Capstone Proposal

Taha Bouhoun
Computational Sciences
Data Science and Statistics

CP192 - Prof. Carter, W@18:00

March 8, 2020

An author's note

The incertitude around my previous project idea on sustainable mobility led me to reconsider another topic within the same theme. I tried to narrow down the literature review scope and focus on the new paradigms of the mobility industry. My aim is to synthesize an updated perspective of the main challenges in light of the growing urbanization trend, as well as learn about past experiences and attempted solutions in managing mobility demand.

The rationale behind the pivot was purely due to timeline and external dependency constraints. Furthermore, the project was a bet on the rising environmental consciousness since multiple experiments proved that "informing" is not a sufficient nudge mechanism for behavioral change similar to how people are desensitized to cigarette package graphic health warnings.

My current project is rather more technical and doesn't have major external dependencies (requires open-source data that is already available for academic research). The metrics for the deliverable are well defined and serve as a basis for a policy recommendation on traffic flow management. (180 words)

Impact of Infrastructure Changes on Traffic Flow

Abstract

Systems involving many interacting agents are prominent in many social and natural phenomena. Assessing causal relations in such non-linear systems offers insights into their behavior and constructs the basis for policy decisions. The aim of this project is to evaluate the impact of infrastructure changes (e.g., due to renovation) on the flow of traffic using adaptive network simulation. Furthermore, the project would use the case of the London Tower Bridge closure as a case study and offer a policy recommendation on managing traffic flow. (95 words)

Contents

1 Background and context

The literature review presented two major ideas that are central to the scope of this project:

- Democratizing road space: a utilitarian approach to managing infrastructure by giving a preferential allocation of road space to means that carry the most people (e.g., exclusive bus lanes). In the context of developing countries, enhancing the road network is a positive externality that is exclusive to car owners, thus, these solutions are implicitly benefiting people based on their socioeconomic levels (i.e., owning a car implies that one's relatively richer) which is problematic.

- Managing demand rather than enhancing supply: The argument raised a concern about policymakers' approaches for mitigating congestion in cities by investing in infrastructure (*spatial thinking*) which contradicts the idea of "democratizing".
  Lewis Mumford said that "Adding highway lanes to deal with traffic congestion is like loosening your belt to cure obesity." (Mumford. L., 2015). However, the challenge was to propose a metric to gauge the concept of democratizing and the way we assess the utility of transportation means.

Mobility is defined as the aggregate result of the choices made by individuals and of the logistic organization of those choices into a mobility chain (Macario, R, 2011). Thus, I think that assuming a linear functional relation when assessing changes in infrastructure might not lead to accurate results. To illustrate the case of the closure of the London Tower Bridge in late 2016, Uber Movement offered open-source data for researchers and policymakers to assess the effect of the renovation. In their article, they stated the average time travel between two given points on different sides of the bridge prior to its closure then compared it against the same metric after the treatment occurred.

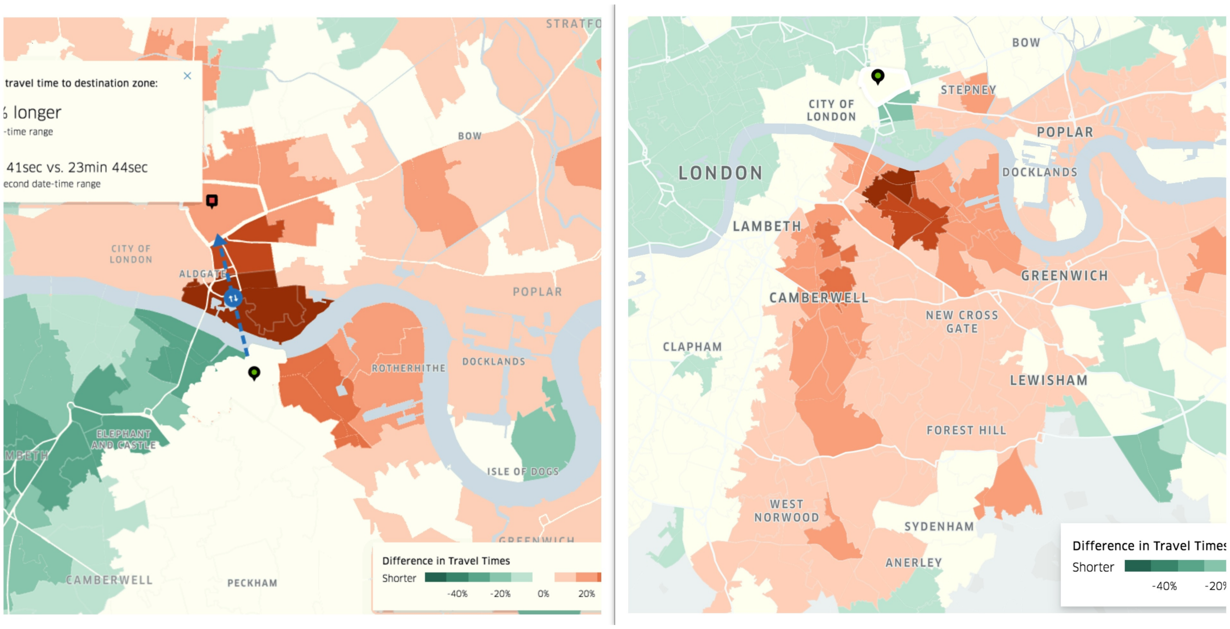

Figure 1. Illustrating the change in the average travel time between two regions on different sides of the bridge (*Source: Uber Movement team, 2018*)

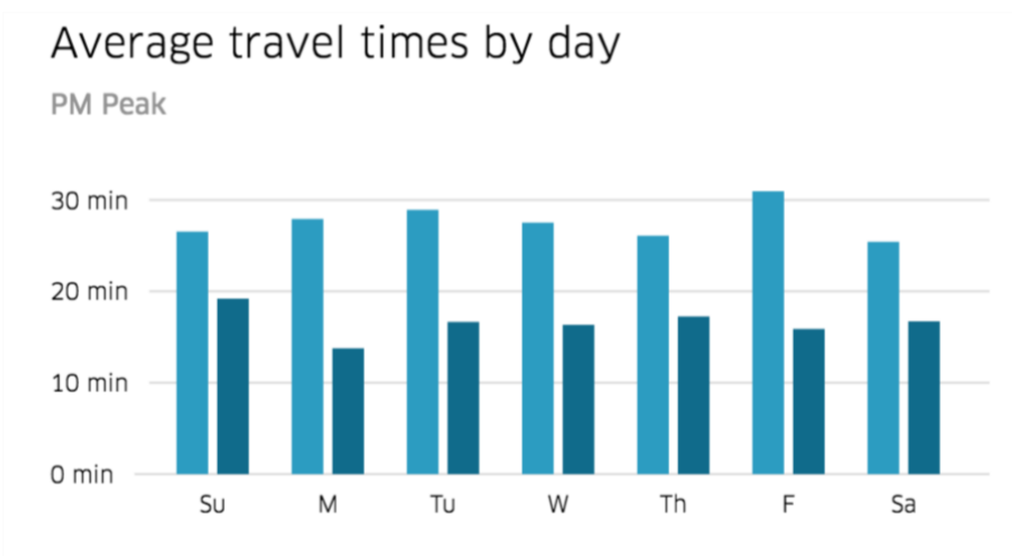

Figure 2. The change in the daily average travel time between two regions on different sides of the bridge (*Source: Uber Movement, 2018*)

The Uber movement team estimated an increase of 65% in the southbound average travel time and roughly 35% on the northbound average travel time between two regions on different sides of the bridge. A closer look at their methodology would raise many questions on the accuracy of their estimates:

1. Shouldn't we control for the number of trips between both sides of the bridge? It's possible that people chose to seek alternative means to move to the other side of the bridge, in such cases, the treatment effect (bridge closure) is underestimated. On the other hand, we can also hypothesize that there is a long-term rising trend of car-sharing over time which contributes to the increase in average travel time, in such cases, the treatment effect is such cases is overestimated.

2. The traffic network is interdependent, so isn't problematic to assume that the rise in the average travel time in a given region is entirely dependent on any other region. To illustrate this point, we can imagine having four regions stacked on top of each other with a bridge passing in the middle to separate them into two regions at each side of the bridge. If we consider the impact of the bridge closure of the far apart regions, we are then adding up the aggregate impact of all regions that link between them.

3. Similar to the second question, the estimates are prone to spurious correlation when two events are associated but not causally related. In such a case, generating a simulation can help test to what extent the closure of the bridge will -on average- delay the commuting between both sides.

(640 words)

2 Description of the proposed project

2.1 Problem statement

Causal inference is pervasive in the analysis of complex systems, and the dynamics between its components serve as a basis for policy making. In the context of mobility systems, disentangling the causal effect of a change in infrastructure cannot be accurately assessed using linear models given that they assume linear functionality and are prone to confounding factors that are hard to control for.

2.2 Approach and methodology

The intended approach is to use adaptive network simulation to better estimate the effect of a change in infrastructure. First, we build a grid of a road network of a given city (e.g., London) where each edge of the network refers to a road segment and the intensity of the edge represents the average flow of traffic along a segment of road. On the other hand, the nodes of the network represent the intersection of two or more road segments.
In such a model, the closure of the London Tower Bridge would be equivalent to dropping one of the edges of the network. Consequently -holding the traffic flow constant- we can then assess the true effect of the loss of the bridge segment by running multiple random simulations of traffic flow.

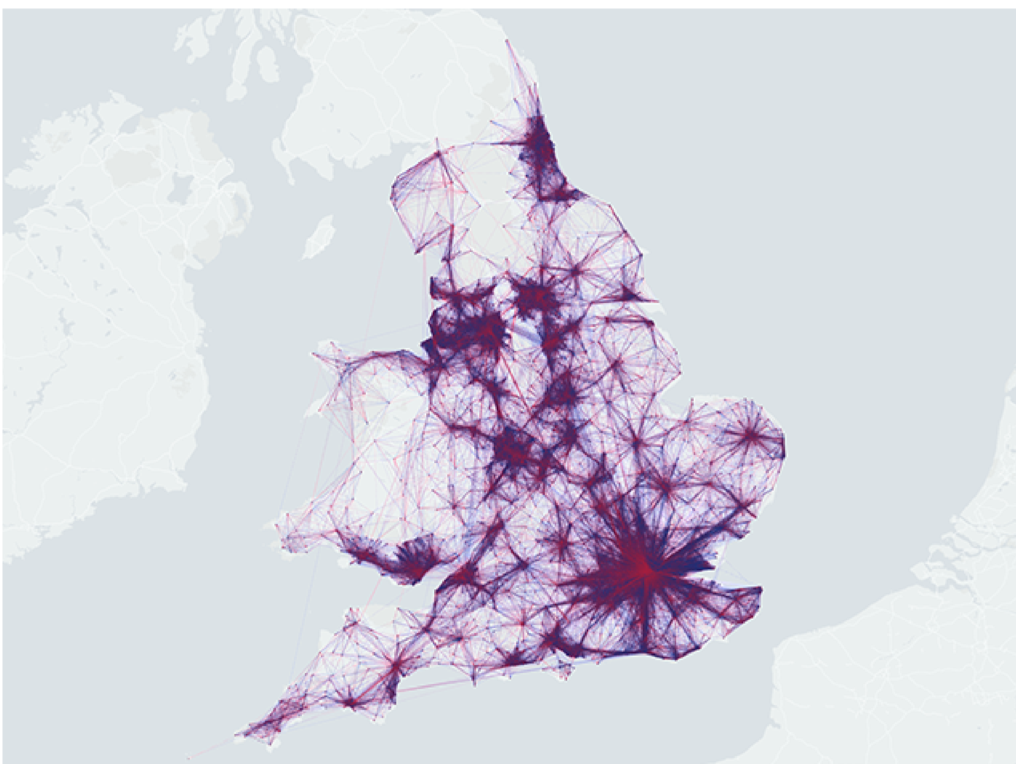

Figure 3. Visualization of the road network in the United Kingdom and how it can be considered an adaptive network (*Source: Kepler, 2018*)

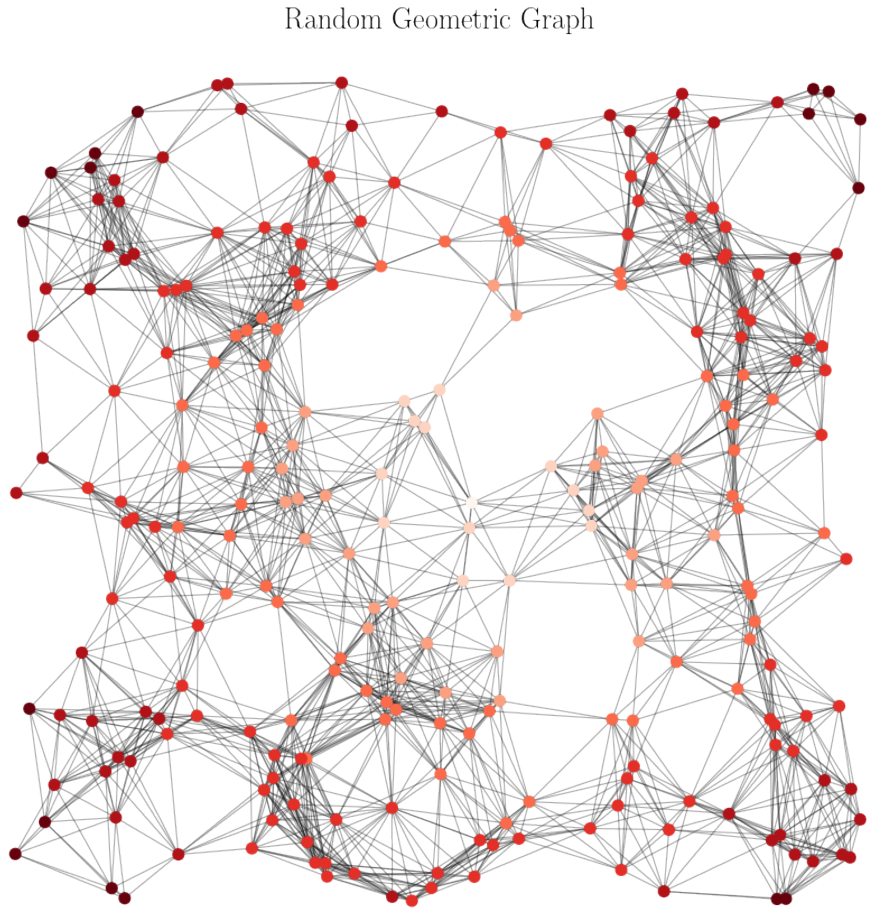

Figure 4. Random Geometric Graph generated using Python package **Networkx**. Notice that the color-coding gets darker relative to the position of the node with respect to the center of the graph which illustrates that it takes more time to reach nodes at the edges starting from the center (*Code in appendix*)

Such graphs can offer an approximation of the road network in a given area. In addition, we can build our analysis based on the dynamics of networks and dynamics on networks (Refer to appendix 2 for definitions). Upon running simulations for random traffic flows (i.e., random requests for moving from a node to another), we can gather a set of recommendations for managing the demand from the point of view of policymakers.

Moreover, simulating adaptive networks allow a holistic overview of the change in the system. As an example of the Random Geographical Graph, I generated two distributions for the shortest path between two random nodes and for the degree distribution (the number of edges connected to each node). These parameters can serve as metrics for my simulation since the smaller the path that links between two nodes, the higher the flow of traffic.

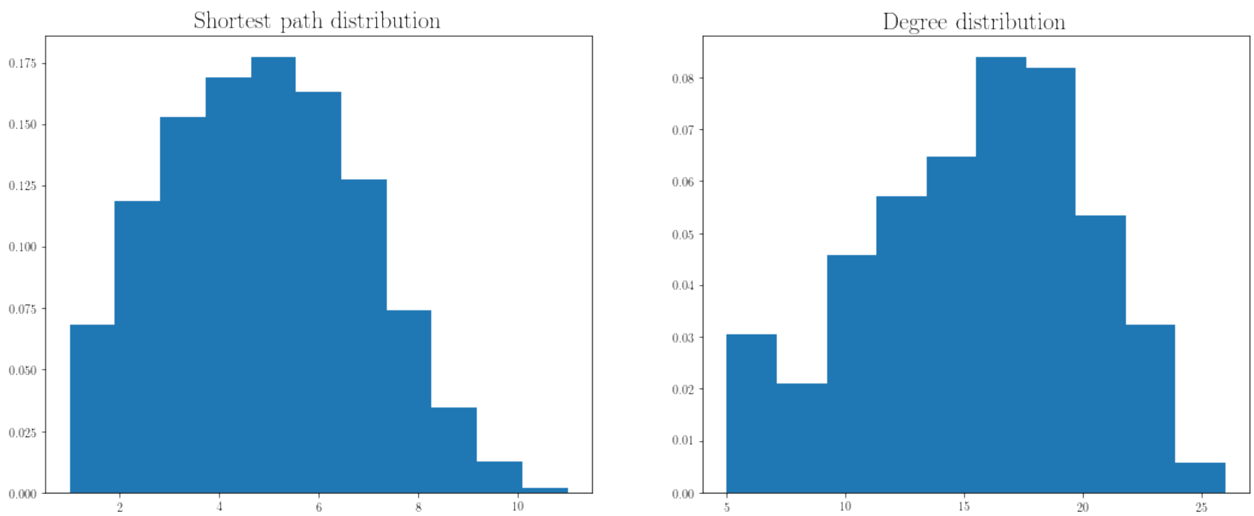

Figure 5. Shortest path distribution and Degree distribution for the Random Geometric Graph (*Code in appendix*)

2.3 Deliverable and outcome

The main purpose of the simulation is to accurately assess the implication of changes in the infrastructure on traffic flow. In addition, these analyses would serve as a basis for urban planning to mitigate the negative effect of these changes (e.g., they don't necessarily have to be due to renovation but they can also be due to festivals, protests, etc). Lastly, I aspire to test new scheme for managing the flow of traffic similar to the "Superblock" experiment conducted in Barcelona where cars were constrained to circulate around Superblocks made of nine blocks (Fig. 6). (597words)

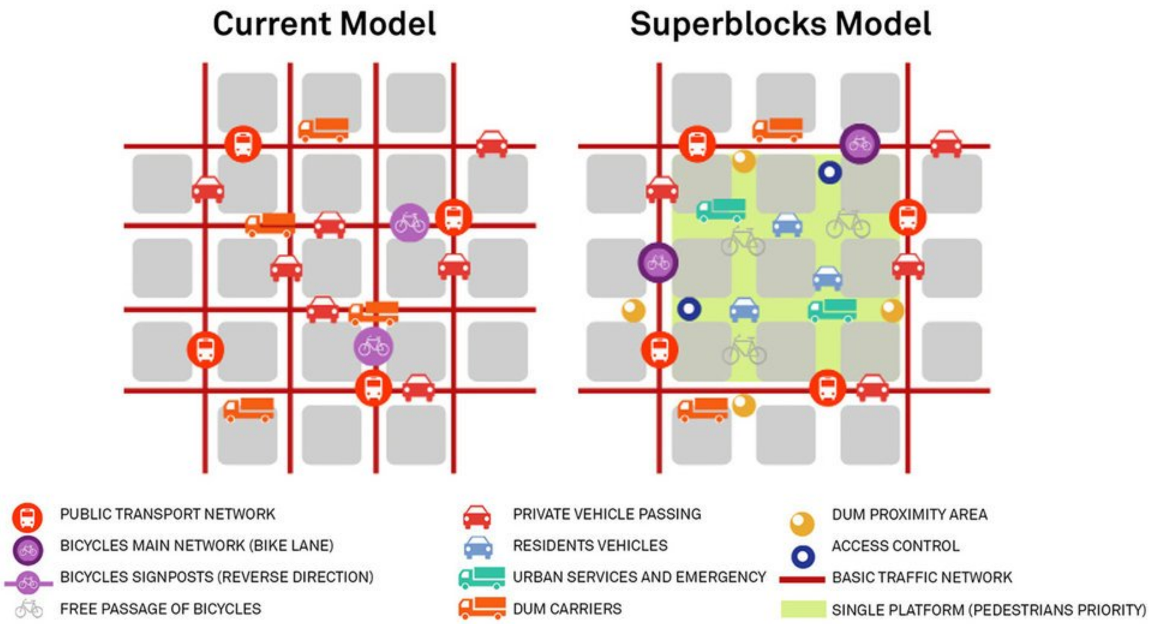

Figure 6. Illustration of the Superblock model compared to the previous ordinary model. A concept coined by BCNecologia led by Salvador Rueda

3 Conclusion

In a nutshell, the motivation of the capstone project is to explore assessing causality in adaptive networks since the assumptions of linear functional relationships cannot hold in a complex system such as traffic in cities. Furthermore, the main claim behind this simulation is to explore ways to better manage demand for mobility rather than enhancing infrastructure, ways that policymakers can focus on moving people rather than moving vehicles.
(68 words)

Appendix

A - Goals, specific objectives, key results or deliverable, and deadlines pertinent to your project.

Timeline	Objectives
April 15th, 2020	• Continue the literature review of causal inference in complex systems • Test the model on fake data to check the validity of the metrics • Discuss the model with Prof. Scheffler as part of the CS166 • Finalize a version of the project proposal
May 30th, 2020	• Finalize the structure of the model specification • Extract the road network of London
July 31st, 2020	[After the AI Lab Internship] • Build the first simulation and compare with Uber's estimates • Gather insight from senior data scientists on the viability of the model
Sept 1st 2020	• Look into examples of policy recommendations • Extend the model to test the impact of bridge closure in New York
Dec 31st 2020	• Recommend ways to manage traffic flow based on the time of the day • Explore ways to include the contribution of micro-mobility

B - LO/HC Application:

#dataviz [HC]: to visualize the results of the simulations.

#modeling [HC]: the process of modeling a version of the actual traffic system.

#networks [HC]: using network to represent the road network.

#algorithms [HC]: simulation is based on algorithms such as Monte Carlo method.

#simulation [HC]: a methodology to consider non-linear functional relations.

#optimization [HC]: optimizing the flow of traffic based on the density of the network.

#case_study [HC]: the closure of the London Tower Bridge as a case study.

#gap_analysis [HC]: bridging the gap between assessing the causal effect in a linear model and using adaptive networks.

#system_dynamics [HC]: highlight the ripple effect of changing infrastructure on traffic flow.

#levels_of_analysis [HC]: analyze the effect of the change in infrastructure for the point of view of commuters, then from the perspective of the whole network.

#network_analysis [LO]: using renormalization group analysis to assess the theoretical thresholds for the dynamics

#python_implementation [LO]: implementing the code for the simulation in Python

#optimization_strategy[LO]: optimize the flow of traffic based on theoretical/mathematical optimization methods

#decision_inference [LO]: turn the findings of the simulation into actionable policy recommendations

#causal_effect [LO]: disentangle causal effect in complex system using simulation

C - Potential sources of relevant support, advice, or feedback:

People	Background, affiliation, and fields of expertise
Prof. Scheffler	PhD in Bayesian Inference from the University of Cambridge and professor of both data sciences and modeling simulation. I aim to discuss my plans for the simulation as well as any assumptions and limitations that my project might face.
Prof. Diamond	Professor of Information-based decisions class, I aspire to get in touch to discuss the causal inference aspect of the course as well as the confounders that coarse-gaining might have.
Prof. Howe	Professor of quantum physics at Minerva. Although I didn't have a chance to take his NS class, I was able to meet up with Prof. Howe in BA and discuss the similarities of the mathematics of complex social systems and the simulations at the quantum level.
AI Lab Team	Having my offer to intern with AI Lab in Japan would be a huge step to work on this project, I plan to connect with senior data scientist and explore their point of view

D - Definitions and Keywords:

- **Dynamics on network**: If there is a dynamic process occurring on the network while the network structure remains static.
- **Dynamics of network**: If the dynamics are limited to the evolution of network structures only.
- **Degree of a node**: is the number of edges connected to the node.

E - Source code [https://github.com/Tahahaha7/Capstone_Project]

```
# FIGURE 4
Graph = nx.random_geometric_graph(250, 0.15)
pos = nx.get_node_attributes(Graph, 'pos')
d_min, n_center = 1, 0
for n in pos:
    x, y = pos[n]
    d = (x - 0.5)**2 + (y - 0.5)**2
    if d < d_min: n_center, d_dim = n, d
p = dict(nx.single_source_shortest_path_length(Graph, n_center))

# FIGURE 5
degree_dist = [degree for node, degree in Graph.degree]
short_path_dist = []
for source, source_lengths in nx.all_pairs_shortest_path_length(Graph):
    short_path_dist.extend(x[1]
    for x in source_lengths.items() if x[0] > source)
```

References

[1] Malik, N., Shi, F. (2017). Adaptive networks in action: Opinion formation, epidemics and the evolution of cooperation. Dynamical Systems Magazine.

[2] Cunningham S., (2018). Causal Inference: The mixtape (V. 1.7).

[3] Uber Movement (2018). Examining the Impact of the London Tower Bridge Closure

[4] Barcelon architecture walks (2016). Ajuntament de Barcelona

[5] Kepler (2018). Make an impact with your location data

[6] Macario, R. (2011). Managing urban mobility systems. Bradford: Emerald Group Pub. (2011)

[7] Mumford., L.(2015). From the ground up: observations on contemporary architecture, housing, highway building, and civic design