



Active
Travel
England

Data Science for Active Travel Planning

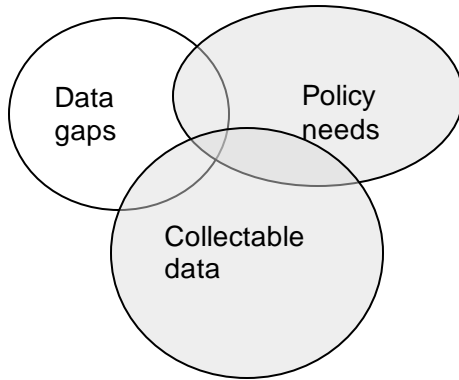
Streets, schemes and networks

Robin Lovelace

Head of Data and Digital, Active Travel England



What we do in Active Travel England



Infrastructure data

- Existing or (e.g. OSM/OS) **planned** infrastructure
- Geometries (space)
- Lifecycle (time)
 - Completion date
 - Planned infrastructure
 - Dates
- Attribute data
 - Infrastructure type

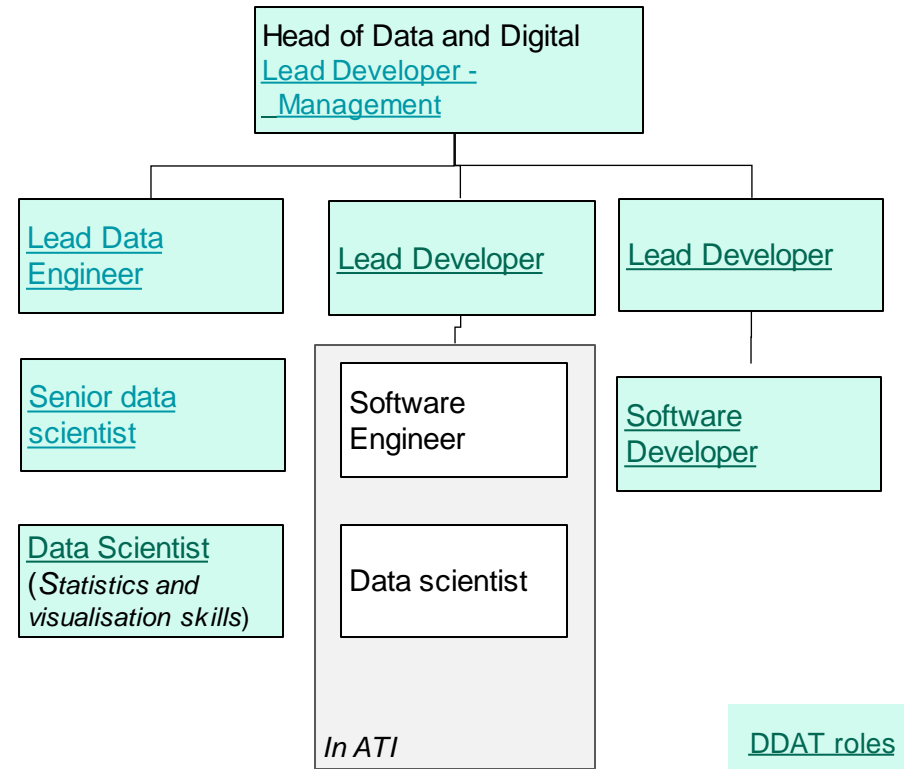
Behaviour data

- Travel survey data
- Preferences
- Movement data
 - GPS data
 - Point/area counts

Outcomes

- Environmental
- Health
- Economic
- ...

Data and Digital (7+2 FTE)



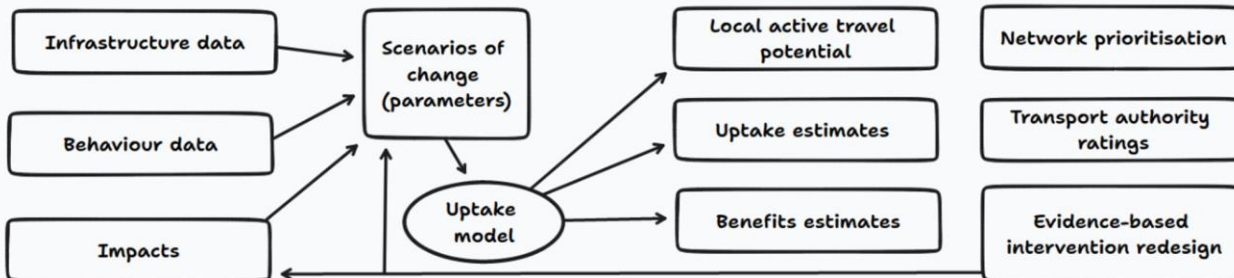
DDAT roles

Input datasets

Policy inputs and modelling

Model and data outputs

Policy outcomes

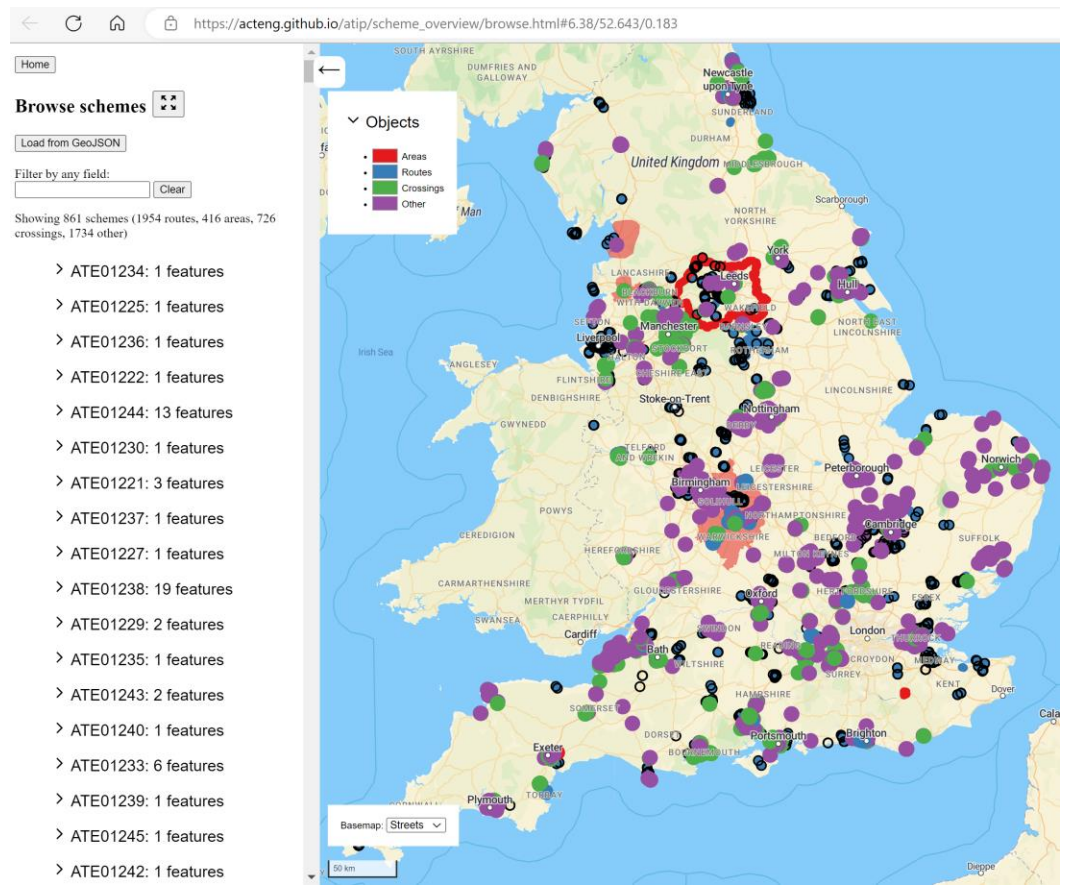


Stakeholders

- Analysis team
- ATE
- Government
- Public

Example project: ATIP

- We now know where schemes are
- Inspectors have a baseline for remote and on site inspections
- Reduced friction between different parts of the transport planning process
- We can do additional analysis, e.g. to identify all schemes in proximity to schools
- **Live Demo!**



"The inspections team at Active Travel England used ATIP to quickly determine location, context and whether the correct approach was being taken with the proposed scheme assets. Furthermore, ATIP gave the team a chance to assess network cohesion across multiple government programmes and compliance with LCWIP (Local Cycling and Walking Infrastructure Plan) priorities."

Brian Deegan, Director of Inspections, Active Travel England

Safe routes to hospitals?

Latent demand

Choose File fastest.geojson

- ☒ Route network
- ☐ Endcaps for routes
- ☐ Origins
- ☐ Destinations

- LTS 1 - suitable for children
- LTS 2 - low stress
- LTS 3 - medium stress
- LTS 4 - high stress

Route segment counts from 1.00 to 2100.00

2335 origins, 1 destinations

Override max for line width styling:

1000

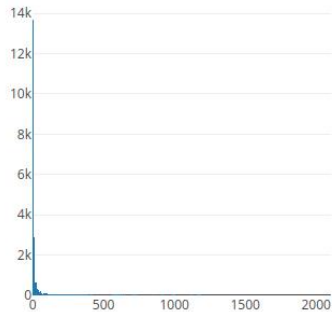
Change origin point size:

3

Change destination point size:

3

Edge counts



Note: LTS model from [BikeOttawa](#)



'Fastest' routes to Alder Hey Hospital in Liverpool based on OD data. Source: <https://dabreegster.github.io/routing-engines/>

Route choices and tradeoffs

Latent demand

Choose File quietest.geojson

- ☒ Route network
- ☐ Endcaps for routes
- ☐ Origins
- ☐ Destinations

- LTS 1 - suitable for children
- LTS 2 - low stress
- LTS 3 - medium stress
- LTS 4 - high stress

Route segment counts from 1.00 to 3095.00

2334 origins, 1 destinations

Override max for line width styling:

1000

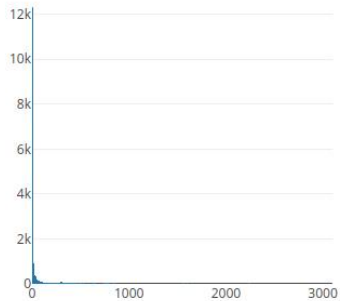
Change origin point size:

3

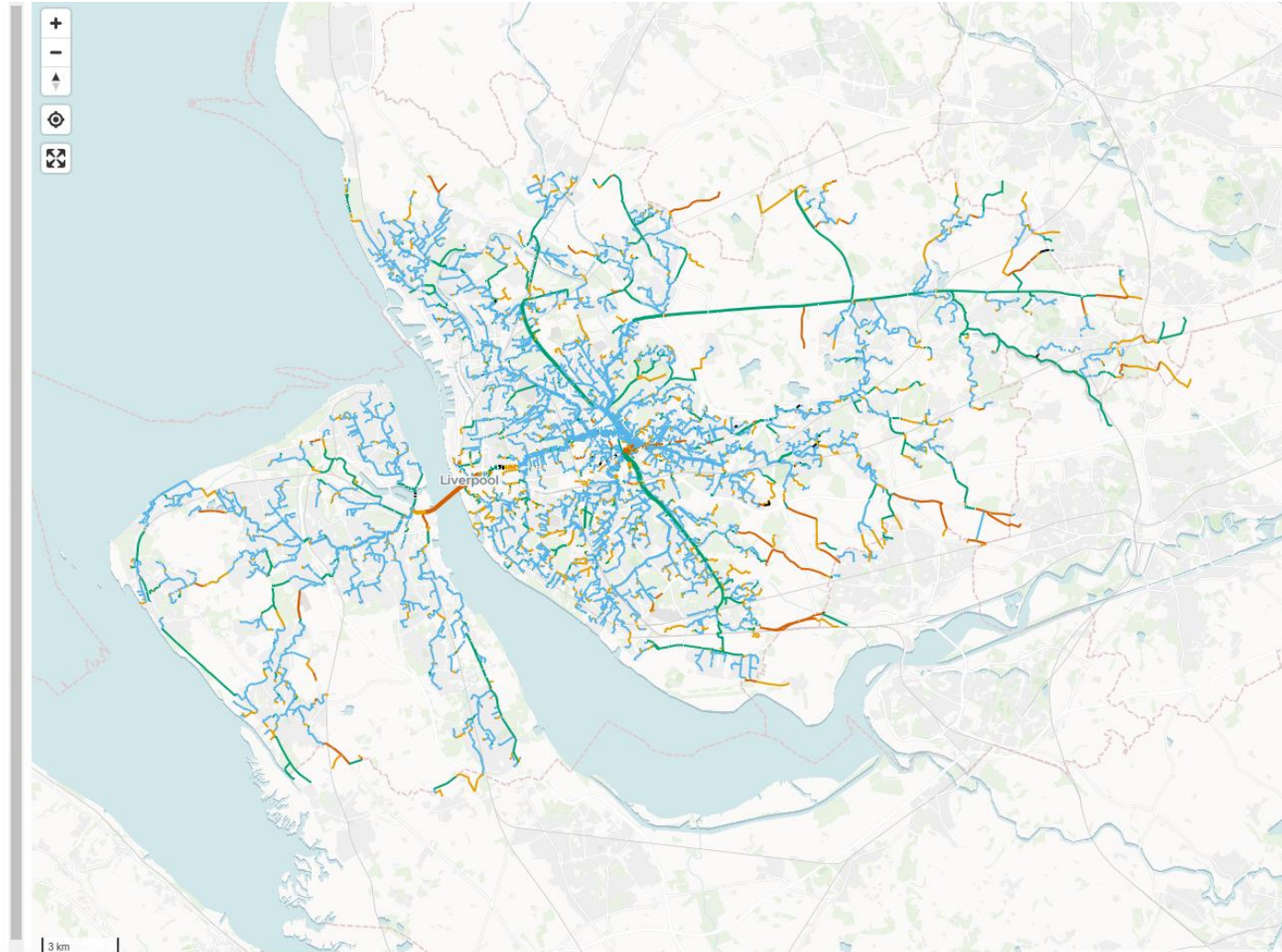
Change destination point size:

3

Edge counts

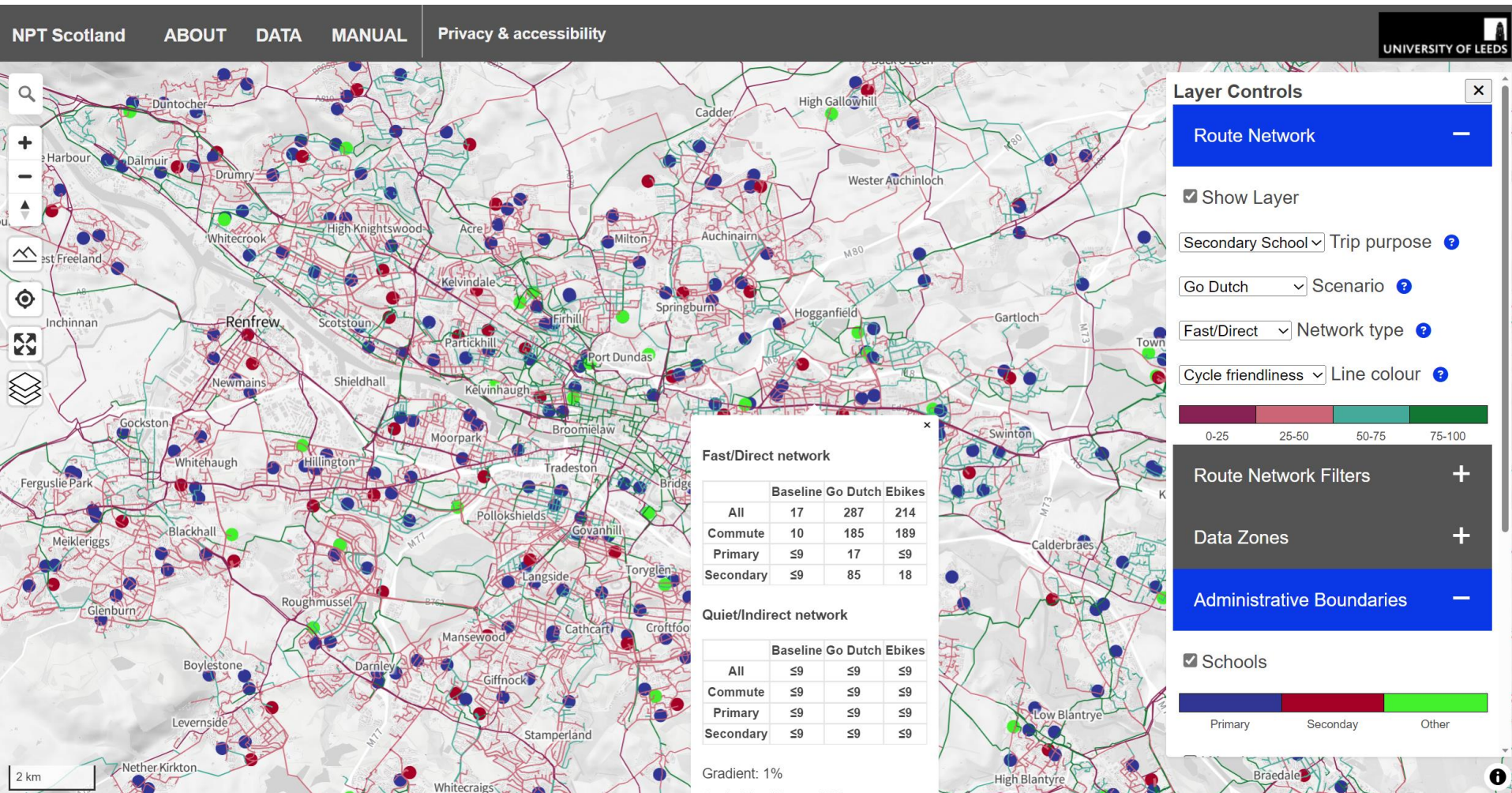


Note: LTS model from [BikeOttawa](#)



'Quietest' routes to Alder Hey Hospital in Liverpool based on OD data. Source: <https://dabreegster.github.io/routing-engines/>

Example project: SchoolRoutes



Network planning tool: <https://nptscot.github.io/>

Elements of Transport Data Science

- **Areal units:** transport patterns can be understood with reference to zonal aggregates, such as the main mode of travel (by car, bike or foot, for example), and average distance of trips made by people living in a particular zone, covered in Section 13.3
- **Desire lines:** straight lines that represent 'origin-destination' data that records how many people travel (or could travel) between places (points or zones) in geographic space, the topic of Section 13.4
- **Nodes:** these are points in the transport system that can represent common origins and destinations and public transport stations such as bus stops and rail stations, the topic of Section 13.5
- **Routes:** these are lines representing a path along the route network along the desire lines and between nodes. Routes (which can be represented as single linestrings or multiple short *segments*) and the *routing engines* that generate them, are covered in Section 13.6
- **Route networks:** these represent the system of roads, paths and other linear features in an area and are covered in Section 13.7. They can be represented as geographic features (typically short segments of road that add up to create a full network) or structured as an interconnected graph, with the level of traffic on different segments referred to as 'flow' by transport modelers (Hollander 2016)

Another key element: Agents

Source: Geocomputation with R

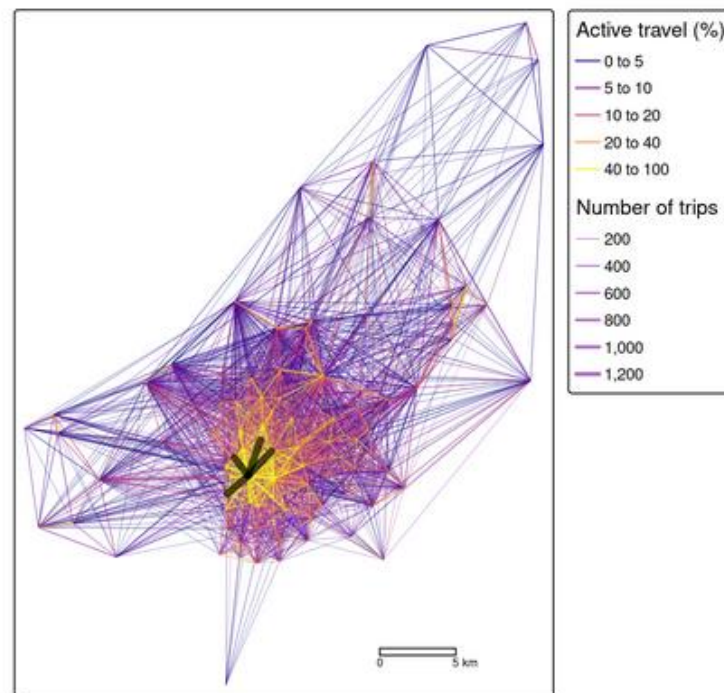


FIGURE 13.3: Desire lines representing trip patterns in Bristol, with width representing number of trips and color representing the percentage of trips made by active modes (walking and cycling). The four black lines represent the interzonal OD pairs in Table 7.1.

Origin-destination (OD) data

```
od_top5 = bristol_od |>  
  slice_max(all, n = 5)
```



o	d	all	bicycle	foot	car_driver	train
E02003043	E02003043	1493	66	1296	64	8
E02003047	E02003043	1300	287	751	148	8
E02003031	E02003043	1221	305	600	176	7
E02003037	E02003043	1186	88	908	110	3
E02003034	E02003043	1177	281	711	100	7


TABLE 13.1: Sample of the top 5 origin-destination pairs in the Bristol OD data frame, representing travel desire lines between zones in the study area.

Source: Geocomputation with R: [Transportation](#)

Routing

3 types of routing from data science environment: 1) in memory, 2) locally hosted routing engine (e.g. OSRM), 3) remote services (e.g. Google Directions)

Locally hosted routing engines include OpenTripPlanner, [Valhalla](#), and R5 (which are multi-modal), and the OpenStreetMap Routing Machine (OSRM) (which is 'uni-modal'). These can be accessed from R with the packages `opentripplanner`, `valhalla`, `r5r` and `osrm` ([Morgan et al. 2019](#); [Pereira et al. 2021](#)). Locally hosted routing engines run on the user's computer but in a process separate from R. They benefit from speed of execution and control over the weighting profile for different modes of transport. Disadvantages include the difficulty of representing complex networks locally; temporal dynamics (primarily due to traffic); and the need for specialized external software.

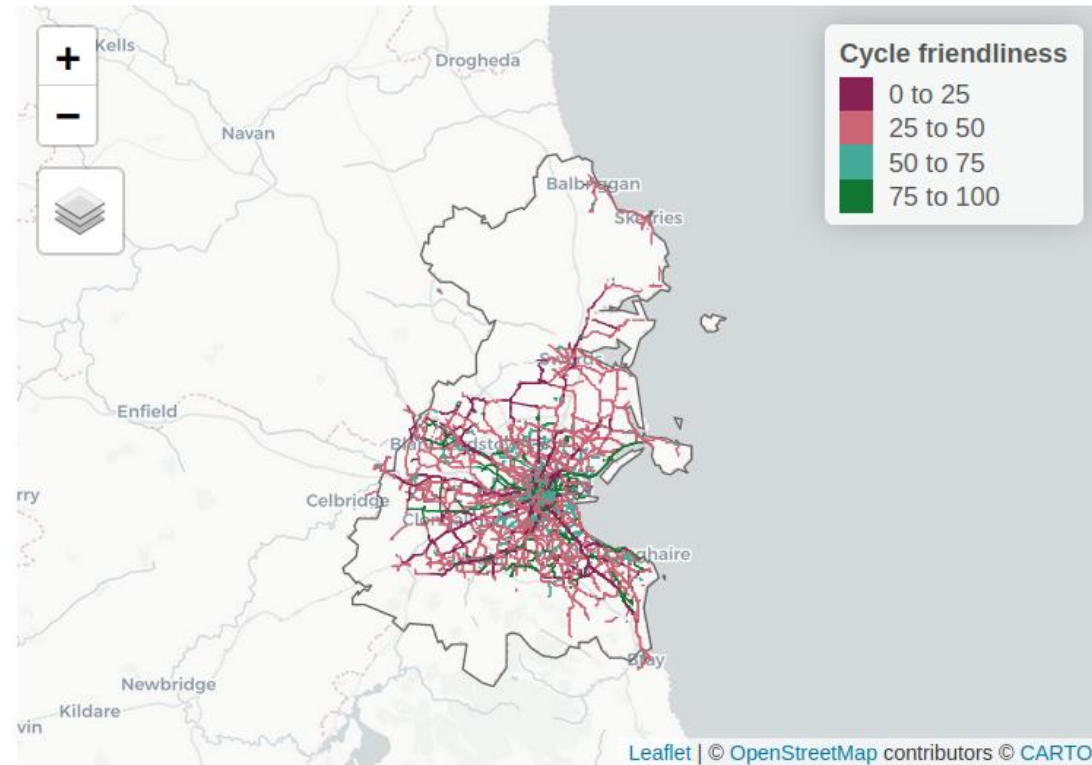
```
routes_short = route(l = desire_lines_short, route_fun = route_osrm,   
                    osrm.profile = "bike")
```

Route network analysis

- What are the key summary statistics we want to know about a route network?
- A question that data science can help answer!
- Network length
- Average flow per link
- Average flow-weighted 'quietness' per link
- Network density
- Network density compared with all available segments
- Distribution of flows
- Spatial inequalities
- Comparisons with other spatially-variable statistics (e.g. population density)
- ...
- Source: open CRUSE tool: <https://cruse.bike/dublin/route-types>

Fastest network

On the fastest network, 29% of the distance cycled occurs in non-hostile segments and 13% in cycle-friendly segments.



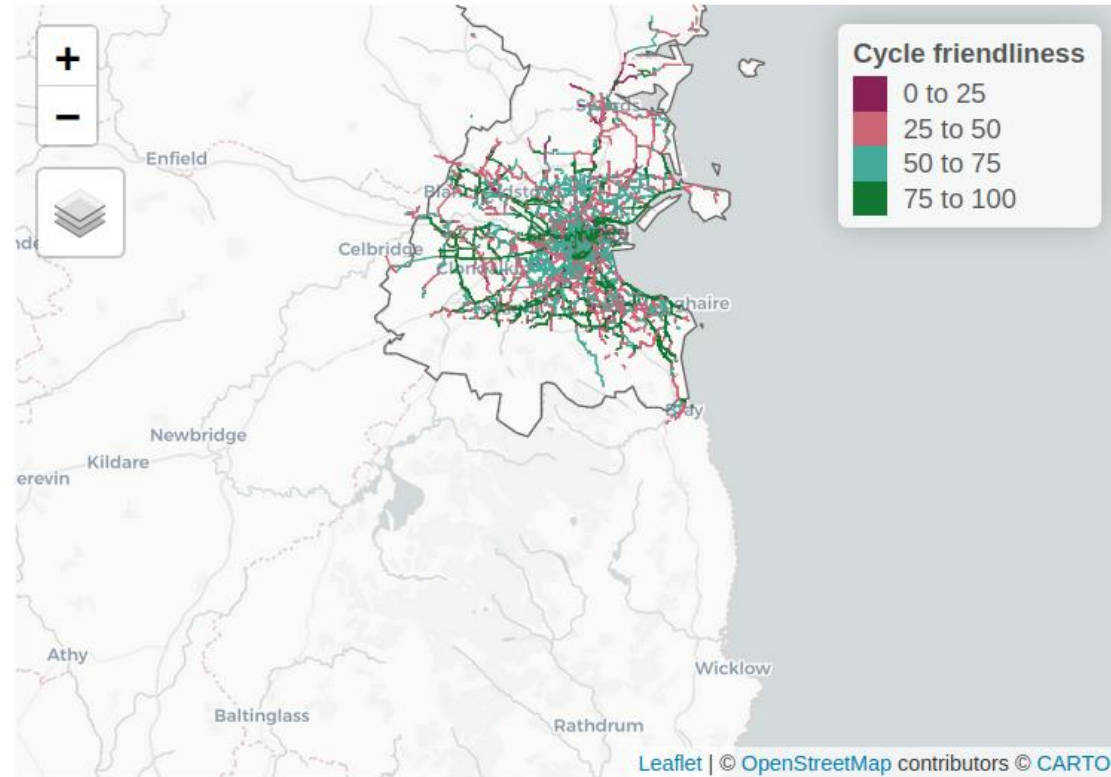
Click [here](#) to see the map full screen.

Route network analysis

- What can change?

Quietest network

Under the baseline scenario, 69% of the distance cycled on the quietest network occurs in non-hostile segments and 32% in cycle-friendly segments.



Click [here](#) to see the map full screen.

On foot

By bike

Active travel planning: details matter

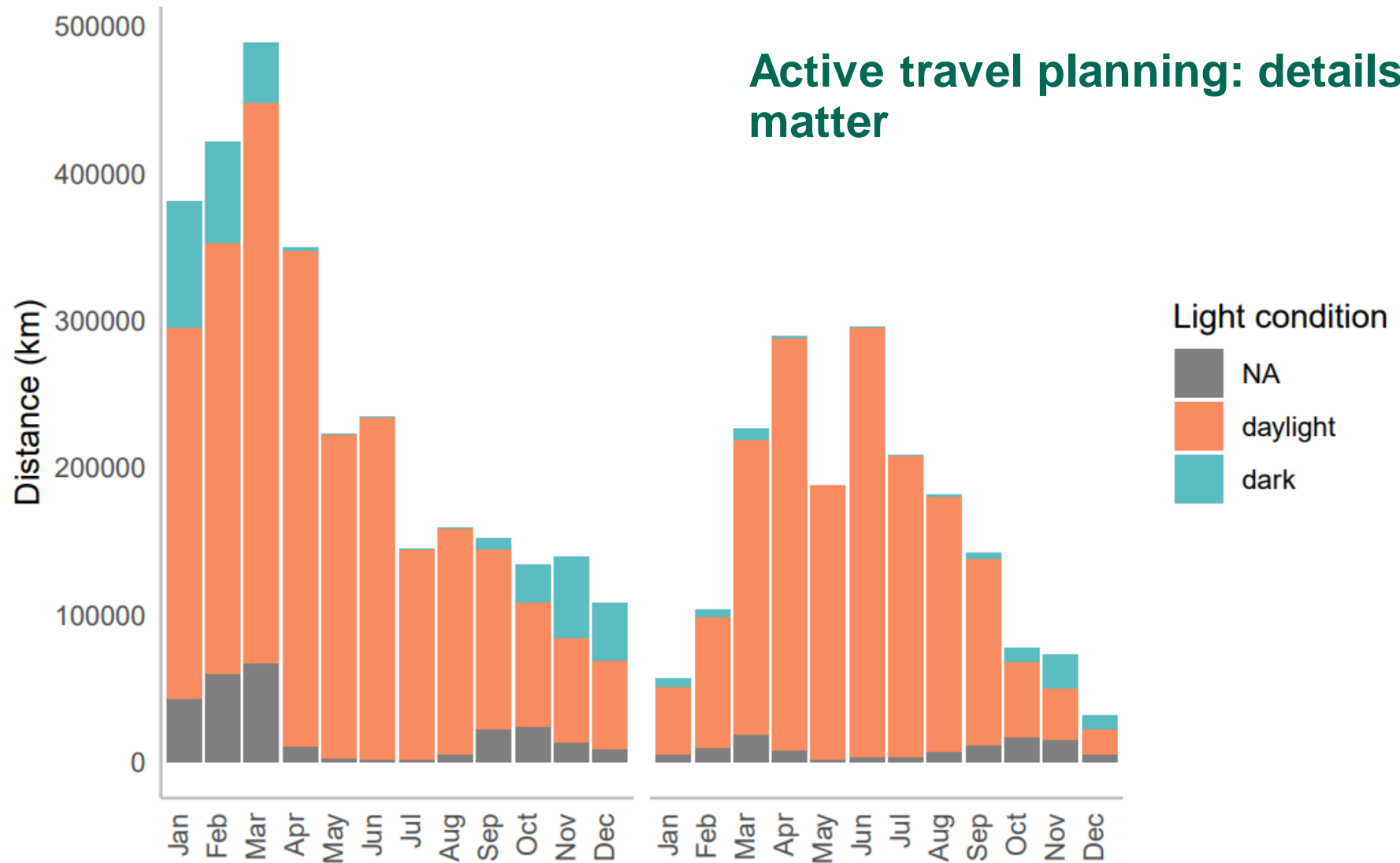


Figure 2: Distance (km) travelled by Strava users per month and light condition, City of Glasgow

Research by Eugeni Vidal, SATURN project. Source: <https://github.com/saturnlighting>

Table 4: Standardized estimates and standard errors of associations between crime, night lighting, their interaction, other relevant spatial factors and distance (km) on foot and by bike in the daylight and dark.

	Daylight		Dark	
	km on foot	km by bike	km on foot	km by bike
Crime	-0.36 ***	-0.36 ***	-0.35 ***	-0.57 ***
	-0.08	-0.08	-0.08	-0.11
Light	0.52 ***	0.60 ***	0.55 ***	0.71 ***
	-0.1	-0.11	-0.1	-0.14
Crime*Light	0.23 *	0.21 *	0.28 **	0.21
	-0.09	-0.1	-0.1	-0.13
Income	1.02 ***	0.56 ***	1.15 ***	0.73 ***
	-0.08	-0.08	-0.08	-0.1
Access	0.28 ***	0.04	0.24 ***	0.15
	-0.06	-0.07	-0.06	-0.09
Density	-0.07	-0.18 **	-0.09	-0.07
	-0.08	-0.07	-0.08	-0.1
Recreational	-0.01	-0.23 **	-0.12	-0.27 **
	-0.13	-0.07	-0.08	-0.09
Commercial	-0.15 *	-0.09	-0.1	-0.12
	-0.08	-0.08	-0.08	-0.11
Industrial	-0.13	-0.02	-0.24 ***	-0.16
	-0.07	-0.08	-0.06	-0.09
Infrastructure	0.92 ***	0.83 ***	0.96 ***	0.95 ***
	-0.12	-0.12	-0.13	-0.16
Quietness	0.43 ***	0.44 ***	-0.03	0.25 **
	-0.06	-0.07	-0.07	-0.09
Gradient	-0.35 ***	-0.41 ***	-0.36 ***	-0.47 ***
	-0.06	-0.06	-0.07	-0.09
N	746	746	746	746
sigma	1.44	1.53	1.51	2.02
logLik	-5836.8	-5606.53	-4341.55	-2449.17
AIC	11701.6	11241.07	8711.1	4926.33
BIC	11766.2	11305.67	8775.7	4990.94

*** p < 0.001; ** p < 0.01; * p < 0.05.

What else can transport data science + geocomputation do?

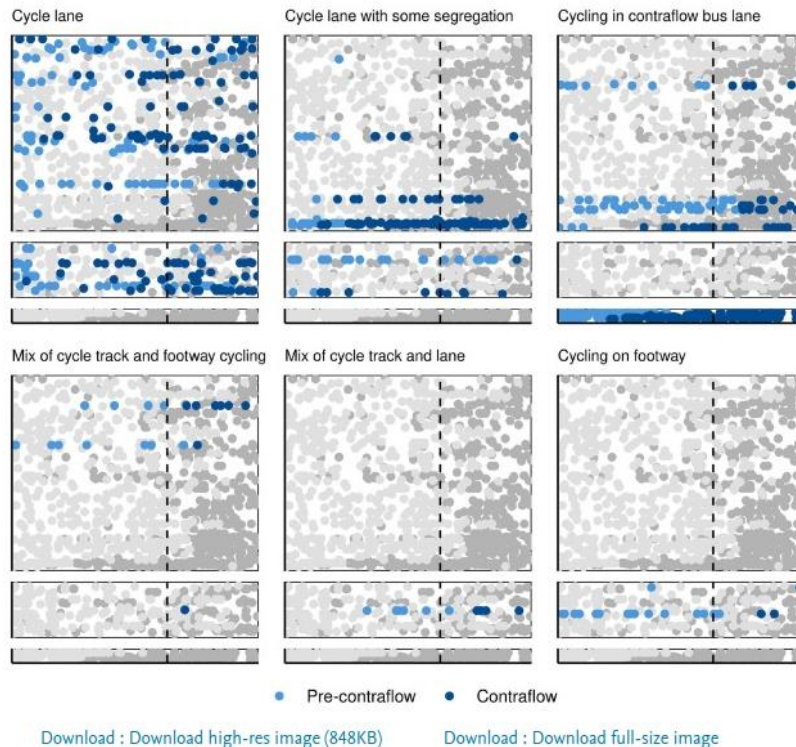
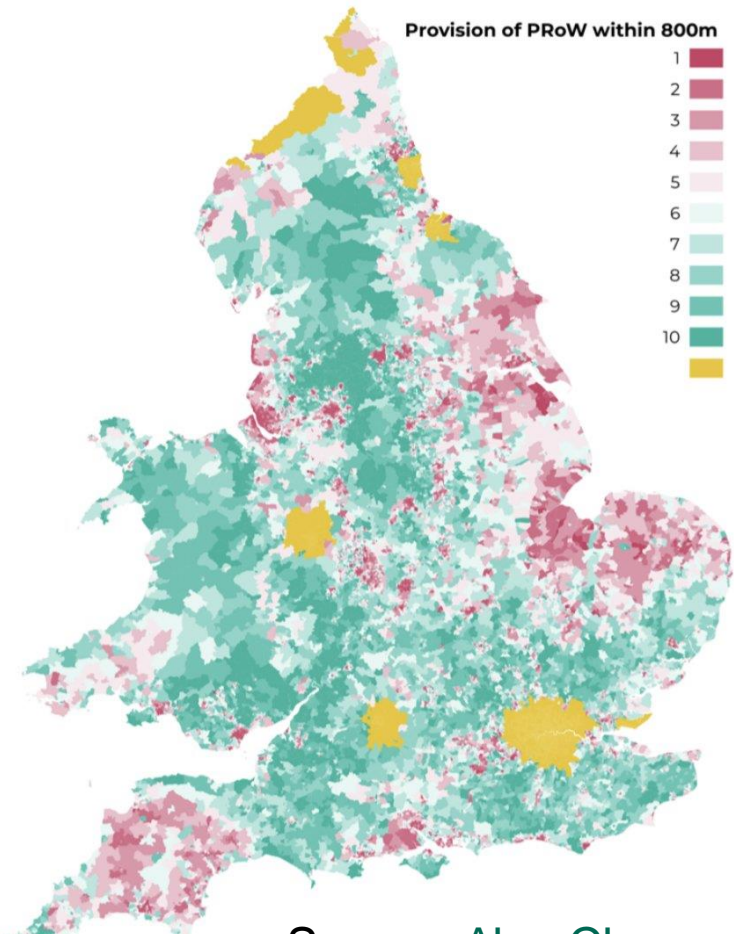


Fig. B1. Dot visualisation of all crashes involving pedal cycles within 10 m of a road segment by unique road segment (vertical position); date of crash (horizontal position); crash segment status (colour); and significant change to road segments (pane). Top visualisation represents all crashes. Lower visualisations highlight crashes by additional cycling infrastructure mentioned in Traffic Regulation Order. The seven contraflow removed crashes have been omitted to aid visualisation.

Source: Caroline Tait ([2023](#)) Accident Analysis & Prevention

MAP 1: LENGTH OF PUBLIC RIGHT OF WAY WITHIN 800M OF A POSTCODE, GROUPED INTO LOWER LAYER SUPER OUTPUT AREA (LSOA) DECILES. AREAS WITH MISSING DATA ARE SHOWN IN YELLOW, WHERE DECILE 1 (PINK) REPRESENTS THE LOWEST LEVELS OF PROW PROVISION.



Source: [Alex Chapman](#)

Source: Authors' analysis of local authority public rights of way datasets



Active
Travel
England

Active Travel England

7/28 GMH, Great Minster House
33 Horseferry Rd, London SW1P 4DR

gov.uk/government/organisations/active-travel-england

