CHAPTER 1 THE LANGUAGE

Statistics is a way of communicating information using numerical data. It involves the collection of data from a source or variety of sources; this data is then analysed, interpreted, and presented in a way that makes it easier for people to make sense of.

Example

Sarah owns a car yard and is looking at sales of Toyotas for the year to help work when she should run a promotion. When she looks at the data for just one two months, the sales figures are small enough for Sarah to compare them by hand. When she looks at a whole year, it takes a while, but she can still compare the figures by hand and see that there were more sales in December than in July. Yet when Sarah looks at Toyota sales over a ten-year period, there is too much data for Sarah to reasonably count Toyota sales per month, then do monthby-month and year-by-year comparisons. This is where statistics comes in — as long as Sarah has kept good records in a spreadsheet on her computer, she can use a set of simple statistical tools to find out which months have higher sales and which months have lower sales. She can then analyse this data to find patterns which can help her make good decisions about when to run promotions.



Newspapers, advertisements, and other media often cite statistics in their reports. Examples might include:

- poll results showing political predictions
- 23% of children assessed were below the national average for literacy and numeracy
- the average rainfall for London is 690 mm per year
- 96% of women saw results in just 8 weeks.

Statistics are also used in many other areas. For example –

- Education in education students will receive grades – 50% for their maths exam. They will also receive report cards. Many educational facilities will also keep other records, such as absences, gender, cultures and races etc.
- Companies will make purchasing decisions based on statistics, which product is the cheapest, how much profit they will make by selling a product in a certain way etc.
- Statistics are used in the prediction of weather which is important for many different groups, such as sailors, fishers, sports people, events organisers, tourist organisations, farmers and more.
- Statistics in sports are also important, from the recording of goals and scores to see who wins, to overall recording to see who wins the finals etc. In elite sports, statistics

can become even more detailed, for example, looking at how often a goalkeeper lets a goal in or how hard a volleyball player is hitting a ball.

There are many more examples than this, but in short, statistics are an undeniable part of everyday life, and having a better understanding of how they work will make it easier to decode new information and make good future decisions.

Important Terms in Statistics

Statistics unsettles some people because it seems like a foreign language. People who are not routinely communicating through numbers will naturally feel more comfortable using words, and words they are familiar with. When people hear the word "statistics" they can feel uncomfortable, even though they may be doing statistics all the time. How often do we go to the shops and think how much something has gone up in price or which product would be cheaper and so on, these are all simple statistics.

When people try to communicate using words that are not commonly heard or used by you, it is understandable that you might feel uncomfortable.

The first step to feeling more comfortable about statistics, therefore, might be to be able to understand a bit more about the language. This section will help make it easier to decode statistical language in reports, media, work documents, and more.

Data Collection Method

Data is individual facts, numbers, statistics or bits of information. A data collection method refers to the tools and techniques used to gather data. For example –

- Surveys A survey might ask people what their preference is for flavoured milk, then look at what percentage of people like chocolate milk, what percentage strawberry milk and so on.
- Recording certain things such as sales figures.
- Counting how many people enter a store within a set time period.

There are many different data collection methods, and some are better than others. Researchers must think very carefully about the data collection method they choose and the tools they use (For example, a watch that counts minutes and seconds, versus a watch that counts only minutes) to ensure they collect the best possible data.

Statistical Method

Raw data is unprocessed data. So going back to the flavoured milk, if we ask 100 people for their preference, we will have 100 answers. A statistical method refers to the tools and techniques used to analyse that raw, unprocessed data. What will we do with it? Most likely we will work out percentages of how many people like chocolate milk, how many like strawberry and so on. This is how researchers and analysts extract information from data. There are many statistical methods available. The two that are most useful for the average person to understand are:

- descriptive statistics
- inferential statistics

Descriptive Statistics

Basically, descriptive statistics do just that, they describe the statistics, the results. Examples of descriptive statistics include measures of central tendency, such as mean, median, mode and measures of variability such as variance and standard deviation. Descriptive statistics are often shown with graphs, charts, tables and so on.

Inferential Statistics

So whilst descriptive statistics describe data, inferential statistics allow you to make predictions from the data – inferences.

The main areas of inferential statistics are -

- Sampling. For example, you can study a sample of the population and then make generalisations about the entire population. When we see statistics in the news, such as 50% have experienced depressive feelings at some point in their lives, this does not mean that they have asked every single human being how they feel. It means that they asked, say 100 people, and found that out of those 100 people, 50% had experienced depressive feelings. They then infer that this means that all humans may feel like this.
- Hypothesis testing We may have a research question, such as whether one headache tablet is better than another or whether music helps learning better than sitting in silence. We can use sample data to try to find out an answer to our research question.

So we might conduct an experiment. We could give Group A headache tablet A and Group B headache tablet B. We then ask them to report how their headache felt over a period of a month. This is using sample (inferential method) as we do not test everyone in the world, only the people in the group.

We collect all the results from their reports.

We can then prepare descriptive statistics about their results, such as – graphs, bar charts, the mean results etc (descriptive statistics).

We then infer something from our results – whether tablet A or tablet B is better (inferential statistics).

Statistical methods are separate to data collection methods, such as surveys. Although these are sometimes spliced together in media reports, researchers and statisticians keep these two things separate. So, the data collection method is simply how we collect that data. The statistics is what we do with the data.

Data Point

A data point is one piece of data. For example, if asked how many children they have a person might reply 3 or 0 or 2. This is a single point of data at that time. Another example, how much did you earn last year? How many cm of rain was there yesterday? 3cm. It is a single point of data.

Observation

An observation is the thing being counted, measured, or recorded for the

statistical study. For example, how many tigers are counted in the wild.

Data Set

A data set is the whole set of data to be analysed. For example, if a health check was being carried out on children in a school, they might record their height, weight, IQ, age in months and years. This would give us several pieces of data about each child, so a data set.

Research Population

The population of research is the people or animals or other factors that are being studied. As we said earlier, it is hard to test every single human being on the planet, so if we studied 1000 people in our research, our research population would be 1000 humans. If we wanted to study koalas in the wild, we may not be able to study all koalas, so we might study the koalas living in a certain national park. So, the research population would be the koalas living in that national park, not all koalas.

