

Monash University

Semester Examination Period Faculty of Information Technology

EXAM CODES:

Sample-

TITLE OF PAPER:

Sample-

EXAM DURATION:

READING TIME:

THIS PAPER IS FOR STUDENTS STUDYING AT: (tick where applicable)

- | | | | | |
|---|---|------------------------------------|---|--|
| <input type="checkbox"/> Berwick | <input checked="" type="checkbox"/> Clayton | <input type="checkbox"/> Malaysia | <input checked="" type="checkbox"/> Off Campus Learning | <input type="checkbox"/> Open Learning |
| <input checked="" type="checkbox"/> Caulfield | <input type="checkbox"/> Gippsland | <input type="checkbox"/> Peninsula | <input type="checkbox"/> Enhancement Studies | <input type="checkbox"/> Sth Africa |
| <input type="checkbox"/> Pharmacy | <input type="checkbox"/> Other (specify) | | | |

During an exam, you must not have in your possession, a book, notes, paper, calculator, pencil case, mobile phone or other material/item which has not been authorised for the exam or specifically permitted as noted below. Any material or item on your desk, chair or person will be deemed to be in your possession. You are reminded that possession of unauthorised materials in an exam is a discipline offence under Monash Statute 4.1.

No examination papers are to be removed from the room.

Attempt all questions. All answers must be **printed neatly** on this paper. Answer questions with concisely expressed factual information. The backs of pages in this paper may be used for any rough work. **Any material written on the backs of pages will not normally be corrected.** If an answer needs to overflow from its designated answer space to a blank page, clearly indicate that this is the case and that the material on the blank page is for correction. This paper must be handed up at the end of the examination, even if no questions are attempted. **There are 8 questions, each of which is worth 10 marks. The total is 80 marks.** This exam counts as 60% of the final assessment for the unit.

AUTHORISED MATERIALS

CALCULATORS

☐ YES☒ NO

OPEN BOOK

☐ YES☒ NO

SPECIFICALLY PERMITTED ITEMS

☐ YES☒ NO

Candidates must complete this section if required to write answers within this paper

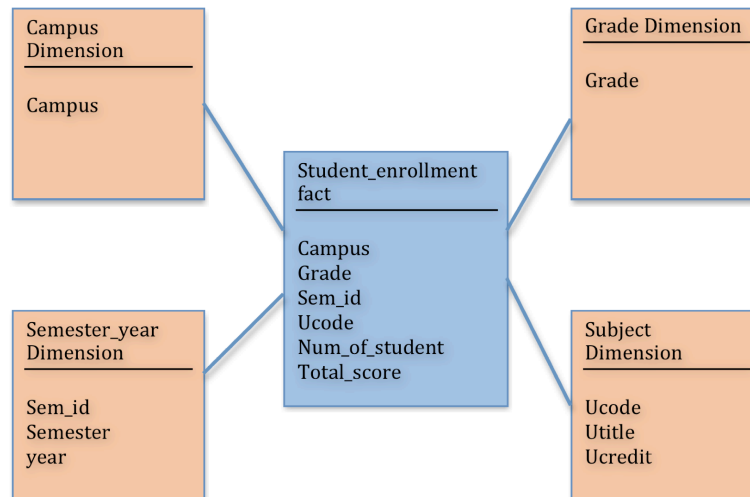
STUDENT ID

DESK NUMBER

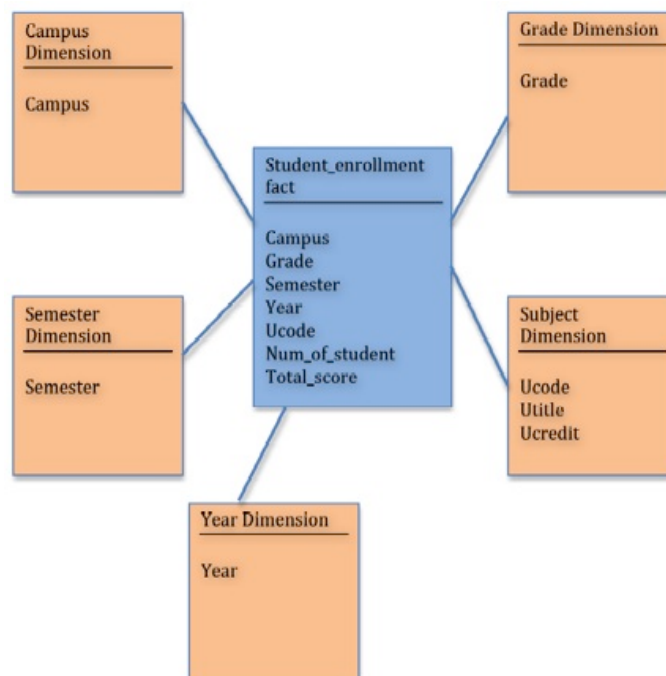
Question 1:

Consider the following two star schemas. Star schema-1 has a dimension called Semester_Year_Dimension, whereas in Star Schema-2, the Semester_Year_Dimension is split into two dimensions (e.g. Semester_Dimension and Year_Dimension)

Star Schema-1 (one semester-year dimension)



Star Schema-2 (two separate dimensions for semester and year)

**Questions:**

- Show the sample contents of tables Semester_Year_Dimension and Fact-1 (from Star Schema-1), and the sample contents of tables Semester_Dimension, Year_Dimension, and Fact-2 (from Star Schema-2)
- Compare and contrast the two star schemas.

Write your answer here:

Continue your answer here:

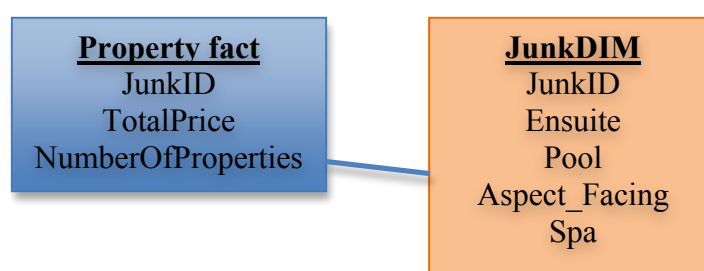
Question 2:

Using the real estate case study, assume we have a junk dimension as follows:

Table JunkDIM:

JUNKID	ENSUITE	POOL	ASPECT_FACING	SPA
1	yes	no	East	yes
2	yes	no	South	yes
3	no	no	North	yes
4	no	no	East	no
5	yes	no	North	no
6	yes	yes	South	yes
7	no	no	West	no
8	yes	yes	East	yes
9	null	no	North	no
10	no	yes	East	no
11	yes	no	South	no
12	yes	no	North	yes
13	yes	no	East	no
14	no	no	North	no
15	no	yes	South	yes
16	yes	yes	North	no
17	no	no	East	yes
18	yes	yes	South	no
19	no	no	South	yes
20	no	yes	West	no
21	yes	yes	North	yes
22	no	yes	North	yes
23	null	no	East	no
24	no	no	South	no
25	nul	no	West	no
26	no	yes	West	yes
27	yes	no	West	no
28	no	no	West	yes
29	yes	yes	West	no
30	no	yes	North	no
31	no	yes	East	yes
32	nul	no	South	no
33	yes	yes	West	yes
34	yes	no	West	yes
35	yes	yes	East	no
36	no	yes	South	no

The star schema using this junk dimension is as follows:



A tempfact table has been created as follows:

```
Create Table TempFactProperty As
Select Ensuite, Pool, Aspect_Facing, Spa,
       sum(Houseprice) as TotalPrice,
       count(*) as NumberofProperties
From dw.Property1
Group By Ensuite, Pool, Aspect_Facing, Spa;
```

A JunkId attribute is added to the TempFactProperty table using the following Later Table command:

```
Alter Table TempFactProperty
Add (JunkID Number(2));
```

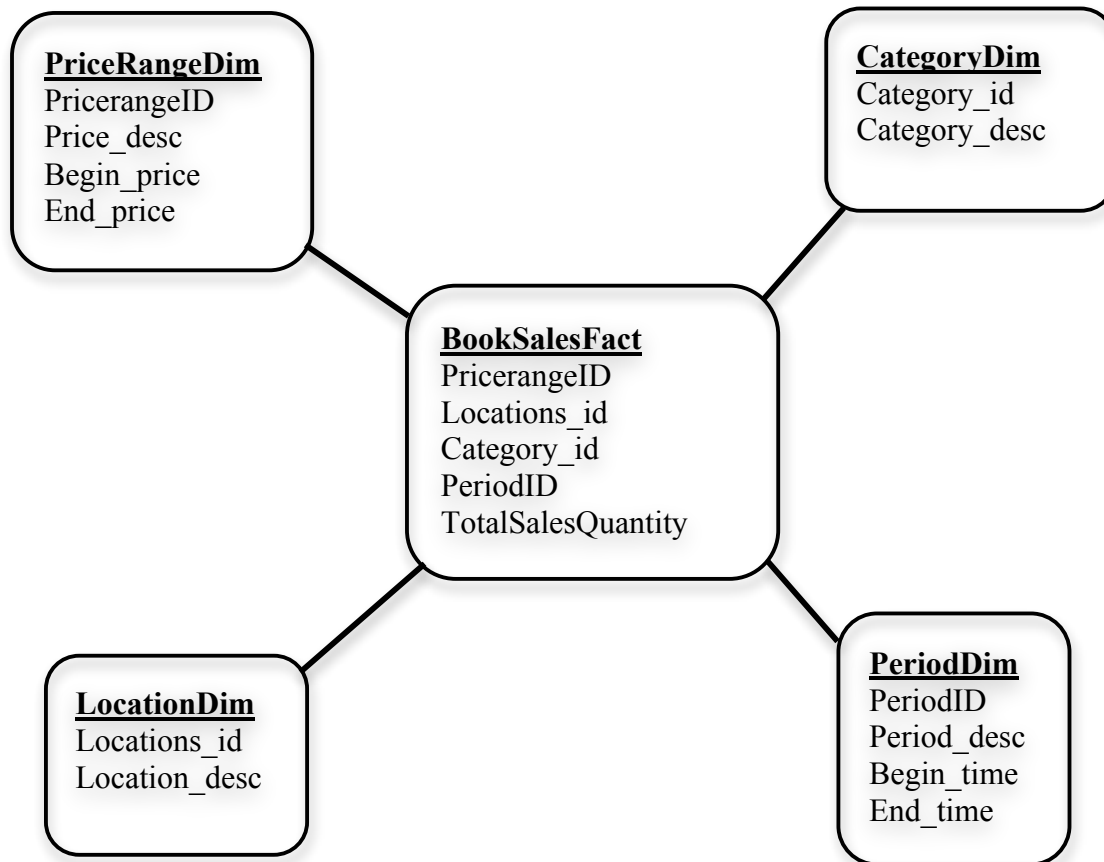
The JunkID column in TempFactProperty is still empty. There are two ways to write an update SQL command to update the JunkID column in the TempFactProperty. Option-1 is to use a cursor in which the update junkid is placed in a loop. Option-2 is to use a subquery in the update TempFactProperty set JunkID.

Question: Write the SQL code to update JunkID column in TempFactProperty to match with the JunkID column in the JunkDim using the above two options.

Write your answer here:

Question 3:

Given the following star schema for a bookshop:



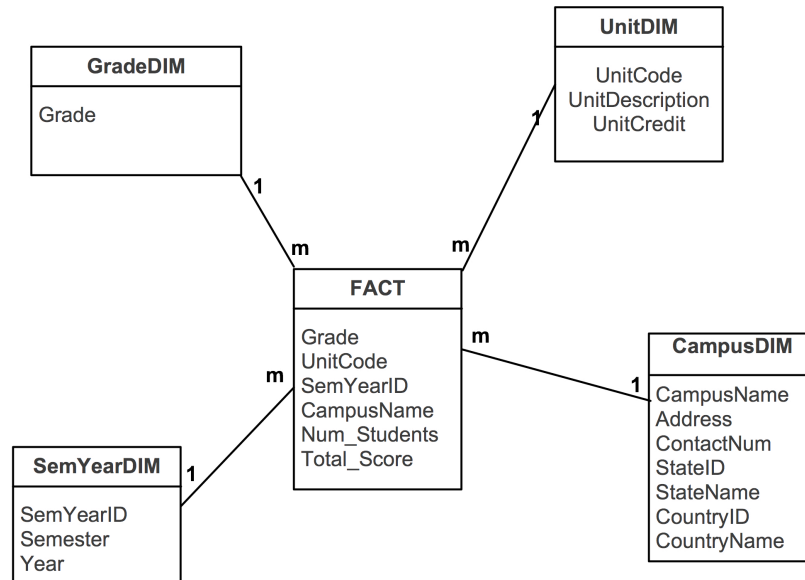
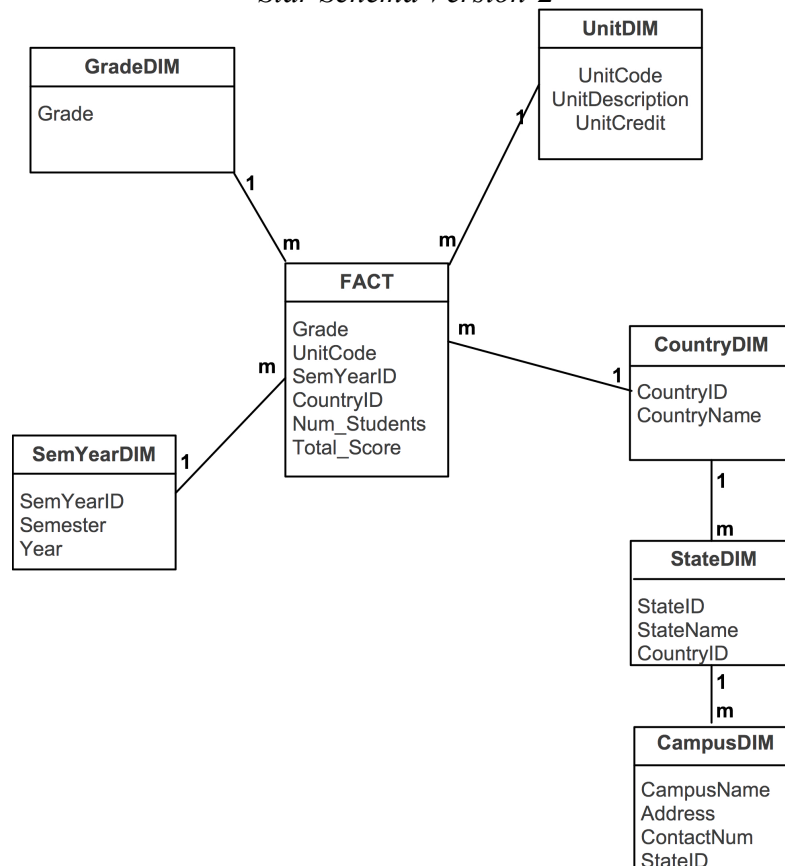
Write the SQL commands to answer the following queries:

- Show total number (quantity) of books, including subtotals, by different transaction periods and customer locations
- Show top 10% book categories from the sales
- Show top 2 book price ranges which include the most number (quantity) of sold books

Write your answer here:

Question 4:

Consider the following Student Enrolment star schemas: Star Schema Version-1 does not have a dimension hierarchy, whereas Star Schema Version-2 has a dimension hierarchy: from country to state, and to campus.

Star Schema Version-1*Star Schema Version-2*

Questions:

- a. In contrasting both star schemas, is there any mistake in any of the two star schemas (Note that Star Schema Version-1 does not have a hierarchy, and Star Schema Version-2 does have)?
 - If yes, state which star schema, and explain your reason.
 - If no, also explain your reason.
- b. Compare both star schemas.
 - If there are mistakes in any (or both) star schemas, you need to draw the correct schema(s) first before comparing between each other.
 - If there are no mistakes in both star schemas, you can immediately compare the two star schemas.

Also, when you compare the two star schemas, you need to use some sample data (in the fact and in certain dimensions) to support your arguments

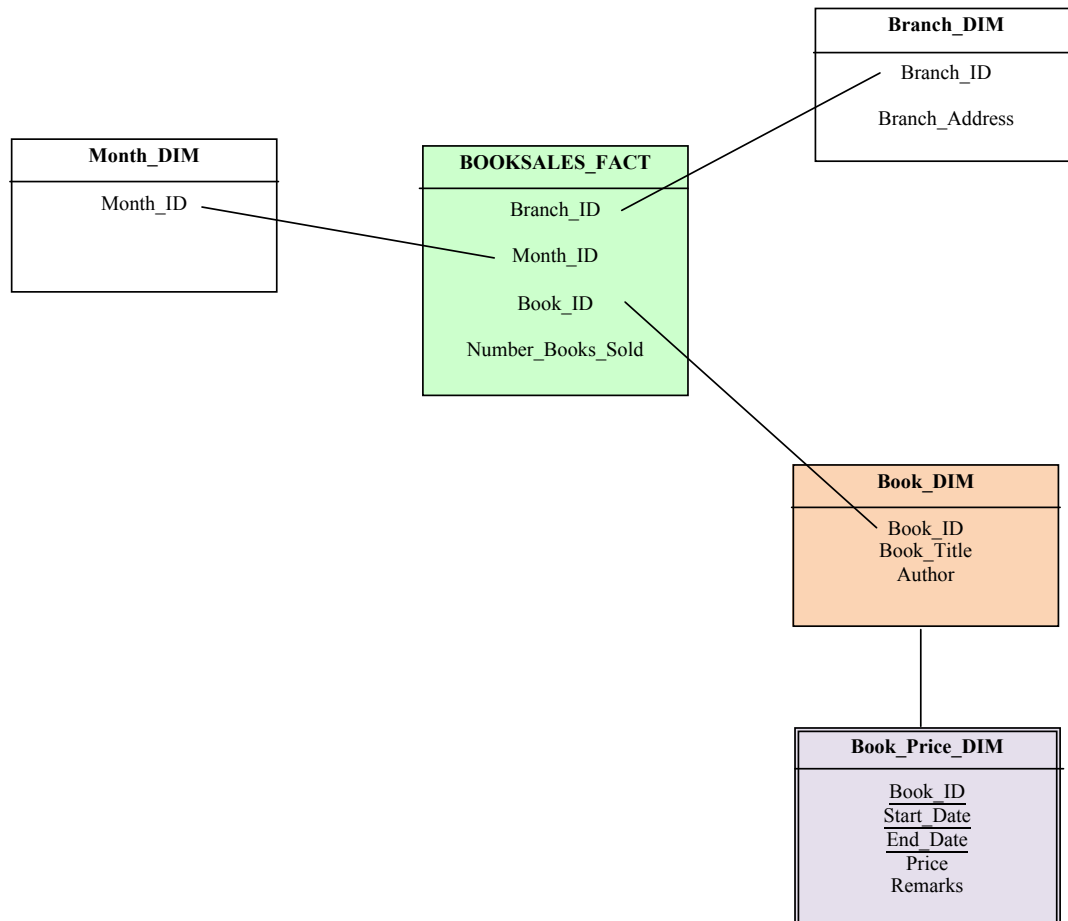
Write your answers here:

Continue your answers here:

Continue your answers here:

Question 5:

This question is taken from the *Bookshop* Case Study on Temporal Data Warehousing. The following shows a star schema shows a fact table (number of books sold) and three dimensions (e.g. Month, Branch, and Book). The Book dimension is temporal dimension, which contains a temporal attribute, called Price, which is book price.



The tables for this star schema have been created and populated from the operational database. The sample data is as follows:

Month_DIM Table

Month_ID
201503
201502
201501
201412
etc

Branch_DIM Table

Branch_ID	Branch_Address
City	Melbourne Central Shopping Centre, Melbourne
Chadstone	285 Dandenong Road, Chadstone
Camberwell	199 Burke Road, Camberwell
etc	

Book DIM Table

Book_ID	Book_Title	Author
C1	CSIRO Diet	CSIRO Team
H6	Harry Potter 6	Rowling
DV	Da Vinci Code	Dan Brown
...

Book Price DIM Table

Book_ID	Start_Date	End_Date	Price	Remarks
C1	201401	201407	\$45.95	Full Price
C1	201408	201410	\$36.75	20% Discount
C1	201411	201501	\$23.00	Half Price
C1	201502	201512	\$45.95	Full Price
H6	201401	201403	\$21.95	Launching
H6	201404	201501	\$30.95	Full Price
H6	201502	201512	\$10.00	End of Product Sale
DV	201401	201512	\$27.95	Full Price
...	

BookSales Fact Table

Month_ID	Branch_ID	Book ID	Number_Books_Sold
201503	City	C1	5
201503	City	H6	15
201503	City	DV	23
201503	City	...	
201503	Chadstone	C1	15
201503	Chadstone	H6	3
201503	Chadstone	DV	2
201503	Chadstone	...	
201503	Camberwell	C1	1
201503	Camberwell	H6	1
201503	Camberwell	DV	2
201503	Camberwell	...	
201503	
...	
201412	City	C1	15
201412	City	H6	6
201412	City	DV	6
201412	City	...	
201412	Chadstone	C1	10
201412	Chadstone	H6	8
201412	Chadstone	DV	1
201412	Chadstone	...	
201412	Camberwell	C1	18
201412	Camberwell	H6	3
201412	Camberwell	DV	2
201412	Camberwell	...	
201412	
...	

Question:

Write the SQL command to produce the following report (10 marks):

Month_ID	Branch_ID	Book_ID	Book_Title	Author	Price	Number_Books Sold
201503	City	C1	CSIRO Diet	CSIRO Team	\$45.95	5
201503	City	H6	Harry Potter 6	Rowling	\$10.00	15
201503	City	DV	Da Vinci Code	Dan Brown	\$27.95	23
201503	City		...			
201503	Chadstone	C1	CSIRO Diet	CSIRO Team	\$45.95	15
201503	Chadstone	H6	Harry Potter 6	Rowling	\$10.00	3
201503	Chadstone	DV	Da Vinci Code	Dan Brown	\$27.95	2
201503	Chadstone		...			
201503	Camberwell	C1	CSIRO Diet	CSIRO Team	\$45.95	1
201503	Camberwell	H6	Harry Potter 6	Rowling	\$10.00	1
201503	Camberwell	DV	Da Vinci Code	Dan Brown	\$27.95	2
201503	Camberwell		...			
201503			
...			
...			
201412	City	C1	CSIRO Diet	CSIRO Team	\$23.00	15
201412	City	H6	Harry Potter 6	Rowling	\$30.95	6
201412	City	DV	Da Vinci Code	Dan Brown	\$27.95	6
201412	City		...			
201412	Chadstone	C1	CSIRO Diet	CSIRO Team	\$23.00	10
201412	Chadstone	H6	Harry Potter 6	Rowling	\$30.95	8
201412	Chadstone	DV	Da Vinci Code	Dan Brown	\$27.95	1
201412	Chadstone		...			
201412	Camberwell	C1	CSIRO Diet	CSIRO Team	\$23.00	18
201412	Camberwell	H6	Harry Potter 6	Rowling	\$30.95	3
201412	Camberwell	DV	Da Vinci Code	Dan Brown	\$27.95	2
201412	Camberwell		...			
201412			
...			

The structures of the above tables are as follows:

SQL> desc Month_DIM;

Name	Null?	Type
MONTH_ID		VARCHAR2 (6)

SQL> desc Branch_DIM;

Name	Null?	Type
BRANCH_ID		VARCHAR2 (15)
BRANCH_ADDRESS		VARCHAR2 (50)

SQL> desc Book_DIM;

Name	Null?	Type
BOOK_ID		VARCHAR2(5)
BOOK_TITLE		VARCHAR2(20)
AUTHOR		VARCHAR2(20)

SQL> desc Book_Price_DIM;

Name	Null?	Type
BOOK_ID		VARCHAR2(5)
START_DATE		VARCHAR2(6)
END_DATE		VARCHAR2(6)
PRICE		NUMBER(6,2)
REMARKS		VARCHAR2(20)

SQL> desc BookSales_Fact;

Name	Null?	Type
MONTH_ID		VARCHAR2(6)
BRANCH_ID		VARCHAR2(15)
BOOK_ID		VARCHAR2(5)
NUMBER_BOOKS_SOLD		NUMBER

Write your answer here:

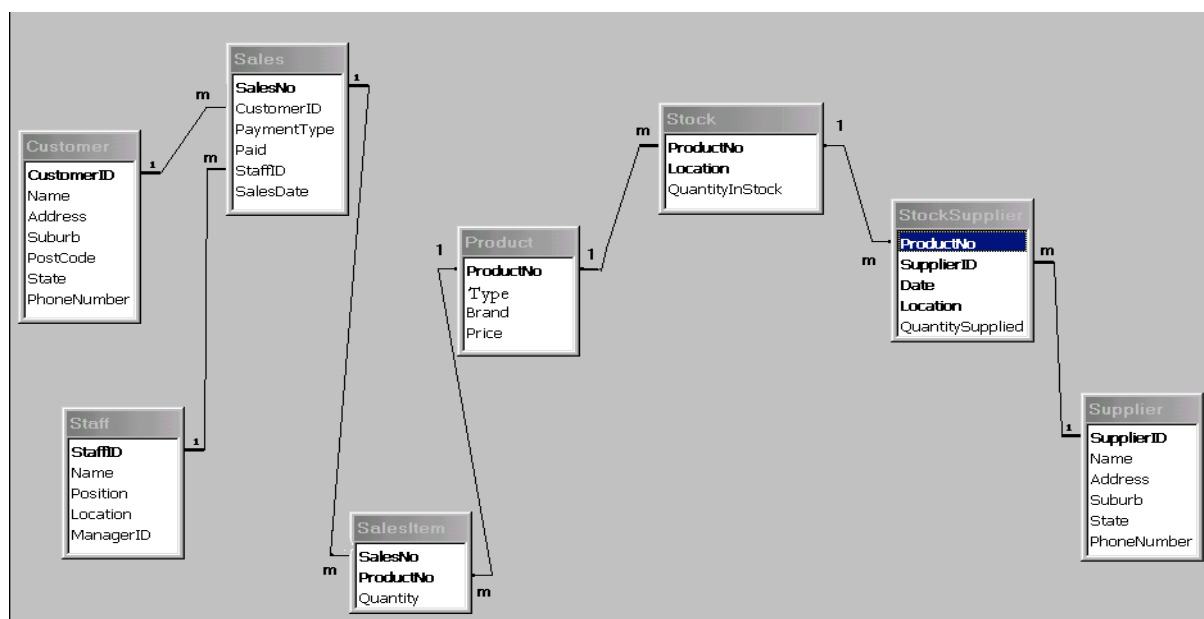
Question 6:

This question is taken from the *Product-Sales-Supplier* Case Study.

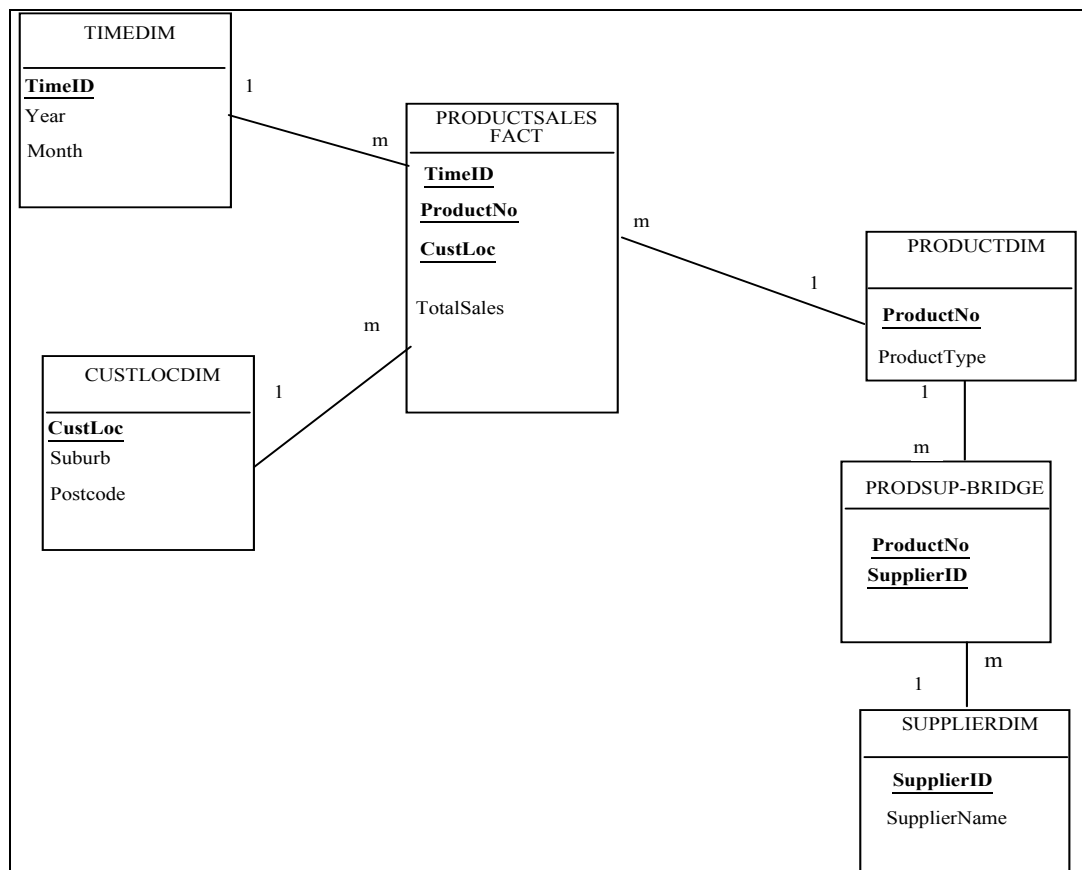
The director of a company is interested in analyzing the statistics of its product sales history. The analysis is needed for identifying which products are popular, which suppliers supply those products, when is the best time to purchase more stock, etc. You are required to design a small Data Warehouse to keep track of the statistics.

The director is particularly interested in analyzing the *total sales* (Quantity * Price) by *product*, *customer locations* (suburbs and postcodes), *sales time periods* (monthly and yearly), and *supplier*.

The operational database currently has the following tables:



Your snowflake schema will have a Bridge Table connecting Product Dimension and Supplier Dimension. A snowflake schema with a Bridge Table as shown below:



The above snowflake schema is missing two attributes: **WeightFactor** attribute, and **ListAGG** attribute.

Questions:

- Draw a new snowflake schema** (call it Snowflake Schema version 2) for the above case study, but this new snowflake schema must **use a WeightFactor attribute (without ListAGG attribute)**. You also need to **show sample records** in the Product Dimension, the Bridge Table, and the Supplier Dimension. The sample data must show the correct values for the Weight attribute. Make sure that in your snowflake schema, the attributes are clearly shown.
- Draw another snowflake schema** (call it Star Schema version 3), which also has a Bridge Table and a WeightFactor attribute. But version-3 snowflake schema has the **ListAGG** attribute. You also need to **show sample records** in the Product Dimension, the Bridge Table, and the Supplier Dimension. The sample data must show the correct values for the Weight and ListAGG attributes.
- Write the **SQL query** to create the ProductDim table for the Star Schema version 3.

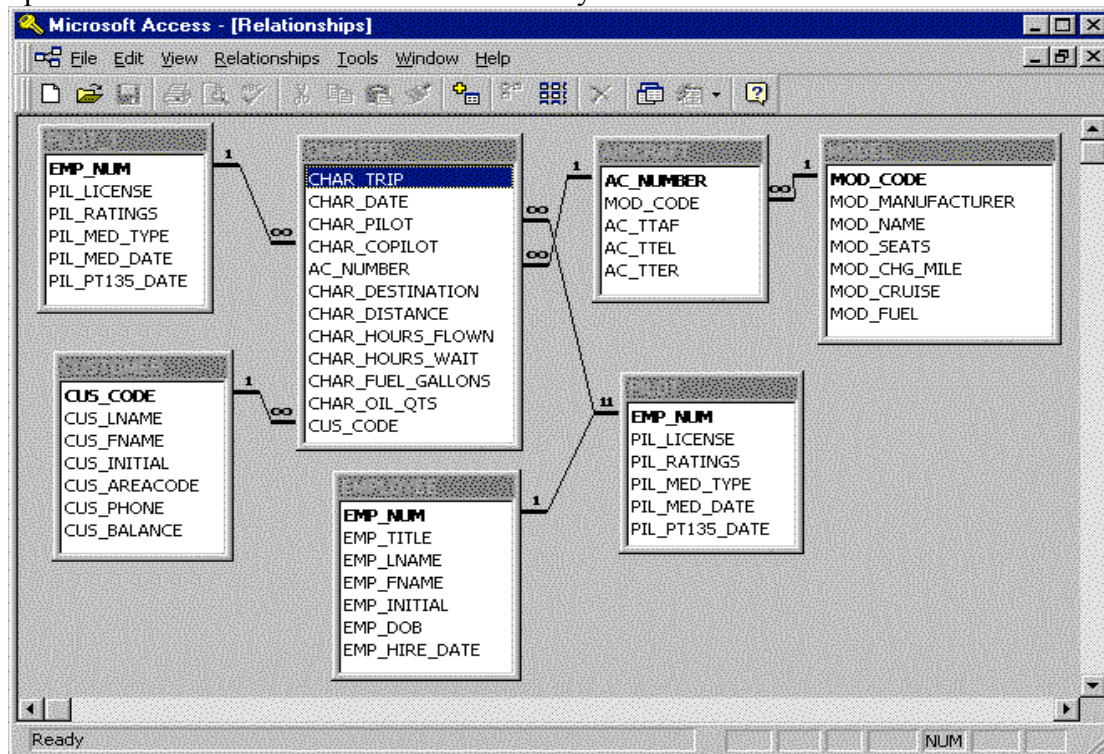
Write your answer here:

Continue your answer here:

Continue your answer here:

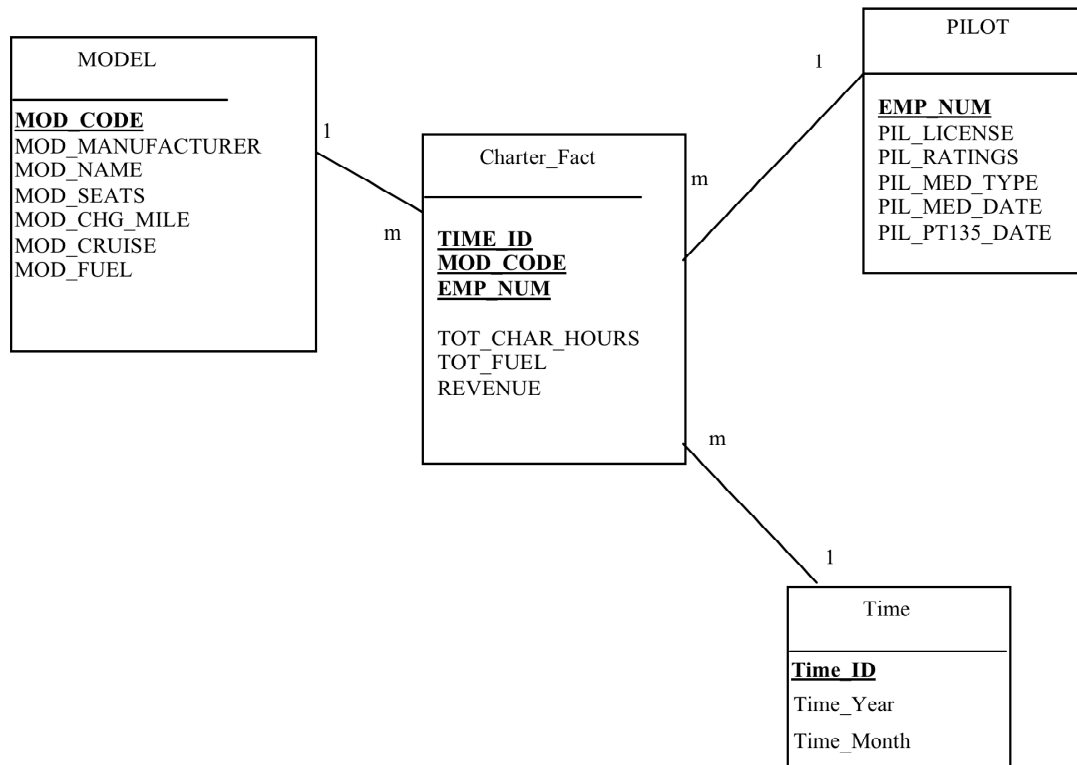
Question 7:

This question is based on the Robcor case study. The following is the E/R diagram of the operational database in the Robcor case study:



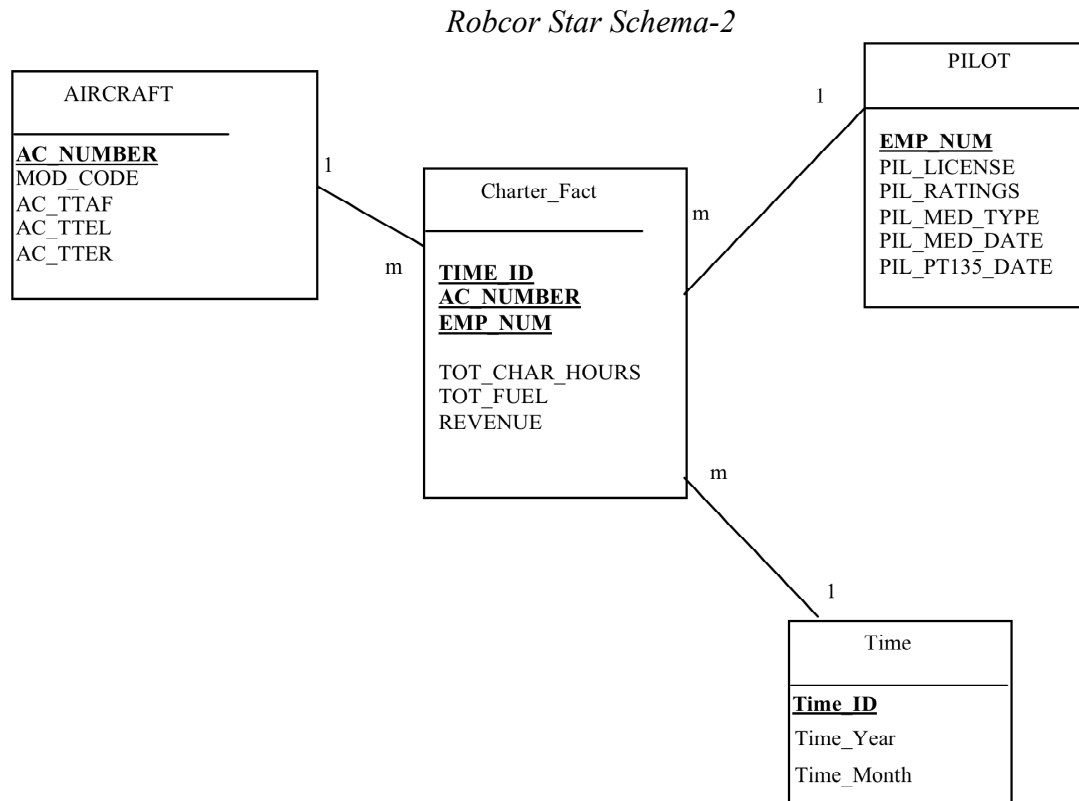
A star schema for the above operational database is shown as follows:

Robcor Star Schema-1



Questions:

- a. Is it possible to determine which level Robcor Star Schema-1 is? If it is possible, state the level and also give the reason. If it is not possible to state the level, then give the reason.
- b. Let's have a look at the following star schema (Robcor Star Schema-2). Between the two star schemas (Robcor Star Schema-1 and Robcor Star Schema-2), which one has a higher level of aggregation? State the name of the star schema, whether it is Robcor Star Schema-1 or Robcor Star Schema-2, and explain the reason.



Write your answers here:

Question 8:

Suppose an operational database contains one table, called *Student*.

Table *Student*

StudentID	StudentName	Suburb	Postcode	Sex
21001	Adam	Caulfield	3162	M
21003	Ben	Caulfield	3162	M
21008	Christine	Chadstone	3148	F
21019	Daisy	Caulfield	3162	F
21033	Edward	Clayton	3168	M
21122	Fred	Caulfield	3162	M
21123	Greg	Chadstone	3148	M

Suppose the star schema contains only two dimensions: SuburbDIM and SexDIM. The fact table has one fact measure, called TotalStudents.

SuburbDIM table can be created using the following SQL command (assume that the operational database is located at dtaniar account):

```
CREATE TABLE SuburbDIM as
SELECT distinct Suburb, Postcode
FROM dtaniar.Student1;
```

Table *SuburbDIM*

Suburb	Postcode
Caulfield	3162
Chadstone	3148
Clayton	3168

SexDIM table can be created the same way:

```
CREATE TABLE SexDIM as
SELECT distinct Sex
FROM dtaniar.Student1;
```

Table *SexDIM*

Sex
M
F

And the Fact table can be created using the following SQL command:

```
CREATE TABLE Fact as
SELECT Suburb, Sex, Count(*) as TotalStudents
FROM dtaniar.Student1
GROUP BY Suburb, Sex;
```

Table *Fact*

Suburb	Sex	TotalStudents
Caulfield	M	3
Caulfield	F	1
Chadstone	M	1
Chadstone	F	1
Clayton	M	1

Suppose we want to implement a surrogate key in the SuburbDIM, which we call SuburbID, for example.

Table *SuburbDIM*

SuburbID	Suburb	Postcode
1	Caulfield	3162
2	Chadstone	3148
3	Clayton	3168

Questions:

(a) Write the SQL query to create SuburbDIM to include a surrogate key called SuburbID as shown in the same data above.

(b) Supposed a tempfact table has been created using the following SQL query:

```
CREATE TABLE TempFact as
  SELECT Suburb, Sex FROM dtaniar.Student1;
```

Write an alter and an update command to alter the Tempfact to include a surrogate key (SuburbID).

(c) Write an SQL query to create a fact table so that the fact table will have the following contents:

Table *Fact*

SuburbID	Sex	TotalStudents
1	M	3
1	F	1
2	M	1
2	F	1
3	M	1

Write your answer here:

THE END