Analyzing the Noise in New York City with 311 Data

HUANG Jina 1155113039

WANG Siyu 1155111346

WANG Yuanyuan 1155114422

GEOGRAPGY & RESOURCE MANAGEMENT, CUHK

GISM5065 Geospatial Big Data for Urban Planning

20/12/2018

Contents

1 Introduction

Noise pollution, air pollution, and water pollution are listed as three major pollutants around the world. In recent years, noise pollution has caused a huge amount of environmental problems in the world. The World Health Organization (WHO) has conducted a survey on noise pollution in the world. It is reported that the noise pollution problem in the United States and developed countries are becoming more and more serious. In the United States, the number of people living with noise pollution above 85 decibels has increased several times in 20 years, while more than 40 percent of residents are affected almost all day by traffic noise in the European Union. However, noise pollution problems are also quite severe in lots of cities among developing countries with a level of 75 to 80 decibels twenty-four hours a day.

According to research published by WHO, which was conducted by a group of scientists, noise pollution does have some negative effects on people. Not only does noise pollution lead to a hypothesis, but also may cause the disorders of the digestive system. Especially, the most direct harm to noise is hearing damage. If people work in a noisy environment for a long time, the incidence of deafness increases significantly. Thus, it is important for governments to develop relevant laws and regulations in order to reduce noise pollution.

311 is a local complaint call in the United States. For any non-emergency situation, you can call 311 to report problems in the city, such as holes in the road, trees collapse, garbage, noise problems, and so on[5]. When you report a problem, relevant departments will take some measures to tackle the problem. Consequently, it is the primary approach of dealing with complaints from citizens. NYC311 deals with a mass of complaints every day, producing a huge number of data. In terms of Local Law 47 of 2005, reports about NYC311's data must be presented to the City Council, the

Public Advocate, community boards, and the public by the Department of Information Technology and Telecommunications every month[4].

According to the public record of complaints on the NYC Open Data website, noise ranks the third place of various types of 311 complaints, which has shown over two million complaints from eight million people in New York City[1].

Based on the 311 Noise Data in New York City, two types of analysis (Time series analytics as well as spatial analytics) were conducted in this project in different platforms (Excel 2016 as well as ArcGIS Pro 2.2) and several conclusions were drawn.

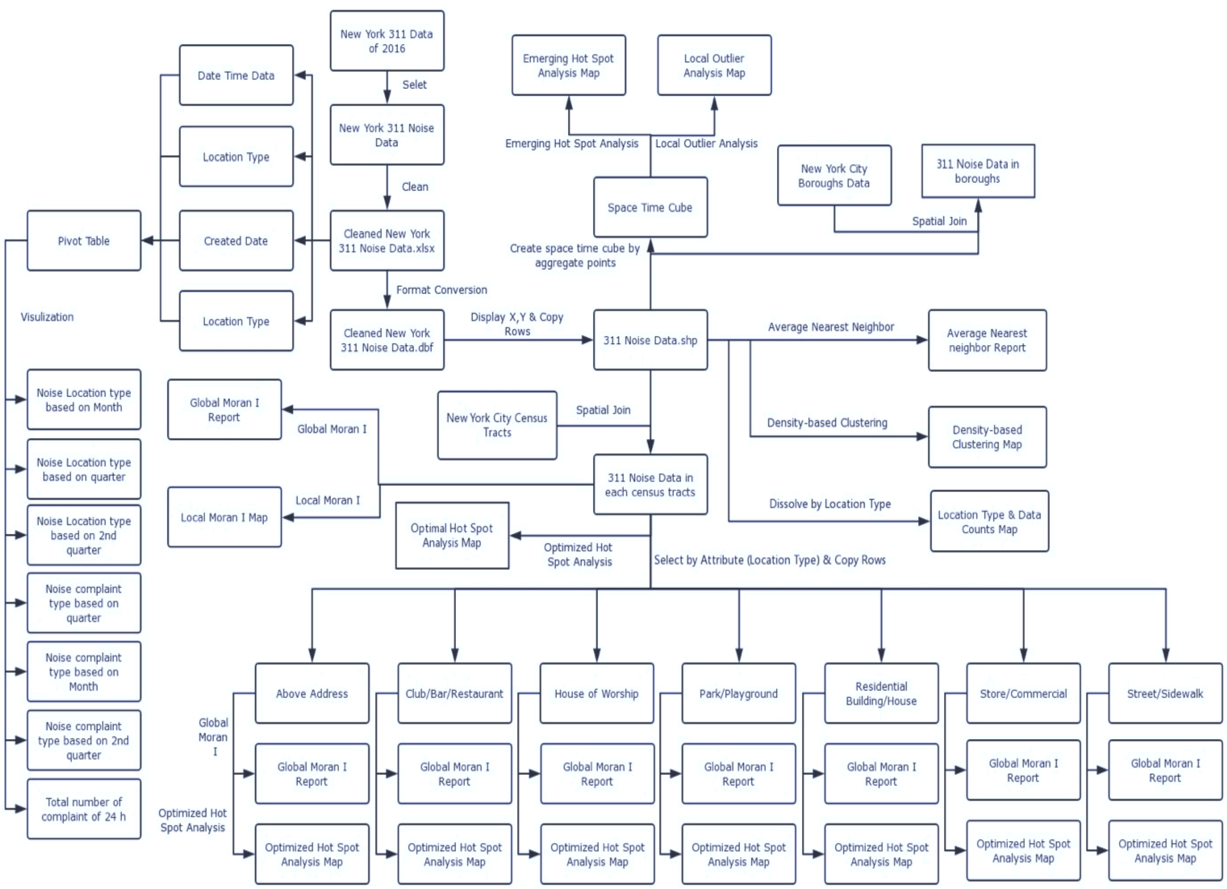

Figure 1.1 The technical flow chart

2 Study Area and Data

2.1 Study Area

The study area in this project is New York City (Fig.2.1), the core of the New York metropolitan area having great influence on the cosmopolitan cities, which is not only the most populous city in the USA, but also one of the world's largest cities for the global economy, business, finance, media, politics, education and entertainment. Subdividing into each census tract, in New York City, there are over two thousand census tracts, which usually have the number of populations 3500. The average area of each census tract is 90 acres. The Census Bureau updates the information of Census Tract each decade who keep changes to a minimum. Therefore, it is important for us to keep in mind that it is the 2016 census tract geography employed by this way.

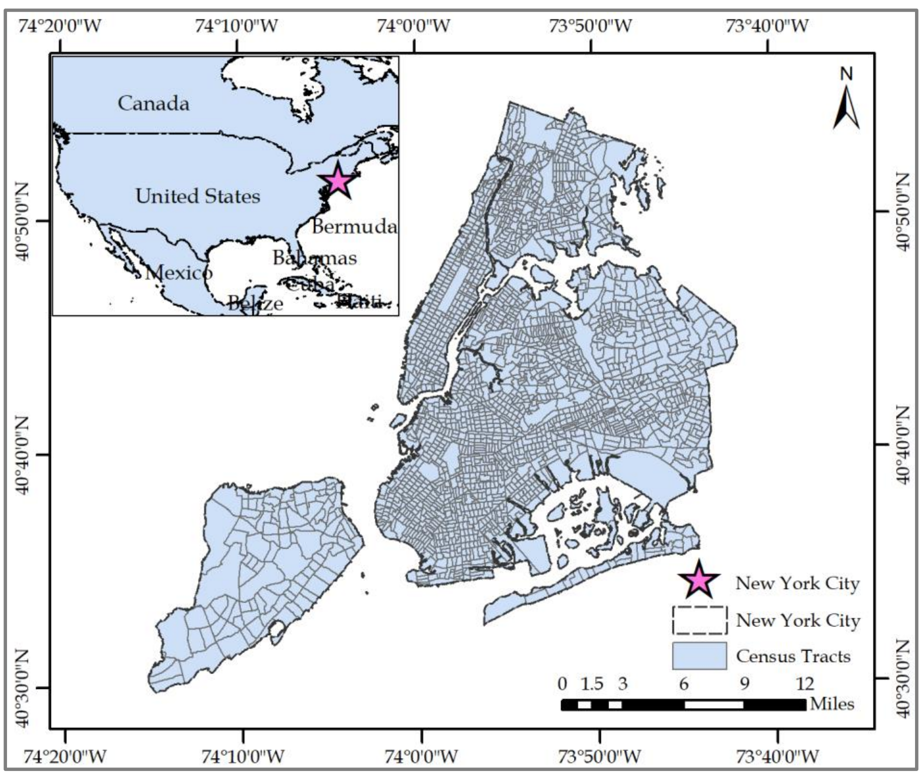

Figure 2.1 The location map of New York City

2.2 Data source and preparation

We obtained 311 complaint data records of New York City for the full year 2016 from the NYC open data portal website (https://data.cityofnewyork.us/browse?q=311) and the population data for the 2016 New York City based on census tract. The small statistical subdivisions of counties are Census Tracts provided by the US Census Bureau. There are 2166 census tracts in New York City. Due to a large amount of data, all types of complaint records are included, and with many unnecessary fields and empty records, so we need to filter and clean the data.

The 311 noise complaints for the entire year of 2016 include seven Location types: Above address, Club/Bar/Restaurant, House of Worship, Park/Playground, Residential Building/House, Store/Commercial and Street/Sidewalk. And also, seven kinds of noise complaint types which are Commercial, Helicopter, House of Worship, Park, Residential, Street/Sidewalk, and Vehicle.

There are 2360812 lines of 311 data records, while only 1048576 lines of data can be read at the same time. The original CSV file is divided into three CSV files using a script file (Free Huge CSV Splitterchs) to simplify the raw data. After screening the complaint type as noise data, blank columns, coordinate-free data, the completion of the complaint in 2016, empty Location type, Park Facility columns and the status as unfinished records have all been removed. Thus, the data of the three periods may be combined into one file and finally get 358162 data records (Fig.2.2).

Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type
32305299	1/1/2016	1/1/2016 1:57	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk
32310343	1/1/2016	1/1/2016 3:12	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32305983	1/1/2016	1/1/2016 3:24	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32305208	1/1/2016	1/1/2016 2:43	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32309484	1/1/2016	1/1/2016 0:28	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk
32309055	1/1/2016	1/1/2016 3:54	NYPD	New York City Police Department	Noise - Residential	Banging/Pounding	Residential Building/House
32310309	1/1/2016	1/1/2016 1:35	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32305945	1/1/2016	1/1/2016 1:30	NYPD	New York City Police Department	Noise - Residential	Banging/Pounding	Residential Building/House
32310278	1/1/2016	1/1/2016 0:10	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32307413	1/1/2016	1/1/2016 5:23	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32305991	1/1/2016	1/1/2016 2:36	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32309413	1/1/2016	1/1/2016 7:15	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32305243	1/1/2016	1/1/2016 2:21	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32306674	1/1/2016	1/1/2016 0:31	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32305260	1/1/2016	1/1/2016 1:46	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House
32307278	1/1/2016	1/1/2016 1:17	NYPD	New York City Police Department	Noise - Residential	Loud Music/Party	Residential Building/House

Figure 2.2 The valid 311 Noise Data Records

3 Methods

Generally, there are two types of methods be used in this project. The first one is time series variation analytics which is finished in Excel 2016. The second one is spatial analytics which is realized in ArcGIS Pro 2.2.

3.1 Time series variation Analytics

Export the Created date column, Complaint Type column, and Location type column of the cleaned data to the new table. The Created date column is split, only the year, month, and day are taken, and the format of the data output must be in the format of text and then converted to date. Otherwise, the time data format is not uniform, and the PivotTable cannot be further created.

The data format of the created date column is unified, so we need to use a time format such as '14/3/2012'. Then, create a PivotTable of Complaint type, Location Type, time-sharing data and the time of the complaint throughout the year (Fig.3.1).

Created Date	Complaint Type	1	Created Date	Location Type
1/1/2016	Noise - Street/Sidewalk	2	1/1/2016	Street/Sidewalk
1/1/2016	Noise - Residential	3	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	4	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	5	1/1/2016	Residential Building/House
1/1/2016	Noise - Street/Sidewalk	6	1/1/2016	Street/Sidewalk
1/1/2016	Noise - Residential	7	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	8	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	9	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	10	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	11	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	12	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	13	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	14	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	15	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	16	1/1/2016	Residential Building/House
1/1/2016	Noise - Residential	17	1/1/2016	Residential Building/House

Figure 3.1 The table of Complaint and Location Type

3.2 Spatial Analysis

The spatial analysis methods used in this project could be divided into two types. The first one is Spatial Statistics. The second type is the Space-Time Analysis.

3.2.1 Spatial Statistics

Spatial Statistics help us to explore the distribution, patterns, relationships of the data. The theories of spatial statistics resemble to the normal one. Nevertheless, the input data of spatial statistics contain information related to space which differs it from the normal statistics.

There are five primary Statistics Analysis methods used in this project. The fundamental introduction of those tools will be introduced in the paragraphs below. Readers could find out more information in the help document of ArcGIS Pro 2.2.

1) Density-based Clustering identifies the clusters of points as well as noise within a mass of points in terms of their distribution in space using a way of machine learning. In our project, we chose Self-adjusting method, which is more suitable for our study, by comparison, to conduct the manipulation[6].
2) Average Nearest Neighbor computes the nearest neighbor index in terms of the median distance between a feature and its nearest neighbor feature. Computed based on a single ratio formula (observed/ expected), when the value of the index is lower than 1 then the input features show a pattern of clustering. On the contrary, if the value is higher than 1 the input features exhibits a dispersed trend[6].
3) Optimized Hot Spot Analysis is a widely used tool to explore the clusters of hot spots and cold spots that are significant in the context of statistics using the Getis-Ord Gi* statistic[6].

4) Cluster and Outlier Analysis (Anselin Local Moran's I) not only calculate the hot/cold spots of your data but also provide you with the information about the outliers[6].
5) Spatial Autocorrelation (Global Moran's I) using Global Moran's I statistic method to explore the spatial autocorrelation in terms of the spatial location of input data as well as attribute at the meantime[6].

3.2.2 Space-Time Series Analysis

In order to better utilize our informative noise data which would give us a deeper understanding of the spatial-temporal distribution and trends of noise in New York City, we conducted several Space-Time tools. First and foremost, we created a Space-Time Cube by aggregating the 311 Noise Data in New York City. The Cube we created lay the foundation for two Space-Time mining tools (Emerging Hot Spot Analysis as well as Local outlier analysis). The basic introduction of the three tools are described below.

1) Create Space Time Cube by Aggregating Points condenses points into a netCDF data structure by assembling them into spatial-temporal bins. The points as well as the attributes in specific fields are counted and assembled in each bin. The trendy of counts and summary attribute values in each bin will be assessed[6].
2) Emerging Hot Spot Analysis analyzes the trends types of change of data value in the Space-Time Cube. The trends types include new, intensifying, diminishing, and sporadic hot and cold spots[6].
3) Local Outlier Analysis finds the space-time clusters as well as outliers which are significant according to statistic. It is the space-time enhanced version of the Anselin Local Moran I statistics[6].

3.2.3 Model Builder

In order to effectively perform spatial analysis, we build several models in ArcGIS Pro 2.2 to conduct the spatial analysis. In terms of the input data, eight models were constructed.

The first model which is also the primary model used two shapefiles as inputs. The model could be divided as two parts. As the figure shown below, in the left part, a polygon shapefile which refers to the noise data in each census tracts were used as input to perform Spatial Autocorrelation (Global Moran's I), Optimized Hot Spot Analysis as well as Cluster and Outlier Analysis (Anselin Local Moran's I). In the right part, a point shapefile which contains all of the 311 Noise Data (358,161records) in New York City was used as input to conduct Average Nearest Neighbor, Density-based Clustering and also used to create a Space-Time Cube which is the foundation for Emerging Hot Spot Analysis as well as Local Outlier Analysis (Fig.3.2).

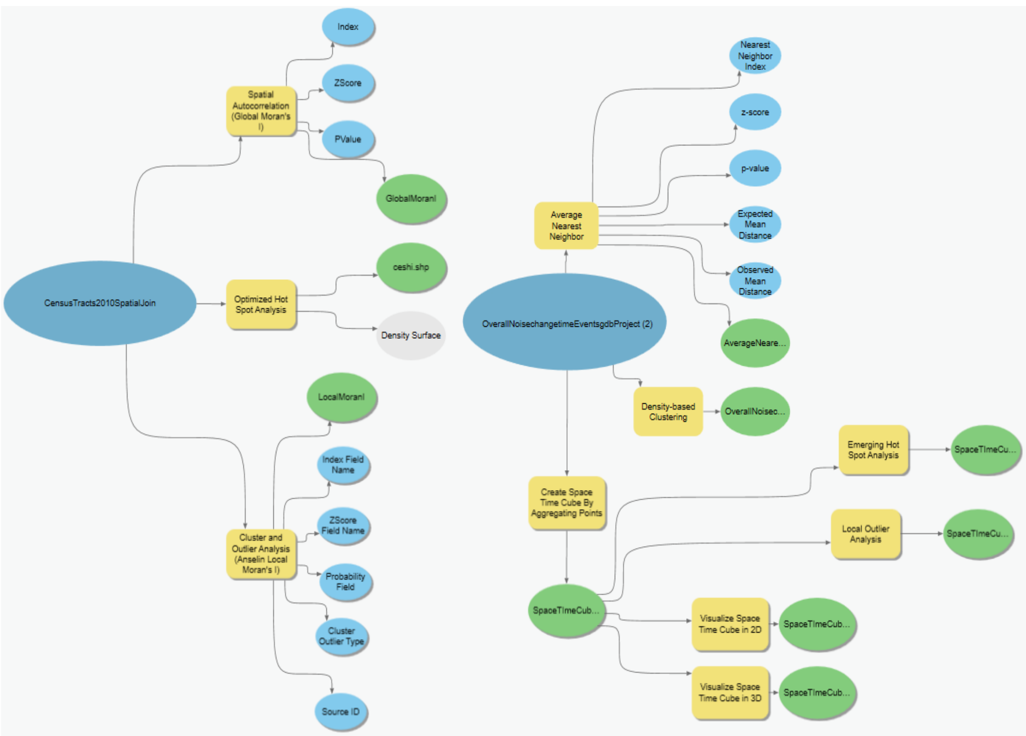

Figure 3.2 The primary model

The minor models used the polygon shapefile which contain different location types' 311 Noise Data in each census tracts in New York City as input. There are seven location types in total: Above Address, Club/Bar/Residential, House of Worship, Park/Playground, Residential Building/House, Street/Sidewalk, Store/Commercial.

Because the seven models use the same tools and just differs in terms of the input data, we would not place them one by one here in order to reduce redundancy (Fig.3.3).

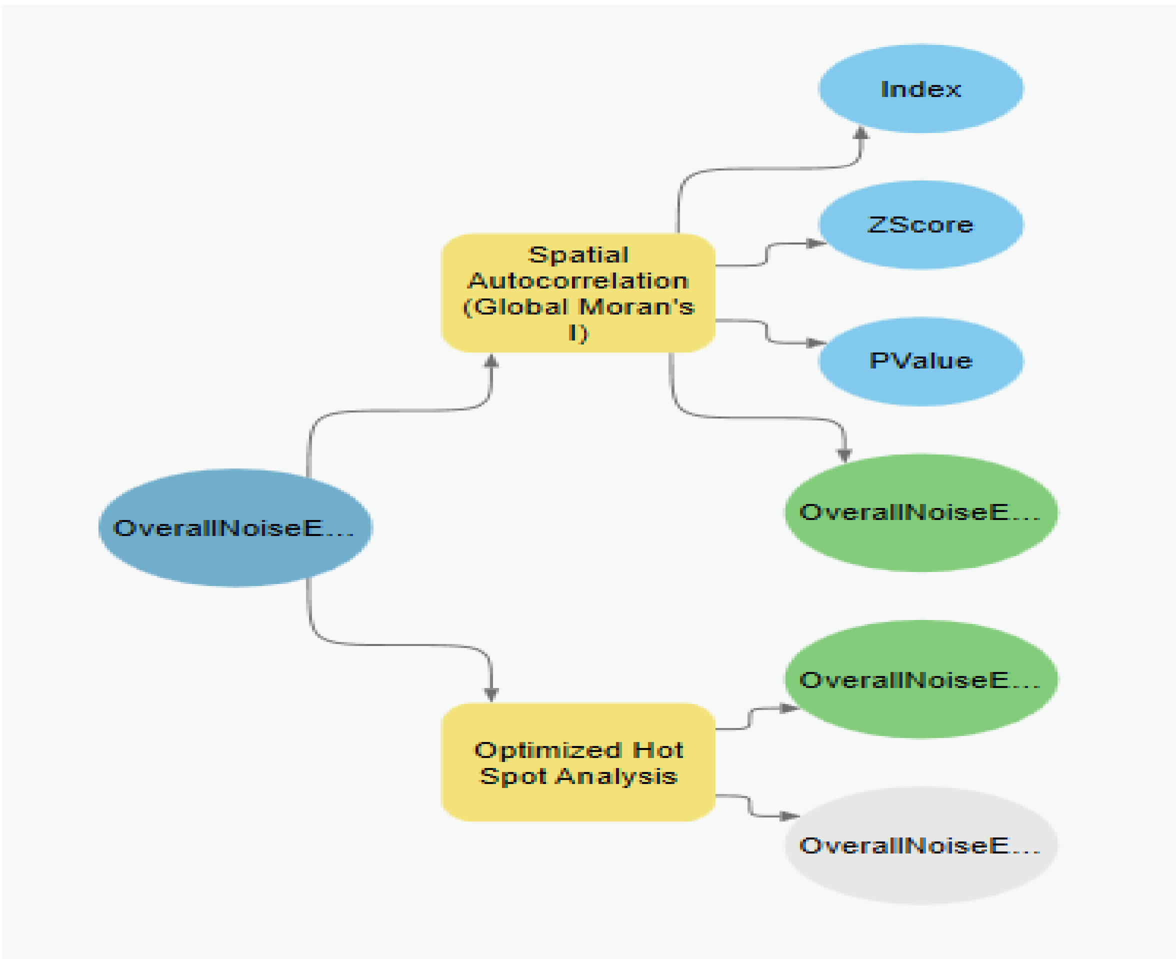

Figure 3.3 The minor model

4 Results Analytics

4.1 Time series variation Analytics

Generate monthly and quarterly time-series timing diagrams based on Location type. In the monthly time chart, May was the month with the highest number of noise complaints, with a total of 38,823 records, followed by June and September. The number of noise complaints in July and August was slightly lower, but overall it was still higher than January, February, November and December. Because New York City is Humid subtropical climate, it is rainy and hot in summer while rainy and mild in winter. We found that the number of noise complaints was positively correlated with the degree of climate variability. In the winter in New York City, the figure of complaints is also small due to fewer crowd activities (Fig.4.1.1).

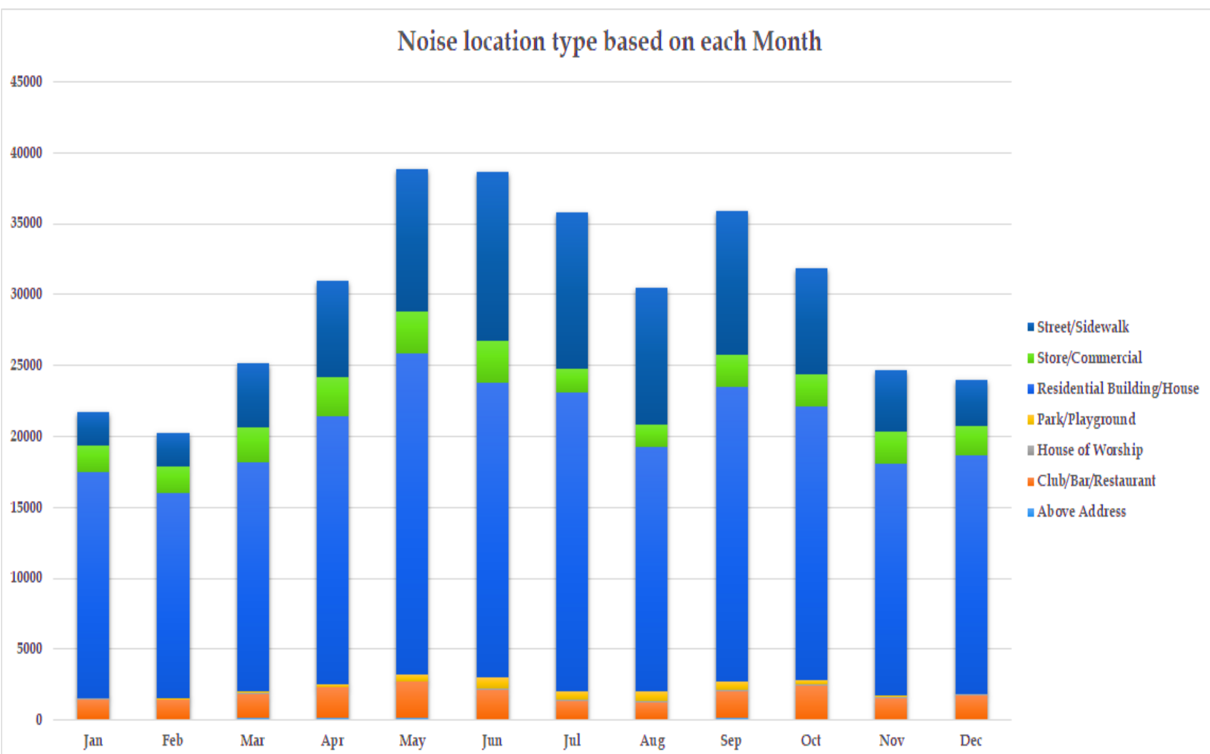

Figure 3.1.1 Noise location type based on each Month

The number of noise complaint of Location type for Residential Building is the highest in each quarter of the year. This part of the noise complaint comes from the door knocking, party opening, decoration and so on (Fig.4.1.2).

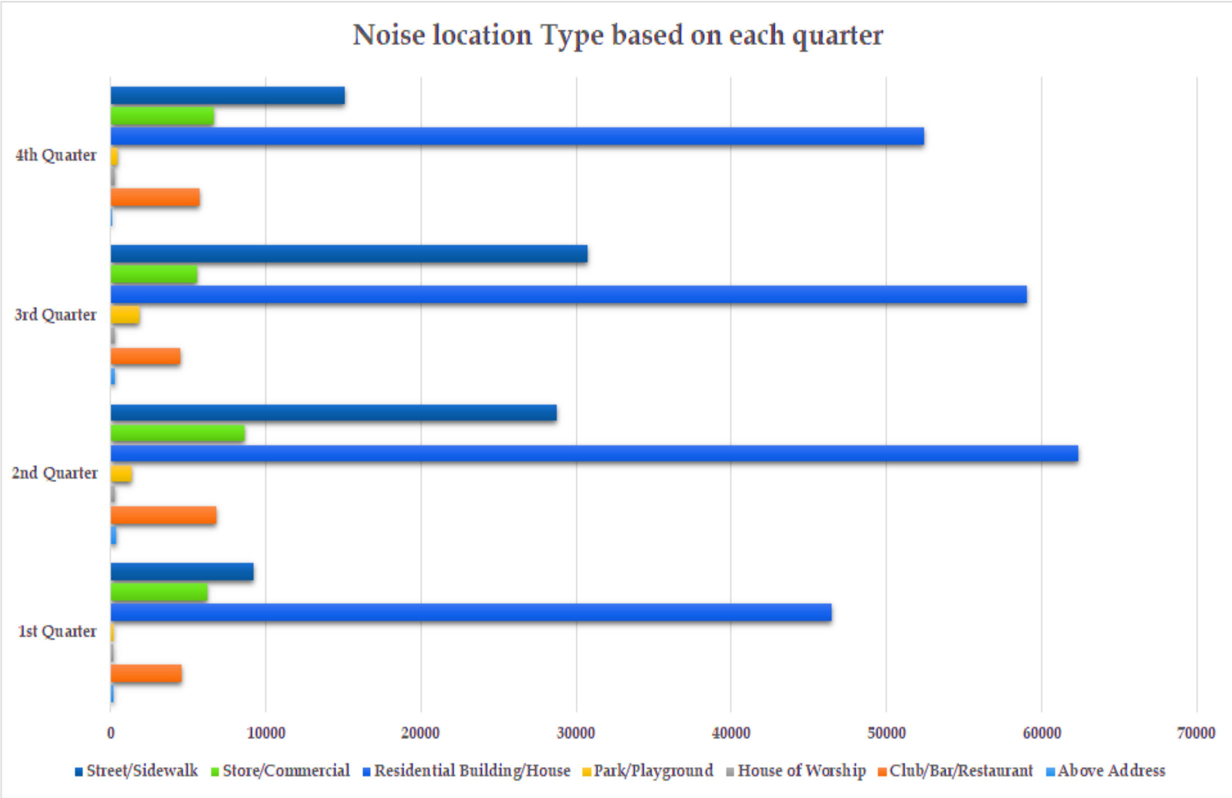

Figure 4.1.2 Noise location type based on each quarter

Using the same method, generate a month based on the Complaint type and a quarterly time chart. The results show that the type of complaint is dominant in every month, followed by the Street/Sidewalk type.

Since the total amount of noise complaints in the second quarter was the largest, we visualized the total number of daily complaints in the second quarter. An interesting phenomenon was found, since the amount of noise complaints on the rest day was significantly higher than the working day, and this difference was most noticeable in June. This shows that people in New York City will make more noise on the day off, and more people will complain (Fig.4.1.3).

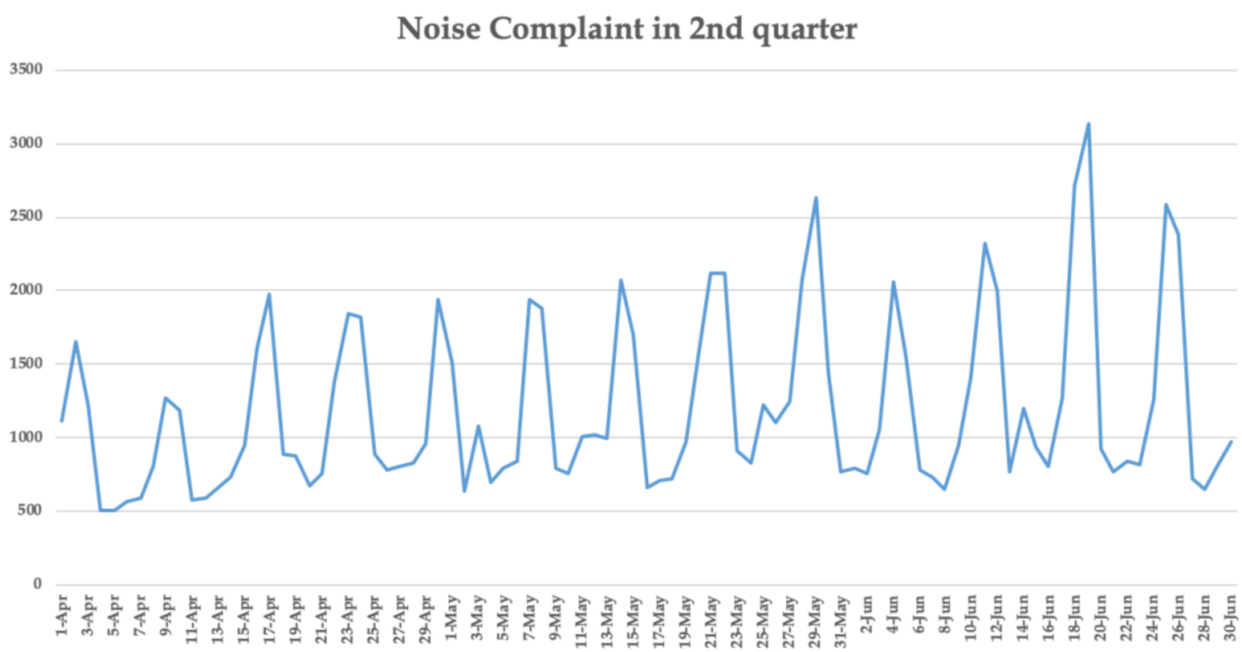

Figure 4.1.3 Noise Complaint in 2nd quarter

Finally, based on the 24 hours of the year, a graph of the total number of complaints per hour is generated. From the picture, around 1AM, 11AM and 11PM may be the top three highest number of complaints, as many clubs are too noisy after nighttime while transportation would be busy at around 11AM. As shown in the figure, the volume of complaints peaked at noon and early morning. During the night, people in New York City need a better rest, so they are particularly sensitive to noise and generate more complaints (Fig.4.1.4).

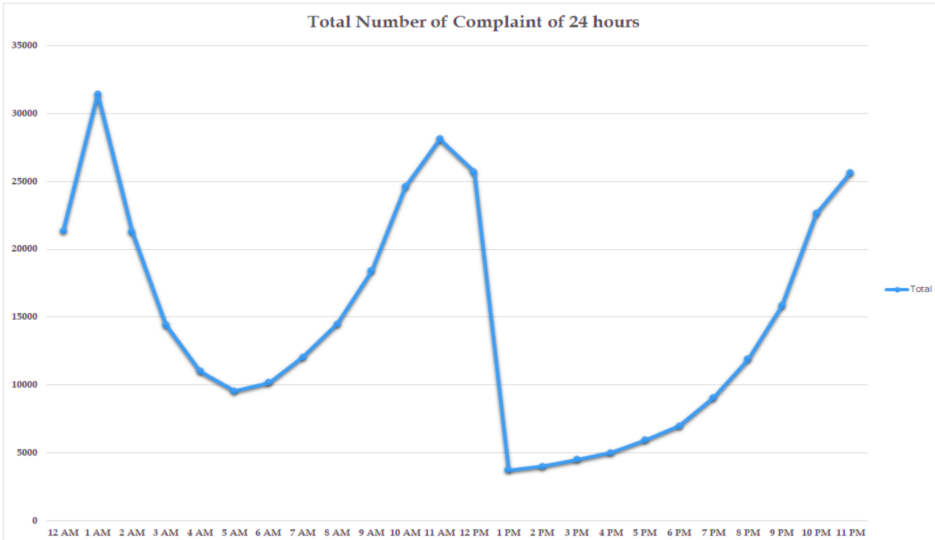

Figure 4.1.4 Total Number of Complaint of 24 hours

4.2 Spatial Analytics

4.2.1 The analysis results of 311 noise data in New York City

1) In terms of the counts in each census tracts results

There is a total of 358161 noise complaints in New York City in 2016, and a total of 2156 census tracts with noise complaint data in the census tracts in New York City. After comparison, we find that the 10 census tracts without noise complaints are parks, islands and airports areas which located in the southern part of Queens and Brooklyn and Jamaica Bay.

The census tract with the highest noise complaints is located in the Fort Greene in Brooklyn with a number of 4133. The second and third one is the Soundview-Bruckner in the Bronx and the Washington Heights North in Manhattan. The higher noise complaints are concentrated in Manhattan and the northern part of Brooklyn. The lower noise complaints occurred in most of the areas in Queens and State Island, the northeast part of the Bronx, and the northern part of Brooklyn (Fig.4.2.1- Fig.4.2.4).

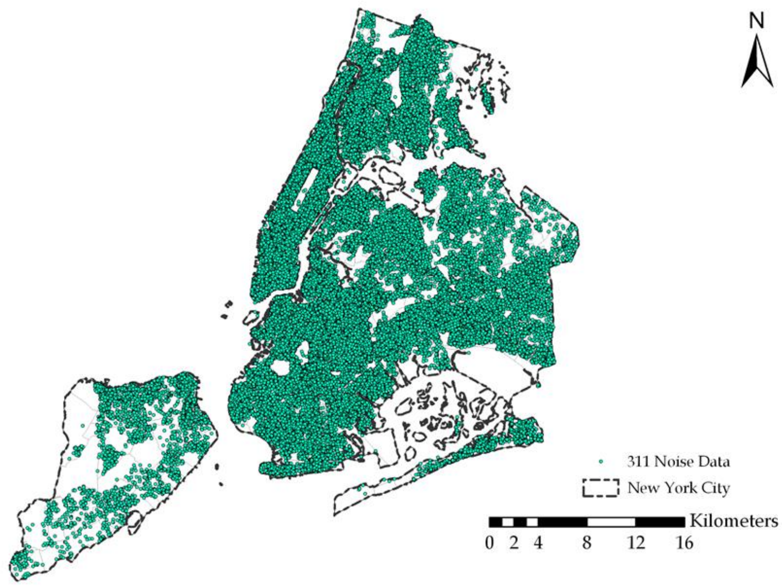

Figure 4.2.1 The distribution map of 311 Noise data in New York City

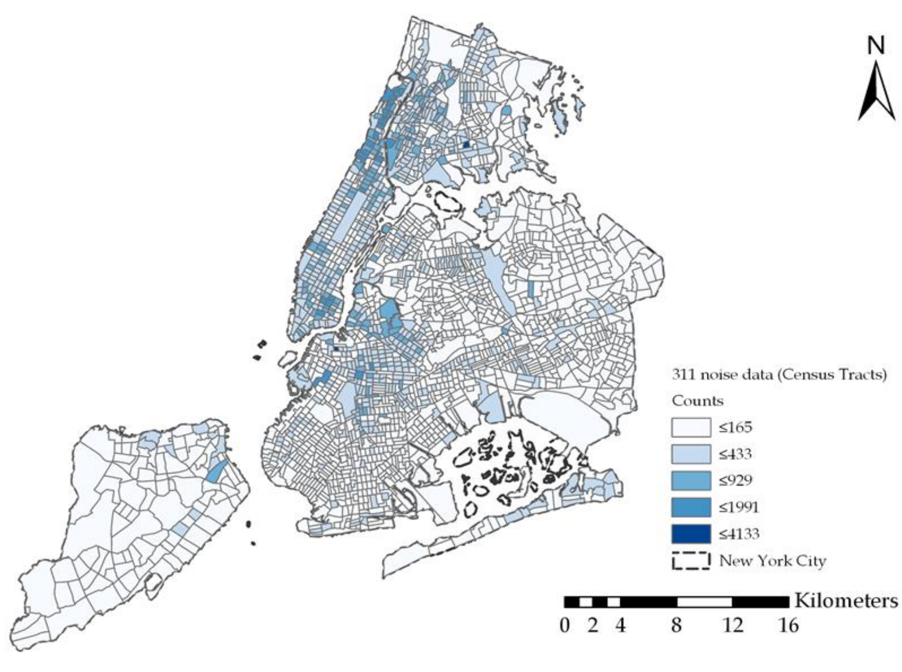

Figure 4.2.2 The distribution map of 311 Noise Data in the census tracts of New York City

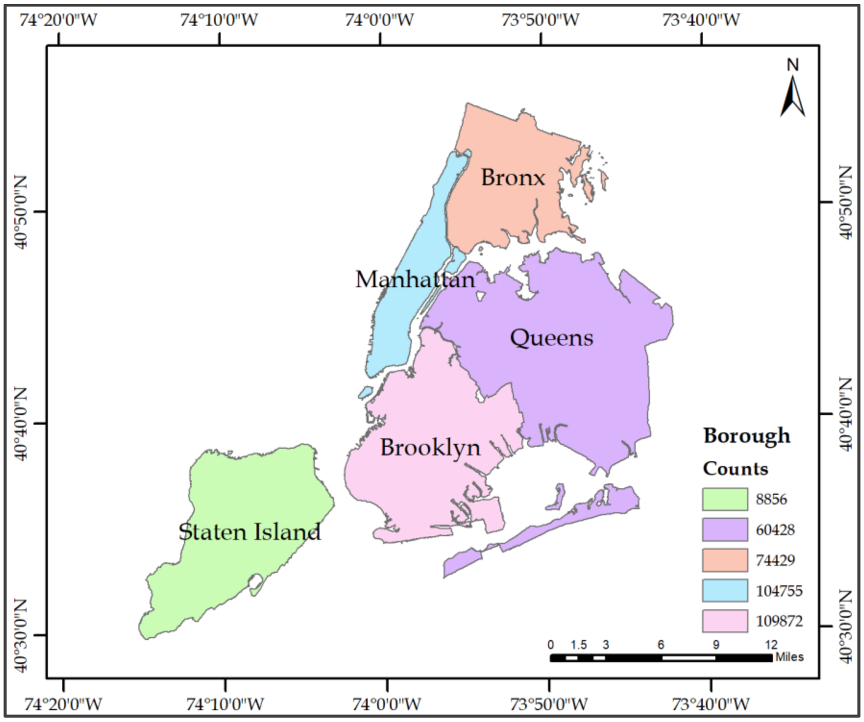

Figure 4.2.3 The distribution map of 311 Noise Data in the boroughs of New York City

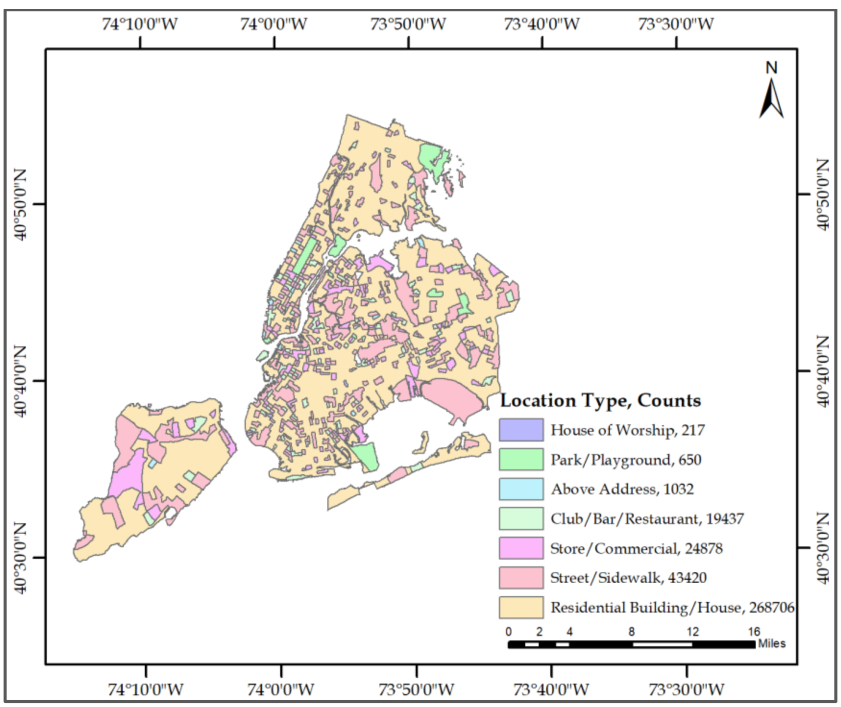

Figure 4.2.4 The distribution map of 311 Noise Data New York City in terms of location types

2) In terms of the Average Nearest Neighbor results

The table below displays the Nearest Neighbor Ratio as 0.104538, Z-Score as -1025.220153, P-Value as 0.000000 which is a very low value, showing that almost the 99% level of the pattern isn't a random spatial processes consequence. There is a less than 1% likelihood given by the z-score of -1025.220153 that this clustered pattern could be a random result. There is clear evidence of clustering for noise data in New York City (Tbl.4.2.1).

Table 4.2.1 Global Moran's I Summary of 311 Noise Data in census tracts level

Moran's Index:	0.300387
Expected Index:	-0.000464
Variance:	0.000013
z-score:	82.844313
p-value:	0.000000

3) In terms of the Optimized Hot Spot Analysis results

The results indicate statistically significant hot spots in the northern part of Brooklyn, the northern as well as southern part of Manhattan, the southwest part of Bronx (99% confidence). The cold spots, as expected, centrally situated in the southern part of Brooklyn, the northern, the eastern as well as northern part of Queens (Fig.4.2.5).

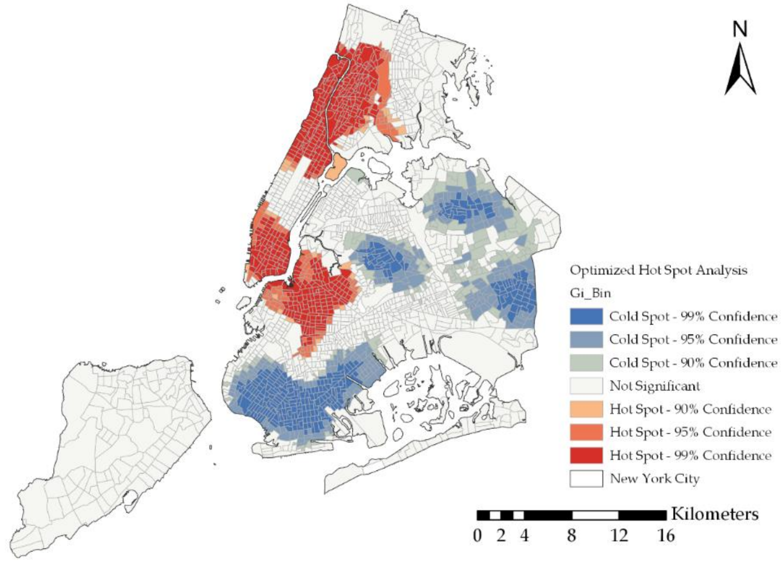

Figure 4.2.5 The Optimized Hot Spot Analysis map of 311 Noise Data in New York City

4) In terms of the Spatial Autocorrelation (Global Moran's I) results

The table below displays the Moran's Index as 0.300387, Z-Score as 82.844313, P-Value as 0.000000 which is a very low value, showing that almost the 99% level of the pattern isn't a random spatial processes consequence. The Z-Score is 82.844313 meaning that there is a less than 1% likelihood that this clustered pattern could be a random process consequence. There is clear evidence of clustering for noise events above address. (Tbl.4.2.2)

Table 4.2.2 Global Moran's I Summary of 311 Noise Data in census tracts level

Moran's Index:	0.300387
Expected Index:	-0.000464
Variance:	0.000013
z-score:	82.844313
p-value:	0.000000

5) In terms of the Cluster and Outlier Analysis (Anselin Local Moran's I) results

The hot spots in yellow color represent the High-High Cluster areas which are concentrated in Manhattan and the northern part of Brooklyn. The cold spots in light green refer to the Low-Low luster areas which are centrally located in most of the areas in Queens and State Island as well as the southern part of Brooklyn.

The grey and purple census tracts reflect the outliers which mainly located in Howard Beach and Flushing Meadows Corona Park, the central part of the Bronx and the northern part of Brooklyn separately. Low-High (LH) category of outlier meaning that the census tract itself is low in terms of the figure of noise complaints but its surrounding census tracts contain a higher noise complaints volume.

As for the High-Low outliers, these are census tracts where contain a higher noise complaints number but the census tracts surrounding them have less or not statistically significant noise complaints data. They have unexpected high noise complaints. The result shows in the below (Fig.4.2.6).

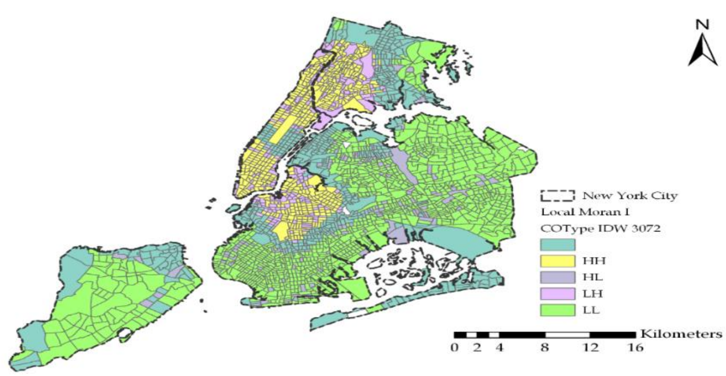

Figure 4.2.6 The Anselin Local Moran's I map of 311 Noise Data in New York City

6) In terms of the Emerging Hot Spot Analysis results

According to the result, most of the census tracts show no patterns. However, it is deserved to notice that there was no cold area. In other words, there was an increasing complaint in New York City. This somewhat shows that the city is getting noisy. By further analysis, we find that the census tracts located between the Bronx and Manhattan show an intensifying hot trend (Fig.4.2.7).

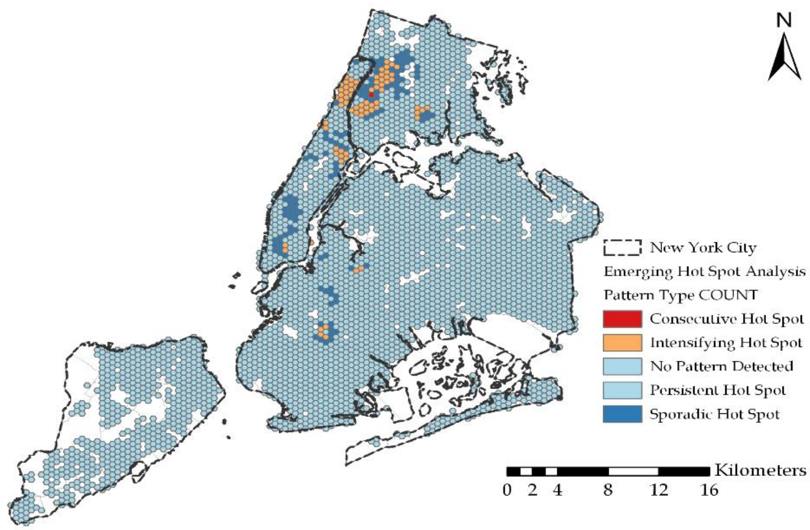

Figure 4.2.7 The Emerging Hot Spot Analysis map of 311 Noise Data in New York City

7) In terms of the Local Outlier Analysis results

According to the result, over 28% of the area shows Low- Low Cluster, which mainly occurred in the area without Manhattan. 48% of the area showed multiple types. There are some outliers which need to be further discussed why they occur. After combing with various types of online information about New York City, we think the race, income, population density may account for the problem (Fig.4.2.8).

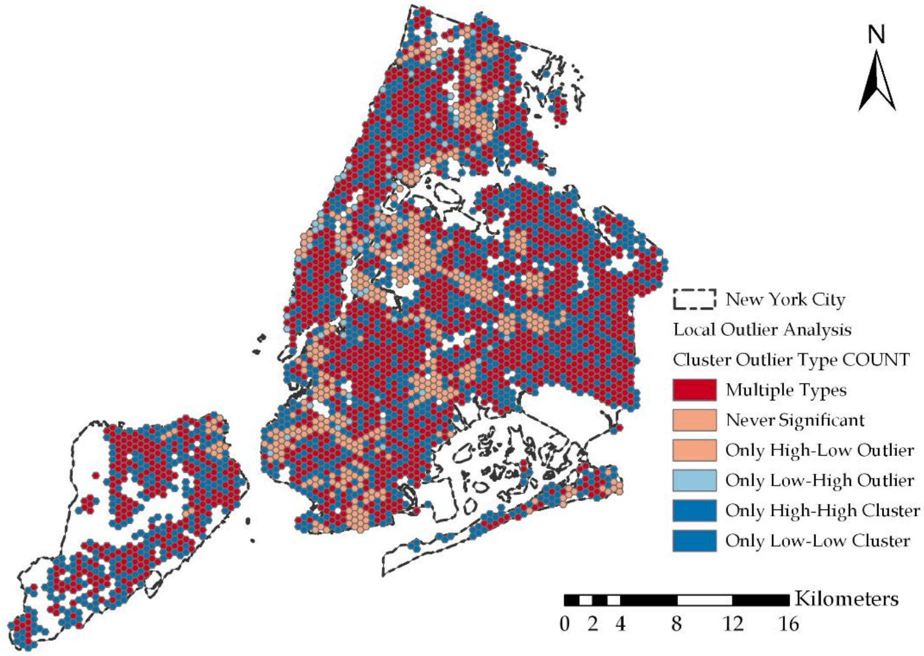

Figure 4.2.8 The Local Outlier Analysis map of 311 Noise Data in New York City

4.2.2 The analysis results of each location type noise data in New York City

4.2.2.1 The analysis results of Above address type noise data in New York City

1) In terms of the counts in each census tracts results

There is a total of 821 noise complaints of the location type for the Above Address in 2016, and a total of 327 census tracts with noise complaint data in the census tracts in New York City. From the comparison map we find that areas with higher complaints are concentrated in Manhattan, northwest of Brooklyn and west of Queens. Most of the complaints of the Above Address type occur in the white area because the per capita income of the white area is higher, and the noise of the Location type, which is the Location type, is mostly generated by helicopters, and the helicopters are usually affordable in areas where the per capita is high. Therefore, the noise address type of the Above address type in the high-income white area is significantly higher than other areas.

The highest number of 105 appears in census tracts 159 (Manchester's Lincoln Square), the second highest value of 40 appears in census tracts 155 (Brooklyn's Park Slope-Gowanus), the third highest A value of 38 appears in census tracts 317.04 (Battery Park City-Lower Manhattan in Manhattan (Battery Park City) is located on the Hudson River, a residential area consisting mainly of high-rise buildings. The casual restaurants and bars in this area will always attract office workers from nearby financial districts. On weekends, the pace of the area is slow, but locals and tourists will head to the contiguous Brookfield Place shopping center. And the Riverside Park Trail for cycling. The lowest number of areas are mainly distributed on Staten Island (similar to Lantau Island), most of the Queens area, central of Brooklyn, south and most of Bronx (Fig.4.2.9).

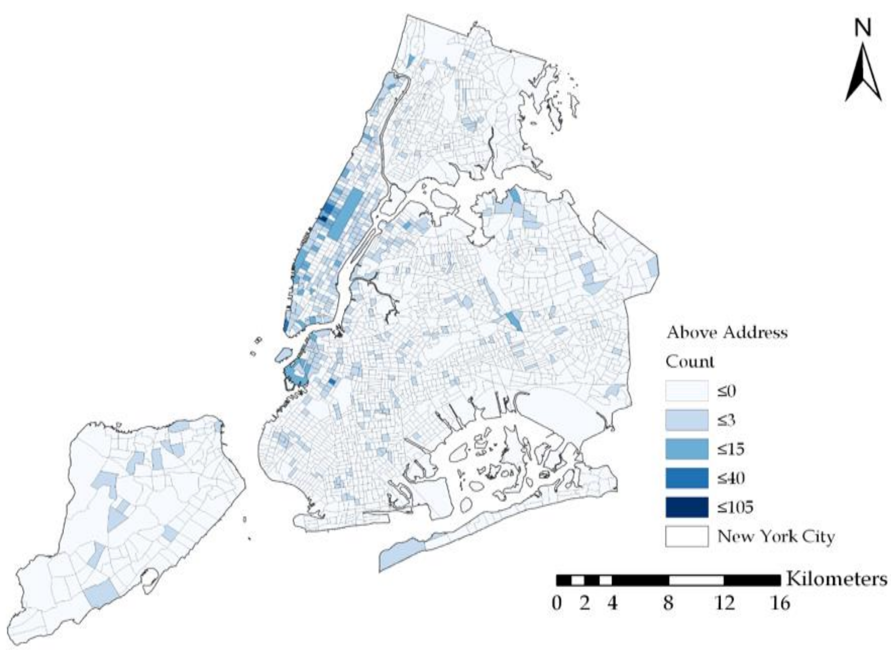

Figure 4.2.9 The distribution map of Above Address type noise data in New York City

2) In terms of the Optimized Hot Spot Analysis results

The results indicate statistically significant hot spots in lower Manhattan and Upper West Side in Manhattan (99% confidence). The cold spots, as expected, are centrally located in the county near the downtown area (Fig.4.2.10).

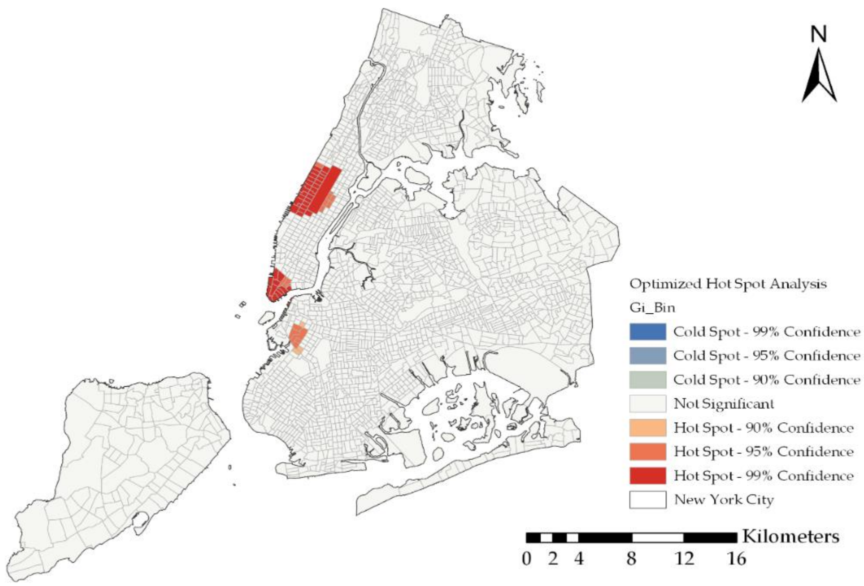

Figure 4.2.10 The Optimized Hot Spot Analysis map of Above Address type noise data in New York City

3) In terms of the Spatial Autocorrelation (Global Moran's I) results

The table below displays the Moran's Index as 0.063653, Z-Score as 24.195414, P-Value as 0.000000. The p-value of 0.000000 is quite low, indicating that we can be confident at the 99% level that the pattern we're seeing is not the result of random spatial processes. Given the z-score of 24.20, there is a less than 1% likelihood that this clustered pattern could be the result of random chance. There is clear evidence of clustering for noise events above address. (Tbl.4.2.3)

Table 4.2.3 Global Moran's I Summary of Above Address

Moran's Index:	0.066861
Expected Index:	-0.003067
Variance:	0.000088
z-score:	7.455487
p-value:	0.000000

4.2.2.2 The analysis results of Club/Bar/Residential type noise data in New York City

1) In terms of the counts in each census tracts results

There is a total of 21,606 noise complaints for the location for Club/Bar/Restaurant in 2016. There are a total of 1082 census tracts with noise complaint data in the census tracts in New York City. From the contrast maps we find that areas with higher complaints are concentrated in Manhattan (wide coverage! Strong), west of Bronx, north of Brooklyn.

The highest number of 382 appears in census tracts 18 (Manhattan's Chinatown (Vibrant Chinatown is a densely populated community where many Chinese and Southeast Asian restaurants attract old ladies and tourists.)), the second highest value of 335 appeared in census tracts 509 (Williamsburg) is a trendy neighborhood in Brooklyn, attracting young people and fashion people to visit the trendy boutiques, trendy cafes and lively restaurants. The bar and concert hall make the nightlife scene lively, the third highest value is 261, appears in census tracts 34. The lowest number of areas are mainly distributed on Staten Island (similar to Lantau Island), in the middle of Queens, in the east, in the south of Brooklyn and in the eastern part of Bronx (Fig.4.2.11).

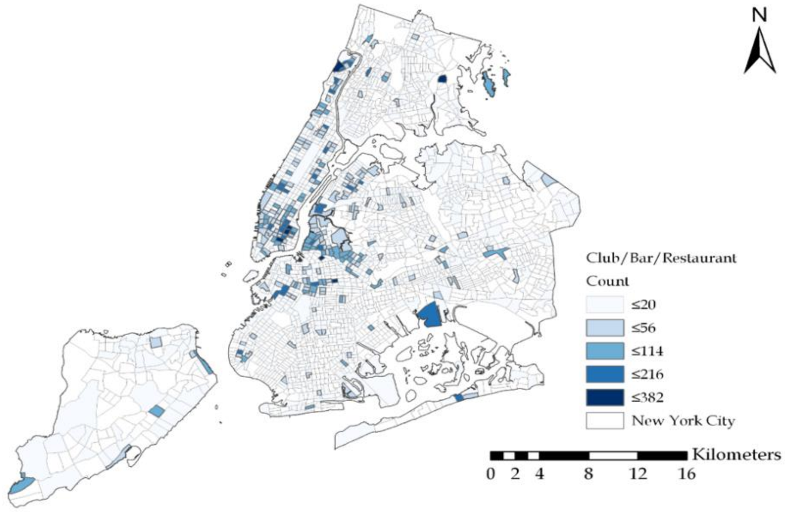

Figure 4.2.11 The distribution map of Club/Bar/Restaurant type noise data in New York City

2) In terms of the Optimized Hot Spot Analysis results

The results indicate statistically significant hot spots in lower Manhattan, East Village, Greenwich Village, Inwood in Manhattan, Bumbo, Vinegar Hill, Williamsburg in Brooklyn (99% confidence). The cold spots are centrally located in East Flatbush in Brooklyn (Fig.4.2.12).

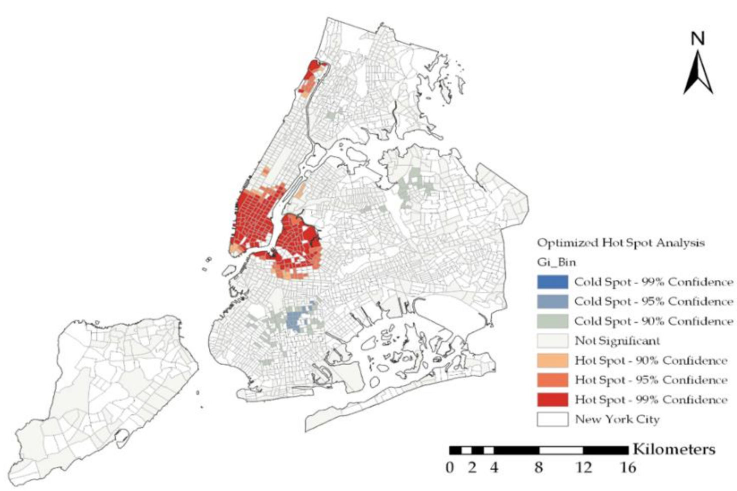

Figure 4.2.12 The Optimized Hot Spot Analysis map of Club/Bar/Restaurant noise data in New York City

3) In terms of the Spatial Autocorrelation (Global Moran's I) results

This table below displays the Moran's Index as 0.187742, Z-Score as 29.895877, P-Value as 0.000000. The p-value of 0.000000 is quite low, indicating that we can be confident at the 99% level that the pattern we're seeing is not the result of random spatial processes. Given the z-score of 29.90, there is a less than 1% likelihood that this clustered pattern could be the result of random chance. There is clear evidence of clustering for noise events occurred in car/bar/restaurant. (Tbl.4.2.4)

Table 4.2.4 Global Moran's I Summary of Car/Bar/Restaurant

Moran's Index:	0.020563
Expected Index:	-0.001326
Variance:	0.000027
z-score:	4.246916
p-value:	0.000022

4.2.2.3 The analysis results of House of Worship type noise data in New York City

1) In terms of the counts in each census tracts results

In 2016, there were a total of 914 noise complaints for the House of Worship, and a total of 372 census tracts with noise complaint data in the census tracts in New York City. From the comparison map we find that the higher complaints are concentrated in the north of Manhattan, in the west of Bronx, in the majority of Brooklyn and in the Midwest of Queens. However, the overall distribution is relatively average, and there is no obvious correlation with factors such as race and income.

The highest number of 83 appears in census tracts 1188 (Cypress Hills City Line, Brooklyn), the second highest value of 63 appears in census tracts 241 (Bedford, Brooklyn), and the third highest value of 29 appears in census tracts 373 (Brex's Tremont) Park). The lowest number of areas are mainly distributed on Staten Island (similar to Lantau Island), most of the Queens area, south of Brooklyn and eastern Bronx (Fig.4.2.13).

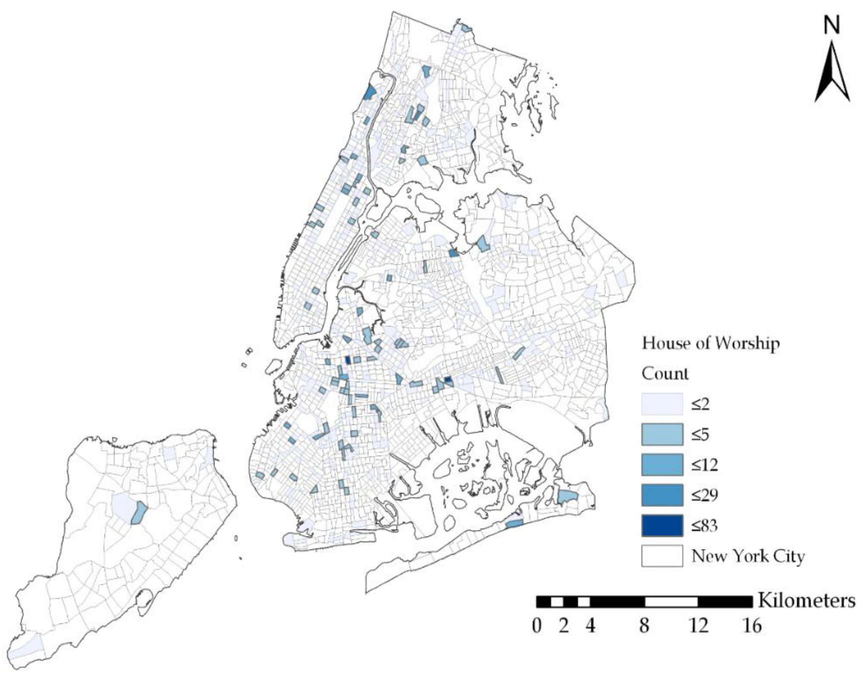

Figure 4.2.13 The distribution map of House of Worship type noise data in New York City

2) In terms of the Optimized Hot Spot Analysis results

The results indicate hot spots in the North West corner of East New York in Brooklyn (90% confidence) (Fig.4.2.14).

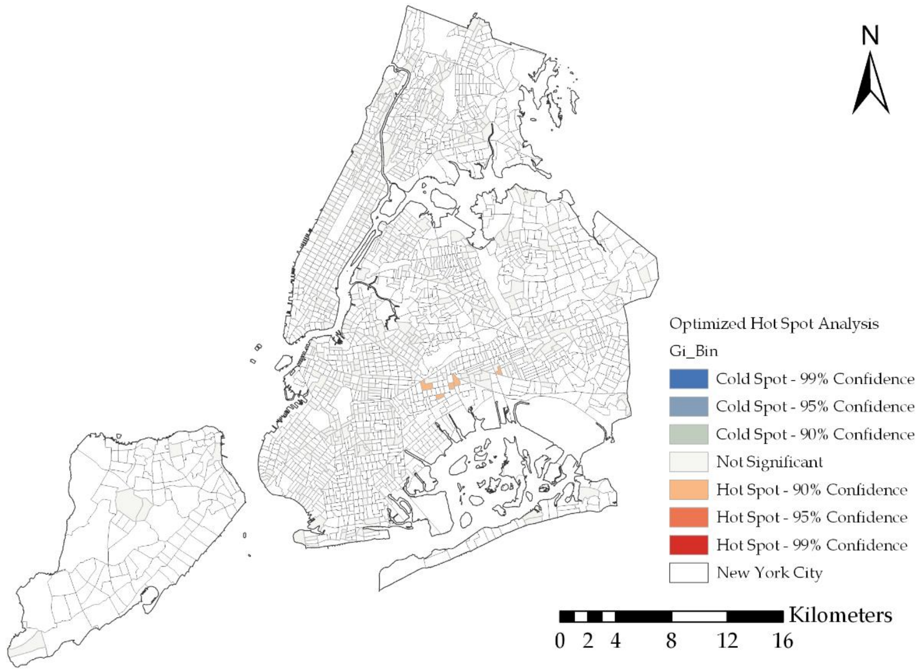

Figure 4.2.14 The Optimized Hot Spot Analysis map of House of Worship noise data in New York City

3) In terms of the Spatial Autocorrelation (Global Moran's I) results

This table below displays the Moran's Index as -0.004852, Z-Score as -0.368342, P-Value as 0.712618. The p-value of 0.712618 is quite high, so given the z-score of -0.37, the pattern would not appear to be significantly different than random. There is clear evidence of random pattern for noise events occurred in house of worship. (Tbl.4.2.5)

Table 4.2.5 Global Moran's I Summary of House of Worship

Moran's Index:	-0.004852
Expected Index:	-0.002695
Variance:	0.000034
z-score:	-0.368342
p-value:	0.712618

4.2.2.4 The analysis results of Park/Playground type noise data in New York City

1) In terms of the counts in each census tracts results

There was a total of 3,820 noise complaints for Park/Playground, and a total of 755 census tracts have noise complaint data in census tracts in New York City. From the comparison map we find that areas with higher complaints were concentrated in Manhattan, north of Brooklyn and central Queens and rockaway peninsula.

The highest number of 239 appears in census tracts 96 (Sunset Park East in Brooklyn (The residents of the neighborhood are mainly Salvadorans and Chinese, as well as other Hispanic Americans, Indians and Norwegians.)), and the second highest value of 119 appears in census tracts 59 (West Village of Manhattan) Designer boutiques and trendy restaurants appeal to fashion trends. This historic art district also has piano bars, dance halls and theaters. The third highest value of 70 appears in census tracts 128.01 (Dyker Heights of Brooklyn). The lowest number of areas are mainly distributed on Staten Island (similar to Lantau Island), east of Queens, south, south of Brooklyn (Fig.4.2.15).

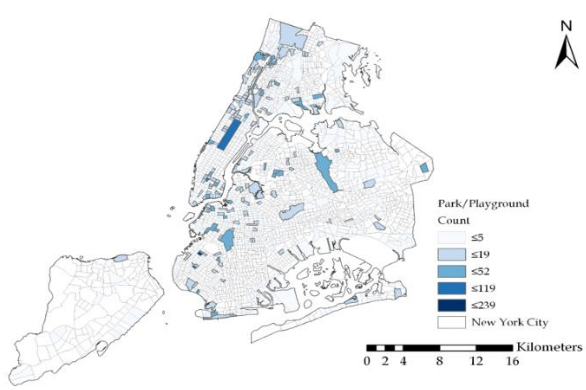

Figure 4.2.15 The distribution map of Park/Playground type noise data in New York City

2) In terms of the Optimized Hot Spot Analysis results

The results indicate statistically significant hot spots in some tract in Sunset Park, Borough Park and Flatbush-Ditmas Park in Brooklyn (99% confidence) (Fig.4.2.16).

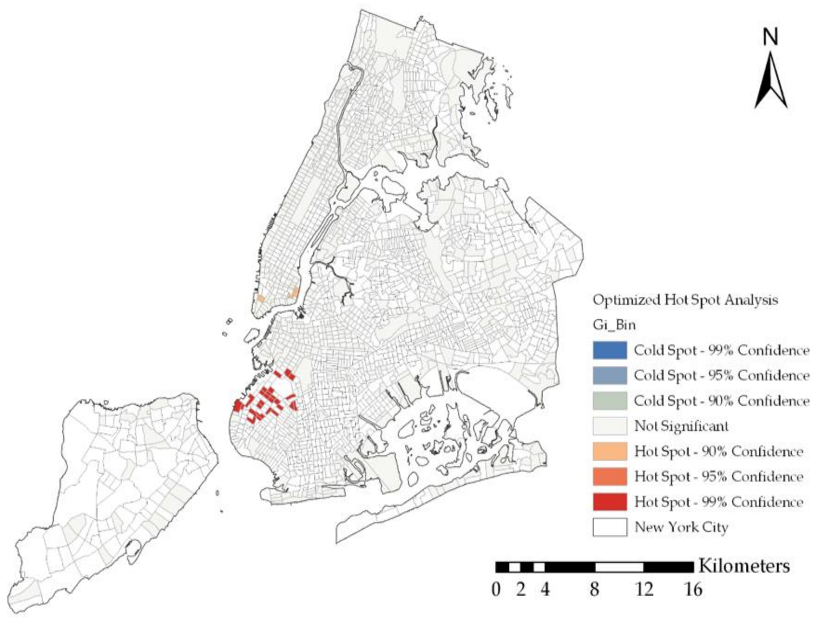

Figure 4.2.16 The Optimized Hot Spot Analysis map of Park/Playground noise data in New York City

3) In terms of the Spatial Autocorrelation (Global Moran's I) results

This table below displays the Moran's Index as 0.020563, Z-Score as 4.246916, P-Value as 0.000022. The p-value of 0.000022 is quite low, indicating that we can be confident at the 99% level that the pattern we're seeing is not the result of random spatial processes. Given the z-score of 4.25, there is a less than 1% likelihood that this clustered pattern could be the result of random chance. There is clear evidence of clustering for noise events occurred in park/playground. (Tbl.4.2.6)

Table 4.2.6Global Moran's I Summary of Park/Playground

Moran's Index:	0.020563
Expected Index:	-0.001326
Variance:	0.000027
z-score:	4.246916
p-value:	0.000022

4.2.2.5 The analysis results of Residential Building/House type noise data in New York City

1) In terms of the counts in each census tracts results

There was a total of 220,272 noise complaints for the Residential location/Residential Building/House in 2016, and a total of 2,133 census tracts with noise complaint data in the census tracts in New York City. Most of Manhattan's northern and Brooks are mainly Latino and black, and from the comparison map we can find that the higher complaint areas are concentrated in the north and south of Manhattan, the majority of Bronx, northeast of Brooklyn and Rockaway peninsula in Queens. The highest number of 4125 appeared in census tracts 29.01 (Brooklyn's Fort Greene (at Downtown Brooklyn, a lot of universities)), and the second highest value of 1314 appeared in census tracts 223.01 (Manhattan's Manttanvile) the third highest value of 1215 appears in census tracts 287 (Manchester's Marble Hill-Inwood (Inwood is an area north of Manhattan, New York City.)). The lowest number of areas are mainly distributed on Staten Island (similar to Lantau Island), most of Queens, the west of Brooklyn, the south and the northeast corner of Bronx, and the south of Brooklyn is a region with more whites and East Asians, combined with previous analysis. It shows that the number of complaints in white and East Asian regions is less than the figure in Latino and black regions (Fig.4.2.17).

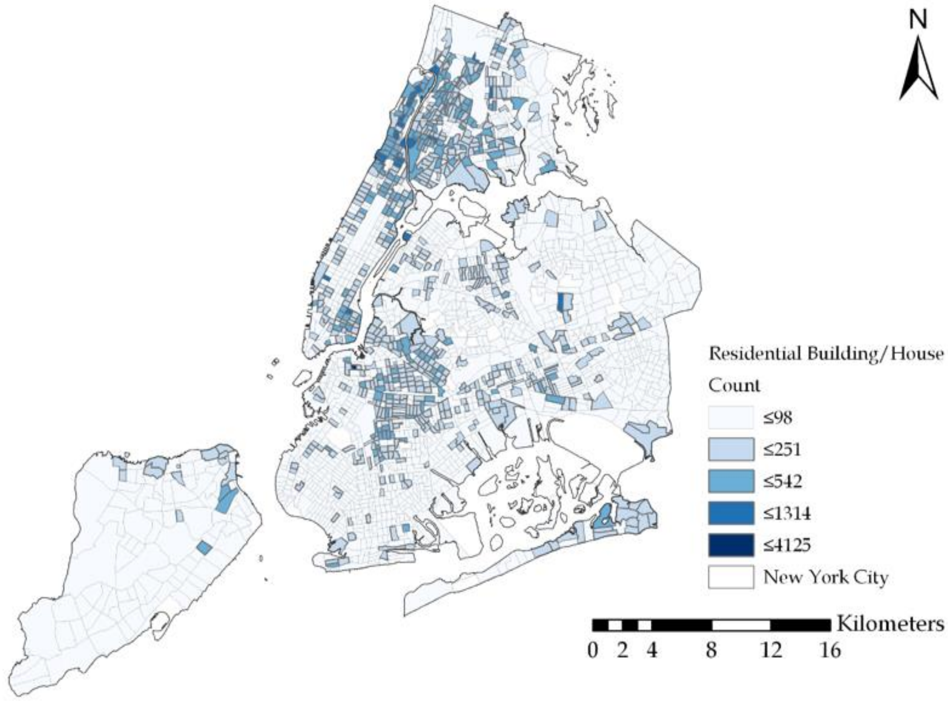

Figure 4.2.17 The distribution map of Residential Building/House type noise data in New York City

2) In terms of the Optimized Hot Spot Analysis results

The results indicate statistically significant hot spots in Inwood, Fort George, Washington Heights in Manhattan, West Bronx in Bronx, Williamsburg, Bedford-Stuyvesant, Brooklyn Heights in Brooklyn (99% confidence). The cold spots, as expected, are centrally located in Bensonhurst, Gravesend, Sheepshead Bay in Brooklyn (99% confidence) (Fig.4.2.18).

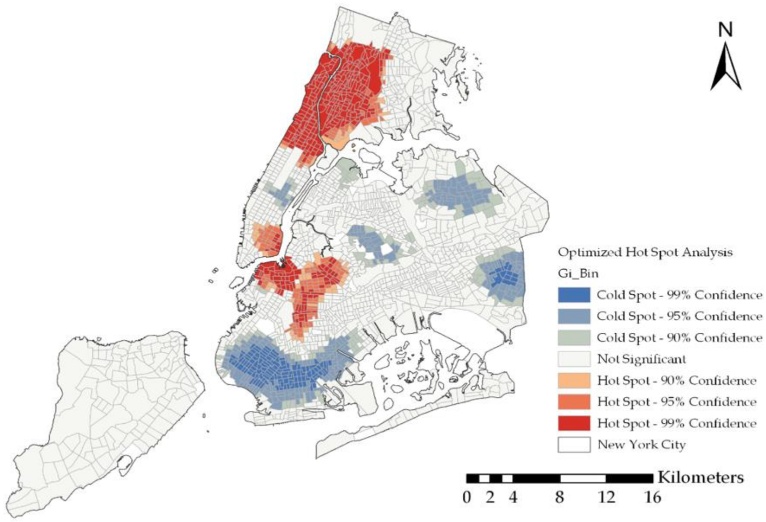

Figure 4.2.18 The Optimized Hot Spot Analysis map of Residential Building data in New York City

3) In terms of the Spatial Autocorrelation (Global Moran's I) results

This table below displays the Moran's Index as 0.212727, Z-Score as 61.237649, P-Value as 0. The p-value of 0is quite low, indicating that we can be confident at the 99% level that the pattern we're seeing is not the result of random spatial processes. Given the z-score of 61.24, there is a less than 1% likelihood that this clustered pattern could be the result of random chance. There is clear evidence of clustering for noise events occurred in residential building/house. (Tbl.4.2.7)

Table 4.2.7 Global Moran's I Summary of Residential Building/House

Moran's Index:	0.212727
Expected Index:	-0.000469
Variance:	0.000012
z-score:	61.237649
p-value:	0.000000

4.2.2.6 The analysis results of Street/Sidewalk type noise data in New York City

1) In terms of the counts in each census tracts results

In 2016, there were a total of 27,021 noise complaints for Store/Commercial, and a total of 1518 census tracts with noise complaint data in census tracts in New York City. From contrast maps we find that areas with higher complaints are concentrated in the north and south of Manhattan, north of Brooklyn. The highest number is 1050 in census tracts 70 (BRONX's Soundview-Bruckner), and the second one is 811 in census tracts 119 (Brooklyn's Gowanus (formerly located around Brooklyn's Gowanus Canal) District, Gowanus is a developing area, mixed with classic townhouses and converted warehouses. After dark, this community is becoming more vibrant as young professionals and crowds emerge from the nearby Barclays Center stadium to the trendy bars, casual restaurants and trendy music venues around Fourth Avenue.), the third highest value 249 (Bedford, Brooklyn).

Combined with the income data distribution map, we can see that the complaints about Store/Commercial are related to the income level (Fig.4.2.19). The highest income area in Manhattan is in the south, and the highest income area in Brooklyn is in the north. This is exactly the same as the high frequency hotspot of Store/Commercial complaints with one-to-one correspondence. The lowest number of areas are mainly distributed on Staten Island (similar to Lantau Island), most of Queens, the central part of Brooklyn, the south and most of Bronx (Fig.4.2.20).

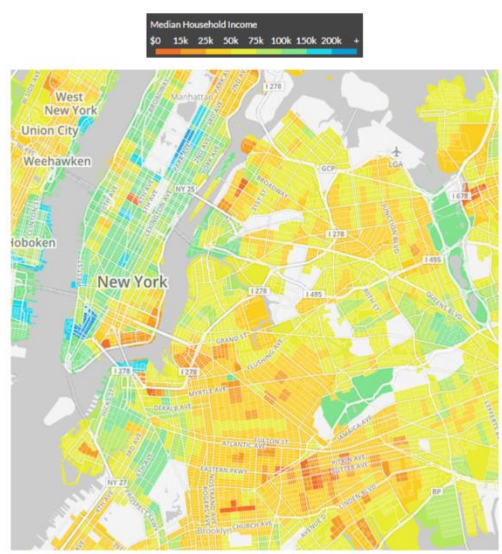

Figure 4.2.19 Distribution of income level in NYC

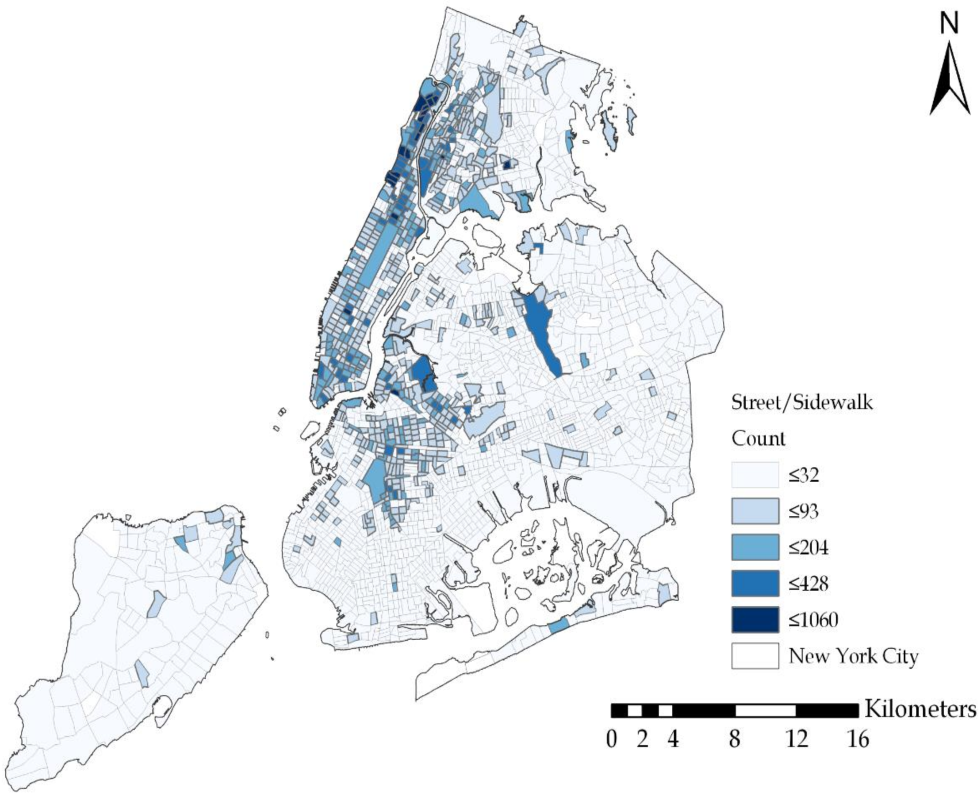

Figure 4.2.20 The distribution map of Street/Sidewalk type noise data in New York City

2) In terms of the Optimized Hot Spot Analysis results

The results indicate statistically significant hot spots in lower Manhattan and Midtown in Manhattan and Greenpoint, East Williamsburg and Williamsburg in Brooklyn (99% confidence). (Fig.4.2.21)

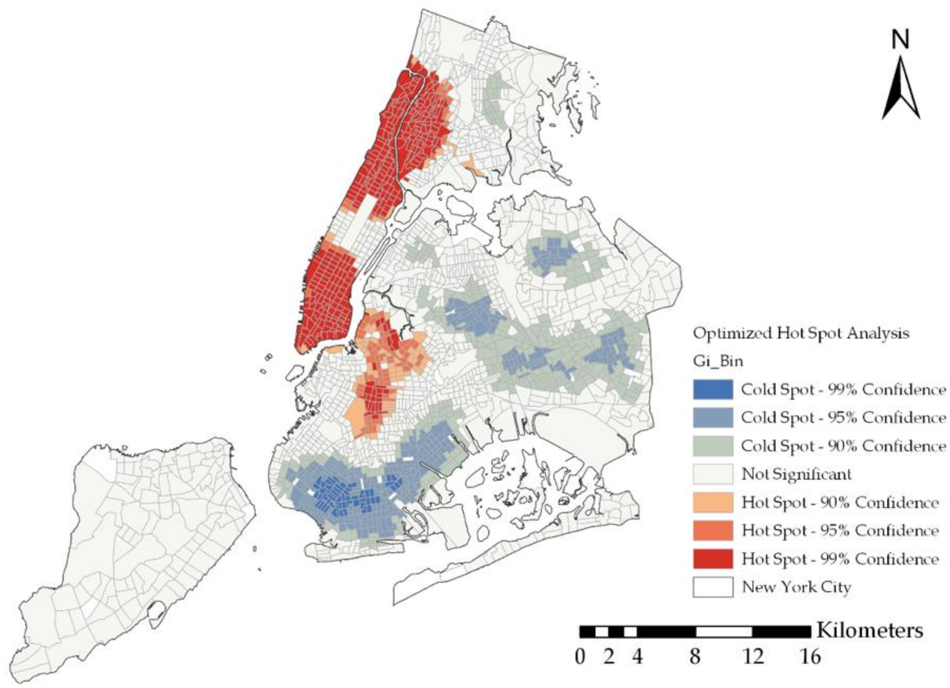

Figure 4.2.21 The Optimized Hot Spot Analysis map of Street/Sidewalk type noise data in New York City

3) In terms of the Spatial Autocorrelation (Global Moran's I) results

This table below displays the Moran's Index as 0.086622, Z-Score as 18.865443, P-Value as 0.000000. The p-value of 0.000000 is quite low, indicating that we can be confident at the 99% level that the pattern we're seeing is not the result of random spatial processes. Given the z-score of 18.87, there is a less than 1% likelihood that this clustered pattern could be the result of random chance. There is clear evidence of clustering for noise events occurred in store/commercial. (Tbl.4.2.8)

Table 4.2.8 Global Moran's I Summary of Street/Sidewalk

Moran's Index:	0.303464
Expected Index:	-0.000472

Variance:	0.000014
z-score:	82.269640
p-value:	0.000000

4.2.2.7 The analysis results of Store/Commercial type noise data in New York City

1) In terms of the counts in each census tracts results

In 2016, there were a total of 838,86 noise complaints for Street/Sidewalk, and a total of 2,118 census tracts with noise complaint data in census tracts in New York City. From the comparison map we can find that the higher complaint areas were concentrated in Manhattan, southwest of Bronx, central and northern Brooklyn, and Flushing Meadows Corona Park in Queens. The data shows that the ethnic composition of the Manhattan area is quite different from other regions, with 41% of the population being Asian, 32% being non-Hispanic, 19% Hispanic and 6% black; and 43% of the population being immigrants. The reason is that Chinatown's population accounts for 55% of Lower Manhattan. From the optimized profile of the Optimized Hot Spot map, it can be seen that Asians, Latinos and blacks in Manhattan are more likely to have noise complaints, which may be related to their lifestyle, such as blacks and Latinos, they prefer to go to dance parties on the street, which creates a higher noise complaint with street/sidewalk location type.

The highest number is 1060 in census tracts 70 (BRONX's Soundview-Bruckner), the second highest is 784 in census tracts 285 (Manhattan's Washington Heights North), and the third highest is 746 in census tracts 287 (Manhattan's Washington Heights North (with restaurant and bar school, etc., see below)). The lowest number of areas are mainly distributed on Staten Island (similar to Lantau Island), most of the Queens area, south of Brooklyn and eastern Bronx (Fig.4.2.22).

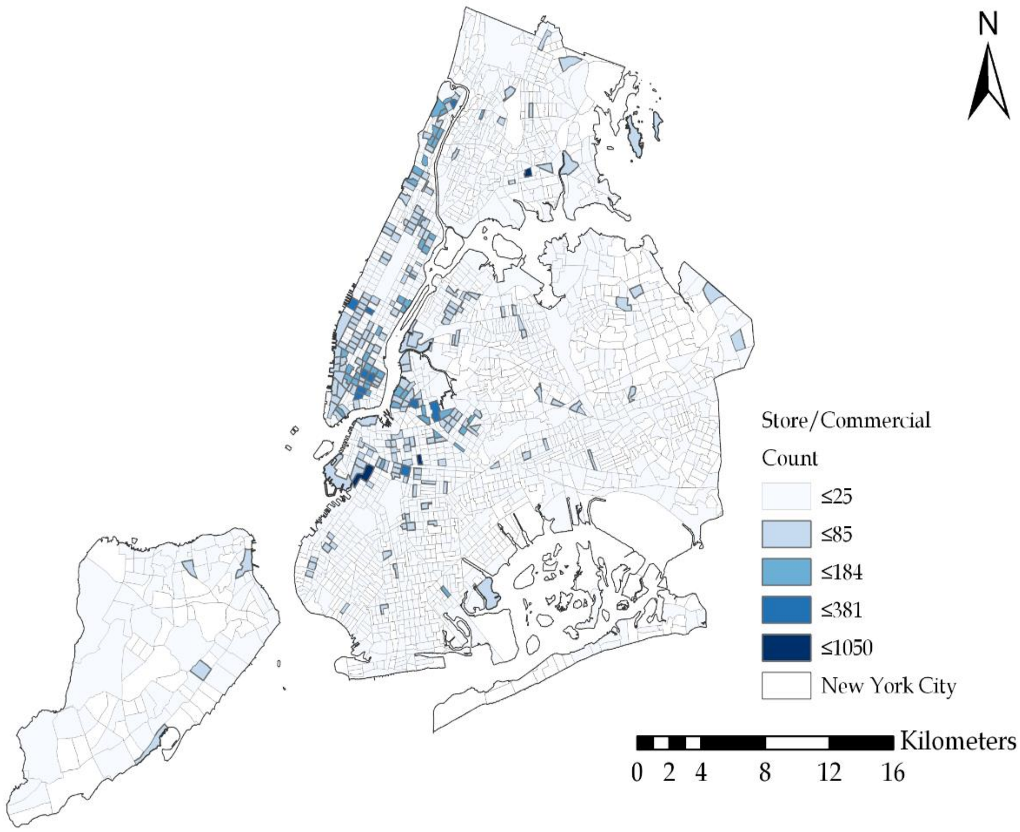

Figure 4.2.22 The distribution map of Store/Commercial type noise data in New York City

2) In terms of the Optimized Hot Spot Analysis results

The results indicate statistically significant hot spots in Lower Manhattan, Midtown, Upper Manhattan in Manhattan, West Bronx in Bronx, Williamsburg, Crown Heights, Bushwick in Brooklyn (99% confidence). The cold spots in Benson Hurst, Gravesend, Sheepshead Bay in Brooklyn (99% confidence) and Murray Hill in Queens (Fig.4.2.23).

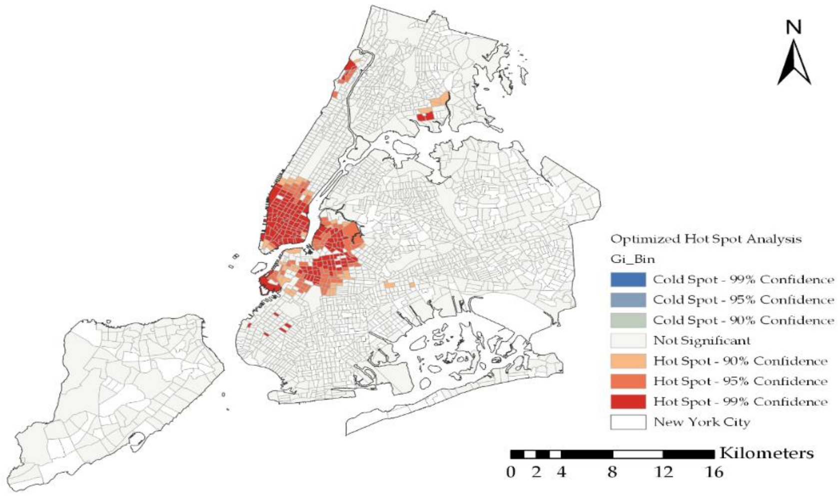

Figure 4.2.23 The Optimized Hot Spot Analysis map of Store/Commercial noise data in New York City

3) In terms of the Spatial Autocorrelation (Global Moran's I) results

This table below displays the Moran's Index as 0.303464, Z-Score as 82.269640, P-Value as 0.000000. The p-value of 0.000000 is quite low, indicating that we can be confident at the 99% level that the pattern we're seeing is not the result of random spatial processes. Given the z-score of 82.27, there is a less than 1% likelihood that this clustered pattern could be the result of random chance. There is clear evidence of clustering for noise events in street/sidewalk. (Tbl.4.2.9)

Table 4.2.9 Global Moran's I Summary of Store/Commercial

Moran's Index:	0.086622
Expected Index:	-0.000659
Variance:	0.000021
z-score:	18.865443
p-value:	0.000000

5 Conclusions

In this project, we use two kinds of methods to analyze the 311-noise data of 2016 and draw X conclusions.

1) Through the timing analysis, we can figure that the noise complaints are the largest in the mild climate period, and the Residential Building/House location type is the main source of complaints. Local governments can focus on enhancing noise monitoring and management for mild climate seasons, such as spring and fall.
2) Through the analysis of the distribution of 7 kinds of Location types, Optimized Hot Spot and Global Moran I, we can figure that the number of complaints in Manhattan is almost the highest for each location type, which is related to the population, income and ethnicity of Manhattan. Latino and black areas are prone to more complaints for Street/Sidewalk and Residential/Building while High level income area domains the Above address and Store/Commercial noise complaints.

6 Discussions

1) The timing analysis section does not perform more relevant analysis due to data defects, such as analysis of the length of complaint resolution and the type of complaint or analysis based on the week. It's necessary to learn more about data processing methods, such as using Python or MATLAB for data visualization.
2) The spatial analysis part does not combine the complaint type but excel has an analysis based on the complaint type. If we combine the analysis of the complaint type in spatial analysis, we may get more relevant conclusions.
3) In-depth analysis is not carried out in combination with accurate population, ethnicity, income, and other factors that may affect people's ability to dial 311 for noise complaints. It is simply a comparison of government-published population, ethnicity, and other maps. Lack of local background knowledge, the analysis is not deep enough. Any research needs to have a good background knowledge of the study area.

7 References

[1]. Zheng Y, Liu T, Wang Y, et al. [ACM Press the 2014 ACM International Joint Conference - Seattle, Washington (2014.09.13-2014.09.17)] Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp \"14 Adjunct - Diagnosing New York city\"s noises with ubiquitous data[J]. 2014:715-725.

[2]. ArcGIS Pro 2.2.4 (2018, November) [DB/OL]. Retrieved December 19, 2018, from http://pro.arcgis.com/en/pro-app

[3]. Thomas P. DiNapoli. Noise in New York City Neighborhoods[Z]. OFFICE OF THE NEW YORK STATE COMPTROLLER, 2018.

[4]. The Official Website of the City of New York [EB/OL]. Retrieved December 19, 2018, from: https://www1.nyc.gov/311

[5]. Streamdata.io (2018) [DB/OL]. Retrieved December 19, 2018, from https://streamdata.io/blog/311-data-important/

[6]. Tool-reference. *ArcGIS Pro 2.2*.4 (2018, November) [DB/OL]. Retrieved December 19, 2018, from http://pro.arcgis.com/en/pro-app/tool-reference