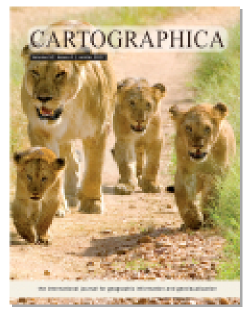
What Is So "Hot" in Heatmap?: Qualitative Code Cluster Analysis with Foursquare Venue



Ilyoung Hong, Jin-Kyu Jung

Cartographica: The International Journal for Geographic Information and Geovisualization, Volume 52, Number 4, Winter 2017, pp. 332-348 (Article)

Published by University of Toronto Press

ARTICLE

What Is So "Hot" in Heatmap? Qualitative Code Cluster Analysis with Foursquare Venue

Ilyoung Hong
Department of GIS Engineering / Namseoul University / Cheonan-si Chungcheongnam-do / South Korea

Jin-Kyu Jung
School of Interdisciplinary Arts and Sciences / University of Washington-Bothell / Bothell / WA / USA

ABSTRACT

Foursquare is a popular Web service and a representative location-based social network (LBSN) service using position data. Heatmap is a widely used means of geovisualization for analyzing social data with locational values. Until now, heatmap analysis of LBSN has focused on identifying quantitative distribution and patterns, with little consideration of the qualitative analysis of data content. Based on a case study of Foursquare venues and user-created content in Seattle, WA, this study conducts analyses assessing both the quantitative spatial distribution and the qualitative characteristics of coffee shops in the Seattle metropolitan area. It specifically proposes a new analytical method referred to as "code cluster," which is designed to employ quantitative and qualitative approaches simultaneously. The significance of this method is its capacity to explain geographical differences in terms of qualitative traits in cluster regions, in addition to analyzing their spatial characteristics and distributions. In introducing this new hybrid approach, our aims are to reflect the original intent and essence of the data throughout the research process and to make further efforts to analyze and interpret the contextualized meanings. This will be possible through integration of advanced spatial analysis, geovisualization, and qualitative research that build on current geographic and geovisual research with big data.

Keywords: code cluster, hybrid approach, geovisualization, spatial analysis, location-based social network, Foursquare

RÉSUMÉ

Foursquare est un service Web populaire, typique des réseaux sociaux géodépendants (RSG) utilisant les données de géolocalisation. La carte de densité de clics est une technique de géovisualisation largement utilisée pour analyser les données sociales ayant des valeurs à référence spatiale. Jusqu'à maintenant, l'analyse des cartes de densité de clics des RSG visait plus particulièrement l'observation de distributions et de profils quantitatifs, l'analyse qualitative du contenu en information suscitant peu d'intérêt. Dans une étude de cas portant sur des endroits de Foursquare et un contenu créé par les utilisateurs à Seattle (Washington), les auteurs procèdent à des analyses visant à évaluer la distribution spatiale quantitative et les caractéristiques qualitatives des cafés-restaurants de la région métropolitaine de Seattle. Ils proposent plus précisément une nouvelle méthode analytique dite des « grappes de codes », conçue pour permettre l'emploi simultané d'approches quantitative et qualitative. L'intérêt de cette méthode tient au fait qu'elle peut expliquer les différences géographiques en termes de caractéristiques qualitatives dans les régions accueillant une grappe, en plus de l'analyse des caractéristiques et des distributions spatiales. En proposant cette nouvelle méthode hybride, les auteurs ont pour but de rendre compte de l'intention et de l'essence initiales des données au fil du processus de recherche et de consacrer davantage d'efforts à l'analyse et à l'interprétation de leurs significations contextualisées. Ce but pourra être atteint grâce à l'intégration de l'analyse spatiale avancée, de la géovisualisation et de la recherche qualitative reposant sur la recherche géographique et géovisuelle actuelle faisant appel aux mégadonnées.

Mots clés : analyse spatiale, Foursquare, géovisualisation, grappe de codes, méthode hybride, réseau social géodépendant

Introduction

Foursquare and Facebook Place are popular location-based social networks (LBSNs) that use position data and Web services. Unlike conventional social media such as Twitter and Facebook, LBSNs allow users to form relationships around particular locations. For example, users can save location data regarding places of interest and leave comments to share the information with their friends and other users. LBSN data have also become an important

  doi:10.3138/cart.52.4.2016-0005

topic of interest for social media analysts and geographers, and particular attention has recently been given to the geographic characteristics and distribution of LBSN contents (Noulas and others 2011; Cheng and others 2011). These researchers all acknowledge the impact of social media and networks, and how they leave geographic traces. Stefanidis and others (2013) viewed this trend as the emergence of "ambient geospatial information (AGI)" that captures people's references to locations that represent momentary social "hotspots." Wilson (2011, 370) also offered some useful observations about the technically mediated "conspicuous mobility" that forms emergent mappings of everyday life through LBSN such as Foursquare. It is the demonstration of the re-conceptualization of "mapping" in the big data world that Kitchin and Dodge (2007, 342) called "emergent cartography," which focused on how the concepts of map and mapping emerged through contingent and context-embedded practices. These studies prompt us to think about a new possibility of identifying and mapping geospatial hotspot emergence, and more importantly, analyzing the contextual meanings of data that might provide valuable insight into an area of interest. A salient point is even made by Foursquare on their Web site inviting users to "Find the best places to eat, drink, shop, or visit any city in the world. Access over 75 million short trips of local experience."[1] It is not surprising that Foursquare claims to show us the trending hotspots where people go for various activities (Valentino-Devries 2011). However, a question still remains as to "why" and "for what" those areas are so hot. This is the starting point of our inquiry. By rethinking and repositioning the epistemology and analysis of mapping geospatial social media and spatial big data research, we would like to suggest an integrative approach that combines quantitative, geovisual, and qualitative analyses to progress beyond simple location mapping, valuing the contextual meanings of LBSN data within a hybrid framework.

A heatmap is frequently used to visualize geographically tagged social data with high spatial densities. To this end, the kernel density method is generally used to smoothly represent the density distribution surfaces of point data according to the determined distance (Chainey and others 2003). However, a heatmap only shows the density of point data, not the temperature innate to the context of the text contained in social media. The meanings and stories behind data, or the "thick description" (Geertz 1973) of the geotagged data, should be taken fully into consideration in seeking a comprehensive understanding of the data and their relations to the people and places. This also speaks for the need to apply qualitative analysis in the research of spatial media and the importance of reflective and reflexive interpretations of the contexts of data, rather than their quantified representation in Euclidean space. We intend to fill this gap by first reviewing the efforts made in previous studies to integrate qualitative methods with quantitative geospatial and geovisual analyses in studying big data hitherto, and then extending the discussion using our own approach based on a case study of Foursquare in Seattle, WA.

By combining geosocial media's quantitative analytical method of the heatmap and the qualitative and visual method of the code cloud, this study proposes code cluster analysis as a means for analyzing *and* visualizing the context of a heatmap. In addition to assessing the spatial distribution and characteristics of social media data with coordinate values, code cluster analysis helps us to evaluate the qualitative characteristics of the clusters. In other words, it not only shows the hotspots of the social media, but also allows us to understand and *see* how and why certain regions or places become "hot" in relation to users' participation and interests and to local characteristics. With the empirical case of coffee shops in Seattle using Foursquare venues and user-created contents, not only the conceptual but also the practical aspects of the code cluster method will be carefully examined.

We will begin by reviewing various existing studies on LBSN and social media, and then discuss the characteristics of Foursquare as an example of LBSN. In the following section, we demonstrate a detailed design and a framework of code cluster from data collection and quantitative cluster analysis to qualitative coding analysis. A case study of Seattle's coffee shops using Foursquare demonstrates how code cluster analysis can be implemented as a hybrid approach combining quantitative and qualitative methods using LBSN data, as well as yielding new insights about each clustering area. In the concluding section, we reflect on the significance of this study by highlighting new possibilities for the geographic analysis of social and spatial media and for spatial big data in general.

Related Studies

LBSN is a social network that provides a location-based service. As more people use LBSN, there is a rapid increase in the number of LBSN data proactively created by smartphone users sharing their location data and information about places. Popular LBSN services include Foursquare, Google Local, and Facebook Place; Foursquare is the most widely used (Zickuhr and Smith 2010). Venue in Foursquare, in particular, has become a major social network medium, and it has emerged as an important research topic as a source of information that can reflect local and regional characteristics. The information stored in a venue includes the attributes of the place, name, tips, photos, and the category code, as well as statistical data such as numbers of check-ins and tips. Unlike conventional social networks such as Twitter and Facebook that only allow friends to share information, LBSN users share information by checking into or writing comments (tips) about specific places. Gao and Liu (2014) explained the characteristics of LBSN with the "3+1" framework: the contents layer (audio, video, photograph, tip), the social layer (social relationships), and the geographical layer (users' check-in activities), plus the timeline along which the activities take place. The geographical layer is an especially distinctive set of data not found in conventional social media.

Research on the geographical approach to social media can be broadly categorized into three major areas. First, there have been studies of the spatial distribution of point information that implement point data using tagged coordinates. These studies have assessed areas with high densities of point data, focusing on their patterns, and analyzed the geographical attributes of the data and their associations with regional characteristics (Li, Goodchild, and Xu 2013a; Mitchell and others 2013). For example, the majority of studies of LBSN have focused on statistical analysis by compiling data regarding tens of millions of venues around the world and their user check-ins (Cheng and others 2011; Li and others 2013b). Noulas and others (2011), for instance, conducted a study with 700,000 Foursquare users over 100 days based on their check-in behaviour, primary usage hours, and spatial patterns of checked-in locations. These studies emphasized collecting big data using an application programming interface (API) and the statistical characteristics of the compiled data.

Second, there have been studies assessing the correlation between geotagged data and the corresponding geographic regions or the characteristics of the residents in the region by applying geographic analytical methods (Ghosh and Guha 2013; Hahmann and others 2014). For example, Hong (2015) analyzed the geographical distribution characteristics of LBSN data and their association with sociodemographic information in his case study of Seoul, Korea, using census data.

Third, there are a relatively small number of qualitative analytical studies with emphasis on social media contents, and only a few synthetic approaches invite integration of qualitative data and analyses and geographic visualization. These qualitative analyses are often intended to identify content of social media and propose means for representing them on a map. To depict textual data contained in social media, the most widely used method is the word cloud, which visualizes the frequencies of the words included in text (Cidell 2010). While the word cloud identifies common keywords that are of interests to users, it also has the limitation of not accounting for the geographical context that arises from the keywords. Drawing on spatially linked social media and applying lessons from qualitative GIS and geographic visualization in particular, Jung (2015) proposed the code cloud method to consider both the spatial and qualitative characteristics of geotagged social media. As a means of analyzing and visualizing the contextual meanings of geosocial media, the code cloud is a unique alternative. As an example of qualitative geovisualization, code clouds guide us to include various forms of quantitative and qualitative data such as numbers, texts, photos, videos, GIS maps, or even hyperlinked multimedia that consider a map by itself as a procedural platform for the navigation. However, we believe this method also has a few limitations that we hope to respond to and improve in this article. In processing large-scale data such as geotweets, promptly integrating the outcomes of qualitative analyses with the spatial processing of geotagged data seemed difficult. In addition, in regard to implementing spatial analysis and geovisualization, the code cloud is also prone to the limitation that it only applies the basic level of GIS analysis, for example, overlaying of points with Census demographic data. Furthermore, there might be an inherent limitation in manually conducting qualitative ethnographic research on big data, because only a small number of data could be analyzed within a limited timeline. In spite of these challenges, we can exert steady efforts to develop new theoretical and practical methods for combining quantitative spatial analysis, geographic visualization, and qualitative analysis to engage with geosocial data. For this purpose, we present a code cluster method as a new way of carrying out an integrated analysis by combining qualitative analyses and enhanced geospatial analyses with the power of (qualitative) geovisualization.

Working with Foursquare

Foursquare provides highly instrumental LBSN data that can be categorized into venues and tips on venues. Until now, most studies using Foursquare have focused on the statistical distribution and the characteristics of tips, and there has been little research on the application of geographical analysis, or on the content aspect, such as categorization of venues and texts. On the other hand, the discussions of combining geographic analysis and qualitative approaches to social media have primarily focused on geotweets, including Cidell's (2010) content cloud and Jung's (2015) code cloud, which we previously reviewed. Both Cidell (2010) and Jung (2015) emphasized the regional context or key issues, which had not been taken into account in earlier quantitative analyses of social media data. However, there is also a difficulty in working with geotweets due to the weak association between geotweet data and their corresponding locations. The weak association stems from the fact that a user's tweet is not necessary relevant to a particular location; for example, a tweet's contents might not be closely related to where the tweet was generated. In this regard, Foursquare provide a unique advantage. Unlike geotweets, Foursquare contains location-specific content in the form of venues (Gao and Liu 2014). A venue can be regarded as a point of interest (POI) constructed based on user participation. Users express their opinions and evaluations of a venue in the form of tips, which allow others to learn about the venue. User opinions can be also interpreted as the level of hotness of a venue. For instance, venues and tips together provide a frame for creating data that can take into account the specific traits of a location. When users write tips about a venue, these characteristics allow them to generate contents with high levels of connection with specific topics or the area where the venue is located.

In this regard, hybrid, interdisciplinary, or cross-disciplinary research is necessary to analyze the characteristics of various types of data simultaneously for a set of data with a strong qualitative trait, such as social media text (Barewald 2010; Sui and DeLyser 2012; DeLyser and Sui 2013). Earlier and recent discussions of critical GIS and qualitative GIS/geovisualization offer several innovative ways to integrate qualitative data and analysis with GIS (Schuurman 1999; Kwan and Ding 2008; Cope and Elwood 2009; Knigge and Cope 2009; Elwood 2011; Burns and Skupin 2013; Schoepfer and Rogers 2014). Qualitative GIS/geovisualization extends the capacity of geovisualization and maps to better represent people's experiential and interpretive knowledge of geographic spaces by showing the inherent impossibility of framing any one research method or form of representation. Qualitative GIS/geovisualization research further provides a theoretical frame in which we can explore innovative and creative possibilities in geovisualization intersected with critical social theories, visual methodologies, arts, and digital spatial humanities (Elwood and Mitchell 2013; Hawkins 2014; Elwood 2015; Harris 2015). The code cluster method, which we will explain further in the following section, is our response to new developments, particularly those that employ a hybrid research model in analyzing the context of the text and geovisualize the social media data to reflect regional characteristics.

Statistical and quantitative analysis are a priority, considering that social media are stored in the form of big data in a database, often on the Internet; however, a qualitative approach is also necessary, since social media have qualitative data (e.g., text data) that contain users' opinions and emotions. For this reason, a hybrid method is integral to working with geosocial data, particularly geovisualizing them. Mutual complementation can take place in the process of analyzing data as well as in final analytical outcomes. Geographic visualization is traditionally regarded as a quantitative approach; however, it also draws on various approaches in many other disciplines, including geography, information visualization, exploratory data visualization, and GIS (Dykes and others 2005). In other words, geographic visualization is a hybrid process that involves both quantitative and qualitative analyses, and we would like to use this hybrid power of geovisualization fully in researching LBSN data.

Following quantitative and qualitative approaches at the same time will be practically challenging, but particularly beneficial for analyzing Foursquare data. This database already contains "codes," based on venue categories divided up for user convenience. For example, a user can collect and analyze data on venues or tips in a particular area or category, which is a merit that is not offered by other social media. Second, the qualitative coding approach enables more profound analysis of data included in non-standardized texts such as tips. This also allows analysis of the context of the data beyond what is provided by the database format of the venue category. Third, the social media text requires an analytical approach for codes that contain human emotions. A qualitative approach is essential for processing numerous emoticons, slang expressions, tweet lingos, symbols, and abbreviations, which well complements the spatial and quantitative analysis of geosocial media data.

Code Cluster Analysis Design

The process of code cluster analysis consists of three key stages: collecting data using a Web API, GIS analysis for finding clusters using the kernel density method, and qualitative analysis for assessing the geographical characteristics of each cluster. Each analysis includes hybrid procedures that combine different approaches, such as computer analysis, GIS spatial analysis, and qualitative text analysis. Major characteristics of each analytical stage are shown in Figure 1.

The data collection stage involves Web programming using API and processing of collected data and the spatial database. A computing approach for converting compiled data into spatial data is primarily used at this stage, which consists of collection of venues and tips using Foursquare API and Python programming, together with implementation of a spatial database using *Spatialite*. To collect Foursquare venue and tip data in specific regions, the Foursquare server was accessed using Python programming. Foursquare provides API-based access to information about venue, tips, and users. The approach to the Foursquare API is available in various programming languages; we used Foursquare Python[2] and developed a program that can collect data such as venue, tips, and check-in statistics. Box 1 shows some of the code to collect the venues and tips.

As shown in the above Python code, Foursquare's search API allows searching for venues and tips around specific locations; however, it limits the number of data that can be collected at one location. For this reason, the data collection was performed with different searching radius values to obtain the detailed information. For example, we use a values of 300 m in the urban area where the venues are concentrated in data collection and applied a search range of 500 m to other areas. After data collection with a Python program, data cleaning is carried out. It consists of converting the text information collected by the Python program into GIS data.

The information on Foursquare can be divided into information on the venue, which is the place where the users check in, and tip information, in which the user can comment on the venue. All of the information includes geographic coordinate values of latitude and longitude. The

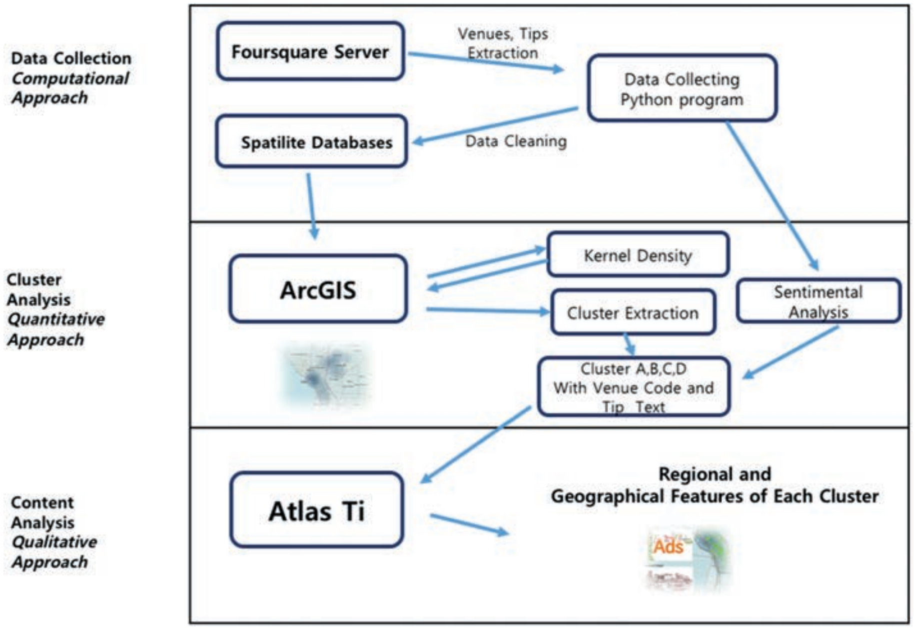


Figure 1. Code cluster analysis flow
Source: Figure by author with the exception of the heatmap image and code cloud image of Cluster A © OpenStreetMap and contributors. CC BY-SA 2.0.

Box 1. Example of the code to collect data in Foursquare Python

```
venues = client.venues.search(params = {'ll': str(llat)+','+str(llong), 'radius': radius})
```

properties of the venue include the name of the venue, the date of the venue creation, the category information of the venue, the address and location information, and statistical values related to the venue, such as the numbers of check-ins and check-in users. Among these, the name of the venue and the category information for the venue are important information for distinguishing characteristics of venues. Check-in is a sign that Foursquare visited the venue and provides a check-in function. Check-in allows users to share check-in information with their friends. A high frequency of check-ins means a lot of people visit and a venue is a well-known place in the area. Foursquare users leave a tip on venue, and a venue's tips include feedback from users about places and interest information. A large number of check-ins means a venue is popular and many people visit, and many tips mean visitors have the intention and interest to share the venue with others.

The search results were presented in JavaScript object notation (JSON). Foursquare offers venues and tips, as well as numerous complex attributes regarding users. Attributes used in the table were venue ID, venue name, category 1, category 2, category 3, category ID, number of check-ins, number of users, number of tips, latitude, and longitude. Venue category consisted of 10 primary categories, and the category values were used to collect venues related to coffee. Next, statistical attributes regarding venues, including number of check-ins, number of users, and number of tips, were used as kernel weight values for cluster analysis. Finally, latitude and longitude values were used to convert venues into spatial coordinates. Data collected regarding tips were tip ID, date of tip creation, tip text, venue name, category 1, category 2, category 3, category ID, latitude, and longitude. In the venue tip table, venue categories were used to select tips related to coffee, and latitude and longitude values were used to convert tips into spatial coordinates. Tip texts were particularly used as data for later qualitative analyses.

The second stage is a heatmap analysis for finding venue clusters among the collected venue data corresponding to a category. Methods for creating maps using point data can generally be categorized into heatmap and hotspot analyses. Although they are similar in terms of creating maps with point data, there are some differences between the two. Whereas a heatmap shows the density and concentration of point data, hotspot analysis assesses whether the level of concentration of point data is statistically significant. For example, hotspot analysis is appropriate for analyzing an area where crime and disease occur more frequently than in other areas, with statistical significance. On the other hand, heatmap analysis is suitable for showing the levels or differences in data when the distribution of point data is already very dense. Density analysis is most widely used for analyzing heatmaps. Density analysis of point data can be also categorized into point density and kernel density. In point density, the number of data around a point is calculated and assigned a cell value, and a weight is added to the point data for a weighed item. In the case of kernel density, a kernel function referred to as a Gaussian curve is applied to calculate the density.

The collected venues are represented point features, and kernel density estimation is applied to present the quantitative spatial distribution of their density. The kernel function measures the density of the point data contained in a certain analysis radius and expresses it as a kernel function *K*, as shown in the equation

$$f(x, y) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{1}$$

where $n =$ total number of points, $h =$ the bandwidth to determine the amount of smoothing, $K =$ kernel function, $d =$ data dimensionality, $x =$ location of estimated point, and $x_i =$ location of known point *i*.

The analysis of the kernel is performed by assigning the highest value to the nearest object and then decreasing the values as the distance becomes greater. This is an efficient analysis method for locating clusters in which venues are concentrated. The bandwidth (*h*) for the search radius is calculated by applying various values to the results and comparing them with the actual values. The result of kernel density is presented as colour: the dark colour indicates the weight of the cell near the venues, and the border of the dense raster data can be converted to polygon data. Using the converted polygons, the venues and tips that were contained in the cluster region were reselected, and the characteristics of these venues were used for the further analysis of each clustered region.

The third stage is qualitative analysis, where geographical characteristics are closely examined based on the venues and tips obtained from the cluster analysis. Whereas the heatmap acquired from cluster analysis identifies areas where specific venues are concentrated in the study area, qualitative analysis is the process of analyzing the nature of hotness based on detailed examination of the tip texts related to the corresponding cluster. The outcome of the cluster analysis in the previous stage is important, since we can only focus on analyzing qualitative data located within the identified clusters. Various qualitative methods are applied. Coding analysis is conducted to identify key themes, topics, or patterns. We particularly followed the *grounded theory* approach (Strauss and Corbin 1997; Charmaz 2008). Grounded theory is a method of generating "theories" from grounded data, and it is often considered as a blanket term for coding and analyzing qualitative data (Urquhart 2012). No code was pre-defined; we developed the codes completely from scratch, and let key topics and themes emerge. Coding analysis can be exploratory, since it has an inductive nature; however, it offers a clear systematic approach to qualitative analysis as well. For this reason, coding analysis is quite an effective way to find out spatial as well as thematic characteristics and patterns between clusters shown in the heatmap. Based on the tips on the corresponding topic and venue, coding analysis focused on analyzing the embedded contexts of tips from the selected areas with characteristics of concentration and density. In addition, sentiment analysis is applied to the text tips, using Python sentiment library. These integrated analytical processes enable analysis with a stronger emphasis on spatial topics among the large amounts of text data distributed in non-specific patterns. In particular, qualitative analysis focuses on finding the context of geographic characteristics that could have been overlooked in conventional text mining, and content analysis counts the frequency of words or phrases using statistics (e.g., the content clouds of Cidell 2010).

Case Study: Code Cluster Analysis of Coffee Shops in Seattle

GENERAL ANALYSIS OF FOURSQUARE VENUES

A case study was conducted with coffee shops in Seattle based on code cluster analysis as a hybrid approach combining qualitative and quantitative methods using LBSN data from Foursquare. We ran the program for two days, 5 and 6 January 2016, to collect all Foursquare venue data that had been generated since the launch of Foursquare in 2009. We collected a total of 34,519 venue data for Seattle created between March 2009 and 6 January 2016. Figure 2 shows the complete procedure of the hybrid approach we used for our case study, from data collection to process and analysis.

Table 1 displays the overall characteristics and the distribution of the venues collected in Seattle. The categories with the largest numbers of venues were Professional, Shop, Service, and Food. The Food category had the largest numbers of check-ins (visits), users, and tips. The Food category includes various types of restaurants, cafes, and coffee shops, and the coffee shops had the largest number of venues. Seattle is famous for coffee chains, such as Starbucks and Tully's. In this study, 701 venues included in the categories of Coffee Shop and Café were used to perform cluster analysis. The overall spatial distribution of the venues located in Seattle is shown in Figure 3. Neighbourhoods such as Capitol Hill, Queen Anne, University District, Industrial, and Central District show high densities. Among the coffee shop venues collected in Seattle, Starbucks had the highest percentage with 122 venues (17%), followed by coffee shops with branches such as Tully's (101 venues, 15%) and independent coffee houses (478 venues, 68%).

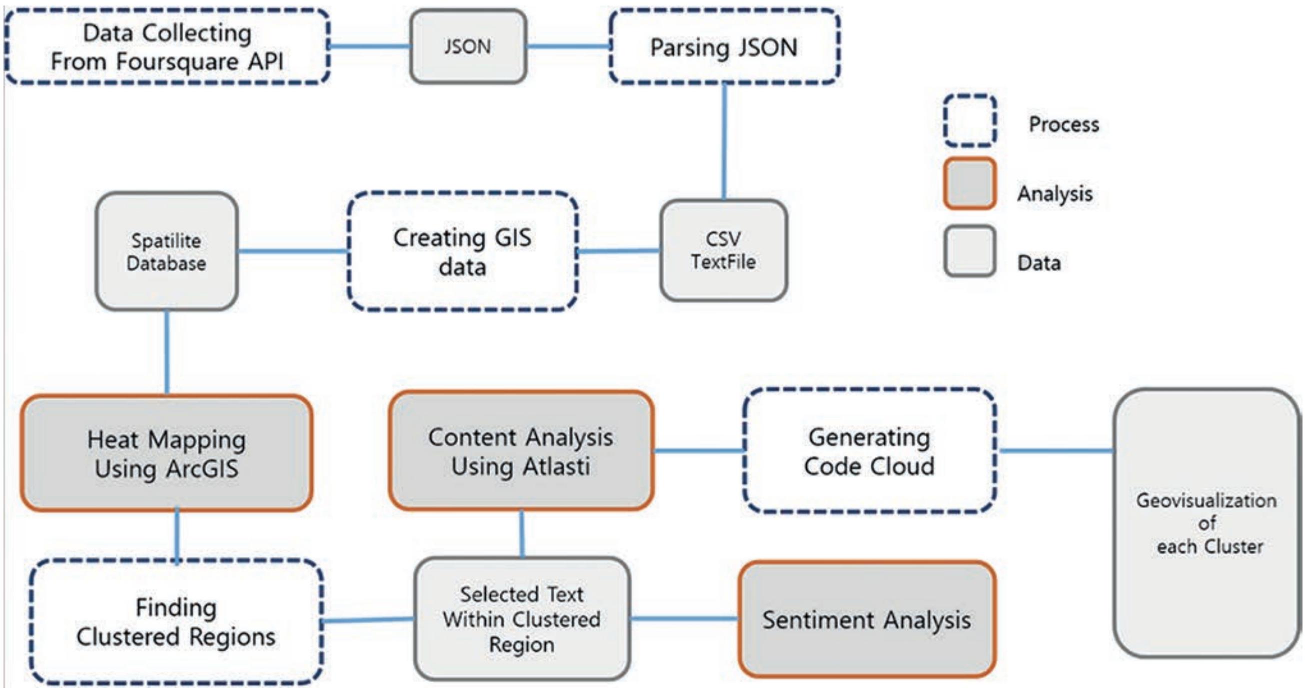


Figure 2. Procedure for processing and analyzing foursquare data

Table 1. Distribution and characteristics of venues in Seattle according to primary categories

Category Name	No. of Venues	%	Average of Check-Ins	Sum of Check-Ins	Average of Users	Sum of Users	Average of Tips	Sum of Tips
Professional	7177	25.69	254	1,820,175	67	477,344	1	5,480
ShopService	6700	23.99	493	3,300,169	151	1,010,476	2	13,684
Food	3697	13.24	861	3,182,207	398	1,471,825	9	33,345
Outdoors	2507	8.98	706	1,770,668	205	513,073	2	4,771
Travel	2442	8.74	444	1,085,149	142	346,336	2	3,868
Art	1689	6.05	655	1,105,569	318	537,663	3	4,752
Residence	1404	5.03	186	260,534	14	20,319	0	512
Nightlife	1346	4.82	595	801,352	266	358,152	5	6,346
College	960	3.44	528	507,227	113	108,741	1	1,078
Event	10	0.04	8	79	7	73	0	0

Heatmap analysis was performed with a total of 701 point data consisting of coffee shops and cafes in Seattle. Kernel density was used for analysis with the number of user check-ins as the weight value. The results of heatmap analysis are shown in Figure 4, which clearly indicates five clusters, two with high densities in the central regions (Cluster A) and three in outer areas (Clusters B, C, and D). The regional characteristics of the five clusters with high concentrations of coffee venues were analyzed by re-categorizing the venues and tips according to regions. We additionally conducted a sentiment analysis, and the results of each cluster are shown in Figure 5. Table 2 summarizes the statistics of four identified clusters.

As discussed in the previous section, content analysis is one way of analyzing qualitative data, and it counts and measures the frequency of qualitative data such as words. Content analysis is useful for initial data exploration and offers a quick overview of the data set; however, it alone does not allow us to fully understand the significance data. For example, the most common keywords in our case study were “coffee” and “Starbucks,” but it does not provide the contextual meanings, in other words, the reasons that was the case. Further analyses and interpreta-tion are needed. For example, we cannot determine whether the word “coffee” has a positive or negative con-notation unless we find out how it is used in a sentence or

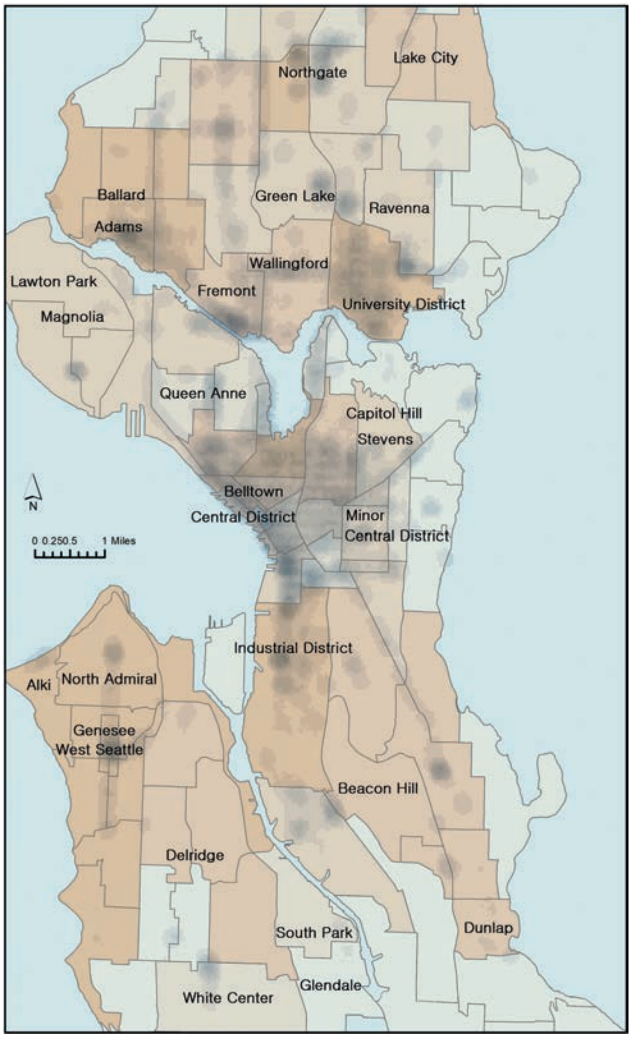


Figure 3. Spatial distribution of coffee venues in Seattle Source: © OpenStreetMap and contributors. CC BY-SA 2.0.

a paragraph. In this regard, qualitative analysis, such as coding analysis, can help us discover the contexts of qualitative data, and the outcome of coding can be combined with the results of spatial analyses (Cope 2005; Altheide and Schneider 2013). Therefore, the coding analysis also allowed us to systematically analyze the Foursquare tips of the geographically focused areas (e.g., four clusters) and to explore multiple interpretations of those areas.

Following the grounded theory approach, we implemented two levels of coding procedures in our research. The initial step was in vivo coding and the second step was analytic coding. The in vivo codes are words or phrases taken directly from the texts, and they are thought of as being emic to the original data. The analytic codes are more interpretive and reflective about the description of data (Cope 2005). Coding does not just add tags to data, but also develops general ideas or key themes that better represent the contextual meanings of data. In addition, we can visualize both in vivo and analytic codes through code clouds as a way of visualizing qualitative data and analyses. In other words, qualitative analysis, and in particular coding analysis, should have preceded the creation of code clouds.

Code clouds are generated in the same way as word clouds, where the importance of each word is represented with a text size or colour that is proportional to the frequency of the word (Ward and others 2010). However, there is a significant difference between the two. Where word clouds use unanalyzed raw data, code clouds make use of qualitative "codes" as input data that enable readers to see what the data actually mean. In other words, instead of showing the words themselves or their frequencies, code clouds are designed to represent the meanings they contain (Jung 2015). Technically, there are various tools for visualizing texts, such as TwitScoop,[3] TagCrowd,[4] Wordle,[5] and Jigsaw.[6] However, we chose Wordle to generate code clouds, for its flexible capacities for visualizing text analysis, such as word counting, and various layout options in terms of placement, shape, and colour choices (Steele and Iliinsky 2010).

We carefully examined the Foursquare tips in each cluster and conducted coding analysis without any pre-determined categories or themes. Instead, we paid particular attention to keywords (i.e., in vivo codes) and explored multiple possibilities and interpretations that are naturally developed from the original tips (i.e., analytic codes). Clusters provided us with geographically focused areas that we could use to systematically analyze raw Foursquare data and started to draw out recurring key themes of each cluster. As a result, we identified the most common codes drawn from the coding analysis. Table 3 shows the top 30 codes generated in each cluster. We will now demonstrate the final result of code cluster analysis by providing the evidence of characterizing themes of each cluster. The integration of code clouds and the mapping of the location of original Foursquare tips data will create code cluster visualizations for each cluster.

CLUSTER A: DOWNTOWN

This is the central business district (CBD) of Seattle, with a total 69 codes. The top 30 codes of cluster A are identified in Table 3. The number next to the code name indicates how many times a particular code was generated in Foursquare tips in the responding cluster. For example, in Cluster A, "ads" was the most common code, which was created 215 times. The second most frequently generated code was "promotion," which was closely related to advertisements as well. We observed many tips related to advertisements and promotions for deals or new products, and they were often from a large coffee company such as Starbucks and Seattle's Best Coffee. Examples included "Hot tip for morning travelers: bring Starbucks VIA Instant

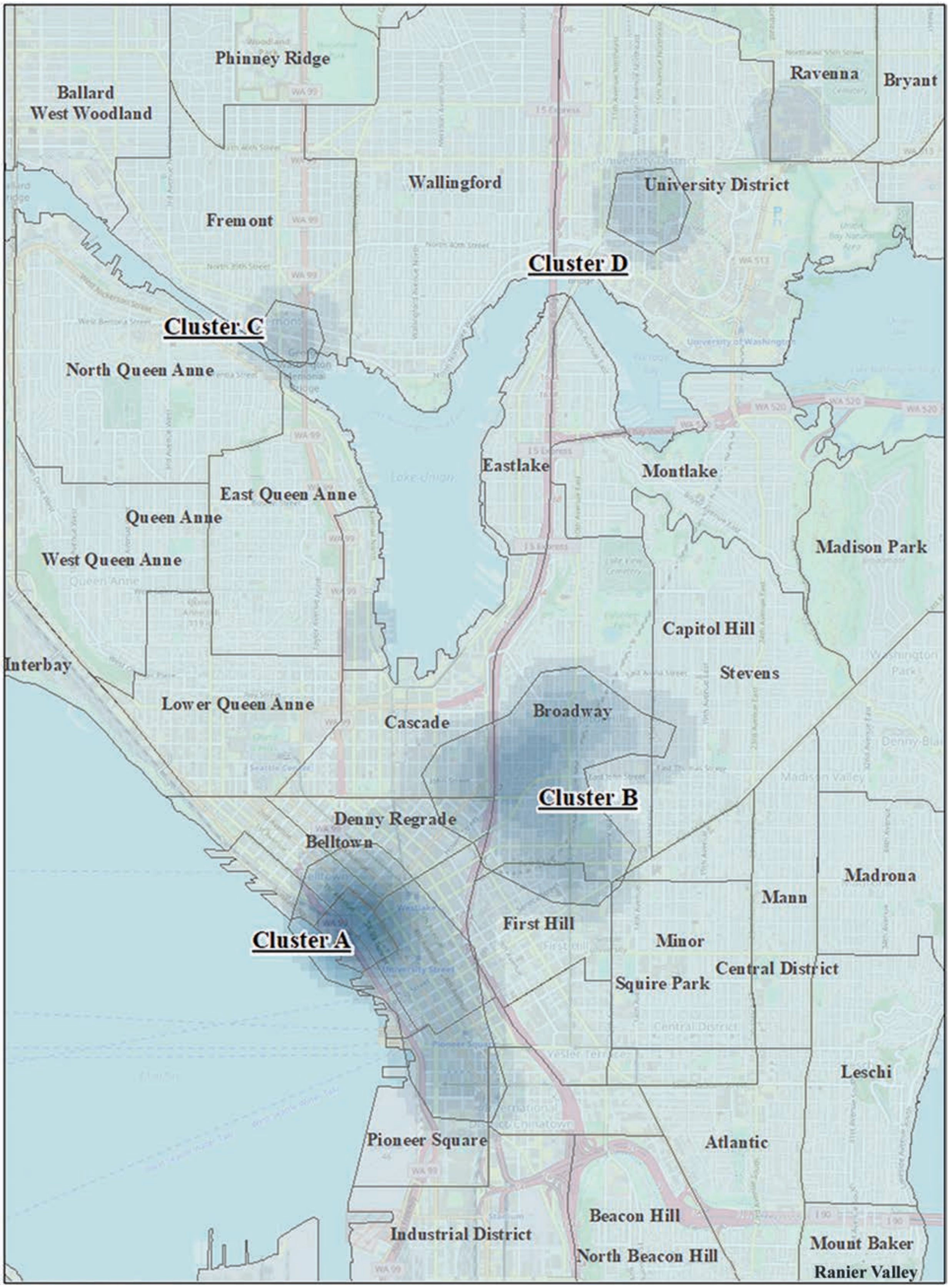


Figure 4. Result of heatmap analysis
Source: © OpenStreetMap and contributors. CC BY-SA 2.0.

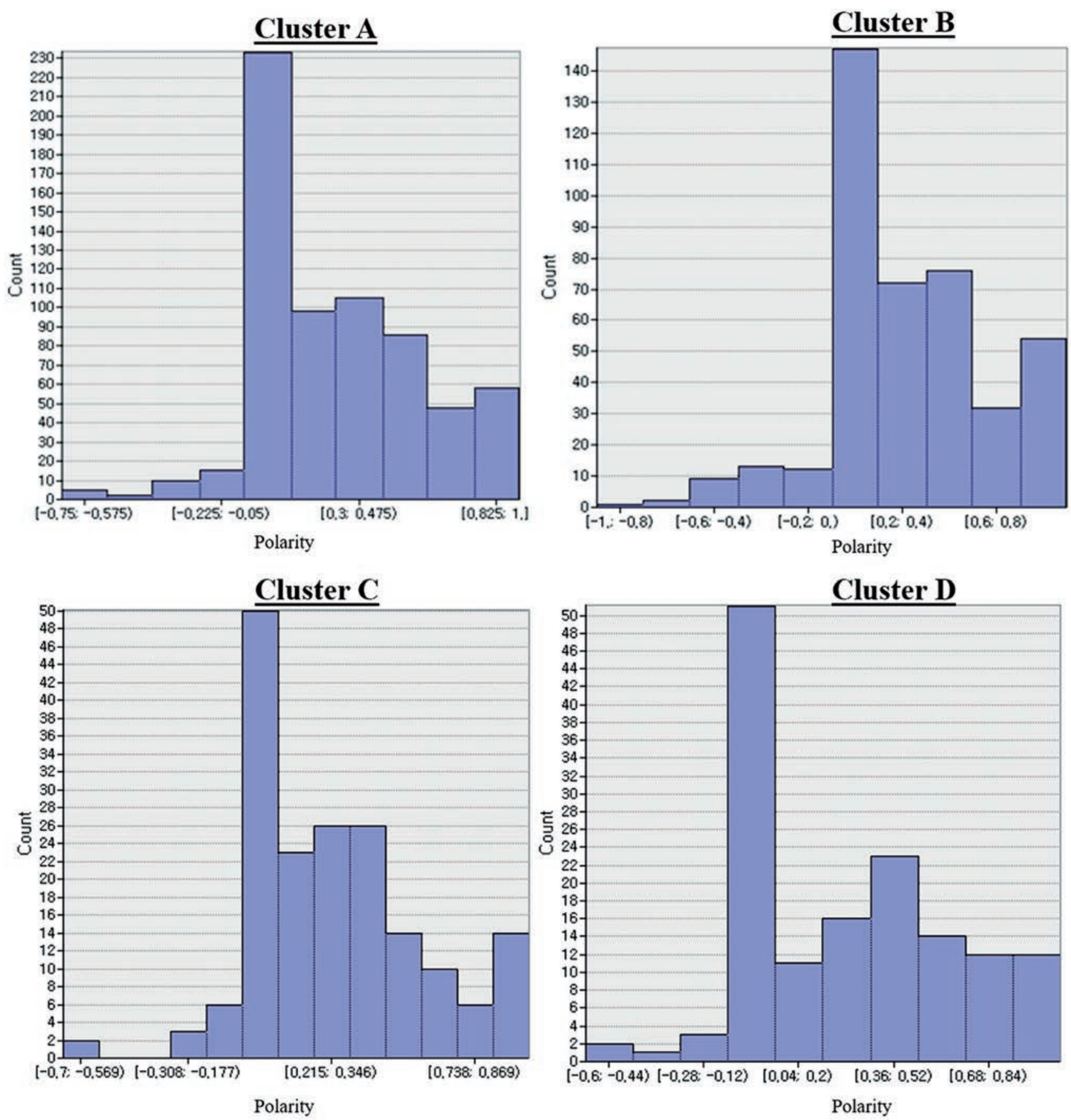


Figure 5. Sentimental analysis results of each cluster

Coffee packets with you in case the bistro coffee machine is out of service" and "Try one of our holiday favorites." People also often commented about the services they received: "Great customer service," "Nice owners and staff," "Great staff – always friendly," and "By far the best staff of any Starbucks." However, there were a similar number of complaints about the services in respect to the slowness of service, a long wait, food, coffee, price, or even Wifi: "The girls here are very distracted, don't make an effort to expedite the order. Always a long wait," "The slowest cooks you will ever encounter. Not to mention overpriced," and "It's so hard to get on wifi!!" Figure 6 shows the final outcome of code cloud analysis of Cluster A, which represents the context of data as a visualized outcome of coding analysis with the location of the Foursquare tips inside cluster A. As we can see from the code clouds image, commercial advertisements and promotions are prominent in this cluster.

Table 2. Statistics of each cluster

	Number of Venues	Number of Users	Number of Check-Ins	Number of Tips	Ave of Polarity	Number of Positive Tips	Ave of Positive Tips	Number of Negative Tips	Ave of Negative Tips
Cluster A	127	113,108	269,198	660	0.29	420	0.49	38	−0.26
Cluster B	62	61,083	202,775	418	0.29	257	0.52	37	−0.33
Cluster C	13	4,630	17,323	145	0.30	88	0.51	7	−0.28
Cluster D	13	10,312	37,643	181	0.29	128	0.43	11	−0.24

Table 3. List of top 30 common codes in each cluster

Rank	Cluster A		Cluster B		Cluster C		Cluster D	
1	**ads**	**215**	**brew method**	**116**	**best**	**14**	**food**	**227**
2	promotion	75	ads.	94	baristas	13	ads.	29
3	food	44	food	33	ads.	12	taste	26
4	service	38	pastry	26	promotion	12	atmosphere	19
5	best	27	best	23	brew method	11	suggestion	16
6	sandwich	27	taste	21	pastry	11	pastry	15
7	suggestion	24	atmosphere	21	cappuccino	7	promotion	11
8	mocha	23	amenity	21	espresso	7	service	11
9	taste	23	promotion	17	new product	7	latte	9
10	complaint	22	service	16	service	7	complaint	8
11	latte	21	mocha	16	amenity	6	amenity	6
12	pastry	21	Seattle	14	mocha	6	history	6
13	baristas	19	latte	14	atmosphere	5	mocha	6
14	praise	15	roast	13	chocolate	5	music	6
15	amenity	14	espresso	12	coldbrew	5	chai	5
16	atmosphere	13	drink	12	comments	5	praise	5
17	espresso	13	baristas	12	complaint	5	best	4
18	Wifi	13	sandwich	11	Fremont	5	tea	4
19	price	12	praise	11	latte	5	Wifi	4
20	roast	12	music	11	aeropress	4	baristas	3
21	brew method	11	work place	10	excellent	4	espresso	3
22	location	9	Wifi	8	roast	4	magazine	3
23	comments	8	complaint	8	Wifi	4	update	3
24	arts	7	coldbrew	8	clover	3	americano	2
25	bagel	5	americano	8	happy hour	3	cheap	2
26	cappuccino	5	location	7	lemonade	3	drink	2
27	chocolate	5	tip	6	nutella	3	hour	2
28	drink	5	suggestion	6	sandwich	3	large portion	2
29	history	5	restroom	6	taste	3	location	2
30	tip	5	cookie	6	tea	3	people-watching	2

CLUSTER B: NEAR CAPITAL HILLS

Many contrasting cultures coexist in the Capital Hills neighbourhood, from century-old mansions to a great number of apartment buildings, trendy bars, and bistros. This area is also well known as the hub of the LGBT community in the city. There were 72 codes. As in cluster A, there were many comments related to the advertisements (“ads”); however, it has a distinctive difference. The most common and noticeable code was “brew method,” and there were many tips about it (see Figure 7). Some examples included “Don’t order a pour over unless you have 10 minutes to spare,” “Get the French press,” “Get the nitro pushed cold brew,” “I love this place! The roasts are incredibly good, esp the French Press,” and “Extremely impressed by this place. A variety of brew methods …” Closely related to that, it was not surprising to see many codes representing varieties of coffee in this cluster, such as “mocha,” “latte,” “roast,” “espresso,” “Americano,”

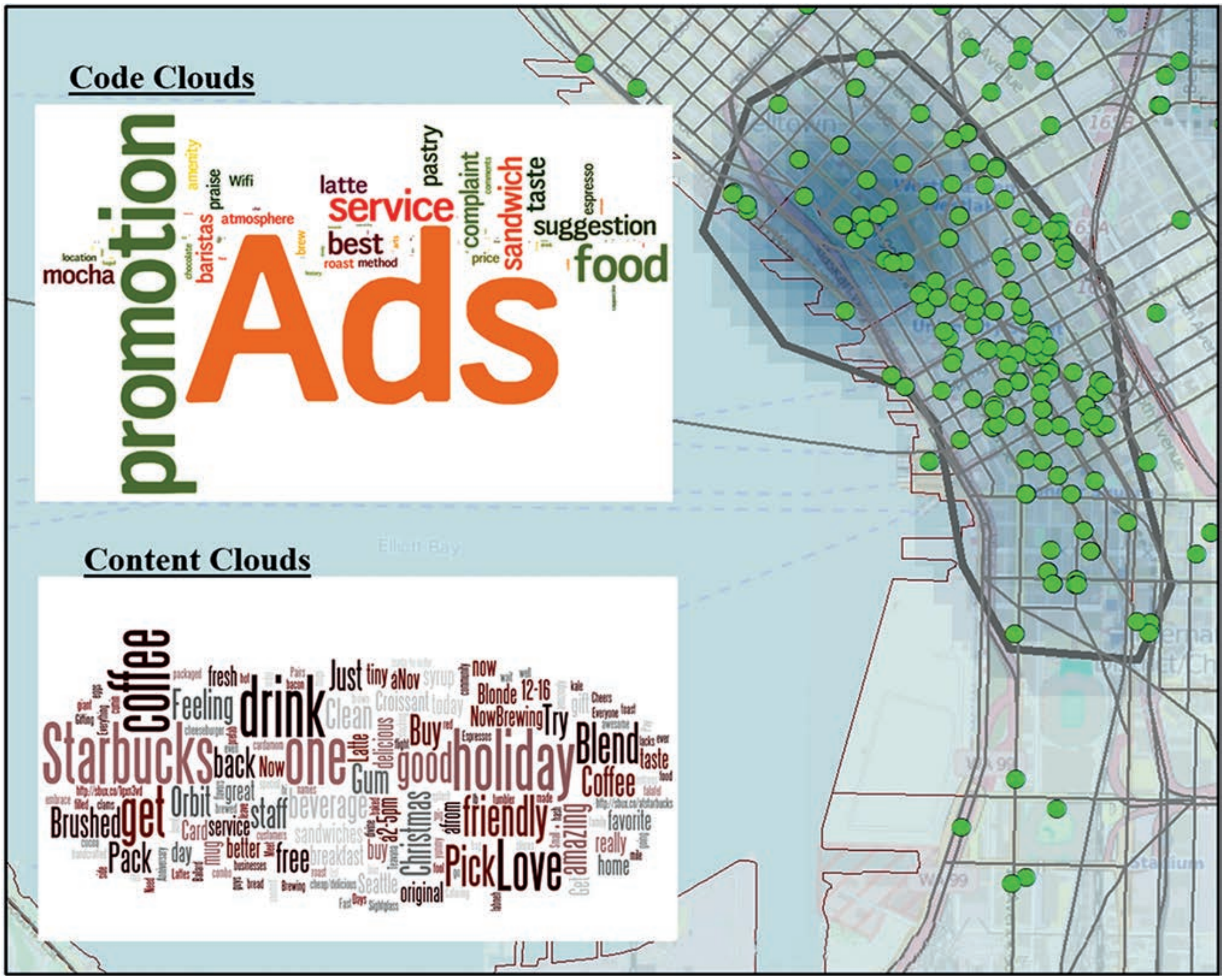


Figure 6. Code cloud of Cluster A
Source: © OpenStreetMap and contributors. CC BY-SA 2.0.

and even “coldbrew.” The importance of “atmosphere” inside coffee shops and “music” played was obviously important as well. For example, “I did the coffee, the chill vibe, and the groovy tunes. My new favorite coffee place in Seattle,” “This place has a really, really good playlist,” and “Jamming tunes, comfortable chairs, bike rack in the front and the best coffee on Seattle … what more could you ask for?” Interestingly, people also mentioned “restroom” and sometimes even shared the passcode. For instance: “The restrooms always smell,” “Women’s bathroom code as of April ’14 is 02468XX,” and “Great coffee and great work spaces … but no public restrooms?!” Overall, many tips in cluster B gave us the impression that they were from coffee connoisseurs who come to a coffee shop for a particular type and taste.

CLUSTER C: FREMONT

Fremont is centrally located in the city of Seattle, and it presents a unique neighbourhood identity and an eclectic vibe, as indicated by its nickname “Center of the University.” Once a neglected community, it has become an attractive place for artists and students and a gentrified neighbourhood for young married couples and singles. Fremont had a total of 52 codes, a smaller number than for the other three clusters, and these codes were much less grounded in the original tips data. Nonetheless, the cluster presented two unique characteristics: social aspects of coffee and people’s love for and attachment to a place. Similarly to Cluster B, there were many comments about brewing methods and the types of coffee (see Figure 8). At the same time, tips generated in Fremont showed a lot of pride and love for their neighbourhoods and coffee shops. In particular, there were numerous comments related to the codes of “best” and “Fremont.” For example, “The best coffee in Seattle,” “The best coffee shop this side of Queen Anne hill,” “Fremont needed a place like Milstead & Co,” and “An authentic Fremont experience.” A strong and positive identity for the neighbourhood was vividly reflected in the digital world.

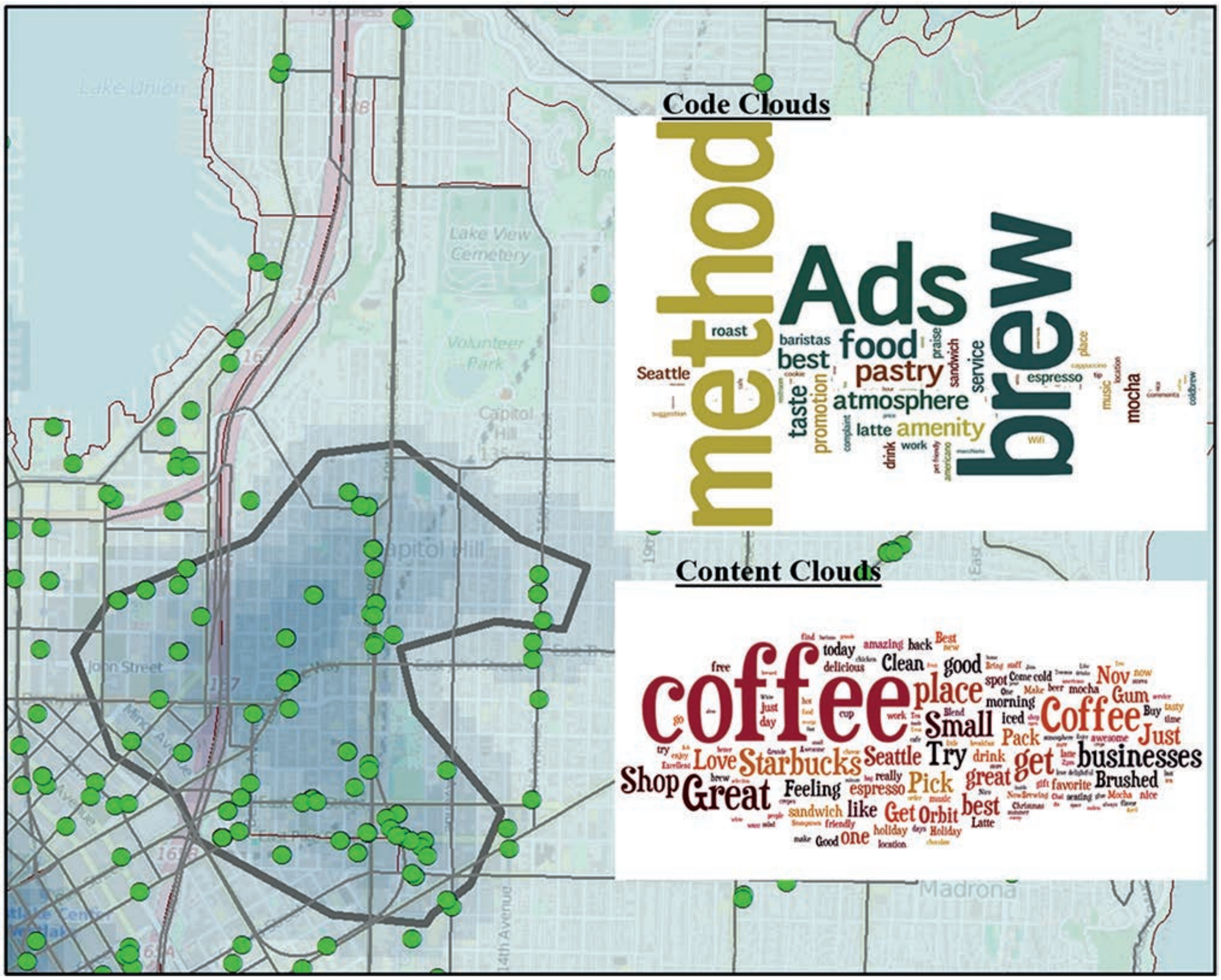


Figure 7. Code cloud of Cluster B
Source: © OpenStreetMap and contributors. CC BY-SA 2.0.

The barista was an important reason to visit a particular coffee place. Customers often knew the names of baristas and acknowledged and appreciated their skills, knowledge, and personalities. Many people had their favourite baristas, and seemed to actively interact with them. For example, "Try to get Nick or Alexa to make your drinks. White Mocha with whole milk and caramel sauce with whip," "Talk to Andrew, he's nice and will chat you up," and "Best baristas evaaaaaar." Cluster C demonstrates that coffee shops have been instrumental in facilitating social interactions, fostering a sense of belonging, and preserving and building strong community identity and pride.

CLUSTER D: UNIVERSITY DISTRICT

Cluster D is located in the University District (U-district), which includes the University of Washington – Seattle campus. U-district is a magnet for the younger population, including college students, but it is also a popular lively neighbourhood for all ages. A wide variety of restaurants, bars, and coffee shops are located especially along University Avenue, which locals call "The Ave." Foursquare tips also well elicit some characteristics of this university village. We initially collected Foursquare tips that were related to the "coffee" category; however, "food" seemed to be more important than coffee in this cluster: oat bar, waffle, samosas, salad, croissant, sandwich, tofu banh mi, and many more (see Figure 9). Not only the variety and taste of food, but also the portion and price seemed to matter. For example, "Their half sandwiches are really whole. Sweet deal," "Get the early bird breakfast – cheap (all under $5), yummy, and pretty filling. Enjoy your morning solitude here," and "Breakfast is nice but a little pricey." The atmosphere of the shops was an important factor, especially attuned to students' tastes, and it seemed quite noticeable for visitors as well. For instance, "Nice place to sit, but too noisy for studying," "Great

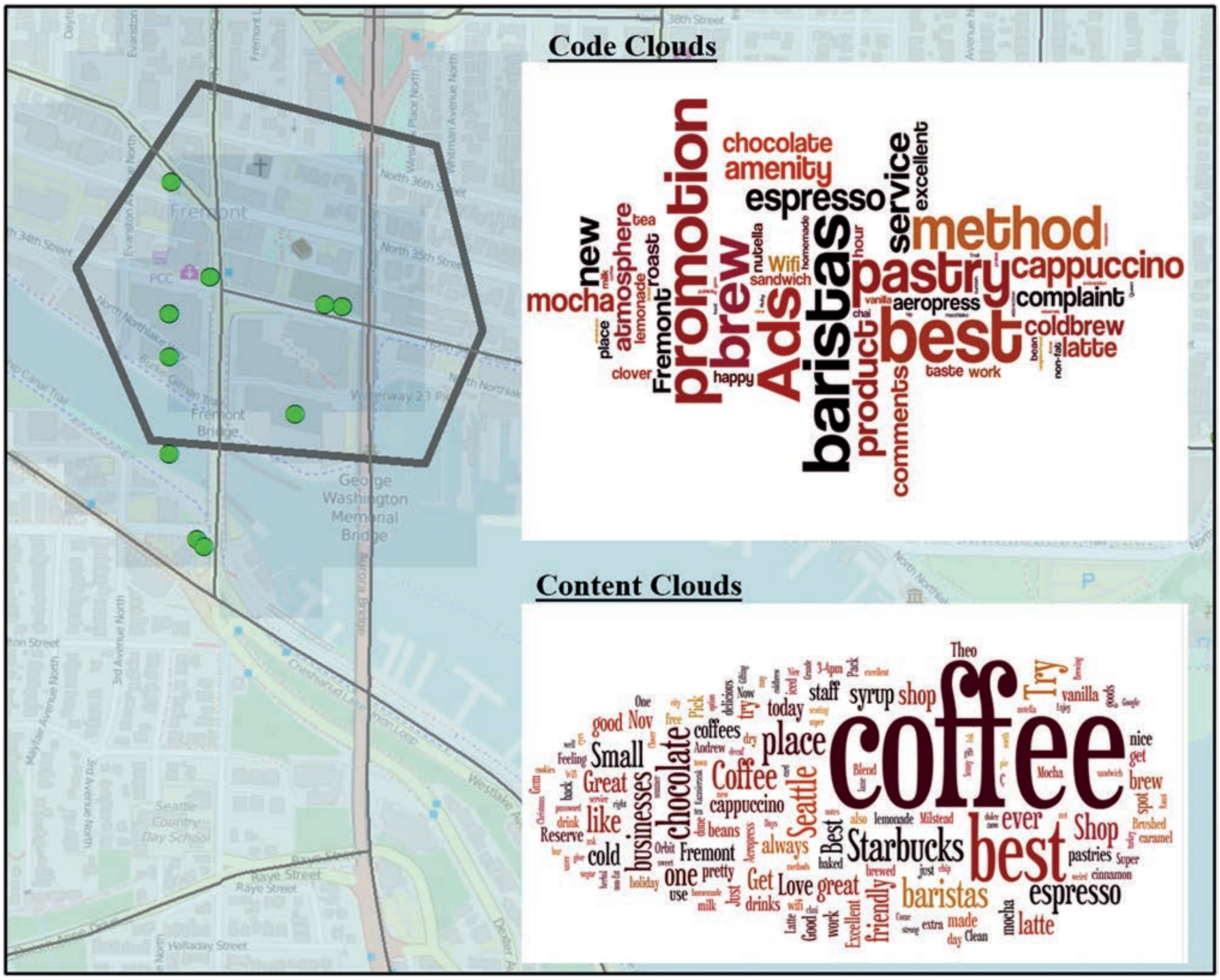


Figure 8. Code cloud of Cluster C
Source: © OpenStreetMap and contributors. CC BY-SA 2.0.

place to get work done. Free refills on some menu items. Open late,” and “Looks like a perfect place for study, pretty young UW crowd.” Amenities related to studying were often mentioned as well, such as the availability of enough power outlets, WiFi, large tables, lights, and quiet. Cluster D clearly shows the characteristics of the geographic place that coffee shops are located in. People who come to coffee places in the U-district seem to look for a cozy and college-student-oriented atmosphere where they can feel at home and comfortably spend time or study, and, more importantly, they can fill up with tasty, good-portion-size food.

Conclusion

This article presents the innovative, hybrid analytical method of code cluster analysis for studying characteristics of LBSN data, especially Foursquare data, which constitutes a new approach to intersecting spatial analysis, geographic visualization, and qualitative research. We acknowledge that the geographic analysis of LBSN data requires understanding the computing environment, such as the network and database, as well as complex applications of various analytical tools, such as the GIS analysis of spatial data that include coordinate values. However, we also want to highlight the importance of a qualitative approach for fully analyzing and interpreting the contextual meanings of qualitative data. The potential of code cluster analysis was particularly validated by a case study of Seattle coffee shops.

In summary, data were collected using Python and Foursquare API to assess the venue distribution characteristics and the statistical characteristics that corresponded to 10 categories. Based on cluster analysis of the venues that correspond to the topic of coffee, four cluster areas were selected. With venue texts that reflect user opinions about the coffee venues in the four clusters, qualitative coding

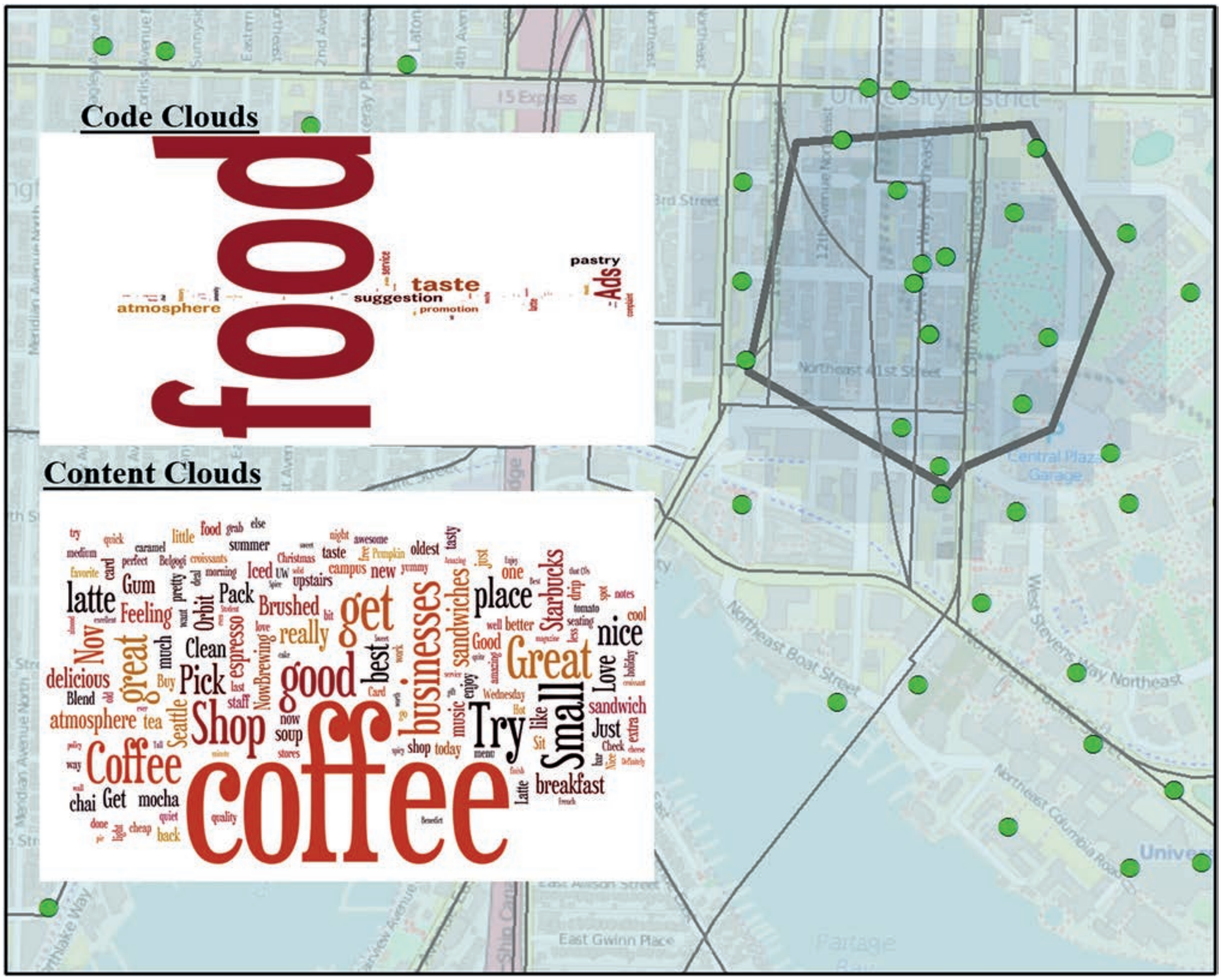


Figure 9. Code cloud of Cluster D
Source: © OpenStreetMap and contributors. CC BY-SA 2.0.

analysis was performed on their specific geographic characteristics expressed by Foursquare tips, and an attempt was made at geovisualization of the data as a code cluster.

Code cluster analysis was not intended to determine the correlation between what was said about a venue people visited and the socio-spatial characteristics of hotspot clusters. Also, it is too early and difficult to make a conclusive statement that code cluster analysis revealed new insights about hotspots, especially in relation to number of check-ins and reasons for hotness. However, our goals were to make strong efforts to draw more meaningful and nuanced insights from the LBSN data and to examine complex relational issues by building a more comprehensive interdisciplinary method that integrates quantitative, qualitative, and geovisual approaches. More specifically, the significance of code cluster analysis is in offering a means of comprehensively analyzing and visualizing the spatial and statistical and revealing the qualitative characteristics of geosocial media. In the building of a hybrid method, each phase of the research process, from data collection and cluster analysis, to qualitative coding analysis, contributes greatly to extracting more contextual and nuanced meanings of geosocial data. This also transverses and transcends often-considered constrained boundaries of computational, quantitative, and qualitative approaches in relation to geographic visualization. Advanced GIS analysis such as cluster analysis allows us to identify hotspot areas that we can focus on studying, and the implementation of various qualitative analyses helped us see the characteristics of Foursquare tips and corresponding clusters. To this end, we could elicit various reasons for hotness in our four identified clusters, from brewing methods or the quality of service and amenities to baristas, music, and food. These are examples of contextual and qualitative meanings that we could draw from the Foursquare data using a code cluster analysis. This is a new call for a more hybrid and "grounded" understanding of geosocial data that opens up a new epistemological and methodological reframing of geographic and geovisual research with LBSN and spatial big data in general.

However, further in-depth research is required. One trajectory can be a method for automating and expediting the analytical process. The development of computational text analysis with the use of advanced algorithmic approaches in linguistics and computer science might provide a fast and automated capacity to extract meaning or contexts from social and spatial media data (Wiedemann 2013; Hahmann and others 2014; Zafarani and others 2014). These text mining and topic modelling geosocial data approaches will be particularly beneficial for tackling the problems of analyzing large volumes of these data and finding linkages of such insights to places and people. On the other hand, we also intend to continue our research efforts at bridging the gap between qualitative, quantitative, and geovisual analyses to preserve the contextual meanings and to cartographically visualize the degrees and reasons of data (e.g., hotness) without any transformation in the research process. Considering the emerging discussions in the geographic analysis of big data, code cluster may provide one way we can articulate and elicit digitally mediated activities that are happening in both material and virtual spaces. It may also provide a potential solution for exploring and identifying thematic patterns and qualitative meanings contained in socio-spatially mediated spaces with a power of visualization (Graham and Shelton 2013; Kelley 2013; Kinsley 2014).

Acknowledgement

The funding for this article was provided by Namseoul University.

Author Information

Ilyoung Hong is an assistant professor in the Department of GIS Engineering, Namseoul University, Cheonan, South Korea. He has a PhD in geography from the University of Buffalo. He earned his master's degree and bachelor's degree in geography education at Seoul National University. His research interests include spatial analysis, the geospatial Web, and digital cartography. E-mail: ilyoung.hong@nsu.ac.kr.

Jin-Kyu Jung is an associate professor in the School of Interdisciplinary Arts and Sciences at the University of Washington—Bothell. He is an urban geographer whose interdisciplinary research program contributes to qualitative Geographic Information Systems (GIS), children's urban geographies, and community-based research and planning. He has particularly focused on developing *new* ways of expanding these critical, qualitative, and creative capabilities of GIS and geographic visualization in conjunction with the social and spatial theorization of (urban) space. E-mail: jkjung5@uw.edu.

Notes

1. As stated at https://foursquare.com.
2. See https://pypi.python.org/pypi/foursquare.
3. See www.twitscoop.com.
4. See https://tagcrowd.com.
5. See www.wordle.com.
6. See https://www.cc.gatech.edu/gvu/ii/jigsaw/.

References

Altheide, D.L., and C.J. Schneider. 2013. *Qualitative Media Analysis*. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781452270043.

Barewald, T.J. 2010. "Prospects for Geography as an Interdisciplinary Discipline." *Annals of the Association of American Geographers* 100(3): 493–501. https://doi.org/10.1080/00045608.2010.485443.

Burns, R., and A. Skupin. 2013. "Towards Qualitative Geovisual Analytics: A Case Study Involving Places, People, and Mediated Experience." *Cartographica* 48(3): 157–176. https://doi.org/10.3138/carto.48.3.1691.

Chainey, S., S. Reid, and N. Stuart. 2003. "When Is a Hotspot a Hotspot? A Procedure for Creating Statistically Robust Hotspot Maps of Crime." In Socio-economic Applications of Geographic Information Science, ed. G. Riggs, and D. Kidner, and S. White, Innovations in GIS, 21–36. London: Taylor and Francis eBooks. https://doi.org/10.4324/9780203301074_chapter_3.

Charmaz, K. 2008. "Grounded Theory in the 21st Century." In *Strategies of Qualitative Inquiry*, ed. N.K. Denzin and Y.S. Lincoln, 203–241. London: Sage.

Cheng, Z., J. Caverlee, K. Lee, and D.Z. Sui. 2011. "Exploring Millions of Footprints in Location Sharing Services." In Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011, ed. L.A. Adamic, R.A. Baeza-Yates, and S. Counts, 81–88. AAAI Press

Cidell, J. 2010. "Content Clouds as Exploratory Qualitative Data Analysis." *Area* 42(4): 514–523. https://doi.org/10.1111/j.1475-4762.2010.00952.x.

Cope, M. 2005. "Coding Qualitative Data." In *Qualitative Research Methods in Human Geography*, ed. I. Hay, 310–324. New York: Oxford University Press.

Cope, M., and S. Elwood. 2009. *Qualitative GIS: A Mixed Methods Approach*. London: Sage. https://doi.org/10.4135/9780857024541.

DeLyser, D., and D. Sui. 2013. "Crossing the Qualitative–Quantitative Chasm II: Inventive Approaches to Big Data, Mobile Methods, and Rhythmanalysis." *Progress in Human Geography* 37(2): 293–305. https://doi.org/10.1177/0309132512444063.

Dykes, J., A.M. MacEachren, and M.-J. Kraak. 2005. *Exploring Geovisualization*. London: Elsevier.

Elwood, S. 2011. "Geographic Information Science: Visualization, Visual Methods, and the Geoweb." *Progress in Human Geography* 35(3): 401–8. https://doi.org/10.1177/0309132510374250.

Elwood, S. 2015. "*Still* Deconstructing the Map: Microfinance Mapping and the Visual Politics of Intimate Abstraction."

Cartographica 50(1): 45–49. https://doi.org/10.3138/carto.50.1.09.

Elwood, S., and K. Mitchell. 2013. "Another Politics Is Possible: Neogeographies, Visual Spatial Tactics and Political Formation." *Cartographica* 48(4): 275–292. https://doi.org/10.3138/carto.48.4.1729.

Gao, H., and H. Liu. 2014. "Data Analysis on Location-Based Social Networks." In *Mobile Social Networking*, ed. A. Chin and D. Zhang, 165–194. New York: Springer. https://doi.org/10.1007/978-1-4614-8579-7_8.

Geertz, C. 1973. "Thick Description: Toward an Interpretive Theory of Culture." In *The Interpretation of Cultures*, ed. C. Geertz, 3–30. New York: Basic Books.

Ghosh, D., and R. Guha. 2013. "What Are We 'Tweeting' about Obesity? Mapping Tweets with Topic Modeling and Geographic Information System." *Cartography and Geographic Information Science* 40(2): 90–102. https://doi.org/10.1080/15230406.2013.776210.

Graham, M., and T. Shelton. 2013. "Geography and the Future of Big Data, Big Data and the Future of Geography." *Dialogues in Human Geography* 3(3): 255–261. https://doi.org/10.1177/2043820613513121.

Hahmann, S., R.S. Purves, and D. Burghardt. 2014. "Twitter Location (Sometimes) Matters: Exploring the Relationship between Georeferenced Tweet Content and Nearby Feature Classes." *Journal of Spatial Information Science* 2014(9): 1–36.

Harris, L.M. 2015. "Deconstructing the Map after 25 Years: Further Engagements with Social Theory." *Cartographica* 50(1): 50–53. https://doi.org/10.3138/carto.50.1.10.

Hawkins, H. 2014. *For Creative Geographies: Geography, Visual Arts and the Making of Worlds*. New York: Routledge.

Hong, I. 2015. "Spatial Analysis of Location-Based Social Networks in Seoul, Korea." *Journal of Geographic Information Systems* 7(3): 259–265. https://doi.org/10.4236/jgis.2015.73020.

Jung, J.-K. 2015. "Code Clouds: Qualitative Geovisualization of Geotweets." *Canadian Geographer/Géographe Canadien* 59(1): 52–68. https://doi.org/10.1111/cag.12133.

Kelley, M.J. 2013. "The Emergent Urban Imaginaries of Geosocial Media." *GeoJournal* 78(1): 181–203. https://doi.org/10.1007/s10708-011-9439-1.

Kinsley, S. 2014. "The Matter of 'Virtual' Geographies." *Progress in Human Geography* 38(3): 364–384. https://doi.org/10.1177/0309132513506270.

Kitchin, R., and M. Dodge. 2007. "Rethinking Maps." *Progress in Human Geography* 31(3): 331–344. https://doi.org/10.1177/0309132507077082.

Knigge, L., and M. Cope. 2009. "Grounded Visualization and Scale: A Recursive Analysis of Community Spaces." In *Qualitative GIS: A Mixed Methods Aapproach*, ed. M. Cope and S. Elwood, 95–114. London: Sage.

Kwan, M.P., and G. Ding. 2008. "Geo-narrative: Extending Geographic Information Systems for Narrative Analysis in Qualitative and Mixed-Method Research." *Professional Geographer* 60(4): 443–465. https://doi.org/10.1080/00330120802211752.

Li, L., M.F. Goodchild, and B. Xu. 2013a. "Spatial, Temporal, and Socioeconomic Patterns in the Use of Twitter and Flickr." *Cartography and Geographic Information Science* 40(2): 61–77. https://doi.org/10.1080/15230406.2013.777139.

Li, Y., M. Steiner, L. Wang, Z.L. Zhang, and J. Bao. 2013b. "Exploring Venue Popularity in Foursquare." In INFOCOM, 2013 Proceedings IEEE, 3357–3362.

Mitchell, L., M.R. Frank, K.D. Harris, P.S. Dodds, and C.M. Danforth. 2013. "The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place." *PLoS One* 8(5): e64417. https://doi.org/10.1371/journal.pone.0064417.

Noulas, A., S. Scellato, C. Mascolo, and M. Pontil. 2011. "An Empirical Study of Geographic User Activity Patterns in Foursquare." In *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*, ed. L.A. Adamic, R.A. Baeza-Yates, and S. Counts, 570–573. AAAI Press.

Schoepfer, I., and S.R. Rogers. 2014. "A New Qualitative GIS Method for Investigating Neighbourhood Characteristics Using a Tablet." *Cartographica* 49(2): 127–143. https://doi.org/10.3138/carto.49.2.1810.

Schuurman, N. 1999. "Critical GIS: Theorizing an Emerging Science." Monograph 53. *Cartographica* 36: 1–99.

Stefanidis, A., A. Crooks, and J. Radzikowski. 2013. "Harvesting Ambient Geospatial Information from Social Media Feeds." *GeoJournal* 78(2): 319–338. https://doi.org/10.1007/s10708-011-9438-2.

Steele, J., and N. Iliinsky. 2010. *Beautiful Visualization*. Sebastopol, CA: O'Reilly.

Strauss, A., and J. Corbin. 1997. *Grounded Theory in Practice*. Thousand Oaks, CA: Sage.

Sui, D., and D. DeLyser. 2012. "Crossing the Qualitative–Quantitative Chasm I: Hybrid Geographies, the Spatial Turn, and Volunteered Geographic Information (VGI)." *Progress in Human Geography* 36(1): 111–124. https://doi.org/10.1177/0309132510392164.

Urquhart, C. 2012. *Grounded Theory for Qualitative Research: A Practical Guide*. London: Sage.

Valentino-Devries, J. 2011. "Where the Young and Tech-Savvy Go?" *Wall Street Journal* (May 19). Available at https://blogs.wsj.com/digits/2011/05/19/a-week-on-foursquare/.

Ward, M., G. Grinstein, and D. Keim. 2010. *Interactive Data Visualization: Foundations, Techniques, and Applications*. Natick, MA: A.K. Peters.

Wiedemann, G. 2013. "Opening Up to Big Data: Computer-Assisted Analysis of Textual Data in Social Science." *Forum Qualitative Social Research* 14(2): 23. http://dx.doi.org/10.17169/fqs-14.2.1949.

Wilson, M.W. 2011. "'Training the Eye': Formation of the Geocoding Subject." *Social and Cultural Geography* 12(4): 357–376. https://doi.org/10.1080/14649365.2010.521856.

Zafarani, R., M.A. Abbasi, and H. Liu. 2014. *Social Media Mining: An Introduction*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139088510.

Zickuhr, K., and A. Smith. 2010. "4% of Online Americans Use Location-Based Services." Washington, DC: Pew Research Center (November 4). Available at http://www.pewinternet.org/2010/11/04/4-of-online-americans-use-location-based-services/.