Creating Partitioned Table

- Create the cases table with multiple partitions. We partition by the department, and then by the severity_of_illness -
    - **create table cases_part(Case_Id int,**

      Hospital_Code int,

      Patient_Id int,

      Ward_Type char(1),

      Ward_Facility_Code char(1),

      Bed_Grade int,

      Type_of_Admission string,

      Visitors_with_Patient int,

      Age_Group string,

      Admission_Deposit float,

      Stay string)

      partitioned by (Department string, Severity_of_Illness string)

      row format delimited

      fields terminated by '$'

      TBLPROPERTIES ("serialization.null.format"="");
- Put file to HDFS -
    - **hdfs dfs -put /home/itv180149/Hive/Partitions/Hospital_Data/case.txt /user/itv180149**

- Check directory -
    - **hdfs dfs -ls /user/itv180149**
- Load data into Hive table -
    - **load data inpath "/user/itv180149/case.txt" into table cases_part;**
- Check directory in HDFS -
    - **hdfs dfs -ls /user/itv180149/warehouse/hospital_db.db/cases_part**
- Notice the new subdirectories that are created. These subdirectories are based on the values in the department column on which we partitioned the table. And each subdirectory contains data specific to that particular department. Compare this to the directories we saw earlier in the module on non-partitioned tables. We can actually look inside one of these subdirectories -
    - **hdfs dfs -ls /user/itv180149/warehouse/hospital_db.db/cases_part/department=anesthesia**
- Inside each sub-directory, we have the data partitioned by the second column on which we partitioned by data, which is the severity_of_illness
    - **hdfs dfs -ls /user/itv180149/warehouse/hospital_db.db/cases_part/department=anesthesia/severity_of_illness=Extreme**
- Inside this we have the actual data. So this data was first partitioned by the department and then by the severity of illness.

- We can check the description and it will show that the table is partitioned on the columns-
    - **describe cases_part;**
- And we can even check each partition value by typing in the following command -
    - **show partitions cases_part;**