OOP and FP homework 6

1

```
                    Sensor ID: Unknown (door sensor)

Location: Fridge     Sensor ID: 20-31

Location: Towel rail   Sensor ID: 20-87

Temperature: 32 deg

Function: Smoke alarm  Sensor ID: 65-4F


              Function: Temperature

  Location: Kitchen                Serviced by: XYZ alarms
```

2

- Big data needs to be processed quickly, for example a credit card company cannot afford to wait to detect if a transaction is suspicious: they need to catch it as it is going through
- Big data is generated very quickly, for example AWS gets thousands of customer logs per second
- There are lots of different types of big data. For example, YouTube generates lots of video, Instagram generates lots of photos, Google drive generates lots of documents
- Some of this data is structured, some is unstructured, for example Google accounts have relational data whereas Google docs are files

3

Using multiple servers in a distributed architecture allows you to divide work between lots of processors and process very quickly at scale. So long as you can reduce bottlenecks and process a lot concurrently, you can avoid having to process all the data on one machine (which would be unfeasible at best)

4a

A processing task applies a map to process data in parallel, and then uses reduce on the result of the map operation in order to calculate the final result to be returned. The map operation can be parallelized over a distributed architecture, so only the reduce operation at the end is a bottleneck

4b

These algorithms are commonly used with big data as they are pure, which means that you can do lots of parallel processing without running into race conditions