Title: Using Big Data to Study Urban Sentiments: Twitter Data vs. Published Meeting Minutes

Presented by: Justin Hollander

Authors: Justin Hollander, Department of Urban and Environmental Policy and Planning, School of Arts and Sciences, Tufts University; Erin Graves, Federal Reserve Bank of Boston

Abstract:

A large body of sociological and psychological research has established that concentrated poverty is detrimental to children and families. A separate body of planning and economic research has established that urban economies in the United States have followed different trajectories, especially in the post-recession years . Yet there has been little work to understand how these two phenomena – concentrated urban poverty and metropolitan economic trends – intersect. In particular, we know very little about how the current context of decreasing urban economies is understood by the people who live in these locales and are affected by planning interventions.

We contribute to the larger literature around neighborhood change, specifically decline, by asking, what ways can we learn about residents' sentiments and civic concerns in the shrinking city context?

To pursue this research question, we use the case of the post-industrial city of New Bedford, Massachusetts. We collected data on the attitudes and opinions of residents of New Bedford by conducting 1) a conventional content analysis of published planning meeting minutes 2) a sentiment analysis of geo-coded Twitter messages in the city.

We reviewed over 300 meeting minutes for several committees in New Bedford from the years 2007-2013 to understand the deliberations of civically engaged New Bedford residents. A 25% sample is analyzed using conventional content-analysis techniques, with two researchers independently coding and then comparing their codes in order to improve inter-rater reliability.

We designed a database to collect data from the Twitter API. Our database stores, indexes and analyzes a continuous stream of data from the Twitter Decahose (a 10% sample of all Tweets) that are geo-tagged to specific locations. Most Twitter posts have little to nothing to do with perceptions of place. We filter the results based on a text search for key terms that pertain to New Bedford, their neighborhoods, or other specific places therein. Then, the ongoing sample of tweets is directly fed into SPSS Modeler to enable a sentiment analysis of how residents perceive their communities, using the same approach employed in analyzing the meeting minutes.