

Digital Learning and Teaching
Victoria

DLTV Resource Kit

For use with the VCE Applied Computing 2020–2024 Study Design

Applied Computing: Data Analytics

Units 3 and 4 Trial Examination for 2021

Sample solution

Copyright: All Victorian Curriculum and Assessment Authority (VCAA) material is copyright. Area of study outcomes, key knowledge, key skills and assessment task descriptions reproduced by permission from the *VCE Applied Computing (2020–2024)* © Victorian Curriculum and Assessment Authority (VCAA). VCE is a registered trademark of the VCAA. The VCAA makes no warranties regarding the correctness or accuracy of this DLTV resource. Teachers are advised to check the [VCAA Bulletin](#) for updates. The current VCE study design and related content can be accessed directly at www.vcaa.vic.edu.au and has been reproduced with permission.

Disclaimer: This resource has been written by the author for Digital Learning and Teaching Victoria (DLTV) to be utilised by teachers and students of VCE Applied Computing. This does not imply that it has been endorsed by the Victorian Curriculum and Assessment Authority (VCAA).

While every care has been taken by the author, they accept no responsibility for the accuracy or advice contained in this task. Teachers are asked to preview and evaluate all resources before using them or distributing them to students.

Section A – Multiple Choice

Multiple Choice Answers

Question	Answer
1	B
2	D
3	C
4	B
5	A
6	C
7	D
8	D
9	A
10	C
11	C
12	B
13	D
14	A
15	B
16	C
17	B
18	A
19	A
20	C

Section B – Short-answer questions

Solutions provided for this exam are example responses. In some cases, additional responses are given as alternative suggestions. A marking allocation guide has also been suggested for some questions.

Question 1 (7 marks)

- a. Identify a suitable type of data visualisation that could be used. 1 mark

Rachel could produce a bar chart.

1 mark for identifying any of the following, or similar: bar chart, histogram, line graph, stream chart – anything suitable for the purpose (not a pie chart)

- b. State the primary purpose of this visualisation. 1 mark

To educate or to persuade.

1 mark allocated for stating either educate or persuade. The purpose here is not to entertain or inform.

- c. Explain why there might be some authenticity problems with the data in this file. 2 marks

Rachel has no idea who created this file or collected the data. She would need to find out who carried out the research, how and when they did it before considering using it.

Response must refer to the questionable origins of the data – authenticity focussed. Any other issues must be link to the authenticity, otherwise 0 marks must be awarded.

1 mark each for:

- *Not knowing who carried out the research*
- *Any one point from:*
 - *Not knowing how the research was conducted*
 - *Not knowing when it was carried out*
 - *Not knowing how long it was carried out for (single day? 5 months?)*
 - *Anything similar and relevant*

- d. Complete the following table by writing three criteria that Rachel could apply to check the integrity of the data in the file she found. 3 marks

Data integrity factor	Suggested criterion
Accuracy	Is the data specific enough to address the current problems with traffic congestion?
Correctness	Does the data reflect the actual peak hour traffic issues?
Timeliness	Can the data be found by council workers when needed?

Criteria may be written as statements or questions. 1 mark each where it is clear the student knows what the data integrity factor means. Be aware that timeliness refers to the data being available when needed, not when the data was originally collected (this will be the most frequently missed mark).

Note – all 3 must specifically relate to Rachel’s traffic project – generic responses are not awarded marks. They should not use the factor name within the criterion.

Question 2 (9 marks)

- a. The data collected must be processed in a way which makes it attractive to the user. Identify whether this requirement is functional or non-functional. 1 mark

Non-functional.

This is not describing any purpose or process, just the appearance. No explanation is necessary for the mark.

- b. Identify one constraint and explain why Robert must incorporate it into his research. 2 marks

Robert must limit his data collection to Victorian rainfall levels. Data from anywhere else is irrelevant and will skew his findings.

1 mark for identifying a constraint such as limiting to Victoria, or 5 recent years

1 mark for explaining why this would need to happen – eg to exclude irrelevant, old data

- c. Explain how Robert could transform his collected raw data into a dynamic data visualisation. 3 marks

Robert could calculate monthly averages for the last 5 years of rainfall. This could be presented in a summary table, then exported to a data visualisation tool where suitable colours and styles could be applied. Once published, users would be able to hover their cursor over particular elements to interact and show changes in rainfall patterns over the years.

1 mark for reference to data summary, eg. Averaging each month’s rainfall

1 mark for describing some kind of processing, eg. Setting up a table of data

1 mark for user needing to interact. eg. Creating a chart which enables the user to interact to show more detail or show changes over time

Note – if the response does not make reference to the question’s scenario (Robert’s rainfall) then it cannot be awarded full marks.

d. Compare the terms data validation and data verification.

3 marks

Validation checks data at the point of being entered into the system and can only check for reasonableness. Verification checks data after it has been entered and is usually manually confirmed. Verification checks that data is correct, not just reasonable.

1 mark each – must include reference to:

- Validation checks data at the point of entry
- Verification checks data, usually manually, after entry (e.g. confirmation by contacting the creator)
- Validation checks for reasonableness, verification checks for correctness

Question 3 (4 marks)

Complete the following table by stating one advantage and one disadvantage for a wired and a wireless network. Provide completely different statements for each network.

Answer samples – one from each box – mark the first only if there are multiple offerings.

	Advantage	Disadvantage
Wired	<ul style="list-style-type: none">• Better security• Setup more expensive• Faster data access/transfer	<ul style="list-style-type: none">• More cumbersome to set up• Fixed locations for stations/nodes• Cables are expensive
Wireless	<ul style="list-style-type: none">• Can accommodate extra users easily• Users can sit where they like• Easier to set up	<ul style="list-style-type: none">• Security harder to control• Slower data access/transfer• Less reliable

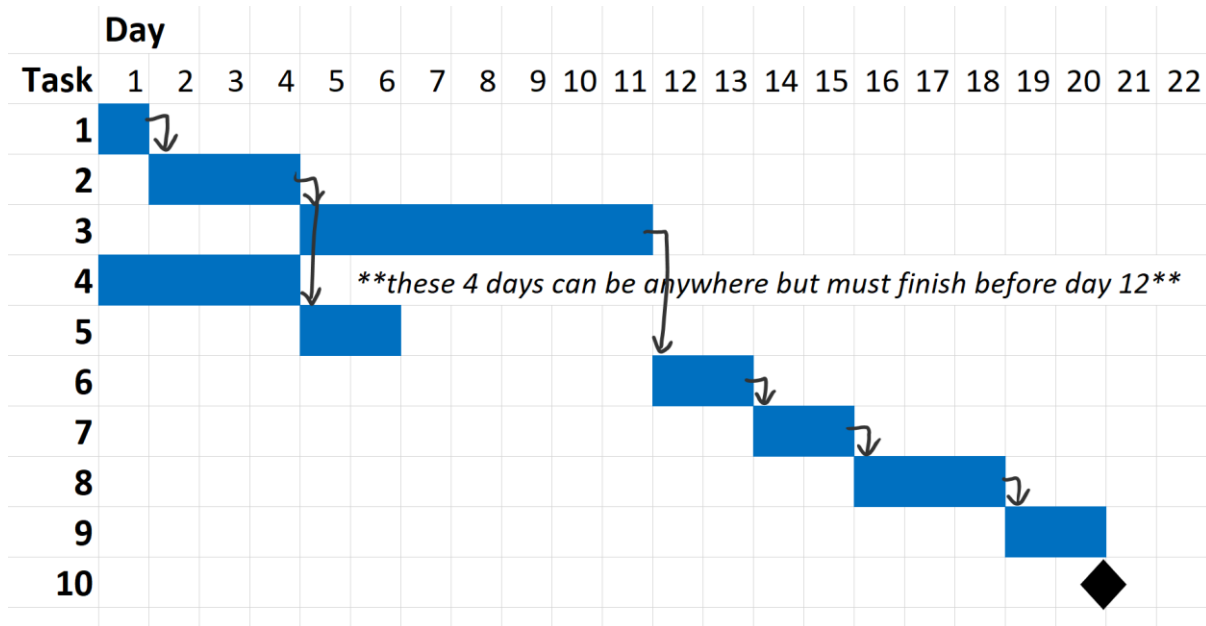
1 mark per box – however, the boxes cannot simply state a corollary of each other, eg. if a student writes that wired is faster, they cannot be awarded a mark for writing that wireless is slower.

Section C – Case study

Solutions provided for this exam are example responses. In some cases, additional responses are given as alternative suggestions. A marking allocation has also been suggested for some questions.

Question 1 (3 marks)

Create a Gantt chart to represent the task list above.



1 mark for correct placement of shading – must fit within given table, Tasks 4 and 5 must be clearly running during other tasks, clear cascading of tasks at end. Can go to 22 days, but not beyond.

1 mark for correct dependencies

1 mark for a shaded diamond at the end indicating a milestone

Question 2 (3 marks)

- a. List one reason why the school wants to acquire data about student attitudes to sport participation. **1 mark**

To gain insight into why students don't participate in the sports

Any relevant reason – to improve participation, to improve student morale... but must be linked to the case study.

- b. Write a suitable research question which could be used for this project. **2 marks**

Why do students participate less in school sporting events as they get older?

There are a range of possible answers to this question. Any relevant question but must be linked to the case study. Examples:

- *How can the school improve student sport participation rates?*
- *Why are students reluctant to take part in school sports?*
- *How can we motivate students to take part in more sporting events.*

Question 3 (3 marks)

Suggest the most appropriate data collection method and justify your choice.

The school should use an online survey. This method will reach a wider number of students in a much shorter period of time. Interviewing would potentially provide richer data, but there would be far less of it due to time constraints.

1 mark for stating online survey (this is clearly the choice based on other questions and the Section C case study)

1 mark for a reason why survey is better (quicker, wider reach, can run out of work hours)

1 mark for why interviews are not preferable (fewer students able to be accessed, can only happen during the school day, take longer)

Note: Justification means students must refer to both options.

Note: If students specify using a specific subset of students that do not attend sporting events, this is a legitimate and possibly more targeted/focused group of people than getting a wide spread of students – in this case interview is a very valuable technique to understand why non sporting students don't attend sports events.

Question 4 (13 marks)

- a. Identify which dataset would be easier to present as a chart. Explain why it would be an easier process. 2 marks**

Dataset 1 would be easier – it is already quantitative and thus easy to add up responses and create a chart.

1 mark each for:

- *Using the term quantitative – or mentioning single discrete value (right concept)*
- *Making reference to adding up/totalling AND creating a chart.*

- b. Explain how the data in Dataset 2 (Table 2) would need to be manipulated before it could be used in a data visualisation. 3 marks**

It would need to be coded in some way – manually (or electronically if they had the right kind of software for this) grouped into categories which could then be counted, and finally those numbers could be presented in a chart of some sort.

1 mark – *coding or grouping themes/ categorising/ classifying*

1 mark – *counting/totalling up the responses in each theme*

1 mark – *creation of a chart*

c. Normalise the data in Table 3 to make it compliant to third normal form.

5 marks

tblStudents	tblResponses
<u>ResponseID</u>	<u>ResponseID</u>
FamilyName	NumSwim
GivenName	NumAths
DoB	Rating
Gender	SwimParticipation
Postcode	AthsParticipation
	Avoid
	Improve
	YearLevel

1 mark each for:

- 2 tables with logical names
- Same primary key in both tables – indicated as primary key (underline, key symbol – something clear)
- Correct fields in each table – YearLevel could go in either table for several reasons.
- Age field removed
- FullName split into 2 fields

Note: No extra marks for showing how they link or indicating the 1:1 cardinality

d. Explain how a database tool could be applied to ensure that the DoB (Date of Birth) field in Table 3 and part c of this question will only accept birth dates for students aged between 10 and 19. 3 marks

There needs to be a range check on the DoB field. This can be done by calculating the age of the student by subtracting the DoB from the current date, then checking to see if it is between 10 and 19. If it is outside that range, it will not be accepted.

Answer must make reference to, for **1 mark each:**

- Range check/validation reference
- Some attempt at a valid formulaic approach (can be described)
- Entries outside this range being rejected or queried

Question 5 (5 marks)

a. Ming and Rosie will create an infographic to display an interpretation of the data. Rosie thinks it would be good to have several design ideas from which to choose. Describe one technique for generating these ideas. 2 marks

Rosie could brainstorm with Ming to come up with some different approaches. This would involve them working together to come up with as many ideas as they can before narrowing them down to the best ones.

1 mark for identifying one of the following: brainstorming, role-play, graphic organisers, mash-ups, the SCAMPER model, micro-pitches

1 mark for providing a specific characteristic of the technique.

Note: Students do not score marks if they list or describe a design tool.

- b. Identify one graphic representation typically used to display the interpretation of data in an infographic. Explain why Ming should use this representation to enhance the presentation of the data in Dataset 2 (Table 2). 3 marks**

Ming should use a bar chart to show the themes that students identified, such as student boredom. This will be easier to see in a chart than people having to read through all of the responses.

1 mark each for:

- *Identifying an element such as chart, graph or word cloud*
- *Explaining how a visualisation is easier to interpret than large amounts of text*
- *Linking to the case study dataset in some way*

Question 6 (2 marks)

Outline two effects on their data if the data integrity were diminished in any way.

If the data doesn't have integrity, then any findings might not be correct, leading to decisions made by the school which don't actually solve the problem. If there is a chance that some students are filling in the survey multiple times, or outsiders are completing it, then that data should be considered useless.

1 mark each for any reference to:

- *Potential lack of authenticity – no way to know if it's only their own students completing the survey*
- *Accuracy might be affected if students are completing it several times*
- *Decisions made based on this data might not be the correct ones as the data might not genuinely reflect the current problems*
- *Or any valid point within the scope of the case study*

Question 7 (6 marks)

It is suggested that the team's files on the school's system be backed up and archived. Explain why it is important that the school backs up and archives its files. 6 marks

Backups refer to when a second copy of data is made and stored in another location in case something happens to the original files. It should be done at least weekly on something like a portable hard drive or the cloud, and stored somewhere else in case the disaster affects the physical building.

Archiving involves relocating old unused files to another place and deleting the original ones to free up storage space for more current work. If the files are needed for something, they can still be accessed.

Backs up files – 3 marks. 1 each for:

- Mentioning frequency, storage medium AND storage location
- Applies to all current files
- Purpose – to have another copy in case of disaster

Archives files – 3 marks. 1 each for:

- Mentioning removal/deletion of original files
- Applies to old, unused files that might still be needed
- Purpose – to free up system space but still have access if needed

Note: if response does not explicitly relate to the case study, it cannot receive full marks

Students should take note of the number of marks allocated to a question to determine the level of detail required.

Question 8 (9 marks)

- a. Identify one convention that has been followed in Design A and explain how it is used to improve readability. 3 marks**

The heading is very clear – large, centred and at the top of the page with strong contrast. This is where headings are expected to be and make it easy for the viewer to read it.

1 mark each for mention of:

- Heading (nothing else really relates to good readability)
- Location – top centre of infographic
- Size

- b. Identify one characteristic of efficiency and one characteristic of effectiveness in Design B. Explain how each is an improvement over Design A. 4 marks**

Efficiency: the design is clearly laid out with the eye drawn to the large stacked column, so viewers can more readily locate the information that most interests them. In Design A, all of the data is mixed in together.

Effectiveness: the design has a much clearer message. Once the viewer has decided what to look at, the data is presented in a clear and easy to understand manner with iconography and few words. This is better than Design A as it is less confusing.

1 mark each for clearly linking Design B to efficiency (time, cost, effort) and effectiveness (clarity, attractiveness, communication of message etc). IF this is incorrect or switched around, award 0 marks.

1 mark each for both giving an example and comparing to Design A. If Design A is not mentioned at all, full marks cannot be awarded

- c. Identify one characteristic of their target audience and explain what they could do with their infographic to ensure that it educates this group appropriately. 2 marks**

The target audience is made up primarily of adults – they are all older. The team should make sure that the infographic is clear, to the point and uses conservative fonts.

The audience is the school council, not the student population. No marks can be awarded if students alone are identified.

There is a 'spoiler' in this question – it answers a previous question (Section B Q1b). Students should be alert and notice little hints like this.

Question 9 (3 marks)

Describe how Rosie could use a software tool to create the infographic.

Rosie could use an online infographic generator. She could prepare the data in a spreadsheet, any text in a word processor and create/edit images beforehand, then upload them. She can use the data to create charts, add images to make it look attractive and include a heading which summarises the key message.

1 mark for mentioning preparation tasks – data, text, images

1 mark for uploading or adding elements such as text boxes, charts, images, icons

1 mark for mention of cohesive / eye-catching elements including a title

Note: this is different to creating a chart, which was an earlier question

Question 10 (3 marks)

Describe a technique that could be used to test that a dynamic aspect of one of the infographic's interactive charts works correctly.

Ming or Rosie could test that the charts are able to display extra information by displaying the chart onscreen and hovering or clicking the relevant sections. They would expect the data to be visible – if not, then there is a problem that needs fixing.

1 mark for identifying the dynamic element (eg. click to show how many students indicated the days were 'boring')

1 mark for stating the prediction AND the actual outcome

1 mark for mentioning the comparison between the prediction and actual result

Note: If no specific reference is made to the case study and an interactive element, the answer cannot receive full marks

Question 11 (6 marks)

- a. Identify the key legislation that affects how the school must handle and store the data it collects. 1 mark**

Privacy and Data Protection Act 2014

No other answers can be accepted. Answer must include the year. Privacy Act 1988 is not acceptable as this is a state government school and does not fall under Federal privacy laws.

- b. Explain the ethical issue around Waterford High School staff giving permission to share this data. 2 marks**

The issue is using student data for a purpose that was not originally intended, data which is old and potentially out-of-date, and the school then receiving benefits ("kickbacks") which might influence further decisions.

1 mark for identifying that the issue is surrendering old data for sporting supplies.

1 mark for more details explaining this point.

- c. **Outline a strategy that could be used to resolve the ethical issue raised in the previous response.** **3 marks**

Rosie, Ming or Alex could contact the ex-students and seek their permission to pass on their contact details. Even though they are no longer students at the school, they still have to right to decide what happens to their data.

1 mark for contacting the ex-students

1 mark for seeking permission to pass on their data

1 mark for relating the response directly to the case study – more than just the character names.

Question 12 (4 marks)

Give an example of an accidental threat that might impact the security of the survey data and explain what could be done to protect it.

The data might be stored on a laptop and while students are being observed at the swimming carnival it could be dropped into the pool. To protect this data, there should be several backup copies, preferably in the cloud, staff should be trained to not take computers near water and the laptop should be stored safely when not in use.

1 mark for naming a reasonable threat – eg dropping a computer in the pool, accidentally moving / deleting files or mis-spelling a file name so that it can't be located

3 marks for any 3 valid points that relate to the identified threat, such as (for file deletion)

- *Have a policy to prevent staff from emptying computer trash cans*
- *Train staff to check files before deleting*
- *Train staff to read all warning messages that appear in dialogue boxes before dismissing them*

Students should choose the discussion point carefully, to ensure there are enough points to discuss for full marks. As always, if the response does not fully relate to the case study, full marks cannot be awarded.