Where To?

Scheduled Final Assessment Preparation (previously Exam)

FIT9132 2021 Semester 2

Application Development

- Database is the data store, need application as well
- Web based front ends
  - Wide range of approaches: PHP, ASP.NET, etc
  - Very Rudimentary:
    - 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";
```


Login to your Oracle Account

User Name:
Password:
Submit Query

Student list UNIVERSITY database

Student ID	Name	Date of Birth	Email
12489379	Gilberto Bwy	30-Aug-1992	Gilberto.Bwy@student.monash.edu
12511467	Francyne Rigney	18-Jan-1992	Francyne.Rigney@student.monash.edu
12609485	Cassondra Sedcole	07-Aug-1990	Cassondra.Sedcole@student.monash.edu
12802225	Friedrick Geist	02-Mar-1997	Friedrick.Geist@student.monash.edu
12842838	Herminia Mendus	25-Apr-1996	Herminia.Mendus@student.monash.edu
13028303	Herculie Mendus	02-Aug-1998	Herculie.Mendus@student.monash.edu
13119134	Shandra Lindblom	20-Apr-2000	Shandra.Lindblom@student.monash.edu

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

```
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;
```
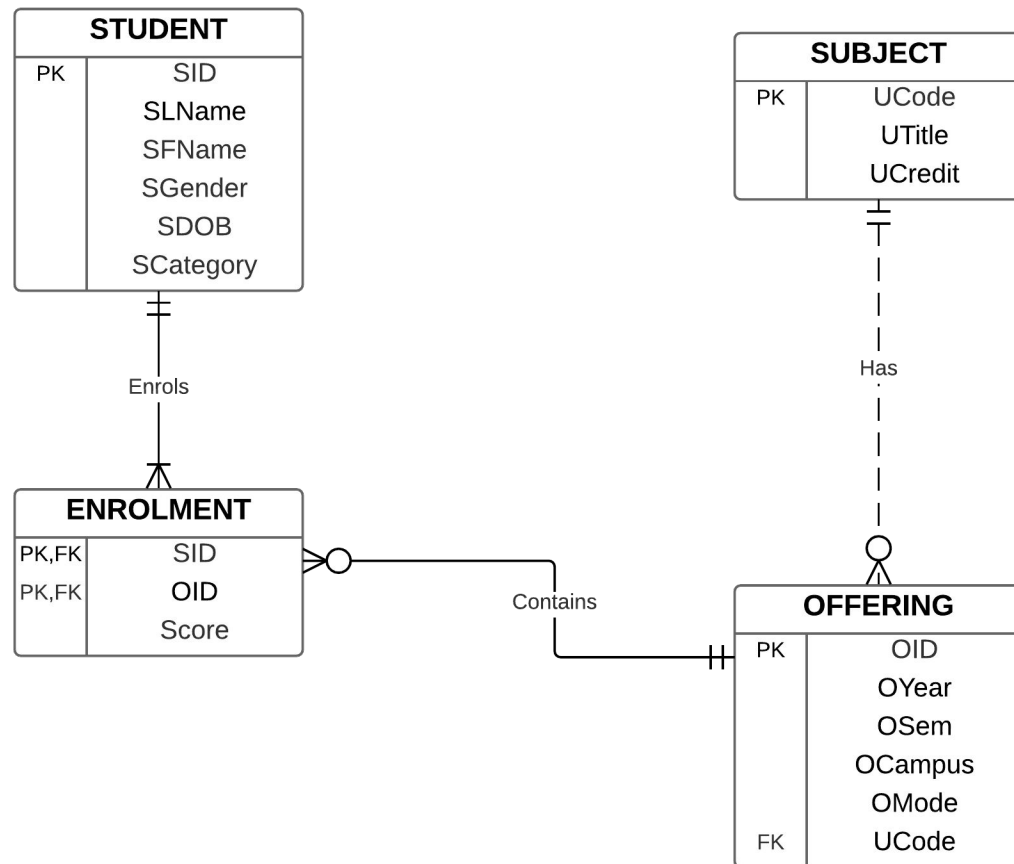
Operational Database - FIT9132 focus
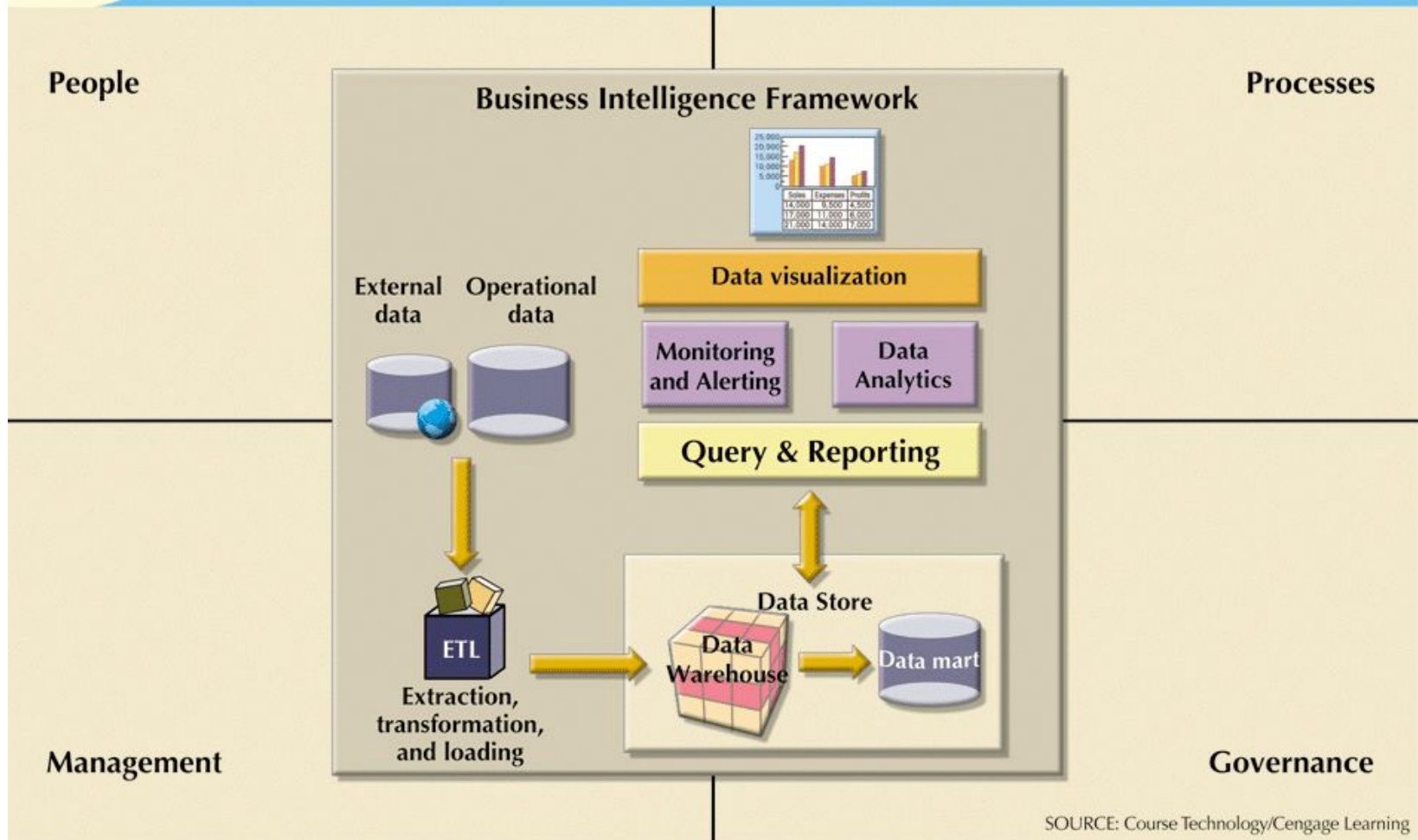
Usage of database

- 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 FIT9132?

- 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.
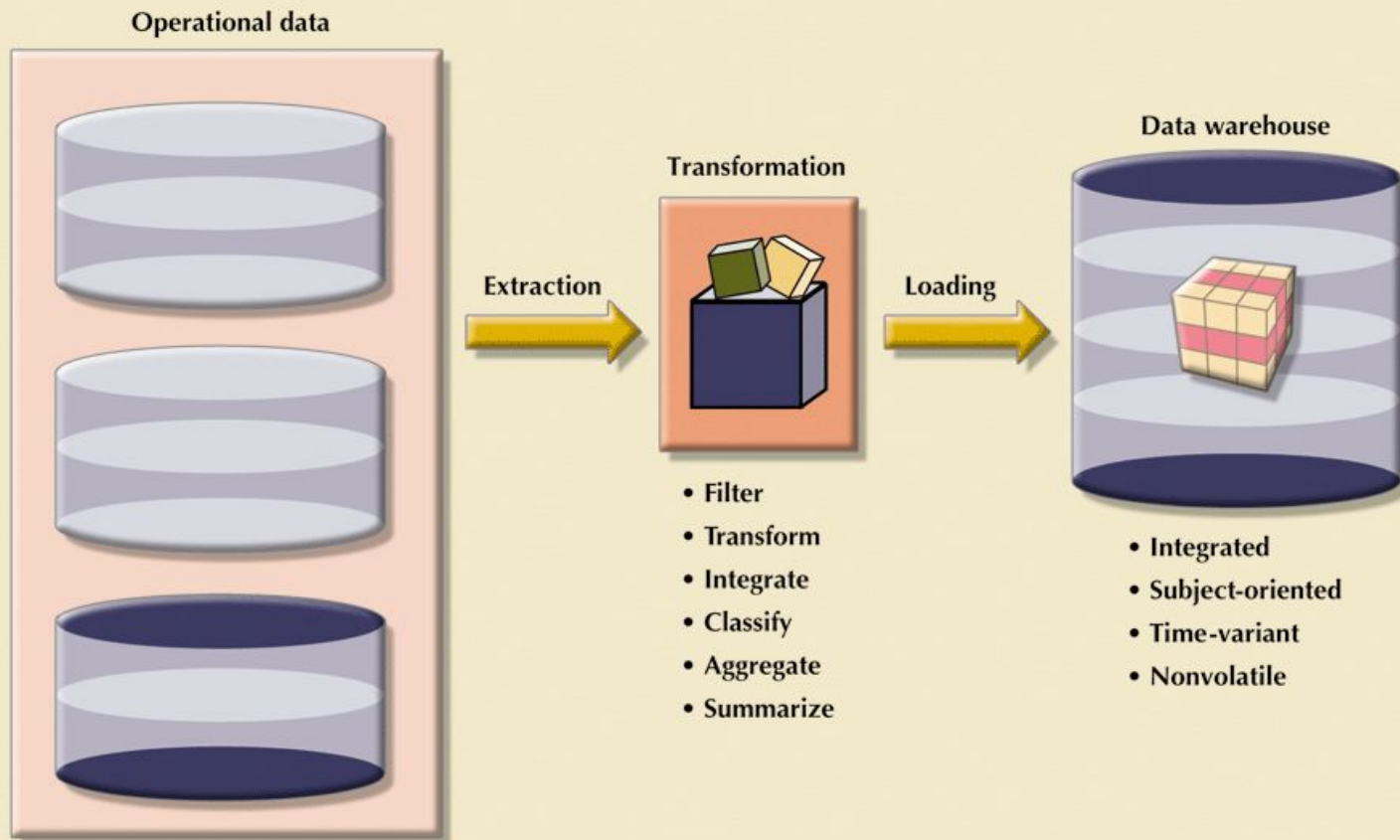
FIGURE 13.1 Business intelligence framework

People

Processes

Business Intelligence Framework

Data visualization

Monitoring and Alerting

Data Analytics

Query & Reporting

External data

Operational data

ETL

Extraction, transformation, and loading

Data Store

Data Warehouse

Data mart

Management

Governance

SOURCE: Course Technology/Cengage Learning

FIGURE 13.4 — The ETL process

Operational data → Extraction → Transformation → Loading → Data warehouse

Transformation:
- Filter
- Transform
- Integrate
- Classify
- Aggregate
- Summarize

Data warehouse:
- Integrated
- Subject-oriented
- Time-variant
- Nonvolatile

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

[FIT5195](FIT5195) - Business intelligence and data warehousing

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

1. describe the scope and application of data warehousing

2. design data warehousing systems

3. design multidimensional data models using star schemas

4. implement data warehousing in relational databases

5. use OLAP in SQL

6. explain the need for data warehousing architecture

7. explain query optimisation and its impact on multi-dimensional design

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

FIT5137 - Advanced Database Technology

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.

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

FIT5202 - Data processing for big data

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

1. identify and explain big data concepts and technologies;

2. write and interpret parallel database processing algorithms and methods;

3. apply common data analytics and machine learning algorithms in a big data envir...

4. use and evaluate streaming methods in big data processing;

5. use big data streaming technologies.

Data Processing Ecosystem

http://www.clearpeaks.com/blog/big-data/big-data-ecosystem-spark-and-tableau
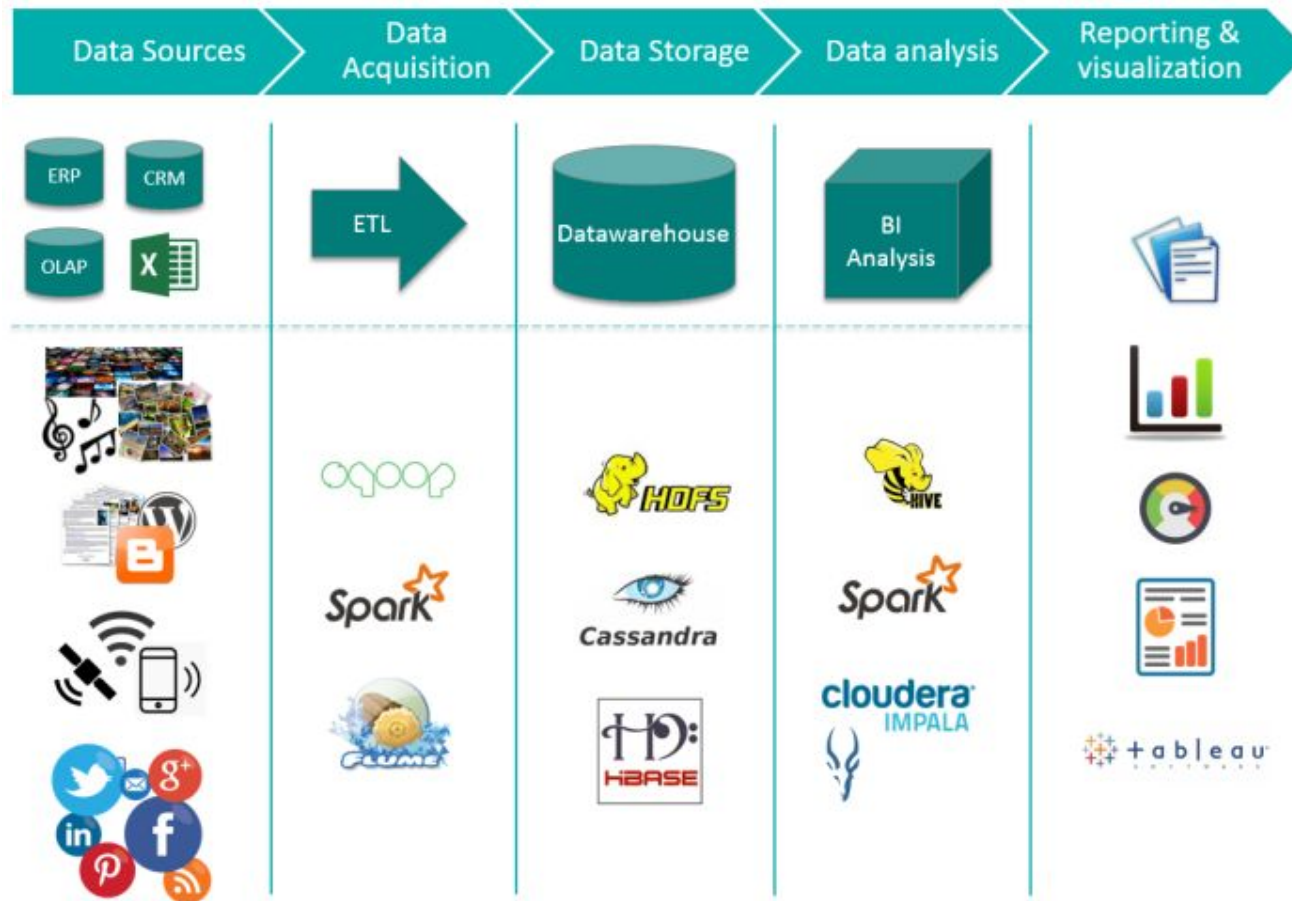
"Horses for Courses"

- 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.

FIT9132 Scheduled Final Assessment

2021 Final Assessment

- **2 HOUR 10 minutes**

  - Time includes reading time

  - **eExam, closed Book, online supervision, no aids (eg. dictionary)**

  - **2 working pages, 1 answer sheet (all three taken in as blank pages)**

  - **Rulers, pens, pencils allowed**

  - Available via the eExam platform

    - https://www.monash.edu/exams/electronic-exams

  - The modelling question makes use of a hybrid question:

    - write answer on paper (***indicate* answered on e-exam**)

      - *submit one page* with complete diagram (one answer sheet)

    - photograph with phone **MUST NOT write after exam time expired**

    - upload via QR code (<u>after paper has been completed</u>) - important that you practice this process

      - Complete the e-exam General Knowledge 2021 practice

2021 Final Assessment Structure

- 100 marks 50% of your final mark in FIT9132.
  - Minimum to pass FIT9132 overall:
    - 45% in-semester, 45% final assessment and 50% overall
  - *Assignment 2B marking will not be finalised before you sit your Final Assessment*
- 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 day, **do not bring** the case study to the Final Assessment or discuss it in the Ed forums
  - No tables/data for the case study will be available on our Oracle server
- Final Assessment when - your responsibility to determine

2021 Final Assessment

- All content specified during your semester of study in FIT9132 is examinable, including but not limited to:

  - Pre-reading (weekly Coronel & Morris chapters)

  - Workshop/Workshop Q&A Slides and Videos

  - Tutorial Notes, and

  - all other Moodle Materials (except where explicitly stated NOT EXAMINABLE).

2021 Mock Final Assessment

- Serves to provide an overview of the general structure of the final assessment paper only.

- Available Thursday, 21st October 2021 6 PM
    - Sample solution available Wednesday, 27th October 9AM (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 - "Scheduled Final Assessment" block/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

MONASH
University

Week 2 and 5 – Data Modelling

- 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

- 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

- 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

- 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, 9 and 10 – SQL

- 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

Week 8 – Transaction Management

- 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 11 – noSQL

- 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/