**ASSIGNMENT 23.2**

**Explain Primary data types and complex data types in Hive with an example in brief**

**Primary Data Types:**

Primary Data Types are further classified into four categories. They are:

• Numeric Types

• String Types

• Date/Time Types

• Miscellaneous Types

**Numeric Data Types :**

* TINYINT :

It is a 1-byte signed integer, from -128 to 127.

* SMALLINT :

It is a 2-byte signed integer, from -32,768 to 32,767

* INT :

It is a 4-byte signed integer, from -2,147,483,648 to 2,147,483,647

* BIGINT :

It is a 8-byte signed integer, from 9,223,372,036,854,775,808 to 9,223,372,036,854,775,807

* FLOAT :

It is a 4-byte single precision floating point number.

* DOUBLE :

It is a 8-byte double precision floating point number.

* DECIMAL:

It is a 17 byte floating point number which has precision upto 38 digits

**String Types:**

* STRING :

String literals can be expressed with either single quotes (') or double quotes (")

* VARCHAR :

Varchar types are created with a length specifier (between 1 and 65355), which defines the maximum number of characters allowed in the character string.

* CHAR :

Char types are similar to Varchar but they are fixed-length meaning that values shorter than the specified length value are padded with spaces but trailing spaces are not important during comparisons.

**Date/Time Types:**

* TIMESTAMP:

TIMESTAMP uses the format yyyy-mm-dd hh:mm:ss. Using timestamp you can do manipulations with time.

* DATE:

DATE values are represented in the form YYYY-MM-DD. Example: DATE ‘2014-12-07’. Date ranges allowed are 0000-01-01 to 9999-12-31.

**Miscellaneous Types:**

Hive supports two more primitive data types, BOOLEAN and BINARY

* Booelan:

Similar to Java’s Boolean, BOOLEAN in hive stores true or false values only.

* Binary:

BINARY is an array of Bytes and similar to VARBINARY in many RDBMSs

**Complex data types:**

1. ARRAY
2. MAP
3. STRUCT
4. UNIONTYPE

ARRAY:

* **Hive Array**behavior is same as Java Array.
* It is an ordered collection of elements.
* The all elements in the array must be of same data type.
* Example array(1,2,3,4,5,6,7,8)
* If you want to access the second element , then array[1] can be used

MAP :

* **Hive Map data type** is one type of Hive complex data types example.
* It is an unordered collection of key-value pairs.
* Keys must be of primitive types.
* Values can be of any type.
* Fields are accessed using array notation of keys (e.g., [‘key’]).
* Example: map(‘a’,1,’b’,2)

STRUCT :

* It is similar to STRUCT in C language.
* It is a collection of elements of different types.
* Moreover It is a record type which encapsulates a set of named fields that can be any primitive data type.
* we can use any data type to specify this struct data type.
* Elements in STRUCT type are accessed using the DOT (.) notation.
* Example: struct(‘a’,1,1.0)

Here I’ve stored a string, integer and float.

UNIONTYPE :

* Uniontype is collection of Heterogeneous data types.
* It is similar to Unions in C.
* At any point of time, an Union Type can hold any one (exactly one) data type from its specified data types