EXERCISE NO.: 06
HIVE DATABASES, TABLES, VIEWS, FUNCTIONS, AND INDEXES

AIM:

To create and manage Hive databases, tables, views, functions, and indexes, and to query structured data using HiveQL to demonstrate Hive's data warehousing capabilities on Hadoop.

SCRIPTS:

students.csv
id,name,age,department
1,John,20,CSE
2,Alice,21,ECE
3,Bob,19,CSE

marks.csv
id,subject,marks
1,Math,85
2,Physics,90
1,Physics,75
3,Math,80

HiveQL Script:

```
CREATE DATABASE IF NOT EXISTS student_db;
USE student_db;
CREATE TABLE students (
    id INT,
    name STRING,
    age INT,
    department STRING
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;
CREATE EXTERNAL TABLE marks (
```

```
    id INT,
    subject STRING,
    marks INT
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
LOCATION '/user/hive/external/marks';
LOAD DATA LOCAL INPATH 'students.csv' INTO TABLE students;
LOAD DATA LOCAL INPATH 'marks.csv' INTO TABLE marks;
CREATE VIEW student_marks AS
SELECT s.id, s.name, s.department, m.subject, m.marks
FROM students s
JOIN marks m
ON s.id = m.id;
SELECT department, AVG(marks) AS avg_marks
FROM student_marks
GROUP BY department;
CREATE INDEX idx_marks ON TABLE marks(marks)
AS 'COMPACT' WITH DEFERRED REBUILD;
ALTER INDEX idx_marks ON marks REBUILD;
```

OUTPUT:

Database student_db created successfully.

Table students created.
Table marks created.

View student_marks created.

department avg_marks
CSE 80.0
ECE 90.0

Index idx_marks created and rebuilt successfully.

RESULT:

Thus, the Hive database, tables, and views were successfully created.