# Discussion Software Use Survey

## Data Collection

▪ In late 2020, PAPI was still used in many (9) PICTs as the main method for data collection. While this traditional approach to data collection was more common in the region at the time of this survey, the situation has now since changed with most[1] Pacific Islands having transformed their data collection methods and processes embracing new technology and committing to a tablet-assisted questionnaire using a computer-assisted personal interview approach (CAPI). What has driven this CAPI adoption?

▪ The digital approach to data collection was recognised by **all** PICT NSO responses as a preference for future statistical development and believe SPC should provide support in CAPI, in particular the majority of NSOs prefer CAPI support in Survey Solutions. How have efficiencies been made in the use of CAPI over traditional approaches to data collection?

## Data Processing

▪ Many of the software used for data processing reflect technical support provision. For Census and Survey SPC is processing data using Stata while UNICEF is processing using CSPro[2] and SPSS for MICS. PAPI census and surveys have traditionally used CSPro for data entry, processing, and tabulation. Stata is the overall preferred data processing software in the region.

▪ There is still widespread use of Excel and Access which are less appropriate for data processing and editing applications. How do capacity constraints with acquisition/adoption of software (human and financial implications) and the provision of technical support involved in data processing approaches between both NSOs and Technical Assistance providers affect selection of Data Processing software?

   For example:

1. World Bank no longer hosts Survey Solutions servers, so SPC has adopted this role.
2. There is a need for more transparent and replicable processing/editing procedures.
3. Data sharing constraints exist with different file formats and proprietary software.
4. Considerable investment has been made into capacity building initiatives in the region.
5. Development of the Pacific Islands Population and Housing Census Editing Handbook

## Data Analysis

▪ There is a wide range of preferred main software and tools for data analyses by Pacific NSOs ranging from Stata, CSPro, Excel, R and SPSS. Why are PICTs more likely to use more than one software in undertaking data analysis and not focus on use of one tool?

▪ Reflecting that choosing the right data analytics tool is challenging, as no tool fits every need. Given there is still widespread use of Excel in the region is there an appetite for more specialised statistical programs such

---

[1] Cook Islands, Kiribati, Marshall Islands, Samoa, Vanuatu, PNG, Tokelau, Niue, FSM, Nauru now use CAPI for their major data collection activities such as Census and Survey

[2] CSPro, Census and Survey Processing System, is a public domain data processing software package developed by the U.S. Census Bureau

Pacific Community (SPC) Headquarters: Noumea, New Caledonia. Regional Offices: Suva, Fiji;
Pohnpei, Federated States of Micronesia; Port Vila, Vanuatu.

www.spc.int   spc@spc.int

Siège de la Communauté du Pacifique (CPS) : Nouméa (Nouvelle-Calédonie). Antennes régionales : Suva (Fidji) ;
Pohnpei (États fédérés de Micronésie) ; Port-Vila (Vanuatu).

www.spc.int   spc@spc.int

as R, Stata and SPSS? If so, what obstacles are there for adoption of some of these software tools for data analysis?  What are some of the data sharing constraints in the region?

- Are the speed of analytical tasks and scripting abilities of such data analysis software being considered in Data Analysis software selection? and what are PICT preferences towards open-source software tools such as GitHub particularly the ability to generate reproducibility of code and replicable analysis to help provide transparency?

**Data Visualisation**

- Why are Excel and Powerpoint the main data visualisation software being used by PICTs when there a many other open-source options available and the PICT preferred software options for data visualisation include R, GIS mapping and specialised data visualisation software?
- When selecting visualisation options how important is it that software is compatible with HTML and CSS which is web friendly?
- Do PICTs consider hard-copy publications still important in data visualisation or less important?