DITT
Diploma in Tax Technology

Module 2
Types of Tax Technology

Session 2.4
Data Visualisation

Learning Outcomes

At the end of this session, you will understand:

- what data visualisation is
- the main types of data visualisation tools
- what makes for good data visualisation
- how data visualisation is used in a tax context
- the limitations of data visualisation

Table of Contents

1. Introduction

1.1 Data visualisation fundamentals

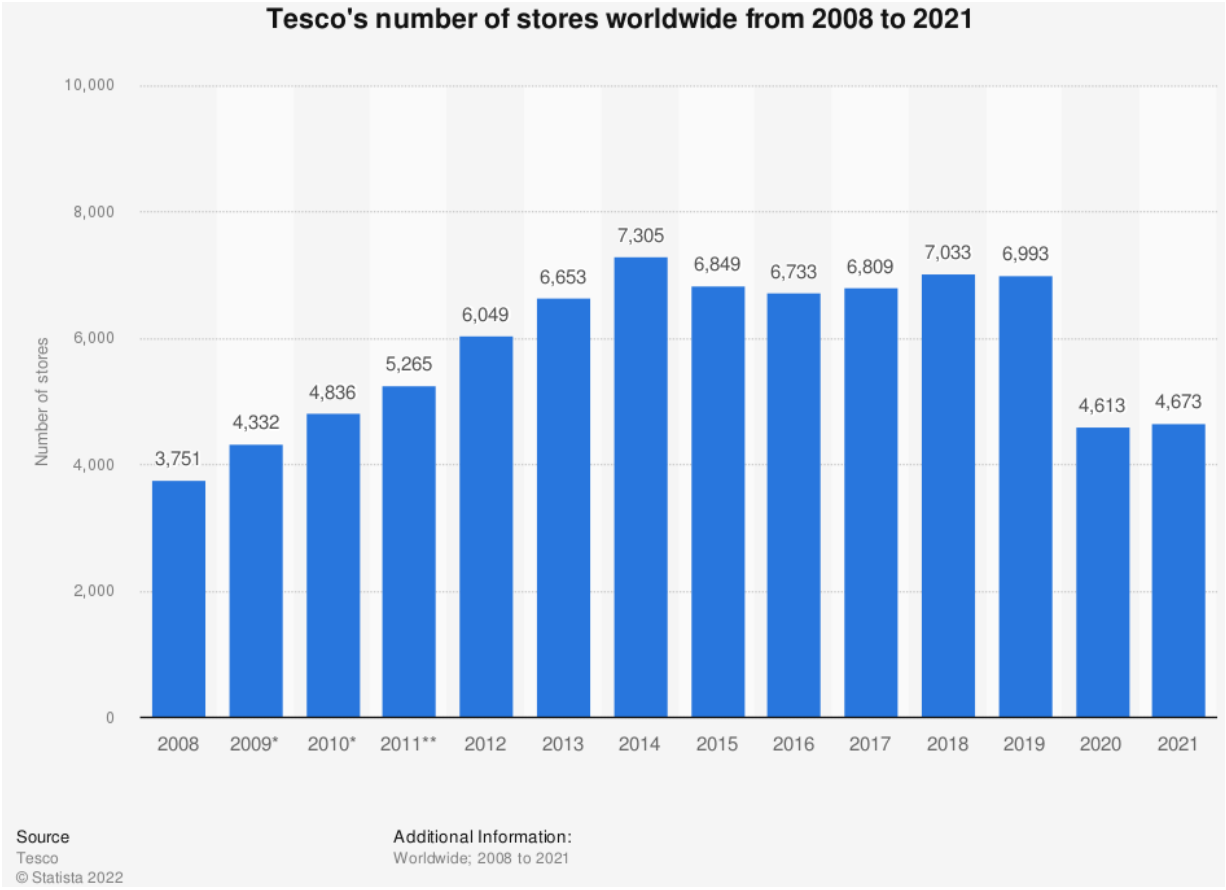
Definition

Data visualisation is the graphical representation of information and data using tools ranging from graphs and charts to animations.

It is easy for most people to understand visual displays of highly complex data, even if they are not an expert in the field. The human brain may struggle with understanding large volumes of data presented in numerical or verbal form but often finds it easier to understand information presented visually. This is because it can usually instantly distinguish colours and shapes from each other.

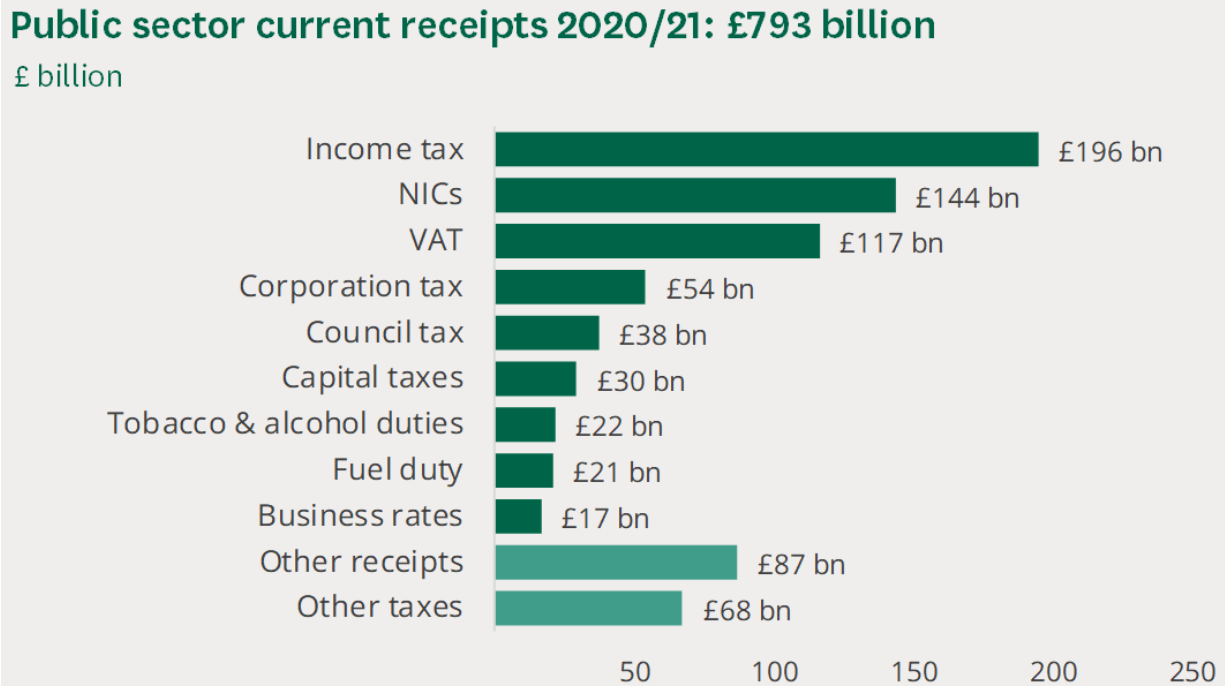
The brain can distinguish and understand the significance of line lengths, shapes, orientation, distances and colours without engaging in a significant amount of processing effort. It is therefore easier for the brain to identify patterns, trends and outliers in large data sets if they are presented in a visual form. It is far more effective to review monthly sales figures by looking at a bar chart than by examining a spreadsheet packed with data.

For example, the bar chart below shows the number of Tesco stores operating between 2008 and 2021. The reader can immediately understand the pattern in the rise and fall of store numbers. In order to make an informed decision about the status of the company – for example, is the fall in store numbers in 2020 and 2021 a result of the COVID-19 pandemic and the switch to online shopping? – the reader would need more information, but a bar chart is able to give an instant first impression.



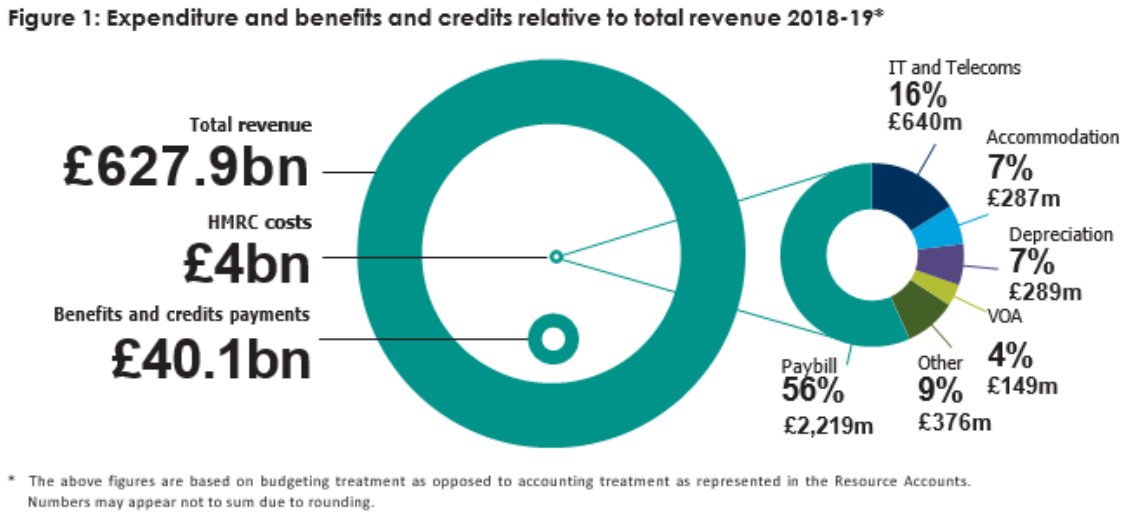
Source: Statista.com

As a second example, looking at the bar chart below from the House of Commons Library Research Briefing ‘Tax statistics: an overview’, the reader can immediately identify the various sources of public sector income. If the monetary data were presented purely as a list of figures, the reader would have to think much more carefully about them to be able to gain an initial understanding.



Source: House of Commons Library Research Briefing, 'Tax statistics: an overview': https://commonslibrary.parliament.uk/research-briefings/cbp-8513/

The same is true of the visualisation below showing a breakdown of HMRC's revenue, costs, and benefits and tax credits payments. The reader can understand immediately how the £4 billion of HMRC costs is split.



Source: HMRC, Annual Report and Accounts 2018-19: https://www.gov.uk/government/publications/hmrc-annual-report-and-accounts-2018-to-2019

The ability to present data visually like this means that it is possible for data scientists to illustrate and explain complex relationships, interactions and data-driven insights in ways that the human brain can understand. Data visualisation tools can analyse enormous volumes of information and present it in a way that enables organisations to make decisions that are driven by data. This means that decisions are less likely to be affected by bias or human perception.

Managers and others can understand visualisations far more easily and quickly than large volumes of raw data presented in a spreadsheet or in a written document. Consequently, one of the main advantages of data visualisation is more effective communication.

1.2 Benefits of data visualisation

Benefits of data visualisation include the following:

Speed	Users can read and understand information quickly, to gain better insights and to make decisions faster.
Comprehension	Users can understand the current situation and use the data to make decisions on strategy.
Engagement	If data is presented in an understandable way, users (eg an audience at a corporate presentation) will remain more focused and engaged.
Distribution	Data can be distributed to all categories of user, expert or not, without the need for further refinement; all employees of an organisation can understand results presented visually regardless of their financial acumen.

2. Data visualisation tools

Historically, non-technological tools such as pie charts, bar charts and graphs drawn by hand were used for the presentation of data. The arrival of computers brought with them the capacity to create such visualisations using data recorded in spreadsheets or other documents.

There are now numerous data visualisation tools, ranging from the simple graph or pie chart created within an Excel spreadsheet, to very intelligent tools which can create interactive methods of presentation, drawing from a wide variety of sources.

2.1 Simple data visualisation tools

Examples of simple data visualisation tools include:

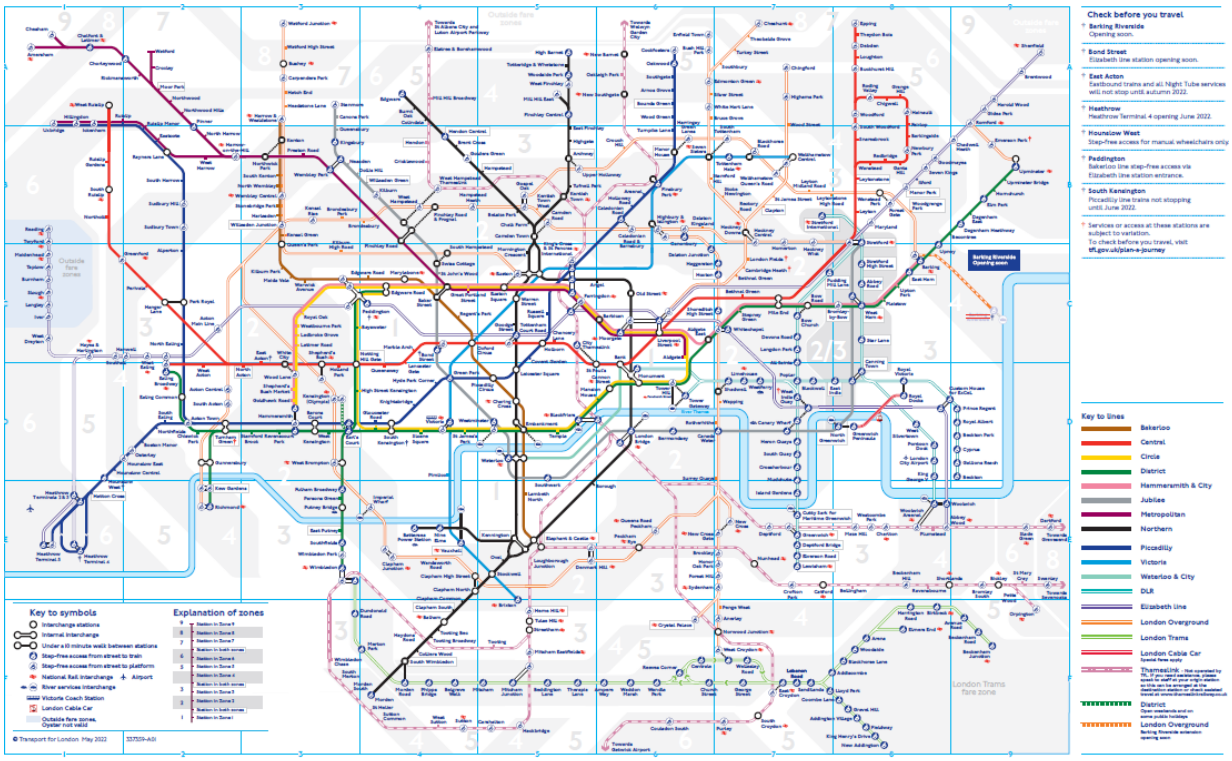
- infographics;
- heat maps;
- fever charts; and
- scatter plots

Each of these are detailed below.

2.1.1 Infographics

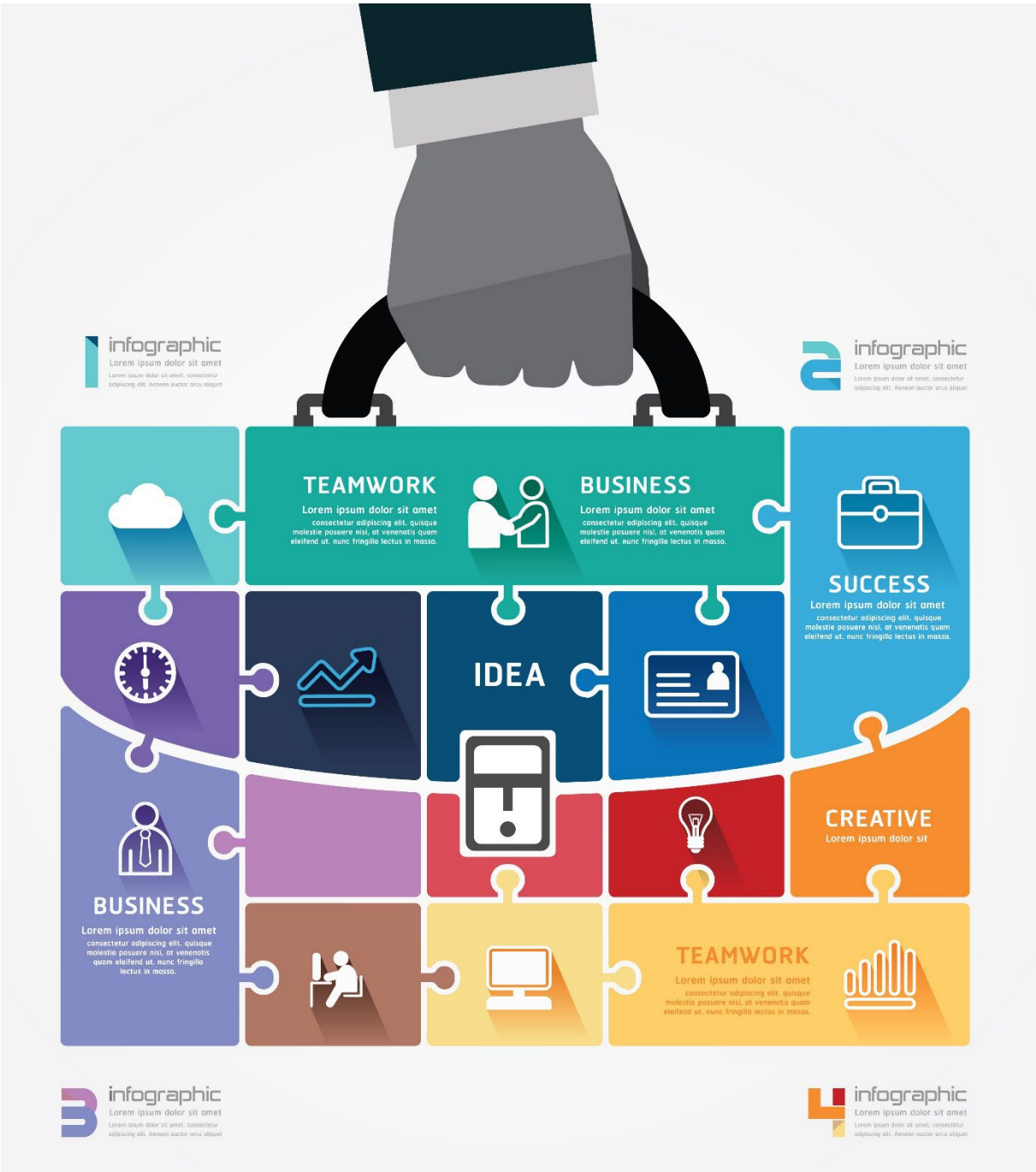
Infographics are a collection of images, data visualisations such as pie charts or graphs, with minimal text, which give an understandable visual description of a topic.

The well-known map of the London Underground, first designed by Harry Beck in 1933 and updated in 2022 to include the Elizabeth line, while not geographically accurate is an example of an extremely useful infographic.

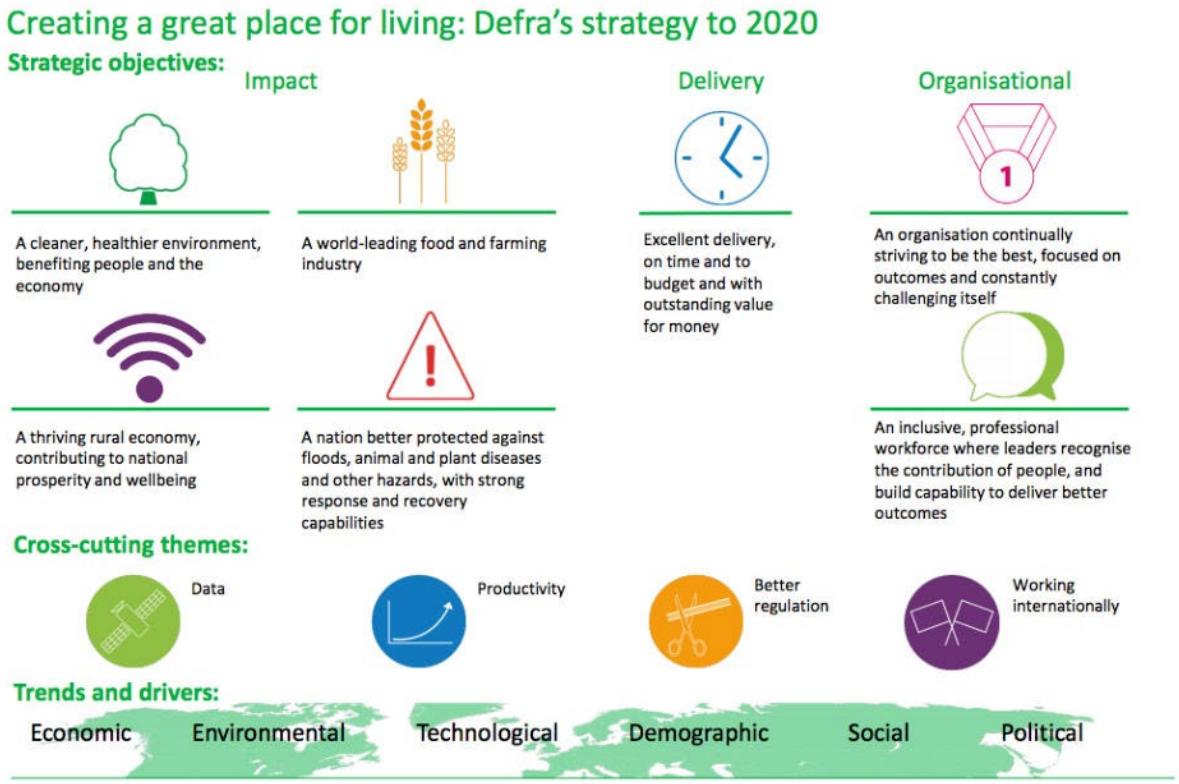


Source: Transport for London: www.tfl.gov.uk

Here is another example of an infographic:



The infographic below shows how images and minimal text can be combined to convey information clearly, with the choice of graphic indicating the subject area of the text.



Source: Department for Environment, Food and Rural Affairs (DEFRA): https://www.gov.uk/government/organisations/department-for-environment-food-rural-affairs

2.1.2 Heat maps

Heat maps are visual depictions of data in which values are indicated by colour. They are often used in the analysis of website usage as they can show user behaviour on particular webpages or templates. For example, they can show where users clicked on a page or how far down the page they scrolled. This can help in the analysis of user behaviour and in the design of improvements to the effectiveness of the site.

Heat maps are useful for identifying outliers, or anomalies in data. Auditors use heat maps as part of data analytics software to identify items of higher risk that may therefore require more attention from the auditor. A heat map can show an analysis of an organisation's transactions, with all transactions that are under a pre-set amount or frequency appearing in a green, or low risk, area.

A heat map can indicate transactions of a significant or unusual amount as outliers in a red area, as seen in the following example.

Source: ICAEW 2022 Guidance Notes

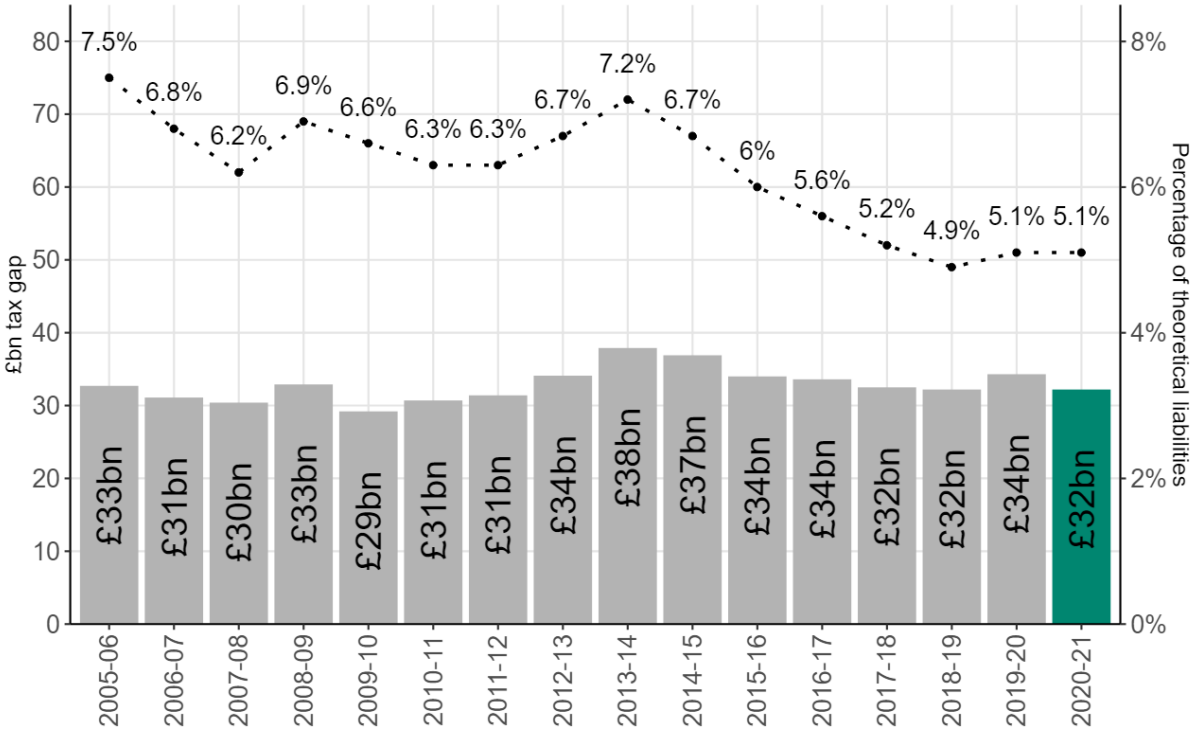
2.1.3 Fever charts (or time series charts)

Fever charts are graphical representations that show a change in data over time. They are often referred to as 'time series charts', as they plot data at particular points in time.

It is easy to see a trend or pattern using a fever chart, so they are used for data that is changing continuously, such as share prices.

The example below shows the changes in the value of the UK tax gap (ie the difference between HMRC's anticipated tax revenue and the amount it actually collected) compared to the percentage of theoretical tax liabilities over a period of years.

Figure 1.1: Tax gap by value and as a percentage of theoretical tax liabilities, 2005 to 2006 up to 2020 to 2021



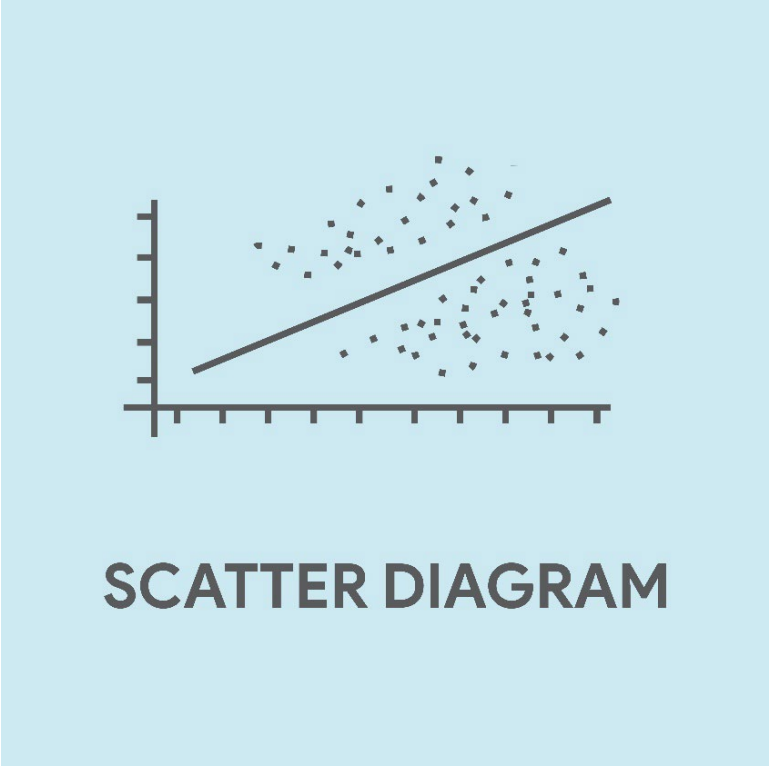
Source: HMRC statistics
https://www.gov.uk/government/statistics/measuring-tax-gaps/1-tax-gaps-summary

2.1.4 Scatter plots

Scatter plots (also known as scatter charts or scatter graphs) display the relationship between two variables. They are presented in the form of an x axis and y axis with dots to represent data points and a line drawn to show the trend. They are used to determine whether two variables have a relationship or correlation.

For example, the variables could be the weight and height of individuals or groups of people. The scatter plot would indicate whether weight and height are connected; for example, the taller a person is, the more they are likely to weigh.

An example is shown below.



2.2 Data visualisation technological developments

The development of technology and more advanced software means that more sophisticated data visualisations can now be produced. Data scientists and designers can easily communicate complex information in a useful and meaningful way.

There are numerous out-of-the-box packages, and some of the most common data visualisation tools are discussed below.

2.2.1 Tableau

Tableau is a commercial data visualisation tool that can create interactive graphs, charts and maps.

Its advantages are as follows:

- Users can connect to different sources of data simultaneously.
- Images are created quickly.
- Collaboration and sharing with other users is possible.

- There is a free-to-use version, Tableau Public, or licence packages, meaning that is a viable solution for all types of business, even if small.

The main disadvantage is that it can take time to become proficient in its use. However, this is the case with all data visualisation software, plus there are many video tutorials available.

https://www.tableau.com/en-gb

2.2.2 QlikView

QlikView is described by some users as a 'data discovery platform' with the ability to reveal new business insights. It also has the ability to combine various types of data sources and produce multifaceted and multicoloured visualisations.

It uses a 'drag and drop' visualisation interface, enabling users to add data from different sources without the need for additional coding.

https://www.qlik.com/us/products/qlikview

2.2.3 Microsoft Power BI

Microsoft Power BI is a range of business intelligence (BI), reporting and data visualisation products and services that integrate with other Microsoft products that can be used to create charts, graphs, reports and dashboards using predictive analytics (see session 2.3).

It can create reports and share insights with others in the organisation via an end-user platform and acts as a centralised repository for all an organisation's business data, which is accessible by all business users.

Users can establish visualisation tools to generate automatic dashboards that track organisational performance across key performance indicators (KPIs) and visually interpret the results.

https://powerbi.microsoft.com/en-gb/

2.2.4 Datawrapper

Specifically designed for newsroom data visualisation, Datawrapper software allows users to create charts and maps directly in their browser by uploading files. It is a free resource, but there are limitations on usage if the user does not enter a licence agreement.

https://www.datawrapper.de/

2.2.5 Google Charts

Google Charts is a free data visualisation tool, specifically for creating interactive charts for embedding online, which publishes to browsers without the use of any additional plugins. Data can be extracted from various sources, such as Google Spreadsheets and Google Fusion Tables.

https://developers.google.com/chart

3. Effective data visualisation

If data visualisations are to be valuable to the user, some thought should go into their preparation, taking into account the needs of the end user and the purpose of the data. If these factors are not considered, the information communicated may be of limited value.

According to Edward Tufte in his book *The Visual Display of Quantitative Information,* an effective graphical display should:

- show the data concerned, but also encourage the reader towards thinking about the substance, ie the message of the data, rather than how it was produced;
- not distort the message being given by the data, for example by making graphics and colour completely dominate, so the user loses sight of the core information;
- present a large volume of data in a compact space using a meaningful method of presentation;
- encourage the eye and the brain to compare different pieces of data, enabling the reader to establish links and correlations;

- enable data to be displayed at different levels of detail, ranging from a high-level overview to a fine level of detail; and

- arguably most importantly, make large data sets coherent and understandable to the end user, enabling decisions to be made as a result.

Points to note

- When an organisation is considering adopting data visualisation techniques, management should identify how the available tools will provide the most benefit for the organisation – for example, in making decisions on fixed assets and depreciation, the approach to indirect taxes or transfer pricing.

- Management should identify processes which are time-consuming, use large volumes of data and require reporting that is difficult to analyse. They should identify the sources of data, decide and clearly record what the data is required to show and identify existing skills sets in the organisation that can start to design and create an analytics platform. It is likely that this will involve capital expenditure, so return on investment calculations and justification of the proposed investment will be needed.

To be effective, visualisations need the data to explain the information clearly and honestly.

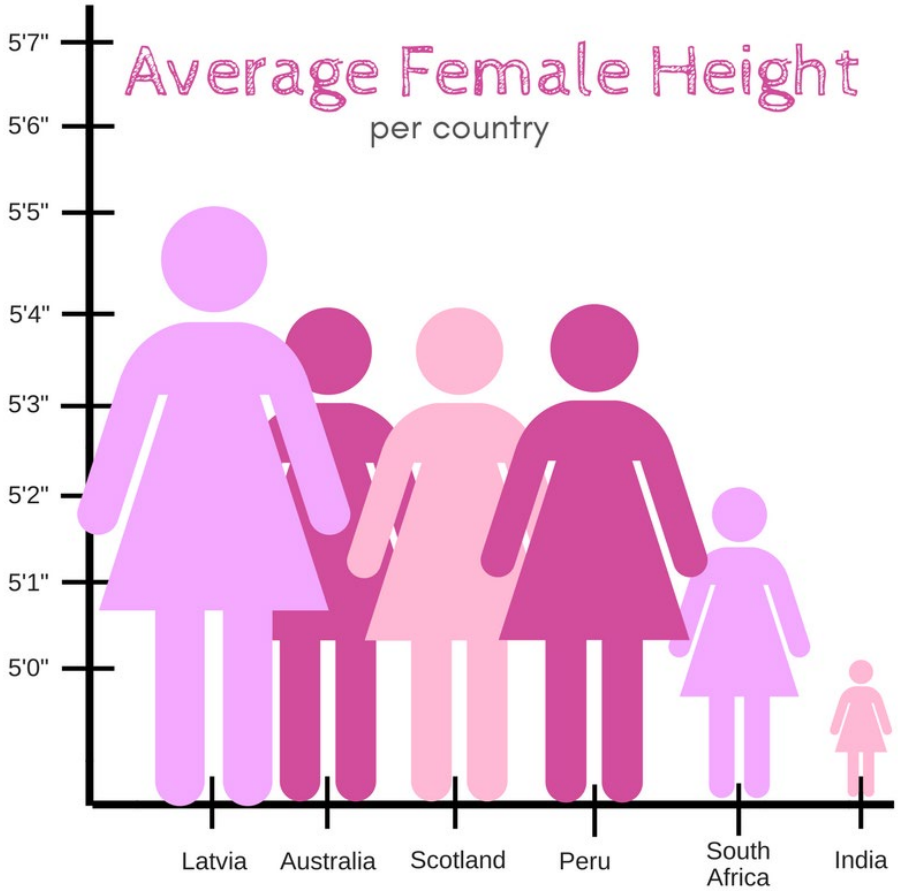
Illustration 1

In spring 2022 the UK government and health authorities produced graphs showing significant falls in the rate of COVID-19 infections.

The graphs did not explain that, with effect from 1 April 2022, free access to COVID-19 tests and the requirement to take a test if displaying symptoms had ended, inevitably causing a reduction in the number of positive cases since far fewer people were taking tests. The data visualisations did not tell the whole story.

The apparent message given by data can be skewed. For example, if one of the axes of a graph is truncated, so the two axes do not correlate, the data can be rendered practically meaningless.

Similarly, if images are not presented consistently, the results lose their impact, as can be seen in this visualisation of the average height of women in different countries.



4. Data visualisation in tax

Large corporate groups can find it challenging to organise and review their tax data due to its volume and complexity and the fact that it may be stored and submitted in different ways in different jurisdictions. It is difficult to identify and then communicate trends and patterns, meaning that risks may not be detected and opportunities for tax planning may be missed.

The development of technology-enabled data visualisation techniques means that highly complex and disparate sets of data can be organised to enable an overview of the position, particularly in relation to compliance, in different jurisdictions.

Challenges, risks and opportunities can be identified and decisions made that are genuinely data-driven, rather than being based on incomplete information and potentially intuition or even guesswork.

As a result, the decision-making skills within the tax departments of firms and their clients, and the ability of tax practitioners to communicate complex tax information to non-experts, all improve.

Example 1

GiantCorp plc is a large multinational conglomerate which owns thousands of fixed assets in different countries, all accounted for in different ways by different departments. Until the development of more sophisticated data visualisation tools, the company would have produced individual spreadsheets or charts relating to depreciation.

Data visualisation tools, such as a heat map, can show the same information visually, enabling high cost or high value outliers requiring further review to be quickly identified.

The accountancy firm Grant Thornton recently produced a report commenting that:

> 'An agile tax department that leverages the power of analytics and visualization makes highly informed decisions, engages leadership more deeply through better understanding, reduces risk, and enhances productivity. The result is not only more informed decisions, but the tax department bringing increased value to its own organization.'

5. Limitations of data visualisation

Although data visualisation tools can be useful for business and tax professionals, and have arguably revolutionised many processes, it is important to be aware of some of the potential drawbacks of using them.

- Poor-quality source data will cause the visualisations to be meaningless or misleading, potentially increasing the risk of poor decision-making.

- Similarly, the source data could be misleading if it has been manipulated, eg by management who want to hide or distort a trend in the results.

- If the data visualisation tool selected is not appropriate for the data concerned, the resulting visualisations could be misleading.

- If the graphics used are badly designed or too complicated, the tax practitioner or client will not be able to understand what they are intended to show.

- If the size of the data set is small, the results may not be particularly conclusive and therefore may not form a good basis for making decisions.

- The visualisation can be so creative, with use of images, colour, interactive features and so on, that the underlying message of the data is obscured.

Additional resources

You may find these resources useful if you would like to explore this subject further:

The Business Journals article, 'Visualizing tax data can boost insights into your business' (2017):
https://www.bizjournals.com/bizjournals/news/2017/10/19/visualizing-tax-data-can-boost-insights-into-your.html

KPMG briefing, 'Tax, data and analytics – moving from control to transformation' (2018):
https://assets.kpmg/content/dam/kpmg/xx/pdf/2018/01/tax-data-and-analytics.pdf

Analysis Function publication, 'Data visualisation: charts' (2022):
https://analysisfunction.civilservice.gov.uk/policy-store/data-visualisation-charts/

References

References from within this session are replicated here in one list:

Tableau, a commercial data visualisation tool that can create interactive graphs, charts and maps:
https://www.tableau.com/en-gb

QlikView, a ‘data discovery platform’ with the ability to reveal new business insights:
https://www.qlik.com/us/products/qlikview

Microsoft Power BI, a range of business intelligence (BI), reporting and data visualisation products and services:
https://powerbi.microsoft.com/en-gb/

Datawrapper, software which allows users to create charts and maps directly in their browser by uploading files:
https://www.datawrapper.de/

Google Charts, Google’s free data visualisation tool, specifically for creating interactive charts for embedding online, which publishes to browsers without the use of any additional plugins:
https://developers.google.com/chart

Review

Are you confident that you understand these concepts and could explain them to a non-technical colleague? Refer to the checklist below to help you assess this:

What is meant by data visualisation	☐
Types of data visualisation techniques	☐
What makes for good data visualisation	☐
The limitations of data visualisation	☐

Self-test questions

Complete the following questions to test your knowledge.

1. Are the following statements **true** or **false**?

1) Data visualisation is always meaningful and useful if a wide variety of colours and images are used.

2) Not all data visualisation techniques are useful in a given situation; organisations should be selective in their use.

2. Complete the sentence.

Heat maps are useful for…

A identifying trends and patterns in data.
B identifying outliers in data sets.
C depicting data as it changes over time.
D explaining the reasons for changes in data.

Solutions to self-test questions

1.

1) False – Too much colour and different imagery can cause data visualisations to be less meaningful and more confusing.

2) True – The most suitable data visualisation for the situation should be selected, eg time series for showing information over time.

2. B – Heat maps are useful for identifying outliers in data sets.