

SMART CITIES

Open Data for Urban Planning

Louisa Ong

Cover Page

# Understanding the Impact of Gentrification on Non-Gentrified Neighborhoods with Brooklyn Crime Data

Professor Michael D. Parrott

Master of Arts Thesis

Columbia University QMSS

15 May 2020

Table of Contents

[Title Page 1](#_Toc33745991)

[1. Introduction & Rationale 4](#_Toc33745992)

[1.1 Smart Cities: Open Data for Urban Planning 4](#_Toc33745993)

[1.2 Research Question 5](#_Toc33745994)

[1.3 Case Study: Brooklyn 6](#_Toc33745995)

[2. Literature Review 7](#_Toc33745996)

[2.1 Definition & Concepts 7](#_Toc33745997)

[2.2 Methodology 8](#_Toc33745998)

[2.3 On Brooklyn 8](#_Toc33745999)

[2.4 Non-Gentrified Neighborhoods 9](#_Toc33746000)

[2.5 Hypotheses 9](#_Toc33746001)

[3. Data & Methodology 10](#_Toc33746002)

[3.1 Datasets 10](#_Toc33746003)

[3.2 Independent Variable: Gentrification 10](#_Toc33746004)

[3.3 Dependent Variable: Violent NYC Crime Rate 11](#_Toc33746005)

[3.4 Time Frame: 2007 - 2018 11](#_Toc33746006)

[3.5 Control Variables 12](#_Toc33746007)

[3.6 Methodology 12](#_Toc33746008)

[4. Expected Results 13](#_Toc33746009)

[5. Discussion 14](#_Toc33746010)

[5.1 Significance of Results for Gentrification­­ 14](#_Toc33746011)

[5.2 How Precise it was to Define Gentrification 14](#_Toc33746012)

[5.3 Using Data to Understand Areas with Less Data 14](#_Toc33746013)

[6. References 15](#_Toc33746014)

# Introduction & Rationale

## Smart Cities: Open Data for Urban Planning

As 68% of the world’s populations are expected to live in cities by 2020 (Revision of World Urbanization Prospects, 2018 - BIBLIO), the concept of “ Smart Cities” was developed as a solution to existing urban planning issues to provide a better quality of life for all citizens. Essentially, the idea behind Smart Cities is to integrate Information and Communication Technologies (ICT) to tackle critical issues relating to cities (OECD, 2017, p. 77). It is a term that is a fusion of ideas and policies and can be understood as a futuristic approach to make cities more efficient, sustainable and livable. These can involve the public sector, the private sector, educational institutions, and also an ordinary data science student.

Many municipalities and private corporations around the world have raced to develop projects and systems around this topic, which has sparked considerable interest. Open Data is an example of one of the possible initiatives under the umbrella of Smart City projects. High quality Open Data entail raw, public data that can be freely used by anyone. Several municipalities do provide public data on their platforms, but the quality of such Open Data is what distinguishes them from others. What is the sample size of the datasets? How many years of data gets published? Is the raw data published, or only the polished findings? These are the questions that researchers would ask while considering whether a dataset would be useful.

Investment into maintaining a rich statistical database and a platform for sharing this data would require a significant amount of resources and political will, undoubtedly. Still, Open Data does seem to be the backbone to developing “Smart Cities” and this research paper seeks to illustrate how Open Data can be beneficial for Urban Planning.

## Research Question

This paper will involve **Using Brooklyn Crime Data to Understand the Impact of Gentrification in Less-Gentrified Neighborhoods**. It consists of extracting the available data on Brooklyn from sources such as the U.S. Census Bureau and New York Public Safety Department and exploring how the borough has evolved overtime.

The data on Gentrification and Crime originate from different data sources but this is an example of where the benefit of Open Data becomes evident – citizens get to independently scrutinize public data and find novel ways to analyze existing information. While demographic and commercial data on Non-Gentrified Neighborhoods in Brooklyn such as Canarsie and Bay Ridge to understand the social well-being of residents in these areas may be less abundant, Crime report data is relatively available consistently throughout New York City. As such, a comparison between the two variables of gentrification and crime could shed light on the impact of gentrification for residents in neighborhoods that have not yet experienced gentrification.

Such information would then especially be useful for urban planning. Gentrification is a hot topic for debate and people tend to harbor strong feelings towards this development, with mixed opinions. The true impact of gentrification is unclear as it is experienced over a long period of time and it is also subjective – some would say that cultural preservation is vital while some would prefer modernization. For this paper, Crime is a variable that is more straightforward, in that most people would agree that having less crime is better, so at least that will be clear for this paper. Whatever it is, more insight on the impact of gentrification would help inform decisions of future projects that will involve gentrifying old neighborhoods not just in Brooklyn but elsewhere in the world. Furthermore, while it is understandable that gentrification may not always be avoidable, perhaps research could be useful for relocating residents as well.

## Case Study: Brooklyn

In this paper, Brooklyn, also known as Kings County, will feature as the setting for the analysis. While gentrification, disruptive urban planning and such social changes have occurred in the entire municipality of New York City and in other cities of the world, what Brooklyn has experienced is more focused and unique.

**Why Brooklyn - it has experienced a great deal of gentrification:** This borough of New York City has undergone extensive cycles of gentrification since the 1990s but still has neighborhoods which have not experienced gentrification. Furthermore, while many original residents were displaced, they remained within the boundaries of Brooklyn (CITE). Using supporting literature of Gentrification taking place in Brooklyn, Loretta Lee’s paper on the super-gentrification process of Brooklyn Heights describes qualitatively the concept of both gentrification and super-gentrification (Lee, 2003). Clearly, this is an area that has undergone a thorough extent of gentrification and in certain areas, multiple phases of gentrification. In fact, she even recounts the Third-wave Gentrification process (Lee, 2003, p. 2490) and this phase had occurred back in the 1990s.

**How this characteristic of Brooklyn relates to my analysis:** Having such extensive gentrification occurring in this borough does indicate that this is one of the most consequential changes in the socio-economic environment of the area. This means that other factors, such as immigration of low-income residents, which could affect this area, and Brooklyn’s crime rates would not be as influential as Gentrification. In other words, there is less noise upon the analysis of how crime rates in non-gentrified areas change.   
Furthermore, since Brooklyn is a large county where some parts are undergoing gentrification and others are not, the impact of gentrification on less-gentrified neighborhoods would not be completely lost and in fact, relatively contained and evident within the geographical area of the county.

All in all, these characteristics of Brooklyn affirm that this county would be an ideal borough, not just within New York City but also worldwide, to carry out a longitudinal geographic study on how Gentrification has affected Crime in the areas that have not yet been gentrified.

**Why I’m Researching Non-Gentrified Neighborhoods:** Non-Gentrified Neighborhoods are equally noteworthy as Gentrified Neighborhoods. The inhabitants are after all, permanent residents of New York and have resided in the city longer than wealthy new immigrants who have just bought the new apartments in for instance, Williamsburg and Greenpoint in Brooklyn. (CITE) However, Non-Gentrified Neighborhoods are not often studied, perhaps because research interests tend to correspond to economic activity and because less data is available. Regardless, using the data that is available somewhat consistently throughout the city, it is possible to get an insight to how Non-Gentrified areas have been impacted as a result of changes in neighboring areas that are undergoing Gentrification. It is important to consider both the impact of Gentrification in Gentrified Neighborhoods and Non-Gentrified Neighborhoods since the effects are more widespread than one would think.

# Literature Review

## **Definition & Concepts**

**Defining Gentrification: The term “Gentrification” represents what this paper seeks to measure using an informed amalgamation of several quantitative variables. While these variables do not explicitly state if an area is gentrified or not, they would be used to construct a Gentrification index. This index will be formulated with references to how researchers have defined this concept in their papers. To start, Arthur O’Sullivan has described this to be “a variety of changes in inner cities caused by the displacement of low-income households by high-income households”** (O’Sullivan, 2005, p.73)**. In another report, the sociologist~~,~~ Ruth Glass, is quoted to have come up with the term Gentrification to describe the changes in neighborhood, be it from income levels, racial demographics, displacement of long-term residents, increase in commercial activity, etc. Similarly, the Encyclopedia of Housing defines this phenomenon as “the process by which central urban neighborhoods that have undergone disinvestments and economic decline experience a reversal, reinvestment, and the in-migration of a relatively well-off, middle and upper middle-class population”** (Smith, 1998, p.198).

**Definition of Crime: When it comes to crime, there are also many categories of crime to consider – Felony Assault, Robbery, Burglary, Rape, Manslaughter, Larceny, etc. On the extreme end would be violent crime, which is explored in the paper by Patrick Sharkey and Gerard Torrats-Espinosa, “The effect of violent crime on economic mobility”** (Sharkey & Torrats-Espinosa, 2017, p.22)**, highlighting that low rates of upward mobility coincides with high rates of violent crime, based on longitudinal data.   
In the dataset from New York Police Department (NYPD), offenses are classified into Felony, Misdemeanor, Violation. Similar to Sharkey and Torrats-Espinosa’s paper, this paper would consider the more severe types of crime since it, based on this literature, seems significant to other aspects of society like socio-economic inequality. Specifically, crimes categorized as a Felony would be used. This category of crimes is defined by NYPD as “a crime that is punishable by a term of imprisonment for more than one year” and include several sub-categories such as homicides, grand larceny, intoxicated driving, kidnapping, rape, robbery, theft-fraud, sex crimes, prostitution, possession of stolen property, gambling, forgery, felony assault, dangerous weapons, dangerous drugs, child abandonment, burglary, arson, abortion and other non-classified crimes.**[[1]](#footnote-1) **(CITE). By considering only the graver crimes, it is likely that underreporting would be lower since these are more conspicuous than less severe crimes classified as Misdemeanor or Violation.**

## **Methodology**

**How gentrification studies were carried out: The way that many studies on gentrification were done, as mentioned in I.G. Ellen and G. Torrats’ Paper, was by using census tracts to stand-in for neighborhoods. Using 2010’s Neighborhood Change Database and 2012-2016’s American Community Survey (ACS), they conducted their research on identifying which low-income city neighborhoods experienced gentrification (Ellen, Torrats, 2019, p. 5).**

**Similarly, a report about School Integration in Gentrifying Neighborhoods: Evidence from New York City (Mordechay, Ayscue, 2019, p.17-18) identified gentrifying neighborhoods based on data primarily from ACS. To conceptualize Gentrification with data from ACS, they defined it as a process that involves the “in-migration” of higher-SES White residents. This is also based on studies suggesting that gentrification in those specific regions is based on a disproportionate stream of higher SES White residents, who tend to be college graduates, into a neighborhood.**

## **On Brooklyn**

**Brooklyn Neighborhoods & Granularity:** Examining the level of gentrification and its impacts in Brooklyn would require us to divide Brooklyn into parts. This can happen through neighborhoods, ethnic enclaves, census tracts, census blocks, police precincts, etc. For this analysis, neighborhoods appear to be the most apt. There are 50 neighborhoods, delineated by the U.S. Census Bureau and named as “Neighborhood Tract Areas”, which is a reasonable quantity to compare one from the other and at the same time, it is plentiful enough to identify any variation present across the neighborhoods. Furthermore, since the impact of gentrification on non-gentrified areas is not easily measured or defined, and since the metrics used are calculated per capita, any increase in granularity would only generate more noise and compromise the analysis. Finally, neighborhoods also have recognizable, known names such as Brighton Beach, Madison, Bushwick North, Bedford, etc., which makes inference more intuitive and referencing more practical, as compared to a numeric ID such as Tract No. 728.

## Non-Gentrified Neighborhoods

Figure 1: Text Scatterplot of all Neighborhoods by Gentrification Score and Change in Felony Crime Rate

**There are positive and negative gentrification scores:** Based on an exploratory data analysis with the constructed Gentrification Index over the period of 2010 - 2018, about half the neighborhoods in Brooklyn have a positive gentrification score and the other, a negative gentrification score, as observed in Fig. 1. A simple cross-reference to the areas that are known to be gentrified, such as Williamsburg and Bedford, and those that aren’t, such as Midwood and Flatbush, verifies that the Gentrification Index has a degree of validity. Hence, this research can be framed to consider neighborhoods with a negative Gentrification Score to represent Non-Gentrified neighborhoods. Furthermore, a clear negative and positive divide underscores the difference between Gentrified and Non-Gentrified neighborhoods, meaning that Gentrification can be perceived as binary, instead of being evaluated on a scale.

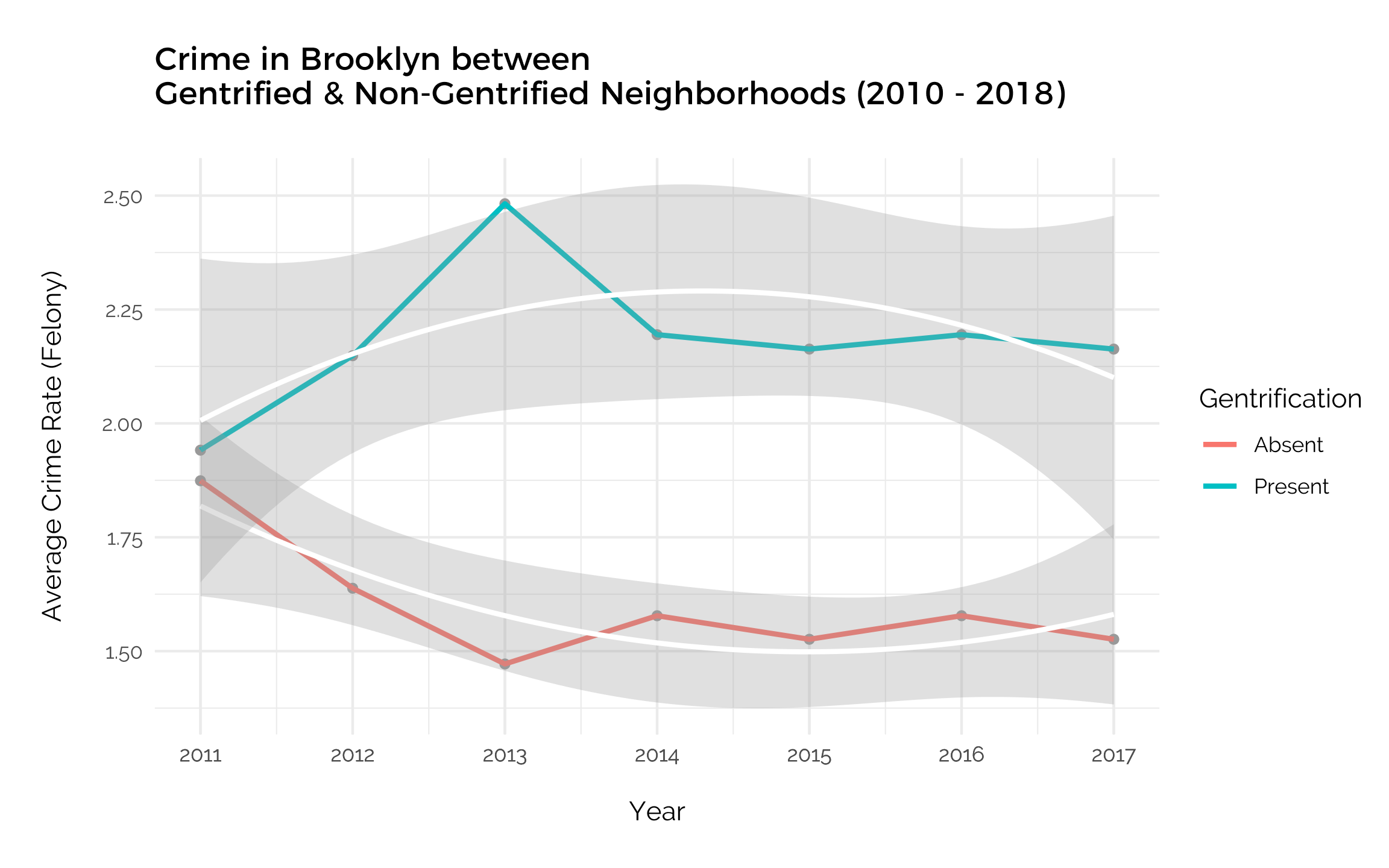


Figure 2: Trend analysis comparison of Felony Crime Rates per Capita between Gentrified and Non-Gentrified neighborhoods for each year from 2010 to 2018.

**There’s a difference in crime rates between the two:** Dividing the neighborhoods into Gentrified and Non-Gentrified is also affirmed when comparing the average yearly Crime Rates per Capita of these two categories of Neighborhoods, as shown in Fig. 2. Each trend line is distinctly different from the other, suggesting a relation between whether Neighborhoods are Gentrified or not with crime rates. While the overall trend is inconclusive and insufficient to conclude the impact of Gentrification on Crime Rates for now, this initial analysis underscores the utility and viability of classifying neighborhoods as Gentrified or Non-Gentrified.

## Hypothesis

**Both are plausible:** Whether or not Felony Crime Rates would increase in Non-Gentrified Neighborhoods as a result of Gentrification is plausible either way. Based on existing research, it could be reasoned that since Gentrification in a county like Brooklyn would displace residents of low SES (Chapple et al., 2018) into the non-gentrified areas, and that since **~~that~~ low rates of upward mobility coincides with high rates of violent crime, (Sharkey, Torrats-Espinosa, 2017), having an influx of more low-income residents in the Non-Gentrified neighborhoods would lead to a rise in felony crime rates. On the other hand, if we were to consider that Gentrification is correlated with falling crime rates in highly Gentrified neighborhoods, as supported by (CITE), it is also possible that this effect ripples to adjacent neighborhoods, thus leading to a decrease in crime rates.**

**I chose one: Based on the strong influence of Gentrification in Brooklyn, I would presume the latter to be true, in that the impact of Gentrification on lower crime rates in Gentrified neighborhoods would also register in Non-Gentrified neighborhoods.**

**What’s more important: Regardless, the possibility of either happening based on theory is likely, and the more pertinent point is to discover whether or not Gentrification actually has an impact on Crime Rates in Non-Gentrified Areas, and how this impact is exerted would be more insightful.**

# Data & Methodology

## **Datasets Imported**

**The main information required for this analysis would be demographic data of Brooklyn neighborhoods to evaluate the extent of Gentrification and the number of Felony crimes committed in the same neighborhood. This was obtained from two datasets:**

### ****A) American Community Survey (ACS) by the U.S. Census Bureau****

**About the Dataset: Initiated in 2005,** the ACS data that is administered by the U.S. Census Bureau (Bureau, n.d.) **consists of data from annual surveys conducted on approximately 3.5 million households in all U.S. States. This data is aggregated into estimates at geographical summary levels such as states, counties, tracts, etc.**

**How to access: A large quantity of data available and only a specific subset of this is required for our analysis. Specifically, demographic data on neighborhoods in Brooklyn over several years is sufficient. Thus, the most efficient way to achieve this is through the Census’ API (Application Programming Interface) accessed with packages in R, such as tidycensus, tigris, choroplethr. The function “get\_acs” allows us to obtain only what we need, by specifying the geographical scale (tract), county (Kings), state (NY), year and each variable of interest that is referenced with a unique ID number.**

**What is used from here: For this analysis, the data used to evaluate the level of gentrification are: year, census tract and demographic variables per tract. The variables of interest used to construct the Gentrification Index include:**

1. **Foreign-born residents (B05012003):   
   This variable estimates the number of residents who were born outside the U.S. in that geographic region for that year..**
2. **White residents (B03002003):   
   This variable estimates the number of White non-Hispanic or Latino residents in that geographic region for that year.**
3. **Residents without a high school diploma (B06009002):**

**This variable estimates the number of individuals ~~over~~ 25 years and over who are not High School graduates in that geographic region for that year.**

1. **Resident population (B01003001):**

**This variable~~s~~ estimates the number of residents in that geographic region for that year.**

**As mentioned in the Literature Review, there is no set definition nor set statistic for Gentrification, so the methodology adopted to define this concept for this paper will be explained in** [3.2.](#_Dependent_Variable:_Violent) **With these selected demographic variables for each Brooklyn neighborhood, it is possible to evaluate the extent of gentrification for that area.**

### ****B) Crime Statistics from the New York Public Safety Department****

**What is in it: Cases of crime and their details such as location, time of occurrence, level of offense, description of offense, classification IDs, suspect and victim descriptions, are shared through the NYC Open Data platform** (Calgary, n.d.)**. Further descriptions of variables and the dataset are included as well. This dataset is extremely large with 2GB of data and it was useful to first subset only Felony data from Brooklyn.**

**Manipulation Required: However, the main data required is not immediately available in this dataset. What is relevant for the analysis is the number of Felonies committed in each Brooklyn neighborhood per year. Such granularity of analysis is sufficient as the demographic data is only available for each year and neighborhoods are the most appropriate scope to investigate, as explained above. Unfortunately, the data available in the raw dataset only shows the location coordinates of each crime reported and are not aggregated by neighborhoods. Thus, other datasets are required to make the correspondence.**

**What this data can tell us: After the necessary data manipulation, it will be possible to compare crime statistics across the years and among the different neighborhoods in Brooklyn. In this way, the impact of gentrification can be observed in these neighborhoods that have not directly experienced gentrification using the Crime statistics over time.**

### ****C) Spatial Compatibility & Manipulation****

**The accessible Census data on Brooklyn delineates the borough into Census Tracts and the NYPD Crime Statistics indicate the specific location of crimes with XY co-ordinates. Hence, two other datasets were required to join the two primary datasets to make the data compatible with one another for subsequent analysis. The sf package in R, which processes the data more quickly than other packages and applications, was utilized for the following steps.**

**Crime: To aggregate Crime statistics into Neighborhood Tract areas, Geospatial JSON data from data.cityofnewyork.us (CITE/LINK)**(*Neighborhood Tabulation Areas*, n.d.)**~~.~~  which comprised Neighborhood Tabulation Areas (NTA), Census Tract numbers and Borough Number is needed. Using the st\_join function and setting the join argument as st\_within, each crime from the NYPD data was assigned to a neighborhood based on its XY co-ordinate. It was thus possible to calculate the number of Felonies per Neighborhood for each year and conduct spatial analysis solely on Crime statistics.**

**Census & Gentrification: ~~Following, a~~A spatial shape file of all tracts from U.S. Census Bureau with their tract numbers, labeled as GEOID, was required to merge the manipulated** **Census data with this spatial data, such that the Census data with a constructed gentrification index and other demographic variables now have spatial attributes. Spatial analysis then becomes possible on Gentrification Index scores of each of the more than 700 Census tracts in Brooklyn. Merging both: To obtain both variables in the same dataset at the same granularity, these datasets were merged by GEOID and the corresponding Gentrification scores of each tract were averaged out to apply to each Neighborhood, so it is now possible to analyze Gentrification and Crime by Neighborhood for each year between 2010 – 2018.**

## **Independent Variable: Gentrification**

**Which variables are used:** The level of gentrification in each neighborhood for each year is the independent variable in this analysis and this needs to be defined quantitatively. With references to the papers (CITE), the variables of Number of White Residents, Number of Foreign-Born Residents, Number of Residents above 25 years who did not graduate from High School per capita of the resident population in each of the 50 Brooklyn Neighborhood are used to compute the Gentrification Index score.

**Gentrification Score calculation:** Specifically, since Gentrification refers to a process, the score compares the variables from one year to a less recent year. The variables are then computed to reflect the percentage change, that is the percentage change of Number of Foreign-Born Residents between 2010 – 2018. If the number of foreign-born residents per capita in 2010 is 2.30, and in 2018 is 4.0, the percentage change would be (4.0 – 2.30) / 4.0. The same calculation is applied to the Number of White residents per capita and the Number of non-High School graduates per capita. Then, these three percentage changes are added up, although the value for Number of non-High School graduates per capita would be given a negative coefficient, since more less High School graduates correspond to an increase in Gentrification, and the other two variables have a positive correlation with Gentrification (CITE). The final value forms the Gentrification Index score.

**Application of my method – no need to be like other indexes:** Other research papers on Gentrification have calculated such Gentrification measures differently, often with more variables and weighted coefficients. This would require more datasets and more data cleaning, which does not justify the pertinence of this score. For this paper, the extent of Gentrification is not the outcome that is being evaluated; the impact on Crime Rate is the variable of interest. Furthermore, the concept of Gentrification itself, as mentioned above, is not precise nor does it bear a universal definition and thus it is impossible to have metrics that represent the extent of gentrification with full accuracy.

**My index is still well thought out**: Despite that, this paper’s relatively simplistic way to measure Gentrification takes into consideration that only the variables which describe the population are being employed, since the per capita value is measured. It means that other corresponding factors such as aggregate income, aggregate rent or number of cafés, are not included since they are not appropriate for measures per capita, particularly when families with children are being considered, or when there are more units of housing available.

As a result, a Gentrification index score based on those three variables per capita was adopted. It establishes whether a neighborhood is Gentrified or not, and that is mostly sufficient for this paper’s analysis.

## Dependent Variable: Brooklyn Felony Crime Rate

**What it is, technically:** The dependent variable to be evaluated measures the change in Felony Crime Rate per capita in each Brooklyn Neighborhood for each year between 2010 – 2018. The data primarily is sourced from the NYPD Crime Dataset, and only crimes categorized as a Felony are taken into account.

**Limitations of Crime Rate:** While the intention of using crime rate is to measure criminal activity in the neighborhood, since this is an indicator of a better or poorer quality of life, it also should be noted that ~~many~~much criminal activity can go unreported, or that reporting is more frequent in certain areas as compared to others. Even then, the level of reporting of crime itself is a variable that is informative in an appropriate context.

## Time Frame: 2010 - 2018

Apart from analyzing the relationship between these variables, this research will also involve a study of them through the period of 2010 – 2018. A time series analysis is necessary because such impact on the social well-being of residents can only be observed over a long duration. The specific time frame of the data is selected as it matches the data of the U.S. Census Bureau that is available. Although, as gentrification started since the 1990s, the effects of Gentrification on Crime are not always immediate (CITE). Furthermore, some neighborhoods in Brooklyn could have also experienced Gentrification to a high degree such that this evolution would not be reflected in this paper’s Gentrification Index, which only includes variables between 2010 – 2018. Despite that, since most of the Gentrification process in Brooklyn has mostly occurred in the last decade, this time frame is suitable for such a study (CITE).

## Control Variables

Control variables that should be considered when evaluating criminal activity in a neighborhood are as follows:

1. Overall New York City Felony Crime Rate   
   If Crime Rate has reduced in the City as a whole, it would be due to stricter NYPD control, anti-crime policies, etc. and not localized to the impact of gentrification on the neighborhood.   
   To account for this, it is only the percentage change in crime rate for each neighborhood from one year and the less recent year that is being evaluated and compared to that of other neighborhoods. The actual crime count for each year is not being used in the subsequent overall analysis.
2. Population of the Neighborhood   
   The variables obtained from the Census dataset to construct the Gentrification Index are descriptive indicators of the population which estimate the number of individuals with a certain characteristic. Hence, it is important that the overall population count of the Neighborhood is taken into consideration, as, for example, the count of White residents could increase just because the overall population in that Neighborhood has increased, while the racial demographic had remained constant.

This control variable is respected by having all variables be measured per capita, to account for variation in population size.

## Methodology

With all these variables at hand, and after the necessary data cleaning, manipulation and merging of datasets, the main analytical models can be operationalized.

### 3.6.1 Affirming Causal Relationship between Gentrification and Crime

Given that this is a Time-Series Analysis, it is possible that the relationship between Gentrification and Crime could be reciprocal, or that it is in fact Crime Rates that influence the extent of Gentrification, rather than Gentrification having an impact on Crime Rates, as is the focus of this paper. A SEM Cross-lagged Model using the R lavaan package was employed to determine the extent of this reciprocal relationship.

### 3.6.2 Determining Granularity of Research

**Neighborhood vs Census Tract:** While it has been stated that Neighborhoods are the level of granularity to review, this decision also comes after exploring the data at the Census Tract level and at the Neighborhood level in the form of choropleth maps with the Crime Rate per capita and Gentrification score variables. It was evident that the choropleth map at the level of Neighborhoods displayed a consistent spread, as opposed to that of Census Tracts. Even if a higher granularity would provide more details, these details may contain more noise, as is the case here. (Should I show the visualization comparison? Or in Appendix?)

**Gentrification as a binary indicator:** With the Gentrification Index used in this paper, it is possible to establish the extent of Gentrification. However, upon review of the data with the aid of the Text Scatterplot of the neighborhoods by Crime Rate and Gentrification score in Fig. 1, it can be observed that the magnitude of the score was not that accurate, especially compared to simply converting this into a binary predictor.

**Time Period:** While the data for this analysis covers the time period of 2010-2018, it is possible to look at the change between each year, or between 2010 and 2018. Both approaches were attempted and from a series of choropleth maps for both Crime Rates and Gentrification Scores from 2010-2011, 2011-2012, etc., compared with a choropleth map with variables considering the change between 2010 and 2018 (refer to Appendix), the latter displayed a more coherent story, particularly in neighborhoods of geographical proximity, as shown in Fig. 3. (Should I show all the choropleth maps?).   
In addition, the trend analysis carried out as shown in Fig. 2, while useful to demonstrate the significant difference in crime rates between Gentrified and Non-Gentrified Neighborhoods, reveals an inconsistent and inexplicable trend in Crime Rates for each year. Evidently, changes between each year are too small to be evaluated, and the overall change in variables are more telling.

### 3.6.3 Identify Non-Gentrified Neighborhoods to Analyze

With a choropleth map analysis of all Neighborhoods according to their Gentrification score from 2010 – 2018, it was evident that the Non-Gentrified Neighborhoods were clustered in the same geographical area.

There were two other outliers with negative gentrification sores, Greenpoint and Brooklyn Heights-Cobble Hill, that can be identified simply by their geographical distance from the other Non-Gentrified Neighborhoods, and by referencing other research that have shown these two Neighborhoods to have gentrified extensively in the years before 2010.

As such, these two outliers were eliminated from the analysis of Non-Gentrified Neighborhoods and only the remaining ones were used for further spatial analysis.

### 3.6.4 Spatial Autocorrelation Statistical Models

Finally, a core part of the analysis entails spatial analysis which uses the variables’ geographic properties. Checking for spatial autocorrelation is helpful since it highlights that each neighborhood is not independent of the other and that the relative outcome of two neighborhoods is related to their distance. For this paper, spatial autocorrelation ~~is in fact~~ assumed, in that the proximity of Gentrified Neighborhoods to Non-Gentrified Neighborhoods would influence Crime Rates in the latter.

To detect autocorrelation, the Moran’s I parametric test using Queen contiguity gives a positive or negative spatial autocorrelation score and a p-value for the level of autocorrelation, by weighting the correlation with inverse distances.

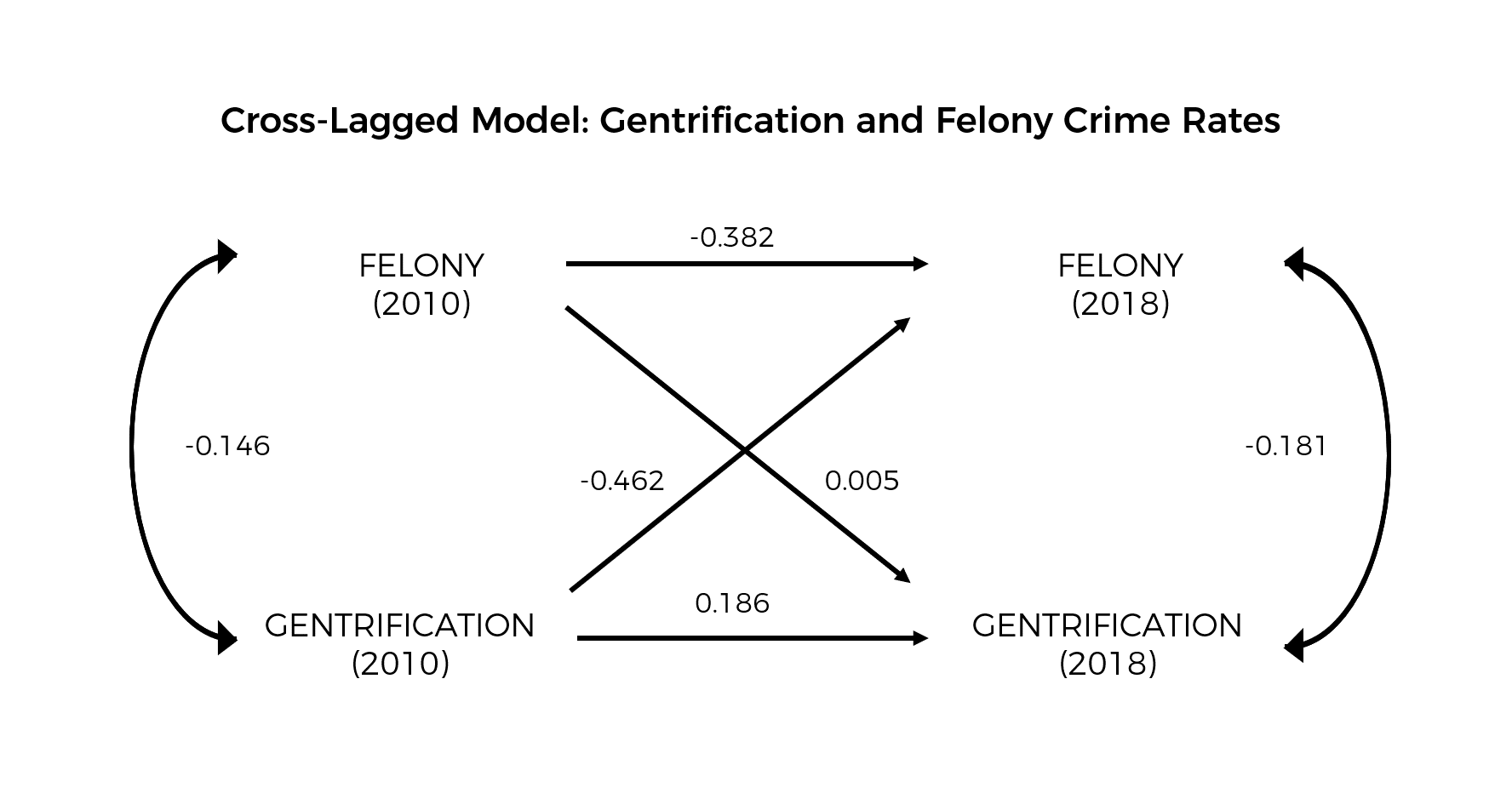
Following, Local Indicators of Spatial Association (LISA) becomes useful in providing a statistic for each location with an assessment of significance by establishing a proportional relationship between the sum of the local statistics and a corresponding global statistic~~s~~. Mapping out any LISA clusters present in the data would provide further insight to spatial autocorrelation, and how Crime Rates are affected due to geographical location.

# Results

## 4.1 Cross-Lagged Model Analysis

To detail the relationship between Gentrification and change in Felony Crime Rates for each neighborhood, the ~~results~~ Cross-Lagged model produced the regression and covariance estimates as shown in this diagram of Fig 3.

Figure 3: Cross-Lagged Model analysis for Gentrification and Change in Felony Crime Rates. The Beta values of -0.462 for Gentrification on Crime and .0005 for Crime on Gentrification indicates that the former has a stronger effect.



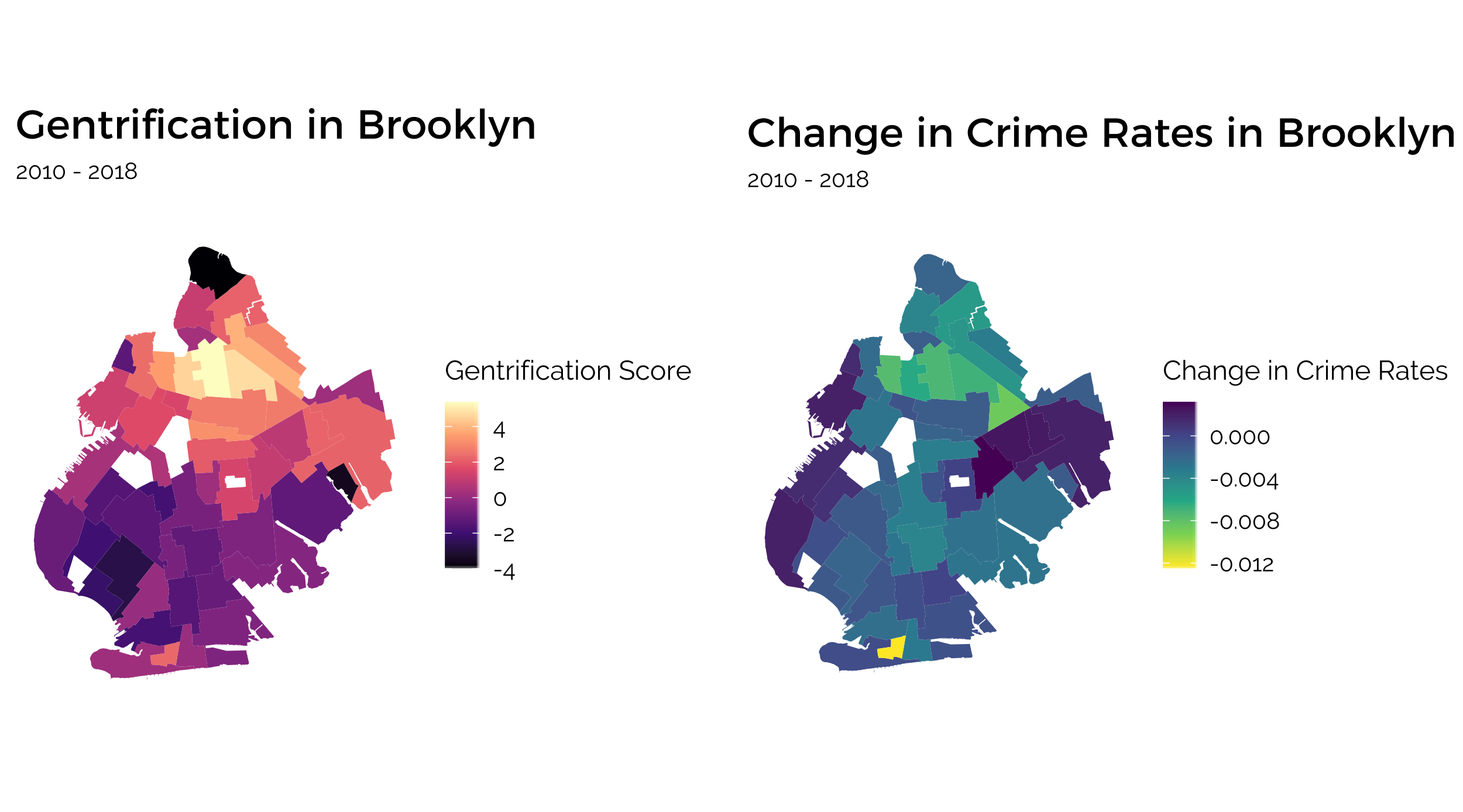
The results show that there is evidence of mutual and reciprocal causation, though the effect of Gentrification on Change in Felony Crime Rates (Beta = 0.462) is 92.4 times stronger than the opposite effect (Beta = 0.005)[[2]](#footnote-2). Hence, this validates that the impact of Gentrification on Crime Rates is relevant for study.

## 4.2 Choropleth Maps

### 4.2.1 All of Brooklyn

Visualizing each variable, Gentrification and Change in Felony Crime Rate, through choropleth maps were helpful in observing spatial correlations between each neighborhood and to affirm the granularity for the analysis. In Fig. 4, the choropleth maps for both variables are shown – the extent of Gentrification is represented in the magma pink palette, where yellow represents a high Gentrification score and dark purple a low Gentrification score. Change in Felony Crime Rates in the viridis turquoise palette indicate higher values in dark blue and lower values in yellow.

Figure 4: Choropleth maps of Gentrification scores and Change in Crime Rate across Neighborhoods in Brooklyn. The maps indicate a spatial relationship for each variable, and also a similarity in spread between both maps.



As can be observed in Fig. 4, there is a general spatial pattern to the distribution, where more gentrified neighborhoods are clustered in the upper-middle portion of Brooklyn, whereas neighborhoods with negative Gentrification scores are located in the lower-middle portion of Brooklyn. Furthermore, the difference between one neighborhood to an adjacent neighborhood is not pronounced, and the gradual spread of colors show that Gentrification levels in neighborhoods are correlated with their adjacent neighborhoods.

This is especially interesting as the same pattern is found in the choropleth map for Change in Crime Rates. Neighborhoods with lower Change in Crime Rates are likewise located in the upper-middle region of Brooklyn, similar to those neighborhoods with higher Gentrification scores. The lower-middle part of Brooklyn shows a higher Change in Crime Rates compared to Gentrified Neighborhoods, although it should be noted that the bulk of these values are still negative, thereby indicating a fall in Crime Rates.

However, the colors on this choropleth map in the lower-middle portion, where Non-Gentrified Neighborhoods are very similar, making it difficult to distinguish the variation in this area. This is mostly because the variation in these neighborhoods is visualized on the same scale as Gentrified Neighborhoods, thus making the difference less pronounced here. Thus, choropleth maps and other analyses focusing on Non-Gentrified Neighborhoods could provide more insight relevant to this paper.

### 4.2.2. Only Non-Gentrified Neighborhoods

Before mapping out the distribution of variables in maps, taking a look at the distribution of its values in this data subset could be useful.

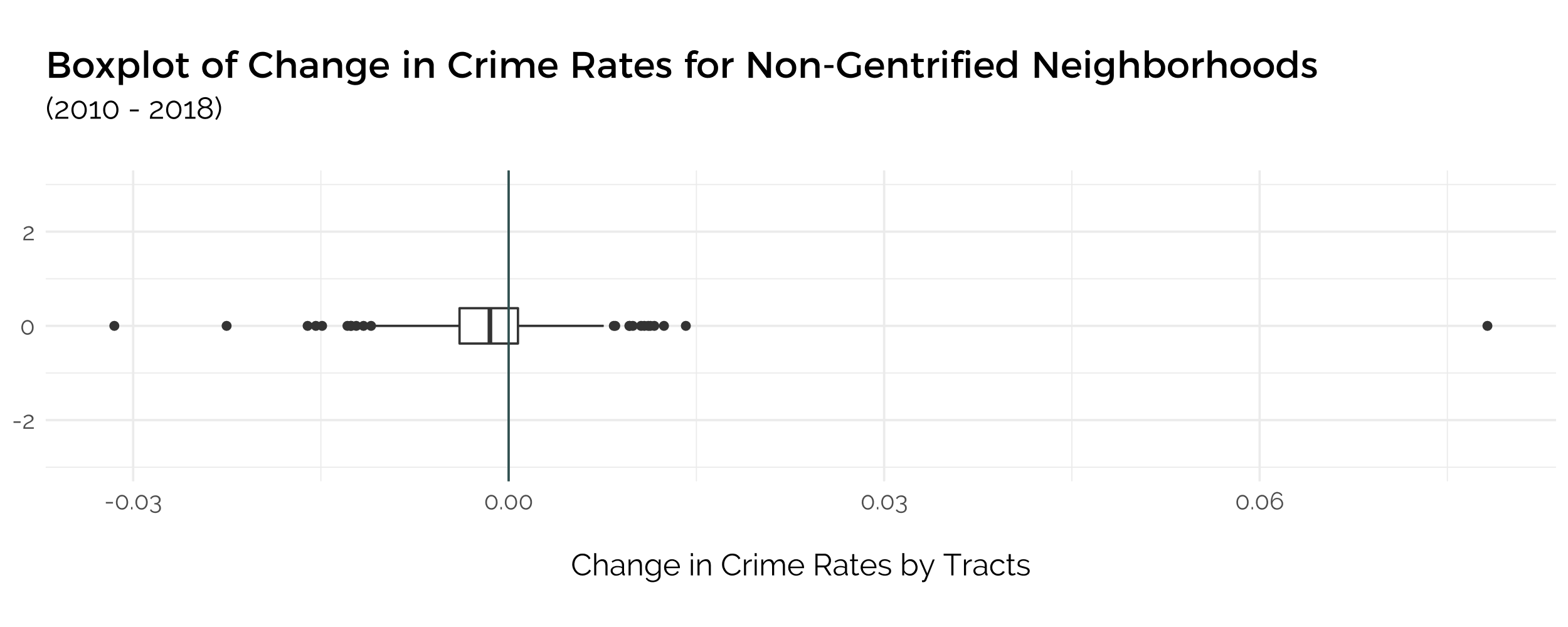


Figure 5.1: Boxplot digram of the Change in Crime Rates for Non-Gentrified Neighborhoods, by Census Tract

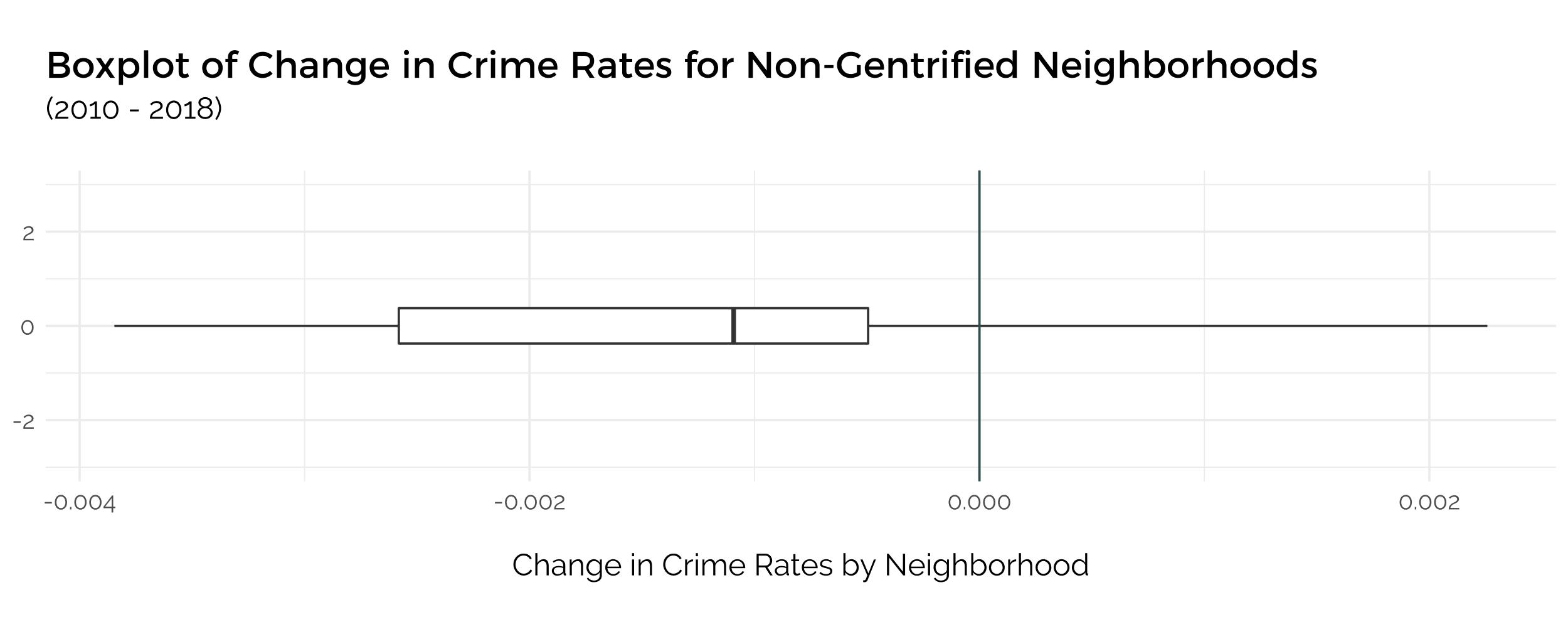


Figure 5.2: Boxplot digram of the Change in Crime Rates for Non-Gentrified Neighborhoods, by Neighborhood

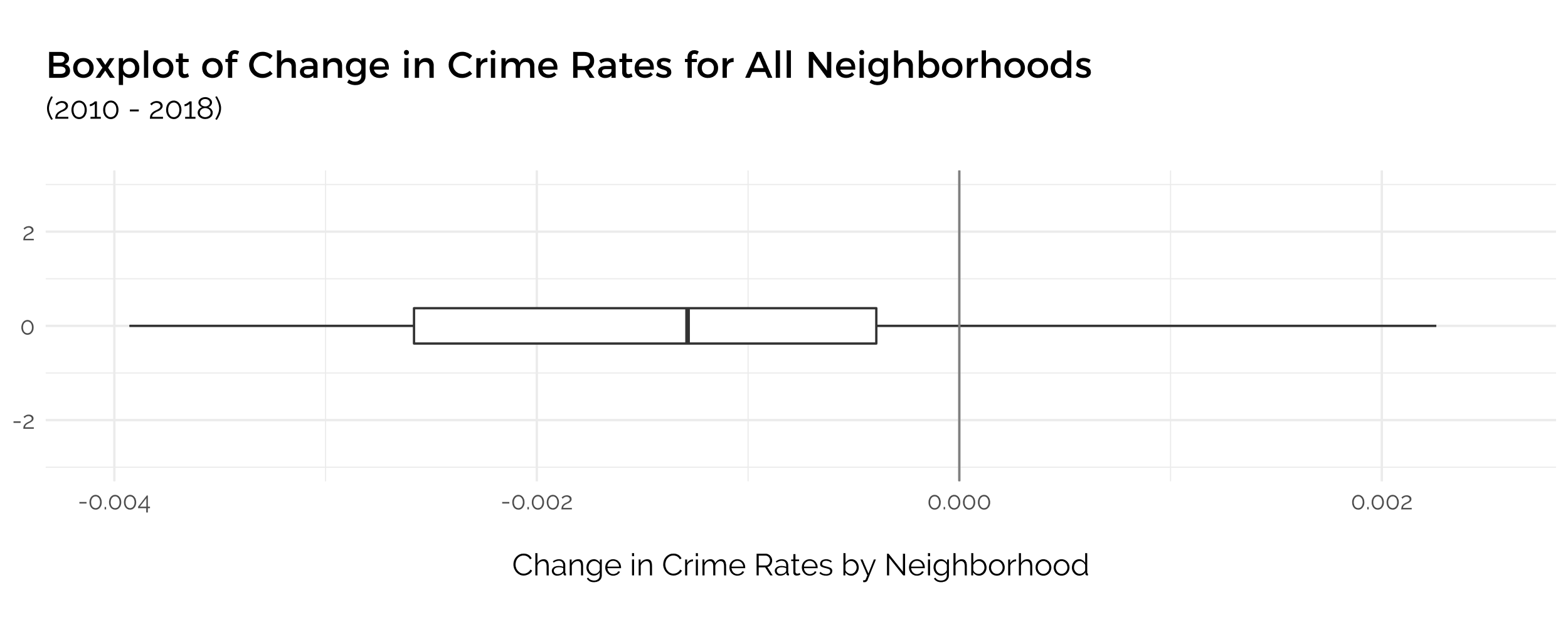


Figure 5.3: Boxplot digram of the Change in Crime Rates for All Neighborhoods, by Neighborhoods

First, the two boxplots of the Change in Crime Rate variable of Non-Gentrified neighborhoods are plotted in two different granularities – by Neighborhood and by Census Tracts, to see if there would be any significant difference. Both graphs indicate a mean below 0, suggesting that **in Non-Gentrified Neighborhoods, there is an average decrease in Crime Rates.** In addition, a comparison of both boxplots shows that plotting by Census Tracts produces a lot more outliers than by Neighborhood, thereby affirming the decision to carry out most of the analysis based on Neighborhoods. Furthermore, the boxplot by Neighborhood also indicates that both upper and lower quartiles fall under 0. It is also important to compare this statistic of Change in Crime Rates to that for all neighborhoods (Fig 5.3), where the mean is also negative, and slightly lower than that for Non-Gentrified Neighborhoods.

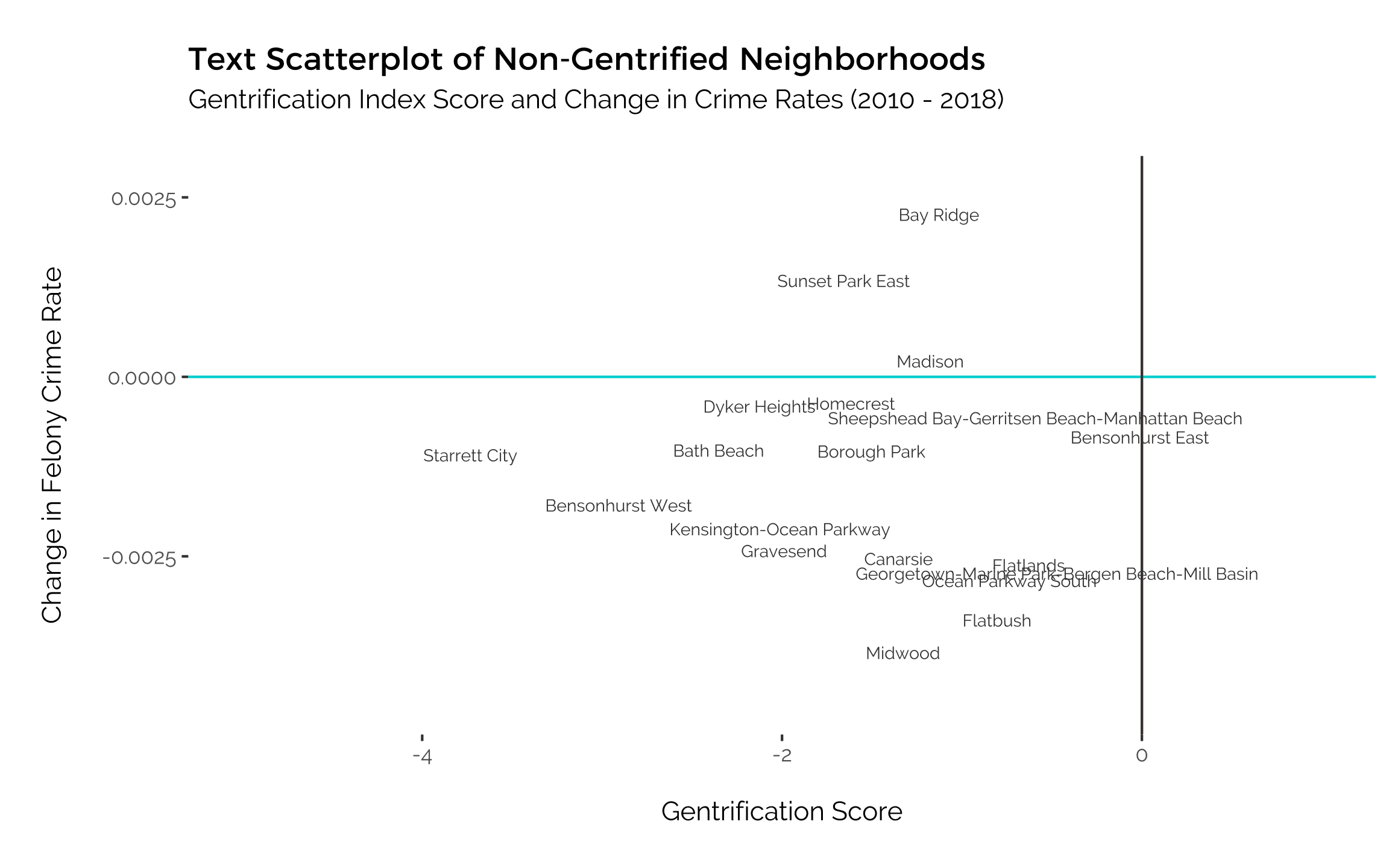


Figure 6: Text scatterplot of only Non-Gentrified Neighborhoods. Most show a decrease in Change in Crime Rates

The same Text scatterplot from above was plotted, though this time with only Non-Gentrified Neighborhoods. This makes it possible to identify exactly which Neighborhoods are anomalies, and to affirm that most of the Non-Gentrified Neighborhoods have experienced a fall in Felony Crime Rates, with the exception of Bay Ridge, Sunset Park East and Madison.

Une image contenant texte

Description générée automatiquement

Figure 7: Linear Regression of Gentrification Score against Change in Crime Rate for Non-Gentrified Neighborhoods

Even though the main interest of this paper is to find out if there is an increase or decrease in Change in Crime Rates in Non-Gentrified Neighborhoods, of which there is a decrease, further detailed exploration into this data could draw some additional insight.

To further detail this relationship, an Ordinary Least Squares Linear Regression was carried out on both variables for Non-Gentrified Neighborhoods. The estimates indicate that for every increase in 1 point of Gentrification Score, there is a -0.000204 decrease in Change in Crime Rates. This translates to meaning that a less negative gentrification score among Non-Gentrified Neighborhoods correlates with lower Crime Rates.

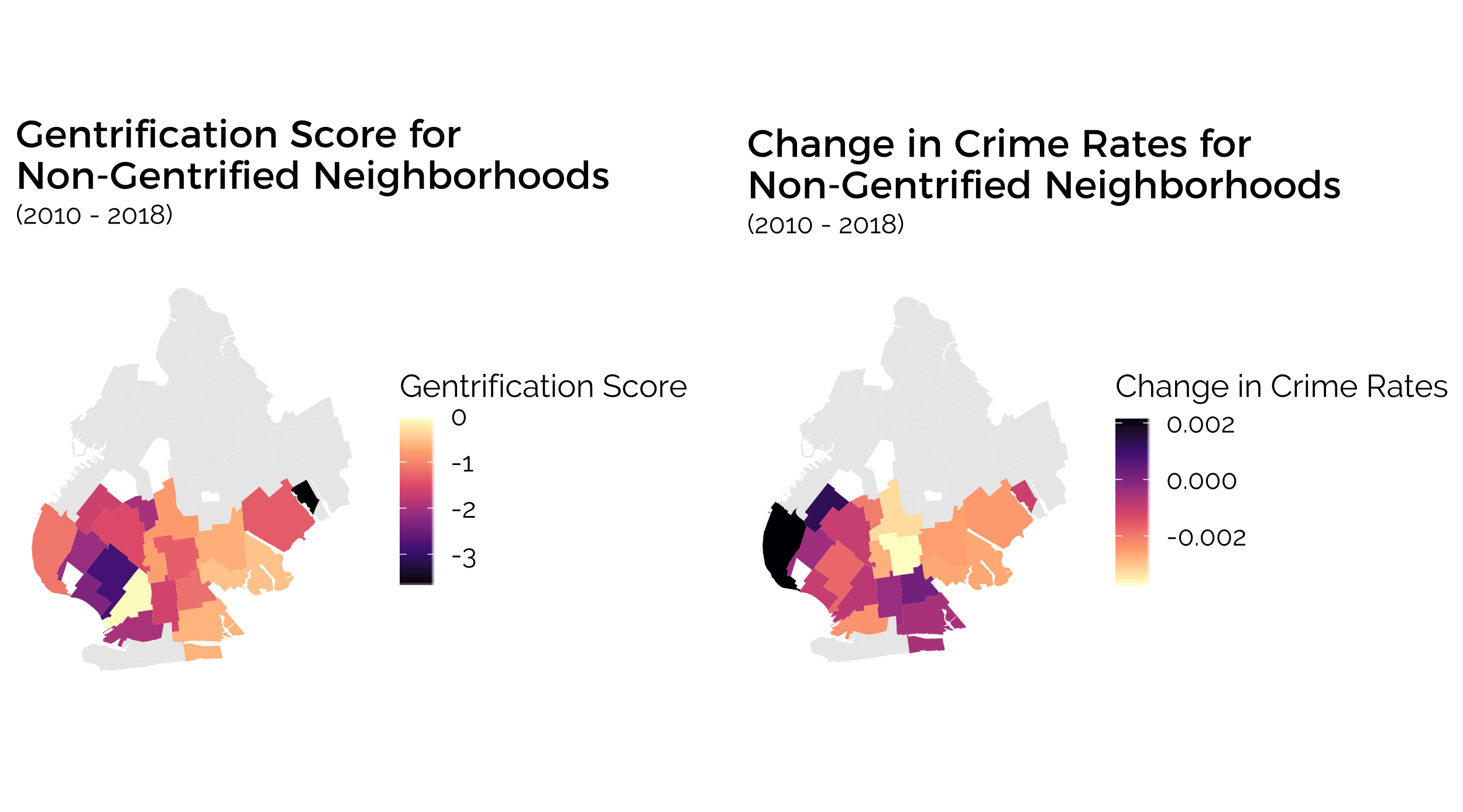


Figure 8: Choropleth Maps comparing the extent of Gentrification and Change in Crime Rates fro Non-Gentrified Neighborhoods

Viewing the distribution of this relationship on choropleth maps would further highlight any spatial relationship present. By only considering Non-Gentrified Neighborhoods, plotting a choropleth map for the Change in Felony Crime Rates indicate the variation between neighborhoods more clearly. Choropleth maps for both Gentrification Scores and Change in Crime Rates in these Non-Gentrified Neighborhoods are displayed together in Fig (?) and with a similar color palette in order to examine the relationship of the spread of both variables.

There are three main observations: first, there is a degree of similarity in how the lighter colors are concentrated in the upper, middle and right areas, while the darker colors are found in the bottom-and left neighborhoods. This implies **a relationship of lower Gentrification and a lower decrease in Crime Rates,** similar to the results from the OLS regression.

Second, the Change in Crime Rates also replicate the trends observed earlier, in that most neighborhoods have similar rates as their neighbors, and where almost all the neighborhoods experience a drop in crime rates. There is also a gradual trend where the neighborhood in the middle has the lowest Change in Crime Rate and this increases gradually around it. This suggests that **the neighborhoods not bordering Gentrified areas have a lower Change in Crime Rate.**

However, the third remark would be that the area on the left, marked by the two neighborhoods in black on the map of Change in Crime Rates, strongly opposes this trend. It has a mild negative gentrification close to the median, but **they have exceedingly high change in Crime Rates by comparison.** The exceptions are the neighborhoods on the left-hand side, which spatial autocorrelation models can confirm.

## 4.3 Spatial Autocorrelation Analysis

In order to understand the relationship between the values of Change in Crime Rate in each Neighborhood and the average value in bordering Neighborhoods, the Moran’s Scatterplot is ideal.

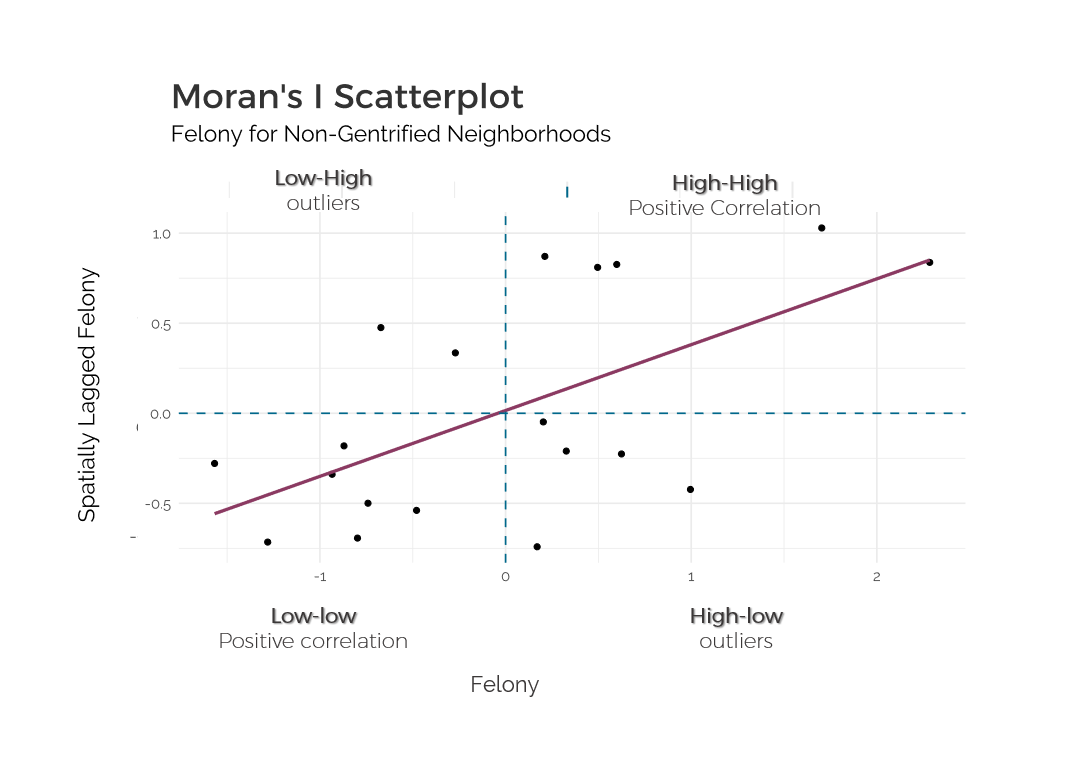


Figure 9: Moran's I Scatterplot analysis for Felony Crime Rates inNon-Gentrified Neighborhoods

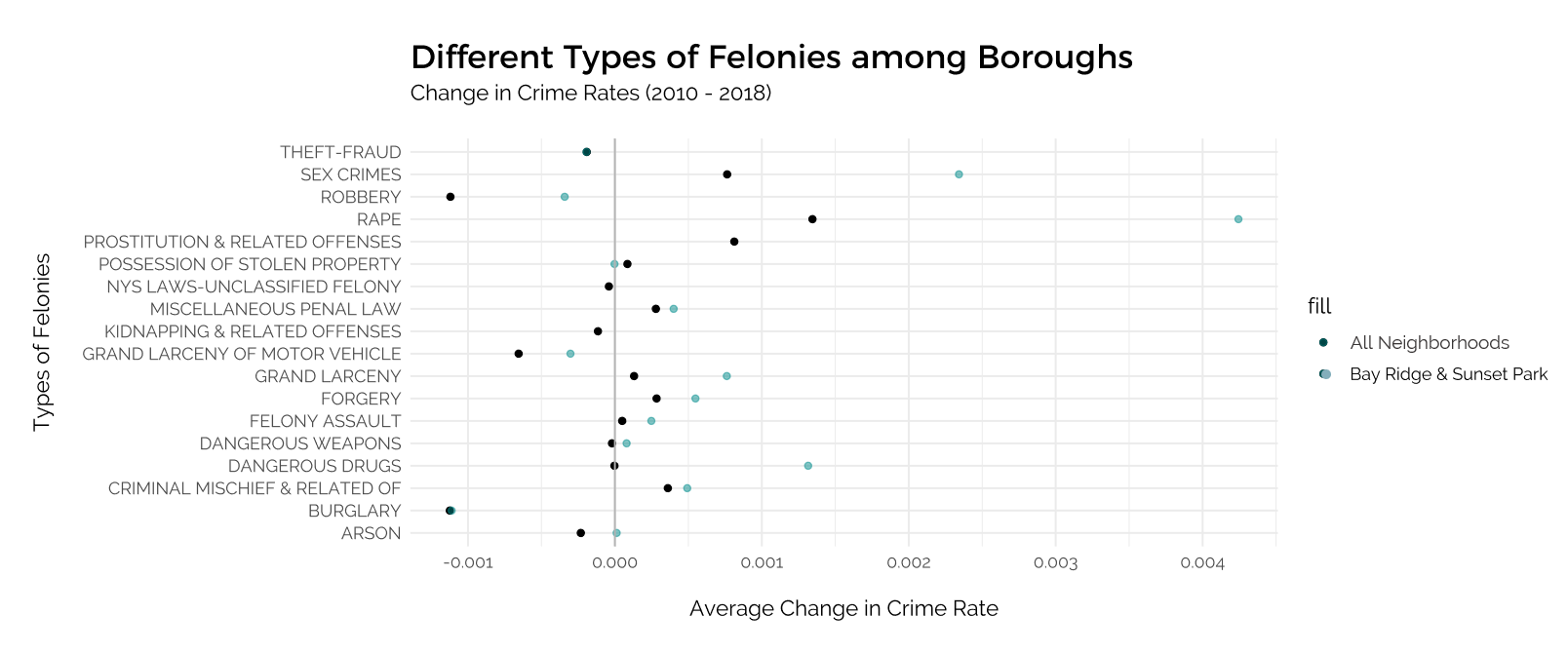
The overall slope of the regression line estimating the global Moran’s I indicates that most of the Neighborhoods with high Change in Crime Rate values are surrounded by those with high values as well, and those with low values are surrounded by those with lower values. To observe whether there are significant high-high, low-low, high-low or low-low values, it is best viewed on a Local Indicators of Spatial Association (LISA) Cluster Map.

Une image contenant carte

Description générée automatiquement

Figure 10: LISA Clusters for Change in Felony Crime Rates of Non-Gentrified Neighborhoods. Four neighborhoods of low-low significance and high-high significance are identified.

The LISA Clusters show a high-high correlation in the neighborhoods on the left, identified as Bay Ridge and Sunset Park East, indicating that both have high values of Change in Crime Rates as compared to their bordering neighborhoods (Fig. 10). These two neighborhoods were also the outliers identified above that opposed the trend of how less negative gentrification scores correspond to lower crime rates. To further understand these exceptions, a deeper analysis into the Change in Crime Rate across the different types of Felonies could prove useful.



While a Felony is already a category of crime that makes reference to the severity of the offence, there still are a range of felonies. In the distribution plot above, the average Change in Crime Rates for the two outliers of Bay Ridge & Sunset Park are compared with the average values for all neighborhoods. The increase in Change in Crime Rates for the outliers are especially high for sex crimes, rape, robbery, grand larceny, dangerous drugs. This informs of how exactly Crime in those two neighborhoods are evolving with respect to all the other neighborhoods.

# Discussion

## Significance of Results for Gentrification on Crime

From the results obtained and the subsequent analysis, several conclusions about Gentrification in Brooklyn can be drawn.

Firstly, **Gentrification has an impact on Crime throughout time**, as proven in the cross-lagged model analysis (Fig. 3). This relationship has been discussed in research (CITE), where some hypothesize that it is Crime which determines the extent of Gentrification, but the results from this analysis proves contrary. Establishing this then provides more credibility and significance for the subsequent findings.

Following, the spatial relationships for each of the variables were explored through choropleth maps, where **a presence of spatial dependence can be observed** (Fig. 4). Certain areas, in the upper portion of Brooklyn, had significantly lower Gentrification and Change in Crime Rates. These include neighborhoods of Bedford, Stuyvesant Heights, Clinton Hill, and Bushwick South. It corresponds to existing reports about how these areas have recently undergone high levels of gentrification (CITE), especially compared to neighborhoods on the top and left, such as Williamsburg, DUMBO, Brooklyn Heights, or Greenpoint. On the other hand, the more southern neighborhoods of Brooklyn have distinctly lower and actually negative Gentrification scores, indicating the lack of Gentrification, as measured by the Gentrification Index used in this paper. The neighborhoods with the most negative Gentrification scores include Bensonhurst West, Bath Beach, and Dyker Heights, where the surrounding neighborhoods also reflect similar low scores of Gentrification. Change in Felony Crime Rates are slightly higher than it is in the more northern parts of Brooklyn, though it is mostly negative.

A closer look into only Non-Gentrified Neighborhoods, which comprise almost half of the Brooklyn neighborhoods, provides a more comprehensive understanding of these areas. It is in these areas that data and research is relatively lacking but having access to the city’s datasets makes it possible to learn more about these areas by using the few datasets available here in combination with data from surrounding areas.

**Non-spatial variable distribution:** Viewing the distribution of the Change in Crime Rate values through boxplot diagrams for Non-Gentrified Neighborhoods and all Neighborhoods confirms that both means are negative, though it is lower for all Neighborhoods than it is for Non-Gentrified Neighborhoods. This indicates that there is a decrease in Change in Crime Rate for Non-Gentrified Neighborhoods, but to a lower extent than when considering this statistic for all of Brooklyn. The decrease thus adheres to the overall trend and comparing it to the overall trend puts ~~it~~ in perspective that the decrease is not an anomaly or particularly significant. Additionally, Fig. 6 confirms the overall decreasing Felony Crime Rates in Non-Gentrified Neighborhoods, with only Bay Ridge, Sunset Park East and Madison experiencing increasing rates, although Madison’s is relatively closer to zero.

**Qualitative relationship:** By exploring the relationship between both Gentrification and Crime variables in Non-Gentrified Neighborhoods through a linear regression, it is observed that a lower decrease in Gentrification Score (i.e. closer to being considered as undergoing Gentrification) corresponds to a higher decrease in Felony Crime Rates (i.e. Crime is decreasing at a greater rate than others). This extends our conclusion that while neighborhoods in non-gentrified areas of Brooklyn have experienced, on average, decreasing Felony Crime Rates, those that are closer to being gentrified experience even greater decreasing crime rates.

**Spatial Qualitative**: This trend is also reflected upon comparison of two choropleth maps of Gentrification and Change in Crime Rates for only Non-Gentrified Neighborhoods (Fig. 8), where the colors are lighter in the middle, and it gradually becomes darker towards the exterior parts. The exterior parts are also the ones closer to Gentrified Neighborhoods, suggesting that Gentrification could lead to an increasing decrease in Crime Rates in the bordering Neighborhoods that have not undergone Gentrification. Two outliers contradict this trend – Bay Ridge and Sunset Park West, which have a Gentrification score closer to zero but a higher increase in Crime Rate. Referencing the Text Scatterplot above for only Non-Gentrified areas, these two neighborhoods are the ones with highest Change in Crime Rates, out of the only three neighborhoods with a positive rate.

**Spatial Outliers:** Upon checking for spatial outliers, the Moran’s I Scatterplot indicates that there are insignificant low-high and high-low values and the global trend still holds a positive correlation between Spatially Lagged Change in Felony Crime Rate values and the original values (Fig. 9). However, the LISA Cluster Map identifies neighborhoods that have exceptionally high-high and low-low relationships. Unsurprisingly, Bay Ridge and Sunset Park West are once again identified as an outlier, having a significant high-high relationship.

## Understanding Outliers

With such outstanding outliers, it seemed important to further discern the reason behind these, if any; they could be statistical outliers, or this could provide more revelations.

Referencing existing literature and news reports (CITE), it appears that these two neighborhoods, Bay Ridge and Sunset Park West, are beginning to experience Gentrification. There are new developments and buildings being constructed, but the population has not yet changed too much, thereby explaining the still negative Gentrification Index Score. Still, current residents are beginning to get displaced (CITE).

This paper (CITE) also surmises that in areas where Gentrification is starting to occur, and where there is an amalgamation of original and new residents of higher SES, crime rates increase. This is explained by having more crimes reported, higher vigilance of police officers, and having more opportunities for criminal activity.

To further explore this reasoning, the Change in Crime Rates for the different types of Felonies for these two neighborhoods were compared with that of all neighborhoods, to investigate whether there is a significant increase in certain types of crime. The increase of sex crimes, rape, grand larceny, dangerous drugs, and robbery were found to be significantly higher than the overall Brooklyn average. (MORE REFERENCE NEEDED PERHAPS)

## Limitations of Research

To be continued   
- Defining Gentrification

- Time Period limit

- Other external factors

## 

# Conclusion

Overall, it is possible to understand the impact of Gentrification on less-studied Non-Gentrified Neighborhoods using crime data that is available consistently throughout New York. This highlights the importance of having such quantitative public datasets ~~to be~~ freely accessible, so that researchers may explore non-obvious relationships and provide more insight to less-researched areas in an economical way, in the context of Urban Planning.

The relationship between Gentrification and Crime in Non-Gentrified Neighborhoods in Brooklyn can be summarized as follows:

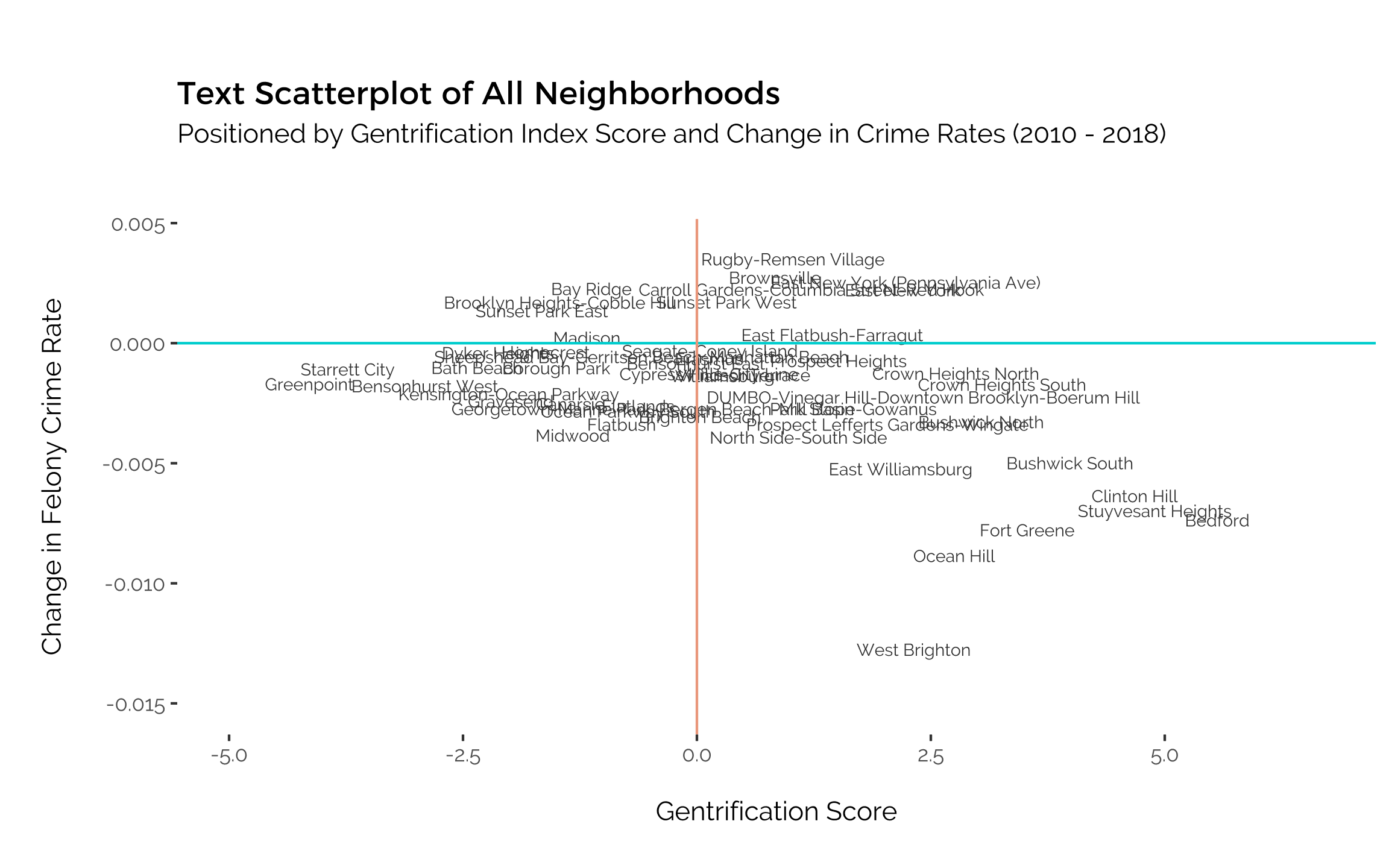
Non-gentrified neighborhoods of Brooklyn which are located mostly in the south (with the exception of Coney Island) experience an overall fall in Felony crime rates, though to a lower extent than the whole of Brooklyn. There is also a trend of neighborhoods which are less non-gentrified having a larger decrease in crime rates, and this is especially so in neighborhoods bordering Gentrified areas. However, two neighborhoods, Bay Ridge and Sunset Park West, distinctly oppose this trend. Identifying these outliers led to further research into these specific areas which suggest that in areas that are starting to gentrify and have a mixture of original and new residents, crime rate increases, likely due to both higher criminal activity but also more reporting.

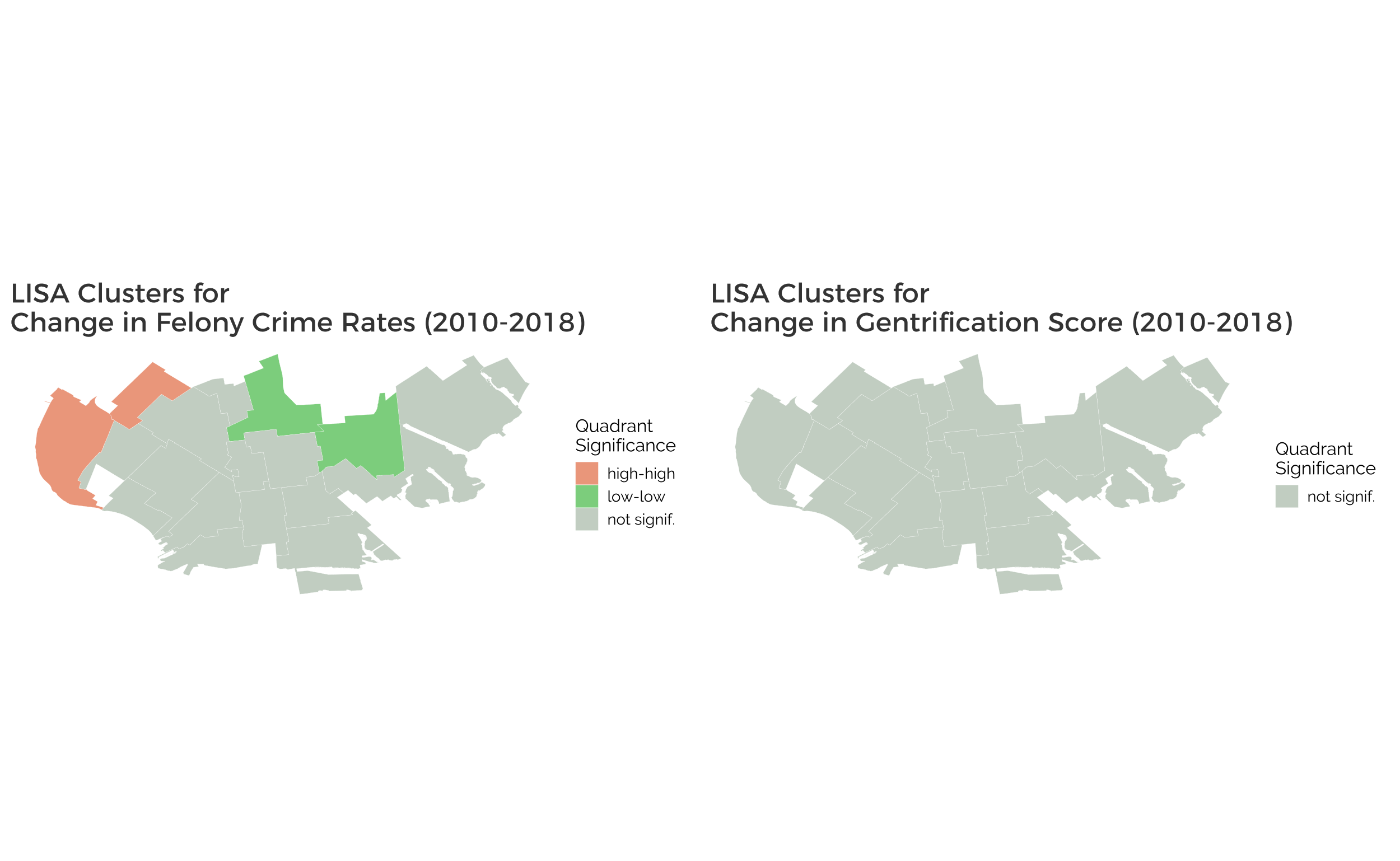
Overall, the modern definition of Gentrification is an urban process that is increasingly present throughout the world, with Brooklyn being an area that has experienced several waves of Gentrification since the 1990. This evolution is arguably one of the most significant socio-economic changes to the region, thereby reducing external variables and thus making it accurate to understand the social impact of gentrification on non-gentrified neighborhoods, not just in this borough, but perhaps worldwide.

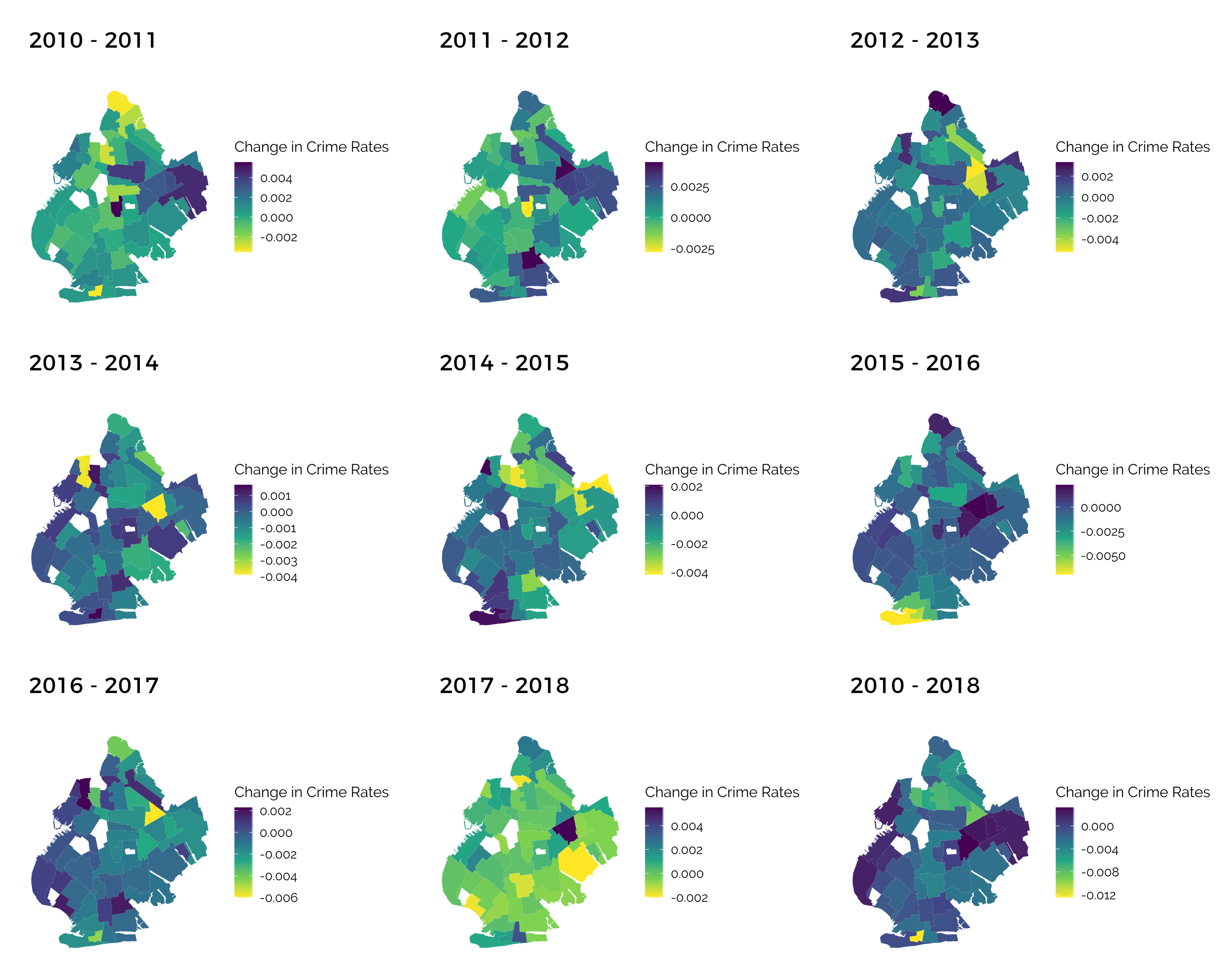
# References

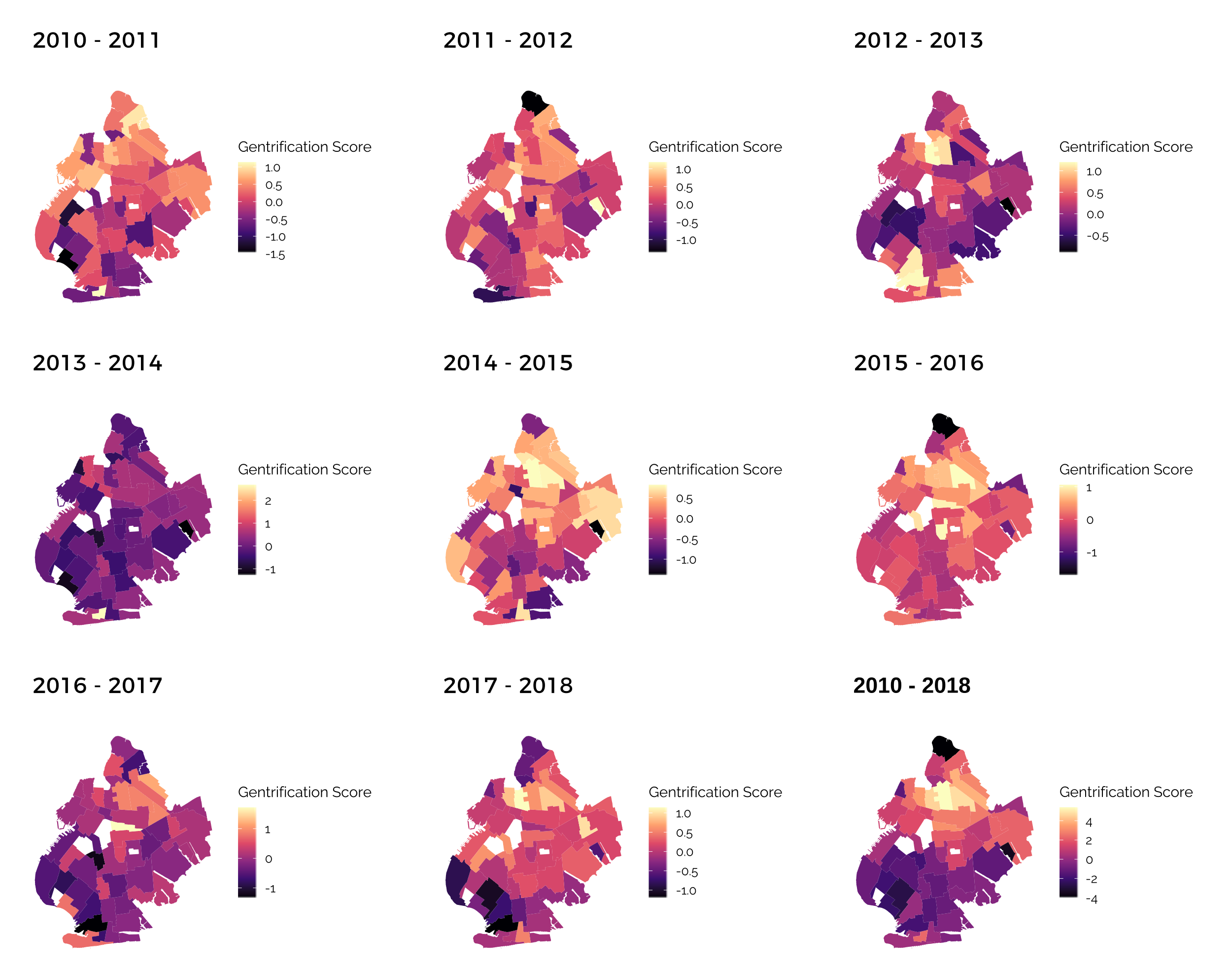
1. Ellen, I. G., & Torrats-Espinosa, G. (2019). Gentrification and Fair Housing: Does Gentrification Further Integration? *Housing Policy Debate*, *29*(5), 835–851. <https://doi.org/10.1080/10511482.2018.1524440>
2. *International Encyclopedia of Housing and Home | ScienceDirect*. (n.d.). Retrieved February 13, 2020, from <https://www.sciencedirect.com/referencework/9780080471716/international-encyclopedia-of-housing-and-home>
3. Lees, L. (2003). Super-gentrification: The Case of Brooklyn Heights, New York City. *Urban Studies*, *40*(12), 2487–2509. <https://doi.org/10.1080/0042098032000136174>
4. Mordechay, K., Ayscue, J. B., & Orfield, F. G. (2019). *School Integration in Gentrifying Neighborhoods:* 53.
5. Sharkey, P., & Torrats-Espinosa, G. (2017). The effect of violent crime on economic mobility. *Journal of Urban Economics*, *102*, 22–33. <https://doi.org/10.1016/j.jue.2017.07.001>
6. Sharkey, P., Torrats-Espinosa, G., & Takyar, D. (2017). Community and the Crime Decline: The Causal Effect of Local Nonprofits on Violent Crime. *American Sociological Review*, *82*(6), 1214–1240. <https://doi.org/10.1177/0003122417736289>

# Appendix









1. <https://www1.nyc.gov/site/nypd/services/victim-services/glossary.page> [↑](#footnote-ref-1)
2. Full statistical results from R would be included in the Appendix [↑](#footnote-ref-2)