GeoCensos

# Mapping responses to COVID-19: How far are National Statistical Offices (NSOs) leading a data revolution?


## Abstract:

*In national statistic offices of countries around the world, geospatial data is needed for a swift response to the COVID-19 emergency. Initiatives aimed at integrating geospatial data with statistics to understand the spatial behaviour of this phenomenon are a great tool for decision taking not only in the short term but also within the 2030 sustainable development agenda endeavour. This report presents a detailed list and an analysis of geo-statistical initiatives undertaken by NSOs in response to coronavirus. The purpose of the produced list is to document basic characteristics to later identify common points that could offer preliminary but valuable lessons, especially regarding data partnerships for the SDGs. This report highlights seven lessons coming from the unusual collaboration between NSOs and third party sources that took place in the corona virus juncture. It suggests their importance to anticipate real progress in the data revolution to motivate NSOs to further improve their performance towards the 2030 Agenda leading innovative data partnerships.*

GeoCensos

## Introduction

Antonio Gramsci, the ill-fated Italian intellectual, said at the beginning of last century that "The old world is dying but the new one is delayed in emerging and in that chiaroscuro monsters emerge". If this thinker had lived in our days, he would have recognized one of those monsters in the impacts of the COVID-19 epidemic. And not only because the virulent coronavirus leaves death in its wake, but also because it especially reveals gaps in the responses of governments. One of the gaps that have been revealed concerns the controversial chord of official data and statistics on the territory.

As a global phenomenon that has evolved over time and space with unprecedented operational conditions, governments have been hampered in their ability to generate, share, match and disseminate geospatial and statistical data in a timely and reliable manner. Indeed, only a small number of national statistical offices and systems have demonstrated expertise in addressing the emergency. However, there are few but valuable geospatial response initiatives to COVID-19 at the global, continental and national levels.

At the global level, the COVID-19 monitoring initiative launched by Johns Hopkins University is well known. It combines statistical and geographical data and tracks the phenomenon of the pandemic in 188 countries. In the now five-month pandemic journey from China to United States, this platform has successfully combined multiple sources of real-time daily data. It is understandable that is was to be cited extensively in national and global news stories on a day-to-day basis.

## Data partnerships and NSO geospatial responses to COVID-19

At the national level, the challenge of understanding the pandemic in its impact on the territory has been taken up by a number of NSOs. The geospatial response to coronavirus has resulted in the launch of solutions with varying added values for decision-making. The following table summarizes a subset of reviewed strategies:

**Some characteristics of geo-statistical responses to COVID-19 by NSOs in the world [1]**

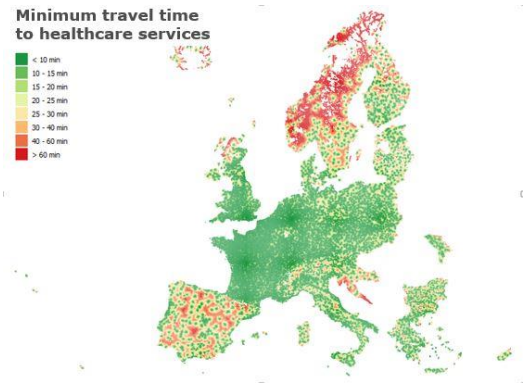
| Leading NSO | Type of Analysis / Visualisation | Kind of Partnership | Source Data Layer | Data reusability |
|---|---|---|---|---|
| **Eurostat** | Travel time to health services | G2G | GDB of health services | Open |
| **DANE Colombia** | Vulnerability Index | G2G | 2018 + Census | Proprietary |
| **INE Spain** | Population Mobility | G2B | Real time | Proprietary |
| **INE Portugal** | Hex Grid of Infected People | G2G | 2018 + Census | Open |
| **INEGI Mexico** | Mapping of infected and Other | G2G+ | N/A | Open |
| **GSS Ghana** | Population Mobility | G2B | Real time | Proprietary |
| **Statistiks Norway** | Hex Grid of Infected People | G2G | Registries | Open |

[1] Source: Own based on contents of links in considered cases. *Kind of Partnership* was characterized based on classification established in Citizen to government data partnerships: What can we learn from and recommend to civil society groups working in the official statistics domain? In source data layer "+" means that there are other layers used. In *Kind of Partnership* "+" means other partners.

The following is a more detailed review of found geo-statistical initiatives in response to the coronavirus phenomenon by NSOs around the world, although it does allow for

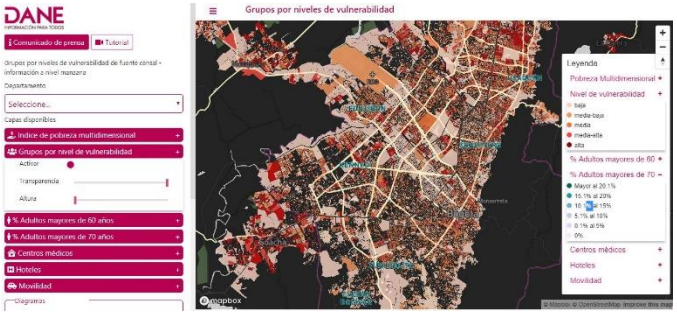
other experiences to enrich it. The purpose of the following list is to document basic characteristics in order to subsequently identify commonalities that may offer preliminary lessons:

- At the continental European level, the European Union Statistics Office (Eurostat) has provided coverage for all the countries of the Union by [mapping hospitals and other health care centres and by hosting resulting databases in the web services of ESTAT-GISCO project] . As a result, it allows having an accurate analysis of the locations of health care points of interest with respect to roads and calculating minimum time of access.




- The [Norwegian] and [Portuguese] statistical agencies have also produced open data sets mapping hexagons at the sub-national level to reflect the distribution of coronavirus infections at the municipal level in those nations.
- The Spanish Institute of Statistics has ordered a study of the location of more than 80% of mobile phones across Spain in order to gain an understanding of the mobility of people in the country at this juncture.. The study was [prepared by the National Statistics Institute (INE) in collaboration with the three major mobile telephone operators] (Orange, Telefónica and Vodafone). Its approach breaks down the territory into mobility areas, where each operator supplies the aggregate number of devices, considered as the resident population in each area. In other words, a person (its mobile device) is classified as resident in the area where his/her phone has remained for the majority of the time from 0h to 6h. In this way, each mobile phone is assigned a residence area and, on this basis, findings are generated with the observations obtained in an aggregate manner.
- The Data Partnership formed by the Irish NSO for geospatial COVID-19 surveillance involved the Central Statistics Office of Ireland (CSO), Ordnance Survey Ireland, the Department of Housing, Planning and Local Government and the All Islands Research Observatory at the University of Maynooth, with the support of Esri Ireland. This partnership developed [the Covid-19 National Data Hub over the GeoHive platform]. The GeoHive is the state's geospatial data platform for the Public Service Data Strategy 2019-2023. Its data format is of a "Statistical Interlinked Open Data" (LOSD) nature which allows a smooth integration.
- The Ghanaian Statistical Service (GSS) conducted a [study using anonymized and aggregated data from mobile network operators in the country]. The initiative aims to understand mobility patterns of its population and strengthen decision making. The data was analysed in real time with the support of a UK-based academic spin-off called Flowminder to produce mobility indicators from data provided by the firm Vodafone. The findings of the initiative report are based on an "origin-destination matrix", a data analysis based on the actual number of users travelling between two locations.

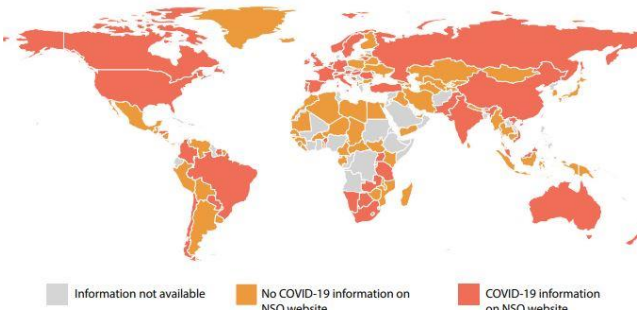
- Another input is that of Colombia's administrative national direction of statistics (DANE, in its Spanish acronym), which uses a viewer to determine, at a disaggregated level, the population susceptible to infection by COVID-19. Its method selects municipalities from all over the country and uses data from the last census, in addition to administrative records collected by two other public agencies. The approach of identifying blocks which have high levels of vulnerability groups links together territorial units with similar demographic characteristics and comorbidities likely to be related to the pandemic. The crossing of layers of different bases calculates an index that is related to centroids that identify different levels of vulnerability.




- Mexico's INEGI has set up an analytical viewer for COVID 19 where it is possible to retrieve an open data layer classifying cases of people infected by the virus at the state level and adding information on health centres, basic commercial activities and supply points of interest. This Institute acknowledges on its website that the measures derived from the sanitary emergency have affected its reporting programs in different ways and releases its responsibility regarding quality, coverage or other problems in the data, thus relegating its explanation to technical notes.
- Finally, Statistics Sweden's mapping tool (SCB in Swedish) contains a number of simple functions for displaying data linked to regional divisions and displays both location of the vulnerable population and tabular statistics. It looks especially at the population aged 70 and over by area of demographic statistics (DeSO). The viewer can calculate statistics by county, municipality and also by more detailed municipal levels. It is possible to filter areas by lower or higher percentages of vulnerable population.
- There are marginal cases such as those in Brazil and Lithuania that have set up Hubs through ESRI's Argis on line platform, showing data on infected populations. There also exists a case in the United Kingdom that has produced a code developed by the Campus of Information Sciences of its ONE (ONS) to "scrape" or retrieve the data from Mobility profiles that the Community of Google publishes and is based on the aggregation of mobile phone mobility information by activating GPS functionality on the Android platform.

## An unexpected revolution in data, but truly led by NSOs?

The purposeful and synergetic association of partners in this report represents in part the self-prophecy that the "data revolution" promises from the 2030 agenda. This initiative, agreed globally by the United Nations, promotes sustainable development, recommends transformative measures for national statistics, and seeks improvements in the way data are produced and used to achieve goals in 17 key areas. Since 2015, the agenda calls on a wide range of actors - not just governments and their national statistical systems - to work together to generate timely and usable data, while facilitating informed decision-making, including that during emergencies.




Nevertheless, while most governments in the world were already taking action to prevent and mitigate COVID-19, by early April 2020 - based on OECD figures - less than 20% of the approximately 186 known official NSO websites published information about the pandemic online. According to the UN hub on coronavirus responses as of the end of May 2020, no more than 14 NSOs from countries around the world had published any documentation or data crossing regarding coronavirus, including some of the initiatives reviewed above.

Unsuspectedly, and despite the fact that Agenda 2030 was led by governments and public officials, its spirit has been difficult to identify in most statistical agencies around the world, at least judging by the communication needs that remained unsatisfied during the pandemic. Since these offices are precisely being called upon by the United Nations to lead the data revolution, users of their data will expect them to provide timely, efficient and quality results in monitoring both this pandemic and each of the 17 Agenda 2030 goals. The level of leadership required by them is such that it facilitates the transformation of reality on a true basis for all possible sustainable development objectives by making valuable information available to interested parties.

A possible explanation for the leadership shortcomings shown by many NSOs could be found in the Policy Brief of the Association of Statistics for Development in the 21st Century, also known as Paris 21. Its approach is that as developing countries adopt more focused containment measures, the COVID-19 crisis causes a double disruption that "squeezes" statistical offices with 1) pressure of the demand-side data required by both governments and citizens and; 2) pressure from the supply side, as a result of limited operational and technical resources in statistical agencies, planning for a more or less stable generation of data in the short and medium terms. This would be precipitating interruptions in the value chain of official data and statistics. And according to some results for certain cases, the pressures would also anticipate results of excellence, such as when internal forces of the mountain turn coal into diamonds.

The reasonable dedication to leadership that the document "a world that counts" reserves for NSOs in the setting of the data revolution is work in progress. That at least is evident in the few cases of initiatives that show a reaction in the world to the emergence of the coronavirus. Needless to say, the initiatives best known to the general public have been contributed by actors not necessarily related to statistics from governments.

Leadership and resilience, especially in the face of external and emerging phenomena, still remain a pending issue for a significant number of NSOs. Reinforcing leadership and resilience of NSOs could not only help in the understanding of global - and predominantly territorial - phenomena such as pandemics but would also underpin an overarching vision of the effective contribution of collaboration with third party sources. After all, any effective leadership will require inclusiveness at its core.

## Seven preliminary lessons from the revised data associations

While a variety of initiatives have been identified from the global, continental and national/subnational levels, there is still ample opportunity to reflect impacts and issues by integrating diverse sources represented through geospatial dimensions. Such innovation could effectively help in the understanding of the territorial behaviour of the "monster" that a pandemic means. Insisting on a successful strategy in this regard will facilitate decision making resulting from the territorial processes involved in the pandemic, especially those that support the return to the new normality in the 188 infected countries. Some general and particular observations are detailed below in relation to the surveyed initiatives:

1. This preliminary survey has not identified - at least in those experiences of public access carried out by NSOs in Africa, the Americas and Europe - real-time geospatial responses revealing challenges related to the coronavirus, including individualized location of infected persons or follow-up of contacts maintained by those affected.
2. No cases have been verified where data integration has resulted in the transformation of existing or ongoing processes. In general, data integration involves data fusion exercises, transitory experiments and data laboratories using the ArcGis Hub, combining official sources with each other or adding new ones.
3. The reviewed experiences and geospatial response laboratories showed in a short time a great potential for integrating sources from different origins, suggesting that the necessary institutional arrangements could be solved also in the short term, especially when the necessary incentives are ensured.
4. Few initiatives involved private actors to consolidate data partnerships coming from NSOs. Although in the case of Vodafone-Orange/Flowminder examples were already available at an experimental level, it should be noted that it is not a common experience for NSOs to enter into data partnerships with business partners, especially not in developing countries. This fact promises opportunities for progress in the statistical field of Government2Business data partnerships.
5. Commendable efforts have been made by statistical offices, in particular in less developed countries, to disaggregate national data using administrative records or surveys from other official sectors of government. Such innovations demonstrate that in the context of emergency it is possible to move forward expeditiously in the data revolution as suggested by the United Nations five years ago.
6. The importance of having a clear and carefully targeted communication campaign to avoid generating further confusion among the population cannot be underestimated. In the case of the vulnerability viewer led by DANE in Colombia, for example, it was necessary a counter campaign using social networks for clarifying that the initiative did not present real time data but an analysis of pre-existing data before the pandemic. Also, in the case of INE in Spain further assurances of proper use of private data were needed.
7. Finally, no cases of data collaboration between civil society and NSOs have been found to provide geospatial responses to COVID-19. This kind of data partnerships are especially recommended in the 2030 agenda and are being actively studied for successful data partnerships. Considering the broad granularity of the analysed territorial problems, the inclusion of this vital partner in future NSO initiatives should be wisely promoted, not only by calling on other

official or private actors, but also academia and the civic sector, both users and producers of data.

The unprecedented collaboration between NSOs and sources from different origins provided by external actors is the most common denominator identified in the above observations. This anticipates real progress in the data revolution and motivates NSOs to further explore data partnerships with third parties to meet any future needs. A likely rationale for success is that these partnerships represent a Pareto Optimum exchange in the sense that they achieve an increase in the well-being of all by making better data available for decision making without decreasing the well-being of the partners providing these data. According to the reviewed research reports, the data partnerships have not implied major sacrifices or additional costs for any of the parties in their response to the coronavirus.

Taking this logic a step further, national statistical offices' leadership could be strengthened through the generalization of pioneering data partnerships. The geostatistical domain should capitalize on these costly lessons from the pandemic. Highly conceived strategies could provide the momentum needed to accelerate the delayed monitoring that will guide the achievement of the sustainable development goals called for in Agenda 2030 with a healthy emphasis on effective visualization and geographic disaggregation.