

03 Hive Built-in Functions - KirkYagmai

1. Standard Functions

Standard functions take a row/columns in a row as arguments.

```
SELECT concat(first_name, " ", last_name) FROM employees;
```

They return a single result for each row.

Example:

```
SELECT concat(first_name, " ", last_name) FROM employees;
-- Result:
-- Kirk Yagami
-- Nikhil Sharma
-- Minku Mathur
```

These include:

- **Mathematical Functions:** EXP(), LN(), SQRT(), POW(), etc.
- **String Functions:** LENGTH(), REVERSE(), REGEXP_REPLACE(), and many others.

2. Aggregate Functions

These functions take multiple rows as input and return a single result.

- **Examples:** SUM(), COUNT(), AVG(), etc.

These functions are normally used with a GROUP BY clause.

Example:

```
SELECT Productname, SUM(Revenue)
FROM sales_data
GROUP BY Productname;
```

3. Table Generating Functions

These functions take 1 row as input and output multiple rows. They normally operate on collection data types like arrays, maps, and structs.

```
SELECT explode(array(1, 2, 3));  
-- Result:  
-- 1  
-- 2  
-- 3
```

Size()

- FINDING THE SIZE OF AN ARRAY/MAP
- HIVE HAS A FEW BUILT IN FUNCTIONS FOR COLLECTION DATA TYPES LIKE ARRAYS, MAPS, STRUCTS ETC
- HIVE FUNCTIONS SIZE() FOR ARRAYS IT RETURNS THE NUMBER OF ELEMENTS
- FOR MAPS IT RETURNS THE NUMBER OF KEY-VALUE PAIRS

```
SELECT SIZE(ARRAY(1,2,3));  
3  
  
SELECT SIZE(MAP("NAME","SWETHA","AGE",30));  
2
```

4. Case statement

```
-- Create Table: EmployeeTenure  
CREATE TABLE EmployeeTenure (  
    Emp_ID INT,  
    EmpName STRING,  
    Tenure INT  
)  
ROW FORMAT DELIMITED  
FIELDS TERMINATED BY ','  
STORED AS TEXTFILE;  
  
-- Insert Data into EmployeeTenure  
INSERT INTO TABLE EmployeeTenure  
VALUES  
    (1, 'Nikhil', 1),  
    (2, 'Minku', 4),  
    (3, 'Pinku', 2),
```

```
(4, 'Coco', 3);
```

```
FROM EmployeeTenure SELECT empname, CASE WHEN TENURE<2 THEN 0 WHEN TENURE>=2 AND  
TENURE<=3 THEN 2 WHEN TENURE>3 THEN 3 END AS EXTRA_VACATION_DAYS;
```

```
SELECT Emp_ID, empname, Tenure, CASE WHEN Tenure < 2 THEN 0 WHEN Tenure BETWEEN 2  
AND 3 THEN 2 WHEN Tenure > 3 THEN 3 END AS ExtraVacationDays FROM EmployeeTenure;
```