

Dr Greg Wadley
David Eccles



INFO 90002

Database Systems & Information Modelling

Lecture 1:

Introduction to INFO90002

Introduction to Relational Databases

Introduction to MySQL Workbench

Lecture 2:

The Database Development Lifecycle



- Tutorials start this week!
 - Learn how to install MySQL Workbench & Server
- Zoom links are available from the LMS



- Lecture 01:
 - Introductions and administration
 - subject overview
 - staff and students
 - learning resources
 - Assessment
 - Academic Integrity
- Lecture 02:
 - Introduction to databases & MySQL RDBMS
 - database technology, past present and future
 - how databases are designed, implemented and used



- Welcome to INFO 90002
Database Systems & Information Modelling

Why this subject matters

- database = key building block in many technology careers
- database = one of the most widely-used technologies
 - embedded within most of the interesting ICT of today
 - social media, apps, websites, banking, scientific research ...
- database have come to influence our culture
 - “The database is the major cultural form of the 21st century in much the same way as the novel was for the 19th and the film for the 20th. ... While retaining the visual and temporal aspects of film, the modality of hypertext or of computer games eschews its linear modality for the modality of the database, in which objects are linked together but their assembly into a narrative experience is in the hands of the audience.”

Dourish and Mazmanian (2011), discussing Manovich (2002) *The Language of New Media*



Forbes (IT)	Forbes (NACE)	AllA	Youth Central (Generic)	Dept of Employment
SQL	Teamwork	Cross-Cultural Competency	Communication	(ICT Sector) Lack of;
Mentoring	Solve Problems	Social Intelligence	Problem Solving	Industry Knowledge
ORACLE	Make Decisions	Novel & Adaptive Thinking	Teamwork	Technical Skills
Collaboration	Communicate	Virtual Collab.	Learning	Other Soft Skills
Process Improvement	Plan, Organise, Prioritise	New Media Literacy	Initiative and Enterprise	Communication Skills
Business Devt.	Obtain and Process Info	Computational Thinking	Self-management	
Decision Making	Quantitative Data Analysis	Design Mindset	Planning and Organising	
Data Analysis	Technical Skills	Resilience	Technical Skills	
SAP	Influencing	Transdisciplinarity		



- Subject coordinators
 - David Eccles (Lecturer)
- Tutors
 - Veronica Torres Pena
 - Fangziyun Tong
 - Fraser McHarg
 - Mina Sorsok
 - Neven Tomov
 - Nick Howard
- Student representative (no extra credit for this task but you can put it on your CV)
 - One Volunteer (maybe 2)
 - Must be able to speak truth to power!
 - Must watch all lectures and attend all tutorials
- Interacting with staff
 - Lecturer consultations (Tuesday 11am Wednesday 7pm.)
 - Canvas Discussion Wall (**All general questions are answered here**)
 - email for *personal* questions only from your student email account
 - (e.g. AEA, personal issues, formal extension requests)
 - include **SUBJECT CODE** (INFO90002) and **STUDENT ID** (987654) otherwise your email will not be acknowledged (**No student ID – No answer from me**)



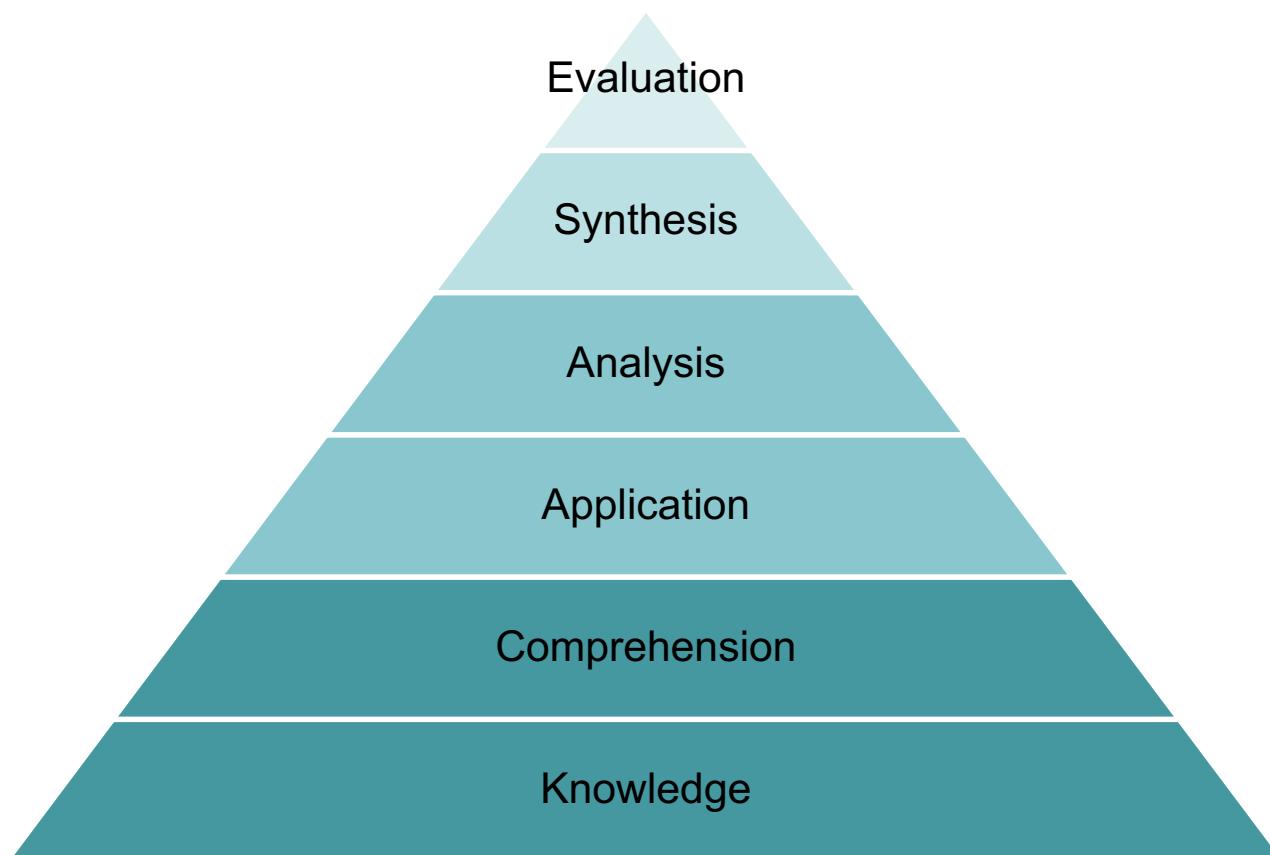
- Prerequisites, not-allowed subjects, credit for experience
 - Have you studied DB already? Don't study it again - get credit!
 - Learn something NEW and EXCITING
- Semester schedule: the big picture
 - Modelling
 - SQL
 - Advanced Database Concepts & Ideas
- Assessment – **New in 2021!**
 - assignment 1: data modelling (20%) .. (group assignment)
 - 3 Individual timed Quizzes in weeks 6, 9 and 12 each worth 10% (30%)
 - end of semester exam (50%, 2 hours) (down from 70%)
- How to succeed in this subject
 - Perfect practice! (especially Data Modelling and SQL skills)
 - use all the learning resources provided

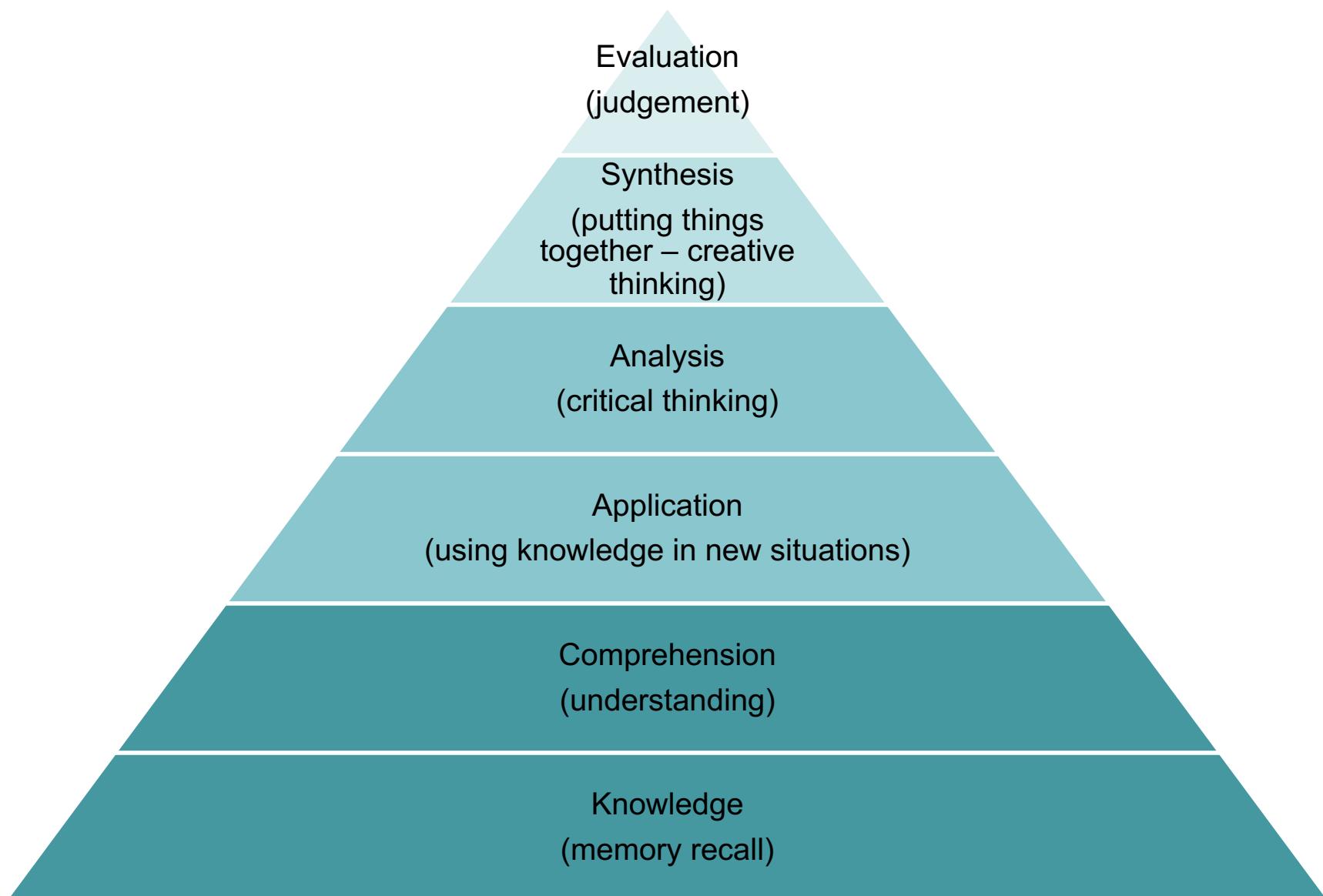


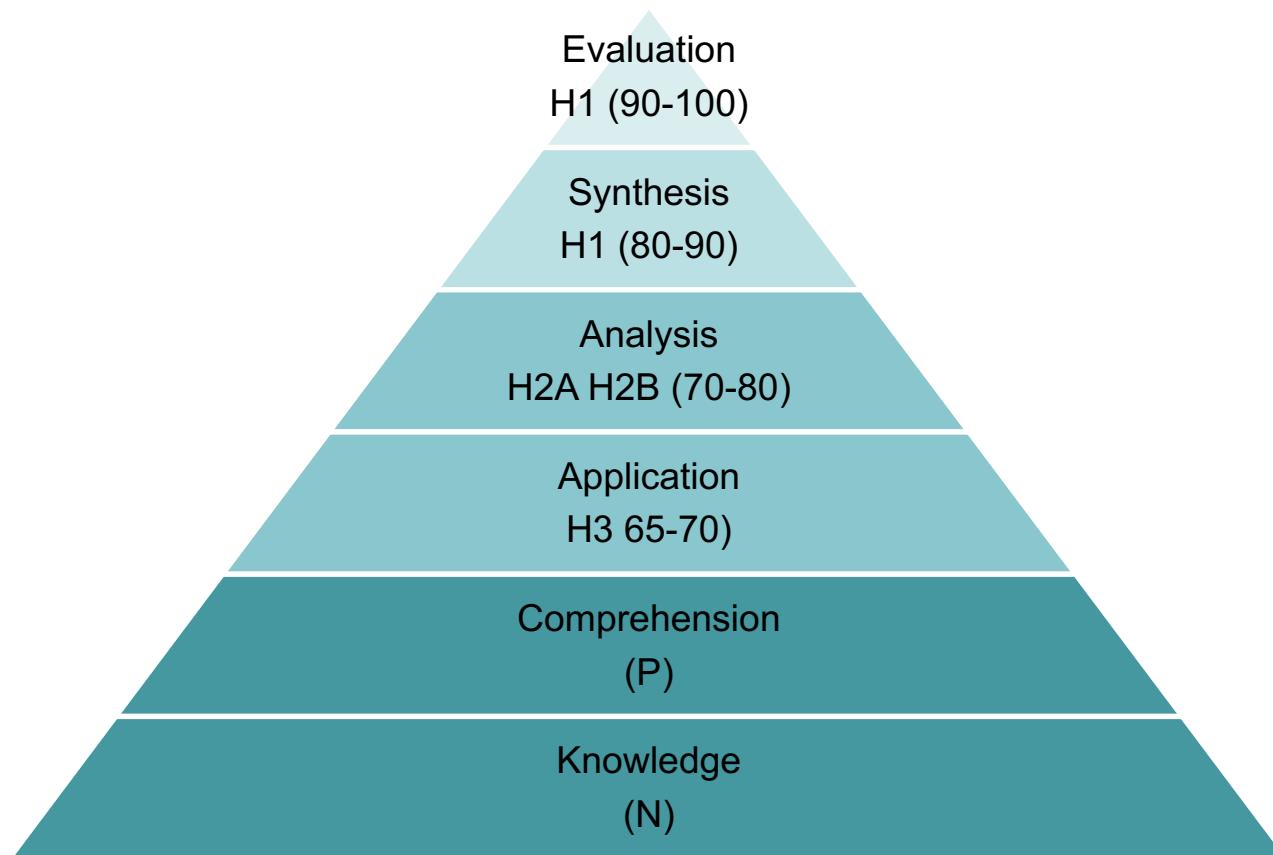
- You should, on average, invest 10 hours per week of work on this subject (this is standard for a Masters-level subject), e.g.:
 - 3 hours in class/tutorial
 - 2-3 hours reading
 - 4-5 hours assessment tasks
- Full time Study and Full time Job
 - You can do full time work well
 - You can do full time study well
 - You can not do full time study and full time work well
 - Please reconsider your study and work commitments now. I've yet to meet anyone who has done well at both



- You may know SQL
- You may know Modelling
- Industry students generally do poorly because:
 - Don't follow the instructions
 - Don't do well in normalisation and advanced theory topics
- Industry experienced students do very well
 - SQL
 - Crows Foot Modelling
- I can only assess what we teach in INFO90002 not what you learned in Industry







You are not here to memorise, you are here to synthesize
Memory recall is insufficient to pass this subject!



Graduates at Masters level will have:

- Specialised knowledge and skills for research, and/or professional practice and/or further learning
- Advanced and integrated understanding of a complex body of knowledge
- Expert, specialised cognitive and technical skills in a body of knowledge or practice
- Ability to apply knowledge and skills to demonstrate autonomy, expert judgement, adaptability and responsibility as a practitioner or learner

See the Australian Qualifications Framework:
<http://www.aqf.edu.au/aqf/in-detail/aqf-levels/>



- I can't (and won't) tell you what to do
- We will give you the necessary knowledge and skills practice to be able to do it
- We can (but won't) tell you if you're right during assessments
- We will tell you if you're going the wrong way (if you ask)
- You will need to synthesize and critical apply the knowledge
- You must demonstrate comprehension, application, synthesis and critical thinking
- Memory and Recall is Junior School (AQF 1-6)
- Application is Undergraduate level (AQF 8)
- You are studying at Level 9 of the (AQF9)



Week	Lecture 1	Lecture 2	Topic	Ramakrishnan & Gehrke reference	Key Dates
Week 1	1. Subject Administration	2. Introduction to Databases and the database lifecycle	1. Installation and set up of MySQL Workbench Get to know your tutors and classmates	Chapter 1 Overview of Database Systems 1.1 Managing Data 1.2 A Historical Perspective 1.3 File systems versus a DBMS 1.4 Advantages of a DBMS 1.5 Describing and Storing Data in a DBMS	
Week 2	3. Conceptual Modelling	4. Logical & Physical Modelling	2. Using Workbench to build a Physical Model	Chapter 2 Introduction to Database Design	Assignment 1 Released
Week 3	5. ER Modelling and data types	6. Normalisation	3. Chen Conceptual Notation and Verb-Noun Analysis	Chapter 3 The Relational Model	Practice Quiz No 1
Week 4	7. ER Modelling - LIVE!	8. SQL 1 Overview DDL DML Joins and Functions	Normalisation	Chapter 5 SQL: Queries, Constraints, Triggers 5.2 The Form of a Basic SQL Query 5.4, Nested Queries	
Week 5	9. SQL 2 Overview DDL DML JOINS FUNCTIONS	10. SQL 3		Chapter 5 SQL: Queries, Constraints, Triggers 5.5.1 The GROUP BY and HAVING clauses 5.6 NULL Values 12.1 The System Catalog	Assignment 2 Database released
5-Apr	NO CLASS	NO CLASS	NO CLASS	MID SEMESTER BREAK	NO CLASS
Week 6	11. Apps & Web Apps	12. Quiz No 1 (10%)	SQL 2	Chapter 5 5.3 UNION INTERSECT and EXCEPT	Assignment 1 Due Practice Quiz No 2
Week 7	13. Transactions Concurrency & Locking	14. Storage & Indexing	SQL 3	Chapter 16 16.1 The ACID properties 16.2 Transactions and Schedules 16.3 Concurrent execution of Transactions 16.4 Lock-Based Concurrency Control 16.6 Transaction Support in SQL Chapter 17. Concurrency Control 17.1 View Serializability 17.2 Introduction to Lock Management 17.6 Concurrency control without Locking	
Week 8	15. Administration & Architecture	16. Distributed Databases	SQL 4		SQL Practice Quiz
Week 9	17. Quiz No 2 (SQL)	18. Quiz No 2 (SQL) 10%	Referential Integrity		
Week 10	19. Data Warehouse	20. Security & Ethics		Chapter 25 25.6 25.7 25.8	
Week 11	21. NoSQL	22. Assignment 1 Feedback	Data Warehousing		
Week 12	23. Quiz No 3 (10%)	24. Subject Review and Exam Prep	No SQL		



- Tutorials run at various times during the week
 - (Tuesday to Friday this semester)
- Check <https://sws.unimelb.edu.au/2020/>
- Tutorials are not assessed
- Attendance is not recorded
- The delivery model is via Zoom
- You can attend as many or as few tutorials for any given week
- Tutorial links are published in the LMS under the Zoom link



- Two lecturer consultation hours
- 11am Tuesdays Melbourne time
- 7pm Wednesdays Melbourne time
- There may also be the additional consultations leading up to Assignment deadlines and post assignment feedback
 - These will be announced with a zoom link via the LMS
- Tutorial consultations are not recorded
 - Please bring your questions!
 - Lecture Number, Slide Number
 - Tutorial Week and Page Number

- We will use *MySQL* as server and client in this subject
- You can use either:
 - the University's database server
 - accessible from labs, or from home via VPN*
 - Your assignment 2 solution must be able to run on this server!
 - or, your own computer
 - do lab exercises and assignments at home
- Server address: `info90002db.eng.unimelb.edu.au` : port 3306
 - your username and password will be given out in first lab
 - not available outside the university without a VPN
- If you want to use your own computer ...
 - download MySQL from <http://dev.mysql.com/downloads/mysql/>
 - get both Server and Workbench – (Covered in Week 1 Tutorial)
 - Binary copies on the resources module on LMS



- Please do not use SER in your assignments
- Please do not be a victim of student study site scams
 - You are blackmailed
 - Universities can and do revoke degrees up to 5 years post graduation
 - I can use Google too!
- Please understand what is and isn't academic dishonesty (what I call cheating)
- You can study in groups, but your work must always be your own
 - not another current, former, or previous students
- See <https://academicintegrity.unimelb.edu.au/>

Dr Greg Wadley
David Eccles

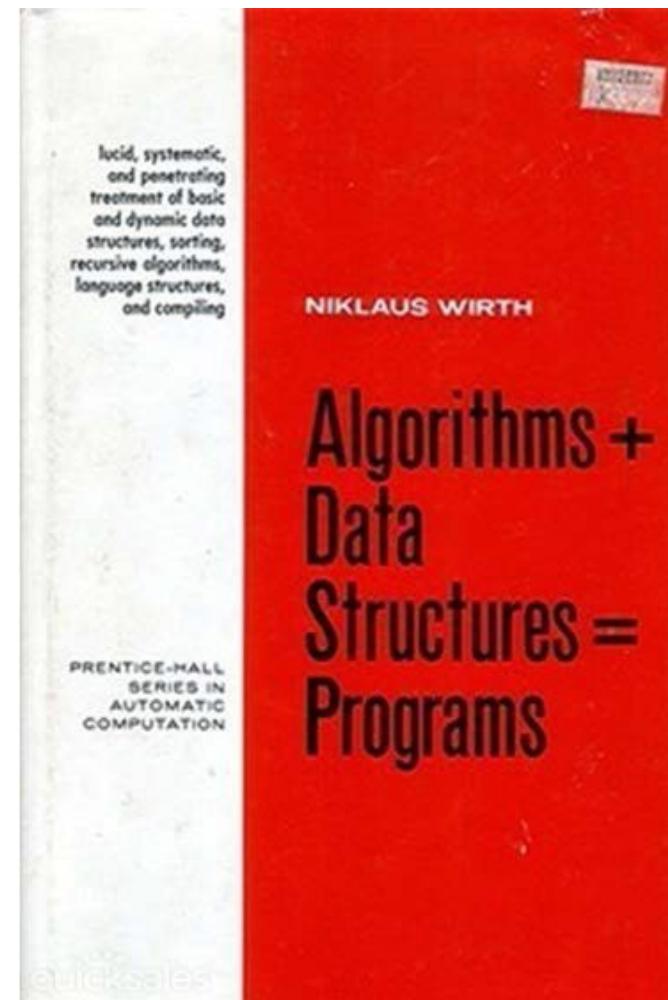


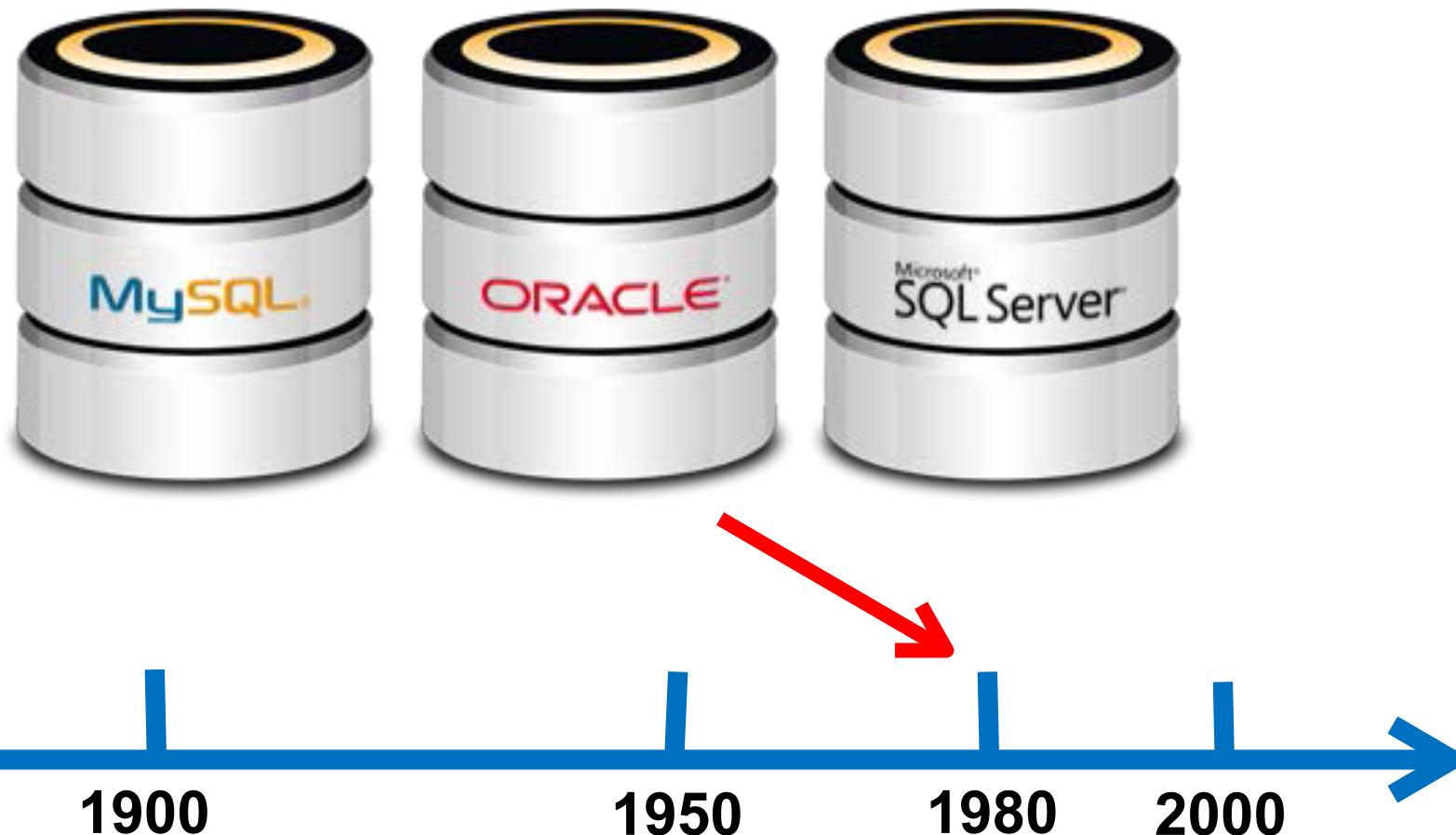
INFO 90002

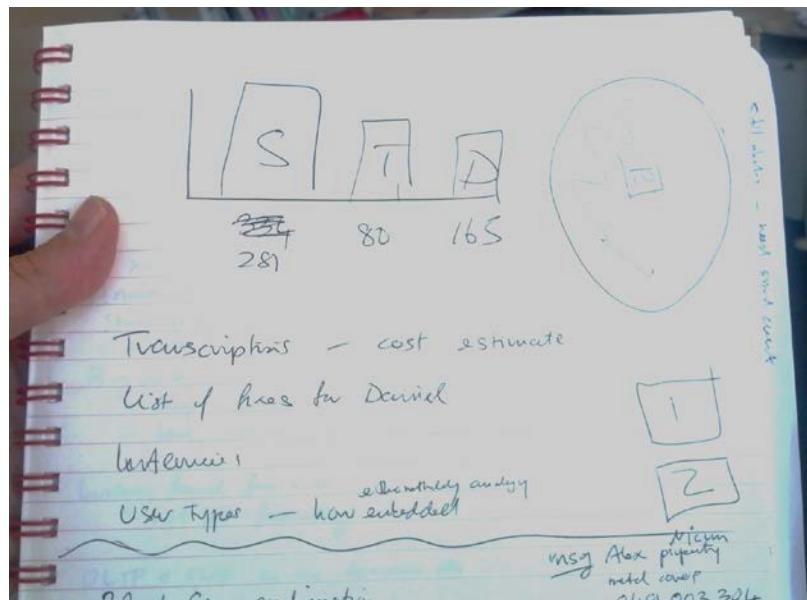
Database Systems & Information Modelling

Lecture 02
Introduction to Databases

- Computer systems consist of *software* (algorithms) working to process *data*.
- You will learn about creating software and algorithms in other subjects (INFO90059)
- This subject is about *data*.
- Focus away from *computation* to *information*







A	B	C	D	E
1	My Holiday Budget ----->			
2				
3	Flights	\$10,000.00		
4	Accommodation	\$5,000.00	< can this be reduced?	
5	Food	\$1,000.00		
6				
7		\$16,000.00	TOTAL	
8				
9				
10				
11				

- handwritten notes
- printed books
- spreadsheets etc

80 NATURAL SELECTION. CHAP. IV.

CHAPTER IV.
NATURAL SELECTION.

Natural Selection—its power compared with man's selection—its power on characters of trifling importance—its power at all ages and on both sexes—Sexual Selection—On the generality of intercrosses between individuals of the same species—Circumstances favourable and unfavourable to Natural Selection, namely, intercrossing, isolation, number of individuals—Slow action—Extinction caused by Natural Selection—Divergence of Character, related to the diversity of inhabitants of any small area, and to naturalisation—Action of Natural Selection, through Divergence of Character and Extinction, on the descendants from a common parent—Explains the Grouping of all organic beings.

How will the struggle for existence, discussed too briefly in the last chapter, act in regard to variation? Can the principle of selection, which we have seen is so potent in the hands of man, apply in nature? I think we shall see that it can act most effectually. Let it be borne in mind in what an endless number of strange peculiarities our domestic productions, and, in a lesser degree, those under nature, vary; and how strong the hereditary tendency is. Under domestication, it may be truly said that the whole organisation becomes in some degree plastic. Let it be borne in mind how infinitely complex and close-fitting are the mutual relations of all organic beings to each other and to their physical conditions of life. Can it, then, be thought improbable, seeing that variations useful to man have undoubtedly occurred, that other variations useful in some way to each being in the great and complex battle of life, should sometimes occur in the course of thousands of generations? If such do occur, can we doubt (remem-



- Each row represents an instance of a set of entities
- Entities might be people, things, events, transactions...
- You can extract useful knowledge via simple, repetitive processes
- A natural fit for business and scientific data.



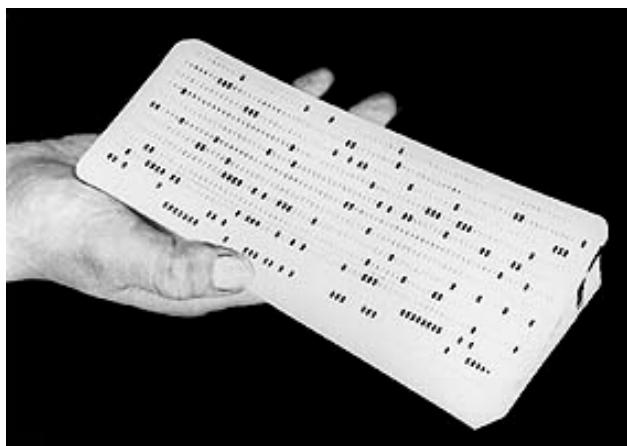
O in V										O in S									
Time from h. m. s.	10	11	12	Ascend.	1	2	3	time from h. m. s.	10	11	12	Ascend.	1	2	3	h. m. s.	II	III	IV
h. m. s.	II	III	IV	ad	m	s	a	h. m. s.	II	III	IV	ad	m	s	a	h. m. s.	ad	m	s
0. 0. 0	0	9	22	26	36	12	3	1. 51. 37	0	9	17	16. 27	4	28					
0. 3. 40	1	10	23	27	17	13	3	1. 55. 27	1	10	18	17. 18	5	29					
0. 7. 20	2	11	24	27	38	14	4	1. 59. 17	2	11	19	17. 48	6	30					
0. 11. 0	3	12	25	28	42	15	5	2. 3. 8	3	12	19	18. 27	7	1					
0. 14. 41	4	13	25	29	17	16	6	2. 6. 59	4	13	20	19. 9	8	2					
0. 18. 21	5	14	26	29	55	16	7	2. 10. 57	5	14	21	19. 49	9	2					
0. 22. 2	6	15	27	02	24	17	8	2. 14. 44	6	15	22	20. 29	9	3					
0. 25. 42	7	16	28	1. 14	18	8	9	2. 18. 37	7	16	22	21. 10	10	4					
0. 29. 23	8	17	29	1. 55	18	9	10	2. 22. 31	8	17	23	24. 52	11	5					
0. 33. 4	9	18	29	2. 33	19	10	11	2. 26. 25	9	18	24	22. 32	11	6					
0. 36. 45	10	19	1. 3	14	20	11	2	3. 30. 20	10	19	24	23. 14	12	7					
								34. 16	11	20	26	23. 65	13	8					
								38. 13	12	21	16	24. 36	14	9					
								42. 10	13	22	27	25. 17	15	10					
								46. 8	14	23	28	25. 58	15	11					
								50. 7	15	24	29	26. 10	16	12					
								54. 7	16	25	29	27. 22	17	12					
								58. 7	17	26	36	28. 4	10	13					
								62. 8	18	26	1	28. 45	10	14					
								66. 9	19	27	2	28. 28	17	15					
								70. 12	20	28	3	6. 12	20	16					
								74. 15	21	29	3	0. 54	21	17					
								78. 19	22	20	4	1. 36	22	19					
								82. 23	23	1	5	2. 20	22	19					
								86. 29	24	2	6	3. 1	23	20					
								90. 35	25	3	7	3. 45	24	21					
								94. 41	26	4	7	4. 20	25	22					
								98. 49	27	5	8	5. 11	26	23					
								102. 57	28	6	9	5. 54	27	24					
								106. 6	29	7	10	6. 29	27	25					

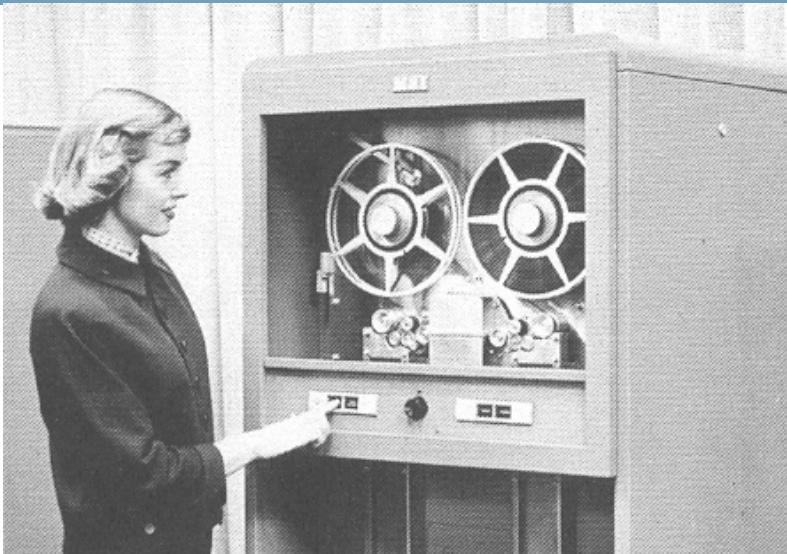


1890 US census

1	2	3	4	CM	UM	Jp	Ch	Oc	In	20	50	80	Dv	Un	3	4	3	4	A	E	L	a	g
5	6	7	8	CL	UL	O	Mi	Qd	Mo	25	55	85	Wd	CY	1	2	1	2	B	F	M	b	h
1	2	3	4	CS	US	Mb	B	M	O	30	60	0	2	Mr	0	15	0	15	C	G	N	c	i
5	6	7	8	No	Hd	Wf	W	F	5	35	65	1	3	Sg	5	10	5	10	D	H	O	d	k
1	2	3	4	Fh	Ff	Fm	7	1	10	40	70	90	4	0	1	3	0	2	St	I	P	e	l
5	6	7	8	Hh	Hf	Hm	8	2	15	45	75	95	100	Un	2	4	1	3	4	K	Un	f	m
1	2	3	4	X	Un	Ft	9	3	i	c	X	R	L	E	A	6	0	US	Ir	Sc	US	Ir	Sc
5	6	7	8	Ot	En	Mt	10	4	k	d	Y	S	M	F	B	10	1	Gr	En	Wa	Gr	En	Wa
1	2	3	4	W	R	OK	11	5	l	e	Z	T	N	G	C	15	2	Sw	FC	EC	Sw	FC	EC
5	6	7	8	7	4	1	12	6	m	f	NG	U	O	H	D	Un	3	Nw	Bo	Hu	Nw	Bo	Hu
1	2	3	4	8	5	2	5	0	n	g	a	V	P	I	Al	Na	4	Dk	Fr	It	Dk	Fr	It
5	6	7	8	9	6	3	0	p	o	h	b	w	q	k	Un	Pa	5	Ru	Ot	Un	Ru	Ot	Un

[https://www.census.gov/history/www/innovations/technology/the hollerith tabulator.html](https://www.census.gov/history/www/innovations/technology/the_hollerith_tabulator.html)





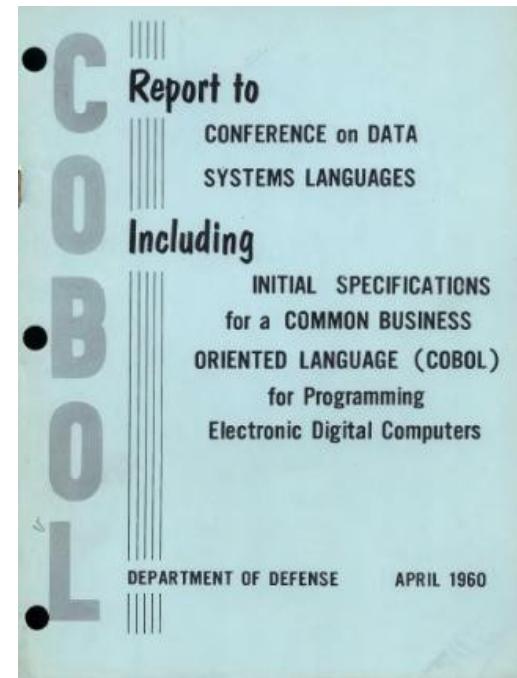
EMP_RECORD_ID	EMP_ID	EMP_REGION	EMP_DEPT	EMP_HIRE_DATE	E1	E2	EMP_SALARY	EMP_NAME	EMP_SKILL_CODE	
68	3715	4	153	09061987	9	6	1987	14000000	IRENE HIRSH	041085
62	39412	1	650	031119590	3	11	9590	167000000	ANN FAHEY	031099
56	1939	2	265	09281988	9	28	1988	21300000	EMILY WILM...	021077
50	3502	2	165	07041985	7	4	1985	19500000	CATHEZINE ...	011015
44	4435	2	117	05141989	5	14	1989	17000000	AGNES KING	00
68	1673	3	138	07021985	7	2	1985	16800000	MARTIN XU	041033
62	4181	3	161	02031988	2	3	1988	15900000	JOHN DURN	030045
56	1443	1	265	12028900	12	2	8900	6000000	PAT DUNN	021055
50	3607	3	127	08072000	8	7	2000	18300000	ANDREA HIN...	011014
44	1775	3	288	02051989	2	5	1989	2700000	PETER JONES	00
68	1209	2	165	05121986	5	12	1986	17300000	DIDRA WLK...	041065

<http://www.computerhistory.org/timeline/memory-storage/>



<http://groups.engin.umd.umich.edu/CIS/course.des/cis400/cobol/cobol.html>

```
DATA DIVISION.  
FILE SECTION.  
FD StudentFile.  
01 StudentRec.  
    88 EndOfFile VALUE HIGH-VALUES.  
    02 StudentId      PIC 9(7).  
    02 StudentName.  
        03 Surname      PIC X(8).  
        03 Initials     PIC XX.  
    02 DateOfBirth.  
        03 YOBirth       PIC 9(4).  
        03 MOBirth       PIC 9(2).  
        03 DOBirth       PIC 9(2).  
    02 CourseCode     PIC X(4).  
    02 Gender         PIC X.
```



Problems with flat-files:

- data access routines must be programmed in detail
- each program must include full detail of data structure
- multiple users cannot simultaneously access data
- multiple copies of data - not centrally managed



- DBMS = DataBase Management System
- Data independence
 - Applications should not be exposed to data representation and storage
 - DBMS provides an abstract view that hides representation & storage
- Efficient access
 - More efficient data storage and retrieval than flat files
- Data integrity and security
 - DBMS enforces data integrity constraints, access controls and govern user access
 - Not reliant on just the operating system



- Uniform data administration
 - Specialist skills in data management and administration
 - Layer of expertise reduces risk to data and data owners
- Concurrent access & crash recovery
 - Schedules concurrent access. Protects data from system failures
- Datasets increasing in diversity and volume
- Data Independence
 - Logical data independence: Protection from changes in logical structure of data.
 - Physical data independence: Protection from changes in physical structure of data
- Reduced Application Development time*



- the first relational databases from Oracle and IBM appear around 1980

ORACLE®

IBM

DB2

- others appear later

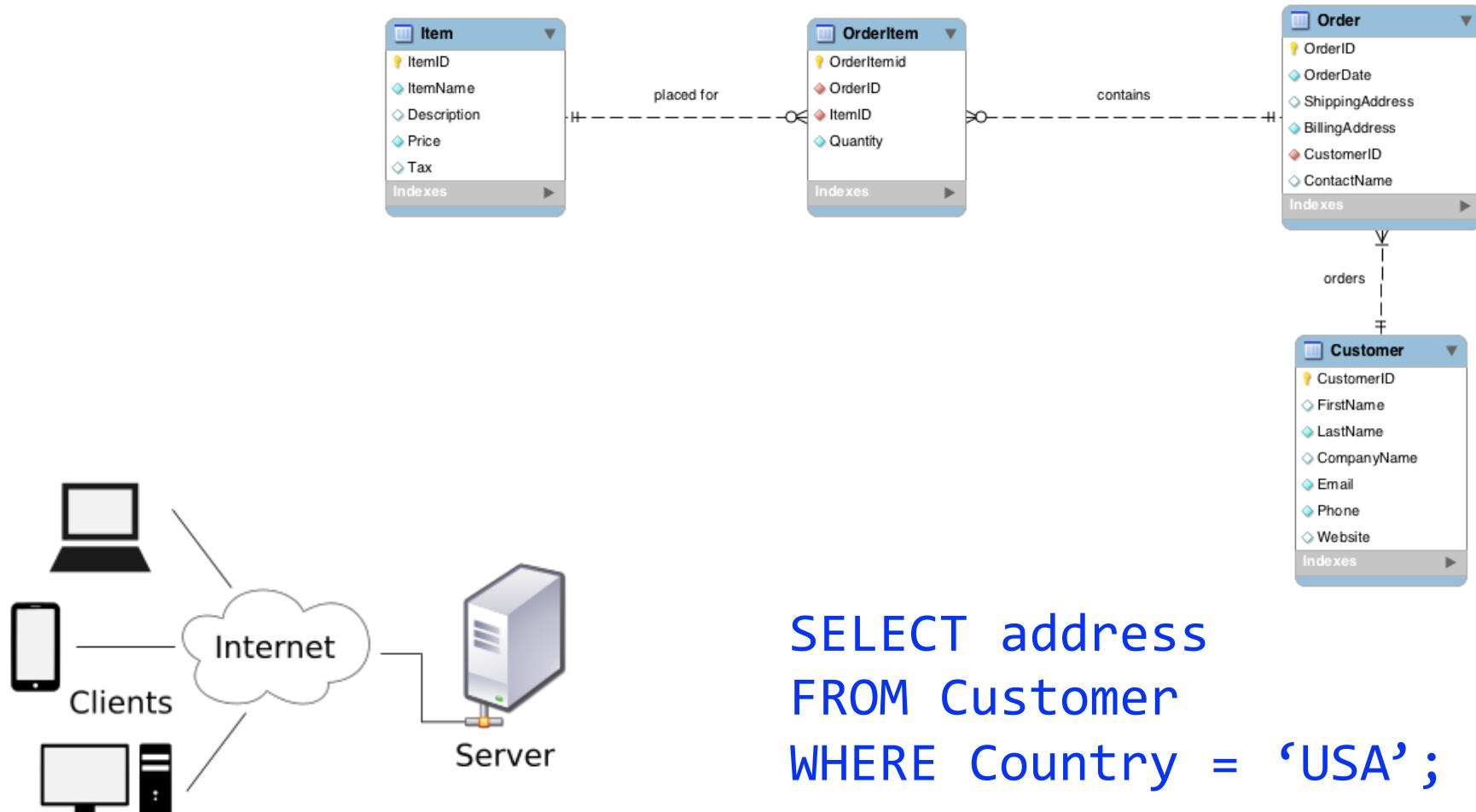


Microsoft®
SQL Server®





- entity-relationship diagrams, client-server architecture, SQL language

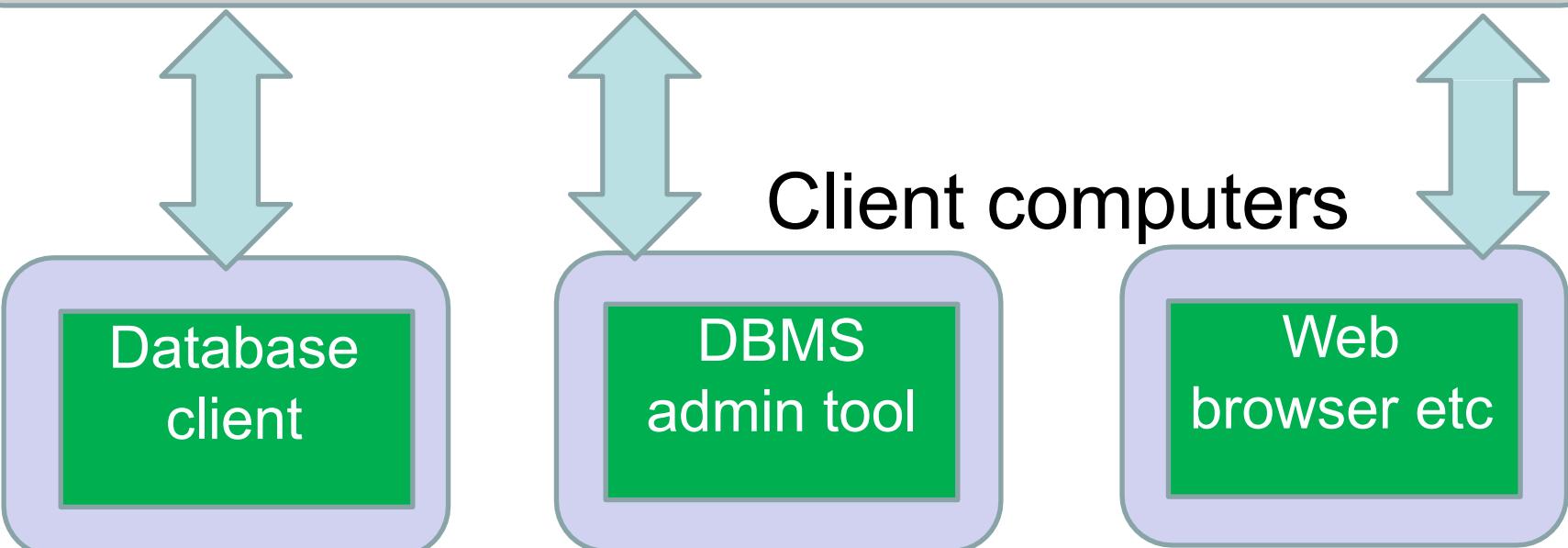
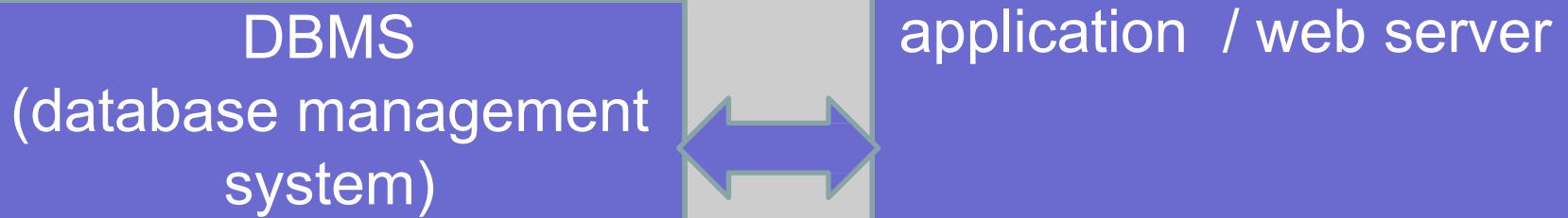


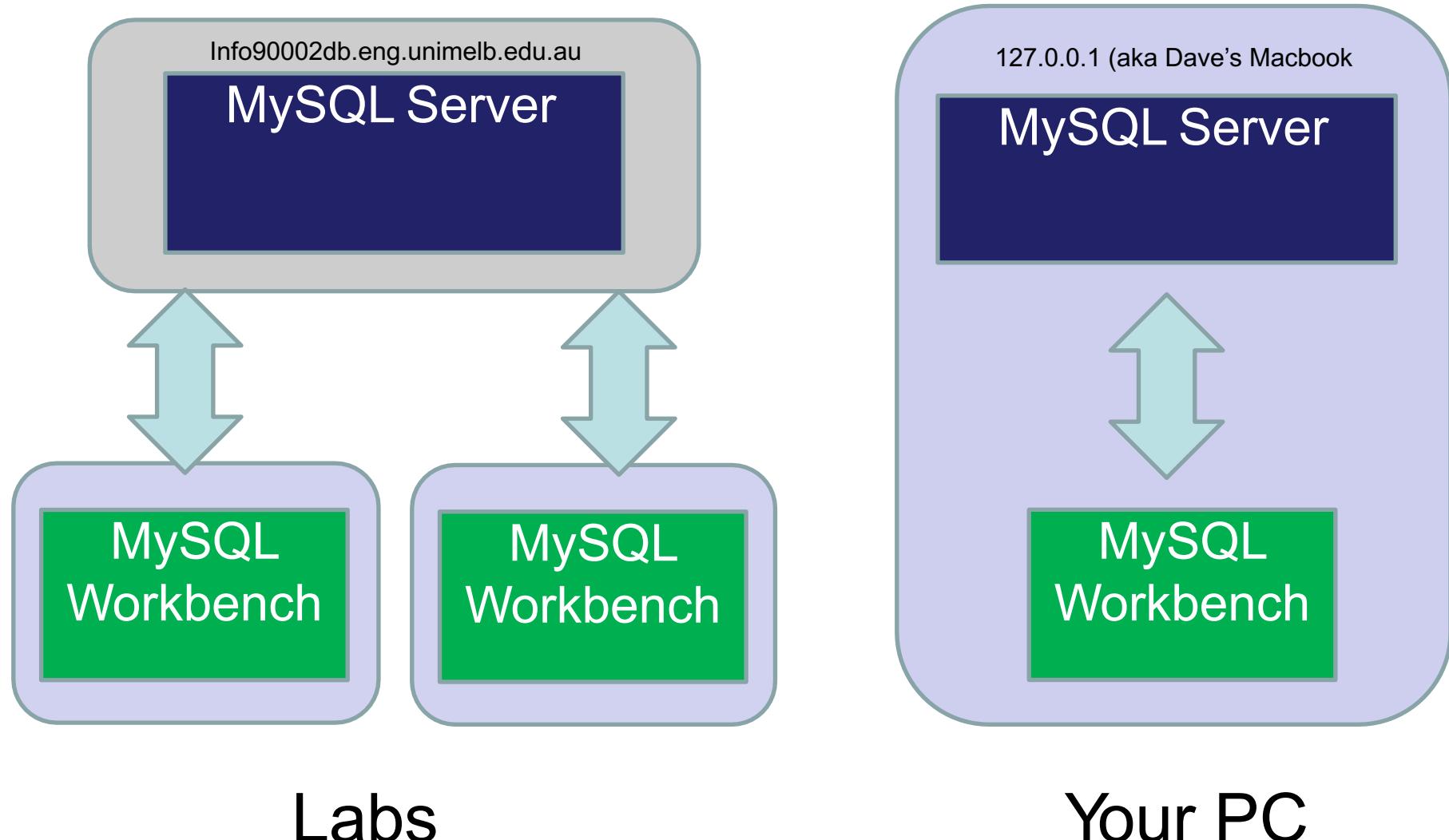


Introduction to MySQL Server and MySQL Workbench

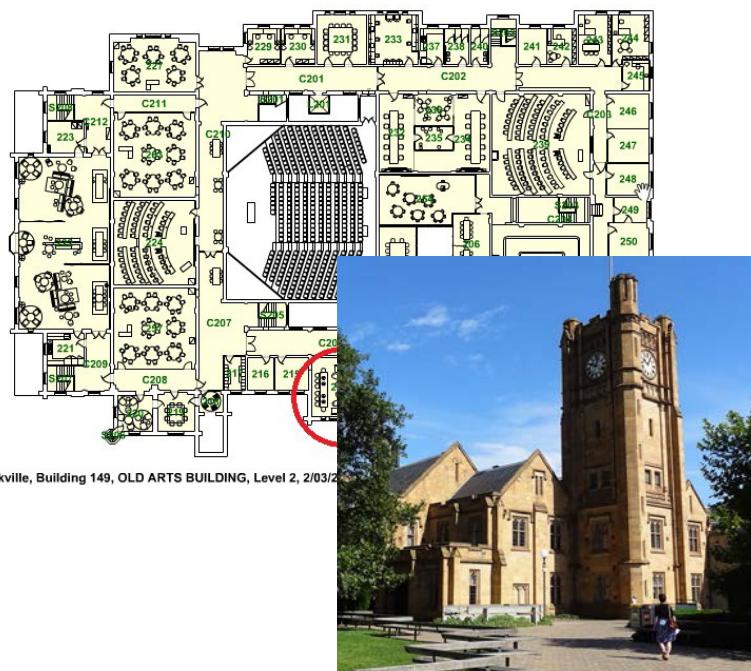


Server computer

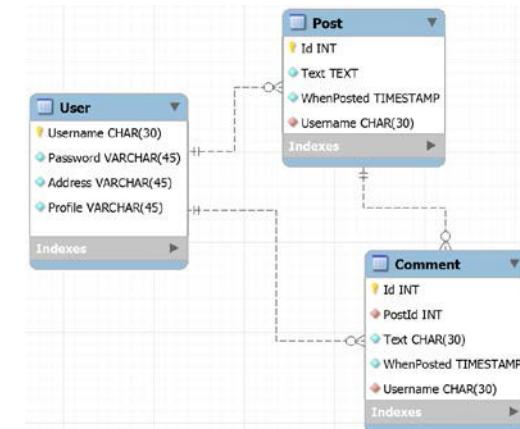


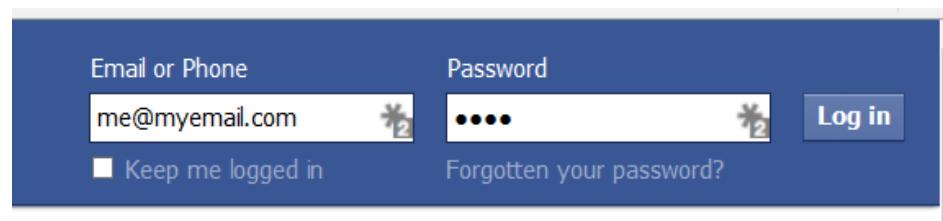


"What is unique about ICT ... In no other discipline is there such an emphasis on developing artefacts (e.g., computer and information systems) which are so abstract and complex and where modelling tools and methods are essential. The systems that ICT professionals deal with cannot be seen or handled in the same simple and direct manner as products of other applied disciplines ..." *from the Australian Computer Society CBOK*



Parkville, Building 149, OLD ARTS BUILDING, Level 2, 2/03/2012





User

Username	Password	Address	Profile
Anne	pass1234	1 Anne st	hi I am Anne
Bill	petsname	2 Bill st	this is Bill's profile
Christine	mystreet	3 Christine st	hi everyone this is Christine!!!
David	childname	4 David st	David's profile data

Post

Id	Text	WhenPosted	Username
1	Here's what I had for lunch	2015-01-24 20:09:02	Bill
2	What's everyone doing tonight?	2015-01-24 20:09:02	Anne
3	check out this great CAT VIDEOZ!!!	2015-01-24 20:09:02	David
4	now look what I had for dinner	2015-01-24 20:09:02	Bill
5	I love Game of Thrones	2015-01-24 20:09:02	Anne
6	I am posting a lot today...	2015-01-24 20:09:02	Anne

Comment

Id	PostId	Text	WhenPosted	Username
1	3	ha ha great video!!	2015-01-24 14:01:33	Christine
2	2	going to the movies	2015-01-24 14:01:33	Bill
3	2	oh that sound nice!	2015-01-24 14:01:33	Anne

Username	Password	Address	Profile
Anne	pass1234	1 Anne st	hi I am Anne
Bill	petsname	2 Bill st	this is Bill's profile
Christine	mystreet	3 Christine st	hi everyone this is Christine!!!
David	childname	4 David st	David's profile data

Id	Text	WhenPosted	Username
1	Here's what I had for lunch	2015-01-24 20:09:02	Bill
2	What's everyone doing tonight?	2015-01-24 20:09:02	Anne
3	check out this great CAT VIDEOZ!!!	2015-01-24 20:09:02	David
4	now look what I had for dinner	2015-01-24 20:09:02	Bill
5	I love Game of Thrones	2015-01-24 20:09:02	Anne
6	I am posting a lot today...	2015-01-24 20:09:02	Anne

Id	PostId	Text	WhenPosted	Username
1	3	ha ha great video!!	2015-01-24 14:01:33	Christine
2	2	going to the movies	2015-01-24 14:01:33	Bill
3	2	oh that sound nice!	2015-01-24 14:01:33	Anne

Instance (above)

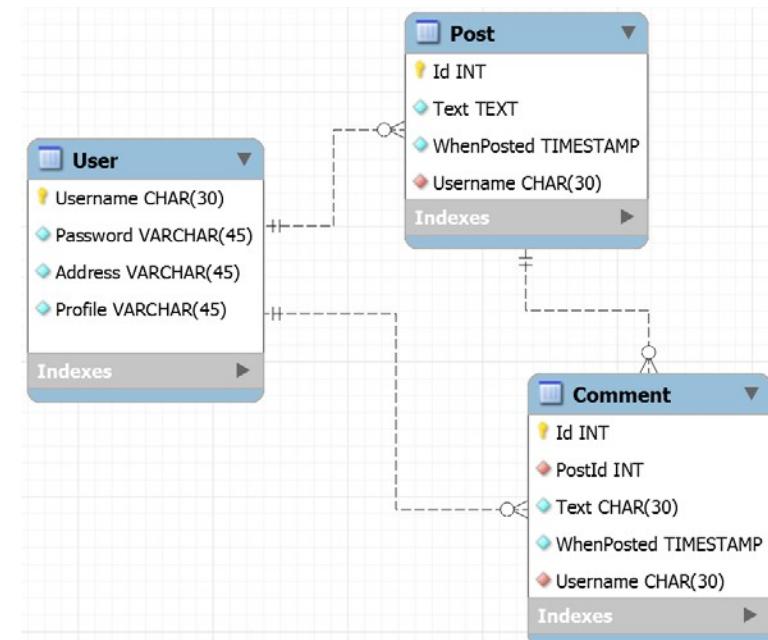
Relational notation (below)

User(username, password, address, profile)

Post(id, text, whenposted, username)

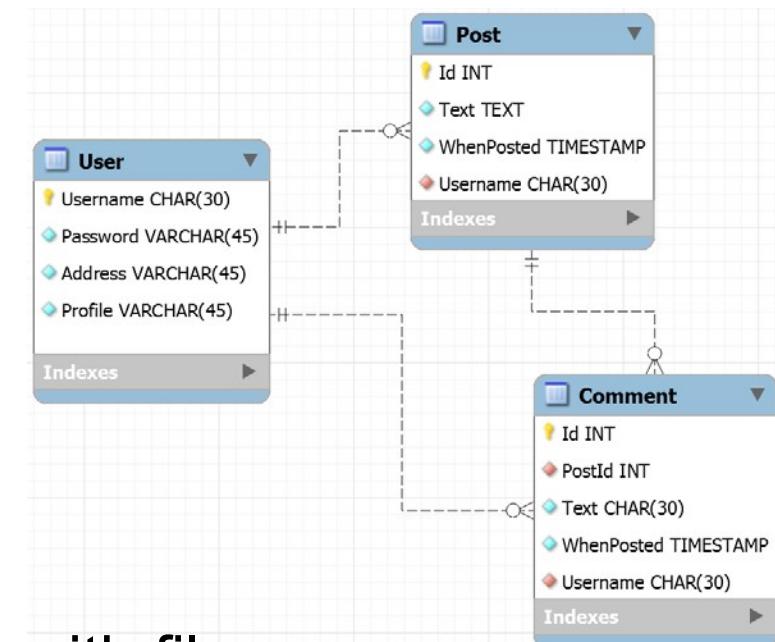
Comment(id, postid, text, whenposted)

Entity-Relationship diagram:





What commands do we need to manipulate the *structure* of our tables?



Working with tables is like working with files – there are 4 things you can do:

CREATE a table

DROP (i.e. delete) a table

ALTER a table (e.g. add a column)

RENAME a table



What commands do we need to manipulate the *contents* of our tables?

User

Username	Password	Address	Profile
Anne	pass1234	1 Anne st	hi I am Anne
Bill	petsname	2 Bill st	this is Bill's profile
Christine	mystreet	3 Christine st	hi everyone this is Christine!!!
David	childname	4 David st	David's profile data

You simply need to be able to:
SELECT, or read, data from the table
INSERT new rows into the table
DELETE existing rows from the table
UPDATE existing rows in the table



What commands do we
need to control *users'*
access to our tables?

(users = I.T. staff here)

Hostname:	info90002db.eng.unimelb.edu.au	Port:	3306
Username:	joebloggs		
Password:	Store in Vault ...	Clear	
Default Schema:			

There are 4 things you can do:

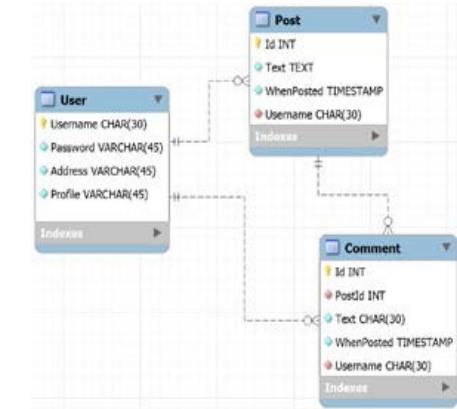
CREATE a user

DROP a user

GRANT a user access rights to a
table **REVOKE** those rights



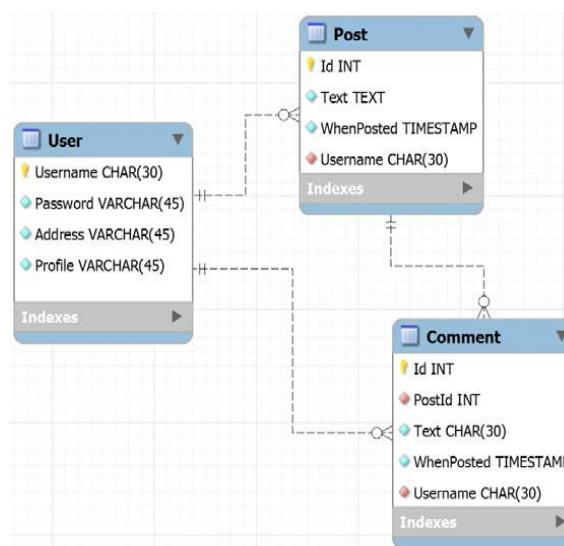
- Design the database
 - data modelling, E-R diagrams
- Implement the database
 - data definition language DDL
 - CREATE
 - DROP
 - ALTER
 - RENAME
 - data manipulation language DML
 - SELECT
 - INSERT
 - UPDATE
 - DELETE
- Data access / programming
 - data manipulation language DML
- Database administration
 - data control language DCL
 - GRANT
 - REVOKE





You will be given requirements such as:

- “We have many users.”
- Users may enter posts into the system. Each post is entered by exactly one user. Users can comment on posts.
- A post can be commented on many times.”



- What are the entities that need to be tracked?
- What attributes will be recorded about each entity?
- What are the relationships between entities?
- What are the cardinalities of relationships?



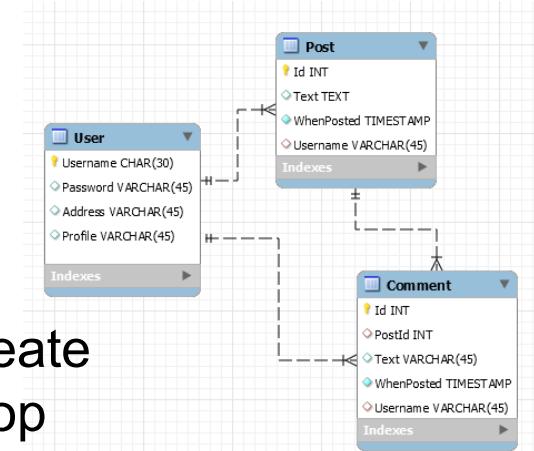
- An investment bank has a number of branches. Within each branch a number of departments operate and are structured in a hierarchical manner. The bank employs around 3000 staff who are assigned to work in the various departments across the branches. There are essentially three types of special employees where extra details required by the system. There are dealers who carry out investments who have limits imposed upon them for how much they can spend. There are IT compliance managers whose Basel2 role is required to be stored and there are HR managers that need have their assessment number recorded (along with other details not specified here).
- We need a database to record staff details including which department and branch they are assigned...

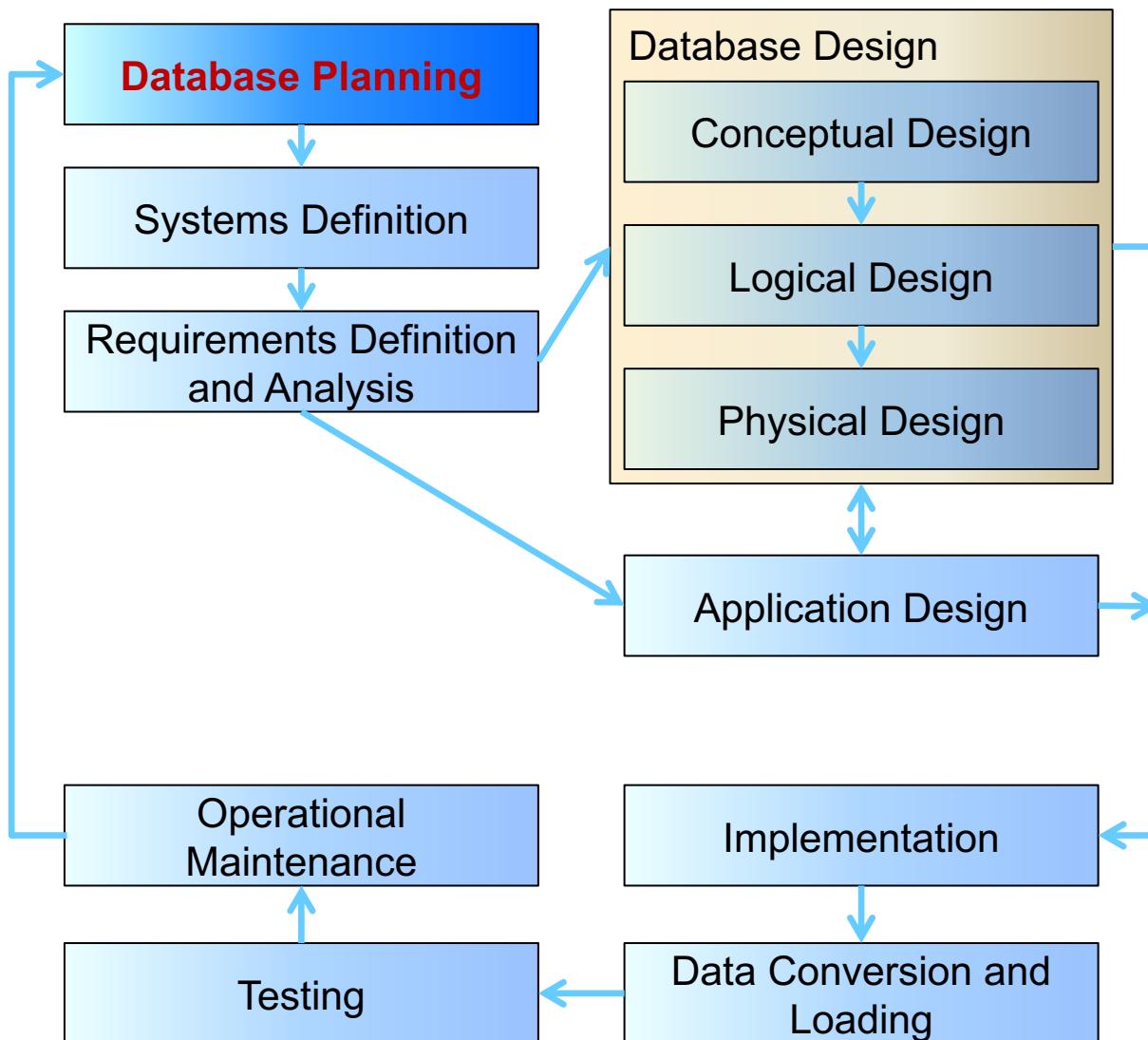


Database Development Lifecycle

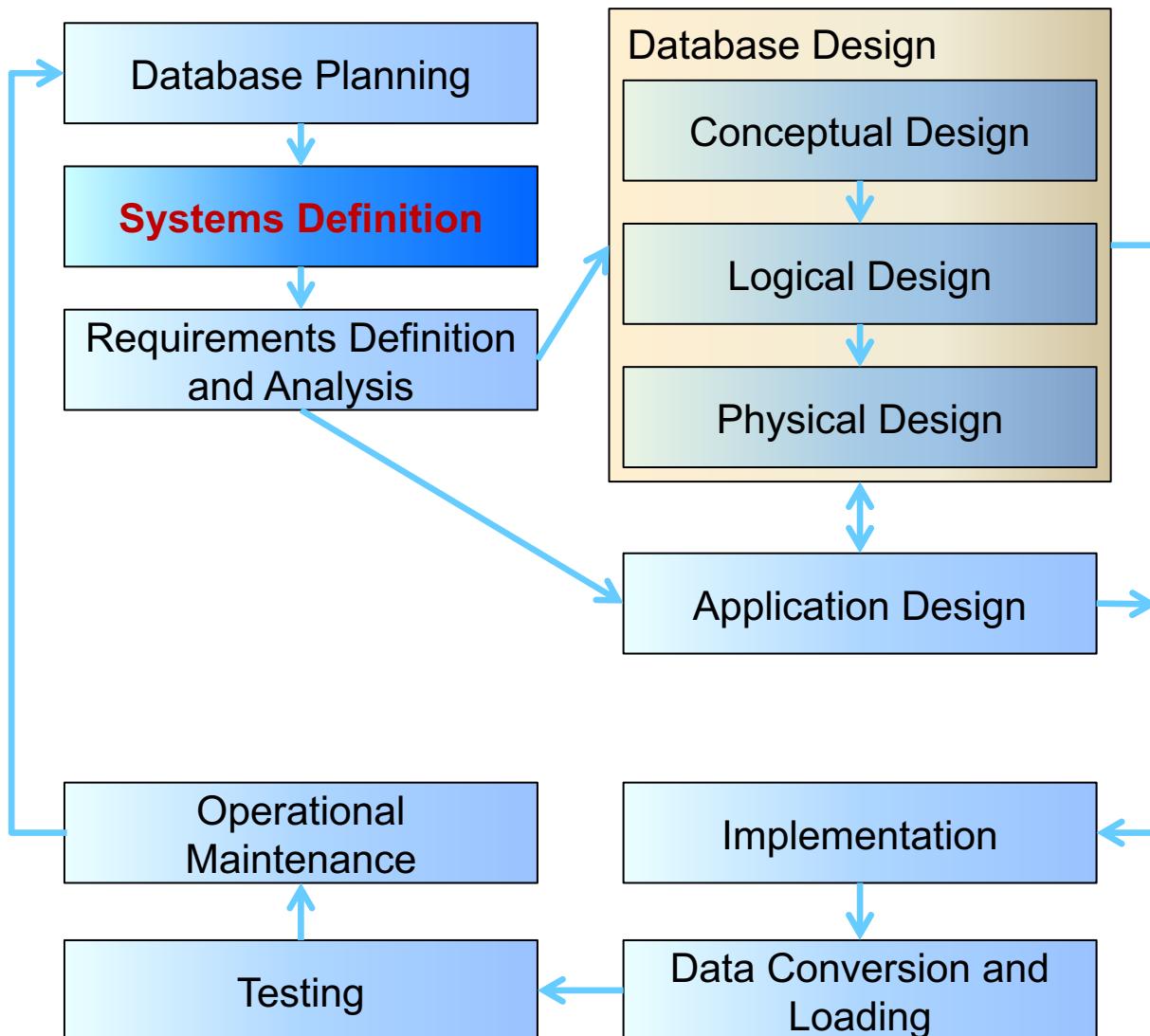


- Design the database
 - data modelling, E-R diagrams
- Implement the database
 - data definition language (DDL)
 - Create
 - Drop
 - Alter
 - Rename
 - data manipulation language (DML)
 - Select
 - Insert
 - Update
 - Delete
- Data access / programming
- Database administration
 - data control language (DCL)
 - Grant
 - Revoke

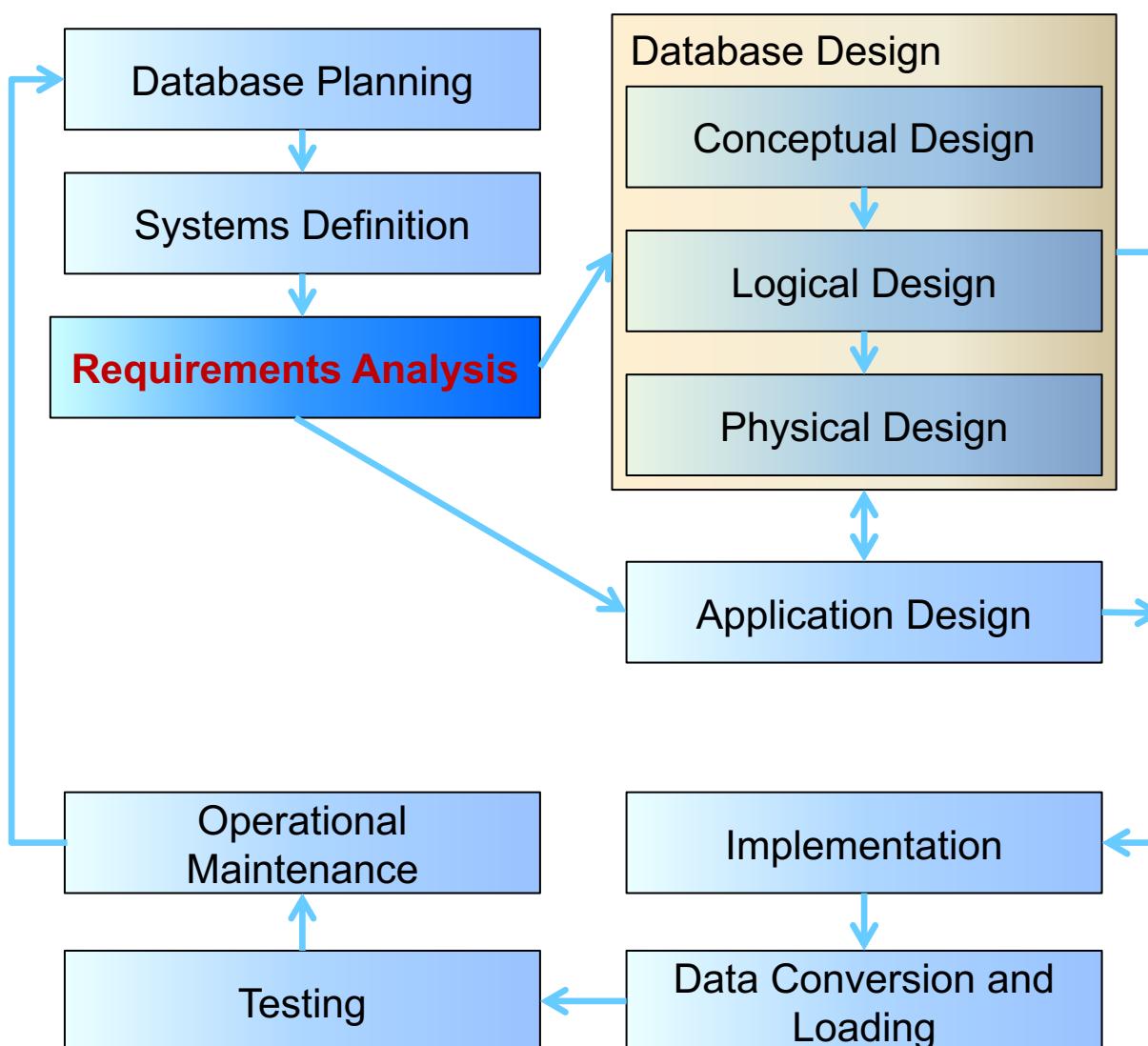




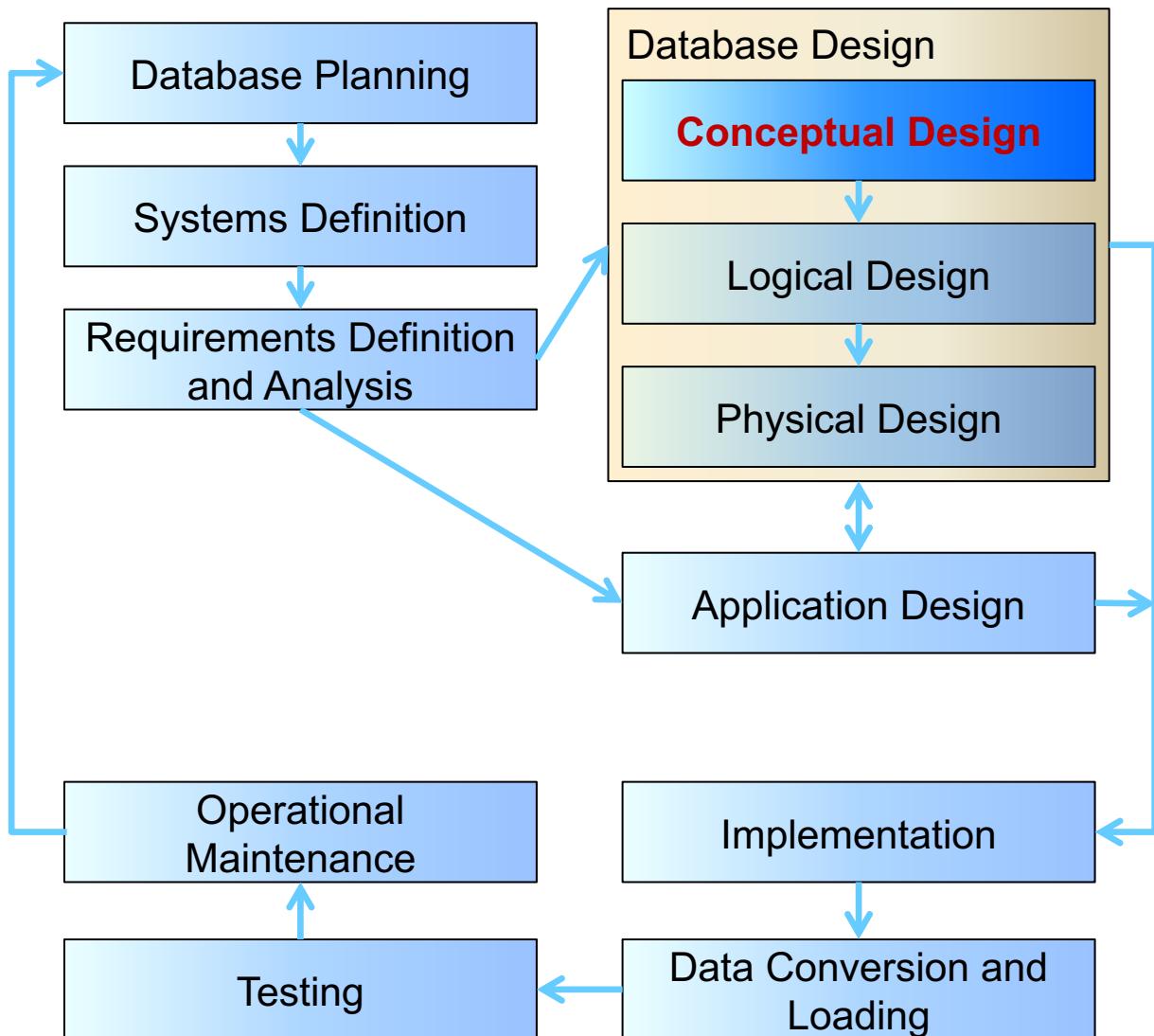
- Planning how to do the project.
 - How does the enterprise work
 - Enterprise data model
- How can the stages be completed efficiently and effectively.
- Outside scope of the course



- Specifying scope and boundaries
 - Users
 - Major user views
 - Application areas
- How does it interact with other systems
- User views – how the system operates from differing perspectives
- Outside scope of the course (slightly)

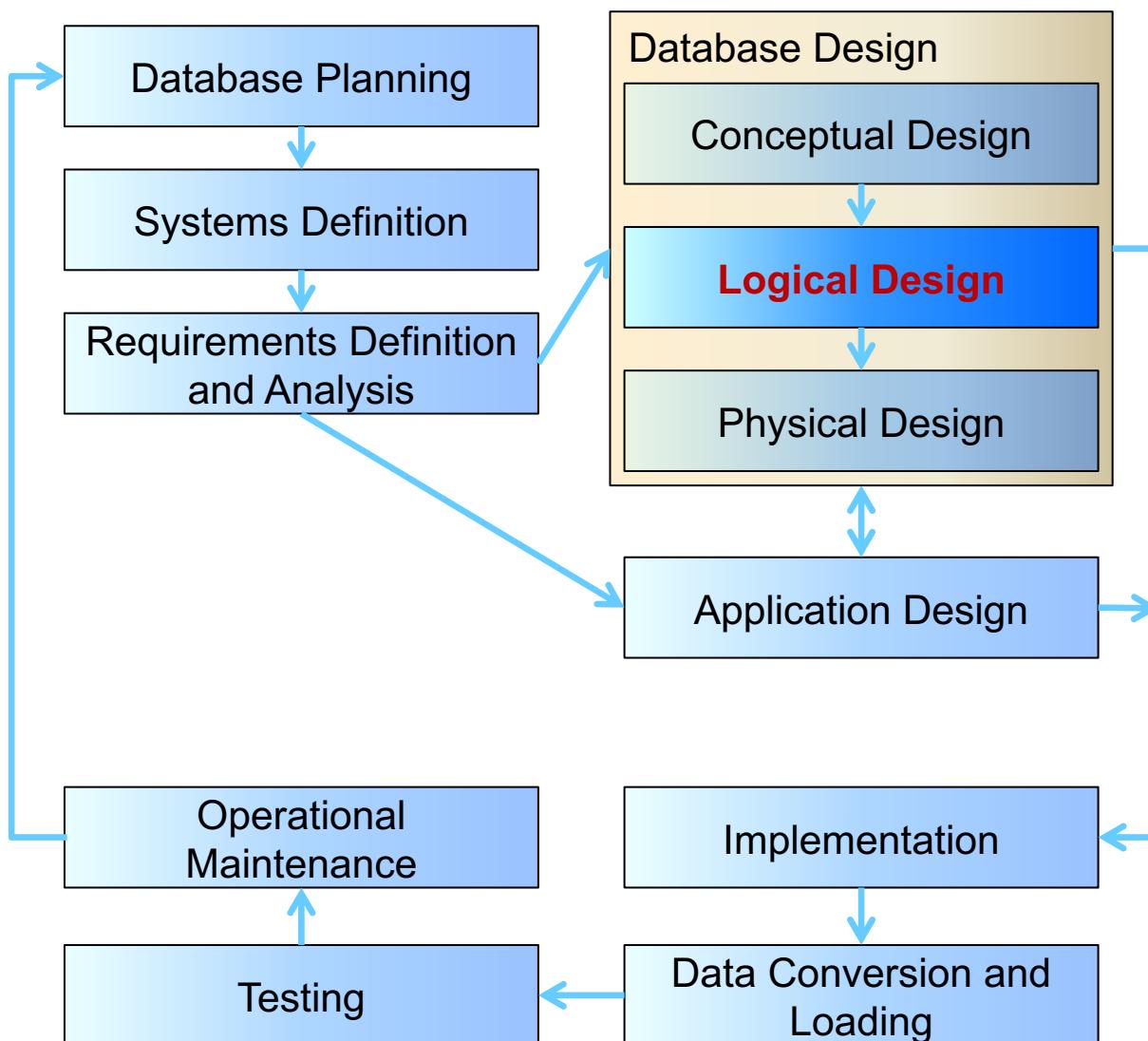


- Collection and analysis of requirements for the new system

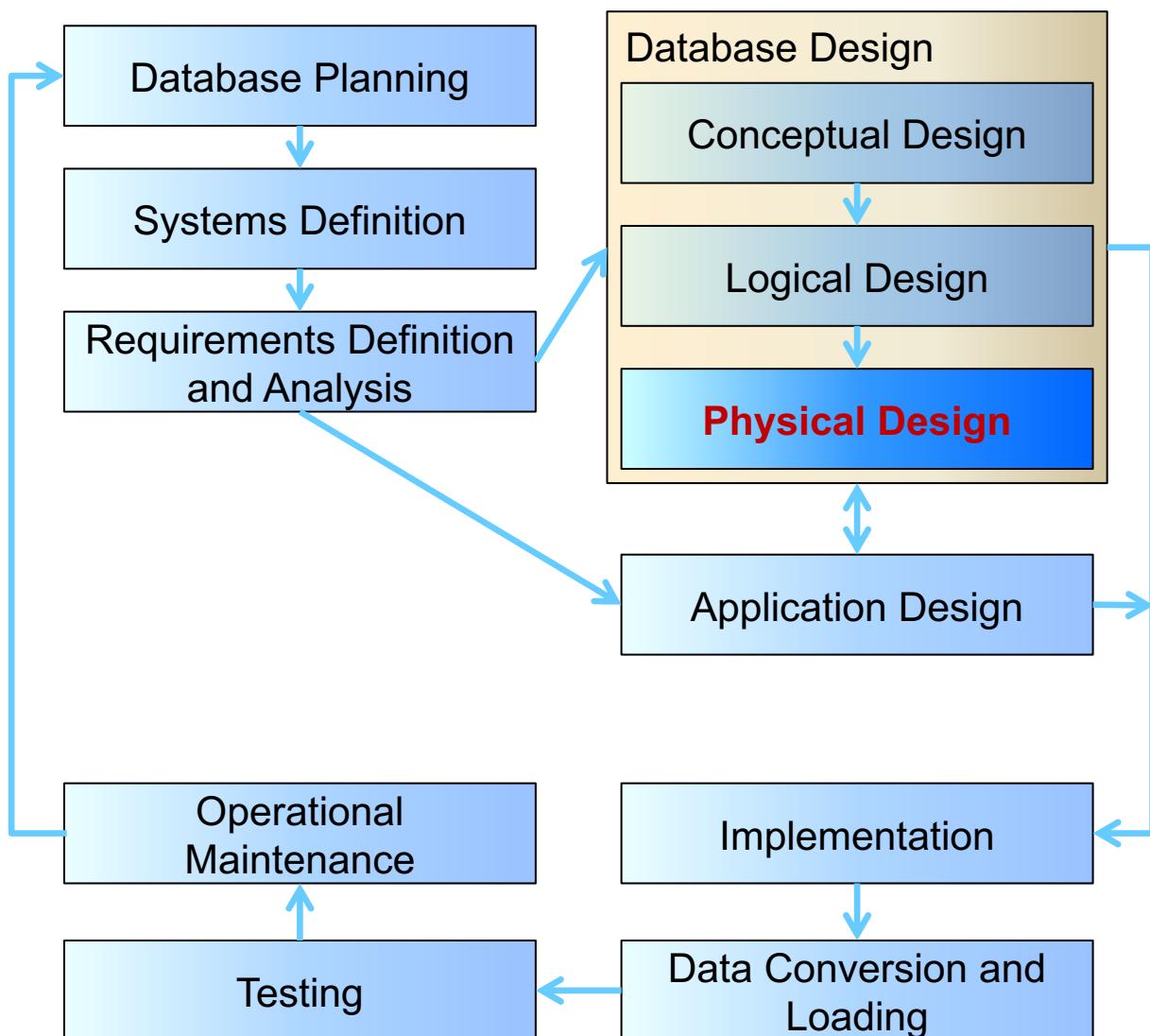


- High-level, first-pass model of entities and their connections
- Typically omits attributes*
- Could potentially be implemented in a non-relational database
- Thus can include many-to-many relationships, repeating groups, composite attributes

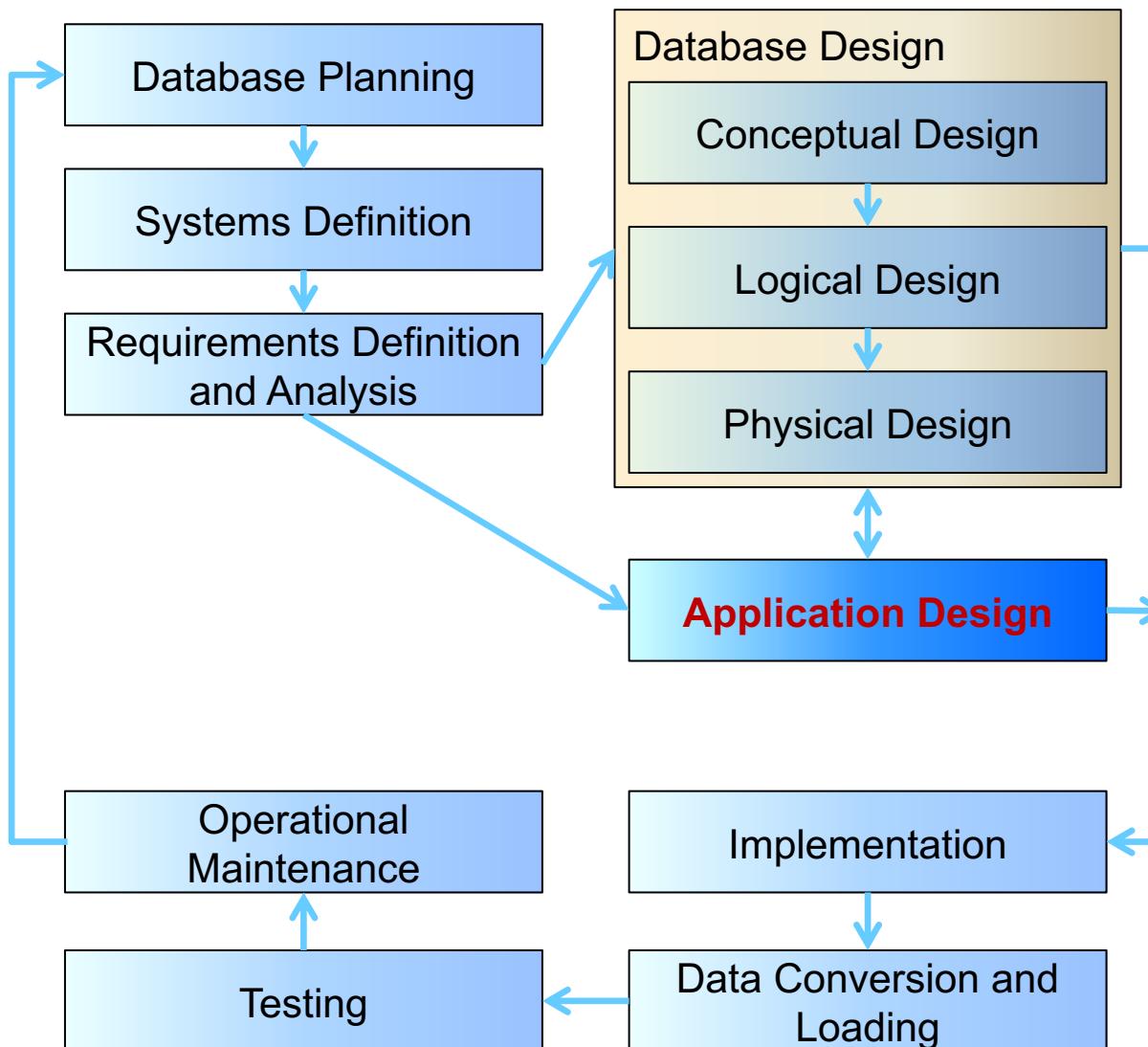
* Typically we list only attributes in the case study



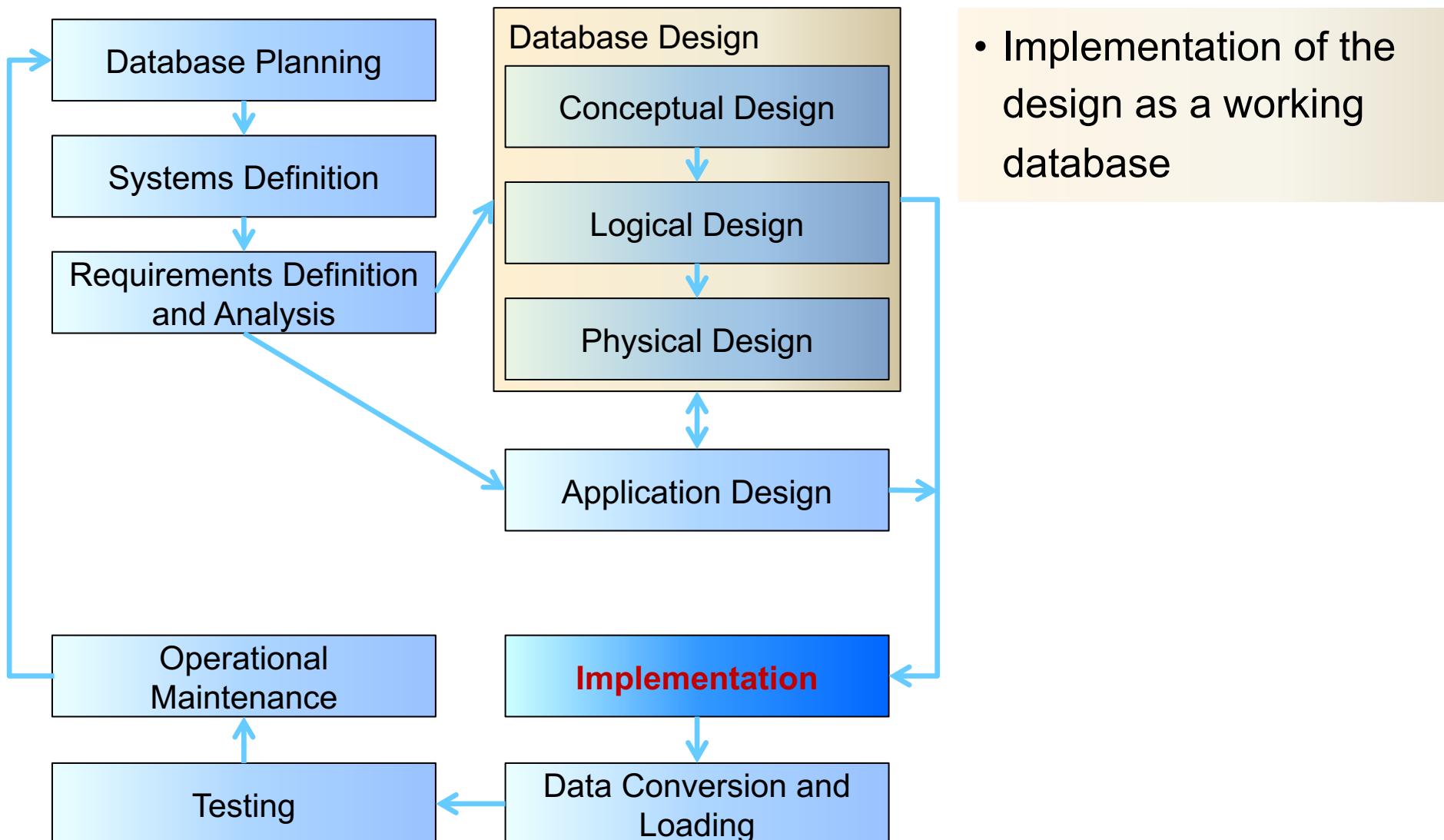
- Builds on the conceptual design
- Designing now for a relational database
- Includes columns and keys
- Independent of a specific vendor and other physical considerations

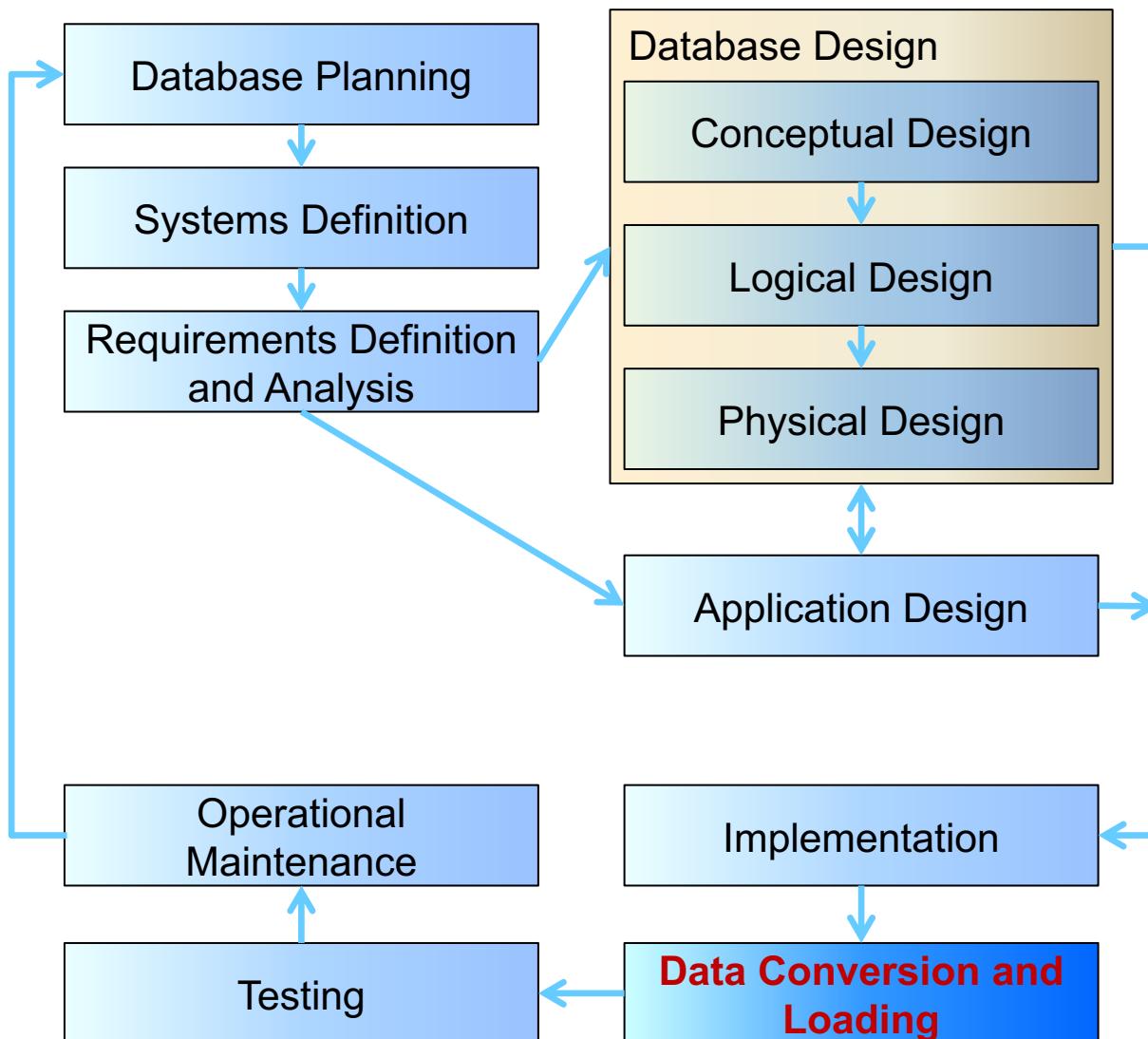


- Implements the logical design for a specific DBMS.
- Describes:
 - Base tables
 - Data types
 - Indexes
 - Integrity constraints
 - File organisation
 - Security measures
- We will cover some aspects of physical design

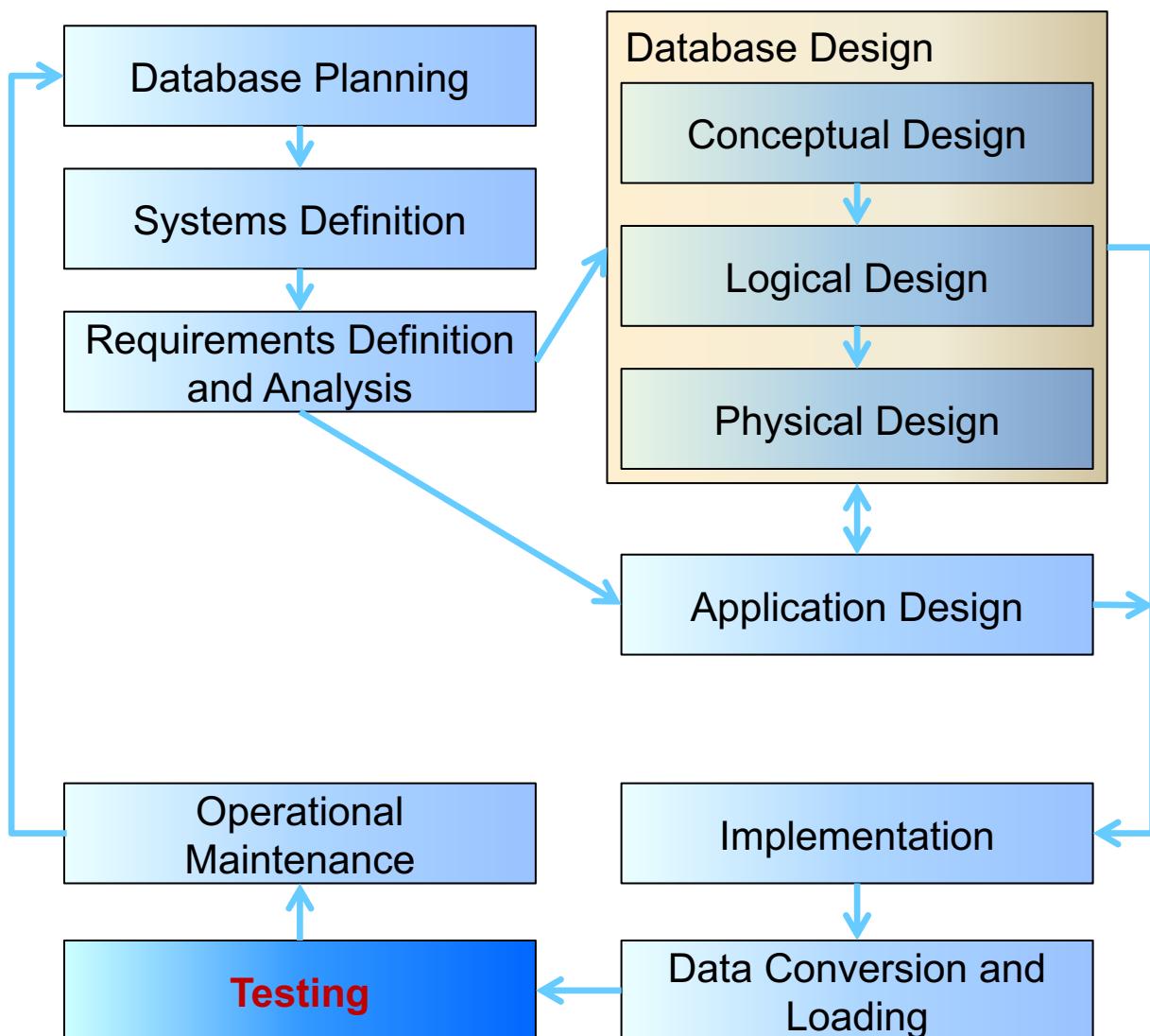


- Done in conjunction with database design
- Design of the interface and application programs that use and process the database
- Mostly outside scope of the course, but discussed in lectures 15 & 16.

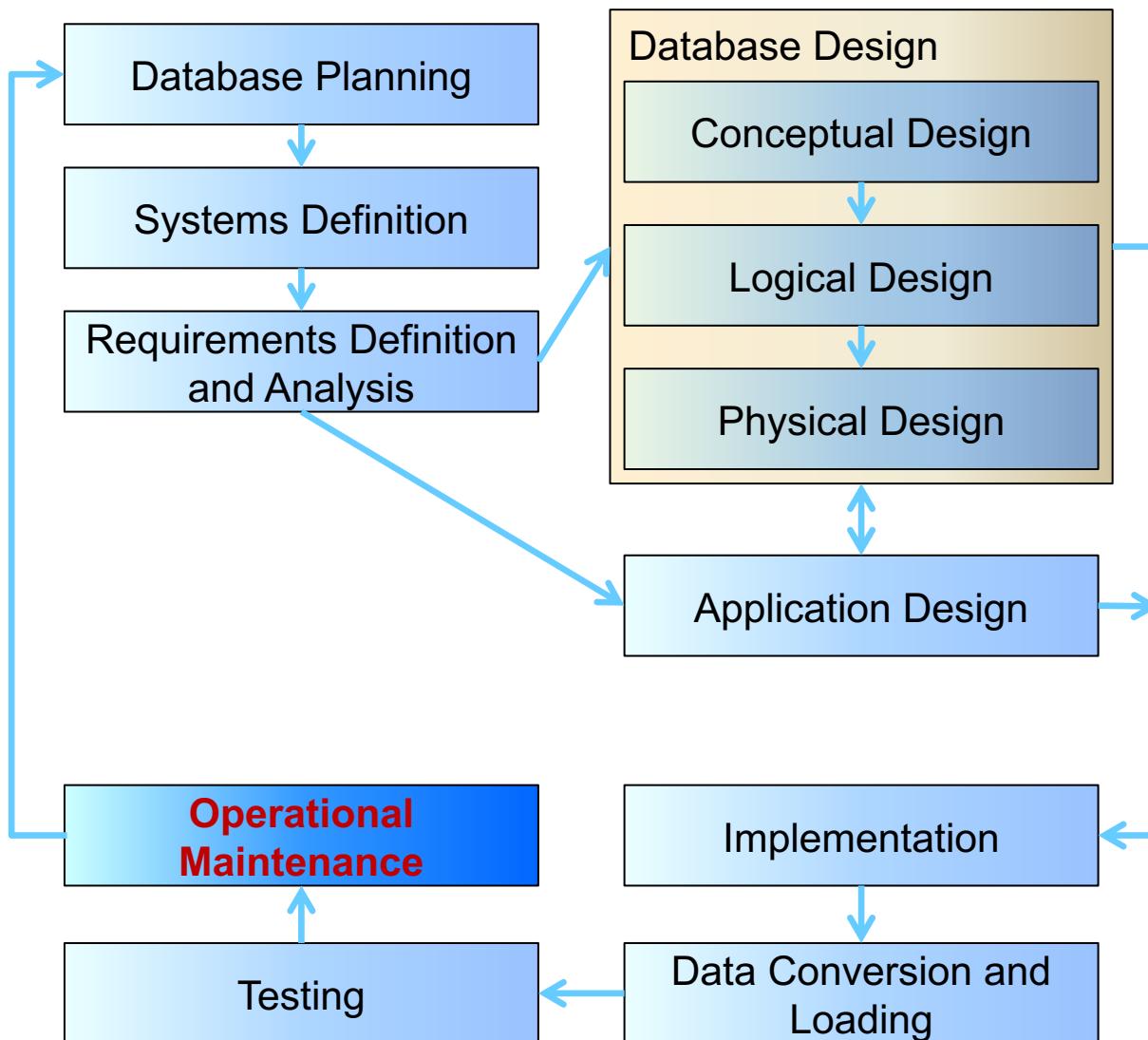




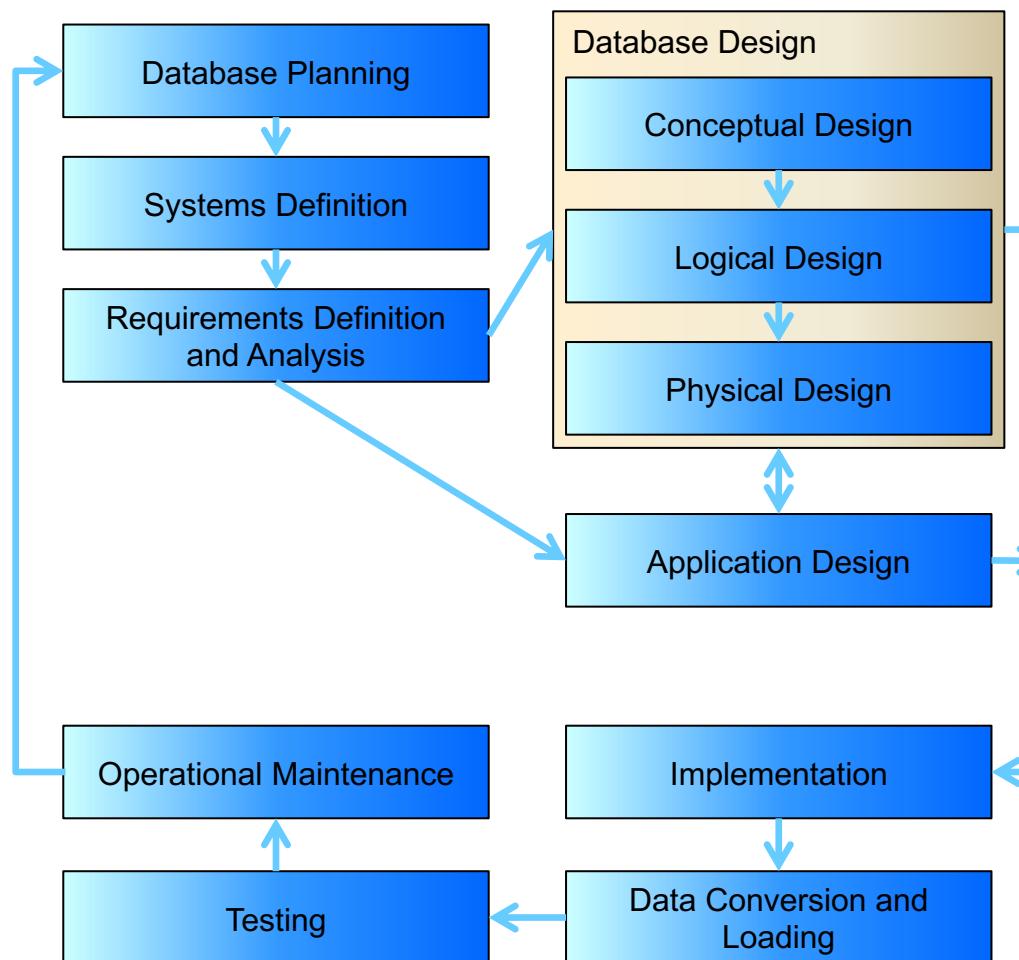
- Transfer existing data into the database
- Conversion from old systems
- Non trivial task
- **Mostly outside scope of the course, (concept covered in the Data Warehouse lecture)**



- Running the database to find errors in the design / setup
- Other issues also
 - Performance
 - Robustness
 - Recoverability
 - Adaptability
 - Security
- Mostly outside scope of the course (see **ISYS90086 Data Warehouseing**)



- The process of monitoring and maintaining the database following its commissioning
 - Monitoring and improving performance
 - Handling changes to requirements
-
- We will touch on some of these topics later in lectures 15, 20





- The database development life cycle
- Three stages of database development (conceptual, logical, physical)
- The ability to analyse a case study

* All material is examinable – these are the suggested key skills you would need to demonstrate in an exam scenario



- Logical & Physical Modelling
- Modelling with MySQL Workbench

Dr Greg Wadley
David Eccles



INFO 90002 Database Systems & Information Modelling

Lecture 1
Introduction to Subject
Introduction to Relational Databases
Introduction to MySQL