>Create db meenu\_cap;

> set hive.exec.dynamic.partition=true;

>set hive.exec.dynamic.partition.mode=nonstrict;

> CREATE TABLE IF NOT EXISTS meenu\_cap.coursedetails\_demo (

coursetype string,

courseid string,

title string,

competency string,

complexity string)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ',';

Chart, text

Description automatically generated

> load data inpath 'capstone\_project/StudentCourseDetails.csv' into table coursedetails\_demo;

Graphical user interface, text, application

Description automatically generated

> CREATE TABLE IF NOT EXISTS meenu\_cap.coursedetails (

courseid string,

title string,

competency string,

complexity string)

partitioned by (coursetype string)

row format delimited

fields terminated by ',';

Text

Description automatically generated

> insert into meenu\_cap.coursedetails

partition(coursetype)

select courseid,title,competency,complexity,coursetype from meenu\_cap.coursedetails\_demo;

Graphical user interface, text, application, email

Description automatically generated

> select \* from coursedetails limit 10 ;

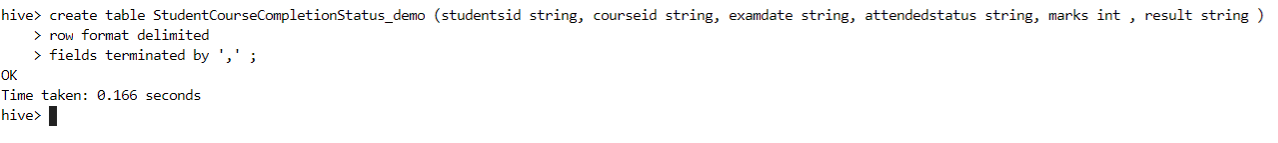
Text

Description automatically generated

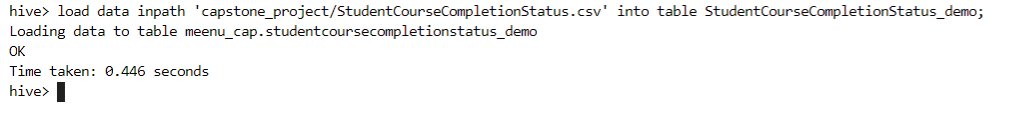
> create table StudentCourseCompletionStatus\_demo (studentsid string, courseid string, examdate string, attendedstatus string, marks int , result string )

row format delimited

fields terminated by ',' ;



> load data inpath 'capstone\_project/StudentCourseCompletionStatus.csv' into table StudentCourseCompletionStatus\_demo;



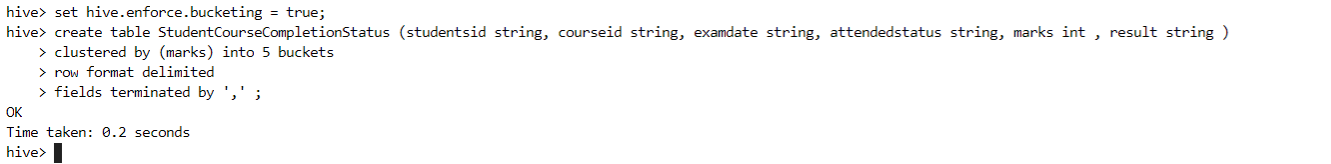
> set hive.enforce.bucketing = true;

> create table StudentCourseCompletionStatus (studentsid string, courseid string, examdate string, attendedstatus string, marks int , result string )

clustered by (marks) into 5 buckets

row format delimited

fields terminated by ',' ;



> insert overwrite table StudentCourseCompletionStatus select \* from StudentCourseCompletionStatus\_demo;

Text

Description automatically generated

> select \* from studentcoursecompletionstatus limit 10 ;

Graphical user interface, table

Description automatically generated

1. List the count of students qualified for various courses

>select count(result) from StudentCourseCompletionStatus where result = 'Qualified' group by result;

Text

Description automatically generated

2. List the courses available.

>select distinct(coursetype) from coursedetails ;

3. Extract the following fields from both table from HDFS directory and save the following result set to HDFS output directory -

a . StudentID

b. CourseTitle

c. Result

> INSERT OVERWRITE

DIRECTORY 'capstone\_project'

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

select e1.studentsid, e2.title,e1.result from StudentCourseCompletionStatus e1 left outer join coursedetails e2 on e1.courseid= e2.courseid;

Graphical user interface, text, application, email

Description automatically generated

Text

Description automatically generated