Lesson 9

Presentation and Display of Quantitative Data

Learning Objectives

- All students **will** identify and present data using appropriate graphs, charts and tables.
- All students **should** be able to justify when different forms of graphical display should be used.
- All students **could** interpret graphical displays to draw conclusions.

Key Terms

- Normal distribution
- Positively skewed distribution
- Negatively skewed distribution
- Bar chart
- Histogram
- Line graph

Extension activities: Complete the extension activity on graphs on page 194 of the *Green Haired Girl* digi-book.

Questions to guide your thinking ...

- What are the characteristics of normal and skewed distributions?
- Explain the difference between a positive and negative skew.
- What information do summary tables typically contain?
- What is the difference between a bar chart and a histogram?
- What kind of data is represented by a line graph?

Descriptive statistics and distributions

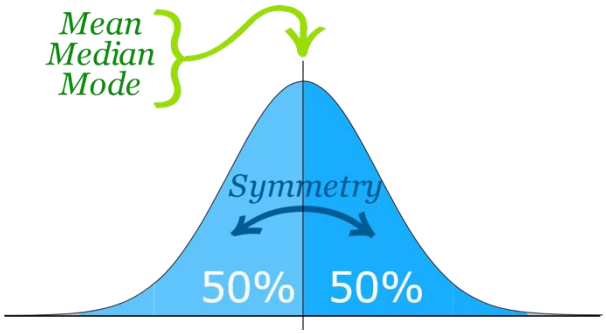
As we saw in the previous lesson, psychologists make use of **descriptive statistics** to summarise and analyse their data. In addition to numerical descriptive statistics, such as the measures of central tendency and dispersion that we have already looked at, a variety of graphs and charts are used.

Psychologists also use frequency distributions to make sense of their data. These are essentially pictures of populations (or samples) that descriptive statistics describe.

Distributions

One of the most common distributions in psychology is the **normal distribution** (also called the 'bell curve' after its shape and the Gaussian distribution after the mathematician who first identified its properties).

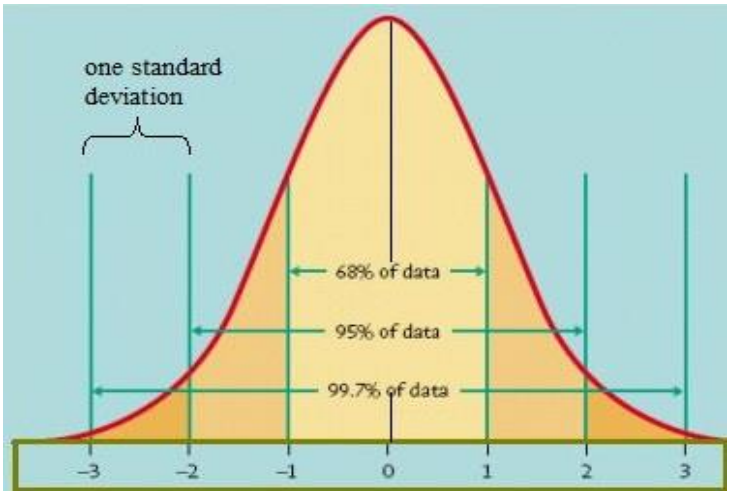
When certain variables (e.g. height, weight, IQ scores) are measured, the frequency of these measurements form a normal distribution, i.e. bell-shaped and symmetrical:



Within a normal distribution, most scores are located in the middle area of the curve, with very few scores at the extreme ends. The mean, median and mode occupy the same mid-point of the curve. The 'tails' of the distribution extend outwards, never touch the x-axis (and therefore never reach zero – more extreme scores are always *theoretically* possible).

Normal distribution curves have important statistical properties. These are linked to two key measurements of the distribution: the mean and standard deviation:

- 68% (68.26%)of the population fall between **one** standard deviation (SD) above and **one** SD below the mean (middle section of the curve).
- 95% (95.44%) of the population fall between **two** SDs above and below the mean.
- 99% (99.73%) of the population fall between **three** SDs above and below the mean value.



- Use your knowledge of the **properties of normal distributions** to answer the following questions:

Question 1

In the total population of the UK, the arithmetic mean of IQ scores is 100 and the standard deviation is 15.

a) What approximate percentage of the UK population have IQ scores between 100 and 115?

b) Between which two scores do approximately 95% of the population of the UK fall?

c) What approximate percentage of the UK have scores above 115? (2)

Question 2

An arithmetic test is given to 200 children. The test has been standardised and scores are normally distributed. The mean of the test is 50 and the standard deviation is 6.

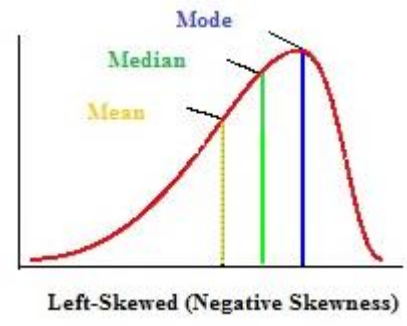
a) Approximately how many children will score between 44 and 56 on this test? (1)

b) Approximately how many children will score above 56? (1)

c) What is likely to be the median score for the 200 children? (1)

Skewed distributions

Some data sets may have a lot of scores at one end or the other. We refer to these distributions as ‘skewed’ distributions because they are unbalanced: one end is longer or more spread out than the other.



Negative skew (tail on the left)

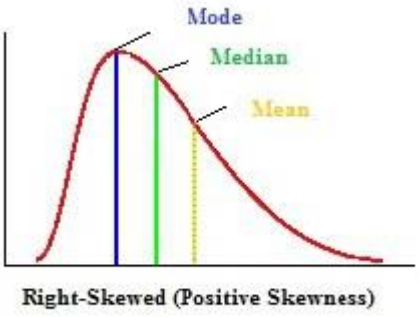
Most of the distribution is concentrated towards the right of the graph, resulting in a long tail on the left e.g. a very easy test in which most people get high marks would produce a negatively skewed distribution.

The mode is still the highest point of the peak, but the median comes next to the left and the mean is far off to the left (if scores are arranged lowest to highest).

Positive skew (tail on the right)

Most of the distribution is concentrated towards the left of the graph, resulting in a long tail on the right e.g. a very difficult test in which most people get low marks would produce a positively skewed distribution.

The mode is again the highest point of the peak, the median is in the middle and the mean is far off to the right (if scores are arranged lowest to highest).



When presenting a table or graph, always have a title and clearly labelled columns/axes.

Tables

Data is often presented in tables. It is displayed in rows and columns

In a report, raw scores are typically converted to and summarised by descriptive statistics (such as means and standard deviations).

For example:

Summary Table Showing Stress Scores in No Exercise and Exercise Conditions

There should be a descriptive title for the table.

	Stress Score in No Exercise Group (Control Condition)	Stress Score in Exercise Group (Experimental Condition)
Mode	31	15.87
Median	30	16
Mean	30.67	16
Range	34	17
SD	8.96	4.41

A summary paragraph is usually included beneath the table to explain the results.

Bar charts

Bar charts are used when data are divided into categories (rather than being measured on a numbered scale). This is known as discrete data.

Categories are placed along the x axis and frequency on the y axis (this can be reversed). The height of each column represents the frequency of that item.

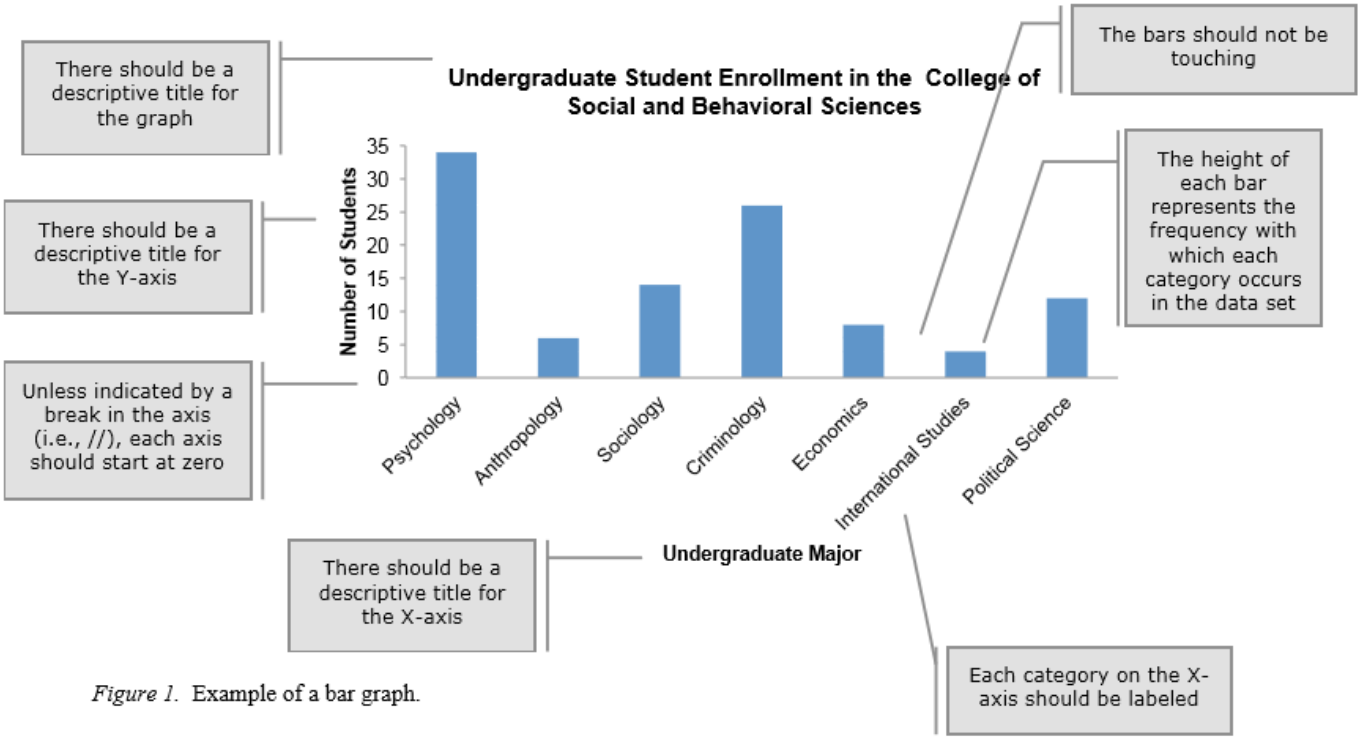


Figure 1. Example of a bar graph.

Histogram

Bars touch each other, which shows that the data is continuous rather than discrete (as in a bar chart).

The x axis is made up of equal-sized intervals of a single category (e.g. percentage scores in a maths test broken down into intervals of 0-9, 10-19, 20-29, etc.).

The y axis represents the frequency (e.g. number of people who scored a certain mark) within each interval. If there is a zero frequency for an interval, the interval remains but without a bar.

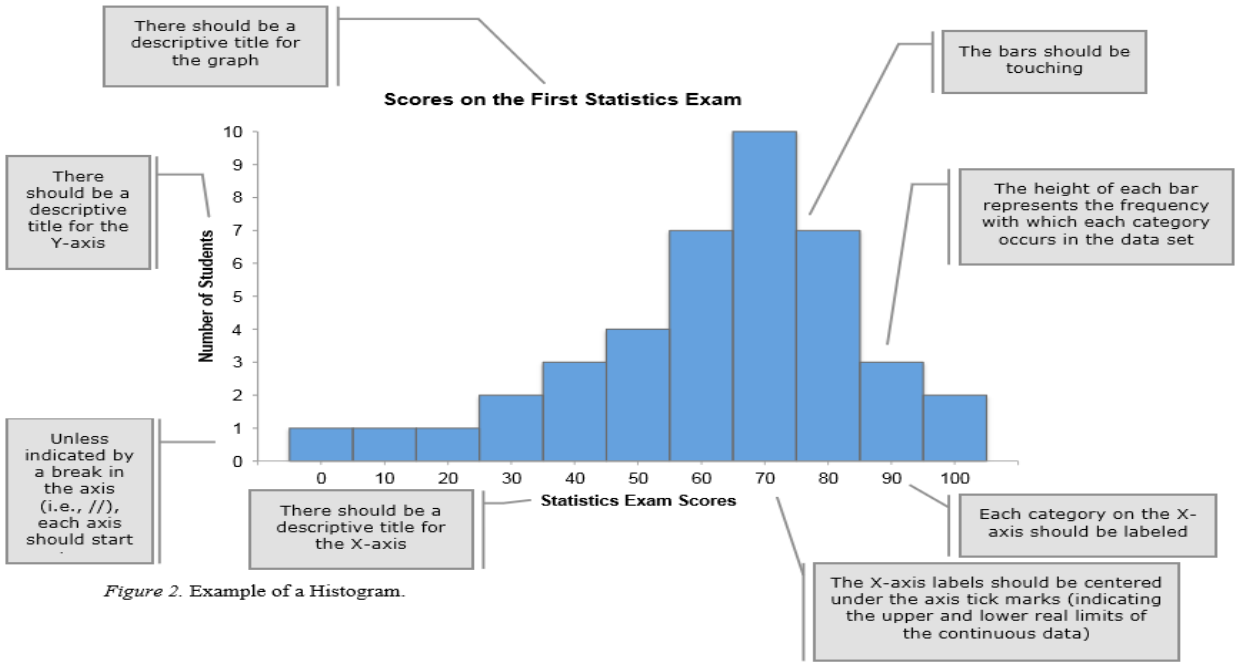
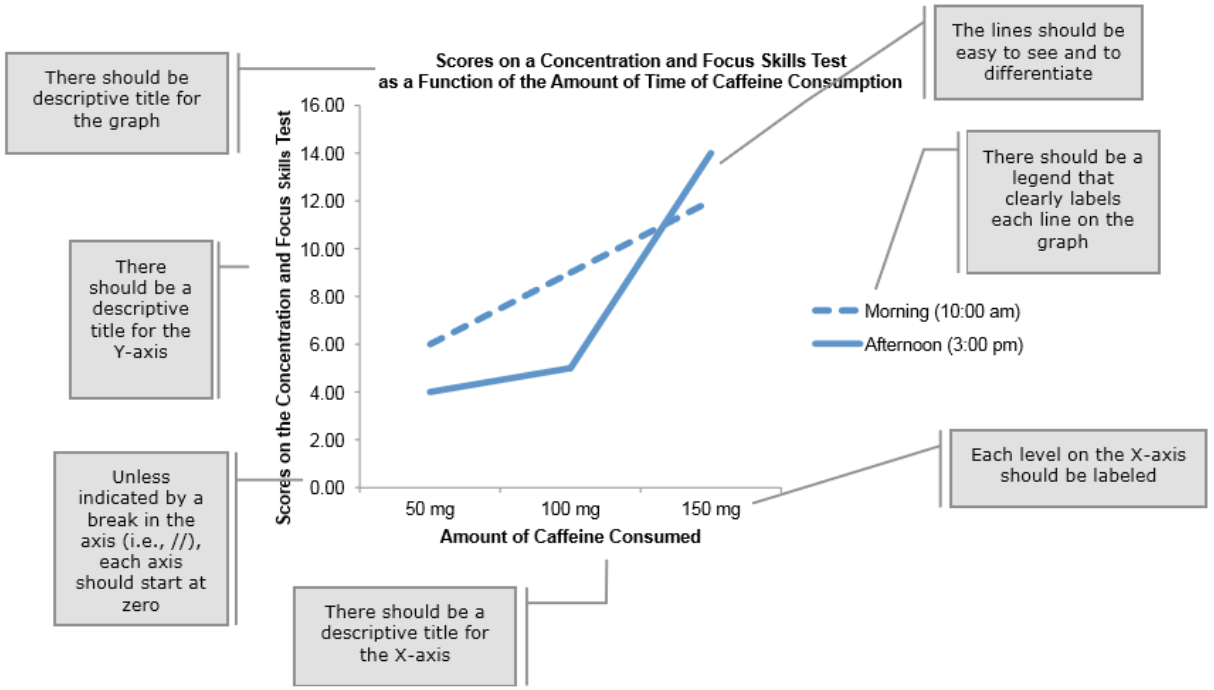


Figure 2. Example of a Histogram.

Line graphs

Line graphs also represent continuous data and use points connected by lines to show how something changes in value, e.g. over time.

Typically, the IV is plotted on the x axis and the DV is plotted on the y axis.



- Apply your knowledge of **descriptive statistics** to the following data sets. For each:
- Identify which type of **graphical display** should be created, justifying your decisions.
- **Create** the appropriate graphical displays.
- Interpret the results of the graphical displays to **draw conclusions**.

Table 1: The frequency of different heights in a class

Height (cm)	Frequency
100-119	5
120-139	10
140-159	15
160-199	4

Table 2: The mean recall score for students who had learned a random word list or an organised word list

	Random word list	Organised word list
Mean score	15	23

Table 3: The frequency of favourite pets in a year group

Favourite pet	Number of people
Rat	28
Dog	52
Cat	73
Budgie	15
Snake	27
Fish	48
Gerbil	5

Exam Practice: Descriptive Statistics and Distributions

Specimen Paper A Level Paper 3 1st set

Read the item and then answer the questions that follow.

A psychologist investigating the investment model of relationships devised a self-report Investment Scale for use with a group of 100 female participants. The scale gave an investment score for each participant on a scale of 0-20, with 0 representing no investment in relationships and 20 representing extreme investment in relationships. The psychologist calculated measures of central tendency for the investment scores. He found that the mean investment score was 8.6, the median investment score was 9.5 and the mode investment score was 13.

8.1 Sketch a graph to show the most likely distribution curve for the investment scores in this study. Label the axes of your graph and mark on it the positions of the mean, median and mode. (3 marks)

8.2 What sort of distribution does your graph show? (1 mark)

A-level Psychology Paper 2 2017

A psychologist wanted to test whether listening to music improves running performance.

The psychologist conducted a study using 10 volunteers from a local gym. The psychologist used a repeated measures design. Half of the participants were assigned to condition A (without music) and half to condition B (with music).

All participants were asked to run 400 metres as fast as they could on a treadmill in the psychology department. All participants were given standardised instructions. All participants wore headphones in both conditions. The psychologist recorded their running times in seconds. The participants returned to the psychology department the following week and repeated the test in the other condition.

The results of the study are given in Table 1 below:

Table 1 Mean number of seconds taken to complete the 400m run and the standard deviation for both conditions

	Condition A (without music)	**Condition B** (with music)
Mean 400m time (s)	123	117
Standard deviation	9.97	14.5

12. Explain why a histogram would not be an appropriate way of displaying the means shown in Table 1. [2 marks]

13. Name a more appropriate graph to display the means shown in Table 1. Suggest appropriate X (horizontal) and Y (vertical) axis labels for your graph choice. [3 marks]

14. What do the mean and standard deviation values in Table 1 suggest about the participants' performances with and without music? Justify your answer. [4 marks]

15. Calculate the percentage decrease in the mean time it took participants to run 400 metres when listening to music. Show your workings. Give your answer to three significant figures. [4 marks]