

Exploring and analysing open transport data

with a focus on SCOOT data and bikeshare data

Overview of the session

- Part I (presentation): 10:00 – 10:45
 - Open data sources by mode:
 - Cars
 - Bikes / pedestrians
 - Public transport
 - How do we obtain them?
 - Applications:
 - A. Popularity of parks
 - **B. COVID and mobility**
 - C. Shared bike as a substitute or complement for subway?
- Part II (practical session – R studio): 11:00 – 11:45

About me chauman.fung@glasgow.ac.uk

- Currently a research associate in transport analytics at the Urban Big Data Centre, University of Glasgow
- Research interests: transport taxes (road pricing), optimal PT supply, accessibility
- Selected publications:
 - *How rural is too rural for transit? Optimal transit subsidies and supply in rural areas* (joint work with Maria Börjesson, and Stef Proost). Journal of Transport Geography, Vol88, October 2020, pp 1-14. <https://doi.org/10.1016/j.jtrangeo.2020.102859>
 - *Do buses hinder cyclists or is it the other way around? Optimal bus fares, bus stops and cycling tolls* (joint work with Maria Börjesson, Stef Proost and Zifei Yan). Transportation Research Part A: Policy and Practice, Vol 111, May 2018, pp 326-346. <https://doi.org/10.1016/j.tra.2018.03.023>
 - *Can we decentralize transport taxes and infrastructure supply?* (joint work with Stef Proost). Economics of Transportation, Vol 9(1), March 2017, pp 1-19. <http://dx.doi.org/10.1016/j.ecotra.2016.10.003>

Open data?



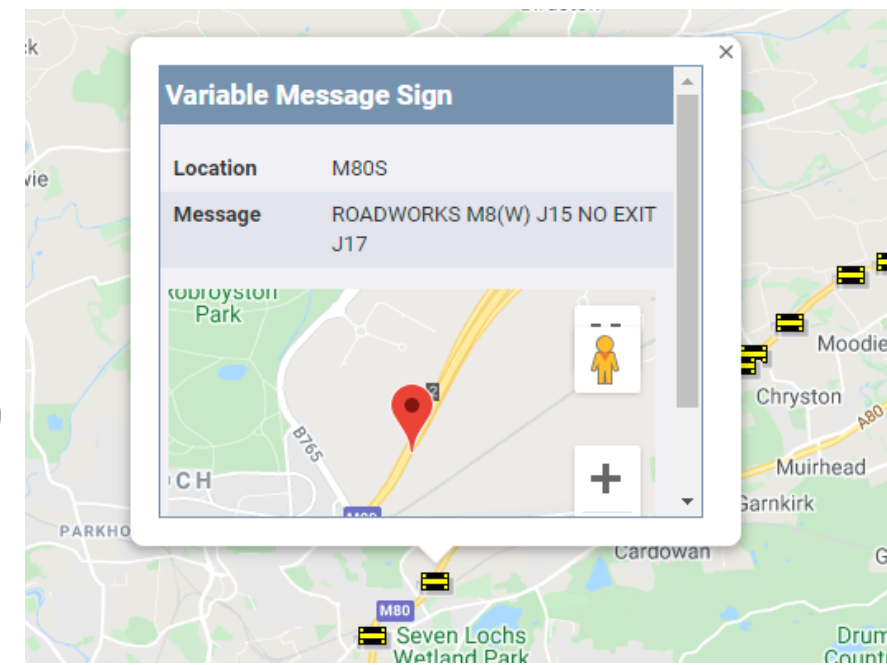
- What is open data?
 - ‘Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.’ (<https://opendatahandbook.org/guide/en/what-is-open-data/>)
- Open Government Licence (OGL)
 - The Open Government Licence (OGL) is a simple set of terms and conditions that facilitates the re-use of a wide range of public sector information free of charge. (<https://www.nationalarchives.gov.uk/information-management/re-using-public-sector-information/uk-government-licensing-framework/open-government-licence/>)
- Beware of licencing terms

Available data sources and how to obtain

- Data sources and tools:
 - [DfT](#), [TfL](#)
 - City council portals: [Manchester](#) , [Leeds](#) , [Scotland](#)
 - [Urban Big Data Centre \(UBDC\)](#)
 - [GitHub](#)
 - Maps: [Google Map](#), [OpenStreetMap](#), [Digimap](#)
- Real time or historical
- How do we obtain the data?
 - Direct download
 - API
 - Web scraping

Cars

- Traffic counts ([Brussels](#), [New Zealand](#), [Glasgow](#))
- Parking availability ([Leeds](#))
- Traffic events
- Variable Message System ([Scotland](#))
- Potholes ([FixMyStreet](#))
- Accidents ([Collision data reports](#))



Bikes/pedestrians

- CCTV (UBDC – [Avenues project API](#))
- [Cycling infrastructure](#)
- Shared bike companies or local authorities
 - Data dump
 - API: real-time bike availability
- A few examples in different parts of the world:
 - [Just Eat Cycles Edinburgh](#)
 - [Blue Bikes Boston](#) (gender, year of birth included)
 - [Citi Bikes NYC](#)

Public transport

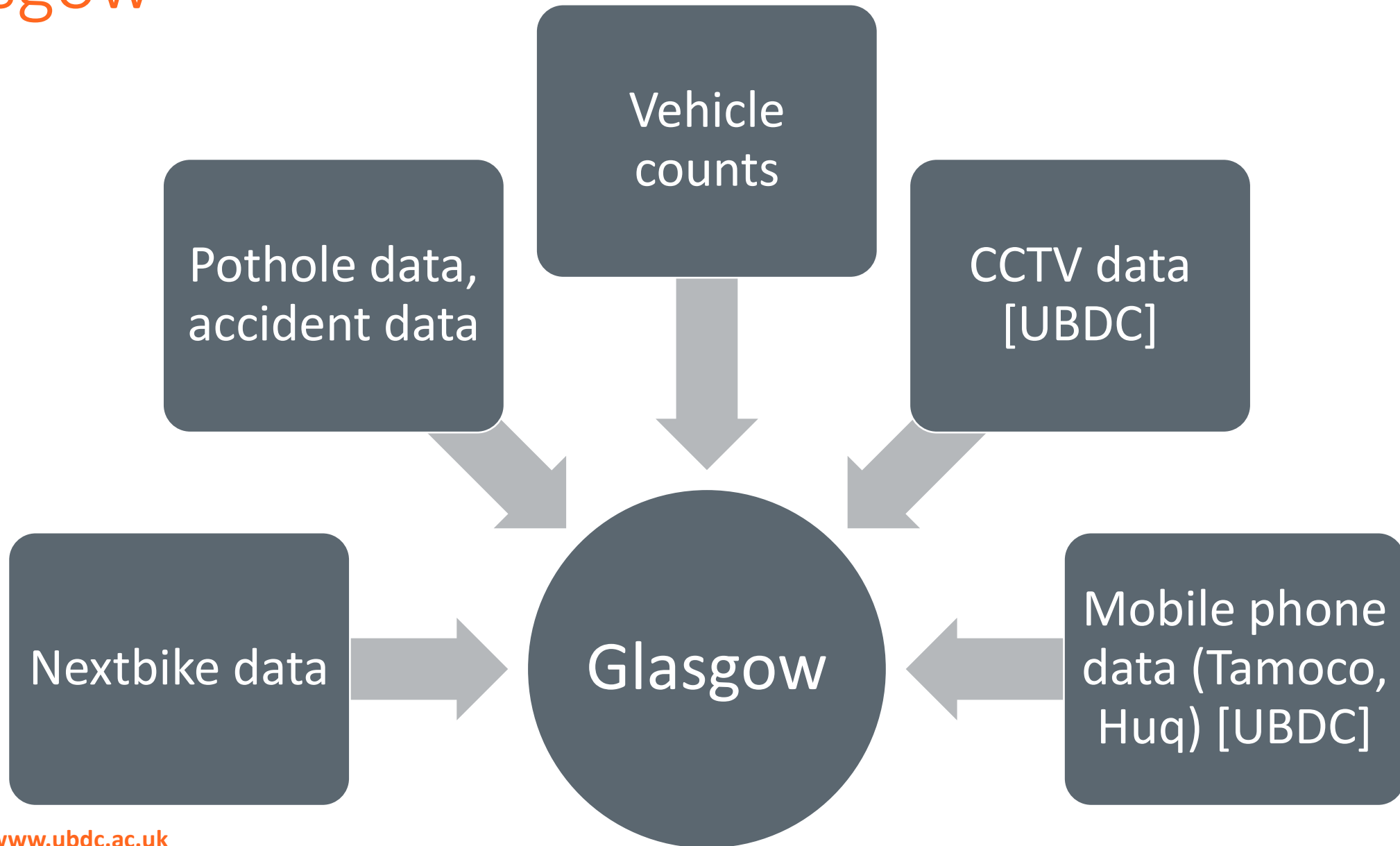
- Timetable in GTFS (General Transit Feed Specification)

[Greater Manchester](#)

- Fares ([Rail Delivery Group](#))
- Transport operators

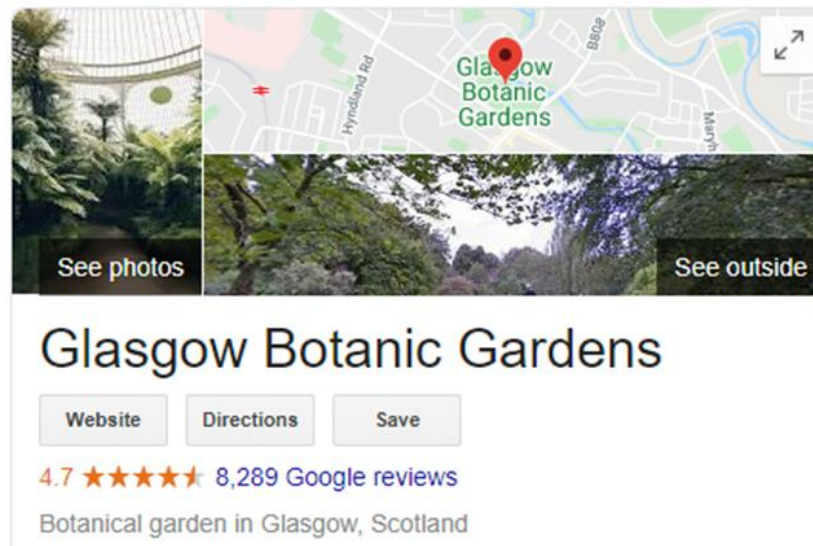


Glasgow



Application (1)

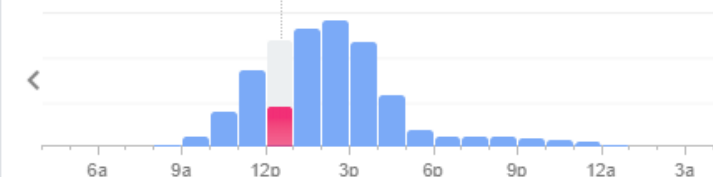
Google Popular Times



Popular times ?

Sundays

Live: Less busy than usual



Send to your phone

Send

Reviews ?

Write a review

Add a photo

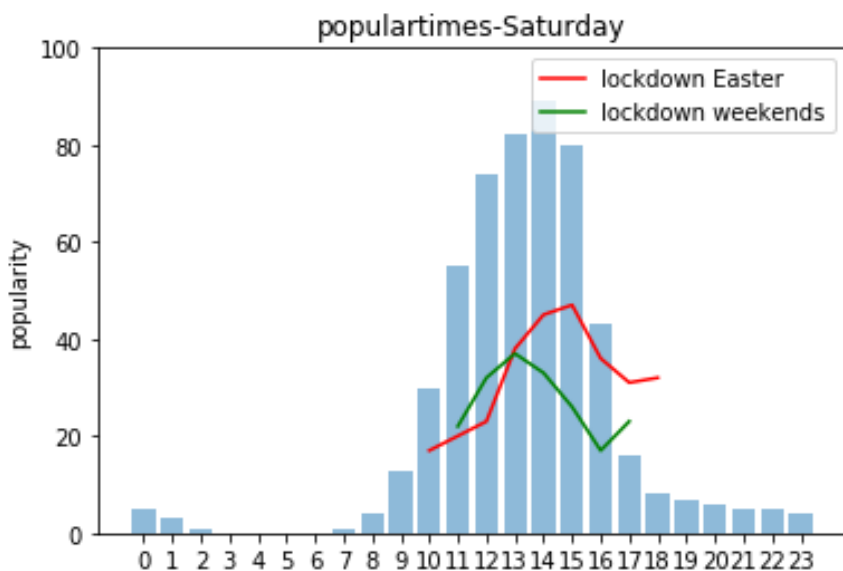
Acclaimed 27-acre park for plants, working vegetable plot and Victorian cast-iron glasshouse - Google

"Waitress service and great food at reasonable prices served by great staff."

"Great size, plenty space to have a walk or sit down on benches"

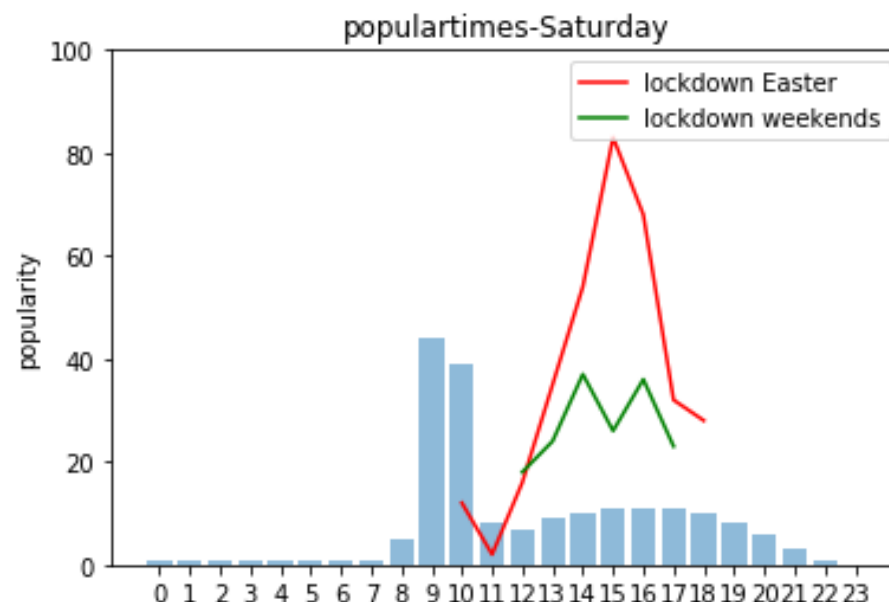
"Parking nearby and the location is close to lovely cafes and restaurants."

[View all Google reviews](#)



A park located at the city centre

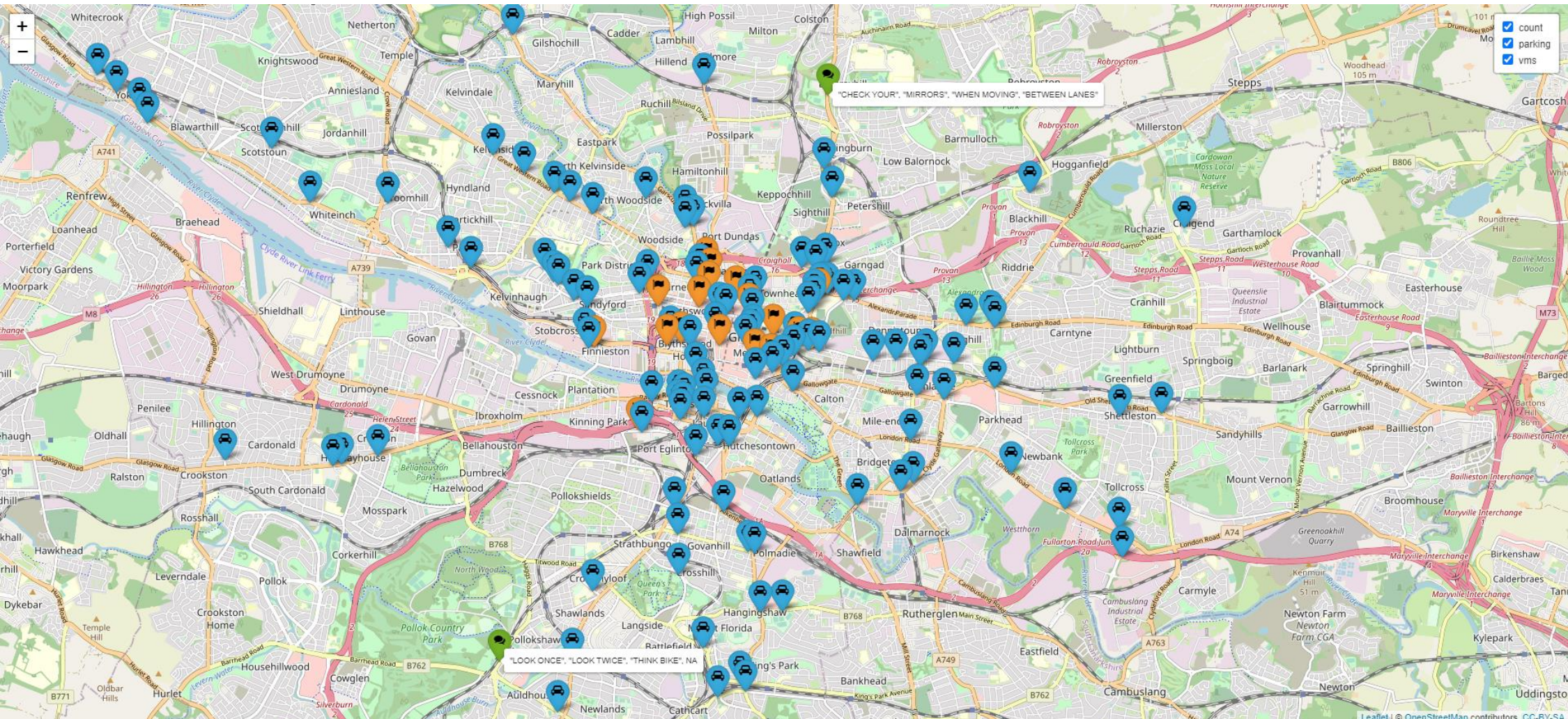
www.ubdc.ac.uk



A park located further from the centre

Application (2) SCOOT data in Glasgow

Blog written by our colleague Jinhyun Hong

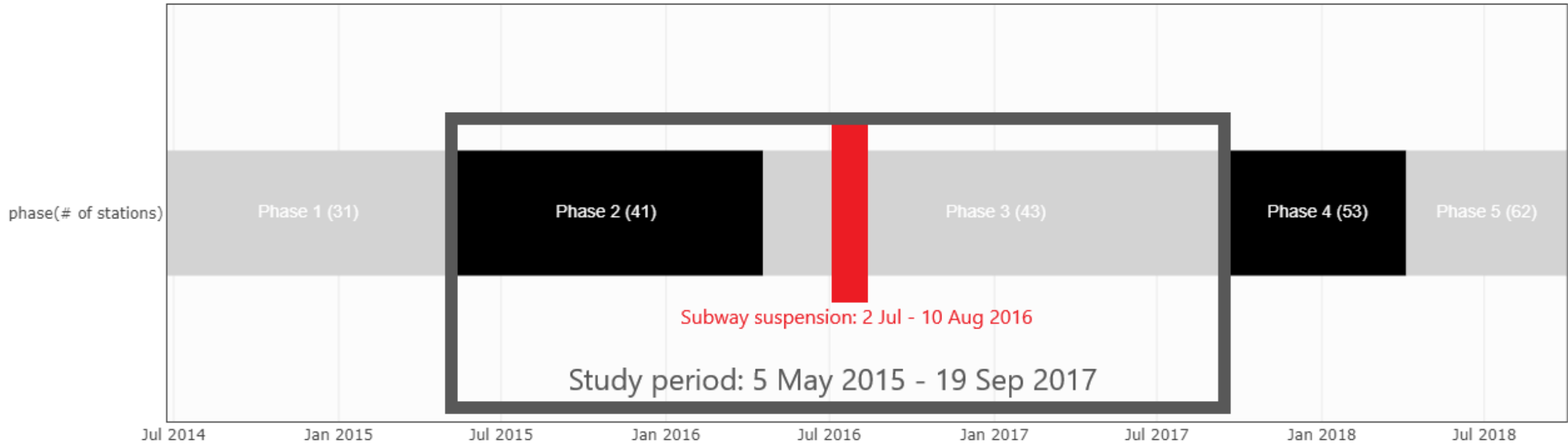


Application (3)

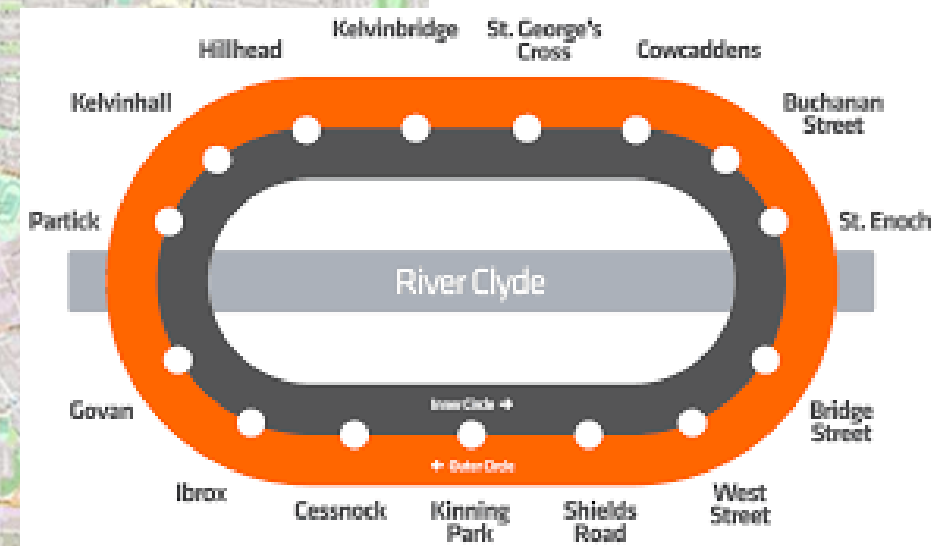
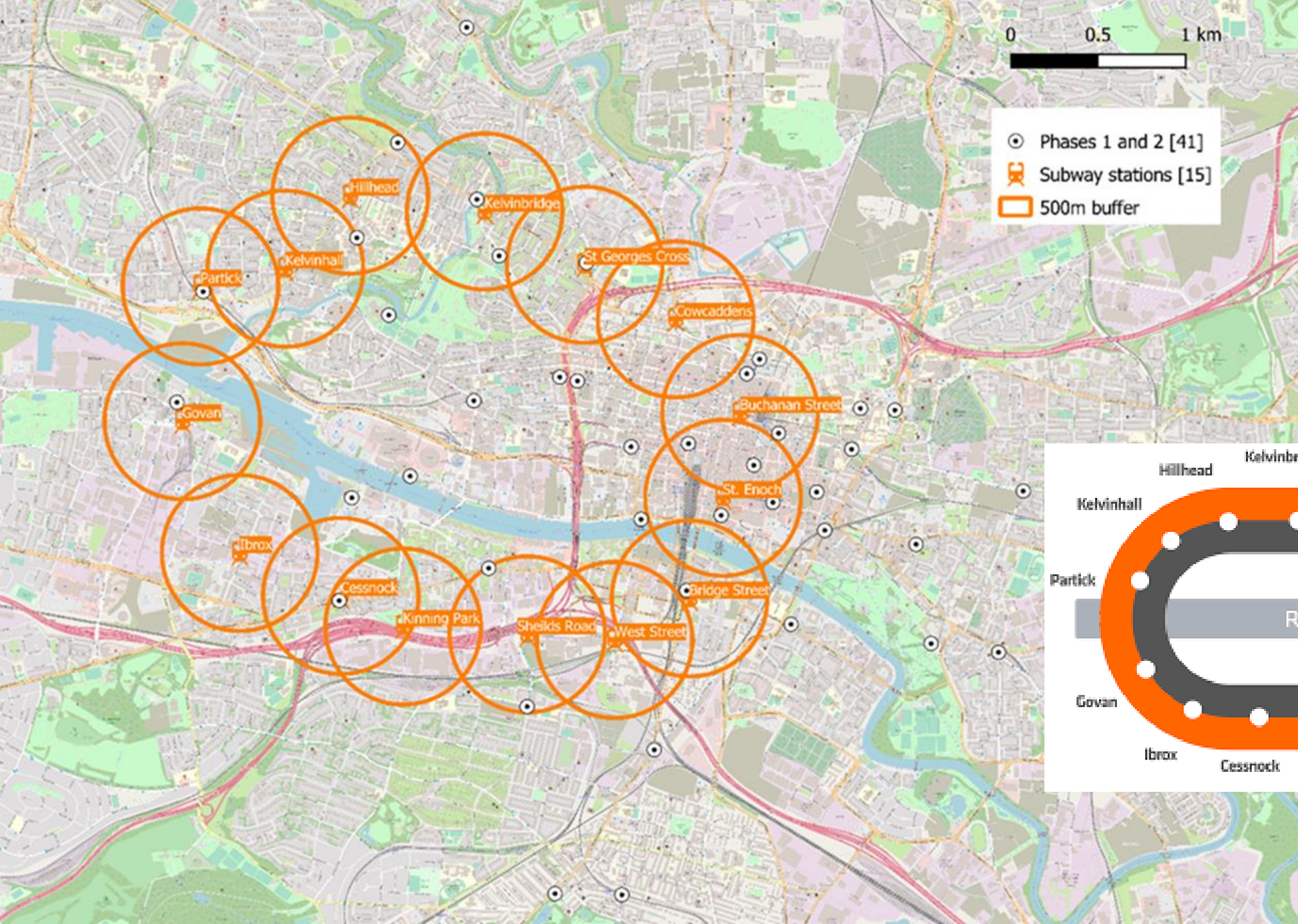
Examining the effects of a temporary subway closure on cycling in Glasgow using bike-sharing data (with David McArthur, Jinhyun Hong)

- Glasgow's subway system is the third oldest subway system in the world
- The service was suspended from the 2nd of July 2016 for essential renewal works to take place
- Nextbike owns 62 docking stations in Glasgow (as of August 2019)
- Objective: To study how different transport modes interact → make use of disruptions which force people to break their usual habits

Background



Timeline created by authors using the *vistime* package (Raabe, 2021) based on data from Mcpherson (2014) and news articles.



Existing studies (1) cross elasticities and diversion factors

- To study how different transport modes interact:

- Cross elasticities:

$$\varepsilon_d = \frac{\% \Delta \text{ in } Q_r}{\% \Delta \text{ in } P_i}$$

- Diversion factors:

when there is one less trip made with the intervention mode due to the intervention, what proportion of the trip shifts to the recipient mode

- Wardman (2014), Fearnley et al. (2018), Dunkerley et al. (2018)

- challenging to generalise or apply the diversion factors involving cycling because they are not symmetric and data on cycle interventions are limited
 - the recommended value of diversion factor with metro being the intervention mode and cycling being the recipient mode is 0.05 in an urban area

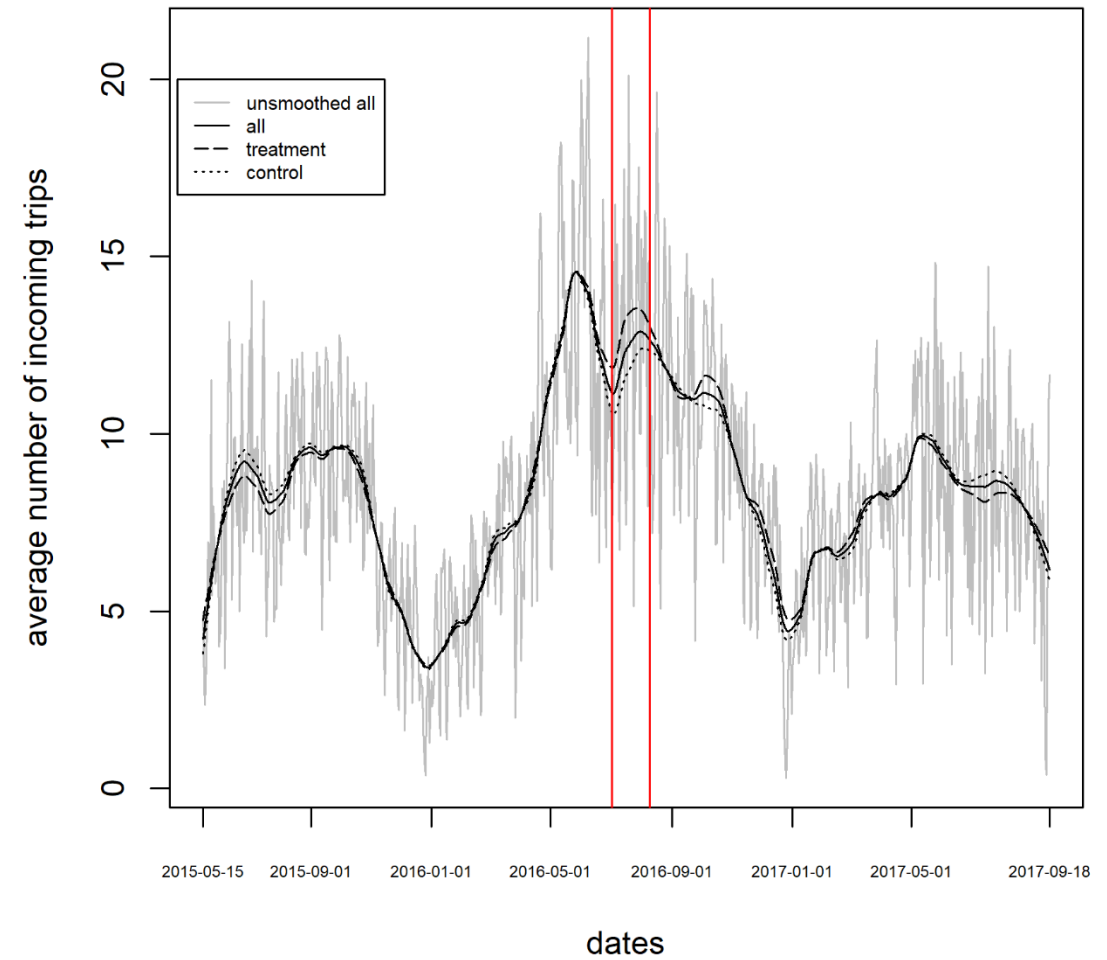
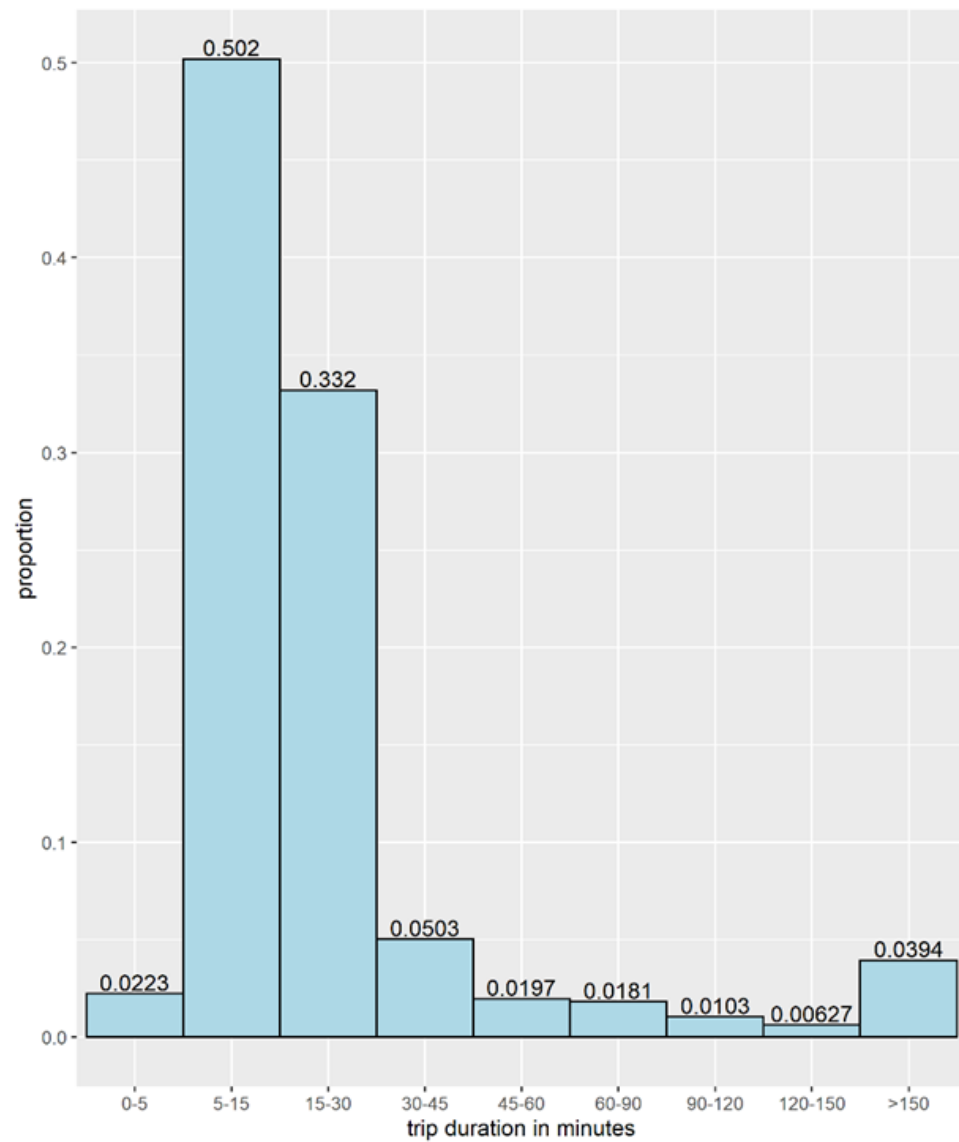
Existing studies (2) substitutability and complementarity of trips

- Hall et al. (2018) found that Uber was a complement for the average-sized public transport agency, and the effect was even stronger for smaller public transport agencies in larger cities
- Martin & Shaheen (2014) In Washington DC, the shift was toward rail and bus in the urban periphery but away from rail and bus in the inner urban areas. The reason is bike-sharing often serves as a first-mile last-mile facilitator in less dense areas.

Method

- Fuller et al. (2019) is the closest to our study: similar intervention (PT strikes vs PT suspension) and similar datasets of shared bike trips and control variables but we do not adopt the interrupted time series techniques
- The station-level data allows us to divide the bike stations into treatment and control stations within the city
- $trip_{it} = FE_i + FE_t + \beta X_{it} + \delta_1 D_{sus} + \theta D_m + \mu_1 D_{sus} D_m + \delta_2 D_{post} + \mu_2 D_{post} D_m + \varepsilon_{it}$

<i>Time period</i>	Treatment	Control
Pre-suspension	θ	0
Suspension	$\delta_1 + \theta + \mu_1$	δ_1
Post-suspension	$\delta_2 + \theta + \mu_2$	δ_2



Hypothesis

Since the Glasgow subway serves a dense city area, city centre subway trips and bike sharing trips are expected to be substitutes. In other words, the coefficient μ is expected to be positive.

Findings

- Subway suspension brought 1.69 more incoming bike-sharing trips and 1.65 more outgoing bike-sharing trips per day than before the suspension for bikeshare stations (treatment group) close to the subway stations
- 20.7% increase in incoming trips and 20.1% of increase in outgoing trips when compared to the average number of incoming trips per day in July in the study period
- Only 12.4% ($=0.21/1.69$) of bikeshare users kept using the service after the resumption of service
- The diversion factor is much smaller than the suggested value of 0.05

Thank you
Feel free to enter your questions/comments in the
chat box

To stay up to date with our research and other activities you can sign up to our newsletter on our website and follow us on social media:

www.ubdc.ac.uk



/urbanbigdata



@UrbanBigData



Urban Big Data Centre