# Worksheet 6 Big data

**Task 1**

1. Explain what is meant by these three characteristics of Big Data:

(a) Volume

(b) Velocity

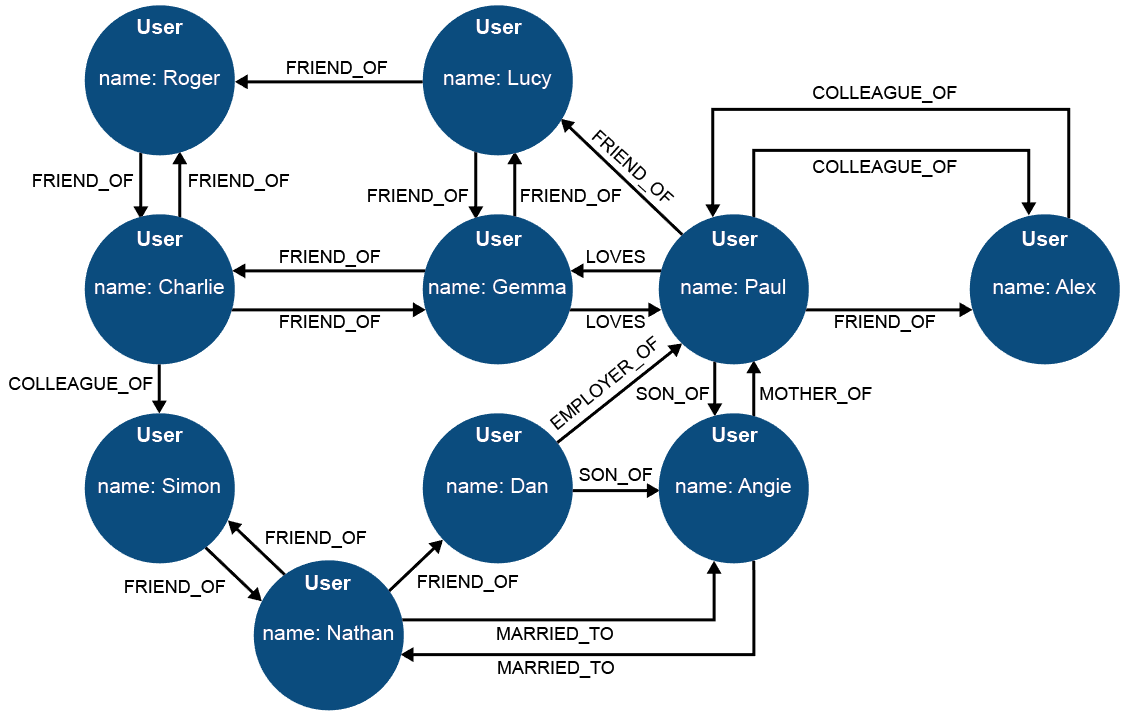
(c) Variety

2. Why can “Big Data” not be held in a traditional database structure?

3. Research examples of Big Data and Smart Cities and make notes about what you can find out about.

Summarise your findings in a one-minute presentation.

**Task 2**



4. How many degrees of separation are there between Roger and Angie?

5. Identify some of the nodes and edges of the above graph.

6. Discussion: can ‘colleague\_of’ work in only one direction? Can ‘employer\_of’ work in both directions? Are human relationships simple to define – or immutable?

**Task 3**

7. Watch the video about Hadoop distributed programming compared to SQL, at <https://youtu.be/MfF750YVDxM> [6m13s]

You are a Big Data consultant. Prepare a one minute presentation to brief a client on which approach they should use, and why, for:

* managing the payroll for a group of UK academy schools.
* researching student attitudes to school uniform policies as expressed through social media.
* managing a traffic light control system in a Smart City, with the aim of both reducing congestion and giving priority to emergency service vehicles responding to incidents.

8. Research the PageRank algorithm at the heart of the Google story. Consider how the scale of web search activity has grown in your lifetime.