Spatial Analysis & Data Insights (SANDI) - Enriching mobile network data using Census, NTS and AI techniques

Tom van Vuren

September 2025

Amey

Life's better connected

With SANDI we enrich anonymised travel movement data obtained from mobile network providers, and provide the ability to visualise and analyse enriched transport data in an intuitive and interactive way.

Amey

Why mobile network data?

Derived from phone masts, not phones
Everyone connected unless actively opted out
Large sample and unbiased

But…

Density of masts determines spatial granularity (particularly rural)
Anonymity requirements limit reported values per movement to 5+
Short distance trips underreported (eg walk and cycle)

And…

Purposes limited to HBW, HBO, NHBW, NHBO
Modes limited to Road, Rail, Active, Freight
No socio-demographic information on the traveller (apart from gender and resident/visitor)

Amey

Understanding the flow of people:

MND data is ubiquitous and nowadays relatively cheap to procure, it remains limited in its representation of who is travelling, how and why.

Informed decision making:

More detailed insights are increasingly valuable to enable project and policy decisions for example:

- Where infrastructure investment may be most beneficial and to whom
- Which trips may be shifted to more sustainable modes
- Gaps in the public transport or active mode network, where demand exists but no services
- Winners and losers and alignment with wider objectives

Digitalised data:

SANDI enriches the raw data, aims to speed up access, making the analysis faster and consistent, and reduces error.

User-focused:

We have built a new data interface that is more intuitive and interactive, enabling comparisons in almost real-time.

How do we do it?

Amey

- SANDI enhances People Movement Data (mobile network data) with data from the Census, National Travel Survey and Office for National Statistics.
- We enrich anonymised movement data with the most likely additional spatial information, on how, why, when and where people move, to visualise and analyse their observed travel patterns.
- The user can visualise people's routes and broader statistics for different modes of transport based on the available transport network, also allowing *what-if* analyses around network changes.
- The tool estimates the socio-demographic characteristics of the people making these trips based on the area's characteristics obtained from the Census.
- Data from either Amey's or the user's databases, national or open sources could be added or merged, e.g. from users' own fleets, or from local transport models, to provide additional insights.

The real techy detail...

- We use multiple open-source solutions, such as Open Street Map and the GraphHopper routing engine.
- We use Artificial Intelligence techniques such as clustering for data enrichment.
- We use cloud computing integrated with Amey's DataLake for specific Databricks Cluster requirements, ensuring optimal performance and scalability for our data processing needs.

Using Statistical Tools and AI, we:

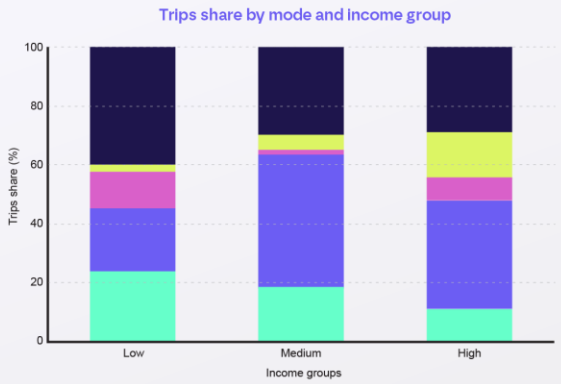
- Created a Synthetic Population with Census and NTS data for the whole UK with a heuristic algorithm for iterative generation.
- Generated a Synthetic Trip Population, adding information on travel behaviour from NTS with Cross-Classification.
- Subsequently enriched MND data using Artificial Intelligence (i.e. Random Forest Classification Technique).

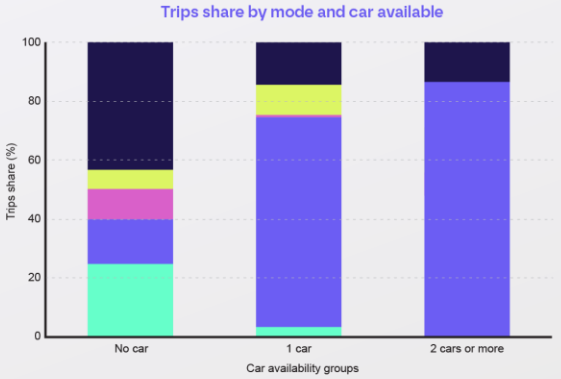
Example visuals

Insights into socio-demographic characteristics

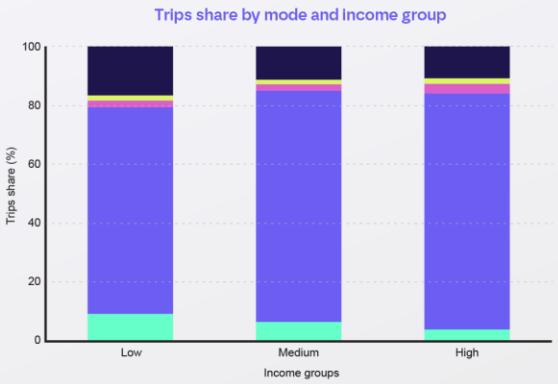
Amey

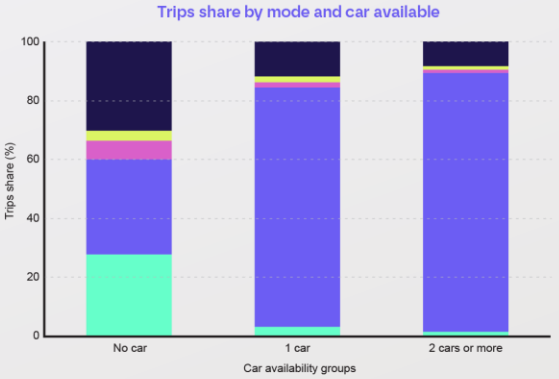
Trips to a Central Zone in Cardiff





Full population





By inferring the socio-demographic characteristics of the traveller, it is possible to assess who is travelling, and hence the winners and losers of potential interventions.

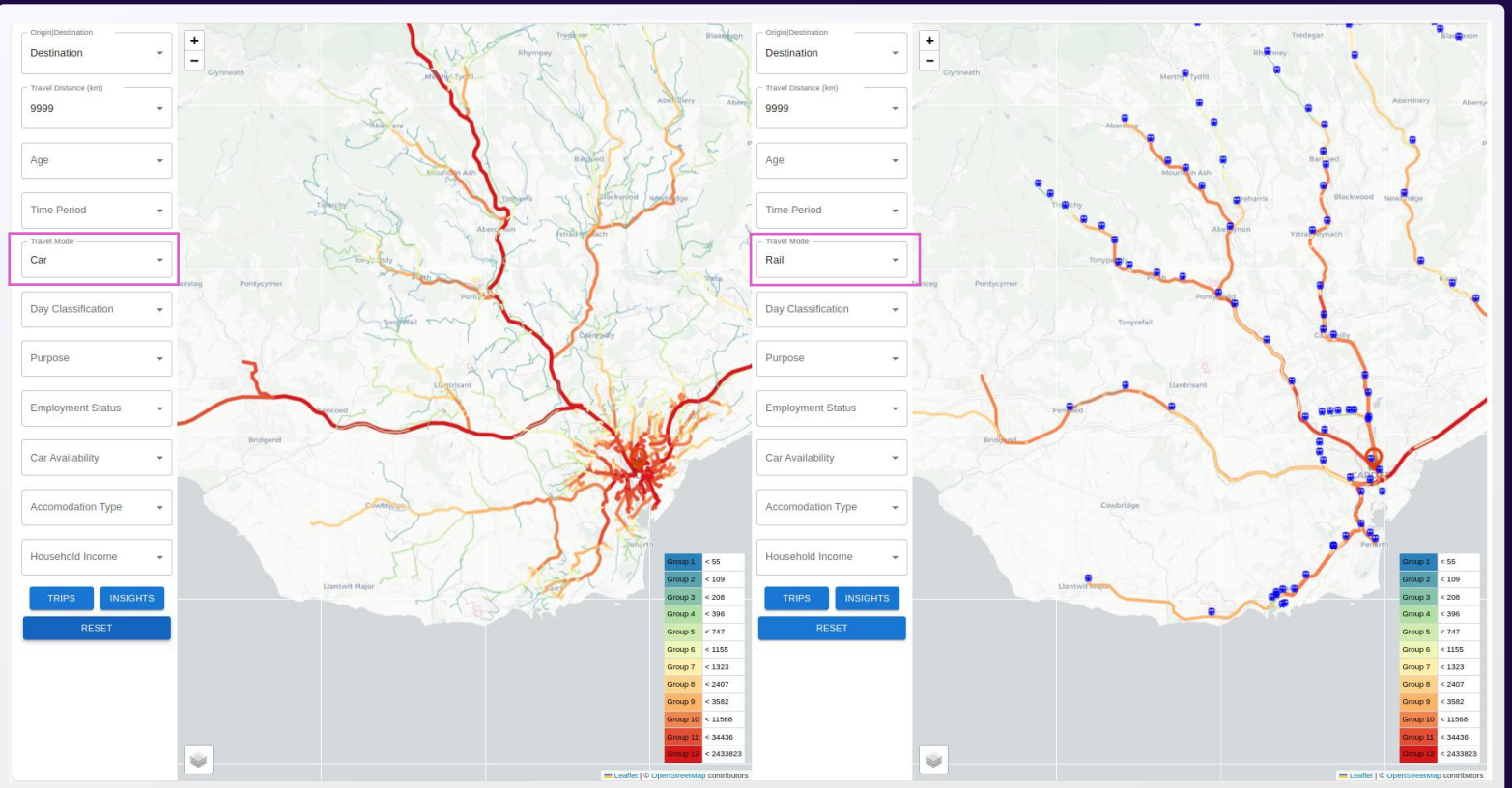
Client: Local authority or central government.

Challenge: A lack of understanding about the demographic characteristics of people making trips.

Approach: Analysis of enriched MND data to establish socio-demographic characteristics of travellers.

Outcome: Granular insights into the impact of income and car availability on mode choice to a city centre zone, also compared with general population characteristics.

Impact of a temporary rail closure



On the left, SANDI shows corridors such as the M4 and A470, as well as short trips within the city centre, where there is a high volume of trips made by road.

On the right, rail patronage is particularly high from Southwest and London, but also from the valleys. The data highlights there was a rail line down due to maintenance between Cardiff and Merthyr Tydfil which could be contributing to the high volume of cars along the A470.

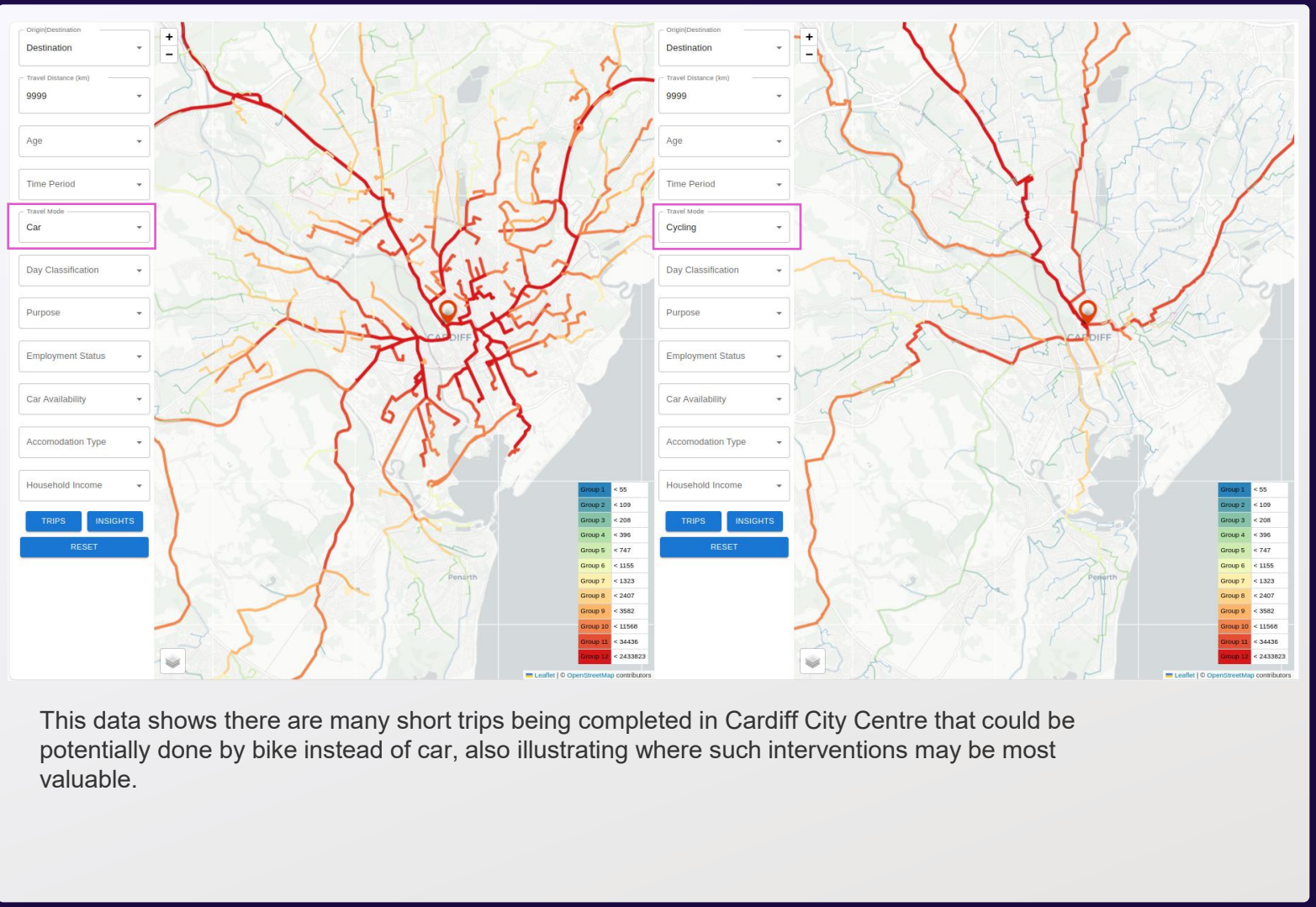
Client: Local authority or rail operator.

Challenge: Temporary rail line closure and unknown impact on travel patterns.

Approach: Explore the impact of a temporary rail closure on the road network.

Outcome: Identified increase in road traffic and consider ways to better manage additional road traffic during times of rail disruption.

Opportunities for modal shift



This data shows there are many short trips being completed in Cardiff City Centre that could be potentially done by bike instead of car, also illustrating where such interventions may be most valuable.

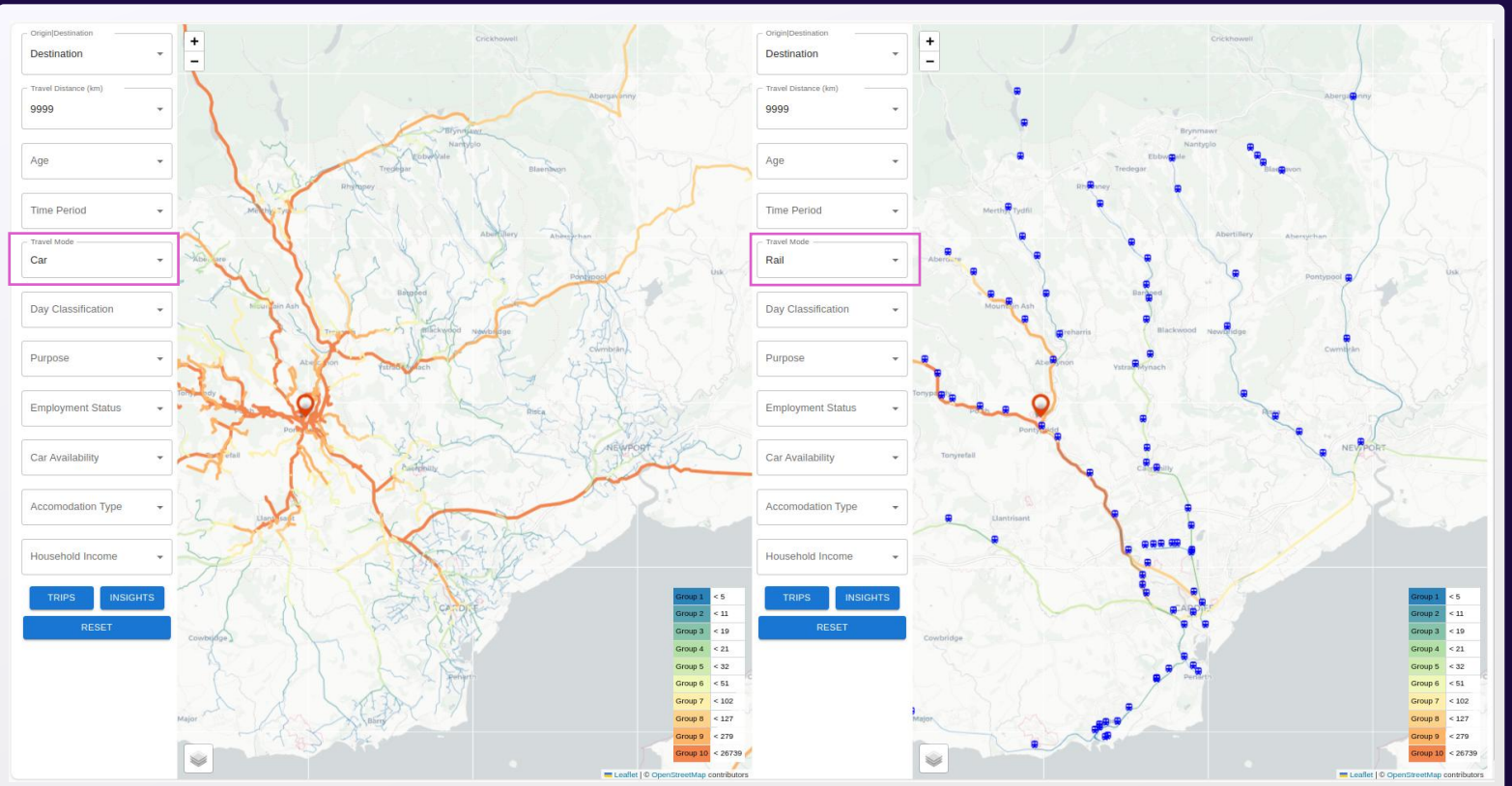
Client: Local authority or bike share operator.

Challenge: There are too many car trips being made in the city centre.

Approach: Compare the car trips and active travel trips that are being made – where they are, by who and for what purpose.

Outcome: Identify opportunities for physical or behavioural interventions to encourage modal shift to active travel.

Opportunity to improve rural connectivity



This demonstrates that many local journeys and journeys across the valleys are completed by car rather than rail. This is likely due to the limitations of the geography and the rail network. Also, there is a high volume of cars that are travelling to the Southwest and London which could be done by rail but would require changing in Cardiff which is likely a deterrent for people to choose public transport.

Client: Local authority or transport operator.

Challenge: High car use from rural areas into the city centre.

Approach: Explore alternative means of transport for people to use to travel into the city centre and the patronage at specific times (during the month/day).

Outcome: Establish there is poor rail connectivity and there is an opportunity for an improved bus service to connect rural journeys to larger transport hubs or stations.

Insights into socio-demographic characteristics

These graphs provide further information on the socio-demographic characteristics of the users of each mode.

Client: Local authority or transport operator.

Challenge: A lack of understanding about the demographic characteristics of people using public transport.

Approach: Analysis of enriched MND data to establish socio-demographic characteristics of travellers.

Outcome: Granular insights into the impact of employment and income on the use of bus and rail, such as the importance of bus for the retired population.

Demo (for a copy of the video, contact Tom van Vuren, Tom.van-Vuren@amey.co.uk)

Amey

Next steps

Further enrichment

Merge observed short distance (active mode) data with synthetic data
Strengthen active modes / analyses
Import more data for comparative analysis

Applications

If you have your own MND dataset….
If you don’t hold mobile network data….

Amey

Want to know more?

Tom.Van-Vuren@Amey.co.uk
Francesco.Manca@Amey.co.uk
Suren.Samarasinghe@Amey.co.uk