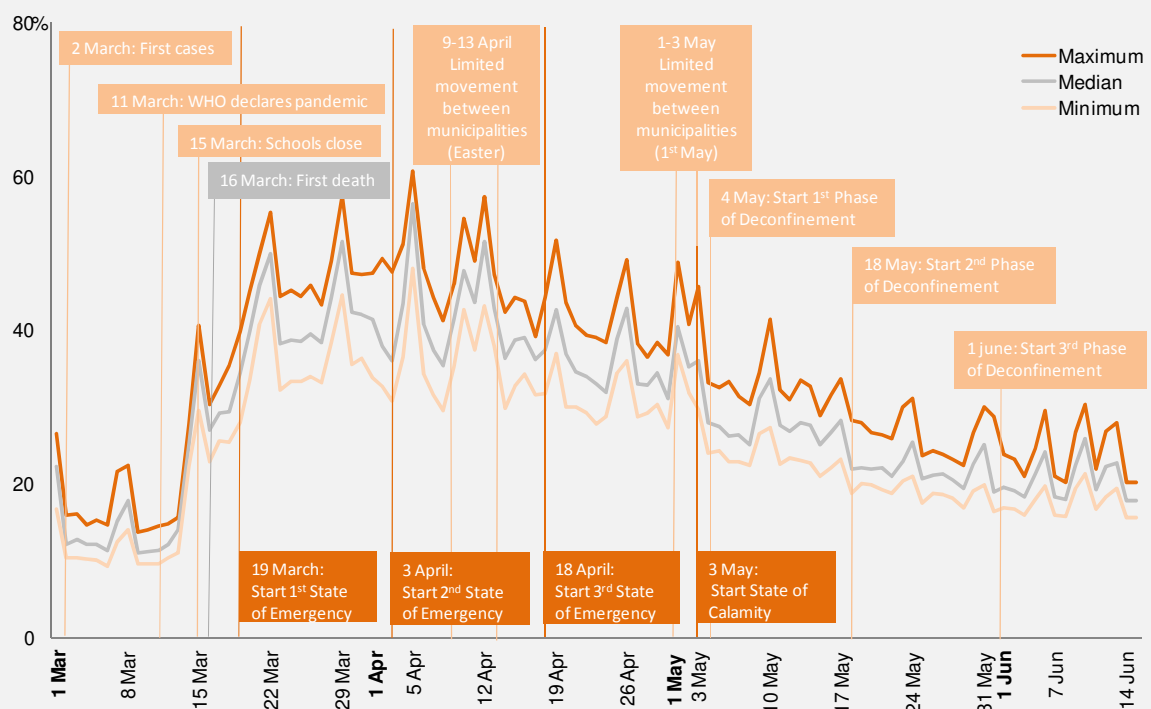


## Population mobility indicators at regional level: an analysis based on information from Facebook's "Data for Good" Initiative

In this box, taking advantage of Facebook's "Data for Good" Initiative, population mobility indicators at NUTS 3 level in the national territory are released.

The data represented in the figure below correspond to the proportion of population "staying put" between March 1<sup>st</sup> and June 15<sup>th</sup>, namely minimum, median and maximum values obtained from the 25 NUTS 3 sub-regions of the country. For a better contextualization of the information, the figure includes the main key moments associated with the COVID-19 pandemic in Portugal.

Proportion of the population "staying put" between March 1<sup>st</sup> and June 15<sup>th</sup> – minimum, median and maximum values of NUTS 3

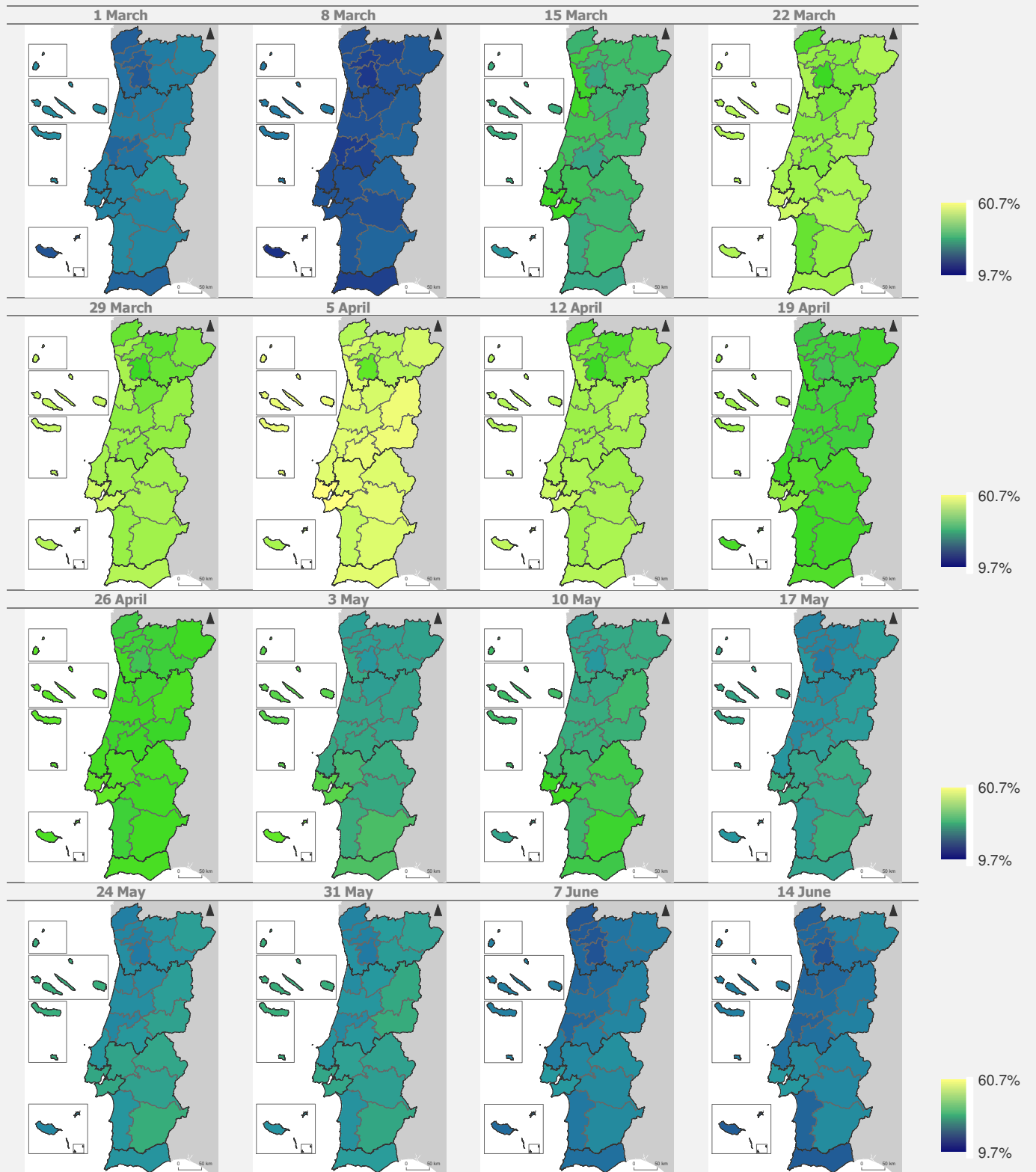


Source: Facebook's "Data for Good" Initiative. Data provided by Carnegie Mellon University.

Note: The dates marked on the graph axis correspond to the first days of the month and Sundays.

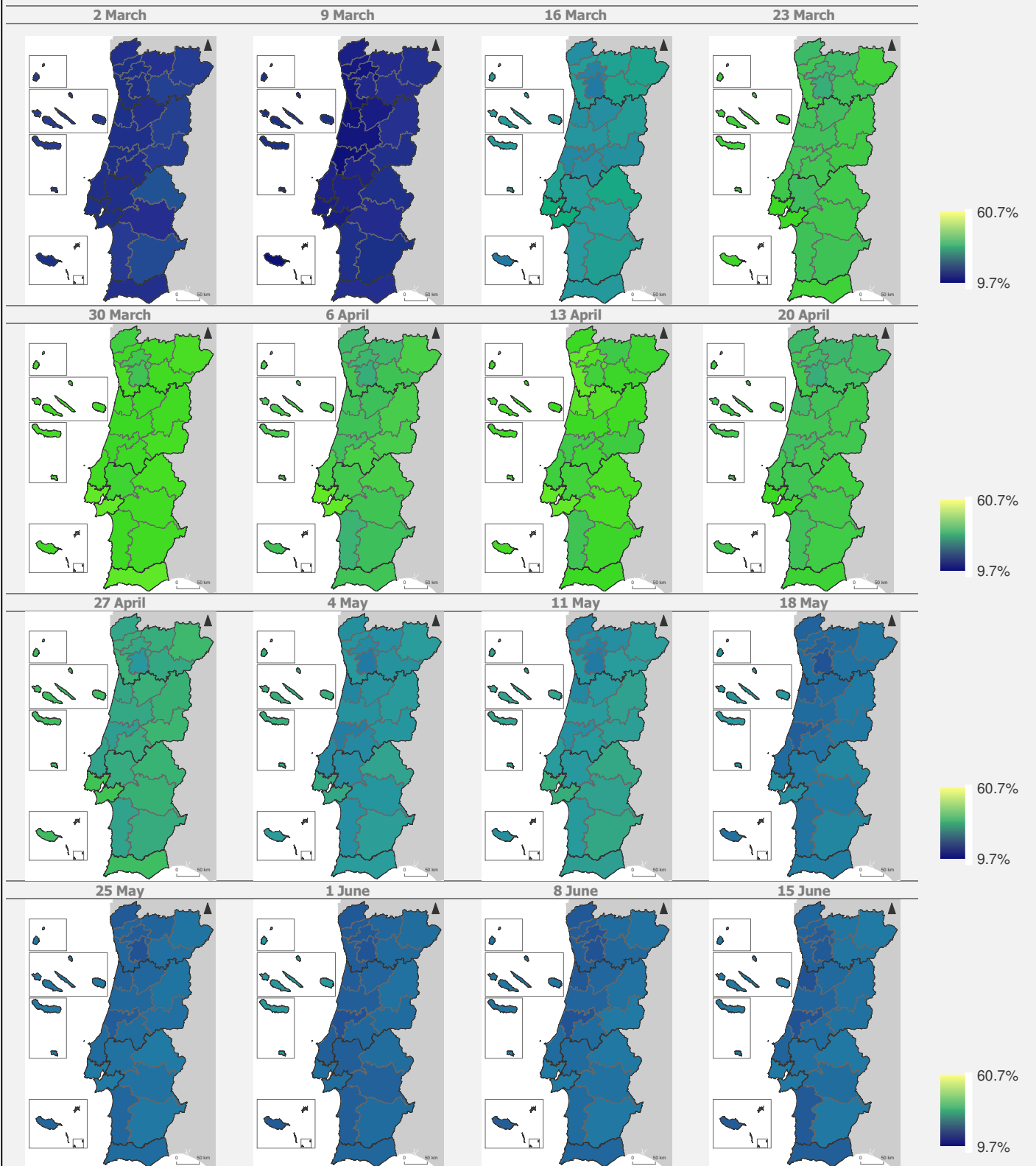
The following figures show this indicator at NUTS 3 level for the days corresponding to Sundays [Figure 15] and Mondays [Figure 16], since the beginning of March. It can be seen that the days corresponding to Sundays indicate, overall, less mobility of the population than the days corresponding to Mondays. In particular, there is a reduction in mobility levels with the beginning of the State of Emergency on March 19 (maps of March 22 and 23). On the contrary, a progressive increase in mobility has been registered with the transition from the State of Emergency to the State of Calamity on May 3, followed by the first phase of implementation of the deconfinement measures (maps on May 3 and 4), and by the second phase of deconfinement on May 18 (maps on May 18, 24, 25 and 31 and June 1), and with the beginning of the third phase of deconfinement on June 1 (maps on June 1, 7, 8, 14 and 15).

Proportion of the population "staying put" on Sundays between March 1<sup>st</sup> and June 14<sup>th</sup> by NUTS 3



Source: Facebook's "Data for Good" Initiative. Data provided by Carnegie Mellon University.

Proportion of the population "staying put" on Mondays between March 2<sup>nd</sup> and June 15<sup>th</sup> by NUTS 3



Source: Facebook's "Data for Good" Initiative. Data provided by Carnegie Mellon University.

**Technical Note:**

The mobility data from Facebook's "Data for Good" Initiative correspond to location updates collected from mobile devices of Facebook application users that have the "location history" option turned on. Only location accuracy (GPS) data of less than 200 meters is considered and if a user has multiple locations resulting from more than one associated mobile device, Facebook only considers the data with the highest location accuracy. Obtaining results for the NUTS 3 level implies a minimum of 300 unique users per sub-region.

The proportion of the population "staying put" is measured by the number of Facebook users associated with a single 600mx600m reference grid during 8am and 8pm on day x, requiring at least three occurrences during that time period. The reference grid, as a "residence" proxy, is measured daily based on the largest number of locations observed between 8pm and midnight on day x-1 and between 0 am and 8 am on day x, requiring at least three occurrences during that time period.

The information associated with the 600mx600m grids is allocated to the respective NUTS 3 sub-region. Since a grid can intercept more than one sub-region, 9 sample points are generated in each grid, assigning 1/9 of the grid population to each point in the sample.

Facebook's "Data for Good" initiative aims to provide data for research on humanitarian issues and has allowed results to be published in scientific articles particularly in the United States. Obviously, Statistics Portugal's use of this data source in the Statslab domain is not motivated by any publicity motive, but by the public interest of the information. Statistics Portugal thanks researcher Miguel Godinho Matos<sup>1</sup> for his support in the analytical preparation of this information

---

<sup>1</sup> Associate Professor at Católica Lisbon School of Business & Economics and visiting research scholar at the Carnegie Mellon University.

## Technical note

### Data sources

Data on [Deaths](#) correspond to general deaths (all causes of death) occurring in the national territory since March 1<sup>st</sup>, 2020 and until the Tuesday of the week prior to publication. The information is preliminary and is obtained from statistical operations of direct and exhaustive collection on deaths occurring in Portuguese territory using facts that are subject to compulsory civil registration (death) in the *Sistema Integrado do Registo e Identificação Civil* (SIRIC). In addition to administrative information obtained from Civil Register Offices, Statistics Portugal collects an additional set of variables identified as statistically relevant to the National Statistical System (NSS) and the European Statistical System (EES). Data are recorded and sent electronically, in compliance with the requirements set out by Statistics Portugal and laid down in liaison with the *Instituto de Registos e Notariado* (IRN) and the *Instituto de Gestão Financeira e Equipamentos da Justiça* (IGFEJ).

Data on the number of confirmed cases are based on those published daily in the [Directorate-General of Health COVID-19 Status Report](#) for the entire country and by municipality. The confirmed cases are referenced to the municipality of occurrence and correspond to the total of clinical notifications in the SINAVE (National System of Epidemiological Surveillance) system. For the reference dates considered in this press release –June 17 – data by municipality corresponded, respectively, to 90% of confirmed cases in the national territory. This proportion reflects data confidentiality by municipality, but also limitations in the process of spatial referencing of information. In fact, when the confirmed cases by municipality are fewer than 3, for confidentiality reasons, data are not disclosed by the Directorate-General of Health.

Data on dwellings sales are based on the use of administrative procedures, namely from anonymised administrative tax data obtained from the Portuguese Tax and Customs Authority (AT) under an agreement signed with Statistics Portugal, on the Municipal Property Transfer Tax (IMT) and the Municipal Property Tax (IMI). The calculation is based on the linking of information from IMT with that from IMI and only sales where the IMT destination code is "Housing" are used and the associated information from IMI is defined as "Housing". The calculations follow the methodology described in the Methodological Document ["House Price Statistics at Local Level"](#). As part of the monitoring of the impact of the COVID - 19 pandemic, Statistics Portugal anticipates the dissemination calendar and calculates the indicator of the number of dwellings sales for each month, preliminary estimates, corresponding to the information recorded in the reference month and the two previous months, i.e. with a reference period of 3 months.

Also, data on new lease agreements are based on administrative procedures, namely from anonymised administrative tax data provided by the Portuguese Tax and Customs Authority (AT) under an agreement signed with Statistics Portugal, on the Statement of Stamp Duty Model 2 - Communication of lease agreements (Model 2) and the Municipal Property Tax (IMI). The calculation is based on the linking between Model 2 information with that of IMI. The first declarations and declarations of substitution of new lease agreements for urban buildings, with a monthly rent period, for which the purpose is permanent housing, and the associated information from IMI is defined as "Housing", are used. The calculations follow the methodology described in the Methodological Document of ["House Rental Statistics at Local Level"](#). As part of the monitoring of the impact of the COVID - 19 pandemic, Statistics Portugal anticipates the dissemination calendar and calculates the indicator of the number of new lease agreements for each month, preliminary estimates, corresponding to the information on the new lease agreements registered in the reference month and the two previous months, i.e. with a reference period of 3 months. In this context of preliminary estimates, an error in the calculation of the information disseminated in the press release of 22 May took place, leading to an underestimation of the number of new lease agreements for the months of January, February and March 2020. The image that resulted from this calculation, in general, did not affect the relative position of the NUTS 2 regions or the trend in the period from January to March 2020.

This press release includes the resident population data as of December 31, 2019 released on June 15.

### Disseminated Indicators

Number of total deaths, by sex or age group

Number of deaths in the last 4 weeks per deaths in the same reference period

Number of confirmed cases of COVID-19 disease per 10 thousand inhabitants

Population density

Number of new confirmed cases of COVID-19 disease in the last 7 days per 10 thousand inhabitants

Proportion of resident population with 75 or more years old

Number of new lease agreements per one thousand conventional dwellings

Number of dwellings sales per one thousand conventional dwellings

Relation between the number of dwellings sales in April 2020 (last 3 months) and sales in the same reference period

Location coefficient

The location coefficient (LC) is obtained using the following formula:

$$LC = \left( \frac{1}{2} \sum_{j=1}^n |x_j - y_j| \right) \times 100 \quad \text{where:}$$

$x_j$  corresponds to the ratio of the number of confirmed cases of COVID-19 in each municipality  $j$  to the number of confirmed cases of COVID-19 for the total country;

$y_j$  corresponds to the ratio between the resident population in each municipality  $j$  and the total resident population in the country.

The Location coefficient varies between 0 and 100, with values closer to 100 reflecting greater inequality in the distribution of confirmed cases of COVID-19 against the total resident population and, in this sense, indicates situations of greater territorial concentration.

The location curve (or Lorenz concentration curve) corresponds to a graphical representation that relates the cumulative distribution of two variables. This representation also includes the straight line of equal distribution, and the greater the distance from it, the greater is the concentration of the variable represented in the ordinate axis (in this analysis, the confirmed cases of COVID-19, by period of reference) versus the variable represented in the abscissa axis (in this analysis, the total resident population).