

KK06 — Characteristics of data sources (TXT, delimited CSV and XML files)

Programming · Comprehensive exam study guide

STUDY DESIGN — YOU MUST KNOW

- structure
- reasons for use

Solutions often read/write external files. You must know the **structure** of plain text (TXT), delimited (CSV) and XML files and the **reasons** each is used — a frequent exam pairing.

The three file types

File	Structure	Reasons for use
TXT (plain text)	Unstructured lines of characters; no fields defined.	Simple to read/write; good for logs or notes; human-readable.
CSV (delimited)	Rows of values separated by a delimiter (usually a comma); first row often headers.	Compact, tabular data; easily opened in spreadsheets; simple to parse.
XML	Hierarchical, with nested tags that label each piece of data, e.g. <code><student><name>...</name></student></code> .	Self-describing; stores complex/nested data; widely used to exchange data between systems.

Choosing a format

- Use **TXT** when structure isn't needed (free text, logs).
- Use **CSV** for simple tables of records where every row has the same fields.
- Use **XML** when data is nested/hierarchical or must be self-describing and portable between systems.

EXAMPLE

CSV: `id,name,paid\n1,Sara,true` — vs XML: `<student><id>1</id><name>Sara</name></student>`.
CSV is smaller; XML is clearer and supports nesting.

Key definitions to memorise

Delimiter	A character (e.g. a comma) that separates fields in a CSV file.
Tag	A label in XML, written in angle brackets, that names a piece of data and can nest other tags.
CSV file	A plain-text file storing tabular data as delimiter-separated values, one record per line.
XML file	A hierarchical text file that uses nested tags to label and structure data.

Exam technique

EXAM TIP

- When comparing, pair a structural point with a reason-for-use point.
- Say XML is 'self-describing' and supports 'nested/hierarchical' data — these are the key marks.

COMMON MISTAKES

- Saying CSV can store nested data — it is flat/tabular; XML handles nesting.
- Describing XML as comma-separated — it uses tags, not delimiters.

Quick revision checklist

- I can describe the structure of TXT, CSV and XML.
- I can give a reason for using each and choose between them for a scenario.

Exam practice — questions & model answers

Q1. Describe the structure of a CSV file.

Describe · 2 marks

Model answer

- It stores records as lines of plain text (1),
- with fields separated by a delimiter such as a comma, often with a header row (1).

Q2. Compare CSV and XML as data sources, referring to structure and one reason for use.

Compare · 4 marks

Model answer

- CSV is flat/tabular with delimiter-separated values (1) and is compact and easy to open in spreadsheets (1).
- XML is hierarchical with nested tags (1) and is self-describing, suited to complex data exchanged between systems (1).

Q3. State one reason a developer might choose a TXT file over a CSV file.

State · 1 mark

Model answer

- When the data is unstructured/free text (e.g. a log) and does not need fields (1).