Background document for the Note by the Secretary-General transmitting the report of the Global Working Group on Big Data for Official Statistics (E/CN.3/2020/24)

Statistical Commission
Fifty-first session

3–6 March 2020
Item 3 (t) of the provisional agenda*
Items for discussion and decision: Big data

Background document
Available in English only

Global assessment of institutional readiness for the use of big data in official statistics

Executive summary

This report outlines findings from Project 1: An assessment of NSO readiness for the use of big data in official statistics from the UN GWG Task Team on Training, Competencies and Capacity Development. The results show that, of those NSOs that responded to the survey, a large proportion are already embracing big data / data science:

- **Key point 1 - Strategic coordination:** Strategic coordination capacities are fairly established. Many NSOs are actively engaged in big data projects. Ethics and quality frameworks are fairly established. Most NSOs view coordination with Big Data source owners inside their NSS as the lowest challenge.
- **Key point 2 – Legal framework:** Overall, respondents are aware of the fundamental role legal frameworks play in establishing big data projects. Many NSOs appear to have well developed legal frameworks that penalize data disclosures and allow accredited to access their data.
- **Key point 3 – IT infrastructure:** The analysis shows a heterogeneous picture over the IT infrastructure. NSOs stated that basic IT infrastructure such as power supply, and air-conditioning mostly met their needs, but they outlined struggles with storage facilities and computing power.
- **Key point 4 – Human Resources:** NSOs recruit significantly more analysts than data scientists and prioritize up-skilling over hiring external staff to perform big data/data science techniques.

Overall, the findings present a positive picture in terms of ensuring that the required foundations are in place and illustrates the ambition across NSOs to incorporate big data / data science into its core business. There are areas in which NSOs may require further information, guidance, development and knowledge to ensure that barriers to working with big data are removed:

- **Key point 1 - Strategic coordination:** Only a third of all NSOs have overarching big data strategies in place and Chief Data Officers only exist in some NSOs. The biggest challenge for NSOs is the collaboration with Big Data source owners outside the government.
- **Key point 2 – Legal framework:** Legal frameworks are still insufficient to regulate big data applications. Only a small share of NSOs rely on legal frameworks that guarantee access to big data.
- **Key point 3 – IT infrastructure:** IT infrastructure appears as central barrier to develop big data capacity; onsite and offsite storage capacity needs improvement for many. Only a few NSOs consider cloud storage a relevant option.
- **Key point 4 – Human Resources:** Most NSOs lack a competency framework to develop new skills to cope with big data (mobile phone, geospatial data) and new methodologies (machine learning).

Recommendations

Based on the analysis conducted by the Task Team, the following recommendations can guide the work of international organizations, development cooperation partners and national statistical offices to adapt to big data requirements:

Strategic Coordination:

- Promote the sharing of training resources on the United Nations Global Platform

- Promote the exchange of Big Data projects from all regions through the United Nations Global Platform
- Advocate big data strategies as one pillar of National Statistical Development Strategies (NSDS)
- Facilitate partnerships and exchange with data owners outside the NSS

Legal frameworks:

- Develop legal frameworks that include data sharing agreements between NSOs and private sector data owners
- Advocate for the importance of data privacy and data protection laws

IT Infrastructure:

- Advocate for cloud storage facilities in countries with necessary pre-requisites

Human Resources:

- Develop an overarching competency framework for big data skills development and HR strategies
- Investigate the potential for defining a data scientist pathway
- Foster partnerships with higher education institutes to design skill profiles for future employees
- Identify training pathways that allow up-skilling of staff available to share their knowledge in their teams (in collaboration with academic institutes)

Background

Big Data is, by definition, different from the traditional data sources used by National Statistical Organisations (NSOs). The new data sources pose new challenges across a range of expert areas, including methodology, quality assurance, technology, security, privacy, legal matters and skills. The breadth of challenge adds to the complexity of incorporating big data into regular research or organisational operations and ensures that the

transition to their use is difficult, or hindered, for many NSOs. The UN GWG Task Team on Training, Competencies and Capacity Development, is tasked with delivering projects in five specific areas:

A. Assessment of institutional readiness for big data in official statistics;

B. Development of a Competency Framework for new data sources in official statistics;

C. Analysis of the supply and profiles of specialists in areas related to the analysis of new data sources and big data;

D. Development of a curriculum and associated training courses;

E. Capacity building and sharing experiences through innovation centres via a global network.

This report presents results from the first project, (A) an assessment of institutional readiness. The project aims are to explore and understand the readiness of NSOs for the use of big data in official statistics, as well as to gather useful insights that might feed project strands (B) – (E). For the purpose of this project, an institution's "readiness" is defined by its maturity within four strategic areas:

1. Strategic Data Science Coordination: The presence of, or future plans for, strategic data science coordination within the NSO and across the National Statistical Service (NSS). This will have also considered the budgetary requirements for financing big data analytics within the organisation.
2. Legal Framework: The presence of, or future plans for, a legal framework for data access and data sharing within the NSO, the NSS, and potentially wider.
3. IT Infrastructure: The extent of, or future plans for, the IT infrastructure to enable big data analytics in a secure environment.
4. Human Resources: The number of data science posts within the NSO/NSS, the skills gaps and the future plans for recruitment and growth. This includes the skills needed to develop and maintain appropriate methodologies.

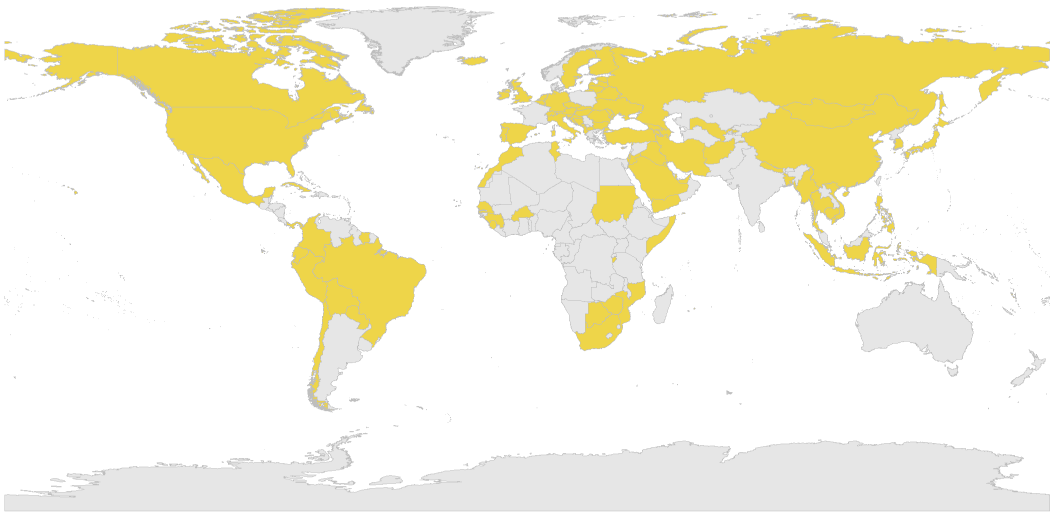
Data collection was undertaken via a questionnaire. This was designed to collect data from across these four areas, to enable an assessment of institutional readiness. The questionnaire was issued to 160 NSOs during the period from 4th October 2019 to 15th November 2019. Responses

were received from 109 statistical organisations. After data cleaning (removal of non-complete responses and larger non-national organisations) 100 National Statistical Organisations (NSOs) were then used for our analysis. The overall response rate was 63%.

In order to support the work of other UN Task Teams, the results of the analysis will be fed across the UN network, to ensure that important findings from the data are shared for constructive use by others.

Survey respondents:

Africa	Americas	Asia.and.Pacific	Europe
Botswana	Antigua and Barbuda	Afghanistan	Albania
Burkina Faso	Bolivia (Plurinational State	Armenia	Austria
Burundi	Brazil	Azerbaijan	Belarus
Cabo Verde	Canada	Bahrain	Belgium
Guinea	Chile	Bangladesh	Bosnia and Herzegovina
Mauritius	Colombia	Brunei Darussalam	Bulgaria
Morocco	Cuba	Cambodia	Croatia
Mozambique	Ecuador	Macao	Czechia
Senegal	Mexico	Hong Kong	Estonia
Sierra Leone	Curaçao	China	Finland
Somalia	Panama	Cyprus	Germany
South Africa	Paraguay	Georgia	Hungary
Sudan	Peru	Indonesia	Iceland
Tunisia	Saint Kitts and Nevis	Iran (Islamic Republic of)	Ireland
Zimbabwe	Suriname	Iraq	Italy
	Montserrat	Israel	Latvia
	United States of America	Japan	Lithuania
		Jordan	Luxembourg
		Kuwait	Montenegro
		Maldives	Netherlands
		Mongolia	North Macedonia
		Myanmar	Portugal
		Nepal	Republic of Moldova
		State of Palestine	Romania
		Philippines	Russian Federation
		Qatar	Slovakia
		Republic of Korea	Slovenia
		Saudi Arabia	Spain
		Singapore	Sweden
		Thailand	Switzerland
		Turkey	Ukraine
		United Arab Emirates	United Kingdom of Great
		Uzbekistan	
		Vanuatu	
		Viet Nam	
		Yemen	



Main report

Strategic Data Science Coordination

The Strategic Data Science Coordination section of the questionnaire aimed to assess the establishment of (or plans for) strategic data science coordination within NSOs and wider (such as their NSS).

1. Big data/data science projects established

Many NSOs provided qualitative information about the type of big data / data science projects that have been established at their organisation. Some of the more common projects involve alternative data sources, for example, web scraped data, mobile phone data and scanner data. More information on projects can be found in the following inventory: https://unstats.un.org/bigdata/inventory/.
Almost half of the respondents undertake big data or data science projects. 47% of NSOs currently undertake big data projects, 32% do not undertake any of those projects, but are trying to establish. Around 21% do not plan to undertake those projects at all.

2. Big data/data science strategy in place

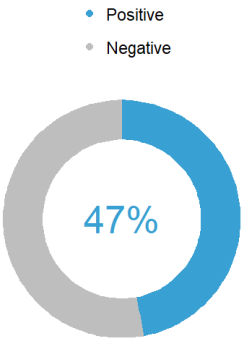
28 NSOs indicate to have a big data/data science strategy in place, with 35% of respondents to the survey indicating that they have implemented such a strategy.
A third of respondents have a strategy for big data in place, almost two thirds (60%) of respondents do report to try to establish a big data strategy in their NSO. Only 5% of respondents do not have any strategy established.

3. Chief Data Officer/Data Science Lead available

20 NSOs indicate to have a designated Chief Data Officer/Data Science Lead in place, with 25% of the respondents to the survey confirming this post.
A quarter of the respondents to the survey have a designated Chief Data

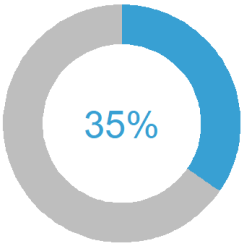
Officer. About 42% trying to establish this post, while about 30% do not plan do so.

NSO has Big Data / data science projects:



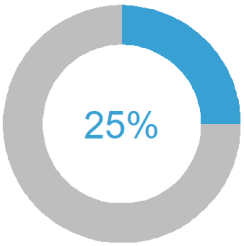
Answer	Count	Percent
Yes	47	47%
No, but trying to establish	32	32%
No, none planned	21	21%

NSO has a strategy for using Big Data in Official Statistics:



Answer	Count	Percent
Yes	28	35%
No, but trying to establish	49	60%
No, none planned	4	5%

NSO has a Data Science Lead / Chief Data Officer:



Answer	Count	Percent
Yes	20	25%
Don't know	2	2%
No, but trying to establish	34	42%
No, none planned	24	30%

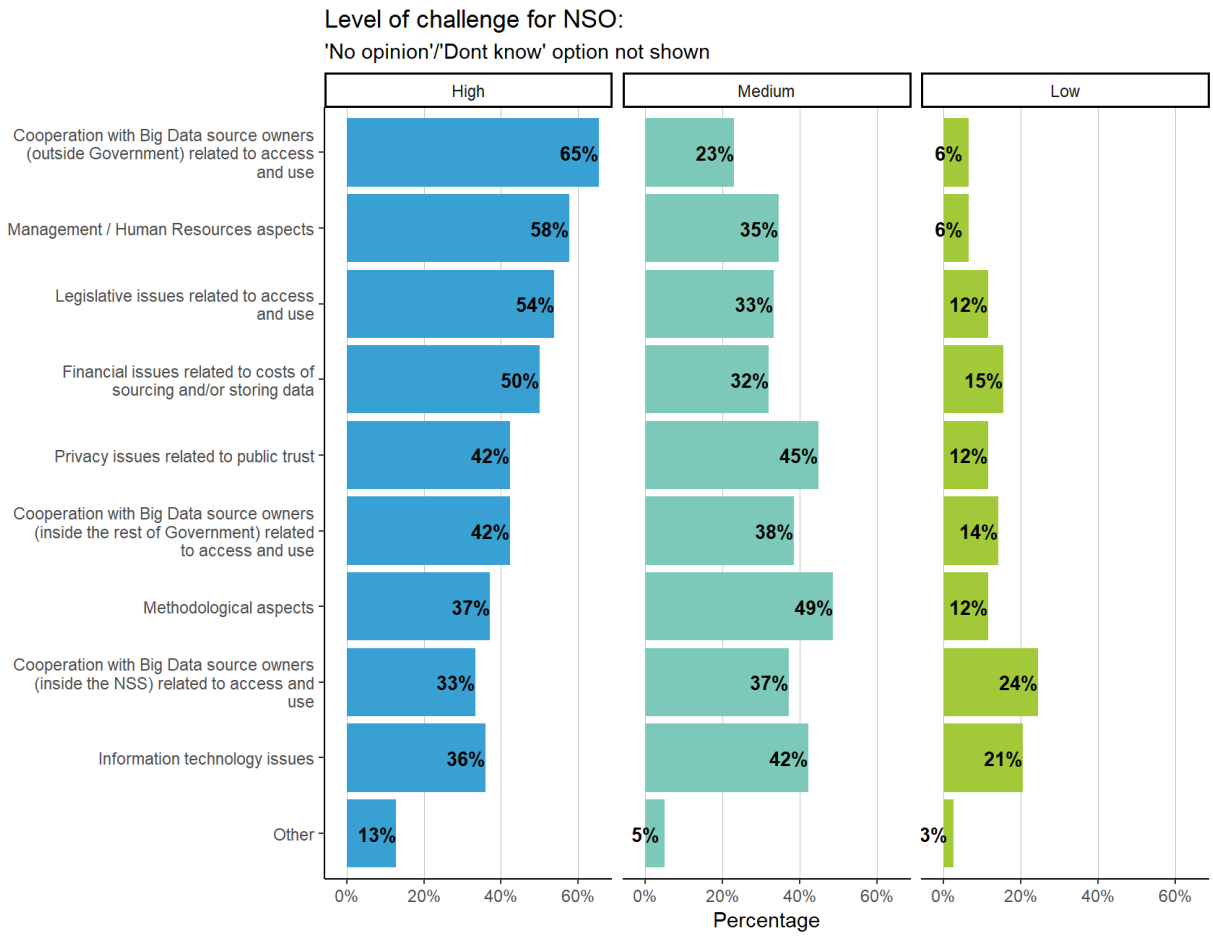
4. Coordination challenges for NSOs

The central challenge for NSOs is collaborating with Big Data source owners outside the government (65% of respondents), followed by human resources (58% of respondents) and legislative issues (54% of respondents).

Privacy issues related to public trust and methodological aspects cause medium level challenges, with 44% of the respondents indicating that privacy issues are difficult and 48% of respondents pointing at methodological aspects.

Cooperation with big data source owners inside the NSS and information technology issues are seen as low-level challenges, with 24% of NSOs indicating that cooperation with data owners inside the NSO are low level

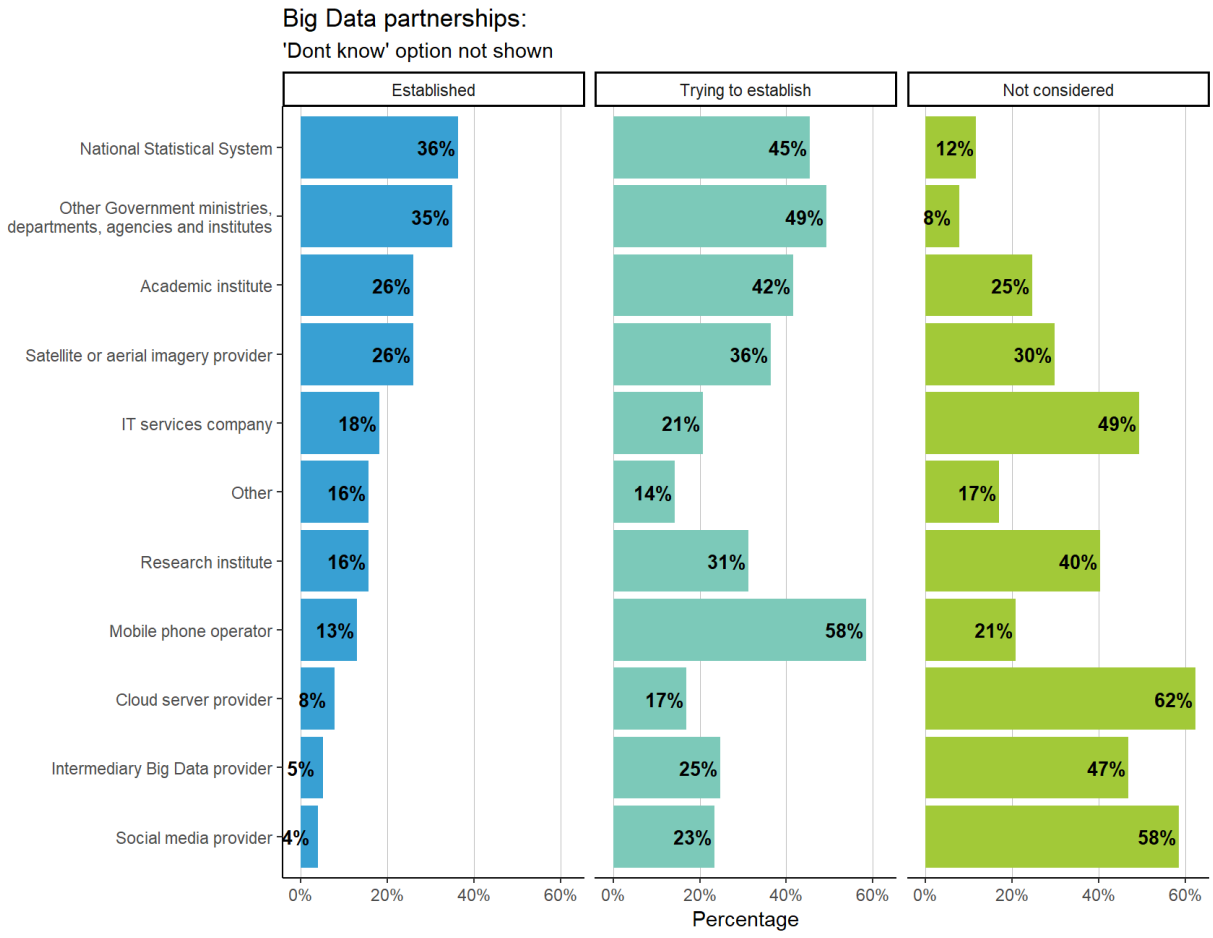
challenges, and 21% of NSOs indicating that IT issues are low level challenges.



5. Big data partnerships established

Partnerships inside the NSS and with the government still dominate the field; however, there is high interest in partnerships with new data sources and providers.

Around 36% of NSOs have established partnerships with the NSS/government ministries. A quarter of NSOs engage with academic institutes and satellite or aerial image provider (ca. 26% respectively). Least importantly appear social media providers – less than 5% have established partnerships with them, and about 58% of respondents do not consider doing so. Partnerships with cloud server providers seem similarly unpopular. Importantly, about 58% of NSOs are trying to establish a partnership with mobile phone operators.



6. Negotiation capacity of NSOs

The majority of NSOs are able to negotiate data provision with their partners.
40 NSOs are able to negotiate data provision with their partners, with 51% indicating to be able to do so. 35% of NSOs trying to establish negotiation capacity and only 8% do not plan to do so.

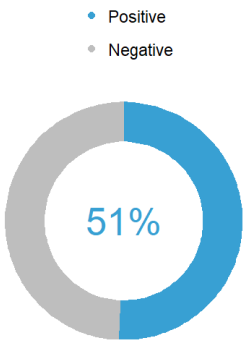
7. Data ethics policy

Almost two thirds of NSOs have a data ethics policy in place.
50 NSOs indicate to have a data ethics policy in place, with 63% of NSOs confirming to do so, and 19% of all respondents trying to establish one. Only 14% of NSOs do not plan to implement an ethics policy.

8. National quality assurance framework

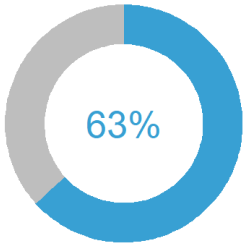
Over a third of respondents do have a national quality framework in place. 32 NSOs indicate to have a national quality framework in place, with 41% of respondents to the survey indicate to have done so, and 44% are trying to establish one. Only 14% of respondents indicate to not plan any to establish any quality framework.

NSO able to negotiate data provision with their partners:



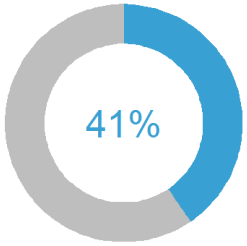
Answer	Count	Percent
Yes	40	51%
Don't know	5	6%
No, but trying to establish	28	35%
No, none planned	6	8%

NSO has a data ethics policy:



Answer	Count	Percent
Yes	50	63%
Don't know	3	4%
No, but trying to establish	15	19%
No, none planned	11	14%

NSO has a national quality assurance framework:



Answer	Count	Percent
Yes	32	41%
Don't know	1	1%
No, but trying to establish	35	44%
No, none planned	11	14%

Legal Framework

The Legal Framework section of the questionnaire aimed to assess the establishment of (or plans for) a legal framework for data access and data sharing within NSOs, their NSS, and potentially wider.

1. Legal framework with access to big data

44% of the responding NSOs have a legal framework that covers access to big data from other Government Departments and big data partnerships.
35 of NSOs indicate that there is a legal framework in place, with 44% confirming to have one and 39% trying to establish one. Only 10% of respondents do not consider a legal framework, and 6% do not know.

2. Legal framework for data access for academia

Over two thirds of NSOs have a legal framework in place that allows academia to safely access their data.
58 NSOs indicate to have a framework for data access for academia in place, with 73% of all respondents confirming to have such a framework, 15% aim to establish a framework, and only 9% do not aim to establish one.

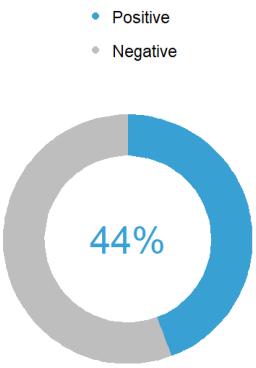
3. Legal framework that penalizes data disclosures

The large majority have a legal framework in place that penalizes data disclosures.
69 respondents indicate to have a legal framework in place that penalizes data disclosures, with 91% of respondents confirming to have one. 3% of respondents are trying to establish one and only 1% of respondents does not plan to implement such a framework.

4. Data protection law

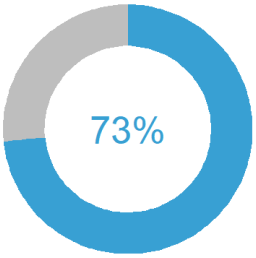
47 of the respondents have established an overarching data protection law (59%), 18% are trying to establish one. 11% of respondents are not considering the establishment of a data protection law.

NSO has legal framework that covers access to Big Data from other Government departments and partnerships:



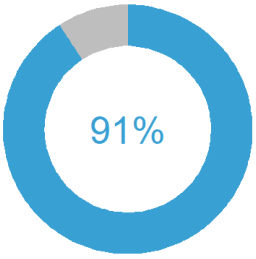
Answer	Count	Percent
Yes	35	44%
Don't know	5	6%
No, but trying to establish	31	39%
No, not considered	8	10%

NSO has a legal framework to make data safely available to accredited researchers:



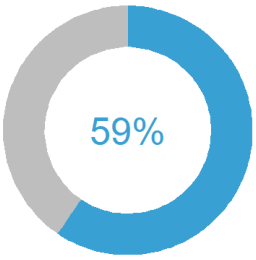
Answer	Count	Percent
Yes	58	73%
Don't know	2	3%
No, but trying to establish	12	15%
No, not considered	7	9%

NSO has a legal framework that will impose penalties for unlawful disclosure of confidential personal data:



Answer	Count	Percent
Yes	69	91%
Don't know	4	5%
No, but trying to include this	2	3%
No, not considered	1	1%

Overarching data protection law exists that can be applied to use of Big Data in Official Statisitcs:



Answer	Count	Percent
Yes	47	59%
Don't know	9	11%
No, but trying to establish	14	18%
No, not considered	9	11%

Does or will the legal framework restrict data from being stored outside the country?

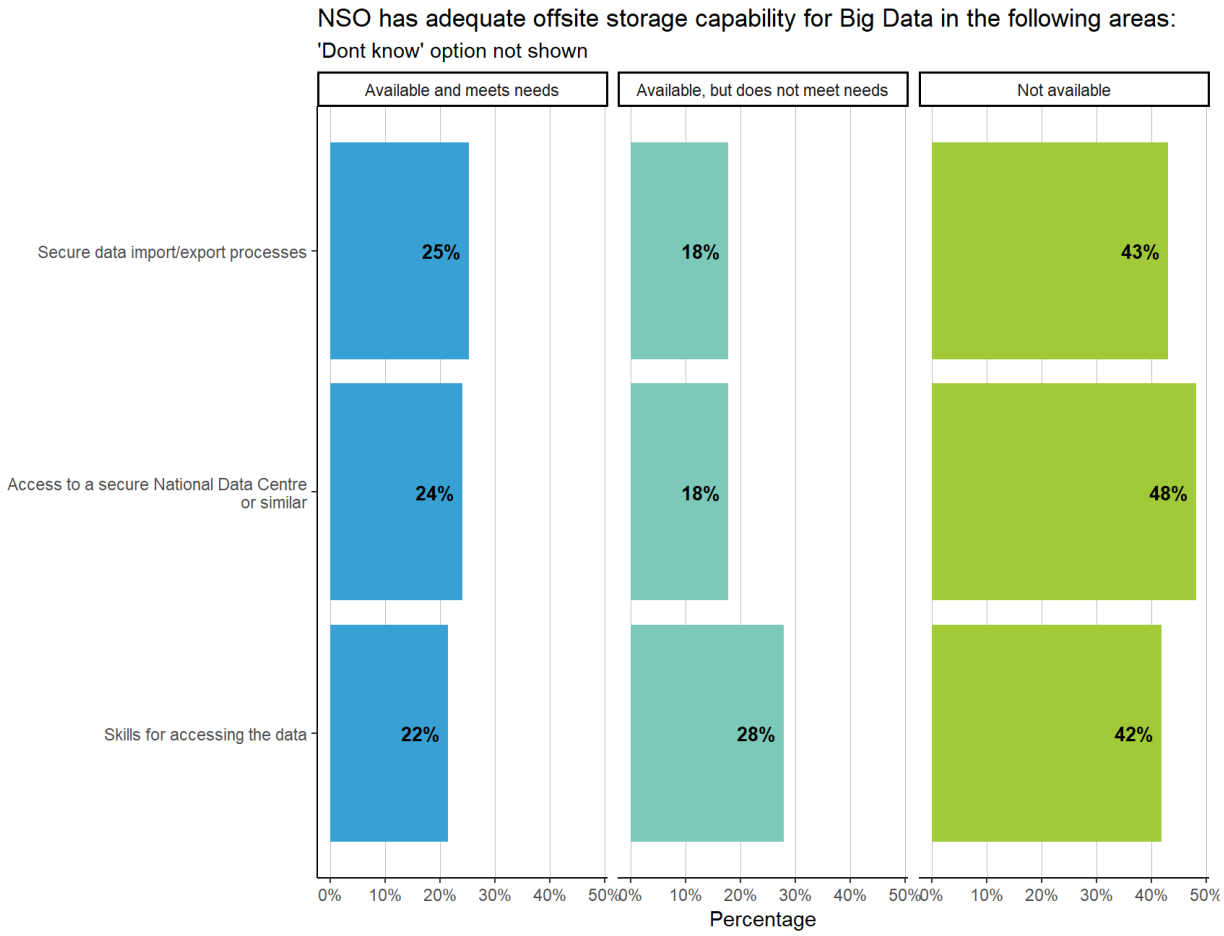
Statistic	Yes	No, but trying to include this	No, not considered	Don't know
Count	33	2	30	11
Percent	43%	3%	39%	14%

IT Infrastructure

The IT infrastructure section of the report outlines the extent of, or future plans for, IT infrastructure within NSOs. This section also assesses how the IT of NSOs enables big data analytics in a secure environment. The IT Infrastructure for many NSOs seems to be presenting more of a challenge for incorporating big data. The below graph depicts the responses to questions around onsite data storage capability, computing power and skills at the NSO. The results show that approximately:

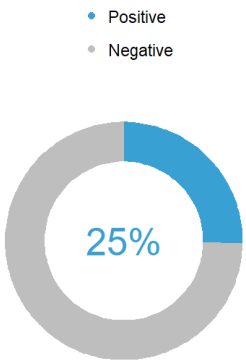
- 42% of NSOs have the appropriate processes in place for secure import / export of the data;
- 52% have adequate (i.e. un-interrupted) power supply;
- 32% have the required skills for accessing the data. Hence, the challenges posed are around the lack of onsite storage and computing power onsite, plus the lack of required skills for accessing the data.

Other data collected shows that around 24% of NSOs have access to, and are using, offsite national data centres. Access to a secure data centre was not available to 48% of the NSOs surveyed.



25% of NSOs reported having secure cloud infrastructure. Secure cloud infrastructure is not being considered by 33% of the NSOs surveyed.

NSO stores and shares data using secure cloud infrastructure:



Answer	Count	Percent
Yes	20	25%
Don't know	5	6%
No, but trying to establish	28	35%
No, not considered	26	33%

Human Resources

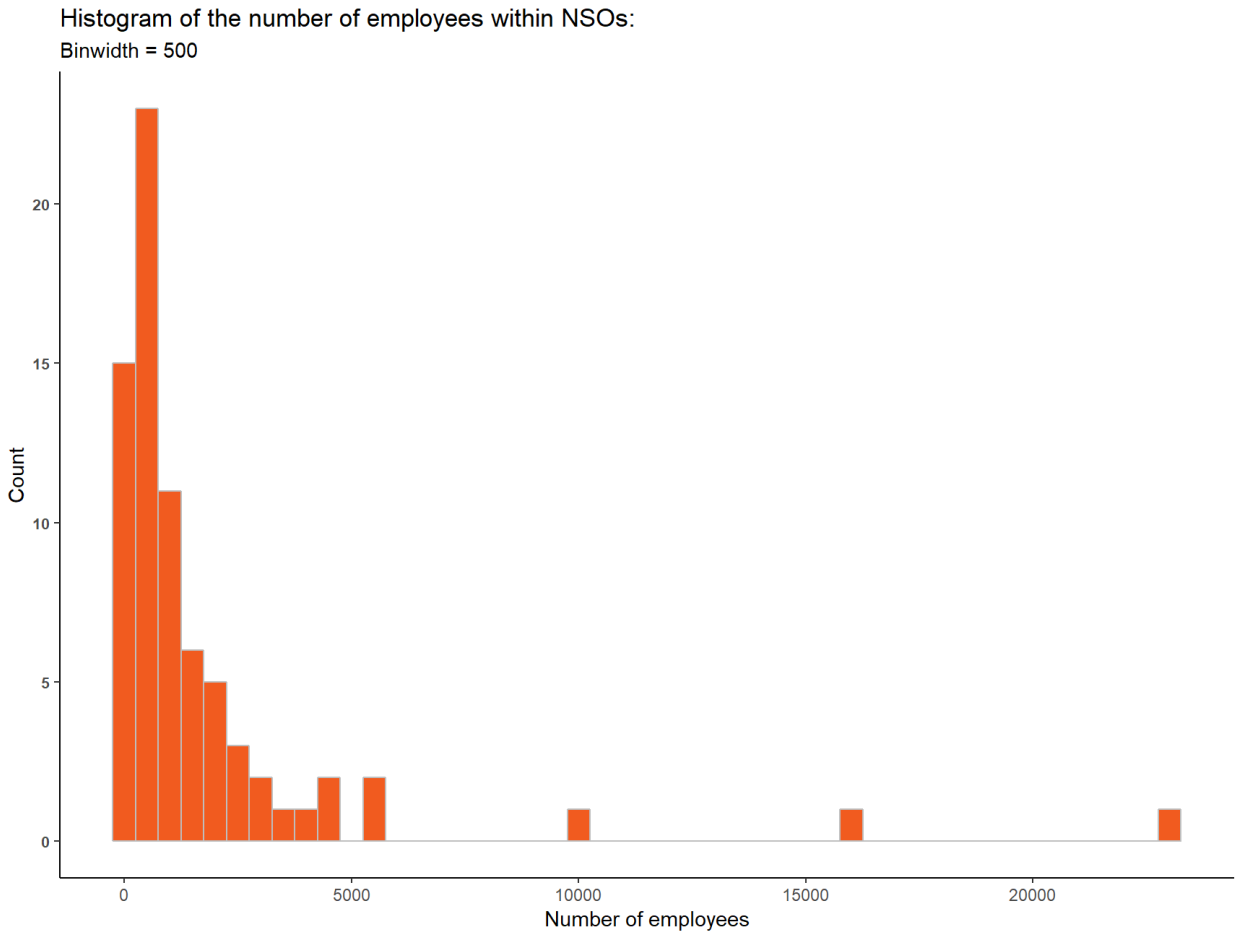
The Human Resources section of the survey asked questions on the number of data science posts and practitioners within each NSO. It also asked questions on skills gaps and the future plans for recruitment and growth. This included the skills needed to develop and maintain appropriate methodologies.

1. Staff numbers

Unsurprisingly, the size of NSOs varied. To categorise them the NSOs were grouped from small to very large. Those with a size less than 500 were categorised as small, between 500 and 2,499 as medium, between 2,500 and 5,000 as large, and, those with more than 5,000 were grouped as very large. There were 75 valid responses that could be used to group NSOs.

Size groups of NSOs

Size	Count
Small	26
Medium	37
Large	6
Very large	5



Summary statistics on the total number of employees

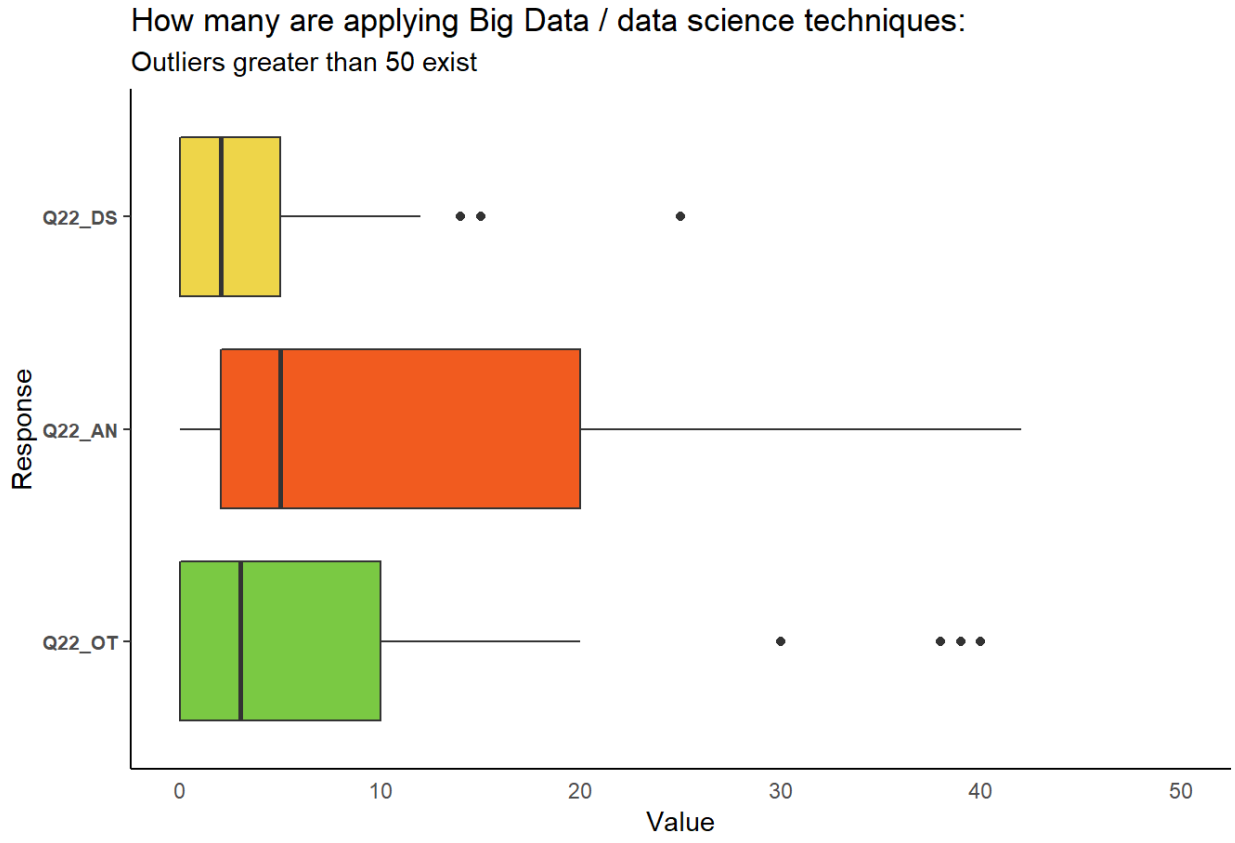
minimum	q1	median	mean	q3	maximum	na
1	334.5	700	1776.27	1886	22969	26

The survey also asked NSOs how many of their employees are applying Big Data / data science techniques. NSOs were asked:

- The number of qualified “Data Scientists” at MSc or PhD level
- The number of analysts who are applying Big Data / data science techniques
- Others who are applying Big Data / data science techniques, such as IT professionals

Summary statistics on the type of employees

Type	minimum	q1	median	mean	q3	maximum	na
Data Scientist	0	0.00	2	10.39130	8	149	54
Analyst	0	2.00	5	20.73913	20	250	54
Other	0	0.25	3	16.02174	10	400	54



Before reading the following please consider the accuracy of responses from NSOs, the staff number given by NSOs may be estimates. It would not be unreasonable to hypothesize that the larger the NSO, the greater number of staff there is applying data science techniques. This appears true for the number of qualified at MSc/PhD level and the number of analysts. Evidence does not exist to suggest that the number of other roles applying techniques has a correlation with total staff number at NSOs.

Evidence suggests a moderate relationship between size of NSO and the amount of qualified data scientists. It also suggests a weak relationship between size of NSO and analysts applying data science skills.

Type	estimate	statistic	p.value	method	alternative
Data Scientist	0.4953287	7161.286	0.0006296	Spearman's rank correlation rho	two.sided
Analyst	0.3082279	9816.247	0.0417956	Spearman's rank correlation rho	two.sided
Other	0.1692327	11788.588	0.2721214	Spearman's rank correlation rho	two.sided

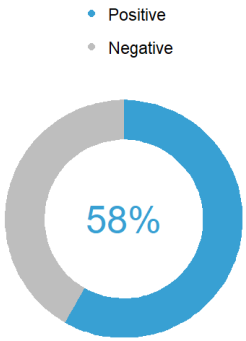
2. External recruitment strategy

Only 33% of the NSOs surveyed reported having a strategy for recruiting external staff. A greater number (42%) are looking to establish a strategy. Reasons cited for difficulty in recruiting include having no coordinated strategy and/or one that is specific to hiring the technical experts needed for Big Data / data science work. There is a feeling that the lack of competitive benefits that can be offered provides NSOs with difficulties when actively looking to employ experts.

3. Internal upskilling strategy

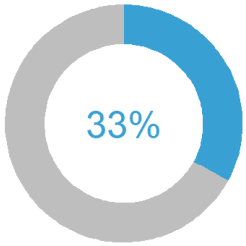
A strategy to upskill current employees appears to be a higher priority for NSOs with 41% having a strategy for this with over half (52%) trying to establish one. Included within these strategies include taught and self taught training programmes ranging from introductory to Masters degree level, international collaborations, R and Python implementation, training roadmaps and curriculums, and also, networking groups. Generally, it appears as though NSOs hold themselves responsible for upskilling staff in this area. There is also the effort by some to partner with academia and other nations to provide the high class training that NSOs may struggle to provide internally.

NSO has employees applying Big Data / data science techniques:



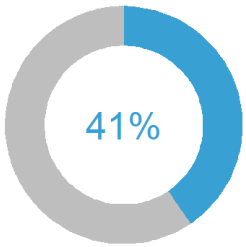
Answer	Count	Percent
Yes	46	58%
Don't know	10	13%
No	23	29%

NSO has a strategy to recruit new employees who are able to apply Big Data / data science techniques:



Answer	Count	Percent
Yes	26	33%
Don't know	7	9%
No, but trying to establish	33	42%
No, not considered	13	16%

NSO has a strategy to develop Big Data / data science skills of current employees:



Answer	Count	Percent
Yes	32	41%
Don't know	2	3%
No, but trying to establish	41	52%
No, not considered	4	5%

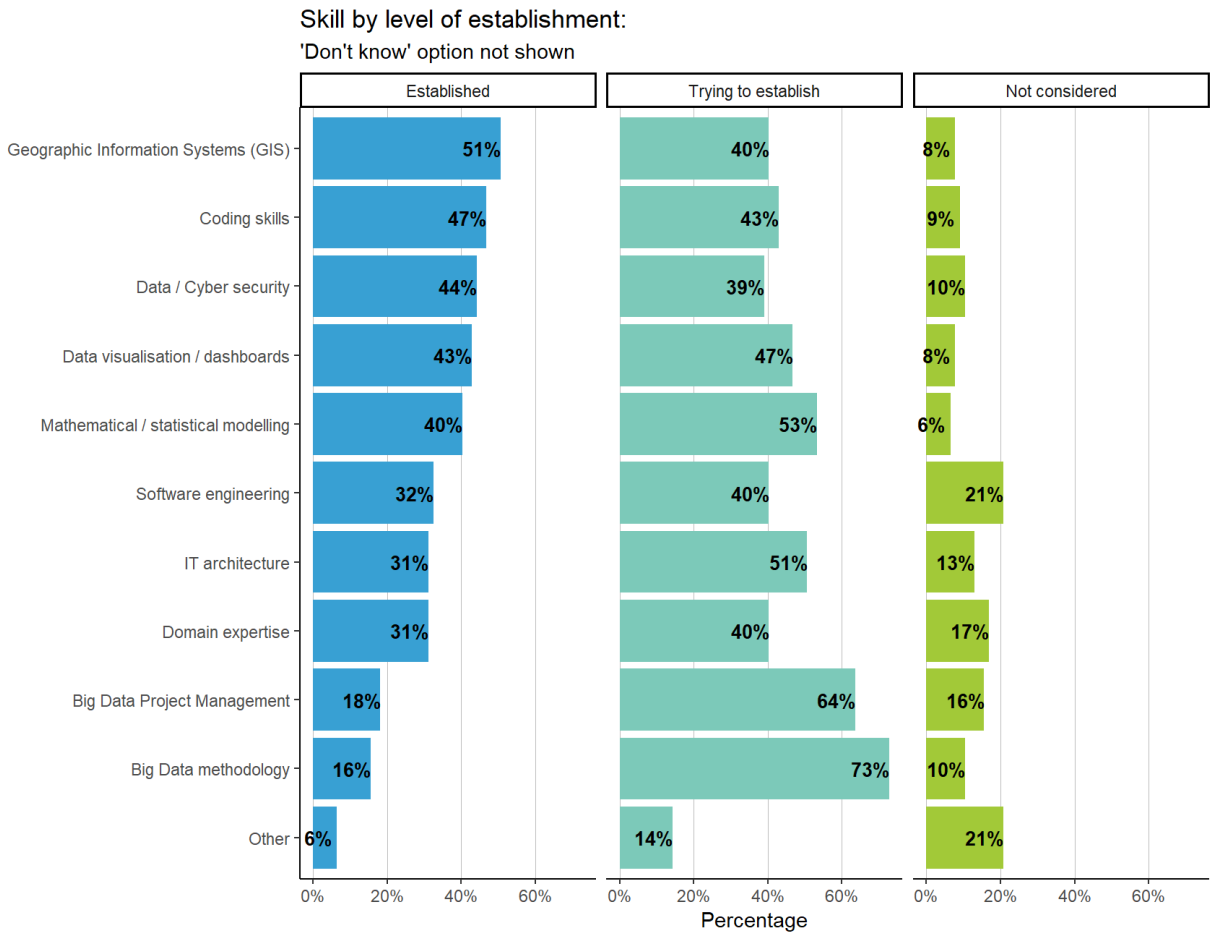
4. Existing Big Data skills

Although many NSOs have already accessed training or are in the process of establishing it, there are still large gaps in big data skills that are apparent from the survey responses.

Of the NSOs that responded to direct questions about big data skills, the most established skills are identified as: Geographic Information Systems (51%) and Coding skills (47%); whilst the least developed skills (or most needed) are: Big Data Methodology (73%), Big Data Project Management (64%), and Mathematical / Statistical modelling (53%). Skills that were written following the 'Other' response included topics such as data engineering and a need for improved domain knowledge.

A point raised by one NSO and something to consider is that because of the rapid development of new tools there exists an ongoing need for training and improvements even in skills that are deemed established.

Turnover of staff is another factor that contributes to a need for ongoing training.



5. Delivery of training

Big data / data science skills and access to training is of high interest to the Task Team, since it is also tasked with developing a Competency Framework against which training will be mapped. Big Data / data science training has been delivered by 47% of NSOs surveyed. 28 (35%) of the NSOs are trying to establish this training. The main focus of this delivered training appears to be on the usage of the programming languages R, Python and SQL. As well as looking to upskill coding skills, NSOs reported delivering training on reproducibility and machine learning.

6. Academic partnerships

Partnerships with academia can prove useful in supporting NSOs with Big Data work. These partnerships were reported to help not just on a project resource basis but also with providing NSOs with access to highly skilled technical experts and domain knowledge expertise. These partnerships have been built by less than half of the NSOs (46%). It appears as though the value of these partnerships is generally well considered with a further 38% trying to establish partnerships. Partnerships with international organisations, other NSOs and international universities were also mentioned. Collaborations with organisations such as Eurostat was mentioned and one NSO reported that their collaboration with an international University on a social media project was highly successful and resulted in national and international press coverage. This project was reported to have helped to accelerate Big Data adoption in the NSO.

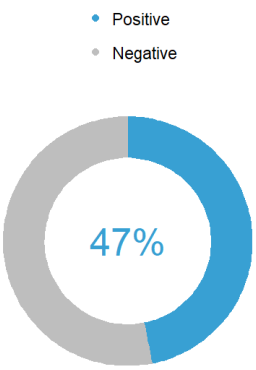
7. Competency framework

The need for a competency framework is emphasised by the results of the survey. Only 8% of NSOs surveyed stated that they had a Big Data / data science competency framework. Over half (52%) of NSOs are looking to establish one.

8. Career pathway

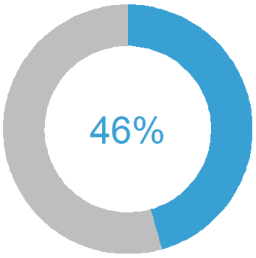
The amount of NSOs with a data scientist career pathway was also reported to be low. The uptake was also 8% but over half (54%) were not considering trying to establish one compared to 30% that are. The pathways may exist but not specifically for data scientist positions.

NSO has delivered Big Data / data science training:



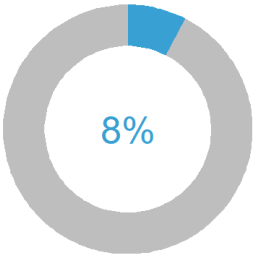
Answer	Count	Percent
Yes	37	47%
Don't know	2	3%
No, but trying to establish	28	35%
No, not considered	12	15%

NSO has partnerships with academia and other NSO to support Big Data work:



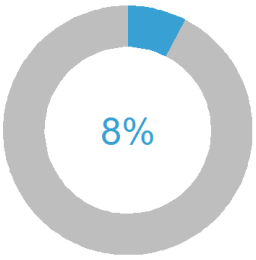
Answer	Count	Percent
Yes	36	46%
Don't know	3	4%
No, but trying to establish	30	38%
No, not considered	10	13%

NSO has a Big Data / data science competency framework:



Answer	Count	Percent
Yes	6	8%
Don't know	5	6%
No, but trying to establish	41	52%
No, not considered	27	34%

NSO has a career pathway for those applying Big Data / data science techniques:



Answer	Count	Percent
Yes	6	8%
Don't know	6	8%
No, but trying to establish	24	30%
No, not considered	43	54%

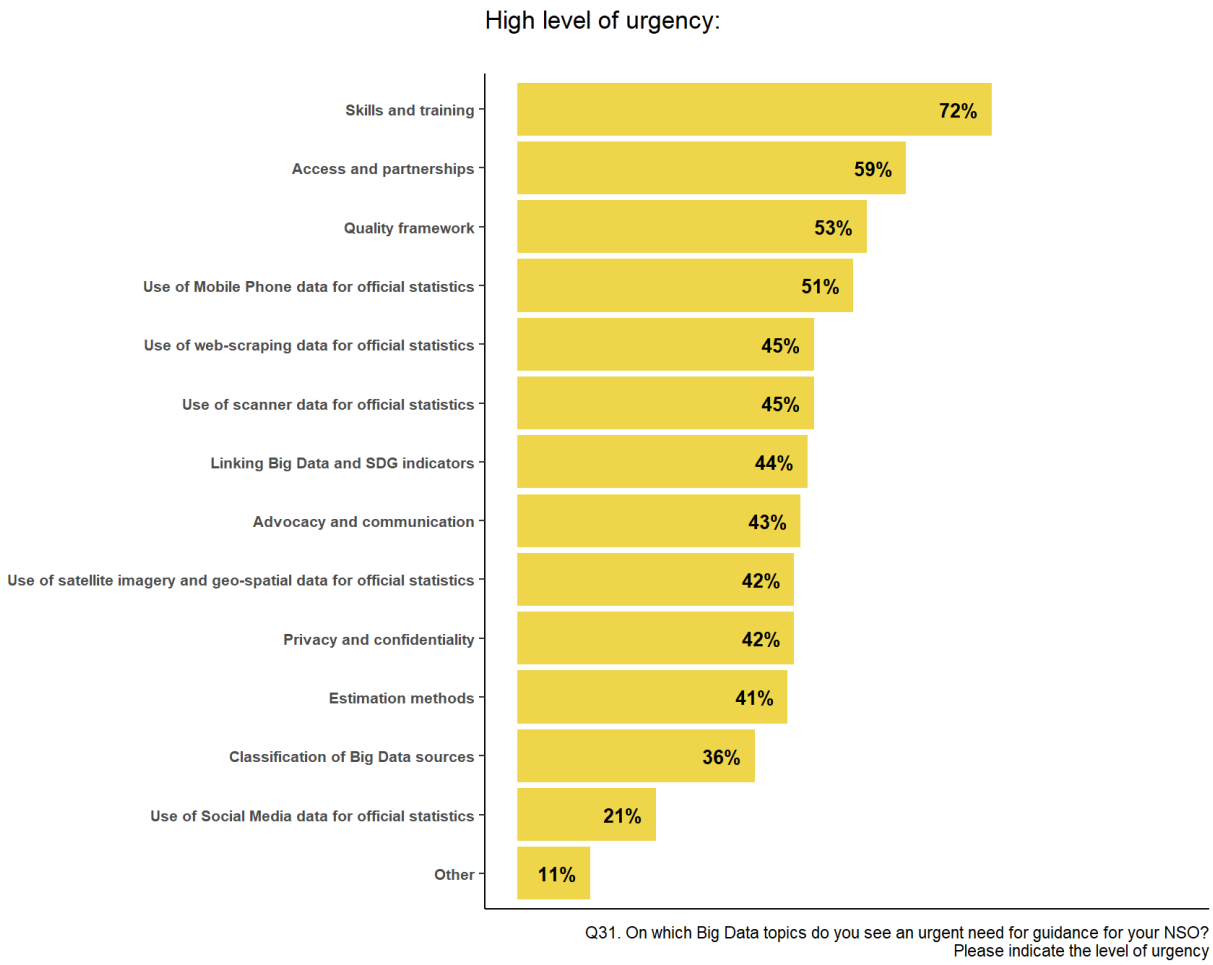
9. Challenges to delivering training

NSOs face a number of challenges to delivering Big Data / data science training to their employees. The most frequently cited reason appears to be budget constraints limiting their ability to fund the necessary training. Another problem is some NSOs have difficulty accessing or hiring technical experts and highly skilled trainers that can upskill their workforce. The need

for some to source this training externally can incur higher costs.
One NSO provides an example that it is difficult to hire staff that tick both boxes of highly skilled computer science and also maths skills. This suggests that even highly skilled graduates in respective fields may require training offers in the field of Big Data and data science.
The spread of skills that are interpreted as being as necessary when working with Big Data / data science may present a challenge to NSOs in where to focus their training.
One NSO cited that a challenge for them was inaccessibility to the best practice of other countries. It could perhaps be considered that encouragement on the sharing of training materials and best practice training is something that may be an appropriate recommendation.

Guidance

The final section of the survey asked NSOs to indicate the level of urgency for guidance on big data topic areas. The below graph highlights responses, with the highest urgency identified as ‘Skills and Training’ (72%). Although in the previous section, it was reported that training had been delivered by 47% of respondents, there are clearly still gaps in the provision and access to training that requires further investigation. The need for guidance is also high for ‘Access and partnerships’ (59%), and ‘Quality frameworks’ (53%).



Q31. On which Big Data topics do you see an urgent need for guidance for your NSO? Please indicate the level of urgency

Annex

Scoring readiness for using Big Data in Official Statistics

Methodology

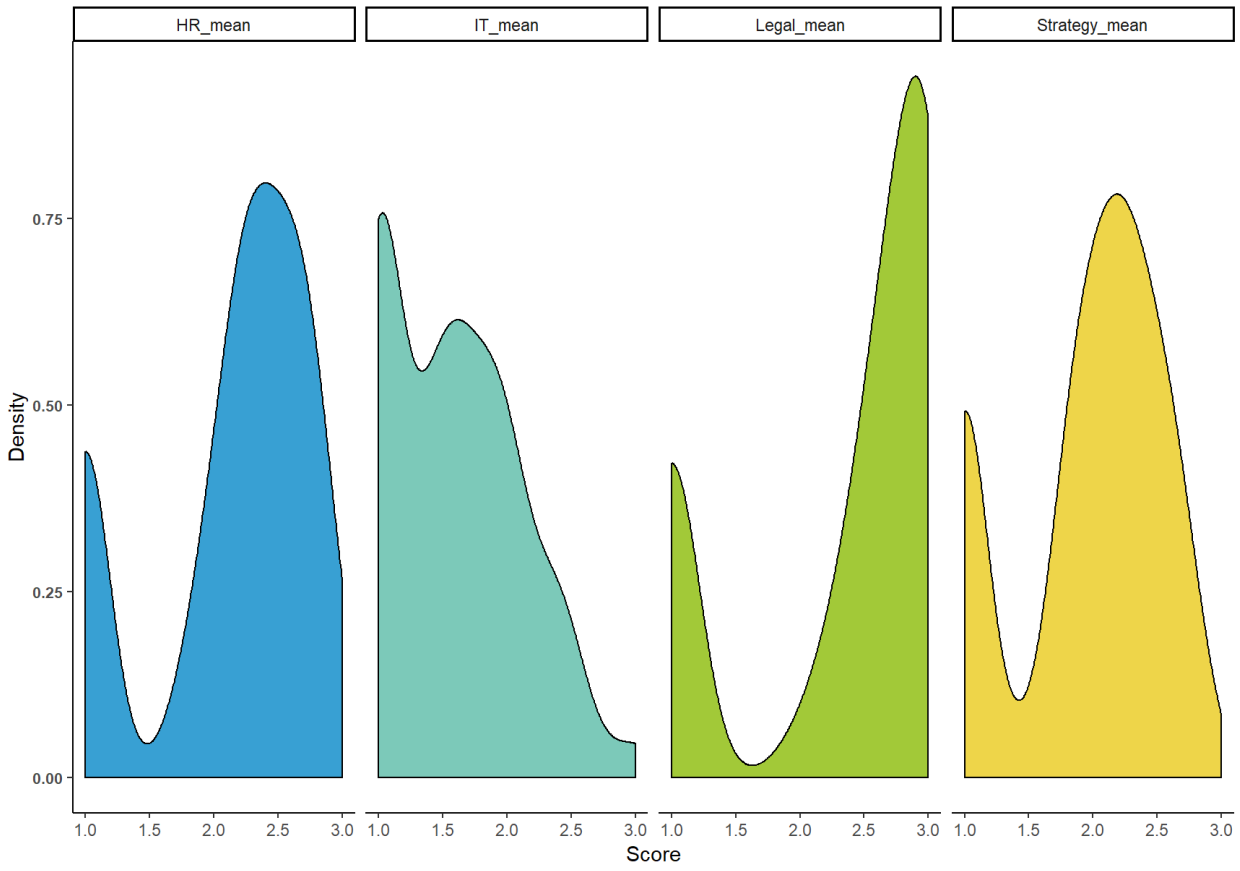
To derive a score for each section we have decided to take the average (mean) of questions that fall within each respective section. Only questions that appear to follow an actual order were included in deriving the scores. Radio grid questions (those that asked for information on a number of topics under the same heading) were averaged (mean) to count only as one question. The method of taking the mean instead of the median was chosen due to the low number of response options in questions. It was felt that this would distinguish between nations better than other measures of central tendency such as the median or mode. It should be noted that nonresponse bias may exist with nations low in terms of Big Data readiness not completing the survey.

Code

Summary statistics for section scores:

Strategy_mean	Legal_mean	IT_mean	HR_mean
Min. :1.000	Min. :1.000	Min. :1.000	Min. :1.000
1st Qu.:1.725	1st Qu.:2.188	1st Qu.:1.000	1st Qu.:1.900
Median :2.067	Median :2.750	Median :1.500	Median :2.296
Mean :1.953	Mean :2.369	Mean :1.575	Mean :2.118
3rd Qu.:2.376	3rd Qu.:3.000	3rd Qu.:2.000	3rd Qu.:2.616
Max. :2.907	Max. :3.000	Max. :3.000	Max. :2.983

Distribution of scores:



NSO readiness by region

The data provides no evidence at present to suggest statistically significant differences in Strategy, IT or Legal between the chosen regional groups of Africa, Asia and Pacific, America, and Europe. There does appear to be significant differences between the regions in regards to scores on Human Resources. It appears that this significance difference exists between Europe and Africa.

Code

Score	statistic	p.value	method
Strategy	3.797101	0.2842235	Kruskal-Wallis rank sum test
Legal	5.815305	0.1209501	Kruskal-Wallis rank sum test
IT	2.931269	0.4023439	Kruskal-Wallis rank sum test
HR	9.897512	0.0194577	Kruskal-Wallis rank sum test

Code

Pairwise Wilcoxon Rank Sum Tests:

group1	group2	p.value
Americas	Africa	0.4637900
Asia and Pacific	Africa	0.1130868
Europe	Africa	0.0368025
Asia and Pacific	Americas	0.3252033
Europe	Americas	0.1130868
Europe	Asia and Pacific	0.1881669

NSO readiness by size of NSO

A weak positive relationship appears to exist between size of NSO and Strategy score. There also appears to be a weak positive relationship between size of NSO and HR score. The extent of our research suggests no relationship between NSO size and the other scored areas.

Code

Score	estimate	p.value	statistic	alternative	method
Strategy	0.2566856	0.0272700	50192.30	two.sided	Spearman’s rank correlation rho
Legal	0.1178838	0.3171635	59564.90	two.sided	Spearman’s rank correlation rho
IT	-0.0115782	0.9220064	68306.81	two.sided	Spearman’s rank correlation rho
HR	0.3313481	0.0039293	45150.72	two.sided	Spearman’s rank correlation rho

NSO readiness by development status

Of the relationships investigated, development status appears to relate most to big data readiness. Our exploration into this returned statistically significant results for Strategy, Legal frameworks, and Human Resources. No significant result was found for IT infrastructure.

Code

Development status of sample	Frequency
Developed	37
Developing	63

Code

Score	statistic	p.value	alternative	method
Strategy	1485.0	0.0221346	two.sided	Wilcoxon rank sum test with continuity
Legal	1526.0	0.0076733	two.sided	Wilcoxon rank sum test with continuity
IT	1208.5	0.7569799	two.sided	Wilcoxon rank sum test with continuity
HR	1556.0	0.0051485	two.sided	Wilcoxon rank sum test with continuity

Tables

Strategic Data Science Coordination

Do you have any Big Data / data science projects within your NSO?

Country	Statistic	No, but trying to establish	No, none planned	Yes
Africa	Count	3	6	6
Africa	Percent	20%	40%	40%
Americas	Count	7	5	5
Americas	Percent	41%	29%	29%
Asia and Pacific	Count	17	6	13
Asia and Pacific	Percent	47%	17%	36%
Europe	Count	5	4	23
Europe	Percent	16%	12%	72%

Do you have any Big Data / data science projects within your NSO?

Country	Statistic	No, but trying to establish	No, none planned	Yes
Total	Count	32	21	47
Total	Percent	32%	21%	47%

Does your NSO currently have a strategy for using Big Data in Official Statistics?

Country	Statistic	No, but trying to establish	No, none planned	Yes
Africa	Count	6	1	3
Africa	Percent	60%	10%	30%
Americas	Count	8	1	4
Americas	Percent	62%	8%	31%
Asia and Pacific	Count	19	1	10
Asia and Pacific	Percent	63%	3%	33%

Does your NSO currently have a strategy for using Big Data in Official Statistics?

Country	Statistic	No, but trying to establish	No, none planned	Yes
Europe	Count	16	1	11
Europe	Percent	57%	4%	39%
Total	Count	49	4	28
Total	Percent	60%	5%	35%

Who does / will the strategy apply to?

Country	Statistic	NSO and the wider NSS	NSO only
Africa	Count	2	1
Africa	Percent	67%	33%
Americas	Count	1	3
Americas	Percent	25%	75%

Who does / will the strategy apply to?

Country	Statistic	NSO and the wider NSS	NSO only
Asia and Pacific	Count	8	2
Asia and Pacific	Percent	80%	20%
Europe	Count	2	9
Europe	Percent	18%	82%
Total	Count	13	15
Total	Percent	46%	54%

Does your NSO currently have a Data Science Lead / Chief Data Officer?

Country	Statistic	Don’t know	No, but trying to establish	No, none planned	Yes
Africa	Count	NA	4	1	4
Africa	Percent	NA	44%	11%	44%

Does your NSO currently have a Data Science Lead / Chief Data Officer?

Country	Statistic	Don’t know	No, but trying to establish	No, none planned	Yes
Americas	Count	NA	2	8	3
Americas	Percent	NA	15%	62%	23%
Asia and Pacific	Count	1	16	6	7
Asia and Pacific	Percent	3%	53%	20%	23%
Europe	Count	1	12	9	6
Europe	Percent	4%	43%	32%	21%
Total	Count	2	34	24	20
Total	Percent	2%	42%	30%	25%

Is your NSO able to discuss or negotiate how data are provided (e.g. in which format) by MOST (at least half) of the Big Data partnerships identified previously?

Country	Statistic	Don’t know	No, but trying to establish	No, none planned	Yes
Africa	Count	1	2	1	5
Africa	Percent	11%	22%	11%	56%
Americas	Count	NA	5	2	5
Americas	Percent	NA	42%	17%	42%
Asia and Pacific	Count	1	15	2	12
Asia and Pacific	Percent	3%	50%	7%	40%
Europe	Count	3	6	1	18
Europe	Percent	11%	21%	4%	64%
Total	Count	5	28	6	40
Total	Percent	6%	35%	8%	51%

Does your NSO have a Data Ethics policy?

Country	Statistic	Don’t know	No, but trying to establish	No, none planned	Yes
Africa	Count	1	NA	1	7
Africa	Percent	11%	NA	11%	78%
Americas	Count	NA	4	2	6
Americas	Percent	NA	33%	17%	50%
Asia and Pacific	Count	NA	5	4	21
Asia and Pacific	Percent	NA	17%	13%	70%
Europe	Count	2	6	4	16
Europe	Percent	7%	21%	14%	57%
Total	Count	3	15	11	50
Total	Percent	4%	19%	14%	63%

Does your NSO have a National Quality Assurance Framework, a regional or national Code of Practice or similar, that covers the use of Big Data in Official Statistics?

Country	Statistic	Don't know	No, but trying to establish	No, none planned	Yes
Africa	Count	NA	3	NA	6
Africa	Percent	NA	33%	NA	67%
Americas	Count	NA	5	2	5
Americas	Percent	NA	42%	17%	42%
Asia and Pacific	Count	NA	16	5	9
Asia and Pacific	Percent	NA	53%	17%	30%
Europe	Count	1	11	4	12
Europe	Percent	4%	39%	14%	43%
Total	Count	1	35	11	32
Total	Percent	1%	44%	14%	41%

Legal Framework

Is there a legal framework for your NSO that covers access to Big Data from other Government departments and Big Data partnerships?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Africa	Count	NA	3	1	5
Africa	Percent	NA	33%	11%	56%
Americas	Count	2	3	3	4
Americas	Percent	17%	25%	25%	33%
Asia and Pacific	Count	1	16	1	12
Asia and Pacific	Percent	3%	53%	3%	40%
Europe	Count	2	9	3	14
Europe	Percent	7%	32%	11%	50%
Total	Count	5	31	8	35

Is there a legal framework for your NSO that covers access to Big Data from other Government departments and Big Data partnerships?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Total	Percent	6%	39%	10%	44%

Is there a legal framework for your NSO to make data available safely and securely to accredited researchers (e.g. academics) for public good research?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Africa	Count	NA	1	NA	8
Africa	Percent	NA	11%	NA	89%
Americas	Count	2	2	2	6
Americas	Percent	17%	17%	17%	50%
Asia and Pacific	Count	NA	4	3	23
Asia and Pacific	Percent	NA	13%	10%	77%

Is there a legal framework for your NSO to make data available safely and securely to accredited researchers (e.g. academics) for public good research?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Europe	Count	NA	5	2	21
Europe	Percent	NA	18%	7%	75%
Total	Count	2	12	7	58
Total	Percent	3%	15%	9%	73%

Does or will the legal framework impose penalties for anyone (including your NSO) who unlawfully discloses confidential personal data?

Country	Statistic	Don't know	No, but trying to include this	No, not considered	Yes
Africa	Count	NA	NA	NA	9
Africa	Percent	NA	NA	NA	100%
Americas	Count	1	1	1	8

Does or will the legal framework impose penalties for anyone (including your NSO) who unlawfully discloses confidential personal data?

Country	Statistic	Don't know	No, but trying to include this	No, not considered	Yes
Americas	Percent	9%	9%	9%	73%
Asia and Pacific	Count	NA	NA	NA	29
Asia and Pacific	Percent	NA	NA	NA	100%
Europe	Count	3	1	NA	23
Europe	Percent	11%	4%	NA	85%
Total	Count	4	2	1	69
Total	Percent	5%	3%	1%	91%

Does or will the legal framework restrict data from being stored outside the country?

Country	Statistic	Don't know	No, but trying to include this	No, not considered	Yes
Africa	Count	1	NA	5	3

Does or will the legal framework restrict data from being stored outside the country?

Country	Statistic	Don’t know	No, but trying to include this	No, not considered	Yes
Africa	Percent	11%	NA	56%	33%
Americas	Count	2	NA	4	5
Americas	Percent	18%	NA	36%	45%
Asia and Pacific	Count	1	1	11	16
Asia and Pacific	Percent	3%	3%	38%	55%
Europe	Count	7	1	10	9
Europe	Percent	26%	4%	37%	33%
Total	Count	11	2	30	33
Total	Percent	14%	3%	39%	43%

Is there an overarching data protection law that can be applied to the use of Big Data in Official Statistics, or for sharing data collected for those purposes?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Africa	Count	NA	1	3	5
Africa	Percent	NA	11%	33%	56%
Americas	Count	4	1	2	5
Americas	Percent	33%	8%	17%	42%
Asia and Pacific	Count	NA	10	4	16
Asia and Pacific	Percent	NA	33%	13%	53%
Europe	Count	5	2	NA	21
Europe	Percent	18%	7%	NA	75%
Total	Count	9	14	9	47
Total	Percent	11%	18%	11%	59%

IT Infrastructure

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following:

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	5	3	NA	1
Africa	Percent	56%	33%	NA	11%
Americas	Count	7	3	1	1
Americas	Percent	58%	25%	8%	8%
Asia and Pacific	Count	6	16	NA	8
Asia and Pacific	Percent	20%	53%	NA	27%
Europe	Count	6	11	3	8
Europe	Percent	21%	39%	11%	29%
Total	Count	24	33	4	18

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following:

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Total	Percent	30%	42%	5%	23%

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Secure data import/export processes

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	4	3	NA	2
Africa	Percent	44%	33%	NA	22%
Americas	Count	5	5	1	1
Americas	Percent	42%	42%	8%	8%
Asia and Pacific	Count	11	14	NA	5
Asia and Pacific	Percent	37%	47%	NA	17%

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Secure data import/export processes

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Europe	Count	13	5	4	6
Europe	Percent	46%	18%	14%	21%
Total	Count	33	27	5	14
Total	Percent	42%	34%	6%	18%

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Power supply

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	7	2	NA	NA
Africa	Percent	78%	22%	NA	NA
Americas	Count	8	3	1	NA

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Power supply

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Americas	Percent	67%	25%	8%	NA
Asia and Pacific	Count	13	11	NA	6
Asia and Pacific	Percent	43%	37%	NA	20%
Europe	Count	13	5	6	4
Europe	Percent	46%	18%	21%	14%
Total	Count	41	21	7	10
Total	Percent	52%	27%	9%	13%

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Computing power

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	4	2	1	2

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Computing power

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Percent	44%	22%	11%	22%
Americas	Count	4	6	1	1
Americas	Percent	33%	50%	8%	8%
Asia and Pacific	Count	8	15	NA	7
Asia and Pacific	Percent	27%	50%	NA	23%
Europe	Count	8	10	3	7
Europe	Percent	29%	36%	11%	25%
Total	Count	24	33	5	17
Total	Percent	30%	42%	6%	22%

Presence of air conditioning

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don’t know	Not available
Africa	Count	8	1	NA	NA
Africa	Percent	89%	11%	NA	NA
Americas	Count	8	2	1	1
Americas	Percent	67%	17%	8%	8%
Asia and Pacific	Count	14	10	NA	6
Asia and Pacific	Percent	47%	33%	NA	20%
Europe	Count	14	4	6	4
Europe	Percent	50%	14%	21%	14%
Total	Count	44	17	7	11
Total	Percent	56%	22%	9%	14%

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Skills for accessing the data

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	1	6	NA	2
Africa	Percent	11%	67%	NA	22%
Americas	Count	7	3	1	1
Americas	Percent	58%	25%	8%	8%
Asia and Pacific	Count	8	16	NA	6
Asia and Pacific	Percent	27%	53%	NA	20%
Europe	Count	9	14	2	3
Europe	Percent	32%	50%	7%	11%
Total	Count	25	39	3	12

Does your NSO have adequate onsite data storage capability for Big Data in terms of the following: Skills for accessing the data

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Total	Percent	32%	49%	4%	15%

Does your NSO have adequate offsite data storage capability for Big Data in terms of the following: Access to a secure National Data Centre or similar

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	2	NA	NA	7
Africa	Percent	22%	NA	NA	78%
Americas	Count	3	1	3	5
Americas	Percent	25%	8%	25%	42%
Asia and Pacific	Count	10	6	1	13
Asia and Pacific	Percent	33%	20%	3%	43%

Does your NSO have adequate offsite data storage capability for Big Data in terms of the following: Access to a secure National Data Centre or similar

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Europe	Count	4	7	4	13
Europe	Percent	14%	25%	14%	46%
Total	Count	19	14	8	38
Total	Percent	24%	18%	10%	48%

Does your NSO have adequate offsite data storage capability for Big Data in terms of the following: Secure data import/export processes

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	2	NA	NA	7
Africa	Percent	22%	NA	NA	78%
Americas	Count	4	NA	4	4

Does your NSO have adequate offsite data storage capability for Big Data in terms of the following: Secure data import/export processes

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Americas	Percent	33%	NA	33%	33%
Asia and Pacific	Count	9	9	1	11
Asia and Pacific	Percent	30%	30%	3%	37%
Europe	Count	5	5	6	12
Europe	Percent	18%	18%	21%	43%
Total	Count	20	14	11	34
Total	Percent	25%	18%	14%	43%

Does your NSO have adequate offsite data storage capability for Big Data in terms of the following: Skills for accessing the data

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Africa	Count	2	NA	NA	7
Africa	Percent	22%	NA	NA	78%
Americas	Count	3	1	3	5
Americas	Percent	25%	8%	25%	42%
Asia and Pacific	Count	6	12	1	11
Asia and Pacific	Percent	20%	40%	3%	37%
Europe	Count	6	9	3	10
Europe	Percent	21%	32%	11%	36%
Total	Count	17	22	7	33

Does your NSO have adequate offsite data storage capability for Big Data in terms of the following: Skills for accessing the data

Country	Statistic	Available and meets needs	Available, but does not meet needs	Don't know	Not available
Total	Percent	22%	28%	9%	42%

Does your NSO store and share data using secure cloud infrastructure? By secure cloud infrastructure, we are referring to private or government cloud environments.

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Africa	Count	NA	3	3	3
Africa	Percent	NA	33%	33%	33%
Americas	Count	1	3	5	3
Americas	Percent	8%	25%	42%	25%
Asia and Pacific	Count	1	16	7	6
Asia and Pacific	Percent	3%	53%	23%	20%

Does your NSO store and share data using secure cloud infrastructure? By secure cloud infrastructure, we are referring to private or government cloud environments.

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Europe	Count	3	6	11	8
Europe	Percent	11%	21%	39%	29%
Total	Count	5	28	26	20
Total	Percent	6%	35%	33%	25%

Human Resources

Are any employees within your NSO applying Big Data / data science techniques?

Country	Statistic	Don't know	No	Yes
Africa	Count	2	6	1
Africa	Percent	22%	67%	11%
Americas	Count	2	4	6

Are any employees within your NSO applying Big Data / data science techniques?

Country	Statistic	Don't know	No	Yes
Americas	Percent	17%	33%	50%
Asia and Pacific	Count	2	11	17
Asia and Pacific	Percent	7%	37%	57%
Europe	Count	4	2	22
Europe	Percent	14%	7%	79%
Total	Count	10	23	46
Total	Percent	13%	29%	58%

Does your NSO have a strategy to recruit new employees, from outside your NSO, who are able to apply Big Data / data science techniques?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Africa	Count	2	5	NA	2

Does your NSO have a strategy to recruit new employees, from outside your NSO, who are able to apply Big Data / data science techniques?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Africa	Percent	22%	56%	NA	22%
Americas	Count	1	5	4	2
Americas	Percent	8%	42%	33%	17%
Asia and Pacific	Count	2	13	4	11
Asia and Pacific	Percent	7%	43%	13%	37%
Europe	Count	2	10	5	11
Europe	Percent	7%	36%	18%	39%
Total	Count	7	33	13	26
Total	Percent	9%	42%	16%	33%

Does your NSO have a strategy to develop Big Data / data science skills of current employees?

Country	Statistic	Don’t know	No, but trying to establish	No, not considered	Yes
Africa	Count	NA	4	1	4
Africa	Percent	NA	44%	11%	44%
Americas	Count	NA	7	2	3
Americas	Percent	NA	58%	17%	25%
Asia and Pacific	Count	1	15	NA	14
Asia and Pacific	Percent	3%	50%	NA	47%
Europe	Count	1	15	1	11
Europe	Percent	4%	54%	4%	39%
Total	Count	2	41	4	32
Total	Percent	3%	52%	5%	41%

Has any Big Data / data science training been delivered within your NSO?

Country	Statistic	Don’t know	No, but trying to establish	No, not considered	Yes
Africa	Count	NA	3	4	2
Africa	Percent	NA	33%	44%	22%
Americas	Count	NA	4	1	7
Americas	Percent	NA	33%	8%	58%
Asia and Pacific	Count	1	11	3	15
Asia and Pacific	Percent	3%	37%	10%	50%
Europe	Count	1	10	4	13
Europe	Percent	4%	36%	14%	46%
Total	Count	2	28	12	37
Total	Percent	3%	35%	15%	47%

Does your NSO have partnerships with academia or with other national and/or international organisations that support and benefit your work with Big Data for Official Statistics?

Country	Statistic	Don't know	No, but trying to establish	No, not considered	Yes
Africa	Count	NA	2	3	4
Africa	Percent	NA	22%	33%	44%
Americas	Count	1	5	1	5
Americas	Percent	8%	42%	8%	42%
Asia and Pacific	Count	1	16	4	9
Asia and Pacific	Percent	3%	53%	13%	30%
Europe	Count	1	7	2	18
Europe	Percent	4%	25%	7%	64%
Total	Count	3	30	10	36
Total	Percent	4%	38%	13%	46%

Does your NSO have a Big Data / data science competency framework?

Country	Statistic	Don’t know	No, but trying to establish	No, not considered	Yes
Africa	Count	1	3	5	NA
Africa	Percent	11%	33%	56%	NA
Americas	Count	NA	8	3	1
Americas	Percent	NA	67%	25%	8%
Asia and Pacific	Count	2	17	9	2
Asia and Pacific	Percent	7%	57%	30%	7%
Europe	Count	2	13	10	3
Europe	Percent	7%	46%	36%	11%
Total	Count	5	41	27	6
Total	Percent	6%	52%	34%	8%

Does your NSO have a career pathway specifically for employees who are applying Big Data / data science techniques?

Country	Statistic	Don’t know	No, but trying to establish	No, not considered	Yes
Africa	Count	1	NA	8	NA
Africa	Percent	11%	NA	89%	NA
Americas	Count	1	4	6	1
Americas	Percent	8%	33%	50%	8%
Asia and Pacific	Count	3	15	8	4
Asia and Pacific	Percent	10%	50%	27%	13%
Europe	Count	1	5	21	1
Europe	Percent	4%	18%	75%	4%
Total	Count	6	24	43	6
Total	Percent	8%	30%	54%	8%