**Step to School**

**1. Introduction**

This project set out to answer a practical question: *How easy is it for students in the ACT to walk to a bus stop that serves their school?* To do this, we built a pipeline that combines open transport data, census geographies (SA1s), and pedestrian networks from OpenStreetMap. The analysis measures walk times from each residential area to the closest bus stop serving each school, and then computes coverage metrics. I also tested “quick-win” scenarios where new stops could be added to fill coverage gaps.

The analysis is designed to give both a broad overview (across the entire ACT) and a fine-grained view (at the SA1 or school level).

**2. Methodology**

The work proceeded in stages, each automated by scripts in the pipeline.

**Step 1: Data Collection**

* **Student origins:** The ABS Statistical Area Level 1 (SA1) centroids were used as proxies for residential locations.
* **School bus stops:** Data on special school services and bus stops was downloaded from open government feeds.
* **School assignments:** Stops were matched to schools using fuzzy text matching on route names, headsigns, and stop descriptions.
* **Pedestrian network:** A full walking network for the ACT was built from OpenStreetMap, with travel times estimated from edge lengths and an assumed walking speed.

**Step 2: Walk Time Calculation**

For each SA1–school pairing, I computed the shortest walk time from the SA1 centroid to the nearest bus stop serving that school. This produced a large table of *SA1 × school × walk time (sec)*.

**Step 3: KPI Computation**

Coverage key performance indicators (KPIs) were calculated in two ways:

* **By SA1:** What percentage of that SA1’s possible school pairings are reachable within 10 minutes? Within 15 minutes?
* **By School:** What percentage of SA1s can access a stop for this school within 10 or 15 minutes?

This ensures we can see both the *local coverage situation* and the *equity across schools*.

**Step 4: Quick-Wins Heuristic**

To identify potential new stops, I ran a simple greedy clustering algorithm:

* Look for SA1s with poor coverage (<100% within 10 minutes).
* Iteratively place a hypothetical stop at the centroid of the densest “uncovered” cluster.
* Stop after placing a set number of candidates or once coverage is improved.

This suggests where new stops could make the biggest difference.

**Step 5: Outputs for Analysis**

The results were flattened into CSVs suitable for Power BI dashboards:

* **SA1 × school walk times**
* **SA1 KPIs**
* **Stops and candidate stops**

This allows interactive visualisation and mapping.

**3. Results**

**Coverage at the SA1 Level**

The majority of SA1s have **no coverage within 10 minutes** of walking. Out of 624 SA1s examined:

* **34 SA1s (5%)** have *full coverage* (all their school pairings ≤10 minutes).
* **23 SA1s (4%)** have *partial coverage* (some schools within 10 minutes, others not).
* **567 SA1s (91%)** have *no coverage* within 10 minutes.

At a more lenient 15-minute threshold, the picture improves slightly:

* **76 SA1s (12%)** have full coverage.
* **54 SA1s (9%)** have partial coverage.
* **494 SA1s (79%)** still have no coverage.

**Visualisation – Distribution of Coverage**  
Below is a histogram showing how SA1s are distributed by percentage of school pairings covered within 10 and 15 minutes.

This shows a heavy skew toward *zero coverage*, meaning that in most SA1s, students would have to walk more than 15 minutes to reach a stop for their school.

**Candidate Stop Locations**

The heuristic for “quick wins” proposed new candidate stop locations. These are concentrated in areas where clusters of SA1s lacked coverage. For example, candidate stops were suggested in:

* Northern suburbs with long gaps between school service stops.
* Outer southern suburbs where existing stops are sparse.
* Pockets in central areas where walking catchments don’t overlap.

These proposed stops are not final recommendations, but highlight where *marginal improvements could yield the most gain in coverage*.

**Stops Dataset**

The existing school bus stops are clustered around certain corridors. Many stops serve multiple schools, but some schools are poorly represented, leaving large residential catchments underserved. Candidate stops aim to “fill in the holes” in this network.

**4. Interpretation**

The results are stark:

* Within **10 minutes of walking**, the vast majority of students have no access to a bus stop serving their school.
* Even within **15 minutes**, nearly 80% of SA1s remain without coverage.
* This suggests that the current school bus network is heavily centralised, with many neighbourhoods left out.

**What This Means**

* **Equity concerns:** Students in most suburbs face a barrier to bus access, potentially leading to car dependency or inequitable transport choices.
* **Policy relevance:** If the target is that “all students should be within a 10–15 minute walk of a bus to their school,” current provision falls far short.
* **Quick wins:** Adding stops in identified gap areas could raise coverage noticeably, especially for schools with broad catchments.

**6. Conclusion**

This analysis provides a baseline picture of school bus accessibility in the ACT. The key takeaway is that most students face a walk longer than 15 minutes to reach a bus stop serving their school. Candidate stop placement can improve matters, but even then, systemic redesign may be needed if the policy goal is equitable access.

The methodology is robust—built on open datasets, reproducible code, and a clear KPI framework—and can be extended to test different scenarios, such as:

* What happens if routes change?
* How would shifting stop locations impact accessibility?
* Can we weight the analysis by student numbers for demand-focused planning?

Ultimately, this gives planners a data-driven tool for evaluating where the network is working, where it is not, and what can be done about it.