**Victorian Certificate of Education
2020**

SUPERVISOR TO ATTACH PROCESSING LABEL HERE

Letter

STUDENT NUMBER

APPLIED COMPUTING: DATA ANALYTICS
Written examination

Wednesday 25 November 2020

**Reading time: 11.45 am to 12.00 noon (15 minutes)
Writing time: 12.00 noon to 2.00 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	5	5	20
C	15	15	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 25 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.

SECTION A – Multiple-choice questions

<div style="border">

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.

</div>

Question 1

Selected data from large repositories can be collected and saved as a file.

These files can be imported into a database if the files are in which format?

A.   .pdf

B.   .csv

C.   .psd

D.   .jpeg

Question 2

Ginny and Alejandro are discussing the best way to store data and information for their DJ business. They decide to utilise network attached storage (NAS).

Which one of the following comments about NAS is correct?

A.   Ginny states that a NAS device is cost-effective, but they will have to purchase extra space for individual users.

B.   Alejandro states that a NAS device will be affected by the failures of local hardware, which might compromise the data stored on the local hardware.

C.   Ginny states that a NAS device will add enough storage space to their local computer and Alejandro states that it will boost the available storage space.

D.   Ginny states that all documents will be available from a central location that can be easily accessed by anyone, but they will still have to email encrypted documents in order to collaborate with their team.

Question 3

When Mohammed tries to log in to his notebook computer, the following appears on his screen.

Click here to pay

Your files have been locked!

To regain access to your files you will need to click the
button and pay the required amount with Bitcoin.
We will then release your files.
If you do not pay the price within 24 hours, all of your
files will be deleted.

Which one of the following would have been the best way for Mohammed to have avoided this threat?

A. always using well-encrypted passwords

B. not downloading files from unreliable websites

C. logging out of his notebook computer after each session

D. not allowing cookies to be stored by a website's browser

Question 4

Bruce has undertaken gender inclusiveness training to ensure that he avoids alienating audiences when creating effective infographics and dynamic data visualisations.

To ensure there is gender neutrality in his work, which one of the following lists of terms should Bruce use?

A. personnel, actress, workers

B. manpower, workers, man-hours

C. workers, personnel, representative

D. spokeswoman, spokesman, representative

Question 5

A dynamic data visualisation solution should enable users to

A. interact with the data visualisation.

B. annotate the data visualisation.

C. create worksheets.

D. create tables.

Question 6

A feature of a good research question is that it must be able to be

A. proven.

B. refuted.

C. supported.

D. supported or refuted.

Question 7

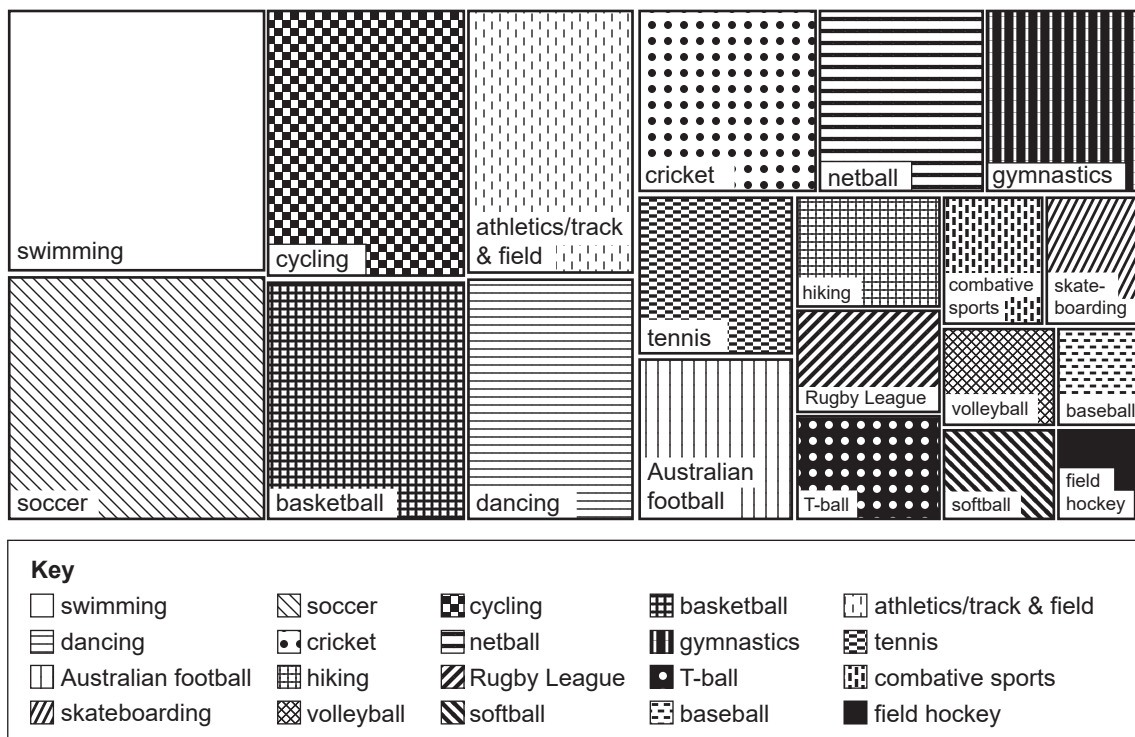Colour choice is important when creating an infographic or a dynamic data visualisation.

Inappropriate colour choice can

A. influence an audience's character.

B. distract the audience from the message.

C. improve clarity between the text and the background.

D. highlight the most important points of a presentation.

Question 8

A survey of sports participation rates among children aged six to 13 has been prepared as a data visualisation, shown below.

Sports participation rates, ages 6–13 (2018)

Data: adapted from Roy Morgan Research, 'Top 20 sports led by walking, swimming and gym training', 13 May 2018, <www.roymorgan.com>

Which one of the following is the best reason to select this type of data visualisation for sports participation rates?

A. to find out how many students were surveyed

B. to quickly find out the most and least popular sports

C. to understand why certain sports are popular with children

D. to evaluate whether summer or winter sports are more popular

Use the following information to answer Questions 9 and 10.

Tilde runs a small, local animal rescue shelter. She wants to expand her services to include all of Victoria, but first she needs to find out more about the animals that need help as well as more about potential pet owners.

Question 9

Tilde wants to download data about pet ownership preferences so that she can create a histogram that will be part of an infographic about rescued pets. Before she begins this process, she should check the suitability of the data she intends to download.

Suitable data would be

A.   visual.

B.   qualitative.

C.   quantitative.

D.   an infographic.

Question 10

Some of the data that Tilde downloads includes text-based descriptions of pet owners' experiences with animals.

One approach to coding this data could be to

A.   create an infographic.

B.   create a dynamic data visualisation.

C.   create categories and count the references to each category.

D.   download data that is exclusively numeric (no text) to create better charts.

Question 11

What is the purpose of software security controls?

A.   to protect data from authorised access

B.   to allow for access restrictions and encryption

C.   to prevent biometric security from functioning

D.   to stop users from completing surveys multiple times

Question 12

Under the *Privacy and Data Protection Act 2014*, when a Victorian Government organisation wants to dispose of personal information that is no longer needed, the organisation must

A.   take reasonable steps to destroy or permanently de-identify the personal information.

B.   refuse to destroy or permanently de-identify the personal information.

C.   provide the personal information to organisations upon request.

D.   sell the personal information to the highest bidder.

Question 13

A new dance company, InDanz Studio, has been preparing for its virtual performances that will be held online at the end of 2020. InDanz Studio does all of its scheduling on spreadsheets. The dance company has spreadsheets for performances, competitions, classes and enrolments. Lately, InDanz Studio has noticed issues with its spreadsheets. Dates and times have been incorrect and sometimes missing. The dance company does not know why.

Which of the following characteristics of data integrity have an impact on the use of spreadsheets by InDanz Studio?

A.   the relevance and timeliness of data

B.   the clarity and completeness of data

C.   the accuracy and authenticity of data

D.   the accuracy and completeness of data

Use the following information to answer Questions 14 and 15.

A national sporting league manages the sale of tickets for its annual awards night by writing the sale information in a book. Members provide their membership number and the number of tickets that they wish to purchase.

Question 14

The manager suggests that the data be stored in tables and that these tables be properly designed using a data dictionary.

The purpose of a data dictionary is to

A. describe all properties of each of the table's fields.

B. show how tables are linked to each other.

C. depict the colours and fonts used.

D. describe the queries to be used.

Question 15

The reception hall for the awards night has tables that seat at most eight people. The organising committee has chosen to allocate seats in stages, initially selecting members who have purchased six or more tickets.

ReceiptNo.	Date	MembershipNo.	NumberOfTickets	Total ($)	TableAllocation
001	29/10/2020	35	4	400	
002	1/11/2020	4	5	500	
003	4/11/2020	56	1	100	
004	4/11/2020	83	2	200	
005	8/11/2020	3	9	900	
006	11/11/2020	12	3	300	
007	11/11/2020	11	8	800	
008	11/11/2020	8	3	300	

Which one of the following query rules will generate a list of members who will be allocated to tables in the initial stage?

A. NumberOfTickets $<= 8$

B. NumberOfTickets $>= 6$

C. NumberOfTickets $>= 6$ OR NumberOfTickets $<= 8$

D. NumberOfTickets $>= 6$ AND NumberOfTickets $<= 8$

Question 16

Evaluating strategies is a critical step during the evaluation stage of a project.

Which one of the following statements about evaluation is true?

A. Evaluation criteria are developed during the design stage of the problem-solving methodology and used again in the evaluation stage.

B. Evaluation criteria are developed during the evaluation stage of the problem-solving methodology.

C. Evaluation refers to checking the accuracy of the output.

D. Evaluation is the same as testing a solution.

Question 17

Which of the following is the most important consideration when selecting interviews as a data collection method?

A. how timely the data will be

B. how well the participants speak English

C. the relevant formats and conventions for collecting the data

D. the time available for questions and the need for follow-up questions

Question 18

Data that has integrity has which of the following characteristics?

A. accuracy, relevance, timeliness

B. relevance, reliability, validation

C. authenticity, correctness, readability

D. reasonableness, timeliness, validation

Question 19

The most appropriate design tool for representing tables in a relational database is

A. a mock-up.

B. brainstorming.

C. a data dictionary.

D. an annotated data structure.

Question 20

Jayde's research on footpath safety in icy conditions will be presented at an international conference and it is important that she is able to present her findings clearly.

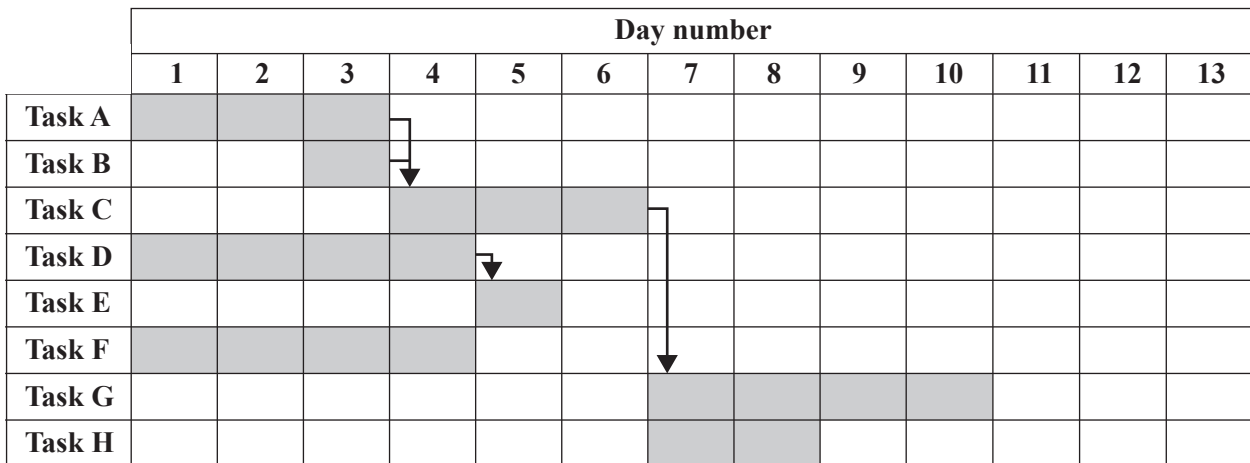Which one of the following is the most important consideration for this presentation?

A. The order of her findings should be random.

B. Only the basic features of presentation software should be used.

C. All research should be included in the order in which it will be accessed.

D. A combination of charts, images and statistics should be used to reveal trends in her findings.

END OF SECTION A

SECTION B – Short-answer questions

<table>
<tr><td>Instructions for Section B</td></tr>
<tr><td>Answer all questions in the spaces provided.</td></tr>
</table>

Question 1 (5 marks)

Consider the Gantt chart below.

						Day number							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Task A	▓	▓	▓										
Task B			▓										
Task C				▓	▓	▓							
Task D	▓	▓	▓										
Task E					▓								
Task F	▓	▓	▓	▓									
Task G							▓	▓	▓	▓			
Task H							▓	▓					

a. Identify **one** feature that is missing from the Gantt chart above.   1 mark

_____

b. List the tasks that are on the critical path in the Gantt chart above.   1 mark

_____

c. Modify the Gantt chart above to show the impact on the project's critical path if Task C takes
two days longer than planned.   1 mark

d. Explain what would happen to the project's critical path if Task C takes two days longer than
planned.   2 marks

_____

_____

_____

_____

Question 2 (5 marks)

Meals by Delivery is a meal delivery service. Customers can use a mobile phone app to order meals, for dinner only, from fine-dining restaurants located in districts around the town hall of any city or town that Meals by Delivery services. The app connects to Meals by Delivery's central database. Currently, Meals by Delivery is restricting deliveries to customers who live within a 3 km radius of a town hall.

Meals by Delivery's sales manager has suggested that the business expand its operations to include deliveries within a 10 km radius, as he believes there are many more families that would enjoy fine-dining meals and would pay for the delivery service.

The sales manager has commissioned a marketing company to gather data from residents about their eating habits when it comes to meal deliveries in order to determine if expanding the company's operations would be worthwhile. This information will be presented as an infographic with data visualisations representing customer and family eating habits.

a. Describe a method that the marketing company could use to gather data to identify constraints for the data visualisations. Suggest why this method would be suitable for this task. 2 marks
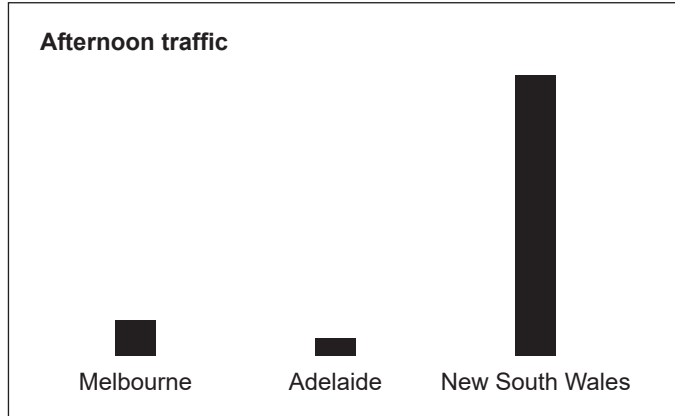
b. Identify **one** example of a constraint for the data visualisations. 1 mark

c. Suggest how the marketing company could generate a list of the most important non-functional requirements for the data visualisations based on the data it gathers from both residents and Meals by Delivery. 2 marks

Question 3 (3 marks)

Amina developed a data visualisation highlighting road congestion in Melbourne, Adelaide and Sydney over a five-year period using her own research. When she presented her findings to a panel, the panel asked her to check her work and verify the data in her presentation. Below is a slide from her presentation on afternoon traffic in Australian cities.

a. Consider the data visualisation presented in the slide above. Give **one** reason why the panel would have asked Amina to check her work and verify the data. 1 mark

_____

_____

b. When Amina checked her work, she discovered issues involving the use of the data in the data visualisation.

Describe **one** technique that Amina could have used to verify the data before developing the data visualisation. 2 marks

_____

_____

_____

_____

Question 4 (4 marks)

Arnis manages a water conservation organisation and is conducting research about Australia's growing drought problem.

a. Outline two reasons why Arnis might want to search for existing data sets from government organisations, such as the Australian Bureau of Statistics (ABS) or the Commonwealth Scientific and Industrial Research Organisation (CSIRO). 2 marks

1. _____

_____

2. _____

_____

b. Explain why acquiring this government data would be important to Arnis's organisation. 2 marks

_____

_____

_____

Question 5 (3 marks)

Paloma is creating an infographic for children, based on her findings about water safety. Her infographic will include several data visualisations based on the research she has conducted, which includes statistics for safety issues recorded at all Victorian beaches and popular swimming venues, including rivers, lakes, public swimming pools and swimming pools at home. This data will include the time at which the safety issue occurred, the year, whether first aid was rendered and the outcome.

As she has been researching this topic for the past four years as part of her university course, Paloma has collected large amounts of data, which she has analysed using a spreadsheet. Her work is heavily text-based and directed at primary and secondary students in schools, to promote safe swimming.

Paloma needs assistance with creating an infographic to present her findings to her intended audience.

Describe **three** types of information, other than text, that should be included in the infographic to educate Paloma's intended audience.

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 (2 marks)

Write a research question that will help Evan present his findings about young travellers.

_____

_____

_____

_____

Question 2 (3 marks)

Huang does not understand why there is a need to determine solution requirements, constraints and scope in order to present findings.

Evan explains to him that doing so will help ensure that Huang collects and analyses the most suitable data.

Complete the table below with the correct category for each element given using the following list:

- non-functional requirement
- constraint
- scope

Element	Category
Only data regarding young travellers is to be used.	
Information displayed should be easy to read.	
Only data from 2015–2019 is to be used.	

Question 3 (3 marks)

Evan will need to collect both primary data and secondary data from a range of sources. He has decided to collect primary data by sending out surveys to tour operators around Australia. The data collected will be about young travellers: their age, country of residence, city, postcode/zipcode, usual travel period, destinations visited, usual accommodation and average amount spent.

a.   State **one** advantage of using surveys as a method of collecting data.   1 mark

_____

_____

b.   Suggest **one** technique for collecting secondary data and explain why this would be useful to
        Huang and Lucinda.   1 mark

_____

_____

_____

c.   Evan wants to ensure that he references the data that he collects.

        Identify a referencing system for Evan to use.   1 mark

_____

Question 4 (4 marks)

Evan has found a study from 2010 about the number of business travellers to Melbourne. The study includes information on landmarks, places of interest such as museums, galleries and theatres, restaurants, healthcare services and sports facilities.

Evan has suggested to Lucinda that this would be a good starting point for identifying the travel needs of the young travellers they wish to attract to their hotel. However, Lucinda has questioned the integrity of this data.

a. Identify **two** reasons why Lucinda has concerns regarding the integrity of the data in the study that Evan has found.   2 marks

_____

_____

_____

_____

b. Recommend **two** techniques for improving data integrity when searching for and analysing data.   2 marks

_____

_____

_____

_____

_____

Question 5 (3 marks)

Evan thinks that the findings should be presented as an infographic rather than a dynamic data visualisation.

Justify why an infographic would be a more suitable way of presenting his findings on young travellers than a dynamic data visualisation.

_____

_____

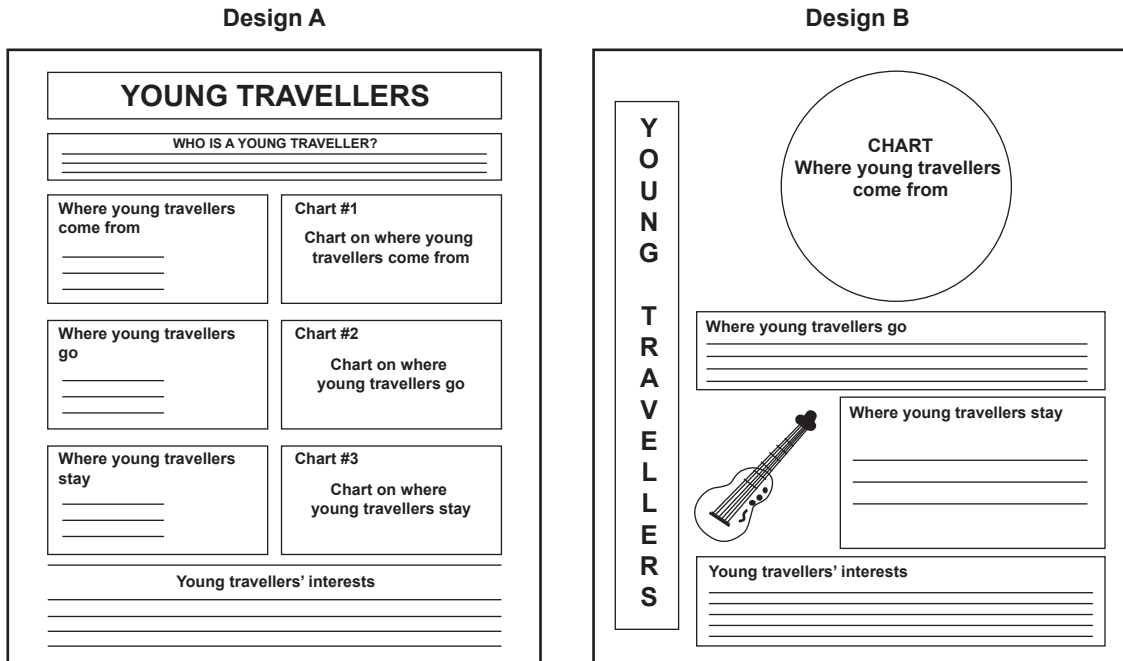_____

_____

_____

Question 6 (2 marks)

Huang found a few images and charts while searching for suitable data sets with Evan. He planned to use these in the infographic until Lucinda pointed out that it would be illegal to do so.

a.   Identify the legal issue that Huang and Lucinda would face if the images and charts were used.   1 mark

_____

_____

_____

b.   If Huang and Lucinda decide to use the images and charts, what will they need to do to ensure they meet legal requirements?   1 mark

_____

_____

_____

Question 7 (8 marks)

Evan has worked closely with Huang and Lucinda to develop two design ideas for the infographic. These ideas were drawn as mock-ups, shown below, and labelled Design A and Design B.

a. Explain why Evan would draw two design ideas for Huang and Lucinda instead of one. 2 marks

b. Suggest **one** criterion Huang and Lucinda could use to determine their preferred design. 1 mark

c.   Which of the two designs – Design A or Design B – should Huang and Lucinda choose?
Justify your response by referring to the criterion suggested in **part b.**   3 marks

_____

_____

_____

_____

_____

_____

d.   When developing the preferred design, Evan will need to follow design principles relating to the appearance of the infographic.

Identify two design principles relating to the appearance of the infographic.   2 marks

Design principle 1 _____

Design principle 2 _____

Question 8 (10 marks)

Some local tour operators have offered to supply data from their past bookings. Evan has decided to import the tour operators' data from the tour operators' files, as well as data from other files, into a database in order to identify any patterns, and then into a spreadsheet for further refinement and the development of charts. After this, he will then develop the infographic.

a. Identify two software functions that would be suitable for creating the database and describe how each will be used to develop the database. 4 marks

Software function 1 _____

Description _____

_____

_____

Software function 2 _____

Description _____

_____

_____

b. Evan exports the data from the database into a spreadsheet. He notices that there is a lot of data in rows and columns, and he wants to make the spreadsheet more readable by applying formats and conventions.

Identify and outline one format and one convention that could be applied to the spreadsheet. 2 marks

Format _____

_____

_____

Convention _____

_____

_____

c.   As Evan looks more closely at the data before manipulating it to produce charts for the infographic, he notices that quite a few cells are blank.

Identify the validation technique that Evan should have used to avoid having blank cells. Describe how this technique would avoid the issue of blank cells.   2 marks

Validation technique _____

Description _____

_____

_____

d.   Describe a test that Evan could perform to ensure that another validation rule will restrict the ages in the data to between 18 and 25.   2 marks

_____

_____

_____

_____

Question 9 (4 marks)

Evan wants to manipulate the data in his spreadsheet so that he can add it to the infographic. He wants to show the percentages for the top five responses to 'where young travellers come from', 'where young travellers go' and 'where young travellers stay'. Charts will then be produced to show this information visually.

Outline two techniques that Evan could use to efficiently and effectively manipulate the data for the infographic.

Technique 1 _____

_____

_____

_____

Technique 2 _____

_____

_____

_____

Question 10 (3 marks)

Evan has now completed the infographic and is ready to test it.

Describe how the infographic should be tested.

_____

_____

_____

_____

_____

SECTION C – continued

Question 11 (2 marks)

Outline a strategy for evaluating the effectiveness of the features of the infographic.

_____

_____

_____

_____

_____

_____

Question 12 (3 marks)

Lucinda has recently become aware that the hotel has been collecting and storing a lot of customer data over the years. She thinks some of this data may be useful to their work with Evan and she places the data into separate files according to the year in which it was collected.

A few weeks later, the police inform Lucinda and Huang that this customer data, including names, addresses, phone numbers and credit card details, has been published on a website. The police also inform them that they believe the customer data was accessed by a guest using the hotel's free wi-fi.

a.   Identify the legislation that has been breached by Notrow's Hotel.   1 mark

_____

b.   Outline a possible legal consequence for Notrow's Hotel.   2 marks

_____

_____

_____

_____

Question 13 (4 marks)

Evan explains to Huang and Lucinda that the issue involving unauthorised access to customer data is a result of a deliberate threat to the security of the hotel's data and information.

Upon closer investigation of the hotel's network security, Evan discovers that:

- guests can access the hotel's wi-fi without a password
- staff do not need to have a username or password to access the system
- there is no firewall
- the anti-virus software has not been updated for two years.

Name a possible type of deliberate threat to the hotel's data and information, and suggest a strategy to improve the hotel's current practices to protect against this type of deliberate threat.

Deliberate threat _____

Strategy _____

_____

_____

_____

_____

Question 14 (4 marks)

Recommend two software security controls and describe how each will improve the hotel's data and information security practices.

Software security control	Description of improved security practice

Question 15 (5 marks)

Evan realises that the backup procedures at Notrow's Hotel do not adequately protect the organisation's data and information. Upon further investigation he discovers that the hotel does not have a disaster recovery plan. Lucinda is very concerned about this.

a. Apart from having a backup strategy, identify any **three** components of a disaster
    recovery plan.   3 marks

b. Recommend an appropriate backup strategy that Notrow's Hotel could implement to protect
    its data and information.   2 marks

END OF QUESTION AND ANSWER BOOK

Insert for Section C – Case study

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

THIS PAGE IS BLANK

Notrow's Hotel

Huang and Lucinda own and run a boutique hotel called Notrow's Hotel, located along a beach on the south coast of regional Victoria. For many years they have attracted and served an older clientele, mostly from Victoria. Business from this market has been decreasing over the last few years, from $7 million a few years ago to $4 million last year.
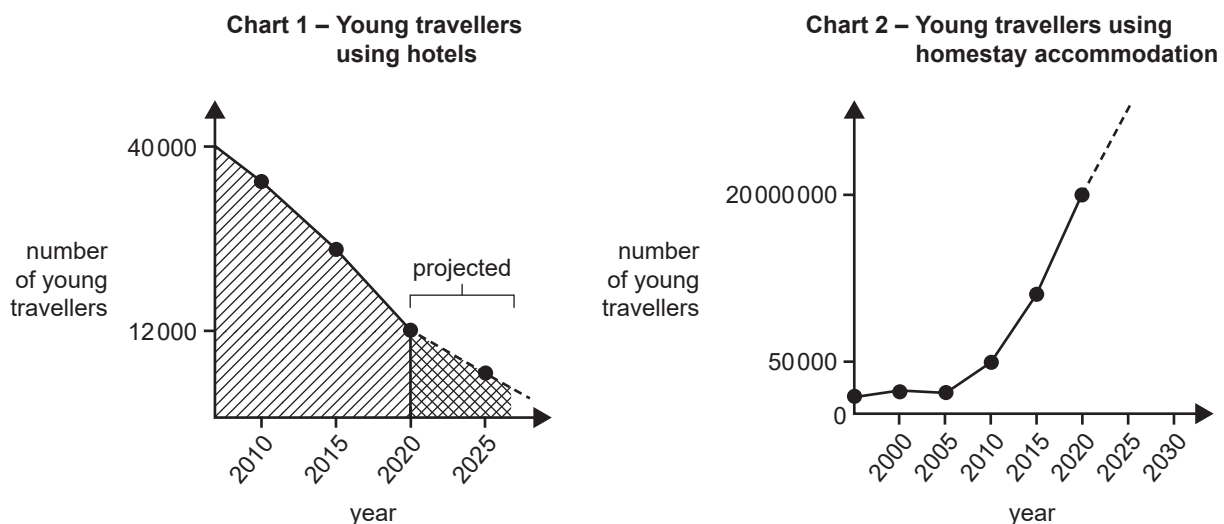
Huang and Lucinda have heard that young travellers aged 18–25 are passionate about experiencing new destinations and require accommodation to suit their needs. They have heard that this group of travellers enjoys travelling to new locations, tends to stay in homestay accommodation and tends to visit local places of interest. They also enjoy taking local tours.

Attracting more young travellers

Huang and Lucinda feel that they will have to reconsider their marketing strategies in order to attract these young travellers. First, they need to understand the needs of these travellers.

Lucinda has recently read a newspaper article about the global homestay accommodation market, which stated that the market for homestay accommodation has grown from just under 50 000 young travellers in 2009 to nearly 20 million in 2019. These statistics have captured Huang and Lucinda's attention and they have decided to seek the assistance of Evan, a project manager for a data and security consultancy firm.

Evan shows them an online article called 'TravelStyle Mate' that was published in 2019. The article is about young travellers and their travelling habits and plans. Of particular interest are two partial charts, Chart 1 and Chart 2, shown below.

Evan explains the importance of this data and information to Huang and Lucinda. They know that this data will help them develop future opportunities for their hotel. Accessing this data could give Huang and Lucinda a competitive advantage in their region.

Huang also expresses his concerns to Evan regarding their network security. He is aware of increased threats in 2020 and wants to protect their data and information. His current security strategies are inadequate.

Evan suggests that he could manage a project for Huang and Lucinda that would involve:
- framing a research question around young travellers' travel habits
- documenting solution requirements
- accessing data sets regarding young travellers, including their travel habits and interests
- developing and evaluating either infographics or dynamic data visualisations to present findings
- investigating the hotel's current data and information security strategies.

END OF INSERT