

STUDENT NUMBER Letter

APPLIED COMPUTING: DATA ANALYTICS

Written examination

Tuesday 31 October 2023

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	4	4	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 23 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.

Use the following information to answer Questions 1 and 2.

The owner of a new luxury cat boarding house, where cat owners leave their cats while on holiday, is setting up a customer database.

Question 1

The most appropriate data type for a mobile telephone number field would be

- A. text.
- B. integer.
- C. Boolean.
- D. numeric.

Question 2

A field in the database contains data about a cat's birthday.

Which is the most appropriate field name to use?

- A. catBirthday
- B. customerBirthday
- C. allTheCatsBirthday
- D. birthday_exact_estimated

Question 3

Jace is designing a spreadsheet to record data about the different types of flowers in his suburb. He is not worried about the appearance of the spreadsheet; however, he is focused on the functionality.

The characteristics of the spreadsheet should include

- A. usability and appearance.
- B. repetition, colour, proportion, text hierarchy, text style and titles.
- C. robustness, flexibility, ease of use, navigation and error tolerance.
- D. navigation, error tolerance, repetition, colour, robustness and ease of use.

Question 4

The owner of a European car sales company is looking for an effective way to display a list of all cars made after 2020 from either Italy or Germany in their database.

This query would be written as

- A. Year < 2020 AND (Country = 'Italy' OR Country = 'Germany').
- B. Year < 2020 AND (Country = 'Italy' AND Country = 'Germany').
- C. Year > 2020 AND (Country = 'Italy' OR Country = 'Germany').
- D. Year >= 2020 AND (Country = 'Italy' AND Country = 'Germany').

Question 5

Examples of digital system components include

- A. a phone, printer and data.
- B. a touch screen, software and a printer.
- C. a virtual reality (VR) headset, hardware and formats.
- D. a smart speaker, people and processes.

Question 6

Consider a dynamic data visualisation solution aimed at a particular audience where their level of expertise needs to be considered.

A typical constraint on the solution will be

- A. social.
- B. usability.
- C. technical.
- D. age appropriateness.

Question 7

A school was concerned about their students' poor academic results and was proposing a research question to investigate. School leaders had initial ideas about how the time spent by students playing video games impacted their sleep.

Which of the following is the best research question to pursue?

- A. Are video games and lack of sleep bad for study results?
- B. What is the impact of playing video games on academic results?
- C. Do students who play video games at night and have reduced sleep get lower academic results?
- D. Can students who play console games achieve higher results than those who play online computer games?

Question 8

Qualitative data is generally

- A. expressed in words, uses closed questions and is collected on a large scale.
- B. expressed numerically, uses open questions and is collected on a large scale.
- C. expressed in words, involves opinion and is collected on a small scale.
- D. expressed numerically, involves opinion and is collected on a small scale.

Question 9

Thanos has been collecting data from her class about the hours that they play online games on the weekend, and she records this in a spreadsheet. The spreadsheet calculates the total hours played on the weekend; however, the spreadsheet does not have data for two members of the class.

Which aspect of data integrity has been affected?

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

Question 10

Which of the following criteria could be used to evaluate both the efficiency and the effectiveness of a dynamic data visualisation?

- A. Does the data visualisation load quickly and is it accessible to users?
- B. Does the data visualisation present accurate information and is it readable to the audience?
- C. Is the information in the data visualisation complete and is the message communicated clearly?
- D. Does the data visualisation take little effort for users to understand, and does it cost less to produce?

Question 11

A team working on a project is having issues finding the correct versions of files that they should be working with, due to inconsistency about where the files have been stored.

What is an appropriate technique that the team could implement to resolve this issue?

- A. Files are archived regularly.
- B. A backup copy is made of current working files.
- C. A file management plan is created and implemented.
- D. Files are protected from unauthorised users, using two-factor authentication.

Question 12

A medical company is developing a dynamic data visualisation, analysing the impact of influenza vaccinations on the spread of influenza during the past three years. The company wants to compare trends in Australia and neighbouring Oceanic countries.

What functionality of the dynamic data visualisation software would best support the most efficient use of the data visualisation?

- A. check boxes to select countries and years
- B. live and continually updating data sources
- C. all charts and information displayed on one interface
- D. a dashboard for each country connected through navigational buttons and a menu

Question 13

Quinn has developed an online infographic to summarise the survey data and the findings of his research question.

What feature could he apply to text elements to communicate his results more effectively to the target audience?

- A. Replace text with uniquely designed icons.
- B. Provide a hyperlink to the complete data set.
- C. Use more informal and familiar language to engage the audience.
- D. Provide a hyperlink to a glossary of terms to explain acronyms.

Question 14

Maria has a project focused on researching the impact of income and population on the life expectancy of people in a country. She wants to display the relationship between these three factors.

Which type of chart should Maria use to create an effective data visualisation?

- A. a scatter chart
- B. a bubble chart
- C. a 3D bar chart
- D. a stacked area chart

Question 15

The security of data files stored on a LAN can be improved by

- A. using SSL encryption.
- B. implementing audit trails.
- C. using virus protection software.
- D. restricting access through user permissions.

Question 16

As a member of the student representative council (SRC), Roberto was asked to survey the student body about their attitudes to the school's sports uniforms and if any changes need to be made to the uniforms. The results of the survey will be presented to the principal as a dynamic data visualisation at the next SRC meeting.

Roberto has finished his dynamic data visualisation and he would like to get student feedback on whether it meets the needs of the student body.

Which method would be the most efficient way for Roberto to seek feedback about his dynamic data visualisation?

- A. Interview users about the dynamic data visualisation.
- B. Observe how users interact with the dynamic data visualisation.
- C. Provide an email link to an online survey with open-ended questions about the solution.
- D. Send all students an email including an online survey that uses Likert scale questions.

Question 17

Which of the following is most commonly included in a disaster recovery plan?

- A. ethical issues
- B. a security plan
- C. a strategic plan
- D. evacuation procedures

Question 18

There was a lightning strike to the building containing a company's network servers, which caused all data to be lost.

This is an example of

- A. a cyber threat.
- B. a deliberate threat.
- C. an accidental threat.
- D. an event-based threat.

Question 19

Which of the following could result from a lack of information security strategies for an organisation?

- A. easier retrieval of information
- B. financial loss and legal consequences
- C. advantage for the organisation over other companies
- D. operational disruptions and an increase in customer satisfaction

Question 20

Physical security controls for preventing unauthorised access to data include

- A. biometric locks.
- B. patching systems.
- C. segmenting networks.
- D. enforcing least privilege access.

DO NOT WRITE IN THIS AREA

**END OF SECTION A
TURN OVER**

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

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

Question 1 (4 marks)

DelvrStuff is a new app that allows people to pay for and receive delivery of furniture between sellers and buyers for a small cost.

DelvrStuff allows the sellers to list items of furniture they want to sell, and it requires them to register their name, email address and their mobile phone number to become a seller.

DelvrStuff lets buyers look for items of furniture, by searching the items' names, their descriptions or their sale prices. Once a buyer has found an item, they can buy it and select the date for delivery. The app then provides the buyer with the set delivery fee amount, the total cost for the transaction, and the date it was bought. The buyer must register to use DelvrStuff by providing their name and delivery address details, including street address, suburb and state.

Draw a normalised design, to third normal form (3NF), to represent the database structure for this app.

Question 2 (6 marks)

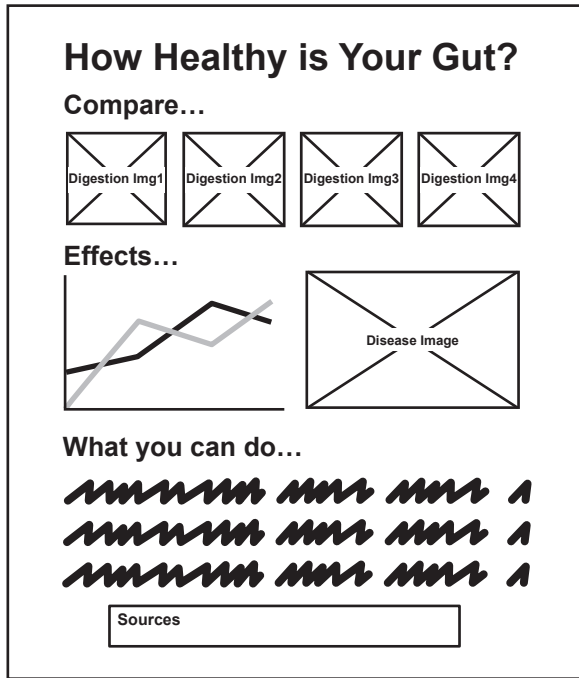
An organisation that promotes healthy eating feels that there is not enough knowledge about the health issues of elderly people related to poor diet and digestion. After collecting and analysing data, the organisation has decided to develop an infographic aimed at educating the elderly about digestive health.

- a. Outline a technique the organisation could use to generate design ideas for the infographic. 1 mark

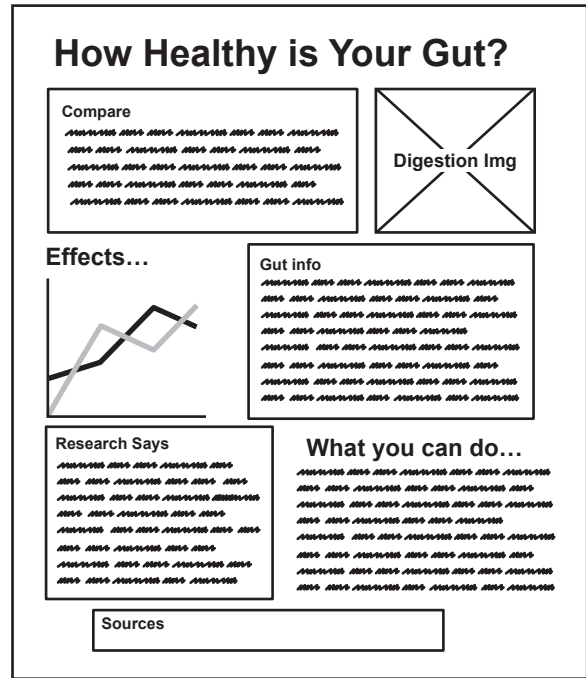
- b. The organisation will need to consider design principles when designing the infographic. Identify and describe **one** design principle related to appearance that is appropriate for the target audience. 2 marks

Two alternative design ideas have been created for the infographic.

Design A



Design B



One criterion for evaluating the alternative designs is stated below:

- The infographic should be able to communicate the intended message to the target audience.
- c. Based on the criterion above, justify which design idea, A or B, would be more appropriate. 3 marks

DO NOT WRITE IN THIS AREA

Question 3 (4 marks)

Caleb has completed his dynamic data visualisation on greenhouse gas emissions. He needs to test his visualisation for navigation and error tolerance of the interface.

- a. Explain the difference between navigation and error tolerance. 2 marks

- b. Provide **one** suitable technique for testing navigation. 1 mark

- c. Provide **one** suitable technique for testing error tolerance. 1 mark

Question 4 (6 marks)

Marina and Parth run an online business that allows people to buy and sell old sneakers and clothes. The data from their business is stored in a database that is hosted on a server in their office. A full backup of their business data is performed at the start of each month, and a differential backup is performed each week on Friday night. All backups are stored on an external hard disk drive (HDD) located in the office. The backups are regularly tested, and all backups are created successfully.

- a. If the server had a hard drive failure, what is the maximum number of days of data that will be lost if the database needs to be restored from backup? Explain your answer. 2 marks

- b. Suggest and justify **two** changes to the backing up strategy to improve data protection and the ability to recover from a disaster. 4 marks

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 (4 marks)

- a.** Identify two items that should be included in the scope for the project. 2 marks

Item 1 _____

Item 2 _____

- b.** Identify two constraints that Ali and his team will have to consider when developing their dynamic data visualisation. 2 marks

Constraint 1 _____

Constraint 2 _____

Question 2 (7 marks)

a. Ali’s team have identified 10 tasks to be completed as part of the project.

Use the data provided in Table 1 in the case study insert and complete the Gantt chart for the project, showing all task durations, dependencies and milestones.

4 marks

		November															
		Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue
		13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Task no	Task																
1	Interview community leaders																
2	Collate and analyse interview results																
3	Develop community survey																
4	Online survey available to communities																
5	Analyse online community survey results																
6	Primary data complete																
7	Complete secondary research into population growth, crop and weather data																
8	Analyse secondary data																
9	Secondary data complete																
10	Hand over to design team																

DO NOT WRITE IN THIS AREA

- b. List the tasks that are not included on the critical path. 1 mark

- c. Explain why these tasks are not included on the critical path. 2 marks

DO NOT WRITE IN THIS AREA

Question 4 (3 marks)

The team used the online community survey to gather data, and a sample of responses is shown below. Ali has told the team that the data is not suitable for use in the dynamic data visualisation. They will need to codify the responses.

Comment on a new grain production facility in your community area
I love it. It would be great for jobs.
I would hate the extra traffic. So, NO!
No, thank you – I wouldn't like the noise and pollution in the area.
More employment for the community would be great.
I don't want bad smells in the air, so I am against it.
Will all the grain processed be from local farmers, so they get the money?
If there are many jobs created, I am all for it.
There are already too many cars and trucks. So, I would not like it.
I am concerned about the river being affected, but the area could do with investment for jobs and infrastructure.
All for it, if farmers get more money?
I would like more jobs in the community, but what about the environment?
Yes – new facilities in the area would be great.

- a. Identify **one** reason why the responses need to be codified. 1 mark

- b. Explain the technique the team could use to codify the data so that it is suitable for the dynamic data visualisation. 2 marks

Question 5 (4 marks)

Grains4Food wants to provide farmers with temperature and rainfall sensors to collect data from the farms. The farmer will upload the data from these sensors onto a computer on the farm's network and then send the data to Grains4Food using their internet connection.

Identify and explain the purpose of two network components that the farmers need to have so they can send the temperature and rainfall data to Grains4Food.

Network component 1 _____

Purpose of network component 1 _____

Network component 2 _____

Purpose of network component 2 _____

Question 6 (4 marks)

The primary data collected at the farms needs to be securely sent back to Grains4Food.

- a. Explain how a software security control can be used to protect the data as it is sent. 2 marks

- b. The collected data will be stored on a 'network attached storage' device at Grains4Food. Outline one advantage and one disadvantage of using a 'network attached storage' device. 2 marks

Advantage _____

Disadvantage _____

Question 7 (2 marks)

A member of Ali's team has found an online blog post by an unnamed user, called 'The Importance of Grain to all Aussies', on the website 'Grains Rule the World'. Justify whether Ali's research team should use this information in the dynamic data visualisation.

Question 8 (1 mark)

The research team must reference all sources they use according to the American Psychological Association (APA) referencing standard.

Give **one** reason why sources should be referenced.

DO NOT WRITE IN THIS AREA

Question 9 (4 marks)

Ali thinks that collecting data through a Geographic Information System (GIS) is a good technique for efficient and effective data collection for each of the regions.

- a. What are **two** components that could make up a GIS data set? 2 marks

- b. Explain how GIS data collected in each of the regions can be used to assist Ali in preparing an effective visualisation. 2 marks

Question 10 (3 marks)

The team needs to select a design tool to represent the designs for the development of the dynamic data visualisation.

List the most appropriate design tool for the following components of the dynamic data visualisation.

Project component	Design tool
Identifying and naming the required data for the database	
Spreadsheet formulas for data manipulation	
Dynamic data visualisation appearance	

Question 11 (2 marks)

Ali has been monitoring the progress of the project since it began. Explain a technique that can be applied to the Gantt chart to record this progress.

Technique _____

Explanation _____

Question 12 (7 marks)

The team will manipulate the data collected from the weather bureau using spreadsheet software. All the data for the years that are being included are required in the spreadsheet with no blank cells. A sample of a section of this spreadsheet is shown in Table 2 in the case study.

- a. Identify and describe two techniques that Ali could use to validate the data in the spreadsheet. 4 marks

Data validation technique 1 _____

Data validation technique 2 _____

- b. Describe **one** technique to test the functioning of the spreadsheet. 3 marks

Question 13 (6 marks)

Ali imports the prepared spreadsheet data for rainfall into the data visualisation software.

- a. Identify and outline one format and one convention that could be applied to produce a data visualisation that would be the most effective in showing the comparison of annual rainfall for the three regions for the past five years. 2 marks

Format _____

Convention _____

- b. Explain why your selected format is preferred to educate the target audience. 2 marks

- c. Outline a feature that could be added to your selected format in order to create a dynamic display of the comparison of annual rainfall for the three regions. 2 marks

DO NOT WRITE IN THIS AREA

Question 14 (4 marks)

Ali has completed the project and wants to assess the effectiveness of the project plan in managing the project.

Propose two measurable factors and discuss how they could have contributed to the effectiveness of the project plan.

Factor 1 _____

Discussion _____

Factor 2 _____

Discussion _____

Question 15 (2 marks)

Sara was attending an industry event when she was approached by a fertiliser company that wanted to support Grains4Food's farmers. The company was willing to offer a 20% discount to all farmers on any fertiliser they bought, but requested that Grains4Food send them the farmers' personal contact information.

Should Sara be worried about this request? Outline a relevant legal issue related to this request.

Insert for Section C – Case study

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

A large Australian multinational company, Grains4Food produces wheat, barley, oats and other grains. Grains4Food have offices throughout Australia and have decided that they need to build a new grain processing facility to support their three farm regions, region 1, region 2 and region 3. They believe this will help them add value to their products and improve their profits.

Ali leads data analytics at the company and has been given the task of developing a dynamic data visualisation by Sara, the CEO of the company. This dynamic data visualisation will be used by Sara and the Board of Directors of Grains4Food to help determine the best location for the new grain processing facility.

Sara has employed two full-time staff to join the team for the project, which needs to be completed by 8 December. Sara, along with the Board, needs to easily use and review the dynamic data visualisation to help determine the feasibility of the grain processing facility proposal. Ali has been asked to present the dynamic data visualisation so that it can be viewed through a browser, and not through a data visualisation application.

Ali will need to make sure that the dynamic data visualisation represents information for each region and covers factors such as population, climate and grain production for each of the three regions for the past five years. Sara has decided that wheat will be the only grain that will be processed at the new facility.

Ali's team have decided on the tasks that need to be completed to make sure that they can present the dynamic data visualisation to the Board on 8 December. The team do not work on weekends. However, team members can work on different tasks at the same time. Part of the team will work with primary data and the rest of the team will work with secondary data. The tasks, durations and dependencies that have been decided are shown in Table 1. The project must begin on 13 November and the handover to the design team must be completed by 28 November.

Table 1

Task No.	Task	Duration (days)	Dependencies
1	Interview community leaders	5	
2	Collate and analyse interview results	1	1
3	Develop community survey	1	2
4	Online survey available to communities	5	3
5	Analyse online community survey results	1	4
6	Primary data complete	0	5
7	Complete secondary research into population growth, crop and weather data	5	
8	Analyse secondary data	2	7
9	Secondary data complete	0	8
10	Hand over to design team	0	6, 9

Ali has requested data from the Grains4Food farm management unit as an initial starting point for his team's research. He has been sent seven years of data from all five regions of different grain production throughout Australia.

Ali has also downloaded rainfall data from the weather bureau website for each of the regions for a period of 10 years to use in the dynamic data visualisation. Table 2 lists a sample of the data for Region 3 from this data set.

Table 2

Region	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual total (mm)
Region 3	2013	122	112	52	69	67	45	47	40	101	97	137	85	703
Region 3	2014	50	121	104	75	94	95	74	90	69	65	63	58	522
Region 3	2015	39	85	74	54	56	156	65	68	96	72	86	65	589
Region 3	2016	33	39	50	90	42	69	45	56	72	61	75	66	427
Region 3	2017	70	78	62	75	65	63	103	59	67	56	94	84	441
Region 3	2018	74	34	57	66	77	96	93	60	129	107	69	107	644
Region 3	2019	50	57	58	128	44	41	42	65	53	58	76	149	550
Region 3	2020	110	42	70	47	100	111	56	61	56	78	138	156	588
Region 3	2021	41	45	36	35	89	88	56	77	68	63	79	37	388
Region 3	2022	130	86	119	160	79	52	58	96	57	94	74	60	742