

DATA2x01 Group Assignment Report

1. Overview

This report summarizes the analysis of socio-economic indicators for the SA4 region of Inner South West, NSW.

The project integrates geospatial and statistical analysis using PostgreSQL, PostGIS, and Python.

2. Data Sources

- SA2 boundaries: ABS shapefile (2021)
- Business counts: NSW government open data
- Population distribution: Demographic breakdown per SA2
- Income statistics: Average and median incomes per SA2
- Points of Interest (POI): Accessed via NSW ArcGIS REST API

3. Methodology

Data was filtered to SA2s within the Inner South West SA4 region. After loading into PostGIS, spatial joins were

performed to assign POIs to regions. Z-scores were calculated for POI count, mean income, total businesses, and

total population. A weighted sum was used to compute the final score per SA2.

4. Scoring Formula

Final Score = $0.3 * z(\text{POI count}) + 0.3 * z(\text{mean income}) + 0.2 * z(\text{businesses}) + 0.2 * z(\text{population})$

5. Tools Used

DATA2x01 Group Assignment Report

- Python (pandas, geopandas, sqlalchemy, scipy, matplotlib)
- PostgreSQL with PostGIS extension
- NSW ArcGIS API (via HTTP request)
- Jupyter Notebook & script automation

6. Output & Visualisation

- SA2 Scores stored in table 'sa2_scores'
- Choropleth map saved as 'output/score_map.png'

7. Conclusion

This project demonstrates the integration of spatial and socio-economic data for regional analysis. The methodology allows identifying high-opportunity areas within the SA4 region based on data-driven scoring.