Name: Siddhesh Dilip Khairnar

Class: **TY**	Division: **B**	Roll No: **372028**

Semester: **6th**

Academic Year: **2023-24**

Subject Name & Code: **Data Science & ADUA32202**

Title of Assignment: Perform following using HiveQL for Student information system

ASSIGNMENT NO: - 7

Aim: Perform following using HiveQL for Student information system
1) Creating Database tables
2) insert new values in the table
3) Join tables with Hive
4) Find the average grades of students per

1) Creating Database tables

```
CREATE TABLE table_name(
    column1 datatype1,
    column2 datatype2,
    column3 datatype3,
    .....
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;
```

This will create a table with the specified column names and datatypes, delimited by commas and stored as a text file.

2) insert new values in the table

```
INSERT INTO TABLE table_name VALUES (value1, value2, ...);
```

```
INSERT INTO TABLE students VALUES (1, 'John', 2022, 85.5);
```


3) Join tables with Hive

To join tables in HiveQL, you can use the JOIN keyword and specify the columns to join on

```
SELECT table1.column1, table2.column2
FROM table1 JOIN table2
ON table1.id = table2.id;
```

This will select column1 from table1 and column2 from table2, where the id column in both tables matches.

4) Find the average grades of students per:

To find the average grades of students per subject, you can use the GROUP BY clause and the AVG() function in HiveQL. For example:

```
SELECT subject, AVG(grade) as avg_grade
FROM student_grades
GROUP BY subject;
```

This will calculate the average grade for each subject in the student grades table.

Conclusion: In conclusion, leveraging HiveQL for a Student Information System proves to be a robust method for managing and analyzing student data. This assignment encompassed various tasks including table creation, data insertion, table joining, and grade analysis.