

**Unit 3 Data Analytics 2025**
**Outcome 1 Data analytics – Template for developing an assessment task – Plan**

<b>Outcome 1</b> On completion of this unit the student should be able to interpret teacher-provided solution requirements and designs, extract data from large repositories, manipulate and cleanse data, conduct statistical analysis and develop data visualisations to display findings.		<b>Assessment task development</b> Create a scenario that is a real-world example that provides students with solution requirements and designs that will enable them to extract authentic data from large repositories using database software tools, manipulate and cleanse the data using spreadsheet software tools, conduct statistical analysis to identify trends, relationships and patterns using spreadsheet software tools and present findings using data visualisations. Key content within the tasks should be based on the targeted key knowledge and key skills. The total number of marks for the outcome is to be out of 100.
<b>Key knowledge</b>	<b>Key skills</b>	
<ul style="list-style-type: none"> <li>• emerging trends in data analytics using artificial intelligence, including:                             <ul style="list-style-type: none"> <li>– integration of artificial intelligence features into software tools</li> <li>– generating data visualisations through the writing and refinement of prompts</li> <li>– machine learning and statistical modelling for making predictions, decisions and recommendations</li> </ul> </li> <li>• characteristics of functional and non-functional requirements, constraints and scope</li> <li>• design tools for representing databases and spreadsheets, including:                             <ul style="list-style-type: none"> <li>– data dictionaries</li> <li>– query designs</li> <li>– layout diagrams</li> <li>– input-process-output (IPO) charts</li> </ul> </li> <li>• design tools for representing data visualisations, including:                             <ul style="list-style-type: none"> <li>– mock-ups</li> <li>– storyboards</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• interpret solution requirements and designs</li> </ul>	<p>Content to be included in the assessment task should introduce students to a scenario. The scenario should indicate the data repositories that students are to use. The scenario should clearly state the solution requirements and designs for the database, spreadsheet and data visualisations solutions and provide students with sufficient opportunities to demonstrate their knowledge and to meet the requirements of the outcome. A range of appropriate design tools are to be used. Students are not to complete designs themselves. Design tools should be appropriate for the software tool used.</p>
<ul style="list-style-type: none"> <li>• characteristics of data types, including:                             <ul style="list-style-type: none"> <li>– text (character, string)</li> <li>– numeric (integer, floating point, date/time)</li> <li>– Boolean</li> </ul> </li> <li>• techniques for identifying, selecting, extracting and validating authentic data stored in large repositories, including:                             <ul style="list-style-type: none"> <li>– downloading datasets in a range of formats</li> <li>– the use of SQL functions to retrieve, filter, sort and link dataset values (SELECT, FROM, WHERE, ORDER BY, INNER JOIN)</li> <li>– the use of Boolean operators (AND, NOT, OR) for WHERE statements</li> <li>– existence checking, type checking and range checking</li> </ul> </li> <li>• structural characteristics of relational database management systems (RDBMS), including:</li> </ul>	<ul style="list-style-type: none"> <li>• identify, select, extract and validate relevant data from large repositories using database software</li> </ul>	<p>Students are to identify, select, extract and validate the relevant data from the large repositories listed in the scenario using database software tools. Data is to be queried using databases and SQL.</p>

## Unit 3 Data Analytics 2025

## Outcome 1 Data analytics – Template for developing an assessment task – Plan

<ul style="list-style-type: none"> <li>– data types and field sizes</li> <li>– data in tables</li> <li>– relationships using primary and foreign key fields</li> <li>– use of SQL to generate queries</li> </ul>		
<ul style="list-style-type: none"> <li>• methods for referencing data sources using the American Psychological Association (APA) referencing system</li> </ul>	<ul style="list-style-type: none"> <li>• use the APA referencing system to acknowledge intellectual property</li> </ul>	Students are to use the APA referencing system to acknowledge intellectual property used within the data visualisations.
<ul style="list-style-type: none"> <li>• techniques for effectively and efficiently manipulating and cleansing data, including: <ul style="list-style-type: none"> <li>– formulas and functions to perform calculations</li> <li>– sorting, filtering and reformatting</li> <li>– identifying and fixing errors</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• manipulate and cleanse data using spreadsheet software</li> </ul>	The scenario with the solution requirements and designs should enable students to determine the appropriate selection and use of features for the spreadsheet software tool to enable them to manipulate and cleanse data.
<ul style="list-style-type: none"> <li>• techniques to statistically analyse data to identify trends, relationships and patterns, including: <ul style="list-style-type: none"> <li>– descriptive statistics (average, median, minimum, maximum, range, standard deviation, count/frequency, sum)</li> <li>– Pearson’s correlation co-efficient (r)</li> <li>– the shape and skew of data</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• conduct statistical analysis to identify trends, relationships and patterns</li> </ul>	Students are to conduct statistical analysis on the data they have selected using spreadsheet software tools to identify trends and patterns.
<ul style="list-style-type: none"> <li>• purposes of data visualisations, including: <ul style="list-style-type: none"> <li>– exploratory data analysis</li> <li>– presentation of information</li> <li>– providing interactive experiences for users to explore data</li> </ul> </li> <li>• types of data visualisations, including: <ul style="list-style-type: none"> <li>– infographics (series or long-form, static)</li> <li>– dashboards (interactive, static or live data)</li> <li>– dynamic data visualisations (interactive, live data)</li> </ul> </li> <li>• formats and conventions applied to data visualisations to improve their effectiveness for intended users, including: <ul style="list-style-type: none"> <li>– use of colours, fonts, images and icons</li> <li>– visual hierarchy and clarity of message</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• select, justify and apply functions, formats and conventions to create effective data visualisations</li> </ul>	The scenario with the solution requirements and designs should enable students to determine the use of appropriate functions, formats and conventions to develop their data visualisations solutions to display findings. Students are to use a data visualisation tool. They are to justify and explain their selection of functions, formats and conventions used to develop their data visualisations. This written justification and explanation could be included as a separate written report.
<ul style="list-style-type: none"> <li>• techniques for testing databases and spreadsheets, including: <ul style="list-style-type: none"> <li>– testing formula and query results</li> <li>– testing validation</li> <li>– test cases comparing expected and actual results in testing tables</li> </ul> </li> <li>• techniques for testing data visualisations, including: <ul style="list-style-type: none"> <li>– visual inspection of the appearance of the data visualisation</li> <li>– confirming that charts and graphs are representative of the data being visualised</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• develop and apply suitable testing techniques to software tools used</li> </ul>	A testing table is to be developed that involves the testing of all validation and processing, such as queries, calculations, etc. The testing table should include columns for expected and actual output and show evidence of tests that work and don’t work.