

8th International Monetary Fund Statistical Forum
Proposal Session No.1 “The economics of a pandemic: What are the data needs”
DANE’s- Per Block Vulnerability Index

Executive Summary:

The Covid-19 situation has accelerated some existing trends among the International Statistical System such as growing data demands, the emergence of new stakeholders and a constantly changing environment. The consolidation of these trends have led to National Statistical Offices to re-think their business model in order to strengthen their relation with the citizenship and other stakeholders which in sum led up to a constant search for innovation under a risk prone position.

From the executive branch’s perspective, taking into consideration the social isolation measures adopted in order to limit the spread of the Covid-19, one of the priorities was to identify the vulnerable populations that could benefit from state programs (whether they were cash transfer based or others).

This paper seeks to present how we managed to create a Per Block Vulnerability Index Geovisor, by combining variables such as sociodemographic conditions (population density and age of the population) and pre-existing comorbidities (e.g. people diagnosed with diabetes, chronic lung disease and others). Moreover, by sharing our specific experience on the matter, we will address the questions of what are the data needs? and how can the pandemic be an opportunity for innovation. The latter will allow us to highlight the need of building resilient and adaptive partnerships with key stakeholders (including the users) that have official statistics as a public asset at its core.

Official Statistics under the Covid Pandemic:

The Data revolution has implied a growing volume of information, a diversity of methodologies under an everchanging scenario in which the need of a solid evidence base for the decision-making process is highlighted. In addition, the unfolding of a global pandemic has reconfigured a context in which factors such as precision, quality and feasibility must be harmonized, and it has been the role of the National Statistical Offices as the country’s Data Steward to assume this task and assure that the user’s needs are satisfied.

Moreover, Covid-19 has transformed the socio-economic outlook and will have long lasting effects in various areas such as demand and consumption patterns, the global value chains, employment dynamics, income distribution, the social structures, and the environment. By understanding the wide scope of the crisis, the value of official statistics as a public asset has been highlighted; nonetheless, to achieve this, National Statistical

Offices must face technological, ethical and communicational challenges (data infrastructure, the integration of new sources such as geospatial information, big data among others). By doing so, NSO's will be able to have statistical information that not only meets the data needs of users but empowers them to create the enabling conditions for a policy dialogue under an economic reactivation context.

Dealing with these challenges requires to establish adaptative cultures as an opportunity to strengthen coordination spaces and mechanisms in the learning community that is the International Statistical Community and under which International Organizations play a major role as they are both scenarios and agents.

Precision, as one of the attributes/dimensions of Quality of official statistics, can be attained by adding methodological thoroughness implemented by researchers, which implies a long and methodic process, it requires also to have interdisciplinary task teams allowing to have an integral perspective. Nonetheless with the Covid-19 pandemic, a new challenge emerged regarding the need to give concrete answers and inputs to a complex issue of which no documentation or prior experience in the LAC region could guide our efforts. As professionals in, and associated with, statistics the clear answer to this was to focus on the data, and let the data, its correlations, locations and trends speak to us. But then again, we encountered a new and wider challenge, which data should we use? which type of data could we use? Which is the correct frequency of the data? And finally, who *owns* this data?

This set of questions were initially addressed by the National Administrative Department of Statistics (DANE in Spanish) thanks to the strengths of the technical team, and afterwards were socialized with relevant stakeholders amid the pandemic such as the National Planning Department and the National Health Institute. Some of the issues that we had to evaluate before engaging with others was the assessment of our storage and processing capacities, the availability and quality of the information that we needed and specially the definition of working schedules and timelines that allowed for an efficient and effective planning.

As a governmental agency and as owner of the official statistics of the country, DANE made available all of its infrastructure, information and human resources to help in the construction of analysis that allowed the decision makers, at the national and local level, to better understand the crisis and create policies that were fit for purpose.

The Per Block Vulnerability Index:

At the beginning of 2019, DANE presented to the country the results of its largest statistical operation, the National Population and Housing Census (CNPV 2018) which marked important milestones such as the design and implementation of an electronic census, the innovation in the logistics domain by designing a "routes method" allowing to increase the coverage of the statistical operation, an ethnic approach and the collection personal identification variables. The latter would become one of the major assets of the

NPHC 2018 due to the ease of integration with other statistical operations such as household surveys and administrative records. In addition to the personal identification, the composition of the households and the places of residence immediately became the information in greatest demand when the pandemic began.

The availability of this input as well as the juncture, allowed for a reinforced collaboration between DANE, the National Planning Department, the Ministry of Health and Social Protection and the Analytical Unit of the Institute for Health Technology Evaluation (IETS) in the characterization of the population that presented a bigger risk if infected with covid-19 and that would require a higher priority in the medical services of hospitalization and intensive care. To this end, we based our efforts in the sociodemographic characteristics of the deceased at the global level to identify patterns regarding age, sex, comorbidities, and conditions of habitability.

In this manner, all adults over 60 years of age were identified as susceptible people; Also, people with preexistence of hypertension, diabetes, chronic respiratory diseases, ischemic heart disease and any type of cancer make up some of the individual characteristics most frequently presented by those who died from Covid-19. Additionally, if a person had these characteristics, the interaction with the groups with the greatest contagion could put them at greater risk, therefore, and thanks to the composition of the household provided by the NPHC 2018 it was possible to locate the elderly living with people between 20 and 30 years old (with a higher risk of contagion) and those who coexisted with people between 30 and 50 years old (with a medium risk of contagion).

Once the isolation measures were taken at the national and local level, older adults who live alone or live with people who belong to their family nucleus would be at risk, since they must necessarily expose themselves to be able to stock up on food. Consequently, interactions of any kind could increase the risk of contagion, so determining the places with the highest demographic density, as well as overcrowded conditions in rooms and bedrooms, could help in planning and distribution of resources.

This understanding and the availability of the NPHC 2018 data, the records of the Unique Database of Health Affiliation (*Base de Datos Única de Afiliación en Salud* -BDUA) and individual records of health service provision (*Registros Individuales de Prestación de Servicios de SALUD* -RIPS), allowed, through deterministic integration (by identification document) the construction of 13 indicators calculated at block levels in all the municipal seats of the 1,122 municipalities in the country where according to the CNPV 2018, 75.5% of the population resides, which represents approximately 36.5 million people. With this information, the next step consisted in determining the areas with the greatest vulnerability, that is, the areas with the highest levels in the constructed indicators, which led to what was called the Vulnerability Index.

This index aims to help determine the volume and location of people with the greatest susceptibility to the virus and therefore prepare for the possible saturation of the health system. Through an interactive geovisualization platform it not only helps decision makers but also contributes in the generation of self-care awareness because the results were

available to all people, where they could easily find their residence block and understand how They and their community could be affected if they virus spread in their area.

The importance and advantages of the spatialization of thematic variables from statistical information

Under the premise that everything that happens, happens somewhere, the relation between the different natural and socio-economic phenomena with the space in which these unfold, results in an indissoluble link, without which it would be impossible to analyze and, therefore, understand the behavior of these phenomena from its distribution in the geographical space, where the interactions between the human being and the surrounding environment are established.

Although the information as a whole is considered a fundamental asset for decision-making, the information resulting from statistical operations (census or samples) is of great relevance due to the rigorous methodologies used to obtain it, from operational and sample designs that allow to determine or infer the main characteristics of a target population.

Depending on the level of geographical / territorial disaggregation, the dissemination of the results obtained from the statistical operations is generally done alphanumerically through tables and graphs that synthesize the information. If the integration of statistical and geospatial information is added to the above, through the spatialization and visualization of statistical information through thematic maps, effective communication instruments are obtained for understanding trends and patterns of the different phenomena of interest.

Moreover, through the development, use and updating of areal statistical frameworks such as the National Geostatistical Framework (MGN) of DANE, we have developed a robust instrument with transversal support throughout the entire statistical process. Therefore, although the MGN is built for statistical purposes, it is not only the basis for collecting data in the sample designs and census operations, but it is seen as the main tool for the spatialized dissemination of statistical information at different geographic scales: from the general aggregations associated with the political-administrative division of the country, to the more specific disaggregations in rural and urban contexts, such as sections and blocks, respectively.

Said disaggregation allows to explore, in a more focused way, the detailed spatial behavior of the different thematic variables constructed from the source information, guaranteeing, in turn, the statistical reserve by means of the anonymization of the information.

Based on the above, one of the greatest achievements obtained by DANE in its mission of producing information for everyone has been to respond both to the needs of different types of users and to the increasingly substantial challenges of the current situation. Under these premises, we have developed an interactive geographic viewer of the Vulnerability Index of census source by block, with the collaboration of the National Planning

On the other hand, the integration of statistical and geospatial information is not only materialized through the visualization of thematic variables through maps but is also enhanced through the analyzes that are derived to understand our current geographical reality. In this sense, spatialized indicators can be exploited analytically, through spatial statistical processes, in order to determine significant spatial groupings where decision-making and public policy could be focused. Therefore, the construction of the indicator and its geo-visualization becomes the first phase of support for new methodological and analytical advances.

It is also important that under the current juncture, as the country's Data Steward we are facing challenges in terms of mindset, as to how can we change how data is perceived; in terms of technological capacity and innovation, as it is important to be able to have the infrastructure to receive, harmonize and dispose strategic information in a safe and efficient way, and to implement new methodologies in statistical processes, understanding the rapid technological change in societies; and from the organization/institutional standpoint as we have to ensure that the functional units of the organization have processes and procedures that allow the transfer of capabilities among the National Statistical System members as well as the quality assurance in the different data sources.

In conclusion, the pandemic can be seen as a window of opportunity for innovation from the statistical perspective, as it has fostered the strengthening of traditional partnerships and highlighted the value of statistics as a public asset.

As we enter in a post Covid-19 scenario, new challenges will arise and some older ones will remain such as the need to design new working arrangements to make statistical systems more resilient in order to pivot towards a more integrated and multidimensional measurement of development and the economy , in particular by including the wellbeing of citizens and the impact of the economy in the environment. Some of the questions that we would seek to discuss with the audience include: How can the Data Stewardship approach foster an appropriate transition from an emergency/urgency situation to the post-Covid scenarios? What are the main challenges regarding the strengthening of our work on experimental statistics? How did the data needs, under an ongoing data revolution, evolved during the Covid-19 pandemic? And how did NSO's adapt to remain fit for purpose.