COMMONWEALTH OF AUSTRALIA

Copyright Regulations 1969

WARNING

This material has been reproduced and communicated to you by or on behalf of Monash University pursuant to Part VB of the Copyright Act 1968 (the Act).

The material in this communication may be subject to copyright under the Act. Any further reproduction or communication of this material by you may be the subject of copyright protection under the Act.

Do not remove this notice.

Week 1

Case Study #1
(STUDENT ENROLLMENT)

Assoc. Prof. David Taniar

Contents

1. Star Schemas

The Star Schema is a data modeling technique used to map multidimensional decision support data into a relational database.

The reason for the star schema's development is that existing relational modeling techniques: ER and normalization, did not yield a database structure that served the advanced data analysis requirements well.

There are **Three** main components of the Star Schema:

1. Facts
2. Dimensions
3. Attributes

A Simple Star Schema

1. Facts

Facts are **numeric measurements** (values) that represent a specific business aspect or activity.

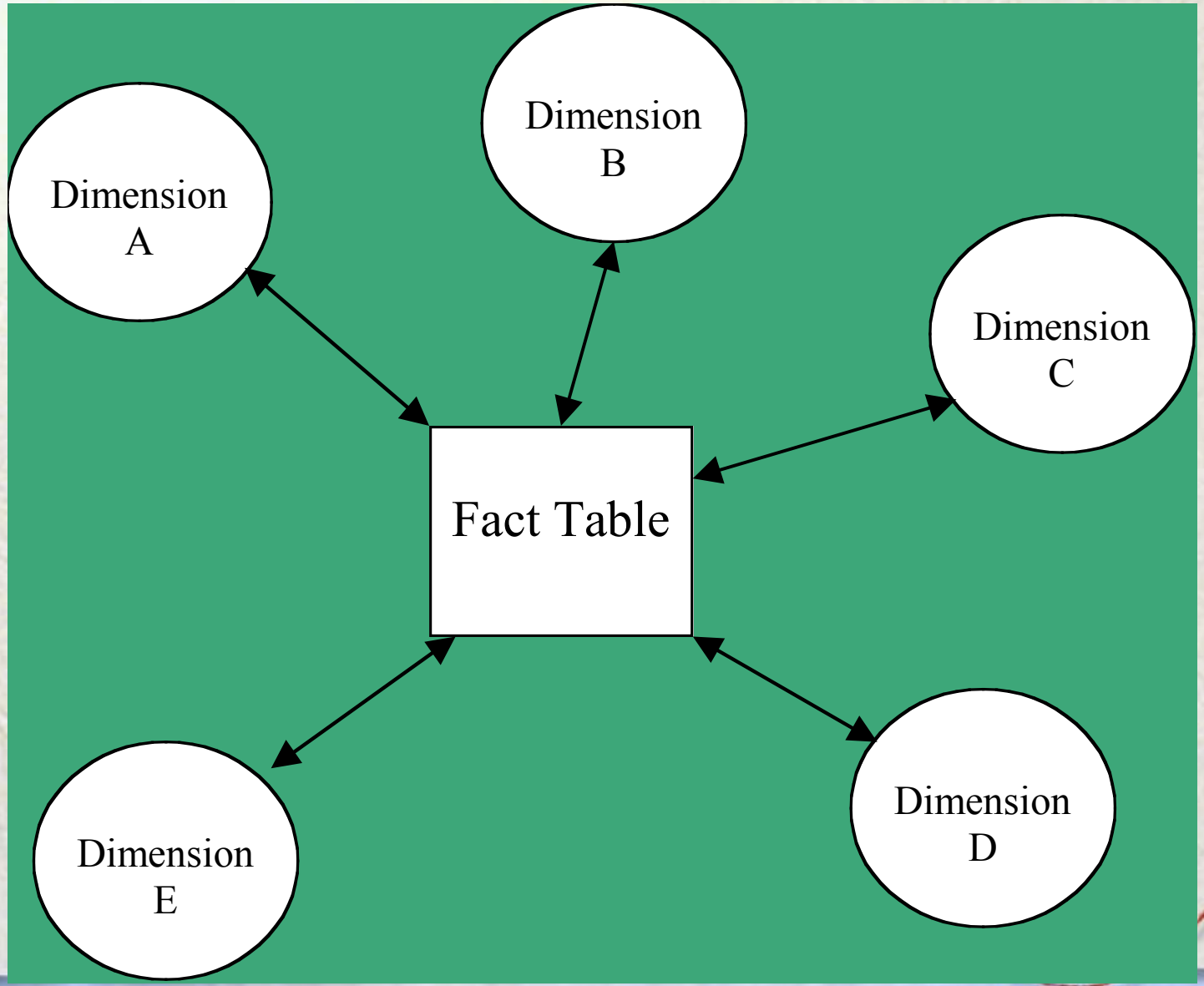
For example, sales figures are numeric measurements that represent product and/or service sales.

2. Dimensions

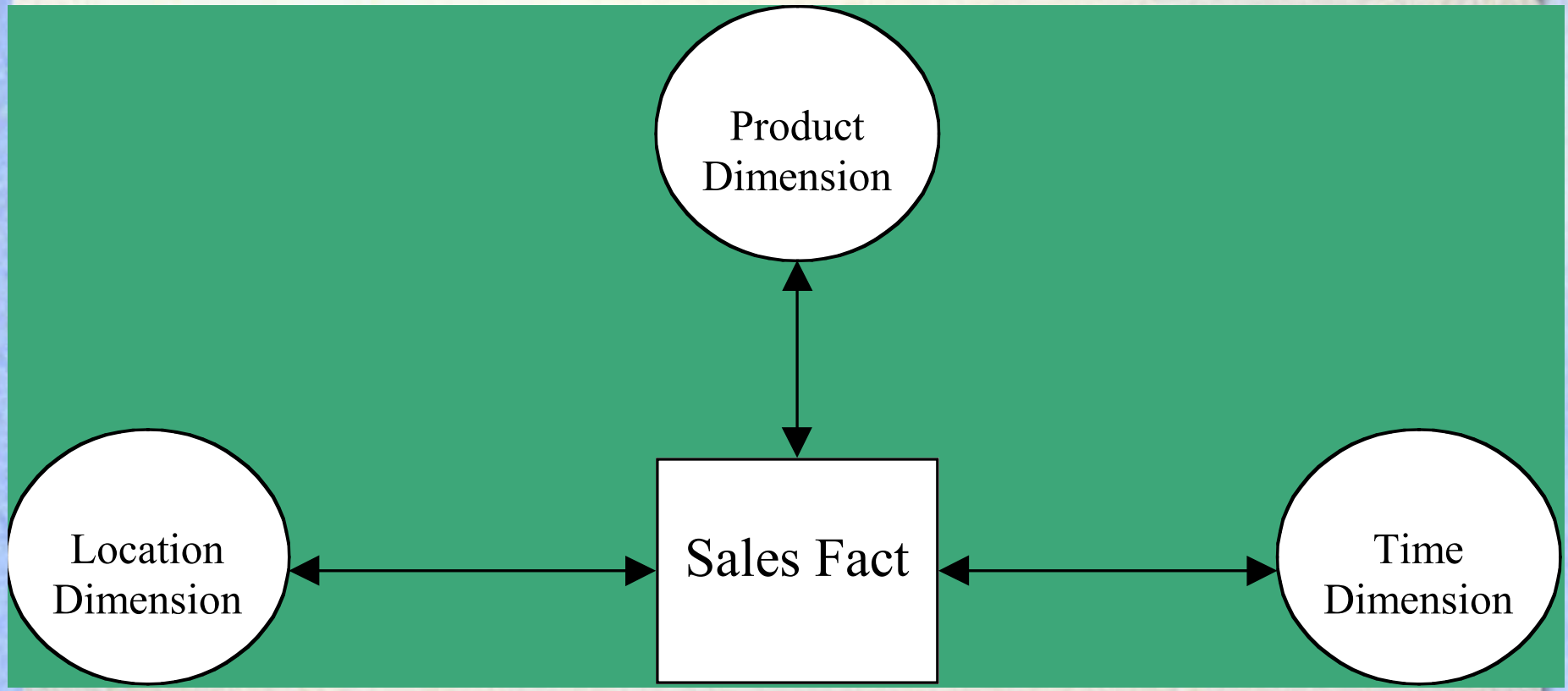
Dimensions are qualifying characteristics that provide **additional perspectives** to a given fact.

For example, sales might be *viewed* from specific dimension(s), such as sales location, sales period, sales product, etc.

A Simple Star Schema



A Simple Sales Star Schema



Star Schemas

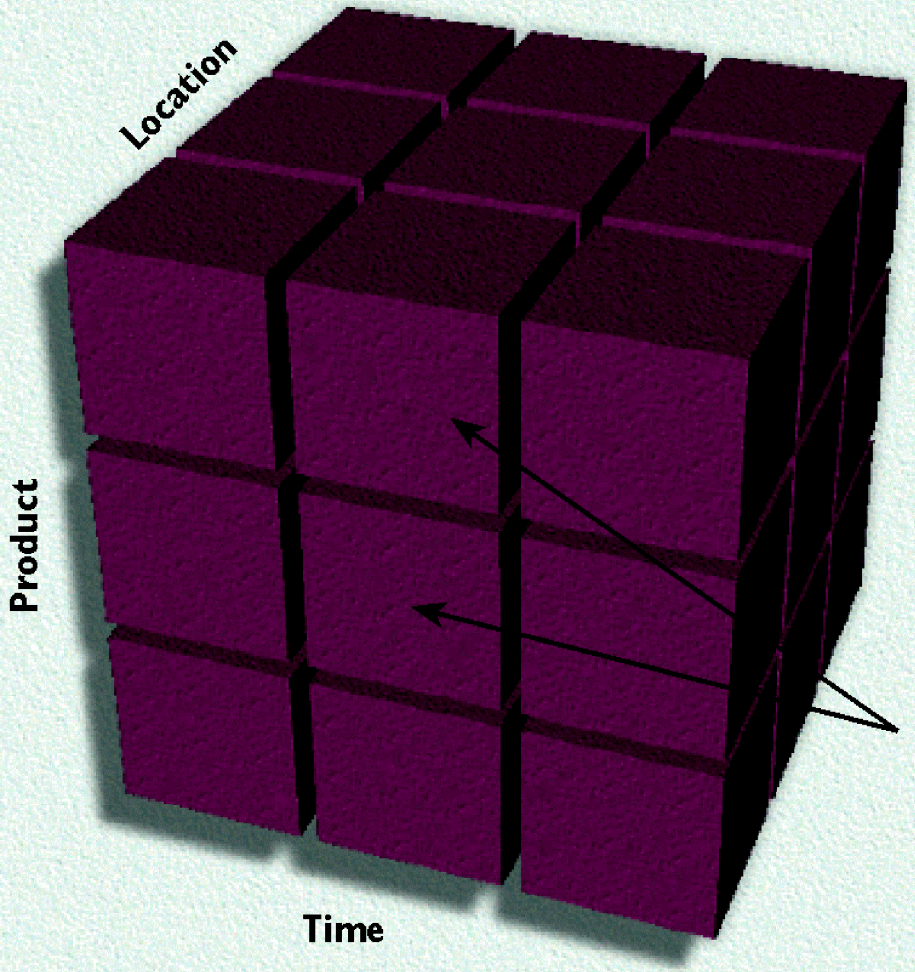
3. Attributes

Each dimension table contains attributes.

For example:

Product dimension:	Prod Type, Description.
Location dimension:	Region, State, City.
Time dimension:	Year, Month.

Star Schemas



Conceptual three-dimensional cube of sales by product, location, and time.

Sales facts are stored in the cells at the intersection of each product, time, and location dimension.

Contents

2. A Case Study

The University Administrator(s) needs to keep track of the number of enrollment for particular unit or campus and the students' performance each year in order to maintain the University performance. The head of admin has assigned you the task of developing a small Data Warehouse in which to keep track the enrollment and performance statistics.

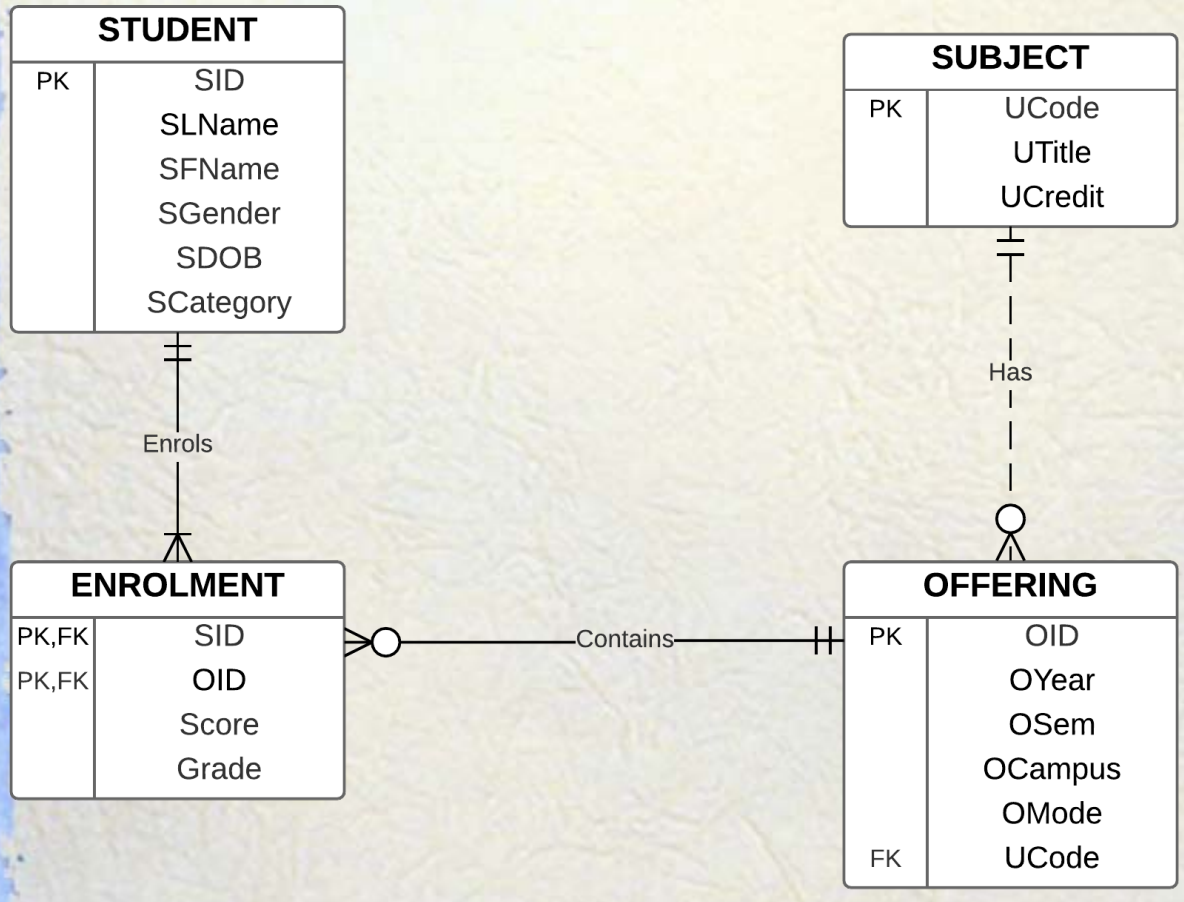
The *main requirements* for this database are to:

1. How many students enrolled in the *Database unit* offered in the *Main campus*?
2. What is the total score of students taking the Database unit in the Main campus?
3. How many students enrolled in the Java unit offered in *Semester 2, 2009*?
4. What is the total score of students taking the Java unit offered in Semester 2, 2009?
5. How many students received HD in the SAP unit offered in Semester 1, 2009?

A Case Study: ER Diagram

Use the provided database which includes the four following tables:

a. STUDENT
b. OFFERING
c. SUBJECT
d. ENROLLMENT



A Case Study: The Data (Student Table)

Table Student

SID	Slname	Sfname	Sgender	Sdob	Scategory
10001	Tan	Mirriam	F	19-Jul-81	112
10002	Murray	Juan	M	10-Jun-83	211
10003	Lay	Andy	M	19-Feb-86	211
10004	Wright	Allan	F	29-Jan-83	211
10005	Simon	Ally	F	24-Aug-83	112
10006	Smith	Ben	M	9-Jul-87	211
10007	Brown	Kate	F	19-Oct-72	112
10008	Miller	Larry	M	22-Jul-73	211
10009	Smith	Leonard	M	26-May-85	211
10010	Brown	Menson	M	12-Jul-83	112

A Case Study: The Data (Subject & Offering)

Table Subject

Ucode	Utitle	Ucredit
IT001	Database	5
IT002	Java	5
IT003	SAP	10
IT004	Network	5
IT005	ASP.net	5

Table Offering

OID	Oyear	Osem	Ocampus	Omode	Ucode
1	2009	1	Main	D	IT001
2	2009	2	City	E	IT001
3	2009	2	DE	E	IT004
4	2009	2	Main	D	IT002
5	2009	1	City	E	IT003
6	2009	1	Main	E	IT002
7	2010	1	Main	D	IT001
8	2010	2	City	E	IT001
9	2010	2	DE	E	IT004
10	2010	2	Main	D	IT002
11	2010	1	City	E	IT003
12	2010	1	Main	E	IT002

A Case Study: The Data (Enrollment Table)

Table Enrollment

SID	OID	Score	Grade
10001	1	81	HD
10001	4	78	D
10002	2	64	C
10002	3	53	P
10003	2	32	N
10004	1	41	N
10005	5	63	C
10006	4	73	D
10006	1	74	D
10007	1	85	HD
10008	1	87	HD
10008	4	64	C
10009	1	75	D
10010	3	52	P
10005	6	65	C
10010	6	47	N

A Case Study

Using the provided data in the tables, complete the following problems.

1. Define the main **fact** to be analyzed,
2. Define and describe the possible **dimensions**,
3. Draw the student enrollment star schema, and
4. Define the **attributes** for each of the dimensions in question 2 above.

Your task is to design the data warehouse and to implement it in the lab.

End of Case Study #1 (STUDENT ENROLLMENT)