

Unit	Principles of Data Management Infrastructure SQA Unit Code – HD54 04	
Code	DAPDMI	<i>QCF:</i>
Level	SCQF: 8	<i>QCF: 4/5</i>
Credit Value	SCQF:18	<i>QCF: tbc</i>
Unit summary	<p>This unit is concerned with management of large, highly dynamic data sets, often referred to as 'Big Data'. These data sets typically include semi- and un-structured data in addition to structured data. The data will be derived from a range of sources and can relate to contexts including business and scientific or social research.</p> <p>The unit covers hardware, software and associated procedures.</p>	
Learning Outcomes The learner will:	Assessment Criteria	
1. Understand the types of hardware and software used to store and manage large data sets	<p>1.1 Evaluate the importance of scalability of capacity and performance to data storage and management</p> <p>1.2 Describe the use of computer clusters and distributed computing in data storage and management</p> <p>1.3 Describe different dynamic data storage models and their implementation including Network-Attached Storage (NAS), Object-based Storage and Hyperscale Storage</p> <p>1.4 Describe different types of software tools for manipulating large data sets including the contemporary frameworks, programming languages and NoSQL databases</p>	
2. Understand data management techniques for large data sets	<p>2.1 Describe how data can become inaccurate, incomplete or erroneous</p> <p>2.2 Evaluate the impact of dirty data on business insight and decision making</p> <p>2.3 Describe how data cleansing can be used to improve the quality of data</p> <p>2.4 Explain how data reduction can be used to increase storage efficiency and the usefulness of data</p> <p>2.5 Explain the capabilities provided by differing data</p>	

	<p>management software systems for:</p> <ul style="list-style-type: none"> • manipulating data • linking data sets • data migration • data abstraction
<p>3. Understand the organisational policies needed for the storage and management of large data sets</p>	<p>3.1 Describe the types of legislation and regulations which impact on data storage and management</p> <p>3.2 Outline the content required for organisational policies and procedures for:</p> <ul style="list-style-type: none"> • management, storage and deletion of data • data sharing and re-use • quality assurance • archiving and backup of data • data protection, ownership and access • compliance with legislation and regulations <p>3.3 Explain the importance of organisational policies and procedures being up-to-date and widely publicised.</p>
Additional information about the unit	
Guidance on approaches to assessment	<p><i>This unit may be assessed by any means which provides evidence that the candidate understands the content. Every effort should be made to relate the content to the candidate's organisation wherever possible.</i></p> <p><i>The assessment criteria are intended to be system neutral and the terminology used may be interpreted to meet the context of the particular data management and analysis system in use.</i></p>
Details of the relationship between the unit and relevant National Occupational Standards or other professional standards	This unit is based on the NOS for Data analytics.
Location of the unit within the subject/sector classification system	IT Professional
Name of the organisation submitting the unit	The Tech Partnership