

Beecham, R. (2015). Using Bikeshare Datasets to Improve Urban Cycling Experience and Research Urban Cycling Behaviour. In: R. Gerike & J. Parkin (Eds.), *Cycling Futures: From Research into Practice*. (pp. 267-283). Farnham: Ashgate. ISBN 978-1-4724-5361-7



**CITY UNIVERSITY
LONDON**

[City Research Online](#)

Original citation: Beecham, R. (2015). Using Bikeshare Datasets to Improve Urban Cycling Experience and Research Urban Cycling Behaviour. In: R. Gerike & J. Parkin (Eds.), *Cycling Futures: From Research into Practice*. (pp. 267-283). Farnham: Ashgate. ISBN 978-1-4724-5361-7

Permanent City Research Online URL: <http://openaccess.city.ac.uk/12660/>

Copyright & reuse

City University London has developed City Research Online so that its users may access the research outputs of City University London's staff. Copyright © and Moral Rights for this paper are retained by the individual author(s) and/ or other copyright holders. All material in City Research Online is checked for eligibility for copyright before being made available in the live archive. URLs from City Research Online may be freely distributed and linked to from other web pages.

Versions of research

The version in City Research Online may differ from the final published version. Users are advised to check the Permanent City Research Online URL above for the status of the paper.

Enquiries

If you have any enquiries about any aspect of City Research Online, or if you wish to make contact with the author(s) of this paper, please email the team at publications@city.ac.uk.

Chapter 13

Using Bikeshare Datasets to Improve Urban Cycling Experience and Research Urban Cycling Behaviour

Roger Beecham

Introduction

With access to public and shared transport systems becoming increasingly digitized, transaction datasets of unprecedented size as well as temporal and spatial precision are automatically generated (Blythe and Bryan 2007; Bagchi and White 2005; Pelletier et al. 2011). Data collected through smartcard payment methods are perhaps the largest and most obvious example. Although introduced for the purpose of improving payment processes, such data provide a detailed view of demand on a transport system, the potential for service improvements to be suggested (Ferrari et al. 2014) and an opportunity for studying individual traveller behaviour (Agard et al. 2006; Morency et al. 2006; Lathia et al. 2013). A substantial benefit of such data over more traditional data collection methods is that a complete and total record of usage for every smartcard customer is automatically generated (Bagchi and White 2005). Problems associated with sampling and recall bias, which make actively collected travel surveys somewhat difficult to administer, are avoided. The two most obvious disadvantages, at least for travel behaviour research, are that those individuals using smartcard technology may not be representative of the total population using that system or navigating a city more generally; and that variables such as individual trip purpose can only be inferred since they are not recorded directly.

The recent growth of large-scale, information-technology based (Fishman et al. 2013), urban bikeshare schemes has produced datasets which have similar benefits and challenges. Contemporary bikeshare schemes consist of networks of docking stations at fixed locations, from which bicycles can be collected or returned. Docking stations most often cover the central area of a city and are distributed relatively evenly, with an average distance between each docking station and its nearest neighbour of between 300 and 400 metres, as calculated from a sample of large-scale schemes (O'Brien et al. 2014). An important feature of such self-service bikeshare systems is that users can return a bike to a

different docking station from which it was collected. In addition, the nature of the bicycles themselves (usually heavy and under-gearred) and pricing structures (the first 30 minutes are free) mean that more recent schemes are generally designed to encourage short, high frequency and one-way journeys (Beroud and Anaya 2012).

In a similar way to smartcards, which record ‘tap ins’ and ‘tap outs’ on major metro systems, information-technology based bikeshare schemes contain technologies that allow the start and end of bikeshare journeys to be tracked. Another important aspect of these schemes is that docking station status, that is, the number of bicycles and empty docking points it contains, is reported to central databases in near real-time. This occupancy data alerts scheme operators to individual docking stations of concern, for example, docking stations that are almost full or empty of bicycles. In addition, much of this automatically collected data has been made available to the public, with information on the number of available bicycles and empty spaces at docking stations relayed via web-based maps. While the underlying data can be harvested from websites for analysis, a number of schemes provide web Application Programming Interfaces (APIs) for accessing this and other information, such as individual origin–destination (OD) journey data. There is, then, a digital record not only for the very large number of short, high frequency journeys made through bikeshare schemes; there is also a continuous measure of a bikeshare system’s performance and usability.

Such large-scale observational datasets on urban cycling are unprecedented. This chapter offers an overview of recent studies that have been carried out by researchers working in the fields of data mining, operations research, information visualization, geography and public health. Whilst many of these studies are related, the chapter distinguishes between two categories of work: analysis that directly informs service-improvement and system design and analysis that reveals wider information about urban cycling behaviour. The next section discusses work that has used both occupancy and journey data to tackle problems of bikeshare fleet management and then provides an overview of early approaches that have attempted to use bikeshare data for policy evaluation. The following section identifies studies of individual-level behaviours and work that may contribute more directly to wider cycling research. The chapter concludes by reflecting on the work discussed, as well as some immediate research possibilities.

Bikeshare Data Analysis for System Improvement and Evaluation

Fleet Management

A challenge facing many urban bikeshare schemes is that of fleet management (OBIS 2011). Bikeshare systems consist of collections of docking stations of finite capacity. High demand during peak times often results in bicycles being disproportionately transported from docking stations located in one part of a city to another. Large numbers of commuters may, for instance, wish to travel from a peripheral to a more central location during the morning rush hour. This act renders many docking stations full in central locations and empty in peripheral locations for the majority of the day. Since most schemes operate in relatively compact city centres, there is often limited space to expand or increase capacity. Thus bikeshare operators must manually redistribute bicycles during these peak times (OBIS 2011). An early application for bikeshare usage data has therefore been to provide information that might support the fleet management process.

Within data mining disciplines, a number of researchers have interrogated docking station occupancy data with this purpose in mind. For example, bikeshare docking stations associated with heavy peak-time use were identified. Those docking stations were then often classified according to their temporal usage characteristics, such as how full or empty they are through the course of the day (Borgnat et al. 2011; Froehlich et al. 2008; Côme and Latifa 2014). Related to this work are approaches that aim to accurately forecast the number of available bicycles and spaces at individual docking stations in advance (Kaltenbrunner et al. 2010; Borgnat et al. 2011; Yoon et al. 2012; Guenther and Bradley et al. 2013). Accurate short-term availability forecasts would clearly provide important insights for those operating the bikeshare schemes. Yoon et al. (2012) also used this style of forecasting to develop a personalized journey planner aimed at users of Dublin's CityRide scheme. When given an origin, destination and departure time, their application suggests an optimum pair of stations for minimizing travel times while also maximizing the probability of collecting and returning bicycles at those locations and times.

An alternative means of engaging with the fleet management problem is to present the station occupancy data in a way that can be quickly analyzed and understood. Many scheme operators have made docking station status data available on the web using 'mashup' graphics (Transport for London 2015a; Paris Velib 2015; Barcelona Bicing 2015). Here, markers appear at docking stations and, by clicking on those markers, information on the status of docking stations is reported. O'Brien (2015) developed a particularly informative online system. In this system, circles are used to represent docking stations and they are sized according to each docking station's capacity (that is, the total number of

docking points). By varying the colour of these circles, readers are immediately informed of docking station status. When stations appear as deep red, they are at or very close to maximum capacity. In contrast, when they appear as deep blue, they are close to being empty. To further emphasize stations that are either empty of available docking points or bikes, a border is drawn around the docking stations that are completely full or empty.

One limitation of the 'mashup' approach is that, within the constraints of a map, it is very difficult to encode additional information on docking station availability over time without some form of interaction. This historical information is important because it allows either operators or cyclists to make informed predictions about the scheme's usability at a given space and time. A possible solution, described by Wood et al. (2011), is to move from a spatial to a semi-spatial, or a spatially ordered, grid layout (Figure 13.1). Here, each rectangle represents a docking station at its approximate spatial position, with the River Thames and London's parks included as spatial reference points. Inside each cell is a line chart which summarizes bicycle availability over the last 24 hours. To further emphasize current (live) status, a blue bar appears representing the proportion of bicycles docked at any given time. The top line of the highlighted set of stations in Figure 13.1 shows that Upper Grosvenor Street and Green Street tend to have a relatively balanced ratio of bicycles to docking units over a 24-hour period. In contrast, the Millennium Hotel and St George Street stations fill up with bicycles during the working day and are relatively empty in the evening.

The work described in this section, then, suggests that analysis of bikeshare usage data can help support scheme operation in two ways: in providing insights that inform the manual redistribution and redeployment of bicycles in a system and in offering real-time information to customers on the current and short-term usability of that system.

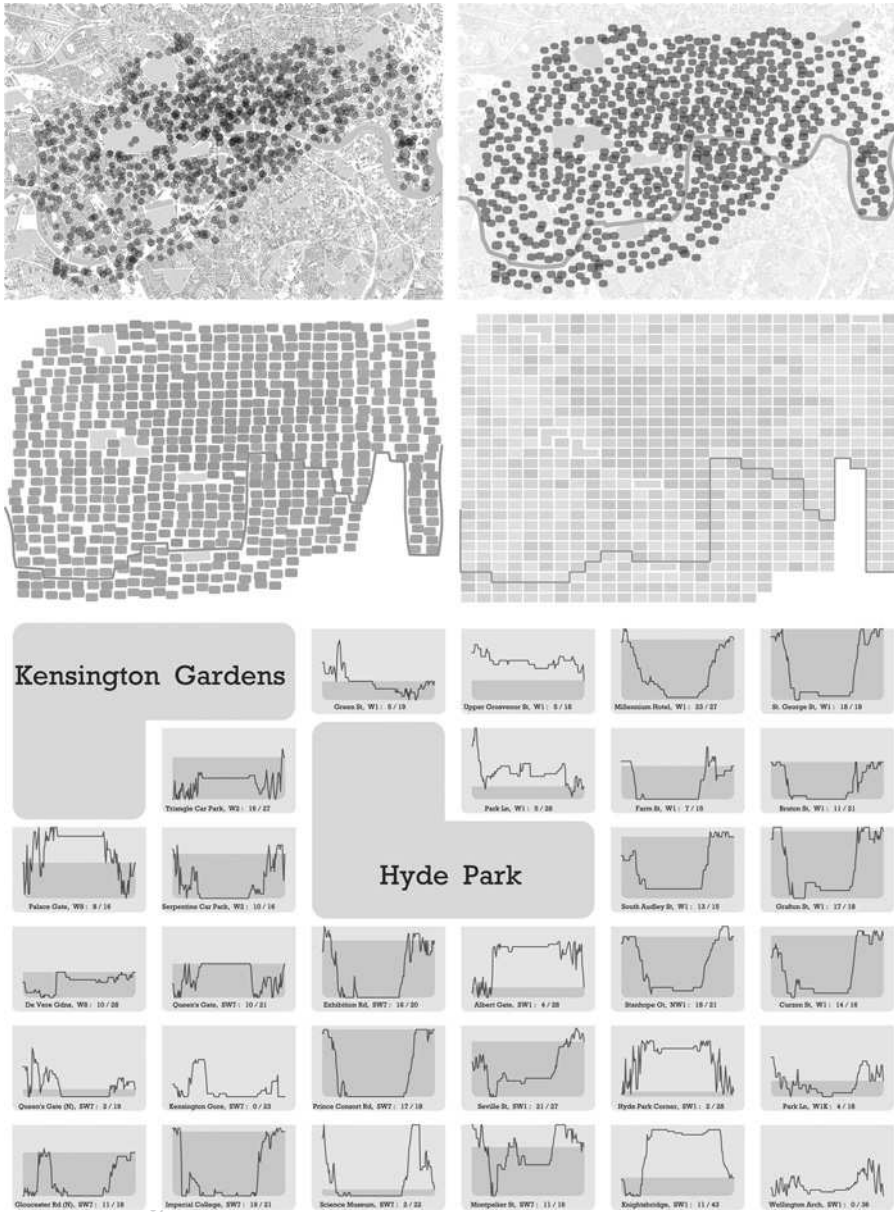


Figure 13.1 Above: Spatial abstraction required in bikegrid to morph from a spatial to semi-spatial grid layout. Bottom: A zoomed in set of stations at 1pm on a weekday

Source: Copyright Jo Wood.

Policy Evaluation

A unique advantage of data collected from modern bikeshare schemes is that the record of usage is continuous and complete. When interrogating historical journey data, it is possible to study the effects of events and interventions on bikeshare usage that are both internal and external to the schemes themselves. This section provides an overview of early studies that have attempted to perform such evaluations.

Working with data from London's bikeshare scheme, Lathia et al. (2012) sought to answer a very specific research question: what was the impact of making the scheme available to casual payment cyclists? For the first four months following its opening on 30 July 2010, the London bikeshare scheme was available only to those formally subscribed as members. However, from 3 December 2010, so-called casual payment customers could also use the scheme. These users release bicycles from docking stations by paying on the day of travel with a credit or debit card. Lathia et al. (2012) took a sample of station occupancy data from a period of time before and after the introduction of casual payment cycling and, controlling for seasonal variables, identified whether and where docking stations were used as a result of the new policy. They found an increase in weekend usage along with a reinforced usage during weekday peak-time. Lathia et al. (2012) then studied changes in the temporal usage characteristics of individual docking stations using clustering techniques. They identified docking stations associated with daytime origins (empty during the day, full at night), daytime destinations (full during the day, empty at night) and stations with both origins and destinations during the daytime. Only very slight changes in station usage were found in most cases between the situation pre- and post- the introduction of casual payment access. However, some stations switched from being associated with daytime origins to daytime destinations and others changed from daytime destinations to daytime origins. This specific information on docking-station-level changes may benefit scheme operators, who might alter redistribution practices at those stations.

Jurdak (2013) later evaluated how cost structures in bikeshare schemes influence usage behaviours. Empirical data from two bikeshare schemes in the USA – Boston and Washington, DC – were analyzed. As with most recent schemes, Boston and Washington, DC operate under a pricing regime in which the first 30 minutes of use is free, with the cost increasing sharply at the 60 and 90 minute points. Jurdak (2013) created a frequency distribution of hires made in both schemes. A very sharp decline in hires just under or around 30 minutes was found, which coincided with the point at which the journeys were no longer free. When hires extended beyond the 30 minute threshold, hire durations were also stretched to the next substantial price increase (at

60 minutes). Jurdak (2013) concluded that, given the clear sensitivity of hire durations to price point boundaries, additional cost structures and incentives might be introduced to positively influence favourable usage patterns. In order to support fleet management, for example, journeys made within the middle of the working day, and which redistribute bikes more naturally after the morning rush hour, might be made cheaper.

While aspects that are internal to the schemes themselves, such as their design, operation, geographic configuration and usability, will impact the way in which those schemes are used, the relative availability of transport alternatives is also likely to affect demand. Fuller et al. (2012) considered interactions between London bikeshare usage and the dominant means of transport in central London, the London Underground. The authors measured the impact of two, day-long London Underground strikes on aggregate usage of the bikeshare scheme and found significant increases in daily bikeshare trip counts during both strikes. Although more modest than on the strike days, the authors also found greater trip counts on the days immediately following the strikes. Fuller et al. (2012) argued that these findings have implications outside of bikeshare schemes themselves: interventions limiting individuals' transport options may help increase uptake of more active travel modes.

Finally, by studying aggregate usage data from Washington, DC's bikeshare scheme and collecting hourly weather data, Gebhart and Noland (2013) considered the impact of adverse weather on the number and duration of bikeshare journeys taken per hour. Additionally, the authors considered journeys that start within close proximity of the city's metro system. They found that while adverse weather did affect aggregate usage, trips taken from bikeshare stations close to Metro stations were disproportionately affected.

Bikeshare Data Analysis for Researching Individual Cycling Behaviour

Although aggregate level usage patterns have been analysed and described in some detail, relatively few studies have considered how regular or returning customers use bikeshare schemes. The shortage in research is perhaps due to a lack of available data; very few schemes are able to make such data publicly available. There are exceptions though. Individual-level data have been released for Boston's bikeshare scheme as part of a data challenge (hubway 2012). Transport for London (TfL), the authority overseeing the London Cycle Hire Scheme (LCHS), has also made such information available to a limited number of researchers. Two datasets have been provided. The first is an origin–destination (OD) journey dataset very similar to that made freely available at TfL's API (Transport for London 2015b). The second is a customer database, which stores information on every customer registered to

use the scheme; it includes a unique customer identifier variable which is also present in the OD journey dataset. Individual customers can thus be linked to their journeys through this customer identifier variable. Together, the journey and customer datasets provide a total record of cyclist-level usage. The next sections describe early analysis of this individual-level data and then reflect on some of the challenges associated with using individual-level data for wider research of urban cycling behaviour.

User Characteristics

Ogilvie and Goodman (2012) were the first researchers to analyse the combined customer and journey datasets and to consider demographics and high-level usage characteristics of LCHS customers. The customer database records the gender and home postcode for every customer subscribing to the scheme. Ogilvie and Goodman (2012) augmented this dataset by matching home postcodes to small area socio-economic indicators and computed distances from postcode centres to their nearest docking station, along with the number of docking stations available within a 250 metre radius of postcodes. Linking the customer and journey datasets, the authors created a linear regression model with 'mean number of trips per month registered' as a primary outcome. Explanatory variables were cyclists' gender, area-level income deprivation and area-level ethnicity. Separate models were created to adjust for relative access to docking stations from customers' home postcodes. The authors found that being female is associated with smaller mean monthly trips and that living outside London is associated with larger mean monthly bikeshare trips. However, after adjusting for access to docking stations (from home postcodes), members living in areas of greater income deprivation were found to be associated with a higher average number of trips per month registered. Access to docking stations within 250 metres of a home postcode was also associated with a greater average number of trips (Ogilvie and Goodman 2012). In a later paper, Goodman and Cheshire (2014) revisited this analysis, using a more recent set of data to identify how usage behaviour changed over the three years since LCHS's launch. They found that many of their earlier findings were consistent, but that the scheme's expansion into areas of greater social deprivation was accompanied by an increase in use amongst customers living in those areas. Additionally, the authors analyzed bikeshare trips made by customers living outside London and who apparently commute into the city by train before using the bikeshare scheme to complete the final leg of their journey. They found that this tendency might be influenced by the relative popularity of cycling in the home towns of those commuters. For example, customers with home postcodes in Cambridge and Oxford (two cities very much associated with cycling) made 2.1 per cent of

such journeys. This was three times as many as expected given the fact their journeys constituted just 0.7 per cent of all commuter journeys into London.

Gender and Bikeshare Usage

Ogilvie and Goodman's (2012) and Goodman and Cheshire's (2014) work not only offers a profile of London bikeshare customers; it begins to consider the possible discriminants of bikeshare usage. Some of the user characteristics they identified might apply to other forms of cycling. Thus, both studies could represent a contribution to the wider literature on urban cycling behaviour. A particularly substantial focus of such literature is on gendered differences in cycling practices (Garrard et al. 2008; Heesch et al. 2012; Emond et al. 2009).

In Beecham and Wood (2014a), spatio-temporal differences in cycling behaviour of male and female LCHS customers were identified and explored in detail. With access to the same dataset as Ogilvie and Goodman (2012), area-level socio-economic indicators were derived and a number of behavioural variables aimed at summarizing individual usage characteristics were calculated. By analysing the spatial and temporal structure of journeys taken by bikeshare cyclists, the authors found noticeable differences in how male and female customers used the scheme. A heavy commuter function was present in men's journeys, whereas an apparent leisure function was found for journeys made by women (Figure 13.2). While Beecham and Wood (2014a) suggested that the different behaviours might relate to differences in the type of men and women subscribing to the scheme, the authors also argued, by controlling for the home location of cyclists and intensity of usage, that the nature of the spatial differences may reflect more fundamental differences in the cycling of men and women (Beecham and Wood 2014a). The authors stated that their research provides empirical support for findings previously identified in social attitude surveys and observed in small-scale GPS-based studies. Perhaps the most obvious is an apparently distinct preference amongst female users for cycling separately from traffic and in parts of the city associated with slower-moving roads.

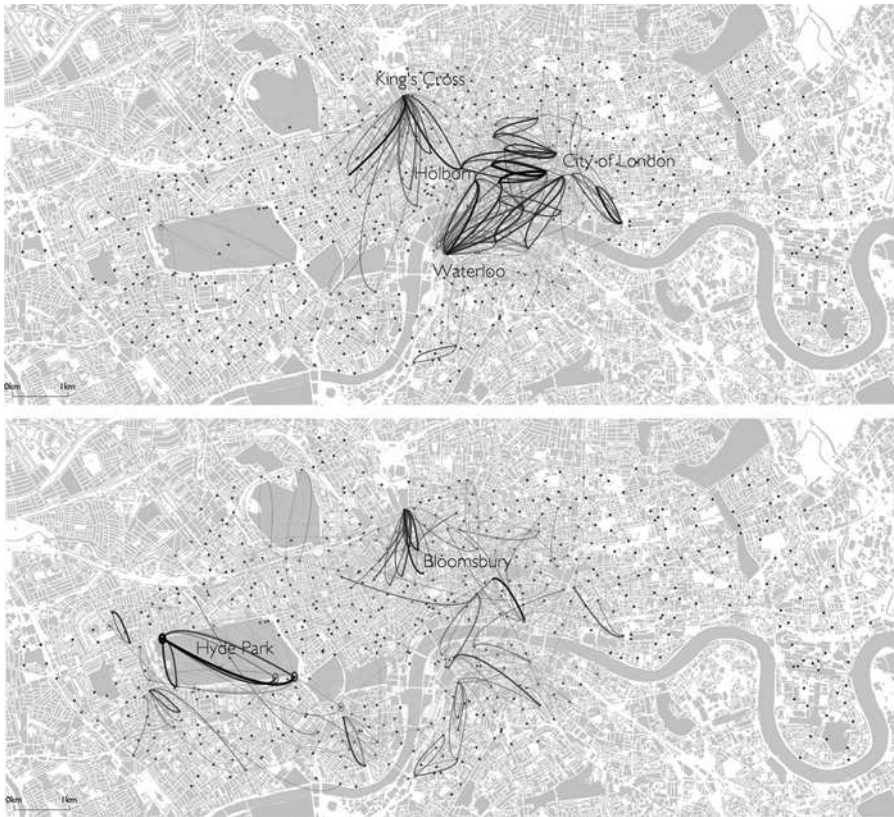


Figure 13.2 Above: All journeys taken by men from the LCHS's inception through to September 2012. Below: All journeys taken by women over this same time period.

Notes: Above: Journeys between London's commuter rail hubs (King's Cross and Waterloo) and employment centres (City of London) dominate. Below: Flows within London's Hyde Park dominate; so too do flows between King's Cross and Bloomsbury. Flow lines representing journeys are weighted according to their relative frequency. To distinguish between origin and destination, they are made asymmetric; straight ends represent the origin, curved ends the destination. See Wood et al. (2011) for a detailed description of this encoding.

Source: Copyright Roger Beecham. Background mapping uses Ordnance Survey data.

Group or Social Cycling

Beecham and Wood's (2014a) exploratory analysis provides large-scale evidence to support findings already discussed in detail using survey-based methods. The authors later argued (Beecham and Wood 2014b) that the size and spatiotemporal precision of bikeshare datasets can also enable under-researched aspects of behaviour, such as group or social cycling, to be analyzed on a larger scale. Here,

the authors approximated group journeys by mining individual customers' journeys. For each customer, instances where that individual made the same journey (matching OD) with another individual at approximately the same time were labelled. Where such 'same journeys' happened with a pair of customers on more than one occasion, they were labelled as bikeshare 'friends' and future journeys made together were identified as group journeys. The authors noted that attempting such an analysis using traditional survey-based methods may be problematic; it would require entire social networks of cyclists to be recruited, provided with GPS devices and monitored over a relatively substantial period of time (Beecham and Wood 2014b).

Beecham and Wood's (2014b) work highlighted two distinct types of group cycling activity: discretionary and imposed. Discretionary group-cycling journeys were found to fit an expected pattern of activity in that they were more likely to take place at weekends, late evenings and lunchtimes; discretionary journeys also appeared to occur within more pleasant parts of the city. By contrast, imposed group cycling coincided with commuting peaks and these journeys were made between very heavily used commuting hub stations, where docking stations are manually replenished with bicycles at peak times. Beecham and Wood (2014b) then studied individual group-cycling networks and found evidence that, especially for women and less experienced users, group cycling may have been a means through which certain types of bikeshare cyclists first began to use the scheme. This latter finding may be prescient given a recent small-scale study which found that, for a group of adult women which had not cycled since childhood, group or social cycling was a motivation for returning to cycling (Bonham and Wilson 2012).

Limits to Studying Individual-Level Bikeshare Cycling Behaviour

So far, this chapter has enumerated the many advantages that bikeshare datasets hold over traditionally collected data. There are nevertheless a number of pitfalls that must be considered. These especially apply to the research described in the two previous sections ('Gender and Bikeshare Usage', 'Group or Social Cycling'), which attempt to make wider inferences from analysing individual behaviour.

It has already been noted that bikeshare schemes offer large and complete population-level datasets, thus providing a clear benefit of scale. In addition, since individuals are not recruited to take part in formal studies, bikeshare data sets avoid problems such as self-selection and social desirability-bias, from which survey-based data typically suffer. The absence of a known, deliberately sampled set of research participants also brings substantial disadvantages. Bikeshare operators record very little demographic information on their customers; only the gender and home postcode of users are directly recorded in the LCHS

database. Ogilvie and Goodman (2012) attempt to provide additional context by linking postcode data to area-level population indicators and in Beecham et al. (2014) a technique is created for estimating spatial locations for bikeshare users' workplaces. However, with little information directly recorded, it is not known how typical bikeshare users are of the wider cycling population, or of the population of the cities in which they operate.

In addition, one of the benefits of studying and comparing individual behaviour within a bikeshare scheme is that factors such as the type of bicycle and (to a lesser extent) an individual's opportunity to access bikes are held constant. It must be noted, however, that bikeshare usage represents a very special case of cycling and it is not obvious how typical an individual's bikeshare cycling might be of their 'normal', non-bike-share cycling behaviour.

The pitfalls above might be partially overcome by surveying a sample of bikeshare customers, recording their demographic characteristics and observing differences between their bikeshare and non-bikeshare cycling behaviour. A more intractable problem is that of measurement error (Goodman and Cheshire 2014). Bikes are typically released from docking stations using access keys and it may be the case that individual cyclists occasionally lend their keys to friends or co-workers. There is therefore no guarantee that the data recorded for an individual user describe only journeys made by that individual.

Finally, this section has discussed the advantages that bikeshare data hold over existing observational datasets, for example GPS-based surveys. However, to the author's knowledge, modern bikeshare schemes do not yet provide technology that allows the location of bikes to be reported whilst in transit and therefore full journey trajectories to be known. Although routing information can be estimated (Woodcock et al. 2014), this fact clearly limits the extent to which important aspects of spatial travel behaviours, such as route preference, can be analyzed.

Discussion and Future Work

The recent expansion of contemporary urban bikeshare schemes has brought new opportunities for those working across a range of data-related disciplines. The largest area of early research has been around fleet management and service improvement. This work has made information on the usability of bikeshare schemes intelligible both to those operating and those wishing to access them. Not discussed here are studies that use historical bikeshare data to create location-allocation models for optimizing the capacity and spatial configuration of bikeshare stations (García-Palomares et al. 2012), as well as models for the optimal redistribution of bicycles in an existing system (Shu et al. 2013).

The fact that data from bikeshare schemes are continuously collected brings numerous opportunities for those wishing to evaluate the impact of specific interventions and events. This is evidenced in the work of Lathia et al. (2012), Fuller et al. (2012), Jurdak (2013) and Gebhart and Noland (2013). Opportunities for furthering this work clearly increase as schemes become more established. For example, bikeshare schemes are generally new facilities and one extension might be to consider whether aggregate responses to events, such as failure in other transport systems, change as users become more aware of the full set of docking stations available to them. Studies of how *individual* users variously respond to events would be particularly instructive; for instance, one might expect certain types of user and usage characteristic to be more sensitive to adverse weather than others. In addition, cities that introduce bikeshare schemes often do so alongside wider improvements in cycling infrastructure. Thus, studying aggregate and individual-level usage both before and after that infrastructure is introduced may provide useful insights into their relative effectiveness.

There is also much potential for using bikeshare data to better understand general cycling behaviour. An obvious benefit, which has been discussed throughout the chapter, is that bikeshare datasets provide a uniquely large record of observed usage. In their study of male and female LCHS users, Beecham and Wood (2014a) analyzed over 80,000 customers making five million journeys in a 12-month period. With this amount of data, usage patterns can be queried at finer temporal and spatial resolutions with fewer concerns about an insufficient sample, a non-trivial point where researchers might wish to identify and study gaps in cycling at particular space-times in a city. Beecham and Wood's (2014b) work also argued that the completeness and spatio-temporal precision of bikeshare datasets make the study of relatively under-researched aspects of behaviour, such as group cycling, possible.

Studying how usage behaviour changes over time might be an immediate way of furthering this early work into individual behaviour. For instance, it might be the case that customers initially use bikeshare schemes for making leisure-oriented journeys before later using bikeshare schemes to commute and make other utilitarian journeys. Large-scale insights into individual 'user trajectories' might then help inform strategies for promoting cycling more generally.

Finally, several of the limitations associated with using bikeshare data to study wider behaviour were outlined. Some of these may be very difficult to overcome; for example, the risk of measurement error due to individual users sharing access keys. As long-term relationships between researchers and bikeshare data owners develop (Wood et al. 2014), it might however be possible to collect more detailed attribute information on scheme users. In addition, it is conceivable that bicycles containing GPS-tracking technology might be introduced to major bikeshare schemes, or that a sample of bicycles might be

fitted with such a technology. This might be a particularly exciting development, as it would be possible to follow approaches so far taken with smaller GPS survey-based datasets (see Broach et al. 2012) in order to investigate how individual cyclists' route choices and preferences vary on a very large-scale, over time and under different conditions.

Summary

The emergence of contemporary, information technology-based bikeshare schemes has resulted in datasets on cycling behaviour of an unprecedented size and spatiotemporal precision. Existing research into such relatively new data can be located within two categories: analysis that directly informs service-improvement and system design and analysis that reveals wider information about urban cycling behaviour. Work that focuses on service improvement has often aimed at solving problems of fleet management from which many bikeshare systems suffer, such as profiling local docking stations according to how they are used over time and creating predictive models of short-term bicycle availability. It has also been possible to directly evaluate the effect of policy changes or, for instance, short-term failure in other transport systems. Research using bikeshare data to better understand cycling behaviour varies in terms of the level of information made available. Perhaps the most promising area of research is around studies of *individual* behaviours. With access to customer data, distinct usage characteristics have been identified and related to existing research findings from more traditionally collected, survey-based datasets. The scale and completeness of these data has also enabled relatively under-researched themes of analysis to be studied.

Many opportunities exist for further using bikeshare data in research and analysis. As the volume of historical usage data grows, so too do opportunities for evaluating in more detail the impact of various events, interventions and environmental conditions. Information on different user types can be inferred from studying individual level data. Further, monitoring these over time may provide important insights to inform strategies for cycling promotion. Finally, bikeshare journey datasets comprise only of simple origin–destination pairs. If journey routes were known, factors such as route preference and selection could be studied on a large scale and over time.

References

- Agard, B., Morency, C. and Trepanier, M. 2006. Mining Public Transport User Behaviour from Smart Card Data. In 12th IFAC Symposium on Information Control Problems Manufacturing, Saint-Etienne, France, 17–19 May 2006.
- Bagchi, M. and White, P.R. 2005. The Potential of Public Transport Smart Card Data. *Transport Policy*, 12(5), 464–74.
- Barcelona Bicing 2015. Barcelona Bicing Scheme. Available at: <https://www.bicing.cat/es/formmap> [accessed: 12 January 2015].
- Beecham, R. and Wood, J. 2014a. Exploring Gendered Cycling Behaviours Within a Large-Scale Behavioural Data-Set. *Transportation Planning and Technology*, 37(1), 83–97.
- 2014b. Characterising Group Cycling Journeys Using Interactive Graphics. *Transportation Research Part C: Emerging Technologies*, 47(2), 194–206.
- Beecham, R., Wood, J. and Bowerman, A. 2014. Studying Commuting Behaviours Using Collaborative Visual Analytics. *Computers, Environment and Urban Systems*, 27(September), 5–15.
- Beroud, B. and Anaya, E. 2012. Private Interventions in a Public Service: An Analysis of Public Bicycle Schemes. *Transport and Sustainability*. Bingley: Emerald, 269–301.
- Blythe, P. and Bryan, H. 2007. Understanding Behaviour through Smartcard Data Analysis. Proceedings of the ICE. *Transport*, 160(4), 173–7.
- Bonham, J. and Wilson, A. 2012. Bicycling and the Life Course: The Start-Stop-Start Experiences of Women Cycling. *International Journal of Sustainable Transportation*, 6(4), 195–213.
- Borgnat, P., Abry, P., Flandrin, P., et al. 2011. Shared Bicycles in a City: A Signal Processing and Data Analysis Perspective. *Advances in Complex Systems*, 14(3), 1–24.
- Broach, J., Dill, J. and Gliebe, J. 2012. Where Do Cyclists Ride? A Route Choice Model Developed with Revealed Preference GPS Data. *Transportation Research Part A: Policy and Practice*, 46(10), 1730–40.
- Côme, E. and Latifa, O. 2014. Model-Based Count Series Clustering for Bike-Sharing System Usage Mining, a Case Study with the Vélib System of Paris. *ACM Transactions on Intelligent Systems and Technology*, 5(3), 1–21.
- Emond, C.R., Tang, W. and Handy, S.L. 2009. Explaining Gender Difference in Bicycling Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, 2125, 16–25.
- Ferrari, L., Berlingiero, M., Calabrese, F. and Reades, J. 2014. Improving the Accessibility of Urban Transportation Networks for People with Disabilities. *Transportation Research Part C: Emerging Technologies*, 45(August), 27–40.
- Fishman, E., Washington, S. and Haworth, N. 2013. Bike Share: A Synthesis of the Literature. *Transport Reviews*, 33(2), 148–65.

- Froehlich, J., Neumann, J. and Oliver, N. 2008. Measuring the Pulse of the City through Shared Bicycle Programs. In Proceedings of the International Workshop on Urban, Community, and Social Applications of Networked Sensing Systems. UrbanSense08. Raleigh, North Carolina, USA, 16–20.
- Fuller, D., Sahlqvist, S., Cummins, S. and Ogilvie, D. 2012. The Impact of Public Transportation Strikes on Use of a Bicycle Share Program in London: Interrupted Time Series Design. *Preventive medicine*, 54(1), 74–6.
- García-Palomares, J.C., Gutiérrez, J. and Latorre, M. 2012. Optimizing the Location Of Stations in Bike-Sharing Programs: A GIS Approach. *Applied Geography*, 35(1–2), 235–46.
- Garrard, J., Handy, S. and Dill, J. 2012. Women and Cycling. *City Cycling*. London: MIT Press, 211–35.
- Garrard, J., Rose, G. and Lo, S.K. 2008. Promoting Transportation Cycling for Women: The Role of Bicycle Infrastructure. *Preventive Medicine*, 46(1), 55–9.
- Gebhart, K. and Noland, R.B. 2013. *The Impact of Weather Conditions on Capital Bikeshare Trips*. In Transportation Research Board 92nd Annual Meeting, Washington DC, USA, 13–17 January 2013.
- giCentre 2015. BikeGrid. Available at: www.staff.city.ac.uk/~jwo/cyclehire/bikegrid/ [accessed: 12 January 2015].
- Goodman, A. and Cheshire, J. 2014. Inequalities in the London Bicycle Sharing System Revisited: Impacts of Extending the Scheme to Poorer Areas but Then Doubling Prices. *Journal of Transport Geography*, 41(December), 272–9.
- Guenther, M.C. and Bradley, J.T. 2013. Journey Data Based Arrival Forecasting for Bicycle Hire Schemes. In: Dudin, A. and Turck, K.D. (eds), *Analytical and Stochastic Modeling Techniques and Applications*. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 214–31.
- Heesch, K.C., Sahlqvist, S. and Garrard, J. 2012. Gender Differences in Recreational and Transport Cycling: A Cross-Sectional Mixed-Methods Comparison of Cycling Patterns, Motivators, and Constraints. *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), 106.
- Hubway 2012. Hubway Data. Available at: <http://hubwaydatachallenge.org/> [accessed: 12 January 2015].
- Jurdak, R. 2013. The Impact of Cost and Network Topology on Urban Mobility: A Study of Public Bicycle Usage in 2 U.S. Cities. *PLOS ONE*, 8(11), p.e79396.
- Kaltenbrunner, A., Meza, R., Grivolla, J., et al. 2010. Urban Cycles and Mobility Patterns: Exploring and Predicting Trends in a Bicycle-Based Public Transport System. *Pervasive and Mobile Computing*, 6(4), 455–66.
- Lathia, N., Smith, C., Froehlich, J. and Capra, L. 2013. Individuals Among Commuters: Building Personalised Transport Information Services from Fare Collection Systems. *Pervasive and Mobile Computing*, 9(5), 643–64.

- Lathia, N., Ahmed, S. and Capra, L. 2012. Measuring the Impact of Opening the London Shared Bicycle Scheme to Casual Users. *Transportation Research Part C: Emerging Technologies*, 22(June), 88–102.
- Morency, C., Trépanier, M. and Agard, B. 2006. Analysing the Variability of Transit Users Behaviour with Smart Card Data. *IEEE Intelligent Transportation Systems Conference*, Toronto, Canada, 17–20 September 2006.
- O'Brien, O. 2015. Bike Share Map | Suprageography. Available at: <http://oobrien.com/bikesharemap/> [accessed: 12 January 2015].
- O'Brien, O., Cheshire, J. and Batty, M. 2014. Mining Bicycle Sharing Data for Generating Insights into Sustainable Transport Systems. *Journal of Transport Geography*, 34(2), 262–73.
- OBIS 2011. Optimising Bike Sharing in European Cities: A Handbook, *OBIS Handbook*.
- Ogilvie, F. and Goodman, A. 2012. Inequalities in Usage of a Public Bicycle Sharing Scheme: Socio-Demographic Predictors of Uptake and Usage of the London (UK) Cycle Hire Scheme. *Preventive Medicine*, 55(1), 40–45.
- Paris Vélib 2015. Paris Vélib scheme. Available at: <http://en.velib.paris.fr/Stations-in-Paris> [accessed: 12 January 2015].
- Pelletier, M.-P., Trépanier, M. and Morency, C. 2011. Smart Card Data Use in Public Transit: A Literature Review. *Transportation Research Part C: Emerging Technologies*, 19(4), 557–68.
- Shu, J., Chou, M., Liu, Q., et al. 2013. Models for Effective Deployment and Redistribution of Bicycles within Public Bicycle-Sharing Systems. *Operations Research*, 61(6), 1346–59.
- Transport for London 2015a. London Cycle Hire scheme. Available at: <https://web.barclayscyclehire.tfl.gov.uk/maps> [accessed: 12 January 2015].
- 2015b. Open Data Users. Available at: <http://www.tfl.gov.uk/businessandpartners/syndication/default.aspx> [accessed: 12 January 2015].
- Wood, J., Beecham, R. and Dykes, J. 2014. Moving Beyond Sequential Design: Reflections on a Rich Multi-Channel Approach to Data Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 2171–80.
- Wood, J., Slingsby, A. and Dykes, J. 2011. Visualizing the Dynamics of London's Bicycle-Hire Scheme. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 46(4), 239–51.
- Woodcock, J., Tainio, M., Cheshire, J., et al. 2014. Health Effects of the London Bicycle Sharing System: Health Impact Modelling Study. *BMJ: British Medical Journal*, 348(g425).
- Yoon, J.W., Pinelli, F. and Calabrese, F. 2012. *Cityride: A Predictive Bike Sharing Journey Advisor*. In 13th International Conference on Mobile Data Management (MDM), Bengaluru, India, 23–26 July 2012.

