HAROKOPIO UNIVERSITY

School of Environment, Geography and Applied Economics
Department of Geography

Postgraduate Programme “Applied Geography and Spatial Planning”
Sector: Geoinformatics

SPATIAL ANALYSIS OF CRIME IN EUROPE

Master Thesis of **Artemis Tsiopa**

Athens, 2018

HAROKOPIO UNIVERSITY

School of Environment, Geography and Applied Economics
Department of Geography

Postgraduate Programme “Applied Geography and Spatial Planning”
Sector: Geoinformatics

Three Member Committee

Stamatis Kalogirou (Supervisor)
Assistant Professor, Department of Geography, Harokopio University

George Mavrommatis
Lecturer, Department of Geography, Harokopio University

Issaak Parcharidis
Associate Professor, Department of Geography, Harokopio University

I, Tsiopa Artemis, declare responsibly that:

1. I am the owner of the intellectual rights of this original thesis, and to the best of my knowledge, my thesis does not slander any physical persons, nor does it offend the intellectual rights of third parties.
2. I accept that the LIC may, without changing the content of my thesis, make it available electronically through its Digital Library, copy it in any medium and/ or any format and hold more than one copy for maintenance and safety purposes.

"There is no society known where a more or less developed criminality is not found under different forms. No people exists whose morality is not daily infringed upon. We must therefore call crime necessary and declare that it cannot be non-existent, that the fundamental conditions of social organization, as they are understood, logically imply it."

Émile Durkheim

Acknowledgements

Upon completing my Master dissertation, I would like to thank my supervisor and Assistant Professor of the Department of Geography, Dr. Stamatis Kalogirou, for the efficient collaboration and for his guidance through every step of the dissertation. Furthermore, I would like to thank my family for their overall support and encouragement throughout my studies.

■ Contents

Περίληψη

Η εγκληματικότητα είναι ένα από τα βασικότερα προβλήματα κάθε κοινωνίας, καθώς περιορίζει το αίσθημα ασφάλειας των πολιτών. Στην συγκεκριμένη εργασία επιχειρείται η χωρική ανάλυση της εγκληματικότητας στην Ευρωπαϊκή Ένωση, σε επίπεδο περιφερειών (NUTS 2). Ειδικότερα, μέσω διαφόρων μεθόδων και δεικτών χωρικής ανάλυσης, θα μελετηθεί η συμβολή διαφόρων κοινωνικών παραγόντων (όπως επίπεδο εκπαίδευσης, ανεργία κ.α.) στην ένταση του φαινομένου. Τα στοιχεία για την εγκληματικότητα κάθε περιφέρειας, καθώς και για κάθε έναν από τους παράγοντες που θα μελετηθούν, προέρχονται από την Ευρωπαϊκή Στατιστική Αρχή (Eurostat).
Η χωρική ανάλυση της εγκληματικότητας θα αναπυχθεί σε ένα διαδραστικό περιβάλλον, μέσω της δημιουργίας μίας web εφαρμογής. Η εφαρμογή αυτή θα δημιουγηθεί με τη χρήση του πακέτου Shiny της γλώσσας προγραμματισμού ανοιχτού κώδικα R. Στην εφαρμογή θα είναι δυνατή η μελέτη χαρτών και διαγραμμάτων που αφορούν το θέμα.

Λέξεις κλειδιά: Εγκληματικότητα, χωρική ανάλυση, Ευρώπη, web maps, Shiny.

Abstract

Crime is one of the most significant problems that every society has to face, because it limits the sense of security of the civilians. This project aims to present the spatial analysis of crime in the European Union, on a regional level (NUTS 2). Specifically, a variety of methods and indexes, aid the study of the contribution of several social factors (such as education, unemployment etc.) in the intensification of the phenomenon. The data for the crime rates of each region, as well as the data for each one of the social factors that will be studied, will be retrieved from Eurostat.

The spatial analysis of crime will be developed in an interactive environment, through the creation of a web application. This application will be created with the usage of the package “Shiny” of the open source programming language R. In this application, the study of maps and charts regarding the subject will be available.

Key words: Crime, spatial analysis, Europe, web maps, Shiny.

List of Figures

List of Tables

1. Introduction

Crime is one of the major social problems that every kind of society has to face. Its gravity is attributed to the thousands of people's lives it affects each year. The term "crime" refers to any act of violence or deception that is made in pursuit of personal interests (Siegmunt, 2016). Nevertheless, this definition does not include the legal strand of crime, which is pivotal. A different definition of the term suggests that crime is any act or negligence that is prohibited by public and criminal law for the insurance of public safety. Such act or omission is committed without justification and is punishable by the judicial system of each society (Malik, 2016). The severity of the problem of crime lies in the generation of a considerable amount of fear that it produces within the communities. This fear follows each criminal act and can result in the restriction of citizens' freedom of movement and, consequently, in the prevention of their full participation in these communities' activities (National Crime Prevention Council, 2003).

According to Arnot and Usborne (2001), what is considered crime is a social construction and has changed tremendously over the years. To give an illustration, women in the Dutch countryside, during the 19th century, never considered that calling irregular midwives to help them in childbirth was an illicit act. Nevertheless, when the doctors of the era decided that this was a field that they should supervise, but not practise due to low financial rewards, they attempted to outlaw irregular midwifery as a professional crime (Arnot and Usborne, 2011). Still, in different places of the world, different activities might be considered criminal and illicit. For example, in Saudi Arabia it is still considered illegal for a woman to drive a car (Rajkhan, 2014) and in Singapore it is illegal to sell and consume non-medical chewing gum (Arnold, 2006).

The study of crime, criminology, arose from the need of every society to find a solution to this pressing concern (Kappeler and Potter, 2017). Criminology tries to explain both crime and criminal behaviour (Wortley and Mazerolle, 2008), because the feasibility of future crimes' prediction can serve as an extremely valuable source of knowledge for the police authorities (Groff and La Vigne, 2002). The most significant crime feature (and the most recently acknowledged one) that aids its study, is its geographical quality. Every crime that occurs, happens at a specific geographical location. Besides that, every criminal, also, comes from a geographical location. This location may be the same with the one where the crime was committed or in a close distance to it. Consequently, it should be pointed out that "place" plays a tremendously significant role in

understanding and reducing crime (Chainey and Ratcliffe, 2005). Place can be an important crime factor, either by influencing and shaping criminal behaviour, or by attracting to a specific area people with the said behaviour (Almeida et al., 2003). According to Wortley and Mazerolle (2008, p.78), "Each element in the criminal event has a historical trajectory shaped by past experience and future intention, by the routine activities and rhythms of life and by the constraints of the environment. Patterns within these complexities (considered over many criminal events) should point us towards understandings of crime as a whole". The understanding of crime as a whole, with its geographic dimension included, can lead to the development of more accurate and effective strategies to combat it.

New technological advances have contributed immensely to the study of crime. One of these advances is the development of spatial analysis, which is closely related to the Geographic Information Systems (GIS) (also one of these advances) and Geoinformatics. Spatial analysis is a field that is being used for the understanding of various social phenomena that might occur in any place. For the development of each method and technique of spatial analysis (that are based on the sciences of mathematics and statistics), coding in a suitable programming language is important. The analysis and the export of the results can be carried out in the environment of different software (such as R). Spatial analysis, as opposed to classic statistics, takes into consideration the location of each observation of the studied variable. This fact is essential for the extrapolation of more and more efficient conclusions (Kalogirou, 2015). The data that are being used during spatial analysis are spatial data. This means that they contain information about the location where every studied phenomenon occurs. The main principle of spatial analysis is Tobler's first law of Geography, according to which "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970, p.236).

The aim of this dissertation, is the spatial analysis of crime in the European Union, on a regional level. The features of these regions are studied thoroughly, in order to examine whether they affect criminal activity in negative or positive ways. In other words, the objective of this analysis is to determine which factors that arise from crime literature, affect the crime rates in the EU regions. In order to achieve this goal, different methods of spatial analysis are being used. These methods include Regression and Geographically Weighted Regression, as well as the calculation of several indexes that are essential for spatial analysis, such as the autocorrelation and the inequality index. The data for the regional crime rates, as well as the data for each one of the social factors that will be studied have been retrieved from Eurostat. The software environment that is

being used for the spatial analysis, is the open source programming environment RStudio, which is based on the programming language R. The code that has been written accomplishes both the implementation of the spatial analysis and the presentation of its results in an interactive environment, which is created with the usage of the package “Shiny”, which develops web applications.

The following chapter of this dissertation examines the theoretical framework of crime. Specifically, the second chapter presents the development of the crime research over time (1800s to present), the crime theories that have been elaborated and the crime trends of the latest years (1980-2016) in different geographic levels. The third chapter focuses on the literature review of papers that have a similar research goal and approach to this study. The fourth chapter presents the spatial data analyzed and the following chapter presents the methodology applied. The results of the analysis are presented and discussed in the sixth chapter of this dissertation. In the final chapter, some conclusion are being drawn that highlight the findings of the analysis.

2. Theoretical Framework of Crime

2.1. Crime Study

2.1.1. Sociological Approach

Criminal activity has, indisputably, always been one of the most worrisome phenomena of every society. The study of crime has, traditionally, been the domain of other sciences, such as sociology and psychology (Chainey and Ratliffe, 2005). Sociological criminology has explained crime based on an individual's social environment and, therefore, social relationships. Psychological criminology has explained crime by studying the behavioral tendencies, which mainly develop in childhood from both social and biological factors (Entner Wright et al., 2001). The study of the correlation between crime rates and the spatial characteristics of an area began in the early 19th century. The first scientist to present the spatial aspect of crime was Quetelet, a Belgian mathematician, who applied a statistical method in crime observation. In 1835 he published the book "Essay on Social Physics: Man and the development of his faculties", in which he noted that crime rates seem to be stable every year. One of his theories regarding criminology suggests than violent crimes occur more often in southern areas and in hotter temperatures than in northern and colder areas. At the same time, Guerry, a French lawyer and amateur statistician, published a book titled "Essay on moral statistics of France". In this book he analyzed statistical data of juridical decisions in France, which he mapped. He was the first scientist to map crimes and study their correlation to demographic, educational and climatological factors (Bakirli, 2005). The two of them together are considered the founders of the French-Belgian Cartographic School of criminology (Kobogianni, 2012) and they suggested that crime varies across different geographical regions. Overall, they not only produced maps to visualize crime and associated it to poverty and education, but, also, they set the foundations for the shift of approaches regarding crime (Bakirli, 2005).

During the early 20th century, other researchers, this time associated with the University of Chicago, studied how other social characteristics affected crime. The main representatives of this approach were Shaw, McKay, Burgess, Park and Thrashers, who are, also, considered the founders of the Chicago School of sociology. The Chicago School brought much attention to the topic and demonstrated the importance of geography in understanding and studying crime (Chainey and Ratcliffe, 2005). Their main belief was that people are social constructions and their behavior

depends on their social environment (Kobogianni, 2012). Specifically, Shaw and McKay, presented an "ecological" approach and they suggested that social disorganization leads to higher crime rates (Goldsmith et al., 2000; Bakirli, 2005; Kobogianni, 2012). Their work on juvenile delinquency, published in 1942, has become a classic example of their theory. Generally, they observed that due to urbanization and industrialization, many communities contain competitive cultures, which lead to social disorganization. They proved that crime rates are different in different areas and they observed that higher crime rates corresponded to communities with higher population density and environmental degradation and to communities that neighbored industrial or mercantile areas (Kobogianni, 2012).

To some extent, the interest in the geographical aspect of crime decreased between the years 1950 and 1980. Nevertheless, in the 70s this interest began to rise again and it peaked during the 80s, but it did not concern only sociologists and criminologists anymore, but, also, architects and city planners. The term "defensible space" was introduced by Newman, who designed crime prevention methods through the environmental design of neighborhoods (for example more public lights are suggested in order to avoid dark and "dangerous" places). At the same time Wilson and Kelling published their "Broken windows theory", which suggests that abandoned and decaying neighborhoods attract people with delinquent tendencies and behaviors (Kobogianni, 2012).

By the 1970s the police had long recognized the importance of the geographical factor in crime study, by sticking pins, which represented criminal events, into maps that were displayed on walls. What was realized in the 70s was that crime could be understood more in depth by exploring its geographic characteristics. To that end, new techniques for crime analysis emerged. These techniques mainly had to do with crime mapping and the Geographic Information Systems (GIS) (Chainey and Ratliffe, 2005).

2.1.2. Geographical Approach

Even though, it is still possible to use basic mapping techniques, such as sticking pins into maps, being able to obtain a variety of spatial and temporal information of criminal events is essential for a thorough analysis. Unless data are computerized and analyzed by specialist software and specific processes, that information will remain underutilized (Ratcliffe, 2010). Analog data and manual processes do not always provide accurate, reliable and comprehensive results (Johnson,

2000). Technological advances are fundamental both in the study of crime and in the creation of prevention policies (Anselin et al., 2000). As it was stated by Johnson (2000, p.2) "Digital maps are the quickest means of visualizing the entire crime scenario". Crime maps can easily display high or low concentrations of criminal events, which can aid the crime analysts in drawing conclusions and in developing appropriate restraining and preventing techniques. This can be attributed to the fact that they are able to understand where and why crimes occur (Johnson, 2000; Ahmadi, 2003; Chainey and Ratcliffe, 2005). Crime analysis uses both qualitative (non-numerical) and quantitative (numerical) data (Boba, 2001).

2.2. Crime Theories

The connection between criminal activities and the socio-economic characteristics of any society is undeniable. Due to the complex nature of the subject, many scientific fields (such as psychology, sociology, geography, criminology and demography) have studied it from their own perspectives (Malik, 2016). This section presents the theories that have been developed in order to study crime from a spatial point of view. These theories do not lead to different types of crime analysis, but they explain crime from different perspectives that should be combined in order study criminal activities thoroughly.

A general theory could be the one defining the four dimensions of each crime. These dimensions are: the offender, the victim or target, the place and time and the law. The first two dimensions suggest that in order for an action to be considered a crime, there should be an offender with a motive and a victim or target. The third dimension includes both the place and the time that a crime occurs, as well as the conducive conditions for the offender to act. The last dimension consists of the social reaction to the crime and the legislation that prohibits it from occurring (Brantingham et al., 2005; Kobogianni, 2012).

2.2.1. Social Disorganization Theory

The Social Disorganization Theory was developed by Shaw and McKay in the 20th century. This theory establishes the correlation between crime and its geographical distribution. A society is considered organized when all of its members have developed some kind of coherence regarding

their common goals and behavioral standards. A disorganized society presents a difficulty in putting common values across its residents and in preserving active social controls[1] (Bakirli, 2005).

According to criminal literature the main factors that lead to social disorganization are:

1) The financial conditions of the society. Even though many researchers argue that there is no direct link between this factor and crime rates, the highest rates of juvenile delinquency have been noticed in areas with the lowest average income.
2) The demographic composition of the population of the society. High crime rates are often associated with the gathering of a great number of foreign immigrants.
3) The population mobility in the society. Due to the gathering of immigrants in certain areas (because they cannot afford to reside in different ones) in many cases, old residents decide to move to more upgraded areas.
4) City growth, and particularly business and industry development in residential areas (Bakirli, 2005; Matsueda and Grigoryeva, 2014).

Shaw and McKay, found these patterns, which they presented, to be regular and consistent and noted that crime rates in a society change slowly or not at all, unless there is another significant disturbing factor (Ratcliffe, 2010).

2.2.2. Pattern Theory

A significant crime theory is the Crime Pattern Theory, which is, also, the most under-researched area of spatial criminology (Ratcliffe, 2010). A crime pattern is, according to the International Association of Crime Analysts (IACA, 2011) a group of at least two crimes that are unique due to specific conditions. These conditions include:

1) Sharing at least one similarity in the type of crime, the behavior or/and the characteristics of the offender or the victim, the property that was taken, or the location where the crime was committed.
2) The absence of any relationship between the victim(s) and the offender(s).
3) The distinction of the specific set of crimes from others, due to the notable similarities.
4) The limited duration of the criminal activities.

[1] The term “social controls” refers to the ability of any society to self-regulate itself on the principles and values that it has determined (Bakirli, 2005).

5) The treatment of the specific set of crimes as one unit of analysis.

The distinction of these patterns can help notably the police forces and the criminologists, due to the fact that they offer more information about the crime and the offender than a singular criminal case would. Every day, police analysts examine data in order to link cases by key factors and export more information, so that they can attempt to prevent and reduce crime (International Association of Crime Analysts (IACA), 2011).

2.2.3. Routine Activities Theory

From every research it is obvious that crime is not spread evenly across space and it is not random either. It clumps in some areas and it is absent in others (Block and Block, 1995; Ackerman and Murray, 2004; Brantingham et al., 2005; Eck et al., 2005; Ratcliffe, 2010). However, in Block's view "the relationship between crime and place is neither uniform nor static", as the patterns may change over time due to a variety of factors (Block and Block, 1995, p.147). Different types of crime are possible to cluster in different areas. This can be the result of specific characteristics of certain areas that attract potential offenders. Such characteristics may be community disorganization, lack of social services in the area, easy road network accessibility or routine activities that define the area (for example certain areas serve as nightlife centers). The greater risk of being a victim of criminal activities in certain areas has led people to avoid them and choose others for residence, services and recreation (Block and Block, 1995; Eck et al., 2005; Congdon, 2013).

In this context, one of the main crime theories suggests that crime rates are based on people's everyday movements. The Routine Activities Theory, indicates that offenders are more active and commit crimes near the areas where they spend most of their everyday time and victims are victimized near areas where they spend most of their time too. Offenders have awareness and their own perceptions of the environments in which they frequent and, in order to benefit themselves, they can identify efficient criminal opportunities from dangerous targets. The structural changes that can influence crime rates, according to this theory, are the presence of motivated offenders, the presence of suitable targets and, mainly, the presence or absence of capable guardians against violations (Cohen and Felson, 1979; Groff and La Vigne, 2002; Brantingham et al., 2005; Brunsdon et al., 2007; Tompson et al., 2009).

2.2.4. Socio-Economic Factors

In criminal literature, many theories and studies have arisen, which describe the relationship of crime rates and specific socio-economic factors. These factors are said to have a direct and extremely significant effect on crime. The majority of these factors are, also, connected with social disorganization.

Inequality functions as one of the most important and general socio-economic factors. It is stated by many criminologists that rates of violence tend to be higher in more unequal and disadvantaged societies (Krivo and Peterson, 1996; Sampson et al, 1997; Cameron, 2001; Rufrancos et al., 2013; Matsueda and Grigoryeva, 2014). Extreme disadvantage in certain areas is linked with extraordinarily high levels of crime, because the conditions that encourage criminal activities are particularly intense. The residents of these areas are more used to viewing and being present in criminal events that, eventually, they become a common aspect of their lives. The need to adapt to environments where crime is predominant leads individuals to use or appear ready to use violence, in order to defend themselves and their properties. As more people adopt defensive to violence attitudes, violence escalates (Krivo and Peterson, 1996). Additionally, inequality, and especially income inequality, causes an augmentation of feelings of unfairness, which lead individuals from areas with low average income, to develop the idea that they can reduce this unfairness through crime. These individuals are more sensitive to inequalities and choose to adopt more risky behaviors when the low-risk or legal activities offer them poor returns (Rufrancos et al., 2013).

The concentration of disadvantages in an area results in fewer informal, community-based networks of control, formed by families, neighbors and other social groups, in order to watch each other's properties, intercede in crimes and supervise suspicious juvenile activities that may evolve to crime. Moreover, in disadvantaged communities, more unemployed and irregularly employed people reside. This means that these people remain idle for large parts of the day, so they are likely to spend notable amounts of time in surroundings where delinquent behaviors may be developed (street corners, local taverns, pool halls) (Krivo and Peterson, 1996).

Since criminal activities are largely motivated by the economic benefits that they may offer (Altindag, 2012), income levels and unemployment play a significant role in crime rates (Gould et al., 2002; Huang et al., 2004). According to studies, crime rates increase when unemployment rates increase and when average wage rates decrease (Lochner, 2007). The decision to take part

in criminal activities prevails in cases that the value of the potential benefits surpasses the benefits that come from legal activities, and on the other hand, in cases that the benefits from legal activities are adequate, individuals do not consider taking part in illegal activities (Raphael and Winter-Ebmer, 2001; Kling et al., 2004; Diamanti, 2010). Is should be noted, that researchers have observed that the number of criminal opportunities may be lower during recessions. This happens, because under these conditions, potential victims have a lower income, which they are more protective of and which they try to defend (Raphael and Winter-Ebmer, 1998).

Other factors that seem to influence crime rates are educational attainment and urbanization. Economic theory suggests that there is a negative correlation between most types of crime and education attainment, which has been documented by many studies (Huang et al., 2004; Lochner, 2007). In other words, low educational attainment affects in a negative way the crime rates. When the educational level of an area increases, crime rates decrease and when there is a decrease in the educational level, the crime rates increase. On the other hand, it is argued that the crime rates of large cities and urbanized areas, are notably higher than the equivalent crime rates in rural areas (Almeida et al., 2003; Zhong et al., 2011; Zakaria and Rahman, 2016; Malik, 2016). This phenomenon occurs due to several socio-economic factors that appear more and more intensely in urban areas. Such factors include environmental characteristics and financial, social, political and demographic conditions (Zakaria and Rahmna, 2016). It is suggested that in every part of the world, over a five year period, two out of three residents of urban areas are victimized at least once (Ackerman and Murray, 2004).

Scientists have, also, studied other factors, such as age, which represents one of an area's demographic characteristics. Age is considered to be the easiest fact about crime to study, because it is always recorded when an incident occurs. The effect of age in crime rates is owed to the fact that younger individuals are characterized by explosive and irresponsible behavior, which is more likely to lead to criminal activities (Diamanti, 2010). Individuals in the age group 15-25 appear to have the highest arrest rates of any other age group (Hirschi and Gottfredson, 1983), while it is noted that crime declines with age (Sampson and Laub, 1993). It is, also, suggested that mothers who give birth for the first time at a very young age (17 years old or younger) are slightly more likely to have delinquent children (Farrington and Welsh, 2007).

2.2.5. General Strain Theory

A different theoretical perspective, which is not directly related to spatial characteristics, attempts to explain why individual level disadvantages may lead to crime. The General Strain Theory (GST) is examined by analyzing the effect of strains on crime and it notes that strains or other stressors increase the possibility of developing negative emotions (like anger and frustration), which can ultimately lead to crime. The term "strain" refers to any relationship in which the individual is not being treated by others the way they would like to be treated. According to Agnew (2001), strains are most likely to result in crime when:

1) They are interpreted as inequitable.
2) They are considered events of great importance.
3) They are linked with low social control.
4) They create pressure or motivation to engage in criminal activities or criminal coping.

Some types of strain that are presumed to have a relationship with crime rates, are parental rejection, harsh parental discipline, child abuse, child neglect, negative school experiences, peer abuse, marginalization, experiences with prejudice and discrimination and the ability to achieve selected goals. However, it should be noted that certain types of strain are not related, or are weakly related, to crime. One of these strains is the failure for educational or professional attainment (Agnew, 2001).

In the same framework, other theories have, also, emphasized on the indirect importance of social ties in crime rates. These theories suggest that antisocial behavior during an individual's childhood, can later disrupt professional and personal relationships and consequently increase the possibility of developing criminal behavior. Other factors that may lead to criminal activities include high testosterone levels that increase aggression and child aggressive behavior that provokes the decrease of parenting quality and of school commitment and the increase of deviant relationships (Entner Wright et al., 2001).

2.2.6. Problems

In the study of crime and its theories, many difficulties arise. One of the most significant ones is the choice of the spatial scale, which is defined as the Modifiable Areal Unit Problem (MAUP) and was described in detail by Stan Openshaw. The problem of spatial scales means that, when

mapped, different geographical boundaries may produce different visual representations (Ratcliffe, 2010; Leitner, 2013).

Additionally, it is argued that some analyses may not be fully possible due to missing data. Missing data, which may occur because of confidentiality issues, can prevent a profound analysis in certain areas (Leitner, 2013). Another reason why some data cannot be studied is that they may not be recorded. Crime statistics only use an unknown share of all the delinquencies that occur, because they are the only ones that have been reported to the police (Entorf and Spengler, 2002). In some cases, only a small proportion of crimes are reported to the police. This is due to the fact that some victims perceive some types of crime as not very serious and they think that they are not worth the burden, and even the embarrassment, of reaching a police station, completing various forms and answering a variety of questions (Del Frate, 1998). In some cases victims avoid to report a crime in lack of confidence in the police, or even in fear of it (International Centre for the Prevention of Crime (ICPC), 2010).

Another difficulty that arises, from crime studies, is the troublesome comparison of international crime data. This problem stems from the differences in the recording and the reporting practices that are being used. Each country's crime rate depends on its justice system and the way the police defines, records and counts crime (Entorf and Spengler, 2002; Gruszczynska, 2004; International Centre for the Prevention of Crime (ICPC), 2010).

2.3. Crime Trends

The immediate and effective confrontation of crime is one of the priorities of every society, because a criminal event can occur anywhere, anytime and in any form (Diamanti, 2010). Historically, crime has preoccupied human societies, since their creation. Nevertheless, the oldest records of criminal activities that exist are from the medieval times and concern Europe, and particularly France and England (Dean, 2001). These records indicate some serious criminal cases that troubled the societies of those centuries.

Violence and crime consist a major threat to every society in the world and function as an obstacle to development. In a global level, it is observed that the highest violent crime rates, and particularly homicides, are found in Central and South America and in Africa. The lowest rates of

homicide are found in the European countries and the lowest rates of all types of crimes are found in Asian developing countries (Del Frate, 1998; Harrendorf et al., 2010). The low rates of Asian developing countries can be attributed to the fact they tend to not report every crime that has been committed. During the decades of 1980 and 1990, the rate of intentional homicide increased by 50% in Central and South America and by 100% in Eastern Europe and Central Asia (Fajnzylber et al., 2002). According to data from the United Nations Office on Drugs and Crime, in 2004, more than 490,000 people were the victims of intentional homicide. This number represents a world average homicide rate of 7.6 per 100,000 people (Harrendorf et al., 2010). The same rate in 2014 and 2015 was significantly lower and specifically it was 5.3 per 100,000 people. In 2015, the countries that had the highest intentional homicide rates were El Salvador (109 per 100,000 people), Honduras (63.8), Venezuela (57.1), Jamaica (43.2), South Africa (34.3), the small states of Caribbean (31.5) and Trinidad and Tobago (30.9). Other countries of South America, also, had notably high rates. Two of these countries were Brazil (26.7) and Colombia (26.5). These high levels of homicide in Central and South America, can be explained by the brutal civil wars that the containing countries had been experiencing for nearly five decades. It should be noted, that the prison population of these countries has grown rapidly since the beginning of the 21st century. In particular, this population increased by 28% between 2003-2005 and 2012-2014. In this area, significantly high rates of crime associated with drugs are, also, observed (Commission on Crime Prevention and Criminal Justice, 2017).

On the other hand, Andorra and Liechtenstein were the only countries, in 2015 (with available data) that had a rate of 0 homicides per 100,000 people. Other countries with rates lower than 1 per 100,000 were: Singapore (0.2), the region of Macao in China (0.2), the region of Hong Kong (0.3), Austria (0.5), Ireland (0.6), the Netherlands (0.6), Czech Republic (0.7), Poland (0.7), Spain (0.7), Switzerland (0.7), the United Arab Emirates (0.7), Germany (0.8), Greece (0.8), Italy (0.8), Croatia (0.9), Iceland (0.9) and Slovakia (0.9) (World Bank Group, 2017). From these data it is obvious that the majority of these countries are located in Europe.

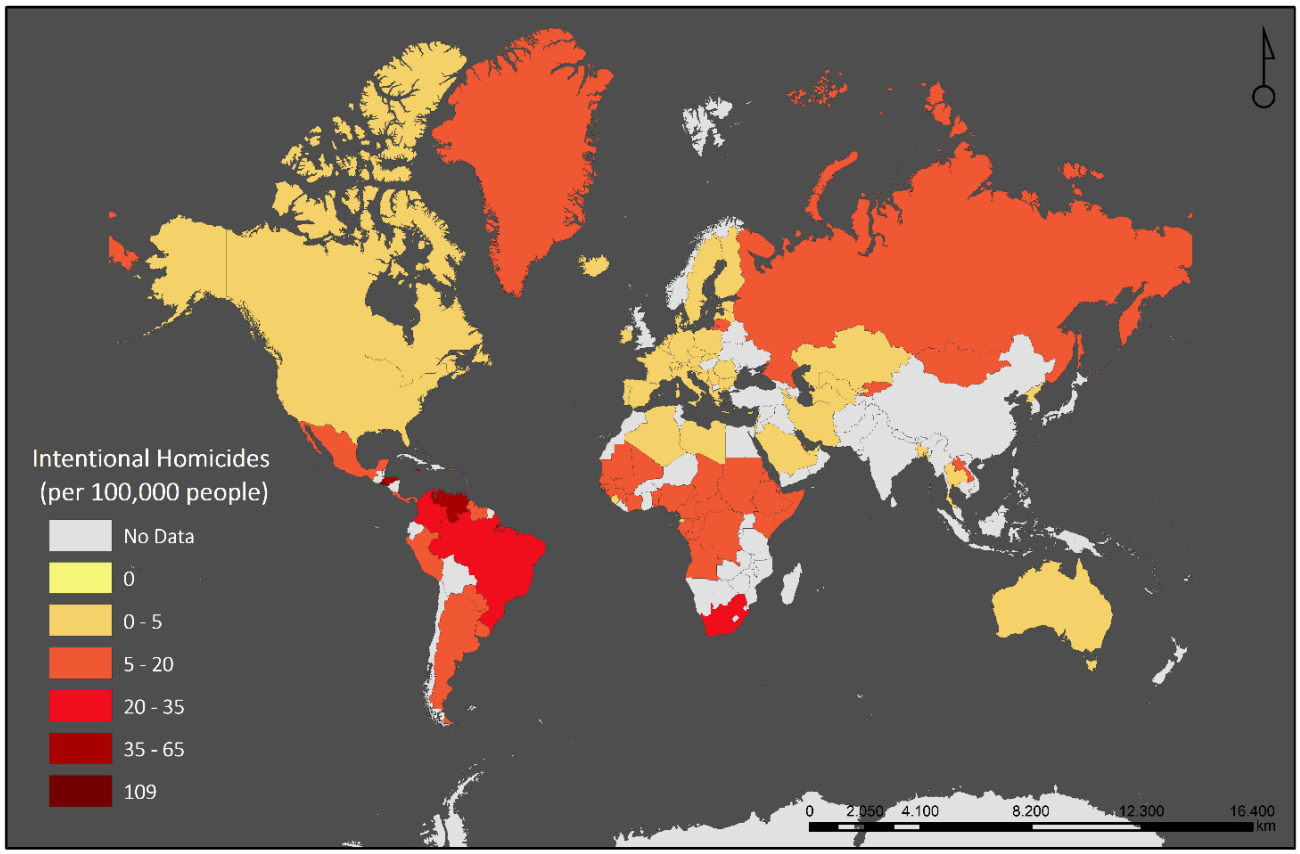

Figure 1. *World map of intentional homicide rates in 2015 (per 100,000 people)*

Data source: World Bank Group, 2017

In Europe, and particularly in Eastern Europe, the transformation that began in the late 1980s, caused a variety of political, economic and social changes. These changes that included the opening of borders, the privatization of public enterprises, the influx of foreign capital and the augmentation of competition which is associated with early capitalism (Commission on Crime Prevention and Criminal Justice, 2017), transformed Europe rapidly. From the criminological point of view, and as it was previously suggested, social changes increase drastically criminal opportunities (Gruszczynska, 2004). Hence, the crime rates in Europe were significantly higher during the 1990s and the early 2000s. This tendency changed from 2007, when the rates decreased notably. In 2015 the crime rate in Europe was 50% lower than it was in 1993.

The prison population of the countries of the European Union rose between the years 2008 and 2012 and then dropped by 3.6% in 2013, by 3.5% in 2014 and by 2.9% in 2015. This rate in 2015 was 6.4% lower than the equivalent rate in 2008 (Eurostat, 2017). In Eastern Europe, the incarceration rates were quite high in the early 2000s, but they dropped by 27% between the periods 2003-2005 and 2012-2014. In the same area, during the period 2010-2012, approximately 42% of the prison population had previously served another prison sentence. The majority of the

sentences, in Eastern Europe, last from 1 to 5 years (46%) and from 5 to 10 years (33%). The most common type of crime in Eastern Europe is property crime (37%), followed by homicide (22%), drug crime (16%) and violent crime (13%) (Commission on Crime Prevention and Criminal Justice, 2017).

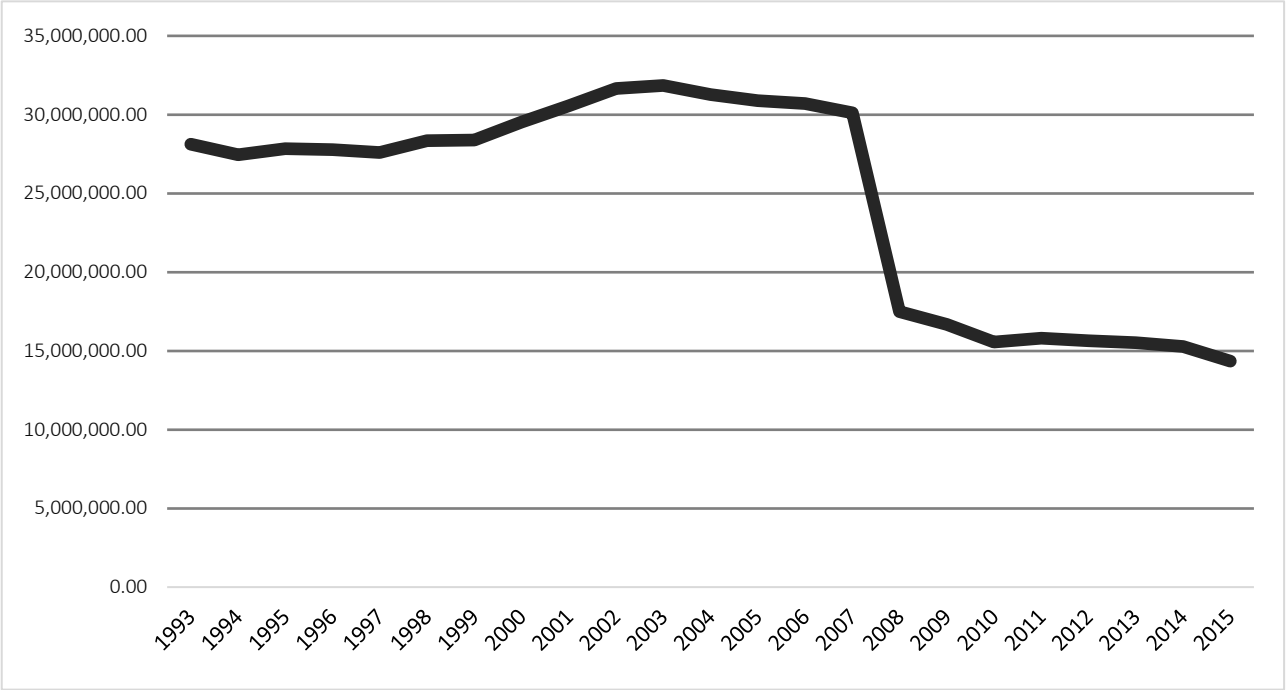

Figure 2. *Number of recorded crimes in Europe during the period 1993-2015*[2]

Data source: Eurostat, 2017

In a country level, the European countries with the highest crime rates for 2015 were the United Kingdom (and particularly England and Wales), France, Germany and Italy. The country with the lowest crime rate was Liechtenstein, which can be explained by the fact that it is a tiny state with a population of approximately 37,000 people. During 2015, and during the previous years in general, the most common crime types in each European country, were thefts. Other common crime types were burglaries and assaults (Eurostat, 2017).

[2] The recorded crime types are: intentional homicide, acts causing harm or intending to cause harm to the person, injurious acts of a sexual nature and acts against property involving violence or threat against a person, robbery, burglary, theft and unlawful acts involving controlled drugs or precursors. The data are collected from 37 European countries: Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Croatia, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, England and Wales, Scotland, Northern Ireland (UK), Iceland, Liechtenstein, Norway, Switzerland, Montenegro, FYROM and Serbia.

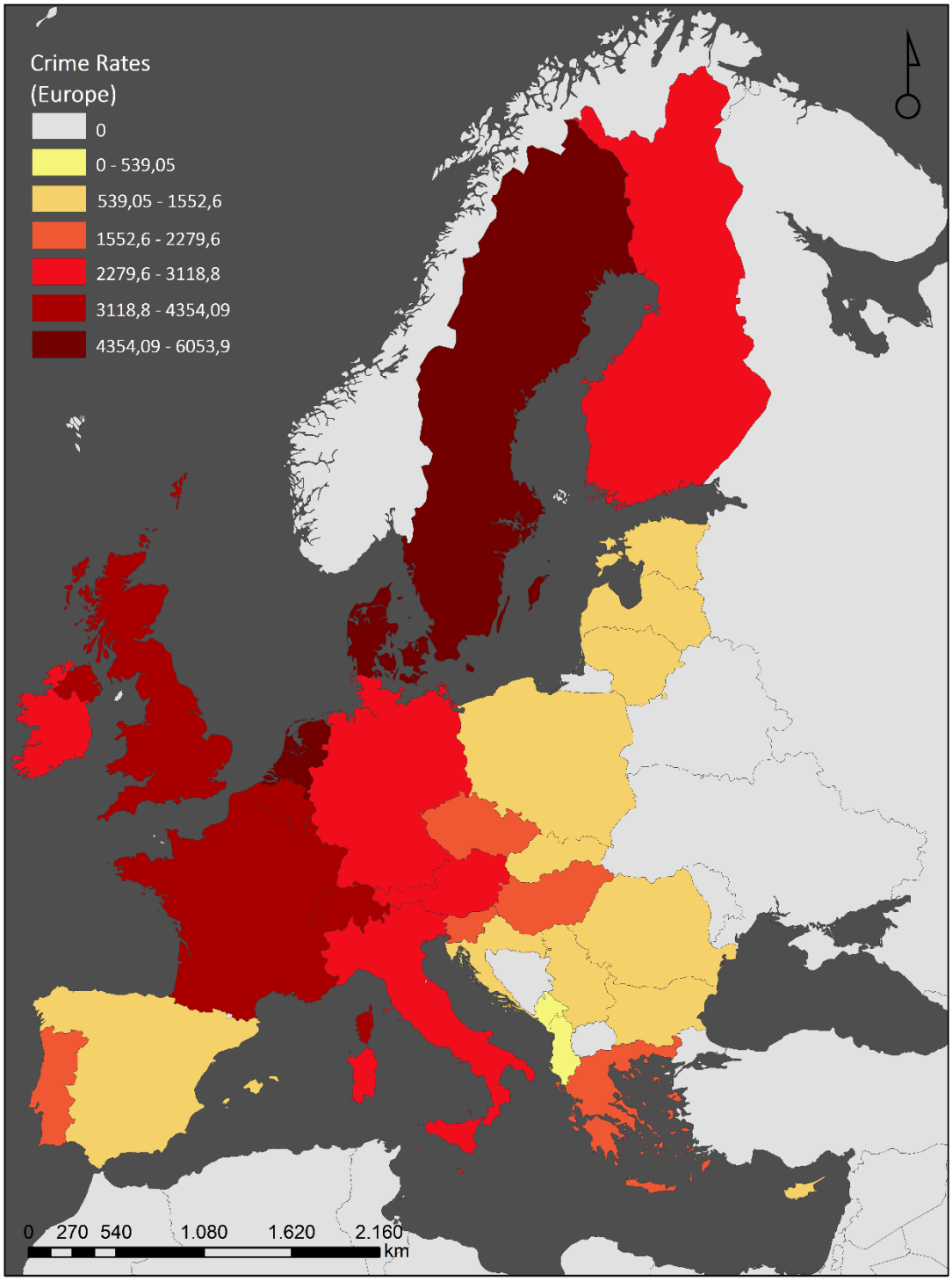

Figure 3. *Map of the crime rates in European countries in 2015[3](per 100,000 people)*

Data Source: Eurostat, 2017

[3] The recorded crime types are: intentional homicide, acts causing harm or intending to cause harm to the person, injurious acts of a sexual nature and acts against property involving violence or threat against a person, robbery, burglary, theft and unlawful acts involving controlled drugs or precursors. The data are collected from 34 European countries: Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Croatia, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, United Kingdom, Iceland, Liechtenstein, Switzerland, Montenegro, Albania and Serbia. The number for each country might differ, because the records for some types of crime and for some countries do not exist. Crime rates are calculated by dividing the number of the recorded crimes by the total population and then multiplying the result by 100,000.

For Greece (for which there are crime data since 1930 (Roinioti, 2009)), the crime trends tend to follow the European example. After 2005 there seems to be a significant decrease to the number of criminal events in the country. From 2010 to 2016 the number of criminal incidents appears to be quite stable, with a slight increase tendency being noticed in 2016.

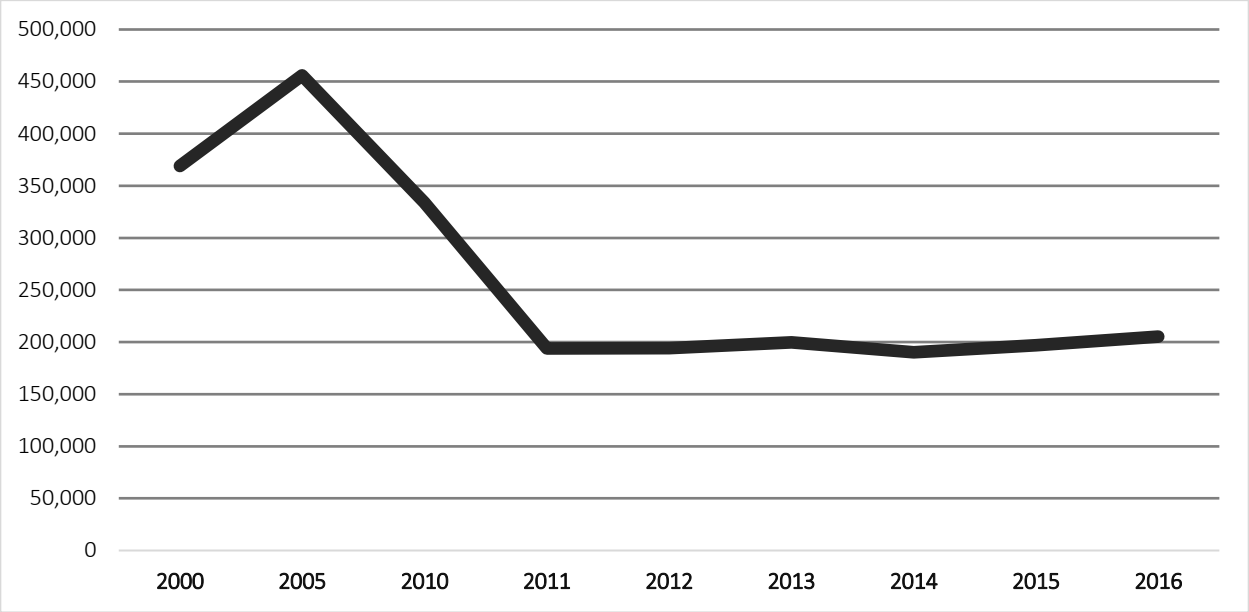

Figure 4. *Number of recorded crimes in Greece during the period 2000-2016*

Data source: Hellenic Statistical Authority, 2017

From the following map, it is obvious that the majority of the crimes that occurred in 2016, seem to be located in the region of Attica and in the region of Central Macedonia. In these two regions resides the majority of the country's population, because they contain the two biggest cities of the country, Athens and Thessaloniki. So this concentration of crime is quite understandable. On the other hand, the regions with the lowest crime rates are the ones with the lowest population density. Specifically, these regions include Epirus, the islands of the northern Aegean Sea and the islands of the Ionian Sea.

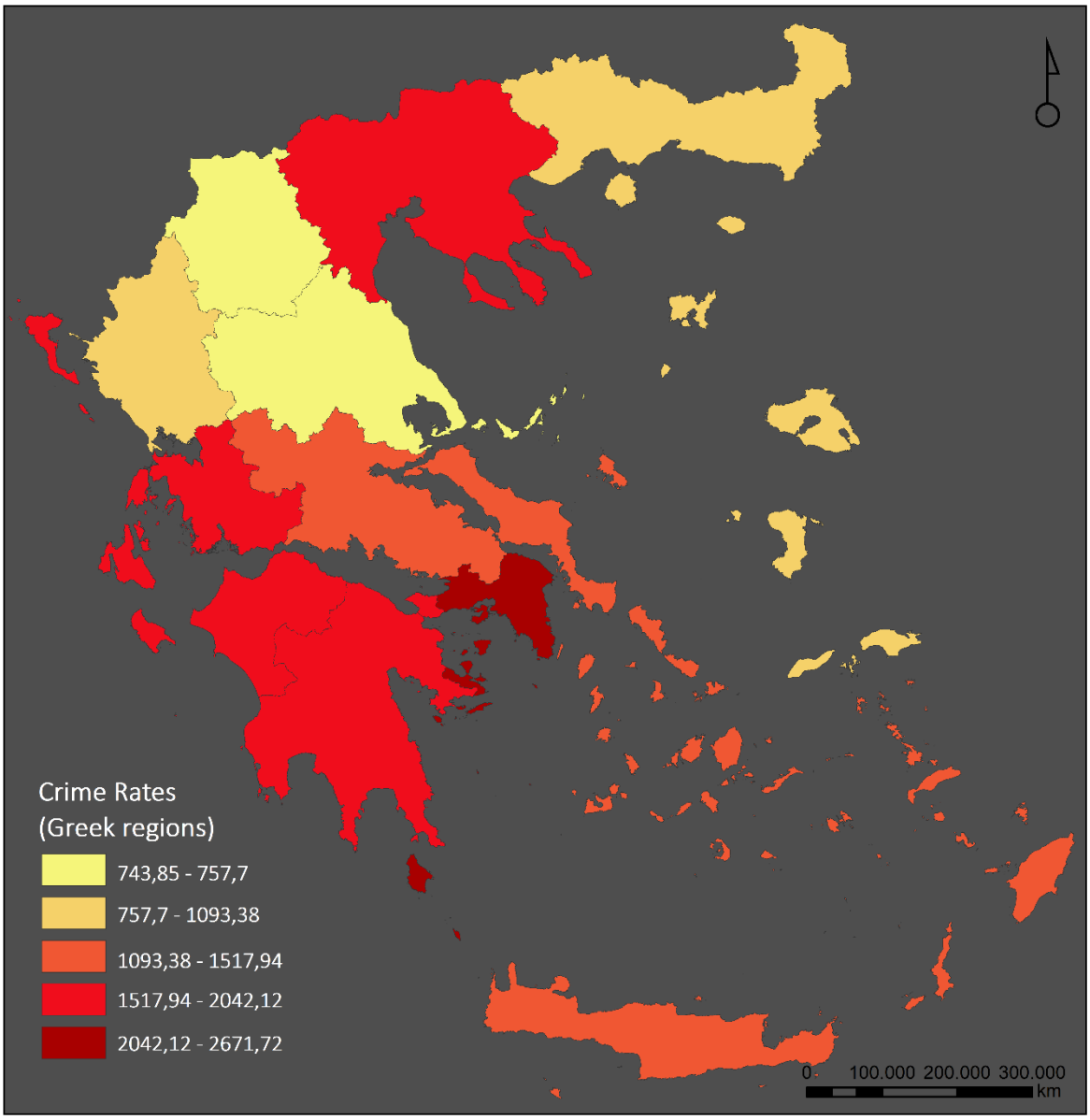

Figure 5. *Map of the crime rates in Greek regions in 2016 (per 100,000 people)*[4]

Data Source: Hellenic Statistical Authority, 2017

[4] Crime rates are calculated by dividing the number of the recorded crimes by the total population and then multiplying the result by 100,000.

3. Literature Review

Despite the fact that crime is a highly significant problem for every society and its study is essential for the development of prevention policies, our knowledge regarding the factors that make some places to have higher crime rates than others is still considered limited (Goldsmith et al., 2000; Fajnzylber et al., 2002). To that end, many social sciences' scholars, from different fields, have studied the subject, so that an abundance of papers has been published and continue to be published. Some of the published papers emphasize on the presentation and the analysis of spatial patterns. For example, Sherman et al. (1989) presented a Poisson model for spatial data, from 323,979 calls to the police in the city of Minneapolis over a one year period. The authors observed that in this specific city, crime was both rare and concentrated and that this concentration varied by offence type. They, also, observed a few hot spots which produced more calls to the police.

The majority of the literature tries to identify the factors that increase or decrease crime rates in specific areas. In two different papers, the authors apply spatial analysis for crime data in Italy. In the first one the authors examine the relationship between crime rates and the share of industrial activities in the areas in consideration, the rate of unemployed young males and the presence of foreigners in these areas. The results confirm the relevance of these factors in the development of criminal activities (Cracolici and Uberti, 2008). In the second paper, the authors claim that the unemployment rates, the rates of employment in the service sector and the number of public workers have a significant effect on crime rates. Additionally, they point out that the average education level of the population, as well as the share of young males in the population, do not have any effect on crime rates (Marselli and Vannini, 1997). In a research, whose study area is close to the previous papers, the author observes the relationship between crime and land use categories on the Maltese islands. From the analysis it is suggested that the residential areas have the highest crime rates, and particularly the highest serious crime rates (Formosa, 2007). In an analysis for the city of Volos, in Greece, the author argues that higher crime rates seem to be related to the population density, the number of large buildings and the educational and professional attainment of the population (Alevizaki, 2010). A similar approach is, also, used in a paper that focuses on Germany. The paper demonstrates that the share of unemployed young persons and the share of foreigners have an important negative influence on crime rates (Entorf and Spengler, 2000).

The authors of a paper published in 2002, examine the impact of income rates and unemployment on crime, in the United States of America (USA). They conclude that both of these factors are significantly related to crime, with the first one being related on a greater level. Specifically, they suggest that lower income and higher unemployment increase crime rates (Gould et al., 2002). In a different paper that focuses on the USA, specifically on Baltimore County, the author notes that higher population densities are related to higher levels of crime (Harries, 2006). In a report, in which the author applies spatial analysis for Appalachia (1980-1990), different regressions are applied using a variety of variables, such as the population size, the poverty and unemployment rates, the per capita income, the growth, the residential mobility, the age structure, the racial/ethnic diversity, the educational attainment, the family stability, the changing household structures and the changing industrial composition. The results indicate a positive correlation to crime, of the share of divorcees, the share of female headed households and of racial diversity. A negative correlation is proved to exist between violent crimes and the High School drop-outs share, as well as the share of unemployed individuals. Residential mobility and poverty levels do not seem to be predictors of violent crime. In addition, the author suggests that the share of persons aged from 15 to 29 years old, is not a significantly correlated factor to crime in metropolitan areas, but its significance increases importantly in non-metropolitan areas. A conclusion that is being emphasized is that the levels of crime in a specific location are strongly influenced by conditions in neighboring locations as well (Cameron, 2001).

According to Carcach (1999), social and economic disadvantages and the share of individuals aged from 15 to 24 years old, are related to firearm-related homicides in Eastern Australian states. Unemployment does not seem to be associated with these crimes and neither does the share of persons that are older than 55 years old. Concluding, the author suggests that factors like unemployment, economic inequality, access to social services, family relationships and others need to be more thoroughly examined. In a study that spatial analysis is applied about the patterns of property crime in Peninsular Malaysia, it is argued that there is a negative correlation between the area's property crime rate and the area's income rates (Zakaria and Rahman, 2016). The authors conclude that the high incidence of poverty led to the augmentation of property crime rates. A study that explores the spatial characteristics of crime in Shanghai, reveals that 92.8% of violent crime and 95.3% of property crime occurred in industrial, traffic and public land, as well as in residential areas. The authors of the study note that neither violent nor property crimes were related to population density (Zhong et al., 2011).

The effect of unemployment on crime rates, which has largely been mentioned in many theoretical approaches of crime explanation, has concerned a lot of researchers. According to Altindag (2012), who uses data from 33 European countries, an increase in the unemployment rate by 1% increases the property crime rate by about 2%. A study by Huang et al. (2004), also, claims that there is a positive correlation between crime rates and unemployment (as well as low levels of educational attainment and poverty). On this correlation emphasize two other papers, both by Raphael and Winter-Ebmer. The papers highlight that even though unemployment seems to be positively correlated to property crimes (2001) in many cases, it has no relationship with violent crimes (and particularly with murder) (1998).

Other factors that seem to be related to crime rates are income inequality, social disorganization and educational attainment. According to a paper published in 2002 (that uses data of developed and developing countries for the period 1970-1994, acquired from the United Nations World Crime Surveys), both economic growth and income inequality affect crime rates (Fajnzylber et al., 2002). In the same study, the authors claim that the average income, the average educational attainment and the degree of urbanization of an area are not significantly and consistently related to crime rates. In a more specific paper, the authors suggest that an augmentation in income inequality increases property crime rates. The same pattern (in a lower level) is, also, noticed in violent crimes, such as homicide and robbery (Rufrancos et al., 2013). In a different study, which focuses on Great Britain, the authors claim that low income, ethnic heterogeneity, residential mobility and family disruption lead to social disorganization, which ultimately increases crime rates. The authors define that the level of social disorganization is "measured in terms of local friendship networks, control of street-corner teenager groups and prevalence of organizational participation" (Sampson and Groves, 1989, p. 774). In an article titled "Extremely disadvantaged neighborhoods and urban crime" the authors use a variety of variables to indicate the relationship between crime and disadvantaged neighborhoods. These variables are the share of families headed by females, the share of civilian noninstitutionalized males older than 16 years old who are either unemployed or not in the labor force, the share of persons older than 16 years old who are employed, the rental occupancy, the vacancy rate, the share of the male population between the ages 15 and 24 and the share of the ethnic heterogeneity. The results indicate that all these variables have a positive relationship with property and violent crime (Krivo and Peterson, 1996). In the same context, in a research published in 2001, data from an experiment, which was based on a randomized housing-mobility, are used, to examine the effects of relocating families from

neighborhoods of high poverty rates to neighborhoods of low poverty rates (and the opposite) on juvenile crime. The findings suggest that providing families with the opportunity to move to neighborhoods with lower poverty rates, reduces juvenile violent criminal activities, but it might cause an increase in property crime offences, at least for a short time period of adjustment (Ludwig et al., 2001). From a different perspective, in a paper that emphasizes on the effects of educational attainment on crime rates, the author suggests that education increases the future average income rates. So individuals who are enrolled in school will be less likely to engage in crime than others. According to the author, school attendance reduces property crime rates but increases violent crime rates among juveniles (Lochner, 2007).

Numerous studies question the relationship between crime rates and other spatial factors that are not affected by social conditions. For example, in a study published in 1982, nine regression models are used in order to examine the contribution of tourism to nine different types of crime, in the USA for the year 1975. The results of this study indicate that tourism does not affect at all five out of the nine crime types and it negatively affects the remaining four in such a low degree that its attribution could be considered insubstantial (Pizam, 1982). Vegetation is, also, a factor that has been considered as a characteristic that might decrease crime rates. Kuo and Sullivan (2001), hypothesize that vegetation, high-canopy trees and grass in particular, defines crime rates in neighborhoods of low average income. Specifically, the authors indicate that higher levels of vegetation can be associated with lower levels of both property and violent crimes. They, also, claim that the presence of such vegetation in an area, encourages the residents to use outdoor spaces more often, which consequently leads to these spaces being under surveillance for more time. Hence less crimes occur in these spaces. It is, also, noted that the presence of vegetation can diminish mental fatigue, which is often described as an affecting factor.

Alcohol consumption and the development of certain type of businesses, are other factors whose effect on crime rates has been considered important by numerous studies. In a paper published in 2004, the authors present that the social and structural factors explain 59% of violent crime rates in Austin and 39% in San Antonio. Adding alcohol as a factor these percentages increased. Specifically, 71% of crime rates in Austin were explained by these variables and 56% in San Antonio (Zhu et al., 2004). On the other hand, in another paper the authors explore the correlation between the crime rates of certain communities in Chicago and the presence of taverns, bars and liquor stores and they present different results. From the analysis, the authors concluded that concentrations of liquor establishments should not be used as a variable in crime rates' analysis,

neither for crimes occurring inside these establishments, nor for crimes occurring in the surrounding areas (Block and Block, 1995). In the same framework, in another paper, focused on Atlantic City, it is indicated that the levels of all crimes appear to be higher in the post-casino years (1978-1984) than in the earlier period (1972-1977), with other factors controlled. In addition, the authors suggest that crime rates reduce as the distance from the city increases (Hakim and Buck, 1989).

The various spatial factors may have an indirect effect on crime rates. A paper published in 2001 explains their effects on social relationships which, subsequently, lead to crime. The authors note that social associations, such as educational and employment relationships, family ties and partnerships, reduce crime rates, while antisocial ties, such as delinquent peers, increase crime rates among individuals with low self-control (Entner Wright et al., 2001). Other factors that have been examined are age (negative correlation) (Moberg, 1953; Hirschi and Gottfredson, 1983), gender (males appear to perform more crimes) (Kling et al., 2004), weather (positive correlation) (Field, 1992; Murataya and Gutiérrez, 2013) and urbanization (positive correlation) (Malik, 2016). In addition, Osgood (2000) suggests as variables the residential instability, the ethnic heterogeneity, the family disruption, the unemployment rate and the proximity to metropolitan counties.

4. Data

As mentioned above, the main goal of this dissertation is to examine and to give prominence to the social and economic factors that have an effect on crime rates. The study focuses on the European Union. The principal source of the used data is the European Statistical Authority (Eurostat). The main dataset, acquired from Eurostat, is a file that contains information about the number of the recorded crimes by the police in the year 2010. These data correspond to subdivisions of the European Union and the European Free Trade Association (EFTA)[5] countries and particularly to their regions in a NUTS 2 level. The term "NUTS" refers to the classification of territorial units for statistics (Nomenclature des unités territoriales statistiques). NUTS 2 is the second lowest level of the hierarchy (the highest is NUTS 1 and the lowest is NUTS 3).

Nevertheless, the dataset that was previously described, does not include the recorded crimes for the regions of every European Union country. Specifically, data for the United Kingdom, Greece and Ireland's regions were included only in a NUTS 1 level. The missing NUTS 2 data were retrieved from the National Statistical Authorities of these countries. In the case of the United Kingdom, separate data were retrieved from the statistical database of England and Wales and the statistical database of Scotland. Northern Ireland consists of only one NUTS 2 region, which is included in one NUTS 1 region, so its data were available from Eurostat.

The number of the recorded crimes consists of the number of the intentional homicides, the number of the robberies, the number of the burglaries of private residential premises and the number of thefts of motorized land vehicles. By definition the term "intentional homicide" refers to the intentional killing of a person (including murder, manslaughter, euthanasia and infanticide, but excluding attempted and unsuccessful homicide, deaths from dangerous driving, abortions and help with suicide). Intentional homicide is considered to be the most serious type of violent crime and that is the reason that contributes to the more effective recording of this type of crime. The other types of crime that are comprehended in this dataset are property crimes. Robbery is a type of crime that involves the use of violence or the threat of violence in order to steal from one person. The term does not include pick-pocketing, extortion and blackmailing. Burglary is described as the result of gaining access to a private or closed dwelling or building by force with

[5] The European Free Trade Association (EFTA) is an intergovernmental organization that aims on the promotion of free trade and on the economic integration of its four member-states. These member-states are Iceland, Liechtenstein, Norway and Switzerland (The European Free Trade Association, 2018).

the intent to steal goods (Harrendorf et al., 2010; Eurostat, 2017). The thefts of motorized land vehicles include stealing all types of land vehicles that have an engine, run the road and are used to move people (cars, motorcycles, buses, construction and agricultural vehicles, trucks, etc.) (Eurostat, 2017).

Nevertheless, the number of the recorded crimes in each one of the NUTS 2 regions is not a measure that can easily be compared between the European Union countries' regions. This can be attributed to the fact that the regions, of the countries that were included in this analysis, do not have the same land areas and consequently, the same population sizes. A specific number of recorded crimes can be considered extremely high in a region whose population is significantly small, but negligible in regions with a larger population size. A comparison of the recorded crimes between regions that differ in this area can be exceptionally inexpedient. To that end, the variable that is examined in this dissertation is not the number of the recorded crimes, but the crime rates of the NUTS 2 regions. A crime rate describes the number of the crimes recorded by law enforcement agencies per 100,000 inhabitants. A crime rate is calculated by dividing the number of the recorded crimes by the total population of the reference area. The result is multiplied by 100,000.

$$Crime\ rates = \frac{Recorded\ crimes}{Total\ Population} \times 100{,}000$$

The independent variables that are used in this dissertation, arose from the reviewed literature, but were limited to their availability. Some of them were transformed into ratios for a better understanding and comparison between the NUTS 2 regions, as stated above, as well as for them to be included in a regression model. The selected variables are:

1) The population density of the regions in 2010, which is measured as the number of inhabitants per km^2. The population density is calculated by dividing an area's total population by the land area in km^2 (or square miles). In this case the total population of each NUTS 2 region in 2010, was divided by its total land area. In general, the population density indicates how many people correspond to one square unit (kilometer, mile, etc.).
2) The ratio of the male population of the regions that, in 2010, was older than 15 years old and younger than 64 years old. This ratio was calculated by dividing the number of the male population of each NUTS 2 region that in 2010 was in that age group, by the total population of the regions in the same year.

3) The ratio of the immigrants in every region in 2010. Specifically, the number of the permanent residents that had a foreign citizenship during this year in each NUTS 2 region, was divided by the total population of the regions. This ratio presents the percentage of the immigrants in the total population of each region.
4) The ratio of the inhabitants of the regions that had received no education (recorded in 2010). In order to measure this ratio, the number of the inhabitants that had received no education in each region was divided by the total population of the regions. Since, the number of individuals with no education was calculated based on the total population of any age group, the ratio was created with the usage of the number of the total population. In each case, the denominator of the fraction changes according to the way that the numerator is measured. For example, the denominator can be the total population that is older than 15 years old, the total economically active population etc.
5) The ratio of the inhabitants of the regions that were unemployed in 2010. This ratio was calculated by dividing the number of unemployed individuals in each region, by the number of the economically active population in the region. The economically active population comprises both employed and unemployed persons. Employed persons are considered all individuals aged 15 and over, who during the reference time period, worked at least one hour for pay or other profit (for example family gain). In this category, individuals over the age of 15 who were not at work, but had a job or business from which they were temporarily absent, are included. Unemployed persons are considered all persons in the age group 15-74, who during the reference time period did not have a job, were available to work and they were actively seeking to work (had taken specific steps in order to seek paid employment or self-employment). Individuals who had found a job that would start within a time period of at most three months, are, also, included in this category (Eurostat, 2018).
6) The ratio of the inhabitants of the regions that were employed in the public sector during 2010. This ratio was created by dividing the number of the individuals that were employed in the public sector, by the total economically active population of each NUTS 2 region. Public employees are considered employees in public administration, defense, education, human health and social work activities (Eurostat, 2018).
7) The Gross Domestic Product (GDP) per capita of the regions in 2010. Gross Domestic Product is called the total market value of all the final goods and services that are produced within a country's borders (or other administrative boundaries), annually. GDP per capita

is called the GDP of a country (or region in this case) divided by the total population of this area (Arnold, 2007). The GDP is one of the most significant monetary measures, because it shows the dynamic of any administrative area. This variable is measured as the total euros that correspond to each inhabitant.

8) The average disposable income that every inhabitant of the regions had in 2010. Disposable income is the amount of money that every household has available for spending and saving. This income is calculated after the income taxes have been accounted for. Allowances (for example disability living allowances etc.), are, also, included in the disposable income.
9) The ratio of the total artificial land cover in each NUTS 2 region in 2010. The ratio was calculated by dividing the total artificial land cover in km^2, by the total land area of each region (km^2). Artificial land includes, not only buildings, but, also, roads, railways and all other types of build-up areas (Eurostat, 2018).

However, some of the values of the datasets that were retrieved from Eurostat, were missing. For this purpose, the missing values were retrieved from the National Statistical Authorities of the corresponding countries. In particular, data concerning the total of the population in 2010 were acquired from the databases of Switzerland and Germany (only for the regions "Chemnitz" and "Leipzig"). The population density of Germany (only for «Chemnitz" and "Leipzig") and Slovenia was found on the equivalent databases. The total of the male population between the ages 15 and 64 was not available for Germany's regions "Chemnitz" and "Leipzig" in the dataset available on the Eurostat website and, so, it was retrieved from Germany's Statistical Authority. Furthermore, data about the number of the people that had not received any education were not available for Denmark, Sweden and the United Kingdom, but they were found on each country's National Statistical Office websites (for the case of the United Kingdom, the databases of both England and Wales and Scotland were accessed). Additionally, for Liechtenstein, the number of the employed persons in the public sector, was acquired from the country's statistics. From Iceland and Switzerland's databases, data about the GDP per capita in 2010 were retrieved. The same way, the average disposable income of Iceland, Norway and Switzerland's regions was, also, retrieved. Finally, from the Statistical Authorities of Norway and Switzerland the area of each land cover category was found.

5. Methodology

Traditionally, geographers and scientists from other related fields, have compared patterns of different phenomena on maps, either by placing them side by side and observing them, or by overlaying them upon one another. With these techniques they have been able to examine them and to draw subjective conclusions. According to Abler et al. (1971), the main problem with this approach is that "the human eye is not always a very precise assessor of the strengths of spatial relationships and it can be an extremely misleading instrument" (p.120). To that end, the correlations that are suspected on theoretical and intuitive grounds should be examined thoroughly, in order to provide the necessary scientific evidence of their existence (Abler et al., 1971; Brunsdon et al., 1996). For this purpose, many different methods and techniques of data analysis have emerged. The enormous advances in the field of the information technologies, have contributed to this cause tremendously, by making the results more accurate and the analysts less aware of the technical procedures, but more aware of the results (Goodchild et al., 2000; Longley et al., 2005). Some of the authors whose papers have been reviewed in this dissertation, have used simple regressions (Pizam, 1982; Marselli and Vannini, 1997; Gould et al., 2002), Poisson regressions (Sherman et al., 1989; Carcach, 1999; Osgood, 2000), or have combined the application of both a simple regression and a geographically weighted regression (Alevizaki, 2010). A different regression method that has been used in some of the reviewed papers is the Ordinary Least Squares (Krivo and Peterson, 1996; Zhu et al., 2004; Altindag, 2012). In other papers, the authors use different methods that are provided through Geographic Information Systems (Zhong et al., 2011), the Exploratory Spatial Data Analysis (ESDA)[6] and the Confirmatory Spatial Data Analysis (CSDA)[7] (Cameron, 2001; Cracolici and Uberti, 2008) and the Normal Mixture Model[8] (Zakaria and Rahman, 2016). Some of the authors of the papers present different models that had been created for the analysis (Huang, 2004), statistical methods (Rufrancos, 2013), or even, theoretical approaches (Hirschi and Gottfredson, 1983). This dissertation uses different methods and techniques of spatial analysis, in order to process the used data and to draw conclusions about

[6] Exploratory Spatial Data Analysis (ESDA) is a process where data are viewed from many different view points, including from their display on maps (Fischer and Getis, 2010). The method is used to discover patterns of spatial association through examining forms of spatial instability or spatial non-stationarity (Longley et al., 2005).

[7] The tools of the Confirmatory Spatial Data Analysis (CSDA) group the quantitative processes of modeling, estimation and validation that are crucial for the analysis of different spatial components (Lopes et al., 2007).

[8] The Mixture Model is a probabilistic model that indicates the presence of subpopulations within a population.

the correlation between crimes and other socio-economic variables. The aim of this chapter is to present these methods and techniques in detail.

Spatial analysis has evolved into a highly significant field for the study of every social phenomenon, since the analysis of space and place has developed as a pivotal component of any social research (Goodchild et al., 2000). Its importance is magnified in the study of extremely crucial social problems with a spatial dimension, such as crime, which consists the main topic of this dissertation. Through spatial analysis, researchers are able to delve into a phenomenon and understand the reasons that cause it in the specific areas, where it occurs. Some researchers have defined spatial analysis as the study of the distribution of points, lines and polygons on a map. Nevertheless, this definition is quite vague. Other researchers have defined the method as the quantitative study of spatial phenomena. A more recent definition suggests that the method is the way of studying the behavior of different phenomena in space and their relationship with other spatial phenomena (Kalogirou, 2015). A different definition presents spatial analysis as "a well-defined subset of methods of analysis available to a project", or "a set of methods useful when the data are spatial, in other words when the data are referenced to a 2-dimensional frame", or "a subset of analytic techniques whose results depend on the frame, or will change if the frame changes, or if objects are repositioned within it" (Longley et al., 2005, p.569). According to Fotheringham and Rogerson (1994), spatial analysis is not a well-defined term, but it consists of a great variety of techniques and procedures of analysis. The main principle of spatial analysis, is Tobler's first law of Geography, which highlights that everything is related to everything else, but near things are more related than distant things (Tobler, 1970; Kalogirou, 2015). The verification or the rejection of this suggestion, is one of the primary objectives of spatial analysis. The other fundamental principles of spatial analysis come from the fields of mathematics, statistics and econometrics, since it requires the use of probability, statistics and econometric techniques (Fischer and Getis, 2010).

One of the first, if not the first, recorded applications of spatial analysis in the study of a social phenomenon, was the study of the distribution of cholera cases in London, during the outbreak of 1854. John Snow was the doctor that conducted the research and proved that the spread of cholera had been caused by a virus transmitted through water. In order to reach this conclusion, he created points on a map that indicated the areas where infected individuals lived. From this map, he noticed that the points were concentrated near a specific water source. Then he examined the water of this source and the water from a different source of an area with the same

social and economic conditions. From this examination he found out that the water source of the first area contained waste and the second was clean. He, also, noticed that the first area had a higher rate of mortality by cholera than the second one (8 times higher). So he concluded that contaminated water was the cause of the plague and not the air, as it had been speculated until then (Winkelstein, 2007; Tsatsaris, 2017).

In the 1950s, the quantitative revolution (or evolution) connected quantitative analysis (which is strongly linked with spatial analysis) with geography and new technological methods. During this decade, statistical and mathematical methods, as well as, theories from other fields, were included in geography. This period came to an end during the 1970s, when new researchers rejected quantitative geography and introduced radical geography. However, quantitative geography and spatial analysis, continued to develop, through new technological advances, such as the Geographical Information Systems (GIS) and Remote Sensing (Longley et al., 2005; Kalogirou, 2015). Nowadays, a newly emerged type of study requires the use of spatial analysis. This study is the big data analysis, whose importance is being magnified constantly, because of the technological characteristics of this era, which leave a great amount of online traces and information. These technological characteristics lead, not only to the creation of spatial data, but, also, to their immediate distribution to analysts or other users through digital databases (Agnew and Livingstone, 2011).

As mentioned previously, spatial analysis is closely related to statistical analysis, but it contains spatial information. This spatial information contains the geographical location where an incident has occurred, the adjacency of the studied area and its distance to other areas that might be of importance. This spatial dimension can transform the analysis and result in more accurate and efficient conclusions about the spatial causes of a phenomenon and its correlation to other spatial characteristics. This way spatial analysis can provide answers to some fundamental questions about “the location of human activities, the construction of social space, and the relationship between social space and physical environment” (Goodchild et al., 2000, p. 142). The application of spatial analysis methods in real data can take place using specialist software, such as commercial and open source GIS software and statistical packages, or programming languages, like R and Python that use libraries, such as lctools and PySAL, respectively (Fischer and Getis, 2010).

5.1. Descriptive Statistics

Descriptive Statistics constitute one of the two main categories of Statistics (the other one is Inferential Statistics). The main goal of this category is the presentation and the description of various statistical data, through tables and charts, as well as the presentation of the distribution of every studied variable. This statistical method leads, not only to a better understanding of the relationship between two or more variables, but, also, to the presentation of every variable's frequency of emergence. The most common measures used for the development of Descriptive Statistics are the Measures of Central Tendency and the Measures of Dispersion.

The Measures of Central Tendency attempt to examine the tendency of the values to cluster around some central value. These measures include the arithmetic mean, the median and the mode. Arithmetic mean is called the summation of all the values of the observations divided by the number of the observations of the particular data set. This measure is the most commonly used statistical measure. In a dataset containing the values x_1, x_2,..., x_n (n is the number of the values), the arithmetic mean $\bar{X}$ is defined by the following formula:

$$\bar{X} = \frac{1}{n}\sum_{i}^{n} x_i$$

Median is called the middle value, which separates the values in half so that half of them are greater than the median and the other half are smaller. Basically the median is the value that is positioned in an array or distribution after the 50% of all the values. Mode is called the most frequent value in a data set. Other Measures of Central Tendency, that are not used often, are the geometric mean, the harmonic mean, the simplicial depth, the geometric median etc.

The Measures of Dispersion examine how much the values of an array or a distribution are dispersed left and right of some central value (Zairis, 2010). These measures include the range, the interquartile range, the variance, the standard deviation and the coefficient of variation (CV). Range is the difference between the maximum and the minimum value of the data set and it is calculated by the formula: $R = X_{max} - X_{min}$. The interquartile range is the part of the array or the distribution which contains the 50% of the observations. This measure is estimated as the difference between the 25th and the 75th quartile and particularly using the formula:

$$Q = \frac{Q_3 - Q_1}{2}$$

Variance is defined as the mean of the squared deviations of the values from their arithmetic

mean. In a dataset of the values X_1, X_2,…, X_N, (N is the number of the values) the variance is defined by the following formula:

$$\sigma^2 = \frac{\sum(x_i - \bar{X})^2}{N} = \frac{\sum X_i^2 - N\,\bar{X}^2}{N} = \frac{\sum X_i^2}{N} - \bar{X}^2$$

Standard deviation is the positive value of the variance's square root, $\sigma=\sqrt{\sigma^2}$. The coefficient of variation expresses the standard variation as a percentage of the arithmetic mean. Additionally, it presents the extent of variability in relation to the mean of the population. The coefficient of variation can be defined by the formula:

$$CV = \frac{\sigma}{\bar{X}}$$

The results of the Descriptive Statistics are presented both in a numeric form through tables and in a graphic form through plots. The most commonly used plots are histograms and boxplots. Histogram is a plot that presents the frequency of the appearance of a variable's values, or the distribution of these values. The histogram can indicate if the variable has a normal distribution or not, because of the shape of the line that is created. A boxplot (or whisker plot) is a plot that presents graphically the maximum and the minimum value of the variable, as well as the arithmetic mean, the values of the 1st and the 3rd quarter and the outliers (Zairis, 2010; Symeonaki, 2015). Different types of data are presented in different plots. The choice of the suitable plot depends on the nature of the data, on what they represent and on which one of their aspects need to be highlighted.

5.2. Spatial Autocorrelation

Spatial patterns can cause several problems, such as dependence and heterogeneity, which render some statistical methods unstable (Lopes et al., 2007). The Exploratory Spatial Data Analysis (ESDA), is a spatial analysis tool that allows the data to be viewed from many different view points, including from their display on maps (Fischer and Getis, 2010). The method is used to identify patterns of spatial association (clusters) and atypical observations (extreme values), through examining forms of spatial instability or spatial non-stationarity (Longley et al., 2005; Lopes et al., 2007).

One of the most important functions of the exploratory spatial data analysis is spatial autocorrelation. Spatial autocorrelation is called the correlation between the values of a variable

in neighboring areas. This correlation is caused completely by the proximity of these values in space (Kalogirou, 2015). Basically, spatial autocorrelation measures how much the values of two neighboring areas resemble each other or not. A different definition of the term suggests that spatial autocorrelation depicts the relationship between spatial units (each unit corresponds to the value of the studied variable) that are located near each other, on the maps (Fotheringham and Rogerson, 2009; Fischer and Getis, 2010). The concept of spatial autocorrelation was developed in the late 1950s, at the University of Washington, by the geographers Michael Dacey, William Garrison and Edward Ullman. In geography, the concept was developed in a statistical framework in 1969, by Cliff and Ord, in the paper "The problems of Spatial autocorrelation". Cliff and Ord, also, developed the most popular index that examines the presence of spatial autocorrelation, the Moran's I, in 1973 (Fischer and Getis, 2010). The Moran's index, which focuses on each observation and not on the mean of all observations (Fischer and Getis, 2010), can be both global and local. The first version of the formula that is used for the calculation of the Global Moran's I is:

$$\mathrm{I} = \frac{\mathrm{n}}{2\mathrm{A}} \frac{\sum_{\mathrm{i}}^{\mathrm{n}} \sum_{\mathrm{j}}^{\mathrm{n}} \mathrm{w_{ij}} (\mathrm{x_i} - \overline{\mathrm{x}})(\mathrm{x_j} - \overline{\mathrm{x}})}{\sum_{\mathrm{i}}^{\mathrm{n}} (\mathrm{x_i} - \overline{\mathrm{x}})^2}$$

where n is the number of the observations, $\overline{x}$ is the arithmetic mean of the $\mathrm{x_i}$ observations, A is the total number of joins in the system and $\mathrm{w_{ij}}$ are the weights that are calculated according to the spatial proximity of the observations (Kalogirou, 2001; Kalogirou, 2015). The formula that is currently used for the calculation of the Moran's index is:

$$I = \frac{n \sum_i^n \sum_j^n w_{ij} (x_i - \overline{x})(x_j - \overline{x})}{(\sum_i^n \sum_j^n w_{ij}) \sum_i^n (x_i - \overline{x})^2}$$

(Kalogirou, 2001; Fischer and Getis, 2010; Kalogirou, 2015).

The Global Moran's I can be in a range from -1 to 1. An index close to -1 indicates a strong negative autocorrelation, namely, areas with high values of the variable tend to be near areas with low values. On the other hand, an index close to 1 indicates a strong positive autocorrelation. That means that areas with high values of the variable tend to be located near areas with similarly high values and areas with low values tend to be near areas that, also, have low values. If the index is nearly or precisely 0, it indicates that there is no spatial autocorrelation and hence no spatial patterns (Kalogirou, 2015).

The Local Moran's I is interpreted in the same way as the Global, but it focuses more on the relationship between neighboring areas and its values are not necessarily between the range -1 to 1. A positive index indicates a spatial concentration of similar values (high or low) and a negative index indicates a spatial concentration of different values (low with high). Through the local Moran's I, it is possible to create a thematic map, in which every spatial entity is classified in one of the following categories:

- High – High: A spatial entity with high values of a studied variable that neighbors a spatial entity with high values of the studied variable.
- Low – Low: A spatial entity with low values that neighbors a spatial entity with low values of the variable.
- Low – High: A spatial entity with low values that neighbors a spatial entity with high values.
- High – Low: A spatial entity with high values that neighbors an entity with low values.
- Statistically not significant local Moran's I (Kalogirou, 2015).

The formula that leads to the calculation of the Local Moran's Index is the following:

$$I_i = \frac{x_i - \overline{x}}{m_2} \sum_{j=1}^{k} w_{ij}(x_j - \overline{x}), \qquad j \neq i$$

where m_2 is a stable value for every local I_i of a variable and equals $\sum_{i=1}^{n}(x_i - \overline{x})^2 / n$ (Fischer and Getis, 2010; Kalogirou, 2015).

5.3. Inequality Index

Inequality indexes are increasingly a part of spatial analysis, since they contribute importantly in the measurement of inequalities among spatial data. Inequality measures are measures of dispersion that present the distribution of a value (Rey and Smith, 2012). The most popular index that measures inequality, is the Gini index, which can be thought of as the ratio of the area of the inequality distribution and the area of the triangle under the Lorentz curve (Zairis, 2010). The value of this index can be in a range from 0 to 1. If the index is equal to 0, there is a perfect equality in the variable, whereas, if the value of the index is 1, there is maximum inequality. The formula that is used for the calculation of the Gini coefficient is:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\bar{x}}$$

where x_i is the value of the variable x that is observed in the location $i = [1, 2, \dots, n]$ and $\bar{x}$ is the arithmetic mean (Rey and Smith, 2012).

The study of inequality presents several serious problems, like the checkerboard problem and the Modifiable Areal Unit Problem (MAUP). The first problem, refers to the cluster of units when they are near each other in a single part of an area. However, these units do not seem to cluster when they are distributed like a checkerboard. The index of dissimilarity is not sensitive when the units are evenly distributed. Spatial autocorrelation measures the degree of clustering, but not the degree on which the distribution is uneven. The modifiable areal unit problem, which was described in 1984 by Openshaw in detail, refers to the fact that a sudden change of the boundary, can change a measure if the distribution of the units is heterogeneous. Except for this, the scale or the spatial precision of the units can, also, change a measure. These problems created the need for the development of a decomposition of the Gini. The aim of this decomposition is to compare the spatial patterns of the clustered units across areas with different levels of assemblage. It should be noted, that without a decomposition, analysts could calculate the Gini for each sub-area, but a reduction in the inequality of one of them may not lower the Gini for the total area. In 1989, Silber created a decomposition of the Gini using matrix algebra that allows the identification of the differences that exist within and between the clusters. In 2004, Dawkins, developed a spatial Gini decomposition measure, which is based on the splitting of the population of the units into mutually exclusive subgroups and several spatial scales (Rey and Smith, 2012).

The statistical significance of the Gini index can be evaluated through the Monte Carlo simulation. In this simulation the data are redistributed spatially in a random way. Then the inequality of non-neighbors is estimated (both for the real and the simulated data) and lastly, the p value is calculated. If the p value is equal or lower than 0.05, the components of the index (that concern both the inequality of the neighbors and the inequality of the non-neighbors) are statistically important. In order for this simulation to be efficient, there are required at least 19 repetitions of the method, but usually more than 99 are recommended (Kalogirou, 2015), for a more precise result. The equation that defines the Monte Carlo simulation is the following:

$$SG = \frac{\sum_{j=1}^{n}\left(1 - w_{i,j}\right)|x_i - x_j|}{2n^2\bar{x}G}$$

In this equation G is the Gini and SG is the share of overall inequality that is associated with non-neighboring pairs of locations (Rey and Smith, 2012). The Gini index and the Monte Carlo simulation, can be calculated in the open source programming language R through the package lctools.

5.4. Regression

One of the core techniques of the quantitative revolution in geography was the development of regression (Fotheringham et al., 2000; Fotheringham and Rogerson, 2009). Regression is frequently used in spatial analysis, due to the fact that it can help the comprehension of how specific factors affect different phenomena. Additionally, it can contribute in the quantification of the cause-effect relation. Moreover, it allows the analysts to calibrate a model and draw empirical conclusions about the explanation of the studied phenomenon. Regression is called the process that estimates the type and the degree of the relationship between one dependent variable and one or more independent variables (Rogerson, 2001; De Smith et al., 2007). Dependent variable (or the response variable or the regressand) is called the one that is being studied and independent (or the predictor variables or the regressors) the one(s) that may or may not explain the values of the dependent (Charlton and Fotheringham, 2009). In this technique the dependent variable is modeled as a linear function of the set of the independent variables (Brunsdon et al., 1996). This model can be considered as a simplification of reality. The importance of regression analysis lays on the extended information that it provides. Particularly, it presents not only a simplified view of the relationships between the variables, but, also, it provides a way of fitting the model with the used dataset. Moreover, it provides a way for the evaluation of the importance of the studied variables and for the correction of the model (Rogerson, 2001). An interest for the creation of this technique started to develop in the early 1950s, even though it was introduced in quantitative geography in the early 1970s by Cliff and Ord and later by Upton and Fingleton (Fotheringham and Rogerson, 2009).

There are several types of regression that are being used depending on the phenomenon that is being studied. The most popular types of regression, which are included in the Generalized Linear Model, are:

- Linear/Gaussian Regression. It is used when the dependent variable is a ratio and follows a normal distribution. The assumption that a linear relationship between the dependent variable and the independent variables already exists, is the main principle of linear regression that allows the analysis to occur. The next step is the fitting of a straight line to the dataset and the following is the interpretation and the analysis of the effects of the independent variables on the dependent variable (Rogerson, 2001). Generally, the assumptions of regression analysis for simple regression are: (1) the relationship between the dependent variable and the independent variable(s) is linear, (2) the errors have a mean that equals zero and a constant variance, (3) the residuals are independent (the value of one error is not affected by the value of another one), (4) for each x value the errors have a normal distribution (Rogerson, 2001; Kalogirou, 2015).

  The simple linear regression, which consists of only one independent variable, can be calculated by the formula:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

  in which y is the dependent variable, x is the independent variable, β_0 is the constant, β_1 is the parameter that shows the relationship between y and x and ε is the error term (Fotheringham and Brunsdon, 1999; Kalogirou, 2001; Rogerson, 2001; Charlton and Fotheringham, 2009).

  In the case of more than one independent variables, the formula that calculates the multiple linear regression is the following:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon$$

  where y is the dependent value, x_1, x_2, x_n are the n independent values, β_0 is the constant, β_1, β_2, …, β_n are the parameters that show the relationship between the dependent value and the corresponding independent value and ε is the error term (Brunsdon et al, 1996; Fotheringham and Brunsdon, 1999; Rogerson, 2001; De Smith et al., 2007; Charlton and Fotheringham, 2009; Fotheringham and Rogerson, 2009).

  In both these formulas, the parameters β_1, β_2, …, β_n can be calculated with the formula:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

  where $\hat{\beta}$ is the vector of the estimated parameters, X is the matrix that contains the values of the independent variables and one column of 1s, y is the vector of the observed values and $(X^T X)^{-1}$ is the inverse of the variance-covariance matrix. In regression, it is common to weight the observations. In this case the previous formula would be written as:

$$\hat{\beta} = (X^T W X)^{-1} X^T W y$$

(Brunsdon et al, 1996; De Smith et al., 2007; Charlton and Fotheringham, 2009). The error term ε can be calculated by the subtraction of the dependent variable from the evaluation of the trend surface at each point i: $\varepsilon_i = \widehat{y_i} - y_i$ (De Smith et al., 2007).

- Poisson Regression. It is used when the dependent variable is a number of rare events and follows a Poisson distribution. This difference in the distribution is the only real difference between the Poisson regression and the linear regression. The Poisson probability distribution with a parameter i is given by the following formula:

$$p_Y(y;\mu) = pr(Y = Y;\mu) = \frac{\mu^y e^{-\mu}}{y!}$$

in which μ is the dependent variable.

The Poisson regression analysis is a maximum likelihood procedure. Its likelihood function is notably more complex than the linear regression's function. One form of this function is the following:

$$L(y;\beta) = \prod_{i=1}^{n} p_{Y_i}(y_i;\beta) = \prod_{i=1}^{n} \left\{ \frac{[l_i\lambda(x_i,\beta)]^{y_i} e^{-l_i\lambda(x_i,\beta)}}{y_i!} \right\}$$
$$= \frac{\{\prod_{i=1}^{n}[l_i\lambda(x_i,\beta)]^{y_i}\} \exp[-\sum_{i=1}^{n} l_i\lambda(x_i,\beta)]}{\prod_{i=1}^{n} y_i!}$$

where $l_i\lambda(x_i,\beta)$ is the dependent variable (Kalogirou, 2001).

- Logistic Regression. It is used when the dependent variable consists of only two values (yes/no translated as 1/0 in the data set) and follows a binomial distribution. Typically, the data are transformed with logit and are used for the conduction of linear regression (De Smith et al., 2007; Kalogirou, 2015). The formula of the logit model that has as a dependent variable the likelihood p_i is:

$$logit(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = a_0 + \sum_{1 \le k \le N} a_k x_{ki} + \varepsilon_i$$

In the formula above, p_i is the likelihood that the dependent variable has the value 1 in the point i, x_{ki} is the value of one of the k independent variables in the point i, a_0 is the constant, a_k is the k parameter of the variable x_k and ε_i is the error (Kalogirou, 2015).

Since the data used in this dissertation correspond to ratios and real numbers, the regression method, which will be used is the multiple linear regression.

One of the most useful indexes that can be retrieved through regression is the correlation coefficient R^2, which shows the percentage of the explanation for the dependent variable that is provided by the independent variables. In other words, this index shows what percentage of the dependent value is explained by the independent. The range of this index is from 0 to 1. If the value of the index is close to 1, the variance of the dependent variable can be predicted perfectly by the independent variable(s). If the value of the index is near 0, there is no linear correlation between the dependent and the independent variable(s). However, the R^2 can be easily increased by adding variables, so the adjusted R^2 is often reported instead. The adjusted R^2 takes into account the number of the independent variables of the model (Charlton and Fotheringham, 2009).

It should be noted that before the calculation of any type of regression, it is essential to calculate the correlation coefficient. This coefficient indicates whether the independent variables are independent from each other. In order to accept any independent variable into the regression model, it has to be verified that they do not present a strong positive or negative correlation with any others. To that end, when two variables have a correlation coefficient that is higher than 0.8 or lower than -0.8, they should be examined further, so that one of them can be removed. This examination happens by calculating the correlation coefficient of both of them with the dependent variable. The independent variable that has the weakest (of the two) correlation (negative or positive) with the dependent variable is removed from the model. There are two methods for calculating the correlation coefficient that are more commonly used. The first one is the Pearson correlation coefficient, which is used when all of the variables have a normal distribution. The formula for this coefficient is:

$$r_i = \frac{{}^1\!/_n \sum_j w_{ij} (x_j - x^*)(y_i - y^*)}{[{}^1\!/_n \sum_j w_{ij} (x_j - x^*)^2]^{1/2} [{}^1\!/_n \sum_j w_{ij} (y_j - y^*)^2]^{1/2}}$$

where r_i is the local correlation coefficient for the point i, x and y are two variables with means x^* and y^* and w_{ij} is the weight of the data at a point j for the calculation of r at a point i (Fotheringham and Brunsdon, 1999). The second commonly used correlation coefficient is the Spearman coefficient, which is chosen when at least one of the studied variables does not have a normal distribution. The formula that is used for this method is:

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

where r is the correlation coefficient, d_i is the difference between the ranks of the variables and n is the number of the observations (Zar, 2005).

5.5. Geographically Weighted Regression

Geographically Weighted Regression (GWR) is a local method, which analyses and presents the spatial modifications of the correlations between variables. The method is being used for the study of the presence (or not) of spatial stability in the correlations, which is always considered existent in classical regression methods. The term was introduced by Brunsdon, Fotheringham and Charlton in 1996, to describe a group of regression models in which the coefficients are allowed to vary spatially. This means that the parameters can be estimated anywhere in the study area, as long as there is a dependent variable and at least one independent variable that have been measured at places whose location is known (De Smith et al., 2007; Charlton and Fotheringam, 2009). According to Brunsdon, Fotheringham and Charlton, GWR is a simple technique that allows the coefficients of the model to have a variance, so that they are specific for every spatial unit that is studied (1996). GWR was built on a non-parametric technique of locally weighted regression. The innovative part of this technique was that it uses a dataset proximate to the model calibration in geographical space, rather than in variable space (Fischer and Getis, 2010).

Along with the theoretical formulation of the technique, a suitable software was, also, developed. The first version of the software, called GWR, was an application that could run on a Unix environment. The next versions started to become more user friendly and could run on a Windows environment (versions GWR 1.0, GWR 2.x and GWR 3.x). In 2015, the version GWR 4.08 was developed to support the Geographically Weighted Generalized Linear Model (Kalogirou, 2015).
The equation that allows the calculation of the GWR is the following:

$$y_i = \beta_0 + \sum_{k=1}^{m} \beta x_{ik} + \varepsilon_i$$

where y_i is the value of the dependent variable at the location i, x_{ik} is the value at the kth covariate at the location i, β_0 is the intercept parameter, β_k is the local regression coefficient for the kth independent variable, m is the number of the independent variables and ε_i is the random error at the location i. The parameters β_k can be calculated with the formula: $\beta = (X^T X)^{-1} X_y^T$, where β represents a vector of global parameters that will be estimated, X is a matrix that consists

of the independent variables and whose elements of the first column are equal to 1 and y represents a vector of observations on the dependent value (Brunsdon et al, 1996; Fotheringham et al., 1998; Fotheringham and Brunsdon, 1999; Kalogirou,2001; De Smith et al., 2007; Charlton and Fotheringham, 2009; Fotheringham and Rogerson, 2009; Fischer and Getis, 2010; Lu et al., 2012; Kalogirou, 2015).

Since GWR allows the calibration of a local model, by calibrating the model around a point i in space, including some (or all) of the m in the dataset weighted by a weighting system, the formulas that can be used instead of the two previous ones are the following:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{m} a_k(u_i, v_i)x_{ik} + \varepsilon_i$$

$$\beta(u_i, v_i) = (X^T W(u_i, v_i)X)^{-1} X^T W(u_i, v_i)y$$

where $W(u_i, v_i)$ is an m by m matrix whose off-diagonal elements are zeros and whose diagonal elements display the geographical weighting of the observed data for the point i (Kalogirou, 2001; De Smith et al., 2007; Fotheringham and Rogerson, 2009; Fischer and Getis, 2010; Lu et al., 2012; Kalogirou, 2015).

The weighting system $W(u_i, v_i)$ is calculated with a kernel function that is based on the proximities between the regression point i and the data points that are located around it (Lu et al., 2012). There are two main categories of kernel, the fixed kernel and the adaptive. The most crucial component of a kernel is its bandwidth, which determines the radius around the point i. This radius defines the area around that point, where the observations will be weighted and therefor included in the regression. In a fixed kernel, this radius is constant, whereas in the case of an adaptive kernel the analysts need to define the number of the nearest neighbours. This number is basically the number of the observations around the point i that determine the radius and will be included in the regression (Kalogirou, 2001; Charlton and Fotheringham, 2009; Lu et al., 2012; Kalogirou, 2015). The choice of the bandwidth impacts in a great scale the results of the GWR. Generally, it can be considered a smoothing parameter of a model. A model that has been over-smoothed will produce parameters that have a similar value across the study area. A model that has been under-smoothed, will produce parameters that will have a lot of local variation, which will make the decision of whether there are any patterns extremely challenging. The best bandwidth is the one that lies in between these two extreme cases (Kalogirou, 2001).

One of the most important indexes that are produced from the GWR is the index AICc (Akaike Information Criterion), which indicates how well the model adapts to the data and hence it shows if we can accept or decline the model. A smaller AICc value represents a better model. The AICc is strongly related to the kernel function, since the selection of the appropriate bandwidth minimizes it. In other words, the AICc is a technique that is able to calculate the ideal bandwidth for the GWR model (Kalogirou, 2001). The equation that leads to the calculation of the AICc is the following:

$$AICc = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n\left(\frac{n + tr(S)}{n - 2 - tr(S)}\right)$$

where n is the number of the observations in the dataset, $\hat{\sigma}$ is the estimate of the standard deviation of the residuals and $tr(S)$ is the trace of the hat matrix (Kalogirou, 2001; Charlton and Fotheringham, 2009; Lu et al., 2012).

In the GWR, the R^2 is, also, an important correlation coefficient, due to the fact that it measures how much of the dependent variable can be explained by the independent one(s). Its range is the same in this type of regression, as well (from 0 to 1). The formula that calculates the R^2 is:

$$R^2 = \frac{\sum_{i-1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where n is the number of the observations, y_i is the observed values of the dependent variable y, $\bar{y}$ is the mean of y and $\hat{y}_i$ is the estimated values of y (Kalogirou, 2001).

5.6. Cartography

Since ancient times, people have been creating and using maps in order to navigate and examine their surroundings. Over time, cartography has evolved into a field of great importance for every type of spatial analysis. Cartography is defined as both the study of maps and the art, science and technology of creating maps. During the last decades, the field has acquired more dimensions, since it does not depend solely on the depiction of the geographic characteristics of any area, but it can provide a variety of information and solutions as well. This evolution of cartography can be attributed to the development of the Geographic Information Systems (GIS), which are used for the insertion, the retrieval, the handling and the output of spatial data (Robinson et al., 2002; Chalkias, 2011).

The aim of cartography is the production of a map. Maps can be defined as depictions of earth in an image and are able to present different phenomena and relationships that occur in any space. The basic attributes of every map are the scale, the projection and the symbolization that they use, which indicate the map's possibilities and limitations (Monmonier, 1996). Cartography uses different symbols and other particular coded graphics to represent different information, which are explained in the legend of the produced map (Robinson et al., 2002; Sidiropoulos, 2006). Maps are, also, created in a specific scale. Scale is the ratio of a distance on a map to the corresponding distance on the ground. It is used to make maps more easily transported and to provide the ability to map large areas, or even the entire Earth (Robinson et al., 2002).

Every type of map that can be created, should have a coordinate system, which defines the exact location of every spot on the map, either through a number or through the latitude and the longitude. Latitude is called the angle that is created between a vertical line, from the studied spot, and the equator. Longitude is called the angle that is created between the meridian of the studied spot and the Prime meridian. The system that uses the latitude and the longitude is the Geographic Coordinate System. This system is based on the ellipsoid shape of the Earth. A different system is the Projected Coordinate System, which presents the ellipsoid on a flat surface. In this system, latitude and longitude are converted in x and y in a two-dimensional level (Longley et al., 2010).

One of the most common types of maps that can be created is the thematic map. A thematic map is used to depict the spatial distribution of natural or human-made phenomena. The most popular thematic map is the choropleth, which presents a phenomenon, with the utilization of different area symbols, such as shading or colors. These area symbols represent the categorized classes of the examined phenomenon (Robinson et al., 2002; De Smith et al., 2007). The main elements of each thematic map are the size and the shape of the geographical areas, the color palette that is used, the number of the categories in which the values of the variable are classified and the method used for the definition of the limits of these categories (such as equal intervals etc.) (Kalogirou, 2015).

5.7. Programming Language R

In the last decade there has been a significant increase to the available spatial data. The main reason for this augmentation is the technological evolution, and particularly, the extensive use of

the Internet. Almost every individual of the world has access to the worldwide web and uses it daily. This online presence leaves online spatial traces. On the other hand, due to this easy access to the web, many spatial data are provided online, for everyone to obtain. This abundance of spatial data has created the need for the development of appropriate software that can edit and examine them. R is a popular open source programming language which provides a great variety of capabilities, regarding the spatial analysis of any data set. R's website, offers a wide range of information about the language, as well as useful open access manuals that are available for new and experienced users and programmers (The R Project for Statistical Computing, 2018). The basic software of the language is quite limited but it can be expanded through a considerable amount of additional packages that can be downloaded and used. These packages offer many different capabilities, including mapping, and can be downloaded either from the R environment, or from the network of servers CRAN (The R Project for Statistical Computing, 2018).

The R language comes from the S programming language, which was created in 1976 by John Chambers. A few years later, in 1995, Ross Ihaka and Robert Gentleman, altered the S language, due to some issues that they had been experiencing, and developed the language R. They named the language R, in order to acknowledge the influence of S and to celebrate their own efforts as well (both of their names start with the letter R) (Ihaka, Gentleman, 1996). In 1995 they released the initial version of the language and in 2000 a beta version. The R language is an open source software, which means that any user is able to download and use it without any subscriptions.

RStudio is an integrated development environment (IDE) for R, which was released in 2011 by JJ Allaire. RStudio is available in both an open source and in a commercial edition and it can run on a desktop or in a browser that is connected to the RStudio server. The main goal of the development of RStudio was not only, to make learning R easier for new users, but, also, to benefit more advance users, through providing them with high productivity tools (Allaire, 2011).

The creation of RStudio made easier the development of multiple additional packages. Some packages allow the processing of a large amount of data (for example the package "dplyr"), while others make reading many types of files possible (like the package "readr"). Additionally, there are several packages that allow the creation of thematic maps and many other packages that allow the processing of spatial information (for example the packages "sp", "spdep", "GISTools", "lctools", "rgdal" etc.). R offers the ability to select the appropriate packages for the analysis of spatial data, through a view with different spatial topics (CRAN Task View: Analysis of Spatial Data).

Moreover, there are packages that allow the generation of dynamic reports with R, like the package "knitr" (R Studio, 2018).

5.8. Package "Shiny"

One of the most unique packages that can be downloaded and used in R, is the package "Shiny". This package was released in 2017 by Cheng and many other authors (including Allaire, the creator of RStudio) and it is used in order to build interactive web applications with R (The R Project for Statistical Computing, 2018). What makes these applications interactive, is the ability to create elements that react to the users' selections and produce different outputs according to them. For this purpose, Shiny offers a wide range of script lines, which correspond to different widgets, such as action buttons, select boxes, sliders etc. (Shiny from R Studio, 2018).

The applications that are created can be hosted either on a webpage or in an R Markdown document. Since Shiny applications are fundamentally webpages, they are usually shared over the web. For this reason, they need to be hosted on a server. Shiny offers two ways to make this possible. The first one is the ability of the users to create their own server, through an open source and free back end program, which is called Shiny Server. This program builds a Linux web server that is specifically designed to host Shiny applications. The second way to share the applications online, is through the server that is maintained by RStudio. This server can be reached by creating an account on shinyapps.io. This server is free, secure, easy to use and scalable, which means that when many people visit the application at the same time, the server handles each person individually, so that they do not experience a slowdown (Shiny from R Studio, 2018). The application that is created for this dissertation, is hosted on the shinyapps.io server on the URL: https://artemistsiopa.shinyapps.io/shiny/.

The script that is used for the creation of applications on Shiny, does not differ tremendously from the typical scripts of R. The main difference, is that the script consists of three main components, the User Interface section (UI), the server function and a call to the shinyApp function. The UI object defines the layout and the appearance of the application. Anything written inside this section

```
library(shiny)
ui<-fluidPage()

server<-function(input, output){}

shinyAp(ui=ui, server=server)
```

Figure 6. *Template for Shiny applications.*

will appear on the map. The server function contains the entire code; all the instructions that the computer needs in order to run the desirable functions. The call to the shinyApp function knits the other two components (UI and server) together to build the application (Shiny from R Studio, 2018).

The application and every other file that it relies upon (such as images), must be saved in the same directory. The main script must strictly be named “app.R”, because the Shiny server will search in the directory for a script with this exact name in order to build an application. Originally, the application code had to be divided into two scripts. The first one would be strictly named “ui.R” and would contain the UI section of the script and the second one would be the “server.R”, which would contain the server section of the application. The usage of these two separate script files is still possible. In case that the Shiny server does not find an “app.R” file in the directory, it will search for a “ui.R” and a “server.R” file. If the creators want to display in their application objects that they already have on their computer, for example images, they have to place them in a folder, strictly named “www”. This folder must be created in the same directory as the script files (Shiny from R Studio, 2018).

The basic elements of a Shiny application are the inputs and the outputs. Inputs are the objects that the users can toggle and use in order to provide values for the application (such as dropdown menus etc.). The Shiny package provides many input functions, such as “radioButtons()”, “selectInput()”, “sliderInput()”, etc., which, basically, include every type of interaction that users can make with a mouse click on a webpage. Outputs are the objects that are displayed on the application (like plots, images, texts etc.) and can respond to any selections the users make on the input objects. The output functions are based on the type of objects that will be displayed, for example some output functions are the “plotOutput()”, the “textOutput()”, the “imageOutput()” etc. Both of these families of functions must be added to the UI section. Their basic syntax is quite similar, since the first argument for the both of them should be the “InputId” and the “OutputId” respectively. In these arguments, every input and output object must be designated with its own unique name, so that no errors occur when the server runs the code. These functions can, also, have other arguments, like a label argument, which will display a title above the input or output object, a height argument and a width argument that will determine the height and the width of the displayed input or output object etc. (Shiny from R Studio, 2018).

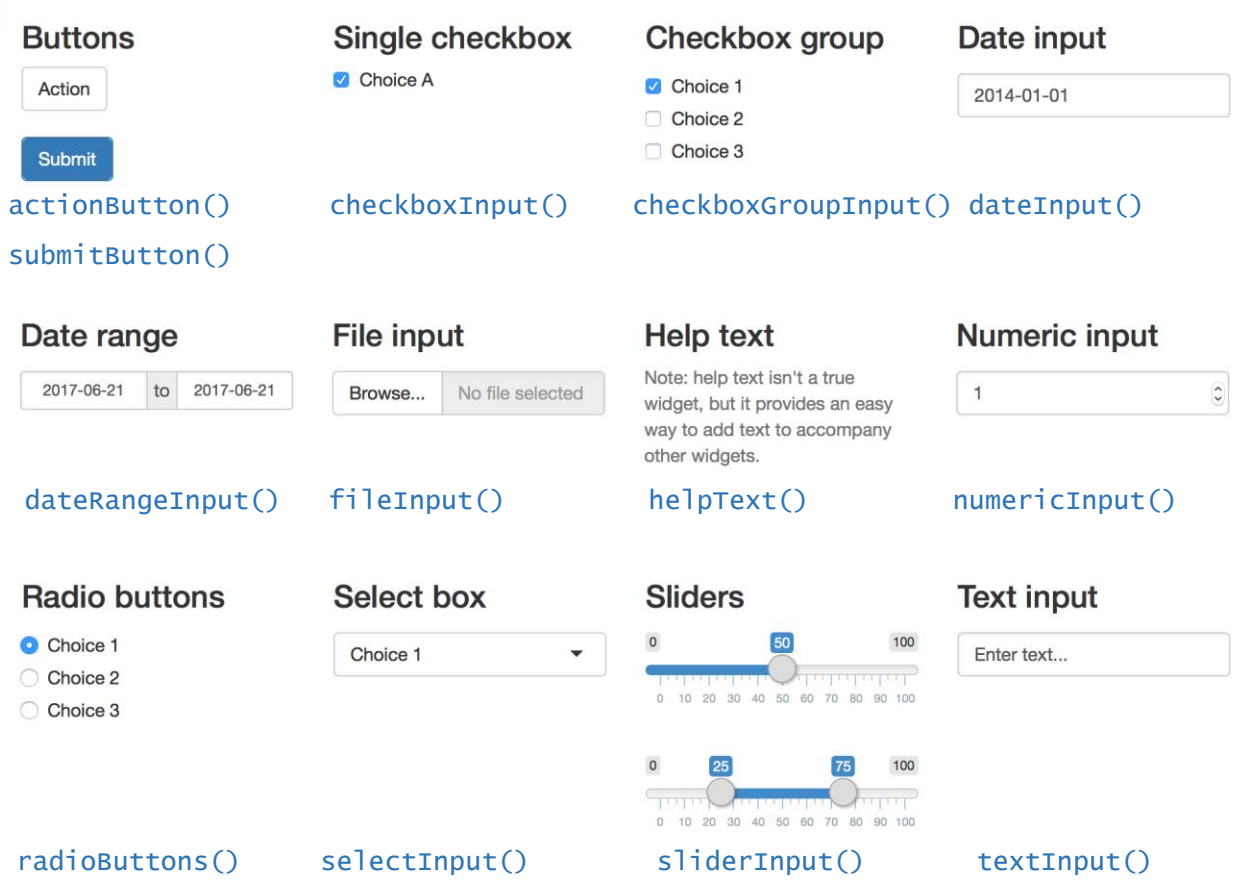

Figure 7. *Input functions.*

Source: Shiny from R Studio, 2018

In order for the code to display any elements that the users want to appear in the application, the output objects that are defined in the UI section, must be connected with the outputs that are created from the code in the server section. This can be achieved with the usage of the render functions, which belong to the reactive family of functions. To use a render function, it is important to save it by writing “output$” and the name (“OutputId”) of the output that was created in the UI section, because that is the way to get the results into the application. The arguments of a render function should begin with a curly bracket and end with a curly bracket ({}). Between these curly brackets there should be the argument that will be rendered to the output. This argument can be one word, if the object that will be displayed has already been created and named in the server section, or an entire chunk of code that will create it. The type of the render function that will be used is determined by the type of output that will be rendered. To give an illustration, if the output object is an image, the render function that will be used is the “renderImage()” function etc. (Shiny from R Studio, 2018).

The applications created with the Shiny package, can be customized in many ways. The creators can add, except for the multiple input and output objects, different features to the applications. They can create applications with multiple tabs, which can display different objects and have completely different characteristics, or multiple panels that can, also, host different inputs and outputs. In addition, creators can change the graphics and colors of the applications that they create. This can happen with two possible ways. The first way is by adding a Cascading Style Sheets (CSS) file, which is a file that contains code written in HTML. This code consists of functions that can change the appearance of webpages and particularly, the colors, the background colors, the fonts, the font sizes and types and generally, allows the customization of everything that appears on a webpage. Creators can write their own CSS files and add them with the function "includeCSS()" that can be written in the UI section. In order for this function to work, the CSS file should be found in the "www" folder, which exists in the same directory as the script (Shiny from R Studio, 2018). The second way is by adding a theme using a code that has already been written and is available through the "shinythemes" package. This package offers 16 different themes, that can be chosen with the function "theme=shinytheme()", which should be placed in the UI section. The argument of this function is the name of the theme that the creator wishes to use, in quotation marks, for example "sandstone", "cerulean", "cyborg" etc. (R Studio, 2018).

6. Results

Crime is an increasingly serious and costly problem that concerns every society, and especially developing societies and societies in transition (Zhong et al., 2011). The concern with crime is well justified due to the fact that it effects economic activities, as well as the quality of life of the people who live in high crime areas. These people are obliged to cope with a significantly reduced sense of personal and proprietary security (Fajnzylber et al., 2002). Scientists use a variety of methods, in order to measure the costs of crime to societies and emphasize on the benefits of crime reduction and prevention (International Centre for the Prevention of Crime (ICPC), 2010). Nevertheless, the study of crime can still be considered inadequate (Zhong et al., 2011). The study of the geographic distribution and the determinants of crime has interested scientists from different fields over the years and it has developed tremendously (Leitner, 2013). Criminologists and sociologists believe that crime is highly affected by demographic and socio-economic conditions. Geographers share that belief, but, also, emphasize on the geographic dimension of crime (Zhang and Peterson, 2007; Zhong et al., 2011). From developing the field of spatial analysis, they have observed that crime is disproportionately distributed across different areas. Additionally, they suggest that crime is exceptionally concentrated in few, small areas that consist the crime hot spots (Zhang and Peterson, 2007). The introduction of spatial analysis is considered fundamental to the crime examination, since it addresses the problem of aspatiality in criminological research (Ratcliffe, 2010). This dissertation, uses methods and techniques of spatial analysis, in order to examine these crime theories. The aim of this chapter is to present and to explicate the results of the used methods and techniques.

6.1. Maps

The study area of this analysis, is the European Union and particularly, its regions, on a NUTS 2 level. The examined phenomenon is the crime rates in these regions during the year 2010. The factors whose contribution to crime rates is studied are the population density, the ratio of the males in the age group 15-64 in the total population, the ratio of the immigrants in the total population, the ratio of the persons that had received no education, the ratio of the unemployed persons, the ratio of the employed individuals in the public sector, the GDP per capita, the average

disposable income and the ratio of the artificial land cover in the total area of each NUTS 2 region. Since the spatial information is fundamental for spatial analysis, it is essential to display the phenomenon, which is being studied, on maps.

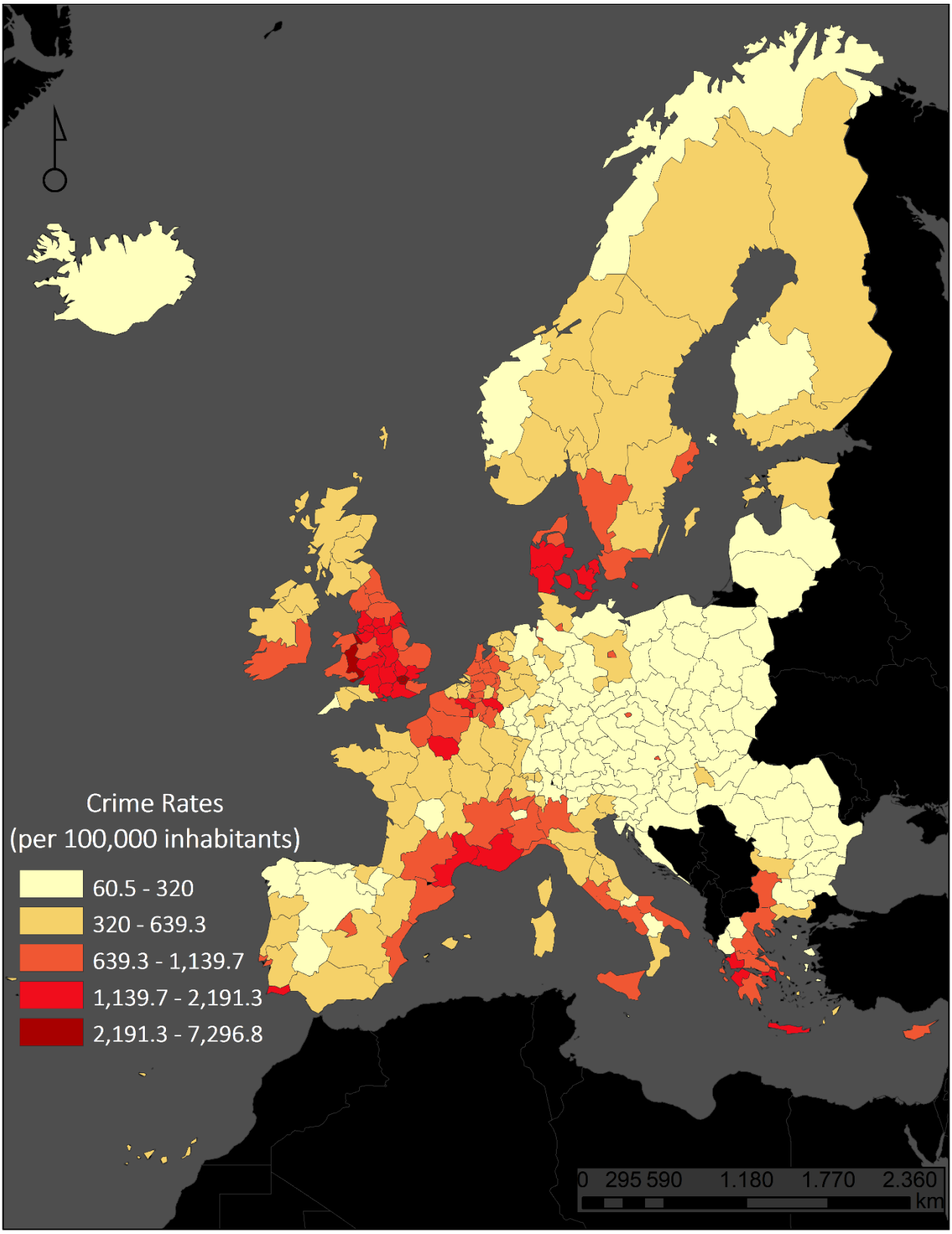

Figure 8. *Map of the crime rates in 2010 (per 100,000 persons).*

Data source: Eurostat, 2017

According to the map that demonstrates the crime rates, most of the NUTS 2 regions have a low crime rate, lower than 500 crimes per 100,000 inhabitants. It is notable that the regions with the highest crime rates (more than 5,000 crimes) are located in the United Kingdom. Particularly, the highest crime rate, which is 7,296.76 crimes per 100,000 inhabitants, corresponds to the region of Inner London, which consists of the West and East Inner London NUTS 3 regions. The second highest rate, 5,581.58 crimes per 100,000 inhabitants, corresponds to Outer London, which includes the East & North East, the South and the West & North West Outer London NUTS 3 regions. In the same context, the third highest rate, 4,053.74 crimes per 100,000 persons, is attributed to the region of East Wales (Monmouthshire & Newport, Cardiff & Vale of Glamorgan, Flintshire & Wrexham and Powys NUTS 3 regions). A potential reason for the high crime rates in these regions, and particularly in the first two regions, is the fact that they include the city of London, which not only, has a great number of inhabitants, but it, also, attracts many tourists throughout every month of the year. On the other hand, the regions with the lowest crime rates (less than 70 crimes per 100,000 individuals) are located in Romania and Poland. Especially, the regions with the lowest rates are North-East Romania (60.53 crimes per 100,000 inhabitants), Podkarpackie in Poland (63.86 crimes per 100,000 inhabitants) and South Muntenia in Romania (66.48 crimes per 100,000). The fact that these regions have significantly low crime rates, can potentially be attributed to the deficient reports of crime incidents to the police forces.

The map that presents the population density of each NUTS 2 region during 2010, shows a quite interesting pattern. The average population density of these regions was 381.3 inhabitants per km^2, which cannot be considered neither extremely high, nor extremely low. The regions that in 2010 had an extremely low population density, lower than 10 inhabitants per km^2, are located in Iceland, Sweden, Norway and Finland. Particularly, Iceland, which consists of only one region, had a population density of only 3.2 citizens per km^2. A quite similar number, 3.3, also, shared the Upper Norrland region of Sweden. The regions that followed were Northern Norway in Norway (4.4), Middle Norrland in Sweden (5.2), North and East Finland in Finland (6.4) and Hedmark and Oppland in Norway (7.5). It is probable that these regions had such a low population density, due to the weather conditions in their locations, which consist of notably low temperatures. The regions that during the same year had significantly high population densities, higher than 5,000 people per km^2, are located in Spain, Belgium and the United Kingdom. The region with the highest population density is Inner London in the United Kingdom, which was the region that presented the highest crime rates during this year, as well. Inner London, had a population density of 9,625.4

citizens per km^2. The region with the second highest population density (of 6,902 people per km^2) was Brussels in Belgium, which includes Brussels, the largest city of the country. The region of Melilla in Spain, also, has an importantly high population density of 5,776.1 people per km^2. This number can be attributed to the extremely small land are of this region (13.4 km^2). Other regions with significantly high population densities are Vienna in Austria (4,289.3), Ceuta in Spain (4,177.4), Berlin in Germany (3,871.6), Outer London (3,827.4) and West Midlands in the United Kingdom (3,008.6).

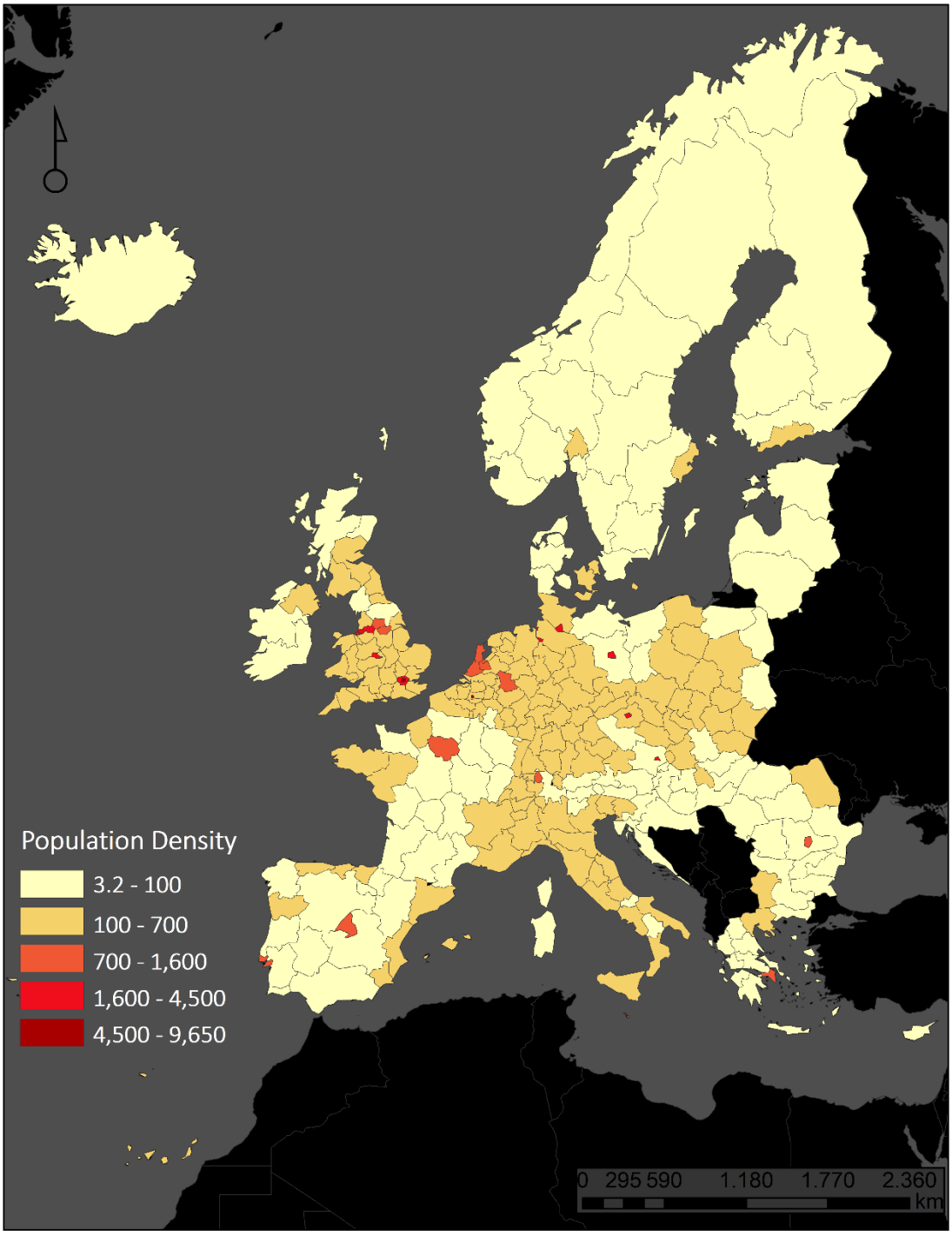

Figure 9. *Map of the population density in 2010.*

Data Source: Eurostat, 2017

The average ratio of the male population in the European Union's regions that during 2010 was older than 15 years old and younger than 64 years old, appears to be 33.88%. The other 66.12% corresponds to women and men that in 2010 were in different age groups, other than the examined one. The regions, which, according to the data that were retrieved from Eurostat, had the highest male rates in these age group are both located in the United Kingdom. The data indicate that the region, which presented the highest male population (71.23%) was the region of Surrey and East and West Sussex. This NUTS 2 regions consists of the Brighton and Hove, the East Sussex CC, the Surrey and the West Sussex NUTS 3 regions. The second highest rate was found in East Wales (69.2%). These high rates can be caused by the concentration of certain types of economic activities in these regions, which attract more men. The region with the lowest percentage of males in the age group 15-64 is Liguria in Italy, with a percentage of 30.21%, which is considered average and not extremely low.

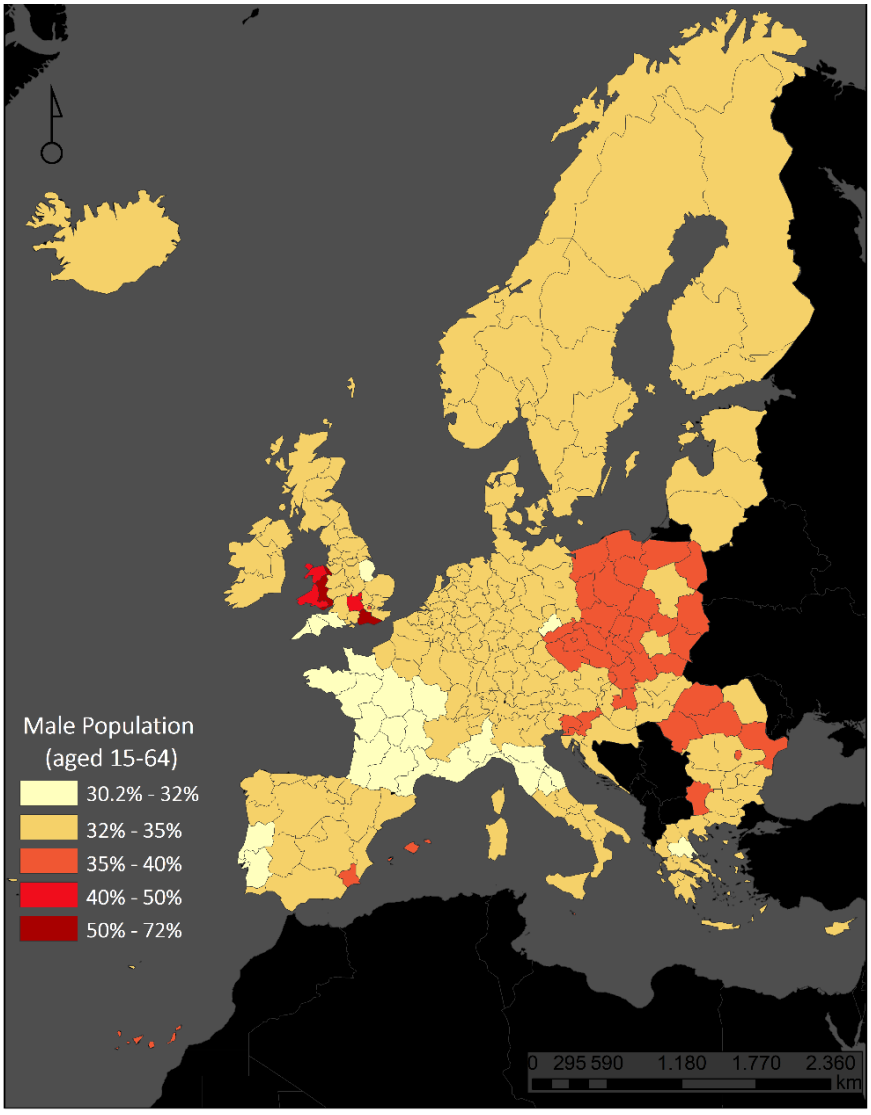

Figure 10. *Map of the percentage of the male population in the age group 15-64 in 2010.*

Data Source: Eurostat, 2017

Owing to the fact that the average percentage of immigrants in the total population of the European Union's regions was 6.4% during 2010, the data indicate several extremely different cases. Particularly, 41 of the regions present a percentage lower than 1%. The regions with the lowest percentage of immigrants, less than 0.1% correspond to Romania and specifically, to the regions South-West Oltenia, South Muntenia and South-East Romania. These percentages are 0.05, 0.06 and 0.09 respectively. The region with the highest percentage of immigrants is the only region of Luxembourg. The percentage of immigrants that reside in Luxembourg is 43.61% of the country's total population. The second and third highest percentages (33.55% and 33.43%) correspond to Brussels in Belgium and to the only region of Liechtenstein respectively.

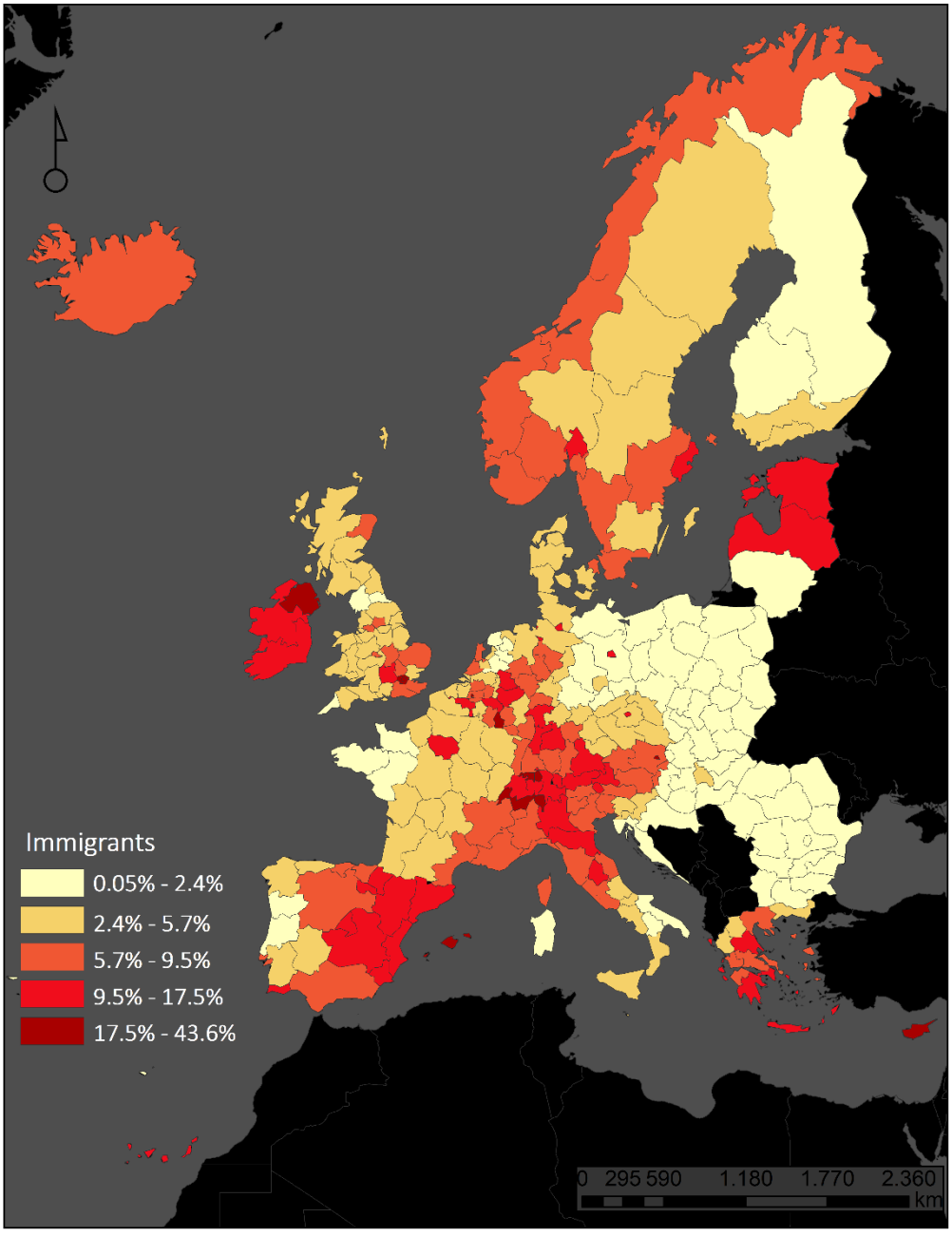

Figure 11. *Map of the percentage of immigrants in 2010.*

Data Source: Eurostat, 2017

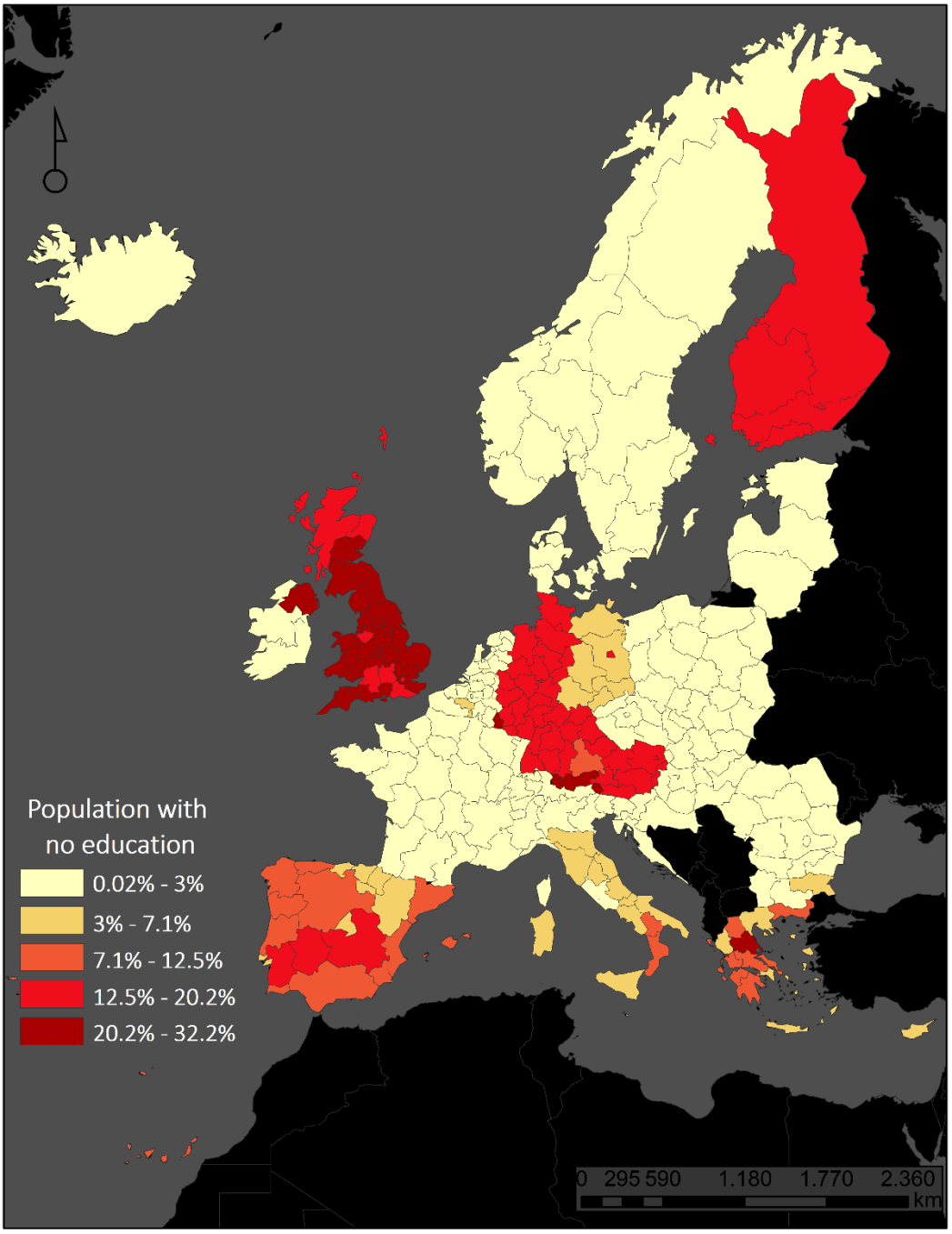

Figure 12. *Map of the percentage of citizens that had received no education in 2010.*

Data Source: Eurostat, 2017

The average percentage of the people that in 2010 had received no education, for the European Union's regions is 8%. This percentage can be considered low. However, some of the studied regions present different results. According to the map that was created, some of the regions present extremely low percentages and others considerably higher than the average. Specifically, the region with the lowest percentage of citizens, who had received no education is the only region of Iceland (0.02%). The other regions whose equivalent percentages are lower than 0.1%, are Bratislava is Slovakia, the region of North Jutland in Denmark, Central Slovakia and the region of

Zealand in Denmark. Their percentages are 0.03, 0.07 for the first two regions and 0.09 for the two last ones. The only regions with a percentage greater than 30%, are both located in the United Kingdom. These regions are Northern Ireland, whose 31.7% of the population had received no education in 2010, and West Midlands, whose 32.2% of the population had not received any education.

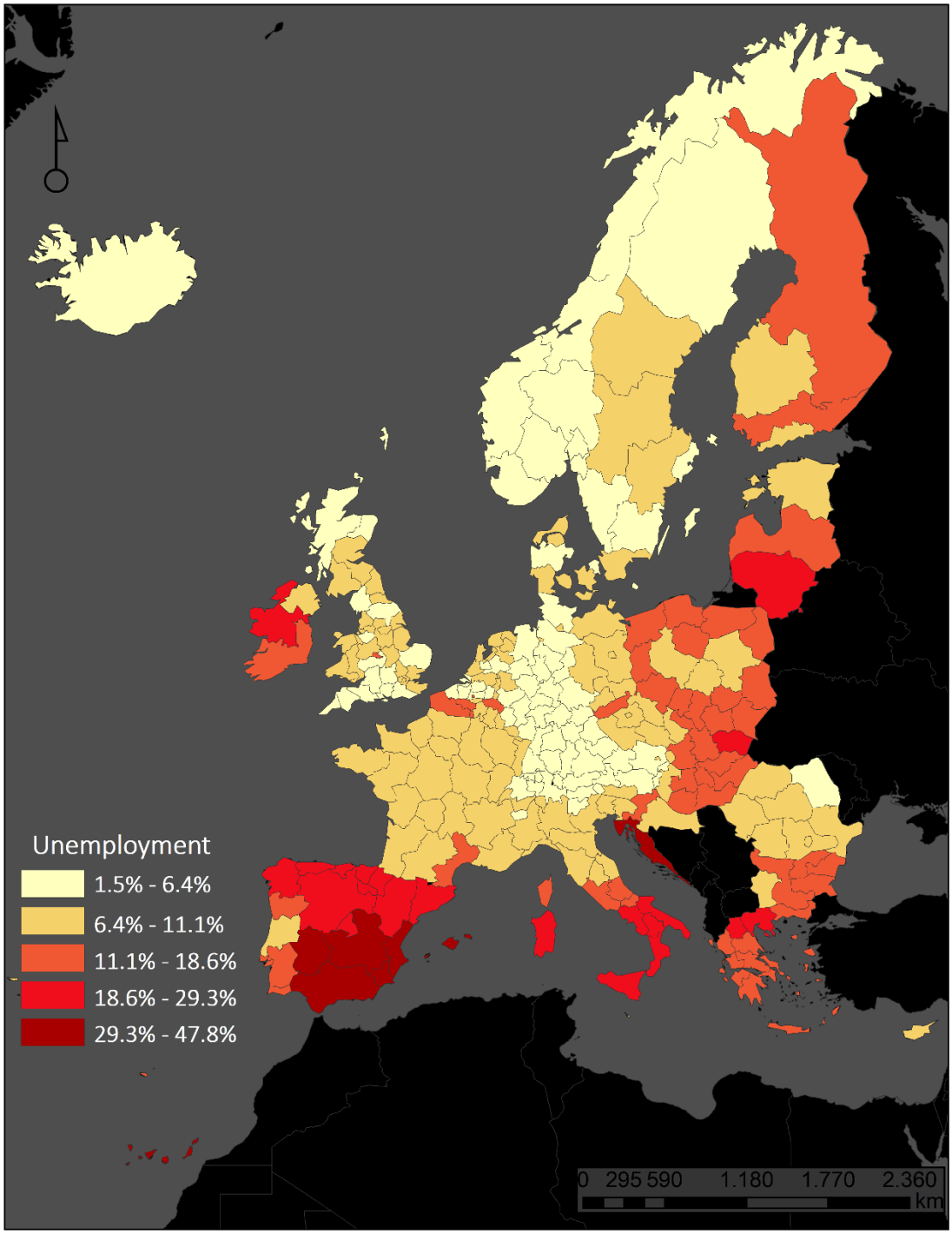

Figure 13. *Map of the percentage of unemployment in 2010.*

Data Source: Eurostat, 2017

According to the retrieved data the average unemployment percentage in the NUTS 2 regions, during 2010, was 10.67%. The regions with the highest unemployment percentages, more than 40% of the economically active population, are all located in Spain. Specifically, the regions Merilla, Ceuta, Andalusia and Extremadura, present the percentages 47.77%, 43.69%, 42.81% and 41.97% respectively. It should be noted that all the Spanish regions have a high unemployment percentage. All have a percentage higher than 20%, but the majority of them has a percentage more than 30%. This phenomenon can be explained by the ongoing financial crisis that began in Spain in 2008. Since the beginning of the crisis the unemployment rates increased extremely. The most concerning fact of the crisis is that the unemployment rates of the young workers, between the ages of 16 and 25, are twice as high as the general rates (Carballo-Cruz, 2011). On the other hand, the regions that present the lowest unemployment rates (lower than 2%) are all located in Norway. Specifically, the regions Western Norway, Trondelag, Agder and Rogaland, Hedmark and Oppland and Northern Norway have unemployment percentages that equal 1.46%, 1.67%, 1.71%, 1.78% and 1.8% respectively. The essence of this variable lies in the fact that it highlights the huge differences among the dynamics and the social and economic conditions of the European Union's countries.

The map that presents the percentage of employees in the public sector for each NUTS 2 region, indicates an average percentage of 23.7%. The regions with the lowest percentages of public employees correspond to Greece and Romania. Particularly, in the region Thessaly of Greece, 8.42% of the economically active population worked in the public sector during 2010. The equivalent percentage of the Romanian region South-West Oltenia, was 9.39%. By contrast, a different Greek region, Epirus, is the one than during 2010, had the highest number of public employees, 42.49%. A region with a similar value is the region of Northern Norway, whose equivalent percentage is 42.45%.

The Gross Domestic Product (GDP) per capita is one of the most important financial indexes. The reviewed data indicate major differences between the GDP of the European Union's regions. The average GDP of all the NUTS 2 regions is 27,722.6 euros. The regions with the lowest GDPs are located in Bulgaria and Romania. The lowest GDP is presented in the region Northwestern Bulgaria and it equals 3,200 euros per citizen. The other regions with extremely low values are Northern Central Bulgaria, Southern Central Bulgaria and North-East Romania. Their GDPs per capita in 2010, were 3,400, 3,600 and 3,800 euros equivalently. On the other hand, the regions with the highest GDPs are located in the United Kingdom and in Liechtenstein. Specifically, the highest

value can be noticed in the only region of Liechtenstein, where 251,709 euros correspond to each citizen. The equivalent GDP for the region of Inner London is 195,000 euros per capita and it consists the second highest value of the variable for 2010.

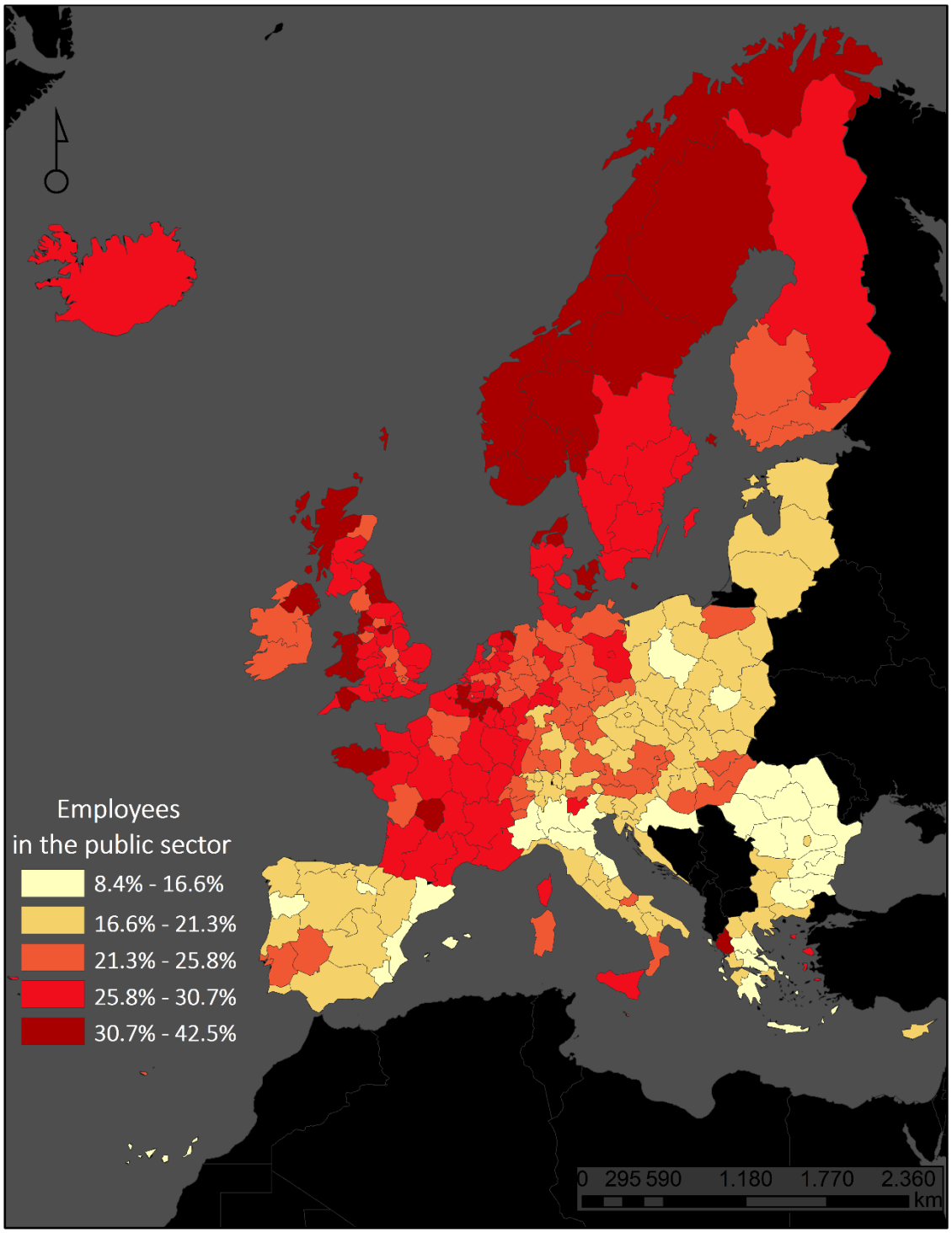

Figure 14. *Map of the percentage of employees in the public sector in 2010.*

Data Source: Eurostat, 2017

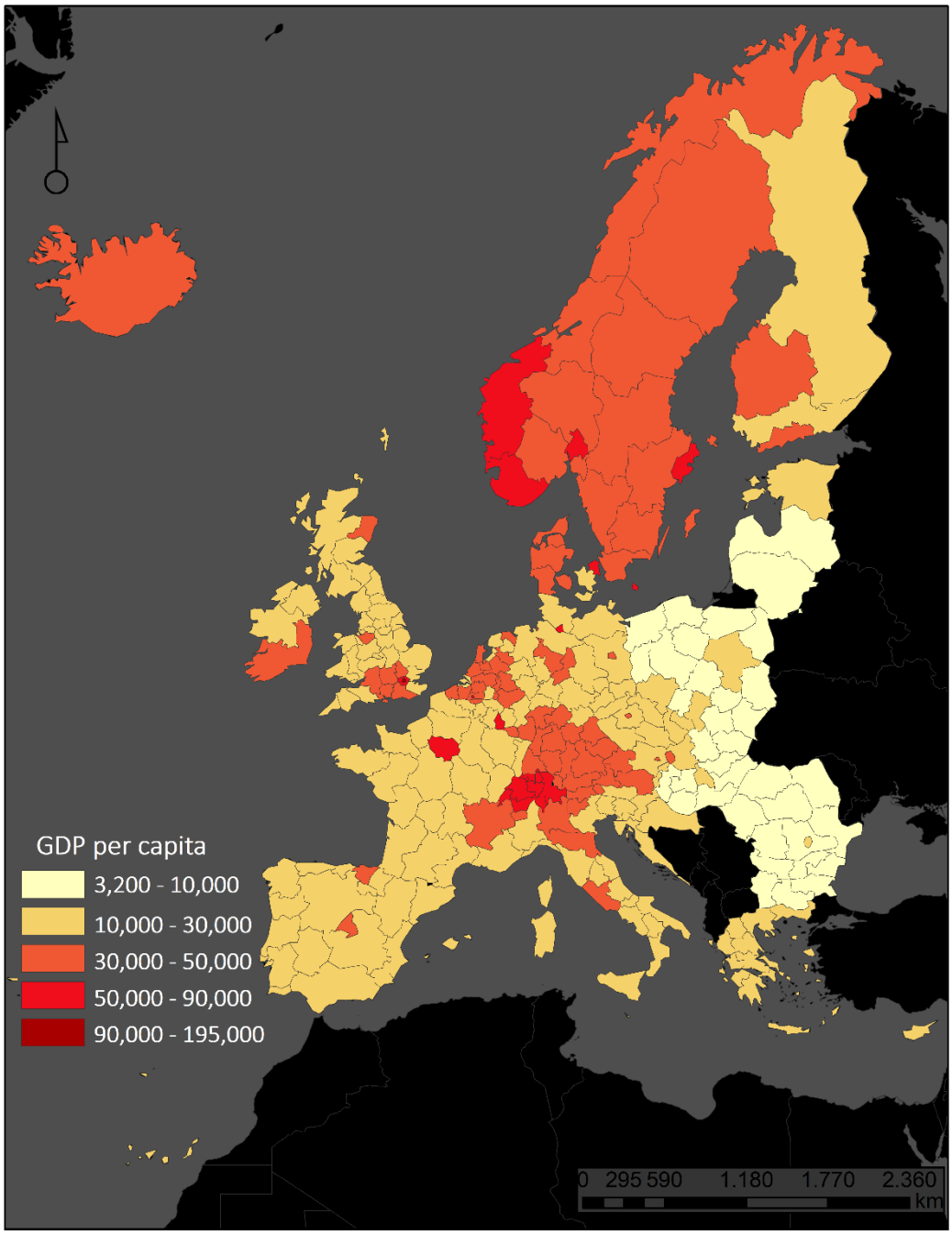

Figure 15. *Map of the GDP per capita (in euros) in 2010.*

Data Source: Eurostat, 2017

The results from the examination of the disposable income of all the studied NUTS 2 regions, are quite similar to the results from the examination of the GDP. The average disposable income of every household in 2010, was 16,570 euros. Nevertheless, the range of the values is notable. The regions with the lowest disposable income are the same regions that during the same year had the lowest GDPs. Specifically, Northwestern Bulgaria and North-East Romania had an average disposable income of 2,300 euros, while Northern Central Bulgaria had an average income of 2,500 euros. In the case of the regions with the highest disposable income, they are all located in

Switzerland. All of the country's regions present an average disposable income that is higher than 55,000 euros. The region with the greatest value is Zurich, whose disposable income is 67,492.23 euros. Except for Switzerland, all the regions of Norway, also, present significantly high values that vary from 39,093 to 43,831 euros.

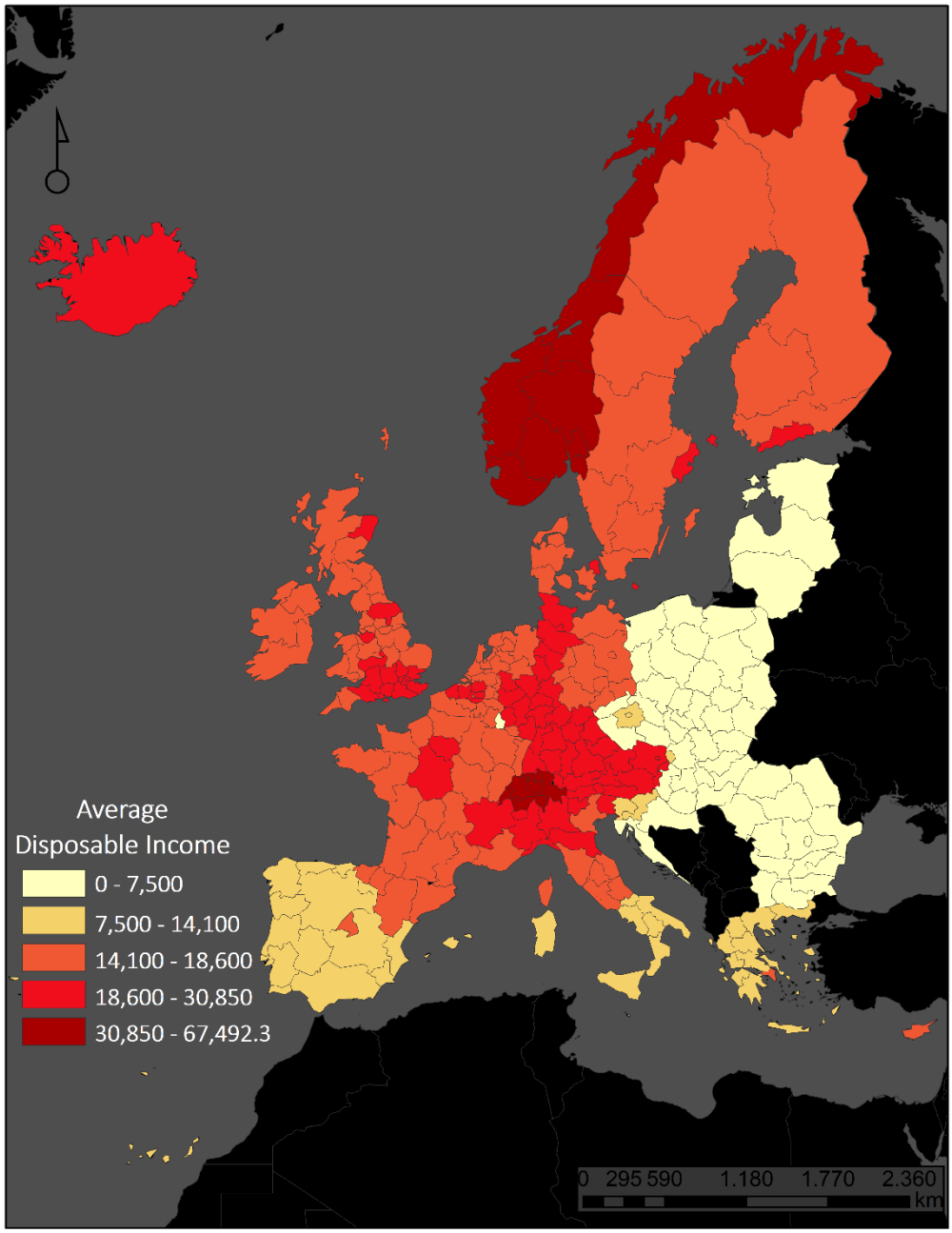

Figure 16. *Map of the average disposable income (in euros) in 2010.*

Data Source: Eurostat, 2017

According to the map that displays the percentage of the total artificial area, the average artificial land cover of the European Union's regions is 7.04%. The region that presents the highest percentage of artificial land cover is Inner London, which was, also, the region with the highest population density during 2010. The 80.49% of the total area of Inner London is covered by

buildings and other artificial constructions. Outer London, also, has a significant percentage of this type of land cover, since the 52.45% of its total area is artificial. In the same context, the regions with the lowest percentage of artificial land cover have a significantly low population density too. These regions consist of Northern Norway (0.6%), the region of Highlands and Islands in the United Kingdom (0.66%), Upper Norrland in Sweden (0.69%), Espace Mittelland in Switzerland (0.79%), Middle Norrland in Sweden (0.88%) and North and East Finland (0.94%). These results are expected, not only because of the population densities of these regions, but, also, because of the fact that they are located in areas with more difficult weather conditions, which do not attract residents.

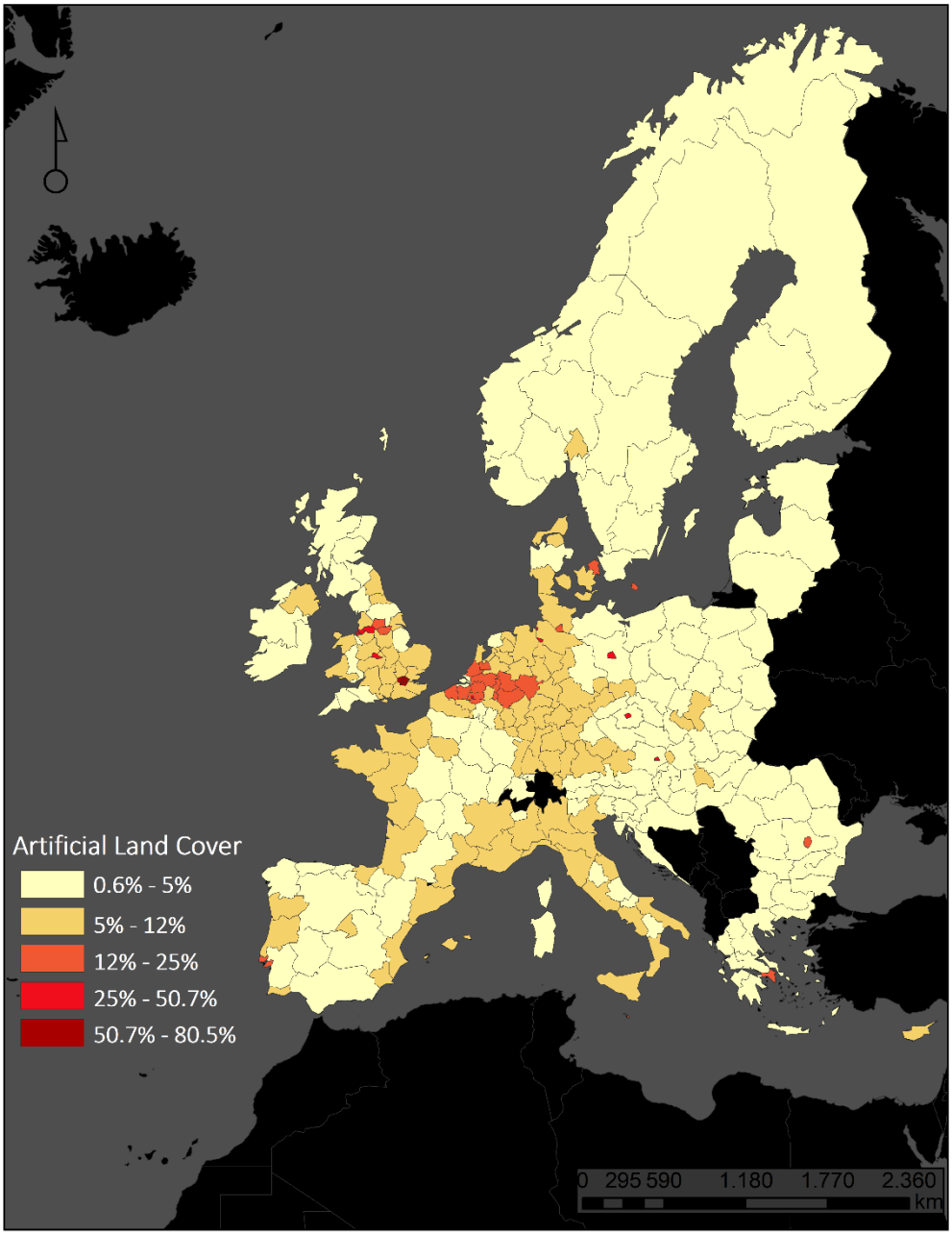

Figure 17. *Map of the percentage of the artificial land cover in 2010.*

Data Source: Eurostat, 2017

6.2. Descriptive Statistics

Descriptive Statistics is a highly significant tool for the understanding of the distribution of a studied variable. This dissertation uses the method, in order to examine and describe the characteristics of the dataset that presents the crime rates of every European Union's region during the year 2010. The technique that is being implemented is the calculation of the measures of central tendency and the measures of dispersion. According to the results, the minimum crime rate is 60.53 and the maximum 7,296.76. These rates indicate the number of the recorded crimes in each region per 100,000 inhabitants. As mentioned previously, this minimum value of the crime rates corresponds to the North-East Romanian region and the maximum value to the region Inner London of the United Kingdom. The average value of the crime rates is 595.7708 crimes per 100,000 people. The median is 423.815, which means that half of the crime rates have a higher value than this and the other half have a smaller value. The great value of the variance suggests that there are large differences between the values of the variable, which is confirmed from the minimum value, the maximum value and the range, as well. The value of the interquartile range, which is fairly high, shows that the data are reasonably spread out from the arithmetic mean.

MEASURES OF CENTRAL TENDENCY						
Arithmetic Mean		*Median*				
595.7708		423.815				
MEASURES OF DISPERSION						
Standard Deviation	*Variance*	*Interquartile Range*	*Minimum Value*	*Maximum Value*	*Range*	*Coefficient of Variation*
682.5617	465,890.4	565.5225	60.53	7,296.76	7,236.23	1.1456

Table 6.1. *Descriptive Statistics of crime rates in NUTS 2 regions.*

These results are, also, presented through the boxplot. The boxplot shows the minimum and maximum value, the arithmetic mean and the values of the 1st and the 3rd quartile. Additionally, it displays the existence of three extreme values, which differ in a large degree from the other values. These extreme values correspond to the three regions with the highest crime rates. As described previously, these regions are Inner London, Outer London and East Wales, all located in the United Kingdom. These outliers are expected to affect the regression model.

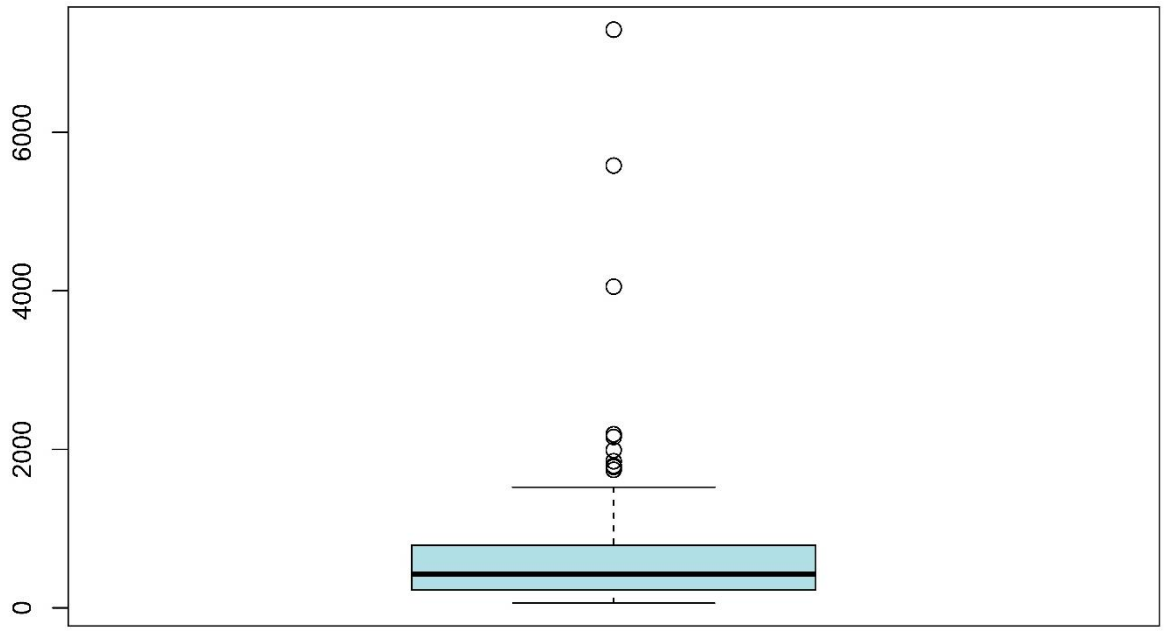

Figure 18. *Boxplot of crime rates in NUTS 2 regions.*

6.3. Spatial Autocorrelation

In order to examine the spatial autocorrelation of the crime rates of every NUTS 2 region that is included in the dataset, during 2010, the R package "lctools" was used. Particularly, the functions "moransI()" and "l.moransI()" were applied for the calculation of the Global and the Local Moran's I respectively.

Bandwidth	*Moran's I*	*Z resampling*	*P-value resampling*	*Z randomization*	*P-value randomization*
3	0.5592	12.6327	1.395e-36	13.7331	6.432e-43
4	0.5143	13.3357	1.434e-40	14.4976	1.255e-47
6	0.466	14.7606	2.623e-49	16.0462	6.081e-58
9	0.4132	16.1327	1.504e-58	17.5368	7.510e-69
12	0.3853	17.5441	6.602e-69	19.0697	4.509e-81
18	0.3403	19.2994	5.434e-83	20.9749	1.111e-97
24	0.3263	21.6519	5.828e-104	23.5295	2.036e-122

Table 6.2. *Global Moran's I of crime rates results.*

In the table produced from the global Moran's I (Figure 20), each row corresponds to a different scenario that was calculated by a repetition of the code. In each scenario, a different number of bandwidth is being used, which means that a different number of possible neighbors is

implemented. The second column displays the value of the Moran's index. It is obvious that the index ranges from 0.3263 to 0.5592. These numbers indicate a positive spatial autocorrelation, which can be weaker when the bandwidth rises and stronger when the bandwidth is lower. The strongest positive autocorrelation is observed between 3 nearest neighbors. While the number of the nearest neighbors increases, the spatial autocorrelation becomes weaker, but it is still positive. The third and the fifth column of this table present the statistical measures Z, which are based on the resampling and the randomization hypotheses. The fourth and the sixth column display the p-values for each Z value, which indicate the statistical importance. The fact that all the p-values are extremely lower than 0.05, demonstrates that the data used in this dissertation are statistically important. This statistical importance can be attributed to the fact that the data concern the total of the population and not a sample.

The object created by the function "l.moransI()", is a scatter plot that contains information that can be mapped. The following scatter plot illustrates the distribution of the normalized values of crime rates and the distribution of the normalized and weighted aggregates of the crime rates of the 20 nearest neighbors of each region. It shows that the majority of the values is concentrated near zero, but they tend to be positive. However, several of the values are spread out in the third quartile, which is positive. These values tend to be higher and, therefore, represent a stronger autocorrelation. The red line that appears in the plot is the regression line.

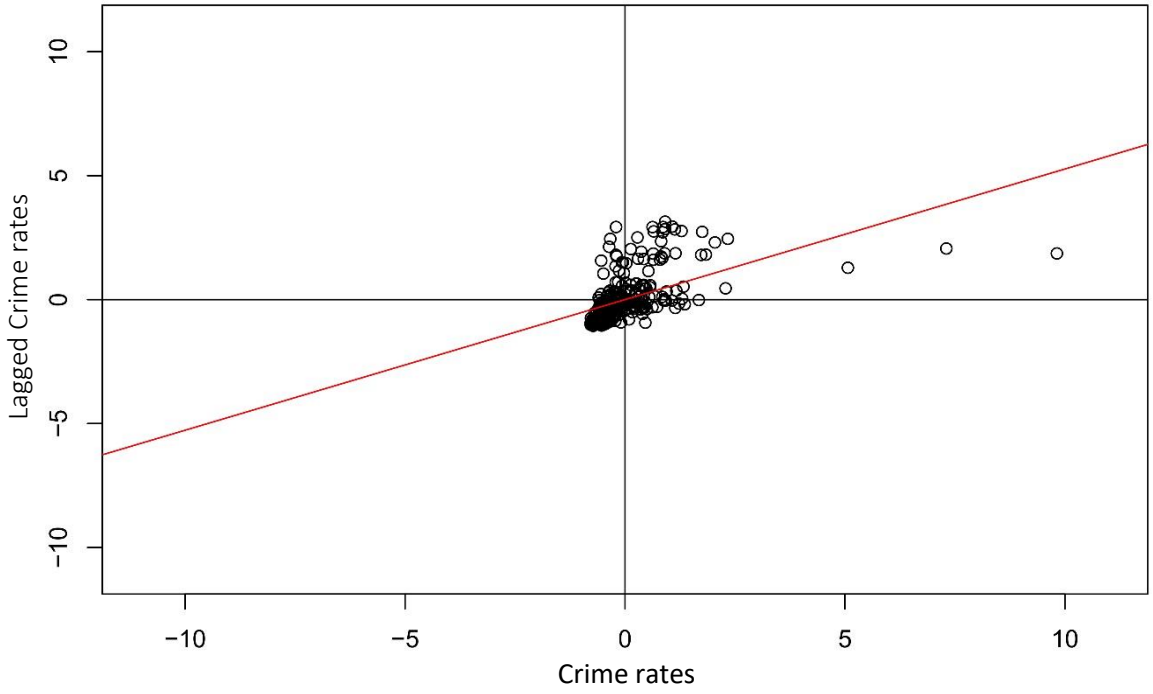

Figure 19. *Local Moran's I Scatter plot.*

The results of the local Moran's I can be used for the creation of maps, as well. Particularly, with the usage of the function "ggplot()", which can be used through the R package "ggplot2", the creation of a cluster map is possible. This map results from the classification of the statistically important local Moran's indexes into four classes, based on the sign of the corresponding paired values of the scatter plot. The produced map indicates that there are two areas, where regions with similarly high crime rates are concentrated. The first area consists of the majority of the south United Kingdom and one Irish region and the second includes two regions of Belgium. On the other hand, there seem to be three areas, where regions with similarly low crime rates are concentrated. The first cluster consists of only Romanian regions and the second includes several Polish regions and one that belongs to Slovakia. The third cluster includes several German regions. The third type of concentration that can be identified from the map, is the one that includes regions with low crime rates that are surrounded by regions with high crime rates.

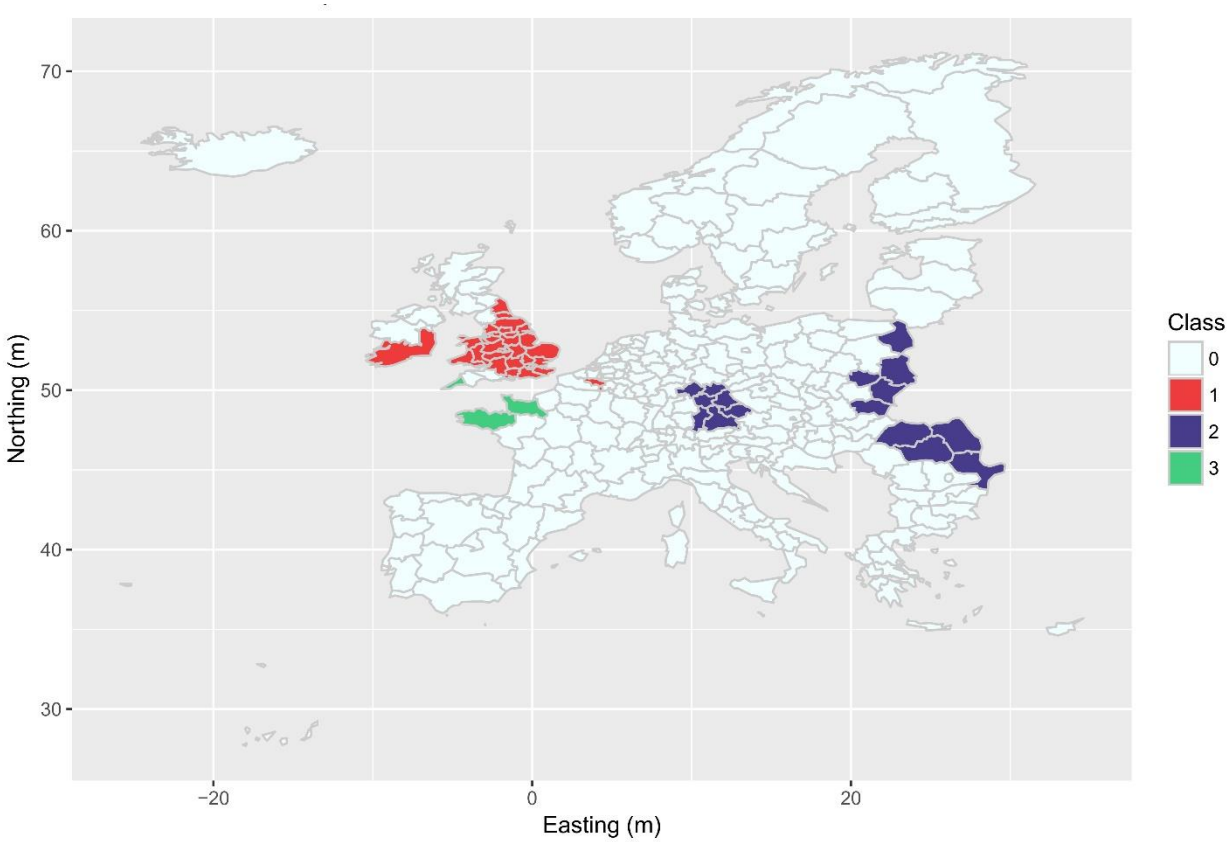

Figure 20. *Moran's cluster map.*

6.4. Inequality Index

For the evaluation of the inequality index, the function “spGini() of the R packages “lctools” was used to calculate the Gini index. The first column of the Table 6.3. presents the number of the nearest neighbors. The numbers of the nearest neighbors that were chosen each time are 2, 3, 25, 37 and 39 (any number can be selected). The second column is the Gini index, which appears to be about 0.45. This number indicates that a considerable inequality between the crime rates of the NUTS 2 regions exists. The third column represents the Gini index between the nearest neighbors and the fourth column, the Gini index between non-neighbors. The fact that the Gini index of neighbors tends to be significantly low and near zero, indicates that there is almost perfect equality between neighboring regions. In other words, regions that are located next to each other, tend to have very similar crime rates. The smaller the bandwidth number gets, the equality becomes more perfect. The Gini of non-neighbors index’s results are quite similar to the Gini’s results. However, it is obvious that as the number of the nearest neighbors increases, the inequality reduces. These three variations of the index are presented graphically in the two following plots.

Bandwidth	*Gini*	*Gini of Neighbors*	*Gini of Non-Neighbors*
2	0.4538	0.0009	0.453
3	0.4538	0.0019	0.452
25	0.4538	0.0287	0.4251
37	0.4538	0.0464	0.4074
39	0.4538	0.0492	0.4046

Table 6.3. *Gini Index results.*

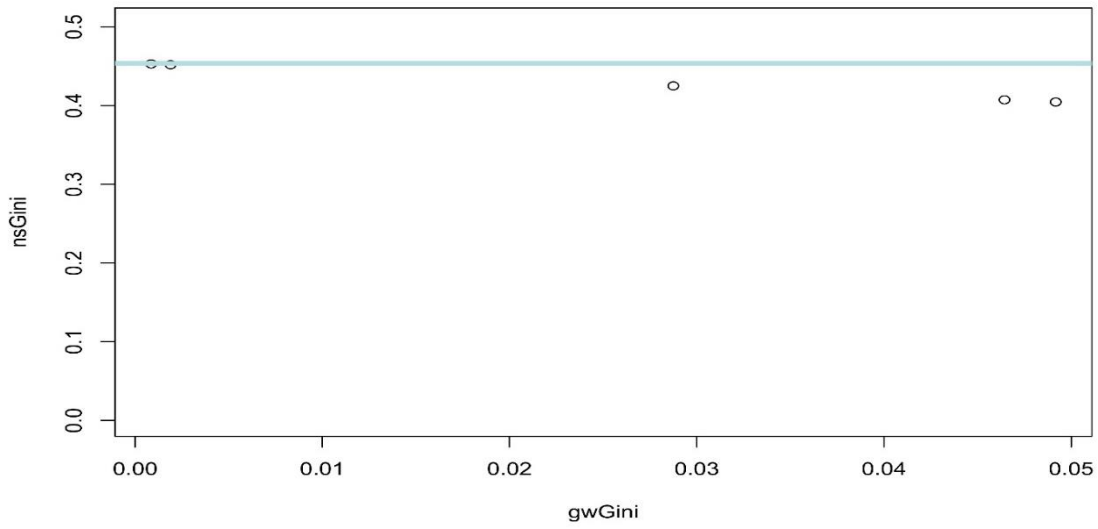

Figure 21. *Plot of the Gini Index (Neighbor vs Non-Neighbor Ginis).*

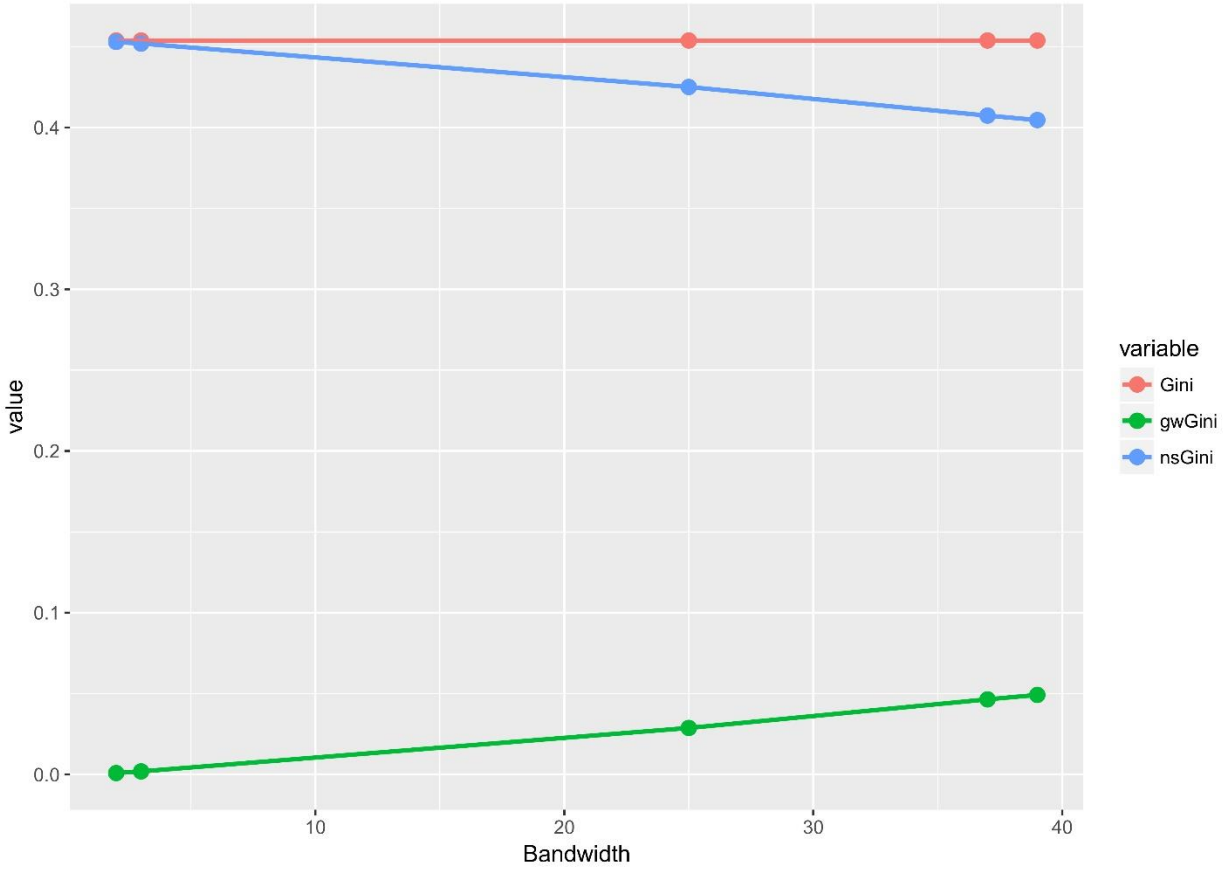

Figure 22. *Plot of the Gini Index.*

6.5. Regression

The first step, before the implementation of any regression method, is the ascertainment of whether the independent variables that will be used, are independent amongst them. To that end, the evaluation of the correlation coefficient is essential. However, in order to select a method for this evaluation, as well, it is important to examine the distribution of the independent variables. This can be accomplished in three ways. The first one is by examining the skewness and the kurtosis of each variable. The conditions that need to be verified so that a variable can be considered to have a normal distribution are: $-1.96 < \frac{Skewness}{std\ error} < 1.96$ and $-1.96 < \frac{Kurtosis}{std\ error} < 1.96$. The second way for examining the normality of a distribution is by observing the histogram that is produced from the variable. A histogram that has a certain type of curve (Figure 23) is considered to present a normal distribution. The third way, which was selected and used in this

Figure 23. *Normal Distribution curve.*
Source: LibreTexts, 2018

dissertation, is the Shapiro-Wilk normality test. The results of the test are defined by the formula:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

where $x_{(i)}$ is the ith order statistic, $\bar{x}$ is the arithmetic mean and a_i are the constants (Shapiro and Wilk, 1965). The important values of the table that is produced from this test, are the p-values. If the p-value is greater than 0.05, then the variable has a normal distribution. The results of the Shapiro-Wilk test of this analysis indicate that only one of the independent variables has a normal distribution. Particularly, the ratio of the employees in the public sector has a p-value, which equals 0.0751. All of the other variables have the same p-value, 2.2e-16, which is significanlty lower than 0.05. The test was implemented in R, with the usage of the function "shapiro.test()", of the package "stats".

Variable	*W*	*p-value*
Population Density	0.3457	2.2e-16
Ratio of Male Population	0.3733	2.2e-16
Ratio of persons with No Education	0.8262	2.2e-16
Ratio of Unemployed persons	0.7909	2.2e-16
GDP per capita	0.6069	2.2e-16
Ratio of employees in the public sector	0.9909	0.0751
Average Disposable Income	0.7368	2.2e-16
Ratio of Artificial Land Cover	0.5789	2.2e-16
Ratio of Immigrants	0.8168	2.2e-16

Table 6.4. *Shapiro-Wilk Test Results of all the variables.*

Since the results of this test, indicate that not all of the independent variables have a normal distribution, the Correlation Coefficient method that is used is the Spearman. In this calculation, the correlation between all the variables, is examined. The function "cor()" that is used, is available from the R package "stats" as well. The results demonstrate a weak positive correlation between the majority of the variables. In some cases there is a weak negative correlation, for example between the ratio of unemployed persons and the population density etc. Nevertheless, two pairs of variables present a very strong positive correlation. One of these pairs is the population density and the ratio of the artificial land cover, whose correlation is 0.8762. The other pair, which has a correlation of 0.8314, is the GDP per capita and the average disposable income. Any regression model whose independent variables have a very strong positive or negative correlation, is not

correct. To that end, one of the variables of each one of these pairs should be removed from the model. Since the population density and the average disposable income are more important to the study of the phenomenon, according to the reviewed literature, the GDP per capita and the ratio of the artificial landcover, will not be used in the regression model.

The correlation coefficient, which was described earlier, can be graphically illustrated in the Figure 24, which was produced using the R package “GGally” and the function “ggpairs()”.

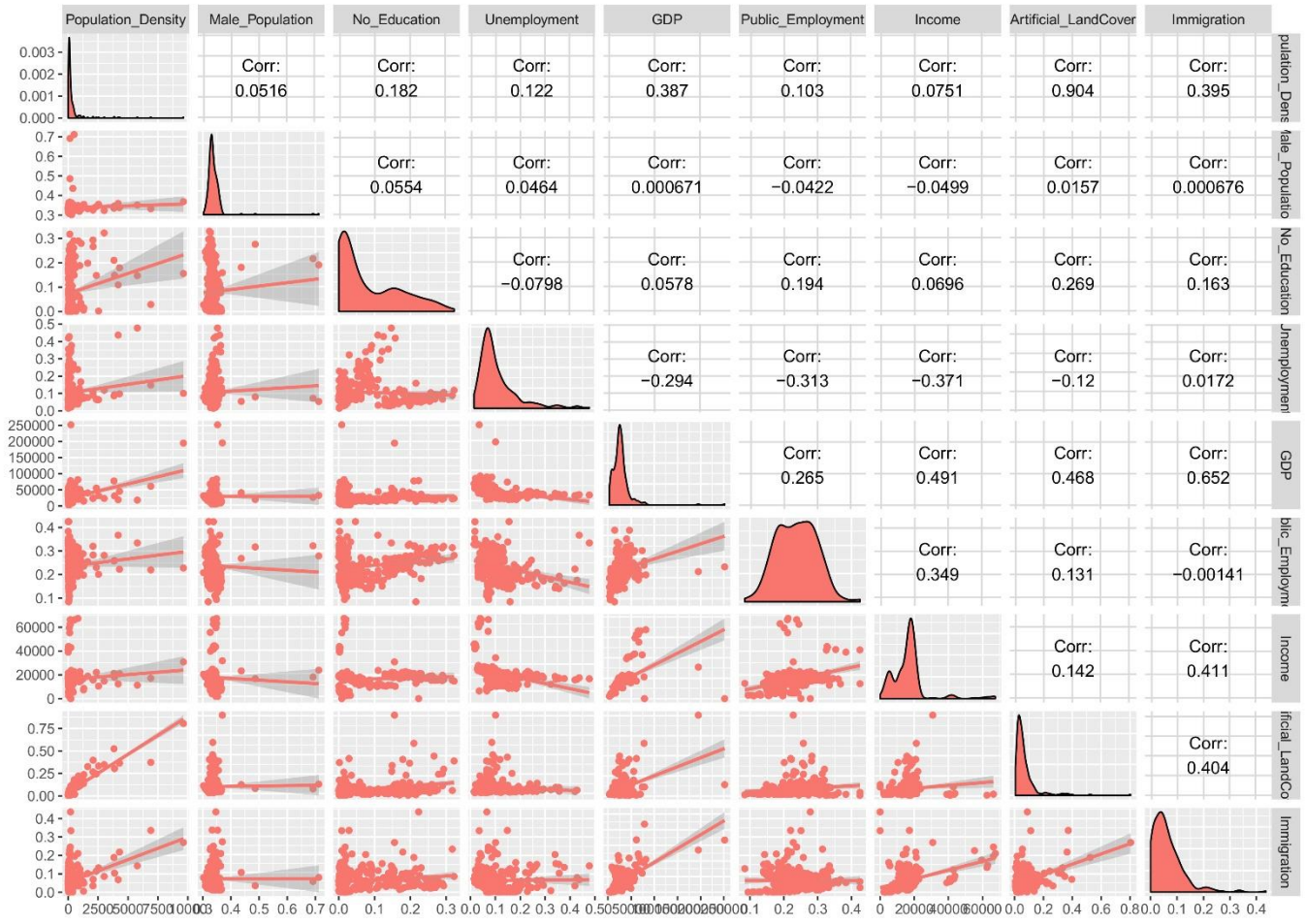

Figure 24. *Plots of the Correlation Coefficient results.*

The linear regression, was calculated with the function “lm()”, which is included in the R package “stats”. The arguments of the function are the dependent variable, the independent variable(s) and the name of the dataset, which includes all these variables. For this dissertation the dependent variable is the crime rates of the European Union’s NUTS 2 regions in 2010. The independent variables that were selected, after the elimination of the two highly correlated variables that was described previously, are the population density of the regions in 2010, the ratio of the males aged 15-64 in the total population, the ratio of the persons that had received no education, the ratio of the unemployed persons in 2010, the ratio of the employed persons in

the public sector, the average disposable income of the regions and the ratio of the immigrants in the total population. The regression results are presented in the Table 6.5.

Coefficients:

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	-1.648e+03	3.291e+02	-5.008	9.84e-07***
Population Density	3.820e-01	3.483e-02	10.967	< 2e-16***
Male Population	4.632e+03	8.475e+02	5.466	1.03e-07***
Persons with no education	9.163e+02	3.576e+02	2.562	0.01094*
Unemployed persons	-2.718e+02	4.292e+02	-0.633	0.52705
Public employees	1.782e+03	5.693e+02	3.131	0.00193**
Disposable Income	-6.542e-04	3.674e-03	-0.178	0.85882
Immigrants	1.142e+03	6.195e+02	1.844	0.06632

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 494.8 on 276 degrees of freedom.
Multiple R-squared: 0.4875.
Adjusted R-squared: 0.4745.
F-statistic: 37.51 on 7 and 276 DF.
P-value: < 2.2e-16

Residuals:

Min: -2157.0
1Q: -226.8
Median: -70.5
3Q: 239.9
Max: 3385.4

Studentized Breusch-Pagan test:
BP = 133.5
df = 7
P-value < 2.2e-16

Table 6.5. *Regression results.*

From the results that are produced from the regression function, a great variety of conclusions can be drawn. In the basic linear model the seven independent variables included, explain about 48% of the variance of crime rates. Particularly, this means that the 48% of the alterations of the crime rates in every NUTS 2 region, can be explained by the population density, the ratio of the male population in the age group 15-64, the ratio of the persons that had receive no education, the ratio of the unemployed persons, the ratio of the employees in the public sector, the average disposable income and the ratio of the immigrants in each region.

A highly significant parameter that is presented with the regression's results, is the column "Estimate" of the produced table. This column consists of the $\beta_0, \beta_1, \beta_2, \dots, \beta_n$, parameters of the regression formula. In the first row of the column there is the constant parameter and in the rest,

the parameters of every independent variable. According to these parameters, the ratio of the unemployed persons and the average disposable income of the inhabitants of a region, affect the crime rates in a negative way, while the other variables have a positive effect. For example, if the ratio of the unemployed persons and/or the average income increase, the crime rates will decrease. Specifically, if the ratio of the unemployed persons increases by 1%, then the crime rates will decrease by 271.8 per 100,000 residents. If the average disposable income increases by 1,000 euros, the crime rates are expected to reduce by 0.6542%. On the other hand, an increase in the population density by 1 person/km^2 is expected to increase the crime rates by 0.382%. A same level of increase in the ratio of the male population in the age group 15-64, the ratio of the persons that had received no education, the ratio of the public employees and the ratio of the immigrants will increase the crime rates by 4,632, 916.3, 1,782 and 1,142 per 100,000 residents respectively. The regression formula that is produced by these results is the following:

Crime Rates= -0.001648 + 0.382*Population Density + 4,632*Male Population + 916.3*Persons with no education – 271.8*Unemployed persons + 1,782*Public employees – 0.0006542*Disposable Income + 1142*Immigrants.

An equally important parameter, is the Pr(>|t|), which represents the p-value. The p-value demonstrates the statistical significance of the correlations presented in the "Estimate" column. If the p-value of a variable is lower than 0.05 (confidence level 95%), then this variable can be considered statistically significant. According to the results the population density of each region and the ratio of the male population in the ages 15-64 are significant, with a 99.9% confidence level. The ratio of the public employees is statistically significant at the level of significance 99%, while the ratio of the individuals that had received no education at the level 95%. The other independent variables are not statistically significant.

The Breusch-Pagan test was implemented using a different function, "bptest()", in order to test for heteroskedasticity in the regression. Heteroskedasticity occurs when the variance of the data is not constant, or in other words when the data are spread unequally. If the p-value, which results from the test is lower than 0.05, then the presence of heteroskedasticity is assumed. The p-value that emerged from the test is $0.22*10^{-15}$ and it indicates the existence of heteroskedasticity. Since one of the assumptions of linear regression is that there is no heteroskedasticity, the results of the regression are biased and a better model should be defined.

The regression results can be plotted. The most frequently created plot is the "Residuals vs Fitted" plot (Figure 25), which is used to detect non-linearity, unequal error variances and outliers. The plot presents the predicted distribution and the real distribution. The fitted values represent the predictions and are illustrated with symbols. The line of the plot represents the real distribution. Since the fitted values are not spread equally along the residuals line, it can be assumed that there is a non-constant variance and there is heteroskedasticity. This assumption is, also, confirmed from the Breusch-Pagan test, which was analyzed previously.

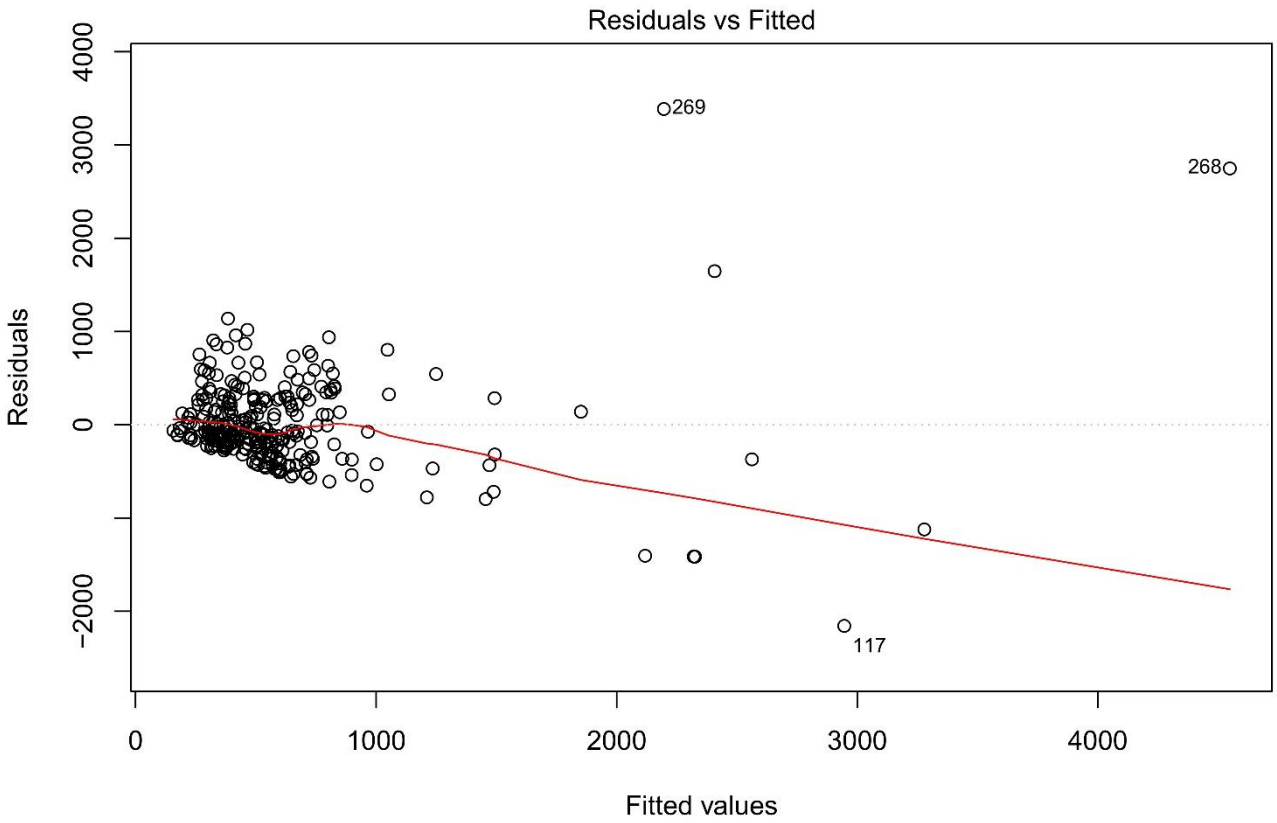

Figure 25. *Residuals vs Fitted plot.*

The Normal Q-Q plot is used to show whether the residuals are normally distributed. In this case, the residuals should follow a straight line well and not deviate severely. From the produced plot (Figure 26), it can be observed that the residuals do follow a straight line for the most part. However, there are several values, which deviate in a great degree. Overall, the residuals seem to follow the line in the middle of the graph, but curve in the extremities. This means that the data have more extreme cases than it would be expected from a normal distribution. For this reason, the distribution can be considered to be a heavy-tailed distribution.

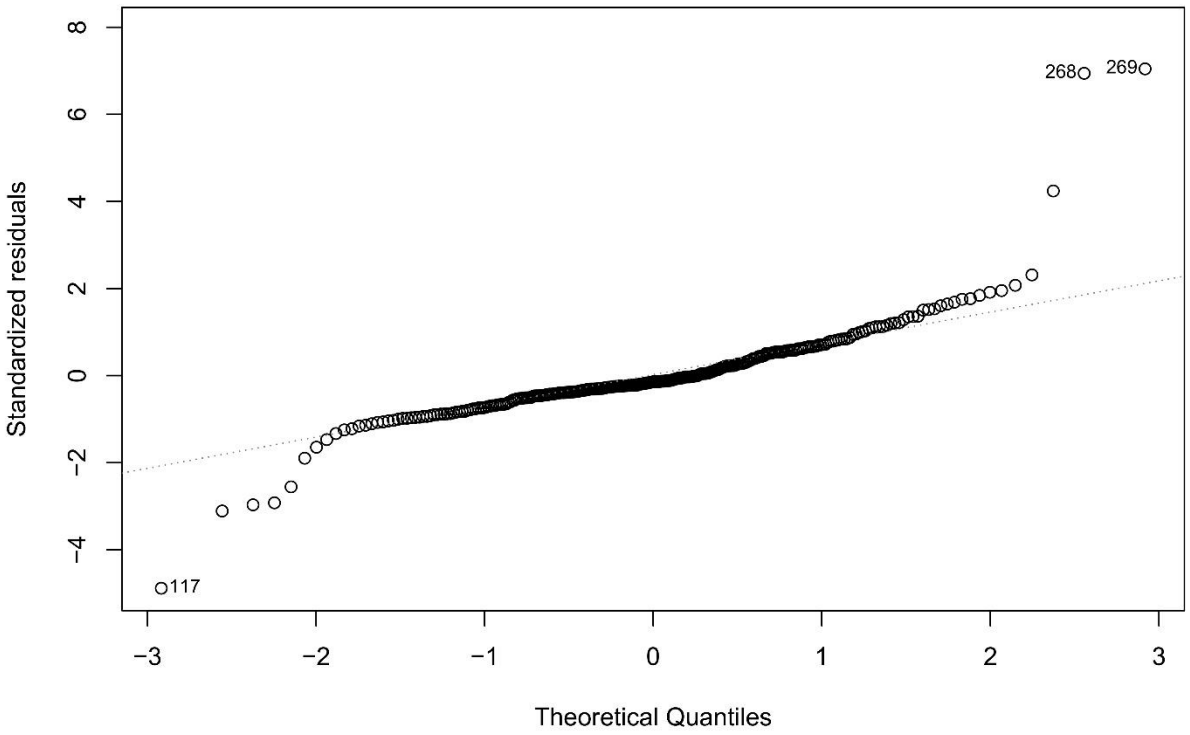

Figure 26. *Normal Q-Q plot.*

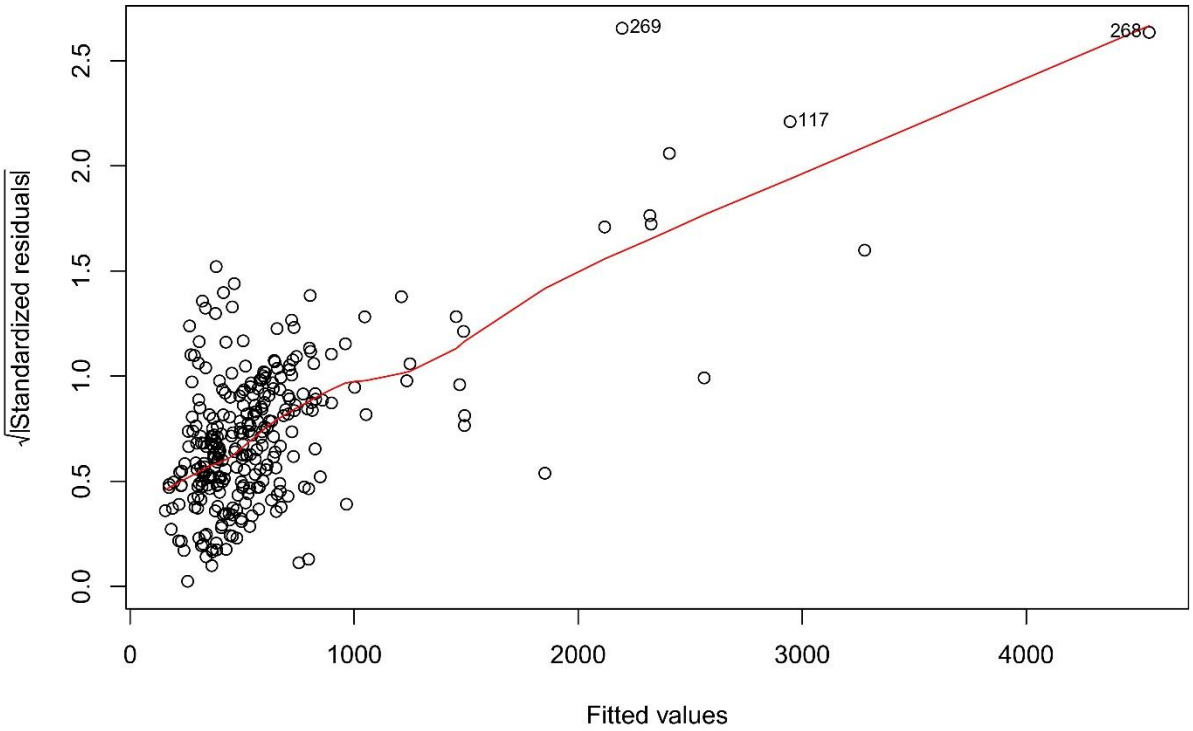

Figure 27. *Scale Location plot.*

The third produced plot, is the Scale Location plot, also called Spread Location plot. The aim of this plot is to show whether the residuals are spread equally along the ranges of the predictors. It is quite similar to the Residuals vs Fitted plot, but it uses the square root of the standardized residuals. The presence of a horizontal line with equally and randomly spread points suggests that

there is homoskedasticity. The plot of the used data, indicates that the residuals begin to spread wider along the x axis as it passes around 800. Due to this fact, the presence of heteroskedasticity is once again confirmed.

The last plot that is produced from the implementation of regression, is the Residuals vs Leverage plot, which identifies influential cases, if any exist. Since many values (even extreme values) do not influence the regression analysis, their inclusion or exclusion would not produce much different regression results. Nevertheless, some values can be notably influential, even if they seem to be within a reasonable range of values. This plot is used to identify these influential values. The patterns of the plot are not relevant. What is examined is the presence of outlying values at the upper and at the lower right corners, because these are the spots, where the influential cases can be found. These cases should be outside of the dashed red line, which is the Cook's distance. When there are cases outside of these lines, it means that they have a high Cook's distance score and they can affect the regression results. On this plot, there can be observed three influential cases, the 117th, the 268th and the 279th.

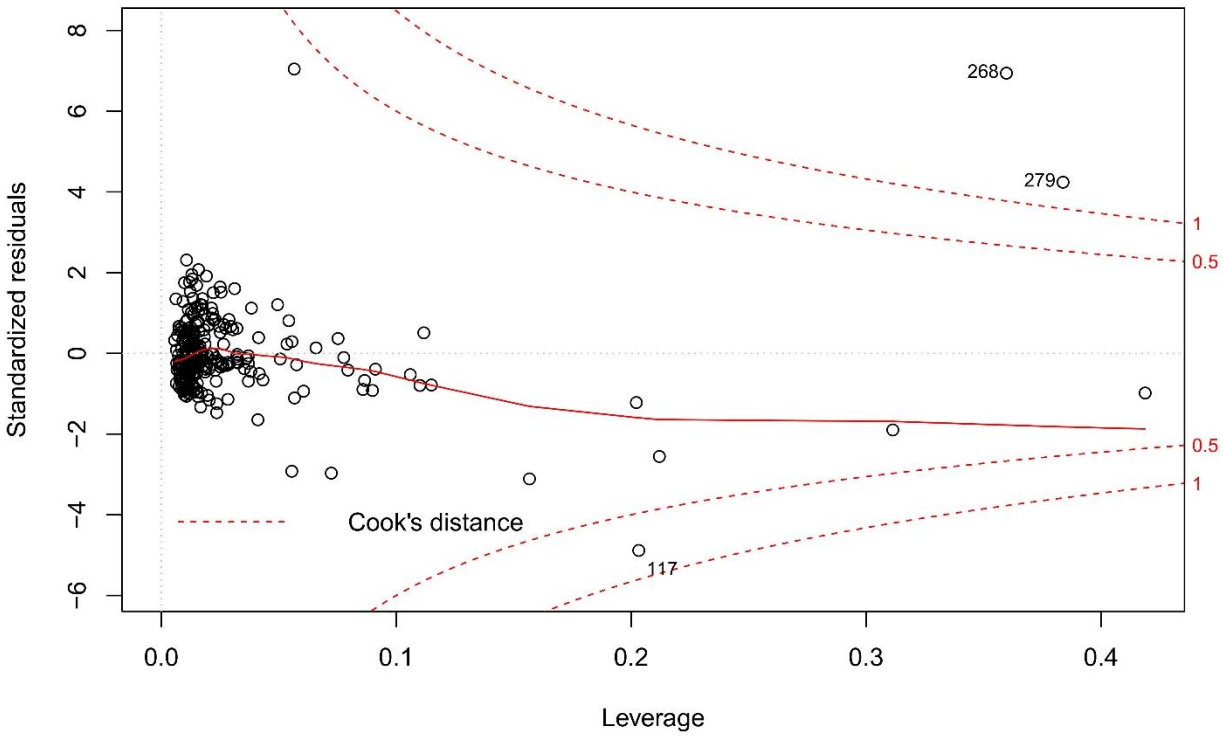

Figure 28. *Residuals vs Leverage plot.*

6.6. Geographically Weighted Regression

The final step of this spatial analysis, is the implementation of the Geographically Weighted Regression (GWR). GWR is one of the most recent regression techniques, which focuses on the examination of local variations in spatial processes. The dependent and the independent variables that are used for the GWR are the same with the ones used in the linear regression. Particularly, the dependent variable is the crime rates in each NUTS 2 region for 2010 and the independent variables are the population density of each region, the ratio of the male population between the ages 15 and 64, the ratio of the individuals that had received no education, the ratio of the unemployed individuals, the ratio of the employed persons in the public sector, the average disposable income and the ratio of the immigrants in the total population.

The function "model.selection.gwr()", which is included in the R package "GWmodel", is the first function that is used for this type of regression. The function runs a diagnostic test for the selection of the best model, by combining and comparing all the variables. The results of this test can be demonstrated in two different plots, both of which present the best model. The best model can be considered the one that has the lowest AICc score. According to the results, the best model is the 28th, which includes all the independent variables.

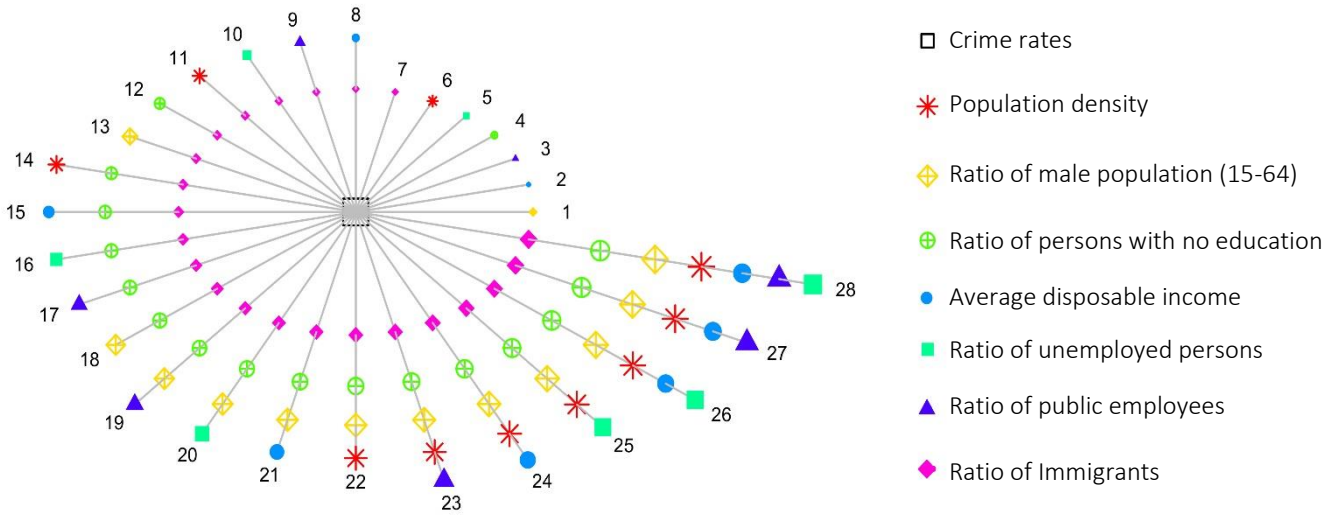

Figure 29. *View of GWR model selection with different variables.*

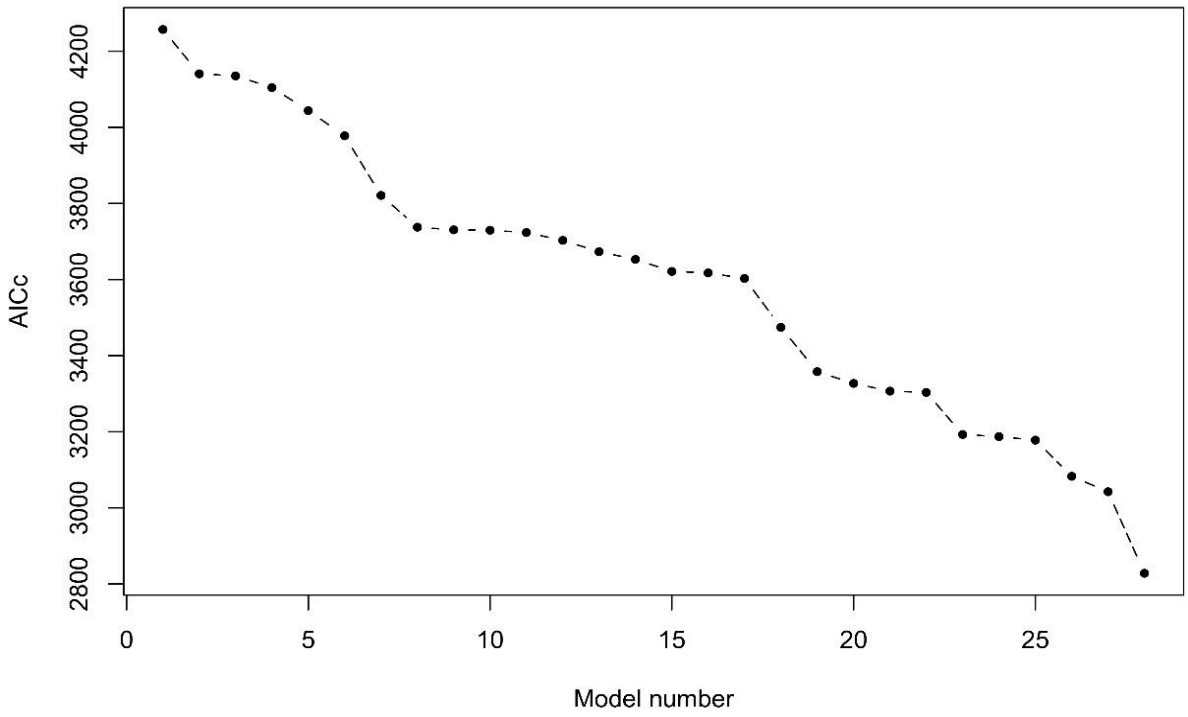

Figure 30. *Alternative view of model selection procedure.*

In the same R package, the function "bw.gwr()", selects automatically the appropriate bandwidth to calibrate the GWR model. By the determination of the argument approach as AICc, the selected number of nearest neighbors is based on the best model, which is the model with the lowest AICc, as it was mentioned previously. In this case, the function suggests that the suitable bandwidth for the model is 43 nearest neighbors.

For the calibration of a basic GWR model, the function "gwr.basic()" is used. The main arguments of the function are the dependent and independent variables, as well as the dataset that contains them and the bandwidth. The GWR function that was written in the script has the following form:

```
GWR<-gwr.basic(crime_rates ~ pop_density + male_pop + no_education + unemployed
+ emp_public + income + immigrants, data=dataset, bw=43, kernel="bisquare",
adaptive=TRUE, F123.test=TRUE)
```

The results of the basic GWR can be divided into two sections. The first one consists of the results of the Global regression, which are the same as the results of the linear regression. The second section contains the results of the GWR.

Results of Global Regression

Residuals:

Min: -2157.0
1Q: -226.8
Median: -70.5
3Q: 239.9
Max: 3385.4

Coefficients:

	Estimate	*Std. Error*	*t-value*	*Pr(>|t|)*
(Intercept)	-1.648e+03	3.291e+02	-5.008	9.84e-07***
Population Density	3.820e-01	3.483e-02	10.967	< 2e-16***
Male Population	4.632e+03	8.475e+02	5.466	1.03e-07***
Persons with no education	9.163e+02	3.576e+02	2.562	0.01094*
Unemployed persons	-2.718e+02	4.292e+02	-0.633	0.52705
Public employees	1.782e+03	5.693e+02	3.131	0.00193**
Disposable Income	-6.542e-04	3.674e-03	-0.178	0.85882
Immigrants	1.142e+03	6.195e+02	1.844	0.06632

Significance stars:

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 494.8 on 276 degrees of freedom.
Multiple R-squared: 0.4875.
Adjusted R-squared: 0.4745 .
F-statistic: 37.51 on 7 and 276 DF.
P-value: < 2.2e-16.

Extra Diagnostic information

Residual sum of squares: 67,566,836.
Sigma(hat): 489.4882.
AIC: 4,339.779.
AICc: 4,340.436.

Results of Geographically Weighted Regression

Model calibration information:

Kernel function: bisquare.
Adaptive bandwidth: 43 (number of nearest neighbours).
Regression points: the same locations as observations are used.
Distance metric: Euclidean distance metric is used.

Summary of GWR coefficient estimates:

	Min.	1st Qu.	Median	3rd Qu.	Max.
(Intercept)	-6.1777e+03	-4.0395e+02	1.6084e+03	2.9204e+03	11,674.5558
Population Density	-4.3724e-02	1.1200e-01	2.2542e-01	4.1339e-01	2.2551
Male Population	-3.4915e+04	-7.2530e+03	-2.2504e+03	4.6485e+03	15,179.006
Persons with no education	-7.4708e+03	-2.4404e+03	-1.5572e+03	-7.7456e+02	3,523.8648
Unemployed persons	-1.3917e+04	-6.3620e+02	1.3735e+03	4.7447e+03	9,061.2531
Public employees	-7.8929e+03	-8.6124e+02	4.2066e+02	1.6566e+03	6,404.2816
Disposable Income	-2.8126e-01	-2.3976e-02	-4.6759e-04	1.0438e-02	0.1206
Immigrants	-4.6739e+03	7.7538e+02	2.5005e+03	6.1639e+03	24,519.6508

Diagnostic information:

Number of data points: 284.
Effective number of parameters (2trace(S) - trace(S'S)): 118.0233
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 165.9767.
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 4,043.052.
AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 3,841.378.
Residual sum of squares: 8,865,189.
R-square value: 0.9327615.
Adjusted R-square value: 0.8846595.

F3 test (Leung et al. 2000):

	F3 statistic	*Numerator DF*	*Denominator DF*	*Pr(>)*
(Intercept)	1.57309	92.62216	194.3	0.004491 **
Population Density	3.08222	22.53825	194.3	1.316e-05 ***
Male Population	1.63343	93.23783	194.3	0.002271 **
Persons with no education	0.99467	66.69524	194.3	0.497531
Unemployed persons	3.20822	84.38842	194.3	1.290e-11 ***
Public employees	1.26670	99.07562	194.3	0.082773
Disposable Income	5.06006	92.57808	194.3	< 2.2e-16 ***
Immigrants	5.30758	49.42308	194.3	< 2.2e-16 ***

Significance stars:
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 6.6. *Basic GWR results.*

According to these results, the local model has a significantly better fit than the global model. This can be concluded from the fact that its AICc is in a great degree lower than the global model's AICc. Particularly, the AICc of the global model is 4,340.436, while the AICc of the local model is 4,043.052. As mentioned previously, a low AICc score indicates a better model fit.

A different parameter, whose value has improved notably in the GWR model, is the R-square. The R-square of the global model is 48.75% and that of the local model is 93.28%. The adjusted R-square of the global model is 47.45% and the equivalent parameter of the local model is 88.47%. It is obvious that these values have almost been doubled in the GWR model, suggesting a local rather than a global effect on the determinants of crime.

The values of the table "Summary of GWR coefficient estimates", show descriptive statistics of the local parameter estimates of each variable. Owing to the fact that the model is local and examines each region separately, the local estimate of each parameter has a fairly large variance. This means that the estimated parameter of each variable is different for each region. The relationship between the crime rates and every variable can be positive for some regions and negative for others. To give an illustration, the increase by 1 person/km^2 of the population density can decrease the crime rates at a region by 0.04% or increase them to 2.26%, at another region. The following

maps, present the results of the local model. Particularly, they show the variances of the statistically significant variables' estimated coefficients, by classifying them. In other words, the maps show how much each variable affects the crime rate of each region.

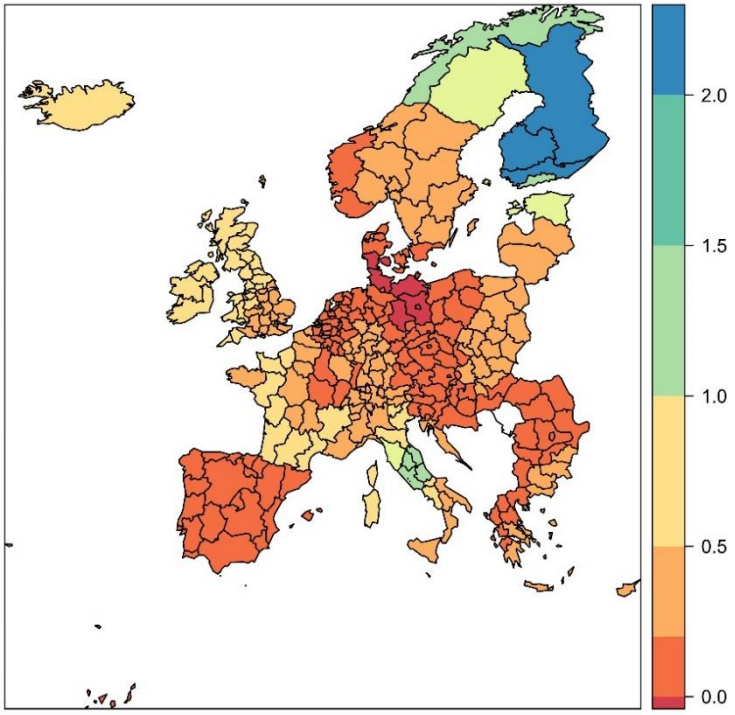

Figure 31. *GWR coefficient estimates for the population density.*

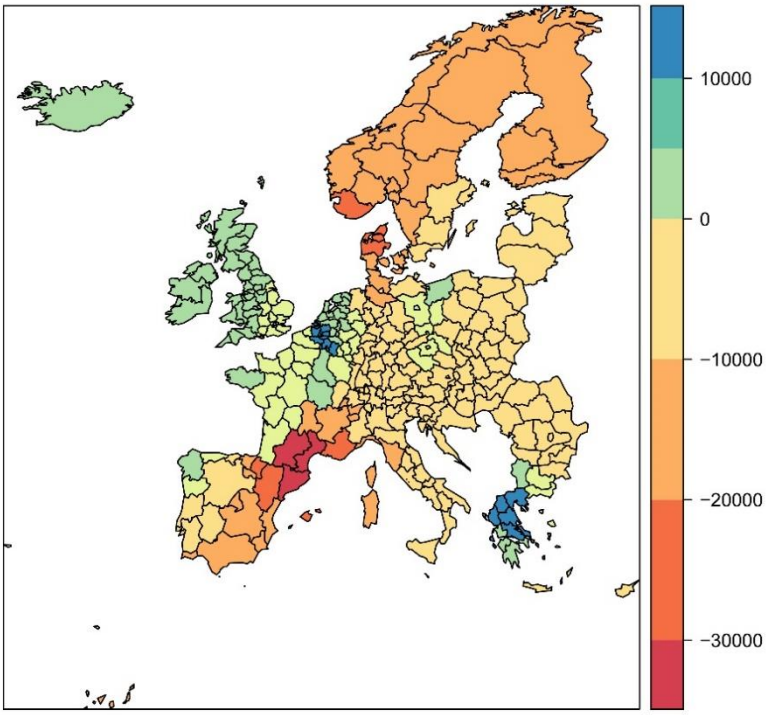

Figure 32. *GWR coefficient estimates for the ratio of the male population.*

These maps indicate that the population density and the ratio of the immigrants affect the crime rates of the majority of the regions in almost the same way that they affect them in the global

model. This means that their estimated variances for most of the regions approach the values of the estimated parameters from the linear regression. On the other hand, the estimated parameters of the ratio of the male population, the ratio of the unemployed persons and the average disposable income, have major differences from the equivalent parameters of the global model.

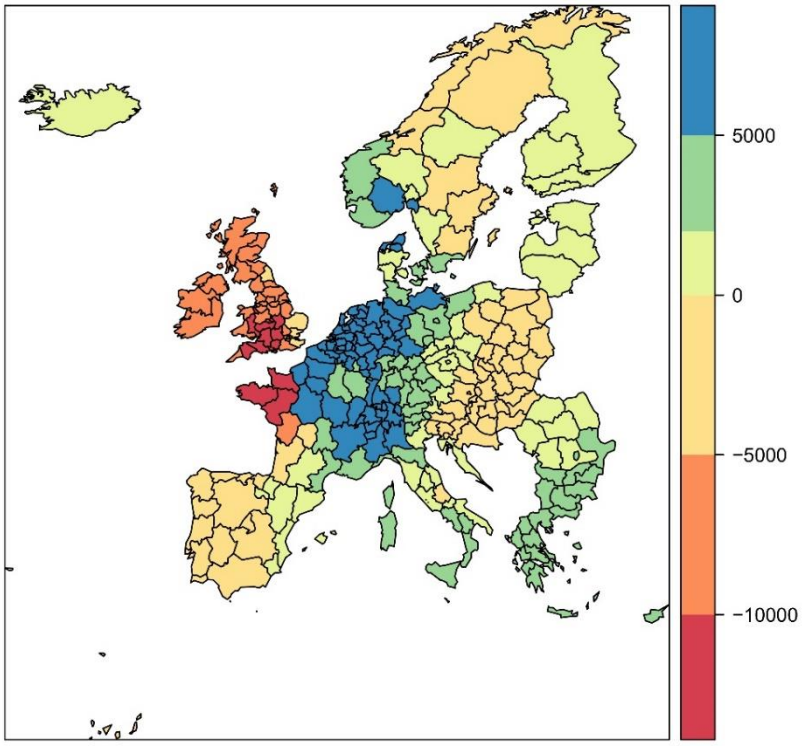

Figure 33. *GWR coefficient estimates for the ratio of unemployed individuals.*

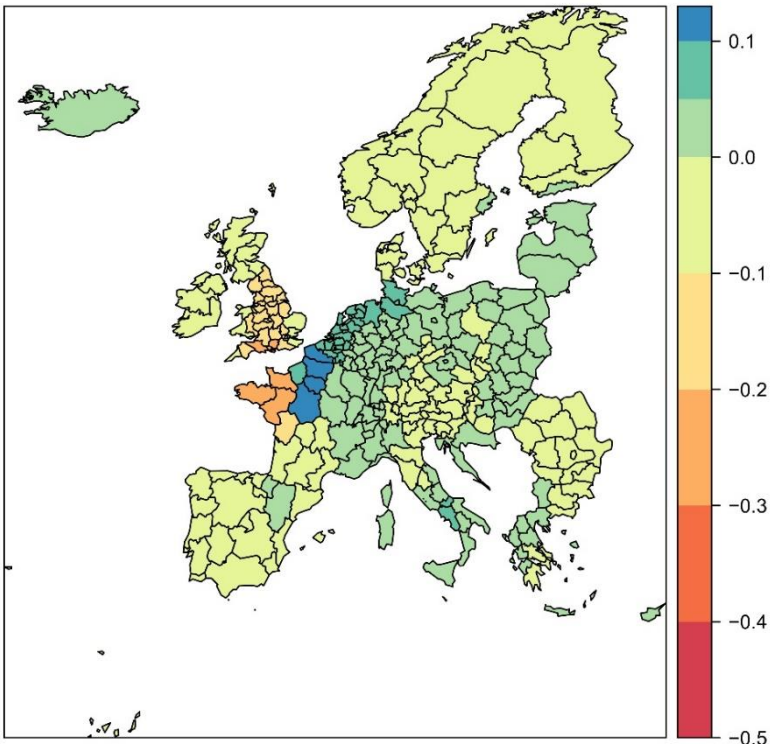

Figure 34. *GWR coefficient estimates for the average disposable income.*

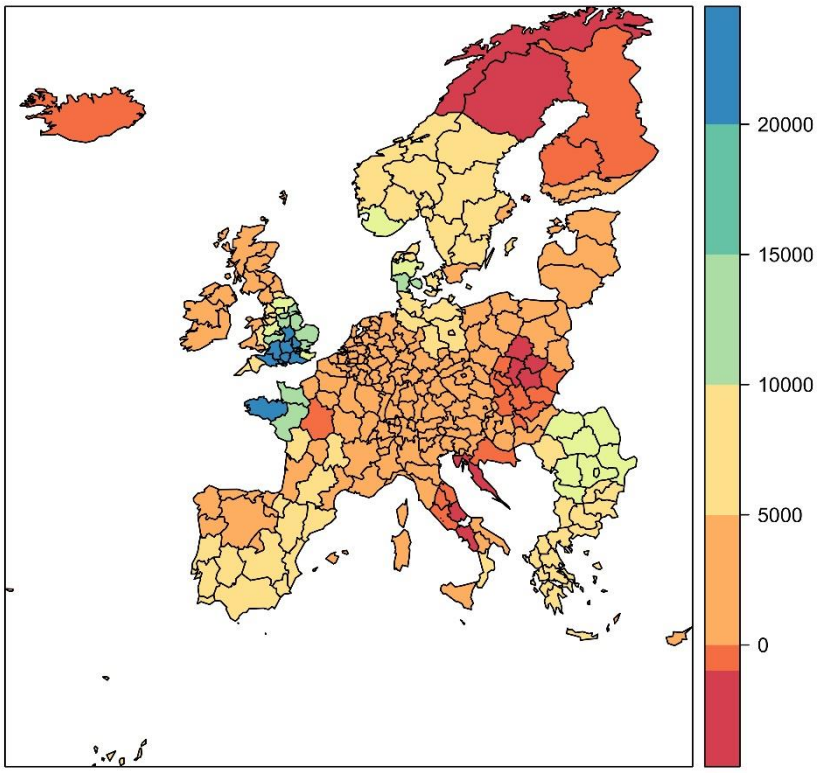

Figure 35. *GWR coefficient estimates for the ratio of immigrants.*

In this type of regression, the statistical significance of every variable's variance can be examined with the Leung tests. According to the results of these tests, the population density, the ratio of the unemployed persons, the average disposable income and the ratio of immigrants are the most statistically significant variables, since their level of confidence is higher than 99.9%. Their p-values are significantly lower than 0.01%. The variable that has a 99% level of confidence, is the ratio of the male population between the ages 15 and 64 (p-value= 0.23%).

Some significant characteristics of the data, such as the presence of outliers and the presence of non-constant variance, can render the implementation of different regression techniques essential. In this case, the robust GWR method is applied, using the R package "GWmodel" and the function "ger.robust()". This method is insensitive to outliers and to high leverage points (Harris et al., 2010; Fox and Weisberg, 2013).

The model that is applied, does not present a better fit than the basic GWR model, since its AICc is higher. Specifically, the GWR's AICc is 4,043.052, while the robust GWR's AICc is 4,212.811. In the same context, the R-squared and the adjusted R-squared, are lower than the equivalent parameters produced from the GWR, as well (although they differ in a smaller degree). In particular, the R-square of the robust model is 87.79% and the equivalent parameter of the GWR is 93.28%. The adjusted R-square that results from the robust model is 79.46% and the GWR's

same parameter is 88.47%. From this number it can be stated that in the robust GWR model, a smaller percentage of the dependent value (crime rates) can be explained by the selected independent values.

In the results of the robust method, each variable's parameters (for the regression formula), present small differences from the equivalent parameters produced from the GWR method. The only difference that can be considered important, occurs in the variable, which corresponds to the ratio of the immigrants. Its variance in the basic GWR model was from -0.0046739 to 24,519.6508 in different regions. This variance changed drastically with the robust GWR model, since, even though it starts from the same lower number, it can, only, reach the number 15,071.9997 in some regions.

Results of Global Regression:

Residuals:

Min: -2157.0
1Q: -226.8
Median: -70.5
3Q: 239.9
Max: 3385.4

Extra Diagnostic information:

Residual sum of squares: 67,566,836.
Sigma(hat): 489.4882.
AIC: 4,339.779.
AICc: 4,340.436.

Coefficients:

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	-1.648e+03	3.291e+02	-5.008	9.84e-07***
Population Density	3.820e-01	3.483e-02	10.967	< 2e-16***
Male Population	4.632e+03	8.475e+02	5.466	1.03e-07***
Persons with no education	9.163e+02	3.576e+02	2.562	0.01094*
Unemployed persons	-2.718e+02	4.292e+02	-0.633	0.52705
Public employees	1.782e+03	5.693e+02	3.131	0.00193**
Disposable Income	-6.542e-04	3.674e-03	-0.178	0.85882
Immigrants	1.142e+03	6.195e+02	1.844	0.06632

Significance stars

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 494.8 on 276 degrees of freedom.
Multiple R-squared: 0.4875.
Adjusted R-squared: 0.4745.
F-statistic: 37.51 on 7 and 276 DF.
P-value: < 2.2e-16.

Results of Geographically Weighted Regression

Model calibration information:

Kernel function: bisquare.
Adaptive bandwidth: 43 (number of nearest neighbours).
Regression points: the same locations as observations are used.
Distance metric: Euclidean distance metric is used.

Summary of GWR coefficient estimates:

	Min.	1st Qu.	Median	3rd Qu.	Max.
(Intercept)	-6.1126e+03	-4.0395e+02	8.7186e+02	2.1446e+03	9,780.6502
Population Density	-4.3724e-02	1.1150e-01	2.1551e-01	4.2801e-01	2.2551
Male Population	-2.9218e+04	-6.1923e+03	-2.3357e+03	2.9094e+03	15,148.8739
Persons with no education	-8.4518e+03	-2.2403e+03	-1.1201e+03	-2.7265e+02	2,185.1668
Unemployed persons	-7.1118e+03	-4.7041e+01	2.1227e+03	4.8311e+03	10,707.5780
Public employees	-5.8011e+03	-1.1896e+03	2.7686e+02	1.6346e+03	6,404.2816
Disposable Income	-9.6244e-02	-6.4954e-03	1.2443e-03	1.3139e-02	0.1922
Immigrants	-4.6739e+03	5.7177e+02	1.8813e+03	5.2717e+03	15,071.9997

Diagnostic information:

Number of data points: 284.
Effective number of parameters (2trace(S) - trace(S'S)): 114.7318.
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 169.2682.
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 4,212.811.
AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 4,010.929.
Residual sum of squares: 16,101,992.
R-square value: 0.8778736.
Adjusted R-square value: 0.7946031.

F3 test (Leung et al. 2000):

	F3 statistic	*Numerator DF*	*Denominator DF*	*Pr(>)*
(Intercept)	0.77408	92.62216	188.14	0.91644
Population Density	1.73284	22.53825	188.14	0.02566 *
Male Population	0.73386	93.23783	188.14	0.95273
Persons with no education	0.73259	66.69524	188.14	0.92911
Unemployed persons	0.63841	84.38842	188.14	0.98981
Public employees	0.76144	99.07562	188.14	0.93422
Disposable Income	0.84066	92.57808	188.14	0.82501
Immigrants	1.16228	49.42308	188.14	0.23661

Significance stars:

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 6.7. *Robust GWR results.*

The accuracy of the GWR's results has been quite controversial. In order to examine this accuracy, a simple test is used. This test is called multi-collinearity diagnostics and aims to investigate the presence of multi-collinearity among the independent variables (Kalogirou, 2015). The term collinearity suggests that the two (or in this case more) independent variables are almost perfect linear combinations of each other. If multi-collinearity is proven to be present, the regression's results are considered unstable and are presumed to have high standard errors (The R Project for Statistical Computing, 2018). The R package that was used for this purpose is the "GWmodel". The function that was included in the script is the "gwr.collin.diagno()", which provides a variety of local collinearity diagnostics for the basic GWR model's independent variables.

In order to render the results defensible and claim that there is no multi-collinearity, the local VIFs should be lower than 10 and the local correlation coefficients should range from -0.8 to 0.8 (Kalogirou, 2015). Neither of these conditions apply completely to the results of the diagnostics of this analysis. The minimum values of the local VIFs are all lower than 10, but their maximum values are notably higher. In the same context, not all the local correlation coefficients are higher than -0.8 and lower than 0.8. The majority of them are, for the most part, in these limits, but there are correlations, which appear to be stronger in a positive or a negative way. To give an illustration, a strong negative correlation can be observed between the ratio of the unemployed persons and the average disposable income (from -0.9654 in one or more regions) and a strong positive one between the ratio of the unemployed persons and the ratio of the persons that had received no education (up to 0.95539 in one or more regions).

	Pop. Den. **VIF**	MalePop. **VIF**	No Edu. **VIF**	Unemp. **VIF**	Income **VIF**	Public Emp. **VIF**	Immigrants **VIF**	Local **CN**
Min.	1.086	1.048	1.056	1.103	1.173	1.079	1.214	60.13
1st Qu.	1.596	1.492	1.711	1.856	2.595	1.726	2.314	175.92
Median	2.946	2.072	2.664	2.615	5.203	2.401	3.683	241.83
Mean	4.202	2.895	3.617	3.800	7.519	2.978	5.452	238.15
3rd Qu.	5.050	3.464	4.359	4.333	8.734	3.451	6.099	292.47
Max.	15.408	11.8	14.864	24.363	56.719	10.131	29.982	512.76

	Intercept **VDP**	Pop. Den. **VDP**	MalePop. **VDP**	No Edu. **VDP**	Unemp. **VDP**	Income **VDP**	Public Emp. **VDP**	Immigrants **VDP**
Min.	0.9290	0.0	0.0000253	0.0000028	0.0000008	0.0000094	0.0000026	0.0000024
1st Qu.	0.9880	0.009577	0.9602439	0.0301967	0.0525423	0.0327176	0.0479186	0.0144076
Median	0.9929	0.043560	0.9850653	0.1313164	0.1376033	0.1776283	0.1517867	0.0541607
Mean	0.9904	0.104623	0.8231779	0.1739089	0.2165381	0.2916835	0.1984943	0.1012766
3rd Qu.	0.9964	0.153072	0.9919803	0.2701334	0.3106525	0.4818653	0.3017067	0.1139111
Max.	0.9996	0.553014	0.9993527	0.6859043	0.8495389	0.9317484	0.8358627	0.6722277

Correlation	Intercept-Pop.Den.	Intercept-MalePop.	Intercept-NoEdu.	Intercept-Unemp.	Intercept-Income	Intercept-PublicEmp.	Intercept-Immigrants.
Min.	0	0	0	0	0	0	0
1st Qu.	0	0	0	0	0	0	0
Median	0	0	0	0	0	0	0
Mean	0	0	0	0	0	0	0
3rd Qu.	0	0	0	0	0	0	0
Max.	0	0	0	0	0	0	0
NA's	282	282	282	282	282	282	282

Correlation	Pop.Den.-MalePop.	Pop.Den.-NoEdu.	Pop.Den.-Unemp.	Pop.Den.-Income	Pop.Den.-PublicEmp	Pop.Den.-Imm.	MalePop.-NoEdu.	MalePop.-Unemp.
Min.	-0.2998	-0.49611	-0.41306	-0.39006	-0.61140	-0.2749	-0.76457	-0.74404
1st Qu.	-0.0282	-0.09774	-0.11436	0.05419	-0.32308	0.1964	-0.34985	-0.20210
Median	0.1009	0.04673	0.09182	0.18695	-0.15004	0.4164	-0.16169	-0.03901
Mean	0.1210	0.09100	0.11158	0.22412	-0.08422	0.4545	-0.14627	0.04935
3rd Qu.	0.2670	0.27123	0.36413	0.38103	0.14279	0.7595	0.06158	0.36030
Max.	0.7336	0.77173	0.52864	0.81674	0.75786	0.9253	0.88158	0.90359

Correlation	MalePop-Income	MalePop.-PublicEmp	MalePop.-Imm.	NoEdu.-Unemp.	NoEdu.-Income	NoEdu.-PublicEmp	NoEdu.-Imm.	Unemp.-Income
Min.	-0.8657	-0.92660	-0.67671	-0.70516	-0.87960	-0.89674	-0.53367	-0.9654
1st Qu.	-0.5317	-0.41075	-0.32409	-0.36585	-0.37518	-0.25689	-0.06189	-0.5929
Median	-0.2137	-0.19043	0.06812	-0.20053	-0.05181	-0.01165	0.28444	-0.3474
Mean	-0.1820	-0.20208	0.05704	-0.06886	0.06712	-0.01205	0.24609	-0.3530
3rd Qu.	0.1564	0.05224	0.40010	0.26629	0.62084	0.27456	0.56041	-0.2065
Max.	0.6285	0.34284	0.85545	0.95539	0.94671	0.74458	0.84578	0.5006

Correlation	Unemp.-PublicEmp	Unemp.-Imm.	Income-PulicEmp.	Income-Imm..	PublicEmp.-Imm.
Min.	-0.91781	-0.760646	-0.70497	-0.7497	-0.8323
1st Qu.	-0.35683	-0.322950	-0.33100	0.1650	-0.3539
Median	0.06910	0.004147	0.05566	0.4700	-0.2006
Mean	-0.03244	-0.009404	0.08120	0.3777	-0.1269
3rd Qu.	0.25318	0.287204	0.46711	0.7354	0.1187
Max.	0.70014	0.795975	0.89144	0.8739	0.6542

Table 6.8. *Multi-collinearity Diagnostics results.*

6.7. Shiny Application

All of the results that are presented and analyzed in this chapter, were produced by writing an R language script in an R Studio environment. This script was uploaded on a server, which is maintained by R Studio for the purpose of uploading Shiny applications. The application that has been created includes all of these outputs and can be accessed through the URL: https://artemistsiopa.shinyapps.io/shiny/.

The application is divided into 6 tabs. Each one of these tabs contains different measures. The first tab consists of the interactive thematic maps of each variable that is used in the analysis. In this tab there is a menu, from which the users are able to select the variable, which they wish to view. Each region's information appears by clicking inside the region's boundaries. The ability to zoom in, zoom out and zoom at the user's location is, also, provided. The second tab contains the descriptive statistics in tables and plots. The third tab presents the results of the spatial autocorrelation and the fourth tab the results of the inequality index. The fifth tab is, also, divided into three tabs, which can be selected from a side menu. The first one includes the results of the correlation coefficient, the second the results of the linear regression and the third tab presents the results of the GWR. The sixth tab of the application consists of information about the subject and the data that are used.

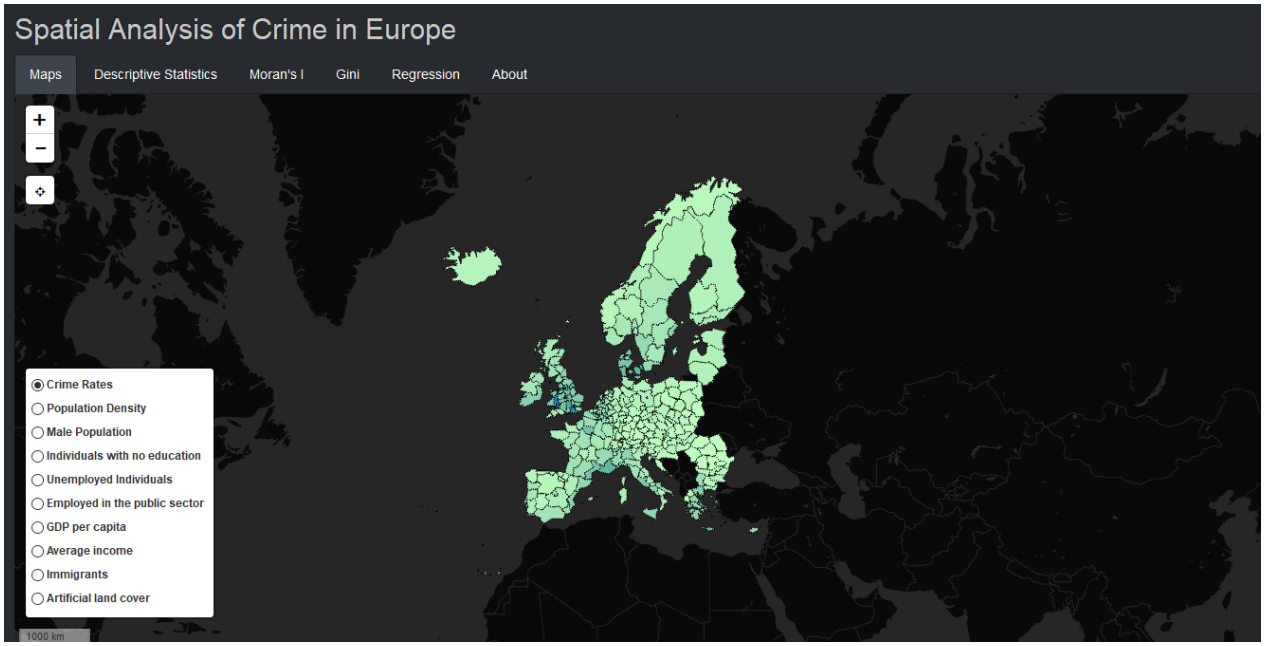

Figure 36. *Shiny application.*

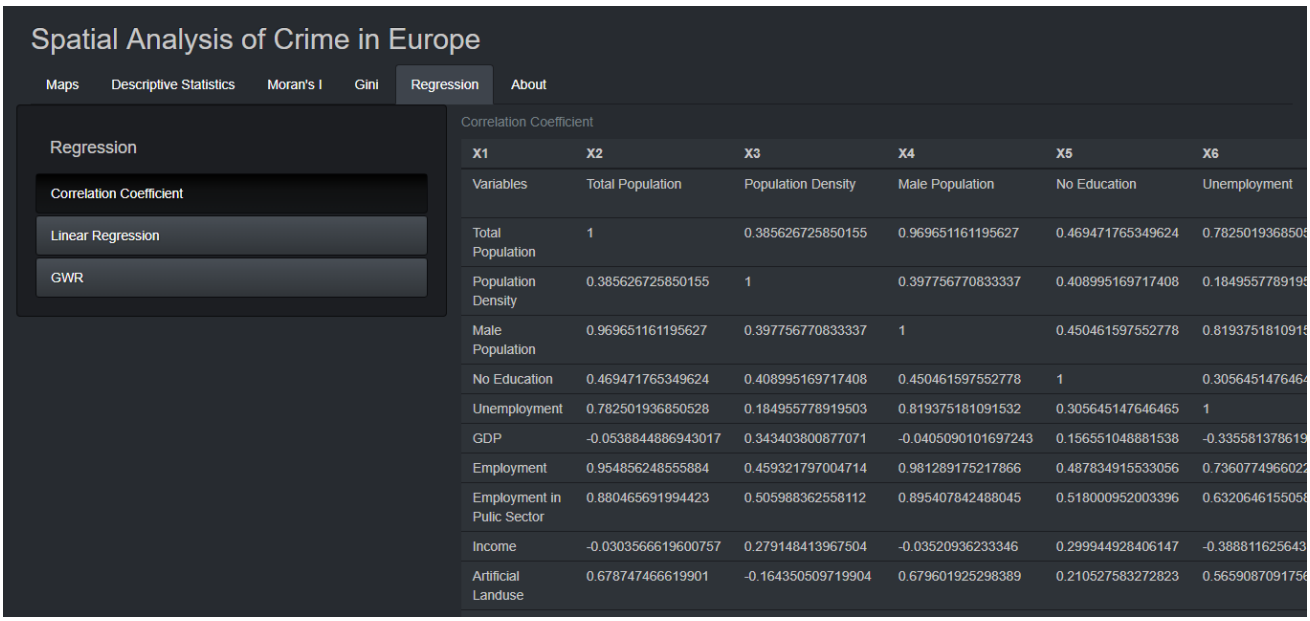

Figure 37. *Shiny application.*

7. Conclusions

Crime is one of the most challenging and concerning problems, which appear in every type of society. The effect that it has on every citizen is tremendous, since its presence creates a perceptible state of unsafety. The importance of the subject has been recognized by the scientific community, which has focused on examining the phenomenon. The different fields that are occupied with crime, concentrate their studies on different aspects of crime. Some of them study the crime patterns and others the criminal behavior. Many scientists focus on the geographic aspect of crime. Furthermore, the identification of prevention and reduction methods is, also, a field of study that stems from crime. What all these different types of study have in common, is the fact that they attempt to identify the factors that cause the increase or the reduction of crime rates. Consequently, many different theories concerning crime have been developed. Most of these theories emphasize on the effects of the environment. These effects can either construct criminal behaviors, or offer the opportunity for criminal activities. According to the reviewed literature, social disorganization, the social and economic characteristics of an area and the strains that they create can affect the crime rates in a great degree. With the evolution of technology, the processes of crime study and analysis have evolved. Furthermore, new methods are able to examine a great amount of data and aid the scientists into drawing more accurate conclusions (for example Big Data analysis, Machine Learning etc.).

This dissertation applies different methods of spatial analysis, in order to study the crime rates in the European Union's NUTS 2 regions, during the year 2010. Spatial analysis is essential because crime is a phenomenon with a strong geographical dimension. This study area consists of 284 regions with completely different characteristics. The characteristics-variables, whose effects on crime rates are studied, are the population density, the ratio of the male population that in 2010 was older than 15 years old and younger than 64 years old, the ratio of the immigrants in the total population, the ratio of the persons that had received no education, the ratio of the unemployed persons, the ratio of the employees in the public sector and the average disposable income. These variables were selected due to the fact that they were frequently used in the reviewed literature. The different methods and techniques used in this dissertation, in order to study the distribution of crime and its spatial structure and in order to provide empirical evidence for its determinants,

are: descriptive statistics, spatial autocorrelation, measures of inequality, regression and the Geographically Weighted Regression.

The first conclusion that can be drawn from the analysis, is that location matters in crime. The results of the spatial autocorrelation analysis, which show a positive autocorrelation, suggest that neighboring regions tend to have similar crime rates. This means that spatial patterns of similarly high and similarly low crime rates can be observed. The southern British regions and the Romanian regions, are examples of these spatial patterns. The first ones are included in a cluster of high crime rates and the latter present low crime rates. The same conclusion can be, also, drawn by the analysis of the inequality index's results. These results highlight the similarity of crime rates in neighboring regions. Regions, which are located in a greater distance from each other tend to have more unequal crime rates, confirming the First Law of Geography.

From the regression analysis, it was given prominence to the fact that two of the variables, the ratio of the unemployed persons and the disposable income, affect crime rates in a negative way, which means that when they increase, they reduce the crime rates. The factors that seem to influence each NUTS 2 region's crime rates in a greater degree are the ratio of the male population (in the age group 15-64), the ratio of the persons with no education, the ratio of the public employees and the ratio of the immigrants. However, not all of these variables are statistically significant. By combining these two parameters, it can be pointed out that the factors that affect the regional crime rates significantly are the presence of males in the age group 15-64, the number of the employees in the public sector and the educational attainment of the population. The population density can, also, be considered a significant factor.

The local regression models is notably superior to the global model, because its explanatory abilities are almost double (+90%). Moreover, the local model allows the spatial modification of the relationships. This is an essential element for the prominence of each European Union region's particular characteristics. One specific variable is able to affect the crime rates of one NUTS 2 region in a negative way and the crime rates of a different reason in a positive way. For example, the average disposable income affects the French region Brittany, in a negative way, but the French region Picardy in a positive way. These variances demonstrate the extremely different characteristics of every NUTS 2 region, and consequently even the differences between every country. In this case, it seems that the majority of the variables can potentially influence the crime rates in a great degree, since most of them are statistically significant. In particular, the most statistically important variables that, also, have a great variance of influence values are the ratio

of the male population (15-64), the ratio of unemployed persons and the ratio of immigrants. From these results, it can be concluded that in a local level the presence of males aged 15-64, the unemployment and the presence of immigrants can influence the crime rates in an exceptionally high degree. The results of the robust GWR, demonstrate that only the influence of the population density can be considered statistically important, even though its dynamic is insignificant.

8. References

Abler, R., Adams, J. S., Gould, P. (1971). *Spatial Organization: The Geographer's view of the world*, USA: Prentice-Hall Inc.

Ackerman, W. V., Murray, A. T. (2004). *Assessing spatial patterns of crime in Lima, Ohio*, Cities, 21(5): 423-437. https://doi.org/10.1016/j.cities.2004.07.008

Agnew, R. (2001). *Building on the foundation of general strain theory: Specifying the types of strain most likely to lead to crime and delinquency,* Journal of research in crime and delinquency, 38(4): 319-361.

Agnew, J. A., Livingstone, D. N. (ed.) (2011). *The SAGE Handbook of Geographical Knowledge*, London: Sage Publications Ltd.

Ahmadi, M. (2003). *Crime mapping and spatial analysis,* International Institute for Geo-information Science and Earth Observation (Master Thesis). [Retrieved October 29, 2017 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.6128&rep=rep1&type=pdf]

Alevizaki, Ch. G. (2010). *The spatial analysis of crime: The case of the city of Volos,* Volos: University of Thessaly (MSc Thesis) (In Greek). [Retrieved November 1, 2017 from http://ir.lib.uth.gr/]

Allaire, J.J. (2011). *RStudio: Integrated Development Environment for R,* The R User Conference, useR! 2011.

Almeida, E. D., Haddad, E. A., Hewings, G. J. D. (2003). *The spatial pattern of crime in Minas Gerais: An explanatory analysis,* Brazil: The University of São Paulo, Regional and Urban Economics Lab.

Altindag, D. T. (2012). *Crime and unemployment: Evidence from Europe,* International review of Law and Economics, 32(1): 145-157. https://doi.org/10.1016/j.irle.2011.10.003

Anselin, L., Cohen, J., Cook, D., Gorr, W., Tita, G. (2000). *Spatial analyses of crime,* Measurement and analysis of crime and justice, 4: 213-262.

Arnold, R. A. (2007). *Economics,* Thessaloniki: Epikentro (In Greek).

Arnot, M. L., Usborne, C. (2001). *Gender and crime in modern Europe,* London: UCL Press.

Bakirli, E. (2005). *Youth and violence in the contemporary society. An ecological analysis of its forms, its effects and its confrontation*, Athens: Panteion University of Social and Political Sciences (MSc Thesis) (In Greek). [Retrieved November 1, 2017 from http://library.panteion.gr/]

Block, R. L., Block, C. R. (1995). *Space, place, and crime: Hot spot areas and hot places of liquor-related crime,* Crime and Place, 145-183. [Retrieved November 1, 2017 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.362.1183&rep=rep1&type=pdf]

Boba, R. (2001). *Introductory guide to crime analysis and mapping,* USA: Department of Justice, Office of Community Oriented Policing Services. [Retrieved October 29, 2017 from https://ric-zai-inc.com/Publications/cops-w0273-pub.pdf]

Brantingham, P. L., Glasser, U., Kinney, B., Singh, K., Vajihollahi, M. (2005). *A computational model for simulating spatial and temporal aspects of crime in urban environments,* Systems, Man and Cybernetics, 2005 IEEE International Conference.

Brunsdon, C., Fotheringham, A. S., Charlton, M. E. (1996). *Geographically Weighted Regression: A method for exploring spatial nonstationarity*, Geographical Analysis, 28(4): 281-298.

Brunsdon, C., Corcoran, J., Higgs, G. (2007). *Visualising space and time in crime patterns: A comparison of methods,* Computers, Environment and Urban Systems, 31(1): 52-75. https://doi.org/10.1016/j.compenvurbsys.2005.07.009

Cameron, J. G. (2001). A s*patial analysis of crime in Appalachia,* National Institute of Justice (NIJ)/ National

Carballo-Cruz, F. (2011). *Causes and consequences of the Spanish economic crisis: Why the recovery is taken so long?,* Panoeconomicus, 3: 309-328. [Retrieved February 5, 2018 from http://www.doiserbia.nb.rs/img/doi/1452-595X/2011/1452-595X1103309C.pdf]

Carcach, C. (1999). *The spatial analysis of crime statistics and crime mapping: Methodological issues,* Australia: Crime in rural Australia Conference. [Retrieved November 1, 2017 from http://www.aic.gov.au/media_library/conferences/rural/carcach.pdf]

Chainey, S., Ratcliffe, J. (2005). *GIS and crime mapping,* England: John Wiley & Sons, Ltd.

Charlton, M., Fotheringham, A. S. (2009). *Geographically weighted regression* (white paper). [Retrieved January 29, 2018 from https://www.geos.ed.ac.uk/~gisteac/fspat/gwr/gwr_arcgis/GWR_WhitePaper.pdf]

Cohen, L. E., Felson, M. (1979). *Social change and crime rate trends: A routine activity approach,* American Sociological Review, 44: 588-608.

Commission on Crime Prevention and Criminal Justice (2017). *World crime trends and emerging issues and responses in the field of crime prevention and criminal justice,* Vienna. [Retrieved November 22, 2017 from https://www.unodc.org/unodc/en/data-and-analysis/statistics/reports-on-world-crime-trends.html]

Congdon, P. D. (2013). *A model for spatially varying crime rates in English districts: The effects of social capital, fragmentation, deprivation and urbanicity,* International journal of Criminology and Sociology, 2: 138-152.

Cracolici, M. F., Uberti, T. E. (2008). *Geographical distribution of crime in Italian Provinces: A spatial econometric analysis,* Review on Regional Research, 29(1): 1-28. https://doi.org/10.1007/s10037-008-0031-1

Criminal Justice Reference Service (NCJRS). [Retrieved November 1, 2017 from https://www.ncjrs.gov/pdffiles1/nij/grants/189560.pdf]

Dean, T. (2001). *Crime in Medieval Europe: 1200-1550,* New York: Routledge.

De Smith, M. J., Goodchild, M. F., Longley, P. A. (2007). *Geospatial analysis: A comprehensive guide to principles, techniques and software tools,* United Kingdom: Troubador Publishing Ltd.

Del Frate, A. A. (1998). *Victims of crime in the developing world,* Italy: United Nations Interregional Crime and Justice Research Institute (UNICRI).

Diamanti, Th. E. (2010). *Unemployment and crime in Greece*, Volos: University of Thessaly (MSc Thesis) (In Greek). [Retrieved November 1, 2017 from http://ir.lib.uth.gr/]

Eck, J. E., Chainey, S., Cameron, J. G., Leitner, M., Wilson, R. E. (2005). *Mapping crime: Understanding hot spots,* National Institute of Justice. [Retrieved October 18, 2017 from http://discovery.ucl.ac.uk/]

Entner Wright, B. R., Caspi, A., Moffitt, T. E., Silva, P. A. (2001). *The effects of social ties on crime by criminal propensity: A life-course model of interdependence,* Criminology, 39(2): 321-352.

Entorf, H., Spengler, H. (2000). *Socioeconomic and demographic factors of crime in Germany: Evidence from panel data of the German states,* International Review of Law and Economics, 20: 75-106.

Entorf, H., Spengler, H. (2002). *Crime in Europe: Causes and consequences,* Berlin: Springer.

Fajnzylber, P., Lederman, D., Loayza, N. (2002). *Inequality and violent crime,* Journal of Law and Economics, 45(1): 1-39, doi: 10.1086/338347.

Fajnzylber, P., Lederman, D., Loayza, N. (2002). *What causes violent crime?,* European Economic Review, 46(7): 1323-1357.

Farrington, D. P., Welsh, B. C. (2007). *Saving children from a life of crime: Early risk factors and effective interventions,* New York: Oxford University Press.

Field, S. (1992). *The effect of temperature on crime,* The British Journal of Criminology, 32(3): 340-351. https://doi.org/10.1093/oxfordjournals.bjc.a048222

Fischer, M. M., Getis, A. (2010). *Handbook of applied spatial analysis: Software tools, methods and applications,* Berlin: Springer.

Formosa, S. (2007). *Spatial analysis of temporal criminality evolution: An environmental criminology study of crime in the Maltese islands,* United Kingdom: University of Huddersfield (PhD Thesis). [Retrieved November 1, 2017 from http://eprints.hud.ac.uk/id/eprint/964/1/sformosathesis2007.pdf]

Fotheringham, A. S., Rogerson, P. (1994). *Spatial analysis and GIS*, United Kingdom: Taylor & Francis Ltd.

Fotheringham, A. S., Charlton, M. E., Brunsdon, C. (1998). *Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis,* Environment and Planning, 30: 1905-1927. [Retrieved January 29, 2018 from https://pdfs.semanticscholar.org/12e2/3a0643b893b84bd3ba7f8642575f1faaf3ea.pdf]

Fotheringham, A. S., Brunsdon, C. (1999). *Local forms of spatial analysis,* Geographical Analysis, 31(4): 340-358. [Retrieved January 29, 2018 from http://www.csiss.org/gispopsci/workshops/2011/PSU/readings/Fotheringham-and-Brunsdon-1999.pdf]

Fotheringham, A. S., Brunsdon, C., Charlton, M. (2000). *Quantitative Geography: Perspectives on spatial data analysis,* London: Sage Publications Ltd.

Fotheringham, A. S., Rogerson, P. (2009). *The SAGE handbook of spatial analysis,* London: Sage Publications Ltd.

Fox, J., Weisberg, S. (2013). *Robust Regression,* USA: University of Minnesota. [Retrieved February 10, 2018 from http://users.stat.umn.edu/~sandy/courses/8053/handouts/robust.pdf]

Goldsmith, V., McGuire, P. G., Mollenkopf, J. H., Ross, T. A. (2000). *Analyzing crime patterns: Frontiers of practice,* California: Sage Publications Inc.

Goodchild, M. F., Anselin, L., Apperbaum, R. P., Herr Harthorn, B. (2000*). Toward spatially integrated social science*, International Regional Science Review, 23(2): 139-159.

Gould, E. D. Weinberg, B. A., Mustard, D. B. (2002). *Crime rates and local labor market opportunities in the United States: 1979-1997*, The Review of Economics and Statistics, 84(1): 45-61. [Retrieved November 1, 2017 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.196.1964&rep=rep1&type=pdf]

Groff, E. R., La Vigne, N. G. (2002). *Forecasting the future of predictive crime mapping,* Crime Prevention Studies, 13: 29-57.

Gruszczynska, B. (2004). *Crime in Central and Eastern European countries in the enlarged Europe,* European Journal on Criminal Policy and Research, 10: 123-136.

Hakim, S., Buck, A. J. (1989). *Do casinos enhance crime?,* Journal of Criminal Justice, 7: 409-416.

Harrendorf, S., Heiskanen, M., Malby, S. (2010). *International statistics on crime and justice,* Vienna: United Nations Office on Drugs and Crime (UNODC).

Harries, K. (2006). *Extreme spatial variations in crime density in Baltimore County, MD,* Geoforum, 37(3): 404-416.

Harris, P., Fotheringham, A. S., Juggins, S. (2010). *Robust Geographically Weighted Regression: A technique for quantifying spatial relationships between freshwater acidification critical loads and catchment attributes,* Annals of the Association of American Geographers, 100(2): 286-306. https://doi.org/10.1080/00045600903550378

Hirschi, T., Gottfredson, M. (1983). *Age and the explanation of crime,* The American Journal of Sociology, 89(3): 552-584.

Huang, C. C., Laing, D., Wang, P. (2004). *Crime and poverty: A search-theoretic approach,* International Economic Review, 45(3): 909-938.

Ihaka, R., Gentleman, R. (1996). *R: A language for data analysis and graphics,* Journal of Computational and Graphical Statistics, 5(3): 299-314.

International Association of Crime Analysts (2011). *Crime pattern definitions for tactical analysis.*

International Centre for the Prevention of Crime (2010). *International report on crime prevention and community safety: Trends and perspectives, 2010.*

Johnson, C. P. (2000). *Crime mapping and analysis using GIS,* Geomatics 2000: Conference on Geomatics in Electronic Governance. [Retrieved November 1, 2017 from http://fac.ksu.edu.sa/sites/default/files/crim_mapping.pdf]

Kalogirou, S. (2001). *The statistical analysis and modelling of internal migration flows within England and Wales,* United Kingdom: Leicester University (PhD Thesis). [Retrieved January 29, 2018 from http://gisc.gr/en/publications/]

Kalogirou, S. (2015). *Spatial Analysis: Methodology and applications with R,* Athens: Hellenic Academic Libraries Link (In Greek).

Kappeler, V. E., Potter, G. W. (2017). *The mythology of crime and criminal justice: Fifth Edition,* Illinois: Waveland Press, Inc.

Kling, J. R., Ludwig, J., Katz, L. F. (2004). *Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment,* The Quarterly Journal of Economics, 120(1): 87-130. https://doi.org/10.1162/0033553053327470.

Kobogianni, S. (2012). *The geography of urban crime: Mapping perpetrators and offences in the city of Volos*, Volos: University of Thessaly (MSc Thesis) (In Greek). [Retrieved November 1, 2017 from http://ir.lib.uth.gr/]

Krivo, L. J., Peterson, R. D. (1996). *Extremely disadvantaged neighborhoods and urban crime,* Social Forces, 75(2): 619-648.

Kuo, F. E., Sullivan, W. C. (2001). *Environment and crime in the inner city: Does vegetation reduce crime?,* Environment and Behavior, 33(3): 343-367. https://doi.org/10.1177/0013916501333002

Leitner, M. (ed.) (2013). *Crime modeling and mapping using geospatial technologies*, London: Springer.

Lochner, L. (2007). *Education and crime*, International Encyclopedia of Education (3rd edition). [Retrieved November 1, 2017 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.709.4910&rep=rep1&type=pdf]

Longley, P. A., Goodchild, M. F., Maguire, D. J., Rhind, D. W. (2005). *Geographical Information Systems: Principles, techniques, management and applications*, USA: John Wiley & Sons Inc.

Longley, P. A., Goodchild, M. F., Maguire, D. J., Rhind, D. W. (2010). *Geographical Information Systems and Science,* Athens: Klidarithmos (In Greek).

Lopes, S. B., Brondino, N. C. M., Rodrigues da Silva, A. N. (2007). *Exploratory and Confirmatory Spatial Data Analysis tools in transport demand modeling,* Brazil: 10th International Conference on computers in urban planning and urban management. [Retrieved January 30, 2018 from http://redpgv.coppe.ufrj.br/index.php/pt-BR/producao-da-rede/artigos-cientificos/2007-1/300-062/file]

Lu, B., Charlton, M., Harris, P., Fotheringham, A. S. (2012). *Geographically weighted regression with non-Euclidean distance metric: A case study using hedonic house price data,* International Journal of Geographical Information Science. DOI: 10.1080/13658816.2013.865739

Ludwig, J., Duncan, G. J., Hirschfield, P. (2001). *Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment,* The Quarterly Journal of Economics, 116(2): 655-679. https://doi.org/10.1162/00335530151144122

Malik, A. A. (2016). *Urbanization and crime: A relational analysis,* IOSR Journal of Humanities and Social Science (IOSR-JHSS), 21(1): 68-74.

Marselli, R., Vannini, M. (1997). *Estimating a crime equation in the presence of organized crime: Evidence from Italy,* International Review of Law and Economics, 17: 89-113.

Matsueda, R. L., Grigoryeva, M. S. (2014). *Social inequality, crime and deviance.* [Retrieved October 18, 2017 from http://faculty.washington.edu/matsueda/]

Moberg, D. O. (1953). *Old age and crime,* Journal of Criminal Law and Criminology, 43(6): 764-776.

Monmonier, M. (1996). *How to lie with maps*, USA: The University of Chicago Press.

Murataya, R., Gutiérrez, D. R. (2013). *Effects of weather on crime,* International Journal of humanities and social science, 3(10): 71-75.

National Crime Prevention Council (2003). *Crime prevention through environmental design (Guidebook).*

Osgood, D. W. (2000). *Poisson-based regression analysis of aggregate crime rates,* Journal of Quantitative Criminology, 16(1): 21-43.

Pizam, A. (1982). *Tourism and Crime: Is there a relationship?,* Journal of travel research, 20(3): 7-10.

Rajkhan, S. F. (2014). *Women in Saudi Arabia: Status, Rights and Limitations,* University of Washington Bothell (Master Thesis). [Retrieved November 22, 2017 from https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/25576/Rajkhan%20-%20Capstone.pdf?sequence=1]

Raphael, S., Winter-Ebmer, R. (1998). *Identifying the effect of unemployment on crime,* University of California. [Retrieved October 18, 2017 from http://escholarship.org/uc/item/5hb4h56g]

Raphael, S., Winter-Ebmer, R. (2001). *Identifying the effect of unemployment on crime,* Journal of Law and Economics, 44(1): 259-283.

Ratcliffe, J. (2010). *Crime mapping: Spatial and temporal challenges,* In: Piquero A., Weisburd, D. (eds) *Handbook of Quantitative criminology*, New York: Springer.

Rey, S. J., Smith, R. J. (2012). *A spatial decomposition of the Gini coefficient,* Letters is Spatial and Resource Sciences, 6(2): 55-70.

Robinson, A. H., Morrison, J. L., Muehrcke, P. C., Kimerling, A. J., Guptill, S. C. (2002). *Elements of Cartography,* Athens: National Technical University of Athens Press (In Greek).

Rogerson, P. A. (2001). *Statistical methods for Geography,* London: Sage Publications Ltd.

Roinioti, E. (2009). *Our knowledge about crime based on quantitative data*, The Greek review of social research, 129(B): 33-59. [Retrieved November 1, 2017 from https://ejournals.epublishing.ekt.gr/index.php/ekke/article/viewFile/6770/6498.pdf]

Rufrancos, H. G., Power, M., Pickett, K. E., Wilkinson, R. (2013). *Income inequality and crime: A review and explanation of the time-series evidence,* Social Criminol, 1(1).

Sampson, R. J., Groves, W. B. (1989). *Community structure and crime: Testing social-disorganization theory,* The American journal of sociology, 94(4): 774-802.

Sampson, R. J., Raudenbush, S. W., Earls, F. (1997). *Neighborhoods and violent crime: A multilevel study of collective efficacy,* Science, 277(5328): 918-924, doi: 10.1126/science.277.5328.918.

Sampson, R. J., Laub, J. H. (1993). *Crime in the making: Pathways and turning points through life,* Cambridge: Harvard University Press.

Siegmunt, O. (2016). *Neighborhood disorganization and social control: Case studies from three Russian cities,* Berlin: Springer.

Shapiro, S. S., Wilk, M. B. (1965). *An analysis of variance test for normality (complete samples),* Biometrika, 52(3/4): 591-611. [Retrieved February 8, 2018 from http://webspace.ship.edu/pgmarr/Geo441/Readings/Shapiro%20and%20Wilk%201965%20-%20An%20Analysis%20of%20Variance%20Test%20for%20Normality.pdf]

Sherman, L. W., Gartin, P. R., Buerger, M. E. (1989). *Hot spots of predatory crime: Routine activities and the criminology of place,* Criminology, 27(1): 27-56, doi: 10.1111/j.1745-9125.1989.tb00862.x.

Symeonaki, M. (2015). *Statistics for everyone with SPSS,* Thessaloniki: Sofia (In Greek).

Tobler, W.R. (1970). *A computer movie simulating urban growth in the Detroit region,* Economic Geography, 46(2): 234-240.

Tompson, L., Partridge, H., Shepherd, N. (2009). *Hot routes: Developing a new technique for the spatial analysis of crime,* Crime Mapping: A journal of research and practice, 1(1): 77-96.

Tsatsaris, A. (2017). *Applications of Geomatics in the Health Sector,* Lecture, Athens: Harokopio University (In Greek).

Winkelstein, W. (2007). *The strange case of the Broad Street pump: John Snow and the mystery of cholera,* JAMA, 297(22): 2529-2533.

Wortley, R., Mazerolle, L. (2008). *Environmental criminology and crime analysis,* New York: Routledge.

Zairis, P.E. (2010). *Statistical Methodology,* Athens: Kritiki (In Greek).

Zakaria, S., Rahman, N. A. (2016). *The mapping of spatial patterns or property crime in Malaysia: Normal mixture model approach,* Journal of Business and Social Development, 4(1): 1-11.

Zar, J. H. (2005). *Spearman rank correlation,* USA: John Wiley & Sons Ltd. [Retrieved January 31, 2018 from ftp://biostat.wisc.edu/pub/chappell/800/hw/spearman.pdf]

Zhang, H., Peterson, M. P. (2007). *A spatial analysis of neighborhood crime in Omaha Nebraska using alternative measures of crime rates,* Internet Journal of Criminology. [Retrieved October 29, 2017 from https://pdfs.semanticscholar.org/139a/6a864b9a30a0b346d7517f138b5059b6c089.pdf]

Zhong, H., Yin, J., Wu, J., Yao, S., Wang, Z., Lv, Z., Yu, B. (2011). *Spatial analysis for crime patterns of metropolis in transition using police records and GIS: A case study of Shanghai, China,* International Journal of digital content technology and its applications, 5(2): 93-105.

Zhu, L., Gorman, D. M., Horel, S. (2004). *Alcohol outlet density and violence: a geospatial analysis*, Alcohol & Alcoholism, 39(4): 369-375, doi:10.1093/alcalc/agh062.

Central Statistics Office [Available at http://www.cso.ie/en/]

Eurostat [Available at http://ec.europa.eu/eurostat]

Federal Statistical Office [Available at https://www.bfs.admin.ch/bfs/en/home.html]

Geodata.gov.gr [Available at http://geodata.gov.gr/el/]

Hellenic Statistical Authority [Available at http://www.statistics.gr/en/home]

Office for National Statistics [Available at http://webarchive.nationalarchives.gov.uk]

Provincial Government, Principality of Liechtenstein [Available at https://www.llv.li/]

R Studio [Available at https://www.rstudio.com/]

Scotland's official statistics [Available at http://statistics.gov.scot/]

Shiny from R Studio [Available at http://shiny.rstudio.com/]

Statistical Office of the Republic of Slovenia [Available at http://www.stat.si/StatWeb/en]

Statistical State Office Saxony [Available at https://www.statistik.sachsen.de/]

Statistics Denmark [Available at http://www.dst.dk/en]

Statistics Iceland [Available at http://www.statice.is/]

Statistics Norway [Available at http://www.ssb.no/en]

Statistics Sweden [Available at http://www.scb.se/en/]

The European Free Trade Association [Available at http://www.efta.int/]

The R Project for Statistical Computing [Available at https://www.r-project.org/]

Thematic Mapping [Available at http://thematicmapping.org/downloads/world_borders.php]

World Bank Group [Available at https://data.worldbank.org/indicator/VC.IHR.PSRC.P5]

Appendix

Country	*Number of NUTS 2 Regions*	*Average Area of NUTS 2 Regions (km^2)*	*Average Population of NUTS 2 Regions*
Belgium	11	2,775.2	985,446
Bulgaria	6	18,500.3	1,236,961
Czech Republic	8	9,858.1	1,307,761
Denmark	5	8,619.7	1,106,948
Germany	38	9,399.2	2,150,810
Estonia	1	45,227	1,333,290
Ireland	2	34,898.5	2,274,714
Greece	13	10,150.5	855,330
Spain	19	26,631.1	2,446,664
France	22	28,629.7	2,852,965
Croatia	2	28,297	2,151,424
Italy	21	14,349.3	2,818,578
Cyprus	1	9,251	819,140
Latvia	1	64,559	2,120,504
Lithuania	1	65,300	3,141,976
Luxembourg	1	2,586	502,066
Hungary	7	13,289.6	1,430,618
Malta	1	316	414,027
Netherlands	12	3,461.9	1,381,249
Austria	9	9,319.9	927,960
Poland	16	19,542.4	2,376,429
Portugal	7	13,173.1	1,510,497
Romania	8	29,798.8	2,682,773
Slovenia	2	10,136.5	1,023,488
Slovakia	4	12,259.3	1,347,603
Finland	5	67,688.1	1,070,285
Sweden	8	55,168.4	1,167,585
United Kingdom	37	6,715.7	1,580,110
Iceland	1	103,000	1,119,494
Liechtenstein	1	160.5	35,894
Norway	7	46,254.6	694,028
Switzerland	7	5,897.8	1,112,258

Correlation Coefficient Results.

	Population Density	*Ratio of Male Population*	*Ratio of persons with No Education*	*Ratio of Unemployed persons*	*GDP per capita*	*Ratio of employees in the public sector*	*Average Disposable Income*	*Ratio of Artificial Land Cover*	*Ratio of Immigrants*
Population Density	1	0.0149	0.2758	-0.2347	0.348	0.1296	0.2687	0.8762	0.2502
Ratio of Male Population	0.0149	1	-0.2536	0.1588	-0.1916	-0.4174	-0.3732	-0.1555	-0.1483
Ratio of persons with No Education	0.2758	-0.2536	1	-0.1387	0.1652	0.1005	0.317	0.2953	0.3339
Ratio of Unemployed persons	-0.2347	0.1588	-0.1357	1	-0.6311	-0.4396	-0.6993	-0.243	-0.2112
GDP per capita	0.348	-0.1916	0.1652	-0.6311	1	0.5011	0.8314	0.4146	0.6248
Ratio of employees in the public sector	0.1296	-0.4174	0.1005	-0.4396	0.5011	1	0.4865	0.2468	0.0978
Average Disposable Income	0.2687	-0.3732	0.317	-0.6993	0.8314	0.4865	1	0.3458	0.5542
Ratio of Artificial Land Cover	0.8762	-0.1555	0.2953	-0.243	0.4146	0.2468	0.3458	1	0.3155
Ratio of Immigrants	0.2502	-0.1483	0.3339	-0.2112	0.6248	0.0978	0.5542	0.3155	1