# Harnessing innovation and smarter data processes

Austen Keating, population health analyst, and Nelson Cheng, data analytics engineer

## Leveraging data pipelines and emerging technologies for efficient, reproducible Health Needs Assessments (HNA): the NWMPHN HNA 2025-28 case study.

### Background

Every Primary Health Network is required to complete a comprehensive region-wide Health Needs Assessment (HNA) every three years. These inform regional health planning, priority setting, and program commissioning.

However, developing a methodologically robust and reproducible HNA presents several challenges, including the volume and complexity of health data, the evolving nature of data sources, and the coordination required across teams and tools.

### Goals

At NWMPHN, we asked: **How can we leverage data engineering principles to improve how we deliver our HNA?**

We aimed to address three persistent challenges:

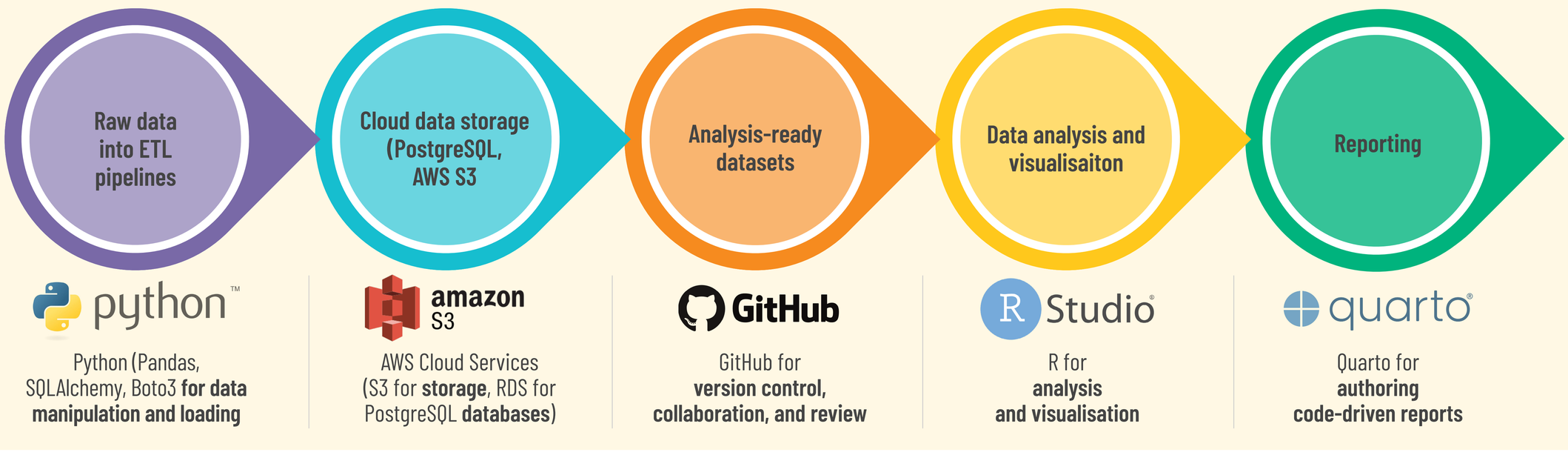
1. **Standardisation**. Align reporting methodologies to reduce variation in data interpretation and application across reports.
2. **Reproducibility**. Facilitate the reuse of code, logic, and datasets to reduce duplication, enhance efficiency, and support continuous improvement.
3. **Scalability**. Lay the groundwork for a flexible system that can grow with our data needs, integrating more sources and delivering insights faster.

### Approach

In 2024, NWMPHN adopted a data-engineering-informed workflow to develop the HNA for 2025–28. This involved using code and reproducible research tools from end to end,from transforming and analysing large datasets (such as the Victorian Admitted Episodes Database and the Emergency Minimum Dataset) to producing descriptive content and visualisations.

The HNA was authored in Quarto, an open-source scientific and technical publishing system, powered by R and Python programming languages. Critical innovations were the use of Extract, Transform, Load (ETL) pipelines to manage data and a version-controlled GitHub repository for collaborative development, version history, and code review.

### Key tools and technologies




### Key wins

- **Improved efficiency**. Datasets became accessible and usable within minutes, instead of hours or days.
- **Reusable workflows**. Processes developed for the HNA now form the basis of other population health and reporting projects.
- **Reproducible outputs**. Each insight and visualisation in the HNA could be regenerated on-demand using code, increasing transparency and consistency.
- **Better collaboration and project management**. Code review and versioning via GitHub meant work was traceable, peer-reviewed, and improved cross-unit communication.

### Lessons learned

- **Upfront investment pays off**. While ETL and infrastructure set-up takes time, the medium-to-long-term gains in speed, accuracy, and scalability are substantial. Further development of data pipelines will increasingly add value to future HNA and other reports.
- **Reusability is a multiplier**. Once data is preprocessed and stored correctly, the same pipelines and datasets can power multiple reports and dashboards.
- **Cultural shift is required**. Moving to code-driven reporting changes how teams collaborate. This requires willingness to upskill and knowledge sharing across teams to improve technical capability.
- **New approach revealed pain points**. Code-based workflows may not align with existing review processes and translating outputs to MS Word complicates version control, feedback and formatting.

### Future directions

We are now expanding our data ecosystem to include additional priority datasets. Next is the Social Health Atlases developed by the Public Health Information Development Unit (PHIDU) at Torrens University. This will further enrich our understanding of health and social indicators at the local level.

Our broader vision is to transition to a full-stack reporting ecosystem, where datasets are:

- continuously updated via automated pipelines
- stored in structured, analysis-ready formats
- accessible to analysts and planners for diverse reporting needs
- reused across multiple projects to increase impact and reduce duplication

### Conclusion

This case-study represents a significant shift in how NWMPHN delivers region-wide HNAs. By embedding data engineering principles and workflows into the development process, we've produced a comprehensive and technically robust report that has laid a scalable foundation to support future work across the organisation.

phn NORTH WESTERN MELBOURNE
An Australian Government Initiative

**For more information**
Contact **hna.admin@nwmphn.org.au**
Or visit **nwmphn.org.au**