABLE**

Create a table in Hive.

CREATE TABLE IF NOT EXISTS employee (

id INT,

name STRING,

salary FLOAT,

department STRING

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE;

**4. SHOW TABLES**

Display all tables in the current database.

SHOW TABLES;

**5. DESCRIBE TABLE**

Get table schema details.

DESCRIBE employee;

**6. ALTER TABLE**

Modify a table's schema or properties.  
**a. Add a column:**

ALTER TABLE employee ADD COLUMNS (age INT);

**b. Rename a table:**

ALTER TABLE employee RENAME TO employee\_details;

**7. DROP TABLE**

Delete a table.

DROP TABLE IF EXISTS employee\_details;

**8. TRUNCATE TABLE**

Remove all rows from a table but keep the schema intact.

TRUNCATE TABLE employee;

**HiveQL DML (Data Manipulation Language)**

DML commands are used to manipulate data stored in Hive tables.

**Common DML Commands:**

1. **SELECT**
2. **INSERT**
3. **LOAD**
4. **EXPORT / IMPORT**

**Examples of HiveQL DML**

**1. LOAD DATA**

Load data into a table from HDFS or a local file.  
**From local file system:**

LOAD DATA LOCAL INPATH '/path/to/local/employee\_data.csv' INTO TABLE employee;

**From HDFS:**

LOAD DATA INPATH '/hdfs/path/employee\_data.csv' INTO TABLE employee;

**2. INSERT INTO TABLE**

Insert data into an existing table.

INSERT INTO TABLE employee

VALUES (1, 'John', 50000, 'HR'),

(2, 'Jane', 60000, 'Finance');

**3. INSERT OVERWRITE TABLE**

Overwrite existing data in a table.

INSERT OVERWRITE TABLE employee

SELECT \* FROM temp\_employee WHERE salary > 40000;

**4. SELECT Query**

Retrieve data from a table.

SELECT name, salary FROM employee WHERE department = 'HR';

**5. EXPORT TABLE**

Export table data to HDFS.

EXPORT TABLE employee TO '/hdfs/path/employee\_backup';

**6. IMPORT TABLE**

Import table data from an export location.

IMPORT TABLE employee FROM '/hdfs/path/employee\_backup';

**Summary Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Command** | **Purpose** | **Example** |
| **DDL** | CREATE TABLE | Create a table | CREATE TABLE employee (...) |
|  | ALTER TABLE | Modify table structure | ALTER TABLE employee ADD COLUMNS (age INT) |
|  | DROP TABLE | Remove a table | DROP TABLE employee |
|  | TRUNCATE TABLE | Remove all rows | TRUNCATE TABLE employee |
| **DML** | LOAD DATA | Load data into a table | LOAD DATA LOCAL INPATH 'file.csv' INTO TABLE employee |
|  | INSERT INTO | Insert data | INSERT INTO employee VALUES (1, 'John', ...) |
|  | INSERT OVERWRITE | Replace table data | INSERT OVERWRITE TABLE employee SELECT ... |
|  | SELECT | Query data | SELECT \* FROM employee WHERE salary > 50000; |
|  | EXPORT / IMPORT | Move data between tables & HDFS | EXPORT TABLE employee TO 'hdfs/path' |