

Using NSO Microdata for Analysis and Research

Thitiwat KAEW-AMDEE

Director of Statistical Standards Group,
Statistical Policy and Techniques Division,
National Statistical Office (NSO)



Agenda

01 NSO Surveys & Microdata

02 Key NSO Survey: SES

03 Analytical Techniques
using Microdata

04 Q&A

A hand is shown typing on a laptop keyboard. Overlaid on the scene is a glowing, spherical network of white nodes and connecting lines, resembling a data visualization or a globe of connections. The background is a soft, blue-toned image of the keyboard and hand.

01

**NSO Surveys
& Microdata**

Why Do We Need Surveys?

“

- To understand people, households, establishments, and societies
- To measure social and economic conditions
- To support evidence-based policymaking
- To provide reliable and (internationally) comparable data

”



<https://www.zimbabwesituation.com/wp-content/uploads/2022/05/IMG-20220421-WA0006-1-780x470-1.jpg>

<https://www.facebook.com/bsp.gu.gov/videos/2020-census-of-guam-enumerators-are-in-your-villages-verifying-addresses-remember/>

Why Surveys Are Still Essential



- Representative of **the entire population**
- Collect variables not available from other data sources
- Measure behavior, income, employment, and opinions
- Use internationally standardized methodologies



Other Data Sources

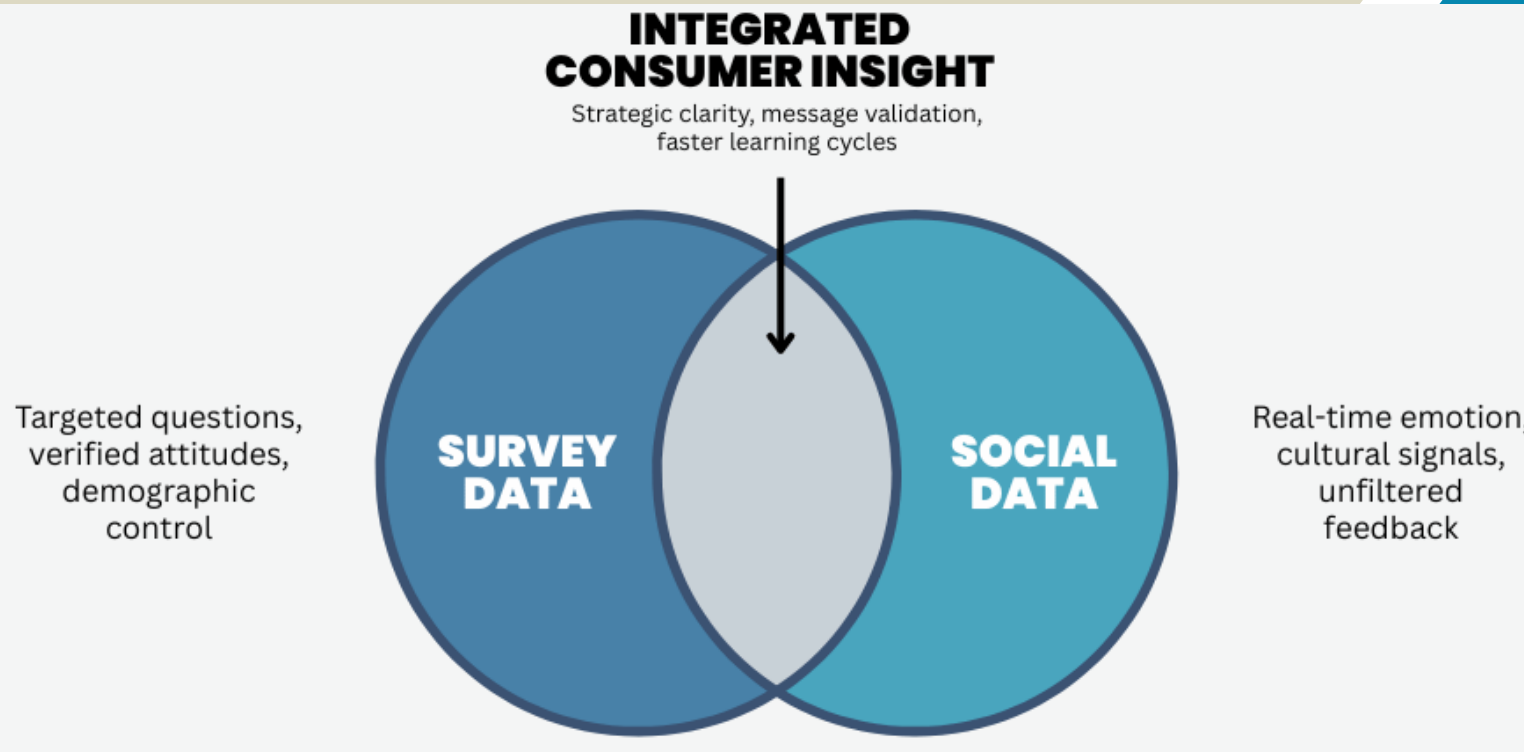
- Administrative data
 - Data from government systems, such as tax or registration data.
- Big data and digital traces
 - Data from digital platforms and online activities

Surveys in the Age of Big Data

“

- Surveys provide **ground truth or a reference point**
- Big data complements but **does not replace surveys**
- Surveys ensure transparency and quality control
- Integration of surveys and big data is the future

”



NSO Economic Censuses

Projects	Frequency	
1. Agricultural Census	Conducted every 10 years (latest round: 2013 [2556 B.E.])	
2. Business trade and Service Census	Conducted every 10 years (latest round: 2012 [2555 B.E.])	} In 2022 [2565 B.E.], integrated under the 2022 Business and Industrial Census
3. Industrial Census	Conducted every 5 years (latest round: 2017 [2560 B.E.])	

Agricultural Census

Data Collected:

1. Operational characteristics and status of agricultural holders
2. Agricultural land area, classified by land use, land tenure, and location
3. Livestock raising; cultivation of rice, rubber, perennial crops, fruit trees, forest plantations, vegetables, herbs, and ornamental plants; field crops; marine salt farming; and freshwater aquaculture
4. Agricultural machinery, tools, and transport equipment
5. Fertilizer use and pest prevention and control
6. Agricultural employees
7. Household members of agricultural holders and their work characteristics
8. Education and membership in agricultural organizations of agricultural holders
9. Agricultural income and debt of agricultural holder households

“

Coverage:

Agricultural holders

engaged in crop cultivation, livestock raising, freshwater aquaculture, and marine salt farming nationwide, both in municipal and non-municipal areas.

”

Business and Industrial Census (Listing)

Coverage:

Establishments nationwide engaged in the following economic activities:

- Trade businesses
- Service businesses
- Manufacturing
- Water supply; sewerage, waste management, and remediation activities
- Construction
- Warehousing and storage facilities
- Information and communication activities

Exclusions:

Market stalls, street vendors, hawkers, and stalls located in shopping malls.



Data Collected:

1. Name and location of establishments
2. Economic activity
3. Legal and economic form of organization
4. Number of persons engaged and number of employees
5. Sale of goods or services via the internet (e-commerce), etc.



Business and Industrial Census (Enumeration): Business trade and Service

Data Collected:

1. General information on the establishment, such as type of business, legal form of organization, economic form of organization, length of operation, registered capital, and foreign investment or shareholding, etc.
2. Persons engaged, employees, and labor compensation
3. Intermediate consumption of the establishment
4. Operating revenue of the establishment
5. Inventories of the establishment
6. Fixed assets of the establishment
7. Establishment-specific information for accommodation establishments, such as number of rooms and room rates, number of guests, number of meeting/seminar or banquet rooms, and types of facilities and services provided, etc.

“

Coverage:

Establishments engaged in wholesale trade, retail trade, and service activities.

”

Business and Industrial Census (Enumeration): Industrial activities

Data Collected:

1. Type of Economic activity
2. Size of the establishment
3. Type of products manufactured
4. Number of persons engaged and employees
5. Labor compensation
6. Production costs and operating expenses
7. Output and revenue
8. Raw materials and inventories at the beginning and end of the year
9. Fixed assets

“

Coverage:

Establishments nationwide engaged in the following economic activities:

- Manufacturing
- Water supply; sewerage, waste management, and remediation activities
- Construction
- Warehousing and storage facilities
- Information and communication activities

”

ISIC vs TSIC



ISIC (International Standard Industrial Classification)

- Developed by the United Nations
- International standard for classifying economic activities
- **Used for:**
 - National accounts
 - Economic censuses
 - Business statistics
 - International comparison
- **Purpose:** Ensure international comparability of industry statistics

Major Changes

- **Digital Economy:** clearer classification of software, platforms, cloud services, and data hosting.
- **Platform & Gig Economy:** better identification of online intermediation services.
- **Green Economy:** expanded coverage of renewable energy, waste management, recycling, and environmental services.
- **Services Sector:** more detailed breakdown of professional services, health and social care, and education services.

Explanatory Notes of the
Standard Industrial Classification of All Economic Activities, Revision 5 (ISIC Rev.5)

(Part of forthcoming UN publication)

Prepared by the Task Team on ISIC

Last update: 11 March 2024

ISIC vs TSIC



TSIC (Thailand Standard Industrial Classification)

- Developed by **Thailand's National Statistical Office (NSO) and Department of Employment.**
- National adaptation of ISIC Rev.4
- Used in:
 - Economic Censuses
 - Establishment surveys
 - Labor Force Survey
 - Business registers / Administrative data
- **Purpose:** Reflect Thailand's economic structure more accurately

DBD Data Warehouse 26201

20201

Found 34 items in juristic person information, sorted by Registration Type (A-Z)

No.	Registered No.	Juristic Person Name	Registered Type	Status	TSIC	
1	0105536126201	CHAI SOONDHORN CO., LTD.	Company Limited	Operating	69100	Legal activities
2	0105568226201	99 MILLION CONSTRUCTION CO., LTD.	Company Limited	Operating	41001	Construction of
3	0105565026201	HORIZON AVIATION CO., LTD.	Company Limited	Operating	30300	Manufacture of
4	0105564009942	T2K GROW TECH CO., LTD.	Company Limited	Operating	26201	Manufacture ar
5	0105537026201	TATIS INDUSTRIAL SUPPLY COMPANY LIMITED	Company Limited	Operating	46593	Wholesale of in
6	0105557126201	C D C MODERN DESIGN COMPANY LIMITED	Company Limited	Operating	13929	Manufacture of
7	0105553026201	DESJOYAUX TOOLS CO.,LTD.	Company Limited	Operating	43909	Other specializ
8	0105549026201	GOLDEN SHIP SUPPLY COMPANY LIMITED	Company Limited	Operating	46591	Wholesale of tr
9	0105548126201	NUCHINPORN COMPANY LIMITED	Company Limited	Operating	55101	Hotels and res
10	0105562026201	IMAGE TRADING COMPANY LIMITED	Company Limited	Operating	46441	Wholesale of ph

2620 Manufacture of computers and peripheral equipment

This class includes the manufacture and/or assembly of electronic computers, such as mainframes, desktop computers, laptops and computer servers; and computer peripheral equipment, such as storage devices and input/output devices (printers, monitors, keyboards). Computers can be analog, digital, or hybrid. Digital computers, the most common type, are devices that do all of the following: (1) store the processing program or programs and the data immediately necessary for the execution of the program, (2) can be freely programmed in accordance with the requirements of the user, (3) perform arithmetical computations specified by the user and (4) execute, without human intervention, a processing program that requires the computer to modify its execution by logical decision during the processing run. Analog computers are capable of simulating mathematical models and comprise at least analog control and programming elements.

This class includes:

- manufacture of desktop computers
- manufacture of laptop computers
- manufacture of main frame computers
- manufacture of hand-held computers (e.g. PDA)
- manufacture of magnetic disk drives, flash drives and other storage devices
- manufacture of optical (e.g. CD-RW, CD-ROM, DVD-ROM, DVD-RW) disk drives

26201 การผลิต/ประกอบคอมพิวเตอร์ ในกิจกรรมนี้ได้แก่

- การผลิตคอมพิวเตอร์ชนิดตั้งโต๊ะ
- การผลิตคอมพิวเตอร์ชนิดแล็ปท็อป (เช่น โน้ตบุ๊ก)
- การผลิตคอมพิวเตอร์เมนเฟรม
- การผลิตคอมพิวเตอร์มือถือ (เช่น พีดีเอ)
- การผลิตคอมพิวเตอร์ชนิดเซิร์ฟเวอร์
- การผลิตคอมพิวเตอร์ลูกข่ายที่ทำงานเฉพาะอย่าง
- การผลิตคอมพิวเตอร์ลูกข่าย เช่น เครื่องบริการถอนเงินอัตโนมัติ (ATM) เทอร์มินอลสำหรับเครื่องชำระเงิน ณ จุดขาย (POS)

ยกเว้น

- การผลิตส่วนประกอบอิเล็กทรอนิกส์และชิ้นส่วนอิเล็กทรอนิกส์ที่ใช้ในคอมพิวเตอร์และอุปกรณ์ต่อพ่วง ได้จัดประเภทไว้ในหมู่ย่อย 2610

26202 การผลิตอุปกรณ์จัดเก็บข้อมูล ในกิจกรรมนี้ได้แก่

- การผลิตเครื่องอ่านหรือบันทึกข้อมูลบนจานแม่เหล็ก (magnetic disk drives) แฟลชไดรฟ์ (flash drives) และอุปกรณ์จัดเก็บข้อมูลอื่นๆ
- การผลิตเครื่องอ่านหรือบันทึกข้อมูลแบบแสง (optical disk drives) เช่น เครื่องอ่านหรือ

Department of Business Development - DBD

NSO Economic Surveys

Projects	Frequency	Projects	Frequency
1. Agriculture Intercensal Survey	every 10 years	8. The Use of ICT (Household Survey)	annually
2. Construction Industry Survey	every 10 years	9. The Use of ICT (Establishment Survey)	annually
3. Commodity Flow Survey	every 5 years	10. The Quarterly Retail Survey	annually (with quarterly and annual results)
4. Non-Profit Organization Survey	every 5 years	11. The Construction Area	annually (with monthly, quarterly, and annual results)
5. Private Hospital Survey	every 5 years		
6. Business Trade and Services Survey	every 2 years		
7. Logistic Capability of Trade Survey	every 2 years		

Non-Profit Organization Survey

Coverage:

Organizations authorized by the **Department of Provincial Administration, Department of Women's Affairs and Family Development, Department of Business Development, Department of Labor Protection and Welfare, Department of Employment, Department of Religious Affairs, and the Thailand International Cooperation Agency (TICA)**, including:

- Organizations primarily engaged in social welfare services (foundations and associations)
- Funeral welfare associations
- Trade associations and chambers of commerce
- Employers' associations, labor unions, and state enterprise employee associations
- Religious organizations
- Foreign non-governmental organizations (NGOs)
- Educational organizations (private schools and private universities)
- Health organizations (private hospitals)



Data Collected:

1. General information of the organization, such as the type of main activities, legal form of organization, methods and channels of assistance (including agencies, individuals, areas, and platforms), and the use of ICT
2. Employment and labor compensation
3. Organizational revenue
4. Organizational expenditure
5. Value of assets and liabilities
6. Operational problems and constraints



Private Hospital Survey

Data Collected:

1. General information, such as type and characteristics of operations
2. Number of patients, number of beds, number of rooms, and room rates per day, ...
3. Number of personnel
4. Medical treatment, administrative, and operational expenditures
5. Revenue from medical services and other income



Coverage:

Private hospitals and medical facilities nationwide that provide inpatient (overnight) services and are established under the **Private Medical Facilities Act B.E. 2541 (1998)**.



The Use of ICT (Household Survey)

Data Collected:

1. General characteristics of household members, such as sex, age, marital status, education, and employment
2. Access to and use of information and communication technology (ICT) by household members, such as:
 1. Internet use by individuals
 2. Mobile phone use by individuals
 3. Household ownership of computers
 4. Household internet access
 5. etc.



Coverage:

Private households

with a household head and household members aged 6 years and above.



The Use of ICT (Establishment Survey)

Data Collected:

1. General information on establishments, such as name, location, and contact address; legal form of organization; economic form of organization; number of persons engaged; and number of employees, ...
2. Use of information and communication technology (ICT) by establishments, including:
 1. Use of computers and the internet
 2. Ownership of a website
 3. Online ordering of goods or services
 4. Online sales of goods or services (electronic commerce or e-Commerce)



Coverage:

All types of **establishments** located in Bangkok and all provinces nationwide, both in municipal and non-municipal areas.



NSO Social Census

Projects	Frequency
----------	-----------

1. Population and Housing Census

every 10 years

- 1960
- 1970
- 1980
- 1990
- 2000
- 2010
- 2025 (PHC-2020 postponed due to COVID19)

Population and Housing Census

Data Collected:

- **PHC2010 (Traditional full enumeration, paper-based with limited technology)**
 - Basic demographics: age, sex, marital status, nationality
 - Education level and school attendance
 - Employment status and occupation
 - Industry (TSIC based)
 - Migration (place of birth, previous residence)
 - Disability (basic categories)
 - Type of dwelling, Construction materials, Housing tenure (owner, renter)
 - Access to utilities: Electricity, Water, Toilet facilities
 - Household electrical appliances ownership
- **PHC2025 (Digital-first approach: Online self-enumeration first, then traditional enumeration)**
 - Basic demographics: age, sex, marital status, nationality
 - Education level and school attendance
 - **Gender (male, female, others, not specified/prefer not to answer)**
 - Type of dwelling, Construction materials, Housing tenure (owner, renter)



Coverage:

- All persons usually residing in Thailand (both Thai and non-Thai) in private households and in collective living quarters (e.g. dormitories, monks, military barracks)
- All types of housing units.

NEXT STEPS: Register-based census



TNSO Social Surveys

Projects	Frequency	Projects	Frequency
1. Household Socio-Economic Survey (SES)	annually	7. Health behavior of population survey	every 3 years
2. Labor Force Survey (LFS)	annually	8. Survey of the Older Persons	every 3 years
3. Informal Employment	annually	9. Disability survey	every 5 years
4. Migration Survey	annually	10. Survey of Population Change (Intercensal Survey)	every 10 years
5. Health and Welfare Survey	every 2 years	etc.	
6. Multiple Indicator Cluster Survey (MICS)	every 3 years		

Labor Force Survey (LFS)



International
Labour
Organization

To collect essential data on the size and characteristics of the labor force in the country.

Data Collected:

- Population and demographic characteristics (age, sex, education, marital status)
- Labor force status (employed, unemployed, not in the labor force)
- Employment characteristics (occupation, industry, employment status, hours worked)
- Unemployment details (job search methods, duration of unemployment)
- Underemployment and informal employment
- Income and wages (for employed persons)
- Working conditions (place of work, multiple jobs)

“

Coverage:

- People living in **private households**, both in municipal and non-municipal areas, in all provinces nationwide.

”

Multiple Indicator Cluster Survey (MICS)



To collect data for the development of indicators to monitor and evaluate the situation of children and women.

Data Collected:

- Coverage and characteristics of households and respondents
- Early childhood development, reproductive health, and maternal health
- Child development, health, and nutrition
- Learning
- Protection from violence and exploitation
- Living in a safe and clean environment
- Social equity



Coverage:

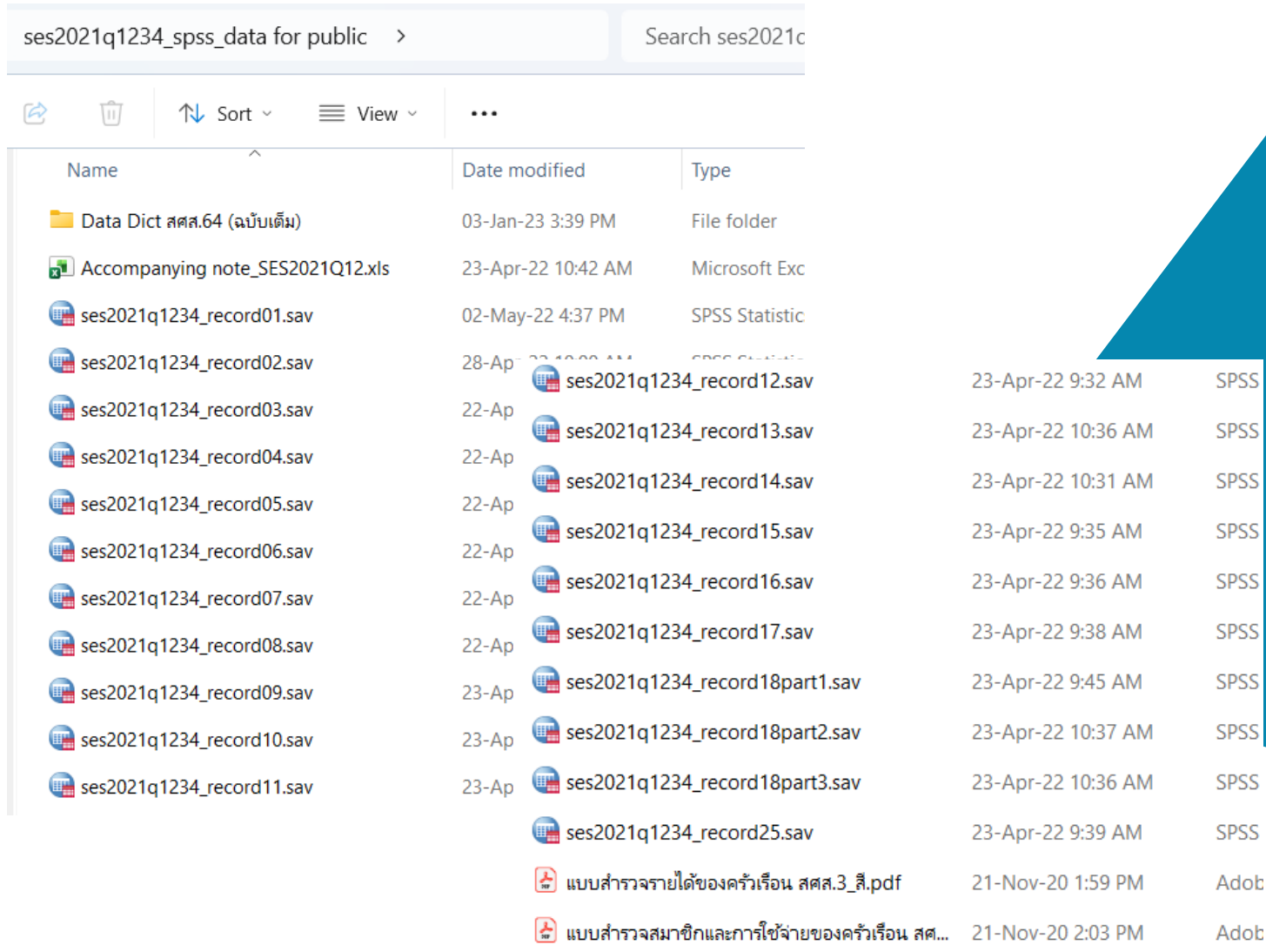
- People living in **private households**, both in municipal and non-municipal areas, in all provinces nationwide.

Questionnaires:

1. Household questionnaire
2. Women aged 15–49 years
3. Men aged 15–49 years
4. Children under 5 years of age
5. Children aged 5–14 years



What is Microdata?



Name	Date modified	Type
Data Dict สศส.64 (ฉบับเต็ม)	03-Jan-23 3:39 PM	File folder
Accompanying note_SES2021Q12.xls	23-Apr-22 10:42 AM	Microsoft Exc
ses2021q1234_record01.sav	02-May-22 4:37 PM	SPSS Statistic
ses2021q1234_record02.sav	28-Apr-22 10:00 AM	SPSS Statistic
ses2021q1234_record03.sav	22-Apr-22 9:32 AM	SPSS
ses2021q1234_record04.sav	22-Apr-22 10:36 AM	SPSS
ses2021q1234_record05.sav	22-Apr-22 10:31 AM	SPSS
ses2021q1234_record06.sav	22-Apr-22 9:35 AM	SPSS
ses2021q1234_record07.sav	22-Apr-22 9:36 AM	SPSS
ses2021q1234_record08.sav	22-Apr-22 9:38 AM	SPSS
ses2021q1234_record09.sav	23-Apr-22 9:45 AM	SPSS
ses2021q1234_record10.sav	23-Apr-22 10:37 AM	SPSS
ses2021q1234_record11.sav	23-Apr-22 10:36 AM	SPSS
ses2021q1234_record12.sav	23-Apr-22 9:32 AM	SPSS
ses2021q1234_record13.sav	23-Apr-22 10:36 AM	SPSS
ses2021q1234_record14.sav	23-Apr-22 10:31 AM	SPSS
ses2021q1234_record15.sav	23-Apr-22 9:35 AM	SPSS
ses2021q1234_record16.sav	23-Apr-22 9:36 AM	SPSS
ses2021q1234_record17.sav	23-Apr-22 9:38 AM	SPSS
ses2021q1234_record18part1.sav	23-Apr-22 9:45 AM	SPSS
ses2021q1234_record18part2.sav	23-Apr-22 10:37 AM	SPSS
ses2021q1234_record18part3.sav	23-Apr-22 10:36 AM	SPSS
ses2021q1234_record25.sav	23-Apr-22 9:39 AM	SPSS
แบบสำรวจรายได้ของครัวเรือน สศส.3_สี่.pdf	21-Nov-20 1:59 PM	Adob
แบบสำรวจสมาชิกและการใช้จ่ายของครัวเรือน สศ...	21-Nov-20 2:03 PM	Adob

“

- Microdata refers to **unit-level data** (for example; at the individual, household, or establishment level).
- **Each row** represents a real person, household, or establishment.
- **Variables** describe the characteristics of each unit.
- Microdata allows **flexible analysis** beyond published statistics.
- It is provided by the National Statistical Office (NSO) under **confidentiality rules**.

”

What is Microdata?



NEW_HH_NO	REG	CWT	AREA	A52	HM01	HM01_1	HM02	HM03	HM04	HM05	HM06	HM07
1	1	10	1	981.4104	1	1	1	2	61	1	1	0
1	1	10	1	981.4104	2	.	3	1	24	1	1	0
8	1	10	1	981.4104	1	1	1	2	61	1	1	0
8	1	10	1	981.4104	2	.	13	2	45	1	1	0
9	1	10	1	981.4104	2	.	13	1	61	1	1	0
9	1	10	1	981.4104	1	1	1	1	58	1	1	0
11	1	10	1	981.4104	2	.	3	1	45	1	1	0
11	1	10	1	981.4104	1	2	1	1	76	1	1	0
13	1	10	1	981.4104	1	1	1	2	56	1	1	0
16	1	10	1	1537.9773	1	1	1	1	75	1	1	0
16	1	10	1	1537.9773	2	.	2	2	71	1	1	0
17	1	10	1	1537.9773	1	2	1	1	62	1	1	0
17	1	10	1	1537.9773	2	.	2	2	63	1	1	0
18	1	10	1	1537.9773	1	1	1	1	61	1	1	0
18	1	10	1	1537.9773	3	.	3	1	28	1	1	0
18	1	10	1	1537.9773	2	.	2	2	58	1	1	0
19	1	10	1	1537.9773	1	1	1	2	63	1	1	0
19	1	10	1	1537.9773	2	.	3	1	37	1	1	0
19	1	10	1	1537.9773	3	.	14	2	5	1	1	0
20	1	10	1	1537.9773	1	1	1	2	52	1	1	0
20	1	10	1	1537.9773	2	.	3	1	24	1	1	0

- Microdata refers to **unit-level data** (for example; at the individual, household, or establishment level).
- **Each row** represents a real person, household, or establishment.
- **Variables** describe the characteristics of each unit.
- Microdata allows **flexible analysis** beyond published statistics.
- It is provided by the National Statistical Office (NSO) under **confidentiality rules**.



ses2021q1234_record02.sav



A hand is shown typing on a laptop keyboard. A glowing, wireframe globe composed of interconnected nodes and lines is superimposed over the keyboard. The background is a soft-focus blue. A dark semi-transparent rectangle is positioned on the right side of the image, containing the text.

02

**Key NSO Survey:
SES**

SES **Socio-economic** S u r v e y

Periodicity: biennial (1986-2004) ; annually since 2006

Questionnaire of Household Members and Expenditures

- Part 1 Household Members
- Part 2 Housing Characteristics
- Part 3 Expenditure on Goods and Services
- Part 4 Expenditure on Food, Beverages and Tobacco

Questionnaire of Household Income

- Part 1 Earning from Wages and Salaries
- Part 2 Income from Non-Farm Business
- Part 3 Income from Farm Business
- Part 4 Income Other Sources
- Part 5 Household Assets and Debt

(Conducted every two years)

1010011010010000
1010100111101110
1101101101010100
0011100101011001
0101001110101000
1010100010110101
1011011010001010
1110001010100010
1000101110101100
0100110100110100
1000010101001111
0111011011011010
1010000111001010
1100101010011101
0100010101000101
1010110110110100



Data items: (Individual)

age, sex, religion, marital status, nationality, disability, staying with parents, education level,

Occupation, Work status, Type of industry (primary & secondary occupation),

medical welfare (Universal health coverage (Gold card), Social Security, Government welfare, Private health insurance, Medical welfare from employer, ...),

benefits from the following government programs (ex. Social pension for the elderly, Welfare card, Free school lunch, (economic stimulus package (digital wallet), ...)

Questionnaire of Household Members and Expenditures

Part 1 Household Members

Part 2 Housing Characteristics

Part 3 Expenditure on Goods and Services

Part 4 Expenditure on Food, Beverages and Tobacco

Part 1 Household Members		Field name
1. Number of members in the household Person		
2. Name and number of member who usually live in this household (Sort by Code Item 3)		HM01
- Number of the respondent	<input type="text"/>	HM01_1
3. Relationship of member to household head		Relationship (HM02)
Head of household = 01	Son/ Daughter - in - law = 06	Brother/ Sister = 12
Spouse = 02	Grandchild = 07	Other relatives = 13
Son/ Daughter = 03	Father/ Mother = 08	Residents = 14
Step child = 04	Parent - in - Law = 09	Servant/ Household employees = 15
Adopted child = 05	Grandparents = 10	Employees in household businesses = 16
(Registered and not Registered)	Uncle/ Aunt = 11	
4. Sex Male = 1 Female = 2		Sex (HM03)
5. Age of last birthday (If less than 1 year, record "000" If over 120 year, record "120")		Age_01 (HM04)
6. Religion Buddhism = 1 Islam = 2 Christian = 3 Others (specify) = 4		HM05
7. Language used to communicate within household		HM06
Thai = 1 Malay/ Yawi = 2 Chinese = 3 Mon/ Burmese = 4 Cambodian/ Souy = 5 Karen = 6 Others (specify) = 7		
8. Nationality Stateless = 0 Thai Nationality = 1 Not of Thai Nationality = 2		HM07
9. Are you disable person?		HM08
No = 0	Yes, since birth	Yes, after birth
	- Physical = 1	- Physical = 4
	- Cerebral Palsey = 2	- Cerebral Palsey = 5
	- Physical and Cerebral Palsey = 3	- Physical and Cerebral Palsey = 6
10. Can you take care yourself (for regular activities) without any help ?		HM09
No = 0 Yes = 1		

Data items: (Household)

Type of Dwelling, Materials of construction,

Tenure (Own dwelling and land, Rent,...), Monthly rent or estimated rental of a house and/ or land,

Number of rooms, Number of bedroom,

Electricity, Cooking fuel, Drinking water, Water supply,

Ownership of personal car, trunk, van, bicycle, E-tan trunk

Ownership of items (Bed, Microwave, Refrigerator, TV, Air conditioner, Water heater, Home Computer, Smart Phones,...)

Questionnaire of Household Members and Expenditures

Part 1 Household Members

Part 2 Housing Characteristics

Part 3 Expenditure on Goods and Services

Part 4 Expenditure on Food, Beverages and Tobacco

Part 2 Housing Characteristics		REC	0	3	SUB	0	1
1. Type of Dwelling	HH01 <input type="checkbox"/>						
Detached house.....	= 1						
Row house/ Commercial building.....	= 2						
Town house/ twinhouse.....	= 3						
Apartment/ flat/ condominium.....	= 4						
Room inside the house.....	= 5						
Temporary shelter.....	= 6						
Others (specify)	= 7						
2. Materials of construction	HH02 <input type="checkbox"/>						
Cement or brick.....	= 1						
Wood.....	= 2						
Half concrete and wood.....	= 3						
Local non-permanent materials.....	= 4						
Re-used materials.....	= 5						
Others (specify)	= 6						
3. Tenure	HH03 <input type="checkbox"/>						
Own dwelling and land.....	= 1						
5. Monthly rent or estimated rental of a house and/ or land (If owned house or occupied rent free)	HH05 <input type="text"/>						
In the case of renting land/ public land, There is no need to estimate the land.							
6. Is part of the building used for business purposes ?	HH06 <input type="checkbox"/>						
No	= 0						
Yes	= 1						
7. Number of rooms							
Occupied by the household	HH07 <input type="checkbox"/>						
Number of bedroom	HH08 <input type="checkbox"/>						
8. Electricity in dwelling (including other sources of electricity connected)	HH09 <input type="checkbox"/>						
No	= 0	Yes	= 1				
9. Cooking fuel (Mainly used)	HH10 <input type="checkbox"/>						
No cooking.....	= 0						
Charcoal.....	= 1						

Data items: (Household)

Record "amount" (If paid) --- in **CASH**
 or "value" (If not paid) --- in **KIND**
 for each group according to **the reference period**

Questionnaire of Household Members and Expenditures

- Part 1 Household Members
- Part 2 Housing Characteristics
- Part 3 Expenditure on Goods and Services
- Part 4 Expenditure on Food, Beverages and Tobacco

A. Household expenditures on consumption

1. Expenditures on dwelling
2. Expenditure on home furnishings, miscellaneous appliances and household operations
3. Expenditures on Servants and house workers
4. Cloth, clothes and clothing material
5. Footwears
6. Personal care
7. Medical and health care
8. Travel and communication expenses
9. Education expense
10. Recreation/ Religious activity expense
11. Special ceremony expenses

		During the past 12 months		Average per month (office use only)		
		Paid in cash	Paid in kind	Field name	Paid in cash	Paid in kind
		(Amount of money)	(Estimated value)		(Amount of money)	(Estimated value)
		(1)	(2)		(3)	(4)
7.2 Medical expenses (outpatient)						
<u>Public hospital/ Health center/ Sub-District Health Promoting Hospital</u>						
	EG52					
- Public hospital	EG52_1					
- Health center/ Sub-District Health Promoting Hospital	EG52_2					
<u>Private hospital/ Clinic</u>						
	EG53					
- Private hospital	EG53_1					
- Clinic	EG53_2					
7.3 Medical services (inpatients)						
<u>Public hospital/ Public health center</u>						
	EG58					
- Public hospital	EG58_1					
- Sub-District Health Promoting Hospital	EG58_2					
<u>Private hospital/ Clinic</u>						
	EG59					
- Private hospital	EG59_1					30
- Clinic	EG59_2					

Data items: (Household)

Record "amount" (If paid) --- in **CASH**
 or "value" (If not paid) --- in **KIND**
 for each group according to **the reference period**

B. Non-consumption expenditure

ex. Tax, Money given to other persons (outside the household), Insurance premium, Lottery & gambling, Lost money, Donation, ...

C. Repaying Household Debt (both in-cash and in-kind)

Amount of repaying debt by type of borrowing purpose
 (For buying house or land, For education, For household consumption, For household non-farm business operation, For household farming operation, ...)

Questionnaire of Household Members and Expenditures

- Part 1 Household Members
- Part 2 Housing Characteristics
- Part 3 Expenditure on Goods and Services
- Part 4 Expenditure on Food, Beverages and Tobacco

Type of borrowing purpose	During the past 12 months		Field name	Average per
	in-cash <i>(amount)</i> <i>(1)</i>	in-kind <i>(value)</i> <i>(2)</i>		in-cash <i>(amount)</i> <i>(3)</i>
- For buying house or land			AE02	<input type="text"/>
- For education			AE03	<input type="text"/>
- For household consumption			AE04	<input type="text"/>
- For buying vehicle			AE04_1	<input type="text"/>
- For the purchase of goods and services (<i>Credit card</i>)			AE04_2	<input type="text"/>
- Other expenses			AE04_3	<input type="text"/>
- For household non-farm business operation			AE05	<input type="text"/>
- For household farming operation (<i>Aquaculture, raise animals, etc.</i>)			AE06	<input type="text"/>
- For other purposes (<i>e.g. guarantee debt, fine debt, etc.</i>)			AE07	<input type="text"/>

Data items: (Household)

A. Only dried food/ canned food which consumed regularly by household

(how long for consumption → average per week)

ex. Rice, Chinese sausage, Dried and salted fishes, Dried shrimp, Fermented fish, Sugar, Coffee, Tobacco, ...

B. All food items consumed by household during the week

ex. Rice noodles, Meat, Fish and seafoods, Vegetables, Ice cream, Alcoholic beverage, ...

Questionnaire of Household Members and Expenditures

- Part 1 Household Members
- Part 2 Housing Characteristics
- Part 3 Expenditure on Goods and Services
- Part 4 Expenditure on Food, Beverages and Tobacco

Items of expenditures	code	Quantity/ unit (gram etc.)	Amount of money or estimated value	How long the items are stored for household consumption ?			Amount of money of estimated value average per week
	cash paid 1		Baht	Month	Week	Day	Baht
	non-cash paid received free or own production 2						
(1)	(2)	(3)	(4)	(5)			(6)
3. Fishes and Seafood							
Dried and salted fishes							
Dried shrimp							
Salted crab							
Fermented fish							
Other processed fish and seafood (specify)							

Data items:

Part 1 for Government , State enterprise or Private employee
 - Wage or Salary, OT, Bonus, employment welfare,...

Part 2 for employer or own-account worker (Non-farm Business)

Part 3 for employer or own-account worker (farm business)
 - amount and value of livestock/product/production
 - operation costs

Part 4 ex. Pensions, Annuities or Welfare,
 Assistance from other persons outside the household,
 Interest from banks, won the lottery, ...

Questionnaire of Household Income

Part 1 Earning from Wages and Salaries

Part 2 Income from Non-Farm Business

Part 3 Income from Farm Business

Part 4 Income Other Sources

Part 5 Household Assets and Debt

(Conducted every two years)

Part 1 Earning from Wages and Salaries <i>(During the past 12 months)</i>	
Only for household member who reported the work status in Q.22 or Q.25, Part 1 SES.2, as Government , State enterprise or Private employee <i>(During the past 12 months)</i>	
	Field name
1. Names and number of household member whose work status was an employee	1W01
<i>(During the past 12 months)</i>	
2. Specify occupations which household member worked as an employee	Occupation code 1W02
	Socio-Economic class 1W03
	Industrial code 1W03_1
3. How long did you work for within the past 12 months ? <i>(If less than 1 year, record "00")</i>	1W04
4. How are wage payments made ? Per item/kg. = 1 Hourly = 2 Daily = 3 Weekly = 4 Monthly = 5 In-kind = 6 Others <i>(specify)</i> = 7	1W05
5. How much did you earn from question 4 ? <i>(In monetary term)</i>	1W06
<i>Ask every occupation</i>	
6. How many days do you work per month ?	1W07
7. How many hours do you work per day ?	1W08

Data items:

A. Value of household assets

- Value (House, Land and Building) owned by household members
- Vehicles
- Financial assets (e.g. Bank deposit, Stocks and Bond, Gold and Jewels, etc.)

B. Household Debt (Both in-cash and in-kind (items))

- amount of debt at the end of last month
by formal sector and informal sector
by objective of borrowing

Questionnaire of Household Income

Part 1 Earning from Wages and Salaries

Part 2 Income from Non-Farm Business

Part 3 Income from Farm Business

Part 4 Income Other Sources

Part 5 Household Assets and Debt

(Conducted every two years)

6. Total amount of debt at the end of last month

Total AD11 Baht

Objective of borrowing:

- For purchase/ hire - purchase of house and land
- For education
- For household consumption
- For buying vehicle
- For the purchase of goods and services (Credit card)
- For other expenses
- For business operation (*not for agriculture*)
- For agricultural operation
- Others (*specify*)

	Loan from formal sector <i>(Bank, financial institution, a cooperative society, village fund scheme, etc.)</i> <i>(Baht)</i>	Loan from informal sector <i>(person outside the household)</i> <i>(Baht)</i>
AD12	<input type="text"/>	AD18 <input type="text"/>
AD13	<input type="text"/>	AD19 <input type="text"/>
AD14		AD20
AD14_1	<input type="text"/>	AD20_1 <input type="text"/>
AD14_2	<input type="text"/>	AD20_2 <input type="text"/>
AD14_3	<input type="text"/>	AD20_3 <input type="text"/>
AD15	<input type="text"/>	AD21 <input type="text"/>
AD16	<input type="text"/>	AD22 <input type="text"/>
AD17	<input type="text"/>	AD23 <input type="text"/>

A hand is shown typing on a laptop keyboard. A glowing, spherical network of white nodes and lines is superimposed over the keyboard and hand. The background is a blurred blue-toned image of a laptop keyboard.

03

Analytical Techniques using Microdata

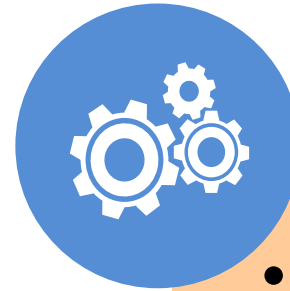
Technique 0: First Steps for using microdata

Step 1: Read the Documentation

- Survey report
- Questionnaire
- Data dictionary or codebook
- Metadata
- Survey Methodology (Understand the Design: Target population, Unit of analysis, Sampling design, Survey weights)

Step 2: Inspect the Data Structure

- Number of observations
- Number of variables
- Data types
- Key identifiers



- Microdata is powerful but complex.
- First steps are critical for correct analysis.
- Understanding before analyzing !

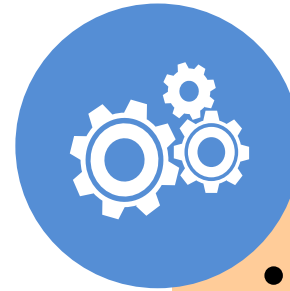
Technique 0: First Steps for using microdata

Step 3: Check Data Quality

- Missing values
 - Codes for unknown, missing values
- Outliers and extreme values
- Logical consistency
- Skip patterns

Step 4: Reproduce Official Statistics

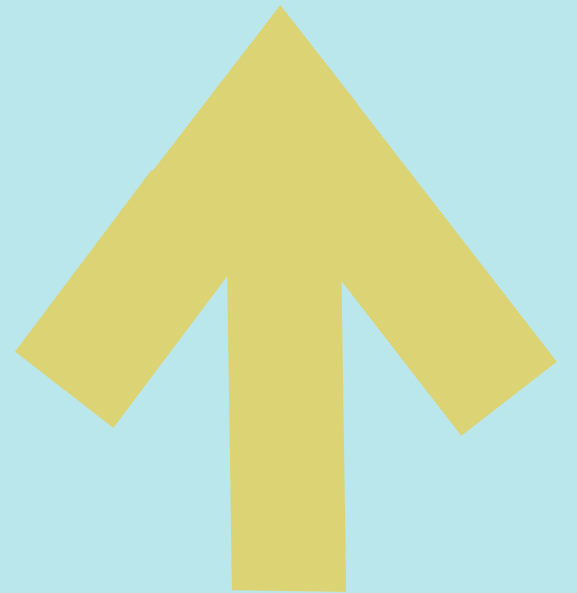
- Calculate basic indicators
- Compare with published results
- Validate understanding of the data
- Build confidence in analysis



- Microdata is powerful but complex.
- First steps are critical for correct analysis.
- Understanding before analyzing !

Technique 1: Grouping or Classification

- Classify households (by key characteristics) into meaningful groups for comparison
- Groups should reflect research objectives
- Examples:
 - socio-economic status
 - Location: Urban vs rural, Region, ...
 - Income quintiles
 - Household size
 - single-person household (one-person household) vs Multi-person household
 - With children vs without children
 - ... etc.

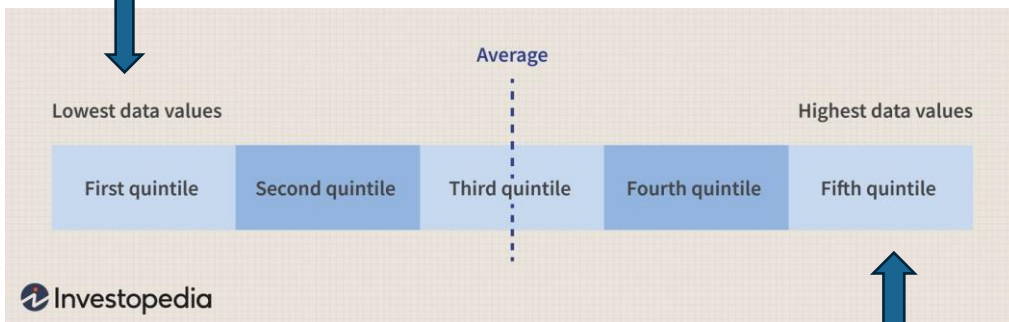


Technique 1: Grouping or Classification

SES Case Study

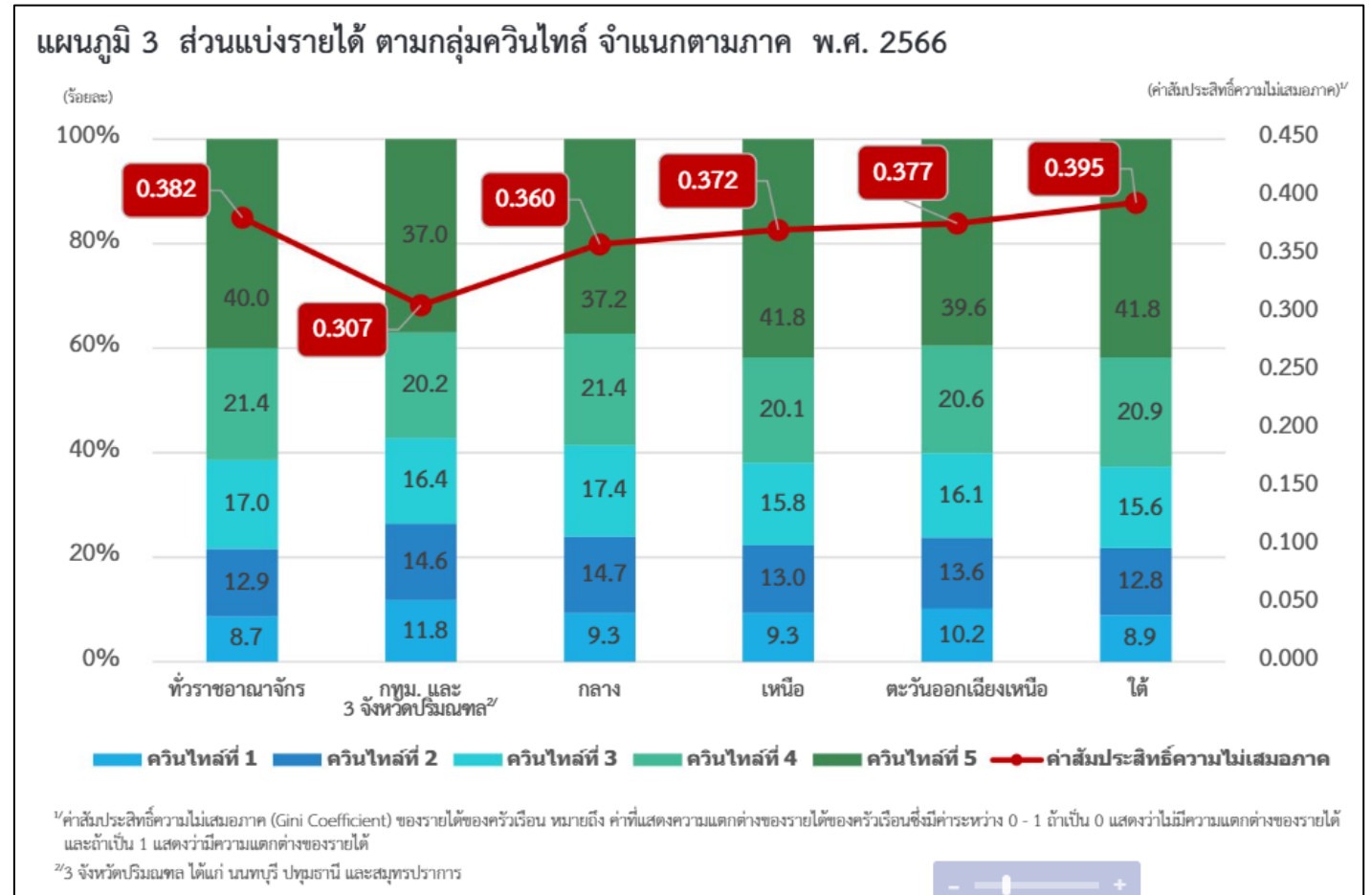
- Divide households into quintiles
- Compare income shares by quintile

the lowest-income group (Quintile 1)



the highest-income group (Quintile 5)

Distribution of household income by income quintile groups across regions of Thailand in 2023



The red line represents the Gini coefficient, which measures income inequality.

Technique 2: Setting Up Benchmarks or Criteria

- Benchmarks define reference points
- Criteria define classification rules
- Essential for comparison and interpretation
(A value alone has little meaning without a reference.)
- They allow us to classify observations into meaningful groups.
 - support targeting of social assistance programs
- Use official or well-defined standards where possible
- Document assumptions clearly; Define clear cut off points
- Examples:
 - National poverty line
 - Minimum wage

Technique 2: Setting Up Benchmarks or Criteria

SES Case Study

- Poverty line → poor vs non-poor

The World Bank's

- Extreme Poverty Line (**\$3.00 per day**)
(based on 2021 Purchasing Power Parity),
- Lower Middle Income Country Poverty Line (**\$4.20 per day**)
- Upper Middle Income Country Poverty Line (**\$8.30 per day**)

Technique 2: Setting Up Benchmarks or Criteria

SES Case Study

- Poverty line → poor vs non-poor

Thailand's poverty line (by NESDC)

- The poverty line represents **the minimum standard of living** required for basic needs.

$$\text{Poverty line} = \text{Food poverty line} + \text{Non-food poverty line}$$

Food poverty line

- Calorie/protein requirements differ across individuals depending on age and sex.
- the calorie cost and protein cost of the poorest 20 percent of the population

ตารางที่ ผ.1 ความต้องการสารอาหาร (แคลอรี) ขั้นต่ำของคนไทย (กรัมต่อวัน)

อายุ	ชาย	หญิง
น้อยกว่า 1 ปี	800	800
1-3	1,000	1,000
4-5	1,300	1,300
6-8	1,400	1,400
9-12	1,700	1,600
13-15	2,100	1,800
16-18	2,300	1,850
19-30	2,150	1,750
31-50	2,100	1,750
51-70	2,100	1,750
ตั้งแต่ 71 ปีขึ้นไป	1,750	1,550

ที่มา: กองโภชนาการ กรมอนามัย กระทรวงสาธารณสุข 2003 (<http://www.anamai.moph.go.th>)

Non-food poverty line

- Expenses beyond food that are essential for basic living.
- **9 categories** such as clothing, housing, fuel and lighting, household goods, healthcare, personal expenses, transportation, communication, and education.
- average expenditure on non-food items:
 - based on the spending patterns of people whose food spending is close to the food poverty line.
 - economy of scale.
 - depended on the area of residence (region and urban–rural area).

Technique 2: Setting Up Benchmarks or Criteria

SES Case Study

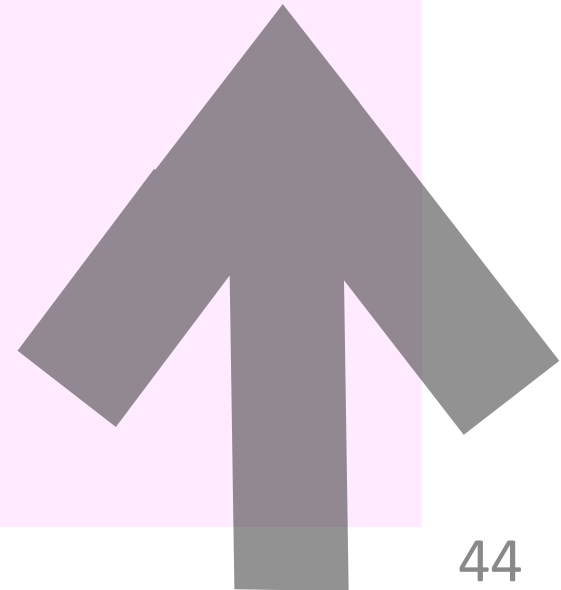
- Food poverty line
- Non-food poverty line

	Poverty line	=	Food poverty line	+	Non-food poverty line
<u>Ex.</u>	3,000	=	1,400	+	1,600

<u>Person 1</u>	3,200	=	1,500	+	1,700	
	non-poor					
<u>Person 2</u>	3,200	=	1,200	+	2,000	
	non-poor					“Hidden food poverty !”
						adequate total expenditure but insufficient food spending
<u>Person 3</u>	3,200	=	2,000	+	1,200	
	non-poor					“Hidden non-food poverty !”
						adequate total spending but insufficient non-food spending

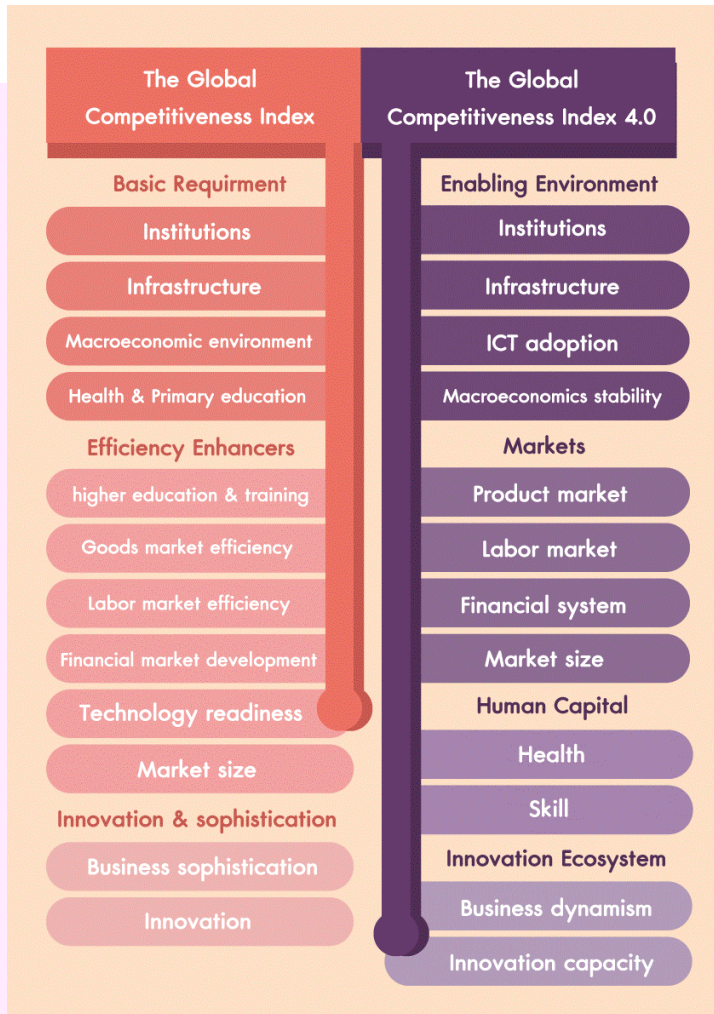
Technique 3: Composite Index Construction

- Combine multiple indicators into “a single measure”.
- Measures complex concepts such as Living standards, Household well-being, Economic vulnerability.
- Define the concept
 - > Select indicators
 - > Normalize variables (ex. Min–Max normalization)
 - > Assign weights
 - > Aggregate into an index
- Applications and Interpretation
 - Compare households or regions
 - Rank provinces or population groups
 - Use index cautiously as a summary measure
 - Always explain assumptions and limitations



Technique 3: Composite Index Construction

Ex.: The Global Competitiveness Index 4.0 (GCI 4.0) by WEF



APPENDIX C The Global Competitiveness Index 4.0 Methodology and Technical Notes

HUMAN CAPITAL
(not used in calculation)

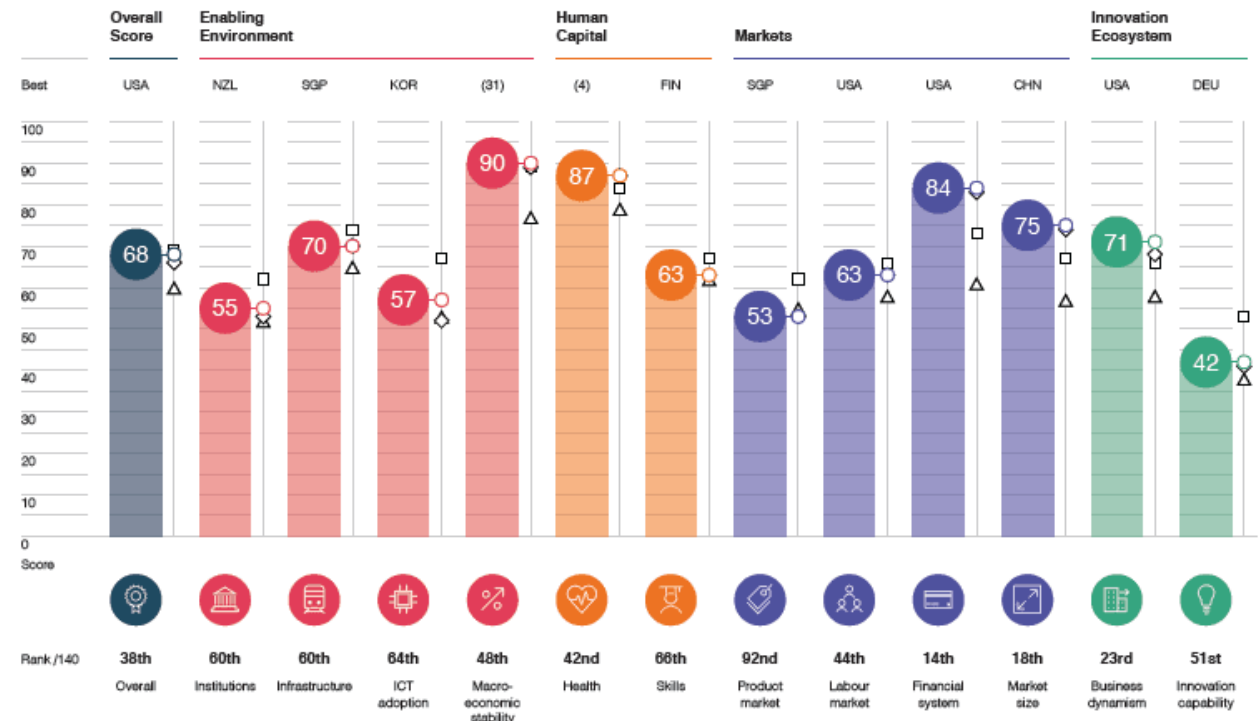
Pillar 5: Health	8.3%
5.01 Healthy life expectancy	
Pillar 6: Skills	8.3%
A. Current workforce	50%
I. Education of current workforce.....	50%
6.01 Mean years of schooling	
II. Skills of current workforce.....	50%
6.02 Extent of staff training	
6.03 Quality of vocational training	
6.04 Skillset of graduates	
6.05 Digital skills among active population	
6.06 Ease of finding skilled employees	
B. Future workforce	50%
I. Education of future workforce.....	50%
6.07 School life expectancy	
II. Skills of future workforce.....	50%
6.08 Critical thinking in teaching	
6.09 Pupil-to-teacher ratio in primary education	

Global Competitiveness Index 4.0 2018 edition

Rank in 2017 edition: 40th/135

Performance Overview Key

◇ Previous edition △ Upper middle income group average □ East Asia and Pacific average



Technique 3: Composite Index Construction

Example: Multidimensional Poverty Index (MPI) by UNDP (United Nations Development Programme)

MPI values range from 0 (no poverty) to 1 (universal poverty).

1 = deprived (does not meet the minimum standard)
0 = not deprived (meets the minimum standard)

MPI INDICATORS

Dimensions of Poverty	Indicator	Deprived if living in the household where...	Weight
Health	Nutrition	Any adult under 70 years of age or any child for whom there is nutritional information is undernourished.	1/6
	Child mortality	Any child under the age of 18 years has died in the family in the five-year period preceding the survey.	1/6
Education	Years of schooling	No household member aged 'school entrance age + six years or older has completed at least six years of schooling.	1/6
	School attendance	Any school-aged child is not attending school up to the age at which he/she would complete class eight.	1/6

Standard of living	Indicator	Description	Weight
	Cooking fuel	The household cooks with dung, wood, charcoal or coal.	1/18
	Sanitation	The household's sanitation facility is not improved (according to SDG guidelines) or it is improved but shared with other households.	1/18
	Drinking water	The household does not have access to improved drinking water (according to SDG guidelines) or improved drinking water is at least a 30-minute walk from home, round trip.	1/18
	Electricity	The household has no electricity.	1/18
	Housing	At least one of the three housing materials for roof, walls and floor are inadequate: the floor is of natural materials and/or the roof and/or walls are of natural or rudimentary materials.	1/18
	Assets	The household does not own more than one of these assets: radio, television, telephone, computer, animal cart, bicycle, motorbike or refrigerator, and does not own a car or truck.	1/18

A person is **multidimensionally poor** if their deprivation score is one-third (1/3) or higher.

Technique 3: Composite Index Construction

Example: Spatial Economic Fundamentals Index (SEFI) by the Fiscal Policy Office (FPO)

➤ 88 indicators across 6 dimensions

- infrastructure,
- public health,
- education,
- economic stability,
- human resources, and
- environment

➤ to assess the strengths and weaknesses of local economies at the district and sub-district levels.

➤ Data from

- Basic Minimum Needs (BMN) data by the Department of Community Development, Ministry of Interior. (31 Indicators)
- Satellite data from various sources (3 Indicators)
- Economic fundamentals data from NSO, NESDC, BOT, MoE, MoPH, ... (53 Indicators)
- Government program data from the Ministry of Finance (used for analysis with SEFI)

the proportion of
State Welfare Card holders
provincial GDP per capita
provincial cost of living

Correlation	SEFI
ผู้ถือบัตรต่อประชากร	-0.62
GPP per capita	0.53
ค่าครองชีพ	0.68

ความสัมพันธ์เชิงพื้นที่ของ SEFI กับ Labour Productivity

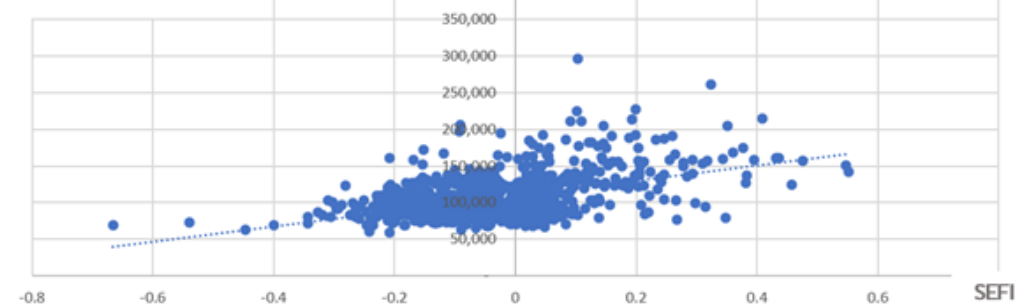
non-agricultural sectors

นอกภาคการเกษตร



agricultural sectors

ภาคการเกษตร



Technique 3: Composite Index Construction

SES Case Study: Thinking ?

If you want to construct a “Household Well-being Index” from SES Microdata, which dimensions/indicators should be included?



Core Principles for Selecting Dimensions

- Income per capita, Expenditure, Saving, Debt
- Housing quality, Access to utilities, Ownership status
- Education
- Employment, Share of working adults
- Health insurance coverage
- Poverty status
- ...
- ...

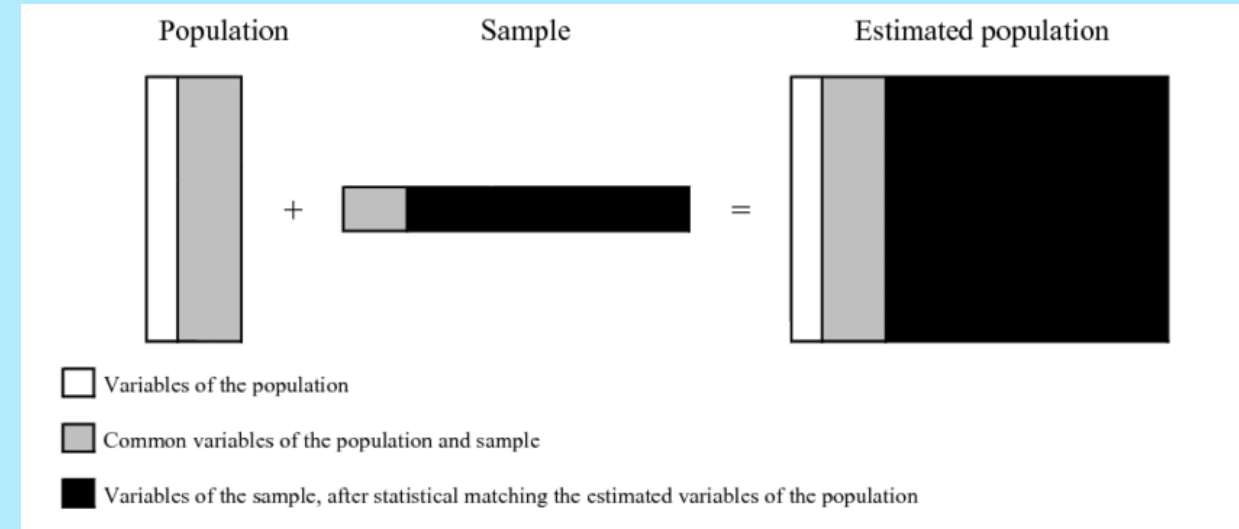
- ❖ Conceptual relevance
- ❖ Policy relevance
- ❖ Data availability
- ❖ Non-redundancy

Technique 4: Using Survey and Census Together

https://www.researchgate.net/figure/Diagram-of-statistical-matching_fig1_237262122

4.1 Statistical Matching

- A method to combine two different datasets by transferring information from a **donor dataset (survey)** to a **recipient dataset (census)** using shared characteristics, without direct record linkage.
- No common unique identifier (No direct record linkage possible)
- Uses shared variables or common variables ex. age, sex, education level, province, household size
- Creates a synthetic dataset (Need to enrich one dataset with variables from another)



“Assign values from donor to recipient”

Strengths of Statistical Matching

No need for unique identifiers
Preserves confidentiality
Flexible and scalable
Widely used in official statistics

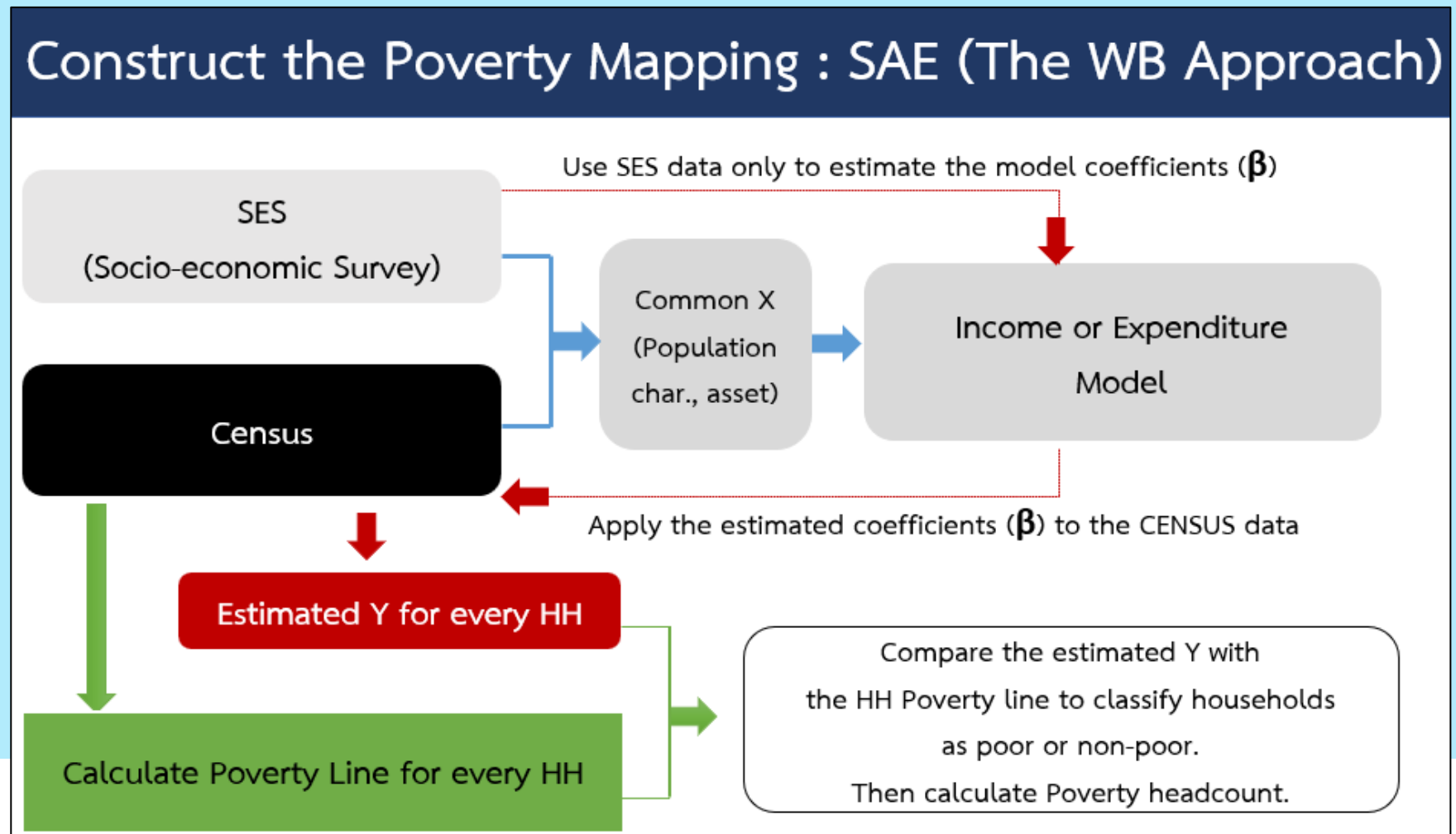
Limitations and Risks

Strong assumptions
Risk of biased estimates
Results depend on matching quality (common variables)
Not suitable for causal inference

Technique 4: Using Survey and Census Together

4.2 Small Area Estimation -> Poverty Mapping in Thailand

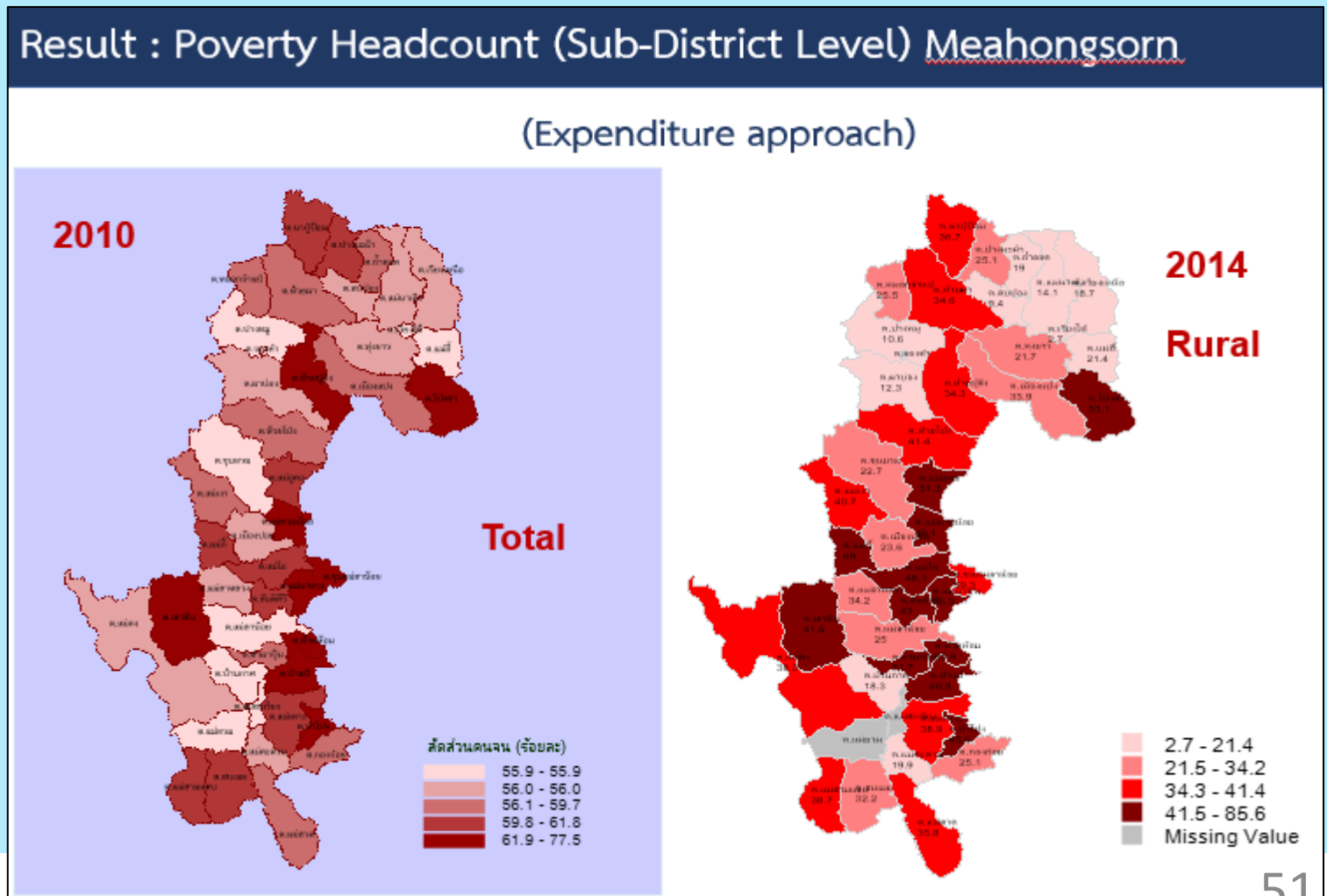
- Use Population and Housing Census 2010, and SES
- Only the variables with same definition and distribution are allowed to be used in the modeling stage.



Technique 4: Using Survey and Census Together

4.2 Small Area Estimation -> Poverty Mapping in Thailand

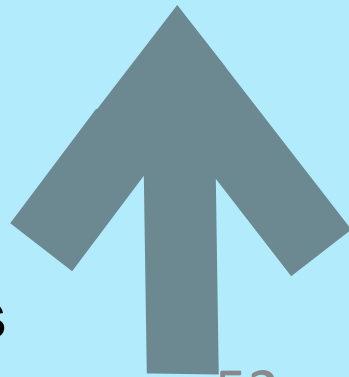
- The Poverty Mapping provides **headcount ratios, poverty gaps, poverty severity, and Gini coefficients** at the national, regional, provincial, district, and sub-district levels.
- Limitations with the decennial Census -> PHC2010 has become outdated, allowing poverty mapping only up to 2017.
- The PHC2025 does not contain enough common X's to support the analysis.



Technique 4: Using Survey and Census Together

4.2 Small Area Estimation

- **Small Area** may refer to
 - Geographical Division
 - Subpopulation ex. age-groups; basic sectors like the farmers, fishermen, senior citizens, and persons with disabilities,...
- **Components of SAE**
 - First Component: Goal or purpose of SAE
 - Second Component: Variable of interest and its form
 - Third Component: Level of disaggregation or the small area/domain
 - Fourth Component: Available data including auxiliary data sets
 - Nationwide Survey Data
 - Administrative Data
 - Satellite Data
 - Census Data
 - Internet Resources
 - Fifth Component: Quality Assessment of the Small Area Statistics



Technique 4: Using Survey and Census Together

4.2 Small Area Estimation

Examples:

- **Estimating Poverty and Wealth:**

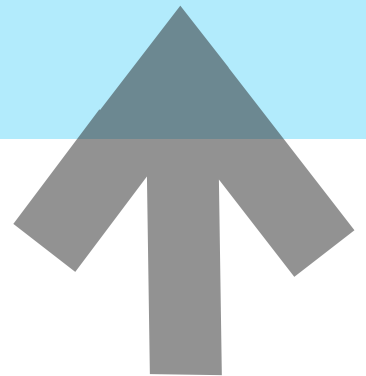
- use satellite imagery (ex. roofing material, presence of swimming pools, nighttime light intensity, road density, share of built-up land)
- Then integrated into SAE models to map poverty levels more accurately than traditional surveys alone.

- **Agricultural Yield Prediction:**

- Use geospatial data from satellites, such as vegetation indices (e.g., Normalized Difference Vegetation Index or NDVI) and weather data
- in SAE models to estimate crop yields in specific small regions.

Limitations of These Examples

- Satellite data are indirect
- Model assumptions matter
- Validation is essential !



Technique 5: Online Data with Survey Microdata

- Combine traditional surveys with online data
 - Surveys: high quality, low frequency
 - Web data: noisy, high frequency
- Enhance analysis using real-time information
- Growing trend in official statistics

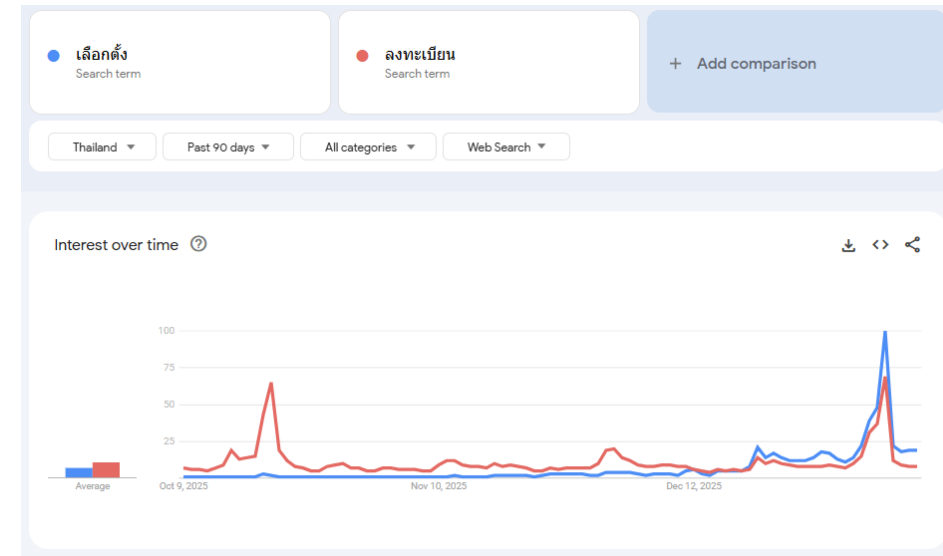
Prediction of Internet Users in Indonesia Using Google Trends Data

- Statistics Indonesia produces the percentage of internet users in Indonesia based on the Indonesian National Socioeconomic Survey (Susenas) annually.
- Using Google Trends as an alternative data source to produce monthly estimates by province.
- Studies show high correlations between Google Trends (e.g., searches for "WhatsApp", "Facebook", digital loans) and actual user numbers or related digital behaviors.
- more accurate, real-time insights than traditional surveys alone.

<https://www.isi-next.org/abstracts/submission/2786/view/>

What Is Google Trends?

- Google Trends is a free online tool that shows how often specific search terms are queried on Google
- Data reflect relative search interest over time and across locations
- Index-based data (0–100), not absolute counts
- Available by time (daily, weekly, monthly)
- Available by geography (country, province, city)
- Near real-time (timely and up-to-date)
- **Suitable for nowcasting**



Technique 5: Online Data with Survey Microdata

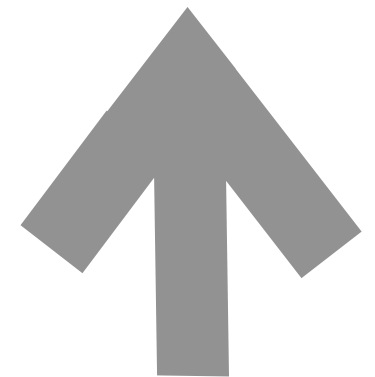
- Combine traditional surveys with online data
 - Surveys: high quality, low frequency
 - Web data: noisy, high frequency
- Enhance analysis using real-time information
- Growing trend in official statistics

What Is Web Scraping?

- Automated collection of online data
- Sources: websites, social media, forums
- Unstructured or semi-structured data

Web-scraping applications at the ONS

<p>Example 1 Data: Food price data scraped from supermarket websites</p> <p>Use: To produce timely measures of food-price inflation</p>	<p>Example 2 Data: Jobs vacancy data scraped from jobs portals</p> <p>Use: to produce timely jobs vacancy statistics and provide a richer source of labour market information</p>	<p>Example 3 Data: Data related to second/holiday homes scraped from holiday lettings and room-sharing websites</p> <p>Use: To help inform census & social survey design & estimation</p>	<p>Example 4 Data: Detailed information on contracts awarded to UK and non-UK companies scraped from procurement websites</p> <p>Use: to evaluate impact of possible changes to procurement rules</p>	<p>Example 5 Data: Data scraped from large numbers of business websites related to whether they conduct e-commerce</p> <p>Use: To produce research and statistics on the digital economy</p>
--	--	--	--	---



A hand is shown typing on a laptop keyboard. A glowing, spherical network of white nodes and lines is superimposed over the keyboard. The background is a blurred blue-toned image of the keyboard. A dark grey rectangular box is positioned on the right side of the image, containing the text '04' and 'Q&A'.

04

Q&A



THANK YOU