Tom Holub
Professor Daniel Rodriguez
Active Transportation, CP 216
12 December 2016

The Target: Improving statistical comparisons of cycling activity

Summary

Analysis of cycling activities is consistently hampered by problems with source data. Some of those problems are intractable, or at least expensive to fix, but some may be addressed by defining a consistent, limited extent. A target circle focused on the heaviest cycling area of the city may assist with geographic and longitudinal data comparisons.

I examined four cities, paired by demographic characteristics, and defined a 4-mile radius circle for each. I then gathered data from multiple sources to test comparability.

Even in the relatively consistent ACS survey data, quality issues hamper full analysis, and some data sources, such as those related to bikeway mileage and bike crashes, are quite poor. Still, this method holds promise in providing new insights as well as better benchmarks for future examination.

Problem statement

A persistent issue with analysis of cycling trends is the lack of adequate and comparable data. Addressing the adequacy of data is difficult and expensive, but improving the comparability of the existing data should be feasible. Geographic extent is a major contributor to problems of comparability, both longitudinally for an individual city, and geographically across different cities. Working on a project requiring comparisons between cities, I needed to devise a mechanism to improve the ability to identify similarities and differences.

The most reliable source of data on cycling rates is the “Means of transportation to work” question from the American Community Survey (US Census 2016). Because the ACS is regularly collected and aggregated, it is commonly used to compare cycling rates between cities. However, raw ACS numbers mask substantial problems with the meanings of the underlying data. They lack geographic comparability because of differences in land area and other physical characteristics of cities, and they lack longitudinal comparability because of changes in city extent over time.

Tom Holub
Professor Daniel Rodriguez
Active Transportation, CP 216
12 December 2016

ACS data is reported for the entire land area of each city. Austin, TX and Minneapolis, MN, two of the places I am investigating as cities with high cycling rates, differ in land area by a factor of six (Figs. 1&2). Austin's central city sees cycling rates comparable to Minneapolis', but its ACS data also includes sprawling suburbs which are excluded from Minneapolis' counts. This bias must be addressed when considering the relative success of the two cities in increasing cycling.

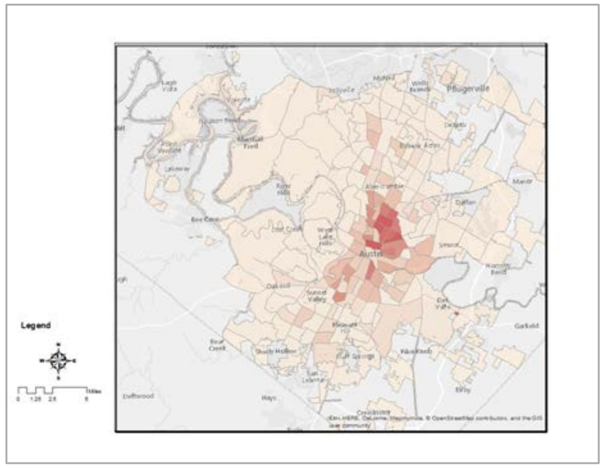

Figure 1: Austin, TX, 1:250,000 scale

Even within a single city, city extent issues confound analysis. Austin is a rapidly growing city which has nearly doubled in extent since 1990 (Fig. 3); longitudinal ACS mode share data report on a city which has changed substantially over time.

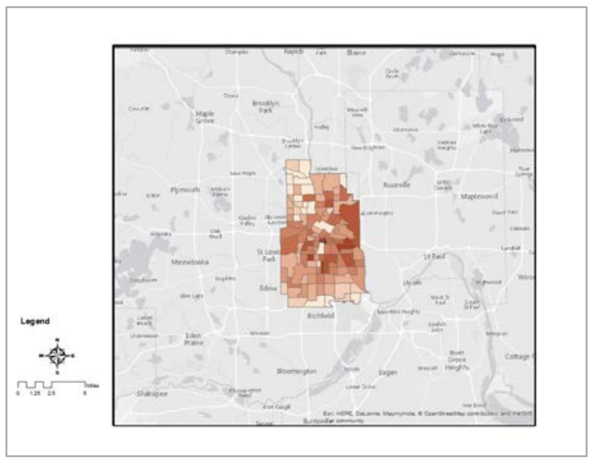

Figure 2: Minneapolis, MN. 1:250,000 scale

Choosing a consistent extent can assist with analysis of cycling data, in much the way that Allan Jacobs' square mile figure/ground maps assist with analysis of city design and morphology (Jacobs 1995). For the purposes of this study, I have chosen a circle of four miles in radius, which will be targeted at the areas of the city with the highest cycling rates. Targeting the circle manually provides a better basis for comparison than, for example, centering it on City Hall or the downtown district, because cycling rates are affected by numerous unique local conditions, notably topography and the presence of schools or universities. The ideal target could be determined computationally, but for most analytical purposes a rough estimate of the ideal placement should be sufficient.

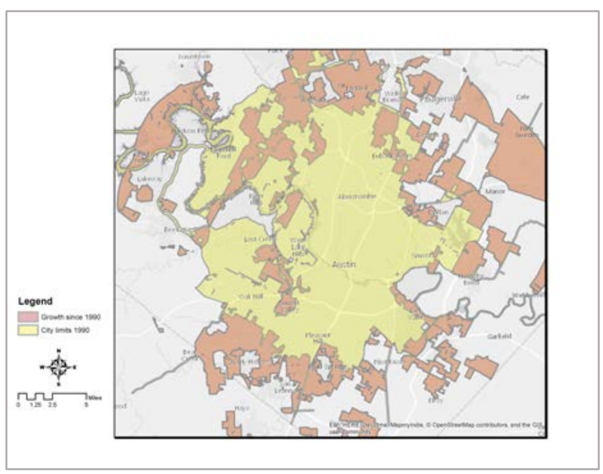

Figure 3: Austin TX extent in 1990 and 2016

Once the target is chosen, it provides for focused analysis of many different data elements, such as population and demography, cycling facilities, and crash data, in addition to mode share, and it provides better benchmarks to track cycling rates within a city over time.

The cities

The four cities under study were chosen in pairs, based on having broadly similar demographic characteristics with very different cycling rates. Austin is paired with Charlotte, NC, and Minneapolis with Columbus, OH. Targets for each city were chosen based on bike mode share data from the ACS 2014 5-year estimates. 5-year estimates chosen to improve sample sizes at the census tract level. Tracts which intersect the circle were used because much of the data available for analysis comes from the U.S. Census.

Due to differences in the way census tracts are laid out, and topological differences between the cities themselves, the target areas differ in extent, by nearly a factor of two (Minneapolis=48.83 mi^2, Austin=84.61 mi^2). Still, this is an improvement, but it may be worth considering other ways to make the areas more closely equivalent (see Limitations, below).

Population and demographic data were obtained from the U.S. Census, using the ACS 2014 5-year estimates for current data, and the 2000 census for historical data (Zellmer 2012). Bikeway mileage and crash data were obtained from sources unique to each city; these data may have issues with differing methodologies (see Limitations) . Still, the resultant maps appear much more comparable after targeting than before (Figs. 4-7).

Tom Holub
Professor Daniel Rodriguez
Active Transportation, CP 216
12 December 2016

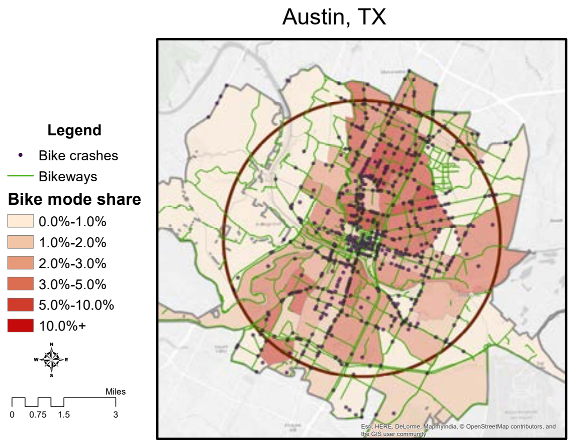

Figure 4: Austin, TX with bikeways and crash data

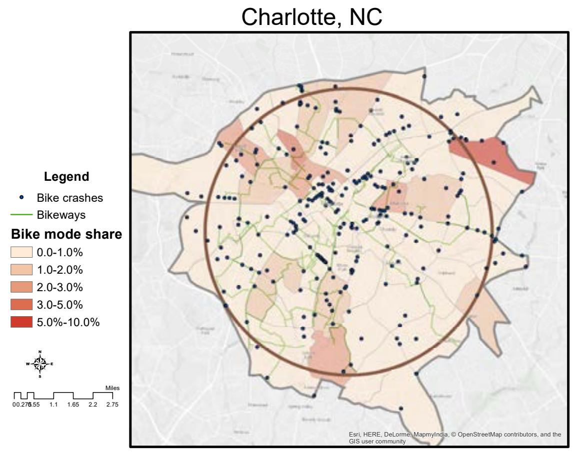

Figure 5: Charlotte, NC with bikeways and crash data

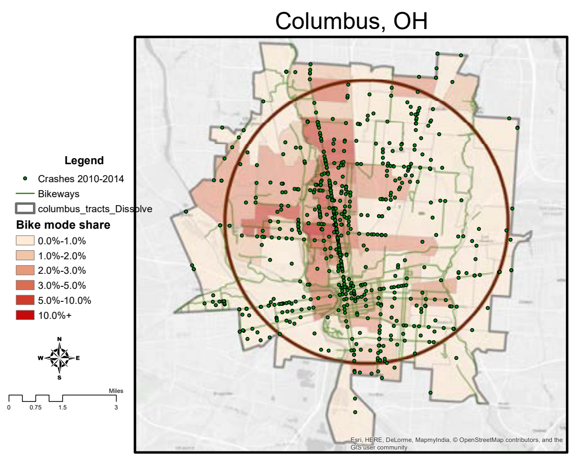

Figure 6: Columbus, OH with bikeways and crash data

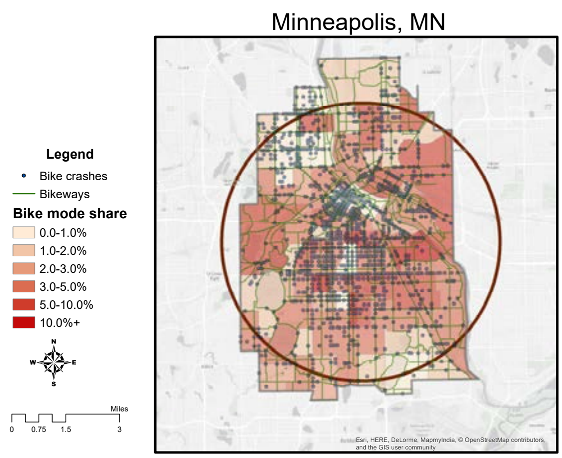

Figure 7: Minneapolis, MN with bikeway and crash data

The data

As expected, all four cities see higher bike mode share rates in the targeted area, but that difference is greater in the low-cycling cities, with Charlotte and Columbus showing increases above a factor of 3 relative to the city-wide estimates for 2014 (Fig. 8) (Donaldson 2014). This difference may be partly sampling error; even using 5-year estimates, ACS data has large error bars at the census tract level, which affects tracts with extremely low cycling rates more than tracts with much active cycling (Fig. 9).

Professor Daniel Rodriguez
Active Transportation, CP 216
12 December 2016

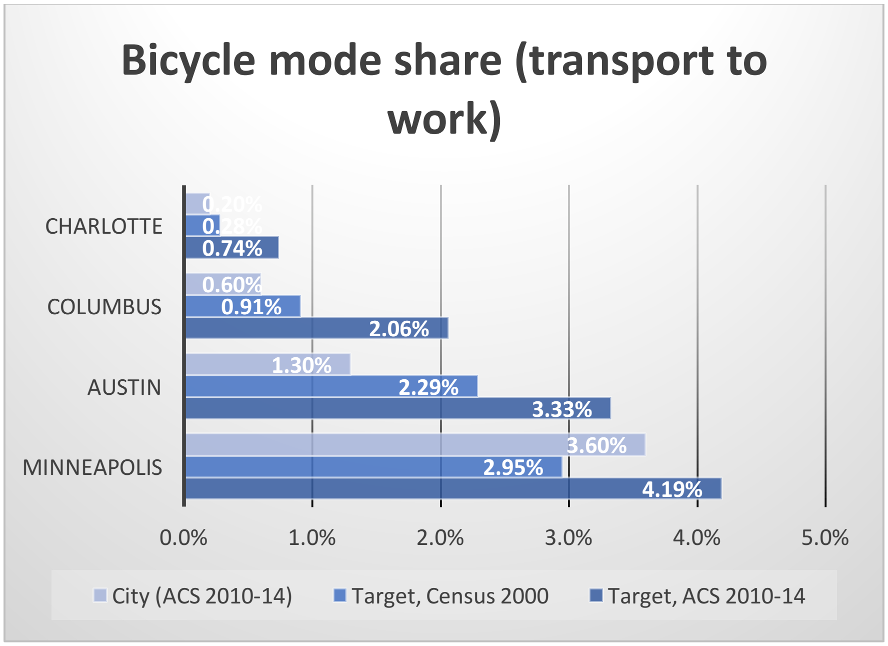

Figure 8: Bicycle mode share in selected cities

Large Margins of Error

	Population	Error	Non-Hispanic whites	Error	Commuters	Error	Cyclists	Error
Austin	304125	31601	153269	21071	159483	20091	5309	4063
Charlotte	318001	34222	134625	18558	147841	20408	1095	2042
Columbus	296188	32016	178017	24712	140238	20698	2883	2847
Minneapolis	365670	35228	227847	26688	198401	24186	8308	5567

Figure 9: Sum of margins of error for ACS 5-year estimates

Trend data shows less change in cycling rates in the target area of high-cycling cities than is been reported in the citywide data, which suggests that greater increase in mode share is occurring outside the center. In the 2000 Census, Austin's target area already showed a mode share of 2.3% (now 3.3%), while Minneapolis was already at 3.0% (now 4.2%). The low-cycling cities have progressed more, with Charlotte tripling its mode share in the target area since 2000, and Columbus doubling it (subject to the same caveat about sample sizes) (Fig. 8).

Bike crash data was obtained from a variety of city and state sources (UNCHSRC 2016; MTCLS 2016; ODOT 2016; Texas 2016). Crash rates per estimated cyclist were much higher in the low-cycling cities, with both Charlotte and Columbus having over three times the crash rates of Minneapolis. This may be due to the well-studied "safety in numbers" phenomenon (Elvik and Bjørnskau 2017), though the effect may be overstated due to methodological differences in data collection. The Minneapolis crash data includes only crashes involving cars (Fig. 10).

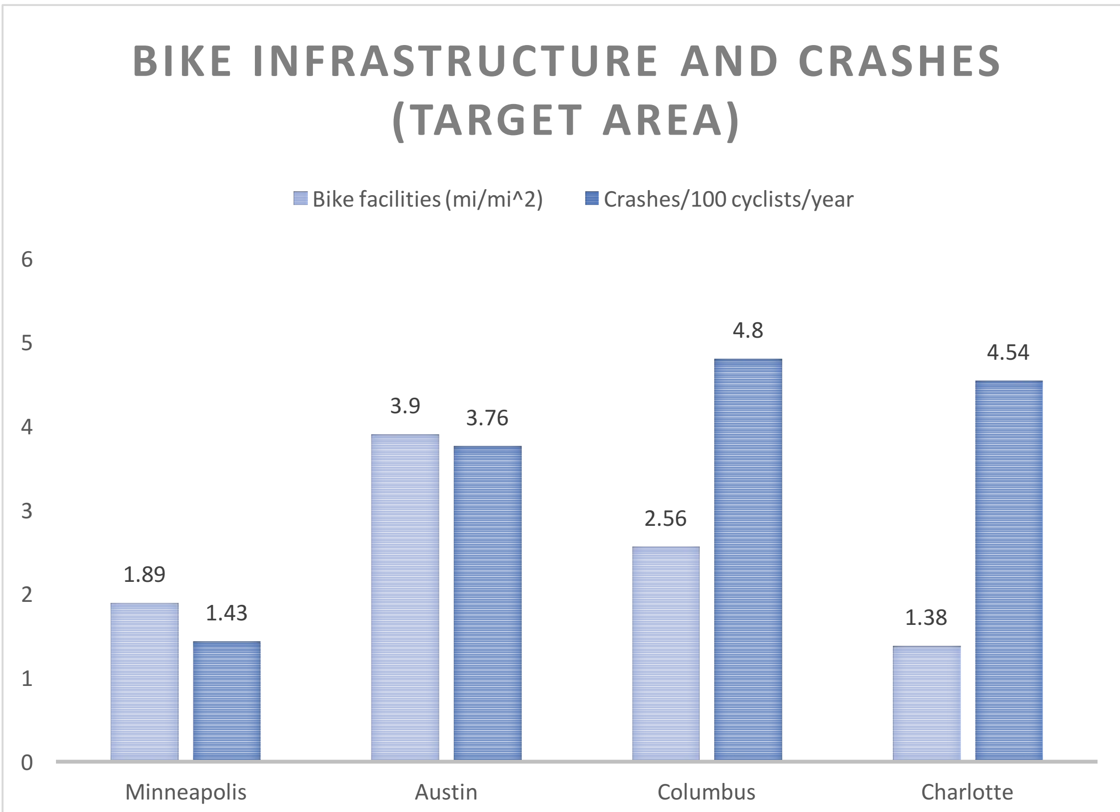

Figure 10: Bike infrastructure and crashes in target areas

Available bike facility data is of low quality. All four cities have GIS-based bike facility data, but given the lack of any clear standards on facility categorization, it is likely that the same facility may be counted differently in different locations, and furthermore, that much of the included bike facility mileage is itself of low quality (ODAustin 2016; CGIS 2016; MNDOT 2016; MOPRC 2016). Even after excluding wide curb lanes and wide shoulders from the Austin data, the city claims 330 miles of bikeway in the target area, more than double the linear miles/mi^2 claimed by Minneapolis (Fig. 11).

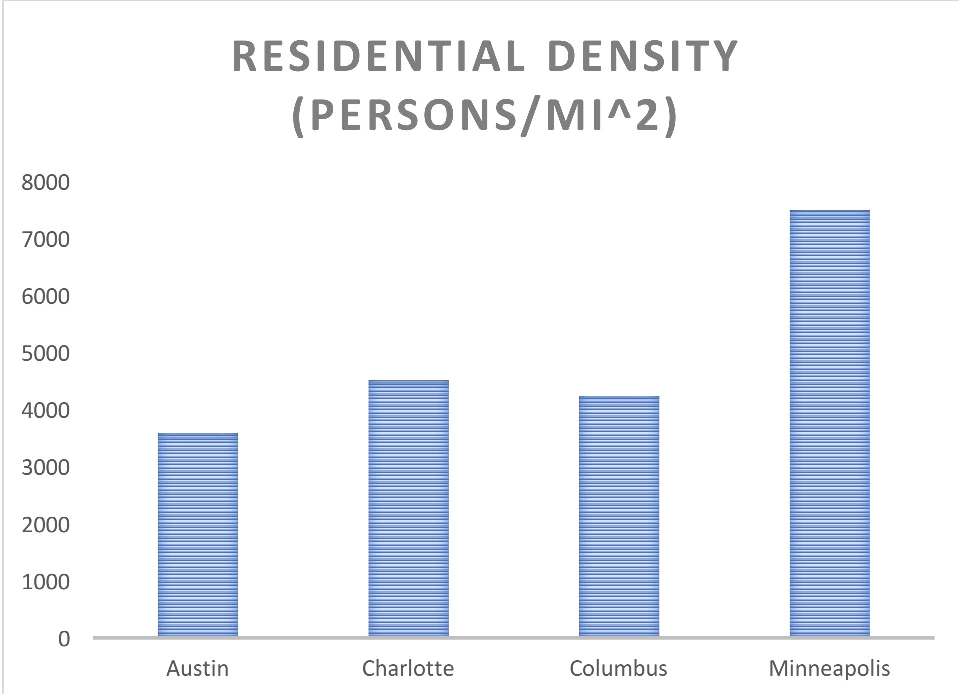

Figure 11: Residential density in target area

Minneapolis has the highest density of the cities, with almost 7500 persons per mi^2 in this sample, more than double Austin's target density (Fig. 10). This is partly due to the fact that the circle diameter is wider than Minneapolis' east-west city limits, so some lower-density areas are not included in the target. Still, this residential density in the central city probably contributes to Minneapolis' high cycling rates.

The target area showed relatively similar ethnic data for the city pairs, consistent with their overall demographic (which was part of the initial selection criteria) (Fig. 9)

Limitations

At the census tract level, bike mode choice data has large error bars. Summing Charlotte's 90 census tracts gives an estimate of 1095 cyclists, +-2042. The data used in aggregate does not have quite that much error, but without access to source data it cannot be recalculated. The small sample sizes probably overstate the cycling rates in tracts and regions where rates are low; tracts with zero reporting cyclists (which included 18 of 90 in Charlotte) report as 0+-12.

The shape of census tracts is a confounding factor; it may be better to increase the radius of the circle to 5mi., and include only those tracts which fall completely within the circle. This would make the land area included more similar between the cities.

Minneapolis is too small to fit a 4mi-radius circle. Areas within the target, but outside the city limits were not included in this analysis. Including these tracts would probably provide more accurate comparisons.

For bike facility mileage data to be useful, it must exclude facilities or segments which do not contribute to cycling safety or mode choice. Data including cyclist stress level, and specific network gaps and hazards within the target area would be more salient.

Crash data is inconsistent between cities and is likely to remain so unless a federal-level agency begins collecting crash data for non-fatal bike and pedestrian crashes. It may be useful for longitudinal analysis within a given city.

Conclusion

This method of improving longitudinal and geographic comparability of data shows some promise. Measuring the activity in the city center is both easier and more interesting than measuring the suburb and exurb activity. The finding that Minneapolis is significantly more population-dense than the other cities, even after controlling for city extent (partially), provides an interesting data point in chicken-and-egg question of whether cycle facilities or cyclists come first.

The method can only address certain aspects of the data challenges presented to bike advocates and planners; where there are significant uncertainties in the underlying data, as with crashes and facilities, those uncertainties will manifest in the targeted area as well.

Bibliography

CGIS. "Charlotte Mecklenburg GIS Online Services." GIS Online Services., accessed Dec 12, 2016, http://charmeck.org/mecklenburg/county/LUESA/GIS/Pages/OnlineServices.aspx.

Donaldson, Catherine. 2014. *Bicycling and Walking in the United States: 2014 Benchmarking Report* Alliance for Biking & Walking.

Elvik, Rune and Torkel Bjørnskau. 2017. "Special Issue Article: Cycling Safety: Safety-in-Numbers: A Systematic Review and Meta-Analysis of Evidence." *Safety Science* 92.

Jacobs, Allan B. 1995. *Great Streets* Cambridge, Mass. : MIT Press, 1995; 1st MIT Press pbk. ed.

MNDOT. "Bikeways - Minnesota Geospatial Commons." Minnesota Geospatial Commons., accessed Dec 12, 2016, https://gisdata.mn.gov/dataset/us-mn-state-metc-trans-bikeways.

MOPRC. "MORPC GIS Files." Mid-Ohio Regional Planning Commission., accessed

Dec 12, 2016, http://www.morpc.org/our-region/data-maps-tools/gis-files/index.

MTCLS. "Traffic Count Management System." Minneapolis Traffic Count and Crash Management System., accessed Dec 12, 2016, http://minneapolismn.gov/traffic/operations/WCMS1P-142538.

ODAustin. "Open Data | City of Austin Texas." Open Data Austin., accessed Dec 12, 2016, https://data.austintexas.gov/browse?limitTo=maps&utf8=%E2%9C%93.

ODOT. "ODOT GIS Crash Analysis Tool (GCAT)." Ohio Department of Transportation., accessed Dec 12, 2016, https://www.dot.state.oh.us/Divisions/Planning/ProgramManagement/HighwaySafety/HSIP/Pages/GCAT.aspx.

Texas, DOT. "Texas Crash Data Analysis and Statistics.", accessed Dec 12, 2016, http://www.txdot.gov/inside-txdot/division/traffic/crash-statistics.html.

UNCHSRC. "North Carolina Crash Data." North Carolina Crash Data., http://nccrashdata.hsrc.unc.edu/datatool.cfm.

US Census. 2016. *American Factfinder [Electronic Resource] : Factfinder.Census.Gov* Washington, D.C.] : U.S. Dept. of Commerce, Economics and Statistics Administration, U.S. Census Bureau, 2009].

Zellmer, L. R. 2012. *National Historical Geographic Information System (NHGIS)* American Library Association CHOICE.