SUPERVISOR TO ATTACH
PROCESSING LABEL HERE

Write your **student number** in the boxes above. **Letter**

Applied Computing: Data Analytics

Question and Answer Book

VCE Examination – Tuesday 12 November 2024

- Reading time is **15 minutes**: 2.00 pm to 2.15 pm
- Writing time is **2 hours**: 2.15 pm to 4.15 pm

Materials supplied

- Question and Answer Book of 24 pages
- Detachable Insert containing a case study for Section C in the centrefold
- Multiple-Choice Answer Sheet

Instructions

- Detach the Insert from the centre of this book during reading time.
- Follow the instructions on your Multiple-Choice Answer Sheet.
- At the end of the examination, place your Multiple-Choice Answer Sheet inside the front cover of this book.

Students are **not** permitted to bring mobile phones and/or any unauthorised electronic devices into the examination room.

Contents

VICTORIAN CURRICULUM AND ASSESSMENT AUTHORITY

VICTORIA State Government

Section A – Multiple-choice questions

Instructions

- Answer **all** questions in pencil in the spaces provided.
- Choose the response that is **correct** or that **best answers** the question.
- A correct answer scores 1; an incorrect answer scores 0.
- Marks will **not** be deducted for incorrect answers.
- No marks will be given if more than one answer is completed for any question.

Question 1

What is the purpose of a query in a database?

A. to format data

B. to validate data

C. to define data types

D. to find specific data or filter data by criteria

Question 2

The most appropriate design tool for representing the appearance of a worksheet in a spreadsheet is

A. a table.

B. a chart.

C. a storyboard.

D. a data dictionary.

Question 3

The stage of the problem-solving methodology when the querying of data stored in large repositories takes place is

A. analysis.

B. design.

C. development.

D. evaluation.

Question 4

Which one of the following is **not** a convention for an infographic?

A. clarity of message

B. use of colours to match the theme

C. use of informative statistics

D. an attention-grabbing heading

Question 5

Which of the following is the best naming convention for a database table that stores the sales data of a company?

A. salesdata

B. SALES DATA

C. tblSalesData

D. Table Sales Data

Question 6

What is the main purpose of an infographic?

A. to generate data from graphical elements

B. to make data look more attractive and colourful

C. to manipulate data to support a certain argument

D. to communicate data insights effectively and efficiently

Question 7

The most suitable technique to improve the effectiveness of an infographic is

A. minimising file size.

B. using easy-to-read fonts.

C. hiring an external designer.

D. uploading it on the company website.

Question 8

Bridget developed a map displaying median house prices for different suburbs. Users can either look at suburb medians or zoom in to the map to get more accurate neighbourhood median prices.

This is an example of

A. a histogram.

B. an infographic.

C. a validation technique.

D. a dynamic data visualisation.

Question 9

Soccer players often use wearable sensors so teams can collect biological and spatial data on the player. Clare is analysing a large set of data, creating a research question, and then adding new data as it comes, to see whether her research question has merit. She currently has over 100 different data files that she has been working with over the last season.

Which method would best help her remain organised?

A. Name each file as an incrementing number.

B. Create a series of folders and save 10 files in each folder.

C. Name files using each player's surname and date of birth.

D. Have a folder structure, organising files by month and player, and files named with date created.

Question 10

Mateo's organisation is seeking government funding to help educate teenagers about the dangers of riding scooters. Mateo is developing a plan for his team to produce a dynamic data visualisation for their meeting with the minister.

Which one of the following should Mateo check when assessing the effectiveness of the team's initial plan?

A. that changes are logged

B. that all annotations are accurate

C. that the graphics used are eye catching

D. that all team members understand their roles and tasks

Question 11

Ngaio has developed two sets of design ideas for an infographic to educate school children on the dangers of random waves around cliffs. He wants to determine which of the two design ideas to select.

Which one of the following is the best method for selecting the most suitable design idea?

A. Show them to his co-workers for their opinions.

B. Run a competition for school children to choose the best design.

C. Apply the evaluation criteria to see which design best meets the requirements.

D. Place each design at different dangerous beaches and see which beach has the fewest incidents.

Question 12

Which of the following is a recommended practice for verifying data?

A. Collect your own data before data entry.

B. Remove any outliers after data entry.

C. Use multiple sources and compare them after data entry.

D. Use only data provided by a registered business before data entry.

Question 13

Klaus is considering installing a wired network in the new clothing store that he manages. His business partner insists that a wireless network would be better, as customers would be able to use it as well.

What is the main advantage of having a wired network?

A. It is more secure.

B. It is quicker to set up.

C. It is easier for the end user.

D. It can be accessed from any location in the store.

Question 14

Traci is a physical therapist who keeps detailed records about her clients' needs. This includes a range of information relating to their physical fitness, flexibility, history of injuries and illnesses, and their progress data.

Traci works by herself and only has a small private business in New South Wales that turns over less than one million dollars a year.

She is approached by a sales representative who asks for her clients' contact details so that they can offer them special deals on fitness supplies.

If Traci gives them this data, she would be in breach of the

A. *Spam Act 2003*.

B. *Privacy Act 1988*.

C. *Health Records Act 2001*.

D. *Data and Privacy Protection Act 2014*.

Question 15

After recently experiencing an events-based security threat to his data, Sasha has been considering improving his protocols to include a disaster recovery plan and data backup.

His best option for securing data via backing up would be to use

A. online cloud-based storage.

B. a flash drive kept in his work bag.

C. a server located in a locked room.

D. internal storage on his laptop, which travels with him.

Question 16

Brett is a teacher who collects data about his Year 12 Data Analytics class at the start of each year. He adds this data to a database and uses this data, along with data from previous years' classes, to make decisions about how he will teach this year's class and what topics he will give extra focus to.

Which of the following sets of characteristics would be most important to ensure the integrity of Brett's data?

A. accuracy, correctness and reliability

B. accuracy, authenticity and timeliness

C. correctness, reliability and timeliness

D. authenticity, reasonableness and relevance

Question 17

When entering data to create charts displaying the comparative heights of famous landmarks, an example of diminished data integrity could be

A. not verifying the final charts.

B. not using a reputable antivirus program.

C. not requiring users to log in before data entry.

D. not checking that all inputted measurements are using metres

Question 18

A school is implementing a new reporting system where student scores are recorded in two different formats – a letter grade for exams and attendance as a count.

In the reporting system, what data types would be most appropriate to handle student scores in letter grades for exams and attendance as a count?

A. string for exam letter grade and string for attendance count

B. string for exam letter grade and integer for attendance count

C. floating point for exam letter grade and integer for attendance count

D. floating point for exam letter grade and floating point for attendance count

Question 19

In a biodiversity study that assesses the population density of different animals in a forest ecosystem, which type of data is more suitable to directly compute averages, conduct statistical analyses, and identify patterns in animals' distribution?

A. quantitative data, such as counts of individual animals

B. qualitative data, such as detailed observation about animal behaviour

C. primary data collected through observations about animal food consumption

D. secondary data obtained from existing ecological literature on behaviour patterns of animals

Question 20

Elise wants to develop an infographic about the influence of pet ownership on teenagers' happiness and has developed four possible research questions.

Which one of the following would be the most appropriate research question to use?

A. Do pets impact teenagers' happiness?

B. Does the type of pet you have make you happy?

C. Is there a correlation between daily pet interaction and teenagers' happiness?

D. What are the various aspects of pet ownership that influence teenagers' happiness?

End of Section A

Section B – Short-answer questions

Instructions

- Answer **all** questions in the spaces provided
- Write your responses in English.

Question 1 (4 marks)

A finance organisation recently completed an audit of its current data and information security strategies.

The consultants made a recommendation to the IT manager to improve the use of physical security controls. This included using biometrics for different departments in the organisation and shredding confidential documents when they were no longer required.

Provide a description of each of the physical security controls and explain why the organisation should implement them.

a. Using biometrics 2 marks

Description _____

_____

_____

Reason for implementing _____

_____

_____

b. Shredding confidential documents 2 marks

Description _____

_____

_____

Reason for implementing _____

_____

_____

Question 2 (5 marks)

An airline is running a promotion for its Gold Class members to upgrade from economy to business class if their rewards points are above 3000. The airline is using a data entry form that contains the three items shown below, to enable a customer to upgrade their ticket on a flight.

Item 1	**Membership number**	*****
Item 2	**Are you a Gold Class member?**	Yes No
Item 3	**Rewards points**	**Text box**

a. Explain why the text box used to input the rewards points should include a range check when the data is entered. 2 marks

b. Develop a query design that would ensure that only valid customers can access this upgrade to business class tickets. 3 marks

Question 3 (3 marks)

Michelle took some unique photographs of a Mount Etna eruption to use as part of an infographic about volcanoes for geography students. She approached her IT-savvy friend, Liz, to ask how to backup these photos.

Liz suggested that Michelle email all of the photos to Liz, who will upload them to the network attached storage (NAS) on her eSport website.

a. If she follows Liz's suggestion, identify **one** way that Michelle could confirm that the backup was successfully completed. 1 mark

b. Explain **two** concerns that you have with Liz's backup strategy. 2 marks

Question 4 (4 marks)

When presenting data, several audience characteristics must be considered. These include age appropriateness, commonality of language, cultural inclusiveness and gender.

The Federal Reserve Bank presents financial data and short-term predictions to home loan providers such as banks every four months.

Identify which audience characteristic is the most important in the information provided. Justify your response.

Question 5 (4 marks)

Oscar works for a company that creates web-based computer games. The games are very successful, and the company plans to release a new game in six months' time. This has generated a lot of excitement in the gaming community, and the company has noticed a significant increase in the number of failed attempts to log in to the developer portal. The failed logins have used usernames in sequence, such as 'Smith01', 'Smith02', 'Smith03' up to 'Smith99'.

Oscar's employer has requested a set of criteria that could be used to evaluate the effectiveness of the company's current data and information security strategies.

a. Identify one criterion that Oscar's employer could use to achieve this purpose. Explain why it should be checked. 2 marks

Criterion _____

Explanation _____

_____

_____

_____

_____

b. Recommend a strategy to ensure that bots and other computer-automated attacks cannot break in to the password-protected portal. 2 marks

_____

_____

_____

_____

End of Section B

Section C – Case study

Instructions

- Write your responses in English.
- Please remove the Insert from the centre of this book during reading time.
- Use the case study provided in the Insert to answer the questions in this section. Answers must apply to the case study.
- Answer **all** questions in the spaces provided.

Question 1 (6 marks)

The snow season begins in June and Tom would like to place ski orders in early April. To assist her dad with placing orders, Georgie has planned to complete her project by 24 February. To keep track of all the tasks for her project, she has decided to create a Gantt chart. She has listed all the key tasks, dependencies and durations in Table 1 in the case study Insert.

a. Identify the three errors in the Gantt chart on the following page by circling and numbering them on the Gantt chart. 3 marks

b. Describe what needs to be done to address each error identified in **part a**. 3 marks

Error 1 correction _____

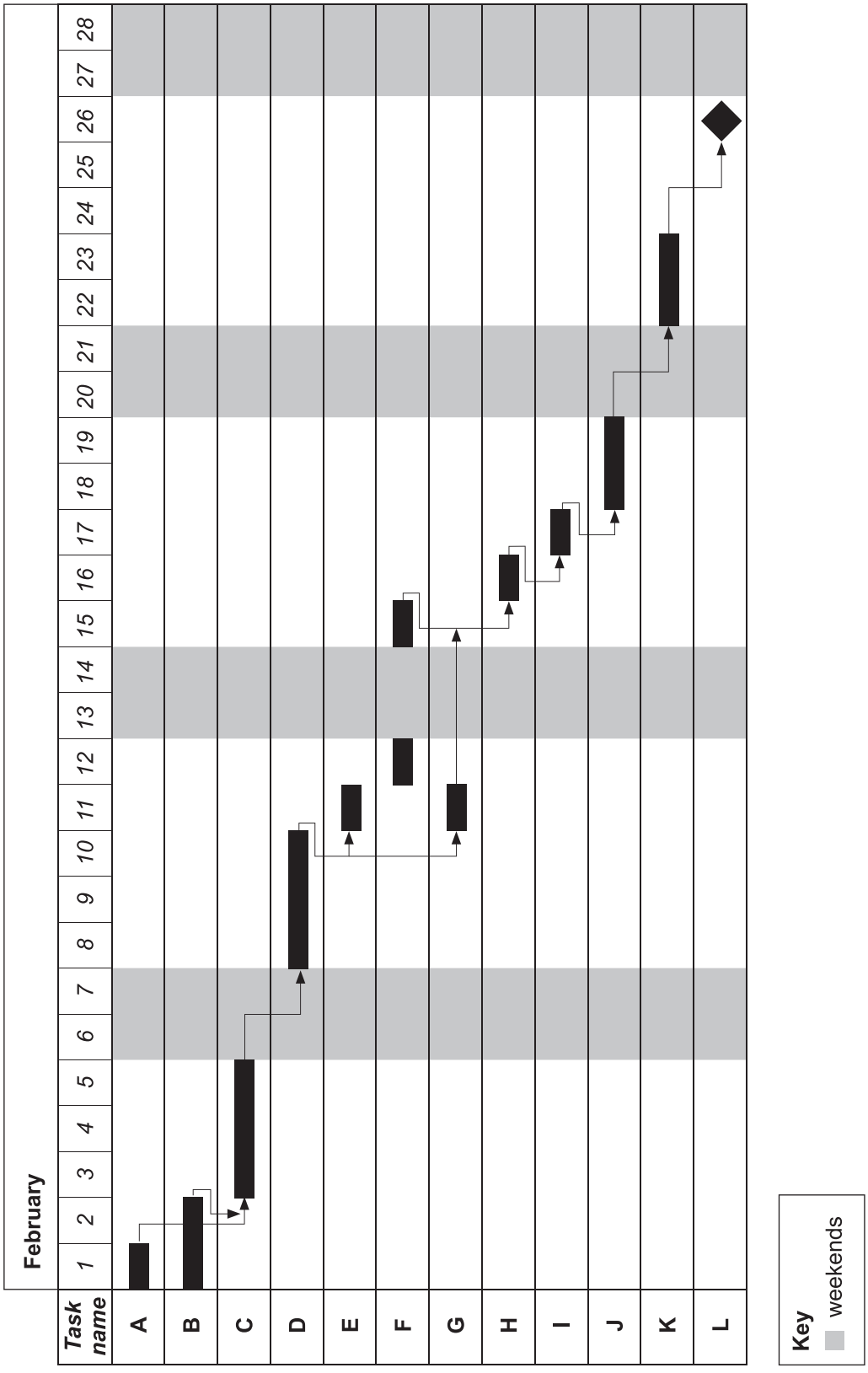
_____

Error 2 correction _____

_____

Error 3 correction _____

_____


February
Task name
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
A
B
C
D
E
F
G
H
I
J
K
L
Key
weekends

Question 2 (8 marks)

Georgie needs to collect and analyse data from the ski shops to which they supply ski equipment and clothing in order to determine the shops' historical purchasing habits.

a. State and describe two suitable methods for Georgie to collect the required data. 4 marks

Method 1

Method 2

b. Describe two advantages of **one** of the methods suggested in **part a**. 4 marks

Method selected

Advantage 1

Advantage 2

Question 3 (5 marks)

Georgie is trying to assist Tom by providing as much data as possible. She has friends who work for ski shops in Canada and the United States and asks them for some historical data regarding their sales.

a. Describe **one** concern about providing too much data. 2 marks

b. Should Georgie use the data from Canada and the United States? Explain your response without reference to scope. 3 marks

Question 4 (6 marks)

As part of her analysis, Georgie needs to determine the solution requirements, constraints and scope.

Complete the table below by identifying **one** example of each of the following categories, along with a description of how it is a relevant example.

Category	Example	Description
Functional requirement		
Constraint		
Scope		

Question 5 (4 marks)

As part of the design process, Georgie must develop evaluation criteria that will be used to evaluate the design ideas for the dynamic data visualisations.

Provide two examples of an evaluation criterion and explain why one is efficient and one is effective.

Evaluation criterion 1 _____

Explanation of how it is efficient _____

_____

Evaluation criterion 2 _____

Explanation of how it is effective _____

_____

Question 6 (4 marks)

Georgie has developed a preferred design for her dynamic data visualisations shown in Figure 2 of the case study.

Identify and explain two design principles used that influence the functionality and appearance of the user controls.

Functionality

Design principle _____

Explanation _____

_____

_____

Appearance

Design principle _____

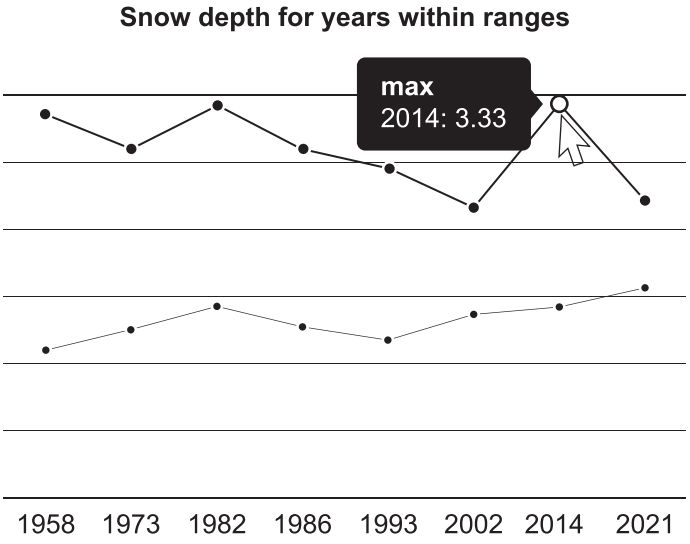
Explanation _____

_____

_____

Question 7 (3 marks)

When Georgie showed one of the dynamic data visualisations (Figure 3) to Tom, Tom was uncertain about the information being shown in the snow depth graph (as shown below) and was unable to correctly interpret the graph.



List three conventions that should be applied to the snow depth graph to ensure that Tom can accurately interpret it.

Convention 1 _____

Convention 2 _____

Convention 3 _____

Question 8 (2 marks)

Georgie needs to select a software tool to develop the dynamic data visualisations.

Identify two different software functions required in the software tool for Georgie to produce the dynamic data visualisations.

Software function 1 _____

_____

_____

Software function 2 _____

_____

_____

Question 9 (6 marks)

Georgie has used an automated tool to extract data from the Snowy Hydro snow depths data website and obtained the following data:

Location	Date of measurement	Snow depth
Deep Creek	15 Jun 2022	82.3 cm
Deep Creek	04 Jun 23	0 cm
Spencers Creek	15/6/23	118.3
Spencers Creek	6/23/2023	–110.7 cm
Spencers ©reek	10 Aug 2023	100 m
Spencers Creek	15 Jun 2023	118.3 cm

After going through the data, Georgie has become concerned about the integrity of the data she obtained.

a. Select one item from this data set and explain why it demonstrates a lack of integrity. 2 marks

b. Describe how Georgie could cleanse and validate this data in an efficient manner to ensure the integrity of the data collected. 4 marks

Cleanse:

Validate:

Question 10 (2 marks)

Georgie will need to test her dynamic data visualisations while she is developing them to ensure they perform as intended.

Describe a testing technique that Georgie could use to test that the dynamic data visualisations are visualising the correct data.

Question 11 (3 marks)

Georgie has now completed the dynamic data visualisations and is considering what to do with the original data, as it is taking up a lot of storage space.

Recommend a procedure for managing those files that are still needed, but not required for the running of the dynamic data visualisations.

Question 12 (2 marks)

Tom looks at the data that Georgie collects from ski shops to ensure that it includes their names, addresses, email addresses and billing information. He also ensures that his data from his suppliers about stock items, prices, safety information and payment options is up to date.

Explain why this data is important to Tom's organisation.

Question 13 (4 marks)

Some time after the dynamic data visualisations have been completed, Tom goes to his office early one day and discovers that the main office desktop computer is on and logged into. A spreadsheet file containing billing information is open and clearly visible. Tom cannot tell if any details have been viewed, changed, copied or deleted. Georgie notices later that day that some data has been deleted.

Recommend two security controls that could be implemented and how they would have prevented this from happening.

Security control 1: _____

_____

_____

_____

Security control 2: _____

_____

_____

_____

Question 14 (5 marks)

As a respected member of the snow-sports industry, Tom has been asked by the Victorian Government to work as a consultant for an industry panel for alpine resorts. As part of this, Tom has been provided with sensitive data that includes the income and expenditure of all ski resorts in Victoria. This also includes salaries of people who manage these resorts. Tom has copied this data onto a large capacity flash drive, which is stored on his keyring for easy access.

a. Identify any relevant legislation that Tom must comply with while having access to this data. 1 mark

b. Explain why the legislation is relevant. 2 marks

c. Outline **one** issue that Tom may have with his treatment of this data. 2 marks

End of Question and Answer Book

Applied Computing: Data Analytics

Insert for Section C

Please remove from the centre of this book during reading time.

VICTORIAN CURRICULUM AND ASSESSMENT AUTHORITY
VICTORIA State Government

Ski stuff for you

Tom is a leading importer specialising in ski equipment and clothing. He is committed to sustainability and driven to provide ethically sourced products consisting of top-quality ski gear to ensure every adventure on the slopes contributes to preserving the beauty of the mountains for future generations.

Tom has struggled to efficiently forecast the ordering of ski equipment and clothing for the next ski season. He shared his concern with his daughter, Georgie, who has recently finished studying Data Analytics. He has identified that the biggest influence on sales is both the maximum and average snow depths for the relevant Australian ski season. Tom would like a software tool that will allow him to forecast the snow falls and determine the ordering of ski equipment and clothing, based on historical data.

Tom has asked Georgie to develop some dynamic data visualisations to assist in forecasting the ordering of ski equipment and clothing. He has requested that this be completed by 26 February. The dynamic data visualisations will help to determine the likely success of the upcoming snow season for Tom's organisation.

Tom and Georgie determine that the dynamic data visualisations should:

- be able to integrate real-time weather data from reliable sources
- use historical weather data for trend analysis
- contain charts or graphs that display correlations between weather patterns and sales of ski clothing and equipment
- have minimal load times for the dynamic data visualisations and data updates

Georgie decides to begin the project on 1 February and complete it by 24 February. She has some restrictions on when she can work on producing the dynamic data visualisations. She can't work on the weekends as she has training to be ready to compete in snowboarding during the snow season, and she will be attending a conference on 5 February. Georgie has decided to create a Gantt chart to keep track of all the tasks. Before she creates the chart, she lists the key tasks, durations and dependencies in Table 1, shown on the following page.

Table 1

Items		Duration in days	Dependencies
A	Design and send survey forms to existing ski customers.	1	None
B	Gather secondary data.	2	None
C	Collect and collate results.	3	A, B
D	Organise, cleanse and manipulate data.	3	C
E	Brainstorm design ideas.	1	D
F	Create mock-up diagrams.	2	E
G	Develop data dictionary.	1	D
H	Receive review and feedback from her friend Angie.	1	F, G
I	Edit design based on feedback.	1	H
J	Create dynamic data visualisations.	2	I
K	Test data visualisations and fix errors.	2	J
L	Complete dynamic data visualisations.	0	K

After conducting research to gather some secondary data, Georgie obtains data that she hopes will enable her to develop the dynamic data visualisations. The data includes:

- historical snow depth data from the Snowy Hydro <https://www.snowyhydro.com.au/generation/live-data/snow-depths/>
- historical Oceanic Niño Index data used to determine the El Niño-Southern Oscillation from the American National Weather Service <https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php>
- long-range temperature forecasts for Victoria from the weather service website
- historical sales data from Tom.

This will then enable Tom to forecast what the upcoming season may be like, and compare it with sales for similar seasons in the past.

Georgie produces the following preferred design for the dynamic data visualisations. This design has been laid out as shown to help Tom understand how the data visualisations will work.

Section 1 – Long-range temperature forecast

This graph is imported from the Victoria weather service website and shows the long-range temperature forecast as of January 2024. It shows the maximum, minimum and predicted temperature difference from the monthly average temperature.

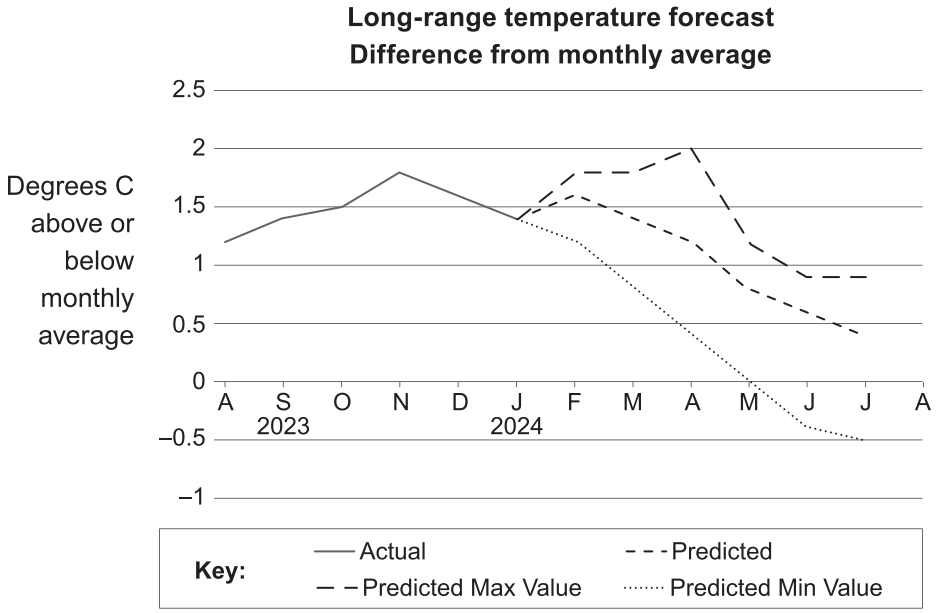


Figure 1

Section 2 – User data entry controls

The user sets maximum and minimum data values for the difference from the monthly average temperatures to select the historical snow and sales data, which will be shown in the dynamic data visualisations summary data.

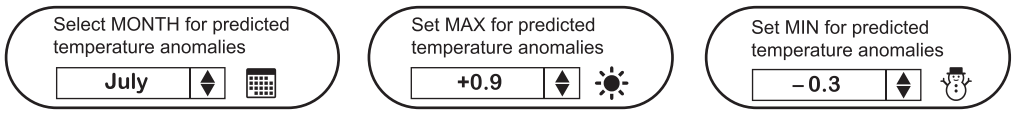


Figure 2

Section 3 – Visualisation of historical data

The summary data for years that fall within the selected data in the user input controls.

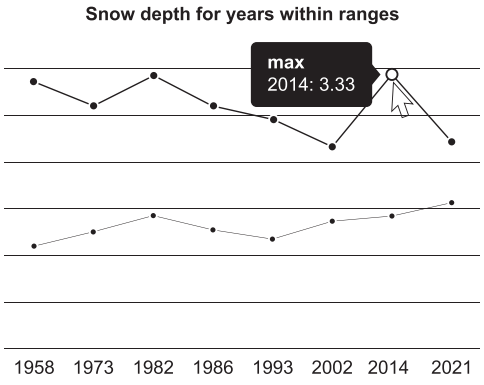


Figure 3

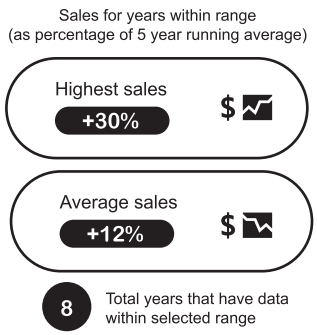


Figure 4

End of Insert