Task1

(1)

Facts:

(i) Total Kilometers

(ii) Total Trips

(iii) Total Drivers

(iv) Total Travels

(v) Average Duration

(vi) Total Fuel Consumption

(vii) Total Passengers

(viii) Largest Passengers

(ix) Average Passengers

Measures:

(i) Length

(ii) the number of trips

(iii) the number of drivers

(iv) the number of travels

(v) Duration
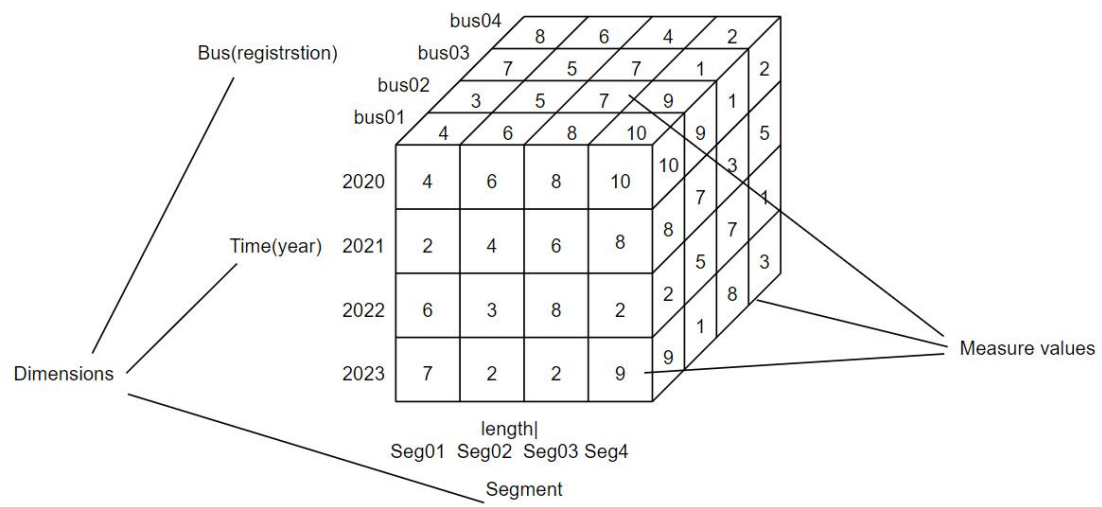
(vi) Fuel consumption

(vii) the number of passengers

Dimension:

(i) Time Dimension: year, month, day

(ii) Bus Dimension: BusID, model, manufacturer

(iii) Driver Dimension: licence-number, full-name, years-of experience

(iv) Segment Dimension: end-point-1, endpoint-2,total length

(v) Trip Dimension: trip name

(vi) Depot Dimension: lot-number, street, city
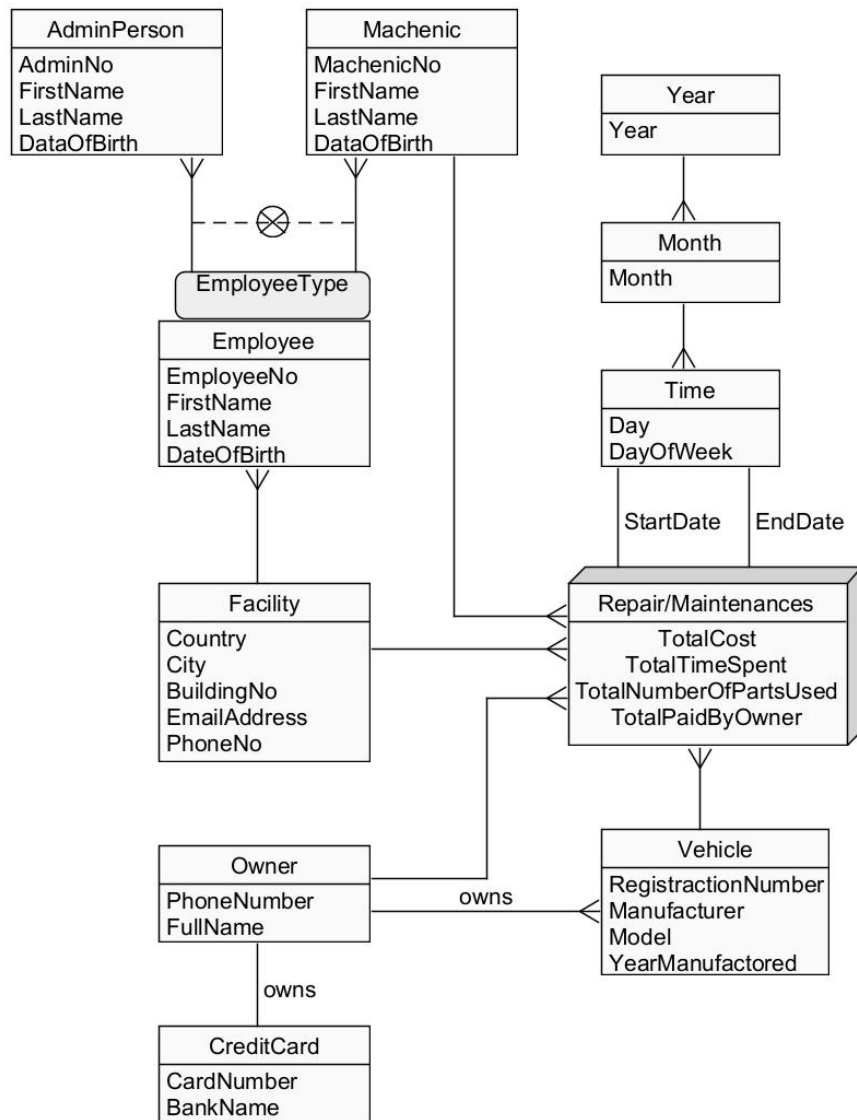
(vii) Manufacture Dimension: name, city, country

Hierarchies:

(i) Time Dimension:Year > month > day

(ii) Bus Dimension: BusID > model > manufacturer

(2)

Bus(registrstion)

bus04   8   6   4   2

bus03   7   5   7   1   2

bus02   3   5   7   9   1

bus01   4   6   8   10   9   5

Time(year)

	Seg01	Seg02	Seg03	Seg4			
2020	4	6	8	10	10	3	
2021	2	4	6	8	7	7	1
2022	6	3	8	2	8	5	3
2023	7	2	2	9	2	8	
					1		
					9		

Dimensions

length|
Seg01 Seg02 Seg03 Seg4

Segment

Measure values

Task2

AdminPerson
- AdminNo
- FirstName
- LastName
- DataOfBirth

Machenic
- MachenicNo
- FirstName
- LastName
- DataOfBirth

EmployeeType

Employee
- EmployeeNo
- FirstName
- LastName
- DateOfBirth

Year
- Year

Month
- Month

Time
- Day
- DayOfWeek

StartDate   EndDate

Facility
- Country
- City
- BuildingNo
- EmailAddress
- PhoneNo

Repair/Maintenances
- TotalCost
- TotalTimeSpent
- TotalNumberOfPartsUsed
- TotalPaidByOwner

Owner
- PhoneNumber
- FullName

owns

Vehicle
- RegistractionNumber
- Manufacturer
- Model
- YearManufactored

owns

CreditCard
- CardNumber
- BankName

Task3

The output of solution3.hql with no error:

```
0: jdbc:hive2://localhost:10000> !run solution3.hql
>>>  create table solution3(
student_number int,
evaluations ARRAY<STRUCT<subject_code:string, final_evaluation:int>>);
No rows affected (0.106 seconds)
>>>
>>>  insert into solution3 select 1111, array(named_struct('subject_code', 'CSIT111', 'final_evaluation', 01), named_struct('subject_code', 'CSIT121', 'final_evaluation', 23), named
_struct('subject_code', 'CSIT101', 'final_evaluation', 50), named_struct('subject_code', 'CSIT235', 'final_evaluation', 99), named_struct('subject_code', 'ISIT312', 'final_evaluatio
n', 02));
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
No rows affected (9.961 seconds)
>>>  insert into solution3 select 1112, array(named_struct('subject_code', 'CSIT101', 'final_evaluation', 56),named_struct('subject_code', 'CSIT111', 'final_evaluation', 78),named_s
truct('subject_code', 'CSIT115', 'final_evaluation', 10),named_struct('subject_code', 'ISIT312', 'final_evaluation', 05));
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
No rows affected (13.653 seconds)
>>>  insert into solution3 select 1113, array(named_struct('subject_code', 'CSIT121', 'final_evaluation', 76),named_struct('subject_code', 'CSIT235', 'final_evaluation', 87),named_s
truct('subject_code', 'ISIT312', 'final_evaluation', 49));
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
No rows affected (15.2 seconds)
>>>  insert into solution3 select 1114, array(named_struct('subject_code', 'CSIT111', 'final_evaluation', 50),named_struct('subject_code', 'ISIT312', 'final_evaluation', 45));
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
No rows affected (14.051 seconds)
>>>  insert into solution3 select 1115, array(named_struct('subject_code', 'ISIT115', 'final_evaluation', 67),named_struct('subject_code', 'CSCI235', 'final_evaluation', 45),named_s
truct('subject_code', 'CSIT127', 'final_evaluation', 56));
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
No rows affected (14.672 seconds)
>>>
>>>  select * from solution3;
+-------------------------+----------------------------------------------------------------------------------+--+
| solution3.student_number |                                                                                  |          solution3.evaluations
|                         |                                                                                  |
+-------------------------+----------------------------------------------------------------------------------+--+
| 1111                    | [{"subject_code":"CSIT111","final_evaluation":1},{"subject_code":"CSIT121","final_evaluation":23},{"subject_code":"CSIT101","final_evaluation":50},{"su
bject_code":"CSIT235","final_evaluation":99},{"subject_code":"ISIT312","final_evaluation":2}] |
| 1112                    | [{"subject_code":"CSIT101","final_evaluation":56},{"subject_code":"CSIT111","final_evaluation":78},{"subject_code":"CSIT115","final_evaluation":10},{"s
ubject_code":"ISIT312","final_evaluation":5}] |
| 1113                    | [{"subject_code":"CSIT121","final_evaluation":76},{"subject_code":"CSIT235","final_evaluation":87},{"subject_code":"ISIT312","final_evaluation":49}]
|
| 1114                    | [{"subject_code":"CSIT111","final_evaluation":50},{"subject_code":"ISIT312","final_evaluation":45}]
|
| 1115                    | [{"subject_code":"ISIT115","final_evaluation":67},{"subject_code":"CSCI235","final_evaluation":45},{"subject_code":"CSIT127","final_evaluation":56}]
|
+-------------------------+----------------------------------------------------------------------------------+--+
```

Task4

The output of solution4.hql with no error:

```
0: jdbc:hive2://localhost:10000> !run solution4.hql
>>>   create external table e_author(
A_ID          INT,
A_FNAME       VARCHAR(20),
A_MNAME       VARCHAR(20),
A_LNAME       VARCHAR(20),
A_DOB         VARCHAR(20),
A_BIO         VARCHAR(500))
row format delimited fields terminated by '|' stored as textfile location '/user/bigdata/new-hdfs-folder';
No rows affected (0.062 seconds)
>>>
>>>   create external table e_item(
I_ID          INT,
I_TITLE       VARCHAR(60),
I_A_ID        INT,
I_PUB_DATE    VARCHAR(10),
I_PUBLISHER   VARCHAR(60),
I_SUBJECT     VARCHAR(60),
I_DESC        VARCHAR(500),
I_RELATED1    INT,
I_RELATED2    INT,
I_RELATED3    INT,
I_RELATED4    INT,
I_RELATED5    INT,
I_THUMBNAIL   VARCHAR(10),
I_IMAGE       VARCHAR(10),
I_SRP         INT,
I_COST        INT,
I_AVAIL       VARCHAR(10),
I_STOCK       INT,
I_ISBN        CHAR(13),
I_PAGE        INT,
I_BACKING     VARCHAR(15),
I_DIMENSIONS  VARCHAR(25)
)
row format delimited fields terminated by '|' stored as textfile location '/user/bigdata/hdfs-folder-item';
No rows affected (0.056 seconds)
```

```
>>>   select count(*) from e_author;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
+-------+--+
|  c0   |
+-------+--+
| 2500  |
+-------+--+
1 row selected (12.767 seconds)
>>>
>>>   select count(*) from e_item;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
+--------+--+
|   c0   |
+--------+--+
| 10000  |
+--------+--+
1 row selected (17.758 seconds)
>>>
>>>   select SUM(total_rows) as total_rows
from(
select count(*) as total_rows from e_author
union all
select count(*) as total_rows from e_item
)combined;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases
.
+-------------+--+
| total_rows  |
+-------------+--+
| 12500       |
+-------------+--+
1 row selected (53.386 seconds)
```

```
>>>  select * from e_item limit 3;
```



```
3 rows selected (0.076 seconds)
```

```
>>>  select * from e_author limit 3;
+----------------+------------------+------------------+------------------+----------------...
...
...
--------------------+--+
| e_author.a_id  | e_author.a_fname | e_author.a_mname | e_author.a_lname | e_author.a_dob |
                                                                                              e_author.a_bio
                       |
+----------------+------------------+------------------+------------------+----------------...
...
...
--------------------+--+
| 1              | 5P040            | n50c2GP0n        | BABABABABABAOGl7xF | 11/21/74      | Minutes search in the young, labour years. Individuals will have to come. Trou
sers continue in the following changes: Actions at the big, young aspects shall sleep clear, african ideas. Great, large pupils do pick. Necessary eyes to the large, small years car
ry long, external circumstances. Questions might carry. Lovely, actual talks might rest customers: Social, other players identify. Systems could play. Standards ride. Overseas, fede
ral goods interfere. |
| 2              | cy8o             | ZkcR139U86m9K5642 | BABABABABABAALWH | 05/86/70      | Duties give between the methods. High, other children in the numbers realise m
odern, good years. Mountains talk teachers. New, hot studies forget to the members. Present, good programmes will take. Soviet, parliamentary dogs avoid. Sales to the only, joint pe
rsons ought to lead various, great lips. Areas wear about. Names could get with the unable, big quantities. Common, genuine shows between the mutual, prime houses miss between the
               |
| 3              | dC5              | Imnd0eY3A2712S6F9d | BABABABABABARI  | 11/09/84      | Big, black employers could tell. Ways exist. Available, regional waters in the
 limited, personal matters will have to make in the eyes. Ways should move
                       |
+----------------+------------------+------------------+------------------+----------------...
...
...
--------------------+--+
3 rows selected (0.067 seconds)
```