# Data science DATS

Applying mathematics, statistics, data mining and predictive modelling techniques to gain insights, predict behaviours and generate value from data.

|  |
| --- |
| **Guidance Notes:**  Data science is typically used for analysing high volume, high velocity and high variety data (numbers, symbols, text, sound and image).  Activities may include, but are not limited to:   * integrating methods from mathematics, statistics and probability modelling using specialised programming languages, tools and techniques * sourcing and preparing data for analysis * identifying, validating and exploiting internal and external data sets generated from a diverse range of processes * developing forward-looking, predictive, real-time, model-based insights to create value and drive effective decision-making * finding, selecting, acquiring and ingesting data sources * integrating and cleaning data to make it fit for purpose * developing hypotheses and exploring data using models and analytics sandboxes * refining requirements, validating, training and evolving models over time to discover deeper insights, make predictions, or generate recommendations * using advanced analytic techniques including, but not limited to, data/text mining, machine learning, pattern matching, forecasting, visualisation, semantic analysis, sentiment analysis, network and cluster analysis, multivariate statistics, graph analysis, simulation, complex event processing, neural networks. |

## Level 2

Under guidance, applies given data science techniques to data.  
Analyses and reports findings and remediates simple issues, using algorithms implemented in standard software frameworks and tools.

## Level 3

Applies existing data science techniques to new problems and datasets using specialised programming techniques.  
Selects from existing data sources and prepares data to be used by data science models.  
Evaluates the outcomes and performance of data science models. Identifies and implements opportunities to train and improve models and the data they use.   
Publishes and reports on model outputs to meet customer needs and conforming to agreed standards.

## Level 4

Investigates the described problem and dataset to assess the usefulness of data science and analytics solutions.   
Applies a range of data science techniques and uses specialised programming languages. Understands and applies rules and guidelines specific to the industry, and anticipates risks and other implications of modelling.  
Selects, acquires and integrates data for analysis. Develops data hypotheses and methods and evaluates analytics models. Advises on the effectiveness of specific techniques based on project findings and comprehensive research.   
Contributes to the development, evaluation, monitoring and deployment of data science solutions.

## Level 5

Plans and drives all stages of the development of data science and analytics solutions.   
Provides expert advice to evaluate the problems to be solved and the need for data science solutions. Identifies what data sources to use or acquire.   
Specifies and applies appropriate data science techniques and specialised programming languages.  
Reviews the benefits and value of data science techniques and tools and recommends improvements. Contributes to developing policy, standards and guidelines for developing, evaluating, monitoring and deploying data science solutions.

## Level 6

Leads the introduction and use of data science to drive innovation and business value.   
Develops organisational policies, standards, and guidelines for data science.  
Sets direction and leads in the introduction and use of data science techniques, methodologies and tools. Leads the development of organisational capabilities for data science.   
Plans and leads strategic, large and complex data science initiatives to generate insights, create value and drive decision-making.