

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676060

LONGPOP

**Methodologies and Data mining techniques for the
analysis of Big Data based on Longitudinal Population
and Epidemiological Registers**

Compilation of different GIS layers on mortality in Madrid and Spain

Deliverable n. 2.2

Disclaimer: This publication reflects only the author's view and the Research Executive Agency is not responsible for any use that may be made of the information it contains.

1 Background

1.1 Evolution of Madrid in the context of the early 20th Century

Unlike many other European capitals, the history of Madrid, as both a capital and large city, is relatively recent. Philip II moved the Habsburg court to the city in 1561, effectively making it the capital of what was then a vast empire. However, Madrid remained small, until the early to mid 1800s, when, after the removal of several in-migration laws, the city began to grow at a rapid rate. In the 19th and 20th centuries, the growth of the city's physical size could not meet the demands of the increased population size—crowding was an issue, and infectious disease epidemics (cholera, small pox, yellow fever, etc.) were frequent [1, 2].

As the city expanded, a plan, called the Ensanche of Madrid, was designed to cope with new migrants by expanding the city through the addition of several neighborhoods, which, in a sense, were intended to divide the city by social class [3]. In the north and east, neighborhoods such as Argüelles, Chamberí, Salamanca, and Retiro were designed for middle and upper class individuals, while those areas such as Arganzuela in the south became home to poorer individuals with less economic and social capital. This division was also not simply of an economic nature, but also reflected the differences in places from where the migrants had come. Many of the poorest in-migrants to Madrid came from both the regions immediately surrounding the city and the south of Spain, particularly Andalusia [4]. Most generally, those migrants from the north and east of Spain had higher professional qualifications of other reasons, beyond purely seeking work and economic relief, to come to the city. This is due to would-be migrants in the north instead being pulled towards other industrial cities, such as Bilbao and Barcelona.

1.2 Spain

Similar to Madrid, the entire country of Spain was quite diverse during the early 20th century and most generally, the country could largely be divided into swaths of poor, rural areas and a few large industrial cities. Also as in the case of Madrid, economic inequality also divided the country into a small wealthy class and many poor individuals [5]. Although it began later than in other European countries, Spain continued to progress through the Demographic Transition as mortality rates declined [6, 7]. Politically, the country was also unstable, and Alfonso XIII's power was overthrown several years following the conclusion of the pandemic, in 1923 by Primo de Rivera who ruled as a dictator. The southern area of the country, particularly the region of Andalusia, suffered severe economic stress and instability, given the high relative importance of agriculture in its economy. Moreover, compared to the rest of Spain, the area was underdeveloped and lacked major industry. As World War I broke out with Spain as a neutral party, industrialized parts of the country began to become major suppliers of goods to other European countries involved in the war. Increasing demand led to an influx of migrants from poorer rural areas to these industrial hubs.

These factors thus resulted in the aforementioned exodus of large segments of the population from the south and regions surrounding Madrid towards the capital city. Other rural areas in the north also witnessed increased out-migration, although these individuals largely went to cities nearby, with greater cultural and political similarities to their own, such as Barcelona and Bilbao [3]. After all, during this time of near constant political crisis, a huge movement of regionalism grew among the population as well. Another group of migrants from Southern and Northwestern Spain as well as Portugal moved from the country to work in mines in southern France that had previously employed men now fighting in the war. All of this migratory movement was facilitated by a large network of private railways that spanned the country [8]. It is the return of these men to their homes, near the end of the war, that is believed to have brought the fall wave of the epidemic to Spain [9, 10, 11]. This transmission is visible in some of the daily mortality maps found in section 4, and was the subject of a large body of work completed by ESR 1 that was disseminated in several conferences and workshops in 2019 and is expected to be published in the future.

1.3 Visualization

The result of the city's growth, to the extent that the Ensanche was executed, means that several variables within the context of the social, demographic, and economic development of Madrid are clearly visible when

presented on a map (as displayed using a GIS layer) at the neighborhood level. The following maps come from two data sources that allow for a broad overview of Madrid from a “birds eye view” at the beginning of the 20th century, when Madrid was administratively divided into 10 districts containing a total of 100 neighborhoods, visible in figure 1. The data represents 91 of these 100 neighborhoods; the 9 remaining neighborhoods are not displayed due to data issues with mortality records, which led to estimation issues that the authors of this report did not feel were accurate.

The data in the following layers is displayed in map form using the `sp` package in R Studio using a blue-green colorblind friendly scale. While the number and place of cut points differ based on measure, the specific data from which the maps have been created is available as layers in a shapefile created from the 1905 geography of Madrid. For more information regarding both the creation of the specific digitized map by administrative boundaries and other similar historical indicators, please refer to HISDI-MAD [12].

1.4 Current work in the LONGPOP project with these layers

ESR 1 (Laura Cilek) has created and used these layers in conjunction with her own work on the Spanish Influenza Pandemic in Madrid throughout the course of the LONGPOP project. In both cases, the mortality graphics created (the animated mortality maps) have been used in conference presentations and in personal (ESR’s own) and professional (LONGPOP) social media accounts to disseminate results and attract interest to the project.

More specifically, the layers related to Madrid (sections 2 and 3) were used to help understand the relationship between each neighborhood’s make up and composition (in such variables as those displayed here) and excess mortality during the Spanish flu. While these layers are extremely spatially correlated, there was no specific spatial correlation in the excess mortality during the pandemic in each neighborhood. Further analyses revealed that in fact, the amount of underlying mortality and excess mortality in each preceding wave was the greatest indicator of excess mortality in the successive waves. That is to say, neighborhoods with high baseline mortality and high excess mortality during the spring and fall waves had higher overall and wave-specific excess mortality during the epidemic. The culmination of this analysis is part of a chapter of the ESR’s dissertation and efforts will be made to publish the paper in an open access, peer reviewed journal in the coming months.

The layers presented in section 4 were also used in a chapter of ESR 1’s dissertation and are part of a paper that will soon be submitted for review. The data behind the layers was used in a breakpoint analysis (segmented regression) in order to statistically find the start, peak, and end dates of the wave in each province, capital, and rural provincial population. The results show that the wave moved from the northeast of the country along railroad lines to Portugal and the southwest part of the Iberian peninsula, and that areas with known spring outbreaks had smaller fall outbreaks than those without a herald wave. Sequence analysis broke each geography into one of five categories based on the amount of time spent in and point of transition to each “state” of the outbreak (before, ascending, descending, and after phases).

2 GIS layers: neighborhood distribution, population and density measures, social indicators

2.1 1915 Padrón

While every year, the city released information on births, deaths, and migratory movements in the Madrid, the 1915 padrón is a particularly unique resource due to its detailed data on specific economic, social, and demographic variables, as well as information about the number of buildings, their purpose, and rental prices in each neighborhood [13]. The subsequent maps provide a visual depiction of this information for 91 of the 100 neighborhoods in the city.

While the total population of each neighborhood is in figure 2, more information regarding the actual age structure of the city can be found in figures 3, 4, and 5. Particularly of note, the highest proportion of children under five can be found in two neighborhoods, Guindalera and Doctor Forquet. Doctor Forquet was home to both the foundling hospital and other places where children may be found, such as orphanages and “asilos.” The

Guindalera neighborhood was less populated and developed, but home to the Fundación Caldeiro, which was a live-in religious school that would likely contribute to a higher proportion of population under five [14, 12].

In this subsection, maps regarding the gross number of buildings per neighborhood, in addition to various density measures in the city can be found (figures 6, 7, and 8). Note that some neighborhoods with particularly large areas, especially in the outside neighborhoods of Madrid, have a higher number of buildings, but when used as a measure of density, the highest levels of population per building or room, with the exception of the Retiro neighborhood, are found in the so-associated “poorer” areas. This is understandable, given the noted overcrowding and reports of the time. In fact, one building in the southwestern district of Latina was reported to have had more than 1,000 people living in it, nearly all from the same nearby city of Toledo [4]. Of note to the population per room density measure, the Retiro barrio, which is primarily occupied by a large park and therefore contains fewer buildings, also housed a large number of military barracks in the southern part below the park. This also likely contributed to the higher density measure in this neighborhood.

A few other economic variables provide insight into both the make up of the city *and* the remarkable differences between north and south, in line with the noted expansion. Unsurprisingly, the density of stores (figure 9) is concentrated in the commercial center, which at the time of the 1915 padrón, was undergoing a large and rapid transition, as the now iconic Gran Vía was being constructed. Outside of the direct center, rent (figure 10) was also highest in the areas of Salamanca, Goya, and in the neighborhood of Retiro, which consists of the area surrounding the park. Rents in the southwest corner of the city were lowest.

The average HISCAM score (figure 11), a measurement often used as a proxy of social class, based on occupation, appears to be a little more diverse. Those areas on the outskirts of the city limits tend to have slightly lower scores, but this may be a relic of occupations of those living there, which may have still been slightly more rural in nature. However, of note are the high average values in the Argüelles, Chamberí, Salamanca, and Goya neighborhoods *relative* to other areas. Thus, while the data may does not show to the extent of the other layers the poorer areas in the south, it does highlight the middle and upper class neighborhoods.

The percentage of literate individuals (figure 12) over the age of nine appears to reflect two things about the city’s development during the time. The first is the lower relative percentage of those who could not read or write in the neighborhoods on the periphery of the cities. This is likely a relic of the remaining countryside and those still living in and working in occupations associated with rural areas as the city sprawl quickly moved outward. Another notable aspect of the distribution of the literate population is the clear clockwise gradient in literacy among neighborhoods; starting from the bottom right corner of the map, those in the poorer southern and south western areas had lower proportions of literate individuals, followed by those in the perceived middle class neighborhoods in the central north and north west, before finally moving to the group of neighborhoods with highest rates (Retiro park aside) in the center and north through Chamberí.

Figure 1: Neighborhoods by District

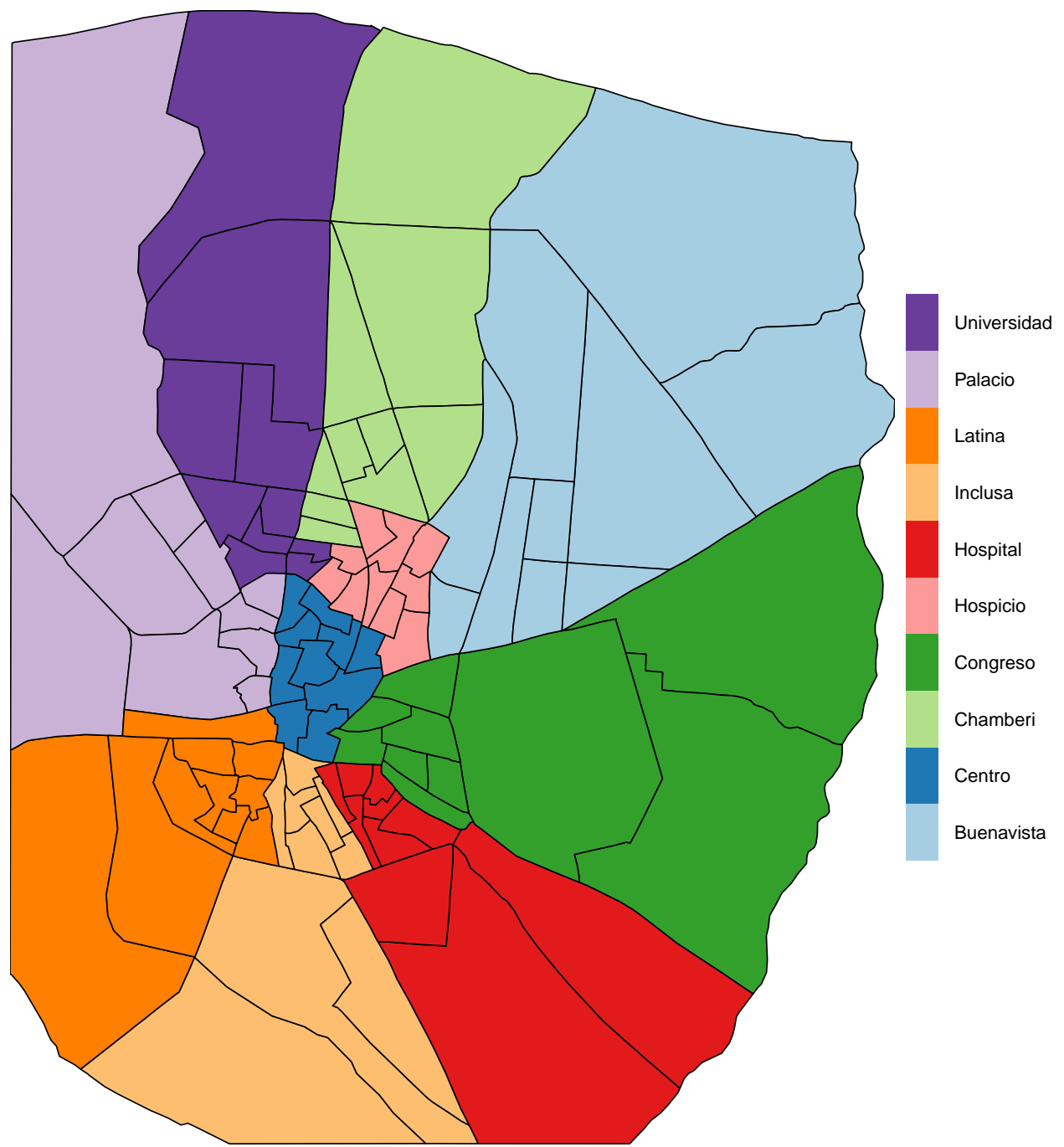


Figure 2: Total Population per Neighborhood

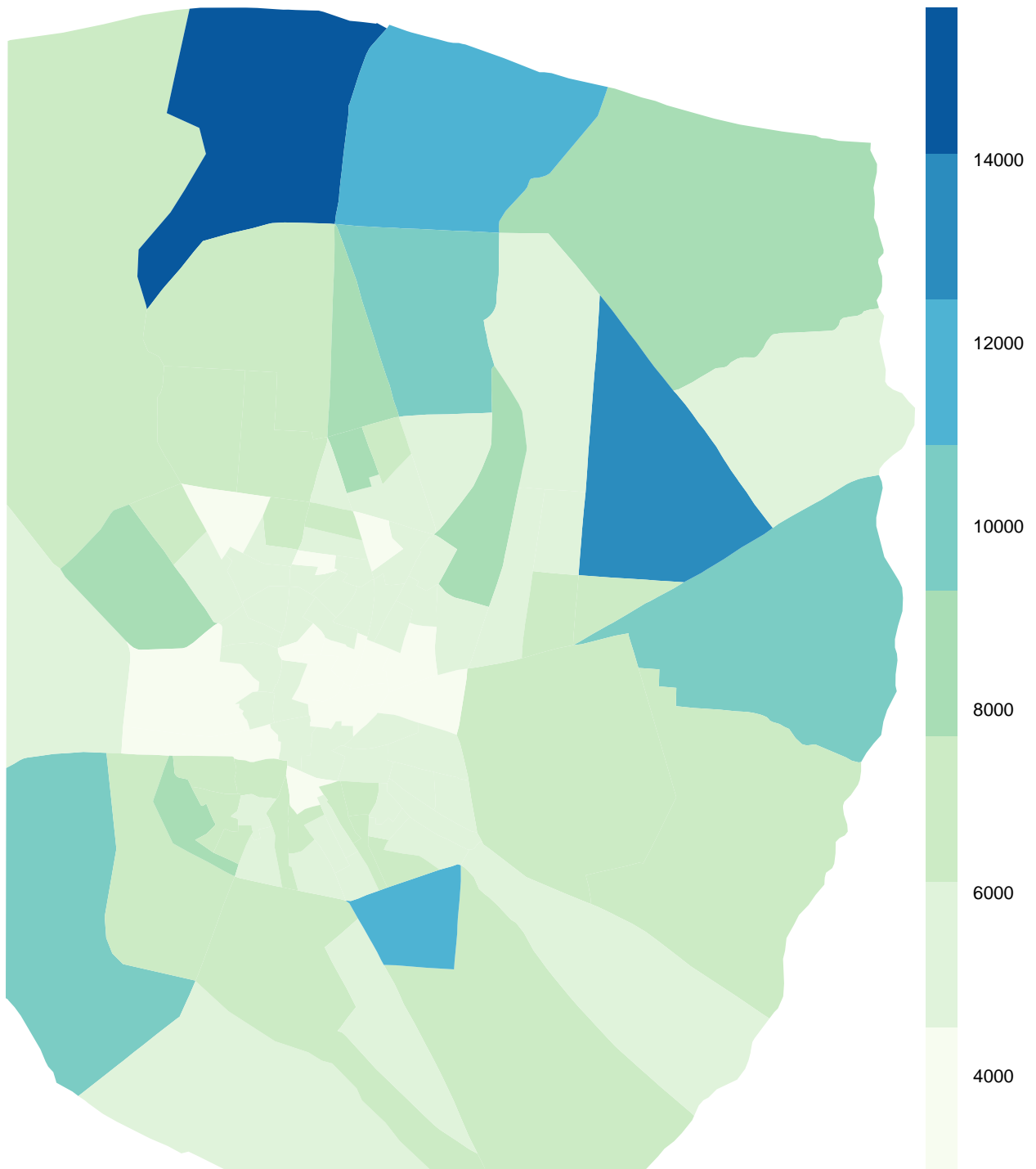


Figure 3: Proportion of Population under 5

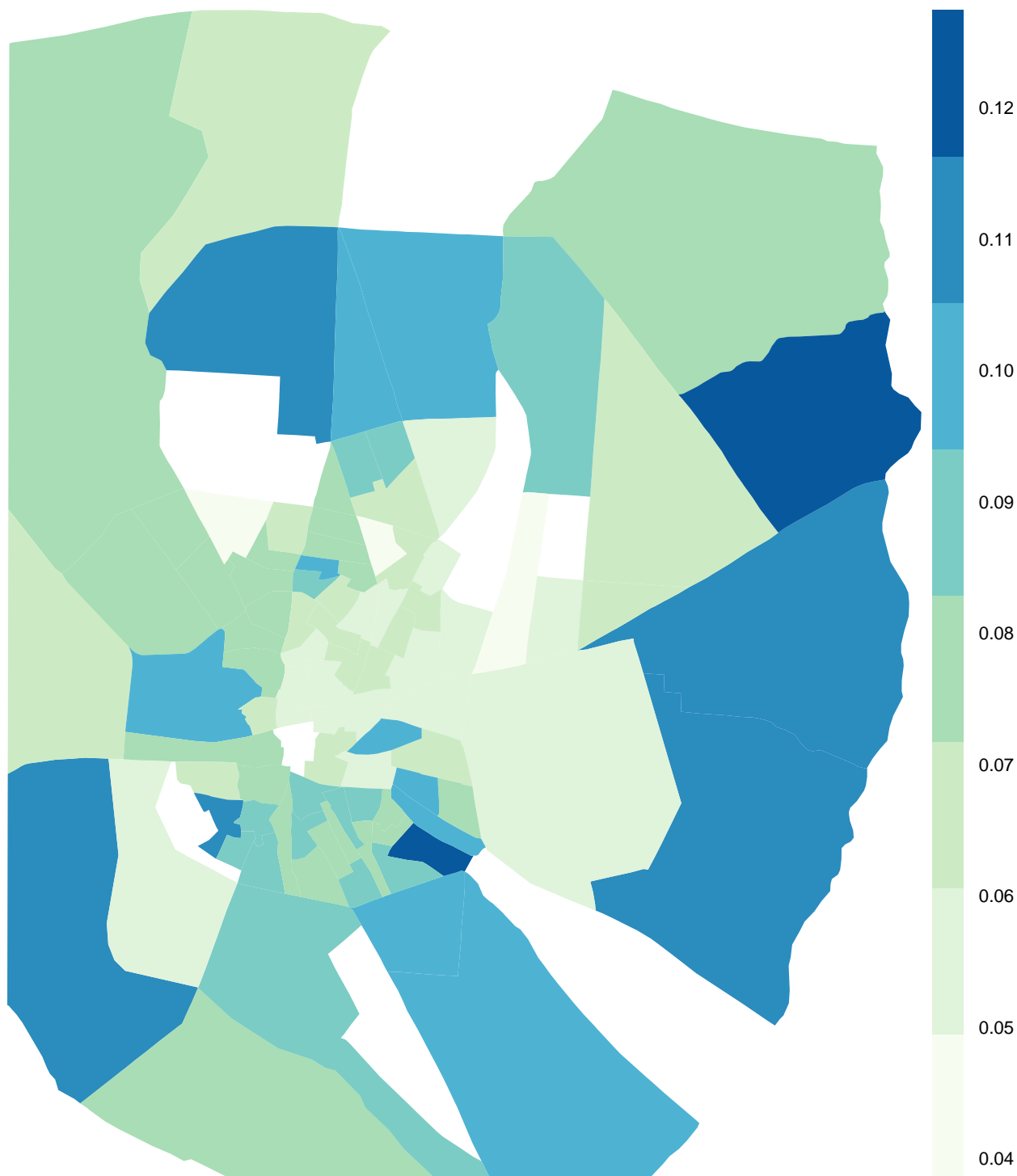


Figure 4: Proportion of Population between 5 and 75

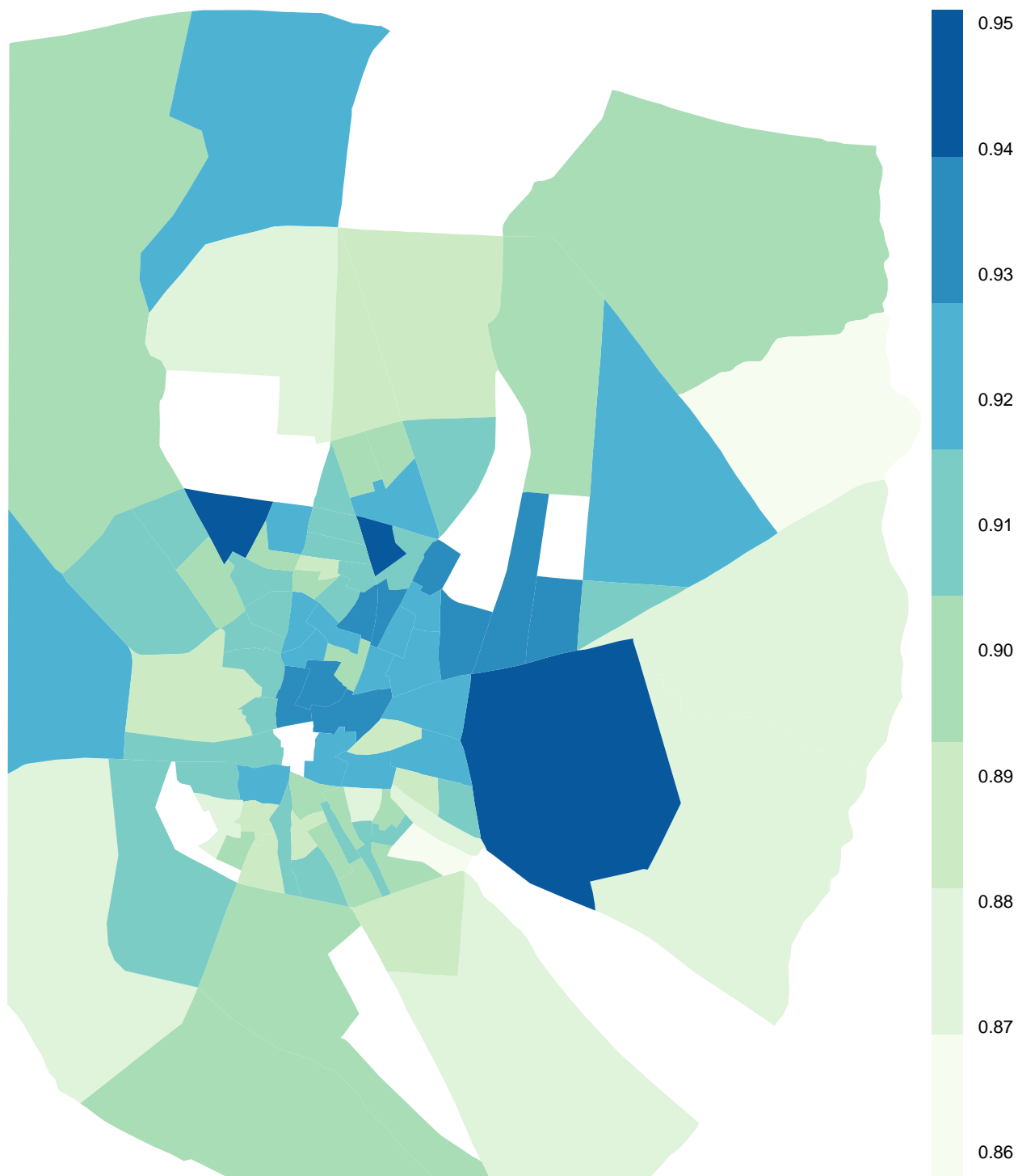


Figure 5: Proportion of Population over 75

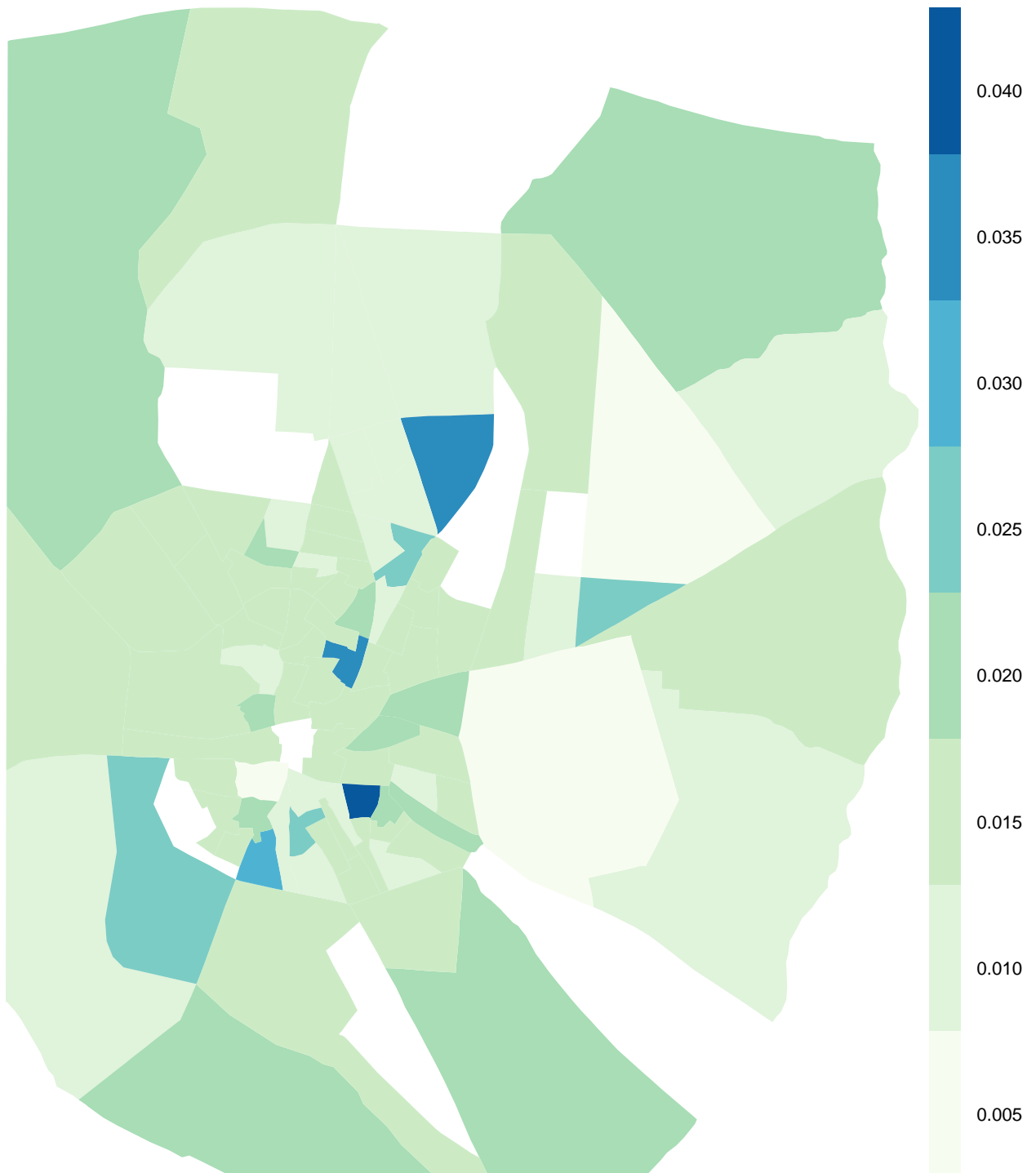


Figure 6: Total Buildings per Neighborhood

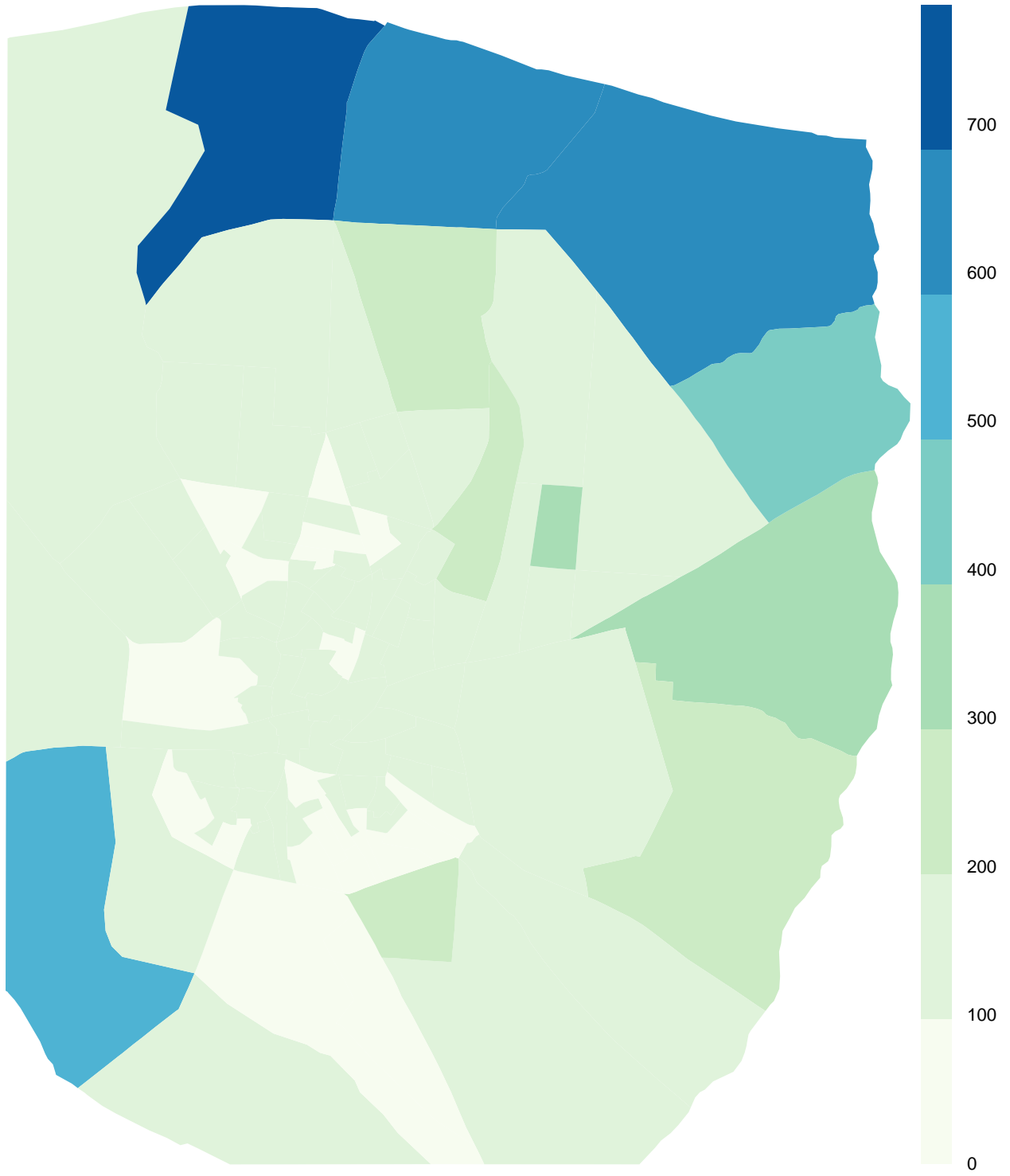


Figure 7: Population Density per Building

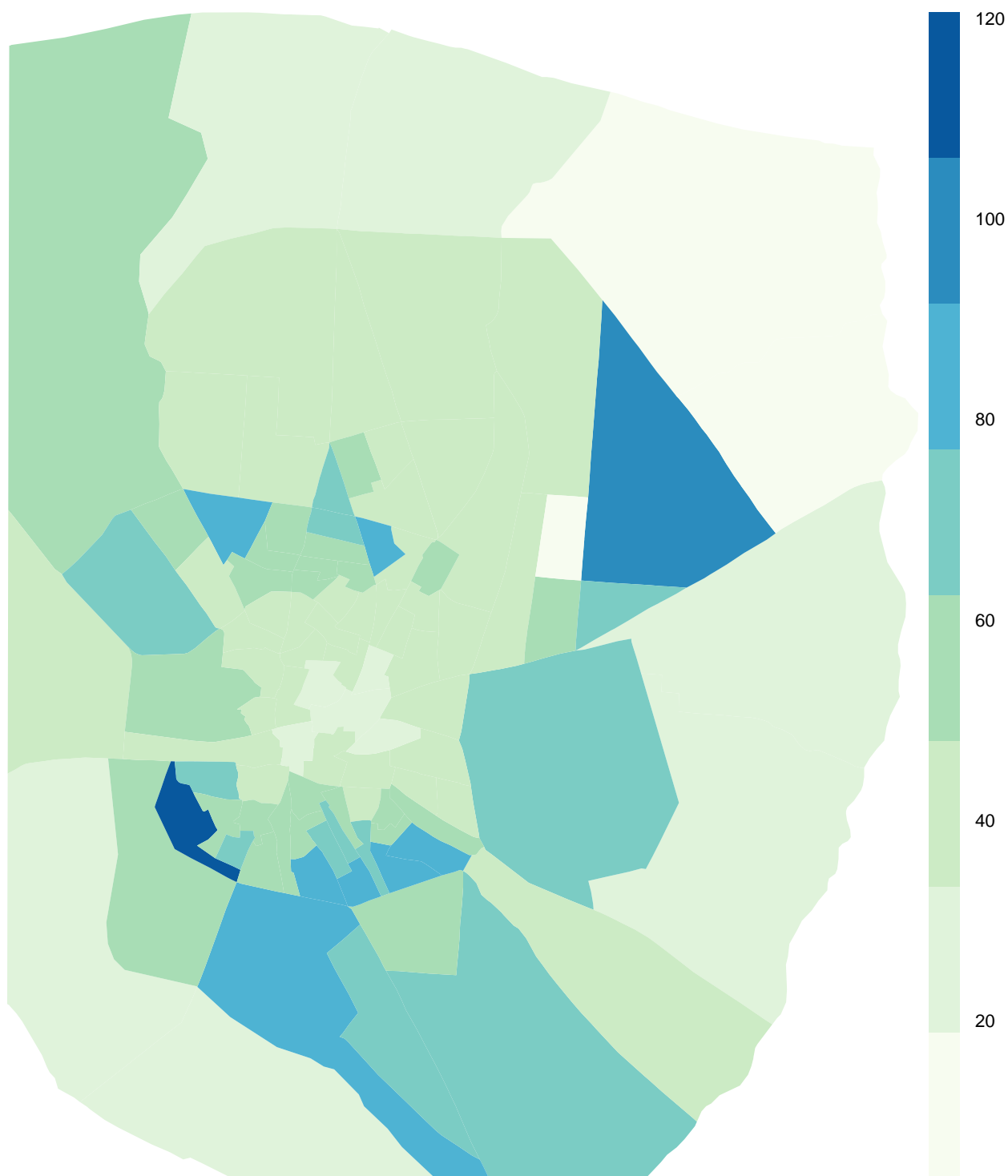


Figure 8: Population Density per Room

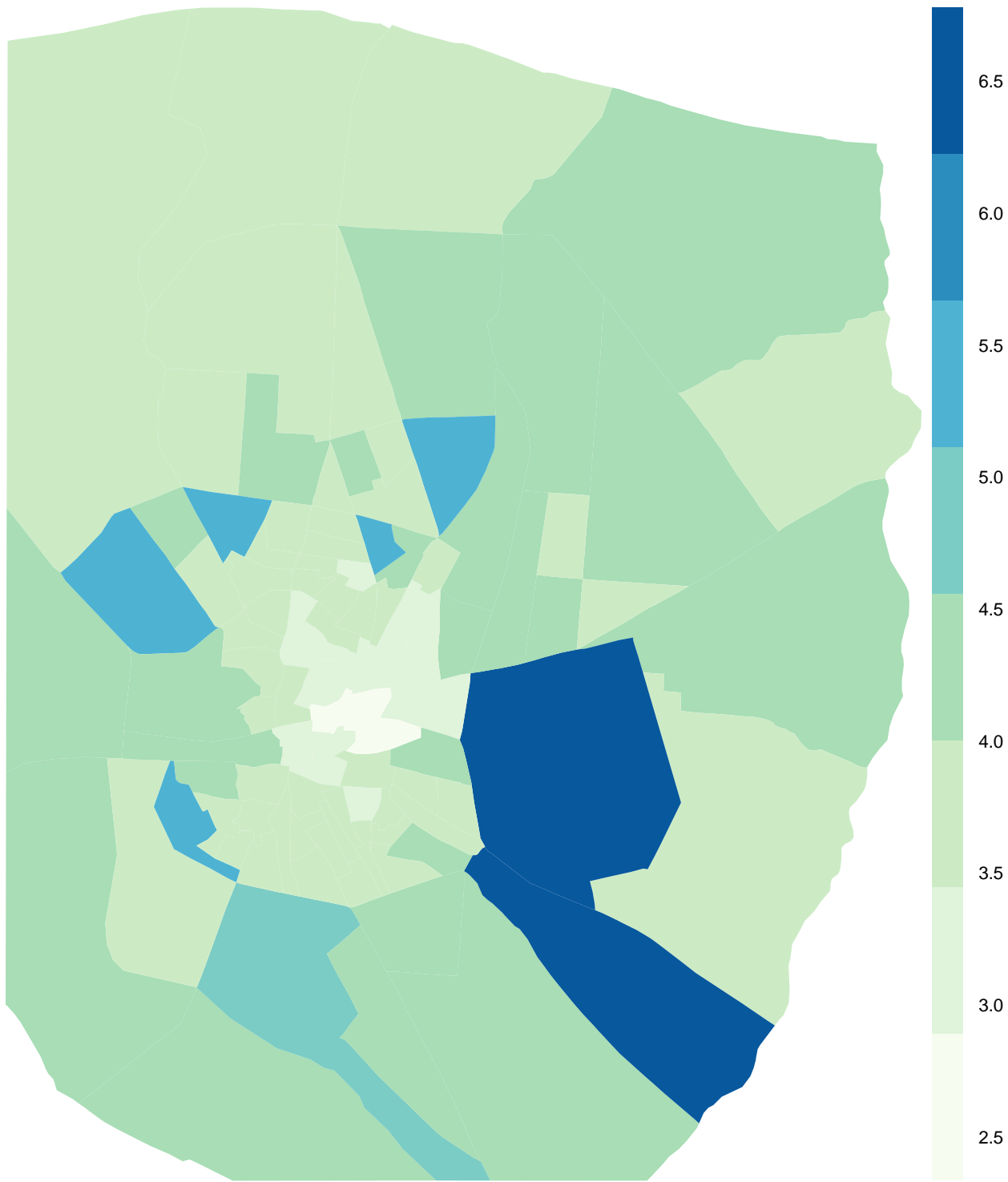


Figure 9: Density of Stores

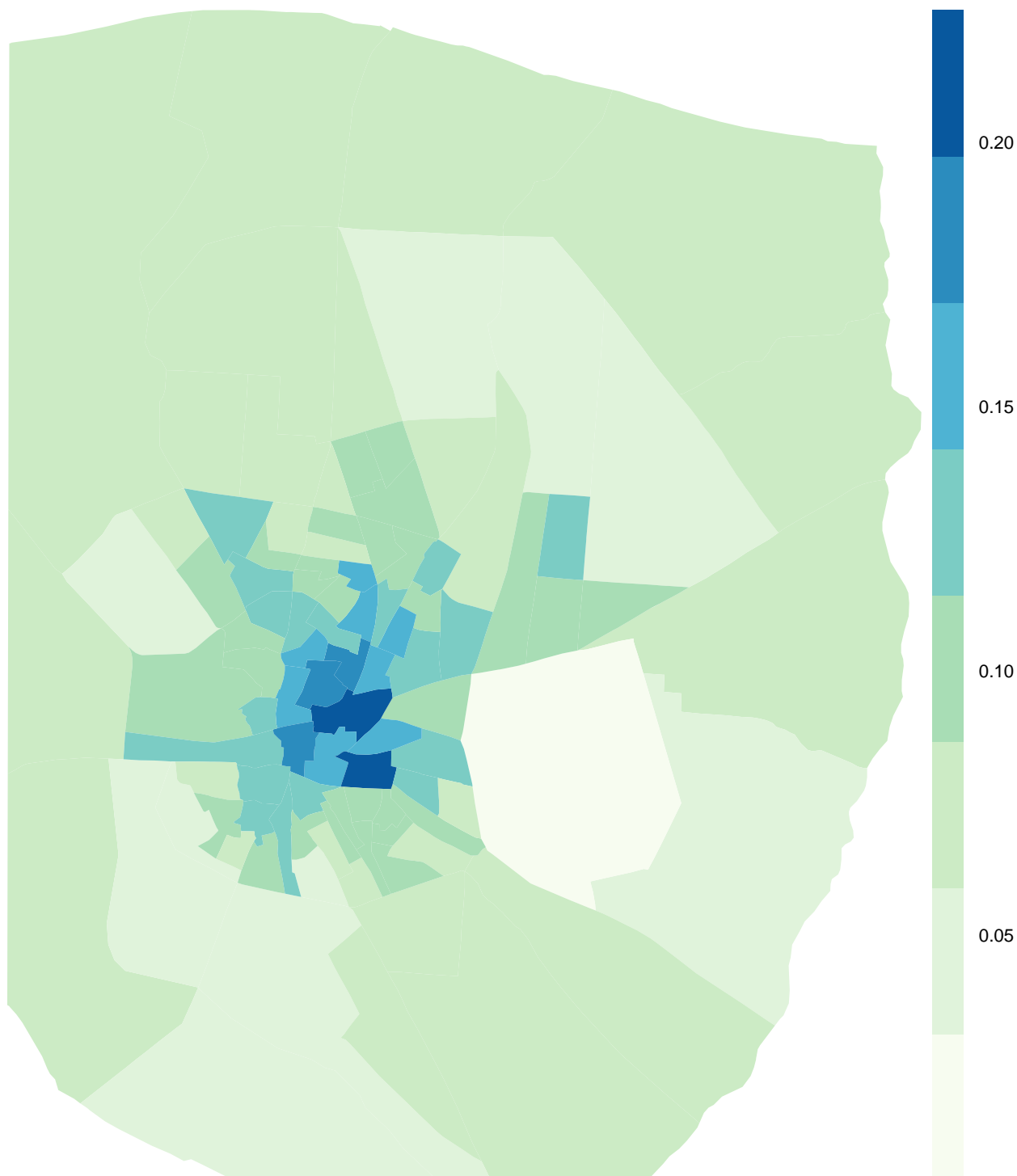


Figure 10: Average Rent (weighted by categories)

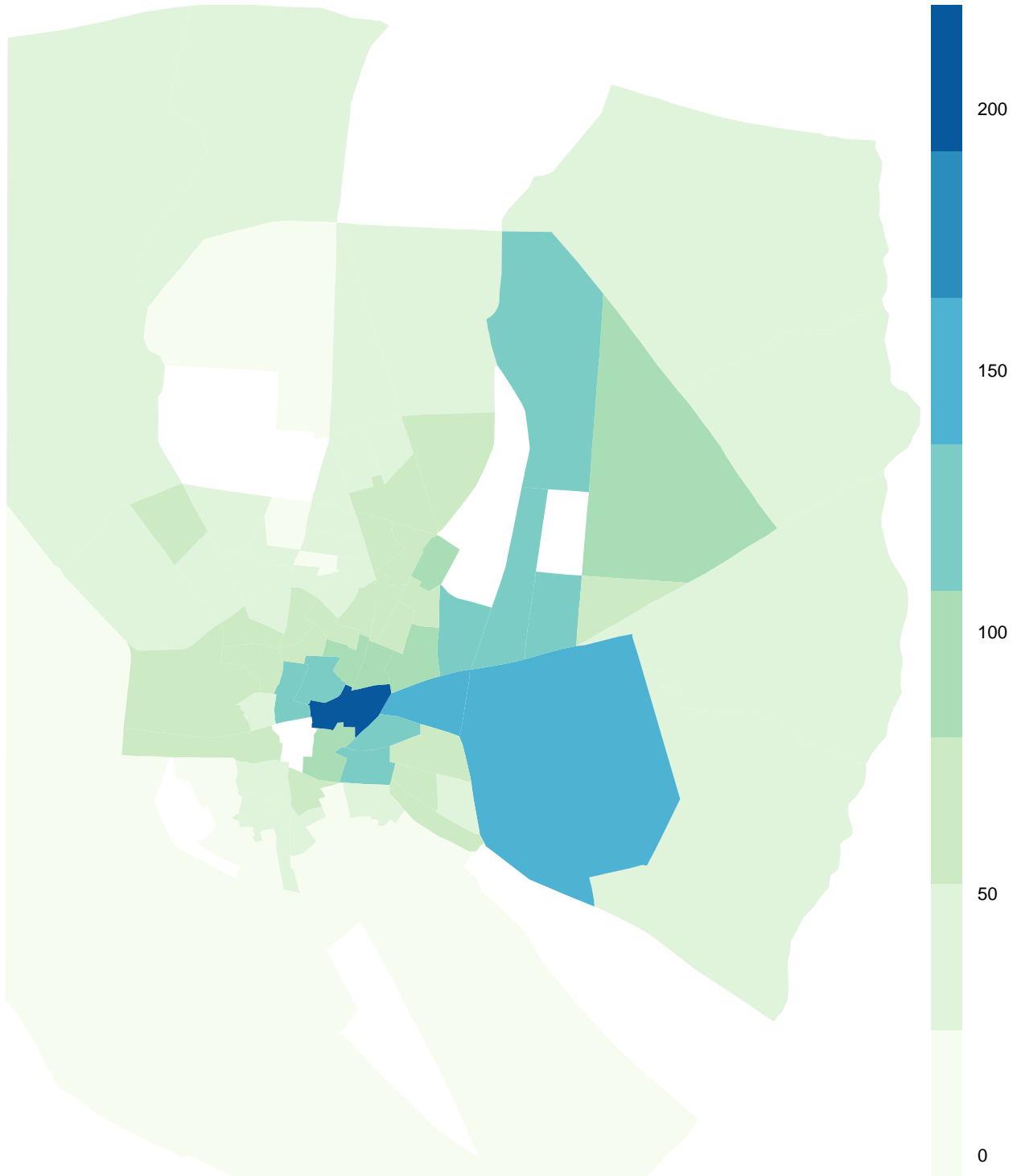


Figure 11: Average HISCAM score

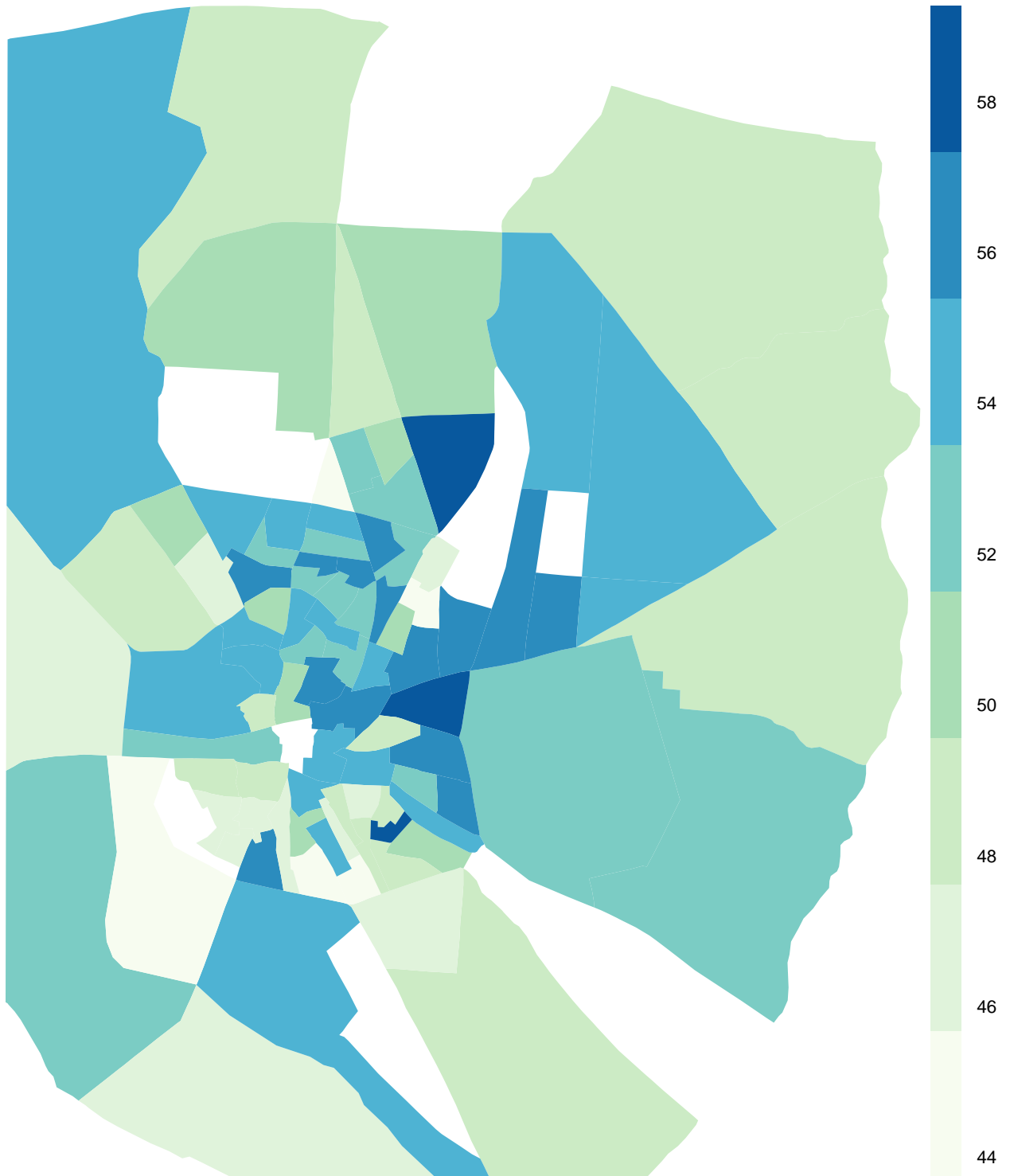
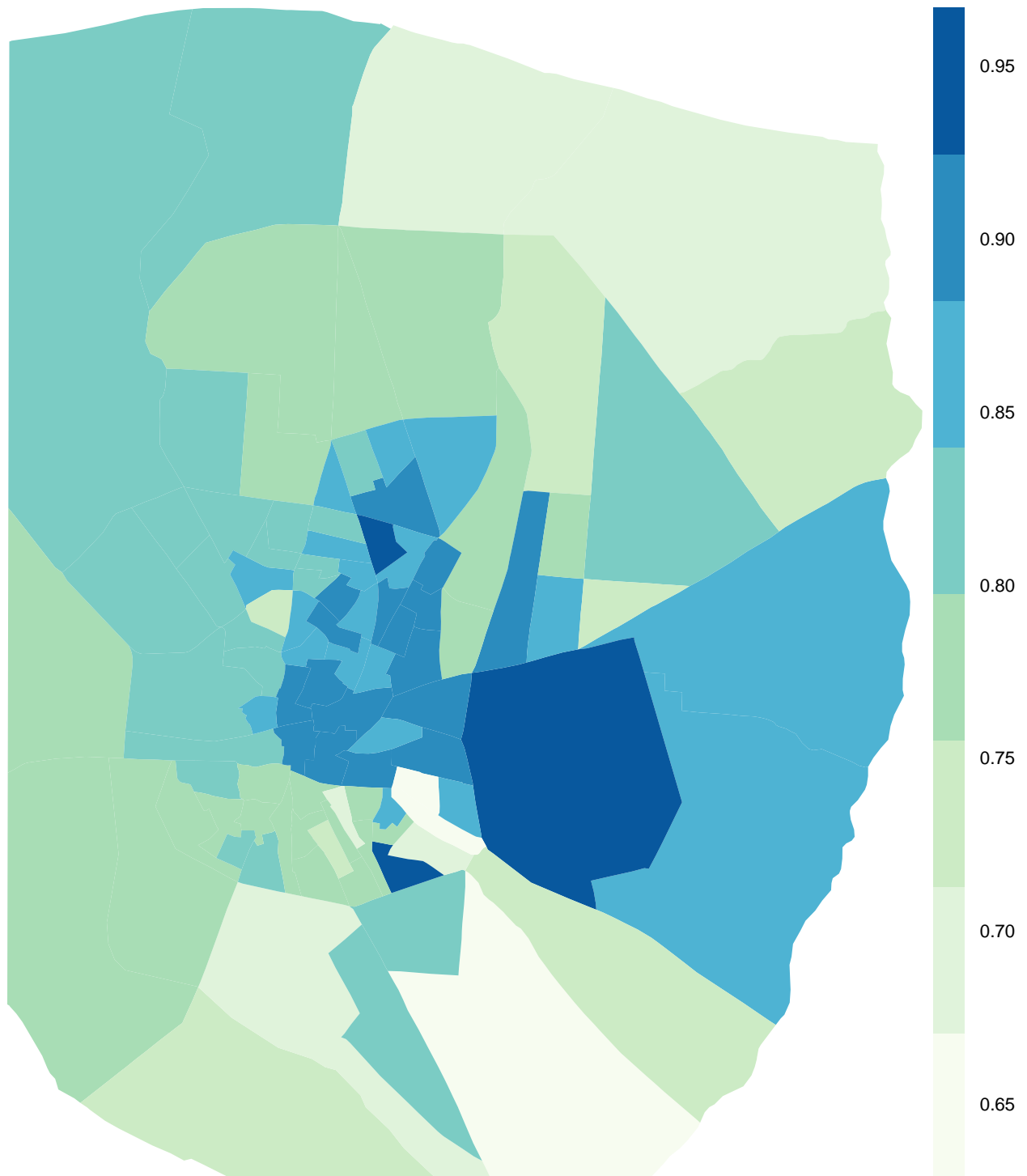


Figure 12: Proportion of Literate (read and write) Population over 9



3 GIS layers: mortality in Madrid

In order to record and visualize weekly mortality information for the city, we use the death records as recorded in the Madrid Civil Registry and retrieved from the Madrid city archives [15]. This second source of data provides detailed mortality information regarding on those who died in the city of Madrid between 1917 and 1922. The geographic information contains the address at death, sometimes mistakenly coded as the residential address, which can then be coded to a neighborhood, as well as the district in which the death was registered, as they were registered in district offices.

Moreover, mortality information is usually interesting in the context of where the deceased was living or worked, which therefore may limit the interpretable ability of any analyses. That is to say, deaths recorded and registered at a hospital address that was not in the same neighborhood in which a person lived would create an added bias to results from any neighborhood-level analysis. Here, we present a side by side comparison between of those deaths occurring in and outside of 140 identified medical centers in Madrid for each week between January 1917 and December 1922 (figure 14). For the calculations used to show the standardized excess mortality (observed / expected level), only deaths occurring *outside* of these medical centers are used (figure 13). As in the previous examples, due to missing and inaccurate information in these death records, only mortality information for 91 of the 100 neighborhoods will be presented in this deliverable, of which the accuracy is presumed to be high.

Moreover, the data was collected and digitized specifically with the intention of quantifying excess mortality during the Spanish Influenza Pandemic in the city of Madrid. Between May 1918 and February 1920, Madrid experienced four separate waves of epidemic influenza, of which the excess mortality for all causes and ages by neighborhood is presented in figure 13. Due to its recrudescence nature, the fall and winter waves of 1918-1919 have been combined under the label “Fall/Winter”, while the first wave (Herald) and last wave (Echo) are also visualized. The top right image shows total excess mortality across all waves.

Figure 14: Weekly mortality rates with (right) and without (left) hospital and medical center deaths

Week 22 1918

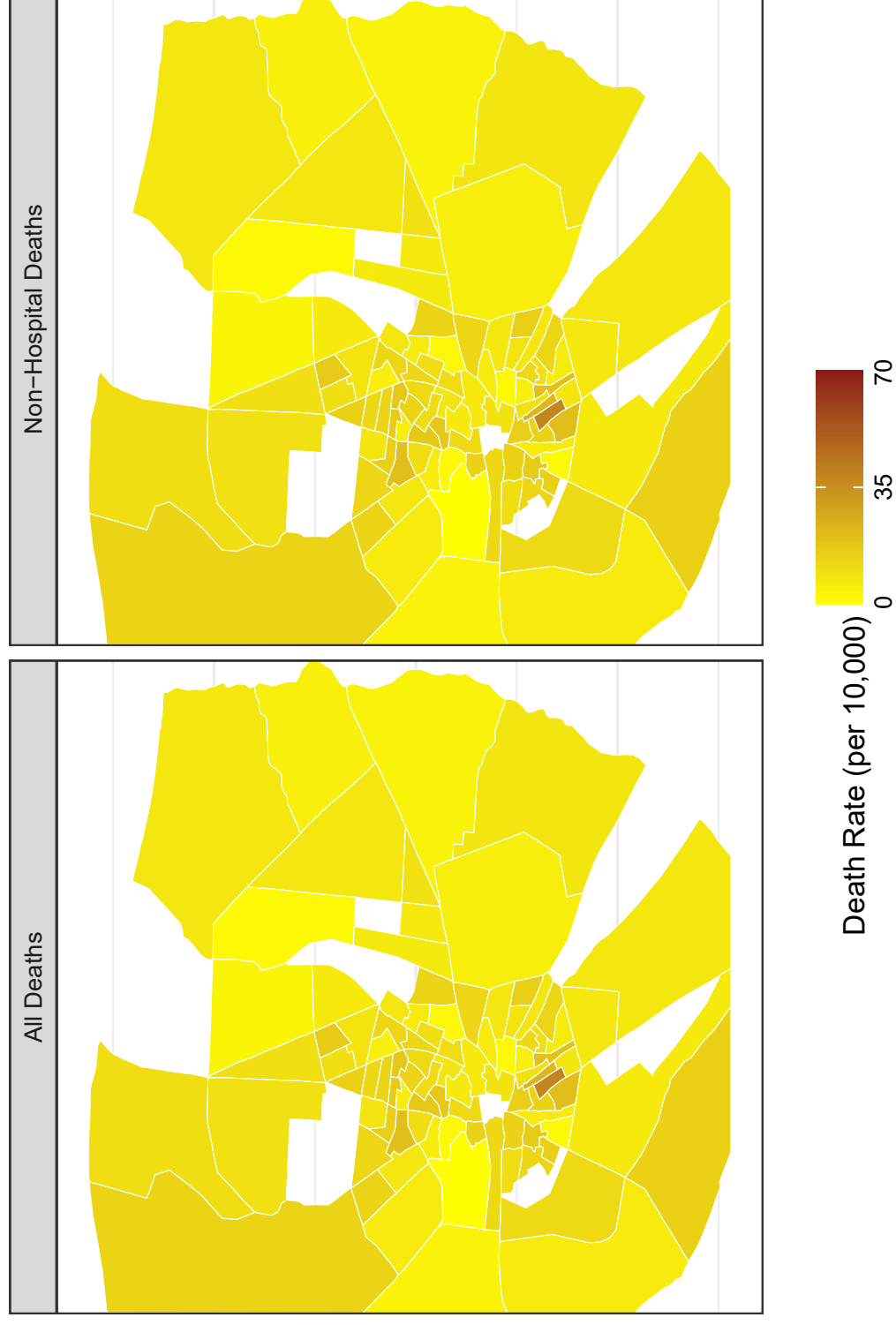
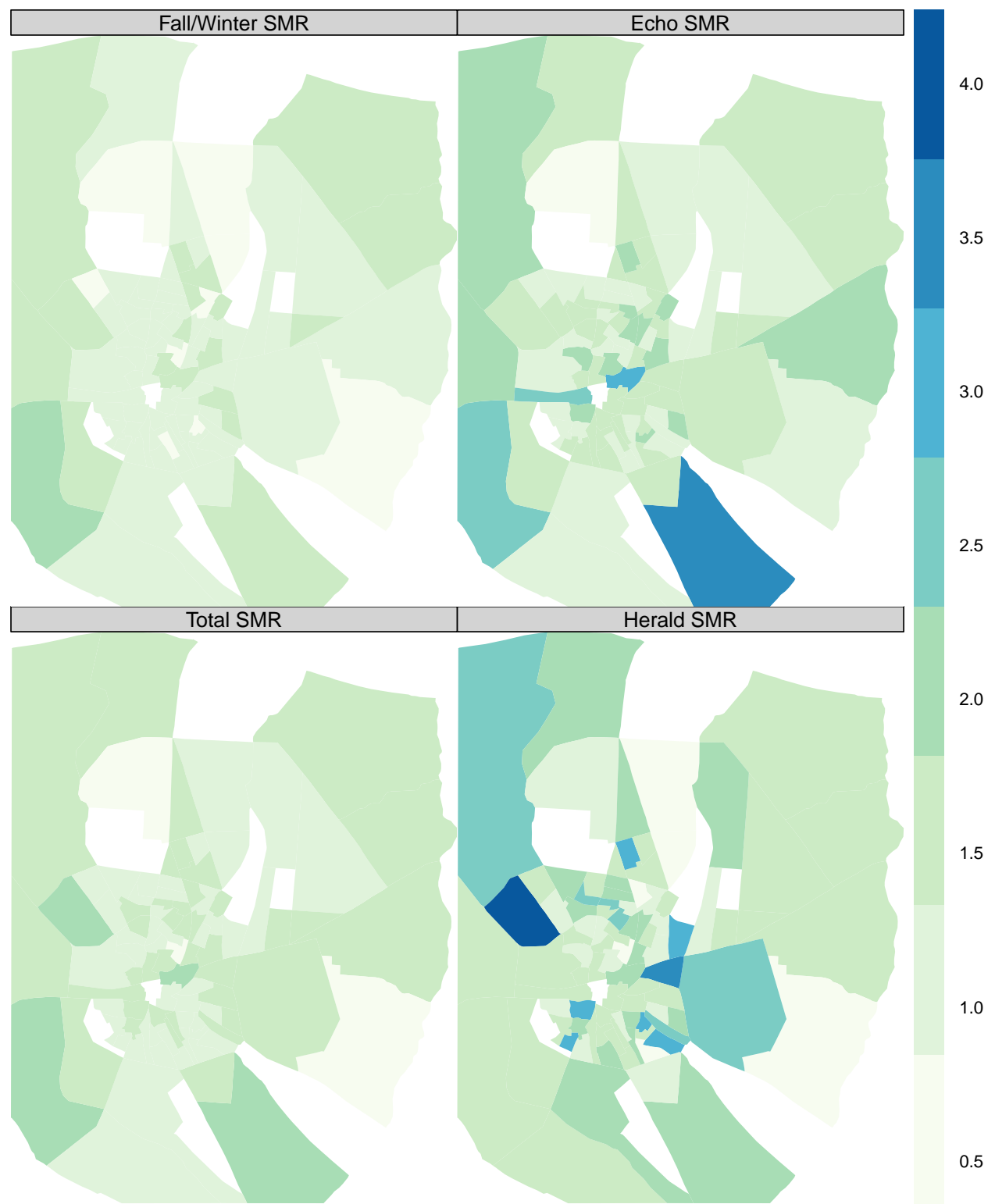


Figure 13: Standardized Mortality Ratio: total pandemic and by wave



4 GIS layers: influenza mortality in Spain

From May through July and September through December 1918, the Spanish Institute of Geography and Statistics collected daily mortality death counts for each provincial capital and entire province in Spain, thus capturing with detail the entirety of the spring wave and most of the immense fall wave throughout the country [16]. Called the “Resumen (Mensual) del Movimiento Natural de la Población de España y de las Capitales de provincia,” or Monthly Spanish Statistics Bulletin, these reports include daily death counts of those who succumbed to influenza. The Spanish National Institute of Statistics (Instituto Nacional de Estadística, or INE) published these reports monthly, after a small delay to collect, aggregate, and prepare the data and report, which covered 100% of the population of Spain. The appearance of the daily death counts is limited to the months noted, as they were part of a special issue of the Boletín, which included this detailed information understanding the outcome of the epidemic waves. Similar issues were published following the especially deadly Cholera outbreaks of 1885.

In the provincial capitals of Spain between September and December 1918, 21,048 deaths were recorded attributed to “Gripe” (influenza). This does not strictly take into account deaths according to respiratory and other diseases normally associated with influenza mortality, such as pneumonia, broncopneumonia, and bronchitis [17, 18]. The figures below display the daily death rates for influenza in the total province, provincial capital, and remaining provincial population of Spain for 49 provinces in May – July and September – December 1918. Two aspects of the graphs are important to highlight. First, it is easy to see that the Herald wave was *not* universal in all areas of Spain. Secondly, and perhaps most notably, the progression of the waves along railways from the northeast of the country to the southeast and west are clearly visible.

Figure 15: Daily mortality rates for the total province population, Spain; May – July 1918

Date: 1918-06-08

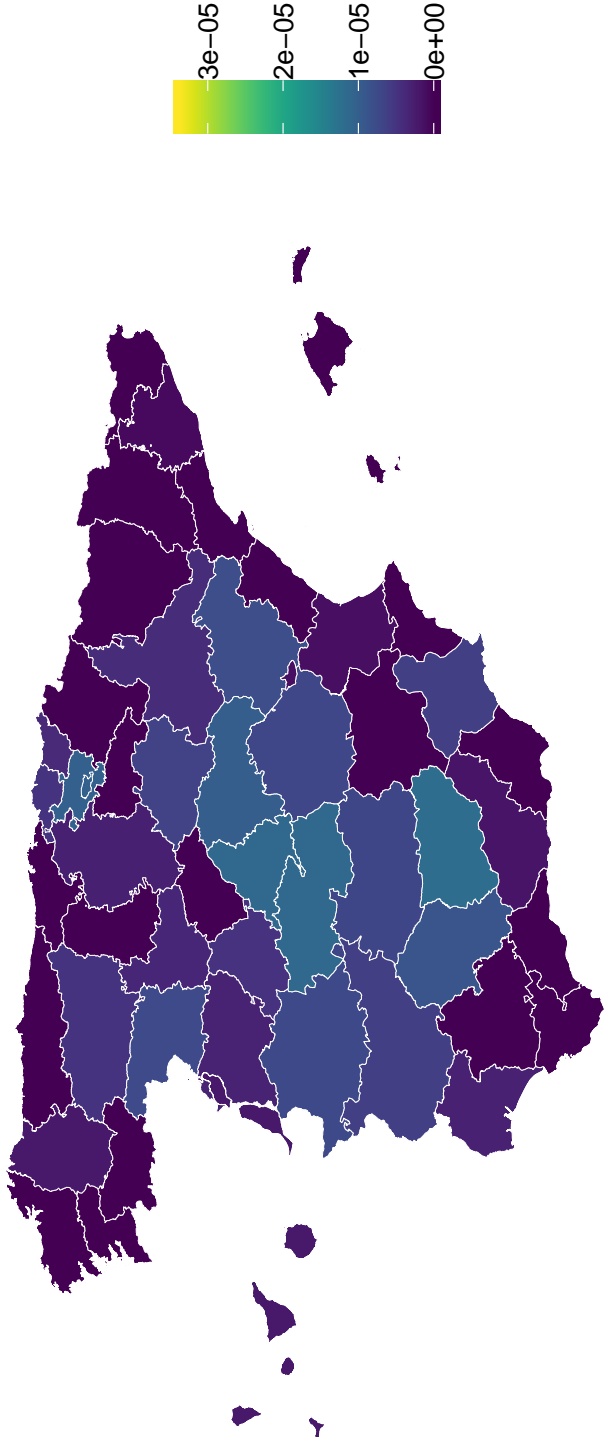


Figure 16: Daily mortality provincial capitals; May – July 1918, Spain. The maps represent the number of deaths in the provincial capitals each day as a proportion of the total number of deaths recorded in that provincial capital during the total period.

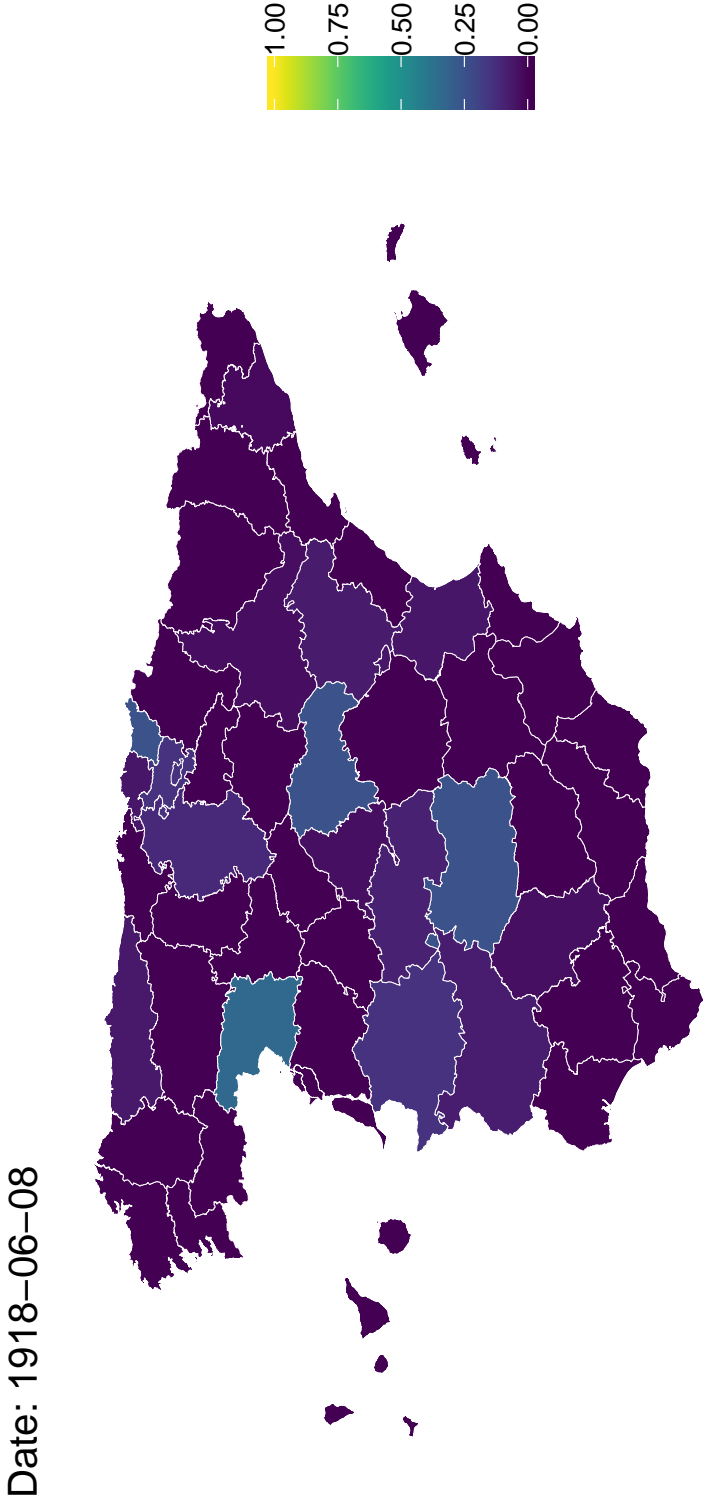


Figure 17: Daily mortality outside of provincial capitals; May – July 1918, Spain. The maps represent the number of deaths outside of the provincial capitals each day as a proportion of the total number of deaths recorded in the same area outside the provincial capital during the total period.

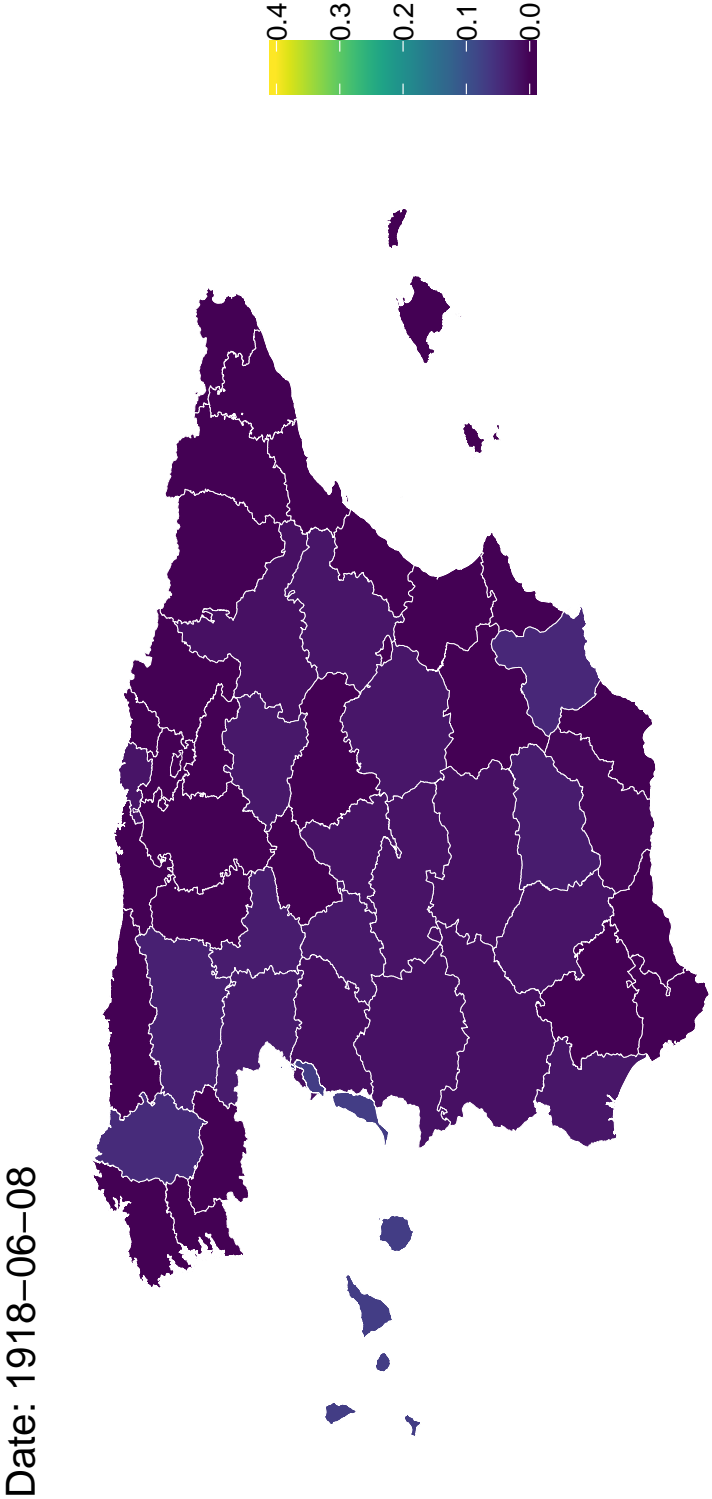


Figure 18: Daily mortality rates for the total province population, Spain

Date: 1918-10-17

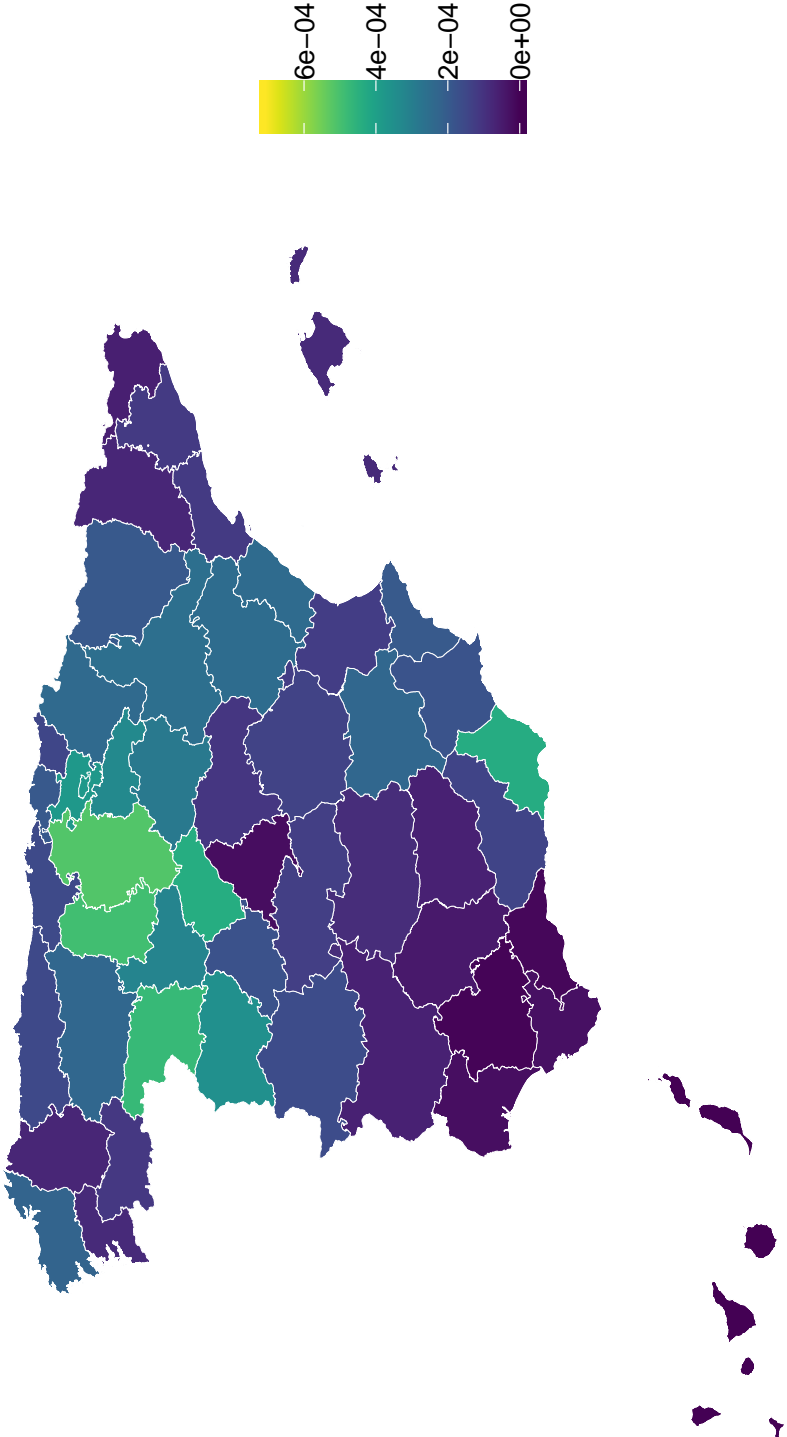


Figure 19: Daily mortality provincial capitals; September – November 1918, Spain. The maps represent the number of deaths in the provincial capitals each day as a proportion of the total number of deaths recorded in that provincial capital during the total period.

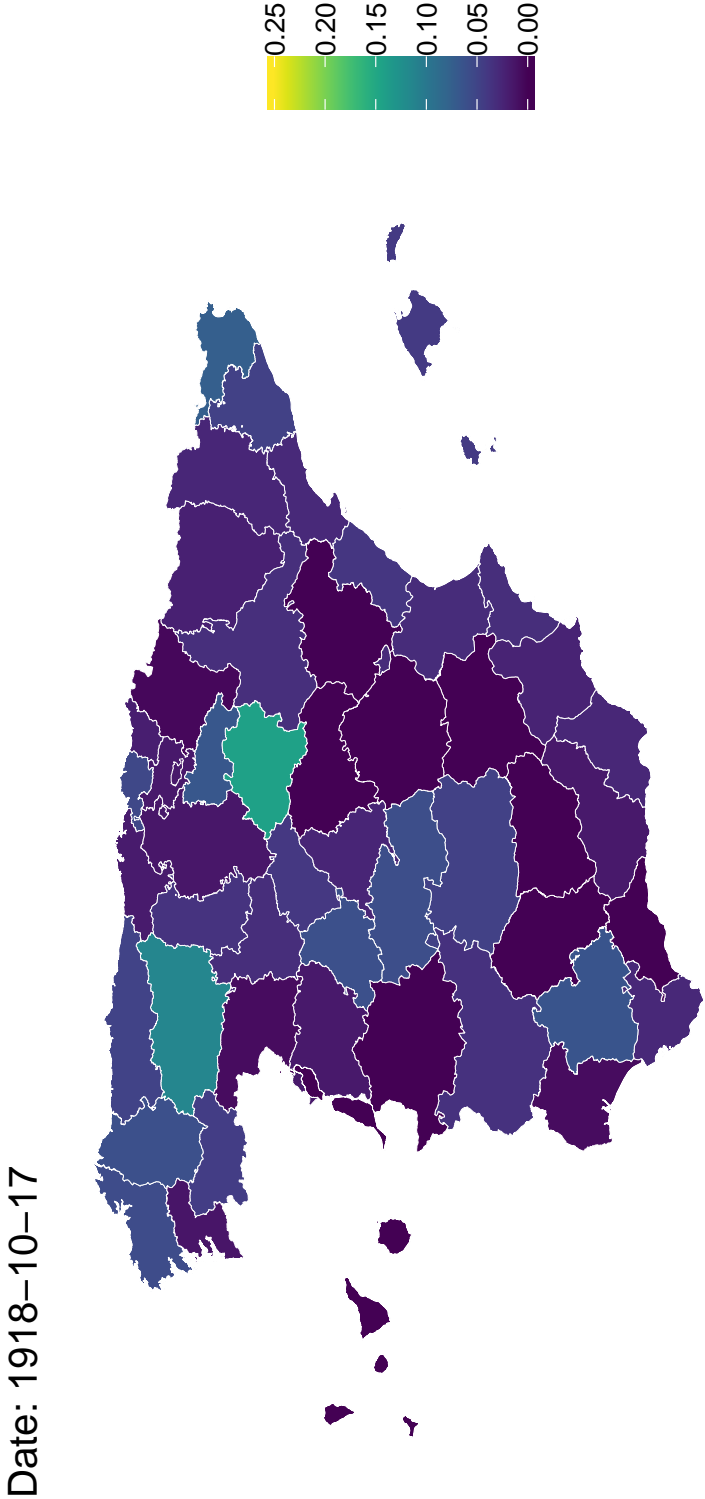
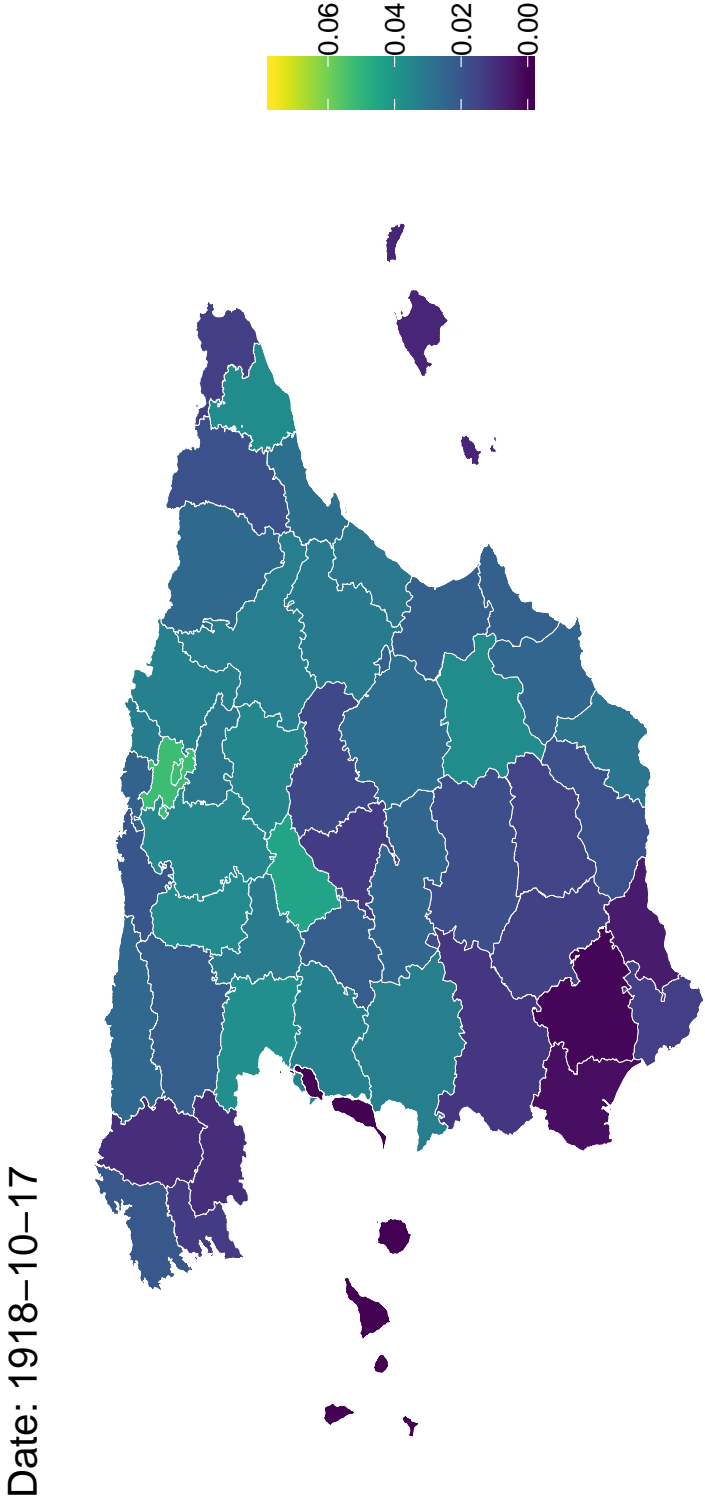


Figure 20: Daily mortality outside of provincial capitals; September – November 1918, Spain. The maps represent the number of deaths outside of the provincial capitals each day as a proportion of the total number of deaths recorded in the same area outside the provincial capital during the total period.



References

- [1] J. A. M. Martín, “Madrid, de villa a metrópoli. las transformaciones del siglo xx,” *Cuadernos de Historia Contemporánea*, no. 22, pp. 225–249, 2000.
- [2] G. G. Bravo, “El delito y la enfermedad en la España del cambio de siglo,” *Cuadernos de Historia Contemporánea*, pp. 113–119, 2007.
- [3] F. V. Albarrán, “Los motores del crecimiento demográfico de Madrid (1860-1930). flujos migratorios y procesos de segregación en los nuevos espacios urbanos,” pp. 259–282, 2011.
- [4] F. Vicente and B. Carballo, “Ser inmigrante en Madrid (1860-1930),” *Ciudad y Modernización en España y México. Granada: Ediciones de la Universidad de Granada*, pp. 441–464, 2013.
- [5] F. Villacorta Baños, “Burguesía y cultura. los intelectuales españoles en la sociedad liberal, 1808-1931,” *Madrid: Siglo XXI*, vol. 105, pp. 1870–1900, 1980.
- [6] D.-S. Reher and F. Dopico, “El declive de la mortalidad en España 1860-1930,” *Asociación de Demografía Histórica (ADEH), Barcelona*. [<http://www.ucm.es/info/geps/424.htm>], 1999.
- [7] A. Sanz Gimeno and D. Ramiro Fariñas, “La caída de la mortalidad en la infancia en la España interior, 1860-1960: un análisis de las causas de muerte,” 2002.
- [8] F. C. Comín, *150 años de historia de los ferrocarriles españoles: La era de las concesiones a las compañías privadas*, vol. 1. Anaya, 1998.
- [9] J. L. Betrán, *Historia de las epidemias en España y sus colonias: 1348-1919*. La esfera de los libros, 2006.
- [10] B. Echeverri Dávila, “Spanish influenza seen from Spain,” *The Spanish Influenza Pandemic of 1918-1919: New Perspectives*, p. 173, 2003.
- [11] A. Trilla, G. Trilla, and C. Daer, “The 1918 Spanish flu in Spain,” *Clinical infectious diseases*, vol. 47, no. 5, pp. 668–673, 2008.
- [12] R. Gutiérrez González, L. Martín-Forero Morente, and I. del Bosque González, “Visualizando el pasado a través de ide históricas. Madrid a principios del s. xx.,” 2012.
- [13] I. N. d. E. Ayuntamiento de Madrid, “Censos de población y renovaciones del padrón municipal de habitantes, 1897–2011,” 2019. data retrieved from Ayuntamiento Estadística.
- [14] F. Caldeiro *Historia - Colegio Fundación Caldeiro*.
- [15] M. C. Registry, *Documentos del Archivo general de la villa de Madrid*, vol. 1. Imp. y lit. municipal, 1917-1922.
- [16] I. G. y Estadístico, “Resumen (mensual) del movimiento natural de la población de España y de las capitales de provincia,” 1918.
- [17] D. M. Morens and A. S. Fauci, “The 1918 influenza pandemic: insights for the 21st century,” *Journal of Infectious Diseases*, vol. 195, no. 7, pp. 1018–1028, 2007.
- [18] D. M. Morens, J. K. Taubenberger, and A. S. Fauci, “Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness,” *Journal of Infectious Diseases*, vol. 198, no. 7, pp. 962–970, 2008.