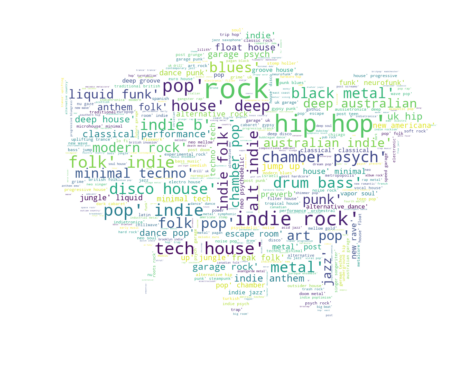
****Fuinki City

雰囲気

Spatial Data Capture, Storage and Analysis

**Msc. Smart Cities and Urban Analytics**

**Word Count: 5000 XXXXX**

**Year 2018**

**Ilja Panic**

**Soma Suzuki**

**Alfred Long**

**Jonathan Cohen**

# Introduction

Through urban computing and analytics, a lot of effort and capital is being directed towards producing a Digital City, where most of its components can be sensed, analysed, predicted and thus induced (Batty et al, 2012). For cultural and societal information this is a more nuanced and complex process.

In this context, UK has identified that its live music scenery generates a distinctive urban personality, expressing social and cultural values and making cities more attractive, and profitable. In recent years, there has been a trend towards attempting to quantify the economic impact of the music industry in a city or region (Homan, 2008; Makkonen, 2014; Titan Music Group, 2015; Boston Consulting Group, 2017; UK Music, 2017). A recent census in the UK has focused solely on the live music industry, showing among other facts that these events enhance social capital and contribute to identity formation (Webster et al, 2017).

In this context, the Greater London Authority is undertaking a “Rescue Plan for London’s Grassroots Music Venues” to ensure that London’s vibrant personality regarding its musical environment is protected and promoted (Greater London Authority, 2017). This report looked into “grassroots venues” and found that “gig listings for grassroots music venues are patchy. Buying tickets often requires knowledge of the music scene that many tourists don’t have” (Greater London Authority, 2017: 9). With this report and the creation of the role of Night Czar, the political leadership in London clearly deems the cultural and societal role of live music to be important.

Another take on the study of the societal and cultural value of music in a city has been qualitative ethnographic studies (Cohen, 2012a; Cohen, 2012b). These studies focused on Liverpool and the history of music and live music performance in the city. An attempt was made to create a GIS map of historic existing, and no longer existing music venues in the city. The map was created to invite reflection on “why live music might thrive in a particular urban area” (2012b: 591) amongst other questions. One of the difficulties come up against in this research was that “given the absence of archival documents it was difficult to trace the history or even the precise location of such venues” (2012b: 593). The research did “prompt reflection on the scope and distribution of live music venues and their embedding in the dynamics of space and time” (2012b, 595). A key point brought out by this paper is that “music venues often provided a physical idiom for defining a particular social group and the relationships involved” (2012b: 595). Another writer on this subject, Malcom Miles, writes that events such as gigs provide the “means to articulate the implicit values of a city when its users occupy the place of determining what the city is” (Miles, 1997, p. 59).

Having identified the lack of a web platform, that provides the user a simple and useful interface of the musical events happening in their city (gigs), the present project will develop a user friendly, but useful interactive tool which will allow the users to navigate through the gigs and their characteristics. Similar GIS mapping techniques have been used to allow users to find out about the geographic element of music (Homan, 2008; Cohen, 2012b), but they have all been static datasets, not taking into account the dynamic nature of live music. This project is a milestone in this research domain and therefore will contribute in generating a detailed dataset of the locations of the gigs and analysing their spatial distribution.

Section 2 will clearly state the objectives of the present research. In the following section a Literature review will provide the necessary state of the art regarding the importance of cultural events and how to measure spatial clustering. Section 4 and 5, will describe the web development and the results of the clustering analysis. Finally, last two chapters will highlight the main findings, future steps and the shortcomings encountered as researching.

## Research Objectives

The project pursues two main objectives. As stated above, there is an opportunity to allow residents of and visitors to a city to explore and discover a music scene. This spatial display of events has not been widely available before. On one hand the project is focused on creating a simple, but direct visualization of London’s live music scene, easily navigable for the lay user.

On the other hand, manipulating the collected data allows a better understanding of the characteristics and spatial distribution of the different events happening in London (GLA). This study aims to be a milestone in the research of cultural events, not only by mapping but also by making a cluster analysis of the location of the gigs.

In sum, by providing a systematic approach towards the data collection, visualization and analysis, this research aims towards a better understanding of the importance of the live music events.

# Literature Review

The cluster analysis of the gig distribution by music genre will use a similar method to Dennett and Page’s (2017) paper. Similarly to that paper it will use the density-based cluster analysis method DBSCAN (Ester et al, 1996). It will be implemented using the sklearn.cluster. DBSCAN package in Python (Pedregosa et al, 2010). DBSCAN is useful as cluster analysis for these purposes as the number of clusters does not need to be specified beforehand. DBSCAN does take parameters of EPS, the maximum distance between two points before they are assumed to be in the same cluster, and minimum samples, the minimum amount of points to be in a group before it is labelled as a cluster. How these are calibrated will be covered in the methodology.

The correctness of the clusters, a rating of their density and tightness will be found using Python’s scikitlearn: sklearn.metrics silhouette score function (Rousseeuw, 1987). This will be used for calibration of the DBSCAN EPS parameter. As in Dennett and Page’s (2017) paper, strong and significant clustering of events in certain locations will indicate that further geographic or location specific factors must have led to this clustering.

In addition, from Regional and Economics theory, there is evidence that supports and explains why some business and economic activities tend to cluster together (Krugman, 1991). As he explains in his pioneering work, as more businesses of the same kind start to agglomerate, they start to enjoy a set of benefits and spill overs that generate positive economics of scale. These are known as Agglomeration Economics. Although, at this point there is no evidence to support that musical events are benefiting from clustering, to understand the underlying factors is a complex research question out of the scope of this study.

Having said this, the first step towards further research in this direction, is to detect the existence of any spatial distribution.

Spatial autocorrelation might be detected using the Local Indicator of Spatial Interaction (LISA) Moran’s I detailed in Aneslin (1995). The statistic will be estimated using GeoDa® software. Among the different vast amount of work done trying to identify spatial patterns, this research project will adapt and follow the methodologies used to explain how the space could affect the risk of obesity (Huang et al, 2015). This is because of the clarity of the work, and the fact that restaurants, and venues hosting gigs, may share some similarities. This research shows how Moran’s I can be comparable to spatial scan and how the strength of spatial relation loses power as other variables as taken into account.

The work done in Indonesia to explore the spatio-temporal relationship of the manufacturing industries provides useful insights on how the temporal component could be relevant in identifying the strength of the relationship (Rothenberg et al, 2017).

Finally, Guo et al. (2013) present a novel methodology to establish the spatial relationship between point and polygons. The methodology presented in this case could have the potential to expand the research presented in this paper.

# Web Development

Our development process was guided by contemporary best practices for web development. Platform-as-a-Service (PaaS) providers were utilised to remove maintenance overhead involved in manually provisioning and managing servers. Mundane tasks like updating server packages, patching security holes and ensuring consistent runtime environments were abstracted so the team could focus on developing the application.

By connecting GitHub repositories to Heroku and Netlify, we were able to provide continuous integration for all the components of our application. This allowed every member of the team to instantly deploy their code changes to the server and continuously monitor how the application behaves in the production. In addition to performance and security benefits, use of PaaS allowed us to split our application into functional modules and thus avoid pitfalls of monolithic applications.

The development was enhanced through the use of online collaborative tools. Team members were assigned tasks and responsibilities via Trello - giving everyone a clear picture of the overall progress and helping avoid duplicate work. By using GitHub, we were all able to collaborate on the same codebase simultaneously without stepping on each other’s toes.

## Technical Detail

The following figure shows a simplified diagram of the project’s functional components and their underlying core technologies. Followed by a more detailed.

|  |
| --- |
| Figure 1: Distribution of Gigs by days |
|  |
| Source: Own source |

## Middleware

The middleware serves as the single point for collecting, processing and storing our dataset. Our chief data source provides 40-120 events per day for London area, with each event requiring 20+ API calls to 4 different API providers. Therefore, the decision to pre-fetch and compile all the data in advance was made in order to offload the client-facing application and provide a smoother user experience. Apart from the original data, the middleware also generates aggregate values, such as min, max, median, that are attached to an event object for easier use in later stages. The server that runs the middleware is configured to fetch data from the following APIs periodically:

* SongKick - 1 call per area - Returns a list of upcoming events for a given location, all the subsequent API calls are based on the data obtained from this call.
* Spotify - 3 calls per artist - For each artist in the event returns biographical information, listeners statistics and musical features of the most played tracks.
* Last.Fm - 3 calls per artist - For each artist in the event returns biographical information, listeners statistics and folksonomy relating to artist’s work.
* Foursquare - 2 calls per event - For each event returns information about the venue, including contact details, rating and user-generated tips.

Python was primarily chosen due to its single-threaded nature that allows for better control over the order in which API calls are executed and processed compared to asynchronous languages such as Node.js that rely on explicitly handling issues arising from concurrent code execution. Moreover, Python’s clear and readable syntax enabled us to rapidly develop the initial phase of the project that gave us access to data early on in the research process.

The middleware is hosted on a single free-tier Heroku dyno which is configured to automatically fetch and store events data for a week in advance every day.

## Database

Data about artists and venues retrieved from the different APIs comes in a JSON format with deeply nested objects and arrays. Data is often fragmented, particularly for lesser known artists and venues. Therefore, MongoDB was chosen as the most suitable database layer for data storage because its object-oriented architecture is flexible enough to support relatively inconsistent streams of data from the middleware. Also, given our choice of Javascript and Node.js in subsequent phases of the project, MongoDB is the optimal choice with regards to current web development best practices and working with objects.

The database cluster is hosted on a free-tier of MongoDB Atlas database-as-a-service provider.

## API

The API serves an interface between the database and the user-facing application. Node.js has been chosen for its capacity to handle multiple concurrent connections making it well-suited for our use-case. Particularly when retrieving data for several artists from events, which may in some cases contain over 10 artists. The API utilises MongoDB’s ‘projection’ functionality to only return data fields that are relevant for the use in the client-facing application, ignoring data intended for local data analysis. The API provides two data endpoints:

- ‘/london/events/:date’ - Returns a JSON array with the events in London for a date passed as a URL parameter.

- ‘/artists/:artistId’ - Returns a JSON array with data for an artist whose Artist ID was passed as a URL parameter.

The API is hosted on a single free-tier Heroku dyno which goes to sleep after 30 minutes of inactivity. A monitoring service that pings the API server every 25 minutes has been set up to prevent dyno from going to sleep, thus avoiding a delay caused by waking it up and providing faster response times and consistent availability.

## API

A widely used Leaflet.js library was used to provide a base map and all the map interactions for the user-facing application. Map tiles layer was custom-designed using MapBox. The application utilises Vue.js library to control DOM elements, trigger interactions, keep track of all the data and manage application’s state. A Webpack configuration would allow for better modularisation of JavaScript files, but due to time constrains a more rudimentary approach using a general-purpose Gulp.js build setup was used to automate tasks such as compiling, concatenating, minifying and optimising assets. SASS preprocessor was used to make CSS modular and stylesheets easier to manage as the application grew. Bootstrap in combination with jQuery were used to provide basic components and common interaction patters.

## Site preview or site highlight

This website intends to serve as a platform where gig-goers can explore the music events that are taking place in their city. While existing music platforms such as Songkick and Resident Advisor allow its users to find various sorts of music events, the events search function is largely based on artist name, artist popularity, or solely city name. To provide users with a different perspective for searching music events, this platform adapts the map interface with each music event visualized as a marker on the map. Each marker describes the information of the artists and venue of the event. Furthermore, this website works as an experimental tool that can visualize the level of the vibes in the city based on the music event data.

|  |
| --- |
| Figure 2: Distribution of Gigs by days |
|  |
| Source: Own source |

# Analysis

## The data set

As described above, the main source of information for this project was SongKick, which was accessed using an API. From this site information about the venues and the bands that were playing was retrieved. Although, there was information about the gig, it was complemented by using Spotify, Last.fm and FourSquare APIs. Consequently, the data from SongKick, could be significantly enriched to satisfy our previously stated objectives. In the Appendix a complete description of the final dataset can be found.

By developing and deploying a set of routines to be executed every day for a month, we were able to build a rich dataset, of the gigs in London. This could easily be done for major cities in the world, such as Tokyo, Buenos Aires, Berlin or Liverpool. As SongKick, Spotify, FourSquare and Last.fm, are services working globally, these datasets could be retrieved for a different set of cities. This information and all the processes here explained and used for London could be deployed for any other geographic region. Of course the quality of the information will vary from country to country, or even among cities, the methodologies developed for these objectives are systematic and scalable.

## London’s music environment

Between the 13th of April and the 26th of May, a total of 1863 gigs were recorded. As the weekday approaches the weekend, the average amount of gigs increases. On average, Fridays host the most events (100), followed by Saturdays with 96. On average, Sundays were less active than the Thursdays by almost 10 gigs. The following graph shows these results.

|  |
| --- |
| Figure 3: Distribution of Gigs by days |
|  |
| Source: https://fuinki.netlify.com/about |

Regarding the different characteristics of the music played by these artists, it became possible to start analysing the ‘vibe’ of the city and its venues. Using the previously described variables retrieved from Spotify, Last.fm and SongKick, the following spider graph was created. By looking at the distributions of these variables, it is possible to identify high dispersion and extreme values, as a result, the median was chosen as the statistic to study. The lines in the graph represent the standardized values of the average values of the different characteristics of the songs played by these artists. According to these variables the bands playing on Fridays have the most followers and have the most “live recordings”. Saturdays, are the most energetic days and Sundays are the days with the most lyrical and danceable music. The most popular bands play on Tuesdays.

|  |
| --- |
| Figure 4: Distribution of Gigs by days |
|  |
| Source: https://fuinki.netlify.com/about |

Before, moving to a spatial analysis it was also necessary to study how frequently the different genres occured. As the spatial analysis will be divided into different types of music, it was important to study the distribution of the different music types. The most frequent labelled genre is Pop music and it was followed by Rock with almost 600 appearances.

Finally, to begin our spatial analysis of how the gigs are spatially distributed, in the about page there will be a public map, produced with Folium technology will show the point densities of the gigs. As this is a visual representation of the density of the gigs, we also included the following bar chart to show the amount of gigs per Boroughs.

|  |
| --- |
| Figure 5: Distribution of Gigs by days |
|  |
| Source: https://fuinki.netlify.com/about |

For more information about descriptive statistics about the distribution of the gigs during the week, type of music or the spatial distribution, visit <https://fuinki.netlify.com/about>.

# Methodology

In order to detect the existence of Spatial Correlation among the location of the gigs during this period of time, two methods are applied: (i) LISA Moran’s I and (ii) DB Scan for cluster detection.

## LISA: Moran’s I

One of the most spatial statistics used to detect clusters, this metric assumes that each ward has a continuous amount of gigs. A weight matrix is used to define the neighbours of each ward. The test is based on probabilistic theory and the clusters can be tested statistically by fixing all observations and then randomly permutating all the observed amounts of gigs. Each time the experiment is executed a distribution is calculated under the null hypothesis of no presence of clusters. The Moran’s I ranges from -1 to 1 were 0 indicates randomness (Zhang and Lin, 2007).

To work with this LISA statistic, the gigs’ locations were joined to the map of the wards, as after a radius of 10 km from the City of London, there were almost no records, the city was sub-setted. Once we had selection of the entire set, different ardency matrices were created and tested. As we were trying to capture the local spatial correlation, we decided to choose 15 nearest neighbours. This ensured that all the adjacent wards were taken into account. The Moran’s I was executed for all the gigs.

|  |  |
| --- | --- |
| Figure 6: Moran’s I – 15 Nearest Neighbours | |
| *Moran’s I INDEX* | *Cluster Map* |
|  |  |
| *Significance Map* | *Local G Cluster Map* |
|  |  |
| *Source: Own production* | |

We found out that Moran’s I index was 0.14, indicating that the spatial interaction for the amount of gigs at a ward level was not strong as expected. Although it was not estimated, we can assume that there is a high correlation between the following maps and the ones of population, density, income or employment.

## Clustering

Subsets of the gig database were created by searching for a genre, for example “reggae”, and returning any event which had an artist playing described in either Spotify or LastFm as a reggae artist. The events were compiled into dataframes for each of the chosen genres.

The minimum sample parameter will be set at 5, assuming that any smaller sample size is not significant enough to be considered a cluster of gigs.

To calibrate the correct EPS distance it is necessary to assess the silhouette score of the clusters at different EPS DBSCAN parameters. Different EPS distances will be set, every 10 meters from 1m to 2000m. 2000m is chosen as at this point the neighbourhood communities searched for would become too large to represent a neighbourhood.

The silhouette coefficient and number of clusters for DBSCAN results will be measured for each genre, for each of the different EPS distances. The silhouette coefficient plot will be used to assess which EPS distance to set for the DBSCAN clustering to create the cluster maps. It will also show which of the genres show strong clustering at different EPS distances. The plot of the number of clusters will show how many “communities” there are of each genre.

After the DBSCAN EPS parameter has been decided the clusters will be plotted on a map and compared against the Moran’s I analysis, this will show what kinds of communities or clusters are picked up by either technique.

If strong clustering relationships occur that are different for different genres, this study allows a starting point for investigating how these dynamic musical communities relate to the other communities sharing the city.

# Results

Moran’s I index was calculated for a selection of points based on the genre of the gigs. The following table of figures shows the results obtained. The differences found in this set of graphs strongly suggest that there are certain wards differentiating from others. It could be that these places are performing actions to promote this or simply that the neighbourhood personality and characteristics promote the development of certain types of events.

For example, it can be appreciated that Jazz gigs are not as central Punk, Techno or House music. These 3 genres might be more popular and therefore they tend to cluster in more central areas.

|  |
| --- |
| Figure 7a: Distribution of Gigs by days |
|  |
| *Source: Own production* |

A similar phenomena was capture when looking at classical music, evidencing again that the more specific the genre is the less centric is agglomerates.

|  |
| --- |
| Figure 8b: Distribution of Gigs by days |
|  |
| *Source: Own production* |

As is visible in the plot below, the silhouette scores of dbscans for clustering of different genres, Reggae is not deemed to be strongly clustered. Techno clusters’ silhouette scores rise as the EPS rises from 500m to 1000m as a group of events around Elephant and Castle, and another around Angel become clusters. Most genres reach an early peak at an EPS of 600m. Almost all of the genres either level off or dip after this point with House, Techno and Reggae as exceptions. House and Techno continue to rise to a peak at an EPS of 1000m, and Reggae only levels off at an EPS value of 1500. All the other genres rise again by the EPS of 1500m, at this point the clusters are growing so large in scale that it becomes less meaningful when working on a city scale.

|  |
| --- |
| Figure 9: Silhouette Metrics at different Eps Values |
|  |
| Source: https://fuinki.netlify.com/about |

|  |
| --- |
| Figure 10: Number of cluster at different EPS Values |
|  |
| Source: https://fuinki.netlify.com/about |

This relationship is also shown by the graph of EPS against the number of clusters by genre, at an EPS of 1500m the largest number of clusters for any single genre included is 7, for Punk and Techno. With EPS values this large it stops becoming meaningful to look for clusters that represent communities or neighbourhoods.

By plotting the clusters as well as the Moran’s I high-high wards, it becomes possible to see the spatial trends of gigs of particular genres. These can then be viewed as musical communities forming/being formed.

In the *Classical Cluster* the gigs are all centrally located. The DBSCAN clustering picks out clusters in the West End, many gigs at a single venue on the Southbank, a few around Kings Cross, the City of London and in Camden. These are well correlated with the Moran’s I highlighting of wards. Due to the relatively low number of gigs, Moran’s I found high-high relationships where the DBSCAN minimum sample size of 5 did not produce clusters. This shows that classical events “communities” occur mainly in central London, though non-clustered gigs are distributed far further both east and west, though not many in South London.

*Folk Cluster* shows Folk gigs are widely spread over many different areas of london. Lots of distinct clusters form in north London, especially to the east. Interestingly Moran’s I does not pick out the large cluster around Mayfair, this is due to the large gap between this cluster and any gigs further west.

*House Cluster* shows the clustering of House gigs in areas that had not been populated by Folk or Classical gigs, Southwark and Brixton. Again, Moran’s I does not pick out the strong clusters further out.

In the *Jazz Cluster* Moran’s I highlights wards far out from the centre of London without many gigs in them. Jazz gigs are highly clustered, with distinct clusters far from one another. Moran’s I also displays this disparate clustering, though highlighting some other areas.

In the *Punk Cluster* the gigs form many distinct clusters in North London, especially in the east.

The *Reggae Cluster* shows Reggae did not have enough events close together to show any strong clustering, as described by the EPS analysis earlier. There is one distinct cluster in Camden but overall it is remarkably evenly distributed.

The *Soul Cluster* shows a large number of gigs, with the majority of clusters in north and central London. Moran’s I highlights wards that correlate very well with the larger gig clusters.

Techno gigs show a few well defined clusters in Camden, Kings Cross, and around Elephant and Castle. Fabric fittingly provides a cluster from the gigs that have happened there alone.

These results are mapped in the following table of images.

|  |
| --- |
| Figure 11:These results are mapped in the following table of images. Figure 12: Cluster analysis by genre |
| Classical Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527059091485_classical-cluster.png |
| Folk Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527059375619_folk-cluster.png |
| House Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527059578848_house-cluster.png |

|  |
| --- |
| Jazz Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527059788195_jazz-cluster.png |
| Punk Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527059903036_punk-cluster.png |
| Reggae Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527060643328_reggae-cluster.png |
| Soul Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527060786710_soul-cluster.png |
| Techno Cluster |
| https://d2mxuefqeaa7sj.cloudfront.net/s_177EA3C7AA9AD0FBCB783ED748A1C4C186C1AF3B720331E87665D8BF754708E1_1527060934082_techno-cluster.png |
| Source: https://fuinki.netlify.com/about |

# Conclusions

Each of the genre types that have been picked out show a distinctive clustering profile that would easily allow further interrogation either through qualitative or quantitative means.

The collection and structuring of data has already addressed the shortcomings in data pointed out in Cohen’s paper (2012b). This provides a way of analysing Miles’ (1997) claim that live music is one avenue for the assertion of a city’s implicit values. These musical communities can now be compared and contrasted with any other form of community sharing the same urban space.

The app has provided one route to overcoming the difficulty in getting to know a music scene by providing a dynamic immersive experience in which a user can navigate using the feel of the music rather than having to know specific artists or venues. This directly addresses the point made in the “Rescue Plan for London’s Grassroots Music Venues” report (Greater London Authority, 2017) about the impenetrability of the local music scene.

By providing a database of gigs in a way that makes quantitative analysis possible it would be very easy to create metrics to improve the musical city algorithm proposed by Baker (2017), this would also be of interest to the initiators of the UK Live Music Census and the institutions behind the economic analyses of the music industry in cities and regions.

# Limitations

The first input source used is the SongKick website. This means that all the data we have collected is initially restricted to those events registered on the SongKick platform. The SongKick data, once collected, was used to call the APIs of Spotify, Lastfm, and Foursquare. Genre information was taken from both Spotify and LastFm, this means that the data was in fact subset again, to those events with artists for which Spotify or LastFm had matching records.

As with any study based on internet sources, what is being analysed is collected events, what is actually happening in the city may be quite different. Many of the most informal events would not be posted on SongKick, perhaps not on any online platform. This may mean our analysis is affected by the digital divide, it is restricted to the kind of music produced and played by people with enough time, education, and money to exploit these kinds of digital platforms.

While analysing the results of Moran’s I and DB Scan, the need to understand the factors influencing the location of the gigs became clear. In other words, this research can be taken as an exploratory exercise of the agglomeration of live music events. The following questions to be answered in the research agenda go in to different directions by understanding the:

1. Effects (positive and negative- direct and indirect) of these activities over the city, to individuals and society.
2. Consequences causes behind the agglomeration of these activities and introducing more sophisticated methodologies that will take into account other demographic, temporal and built environment variables.

# Future steps

Historic data could provide an opportunity to generate spatio-temporal analysis of the musical events in London and any other city. Retrieving this information could be vital to understanding how the dynamics of gentrification are affecting London's night environment.

Moreover, the effects of the night tube, over the spatial distribution of the venues could be studied in greater detail if historical data was provided.

Finally, the tool and the analysis generated for this specific project can potentially be expanded to other cities. Doing this will contribute to gain comprehension of how the music environment is in other cities. This will allow to make a comparison between different urban areas.

# Bibliography

Anselin, L., 1995. Local Indicators of Spatial Association—LISA. Geographical Analysis 27, 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>

Baker, A.J., 2017. Algorithms to Assess Music Cities: Case Study—Melbourne as a Music Capital. SAGE Open 7, 215824401769180. <https://doi.org/10.1177/2158244017691801>

Batty, M., Axhausen, K.W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., Portugali, Y., 2012. Smart cities of the future. Eur. Phys. J. Spec. Top. 214, 481–518. <https://doi.org/10.1140/epjst/e2012-01703-3>

Boal, I., Herrero, L.C., n.d. Where are the artists? Analysing economies of agglomeration in Castile and León, Spain. Papers in Regional Science 0. <https://doi.org/10.1111/pirs.12314>

Boston Consulting Group, 2017. Economic Impact, Trends, and Opportunities: Music in New York City. New York City Mayor’s Office of Media and Entertainment, New York City.

Buczkowski, A., 2016. Mapping the urban sounds. Geoawesomeness.

Cohen, S., 2012a. Bubbles, Tracks, Borders and Lines: Mapping Music and Urban Landscape. J. Roy. Music. Assoc. 137, 135–170. <https://doi.org/10.1080/02690403.2012.669939>

Cohen, S., 2012b. Live music and urban landscape: mapping the beat in Liverpool. Soc. Semiot. 22, 587–603. <https://doi.org/10.1080/10350330.2012.731902>

Dennett, A., Page, S., 2017. The geography of London’s recent beer brewing revolution. The Geographical Journal 183, 440–454. <https://doi.org/10.1111/geoj.12228>

Doi, C., 2017. Connecting Music and Place: Exploring Library Collection Data Using Geo-visualizations. Evid. Based Lib. Inf. Pract. 12, 36–52. <https://doi.org/10.18438/B86078>

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. KDD-96 Proceedings 226–31.

Florida, R., n.d. Which Cities Are the Musical Trend Setters? [WWW Document]. CityLab. URL <http://www.theatlanticcities.com/arts-and-lifestyle/2012/04/which-cities-are-musical-trend-setters/1799/> (accessed 5.14.18).

Fu, W.J., Jiang, P.K., Zhou, G.M., Zhao, K.L., 2014. Using Moran’s I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. Biogeosciences 11, 2401–2409. <https://doi.org/10.5194/bg-11-2401-2014>

Greater London Authority, 2017. Rescue Plan for London’s Grassroots Music Venues: Making progress.

Guo, L., Du, S., Haining, R., Zhang, L., 2013a. Global and local indicators of spatial association between points and polygons: A study of land use change. International Journal of Applied Earth Observation and Geoinformation 21, 384–396. <https://doi.org/10.1016/j.jag.2011.11.003>

Guo, L., Du, S., Haining, R., Zhang, L., 2013b. Global and local indicators of spatial association between points and polygons: A study of land use change. International Journal of Applied Earth Observation and Geoinformation 21, 384–396. <https://doi.org/10.1016/j.jag.2011.11.003>

Homan, S., 2008. A portrait of the politician as a young pub rocker: live music venue reform in Australia. Pop. Music 27, 243–256. <https://doi.org/10.1017/S0261143008004030>

Huang, R., Moudon, A.V., Cook, A.J., Drewnowski, A., 2015. The spatial clustering of obesity: does the built environment matter? Journal of Human Nutrition and Dietetics 28, 604–612. <https://doi.org/10.1111/jhn.12279>

J Borowiecki, K., 2013. Agglomeration economies in classical music. Papers in Regional Science. Papers in Regional Science 94. <https://doi.org/10.1111/pirs.12078>

Krugman, P., 1991. Increasing Returns and Economic Geography. Journal of Political Economy 99, 483–499.

Makkonen, T., 2014. Tales from the thousand lakes: placing the creative network of metal music in Finland. Environ. Plan. A 46, 1586–1600. <https://doi.org/10.1068/a130020p>

Malcolm. Miles, 1997. Art, space and the city: public art and urban futures / Malcolm Miles. Routledge, London ; New York, London.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825−2830.

Pedro, J., 2017. Musical Cartographies of Madrid: City, Popular Music and New Digital Technologies. CIC-Cuad. Inf. Comun. 22, 169–185. <https://doi.org/10.5209/CIYC.55973>

Rothenberg, A., Bazzi, S., Nataraj, S., Chari, A., 2017. Assessing the Spatial Concentration of Indonesia’s Manufacturing Sector: Evidence from Three Decades. RAND Corporation. <https://doi.org/10.7249/WR1180>

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20, 53–65. <https://doi.org/10.1016/0377-0427(87)90125-7>

Santoso, E., Wilantari, R.N., 2018. THE SPATIAL DISTRIBUTION OF INDONESIA’S MANUFACTURING INDUSTRIES: AN EXPLORATORY SPATIAL DATA ANALYSIS. UNEJ e-Proceeding 131–138.

Titan Music Group, 2015. The Austin Music Census: A data-driven assessment of Austin’s commercial music economy. The City of Austin Economic Development Department’s Musci & Enterntainment Division.

Trolliet, A.S., 2014. “Buenos Aires beat”: a topography of rock culture in Buenos Aires, 1965-1970. Urban Hist. 41, 517–536. <https://doi.org/10.1017/S0963926813000722>

UK Music, 2017. Measuring Music. UK Music.

Webster, E., Brennan, M., Behr, A., 2017. Valuing live music: The UK Live Music Census 2017 report 286.

Zhang, T., Lin, G., 2007. A decomposition of Moran’s I for clustering detection. Computational Statistics & Data Analysis 51, 6123–6137. <https://doi.org/10.1016/j.csda.2006.12.032>