HQL script creating the data warehouse

```
CREATE TABLE route (

    route_id INT,

    route_name STRING,

    metrics MAP<STRING, INT>

)

CLUSTERED BY (route_id) INTO 10 buckets

STORED AS ORC;


CREATE TABLE date_dim (

    date_id INT,

    date_format DATE,

    year INT,

    month STRING

)

PARTITIONED BY (month_no INT, day_type STRING)

CLUSTERED BY (date_id) INTO 30 BUCKETS

STORED AS PARQUET;


CREATE TABLE time_dim (

    time_id INT,

    hour INT,

    minutes INT

)

PARTITIONED BY (time_of_day STRING)

STORED AS PARQUET;
```

```sql
CREATE TABLE travel (
    bus_id INT,
    route_id INT,
    departure_time INT,
    arrival_time INT,
    tickets_validated INT,
    bus_capacity INT,
    avg_satisfaction_level_received INT,
    satisfaction_surveys_number INT,
    junk_id INT
)
PARTITIONED BY (travel_date INT)
CLUSTERED BY (route_id, bus_id) INTO 50 BUCKETS
STORED AS ORC;


CREATE TABLE junk (
    junk_id INT,
    satisfaction_level_category STRING,
    occupation_level_category STRING
)
STORED AS PARQUET;


CREATE TABLE bus (
    bus_id INT,
    bus_registration STRING,
    bus_office_id INT,
    additional_equipment ARRAY<STRING>
)
PARTITIONED BY (bus_type STRING)
CLUSTERED BY (bus_id) INTO 10 BUCKETS
STORED AS ORC;
```

```
CREATE EXTERNAL TABLE bus_office (

    bus_office_id INT,

    bus_office_name STRING,

    city STRING,

    size_category STRING

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE

LOCATION 'hdfs:///user/mmajewska/database';
```