An introduction to data analysis

Course Overview

Summary

Having spent a lot of time gathering and preparing data, the temptation is to immediately begin to analyse the data using any technique that seems relevant. First, however, it is important to explore the data to discover its shape and any limitations it might have. We also need to understand the benefits and risks of applying each statistical technique to data.

So before we perform any analysis, this course looks at three key things:

1. How numbers and summary statistics can lead us to incorrect conclusions in "Seeing through a world of numbers"
2. Why establishing "the shape of data" is critical to discover its limitations and biases.
3. Why plotting "Trends in data" is more challenging than taking a ruler to a plot.

Following this we'll then look at how some key statistical techniques and tools can be applied to data.

In this course, we'll be analysing the performance of the London Fire Brigade and looking at the impact on the service of closing 10 fire stations in London in 2014 in order to cut back on spending. Was it a good idea? What has happened since? Is the performance good or bad? Is it fair? What else should happen now?

This course introduces the skills to perform these tasks in order to see what can be discovered from the data.

Learning Outcomes

By the end of this course, you will have the skills to evaluate data from different sources in order to establish its quality, and carry out exploratory data analysis to produce insights.

To achieve this, you will:

- Evaluate the benefits and risks of applying different statistical techniques to data
- Apply several exploratory data analysis techniques to create summary statistics and visual representations
- Practise the skills to establish the shape of data and examine what it means, and how it can be used to determine the state of data
- Use tools to filter and explore trends in data
- Evaluate the difference between qualitative and quantitative data analysis

Learning Experience

Number of modules	6 (+ reflective workbook)
Modality	Asynchronous / Self-directed / Online
Notional learning hours	3.5 hours (total)
Assessment	Formative
Certificate	Certificate of completion

Each module contains learning content that introduces the key concepts in the module, providing examples and case studies that demonstrate these concepts in practice. Each module contains a series of formative questions to support your learning. Learning is applied in activities throughout, and provides a basis for undertaking exploratory data analysis, producing summary statistics and establishing the shape of data. These skills are then applied through an activity that asks the learner to analyse the performance of the London Fire Brigade.

Module Summary

Module Name	Description
Seeing through a world of data	Public spending, health risks, environmental disasters, who is rich, who is poor, pensions, the best and worst schools and hospitals, immigration – life comes in numbers. The trick is seeing through them. To some, numbers can be intimidating. They can also create illusions and lead us to see what is not actually there. This module reveals what the numbers really show and exposes the illusions. This module looks at: • Counting, numbers and chance • Fluctuation, averages and targets • Sampling and hypothesis testing • Correlation and percentages
The shape of data	Data, it turns out, has shape. That shape has meaning. The shape of data tells you everything you need to know about your data from its obvious features to its deepest secrets. Establishing the shape and meaning of data prior to using it in analysis is essential to avoid basic mistakes that can lead to harmful impacts. This module looks at some of the key methods to look at the shape and characteristics of quantitative (numeric) data including: • Averages • 5 number summary • Distributions and histograms
Trends in data	One of the key benefits of data is that of finding insight, spotting patterns and being able to make predictions.

	However, there are risks with mixing descriptive statistics, used for summarising existing data, and inferential statistics which focuses on drawing conclusions, making predictions, or generalising from a sample to a larger population. This module explores: • The difference between descriptive and inferential statistics • Inferential statistics applied to data • How to avoid overfitting • Working with confidence
Filtering and pivot tables	Exploring data in spreadsheet applications can often be overwhelming, especially in the cases where the data is multi-dimensional (lots of columns) or contains a lot of data (by rows). In this module we will learn how to: • Sort and filter data in a spreadsheet • Apply formula to generate simple statistics • Create a number of pivot tables • Create a number of summary graphs
Analyse the performance of the London Fire Brigade	In this module you will be focusing on data analysis through the lens of the London Fire Brigade's operations and the impact of the 2014 station closures. You will be guided through real-world data, enhancing your analytical skills and your ability to interpret and present complex information.
Introduction to qualitative data analysis	Not all data comes in the form of structured tables or accurately located geographic data. Often it can be qualitative data that is the hardest to analyse and work with. Qualitative data is information that cannot be measured and is very subjective as a result. Even colour can be subjective. The aim of qualitative data analysis is to reduce and make sense of vast amounts of information, often from different

	sources, and to offer an explanation, interpretation or thematic summary of the data. Inputs to qualitative analysis can take many forms including interview transcripts, documents, blogs, surveys, pictures or videos. Qualitative data analysis is a more natural process for humans who naturally seek to distil inputs into themes and key outcomes, as is especially true of meetings or focus groups. People will often use mind-mapping or post-it based thought maps to help group together and categorise wide-ranging discussion into key themes. Qualitative data analysis ought to pay attention to the spoken word's context, consistency and contradictions of views, frequency and intensity of comments, their specificity as well as emerging themes and trends.