



UNITED NATIONS



Economic Commission for Latin America and the Caribbean
Subregional Headquarters for the Caribbean

Expert group meeting to examine the potential for integrating
big data within statistical data production in the Caribbean
24 August 2015
Port of Spain, Trinidad and Tobago

LIMITED
LC/CAR/L.475
22 September 2015
ORIGINAL: ENGLISH

**REPORT OF THE EXPERT GROUP MEETING TO EXAMINE THE
POTENTIAL FOR INTEGRATING BIG DATA WITHIN STATISTICAL
DATA PRODUCTION IN THE CARIBBEAN**

This report was reproduced without formal editing.

CONTENTS

| | |
|--|---|
| A. CONCLUSIONS AND RECOMMENDATIONS | 1 |
| B. ATTENDANCE AND ORGANIZATION OF WORK..... | 1 |
| 1. Place and date of the meeting | 1 |
| 2. Attendance | 1 |
| 3. Agenda | 1 |
| C. SUMMARY OF PROCEEDINGS | 2 |
| 1. Opening of meeting..... | 2 |
| 2. Objectives of the meeting..... | 2 |
| 3. Policy and implementation issues pertaining to big data and its use for official statistics in the Caribbean | 2 |
| 4. Open discussion on the way forward: crafting a big data strategy for the Caribbean..... | 6 |
| 5. Closing remarks | 7 |
| Annex I List of participants..... | 8 |

A. CONCLUSIONS AND RECOMMENDATIONS

1. The Caribbean subregion should adopt a contextually suitable definition of big data.
2. Big data solutions must be suited to the context of the Caribbean subregion. Therefore, these solutions cannot simply be imported from other parts of the world.
3. A regional body should be established to serve as a repository of Caribbean big data.
4. A centre of excellence in big data analytics should be established to promote regional expertise in the methodologies and tools of big data analytics.
5. A survey of big data producers in the Caribbean should be carried out in order to have a better understanding of their perceptions of big data and a more robust idea of the purposes for which data producers will use big data.
6. A sustainable financing strategy that would ensure the long term viability of big data initiatives is crucial. The Caribbean subregion should therefore develop a case for financing from donor agencies including the United Nations, World Bank, and Caribbean Development Bank (CDB).
7. The subregion should also consider public-private partnerships and other means for long term financing of big data initiatives.

B. ATTENDANCE AND ORGANIZATION OF WORK

1. Place and date of the meeting

8. The expert group meeting to examine the potential for integrating big data within statistical data production in the Caribbean was held on 24 August 2015 at the Economic Commission for Latin America and the Caribbean (ECLAC) subregional headquarters for the Caribbean, Port of Spain, Trinidad and Tobago.

2. Attendance

9. The meeting was attended by representatives of government entities, specifically the Central Statistical Office of Grenada, the Statistics Department of Saint Lucia, the Statistical Institute of Jamaica (STATIN), and the Telecommunications Services of Trinidad and Tobago (TSTT). Representatives of international organizations namely, the CDB and the Organization of Eastern Caribbean States (OECS) were also present. The ECLAC consultant and various ECLAC staff members were also in attendance.

3. Agenda

1. Opening of the meeting.
2. Objectives of the meeting.
3. Presentation 1: Why the big data push?
4. Presentation 2: Big data: challenges and opportunities.

5. Presentation 3: Integration of big data into official statistics in the Caribbean: realistic or optimistic?
6. Open discussion on the way forward: crafting a big data strategy for the Caribbean.
7. Conclusions and recommendations.
8. Closing remarks.

C. SUMMARY OF PROCEEDINGS

1. Opening of meeting

10. Welcome remarks were made by the Coordinator of the Statistics and Social Development Unit of ECLAC subregional headquarters for the Caribbean on behalf of the Director of ECLAC subregional headquarters for the Caribbean. The opening remarks conveyed the importance of exploring big data in the Caribbean subregion considering the opportunities and challenges that it presents.

11. He noted that the experience in the monitoring of the Millennium Development Goals (MDGs) revealed deficiencies in the capacity of countries to promptly report on the progress in the attainment of goals. This shortcoming, he observed, was acknowledged in the development process of the sustainable development goals (SDGs) and has resulted in a call for the Data Revolution for Sustainable Development. This data revolution will require harnessing new data sources and modern technology to advance the post-2015 development agenda. He emphasized that big data is an integral element of the data revolution, therefore, the Caribbean should not be left behind in exploring big data.

12. In contributing to the United Nations initiatives on big data, the Coordinator indicated that ECLAC has commissioned the study on which the expert group meeting was being convened to examine the state-of-affairs with big data in the Caribbean subregion.

2. Objectives of the meeting

13. The Coordinator of the Statistics and Social Development Unit of ECLAC stated that the purpose of the meeting was to explore the issues surrounding big data in order to devise strategies on the way forward in incorporating big data in official statistics in the Caribbean subregion. He also stated that the meeting would highlight the potential benefits and challenges of big data and explore some solutions to big data obstacles in the subregion.

3. Policy and implementation issues pertaining to big data and its use for official statistics in the Caribbean

Data revolution: why the big data push?

14. The ECLAC consultant made a presentation on “Data revolution: why the big data push?” She described big data as data with high volume (large size), generated with high velocity, in great variety and veracity. Her presentation highlighted the importance of big data for the Caribbean subregion. She stated that many of the world’s poorer nations lagged behind in reporting on the MDG indicators and highlighted that the reporting performance of the subregion has worsened. She suggested that big data would assist in addressing these data inadequacies.

15. She further stated the importance of data monitoring and reporting in the move from the MDGs to the SDGs. She suggested that traditional data management tools are inadequate for big data analysis. The

data revolution therefore represents an opportunity to expand on the variety of data that are being collected and reported. Since many projects implemented globally have shown that big data can be incorporated in official statistics, the ECLAC consultant suggested that national statistical offices should be an essential part of big data in the Caribbean subregion. She indicated that big data has many roles, one of which should be to supplement official statistics.

Discussion

The Caribbean context

The EGM participants noted the following:

16. While there are examples of countries using big data from other parts of the world, it is important that the Caribbean subregion creates strategies that are relevant to its context. The subregion should also closely examine what it stands to gain from big data.

17. Regional organizations such as ECLAC can act as a focal point for big data in the Caribbean. The big data focal point would keep a record of big data projects in the subregion and facilitate knowledge-sharing in big data best practices among Member States.

Role of big data

The EGM participants agreed on the following:

18. Big data, with its many opportunities and benefits, will shape how data are collected and processed in the future.

19. There are a few examples of how big data has been applied in the Caribbean subregion, particularly in post-disaster response and public health surveillance.

Concerns

The EGM participants raised the following concerns:

20. Neither the infrastructure nor the expertise necessary for a successful big data initiative exists in the national statistical offices.

21. Privacy and other policy issues arising from big data are a source of concern.

22. Limited access to big data and lack of influence on big data producers may constitute obstacles to big data projects.

23. Lack of diversity in the sources of big data is a potential problem as telephone and text messages may be replaced, in the near future, by less easily tracked applications like Skype and Whatsapp.

Open data

The EGM participants noted the following:

24. There is interest in open data. Open data represents data that can be freely accessed, used and shared and open data does not have to be big data. Public-private partnerships (PPP) would be necessary if big data projects are to be implemented. It is also important to be clear and as specific as possible in the scope of big data that would be considered in the Caribbean.

Big data: challenges and opportunities

25. The Coordinator of the Statistics and Social Development Unit delivered a presentation on “Big data: challenges and opportunities”. He stated that technology is dynamic and that the Caribbean subregion needed to be on board and be able to manage these changes. He suggested that the countries of the subregion work together in order to better position the subregion to work with big data-producing companies like Google. This is particularly critical since there is a large and generally growing number of people in the subregion who use mobile phones and have access to the Internet. It was noted, however, that certain portions of society have limited mobile phone and Internet access including the young, the elderly, and other vulnerable groups.

26. He disclosed that because of the widespread access to the Internet, big data is cheaper than traditional surveys. He also suggested that by using big data as an official source of data, the timeliness of data could be improved. Big data has the potential to supplement, benchmark, and provide early estimates of official statistics.

27. Despite the many positive opportunities that big data presents, a major challenge in using big data is the concern regarding likely breach of privacy. Specifically, a government’s access to its citizen’s private data constitutes a privacy concern. National statistical offices also need to consider the challenges which may be encountered in efforts to acquire the necessary big data, technology, expertise, and funding to support the initial stages of a big data initiative.

Discussion

Country level data

ECLAC made the following suggestions:

28. Meeting participants were directed to consider annex III of the Big Data Study in which big data projects that have been undertaken largely at the country level are documented. There are important lessons to be learned from countries like China and Italy, as well as organizations like the Economic and Social Commission for Asia and the Pacific (ESCAP). However, the Caribbean subregion must keep in mind the differences in context and culture.

29. Further, rather than work at the country level, the countries of the Caribbean subregion should consider collaborating with one another and with the big data providers. A regional body could be established in the Caribbean subregion to acquire big data, which could later be disaggregated by country.

Opportunities

ECLAC made the following suggestions:

30. The use of big data can present valuable opportunities for the Caribbean subregion including timeliness and cost savings.

31. Big data can also be useful in areas such as health, specifically in managing response to epidemic. Big data has shown to provide reliable information ahead of official statistics from government sources. The example of Chikungunya tracking in Trinidad and Tobago was cited as a success case in the use of big data.

32. Governments should take advantage of the opportunity to raise awareness and highlight the benefits of big data and how it can be efficiently used to benefit the population. In order to realise the benefits and efficiencies, however, checks and balances must be put in place to safeguard privacy and prevent unauthorized access. Benchmarking of statistics derived from big data to those from traditional sources is also critical.

Concerns

The EGM participants observed the following:

33. Big data is not representative of the population and all sectors would not benefit equally from big data.

34. Access to private data: The private sector may be wary of competition and therefore unwilling to share big data that they generate, for example scanner data necessary for some big data applications. This has been the experience in Saint Lucia. Similarly, in Jamaica, mobile providers have been unwilling to share big data. It may therefore be necessary for governments to implement policies to encourage data sharing by private entities.

35. The issues of privacy and trust are important ones. Persons do not intrinsically trust government with their private information and may be uncomfortable with the idea of government having access to their personal data.

36. Cost of big data: Big data frameworks require financing. It is still unclear how the big data initiative in the Caribbean subregion will be financed. Governments will have to decide whether to buy data from big data producers or introduce legislation that ensures that big data are made available for authorized use.

Technical expertise

The EGM participants noted the following:

37. There are several technical aspects that should be taken into account with the big data initiative. For example, there is potential for errors if the sample reference frame is not properly designed; the sources of data must be proven and assured to be reliable and valid to ensure the integrity of the data; and acceptable margins of error must be established. All these require technical expertise in big data analytics which is lacking or limited. The need for technical expertise becomes more important where data acquired through social media are to be used to generate official statistics.

Integration of big data into official statistics in the Caribbean: realistic or optimistic?

38. The Coordinator of the Statistics and Social Development Unit of ECLAC made a presentation on the topic: “Integration of big data into official statistics in the Caribbean: realistic or optimistic?” He summarized the results of a survey of national statistical offices in the Caribbean subregion conducted by ECLAC. The purpose of the survey was to gain a better understanding of the national statistical offices’ perceptions of big data.

39. The Coordinator pointed to the survey’s 35 per cent response rate as an issue of major concern. In addition, he noted that the term “big data” was not well understood by some respondents. The survey revealed dearth of big data projects in the subregion with none of the national statistical offices reporting being aware of any big data projects undertaken in the Caribbean subregion. Moreover, while all the national statistical offices thought that big data could be used to supplement official statistics, the majority did not see a role for big data in the execution of their work.

40. As part of a big data strategy for the Caribbean subregion, the Coordinator suggested the creation of a centre of excellence for big data analytics and a Caribbean big data clearing house to provide public goods for the Caribbean subregion in the area of big data. The centre of excellence could develop expertise in processing big data and provide training to statisticians from the national statistical offices and other institutions in big data analytics. The clearing house would serve as an independent repository of Caribbean big data. The Coordinator elaborated that an academic institution would perhaps be better positioned to serve these purposes.

41. He suggested that public-private collaboration would be useful in addressing the issue of funding. For example, data providers, such as telecommunication companies, could contribute by sharing their data. Additionally, major regional and international bodies including the United Nations and the Eastern Caribbean Central Bank (ECCB), could be approached to provide funding.

Discussion

Costs and funding

The EGM participants agreed to the following:

42. A sustainable funding source is crucial if the subregion is to establish a clearing house for big data. Donor support may prove helpful in the initial stages, however, sustainable and long term funding must be found. Partnership with the geographic information system (GIS) community with regards to data sharing, among others, may be helpful. National statistical offices are increasingly using GIS data and can also contribute basic community maps through open data portals.

43. Regional partnerships should be encouraged as should local partnerships between the public and private sectors. The inclusion of the private sector should be taken into account when funding is considered.

44. Because of the costs associated with setting up a big data project, the participants suggested that it would be prudent to first consider issues such as: Is there demand for big data? Who needs it? For what purposes? What is the frequency of this need? Are end-users willing to pay for the big data and how much?

Defining big data

ECLAC made the following observation:

45. A strict and precise definition of the term “big data” may not be necessary since big data may be conceptualised as an extension of the work that national statistical offices already undertake. Specifically, big data may be viewed as another form of administrative data arising from the advances in technology and the Internet.

Big data, open data, geographic information system, and public-private partnerships (PPPs)

The EGM participants noted the following:

46. The issues of big data, GIS data, and open data should converge. The OECS seems to be ahead in this regard as in addition to engaging in talks on big data, the OECS has secured funding for GIS data from CDB and is in talks on open data with the World Bank.

47. The CDB sees the potential to support the development of legislative frameworks as part of the PPP framework. The organisation is willing to explore more into exactly how to integrate big data in its work.

4. Open discussion on the way forward: crafting a big data strategy for the Caribbean

Regional centre of excellence

48. ECLAC reiterated the need for the creation of a centre of excellence dedicated to big data in the subregion. One of the many responsibilities of this centre would be to train staff of national statistical offices. However, to more accurately determine the need and scope of the proposed centre of excellence, the Caribbean subregion would have to first determine the need for big data. To this end, ECLAC encouraged the national statistical offices to complete the ECLAC questionnaire on big data earlier sent to

them. ECLAC also indicated that a survey targeting big data producers would be commissioned in order to gauge support for big data projects, to determine whether big data producers have a strategic plan for the data that they generate, and to seek their opinion on the proposed centre of excellence.

Financing

49. The meeting acknowledged that it is crucial to have a sustainable financing strategy that would ensure the long term work on big data. The meeting agreed that an important next step, therefore, would be to develop a refined and enhanced justification or case for financing, which would address privacy issues and other sensitive matters, to deliver to donor agencies including the United Nations, World Bank, and CDB showcasing the needs of the subregion.

Assessments

50. The meeting agreed that legislative constraints in the Caribbean subregion should be assessed. The sectors that would benefit most from big data should also be assessed.

51. The meeting recommended that the subregion should make determinations as to what kind of data are included in big data, what it wishes to get from these datasets, and the kinds of strategy to employ in order to attain these data.

National statistical offices

52. ECLAC affirmed that its objective is to help national statistical offices use big data to supplement national statistics which could assist in monitoring the SDGs. ECLAC, therefore, recommended that big data be adequately investigated to ascertain how it can be effective in helping to monitor and report on the SDGs within the post-2015 development agenda.

53. The meeting recommended that Member States explore the technical skills of their human resources and make efforts to improve upon their technical expertise and their data science skills. To this end, it may be useful for statistical officers to be familiar with the statistical package R.

5. Closing remarks

54. Closing remarks were delivered by the Coordinator of the Statistics and Social Development Unit on behalf of the Director of the ECLAC subregional headquarters for the Caribbean. He expressed appreciation for the contributions made and thanked everyone for attending and participating in the meeting, adding that he looks forward to their continued participation in and invaluable support for the big data initiative.

Annex I**LIST OF PARTICIPANTS**

Marsha Caddle, Operations Officer, Technical Cooperation Division, Caribbean Development Bank (CDB), Barbados. E-mail: caddlem@caribank.org

Carol Coy, Director-General, Statistical Institute of Jamaica (STATIN), Jamaica. E-mail: ccoy@statinja.gov.jm

Tiemonne Charles, Statistician/Software Engineer, Central Statistical Office, Ministry of Finance, Planning, Economy, Energy, and Cooperatives, Grenada. E-mail: tiemonne@gmail.com

Alecia Evans, Consultant, Economic Commission for Latin America and the Caribbean (ECLAC). E-mail: evans.alecia@gmail.com

Sean Mathurin, Economist/Statistician, Organisation of Eastern Caribbean States (OECS), Saint Lucia. E-mail: smathurin@oecs.org

Marlon Morris, Chief Performance Officer, Telecommunications Services of Trinidad and Tobago. E-mail: mmorris@tstt.co.tt

Edwin St. Catherine, Director, Statistics Department, Ministry of Economic Affairs and Physical Development, Saint Lucia. E-mail: edwins@stats.gov.lc

**Economic Commission for Latin America and the Caribbean
Subregional headquarters for the Caribbean**

Abdullahi Abdulkadri, Coordinator, Statistics and Social Development Unit. E-mail: Abdullahi.abdulkadri@eclac.org

Francis Jones, Populations Affairs Officer, Statistics and Social Development Unit. E-mail: francis.jones@eclac.org

Robert Williams, Associate Information Management Officer, Caribbean Knowledge Management Centre. E-mail: robert.williams@eclac.org

Tanisha Ash, Statistical Assistant, Statistics and Social Development Unit. E-mail: tanisha.ash@eclac.org

Candice Gonzales, Research Assistant, Statistics and Social Development Unit. E-mail: candice.gonzales@eclac.org