

STUDENT NUMBER Letter

APPLIED COMPUTING: DATA ANALYTICS

Written examination

Monday 7 November 2022

Reading time: 3.00 pm to 3.15 pm (15 minutes)

Writing time: 3.15 pm to 5.15 pm (2 hours)

QUESTION AND ANSWER BOOK

Structure of book

Section	Number of questions	Number of questions to be answered	Number of marks
A	20	20	20
B	6	6	20
C	13	13	60
			Total 100

- Students are permitted to bring into the examination room: pens, pencils, highlighters, erasers, sharpeners and rulers.
- Students are NOT permitted to bring into the examination room: blank sheets of paper and/or correction fluid/tape.
- No calculator is allowed in this examination.

Materials supplied

- Question and answer book of 27 pages
- Detachable insert containing a case study for Section C in the centrefold
- Answer sheet for multiple-choice questions

Instructions

- Detach the insert from the centre of this book during reading time.
- Write your **student number** in the space provided above on this page.
- Check that your **name** and **student number** as printed on your answer sheet for multiple-choice questions are correct, **and** sign your name in the space provided to verify this.
- All written responses must be in English.

At the end of the examination

- Place the answer sheet for multiple-choice questions inside the front cover of this book.
- You may keep the detached insert.

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

THIS PAGE IS BLANK

DO NOT WRITE IN THIS AREA

SECTION A – Multiple-choice questions**Instructions for Section A**

Answer **all** questions in pencil on the answer sheet provided for multiple-choice questions.

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

Minoo is writing a survey to collect primary data from her customers about their opinions regarding potential new products for her store. One question she plans to ask is if the survey participant is interested (I), very interested (V) or not interested (N) in a particular new product.

Which one of the following is the most appropriate data type for Minoo to use when storing participant responses?

- A. binary
- B. numeric
- C. Boolean
- D. character

Question 2

An online printing business creates photo books and albums for their customers. They have noticed that their online sales have dropped by 25% in the last three months. They have asked their customers for feedback about the quality of their photo books and albums.

This data collection will help the business to

- A. understand how their customers behave online.
- B. determine if they need to modify their products.
- C. identify if their organisational structure needs improving.
- D. decide how to improve customer experience on their website.

Question 3

Robbi is a university student who wants to conduct some research on the migration patterns of native Australian birds. She finds a very large database published online that contains relevant data about birds in Australia.

After using this data in her research, Robbi must

- A. publish the data herself.
- B. seek consent from her university to use the data set.
- C. check each record in the data set to make sure it is accurate.
- D. ensure that the data set is referenced using a standard referencing system.

Use the following information to answer Questions 4–6.

Jumpy Joey Gym is a new fitness centre for children aged six months to 16 years. The owner, Virginia, wants to advertise to parents and caregivers the benefits of their children participating in physical activities that are supervised, safe and fun.

Virginia is currently storing membership details in a spreadsheet, as shown below.

FirstName	LastName	Address	DateOfBirth	MemberType	LastVisit	Active
Jaya	Singh	9056 Nisal Rd, Neque VIC 3100	8/8/2018	casual	23/9/2022	yes
John	White	NULL	25/7/2015	frequent	12/10/2022	no
Dilay	Aksoy	5 Magnis Court, Elementum VIC 3101	19/11/2019	ultimate	25/10/2022	yes
Sophy	Singh	9056 Nisal Rd, Neque VIC 3100	8/8/2018	casual	23/9/2022	yes
John	White	64 Miranda Drive, Elementum VIC 3101	31/1/2021	casual	2/7/2022	yes

Question 4

Virginia is creating an infographic that will include a chart that displays different physical activities that children participate in at Jumpy Joey Gym. This chart would visualise the amount of time spent on each of those activities relative to the total physical activity time, as a percentage. Virginia typically has children participate in five physical activities.

The most appropriate chart type for Virginia to use within the infographic is a

- A. pie chart.
- B. bar chart.
- C. line chart.
- D. density map.

Question 5

Virginia is expecting membership numbers to increase after her advertising campaign, and she wants to convert her current membership spreadsheet into a membership database so that she can extract and manipulate the data and generate reports more easily.

Which one of the following would be the most appropriate to use (after normalisation) to represent the relationship between users and their addresses?

A.

User
<u>UserID</u>
FirstName
LastName
<u>AddressID</u>
DateofBirth
MemberType
LastVisit
Active

Address
<u>AddressID</u>
Street
Suburb
State
Country

B.

User
<u>UserID</u>
FirstName
LastName
AddressID
DateofBirth
MemberType
LastVisit
Active

Address
<u>AddressID</u>
Street
Suburb
State
Country

C.

User
<u>FirstName</u>
<u>LastName</u>
Address
DateofBirth
MemberType
LastVisit
Active

Address
Street
Suburb
State
Country

D.

User
<u>FirstName</u>
<u>LastName</u>
Street
Suburb
State
Country
DateofBirth
MemberType
LastVisit
Active

Question 6

Virginia wants to find out the top 10 suburbs where her customers live so that she can target the nearest shopping centres when purchasing advertising space to display her infographic. She needs to use a query to do this.

An appropriate query on customer suburbs would contain

- A. a sum and filter only.
- B. a sum, sort and filter.
- C. a count, sort and filter.
- D. a search, count and filter.

Question 7

The Australian Bureau of Statistics (ABS) directly collects data from households around Australia as part of the Census every five years.

The data that is collected by the ABS is known as

- A. new data.
- B. existing data.
- C. primary data.
- D. secondary data.

Use the following information to answer Questions 8 and 9.

The Director of Administration in a large secondary school is preparing data for Years 8 to 12 teachers that includes student scores and comments for all the assessment tasks and examinations that students have completed during 2022. The summarised data will be used to inform homeroom teachers for 2023 of their students' progress during 2022. A meeting for all 2023 homeroom teachers is planned for December 2022.

Question 8

The Director of Administration has noticed that incorrect scores and comments have been recorded in the data set for some students. This has had an impact on their average marks, which have been calculated over the course of the year.

A criterion to check the integrity of the impacted data is

- A. accuracy.
- B. relevance.
- C. timeliness.
- D. authenticity.

Question 9

The Director of Administration will need to go through each individual student's statement of results to ensure the recorded data is correct and, if not, to correct any errors. She estimates that it will take her staff until late February 2023 to complete this process.

The criterion to check the integrity of data that has been affected is

- A. relevance.
- B. timeliness.
- C. correctness.
- D. reasonableness.

Question 10

A student studying Data Analytics is interested in pursuing further studies in a science, technology, engineering and mathematics (STEM) career when leaving school. They are also interested in the popularity of STEM-related studies in Years 11 and 12 at their school. The student would like to collect some primary data about the numbers of students in STEM-related classes by going to each of the classes and counting the number of students in the class.

This method of collecting data is an example of

- A. counting students.
- B. observing students.
- C. surveying students.
- D. interviewing students.

Question 11

The scope of a data visualisation solution describes

- A. the conditions or limitations when designing the solution.
- B. the boundaries or parameters of the solution.
- C. the quality attributes of the solution.
- D. what the solution should do.

Question 12

Database normalisation is a process that

- A. minimises duplication of data across multiple database tables.
- B. increases duplication of data across multiple database tables.
- C. reduces the number of tables in a database.
- D. uses formulas and filters to extract data.

Question 13

Michael is deciding how to secure files that contain sensitive information on his laptop.

Michael could secure these files using

- A. password protection and malware.
- B. password protection and encryption.
- C. password protection and VPN access.
- D. password protection and internal documentation.

Question 14

Which one of the following is **not** an important feature of software used to develop dynamic data visualisations?

- A. updating charts based on live data
- B. exporting to PDF or JPEG file formats
- C. producing reports based on set criteria
- D. enabling users to interact with the charts

Use the following information to answer Questions 15 and 16.

Alana is a technology consultant for medium to large businesses in Victoria. She is currently working with a national furniture business. The business has a number of staff who are sometimes required to work at different sites on different days of the week, as well as some staff who travel interstate from time to time.

The business wants to minimise the number of files and data that need to be emailed or transferred via portable USB hard drives. There have been complaints about files being out of date, errors occurring within them and USB hard drives being lost.

Question 15

Alana suggests to the business that they should consider a cloud-based storage solution, in order to minimise the number of files and data being emailed between factories or distributed via USB hard drives.

The main advantage of using cloud-based storage in this situation is that

- A. version control for the files and data can be more easily maintained.
- B. access to the data and files can be monitored.
- C. the files and data can be disposed of securely.
- D. the files and data can be archived.

Question 16

Loss of a USB hard drive is an example of

- A. an external threat to the business's data.
- B. a deliberate threat to the business's data.
- C. an accidental threat to the business's data.
- D. an events-based threat to the business's data.

Question 17

Johnny has purchased second-hand USB drives from a company. He has discovered that some of the company's files and data are still accessible on some of the drives. Johnny has decided to publish the company's files and data online.

What is a potential consequence for the company if the files and data are published online?

- A. an increase of trust and improved reputation in the community
- B. decreased risk of being targeted by cybercriminals in the future
- C. a need to develop a risk management plan to prevent this occurring again
- D. loss of trust and reputation within the community, and a potential loss of clients

Question 18

A data analyst has been asked to verify the data presented in a chart (as part of a data visualisation) before it is published.

Which one of the following would be the least efficient method of verifying that the chart is presenting the data accurately, if the data presented in the chart represents 1000 records?

- A. reproducing the chart in a spreadsheet
- B. reproducing the chart using data visualisation software
- C. summarising the data and reproducing the chart by hand
- D. checking the values in the chart are correct using relevant formulas

Question 19

A music festival is being held in Melbourne. It is anticipated that 20 000 people will attend the festival. The festival organisers decide to establish an unsecured 802.11ac wireless network that will be available to all festival attendees and performers.

One of the key reasons for choosing 802.11ac in this situation is the

- A. high level of data security.
- B. low cost of implementation.
- C. high level of device compatibility.
- D. number of users able to connect to the network at once.

Question 20

When incorporated as part of a multi-factor authentication procedure, using a password, biometrics and a text message that is sent to a previously enrolled mobile phone number is one way that

- A. data can be disposed of securely.
- B. malware can be transferred onto a network.
- C. dynamic data visualisations can be generated.
- D. unauthorised access to systems and data can be prevented.

SECTION B – Short-answer questions**Instructions for Section B**

Answer **all** questions in the spaces provided.

Question 1 (4 marks)

An Australian athletics official has collected thousands of track and field event results for athletes all around Australia in the last 10 years. This data has been collected from a range of data sets and then imported into a spreadsheet. Before the official can present findings using a data visualisation, they will have to first cleanse the data.

- a. Explain why cleansing data is a necessary process during data manipulation and analysis. 2 marks

- b. Describe one technique the official can use to cleanse data in the spreadsheet, and provide an example. 2 marks

Description _____

Example _____

DO NOT WRITE IN THIS AREA

Question 2 (2 marks)

A seafood restaurant owner, Raf, is implementing a cloud-based ordering platform that uses QR codes so customers can place orders while they are sitting at their tables. A customer places their order by using a mobile phone to connect to the restaurant’s guest wi-fi, scanning the QR code on their table and then choosing menu items. Once an order is placed, it is sent to tablet-based terminals in the restaurant’s kitchen and bar areas (also connected via wi-fi). The ordering platform is linked with the restaurant’s point-of-sales systems (stored locally and connected via ethernet cable) in order to make payment for orders easy for customers and staff.

Select two digital system components from the following list and explain the role of each component at the restaurant.

- local server
- cloud server
- switch
- wireless access point
- router

Component 1 _____

Explanation _____

Component 2 _____

Explanation _____

DO NOT WRITE IN THIS AREA

Question 3 (8 marks)

Alexis is developing a range of spreadsheet, database and data visualisation solutions for a banking and finance client.

- a. Identify and explain a design tool that Alexis could use to describe the fields within the spreadsheets and database. 2 marks

Alexis's client wants her to design a dynamic data visualisation. The client's employees should be able to interact with and manipulate the data being visualised. The client wants to see patterns and trends relating to loan amounts, profit margins and spending data based on different criteria for locations such as state, region or suburb.

- b. Explain how Alexis can use software to develop a dynamic data visualisation to meet the client's requirements. 3 marks

- c. In preparation for designing the dynamic data visualisation, Alexis has asked the client whether they would like her to use a single annotated mock-up or a storyboard.

Which **one** of these design tools should the client ask Alexis to use? Justify your response. 3 marks

Question 4 (2 marks)

Narelle is preparing to normalise a database to third normal form (3NF).

Outline the conditions required for a database to be considered as being in third normal form.

Question 5 (2 marks)

An organisation involved in offering financial planning services to businesses has recently completed an audit of their data and information security strategies. They have been advised to improve their current data and information security strategies, particularly with respect to their disaster recovery plans. Issues were identified with their backing up of data and restoration plans. The manager of the organisation thought that restoration was part of backing up.

Explain the difference between a plan for backing up data and a restoration plan.

DO NOT WRITE IN THIS AREA

Question 6 (2 marks)

A contractor working with Victoria Police has been asked to email a colleague data that they have collected that involves speeding infringements in a particular suburb. The data that has been requested includes licence plate information and the personal details of the owners of vehicles.

Which piece of legislation does the contractor need to consider before communicating this data to their colleague? State the name of the Act and explain the legal issue.

Act _____

Explanation _____

DO NOT WRITE IN THIS AREA

SECTION C – Case study

Instructions for Section C

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 (1 mark)

Propose a research question that will guide Alison, Thomas and Hong Mei’s work in developing infographics on the perceptions and the issues associated with the Victorian transport network.



Question 2 (3 marks)

Alison develops a Gantt chart identifying the tasks that her team needs to complete for their research.

A partial and incomplete Gantt chart for part of the project is shown below.

Complete the Gantt chart below by showing:

- time allocations for tasks
- dependencies
- milestones.

Task	Duration (days)	1	2	3	4	5	6	7	8
									
7	Create draft survey questions								
8	Meet to refine questions								
9	Survey questions finalised								
10	Create survey								
11	Send survey to registered vehicle owners								
12	Cleanse survey data								
									

DO NOT WRITE IN THIS AREA

Question 3 (3 marks)

Alison has assigned Thomas to determine the constraints as part of the analysis process. Thomas has determined that there are three major constraints; however, he is not sure how to group them.

Assist Thomas by completing the table below, writing each of the constraints from the list beside the correct description.

economic usability legal

Constraint	Description
	Consideration of the ownership of the data used and the privacy of data requirements
	Consideration of the cost of the research and the time taken to complete it
	Consideration of the usefulness of the infographic to the targeted audience

Question 4 (2 marks)

Thomas is considering what to include in the scope for the infographics. He wants to be clear about what should and what should not be included.

Identify what should be included in the scope for the infographics and what should not be included.

Included _____

Not included _____

Question 5 (6 marks)

Once Hong Mei receives the survey results, she must conduct validation and verification on the data before it is processed.

- a. Hong Mei's first check of the data indicates that a text entry field was used rather than a date/time selector for the DateOfBirth question (see Figure 1 on page 2 of the insert). This has resulted in respondents providing the data in a range of formats (see Figure 2 on page 2 of the insert).

Identify an appropriate format that Hong Mei could utilise when converting the data stored in the DateOfBirth field in Figure 2 and provide an example of the format identified.

2 marks

- b. Suggest one technique that Hong Mei could use to **validate** the data. The technique should not be related to the DateOfBirth field in Figure 2. Recommend how she might resolve issues that may arise from using this validation technique.

2 marks

Technique _____

Recommendation _____

- c. Suggest one technique Hong Mei could use to **verify** the data. Recommend how she might resolve issues that may arise from using this verification technique.

2 marks

Technique _____

Recommendation _____

Question 6 (4 marks)

The survey was sent out to the owners of registered vehicles and data was collected.

One of the survey questions asked about how the transport network could be improved. Responses included text-based data in the form of single words, sentences and dot points. Hong Mei would like the data to be coded in order to identify patterns and relationships within the data set.

Outline a process for Hong Mei to code this qualitative data to support the manipulation of the data set.

DO NOT WRITE IN THIS AREA

Question 7 (6 marks)

Alison has tasked Thomas with collecting and collating data on Melbourne’s major roads and how they perform at different times of day and different times of year. He is investigating how to acquire data from the mapping software provider to see if it could assist the team.

Thomas discovers that the approval process to access the data via an online portal appears quite complex. He decides it would be easier to organise a group of university student volunteers to observe Melbourne’s major roads and freeways at different times of day.

Alison disagrees. She thinks that the observation data set would not be as effective or efficient compared to the data they could access from the mapping software provider’s data portal.

- a. Explain why Alison is correct to disagree, with reference to efficiency and effectiveness. 2 marks

Alison and her team have now received approval to utilise the traffic data from the mapping software provider’s data portal. The data has been provided in a CSV file. Thomas has been directed to start identifying patterns within the data. He chooses to view the CSV file using a spreadsheet application.

- b. A sample of the CSV file has been provided in the insert (see Figures 3 and 4 on page 3 of the insert).

Describe two techniques for manipulating the data in the CSV file to meet the focus of the infographics. 4 marks

Technique 1 _____

Description _____

Technique 2 _____

Description _____

DO NOT WRITE IN THIS AREA

Question 8 (4 marks)

Thomas and Hong Mei have been informed that the infographics will be published publicly as a way of highlighting the Victorian Government's roads and public transport initiatives heading towards 2050. Thomas and Hong Mei have made Alison aware of this.

Given the proposed publication of the infographics, outline two audience characteristics that should be incorporated into the infographics and describe how each of the audience characteristics might be considered within the design.

Audience characteristic 1 _____

Description _____

Audience characteristic 2 _____

Description _____

DO NOT WRITE IN THIS AREA

Question 9 (8 marks)

Hong Mei has generated some alternative design ideas by completing some quick sketches of a range of infographics. These are shown below.

Key

~~~~ heading

□ text

■ chart/visual

**Design 1**

Transport users' preferences

Weekly use of transport

Travel during day/night

Transport issues and perceptions

How transport network could be improved

**Design 2**

Transport users' preferences  
+  
Weekly use of transport  
+  
Travel during day/night

Transport issues and perceptions  
+  
How transport network could be improved

**Design 3**

Transport users' preferences  
+  
Weekly use of transport  
+  
Travel during day/night  
+  
Transport issues and perceptions  
+  
How transport network could be improved

a. Identify the design that Hong Mei should select based on the evaluation criterion 'The layout of the elements in the infographics is visually appealing'. Justify your choice. 3 marks

---



---



---



---



---



---

SECTION C – Question 9 – continued

DO NOT WRITE IN THIS AREA

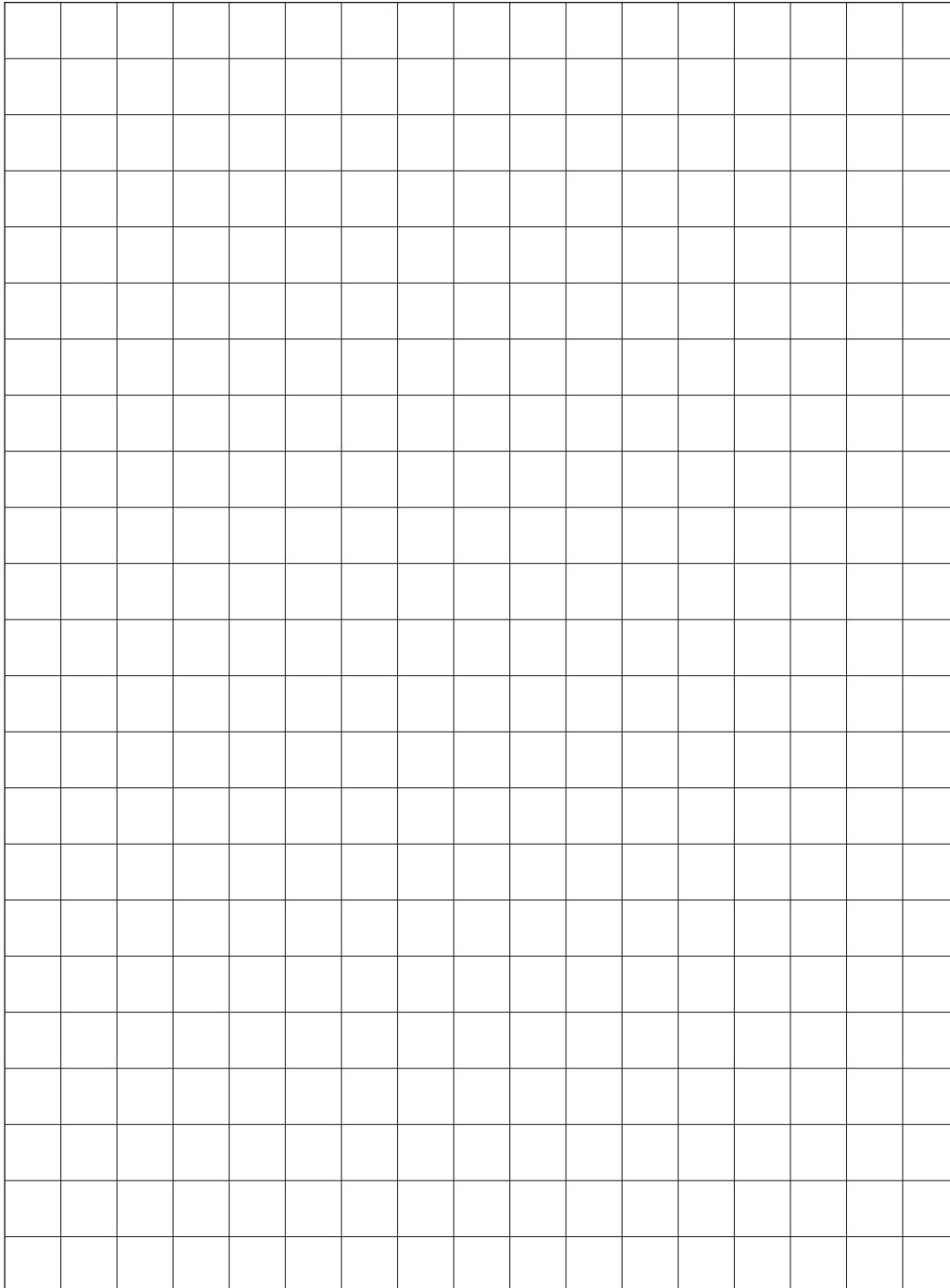
- b. Hong Mei is about to produce her detailed designs for the infographics. She is not too sure what she needs to do to complete these mock-ups, but she is aware that she needs to follow the design principles of alignment and balance.

On the grid below, draw a partial design that follows the required design principles and includes each of the following features:

- the title 'Transport Users' Preferences'
- text to describe the charts and to reference the sources of the data
- four charts to show each of the transport preferences

Ensure each feature of the design is clearly annotated.

5 marks



**Question 10** (2 marks)

Hong Mei will develop the infographics to show the findings of Alison’s research based on her preferred designs.

Before developing the infographics, she needs to consider the functional capabilities of the software she will use to create the infographics.

Outline two functions of the software that Hong Mei will need to create the infographics.

Software function 1 \_\_\_\_\_

\_\_\_\_\_

Software function 2 \_\_\_\_\_

\_\_\_\_\_

DO NOT WRITE IN THIS AREA

**Question 11** (6 marks)

Thomas imports the data from the CSV file from the mapping software provider (see Figures 3 and 4 on page 3 of the insert) into a database. He extracts the data that has been collected on weekdays. Thomas extracts the data using the query below.

|                             |            |           |            |
|-----------------------------|------------|-----------|------------|
| <b>Field</b>                | All fields | MetroArea | DayOfWeek  |
| <b>Sort</b>                 |            |           |            |
| <b>Show in query result</b> | Yes        | Yes       | Yes        |
| <b>Criteria</b>             |            | = True    | < 2 OR > 6 |

In order to check that the data is being extracted properly, Thomas breaks up the sample data into smaller sets of records (four per test) to test the query using each set.

- a. Complete the test table below, based on the information and the query provided above. Test X has been completed as an example.

3 marks

| Test no. | Test data    | Expected records returned | Actual records returned |
|----------|--------------|---------------------------|-------------------------|
| X        | Records W–Z  | Records W and X           | Records Y and Z         |
| 1        | Records 1–4  | Records 2 and 3           |                         |
| 2        | Records 5–8  | Records 6, 7 and 8        |                         |
| 3        | Records 9–12 | Records 9 and 10          |                         |

- b. Identify the error in the query and describe how it leads to the incorrect data being extracted.

2 marks

---



---



---

- c. State the modification to the query that is required to correct the error.

1 mark

---



---

**Question 12** (6 marks)

Alison is planning the evaluation of her project. She will complete this in two stages. She is aware that she needs to evaluate the infographics solution against evaluation criteria. She is also aware that she needs a strategy to evaluate the infographics after they have been published.

- a. Explain why it is important for Alison to evaluate her infographics solution against the evaluation criteria. 2 marks

---

---

---

---

---

- b. Describe two considerations for Alison to evaluate the infographics after they have been published, as part of her evaluation strategy. 2 marks

Consideration 1 \_\_\_\_\_

---

---

Consideration 2 \_\_\_\_\_

---

---

- c. Alison will assess the effectiveness of her project plan in managing the research project. Explain **one** strategy that Alison could use to do this. 2 marks

---

---

---

---

---

DO NOT WRITE IN THIS AREA



DO NOT WRITE IN THIS AREA

**CONTINUES OVER PAGE**

**Question 13** (9 marks)

Martin is the consultancy firm's IT Manager. He has noticed a few security issues arising since the firm has adopted flexible work arrangements, including the following:

- an increased number of attempts to gain access to the company's servers by unauthorised individuals or groups. Some of these attempts have been successful and the intruders have gained access to the personal details of clients and employees, as well as other sensitive data and confidential information
- staff using their devices for personal use, leading to issues with unauthorised software being installed on their devices
- an increase in the number of phishing and spam emails being sent to employees

Martin has raised these security issues with the firm's management. They have decided to handle the unauthorised access to their servers confidentially, without notifying any stakeholders. Martin has indicated to management that if they do this, they will not be following current legislation. Management have responded by asking Martin what legislation is involved.

- a. Identify the legislation that Martin is referring to. 1 mark

---

- b. Explain why the company would not be complying with the legislation identified in **part a.** if they confidentially handled the unauthorised access to their servers. 3 marks

---

---

---

- c. Martin decides to develop a strategy to improve the company's data and information security practices. He wants to focus on mitigation actions that can be implemented within an appropriate timeframe.

Recommend **two** actions that could form part of an overall strategy to improve the current data and information security practices. 2 marks

---

---

---

---

- d. Justify **one** of your recommended actions from **part c.** by highlighting a relevant possible consequence if the action is not taken in a timely manner.

3 marks

---

---

---

---

DO NOT WRITE IN THIS AREA



## **Insert for Section C – Case study**

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

### **The future of Melbourne’s transport network**

Melbourne currently has a population of 5.15 million people and it is expected to grow to 8.5 million people by 2050. Alison is a researcher who works for a consultancy firm that works with federal and state governments in Australia. She specialises in urban development and transport planning. Over the last few years, she has become increasingly interested in Melbourne’s urban growth and the future of its transport network.

### **Research into Melbourne’s transport network**

Recently, Alison was awarded a research grant by the Victorian Government to investigate the current Victorian transport network. The purpose of this research is for Alison to write an advisory report regarding the perceptions and issues associated with the Victorian transport network. Alison will present her findings to the Victorian Government in six months. Infographics are to be included as part of the advisory report. The Victorian Government will then use this research to help them plan for the future needs of the transport network.

Thomas and Hong Mei have been appointed as project officers to support Alison in her research.

The research will focus on the following:

- individual transport user’s preferences for the following modes of transport: driving, cycling, use of ride-share services/taxis, public transport
- how often transport users use each mode of transport each week
- understanding the issues of Melbourne’s major roads and public transport routes at different times of the day and different times of the year
- the types of vehicles being used, such as bikes, scooters, motorcycles, cars, buses, trains and trams
- perceptions of whether the transport network has improved or deteriorated over the last 10 years
- how transport users think the transport network could be improved

Alison, Thomas and Hong Mei have determined that they will need to collect data through a variety of methods. This will include a survey to be sent out electronically to the owners of all registered vehicles in Victoria (approximately 5.1 million vehicles, according to the Australian Bureau of Statistics Motor Vehicle Census in 2021). Part of this survey is shown in Figure 1 (on page 2 of this insert).

A sample of the data received as part of the survey is contained in Figure 2 (on page 2 of this insert).

**Registered Vehicle Owners Survey**

Date of birth (for age calculations):\*

Gender:\*  ▼

Vehicle manufacturer:  ▼

Vehicle model:  ▼

Vehicle is driven at the following times (on a typical day):

12 am–6 am     12 pm–3 pm     9 pm–12 am

6 am–9 am     3 pm–6 pm

9 am–12 pm     6 pm–9 pm

How could the transport network be improved?\*

**Figure 1 – Part of the survey sent to registered owners**

| <b>DateOfBirth</b> | <b>Gender</b> | <b>V_Man</b> | <b>V_Model</b> | <b>Drive_time</b>       | <b>TN_Improve</b>              |
|--------------------|---------------|--------------|----------------|-------------------------|--------------------------------|
| 2001-09-22         | M             | Balksedan    | Roach          | 6 am–9 am<br>3 pm–6 pm  | More trains,<br>more regularly |
| 26/12/96           | F             | Batata       | GizmodaX       | 12 am–6 am              | NA                             |
| Jan 1, 1978        | X             |              |                | 9 am–12pm<br>12 pm–3 pm | Comfortable<br>seats           |
| 1/2/1982           | F             | Zuvda        |                | 3 pm–6 pm<br>6 pm–9 pm  | Bike-friendly<br>options       |

**Figure 2 – Part of the survey sent to registered owners**

Alison also intends to utilise real-time traffic data sourced from a market-leading mapping software provider, which allows users to access their mapping software via web browsers and their dedicated mobile phone app. The app provides users with maps, the ability to seek directions, and real-time traffic updates. It also collects real-time data from its users (even while the app is not being used) to assist in the visualisation of traffic volume (using red for busy and green for not busy) and determination of the fastest route to a destination.

A sample of the data supplied by the mapping software provider through their data portal has been provided in Figure 3. Note: This data set is simply representative of the millions of records contained within the CSV file. The records for inclusion were selected at random. Figure 4 contains supporting information required to translate the values in the DayOfWeek and MetroZone fields to days of the week and areas in Victoria respectively.

### Current information security and management practices

The consultancy firm that Alison works for has standard security measures in place, largely due to the sensitive work they do with federal and state governments, as well as a number of large businesses around the country.

All employees at the company use company-provided laptops, which require individual usernames and passwords to log in. Passwords are reset every six months and must meet length requirements (minimum 10 characters) and structure requirements (a combination of upper-case and lower-case letters, and at least two numbers and/or non-alphanumeric characters). Users cannot keep their password from the previous month; however, they can use older passwords if they wish.

Recently, the firm adopted flexible work arrangements that enable their employees to work from anywhere they choose, provided they are meeting their work obligations. This has meant that employees are accessing company servers and critical data via their home and mobile internet connections or through public wi-fi networks at places such as libraries, hotels, public workspaces and cafes. It has also meant that the company's IT department is increasingly seeing issues arising from personal use of the company-provided laptops.

| Record number | Year | Month | DayOfWeek | Time period (24-hr time) | MetroArea | MetroZone | Suburb          | Traffic volume (0–10) |
|---------------|------|-------|-----------|--------------------------|-----------|-----------|-----------------|-----------------------|
| 1             | 2016 | Jan.  | 1         | 1100–1200                | True      | E         | Doncaster       | 3.5                   |
| 2             | 2016 | Oct.  | 4         | 0400–0500                | True      | W         | Keilor          | 1.8                   |
| 3             | 2017 | Dec.  | 5         | 0700–0800                | True      | W         | Deer Park       | 6.5                   |
| 4             | 2018 | Jun.  | 6         | 1500–1600                | False     | NA        | Sebastopol      | 4.2                   |
| 5             | 2018 | Aug.  | 7         | 1500–1600                | True      | I         | North Melbourne | 2.1                   |
| 6             | 2019 | Feb.  | 2         | 0300–0400                | True      | I         | Brunswick       | 0.5                   |
| 7             | 2019 | Mar.  | 4         | 1200–1300                | True      | N         | Epping          | 5.7                   |
| 8             | 2020 | Jan.  | 5         | 1400–1500                | True      | SE        | Frankston       | 5.3                   |
| 9             | 2020 | Nov.  | 3         | 0800–0900                | True      | SW        | Werribee        | 7.1                   |
| 10            | 2021 | Jul.  | 2         | 1700–1800                | True      | N         | Strathmore      | 8.1                   |
| 11            | 2021 | Sept. | 1         | 2000–2100                | True      | I         | Carlton         | 4.2                   |
| 12            | 2022 | Jan.  | 7         | 1800–1900                | False     | NA        | Birchip         | 2.1                   |

Figure 3 – Sample from the CSV file downloaded from the mapping software provider's data portal

| Legend – DayOfWeek |           | Legend – MetroZone |               |
|--------------------|-----------|--------------------|---------------|
| Value              | Day       | Value              | Zone          |
| 1                  | Sunday    | E                  | East          |
| 2                  | Monday    | I                  | Inner         |
| 3                  | Tuesday   | N                  | North         |
| 4                  | Wednesday | NA                 | Regional Vic. |
| 5                  | Thursday  | SE                 | South-East    |
| 6                  | Friday    | SW                 | South-West    |
| 7                  | Saturday  | W                  | West          |

Figure 4 – Supporting information for the DayOfWeek and MetroZone fields in the CSV