INFO20003 Database Systems

Dr Renata Borovica-Gajic

Lecture 01

What are Database Systems?

Week 1

Data vs Information

- Data
  - known facts stored and recorded
  - can include: text, numbers, dates, plus images, sound, video, and other complex objects
- Information
  - Data presented in context (can be summarised data)
  - Data that has been processed increasing the users knowledge
- Data vs Information
  - *Data* is known and available; *Information* is processed and more useful

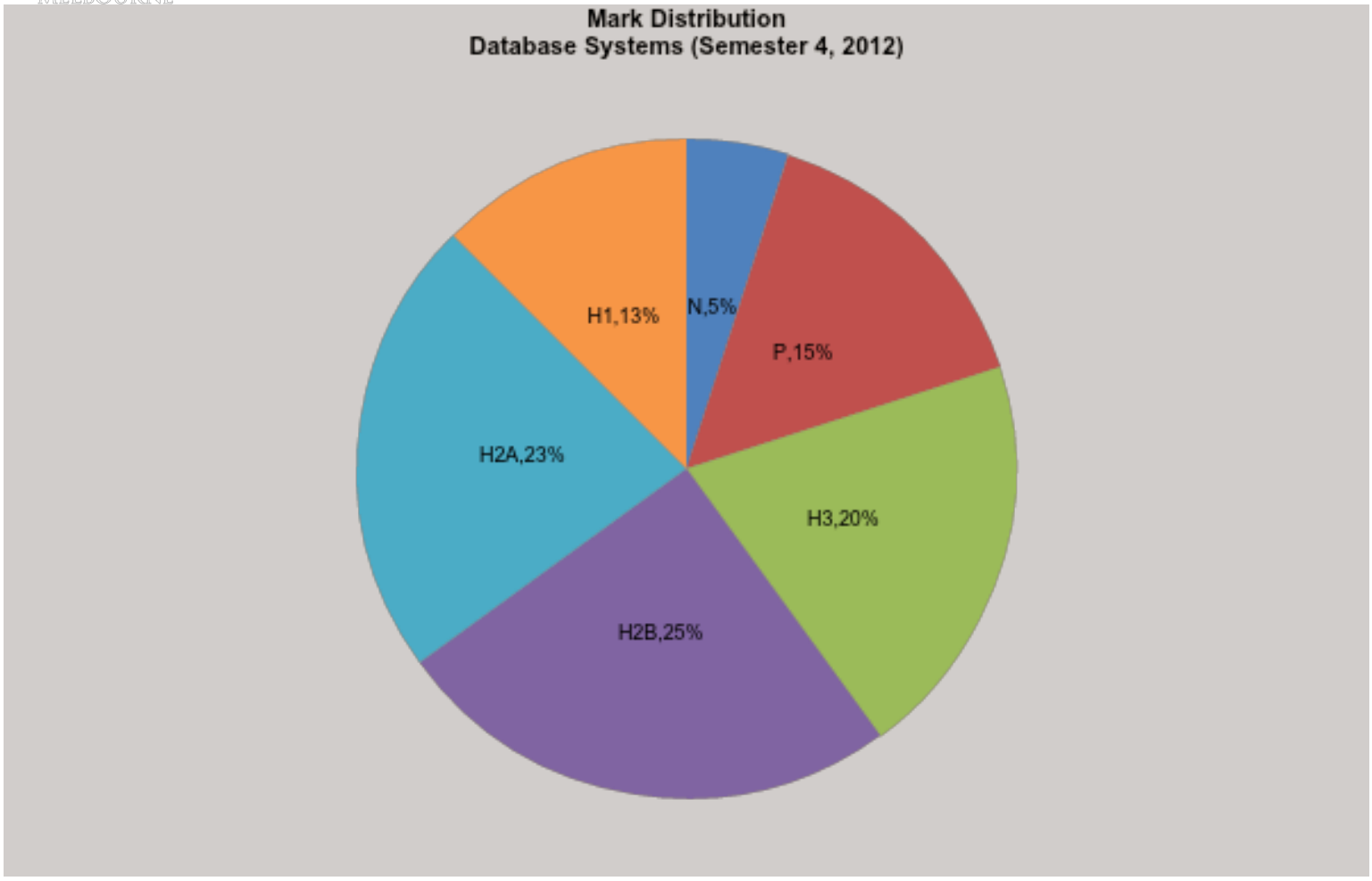
Baker, Kenneth D.	324917628
Doyle, Joan E.	476193248
Finkle, Clive R.	548429344
Lewis, John C.	551742186
McFerran, Debra R.	409723145
Sisneros, Michael	392416582

Information: data in context

Database Systems Assignment 4 Marks
Semester 3 2014

Student Name	Student ID	Grade
Baker, Kenneth D.	324917628	H1
Doyle, Joan E.	476193248	H2B
Finkle, Clive R.	548429344	H3
Lewis, John C.	551742186	H2A
McFerran, Debra R.	409723145	P
Sisneros, Michael	392416582	H3

Information: Summarisation!



Metadata - Data about data

Name	Type	Length	Description
Course	Alphanum	30	Course ID
Tutorial	Integer	2	Tutorial number
Name	Alphanum	30	Student name

- Can include:
  - structure, rules, constraints
- Why do we need Metadata?
  - Consistency
  - Meaning
- We generate a **data dictionary** as part of the analysis of system requirements

data is nice and clean eg. when ask tutorial, people may say is tutorial number
Mondy 11 a.m. or workshop1
→ give meaning to an attribute eg. the meaning of tutorial
⇒ not precise what I expect to get

What is a "database" ?

database

a large, integrated, structured collection of data

- **Usually intended to model some real-world enterprise**
- **Example: a university**
  - *Entities* … such as courses, students, professors
  - *Relationships* … such as enrollment, teaching

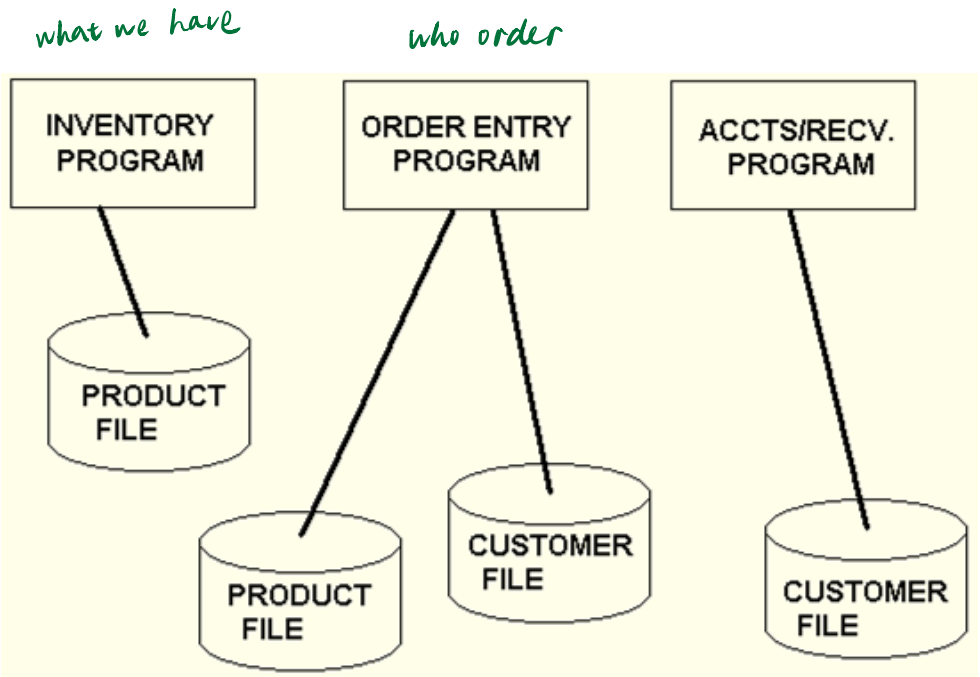
What Is a Database System?

A *Database Management System (DBMS)* is a software system designed to store, manage, and facilitate access to databases.

DBMS like a database system manager

set of programs which you write so that you can easily manipulate, store, retrieve data. (to interact with database)

File Processing Systems



- What are the problems you can see with this?
- (Diagram adapted from Hoffer p. 42)

What's Wrong Here?

- Program-data dependence
  - If the file structure changes, so does the program
  - What if you change data structure for one program
- Duplication of data
  - wasteful, inefficient, loss of data integrity
- Limited data sharing
  - data tied to application, hard/slow to create adhoc reports
- Lengthy development times
  - application has to do low level data management, figure out file format each time
- Excessive program maintenance
  - up to 80% of development time in traditional file based organisations is for maintenance

Database Systems as Solution

- Manage data in a **structured** way
- Many models (hierarchical, network, etc), but **relational** dominant since ~1980
  - Relational Model
    - –Rows & Columns forming Relations
    - –Keys & Foreign Keys to link Relations

Enrolled

sid	cid	grade
53666	Carnatic101	5
53666	Reggae203	5.5
53650	Topology112	6
53666	History105	5

Students

sid	name	login	age	gpa
53666	Jones	jones@cs	18	5.4
53688	Smith	smith@eecs	18	4.2
53650	Smith	smith@math	19	4.8

Database Advantages

- Data independence
  - separation of data and program, application logic
  - central data repository, central management
- Minimal data redundancy
  - redundancy can be controlled (normalization)
- Improved data consistency
  - single store: no disagreements, update problems, less storage space
- Improved data sharing
  - data is shared, a corporate resource, not a necessity for an application
  - external users can be allowed access
  - multiple views of data, arbitrary views of data
- Reduced program maintenance
  - data structure can change without application data changing
- Novel ad hoc data access 'without programming'
  - SQL

What's Examinable?

- Difference between Data and Information
- Being able to discuss the advantages of Databases vs File Processing Systems

data → fact

information → data put into context
or summarize data

DBMS used to maintain, and query large datasets
- can manipulate data and exploit semantics

Other benefits include
- recovery from system crashes
- concurrent access
- quick application development
- data integrity and security

Next Lecture

- The database system lifecycle
  - With a focus on the design stage
    - Conceptual design
    - Logical design
    - Physical design