

# Bias and Fairness in Data and AI

## *Course Overview*

### Summary

Ensuring ethical data collection is critical, and a significant level of scrutiny revolves around the data's origins and the decisions made during the collection phase. Nevertheless, even after diligent efforts to mitigate bias during collection, we may still experience bias and other ethical concerns. It's important not to rush into utilisation without subjecting data to a thorough examination of its quality and integrity, where you could reveal concealed biases.

In Bias and Fairness in Data and AI, you explore the fundamental concept of bias and its implications in data collection, analysis, and AI applications. You will develop the skills to detect bias at various stages of the data lifecycle, understanding the deceptive power of numbers, and statistical techniques to uncover hidden truths.

Whether you're a data scientist, a machine learning engineer, a policymaker, or simply curious about the impact of data and AI on society, this course is tailored to provide valuable insights and practical strategies.

### Learning Outcomes

The purpose of this module is to demonstrate how bias can exist within data. Through this, you will develop skills in the practices that promote fairness in automated systems, such as AI.

To achieve this, you will:

- Develop your understanding of bias in data and algorithms and actively identify instances of bias in real-world datasets
- Assess the impact and consequences of bias in automated decision-making



- Apply practical techniques to mitigate bias and promise fairness in models to support ethical data practices

## Learning Experience

<b>Number of modules</b>	6
<b>Modality</b>	Asynchronous / Self-directed / Online
<b>Notional learning hours</b>	3 hours (total)
<b>Assessment</b>	Formative
<b>Certificate</b>	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 consolidated in the creation of a classification system in our Machine-Learning simulation.

## Module Summary

Module Name	Description
<b>Bias and the data lifecycle</b>	In this module, we will delve into the crucial task of detecting bias in data, with a primary focus on identifying bias during the data preparation phase. Bias in data can arise from various sources throughout its lifecycle, from collection to retention.

<p><b>Seeing through a world of numbers</b></p>	<p>Numbers, often regarded as daunting by some, have the potential not only to inform but also to deceive. They possess the power to craft illusions and steer us towards conclusions that may not align with reality. This module is designed to unveil the true narratives hidden within the numbers and dispel any misconceptions:</p> <ul style="list-style-type: none"> <li>● Counting, numbers, and chance</li> <li>● Fluctuation, averages, and targets</li> <li>● Sampling and hypothesis-testing</li> <li>● Correlation and percentages</li> </ul>
<p><b>Descriptive statistics and bias detection</b></p>	<p>Understanding the shape and characteristics of data is essential in the realm of machine learning and AI, especially when it comes to detecting potential bias in data.</p> <p>Failing to grasp these fundamentals can lead to erroneous conclusions and potentially harmful consequences for individuals and society as a whole. In this module, we'll explore the essential methods for examining the shape of quantitative (numeric) data including:</p> <ul style="list-style-type: none"> <li>● Averages</li> <li>● 5 number summary</li> <li>● Distributions and histograms</li> </ul>
<p><b>Inferential statistics: making inferences from data</b></p>	<p>Data analysis offers valuable insights, pattern recognition, and predictive capabilities, but it also presents risks when confusing descriptive and inferential statistics. This module delves into the distinction between descriptive and inferential statistics, emphasising the importance of responsible data ethics and considerations for AI applications:</p> <ul style="list-style-type: none"> <li>● The difference between descriptive and inferential statistics</li> <li>● Inferential statistics techniques</li> <li>● How to avoid overfitting</li> <li>● Working with confidence</li> </ul>

<p><b>Building your own classification system</b></p>	<p>In the real world, data-driven decision-making is at the heart of many critical choices, from public policies to business strategies.</p> <p>In this challenge we bring you into a scenario where you must create a decision tree—a fundamental tool in machine learning—to classify properties as either being in New York or San Francisco.</p> <p>Your decision-making process will be based on a limited set of characteristics. The previous modules have equipped you with the essential knowledge to tackle this challenge. You've learned how to detect bias in data, interpret numerical insights, understand data shapes, and distinguish between descriptive and inferential statistics. These skills are not just theoretical; they have real-world applications, especially when it comes to building and evaluating AI systems.</p>
<p><b>Bias and fairness: is it possible to be fair?</b></p>	<p>In this module we will look at whether it is possible to mitigate bias and promote fairness in models to support responsible data practices. We will use a case study developed by Microsoft and EY that shows the technical approaches to address fairness in automated systems.</p> <ul style="list-style-type: none"> <li>● Analysing for bias in a data and AI system</li> <li>● The types of bias mitigation techniques</li> <li>● The trade-offs and challenges of using fairness techniques</li> <li>● Application of two technical approaches to ensure fairness in a use case</li> <li>● Choosing a fairness approach</li> </ul>