

Module Synopses

Semester One

1. Probability & Statistics

Covers basic probability and statistical concepts and principles that lay the foundation for students to understand and learn the statistical procedures and methods in the subsequent modules. Topics include the following: data exploration and summary, elementary probability theory, functions of random variables, discrete and continuous probability distributions, sampling and sampling distributions, point and interval estimation and hypothesis testing.

2. Data Mining

Introduces students to the key steps in data mining, including data exploration, data preparation, and building both unsupervised and predictive models. One focus is on developing and evaluating classification models. Modelling techniques covered include k-nearest neighbours, linear regression, logistic regression, classification and regression trees, neural networks, association rules, and k-means clustering.

Semester Two

3. Applied Statistics

Covers multivariate statistical methods commonly used in the analysis of data in industry, and introduces concepts of experimental design. Topics include the following: matrix algebra, ANOVA, principal components analysis, factor analysis, discriminant analysis and cluster analysis.

4. Regression Analysis

Covers the relevant theory and applications of selected data mining techniques and statistical analysis. The focus will be on methods that deal with data modelling with an equal emphasis on statistical rigour and application. The key topics that will be discussed are Linear Regression Models, Indicator Variable Regression, Regression of Heteroscedastic Data, Logistic Regression, and Time Series Analysis.