bikedata

Mark Padgham[1] and Richard Ellison[2]

1 Department of Geoinformatics, University of Salzburg, Austria **2** Institute of Transport and Logistics Studies, The University of Sydney, Australia

Summary

The R package `bikedata` collates and facilitates access to arguably the world's largest open ongoing dataset on human mobility. All other comparable sources of data (such public transit data, or mobile phone data) are either not publicly available, or have been released only at single distinct times for single distinct purposes. Many public hire bicycle systems in the U.S.A., along with Santander Cycles in London, U.K., issue ongoing releases of their usage data, providing a unique source of data for analysing, visualising, and understanding human movement and urban environments (Austwick 2013; Borgnat et al. 2011; Padgham 2012). Such data provide an invaluable resource for urban planners, geographers, social and health scientists and policy makers, data visualisation specialists, and data-affine users of the systems themselves. The `bikedata` package aims to provide unified access to usage statistics from all public hire bicycle systems which provide data. These currently including Santander Cycles in London, U.K., and from the U.S.A., citibike in New York City NY, Divvy in Chicago IL, Capital Bikeshare in Washington DC, Hubway in Boston MA, Metro in Los Angeles LA, and Indego in Philadelphia PA. Additional systems will be added on an ongoing basis. The package facilitates the three necessary steps of (1) downloading data; (2) storing data in a readily accessible form (in this case in a single SQLite3 database); (3) extracting aggregate statistics. The two primary aggregate statistics are matrices of numbers of trips between all pairs of stations, and daily time series. Both forms of aggregation may be extracted for specific dates, times, or demographic characteristics of cyclists.

References

Austwick, Oliver AND Strano, Martin Zaltz AND O'Brien. 2013. "The Structure of Spatial Networks and Communities in Bicycle Sharing Systems." *PLoS ONE* 8 (9): e74685. doi:10.1371/journal.pone.0074685.

Borgnat, P., C. Robardet, J.-B. Rouquier, E. Fleury, P. Abry, and P. Flandrin. 2011. "Shared Bicycles in a City: A Signal Processing and Data Analysis Perspective." *Advances in Complex Systems* 14: 415–38. doi:10.1142/S0219525911002950.

Padgham, Mark. 2012. "Human Movement Is Both Diffusive and Directed." *PLoS ONE* 7 (5): e37754. doi:10.1371/journal.pone.0037754.