**Data Collection (AS)**

|  |  |
| --- | --- |
| **K1** | Understand and use the terms ‘population’ and ‘sample’  Use samples to make informal inferences about the population  Understand and use sampling techniques, including simple random sampling and opportunity sampling  Select or critique sampling techniques in the context of solving a statistical problem, including understanding that different samples can lead to different conclusions about the population |

**Commentary**

Whenever one conducts a statistical experiment we might ask the following questions:

* What problems am I going to address?
* What data do I need to collect?

Once these decisions have been made the key question is then:

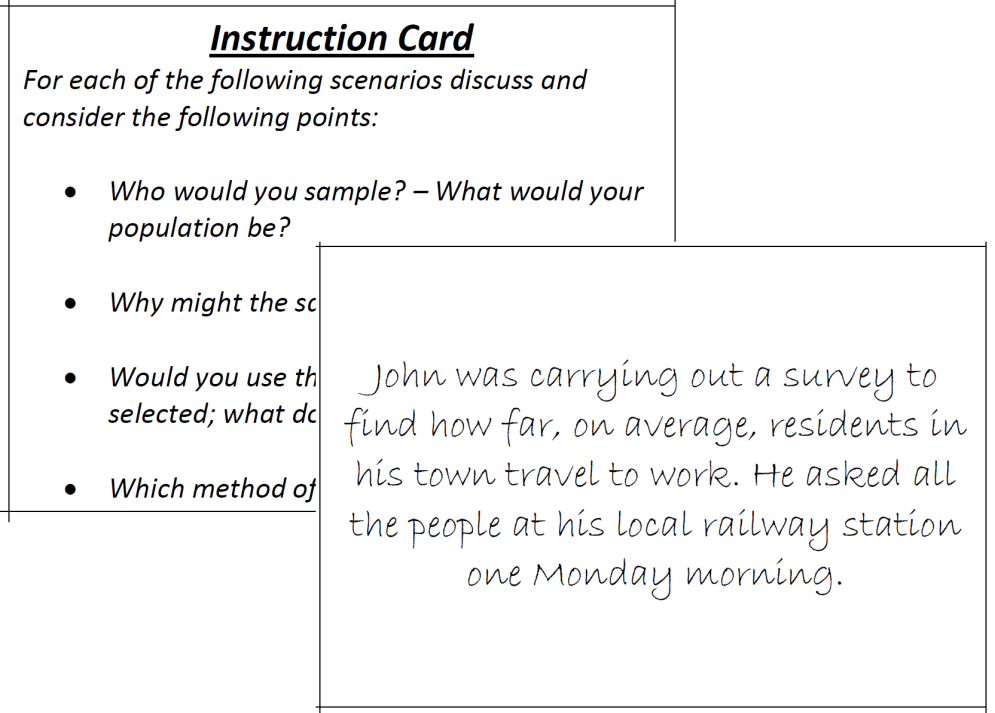
* How will I collect the data?

Information collection often involves taking a sample from all the possible data (the parent population). However, sometimes you are able to collect the whole population; such a 100% sample is called a census. There are many different ways of collecting samples.

A sample typically provides a set of data values of a random variable, drawn from all such possible values, the parent population (often just called the population). The parent population can be finite, such as all professional golf players, or infinite such as the points where a dart can land on a dart board. A sample is intended to give information about the parent population so it must be representative of it.

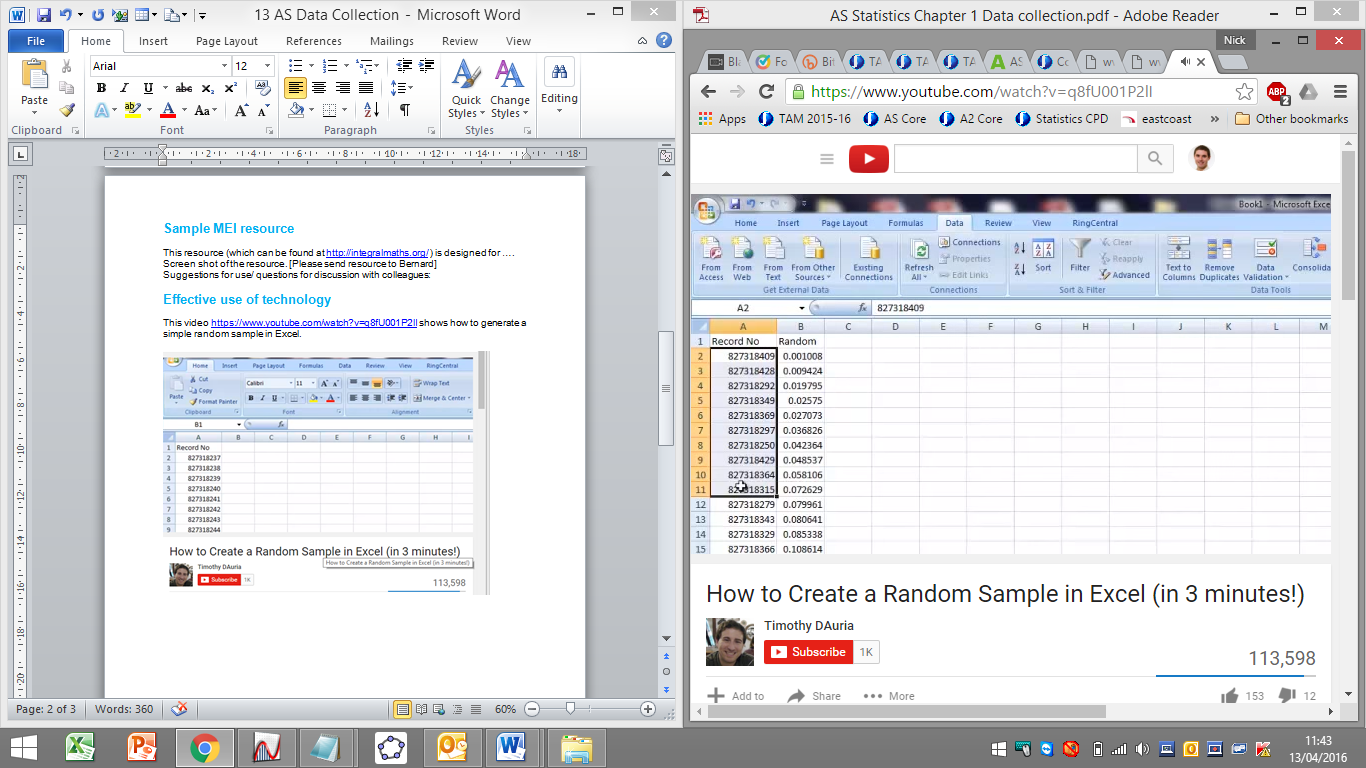
**Sample resource**

‘Sampling techniques’ (which can be found at <https://my.integralmaths.org/integral/sow-resources.php>) encourages students to think carefully about different sampling techniques if they are carrying out different statistical investigations.

****

**Effective use of technology**

‘Excel random sample’ is a short video (which can be found at [www.mei.org.uk/integrating-technology](http://www.mei.org.uk/integrating-technology) ) showing how to generate a simple random sample in Excel.



|  |  |
| --- | --- |
| **Title** | **Time allocation:** |
| **Pre-requisites**   * GCSE: Collecting data and spotting bias in questions | |
| **Links with other topics**   * Hypothesis testing: Having a representative sample to work with is essential for conclusions to be valid. | |
| **Questions and prompts for mathematical thinking**   * What is the same and what is different about random sampling and opportunity sampling? * Give me an example of a data set with an outlier which should not removed * . | |
| **Applications and modelling**   * Which colour Smartie is the most common? * Can you predict the winner of the student council elections by taking a sample? | |
| **Common errors**   * Removing every item of data which looks like an outlier * Understanding how bias can be introduced in sampling and its effects on data analysis. * Making rash comments such as ‘there *are* outliers in the data’ rather than ‘there could be some outliers in the final class’. | |