Year 1

~ Statistics ~

D A T A

C O L L E C T I O N

Data Collection

Lesson 1

A-Level Statistics : Data Collection : Year 1

1.1 Introduction

Many organisations busy themselves gathering data about us. In Britain the Government, for example, gathers information about each person in the country every ten years in what is termed a census. Simple and straight forward information, such as each citizen's age and where they live, help government plan and provide health and school services in each part of the country. Supermarkets have invested large amounts of time and effort building up profiles on their customers, not simply observing what they buy but, for example, how they respond to a variety of special offers. The motivation is to better provide what their customers are likely to respond favourably to. Happy customers will return and so boosts profits.

Often, it's impractical or too expensive to carry out a census on a population and so a sample is taken. We shall look at a variety of ways of taking such a sample, of ***Data Collection***, and the advantages and disadvantages of each method.

Even within a sample, the quantities of numbers, the data, generated can be enormous. To look through pages and pages of numbers does not tell a typical human being much. A statistic is a single number that tries to see into the character of the larger set of sample data in an informative way. We often then want to know how this statistic of the sample data is likely to relate to that of the parent population from which the sample was taken. Some statistical numbers that you will have looked at previously include mean, median, mode, range, interquartile range, standard deviation and skew.

Statisticians are also very aware that a picture is an excellent way of passing impressions of what the "sea of numbers", or sample data, is saying.

There is a dark side to statistics. Statisticians see themselves as neutral observers but sometimes the act of measuring something changes the thing being measured. More infamously, politicians often only report those statistics that support their vested interest, and don't mention any conflicting information.

A best selling mathematics text book is called, "How to lie with statistics". In it the many ways in which politicians, drug companies and marketing managers have in the past used graphs and statistical numbers to mislead and give wrong impressions are explored. We'll look at some of these.

1.2 Exercise

Explain what a 'vested interest' is and give an example.