

Victorian Certificate of Education
2021

SUPERVISOR TO ATTACH PROCESSING LABEL HERE

STUDENT NUMBER

									Letter
--	--	--	--	--	--	--	--	--	--------

APPLIED COMPUTING:
DATA ANALYTICS

Written examination

Thursday 28 October 2021

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	16	16	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 29 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

Tjandamurra is the marketing director of a beverage company. He would like to display the company's recent share of the beverage industry as a percentage in a visual form for the sales department.

Which one of the following would be the best type of chart to use as a data visualisation?

- A. pie chart
- B. line graph
- C. bubble chart
- D. box and whisker plot

Question 2

Latifa wants to create a database for her dog grooming business, which will contain the name, address, telephone number, postcode and email address of each of her customers.

Which one of the following data types should be used for the postcode field?

- A. text
- B. date
- C. numeric
- D. character

Question 3

A large retail store does not have the required software to create dynamic data visualisations.

This is an example of

- A. scope.
- B. design.
- C. a constraint.
- D. a solution requirement.

Question 4

The Australian Bureau of Statistics (ABS) is a good source of authentic data.

Statistics collected provide information on a wide range of

- A. social media issues, which is used by the local community.
- B. qualitative and quantitative data that is mainly used by local communities.
- C. environmental issues, which is used by the government to make improvements.
- D. economic, environmental and social issues, which is used by governments and the local community.

Question 5

Arsenio and Innuka work for a community bank that plans to rebrand its financial products. The bank decides to collect data directly from its customers. Arsenio has been asked to interview customers who walk into the bank and Innuka has been asked to send all customers an online survey consisting of 10 questions. Customers will answer the questions in the online survey by giving a rating between 1 and 10, with 1 being very low and 10 being very high.

Collecting the data for manipulation will involve

- A. Arsenio collecting qualitative data and Innuka collecting qualitative data.
- B. Arsenio collecting qualitative data and Innuka collecting quantitative data.
- C. Arsenio collecting quantitative data and Innuka collecting quantitative data.
- D. Arsenio collecting quantitative data and Innuka collecting qualitative data.

Question 6

A sample record from a student enrolment database is shown below.

StudentID	Subject name	Date commenced	Result	Fees paid
BA78635	Hospitality	29/02/2020	98	Y

Which one of the following data types has been used for the Fees paid field?

- A. string
- B. integer
- C. Boolean
- D. floating point

Question 7

Emiko has collected feedback from customers who have visited her cafe. She has grouped each review comment into one of three general categories: food, service and wait time.

Customer feedback	Category
food is overpriced	food
wait is too long on Friday nights	wait time
not enough staff working on Saturday nights	service
food was delicious	food
wait is too long on some nights	wait time
staff are occasionally rude	service

Which one of the following techniques is Emiko using for data manipulation?

- A. coding data
- B. inputting data
- C. cleansing data
- D. validating data

Question 8

The most appropriate design tool for representing a dynamic data visualisation is

- A. a storyboard.
- B. an IPO chart.
- C. a Gantt chart.
- D. a data dictionary.

Question 9

Which one of the following is a technique for generating alternative design ideas?

- A. mind mapping
- B. conducting a survey
- C. creating a prototype
- D. searching for and downloading data sets

Question 10

A dynamic data visualisation has been created to show the location and movement of traffic in real time on Melbourne roads.

Which one of the following is a criterion for evaluating an alternative design idea to measure the efficiency of this dynamic data visualisation?

- A. the text is readable
- B. information is presented accurately
- C. the time taken to access a street address
- D. plug-ins are compatible with different browsers

Question 11

An organisation needs to protect itself from threats to the security of its data and information. Diminished data integrity can have negative consequences for the organisation.

The characteristics of data integrity include

- A. accuracy, authenticity, completeness and correctness.
- B. authenticity, correctness, reasonableness and relevance.
- C. accuracy, attractiveness, completeness and readability.
- D. reasonableness, relevance, readability and timeliness.

Use the following information to answer Questions 12 and 13.

Kontrol College is a secondary school that was founded in 1902. The Friends of Kontrol College was formed with the purpose of maintaining an association with past students of the college. All past students automatically become members of this organisation. The following online registration form was developed to collect the names of past students.

Kontrol College – Registration

Title*

Please select ▼

First name*

Surname (current)*

Surname (at time of enrolment, if different)

The year you graduated from Kontrol College (Year 12)

The table below displays data collected by the Friends of Kontrol College.

GradID	Title	FirstName	SurnameCurrent	SurnameEnrolled	Year
1	Mr	Hovhannes	Ohanyan	Ohanyan	1959
2	Dr	Mendal	Kohen	Cumberton	1958
3	Miss	Kamala	Patel	Mitra	1951
4	Mr	Alveraz	Hernandez	Hernandez	1962
5	Ms	Dongmei	Lau	Lau	1956
6	Mrs	Mani	Jones	Loch	1963
7	Mr	Tim	Wang	Wang	1965

Question 12

The Friends of Kontrol College want to generate a list of students who graduated between 1958 and 1963. The following query rule was written.

Year > 1958 AND Year <= 1963

Which one of the following displays the GradIDs that would result from the query rule above?

- A. 1, 4, 6
- B. 2, 3, 4
- C. 1, 2, 3, 4
- D. 1, 2, 4, 5

Question 13

A celebration is being organised for students who left Kontrol College 60 or more years ago. The secretary of the organisation has suggested a number of ways that these past students can be identified.

Which one of the following would be considered the most efficient technique for identifying these past students?

- A. Sort the Year field in ascending order.
- B. Create a query referencing the Year field where Year ≥ 60 .
- C. Create a query with the calculated field, (2021 – Year), and the filter, filter ≥ 60 .
- D. Sort the Year field of the student database in descending order and count the first 60 names.

Question 14

To ensure that an infographic caters to an audience of elderly people

- A. use technical words.
- B. use a wide range of colours.
- C. use dark text on a dark background.
- D. ensure that the font size is large and easy to read.

Question 15

There are a number of software tools for creating a dynamic data visualisation.

When selecting a software tool, a developer must ensure that the software tool has features that

- A. will create moving images.
- B. manipulate data in real time.
- C. will arrange slides with charts in logical order.
- D. effectively search for and replace text in a document.

Question 16

Hymie has collected data on the reasons why students are absent from school. He wants to present this data as percentages in a table that he will later place in an infographic. Before creating the infographic, Hymie uses a calculator to determine that the percentages add up to 100%.

Hymie is

- A. validating the contents of the table.
- B. evaluating the contents of the table.
- C. testing the correctness of the contents of the table.
- D. analysing the correctness of the contents of the table.

Question 17

Dino Graphics is a graphic design business that has recently moved into a new office space. Most of the walls and doors are made of aluminium and glass. The staff need greater bandwidth when working at their desks and portability when meeting clients in the meeting rooms.

Which type of network would work best in this office space?

- A. a wired network
- B. a mobile network
- C. a satellite network
- D. a combination of a wired and a wireless network

Question 18

To ensure the integrity and security of data that is transferred on any network, it is recommended that HTTPS protocol is used.

The HTTPS protocol creates a secure connection by

- A. encrypting data.
- B. authenticating data.
- C. using twisted cables.
- D. creating a firewall between the sender and the receiver.

Question 19

Common threats to the integrity of data include

- A. user error, acts of sabotage and lost hardware.
- B. legal obligations, digital systems policies and user error.
- C. theft of hardware, biometric hardware and anti-virus software.
- D. acts of sabotage, biometric identification and legal obligations.

Question 20

Restoration, as part of a disaster recovery plan, involves

- A. testing the disaster recovery plan.
- B. completing a daily, incremental backup.
- C. acquiring and installing replacement equipment.
- D. conducting appropriate risk assessment as part of the disaster recovery plan.

DO NOT WRITE IN THIS AREA

CONTINUES OVER PAGE

TURN OVER

SECTION B – Short-answer questions

Instructions for Section B

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

Question 1 (4 marks)

The Statewide Basketball League has developed the design below for the input form of its player database.

text labels

- *right aligned*
- *Times New Roman*
- *12 pt*

format mask

- *dd/mm/yy*

headshot

- *jpg format*
- *35 × 45 mm*

Player

Surname

First name

Date of birth

Height

Handspan

Hometown

Nickname

University

Player photo

Upload

text labels

- *right aligned*
- *Times New Roman*
- *12 pt*

format mask

- *dd/mm/yy*

headshot

- *jpg format*
- *35 × 45 mm*

background colour

- *white*

drop-down

- *list of universities*

numeric

upload button

- a. Identify the design tool that has been used to create this input form.

1 mark

- b. Identify **one** design principle relating to the appearance of this input form. 1 mark

- c. Data from the input form has been exported into a spreadsheet, part of which is shown below. One of the team managers would like to know the average height and handspan of the players on their team.

Surname	First name	Date of birth	Height	Handspan	Hometown	Nickname	University
Constranza	Kyrie	15/05/00	203 cm	26 cm	Korfonia	Dribbles	University of Korfu
Russo	Vince	27/12/01	2.01 m	24 cm	Mentona	Vinnie	University of Mentona

Describe **two** techniques for cleansing data, which can be used before the data is manipulated and any calculations made.

2 marks

Question 2 (5 marks)

A travel agency wants to create a data visualisation to inform its customers about the 10 most popular travel destinations in Australia. The manager is aware that she can choose between an infographic or a dynamic data visualisation. She does not fully understand the difference between an infographic and a dynamic data visualisation but she wants a poster with text, statistics, images and charts.

- a.** Which type of data visualisation should the manager choose? Justify your response. 3 marks

- b.** Outline one format and one convention that could be used to increase the effectiveness of this type of data visualisation for the travel agency's customers. 2 marks

Format _____

Convention _____

Question 3 (3 marks)

A car manufacturer has decided to install a Global Positioning System (GPS) in its new cars to collect data about customers' driving patterns. The customers' data will be stored in a centralised database in the car manufacturer's server room at its head office in Sydney.

Identify the key legal requirement that the car manufacturer needs to meet with regard to this data and describe what the manufacturer will need to do to meet this requirement.

Legal requirement _____

Description _____

Question 4 (4 marks)

- a. Explain why it is important to generate more than one design idea. 2 marks

- b. Describe how evaluation criteria could be developed to help determine a preferred design. 2 marks

Question 5 (4 marks)

An organisation would like to collect medical data from a range of primary and secondary sources around Australia. The organisation wants to query a range of data sets, identify relationships and patterns, and manipulate and cleanse data. It would then like to sell this data to health insurance companies.

- a. At some point in the next few years, the organisation will have to consider archiving the files and disposing of the files on its servers.

Explain the difference between archiving files and disposing of files.

2 marks

- b. Consideration is being given to enabling health insurers to access this data via cloud storage.

State one advantage and one disadvantage of using cloud storage for storing data.

2 marks

Advantage _____

Disadvantage _____

DO NOT WRITE IN THIS AREA

CONTINUES OVER PAGE

TURN OVER

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 assist the Vice President of Sales, Kamal, to present his findings to the board.

Question 2 (2 marks)

Provide **one** reason why XPLUSMISC would want to acquire customer data through the use of a new loyalty card.

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

- DO NOT WRITE IN THIS AREA**

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

- DO NOT WRITE IN THIS AREA**

DO NOT WRITE IN THIS AREA

Question 4 (4 marks)

Kamal knows that introducing the new loyalty card will help him to learn the spending habits of XPLUSMISC customers. However, he needs to determine how to collect data from his staff and the customers regarding what new items they would like to have introduced and whether they would support a barista service.

Complete the table below by providing an example for each method of data collection and justify how suitable this example would be for collecting data for XPLUSMISC.

Method of data collection	Example	Justification
surveys		
observation		

Question 5 (4 marks)

Kamal will be collecting a lot of staff and customer data. He needs to ensure the integrity of the data as the Chief Executive Officer (CEO) will be making a decision on the future of the business based on Kamal's findings.

Identify and describe two criteria for checking the integrity of the collected data.

Criterion 1 _____

Description _____

Criterion 2 _____

Description _____

Question 6 (4 marks)

Before Kamal can start designing the dynamic data visualisation, he needs to determine its functional and non-functional requirements as part of the analysis.

- a. Outline two **functional** requirements of the dynamic data visualisation. 2 marks

Functional requirement 1 _____

Functional requirement 2 _____

- b. Outline two **non-functional** requirements of the dynamic data visualisation. 2 marks

Non-functional requirement 1 _____

Non-functional requirement 2 _____

Question 7 (6 marks)

- a. Data collected from each store will be stored in a database. The data will be used to create the dynamic data visualisation.

Kamal starts planning the database solution. He enters the following field names into a table to get an idea of how they will look: Reward_ID, Surname, FirstName, Postcode, Mobile_Number, NumberOfPoints, Transaction_ID, DateOfPurchase, Store, Time, TotalCost, ProductPurchased_ID, ProductName, ProductCost, ProductQuantity.

The table is currently in first normal form (1NF).

In the space provided below, normalise the table to third normal form (3NF).

3 marks

DO NOT WRITE IN THIS AREA

- b. Next, Kamal completes a data dictionary in order to determine the correct data types to use for the field names. However, he is struggling with three field names.

Complete the table below by identifying the most appropriate data type for each of the field names listed.

3 marks

Field name	Data type	Description
Reward_ID		stores the customer's Reward_ID ID is made up of letters and numbers
Mobile_Number		stores the customer's mobile number Mobile_Number is made up of four numbers beginning with a zero, followed by a space, followed by three numbers, followed by a space, followed by three numbers
NumberOfPoints		stores the customer's accrued points NumberOfPoints is made up of numbers

Question 8 (4 marks)

After designing the database solution, Kamal moves on to designing the dynamic data visualisation.

Identify and describe two design tools for representing the appearance and functionality of the dynamic data visualisation.

Design tool 1 _____

Design tool 2 _____

DO NOT WRITE IN THIS AREA

Question 9 (3 marks)

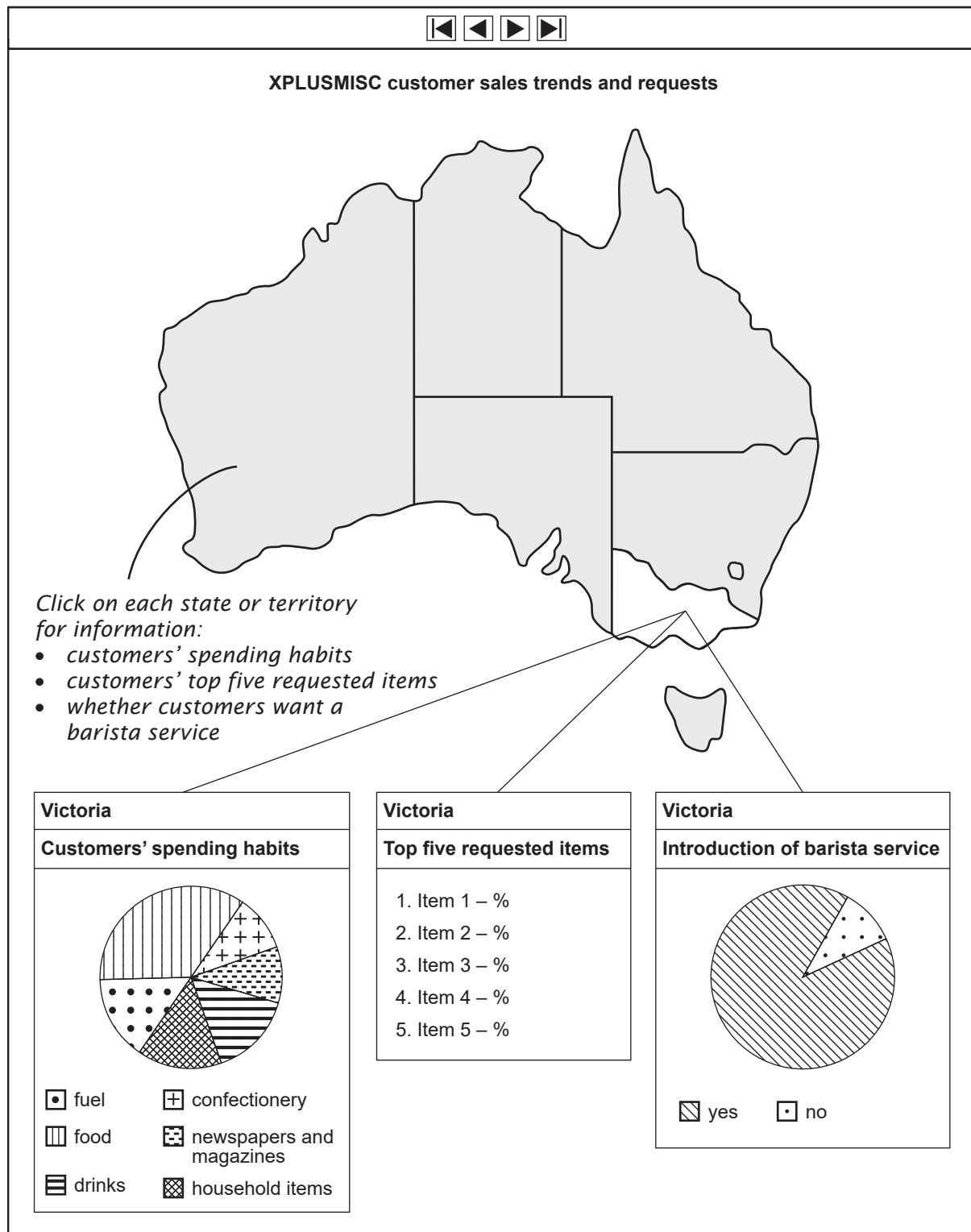
Kamal will be collecting many data sets from a range of primary and secondary sources. This data will be imported into the database file.

The Chief Information Officer (CIO) tells Kamal that he is going to have to consider the security of the files as they contain sensitive corporate and private data.

Propose a set of procedures for Kamal to implement to manage the security of the files.

Question 10 (6 marks)

Kamal has produced his preferred design for the dynamic data visualisation that he will use to present his findings to the board.



The dynamic data visualisation will display customers' spending habits, the top five requested items and whether customers want a barista service.

This data will be displayed for all states and territories in Australia. There will also be a breakdown for each metropolitan and regional area in each state and territory, along with a comparison of XPLUSMISC's data and competitors' data.

- a. Because of the amount of data to be displayed, Kamal will have to consider how to navigate through the dynamic data visualisation.

Describe two suitable techniques for navigating through the dynamic data visualisation.

2 marks

Navigation technique 1 _____

Navigation technique 2 _____

- b. Identify and describe two software functions that would be suitable for developing the dynamic data visualisation.

4 marks

Software function 1 _____

Description _____

Software function 2 _____

Description _____

Question 11 (5 marks)

When developing the software solution, Kamal uses a range of techniques for validating and verifying data.

- a.** Explain the difference between validating and verifying data. 2 marks

- b.** Provide two examples of validating data in the database. 2 marks

Example 1 _____

Example 2 _____

- c.** Provide **one** example of how data can be verified in the dynamic data visualisation. 1 mark

Question 12 (2 marks)

Kamal has completed the dynamic data visualisation and wants to test it to check that it performs as expected.

Outline **one** suitable testing technique for navigating through the dynamic data visualisation.

Question 13 (2 marks)

Outline **one** strategy that Kamal could use to evaluate whether the dynamic data visualisation meets all the functional and non-functional requirements.

Question 14 (3 marks)

In recent months, power outages near stores around Australia have resulted in an inability to collect data and information from customers and to process transactions. The CIO has become increasingly concerned about these interruptions.

What type of threat is a power outage to XPLUSMISC's data and information security? Recommend one strategy that will enable XPLUSMISC to protect its stores against this type of threat.

Type of threat _____

Strategy _____

Question 15 (4 marks)

While Kamal has been collecting data, the CIO has been visiting stores around Australia to determine the effectiveness of their data and information security strategies.

The CIO has noticed the following issues:

- managers and staff accessing customers' details out of curiosity, as there are no access restrictions
- server rooms at stores being left open, potentially allowing customers to enter
- customers in some stores can get behind the counters and use the computers
- customer data is emailed to the head office in spreadsheets with no access restrictions

The CIO realises that XPLUSMISC needs to update its physical and software security controls.

Complete the table below by providing two types of physical security controls and two types of software security controls. Describe how each type of control will improve XPLUSMISC's current data and information security practices.

	Type of control	Description of how it will improve practices
Physical security control 1		
Physical security control 2		
Software security control 1		
Software security control 2		

Question 16 (3 marks)

Recommend an appropriate backup strategy to protect the data and information of XPLUSMISC's stores.

Insert for Section C – Case study

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

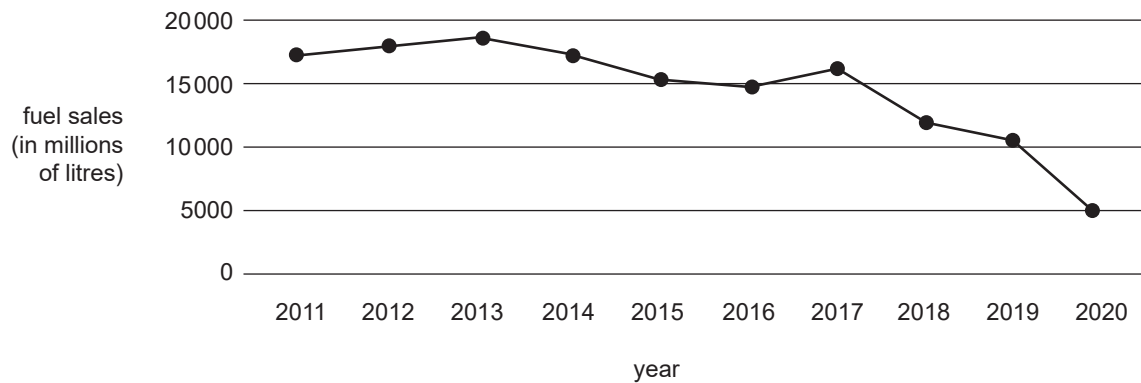
XPLUSMISC convenience stores

XPLUSMISC is an Australian convenience store company with 800 stores located across all Australian states and territories, and in metropolitan and regional areas. XPLUSMISC stores are open 24 hours a day, seven days a week, every day of the year. All XPLUSMISC stores sell newspapers and magazines, confectionery, food, drinks, a range of household items and fuel.

Changes in sources of revenue

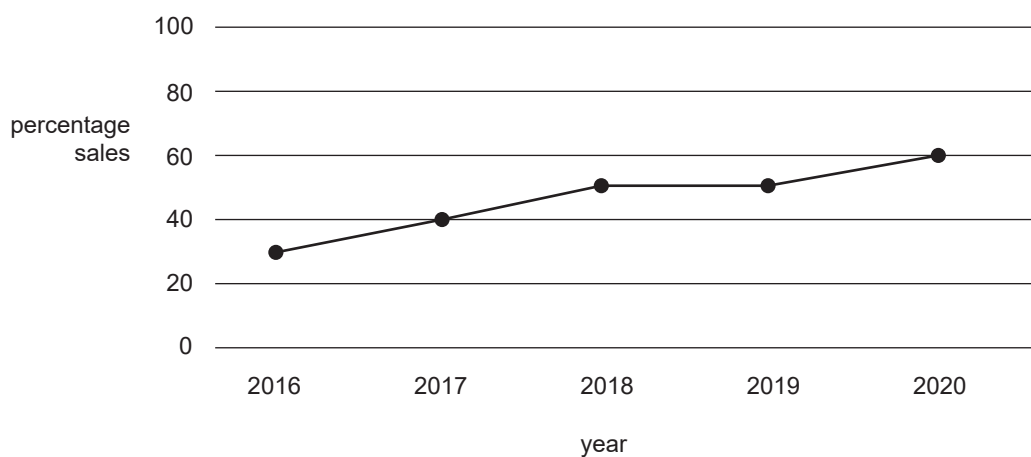
The Vice President of Sales, Kamal, is concerned about the decrease in fuel sales over the last 10 years (Chart 1). He is concerned about the long-term impact this will have on XPLUSMISC.

Chart 1 – Fuel sales (2011–2020)



Kamal has also noticed an increasing trend in the sales of non-fuel items (confectionery, food, drinks and household items) as a percentage of all sales. This is shown in Chart 2.

Chart 2 – Percentage sales of non-fuel sales



XPLUSMISC's response to change

Kamal has had many conversations with store managers around Australia and with industry colleagues. Store managers have reported customers wanting a barista (a person trained in the preparation of different types of coffee) service. Industry colleagues have stated that the introduction of a barista service and increasing the variety of food, drinks and household items available could help to increase XPLUSMISC's overall revenue. This will enable the company to reduce its dependence on fuel sales or even move away from fuel sales altogether.

Kamal has spoken to the Chief Executive Officer (CEO) to see if this is something they should consider. Currently, XPLUSMISC does not know the spending habits of its customers and will require this data to assist with decision-making. XPLUSMISC will also need to know if customers would use a barista service.

Kamal explains to the CEO that they currently have a cardboard loyalty card, which customers can get stamped each time they spend over \$20 at XPLUSMISC. When they have 10 stamps, they get 10% off their next purchase. He further explains to the CEO that they have purchased new loyalty cards that could be given to customers for them to scan at the register to receive points for redeeming discounts. These could be modified to collect data on customers' spending habits.

The CEO is very keen on this idea and tells Kamal that they would also need to collect data from their staff and customers on the new food, drink and household items they would like to have introduced.

The CEO has asked Kamal to investigate this idea. She wants Kamal to collect the data and present the findings as a dynamic data visualisation to the board in 12 months. She also wants Kamal to liaise with the Chief Information Officer (CIO).