

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Travel and us: The impact of mode share on sentiment using geo-social media and GIS

Abstract:

Commute stress is a serious health problem that impacts nearly everyone. Considering that microblogged geo-locational information offers new insight into human attitudes, the present research examined the utility of geo-social media data for understanding how different active and inactive travel modes affect feelings of pleasure, or displeasure, in two major U.S. cities: Chicago, Illinois and Washington D.C. A popular approach was used to derive a sentiment index (pleasure or valence) for each travel Tweet. Methodologically, exploratory spatial data analysis (ESDA) and global and spatial regression models were used to examine the geography of all travel modes and factors affecting their valence. After adjusting for spatial error associated with socioeconomic, environmental, weather, and temporal factors, spatial autoregression models proved superior to the base global model. The results showed that water and pedestrian travel were universally associated with positive valences. Bicycling also favorably influenced valence, albeit only in D.C. A noteworthy finding was the negative influence temperature and humidity had on valence. The outcomes from this research should be considered when additional evidence is needed to elevate commuter sentiment values in practice and policy, especially in regards to active transportation.

Keywords: geo-social media, GIS, sentiment analysis, ESDA, travel behavior

1. Introduction

Commuting stress, an ever-growing social problem, continues to be studied by researchers in public health, planning, social sciences, engineering, economics, and business (Novaco and Gonzalez 2009). Daily commute mode choice decisions have resulted in a large portion of the human population being affected by stress (Legrain, Eluru, and El-Geneidy 2015; Novaco and Gonzalez 2009). Kahnemann et al., (2004) found that commuting was among the lowest ranked factors contributing to net affect (i.e., experienced utility). This has consequences on mental and physical health, including: increased heart rate, blood pressure, back problems, as well as certain types of cancer (Hoehner et al. 2012). Commute stress also affects mode-choice, especially

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

sustainable travel modes (i.e., walking and bicycling) (Legrain, Eluru, and El-Geneidy 2015), which is concerning because these modalities have been linked to elevated satisfaction levels and improved public health outcomes (Olafson 2014; Furie and Desai 2012). Measuring the links between stress and transportation is a growing area of research due to past findings that a person's travel experience is a strong determinant of future behavior and well-being (Eagly 1994; Mokhtarian 2001). Although a multitude of factors have been investigated, research gaps remain (De Vos et al. 2013).

Measuring trip utility and satisfaction is traditionally determined by summarizing human responses to questionnaires, counters, interviews, focus groups, observations, interviews, photography, cognitive maps, and hypothetical scenarios (Berg 2004; Latham 2003). These methods are flawed in large part because they do not measure a traveler's true experience, are impractical, untimely, costly, and biased (Collins, Hasan, and Ukkusuri 2013; Neutens 2011; Luo et al. 2016; Salon 2016; Morris and Guerra 2015). In addition, the spatial granularity of the data is coarse and typically aggregated to government units such as U.S. Census Tracts, thereby minimizing the impact of the results at a local scale (Sun 2017). With the rise of sensor technology, new opportunities to study human-scaled mobility patterns and behavior are available.

Information Communication Technology (ICT), which includes smart phones and the internet, has witnessed a spectacular increase in societal usage, in large part due to the attractiveness of social networking. Location-aware devices are able to collect semantic information with geo-locational attributes. A popular type of ICT that has attracted researchers interested in measuring travel behavior is geo-social media (GSM) (Andrienko et al. 2013). The

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

proliferation is marked by the ubiquitous penetration of mobile devices and accessibility of shared content, which has resulted in new insights and methods for studying the link between well-being and travel. The main advantages of GSM data include its human-scaled resolution, low cost, precise time-stamp, and comments which allows for individualized insight into human-environmental behavior. The current research attempts to investigate GSM data using GIS and contribute to the scant amount of work on spatially exploring human-locational data obtained from ICT and provide nuanced insight into the interrelations among travel mode, sentiment, socioeconomic status (SES), time, and environment (Luo et al. 2016; Lichman and Smyth 2014). The results of this study will assist planners and policy-makers make informed decisions on increasing travel satisfaction, especially for active modes of transportation such as bicycling, walking, and mass-transit.

2. Objectives

The current research set out to examine the relationship between valence (i.e., sentiment) and travel mode using GSM. In doing so, it had two objectives: a) visualize the spatial characteristics of active (walking, bicycling, and mass-transit) and non-active (automobile and water travel) travel mode densities using ESDA and geographic information systems (GIS) in two major metropolitan areas: Chicago and Washington, District of Columbia; b) measure the associations between transportation mode and valence using a global and spatial model, while adjusting for neighborhood character, time, and weather. We believe this study contributes to two broader research areas: the visualization of GSM derived human-scaled travel patterns, and understanding the links between travel mode and sentiment. The theoretical framework used in

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

this research is based on the ecological modeling concept, which states that a combination of environmental and psychosocial variables can aptly explain physical activity, including transportation (Sallis 2002).

The remainder of the paper is structured as follows. Section 3 highlights previous works and section 4 describes the study area and outlines the data sources. Section 5 introduces the methodology leveraged in this research, including Twitter sentiment analysis, GIS, and modeling protocols. The paper concludes with describing the results and conclusion, in sections 6 and 7, respectively.

3. Related Work

3.1 Measuring travel behavior

Considerable research has now established that travel has a positive utility and there are clearly links among travel mode, happiness, and subjective well-being (Reardon and Abdallah 2013; Morris and Guerra 2015; Mokhtarian 2001). The impetus for this work is derived in part from the field of social psychology, which states that behavioral choices are attributed to internal factors (Van Acker, Van Wee, and Witlox 2010; Mokhtarian 2001). This evidence has resultantly exposed a problem with traditional travel-demand models: travel isn't necessarily a cost and may be a sought after event in its own right (Morris and Guerra 2015). Much of the past research on well-being and travel behavior remain unclear, however. For instance, while the reduced health outcomes of commuting are well-documented (Novaco and Gonzalez 2009), other works have found benefits, especially concerning mental health (Redmond and Mokhtarian 2001). Olsson et al., (2013) found that people in three of the largest urban areas of Sweden were generally happy

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

while commuting. The link between well-being and mode-shift has also received significant attention in the literature (Abou-Zeid et al. 2012). Considering that mass-transit is generally the least satisfying transportation mode (Ory et al. 2004), and bicycling and walking the most satisfying, (Morris and Guerra 2015), additional information on why, when, and where travelers are most satisfied may be essential in fully understanding an efficient means to promote active travel.

Traditionally, assessing travel behavior has been through the use of bottom-up approaches which engage participants directly. Questions such as: “how safe was your journey” and “how does it (i.e., travel) affect your overall life quality” are typical requests (Abou Zeid 2009; Zhou and Zhang 2016). The use of qualitative survey instruments are the predominant tool used to collect this information. For instance, Schafer (2000) distributed 30 surveys to ascertain the spatiotemporal patterning of human travel, and St-Louis (2014) disseminated a large survey to faculty, staff, and students at McGill University with the goal of examining travel mode happiness. The downfall with these methods are the high cost, labor intensity, limited sample size, and their inability to adequately decipher complicated human behavior (Steiger et al. 2015; Luo et al. 2016). Geo-located mobile communications have been looked at as new way to collect more accurate responses regarding travel behavior, with the added benefit of locational attributes.

3.2 Geo-social media analysis

Location-based-social networks (LBSN) and GSM data have brought new insights into geographical science, spatiotemporal analysis of humans, and social interactions (Giannotti and

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Pedreschi 2008; Luo et al. 2016; Sui and Goodchild 2011). Platforms such as Twitter, Facebook, and Foursquare are the primary sources of GSM data (Huang and Wong 2016; Luo et al. 2016). Twitter, in particular, is a microblogging application that records text (280 characters or less), time, and geospatial coordinates: allowing for linkages among place, time, and semantics (Tsukayama 2017). The majority of previous travel behavior research using GSM data has been dominated by investigations on human mobility (Cheng et al. 2011; Huang and Wong 2015), activity-social patterns (Steiger et al. 2015; Huang and Wong 2016), and physical activity levels in neighborhood contexts (Nguyen et al. 2016). For instance, Schweitzer (2014) examined how transit agencies and patrons engage with Twitter as a means of communicating service quality and Collins et al., (2013) examined the utility of Twitter for aggregating transit rider opinions. While Twitter doesn't dynamically track movement, it can connect with Foursquare check-ins to link activity space (i.e., geography), sentiment, and transportation (Noulas et al. 2012; Wu et al. 2014; Mondschein 2015). For example, Mitchell, et al., (2013) used check-ins to investigate happiness levels obtained from Twitter throughout the U.S. and Esmine et al., (2012) investigated sports-related sentiments using a similar approach. Quantifying the psychology of "place" may show promise for understanding the probability of adopting new behaviors. Learning takes place in social context, where behaviors are adopted based on how it's performed, perceived, and expected outcomes (Bandura 1977; Nguyen et al. 2016). The current research infers that combining social media sources like Twitter and Foursquare with geography could help analyze attitudes towards differing travel modes, and circumvent the pitfalls of traditional methods. One of the first steps in doing so requires the extraction of the implied opinions from social media.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Sentiment analysis offers researchers the opportunity to link the opinions of users to geography (Mitchell et al. 2013). The task of mining data for sentiments is arduous due to the short length and irregular structure of the user-produced content (Saif, He, and Alani 2012). Resultantly, many methods have been devised and applied in several research areas such as: predicting daily box office revenues (Rui and Whinston 2011) or assessing short-term stock market performance (Bollen, Mao, and Zeng 2011). The sentiment analysis methods used have been both lexicon-based approaches and machine-learning algorithms. Das and Chen (2001) utilized a lexicon-based classification algorithm to extract market emotions from stock message boards, which was further used for decision on whether to buy or sell a stock. Lei et al., (2014) constructed a word list or dictionary for emotion detection. Similarly, Turney and Littman (2003) utilize an unsupervised learning algorithm to classify emotional content of users' reviews of movies, travel destinations, automobiles and banks. A recent machine-based classification algorithm that has been proven useful for analyzing Tweets is LabMT. The method assigns happiness values based on 10,000 of the most popular English words (Frank et al. 2013). Other works which use linguistic structures or a dictionary of words to classify the emotions of words or phrases include SemEval, which introduced the task of "Affective Text" (Strapparava and Mihalcea 2007), SWAT (Katz, Singleton, and Wicentowski 2007), Subjectivity Wordlist (Mihalcea, Banea, and Wiebe 2007), WordNet-Affect (Strapparava and Valitutti 2004), and SentiWordNet (Baccianella, Esuli, and Sebastiani 2010). Despite the limitations of semantic ambiguity, we used a Lexicon based approach in this research due to its simplicity of deployment and successful implementation in previous research (Kim et al. 2017).

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

4. Study Area and Data

4.1 Study area

The study areas in this research were Chicago, Illinois and Washington, District of Columbia (D.C.), USA (figure 1). The locations were chosen because of the large number of Foursquare check-ins and differences in SES, climatic, intermodal transport options, walkability, bicycle-friendliness, and density of recreational opportunities such as the National Mall in D.C. and Lake Michigan in Chicago. Both cities contain multimodal transportation systems and are highly ranked nationally for their bicycle “friendliness.” *Bicycling Magazine* recently ranked D.C. seventh in the nation and Chicago first (Dille 2016). Table 1 highlights additional commonalities and differences in regards to several SES indicators in each city.

Table 1. General statistics for each study city (2010)

	Chicago	Washington D.C.
City Area	227.6 mi ²	68.3 mi ²
Population	2,712,608	633,736
Population Over 18	2,102,746	500,908
Median Age	33.4	33.7
Median Income	\$47,831	\$69,235
High school Diploma or Higher	81.6%	88.9%
Below Poverty Level	22.7%	18.2%
White	48.4%	38.5%
African American	31.9%	50.7%
American Indian	0.3%	0.3%
Asian	5.7%	3.5%
Native Hawaiian/Pacific Islander	0.0%	0.1%
Other	11.5%	4.1%

Source: U.S. Census Bureau

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>



Figure 1. Study areas and spatial distribution of Tweet check-ins

4.2 Twitter data

The first step in this research was to retrieve GSM data from the Twitter and Foursquare API (application programming interface). We collected Tweets from April 2012 through June 2013. Foursquare was chosen because it is the most popular LBSN application and has surpassed 10 billion check-ins as of December 12, 2015 (Smith 2016). The Foursquare check-ins function as “sensors” revealing the geography of human dynamics at different times of the day (Banerjee et al. 2013). The check-ins occur at the business locations through the Foursquare application using a person’s mobile device (i.e., smartphone) and were retrieved from the broad “Travel and

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, Journal of Location Based Services, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Transport” category. In particular, we tracked consumers checking-in at transport locations such as: bus stations, train stations, hotels, piers, among many other transport related locales as these are major transportation destinations and origins within urban areas (table 2). It should be noted that although there may be different reactions to differing mass-transit modes (e.g. bus versus subway), previous research has found that it is generally the least desirable travel mode (Schweitzer 2014), hence we treated all transit check-ins as one mode in this study. The final transportation related Tweets are depicted in table 2 and their frequencies are illustrated in figure 2.

Table 2. Twitter filtering stages and results

Broad category	Specific category	Final travel mode classification
Travel and Transport	Boat/Ferry, Bus Station, General Travel, Light Rail, Pier, Rental Car, Road, Subway, Train Station, and Bike Rental/Bike Share	Mass transit, Bike, Pedestrian, Water Travel, Auto

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

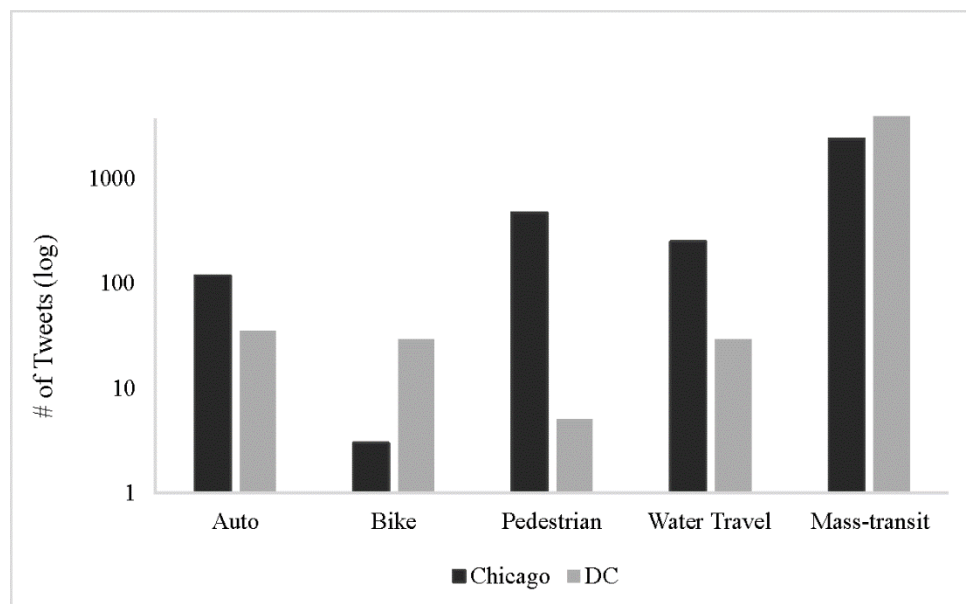


Figure 2: Travel mode Tweet frequencies per city

4.3 Socioeconomic, environmental, and weather variables

A large body of research has established connections among SES, environment, and travel behavior (Cervero and Kockelman 1997; Saelens, Sallis, and Frank 2003). To remain in-line with this work, several related factors were included in this research. We collected data from the 2010 U.S. Census Bureau and the Environmental Protection Agency's (EPA) Smart Location database (SLD). The variables included in this dataset are comprised of 90 different SES and environmental correlates aggregated to the Census block group (CBG) level for the entire nation (Ramsey 2014). The factors were measured and categorized in accordance with known travel behavior and quality of life metrics such as: development density, land-use diversity, network design, accessibility, demographics, and employment. We considered all of these in this research.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Weather is a strong predictor of active travel behavior and air quality is an ever-growing global public health concern (Bocker, Dijst, and Prillwitz 2012; Rybarczyk and Gallagher 2014). Therefore, daily weather and air quality data was collected from the U.S. Environmental Protection Agency's Air Quality System (AQS) (<https://www3.epa.gov/airdata/index.html>) and a commercial entity: Weather Underground (<https://www.wunderground.com/>). The variables integrated into this research included: temperature, precipitation, wind speed, humidity, and air quality (i.e., Ozone). We downloaded air quality data (i.e., AQS) for Chicago and Washington D.C. originally collected from seven and six monitor stations, respectively. Additionally, we gathered precipitation, temperature, humidity, and wind speed from 19 Weather Underground stations in Chicago and 13 in Washington, D.C.

Variations in time can affect travel behavior and mobility patterns due to workplace and life circumstances (Yuan, Raubal, and Liu 2012; Järv, Ahas, and Witlox 2014). To explore how time impacted valence using Tweets, we used the date and time fields from this database for further categorization and analysis. The groups consisted of: hour ranges, am/pm, day, month, and season and were aggregated to ease interpretation of the relations between travel mode, valence, and time.

5. Methods: A multistage procedure

A series of sequential data analysis steps were required in the current analysis. Three main tasks outlined in figure 4 were: 1) Mining and collecting the GSM posts for sentiment analysis, 2) geolocating Foursquare check-ins and merging them to neighborhood level (i.e., CBG) factors using GIS, 3) explore the travel mode trends spatially using ESDA and implementing two types of regression models. The forthcoming paragraphs outline the steps involved in each task.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

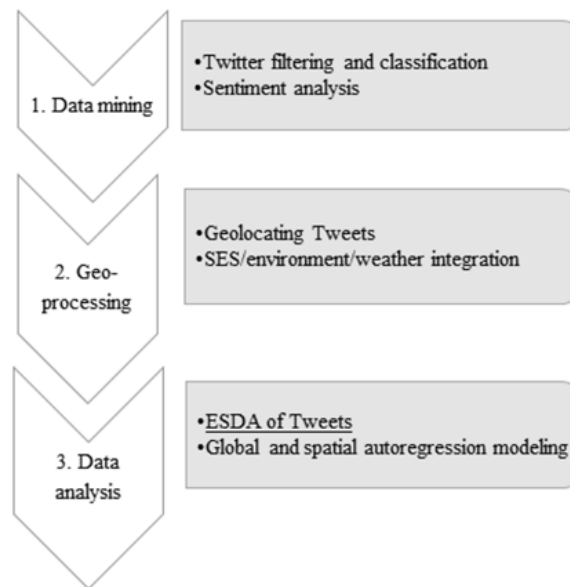


Figure 3: The methodological flowchart

5.1 Data mining

We mined Twitter data from consumers who checked in to different Foursquare venues and then tweeted via the Foursquare application, in the two study cities. We invoked this approach to reduce the typical “noisiness” of Big Data, obtain a sufficient sample, and reveal important spatial patterns, and associations to place. In order to illustrate how users were situated within specific travel modes, we will explain the data sharing process. When users check into a Foursquare venue and tweet, the exact latitude-longitude of the venue is passed on to Twitter which is then accessible via their application programming interface. We collected tweets from the API along with location details of the venues. The details captured included address, establishment

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

categories such as a house, coffee shop, or movie theater. We collected the Tweet ID, Tweet text, time, and user ID. We accumulated Twitter check-in data (via Foursquare) within a 2.5 mile buffer around major locations such as Union Station in Chicago and D.C.(see figure 1).

We considered three main types of Foursquare venues in this research. First, fixed establishments such as a bus station or train station were assessed. The second were transportation related land-uses such as roads or piers (i.e., Lake Shore Drive in Chicago) and the third type were broad categories like “general travel” (table 3). While the first was sufficient to indicate the travel mode accessed by a user, information from second and third categories were combined with tweet content to identify the travel mode used or considered. In other words, a person who checked in on an expressway often revealed they were in a taxi heading to the airport. A person checking in on the Navy Pier revealed they were having a nice walk. Accordingly, users were assigned to auto, bike, mass-transit, pedestrian and water travel. In this research, approximately 5% of the Tweets collected in the area surrounding the reported location were using Foursquare check-ins. While this statistic is low, our sample included enough Tweets with valid geographic information and syntax to conduct statistical and spatial analysis. The final quality assurance step consisted of manually verifying the Tweets to ensure that the check-ins correctly represented the travel mode. The process entailed the comparison of the check-in travel groups with the Tweet content, as well a visual comparison against the most recent geospatial imagery and land use data in a GIS. Approximately 1% percent of the Tweets were removed due to errors. We amassed a total sample size of 4,399 and 3,988 for Chicago and D.C., respectively. To give an idea about how Foursquare

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, Journal of Location Based Services, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Venue categories converted into travel modes: The percentage check-ins from each Venue type that were classified as different travel mode are displayed in table 3.

Table 3. Venue type and travel-related Foursquare check-ins

Venue	Auto	Bike	Mass-transit	Ped	Water Travel	Total
Bike Rental / Bike Share (only DC)	0.00	83.87	0.00	0.20	0.00	0.32
Boat or Ferry	0.00	0.00	0.00	0.00	100.00	3.80
Bus Station	0.00	0.00	12.03	0.00	0.00	10.47
General Travel	9.20	9.68	0.10	0.00	0.00	0.39
Light Rail	0.00	0.00	0.33	0.00	0.00	0.29
Pier (Only Chicago)	0.00	0.00	0.00	91.19	0.00	5.31
Rental Car (Only DC)	2.40	0.00	0.00	0.00	0.00	0.07
Road	88.40	0.00	0.00	6.56	0.00	3.02
Subway	0.00	3.23	35.06	1.64	0.00	30.61
Train Station	0.00	3.23	52.49	0.41	0.00	45.71

As we explained above, venues that indicated establishments, such as Boat or Ferry docks, bus, train stations, light rails and subway universally indicated one form of travel mode (either water travel, or mass-transit). However, the Pier, general travel and road check-ins could be auto, bike, mass-transit, pedestrian and water travel. For example, 87% of the people who checked in on the “road” were traveling in an auto and 13% were pedestrians. These travel modes were verified/determined from the tweet content using the above-mentioned methodology.

The current research used the popular lexicon based sentiment analysis algorithm entitled Affective Norms from English Words (ANEW) developed by Bradley et al., (1999). The method uses a scale to classify all Tweets and derive an overall measure of pleasure score for each check-

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

in. The ANEW scale is based on the assumption that some of the most common words in the English language elicit different levels of pleasure, arousal and dominance and the scale rates those words with a specific level of emotion. Using a list of 2,476 words, individuals responded to them using a self-assessment manikin (i.e., a non-verbal pictorial) that relates to measures of pleasure, arousal, and dominance. The process resulted in a numerical index for each Tweet based on the sum of the keyword scores. Tweet pre-processing was not conducted prior to sentiment analysis. Among the sentiments relevant to effects that the environment has on human emotion and behavior, there are three as suggested by Mehrabein and Russell (1974) and Turley and Milliman (2000): pleasure, arousal and dominance. Pleasure is akin to valence, and is similar to the motion of happiness experienced while immersed in a physical environment. We utilized valence in the current research to assess travel mode sentiment (i.e. happiness). Table 4 lists the valence summary statistics for all travel modes in each city.

Table 4. Travel mode valence descriptive statistics

Variable	Max	Mean	SD
<i>Chicago</i>			
Auto	39.77	5.10	6.52
Mass-transit	62.01	6.69	6.69
Water travel	43.18	13.53	9.36
Pedestrian	47.85	10.35	6.36
Bicycling	13.11	8.51	3.98
<i>Washington D.C.</i>			
Auto	35.23	7.58	8.89
Mass-transit	43.88	6.71	6.70
Water travel	33.55	10.78	9.46
Pedestrian	32.77	10.29	11.95
Bicycling	26.55	6.73	7.29

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

5.2 Geoprocessing and data preparation

From the Foursquare latitude and longitude coordinates of each Tweet, we first created a XY event layer in GIS and then exported them to a point shapefile using ESRI (Environmental Systems Research Institute) ArcGIS software version 10.4. A point-in-polygon geoprocessing method was used to intersect the Foursquare check-in and associated Tweet content with the U.S. Census Block Group (CBG) polygons. By merging these two spatial databases, we could then examine and account for underlying neighborhood conditions. The GIS procedure is commonly used when area scale data needs to be affixed to point scale data (Cromley 2012). Approximately 99% of the check-ins were successfully mapped and joined to the CBG's. Those which were spatially misplaced or erroneous were removed from the analysis. The previously discussed weather factors were joined to the individual Tweets using an attribute join procedure in GIS.

During all modeling efforts, we considered valence as the dependent variable. Prior to model development we transformed this factor using the square root in order to bring it into normality. The primary independent variable in each model was travel mode. Dummy coding was applied and the automobile travel mode was the reference category. A correlation matrix was examined for statistical correlations among the dependent variable and all independent variables (i.e., SES, environment, weather, and time). Factors that were statistically significant ($p < 0.05$) and uncorrelated, using the variance inflation factor (VIF), were ultimately retained for model inclusion. A (VIF) of less than 10 was considered the threshold as values greater than this violates the parametric test of assumption of covariate independence (Shaker 2016). The same independent

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

variables for each city were used during model construction. The final independent variables are displayed in table 5.

Table 5. All model variables with definitions

Variable	Description
Dep. variable: Valence (sq rt)	Unpleasant-Pleasant
<i>Travel modes</i>	
Auto	Dummy: yes = 1, no = 0 (reference)
Boat	Dummy: yes = 1, no = 0
Mass-transit	Dummy: yes = 1, no = 0
Pedestrian	Dummy: yes = 1, no = 0
Water travel	Dummy: yes = 1, no = 0
<i>Temporal, Socioeconomic, Environment, Weather</i>	
Morning	6:00 - 11:59, Dummy: yes = 1, no = 0
Night	20:01 - 5:59, Dummy: yes = 1, no = 0
Population density	Gross residential density (Housing units/acre) on unprotected land
2-car households	Quantity of 2-car households per CBG
Land-use mix	8-tier employment sector
Street intersections	Auto-orientated street intersection density
Humidity	Ave % atmospheric water/day
Temperature	Ave temperature per day

5.3 Data analysis: spatial distribution of travel mode Tweets

To understand the spatial patterns of travel mode check-ins, we implemented a common ESDA technique, kernel density estimation (KDE) using ArcGIS software (see objective “a”). The KDE method approximates the intensity of points per unit area from the creation of a smoothed surface using a kernel function and their distance from each point (Bailey 1995). The inputs are the x, y, z coordinates and outputs a raster model where the cells represent the density of events, often labeled as “hot-spots.” The choice of the kernel has found to be not as important as the

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

bandwidth (i.e., window width) in assessing the contribution of near-by points (Downs 2010). The KDE method has been suggested and used in several past studies visualizing the spatial trends of GSM data (Lichman and Smyth 2014; Han, Tsou, and Clarke 2015; Li, Goodchild, and Xu 2013; Wu et al. 2018; Gerber 2014; Sun et al. 2016; Zhen et al. 2017; Widener and Li 2014; Poorthuis et al. 2016). To remain in-line with these works, KDE was used to transform travel mode check-ins to a continuous surface density. We selected a bandwidth of one kilometer, after multiple rounds of testing other radii, and a cell size of 30 meters. Although several kernel functions are available, the quadratic kernel function, as described in Silverman (1986), was used. The result is a raster surface depicting the spatial distribution of travel mode check-in densities in each city.

5.4 Global and spatial regression model development

The inferential statistical analyses of this study followed three steps: global multiple regression, spatial multiple regression, and model accuracy assessment using a common autocorrelation index, Moran's *I*. Two global ordinary least squares regression models (OLS) were first constructed to explain the relationship between the five travel modes (i.e., auto, walking, bicycling, mass-transit, and water travel) and valence using SPSS (IBM Inc.) version 22 software. All explanatory variables (see table 5) were entered simultaneously into each model. The standardized residuals from both models were tested for autocorrelation using Global Moran's *I*. This procedure was enacted because if spatial autocorrelation is present, the independence of observations assumption is violated (Wagner and Fortin 2005). The method produces an index ranging from -1 to +1. Statistically significant positive values indicate clustering and negative values indicate dispersion (Burt 2009).

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

A simultaneous autoregressive regression (SAR) modeling methodology was employed to correct for spatial dependence among the response and predictor variables. This inferential statistical technique was chosen over other autoregressive methods, such as conditional autoregressive models, primarily because it has been proven effective in other transport related studies (Rybarczyk 2012). The SAR method is a spatial modeling technique that uses a variance-covariance matrix based on the non-independence of spatial observations (Kissling and Carl 2008). The model address spatial autocorrelation by estimating how much the response or predictor variable at any one site reflects the response or predictor values at surrounding sites; this is achieved by adding a distance-weighted function of neighboring response values to the model's explanatory variables (Dormann 2007). The reader is directed to Anselin (2013) for further details on the SAR model. The publically available software Spatial Analysis in Macroecology (SAM), version 4 was used for model development (Rangel 2010).

6. Results and discussion

The spatial distribution of travel mode densities were examined using KDE and GIS to visualize trends in each city first. This section is followed up by the results of the two independent multivariate global and spatial regression models. The global models summarized the directionality and degree of influence of travel mode on valence, while adjusting for SES, time, environment, and weather; the SAR models analyzed the same relationships, while accounting for spatial autocorrelation.

6.1 Travel mode visualizations

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

The initial objective (“a”) in this research was to visualize the spatiality of travel mode densities using KDE, where the darker hues represent hot-spots of travel mode check-ins for each city. The highest density of automobile modes were found near recreational and natural resources (i.e., parks) in Chicago (figure 4a); and adjacent to waterways and the National Mall in D.C. (figure 5a). Expectantly, the largest concentration of mass-transit mode check-ins stood out near the main transit hubs in Chicago and Washington D.C. (figure 4b and 5b). Figures 4c and 5c depict the density of water travel modes in Chicago and D.C., respectively. The clusters were found near waterways and recreational land-uses. The densities of active transportation modes, walking and bicycling, in Chicago are shown in figures 4d and 4e. Pedestrian activity was highly concentrated near Navy Pier, which is logical considering that this area is largely designed for walking. At the same time, bicyclist activity was mainly situated downtown and adjacent to Millennium Park. We found that a hot-spot of pedestrian activity was linearly diffused near the National Mall in D.C. (figures 5d). Figure 5e indicates a minute, however significant, hot-spot of bicycling activity in a neighborhood northeast of Union Station. This area is adjacent to DuPont Circle, an area containing many bicycle and pedestrian facilities. Overall, the visualizations depict nuanced travel mode activities in each city with high spatial granularity.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

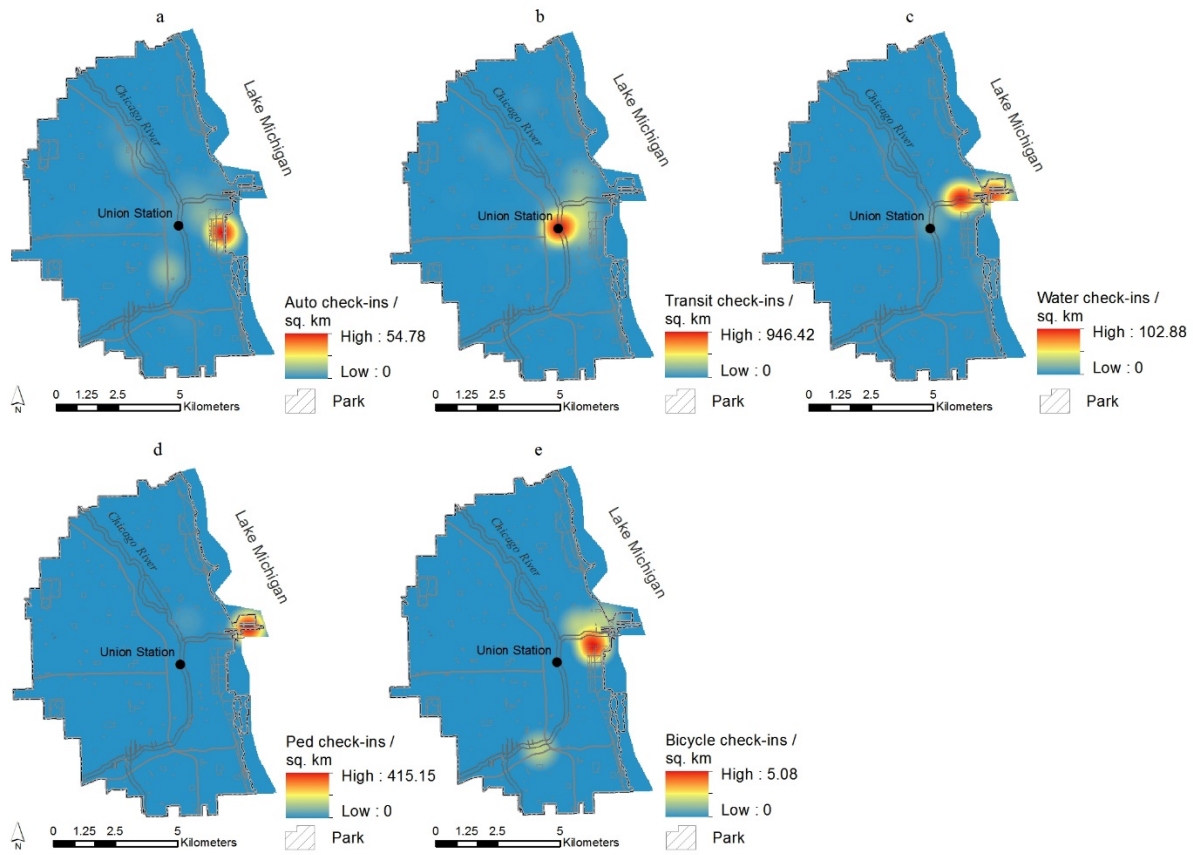


Figure 4: Kernel density estimation of Chicago travel mode check-ins for a) automobile, b) mass-transit, c) water, d) walking, e) bicycle

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

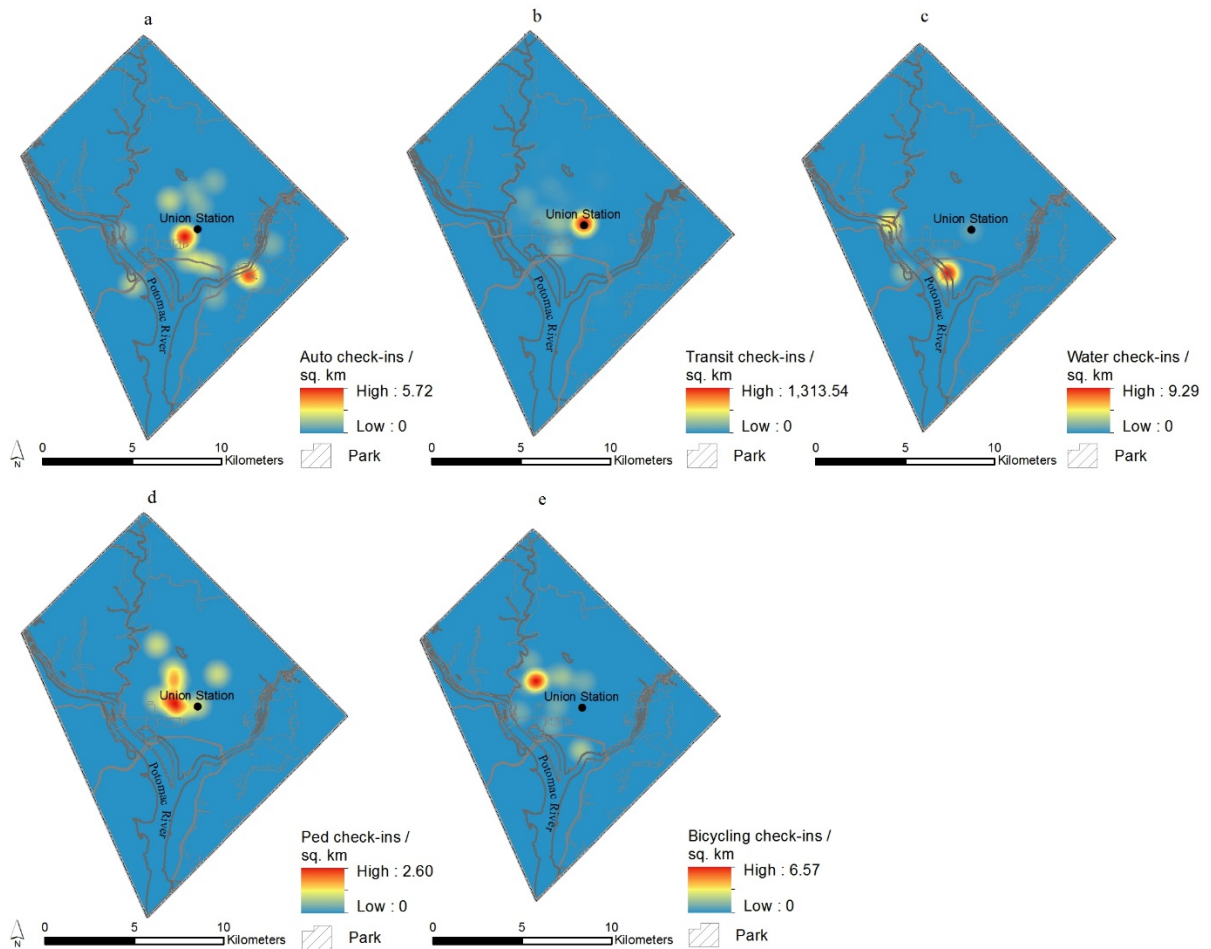


Figure 5. Kernel density estimation of Washington D.C. travel check-ins for a) automobile, b) mass-transit, c) water, d) walking, e) bicycle

6.2 Influence of travel mode on valence: global and spatial model results

Our second objective (“b”) in this research was to measure the associations between transportation mode and valence, while accounting for known predictors of travel. We implemented two global and spatial models to carry this out. The F-statistics and p-values showed

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

that each city's OLS model was statistically significant (p -value < 0.001); however, the explanatory power metrics (i.e., adjusted R^2 and AICc) showed that the models were weak (see tables 6 and 7). The result of using 12 independent variables in the global models explained 7% and 4% of the variance (adjusted R^2) in Chicago and D.C., respectively. The discouraging model diagnostics have been found in previous travel behavior studies. St-Louis et al., (2014) examined travel mode satisfaction among 3,377 commuters and found adjusted R^2 values ranged between .13 and .25 and in a similar study by Morris et al., (2015) the authors found OLS R^2 values between .001 and .139. The autocorrelation of the standardized residuals from each model were analyzed using Moran's I . Positive autocorrelation was found among the residuals for each city (Chicago Moran's $I = .012$, p -value < 0.001 ; D.C. Moran's $I = .03$, p -value < 0.001), indicating spatial nonstationarity. Although marginal, the implication is that the model's estimation reliability is compromised and can reduce our understanding of the links among valence, travel mode, and remaining explanatory factors.

Tables 6 and 7 display the goodness-of-fit measures (i.e., R^2 and AICc) for predicting valence in each city. The implementation of the SAR models improved upon the OLS results and corroborated the initial findings. In Chicago, there was an approximate 1% decrease in AICc and a R^2 increase of 31% (table 6). The SAR model results from D.C. echoed Chicago's; we observed an approximate 1% reduction in AICc and a 45% increase in R^2 was observed (table 7). The comparison of the AICc and R^2 values suggests the SAR models were examining a non-stationary feature in the relationship among the dependent and independent variables not addressed in the global OLS models. The *ex post facto* analysis of SAR standardized residuals from each model

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

(i.e., Chicago and D.C.) using a Moran's I correlogram largely indicated spatial randomness (figure 6). The reader is referred to Appendix A for a detailed Moran's I non-graphical output of the SAR residuals. In sum, the SAR model outputs showed moderate improvement over the OLS diagnostics and the covariate's directionality and statistical significance remained consistent.

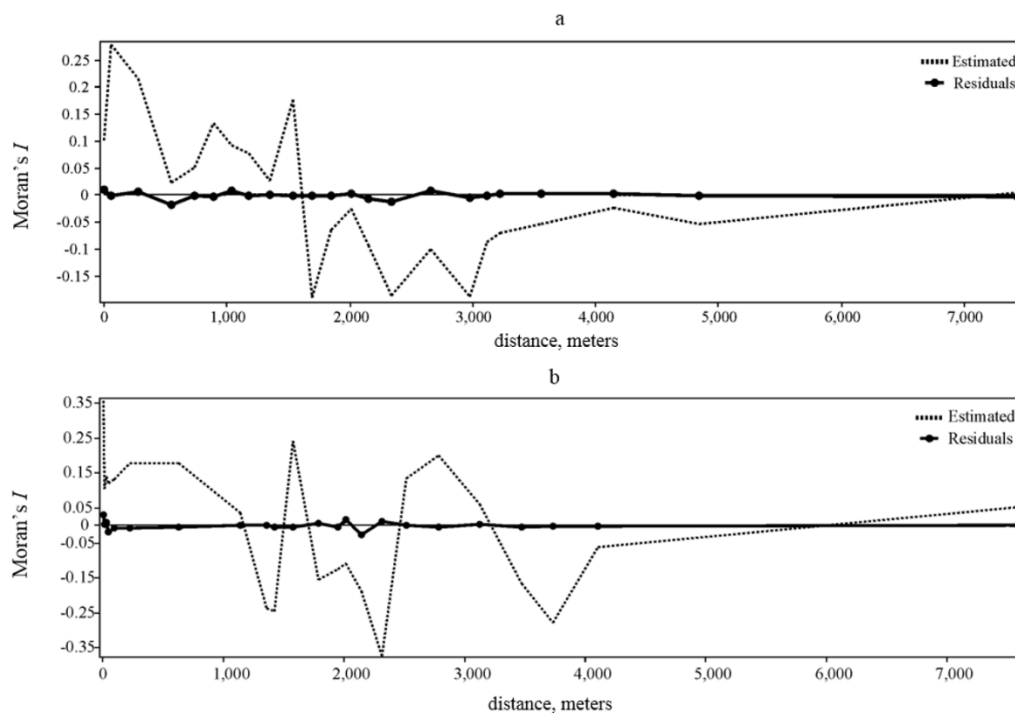


Figure 6. Chicago (a) and D.C. (b) spatial correlogram displaying the Moran's I for the SAR model residuals and estimated valance values.

The coefficients and significance levels for all four model estimates are displayed in tables 6 and 7. Due to the comparative strength of the SAR models, we highlighted the coefficients from these models. Across all models, water travel had the strongest positive influence on valence. The greatest impact was found in Chicago (SAR std. coeff. = 0.275, $p < 0.001$) and to a lesser degree

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

in D.C. (SAR std. coeff. = 0.085, $p = 0.002$). The result leads us to believe that the trip purpose is likely for recreational purposes. We verified this by scanning the database and discovering many Tweets indicating “boat tour.” The elevated valences for this mode are plausible and confirm prior research which has found mental health benefits associated with exposure to nature in urban areas (Bakolis et al. 2018). Pedestrian travel universally showed positive impacts on valences and was statistically significant. This mode induced the highest valence in Chicago (table 6), yet bordered statistical significance (SAR std. coeff. = 0.075, $p = 0.094$); the relationship in D.C. (table 7) was also positive and held greater statistical significance (SAR std. coeff. = 0.062, $p < 0.001$). The positive relationship between walking and valence is aligned with prior works. For instance, Gatersleben et al., (2007) found that walkers found their commute relaxing and later research from Daurte et al., (2010) posited that the exercise benefits of walking and bicycling outweighed the benefits of using the automobile. Table 7 showed that bicycling had a marginally positive affect on valence in D.C. (SAR std. coeff. = .047, $p = 0.066$). This finding is consistent with previous literatures and may also be reflective of the city’s recent national “bicycle-friendliness” ranking (St-Louis et al. 2014). Although bicycling positively affected valence in the City of Chicago as well, it was not statistically insignificant ($p = 0.894$) (table 6). The collective findings are especially encouraging on a broader scale. Given that walkable and bikeable neighborhoods promote social capital and increase health (Rogers et al. 2011; Celis-Morales et al. 2017), practitioners should be encouraged to use GSM data for the promotion of these modes, especially in marginalized communities. For instance, government officials may be able to record bicycling

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

and walking satisfaction in real-time to gauge travel and neighborhood happiness: a suggestion that has been advanced for elevating transit service quality (Collins, Hasan, and Ukkusuri 2013).

The majority of SES, environmental, weather, and temporal factors had an adverse impact on valence in each city (tables 6 and 7). Across both SAR models, weather was negatively associated with valence. Temperature was statistically significant and had a marginal effect on valence in Chicago (SAR std. coeff. = -0.056, $p = 0.003$); humidity was also slightly influential, however statistically less significant (SAR std. coeff. = -0.028, $p = 0.094$). In D.C. only temperature (SAR std. coeff. = -0.087, $p < 0.001$) was significant. The findings here recall past research emphasizing the linkages among temperature, humidity, travel behavior, and sentiment (Hannak et al. 2012; Saneinejad, Roorda, and Kennedy 2012; Miranda-Moreno and Nosal 2011). Additionally, the outputs here are noteworthy as they should assist stakeholders make improved decisions on how to promote active transportation modes while considering weather. For instance, providing air conditioning on mass-transit trains or busses may incentivize their use during extreme heat events and incorporating weather factors into travel-demand models may better predict timely (i.e., hourly, daily, etc.) city-wide walking, bicycling, and mass-transit rates. The morning temporal explanatory variable possessed a marginally negative association to valence and bordered statistical significance (SAR std. coeff. = -0.028, $p = 0.062$) in Chicago (table 6). We can infer that this time-frame reflects the morning commute, and unsurprisingly, has been previously associated with elevated stress (Legrain, Eluru, and El-Geneidy 2015). Therefore, interventions such as congestion pricing or educational campaigns focused on increasing mode shifts to bicycling or walking, should be looked upon as ways to increase sentiments during this time. The

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

density of street intersections showed a statistically significant, yet marginal, downward effect on valance in D.C. (SAR std. coeff. = -0.083, $p = 0.034$) (table 7). The relationship indicates that an increase in urban density results in a lower valence (i.e., sentiment). This result is corroborated by past works; for instance, Melis, et al., (2015) posited that urban density and antidepressant drug use were positively correlated, and Rocha et. al., (2012) found that stress and aggression were linked to traffic noise and pollution. Our results have added to these works, and suggest that GSM data could be used to ameliorate these effects by providing real-time geolocated sentiments on where, and to what extent, elevated congestion and noise levels are occurring for timely mitigation strategies. Additionally, minimizing these externalities could have positive implications for increasing walking and bicycling, as recent research has posited that these modes are used less when congestion related air pollution is elevated (Li and Kamargianni 2017).

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, Journal of Location Based Services, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Table 6. Chicago model diagnostics and standardized coefficients

Diagnostic statistics	Models	
	OLS	SAR
Akaike's Information Criterion (AICc)	8016.10	7915.49
R-square (adjusted)	0.071	0.103
F statistic	21.321	20.977
Independent variables		
Travel mode		
Bicycle	-.001	0.475
Mass-transit	.042	0.028
Pedestrian	.078**	0.075*
Water travel	.293***	0.275***
Socioeconomic		
Population density	-.007	0.006
2-car households	.007	-0.005
Environment		
Land-use mix	.027	0.028
Street intersections	-.024	-0.02
Time		
Morning	-.020	-0.037*
Night	.026	.01
Weather		
Humidity	-.035**	-0.028*
Temperature	-.039**	-0.056**

Adjusted OLS and SAR models includes all predictors simultaneously

SAR Alpha = .79, rho = .339

N= 3,218

OLS and SAR model p-values < 0.001

* p < 0.10

** p < 0.05

*** p < 0.001

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, Journal of Location Based Services, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Table 7. Washington D.C. model diagnostics and standardized coefficients

Diagnostic statistics	Models	
	OLS	SAR
Akaike's Information Criterion (AICc)	6935.736	6842.53
R-square (adjusted)	0.044	0.08
F statistic	11.732	11.346
Independent variables		
Travel mode		
Bicycle	0.045*	0.047*
Mass-transit	0.007	0.002
Pedestrian	0.065***	0.062***
Water travel	0.086***	0.085**
Socioeconomic		
Population density	0.02	0.006
2-car households	0.034	0.04
Environment		
Land-use mix	0.055**	0.04
Street intersections	-0.083**	-0.083**
Time		
Morning	-0.028	-0.028
Night	0.074***	0.051
Weather		
Humidity	0.028	0.019
Temperature	-0.063***	-0.087***

Adjusted OLS and SAR models includes all predictors simultaneously

SAR Alpha = .79, rho = .292

N= 2,804

OLS and SAR model p-values < 0.001

* p < 0.10

** p < 0.05

*** p < 0.001

7. Conclusion

The proliferation of geolocated social media data has provided additional channels to examine human travel behavior. In this paper, we made two significant contributions. First, we used exploratory spatial data analysis (i.e., hotspot mapping) to survey the spatiality of travel mode trends in two major U.S. cities. Second, we combined GSM data with neighborhood

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

indicators and extracted emotions from their written posts to understand their states of mind while traveling using global and spatial models. All our above contributions provide actionable insights and tools for policy makers to assess and enact efficient plans and policies set on reducing commute stress and promoting active transportation modes.

The results of the ESDA assessments illustrated that there were distinct spatial patterns of travel mode densities in each city using a common visualization technique, KDE (see objective “a”). We discovered interesting relationships between the water and automobile travel modes, and their proximity to natural resources. Noteworthy visualizations were also observed for walking and bicycling in each city. The spatial patterns of the mode check-ins indicated that these activities were occurring near tourist destinations, parks, and pedestrian/bicycle friendly neighborhoods. The outputs lend credence to the utility of using GSM data to understand transportation phenomenon at a high spatial resolution. Policy makers and planners could benefit from the results by inventorying the conditions in these areas and developing context-sensitive strategies to encourage a mode-shift from the automobile in neighborhoods where bicycling and walking is low.

Furthermore, in addressing our second objective (“b”), the current study found significant inferential relationships between travel mode and valence measures. Interesting associations were also observed among valence levels, environmental, and temporal attributes. We implemented the global OLS model as a base model. The SAR model proved superior in both cities and produced noteworthy coefficients. An important finding was that elevated valence scores were universally associated with water and pedestrian modes; along with bicycling, albeit in D.C. Our findings lend

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

additional evidence that travel may in of itself be a derived demand (Mokhtarian 2001). Furthermore, if a low-stress mode is utilized-which in this study included water travel, walking, and bicycling-it may possess a higher probably of being utilized again, thus reducing automobile usage and its associated externalities. Understanding when, how, and why stress is reduced, therefore, is essential for encouraging such mode choices in other areas. However, the evidence found in this research tells us that this is not a straightforward task. The results point to the degree in which valence was affected by time, (i.e., the morning commute), weather (i.e., humidity and temperatures), and environmental variables (i.e., urban density). From a policy perspective, these factors need to be accounted for, especially if GSM data is going to be used to quantify travel mode satisfaction in real-time. Overall, the outcomes represent a broader call for municipalities to utilize GSM to monitor and address public concerns in a timely manner through short term, and long term, planning strategies. As an example, real-time monitoring of human sentiments from Twitter could be used to proactively monitor mass-transit service provisioning or weather induced human stress. While strategies such as this deserve attention, the use of GSM requires acknowledgment of certain limitations.

There are two main concerns regarding using GSM data, of which one is selective sampling. Smith et al., (2012) showed that only 15 % of Twitter users are adults and the groups that Tweet mostly consist of: young adults, African Americans, urban/suburbanites, and those with smartphones. Despite past research linking environment to place (Bargh and Chartrand 1999), we must be cognizant of the limitation of GSM data in that check-in locales and Tweets may not always coincide. In other words, people may be Tweeting about their experience much later or

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

geographically further than where they were when the thoughts were manifested. In this study, we merged the Tweets with area level census data based on where they were sent. Therefore, caution should be noted when statements about the linkages among Tweets and underlying factors are made. A final limitation in the current research is there are likely other factors not captured which could contribute to valence fluctuations. For instance, travel purpose (i.e., recreation versus work) and the level of social-interaction were not measured and may influence travel mode sentiment (van den Berg, Sharmeen, and Weijs-Perrée 2017; Olsson et al. 2013).

These limitations notwithstanding, the overall findings of this work point to a convincing argument for the integration of GSM data in studying dynamic travel mode sentiments for reducing commute stress and increasing the pleasure of traveling, especially using active transportation modes. Considering the proliferation of this data, understanding how pleasure or displeasure (i.e. valence) affect travel behavior remains a fast growing research area. Therefore, a follow-up study utilizing more progressive sentiment algorithms such as LabMT or Probabilistic Latent Semantic Analysis (PLSA) will be considered for the next step in this research. In addition, a stronger time-series based predictive model, such as Granger Causality analysis, will be considered in a forthcoming study. It is hypothesized that stronger analytical frameworks such as these will provide stronger predictive capabilities and temporal accountability when GSM data is used. Ultimately, the present study has provided a first step in understanding the intimate geographical connection between travel mode and sentiment, using easily obtainable GSM data, and a transferable methodology which will incite others to acknowledge and bring into practice.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

7. Bibliography

- Abou-Zeid, Maya, Regina Witter, Michel Bierlaire, Vincent Kaufmann, and Moshe Ben-Akiva. 2012. "Happiness and travel mode switching: findings from a Swiss public transportation experiment." *Transport Policy* 19 (1):93-104.
- Abou Zeid, Maya. 2009. "Measuring and modeling activity and travel well-being." Massachusetts Institute of Technology.
- Andrienko, Gennady, Natalia Andrienko, Harald Bosch, Thomas Ertl, Georg Fuchs, Piotr Jankowski, and Dennis Thom. 2013. "Thematic patterns in georeferenced tweets through space-time visual analytics." *Computing in Science and Engineering* 15 (3):72-82.
- Anselin, Luc. 2013. *Spatial econometrics: methods and models*. Vol. 4: Springer Science & Business Media.
- Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. Paper presented at the LREC.
- Bailey, T.C., A.C. Gatrell. 1995. *Interactive Spatial Data Analysis*. London: Longman.
- Bakolis, Ioannis, Ryan Hammoud, Michael Smythe, Johanna Gibbons, Neil Davidson, Stefania Tognin, and Andrea Mechelli. 2018. "Urban Mind: Using Smartphone Technologies to Investigate the Impact of Nature on Mental Well-Being in Real Time." *BioScience*.
- Bandura, A. 1977. *Social Learning Theory*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Banerjee, Syagnik, Vijay Viswanathan, Kalyan Raman, and Hao Ying. 2013. "Assessing prime-time for geotargeting with mobile big data." *Journal of Marketing Analytics* 1 (3):174-83.
- Bargh, John A, and Tanya L Chartrand. 1999. "The unbearable automaticity of being." *American psychologist* 54 (7):462.
- Berg, B.L., H. Lune. 2004. *Qualitative Research Methods for the Social Sciences*. Vol. 5. Boston, MA: Pearson.
- Bocker, Lars, Martin Dijst, and Jan Prillwitz. 2012. "Impact of Everyday Weather on Individual Daily Travel Behaviours in Perspective: A Literature Review." *Transport Reviews* 33 (1):1-21.
- Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. "Twitter mood predicts the stock market." *Journal of Computational Science* 2 (1):1-8.
- Bradley, Margaret M, and Peter J Lang. 1999. "Affective norms for English words (ANEW): Instruction manual and affective ratings." In *Technical report C-1*. Florida: The center for research in psychophysiology, University of Florida.
- Burt, J., G. Barber, D. Rigby. 2009. *Elementary statistics for geographers*. 3rd ed. New York: The Guilford Press.
- Celis-Morales, Carlos A, Donald M Lyall, Paul Welsh, Jana Anderson, Lewis Steell, Yibing Guo, Reno Maldonado, Daniel F Mackay, Jill P Pell, and Naveed Sattar. 2017. "Association between active commuting and incident cardiovascular disease, cancer, and mortality: prospective cohort study." *bmj* 357.
- Cervero, R., and K. Kockelman. 1997. "Travel demand and the 3Ds: density, diversity, and design." *Transportation Research Part D* 2 (3):199-219.
- Cheng, Zhiyuan, James Caverlee, Kyumin Lee, and Daniel Z Sui. 2011. "Exploring Millions of Footprints in Location Sharing Services." *ICWSM 2011*:81-8.
- Collins, Craig, Samiul Hasan, and Satish V Ukkusuri. 2013. "A novel transit rider satisfaction metric: Rider sentiments measured from online social media data." *Journal of Public Transportation* 16 (2):2.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Cromley, E.K., S.L., McLafferty. 2012. *GIS and Public Health*. 2nd ed. New York: The Guilford Press.

Das, Sanjiv, and Mike Chen. 2001. Yahoo! for Amazon: Extracting market sentiment from stock message boards. Paper presented at the Proceedings of the Asia Pacific finance association annual conference (APFA).

De Vos, Jonas, Tim Schwanen, Veronique Van Acker, and Frank Witlox. 2013. "Travel and subjective well-being: a focus on findings, methods and future research needs." *Transport Reviews* 33 (4):421-42.

Dille, I. 2017. "The 50 Best Bike Cities of 2016." *Bicycling Magazine*, Accessed May 17, 2017. <http://www.bicycling.com/culture/news/the-50-best-bike-cities-of-2016>.

Dormann, C.F, J.M McPherson, M.B. Araujo, R. Bivand, J. Bolliger, G. Carl, G Davies, A. Hirzel, J. Walter, D. Kissling, I. Kuhn, R. Ohlemuller, P. Peres-Neto, B. Reineking, B. Shroder, F. M. Schurr, R. Wilson. 2007. "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review." *Ecography* 30 (5):609-28.

Downs, Joni A. 2010. "Time-geographic density estimation for moving point objects." In *Geographic Information Science*, 16-26. Springer.

Duarte, André, Camila Garcia, Grigoris Giannarakis, Susana Limão, Amalia Polydoropoulou, and Nikolaos Litinas. 2010. "New approaches in transportation planning: happiness and transport economics." *Netnomics* 11 (1):5-32.

Eagly, A. H., Mladinic, A., Otto, S. 1994. "Cognitive and affective bases of attitudes toward social groups and social policies." *Journal of Experimental Social Psychology* 30:113-37.

Esmín, Ahmed AA, Roberto L De Oliveira Jr, and Stan Matwin. 2012. Hierarchical classification approach to emotion recognition in twitter. Paper presented at the Machine Learning and Applications (ICMLA), 2012 11th International Conference on Machine Learning and Applications.

Frank, Morgan R, Lewis Mitchell, Peter Sheridan Dodds, and Christopher M Danforth. 2013. "Happiness and the patterns of life: A study of geolocated tweets." *Scientific reports* 3.

Furie, Gregg L, and Mayur M Desai. 2012. "Active transportation and cardiovascular disease risk factors in US adults." *American journal of preventive medicine* 43 (6):621-8.

Gatersleben, Birgitta, and David Uzzell. 2007. "Affective appraisals of the daily commute comparing perceptions of drivers, cyclists, walkers, and users of public transport." *Environment and behavior* 39 (3):416-31.

Gerber, Matthew S. 2014. "Predicting crime using Twitter and kernel density estimation." *Decision Support Systems* 61:115-25.

Giannotti, Fosca, and Dino Pedreschi. 2008. *Mobility, data mining and privacy: Geographic knowledge discovery*: Springer Science & Business Media.

Han, Su Yeon, Ming-Hsiang Tsou, and Keith C Clarke. 2015. "Do global cities enable global views? Using Twitter to quantify the level of geographical awareness of US cities." *PloS one* 10 (7):e0132464.

Hannak, Aniko, Eric Anderson, Lisa Feldman Barrett, Sune Lehmann, Alan Mislove, and Mirek Riedewald. 2012. Tweetin'in the Rain: Exploring Societal-Scale Effects of Weather on Mood. Paper presented at the ICWSM.

Hoehner, Christine M, Carolyn E Barlow, Peg Allen, and Mario Schootman. 2012. "Commuting distance, cardiorespiratory fitness, and metabolic risk." *American journal of preventive medicine* 42 (6):571-8.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

- Huang, Qunying, and David WS Wong. 2015. "Modeling and Visualizing Regular Human Mobility Patterns with Uncertainty: An Example Using Twitter Data." *Annals of the Association of American Geographers*:1-19.
- . 2016. "Activity patterns, socioeconomic status and urban spatial structure: what can social media data tell us?" *International Journal of Geographical Information Science* 30 (9):1873-98.
- Järv, Olle, Rein Ahas, and Frank Witlox. 2014. "Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records." *Transportation Research Part C: Emerging Technologies* 38:122-35.
- Kahneman, Daniel, Alan B Krueger, David Schkade, Norbert Schwarz, and Arthur Stone. 2004. "Toward national well-being accounts." *The American Economic Review* 94 (2):429-34.
- Katz, Phil, Matthew Singleton, and Richard Wicentowski. 2007. Swat-mp: the semeval-2007 systems for task 5 and task 14. Paper presented at the Proceedings of the 4th international workshop on semantic evaluations.
- Kim, Sunghwan Mac, Kyo Kageura, James McHugh, Surya Nepal, Cécile Paris, Bella Robinson, Ross Sparks, and Stephen Wan. 2017. Twitter Content Eliciting User Engagement: A Case Study on Australian Organisations. Paper presented at the Proceedings of the 26th International Conference on World Wide Web Companion.
- Kissling, W Daniel, and Gudrun Carl. 2008. "Spatial autocorrelation and the selection of simultaneous autoregressive models." *Global Ecology and Biogeography* 17 (1):59-71.
- Latham, A. 2003. "Research, performance, and doing human geography: some reflections on the diary-photograph, diary-interview method." *Environment and Planning A* 35 (11):1993-2017.
- Legrain, Alexander, Naveen Eluru, and Ahmed M El-Geneidy. 2015. "Am stressed, must travel: The relationship between mode choice and commuting stress." *Transportation research part F: traffic psychology and behaviour* 34:141-51.
- Lei, Jingsheng, Yanghui Rao, Qing Li, Xiaojun Quan, and Liu Wenyin. 2014. "Towards building a social emotion detection system for online news." *Future Generation Computer Systems* 37:438-48.
- Li, Linna, Michael F Goodchild, and Bo Xu. 2013. "Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr." *Cartography and Geographic Information Science* 40 (2):61-77.
- Li, Weibo, and Maria Kamargianni. 2017. "Air pollution and seasonality effects on mode choice in China." *Transportation Research Record: Journal of the Transportation Research Board* (2634):101-9.
- Lichman, Moshe, and Padhraic Smyth. 2014. Modeling human location data with mixtures of kernel densities. Paper presented at the Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, New York.
- Luo, Feixiong, Guofeng Cao, Kevin Mulligan, and Xiang Li. 2016. "Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago." *Applied Geography* 70:11-25.
- Mehrabian, Albert, and James A Russell. 1974. *An approach to environmental psychology*: the MIT Press.
- Melis, Giulia, Elena Gelormino, Giulia Marra, Elisa Ferracin, and Giuseppe Costa. 2015. "The effects of the urban built environment on mental health: A cohort study in a large northern Italian city." *International journal of environmental research and public health* 12 (11):14898-915.
- Mihalcea, Rada, Carmen Banea, and Janyce Wiebe. 2007. Learning multilingual subjective language via cross-lingual projections. Paper presented at the Annual meeting-association for computational linguistics.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

- Miranda-Moreno, Luis, and Thomas Nosal. 2011. "Weather or not to cycle: Temporal trends and impact of weather on cycling in an urban environment." *Transportation Research Record: Journal of the Transportation Research Board* (2247):42-52.
- Mitchell, Lewis, Morgan R Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M Danforth. 2013. "The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place." *PloS one* 8 (5):e64417.
- Mokhtarian, P.L., I. Salomon. 2001. "How derived is the demand for travel? Some conceptual and measurement considerations." *Transportation Research Part A - Policy and Practice* 35 (8):695-719.
- Mondschein, Andrew. 2015. "Five-star transportation: using online activity reviews to examine mode choice to non-work destinations." *Transportation*:1-16.
- Morris, Eric A, and Erick Guerra. 2015. "Mood and mode: does how we travel affect how we feel?" *Transportation* 42 (1):25-43.
- Neutens, T., T. Schwanen. 2011. "The prism of everyday life: Towards a new research agenda for time geography." *Transport Reviews* 31 (1):25-47.
- Nguyen, Quynh C, Suraj Kath, Hsien-Wen Meng, Dapeng Li, Ken R Smith, James A VanDerslice, Ming Wen, and Feifei Li. 2016. "Leveraging geotagged Twitter data to examine neighborhood happiness, diet, and physical activity." *Applied Geography* 73:77-88.
- Noulas, Anastasios, Salvatore Scellato, Renaud Lambiotte, Massimiliano Pontil, and Cecilia Mascolo. 2012. "A tale of many cities: universal patterns in human urban mobility." *PloS one* 7 (5):e37027.
- Novaco, Raymond W, and Oscar I Gonzalez. 2009. "Commuting and well-being." In *Technology and well-being*, edited by Y. Amichai-Hamburger, 174-205. Cambridge University Press.
- Olafson, K. 2014. "Can everyday bicycling make you happier." In *Momentum Magazine*.
- Olsson, Lars E, Tommy Gärling, Dick Ettema, Margareta Friman, and Satoshi Fujii. 2013. "Happiness and satisfaction with work commute." *Social indicators research* 111 (1):255-63.
- Ory, David T, Patricia L Mokhtarian, Lothlorien S Redmond, Ilan Salomon, Gustavo O Collantes, and Sangho Choo. 2004. "When is commuting desirable to the individual?" *Growth and Change* 35 (3):334-59.
- Poorthuis, Ate, Matthew Zook, Taylor Shelton, Mark Graham, and Monica Stephens. 2016. "Using geotagged digital social data in geographic research." In *Key methods in geography*, edited by S. French N. Clifford, M. Cope, T. Gellespie. Thousand Oaks: Sage.
- Ramsey, K., Bell, A. 2014. "Smart Location Database, Version 2.0 User Guide." In, edited by U.S. Environmental Protection Agency, 1-52.
- Rangel, T.F.L.V.B, Diniz-Filho, J.A.F and Bini, L.M. 2010. "SAM: a comprehensive application for Spatial Analysis in Macroecology." *Ecography* 33:46-50.
- Reardon, Louise, and Saamah Abdallah. 2013. "Well-being and transport: Taking stock and looking forward." *Transport Reviews* 33 (6):634-57.
- Redmond, Lothlorien S, and Patricia L Mokhtarian. 2001. "The positive utility of the commute: modeling ideal commute time and relative desired commute amount." *Transportation* 28 (2):179-205.
- Rocha, Kátia, Katherine Pérez, Maica Rodríguez-Sanz, Jordi E Obiols, and Carme Borrell. 2012. "Perception of environmental problems and common mental disorders (CMD)." *Social psychiatry and psychiatric epidemiology* 47 (10):1675-84.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

- Rogers, Shannon H, John M Halstead, Kevin H Gardner, and Cynthia H Carlson. 2011. "Examining walkability and social capital as indicators of quality of life at the municipal and neighborhood scales." *Applied Research in Quality of Life* 6 (2):201-13.
- Rui, Huaxia, and Andrew Whinston. 2011. "Designing a social-broadcasting-based business intelligence system." *ACM Transactions on Management Information Systems (TMIS)* 2 (4):22.
- Rybarczyk, G. 2012. Forecasting bicycling risk factors within neighborhoods: A spatial autoregressive approach. Paper presented at the Papers of the Applied Geography Conferences.
- Rybarczyk, G., and L. Gallagher. 2014. "Measuring the potential for bicycling and walking at a metropolitan commuter university." *Journal of Transport Geography* 39:1-10.
- Saelens, B., J. F. Sallis, and L. D. Frank. 2003. "Environmental Correlates of Walking and Cycling: Findings From the Transportation, Urban Design, and Planning Literatures." *Annals of Behavioral Medicine* 25 (2):80-91.
- Saif, Hassan, Yulan He, and Harith Alani. 2012. "Semantic sentiment analysis of twitter." *The Semantic Web-ISWC 2012*:508-24.
- Sallis, J.F., N. Owen. 2002. "Ecological models of health behavior." In *Health Behavior and Health Education: Theory, Research, and Practice*, edited by K. Glanz, FM Lewis, BK Rimer, 462-84. San Francisco, CA: Jossey-Bass.
- Salon, Deborah. 2016. "Estimating pedestrian and cyclist activity at the neighborhood scale." *Journal of Transport Geography* 55:11-21.
- Saneinejad, Sheyda, Matthew J. Roorda, and Christopher Kennedy. 2012. "Modelling the impact of weather conditions on active transportation travel behaviour." *Transportation Research Part D: Transport and Environment* 17 (2):129-37.
- Schafer, Andreas. 2000. "Regularities in travel demand: an international perspective." *Journal of transportation and statistics* 3 (3):1-31.
- Schweitzer, Lisa. 2014. "Planning and social media: a case study of public transit and stigma on Twitter." *Journal of the American Planning Association* 80 (3):218-38.
- Shaker, Richard Ross. 2016. "Examining sustainable landscape function across the Republic of Moldova." *Habitat International*:1-15. doi: <https://doi.org/10.1016/j.habitatint.2016.11.002>.
- Silverman, Bernard W. 1986. *Density estimation for statistics and data analysis*. Vol. 26: CRC press.
- Smith, A., J. Brenner. 2012. "Twitter use 2012." In. Washington D.C. : Pew Research Center.
- Smith, C. June 2, 2016. "By the Numbers: 20 Important foursquare Stats." <http://expandedramblings.com/index.php/by-the-numbers-interesting-foursquare-user-stats/>.
- St-Louis, Evelyne, Kevin Manaugh, Dea van Lierop, and Ahmed El-Geneidy. 2014. "The happy commuter: A comparison of commuter satisfaction across modes." *Transportation research part F: traffic psychology and behaviour* 26:160-70.
- Steiger, Enrico, René Westerholt, Bernd Resch, and Alexander Zipf. 2015. "Twitter as an indicator for whereabouts of people? Correlating twitter with uk census data." *Computers, Environment and Urban Systems* 54:255-65.
- Strapparava, Carlo, and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. Paper presented at the Proceedings of the 4th International Workshop on Semantic Evaluations.
- Strapparava, Carlo, and Alessandro Valitutti. 2004. WordNet Affect: an Affective Extension of WordNet. Paper presented at the LREC.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, *Journal of Location Based Services*, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

- Sui, Daniel, and Michael Goodchild. 2011. "The convergence of GIS and social media: challenges for GIScience." *International Journal of Geographical Information Science* 25 (11):1737-48.
- Sun, Y., D. Yunan, Y. Wang, L. Zhuang. 2017. "Examining associations of environmental characteristics with recreational cycling behavior by street-level Strava data." *International journal of environmental research and public health* 14:1-12.
- Sun, Yeran, Hongchao Fan, Ming Li, and Alexander Zipf. 2016. "Identifying the city center using human travel flows generated from location-based social networking data." *Environment and Planning B: Planning and Design* 43 (3):480-98.
- Tsukayama, Hayley. 2017. "Twitter is officially doubling the character limit to 280." In *The Washington Post*. Washington D.C.
- Turley, Lou W, and Ronald E Milliman. 2000. "Atmospheric effects on shopping behavior: a review of the experimental evidence." *Journal of business research* 49 (2):193-211.
- Turney, Peter D, and Michael L Littman. 2003. "Measuring praise and criticism: Inference of semantic orientation from association." *ACM Transactions on Information Systems (TOIS)* 21 (4):315-46.
- Van Acker, Veronique, Bert Van Wee, and Frank Witlox. 2010. "When transport geography meets social psychology: toward a conceptual model of travel behaviour." *Transport Reviews* 30 (2):219-40.
- van den Berg, Pauline, Fariya Sharmeen, and Minou Weijs-Perrée. 2017. "On the subjective quality of social interactions: Influence of neighborhood walkability, social cohesion and mobility choices." *Transportation Research Part A: Policy and Practice* 106:309-19.
- Wagner, Helene H, and Marie-Josée Fortin. 2005. "Spatial analysis of landscapes: concepts and statistics." *Ecology* 86 (8):1975-87.
- Widener, Michael J, and Wenwen Li. 2014. "Using geolocated Twitter data to monitor the prevalence of healthy and unhealthy food references across the US." *Applied Geography* 54:189-97.
- Wu, Chao, Xinyue Ye, Fu Ren, and Qingyun Du. 2018. "Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China." *Cities*. doi: <https://doi.org/10.1016/j.cities.2018.01.017>.
- Wu, Lun, Ye Zhi, Zhengwei Sui, and Yu Liu. 2014. "Intra-urban human mobility and activity transition: Evidence from social media check-in data." *PloS one* 9 (5):e97010.
- Yuan, Yihong, Martin Raubal, and Yu Liu. 2012. "Correlating mobile phone usage and travel behavior—A case study of Harbin, China." *Computers, Environment and Urban Systems* 36 (2):118-30.
- Zhen, Feng, Yang Cao, Xiao Qin, and Bo Wang. 2017. "Delineation of an urban agglomeration boundary based on Sina Weibo microblog 'check-in' data: A case study of the Yangtze River Delta." *Cities* 60:180-91.
- Zhou, Xiaolu, and Liang Zhang. 2016. "Crowdsourcing functions of the living city from Twitter and Foursquare data." *Cartography and Geographic Information Science*:1-12.

Please cite as:

Greg Rybarczyk, Syagnik Banerjee, Melissa D. Starking-Szymanski & Richard R. Shaker (2018) Travel and us: the impact of mode share on sentiment using geo-social media and GIS, Journal of Location Based Services, 12:1, 40-62, DOI: 10.1080/17489725.2018.1468039

To link to this article: <https://doi.org/10.1080/17489725.2018.1468039>

Appendix A

SAR Moran's *I* significance and distance values

Chicago				Washington D.C.			
bin	Distance (m)	Moran's <i>I</i>	p-val	bin	Distance (m)	Moran's <i>I</i>	p-val
1	0.55	0.010	<.001*	1	7.04	0.028	<.001*
2	63.54	0.000	0.290	2	20.05	0.006	<.001*
3	296.58	0.000	0.120	3	34.19	0.007	<.001*
4	573.60	-0.010	<.001*	4	63.61	-0.024	<.001*
5	767.94	0.000	0.300	5	127.27	-0.004	0.039*
6	943.41	0.000	0.320	6	444.09	-0.014	<.001*
7	1,094.86	0.000	0.030*	7	995.42	<.001	0.791
8	1,246.96	-0.010	<.001*	8	1318.75	0.003	0.098
9	1,438.56	<.001	0.900	9	1,398.31	<.001	0.473
10	1,621.91	<.001	0.610	10	1,548.47	-0.004	0.033*
11	1,767.64	<.001	0.800	11	1,793.19	-0.002	0.477
12	1,919.58	0.000	0.100	12	1,957.78	0.002	0.121
13	2,101.79	-0.010	<.001*	13	2,032.34	0.008	<.001*
14	2,282.45	0.010	0.010	14	2,175.83	-0.008	<.001*
15	2,539.30	-0.010	<.001*	15	2,374.48	0.009	<.001*
16	2,876.61	0.010	<.001*	16	2,632.51	-0.003	0.196
17	3,083.71	-0.010	<.001*	17	2,923.52	-0.009	<.001*
18	3,175.71	0.000	0.090	18	3,321.10	-0.005	0.020*
19	3,454.55	0.000	0.110	19	3,696.15	0.001	0.305
20	4,045.97	0.010	<.001*	20	4,064.99	0.002	0.192
21	4,786.99	-0.010	<.001*	21	7,568.00	-0.003	0.118

* p-val < 0.05