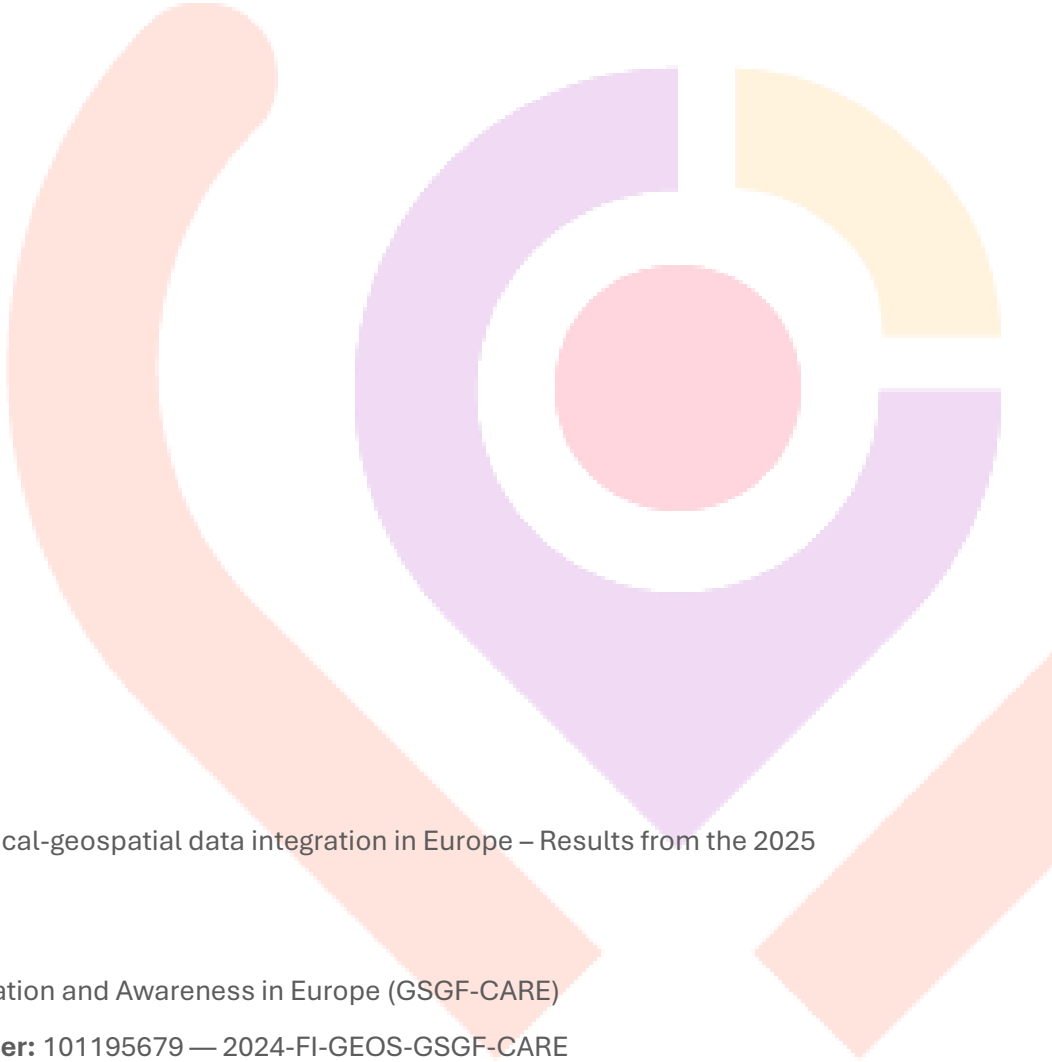




Maturity for statistical- geospatial data integration in Europe

Results from the 2025 GSGF-
CARE Survey



Title: Maturity for statistical-geospatial data integration in Europe – Results from the 2025 GSGF-CARE Survey

Author: GSGF-CARE

Project: GSGF Collaboration and Awareness in Europe (GSGF-CARE)

Grant agreement number: 101195679 — 2024-FI-GEOS-GSGF-CARE

It is permitted to copy and reproduce the content in this report. When quoting, please state the source.

© GSGF-CARE 2025

Acknowledgments

On behalf of the GSGF-CARE project and the European Forum for Geography and Statistics (EFGS), we would like to express our sincerest gratitude to all of you who kindly responded to the survey. We recognise that some questions may have been ambiguous or challenging to answer, and we truly appreciate the patience and cooperation of those who also responded to follow-up questions and requests for clarification.

Thanks to your dedicated support and valuable input, we achieved a strong response rate for this third round of the survey - following the previous rounds conducted by the GEOSTAT1 project in 2011, the GEOSTAT 2 project in 2015 and the GEOSTAT 4 in 2020. This enables us not only to provide a comprehensive overview of the current situation across Europe, but also to track progress over time.

We believe the insights gained from this survey will be of lasting value to the statistical-geospatial community—well beyond the GSGF-CARE project team.

With sincere thanks,

The GSGF-CARE Consortium

Content

Acknowledgments	1
Introduction.....	3
Survey Results	5
Question 1A – Availability and quality of authoritative and fundamental geospatial data	5
Question 1B – Development of authoritative and fundamental geospatial data	10
Question 1C – Usage of authoritative and fundamental geospatial data	16
Question 2 – Geospatial enabling capacity	20
Question 3 – Geospatial data management in statistical production integration	23
Question 4 – Integration of geospatial information in statistical production	25
Question 5 – Geospatial data management and processing.....	26
Question 6 – Obstacles for data integration	28
Question 7 – Administrative data sources	30
Question 8 – Building capacity for statistical-geospatial data integration	31
Question 9 – Areas for improvement.....	33
Question 10 – Advancements and good practice.....	35
Question 11 – Production infrastructure for statistical-geospatial data integration	36
Conclusions	38
Data availability and data quality	38
Streamlining of geospatial information in statistical production	39
Obstacles and capacity needs.....	40
References	42

Introduction

As part of the GSGF-CARE project, a survey was conducted in May 2025 to evaluate the capabilities for integrating geospatial and statistical information across the European Statistical System (ESS) and to help identify the challenges and obstacles that both organisations and countries face in advancing statistical-geospatial data integration. In particular, this survey was carried out as part of the project's activities aimed at monitoring and assessing the maturity for statistical-geospatial data integration in the ESS.

Furthermore, the results of the survey constitute a key output of the project and serve as input that directly informs and shapes subsequent project activities, especially those related to capacity building and capability development, contributing to progress in this field of action across the European context.

A web-based questionnaire (provided by Statistics Finland) was sent to all the National Statistical Institutes (NSIs) representatives of the GISCO Working Group. The questionnaire comprised 13 questions with fixed response options. Some of the questions were designed to allow a follow up on surveys conducted previously by the GEOSTAT projects and to measure progress made in organisations and/or countries. However, there were also newly designed questions not being asked before that enabled to better provide insights and identify trends about the current situation on statistical-geospatial data integration.

Although the survey was primarily addressed to NSIs, respondents were encouraged to consult with their geospatial counterparts in National Geospatial Information Agencies (NGIAs) before submitting their responses, particularly regarding the first questions on the availability and quality of authoritative geospatial data

The survey targeted all members of the ESS, European Union (EU) candidate countries and the former EU Member State United Kingdom. In total, 42 countries were included in the target population. Since the GEOSTAT 4 survey, the target population has been extended to Ukraine and Georgia.

The overall response rate was very high, with 39 respondent countries out of 42 target countries. The following non-responding countries in the target group were Liechtenstein, Malta and Montenegro. Among the EU Member States (EU27), the response rate was 96%. Malta was the only non-responding country in this group. Among the 31 ESS countries (EU27 + EFTA: Iceland, Lichtenstein, Norway and Switzerland) the response rate was 94%. In this group, Malta and Lichtenstein were the only non-responding countries.

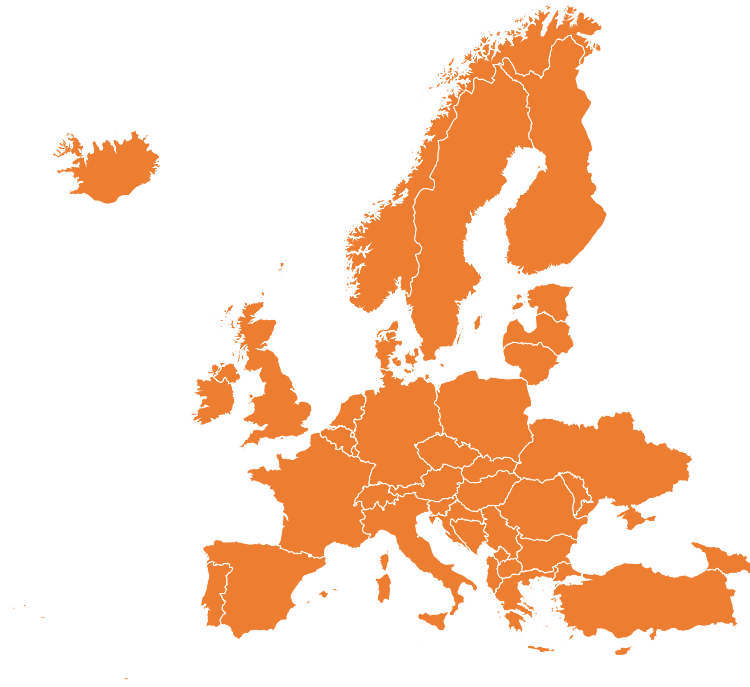


Figure 1. Map of the responding countries as of June 13, 2025.

Compared to the GEOSTAT 4 survey, conducted in 2020 (GEOSTAT 4, 2020), the response rate of the 2025 survey was higher for the whole target population. In the GEOSTAT 4 survey 6 countries out of the 40 countries in the target population did not respond (GEOSTAT 4, 2020).

Survey Results

Question 1A – Availability and quality of authoritative and fundamental geospatial data

The availability and quality of authoritative and fundamental geospatial data have historically been a challenge.

Previous surveys conducted within the series of GEOSTAT projects (GEOSTAT 1 in 2011, GEOSTAT 2 in 2016 and GEOSTAT 4 in 2020) as well as discussions within the GISCO Working Group have consistently highlighted that the lack of high-quality geospatial data is a major barrier to progress in data integration.

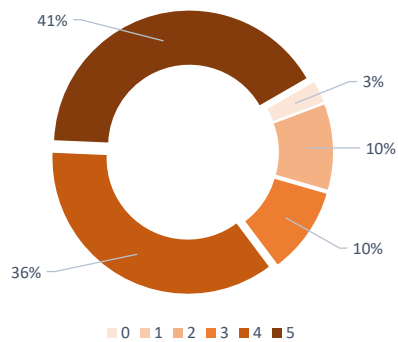
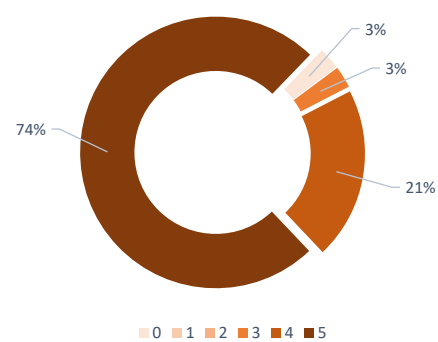
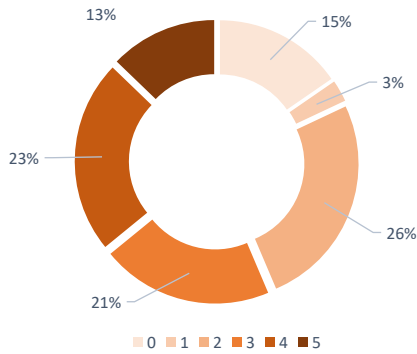
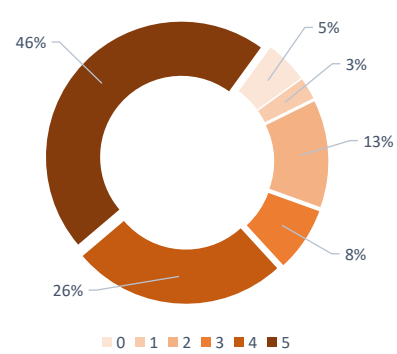
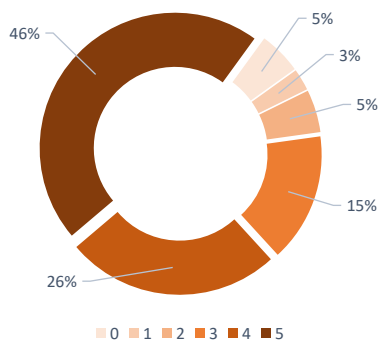
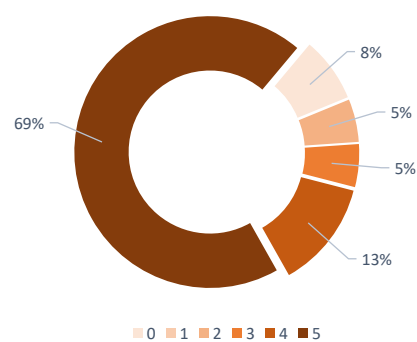
In **Question 1A**, respondents were asked to rate, on a scale from 0 to 5, the current availability and quality of authoritative fundamental geospatial data in their country. Each fundamental data theme was to be scored based on the list of Core data provided by UN-GGIM: Europe (UN-GGIM: Europe, 2016). The following **scoring guide** was provided:

- 0 = Not aware/Not relevant
- 1 = Does not exist
- 2 = Partial coverage exists with quality issues
- 3 = Partial coverage exists with good quality
- 4 = Full coverage exists with quality issues
- 5 = Full coverage exists with good quality

Table 1. Availability and quality of authoritative and fundamental geospatial data by core data theme

Data Themes	0	1	2	3	4	5
Addresses	3%	0%	10%	10%	36%	41%
Administrative Units	3%	0%	0%	3%	21%	74%
Basic Services	15%	3%	26%	21%	23%	13%
Buildings	5%	3%	13%	8%	26%	46%
Cadastral Parcels	5%	3%	5%	15%	26%	46%
Elevation	8%	0%	5%	5%	13%	69%
Geographical Names	3%	0%	3%	8%	21%	67%

Hydrography	13%	0%	8%	3%	28%	49%
Land Cover	10%	3%	5%	8%	31%	44%
Land Use	3%	3%	21%	15%	28%	31%
Ortho Imagery	3%	0%	3%	8%	21%	67%
Regulated or Managed Zones	31%	5%	15%	5%	28%	15%
Statistical Units	3%	5%	5%	3%	8%	77%
Transport Networks	13%	0%	18%	5%	28%	36%

Addresses

Administrative Units

Basic Services

Buildings

Cadastral Parcels

Elevation


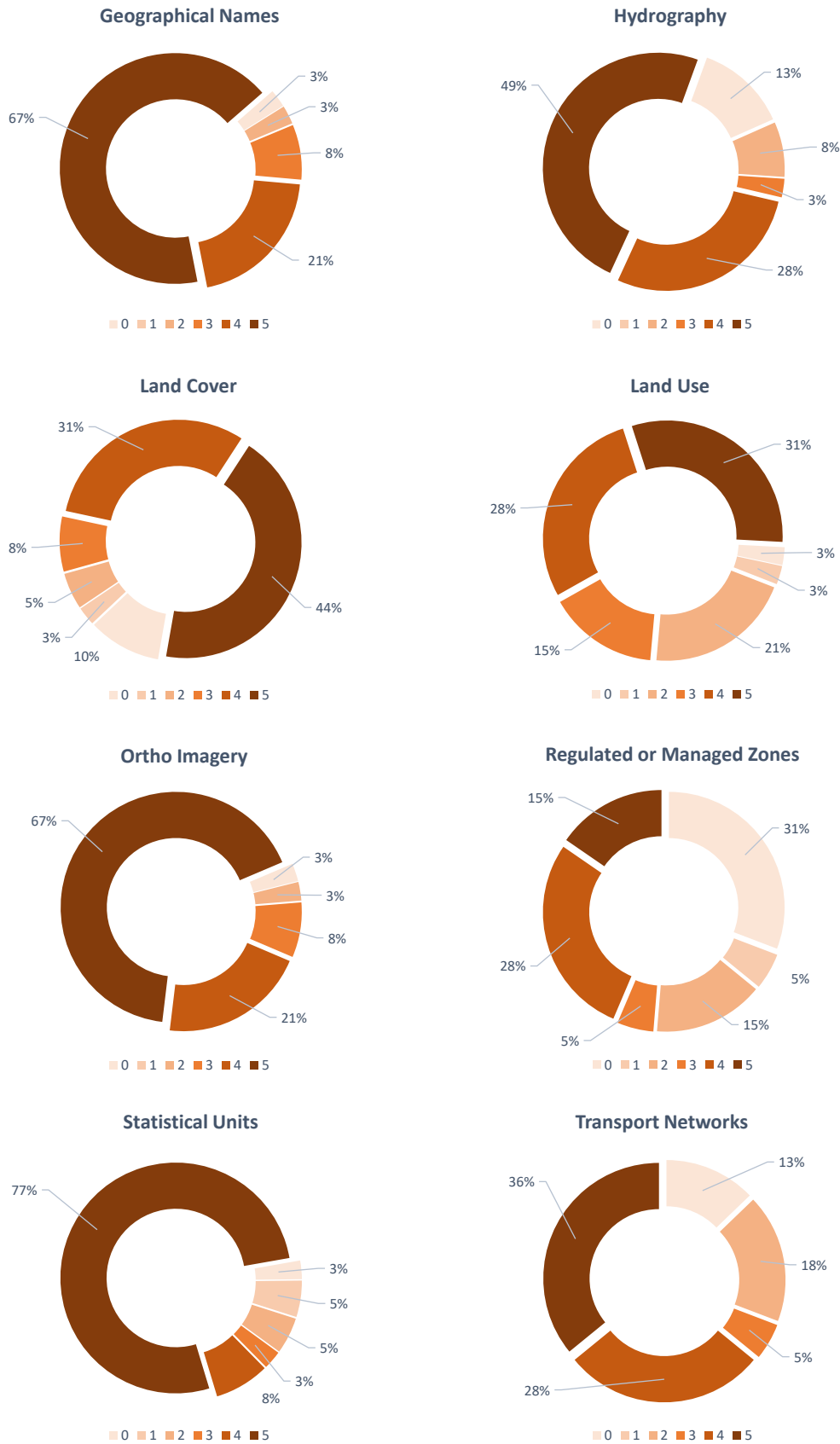


Figure 2. Availability and quality of authoritative and fundamental geospatial data by core data theme

From the responses it is suitable to claim that the availability and quality of many data themes are fairly good (10 out of 14 have a score equal or more than the average, 3.8). The two response options indicating full coverage - whether with quality issues or good quality (4 = Full coverage exists with quality issues and 5 = Full coverage exists with good quality) - account for 70% or more of the responses for the majority of data themes. Only *Basic Services* and *Regulated and Managed Zones* received scores below 50%.

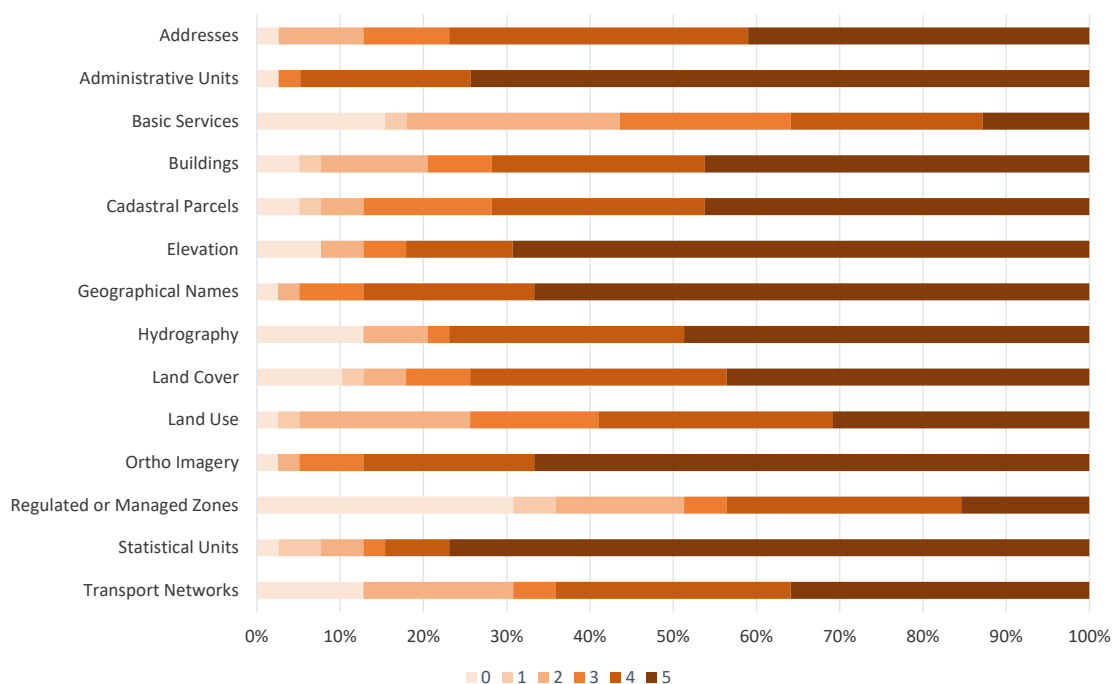


Figure 3. Availability and quality of authoritative and fundamental geospatial data by core data theme

Although, what is a bit concerning is that data themes most critical for NSIs to geospatially enable micro data, such as *Addresses*, *Buildings* and *Cadastral Parcels* objectively have quality issues. Only 41% of the respondents reported that *Address* data have full coverage and good quality, and *Buildings* and *Cadastral Parcels* show similar values (46%).

The data themes with the highest scores in terms of availability and quality were *Administrative Units* and *Statistical Units*, these data themes are considered by 74% and 77%, respectively, of respondents to have full coverage with good quality. Also, *Elevation Data* and *Ortho Imagery* score relatively high score (69% and 67%, respectively). These scores are very promising results as observation data is being considered an important data source in official statistics and increasingly used to support and modernise statistical production, namely by filling data (quality) gaps issues (e.g., spatial resolution and

temporal coverage), reducing costs and respondent burden and enabling faster and timely statistics.

In contrast, the *Regulated or Managed Zones* and *Basic Services* themes stood out with exceptionally low scores, wherein only 15% and 13% of respondents, respectively, indicated that these themes have full coverage and good quality. According to the UN-GGIM Europe specification (UN-GGIM: Europe, 2018), *Basic Services* encompasses a range of different feature types¹. It is likely that this theme presents particular challenges because it typically does not consist of single-source data. Instead, it requires collaboration among various public authorities and governmental organisations, as well as the integration of data from administrative sources (e.g., harmonisation challenges). However, this data theme has better availability and quality conditions compared to *Regulated or Managed Zones*, as more respondents scored between 2 and 4 (partial coverage to full coverage with some quality issues) for the theme *Basic Services*.

It is important to highlight that *Regulated or Managed Zones* was the data theme that most of the respondents (31%) indicated that they are not aware of the related availability and quality, or concluded that such data is not relevant for statistical purposes (e.g., option 0). For *Basic Services*, the corresponding figure was 15% followed by *Hydrography* and *Transport Network* with 13%.

¹ Core feature type Basic Service should include at least the energy supply services (power plants of various sources: thermic, nuclear, hydro-electric, etc.), waste management facilities (disposal sites and landfills, waste water management plants, waste management plants), health services (hospitals, emergency medical services), education services (from pre-primary schools to Universities), safety services (police station, fire station, rescue station, defence site, civil protection site, court or tribunal, prison), leisure services (sport facility, camping site, public park, zoo, library, museum, cultural centre, recreation facility, etc.), government services (such as town halls or embassies), social services (cemeteries, residence for elders, etc.) and economic services (post offices, etc.).

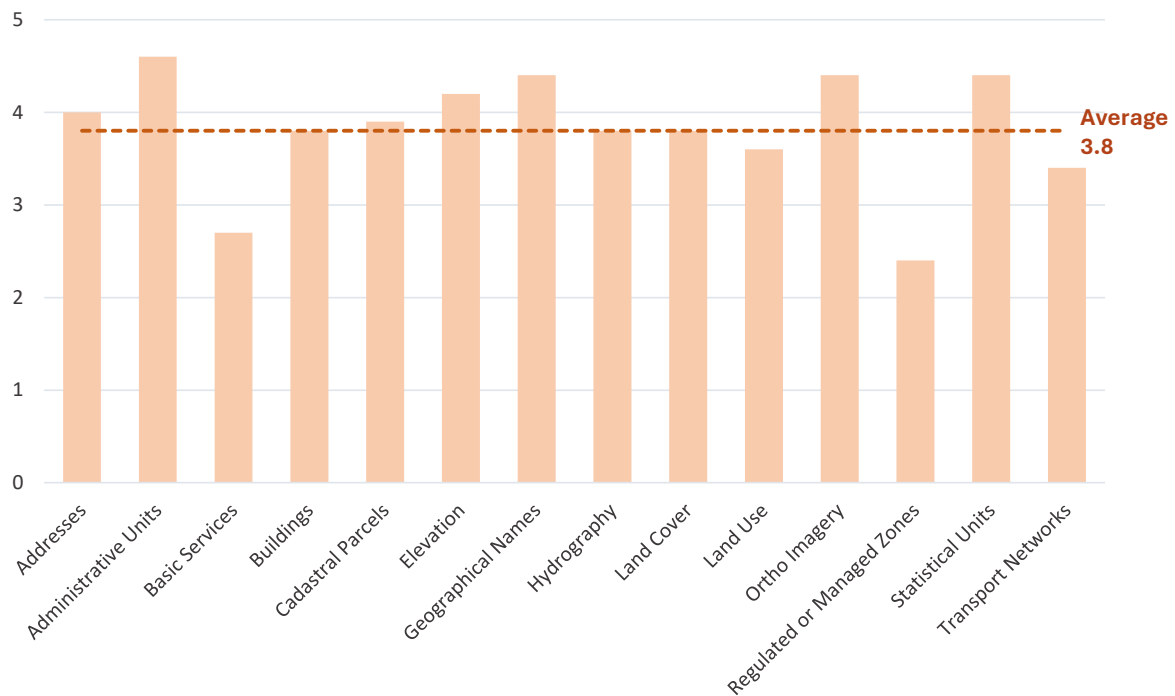


Figure 4. Average on the availability and quality of authoritative and fundamental geospatial data by core data theme

In terms of average availability and quality by data theme, *Basic Services*, *Land Use*, *Regulated or Managed Areas* and *Transport Networks* were the ones that registered average score values below the overall average for all data themes (3.8).

Question 1B – Development of authoritative and fundamental geospatial data

Question 1B is related to Question 1A and aims to capture how the situation has developed over time regarding authoritative and fundamental geospatial data. The respondents were asked to score (between 0 and 4), the changes in coverage and quality of authoritative and fundamental geospatial data over the past 5 years.

The respondents were asked again to score each fundamental data theme according to the list provided by UN-GGIM (UN-GGIM: Europe, 2016). The following **scoring guide** was provided:

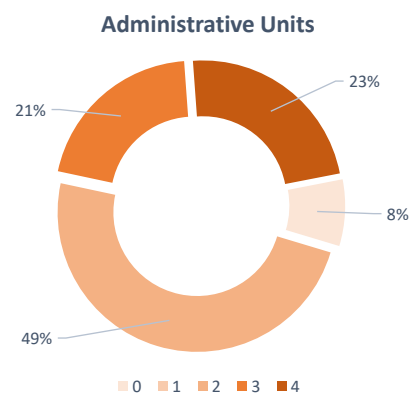
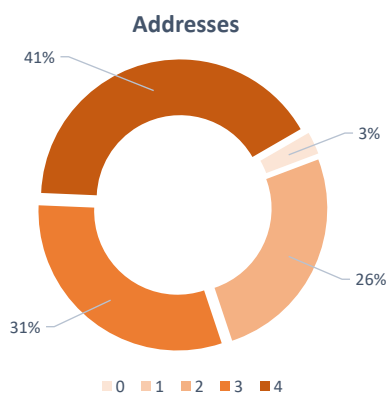
- 0 = Not aware/Not relevant
- 1 = Situation is worse compared to 5 years ago
- 2 = Situation is unchanged compared to 5 years ago

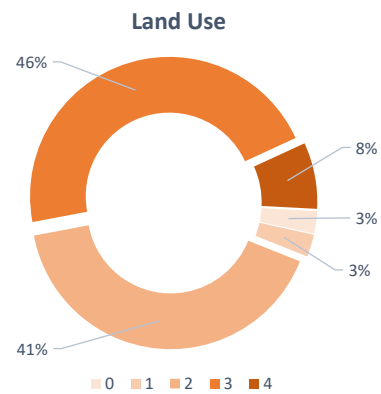
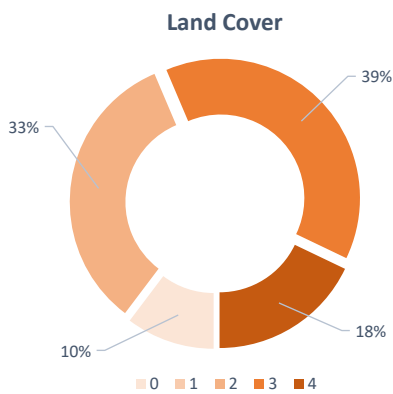
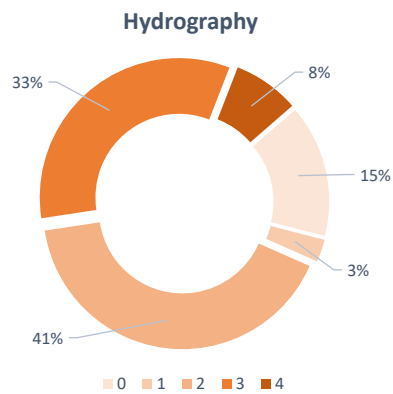
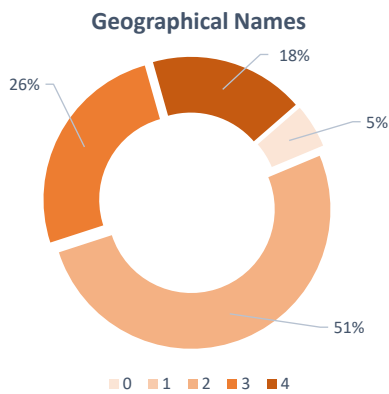
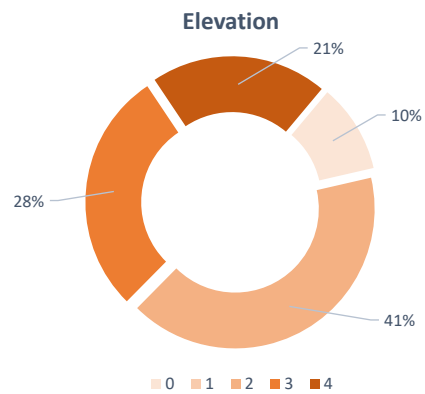
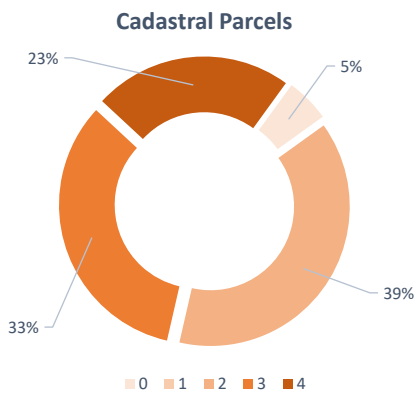
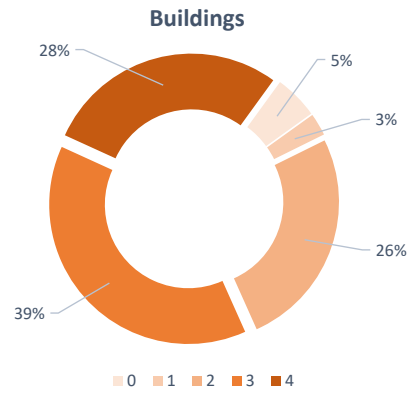
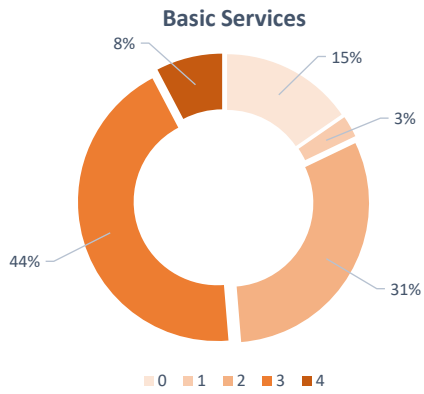
3 = Some improvements can be noted compared to 5 years ago

4 = Significant improvements can be noted compared to 5 years ago

Table 2. Development of authoritative and fundamental geospatial data by core data theme

Data Themes	0	1	2	3	4
Addresses	3%	0%	26%	31%	41%
Administrative Units	8%	0%	49%	21%	23%
Basic Services	15%	3%	31%	44%	8%
Buildings	5%	3%	26%	39%	28%
Cadastral Parcels	5%	0%	39%	33%	23%
Elevation	10%	0%	41%	28%	21%
Geographical Names	5%	0%	51%	26%	18%
Hydrography	15%	3%	41%	33%	8%
Land Cover	10%	0%	33%	39%	18%
Land Use	3%	3%	41%	46%	8%
Ortho Imagery	5%	0%	33%	39%	23%
Regulated or Managed Zones	31%	0%	26%	31%	13%
Statistical Units	8%	0%	41%	23%	28%
Transport Networks	15%	0%	31%	36%	18%





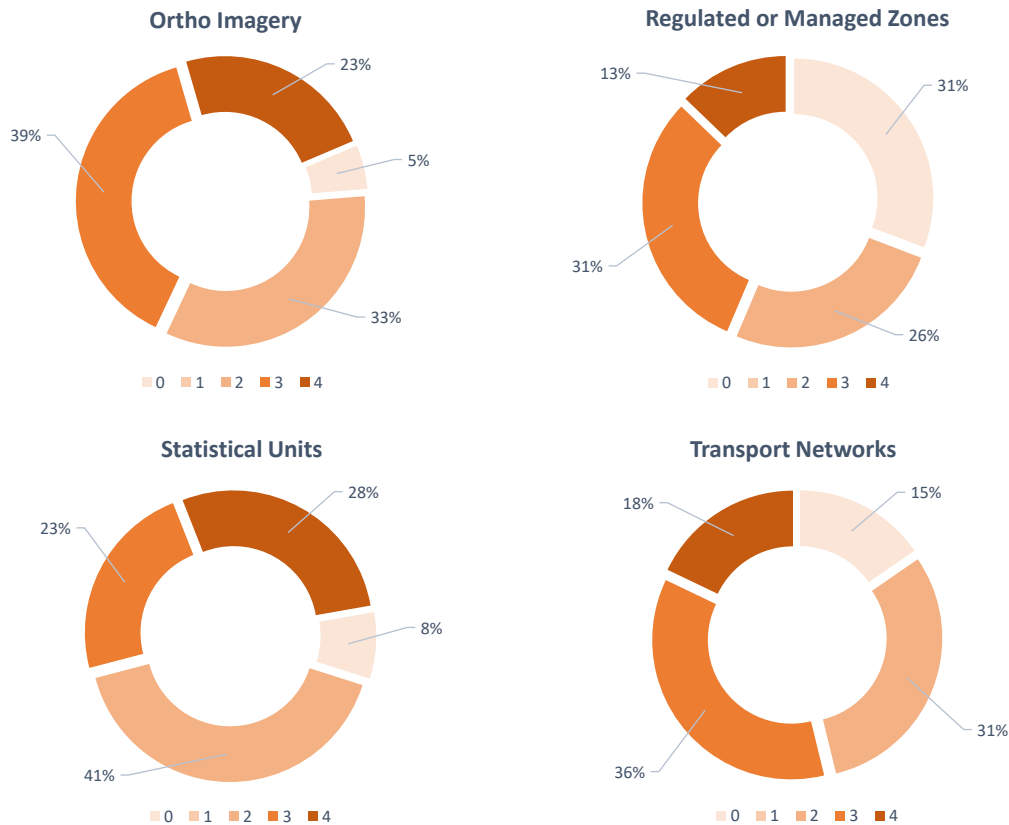


Figure 5. Development of authoritative and fundamental geospatial data by core data theme

The data theme that has been subject to the most significant improvement over the last 5 years is *Addresses*, followed by *Buildings* and *Statistical Units*. Around 41% of the respondents reported that they have noted *significant* improvements in address data compared to 5 years ago. Another 31% reported that they have seen *some* improvement. Summing up both *significant* improvements and *some* improvement, there is an improvement trend in 72% of the responding countries.

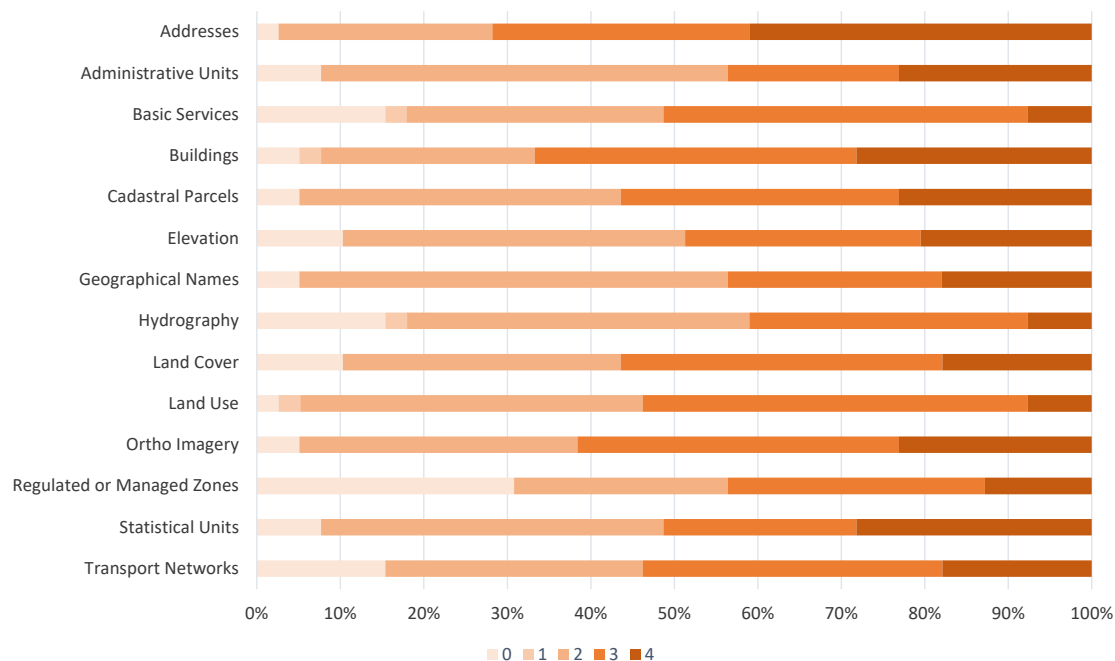


Figure 6. Development of authoritative and fundamental geospatial data by core data theme

Taking into account that only 41% of the respondent in Question 1A consider address data in their country currently to have full coverage and good quality, this must be considered as a very promising development for this data theme

In the scope of location data used to geocode statistics, building data also seems to have a positive development over the past 5 years. The share of countries that has seen a *significant* improvement is less compared to address data (28% compared to 41%), but if summing up both *significant* improvements and *some* improvement, there is an improvement trend in 66% of the responding countries. It is, by all means, very promising that *Addresses, Buildings, and Cadastral Parcels* are data themes that have shown a strong development in the last 5 years, as they are geospatial datasets with high priority for official statistics, namely to geocode statistical information and administrative registers at the unit record level (micro data). However, it should be noted that the improvements relate to a rather weak baseline. Even after improvements, the coverage and quality of these data themes remain questionable in many countries. It is absolutely critical that improvements of coverage and quality continue in the years to come.

Despite seemingly going in the right direction, some data themes are more problematic because of the weak baseline and lack of development support. For example, according to responses on Question 1A, only 31% of the countries have *Land Use* data with full coverage and good quality. Bearing this in mind, a *significant* improvement of *Land Use* data in only

8% of the countries and some improvements in additionally 46% of the respondents is a bit concerning. *Transport Networks* data showed a similar pattern having 18% that noted *significant* improvements and 36% *some* improvements

In addition, although *Basic Services* registered the low scores in the Question 1A, 44% of the respondents noted some improvements in this data theme compared to 5 years ago, which demonstrates a very positive change. Summing up, 51% of the respondents noted *significant* improvements and *some* improvements in this data theme, whereas 44% applied to the *Regulated or Managed Zones* data.

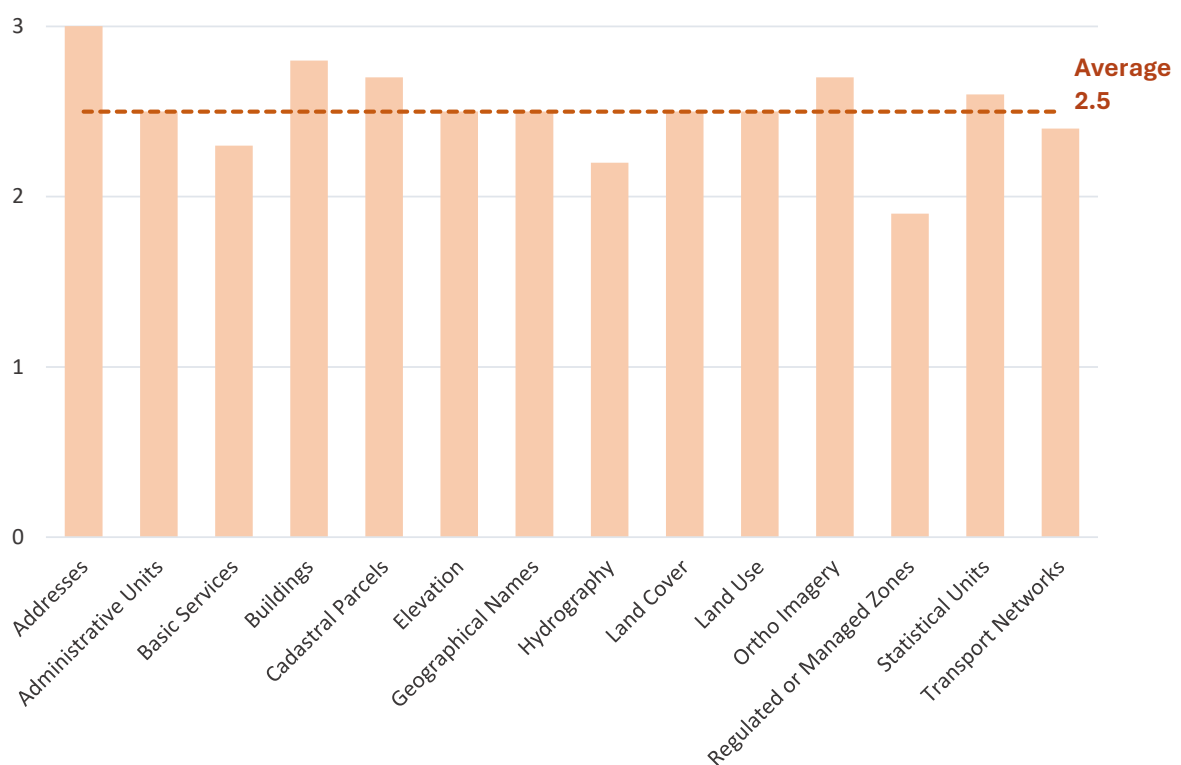


Figure 7. Average on the development of authoritative and fundamental geospatial data by core data theme

In terms of average on development by data theme, *Basic Services*, *Hydrography*, *Regulated or Managed Areas* and *Transport Networks* were the ones that registered average score values below the overall average for all data themes (2.5), highlighting the lowest average on the *Regulated or Managed Areas* (1.9). *Administrative Units*, *Elevation*, *Geographical Names*, *Land Cover* and *Land Use* are the data themes that reported being in overall average on the development of fundamental geospatial data (2.5).

Question 1C – Usage of authoritative and fundamental geospatial data

Question 1C is also related to the previous questions (1A and 1B) and aims to explore to what degree authoritative and fundamental geospatial data is effectively used by NSIs in statistics production processes. Respondents were asked to score to what extent their organisations are using and integrating authoritative and fundamental geospatial data in their statistical production processes. Again, the respondents were asked to score each fundamental data theme according to the list provided by UN-GGIM. The following **scoring guide** was provided:

0 = Not aware/Not at all

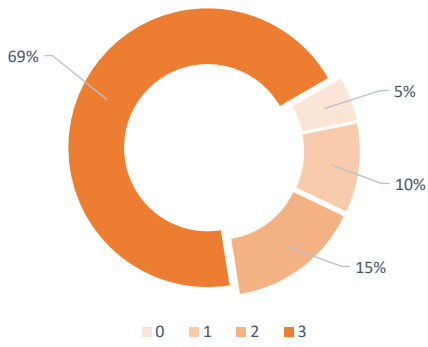
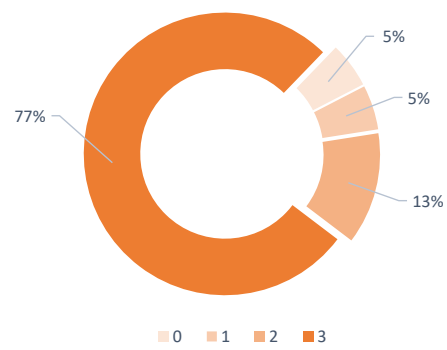
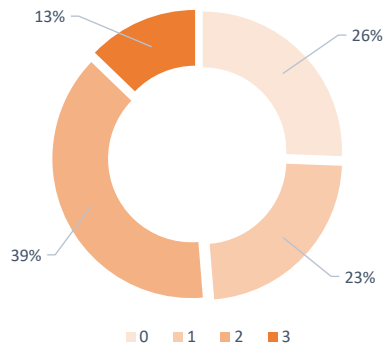
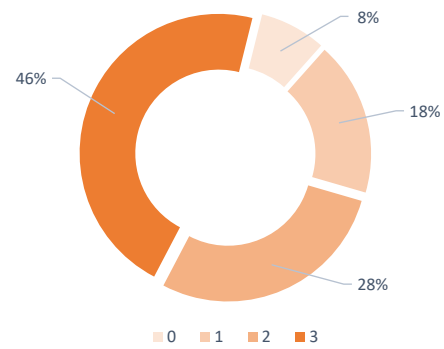
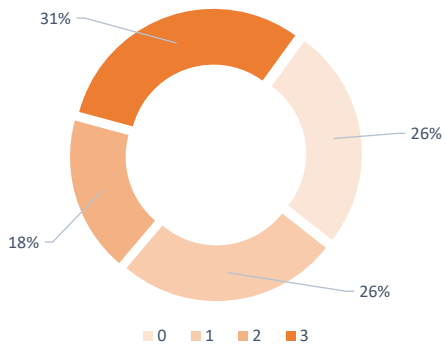
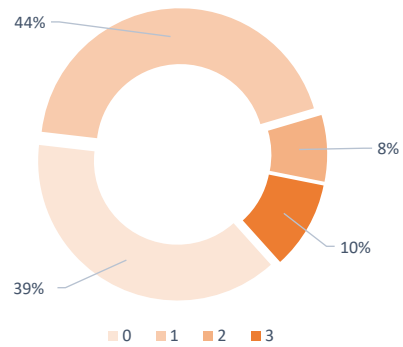
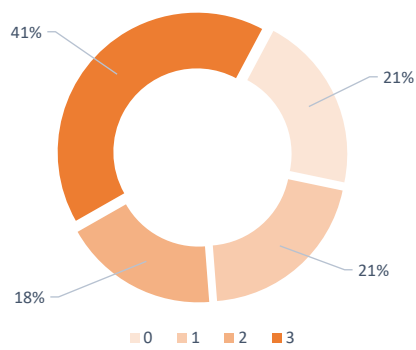
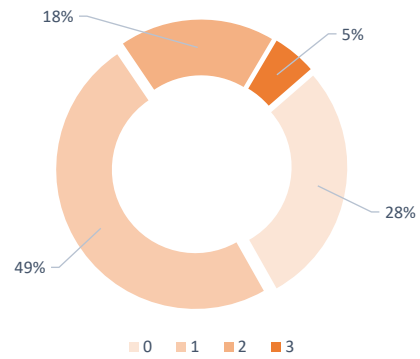
1 = Limited, occasional and/or tailored use

2 = Frequent use, mostly on an experimental and/or project-based level

3 = Frequent use, streamlined into the business production model and/or integrated into statistical outputs

Table 3. Usage of authoritative and fundamental geospatial data by core data theme

Data Themes	0	1	2	3
Addresses	5%	10%	15%	69%
Administrative Units	5%	5%	13%	77%
Basic Services	26%	23%	39%	13%
Buildings	8%	18%	28%	46%
Cadastral Parcels	26%	26%	18%	31%
Elevation	39%	44%	8%	10%
Geographical Names	21%	21%	18%	41%
Hydrography	28%	49%	18%	5%
Land Cover	18%	33%	31%	18%
Land Use	18%	36%	26%	21%
Ortho Imagery	10%	28%	28%	33%
Regulated or Managed Zones	44%	28%	23%	5%
Statistical Units	5%	8%	10%	77%
Transport Networks	21%	31%	28%	21%

Addresses

Administrative Units

Basic Services

Buildings

Cadastral Parcels

Elevation

Geographical Names

Hydrography


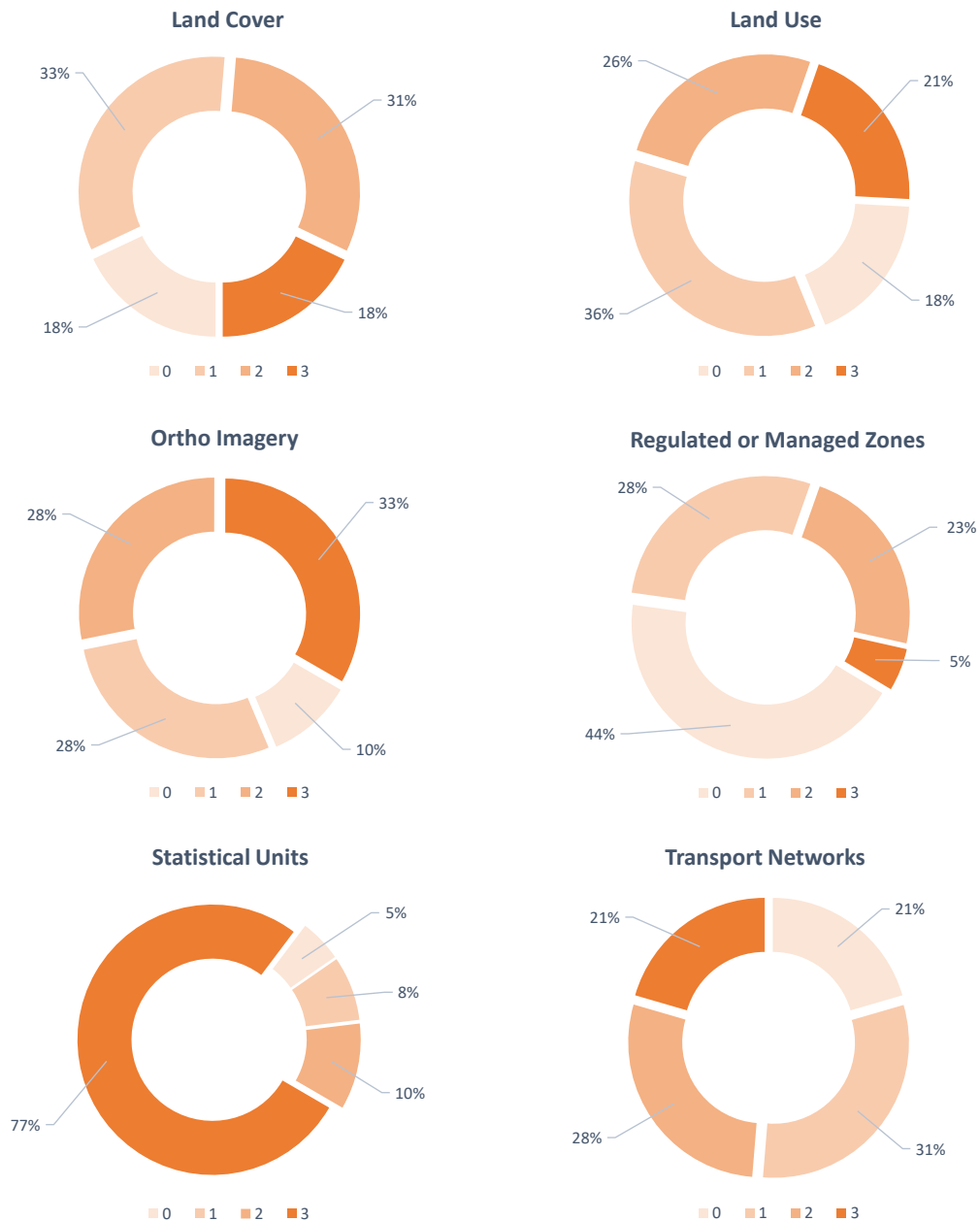


Figure 8. Usage of authoritative and fundamental geospatial data by core data theme

Not surprisingly, the fundamental and authoritative geospatial data most frequently and systematically integrated into the statistical production process are *Administrative Units* (77%), *Statistical Units* (77%), and *Addresses* (69%). The use of these data themes reflect a core capability present in almost all NSIs: **geocoding** (based on address data) and **aggregation** (based on common geographies, including administrative and statistical units) of unit record data - most notably population and census data, as well as **dissemination** (using output geographies to represent geospatial statistics).

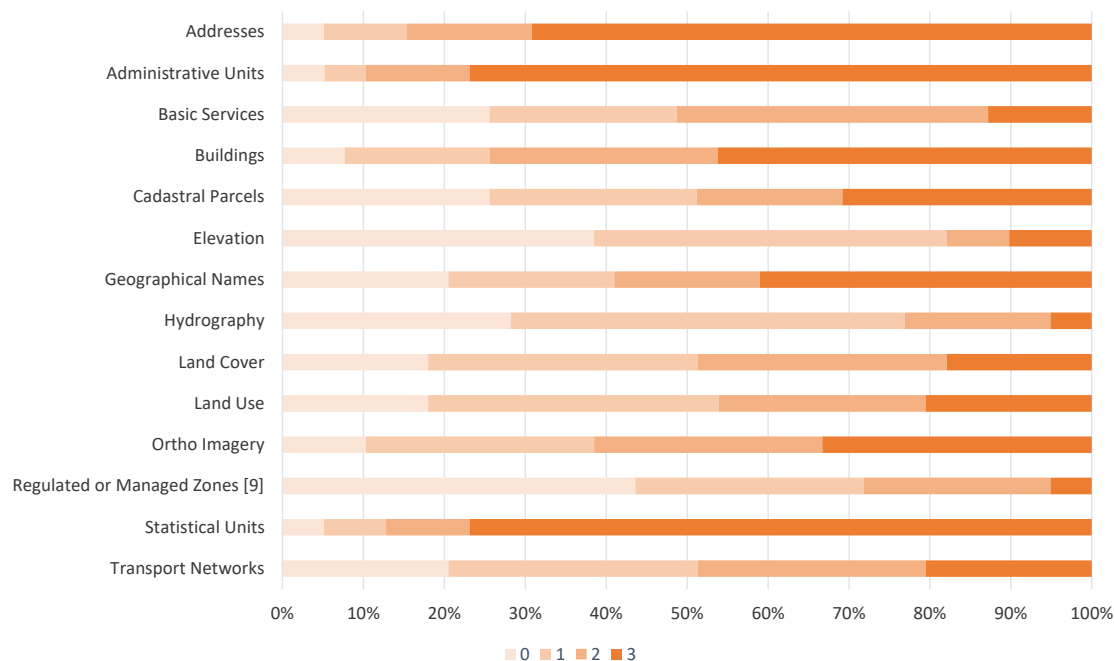


Figure 9. Usage of authoritative and fundamental geospatial data by core data theme

The relatively low scores for data themes, such as *Land Use*, *Land Cover*, *Hydrography*, and *Regulated or Managed Zones*, likely reflect the fact that environmental statistics are not within the mandate of all NSIs, but rather within the mandate of other national organisations. In this regard, NSIs that do not systematically produce environmental statistics and related products, have little reason to integrate these types of data into their business production models. In opposite, more unexpected is the relatively low systematic use of *Buildings* (46%) and *Transport Networks* (21%), the first one as standard location data typically used for geocoding and aggregation of statistical or administrative data, and the second one typically used to support routing and accessibility calculations.

The values of the *Cadastral Parcels* and *Ortho Imagery* (31% and 33%, respectively) were quite surprising when analysed together as the first data theme was expected to be significantly more used for statistical production than the second one due to its key relevance to statistics (e.g., geocoding and value of the geospatial object that goes beyond its location) and for the (national) geospatial infrastructure.

On one hand, it clearly emphasises the growing importance of Earth Observation (EO) data in official statistics and its use by the NSIs to address current and emerging data needs and develop innovative statistical products (e.g., agricultural statistics and environmental accounting). On the other hand, it indicates that more work needs to be done on streamlining cadastral parcel data into statistical production to expand and improve the

location data framework for statistics by integrating this data with other primary types of location data than geospatially enabled statistical/administrative unit records (e.g., address and buildings).

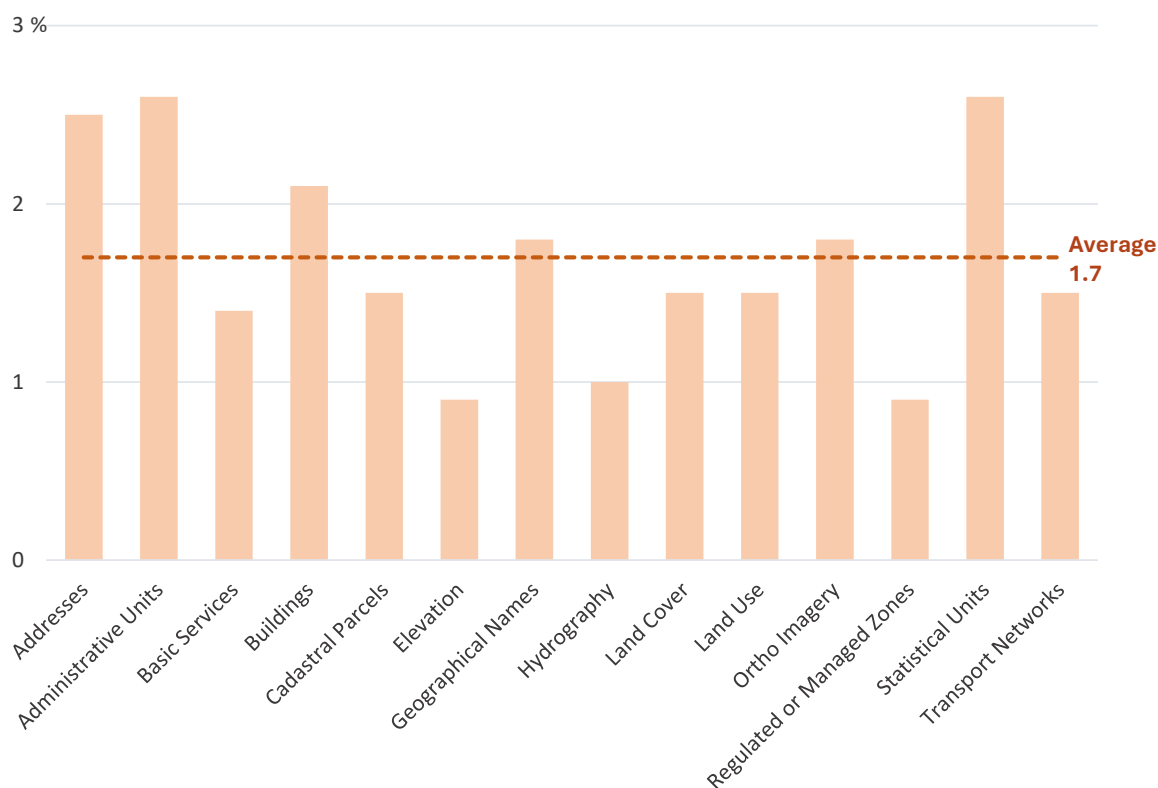


Figure 10. Average on the usage of authoritative and fundamental geospatial data by core data theme

Half of the data themes were below of the overall average on the usage of authoritative and fundamental geospatial data theme (1.7). It includes *Elevation*, *Regulated or Managed Zones*, *Hydrography*, *Basic Services*, *Cadastral Parcels*, *Land Cover*, *Land Use* and *Transport Networks*, listed in ascending order. Although with very different applications to official statistics, *Geographical Names* and *Ortho Imagery* shared the same average of 1.8, just below *Buildings* (2.1).

Question 2 – Geospatial enabling capacity

Question 2 aims to assess the extent to which NSIs are able to geospatially enable - or geocode - statistical and/or administrative micro data down to the level of x and y coordinates, i.e. georeferenced points (high-precision geocodes) that constitute the accurate location of statistical/administrative unit records.

While similar questions have been included in previous surveys, this is the first time respondents were asked to evaluate their geocoding capabilities across specific data domains in order to identify more specific data gaps and geocoding shortcomings. This question serves as a follow-up on the implementation of the Global Statistical Geospatial Framework (GSGF) and the objectives outlined in GSGF Europe, which advocate for geocoding of *all* micro data using coordinate-level precision (point-based location data linked to statistical/administrative data at the unit record level).

Respondents were asked to rate their ability to geocode statistical and/or administrative micro data to the level of x and y coordinates across a set of key data domains for statistics.

The following **scoring guide** was provided:

0 = Not aware/Our organisation does not handle this type of data/information.

1 = We cannot geospatially enable any micro data in this option to the level of x and y coordinates.

2 = We can geospatially enable some micro data in this option to the level of x and y coordinates.

3 = We can geospatially enable most/all micro data in this option to the level of x and y coordinates but with quality and technical issues.

4 = We can geospatially enable most/all micro data in this option to the level of x and y coordinates with good quality.

Table 4. Geospatial enabling capacity by data domain

Data Domains	0	1	2	3	4
Population data	0%	3%	5%	18%	74%
Business data	8%	10%	26%	28%	28%
Agricultural data	8%	13%	18%	31%	31%
Environmental data	49%	18%	15%	18%	0%
Energy data	33%	18%	21%	23%	5%
Building and construction data	26%	13%	26%	21%	15%

Unsurprisingly, population data emerged as the data domain with the highest level of capacity to geospatially enable data. Approximately 74% of respondents reported that their organisations are capable of geospatially enabling *most* or *all* micro data in this domain to

the level of x and y coordinates with good quality. An additional 18% indicated they can do so, although with some quality and technical issues.

This high level of maturity aligns with findings from previous surveys and is probably largely driven by legal obligations to provide census data variables on 1x1 km grid cells, which has significantly advanced geocoding capabilities to produce population grid statistics. In this regard, it is worth to note the methodological efforts and advancements made by several organisations to developed 1km² grid datasets from the population and housing census and/or administrative sources and produce grid statistics, as a complement to the traditionally used hierarchical system of administrative units.

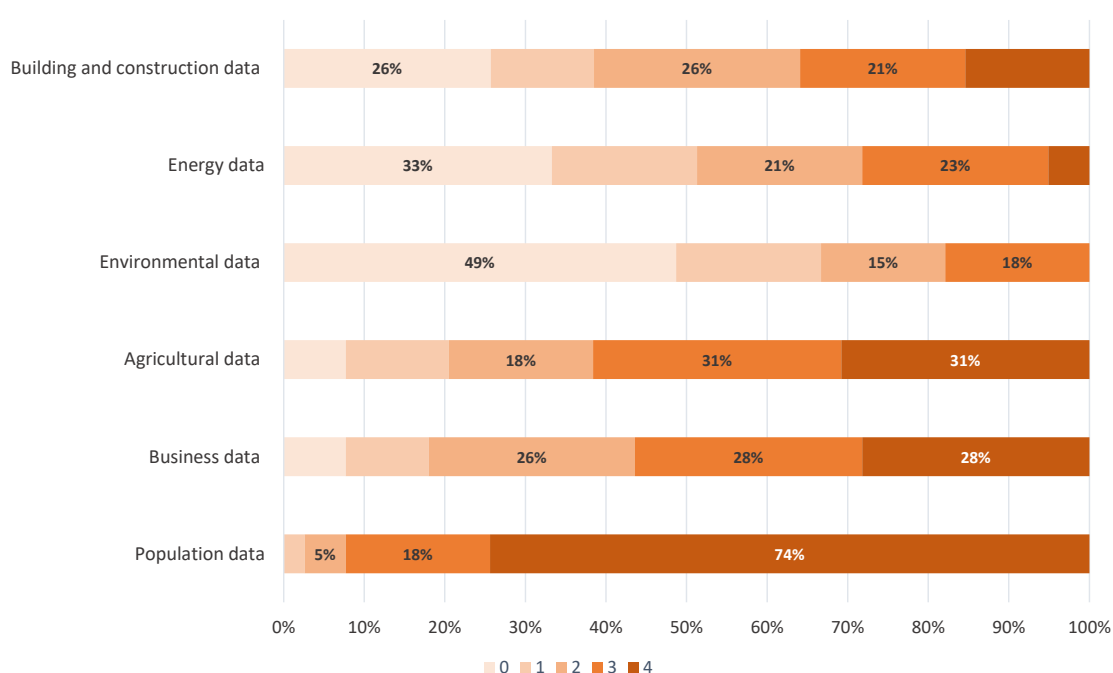


Figure 11. Geospatial enabling capacity by data domain

In contrast, data domains such as environmental data, energy data, and building and construction data received lower scores. This can be partly attributed to the fact that many NSIs do not produce statistics in these thematic areas. However, the low scores in domains like business data and agricultural data are more concerning.

On one hand, only 28% of responding countries reported the ability to geospatially enable most or all micro data in the business data domain to the level of x and y coordinates with good quality. On the other hand, only 5% of the respondents claimed to fully geospatially enabled energy data to the point level (x,y coordinates) with good quality, followed by the environment data that registered 0% in which almost half of the responding countries are

not aware or their organisation does not handle this type of data/information in their business production models.

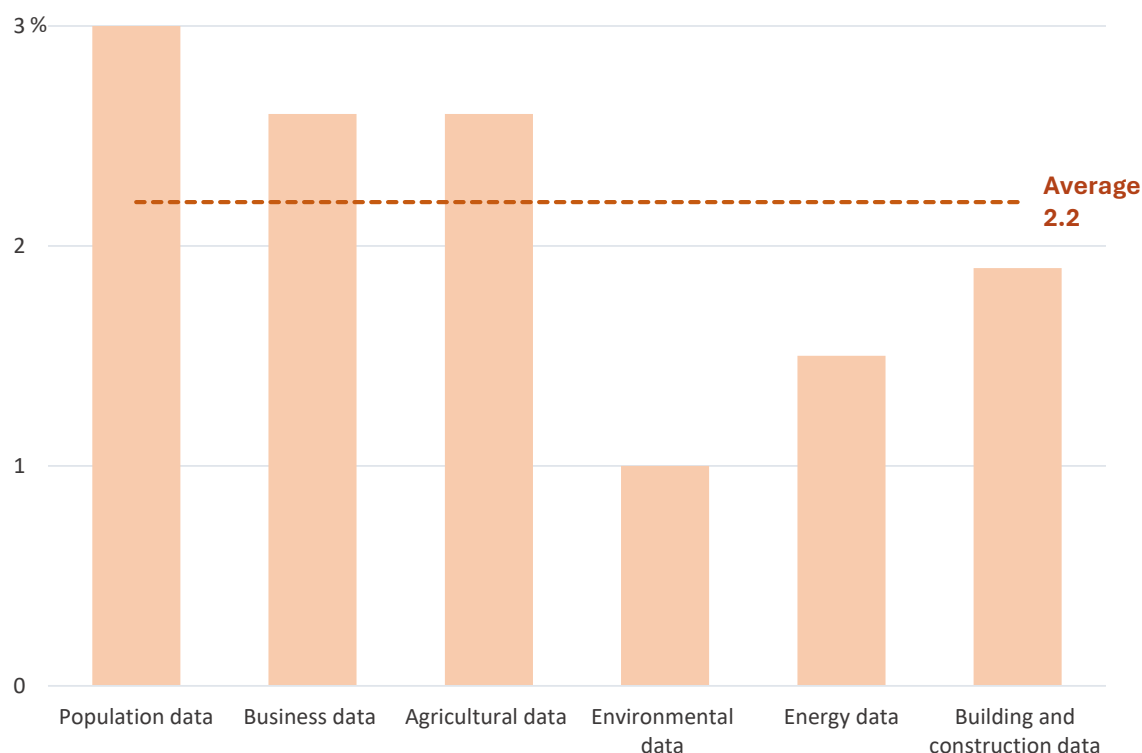


Figure 12. Average on the geospatial enabling capacity by data domain

Based on the above-described analysis, it is not surprising that the environmental data, energy data and building and construction data are the data domains with scores below the overall average (2.2), listed in ascending order (1.0, 1.5 and 1.9, respectively). Business data and agricultural have the same average of 2.6, although when analysing the results by score range it is noted that agricultural data showed higher percentages in the 3 and 4 options (31% against 28% registered for the business data).

Question 3 – Geospatial data management in statistical production integration

Question 3 is designed to assess the extent to which geospatial data management processes and capabilities are fully integrated into the statistical business processes of NSIs. The question addresses both technical and non-technical aspects supporting geospatial data management within the statistical production.

Over the past two decades, several important initiatives have been launched to strengthen the capacity of NSIs to use and manage geospatial information and improve its integration in the statistical production process, namely through innovative solutions to geocode various types of data and link/join statistical and geospatial data (e.g., GEOSTAT projects and GEOS grants). However, concerns remain that these capabilities are not yet systematically embedded within core statistical workflows. The following **response options** were provided:

- Yes, completely.
- Almost, but some elements still missing.
- Not fully integrated, but enough to conduct our business with a satisfactory performance.
- No, only partially integrated which is an obstacle for efficient data integration.
- No, not integrated at all, our geospatial business processes are mostly separated from the rest of the statistical business processes.
- No, we don't even have geospatial data management processes and capabilities.
- Not aware.



Figure 13. Geospatial data management in statistical production integration

In order to facilitate the analysis and provide clear insights, the results can be divided into **three main groups**:

- The **first group** (28%) consists of countries/organisations that consider geospatial data management to be well integrated into statistical production. While no respondent reported full integration (“*completely*”), these countries/organisations generally view geospatial data as more or less streamlined within their statistical

business processes and embedded in their enterprise architecture, production model and information infrastructure.

- The **second group** (33%) includes countries/organisations that do not consider geospatial data management processes and capabilities fully integrated into statistical production. On the other hand, they do not see this as a significant obstacle, as they are still able to carry out their work with satisfactory performance.
- The **third group** (39%) comprises countries/organisations that do not consider geospatial data management processes and capabilities fully integrated (or not at all) into statistical production and in addition they view the lack of integration as a barrier to efficient data integration.

Question 4 – Integration of geospatial information in statistical production

Question 4 builds on Question 3, continuing the examination of how geospatial information management is integrated into the business production model of official statistics. Specifically, it asks whether respondents' organisations have any reference documentation that explicitly addresses the integration of geospatial information and its role in statistical production (e.g. documentation that may have an impact on how organisations are mapping, describing or conducting its business production model, such as strategies, work plans, action plans and/or guidance and technical/methodological materials).

The underlying assumption is that the existence of reference and systematic documentation indicates a more formal recognition of geospatial information as a key component of statistical production in order to facilitate the alignment, streamlining and monitoring both statistical and geospatial data, business processes and services. Thus, it is important to provide a structure for documenting not only statistical processes, but also geospatial processes to ensure harmonisation between both data domains, including standards, good practices and terminology. Respondents were asked to choose one of the following **options**:

- Yes, and it has a true impact on what and how we do things.
- Yes, but it is mostly paper without much impact on what and how we do things.
- No, but we don't think it is needed.

- No, but we would need it.
- Not aware.

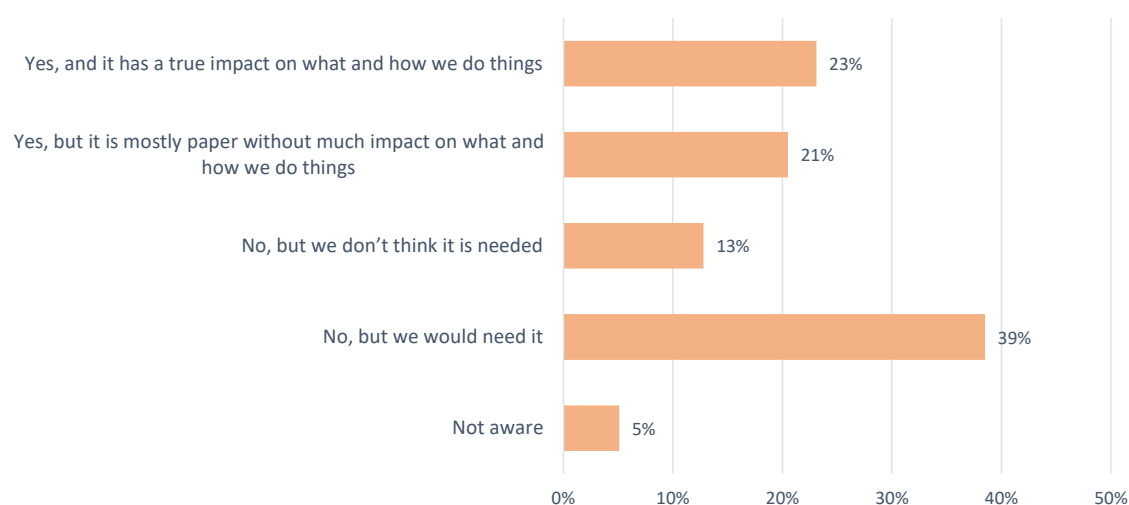


Figure 14. Integration of geospatial information in statistical production

Most respondents indicated that their organisations do not have (or are not aware of) any reference documentation explicitly addressing the integration of geospatial information and its role in statistical production. In total, 51% of the NSIs report that they lack such documentation. Among these, approximately one quarter do not believe that it would contribute to improve their production process.

Almost 45% of the NSIs reported having reference documentation that explicitly addresses the integration of geospatial information into statistical production. However, nearly half of these respondents considered the existing documents and materials to be largely theoretical paper products with little true impact on the statistical production process.

Question 5 – Geospatial data management and processing

Question 5 asks respondents whether they believe the needs and challenges associated with geospatial data management and processing are fully understood and recognised within their organisations, including at the strategic, leadership, and top management levels.

Geospatial data management often requires specialised tools, infrastructure, and technical solutions that differ from those typically used in statistical processes across a traditional statistical production. In this regard, this question aims to understand the level of geospatial awareness and acknowledgement across the organisation in all

organisational levels - from corporate to operational levels –essential for effective and successful statistical-geospatial integration.

Respondents were presented to the following **options**:

- Yes!
- Not fully understood, but to a satisfactory level.
- No, only partially understood.
- No, not understood at all.

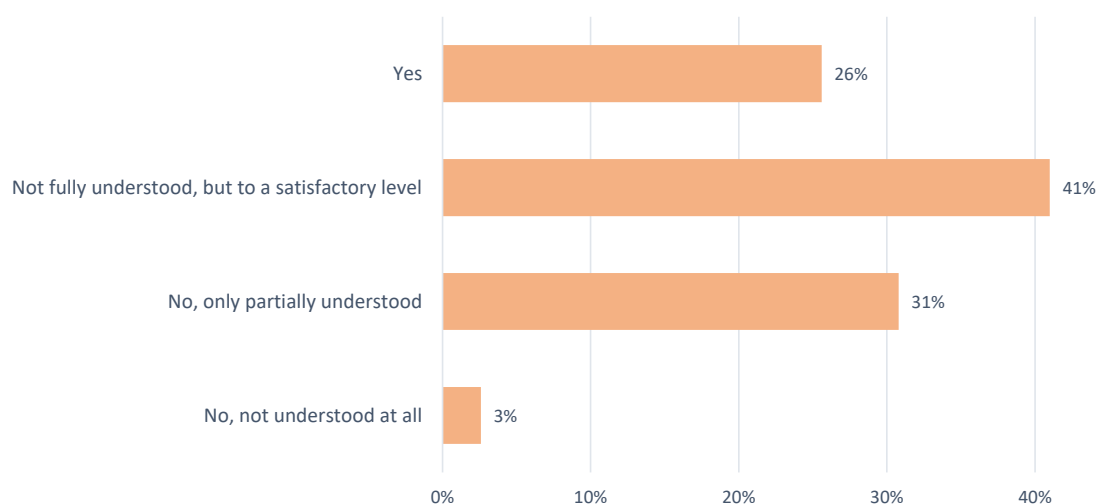


Figure 15. Geospatial data management and processing

Around a quarter of the countries/organisations (26%) reported that the needs and challenges related to geospatial data management and processing are *fully* understood and acknowledged within their organisations, including at the corporate level (e.g., leadership, top managers, etc.).

The largest part of respondents noted that the level of understanding is “satisfactory”, let alone not complete (41%), followed by the responses indicating that the needs and challenges are only partially understood (31%). In total, a considerable number of responses – around one third - expressed that needs and challenges related to geospatial data management and process are *not* yet fully (or not at all) recognised by their organisation.

The term “satisfactory” is somewhat vague and open to interpretation as it encompasses subjective criteria and empirical measures that differs from individual and collective perceptions and experiences. It may suggest that, although fully formalised recognition is lacking, the current level of awareness does not significantly hinder the organisation’s

ability to manage and process geospatial data as well as integrate geospatial process components into statistical production. However, the survey does not clarify whether “satisfactory” reflects an acceptable *status quo* or if respondents would prefer greater organisational engagement to enhance geospatial data management and processing capabilities.

In addition, the condition of “partially understood” may also differ from one NSI to another as being recognised at a high level (governance or strategic level involving areas related to leadership, capability development and corporate support) is different from being understood by the operational and technical levels, depicting distinct organisational situations concerning statistical-geospatial integration, both now and in the near future.

Question 6 – Obstacles for data integration

Question 6 explores what respondents perceive as the main obstacles to effective and successful data integration (statistical, administrative and geospatial data).

It is a slightly modified version of a question previously used in the GEOSTAT 4 survey (see question 2.5 in GEOSTAT 4a). While the primary aim is to identify current gaps and challenges, comparing the results with those from the earlier survey may offer insights into whether the nature of these obstacles has remained or evolved over time, or if new ones have emerged. It also helps to clarify if some of the past identified obstacles were overcome by the organisations during the last years.

The respondents were faced with the following fixed **response options** from which they were allowed to choose maximum two options:

- National authoritative and fundamental geospatial data does not exist, is incomplete, poorly maintained and/or inaccessible.
- Poor semantic and/or technical interoperability between different data sources or across data domains (e.g. lack of consistent identifiers to link data or inconsistent data models from the absence of common standards and guidelines, etc.).
- Lack of coordination and unclear mandates, roles and responsibilities (no data governance policies, custodianship guidelines or compliance mechanisms).
- Legal issues (lack of appropriate legal frameworks or outdated laws) preventing effective production, management, maintenance, exchange and sharing of data.

- Lack of enterprise-level architectures supporting capabilities for integrating geospatial data into statistical production.
- Lack of a technical infrastructure suitable to manage and process geospatial data in statistical production.
- Lack of capacity (know-how and skills) and/or human resources.
- A bit of all or some of the above-mentioned, but no major obstacles.
- No particular obstacle at all, things work quite smoothly!
- Other.

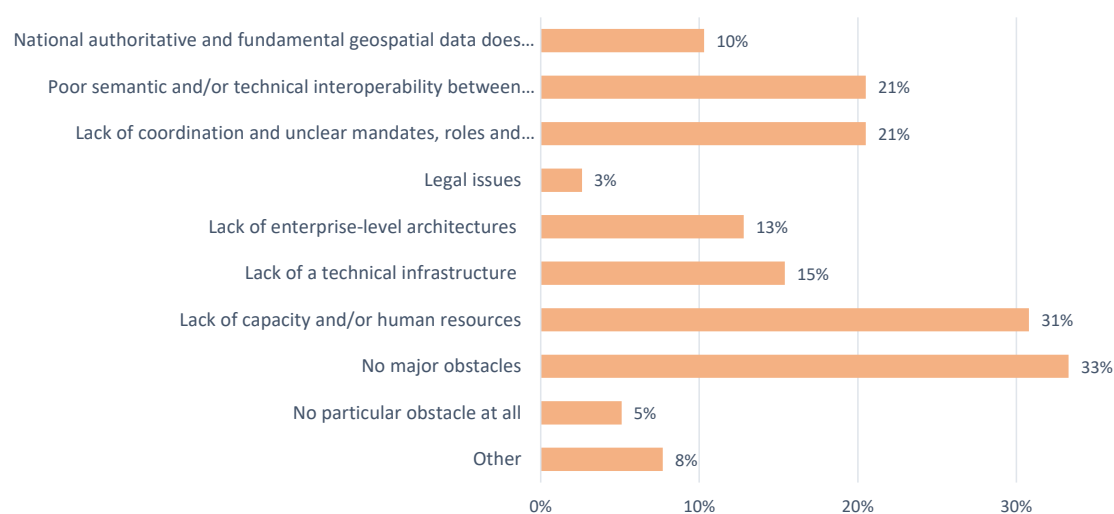


Figure 16. Obstacles for data integration

Two response options stood out in this question. The first is: “A bit of all or some of the above-mentioned, but no major obstacles” (33%). Unsurprisingly, this option scored highly, reflecting that many countries/organisations face challenges across various topics and action areas, though not to the extent that any obstacle can be clearly identified or pose a big challenge or threat.

The second high-scoring option is: “Lack of capacity (know-how and skills) and/or human resources.” However, this response does not clarify whether the primary issue is general understaffing or a lack of (geospatial) expertise with specialised skills among existing staff - likely, it is a combination of both gaps.

Although not directly comparable, it is interesting to contrast these findings with those from the GEOSTAT 4 survey, which produced very different results. In GEOSTAT 4, the top two challenges identified were the lack of nationally standardised data (e.g., address records, building registers) and poor semantic or technical interoperability between data sources.

Neither of these issues scored particularly high in the current survey. Nevertheless, poor semantic and/or technical interoperability between different data sources or across data domains presented a reasonable expression (21%), along with the lack of coordination and unclear mandates, roles and responsibilities that had the same score.

This suggests that the data quality landscape may have changed significantly over the past five years -a shift also reflected by responses to Question 1B in this survey. This remark is also underpinned by the relatively low score (10%) on the incompleteness, poor maintenance and/or lack of access to national authoritative and fundamental geospatial data.

Other reported obstacles ('Other' option) included lack of geospatially enabled data from administrative sources, lack of time and the absence of need/urgency to change or update ongoing workflows in order to integrate geospatial data and related capabilities.

Question 7 – Administrative data sources

Question 7 explores to what extent administrative data sources are used in the production of geospatial statistics. This question is more or less identical to a question used in the GEOSTAT 4 survey, allowing direct comparisons with previous surveys. It enables to assess if there is an increased use of administrative data and registers for data integration purposes (e.g., geocode administrative data and linking administrative registers with authoritative location data). It also aims check the administrative data-based context in supporting the production of geospatial statistics and overall business production model.

The respondents were faced with the following **response options**:

- Administrative data sources have already been implemented in the business production model of one or more of our statistical-geospatial outputs.
- Not yet implemented in our business production model, but we are currently looking into it or have plans to do it in the near future.
- Not implemented in our business production model. We have no plans to do it, and we do not foresee being able to use administrative data sources in the near future.
- Other.

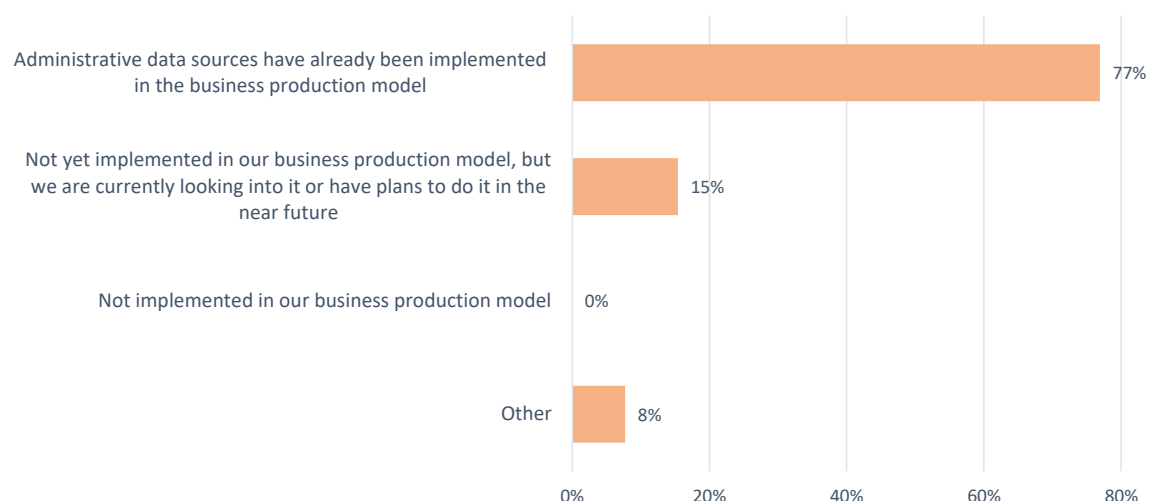


Figure 17. Administrative data sources

A significant majority of respondents (77%) reported that their organisations have already integrated administrative data sources into the production of one or more statistical-geospatial outputs. This marks an increase from the 2020 GEOSTAT 4 survey, where the corresponding figure was 68% (an increase of almost 10%). Although the frame of responding countries differs slightly between the two surveys, the results suggest a clear progress in the use of (geocoded) administrative data for generating statistical-geospatial outputs, thereby assuming a better capacity to integrate such secondary data sources with the geospatial data for statistical purposes. It also may suggest that the access to administrative data sources has improved in recent years, as legal issues are no longer a key obstacle based on the results of Question 6 (3%).

On the other, 15% of respondents indicated that they are currently exploring the integration of administrative data sources or have plans to implement it in the near future, showing some degree of action or commitment. In reality, this share is likely somewhat higher, as several responses categorised under “Other” included details about ongoing initiatives involving administrative data sources, such as for environmental statistics.

Question 8 – Building capacity for statistical-geospatial data integration

Question 8 explores the need for capacity building and training in the integration of statistical and geospatial data. In this context, capacity building should be understood as focusing on technical and practical aspects – developing, enhancing and strengthening the

skills (up-skilling), tools, and methodologies required for effective data integration, including professional training in core capabilities

The question encompassed the following **response options** from which maximum 2 options could be chosen:

- Basic training, i.e. general skills in using Geographic Information Systems (GIS) and geospatial data management.
- Advanced training, i.e. scripting and geospatial programming (e.g., SQL, Python, R, etc.) for increased automation, custom tools and modular services in statistical production.
- Applied methods for geospatial information in statistical production (e.g. data collection, geocoding, data aggregation, geospatial data and metadata management, managing disclosure risks arising from geospatial data, spatial analysis, etc.).
- Methods for data dissemination (Web mapping tools and geospatial services, semantic issues, machine-readable solutions, such as APIs, open technologies/standards for data exchange, cataloguing and tagging, etc.).
- Use of Earth Observation data and other alternative data sources to produce official statistics, highlighting experimental statistics.
- Other.
- No need for training or any other capacity building and capability development actions.

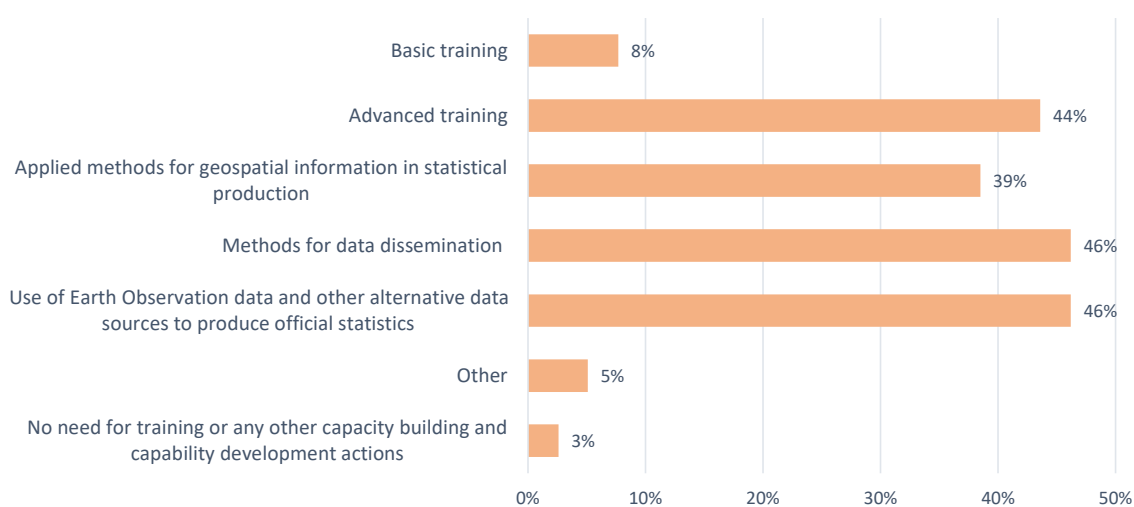


Figure 18. Building capacity for statistical-geospatial data integration

The most frequently selected options were “*Use of Earth Observation data and other alternative data sources to produce official statistics (...)*” (46%), “*Methods for data dissemination (...)*” (46%) and “*Advanced training, i.e. scripting and geospatial programming for increased automation (...)*” (44%). In contrast, the option “*Basic training, i.e. general skills in using Geographic Information Systems (GIS) and geospatial data management*” received only 8% of the responses, indicating that this type of foundational geospatial knowledge, expertise and skills are already retained by the staff of NSIs, namely the respective geospatial teams.

Instead, there appears to be a growing demand to advance and upgrade these skills further, particularly to support workflow automation through learning scripting and programming languages (e.g., Python, etc.) and to better leverage emerging data sources like EO data, following new and emerging trends in official statistics and geospatial technologies. In addition, shortage of such capabilities may also explain the fact the many respondents identified “*Lack of capacity (know-how and skills) and/or human resources*” as one of the main obstacles in advancing statistical-geospatial data integration (Question 6).

Question 9 – Areas for improvement

In **Question 9** respondents were asked to define action areas where they see a need for their country and/or organisation to make improvements and where they think they could benefit from an exchange of experiences and best practices from other countries and/or organisations.

The following **response options** were provided (multiple-choice allowed):

- How to organise a well-functioning collaboration and cooperation model on the national level with data governance guidelines between stakeholders (NSIs, NMCA and other data providers).
- How to build well-structured and interoperable national data ecosystems to facilitate effective statistical-geospatial data integration.
- How to build internal data management environments and location-centred architectures to facilitate effective statistical-geospatial data integration.
- How to streamline (organisationally and technically) geospatial data into statistical production.

- How to build good solutions for data dissemination, including creation of interactive web maps/atlasses, setting up machine-readable services (APIs) or any other custom geospatial service for statistical-geospatial outputs.
- How to implement the GSGF (Europe) on the national and/or organisational levels in all its components (5 Principles and 4 Key Elements).
- No need for development actions and improvements.

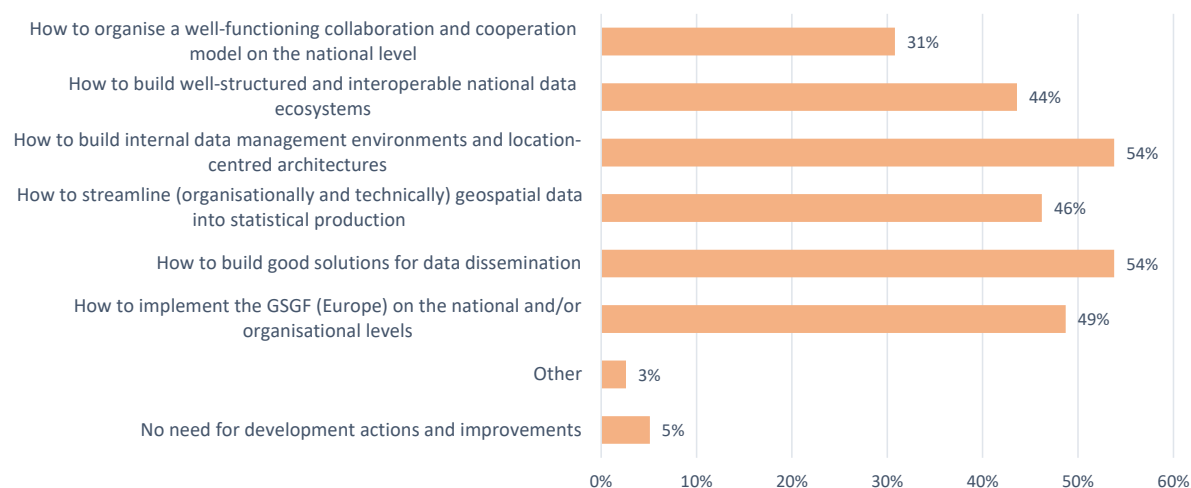


Figure 19. Areas for improvement

As the question allows for multiple choices (no limitation in the number of options), the percentages does not reflect the number of respondents but rather responses.

The two top equally scored categories were “*How to build internal data management environments and location-centred architectures (...)*” and “*How to build good solutions for data dissemination*”, recording 54% of the responses. In terms of data dissemination methods and innovative solutions to publish statistical-geospatial products, there is a clear alignment with the previous question on building capacity, highlighting that this action area should be a priority for improvement.

Though there is not possible to directly compare the results of this question with previous surveys because the design of the questions have changed, the current survey reiterates the needs expressed by respondents already in the GEOSTAT 4 survey, to modernise internal data management environments to support a higher degree of automation (e.g., see question 2.4 in the GEOSTAT 4 survey where the vast majority of responses indicated a need for improvements). Establishing data management environments building on

location-centred architectures can be considered one of the key provisions of Principle 2 of the Global Statistical Geospatial Framework (GSGF).

Nevertheless, the other options that had responses around 40-50% cannot be overlooked, namely to build well-structured and interoperable national data ecosystems (44%) – related to GSGF Principle 1 -, to streamline geospatial data in statistical production (46%) and most importantly, to implement the GSGF (Europe) in the organisation and/or country.

Question 10 – Advancements and good practice

Question 10 is linked to Question 9, but here respondents were asked for areas where they think their country/organisation has made advancements that could serve as guideline and good practice for other countries and/or organisations. The idea behind these two questions is to try to match needs expressed in Question 9 with countries and/or organisations that are willing to share their progress and achievements according to Question 10.

Similar to the previous question, Question 10 also allowed respondents to choose multiple options. The same **response options** as in Question 9 were used.

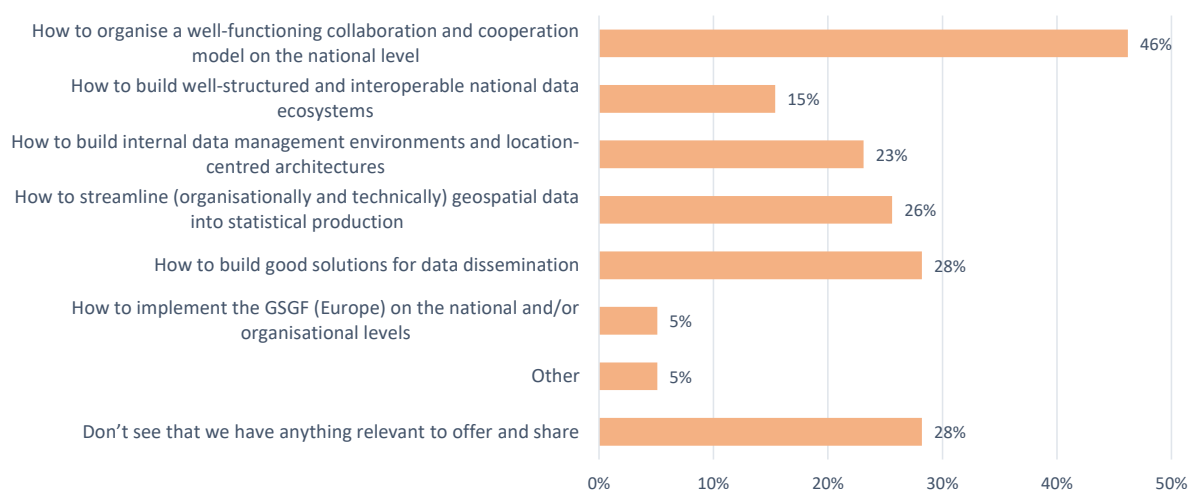


Figure 20. Advancements and good practice

The most significant advancement that countries and/or organisations have made -and are willing to share -is the development of “*well-functioning collaboration and cooperation model on the national level with data governance guidelines between stakeholders (NSIs, NMCAAs and other data providers).*”, recording almost half of the responses (46%). This strongly suggests notable progress in national-level collaboration and inter-institutional

cooperation, which is particularly interesting given that previous surveys and discussions within the GISCO Working Group that have consistently identified a lack of institutional collaboration as a major challenge in data integration.

Not surprisingly, good solutions for data dissemination had a reasonable ratio (28%) following the results from the previous Questions 8 and 9 on building capacity and areas for improvement, respectively. The same ratio was noted in the set of countries/organisations considering that they do not have anything relevant to offer and share. The same analysis is applied to the option *“how to streamline (organisationally and technically) geospatial data into statistical production”*, wherein 26% of the responses indicated advancements and capability development this action area and have related guidelines and good practice cases.

At the same time, the responses reveal a mismatch between demand and supply in certain areas. Some topics identified as high priority for improvement in Question 9 are not matched by a corresponding provider of guidance and good practices. A key example is the implementation of the GSGF at national and/or organisational levels as many respondents expressed a clear need for guidance in this area, only a limited number of countries/organisations reported having relevant experiences or best practices to share. Lastly, data quality was mentioned as another action area (outside the predefined options) encompassing guidelines and good practices to be shared among countries and/or organisations.

Question 11 – Production infrastructure for statistical-geospatial data integration

Question 11, the final question in the survey, is qualitative in nature although with a quantitative response scale. It invites respondents to provide a general self-assessment of their organisation’s production infrastructure for statistical-geospatial data integration, without predefined specifications or direct response options.

Respondents were firstly asked to imagine an ideal scenario for their organisation, considering aspects such as technical solutions, legal frameworks, financial and human resources, data access and quality, and collaboration with other stakeholders. Based on this thinking exercise, they were asked to rate how close they are to this ideal situation on a scale from 0 to 5, where:

- 0 indicates “we have a very long way to go”, and
- 5 indicates “we are more or less in an ideal situation already.”

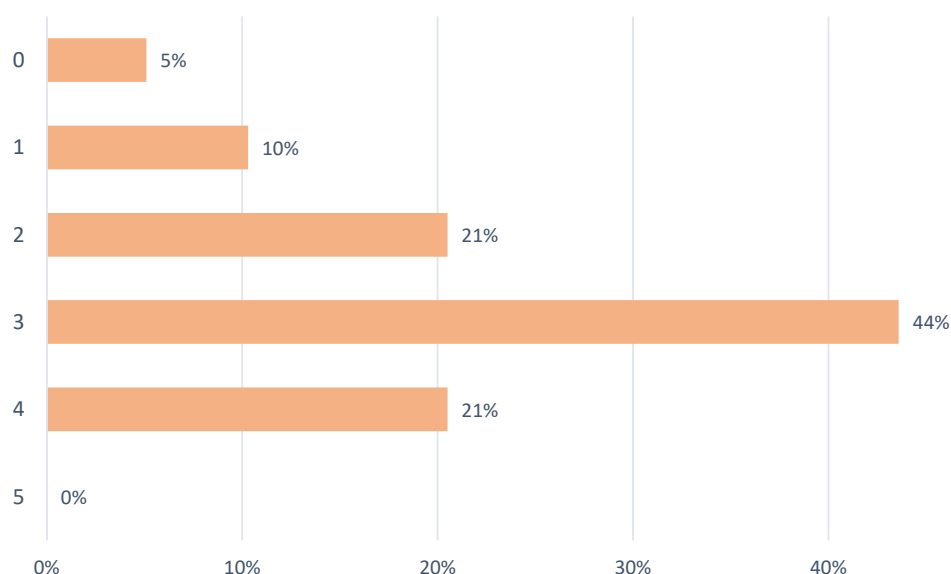


Figure 21. Production infrastructure for statistical-geospatial data integration

The majority of countries/organisations were clustered around the midpoint of the scale with an average rating of 3 (44% of the respondents). This suggests that while most do not consider their current situation ideal, they also do not see themselves in a particular bad spot or at a very early stage. Notably, no respondents rated their situation as a 5 - “more or less an ideal situation already.” However, 21% rated their position as 4, indicating that they are quite close to achieving their envisioned ideal. The same ratio was recorded in 2 demonstrating that around 2/5 of the respondents are either behind the intermediate situation or right after it and getting closer to the ideal situation.

On the other side of the spectrum, 5% of the respondents have rated their situation as a 0 - “we have a very long way to go”, and 10% as a 1, underlining that around 15% of the respondents are very unsatisfied with the situation on statistical-geospatial data integration and have a long walk ahead.

Conclusions

The results from the survey - as well as from previous ones in other projects - indicate steady progress in the field of statistical-geospatial integration over the past decade, albeit slow in some areas. The findings suggest that statistical-geospatial data integration within European NSIs (inside and outside the European Statistical System) is achieving a maturing domain, wherein the operating environment has significantly evolved from its state at the beginning of the century.

However, as maturity increases, so do expectations for further advancements, constant modernisation as well as the need and ambition to adapt to a rapidly changing digital transformation, technological landscape and societal trends. A decade ago, geospatial data management was largely the domain of geographers using traditional desktop GIS tools to spatialise data, perform basic spatial analysis and prepare statistical maps. Today, emerging data sources (e.g., EO data), cloud computing, open-source technologies, and new programming languages pose both exciting opportunities and new challenges for NSIs, in particular to integrate statistical and geospatial data and produce geospatial statistics output.

Data availability and data quality

Most significantly, the availability of data appears to have gradually improved. In 2015, according to the GEOSTAT 2 survey, lack of data or restricted access to data was a recurring issue across multiple questions. Where there was once more or less a complete data gaps, we now see emerging initiatives aimed at building national fundamental geospatial datasets to support policy agendas and strategic priorities. It can be assumed that European legislation (EU law), such as the INSPIRE Directive and the High-Value Dataset (HVD) Implementing Regulation, has played a key role in driving this development in the overall data ecosystem, namely in terms of open and public access and sharing.

High-quality geospatial reference data with broad coverage and good quality, such as addresses, buildings, and cadastral parcels, are absolutely essential for NSIs to advance their data integration efforts and produce innovative statistical outputs in the scope of the modernisation of official statistics. In light of this, it is somewhat concerning that, although address data and building data are now emerging in many countries, partial coverage and quality issues still limits its effective use for tasks such as geocoding of micro data. The

absence of a regulatory/legal framework, data governance policies, custodianships mandates with roles and responsibilities may be possible reasons. Also, the lack of common data (and metadata) standards, agreed quality requirements and a harmonised coding system may be hindering the quality of authoritative geospatial data, especially its integrity.

It is essential that the ESS clearly articulates the need for complete, high-quality geospatial reference data across all countries and align the statistical needs with the geospatial capabilities. While all fundamental geospatial datasets are important, this survey clearly highlights where priorities should lie to maximise impact on the statistical system as a whole.

Streamlining of geospatial information in statistical production

One of the key purposes of the survey was to explore to what extent geospatial data management and processing are integrated into the statistical production processes of NSIs. This is a complex topic to assess, as there are few objective criteria to rely on and each NSI has its own organisational context. What does “streamline” truly mean? To what degree should geospatial data be embedded in processes and what level of repeatability or automation is desirable? Perceptions and expectations vary as organisations operate under different conditions and to execute the same process to get the same output, they can adopt distinct approaches and solutions. What one NSI considers well-integrated may be perceived differently by another depending on both internal and external factors.

Nevertheless, there appears to be a common perception that geospatial data management and processing - to a significant degree - are still separate from core statistical business processes and respective workflows. Is this separation between statistical and geospatial production components problematic? Opinions differ on which some argue it is not, while others believe bridging this gap is essential to fully harness the potential of geospatial data in statistical production.

There is a compelling case for the ESS to address this issue at a strategic level. The ESS, along with its members and partners, stands to benefit significantly from advancing the integration of geospatial and statistical information at the enterprise level, i.e. promote a greater alignment of the business process, information, applications, practices, technologies and infrastructures from both domains. Achieving this requires the

development of clear recommendations and guidelines to help national statistical systems establish a common baseline in order to leverage the disparities between countries.

A crucial first step is the formal recognition of geospatial information and its role within statistical systems, raising a high-level commitment and engagement to increasingly embody the geospatial dimension into the statistical strategies and programmes. Nevertheless, while declared strategies and policy documents alone may not drive effective change, increasing the visibility of geospatial data and showcasing its added value for statistics, can help elevate its importance within organisational priorities.

Obstacles and capacity needs

What stands out prominently in this survey, compared to previous ones, is the widespread lack of capacity in terms of know-how, skills, and/or human resources reported by many countries/organisations. The survey does not clearly indicate whether this challenge stems from understaffing in the units responsible for geospatial data management, or whether existing staff require additional training to keep pace with rapid technological advancements. It may also be due budgets shortfalls, organisational aspects and/or problems on the capacity of attract or retain staff with the required professional qualification and expertise (e.g., high staff turnover).

Nevertheless, the findings point to a strong demand for advanced technical training, such as scripting and geospatial programming (e.g., SQL, Python, R), to support increased automation, the development of custom tools, and modular services in statistical production. There is also a particular need for training in the use of EO data and other alternative data sources. This type of demand goes beyond the traditional geography/geospatial science education and poses some challenges related to professional training and capacity building to acquire new core capabilities and skills tackling emerging data and production needs.

While many NSIs already employ highly technical capable personnel, such as data scientists and data engineers, these experts may not be directly involved in geospatial data management, highlighting the clear geospatial capacity gap. Because of organisational reasons or different academic backgrounds, often it appears to be limited cross-fertilisation between traditional methodological expertise, data science and geospatial science within NSIs. Enabling a better fusion of competencies and expertise and building work synergies between different areas of knowledge and professional experience are

critical to streamline geospatial information in statistical production. This is very important for several stages and processes of statistical production, such as to build a sample frame by using geocoded data, to design data integration methods, exploring and assessing new data sources and to address confidentiality issues by developing disclosure control methods that handle geospatial data (e.g., geographic differencing), just to mention a few examples.

While the lack of capacity and human resources can be addressed and mitigated within the statistical domain, other obstacles require cross-domain solutions and long-term policy-level action. Respondents highlighted several challenges, including a lack of coordination and unclear mandates, roles, and responsibilities, such as the absence of data governance policies, custodianship guidelines, or compliance mechanisms. Additionally, poor semantic and/or technical interoperability between different data sources or across data domains was noted, emphasising the long way to go towards the adoption of common standards, and harmonised and interoperable data.

These obstacles are closely linked to the broader trends on digitalisation of the public sector and the development of data spaces for public sector data, both geospatial and non-geospatial. Although the statistical community is just one of many stakeholders, it plays a crucial role – as a beacon of inter-institutional collaboration and coordination - in advocating for robust data governance policies and interoperability efforts to ensure that data from various domains can be effectively integrated, shared and used.

References

GEOSTAT 1a (2011). GEOSTAT1A Final report, Annex I: User Needs Survey and Annex II: Data Providers' Survey.

<https://www.efgs.info/wp-content/uploads/geostat/1a/GEOSTAT1A-final-report.pdf>.

GEOSTAT 2 (2016). Spatialising Statistics in the ESS. Results from the 2015 GEOSTAT 2 survey on geocoding practices in European NSIs. EFGS and Eurostat.

<https://www.efgs.info/wp-content/uploads/2017/03/GEOSTAT2-Spatialising-statistics-in-the-ESS-2015.pdf>.

GEOSTAT 4 (2020). Findings from the 2020 GISCO / GEOSTAT 4 survey. EFGS and Eurostat. [https://www.efgs.info/wp-](https://www.efgs.info/wp-content/uploads/2025/07/GEOSTAT4_Findings_survey_2020.pdf)

[content/uploads/2025/07/GEOSTAT4_Findings_survey_2020.pdf](https://www.efgs.info/wp-content/uploads/2025/07/GEOSTAT4_Findings_survey_2020.pdf).

GEOSTAT 4 (2021). GSGF Europe. Eurostat ESSnet grant project GEOSTAT 4, 2019-2022. https://www.efgs.info/wp-content/uploads/2022/03/GSGF_Europe.pdf.

Rademacher, W. (2012). Population Grid Statistics from Hybrid Sources. In: Reality, Data and Space International Journal of Statistics and Geography, 3(3). <https://rde.inegi.org.mx/index.php/2012/09/12/population-grid-statistics-from-hybrid-sources/>.

UN-GGIM: Europe (2016). Core Data Scope. Working Group A - First Deliverable of Task 1.a. Version 1.2 - 2016-07-27. https://un-ggim-europe.org/wp-content/uploads/2018/11/UN-GGIM-Europe-WGA-Core_Data_Scope-v1.2.pdf.

UN-GGIM: Europe (2018). Core Spatial Data Theme "Basic Services" Recommendation for Content. Working Group A - Deliverable of Task 1.b. Version 1.0 - 2018-01-17. https://un-ggim-europe.org/wp-content/uploads/2023/03/UN-GGIM-Europe_WGA_Recommandation_Content_Theme_Basic_Services_v1.1.pdf.

UNSC and UN-GGIM (2019). The Global Statistical Geospatial Framework. GSGF. https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf.



2024-FI-GEOS-GSGF-CARE

© GSGF-CARE 2025