

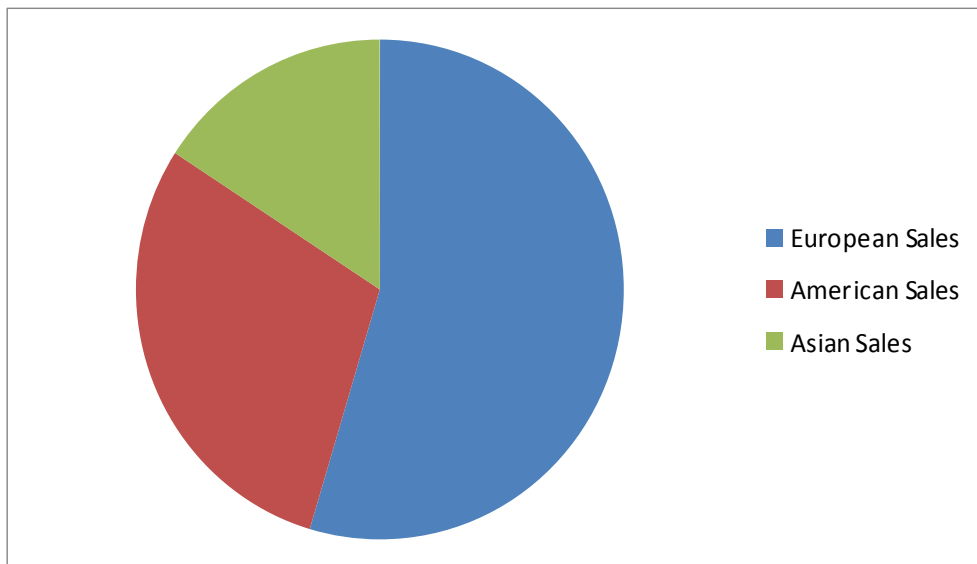
Chapter 1

Data analysis

The time in which we now live is often referred to as the **Information Age** and one of the most important skills required in business is the ability to understand and handle various types of data. It has been said that if all the data generated in the world in one day was put onto DVDs, and they were piled on top of each other, they would reach the moon and back. Under these circumstances the ability to select **relevant** data, and ignore what is referred to as 'background noise' i.e. irrelevant data, is crucial.

Data comes in many forms including tables, charts and graphs. Data is published in many places: newspapers, magazines and journals; company annual reports and company websites; government websites; reports from market research agencies etc. Part of the skill of the researcher is to find the relevant information in the first place. Fortunately, search engines on the internet have made this process easier than it used to be. Once information has been researched by a business, it may need to be presented in a different format (e.g. from a table to a pie chart) ready to be interpreted and analysed by business analysts.

Pie charts



A pie chart is a circular chart that is split into sectors to show percentages or the relative value of different categories of data. For example, the pie chart above shows the sales of a business in different parts of the world (Europe, America and Asia).

A pie chart gives a good visual representation of the relative sizes or shares of a whole part, which can make the different categories of data easier to compare than if the same data was presented in a table.

A pie chart can be used to show market income, market share for products or something more specific such as the sales revenue for different types of pizza sold by a pizza chain. Pie charts alone do not give very detailed information, but they do give an overall picture of the information being presented.

Pie charts can help businesses to make decisions. For example, if the sales of a specific pizza with a mushroom, tuna and pineapple topping had proportionately fewer sales than any other pizza that a business was selling, then they might consider withdrawing the product or promoting it more.

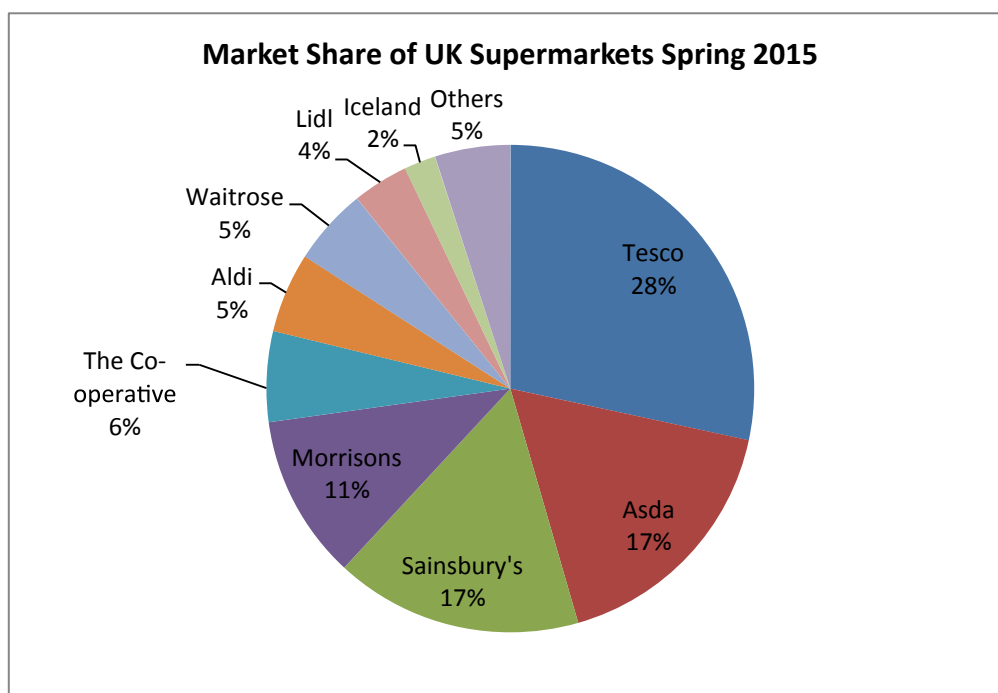
However, pie charts are not very effective for showing increases or decreases of proportions over time as trends are not shown and data cannot be extrapolated. Furthermore, they do not show causal relationships such as the impact of advertising spend on sales revenue, meaning that additional information would be needed (perhaps in narrative form) to support data presented in pie charts.

The following table gives the market share of the main supermarkets at the beginning of 2015 (rounded to the nearest percent):

Supermarket	Market share (Spring 2015)
Tesco	28%
Asda	17%
Sainsbury's	17%
Morrisons	11%
The Co-operative	6%
Aldi	5%
Waitrose	5%
Lidl	4%
Iceland	2%
Others	5%

N.B.: The total proportions within a pie chart should add up to 100%, in order for the 'pie' to be whole.

The same information in the table above can be presented in the form of a pie chart, as seen below:



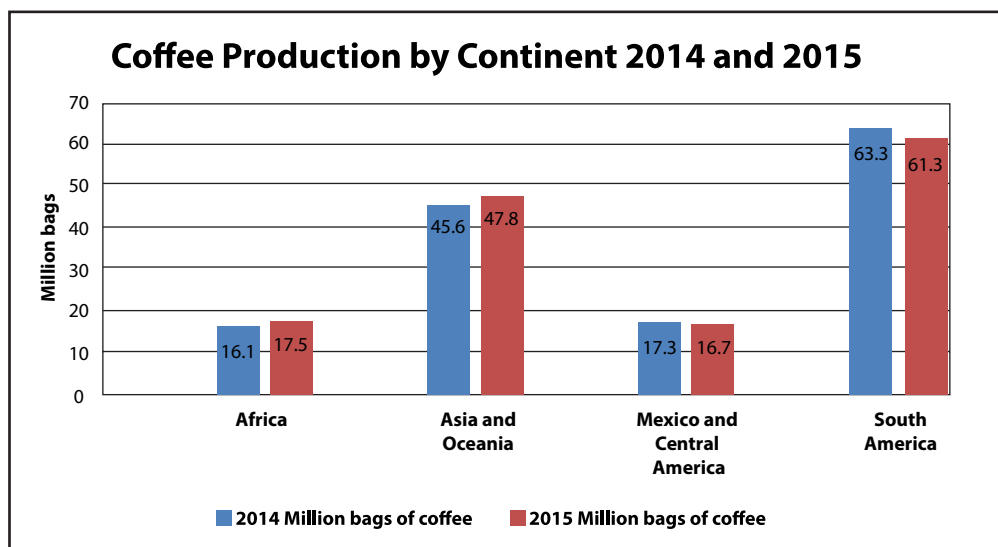
If you compare the data in the pie chart with the data in the table, you should notice that the pie chart is easier to understand as it presents the information more clearly and visibly. At a glance it is possible to answer questions such as “which supermarket has the largest/smallest share? Which is bigger Aldi or Lidl? What is Morrisons’ share of the market?” etc.

Bar charts

Bar charts are another visual way of presenting data. Data is usually grouped into categories using rectangular bars with the height of the bar representing the frequency for the category. These bars can be presented vertically or horizontally. One axis will show the categories being compared and the other axis will show the frequency. The table below gives details of world coffee production by continent for 2014 and 2015:

Continent	2014 <i>Million bags of coffee</i>	2015 <i>Million bags of coffee</i>
Africa	16.1	17.5
Asia and Oceania	45.6	47.8
Mexico and Central America	17.3	16.7
South America	63.3	61.3

The same information, when presented in the form of a bar chart, as below, is much easier to interpret:



It is immediately clear which continent is the most important and which is the least important. The chart also shows which continent's production increased and which fell, which is a useful comparison to make.

The columns are presented in the alphabetical order of the countries but they could actually be ordered in any way at all. Another way to present them could be by ascending order with the countries with the lowest production on the left rising to those with the highest on the right. Whatever way the data is presented, the bar chart would give the same information.

Bar charts allow data to be presented in a clear format. They are particularly useful to summarise a large amount of data in a visual format. Key values can be highlighted very quickly and they are used widely throughout the business world to show key financial data. However bar charts can oversimplify data and further explanations may be needed to give an accurate analysis of the data. Data can also be manipulated to show false results and patterns.

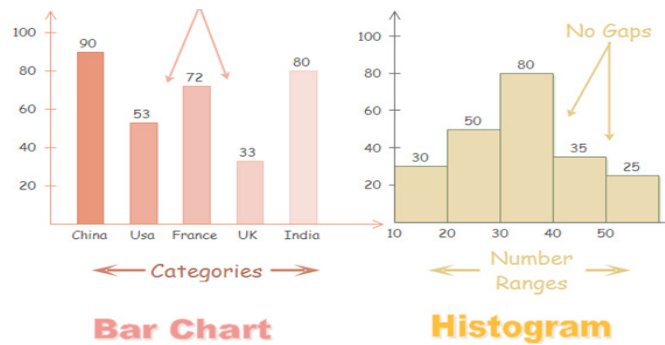
Histograms

Histograms look visually similar to bar charts. The main difference is that histograms show quantitative data and bar charts tend to show categorical data.

The data in a histogram is continuous so there are no gaps between the bars which represent the different intervals.

The area of each bar is proportional to the frequency for each interval. Both the x-axis and y-axis have a scale.

In a bar chart the height of the bar indicates the frequency of the category, and only the y-axis has a numerical scale.

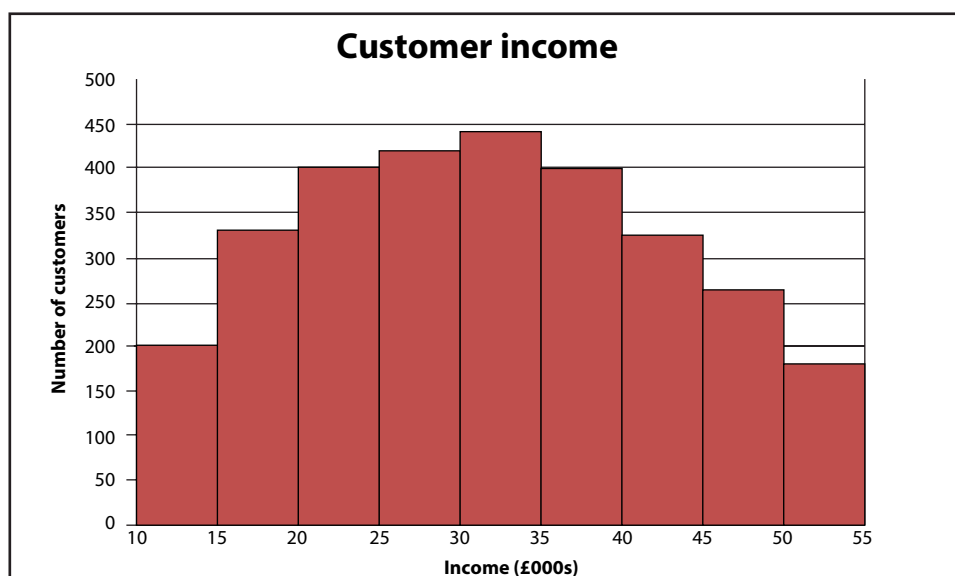


The correct name for the table from which the histogram is drawn, is a frequency table. The following table shows the frequency of results for a survey on customer incomes.

Results from a customer income survey:

Income in £000's	Frequency
10-14	200
15-19	330
20-24	400
25-29	420
30-34	440
35-39	400
40-44	325
45-49	265
50-54	180

The information in the table above can be presented in a histogram, as seen below:



The main advantage of using a histogram to display data is that it shows the shape of the distribution for a large set of data. Histograms are excellent when displaying data which has chronological categories or numerical groupings. Using histograms to display data also helps depict large differences in shape or symmetry of the data collected.

Histograms do have weaknesses when used for displaying data. Firstly histograms cannot be used for exact values as the data is grouped into intervals. Also the effectiveness of data decreases when the range of data is too wide. What would be the point of a histogram used to analyse the age of a magazines readership if the range was 0-20 years old, 21-40 years old and so on? The data is perhaps less meaningful if the groups are very large.

Line graphs

Line graphs are used to compare two variables. The X axis represents a continuous variable (e.g. time) and the Y axis represents the second variable (e.g. quantity or value). A line graph is plotted as a series of points and then joined to produce a continuous line. The line graph shows the relationship between the two variables.

The line graph below shows the price of coffee throughout the period from September 2013 to September 2014. If this information had been presented as a table it would have been difficult to easily identify the trend. As a line graph the visual representation is much clearer and raises questions such as “Why did the price of coffee rise so dramatically after February 2014?”

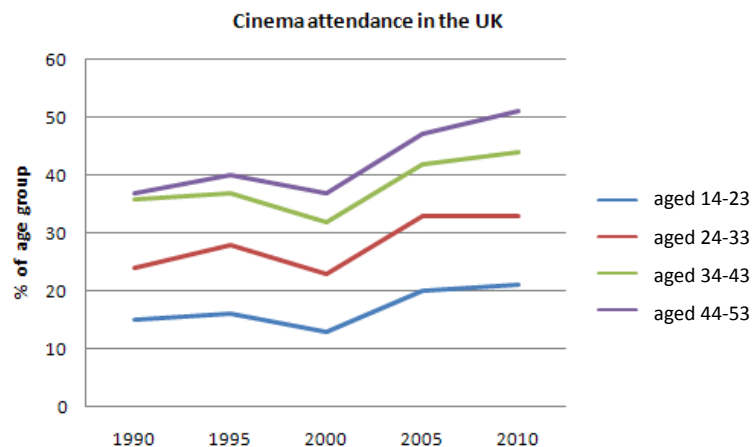
International Coffee Organisation (ICO) composite indicator daily prices (from September 2013 to September 2014)

Graph 1: ICO composite indicator daily prices

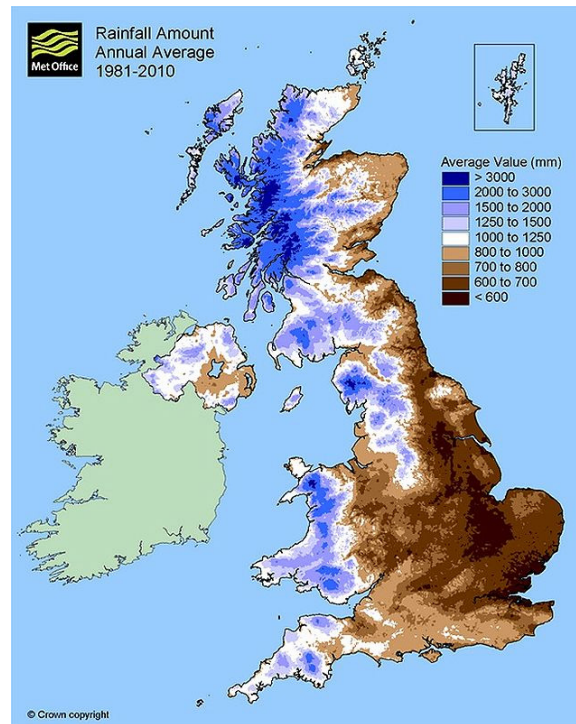


Source: www.ico.org

Line graphs can also represent data for a number of different categories. The following line graph below clearly shows the value and pattern of different age groups going to the cinema in the UK:



Line graphs are a very useful presentational tool for businesses, a vast range of market, economic and financial data can be presented to a number of stakeholders to show business performance and to analyse market trends.

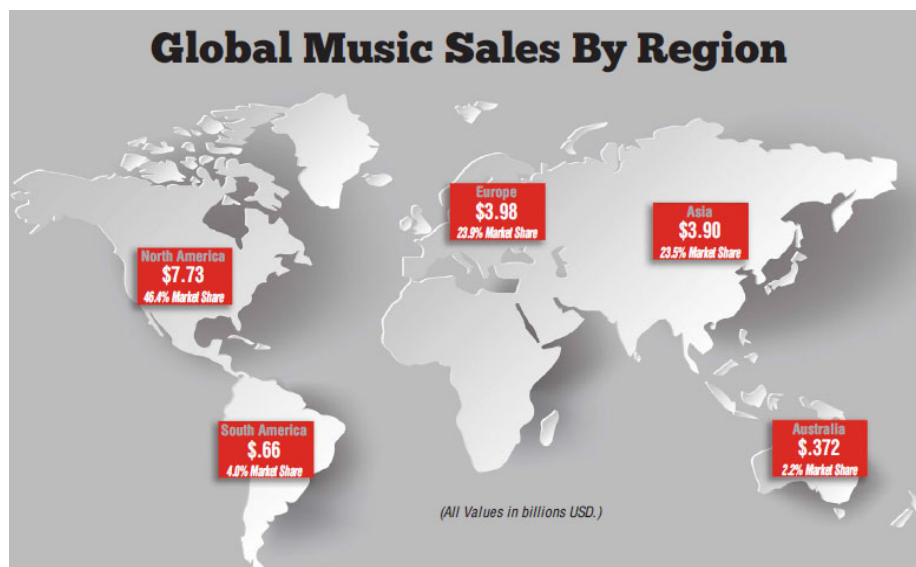


<http://www.metoffice.gov.uk/learning/rain/how-much-does-it-rain-in-the-uk>

Maps also represent a useful way of presenting information visually. The map above shows the average amount of annual rainfall in the United Kingdom between 1981 and 2010. A brief glance at the map shows that the western side of the UK received the highest levels of rainfall and with a little understanding of geography it can also be seen that it was the areas of highest ground, such as Snowdonia in Wales, which were the wettest.

This information might be useful for a business looking for somewhere relatively dry to relocate to or for a business selling umbrellas! Although the latter would be well advised to also look at a map showing the distribution of population in the UK as well, thereby revealing that additional information is needed to make informed decisions, rather than making decisions based on limited data.

The following map is a very simple yet effective way of showing the global music sales by geographical region. Although using maps are not always suitable, but when possible they can present business data in an attractive and effective way.



<http://guitarcan.com/all/music-industry-markets-sales-data/>

Index numbers

Index numbers are used in order to make numerical data easier to understand. They are a useful method of showing changes over time in data such as price levels, consumer income levels or economic output. Using index numbers allows data to be standardised over time so that the data is easily comparable. The example below explains how this works.

Let's assume that in the year 2000, the average weekly sales for a business selling ice-creams was £400. We will use the year 2000 as the 'base index number', which is most often chosen as 100. To convert a value into an index number, we use the following formula:

$$\text{Index number for any time period} = \frac{\text{Value in period}}{\text{Value in base period}} \times 100$$

So for the first year, the year 2000, (our index year) the time period is the same as the base period (the index year), which will give us:

$$\frac{400}{400} \times 100 = 100$$

In 2010 the average weekly sales had increased to £465. To turn the value of average weekly sales for 2010 into an index number we use the same formula. This gives the following:

$$\frac{465}{400} \times 100 = 116.25$$

By comparing the index numbers from the year 2000 (the base year of 100) and 2010 (116.25) we can interpret this as weekly average income increased in pure monetary terms by 16.25% (116.25-100).

Index numbers are useful in business because managers are often concerned with the way in which values such as prices paid for raw materials, numbers of employees and customers, sales, output, productivity and profits etc. may change over time. Index numbers are often used for economic analysis. Perhaps the most recognised of this sort of data are measures for inflation such as the Consumer Price Index (CPI) or the Retail Price Index (RPI).

One advantage of using index numbers is to compare changes over time and to make the value of these changes clear. For example when measuring changing prices, the price levels in the current year (indexed figure) is compared to price levels in the previous year (previous years indexed figure). This will then give us the rate of price increase – inflation. This sort of data is useful to managers as it can potentially indicate how much wages and prices should change by.

Let us consider another example using information about prices. We will assume that in the year 2013 a basket of goods cost £250. Let's have this as a base year (so the index number is 100). The following year the same basket of goods cost £260. In order to calculate the index value for 2014 we use the formula below:

$$\frac{£260}{£250} \times 100 = 104$$

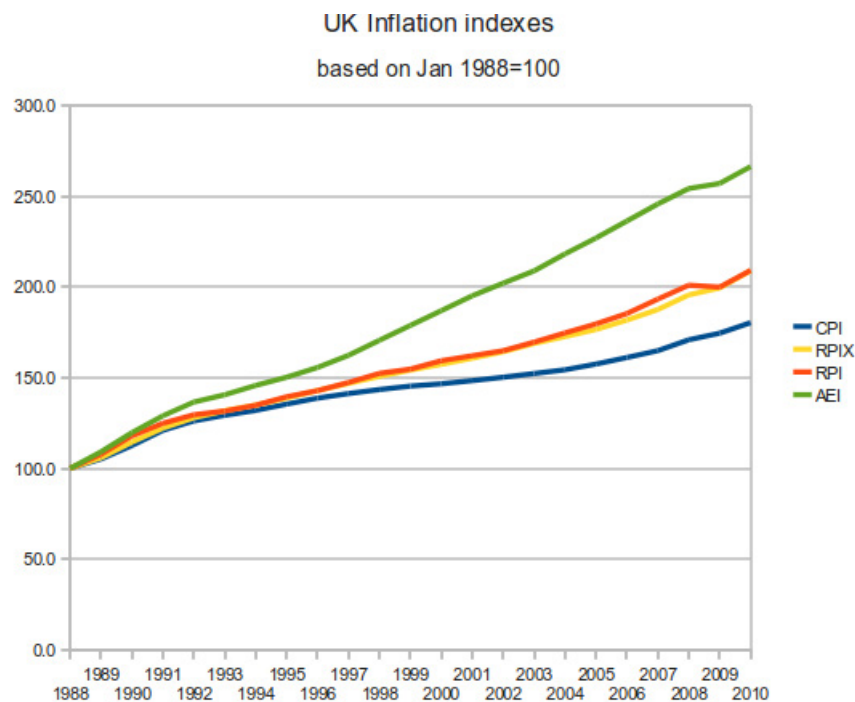
The same process can be applied to the following year (2015). If the basket of goods now cost £265 then once again we use the formula (remembering to divide always by the base year value):

$$\frac{£265}{£250} \times 100 = 106$$

The index numbers calculated above are tabulated below:

	Year 1 (2013) (Base year)	Year 2 (2014)	Year 3 (2015)
Cost of basket	£250	£260	£265
Index number	100	104	106

In this case there has been an increase in the cost of a basket of goods every year and with the use of index numbers it is easy to see that by Year 3 (2015) the prices had increased by 6 percent from the base year of 2013 (106 – 100). Index numbers can also be presented in the form of a graph, shown in the UK inflation indexes below:



https://en.wikipedia.org/wiki/Retail_price_index#/media/File:UKinflation.png

The combination of using index numbers and a line graph gives a clear presentation of the data and helps to show a clear pattern. The use of index numbers allows different sets of prices to be compared in the same graph, which would not have been possible with raw data.

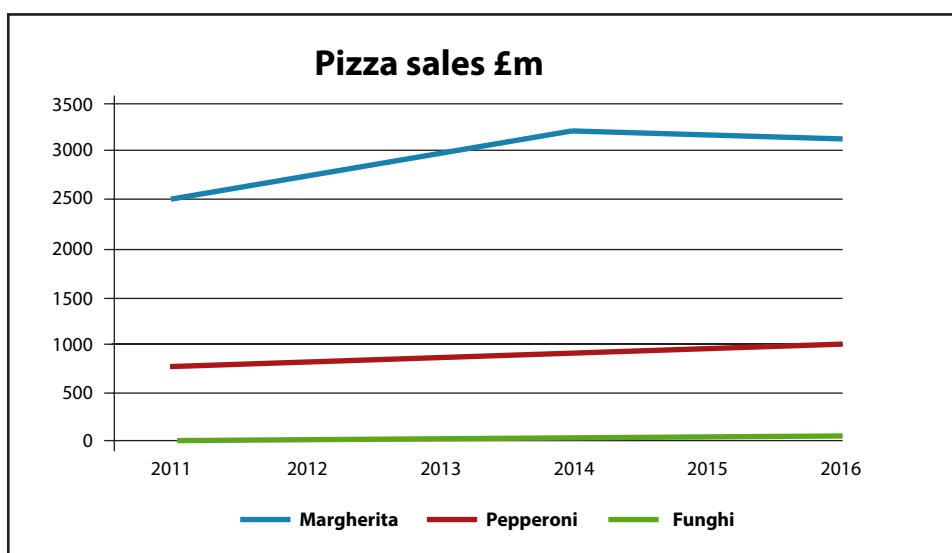
One feature of index numbers is that they allow data with different starting amounts to be easily compared so that rates of change can easily be seen.

Imagine that the pizza company mentioned previously was interested in comparing the rates of growth of different types of pizza toppings and wanted to be able to see this visually. There are a variety of ways of doing this – working out percentage changes year on year would be one – but index numbers works well.

Imagine the sales are as follows:

All figures £m	2011	2012	2013	2014	2015	2016
Margherita	£2 500	£2 750	£3 000	£3 200	£3 150	£3 100
Pepperoni	£750	£800	£850	£900	£950	£1 000
Funghi	£10	£12	£15	£18	£21	£24

If we show this as a line graph, it is very difficult to see what is going on because of the very different start figures:



If instead we use index numbers, then we set all the figures in 2011 to be the base for each pizza. Using the formula given, the index for Margherita pizzas in each year will be:

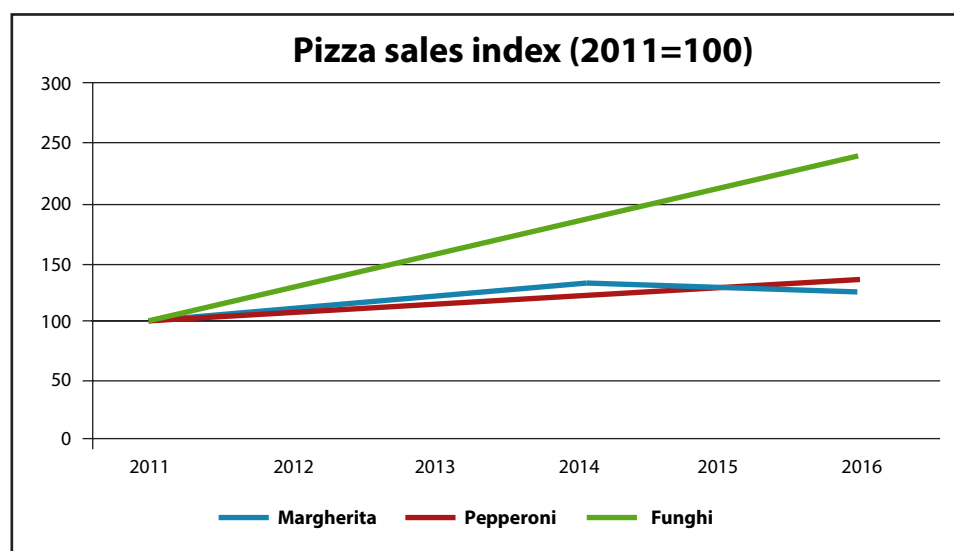
2012	2013	2014	2015	2016
$\text{£}2750/\text{£}2500$	$\text{£}3000/\text{£}2500$	$\text{£}3200/\text{£}2500$	$\text{£}3150/\text{£}2500$	$\text{£}3100/\text{£}2500$
X100	X100	X100	X100	X100
=110	=120	=128	=126	=124

(2011=100 base year)

Doing the same for the other pizzas, we end with:

All figures £m	2011	2012	2013	2014	2015	2016
Margherita	100	110	120	128	126	124
Pepperoni	100	106.7	113.3	120	126.7	133.3
Funghi	100	120	150	180	210	240

It is clear from the table that the best **relative** performer is the Funghi pizza – although this highlights one of the limitations of index numbers in that they focus on the change in data, not on the absolute levels. Margherita pizzas are still far and away the best seller but it's just that the sales of Funghi pizzas are growing fastest:



Discussion themes

Using the internet search for examples of business data presented as:

- Pie charts
- Histograms
- Line charts
- Index numbers

Copy the data into a word document and briefly interpret the data.

Carry out some primary research from your classmates to gather some categorical data, such as their favourite television programme, where they went on holiday, the colour of their eyes, favourite take-away food, etc. Collect this data in a table then produce a bar chart to display the data.

Carry out some further primary research to gather some quantitative variables, such as height, distance they live from school, amount of cups of tea they drink in a week, and so on.

Produce a frequency table and a histogram to show the results.

Data presented in tables and graphs are always more useful to business decision-makers than data presented in written form. Discuss this statement with reference to appropriate examples.

Explain why index numbers are useful to present business data.

Explain why index numbers could be misleading.

Understanding index numbers:

http://www.economicsonline.co.uk/Competitive_markets/Index_numbers.html

http://www.economicshelp.org/macroeconomics/definitions/index_numbers/