Master Thesis in Geographical Information Science nr 148

The Influence of Natural Radon Occurrence on the Severity of the COVID-19 Pandemic in Germany: a Spatial Analysis

Ignatius Gabriel Aloysius Maria Perera

2022
Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Ignatius Gabriel Aloysius Maria Perera (2022). The Influence of Natural Radon Occurrence on the Severity of the COVID-19 Pandemic in Germany: a Spatial Analysis
Master degree thesis, 30/ credits in Master in Geographical Information Science
Department of Physical Geography and Ecosystem Science, Lund University

The Influence of Natural Radon Occurrence on the Severity of the COVID-19 Pandemic in Germany: a Spatial Analysis

Ignatius Gabriel Aloysius Maria Perera
Master thesis, 30 credits, in Geographical Information Sciences

Supervision:

Dr. Lina Eklund
Department of Physical Geography and Ecosystem Science, Lund University

Prof. Dr. Frank Sirocko
Institute of Geosciences,
Johannes Gutenberg University Mainz

Abstract

The naturally occurring, radioactive noble gas radon is known to cause permanent damage in lung tissue. As the COVID-19 pandemic showed largely unexplained high spatial variation of mortality and infection rates, it was suggested that the increased exposure to radon gas damages the lung to a degree, that amplifies an infection with SARS-CoV-2 and causes a severe progression of COVID-19, which could show as an increase of infections and severe COVID-19 progressions in regions of high radon occurrence. Three different time periods were analysed at a district level, to test for spatial correlations of case fatality ratio and infection rate to radon soil gas concentration. In each period case fatality ratios and infection ratios of the total population, population equal to or greater than 60 years and population below 60 years were tested for correlation to mean age, mean income, mean living space, and mean radon soil gas concentration. An additional spatial regression analysis supported the suspected associations, and a possible indication of a weak positive spatial correlation between case fatality ratio and radon occurrence was demonstrated. Additionally, a weak but significant spatial correlation between COVID-19 infection rates and radon occurrence was revealed.
More sophisticated methodology was suggested, due to several limitations of the study project. The need for further research to support the findings and accept the hypothesis of spatial correlation between COVID-19 and radon exposure was pointed out.

Keywords

GIS, Epidemiology, Germany, SARS-CoV-2, COVID-19, Radon, Spatial regression

Abstract

Das natürlich vorkommende, radioaktive Edelgas Radon ist bekannt dafür, langfristige Schäden des Lungengewebes zu verursachen. Da Fallsterblichkeit und Infektionsraten der COVID-19 Pandemie große, unerklärte räumliche Unterschiede aufzeigten, wurde als mögliche Erklärung vermutet, dass Radon-Gas die Lunge in einem Ausmaß belastet, welches eine Infektion mit SARS-CoV-2 begünstigt und einen schweren COVID-19 Verlauf verursacht, was sich an höheren Infektionszahlen und schwereren COVID-19 Verläufen in Regionen erhöhten Radonvorkommens zeigen würde. Um Korrelationen von Fallsterblichkeit und Infektionsraten zur Radon-Konzentration in der Bodenluft zu analysieren, wurden drei verschiedene Zeiträume auf Landkreisebene untersucht. In jedem Zeitraum wurden Fallsterblichkeit und Infektionsraten der gesamten Bevölkerung, der Bevölkerung älter oder gleich 60 Jahren und der Bevölkerung unter 60 Jahren auf Korrelation mit dem Altersdurchschnitt, dem Durchschnittseinkommen, der durchschnittlichen Wohnraumfläche und der durchschnittlichen Bodenluft Radon-Konzentration geprüft. Die vermuteten Zusammenhänge wurden durch eine zusätzliche räumliche Regressionsanalyse unterstützt und es konnten Hinweise auf einen möglichen Zusammenhang von Fallsterblichkeit sowie ein schwacher, aber signifikanter Zusammenhang von Infektionsraten zu Radonvorkommen aufgedeckt werden.
Wegen verschiedener Einschränkungen des Forschungsprojekts wurde ein Erweitern der angewendeten Methodik und weitere Untersuchung vorgeschlagen um die Hypothese eines räumlichen Zusammenhangs von Radonbelastung und COVID-19 zu bestätigen.

Keywords

GIS, Epidemiologie, Deutschland, SARS-CoV-2, COVID-19, Radon, Räumliche Regression

Table of Contents

List of Abbreviations

ARS:	Acute Radiation Syndrome
ARDS:	Acute Respiratory Distress Syndrome
BfS:	Bundesamt für Strahlenschutz (German Federal Office for Radiation Protection)
BRD:	Bundesrepublik Deutschland (Federal Republic of Germany)
CFR:	Case Fatality Ratio
COPD:	Chronic Obstructive Pulmonary Disease
COVID19:	Coronavirus Disease 2019
CRS:	Coordinate Reference System
CSV:	Comma-separated values, file format for text data
Destatis:	Deutsches Statistisches Bundesamt (German Federal Statistical Office)
DIVI:	Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin (German Interdisciplinary Association for Intensive Care and Emergency Medicine)
EPSG:	European Petroleum Survey Group Geodesy
GIS:	Geographical Information System
GeoJSON:	see JSON, for spatial data.
GRP:	Geogenic Radon Potential
ICU:	Intensive Care Unit
JSON:	JavaScript Object Notation, open standard file and data interchange format
MSE:	Mean Square Error
RDBMS:	Relational Database Management System
RKI:	Robert-Koch-Institute
Rn:	Radon
RT:	Radiotherapy
SARS-CoV-2:	Severe Acute Respiratory Syndrome Coronavirus Type 2
SPLOM:	Scatter Plot Matrix
SRID:	Spatial Reference Identifier
SQL:	Structured Query Language
INSPIRE:	INfrastructure for SPatial InfoRmation in Europe

List of Tables

List of Figures

1 Introduction

The COVID-19 pandemic has influenced the world's majority during the past two years. COVID-19 is an infectious disease caused by the novel SARS-CoV-2 virus which has spread within a few weeks to every continent and almost every country (World Health Organization, 2022). While many infected persons experience none or only mild to moderate symptoms of the respiratory system, like a dry cough or sore throat, other patients experience life threatening symptoms, the most prominent ones being lung complications such as pneumonia or the acute respiratory distress syndrome (ARDS).

Environmental factors such as temperature, relative humidity, wind speed and precipitation (Ganslmeier et al. 2021) which influence the transmission of SARS-CoV-2 virus and the behaviour of the population have been studied by various scientists. Radon, which is a radioactive and naturally occurring gaseous element, can be found in varying concentrations everywhere in nature and is known to be the primary natural cause of lung cancer and thus responsible for most cases of lung cancer after smoking (World Health Organization 2021). The exposure to radon does not necessarily result in lung cancer, but can cause a variety of respiratory diseases, like pneumonitis, lung fibrosis or chronic obstructive pulmonary disease (COPD) and a respiratory system prone to infections (Turner et al. 2012). Patients with various pre-existing lung diseases have been early identified to be at high risk for severe COVID-19 disease progression and an increased mortality (Maragakis 2021). The impact of meteorological driven exhalation of radon gas into the atmosphere (Mudelsee et al. 2020; Kulali et al. 2017; Tchorz-Trzeciakiewicz and Kłos 2017; Rey et al. 2022) on the severity of a COVID-19 progression or on the transmission of the SARS-CoV-2 virus is unknown.

Although infections appeared in all regions, mortality rates show large spatial variations (Figure 7, Section 3.2.1; Figures 13-20, Appendix), which could be correlating with radon occurrence (Figure 1, Section 1.1.3). As the COVID-19 disease can have a heavy impact on a weakened lung and on a damaged respiratory system, it is of interest to analyse if areas of geologically high radon potential or times of meteorologically caused high radon exhalation can be linked to higher mortality, hospitalization rates or infection rates.

1.1 Background

1.1.1 COVID-19 and SARS-CoV-2

COVID-19 is the infectious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), first identified in December 2019 in Wuhan, China (Huang et al. 2020). The diseases global and fast spread motivated the World Health Organization to declare COVID-19 a pandemic disease on March

11, 2020 (World Health Organization 2020). At that point of time, more than 118 000 cases in 114 countries had been registered.
SARS-CoV-2 is primarily transmitted via respiratory droplets and aerosols, when contact to oral, nasal or conjunctival mucosa occurs (Cevik et al. 2020). The virus binds to the human receptor angiotensin-converting enzyme 2 (ACE2), which is found in lungs, kidneys, heart and intestine (Yan et al. 2020). According to a meta-analysis by Quesada et al. (2021), the mean incubation period ranges from 5.6 to 6.7 days, with increasing incubation time at increasing age. While mild symptoms like fever, cough, tiredness or the loss of smell or taste are common, COVID-19 can also manifest as headache, gastrointestinal disorder, sore throat, pain or in severe cases as shortness of breath, confusion, chest pain, respiratory failure, or multiple organ dysfunction (National Institutes of Health 2022; World Health Organization 2022). The case fatality rate (CFR), which is defined as the ratio between total confirmed deaths and total confirmed cases ranges between 0.4 and 15 %, is heavily influenced by multiple factors, such as pre-existing diseases, available hospital beds, health care expenditure or air pollution (Rajgor et al. 2020), and thus is spatially varying between and within countries. Spatial variation could also be observed in Germany (Figure 7, Section 3.2.1; Figures 13-20, Appendix). Long-term effects after an infection, also known as long COVID or post-COVID-19 syndrome, were at first not noticed, but have shown to range from tissue damage and unresolved inflammations to autoimmune disorders (Yong 2021).
A major factor in containing the pandemic in terms of reducing transmission rates and reducint mortality was the successfull development of multiple vaccinations. Approximately one year after the beginning of the pandemic, the first vaccination against the SARS-CoV-2 virus was conditionally approved by the European Commission, soon to be followed by other vaccinations. The effectiveness of the vaccines have been assessed to be varying, due to virus mutations, age of patients or vaccine manufacturer and vaccination type, but all revealed a partial protection against death, effectively reducing mortality (Jabłońska et al. 2021).

1.1.2 Meteorology, Demography, Socioeconomics

Various studies and research projects have studied the links between meteorological parameters and the pandemic. Mean air humidity and COVID-19 mortality have shown to be correlated (Biktasheva 2020), but also the influence of temperature and wind speed on the spread of the virus have been revealed (Ganslmeier et al. 2021). For the first quarter of 2020, COVID-19 cases showed positive associations with temperature and negative associations with wind speed, when a 14-day lag has been applied (Islam et al. 2021). An absolute humidity of 5-10 g/m³ was suspected to be ideal for viral transmission. However, the strongest driver for incidence rate is previous incidence rate. The coefficient of determination in a regression only using lagged (previous) incidence rate is much higher than in a regression where significant lagged humidity or temperature have been used. (Ganegoda et al. 2021).

As the spread of the SARS-CoV-2 virus as well as the severity of a COVID-19 progression have shown correlations to multiple socioeconomic, behavioural and demographic factors (Mollalo et al. 2020; Drefahl et al. 2020), it is necessary to normalise and control assumed controlling variables before statistical analysis.

For example increasing age, lower income and education levels have shown strong correlations to death from COVID-19 in Swedish data (Drefahl et al. 2020).The clearly dominating effect of crowding on the infection rates has been identified by evaluating New York City's mean household sizes for each zip code in comparison to infection rates (Federgruen and Naha 2021). Another research identified a living space of less than 9 m² per individual as the relevant threshold (Universitätsmedizin and der Johannes Gutenberg-Universität Mainz 2021).

1.1.3 Radon

The recently (January 2022) updated soil radon concentration map (Figure 1) of Germany is based on over 4000 soil measurements performed over the last 25 years. These field measurements were taken all over Germany with varying density, in dependence of geology and lithostratigraphic aspects, to map the spatial varying concentrations of naturally occurring radon (Kemski et al. 2001). Radon originates as intermediate product of the radioactive decay chains from uranium-235, uranium-238 and thorium-232, which's concentrations in rock and soil depends on geology. Radon can be found everywhere in Germany's soils and is known for its radioactivity. Radioactivity in general has multiple effects on cells, DNA and organs (Nanduri et al. 2021). Lung cancer is the most prominent symptom of long-term radon exposure (Al-Zoughool and Krewski 2009), but also an increased mortality from chronic obstructive pulmonary disease (COPD) are known to result from radon based exposure (Kang et al. 2019). As the progenies from radon's radioactive decay are not gaseous but solid, these particles can adhere to the sensitive lung tissue when inhaled, while radon gas itself is exhaled again (Al-Zoughool and Krewski 2009). Thus, radon exposure can result in respiratory tract accumulations of radioactive decay products: extremely toxic (Ansoborlo 2014; Slikkerveer and de Wolff 1989; Wani et al. 2015) and radioactive isotopes of Polonium, Lead and Bismuth. After the decay chain finishes, the stable, non-radioactive product is toxic lead, which again poses a health risk (Wani et al. 2015).

At the time of this research (March 2022), there have not been conducted any studies pointing to the possible influence of spatially varying radon exposure as a factor contributing to the severity of a COVID-19 disease progression or to increased infection rates. The only researched existing link between SARS-CoV-2 or COVID-19 and radon is the increased risk of radon exposure since the beginning of the pandemic, due to counter measurements such as social distancing, quarantine, and lockdowns: more time spent inside housing, especially when lower ventilated in cold months, increases exposure to air of bad quality and potentially higher radon-concentrations (Maya et al. 2020; Verde et al. 2022).

The potential link to COVID-19 has not been researched yet, but patients with various pre-existing lung diseases have been early identified to be at high risk for severe COVID-19 disease progression and high mortality (Maragakis 2021).

COVID-19 patients with a history of radiation therapy (RT) have been found to have a poor prognosis, as mortality appeared correlated to the extent of prior lung irradiation (Kabarriti et al. 2020) and a link between radiotherapy and a risk of hospitalization due to COVID-19 has been revealed (LaPlant et al. 2021). Immunological commonalities between a COVID-19 disease and a radiation injury have been pointed out, as symptoms and the underlying pathogenesis of the multi-organ injury caused by SARS-CoV-2 in a severe progression resemble the one's of an acute ionizing radiation exposure (Rios et al. 2020).

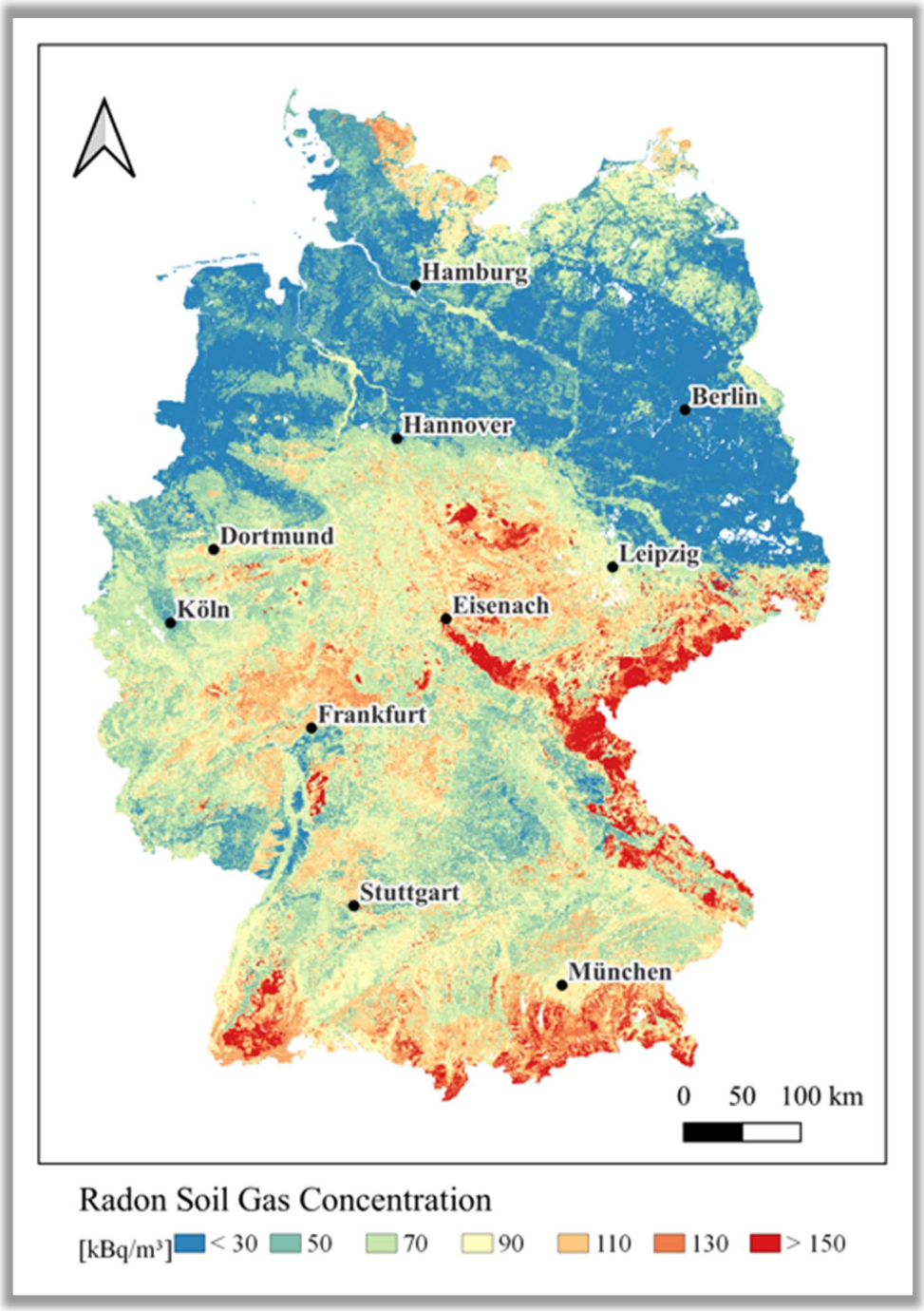


Figure 1 - Radon Soil Gas Concentration, Data: Bundesamt für Strahlenschutz (2022a)

1.1.4 Geographical Information Science in Public Health

Already in the year 1854, geographic information was used in a public health context, when John Snow mapped a cholera outbreak in London, associating disease with spatial data (Snow 1855). Snow showed on a map, that cholera mortality peaked in the supply area of a certain drinking water pump, which's decommission by the authorities swiftly ended the cholera outbreak.

Geographical information science (GIS) acts as a tool to store, analyse, archive, modify and visualise spatial and temporal data of various fields such as demography, economy, meteorology, geology, environment, infrastructure or health and diseases. With today's increasing availability of spatial and temporal data, GIS in public health becomes an important tool to increase the "understanding of the prevalence, aetiology, transmission, and treatment of many diseases" (Richardson et al. 2013). Through sophisticated spatial and spatiotemporal statistical analysis, correlations between different factors and diseases can be identified and thus contribute to better interventions or prevention measurements to enhance public health conditions.

Not only in disease control and prevention, but also in general health management, GIS has contributed to public health, for example by calculating adequate placement of medical institutions or by introducing public information systems like pollen forecasts.

While spatial analysis has been applied for a few decades in the field of health, the addition of a time dimension to geographic data has become increasingly important, especially in epidemiology (Musa et al. 2013). The proceeding digitalisation in all fields of life revealed many sources of geo-located big data, such as social media meta data or cellular mobility data, making spatial data analysis a field with large potential for future health research.

Regarding the COVID-19 pandemic, from early on the beginning of the global spread, there have been conducted many analysis in GIS with varying focuses such as mobility, politics, environmental variables, social geography, or epidemiological prediction (Franch-Pardo et al. 2020).

1.2 Scope of Study

This research project aims on assessing the correlation between radon occurence and various measures of COVID-19 severity like number of COVID-19 patient occupied intense care units (ICU) or the number of death cases per infected, also referred to as the case-fatality-ratio (CFR). As the potential drivers and influences for increased COVID-19 transmission rates and mortality are numerous and could not all be researched, this research project's scope is spatially and temporally limited and focuses only on radon soil gas concentration. Other potentially influencing variables like income as a health care measure and population density or mean living space as a measure for increased exposure to infected and consequently higher transmission risk are additionally evaluated.

1.2.1 Study Area

The area of this research project is limited to the inhabited area of the Federal Republic of Germany, which is located centrally on the European continent. The humid westerlies dominated climate of the country is partly temperate (Köppen-Geiger class: Cfb) and partly cold (Köppen-Geiger class: Dfb) with warm summers and without dry seasons (Beck et al. 2018). While temperatures in summer reach on average around 17°C, winter temperatures average around 0°C.

Germany has sea connection to the Atlantic Ocean via the North Sea and the Baltic Sea in the north of the country, which consists of some wetlands and flat terrain, formed by the last glacial period. Central Germany's rougher uplands were partly formed by ancient activity of volcanism in the west and merge in the east with the Ore Mountains. Topography towards the south shows gradually rising altitudes until the high mountains of the Alps are reached at the south borders of Germany.
While surface geology of the flat north consists primarily of Quaternary sedimentary rock deposits, the central and southern regions are rich in granites, Palaeozoic black shales and silica-rich volcanic rocks, as well as Triassic sandstones and marls (Kemski et al. 2001). The contents of radionuclide in these individual geological formations cause spatial variations of radon occurrence, which's soil concentrations range from below 20 kilobecquerel/m³ (kBq/m³) up to above 1 megabecquerel/m³ (mBq/m³) (Bundesamt für Strahlenschutz 2022a).

Germany's central position on the European continent, Frankfurt Airport being one of the busiest airports in Europe and being a nation, which has strong international economic and touristic exchange introduced SARS-CoV-2 at an early stage of the pandemic into the country. Within less than eight weeks of the first confirmed SARS-CoV-2 infection, the virus spread in over 95% of all German districts (Robert Koch-Institut 2020).
Every identified infection with SARS-CoV-2 must be reported to the public health department of the district and all of the over 400 districts report daily to the Robert Koch-Institute (RKI), the agency and research institute of the German federal government responsible for disease control and prevention. The resulting reports of all anonymised infection and death cases are available to the public. A high-resolution radon soil gas concentration map (Bundesamt für Strahlenschutz 2022a) and the source data is also publicly accessible, making it possible to perform this research.
The SARS-CoV-2 records of Germany and other countries revealed a high positive correlation between age and mortality (O'Driscoll et al. 2021). Reported COVID-19 related death cases were extraordinarily high in the southern districts of the states Saxony and Thuringia, as well as in the north-eastern districts of the state Bavaria, and show a similar pattern (Figures 18-20, Appendix) to the distribution of higher radon concentrations (Figure 1), which makes it a very interesting study area.

1.2.2 Study Period

To research different conditions the analysis will be performed on three time periods:

I. a subset of the reports from 01.01.2020 until 31.12.2020, limited to the period before the first vaccinations in Germany, which became effective within the first week of 2021. As the vaccinations were not evenly distributed within all districts and age groups and there are great spatially varying vaccination rates, the time before applied vaccinations is the most relevant to perform an analysis.

II. a second subset limited to the 6 weeks of and after a nationwide low-pressure event ("Sturmtief Ignatz") from 19.10.2021 until 30.11.2021. Low atmospheric pressure increases exhalation of radon from soil into the environment (Mudelsee et al. 2020; Nazaroff 1992) and consequently causes a higher radon exposure, which could be shown in increased COVID-19 indicators in the weeks after a low-pressure event. Considering the average incubation time of 6 days and an average serial interval of over 4 days, as well as an average time of 11 days after first symptoms until the time of death in severe cases (Robert Koch Institut 2021; Quesada et al. 2021), a relevant study period of 6 weeks after the low-pressure event is assumed.

III. the full dataset containing all reported COVID-19 associated ICU cases, infections, and deaths from 01.01.2020 until 01.03.2022. The two subsets of period I and period II only cover times of comparatively low nationwide infections and subsequently pose a much smaller sample size than the full dataset, which includes the very high nationwide infection-rates of the fifth wave in Germany.

1.3 Specific Aim and Added Value

Of uranium's and thorium's decay chains, radon is the only element in gaseous form, making it possible to exhale from soil and accumulate in the air, where radon poses a danger to the lungs.
The COVID-19 pandemic has shown unexplained spatially and temporally varying case-fatality-ratios (Sorci et al. 2020) and hospitalization rates, motivating scientists to investigate various potential risk factors. As radon causes diseases of the respiratory system and considering the research results mentioned in 1.1.3, which are pointing to a link between radon exposure and severity of COVID-19, I wondered, if the damage caused by radon exposure poses a risk to be higher susceptible to an infection with SARS-CoV-2 or to the development of a severe progression of COVID-19 and possibly even cause a high mortality risk.

If the possibility exists, that radon exposure caused lung tissue and health damage has an impact on the spread of SARS-CoV-2 or on the severity of a COVID-19 progression, can these impacts be detected on spatial and temporal distribution of death cases and infection rates?
To produce reliable results, I will test, whether possible associations can be isolated from other potentially influencing factors such as age distribution, population density or low health care indicated by low income.
I will test the following hypothesis:

1: High COVID-19-associated case-fatality-rate is spatially linked to high radon concentration in soil air.
2: High COVID-19-associated ICU admission rate is spatially linked to high radon concentration in soil air.
3: High COVID-19 infection rate is spatially linked to high radon concentration in soil air.

The results of this project can be used for further research and could consequently help individuals to reduce health risks, by focusing on the use of radon exposure mitigation strategies, like periodic ventilation, radon insulation and sealing or installing radon pumps, which suck radon out of the soil, before it can enter and accumulate in a building,
Additionally, this project could be used by public health authorities to improve strategies for reducing health impacts or assessing other environmental health risks.

2 Data

In this section I describe the structure, content, and source of the data used to perform the analysis. Germany's governmental agencies provide most non-personal data to the public according to an 'Open Data Act', the E-Government-Gesetz EGovG (Bundesrepublik Deutschland 2013). Especially spatial data is largely available due to the INSPIRE Directive (INfrastructure for SPatial InfoRmation in Europe), which acts as a framework for boundary-crossing utilization of spatial information in Europe (Bundesamt für Kartographie und Geodäsie 2022).

2.1 Administrative boundaries

Administrative boundary data (Bundesamt für Kartographie und Geodäsie 2020) were acquired from the German Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie, short BKG). The administrative level used in this research is the "KREISLEVEL" (Figure 3). A "Kreis" or "Landkreis" corresponds to the EU's Nomenclature of Territorial Units for Statistics (NUTS) Level 3 (Districts).

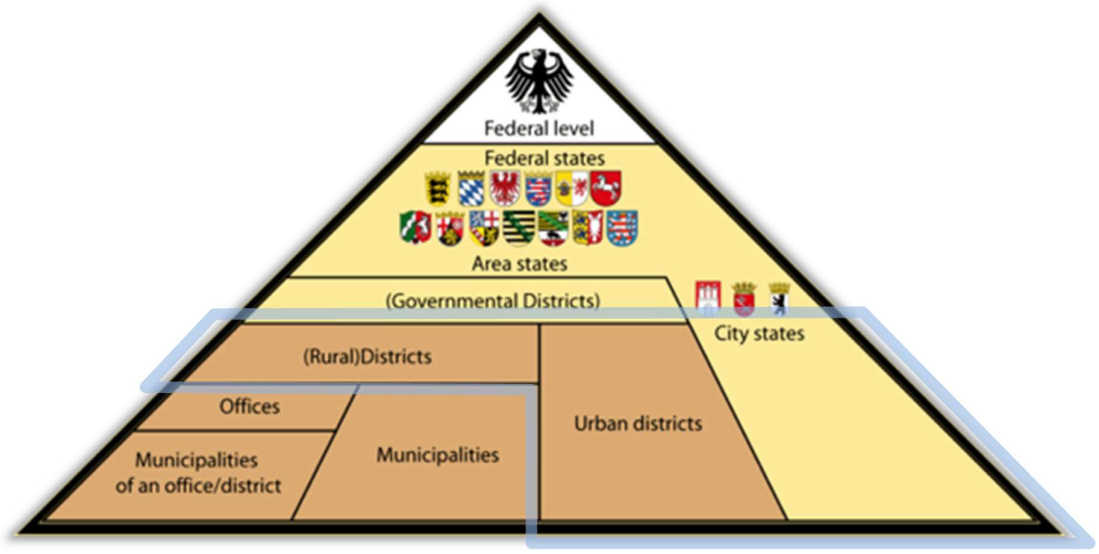


Figure 2 - Administrative Divisions of Germany (David Liuzzo, CC BY-SA 2.0 DE), relevant Level highlighted.

The European Union uses the NUTS classification system to hierarchically divide its territory for the application of uniform statistics and comparable regions. NUTS 1 population sizes are usually between 3 and 7 million, NUTS 2 regions have between 800 000 and 3 million inhabitants and NUTS 3 regions between 150 000 and 800 000 inhabitants. In Germany NUTS Level 1 corresponds to the federal States ("Bundesländer"), NUTS Level 2 to governmental districts ("Regierungsbezirke") and NUTS Level 3 to rural and urban districts, as well as the

three special cases of city states (Berlin, Bremen, Hamburg), which are NUTS 1, 2 and 3 at the same time (Figure 2).

Figure 3 - NUTS3 Regions of Germany (grey) and adjacent countries; Data: European Commission – Eurostat/GISCO (2021).

The used district shapefile contains spatial extend, district name, the NUTS 3 code, cadastral area, and population data, for the reference date 31.12.2020. The data is published by the German Federal Statistical Office, also referred to as Deutsches Statistik-Informationssystem (Destatis).
Additionally, administrative boundaries with population data from 31.12.2019, just before the start of the pandemic, is available in BKG's archive. All data is referenced in the coordinate reference system (CRS) EPSG:4326.

2.2 Demographics and Socioeconomics

High resolution census data, gathered by Destatis for the year 2011 is available for every 1km² as *.csv dataset, showing amongst others, the number of inhabitants. Each km² is uniquely identified and corresponds to a grid cell of a regular EPSG:3035 referenced grid. Due to privacy, Destatis performed some alteration of

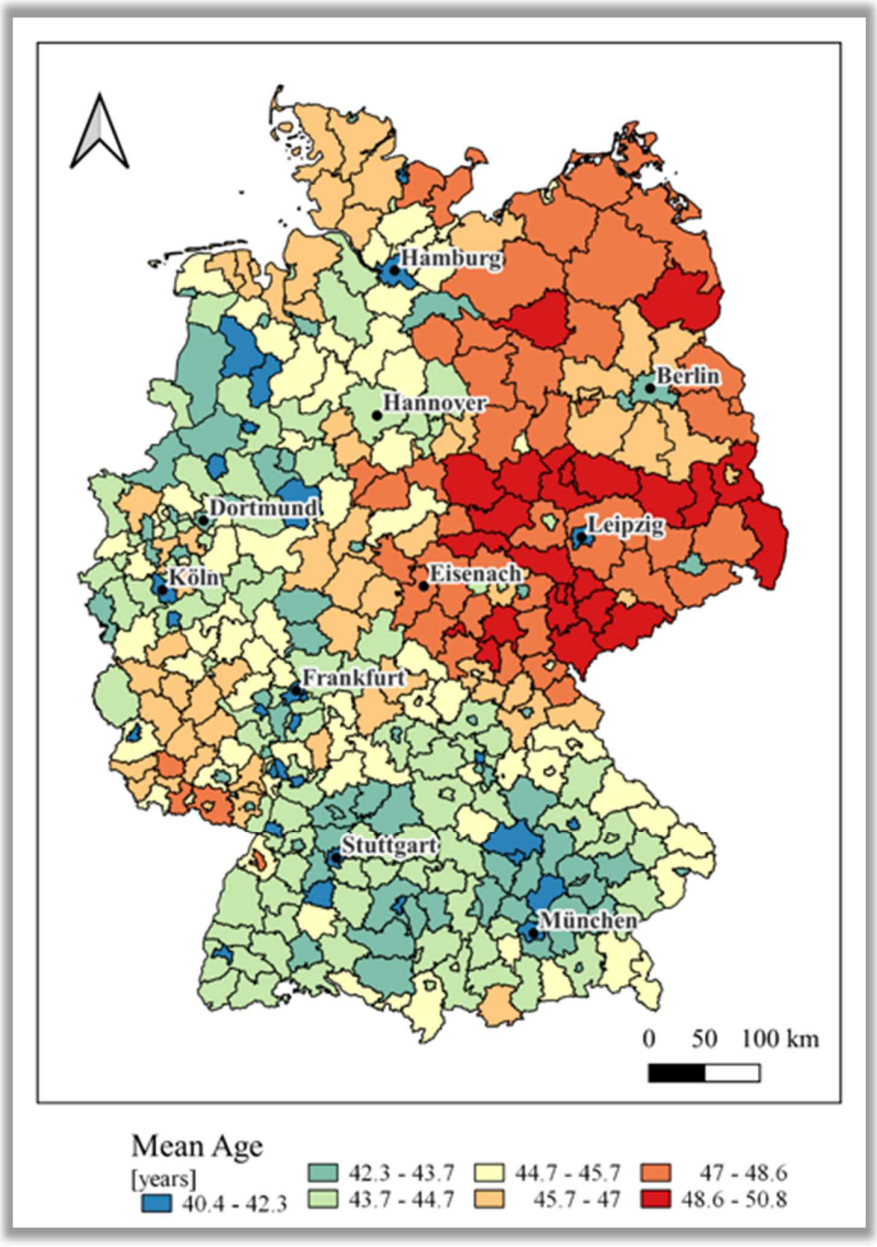


Figure 4 - Mean Age of Population per District.

individual cells, but the overall population is not changed: only cells with one person are shown as cells without inhabitant (value= -1), cells with two inhabitants as cells with 3 inhabitants. This dataset can be used to identify demographic

disparities and create weightings based on population distribution within single districts.

The German Federal Statistical Office calculates annually mean population age (Figure 4) derived from the 2011 census by assessing annual new births, deaths, and migration. The dataset can be acquired from the German Statistical Bureau and contains the mean age of the population for every NUTS3 level (Statistische Ämter des Bundes und der Länder 2019).

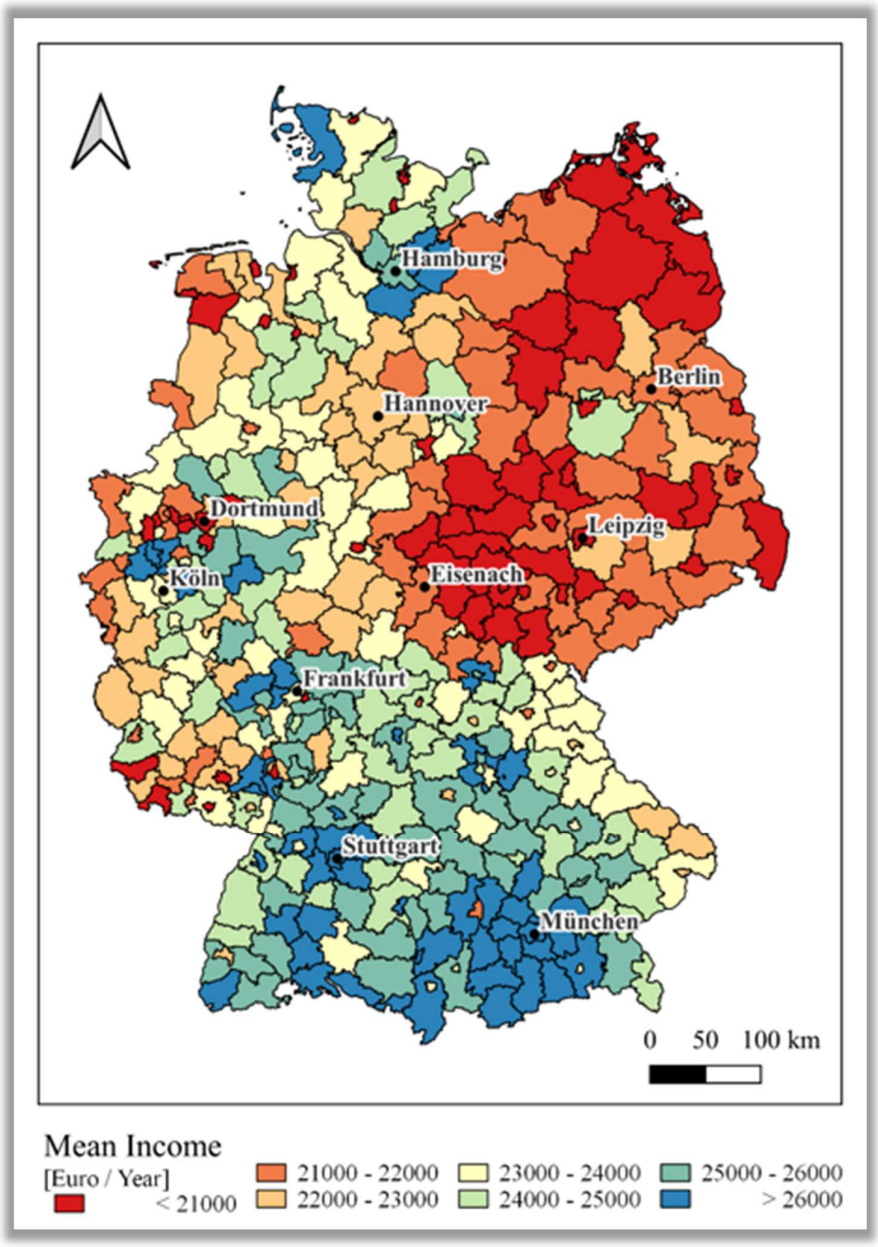


Figure 5 - Mean Income per Inhabitant in 2019; Data: Statistisches Landesamt Baden-Würtenberg (2020).

Additionally more detailed population data for specific age-groups was used (Statistisches Bundesamt 2020) for analysing case fatality ratio of specific age-groups and account for a spatially heterogeneous distributed demography. Mean income per inhabitant (Figure 5) is available from Destatis (Statistisches Landesamt Baden-Würtenberg 2020). Age and income will be used as a potentially influencing factor, as Swedish data have shown a higher risk of death from COVID-19 for high age groups and lower income households (Drefahl et al. 2020) .

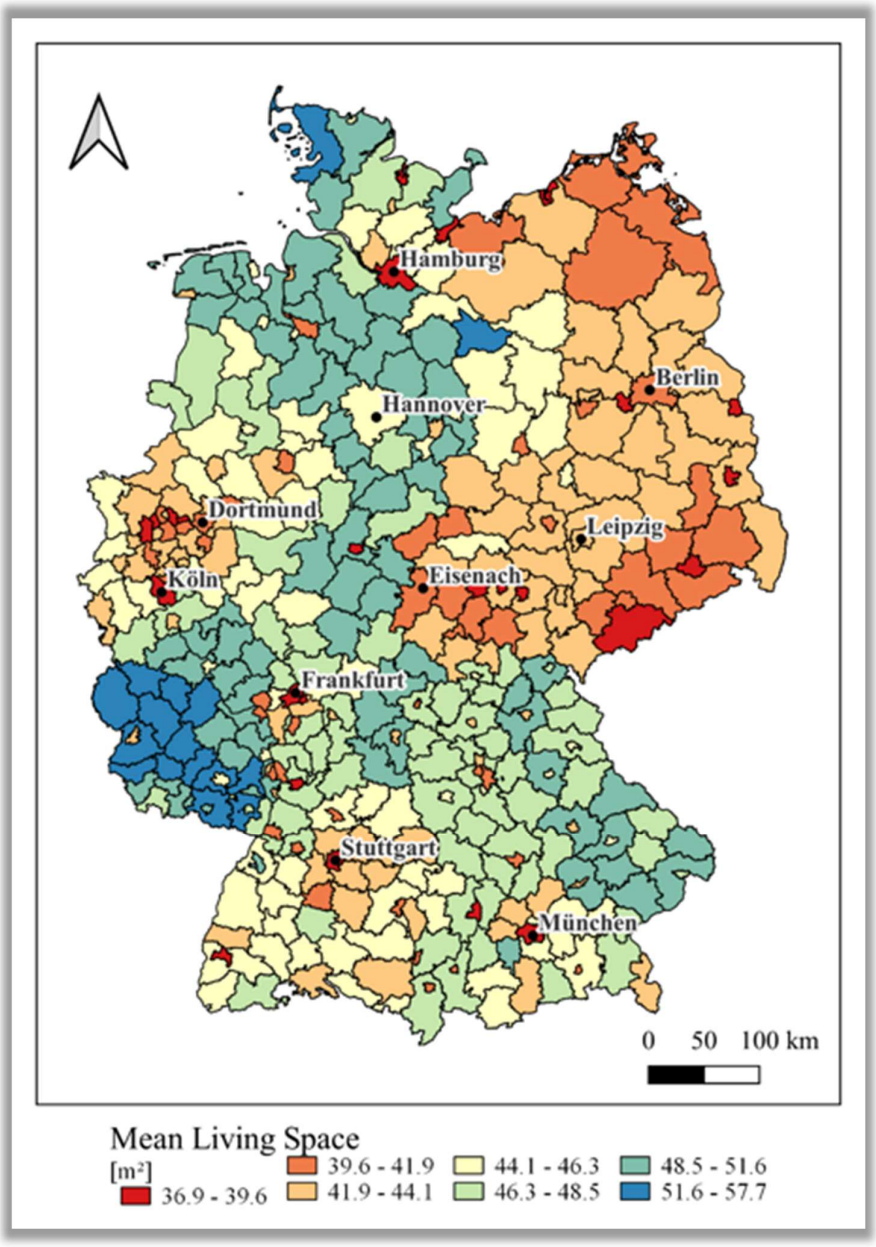


Figure 6 - Mean Living Space per Inhabitant in m²; Data: Statistische Ämter des Bundes und der Länder (2022).

As the available living space has been shown to significantly pose a risk for infection with SARS-CoV-2 (Federgruen and Naha 2021; Universitätsmedizin and der Johannes Gutenberg-Universität Mainz 2021), the district level aggregated mean living space per inhabitant (Figure 6) is incorporated in this research.

2.3 Radon

Geogenic radon potential (GRP) has been defined in several ways (Petermann et al. 2021), but expressing GRP as function or radon concentration in soil gas and soil gas permeability (Neznal et al. 2004) is widely applied and also used in the updated geogenic radon potential map of the German Federal Office for Radiation Protection (BfS).The database of GRP measurements in Germany, performed between 1992 and 2020 (Bundesamt für Strahlenschutz 2022b), counts 6293 locations, with data of a soil gas permeability and radon concentration in soil gas, derived from a standardized in-situ measuring method (Kemski et al. 2001). Combined with geological data, soil characteristics and climate, a machine learning approach (Petermann et al. 2021) was used to estimate radon concentrations in Germany at a 1km² resolution (Figure 1, Section 1), represented in thousands Becquerel/m³ (Bq/m³). The dataset is published by the BfS as GeoJSON dataset (Bundesamt für Strahlenschutz 2022a).

2.4 SARS-CoV2 / COVID-19

For each district, age group and gender, the counts of new confirmed COVID-19 infections, recoveries and fatalities are daily reported (Appendix 8.6) and available as CSV dataset with over 4 million lines (March 2022) at the Robert-Koch-Institute's (RKI) published dashboard on opendata.arcgis.com.
A COVID-19 fatality counts as such, if both, the SARS-CoV-2 has been positively tested and the death could not be excluded to be caused by the infection and its consequences.
From this relatively large dataset, infection rates, mortality rates, transmission rates, for each district, age group or gender for every desired time step can be processed.

Additionally to infection, mortality and recovery numbers, the German Interdisciplinary Association for Intensive Care and Emergency Medicine (DIVI) constantly monitors the occupancy of ICU by COVID-19 patients. The CSV dataset contains for every district daily numbers of COVID-19 patients in ICU and the number of ventilated COVID-19 patients in ICU, i.e. cases of moderate and of severe disease progressions (Robert Koch-Institut & DIVI e.V. 2020). The DIVI dataset does not contain age or gender information of treated patients, thus cannot be used for an age- or gender-specific analysis.

2.5 Vaccinations

Vaccination data, similarly to infection and fatality data published by the RKI (Robert-Koch-Institut 2022) as counts for every district, age-group and vaccination status (Appendix 8.5) is used for the study periods II and III, in which vaccinations have been administered. In study period I, there have not been any vaccinations administered, and vaccination quotas per district and age group can be neglected. This dataset can only give indications about the vaccination quota in a district, as the report contains the district, where a vaccination has been administered not where the patient is living in.

All spatially, temporal and content wise relevant data for this research is available to the public at no cost, but for detailed statistical analysis and testing, the data will need further processing and preparation.

3 Methods

In this section, the data import, preparation, creation, as well as handling of data inconsistency is documented, followed by the description of used exploratory and explanatory statistical methods and the regression model.

3.1 Data pre-processing

3.1.1 Importing & preparing data

As the study uses multiple large tables, which are too large for editing in conventional spreadsheet software like Microsoft Excel, the datasets are imported in a PostgreSQL database management system. PostgreSQL is an Open-Source relational database management system (RDBMS) which is compliant to the quasi-standard Structured Query Language (SQL) for creating, managing, manipulating, or querying complex or large datasets in databases
PostgreSQL uses the PostGIS extension, to handle spatial information of datasets. Shapefiles are imported via a dedicated Shapefile-to-PostGIS function. Tables were created via SQL-query and filled with data from CSV files.
For visualising in QGIS, an Open-Source GIS software, datasets in the PostgreSQL database can be selected as layer-source, making it possible to access and edit large datasets directly within the GIS software.

After import the datasets needed to be homogenized: reported infections, ICU numbers and population datasets all use different nomenclature and codes for the districts. The relevant tables were joined in the SQL database for further processing.

Special care was given to avoid problems caused by common (German) special characters in file or path names or an inconsistent use of decimal separator in datasets or calculations.

Radon soil concentration data, which is available as a regular-grid structured GeoJSON shapefile, was rasterized, to match the 1 km² grid of the census dataset, containing the demographic information. This included a reprojection, as the German national standard projection system used for the radon soil concentration data differs from the EU standard projection system used for census data.

The census 1x1 km regular-grid shapefile was joined with the data from tables containing population counts and age distribution data for each km². . After the preparation of the grid-shapefile with census information, the data was imported into the SQL database via PostGIS Shapefile Import Manager.

The shapefile with administrative boundaries containing the polygons of each district was imported to the PostgreSQL database in the same way and was updated with various new columns, which will be discussed in section *3.1.2 Data creation*. City-district Eisenach (NUTS3-ID: 16056) within rural district Wartburg united to

one district with the surrounding district „Landkreis Wartburg “(NUTS3-ID:16063) in the year 2019. As population counts were taken from before the COVID-19 pandemic (year 2019), the population dataset from Destatis was still based on the old district organization, but the RKI-corona infection and mortality register uses the new administrative districts. I solved this case of data inconsistence by summing up the population counts of the originally two districts and merging the polygons during the process of data preparation.

District aggregated mean living space per inhabitant (Statistische Ämter des Bundes und der Länder 2022) from the year 2011 nationwide census is acquired as a shapefile and joined according to the NUTS3 code to the district table in the SQL database. The special cases of the City-States (Berlin, Hamburg, Bremen) use in the mean living space dataset their NUTS1 code and must be manually transformed to NUTS3 code for a successful table join.

Since the last census a large restructuring and merging of districts has happened in the state of Mecklenburg-Vorpommern in Germany, which made it necessary to manually assign the mean living space to be the mean of the merged districts.

The mean age per district table needed the same manual adaption for the three City-States like the living space per inhabitant data.

Income data (Statistisches Landesamt Baden-Würtenberg 2020) for each district was extracted from an *.xlsx spreadsheet and imported as a new table in the PostgreSQL database, where the data was joined with the district dataset according to the NUTS3 code. Like the mean living space per inhabitant or the mean age per district datasets, the City-States needed special attention.

The vaccination dataset contains information for each district about age-group, vaccination status, date, and count of vaccinations. The vaccination status is either 1 for first vaccination, 2 for a full vaccination and 3 for a booster vaccination. Additionally, there are two special cases for the district ID value:
17000, meaning the vaccination has been applied by a governmental institution which does not report district ID; ‘u’ for unknown or not reported. These special cases make up 0.4% of all vaccinations. The data for governmental institutions and unknown locations were not used in this research which introduced a neglectable error.

The RKI does for reasons of privacy not publish data when a reported count is below 5. Missing counts will be aggregated until the day they exceed the count of 5 and will be assigned to that date. This minor data bias can be neglected, because no single but multiple days were evaluated. As the vaccination reports are split into various age-classes, the following two relevant groups will be merged and evaluated: greater than or equal to 60 years and below 60 years.

3.1.2 Data creation

For statistical analysis, a variety of potentially dependent and independent explanatory variables was deduced from the existing datasets on a district scale and added as new columns to the administrative boundary shapefile:

Independent variables:

- Mean Rn soil concentration and cumulative sum for each district, as described in more detail in section 3.1.2.1
- A population-count weighted Rn soil concentration mean, as described in more detail in section 3.1.2.1
- Population count

For comparison, other potential independent variables will be added:

- Mean age
- Mean population density
- Mean income per inhabitant
- Mean living space per inhabitant
- Full vaccinations per inhabitant
- Booster vaccinations per inhabitant

Dependent Variable columns are created by SQL queries:

- Infection counts per inhabitant
- Case fatality ratio (CFR) defined as deaths per registered infection
- Counts of ICU per registered infection
- Counts of ventilated ICU per registered infection

The variables case fatality ratio, infections per inhabitant and vaccinations are created for the total population, for the population greater than or equal to 60 years and below 60 years. The required population data is available at Destatis, and the Robert-Koch-Institute's register has all infection, death case and vaccination reports classified in corresponding age groups.

The Divi register for ICU numbers does not contain information about the patients age and will not be evaluated for specific age groups, but only for the total population.

3.1.2.1 Independent or explanatory variables

To weight radon soil concentration with population, it is necessary to create a grid with each km² percentage (weight) of the total district population.
This is calculated by rasterizing the districts with the values of their total population and then dividing the census grid with it.

The result is a district - population weight raster, which can be multiplied with the radon dataset to receive a weighted radon mean for each district's polygon.
The use of a population density adjusted radon soil concentration incorporates the effects of heterogeneity in space. By weighting with population, bias for the mean radon concentration will be reduced. For example, a cell containing 20% of a district's population will proportionally impact the district's mean radon soil concentration.

The original radon data is converted from a json-shapefile with a regular 1km² grid to a raster dataset, conserving its original resolution, but reprojecting to the population raster's projection. The unit of the radon concentration is thousands Bq/m³.

To add the sum or mean of the radon data to the district shapefile, a zonal statistic tool is used, which creates new columns for each method (sum, mean, max, min…). QGIS can perform this process directly on the PostgreSQL database.

Additionally, mean age of a district's population, mean living space, mean income and population density are added as new columns to the district shapefile, to account for these variables.

3.1.2.2 Dependent Variables:

Variables for each district were calculated and added with pgAdmin, the graphical user interface (GUI) for administrating PostgreSQL databases.

Samples of SQL-queries, for example to derive the sum of total death cases per district from the Robert-Koch-Institute's dataset and fill the result in a new column are attached in Appendix 8.4.

With SQL-queries, for each study period new columns are created and filled with processed data for:

- Total and each age groups sum of death cases,
- Total and each age groups ratio of infections per inhabitant,
- Total and each age groups case fatality ratio (CFR),
- Total sum and ratio of ICU cases per infection
- Total sum and ratio of ventilated ICU cases per infection

Another special case, that needed manual adaption, was caused by the fact, that the RKI- COVID-19 reports split the special case of City-State Berlin into multiple various NUTS3 codes, which don't exist in the administrative boundary dataset. The Berlin counts had to be summed up and filled into the table by an individual SQL-query (Appendix 8.4).

3.2 Analysis

To perform statistical analysis of the existing and created spatial datasets, the powerful Open-Source spatial analysis software GeoDa was used. GeoDa, is a leading tool for exploratory analysis, to create scatterplots and scatterplot matrices, assess spatial autocorrelation and perform a spatial regression. GeoDa's provides a graphical user-friendly interface with a focus on spatial data analysis (Center for Spatial Data Science 2022) and can access directly the data within the PostgreSQL database.

For a spatial analysis and assessing spatial autocorrelation, a weight matrix, defining the weight of interactions between neighbours, was created. In this study, one contiguity based and two distance based approaches were chosen:

- k-nearest neighbour, where k = 4
- Euclidian distance neighbour, where d = 30km
- first order queen contiguity

In spatial analysis, to assess the similarity of a point of interest to its neighbours (spatial autocorrelation, Section 3.2.1.1), the neighbourhood can either be defined by the sharing of borders (contiguity decay) or by (centroids) being within a certain distance (distance decay).
The queen continuity defines polygons which share vertices as a neighbourhood. A first order contiguity includes only direct neighbours. The chosen distance of 30 km is the maximum distance of 80 % of German commuters (Bundesministerium für Bau- Stadt - und Raumforschung 2019) and yielded comparable results to the distance decay of 4 nearest neighbours and first order queen contiguity. Converging results from both approaches have been remarked by others (Aturinde 2020; Anselin 2020).

The COVID-19 CFR and infection rates of the total population, of the age group greater than or equal to 60 years and of the age group less than 60 years for time periods I, II, and III will be analysed.

3.2.1 Exploratory Analysis and Geovisualisation

After preparation and creation of the datasets mentioned in 3.1, GeoDa Software was used to discover patterns and perform statistical analyses, while corresponding maps in QGIS were designed for a visual comparison of the distributions in Germany: Where are districts of high mean age (Figure 4, Section 2), low mean income (Figure 5, Section 2)? Where are districts of high infections rates (Figure 8; Figures 21-28, Appendix), high case fatality ratios, (Figure 7; Figures 13-20,

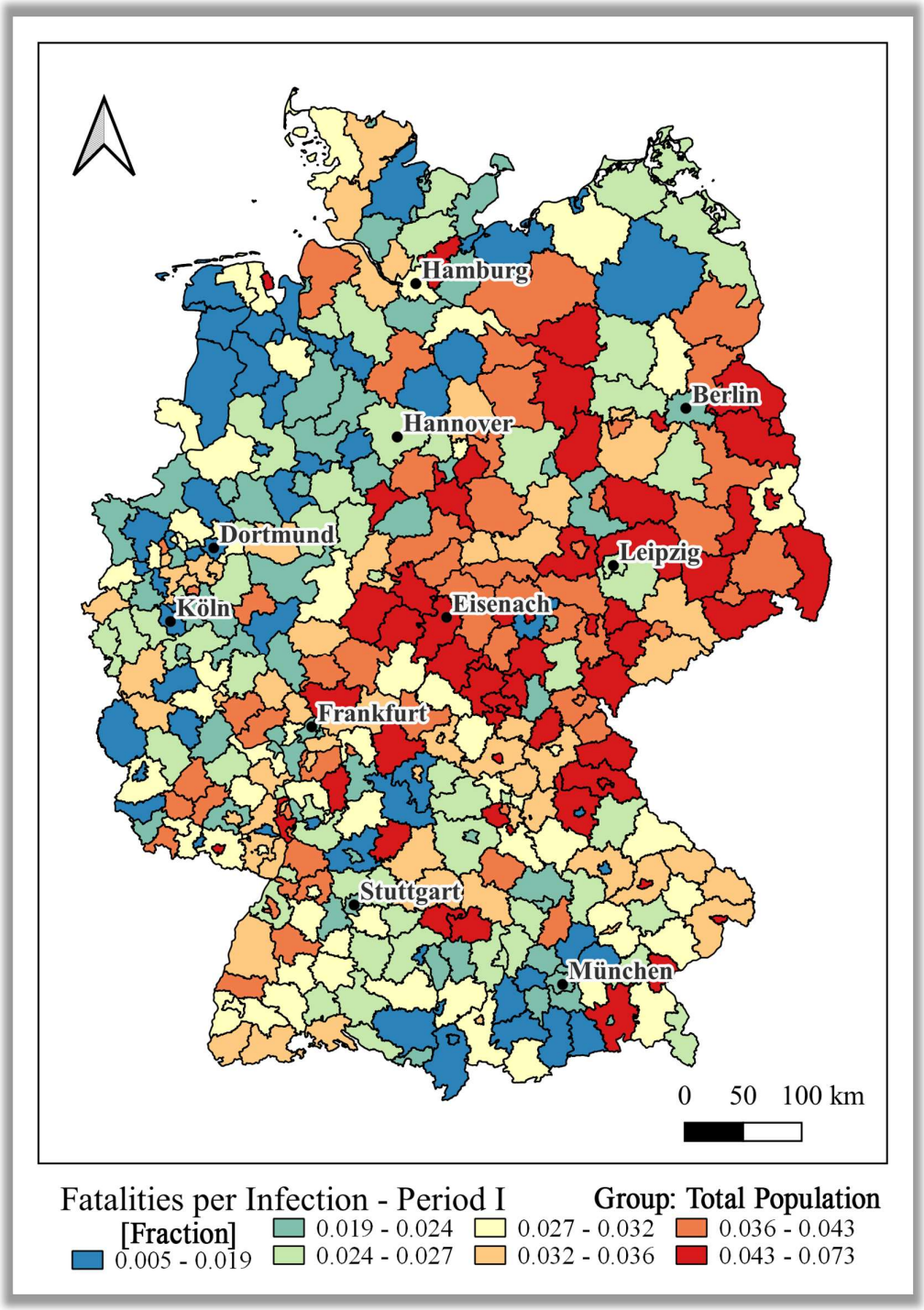


Figure 7 - Case Fatality Ratio of Total Population in Period I.

Appendix) or regions of high counts of COVID-19 caused occupation of ICU (Figure 11, Section 4.1)?
Which district have a high mean radon soil concentration (Figure 30, Appendix), which districts have a high population weighted mean radon soil concentration

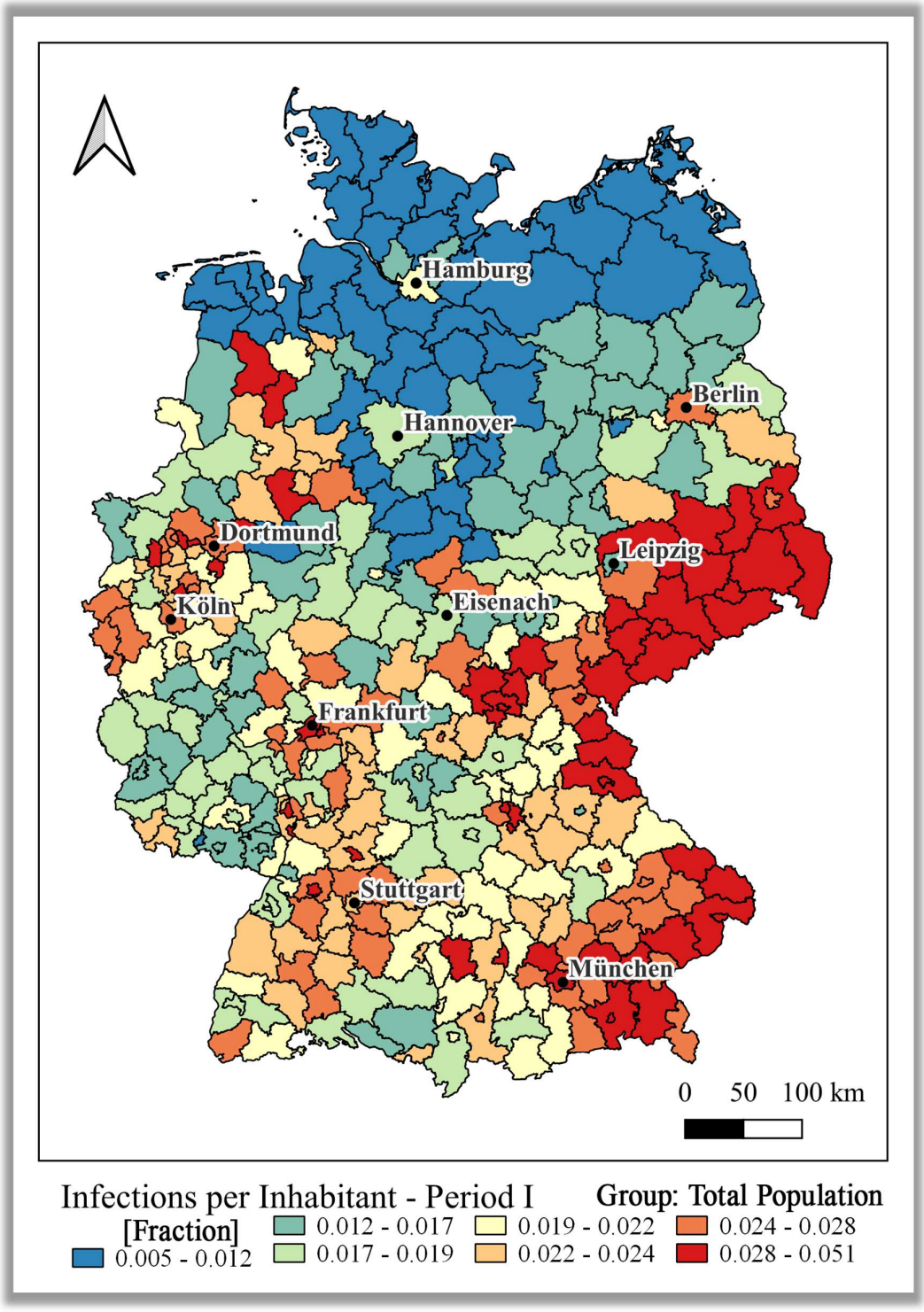


Figure 8 - Infections per Inhabitant of Total Population in Period I.

(Figure 29, Appendix)? Chosen histograms, scatter plots and tables of basic statistical properties are presented in the results and the appendix.

3.2.1.1 Global Spatial Autocorrelation

Spatial autocorrelation is a phenomenon also described as spatial dependence. Methods for testing in nonspatial statistics commonly rely on the hypothesis, that samples are randomly chosen, and the observations are independent, but in spatial data this hypothesis must often be rejected (Grekousis 2020).
Tobler's first law of geography states that:

> *"Everything is related to everything else, but near things are more related than distant things"*
>
> (Tobler 1970)

This law is the foundation of spatial autocorrelation and the phenomenon, that observations at locally close samples can be influenced by neighbouring samples.

Datasets in which I suggested clustering or hotspots, were checked to verify the rejection of a null hypothesis of spatial randomness. An ideal statistical index to assess global spatial autocorrelation, i.e. influencing neighbourhoods, is Moran's I (Oden 1995). Moran's I compares each observation with the observations at its spatial neighbourhood and indicates the mean of all these comparisons in a dataset (Equation 1). A perfect positive spatial autocorrelation results in a Moran's I of 1, spatial randomness in a Moran's I of 0 and a perfect negative spatial autocorrelation in a Moran's I of -1 (Figure 9).

$$I = \frac{n}{W} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{X})(x_j - \bar{X})}{\sum_{i=1}^{n}(x_i - \bar{X})^2}$$

I	Moran's criterion,
n	number of objects in space,
$\bar{X}$	a mean value of an attribute,
x_i, x_j	values of an attribute for objects i and j,
w	a spatial weight for a pair of objects,
W	the sum of spatial weights.

Equation 1 - Calculation for Moran's Criterion.

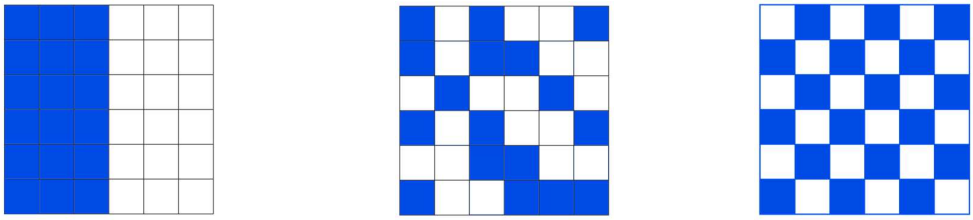


Figure 9 – Spatial Autocorrelation from Perfect Positive via Spatial Randomness to Perfect Negative Spatial Autocorreleation}.

Typical epidemiological studies focus on identifying and explaining spatial and temporal hotspots and outbreaks (Shaweno et al. 2018). Due to the high attention COVID-19 has received, hotspots and outbreaks were constantly monitored through the published incidence rates of each district and the focus in this study should lay on explaining hotspots and the reasons for clustering.

I refrained from the assessment of Local Spatial Autocorrelation, as individual districts correlation to their neighbours is not relevant in this project.

3.2.1.2 Scatter Plot

A Scatter Plot or multiple ones in a Scatter Plot Matrix (SPLOM) can be used to visualise bivariate relationships between various pair combinations of variables and can give indices to explanatory variables or the strength of correlations. Although several variables can be explored, this method is not multivariate but rather bivariate (Anselin 2020). This method was useful to choose other variables than radon soil concentration to perform correlation tests with mortality and infections on.
GeoDa also implements a linear regression fit to the scatter plot and can additionally add a smoothed local, non-linear regression fit, which helps identify non-linear relationships (Cleveland 1979). A Parallel Coordinate Plot (PCP), suggested to visualise multivariate clusters (Inselberg 1985) and a common tool and method in multivariate analysis (Anselin 2020), did not produce clear patterns for the clusters and I refrained from the further use.

3.2.1.3 Hypotheses Test

To verify the three hypotheses:

1: High COVID-19-associated case-fatality-ratio is spatially linked to high radon concentration in soil air.
2: High COVID-19-associated ICU admission rate is spatially linked to high radon concentration in soil air.
3: High COVID-19 infection rate is spatially linked to high radon concentration in soil air.

The null-hypothesis of a not statistically significant Pearson Correlation needs to be rejected, which is allowed if the p-value is below the chosen significance level of 0.05.
The Pearson Correlation is used to measure the strength of a linear relationship between two variables. As the relationship between radon and COVID-19 severity indicators could be non-linear, the Scatter Plots helped to give further indications about potential non-linear correlation types and significance of correlations, which were then used for modelling, described in the next chapter.

While the correlation does not give information about causality, the directional relationship between the variables is apparently assumed as stated in the introduction.

3.2.2 Spatial Modelling

3.2.2.1 Regression

Regression models are commonly used in statistical analysis to quantify the relationships between multiple predicting variables and a dependent variable. For a regression model, it is crucial, that predicting independent variables show very little to no collinearity (multicollinearity), to reduce the standard error. Multicollinearity is present, when predicting variables are correlated to each other.
In the field of geographical information, a common approach for a regression analysis is the spatial regression, because a regular, non-spatial regression model neglects the effects of spatial autocorrelation.
Based on the findings in the exploratory analysis a spatial regression was used to predict for each study period the COVID-19 severity indicators case-fatality-ratio and infections per inhabitant. To account for the identified spatial autocorrelation, a spatial lag model was chosen. The neighbourhood weight matrix was defined by a first order queens contiguity approach as described in section 3.2.

4 Results

In this section, I will present the results of spatial autocorrelation and correlation analysis of independent and dependent variables. Furthermore, I will present the results of the performed regressions.

4.1 Global Moran's I

The assessment of each independent and dependent variable's spatial autocorrelation revealed if data is randomly distributed, or patterns of clustering are present.

Table 1 - Moran's I of independent Variables.

	Moran's I	Neighbourhood Model		
	Variable	4 Nearest Neighbours	30km Distance Decay	Queen 1st Order
Independent Variables	Mean Radon Soil Gas Concentration	0.7	0.7	0.72
	Population Weighted Mean Radon	0.68	0.71	0.69
	Mean Income per Inhabitant	0.52	0.5	0.48
	Mean Age of Population	0.53	0.45	0.56
	Mean Living Space per Inhabitant	0.5	0.45	0.54

All independent variables showed a relatively strong positive spatial autocorrelation, which justified the use of a spatial regression instead of a standard linear regression.

Vaccination rate indicators, which were defined as applied second vaccinations (fully vaccinated) per inhabitant and applied booster vaccinations (boosted) per inhabitant, showed some weak negative spatial autocorrelation (Table 2).

Table 2 - Moran's I of Vaccination Rates.

		Moran's I		Neighbourhood Model		
				4 Nearest Neighbours	30km Distance Decay	Queen 1st Order
Administered Vaccinations / Inhabitant	Period II	Total Population	fully vaccinated	-0.15	-0.29	-0.18
			boosted	0.13	0.08	0.14
		Population >= 60 years	fully vaccinated	-0.17	-0.28	-0.20
			boosted	0.17	0.13	0.16
		Population < 60 years	fully vaccinated	-0.12	-0.26	-0.14
			boosted	0.12	0.06	0.14
	Period III	Total Population	fully vaccinated	-0.15	-0.29	-0.19
			boosted	-0.05	-0.14	-0.07
		Population >= 60 years	fully vaccinated	-0.17	-0.28	-0.2
			boosted	-0.05	-0.14	-0.08
		Population < 60 years	fully vaccinated	-0.12	-0.27	-0.16
			boosted	-0.03	-0.13	-0.05

The data source for vaccinations does not contain the residency district of vaccinated individual, but the district of vaccination administration. As public large vaccination centres, which accounted for a large quantity of administered vaccinations, were not offered in every district, vaccination rates were higher in districts with vaccination centres and lower in the surrounding or neighbouring districts (Figure 10. Due to the fact, that some districts have higher numbers of administered vaccinations than numbers of inhabitants, values above 1 are present.

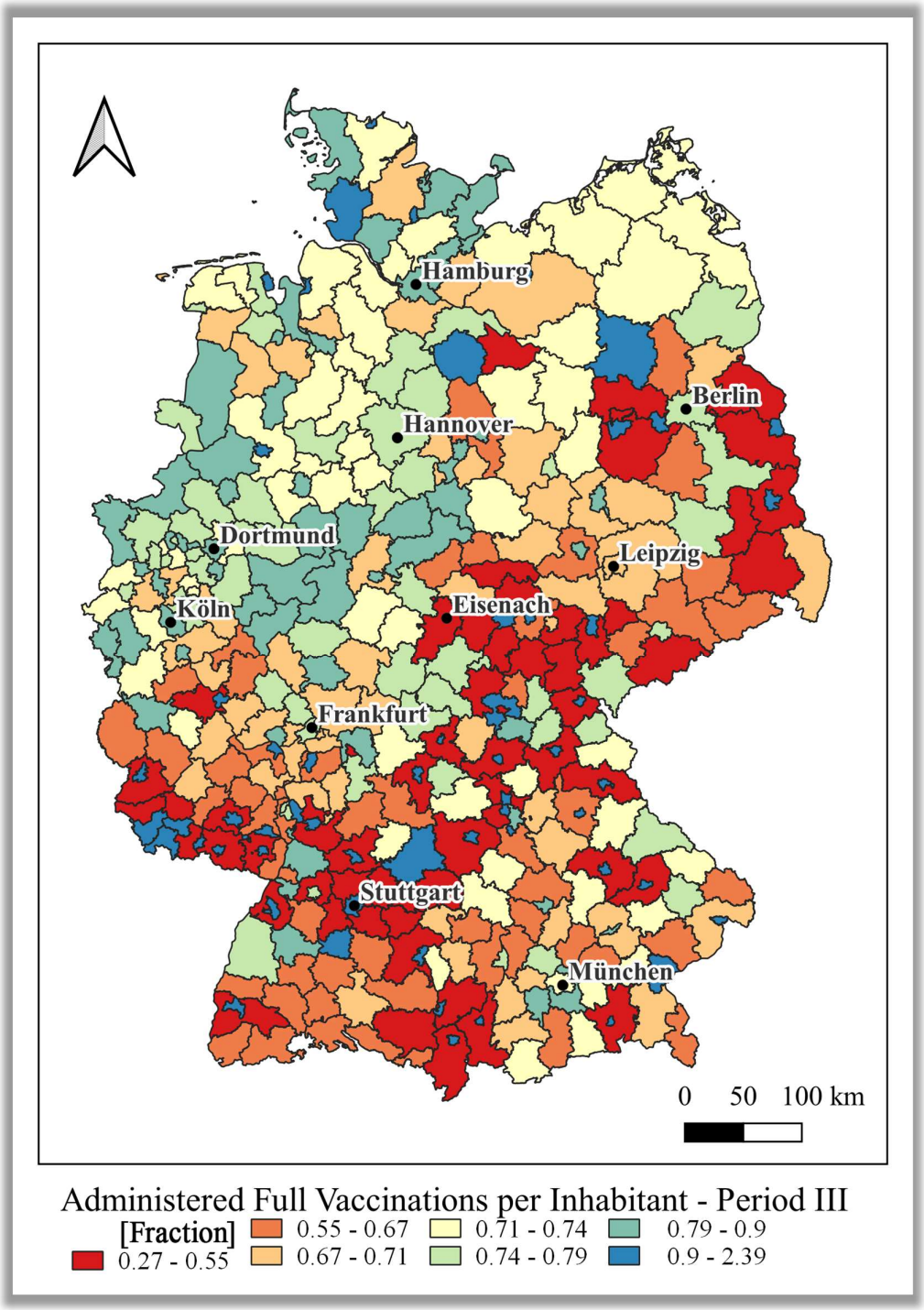


Figure 10 - Administered Full Vaccinations in Period III.

Both ICU admission rates, regular ICU and invasive ICU, also showed a weak negative spatial autocorrelation (Table 3), resulting in a checkerboard-like pattern of high and low ICU admission rates between neighbouring districts. This is caused by the fact, that patients are often not admitted to ICU in the districts they are residing, but in neighbouring districts with larger hospitals and more available ICUs (Figure 11).
ICU counts and vaccination rates therefore were not further used in this research.

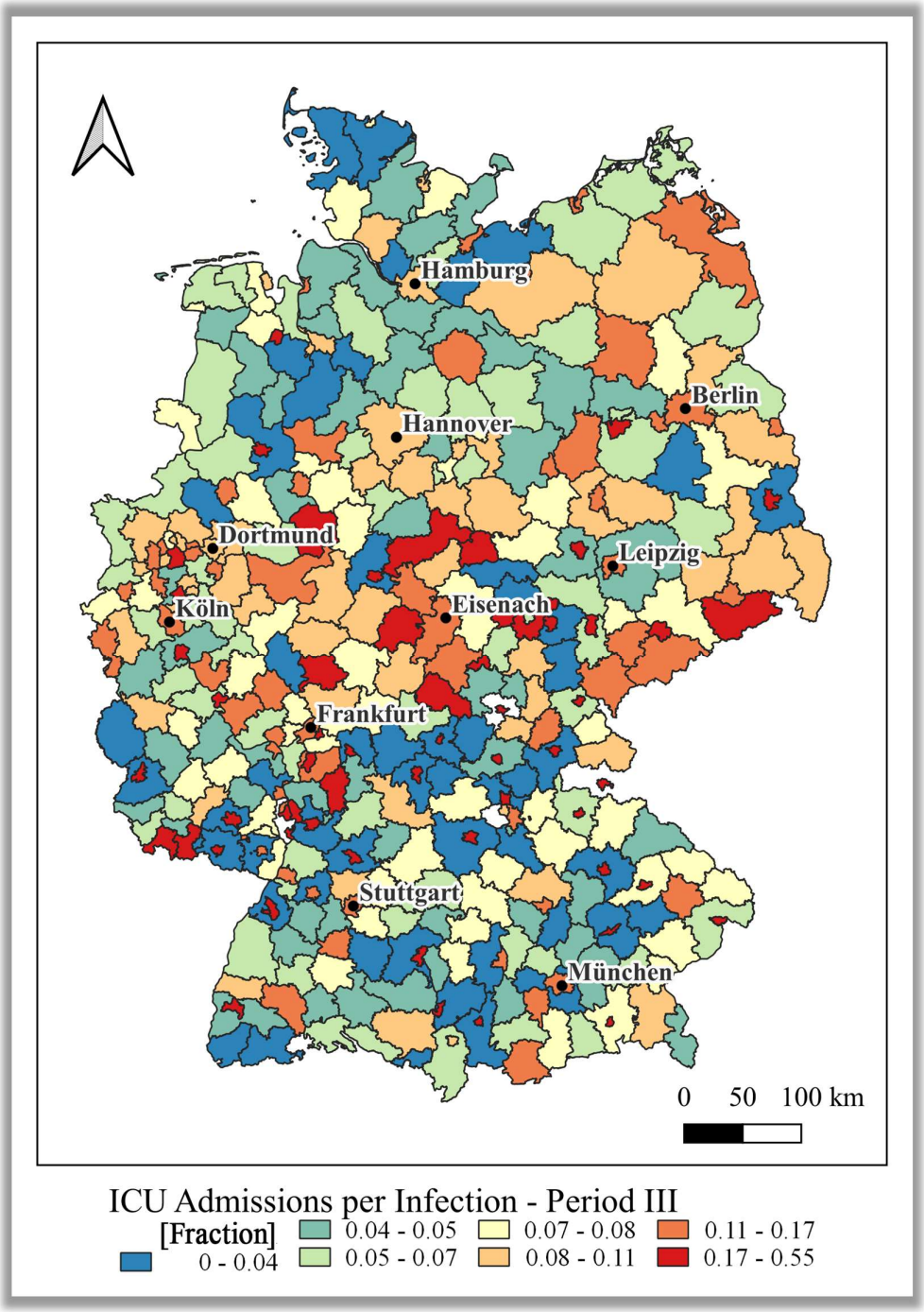


Figure 11 - ICU Admissions in Period III.

Table 3 - Moran's I of ICU and Invasive ICU Admissions of COVID-19 Patients.

	Moran's I		Neighbourhood Model		
			4 Nearest Neighbours	30km Distance Decay	Queen 1st Order
ICU admissions	**Period I**	regular ICU	-0.1	-0.21	-0.1
		invasive ICU	-0.09	-0.21	-0.09
	Period II	regular ICU	-0.08	-0.21	-0.05
		invasive ICU	-0.05	-0.15	-0.03
	Period III	regular ICU	-0.06	-0.18	-0.06
		invasive ICU	-0.06	-0.17	-0.05

Table 4 - Moran's I Infections per Inhabitant.

	Moran's I	Variable	Neighbourhood Model		
			4 Nearest Neighbours	30km Distance Decay	Queen 1st Order
Infections per Inhabitant	**Period I**	Total Population	0.69	0.68	0.7
		Population >= 60 years	0.68	0.68	0.69
		Population < 60 years	0.67	0.66	0.69
	Period II	Total Population	0.81	0.77	0.87
		Population >= 60 years	0.79	0.75	0.84
		Population < 60 years	0.8	0.76	0.86
	Period III	Total Population	0.73	0.68	0.79
		Population >= 60 years	0.76	0.71	0.81
		Population < 60 years	0.72	0.67	0.78

Infection rates showed a strong spatial autocorrelation and clustering in all periods and age groups (Table 4), in contrast to the case fatality ratio, which showed varying spatial autocorrelation in the different study periods and age groups (Table 5).

Table 5 - Moran's I Case Fatality Ratios.

	Moran's I	Variable	Neighbourhood Model		
			4 Nearest Neighbours	30km Distance Decay	Queen 1st Order
Case Fatality Ratio	**Period I**	Total Population	0.3	0.32	0.27
		Population >= 60 years	0.16	0.18	0.134
		Population < 60 years	0.09	0.1	0.04
	Period II	Total Population	0.26	0.31	0.26
		Population >= 60 years	0.13	0.15	0.14
		Population < 60 years	0.03	0.05	0.01
	Period III	Total Population	0.55	0.6	0.55
		Population >= 60 years	0.33	0.38	0.32
		Population < 60 years	0.29	0.39	0.25

4.2 Correlations

In this section, the main findings of the correlation analysis will be presented. Correlation factors with a significance level of $P < 0.05$ are marked with one asterisk, a P-value < 0.01 is marked with two asterisks

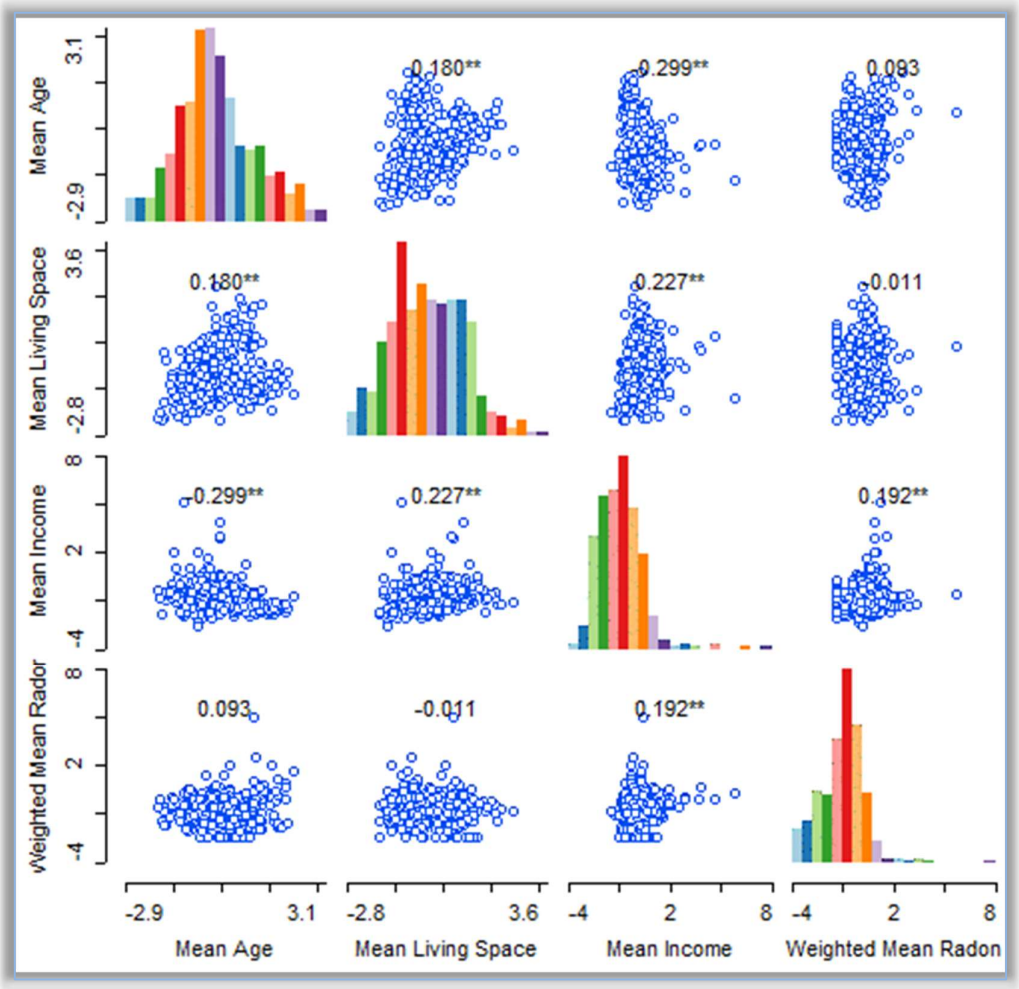


Figure 12 - Scatter Plot Matrix of Standardised Independent Variables and Correlation Coefficient: Weighted Mean Radon Soil Gas Concentration, Mean Income, Mean Age, Mean Living Space; Units are standard deviations (σ) . Districts represented by Blue Dots; Value Range in Histograms Split into Twenty Differently Coloured Bars.

In the scatter plot matrix (Figure 12) weak correlations between the four chosen independent variables were identified at significant levels. For better comparison, each variable values have been standardized so that the mean equals zero and the standard deviation equals 1. Consequently the units in figure 12 are standard deviations (**σ**). Population weighted mean radon soil gas concentration and mean income have a significant, but low correlation of 0.192, while the non-weighted mean radon soil gas concentration has a correlation of 0.230. Especially the variables mean age, mean income and mean living space per inhabitant indicate a

weak multicollinearity. The histograms revealed positive skewed distributions of radon and income.

Table 6 - Correlations between Infections per Inhabitant and Radon Soil Gas Concentration.

			Radon Soil Gas Concentration	
		Correlation	Mean	population weighted mean
Infections per Inhabitant	**Period I**	Total Population	0.37**	0.41**
		Population >= 60 years	0.39**	0.42**
		Population < 60 years	0.36**	0.39**
	Period II	Total Population	0.38**	0.38**
		Population >= 60 years	0.4**	0.4**
		Population < 60 years	0.37**	0.37**
	Period III	Total Population	0.35**	0.38**
		Population >= 60 years	0.41**	0.45**
		Population < 60 years	0.34**	0.36**

Compared to the mean radon soil gas concentration, the use of the population weighted mean radon soil gas concentration yielded in all ages and periods a higher correlation to infections per inhabitant (Table 6). The coefficients suggest a weak, but significant correlation between radon and infection rates.
In comparison to the unweighted radon mean, the weighted radon mean also showed slightly stronger correlations to the case fatality ratio (Table 7). Although the correlations increase in period II and III, the coefficients indicate no to very low correlations in all periods.

Table 7 - Correlations between Case Fatality Ratio and Radon Soil Gas Concentration.

			Radon Soil Gas Concentration	
		Correlation	Mean	population weighted mean
Case Fatality Ratio	**Period I**	Total Population	0.18**	0.2**
		Population >= 60 years	0.09	0.1
		Population < 60 years	-0.01	-0.04
	Period II	Total Population	0.21**	0.23**
		Population >= 60 years	0.16*	0.16**
		Population < 60 years	0.04	0.07
	Period III	Total Population	0.28**	0.31**
		Population >= 60 years	0.2**	0.21**
		Population < 60 years	0.19**	0.22**

There were no significant or only weak negative correlations between mean living space per inhabitant and case fatality ratio, the strongest being -0.2 during period II in the age group greater than or equal to 60 years.
In contrast, but in line with other research (Universitätsmedizin and der Johannes Gutenberg-Universität Mainz 2021), the negative correlation between mean living space and infection rate was stronger and in all periods significant, ranging between -0.251 up to -0.42 (Appendix 8.1.3).

Table 8 - Correlations between Case Fatality Ratio and Mean Age.

		Correlation	Mean Age
Case Fatality Ratio	**Period I**	Total Population	0.43**
		Population >= 60 years	0.02
		Population < 60 years	0.14*
	Period II	Total Population	0.33**
		Population >= 60 years	0.1
		Population < 60 years	0.02
	Period III	Total Population	0.6**
		Population >= 60 years	0.29**
		Population < 60 years	0.3**

The scatterplots of mean age versus case fatality ratio (Appendix 8.1.5) indicated especially in the total population sample of each period a medium strong positive correlation of up to 0.599 (Table 8), whereas the correlation to infections per inhabitant showed no consistency within age groups and periods (Table 9). Only period II showed a significant correlation between mean age and infections per inhabitant.

Table 9 - Correlations between Infections per Inhabitant and Mean Age.

		Correlation	Mean Age
Infections per Inhabitant	**Period I**	Total Population	-0.07
		Population >= 60 years	0.03
		Population < 60 years	-0.09
	Period II	Total Population	0.25**
		Population >= 60 years	0.22**
		Population < 60 years	0.31**
	Period III	Total Population	-0.03
		Population >= 60 years	0.14*
		Population < 60 years	0.09

Table 10 - Correlations between Infections per Inhabitant and Mean Income.

		Correlation	Mean Income
Infections per Inhabitant	**Period I**	Total Population	0.1*
		Population >= 60 years	0.07*
		Population < 60 years	0.11*
	Period II	Total Population	0.01
		Population >= 60 years	0.04
		Population < 60 years	-0.02
	Period III	Total Population	0.11*
		Population >= 60 years	0.03
		Population < 60 years	0.07

Table 11 - Correlations of Case Fatality Ratios to Mean Income.

		Correlation	Mean Income
Case Fatality Ratio	**Period I**	Total Population	-0.16**
		Population >= 60 years	-0.04
		Population < 60 years	-0.2**
	Period II	Total Population	-0.24**
		Population >= 60 years	-0.18**
		Population < 60 years	-0.15**
	Period III	Total Population	-0.31**
		Population >= 60 years	-0.24**
		Population < 60 years	-0.29**

While mean income showed no significant correlation to infections per inhabitant, a negative correlation to case fatality ratio was present in all three periods (Tables 10, 11). Contrary to other research findings (Wong and Li 2020; Rocklöv and Sjödin 2020), a consistent correlation of population density and infections per inhabitant could not be identified (Table 12).

Table 12 - Correlations of Infections per Inhabitant to Population Density.

		Correlation	Population Density
Infections per Inhabitant	**Period I**	Total Population	0.24**
		Population >= 60 years	0.12*
		Population < 60 years	0.27**
	Period II	Total Population	-0.17*
		Population >= 60 years	-0.18**
		Population < 60 years	-0.18**
	Period III	Total Population	0.09
		Population >= 60 years	0
		Population < 60 years	0.04

Especially in the total population and in the age group equal to or greater than 60 years infections per inhabitant showed weak correlations to case fatality ratio in all periods (Table 13).

Table 13 - Correlations of Infections per Inhabitant and Case Fatality Ratio.

		Correlation	Case Fatality Ratio
Infections per Inhabitant	**Period I**	Total Population	0.25**
		Population >= 60 years	0.33*
		Population < 60 years	0.03
	Period II	Total Population	0.33**
		Population >= 60 years	0.28**
		Population < 60 years	0.06
	Period III	Total Population	0.27**
		Population >= 60 years	0.37**
		Population < 60 years	0.21**

4.3 Regression Analysis

Based on the findings from the correlation analysis and the assessment of spatial autocorrelation, I chose to perform spatial regressions:

- to predict total case fatality ratio with the predictors:
  - mean income, mean age, weighted mean radon soil gas concentration

- to predict total infections per inhabitant with the predictors:
  - mean living space and weighted mean radon soil gas concentration

Detailed regression reports can be found in the Appendix (Section 8.3).

Variation of the dependent variable case fatality ratio could be explained between 22% and 59% by the predictors, depending on the study period (Table 14). In period I, the variable mean income showed a non-significant coefficient, in period II, the constant-variable of the regression model had an insignificant coefficient, but in period III, all variable coefficients were significant at least at the 0.01 level (Appendix 8.3.1).

Table 14 - Regression Indicators for Total Case Fatality Ratios.

	Case Fatality Ratio		
Regression	**Period I**	Period II	Period III
Akaike Info Criterion	-2559	-3408	-3669
R-Squared	0.27	0.23	0.59
Log Likelihood	1285	1709	1840

Adding the infections per inhabitant variable as a predictor to predict CFR improved in all periods the regression model at a neglectable level.

When analysing the population equal to or greater than 60 years, the regression model to predict CFR without the predictor mean age (Table 15) yielded less robust results in comparison to the regression model with infections per inhabitant added as predictor (Table 16)

Table 15 - Regression Indicators for CFR in Age Group greater than or equal to 60 Years.

	CFR Age Group greater than or equal to 60 Years		
Regression	Period I	Period II	Period III
Akaike Info Criterion	-1702	-2092	-2329
R-Squared	0.06	0.1	0.25
Log Likelihood	855	1050	1169
(Predictors: Population Weighted Radon Mean, Mean Income)			

Table 16 - Regression Indicators for CFR in Age Group greater than or equal to 60 Years; Infection Rate as Additional Predictor.

	CFR Age Group greater than or equal to 60 Years		
Regression	Period I	Period II	Period III
Akaike Info Criterion	-1732	-2109	-2356
R-Squared	0.13	0.13	0.3
Log Likelihood	871	1059	1183
(Predictors: Population Weighted Radon Mean, Mean Income, Infections per Inhabitant)			

In all study periods over 70% of the variability in infections per inhabitant could be explained with the independent variables mean living space and population weighted mean radon soil gas concentration (Table 17). The variable's coefficients were all significant at the 0.01 level except in period II, where they were significant at the 0.05 level (Appendix 8.3.2).

Table 17 - Regression Indicators for Infections per Inhabitant in All Periods.

	Infections per Inhabitant		
Regression	Period I	Period II	Period III
Akaike Info Criterion	-3192	-2952	-1922
R-Squared	0.72	0.84	0.76
Log Likelihood	1600	1480	966

The regressions to predict the separate age group's infections per inhabitant produced similar results. In period II the coefficient of determination was the highest and in the age group greater than or equal to 60 years, the model showed stronger results than in periods I and III.

Table 18 - Regression Indicators for Infections per Inhabitant in Age Group greater than or equal to 60 Years.

	Infections per Inhabitant greater than or equal to 60 Years		
Regression	Period I	Period II	Period III
Akaike Info Criterion	-3145	-3336	-2345
R-Squared	0.69	0.8	0.79
Log Likelihood	1576	1672	1176

Table 19 - Regression Indicators for Infections per Inhabitant in Age Group below 60 Years.

	Infections per Inhabitant below 60 Years		
Regression	Period I	Period II	Period III
Akaike Info Criterion	-3153	-2772	-1757
R-Squared	0.7	0.83	0.75
Log Likelihood	1581	1390	883

Hypothesis:

2: High COVID-19-associated ICU admission rate is spatially linked to high radon concentration in soil air.

was not further tested, because available intensive care unit admission data was biased and did not support conclusions.

The hypotheses:

1: High COVID-19-associated case-fatality-rate is spatially linked to high radon concentration in soil air.

and

3: High COVID-19 infection rate is spatially linked to high radon concentration in soil air.

cannot be rejected. Weak influences of radon soil gas concentration on case fatality ratio are indicated in period III and supported by the regression analysis. Weak correlations of radon and infection rate exist in all periods and age group at a

significant level. Especially for period II, the regression model using radon soil gas concentrations yielded better results when predicting infections per inhabitant.

The regression models to predict CFR in the separate age groups and periods had all relatively low coefficients of determination, except in period III, indicating a correlation between radon and CFR.

While population weighted mean radon soil gas concentration significantly improved the regressions for both, predicting case fatality ratio and predicting infections per inhabitant, the improvement was relatively low in comparison to the predictors mean living space for infection rate and to the predictors mean age and infection rate for case fatality ratio.

In the following, I will discuss my interpretation of the results and limitations of this research.

5 Discussion

In accordance with results of previous studies, I identified cramped or crowded housing conditions as a key driver for increase infection rate (Federgruen and Naha 2021; Universitätsmedizin and der Johannes Gutenberg-Universität Mainz 2021). My findings of income and age being correlated to case fatality ratio are consistent with the results of studies on Swedish data (Drefahl et al. 2020) and multi-national data (O'Driscoll et al. 2021). My approach of additionally accounting for radon soil gas concentration increased the predictive power of the regression models in all periods for both, case fatality ratio and infection rate.

The weak but significant correlations between radon soil gas concentration and infections per inhabitant (Table 6), can be interpreted as an indication of a higher infection rate caused by higher radon exhalation. Although the higher coefficients of determination of the regressions in period II are suggesting a higher infection rate after a nationwide low-pressure phenomenon ("Sturmtief Ignatz"), the smaller correlation coefficients between infection rate and radon soil gas concentrations in period II suggest no increase in radon exhalation caused by the low-pressure phenomenon.

The regression to predict case fatality ratio with the predictors mean age, mean income, and population weighted mean radon soil gas concentration showed a relatively high coefficient of determination in period III (Table 14). The much larger sample number of period III in comparison to periods I and II can be interpreted as a support of the conclusion, that effects of radon on the case fatality ratio are visible especially in a large scale and long-term analysis.

I consider both findings reasonable and the conclusions suggest a support of hypotheses 1 and 3. I refrained from further testing hypothesis 2, because the underlying data for ICU-admissions was biased at the district level.

However, the identified existing weak positive correlations between infections per inhabitant and case fatality ratio (Table 13) indicate, that in regions of high overall infection rates either the health system was overburdened, or mild progressions were less reported and the number of unreported infections was much higher, or that a combination of both conditions was present. This leads me to discuss the limitations of the research methodology and the used data.

The COVID-19 reports used in this research are not the total population of COVID-19 cases, but only observations (registered cases). The number of unreported or undetected cases for both, infections and death cases is not known. One large study (Universitätsmedizin and der Johannes Gutenberg-Universität Mainz 2021) conducted by the University Medical Center of the Johannes Gutenberg University Mainz in Germany identified over 40% of unregistered infections within participants. An additional finding of said study was a more frequent testing among younger population, resulting in a bias of detected COVID-19 infections depending on age groups.

As multiple factors influence COVID-19 transmission, disease progression and mortality, the exclusion of other controlling factors can result in wrong indications (data artifacts). Transmission has been shown to be dependent on multiple socioeconomic factors (Qiu et al. 2020). In this research I did not account individual lifestyle, amount of social interaction, or acceptance of virus transmission counter measurements like wearing filtering face pieces and face covering as well as keeping physical distance and getting regularly tested for SARS-CoV-2.
Areas of high incidence and mortality in Germany have been shown to be highly correlated to areas of high voting quotas of a right-wing political party (Richter et al. 2021). This party is known to be voted from people with low acceptance of COVID-19 counter measurements and a distrust in governmental structures, suggesting a bias. Interestingly, in some districts, mobility data has shown an increase in mobility during lockdowns and travel restriction times (Balterman et al. 2021), pointing to a high denial of lockdown and countermeasure rules, which could have led to higher infections, possibly unregistered and thus a higher number of fatalities or a higher case fatality ratio.

I had to refrain from accounting for vaccinations in study periods II and III, because the vaccination data is not representative for the districts vaccination rate.
High CFR and infection rates could entirely be caused by and within the group of not-vaccinated individuals. The vaccination quotas are varying between States by up to 25% (Robert Koch-Institut 2022). The varying transmission and mortality of virus variants (SeyedAlinaghi et al. 2021) is also not accounted for and could be cause of a large uncertainty. Before drawing any conclusion from the indication of correlations between radon soil gas concentration and case fatality further analysis is necessary.

Although associations of mortality and transmission to meteorological and climatic factors have been identified, the findings were inconsistent (Kerr et al. 2021). Considering, that Germany is separated in two different Köppen-Geiger climate classifications, the climatic differences should not be neglected. For a detailed analysis of meteorological effects, the temporal dimension should be included, which would have exceeded the scope of this thesis. The missing time component is also critical, as infection rates are strong associated to previous infection rates (Ganegoda et al. 2021). This autoregressive characteristic has been neglected in this research. The increased infection rate in period II could also be caused by a previous high infection rate and does not prove increased infection rates after an event of increased radon exhalation.

While I analysed not only the total population of each district, but also the age groups equal to or greater than 60 years and below 60 years, the demographic within these age groups can still vary greatly. With the publicly available data, a more age-specific analysis was not possible, due to varying age classifications of COVID-19 reports and census data. Median age is a better indicator in demographics than mean age, but the use as independent variable instead of mean age was due a lack of data not possible.

Data for mean living space per inhabitant and for the grid of population counts per 1km² was based on the census 2011 and were at the point of this research outdated

since a decade. The bias introduced through potential change in the structure of living situations and population distribution was neglected.

A significant clustered surplus of men, especially in younger generations, has developed within the last decades in the states of former East-Germany (Institut der Deutschen Wirtschaft Köln 2017). Researchers early identified male sex as a risk factor for death and ICU admission, after meta-analysing over 3 million cases from 46 countries (Peckham et al. 2020). This sex-bias was suggested to be caused by multiple sex-based differences like amongst others immune response suppressive effects of testosterone or an accelerated immune aging in males, as well as behavioural or socio-cultural differences based on gender identity. To account for the influence of sex, the genders of each age group should be separately analysed. The open available data for gender and age-group specific population counts at district level was recorded at the 2011 nationwide census and will be updated with the postponed 2021 census (Statistisches Bundesamt 2021) after the deadline of this project. Consequently, a gender specific analysis was not performed, and the gender mortality bias could not be accounted for.

In this research, I equated radon soil gas concentration with radon exposure. Radon soil gas concentration does not necessarily mean higher radon exhalation into the environment. Exhalation into surface air and into the atmosphere is largely dependent on geological properties like soil permeability and porosity (Nazaroff 1992), but is also weather dependent. As radon exhalation from soil into the atmosphere is driven by meteorological parameters like surface pressure or rainfall (Mudelsee et al. 2020) radon exposure is not constant but prone to hourly, daily, and seasonal variations of weather conditions (Tchorz-Trzeciakiewicz and Kłos 2017; Sesana et al. 2003). Neither a time-component, nor actual radon exhalation are included in this research.

The exposure to radon is influenced on additional factors:
Sealing of soils can cause a barrier for radon, which again can cause increased exhalation in areas of better gas exchange and permeability. Radon usually intrudes buildings through the floor, but also through the walls of subsurface basement rooms. When radon exhalation happens in non-confined space, radon concentration quickly decreases, depending on wind speed, and mixes with atmospheric air. In confined spaces, radon concentration increases, especially in colder months or in the heating period, in which ventilation with cold outside air is reduced and the stack effect creates a pressure gradient that facilitates intrusion from soil air (Rey et al. 2022). Many, but not all houses in Germany have an underground level, which can act as a large-scale entry for radon into buildings. Radon concentration in confined space is not only restricted to basement or ground floors. Depending on the design of a house, installed ventilation or fireplaces, radon can be sucked into all levels when pressure gradients between inside and outside favour an intrusion (Nazaroff 1992). It is common, that a house consists of different flats in each level. This results in unequal exposition of inhabitants, depending on the level of their flat.

These multiple factors suggest a high variance in individual exposure to radon at only little changed circumstances in areas of similar radon soil concentration.

Individual data like migration, living habits and housing situation needed to be evaluated to account for a radon exposure history.

As the histograms of the independent variables revealed skewed data and outliers, a log transformation should be applied. To account for vaccinations, all data should be aggregated to a larger scale. Working at a higher NUTS level, could also reduce the bias in ICU admission locations. But changing the scale and shape of data aggregation can produce very different results. Statistically significant relationships of the data at an aggregated higher NUTS level, does not necessarily reproduce in small local scale or individual level. That a relationship is true at the detailed level because it is true at the aggregated data level, is a common misconception (Grekousis 2020). This is also referred to as the modifiable area unit problem, which also poses a restriction at the scope of this research. The results in this study are not immune to this phenomenon and could suggest wrong conclusions.

6 Conclusion

In this research, I tested spatial correlations between mean radon soil gas concentrations and COVID-19 case fatality ratio as well as COVID-19 infection rates. A testing for correlation between radon soil gas concentrations and COVID-19 caused ICU admissions could not be performed, due to the lack of quality in the underlying data source.
Although I could not show strong correlations to case fatality ratio or infection rate, I demonstrated, that infection rates are weak but significant positive correlated to radon soil gas concentration in all analysed time frames. I further identified an indication of a significant positive correlation between radon and case fatality ratio in the period 01.2020- 02.2022.
However, the results of analysis in this period must be interpreted with caution, as vaccination quotas, a strong influencing factor, were neglected in that period.
Other large sources of uncertainty are the neglect of multiple unaccounted factors and the exclusion of a temporal dimension, necessary to account for the autoregressive characteristics of epidemiological processes.
I suggest the results of this research to be used as a foundation for further research. My demonstrated limitations should be considered in the development of more elaborated methods for testing and quantifying the impact of radon exposure to the severity of the COVID-19 pandemic.
The additional risk, that is possibly posed to the public health by radon exposure should be determined to raise awareness of the importance of radon exposure mitigation strategies and reduce not only cases of COVID-19 related infections and death, but also other respiratory diseases, such as pneumonia, COPD or lung cancer.

7 References

Al-Zoughool, M., and D. Krewski. 2009. Health effects of radon: A review of the literature. *International Journal of Radiation Biology,* 85**:** 57-69. DOI: 10.1080/09553000802635054

Anselin, L. 2020. GeoDa Workbook.

Ansoborlo, E. 2014. Poisonous polonium. *Nature Chemistry,* 6**:** 454-454. DOI: 10.1038/nchem.1928

Aturinde, A. 2020. GIS and Health: Enhancing Disease Surveillance and Intervention through Spatial Epidemiology.

Author. 2021. Die Mobilität in Deutschland geht zurück – aber nicht überall. *Tagesspiegel*.

Biktasheva, I. V. 2020. Role of a habitat's air humidity in Covid-19 mortality. *Science of The Total Environment,* 736**:** 138763. DOI: https://doi.org/10.1016/j.scitotenv.2020.138763

Bundesamt für Kartographie und Geodäsie. 2020. GeoBasis-DE VG250. ed. Bundesamt für Kartographie und Geodäsie.

Bundesamt für Kartographie und Geodäsie. 2022. INSPIRE. Retrieved 03.03.2022 2022, from. https://www.gdi-de.org/en/INSPIRE

Bundesamt für Strahlenschutz. 2022a. Radon-222 in Soil (90th Percentile, Prognosis).

Bundesamt für Strahlenschutz. 2022b. Radon in der Boden-Luft in Deutschland. Retrieved 22.02.2022, from. https://www.bfs.de/DE/themen/ion/umwelt/radon/karten/boden.html

Bundesministerium für Bau- Stadt - und Raumforschung. 2019. Pendeldistanzen und Pendlerverflechtungen. Retrieved, from. https://www.deutschlandatlas.bund.de/DE/Karten/Wie-wir-uns-bewegen/100-Pendlerdistanzen-Pendlerverflechtungen.html#_niktke4gm

Bundesrepublik Deutschland. 2013. Gesetz zur Förderung der elektronischen Verwaltung (E-Government-Gesetz - EGovG).

Center for Spatial Data Science. 2022. Introducing GeoDa 1.20. Retrieved, from. https://geodacenter.github.io

Cevik, M., K. Kuppalli, J. Kindrachuk, and M. Peiris. 2020. Virology, transmission, and pathogenesis of SARS-CoV-2. *BMJ,* 371**:** m3862. DOI: 10.1136/bmj.m3862

Cleveland, W. S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association,* 74**:** 829-836. DOI: 10.1080/01621459.1979.10481038

Drefahl, S., M. Wallace, E. Mussino, S. Aradhya, M. Kolk, M. Brandén, B. Malmberg, and G. Andersson. 2020. A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. *Nature Communications,* 11**:** 5097. DOI: 10.1038/s41467-020-18926-3

Federgruen, A., and S. Naha. 2021. Crowding Effects Dominate Demographic Attributes in COVID-19 Cases. *International Journal of Infectious Diseases,* 102**:** 509-516. DOI: https://doi.org/10.1016/j.ijid.2020.10.063

Franch-Pardo, I., B. M. Napoletano, F. Rosete-Verges, and L. Billa. 2020. Spatial analysis and GIS in the study of COVID-19. A review. *Science of The Total Environment,* 739**:** 140033. DOI: https://doi.org/10.1016/j.scitotenv.2020.140033

Ganegoda, N. C., K. P. Wijaya, M. Amadi, K. Erandi, and D. Aldila. 2021. Interrelationship between daily COVID-19 cases and average temperature as well as relative humidity in Germany. *Sci Rep,* 11**:** 11302. DOI: 10.1038/s41598-021-90873-5

Ganslmeier, M., D. Furceri, and J. D. Ostry. 2021. The impact of weather on COVID-19 pandemic. *Scientific Reports,* 11**:** 22027. DOI: 10.1038/s41598-021-01189-3

Grekousis, G. 2020. *Spatial Analysis Methods and Practice: Describe – Explore – Explain through GIS*. Cambridge: Cambridge University Press.

Huang, C., Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, et al. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The lancet,* 395**:** 497-506.

Inselberg, A. 1985. The plane with parallel coordinates. *The Visual Computer,* 1**:** 69-91. DOI: 10.1007/BF01898350

Institut der Deutschen Wirtschaft Köln, 2017. Geschlechterverhältnisse und Geburten in den deutschen Regionen. Report. [in Swedish, English summary]

Islam, N., Q. Bukhari, Y. Jameel, S. Shabnam, A. M. Erzurumluoglu, M. A. Siddique, J. M. Massaro, and R. B. D'Agostino. 2021. COVID-19 and climatic factors: A global analysis. *Environmental Research,* 193**:** 110355. DOI: https://doi.org/10.1016/j.envres.2020.110355

Jabłońska, K., S. Aballéa, and M. Toumi. 2021. The real-life impact of vaccination on COVID-19 mortality in Europe and Israel. *Public Health,* 198**:** 230-237. DOI: https://doi.org/10.1016/j.puhe.2021.07.037

Kabarriti, R., N. P. Brodin, M. I. Maron, W. A. Tomé, B. Halmos, C. Guha, S. Kalnicki, M. K. Garg, et al. 2020. Extent of Prior Lung Irradiation and Mortality in COVID-19 Patients With a Cancer History. *Advances in Radiation Oncology,* 5**:** 707-710. DOI: 10.1016/j.adro.2020.04.028

Kang, J. K., S. Seo, and Y. W. Jin. 2019. Health Effects of Radon Exposure. *Yonsei Med J,* 60**:** 597-603. DOI: 10.3349/ymj.2019.60.7.597

Kemski, J., A. Siehl, R. Stegemann, and M. Valdivia-Manchego. 2001. Mapping the geogenic radon potential in Germany. *Sci Total Environ,* 272**:** 217-230. DOI: 10.1016/s0048-9697(01)00696-9

Kerr, G. H., H. S. Badr, L. M. Gardner, J. Perez-Saez, and B. F. Zaitchik. 2021. Associations between meteorology and COVID-19 in early studies: Inconsistencies, uncertainties, and recommendations. *One Health,* 12**:** 100225. DOI: https://doi.org/10.1016/j.onehlt.2021.100225

Kulali, F., I. Akkurt, and N. Özgür. 2017. The effect of meteorological parameters on radon concentration in soil gas. *Acta Phys. Pol. A,* 132**:** 999-1001.

LaPlant, Q., M. Thor, N. Shaverdian, J. Y. Shin, P. Gilbo, J. Luo, D. R. Gomez, and D. Y. Gelblum. 2021. Association of prior radiation dose to the cardiopulmonary system with COVID-19 outcomes in patients with cancer. *Radiotherapy and Oncology,* 161**:** 115-117. DOI: 10.1016/j.radonc.2021.06.002

Maragakis, L. (2021). https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/coronavirus-and-covid19-who-is-at-higher-risk (last accessed 02.03.2022).

Maya, J., L. Mohamadou, S. Mbida Mbembe, A. Likéné, B. Mbembe, and B. Maimounatou. 2020. Radon Risks Assessment with the Covid-19 Lockdown Effects. *Journal of Applied Mathematics and Physics,* 08**:** 1402-1412. DOI: 10.4236/jamp.2020.87106

Mollalo, A., B. Vahedi, and K. M. Rivera. 2020. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of The Total Environment,* 728**:** 138884. DOI: https://doi.org/10.1016/j.scitotenv.2020.138884

Mudelsee, M., J. Albert, and F. Sirocko. 2020. Weather control in radon flux time series from Schleswig-Holstein, Germany. *GEM - International Journal on Geomathematics,* 11**:** 23. DOI: 10.1007/s13137-020-00156-w

Musa, G. J., P.-H. Chiang, T. Sylk, R. Bavley, W. Keating, B. Lakew, H.-C. Tsou, and C. W. Hoven. 2013. Use of GIS Mapping as a Public Health Tool-From Cholera to Cancer. *Health services insights,* 6**:** 111-116. DOI: 10.4137/HSI.S10471

Nanduri, L. S. Y., P. K. Duddempudi, W.-L. Yang, R. Tamarat, and C. Guha. 2021. Extracellular Vesicles for the Treatment of Radiation Injuries. *Frontiers in pharmacology,* 12**:** 662437-662437. DOI: 10.3389/fphar.2021.662437

National Institutes of Health. 2022. Coronavirus Disease 2019 (COVID-19) Treatment Guidelines. COVID-19 Treatment Guidelines Panel. Retrieved 2022-03-01, from. https://www.covid19treatmentguidelines.nih.gov/

Nazaroff, W. W. 1992. Radon transport from soil to air. *Reviews of Geophysics,* 30**:** 137-160. DOI: https://doi.org/10.1029/92RG00055

Neznal, M., M. Neznal, M. Matolin, I. Barnet, and J. Miksova. 2004. *The new method for assessing the radon risk of building sites*. Czech Geological Survey Prague.

O'Driscoll, M., G. Ribeiro Dos Santos, L. Wang, D. A. T. Cummings, A. S. Azman, J. Paireau, A. Fontanet, S. Cauchemez, et al. 2021. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature,* 590**:** 140-145. DOI: 10.1038/s41586-020-2918-0

Oden, N. 1995. Adjusting Moran's I for population density. *Statistics in Medicine,* 14**:** 17-26. DOI: https://doi.org/10.1002/sim.4780140104

Peckham, H., N. M. de Gruijter, C. Raine, A. Radziszewska, C. Ciurtin, L. R. Wedderburn, E. C. Rosser, K. Webb, et al. 2020. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nature Communications,* 11**:** 6317. DOI: 10.1038/s41467-020-19741-6

Petermann, E., H. Meyer, M. Nussbaum, and P. Bossew. 2021. Mapping the geogenic radon potential for Germany by machine learning. *Science of The Total Environment,* 754**:** 142291. DOI: https://doi.org/10.1016/j.scitotenv.2020.142291

Qiu, Y., X. Chen, and W. Shi. 2020. Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. *Journal of Population Economics,* 33**:** 1127-1172. DOI: 10.1007/s00148-020-00778-2

Quesada, J. A., A. López-Pineda, V. F. Gil-Guillén, J. M. Arriero-Marín, F. Gutiérrez, and C. Carratala-Munuera. 2021. Incubation period of COVID-19: A systematic review and meta-analysis. *Revista Clínica Española (English Edition),* 221**:** 109-117. DOI: https://doi.org/10.1016/j.rceng.2020.08.002

Rajgor, D. D., M. H. Lee, S. Archuleta, N. Bagdasarian, and S. C. Quek. 2020. The many estimates of the COVID-19 case fatality rate. *The Lancet Infectious Diseases,* 20**:** 776-777. DOI: 10.1016/S1473-3099(20)30244-9

Rey, J. F., S. Goyette, M. Gandolla, M. Palacios, F. Barazza, and J. Goyette Pernot. 2022. Long-Term Impacts of Weather Conditions on Indoor Radon Concentration Measurements in Switzerland. *Atmosphere,* 13**:** 92.

Richardson, D. B., N. D. Volkow, M.-P. Kwan, R. M. Kaplan, M. F. Goodchild, and R. T. Croyle. 2013. Spatial Turn in Health Research. *Science,* 339**:** 1390-1392. DOI: doi:10.1126/science.1232257

Richter, C., M. Wächter, J. Reinecke, A. Salheiser, M. Quent, and M. Wjst. 2021. Politische Raumkultur als Verstärker der Corona-Pandemie? Einflussfaktoren auf die regionale Inzidenzentwicklung in Deutschland in der ersten und zweiten Pandemiewelle 2020. *ZRex–Zeitschrift für Rechtsextremismusforschung,* 1.

Rios, C., D. Cassatt, B. Hollingsworth, M. Satyamitra, Y. Tadesse, L. Taliaferro, T. Winters, and A. DiCarlo. 2020. Commonalities Between COVID-19 and Radiation Injury. *Radiation Research,* 195. DOI: 10.1667/RADE-20-00188.1

Robert-Koch-Institut. 2022. 2022-02-14_Deutschland_Landkreise_COVID-19-Impfungen.csv.

Robert Koch-Institut, 2020. Situationsbericht des Robert Koch-Instituts vom 27.3.2020 zu COVID-19. Robert Roch-Institut Report. [in Swedish, English summary]

Robert Koch-Institut. 2022. Digitales Impfquotenmonitoring zur COVID-19-Impfung. Retrieved, from. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Impfquoten-Tab.html

Robert Koch-Institut & DIVI e.V. 2020. COVID-ITS-Fälle und ITS-Kapazitäten. www.intensivregister.de.

Robert Koch Institut. 2021. Epidemiologischer Steckbrief zu SARS-CoV-2 und COVID-19. Retrieved, from. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html

Rocklöv, J., and H. Sjödin. 2020. High population densities catalyse the spread of COVID-19. *Journal of Travel Medicine,* 27. DOI: 10.1093/jtm/taaa038

Sesana, L., E. Caprioli, and G. M. Marcazzan. 2003. Long period study of outdoor radon concentration in Milan and correlation between its temporal variations and dispersion properties of atmosphere. *Journal of Environmental Radioactivity,* 65**:** 147-160. DOI: https://doi.org/10.1016/S0265-931X(02)00093-0

SeyedAlinaghi, S., P. Mirzapour, O. Dadras, Z. Pashaei, A. Karimi, M. MohsseniPour, M. Soleymanzadeh, A. Barzegary, et al. 2021. Characterization of SARS-CoV-2 different variants and related

morbidity and mortality: a systematic review. *European Journal of Medical Research,* 26**:** 51. DOI: 10.1186/s40001-021-00524-8

Shaweno, D., M. Karmakar, K. A. Alene, R. Ragonnet, A. C. A. Clements, J. M. Trauer, J. T. Denholm, and E. S. McBryde. 2018. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. *BMC Medicine,* 16**:** 193. DOI: 10.1186/s12916-018-1178-4

Slikkerveer, A., and F. A. de Wolff. 1989. Pharmacokinetics and Toxicity of Bismuth Compounds. *Medical Toxicology and Adverse Drug Experience,* 4**:** 303-323. DOI: 10.1007/BF03259915

Snow, J. 1855. *On the mode of communication of cholera*. John Churchill.

Sorci, G., B. Faivre, and S. Morand. 2020. Why does COVID-19 case fatality rate vary among countries?

Statistische Ämter des Bundes und der Länder. 2019. 12411-07-01-4: Durchschnittsalter der Bevölkerung - Stichtag 31.12. - regionale Tiefe: Kreise und krfr. Städte.

Statistische Ämter des Bundes und der Länder. 2022. Zensus 2011 - Durchschnittliche Wohnfläche je EW - License: dl-de/by-2-0 www.govdata.de/dl-de/by-2-0.

Statistisches Bundesamt. 2020. Interaktive Karte mit Altersgruppen auf Kreisebene 2020. Retrieved, from. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/karte-altersgruppen.html

Statistisches Bundesamt. 2021. Census postponed to 2022 - New census reference date: 15 May 2022.

Statistisches Landesamt Baden-Würtenberg. 2020. Einkommen der privaten Haushalte in den kreisfreien Städten und Landkreisen der Bundesrepublik Deutschland 1995 bis 2019 Reihe 2, Kreisergebnisse Band 3. ed. S. L. Baden-Würtenberg.

Tchorz-Trzeciakiewicz, D. E., and M. Kłos. 2017. Factors affecting atmospheric radon concentration, human health. *Science of The Total Environment,* 584-585**:** 911-920. DOI: https://doi.org/10.1016/j.scitotenv.2017.01.137

Tobler, W. R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography,* 46**:** 234-240. DOI: 10.2307/143141

Turner, M. C., D. Krewski, Y. Chen, C. A. Pope, S. M. Gapstur, and M. J. Thun. 2012. Radon and COPD mortality in the American Cancer Society Cohort. *European Respiratory Journal,* 39**:** 1113. DOI: 10.1183/09031936.00058211

Universitätsmedizin, and der Johannes Gutenberg-Universität Mainz. 2021. Gutenberg COVID-19 Studie - Ergebnisse aus einer bevölkerungsrepresentativen Studie.

Verde, G., V. Artiola, M. Commara, V. D'Avino, L. Angrisani, G. Sabatino, and M. Pugliese. 2022. COVID-19 and the Additional Radiological Risk during the Lockdown Period in the Province of Naples City (South Italy). *Life,* 12**:** 246. DOI: 10.3390/life12020246

Wani, A. L., A. Ara, and J. A. Usmani. 2015. Lead toxicity: a review. *Interdisciplinary toxicology,* 8**:** 55-64. DOI: 10.1515/intox-2015-0009

Wong, D. W. S., and Y. Li. 2020. Spreading of COVID-19: Density matters. *PLOS ONE,* 15**:** e0242398. DOI: 10.1371/journal.pone.0242398

World Health Organization. 2020. Media Briefing on COVID-19, 11 March 2020.

World Health Organization. 2021. Radon and Health. Retrieved, from. https://www.who.int/news-room/fact-sheets/detail/radon-and-health

World Health Organization. 2022. Coronavirus Disease (COVID-19). Retrieved 2022-04-03, from. https://www.who.int/health-topics/coronavirus#tab=tab_3

Yan, R., Y. Zhang, Y. Li, L. Xia, Y. Guo, and Q. Zhou. 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science,* 367**:** 1444-1448. DOI: doi:10.1126/science.abb2762

Yong, S. J. 2021. Long COVID or post-COVID-19 syndrome: putative pathophysiology, risk factors, and treatments. *Infectious Diseases,* 53**:** 737-754. DOI: 10.1080/23744235.2021.1924397

8 Appendices

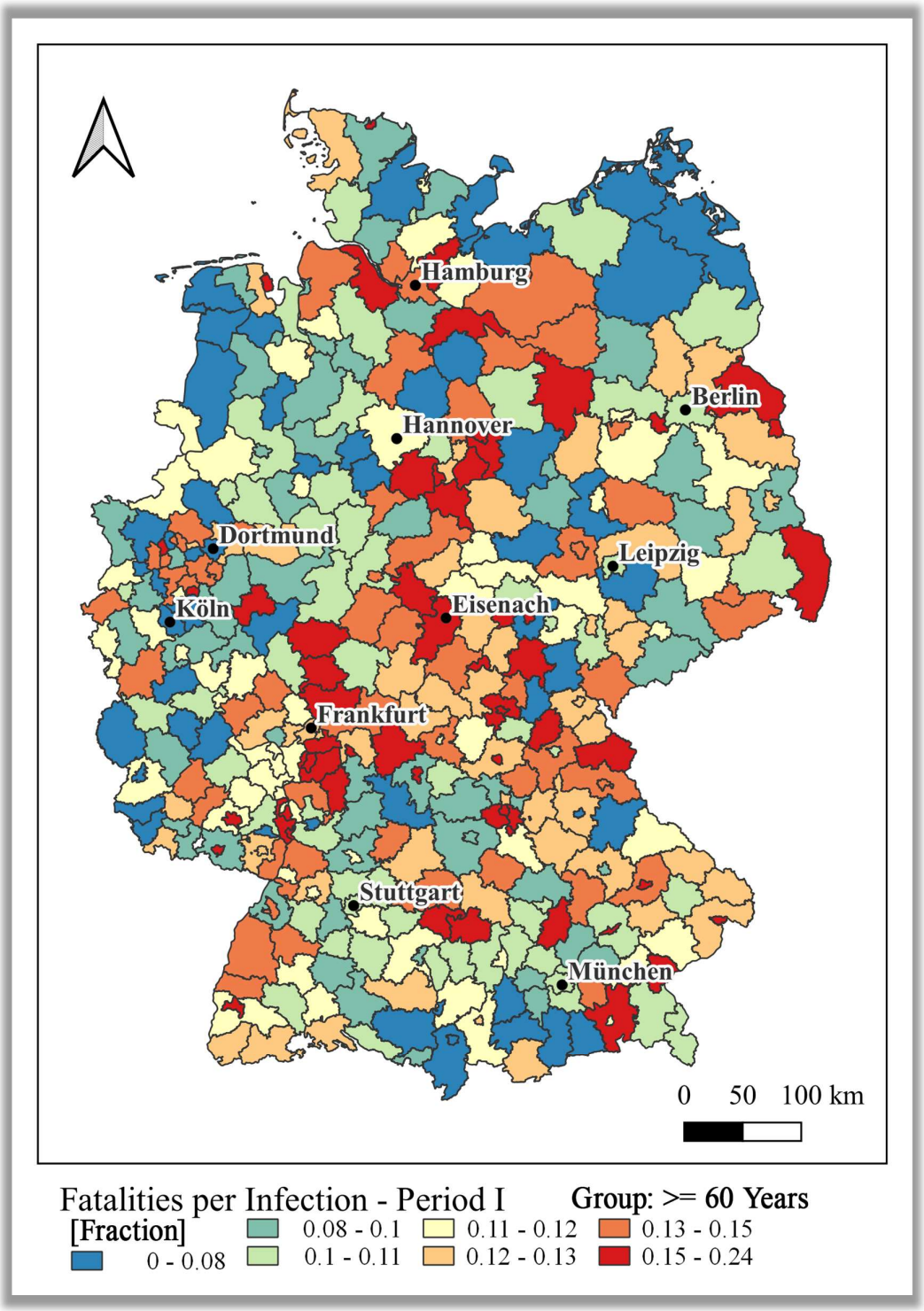


Figure 13 - Case Fatality Ratio of Age Group greater than or equal to 60 Years in Period I.

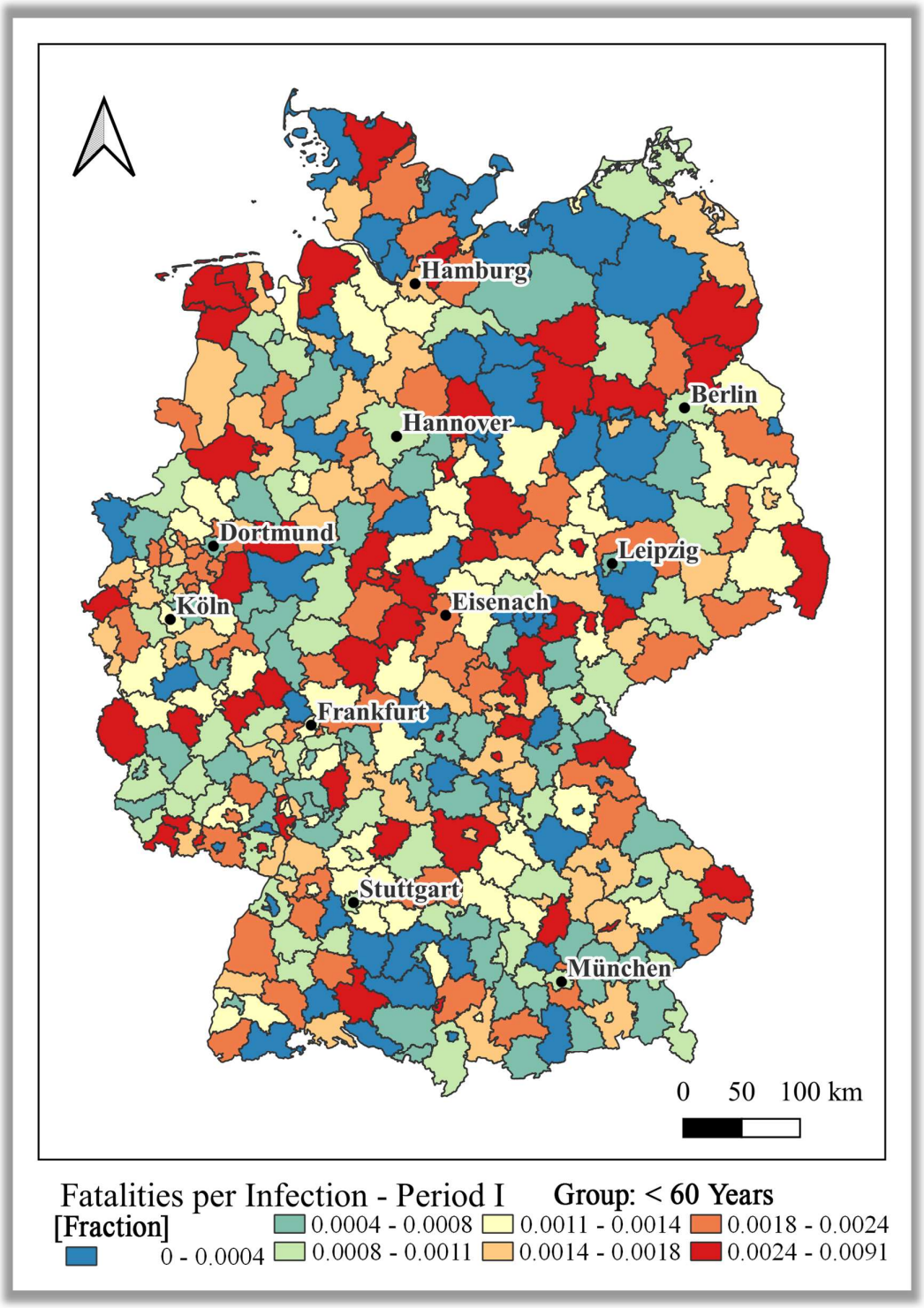


Figure 14 - Case Fatality Ratio of Age Group below 60 Years in Period I.

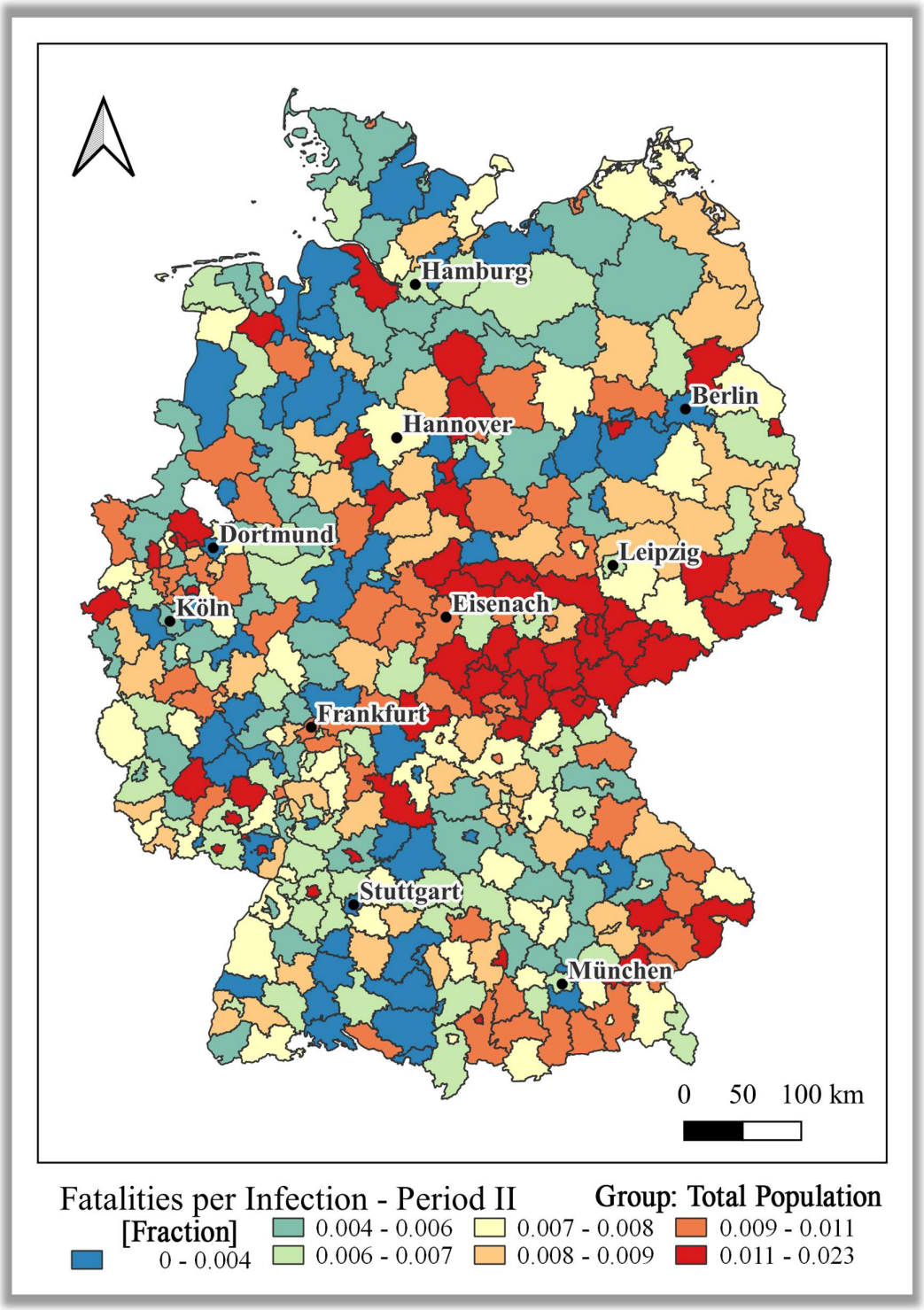


Figure 15 - Case Fatality Ratio of Total Population in Period II.

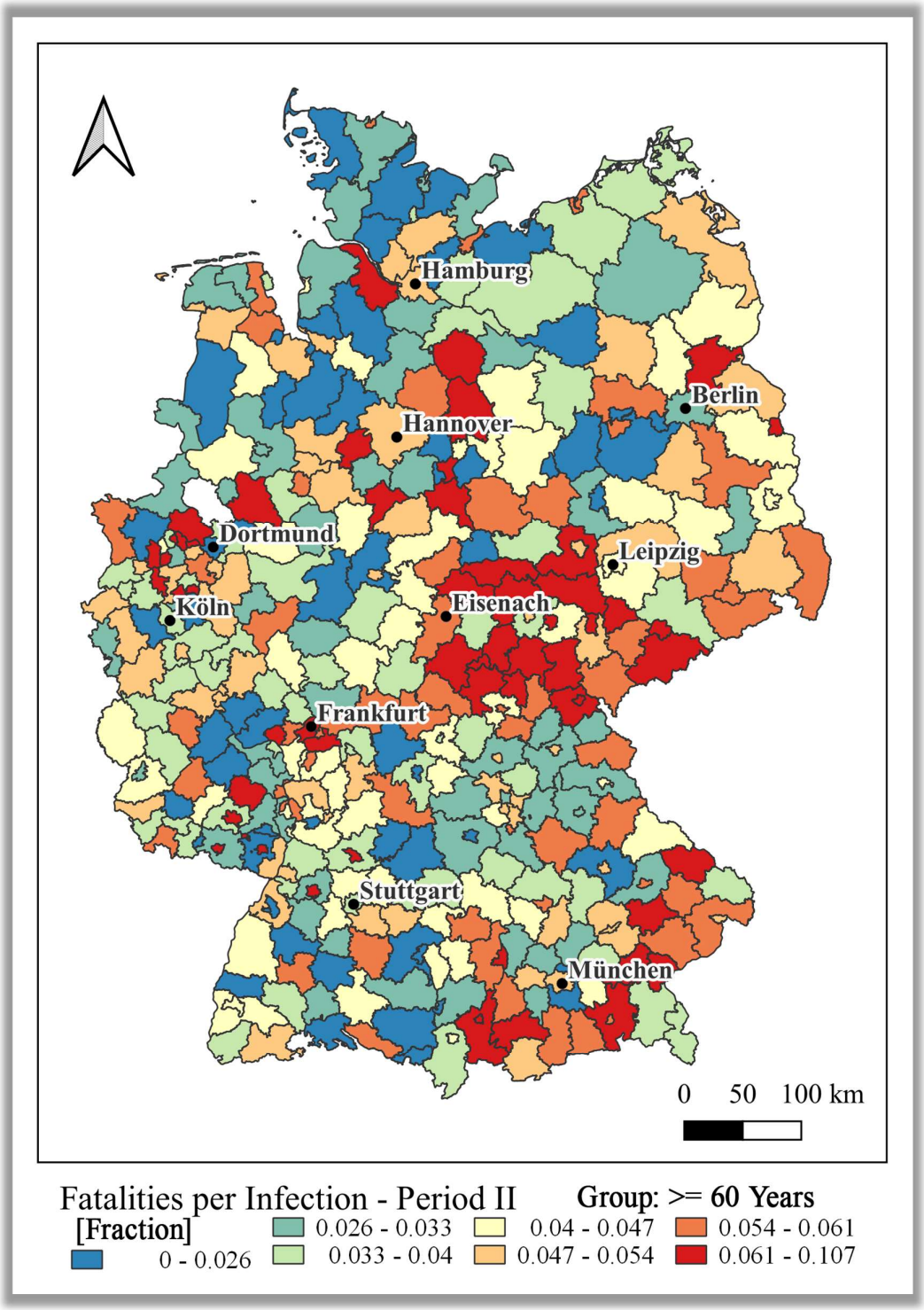


Figure 16 - Case Fatality Ratio of Age Group greater than or equal to 60 Years in Period II.

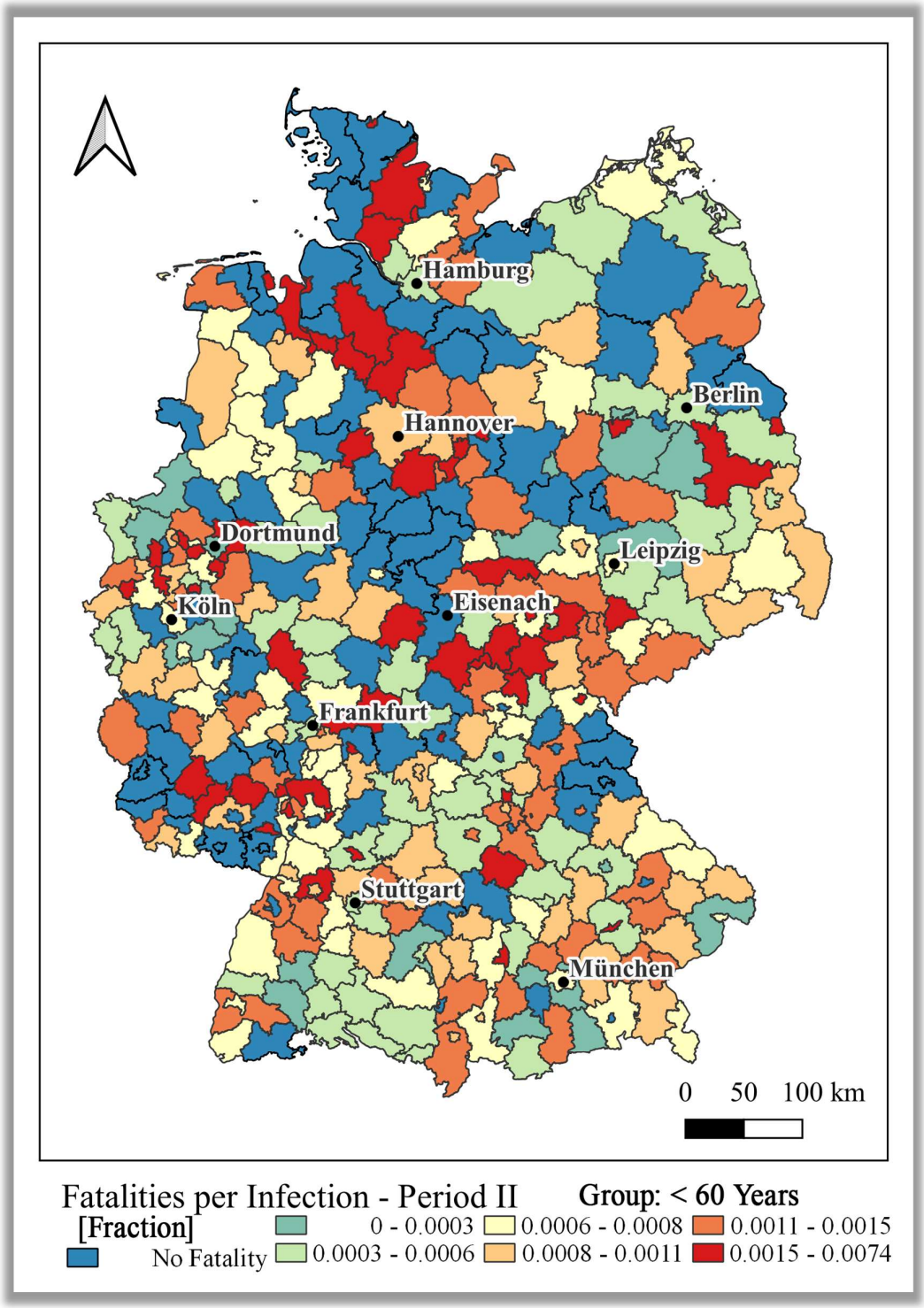


Figure 17 - Case Fatality Ratio of Age Group below 60 Years in Period II.

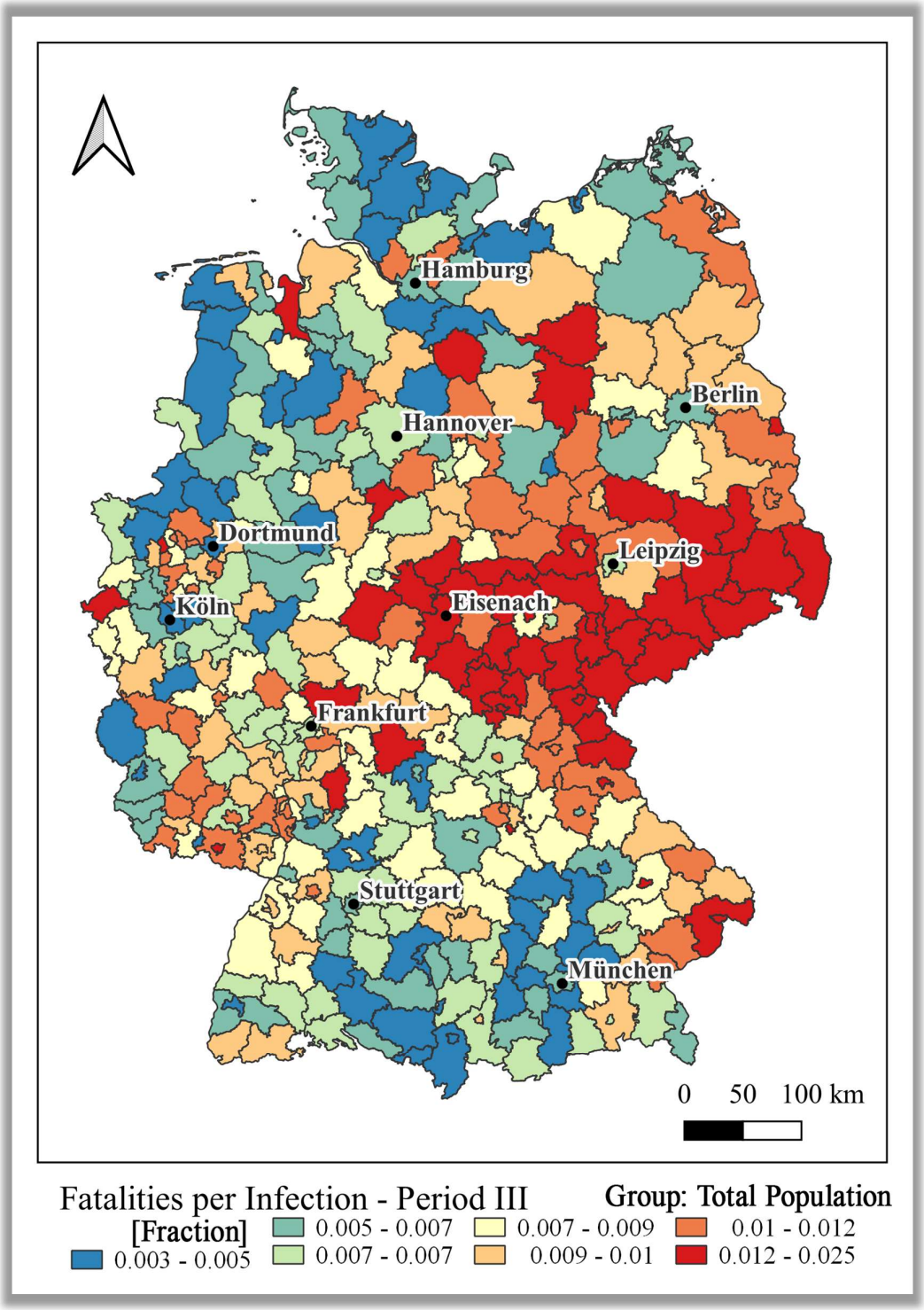


Figure 18 - Case Fatality Ratio of Total Population in Period III.

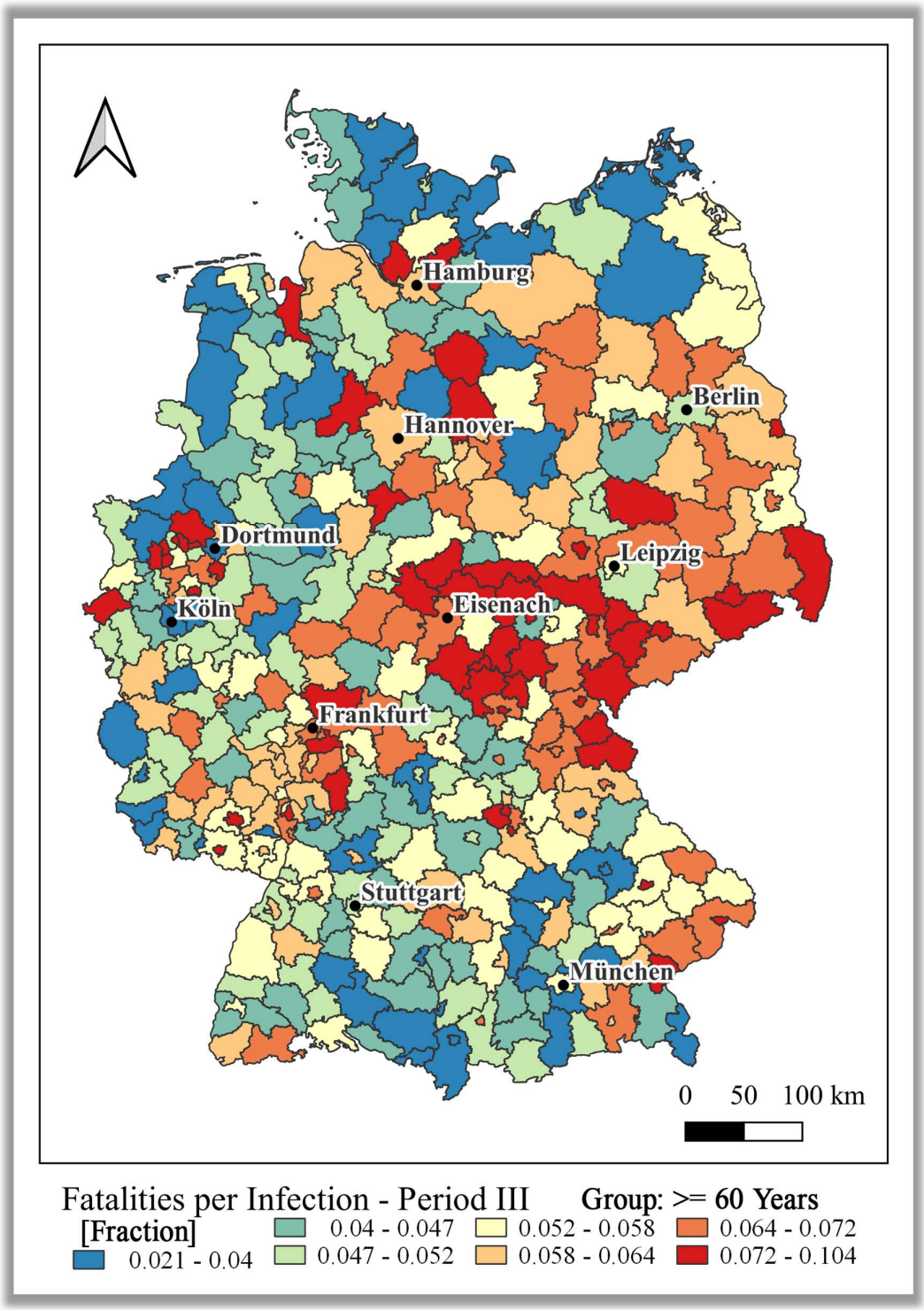


Figure 19 - Case Fatality Ratio of Age Group greater than or equal to 60 Years in Period III.

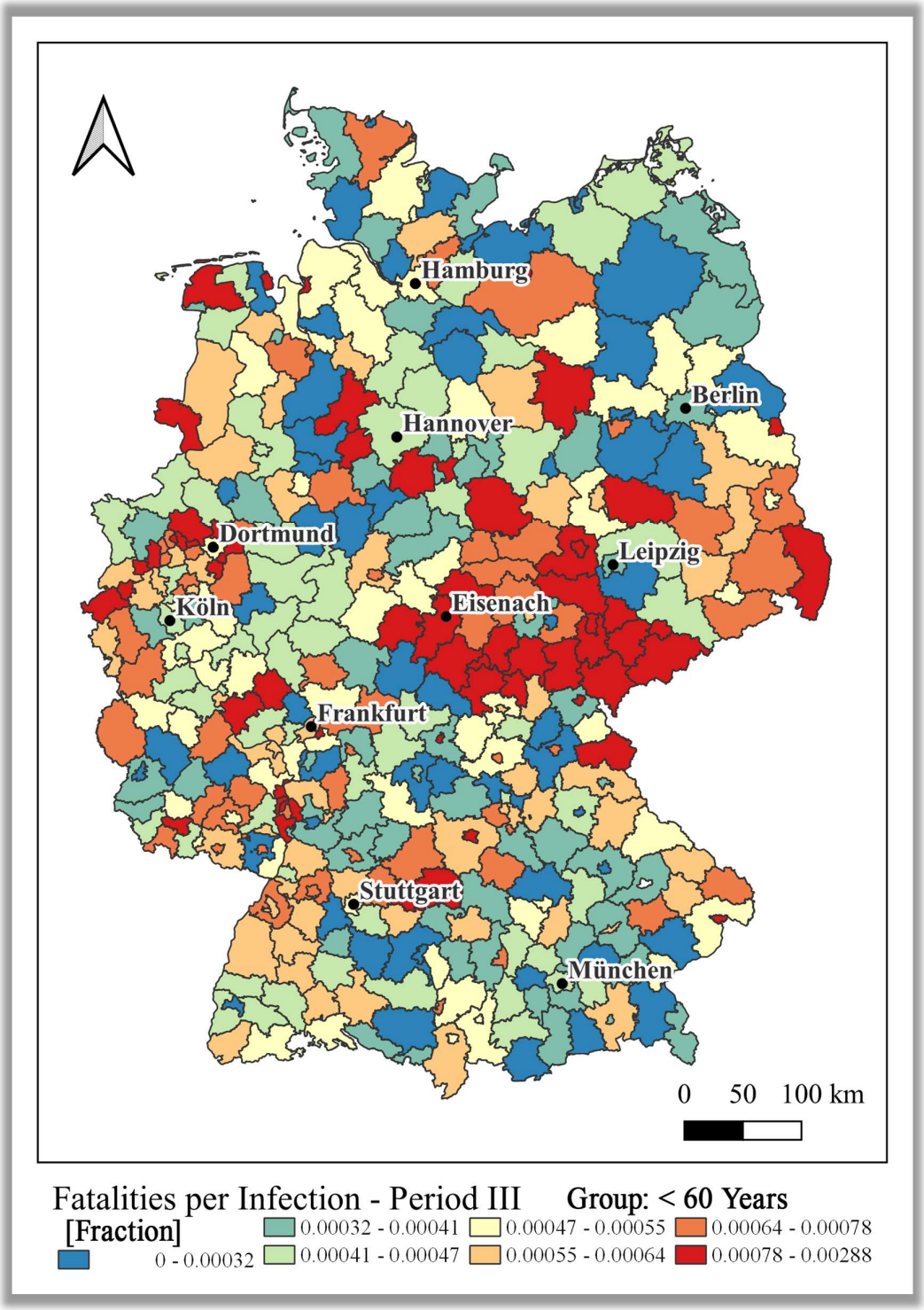


Figure 20 - Case Fatality Ratio of Age Group below 60 Years in Period II.

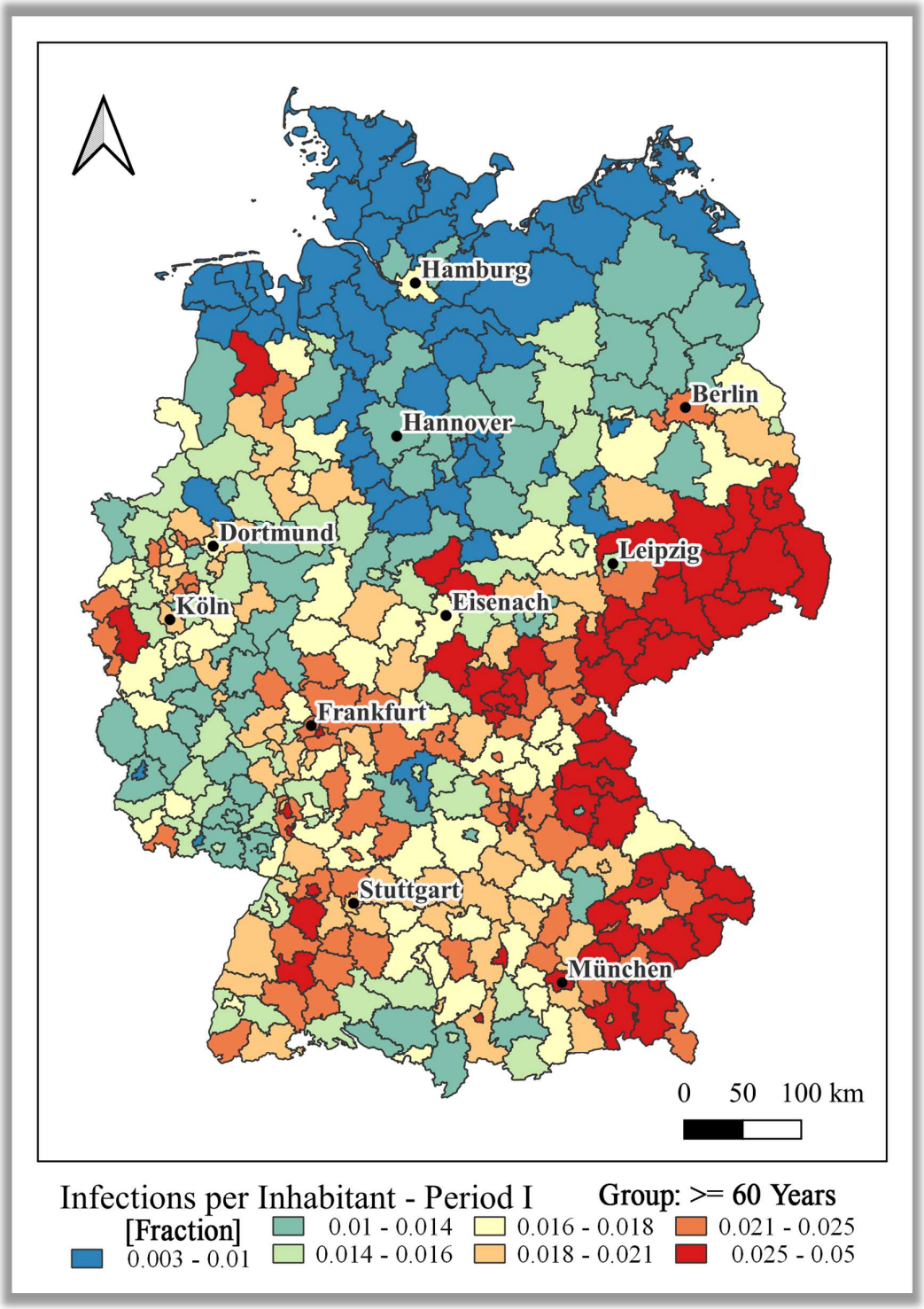


Figure 21 - Infection Rate of Age Group greater than or equal to 60 Years in Period I.

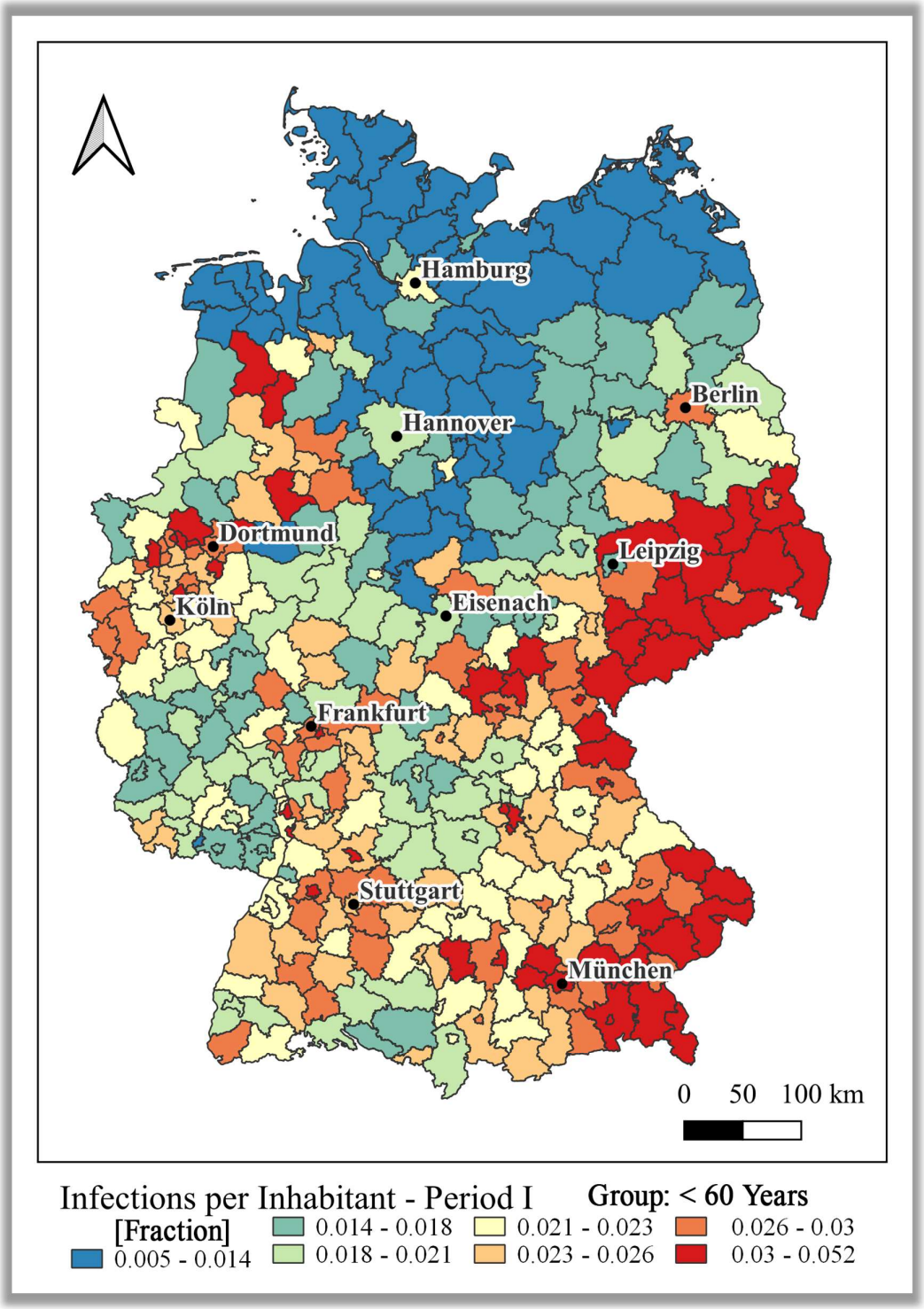


Figure 22 - Infection Rate of Age Group below 60 Years in Period I.

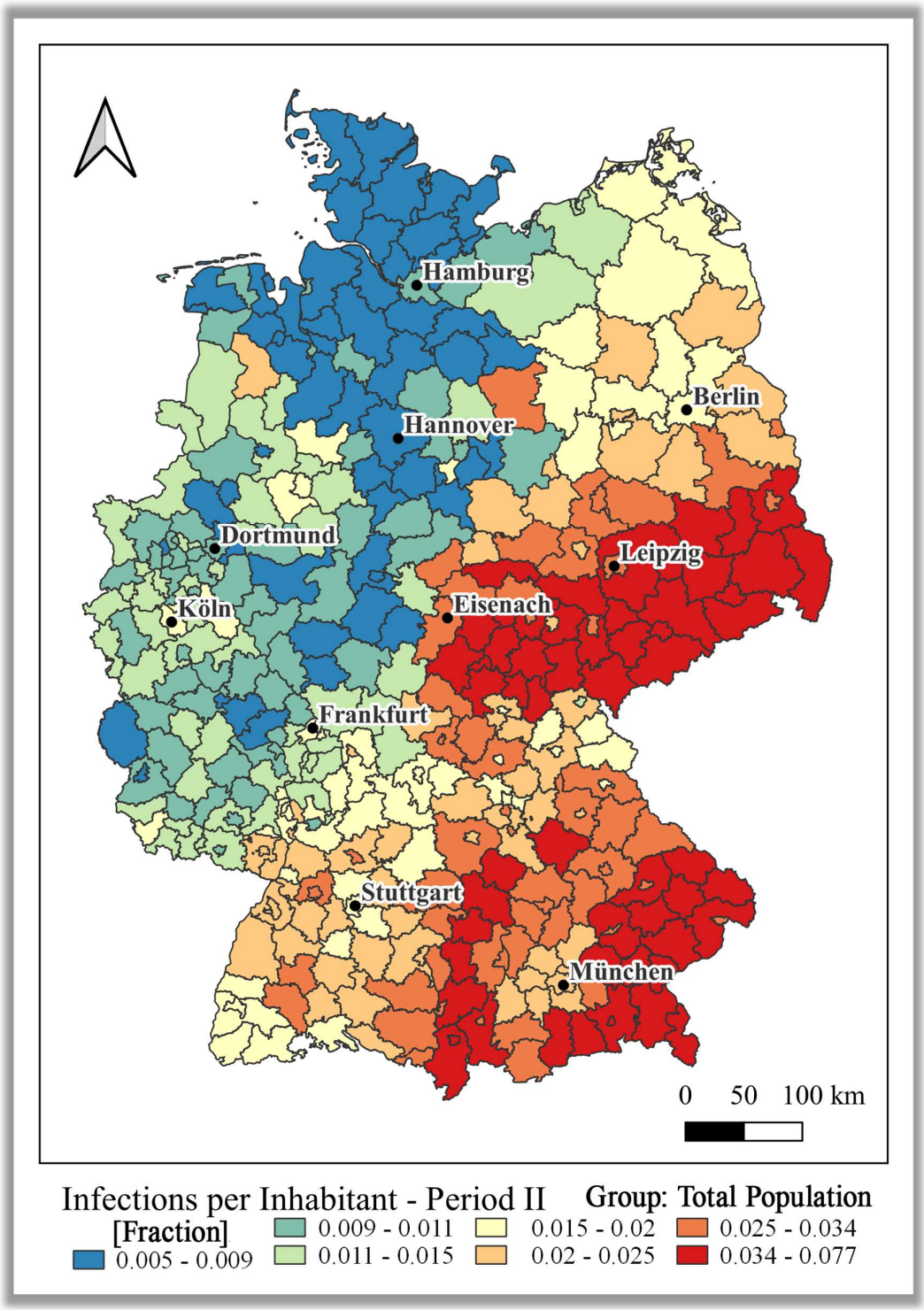


Figure 23 - Infection Rate of Total Population in Period II.

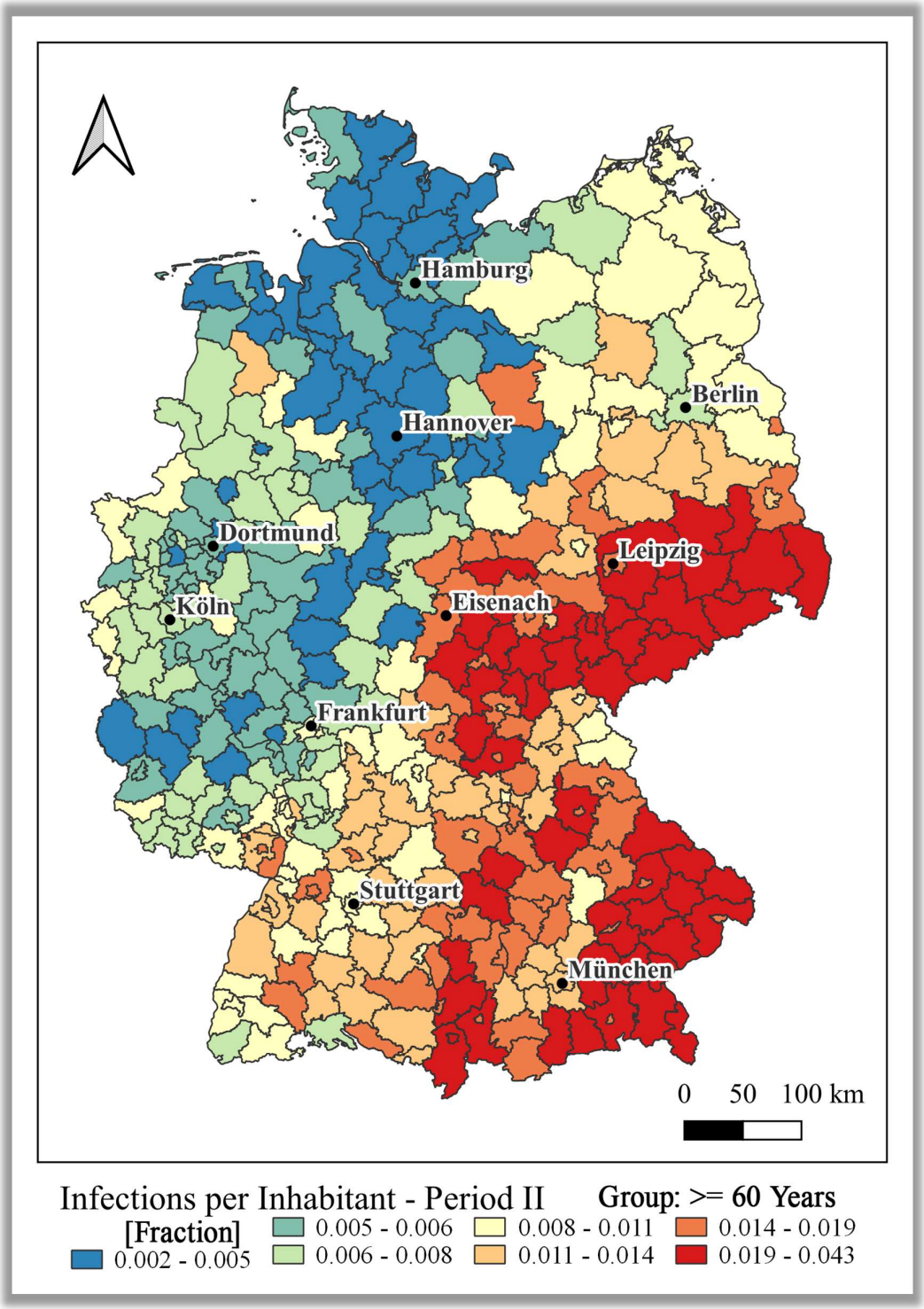


Figure 24 - Infection Rate of Age Group greater than or equal to 60 Years in Period II.

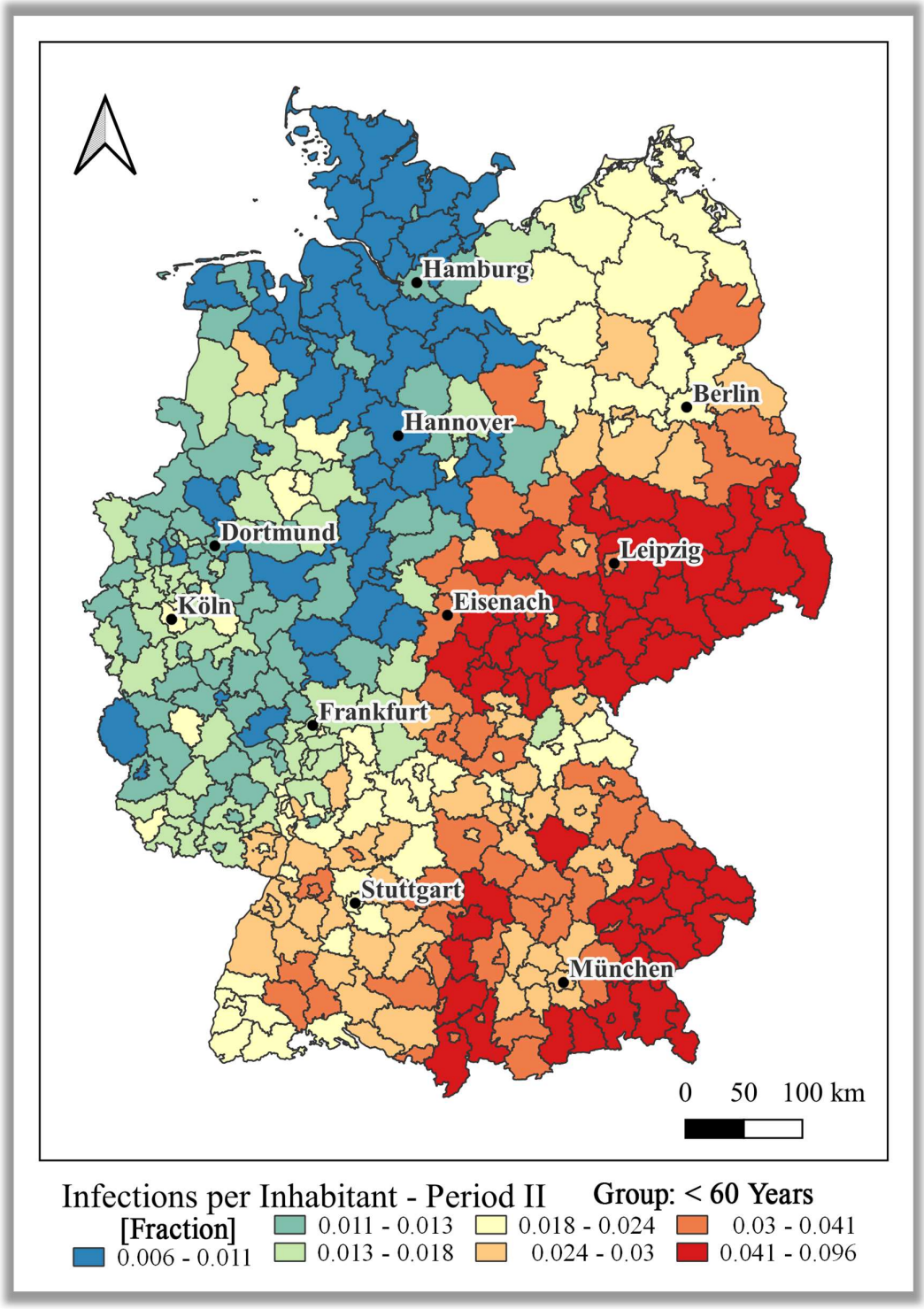


Figure 25 - Infection Rate of Age Group below 60 Years in Period II.

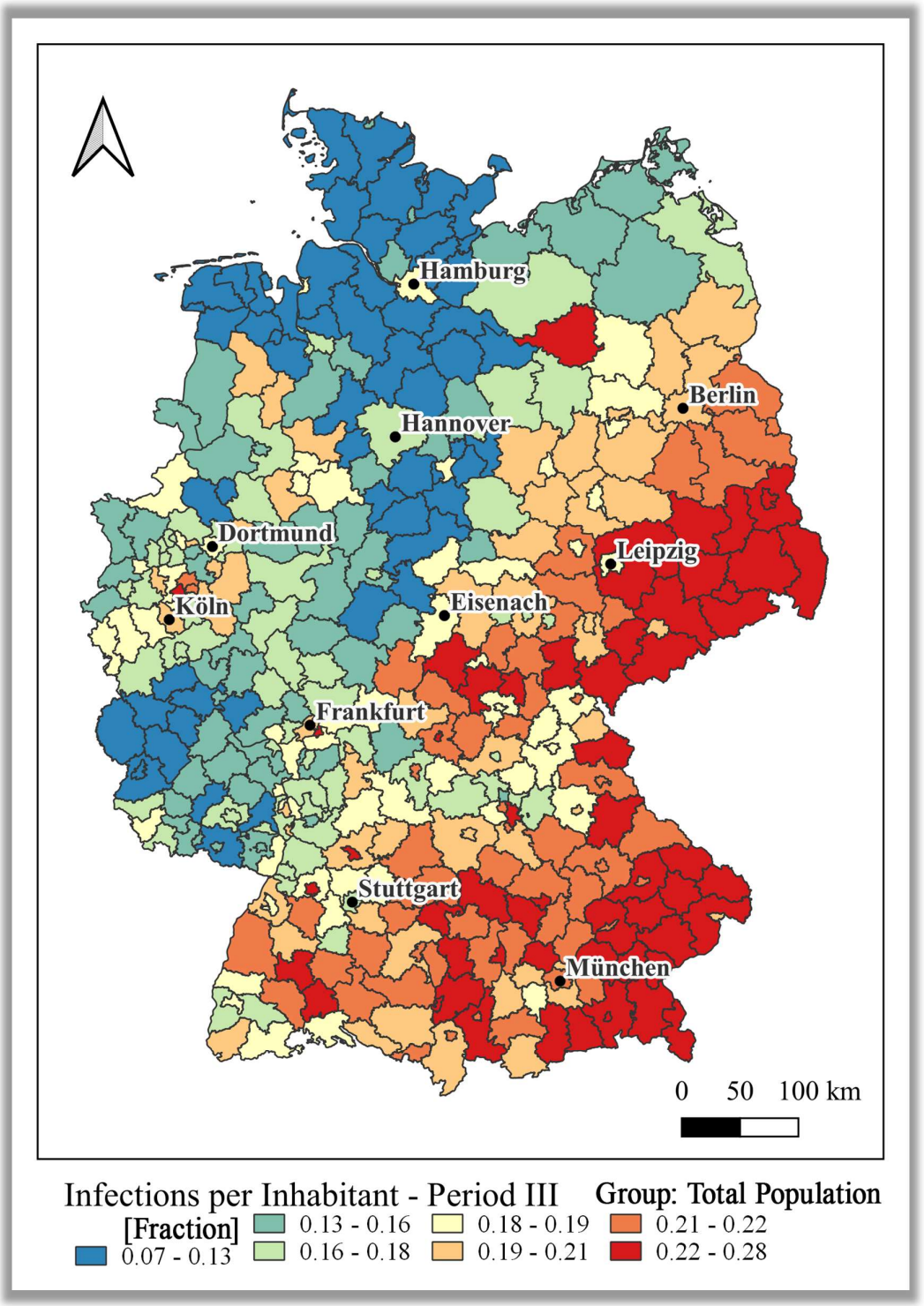


Figure 26 - Infection Rate of Total Population in Period III.

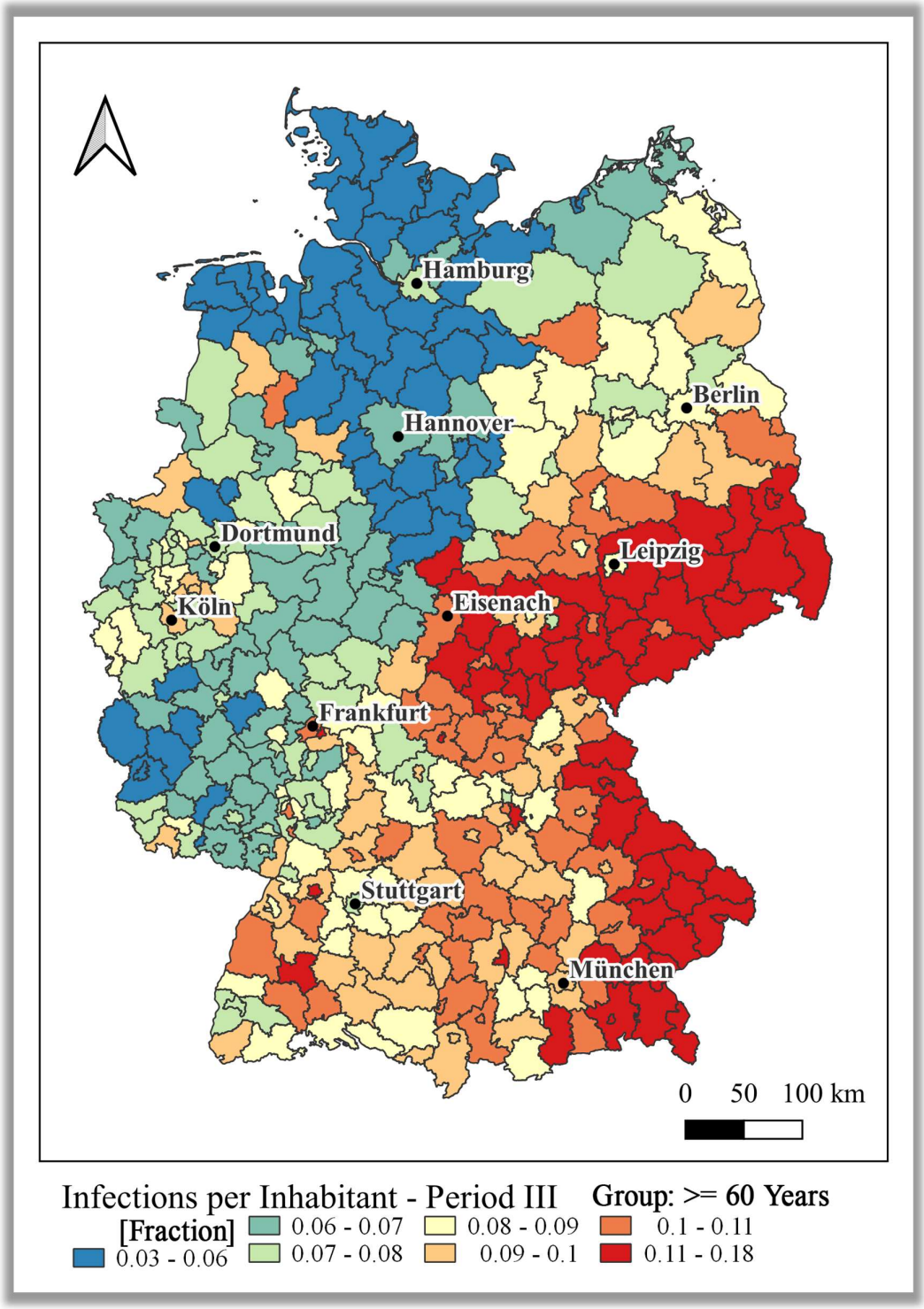


Figure 27 - Infection Rate of Age Group greater than or equal to 60 Years in Period III.

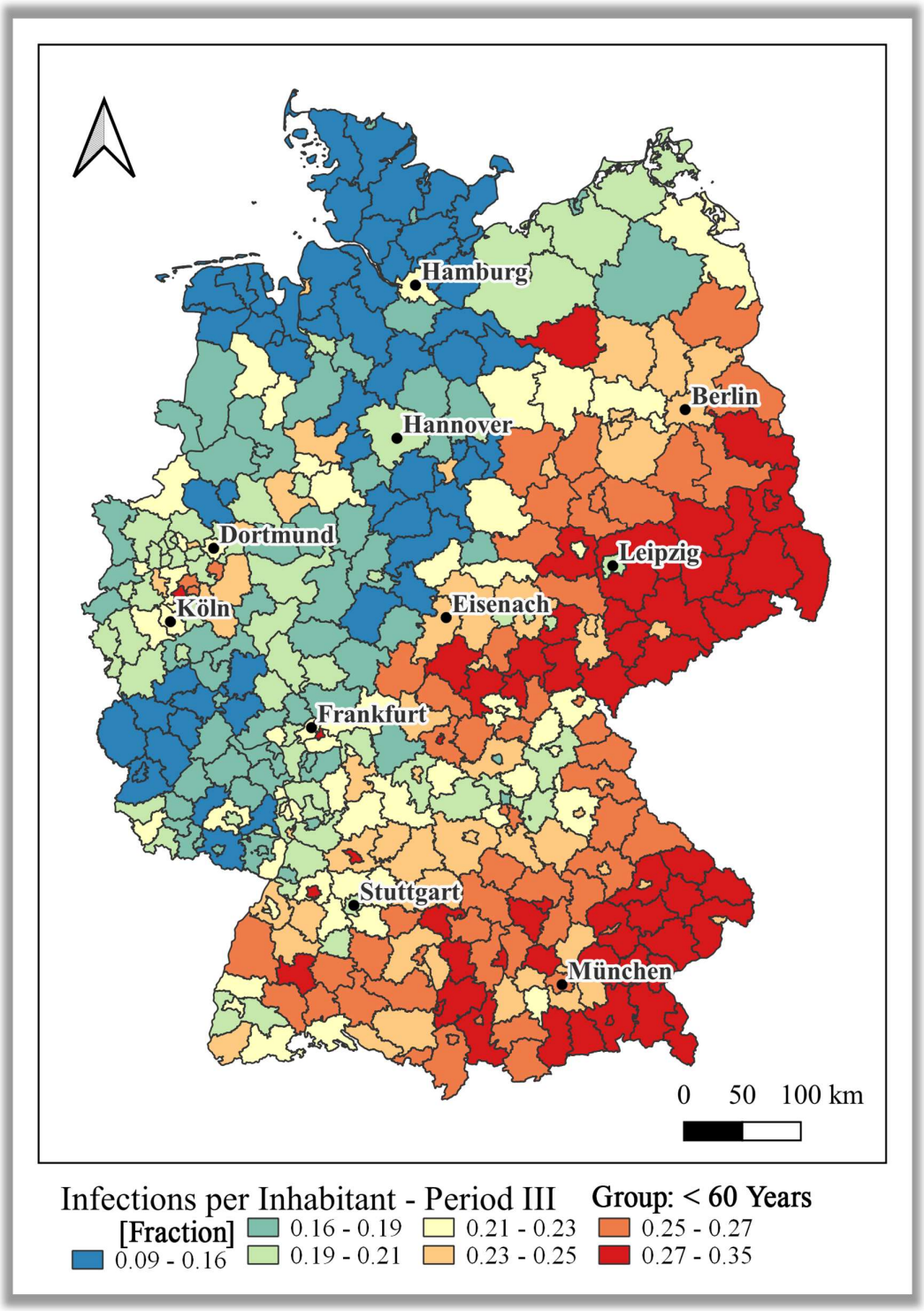


Figure 28 - Infection Rate of Age Group below 60 Years in Period III.

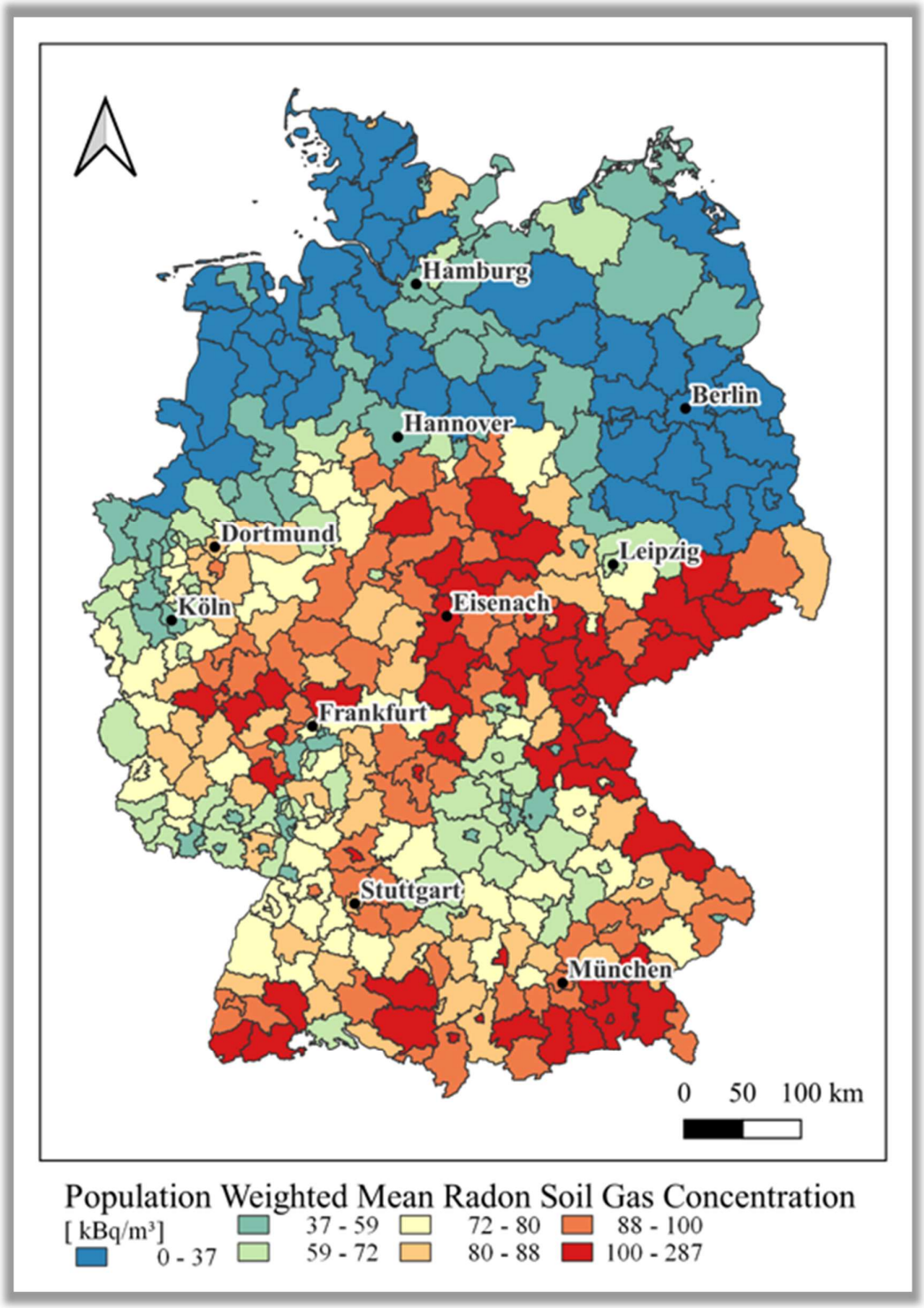


Figure 29 - Population Weighted Mean Radon Soil Gas Concentration.

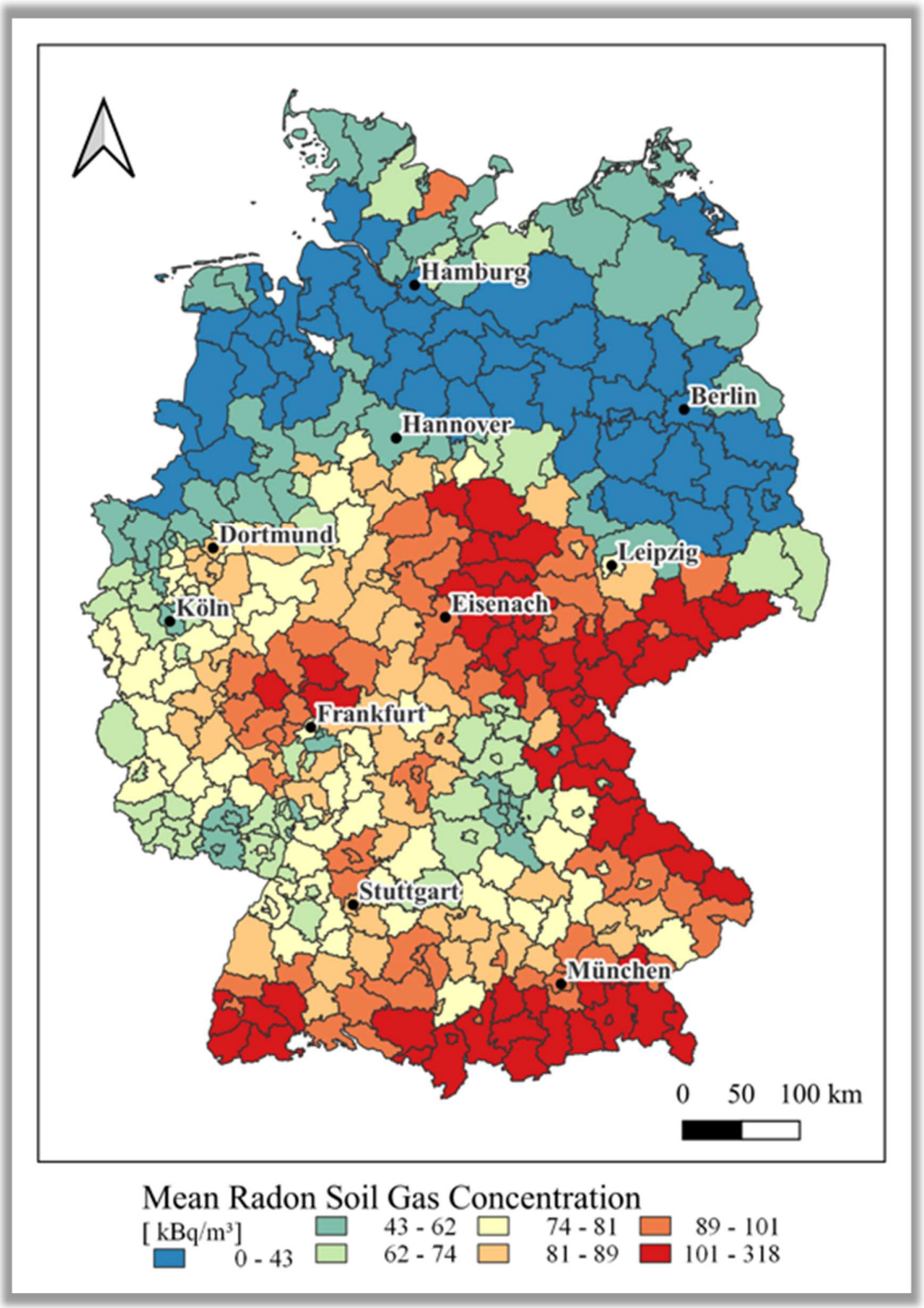


Figure 30 - Mean Radon Soil Gas Concentration.

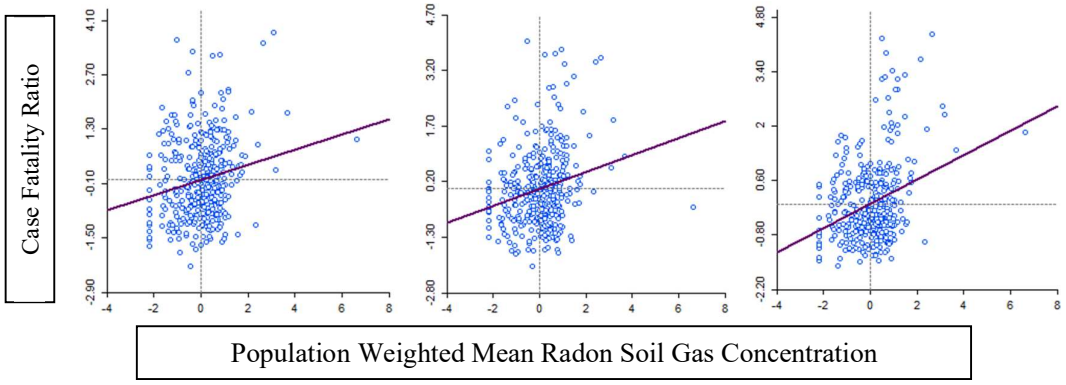
8.1 Scatterplots & Correlations

$p < 0.05$ significance marked with one asterisk
$p < 0.01$ significance marked with two asterisks

8.1.1 Population Weighted Mean Radon Soil Gas Concentration / Case Fatality Ratio

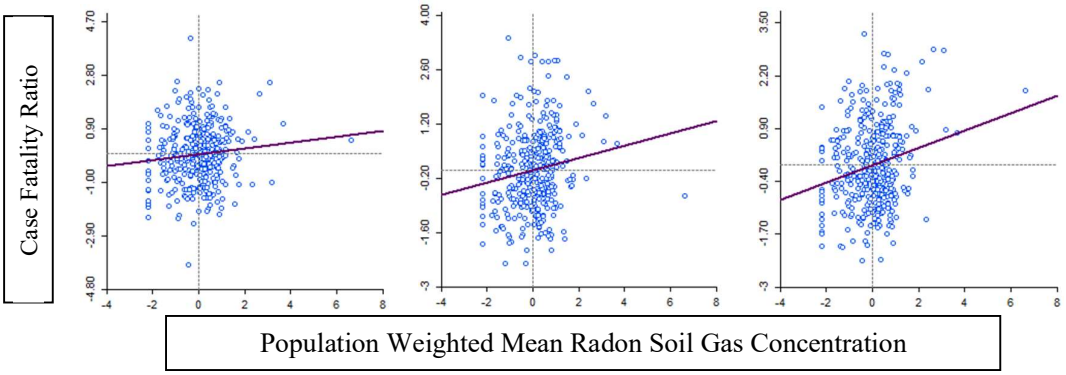
-Total Population:

Period I: 0.196** Period II: 0.229** Period III: 0.313**



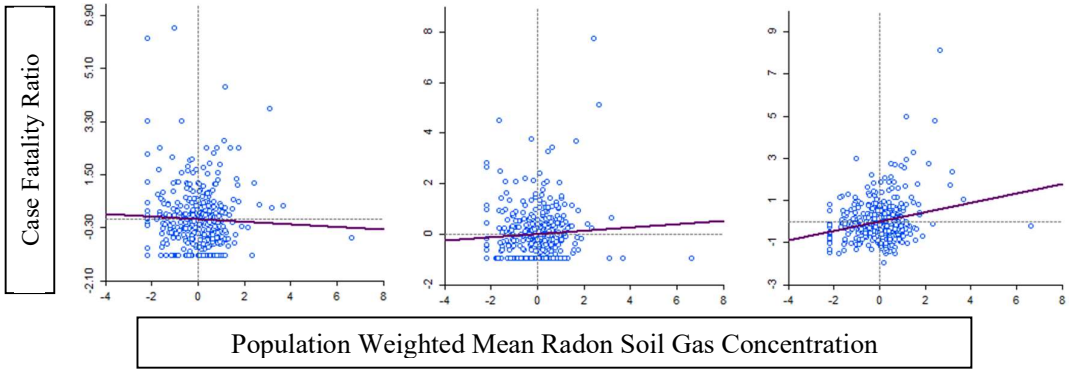
-Age Group greater than or equal to 60 Years:

Period I: 0.103 Period II: 0.160** Period III: 0.213**



-Age Group < 60 Years:

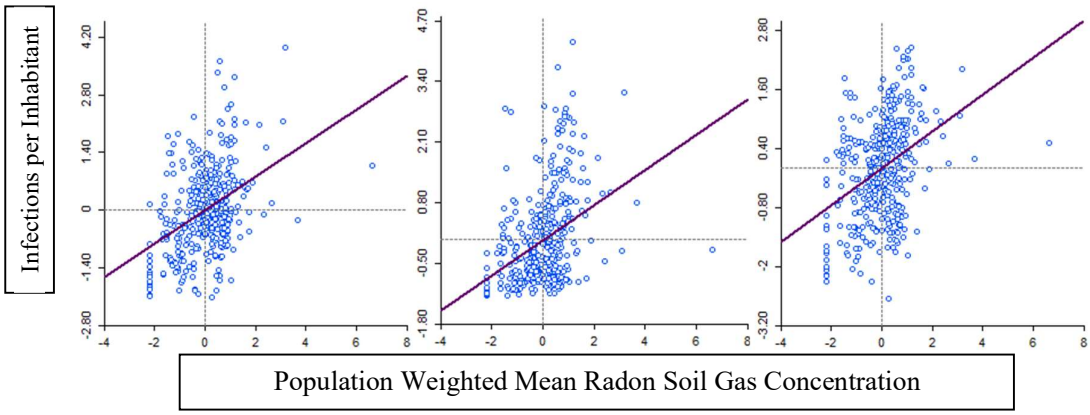
Period I: -0.044 Period II: 0.067 Period III: 0.220**



8.1.2 Population Weighted Mean Radon Soil Gas Concentration / Infections per Inhabitant

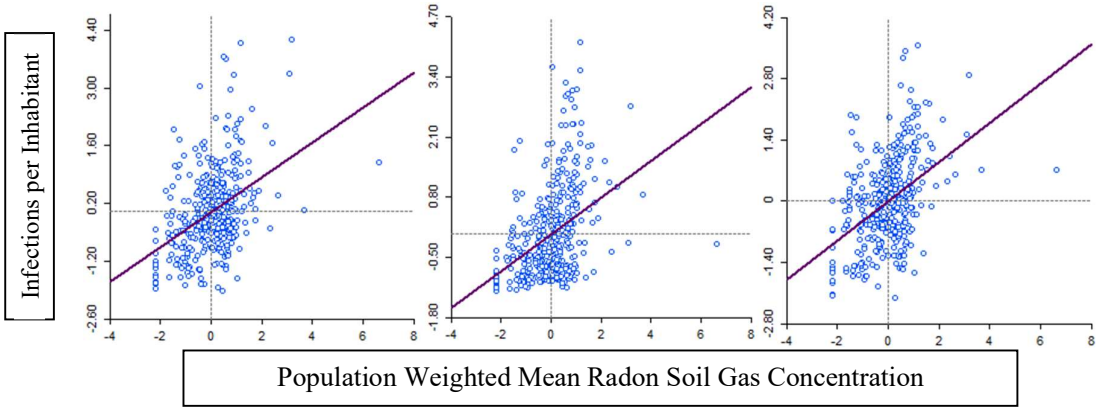
-Total Population:

Period I: 0.409** Period II: 0.377** Period III: 0.376**



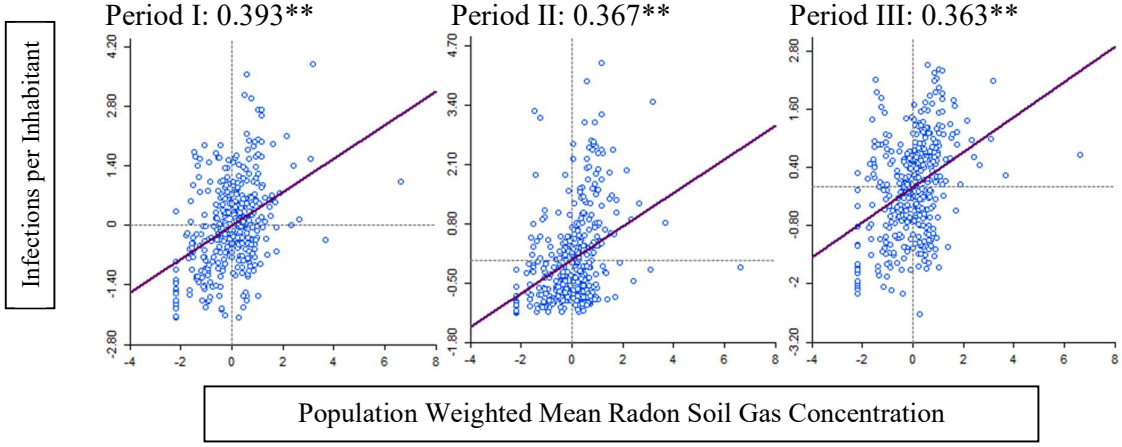
-Age Group greater than or equal to 60 Years:

Period I: 0.423** Period II: 0.397** Period III: 0.447**



-Age Group below 60 Years:

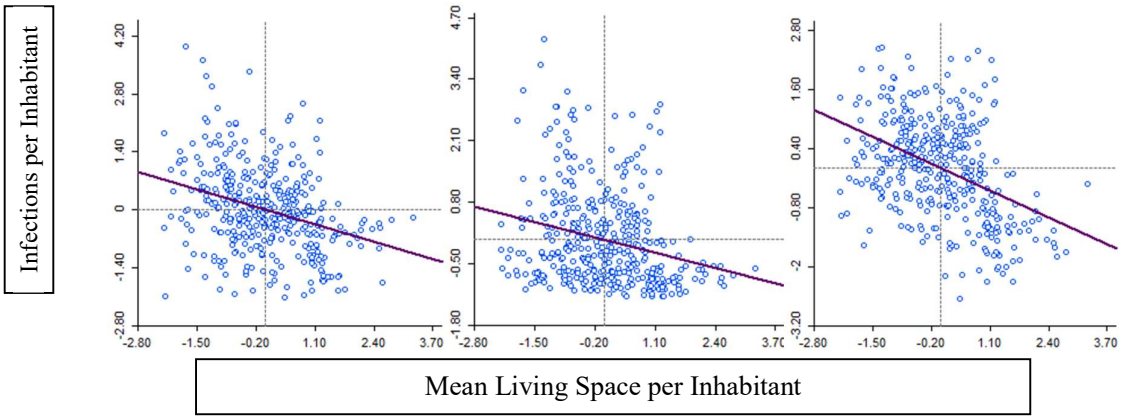
Period I: 0.393** Period II: 0.367** Period III: 0.363**



8.1.3 Living space per Inhabitant / Infections per Inhabitant

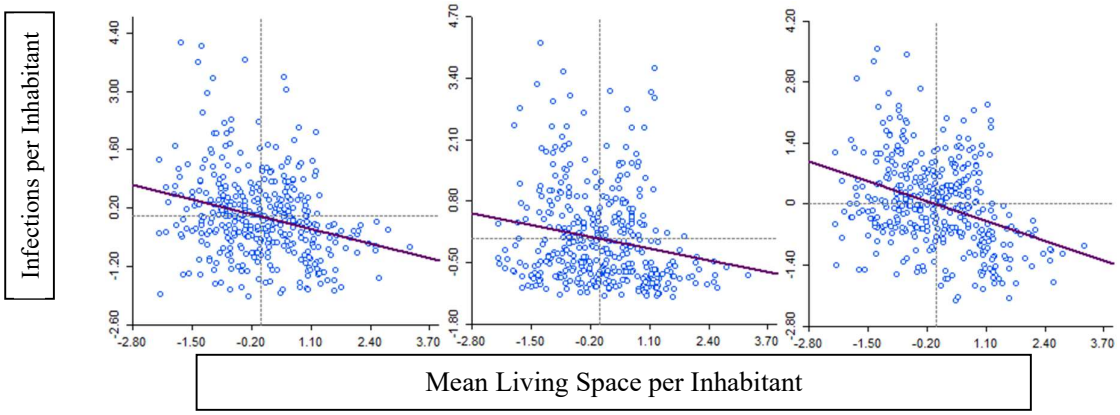
-Total Population:

Period I: -0.325** Period II: -0.251** Period III: -0.420**



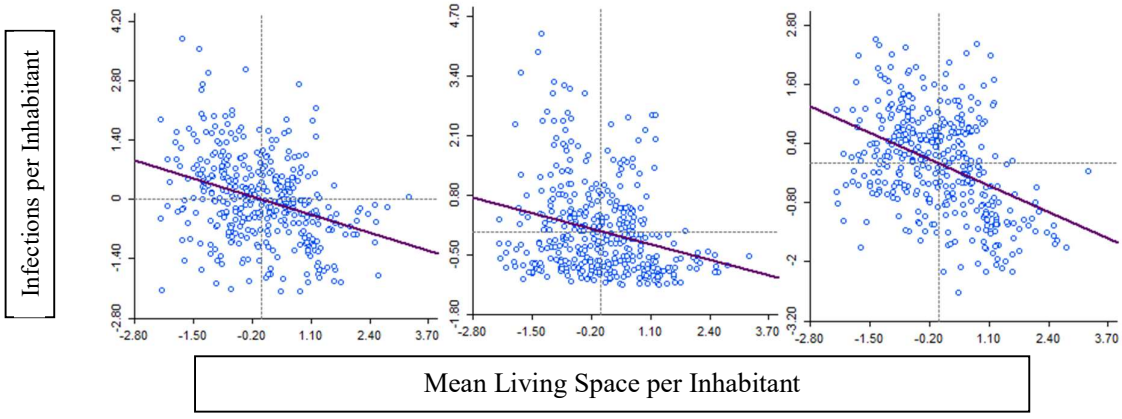
-Age Group greater than or equal to 60 Years:

Period I: -0.271** Period II: -0.193** Period III: -0.350**



-Age Group below 60 Years:

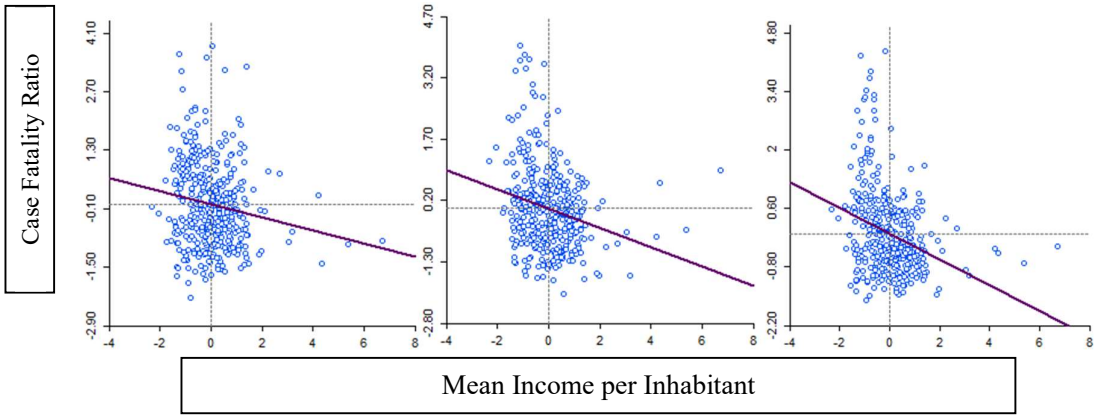
Period I: -0.331** Period II: -0.260** Period III: -0.411**



8.1.4 Mean Income / Case Fatality Ratio

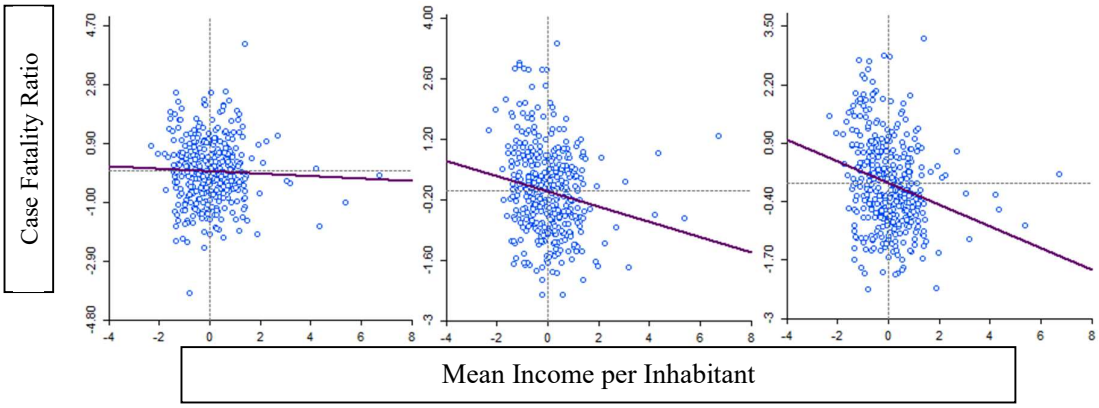
-Total Population:

Period I: -0.155** Period II : - 0.236** Period III: -0.307**



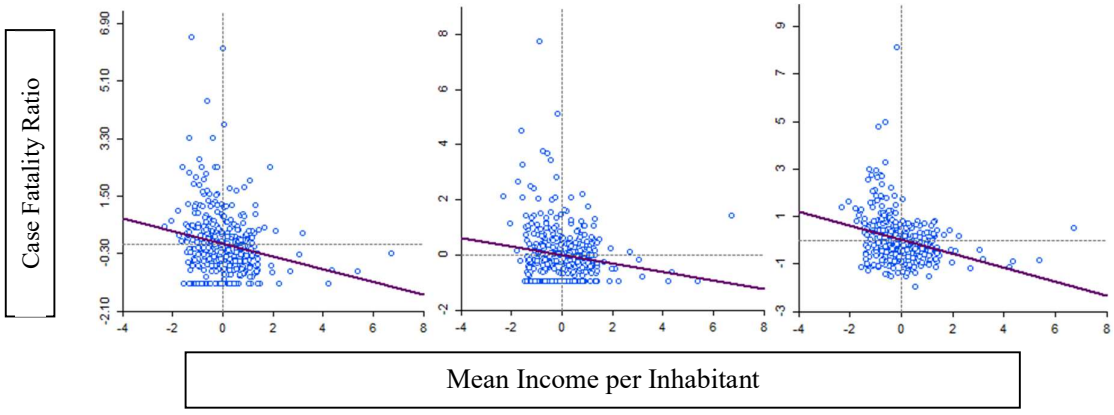
-Age Group greater than or equal to 60 Years:

Period I: -0.038 Period II : -0.176** Period III: -0.240**



-Age Group below 60 Years:

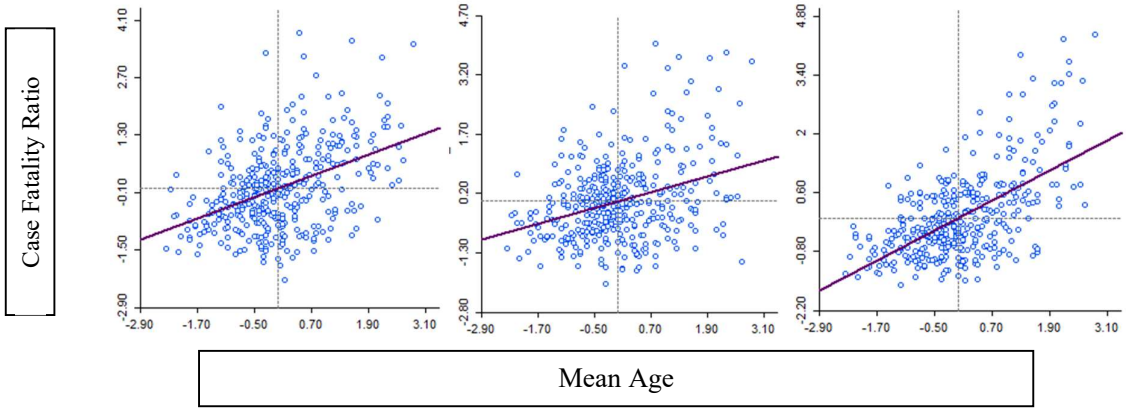
Period I: -0.200** Period II : -0.153** Period III: -0.293**



8.1.5 Mean Age / Case Fatality Ratio

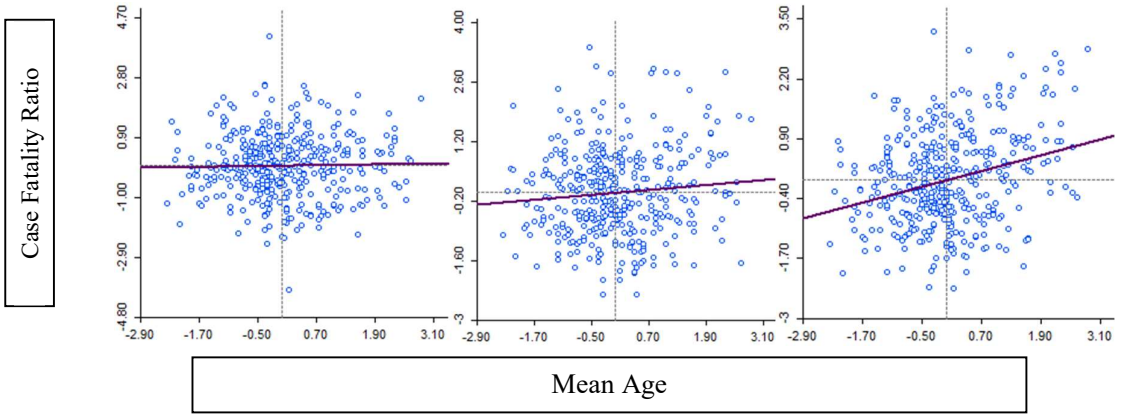
-Total Population:

Period I: 0.433** Period II: 0.330** Period III: 0.599**



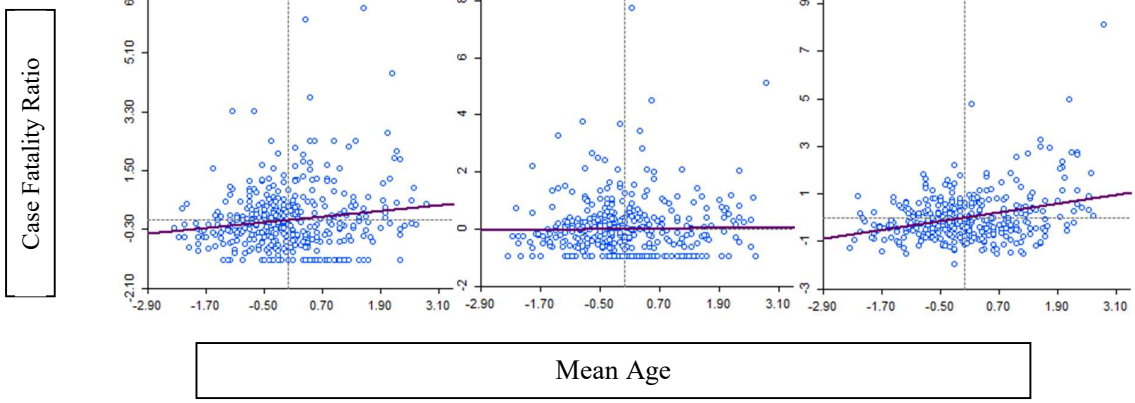
-Age Group greater than or equal to 60 Years:

Period I: 0.022 Period II: 0.098 Period III: 0.288**



-Age Group below 60 Years:

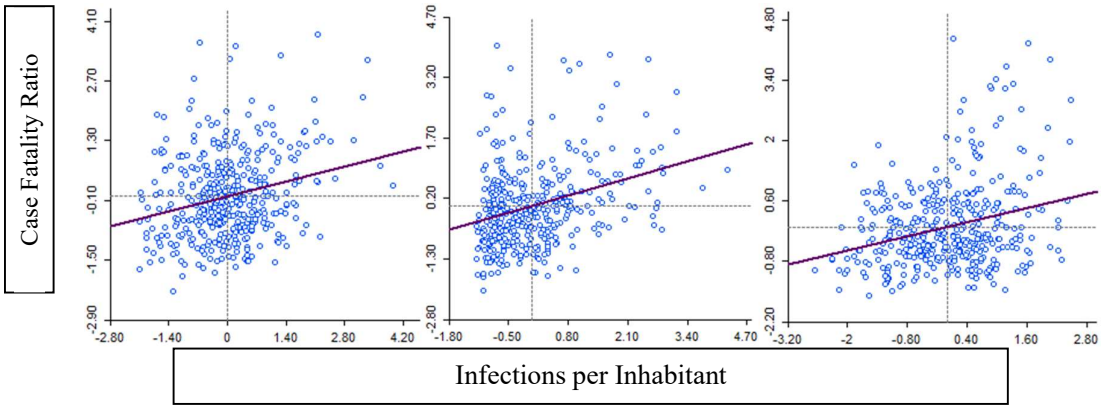
Period I: 0.144** Period II: 0.023 Period III: 0.303**



8.1.6 Infections per Inhabitant / Case Fatality Ratio

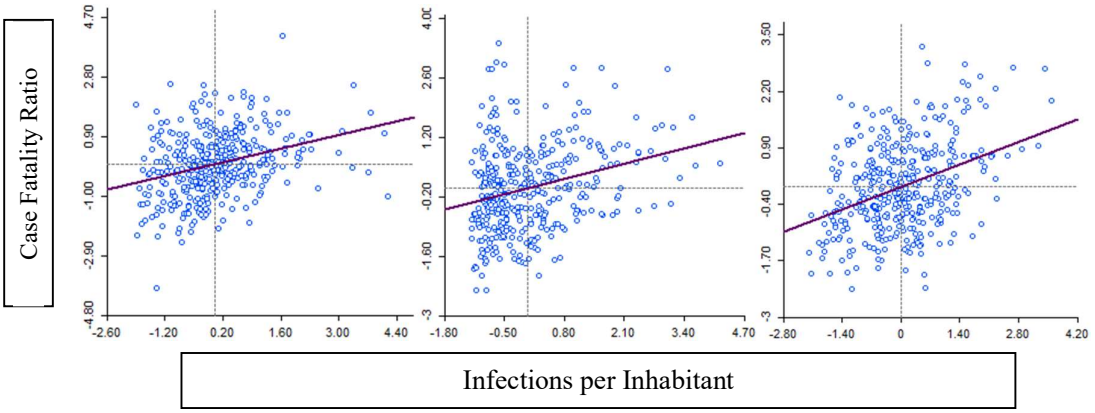
-Total Population:

Period I: 0.248** Period II: 0.327** Period III: 0.270**



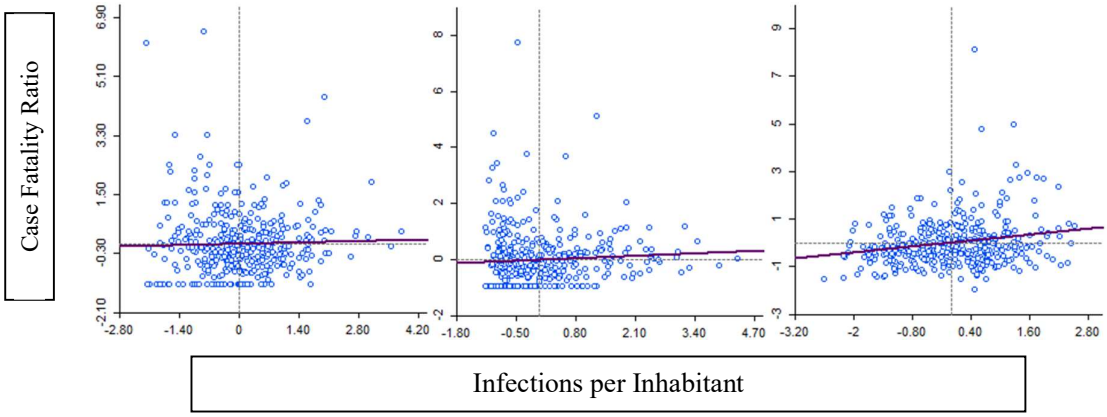
-Age Group greater than or equal to 60 Years:

Period I: 0.331** Period II: 0.280** Period III: 0.372**



-Age Group below 60 Years:

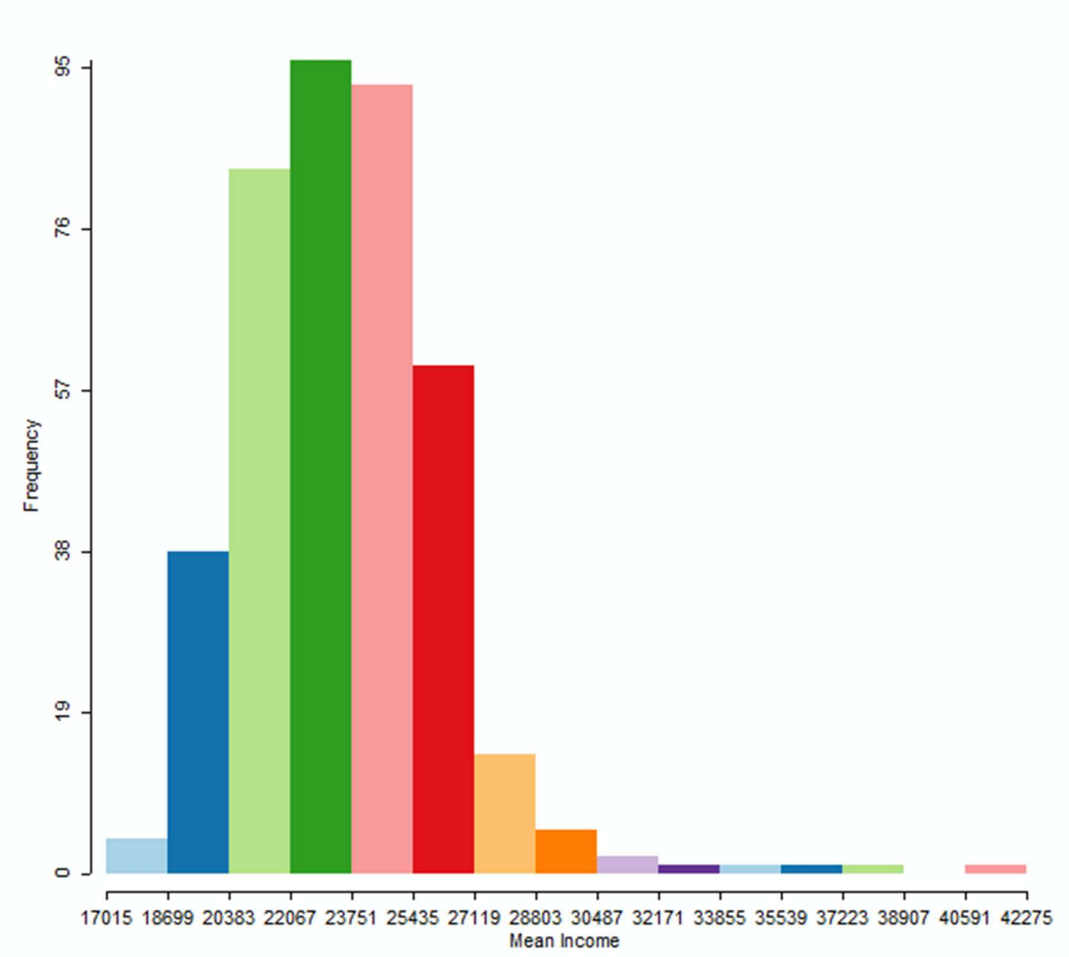
Period I: 0.028 Period II: 0.063 Period III: 0.209**



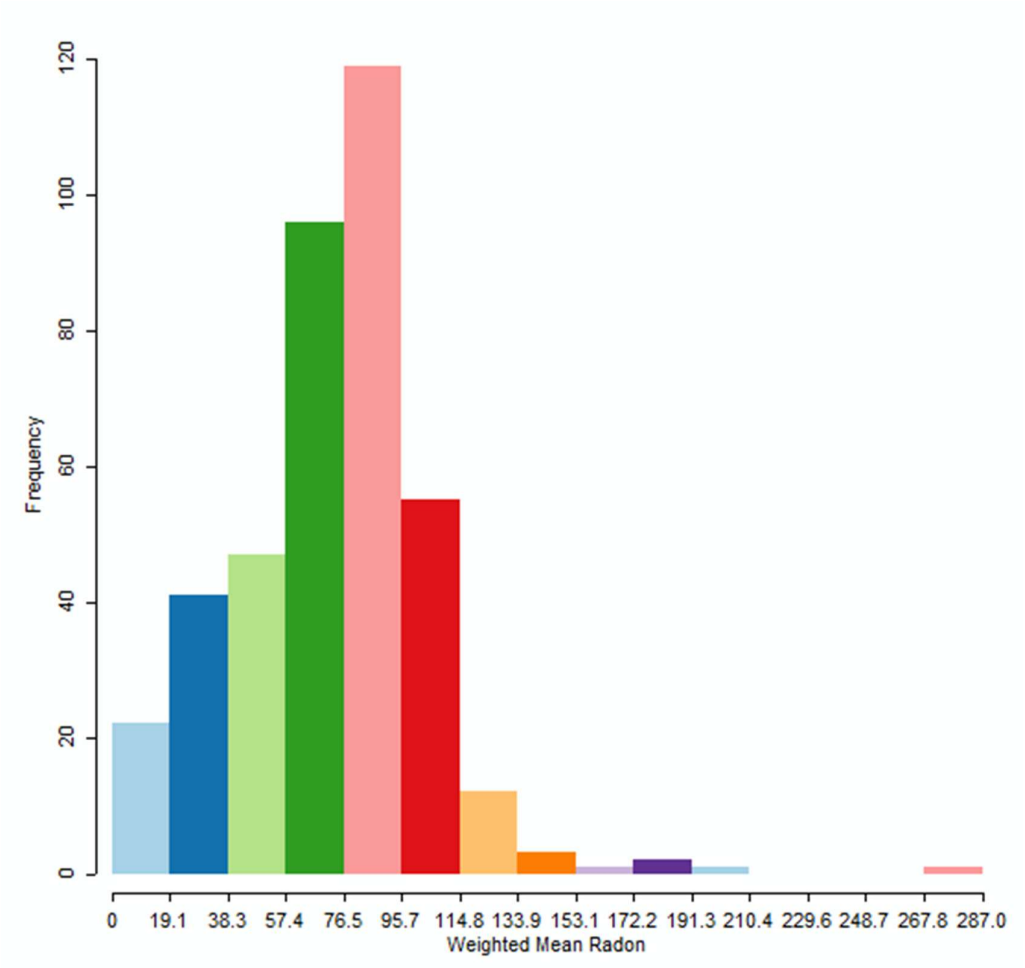
8.2 Histograms

8.2.1 Independent Variables

Mean Income per Inhabitant



Population Weighted Mean Radon Soil Gas Concentration



8.3 Regression Reports

8.3.1 Case Fatality Ratio

8.3.1.1 Total Population

Period I

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :    mort_inf  Number of Observations:  400
Mean dependent var  :   0.0307698  Number of Variables   :    5
S.D. dependent var  :   0.0112862  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.301361

R-squared           :    0.267327  Log likelihood        :     1284.66
Sq. Correlation     : -            Akaike info criterion :    -2559.32
Sigma-square        : 9.33263e-05  Schwarz criterion     :    -2539.36
S.E of regression   :  0.00966055
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
        W_mort_inf      0.301361      0.0623372        4.83437     0.00000
          CONSTANT    -0.0644504      0.0134753       -4.78285     0.00000
rn_mean..pweighted   4.17143e-05    1.56161e-05        2.67123     0.00756
        income2019  -1.07496e-07    1.89846e-07      -0.566226     0.57124
           avg_age    0.00189545    0.000267983        7.07302     0.00000
```

With Infection Rate:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :    mort_inf  Number of Observations:  400
Mean dependent var  :   0.0307698  Number of Variables   :    6
S.D. dependent var  :   0.0112862  Degrees of Freedom    :  394
Lag coeff.   (Rho)  :    0.246541

R-squared           :    0.303616  Log likelihood        :     1296.07
Sq. Correlation     : -            Akaike info criterion :    -2580.13
Sigma-square        : 8.87039e-05  Schwarz criterion     :    -2556.19
S.E of regression   :  0.00941828
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
        W_mort_inf      0.246541      0.0630477        3.91038     0.00009
          CONSTANT    -0.0764353      0.0134132       -5.69852     0.00000
           avg_age    0.00210372     0.00026667        7.88887     0.00000
        income2019  -1.28137e-07    1.85208e-07      -0.691855     0.48903
rn_mean..pweighted   1.18137e-05     1.6499e-05       0.716025     0.47398
           inf_pop      0.331923      0.0700668        4.73724     0.00000
```

Period II

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v2mort_inf  Number of Observations:  400
Mean dependent var  :  0.00801468  Number of Variables   :    5
S.D. dependent var  :  0.00381278  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.267816

R-squared           :    0.227129  Log likelihood        :     1708.86
Sq. Correlation     : -            Akaike info criterion :    -3407.73
Sigma-square        : 1.12354e-05  Schwarz criterion     :    -3387.77
S.E of regression   :  0.00335193
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_v2mort_inf      0.267816      0.0642255        4.16992     0.00003
          CONSTANT    -0.0065685     0.00466916       -1.40679     0.15949
rn_mean..pweighted   2.14671e-05    5.45915e-06        3.93231     0.00008
        income2019  -2.39004e-07    6.61157e-08       -3.61493     0.00030
           avg_age   0.000366948    9.14811e-05        4.01119     0.00006
```

With Infection Rate:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v2mort_inf  Number of Observations:  400
Mean dependent var  :  0.00801468  Number of Variables   :    6
S.D. dependent var  :  0.00381278  Degrees of Freedom    :  394
Lag coeff.   (Rho)  :    0.215047

R-squared           :    0.249040  Log likelihood        :     1715.66
Sq. Correlation     : -            Akaike info criterion :    -3419.31
Sigma-square        : 1.09169e-05  Schwarz criterion     :    -3395.36
S.E of regression   :  0.00330408
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v2mort_inf      0.215047      0.0657172        3.27231     0.00107
          CONSTANT   -0.00387738     0.00466886      -0.830476     0.40627
           avg_age    0.00030861    9.18426e-05        3.36021     0.00078
        income2019  -2.51652e-07    6.49681e-08       -3.87346     0.00011
rn_mean..pweighted    1.5335e-05    5.68073e-06        2.69948     0.00694
         v2inf_pop     0.0527458      0.0144709        3.64495     0.00027
--------------------------------------------------------------------------------
```

Period III

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v3mort_inf  Number of Observations:  400
Mean dependent var  :  0.00874164  Number of Variables   :    5
S.D. dependent var  :   0.0036832  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.507299

R-squared           :    0.586951  Log likelihood        :     1839.63
Sq. Correlation     : -            Akaike info criterion :    -3669.27
Sigma-square        : 5.60341e-06  Schwarz criterion     :    -3649.31
S.E of regression   :  0.00236715
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v3mort_inf      0.507299       0.048081        10.5509     0.00000
          CONSTANT    -0.0218534     0.00335303       -6.5175     0.00000
rn_mean..pweighted   1.98028e-05    3.96904e-06        4.98931     0.00000
        income2019  -1.19302e-07    4.81565e-08       -2.47739     0.01323
           avg_age    0.00061105    6.93428e-05        8.81202     0.00000
--------------------------------------------------------------------------------
```

With infection rate:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v3mort_inf  Number of Observations:  400
Mean dependent var  :  0.00874164  Number of Variables   :    6
S.D. dependent var  :   0.0036832  Degrees of Freedom    :  394
Lag coeff.   (Rho)  :    0.456152

R-squared           :    0.599260  Log likelihood        :     1848.04
Sq. Correlation     : -            Akaike info criterion :    -3684.09
Sigma-square        : 5.43642e-06  Schwarz criterion     :    -3660.14
S.E of regression   :  0.00233161
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v3mort_inf      0.456152      0.0495079        9.21373     0.00000
          CONSTANT    -0.0251228     0.00343351       -7.31695     0.00000
         v3inf_pop     0.0131293     0.00325047        4.03921     0.00005
           avg_age   0.000659068    7.01319e-05        9.39756     0.00000
        income2019  -1.40485e-07    4.72896e-08       -2.97074     0.00297
rn_mean..pweighted   1.50644e-05    4.09074e-06        3.68255     0.00023
```

8.3.1.2 Age Group equal to or greater than 60 years:

Period I

Without Age:

```
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  mort_inf60plus
Number of Observations:  400
Mean dependent var  :    0.115508  Number of Variables   :    4
S.D. dependent var  :   0.0292931  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.278712

R-squared           :    0.064606  Log likelihood        :     854.848
Sq. Correlation     : -            Akaike info criterion :     -1701.7
Sigma-square        : 0.000802647  Schwarz criterion     :    -1685.73
S.E of regression   :    0.028331
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_mort_inf60      0.278712      0.0663749        4.19905     0.00003
          CONSTANT     0.0903095      0.0143728        6.28335     0.00000
rn_mean..pweighted   7.85176e-05    4.47389e-05        1.75502     0.07926
       income2019  -5.18219e-07    5.17113e-07       -1.00214     0.31628
------------------------------------------------------------------------------
```

With Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  mort_inf60plus
Number of Observations:  400
Mean dependent var  :    0.115508  Number of Variables   :    5
S.D. dependent var  :   0.0292931  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.278644

R-squared           :    0.064629  Log likelihood        :     854.854
Sq. Correlation     : -            Akaike info criterion :    -1699.71
Sigma-square        : 0.000802627  Schwarz criterion     :    -1679.75
S.E of regression   :   0.0283307
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_mort_inf60      0.278644      0.0663881         4.1972     0.00003
          CONSTANT     0.0946283      0.0400836        2.36077     0.01824
rn_mean..pweighted   7.93513e-05     4.5304e-05        1.75153     0.07985
       income2019  -5.38548e-07    5.46353e-07      -0.985715     0.32427
          avg_age   -8.6276e-05    0.000751243      -0.114844     0.90857
```

With Infections / Inhabitant:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  mort_inf60plus
Number of Observations:  400
Mean dependent var  :    0.115508  Number of Variables   :    5
S.D. dependent var  :   0.0292931  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.227334

R-squared           :    0.132110  Log likelihood        :     870.896
Sq. Correlation     : -            Akaike info criterion :    -1731.79
Sigma-square        : 0.000744723  Schwarz criterion     :    -1711.83
S.E of regression   :   0.0272896

------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_mort_inf60      0.227334      0.0665344        3.41679     0.00063
          CONSTANT     0.0838094      0.0137443        6.09774     0.00000
        income2019   -5.0057e-07    4.98134e-07       -1.00489     0.31495
rn_mean..pweighted  -3.21657e-05    4.72085e-05      -0.681353     0.49565
     inf_pop60plus       1.12706       0.202332        5.57035     0.00000
```

Period II

Without Age

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v2mort_inf60plus
Number of Observations:  400
Mean dependent var  :   0.0441911  Number of Variables   :    4
S.D. dependent var  :   0.0183542  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :      0.2041

R-squared           :    0.096373  Log likelihood        :     1050.23
Sq. Correlation     : -            Akaike info criterion :    -2092.45
Sigma-square        :  0.00030441  Schwarz criterion     :    -2076.49
S.E of regression   :   0.0174473
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v2mort_inf        0.2041       0.067881        3.00674     0.00264
          CONSTANT     0.0587082     0.00833138        7.04664     0.00000
rn_mean..pweighted   9.15079e-05    2.78189e-05        3.28942     0.00100
        income2019  -1.27072e-06    3.21782e-07       -3.94902     0.00008
```

With Age

```
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v2mort_inf60plus
Number of Observations:  400
Mean dependent var  :   0.0441911  Number of Variables   :    5
S.D. dependent var  :   0.0183542  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.205701

R-squared           :    0.096537  Log likelihood        :     1050.24
Sq. Correlation     : -            Akaike info criterion :    -2090.47
Sigma-square        : 0.000304355  Schwarz criterion     :    -2070.52
S.E of regression   :   0.0174458

--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v2mort_inf      0.205701      0.0680539        3.02262     0.00251
          CONSTANT     0.0619069      0.0243572        2.54162     0.01103
rn_mean..pweighted   9.19603e-05    2.81131e-05        3.27108     0.00107
        income2019  -1.28505e-06    3.38482e-07       -3.79652     0.00015
           avg_age  -6.56662e-05    0.000463923      -0.141545     0.88744
```

With Infections per Inhabitant:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v2mort_inf60plus
Number of Observations:  400
Mean dependent var  :   0.0441911  Number of Variables   :    5
S.D. dependent var  :   0.0183542  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.137687

R-squared           :    0.133363  Log likelihood        :     1059.48
Sq. Correlation     : -            Akaike info criterion :    -2108.97
Sigma-square        : 0.000291949  Schwarz criterion     :    -2089.01
S.E of regression   :   0.0170865

--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v2mort_inf      0.137687      0.0691746        1.99042     0.04654
          CONSTANT     0.0583365     0.00811315        7.19036     0.00000
        income2019  -1.26394e-06    3.15068e-07       -4.01165     0.00006
rn_mean..pweighted   4.71734e-05    2.93565e-05        1.60692     0.10807
   v2inf_pop60plus      0.551143       0.129245        4.26432     0.00002
```

Period III

Without Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v3mort_inf60plus
Number of Observations:  400
Mean dependent var  :   0.0560841  Number of Variables   :    4
S.D. dependent var  :   0.0147713  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.428753

R-squared           :    0.251182  Log likelihood        :     1168.57
Sq. Correlation     : -            Akaike info criterion :    -2329.14
Sigma-square        : 0.000163385  Schwarz criterion     :    -2313.17
S.E of regression   :   0.0127822
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v3mort_inf      0.428753      0.0582116        7.36542     0.00000
          CONSTANT     0.0498807     0.00672526        7.41692     0.00000
rn_mean..pweighted   8.00399e-05    2.06584e-05        3.87446     0.00011
       income2019  -9.86538e-07     2.3711e-07       -4.16067     0.00003
```

With Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v3mort_inf60plus
Number of Observations:  400
Mean dependent var  :   0.0560841  Number of Variables   :    5
S.D. dependent var  :   0.0147713  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.391245

R-squared           :    0.258237  Log likelihood        :     1171.85
Sq. Correlation     : -            Akaike info criterion :     -2333.7
Sigma-square        : 0.000161846  Schwarz criterion     :    -2313.75
S.E of regression   :   0.0127219
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v3mort_inf      0.391245      0.0597913        6.54351     0.00000
          CONSTANT    0.00813417      0.0177438       0.458423     0.64665
rn_mean..pweighted   7.51735e-05    2.07148e-05        3.62898     0.00028
       income2019  -8.24726e-07    2.49539e-07         -3.305     0.00095
          avg_age   0.000894449    0.000343603        2.60314     0.00924
```

With Infections per Inhabitant:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v3mort_inf60plus
Number of Observations:  400
Mean dependent var  :   0.0560841  Number of Variables   :    5
S.D. dependent var  :   0.0147713  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.365677

R-squared           :    0.295085  Log likelihood        :      1182.9
Sq. Correlation     : -            Akaike info criterion :     -2355.8
Sigma-square        : 0.000153806  Schwarz criterion     :    -2335.84
S.E of regression   :   0.0124018

--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_v3mort_inf      0.365677      0.0595299        6.14274     0.00000
          CONSTANT     0.0437696     0.00645453        6.78122     0.00000
        income2019  -9.92555e-07    2.29891e-07       -4.31751     0.00002
rn_mean..pweighted   3.18644e-05    2.19162e-05        1.45392     0.14597
   v3inf_pop60plus      0.151621      0.0294284        5.15221     0.00000
```

8.3.1.3 Age Group below 60 years:

Period I

Without Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  mort_inf_bl60
Number of Observations:  400
Mean dependent var  :  0.00146717  Number of Variables   :    4
S.D. dependent var  :  0.00117616  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :   0.0387633

R-squared           :    0.041064  Log likelihood        :     2138.95
Sq. Correlation     : -            Akaike info criterion :    -4269.91
Sigma-square        : 1.32656e-06  Schwarz criterion     :    -4253.94
S.E of regression   :  0.00115176
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_mort_inf_b     0.0387633       0.072716       0.533077     0.59398
          CONSTANT    0.00335614    0.000521214        6.43908     0.00000
rn_mean..pweighted  -1.88772e-07    1.80942e-06      -0.104328     0.91691
       income2019   -8.2046e-08    2.12095e-08       -3.86836     0.00011
```

With Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  mort_inf_bl60
Number of Observations:  400
Mean dependent var  :  0.00146717  Number of Variables   :    5
S.D. dependent var  :  0.00117616  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :   0.0307465

R-squared           :    0.048949  Log likelihood        :     2140.63
Sq. Correlation     : -            Akaike info criterion :    -4271.25
Sigma-square        : 1.31565e-06  Schwarz criterion     :    -4251.29
S.E of regression   :  0.00114702
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_mort_inf_b     0.0307465      0.0730449       0.420925     0.67381
          CONSTANT    0.00058674     0.00159613       0.367602     0.71317
rn_mean..pweighted  -7.29497e-07    1.82648e-06      -0.399401     0.68960
           avg_age   5.58398e-05    3.06892e-05        1.81953     0.06883
        income2019  -6.92701e-08    2.22965e-08       -3.10677     0.00189
```

Period II

Without Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v2mort_inf_bl60
Number of Observations:  400
Mean dependent var  :  0.00082084  Number of Variables   :    4
S.D. dependent var  : 0.000849797  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :   0.0255878

R-squared           :    0.033364  Log likelihood        :     2267.39
Sq. Correlation     : -            Akaike info criterion :    -4526.78
Sigma-square        : 6.98062e-07  Schwarz criterion     :    -4510.82
S.E of regression   : 0.000835501

------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_v2mort_inf     0.0255878      0.0730734       0.350165     0.72621
          CONSTANT    0.00184323    0.000367384        5.01719     0.00000
rn_mean..pweighted   2.59908e-06    1.31742e-06        1.97286     0.04851
       income2019   -5.2219e-08     1.5339e-08       -3.40432     0.00066
------------------------------------------------------------------------------
```

With Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v2mort_inf_bl60
Number of Observations:  400
Mean dependent var  :  0.00082084  Number of Variables   :    5
S.D. dependent var  : 0.000849797  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :   0.0282201

R-squared           :    0.035049  Log likelihood        :     2267.74
Sq. Correlation     : -            Akaike info criterion :    -4525.47
Sigma-square        : 6.96845e-07  Schwarz criterion     :    -4505.51
S.E of regression   : 0.000834772

------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_v2mort_inf     0.0282201       0.073136       0.385858     0.69960
          CONSTANT    0.00275881     0.00116219        2.37381     0.01761
rn_mean..pweighted   2.77386e-06    1.33242e-06        2.08182     0.03736
        income2019  -5.65235e-08    1.61386e-08       -3.50237     0.00046
           avg_age  -1.83702e-05    2.21545e-05      -0.829184     0.40700
------------------------------------------------------------------------------
```

Period III

Without Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v3mort_inf_bl60
Number of Observations:  400
Mean dependent var  : 0.000556737  Number of Variables   :    4
S.D. dependent var  : 0.000284648  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.370319

R-squared           :    0.242582  Log likelihood        :     2748.06
Sq. Correlation     : -            Akaike info criterion :    -5488.11
Sigma-square        : 6.13692e-08  Schwarz criterion     :    -5472.15
S.E of regression   : 0.000247728
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_v3mort_inf      0.370319      0.0596404         6.2092     0.00000
          CONSTANT   0.000864634    0.000117697        7.34627     0.00000
rn_mean..pweighted   1.96591e-06    4.04638e-07        4.85843     0.00000
        income2019  -2.75524e-08    4.66023e-09       -5.91225     0.00000
```

With Age:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  v3mort_inf_bl60
Number of Observations:  400
Mean dependent var  : 0.000556737  Number of Variables   :    5
S.D. dependent var  : 0.000284648  Degrees of Freedom    :  395
Lag coeff.   (Rho)  :    0.325265

R-squared           :    0.254258  Log likelihood        :     2752.52
Sq. Correlation     : -            Akaike info criterion :    -5495.04
Sigma-square        : 6.04232e-08  Schwarz criterion     :    -5475.09
S.E of regression   : 0.000245811
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_v3mort_inf      0.325265      0.0615138        5.28766     0.00000
          CONSTANT  -9.68992e-05    0.000342408      -0.282993     0.77718
rn_mean..pweighted   1.84005e-06    4.03277e-07        4.56274     0.00001
        income2019  -2.37889e-08    4.87527e-09       -4.87951     0.00000
           avg_age   2.00844e-05    6.70068e-06        2.99737     0.00272
```

8.3.2 Infections per Inhabitant

Period I

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :     inf_pop  Number of Observations:  400
Mean dependent var  :   0.0208292  Number of Variables   :    4
S.D. dependent var  :   0.0077055  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.792601

R-squared           :    0.719961  Log likelihood        :     1600.05
Sq. Correlation     : -            Akaike info criterion :    -3192.09
Sigma-square        : 1.66272e-05  Schwarz criterion     :    -3176.13
S.E of regression   :  0.00407765
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
         W_inf_pop      0.792601       0.030727        25.7949     0.00000
          CONSTANT     0.0172274      0.0028797        5.98235     0.00000
rn_mean..pweighted   2.65188e-05     6.6958e-06        3.96051     0.00007
avg_liv..nhabitant  -0.000324237    5.78708e-05       -5.60277     0.00000
```

Age >= 60 yrs:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  I: Inf/Inhab. >60 yrs
Number of Observations:  400
Mean dependent var  :   0.0176936  Number of Variables   :    4
S.D. dependent var  :  0.00779277  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.776778

R-squared           :    0.688965  Log likelihood        :     1576.34
Sq. Correlation     : -            Akaike info criterion :    -3144.68
Sigma-square        : 1.88883e-05  Schwarz criterion     :    -3128.72
S.E of regression   :  0.00434607
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_I: Inf/Inh      0.776778      0.0331864        23.4066     0.00000
          CONSTANT     0.0132216     0.00289988        4.55936     0.00001
 Mean Living Space  -0.000250177    5.97517e-05       -4.18694     0.00003
Weighte..ean Radon   2.95128e-05    7.18563e-06         4.1072     0.00004
```

Age below 60 yrs:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  I: Inf/Inhab. <60 yrs
Number of Observations:  400
Mean dependent var  :   0.0221321  Number of Variables   :    4
S.D. dependent var  :  0.00786177  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.779408

R-squared           :    0.701076  Log likelihood        :     1580.47
Sq. Correlation     : -            Akaike info criterion :    -3152.93
Sigma-square        : 1.84757e-05  Schwarz criterion     :    -3136.97
S.E of regression   :  0.00429834
------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_I: Inf/Inh      0.779408      0.0319888         24.365     0.00000
          CONSTANT     0.0187048     0.00305686        6.11895     0.00000
 Mean Living Space  -0.000345281    6.10221e-05       -5.65829     0.00000
Weighte..ean Radon   2.71111e-05     7.0257e-06        3.85885     0.00011
```

Period II

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :   v2inf_pop  Number of Observations:  400
Mean dependent var  :   0.0207946  Number of Variables   :    4
S.D. dependent var  :   0.0132249  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.863313

R-squared           :    0.834825  Log likelihood        :      1480.1
Sq. Correlation     : -            Akaike info criterion :    -2952.21
Sigma-square        : 2.88886e-05  Schwarz criterion     :    -2936.24
S.E of regression   :  0.00537481
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
       W_v2inf_pop      0.863313      0.0255538        33.7841     0.00000
          CONSTANT    0.00760159     0.00337845        2.25002     0.02445
rn_mean..pweighted   2.36797e-05    8.62027e-06        2.74698     0.00601
avg_liv..nhabitant   -0.00014994     7.1692e-05       -2.09145     0.03649
```

Age >= 60 years:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  II: Inf/Inhab. >60 yrs
Number of Observations:  400
Mean dependent var  :   0.0113941  Number of Variables   :    4
S.D. dependent var  :  0.00753447  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :     0.84166

R-squared           :    0.802048  Log likelihood        :     1672.13
Sq. Correlation     : -            Akaike info criterion :    -3336.26
Sigma-square        : 1.12374e-05  Schwarz criterion     :     -3320.3
S.E of regression   :  0.00335222
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_II: Inf/In       0.84166      0.0281467        29.9026     0.00000
          CONSTANT     0.0038437     0.00209022         1.8389     0.06593
Weighte..ean Radon     1.496e-05    5.37291e-06        2.78435     0.00536
 Mean Living Space  -7.36595e-05    4.44987e-05       -1.65532     0.09786
```

Age below60 years:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  II: Inf/Inhab. <60 yrs
Number of Observations:  400
Mean dependent var  :   0.0250391  Number of Variables   :    4
S.D. dependent var  :   0.0163784  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.863344

R-squared           :    0.830989  Log likelihood        :     1389.96
Sq. Correlation     : -            Akaike info criterion :    -2771.92
Sigma-square        : 4.53374e-05  Schwarz criterion     :    -2755.96
S.E of regression   :   0.0067333
--------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
--------------------------------------------------------------------------------
      W_II: Inf/In      0.863344      0.0255859        33.7429     0.00000
          CONSTANT     0.0081691     0.00420693        1.94182     0.05216
Weighte..ean Radon    3.0595e-05    1.08257e-05        2.82615     0.00471
 Mean Living Space  -0.000162511    8.95377e-05         -1.815     0.06952
```

Period III

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :   v3inf_pop  Number of Observations:  400
Mean dependent var  :    0.181915  Number of Variables   :    4
S.D. dependent var  :   0.0410346  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :     0.75803

R-squared           :    0.759152  Log likelihood        :     965.006
Sq. Correlation     : -            Akaike info criterion :    -1922.01
Sigma-square        : 0.000405549  Schwarz criterion     :    -1906.05
S.E of regression   :   0.0201382

------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
       W_v3inf_pop       0.75803      0.0327562        23.1415     0.00000
          CONSTANT      0.129922      0.0165451        7.85259     0.00000
rn_mean..pweighted   0.000115873    3.24757e-05        3.56801     0.00036
avg_liv..nhabitant   -0.00206708    0.000295005       -7.00694     0.00000
------------------------------------------------------------------------------
```

Age >= 60 years:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  III: Inf/Inhab. >60 yrs
Number of Observations:  400
Mean dependent var  :   0.0871758  Number of Variables   :    4
S.D. dependent var  :   0.0254661  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.792822

R-squared           :    0.786733  Log likelihood        :     1176.33
Sq. Correlation     : -            Akaike info criterion :    -2344.67
Sigma-square        : 0.000138308  Schwarz criterion     :     -2328.7
S.E of regression   :   0.0117604

------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_III: Inf/I      0.792822       0.030535        25.9644     0.00000
          CONSTANT     0.0540796      0.0086732        6.23525     0.00000
Weighte..ean Radon   8.51519e-05    1.95644e-05        4.35239     0.00001
 Mean Living Space  -0.000930449    0.000166732        -5.5805     0.00000
```

Age below 60 years:

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : vg250_krs
Spatial Weight      : vg250_krs_queen1order
Dependent Variable  :  III: Inf/Inhab. <60 yrs
Number of Observations:  400
Mean dependent var  :    0.222853  Number of Variables   :    4
S.D. dependent var  :   0.0490151  Degrees of Freedom    :  396
Lag coeff.   (Rho)  :    0.757361

R-squared           :    0.745103  Log likelihood        :     882.652
Sq. Correlation     : -            Akaike info criterion :     -1757.3
Sigma-square        : 0.000612384  Schwarz criterion     :    -1741.34
S.E of regression   :   0.0247464

------------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
------------------------------------------------------------------------------
      W_III: Inf/I      0.757361      0.0337697        22.4273     0.00000
          CONSTANT      0.142466      0.0198757        7.16785     0.00000
Weighte..ean Radon   0.000139358    3.99193e-05          3.491     0.00048
 Mean Living Space   -0.00216206    0.000354432       -6.10006     0.00000
```

8.4 SQL-Query Samples

//update table and add to each district the corresponding sum of all fatalities before 01.01.2021:

```
ALTER TABLE public.vg250_krs
ADD summort integer;
UPDATE public.vg250_krs
SET summort=t.summort
FROM
(       SELECT SUM("AnzahlTodesfall") as summort,
  "IdLandkreis" as id
   FROM public.rki_covid19
   WHERE "AnzahlTodesfall" >= 1 --Todesfälle
   and "Refdatum" <='2021-01-01' --vor dem 01.01.2021
       GROUP BY "IdLandkreis"
   ) t
WHERE cast(vg250_krs.ars AS INT) = t.id
```

//assign the sum of all fatalities in Berlin RKI district IDs (between 11000 and 12000) (Berlin districts) to the Berlin district ID (ID: 11000) of the Administration Boundary Table (vg250_krs)

```
UPDATE public.vg250_krs
SET summort=t.summort
FROM
(       SELECT SUM("AnzahlTodesfall") as summort
   FROM public.rki_covid19
   WHERE "AnzahlTodesfall" >= 1 --Todesfälle
   and "Refdatum" <='2021-01-01' --vor dem 01.01.2021
   and "IdLandkreis" >= 11000 and "IdLandkreis" < 12000
   ) t
WHERE cast(vg250_krs.ars AS INT) = 11000
```

8.5 Vaccination Dataset

Translated:
Source: https://github.com/robert-koch-institut/COVID-19-Impfungen_in_Deutschland

Variable	Type	Values	Description
Datum	Date	YYYY-MM-DD	Date of Vaccination
LandkreisId_Impfort	Text	01001 bis 16077: district ID 17000 : Federal Deparment u: unknown	District ID basing on „ amtlicher Gemeindeschlüssel“ (AGS). Vaccination of the Federal Department are listed seperately, as the federal vaccination centres don’t report an exact location.
Altersgruppe	Text	05-11: Ages 5 to 11 years 12-17: Ages 12 to 17 years 18-59: Ages 18 to 59 years 60+: Ages 60 years and older	Age groups of reported vaccination cases based on The National Association of Statutory Health Insurance Physicians (NASHIP; KBV: Kassenärztliche Bundesvereinigung)
Impfschutz	integer	1: incomplete vaccination 2: complete vaccination 3: booster vaccination	Vaccination protection information: Complete vaccination protection is given with two vaccinations, vaccination with Janssen, single vaccination with recovered SARS-CoV-2 infection.
Anzahl	integer	≥5	Count of vaccinations of reported group

8.6 RKI-COVID JSON Table:

The included dataset is presented in CSV and JSON format, and features the following columns:

- FID (ObjectId): Unique ID for reach entry in the dataset
- IdBundesland (StateId): Id of the federal state of the case
- Bundesland (State): Name of state
- Landkreis (County): Name of the county
- Altersgruppe (AgeGroup): Age group of the case from the 6 groups 0-4, 5-14, 15-34, 35-59, 60-79, 80+, and unknown
- Geschlecht (Gender): Gender of case
- AnzahlFall (NumberOfCases): Number of cases in the corresponding group
- AnzahlTodesfall (NumberOfDeaths): Number of deaths in the corresponding group
- Meldedatum (RegistrationDate): Date when the case became known to the health department
- IdLandkreis (CountyId): Id of the county of the case
- Datenstand (DateStatus): Date when the data record was last updated
- NeuerFall (NewCase):
    - 0: Case is included in the publication for the current day and in the one for the previous day
    - 1: Case is only included in the current publication
    - -1: Case is only included in the previous day's publication
- NeuerTodesfall (NewDeath):
    - 0: In the publication, there is one death for the current day and one for the previous day
    - 1: In the current publication, the case is a death, but not in the previous day's publication
    - -1: The case is not a death in the current publication, but it was a death in the previous day's publication
    - -9: The case is neither a death in the current publication nor in the previous day
- Refdatum (RefDate)
- NeuGenesen (NewRecovery):
    - 0: The case is in the publication for the current day and in the one for the previous day
    - 1: Case is recovered in the current publication, but not in the previous day's publication
    - -1: Case is not recovered in the current publication, but it was recovered in the publication of the previous day
    - -9: The case is neither recovered in the current publication nor in the previous day
- AnzahlGenesen (NumberOfRecoveries): Number of recoveries in the corresponding group
- IstErkrankungsbeginn (IsOnsetOfIllness): 1 if Refdatum is the onset of illness, 0 otherwise
- Altersgruppe2 (AgeGroup2): Age group of the case from 5-year groups 0-4, 5-9, 10-14, ..., 75-79, 80+, and unknown

Series from Lund University

Department of Physical Geography and Ecosystem Science

Master Thesis in Geographical Information Science

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland (2008).

2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research (2008).

3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing (2009).

4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia (2009).

5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).

6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).

7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010).

8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010).

9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010).

10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011).

11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011).

12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).

13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).

14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011).

15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011).

16. *David Kanyari*: Nairobi City Journey Planner: An online and a Mobile Application (2011).

17. *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012).

18. *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012).

19. *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012).

20. *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013).

21. *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013).

22. *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013).

23. *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013).

24. *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013).

25. *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013).

26. *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014).

27. *Inese Linuza:* Accuracy Assessment in Glacier Change Analysis (2014).

28. *Agnieszka Griffin:* Domestic energy consumption and social living standards: a GIS analysis within the Greater London Authority area (2014).

29. *Brynja Guðmundsdóttir:* Detection of potential arable land with remote sensing and GIS - A Case Study for Kjósarhreppur (2014).

30. *Oleksandr Nekrasov:* Processing of MODIS Vegetation Indices for analysis of agricultural droughts in the southern Ukraine between the years 2000-2012 (2014).

31. *Sarah Tressel:* Recommendations for a polar Earth science portal in the context of Arctic Spatial Data Infrastructure (2014).

32. *Caroline Gevaert:* Combining Hyperspectral UAV and Multispectral Formosat-2 Imagery for Precision Agriculture Applications (2014).

33. *Salem Jamal-Uddeen:* Using GeoTools to implement the multi-criteria evaluation analysis - weighted linear combination model (2014).

34. *Samanah Seyedi-Shandiz:* Schematic representation of geographical railway network at the Swedish Transport Administration (2014).

35. *Kazi Masel Ullah:* Urban Land-use planning using Geographical Information System and analytical hierarchy process: case study Dhaka City (2014).

36. *Alexia Chang-Wailing Spitteler:* Development of a web application based on MCDA and GIS for the decision support of river and floodplain rehabilitation projects (2014).

37. *Alessandro De Martino:* Geographic accessibility analysis and evaluation of potential changes to the public transportation system in the City of Milan (2014).

38. *Alireza Mollasalehi:* GIS Based Modelling for Fuel Reduction Using Controlled Burn in Australia. Case Study: Logan City, QLD (2015).

39. *Negin A. Sanati:* Chronic Kidney Disease Mortality in Costa Rica; Geographical Distribution, Spatial Analysis and Non-traditional Risk Factors (2015).

40. *Karen McIntyre:* Benthic mapping of the Bluefields Bay fish sanctuary, Jamaica (2015).

41. *Kees van Duijvendijk:* Feasibility of a low-cost weather sensor network for agricultural purposes: A preliminary assessment (2015).

42. *Sebastian Andersson Hylander:* Evaluation of cultural ecosystem services using GIS (2015).

43. *Deborah Bowyer:* Measuring Urban Growth, Urban Form and Accessibility as Indicators of Urban Sprawl in Hamilton, New Zealand (2015).

44. *Stefan Arvidsson:* Relationship between tree species composition and phenology extracted from satellite data in Swedish forests (2015).

45. *Damián Giménez Cruz*: GIS-based optimal localisation of beekeeping in rural Kenya (2016).

46. *Alejandra Narváez Vallejo:* Can the introduction of the topographic indices in LPJ-GUESS improve the spatial representation of environmental variables? (2016).

47. *Anna Lundgren:* Development of a method for mapping the highest coastline in Sweden using breaklines extracted from high resolution digital elevation models (2016).

48. *Oluwatomi Esther Adejoro:* Does location also matter? A spatial analysis of social achievements of young South Australians (2016).

49. *Hristo Dobrev Tomov:* Automated temporal NDVI analysis over the Middle East for the period 1982 - 2010 (2016).

50. *Vincent Muller:* Impact of Security Context on Mobile Clinic Activities A GIS Multi Criteria Evaluation based on an MSF Humanitarian Mission in Cameroon (2016).

51. *Gezahagn Negash Seboka:* Spatial Assessment of NDVI as an Indicator of Desertification in Ethiopia using Remote Sensing and GIS (2016).

52. *Holly Buhler:* Evaluation of Interfacility Medical Transport Journey Times in Southeastern British Columbia. (2016).

53. *Lars Ole Grottenberg*: Assessing the ability to share spatial data between emergency management organisations in the High North (2016).

54. *Sean Grant:* The Right Tree in the Right Place: Using GIS to Maximize the Net Benefits from Urban Forests (2016).

55. *Irshad Jamal:* Multi-Criteria GIS Analysis for School Site Selection in Gorno-Badakhshan Autonomous Oblast, Tajikistan (2016).

56. *Fulgencio Sanmartín:* Wisdom-volkano: A novel tool based on open GIS and time-series visualization to analyse and share volcanic data (2016).

57. *Nezha Acil:* Remote sensing-based monitoring of snow cover dynamics and its influence on vegetation growth in the Middle Atlas Mountains (2016).

58. *Julia Hjalmarsson:* A Weighty Issue: Estimation of Fire Size with Geographically Weighted Logistic Regression (2016).

59. *Mathewos Tamiru Amato:* Using multi-criteria evaluation and GIS for chronic food and nutrition insecurity indicators analysis in Ethiopia (2016).

60. *Karim Alaa El Din Mohamed Soliman El Attar:* Bicycling Suitability in Downtown, Cairo, Egypt (2016).

61. *Gilbert Akol Echelai:* Asset Management: Integrating GIS as a Decision Support Tool in Meter Management in National Water and Sewerage Corporation (2016).

62. *Terje Slinning:* Analytic comparison of multibeam echo soundings (2016).

63. *Gréta Hlín Sveinsdóttir:* GIS-based MCDA for decision support: A framework for wind farm siting in Iceland (2017).

64. *Jonas Sjögren:* Consequences of a flood in Kristianstad, Sweden: A GIS-based analysis of impacts on important societal functions (2017).

65. *Nadine Raska:* 3D geologic subsurface modelling within the Mackenzie Plain, Northwest Territories, Canada (2017).

66. *Panagiotis Symeonidis*: Study of spatial and temporal variation of atmospheric optical parameters and their relation with PM 2.5 concentration over Europe using GIS technologies (2017).

67. *Michaela Bobeck:* A GIS-based Multi-Criteria Decision Analysis of Wind Farm Site Suitability in New South Wales, Australia, from a Sustainable Development Perspective (2017).

68. *Raghdaa Eissa*: Developing a GIS Model for the Assessment of Outdoor Recreational Facilities in New Cities Case Study: Tenth of Ramadan City, Egypt (2017).

69. *Zahra Khais Shahid*: Biofuel plantations and isoprene emissions in Svea and Götaland (2017).

70. *Mirza Amir Liaquat Baig*: Using geographical information systems in epidemiology: Mapping and analyzing occurrence of diarrhea in urban - residential area of Islamabad, Pakistan (2017).

71. *Joakim Jörwall*: Quantitative model of Present and Future well-being in the EU-28: A spatial Multi-Criteria Evaluation of socioeconomic and climatic comfort factors (2017).

72. *Elin Haettner*: Energy Poverty in the Dublin Region: Modelling Geographies of Risk (2017).

73. *Harry Eriksson*: Geochemistry of stream plants and its statistical relations to soil- and bedrock geology, slope directions and till geochemistry. A GIS-analysis of small catchments in northern Sweden (2017).

74. *Daniel Gardevärn:* PPGIS and Public meetings – An evaluation of public participation methods for urban planning (2017).

75. *Kim Friberg:* Sensitivity Analysis and Calibration of Multi Energy Balance Land Surface Model Parameters (2017).

76. *Viktor Svanerud:* Taking the bus to the park? A study of accessibility to green areas in Gothenburg through different modes of transport (2017).

77. *Lisa-Gaye Greene*: Deadly Designs: The Impact of Road Design on Road Crash Patterns along Jamaica's North Coast Highway (2017).

78. *Katarina Jemec Parker*: Spatial and temporal analysis of fecal indicator bacteria concentrations in beach water in San Diego, California (2017).

79. *Angela Kabiru*: An Exploratory Study of Middle Stone Age and Later Stone Age Site Locations in Kenya's Central Rift Valley Using Landscape Analysis: A GIS Approach (2017).

80. *Kristean Björkmann*: Subjective Well-Being and Environment: A GIS-Based Analysis (2018).

81. *Williams Erhunmonmen Ojo*: Measuring spatial accessibility to healthcare for people living with HIV-AIDS in southern Nigeria (2018).

82. *Daniel Assefa*: Developing Data Extraction and Dynamic Data Visualization (Styling) Modules for Web GIS Risk Assessment System (WGRAS). (2018).

83. *Adela Nistora*: Inundation scenarios in a changing climate: assessing potential impacts of sea-level rise on the coast of South-East England (2018).

84. *Marc Seliger*: Thirsty landscapes - Investigating growing irrigation water consumption and potential conservation measures within Utah's largest master-planned community: Daybreak (2018).

85. *Luka Jovičić*: Spatial Data Harmonisation in Regional Context in Accordance with INSPIRE Implementing Rules (2018).

86. *Christina Kourdounouli*: Analysis of Urban Ecosystem Condition Indicators for the Large Urban Zones and City Cores in EU (2018).

87. *Jeremy Azzopardi*: Effect of distance measures and feature representations on distance-based accessibility measures (2018).

88. *Patrick Kabatha*: An open source web GIS tool for analysis and visualization of elephant GPS telemetry data, alongside environmental and anthropogenic variables (2018).

89. *Richard Alphonce Giliba*: Effects of Climate Change on Potential Geographical Distribution of Prunus africana (African cherry) in the Eastern Arc Mountain Forests of Tanzania (2018).

90. *Eiður Kristinn Eiðsson*: Transformation and linking of authoritative multi-scale geodata for the Semantic Web: A case study of Swedish national building data sets (2018).

91. *Niamh Harty*: HOP!: a PGIS and citizen science approach to monitoring the condition of upland paths (2018).

92. *José Estuardo Jara Alvear*: Solar photovoltaic potential to complement hydropower in Ecuador: A GIS-based framework of analysis (2018).

93. *Brendan O'Neill*: Multicriteria Site Suitability for Algal Biofuel Production Facilities (2018).

94. *Roman Spataru*: Spatial-temporal GIS analysis in public health – a case study of polio disease (2018).

95. *Alicja Miodońska*: Assessing evolution of ice caps in Suðurland, Iceland, in years 1986 - 2014, using multispectral satellite imagery (2019).

96. *Dennis Lindell Schettini*: A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017 (2019).

97. *Damiano Vesentini*: The Po Delta Biosphere Reserve: Management challenges and priorities deriving from anthropogenic pressure and sea level rise (2019).

98. *Emilie Arnesten*: Impacts of future sea level rise and high water on roads, railways and environmental objects: a GIS analysis of the potential effects of increasing sea levels and highest projected high water in Scania, Sweden (2019).

99. *Syed Muhammad Amir Raza*: Comparison of geospatial support in RDF stores: Evaluation for ICOS Carbon Portal metadata (2019).

100. *Hemin Tofiq*: Investigating the accuracy of Digital Elevation Models from UAV images in areas with low contrast: A sandy beach as a case study (2019).

101. *Evangelos Vafeiadis*: Exploring the distribution of accessibility by public transport using spatial analysis. A case study for retail concentrations and public hospitals in Athens (2019).

102. *Milan Sekulic*: Multi-Criteria GIS modelling for optimal alignment of roadway by-passes in the Tlokweng Planning Area, Botswana (2019).

103. *Ingrid Piirisaar*: A multi-criteria GIS analysis for siting of utility-scale photovoltaic solar plants in county Kilkenny, Ireland (2019).

104. *Nigel Fox*: Plant phenology and climate change: possible effect on the onset of various wild plant species' first flowering day in the UK (2019).

105. *Gunnar Hesch*: Linking conflict events and cropland development in Afghanistan, 2001 to 2011, using MODIS land cover data and Uppsala Conflict Data Programme (2019).

106. *Elijah Njoku*: Analysis of spatial-temporal pattern of Land Surface Temperature (LST) due to NDVI and elevation in Ilorin, Nigeria (2019).

107. *Katalin Bunyevácz*: Development of a GIS methodology to evaluate informal urban green areas for inclusion in a community governance program (2019).

108. *Paul dos Santos*: Automating synthetic trip data generation for an agent-based simulation of urban mobility (2019).

109. *Robert O' Dwyer*: Land cover changes in Southern Sweden from the mid-Holocene to present day: Insights for ecosystem service assessments (2019).

110. *Daniel Klingmyr*: Global scale patterns and trends in tropospheric NO2 concentrations (2019).

111. *Marwa Farouk Elkabbany*: Sea Level Rise Vulnerability Assessment for Abu Dhabi, United Arab Emirates (2019).

112. *Jip Jan van Zoonen*: Aspects of Error Quantification and Evaluation in Digital Elevation Models for Glacier Surfaces (2020).

113. *Georgios Efthymiou*: The use of bicycles in a mid-sized city – benefits and obstacles identified using a questionnaire and GIS (2020).

114. *Haruna Olayiwola Jimoh*: Assessment of Urban Sprawl in MOWE/IBAFO Axis of Ogun State using GIS Capabilities (2020).

115. *Nikolaos Barmpas Zachariadis*: Development of an iOS, Augmented Reality for disaster management (2020).

116. *Ida Storm*: ICOS Atmospheric Stations: Spatial Characterization of CO2 Footprint Areas and Evaluating the Uncertainties of Modelled CO2 Concentrations (2020).

117. *Alon Zuta*: Evaluation of water stress mapping methods in vineyards using airborne thermal imaging (2020).

118. *Marcus Eriksson*: Evaluating structural landscape development in the municipality Upplands-Bro, using landscape metrics indices (2020).

119. *Ane Rahbek Vierø*: Connectivity for Cyclists? A Network Analysis of Copenhagen's Bike Lanes (2020).

120. *Cecilia Baggini*: Changes in habitat suitability for three declining Anatidae species in saltmarshes on the Mersey estuary, North-West England (2020).

121. *Bakrad Balabanian*: Transportation and Its Effect on Student Performance (2020).

122. *Ali Al Farid*: Knowledge and Data Driven Approaches for Hydrocarbon Microseepage Characterizations: An Application of Satellite Remote Sensing (2020).

123. *Bartlomiej Kolodziejczyk*: Distribution Modelling of Gene Drive-Modified Mosquitoes and Their Effects on Wild Populations (2020).

124. *Alexis Cazorla*: Decreasing organic nitrogen concentrations in European water bodies - links to organic carbon trends and land cover (2020).

125. *Kharid Mwakoba*: Remote sensing analysis of land cover/use conditions of community-based wildlife conservation areas in Tanzania (2021).

126. *Chinatsu Endo*: Remote Sensing Based Pre-Season Yellow Rust Early Warning in Oromia, Ethiopia (2021).

127. *Berit Mohr*: Using remote sensing and land abandonment as a proxy for long-term human out-migration. A Case Study: Al-Hassakeh Governorate, Syria (2021).

128. *Kanchana Nirmali Bandaranayake*: Considering future precipitation in delineation locations for water storage systems - Case study Sri Lanka (2021).

129. *Emma Bylund*: Dynamics of net primary production and food availability in the aftermath of the 2004 and 2007 desert locust outbreaks in Niger and Yemen (2021).

130. *Shawn Pace*: Urban infrastructure inundation risk from permanent sea-level rise scenarios in London (UK), Bangkok (Thailand) and Mumbai (India): A comparative analysis (2021).

131. *Oskar Evert Johansson*: The hydrodynamic impacts of Estuarine Oyster reefs, and the application of drone technology to this study (2021).

132. *Pritam Kumarsingh*: A Case Study to develop and test GIS/SDSS methods to assess the production capacity of a Cocoa Site in Trinidad and Tobago (2021).

133. *Muhammad Imran Khan*: Property Tax Mapping and Assessment using GIS (2021).

134. *Domna Kanari*: Mining geosocial data from Flickr to explore tourism patterns: The case study of Athens (2021).

135. Mona Tykesson Klubien: Livestock-MRSA in Danish pig farms (2021).

136. *Ove Njøten*: Comparing radar satellites. Use of Sentinel-1 leads to an increase in oil spill alerts in Norwegian waters (2021).

137. *Panagiotis Patrinos*: Change of heating fuel consumption patterns produced by the economic crisis in Greece (2021).

138. *Lukasz Langowski*: Assessing the suitability of using Sentinel-1A SAR multi-temporal imagery to detect fallow periods between rice crops (2021).

139. *Jonas Tillman*: Perception accuracy and user acceptance of legend designs for opacity data mapping in GIS (2022).

140. *Gabriela Olekszyk*: ALS (Airborne LIDAR) accuracy: Can potential low data quality of ground points be modelled/detected? Case study of 2016 LIDAR capture over Auckland, New Zealand (2022).

141. *Luke Aspland*: Weights of Evidence Predictive Modelling in Archaeology (2022).

142. *Luís Fareleira Gomes*: The influence of climate, population density, tree species and land cover on fire pattern in mainland Portugal (2022).

143. *Andreas Eriksson*: Mapping Fire Salamander (Salamandra salamandra) Habitat Suitability in Baden-Württemberg with Multi-Temporal Sentinel-1 and Sentinel-2 Imagery (2022).

144. *Lisbet Hougaard Baklid*: Geographical expansion rate of a brown bear population in Fennoscandia and the factors explaining the directional variations (2022).

145. *Victoria Persson*: Mussels in deep water with climate change: Spatial distribution of mussel (Mytilus galloprovincialis) growth offshore in the French Mediterranean with respect to climate change scenario RCP 8.5 Long Term and Integrated Multi-Trophic Aquaculture (IMTA) using Dynamic Energy Budget (DEB) modelling (2022).

146. *Benjamin Bernard Fabien Gérard Borgeais*: Implementing a multi-criteria GIS analysis and predictive modelling to locate Upper Palaeolithic decorated caves in the Périgord noir, France (2022).

147. *Bernat Dorado-Guerrero*: Assessing the impact of post-fire restoration interventions using spectral vegetation indices: A case study in El Bruc, Spain (2022).

148. *Ignatius Gabriel Aloysius Maria Perera*: The Influence of Natural Radon Occurrence on the Severity of the COVID-19 Pandemic in Germany: A Spatial Analysis (2022).