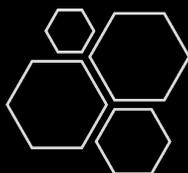


JEN SCHELLINCK    TRISTAN SHAEEN  
PATRICK BOILY    STEPHEN DAVIES    JOHN STROUD

# THE LITTLE HANDBOOK OF DECISION- MAKING WITH DATA

BRINGING YOU INTO THE WORLD OF DATA



SYSABEE



SCHELLINCK, SHAEEN, BOILY, DAVIES, STROUD

THE LITTLE HANDBOOK  
OF DECISION-  
MAKING WITH DATA

BRINGING YOU INTO THE WORLD OF DATA

SYSABEE

First edition, 2026

Schellinck, Shaen, Boily, Davies, Stroud

This work is licensed under a [Creative Commons](https://creativecommons.org/licenses/by-sa/4.0/) “Attribution-ShareAlike 4.0 International” license.



Published by Sysabee

# PRELUDE: DATA AND THE DECISION LANDSCAPE

We want this handbook to serve several purposes. Specifically, we want it to help people:

- become better at decision-making and at using data and analysis to support decisions;
- understand how data and analysis fit into the broader decision-making context;
- know when and how it is appropriate to use data and data analysis in decision-making;
- understand how to work with others to make decisions, and how different roles, including data and analysis roles, fit into the larger picture; and
- strengthen their decision theory expertise.

In serving these purposes, this handbook tackles two main questions:

1. **How** do I make evidence-, data-, or information-based decisions, individually or in a group?
2. **When** should I use evidence-, data-, or information-based decision-making?

With this in mind, the handbook has two intended audiences: (1) decision-makers who want to learn how data and data work can support

good decisions, and (2) people who work with data and want more insight into how data and data work fits into the decision-making process, and how best to support decision-makers.

To make good decisions with data, it's important to understand decision-making itself. Accordingly, the handbook is roughly 40% about decision-making, 20% about data and related topics (e.g., analysis, data engineering), and 40% about combining decision-making with data to make good decisions. The book is organized as a set of reference chapters, each providing a brief introduction to key concepts related to its main topic. The material is intended as a high-level starting point: choose the topic that seems most relevant and go to that chapter or section for an overview. At the same time, the chapters are arranged so the book can be read cover-to-cover, with each chapter building on earlier concepts while introducing new ones.

We hope this handbook will give you tools and ideas to think about decision-making in a new and comprehensive way, and to integrate data into your decision-making practices in order to arrive at good decisions. In presenting these topics, we provide our own opinions and theories about data-supported decision-making, while also drawing on the work of others in this space.

To fully understand the role of data in decision-making, a conceptual journey is required. You must first travel through the land of decision-making itself, before reaching the mountain of data and then finally the peak of analysis, at which point you can descend back down to decision-making again. Along the way, you may discover that you are already an accomplished decision-maker, perhaps without even realizing it. You may simply have needed labels and concepts that help reveal and communicate that existing know-how. We hope this handbook supports you on your journey.



# CONTENTS

<i>Prelude: Data and The Decision Landscape</i>	iii
<i>I. The History and People of Decision-Making</i>	I
History of Modern Decision-Making . . . . .	1
Decision-Making for Management . . . . .	1
Decision Makers are Not Perfect . . . . .	2
Decision Analysis . . . . .	3
Evidence-Based Decision-Making . . . . .	5
Modern Decision Science: Key Players and Study Areas . .	5
Chapter Essentials . . . . .	7
References . . . . .	8
<i>II. Three Approaches to Decision-Making</i>	9
Armchair Decision-Making . . . . .	10
Data-Inspired Decision-Making . . . . .	12
Data-Driven Decision-Making . . . . .	13
When to Use Data-Supported Decision-Making . . . . .	13
No-Data, No-Thought Decision-Making . . . . .	14
Kozyrkov’s Take on When Data Should Get Involved in Decision-Making . . . . .	15
Our Take on When Data Should Get Involved: The Mustard in the Sandwich . . . . .	17
Chapter Essentials . . . . .	18
References . . . . .	18
<i>III. Decision-Making States and Elements</i>	19
Three Decision-Making States . . . . .	19

Core Elements of Active Decision-Making . . . . .	20
Element: Problem, Goal, Target . . . . .	21
Element: Choices, Possible Actions . . . . .	22
Element: Outputs, Results, Outcomes, Consequences . . . . .	24
Element: Information, Data, Knowledge . . . . .	25
Chapter Essentials . . . . .	31
References . . . . .	32
<i>IV. Analysis Techniques</i> . . . . .	33
Data Visualization and Basic Analytics . . . . .	34
Metrics . . . . .	35
Statistics and Modeling . . . . .	37
Optimization . . . . .	40
Machine Learning . . . . .	41
Generative AI as an Analysis Technique? . . . . .	42
Analytics as Question Answering . . . . .	43
Chapter Essentials . . . . .	46
References . . . . .	48
<i>V. Group Decision-Making Frameworks</i> . . . . .	49
A Selection of Decision-Making Frameworks . . . . .	50
DECIDE Framework . . . . .	51
ACE Framework . . . . .	52
Government of Canada / Health Canada Framework . . . . .	52
Synthesizing Frameworks . . . . .	54
Decision-Making Roles . . . . .	55
Chapter Essentials . . . . .	56
References . . . . .	57
<i>VI. Human Factors</i> . . . . .	59
Cognitive Biases . . . . .	60
The Anchoring Bias . . . . .	62
Bandwagon Effect . . . . .	63
Sunk Cost Fallacy . . . . .	63
Decision-Making When There Is No (or Not Enough) Data . . . . .	65
Bad Faith Decision-Making . . . . .	66
References . . . . .	67
Chapter Essentials . . . . .	68

<i>VII. Data Storytelling</i>	69
When to Use Stories . . . . .	70
Stages of Data Storytelling Development . . . . .	70
Data Visualizations for Storytelling . . . . .	72
Data Storytelling Presentations . . . . .	73
Data Storytelling Dashboards . . . . .	74
Data Storytelling Risks . . . . .	76
Chapter Essentials . . . . .	78
References . . . . .	78
 <i>VIII. Data Engineering for Decision-Making</i>	 79
Data Pipelines . . . . .	79
Desktop Data Pipelines . . . . .	81
Enterprise Data Pipelines . . . . .	82
Data Pipeline Roles . . . . .	82
Data Governance . . . . .	84
Data Use Principles . . . . .	85
References . . . . .	86
Chapter Essentials . . . . .	87



# I

## THE HISTORY AND PEOPLE OF DECISION-MAKING

### I.1 HISTORY OF MODERN DECISION-MAKING

There are three historically distinct but conceptually overlapping strands in the modern decision-making space:

- decision-making in management;
- decision analysis and decision intelligence;
- evidence-based decision-making.

We will consider each of these in turn, before focusing on some current researchers in the decision science space.

#### I.1.1 DECISION-MAKING FOR MANAGEMENT

In their article *A Brief History of Decision-Making*, Leigh Buchanan and Andrew O’Connell describe how, in mid-20th century management circles, there was a shift towards viewing decision-making as a specific activity of management. Prior to this, management was viewed as primarily focusing on administrative activities such as resource allocation and policy making, coordinating and logistics. While managers

would certainly be making decisions in these various contexts, decision-making was not considered a separate focus or skillset.

Buchanan and O’Connell note that this changed when Chester Barnard, a key decision theorist of the time, introduced the term “decision-making” as a label for a specific type of managerial activity. Subsequently, an explicit goal of management would be to make deliberate, definitive decisions on behalf of their organization.

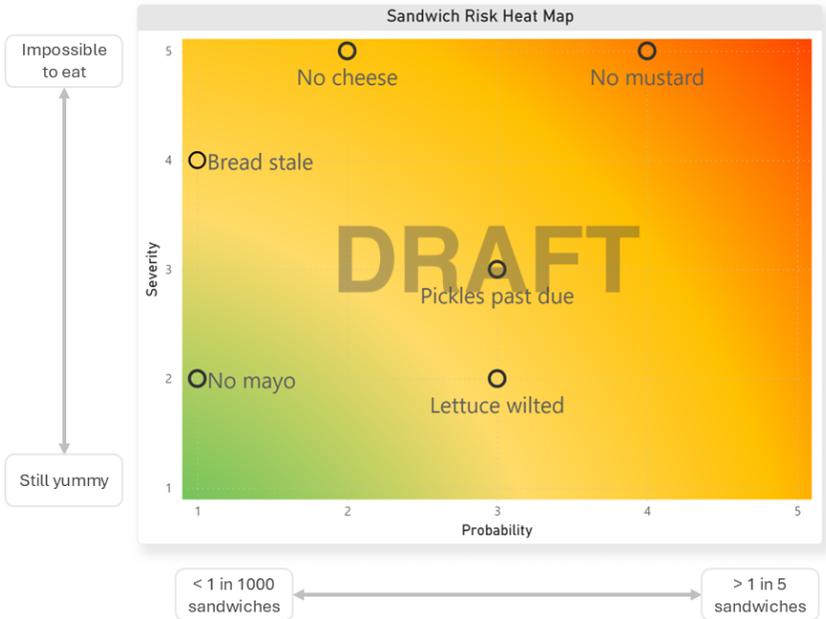
Other researchers, such as James March and Herbert Simon followed on from Barnard. James March advanced the Garbage Can Theory of organizational decision-making, which suggested that organizational decision-making was frequently disorganized and fluid. Crucially for future research in the area, Simon developed the idea of bounded rationality, which describes the cognitive constraints faced by decision makers. As a result, these decision makers could not be expected to make perfectly rational decisions.

### I.1.2 DECISION MAKERS ARE NOT PERFECT

Buchanan and O’Connell further describe how, building on Simon’s concept of bounded rationality, researchers like Daniel Kahneman, Amos Tversky, and Antonio Damasio explored the space and identified various psychological factors that contributed to an inability to make perfectly rational decisions.

Damasio considered the role of emotions in decision-making, while Tversky and Kahneman explored the factors that lead people to make decisions that go against their economic interests.

Coming from this research, there was a growing recognition that humans suffer from a suite of cognitive biases. New researchers considered strategies for making good decisions within this cognitive context. Gerd Gigerenzer promoted the use of simple straightforward heuristics, while Amitai Etzioni proposed that people focus on “humble” decision-making tactics.



*Using a heatmap to determine the risk of eating a particular sandwich.*

## NOW: MODERN TOOLS AND RISK MANAGEMENT

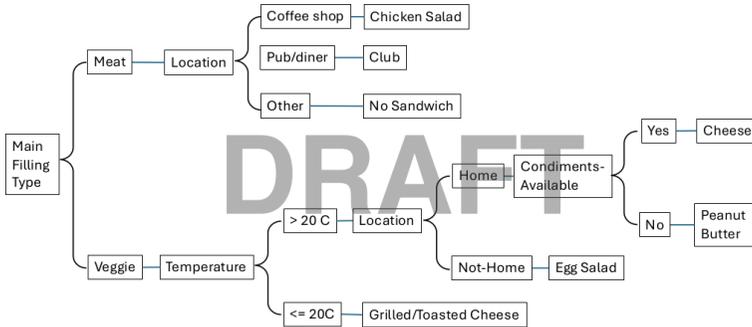
With the advent of computers and digital representations of data, as well as maturing statistical techniques, the focus of decision research in this space turned towards how the findings from researchers like Gigerenzer and Etzioni could be combined with data and data focused techniques to inform different aspects of data supported decision-making.

This, it is hoped, will free people from cognitive limitations and allow for more rigorously defined decision-making strategies. Risk management is one such approach, the goal of which is to quantify the risk associated with various events and choices in order to optimize the outcome of decision-making in managerial contexts.

### I.1.3 DECISION ANALYSIS

Moving alongside this first stream of decision research, which was dominated by social science researchers, the field of decision analysis, starting in the 1960s, brought a mathematical perspective to the space. In *A Brief*

*History: From Decision Analysis To Decision Intelligence*, Ole Delgado provides an overview of key periods in the history of decision analysis and decision intelligence. He begins by noting that a key moment for this stream of decision-making research was the “Choices Under Uncertainty” lecture by Howard Raiffa in 1968. Additional researchers in mathematics, operations research and engineering contributed methodological framings, mathematical formalisms and conceptual models.



*Example decision tree for selecting a type of sandwich.*

This field was further supported in the 1980s by the development of computer-based decision support systems. Decision trees and risk benefit analysis were already in play at this point, but having computers for support allowed for the modelling of more complex scenarios. Expert based systems were also launched. These systems were used in business intelligence as well as in other (e.g., military, medical) intelligence and decision-making contexts.

By the 1990s decision analysis techniques were increasingly being supported by larger and larger amounts of data, which was being managed via data warehouses and other information technology (IT) structures and techniques (e.g., online analytical processing, OLAP). These techniques also supported merging data from different sources. In the 2000s, web-based business intelligence and collaborative decision support systems became widely available. By the 2010s business intelligence platforms were increasingly common, and artificial intelligence (AI) and machine learning (ML) techniques were making inroads.

The 2020s brought a focus on creating specific decision-making processes supported by data. Now, real time data is increasingly available through Internet of Things (IoT) technology, which further supports situational awareness and in-the-moment decision-making.

#### I.1.4 EVIDENCE-BASED DECISION-MAKING

Unlike the previous two streams, which focus to a large extent on business related decision-making, evidence-based decision-making came from the evidence-based medicine movement in the 1990s. The term “evidence-based medicine” was first used in 1991 by Gordon Guyatt, who argued that medical professionals should make decisions based on evidence, rather than traditional wisdom or personal experience. The concept of evidence-based decision-making gained even more momentum in the 2000s, coinciding with a rise in the availability of medical data. This idea then extended into other spheres, and the concept of evidence-based policy or governance quickly followed. The adjective “data-driven” also started to be popular around this time, although the term “data journalism” had been used by Ben Wattenberg in 1960.

### I.2 MODERN DECISION SCIENCE: KEY PLAYERS AND STUDY AREAS

As the previous history review makes clear, many researchers and practitioners over the past seventy five years or so have studied decision-making under the banners of psychology, economics and management studies, as well as mathematics, operations research, engineering and medicine. Nonetheless, the decision science field is still young, with the terms ‘Decision Scientist’ and ‘Decision Intelligence’ being popularized by Cassie Kozyrkov circa 2018.

Here we highlight six key contributors to current decision science, along with their areas of contribution.

#### DANIEL KAHNEMAN

A Nobel Prize winning psychologist known for his work in behavioral economics and decision-making under uncertainty. A psychologist and economist, his major field of research was behavioural economics, with a focus on decision-making, judgment, heuristics and biases. He is perhaps most well known by the general public for his popular book *Thinking, Fast and Slow*, published in 2011.

GARY KLEIN

A cognitive psychologist renowned for his work on naturalistic decision-making and expertise. He is a research psychologist whose focus is on naturalistic decision-making. Among other research areas, he studied firefighters in their natural environment and developed a new decision-making framework. He published *Sources of Power: How People Make Decisions*, in 2017.

RICHARD THALER

A behavioral economist who won the Nobel Prize in Economics for his contributions to understanding irrational decision-making. He has carried out research on the concept of nudges, small actions that push people to make decisions in certain directions, and co-authored with Cass Sunstein in 2008 the popular book *Nudge: Improving Decisions About Health, Wealth, and Happiness*. This led to nudge departments in governments all over the world, focusing on the concept of choice architecture.

SHEENA IYENGAR

A psychologist whose work focuses on understanding how people make choices and the implications of choice overload. She teaches at Columbia University in the Business school and is an expert on choice: why people want choice, what affects how and what we choose, and how to improve our decision-making. She authored *The Art of Choosing* in 2010.

ANNIE DUKE

A former professional poker player turned decision strategist and author. Her work focuses on the application of decision theory and cognitive psychology in business and personal life. She is an author in cognitive behavioural decision science and decision-making, holds a World Series of Poker gold bracelet and has authored numerous books on poker as well as books on decision-making, the most well known of which is *How to Decide: Simple Tools for Making Better Choices*, published in 2020.

CASSIE KOZYRKOV

A data scientist, Chief Decision Scientist at Google until 2023 and popularizer of the terms “decision science” and “decision intelligence”. She is viewed as the first official decision scientist, and is a prominent advocate for decision intelligence. According to Kozyrkov, “A lot of the training that data scientists have assumes that the decision maker knows exactly what they need and the question and problem are framed perfectly.” She further argues, from her experiences at Google and with other organizations, that this is not usually the case. Kozyrkov’s perspectives on different aspects of decision science will be a focus of this book. Her concept of the decision scientist, and views on the need for decision scientists, arose from Google’s adoption of machine learning.

In this context, Kozyrkov observed that an algorithm is only as good as the data it is trained on, which itself is implicitly full of human decisions. Therefore, we need a principled approach to human decision-making: decision science. Otherwise, decisions may be steered by unconscious bias rather than the data, or by data that is itself biased.

## CHAPTER I ESSENTIALS

### DECISION-MAKING RESEARCH STREAMS

**Decision-Making for Management:** Starting in the 1950s, decision-making became a focused skillset for managers. This stream is psychology and social science focused.

**Decision Analysis:** Starting in the 1960s, this involves the application of mathematical, computational and engineering approaches for decision-making optimization.

**Evidence Based Decision-Making:** Starting in the 1990s and coming out of the medical field, the emphasis of this stream is on acting based on facts generated by solid research methodologies.

### KEY DECISION SCIENCE RESEARCHERS

Daniel Kahneman, Gary Klein, Richard Thaler, Sheena Iyengar, Annie Duke, Cassie Kozyrkov.

## REFERENCES

Buchanan, L. and O'Connell, A. (2006). [A Brief History of Decision-Making](#) [↗](#) . Harvard Business Review.

Wikipedia contributors. [Garbage can model](#) [↗](#) . Retrieved on 29-Jul-2025.

Cohen, M.D., March, J.G., Olsen, J.P. (1972). [A garbage can model of organizational choice](#) [↗](#) . *Administrative Science Quarterly*. 17 (1): 1-25.

Delgado, O. (2023). [A Brief History: From Decision Analysis To Decision Intelligence](#) [↗](#) . LinkedIn.

Raiffa, H. (1968). *Decision analysis: introductory lectures on choices under uncertainty*. Addison-Wesley.

Sur R.L., Dahm P. (2011). [History of evidence-based medicine](#) [↗](#) . *Indian J Urol*. 2011 Oct;27(4):487-9. doi:10.4103/0970-1591.91438.

Wikipedia contributors. [Data Journalism](#) [↗](#) . Retrieved on 25-May-2025.

Byrne, C. (2018). [Why Google defined a new discipline to help humans make decisions](#) [↗](#) . Fast Company.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.

Klein, G. (2001). [Sources of Power: How People Make Decisions](#) [↗](#) . 10.1061/(ASCE)1532-6748(2001)1:1(21).

Thaler, R. H., Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Iyengar, Sheena. (2010). *The art of choosing*. New York, NY: Twelve.

Duke, A. (2020). *How to decide: simple tools for making better choices*. Portfolio/Penguin.

## II

# THREE APPROACHES TO DECISION-MAKING

In this chapter we will consider three different levels of data-supported decision-making:

- Armchair
- Data-Inspired
- Data-Driven

All three of these levels of decision-making involve a problem (or goal), choices, and outcomes, but information and data are incorporated differently in each approach.

Two of these levels come from Cassie Kozyrkov. In the Hacker-  
Noon article *Data-driven? Think again*, she introduces the terms data-inspired and data-driven decision-making to describe certain approaches to data-supported decision-making that she considers to be prevalent in organizations.

Armchair decision-making is our own description for a third approach to decision-making, which is also discussed in some detail by Philip E. Tetlock in the book *Superforecasting: The Art and Science of Prediction*.

## II.1 ARMCHAIR DECISION-MAKING

Think about a news show with a bunch of talking heads, discussing a particular issue and what should be done. This is quintessential armchair analysis. Where are the opinions of these talking heads coming from? They seem to be relying on what might be called “facts” but are often instead offering up beliefs and opinions, residing in their brains.

People carrying out armchair analysis and subsequent armchair decision-making could be subject matter experts on the topic they’re discussing, which perhaps gives their opinions some weight. But in any given situation, it is worth asking: are they, in fact, experts in the field they are discussing? It’s also worth considering what cognitive biases and other thought errors (for example, mis-remembered facts) might come into play.

Tetlock vividly captures the attitude that can accompany this type of decision-making support:

“Open any newspaper, watch any TV news show, and you find experts who forecast what’s coming. Some are cautious. Most are bold and confident. A handful claim to be Olympian visionaries able to see decades into the future. With few exceptions, they are not in front of the camera because they possess any skill at forecasting. Accuracy is seldom even mentioned. Old forecasts are like old news, soon forgotten. The one undeniable talent that talking heads have is their skill at telling a compelling story with conviction, and that is enough. Many have become wealthy peddling forecasting of untested value to corporate executives, government officials and ordinary people who would never think of swallowing medicine of unknown efficacy and safety but who routinely pay for forecasts that are as dubious as elixirs sold from the back of a wagon.”

p. 5, Tetlock, *Superforecasting: The Art and Science of Prediction*

How common is armchair analysis and its use in decision-making? If we include people sitting around a campfire back in days long ago, it has

arguably been the most common decision-making strategy of humans since humans were able to discuss their decisions. Even today, based on our own experiences, it's quite a typical approach to decision-making, in both organizational and personal contexts.

Fast	Slower	Slower
Low Grounding	Medium Grounding	High Grounding
High Risk	Medium Risk	Medium Risk
Hard to Automate	Hard to Automate	Easier to Automate
Opinion-Focused	Fact-Focused	Data-Focused

---

Armchair	Data Inspired	Data Driven
----------	---------------	-------------

*A comparison of armchair, data-inspired, data-driven decision making.*

In some circumstances (for example, what sandwich should I eat for lunch?) it might be entirely appropriate. Perhaps your friend, who is a real sandwich connoisseur, will have some great advice. In other situations (for example, should a city institute school lunches for all children?) the armchair approach may not be the best strategy for making a good decision, which is to say one that has a high probability of a desired outcome.

Under some circumstances, people may genuinely prefer their decision-making to be fact-free (facts will be defined more clearly in a later chapter). They may resist incorporating facts into their decision-making process by saying things like “I don’t want to hear it!” or “The facts will just get in the way”. This might happen if they prefer to be mainly driven by values or preferences (they want what they want), or they feel overwhelmed by trying to incorporate facts into their process, or they trust others to make the decision for them, or they feel the decision must be made in a hurry and there is not time to think too much, or they don’t care too much about the decision one way or the other.

Many people, however, do want their decision-making activities, particularly the important ones, to be informed by facts. They recognize that grounding their choice of actions in a true understanding of the world and what is actually the case will increase the probability of their actions resulting in the achievement of their goals.

What can be helpful for such fact-supported decision-making is a giant, and preferably relevant, “grab bag of facts” to draw on. Is this

what we really mean when we advocate for evidence-based decision-making? How does this bag of facts connect to data and analysis? Do we always need data sets and analytic activities in order to achieve evidence-based or fact-driven decision-making? At the very least, it seems like our decision-making will be better in the presence of relevant information from reliable sources. The next level of data-supported decision-making, data-inspired decision-making, captures this idea.

## II.2 DATA-INSPIRED DECISION-MAKING

As might be expected, data-inspired decision-making has a greater reliance on information and data than armchair decision-making. According to Kozyrkov, again in her article *Data-driven? Think again*, “[Data-inspired decision-making is] where we, like whales encountering plankton, swim around in some numbers, and then reach an emotional tipping point and... decide. There are numbers near our decision somewhere, but those numbers do not drive it.”

Here, data is adjacent to the process.

Kozyrkov takes a dim view of data-inspired decision-making because here data is still not the primary driver of the decision-making activity. There are other influencers involved, such as emotions and opinions and personal experience. Cognitive biases often come into play and the presence of data may not be able to mitigate these biases.

We are more of a fan of some variants of data-inspired decision-making than Kozyrkov. This is because we think that the impact of cognitive biases can be mitigated through the awareness of biases and bias-mitigating strategies, and also that, on the flip side, always using the data-driven decision-making approach, discussed next, can have its own issues. We also believe that values have a specific, relevant role to play in decision-making, and sometimes emotions are signals about relevant values.

## II.3 DATA-DRIVEN DECISION-MAKING

In the article, Kozyrkov describes data-driven decision-making thus:

“For a decision to be data-driven, it has to be the data, as opposed to something else entirely, that drive it. If the first thing we do is poke around in our data, our decision will be, at best ... data-inspired... Solution: set the goalposts in advance and resist temptation to move them later.”

At its most extreme this is essentially automated decision-making. A process is set up, a decision point is determined, data/information pours in and then the decision is taken (choice made, action taken) automatically based on this data. For example, suppose that I have decided to use data-driven decision-making to buy my sandwich for lunch. Perhaps I decide that I will buy whatever sandwich my dataset tells me is the most nutritious. That is the decision-making criteria I have defined in advance to be the optimal one for me, and that is the one I will follow. The nutritious sandwich is selected, and I eat my lunch.

Importantly, in this case, by prioritizing nutrition and using it as the only criteria for my decision-making, I'm effectively deciding that it does not matter whether or not I actually like the sandwich that has been chosen. Rather, I have presumably decided in advance that taste is not a factor, or at least not one that is being considered or prioritized. And, if I have carefully decided and sincerely committed to always value nutrition over taste, that might be totally fine. I will eat my perhaps untasty sandwich and be pleased with the outcome of my decision-making.

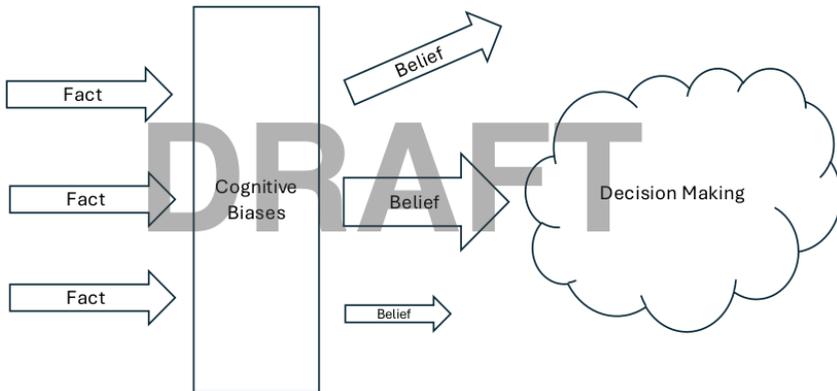
## II.4 WHEN TO USE DATA-SUPPORTED DECISION-MAKING

All three of the approaches above do involve some amount of thought. Is it possible for us to make decisions without thinking at all? To some extent.

## II.4.1 NO-DATA, NO-THOUGHT DECISION-MAKING

When faced with a decision, we can choose to flip a coin or consult the magic eight ball and choose an action based on the result. There is no reasoning involved in the choice of action, and the decision-making is not informed by data. To use such decision aids, you do still have to know the choices that are before you, and set up a process (if the coin lands heads, I will do A, but if the coin lands tails I will do B).

Once you have done that, however, you could argue that the coin tossing or eight ball consulting approach cuts out everything else, leaving the choice to fate. This might be a fine strategy if all of the choices available seem equally good (or bad). In other cases, would asking an all-knowing oracle, or perhaps an all-knowing AI, and then deferring to them, be a better approach? Perhaps, but no such things exist. This means that we ourselves must make and take responsibility for our choices. It also means that we will be making them under a condition of uncertainty, where some data and information is lacking, and we will still have to take responsibility for their outcomes.



*How cognitive biases impact fact interpretation.*

A main guiding principle of this handbook is that, even under conditions of uncertainty, we're more likely to make decisions that have good outcomes when these decisions are supported by quality data and skillful analysis.

## II.4.2 KOZYRKOV'S TAKE ON WHEN DATA SHOULD GET INVOLVED IN DECISION-MAKING

It is perhaps surprising, given her positions above, that Kozyrkov does not think it's always practically worthwhile for data to be involved in decision-making. Her position in *Data-driven? Think again* is that if you, the decision-maker, already have your preference, and there is no data or information that you can feasibly access that would, in fact, change your mind, there's no need to turn to the data. That would just be a waste of effort.

The argument here is essentially: be honest. If your mind cannot be changed by the facts, then make your decision based on your preference, and call it a day. She is not necessarily saying this is a good thing, or that it will lead to a good outcome, but she encourages people to be pragmatic about the reality of the situation.

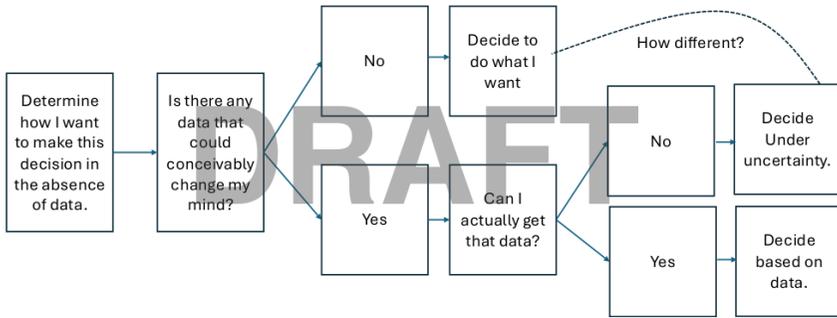
How does this work in practice and get incorporated into a broader decision-making context? In her article, Kozyrkov provides a shopping example, which we have expanded and adapted here as follows:

Consider a situation where you are shopping with your friend and a brand-new shirt catches your eye. Do you:

1. Buy the shirt without bothering to look at the price?
2. Look at the price before making your decision? And then perhaps do some armchair or data-inspired analysis before deciding what to do?
3. Decide what the price would have to be to make your decision one way or the other, and only then look at the price?

Our interpretation of Kozyrkov's take on how to decide if it's worth involving data (and by extension the work to obtain and then analyze it) in the decision-making process is as follows:

1. Determine how you really want to make the decision in the absence of any additional information: What is your default choice before knowing the price of the shirt and other facts? Do you really want to buy the shirt, or not buy the shirt?



*A high-level process for data-driven decision-making.*

2. Be honest: Is there anything that would really, truly, convince you to make a different choice, that you could not then talk your way out of? Is there some price threshold? Is the issue mainly the price? Does the material of the shirt matter? Or where it was made? What are the relevant factors or criteria? Define how you would make the decision if you had access to any information you wanted, as well as what would truly convince you to decide against what you want now (you can think of it as a decision tree).
3. Look at whether or not you actually can access the type of data you would need to change your mind (price, washing instructions, etc.).
4. If you cannot access this mind-changing data, why not just go with your default choice right off the bat? If you would buy the shirt at any price (perhaps you are a billionaire for whom the price of shirts is trivial), or regardless of the washing instructions and shirt's material (perhaps you don't care if the shirt will fall apart after you wear it once), then just take the shirt to the cash and hand them your credit card. Don't bother expending the energy to look at the tags and debating about what to do. In general, do not waste time and energy debating or analyzing the situation unless it will honestly make a difference.
5. On the other hand, if price and/or washing instructions truly do matter to you, and can budge you from your default choice, then you must still vigorously guard against your cognitive biases. Set your goalposts, or decision boundaries, before looking at the

price tag or washing instructions: I will only buy the shirt if the price is less than this value and the shirt does not need to be hung to dry.

6. Then stick with these decision boundaries when the facts are revealed, and take your action based upon them.

### II.4.3 OUR TAKE ON WHEN DATA SHOULD GET INVOLVED: THE MUSTARD IN THE SANDWICH

Our general position is that we can think of data and information (their relationship to each other to be discussed shortly) as the mustard in the sandwich (or pick the favourite sandwich spread of your choice). The mustard is perhaps not strictly necessary, in every situation, but when used with skill it takes a sandwich from “Meh” to “Amazing”. Similarly, involving data and analysis in decision-making can increase the quality of the decision, and by extension increase the positive impact and minimize negative side effects.

This means that moving from armchair to data-inspired decision-making is still a valuable step, in our opinion. We agree with Kozyrkov that data-inspired decision-making has its weaknesses, but we also think that armchair decision-making is still very prevalent, and we believe that in some circumstances data-inspired decision-making will at least be more successful than armchair analysis from a good-outcome point of view, possibly quite a bit more if some precautions are taken. Data-driven decision-making can also be great in clear-cut situations where all of the required data is available, and there are few possible edge cases or unanticipated consequences. However, that may not be the situation at hand.

The three levels of data-supported decision-making (armchair, data-inspired, data-driven) highlight the general role that data and information can play in decision-making. So far not much has been said, however, about the specific mechanics, components, and activities of decision-making, which could arise in any of these three contexts. The subsequent reference chapter will get more into examining these components. The final part of the next chapter then returns to data and

information, considering how they integrate with the other elements and activities discussed.

## CHAPTER 2 ESSENTIALS

### DECISION-MAKING APPROACHES

**Armchair Decision-Making:** Carrying out analysis and decision-making in an informal fashion, relying on opinions, experiences, and “facts in the head”.

**Data-Inspired Decision-Making:** A semi-structured decision-making approach that incorporates data-supported facts but is also influenced by values, emotions, and biases.

**Data-Driven Decision-Making:** An approach that centers the output of data analysis as a predetermined trigger for specific decisions. This makes decision-making much easier to automate. Good results rely on high quality data and upfront buy-in from decision-makers.

**Kozyrkov’s Caveat:** Only go to the effort of data-driven decision-making if it will truly influence the choices made and actions taken.

## REFERENCES

Kozyrkov, C. (2018). *Data-driven? Think again*. HackerNoon. [Data-driven? Think again](#) ↗.

Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: The art and science of prediction*. Crown Publishers/Random House.

# III

## DECISION-MAKING STATES AND ELEMENTS

### III.I THREE DECISION-MAKING STATES

From a decision-making point of view, for any given decision-making situation, people are always in one of three decision-making states:

- Pre-decision-making
- (Active) decision-making
- Post-decision-making

Within the pre-decision-making state, situational awareness allows for a recognition of problems or opportunities. As problems become more prominent or goals more desirable, or the situation changes in some key way, this situational awareness can prompt a move into an active decision-making state.

The active decision-making state can last a moment, or years, depending on the circumstance. After decisions have been made and actions taken, people enter into the post-decision-making state, where they experience the outcomes and consequences of their actions. This is also the time for evaluating the decision-making process and results.

## III.2 CORE ELEMENTS OF ACTIVE DECISION-MAKING

Once in the active decision-making space, the core elements of the decision-making experience include:

- the problem to be solved (and/or the goal or target to be achieved),
- choices and possible actions,
- outcomes, products, results, consequences
- and information.

We'll discuss each of these elements in turn and consider which decision-making activities can be carried out in connection with the element. The more skill a person develops in these activities, arguably the better the decision-making process will be for them.

During this discussion, we'll sometimes loosely refer to "the decision-making process", but we don't intend to imply that there is necessarily a well-defined path proceeding linearly from problem identification to decision outcomes during decision-making. As has been noted by Henry Mintzberg and Frances Westely, decision-making in the wild is often an organic, informal, disorganized process.



*Decision-making states.*

However, we believe that the above elements will be present in some fashion during decision-making, even if they can perhaps only be identified retrospectively. Consequently, the skills developed in connection with each element will be beneficial in both formal and informal decision-making contexts.

### III.2.1 ELEMENT: PROBLEM, GOAL, TARGET

A problem, goal or target kicks off the decision-making process proper. It's what we strive to solve, achieve, attain (or in the case of a negative outcome, avoid) as a product of the decision-making process. As such it can be thought of as appearing at both the beginning and end of a decision-making exercise.

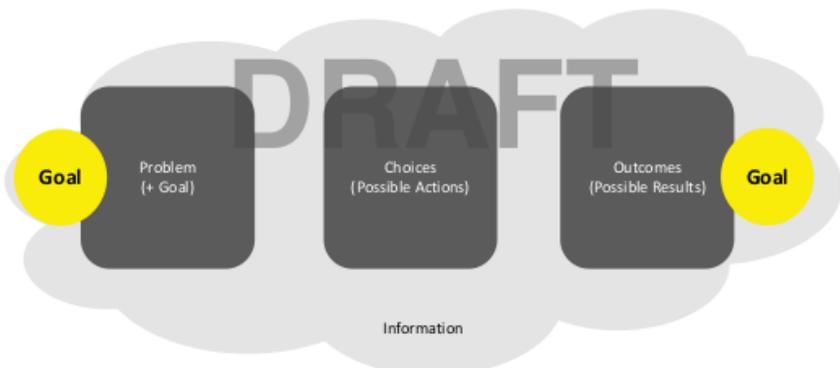
When identifying or receiving the goal, problem or target, we are moving from the pre-decision-making state (or space), into an active decision-making space. We perhaps find ourselves in a state that is not stable or satisfactory. We want things to be better, or at least not get worse. Or maybe we're in a situation that simply must evolve in some direction, and we want to have agency in that context. We start to consider what we could do under the circumstances.

Relevant decision-making activities and skills associated with the problem, goal or target are:

**Describing:** Describe the current state of affairs. What situation are you in now? What about the current state cannot stay the same, or is no longer the same? What do you want about the situation to change, or stay the same?

**Defining/Stating:** State clearly the problem or goal that you have identified or have been given by someone else.

Situational awareness is important at this stage, as it allows us to detect that we are in a situation that needs some decision-making (or some action) and to gain a better appreciation of just what that situation is.



*The core elements of decision-making.*

We might ask ourselves:

- What exactly is my situation right now? (Do I have data to tell me?)
- What factors are contributing to my situation? (cause and effect)
- What factors might cause or are causing my situation to change? (cause and effect)

Here, by factors we mean those aspects or properties of the situation that are constraining or defining the situation. If these were different, or had different values, then the situation would be importantly different.

We may not know about problems unless we've been monitoring our environment in the pre-decision-making state. This monitoring also provides an opportunity for data and information gathering. If we arrive at active decision-making with no relevant data, there will be increased pressure to just make a choice via armchair decision-making, and get on with things.

Often this aspect of decision-making will be more emotion, preference or value driven than data-driven. As the decision maker, whether or not we want to stay in and maintain the current state, or make a change, will depend on how we feel about it, and how strongly we feel. Whether or not we consider some situation to be a problem, or what we consider to be a worthwhile goal, will depend on our values and preferences. An understanding of the emotions, values, preferences and priorities at play during decision-making is critical.

### III.2.2 ELEMENT: CHOICES, POSSIBLE ACTIONS

If we agree that we wish to act, or must act, we can start by asking:

- What are our possible choices of actions?
- What are the constraints on these choices or actions?
- What are the factors at play? How do these influence our decision-making? How do they differentiate or impact one possible course of action (or one possible outcome of the action) from another?

When we talk about factors in connection with choices and actions, we mean those aspects or properties of the situation that we can imagine being different, or having multiple different values, where these different values are connected to different choices and outcomes.

In the shirt example, factors included the price of the shirt, the material of the shirt and where it was made. At a music concert, factors could include the mode of transportation taken to get there, the time (year, day) of the concert, the number of people at the event, and the location and the type of music.

Relevant decision-making activities and skills associated with choices and possible actions are:

**Imagining:** Imagine possible choices and actions.

**Identifying:** Identify relevant factors. What impacts or influences choices? Alternatively, identify constraints on our choices.

When we are considering our choices, we may want to avoid choice tunnel vision, or the appearance of forced choice, where the perception is that there are a limited number of choices that you must consider, or can't avoid, and thus one of these must necessarily be chosen.

Different techniques can help you break away from the perception of forced choice. For example, Effects Based Planning is a technique used by some military organizations to help planners be more aware of the effects of their choices and break out of a forced choice mindset.

To help avoid choice tunnel vision:

- Identify a particular pain point in your situation that could, in turn, become a decision-making point.
- Take a moment to brainstorm: How do you characterize the choices you see.
- Deliberately ask yourself: Did you throw yourself into (or have you been artificially placed into) a forced choice situation?

At the same time, we want to avoid analysis paralysis, where there are too many choices and we cannot pick one in order to take action, even though not acting could be causing its own substantive issues.

The *Indeed* article “What is Analysis Paralysis” identifies several strategies for overcoming analysis paralysis, including:

- Recognize that you are in it
- Prioritize taking a decision-making decision
- Take a break
- Ask for advice
- Make small, quick decisions rather than one large one
- Set a deadline
- Understand your goals
- Limit your information intake

### III.2.3 ELEMENT: OUTPUTS, RESULTS, OUTCOMES, CONSEQUENCES

What are the different outcomes and consequences that could happen due to a particular action? When considering outcomes and consequences, it’s useful to draw on the concept of possible worlds. Each different possible outcome can be thought of as representing a different possible world. A single choice could have multiple possible outcomes, in which case each would be imagined as its own possible world.

For each pair of (possible course of action, imagined outcome), we can then consider pros and cons, risks and benefits. This returns us to our consideration of factors, but in a different sense: What are the factors (effects) that lead to different possible worlds depending on the value of the factor. For example, considering the factor of transportation, if we take the bus to the music concert or we take our car to the concert, what are different possible worlds that arise from one option versus another? Or, if funding is a factor in outcomes for a project, what are the possible worlds where the funding is \$2,000 versus where the funding is \$10,000?

Between our choices of action and possible outcomes, we can also think about the direct outputs, products or immediate results of our

choice of actions, the arrival time by bus or car, say. These can be used as a stepping stone to help further consider the broader and more contextual outcomes and consequences of the initial results or impacts.

Relevant decision-making activities and skills associated with outputs, products and outcomes are:

**Imagining:** Imagine possible outcomes as possible worlds.

**Identifying:** Identify factors (impacts), identify constraints, identify probabilities of each outcome (probability, cause and effect).

### III.2.4 ELEMENT: INFORMATION, DATA, KNOWLEDGE

We have been throwing around the term “information” quite a bit so far in this handbook without properly defining it. Now it’s time to get more precise about what we mean. But before we do that, we want to list up front the relevant decision-making activities associated with information and data.

Relevant activities and skills associated with information are:

**Identifying:** Identify sources of information and data. Also, identify which parts of the decision-making process would benefit from particular statements or types of statements.

**Transforming:** Transform data and information, through analysis, into statements about certain topics, and with certain structure.

#### THE FORM OF INFORMATION IN DECISION-MAKING

Although information has different technical definitions in different contexts, some more fundamental than others, in the context of decision-making the focus is typically on representing some aspect of our reality using symbols and language. Because of this, we suggest that the statement (or proposition) is a fundamental form of information in decision-making contexts.

Drawing on traditions in analytic philosophy, we will define a statement, for the purposes of this handbook, as a piece of language, spoken or written, that represents an actual or potential aspect of reality. For example, “It is raining right now.” is a statement.

A statement may accurately (truly) or inaccurately (falsely) represent reality. A fact is then defined as a true statement. As can be seen from the example statement above, whether or not the statement “It is raining right now.” is true depends partially on the statement, and partially on the state of the world. “It is raining right now.” could be a true statement in the morning and a false statement in the afternoon.

How does this connect to what’s in our heads? For the purposes of this handbook, a belief can be defined, loosely, as a statement (which might be true or false) that we have in our mind. Knowledge is often defined as a belief that, first, is true, and second, we are justified in believing to be true. The reason for this extra aspect, of justification, is that we might believe something that is actually true, but we believe it due to bad reasoning on our part. We end up with the right conclusion, but in the wrong way, and that doesn’t seem like it should be called knowledge.

Even justification is insufficient for something to count as knowledge according to some. For example, if we look at a broken clock at exactly the right time, not knowing that the clock is broken, we might end up with a true belief about the time, and our reasoning is fairly sound (clocks are built to tell time) but in reality knowing the correct time is still just luck, so technically in this case we also shouldn’t call our belief knowledge (this is known as the Gettier problem). However you slice it, though, the important takeaway from this debate is that all of these definitions place knowledge in our heads, rather than in our books.

Data is the last piece of the puzzle to be addressed. For the purposes of decision-making, we define data as a structured collection of measurements, with a datum being the representation of the result of a single measurement. For example, a measurement of your age at the time of reading this sentence would be a datum. Each datum (which can be more concretely thought of as the value in one cell of a data table) can then be viewed as representing, or implying, its own very specific and localized statement about the entity or object that has been measured: “At 10:52 AM on Oct. 5th, 2025, my age was ...”

In the context of a data table, the object or entity being measured is represented by an entire row of the table. Each cell of that row then contains a very specific fact about that object or entity. For example, if a row is about a person, statements implied by the cells in the row might include “At the time of data collection, my age was 37 years.” and “At the time of data collection my height was 162.56 cm.” Analysis techniques then act over data to generate new statements from this collection of very specific statements (represented by the data).

This definition of data, focusing as it does on measurement, might seem to exclude what is often called unstructured data. For our purposes, we will simply note that unstructured material or similar digital objects typically must be further structured prior to use in analysis. Because of this, these other sources of information could be viewed as the raw material for generating some structured, more measurement-oriented data, which could then be used in analysis and decision support. Alternatively, text that represents statements could be viewed not as data but as a source of statements that are a product of previous data analysis, or as an expression of opinion residing in someone’s brain.

### WHERE DO THE FACTS FIT?

If we choose a decision-making approach that incorporates information and data (for example, we aren’t just using a coin flip), this information and data must directly connect with, and influence, one or more of the decision-making elements discussed above. How does this work more precisely? What are the specific mechanisms that connect information and data to goals, choices and outcomes?

Understanding the form, or structure, that statements have makes it possible to connect them to particular decision-making elements and use them to directly support decision-making. In the next chapter we will discuss how different types of analysis turn data into statements with different structures, or forms. The specific form of a statement then allows it to support particular elements of active decision-making. High quality data, that accurately represents the world, is more likely to lead to true statements that support good decision-making. Before we can get to the good decision-making part, however, we need to have a better understanding of what we mean by statement structure.

THE FORM AND STRUCTURE OF STATEMENTS

From a pragmatic decision-making perspective, we suggest it's useful to consider three main types of statements: descriptive, causal and value. In terms of their structure, the base format of each of these types of statement are as follows:

**Descriptive:** It is the case that A.

**Causal:** If A then B.

**Value:** A is [value].

Here, the letters are placeholders for more specific parts of a statement. For example: It is the case that it is raining. If I jump up, then I will come down. Eating ice cream is [nice].

Descriptive statements support situational awareness and also increase an understanding and description of the problem or goal (i.e., this is the situation, in which I will act; the problem is as follows). Descriptive statements can also be used to support the generation of a list of possible choices (as in: one possible action is the following).

Causal statements provide information about cause and effect, action and outcome. They can be chained together to support an understanding of the relationship between choices made and the possible outcomes of these choices (for instance: if I act this way, then this is what could happen).

Value statements might technically fall under the descriptive umbrella, but they are descriptive statements that include expressions of value, preference or personal experience. They help determine which goals will be pursued and how outcomes will be evaluated (i.e., I value this, so I will pursue this goal; I value this, so it's the best outcome).

Consider this example of a thought process leading to decision-making and action, which incorporates all three types of statements:

I'm hungry, and that's a problem. Eating ice cream is always a good thing to do when I'm hungry and there's an ice cream store down the street. That means that if I walk down the street, then I can buy and eat some ice cream. I'm going to get my shoes on and walk down the street to the ice cream shop.

Here's the same example with the statement types highlighted:

I'm hungry, and that's a problem (descriptive statement, value statement). Eating ice cream is always a good thing to do when I'm hungry (value statement) and there's an ice cream store down the street (descriptive statement). That means that if I walk down the street, then I can buy and eat some ice cream (causal statement). I'm going to get my shoes on and walk down the street to the ice cream shop (descriptive statements).

This example sounds a bit robotic. People don't usually think so methodically, but it gives a flavour of how these types of statements might be involved in a decision-making activity.

In day-to-day decision-making activities, most people carry out such reasoning intuitively, and probably unconsciously. In many cases this is fine, but as Kahneman discusses in *Thinking, Fast and Slow*, it can also lead to being influenced by innate cognitive biases.

#### STATEMENT TYPES: THE DETAILS

As can be seen in the ice cream shop example, there are many variants of the base forms in each statement category, including:

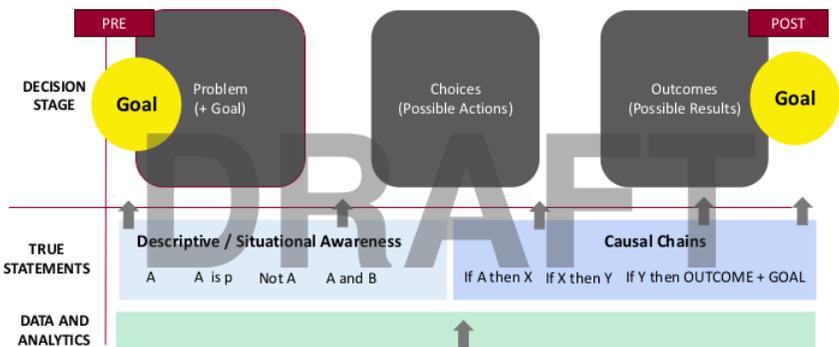
- **descriptive:**
  - A, It is the case that A, A is happening, A is the situation
  - NOT A
  - EITHER A OR B but NOT BOTH, A AND B
  - As are x, some As are x, As are often x
- **causal:**
  - IF A THEN B
  - As are x (but with a causal implication)
  - IF A goes up, then B goes down
  - When A goes up, B goes up

- It seems like A causes B
- It seems like A doesn't influence B
- **value:**
  - A is good
  - B is bad
  - We want A
  - We don't want B

Becoming practiced in recognizing the different forms of these statements can help when incorporating them into different parts of the decision-making process.

### DATA AS STATEMENT CREATION FUEL

Analysis techniques use data as fuel to generate useful statements. Different analysis techniques generate different types of statements, with different structures. This is an important practical differentiator between different types of analysis techniques. Generated statements can then, in turn, act as input into different parts of the decision-making process. If applied “naively”, analyses can result in an unfocused grab bag of statements. If applied in a more focused manner, they can lead to statements relevant to a particular decision-making decision.



*How data and analytics integrate into decision-making.*

If a decision maker understands which types of statements are generated by particular analysis techniques, then they don't need to know the

details of how particular techniques work in order to ask for a particular type of analysis to be carried out. Subject matter experts are important in this environment. They can suggest which type of statement would be most useful, prior to analysis. They can also interpret the “concentrated” statements generated by data and put them into relevant context. Experts answer the question “So what?”

Values and priorities are also essential. Facts generated by data can’t directly help with the value (or preference or priority) part. However, facts can connect information to the value part. For example, consider the following series of statements: “A is good. If B, then A. C usually leads to B.” The conclusion here is, let’s do C. Alternatively, facts can help better understand values, or the relationship between actions and values: “If A then B. B is not good. Therefore, A is also not good.”

Within analysis itself, there are many techniques. How do their outputs inform the decision-making process? How do we map analysis techniques to different elements of the decision-making landscape? This is the subject of the next reference chapter.

## CHAPTER 3 ESSENTIALS

### STATES, ELEMENTS, ACTIVITIES

Improving skills in these areas improves decision-making.

**States:** for any given decision-making context, you will always be in one of three decision-making states:

- Pre-decision-making: the decision is on the horizon.
- (Active) decision-making: considering choices and consequences.
- Post-decision-making: experiencing and evaluating the outcome of the decision.

**Elements:** the main decision-making elements include:

- the problem to be solved (and/or the goal or target to be achieved),

- choices and possible actions,
- outcomes, products, results, consequences
- and information.

**Activities:** working with the decision-making elements, decision-making activities include: describing, defining, stating, imagining, identifying, transforming/analysing, choosing, and deciding.

### THREE TYPES OF STATEMENTS

**Descriptive:** It is the case that A.

**Causal:** If A then B.

**Value:** A is [value].

## REFERENCES

Mintzberg, H., Westley, F. (2001). [Decision-Making: It's Not What You Think](#) [↗](#). *MIT Sloan Management Review*.

Indeed Editorial Team (2025). [What Is Analysis Paralysis? \(Plus 8 Tips To Overcome It\)](#) [↗](#).

Gettier, E. (1963). [Is Justified True Belief Knowledge?](#) [↗](#). *Analysis*. 23 (6): 121-123. doi: 10.2307/3326922. JSTOR 3326922.

Wikipedia contributors. [Effects-Based operations](#) [↗](#). Retrieved on 14-Apr-2025.

# IV

## ANALYSIS TECHNIQUES

In this chapter we will consider five (plus one) types of analysis that use data to generate decision-making support statements:

1. Data Visualization and Basic Analytics
2. Metrics
3. Statistics and Modeling
4. Optimization
5. Classic Machine Learning
6. (+1) Generative AI?

The “plus one”, generative AI, is relatively new to the scene at the time of writing, and there are questions surrounding its potential roles in decision-making which we will discuss near the end of the chapter.

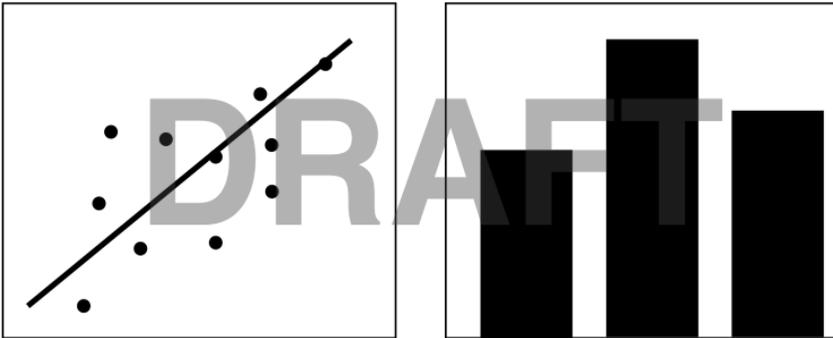
We provide an overview of these techniques from a functional perspective, without discussing how to carry them out. The goal of this chapter is not to increase people’s technical proficiency; the focus is on the role of the technique in the larger decision-making process, with respect to the types of statements generated by the technique.

Following a review of the five plus one analysis options, we’ll switch focus and briefly review a commonly presented framework that characterizes the results of analysis as being either descriptive, diagnostic, predictive, or prescriptive. We’ll connect this popular framework to the decision-making elements and statement structures we have already reviewed.

## IV.1 DATA VISUALIZATION AND BASIC ANALYTICS

Data visualization techniques present many pieces of data to the viewer of the visualization simultaneously, by mapping them onto a two- or three-dimensional space. Presenting many data points together in this way takes advantage of our visual systems and associated cognition, which means that we can more readily perceive and understand how each datum relates to other data. As a result, it becomes easier to synthesize and extract insights from the data. The data used in a visualization may be raw data, summary data or data that has been transformed in some other way.

Analytics is a broad term that encompasses any analysis techniques and technologies that are used for business support. As such, analytics goes hand in hand with data visualization. Basic analytics calculates useful summary numbers (e.g., totals and averages) from raw data to generate statements (i.e., “we sold 56 books in 2012”), and is a common precursor to data visualization. It provides a more general or abstract perspective on a given situation than the raw data does.



*The same data presented first in a data exploration and then in a data storytelling fashion.*

**Types of statements generated:** Basic analytics and data visualizations can be used to generate and communicate a wide variety of descriptive statements, and to a lesser extent, causal statements. For example, by looking at a chart of total sales per month relative to average monthly temperature, it might be possible to extract a descriptive statement like

“Over the last five years, mitten sales have remained relatively constant.” or a causal statement like “If the weather is colder, mitten sales could go up in that month.”

But caution must be exercised. The process of deriving statements from visualizations, particularly causal statements, is not very rigorous, so there is a real danger that the statements extracted directly from data visualizations may not be precise, accurate, or even true at all.

As a result, extracting statements from visualizations can best be viewed as a jumping off point for analysis, with the analysis carried out to verify the accuracy and precision of the statements. Alternatively, post-analysis visualizations can be created to highlight key facts that have already been generated by other types of analysis.

In this way, visualization may be used either for exploring data prior to a more formal analysis or for presenting curated data to decision makers in a way that allows them to extract statements from the visualization that are relevant to their decision-making process. During data exploration, aesthetics and storytelling in the data visualization will not be a priority. After the data has been analyzed, transformed, and curated, however, visualizations that are well designed to highlight appropriate aspects of the data can be created to assist decision makers in drawing useful conclusions from the data. This will be discussed further in the data storytelling chapter.

To become a professional data visualizer, in addition to having an appreciation of the analysis itself, the creator of data visualization charts must have skills in design thinking, both in graphic design, potentially including areas like colour theory and gestalt principles, and in user experience design, which could include the ability to carry out user experience studies and apply design thinking principles to particular data visualization or data storytelling situations. This hybrid set of skills can be hard to find, which is why it's not unusual to come across poorly constructed data visualizations.

## IV.2 METRICS

The term ‘metric’ can have different meanings in different contexts. In a business context, metrics are abstract measures that represent or de-

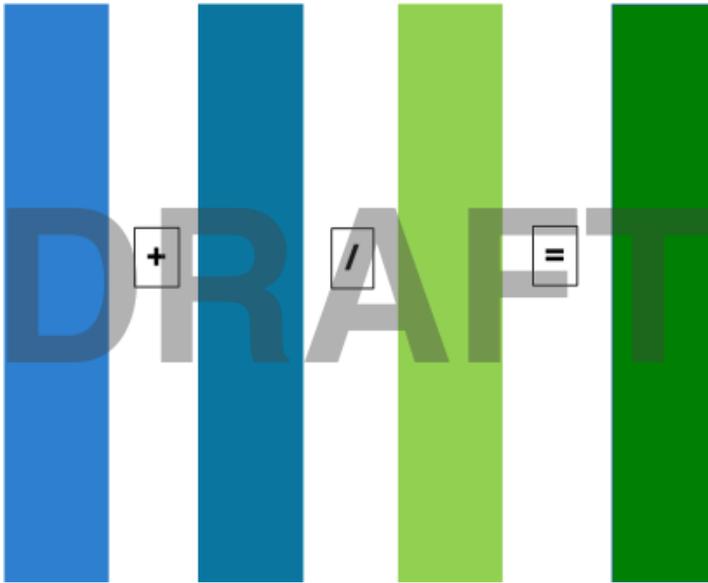
scribe some situation of interest. For example, they might represent or describe some aspect of an organization, like organizational efficiency (e.g., on a scale of one to five) or employee satisfaction (e.g., low, medium, high). Key performance indicators (KPI) are a classic example of business metrics. Metrics are similar in concept to indices in statistics, such as the Gross National Happiness Index.

In a previous chapter we defined data as collections of measurements, each of which represents some aspect or property of an object. Metrics can be viewed as a strategy for transforming this raw data into a measure that is more complex, abstract or meaningful.

Metrics are constructed by combining two or more existing measures, often raw measures. These more basic measures are related to each other using mathematical operations, which allow them to be combined into a single metric. The resulting metrics are intended to represent more abstract or complex concepts, relative to the raw measures that were used to generate them. As such, metrics are effectively small mathematical models, even if they are not always recognized as such by the people who create them.

In a data table, or spreadsheet, each column can (usually) be thought of as a collection of measurement values of a certain type (i.e., in the height column there are measurements of height, with numbers representing the results of these measurements; in the favourite colour column there are colour values like red, blue and orange representing people's favourite colours). Building on this, in the context of a data table, a metric is constructed by combining two or more data columns, with the columns joined by mathematical operations.

For some interesting reading and best practices in uses of metrics, we recommend *The Tyranny of Metrics* by Jerry Z. Muller. Muller is not entirely anti-metric, but notes that, without following best practices (discussed in detail in the book), metrics can easily be used in ways that run counter to their intended purpose. He notes that people who use metrics must be aware of Goodhart's law, which is often stated as: "When a measure becomes a target, it ceases to be a good measure." This highlights the tendency of people to tune their work to the metric, rather than allowing it to measure their work in some more objective fashion. A more extreme version of this occurs when people try to "game" the metric, once they become aware of it.



*Combining data columns into metrics by applying basic mathematical operations to the data columns.*

**Types of statements generated:** When used in a standalone fashion, metrics generate descriptive statements. In the pre-decision-making state they can be used to give insight into goals and increase situational awareness. During decision-making they can serve as impact evaluators for outcomes of interest or value. Post-decision-making they can be used to aid in the evaluation of decision-making outcomes.

In a business context, metrics also sometimes support ad hoc experimentation: “If we change our organization like this, does the metric go up? Or down?” When used in this context, they can generate causal statements (A causes B, if A then B). The truth of these statements will depend, among other things, on the rigor of the experiment and the validity of the metric.

## IV.3 STATISTICS AND MODELING

Statistical techniques can be used to accurately generalize from facts known about a small group (the sample), to likely true statements about a larger group with the same characteristics (the population). For example, if we study a small group of cats (the sample) and learn that 90% of

them have multi-coloured fur, we could generate the general statement “90% of all cats have multi-coloured fur.” (the population here is the population of all cats). For this generalizing to work well, appropriate data collection conditions (i.e., the amount of data collected, and how the data was collected) must also be met.

Descriptive statistics describe the properties of the sample. Inferential statistical techniques can then be used to infer both the properties of the larger group (population) as well as general relationships between two or more properties of the larger group (for instance, inferential statistics might be used to generate the statement “People who eat apples also have fewer trips to the doctor, relative to those who do not.”).

If the main group can be further divided into subgroups, for example, in a group of “people who play sports”, it could be possible to create subgroups based on types of sports played by the people in the group, statistics can uncover relationships between these subgroups (i.e., “this subgroup of basketball players is on average taller than this other subgroup of soccer players”). The techniques can then be extended to make predictions based on the generated statements (i.e., “you belong to this subgroup of basketball players and I belong to that subgroup of soccer players, so we predict that you will be taller than me”).

In addition to generating the statements themselves, statistics also generates an estimate of how reliable the statement is, based on the details of how it has been generated (i.e., “there were only a few soccer players in the group that was analyzed, so we are not that confident about our general conclusions about soccer players”).

**Types of statements generated:** Statistical techniques typically generate either descriptive statements or quasi-causal statements, depending on the technique. Common forms of general statements generated by statistical techniques are:

- A is also B (i.e., “Sandwich eaters are also happy.”).
- A is often X (i.e., “Vegetarians often have low cholesterol.”).
- A has the property X (i.e., “Canadian women are 164cm tall.”).
- When A, typically B (i.e., “When people eat sandwiches, typically they are happy the next day.”).

Some of these statements can be used to derive or lend weight to subsequent causal statements (i.e., the statement “vegetarians often have low cholesterol” might lend weight to statements like “If someone is a vegetarian, then they will have low cholesterol” or “If we want to promote low cholesterol then we should promote being a vegetarian”). At the same time, many people have been taught the saying “Correlation does not equal causation!”

Interestingly, while this is certainly true, it would also be fair to say that there is a lot of upheaval going on in statistics right now, and this is leading people to take a look at this maxim with fresh eyes. Frequentists, which is to say statisticians who subscribe to the frequentist theory of statistics, are somewhat under attack (in a scholarly fashion) from Bayesians and causal modellers, who have different ways of interpreting statistical results.

These groups all agree that it is a problem if we erroneously conclude causation from correlation. However, some among them suggest that it is also a problem if we never admit that correlation is a very strong clue that there could be something causal at play. They urge people not to throw the baby out with the bathwater when it comes to using statistics in applied settings. Adding causal modelling, simulations, and other modelling techniques into the mix can improve our understanding of the causal forces potentially at play.

## CONSIDER THE DECISION MAKER PERSPECTIVE

Our discussion of statistics provides a good opportunity to consider more broadly the situation of the decision maker in relation to these analysis techniques. The decision maker most likely does not know about (and should not have to know about) the details of the statistical or analytical techniques, but likely knows the type of statements, descriptive statements and causal statements, that would be helpful to the decision-making process.

Working backwards from this, if decision makers understand the type of statements generated by different analysis techniques, they should be able to communicate their requirements or wish list of statements, or types of statements, to the data analyst. The data analyst can

then determine if the appropriate data is available for generating the desired type of statement using statistical or some other analysis technique.

## IV.4 OPTIMIZATION

As the name suggests, optimization is a group of analysis techniques that seeks to find the optimal solution to a particular problem. More specifically, optimization seeks to answer questions like:

- What is the maximum or minimum possible in a certain circumstance?
- What is the best possible solution, given a set of constraints and relationships?

Optimization techniques generate statements of the form:

- “The best possible A is B (approximately), and that happens when C.”
- “The worst possible A is B (approximately), and that happens when C.”

For example, suppose you are trying to select which bus route to regularly use to commute home from work, out of several possible options. This is a relatively straightforward decision-making scenario. To use optimization techniques to help, you would first want to define what makes a bus route good, or the best. This could involve the length of the trip, the number of transfers, whether or not there were heated bus shelters, and so on. Optimization techniques could then tell you the best possible result based on your criteria, and which bus routes best met that criterion.

Optimization is most useful when there are a very large number of choices (think millions of possible bus routes) and where going over all of the choices manually is not feasible. In such a case, optimization might not be able to tell you what the absolute best bus route was, but it could provide results that are probably very nearly the best bus routes.

One of the major challenges of optimization is determining in a concrete and quantifiable fashion what counts as “the best” or “the worst”. Although it is a technique that focuses on data, it also requires a focus on organizational values and goals, as well as metrics that have been developed to measure these. It further requires an ability to imagine possible outcomes and categorize them as good or bad outcomes.

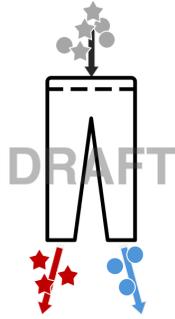
## IV.5 MACHINE LEARNING

Many classic machine learning techniques, in one fashion or another, label objects. Although this sounds unassuming on the face of it, it turns out that labelling is a fairly fundamental cognitive activity, which might explain why children learning language are so motivated to do it. This labelling focus of classic machine learning means that, unlike in previous techniques (i.e., statistics), the lens of machine learning is more firmly on the individual.

Classic machine learning techniques include classification, which is the term used if there are pre-existing labels involved in the labelling activity, and clustering (or categorization) if the goal of the technique is to first create new categories, create labels for these categories, and then apply these new labels to individuals or objects.

Machine learning techniques typically proceed in two steps. First, data is used, in combination with a computer program, to create a model that is specific to the situation represented by the data. This structure can then be incorporated into a second program that can be run in order to generate a useful result (you have programs creating parts of programs here). This is where the learning takes place: the first program effectively teaches, or trains, a crucial part of the second program, which then embodies and uses what has been learned in new tasks. Sometimes the programs produced by these techniques are referred to as engines.

Frequently, the situation-specific program is used to provide labels for new data. In this case, the criteria for success during training is: if the trained engine is presented with a new object that it has never seen before, can the engine successfully label it, and thus classify it into one of several predefined categories?



*When imagining a classification engine we can picture “the classification pants”.*

Some classifiers (i.e., neural networks, commonly shortened to neural nets) are very “black box”. This is because, although the general process used to create a particular engine might be well understood, the particular engines that are then created by the process in specific instances are very difficult to interpret. As a result, a neural net classifier might be good at classifying, but you do not know why. This might be acceptable in situations where a decision does not need to be justified or explained, but problematic if that is not the case.

**Type of statements generated:** Classic machine learning techniques frequently provide statements of the form “A is an X.” For example, “This picture is a picture of a cat.”

## IV.6 GENERATIVE AI AS AN ANALYSIS TECHNIQUE?

Generative AI is a type of neural network machine learning. Speaking in the broadest of terms, generative AI techniques create AI engines that transform some provided digital (symbolic) input into a certain desired type of digital (symbolic) output. The particular behaviour of the generative AI engine is determined by the training set and training techniques that have been used to create it. For example, some generative AI engines will transform text into images, where the image is meaningfully related to the text, that is, the text “a cute orange cat” will generate

an image with some relation to cute orange cats. We can think of this as the engine transforming the text into the image.

This has similarities to other analysis techniques we have discussed, such as statistics or classic machine learning, which also transform input data into output of a certain form. For example, a statistical technique might transform sample data into generalized statements about a population of interest. But generative AI techniques are more general and flexible with respect to the nature of the output relative to the input.

Importantly, as noted above, in the case of generative AI engines, the transformation results in an output that is semantically related (related by meaning) to the input, but in only a partially specified fashion. If the input is the text “a furry orange cat”, a generative AI engine trained to turn text into an image might output an image that looks like some kind of furry orange cat, but the exact nature and details of the orange furry cat image will not be readily predictable in advance.

Why this exact image was produced will be difficult to explain, due to the black box nature of the engine.

This can cause problems when using such engines for analysis activities. Typically, analysis is strongly invested in truth. As well, the goal is often to have the analysis be replicable, verifiable, and explainable. These qualities are not strengths of generative AI techniques, although much effort is currently being devoted to improving this situation.

Some success has been recently achieved in creating engines that, possibly in conjunction with other technologies, can provide more reliably true statements. However, at the time of this writing, many generative engines are only partially grounded in real-world data and information, and as a result their statements should not be considered to be true at face value. Until these aspects of the technology substantially improve, it is difficult for us to recommend using generative AI as an analysis technique.

## IV.7 ANALYTICS AS QUESTION ANSWERING

We have suggested thinking about analysis techniques in terms of the types of statements they generate, and we have offered three categories

of statements, descriptive statements, causal statements, and value statements, to help contextualize the role that these techniques and the subsequent statements can play in decision-making.

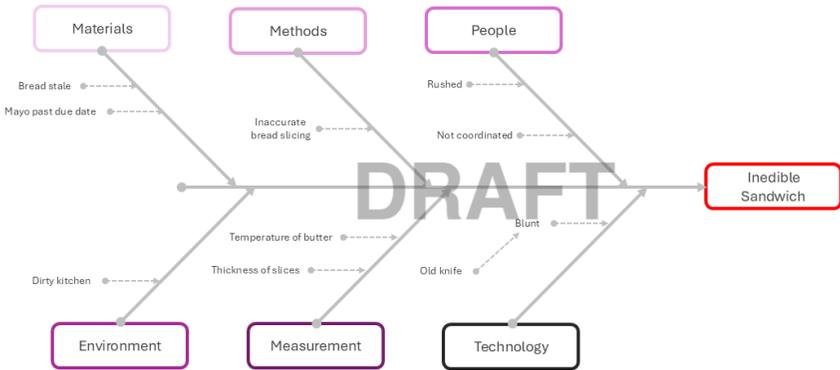
However, many aspects of decision-making can also be framed in terms of asking questions. Relevant to this, within the business intelligence analytics community another commonly used framework characterizes analytic techniques, and their outputs, in terms of the questions that different techniques can answer. Here there are four question-related categories:

- **Descriptive:** What is happening?
- **Diagnostic:** Why is this happening?
- **Predictive:** What will happen?
- **Prescriptive:** What should happen?

**Descriptive techniques:** these use data and analysis to answer the questions “What is happening?” or “What has happened?” The results of descriptive techniques are typically presented in the form of data visualizations like charts, reports, and dashboards. The intelligence produced usually lags behind the current situation to some extent, and is usually focused on a subset of data or a single data set. The data sets involved are very often not appropriate or large enough to infer statistical significance or to perform reliable tests.

As the name suggests, this category of techniques produces descriptive statements. An example of the question here would be “How much money do I currently have in my budget?” Such statements might be used, for example, to help a business understand how it is performing by providing some necessary context.

**Diagnostic techniques:** these provide a deeper analysis, relative to purely descriptive techniques, in order to answer the questions “Why did this happen?” or “Why is this happening?” These techniques very often use formalized root cause analysis tools (i.e., 5 Whys, Ishikawa diagrams, Taproot Root Cause Analysis, Causal Trees). Not surprisingly, given this, this category of techniques produces causal statements.



*Root cause analysis diagram showing causes involved in getting an incredible sandwich.*

Diagnostic analysis is difficult to fully automate and prone to human variability. It can run into issues when it conflates correlation with causality. If the data sets are appropriate, artificial intelligence and machine learning can provide potentially significant value (i.e., decision trees). An example result of using this type of technique would be finding an answer to the question “Why do I have less in my budget than I thought?”

**Predictive techniques:** these apply various methods (statistics, machine learning, etc.) to model data in order to answer the question “What will happen?” Models are first generated using past data. These models are then applied to current data in order to predict what might happen next, or what could happen under certain specified conditions.

Predictive intelligence can suffer from model drift, inappropriate selection of past data for modeling, and many other issues (including size of data set). An example predictive question would be “How much will I have in my budget next quarter?” This category of techniques provides both descriptive statements and causal statements, descriptive statements in the sense of “Here is a description of what could happen,” and causal statements in the sense of “If these were to be the circumstances, this is then what could happen.”

**Prescriptive techniques:** these take advantage of approaches that can suggest various courses of action in order to answer the question “What should happen?” Prescriptive techniques incorporate

and build on predictive intelligence, which on its own is very often probabilistic and uncertain (i.e., “There is a 65% chance of rain.”). Prescriptive techniques attach this predictive intelligence to advice and recommendations (i.e., “Based on that we suggest you take an umbrella.”).

Prescriptive intelligence is very often used as part of an augmented decision-making framework. An example question would be “How should I change my spend profile to avoid deficit?” This category of techniques provides statements that are tied to values or goals.

There is another very relevant ability that is worth considering in this question asking and answering context, although it is arguably not an analytics technique, nor a direct output of analytics, nor an answer to questions. That ability is persuasion. For example, once you have generated a description of a situation through analysis, how do you convince people of the “rightness” of the analysis? How do you convincingly present the resulting answer to a question that was asked? In a business context, it is not enough to know (or suspect) that something is or will be the case, and providing technical details will not help. Rather, you need to communicate and convince other people to listen to and to trust your answer. We will consider this aspect of decision support further in subsequent reference chapters.

## CHAPTER 4 ESSENTIALS

### ANALYSIS TECHNIQUES AND OUTPUT STATEMENTS

#### **Data Visualization and Basic Analytics**

Output Statements: Preliminary versions of descriptive statements, and to a lesser extent, causal statements, but these must be backed up by other techniques.

#### **Metrics**

Output Statements: Descriptive statements that support situational awareness and that can also act as hypotheses for ad hoc experimentation.

#### **Statistics and Modeling**

Output Statements:

- A is also B .
- A is often X.
- A has the property X.
- When A, typically B.

### **Optimization**

Output Statements:

- The best possible A is B (approximately), and that happens when C.
- The worst possible A is B (approximately), and that happens when C.

### **Classic Machine Learning**

Output Statements: Variable, but often of the form “A is an X.” For example, “This picture is a picture of a cat.”

### **Generative AI?**

Output Statements: Can produce a wide variety of types of statements, but caution is advised.

## MATCHING QUESTIONS WITH TECHNIQUES

Decision-making support is also framed by asking questions.

- **Descriptive:** What is happening?
- **Diagnostic:** Why is this happening?
- **Predictive:** What will happen?
- **Prescriptive:** What should happen?

Each of these questions aligns with particular analysis techniques.

## REFERENCES

Ricks, E. (2021). [Five questions for better data communications](#) [↗](#) . *Storytelling With Data*.

Muller, J. Z. (2019). [The Tyranny of Metrics](#) [↗](#) . Princeton: Princeton University Press. Project MUSE.

Kashef, M. (2022). *Analytics for Decision-Making*. DataScience2Go.

Boily, P. (2023-2026). [Data Understanding, Data Analysis, Data Science](#) [↗](#) , volumes 1-5. IACS/Quadrangle.

Boily, P., Davies, S., Schellinck, J. (2023). [The Practice of Data Visualization](#) [↗](#) . Data Action Lab.

# V

## GROUP

### DECISION-MAKING

### FRAMEWORKS

We must often make decisions in situations where we ourselves do not have expertise. For example, a doctor asks us to choose one treatment from amongst several options. Even if we have less knowledge than the doctor, it is our responsibility to make the final decision about what to do. More generally, it is not uncommon for an expert to come to us and provide us with information we may not fully understand.

Nonetheless, we are the ones with the authority to make a decision related to this information. These experts are acting as our decision support, but we are the decision maker. In such a situation, it may not always be clear how to successfully integrate their expertise and information into our decision-making process. How do we successfully get more people involved?

Decision-making frameworks are one option for structuring group decision-making. They can add rigour and consistency to decision-making, particularly in a group context. They act as a checklist: did we miss any crucial stages? They act as a guide: here are the stages to go through, and here are the stages where we can add in evidence, information, and data. They can provide built-in strategies for avoiding cognitive biases. In these ways, they can help if we are in a situation where we must make a decision with little expertise or knowledge, but have decision support available that we can integrate into the process.

## V.1 A SELECTION OF DECISION-MAKING FRAMEWORKS

Some common decision-making frameworks are:

- DECIDE: Define the problem, establish the criteria, consider alternatives, identify best, develop and implement a plan, evaluate the plan.
- GOFER: Goal clarification, options generation, facts-finding, effects, review.
- ACE: Assess, choose, execute.
- The 7 Steps: Identify the decision, gather information, identify alternatives, weigh evidence, choose, take action, review.
- Government of Canada / Health Canada Framework: Issue identification (identify issue and context), risk assessment (assess risks and benefits), risk management (identify and analyze options, select a strategy, implement the strategy, and monitor and evaluate the results).

All of these frameworks have commonalities, but different frameworks choose to prioritize, emphasize, or focus on different aspects of the decision-making process. To get a flavour of this, we have selected three of these, each quite different from each other, to discuss more closely. As will become apparent, one challenge with all of these frameworks is that they have not been specifically structured to provide obvious points of involvement for data and analysis.

To assist with this, we briefly discuss where the data and analysis piece fits into each of these frameworks. Then we synthesize the common elements across these frameworks and relate these common elements to the three main decision-making states we have already introduced, namely pre-decision-making, active decision-making, and post-decision-making.

## V.I.I DECIDE FRAMEWORK

A popular decision-making process model is the DECIDE model, developed in 2008 by Hawaiian educator Kristina Guo. It has similar steps to the GOFER (goal clarification, options generation, fact-finding, consideration of effects, review and implementation) model, and Guo developed it to help healthcare managers make better decisions. The acronym stands for:

1. **D**efine the problem.
2. **E**stablish the criteria.
3. **C**onsider the alternatives.
4. **I**dentify the best alternative.
5. **D**evelop a plan and implement the plan of action.
6. **E**valuate and monitor the solution and feedback.

As noted, the first five of these steps are similar to the GOFER model, but the DECIDE model has an additional step, evaluate, which is crucial for improving decision-making processes in the long run. Evaluating our choice will help us learn how successful it was and whether we should make the same decision in the future.

To see how data and analysis fits in, it is useful to recall, first, that statements are produced (or verified) by data and analysis, and second, that there are three different types of statements to keep in mind: descriptive, causal, and value. In the context of GOFER, descriptive statements, derived from data, are useful for more clearly, concretely, and precisely defining the problem, which in turn allows for a full understanding of the problem.

Causal statements can then be used to establish criteria and alternatives that are causally realistic, helping to identify alternatives and evaluate them relative to identified criteria for what counts as a good outcome. Causal statements also support the creation of a realistic plan of action that has a higher chance of being carried out successfully, for example: “According to this causal statement, if I take this action I can realistically expect the following result.”

## V.I.2 ACE FRAMEWORK

The ACE framework is a very streamlined framework put forward by Hasard Lee in his book *The Art of Clear Thinking: A Stealth Fighter Pilot's Timeless Rules for Making Tough Decisions*. As the title of the book suggests, it is intended for use in fast-moving, rapidly changing contexts. The three steps are:

1. Assess.
2. Choose.
3. Execute.

Although this may seem like a very minimal list of steps for a decision-making process, having these three short stages spelled out still draws attention to them in heat-of-the-moment situations.

Here, descriptive facts produced by data and analysis help to more clearly assess the situation at hand and increase situational awareness, while causal statements identify realistic choices relative to the situation and support execution.

## V.I.3 GOVERNMENT OF CANADA / HEALTH CANADA FRAMEWORK

The Government of Canada / Health Canada Framework (GoC Framework) emphasizes risk assessment and management, involvement of those impacted, and the importance of documenting the decision-making process for future reference and evaluation. The Government of Canada (GoC) public website provides a thorough discussion of the approach, which we summarize here.

The framework consists of a series of interconnected and interrelated steps grouped into three phases:

1. Issue identification (identify the issue and put it into context),
2. risk assessment (assess risks and benefits),
3. risk management (identify and analyze options, select a strategy, implement the strategy, and monitor and evaluate the results).

Each step involves a decision point as to whether to proceed to the next step, revisit a previous step, or end the process.

Those involved in the decision-making process (for example, in the case of the GoC this could be partners, the public, and other stakeholders) all play a key role in issue identification, risk assessment, and risk management. The intent is for everyone to provide valuable information, knowledge, expertise, and insights throughout the process.



*Quadrant diagram comparing time + effort with other factors.*

The communication of risk-related information is set up to be an integral part of the process, because the goal of the process is to ensure that both the substance and the process of the decisions taken are acceptable to a broad range of interested and affected parties.

Documentation is also a key aspect of the process. Two types of documentation are prescribed for each step:

1. a description of how the step should be undertaken, including data requirements, assumptions, considerations, and how decisions should be arrived at;
2. a summary of how the step was actually undertaken, the assumptions used, the uncertainties that exist, and how decisions were made, with an explanation given of any plan modification.

The GoC notes that the level of detail of the documentation may vary depending on the issue being addressed and its context.

This is a more elaborate framework than the previous two, with an emphasis on group participation, careful consideration of all aspects of the decision as well as the decision-making process itself, and a focus on the risks and benefits associated with different possible actions. Although it requires substantial time and energy to carry out, it could be appropriate for major decisions that are likely to impact many people.

A major focus of this framework is evaluating and managing risk, and this is an obvious place for data and analysis to come into play. Risk assessment involves a consideration of both potential outcomes and the potential impacts of those outcomes. A clear picture of the starting conditions for this assessment can be provided by descriptive statements, and then causal statements can be used to determine potential outcomes and impacts.



*The three main roles in decision-making*

#### V.1.4 SYNTHESIZING FRAMEWORKS

Looking at these frameworks, it is possible to identify many commonalities and similarities across them. The similarities between framework process components and the decision-making elements reviewed in Chapter 3 are also apparent. Framework components can be grouped together relative to the three states of decision-making discussed in Chapter 3 to provide a high-level synthesis of decision-making considerations at each stage of a decision-making process.

##### **Pre-Decision-Making Activities:**

1. Define knowledge, culture, and goals.
2. Recognize opportunity, values, and past experience.

3. Maintain awareness of the current situation.
4. Define vision, problems, and objectives.

**Decision-Making Activities:**

5. Choose problem/goal, choices, and outcomes.
6. Develop or recognize options.
7. Simulate worlds of consequences (possible worlds).
8. Decide.

**Post-Decision-Making Activities:**

9. Evaluate outcomes.
10. Re-start (if necessary).

The major output from the pre-decision-making stage is a problem, vision, or objective. Sometimes these can be implicit or explicit in the pre-decision-making space, but it is better to make these components explicit going into decision-making if at all possible.

This is particularly important in a group context. The output of active decision-making is taking an action and then the subsequent outcome and consequences of that action. This then gets fed as input into post-decision-making. The output of post-decision-making is an evaluation of the results of the decision.

## V.2 DECISION-MAKING ROLES

Arguably, across all of these frameworks, decision-making responsibilities can be distilled into three basic decision-making roles:

- Problem identifier
- Decision supporter
- Decision maker

Focusing in on the decision maker, the Decision Lab's Dan Pilat and Sekoul Krastev have identified several different types of decision makers:

- **Directive:** decision makers who rely on their own rationality and personal knowledge, disregarding other opinions.
- **Conceptual:** decision makers who explore multiple angles, brainstorm alternatives, and seek creative solutions, though this approach may be time-consuming.
- **Analytical:** decision makers who gather extensive information for well-informed and objective choices, sometimes limiting innovation for the sake of data-backed decisions. Choices and decisions are well-informed and objective.
- **Behavioral:** group-oriented decision makers who present options for discussion, requiring a decisive leader to conclude deliberations effectively.

Different decision-making styles influence the role that facts play in the decision-making process. This in turn ties into the role of the giant grab bag of facts, and data and analysis techniques. The decision-making styles of decision makers and decision supporters also connect to questions about how the individuals involved in a decision-making activity can influence the quality of that exercise. This will be considered through the lens of cognitive biases in the next reference chapter.

## CHAPTER 5 ESSENTIALS

### DECISION-MAKING FRAMEWORKS: A SAMPLE

**DECIDE:** Define the problem, establish the criteria, consider alternatives, identify best, develop and implement a plan, evaluate the plan.

**GOFER:** Goal clarification, options generation, facts-finding, effects, review.

**ACE:** Assess, choose, execute.

**The 7 Steps:** Identify the decision, gather information, identify alternatives, weigh evidence, choose, take action, review.

**Government of Canada / Health Canada Framework:** Issue identification (identify issue and context), risk assessment (assess risks and benefits), risk management (identify and analyze options, select a strategy, implement the strategy and monitor and evaluate the results).

All of the frameworks have common elements and actions (see details in this chapter), but information fits into each of these frameworks in different ways; it is important to identify where and how information connects to the process for a given framework.

#### KEY DECISION-MAKING ROLES

Determine which one of these you are for a given decision-making situation:

- **problem identifier**
- **decision supporter**
- **decision maker**

## REFERENCES

Pilat, D., Krastev, S. (2022). [Decision-making process](#) . Decision Lab.

Guo, K. L. (2020). [DECIDE: A Decision-Making Model for More Effective decision-making by Health Care Managers](#) . *The Health Care Manager*, 39(3), 133–141.

Author unknown (2020). [7 steps of the decision-making process](#) . Lucidchart.

Mann, L., Harmoni, R., Power, C. (1991). *The GOFER course in decision-making*. In J. Baron & R. V. Brown (Eds.), *Teaching decision-making to adolescents* (pp. 61–78). Lawrence Erlbaum Associates, Inc.

Lee, H. (2023). *The Art of Clear Thinking: A Stealth Fighter Pilot's Timeless Rules for Making Tough Decisions*. St. Martin's Publishing Group.

Health Canada contributors (2000). [Health Canada Decision-Making Framework for Identifying, Assessing, and Managing Health Risks, August 1, 2000](#) . Government of Canada.

# VI

## HUMAN FACTORS

Previously, we noted that, in the absence of an all-knowing oracle or all-powerful AI who can make decisions for us, humans have to be involved in making decisions. We also noted that those decisions must be made under conditions of uncertainty. Thus, as humans, we must grapple with the human factors involved in making decisions, even when decision-making is supported by data or facts derived from data.

The humans in the loop who are an integral part of the decision-making process will provide the values driving the process and take responsibility for the end result. In between these two points, cognitive biases can prevent the combined human plus technology approach to decision-making from having the desired impact. A key strategy for dealing with cognitive biases is awareness and recognition at each stage of the decision-making process. Empathy and understanding the reality of the other roles in the decision-making process are also important for attaining optimal decision-making outcomes.

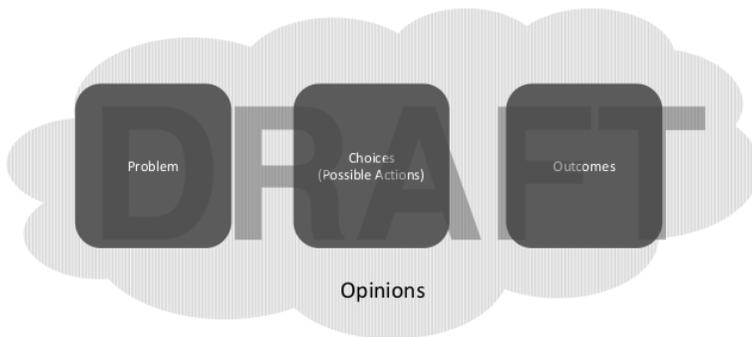
Returning to Kozyrkov’s concerns about data-inspired decision-making, as well as our own concerns about armchair decision-making, where people are influenced by opinions instead of facts, and choices of action are overly influenced by human biases, how can data analysis approaches serve to mitigate these issues?

Suppose that people are making decisions based on a “bag of opinions” in their head, rather than a bag of facts. In such a scenario, one option is to use analysis to convert this bag of opinions into a bag of facts, where opinions that are supported by evidence are kept in the bag and opinions that are not are ejected. How can this exercise be carried

out, practically speaking? There are several options, if (although this could be a big if) people are open to them. People can:

- become aware of and make explicit (either to themselves or others) their opinions, values, priorities, and preferences;
- challenge their own and others' opinions;
- expand and enhance the availability of other opinions;
- validate the truth of their opinions using data and analysis, and be open to the possibility, in the process, that their opinions might be wrong.

What can prevent these opinion evaluating activities from being successful? One issue: cognitive biases!



*A bag of opinions?*

## VI.1 COGNITIVE BIASES

Cognitive biases are patterns of thinking that people reflexively adopt even if these thought patterns often lead people away from objective reality, which is sometimes referred to as the “ground truth”, or “what is actually the case”. Cognitive biases are not the same as biases in the sense of preferences or prejudices, but rather, more broadly, are ways of thinking that are not optimally rational in the moment.

Although cognitive biases may have evolutionarily adaptive origins, their adaptiveness to many situations should be deeply questioned. Be warned! They feel very powerful and they are instinctual. Overcoming them or mitigating them takes dedication and practice.

Researchers continue to identify new cognitive biases that prevent humans from making optimal decisions. A partial list of cognitive biases, presented in the form of questions (in order to help you start to question them), includes:

- **Anchoring Bias:** Why do we compare everything to the first piece of information we have received?
- **Action Bias:** Why do we prefer doing something to doing nothing?
- **Bandwagon Effect:** Why do we support opinions as they become more popular?
- **Confirmation Bias:** Why do we favour our existing beliefs?
- **Belief Perseverance:** Why do we maintain the same beliefs, even when they are proven wrong?
- **Bounded Rationality:** Why are we satisfied by good enough?
- **Choice Overload:** Why do we have a harder time choosing when we have more options?
- **Commitment Bias:** Why do we favour our existing beliefs?
- **Decision Fatigue:** Why do we make worse decisions at the end of the day?
- **Empathy Gap:** Why do we mispredict how much our emotions influence our behaviour?
- **Framing Effect:** Why do our decisions depend on how options are presented to us?
- **IKEA Effect:** Why do we place disproportionately high value on things we helped to create?
- **The Sunk Cost Fallacy:** Why are we likely to continue with an investment even if it would be rational to give it up?
- **Zero Risk Bias:** Why do we seek certainty in risky situations?



*Some common cognitive biases.*

Cognitive biases can have an impact in all three decision-making states: pre-decision-making, active decision-making, and post-decision-making. Below we'll showcase three cognitive biases from the list above, Anchoring Bias, Bandwagon Effect, and the Sunk Cost Fallacy, and consider how each one might be mitigated.

### VI.1.1 THE ANCHORING BIAS

The first piece of information we receive on a topic acts as a reference point from which we interpret new information. For example, if a customer first sees an item at its “original price”, and then subsequently sees the item at a discounted price, they will evaluate the new price as a good deal, separate from the actual amount of money.

This bias could prevent us from correctly adjusting our plan timeline when we discover that our initial timeline estimations were inaccurate. We will be likely to adjust the timeline relative to the original time estimates, even if they have been shown to be wildly inaccurate.

**How can we mitigate the effects?** Recognize the existence of the anchor. Come up with concrete reasons why that anchor is inappropriate. This has been shown to reduce the anchor effect in experiments. Consider multiple options. Use red teaming, where people are designated to play “devil’s advocate”, and challenge the ideas of the group. Exercises that expose the weaknesses of alternatives can inoculate peo-

ple against any hidden anchoring influence (remember that biases are unconscious!).

### VI.1.2 BANDWAGON EFFECT

This refers to our tendency to go along with the majority opinion or behaviour (“go along to get along”, “if you can’t beat ’em, join ’em”, etc.). This overrides reasoning faculties that are essential for critical thinking. It might be a helpful evolutionary heuristic, but it is not always adaptive.

**How can we mitigate the effects?** Allow some time and space before you make your decision or take your position when the information was conveyed to you in such a way that made the majority opinion clear. Try to put yourself in an environment where social pressure is reduced. Make it a habit to also “think for yourself”!

### VI.1.3 SUNK COST FALLACY

This is our tendency to persist in an endeavour simply because we have already invested resources into it (for example, time, money, energy, etc.). We cannot do anything about past irrecoverable costs, but we can do something about the future. We are biased away from looking rationally at current alternatives.

**How to mitigate its effects?** Rather than fixating on irrecoverable past losses, focus on the present and the future with a fresh perspective. Studies have shown that when we are deterred from making decisions based off emotions, this bias is significantly reduced. Could turning to AI or automated decision-making also help with this? For example, automated notifications could be generated at regular intervals during a project suggesting a cost benefit analysis of continuing versus stopping.

## MATCHING BIASES TO DECISION-MAKING STAGE

Having taken a close look at a small selection of cognitive biases and how they might impact decision-making, we’ll conclude with a survey of which biases we suspect might have the most impact at different stages of decision-making.

### **Taking in new statements (beliefs/facts)**

- Anchoring Bias
- Bandwagon Effect
- Base Rate Fallacy

### **Handling currently held statements**

- Confirmation Bias
- Belief Perseverance
- Commitment Bias

### **Choosing/selecting among options**

- Action Bias
- Choice Overload
- Bounded Rationality
- Decision Fatigue
- Empathy Gap
- Zero Risk Bias
- Decision Evaluation
- IKEA Effect
- The Sunk Cost Fallacy

In general, what can be done to mitigate the influence of cognitive biases in decision-making? Is there an approach that can increase the influence of data relative to biased opinions and choices? Is there a way that data support roles can combat bias, maybe by fighting fire with fire? In the next reference chapter we will consider how data storytelling might be used to influence people away from their biased positions.

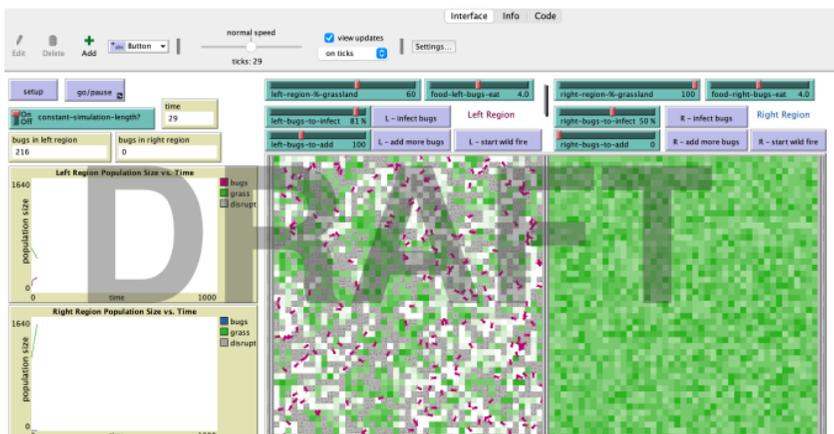
However, this use of data storytelling presumes both that data is available and also that participants in the decision-making process are acting in good faith. So, before moving on to this topic, it is important to briefly consider what happens when these elements are not present.

## VI.2 DECISION-MAKING WHEN THERE IS NO (OR NOT ENOUGH) DATA

The emphasis on using facts generated from data to support good decision-making presumes either that we have access to the relevant data and can carry out the relevant analysis on that data, or that someone previously had access to that data, carried out the appropriate analysis, and we are fortunate recipients of the resulting facts.

However, we have already noted that decision-making always proceeds under a condition of at least some uncertainty, broadly speaking. Within this context, it is very often the case that “the right data” is not currently available, and may be difficult or almost impossible to get.

What should be done when a person feels they should have access to certain facts in order to carry out good evidence-based decision-making, but the data that can provide this evidence is not available? In the most extreme version of this, there are situations or events that have never, or almost never, happened before, and so it is not possible to have much data on them. These are sometimes referred to as “black swan events”, a reference to the non-zero yet relative rarity of black swans.



*A simulation of a situation, used to generate synthetic data.*

Even in conditions of (possibly extreme) uncertainty, it is rarely the case that absolutely no information is (or could prove potentially) relevant to the decision-making situation.

Given this, a good start is to explicitly identify statements that we would like to be able to judge as true or false, even if we do not currently have the data to allow us to do this. Then, we can also explicitly state the current assumptions we are making about the truth and falsity of these statements, in the face of our lack of knowledge. These assumptions can then be updated in a principled fashion if or when more information becomes available.

As well, it may be possible to generate what is referred to as synthetic data, which is data produced by models or simulations that have been created based on more fundamental or basic facts. This synthetic data can then be used to support the assumptions that have been made, and increase the probability that these assumptions are valid. In this way, even in the absence of exactly the right data, the decision-making process can still proceed in a principled fashion, again with the understanding that assumptions will be updated during the decision-making process when new information comes in.

## VI.3 BAD FAITH DECISION-MAKING

In everything we have discussed so far, we have been assuming that people wish to engage in the decision-making process in good faith. By that we mean that those involved in a decision-making activity are committed to carrying out or supporting the decision-making task at hand in order to achieve results that are as optimal as possible relative to the goals and values held by the people who have instigated the decision-making process and/or who have been granted the authority to make or participate in the decision and take responsibility for the result.

However, it can also be the case that people are not engaging in a decision-making process in good faith. If someone will personally benefit (or believe they will personally benefit) from one choice of action over another, and thus have a vested interest in a particular choice being made, they may use strategies to have that choice be the one selected.

Or, if someone does not agree with the particular goal that has been set, preferring a different goal, or does not agree that what has been defined as a problem is actually a problem that requires some action, or has different values than those driving the decision-making process,

this can motivate that person to push for choices of action that may not have the intended or stated outcome.

The diverting, influencing, or even sabotaging of decision-making activities that can arise from these situations is different from a situation where people may (or more realistically will) be operating under unconscious cognitive biases but still have good intentions and want to honour the stated goal and values. In the case of bad faith decision-making, people might deliberately withhold or misinterpret information in order to manipulate or coerce people into making certain decisions, rather than others. They may even feel this is the right thing to do depending on how their values line up with the different options.

This must be kept in mind when engaging in activities like data analysis, and also when presenting the results of these analyses. Data storytelling, the topic of the next reference chapter, can be a powerful tool for compelling action, so care must be taken to prevent bad actors from using such approaches to lead away from data-supported decision-making. Why data storytelling is powerful, and how it can be appropriately grounded in data and facts will be the topic of the next reference chapter.

## REFERENCES

Pilat, D. and Krastev, S. (2019), [Cognitive Biases](https://thedeclarationlab.com/biases) . The Decision Lab. <https://thedeclarationlab.com/biases>.

## CHAPTER 6 ESSENTIALS

### COGNITIVE BIASES

Cognitive biases are patterns of thinking that people reflexively adopt even if these thought patterns often lead people away from objective reality, which is sometimes referred to as the “ground truth”, or “what is actually the case”. Some well-known examples: the Bandwagon Effect, Confirmation Bias, and the Sunk Cost Fallacy.

There is a long list of known cognitive biases (see this chapter), with more being discovered all the time. We can combat these biases by being aware of them and taking deliberate steps to mitigate them.

### DECISION-MAKING WITHOUT ENOUGH DATA

It is very common to be in a situation where you do not have enough of the right kind of data. In this case: identify statements that would be valuable if their truth status were known, look for more general data that could be connected to these statements, and consider using simulations and simulated data to generate insight.

# VII

## DATA STORYTELLING

In her article “*Data Storytelling: How to Effectively Tell a Story with Data*”, Catherine Cote defines data storytelling as the ability to effectively communicate insights from a dataset using narratives and visualizations.

Data stories help either to tell stories about situations that exist or, if using forecasting or synthetic data, to articulate stories that reflect a potential (future) situation. They can be used to put data insights into context for your audience and inspire actions.

According to Cote, a data story has three key components:

1. **data:** the foundation of the data story,
2. **narrative:** the storyline used to communicate the insights recommended actions gleaned from data and context, and
3. **visuals:** representations of data, analysis results, and narratives, which are used to communicate stories clearly and memorably (charts, graphs, diagrams, pictures, or videos).

Data storytelling shines in situations where there is the opportunity to carry out multiple iterations of decision-making activities in advance of the decision, flesh out the various options and then present one or more to decision makers who have been less directly involved in the decision-making process prior to taking their decision.

## VII.1 WHEN TO USE STORIES

In general, when does it make sense to use storytelling? And why is storytelling effective? In his book *“Sapiens: A Brief History of Humankind”*, Yuval Noah Harari argues that storytelling is effective because humans default to thinking in stories.

Humans use stories for many reasons, including explaining, persuading, and building trust. All of these are relevant to the decision-making process.

Consider a situation where you have played a supporting role in a decision-making process. What did you wish the decision makers knew? Did you want them to know how things were going in a certain situation? Something new or interesting? Something problematic or unfair? As a decision supporter, one of the challenges of briefing up to the decision makers is being succinct about all of these types of updates without also omitting key details. Crafting and telling stories make it easier to do this. The process also helps the storytellers themselves understand what the person receiving the story would consider to be the key details in the decision-making situation, and determine how to best communicate those details.

## VII.2 STAGES OF DATA STORYTELLING DEVELOPMENT

When creating a story, data or otherwise, it is recommended that the storyteller start by thinking about the audience. Who is the audience, and what do they want? For example, do they want confidence, for example, peace of mind that they are making the right decision? Do they want opportunity, for example, the chance to be the “hero” in their own story? Do they want foresight, for example, early warning of risk-related information? Or do they want something else? Some further questions you can ask to gain insight into the audience are:

- What are their needs?
- What are they seeking?

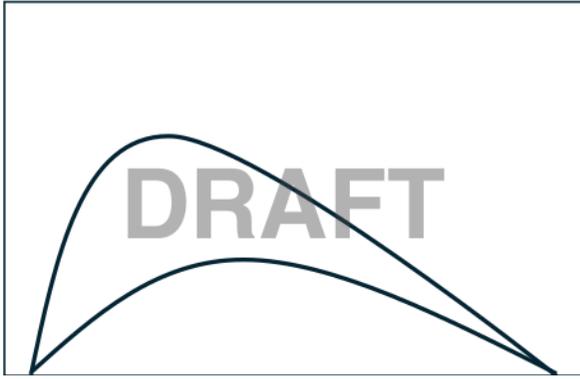
- Why are they seeking this?
- What decisions do they need to make?
- What data and analysis can you provide?
- Why are data visualization and data storytelling important to you?

As the Tetlock quote from the second reference chapter highlights, storytelling with data is importantly different from storytelling without data. Storytelling without data (or facts) is markedly less constrained. Unscrupulous actors who are good storytellers can use storytelling to convince and persuade even in the absence of facts. Alternatively, storytelling with data is constrained by facts, and that makes the storytelling process a bit different.

One challenge here is that the data required to support or inform the stories we hope to tell about the real world might not actually be available. Or we may be using synthetic data that reflects possible worlds rather than actual worlds. We need to keep all of this in mind when crafting data stories, and carefully distinguish between ones that are intended to portray what is versus ones portraying what could be.

Analysis and visualizations are cornerstones of data storytelling. In general, when working with data, we carry out analysis and create visualizations iteratively. From a storytelling perspective, this is reminiscent of the process behind investigative journalism:

1. Initially, we scope out the area of investigation (data collection, story scope).
2. Then we explore the situation and the data we have collected about it. This will help us further develop potential stories.
3. We may use the outcome of this exploration to explain the situation to our own satisfaction, which will further shape the story.
4. We then reveal this to others, possibly to persuade them about some course of action that should be taken with respect to the situation using the story or stories we have developed.



*A visualization that strongly illustrates a particular data story related to number of cases during the pandemic.*

Importantly, data analysts have agency throughout this process. They select:

- the exact question to answer,
- what data to collect,
- how to clean that data,
- which analytical method(s) to use, and
- on what part(s) of the data to focus.

This impacts the stories that will be told with the data, relative to all of the stories that possibly could be told about the situations and events that are represented by the data.

## VII.3 DATA VISUALIZATIONS FOR STORYTELLING

It is our position that visualizations are the illustrations to the data story, not the story itself. Think of a visualization with no text added to it as being like a photograph with no caption. Although there is the saying “a picture is worth a thousand words”, we suggest that, even if it would take a thousand words to describe the picture, the picture itself is not a

story, but rather a snapshot in time of a situation that could be related to a story, or used to inspire a story.

Data visualizations are effectively pictures of data, and data visualizing requires a very particular skillset, involving an understanding of data, Aesthetics, graphic design, and illustration principles. Data visualizers must first select appropriate visualization types out of a large catalogue of types of visualizations. Consideration must then be given to Aesthetics and cognitive load, so the viewers of the visualization can extract something meaningful from it. Visualizers must know which parts of a larger data story the visualization is intended to support, so they can structure it to highlight relevant aspects of the data and possibly add text to supplement, or even outright provide, the data story being told.

## VII.4 DATA STORYTELLING PRESENTATIONS

Suppose that, as a decision supporter, you have done your analysis and are now looking to influence decision-makers using a data storytelling presentation. You yourself must now decide which analysis results to include and which to omit in the presentation, as well as how to structure the presentation narrative. McKinsey and Company and the Boston Consulting Group have created a well-known, well-crafted method for telling compelling and easy-to-understand data stories in presentation form that can help with this.

First, in structuring the storyline, consider three key elements: situation (Why is this important? What do you need to know to understand the problem?), complication (What went wrong? Why is the problem a terrible thing?), and resolution (What is the solution?).

Once you have thought through these elements, create the presentation as follows:

1. Start in a text editor, not a slide editor. It is easier to see the storyline and logical connections on a single page, instead of flipping through slides. Drafting is an iterative process, and a text document is easier to modify.

2. Use the Dot-Dash Method to structure slides once you have the material created in the slide editor.
3. Dots:
  - Slide title
  - The key statement
4. Dashes:
  - The supporting data

When considering the skeleton of the presentation, try reading the slides in a “horizontal” manner, looking just at the titles of the slides (the dots). It should be possible to understand the argument by just reading the titles. Then, the “vertical” component of the presentation, the data and visualizations (the dashes) on a slide, supports the title of the slide.

## VII.5 DATA STORYTELLING DASHBOARDS

When decision supporters are tasked with creating data dashboards, they are often instructed to make sure there is data storytelling involved. However, data dashboards can have functions that do not require or lend themselves to data storytelling. Because of this, the presumption that every dashboard should have a storytelling element, and subsequent efforts to incorporate data storytelling into every dashboard, can lead to a confusing, frustrating, and ultimately unsuccessful dashboard creation experience. In order to avoid this, it’s important to determine in advance which of three main types of dashboard, data exploration, situational awareness, or storytelling, you are dealing with.

**Data Exploration Dashboard:** Dashboards are often used for data exploration during the preliminary stages of analysis. Visualizations may be grouped together on dashboard canvases and collectively used to extract patterns from the data. These patterns might then be used to develop a narrative, but they may not have a narrative themselves.

In such cases, it is expected that the visualizations and dashboard have:

- a high level of interactivity,
- a high level of detail,
- potentially all, or many, aspects of data represented (tables, columns, calculations, etc.),
- a low level of graphic design,
- no annotations or explanations required.

**Situational Awareness Dashboard:** Dashboards can also be used for situational awareness purposes. This is arguably the original *raison d'être* for dashboards. In this case, the data dashboard acts like the dashboard in a car, where the purpose is to report in real or near real time key aspects of what is happening in the environment (internal or external). The dashboard will likely contain simple dynamic visualizations that change over time as the situation changes. Frequently, the visualizations will be created to indicate when particular threshold values have been reached (for example, a red line indicating when a particular value in the environment has passed a threshold of concern).

Decisions supported by this type of dashboard will likely be carried out quickly (the car is going over the speed limit; I will step on the brakes to slow it down). As a result, this type of dashboard is not a good candidate for data storytelling because it has been designed so that choices are clear-cut and decision-making will be rapid. However, it is possible that some dashboard outputs could be saved for later use in data storytelling (for example, later provided in a presentation format).

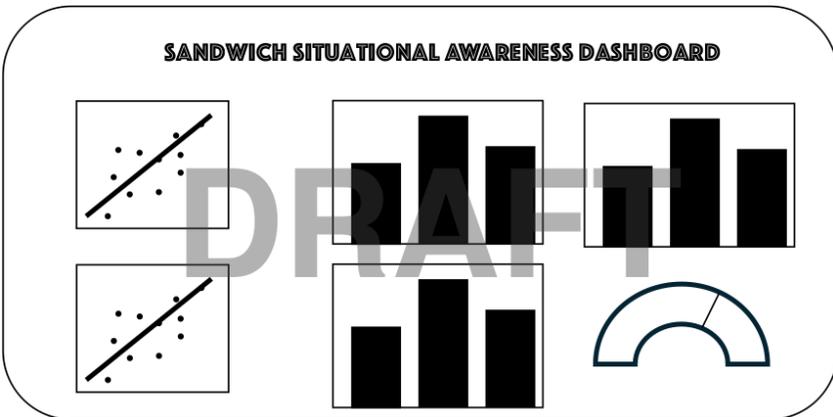
**Storytelling Dashboard:** Lastly, but not least, a dashboard can act as a data storybook. In this case, the visualizations in the dashboard are used as a tool to explain the data and support the story that is being told. The expectation here is that there will be:

- a low level of interactivity,
- a low level of detail,

- a slow rate of change (relative to a situational awareness dashboard),
- key aspects of the data represented,
- a high level of graphic design,
- annotations and explanations that drive the “story”.

The storytelling dashboard will often consist of multiple panels or canvases, where each panel could be thought of as a chapter in the story. Multiple visualizations and accompanying text will likely be added to each panel, and the audience will navigate through these collections of visualization.

When using this approach, Scott McCloud’s book “*Understanding Comics*” is an excellent reference guide. Although it focuses on the medium of comics, much of it will be applicable to storytelling with dashboards.



*An example of a dashboard.*

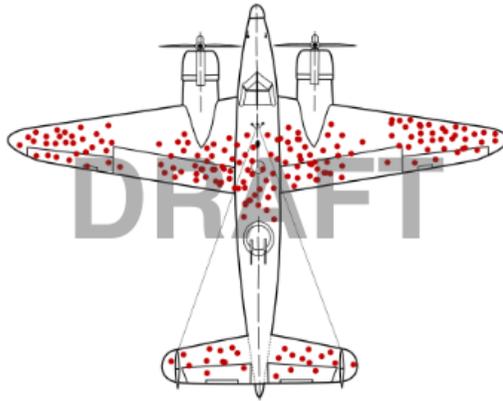
## VII.6 DATA STORYTELLING RISKS

A good story can help to provide insights on a situation, but storytelling requires choices, and the outcome is affected by what is included and what is omitted in the story. With data stories, there is an additional complication relative to plain-jane stories: we want to tell a story that is supported by data, but we only have access to the available data. The

data that was not collected, no doubt for a variety of reasons, is by definition not accessible to us. This can lead to compelling yet fundamentally flawed data stories, which are either not truly supported by the data or which tell a story about something that is only adjacent to the decision-making process.

A very popular example of how care must be taken during data storytelling happened during WWII. Mathematician A. Wald undertook a study to help protect British bombers flying over enemy territory. Data included the number and location of bullet holes on returning aircraft, and the goal was to use this information to determine where to add armor to best protect the plane's structure.

A chart was created to show where the maximum number of bullet holes were located on returning aircraft. This chart showed greatest damage on the aircraft extremities, not on the main wing and tail spars, engines, and core fuselage areas. As such, the Air Ministry wanted to add armor to the extremities.



*A diagram showing the location of bullet holes on returning planes.*

Fortunately, Wald was on hand and suggested they were dead wrong. In fact, to avoid survivorship bias, armor should be added to the areas with the fewest holes: if no returning planes had holes in their wing spars and engines, then even a few holes in those locations were deadly.

The takeaway is that the data that is missing may be as, or more, important to the story than the data that is there; storytelling is not always an obvious endeavour. As can be seen from the information in this chapter, data storytelling is powerful and valuable, but must be used with caution and attention to the true story of the data.

## CHAPTER 7 ESSENTIALS

### **Three Main Components of Data Storytelling**

- data: the foundation
- narrative: the storyline
- visuals: representations of data, illustrate the story

### **Three Different Types of Dashboards**

- Situational Awareness
- Data Storytelling
- Exploratory

**Structuring Data-Heavy Presentations:** (digestible, motivating)  
Dot-Dash Method of McKinsey and Boston Consulting Group.

## REFERENCES

Knaflic, N. (2015). *Storytelling with Data*. John Wiley & Sons.

Cote, C. (2021). [Data Storytelling: How to Effectively Tell a Story with Data](#) . Harvard Business School Online.

Harari, Y. N. (2015). *Sapiens: A Brief History of Humankind*. Harper-Collins Publishers.

Galletta, D. (2023). [Storytelling in PowerPoint: Learn McKinsey's 3-Step Framework](#) . YouTube.

McCloud, S. (1993). *Understanding Comics: The Invisible Art*. Tundra.

Moll, C. (2022). [Abraham Wald and the airplane diagram with red bullet holes: here's the origin story](#) . Authentic Boredom.

# VIII

## DATA ENGINEERING FOR DECISION-MAKING

Kozyrkov's concept of data-driven decision-making readily lends itself to automated decision-making, where a decision-making process is defined in advance, analysis engines are created and embedded in a data pipeline and data is then fed into that pipeline to generate results that can trigger appropriate decisions and automated actions, depending on pre-determined types of outputs of the engines. This type of fully automated decision-making system could be viewed as a fairly extreme version of data-driven decision-making, where cognitive biases are essentially dealt with by removing the most obvious source of the biases, humans.

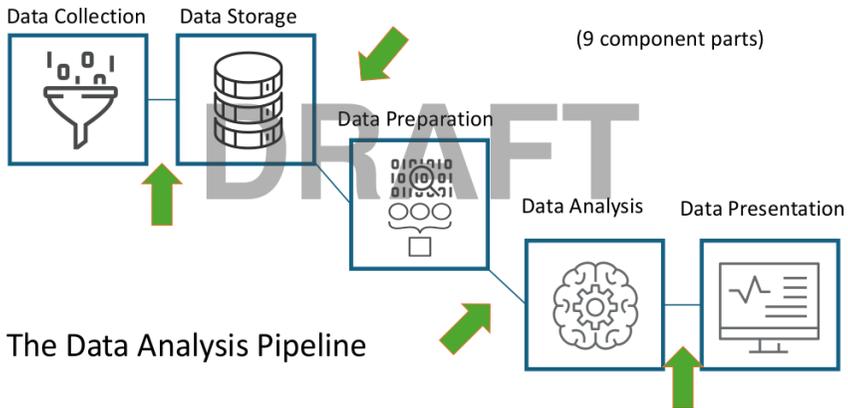
Even if decision-making is not automated to this extent, more human-in-the-loop versions of data-driven decision-making will still benefit from a well set up data infrastructure, designed by data architects, implemented by data engineers, and supported by data managers and good data governance. In this chapter we will consider all of these aspects of modern data-supported decision-making.

### VIII.1 DATA PIPELINES

A data pipeline is an interconnected series of digital components through which data flows and is transformed. Data pipeline components and associated activities may be a combination of manual and automated.

Data pipelines typically have five main components (plus the connectors between these components, which are important for allowing data to flow between the components). Pipeline components provide the following functionality.

- **Data Collection:** data is brought into the pipeline through connections to multiple endpoints, which could be some combination of application programming interfaces (APIs), data dumps, downloads/uploads, forms, data collection applications, activity logs, etc. At the point of collection, data is referred to as raw data.
- **Data Storage:** data is stored in structured data repositories. Some of this structure might reflect how the data was collected and some of this structure might reflect how the data will be used. Data may be stored in multiple locations, may be duplicated across locations, and may be transformed while being transferred into or out of storage.
- **Data Preparation:** data preparation is the transformation of previously collected data into a format that can be used as high-quality input for a particular analysis technique or analytical engine. Some of the transformation may relate to the form or the type of the data (for example converting string data into numeric data). Some of the transformation may relate to making the data more correctly represent relevant aspects of the world, or deleting data that clearly does not (for example, deleting a row of data if an age is listed as a negative number). Data preparation might even include using the data to create models or AI engines that are used to further refine the data prior to data analysis. Data preparation is typically the most labour-intensive aspect of the data pipeline, and may be highly complex and iterative.
- **Data Analysis:** the transformation of data into statements or other data presentation components, via application of techniques, models or engines. There can be overlap and iteration between data preparation and data analysis.
- **Data Presentation:** the generation of reports, visualizations and other outputs, to be used by consumers of these data pipeline results.



Typical data pipeline creation phases are similar to the phases seen in other software engineering projects, and include, at a high level:

1. Research + Design
2. Implementation
3. Testing
4. Production + Management
5. Back to Research + Design

Although these steps are presented sequentially here, in practice it is not unusual for them to be occurring somewhat in parallel, particularly once a pipeline has been launched.

### VIII.1.1 DESKTOP DATA PIPELINES

Over the past decade or more, business intelligence needs have pushed forward the development of desktop data analysis tools and pipelines, which may be provided either as desktop applications or as software as a service (SaaS) platforms accessible through web browsers. The ready availability of these tools is leading to further democratization of data, as well as increases in data and digital literacy among people who are not data specialists.

This increasing ability to access and work with data is likely going to make it easier for organizations of all sizes to carry out data-supported or evidence-based decision-making.

Desktop data analysis, with its associated small scale data pipelines, is not necessarily a substitute for “industrial”, “professional”, or enterprise data pipelines, but it is a good start.

### VIII.1.2 ENTERPRISE DATA PIPELINES

Getting a solid enterprise business intelligence pipeline, which carries out basic data analytics and data visualizations, up and running is a major stepping stone in an organization. Moving forward from that, artificial intelligence/machine learning/data science (AI/ML/DS) engines can also be integrated into the pipeline, although it is worth noting here that the data architecture and tools needed for AI/ML/DS analysis may not be the same as those required for more straightforward business analysis needs, supported by basic analytics and data visualizations.

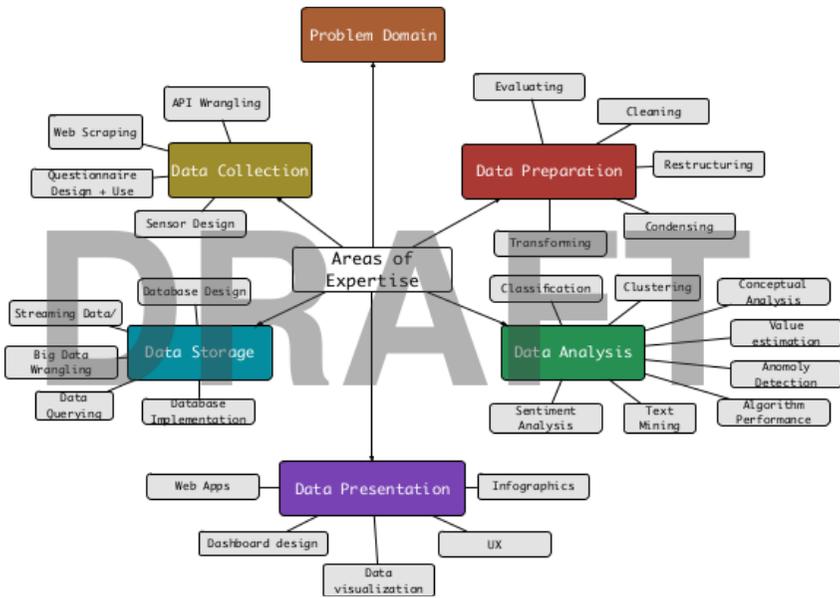
In particular, if the addition of AI/ML/DS engines are made to the pipeline, some parts of the pipeline may need to be redesigned to support new AI/ML/DS requirements. Database architecture may need to change, and it could be worth considering less traditional data infrastructure (for example, data lakes, NoSQL) at this point.

### VIII.1.3 DATA PIPELINE ROLES

Data infrastructure does not build itself, data governance does not happen by itself, and many different parts of preparation and analysis still cannot be automated. As a result, a team of people with the right skills are needed to support different parts of data-driven decision-making. Analysts often view themselves as the stars of these teams, but a rule of thumb for the ratio of other team members to analysts for an enterprise level data analysis pipeline in practice is around 8:1.

Across the pipeline, data roles include:

- **Data Engineer:** Focus on data infrastructure design and implementation, with a strong emphasis on IT and DevOps.



*Key data pipeline roles.*

- **Data Collector:** Focus on design of data collection strategies and implementation of data collection tools.
- **Data Architect:** Focus on data storage and data architecture design.
- **Data Manager:** Focus on day-to-day management of the data and data infrastructure.
- **Data Preparation:** Works hand in hand with the data collection, the data managers and data analysis members of the team to get the data into a state where it is ready for analysis. To automate this requires the design and implementation of processes to carry out all of the relevant steps. This is a pivotal position on the team.
- **Data Analyst:** Analysts determine what analysis can work with the information available, and can give insight that is relevant and useful. The analyst designs algorithms that can be used to automate these analyses.
- **Data Pipeline UX Expert:** The champion of interface design and user experience. Often an advocate for the end-user, which is critical to success.

- **Data Communication:** Focus on data visualization and data presentation more generally.
- **Subject Matter Expert:** Knows a lot about the situation, understands what is important, what data could provide insight, and how to interpret and apply the results of the analysis.
- **Business or Organization Strategy Expert:** Holds the big picture and knows where the organization wants to head. The strategy expert needs to provide this information to the team.
- **Project Lead:** The focus is on keeping everyone on track and working together.
- **Data Translator:** Knows how the different pieces of the pipeline work at a high level, and knows something about the subject matter and business relevance. Good at connecting people and helping them talk to each other.

## VIII.2 DATA GOVERNANCE

Keeping data in an unorganized digital pile is not much more useful than having no data at all. Such data cannot easily be used to support data-driven decision-making. This means that data governance is critical if data is to be used effectively for decision-making. The Data Management Association (DAMA) is a well-recognized data governance professional organization, and they have produced a useful reference for data governance practices, the DMBOK (Data Management Body of Knowledge), which is currently on its second edition.

The DMBOK provides a substantial list of data governance roles, which are distinct from the data pipeline roles described above. Here we have reduced the list to a smaller number, to give a general sense of the type of governance roles that are important to consider in this space.

- **Data Owner:** A person who has governance and compliance responsibility for a set of data assets. They are accountable for compliance, risk, oversight, approval, championing governance and issue resolution. This person is often in a managerial or executive position in an organization.

- **Data Steward:** A person who has business accountability for a set of data assets. They are accountable for accuracy, consistency, business requirements, metadata definition and management, quality, curation, fitness for purpose, governance, inventory, research data management and role management. This person is often in an operational/business position in an organization and may or may not be working with the data themselves as a part of their responsibilities, alongside the data users and consumers.
- **Data Custodian:** A person who has technical accountability for a set of data assets. They are accountable for security, access management, availability, capacity, continuity, safeguarding, implementation, technical standards, configuration, control, modeling, versioning, change management. This person is often in an IT/data related position in an organization.
- **Data Contributor:** A person who creates or collects data that is relevant to the organization. They are accountable for acquisition and entry, quality, metadata preparation, ethical and secure gathering of data, identification of issues, identification of new data sources. This person may be in an operational/business position in an organization or an IT/data related position.
- **Data User/Data Consumer:** A person who uses data to enable business outcomes and make decisions. They are accountable for ethical use, reporting on quality, reporting on fitness for purpose, identification of business and data rules, identification and reporting of data control, use in line with governance. This person will likely be in an operational/business position, and may also be in a managerial or executive position.

### VIII.3 DATA USE PRINCIPLES

When organizations are working with data, in a pipeline or otherwise, they must be mindful of ethical considerations. Some basic principles include:

- **“Do No Harm”:** Data collected from an individual should not be used to harm the individual.

- **Informed Consent:**
  - Individuals must agree to the collection and use of their data.
  - Individuals must have a real understanding of what they are consenting to, and of possible consequences for them and others.
- **Opt-In/Opt-Out:** Informed consent requires the realistic ability to opt out.
- **Respect Privacy:** Only collect the amount of data necessary and anonymize data whenever possible. It is fair to say that data privacy can be hard to maintain in the age of constant trawling of the Internet for personal data, but it is still important.
- **Open Data:** When possible, data should be viewed as a public resource and made available to the public. This must be balanced at all times by the principle of respecting privacy.

## REFERENCES

Dama International. (2017). *DAMA-DMBOK: Data Management Body of Knowledge* (2nd Edition). Technics Publications, LLC, Denville, NJ, USA.

## CHAPTER 8 ESSENTIALS

### **Data Pipeline Components**

- Data Collection
- Data Storage
- Data Preparation
- Data Analysis
- Data Presentation
- + Connectors Between Components

### **Data Pipeline Roles**

- Data Preparation
- Data Analyst
- Data Pipeline UX Expert
- Data Communication
- Subject Matter Expert
- Business or Organization Strategy Expert
- Project Lead
- Data Translator

### **Data Governance Roles**

- Data Owner
- Data Steward
- Data Custodian
- Data Contributor
- Data User/Data Consumer

*Schellinck, Shaeen, Boily, Davies, Stroud*

# INDEX

- ACE, 50, 52, 57
- actions, 6, 11, 13, 14, 17–19,  
    21–25, 28, 31, 40, 45,  
    50–52, 55, 57, 66, 67,  
    69, 71, 79
- actions, possible, 20, 22–24, 28,  
    32, 40, 54
- actions, possible choices of, 22,  
    40, 59, 66, 67
- Aesthetics, 35, 73
- AI engines, 42, 80, 82
- algorithm, 7
- algorithms, 83
- Amitai Etzioni, 2, 3
- Amos Tversky, 2
- analysis paralysis, 23, 24
- analysis techniques, 4, 12, 14–18,  
    27, 30, 31, 33–35, 39, 40,  
    43, 44, 46, 47, 56, 59,  
    60, 63, 65, 67, 72, 80
- analysis, armchair, 9, 10, 15, 17,  
    18
- analysis, cost benefit, 45, 63
- analysis, risk benefit, 4, 24, 45
- analysis/analytic engines, 4, 79,  
    80
- analytic engines, 79, 80
- analytics, 30, 33, 34, 44, 46, 48,  
    82
- analytics Basic, 34
- analytics, Basic, 33, 34, 46, 82
- analytics, descriptive, 33, 44, 47
- analytics, diagnostic, 33, 44, 47
- analytics, predictive, 33, 44, 47
- analytics, prescriptive, 33, 44,  
    47
- Andrew O’Connell, 1, 2, 8
- Annie Duke, 6–8
- Antonio Domasio, 2
- artificial intelligence/AI, 4, 14,  
    45, 59, 63, 80, 82
- assumptions, 7, 53, 66
- audience, 69, 70, 76
- Bayesians, 39
- belief, 10, 26, 61, 64
- Ben Wattenberg, 5
- Bias, 2, 5, 7, 10, 12, 16, 18, 29,  
    59–64, 68, 77, 79
- bias, action, 61, 64
- bias, anchoring, 61, 62, 64
- bias, cognitive, 2, 5, 10, 12, 16, 18,  
    29, 49, 56, 59–64, 67,  
    68, 79

- bias, commitment bias, 61, 64
- bias, confirmation bias, 61, 64, 68
- bias, zero risk, 61, 64
- black box, 42, 43
- black swans, 65
- bounded rationality, 2, 61, 64
- briefing up, 70
- business intelligence, 4, 44, 81, 82, 84
- Cassie Kozyrkov, 5, 7, 9, 12, 13, 15, 17, 59, 79
- categorization, 41
- causation, 22, 25, 28, 37, 39, 45
- charts, 34, 35, 44, 69, 77
- Chester Barnard, 2
- choice, 3, 6, 8, 9, 11, 13, 14, 16, 18, 22, 24, 25, 32, 40, 50–52, 55–57, 59, 61, 62, 64, 66, 67, 75, 76
- choice overload, 6, 61, 64
- choice tunnel vision, 23
- choice, forced, 23
- choices, 20, 22–24, 27, 28, 31, 32
- classification, 41, 42
- classification pants, 42
- clustering, 41
- cognitive load, 73
- colour theory, 35
- consequences, 17, 19, 20, 24, 25, 31, 32, 55, 86
- constraints, 2, 16, 17, 22, 23, 25, 40, 71
- correlation, 39, 45
- criteria, 13, 16, 40, 41, 50, 51, 56
- Dan Pilat, 56, 57, 67
- Daniel Kahneman, 2, 5, 7, 8, 29
- dashboard, data exploration, 74, 78
- dashboard, situational awareness, 44, 74–76, 78
- dashboard, storytelling, 74, 78
- data, 3–5, 7–9, 12–18, 22, 25–27, 30, 31, 33–38, 40, 41, 43–46, 48–54, 56, 59, 60, 64–87
- data analysis, 12, 14, 15, 17, 18, 27, 50–52, 54, 59, 67, 69, 71, 73, 74, 80, 82–84, 87
- data analysis, desktop, 81, 82
- data analyst, 39, 72, 82, 83, 87
- data and information gathering, 22, 85
- data architect, 79, 83
- data architecture, 82, 83
- data architecture design, 79, 83
- data as fuel, 30
- data assets, 84, 85
- data collection, 27, 38, 71, 80, 83, 85–87
- data collection strategies, 83
- data collection tools, 80, 83
- data collector, 83
- data column, 36, 37
- data communication, 34, 39, 46, 48, 84, 87
- data consumer, 80, 85, 87
- data contributor, 85, 87
- data custodian, 85, 87
- data engineer, 79, 82
- data engineering, 79

- data exploration, 34, 35
- data gathering, 22, 85
- data governance, 79, 82, 84, 85, 87
- data governance roles, 84, 85, 87
- data infrastructure, 79, 82, 83
- data infrastructure design, 82, 83
- data journalism, 5, 8
- data literacy, 81
- Data Management Association (DAMA), 84
- data manager, 79, 81, 83, 85
- data owner, 84, 87
- data pipeline, 79–85, 87
- data pipeline roles, 83, 84, 87
- data pipelines, desktop, 81, 82
- data pipelines, enterprise, 82
- data pipelines, small scale, 82
- data pipelines, ‘industrial’ or ‘professional’, 82
- data preparation, 80, 82, 83, 85, 87
- data presentation, 80, 84, 87
- data privacy, 86
- data row, 27, 80
- data steward, 85, 87
- data storage, 80, 83, 87
- data storage design, 83
- data storytelling, 34, 35, 64, 67, 69, 71–78
- data storytelling presentations, 73–75
- data table, 26, 27, 36
- data translator, 84, 87
- data use principles, 85
- data user, 85, 87
- data visualization, 33–35, 44, 46, 69, 71–76, 80, 82, 84
- data visualizer, 35
- data warehouse, 4
- data, democratization of, 81
- data, high quality, 14, 18, 27, 80
- data, merging, 4
- data, metadata, 85
- data, open, 86
- data, synthetic, 65, 66, 69, 71
- data, unstructured, 27
- data-driven, 5, 9, 12, 13, 15, 17, 18, 22, 79, 82, 84
- datum, 26, 34
- DECIDE, 50, 51, 56, 57
- decision analysis, 1, 3, 4, 7, 8
- decision fatigue, 61, 64
- decision intelligence, 1, 4, 5, 7, 8
- decision maker, 2, 7, 15, 18, 22, 30, 35, 39, 49, 55–57, 69, 70, 73
- decision science, 1, 5–7
- decision support, 27, 46, 47, 49
- decision support systems, 4
- decision supporter, 55–57, 70, 73, 74
- decision trees, 4, 16, 45
- decision-making, 1–15, 17–33, 37, 39, 40, 42, 44, 46–59, 61–70, 75, 77, 79, 82, 84, 85
- decision-making activities, 11, 12, 17, 18, 20, 21, 23, 25, 29, 32, 55, 66, 67, 69

- decision-making approach, 9, 11–14, 18, 27
- decision-making elements, 20, 21, 27, 31–33, 54
- decision-making exercise, 21, 56, 59
- decision-making frameworks, 6, 46, 49, 50, 54–57
- decision-making frameworks, synthesizing, 50, 54
- decision-making outcomes, optimal, 59, 66
- decision-making process, 4, 11–15, 19–21, 25, 30, 31, 33, 35, 39, 49–54, 56, 57, 59, 64, 66, 69–72, 77, 79
- decision-making stage, 59, 63
- decision-making state, 19–22, 31, 37, 50, 54, 62
- decision-making system, automated, 13, 18, 79
- decision-making tactics, 2
- decision-making, active, 19–22, 27, 31, 50, 55, 62
- decision-making, armchair, 9–12, 17, 18, 22, 59
- decision-making, automated, 13, 18, 63, 79, 82, 83
- decision-making, bad faith, 66, 67
- decision-making, data-driven, 3, 9, 12, 13, 17, 18, 22, 79, 82, 84
- decision-making, data-inspired, 9, 12, 13, 15, 17, 18, 59
- decision-making, evidence-based, 12
- decision-making, good, 2, 27, 64–66
- decision-making, group, 49, 54–56
- decision-making, organizational, 2
- decision-making, post, 19, 31, 37, 50, 55, 62
- decision-making, pre, 19, 21, 22, 31, 37, 50, 54, 55, 62
- default choice, 15, 16
- descriptive techniques, 44
- design thinking, 35
- DevOps, 82
- diagnostic techniques, 44
- digital literacy, 81
- DMBOK (Data Management Body of Knowledge), 84
- Do No Harm, 85
- documentation, 52, 53
- Dot-Dash Method, 74, 78
- Dr. Sekoul Krastev, 56, 57, 67
- DUDADS, 48
- effect, bandwagon, 61–64, 68
- effect, framing, 61
- effect, IKEA, 61, 64
- emotion, 2, 12, 18, 22, 61, 63
- emotion, role of, 2, 12, 18, 22, 61, 63
- empathy, 59
- empathy gap, 61, 64
- engine, 41–43, 79, 80, 82
- evaluating, 19, 28, 31, 37, 50–52, 54–57, 60, 64

- evidence-based decision
  - making, 59, 65, 67, 82
- evidence-based
  - decision-making, 1, 5, 7, 12, 49, 50, 57
- evidence-based medicine, 5, 8
- expert based systems, 4
- fact, 11, 12, 26, 27, 35, 37
- factors, 2, 13, 16, 22–25
- facts, 7, 10–12, 15, 17, 18, 27, 31, 50–52, 56, 59, 64–67, 71
- facts, giant grab bag of, 11, 56, 59
- falsity/false, 26, 66
- Frances Westely, 20, 32
- frequentists, 39
- Garbage Can Theory, 2, 8
- Gary Klein, 6–8
- Generative AI, 33, 42, 43, 47
- Gerd Gigerenzer, 2, 3
- gestalt principles, 35
- goal, 9, 11, 19–22, 24, 27, 28, 31, 37, 41, 46, 50, 51, 53–56, 66, 67, 77
- GOFER, 50, 51, 56, 58
- Gordon Guyatt, 5
- Government of
  - Canada/Health
  - Canada Framework, 50, 52, 57, 58
- graphic design, 35, 73, 75, 76
- ground truth, 60, 68
- grounding, 11, 43, 67
- Henry Mintzberg, 20, 32
- Herbert Simon, 2
- heuristics, 2, 5, 63
- Howard Raiffa, 4, 8
- human factors, 59
- human-in-the-loop, 59, 79
- illustrations, 72, 73
- implementation, 50–52, 56, 57, 81–83, 85
- index/indices, 36
- information, 4, 9, 12–18, 20, 22, 24, 25, 27, 28, 31, 32, 43, 49, 50, 53, 56, 57, 61–63, 65–67, 77, 83, 84
- information gathering, 22
- informed consent, 86
- interface design, 83
- Internet of Things (IoT), 4
- issue identification, 50, 52, 53, 57, 85
- James March, 2, 8
- Jerry Z. Muller, 36
- knowledge, 25, 26, 49, 53, 54, 56, 66
- Kristina Guo, 51, 57
- labelling, 41
- language, 25, 26, 41
- learning, 41
- Leigh Buchanan, 1, 2, 8
- machine learning, 4, 7, 33, 41–43, 45, 47, 82
- machine learning, Classic, 33, 41–43, 47
- mathematics, 3–5, 7, 36, 37

- McKinsey and Company and the Boston Consulting, 73
- McKinsey and Company and the Boston Consulting Group, 78
- measurement, 26, 27, 36
- measures, raw, 36, 80
- metric, 35–37
- metrics, 33, 35–37, 41, 46
- model, 4, 33, 36, 39, 41, 45, 46, 66, 80, 85
- model, mathematical, 36
- modeller, causal, 39
- narrative, 69, 73, 74, 78
- neural net classifier, 42
- neural networks, 42
- Ole Delgado, 4, 8
- operations research, 4, 5
- opinion, 10, 12, 18, 27, 56, 59–61, 63, 64
- opinions, bag of, 59, 60
- Opt-In/Opt-Out, 86
- optimization, 3, 7, 33, 40, 41, 47
- options, 69
- outcome, 3, 9, 11, 13–15, 17, 19–25, 27, 28, 31, 32, 37, 41, 51, 54, 55, 67, 71, 76
- persuasion, 46, 70, 71
- Philip E. Tetlock, 9, 10, 18, 71
- planning, effects based, 23
- population, 37, 38, 43
- possible worlds, 24, 25, 55, 71
- predictive intelligence, 45, 46
- predictive techniques, 45
- preference, 11, 15, 22, 28, 30, 31, 60
- prescriptive intelligence, 46
- prescriptive techniques, 45, 46
- priorities, 13, 22, 31, 60
- problem, 7, 9, 19–22, 28, 29, 31, 40, 50, 51, 55, 56, 66, 73
- problem identifier, 55, 57
- production, 81
- products, 20, 21, 24, 25, 32
- project lead, 84, 87
- property, 22, 23, 36, 38, 47
- property, object, 36
- proposition, 25
- question answering, 46
- reality, 25, 26, 59, 60, 68
- reasoning, 14, 26, 29
- representation, 3, 80
- representing, 25, 26, 80
- research and design, 81
- results, 19, 20, 24, 25, 32, 50–52, 55, 57, 59, 66, 67, 69, 73, 79, 80, 84
- Richard Thaler, 6–8
- risk assessment, 3, 50, 52–54, 57
- risk management, 3, 50, 52–54, 57, 84
- risk-related information, 3, 53, 70
- sample, 37, 38, 43
- Scott McCloud, 76
- Sheena Iyengar, 6–8
- simulation, 39, 55, 65, 66, 68
- situational awareness, 4, 19, 21, 28, 37, 46, 52, 55, 74–76, 78

- stakeholder, 53
- statement, 21, 25–35, 37–40, 42–47, 51, 64, 66, 68, 74, 80
- statement, causal, 28, 29, 32, 34, 35, 37–39, 44–46, 51, 52, 54
- statement, descriptive, 28, 29, 32, 34, 37–39, 44–46, 51, 52, 54
- statement, value, 28–30, 32, 44, 51
- statistical techniques, 3, 37
- statistics, 33, 36–41, 43–46
- statistics, descriptive, 38
- statistics, inferential, 38
- storytelling, 35, 69–78
- strategy expert, 84, 87
- subject matter expert, 10, 31, 84, 87
- target, 20, 21, 31, 36
- testing, 44, 81
- The 7 Steps, 50, 57
- The Practice of Data
  - Visualization, 48
- The Sunk Cost Fallacy, 61–64, 68
- The Tyranny of Metrics, 36, 48
- trained engine, 41, 43
- training, 7, 41, 42
- true/truth, 26, 27, 35, 37, 39, 43, 60, 66, 68
- uncertainty, 4, 5, 8, 14, 46, 53, 59, 65
- unconscious bias, 7, 63, 67
- Understanding Comics, 76
- user experience, 35, 83
- user experience expert, 83, 87
- user experience studies, 35
- values, 11, 12, 18, 22–24, 28, 30–32, 37, 41, 45, 46, 54, 59–61, 66, 67



Making good decisions is hard. Making good decisions with data can be even harder... but also easier when approached properly. This is a practical guide to evidence-, data-, and information-based decision-making for two audiences: decision-makers who want to use data well, and data practitioners who want to support decisions effectively.

In concise reference chapters, we cover decision-making and decision theory, survey essential data and analysis ideas, and focus on how to combine them in practice. You will learn when data helps, when it does not, and how to collaborate across roles without losing sight of the decision itself.

Read it cover-to-cover or dip in by topic for a clear framework, a shared vocabulary, and tools for better decisions.

