

# new\_report

May 18, 2025

## 1 DATA2001 Assignment - Team 3 Report

**Team Members:** Michella Krishna, Yujia Liu, Katherine Kong

**Region Focus:** City and Inner South, North Sydney and Hornsby, Inner West

### 1.1 1. Dataset Description

Several data sources were used to compute accessibility scores across SA2 regions, each selected for their relevance to local resources.

- **Businesses**  
Provided by the [Australian Bureau of Statistics](#). This file contains business counts by industry and turnover range for each SA2. The `total_businesses` field was used to calculate business density per 1,000 people in selected sectors.
- **Income**  
Sourced from the [Canvas](#). Includes variables such as `median_income` by SA2. These fields were used to assess the economic situation of each region.
- **Population**  
Sourced from [Canvas](#), reports population counts by age groups. Fields including `total_people` and youth segments (`10_14_people`) were used to support normalization and school-related metrics.
- **Stops**  
Sourced from [Transport for NSW](#). Includes public transport stops with geographic coordinates and SA2 codes. Stop counts were calculated per region.
- **School Catchments**  
Provided by the [NSW Department of Education](#). Includes shapefiles for primary, secondary, and future school catchment zones. These spatial layers were used to evaluate school accessibility.
- **POI**  
Extracted using the [NSW Points of Interest API](#). Contains locations of community facilities identified within each SA2 bounding box. Relevant POIs were counted to represent local amenity access.

## 1.2 2. Database Description

## 1.3 3. Result Analysis

### 1.3.1 3.1 Formula Rationale

To calculate the final score for each SA2 region, four indicators were chosen: **stops**, **schools**, **parks as POI**, and **businesses (retail industry)**. These factors were selected because they reflect important aspects of **local accessibility**, **education**, **community facilities**, and **economic activity**.

Each metric was converted into a **z-score** to standardize their values and allow fair comparison across different scales. The four z-scores were then **added together**, assuming **equal importance** for each component.

Finally, a **sigmoid function** was applied to the total score to scale it between 0 and 1, making the results easier to interpret and compare across regions.

### 1.3.2 3.2 Analysis for Inner West

**3.2.1 Component Impact on Final Scores** The correlation values show how each metric influences the final score. In Inner West, **stops**, **parks**, and **businesses** have the strongest correlations with the final score, meaning they contribute more to score variation. **Schools** has a weak correlation, suggesting school access is more evenly spread across regions.

### 3.2.2 Distribution of Scores

<br>  
<p style="max-width: 300px; margin: 8px auto;">

The distribution of final scores in Inner West appears relatively even, with no extreme clusters.  
</p>

<br>  
<p style="max-width: 300px; margin: 8px auto;">

The map visual shows how final resource scores are distributed across SA2 regions in Inner West.  
</p>

**3.2.3 Notable Trends or Regions** The table shows the top three and bottom three SA2 regions by final score. These highlight the range of resource access across Inner West, with noticeable gaps between the highest and lowest scoring areas. This reflects particular scores of interest and uneven distribution worth noting in the analysis.

## 1.4 4. Correlation Analysis (2 marks)

### 1.4.1 4.1

### 1.4.2 4.2

### 1.4.3 4.3 Inner West

To explore the relationship between final scores and income, a basic correlation check was performed between each SA2 region's median income and its computed score. The result showed no

meaningful linear correlation (correlation coefficient = 0), suggesting that the score constructed in this assignment may not reflect the economic situation of Inner West.

#### 1.4.4 Limitations

The result highlights a key limitation in how the final score was calculated. Although stops, schools, parks, and retail businesses are relevant indicators, they may not fully capture the economic or social dynamics of each SA2 region. Other factors — such as housing affordability, health services, or population density — were not included in the score and may play a significant role in shaping access to resources.

#### 1.5 5. Conclusion

Key findings and what they suggest.

#### 1.6 6. References

- Australian Bureau of Statistics – Counts of Australian Businesses  
<https://www.abs.gov.au/statistics/economy/business-indicators/counts-australian-businesses-including-entries-and-exits/latest-release#data-downloads>
- NSW Department of Education – School Catchment Areas  
<https://data.cese.nsw.gov.au/data/dataset/school-intake-zones-catchment-areas-for-nsw-gov>
- NSW Points of Interest API  
<https://data.nsw.gov.au/data/dataset/nsw-points-of-interest-poi/resource/f4b643b4-a872-4b63-bbb5-5b768bacbc78>
- Transport for NSW – GTFS Timetable Data  
<https://opendata.transport.nsw.gov.au/dataset/timetables-complete-gtfs>

#### 1.7 7. Appendices

- Supplementary tables or extended visuals (if needed).