COMP 7027 ADVANCED DATA ENGINEERING

Credit Points 10

Description Data Engineers play a crucial role in designing data systems tailored to certain specifications. While several software modules can be used for constructing data systems, a Data Engineer must be able to identify the inner mechanisms underlying these systems, thereby gauging their suitability. In this subject, students will delve into the intricacies of modern-day data system design and algorithms, ensuring these systems' reliability, scalability, and ease of maintenance. Furthermore, we will explore common problems affecting data systems in the hardware and software space and identify their implications. In addition, students will investigate algorithms and data structures that optimise the storage and flow of data through data systems.

School Computer, Data & Math Sciences

Student Contribution Band HECS Band 2 10cp

Check your fees via the Fees (https://www.westernsydney.edu.au/currentstudents/current_students/fees/) page.

Level Postgraduate Coursework Level 7 subject

Pre-requisite(s) COMP 7026

Learning Outcomes

1. Compare the reliability, scalability and maintainability of data systems.
2. Given a set of specifications, design a suitable data system from existing modules.
3. Identify the challenges when using distributed computing.
4. Implement replication, partitioning/sharding and transactions on distributed systems.
5. Write scripts to construct derived databases from existing data.

Subject Content

Data Engineering Design

- Reliable, Scalable, and Maintainable Applications
- Data Models and Query Languages
- Storage and Retrieval Algorithms
- Encoding and Evolution

Distributed Data

- Data Replication
- Data Partitioning
- Transactions
- The Trouble with Distributed Systems
- Consistency and Consensus

Derived Data

- Integrating Data using Batch Processing
- Stream Processing
- The Future of Data Systems

Assessment

The following table summarises the standard assessment tasks for this subject. Please note this is a guide only. Assessment tasks are regularly updated, where there is a difference your Learning Guide takes precedence.

Type	Length	Percent	Threshold	Individual/ Group Task	Mandatory
Quiz	30 minutes (per Quiz)	20	N	Individual	Y
Case Study	3 weeks	30	N	Individual	Y
Final Exam	2 hours	50	N	Individual	Y