COnhecer **M**ais **P**a**R**a **I**ntervir **ME**lhor – COMPRIME
Mapping Municipal Level Determinants of COVID-19 Transmission in Portugal in different epidemic phases
(ID: 596685735)

Final Report

October 2020

Consortium

The project is a consortium headed by ENSP-CISP-UNL, in partnership with IGOT-CEG-UL; FF-UL; FM-UL, ACES Lisboa Central; ACES Oeste Sul e Câmara Municipal de Torres Vedras.

The Team includes multidisciplinary researchers and professionals from distinct areas of public health, social sciences and statistics: Paulo Sousa (Coordinator); Adalberto Campos Fernandes; Ana Umbelino; André Alves; Andreia Leite; António Martins; Carlos Freitas; Eduarda Marques da Costa; Filipa Duarte Ramos; Guilherme Romana; Jorge Rocha; Nelson Mileu; Nuno Marques da Costa; Rogério Gaspar; Rui Portugal; Sónia Dias; Vasco Ricoca Peixoto.

Index

1. Background

The pandemic associated with the spread of SARS-CoV2 represents a global threat to humanity, and economic and social security. The pandemic shows the importance of obtaining and analysing, credible, robust and timely information, available to support studies and decision processes in different levels – public health measures, at the national, regional and local level using available epidemiological data from implemented surveillance systems. The project **CO**nhecer **M**ais **P**a**R**a **I**ntervir **ME**lhor – COMPRIME (ID: 596685735) is one of the Projects of fast implementation and innovative solutions approved under the FCT call Research 4 COVID-19.

Recently it has been recommend that epidemiological studies that consider multi-level investigations of reliable and representative environmental, societal, and population determinants.[1]

Environmental and social epidemiology studies are now required to generate stronger hypothesis and infer association or causality by adequately controlling for the determining factors of COVID-19 transmission. Behavioural, societal, and community factors and control measures, socioeconomic factors and the effects of population mixing, multiple environmental determinants, and the use of appropriate spatial and temporal resolution and time frames, need to be carefully investigated. Importantly, current and future epidemiological studies should account for the differences, and varying accuracy, in the COVID-19 the timing of, or delay in reporting, the evolutionary phases of the pandemic, and the differences in data availability between and within regions, countries, and communities, and with time.[1]

The community level determinants of COVID-19 Transmission are still uncertain. They vary in different contexts, regions and moments in time. Very few research has attempted to identify municipality level determinant of transmission while adjusting for confounding or using innovative methods such as Neural Network Analysis. One study mapped community-level determinants of COVID-19 transmission in nursing homes in the USA and found that municipality-level factors like per-capita income, average household size, population density, and minority composition were significant predictors of COVID-19 cases in the nursing home[2].

The Social determinants of Health are interrelated and played major role during COVID-19 pandemic. Education level of an individual can impact his or her occupation, which determines economic stability and income level, which can impact the type of healthcare and health seeking behaviour and what neighbourhood the individual lives in, which then impacts the social and community context the individual is surrounded by and those factors played important role in current COVID-19 transmission dynamics.[3] However this intricate network makes adjustment for confounding difficult and must be considered with caution.

Pandemics are often, more of a social problem than a healthcare problem. The population that lives in poverty and in neighbourhoods that are overcrowded with poor maintenance and sanitation is being disproportionately affected by COVID-19. Lower income has been associated with poor dietary intake and habits. Minority groups, are at a disadvantage due to individual and structural discrimination.

In this research, we aim to explore which municipal level factors are associated with higher number of reported cases in different pandemic moments with different public health measures and population behaviour to raise hypothesis on community drivers, that promote targeted and context sensitive intervention and further research as necessary.

2. Objectives/aims

The project **CO**nhecer **M**ais **P**a**R**a **I**ntervir **ME**lhor - COMPRIME, aims to identifying the dynamics of the spread of the SARS Cov2 virus, in its relation with the demographic and socioeconomic profiles of territories, at the county/municipality level, identifying the determinants of different incidences in different moments in time. The projects also aims to facilitate the visualization of such results in time and space in an publicly accessible online dashboard.

3. Methodology

The work was developed in six stages.

The first stage corresponds to the statistical and cartographic representation of COVID-19 cases epidemiological data, available in two dashboard provided in Portuguese and English. The dashboards are organized in four levels, representing four scales of geographical analysis and include COVID-19 cases and Municipal level Indicators:

Main indicators and data sources:

COVID-19 reported Cases:

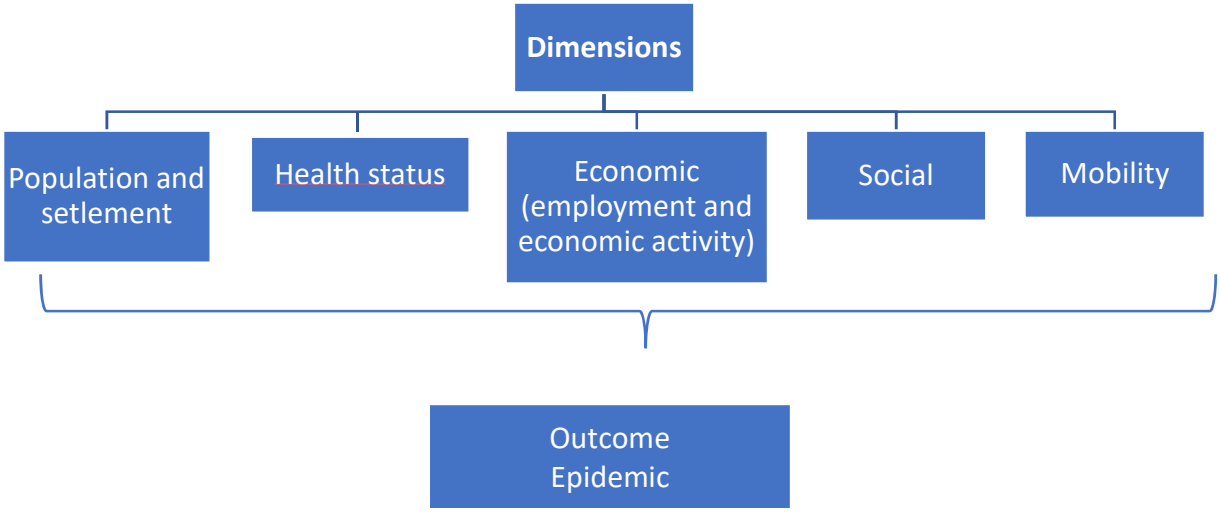
- World data, using WHO COVID case data and the "Our world in data" information (elaborated with the European Centre for Disease Prevention and Control (ECDC) data and available in https://ourworldindata.org/coronavirus);
- National data;(source: DGS[4])
- Regional data;(Source: DGS[4])
- Municipal data (Source: DGS[4])

Municipal level Indicators:

- Municipal level indicators, publicly available from National Institute of Statistics with main a selection of community level determinants based on relevant literature. (Source: INE, Instituto Nacional de Estatísitica[5])
- Primary Care Health Centres, Clinical Diagnosis data: ICPC-2 comorbidities prevalence by Municipality - Source: ACSS (by request)

The second stage corresponds to the categorizations of indicators in distinct dimensions. We considered five dimensions of analysis in a total of 60 indicators (see Annex1): population and settlement; diseases; economic; social; mobility. Considering relevant literature regarding social determinant of health and of infectious respiratory disease transmission, these indicators were selected as exposures at the municipal level to explore association with four outcome indicators (number of cases; number of new daily cases; number of cases/10000 inhabitants and the number of new daily cases/10000 inhabitants).

Figure 1 – Dimensions of Analysis



Source: authors elaboration

The 3rd step corresponds to the "key moments" selection. The identification of these moments is crucial to identify the relation between the selected dimensions and the spatial diffusion of the pandemic. The initially defined moments correspond to the implementation of national restriction measures in Portugal. The dates defining the beginning and end of each period were selected considering main changes in groups of public health measures in place considering government decisions and resolutions/ legal documents. Different periods can be analysed with similar methods in the future to explore municipal level determinant of transmission for any time period. The Project will continue to run analysis on future periods considering relevant changes in public health measures, enforcement and people´s compliance that will be available on the website dashboard.

Table 1 – Important dates during March to June and main changes in Public Health Measures[6]

Important dates	Territorial impact	Description
12-03-2020	National	Closing Universities
16-03-2020	National	Closing schools
18-03-2020	National	State of Emergency announcement
22-03-2020	National	State of Emergency start/Lockdown
28-04-2020	National	State of Emergency end/end of Lockdown

04-05-2020	National	1st phase of post-lockdown (re-start of small enterprises of trade and services)
18-05-2020	National	2nd phase of post-lockdown (re-start of schools of 11º and 12º level)
01-06-2020	National, with exception of Lisbon Metropolitan Area	3rd phase of post-lockdown
15-06-2020	Lisbon Metropolitan Area	3rd phase of post-lockdown
01-07-2020	National	Frontiers opening

To choose the moments of analysis, we took in consideration two aspects: the date of publication of the national orientations and the maximum number of cases (determined by the 3 days moving average) that occurred in the following two weeks, to select the moment of analysis: March 23, corresponds to the 1st day when the information was published per county (lockdown phase); 2nd moment, on May 28th - to measure the effects of the 1st phase of the gradual resumption of activities; the 3rd moment, 8 June - intends to evaluate the effects of the 2nd phase of the gradual resumption of activities; the 4th moment- 27 June, after the 3rd phase the gradual resumption of activities.

Table 2 – Selected moments for the statistical analysis

Moment of analysis selection
23/03/2020
28/05/2020
08/06/2020
27/06/2020

The following two steps corresponds to the application of a linear model (multiple regression analysis) and a nonlinear model (Artificial Neural Networks) in the four selected moments of analysis.

The multiple linear regression analysis, is used to explore the relationship between one continuous dependent/outcome variable (daily number of cases) and multiple independent/explanatory variables. With the analysis, we identify the strength of association that the independent variables have on a dependent variable while

adjusting for confounding. For the study, we assume that the outcome was the “Daily Number of cases” (representing the epidemiological dimension) in the four-selected moment of analysis and the independent one corresponds to the other 60 indicators distributed by the five dimensions (population, dimension, demographics and settlement), diseases; economic; social; mobility). The variable presented in the final Linear Regressions were selected by backward elimination until all variables had a p-value<0.05.

The Artificial Neural Networks (ANN) are seen as nonlinear parametric models. This methodology has the advantage of implicitly detecting any non-linear relationship between the outcome variable and the explanatory ones. AAN have been used to identify risk factor for different health outcomes, including for COVID-19 reported incidence at county/ municipality level and can be compared to more traditional methods to assess consistency in complex causality networks[7-9]. The logistic regression model can also be used to model a non-linear relationship between the response variable and the explanatory variables; however, this non-linear relationship has to be explained by the user.

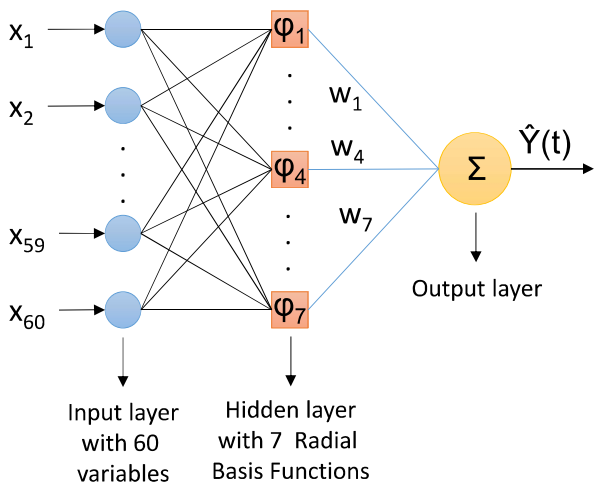
Due to the fact that there is no need for independence and normality of the variables under study, as well as their great capacity for learning from the environment, the application of ANN in the statistical analysis of epidemiological data is attractive. In addition, neural processing is able to extract relationships from input variables directly over the high-dimensional spaces that typically characterize them, making such processing a valuable tool in complex pattern recognition problems.

The two most used ANN architectures are Multilayer Perceptron (MLP) and Radial Base Function (RBF). We used the later because usually the RBF network required less time to reach the end of training compared to MLP. The training of the RBF model was terminated once the calculated error reached 0.01 or 500 training iterations were completed. The selected nonlinear approximation implemented has seven neurons within a hidden layer (see figure) and can be described as:

$$y_{t+h} = \sum_{i=0}^{7} w_i \varphi(\|Z_t - C_i\|)$$

Where $\varphi(.)$ is the RBF usually defined as a Gaussian function, e.g. $\varphi(x) = \exp(-x^2/2\sigma^2)$ and C_i is the centre of the ith RBF. The least squares solution for the weights w_i satisfies the relation $(A^TA)w = A^Tb$ where A is the matrix with elements A_{ij} representing the output element of the j-neuron for the ith input. On the present, one used a classical k-means as clustering method.

Figure 2 – Neuronal Clustering Method



Source: author's elaboration

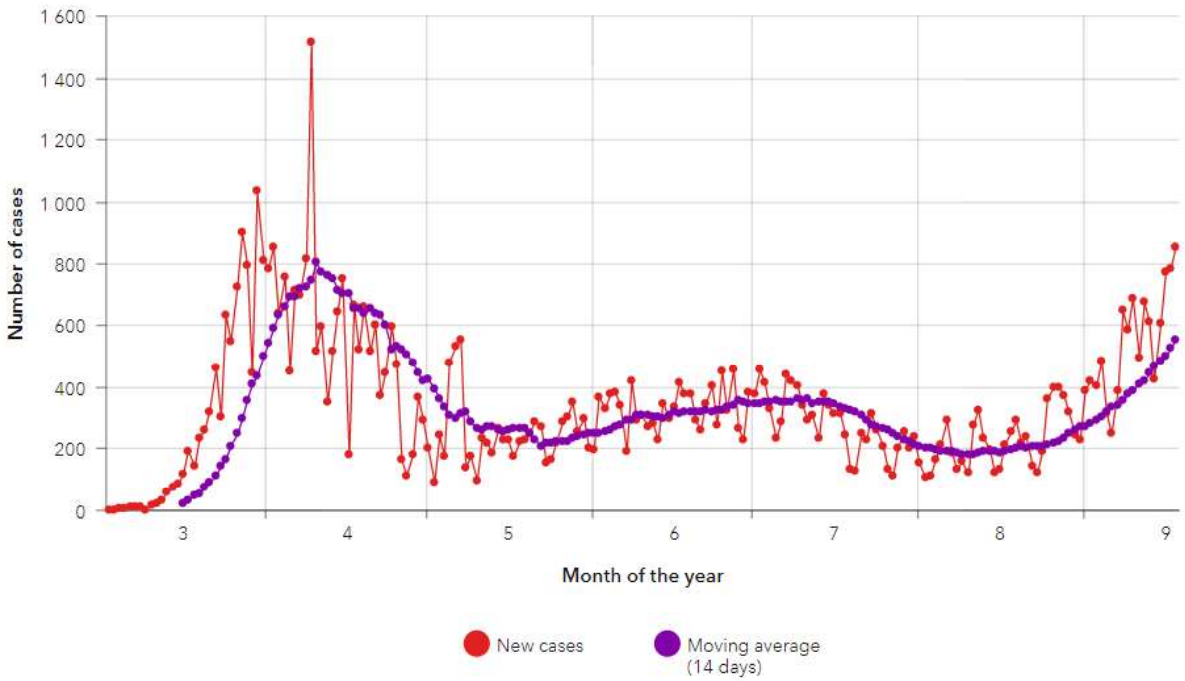
4. Results and discussion

The main results are grouped in 3 components:

4.1. Time series and cartographic representation of the COVID-19 reported cases in different moments of analysis (available in the dashboard (https://www.comprime-compri-mov.com/dashboards.html/ https://www.comprime-compri-mov.com/dashboardsenglish.html?lang=en)

The pattern of regional evolution started in metropolitan areas, but quickly extended to the municipalities of the non-metropolitan coast north of Lisbon and the northern and central coastal territories

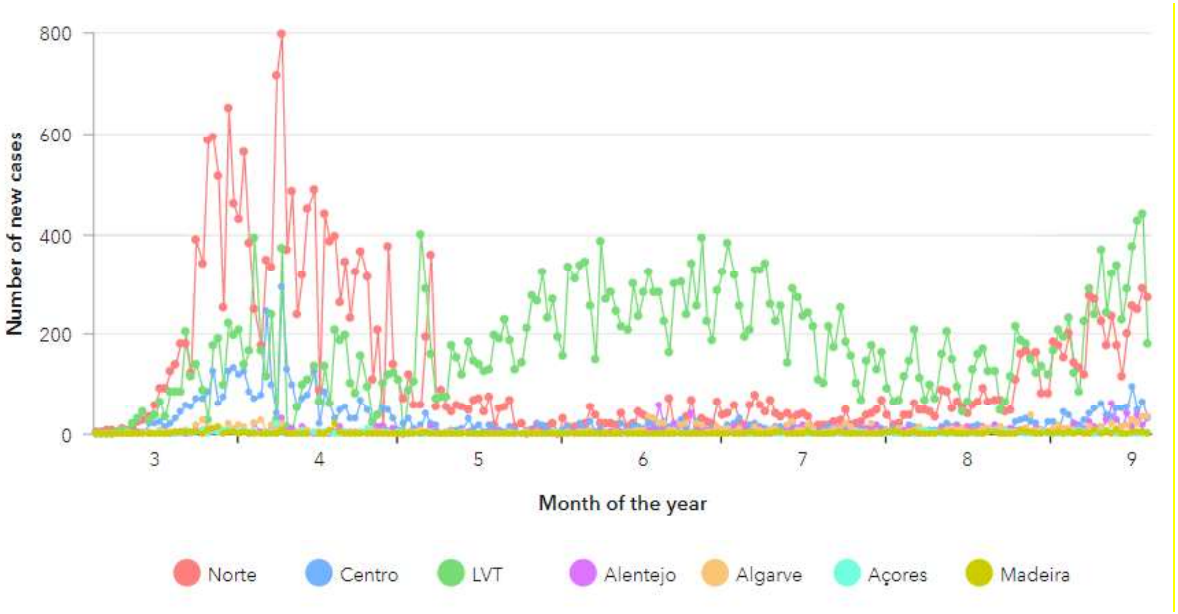
Figure 3 - Evolution of new confirmed cases, until 20th September



Source: DGS data, own elaboration available in https://comprime.weebly.com/dashboards.html

As main preliminary conclusions we highlight. The pattern of regional evolution started in metropolitan areas, but it quickly extended to the municipalities of the non-metropolitan coast north of Lisbon and the northern and central coastal territories.

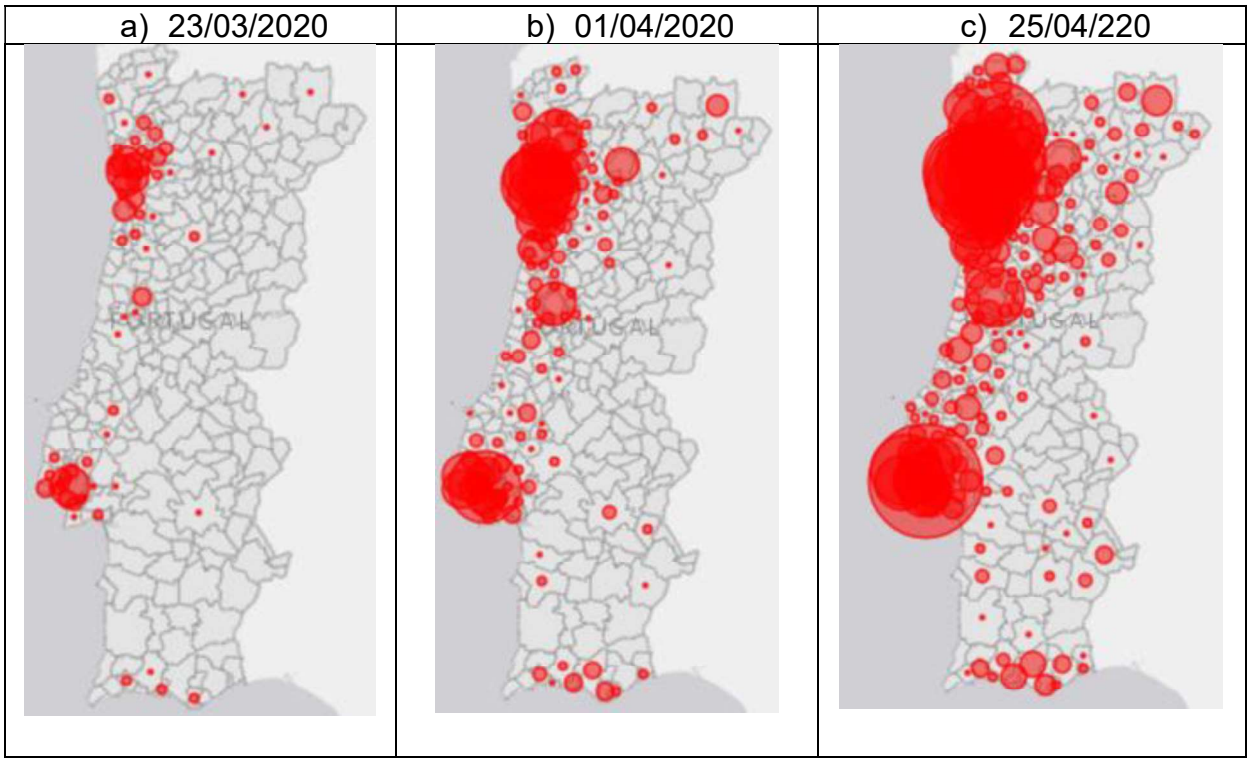
Figure 4 - Evolution of confirmed cases by region per 10 thousand inhabitants, until 20th September



Source: DGS data, own elaboration available in https://comprime.weebly.com/dashboards.html

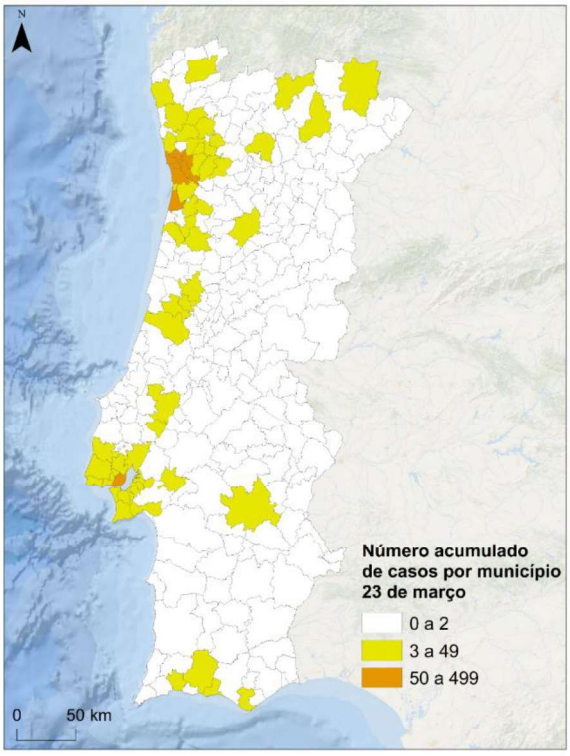
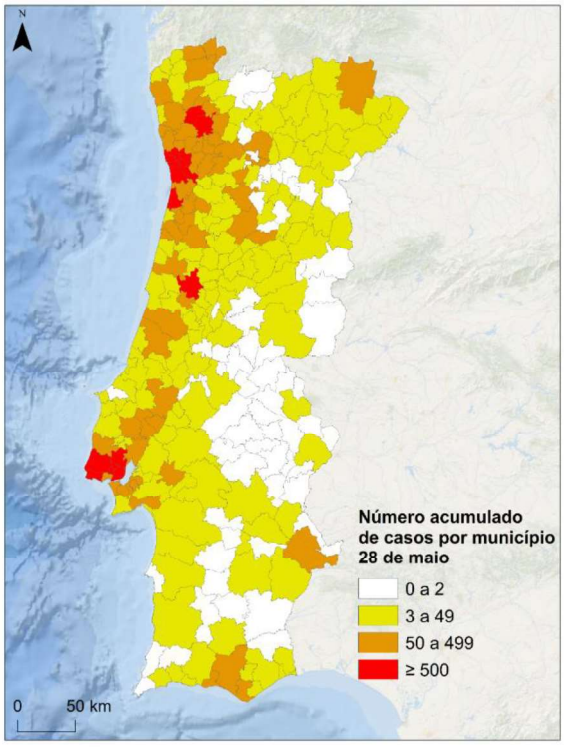
The Alentejo and the Algarve has registered, only in the most recent period, higher incidence of the phenomenon.

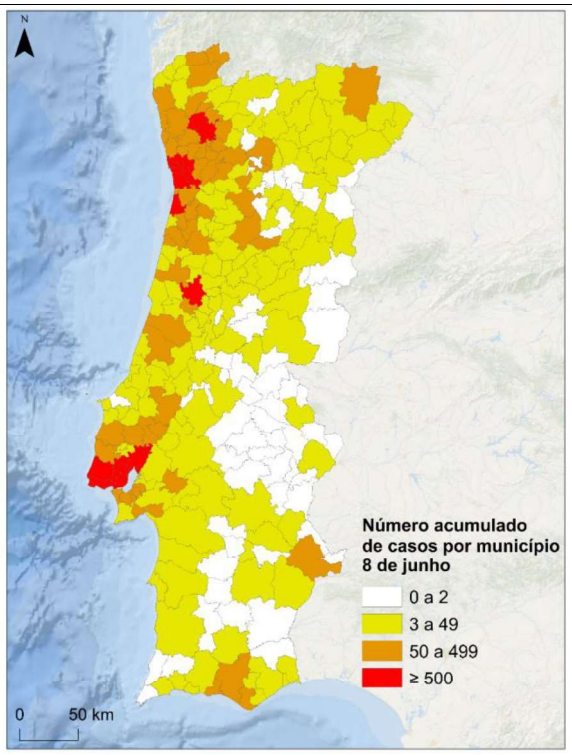
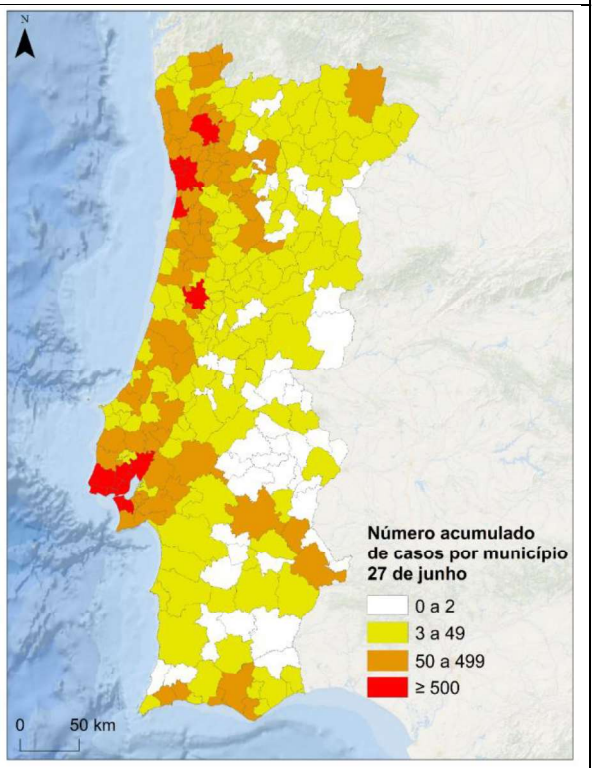
Figure 5 - Evolution of cases by municipalities in the first 3 months



Source: DGS data: own elaboration https://comprime.weebly.com/dashboards.html

Figure 6. Number of cumulative cases by municipality in different moments in time selected for analysis of municipal level associated factors

23/03/2020	28/05/2020
	
Source: DGS	Source: DGS

08/06/2020	27/09/2020
	
Source: DGS	Source: DGS

4.2. Results of linear regressions

The linear regression shows an increasing level of adjustment with the pandemic geographical diffusion. This means that explanatory variables associated are better explaining the number of cases as we move forward in time period.

Table 3 – Results of linear multiple regression
R, R^2 and the R^2 adjusted

23/march	28/may	08/june	27/june
R =0,855267 R^2= 0,731481 R^2 adjusted = 0,721349	R = 0,937844 R^2 = 0,879551 R^2 adjusted = 0,873575	R = 0,954522 R^2 = 0,911113 R^2 adjusted = 0,907057	R = 0,961773 R^2 = 0,925006 R^2 adjusted = 0,922759

Source: own elaboration with DGS (2020), Situation reports (https://covid19.min-saude.pt/relatorio-de-situacao/) and INE (2011) RGP, INE 2018, Statistical yearbook

Explanatory variables with stronger association varied in different moments of analysis.

Linear Regression – Moment 1 –23/03/2020

$Cases\ 23-03 =$ *12,006424* + *0,000165 x* ***Resident Population*** *2018* + *0,093672 x* ***Rate of Higher Education*** - *0,015045 x* ***Secondary Education students per 1000*** - *0,161005 x* % ***Working in parish of same Municipality*** + *308,554602 x* ***Exportations*** + *0,002468 x* ***Population Density*** + *2,178256 x* ***Lodging and Restaurants*** + *0,382329 x* ***patients GCDV 19*** + *0,001179 x* ***Overnights (hotels etc)*** + *0,148024 x* ***% Working outside the municipality***

R = 0,855267 R^2 = 0,731481 R^2 adjusted = 0,721349

In the first moment, Factors More Strongly Associated with higher number of cases were larger resident population higher population density, higher exportations, higher rate of higher education, higher proportions of Patients with a registry of Mental Health or substance abuse disease in Primary Care Registries. The first two variables are obvious and fundamental for adjustment since naturally a higher number of residents will explain a large part of the difference in number of cases in different municipalities. This analysis suggests that in the first moments of transmission within the country. Higher exportations, Lodging and Restaurants and Overnights may be associated because they may also imply stronger initial importation of cases in certain areas that may have gone undetected in earlier phases giving rise to initially undetectable transmission chains. Working outside the municipality was also significantly associated.

Linear Regression – Moment 2 –28/05/2020

$Casos\ 28-05 =$ - 66,864739 + 0,004350 x **Resident Population** 2018 -0,639928 x **Renting Price** + 0,021974 **Population Density** - 0,203452 *x* ***Secondary Education p 1000*** + *3,551664 x* ***Patients with GCDV 19*** + *33,382222* ***Net Migration Rate*** + *0,015484 x* **Overnights** + 1,608499 *x* ***Working outside the municipality***- 29,158 *x*

Construction + 6,380067 *x* **Age 25-65 +** 35,707433 ***x*** **Restaurants** + 2,438806 *x* **Beverages Industry** + 5,220509 **Electric Equipment production**

R = 0,937844 R^2 = 0,879551 R^2 adjusted = 0,873575

In Moment 2, after the second phase of lockdown lifting of restriction took place in the beginning of May, different variables came as most strongly associated. However, this was a period with low circulation and number of cases was mostly related to Population and population density. It is hard to discuss how other identified variables as came about as strong associations and limitations of the study in terms of confounding are to be considered. New factors that become significant were Net Migration Rate, Construction Coefficient, larger younger population among other.

Linear Regression – Moment 3 –08/06/2020

$Casos\ 08 - 06 =$ *208,957877* + *0,004969* *x* **Resident Population** *2018* *-0,668905* *x* **Renting Price** + 0,047487 *x* **Population Density** + 62,057238 *x* ***Net Migration Rate*** - 22,817706 *x* **Income** - 19,512999 *x* **Construction** + 39,244546 *x* **Restaurants** + 4,649448 *x* **Electric Equipment production** + 1337,200038 *x* **Exportations** + 12,615446 *x* Patients with **GCD X** + 2,238563 + 2,238563 *x* **Beverages Industry** - 26,843313 *x* **Education**

R = 0,954522 R^2 = 0,911113 R^2 adjusted = 0,907057

In moment 3 besides Population and population density, higher Net Migration Rate was positively associated with more cases reported, this is compatible with higher transmission detected in some high migration neighbourhoods and municipalities in Lisbon Suburbs. In one situation disease prevalence from Primary Health Care Registries may be importantly influenced by registry culture and practices and it may be spuriously associated because in some urban centres diseases are more often registered in informatics systems. However, GCD X (respiratory diseases) seem to be positively associated with higher cases counts even though this association could be spurious.

In this moment, a higher income was associated with lower number of cases. This is consistent with municipalities with higher migrant population, more rowed housing, higher population density and more work outside the municipality of residency. The Lisbon suburbs were important outliers in transmission during this period being responsible for a large part of new cases in Portugal.

Linear Regression – Moment 4 –27/06/2020

$Casos\ 27-06 =$ *313,168506 + 0 0,006356 x* **Resident Population***2018 + 0,064725 x* **Population Density***- 1,816541 x* ***% Working outside the municipality****- 0,688449 x* **Renting Price** *+ 80,881422 x* ***Net Migration Rate*** *- 24,124363 x* ***Income*** *- 36,672949 x* ***Education*** *- 24,124363 x* ***Food Industries***

R = 0,961773 R^2 = 0,925006 R^2 adjusted = 0,922759

In moment 4, after specify measures were applied to Lisbon Metropolitan Region July 15 factors more strongly associated with higher case counts were population, population density, lower income and higher net migration rate and lower education level.

In the 4 conducted linear regression analysis we show the relative importance of the variables by dimension and in each moment: In general (any of the analyzed moments) , from initially introduced variables we have that in any moment:

- population and settlement: from 9 indicators, there are 4 with significance;
- diseases: from 9 there are 2 with significance;
- economic: from 23 there are 11;
- social: from 15 there are 5;
- mobility: from 4 there are 2

A Summary Table of results of Linear Regressions for the four moments are found in Annex 1.

4.3. Nonlinear model of Neuronal Network Analysis(NNA) Results

The prediction model adds new data, which allows to identify the counties where the verified values are above the expected values, highlighting the question of population density, which adds to the mobility of the population outside the municipality and the presence of migrant communities. In different moments, the variables that explain the variability between municipalities vary. The estimated values are above the expected ones, mainly in the metropolitan Lisbon area, including the municipalities of Lisbon, Amadora, Odivelas, Sintra, Loures.

Population density is of course a relevant explanatory variable.

Using NNA we found, variables that were significant in multiple regression. For the 1st moment variables found to be of relevance in both methods were: Population, Population Density, Exportations, Rate of Higher Education

In this 1st moment transmission was defined largely in period before any strong control measures were implemented the transmission had few opposing measures. Many possible confounders cannot be ruled out. As such, naturally, areas with more population density, mobility and contacts and with more movement to outside the country were more affected.

Variables that were relevant in both methods were Resident Population, Population Density, population working outside municipality of residency, income, net migration rate/ % illegal immigrants.

In the second moment, interestingly exportations remain relevant in the list but possibly due to remaining confounding as these correspond to metropolitan settings. Working outside the Municipality also remained a strong explanatory variable but can be confounded by other aspects of urban metropolitan mobility. We cannot rule out a relevance of public transportation mobility.

A lower income is also associated with poorer and more crowed living conditions, different community dynamics of cooperation within neighbourhoods, many time

having job that cannot work from home and possibly in some contexts lower health literacy and compliance with measures.

In the 3rd moment NNA showed Social habitation, working outside the municipality and income as relevant factors.

Table 4. Results of NNA in the 4th analysed moments after the 3rd phase of lifting lockdown restrictions.

June 27

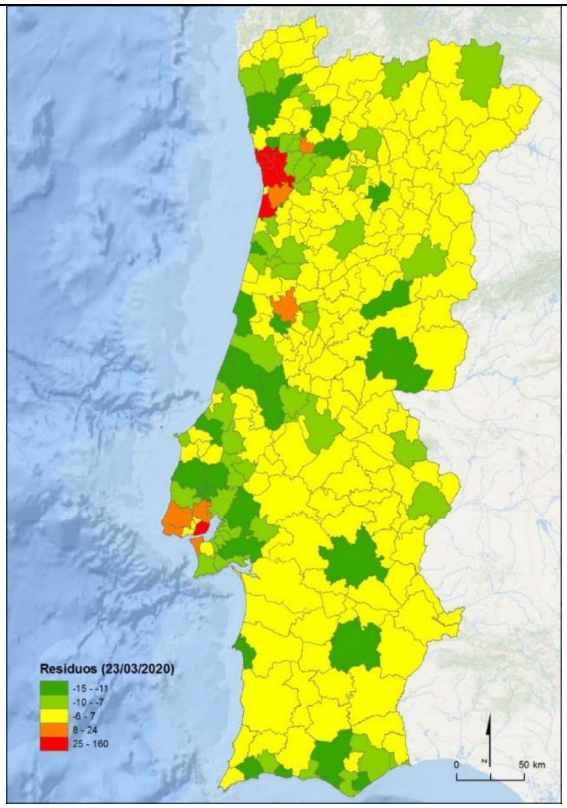
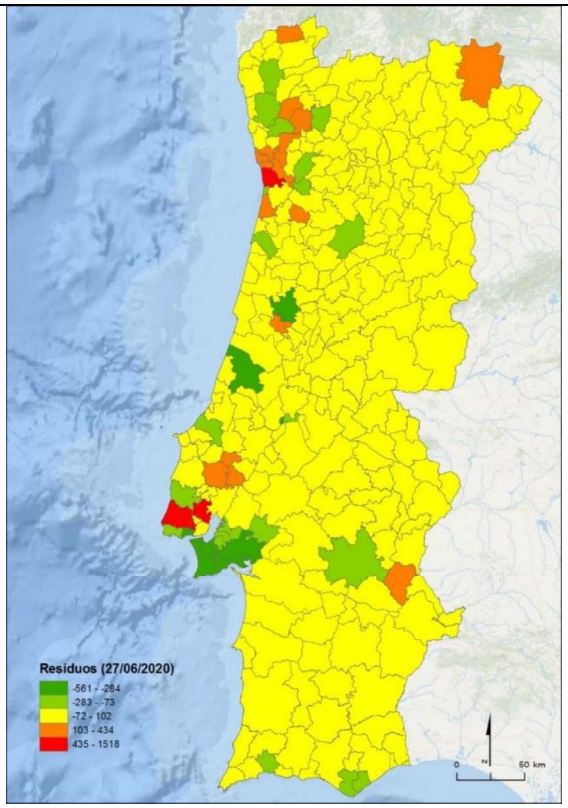
	Importance	Normalized importance
GCDI19	0,036	100,0%
Population density	0,036	99,9%
Resident population	0,032	89,9%
Buildings without needs of repair	0,031	87,5%
Buildings in need of minor repair	0,027	76,0%
Work Outside Municipality	0,027	76,0%
Buildings in need of repair	0,025	70,3%
Buildings in need of medium repairs	0,023	63,6%
Production services	0,023	62,9%
Income	0,022	60,7%
Higher Education students per 10000 inhabitants	0,021	60,0%
Percentage of population using car	0,020	56,8%
Buildings in need of major repairs	0,020	54,9%
1st cycle of basic education students per 10000 inhabitants	0,018	51,5%
Percentage of legal immigrants	0,018	50,9%
Exports	0,018	50,9%
Urbanization rate	0,017	47,1%
GCDIII19	0,017	47,0%

In the 4th moment, besides Population and Population Density, working outside the municipality, production services, Income, number of higher education students per capita and the proportion of legal immigrants, exportations and urbanization were among factors with more relevance in the NNA.

Full results of NNA for each moment can be found in Annex2.

Figure 7 - Predictive model based on artificial neural network – estimated/expected values showing lowest and highest values

(green color represents the number of cases lowest than predicted;
red color represents the number of cases highest than predicted)

March 23rd	June 27th
	

Source: Elaborated from DGS (2020), Situation reports (https://covid19.min-saude.pt/relatorio-de-situacao/) and INE (2011) RGP, INE 2018, Statistical yearbook

5 – Reflections, Findings and Future Steps

The preliminary results of the project **CO**nhecer **M**ais **P**a**R**a **I**ntervir **ME**lhor - COMPRIME, attempt to demonstrate relations between the dynamics of the spread of the SARS Cov-2 virus with the demographic and socioeconomic profiles of territories, at the municipality level, identifying potential determinants of transmission. The knowledge resulting from this study shows the importance of context specific public health measures and their adaptation to territorial characteristics, allowing also to support the preparedness and response strategy to new pandemic waves.

This study also demonstrates how complex social and municipal level interactions can be when to comes to associations with outcomes which warrants more research, different approaches and methods. The results and methods of this project will allow the team to refine the analysis plan to answer specific questions related to broader generated hypothesis in this phase.

The diversity of the team allowed for innovative methods and approaches and a richer interpretation and discussion of findings using indicators that come from different information systems. Knowledge of these was relevant for interpreting complex findings and identifying limitations.

This study raises awareness for relevant hypothesis in terms of drivers of transmission in different moments in time , using municipal level exposures and outcomes.

The methodological approach to COVI-19 transmission risk factors is innovative for two main reasons:

1. Exploring risk factors for transmission at the community/municipality level
2. Using traditional (linear regression) and more recent methodological approaches(Neural Network Analysis) to explore risk factors for an health outcome and address consistency of results.

Although difficult to draw strong conclusions from explorative methods, some hypothesis were generated and can be considered when planning measures that should maximize prevention with minimal external effects on society functioning.

Besides population and population density different factors have standed out in association with higher number of cases in linear regressions and in NNA in more than one moment.

Relevant risk factors standing out in the analysis:

- Municipalities with higher percentage of residents working in other Municipality/parish: This was a relevant factor in moments 1, 2 and 4 in regression analysis and in NNA in moments 3 and 4. This should promote further research in the role of these variable and other associated factors due to:
    - Possible importance of transmission in longer and more crowded public transportation contexts for those working in another municipality
    - The existence of other confounding's related to behaviour and conditions of life of those living in suburban areas.

- Municipalities with higher net migration rate/other high migrant population proxy**:** net migration rate can be considered a proxy of proportion of migrant population and population density. This was a relevant factor in moments 2, 3 and 4 in linear regression and in moment 4 in NNA. More research and targeted intervention is to be considered in Municipalities with high proportions of migrant population due to:
    - Possibly relevant working, transportation, housing and behavioural/cultural determinants of infection among migrant populations with need of targeted interventions.

- Municipalities lower income per capita: Lower income has been associated with higher number of reported cases in linear regressions in moments 3 and 4 (post-lockdown). In NNA it was found to be a relevant factor in moments 3 and 4 were social housing was also a relevant factor. More research and targeted

intervention is to be considered in Municipalities and areas with more population with lower income. This is relevant due to: possible association of lower income with housing (crowded living conditions), work (work with higher risk/more contacts), mobility (more use of crowded public transportation for longer distances).

- Municipalities with higher number of higher education students per capita: This has been an interestingly relevant factor in Neural Network Analysis in moments 1 and 4 (moments corresponding to closer to normal situations). This should promote further research in the role of higher education student's due to possible:
    - Higher number of close social contact while still contacting with family.
    - Lower detection of mild cases among these groups due to milder symptoms, asymptomatic cases and health care avoidance in the presence of mild symptoms, facilitating undetected transmission were case identification, contact tracing and isolation is not possible.
    - Possible confounding due to other factors associated (more urbanization, more contacts in larger metropolitan areas etc.)

- Stronger potential implications of the preliminary interpretation of findings and of this study in the Portuguese context are:

- The need to address vulnerabilities in lower income and high immigration neighbourhoods through inclusive communication strategies and social support in terms of housing, food, work and financial protection facing COVID-19 infection.
- The need to further investigate the importance on public transportation in transmission and other ways through which working outside the municipality of residency may influence transmission.
- The need to address higher education students and eventually young adults in general in terms of targeted and group/context sensitive communication strategies and to further investigate a potentially relevant role in transmission in the Portuguese context considering types of social contacts, symptomatic presentation and under-reporting of mild and asymptomatic cases.

References

1. Zeka A, Tobias A, Leonardi G, et al. Responding to COVID-19 requires strong epidemiological evidence of environmental and societal determining factors. *Lancet Planet Heal*. 2020;4(9):e375-e376. doi:10.1016/S2542-5196(20)30169-8
2. Sugg MM, Spaulding TJ, Lane SJ, et al. Mapping community-level determinants of COVID-19 transmission in nursing homes: A multi-scale approach. *Sci Total Environ*. 2021;752:141946. doi:10.1016/j.scitotenv.2020.141946
3. Singu S, Acharya A, Challagundla K, Byrareddy SN. Impact of Social Determinants of Health on the Emerging COVID-19 Pandemic in the United States. *Front Public Heal*. 2020;8:406. doi:10.3389/fpubh.2020.00406
4. DGS. Relatório de Situação - COVID-19. https://covid19.min-saude.pt/relatorio-de-situacao/. Accessed September 29, 2020.
5. Instituto Nacional de Estatísitca. Portal do INE. https://www.ine.pt/xportal/xmain?xpgid=ine_main&xpid=INE&xlang=pt. Accessed September 29, 2020.
6. INSA. Covid-19: curva epidémica e parâmetros de transmissibilidade Categoria - INSA. http://www.insa.min-saude.pt/category/areas-de-atuacao/epidemiologia/covid-19-curva-epidemica-e-parametros-de-transmissibilidade/. Accessed September 29, 2020.
7. Chen J, Pan QS, Hong WD, et al. Use of an artificial neural network to predict risk factors of nosocomial infection in lung cancer patients. *Asian Pacific J Cancer Prev*. 2014;15(13):5349-5353. doi:10.7314/APJCP.2014.15.13.5349
8. Sherriff A, Ott J. Artificial neural networks as statistical tools in epidemiological studies: Analysis of risk factors for early infant wheeze. *Paediatr Perinat Epidemiol*. 2004;18(6):456-463. doi:10.1111/j.1365-3016.2004.00592.x
9. Mollalo A, Rivera KM, Vahedi B. Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States. *Int J Environ Res Public Health*. 2020;17(12):1-13. doi:10.3390/ijerph17124204

Annexes

Annex 1

List of indicators and Multiple Regression Results by Moment of Analysis

Dimensions	Indicators	23/03	28/05	08/06	27/06
Disease	1. GCDI19 Infectious Diseases 2019				
	2. GCDII19 Neoplasms 2019				
	3. GCDIII19 Blood disorders and other diseases of the immune system 2019				
	4. GCDIV19 Endocrine, nutritional and metabolic diseases 2019				
	5. GCDIX19 Cardiovascular diseases 2019				
	6. GCDV19 Mental illness and substance abuse 2019	x	x		
	7. GCDX19 Respiratory diseases 2019			x	
	8. GCDXI19 Digestive diseases 2019				
	9. GCDXIV19 Diseases of urinary and genital systems 2019				
Economic	10. Exports (% of the municipality over the country total) 2019	x		x	
	11. Overnights 2019	x	x		
	12. Food Industries Location Quotient 2018				x
	13. Beverage Industry Location Quotient 2018		x	x	
	14. Wholesale and retail trade; repair of motor vehicles and motorcycles Location Quotient 2018				
	15. Sale, maintenance and repair of motor vehicles and motorcycles Location Quotient 2018				
	16. Wholesale commerce Location Quotient 2018				
	17. Transport and storage employment Location Quotient 2018				
	18. Storage and similars Location Quotient 2018				
	19. Accommodation, catering and similar activities Location Quotient 2018 2018	x	x		
	20. Catering and similar activities Location Quotient 2018			x	
	21. Civil Construction Location Quotient 2018		x	x	x
	22. Agriculture, animal production, hunting, forestry and fishing Location Quotient 2018				
	23. Textile manufacturing Location Quotient 2018				
	24. Quociente de Localização da Fabricação de produtos químicos e de fibras sintéticas ou artificiais, exceto produtos farmacêuticos 2018				
	25. Manufacture of electrical equipment Location Quotient 2019		x	x	
	26. Education Location Quotient 2019			x	x
	27. Social work activities with accommodation Location Quotient 2018				
	28. Employment in Agriculture, animal production, hunting, forestry and fishing (% at national level) 2018				
	29. Employment in manufacturing (% at national level) 2018				
	30. Employment in traditional industries (% at national level) 2018				
	31. Employment in Accommodation, catering and similar (% at national level) 2018				

	32. Employment in production services (% at national level) 2018				
Dimensions	33. % Population working in the parish of residence 2011	23/03	28/05	08/06	27/06
	34. % Population working in another parish in the municipality 2011				
	35. % of population working outside the municipality of residence	x			x
	36. % of trips with the use of a car 2011	x	x		
Mobility	37. GCDI19 Infectious Diseases 2019				
	38. % Pop. Age 25 - 64 years 2019		x		
	39. % Pop. Age 65 - 74 years 2019				
	40. % Pop. Age 75 and more 2019				
	41. Aging index 2019				
	42. Migration balance rate 2019		x	x	x
	43. % Legal immigrants in the total population 2019				
	44. Resident population 2019	x	x	x	x
	45. Population density 2019	x	x	x	x
População e Povoamento	46. Urbanization rate 2013				
	47. Number of pensioners / 1000 inhabitants 2018				
	48. Gross income paid by taxable person (€) 2018 (€) 2018			x	x
	49. Nº of students in pre-school /1000 inhabitants, 2018/19				
	50. Nº of students in the 1st cycle of Basic Education / 1000 inhabitants, 2018/19				
	51. Nº of students in the 2nd cycle of Basic Education / 1000 inhabitants, 2018/19				
	52. Nº of students in the 3rd cycle of Basic Education / 1000 inhabitants, 2018/19				
	53. Nº of students in with Secondary Education	x	x		
	54. Nº of students in Higher Education 2018/2019				
	55. Basic education rate 2011				
	56. Higher education rate 2011	x			
	57. Medium Rental value 2011		x	x	x
	58. % population with socially valued professions 2011				
	59. Average household size 2011				
	60. Buildings in need of repairs in total 2011 (%)				
Social	61. Social housing buildings in total 2011 (%)				
Number of independent variables from the model		*10*	*13*	*12*	*8*
Epidemic (dependent) - Nº of cases		x	x	x	x
		R =0,855267 R^2= 0,731481 R^2 ajustado= 0,721349	R =0,937844 R^2 = 0,879551 R^2 ajustado = 0,873575	R = 0,954522 R^2 = 0,911113 R^2 ajustado = 0,907057	R = 0,961773 R^2 = 0,925006 R^2 ajustado = 0,922759

Annex 2

Results of Neuronal Network Analysis

March 23rd

Indicators	Importance	Standardized importance
Urbanization rate 2013	0,036	100,0%
Buildings without need of repair 2011	0,036	100,0%
Buildings in need of minor repairs 2011	0,032	89,4%
Buildings in need of repair 2011	0,031	85,7%
Buildings in need of medium repairs 2011	0,029	79,9%
Buildings that need major repairs 2011	0,028	78,0%
Resident population 2019	0,028	77,4%
Population density 2019	0,026	72,1%
Buildings in need of extreme repair 2011	0,026	71,0%
GCDI19 Infectious Diseases 2019	0,025	70,2%
Nº of students in Higher Education 2018/2019	0,025	69,4%
Higher education rate 2011	0,024	67,0%
Nº of students in 1st cycle of Basic Education / 1000 inhabitants, 2018/19	0,023	64,0%
Income per inhabitant	0,023	63,6%
% People employed in production services (compared to the total in the country) 2019	0,023	63,2%
% population with socially valued professions 2011	0,023	62,7%
% of exports of the municipality in the total of the country 2019	0,020	56,2%
% of people working in the parish of residence	0,018	50,3%

May 28th

	Importance	Standardized importance
Population density 2019	0,029	100,0%
GCDI19 Infectious Diseases2019	0,028	98,6%
Resident population 2019	0,028	97,5%
% of population that work outside the municipality of residence 2011	0,025	86,0%
Buildings without need of repair	0,023	79,1%
Gross income paid by taxable person (€) 2018 (€) 2018	0,022	76,9%
Buildings in need of minor repairs	0,021	74,7%
Higher education rate 2011	0,021	73,3%
Production services	0,021	72,2%
Social housing buildings with 2 and more household 2011	0,021	72,2%
Social housing units	0,021	71,6%
Social housing buildings with 1 household 2011	0,020	70,7%
Buildings in need of repair	0,020	70,5%
Nº of students in Higher Education 2018/2019	0,020	69,9%
% of people working in the parish of residence	0,019	66,9%
Buildings in need of medium repairs 2011	0,019	65,4%
% of exports of the municipality in the country total	0,019	65,4%
% Legal immigrants in the total population, 2018	0,019	65,0%

Results of Neuronal Network Analysis

June 8th

	Importance	Standardized importance
Social housing buildings with 2 and more household 2011	0,041	100,0%
Resident population, 2019	0,037	90,6%
Social housing units	0,035	85,8%
Buildings in need of minor repairs 2011	0,030	74,9%
Buildings in need of repair 2011	0,030	73,6%
Buildings in need of medium repairs 2011	0,029	71,0%
Buildings without need of repair 2011	0,028	69,3%
% of population working outside the municipality of residence	0,028	68,5%
Population density	0,027	67,4%
Infectious Diseases	0,026	64,4%
Buildings that need major repairs 2011	0,025	60,7%
GCDIII19	0,024	59,5%
Social housing buildings with 1 household 2011	0,023	57,8%
Buildings in need of extreme repair 2011	0,022	54,2%
Production services employment - QL	0,020	48,0%
Textile manufacturing employment - QL	0,019	46,2%
Exports	0,019	45,8%
Income	0,017	42,5%

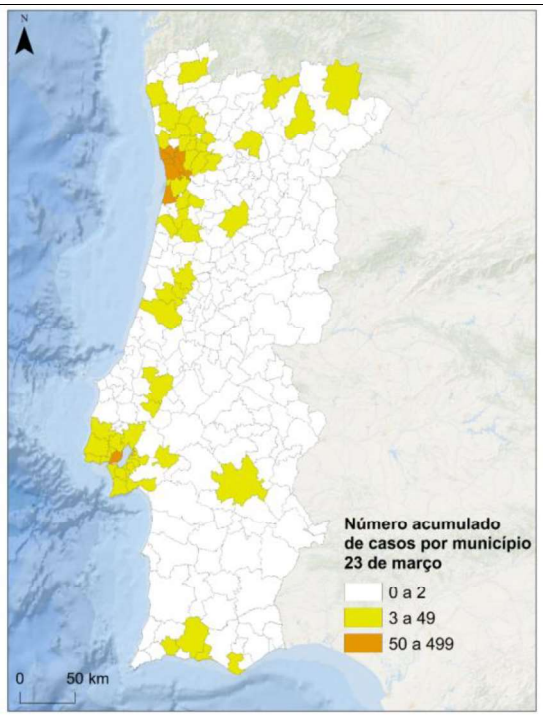
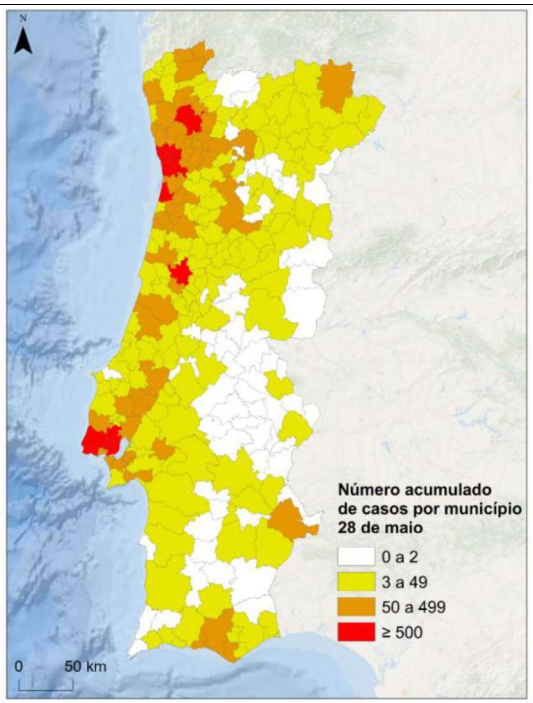
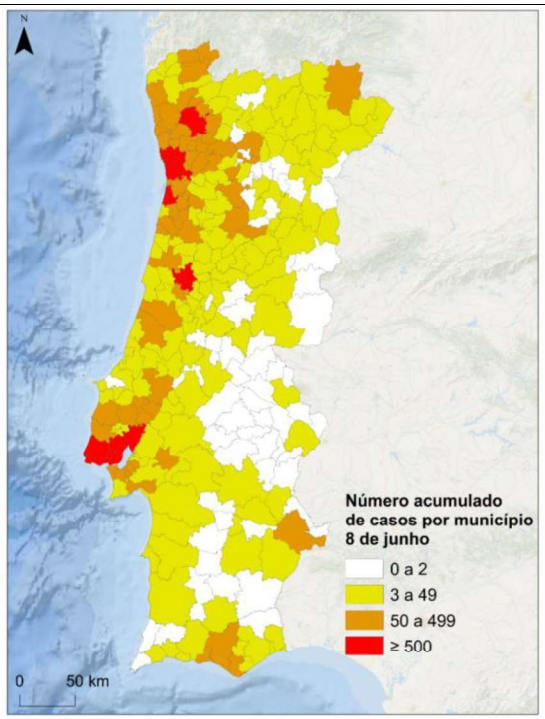
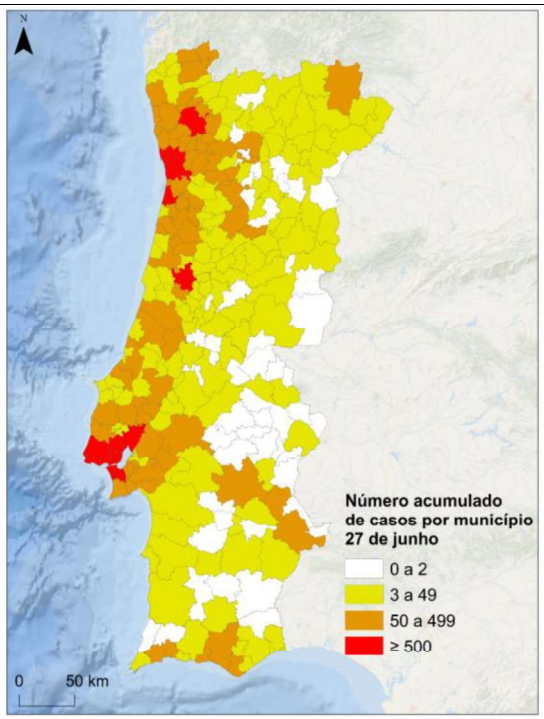
July 27th

	Importance	Standardized importance
Infectious Diseases	0,036	100,0%
Population density	0,036	99,9%
Resident population 2019	0,032	89,9%
Buildings without need of repair 2011	0,031	87,5%
Buildings in need of minor repairs 2011	0,027	76,0%
% of population working outside the municipality of residence	0,027	76,0%
Buildings in need of repair 2011	0,025	70,3%
Buildings in need of medium repairs 2011	0,023	63,6%
Production services	0,023	62,9%
Income	0,022	60,7%
Nº of students in Higher Education /1000 inhabitants, 2011	0,021	60,0%
% of population using car	0,020	56,8%
Buildings that need major repairs 2011	0,020	54,9%
Nº of students in the 1st cycle of Basic Education / 1000 inhabitants, 2018/19	0,018	51,5%
% Legal immigrants in the total population, 2018	0,018	50,9%
Exportation	0,018	50,9%
Urbanization rate	0,017	47,1%
Blood disorders and other diseases of the immune system	0,017	47,0%

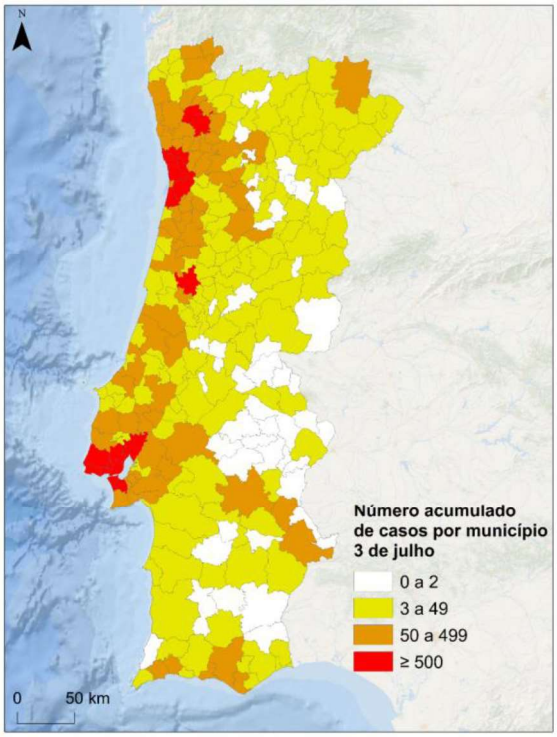
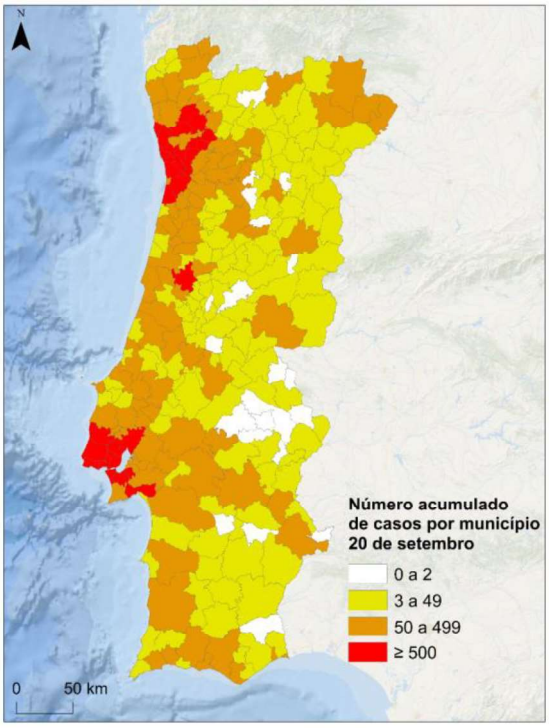
Annex 3

Atlas of Indicators by dimension of analysis

3.1. Cases by municipality

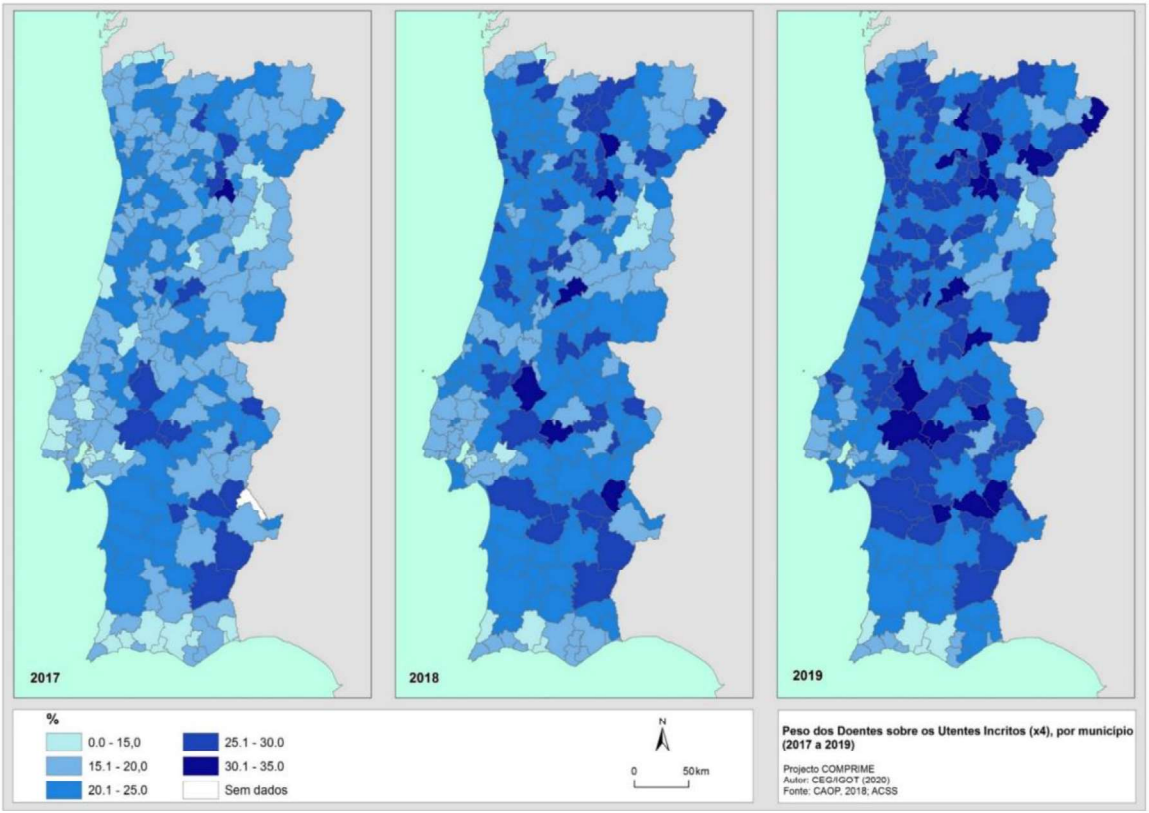
23/03/2020	28/05/2020
	
Source: DGS	Source: DGS
08/06/2020	**27/09/2020**
	
Source: DGS	Source: DGS

5.1. Cases by municipality

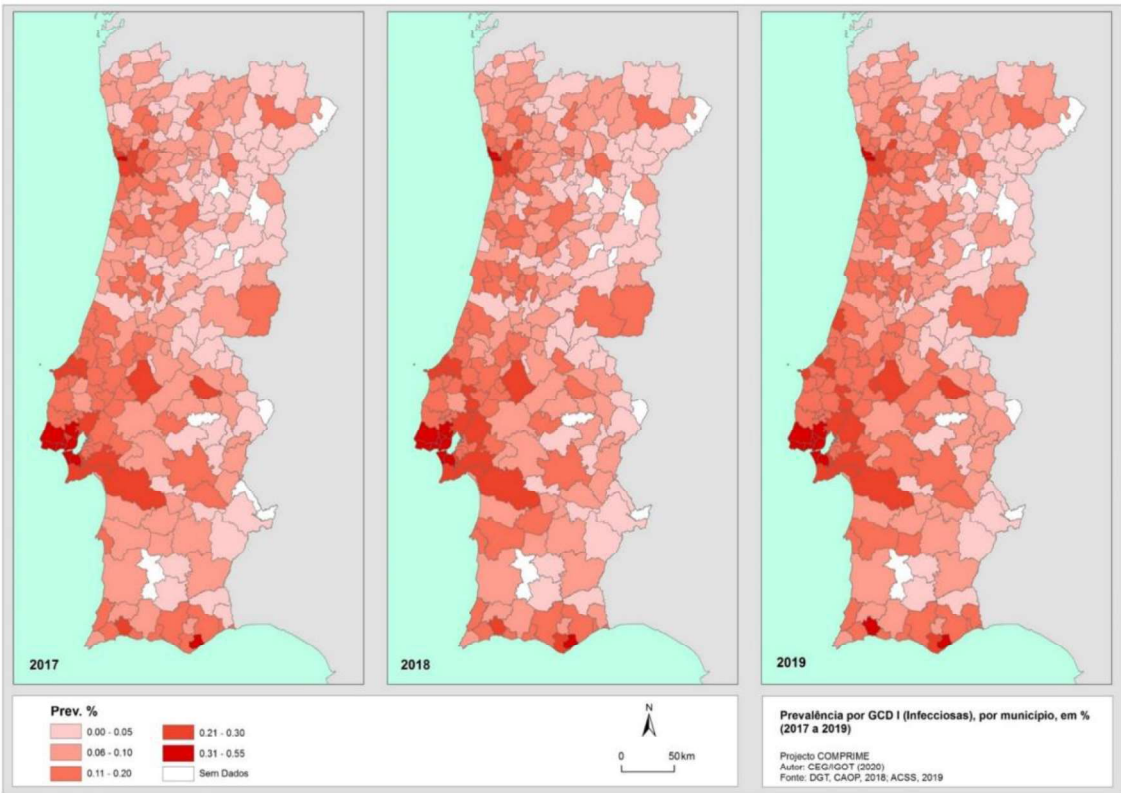
03/07/2020	20/09/2020
	
Source: DGS	Source: DGS

3.2. Diseases

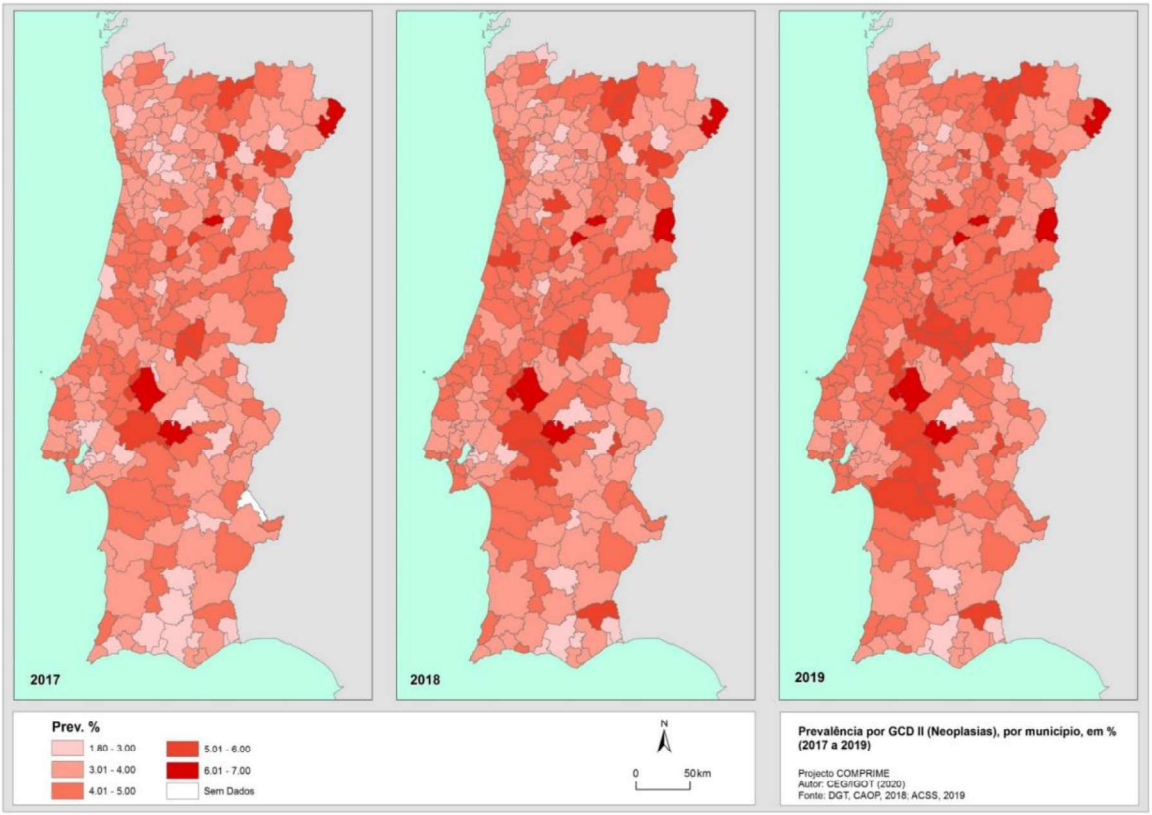
Rate of registered consultations/ total of registered population



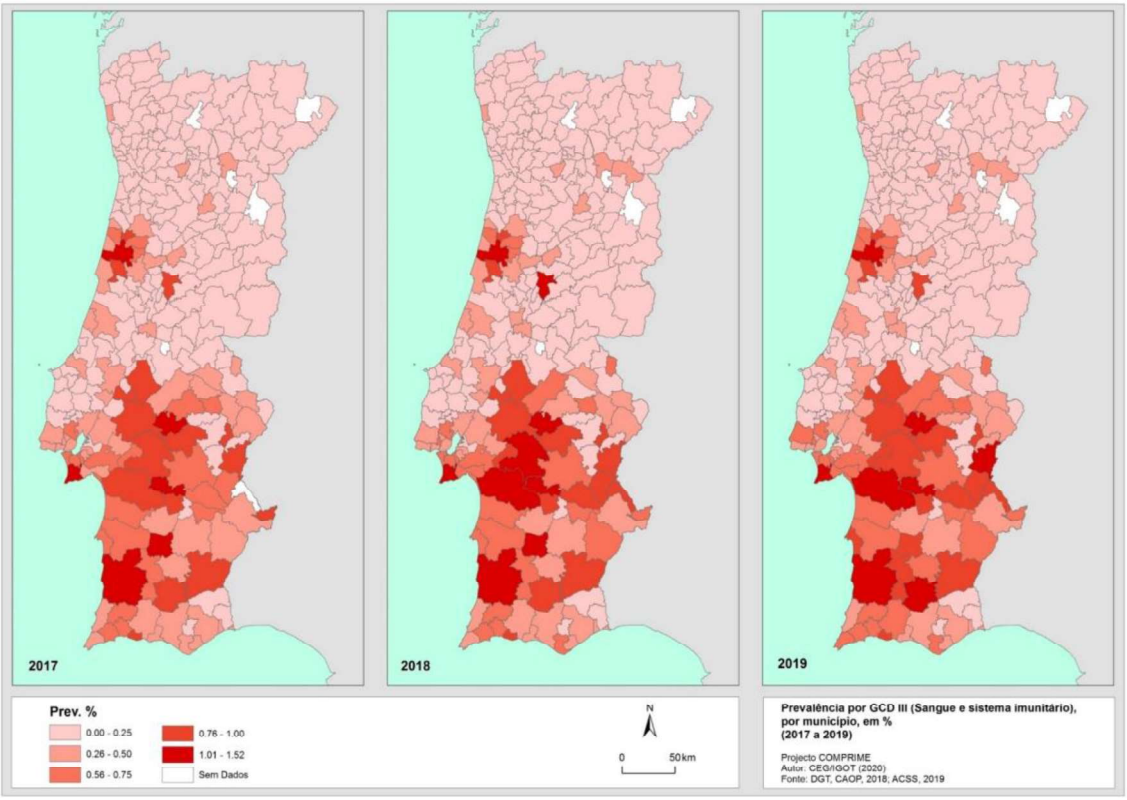
Infectious Diseases



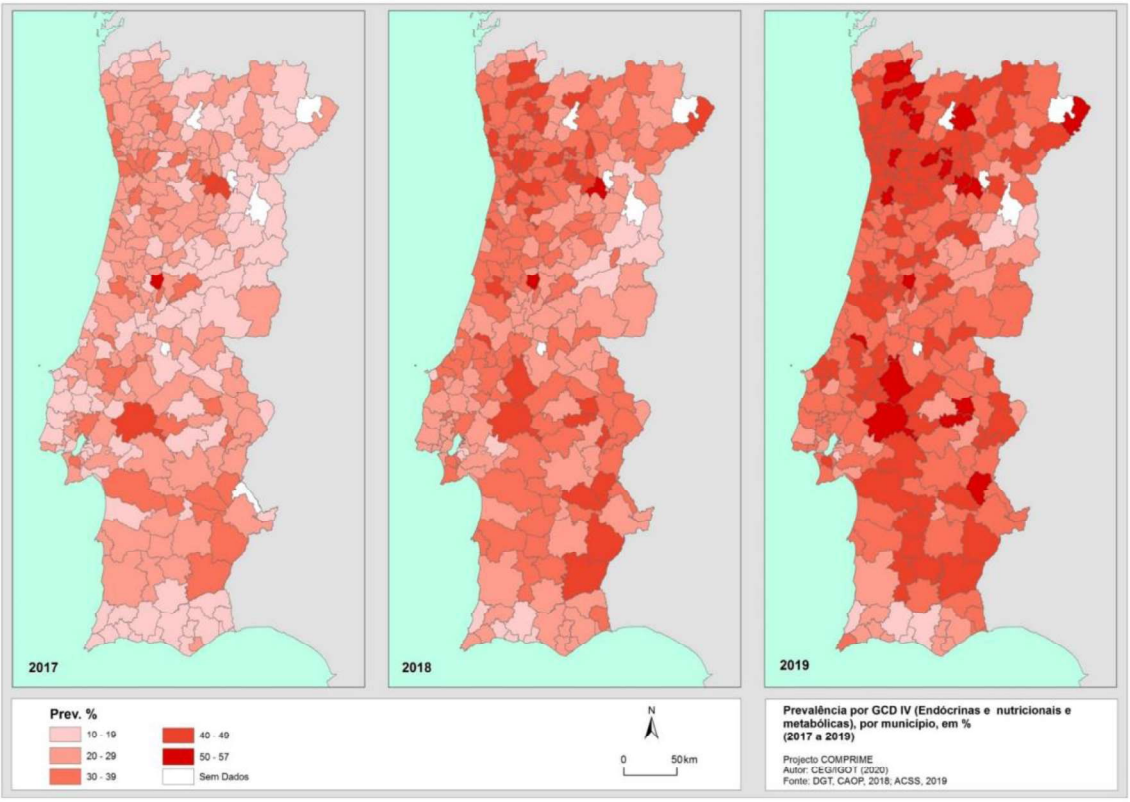
Cancer



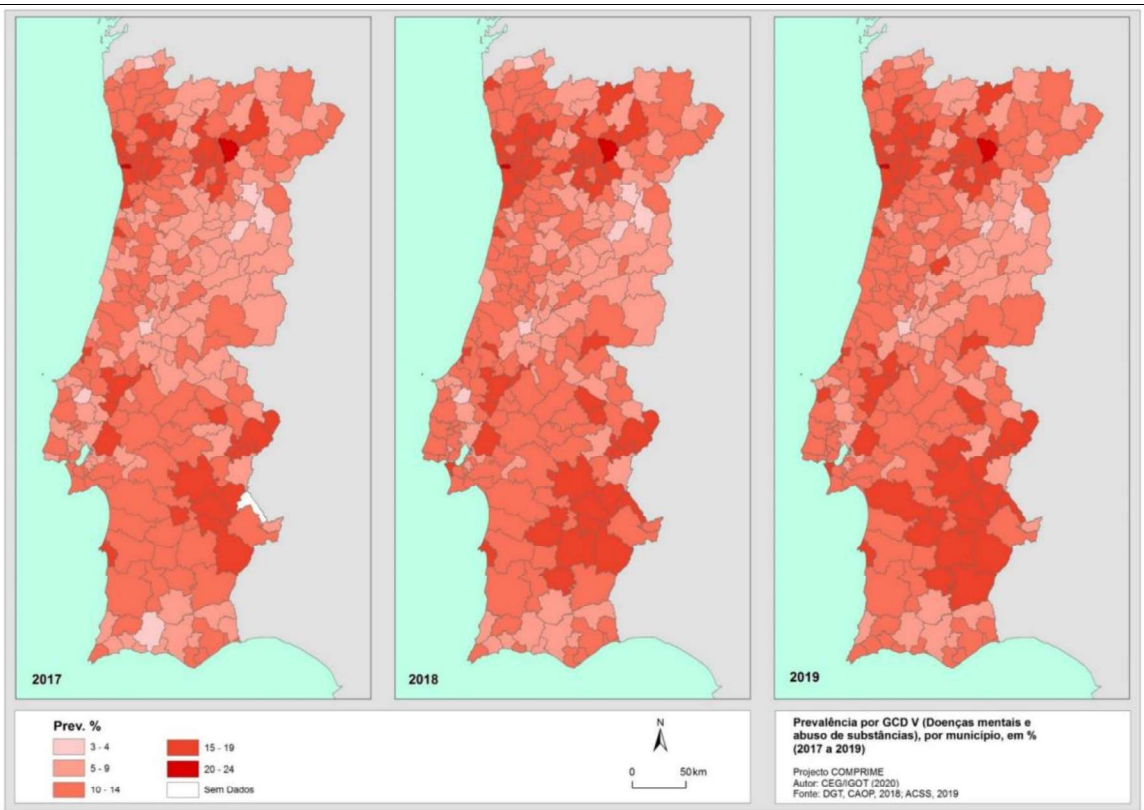
Blood disorders and other diseases of the immune system



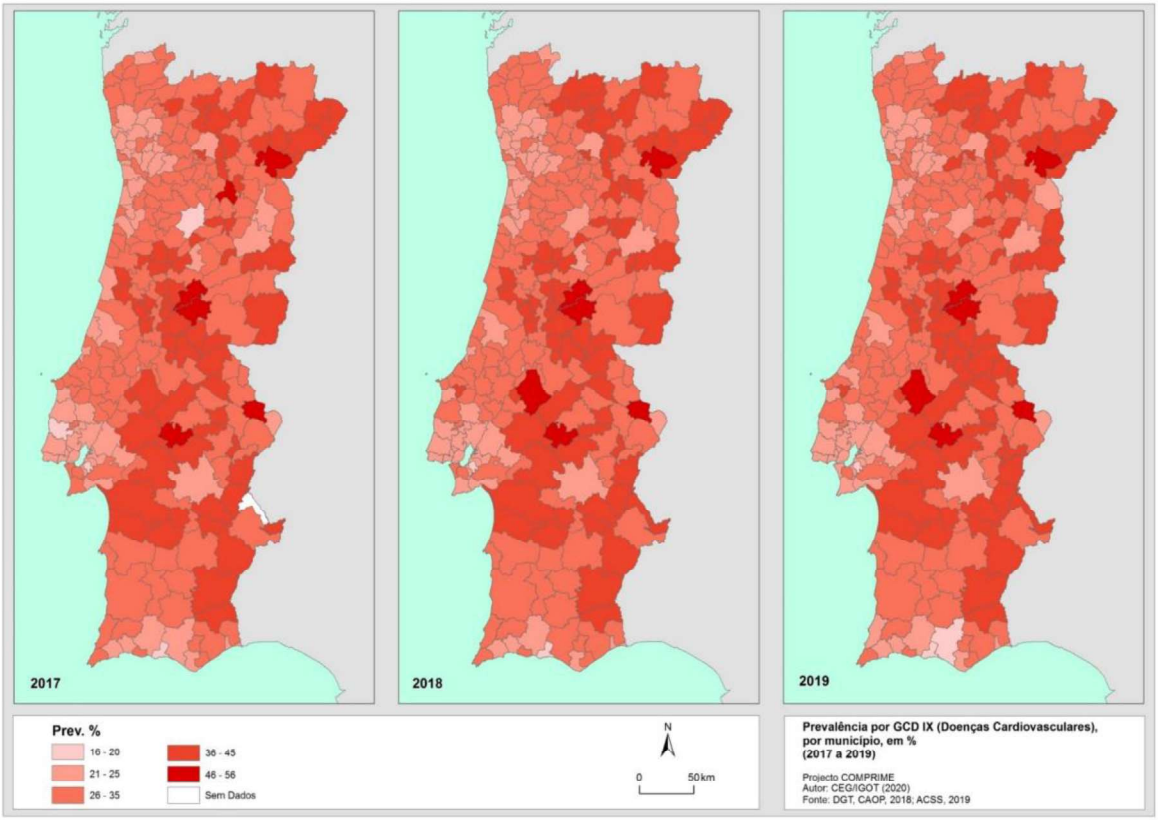
Endocrine, nutritional and metabolic diseases



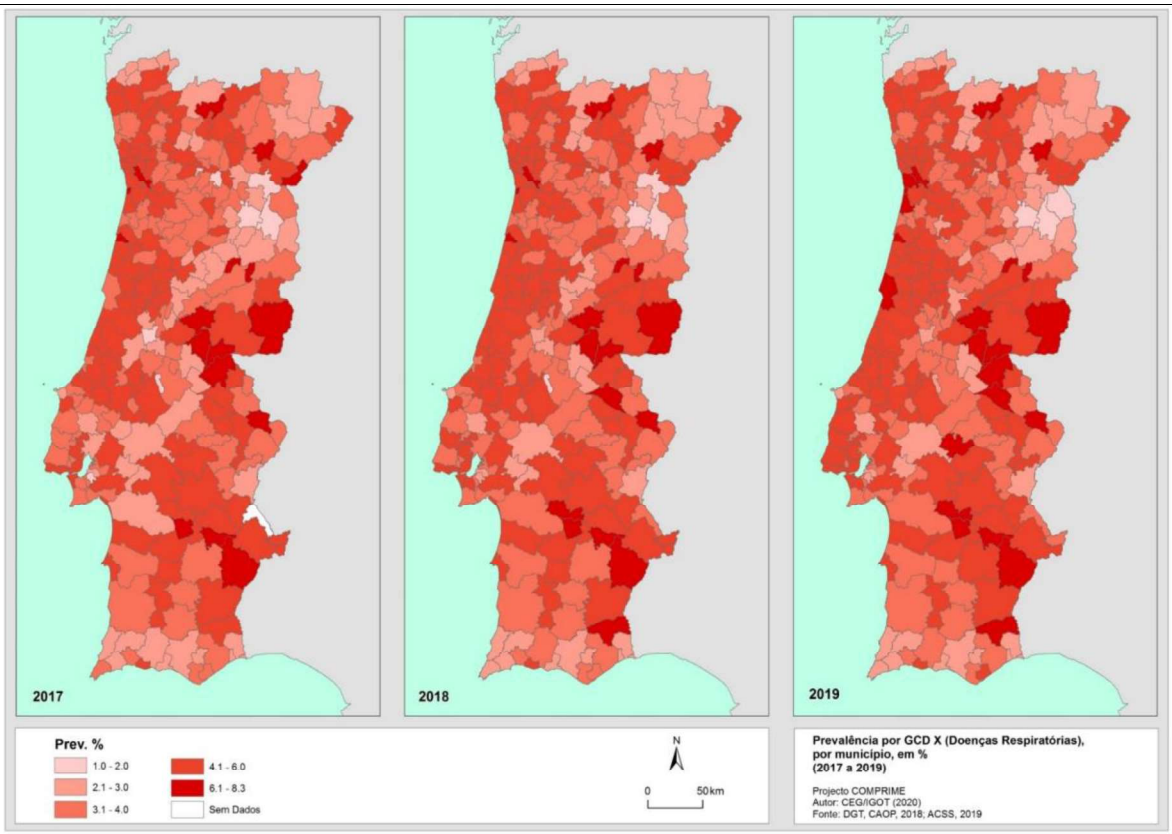
Mental illness and substance abuse



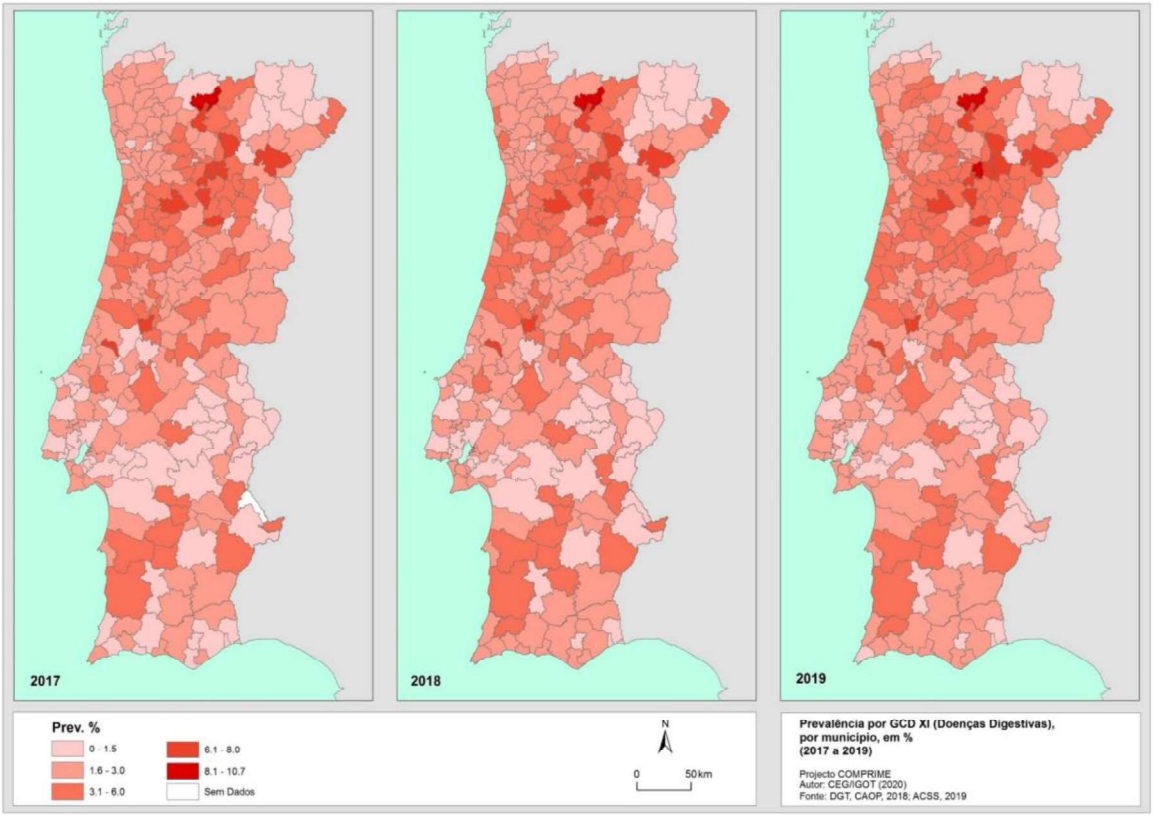
Cardiovascular diseases



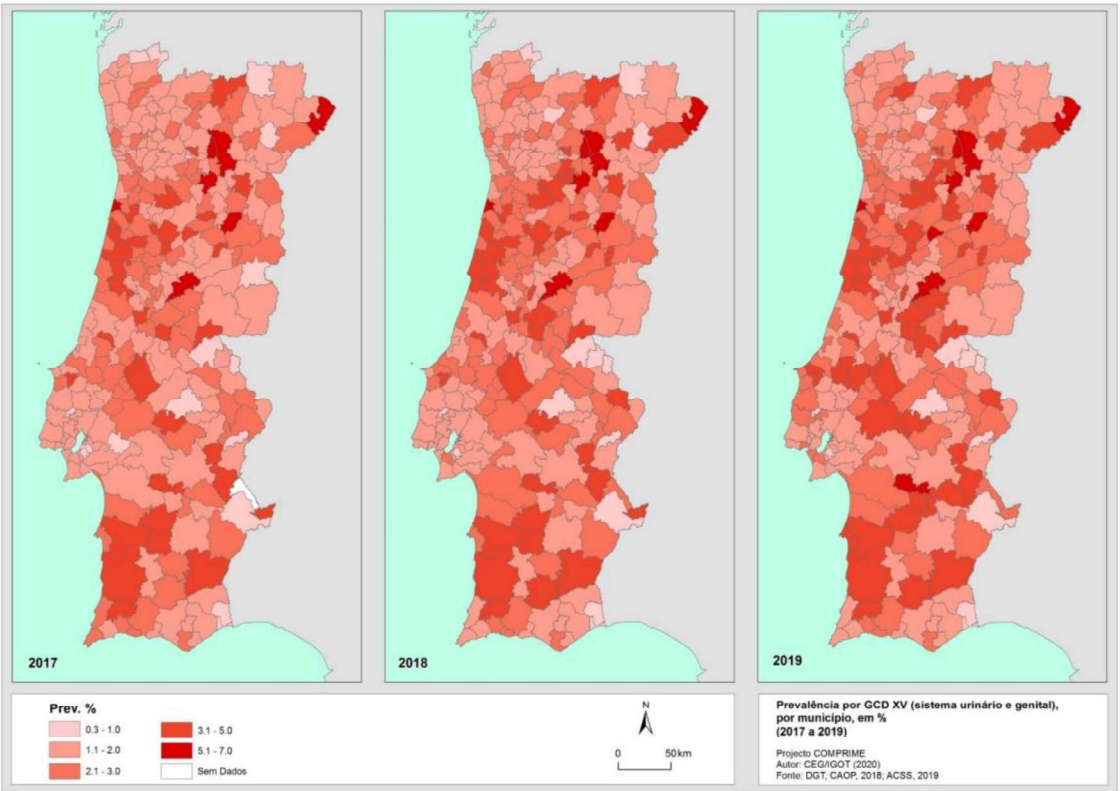
Respiratory diseases



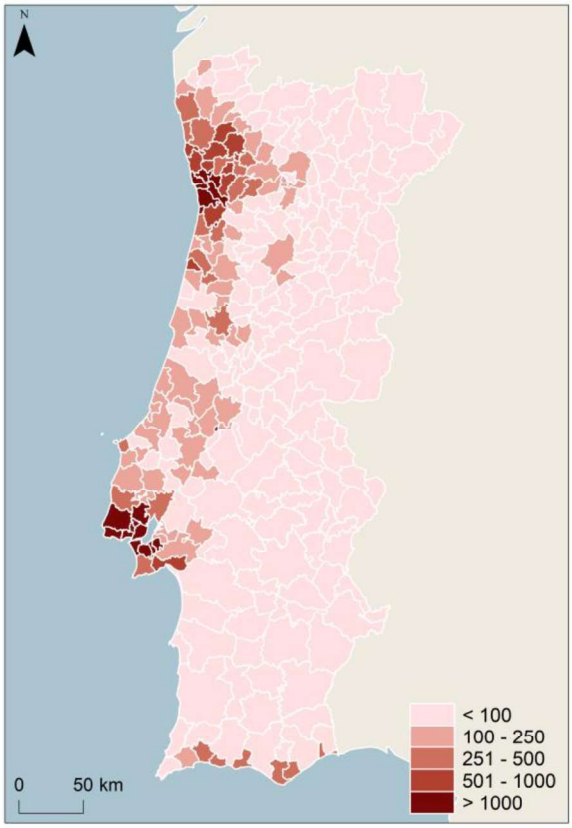
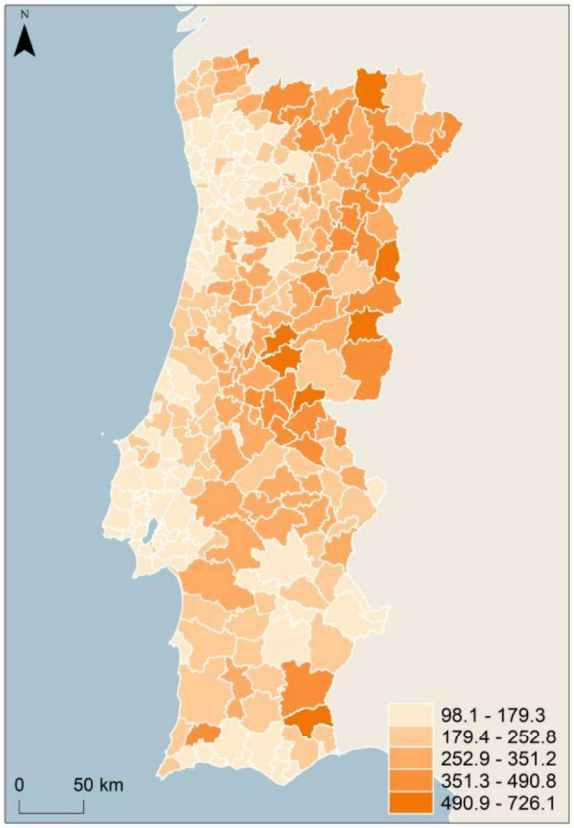
Digestive diseases



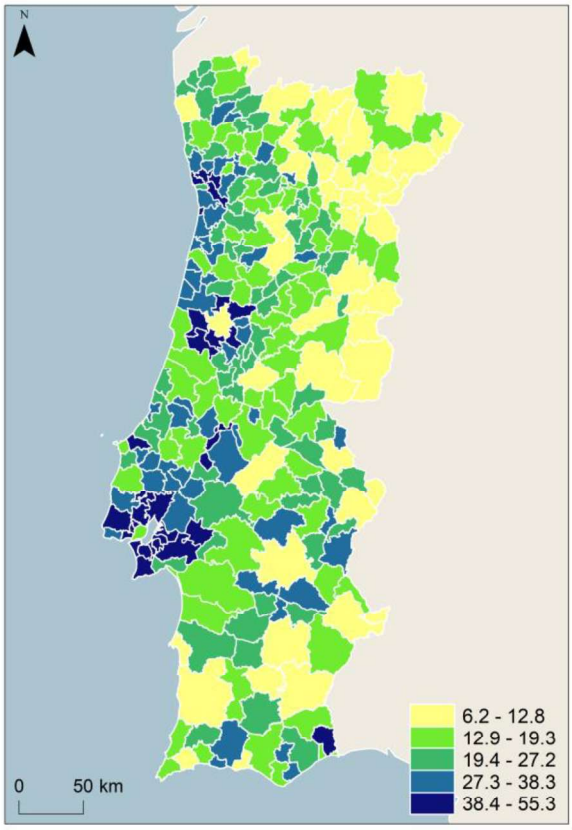
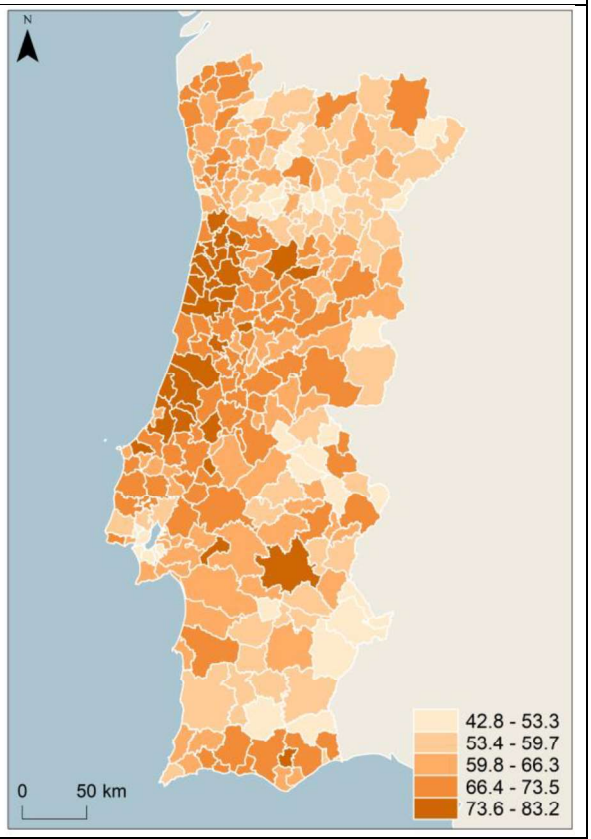
Diseases of urinary and genital systems

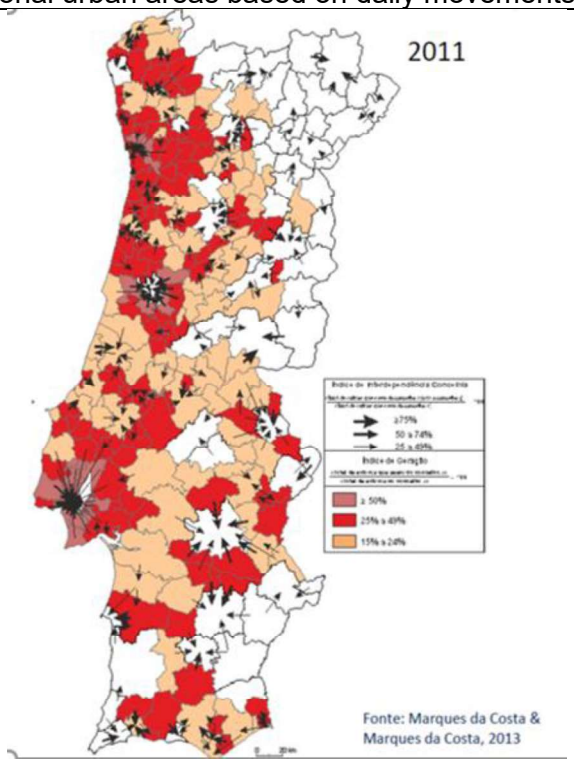


3.3. Population and settlement

Population density, 2019	Aging index, 2019
	
Source: INE	Source: INE

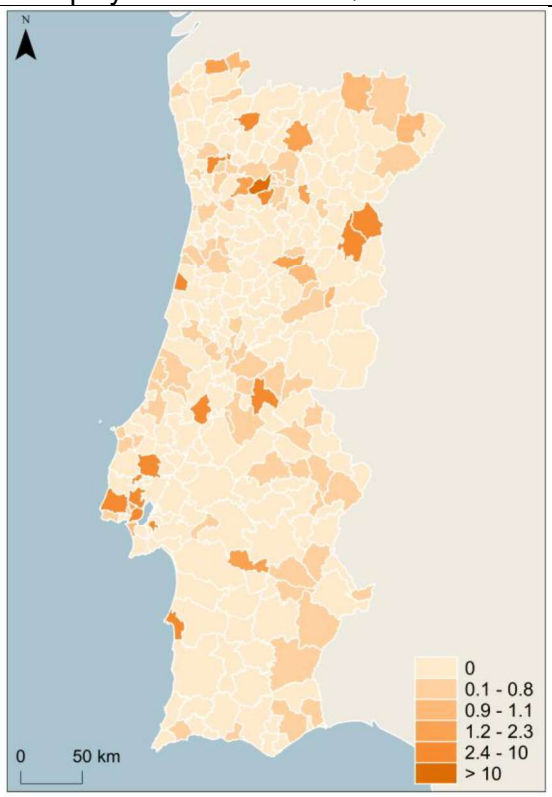
3.4. Mobility

% of population working outside the municipality of residence, 2011	% of population using car, 2011
	
Source: INE, RGP, 2011	Source: INE, RGP, 2011

Functional urban areas based on daily movements, 2011

Source: Marques da Costa e Marques da Costa, 2013

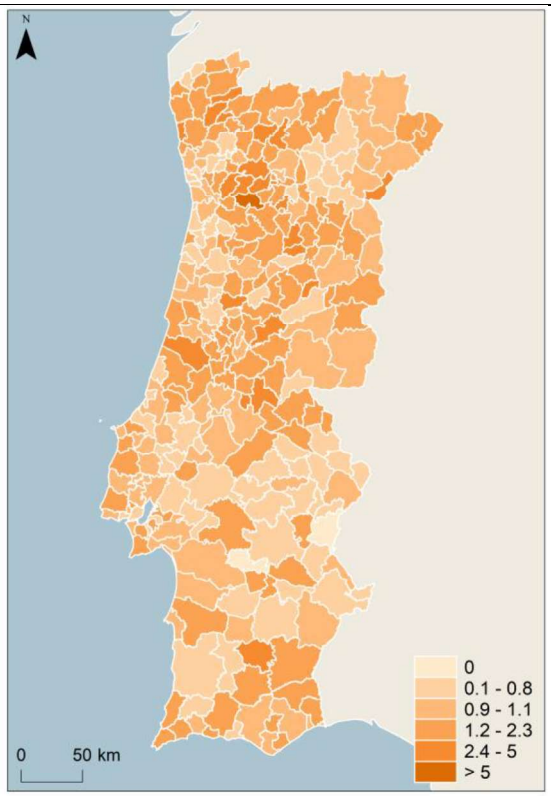
3.5 Economic dimension

Storage and transport activities employment - Location Quotient 2018



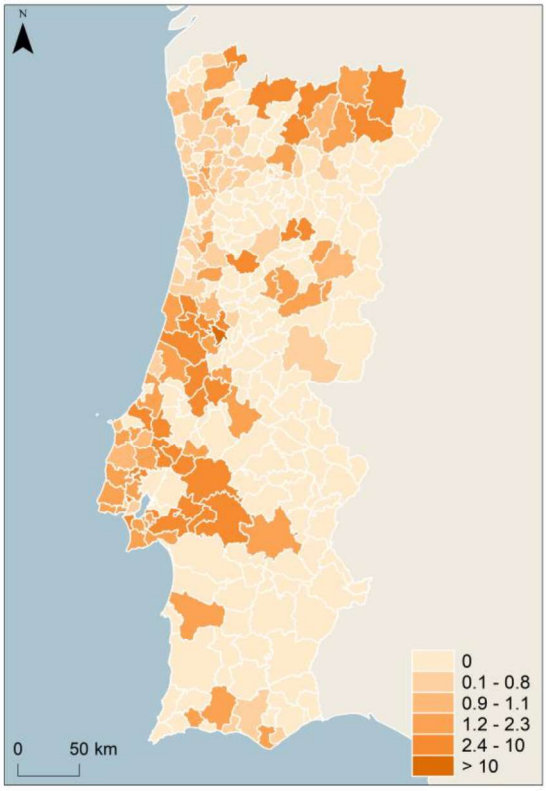
Source: INE

Civil Construction employment - Location Quotient 2018



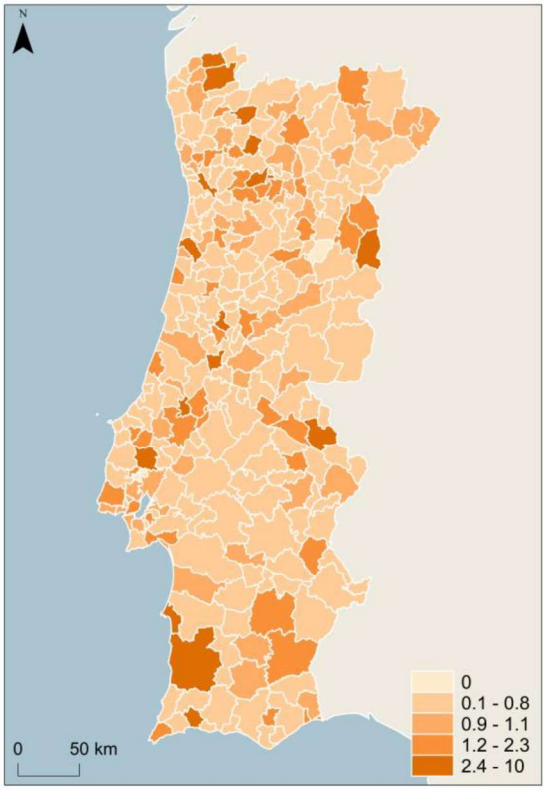
Source: INE

Social support activities employment – Location Quotient 2018

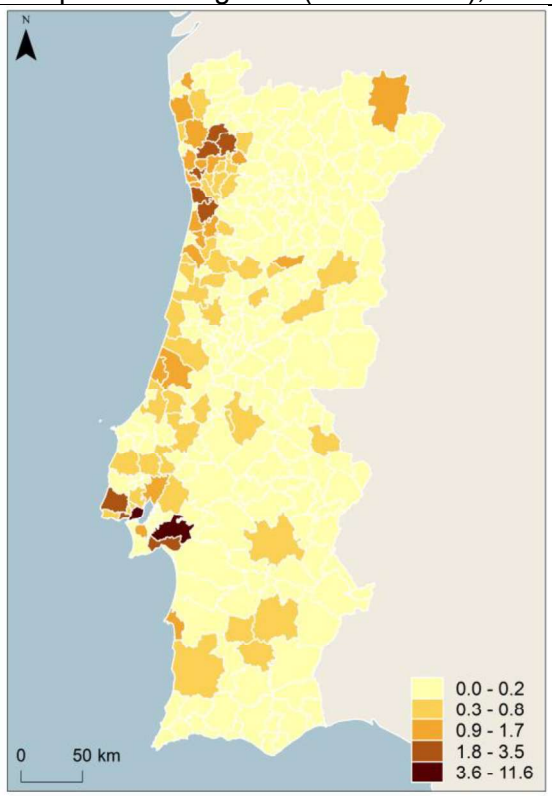


Source: INE

Transport and storage employment – Location Quotient 2018

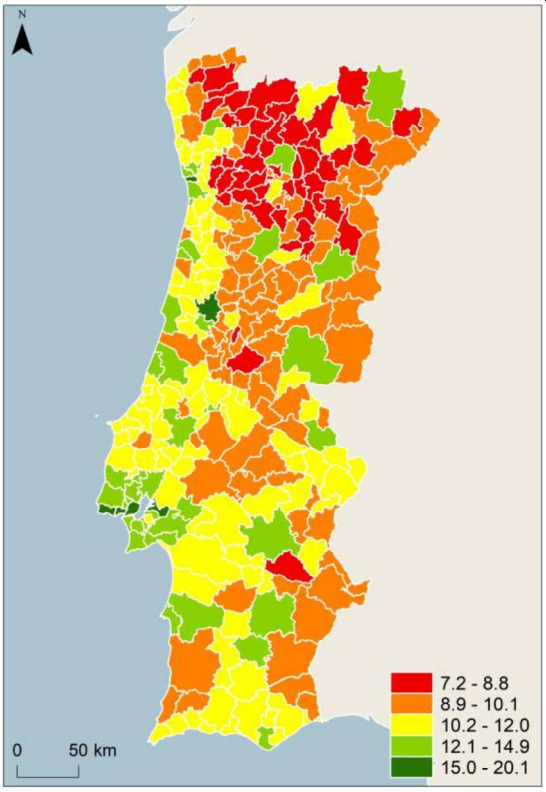


Source: INE

Exportation of goods (% over total), 2019	
	
Source: INE	

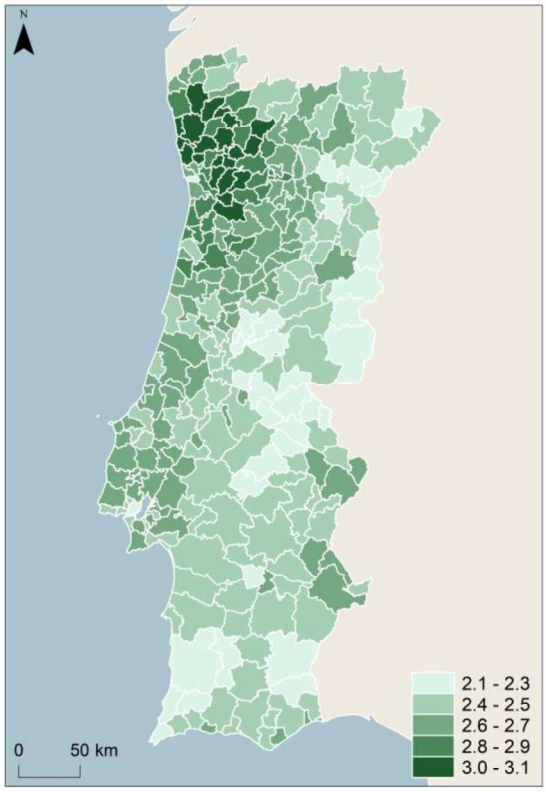
3.6 Social Dimension

Gross income by taxable person (€) 2018



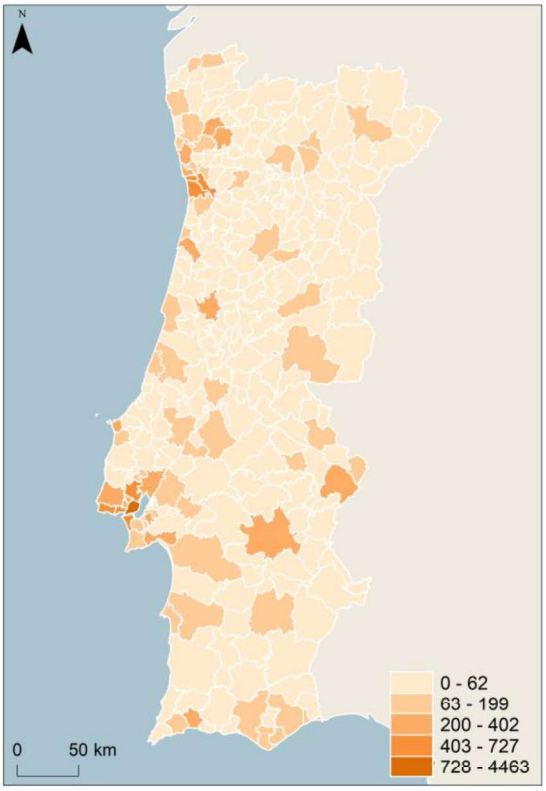
Source: INE

Average household size, 2011



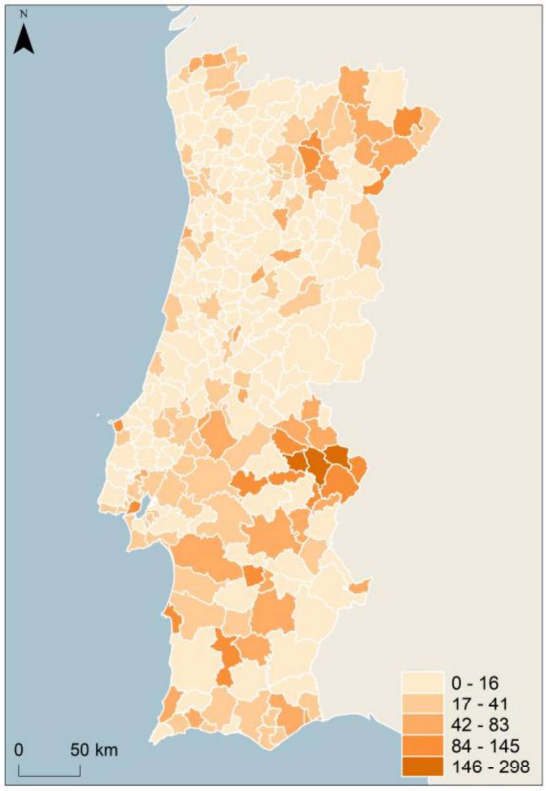
Source: INE

Number of social housing units, 2011



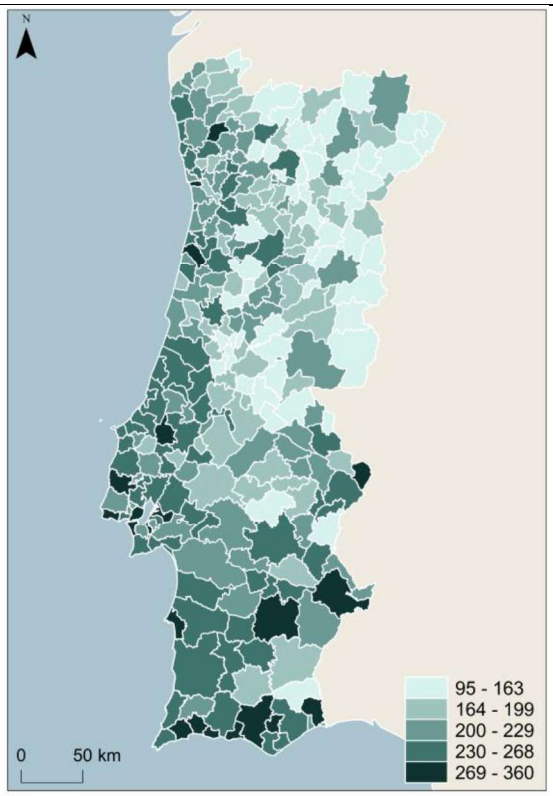
Source: INE

Number of social housing units per 10 thousand inhabitants, 2011



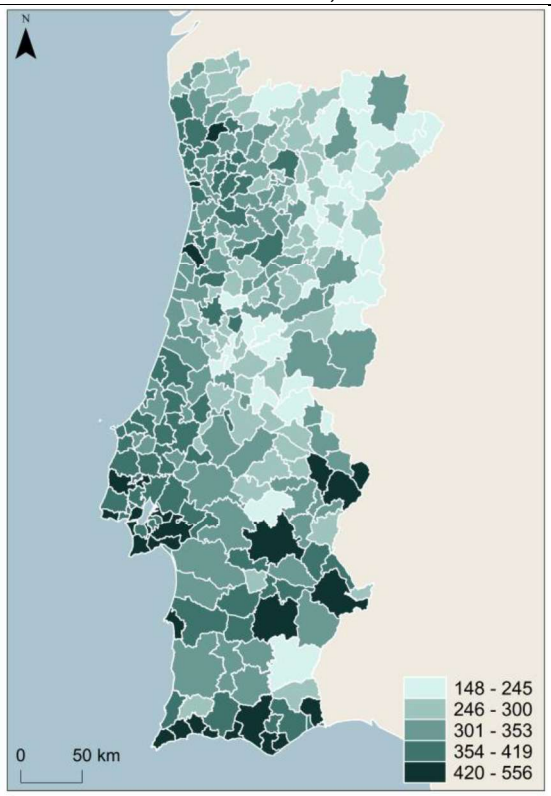
Source: INE

Students in pre-school /1000 inhabitants, 2018/19



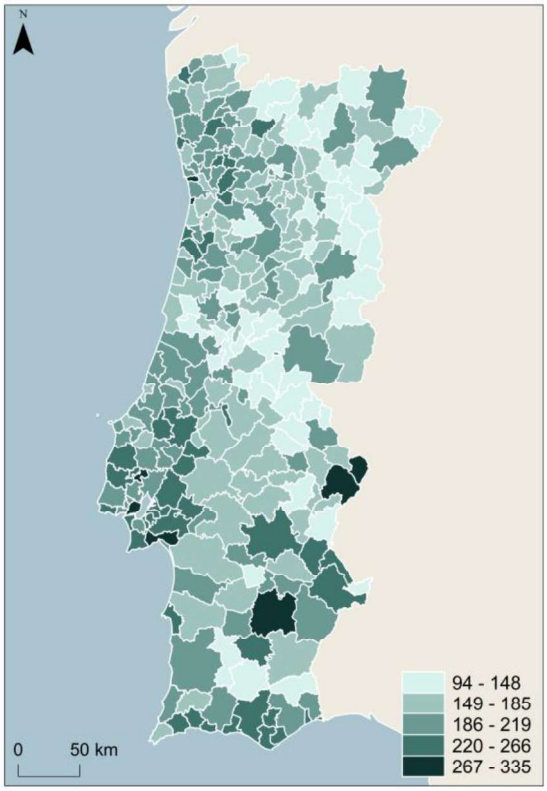
Source: INE

Students in the 1st cycle of Basic Education / 1000 inhabitants, 2018/19



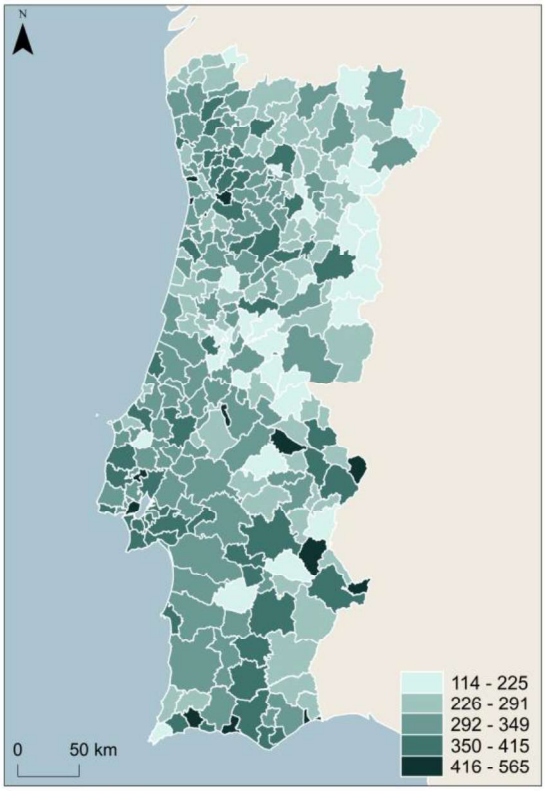
Source: INE

Students in 2nd cycle of Basic Education / 1000 inhabitants, 2018/19



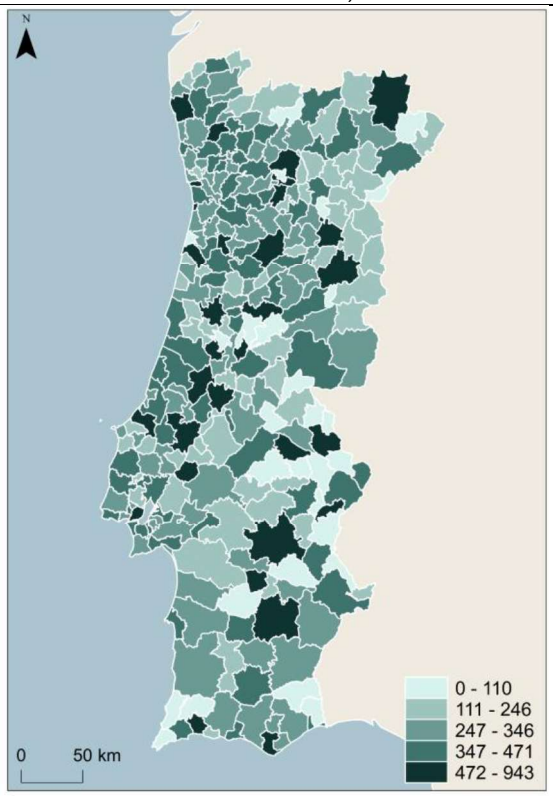
Source: INE

Students in 3rd cycle of Basic Education / 1000 inhabitants, 2018/19



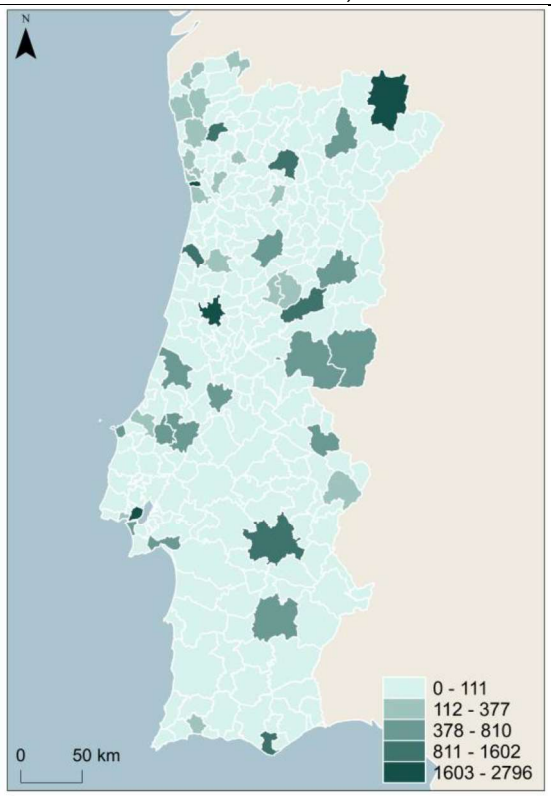
Source: INE

Students in Secondary Education /1000 inhabitants, 2011



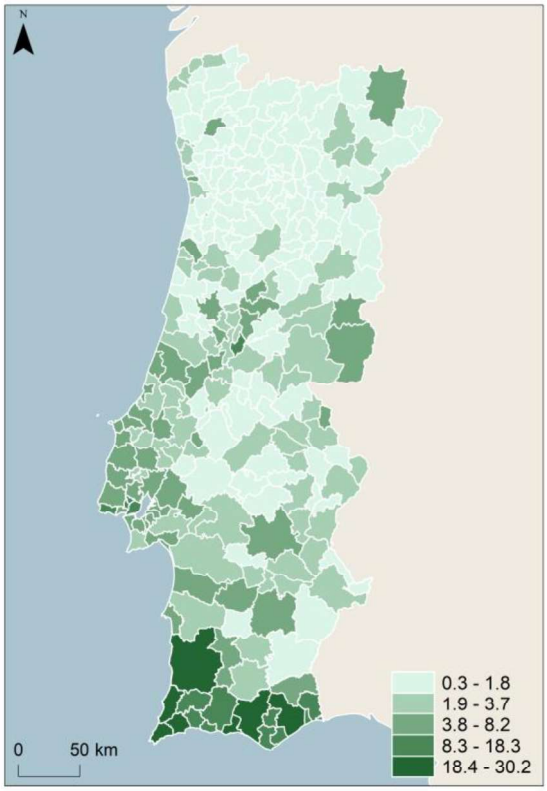
Source: INE

Students in Higher Education /1000 inhabitants, 2011



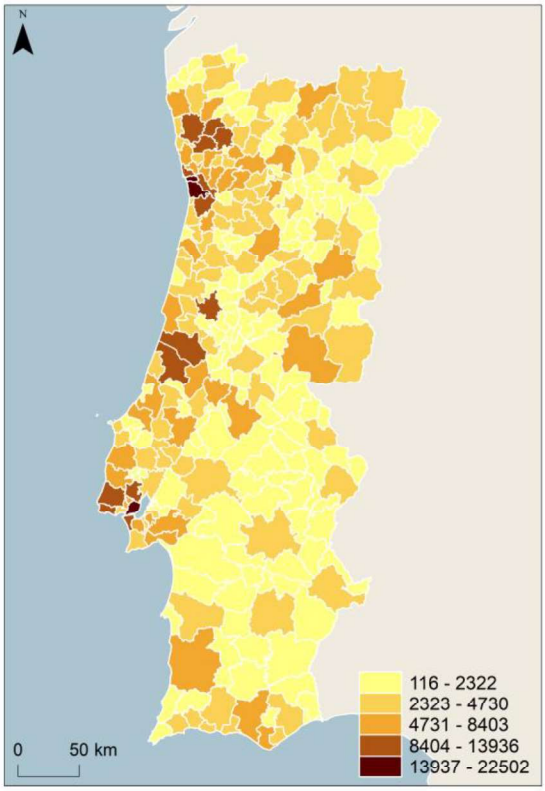
Source: INE

% Legal immigrants in the total population, 2018



Source: INE

Nº of buildings in need of repair, 2011



Source: INE