

3.1A-I · SUPPLIED CHAPTER PP. 57–80 · FOUR TEACHING ARCS; CHOOSE TO FIT
PRIOR LEARNING

Data records part of a situation

Identify three kinds of data in the scene. Which reason for watching would remain an inference?

AI conceptual illustration · fictional scene, not case evidence



One reading becomes four different claims

01

DATA: 31



02

INFORMATION:
room
temperature = 31
°C at noon



03

KNOWLEDGE:
repeated
readings indicate
a hot room



04

WISDOM:
choose
ventilation after
checking safety
and alternatives

Fictional room example. DIKW is a thinking model, not an automatic machine pipeline.

Context and judgement add meaning; more data alone is not wisdom.

The question determines which data you need

Context	Recorded example	Useful question
Financial	Sale amount: \$12	How much revenue?
Medical	Patient report: tired	What experience needs investigation?
Weather / geographical	Rain: 8 mm / site coordinates	Where and when did it rain?

Scientific observations can use these forms too. A category is not a guarantee of accuracy.

Quantitative and qualitative data describe different aspects of a situation.

A photo has content and metadata

IMAGE CONTENT

A photograph of a garden

Illustrative record, not real personal metadata. Location is not present in every file.

Metadata can reveal something the visible image does not.

METADATA

Capture time, file size, creator field; location if recorded

A sensor measurement becomes a digital record

01

Temperature
changes
continuously



02

Sensor produces
a signal



03

ADC samples
and quantizes
the signal



04

Digital value
stored with unit
and time

Simplified analogue-sensor route; some sensors produce digital output internally.

Sampling and conversion create a representation of a physical quantity.

Find a pattern, or link records about an entity?

MINING

Find repeated associations
between products in purchases

Fictional shop records. A wrong or duplicate identifier can create a false match.

Data mining and data matching answer different questions.

MATCHING

Join sales and returns using the
same transaction ID

Check A–C: explain the meaning of a record

RECORDED

12

What unit and context?

LINKED

**Same order ID in
two tables**

What establishes the match?

INFERRED

**Customers prefer
this product**

What evidence is still missing?

Individual first attempt: 3 minutes. Export/print and hand to the teacher when this check is assigned.

Use terminology to justify a distinction, then correct one weak inference.

Five stages manage a record over time

01

CREATE:
event
booking
submitted



02

STORE:
database
with
assigned
access



03

USE:
prepare
attendance
list



04

PRESERVE:
retain
evidence for
a stated
need



05

DESTROY:
remove after
justified
retention
ends

Fictional community event. Retention depends on actual obligations; no fixed legal period is invented.

Preservation and destruction are distinct decisions.

Primary or secondary depends on the investigation

PRIMARY DATA

Our event survey collected for this investigation

The same dataset can be primary to its collector and secondary to a later investigator.

The relationship to collection matters, not whether the file is online.

SECONDARY DATA

An earlier survey dataset reused for a new question

An entity becomes records with typed fields

Field	Type	One record
BookID	String identifier	B017
Copies	Integer	3
Available	Boolean	True

Fictional Books table. Rows = records; columns = fields; each book is an entity instance.

A database schema states what each field can hold.

Keys link entities without repeating every detail

Books	Loans	Relationship
BookID B017 (primary key)	LoanID L101 (primary key)	Loan identifies one borrowing event
Title: Coast Birds	BookID B017 (foreign key)	Loan refers to a Books record
Copies: 3	Return date: 2026-10-20	One book can appear in many loans

Fictional records. A relational database links related tables; having several unrelated tables is not sufficient.

The primary key identifies a record; a foreign key refers to another table.

Textbook layer: Hodder 2022, supplied chapter pp. 65–67. Original classroom examples are fictional.

Suitable is different from accurately copied

VALIDATION

Age 17 passes an allowed 0–120 range

Fictional records. A range check cannot confirm the true age.

A valid value can still be the wrong value.

VERIFICATION

Source says 71: compare or double-enter to detect the mismatch

Classify risk before deciding access

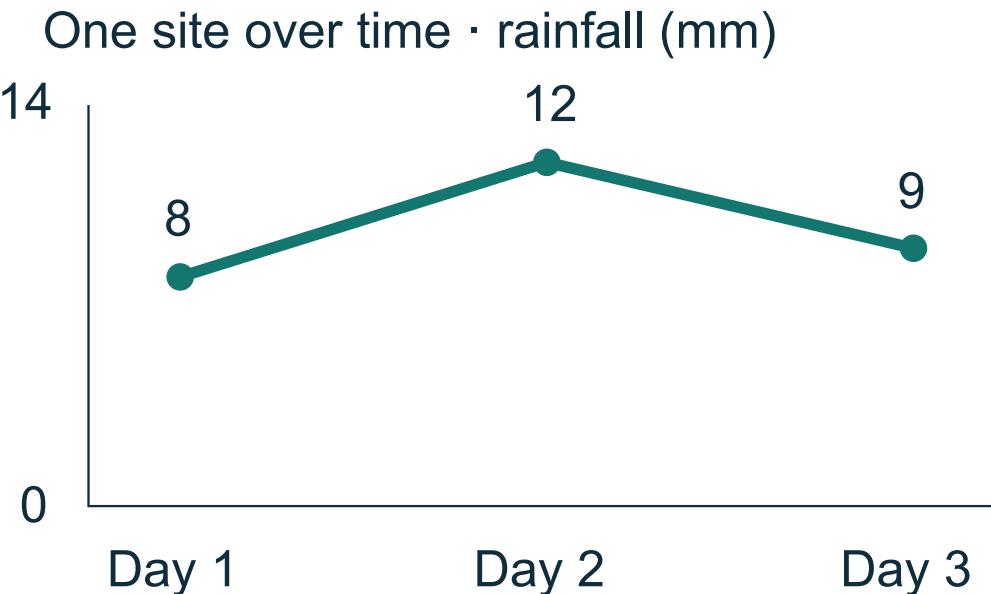
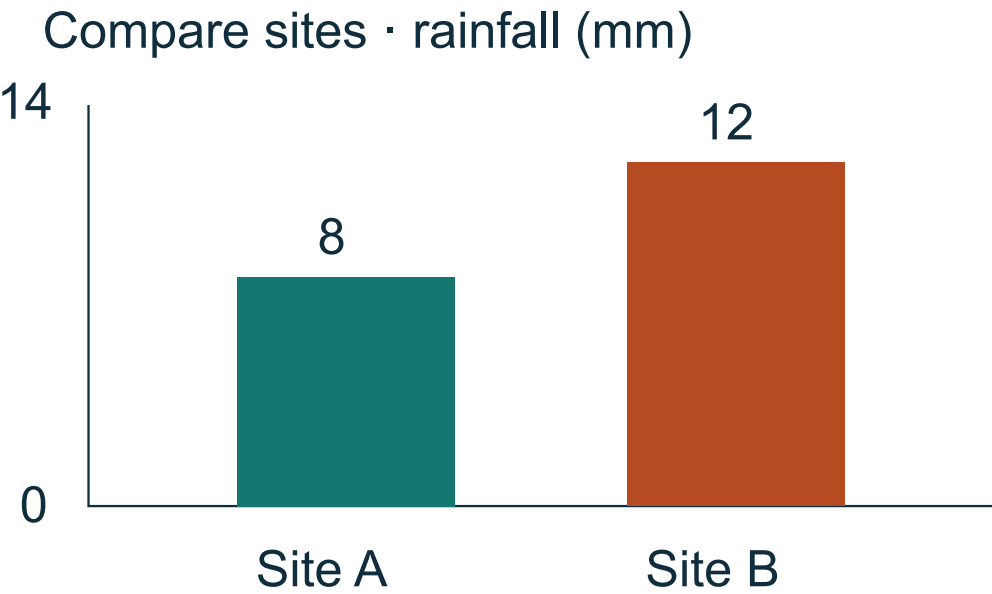
Record	Classification concern	Access decision
Public opening hours	Low confidentiality concern	Publish with accuracy checks
Payment account	Confidential financial details	Restrict to authorized role
Health accommodation request	Sensitive personal information	Minimize and separate access

Fictional event records. Personal, sensitive, confidential and public are not mutually exclusive legal categories everywhere.

Sensitivity and purpose determine permissions; categories may overlap.

Textbook layer: Hodder 2022, supplied chapter pp. 66–67. Original classroom examples are fictional.

Choose a chart that fits the question



Synthetic data · labelled axes, stated units and a common 0–14 scale. A line connects time observations; it does not establish cause.

A chart needs a suitable comparison, units and honest scale.

Check D–F: build a defensible data record

DESIGN

One entity with three typed fields

Identify a unique key

CHECK

A plausible wrong value that passes validation

Add verification

REPRESENT

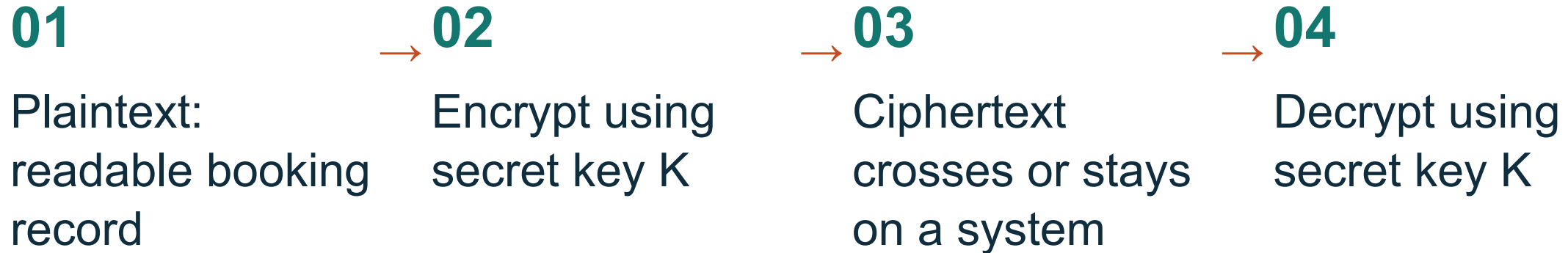
Choose a labelled chart

Match the question

Individual 4-minute sketch; no real learner records. Export/print and hand to the teacher when assigned.

Explain a lifecycle, a relationship and a representation.

Encryption changes readable data into ciphertext



Conceptual model, not runnable cryptography. The key must remain protected.

Symmetric encryption uses the same secret key to encrypt and decrypt.

Public-key encryption separates the two key roles

SENDER

Uses the recipient's public key
to encrypt

RECIPIENT

Keeps the corresponding
private key secret to decrypt

Conceptual encryption roles. Digital signatures are a different operation; real TLS combines mechanisms.

Encrypt for the recipient with their public key; decrypt with their private key.

Masking and encryption solve different problems

MASKED TEST RECORD

Real name replaced with a fictional value for testing

Illustrative approach. Dates, locations or linked records may still identify someone.

Replacing visible details does not automatically remove re-identification risk.

ENCRYPTED STORED RECORD

Readable record transformed; recovered with a permitted key

Deletion is not a disposal guarantee

ORDINARY DELETION

A file disappears from the usual view; copies may remain

VERIFIED SANITIZATION

Applicable clear, purge or destroy process, with verification

NIST 2025 update. Overwriting and degaussing are not universal solutions for every device.

Choose sanitization for the storage medium, sensitivity and required assurance.

A linked ledger protects record history, not truth at entry

01

Participant
submits a
transaction



02

Network applies
its validation
rules



03

Transaction is
recorded in an
agreed block



04

Later block
includes a link to
prior block

Simplified blockchain. Different networks use different consensus/access rules; transactions are often grouped into blocks.

Hash links and network agreement make some changes detectable or difficult.

Check G: match the safeguard to the failure

Failure to analyse	Candidate response	Remaining question
Intercepted readable file	Encryption	Who controls the keys?
Testers receive private details	Mask selected data	Can records be linked?
Retired drive contains records	Verified sanitization	Does the process fit the medium?

Individual 3-minute first attempt. Keep corrections; export/print and hand to the teacher when assigned.

Use a safeguard’s mechanism and limit, not just its name.

Textbook layer: Hodder 2022, supplied chapter pp. 68–72. Original classroom examples are fictional.

Big data and data analytics

Volume: amount

Velocity: arrival/processing rate

Variety: forms

Veracity: trustworthiness

A useful dataset fits the question; more records alone cannot repair missing context or systematic omission.

Big data involves large, varied collections of records

Volume and velocity

Situation	Characteristic to examine
A ten-year archive	Volume
Thousands of new events each second	Velocity
A daily batch analysed overnight	Arrival and processing delay

Volume describes the amount of data

Variety and veracity

VARIETY: DIFFERENT FORMS

Video + caption + device
setting + interaction log

More formats and more rows do not repair the same underlying error.

Variety concerns formats and types of data

VERACITY: TRUSTWORTHINESS

Correct units? Duplicates?
Missing fields? Outdated
labels?

Small and clean, or large and messy?

A · 40 VERIFIED RECORDS

One class. Useful for that class
—but does it represent the
school?

B · 40,000 IMPERFECT RECORDS

Duplicates and missing fields.
Wider reach—but what errors
remain?

Synthetic datasets. Choose against a defined question, not size or cleanliness alone.

Dataset A: 40 verified records from one class

A missing value is an unresolved question

SAME RECORDS. DIFFERENT ASSUMPTION.

8 **blank**

Exclude blank: $8 \div 1 = 8$. **Assume zero:** $(8 + 0) \div 2 = 4$.

The interface cannot tell you which assumption is justified. Investigate the missing record.

Blank duration does not mean zero duration

An association does not establish a cause

1. Observed association →

2. Alternative explanation →

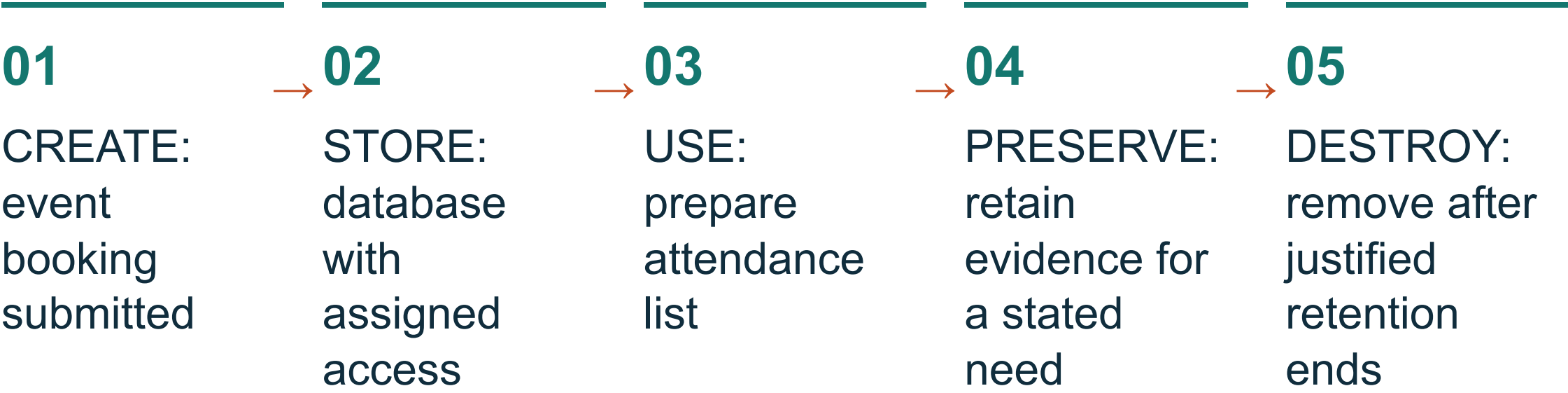
3. Comparison evidence →

4. Qualified causal claim

A possible explanation requires testing. Correlation alone cannot identify intention or establish cause.

A pattern may be useful for prediction

The data lifecycle creates decisions at every stage



Fictional community event. Retention depends on actual obligations; no fixed legal period is invented.

Create or collect data

Collection can shape the eventual result

1. Defined purpose →

2. Choose relevant fields →

3. Check omitted groups →

4. Decide retention

Collection choices shape both evidence and privacy costs.

Collect data that fits a defined purpose

Analytics can support several purposes

DESCRIBE

What happened?

Last month's activity

DETECT

What looks unusual?

A possible false alert

PREDICT

What may happen next?

Compare later observation with the estimate

Different questions create different errors and evidence needs.

Describe a past pattern

A bounded big-data inquiry

01

Question: are new creators represented?



02

Inspect collection, formats and missing records



03

Tie each claim to a source and a limit

A provider account and regulator account do not supply a complete dataset of individual experiences.

Question: whose experience does this dataset describe?

Reliability and integrity are separate tests

RELIABILITY

Are values complete, accurate and current for the decision?

Fictional contact list: an outdated correct entry differs from a tampered current entry.

Accurate, complete data can become unreliable through unauthorized changes.

INTEGRITY

Has the record been changed or compromised improperly?

Privacy and anonymity are different conditions

PRIVATE, IDENTIFIABLE

Named medical record
restricted to authorized care

APPARENTLY ANONYMOUS

Unnamed post whose
time/location reveals the author

Conceptual examples. Privacy includes information control, freedom from intrusion and freedom from surveillance.

A person can be identifiable while their information remains restricted.

Data protection starts before collection

GDPR principle group	Event-booking example
Lawfulness, fairness, transparency	Explain the basis, use and recipients
Purpose limitation / minimization	Collect needed booking fields; do not reuse by default
Accuracy	Provide a route to correct wrong details

EU framework, not a universal school policy. This table covers four of seven principle groups.

Purpose, minimization and accuracy constrain the data pipeline.

Retention and responsibility finish the protection plan

GDPR principle group	Decision to document
Storage limitation	How long is justified, and for what need?
Integrity and confidentiality	Who may access or alter the record?
Accountability	Who can demonstrate that rules are followed?

Remaining three of seven principle groups. Erasure rights have conditions and exceptions.

A protection promise needs a retention rule, safeguards and accountable decisions.

A result difference does not establish its cause

2019

NIST · face recognition
evaluation

Source paraphrase · read the cited original

SUPPORTS THIS CLAIM

Performance differences varied by algorithm, application and data.

DOES NOT ESTABLISH

The study did not establish a single cause for those differences.

Compare what the report measured with what the textbook infers.

Check I: make a conditional data decision

OPPORTUNITY

Use booking records to plan access

Name a benefit and its evidence

DILEMMA

Reuse individual records for advertising

Explain purpose and control

ADJUSTMENT

Aggregate counts; limit access and retention

State a remaining risk

Individual 4-minute practice. P1: apply knowledge; P2: distinguish chapter claims from NIST evidence. No marks assigned.

A defensible judgement joins mechanism, evidence, stakeholders and a safeguard.

Original excerpt check: pp.72–74

RETAINED ORIGINAL EXCERPT QUESTION

Answer for pp.72–74 only.

The question, choices and saved response ID remain unchanged. The expanded chapter has a separate Check I.

For this retained check, “this deck” means the original pp.72–74 excerpt