**Capstone Project – Part 1 & Part 2**

**Task 1: Defining and Creating Tables**

Code:

set hive.metastore.warehouse.dir=/user/ana002683/capstone;

create table status(

studentsid string,

courseid string,

examdate string,

attendedstatus string,

marks int,

result string)

row format delimited

fields terminated by ',';

load data local inpath '/home/ana002683/capstone\_project/completionstatus.csv' into table status;

create table statusbucket(

studentsid string,

courseid string,

examdate string,

attendedstatus string,

marks int,

result string)

clustered by (marks)

sorted by (marks)

into 5 buckets;

insert into statusbucket

select \* from status;

create table details(

coursetype string,

courseid string,

title string,

competency string,

complexity string)

row format delimited

fields terminated by ',';

load data local inpath '/home/ana002683/capstone\_project/CourseDetails.csv' into table status;

create table detailspart(

courseid string,

title string,

competency string,

complexity string)

partitioned by (coursetype string);

insert into detailspart

select courseid,title,competency,complexity,coursetype from details;

Graphical user interface, text, email

Description automatically generatedGraphical user interface, text, application

Description automatically generatedGraphical user interface, text, application

Description automatically generatedGraphical user interface, text, application

Description automatically generated

**Task 2: Data Analysis using Hive**

Code:

1. select count (distinct studentsid) as qualified\_students

from statusbucket

where result="Qualified";

1. select title from detailspart;
2. INSERT OVERWRITE LOCAL DIRECTORY '/user/ana002683/capstone/output'

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

select studentsid,title,result from statusbucket a join detailspart b on a.courseid=b.courseid;

Graphical user interface, text, application

Description automatically generatedGraphical user interface, text

Description automatically generatedGraphical user interface, text, application

Description automatically generated