



# COMP1140: Database and Information Management

---

Lecture Note – Week 01

*Dr Suhuai Luo*

School of SEEC

University of Newcastle



# This lecture...

---

- Introduction to course
- Introduction to DBMSs
- DBMS in Perspective
- Database Design Process
- More on Requirements Gathering



# Introduction to course - Teaching Group

## ■ Coordinator/Lecturer/Tutor:

- Suhuai Luo
- Email: [suhuai.luo@newcastle.edu.au](mailto:suhuai.luo@newcastle.edu.au)
- Office & Phone: ICT3017, 49854508
- Contact times: Tue 2-3 pm (appointment)

## ■ Other members:

- Ananda Kalutantiri ([kalutantiri@newcastle.edu.au](mailto:kalutantiri@newcastle.edu.au))
- Hayden Cheers ([hayden.cheers@uon.edu.au](mailto:hayden.cheers@uon.edu.au))
- Daniel Bell ([daniel.p.bell@newcastle.edu.au](mailto:daniel.p.bell@newcastle.edu.au))



# Introduction to course

---

## Objective...

- Design conceptual models encapsulating data requirements for business and organisational scenarios
- Convert conceptual models to logical data models
- Implement a database solution using contemporary database management systems (DBMS)
- Develop expertise in database languages (e.g. SQL) including the ability to develop sophisticated queries to extract information from large datasets
- Understand data security and data quality management



# Introduction to course

---



[www.datacentreknowledge.com](http://www.datacentreknowledge.com)

- What is database:
  - Collection of related data and information
  
- Database examples:
  - A **university**
  - A **bank**
  - A **bookstore**



# Introduction to course

---



[www.datacentreknowledge.com](http://www.datacentreknowledge.com)

- Applications of database:
  - Store and manage conventional data (e.g., bank statement)
  - Store and manage multimedia data (e.g., audio, video of surveillance camera; YouTube)
  - Store and retrieve Web data (e.g., Google)
  - Store and analyse big data (e.g., customer purchasing behaviour)
  - Monitor system, automatically take action when required (e.g., warning of bank transaction)

# Introduction to course - Why Database

DB is an important part of IT area skills

- Technical
  - Business domain
  - Project management
- 
- DB is part of your degree footpath
  - It is the market and direction
    - Job; research; products; big data; ...



Tech-kid.com

# Introduction to course –

## What you'll get/learn out of the course?

- Theoretical and practical knowledge of database design approaches
- Job-ready skills in DB design
- Potential/capability to work in BI







# Introduction to course.

---



[Toolkit.smallbiz.nsw.gov.au](http://Toolkit.smallbiz.nsw.gov.au)

- How to achieve objectives?
  - Lectures provide...
    - Necessary background knowledge
  - Pracs/Tutes provide...
    - Hands-on application of knowledge acquired
    - Will design, develop and implement a database for an application domain following the database design process

# Introduction to course

- Evaluation of learning:

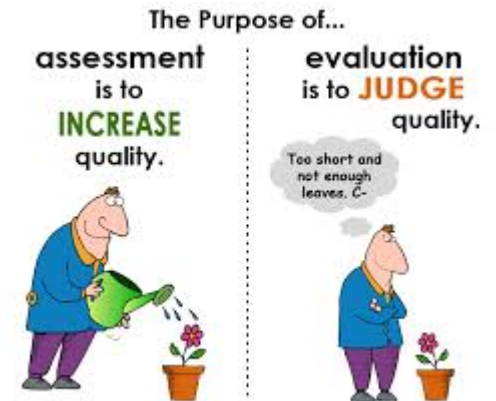
- Assessments

- Assignment -

- 3 assignments , subtotal 35%
      - Beware of lateness and plagiarism – read the university policies & course outline

- Examinations –

- Test – lab test in Week 9, 15%
      - Final exam, 50%



# Introduction to course...

## Assignments

---

- Popular Pizza Project (individual work)
  - Assignment 1 – Description & EER model, due Wk5 , 10%
  - Assignment 2 - Logical Database Design, due Wk8, 10%
  - Assignment 3 - Physical Database Design, due Wk12, 15%
- Submission:
  - Both hardcopy & softcopy
  - Submit before lecture/lab on the due week
    - Softcopy: submit via BB
    - Hardcopy: hand in to lecturer/instructor
  - You MUST be present at the lab of the week 12 to have your A3 marked



# Introduction to course...

- Lectures: Tuesday 10am – 12pm, GP101

- Labs/Tutorials:

■ Tuesday	12pm - 2pm	CT306
■ Tuesday	12pm - 2pm	CT304
■ Tuesday	4pm - 6pm	CT304
■ Tuesday	6pm - 8pm	CT311
■ Wednesday	8am - 10am	CT218
■ Wednesday	5pm - 7pm	CT218
■ Thursday	2pm-4pm,	ES205
■ Thursday	4pm- 6pm,	CT218
■ Friday	1pm – 3pm	CT314
■ Friday	1pm – 3pm	CT220
■ Friday	3pm – 5pm	CT220

- Download SQL Server from DreamSpark is arranged



# Introduction to course...

---

- Blackboard (UoNline) Use
  - for communication, messaging and updates.
  - Lecture notes will be available on Blackboard (but details as per class lecture)
  - Announcements and Assignments will be available for download from Blackboard.
  - Assessments submission
  - Check it regularly along with your student email account. Do not use external email accounts.



# Introduction to course...

---

- Reference Materials:
  - Text:
    - Thomas Connolly, Carolyn Begg, *Database Systems: A Practical Approach to Design, Implementation & Management*, Addison Wesley, 6th Edition (Global), 2015, ISBN-10 1292061189, ISBN-13 9781292061184
  - Other references in course outline
- Course outline provides details...
- Course content
- Qs?

[Course](#) Outline



# Introduction to DBMSs

---

- What is data, information, databases, and DBMS?
- Requirements of today's information systems
- DBMSs and its features
- Files vs. DBMSs
- When to use DBMSs? When not to use DBMSs?



# Data, Information, Databases and DBMSs

---

- **Data** are facts.  
E.g. John Stuart
- Data are processed to create **information**.  
E.g. John Stuart's grades

Name:	John Stuart
Subject/Grades:	INFT2040    A
	INFT1004    A-
	...





# Data, Information, Databases and DBMSs (contd.)

---

- **Database**

Collection of related data and information.

- **Database Management System**

a software package that is used to create, access and maintain databases



# Organisations, data and information

---


- Every organisation collects and manipulates data and information
  - Examples.
    - A **university** consists of a collection of student names, student grades and courses offered
    - A **bank** consists of information of a collection of accounts, customers and transactions
    - A **bookstore** consists of books, pricing information, data items sold and stocks available
    - **Human resource:** consists of information about employees, salaries, payroll taxes and benefits, and generation of paychecks
- As the list illustrates, databases form an essential part of almost all enterprises/daily life today



# Today's information systems

---

- Today's information systems collect and manipulate large amount of data.
  - Example
    - The university's student information management system



# Requirements of today's Information Systems

---

- Store large amount of data
  - Example: bank database requires to collect, store and manipulate data of thousands of customers, millions of transactions, ...
- Efficient (fast) access & manipulation of data
  - E.g.: Customers of a bank wants to access their account information, transfer funds, pay bills , etc. in real-time
- Concurrent access by a large number of users
  - E.g.: Many customers, bank staff, vendors may simultaneously access the bank's database
- Security
  - E.g.: Bank's database must be highly secure prohibiting unauthorised access



# Requirements... (contd.)

---

- Reliability

- E.g.: The bank's database must be highly reliable and not lose data even during system crashes

- Data integrity

- E.g.: The bank's database must maintain consistent and correct data at all times

- Transaction processing

- E.g.: Some actions such as transfer of funds should occur atomically (as a single unit of work)



# Files

---

- On computers, data are stored using different mechanisms.
- A popularly used mechanism is the Operating System's (OS) file system.
- You can create a file (text, word etc.) and store data



# Files (contd.)

---

- Why don't we store data in OS files in an enterprise (e.g. bank database)?
- Files have certain limitations
  - **Sequential Access:** Files are accessed sequentially (hence not efficient for large amounts of data)
  - **Non-Concurrent access:** Files do not allow multiple users to update concurrently
  - **Security:** When you open a file, all the contents are visible to the user. Files have basic read, write, execute permissions.
  - **Cannot span disks:** Files cannot span multiple disks. Sometimes all the data might not fit on a single disk.



# Files (contd.)

---

- The file age/ the importance of DB
  - Imagine time savings, storage space savings, information availability/updating
    - to stay in touch and remember what was said/done in dealing with thousands of customers.
    - to keep the state of training or history of hundreds of employees.
    - to imagine how many people would have to be employed just making notes on paper, instead of being employed in creating new products and services





# DBMSs



- DBMSs is a complex software package that creates and maintains databases.
- DBMSs provide features that are essential for today's information systems:
  - Storage of large amounts of data
  - Data independence
  - Powerful query capabilities
  - Efficient access
  - Security
  - Concurrent access
  - Easy data administration
  - Transaction Processing
  - Reliability
- Therefore, today we find DBMSs are used in almost every enterprise information system



# DBMSs vs. Files

---

- With all its advantages, why do we use files then?
  - E.g.
    - Your assignment is saved as a Word file
    - Your email is saved as a text file
  - Databases are expensive software packages
  - They take up resources (memory, disk space etc.)
  - Above all, the overhead of using a DBMS is not needed for these types of applications as...
    - Small amounts of data are stored (assignment or email is small)
    - Single user applications (not many people want to concurrently update your assignment!)
    - Overhead for concurrency control, query capabilities, security etc. are not needed for these types of applications



# DBMS in Perspective - Evolution of DBMSs

---

- In early 1960s, Charles Bachman proposed the *Network Data Model*
- In the late 1960's, IBM developed Information Management System (IMS) based on the *Hierarchical Data Model*
- In 1970, E.F. Codd of IBM - San Jose Research Laboratory, proposed the *Relational Data Model*
- In 1990, came the **object-oriented DB; Data-warehousing; Web-enabled DB**



## Evolution... (contd.)

---

- Benefits of the relational model was widely accepted & became the de-facto standard in commercial databases
- The System R Project at IBM led to the development of **Structured English Query Language** (Sequel) language which later was standardized to be SQL (Structured Query Language)



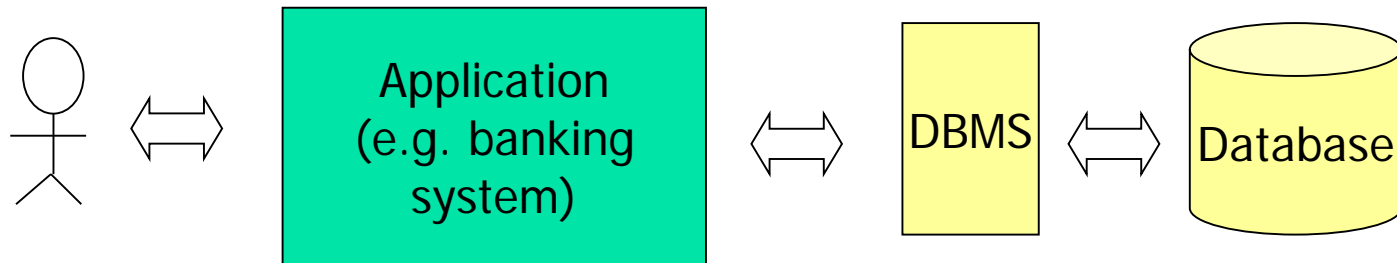
# Evolution of DBMSs... (contd.)

---

- DBMSs are now an essential part of any Enterprise Information System
- Today, extending DBMSs in a variety of ways...
  - Storing complex data types (multi-media, spatial data etc.)
  - Adding tools and techniques to build vast data warehouses and mine data (Decision Support Systems)... etc.

# DBMS Environment

- There are many different types of persons in developing, maintaining and using a database in an organisation



## End-Users

Include all users for which the system was developed (for e.g. customers of a bank, bank employees etc.)

## Application Developers

Develop programs that access the database

## Database Designer

Involved with designing the database for an application domain

## DBA

Create and maintains database and system resources required for database



# DBMS Environment (contd.)

---

- Database Administrators (DBA) are responsible to creating and maintaining databases and resources.
- DBAs use a variety of tools provided by the DBMS to monitor, manage, optimize and maintain the databases.
- DBA's duties include
  - optimize database for performance
  - plan, develop and implement backup and recovery procedures
  - ensure necessary resources such as disk space, memory etc. are met
  - implement security procedures



# DBMS Environment (contd.)

---

- Database Designers are responsible for designing, the database
- Database Designers duties include
  - understand the application domain and data requirements
  - Design the database in a target DBMS, which includes:
    - Understanding application domain requirements
    - Modelling data requirements in a conceptual model
    - Mapping the requirements to a data model (logical database design)
    - Selecting structures and methods for storage and fast retrieval (physical database design)





# DBMS Environment (contd.)

---

- Application Developers develop the programs that access the database providing the functionality of the information system.
- End-users include all users for which the Information System is built.



# Database Design Process – Our course objectives...

---

- This course focuses on gaining knowledge and skills to design, develop and implement a relational database for an application domain



# Development methodologies: What, Why and When?

---

- Developing large scale software is a complex process requiring careful design and management
- Certain methodologies have been proposed to ensure time, reliable and correct software products are built



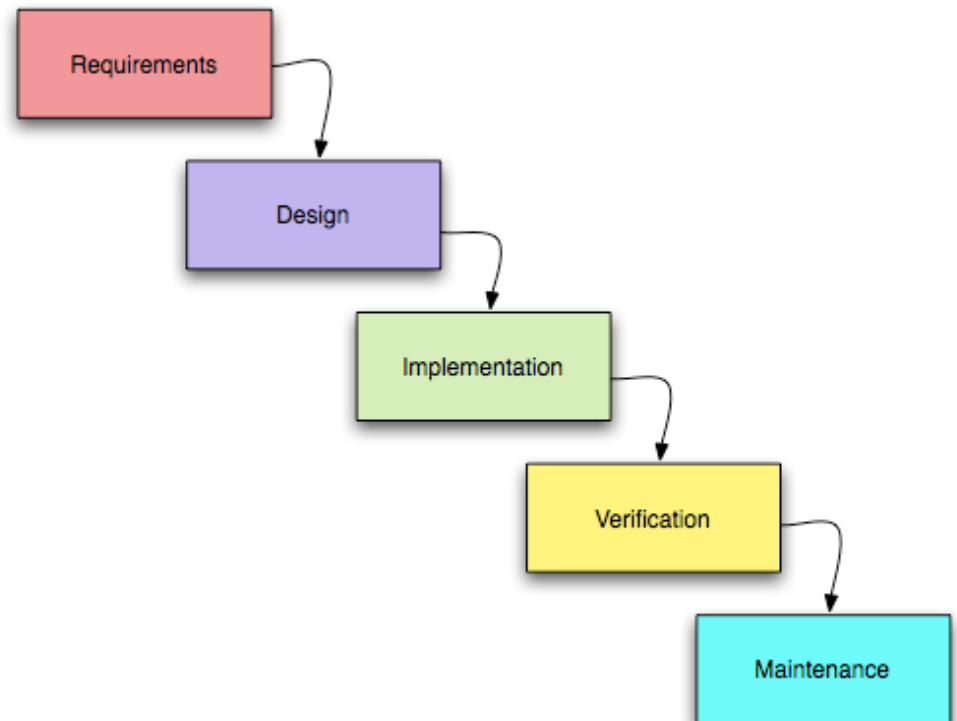
# Development methodologies: Why?

---

- Developing large scale software is complex and has had a high rate of failure (popularly known as the *Software Crisis*)
- Standish Report - The CHAOS Report
  - 24% of projects will be cancelled before they ever get completed
  - 44% of projects will cost more than their original estimates
  - 32% of software projects that are completed on-time and on-budget

# Methodologies for software development

- To enable better productivity & manage large scale software development, software development processes are suggested and followed.
- A popularly known methodology is **Software Development Life Cycle (SDLC)** and its variants.





# Database Design Process

---

- Database is a fundamental component of any enterprise information system.
- Information Systems development incorporates design and development of the database
- This course focuses on the proposed set of steps for developing databases, known as the **Database Design Process**
- We will follow these steps in completing this semester's assignment



# Database Design Process

---

- Database design process consists of the following main steps:
  - Requirements Analysis
  - Conceptual Database Design
  - Logical Database Design
  - Physical Database Design



# Requirements Analysis

---

- In this phase, you will understand and produce documentation outlining the detail requirements for the database system
- Fact finding techniques: interviews, research on existing systems etc. are used at this stage
- This step answers the following basic questions:
  - What types of data are stored in the databases?
  - What are types of services are implemented?
  - What types of business rules exist?





# Conceptual Database Design

---

- In this step, a high-level conceptual model is designed encapsulating the data requirements
- Entity-Relationship (ER) Model is a very popular & widely used conceptual model
- ER Model is enhanced with object-oriented features & known as the Enhanced Entity-Relationship (E-ER) Model
- We'll use EER model in developing a conceptual model for the retail database



# Logical Database Design

---

- At this stage, we convert the conceptual model developed in the previous step to a data model
- We will learn the relational data model and convert the EER conceptual model of database to relational data model (known as the **relational database schema**)
- At this stage, we further refine the relational database schema eliminating any redundancies that may exist. This process is called normalisation.



# Physical Database Design

---

- In this stage, the logical database design is realised in a physical implementation targeting a specific DBMS
- Storage structures and access methods are considered for efficient access
- Design considerations for protected access are also considered



# Requirements Analysis

---

- Fact finding techniques...
  - Examining documentation
    - (e.g. website of library for forms and services)
  - Interviewing
    - (e.g. library staff)
  - Observing the enterprise in operation
    - (e.g. observe the front desk of library)
  - Research
    - (e.g. existing and similar systems)
  - Questionnaires

# Requirements Analysis (contd.)



---

- Results in a document which outlines
  - What data to be stored in the database and their relationships?
  - What are the operations that the database needs to support? (i.e. inserts/updates/queries)
  - What are the rules that the data must follow? (business rules)



# Requirements Analysis – The DreamHome Case Study

- DreamHome specializes in property management, taking an intermediate role between owners who wish to rent out their furnished property and clients of DreamHome who require to rent furnished property for a fixed period. It currently has about 2000 staff working in 100 branches.  
(reference: chapter 11)

'The purpose of the *DreamHome* database system is to maintain the data that is used and generated to support the property rentals business for our clients and property owners and to facilitate the cooperation and sharing of information between branches.'

# Requirements Analysis – The DreamHome Case Study (cont'd)

- To register a staff with the company

<i>DreamHome</i> Staff Registration Form	
<b>Staff Number</b> <u>SG5</u> <b>Full Name</b> <u>Susan Brand</u> <b>Sex</b> <u>F</u> <b>DOB</b> <u>3-Jun-40</u>	<b>Branch Number</b> <u>B003</u> <b>Branch Address</b> <u>163 Main St, Glasgow</u> <b>Telephone Number(s)</b> <u>0141-339-2178 / 0141-339-4439</u>
<b>Position</b> <u>Manager</u> <b>Salary</b> <u>24000</u>	
<b>Enter details where applicable</b> <b>Supervisor Name</b> _____	<b>Manager Start Date</b> <u>01-Jun-90</u> <b>Manager Bonus</b> <u>2350</u>

# Requirements Analysis – The DreamHome Case Study (cont'd)

- To list details of all staff at a branch

<i>DreamHome</i> Staff Listing		
<b>Branch Number</b> <u>B003</u>		<b>Branch Address</b> <u>163 Main St, Glasgow</u> <u>G11 9QX</u>
<b>Telephone Number(s)</b> <u>0141-339-2178 / 0141-339-4439</u>		
<b>Staff Number</b>	<b>Name</b>	<b>Position</b>
SG5	Susan Brand	Manager
SG14	David Ford	Supervisor
SG37	Ann Beech	Assistant
SG112	Annet Longhorn	Supervisor
SG126	Chris Lawrence	Assistant
SG132	Sofie Walters	Assistant

Page 1



- 

<h2 style="text-align: center;">DreamHome</h2> <h3 style="text-align: center;">Property Registration Form</h3>	
<p><b>Property Number</b> <u>PG16</u></p> <p><b>Type</b> <u>Flat</u>      <b>Rooms</b> <u>4</u></p> <p><b>Rent</b> <u>450</u></p> <p><b>Address</b></p> <p><u>5 Novar Drive,</u></p> <p><u>Glasgow, G12 9AX</u></p> <p><u> </u></p> <p><u> </u></p> <p><u> </u></p>	<p><b>Owner Number</b> <u>C093</u> (If known)</p> <p><b>Person/Business Name</b></p> <p><u>Tony Shaw</u></p> <p><b>Address</b> <u>12 Park Pl,</u> <u>Glasgow G4 0QR</u></p> <p><b>Tel No</b> <u>0141-225-7025</u></p>
	<p><b>Enter details where applicable</b></p> <p><b>Type of business</b> _____</p> <p><b>Contact Name</b> _____</p>
<p><b>Managed by staff</b></p> <p><u>David Ford</u></p>	<p><b>Registered at branch</b></p> <p><u>163 Main St, Glasgow</u></p>

# Requirements Analysis – The DreamHome Case Study (cont'd)

- To register a client

<i>DreamHome</i> Client Registration Form	
<b>Client Number</b> <u>CR74</u> (Enter if known) _____  <b>Full Name</b> <u>Mike Ritchie</u>	<b>Branch Number</b> <u>B003</u>  <b>Branch Address</b> <u>163 Main St, Glasgow</u> _____ _____
<b>Enter property requirements</b>  <b>Type</b> <u>Flat</u>  <b>Max Rent</b> <u>750</u>	<b>Registered By</b> <u>Ann Beech</u>  <b>Date Registered</b> <u>16-Nov-02</u>

# Requirements Analysis – The DreamHome Case Study (cont'd)

- To list properties currently available

<b>DreamHome</b> <b>Property Listing for Week beginning 01/06/04</b>																																							
If you are interested in viewing or renting any of the properties in this list please contact our branch office as soon as possible.																																							
<b>Branch Address</b>		<b>Telephone Number(s)</b>																																					
<u>163 Main St, Glasgow</u>		<u>0141-339-2178 / 0141-339-4439</u>																																					
<u>G11 9QX</u>																																							
<table border="1"><thead><tr><th>Property No</th><th>Address</th><th>Type</th><th>Rooms</th><th>Rent</th></tr></thead><tbody><tr><td>PG4</td><td>6 Lawrence St, Glasgow</td><td>Flat</td><td>3</td><td>350</td></tr><tr><td>PG36</td><td>2 Manor Rd, Glasgow</td><td>Flat</td><td>3</td><td>375</td></tr><tr><td>PG21</td><td>18 Dale Road, Glasgow</td><td>House</td><td>5</td><td>600</td></tr><tr><td>PG16</td><td>5 Novar Drive, Glasgow</td><td>Flat</td><td>4</td><td>450</td></tr><tr><td>PG77</td><td>100A Apple Lane, Glasgow</td><td>House</td><td>6</td><td>560</td></tr><tr><td>PG81</td><td>781 Greentree Dr, Glasgow</td><td>Flat</td><td>4</td><td>440</td></tr></tbody></table>					Property No	Address	Type	Rooms	Rent	PG4	6 Lawrence St, Glasgow	Flat	3	350	PG36	2 Manor Rd, Glasgow	Flat	3	375	PG21	18 Dale Road, Glasgow	House	5	600	PG16	5 Novar Drive, Glasgow	Flat	4	450	PG77	100A Apple Lane, Glasgow	House	6	560	PG81	781 Greentree Dr, Glasgow	Flat	4	440
Property No	Address	Type	Rooms	Rent																																			
PG4	6 Lawrence St, Glasgow	Flat	3	350																																			
PG36	2 Manor Rd, Glasgow	Flat	3	375																																			
PG21	18 Dale Road, Glasgow	House	5	600																																			
PG16	5 Novar Drive, Glasgow	Flat	4	450																																			
PG77	100A Apple Lane, Glasgow	House	6	560																																			
PG81	781 Greentree Dr, Glasgow	Flat	4	440																																			

# Requirements Analysis – The DreamHome Case Study (cont'd)

- To report comments by clients

<i>DreamHome</i> Property Viewing Report																							
Property Nummer <u>PG4</u>		Property Address <u>6 Lawrence St, Glasgow</u>																					
Type <u>Flat</u>																							
Rent <u>350</u>																							
<table border="1"><thead><tr><th>Client No</th><th>Name</th><th>Date</th><th>Comments</th></tr></thead><tbody><tr><td>CR76</td><td>John Kay</td><td>20/04/04</td><td>Too remote.</td></tr><tr><td>CR56</td><td>Aline Stewart</td><td>26/05/04</td><td></td></tr><tr><td>CR74</td><td>Mike Ritchie</td><td>11/11/04</td><td></td></tr><tr><td>CR62</td><td>Mary Tregear</td><td>11/11/04</td><td>OK, but needs redecoration throughout.</td></tr></tbody></table>				Client No	Name	Date	Comments	CR76	John Kay	20/04/04	Too remote.	CR56	Aline Stewart	26/05/04		CR74	Mike Ritchie	11/11/04		CR62	Mary Tregear	11/11/04	OK, but needs redecoration throughout.
Client No	Name	Date	Comments																				
CR76	John Kay	20/04/04	Too remote.																				
CR56	Aline Stewart	26/05/04																					
CR74	Mike Ritchie	11/11/04																					
CR62	Mary Tregear	11/11/04	OK, but needs redecoration throughout.																				
Page 1																							

# Requirements Analysis – The DreamHome Case Study (cont'd)

- Lease to a property

<b>DreamHome Lease Number 00345810</b>	
<b>Client Number</b> <u>CR74</u> (Enter if known)  <b>Full Name</b> <u>Mike Ritchie</u> (Please print)  <b>Client Signature</b> _____	<b>Property Number</b> <u>PG16</u>  <b>Property Address</b> <u>5 Novar Dr, Glasgow</u>
<b>Enter payment details</b>  <b>Monthly Rent</b> <u>450</u>  <b>Payment Method</b> <u>Cheque</u>  <b>Deposit Paid (Y or N)</b> <u>Yes</u>	<b>Rent Start</b> <u>01/06/04</u>  <b>Rent Finish</b> <u>31/05/05</u>  <b>Duration</b> <u>1 year</u>



# Assignment

---

- Assignment 1 handout outlines the main features of the system. Will be released & discussed next week.
- Sample requirements documents are available in Appendices A and B of the textbook (placed on Short Loans)



# Summary

---

- Introduction to course
  - Objectives
  - Strategy
  - Lecturer info
  - Assessments
- Introduction to DBMS
  - Data, information, database and DBMS
  - Data needs of today's information systems
  - Files and their limitations
  - DBMS features
  - When and when not to use DBMSs



# Summary (contd.)

---

- DBMS in Perspective
  - Evolution of database technology
  - Roles of persons in a database environment
  - Revisit course objectives
- Database Design Process
- More on Requirements Gathering





# Summary Qs

---

- 1. What are the names and definitions of the four database design processes?
- 2. List nine features that a DBMS can provide that are essential for today's information systems.
- 3. Name and briefly describe the four types of persons in developing, maintaining and using a database in an organisation.