Explain with an example in brief.

● Hive Data Definitions

The hive data definitions relates to all the commands which used express the information about the data.

Different data definitions commands. HDL is a subset of Hive SQL statements that describe the data structure in Hive by creating, deleting, or altering schema objects such as databases, tables, views, partitions, and buckets.

CREATE, DROP, TRUNCATE, DESCRIBE, AND ALTER.

Examples

CREATE TABLE HIVE\_TABLE (foo INT, bar STRING); // used for creating tables

DROP ‘TABLE NAME’; // deleting the table or database;

TRUNCATE is to delete the structure of the table;

ALTER it is used to for changing the structure of the table;

● Hive Data Manipulations

Hive data manipulations commands are used for modifying the values in the table or extracting data from the table.

E.g. LOAD, INSERT, UPDATE, DELETE

LOAD is used for loading the data into a table.

Update is to change the entries in the table.

● HiveQL Manipulations

Different hiveql manipulations are of three types

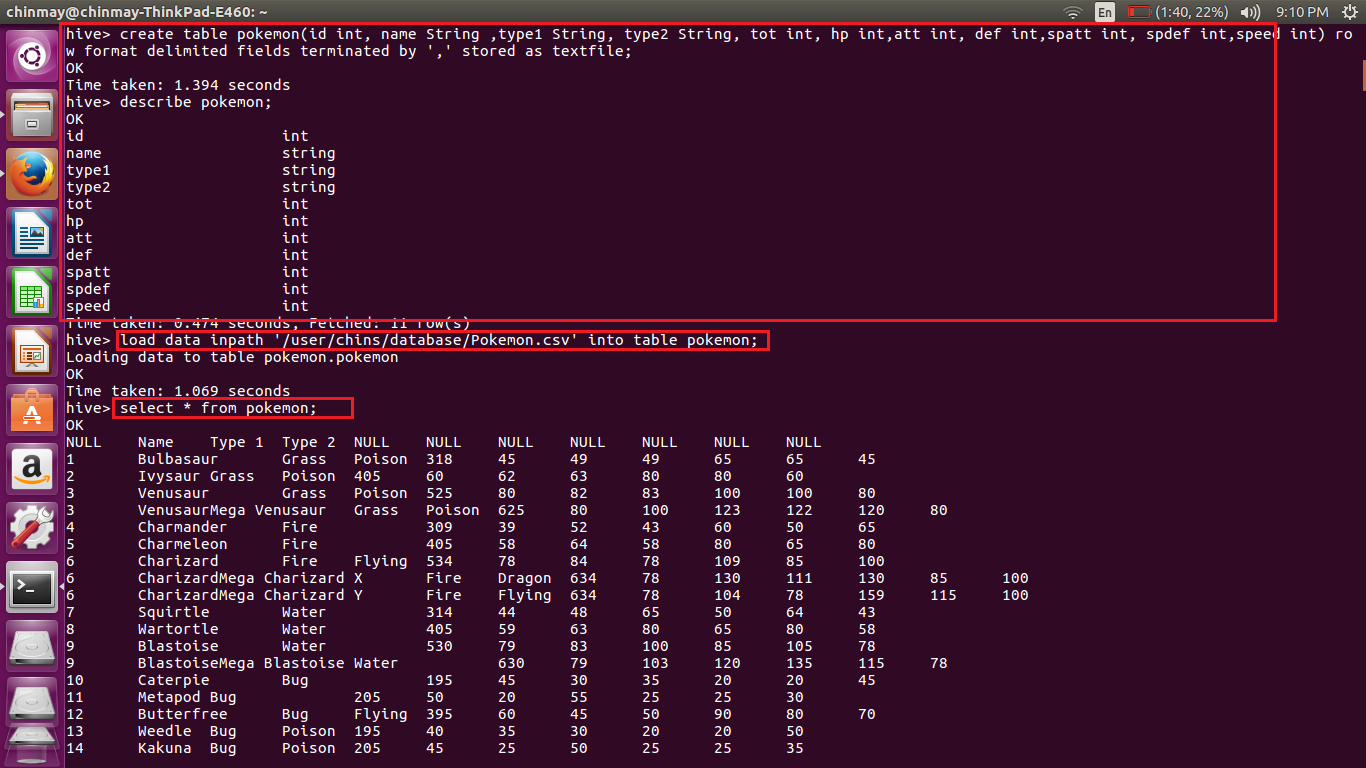
1. Select where
2. Select order by /sort by
3. Select group by

Select where is used to filter the dataset with the where clause and apply conditions.

The select order by clause is used to arrange the dataset in the ascending and descending order by one field or multiple fileds.

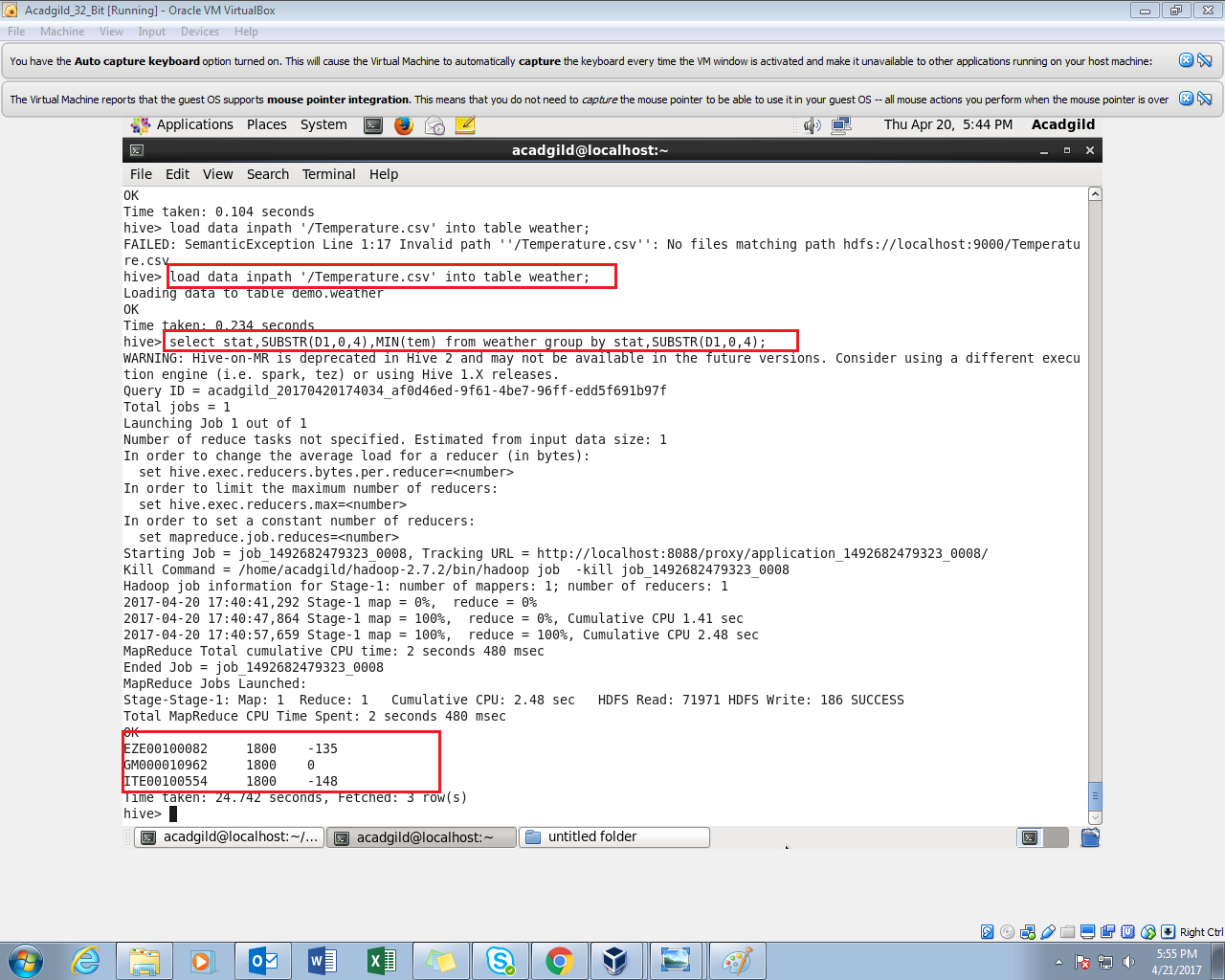
Select group by is used to form subsets of the database using different fields of data.

Below are the examples of the DDL DML and HiveQL



Select \* displays all the data from the table.

Describe command is used to get the structure of the table.



In this above example we are grouping by the field ‘stat’ and substring of the field D1.