



**Gosford High School**  
**Year 9 Computing Technology**  
**Assessment Task 2025**

<b>Number and Title of Task</b>	Assessment Task 3: In-class Project (20%)
<b>Topic</b>	Module 2: Data Science
<b>Weighting/Marks</b>	20% (100 marks)
<b>Due Date</b>	Term 3 Week 5, Friday 22 <sup>nd</sup> August 2025, 09:00am

**Outcomes to be Assessed (Syllabus outcomes being assessed by task)**

**A student:**

**CT5-DPM-01** applies iterative processes to define problems and plan, design, develop and evaluate computing solutions

**CT5-EVL-01** understands how innovation, enterprise and automation have inspired the evolution of computing technology

**CT5-DAT-01** explains how data is stored, transmitted and secured in digital systems and how information is communicated in a range of contexts

**CT5-COM-01** communicates ideas, processes and solutions using appropriate media

**CT5-THI-01** applies computational, design and systems thinking to the development of computing solutions

**CT5-DAT-02** acquires, represents, analyses and visualises simple and structured data

**Introduction**

In this project, you'll navigate the complete data analysis process, from defining the problem to presenting your findings. This project offers invaluable practical experience with real-world data and various data analysis tools. By the end, you'll have developed a robust skill set in data analysis and advanced Python.

**Data Analysis Project (100 marks)**

This project will take you through the entire process of conducting a data analysis, from defining the problem to presenting your findings. You'll gain practical experience in handling real-world data and using various tools and technologies. Here are your instructions for each phase of the project:

**Design Brief:** Identify a complex societal issue (e.g. housing, natural disasters) or a need within the school community. Form a hypothesis related to the issue. Design a system that processes and visualises data to test your hypothesis.

**Phase 1: Identifying and Defining**

Consider a wicked problem in society (e.g. housing crisis, bushfires, floods), or alternatively, a need or opportunity in the school community.

- **Mind Map:** Draw a three-tiered mind map of data you could collect to investigate either from a publicly available dataset, or your own primary research (if looking at school community). On the 2<sup>nd</sup> tier put the wicked problems, on the 3<sup>rd</sup> tier get more specific on each problem.
- **Define Your Purpose:** Before finding a dataset or collecting data, develop a hypothesis on a specific area. For example:
  - *"Suburbs with more public transport options have lower car ownership rates."*
  - *"Rainfall in NSW has become more extreme over the past 10 years."*
  - *"Not enough houses are being built in Australia to allow first-home buyers into the market"*
  - *"Bushfires are getting more frequent and more intense in NSW as global temperatures rise"*
  - *"Supermarkets in Australia are taking advantage of inflation to overcharge on necessary goods"*
  - *"Students who get involved in extracurricular activities are happier at Gosford High School"*
  - *"Students who are active at lunchtime have an improved sense of wellbeing in Period 5"*
- **Requirements Outline:** List down the functional (what your project should do) and non-functional (qualities like speed or security) requirements for your project. Be specific.

## Phase 2: Researching and Planning

- **Research Your Chosen Issue:** What's already available on the topic? Find news articles, journal articles, or interview students / teachers (if local issue) to get a sense of what people think of the topic.
- **Discuss the findings:** Discuss the above information in at least one SEE-I paragraph.
- **Find Your Data:** Acquire any data you need, whether it be through <https://www.kaggle.com/> (Secondary Research), Google Forms (Primary Research) or another source.
- **Planning:** Create a data dictionary to show the type of data and parameters required for your dataset.

## Phase 3: Producing and Implementing

- **Python, Pandas and Matplotlib:** Use Python Pandas to read and manipulate data. Create relevant charts and visualisations using Matplotlib.
- **User Interface:** How does the user find and generate the data to assess your findings? In a separate Python program, create an intuitive text-based or graphical user interface (GUI) which allows for an accessible User Experience (UX). Call the program 'main.py' for the user interface, then move any functions to another file to import (e.g. 'bushfire\_data.py').
  - **Note:** When you move over to the .py files, if you need to output a dataframe you **need to remember to use the print function**, for example:
  - `print(big_mac_df)`

## Phase 4: Testing and Evaluating

- **Analyse and Conclude:** Analyse your findings using **at least one SEE-I paragraph**. If you spot areas needing more research, make a note of them in your findings. Make sure to provide a **conclusion** on your hypothesis.
- **Test Your Analysis:** Make sure your analysis works correctly. Check and double-check your algorithms to ensure they provide accurate results.
- **Peer Verification:** Exchange your work with a classmate. Verify each other's datasets, calculations, and outputs. Provide feedback – PMI tables are helpful here (Plus, Minus, Implication).
- **Evaluate Your Project:**
  - Evaluate your system and results in relation to your **Requirements Outline**.
  - Evaluate your system in relation to **peer feedback**.
  - Evaluate your project in relation to **project management**.
  - Evaluate your system in relation to its **data and security**. Is the data valid, accurate and timely? Is it unbiased? Do we need to improve its security – if so, how? Could the UX be more accessible – how?

## Project Deliverables:

**1. Markdown (.md) File:** You should have a well-documented Markdown file that showcases your entire project, including all planning, research, production, implementation, testing and evaluating.

**2. Dataset:** Provide any datasets which are relevant to your chosen data scenario.

**3. Visualisations:** Include data visualisations and plots that support your analysis.

**4. User Interface:** The user must be able to interact with the program in some capacity, either through a graphical user interface (GUI), ASCII user interface or text-based user interface.

**4. Evaluation:** Write a detailed report evaluating your project against the criteria mentioned earlier.

**5. Peer Reviews:** Don't forget to provide feedback on your classmates' projects.

### Additional Information

You will be provided with all required resources including tutorials and study notes in class.

If you require additional time to study to complete your project's documentation, you may need to do some of this for homework. *No Recess or Lunch Times will be provided in CR1 or CR2.*

Each class will be given a specific date and time to complete the in-class quiz. Questions will be varied from class to class to ensure fairness.

***If you are unsure of any element of the task, it is YOUR responsibility to seek advice and support in a timely manner. Do NOT wait until the week the task is due to seek clarification!***

### Subject Specific Terminology

**Cyber Security:** The protection of information technology elements, including hardware and software, data or network services.

**Data:** A discrete representation of information using number codes. Data may include characters (eg alphabetic letters, numbers and symbols), images, sounds and/or instructions that, when represented by number codes, can be manipulated, stored and communicated by digital systems. For example, characters may be represented using ASCII code or images may be represented by a bitmap of numbers representing each 'dot' or pixel.

**Database:** A collection of data organised by records and fields that can be easily stored, accessed, managed and updated.

**Dataset:** A collection of data combined for a specific purpose.

**Information:** The presentation of data in a manner that is readily understood.

### Feedback (How you will receive feedback and reflect on the task)

#### Teacher feedback will include:

- Verbal formative feedback on project development
- Individual written summative feedback on in-class project
- Verbal summative feedback as a class group

#### Students will reflect on their task:

- Completing detailed documentation of their processes
- Through peer evaluation of project success

- Evaluating their project's success, performance, and necessary improvements

### **Student Acknowledgement of Academic Integrity**

**By submitting the task for marking, I acknowledge the following:**

1. The work submitted is my own work and appropriately acknowledges of all sources has been made.
2. I have not used generative AI in the construction of the task.
3. If there is an allegation of malpractice, I will be required to show my drafts/ planning to prove the task is all my own work.
4. I am aware that the work may be submitted to plagiarism detection processes.
5. Where the work of others is used and not acknowledged, a finding of plagiarism will be made and a mark of zero awarded and I will have to resubmit the task.

## Data Analysis Project

### Phase 1: Identifying and Defining (5 marks)

/5	5 Marks (Per row)	4 Marks (Per row)	3 Marks (Per row)	2 Marks (Per row)	1-0 Marks (Per row)
Problem Definition & Hypothesis CT5-DPM-01	Exceptionally well-defined problem with precise objectives and a clear focus, including highly detailed and divergent brainstorming. Extremely clear hypothesis.	Well-defined problem with clear objectives, including detailed brainstorming. Detailed hypothesis.	Expected level of definition of the problem that is reasonably clear, including some brainstorming. Expected level of hypothesis.	Basic definition of the problem, may not include brainstorming. Basic hypothesis. Both may be a little unclear or too general.	Limited definition of the problem or brainstorming included. Limited or no hypothesis provided.

### Phase 2: Researching and Planning (15 marks)

/10	10 Marks (Per row)	8 Marks (Per row)	6 Marks (Per row)	4 Marks (Per row)	2-0 Marks (Per row)
Research CT5-DAT-02	Comprehensive research with extensive and thoughtful discussion on available information.	Thorough research with detailed discussion on available information.	Expected research with reasonable explanation of available information.	Basic research with some description of available information.	Limited evidence of research provided.

/5	5 Marks (Per row)	4 Marks (Per row)	3 Marks (Per row)	2 Marks (Per row)	1-0 Marks (Per row)
Planning CT5-DPM-01	Extraordinarily detailed data dictionary with no errors and all fields explained appropriately.	Well-constructed data dictionary with most fields explained appropriately.	Expected level of data dictionary provided which may contain a few inaccuracies.	Basic or unclear data dictionary which may contain many inaccuracies.	Limited or incomplete data dictionary which contains significant inaccuracies and errors.

### Phase 3: Producing and Implementing (60 marks)

/60	10 Marks (Per row)	8 Marks (Per row)	6 Marks (Per row)	4 Marks (Per row)	2-0 Marks (Per row)
Fundamental Programming Concepts	Exceptional use of fundamental programming concepts such as loops, functions, lists and dictionaries to create an efficient system.	Proficient use of fundamental programming concepts such as loops, functions, lists and/or dictionaries to create a reasonably efficient system.	Expected level of use of fundamental programming concepts, but system may lack efficiency in some areas.	Progressing use of some fundamental programming concepts, but system highly inefficient and may contain logical errors.	Limited or no use of fundamental programming concepts and may not function due to syntax errors.
Pandas Data Processing	Outstanding use of Pandas for filtering and processing data from external files, with no anomalies or data redundancy.	Proficient use of Pandas for filtering and processing data from external files with few anomalies and/or little data redundancy.	Expected level of use of Pandas for filtering and processing data from external files, with some anomalies and/or data redundancy.	Basic use of Pandas for data processing which may contain errors, anomalies and/or data redundancy.	Limited or no use of Pandas for data processing which may contain significant errors.
Matplotlib Visualisation	Outstanding use of Matplotlib to create visually compelling and informative visualisations which are formatted appropriately.	Proficient use of Matplotlib to create effective visualisations which are mostly formatted appropriately.	Expected level of use of Matplotlib for visualisations which may need further formatting.	Basic use of Matplotlib for visualisations which may contain errors and/or be unclear.	Limited or no use of Matplotlib for visualisations which may contain significant errors.
User Interface (UI) and User Experience (UX)	System UI provides seamless UX, providing clear and accessible results for users wishing to verify findings.	System UI provides an effective UX, with mostly intuitive navigation and accessible results that allow users to verify findings with minor difficulty.	System UI provides expected level of UX. Some elements may be confusing or incomplete, but results are generally accessible for verification.	System UI provides basic UX, with limited usability. Access to results is inconsistent or difficult to interpret for verification purposes.	System lacks a functional UI. Results are unclear, inaccessible, or not provided in a way that supports verification.
Markdown (.md) Documentation	Sophisticated use of Markdown to perform and clearly document data analysis. Markdown is well-structured, polished, and enhances readability and understanding.	Proficient use of Markdown to perform and document data analysis. Markdown is used effectively, though formatting could be more refined for clarity.	Expected use of Markdown with mostly correct data analysis. Markdown is present but lacks structure or clarity in explanation.	Basic use of Markdown. Documentation is minimal or unclear; Markdown formatting is poorly applied.	Markdown is either not used or contains limited relevant analysis or documentation.
Version Control	Extensive and consistent use of GitHub to document development. Commits are meaningful, frequent, and clearly describe progress and iterations.	Proficient and consistent use of GitHub to document development. Most commits are relevant and track progress adequately.	Expected use of GitHub commits are made to document development. However, commit messages may lack clarity or consistency.	Progressing use of GitHub. Few commits are made, and commit messages are often unclear or unrelated to specific stages of development.	Little to no meaningful use of GitHub. Commits are absent, minimal, or do not reflect actual development progress.

**Phase 4: Testing and Evaluating (20 marks)**

<b>/20</b>	<b>10 Marks (Per row)</b>	<b>8 Marks (Per row)</b>	<b>6 Marks (Per row)</b>	<b>4 Marks (Per row)</b>	<b>2-0 Marks (Per row)</b>
Data Analysis and Conclusion CT5-DPM-01 CT5-COM-01	Comprehensive data analysis leading to strong, well-supported conclusions.	Thorough data analysis with robust conclusions.	Satisfactory data analysis with reasonably clear conclusions.	Elementary data analysis with unclear conclusions.	Limited data analysis provided.
Evaluation CT5-DPM-01	Comprehensive project evaluation, considering all aspects and criteria, including feedback from peer reviews.	High-quality project evaluation, considering all aspects and criteria, including feedback from peer reviews.	Satisfactory project explanation, addressing most aspects and criteria, including feedback from peer reviews.	Basic project description which may not address peer reviews.	Limited evaluation or description of any aspect of the project provided.

**Total      /100      Feedback:**