



DATA AVAILABILITY
AND QUALITY REPORT
ARTECHE

SEPTEMBER 2022

DEVELOPMENT DATA LAB



Local governments are at the forefront of redefining public service delivery, but people should be at the core of every innovation. Over the years, it has become widely accepted that local government units (LGUs) should not only strive to improve the type and quality of government services but also transform the culture of government to make it more inclusive, transparent, and accountable. Generally, there are two ways by which LGUs can satisfy this objective:

1. Having a deep understanding of the needs of the citizens and stakeholders as direct beneficiaries of public sector policies and services and

2. Cultivating an active public participation in designing, delivering, and monitoring policies and services that affect their lives.

Central to the approach is the intensive and extensive use of data along with the adoption of appropriate technological tools. On the one hand, LGUs should be able to effectively leverage data to inform strategies in providing services that are attuned to people's needs and well-being. On the other hand, people should be aware of the availability of data and are able to actually use data to influence how they should be governed and yield benefits to their communities. In other words, effective local governance is one that employs a data-informed participatory decision-making process. This

type of governance is more likely beneficial to human development.

It is through the appropriate use of data that duty-bearers can fully grasp the realities on the ground, strategically respond to public concerns, and allocate resources efficiently and effectively. Unfortunately, establishing a culture that values data-driven decision-making and a good uptake of open data by citizens remain elusive, in many local governments in the country. The Data-Driven Development in the Philippines (PH3D) Project helps all its project sites strengthen the culture of data in the local community and embrace innovations towards a people-centered local governance. The Municipality of Arteche is one of them. A preliminary step in this intervention is to make sure the data are available and of good quality. This report provides an assessment of data collected and held by the local government of Arteche, in terms of availability and quality, to effectively ensure that the LGU can make competent decisions for the people and draw policy lessons for good governance. Towards the end of this report some recommendations and solutions are suggested.

DATA AVAILABILITY (MANDATED VS. HELD) & QUALITY ASSESSMENT

National legislations on the collection of local data have increased notably in recent periods (see Appendix 1). In particular, various laws mandate the collection of local data in the following sectors: Disaster Risk Reduction and Management (DRRM), Health, Environment, Labor, Agriculture, Social Services, Business, Education, Human Rights, Gender, Urban Development, and others. Although not exhaustive, Table 1 provides the important sets of data required for collection in the LGU that are deemed sufficient to aid in crafting policies to promote local development. The data are clustered into five sets:

Community-Based Monitoring System (CBMS), Health, Business, Geospatial including Risk Management, and Human Rights Violations Case Monitor. As shown in Table 1, the LGU of Arteché holds data that can potentially cover the sectors of DRRM, Health, Labor, Agriculture, Business, and Social Services. However, no data on monitoring human rights violation cases is available, yet. Table 1 also provides the conditions of the collected data particularly format, disaggregation, and frequency of data collection.



**TABLE
1**

AVAILABILITY OF LOCAL DATA IN ARTECHE

MANDATED DATA	PURPOSE	LGU COLLECTED DATA	YEAR
CBMS HOUSEHOLD SURVEY DATA	Derive the multidimensional poverty index and indicators on: <ul style="list-style-type: none"> • <i>Health, Nutrition, Housing, Water and sanitation, Education, Income and Employment</i> 	CBMS	2019
HEALTH DATA	Gather relevant data on: <ul style="list-style-type: none"> • <i>PWDs</i> • <i>Responsible reproductive health and parenthood</i> • <i>Health facilities and skilled health professionals</i> • <i>Mental health</i> • <i>Universal Health Care</i> • <i>Cancer Cases</i> • <i>Notifiable diseases</i> • <i>Health events of public concern</i> 	Various Health-related Data (derived from the Comprehensive Land Use Plan) <ul style="list-style-type: none"> • <i>PWD, Senior Citizens, LGU-owned facilities, Private health facilities, General Health Indicators</i> 	2016, 2017, 2018, 2019
		Confirmed COVID-19 Cases	2020, 2021
BUSINESS DATA	Identify businesses operating in the locality and status of their compliance to environmental monitoring and compliance	List of business establishments	2022
GEOSPATIAL AND RISK MANAGEMENT DATA	Get information on climate change and disaster risk management, vulnerability of the agriculture and fishery sector, vulnerability of flooding, drought and other calamities, and hunger, and the extent of disaster preparedness	Office of the Municipal Agricultural Services (OMAS) database	2018
		Masterlist of Corn Farmers	2021
		Masterlist of Cassava Farmers	2021
		Farm Mechanization	2021
		Masterlist of Rice Farmers	Not specified in the actual file
		Initial Disaster Report (Typhoon BisingPH)	2021
		Maps shapefiles	Not specified in the actual file
CASE MONITOR DATA	Monitor human-rights cases: <ul style="list-style-type: none"> • <i>Extralegal killings</i> • <i>Enforced disappearances</i> • <i>Torture</i> • <i>Trafficking</i> • <i>Sexual harassment</i> • <i>Programs under the Juvenile Justice and Welfare Act</i> 	None	Not applicable

MANDATED DATA	FORMAT	DISAGGREGATION	FREQUENCY OF DATA AVAILABILITY
CBMS HOUSEHOLD SURVEY DATA	xls	Location/Barangay Age, Sex, Marital Status Educational attainment Household size Income level, etc.	Once every three years
HEALTH DATA	doc	Not applicable (Aggregate Data)	Consolidated Annually
BUSINESS DATA	xls	Location/Barangay Vaccination Status Vaccination Type	Daily (since the start of COVID-19 outbreak)
	xls	Location/Barangay Status of Registration No. of Employee	Cannot be determined (no historical data)
GEOSPATIAL AND RISK MANAGEMENT DATA	xls	Location/Barangay	Cannot be determined (no historical data)
	xls	Location/Barangay	Cannot be determined (no historical data)
	xls	Location/Barangay	Cannot be determined (no historical data)
	xls	Location/Barangay Type of Equipment Number of units acquired	Cannot be determined (no historical data)
	xls	Location/Barangay Farm address Tenure/Status	Cannot be determined (no historical data)
	xls	Location/Barangay Kind of Assistance Given/Action Taken	After disaster event
	shp files	Location/Barangay	Cannot be determined
	CASE MONITOR DATA	Not applicable	Not applicable



A noticeable weakness in the LGU's data assets is the very scarce health data. At present, the digitally available health and health-related data are aggregated data that are used in various LGU plans or reports like the Comprehensive Land Use Plan (CLUP). Although local health units collect the Field Health Service Information System (FHSIS) data which is the designated national health statistics by the Department of Health (DOH) the disaggregated data are not available. The FHSIS is critical data for health policy formulation, health system advances and monitoring and evaluating of DOH and LGU programs on health. The data are collected from public primary healthcare facilities and hospitals following the standards set by the DOH. Unfortunately, there is no certainty that whenever data are collected daily, they are orderly consolidated weekly, quarterly and/or annually per health facility. Moreover, it cannot be determined whether there are digital copies of said collected FHSIS data.

Storing digitized FHSIS data will definitely enhance data quality and therefore, the quality of insights that can be derived from it.

On the other hand, there are some good practices in the LGU's data collection efforts that enhance the quality of local data and its usefulness for crafting local policies. For the purpose of discussion, some of the collected data are assessed in terms of the five dimensions of data quality – precision, accuracy, reliability, timeliness, and integrity. The detailed assessment of the quality of local data is displayed in the Data Processing template (Appendix 2). A summary of the result is shown in Table 2.

TABLE
2

QUALITY OF LOCAL DATA IN ARTECHE

DATASET	PRECISION	ACCURACY	RELIABILITY	TIMELINESS	INTEGRITY
CBMS DATA	Captures household-level data which supports the strategies of designing a well-targeted poverty alleviation policies and programs	Data undergo a systematic way of acquiring, processing, and verifying; allows to derive indexes and indicators for the intended purposes	Protocols for data collection do not change. Reliability of data is enhanced since data collection is conducted using ICT tools.	Data are collected every three years which is determined to be the ideal period for data collection to be responsive to the needs of the local government. It can also be conducted at shorter intervals for purposes peculiarly useful for the LGU.	Makes use of ICT tools which reduces human errors in collecting, processing, and validating data.
HEALTH DATA (AGGREGATED LIST OF PWD)	Contains the magnitude and proportion of persons-with-disability per barangay classified per type of disability, sex, etc.	Important individual data useful for policymaking are missing like age, educational attainment, employment status, marital status, skills, etc.	Aggregate data that provides only the magnitude and/or proportion of the population who are PWDs have limited applicability and usability. It cannot be used for the purpose of identifying the vulnerable individuals who will receive cash or in-kind transfers when the need arises.	Data are collected yearly as determined to be an ideal period for data collection. But no historical data are available.	Aggregated data makes data validation impossible since the profiles of every individual are missing. There is no way to counter-check the reported numbers.
BUSINESS DATA	Contains a list of business establishments in every barangay classified status of registration.	The dataset has a number of missing data like the name of the establishment. Although the format of the variables are consistent. Other variables like owner's contact no. and email address are missing. The no. of employees are disaggregated not by sex. The type or nature of business and business size are not collected.	Data are generated from regular licence issuance operations of the LGU. But, the system is still new and needs to be further modified or improved.	Data are up to date and taken at the point of application.	Data are collected electronically and reduces human errors.
GEOSPATIAL AND RISK MANAGEMENT DATA (GEOSPATIAL DATA)	Includes GIS data (like roads, residential and building expansion, evacuation centers, health facilities, etc.) per barangay	The data provide accurate information on where the land and water areas are, roads, administrative boundaries, high density areas, remote areas, place names, etc.	Follows standard protocols for data collection. Data collections are conducted using ICT tools.	Data are collected or accumulated over the years.	Collected data are validated.

COMMUNITY-BASED MONITORING SYSTEM (CBMS) DATA

In 2019, the mandated Community-based Monitoring System (CBMS) data was implemented in Artech. The conduct of the CBMS generated data on population and household size, health and nutrition, housing, water and sanitation, basic education and literacy, income and livelihood, and peace and order of the municipality.

This household-level data can be used by the LGU to conduct a comprehensive poverty analysis to shape policies and interventions targeting the appropriate households.

The LGU can also obtain information from the collected CBMS data for disaster risk reduction and management. The CBMS data are to be collected every three years making it a useful tool for monitoring the impact of future poverty alleviation and development programs. Unlike the traditional paper-based ones, CBMS data collected digitally has many advantages mainly because they have better quality as presented in Table 2. The data undergo a systematic way of acquiring, processing, and verifying, the protocols for data collection do not change, and the use of ICT tools from collection to validation stage reduces human errors.



HEALTH DATA

Aside from the FHSIS data, the LGU reports many health and health related data like data on persons with disability (PWDs). The dataset contains only the magnitude and proportion of persons-with-disability per barangay classified per type of disability, sex, etc. The problem with aggregate data is that it provides only the magnitude and/or proportion of the population who are PWDs. Therefore, it has limited applicability and usability. It cannot be used for the purpose of identifying the vulnerable individuals who will receive cash or in-kind transfers when the need arises. Aggregated data also makes data validation impossible since the profiles of every individual are missing. There is no way to counter-check the reported numbers.

An ideal practice, which is already adopted by other LGUs, is to collect timely individual-level data of PWDs.

In addition to the collection of individual's type of disability and sex, educational attainment, employment status, marital status and skills are some other important variables to be considered. A good list should also be cumulative or updated yearly. Registered PWDs in the previous years should be included in the next periods and not that only the registrants in the applicable year are in the list.



BUSINESS DATA

The LGU maintains a list of business establishments operating in the municipality in Excel format. The data are disaggregated per location or barangay and status of registration.

The data can be used to identify the nature of businesses thriving in the area. Most establishments are micro and small businesses that are involved in food production and retail sale (like restaurants, bakery, sari-sari stores, etc.).

A good feature of the dataset is it includes the number of employees for each establishment which can serve as an indirect measure of employment in the area. However, caution should be exercised when using the dataset because there may still be informal businesses which make up a significant portion of active businesses in the locality that are not covered. Also, the data can be used to track compliance of the establishments on application and renewal of licences. Unfortunately, it does not contain data on environmental compliance and clearances of registered establishments as required by law to be monitored by the LGU. Still, from Table 2 timeliness appears to be the best characteristic of this dataset because the data are up to date and taken at the point of application. Data are collected electronically such that it reduces encoding errors.



GEOSPATIAL AND RISK MANAGEMENT DATA

Geospatial data, data on farmers and agricultural production are some of the helpful sources of information on disaster risk management, determining the vulnerability of the agriculture sector, vulnerability of the populace to flooding, drought and other calamities, and hunger, and the extent of disaster preparedness of the municipality. These are collected by the LGU and are available in digital format. The LGU also prepares an initial disaster report after a disaster event has occurred like when Typhoon Bising hit the area in April 2021. It contains barangay-level data on how many families and individuals were affected, the number of casualties, damaged houses, evacuees and designated evacuation centers per barangay. This is a good practice that can help facilitate faster disaster response operations in barangays where

help is most needed. Another set of interesting data held by the LGU is its geospatial data. The municipality collects geospatial data from all its barangays that provide accurate information where its land and water areas are, as well as administrative boundaries, roads, residential and building areas, evacuation centers, health facilities, high density areas, remote areas, etc.

For a disaster-prone municipality like Arteche, collecting a good amount of geospatial data is definitely necessary for efficient disaster risk reduction and management.





Data-Driven Development
in the Philippines

GENERAL RECOMMENDATIONS

The availability, conditions, and characteristics of local data are critical to draw valuable insights. High quality data translates into better opportunities to transform the lives of the people. However, based on the previous discussions, the municipality's data assets are still lacking important elements to further progress on data driven development. To ensure better data availability and quality of local data, the following are suggested:

1 STRENGTHEN DIGITAL DATA COLLECTION

Digitizing data or collecting data digitally is vital to ensure data accessibility and quality. Basically, it increases data availability that allows for readily performing computations and data analysis and enhances data quality so that derived insights are reliable and effective. Digital data collection tools facilitate faster and more reliable data collection. There are free or open-source tools that are secured like KoBo Toolbox which the LGU can use. Data are collected real-time and made available in machine readable format. It also readily provides basic analysis. Hence, shifting from paper-based to digital collection reduces costs and increases efficiency in governance.

Digital methods can best be explored to expand business data and geospatial and risk management data, in the hope of collecting more comprehensive, accurate and timely data necessary to make richer analysis and derive more valuable insights in these areas. In terms of health, the LGU can start digitizing FHSIS data for faster and more reliable consolidation of health and health-related data. Electronic Medical Records (EMRs) are also not

available in the LGU. Acquiring an EMR system is a good example of health care innovation that enhances data quality because unlike the traditional paper-based ones, using the digitally collected and stored patient data like the EMR tend to be more precise, accurate, reliable, timely and less prone to human errors.

The CBMS is to be implemented consistently and data should be made available immediately. An important reminder in acquiring and storing survey data like the CBMS is to have proper documentation of the entire process from survey conceptualization to implementation. Some important elements of the documentation should include guides for data users to track details and understand how and what data were collected like the metadata, codebook or coding information, questionnaires, and so on. These documents can make it easier to retrieve datasets and facilitate accurate use and reuse of collected data. It is a standard discipline in data management which the municipality can adopt to ensure the integrity of local data.

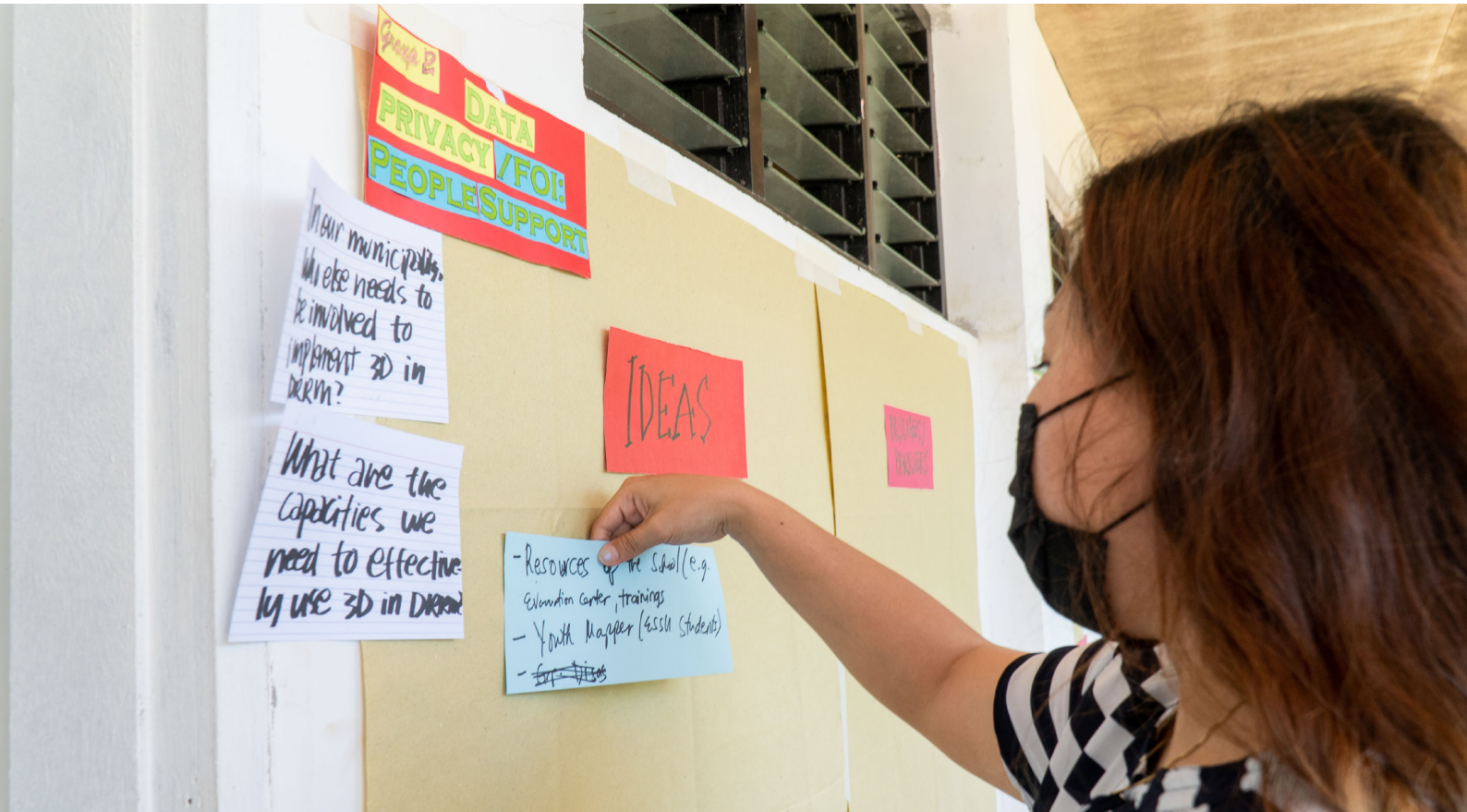
2

DETERMINE AND COLLECT ALL RELEVANT VARIABLES

There are missing or uncollected variables useful to influence decisions and policies. Consider geospatial data, for example. An accurate, complete and up-to-date geospatial data are needed to make timely decisions and appropriate actions, in any disaster situation. It can inform the overall extent of damages caused by the calamity and the affected communities and individuals thereby, helps in assessing the individual needs and determining the type and scale of assistance to be provided. For this reason, it is essential for the LGU to collect geospatial data that will guide decisions for disaster prevention (like topographical maps, communities and settlements, etc.), preparedness (location of facilities, evacuation centers, relocation sites, etc.), mitigation (land use, zoning, etc.), and response (accessible roads, feasible transport routes, etc.) initiatives. Geospatial data also provide critical information for recovery or post-disaster rehabilitation, reconstruction and government administration for sustainable community planning.

In the area of health, aside from the lack of good data from public health facilities there is uncollected data from private facilities which are beneficial to overall health system operations. Health-related data that are important in coming up with effective health programs or strengthening the existing health programs are unavailable too. For example, data on health appropriations and expenditure, health care products and facilities utilization, and human resources for health and their corresponding competencies. There is no data to measure health care worker knowledge like the skilled health workers' continuing professional development. Such data will help ensure that the health workers' knowledge and skills stay relevant, and their capacities keep pace with current health standards. Good quality health and health worker data can be a reliable source of information for LGU's disaster management.





3

AVOID REDUNDANCIES AND KEEP HISTORICAL DATA

The implementation of the CBMS, presents an opportunity for the LGU to synchronize data collection efforts to minimize the costs and maximize the benefits of data collection. Activities like securing a list of households per barangay may no longer be a stand-alone activity. Historical data should also be maintained because it can provide critical information with high value to society. For example, the analysis of health data from previous years

can provide insights about disease trends and risk factors, patterns of care, health care use and its costs, and the outcomes of public health interventions in the municipality. Similarly, an updated list of vulnerable individuals will facilitate faster, more reliable and transparent distribution of cash and in-kind assistance to eligible beneficiaries.

4

ENRICH LOCAL DATA THROUGH PARTNERSHIPS WITH DIFFERENT SECTORS

As mentioned earlier, peoples' participation is the sufficient condition for reforms in governance. Starting from data publication, the LGU can advance work to data dissemination through intentional coalition building. Data should be made easily accessible to spark innovation. People should know data are available and use it to yield

benefits to the community. Eliminating the horizontal (between LGU and the public) and vertical (between LGU and national government offices and agencies) barriers to data access is the first big step towards free flow of data that will potentially broaden the database in the locality.

APPENDIX

1

NATIONALLY MANDATED DATASETS FOR LOCAL GOVERNMENT

SECTOR	DATASET	DESCRIPTION	LEGISLATION/ AGENCY DIRECTIVE	
MULTIPLE	Community-based Monitoring System	Generates updated and disaggregated data necessary in targeting beneficiaries, conducting more comprehensive poverty analysis and needs prioritization, designing appropriate policies and interventions, and monitoring impact over time.	Community-Based Monitoring System Act (RA 11315)	
	Climate Disaster Risk Assessment (CDRA)	Considers the "risks and vulnerabilities of exposed elements, namely the people, urban area, agriculture, forestry and fishery production areas, critical points facilities, and lifeline infrastructure associated with natural hazards and climate change."	HLURB Board of Commissioners Resolution No. 915 (2014), "Approving the Supplemental Guidelines for Mainstreaming Climate Change Adaptation and Disaster Risk Reduction in the Comprehensive Land Use Plan"	
	Local Disaster Risk Reduction and Management Fund (LDRRMF) Utilization	Contains the specific projects an LGU undertakes relative to DRRM, and how the funding for such projects have been utilized.	Philippine Disaster Risk Reduction and Management Act of 2010 (RA 10121)	
DISASTER RISK REDUCTION & MANAGEMENT (DRRM)	Disaster Management Services	Includes up-to-date information about the LGU in times of disasters (contact information of the LGU and agencies required to respond to emergency situations, evacuation areas, relief sites, and pick-up points, etc.)	The Free Mobile Disaster Alerts (RA 10639)	
	Persons with Disabilities (PWDs)	Contains a list of PWDs in the LGU, and the relevant data including specific disabilities of the PWDs	Rep. Act No. 10070, amending Rep. Act No. 7277, otherwise known as The Magna Carta for Disabled Person	
	Responsible Parenthood and Reproductive Health	Data on women's health and safe motherhood and health services in the LGU (list of deceased mothers due to maternal deaths, deceased infants, hospital and health centers in the LGU, assets relevant to dealing with maternity issues)	The Responsible Parenthood and Reproductive Health Act of 2012 (RA 10354)	
	Excise Tax Allocation and Utilization	Contains a list of programs for tobacco farmers and workers and funding support to health facilities	Implementing Rules and Regulations of RA 10351, DOF-BIR-DOH-DBM-DA Joint Circular No. 01-14	
	Mental Health Registry	A registry of people who have been attended to and/or served, the respective kinds of mental illness or disability, duration and result of the treatment, and patients/service profile.	Mental Health Act (RA 11036)	
	Universal Health Care programs	Contains policies and programs undertaken to strengthen and broaden existing health policies to support universal health care implementation	Universal Health Care Act (RA 11223)	
	Disease Surveillance	Report on cases of infection or exposure to outbreak/pandemic/epidemic or any health event of public concern.	IRR of the Mandatory Reporting of Notifiable Diseases and Health Events of Public Health Concern Act (RA 11332) 2020 Revised IRR of RA 11332	
	HEALTH			

APPENDIX

1

NATIONALLY MANDATED DATASETS FOR LOCAL GOVERNMENT

SECTOR	DATASET	DESCRIPTION	LEGISLATION/ AGENCY DIRECTIVE
BUSINESS	Business Establishments	A list of business establishments issued a business licence, clearance, permit, certification, or authorization including information of applicant or requesting party by various local government departments, such as local taxes and clearances, building clearance, sanitary permit, zoning clearance, and other specific LGU requirements, including the fire clearance.	Promoting Ease of Doing Business and Efficient Delivery of Government Services, Amending for the Purpose R.A. No. 9485 (RA 11032)
	Fisheries Information	A database of who the fisherfolk are in the LGU, their equipment such as boats, nets, etc., where they conduct fishing, where they sell their catch, etc.	Adopting a National Plan of Action to Prevent, Deter and Eliminate Illegal, Unreported, and Unregulated Fishing, and for Other Purposes, Executive Order No. 154
ENVIRONMENT	Mining Industry	A database of local ordinances and contracts related to mining for review of the Mining Industry Coordinating Council	Institutionalising and Implementing Reforms in the Philippine Mining Sector, Executive Order No. 79
	Retiring Government Employees	A database of government officials and employees who shall compulsorily retire in the next succeeding fiscal year	RA 10154 & Implementing Rules and Regulations of Republic Act No. 10154 (An Act Requiring All Concerned Government Agencies to Ensure the Early Release of the Retirement Pay, Pensions, Gratuities and Other Benefits of Retiring Government Employees)
LABOR	JobStart Program Registry	A database of all registrants under the Jobstart Philippines Act, and employers who are willing to facilitate qualified registrants for the program	JobStart Philippines Act (RA No. 10869)
	Healthcare Workers	A database of healthcare workers in the LGU to identify those eligible for life insurance, accommodation, transportation, and meals.	Implementing Guidelines on the Provision of Life Insurance, Accommodation, Transportation, and Meals to Public and Private Health Workers under RA 11494, DOH Administrative Order No. 0054-20
AGRICULTURE	Labor Market Information	A registry of establishments operating in the locality, present number and nature of jobs, projection of jobs that the establishment will offer in the next five (5) years	PESO Act (RA 8759), as amended by Rep. Act No. 10691
	Local Producers	A masterlist of farmers, farmworkers, and rice cooperative and associations who are eligible beneficiaries of agriculture programs	Liberalizing the Importation, Exportation and Trading of Rice, Lifting for the Purpose the Quantitative Import Restriction on Rice (RA 11203)
	Coconut Farmers	A registry of coconut farmers and their organization	Coconut Farmers and Industry Trust Fund Act (RA 11524)
	Corn Production	A database of corn farms and corn production in the LGU	General Guidelines on the Implementation of the Fall Armyworm (FAW) Management Project under the Bayanihan to Recover as One Act (RA 11494), DA Memorandum Order No. 26-21

SECTOR	DATASET	DESCRIPTION	LEGISLATION/ AGENCY DIRECTIVE
SOCIAL SERVICES	Pantawid Pasada Program	A database of all tricycle drivers within the LGU eligible for Public Transport Assistance Program	Instituting the Public Transport Assistance Program — Pantawid Pasada, Executive Order No. 32
	Enhanced Comprehensive Local Integration Program	A database of former rebels/extremists in their locality or spotted in the locality.	This is the Revised IRR of AO No.10, s. 2018 entitled: Centralizing All Government Efforts for the Reintegration of Former Rebels and Creating for the Purpose an Inter-Agency Task Force aimed at empowering individual former rebels, former violent extremists, their families and communities.
	Emergency Subsidies Beneficiaries	A list of citizens in the locality who have received emergency subsidies under the Bayanihan to Recover as One Act.	Special Guidelines on the Implementation of the Provisions of Republic Act No. 11494 or the "Bayanihan to Recover as One Act" Concerning DSWD, DSWD Memorandum Circular No. 030-20
PUBLIC UTILITIES ENERGY	Energy Consumption	Includes the amount of energy consumption by the LGU	DOE DEPARTMENT CIRCULAR NO. DC2019-11-0014, Implementing Rules and Regulations of Republic Act No. 11285 (Energy Efficiency and Conservation Act)
	MULTIPLE	Human Rights Violations	A database of extra-legal killings, enforced disappearances, torture, and other grave violations of the right to life, liberty, and security of persons that might have occurred in that locality as wells as cases of trafficking in persons.
Traffic Violations		A database of traffic violation receipts issued by the LGU.	IRR of RA 10930, An Act Rationalizing and Strengthening the Policy Regarding Driver's License amending Section23 of the Land and Transportation and Traffic Code (RA 4136)
Urban Development		A database of areas that were subject to housing and urban development, and further, what the nature of the development was.	IRR of Department of Human Settlements and Urban Development Act (RA 11201)
Gender		A database of sexual harassment in streets and public spaces.	Safe Spaces Act (RA 11313)
Fire Prevention and Safety		Database of establishments and homes, within which those who comply or not comply with fire safety measures are marked accordingly.	Revised IRR of the The Fire Code of the Philippines Of 2008 (RA 9514)
Alternative Learning System (ALS)		Database of ALS learners (out-of-school children in special cases and adults), both prospective and current.	Alternative Learning System Act (RA 11510)

APPENDIX 2

DATA PROCESSING TEMPLATE

(FULL VERSION AVAILABLE UPON REQUEST)

DATASET NAME	YEAR	FILE NAME	FILE TYPE	SHEET NAME	COLUMN NAME	COLUMN DATA TYPE	ENUM VALUES	VALID RANGE	NO. OF RECORDS	NO. OF NULL VALUES
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	NO	INTEGER	N/A	N/A	449	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	NAME	STRING	N/A	N/A	449	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	BIRTHDAY	DATE	N/A	N/A	242	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	GENDER	STRING	M/F	M/F	0	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	FARM ADDRESS	STRING	N/A	N/A	501	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	AREA (HAS)	DECIMAL	N/A	N/A	501	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	AREA- ECOSYSTEM	DECIMAL	RF	RF	52	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	TENURE/STATUS	STRING	T/O	T/O	50	N/A
ARTECHE RICE FARMERS	NOV 25 2021	ARTECHE RICE FARMERS	XLSX	SHEET1	RESIDENCE BARANGAY	STRING	N/A	N/A	501	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 15 2021	MASTERLIST OF CASSAVA FARMERS 2021	XLSX	TAWAGAN	NO	INTEGER	N/A	N/A	37	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 16 2021	MASTERLIST OF CASSAVA FARMERS 2022	XLSX	TAWAGAN	NAME OF FARMER - LAST NAME	STRING	N/A	N/A	37	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 17 2021	MASTERLIST OF CASSAVA FARMERS 2023	XLSX	TAWAGAN	NAME OF FARMER - MIDDLE NAME	STRING	N/A	N/A	37	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 18 2021	MASTERLIST OF CASSAVA FARMERS 2024	XLSX	TAWAGAN	NAME OF FARMER - FIRST NAME	STRING	N/A	N/A	32	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 19 2021	MASTERLIST OF CASSAVA FARMERS 2025	XLSX	TAWAGAN	BIRTHDATE	STRING	N/A	N/A	0	N/A

DATASET NAME	% NULL VALUES	NO. OF INVALID VALUES	% INVALID VALUES	% TOTAL INVALID+ NULL	COMMENTS	DO THE VALUES NEED TO BE PROCESSED OR CLEANED?	IF YES, DESCRIBE NATURE OF PROCESSING OR CLEANING DONE.	IS IT A PII COLUMN? DOES IT (POTENTIALLY) NEED TO BE MASKED?	FINAL VARIABLE NAME	FINAL VARIABLE TYPE
ARTECHE RICE FARMERS		89.62%	0.00%	89.62%		NO		NO	NO	INTEGER
ARTECHE RICE FARMERS		89.62%	0.00%	89.62%	some entries have no farmer name despite having other complete columns in the row	NO		YES	NAME	STRING
ARTECHE RICE FARMERS		48.30%	0.00%	48.30%	ASSUME MM/DD/YYYY FORMAT	NO		YES	BIRTHDAY	DATE
ARTECHE RICE FARMERS		0.00%	0.00%	0.00%		NO		NO	GENDER	STRING
ARTECHE RICE FARMERS		100.00%	0.00%	100.00%		NO		NO	FARM ADDRESS	STRING
ARTECHE RICE FARMERS		100.00%	0.00%	100.00%		NO		NO	AREA - HECTARES	DECIMAL
ARTECHE RICE FARMERS		10.38%	0.00%	10.38%		NO		NO	AREA- ECOSYSTEM	DECIMAL
ARTECHE RICE FARMERS		9.98%	0.00%	9.98%		NO		NO	TENURE STATUS	STRING
ARTECHE RICE FARMERS		100.00%	0.00%	100.00%		NO		NO	RESIDENCE BARANGAY	STRING
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%		NO			NO	INTEGER
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%		NO			NAME OF FARMER - LAST NAME	STRING
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%	FIRST NAMES AND MIDDLE NAMES ARE MIXED UP	YES	PUT MIDDLE AND FIRST NAMES IN PROPER COLUMNS	YES	NAME OF FARMER - MIDDLE NAME	STRING
MASTERLIST OF CASSAVA FARMERS		13.51%		13.51%	FIRST NAMES AND MIDDLE NAMES ARE MIXED UP	YES	PUT MIDDLE AND FIRST NAMES IN PROPER COLUMNS	YES	NAME OF FARMER - FIRST NAME	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%				YES			BIRTHDATE	STRING

APPENDIX

2

DATA PROCESSING TEMPLATE*(FULL VERSION AVAILABLE UPON REQUEST)*

DATASET NAME	YEAR	FILE NAME	FILE TYPE	SHEET NAME	COLUMN NAME	COLUMN DATA TYPE	ENUM VALUES	VALID RANGE	NO. OF RECORDS	NO. OF NULL VALUES
MASTERLIST OF CASSAVA FARMERS	NOV 20 2021	MASTERLIST OF CASSAVA FARMERS 2026	XLSX	TAWAGAN	AREA PLANTED (HA)	DECIMAL	N/A	0-1	37	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 21 2021	MASTERLIST OF CASSAVA FARMERS 2027	XLSX	TAWAGAN	FARM LOCATION	STRING	N/A	N/A	37	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 22 2021	MASTERLIST OF CASSAVA FARMERS 2028	XLSX	TAWAGAN	VARIETY	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 23 2021	MASTERLIST OF CASSAVA FARMERS 2029	XLSX	TAWAGAN	STAGE OF CASSAVA	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 24 2021	MASTERLIST OF CASSAVA FARMERS 2030	XLSX	TAWAGAN	TENURIAL STATUS - OWNER	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 25 2021	MASTERLIST OF CASSAVA FARMERS 2031	XLSX	TAWAGAN	TENURIAL STATUS - TENANT	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 26 2021	MASTERLIST OF CASSAVA FARMERS 2032	XLSX	TAWAGAN	TENURIAL STATUS - NAME OF OWNER	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 27 2021	MASTERLIST OF CASSAVA FARMERS 2033	XLSX	TAWAGAN	RSBSA NO.	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 28 2021	MASTERLIST OF CASSAVA FARMERS 2034	XLSX	TAWAGAN	PRODUCTION - KG	DECIMAL	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 29 2021	MASTERLIST OF CASSAVA FARMERS 2035	XLSX	TAWAGAN	PRODUCTION - INCOME	DECIMAL	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	NOV 30 2021	MASTERLIST OF CASSAVA FARMERS 2036	XLSX	CATUMSAN	NO	INTEGER	N/A	N/A	48	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 1 2021	MASTERLIST OF CASSAVA FARMERS 2037	XLSX	CATUMSAN	NAME OF FARMER - LAST NAME	STRING	N/A	N/A	48	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 2 2021	MASTERLIST OF CASSAVA FARMERS 2038	XLSX	CATUMSAN	NAME OF FARMER - MIDDLE NAME	STRING	N/A	N/A	48	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 3 2021	MASTERLIST OF CASSAVA FARMERS 2039	XLSX	CATUMSAN	NAME OF FARMER - FIRST NAME	STRING	N/A	N/A	20	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 4 2021	MASTERLIST OF CASSAVA FARMERS 2040	XLSX	CATUMSAN	BIRTHDATE	STRING	N/A	N/A	0	N/A

DATASET NAME	% NULL VALUES	NO. OF INVALID VALUES	% INVALID VALUES	% TOTAL INVALID+ NULL	COMMENTS	DO THE VALUES NEED TO BE PROCESSED OR CLEANED?	IF YES, DESCRIBE NATURE OF PROCESSING OR CLEANING DONE.	IS IT A PII COLUMN? DOES IT (POTENTIALLY) NEED TO BE MASKED?	FINAL VARIABLE NAME	FINAL VARIABLE TYPE
MASTERLIST OF CASSAVA FARMERS		0.00%	0	0.00%		YES	REMOVE 'HA'		AREA PLANTED	DECIMAL
MASTERLIST OF CASSAVA FARMERS		0.00%	0	0.00%		NO			FARM LOCATION	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE SHEETS	NO			VARIETY	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			STAGE OF CASSAVA	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			TENURIAL STATUS- OWNER	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			TENURIAL STATUS- OWNER	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO		YES	TENURIAL STATUS- NAME OF OWNER	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NO VALUES	NO			RSBSA NO	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			PRODUCTION KG	DECIMAL
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			PRODUCTION INCOME	DECIMAL
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%					NO	INTEGER
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%					NAME OF FARMER - LAST NAME	STRING
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%	FIRST NAMES AND MIDDLE NAMES ARE MIXED UP	YES	PUT MIDDLE AND FIRST NAMES IN PROPER COLUMNS	YES	NAME OF FARMER - MIDDLE NAME	STRING
MASTERLIST OF CASSAVA FARMERS		58.33%		58.33%	FIRST NAMES AND MIDDLE NAMES ARE MIXED UP	YES	PUT MIDDLE AND FIRST NAMES IN PROPER COLUMNS	YES	NAME OF FARMER - FIRST NAME	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%				NO			BIRTHDATE	STRING

APPENDIX 2

DATA PROCESSING TEMPLATE

(FULL VERSION AVAILABLE UPON REQUEST)

DATASET NAME	YEAR	FILE NAME	FILE TYPE	SHEET NAME	COLUMN NAME	COLUMN DATA TYPE	ENUM VALUES	VALID RANGE	NO. OF RECORDS	NO. OF NULL VALUES
MASTERLIST OF CASSAVA FARMERS	DEC 5 2021	MASTERLIST OF CASSAVA FARMERS 2041	XLSX	CATUMSAN	AREA PLANTED (HA)	DECIMAL	N/A	0-1	48	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 6 2021	MASTERLIST OF CASSAVA FARMERS 2042	XLSX	CATUMSAN	FARM LOCATION	STRING	N/A	N/A	47	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 7 2021	MASTERLIST OF CASSAVA FARMERS 2043	XLSX	CATUMSAN	VARIETY	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 8 2021	MASTERLIST OF CASSAVA FARMERS 2044	XLSX	CATUMSAN	STAGE OF CASSAVA	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 9 2021	MASTERLIST OF CASSAVA FARMERS 2045	XLSX	CATUMSAN	TENURIAL STATUS - OWNER	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 10 2021	MASTERLIST OF CASSAVA FARMERS 2046	XLSX	CATUMSAN	TENURIAL STATUS - TENANT	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 11 2021	MASTERLIST OF CASSAVA FARMERS 2047	XLSX	CATUMSAN	TENURIAL STATUS - NAME OF OWNER	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 12 2021	MASTERLIST OF CASSAVA FARMERS 2048	XLSX	CATUMSAN	RSBSA NO.	STRING	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 13 2021	MASTERLIST OF CASSAVA FARMERS 2049	XLSX	CATUMSAN	PRODUCTION - KG	DECIMAL	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 14 2021	MASTERLIST OF CASSAVA FARMERS 2050	XLSX	CATUMSAN	PRODUCTION - INCOME	DECIMAL	N/A	N/A	0	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 15 2021	MASTERLIST OF CASSAVA FARMERS 2051	XLSX	CAGSALAY	NO	STRING	N/A	N/A	12	
MASTERLIST OF CASSAVA FARMERS	DEC 16 2021	MASTERLIST OF CASSAVA FARMERS 2052	XLSX	CAGSALAY	NAME OF FARMER - LAST NAME	STRING	N/A	N/A	12	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 17 2021	MASTERLIST OF CASSAVA FARMERS 2053	XLSX	CAGSALAY	NAME OF FARMER - MIDDLE NAME	STRING	N/A	N/A	12	N/A
MASTERLIST OF CASSAVA FARMERS	DEC 18 2021	MASTERLIST OF CASSAVA FARMERS 2054	XLSX	CAGSALAY	NAME OF FARMER - FIRST NAME	STRING	N/A	N/A	12	N/A

DATASET NAME	% NULL VALUES	NO. OF INVALID VALUES	% INVALID VALUES	% TOTAL INVALID+ NULL	COMMENTS	DO THE VALUES NEED TO BE PROCESSED OR CLEANED?	IF YES, DESCRIBE NATURE OF PROCESSING OR CLEANING DONE.	IS IT A PII COLUMN? DOES IT (POTENTIALLY) NEED TO BE MASKED?	FINAL VARIABLE NAME	FINAL VARIABLE TYPE
MASTERLIST OF CASSAVA FARMERS		0.00%	0	0.00%		YES	REMOVE 'HA'		AREA PLANTED	DECIMAL
MASTERLIST OF CASSAVA FARMERS		2.08%	0	2.08%		NO			FARM LOCATION	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE SHEETS	NO			VARIETY	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			STAGE OF CASSAVA	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			TENURIAL STATUS- OWNER	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			TENURIAL STATUS- OWNER	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO		YES	TENURIAL STATUS- NAME OF OWNER	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NO VALUES	NO			RSBSA NO	STRING
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			PRODUCTION KG	DECIMAL
MASTERLIST OF CASSAVA FARMERS		100.00%	0	100.00%	NEW COLUMN NOT PRESENT IN REST OF THE CELLS	NO			PRODUCTION INCOME	DECIMAL
MASTERLIST OF CASSAVA FARMERS		0.00%				NO			NO	STRING
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%		NO			NAME OF FARMER - LAST NAME	STRING
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%	FIRST NAMES AND MIDDLE NAMES ARE MIXED UP	NO	PUT MIDDLE AND FIRST NAMES IN PROPER COLUMNS	YES	NAME OF FARMER - MIDDLE NAME	STRING
MASTERLIST OF CASSAVA FARMERS		0.00%		0.00%	FIRST NAMES AND MIDDLE NAMES ARE MIXED UP	NO	PUT MIDDLE AND FIRST NAMES IN PROPER COLUMNS	YES	NAME OF FARMER - FIRST NAME	STRING

PH 3D



Data-Driven Development in the Philippines



Co-funded by
the European Union



Unit 1403 West Trade Center, 132 West Avenue,
Quezon City, Philippines 1104

Phone: (02) 7373 1198 | **Mobile:** (0999) 489 2242
Email: aer@aer.ph | **Website:** aer.ph