

02 Hive Basic Commands-KirkYagami

1. Create Database

Creates a logical namespace for organizing tables.

This database is part of the Hive metastore, which is a central repository storing metadata about Hive tables, such as their schema and location.

```
CREATE DATABASE D1;  
CREATE DATABASE IF NOT EXISTS D1;
```

```
DESCRIBE DATABASE D1;
```

```
+-----+-----+-----+-----+  
-----+-----+-----+-----+  
| db_name | comment | location |  
managedlocation | owner_name | owner_type | connector_name |  
remote_dbname |  
+-----+-----+-----+-----+  
-----+-----+-----+-----+  
| d1 | | file:/opt/hive/data/warehouse/d1.db |  
file:/opt/hive/data/warehouse/d1.db | hive | USER |  
| |  
+-----+-----+-----+-----+  
-----+-----+-----+-----+  
1 row selected (0.049 seconds)
```

```
CREATE DATABASE IF NOT EXISTS D2 COMMENT " My Second db";
```

```
DESCRIBE DATABASE EXTENDED D2;
```

```
+-----+-----+-----+-----+  
-----+-----+-----+-----+  
---+-----+  
| db_name | comment | location |  
managedlocation | owner_name | owner_type | connector_name |  
remote_dbname | parameters |  
+-----+-----+-----+-----+  
-----+-----+-----+-----+  
---+-----+  
| d2 | My Second db | file:/opt/hive/data/warehouse/d2.db |
```

```
file:/opt/hive/data/warehouse/d2.db | hive | USER |
| | |
+-----+-----+-----+-----+
-----+-----+-----+-----+
---+-----+
1 row selected (0.033 seconds)
```

Database Properties

```
CREATE DATABASE IF NOT EXISTS D3 WITH DBPROPERTIES ('CREATOR'='NIKHIL',
'DATE'='2024/27/05');
```

```
DESCRIBE DATABASE EXTENDED D3;
+-----+-----+-----+-----+
-----+-----+-----+-----+
-----+
| db_name | comment | location |
managedlocation | owner_name | owner_type | connector_name |
remote_dbname | parameters |
+-----+-----+-----+-----+
-----+-----+-----+-----+
-----+
| d3 | | file:/opt/hive/data/warehouse/d3.db |
file:/opt/hive/data/warehouse/d3.db | hive | USER |
| | {CREATOR=NIKHIL, DATE=2024/27/05} |
+-----+-----+-----+-----+
-----+-----+-----+-----+
-----+
1 row selected (0.033 seconds)
```

```
SHOW DATABASES;
+-----+
| database_name |
+-----+
| d1 |
| d2 |
| d3 |
| default |
+-----+
4 rows selected (0.028 seconds)
```

2. Create Table

What are the different Types of Tables available in Hive?

In Hive, there are mainly two types of tables:

1. **Managed (Internal) Tables:** These tables are managed by Hive, and Hive assumes full control over their lifecycle, including data storage and deletion. Data for managed tables is typically stored in a Hive-specific directory in Hadoop Distributed File System (HDFS).
2. **External Tables:** These tables are linked to data that exists outside of Hive's control, such as data stored in HDFS or other storage systems like Amazon S3. Unlike managed tables, Hive does not manage the data lifecycle of external tables, and dropping an external table does not delete the underlying data.

```
-- Basic Syntax
CREATE TABLE IF NOT EXISTS TABLE1 (
    COL1 STRING,
    COL2 ARRAY<STRING>,
    COL3 STRING,
    COL4 INT)

ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
COLLECTION ITEMS TERMINATED BY ':'
LINES TERMINATED BY '\n'
STORED AS TEXTFILE;
```

Internal/Managed Table

```
0: jdbc:hive2://localhost:10000/> CREATE TABLE IF NOT EXISTS TABLE1 (
. . . . . > Name STRING,
. . . . . > Random ARRAY<STRING>,
. . . . . > Country String,
. . . . . > Age INT)
. . . . . > ROW FORMAT DELIMITED
. . . . . > FIELDS TERMINATED BY ','
. . . . . > COLLECTION ITEMS TERMINATED BY ':'
. . . . . > LINES TERMINATED BY '\n'
. . . . . > STORED AS TEXTFILE;
```

```
LOAD DATA LOCAL INPATH '/hive_custom_data/data.txt' INTO TABLE TABLE1;
-- this file is present in windows folder which I mounted while running the
```

container.

Contents of data.txt file:

```
John,apple:banana:cherry,USA,30
Doe,car:bike:bus,Canada,25
Alice,cat:dog:hamster,UK,35
Bob,pen:paper:notebook,Australia,40
Eve,grape:melon:orange,USA,28
Frank,train:plane:boat,Canada,32
Grace,hat:scarf:gloves,UK,27
Hank,shoe:sock:boot,Australia,45
Ivy,phone:tablet:computer,USA,29
Jack,coffee:tea:juice,Canada,34
Kara,pen:marker:pencil,UK,22
Leo,ball:bicycle:skateboard,Australia,33
Mia,shirt:jeans:jacket,USA,31
Nina,camera:microscope:telescope,Canada,29
Oscar,book:magazine:newspaper,UK,36
Pia,ring:necklace:bracelet,Australia,38
Quinn,drum:guitar:piano,USA,26
Rita,painting:sculpture:photograph,Canada,30
Sam,shampoo:soap:toothpaste,UK,32
Tina,hat:glasses:watch,Australia,24
Uma,laptop:keyboard:mouse,USA,30
Victor,balloon:partyhat:confetti,Canada,31
Wendy,bag:sunglasses:wallet,UK,33
Xander,shoes:shirt:belt,Australia,39
Yara,pen:highlighter:eraser,USA,25
Zane,drone:robot:car,Canada,28
Ada,clock:calendar:thermometer,UK,30
Ben,chair:table:lamp,Australia,34
Cleo,plant:vase:flowerpot,USA,27
Dan,suit:tie:shirt,Canada,37
Ella,glove:hat:scarf,UK,29
```

```
SELECT * FROM TABLE1 LIMIT 4;
```

table1.name	table1.random	table1.country	table1.age
John	["apple","banana","cherry"]	USA	30

Doe	["car", "bike", "bus"]	Canada	25
Alice	["cat", "dog", "hamster"]	UK	35
Bob	["pen", "paper", "notebook"]	Australia	40

4 rows selected (0.169 seconds)

Creating an external table:

```
CREATE EXTERNAL TABLE ExtEmployee (  
    id INT,  
    name STRING,  
    salary INT  
)  
ROW FORMAT DELIMITED  
FIELDS TERMINATED BY ','  
STORED AS TEXTFILE  
LOCATION '/hive_custom_data/'  
TBLPROPERTIES ("skip.header.line.count"="1");
```

Note: You should have emp.csv in the custom_hive_data folder.

3. Insert Vs Overwrite

- **INSERT INTO:** This command appends new data to an existing table without affecting the existing records. In the example, we use INSERT INTO to add two new employees (Frank and Grace) to the Employees table.
- **OVERWRITE:** This command replaces all existing data in a table with new data. In the example, we use INSERT OVERWRITE to replace all records in the Departments table with a new set of departments, including a new 'Human Resources' department.

```
Create database nik_hive;  
use nik_hive;  
  
-- Create Table 1: Employees  
CREATE TABLE Employees (  
    employee_id INT,  
    name STRING,  
    department STRING,  
    salary DECIMAL(10, 2)  
)
```

```

ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;

-- Insert Data into Employees
INSERT INTO TABLE Employees
VALUES
    (1, 'Alice', 'Engineering', 60000),
    (2, 'Bob', 'Sales', 50000),
    (3, 'Charlie', 'Marketing', 55000),
    (4, 'David', 'Engineering', 65000),
    (5, 'Eva', 'Sales', 52000);

-- Create Table 2: Departments
CREATE TABLE Departments (
    department_id INT,
    department_name STRING,
    location STRING
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;

-- Insert Data into Departments
INSERT INTO TABLE Departments
VALUES
    (1, 'Engineering', 'San Francisco'),
    (2, 'Sales', 'New York'),
    (3, 'Marketing', 'Chicago');

-- Create Table 3: Projects
CREATE TABLE Projects (
    project_id INT,
    project_name STRING,
    start_date STRING,
    end_date STRING,
    department STRING -- Common column for joining later
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;

-- Insert Data into Projects
INSERT INTO TABLE Projects
VALUES

```

```
(101, 'Website Redesign', '2024-01-15', '2024-06-30', 'Engineering'),
(102, 'Product Launch', '2024-03-01', '2024-09-30', 'Sales'),
(103, 'Marketing Campaign', '2024-02-15', '2024-07-31', 'Marketing'),
(104, 'Infrastructure Upgrade', '2024-04-01', '2024-11-30', 'Engineering'),
(105, 'Sales Training Program', '2024-02-01', '2024-08-31', 'Sales');

-- Demonstrate INSERT INTO
INSERT INTO TABLE Employees
VALUES
    (6, 'Frank', 'Marketing', 58000),
    (7, 'Grace', 'Engineering', 62000);

-- Demonstrate OVERWRITE
INSERT OVERWRITE TABLE Departments
VALUES
    (1, 'Engineering', 'Seattle'),
    (2, 'Sales', 'Los Angeles'),
    (3, 'Marketing', 'Boston'),
    (4, 'Human Resources', 'Denver');

-- Create a new table to demonstrate the difference
CREATE TABLE temp_employees LIKE Employees;

-- Use INSERT INTO to add some data
INSERT INTO TABLE temp_employees
SELECT * FROM Employees WHERE department = 'Engineering';

-- Use OVERWRITE to replace all data
INSERT OVERWRITE TABLE temp_employees
SELECT * FROM Employees WHERE department = 'Sales';
```