

Covering Note

The Challenge of Maximizing the Use of Microdata in PICTs

Prepared for the 10th Meeting of the PSMB,
Auckland (NZ) 17-18 October 2022

Ed. by Andrea Borlizzi and Michael Sharp

Contents

1. The Pacific Regional Data Dissemination Strategy (PRDDS)	3
1.1 The PRDDS in a nutshell	3
1.2 Background	4
1.3 The implementation of the strategy	4
2. The PACSTAT project consultancy	5
3. The report on microdata use and dissemination: key findings	6

1. The Pacific Regional Data Dissemination Strategy (PRDDS) ¹

1.1 The PRDDS in a nutshell

The 5th Regional Conference of Heads of Planning and Statistics (HOPS) held in November 2017, endorsed a proposal for a Regional Data Dissemination Framework. This original framework has then been developed into a full Pacific Regional Data Dissemination Strategy (PRDDS). The PRDDS was endorsed by the Pacific Statistics Standing Committee (PSSC) in November 2019.

The PRDDS comprises four components:

- governance around Pacific data acquisition, archiving and dissemination;
- a national data dissemination strategy framework;
- a regional data dissemination platform; and
- a coordinated support program with development partners for Pacific data dissemination.

The PRDDS acknowledges that there are two overriding considerations that need to be taken into account in relation to a strategy for data dissemination:

- the ownership of statistical data and associated metadata always resides with the national producers, in this case mostly the national statistics offices (NSOs); and
- the confidentiality of an individual respondent's information provided during a census, survey or other collection, must be protected, in line with requirements set out in each member country's Statistics Act or other relevant legislation.

The four components of the regional data dissemination strategy therefore address key issues for both data producers and data users as follows:

- the requirement of data producers for data protection and appropriate data governance, particularly for microdata access;
- the need for improved data dissemination strategies for NSOs to better meet the statistical needs of national users of data, notably government policy-makers and planners, as well as the private sector, students, the media and the general public; and
- to meet the statistical needs of regional and international users through a more coordinated effort among regional and international agencies such as SPC, Universities, UN agencies and their regional offices such as ESCAP, World Bank and IMF PFTAC, and others.

Capacity building is essential if the regional strategy is to be successful; the capacity building should be focused on:

- lifting statistical standards,
- improving quality and timeliness of statistical outputs,
- improving metadata and associated documentation,
- harmonising classifications, and
- developing a "first class" data dissemination service.

For the PRDDS to be successfully implemented, additional resources are required at both the national level and in the Statistics for Development Division (SDD) of SPC. In furtherance of the PRDDS,

¹ This section is mostly based on the Pacific Regional Data Dissemination Strategy (PRDDS), available [here](#).

and in line with its Business Plan, SPC-SDD will, as far as resources allow, continue to channel statistical expertise into key components of the strategy.

1.2 Background

Improving access to Pacific data is essential to achieving the Vision for Pacific statistics, as set out in SPC-SDD's 2018-2020 Business Plan:

Trusted Pacific data supporting well informed policy decisions that help improve Pacific people's lives.

This Vision requires the widespread availability, accessibility and appropriate use of the Pacific data.

Regional data dissemination includes both the Pacific Island Countries and Territories (PICTs) producing and disseminating (either in-house) an agreed core set of high-quality economic, social and environmental statistics in a timely and user-friendly manner and in line with national priorities and the dissemination of data about the Pacific by both PICTs and the regional and international organisations to meet other needs and commitments.

1.3 The implementation of the strategy

For each country/NSO the following actions will be required for the successful implementation of the Regional Dissemination Strategy:

- Action 1: Agree Memorandum(s) of Understanding with SPC, covering data supply to SPC-SDD databases, SPC-SDD support for statistical activities, and Data Licence Agreements as required for specific micro-datasets.
- Action 2: Develop user profiles and implement strategies for national data dissemination, building on their NSDS.
- Action 3: Upload microdata in agreed formats to dissemination platforms.
- Action 4: Work with SPC to ensure websites are updated and maintained as new data becomes available.

Regional agencies will also be required to:

- Develop agreements with SPC-SDD on data sharing and coordination of statistical capacity building.
- Ensure that all plans for statistical activities are communicated to SPC, through the DDPG and PSSC, for coordination and on-communication to NSOs.
- Adhere to the dissemination protocols for each national dataset to be accessed.

2. The PACSTAT project consultancy

Within the context of the Pacific regional data dissemination strategy, an expression of interest for a “Microdata dissemination advisor” consultancy was open by SPC in 2021. The overall objective of the assignment was to provide information to the Pacific Statistics Methods Board (PSMB) and other Pacific statistics decision makers to guide decision making in relation to dissemination and use of microdata. The consultancy falls within the scope of the [PACSTAT](#) Project, an International Development Association ([IDA](#)) – funded project with the objective to *improve the quality of welfare data collection and its accessibility*. In fact, public access to data is a challenge still facing the region, and data deprivation hinders efforts to track and monitor progress towards relevant indicators and to implement consistent policies.

National Statistics Offices (NSOs) in the Pacific region lack funding to meet the costs of data collection, and the small populations often lack capacity in terms of budget, labour qualification, and other technical constraints associated with technology and statistics across the vast Pacific area. Statistics development in the Pacific, nonetheless, has been making steady progress over time and this project builds on existing developments. For example, The Pacific Statistics Governance Framework is in place, the Ten-Year Pacific Statistics Strategy (TYPSS) review is underway, there is good local capacity in collecting data and implementing face-to-face surveys and, in most cases, a national statistics development strategy is in place.

Numerous nationally representative census and surveys (<https://sdd.spc.int/census-and-survey-calendar>) have been conducted in the Pacific region; data, however, remain underutilized. In many cases, upon completion of the census/survey report, these datasets are not made available for further use. This is a missed opportunity for the economic and social development of the Pacific region and hinders evidence-based policy derivation and monitoring.

There are numerous factors as to why data are not made available for secondary analysis, which range from unclear mandate of statistics decision making in statistics legislation in relation to microdata dissemination, to data sovereignty, capacity limitations, and risk of individual disclosure. This assignment aims to identify the pros and cons of microdata dissemination, and other related matters, such as the limitations of current statistics legislation.

3. The report on microdata use and dissemination: key findings

The PICTs are lagging behind

1. According to international standards and openness rankings such as ODIN, PICTs are generally lagging behind other comparable developing countries globally. This significant underutilization of the PICTs microdata is reflected in their NSOs' response to the survey questionnaire: they would like to see more analytical reports resulting from their HIES microdata sets.

The Pacific Data Hub Microdata Library

2. The Pacific Data Hub Microdata Library (PDH-ML) is a valuable platform for resource-scarce PICTs, because it not only helps preserve data for posterity², but also facilitates a greater level of access to the microdata for global, regional, national and independent researchers.
3. The performance of the PDH-ML Repository may be assessed by:
 - (a) the percentage of all generated micro datasets actually stored at the PDH-ML, except for years prior to the availability of computer data storage (e.g.: the 1980's);
 - (b) The historically small number of requests for usage of datasets; and
 - (c) the low percentage of approvals of the requests (35%).Such Key Performance Indicators (KPI) ought to be the subject of Internal review by NSOs and PICT governments as to why permissions were not granted.
4. The PSMB should note that the rate of approval for microdata access is low not only for independent researchers and regional academic institutions, but also for multilateral organizations such as the World Bank and FAO, whose *bona fide* would not be in doubt. The PSMB might wish to initiate an inquiry with the PDH-ML and PICTs' NSOs on the process for granting access to the microdata sets and how the approval rate can be improved.

The microdata collections in general: prioritizing surveys

5. The PDH-ML can only store microdata sets if they have been generated by the PICTs. In recent years, many microdata sets have been generated as a result of a new focus on the Pacific and particular PICTs by international organizations with their relatively abundant resources.
6. It is recommended that PSMB place on its agenda a discussion about PICTs' priorities (for example, in the areas of economics, business, labour force and job creation, health, climate change, depending on the PICT) for the conduct of regular national surveys and censuses. Such programs ought to also guide the international donors and multilateral organizations in funding microdata collection exercises.

² Natural disasters such as floods and storms are frequent in the Pacific and can cause the permanent loss of data repositories.

The storing of the microdata sets at PDH-ML

7. Given the low percentage³ of existing microdata sets actually deposited at the PDH-ML, the PSMB should consider the implementation of initiatives to encourage PICTs to deposit the missing microdata sets - including those of HIES conducted over the years in their areas of priority⁴ - as well as the related publications, questionnaires and manuals.
8. There is a potential for generating evidence-based policy analyses by independent academic researchers analyzing a wide variety of HIES and EUS microdata sets. Such initiatives can be facilitated by a confluence of funds from government and donors and quality academic researchers willing to work in the designated areas and producing the policy oriented reports.

Blockers for greater microdata use

9. Different PICTs have different dynamics at work to explain the relative under-utilization of their microdata sets and their lack of approvals to requests. Some of them mentioned ambiguous legislation⁵ not helping their decision-making; some others saw the data requests as not backed by reputable institutions; some thought that the research was of no use to their country and people; risks of compromising confidentiality were also articulated⁶.
10. The NSO responses to the survey questionnaire suggest that:
 - (a) Most of the responding PICTs have ambiguous legislation about their ability to make their microdata available to external researchers, and some even to other Government departments;
 - (b) Most of them have a scarcity of technical and analytical staff;
 - (f) most of them would like to receive assistance on data anonymization and data analysis.

Completing the virtuous cycle: holding workshops

11. NSOs (and partners) often consider the job done when the Reports are produced. However, the PSMB should consider encouraging the PICT NSOs to complete the "virtuous cycle" whereby:
 - (a) funds are made available for national workshops and policy discussions among the relevant government ministries, NGOs and donor stakeholders using the survey results.
 - (b) an adequate media discussion of the survey results is ensured, to improve the goodwill of the public who responded to the NSO household surveys, often revealing sensitive information about their households' to NSO interviewers. This goodwill cannot but make it easier for the NSOs to implement any subsequent survey.

³ NB: low values are sometimes due to old microdata built prior to the availability of computer data storage.

⁴ An initiative going in this direction is the new capacity development programme being initiated by SDD.

⁵ A number of PICTs simply cannot put record files into repositories located abroad because their legislation does not allow for this. The issue of outdated legislation is a very important one; advocating for changes in legislation and getting changes implemented, however, is a time-consuming process which takes place in the long-run.

⁶ Another important issue is the sovereignty over microdata and the research outcomes obtained from them. The "5-safes" approach should incorporate a series of protocols to explicitly respect the subject of the data. In this view, one of the 5 implementation criteria of the "Theory of change" included in TYPSS II is: "Working in culturally, contextually responsive and people-centred ways". The Pacific might like to consider a version of the "5-safe" approach as a basis for incorporation into DLAs that will protect culturally sensitive issues.

12. The public perception of the benefits of this “virtuous cycle” also helps build NSOs’ and Governments’ trust in the PDH-ML as a central Repository, which is of great benefit to PICT NSOs and PICT taxpayers who ultimately pay for the household surveys that lead to the generation of the microdata sets.
13. Donors and development partners could also play a greater role in promoting openness by requiring that any microdata generated with external funding be archived with PDH-ML and that Data License Agreements (DLAs) be a requirement for obtaining financial support.

Critical pathways forward: The 4 Tier Structure

14. The PSMB can assist the PICTs to be in control of the pace and *modus operandi* with which their datasets are made available to the public through a tiered approach:

- | | |
|---------|---|
| Tier 1: | The secure archiving of all microdata with the PDH-ML, with no necessary access to the public (unless agreed to). The target Key performance Indicator (KPI) for this would be: 100% of all microdata historically created is archived. |
| Tier 2: | Microdata are made available by PICTs for a short-term Special Project managed by SDD, for analysis and writing of reports on topics prioritized by the NSOs, such as, but not limited to, poverty, food security, climate change and gender inequalities. For this exercise, SDD can create a call for expressions of interest from multilateral organizations, universities (international and regional) and research organizations, and recognized expert independent researchers to collaborate with the specific NSOs and their designated civil servants for the analysis of the specific microdata and the completion of the related Reports.
Following every publication of such a Report, the SDD/NSO can facilitate a national workshop bringing together all the national and global stakeholders in the policy issues discussed. Eventually, an over-arching regional conference/workshop can be organized, to gather these Pacific-wide findings and publish them in a monograph. |
| Tier 3 | Microdata sets which PICT NSOs make available <i>upon application</i> by external researchers, with the PDH-ML coordinating the process of approval by the PICTs, who have sovereignty over that particular data set. |
| Tier 4: | Microdata sets that PICT NSOs make freely available for download, with strict conditions being met through binding terms of use. |

Transparency of Approvals/Rejection Process and Reporting

15. A formal reporting process whereby the PDH-ML informs every year PICT NSOs of all Requests, Approvals and Rejections (with clear explanations for rejections) should be put in place.
16. The PDH-ML should also post on its website all Reports and publications that have resulted from access to the microdata sets in the PDH-ML, which should be freely downloadable, unless specific restrictions have been put in place by the originating authority and agreed to by the PDH-ML and the concerned PICT.
17. The PDH-ML should then inform over time further changes in the structure of the Tiers and the conditions under which their datasets may be made available to researchers.