# ICT394 Business Intelligence Application Development

## Topic 03: Data Warehousing

### Introduction

In your previous database studies, you most likely would have focused primarily on relational databases, perhaps with an introductory topic in data warehousing. This topic will provide some introduction to data warehousing as the basis for the next two topics, which will concentrate on design and implementation issues associated with data warehousing. This topic sets out to describe the differences between the operational databases with which you are familiar, and the analytical databases that are the focus of this unit.

### What you need to do

#### Pre-Workshop

1. Watch the Chrysler Data Quality Management Case Study, paying particular note to the following:
   1. Problems they were trying to address
   2. Issues with their current systems
   3. Problems with interpretation of current reports
   4. Main features of the solution they implemented
2. Read over your notes or textbook from your previous databases study to refresh your memory about data warehousing. If you do not have access to those resources, just about any databases text will have a section on data warehousing that you will be able to read.
   1. There is a link to Connolly & Begg (2015) in the Topic 03 Readings Link – see Chapter 31.
3. Watch the Topic 03 Lecture Videos/read the Lecture Slides

#### Workshop

1. Download the Tutorial 03 Worksheet and work through the exercises there
2. Complete the Computer Lab

#### After the Workshop

1. Complete the Moodle quiz for Topic 03
2. Prepare practice answers for the Sample Exam Questions for this topic

### Learning outcomes

This topic contributes to the following Unit Learning Outcomes:

1. Describe the common data sources that exist in organisations and their use in BI
2. Demonstrate practical skills in the processes associated with extraction, transformation and loading (ETL) of organisational data
3. Design and implement a simple data warehouse environment

At the completion of this topic, you should be able to:

* Provide a definition of a data warehouse, including examples as to how and in what circumstances it would be used
* Describe and provide examples as to the difference between operational and analytical databases
* Discuss the components of a data warehouse system
* Discuss the basic steps in the development of a data warehouse

### Key concepts

Data warehouse, operational vs analytical databases, subject-oriented, time-variant, integrated, data warehouse design as an iterative process

### Resources for this topic

To undertake study for this topic, you will need:

#### Essential reading:

#### Case Study:

2011, *Chrysler’s Data Quality Management Case Study*, <https://www.youtube.com/watch?v=N78lHpiCD0k>

#### Recommended reading

Read over your notes or textbook from your previous databases study to refresh your memory about data warehousing. If you do not have access to those resources, just about any databases text will have a section on data warehousing that you will be able to read.

There is a link in the Topic 03 Readings to Connolly & Begg (2015). Chapter 31 of that book provides a very good background to Data Warehousing.

### Computer Lab

This week’s lab will involve you working through some SQL relevant to the creation of BI queries.

### Sample Exam Questions for this Topic

1. Explain the following parts of the data warehouse definition:
   1. Structured repository
   2. Integrated
   3. Subject oriented
   4. Enterprise-wide
   5. Historical
   6. Time-variant
   7. Developed for the retrieval of analytical information
   8. May include data at the fine level of detail of summary data or both
2. Explain, using examples as appropriate, why a data warehouse would be created as a separate data store?
3. What are the major components of a data warehouse?
4. Briefly describe the process of developing data warehouse front-end applications.
5. Explain, using examples, why the data warehouse design process is an iterative process.
6. Explain how the SQL GROUP BY, ROLLUP and CUBE commands could be used for BI.