



Logical Modelling of Statistical Data Using the SDMX Standard: Case Study on the Quarterly Gross Regional Domestic Product Table

K Amandasari^{1*}, N Y Pratama², F S Aditama², W Marsisno¹

¹ Politeknik Statistika STIS, Jakarta, Indonesia

² BPS Statistics, Jakarta, Indonesia

*Corresponding author's email : kartikamandasari49@gmail.com

Abstract. Poverty, as a national issue, necessitates data-driven policy planning informed by accurate and consistent statistics. To ensure the optimal quality and consistency of statistical data reporting across diverse regions, the adoption of an international standard is crucial. The Statistical Data and Metadata Exchange (SDMX) standard facilitates the structured exchange of data and metadata. This study aims to design and implement a statistical indicator data model using the SDMX standard to improve table consistency. We utilized Quarterly Provincial Gross Regional Domestic Product (GRDP) data as a case study and applied the Design Science Research Method (DSRM) as the methodology. The results demonstrate that modeling the GRDP data using SDMX yields a uniform and highly consistent table structure, significantly enhancing the consistency of statistical data reporting across regions.

Keyword: Statistical Data, Data Modelling, SDMX.

1. Introduction

Economic development in Indonesia is currently challenged by poverty. This social issue is a primary focus of government attention and serves as a critical indicator of the inclusiveness and impact of economic growth [1]. Poverty alleviation efforts must be executed comprehensively, particularly through effective planning. To facilitate data-driven decisions, it is essential to ensure that the data utilized is of high quality [2]. In this context, Gross Domestic Product (GDP) data is commonly used as a benchmark for the economic conditions of a region within a specific period [3]. Generally, data is utilized for public policy planning, business value enhancement, serving as a research reference, and for decision-making [4]. Various parties, including governments, organizations, and private companies, are increasingly opening up data access to facilitate its widespread use. Data transparency and openness are instrumental in fostering public trust [5].



A survey conducted by the Badan Pusat Statistik (BPS) Indonesia indicates a 5.36% increase in BPS data users from 2021 to 2023 [6]. The growing need for data in Indonesia highlights its increasing importance [7]. As a reliable data source, BPS must ensure that the data presented is easily comprehensible and usable by data consumers. Accurate and relevant data generates information that aids in formulating targeted policies, and also enables the public to comprehend the socio-economic and environmental trends in their surroundings. Data quality is a crucial aspect [8], defined as the extent to which data fulfills established objectives [9]. Various characteristics serve as quality references for statistical data, including accuracy, timeliness, and comparability. Comparability, however, plays a paramount role in international statistics due to the significant challenges in cross-country statistical comparison [10]. Despite the increasing data needs, consumers often face difficulties obtaining official government data. This is attributable to the wide variety of data generated, limited data access rights, and non-standardized data formats [7]. Consequently, the establishment of standards in statistical data production is imperative.

The international standard for statistical data and metadata exchange is the Statistical Data and Metadata Exchange (SDMX). SDMX is an ISO standard designed to describe statistical data and metadata, thereby standardizing and improving the efficiency of data sharing among statistical organizations and related institutions [11]. Eight international organizations support SDMX: the Bank for International Settlements (BIS), the European Central Bank (ECB), Eurostat, the International Labor Organization (ILO), the International Monetary Fund (IMF), the Organization for Economic Co-operation and Development (OECD), the United Nations Statistical Division (UNSD), and the World Bank. SDMX produces technical standards and facilitates the alignment of statistical terms, classifications, and concepts [11].

Regarding SDMX implementation, related research titled "Statistical Data Governance Based on the SDMX" [12] is notable. This study discusses the development of a statistical data governance framework based on the SDMX standard. The findings indicate that SDMX implementation achieves a simpler representation and conserves storage space by reducing redundancy. Another relevant study is titled Integrated Statistics: A Journey Worthwhile [13]. This study highlights the importance of integrated statistical data amidst increasing demand for data disaggregation, supporting the principle of open data. The findings suggest that implementing international standards such as SDMX allows for more efficient data and metadata exchange within the National Statistical System (NSS). Data and metadata are well-structured in a reusable format, establishing interconnectedness between the data and metadata. Furthermore, research titled "Quality Assurance From an Internationally Standardized and Generic Data Validation Ecosystem" [14] found that the implementation of SDMX in this system improves international data validation standards and facilitates inter-organizational harmonization.

SDMX offers three main advantages for statistical organizations: (1) efficiency in data sharing through standardization and automation that meets international data reporting specifications; (2) data consistency and comparability via information models that link data and metadata for enhanced contextual understanding; and (3) extensive technological support through an ecosystem of open-source tools and flexibility in data formats (e.g., XML, JSON, and CSV), which simplifies data implementation and exchange [15]. By implementing SDMX on BPS's GDP tables, statistical data is expected to be more readily analyzed, interpreted, and shared in a standardized and consistent format.



2. Research Method

This research uses the Design Science Research Method (DSRM). This method integrates practice, research, and design into a set of goals to create digital artifacts and gain design knowledge through a systematic procedure [20]. DSRM is commonly used in information systems research [17]. The systematic procedure of DSRM applied in this study is illustrated in Figure 1.

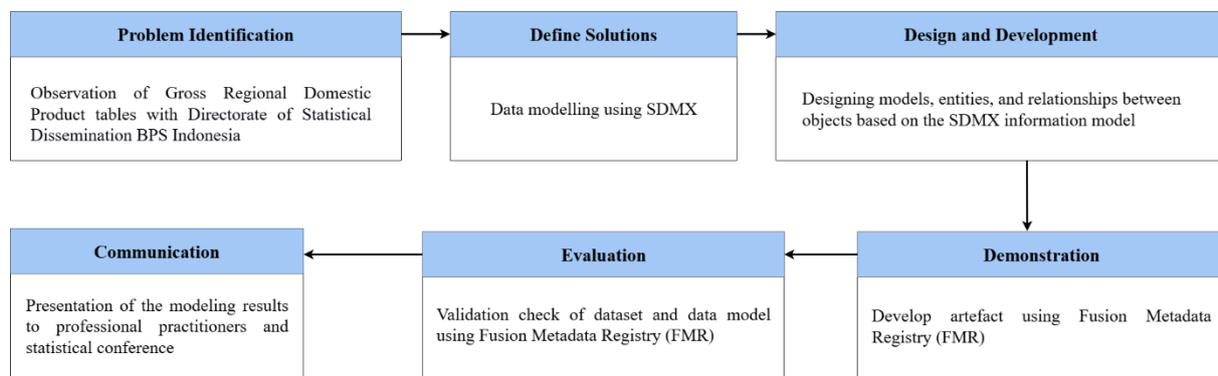


Figure 1. Research methodology

The first stage is problem identification. In this study, problem identification was carried out through observation with the Directorate of Dissemination of BPS RI. At this stage, it was found that BPS tables, specifically quarterly GDP data across several provinces, exhibited format inconsistencies, which complicated data analysis and comparison. Following problem identification, the second stage is determining the solution, which is the implementation of SDMX. This is followed by a literature study on the SDMX implementation as applied to the Provincial GDP Table.

The third stage is data modeling, which is the process of visualizing data interconnectedness within an information system, thereby clarifying the relationship between the data and its structure [18]. In this study, a data model was created in SDMX-ML format (an XML implementation based on the SDMX standard). The creation of data model elements was based on the SDMX Information Model. The fourth stage is demonstration or trial. At this stage, statistical tables are compiled, a data presentation format consisting of rows and columns [19]. The table is organized based on the data model elements developed in the previous stage. The fifth stage is evaluation. At this stage, both the data model entities and the dataset will be validated using the FMR application. In the final stage, the importance of the problem, the developed data model, its usefulness, novelty, and effectiveness will be disseminated to the audience, including researchers, practitioners, and conference attendees.

3. Result and Discussion

3.1 Data Model Design



This research is classified as logical modelling, which defines concepts and relationships among elements in the domain under study [18]. The modelling in this study was based on the SDMX Information Model to examine the National Accounts domain, specifically the Gross Regional Domestic Product (GRDP). The dataset utilized was the Quarterly GRDP by Economic Activity for North Sumatra, DKI Jakarta, and East Java Provinces. The structure of the source data tables for each province is illustrated in Figure 2 (North Sumatra), Figure 3 (DKI Jakarta), and Figure 4 (East Java). The initial design stage involved the identification of the required concepts.

Kategori Lapangan Usaha	[Seri 2010] [Triwulanan]PDRB Provinsi Sumatera Utara Atas Dasar Harga Berlaku Menurut Lapangan Usaha Triwulanan (Milyar Rupiah)				
	2025				
	Triwulan I	Triwulan II	Triwulan III	Triwulan IV	Jumlah
A Pertanian, Kehutanan, dan Perikanan	75.488,89	79.392,33	-	-	-
B Pertambangan dan Penggalian	3.411,70	3.569,82	-	-	-
C Industri Pengolahan	54.764,29	56.507,16	-	-	-
D Pengadaan Listrik dan Gas	282,24	295,95	-	-	-
E Pengadaan Air, Pengelolaan Sampah, Limbah dan Daur Ulang	256,17	270,48	-	-	-
F Konstruksi	36.092,11	36.312,54	-	-	-
G Perdagangan Besar dan Eceran, dan Reparasi Mobil dan Sepeda Motor	55.376,92	56.955,59	-	-	-
H Transportasi dan Pergudangan	15.648,92	16.105,82	-	-	-
I Penyediaan Akomodasi dan Makan Minum	6.437,39	6.548,22	-	-	-
J Informasi dan Komunikasi	6.700,95	6.831,31	-	-	-
K Jasa Keuangan dan Asuransi	8.109,44	8.119,54	-	-	-
L Real Estate	12.434,72	13.371,82	-	-	-
M,N Jasa Perusahaan	3.254,84	3.296,53	-	-	-
O Administrasi Pemerintahan, Pertahanan dan Jaminan Sosial	8.918,31	9.452,87	-	-	-
P Jasa Pendidikan	5.020,67	5.182,35	-	-	-
Q Jasa Kesehatan dan Kegiatan Sosial	2.716,01	2.863,16	-	-	-
R,S,T,U Jasa Lainnya	1.576,89	1.663,29	-	-	-
PRODUK DOMESTIK REGIONAL BRUTO	296.490,46	306.738,78	-	-	-
PRODUK DOMESTIK REGIONAL BRUTO TANPA MIGAS	296.315,34	306.569,74	-	-	-

Figure 2. Table of Quarterly Gross Regional Domestic Product at Current Prices by Economic Activity in North Sumatra Province



Lapangan Usaha	PDRB Triwulanan Seri 2010 Atas Dasar Harga Berlaku Menurut Lapangan Usaha (Milyar Rupiah)				
	2025				
	Triwulan I	Triwulan II	Triwulan III	Triwulan IV	Tahunan
Pertanian, Kehutanan dan Perikanan	734,32	731,76	-	-	-
Pertambangan dan Penggalian	1.368,40	1.184,33	-	-	-
Industri Pengolahan	107.289,80	106.373,41	-	-	-
Pengadaan listrik dan gas	1.231,83	1.207,54	-	-	-
Pengadaan air, pengelolaan sampah, limbah dan daur ulang	283,78	283,48	-	-	-
Konstruksi	99.561,75	101.769,02	-	-	-
Perdagangan besar dan eceran; Resparasi mobil dan sepeda motor	170.762,74	177.174,29	-	-	-
Transportasi dan pergudangan	43.970,79	46.108,57	-	-	-
Penyediaan akomodasi dan makan minum	47.354,84	48.871,92	-	-	-
Informasi dan komunikasi	90.566,87	90.660,96	-	-	-
Jasa keuangan dan asuransi	107.831,74	112.284,30	-	-	-
Real Estate	50.524,52	50.932,32	-	-	-
Jasa perusahaan	85.699,42	87.174,39	-	-	-
Administrasi pemerintahan, pertahanan dan jaminan sosial wajib	43.874,79	42.283,10	-	-	-
Jasa pendidikan	43.145,45	44.001,26	-	-	-
Jasa kesehatan dan kegiatan sosial	21.431,81	22.128,61	-	-	-
Jasa lainnya	41.865,95	43.460,91	-	-	-
PDRB	957.498,80	976.630,19	-	-	-

Figure 3. Table of Quarterly Gross Regional Domestic Product at Current Prices by Economic Activity in DKI Jakarta Province



Lapangan Usaha 17 Kategori	[Seri 2010] PDRB Menurut Lapangan Usaha (17 Sektor) Triwulanan (Miliar Rupiah)									
	Harga Berlaku					Harga Konstan 2010				
	2025					2025				
	Triwulan I	Triwulan II	Triwulan III	Triwulan IV	Tahunan	Triwulan I	Triwulan II	Triwulan III	Triwulan IV	Tahunan
A Pertanian, Kehutanan dan Perikanan	83.714,51	97.684,85	-	-	-	43.047,25	50.162,06	-	-	-
B Pertambangan dan Penggalian	28.804,65	27.841,95	-	-	-	18.359,00	18.388,26	-	-	-
C Industri Pengolahan	257.397,81	264.103,24	-	-	-	151.994,15	155.066,14	-	-	-
D Pengadaan Listrik dan Gas	3.018,47	2.788,13	-	-	-	1.795,98	1.656,40	-	-	-
E Pengadaan Air, Pengelolaan Sampah, Limbah, dan Daur Ulang	661,69	677,80	-	-	-	472,53	482,93	-	-	-
F Konstruksi	69.589,39	72.843,70	-	-	-	44.354,06	46.350,43	-	-	-
G Perdagangan Besar dan Eceran; Reparasi dan Perawatan Mobil dan Sepeda Motor	153.196,88	154.552,88	-	-	-	93.731,17	94.343,63	-	-	-
H Transportasi dan Pergudangan	34.579,65	36.090,91	-	-	-	17.222,72	17.867,87	-	-	-
I Penyediaan Akomodasi dan Makan Minum	51.140,24	52.315,73	-	-	-	28.989,56	29.510,11	-	-	-
J Informasi dan Komunikasi	41.433,23	41.692,43	-	-	-	35.697,04	36.019,50	-	-	-
K Jasa Keuangan dan Asuransi	21.411,97	20.921,22	-	-	-	12.166,17	11.827,94	-	-	-
L Real Estat	12.328,98	12.893,09	-	-	-	8.441,21	8.668,82	-	-	-
M,N Jasa Perusahaan	6.857,15	7.417,23	-	-	-	3.974,57	4.231,13	-	-	-
O Administrasi Pemerintahan, Pertahanan dan Jaminan Sosial Wajib	17.634,74	17.921,55	-	-	-	9.675,80	9.696,00	-	-	-
P Jasa Pendidikan	20.502,21	20.673,10	-	-	-	13.395,21	13.505,96	-	-	-
Q Jasa Kesehatan dan Kegiatan Sosial	5.719,64	5.846,83	-	-	-	3.722,64	3.781,22	-	-	-
R,S,T,U Jasa Lainnya	11.308,24	13.034,57	-	-	-	7.148,65	7.895,52	-	-	-
Produk Domestik Regional Bruto	819.299,46	849.299,20	-	-	-	494.187,72	509.453,91	-	-	-

Figure 4. Table of Quarterly Gross Regional Domestic Product at Current Prices by Economic Activity in East Java Province

Analysis of the GRDP tables for the three regions above reveals format inconsistencies, necessitating data modelling. In designing the data model elements, concepts were identified to specify the required components. Concepts are labels describing the characteristics of the statistical data. The identified concepts are presented in Table 1.

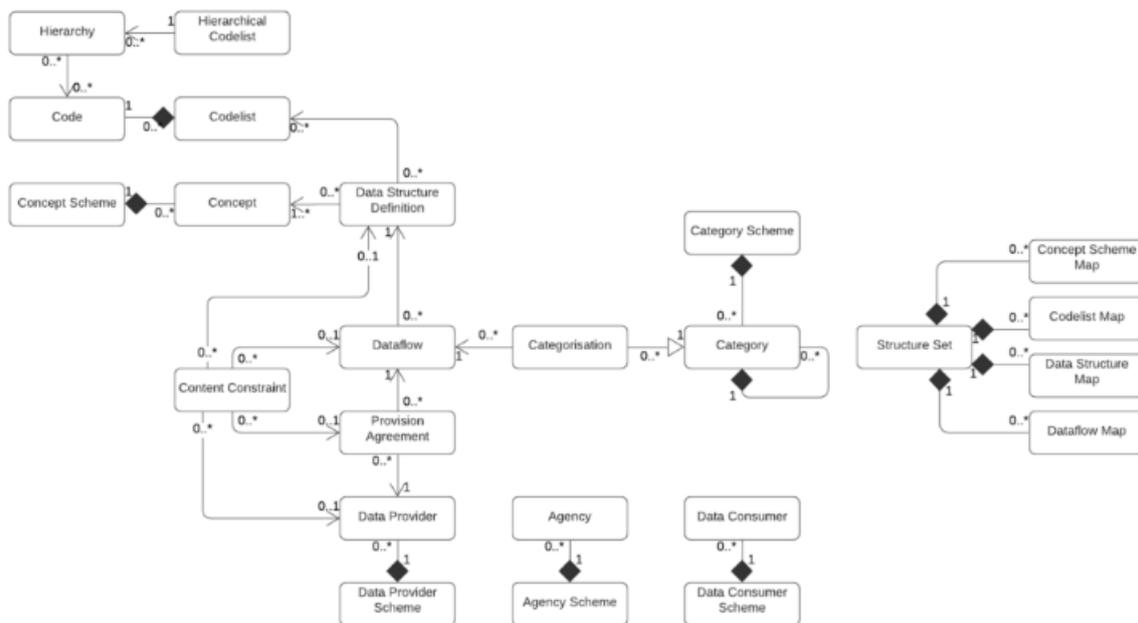
Table 1. Concept Identification

Label	Concept
Kategori Lapangan Usaha (Economic activity)	Economic activity
PDRB Triwulanan Atas Dasar Harga Berlaku Menurut Lapangan Usaha (Quarterly Gross Domestic Regional Product by Economic activity)	Indicator
Tahunan (Annual)	Frequency
Milyar (Billion)	Unit Multiplier
Rupiah	Unit Measure
2024	Time Period



3.2 Data Modelling

Data modeling is the process of visualizing how data elements are interconnected within an information system, thereby making the relationship between the data and its structure easier to comprehend [18]. In this study, the statistical indicator data model is created based on the elements of the SDMX Information Model (SDMX-IM), as illustrated in Figure 5. The research objective is to map the elements of the BPS Provincial statistical table to the SDMX-compliant data model design to establish a consistent



format.

Figure 5. SDMX Information Model

The implementation process begins with the creation of structural metadata artifacts based on the SDMX-IM. These artifacts define the authoritative organizations and the standardized vocabularies necessary for data exchange.

3.2.1 Agency

An agency is defined as an organization responsible for maintaining artifacts, while a data provider is an organization that produces or reports statistical data [20]. The data modeling stage begins with creating an agency in the Fusion Metadata Registry (FMR)



application. In this study, the agency is BPS Indonesia, acting as the owner and manager of the data model. The data provider may be the same as, or distinct from, the agency.

3.2.2 Codelist

The next step, following the determination of the agency and data provider, is the creation of a codelist, which is a set of enumeration codes representing a concept [20]. Codelists function as the standardized vocabulary for the data model. Two types of codelists are utilized: local codelists, which are created to meet the specific needs of the BPS environment and are not available in the SDMX Registry (e.g., codes for national account indicator and base period). The second type, cross-domain codelists, are owned by international agencies and can be generally applied (e.g., economic activity, unit multiplier, frequency, and unit of measure,). Tables 2 and 3 present the codelists utilized in the Quarterly GDP Table at Current Prices by Economic Activity.

Table 2. Cross-domain codelist

Agency	Codelist ID	Codelist Name
IAEG-SDGs	CL_ACTIVITY	Economic Acitivity
SDMX	CL_UNIT_MULTIPLIER	Unit Multiplier
SDMX	CL_FREQ	Frequency
UNSD	CL_UNIT_MEASURE	Unit of measure

Table 3. Local codelist

Agency	Codelist ID	Codelist Name
BPS	CL_INDICATOR_NA	National Account Indicator
BPS	CL_BASE_PERIOD	Base Period

3.2.3 Concept

Following the completion of the codelist, the next step is to define the concepts stored in the concept scheme. A concept describes the structure of multidimensional statistical tables. Concepts are represented in three ways [20]: coded (utilizing a codelist), uncoded with format (lacking codes but following a specific format, such as a time period), and uncoded free-text (without a specific format, such as notes). In this study, all concepts are stored in a single main concept scheme, as detailed in Table 4.

Table 4. Concept of Quarterly GRDP by Economic Activity



Concept ID	Name	Representation
INDICATOR_NA	Indicator of national account	CL_NA
NA_SEC	Economic activity	CL_ACTIVITY
BASE_PERIOD	Base period	CL_BASE_PERIOD
FREQ	Frequency	CL_FREQ
UNIT_MEASURE	Unit of measure	CL_UNIT_MEASURE
UNIT_MULT	Unit multiplier	CL_UNIT_MULT
TIME_PERIOD	Time period	Time dimension (uncoded)
LAST_UPDATED	Last update	Time dimension (uncoded)
OBS_VALUE	Observation value	String (uncoded)

3.2.4 Data Structure Definition

A Data Structure Definition (DSD) is a set of explanations detailing the framework or structure of the dataset to be constructed [20]. The DSD acts as the multidimensional blueprint for the GRDP data, explicitly defining all mandatory dimensions, attributes, and measures required for consistent reporting. DSD is created based on predefined concepts. DSD is created based on the statistics subdomain. For instance, if a table is categorized under the National Account subdomain, the corresponding National Account DSD (DSD_NA) is constructed.

3.2.5 Dataflow

Dataflow defines the flow of data, representing a mapping from the DSD [20]. The dataflow is created based on the existing DSD, specifically the National Account DSD (DSD_NA). The complete specification of this dataflow is detailed in Table 5.

Table 5. Dataflow of Quarterly GRDP by Economic Activity

Concept ID	Nama	Component type
INDICATOR_NA	Indicator of national account	Dimension
NA_SEC	Economic activity	Dimension
BASE_PERIOD	Base period	Dimension
FREQ	Frequency	Dimension
AREA	Province and region	Dimension
TIME_PERIOD	Time period	Dimension



UNIT_MEASURE	Unit of measure	Attribute
UNIT_MULT	Unit multiplier	Attribute
LAST_UPDATED	Last update	Attribute
OBS_VALUE	Observation value	Measure

3.2.6 Content Constraint

Table 6 lists the implemented content constraints, which function to regulate or limit the displayed codes. Content constraints are typically applied to dataflows, but can also be utilized for other entities such as provision agreements, data providers, and data structures [20].

Table 6. Constraint of Quarterly GRDP by Economic Activity

ID	Value	Description
INDICATOR_NA	GDPI_VAL	GDPI Value
NA_EXP	_Z	Not Applicable
BASE_PERIOD	2010	2010
FREQ	Q	Quarterly
AREA	_Z	Not Applicable
NA_SEC	GDP	Gross Domestic Product
	ISIC4_A	Agriculture, forestry and fishing
	ISIC4_B	Mining and quarrying
	ISIC4_C	Manufacturing
	ISIC4_D	Electricity, gas, steam and air conditioning supply
	ISIC4_E	Water supply; sewerage, waste management and remediation activities
	ISIC4_F	Construction
	ISIC4_G	Wholesale and retail trade; repair of motor vehicles and motorcycles
	ISIC4_H	Transportation and storage
	ISIC4_I	Accommodation and food service activities
	ISIC4_J	Information and communication
	ISIC4_K	Financial and insurance activities



ISIC4_L	Real estate activities
ISIC4_O	Company service
ISIC4_P	Public administration and defence; compulsory social security
ISIC4_Q	Education
ISIC4_S	Human health and social work activities
MN	Other service activities

3.2.7 Provision Agreement

The data provision agreement regulates the formal relationship between the data provider and the data type supplied [20]. Table 7 summarizes the key components of the data provision agreement utilized in this study.

Table 7. Provision agreement of Quarterly GRDP by Economic Activity

ID	Data Provider	Dataflow
DF_G_NA_PDRB_BLU_BPS	BPS	DF_G_NA_PDRB_BLU

3.2.8 Reporting Template

A reporting template is a template file used to facilitate the input of observation values into a table, though it is not part of the SDMX Information Model entity. Reporting templates are created in the Fusion Metadata Registry (FMR) application following the creation of all data model elements. An example of the reporting template generated by the FMR application is presented in Figure 6.

INDICATOR_NA BASE_PERIOD FREQ AREA38 NA_EXP UNIT_MULT UNIT_MEASURE	INDICATOR_NA BASE_PERIOD FREQUENCY AREA38 GDP EXPENDITURE UNIT MULTIPLIER UNIT MEASURE	GDPI_VAL 2010 Q _Z _Z \$ IDR	Nilai PDRB Lapangan Usaha 2010 Quarterly Not Applicable Not Applicable Billions Rupiah Indonesia	
TIME_PERIOD	2020-Q1	2020-Q2	2020-Q3	2020-Q4
LAPANGAN USAHA	GDPI_VAL:2010:Q1_Z:2020-Q1	GDPI_VAL:2010:Q2_Z:2020-Q2	GDPI_VAL:2010:Q3_Z:2020-Q3	GDPI_VAL:2010:Q4_Z:2020-Q4
Agriculture, forestry and fishing	219415.64	247947.51		260263.3
Mining and quarrying	11584.9	12280.21		13308.37
Manufacturing	182735.66	193742.57		210615.49
Electricity, gas, steam and air conditioning supply	1061.12	1097.83		1125.19
Water supply; sewerage, waste management and remediation activities	897.66	935.12		1001.63
Construction	126443.91	138780.73		148961.01
Wholesale and retail trade; repair of motor vehicles and motorcycles	181413.59	200530.11		215493.94
Transportation and storage	43388.49	53105.13		57512.42
Accommodation and food service activities	19505.88	22159.77		25330.31
Information and communication	22218	24164.92		26124.96
Financial and insurance activities	28358.82	30279.69		31826.43
Real estate activities	46639.19	46951.79		52204.45
Professional, scientific and technical activities	10072.91	11129.96		11970.07
Public administration and defence; compulsory social security	30243.09	31525.69		34433.7
Education	16880.47	18414.73		19784.2



Figure 6. Reporting template for Quarterly Regional Gross Domestic Product (PDRB) Dataset by Economic Activity

The reporting template is composed of all the data model elements. Consequently, the generated table format will be uniform and consistent across all regions. Subsequently, a validity check was performed on the table containing the observation values within the FMR application. If the dataset is valid, the standardized dataset is ready for sharing and reporting across regions and organizations.

3.3 Validation

Following the creation of all data model elements and the population of the reporting templates with the appropriate observation values, validation is performed based on the validation rules in the FMR application, as shown in Table 8.

Table 8. FMR Validation Rules

Concept ID	Description
Semantically Compliant	XML, JSON, CSV, or formatted Excel Values are unique
Duplicate Observations	Only one observation value is reported for each time period
Mandatory Components	All attributes must be reported (included)
Obs Status	Observation status is consistent with the observation value
Time Period Format	(Example: $FREQ=M$ (monthly) means the time format must be YYYY-MM)
Valid Calculations	Calculations are consistent with the equations from the validation scheme
Valid Constraint	The data is within the limits defined by the constraint
Valid Representation	Each component corresponds to the representation defined in the DSD
Valid Structure	Dimensions and attributes are consistent with DSD

Figure 7 shows that the Quarterly GDP dataset by Economic Activity has met nine validation rules.

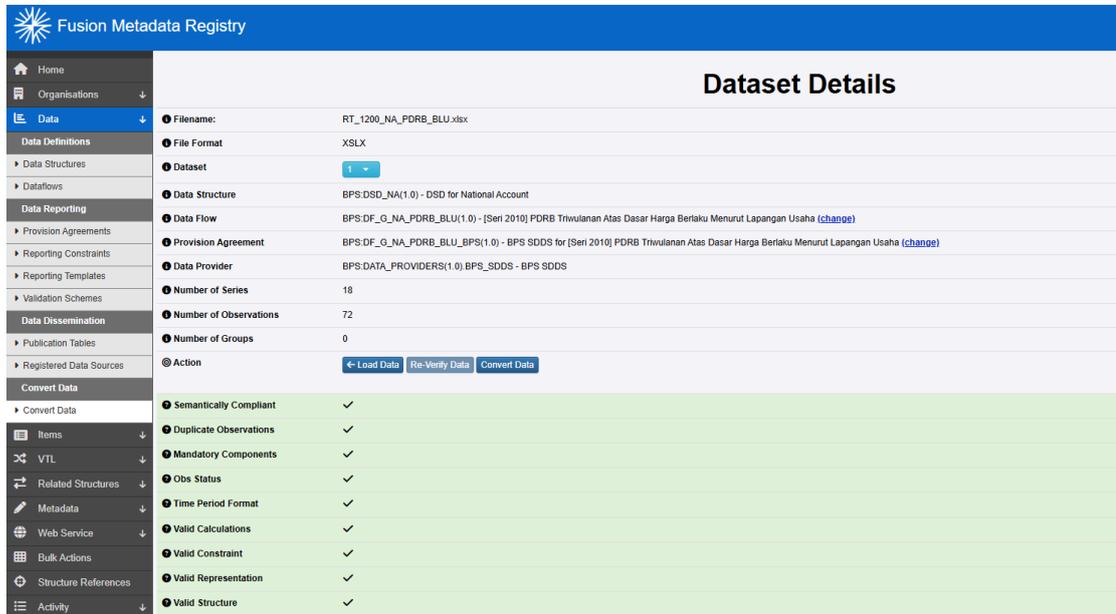


Figure 7. Validation of Quarterly GDP Dataset by Economic Activity

Following successful validation, the dataset is confirmed to be SDMX-compliant and is ready for dissemination. The final dataset, demonstrating the output in SDMX-ML format, is presented in Figure 8.



```

<?xml version="1.0" encoding="UTF-8"?>
<message:CompactData
  xmlns:ns1="urn:sdmx:org.sdmx.infomodel.registry.ProvisionAgreement=BPS:DF_G_NA_PDRB_BLU_BPS(1.0):compact"
  xmlns:common="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/common"
  xmlns:message="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="
    http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message
    https://registry.sdmx.org/schemas/v2_0/SDMXMessage.xsd
    urn:sdmx:org.sdmx.infomodel.registry.ProvisionAgreement=BPS:DF_G_NA_PDRB_BLU_BPS(1.0):compact
    http://localhost:8080/ws/public/sdmxapi/rest/schema/provisionagreement/BPS/DF_G_NA_PDRB_BLU_BPS/1.0?format=sdmx-2.0">
  <message:Header>
    <message:ID>IREF620309</message:ID>
    <message:Test>false</message:Test>
    <message:Prepared>2025-06-23T07:00:33Z</message:Prepared>
    <message:Sender id="BPS_SDDS"/>
    <message:KeyFamilyRef>DSD_NA</message:KeyFamilyRef>
    <message:KeyFamilyAgency>BPS</message:KeyFamilyAgency>
    <message:DataSetID>a3bb12501-d10a-4d92-8233-f8f16d563ce2</message:DataSetID>
    <message:DataSetAction>Information</message:DataSetAction>
    <message:Source xml:Lang="en">Fusion Metadata Registry</message:Source>
  </message:Header>
  <ns1:DataSet
    dataProviderSchemeAgencyId="BPS"
    dataFlowID="DF_G_NA_PDRB_BLU"
    dataProviderSchemeId="DATA_PROVIDERS"
    dataFlowAgencyID="BPS"
    UNIT_MULT="9"
    UNIT_MEASURE="IDR"
    dataProviderID="BPS_SDDS">
    <ns1:Series NA_SEC="GDP" INDICATOR_NA="GDPI_VAL" BASE_PERIOD="2010" FREQ="Q" AREA="ID_1200" NA_EXP="_Z">
      <ns1:Obs TIME_PERIOD="2023-Q1" OBS_VALUE="885.51"/>
      <ns1:Obs TIME_PERIOD="2023-Q2" OBS_VALUE="753.81"/>
      <ns1:Obs TIME_PERIOD="2023-Q3" OBS_VALUE="799.51"/>
      <ns1:Obs TIME_PERIOD="2023-Q4" OBS_VALUE="950.77"/>
    </ns1:Series>
    <ns1:Series NA_SEC="ISIC4_A" INDICATOR_NA="GDPI_VAL" BASE_PERIOD="2010" FREQ="Q" AREA="ID_1200" NA_EXP="_Z">
      <ns1:Obs TIME_PERIOD="2023-Q1" OBS_VALUE="9801.35"/>
      <ns1:Obs TIME_PERIOD="2023-Q2" OBS_VALUE="9481.22"/>
      <ns1:Obs TIME_PERIOD="2023-Q3" OBS_VALUE="9796.12"/>
      <ns1:Obs TIME_PERIOD="2023-Q4" OBS_VALUE="10307.39"/>
    </ns1:Series>
    <ns1:Series NA_SEC="ISIC4_B" INDICATOR_NA="GDPI_VAL" BASE_PERIOD="2010" FREQ="Q" AREA="ID_1200" NA_EXP="_Z">
      <ns1:Obs TIME_PERIOD="2023-Q1" OBS_VALUE="2148.77"/>
      <ns1:Obs TIME_PERIOD="2023-Q2" OBS_VALUE="2193.58"/>
      <ns1:Obs TIME_PERIOD="2023-Q3" OBS_VALUE="2337.78"/>
      <ns1:Obs TIME_PERIOD="2023-Q4" OBS_VALUE="2482.78"/>
    </ns1:Series>
    <ns1:Series NA_SEC="ISIC4_C" INDICATOR_NA="GDPI_VAL" BASE_PERIOD="2010" FREQ="Q" AREA="ID_1200" NA_EXP="_Z">
      <ns1:Obs TIME_PERIOD="2023-Q1" OBS_VALUE="4943.13"/>
      <ns1:Obs TIME_PERIOD="2023-Q2" OBS_VALUE="4857.68"/>
      <ns1:Obs TIME_PERIOD="2023-Q3" OBS_VALUE="4738.09"/>
      <ns1:Obs TIME_PERIOD="2023-Q4" OBS_VALUE="5361.58"/>
    </ns1:Series>
  </ns1:DataSet>
</message:CompactData>

```

Figure 8. Quarterly Regional Gross Domestic Product (PDRB) Dataset by Economic Activity in SDMX-ML format.



4. Conclusion

Based on the results and discussion, this study concludes that designing data model elements based on the SDMX Information Model facilitates the exchange and reporting of quarterly regional GDP data in a consistent and SDMX-compliant format.

5. Limitation

The current SDMX data model implementation has several limitations. In terms of scope, the developed model is restricted to Quarterly Regional Gross Domestic Product (GRDP) indicators by economic activity, and has only been validated using data from three provinces (North Sumatra, DKI Jakarta, and East Java). Furthermore, this model primarily focuses on the design of the Data Structure Definition (DSD) and has not been extended to incorporate the Metadata Structure Definition (MSD) or Structure Set. This limitation consequently restricts the model's ability to fully map relationships between various DSDs and metadata flows.

6. Future Work

This consistent model is recommended for immediate replication and implementation across all provinces in Indonesia to achieve uniformity in national GRDP data reporting. Furthermore, the designed Concept Scheme and Codelist demonstrate high extensibility, enabling them to integrate other macroeconomic indicators (e.g., inflation or unemployment rates) under the same National Accounts domain. For future research, it is suggested to integrate this statistical indicator data model with an Application Programming Interface (API) connecting directly to the SDMX Dataflow to automate the data reporting process from the source.

Acknowledgement

Directorate of Statistical Dissemination BPS Indonesia.

References



- [1] D. I. Suhada, D. Rahmadani, M. Rambe, M. A. Fattah, P. F. Hasibuan, S. Siagian, and S. Wulandari, "Efektivitas para pelaku ekonomi dalam menunjang pertumbuhan ekonomi Indonesia," *Jurnal Inovasi Penelitian*, vol. 2, no. 10, pp. 3201–3208, 2022.
- [2] K. Grueneberg, S. Calo, P. Dewan, D. Verma, and T. O’Gorman, "A policy-based approach for measuring data quality," in *Proc. 2019 IEEE Int. Conf. Big Data (Big Data)**, Dec. 2019, pp. 4025–4031.
- [3] M. I. Muslimah and L. R. Indrawati, "Pengaruh inflasi, ekspor dan penyerapan tenaga kerja terhadap pertumbuhan ekonomi Jawa Tengah," *Jurnal Paradigma Multidisipliner (JPM)*, vol. 3, no. 2, pp. 99–111, 2022.
- [4] L. Koesten, E. Simperl, E. Kacprzak, T. Blount, and J. Tennison, "Everything you always wanted to know about a dataset: Studies in data summarisation," *Int. J. Human-Computer Studies*, vol. 135, 102377, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijhcs.2019.10.004>
- [5] A. Purwanto, A. Zuiderwijk, and M. Janssen, "Citizens’ trust in open government data: a quantitative study about the effects of data quality, system quality and service quality," in *Proc. 21st Annu. Int. Conf. Digital Government Research**, Jun. 2020, pp. 310–318.
- [6] Badan Pusat Statistik, *Persentase Pengguna Data yang Menggunakan Data BPS Sebagai Dasar Perencanaan, Monitoring dan Evaluasi Pembangunan Nasional (Persen), 2021–2023*, Statistik Indonesia, 2023. [Online]. Available: <https://bps.go.id>
- [7] D. Sarihastuti, "Optimalisasi Penyelenggaraan Statistik Sektorial Sebagai Upaya Pemenuhan Data Statistik Berkualitas di Indonesia," *Jurnal Syntax Admiration*, vol. 5, no. 10, pp. 4343–4363, 2024.
- [8] L. M. Koesten, E. Kacprzak, J. F. A. Tennison, and E. Simperl, "The trials and tribulations of working with structured data: A study on information seeking behaviour," in *Proc. 2017 CHI Conf. Human Factors in Computing Systems*, 2017, pp. 1277–1289. [Online]. Available: <https://doi.org/10.1145/3025453.3025838>
- [9] Z. Zainuddin and E. A. P. Akhir, "Systematic literature review of data quality in open government data: Trend, methods, and applications," *IEEE Access*, 2024.
- [10] B. Sundgren, C. Androvitsaneas, and L. Thygesen, *Towards an SDMX User Guide: Exchange of Statistical Data and Metadata between Different Systems, National and International*. OECD Expert Group on Statistical Data and Metadata Exchange, 2006.
- [11] SDMX, *What is SDMX?*, 2023. [Online]. Available: <https://sdmx.org/>
- [12] H.-D. Biao and E. Assogba, "Statistical data governance based on the SDMX," *Department of Computer Engineering and Telecommunications, EPAC / University of Abomey-Calavi*, 2024.
- [13] UN ESCAP, *Integrated statistics: a journey worthwhile*, 2019. [Online]. Available: <https://hdl.handle.net/20.500.12870/946>
- [14] O. Bosch and M. van der Loo, "Quality assurance from an internationally standardized and generic data validation ecosystem," in *Proc. European Conf. Quality in Official Statistics (Q2022)*, Vilnius, 2022.
- [15] C. Jones, *Statistical implementation standards in the context of GSBPM*. United Nations Economic Commission for Europe, 2024. [Online]. Available: <https://coilink.org/20.500.12592/4hdq16r>
- [16] V. C. Storey and R. L. Baskerville, "The digital science field of design science research," in *Proc. 16th Int. Conf. Design Science Research in Information Systems and Technology (DESRIST 2021)*, pp. 343–355. 2021.
- [17] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manage. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007. [Online]. Available: <https://doi.org/10.2753/MIS0742-1222240302>
- [18] IBM Analytics, "What is data modeling?," IBM, 2021. [Online]. Available: <https://www.ibm.com/think/topics/data-modeling>
- [19] Z. S. Hutabarat and N. Sari, *Statistik pendidikan*. Jakarta: Eureka Media Aksara, 2022. [Online]. Available: <https://repository.penerbteureka.com/media/publications/559396-statistik-pendidikan-e21181a3.pdf>
- [20] Eurostat, Directorate B: Statistical Methodologies and Tools, Unit B-5: Statistical Information Technologies, *SDMX Self-Learning Package No. 2: Student Book – The SDMX Information Model*. 2010