School of Computer and Mathematical Sciences

Master of Computer and Information Sciences

Data Warehousing and Big Data

COMP810

Semester 02, 2015

1 COURSE INFORMATION

Course Title:	Data Warehousing and Big Data
Course Code:	COMP810
Prerequisites:	None
Co-requisites:	None
Level:	8
Points:	15

1.1 Course prescription

Discusses the main concepts and solutions for Data Warehousing and managing Big Data The key concepts underpinning the logical design, physical design and implementation of data warehouses are appraised. Data collection, data extraction, cleansing, transformation and loading methods are considered along with query optimization techniques. Data warehousing architectures, and applications and solutions for managing Big Data are reviewed.

1.2 Content

Content Outline

Week 1: Introduction to course, Introduction to Data Warehouse

Week 2: Lecture on Course Assessments

Week 3: DW Life Cycle and Basic Architecture

Week 4: DW Architecture (Continued)

Week 5: Logical Model

Week 6: Physical Model

Week 7: Indexes

Week 8: Optimizations

Week 9: OLAP Operations and Queries

Week 10: Building the DW

Week 11: Real-Time DW

Week 12: Big Data

Note: Structured Query Language (SQL) and DW project will be covered in Labs.

1.3 Learning Outcomes

On successful completion of this paper students will be able to:

1. Gain an appreciation of the role that Data Warehouses (DW) and Big Data play in enhancing the decision making process.
2. Gain an understanding of the fundamental concepts of the Star and the Snowflake Schema; learn how to design the schema of a DW based on these two models.
3. Understand the architecture of DW Systems and be able to specify the advantages and potential problem areas.
4. Use Analytic SQL to aggregate, analyse and report, and model data
5. Conduct a critical appraisal of concepts and applications of Big Data management technologies.
6. Appraise ethical issues involved in data collection and processing.

1.4 Learning and Teaching Strategies

- *Lectures* – The course is structured to give students a conceptual framework for DW design and implementation that can be applied to business systems in general. There will be teacher directed lectures. Participatory teaching methods will also be emphasized. Many concepts will be developed through problem-based learning, discussion and analysis.

- *Tutorials/laboratories* – To enable the students to become familiar with the framework they will be required to apply it in exercises on several small and medium size scenarios. Group work will be used to encourage interpersonal skills and group management.

- *Review and research* – Students will review and research key concepts introduced by lecturers, to broaden base knowledge. Students will be expected to carry out supporting reading.

- *AUTOnline* – AUTOnline will be used for posting assessments, course materials, making important announcements, sending e-mail notices to the class, and for other purposes that may arise. Students are required to use AUTOnline for this course and are responsible for checking it regularly, preferably at least once per day. During the first week of the semester, each student is responsible for checking his or her account details in AUTOnline and anyone who cannot access the course must notify the lecturer immediately.

1.5 Class Attendance

All classes should be attended to ensure your success in this course. The Programme Leader may withdraw a student's enrolment in a course if the student is absent for a consecutive period that constitutes more than 20% of the scheduled hours for that course**.**

1.6 Duration & Student Workload

Item	Weeks	Hours
On campus learning Teaching and lab practical	12 weeks * 2 hr per week	24
Students direct learning Completing of exercises and project	12 weeks * 10.5 hrs	126
Total		150
Independent Study:	7-8 hours weekly through-out the semester; In total 150 hours of student study time is expected.	

1.7 Teaching Team

i) Dr Muhammad Asif Naeem (leader & lecturer)

Email: mnaeem@aut.ac.nz

ii) Ali Haider Hussein Ghazala (TA)

Email : ahaiderh@aut.ac.nz

1.8 Class Times and Places

The timetable for this course can be obtained from the Arion website (https://arion.aut.ac.nz). You can access your personal timetable using your Arion student login name and password.

1.9 Communication via Email

It is our policy that all email communication with students is conducted using their AUT email addresses. Please ensure that your AUT email address is regularly checked as important information might be sent to you during the semester. It is the student's responsibility to check his or her email box and/or arrange for forwarding AUT email correspondence to another [private] address.

2 OVERVIEW OF THE SEMESTER

The course calendar outlines the topics for each week. The calendar as presented is subject to change. Notification will be given if this occurs.

2.1 Course Calendar

WEEK	WEEK BEGINN	LECTURE	LAB
1	20th Jul	Introduction to course, Introduction to Data Warehouse	Writing Basic SQL statements;
2	27th Jul	Lecture on Course Assessments	SQL Operators – Restricting and Sorting Data
3	3rd Aug	DW Life Cycle and Basic Architecture	Retrieving Data from Multiple tables
4	10th Aug	DW Architecture (Continued)	SQL functions and Aggregating data using Group Functions
5	17th Aug	Logical Model	Table creation, alteration and applying constraints
6	24th Aug	Physical Model	Data Manipulation Language
Mid Semester Break			
7	14th Sep	Indexes	SQL practice
8	21th Sep	Optimizations	Creating of DW Schema
9	28th Sep	OLAP Operations and Queries	Working on Project
10	5th Oct	Building the DW	Working on Project
11	12th Oct	Real-Time DW	Working on Project
12	19th Oct	Big Data	No Lab

3 ASSESSMENT

3.1 Assessment Items Summary

Item	Assessment details	Weight %	Date	LO
Assessment 1	Research Report	20%	To be confirmed	1, 5, 6
Assessment 2	Project	80%	To be confirmed	1, 2, 3, 4, 5, 6
Grade Map	Grade Map 1: A+ A A- Pass with Distinction B+ B B- Pass with Merit C+ C C- Pass D Fail			

3.2 Pass Mark

To pass the paper, student needs to attempt both assessments and obtain a C- grade overall.

4 READINGS

4.1 Text book

Building the Data Warehouse *by William H. Inmon*

Wiley, ISBN 978-0-7645-9944-6

4.2 Additional Reference and Supplementary Reading

- **The Data Warehouse Toolkit** *by Ralph Kimball & Margy Ross*

  Wiley, ISBN 978-1-118-53080-1
- **Big Data** *by Nathan Marz, James Warren*

  Wiley, ISBN 978-1-617290-34-3

ORACLE Online documentation