



4.11 Big Data

1) One characteristic of a data set that might result in it being classified as Big Data is that it contains a variety of different forms of information.

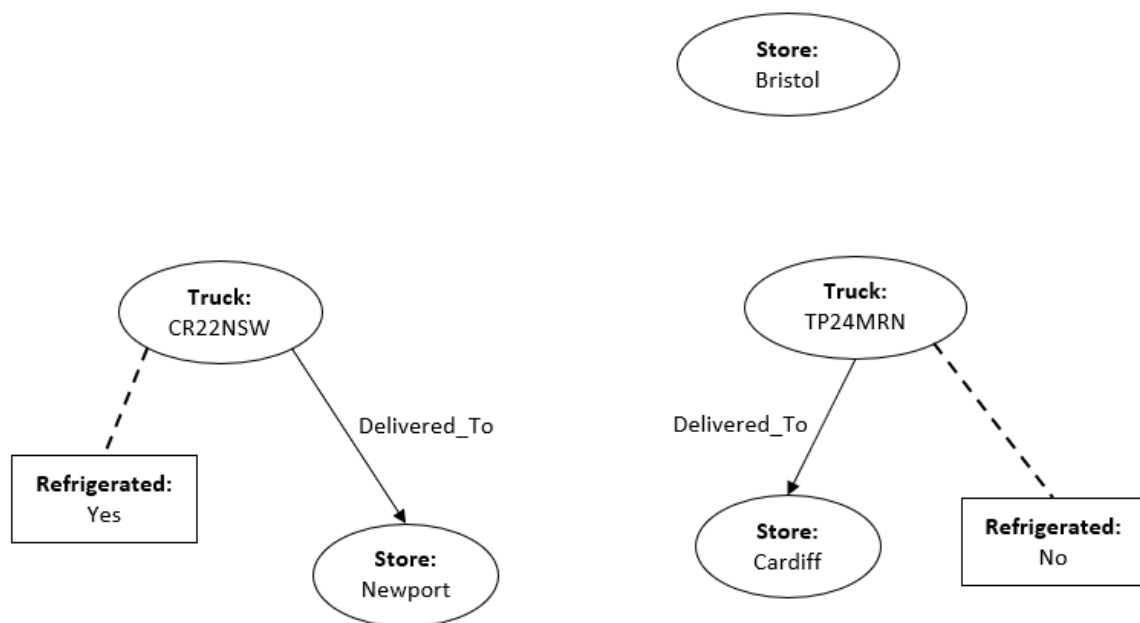
a) Describe **two other** characteristics that might result in a data set being classified as Big Data.

[2 marks]

b) In a fact-based model, data is represented as atomic facts, which are immutable (i.e. will never change). Fact-based models can be represented visually using a graph schema.

Figure 1 shows part of a graph schema for a data set about deliveries made to stores by trucks.

Figure 1



Complete the graph schema in **Figure 1** to represent the following additional facts.

- Truck TP24MRN has made a delivery to the Bristol store.
- Truck CR22NSW was last serviced on 17 November 2025 and truck TP24MRN was last service on 01 October 2025.
- Both trucks are owned by a haulage company called Trevor's which has 12 employees and has a head office in Bath.

[3 marks]



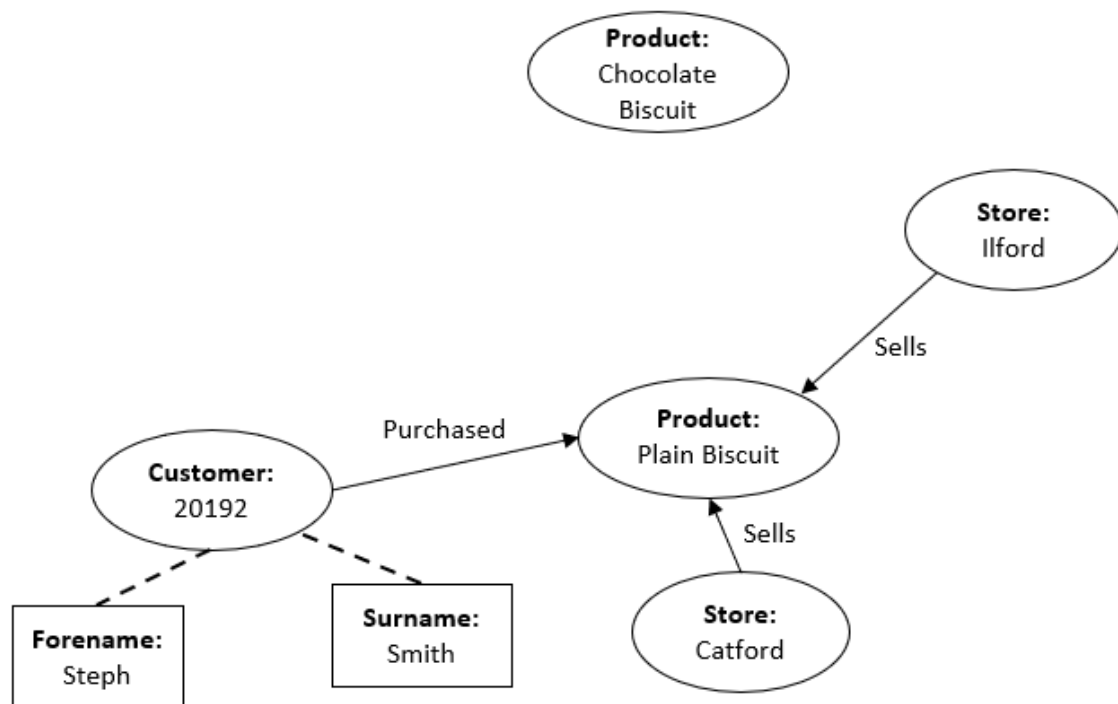


- 2) A supermarket chain uses a system which stores details of all of the products that it sells and the sales that it makes.

The data that the supermarket stores is Big Data.

- a) The graph schema in **Figure 2** represents part of a fact-based model of the dataset that the supermarket has built.

Figure 2



Modify the graph schema in **Figure 2** to represent the following additional facts.

- The Ilford store sells chocolate biscuits.
- There are 15 individual biscuits in a packet of plain biscuits and each packet costs £1.
- Both chocolate biscuits and plain biscuits are made by the company Snack City. The company has 65 employees and also makes Pop tarts.

[3 marks]

- b) One approach to dealing with Big Data is to write code that can be distributed to run across more than one server.

State **two** features of functional programming languages that make it easier to write code that can be distributed to run across more than one server.

[2 marks]





- 3) Big Data is increasingly used by organisations to store and analyse large amounts of information.

Discuss the characteristics of Big Data, the challenges involved in storing and processing it, and the techniques that can be used to overcome these challenges.

You should also consider ethical and legal issues associated with storing and processing personal data.

[12 marks]

