Table of Contents

# 1.3 Lifecycle, Quality & Quantity of Data

## Introduction

Successful analytics projects follow a structured lifecycle and depend on high‑quality, sufficient data. This section outlines the lifecycle phases and discusses data quality and quantity considerations.

## Data analytics lifecycle

* **Define the problem** – Clarify objectives and key questions.
* **Collect data** – Acquire relevant data from databases, surveys or sensors.
* **Prepare data** – Clean, merge and transform raw data into an analysable form.
* **Analyse** – Apply statistical and computational methods to extract insights.
* **Interpret and communicate** – Visualise results and draw conclusions for stakeholders.
* **Act and monitor** – Implement recommendations and monitor outcomes for continuous improvement.

## Data quality

Data must be accurate, complete, consistent, timely and reliable. Poor data quality can lead to misleading conclusions. Data cleaning (handling missing values, duplicates and errors) improves quality and should be performed before analysis.

## Data quantity

The amount of data affects the reliability of insights. Too little data results in high uncertainty, while very large datasets require scalable tools (big data platforms). Sampling methods may be used when analysing entire populations is impractical.

## Example

In healthcare analytics, continuous monitoring devices collect high‑frequency physiological data. Analysts must ensure the data is complete (no missing vital signs) and accurate (calibrated sensors) before using it to predict patient deterioration. Adequate sample size ensures the predictions are reliable.

## Summary

The data analytics lifecycle guides the process from problem definition to action. Ensuring data is of high quality and adequate quantity is essential for trustworthy results.

## Reflection questions

1. Why is data cleaning necessary before analysis?
2. What risks arise when data quantity is insufficient?
3. Describe a situation where data quality affected an analytics outcome.

## References