

Data Analytics MSc module details

Block 1: Data Analytics Infrastructure

The module introduces Business Intelligence/Analytics and Data Infrastructure, as a foundation for understanding and applying these concepts and methods in an organisation. The module is delivered in two parts:

- **Part 1: Foundations of Business Intelligence and Analytics in the Organisation**
- **Part 2: Data Warehouse design**

Part 1 introduces Business Intelligence systems, which other modules in the programme can draw upon when studying a more detailed BI system component or the development of such systems. Specifically, it introduces students to the Business Intelligence (BI) system concept and its application within organisations. Examples of some topics covered include: Defining and Framing Business Intelligence; the Architecture of Business Intelligence; Business Metrics, Types of Data, and the Analytics Continuum; Aligning the Data Strategy with the Business Strategy, etc.

Part 2 focuses on the design of data warehouses. It builds on the student's prior knowledge of Relational Databases and Relational Database Management Systems (DBMS) to consider the data requirements, underpinned by an appropriate technical infrastructure, for a data warehouse in response to a particular business situation. The role, functionality and architecture of a data warehouse facility will be discussed. Examples of some topics covered include: data warehouse overview, ER modelling concepts, dimensional modelling, design considerations for data warehouses; current limitations of research issues, and new development in data warehousing.

Block 2: Big Data Applications

This module will introduce students to state-of-the-art approaches to Big Data problems. It will utilise the Hadoop Distributed File System (HDFS) and Apache Spark to demonstrate data mining and machine learning algorithms for knowledge discovery and for presenting the newly acquired information in meaningful ways. It will also cover graph processing using Graph Frames for analysing data in graphical forms such as social networks and trees, Spark SQL, Spark Streaming for real time analytics and some natural language processing such as sentiment analysis. Parallel computing in the cloud will be a key aspect incorporated throughout. Business reporting tools for big data and timeseries analysis using indicative tools such as SAS Enterprise Guide will also be covered in this module.

Block 3: Data Mining Techniques and Methods

The aim of this module is to review the data mining methods and techniques available for uncovering important information from large data sets and to know when and how to use a particular technique effectively. The module will enable the student to develop an in-depth knowledge of applying data mining methods and techniques and interpreting the statistical results in relevant problem domains.

Current application areas and research topics in data mining will also be discussed and students will be expected to contribute to these discussions to increase their background knowledge and understanding of issues and developments associated with data mining.

The module uses the data mining tool SAS Enterprise Miner and BASE SAS programming language, open source alternatives are available such as Weka and R.

Block 4: Project proposal, Planning and project management (PPPM)

This module provides grounding in statistical and research methods required at MSc level, looking at both quantitative and qualitative approaches including laboratory evaluation, surveys, case studies and action research. The Statistical content of this module will support the research process, sampling methods, sample size, bias, central limit theorem, probability, distributions families, confidence levels and will include parametric and non parametric analysis. The students will be equipped with skills to carry out their own research analysis, and have the statistical literacy to the critical evaluation of the results, analysis of published work. There will be a strong practical element to this component of the course the student will get a grounding in using a statistical application, such as SAS studio.

Blocks 5 & 6: PGT Project

The aim of the project/dissertation is to provide students with the opportunity to carry out a self-managed in-depth study involving design, fact finding, analysis, synthesis and integration of complex ideas which are sometimes based on incomplete and contradictory data or requirements. The project is likely to demonstrate the application of skills acquired from the taught course to the solution to a particular problem or research topic. Normally the project is a self-contained piece of work of considerably greater depth than can be accommodated within a taught module and may reflect and build on the entire breadth of material studied by the student.