

## Reflection Section

This module took me through the full journey of handling data, right from extraction and cleaning to building proper databases, working in teams, tackling data compliance, and exploring machine learning. It's one of the most well-rounded experiences I've had in data so far, and I've picked up both technical and personal lessons along the way.

### **What?**

In the first few units, I got hands-on with data wrangling—working with messy, unstructured datasets and learning how to clean them up using Python. I used tools like Pandas and jsonschema to structure and validate data, which helped me understand what makes a dataset trustworthy. We looked at file types like JSON, XML, and SQL and discussed how each one is used in real-world data systems (McKinney, 2022). I also became more confident identifying data quality issues like missing values, duplication, and inconsistent formatting.

One of the major highlights was our group project in Unit 6, where we had to design a relational database system for a logistics company. My focus was on cleaning the dataset, normalising the tables to 3NF, and preparing the ER diagram. We also submitted a project proposal outlining the tools, assumptions, and logic we'd follow—choosing PostgreSQL for its scalability and support for ACID-compliant transactions (Connolly and Begg, 2015). Then, in Unit 11, we completed the final deliverable—a professional executive summary comparing SQL and NoSQL options, reflecting on GDPR compliance, and recommending how the system could be used to support business goals.

Along the way, we also explored **compliance laws** like the **GDPR** and Nigeria's **NDPA (2023)**. These helped me think beyond the tech and focus on how data is protected and used responsibly. In my comparison, I found that while GDPR and NDPA share a lot in terms of rights and structure, the NDPA is still developing its enforcement strength (NDPA, 2023; ICO, n.d.). I also joined peer discussions comparing data laws in Saudi Arabia, Thailand, and the UK—highlighting how enforcement and exemptions differ from place to place (Alshammari and Simpson, 2023; DLA Piper, 2025).

In the later units, I explored API security. I wrote a short security specification focused on protecting JSON-based APIs—using input validation, role-based access, safe parsing, and rate limiting to

reduce risks like injection and data leakage (OWASP, 2017). Then finally, in Unit 12, we looked at the future of big data—particularly machine learning and its role in automation and decision-making. I found this part especially interesting because it showed me how everything we’ve learned fits into the bigger picture of building smart, adaptive systems.

### **So what?**

This course reminded me that working with data isn’t just about writing code or running models—it’s about thinking critically, designing with intention, and understanding the legal and ethical responsibilities that come with handling information. I used to focus mostly on “getting it done,” but now I see how much structure and planning goes into doing it right.

The group project was a challenge, mostly because not all contributions were equal. Some team members delivered solid work, others didn’t. That was frustrating at times, but it taught me something important: collaboration means letting go of the idea that everyone will approach the task the same way. As someone who runs a small business and is used to doing things my own way, this felt familiar. It reminded me of working with external vendors who sometimes miss deadlines or don’t deliver the quality you expect. You still have to adapt, keep moving, and find ways to get the job done.

Understanding compliance laws like GDPR and NDPA also gave me a more structured way to think about data governance. These laws don’t just exist in theory—they shape how systems are built, what features are prioritised, and how users are protected. I now understand that data security isn’t just a backend task. It affects how users experience a product and whether they can trust the platform (ICO, n.d.; NDPA, 2023).

Designing and building a normalised database also helped me understand why structure matters. Removing redundancy, defining relationships, and enforcing constraints aren’t just for neatness—they’re about ensuring data integrity, avoiding anomalies, and building systems that scale without falling apart (Connolly and Begg, 2015).

## **Now what?**

This module has made me more confident in my ability to work with data end-to-end. I know how to extract, clean, structure, and analyse it in a way that's useful and responsible. Going forward, I want to build on this by:

- Getting deeper into Python automation, especially for validation and ETL pipelines.
- Practising with more real-world datasets and open-source projects to sharpen my modelling and reporting skills.
- Staying on top of evolving data laws and privacy standards—especially as I plan to build or support data-focused tools for business use.

This experience has also taught me to document my process better. Whether I'm working alone or with a team, being able to trace decisions, note gaps, and reflect on what worked makes a huge difference.

Overall, this module pulled together everything I'd hoped to learn and pushed me to improve how I work, think, and collaborate. It wasn't just about learning tools—it was about learning how to use them well.

## References

Alshammari, M. and Simpson, A. (2023) 'PDPL vs. GDPR: A Comparative Analysis of Data Protection Laws in Saudi Arabia', *Journal of Information Policy*, 13(1), pp. 85–104.

Connolly, T. and Begg, C. (2015) *Database Systems: A Practical Approach to Design, Implementation and Management*. 6th edn. Harlow: Pearson Education.

DLA Piper (2025) *Data protection laws of the world: Thailand*. Available at: <https://www.dlapiperdataprotection.com/?c=TH> (Accessed: 21 July 2025).

ICO (n.d.) *Security*. Available at: <https://ico.org.uk/for-organisations/guide-to-data-protection/security/> (Accessed: 21 July 2025).

McKinney, W. (2022) *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter*, 3rd edn. Sebastopol, CA: O'Reilly Media.

NDPA (2023) *Nigeria Data Protection Act*. Available at: <https://placng.org/i/wp-content/uploads/2023/06/Nigeria-Data-Protection-Act-2023.pdf> (Accessed: 21 July 2025).

OWASP (2017) *OWASP REST Security Cheat Sheet*. Available at: [https://cheatsheetseries.owasp.org/cheatsheets/REST\\_Security\\_Cheat\\_Sheet.html](https://cheatsheetseries.owasp.org/cheatsheets/REST_Security_Cheat_Sheet.html) (Accessed: 21 July 2025).

Rolfe, G., Freshwater, D. and Jasper, M. (2001) *Critical reflection for nursing and the helping professions: a user's guide*. Basingstoke: Palgrave Macmillan.