HIVE operations

0. Hadoop start and enter command

   docker container start -i hadoop;

1. Transfer file to hadoop file system

   hdfs dfs -put /sagor/student.txt   /sagor

2. Create a dummy table for inserting into Student partitioning table

```
CREATE EXTERNAL TABLE IF NOT EXISTS StudentDummy (
   id int,
   name STRING,
   age INT,
   Gender string
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY  ','
STORED AS TEXTFILE
LOCATION '/data/output';
```

3. Load data to StudentDummy

   LOAD DATA INPATH '/sagor/students.txt' INTO TABLE StudentDummy;

4. Create a partition table

```
CREATE EXTERNAL TABLE IF NOT EXISTS Student (
   id int,
   name STRING,
   age INT
)
PARTITIONED BY (gender STRING)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY  ','
STORED AS TEXTFILE
LOCATION '/data/output';
```

5. Insert data to partitioned table **(Dynamic partition)**

   ```
   SET hive.exec.dynamic.partition=true;
   SET hive.exec.dynamic.partition.mode=nonstrict;

   INSERT INTO TABLE Student PARTITION (gender)
   SELECT id, name, age, gender FROM StudentDummy;
   ```

Not needed, step 5 do this work;

---

6. Insert class-1(male) data to Student table

   ```
   INSERT INTO TABLE Student PARTITION (gender='male')
   SELECT id,  name, age FROM StudentDummy WHERE gender='male';
   ```

7. Insert class-2(female) data to Student table

   ```
   INSERT INTO TABLE Student PARTITION (gender='female')
   SELECT id,  name, age FROM StudentDummy WHERE gender='female';
   ```

---

8. Test out Student table

   ```
   Select * from Student;
   ```

   ```
   Select *
   Form Stud as a cross join Stud as b
   Where a.gender=’male’ and b.gender=’female’ and a.age=b.age
   ```