MONASH University

MONASH INFORMATION TECHNOLOGY

Week 12 – Where To?

Scheduled Final Assessment Preparation (previously Exam)

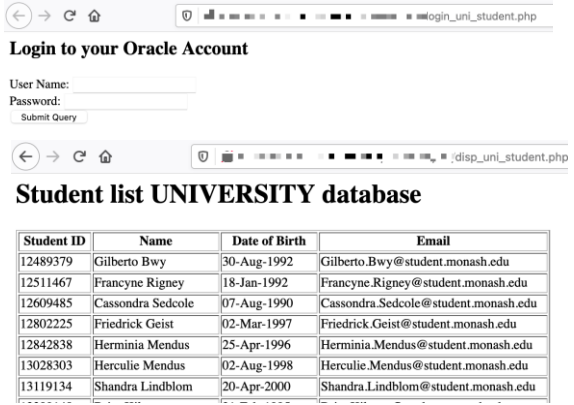
FIT3171 Databases
Semester 1 2022

Malaysia Campus

Application Development

- Database is the data store, need application as well
- Web based front ends
  - Wide range of approaches: PHP, ASP.NET, Oracle ORDS REST API's, etc
  - PHP OCI8 Package:
    - Simple example: web page for a user to log in and then return a list of students in the enrolment database

```php
<?php
//SQL query statement
$query = "SELECT studid,
 rtrim(studfname) || ' ' || rtrim(studlname) as sname,
 to_char(studdob,'dd-Mon-yyyy') as sbdate,
 studemail
 FROM uni.student
 ORDER BY studid";
```



- PL/SQL
  - backend development
  - Triggers, functions, procedures and packages
  - Procedure to change employee departments (empno, new dept)
    - **`move_employee (101, 2)`**

PL/SQL Procedure

```
create or replace
procedure move_employee (
    arg_empno in employee.emp_no%type,
    arg_newdeptno in department.dept_no%type)
as

INVALID_EMPLOYEE exception;
INVALID_DEPARTMENT exception;
dept_count number;
emp_count number;
currentempdeptno department.dept_no%type;

begin

    select count(*) into emp_count
        from employee where emp_no = arg_empno;

    if emp_count = 1 then

        select count(*) into dept_count
               from department where dept_no = arg_newdeptno;

        if dept_count = 1 then

          -- get employees current department number
          select dept_no into currentempdeptno
             from employee
             where emp_no = arg_empno;

          -- change employees department number
          update employee set dept_no = arg_newdeptno
                    where emp_no = arg_empno;

                -- decrement old department counter
                update department set dept_empcnt = dept_empcnt - 1
                   where dept_no = currentempdeptno ;

                -- increment new department counter
          update department set dept_empcnt = dept_empcnt + 1
                   where dept_no = arg_newdeptno ;

                commit;
          dbms_output.put_line ('Employee successfully moved');

        else

            raise INVALID_DEPARTMENT;

        end if;

    else

        raise INVALID_EMPLOYEE;

    end if;
```

FIT2104 - Web database interface

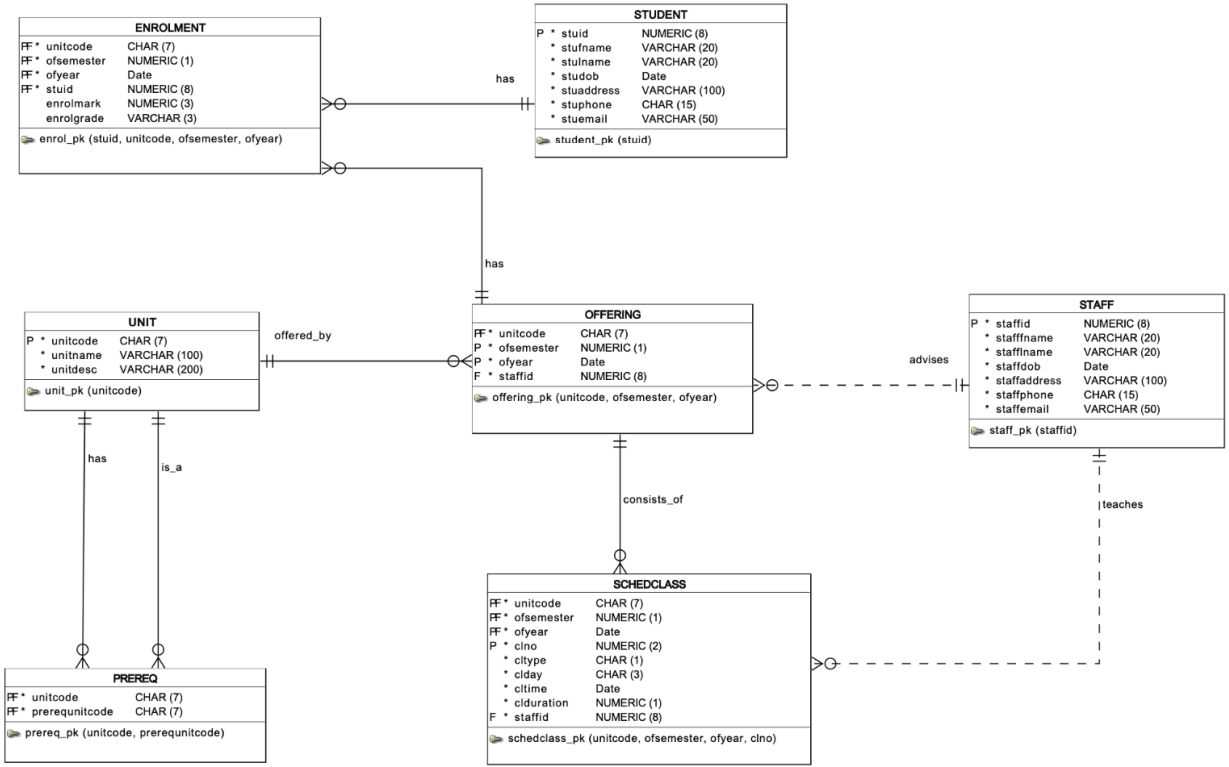
Learning outcomes

Expand all

On successful completion of this unit, you should be able to:

1. Explain the need and importance for application developers to have skills in this area of IT applications;
2. Describe and compare the key basic technologies which underly the development of web database applications;
3. Evaluate and assess the key technological issues confronting developers when building applications of this type;
4. Implement the key features of programming languages which are commonly used for developing web database applications;
5. Analyse, design, develop and implement a web database application using a commonly used programming language;
6. Evaluate and critique proposed web database solutions to a business problem.

Operational Database - Our main focus for the unit



Usage of database

- Example of a supermarket
- Decision making
  - Operational level
    - How often do we need to re-stock X-item?
  - Strategic and tactical level
    - Is there any branch that performs worse than the state average?
    - What is the total sales made by each state each year and across a number of years?

Operational Data vs. Decision Support Data

- Operational data
  - Mostly stored in relational database
  - Optimized to support transactions representing daily operations
  - Example:
    - How many students enrolled in a particular unit?
- Decision support data differs from operational data in three main areas:
  - Time span
  - Granularity
  - Dimensionality
  - Example:
    - What is the total number of students in the foundation units in each year (subtotal of the two semesters numbers) and the total across years, across a single unit.

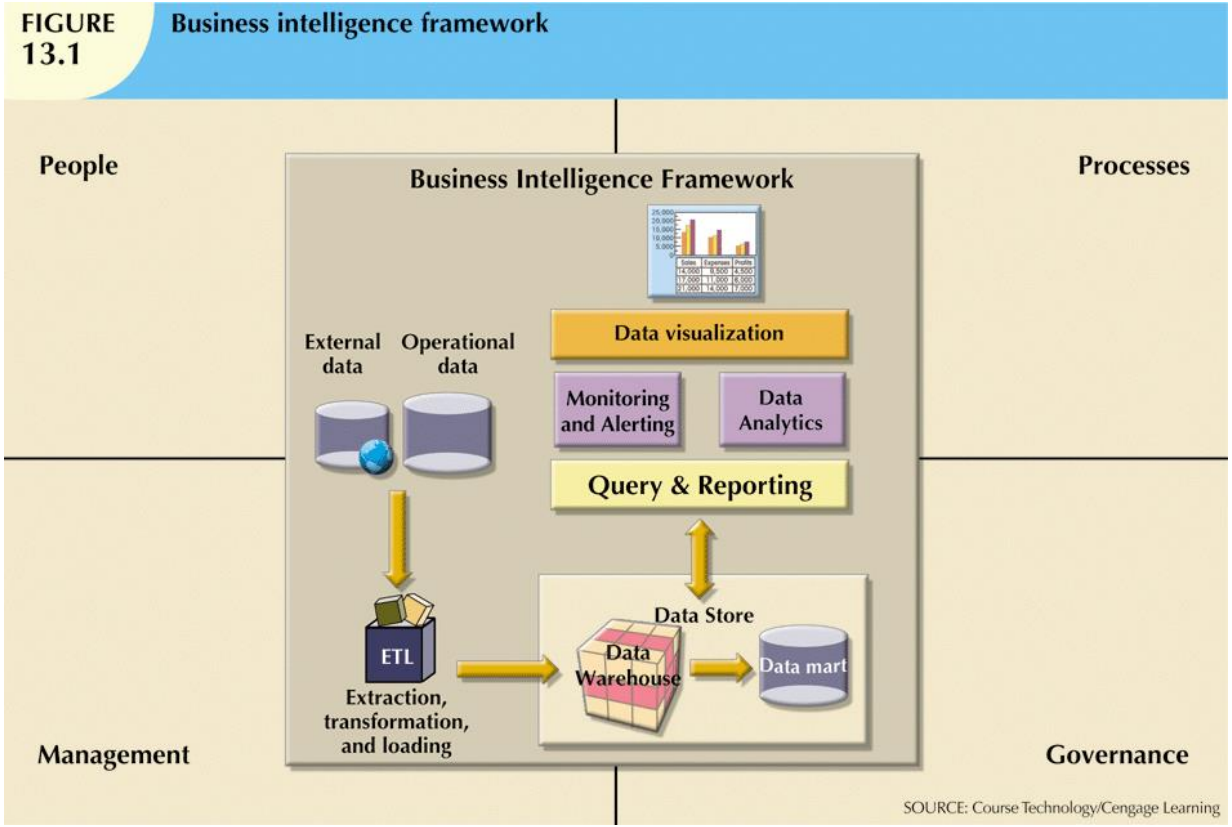
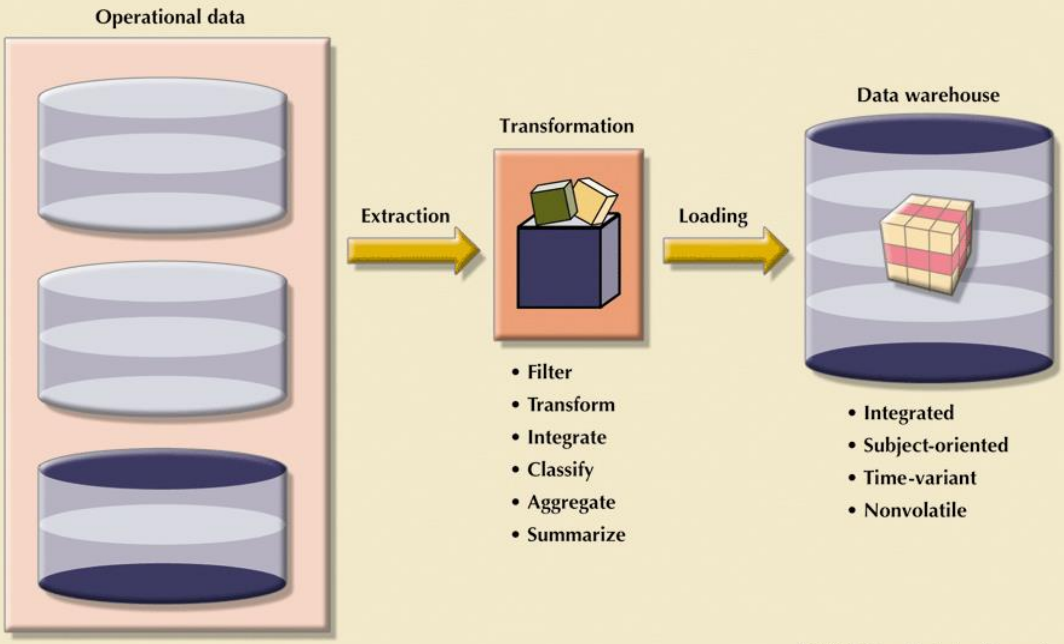


FIGURE 13.1
Business intelligence framework
People
Processes
Business Intelligence Framework
External data
Operational data
Data visualization
Monitoring and Alerting
Data Analytics
Query & Reporting
Data Store
ETL
Data Warehouse
Data mart
Extraction, transformation, and loading
Management
Governance
SOURCE: Course Technology/Cengage Learning

FIGURE 13.4 The ETL process



SOURCE: Course Technology/Cengage Learning

TABLE 13.5 Contrasting Operational and Decision Support Data Characteristics

CHARACTERISTIC	OPERATIONAL DATA	DECISION SUPPORT DATA
Data currency	Current operations Real-time data	Historic data Snapshot of company data Time component (week/month/year)
Granularity	Atomic-detailed data	Summarized data
Summarization level	Low; some aggregate yields	High; many aggregation levels
Data model	Highly normalized Mostly relational DBMSs	Non-normalized Complex structures Some relational, but mostly multidimensional DBMSs
Transaction type	Mostly updates	Mostly query
Transaction volumes	High update volumes	Periodic loads and summary calculations
Transaction speed	Updates are critical	Retrievals are critical
Query activity	Low to medium	High
Query scope	Narrow range	Broad range
Query complexity	Simple to medium	Very complex
Data volumes	Hundreds of gigabytes	Terabytes to petabytes

The Data Warehouse

- Database size
  - 2014 world's largest data warehouse SAP at 12.1 Petabytes (around 12,400 Terabytes)
  - DBMS must support very large databases (VLDBs)
- Integrated, subject-oriented, time-variant, and nonvolatile collection of data
  - Provides support for decision making
- Usually a read-only database optimized for data analysis and query processing
- Requires time, money, and considerable managerial effort to create

FIT3003 - Business intelligence and data warehousing

Learning outcomes

Expand all

On successful completion of this unit, you should be able to:

1. Design multi-dimensional databases and data warehouses;
2. Use fact and dimensional modelling;
3. Implement online analytical processing (OLAP) queries;
4. Explain the roles of data warehousing architecture and the concepts of granularity in data warehousing;
5. Create business intelligence reports using data warehouses and OLAP.

IOT - the explosion - Data, Data, Data …..

https://youtu.be/bUnOo10z42k

Issue 1: Data Volume

FIT3176 - Advanced database design

1. Describe various types of non-relational database systems, including NoSQL;
2. Design and model document-store and wide column-store databases;
3. Compare and contrast between relational and non-relational database modelling;
4. Explain the concepts of transactions in non-relational systems;
5. Implement document-store and wide column-store systems;
6. Construct applications using a graph database system;
7. Demonstrate graph query processing.

Issue 2: Data Processing

Big Data Processing

- Computer systems
  - –Parallel computer
    - •A single machine with massive number of CPUs.
  - –Cluster of computers
    - •Multiple machines connected via network.
    - •Commodity computer.
- Database structure
  - –Non-relational database (NoSQL)
    - •No update, append only.
    - •Optimised for a ’main’ operation
    - •Examples: MongoDB, Cassandra
  - –Distributed File Systems
    - •HDFS (Hadoop File Systems)
    - •Parquee File Systems
- Parallel data processing
  - –Hadoop
  - –Spark
- In Memory database

FIT3182 - Big data management and processing

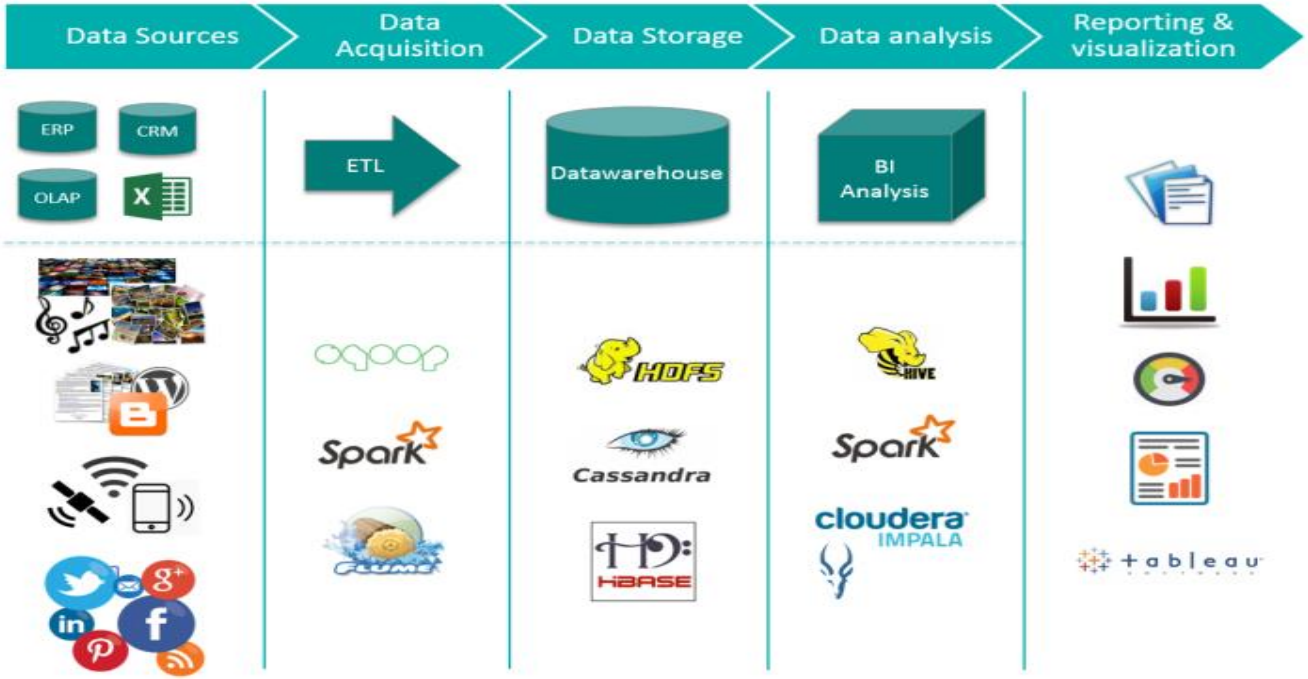
Learning outcomes

Collapse all

On successful completion of this unit, you should be able to:

1. identify big data concepts and technologies;
2. write and interpret parallel database processing algorithms and methods;
3. use big data processing frameworks and technologies;
4. describe and compare NoSQL technologies;
5. use big data streaming technologies.

Data Processing Ecosystem



http://www.clearpeaks.com/blog/big-data/big-data-ecosystem-spark-and-tableau

"Horses for Courses"

- Conventional RDBMS will continue play an important and significant role in OLTP (Online Transactions Processing)
- Increasingly now a *range* of database products are available, need to select appropriate product/model for task at hand.

Unit Scheduled Final Assessment

2022 Final Assessment

- **2 HOUR 10 minutes**
  - Time includes reading time
  - **eExam, Closed Book with permitted sites/software, online supervision, no aids (eg. dictionary)**
    - **Full access to the Unit Moodle teaching material, unit software and databases**
    - Throughout the exam the screen of your device will be recorded (supervisor watching, AI supervision and post exam analysis)
    - Must not access multiple devices or screens (may only use one screen)
    - Access to any site or software which has not been approved will result in exam being cancelled
    - Posting of questions to other sites/users will/has resulted in suspension/expulsion from course (Chegg and such sites are monitored)

2022 Final Assessment continued

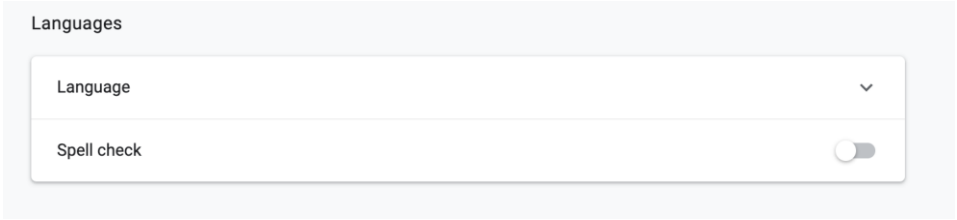
- Paper will be available via the eExam platform
  - https://www.monash.edu/exams/electronic-exams
  - Note Academic Staff for this unit have NO control over this platform, all issues/concerns MUST be raised with the eExam staff
    - Getting help and support - Current students (monash.edu)
    - This link should be used for both pre exam, during exam and post exam issues
  - Complete the e-exam General Knowledge Practice to see form of eExam papers

2022 Final Assessment continued

- During the exam you can highlight text as you read the question for emphasis

- Spelling check is turned off in the exam paper, however you also need to turn it if off in your browser

  eg. for Chrome and Edge:



2022 Final Assessment Structure

- 100 marks 40% of your final mark for this unit.
  - Minimum to pass - overall:
    - 45% in-semester, 45% final assessment and 50% overall
- Questions:
  - Cover theory and application
  - Timing - 100 marks, 120 mins - **1 mark/minute target**
  - SQL/NoSQL case study will be released on Moodle the day before the scheduled exam, **do not bring** the case study to the Final Assessment or discuss it in the Ed forums
  - Oracle tables for the case study will be available during the exam time **ONLY - tables must not be created in your account**
- Final Assessment when - **your responsibility to determine**

Permitted Resources during Exam

- **Online Resources**
  - FIT3171 2022 Semester 1 Moodle site:
    - https://lms.monash.edu/course/view.php?id=136243
  - Oracle ORDS Server:
    - https://ora-fit.ocio.monash.edu:8441/ords/sql-developer
  - MoVE (File Explorer and SQL Developer only):
    - https://move.monash.edu/
  - Visual Studio Code for the Web
    - https://vscode.dev/
  - MongoDB Shell:
    - https://www.mongodb.com/docs/manual/tutorial/getting-started/
- No collaboration tools or websites other than those listed above are permitted

Permitted Resources during Exam continued

- **Locally installed software**
  - Oracle SQL Developer
  - MongoDB Shell
- Copy and Paste is permitted **ONLY**
  - where your answer is developed in the Oracle or MongoDB environment, and
  - where indicated permitted for the particular question on the exam paper
- Copy and Paste must always be your own content, and not the duplication of anyone else's work

Exam cover sheet - Instructions

Instructions

- This is an **Closed Book with Specifically Permitted Items** exam.
  - **ONLY** the following online resources may be used:
    - FIT9132 2022 Semester 1 Moodle site: https://lms.monash.edu/course/view.php?id=136238
    - Oracle ORDS Server: https://ora-fit.ocio.monash.edu:8440/ords/sql-developer
    - MoVE (File Explorer and SQL Developer only): https://move.monash.edu/
    - Visual Studio Code for the Web https://vscode.dev/
    - MongoDB Shell: https://www.mongodb.com/docs/manual/tutorial/getting-started/
  - **ONLY** the following locally installed software may be used
    - Oracle SQL Developer
    - MongoDB Shell
  - No collaboration tools or websites other than those listed above are permitted
  - Your answer for Q3 - Q8 (Section B and Section C) should be developed in the appropriate database environment. Copy and Paste is permitted **ONLY** for these answers, but must always be your own content, and not the duplication of anyone else's work.
- Please answer **ALL** questions.

Authorised Materials

Closed Book with Specifically permitted items

No Calculator

Websites: FIT9132 Sem 1 Moodle Site; Oracle ORDS Server; MoVE; VS Code for Web; MongoDB Shell
Software: Oracle SQL Developer; MongoDB Shell.

Rules

During an exam, you must not have in your possession any item/material that has not been authorised for your exam. This includes books, notes, paper, electronic device/s, mobile phone, smart watch/device, calculator, pencil case, or writing on any part of your body. Any authorised items are listed above. Items/materials on your desk, chair, in your clothing or otherwise on your person will be deemed to be in your possession.

You must not retain, copy, memorise or note down any exam content for personal use or to share with any other person by any means following your exam.

You must comply with any instructions given to you by an exam supervisor.

As a student, and under Monash University's Student Academic Integrity procedure, you must undertake your in-semester tasks, and end-of-semester tasks, including exams, with honesty and integrity. In exams, you must not allow

2022 Final Assessment

- All content specified during your semester of study in this unit is examinable either directly or may be required in the context of a question. This includes but is not limited to:
  - Pre-reading (weekly Coronel & Morris chapters), Pre-recorded Videos, HP5 review exercises
  - Forum Slides and Videos
  - Tutorial Session Notes/Solutions, and
  - all other Moodle Materials (except where explicitly stated NOT EXAMINABLE).

2022 Mock Final Assessment

- Serves to provide an overview of the general structure of the final assessment paper only.
- Available Friday 27th May at 3 PM Malaysia time.
  - Sample solution available Wednesday, 8th June 7AM Malaysia time (please attempt the mock paper before accessing the sample solution)
  - Both sample paper and solution close 6 hrs before Scheduled Final Assessment (Exam)
  - Forums also close 6 hrs before Scheduled Final Assessment (Exam)
- Link on Moodle under - "Scheduled Final Assessment" on Assessments page
- To protect the integrity of the paper: NO ACTUAL FINAL ASSESSMENT PAPER QUESTIONS are included; and the COMPOSITION OF THE SUB QUESTIONS are SUBJECT TO CHANGE.
- Suggest you leave it until you can attempt it under time limit of 2 hours 10 minutes (**time control is critical**)
- Allows multiple attempts

Week 2 & 5 Data Modelling

Not directly examined

- Conceptual vs Logical Level
- Entity
  - Strong vs weak
  - Associative entity
- Types of attributes
- Relationship
  - Connectivity - type : one-to-one, one-to-many, many-to-many
  - Cardinality
  - Participation
  - Identifying vs Non-identifying.
- Mapping from Conceptual to Logical
  - E.g. Mapping many-to-many

Week 3 – Relational Model EXAMINED

- Relational model properties.
- Keys
  - Superkey, Candidate Key, Primary Key
  - Foreign Key
- Data Integrity
  - Entity integrity
  - Referential Integrity
- Relational Algebra
  - Understanding of efficiency

Week 4 – Normalisation Not directly examined

- UNF to 3NF
  - **Mapping form to UNF**
  - UNF to 1NF – remove repeating group.
  - 1NF to 2NF – remove partial dependency (***general*** definition)
  - 2NF to 3NF – remove transitive dependency.
- Dependency diagrams
  - For example: cust_id -> cust_name, cust_phone, cust_balance
- Be careful in choosing the PK!
- Mapping a set of 3NF relations to a logical model

Week 6 – Data Definition Language Not directly examined

- CREATE TABLE statements
  - Primary key definition
  - Foreign key definition
  - Other Constraints
- ALTER
- INSERT
  - Adherence to referential integrity constraints
    - Order of insertion
- Oracle Sequence
- UPDATE (DML)
- DELETE (DML)

Week 7 and 9 – SQL EXAMINED

- Single table retrieval with predicate
- Join
  - Natural join
  - Outer join
- Aggregate functions
- Set Operators
- Subquery
- Oracle functions
  - TO_CHAR, TO_DATE, NVL, UPPER, LOWER, ROUND, RTRIM/LTRIM, LPAD/ RPAD
- PAYROLL SQL revision available on Moodle

Week 8 – Transaction Management EXAMINED

- Transactions
  - transaction boundaries (start and end)
  - use of commit/rollback
- ACID properties.
- Transaction problems.
- Transaction management with locks.
- Wait For Graphs
- Two-Phase Locking
- Restart and Recovery using Transaction Log.

Week 10 – PL/SQL Not directly examined

Trigger:

- Before vs After
- Row level vs statement level
- Use trigger for update cascade
- Use trigger to maintain data integrity
- Test harness

Procedure:

- Input and output arguments
- Use procedure to encapsulate business logic
- Test harness

Week 11 – noSQL EXAMINED

- Characteristics of Big Data
- Four major noSQL models
- Role of Hadoop (basic only)
- Using SQL to generate JSON document
- MongoDB CRUD commands
  - C: insert (One and Many)
  - R: find (predicates, count(), pretty())
  - U: update (One and Many)
  - D: delete (One and Many)

Week 12 – Database Future Directions

- The content of week 12's lecture
  - Database Trends
  - Future directions

Is NOT examinable (questions related to this week's new content will not appear on the paper)

Consultations for Scheduled Final Assessment

- Online consultation sessions will be provided.
  - –Details to be posted on Moodle

- Don’t come to consultations in the hope of obtaining some 'extra' information about the paper
  - Session intended to clear up any issues *YOU* find as you prepare for the Scheduled Final Assessment

http://blog.proqc.com/administrative-professionals-quality-thank-you/