Break-out session 1

Session 1: Please discuss your national experience of transforming statistical production processes from data collection to dissemination in the context of the digital and technological revolution

What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

- Administrative Data
- i – Advantages
  - Dominica, Barbados & Montserrat – Readily available, get it from government departments that are more organised – trade, health and social security.
  - SVG – Reduces respondent burden
  - Jamaica- Dept of government can provide in electronic format – establish file protocols
  - Montserrat – Improves the relationship with agency
  - T&T – Cheaper than surveys and census

What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

- Administrative Data
- Disadvantages
  - Anguilla - Not fit for use by the statistical agency, was collected only for the purpose of the agency that collected it. (eg; classifications)
  - Dominica - Data is aggregated at too high a level – not detailed enough

What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

Geospatial

ii Advantages

- Anguilla – Most persons grasp concepts with pictures and it allow for easy understanding of the data being presented
- T&T – Easy to make link with socio-economic data

What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

Geospatial

ii Disadvantages

- T&T – Cost of software is expensive
- T&T – Requirement of technical persons

What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

Big Data

ii Advantages

- T&T – Real time
- T&T - Large coverage

What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

- Disadvantages
  - Data scientists required to mind the large volumes of data
  - Few standards that govern big data
  - Anguilla - Lack of classifications and approaches used in official statistics
  - T&T – Invite the invasion of privacy

Q1 What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

Advantages

i) Administrative data

- The scope and comprehensiveness of the statistical systems is strengthened.
- It reduces cost of data collection

ii) Big data

- There are opportunities for NSS to utilise big data for the production of national statistics.

iii) Geospatial information

- Geospatial data enhances the granularity process, which improve the usability and allows for analytical possibilities.

iv) Population and housing census

- The integration of the various data source allows for the possibility of shortened census questionnaires.

Q1 What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

Advantages (Cont'd)

iv) Population and housing census

- The census provides benchmarks for future surveys
- The census also provides detail information that can only be accessible during the census.

Disadvantages

- **Administrative data** – There are inherent challenges that affect quality of the data.
- **Big data -** The issues is on accessibility and capabilities to mine big data
- **Geospatial data -** The initial costs in the infrastructure and training of individuals to maximise use of geospatial data is high
- There would be need for a national spatial data infrastructure, since most statistical systems in member do that have this in place.

Q1 What could be the advantages and disadvantages of a broader integration of i.) administrative data, ii.) big data, iii.) geospatial information and iv.) population & housing census in the production of official statistics and indicators? Are there any issues in using these 4 categories of data sources? If yes, name the most important issues.

- **Disadvantage cont'd**
- **Population & housing census** – Using the example of labour force data, censuses provide limited ability to delve deeper into labour force related data, since it does not allow for detail labour force data.
- The population housing census is often too lengthy.

To what extent have innovative technologies been deployed in your respective production processes, starting at data collection to the dissemination of official statistics? Please explain.

- All – Use of CAPI – LFS, upcoming censuses, price collection
- Tools for dissemination – SDMX
- Eurotrace – Webcomex
- Radatum
- DataZoa
- Teleform – fit for certain Surveys
- PIPS
- CSPro
- Survey Solutions

Survey 1-2-3

Q2 To what extent have innovative technologies been deployed in your respective production processes, starting at data collection to the dissemination of official statistics? Please explain.

- Most of the survey are conducted using tablets - CAPI
- There has been work in the region to improve/modernised dissemination data through the use of website and social media.
- Use of GIS and integration of GIS in the production of official statistics.

Can you provide any relevant success stories, new initiatives or best practices regarding question 1 and 2 above? Examples of failed attempts may also be helpful.

- Admin. Data
  - SVG – using admin data for years
  - Some countries have successfully used the CAPI (Montserrat – Labour Force Census)

Q3 Can you provide any relevant success stories, new initiatives or best practices regarding question 1 and 2 above? Examples of failed attempts may also be helpful.

- Examples of failed attempts – The use of Windows tablet when using CS-Pro.
- The cost associated with replacing of tablets and software.
- There is need for security policies within the NSO to be considered within our system.

What additional support would you wish to receive in order to better grasp the digital revolution in the production of statistics and indicators?

- Training in working with big data
- Mont - GIS – For some offices training in GIS
- Jam - Consistent funding for capital funding
- Ang - Coordinated approach IDPs
- Dom - Support from government – financially

Q4 What additional support would you wish to receive in order to better grasp the digital revolution in the production of statistics and indicators?

- There is need for support to build the IT specifications and infrastructure to support the new business processes of an integrated national statistical system.
- There is need to develop capacity of data science, to be integrated in the work of NSOs; need for training in GIS.
- As it relates to big data, there is need for additional support in capacity building for NSO to collect, use and analyse big data.

Q5 What is missing from the discussion on this session that should be covered in the Handbook?

- South South Cooperation
- Organisational size and structure requires contextualisation
- Best practices must be given based on organisational size – eg data dissemination

Q5 What is missing from the discussion on this session that should be covered in the Handbook?

- Backup system, recovery plan or alternative options to support the transformation of statistical production processes from the data collection to dissemination in the context of the digital and technological revolution.

- Recommendations specific to SIDS context.