Deliverable 5 · Data and security brief

Your project company already has a map of its IT and OT systems, a technology report, a cloud architecture and an AI plan. All of them run on data. Deliverable 5 answers four questions about those data: which data the company has and how they live, what it could learn from them, where they are kept safe, and how well its public front door is locked. The group then decides which laws apply and puts everything in one risk table.

In the city-builder game, this is the chapter where you list every info view the city depends on, check that the archive has copies in another town, walk around the walls looking for open gates (without opening any), and read the laws the city must follow.

What you will hand in

- One PDF per group on Google Classroom, by **Monday 03/05/2027 at 23:59**
- Four individual, signed sections (A, B, C and D) and a group part
- An individual **defence** of 2–3 minutes in session 24 (03/05/2027)

The rules

- **Facts must be cited.** Everything you state as a fact about the company (what data it collects, which cloud it uses, what its website sends) must have a source: the company's website or privacy policy, an annual report, the press, Ibestat, the BOE, a screenshot with date and time. What you cannot find is labelled **hypothesis**, and you explain why it is reasonable.
- **Passive only.** The security audit only **looks** at what the public website shows to any visitor: headers, certificate, cookies, banner, privacy policy. Never scan ports, try logins or send strange input to a real company's systems.
- **No generative AI** in any part of this deliverable, unless the teacher authorises it for a specific task. Unauthorised AI or plagiarism means a 0.
- **Real data only** in section B: a real open dataset, cited, with its download date. No invented numbers.

Sections

Each member owns and signs **one** section (put your name under its title). Pick one you have not done before in other deliverables. With 3 members, section D goes to the group part.

Section	Criteria	What to write	Owner
A · Data inventory and lifecycle	a, b	A table of at least **six kinds of data** the company handles (office and plant): what it is, where it comes from, personal or not, which **information** or decision it allows, found or hypothesis. Then the **six stages of the lifecycle** of the most important dataset, from capture to deletion, with the retention period if the privacy policy gives it. Mark where data become information and the riskiest stage.	Individual

Section	Criteria	What to write	Owner
B · Mini data-science pipeline	c, d, e, h	A question that matters to the company, answered with a **real open dataset related to its sector** (Ibestat, Balearic Government open data, datos.gob.es, INE…). Show each stage: question, source and download date, cleaning problems and decisions, the analysis (a pivot table or formulas), **one chart**, one conclusion and one limitation. Then: which **objective** of data science this serves, whether the company has a **big data** problem (justified with at least two V's) and whether ML would add anything. Link the spreadsheet.	Individual
C · Cloud storage and backups	f, g	For three of the company's datasets (from section A): which **kind of storage** (object, block, file, managed database, warehouse, lake), which deployment model (Unit 3) and **region**, and why. A **3-2-1 backup plan**: where each copy lives, how often it is made, how long it is kept, who can restore it and how a restore is tested. Explain why the cloud matters for this company, with one real price or feature from a provider's page as evidence.	Individual
D · Passive security audit of the website (4th member)	i	The audit of the company's public website: certificate, **securityheaders.com** grade and missing headers, **Mozilla HTTP Observatory** score and two failed tests, cookies set **before consent**, the banner, and the privacy-policy checklist. Screenshots with date and time. For each finding: which pillar (confidentiality, integrity, availability) it affects and one recommended measure. Polite wording: findings, not accusations.	Individual (group if 3 members)
Group · Regulation	i	Which of **RGPD**, **LOPDGDD** and **NIS2** apply to the company and why (sector, size, personal data), with the official references. Spain's NIS2 transposition status, checked in the BOE on the date you write it. Two concrete consequences for the company's applications.	Group
Group · Risk table	i, b	At least **five risks** across sections A–D: the risk, the asset (data or system), the pillar it threatens, the level (device, system, global), the measure, and who in the company would own it.	Group
Group · Integration	all	A short introduction, one diagram or table that connects the four sections (the dataset from A, analysed in B, stored in C, exposed through the website in D), conclusions, the sources list, format and spelling.	Group
Group · Group log	—	Who owned which section, who had which role, what was decided and when.	Group

Roles

Rotate them: nobody repeats the role they had in Deliverable 4. With 3 members, the Editor role merges into the Coordinator's.

Role	In this deliverable
Coordinator	Plans the work back from 03/05, chases deadlines, hands in the PDF

Role	In this deliverable
Source checker	Opens every link, checks that every fact has a source and every hypothesis is labelled; checks the BOE for NIS2
Technician	Helps with the hands-on parts: the spreadsheet of section B, the DevTools steps of section D
Editor	One format and one voice, spelling, numbered figures, the sources list in a single style

Assessment

Deliverable 5 is the 70% of the RA5 mark. The other 30% is the class activities of sessions 20 to 24.

Part	Weight	What is assessed
Your own section	40%	The section you signed, using the rubric below
Your defence	20%	Your answers in the 2–3 minute conversation in session 24
Group part	40%	Regulation, risk table, integration and coherence, sources, format, group log

Rubric for the sections

Section	Excellent (9–10)	Adequate (5–6)	Insufficient (0–4)
A	Six or more kinds of data, clearly separated from the information they give, mostly found and cited; a lifecycle specific to the company, with deletion and retention resolved	The inventory and the six stages are there, but generic or with few sources	Confuses data with information, misses stages, or does not talk about this company
B	A real business question, a cited dataset, cleaning decisions explained, a correct chart, an honest conclusion and limitation; V's and objective well argued	The pipeline is complete but the question is vague or the cleaning is not explained	Invented or uncited data, missing stages, or a chart that does not answer the question
C	A reasoned storage choice per dataset, a complete 3-2-1 plan with frequency, retention and restore test, cloud importance backed by evidence	Correct choices with little justification; the 3-2-1 plan has gaps	No backup plan, or storage choices that do not fit the data
D	All checks done passively, with dated evidence, findings linked to the triad and a sensible measure for each	Most checks done, evidence or measures partial	Checks missing, no evidence, or anything other than a passive check
Group part	Correct and justified application of the three laws, a risk table that uses all four sections, one coherent document, every fact cited, complete log	Laws identified but weakly justified, risk table generic, some inconsistencies	Laws wrong or missing, no risk table, sections that do not fit together

The intermediate marks (7–8 and 3–4) are for work between two columns.

Deadline and how to hand in

- **One PDF per group**, uploaded to the Classroom assignment by the Coordinator, by **Monday 03/05/2027 at 23:59**. Include the link to the section B spreadsheet, shared with the teacher.
- Late work: maximum mark 5; more than 48 hours late, 0.
- Keep the shared document's **version history**: it shows who wrote what.

The defence

In session 24 the group works on the final version while the teacher calls each of you, one at a time, for **2–3 minutes**. No paper, no computer: a conversation. You will get two questions about **your own section** (why you chose that dataset, what a header does, why that copy is off-site) and one about **another member's** section, to check that the group really put the document together. If you wrote your section yourself, you will be fine.

Final check before you hand in

Check	Done?
Every section has its owner's name	
Every fact about the company has a source; every hypothesis is labelled	
Section B uses a real open dataset, cited with its download date, and the spreadsheet link works	
Section C has all three copies of the 3-2-1 plan and a restore test	
Section D has dated screenshots and only passive checks	
The group part says which laws apply and why, with the NIS2 status checked in the BOE	
The risk table has at least five risks drawn from the four sections	
The group log says who did what and which role each person had	
No generative AI has been used	