Sleman HDSS Manual Series

Computing Wealth Index as a Measure of Household Socio-Economic Status using Principal Component Analysis (PCA)

Authors:
Septi Kurnia Lestari
Kadarmesthan Gilang Pratama
Ratri Kusuma Wardani

UNIVERSITAS GADJAH MADA
FACULTY OF MEDICINE
PUBLIC HEALTH AND NURSING

SLEMAN HEALTH AND DEMOGRAPHIC
SURVEILLANCE SYSTEM
FMPHN UGM

Sleman HDSS Manual Series

Computing Wealth Index as a Measure of Household Socio-Economic Status using Principal Component Analysis (PCA)

Septi Kurnia Lestari, S.Gz., M.Med.Sc.PH, Ph.D.
Kadharmestan Gilang Pratama, S.Stat.
Ratri Kusuma Wardani, S.Gz.

Faculty of Medicine, Public Health and Nursing
Universitas Gadjah Mada
Yogyakarta-Indonesia
2023

Title:

Sleman HDSS Manual Series:
Computing Wealth Index as a Measure of Household Socio-Economic Status using Principal Component Analysis (PCA)

Supervisor:
dr. Ifta Choiriyyah, MSPH, Ph.D.

Authors:
Septi Kurnia Lestari, S.Gz., M.Med.Sc.PH, Ph.D.
Kadharmestan Gilang Pratama, S.Stat.
Ratri Kusuma Wardani, S.Gz.

Cover designer:
Feby Nurul Wahyuni, S.Gz.

Layout:
Septi Kurnia Lestari, S.Gz., M.Med.Sc.PH, Ph.D.

Proofreader:
Naufal Farah Azizah, S.KM.

Published by:
Faculty of Medicine, Public Health and Nursing,UGM
Farmako Street, Sekip Utara, Sleman
Special Region of Yogyakarta, Indonesia 55281
Telp./Fax. (0274) 560300
Email: iro.fk@ugm.ac.id

Publisher in collaboration with:
Sleman Health and Demographic Surveillance System (HDSS)
Radiopoetro Building 1st Floor, West Wing, Farmako Street, Sekip Utara,
Special Region of Yogyakarta 55281
Email: hdssjogja.fkkmk@ugm.ac.id | Website: https://hdss.fk.ugm.ac.id
First printing August 2023

This book can be cited as:
Lestari SK, Pratama KG, Wardani RK. Sleman HDSS Manual Series:
Computing Wealth Index as a Measure of Household Socio-Economic Status using Principal Component Analysis (PCA). Sleman (ID): Universitas Gadjah Mada; 2023.

Table of Contents

List of Figures

List of Tables

Preface

The primary objective of this book is to serve as a practical guide for researchers, students, and practitioners interested in employing the principal component analysis (PCA) method to calculate the wealth index as an indicator of household socio-economic status. Within its pages, readers will find comprehensive explanations of the step-by-step process involved in conducting PCA, encompassing data preparation, variable selection, testing the appropriateness of variables, and interpreting the outcomes. Additionally, the book includes syntax for calculating the Wealth Index using the PCA method with Sleman HDSS data on Stata software version 17.

We firmly believe that this book will prove valuable to its readers and make a significant contribution to the advancement of knowledge. We acknowledge that the book may still have certain limitations and shortcomings, thus we earnestly welcome feedback and suggestions from readers to facilitate future improvements. Furthermore, we trust that this publication will serve as a valuable resource for individuals keen on acquiring knowledge and applying PCA in their research endeavours.

In conclusion, we extend our heartfelt gratitude to all parties who have provided their assistance and support throughout the process of compiling and publishing this book. Special acknowledgment goes to the Sleman HDSS team for their diligent efforts in collecting and managing the data utilised in this work.

Sleman, August 2023

Chairperson of Sleman HDSS

Ifta Choiriyyah, MD, MSPH, Ph.D

NIP. 111198412201101201

Background

Measuring household economic status poses unique challenges, especially in Low-Middle-Income Countries like Indonesia. Valid data on economic status, such as income, expenditure, and asset ownership, is scarce in Indonesia. Various methods can be employed to measure socio-economic status, and one commonly used alternative measure is the wealth index. This index is typically calculated using principal component analysis (PCA), based on several indicators that reflect the fulfilment of basic living needs, such as asset ownership (e.g., motorbike, computer), physical characteristics (e.g., flooring type, roofing, framework), and household facilities (sanitation system, waste disposal, water source) (Vyas & Kumaranayake, 2006). Obtaining and validating this information tends to be easier compared to data on income and expenditure from respondents. The wealth index is also considered reasonably valid when compared to socio-economic status measurements based on income and household consumption (Howe et al., 2012; Poirier et al., 2020).

Various household surveys, both domestic and international, such as the Demographic and Health Survey (DHS), Multiple Indicator Cluster Survey (MICS), and Indonesia's Basic Health Research (Riskesdas) (Depkes, 2013), utilise the wealth index to determine household economic status. Sleman HDSS collected data related to the wealth index in the first wave of 2015. This data was updated in the fourth wave (2018) and the eighth wave (2022). The household wealth index in Sleman HDSS is calculated by the Sleman HDSS data team. Based on its wealth index, households in the Sleman HDSS population are categorised into five quintiles, indicating low, low-middle, middle, high-middle, and high economic groups. Indicators of household economic status are available in the Sleman HDSS dataset and can be readily used by data users. However, it is important for Sleman HDSS data users to understand the process of calculating the wealth index.

This book is designed to provide an overview of the various stages of data analysis conducted to obtain the household wealth index in Sleman HDSS. However, it does not provide an in-depth explanation of the math behind PCA. Readers interested in delving deeper into these aspects can refer to the literature cited in this book.

Understanding Principal Component Analysis (PCA)

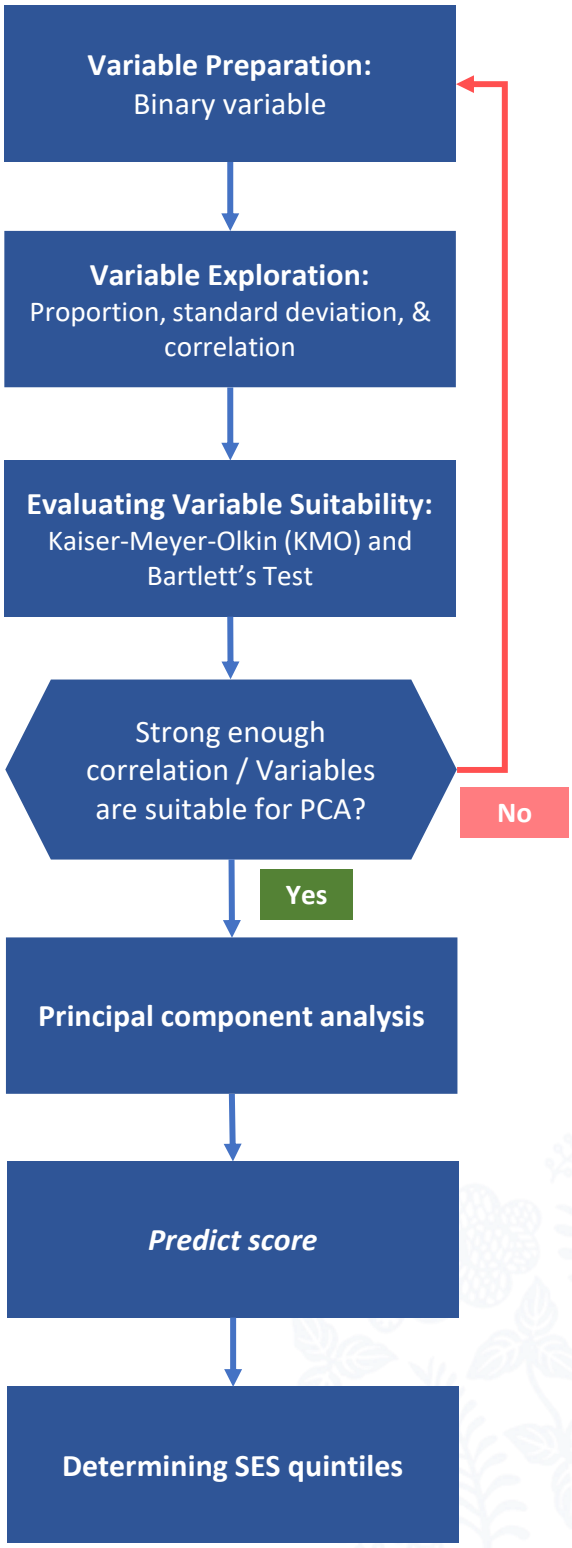

Figure 1. Stages of Wealth Index Analysis

PCA is a statistical technique used to reduce the dimensionality of a dataset without losing its valuable information (Jolliffe, 2002). It effectively summarises a large dataset or multiple variables into a smaller set of summary indices.

The process of PCA involves two stages: understanding and summarising the information. PCA identifies important information in the data by examining a statistical parameter known as variance. Variance measures the spread of data and how each value deviates from its mean. Higher variances in the original variables indicate more valuable information that can be learned.

PCA achieves data summarisation by decomposing the dataset into principal components, which are linear combinations of the original variables. These components capture the maximum amount of variance in the data. The first principal component explains the majority of the variance, while subsequent components explain less. The number of principal components generated is equal to the number of variables analysed.

The goal of PCA is to select a subset of principal components that can effectively explain most of the variance in the data. This allows for a meaningful reduction in dimensionality without significant loss of information. The explanatory power of the principal components is determined by their eigenvalues, which indicate their ability to account for variance in the data.

Calculating Wealth Index of Sleman HDSS

The Wealth Index and household economic status of Sleman HDSS are determined using the Principal Component Analysis (PCA) method (Vyas & Kumaranayake, 2006). However, before conducting PCA, several steps need to be taken (See **Figure 1**). These sequential steps highlight the importance of selecting the appropriate variables as indicators of socio-economic status. The use of relevant variables is crucial in generating a valid index. Explanation of each step in the wealth index calculation process is as follows:

1. Variable Preparation

There are 29 variables related to household characteristics, housing facilities, asset ownership, and livestock that have the potential to reflect the economic status of Sleman HDSS households (**Table 1**). The selection of these variables is based on a literature review and the researchers' knowledge on the geographic, demographic, social, and economic conditions in Sleman Regency.

The Wealth Index calculation is typically conducted for the entire population, such as a country or, in the case of Sleman HDSS, the population of Sleman Regency. However, this approach has limitations as the quality of housing facilities is generally better in urban areas compared to rural areas. Consequently, more households in urban areas tend to be categorised as having a higher socio-economic status (Vyas & Kumaranayake, 2006). The decision to calculate the wealth index for the entire population or subpopulations should be based on the research objectives and needs.

Sleman HDSS calculates the wealth index for the entire population to determine the socio-economic status of a household relative to other households in Sleman Regency. To ensure the representation of economic indicators for the rural population, variables such as livestock ownership, paddy fields, gardens, and yards are used as potential predictors of socio-economic status. These predictor variables can be continuous, binary, or categorical. PCA requires continuous or binary variables. Therefore, all these variables are transformed into binary variables. This process yields 41 binary variables. The definitions of these variables can be found in **Table 1**.

Table 1. Operational definition of potential variables of household socio-economic level

	Original variable	Original options	New variable	New options
hrt01a	Number of televisions (TV) owned		tv	0 (hrt01a=0); 1 (hrt01a>0)
hrt01b	Number of subscribed TV services owned		tvkabel	0 (hrt01b=0); 1 (hrt01b>0)
hrt01c	Number of refrigerators owned		kulkas	0 (hrt01c=0); 1 (hrt01c>0)
hrt01d	Number of air conditioners (AC) owned		AC	0 (hrt01d=0); 1 (hrt01d>0)
hrt01e	Number of washing machines owned		mesin_cuci	0 (hrt01e=0); 1 (hrt01e>0)

	Original variable	Original options	New variable	New options
hrt01f	Number of mobile phones owned		telpon_seluler	0 (hrt01f=0); 1 (hrt01f>0)
hrt01g	Number of water heaters owned		pemanas_air	0 (hrt01g=0); 1 (hrt01g>0)
hrt01h	Number of laptops/net-books/tablets owned		laptop	0 (hrt01h=0); 1 (hrt01h>0)
hrt01i	Number of personal com-puters (PC) owned		pc	0 (hrt01i=0); 1 (hrt01i>0)
hrt01j	Number of bicycles owned		sepeda	0 (hrt01j=0); 1 (hrt01j>0)
hrt01k	Number of motorcycles owned		motor	0 (hrt01k=0); 1 (hrt01k>0)
hrt01l	Number of cars owned		mobil	0 (hrt01l=0); 1 (hrt01l>0)
hrt01m	Number of trucks owned		truk	0 (hrt01m=0); 1 (hrt01m>0)
hrt01n	Number of buses owned		bus	0 (hrt01n=0); 1 (hrt01n>0)
hrt02	The primary fuel used in the kitchen:	1. Electricity, 2. 3kg Gas, 3. 12kg Gas, 4. Wood, 5. Charcoal, 6. Coal, 7. Kerosene, 8. Biogas, 9. 5kg Brightgas, 10. 12kg Brightgas, 11. 5kg Bluegas, 95 others.	gas_nonsubsidi	1 (hrt02=3 9 10 or 11); 0 (hrt02=other categories)
hrt03	Does the household own livestock?	0 No; 1 Yes	ternak	0 No; 1 Yes
hrt05	Does the household own paddy field?	0 No;1 Yes	sawah	0 No; 1 Yes
hrt09	Does the household own a garden/plot of land?	0 No; 1 Yes	kebun	0 No; 1 Yes
hrt13	Does the household have a yard?	0 No; 1 Yes	pekarangan	0 No; 1 Yes
kr01	Ownership status of the house or residential build-ing	1. Own, 2. Parents/in-laws/siblings, 3. Rent, 4. Official residence, 95 others.	rumah	1 (kr01= 1); 0 (kr01=other cat-egories)
kr03	The primary type of floor-ing	1. Ceramic, 2. Tiles, 3. Marble, 4. Planks, 5. Bamboo, 6. Cement, 7. Soil.	lantai_tanah	1 (kr03= 7); 0 (kr03=other cat-egories)
			lantai_papanbambu	1 (kr03= 4 or 5); 0 (kr03=other cate-gories)
			lantai_semen	1 (kr03= 6); 0 (kr03=other cate-gories)
			lantai_keramikubin	1 (kr03= 1 or 2); 0 (kr03=other cate-gories)
			lantai_marmer	1 (kr03= 3);

	Original variable	Original options	New variable	New options
				0 (kr03=other categories)
kr04	The primary type of wall	1. Brick, 2. Wood, 3. Planks, 4. Plywood, 5. Bamboo, 6. Woven bamboo, 7. Zinc	dinding_tembok	1 (kr04= 1); 0 (kr04=other categories)
			dinding_papan~u	1 (kr04= 2,3,4, 5, or 6); 0 (kr04=other categories)
			dinding_seng	1 (kr04= 7); 0 (kr04=other categories)
kr06	The primary roof framework type	1. Concrete, 2. Steel, 3. Lightweight steel, 4. Wood, 5. Bamboo.	rangka_betonbaja	1 (kr06= 1,2, or 3); 0 (kr06=other categories)
			rangka_kayu	1 (kr06= 4); 0 (kr06=other categories)
			rangka_bambu	1 (kr06= 5); 0 (kr06=other categories)
kr07	The primary roof type of the house	1. Ceramic roof tiles, 2. Concrete roof tiles, 3. Clay roof tiles, 4. Asphalt roof tiles, 5. Metal roof tiles, 6. Wood shingles, 7. Zinc, 8. Concrete plates, 9. Glass plates, 10. Coconut fibre, 11. Leaves, 12. Asbestos.	atap_genteng	1 (kr07= 1,2,3,4, or 5); 0 (kr07=other categories)
			atap_ijuksirapkayu	1 (kr07= 6, 10, or 11); 0 (kr07=other categories)
			atap_sengplatasbes	1 (kr07= 7, 8,9, or 12); 0 (kr07=other categories)
kr09	Electric power wattage owned/type of electricity service	1. 450 watts, 2. 900 watts, 3. 1300 watts, 4. 2200 watts, 95 others.	listrik_nonsubsidi	0 (kr09=1); 1 (kr09= other categories)
kr10	Ownership of proper sanitation facilities	1. Own, 2. Joint ownership, 3. Public, 4. None.	fasilitas_bab	1 (kr10=1); 0 (kr10= other categories)
kr11	Type of toilet used	1.Unfloored pit, 2. Floored pit, 3. Gooseneck, 4. Raised platform.	kloset	1 (kr11=3); 0 (kr11= other categories)
kr12	Final disposal site for waste	1. Septic tank, 2. Sewage treatment plant, 3. Pond/paddy field, 4. River/lake, 5. Ground pit, 6. Open land/garden, 95 others,	tpa_septictank	1 (kr12=1); 0 (kr12= other categories)
			tpa_spal	1 (kr12=2); 0 (kr12= other categories)
			tpa_kolamdll	1 (kr12= 3,4,5, or 6); 0 (kr12= other categories)

Original variable		Original options	New variable	New options
kr13	Main source of water in the household for drinking and cooking	1. Bottled water, 2. Refillable water, 3. Tap water/municipal water supply, 4. Retail tap water purchase, 5. Bored well, 6. Pumped well, 7. Covered dug well, 8. Uncovered dug well, 9. Protected spring, 10. Unprotected spring, 11. Rainwater harvesting, 12. River water, 13. Lake water, 14. Irrigation, 95 others.	air_layak	1 (kr13=1-7, 9, or 11); 0 (kr13= other categories)

2. Variable Exploration

This stage is crucial to determine which variables should be included in PCA. PCA works best when the analysed variables are correlated and have varied distributions among households. Therefore, a descriptive analysis (mean and standard deviation) is first conducted on the potential variables. In PCA, variables with high variation are given more weight. Variables that do not vary among households are assigned a weight of zero in PCA, thus having no influence in differentiating household socio-economic status (Vyas & Kumaranayake, 2006). These variables are characterised by low standard deviation (approaching zero) and very high or very low percentages.

The results of the descriptive analysis of potential variables from Sleman HDSS Wave 1, 4, and 8 are presented in **Table 2**. Since the analysed variables are binary, the mean of these variables can also be considered as proportions. In Wave 1, almost all households had at least a television (tv), and their houses were made of brick walls (dinding_tembok) or tiled roofs (atap_genteng). On the other hand, there were hardly any households that own a truck (truk) or bus. There were also few houses with wooden or bamboo floors (lantai_papanbambu), marble (lantai_marmer), or zinc walls (dinding_seng). These overly homogeneous variables would not contribute to differentiating the socio-economic status of families. Therefore, these variables were excluded in the correlation checking stage (refer to **Table 3** and **Table 4**). The same evaluation approach was applied to the variable sets of Waves 4 and 8.

Table 2. Descriptive statistics of the Wealth Index potential predictors

Variable name	Wave 1 (N=5139)		Wave 4 (N=4965)		Wave 8 (N=4967)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
tv	0.95	0.22	0.94	0.24	0.77	0.42
tvkabel	0.02	0.15	0.03	0.18	0.18	0.38
kulkas	0.60	0.49	0.71	0.45	0.82	0.39
AC	0.06	0.23	0.07	0.25	0.08	0.27
mesin_cuci	0.35	0.48	0.43	0.50	0.54	0.50
telpon_seluler	0.87	0.34	0.90	0.30	0.91	0.29
pemanas_air	0.15	0.35	0.09	0.29	0.05	0.22
laptop	0.32	0.47	0.34	0.47	0.33	0.47
pc	0.13	0.34	0.10	0.30	0.06	0.24
sepeda	0.70	0.46	0.71	0.45	0.65	0.48
motor	0.86	0.34	0.90	0.30	0.91	0.29
mobil	0.19	0.39	0.23	0.42	0.24	0.43
truk	0.00	0.07	0.01	0.08	0.01	0.08
bus	0.00	0.02	0.00	0.04	0.00	0.03
gas_nonsubsidi	0.13	0.34	0.12	0.32	0.06	0.25
ternak	0.52	0.50	0.51	0.50	0.41	0.49
sawah	0.31	0.46	0.31	0.46	0.29	0.45
kebun	0.16	0.36	0.18	0.39	0.14	0.35
pekarangan	0.39	0.49	0.38	0.48	0.22	0.41
rumah	0.77	0.42	0.73	0.45	0.68	0.47
lantai_tanah	0.02	0.13	0.02	0.13	0.02	0.12
lantai_papanbambu	0.00	0.02	0.00	0.00	0.00	0.02
lantai_semen	0.31	0.46	0.25	0.43	0.21	0.40
lantai_keramikubin	0.67	0.47	0.74	0.44	0.78	0.42
lantai_marmer	0.00	0.05	0.00	0.03	0.00	0.03
dinding_tembok	0.98	0.14	0.98	0.12	0.99	0.12
dinding_papan~u	0.02	0.14	0.02	0.12	0.01	0.11
dinding_seng	0.00	0.01	0.00	0.01	0.00	0.04
rangka_betonbaja	0.02	0.14	0.03	0.16	0.06	0.24
rangka_kayu	0.84	0.37	0.78	0.42	0.79	0.41
rangka_bambu	0.14	0.35	0.20	0.40	0.15	0.36
atap_genteng	0.97	0.18	0.95	0.22	0.94	0.24
atap_ijuksirapkayu	0.00	0.02	0.00	0.02	0.00	0.02
atap_sengplatasbes	0.03	0.18	0.05	0.22	0.06	0.24
listrik_nonsubsidi	0.56	0.50	0.59	0.49	0.62	0.49
fasilitas_bab	0.91	0.28	0.93	0.25	0.95	0.22
kloset	0.92	0.27	0.96	0.20	0.98	0.14
tpa_septictank	0.88	0.32	0.88	0.32	0.87	0.34
tpa_spal	0.02	0.14	0.03	0.18	0.06	0.24
tpa_kolamdll	0.09	0.29	0.08	0.27	0.07	0.26
air_layak	0.85	0.36	0.94	0.23	0.96	0.20

Note: The red highlight: variables that are excessively homogeneous (with a standard deviation approaching zero or a high mean), therefore they will not be included in the subsequent analysis stage.

Table 3. Correlation coefficient strength interpretation

Correlation coeficient	Interpretation
0	None
-0.3 < r < 0.0 dan 0.0 < r < 0.3	Poor
-0.6 < r ≤ -0.3 dan 0.3 ≤ r < 0.6	Fair
-0.8 < r ≤ -0.6 dan 0.6 ≤ r < 0.8	Moderate
-1.0 < r ≤ -0.8 dan 0.8 ≤ r < 1.0	Very strong
-1 atau 1	Perfect

Next, a correlation test called tetrachoric correlation test (Ariawan, 2006) was conducted to examine the correlation between variables. It is essential for the predictor variables used in PCA to exhibit strong correlations. Therefore, predictor variables with low correlation values to other predictor variables will be excluded, and another round of correlation testing will be conducted. This process ensures that the remaining predictor variables are sufficiently correlated with each other. The determination of correlation strength is based on the guidelines provided by Akoglu (2018).

The results of the tetrachoric correlation test for the sets of variables in Waves 1, 4, and 8 are presented in **Table 5**, **Table 6**, dan **Table 7**. These tables display the correlation coefficients of all possible pairs formed from the analysed set of variables. Across all Waves, variables such as livestock ownership, paddy fields, gardens, yards, and houses showed weak correlations with other variables and were thus excluded from the analysis. On the other hand, car ownership displayed a significant correlation with other asset ownerships, such as refrigerators, washing machines, and non-subsidised gas usage. Consequently, these variables were retained for further analysis. The list of variables utilised in the subsequent stage can be found in **Table 4**.

Table 4. List of variables selected for correlation analysis and PCA

Wave 1 (N=5139)		Wave 4 (N=4965)		Wave 8 (N=4967)	
Correlation test	**PCA**	**Correlation test**	**PCA**	**Correlation test**	**PCA**
		tv		tv	
tvkabel	tvkabel	tvkabel		tvkabel	
kulkas	kulkas	kulkas	kulkas	kulkas	kulkas
AC	AC	AC	AC	AC	AC
mesin_cuci	mesin_cuci	mesin_cuci	mesin_cuci	mesin_cuci	mesin_cuci
telpon_seluler	telpon_seluler	telpon_seluler	telpon_seluler	telpon_seluler	telpon_seluler
pemanas_air	pemanas_air	pemanas_air		pemanas_air	pemanas_air
laptop	laptop	laptop	laptop	laptop	laptop
pc	pc	pc	pc	pc	pc
sepeda	motor	sepeda		sepeda	
motor	mobil	motor	motor	motor	
mobil		mobil	mobil	mobil	mobil
gas_nonsubsidi	gas_nonsubsidi	gas_nonsubsidi	gas_nonsubsidi	gas_nonsubsidi	gas_nonsubsidi
ternak		ternak		ternak	
sawah		sawah		sawah	
kebun		kebun		kebun	
pekarangan		pekarangan		pekarangan	
rumah		rumah		rumah	
lantai_tanah	lantai_tanah	lantai_tanah	lantai_tanah	lantai_tanah	lantai_tanah
lantai_semen		lantai_semen		lantai_semen	
lantai_keramik~n	lantai_keramik~n	lantai_keramik~n	lantai_keramik~n	lantai_keramik~n	lantai_keramik~n
dinding_papan~u	dinding_papan~u	dinding_papan~u		dinding_papan~u	
rangka_betonbaja	rangka_betonbaja	rangka_betonbaja		rangka_betonbaja	
rangka_kayu		rangka_kayu		rangka_kayu	
rangka_bambu	rangka_bambu	rangka_bambu	rangka_bambu	rangka_bambu	
				atap_genteng	
atap_seng~s		atap_seng~s		atap_seng~s	
listrik_nonsubsidi	listrik_nonsubsidi	listrik_nonsubsidi	listrik_nonsubsidi	listrik_nonsubsidi	listrik_nonsubsidi
fasilitas_bab	fasilitas_bab	fasilitas_bab	fasilitas_bab		
kloset					
tpa_septictank		tpa_septictank		tpa_septictank	
tpa_spal		tpa_spal		tpa_spal	
tpa_kolamdll		tpa_kolamdll		tpa_kolamdll	
air_layak		air_layak			

Notes: Red highlight indicate variables that do not vary and are excluded from the tetrachoric analysis.

Table 5. Correlation coefficient between potential variables in Wave 1

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1 tvkabel	1.00	0.48	0.58	0.47	0.39	0.29	0.44	0.41	0.04	0.21	0.56	0.59	-0.19	-0.18	-0.12	-0.07	0.12	-0.52	-0.28	0.46	-0.15	0.29	0.09	-0.29	0.07	0.48	0.32	0.18	0.37	-0.52	-0.01	0.37
2 kulkas	0.48	1.00	0.58	0.77	0.58	0.42	0.70	0.58	0.16	0.56	0.68	0.65	-0.16	-0.05	0.01	-0.04	0.03	-0.37	-0.49	0.56	-0.29	0.24	0.27	-0.37	0.04	0.63	0.54	0.41	0.30	0.02	-0.34	0.25
3 AC	0.58	0.58	1.00	0.57	0.42	0.39	0.47	0.45	0.06	0.22	0.60	0.62	-0.23	-0.18	-0.03	-0.05	0.19	-0.54	-0.41	0.57	-0.67	0.44	0.18	-0.40	0.03	0.53	0.72	0.30	0.20	0.08	-0.32	0.35
4 mesin_cuci	0.47	0.77	0.57	1.00	0.50	0.42	0.64	0.54	0.13	0.48	0.68	0.61	-0.17	-0.09	0.04	-0.01	0.05	-0.34	-0.48	0.52	-0.29	0.30	0.21	-0.38	0.03	0.63	0.49	0.33	0.30	0.01	-0.34	0.25
5 telpon_seluler	0.39	0.58	0.42	0.50	1.00	0.45	0.64	0.60	0.18	0.80	0.55	0.38	0.01	-0.14	-0.02	-0.04	-0.24	-0.38	-0.29	0.37	-0.28	0.20	0.19	-0.24	0.10	0.38	0.35	0.28	0.18	0.03	-0.23	0.14
6 pemanas_air	0.29	0.42	0.39	0.42	0.45	1.00	0.41	0.34	0.07	0.36	0.39	0.36	-0.09	-0.05	0.04	0.02	0.00	-0.31	-0.24	0.31	-0.21	0.21	0.10	-0.24	-0.11	0.33	0.30	0.25	0.06	0.09	-0.17	0.10
7 laptop	0.44	0.70	0.47	0.64	0.64	0.41	1.00	0.58	0.18	0.64	0.66	0.56	-0.11	-0.03	0.07	0.05	0.08	-0.35	-0.40	0.46	-0.19	0.23	0.23	-0.36	0.01	0.55	0.45	0.32	0.22	0.07	-0.32	0.25
8 pc	0.41	0.58	0.45	0.54	0.60	0.34	0.58	1.00	0.21	0.56	0.52	0.50	-0.08	0.00	0.12	0.03	0.08	-0.59	-0.25	0.52	-0.23	0.25	0.18	-0.26	0.03	0.48	0.34	0.26	0.16	0.07	-0.25	0.21
9 sepeda	0.04	0.16	0.06	0.13	0.18	0.07	0.18	0.21	1.00	0.22	0.16	0.08	0.09	0.08	-0.04	0.04	0.00	-0.10	-0.07	0.10	-0.08	0.03	0.05	-0.05	-0.05	0.05	0.10	-0.01	0.04	0.01	-0.04	-0.06
10 motor	0.21	0.56	0.22	0.48	0.80	0.36	0.64	0.56	0.22	1.00	0.46	0.24	0.07	0.04	0.05	0.00	-0.11	-0.33	-0.24	0.41	-0.26	0.20	0.20	-0.26	-0.02	0.34	0.37	0.32	0.22	0.04	-0.23	0.14
11 mobil	0.56	0.68	0.60	0.68	0.55	0.39	0.66	0.52	0.16	0.46	1.00	0.65	-0.14	-0.01	0.14	0.00	0.21	-0.34	-0.48	0.49	-0.29	0.33	0.18	-0.41	-0.01	0.57	0.47	0.29	0.23	0.00	-0.30	0.32
12 gas_nonsubsidi	0.59	0.65	0.62	0.61	0.38	0.36	0.56	0.50	0.08	0.24	0.65	1.00	-0.21	-0.10	0.02	0.00	0.17	-0.61	-0.34	0.61	-0.24	0.36	0.17	-0.40	0.00	0.54	0.42	0.25	0.15	0.07	-0.28	0.24
13 ternak	-0.19	-0.16	-0.23	-0.17	0.01	-0.09	-0.11	-0.08	0.09	0.07	-0.14	-0.21	1.00	0.36	0.27	0.27	0.08	0.06	0.13	-0.11	0.01	-0.17	0.00	0.10	-0.16	-0.22	0.01	-0.02	-0.06	-0.03	0.08	-0.13
14 sawah	-0.18	-0.05	-0.18	-0.09	-0.14	-0.05	-0.03	0.00	0.08	0.04	-0.01	-0.10	0.36	1.00	0.36	0.30	0.26	0.05	0.07	-0.06	-0.11	-0.09	0.03	0.03	-0.33	-0.18	0.04	-0.01	-0.04	-0.13	0.15	-0.05
15 kebun	-0.12	0.01	-0.03	0.04	-0.02	0.04	0.07	0.12	-0.04	0.05	0.14	0.02	0.27	0.36	1.00	0.15	0.17	0.00	0.03	-0.05	-0.15	0.01	0.06	-0.08	-0.15	-0.10	0.10	0.07	0.01	-0.04	0.04	0.11
16 pekarangan	-0.07	-0.04	-0.05	-0.01	-0.04	0.02	0.05	0.03	0.04	0.00	0.00	0.00	0.27	0.30	0.15	1.00	0.19	-0.03	0.05	-0.04	-0.15	-0.06	0.01	0.05	-0.28	-0.05	0.11	0.12	0.04	-0.06	-0.01	-0.07
17 rumah	0.12	0.03	0.19	0.05	-0.24	0.00	0.08	0.08	0.00	-0.11	0.21	0.17	0.08	0.26	0.17	0.19	1.00	-0.05	-0.09	0.06	-0.13	0.07	0.00	-0.07	-0.20	-0.01	0.24	0.05	-0.03	0.01	-0.01	0.11
18 lantai_tanah	-0.52	-0.37	-0.54	-0.34	-0.38	-0.31	-0.35	-0.59	-0.10	-0.33	-0.34	-0.61	0.06	0.05	0.00	-0.03	-0.05	1.00	-0.27	-0.49	0.40	-0.46	-0.09	0.29	-0.01	-0.33	-0.40	-0.27	-0.19	0.00	0.24	-0.12
19 lantai_semen	-0.28	-0.49	-0.41	-0.48	-0.29	-0.24	-0.40	-0.25	-0.07	-0.24	-0.48	-0.34	0.13	0.07	0.03	0.05	-0.09	-0.27	1.00	-0.57	0.22	-0.13	-0.17	0.31	0.01	-0.40	-0.28	-0.18	-0.20	0.01	0.21	-0.22
20 lantai_keramikubin	0.46	0.56	0.57	0.52	0.37	0.31	0.46	0.52	0.10	0.41	0.49	0.61	-0.11	-0.06	-0.05	-0.04	0.06	-0.49	-0.57	1.00	-0.41	0.29	0.28	-0.36	-0.06	0.43	0.39	0.29	0.26	0.02	-0.32	0.20
21 dinding_papan~u	-0.15	-0.29	-0.67	-0.29	-0.28	-0.21	-0.19	-0.23	-0.08	-0.26	-0.29	-0.24	0.01	-0.11	-0.15	-0.15	-0.13	0.40	0.22	-0.41	1.00	-0.57	-0.15	0.47	0.17	-0.27	-0.45	-0.28	-0.19	-0.01	0.18	-0.11
22 rangka_betonbaja	0.29	0.24	0.44	0.30	0.20	0.21	0.23	0.25	0.03	0.20	0.33	0.36	-0.17	-0.09	0.01	-0.06	0.07	-0.46	-0.13	0.29	-0.57	1.00	-0.40	-0.48	0.29	0.31	0.20	0.15	0.18	0.02	-0.23	0.11
23 rangka_kayu	0.09	0.27	0.18	0.21	0.19	0.10	0.23	0.18	0.05	0.20	0.18	0.17	0.00	0.03	0.06	0.01	0.00	-0.09	-0.17	0.28	-0.15	-0.40	1.00	-0.55	-0.13	0.18	0.24	0.23	0.14	0.08	-0.16	0.19
24 rangka_bambu	-0.29	-0.37	-0.40	-0.38	-0.24	-0.24	-0.36	-0.26	-0.05	-0.26	-0.41	-0.40	0.10	0.03	-0.08	0.05	-0.07	0.29	0.31	-0.36	0.47	-0.48	-0.55	1.00	-0.16	-0.33	-0.29	-0.23	-0.20	-0.07	0.26	-0.27
25 atap_sengplatasbes	0.07	0.04	0.03	0.03	0.10	-0.11	0.01	0.03	-0.05	-0.02	-0.01	0.00	-0.16	-0.33	-0.15	-0.28	-0.20	-0.01	0.01	-0.06	0.17	0.29	-0.13	-0.16	1.00	0.13	-0.12	0.01	-0.05	0.15	-0.09	0.04
26 listrik_nonsubsidi	0.48	0.63	0.53	0.63	0.38	0.33	0.55	0.48	0.05	0.34	0.57	0.54	-0.22	-0.18	-0.10	-0.05	-0.01	-0.33	-0.40	0.43	-0.27	0.31	0.18	-0.33	0.13	1.00	0.33	0.31	0.27	0.05	-0.36	0.25
27 fasilitas_bab	0.32	0.54	0.72	0.49	0.35	0.30	0.45	0.34	0.10	0.37	0.47	0.42	0.01	0.04	0.10	0.11	0.24	-0.40	-0.28	0.39	-0.45	0.20	0.24	-0.29	-0.12	0.33	1.00	0.70	0.52	0.11	-0.63	0.25
28 kloset	0.18	0.41	0.30	0.33	0.28	0.25	0.32	0.26	-0.01	0.32	0.29	0.25	-0.02	-0.01	0.07	0.12	0.05	-0.27	-0.18	0.29	-0.28	0.15	0.23	-0.23	0.01	0.31	0.70	1.00	0.74	-0.03	-0.69	0.08
29 tpa_septictank	0.37	0.30	0.20	0.30	0.18	0.06	0.22	0.16	0.04	0.22	0.23	0.15	-0.06	-0.04	0.01	0.04	-0.03	-0.19	-0.20	0.26	-0.19	0.18	0.14	-0.20	-0.05	0.27	0.52	0.74	1.00	-0.46	-0.56	0.04
30 tpa_spal	-0.52	0.02	0.08	0.01	0.03	0.09	0.07	0.07	0.01	0.04	0.00	0.07	-0.03	-0.13	-0.04	-0.06	0.01	0.00	0.01	0.02	-0.01	0.02	0.08	-0.07	0.15	0.05	0.11	-0.03	-0.46	1.00	-0.46	0.12
31 tpa_kolamdll	-0.01	-0.34	-0.32	-0.34	-0.23	-0.17	-0.32	-0.25	-0.04	-0.23	-0.30	-0.28	0.08	0.15	0.04	-0.01	-0.01	0.24	0.21	-0.32	0.18	-0.23	-0.16	0.26	-0.09	-0.36	-0.63	-0.69	-0.56	-0.46	1.00	-0.17
32 air_layak	0.37	0.25	0.35	0.25	0.14	0.10	0.25	0.21	-0.06	0.14	0.32	0.24	-0.13	-0.05	0.11	-0.07	0.11	-0.12	-0.22	0.20	-0.11	0.11	0.19	-0.27	0.04	0.25	0.25	0.08	0.04	0.12	-0.17	1.00

Note:
Red colour signifies a weak correlation, whereas green colour signifies a strong correlation between two variables.
The blue colour indicates the correlation value of a variable with itself.
The numbers on the column header represent the variable's sequential order, referring to the variable list in the column.

Table 6. Correlation coefficient between potential variables in Wave 4

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	tv	1.00	-0.55	0.49	0.08	0.35	0.49	0.08	0.24	0.11	0.21	0.51	0.26	0.15	0.07	0.08	0.07	0.10	0.00	-0.22	-0.18	0.32	-0.20	0.05	0.14	-0.13	0.02	0.20	0.40	0.14	-0.07	-0.09	0.07
2	tvkabel	-0.55	1.00	0.22	0.54	0.28	0.16	0.36	0.38	0.33	0.13	0.17	0.43	0.38	-0.08	-0.01	0.03	0.00	0.12	-0.15	-0.31	0.30	-0.14	0.22	0.18	-0.38	-0.08	0.23	0.13	0.00	0.09	-0.11	0.17
3	kulkas	0.49	0.22	1.00	0.57	0.71	0.61	0.30	0.58	0.49	0.18	0.58	0.62	0.54	-0.13	-0.05	-0.01	0.05	0.03	-0.37	-0.44	0.51	-0.32	0.23	0.27	-0.36	0.01	0.55	0.52	0.18	0.01	-0.21	0.31
4	AC	0.08	0.54	0.57	1.00	0.57	0.51	0.47	0.47	0.52	0.11	0.30	0.63	0.67	-0.22	-0.11	-0.09	-0.02	0.11	-0.58	-0.30	0.57	-0.13	0.25	0.19	-0.41	0.11	0.50	0.46	0.16	0.00	-0.19	0.33
5	mesin_cuci	0.35	0.28	0.71	0.57	1.00	0.57	0.35	0.57	0.49	0.15	0.52	0.62	0.53	-0.13	-0.06	0.04	0.05	0.08	-0.41	-0.35	0.45	-0.32	0.22	0.25	-0.34	0.01	0.50	0.49	0.13	0.06	-0.21	0.25
6	telpon_seluler	0.49	0.16	0.61	0.51	0.57	1.00	0.27	0.68	0.49	0.25	0.81	0.49	0.35	-0.03	-0.11	0.03	0.01	-0.12	-0.37	-0.27	0.34	-0.30	0.21	0.16	-0.23	0.01	0.38	0.36	0.16	0.02	-0.20	0.17
7	pemanas_air	0.08	0.36	0.30	0.47	0.35	0.27	1.00	0.35	0.28	0.09	0.24	0.39	0.39	-0.02	-0.05	0.08	0.08	0.10	-0.20	-0.25	0.26	-0.17	0.23	0.06	-0.24	-0.02	0.29	0.22	0.06	-0.06	-0.05	0.10
8	laptop	0.24	0.38	0.58	0.47	0.57	0.68	0.35	1.00	0.52	0.22	0.65	0.60	0.54	-0.13	0.01	0.10	0.08	0.09	-0.31	-0.39	0.43	-0.22	0.20	0.26	-0.37	-0.02	0.44	0.46	0.15	0.05	-0.21	0.20
9	pc	0.11	0.33	0.49	0.52	0.49	0.49	0.28	0.52	1.00	0.16	0.39	0.47	0.52	-0.10	0.01	0.06	0.05	0.04	-0.60	-0.15	0.42	-0.25	0.24	0.14	-0.31	0.05	0.42	0.33	0.10	0.01	-0.14	0.22
10	sepeda	0.21	0.13	0.18	0.11	0.15	0.25	0.09	0.22	0.16	1.00	0.23	0.20	0.08	0.11	0.14	-0.01	0.11	0.06	-0.12	-0.08	0.13	-0.01	0.03	0.04	-0.08	-0.07	0.09	0.13	0.02	-0.01	-0.02	0.07
11	motor	0.51	0.17	0.58	0.30	0.52	0.81	0.24	0.65	0.39	0.23	1.00	0.51	0.24	0.07	-0.02	0.08	0.08	-0.06	-0.30	-0.24	0.36	-0.24	0.10	0.18	-0.17	-0.06	0.35	0.36	0.13	0.03	-0.16	0.16
12	mobil	0.26	0.43	0.62	0.63	0.62	0.49	0.39	0.60	0.47	0.20	0.51	1.00	0.62	-0.09	0.05	0.14	0.11	0.17	-0.29	-0.49	0.49	-0.23	0.24	0.25	-0.39	0.00	0.50	0.54	0.19	0.01	-0.24	0.25
13	gas_nonsubsidi	0.15	0.38	0.54	0.67	0.53	0.35	0.39	0.54	0.52	0.08	0.24	0.62	1.00	-0.20	-0.02	0.06	0.06	0.14	-0.60	-0.29	0.58	-0.23	0.23	0.24	-0.41	-0.05	0.50	0.40	0.20	-0.04	-0.20	0.28
14	ternak	0.07	-0.08	-0.13	-0.22	-0.13	-0.03	-0.02	-0.13	-0.10	0.11	0.07	-0.09	-0.20	1.00	0.32	0.34	0.27	0.07	0.03	0.13	-0.11	-0.06	-0.11	-0.07	0.12	-0.25	-0.17	0.05	-0.02	-0.14	0.14	-0.10
15	sawah	0.08	-0.01	-0.05	-0.11	-0.06	-0.11	-0.05	0.01	0.01	0.14	-0.02	0.05	-0.02	0.32	1.00	0.33	0.34	0.21	0.00	0.04	-0.05	-0.08	-0.08	0.02	-0.01	-0.34	-0.17	0.12	-0.07	-0.10	0.14	-0.09
16	kebun	0.07	0.03	-0.01	-0.09	0.04	0.03	0.08	0.10	0.06	-0.01	0.08	0.14	0.06	0.34	0.33	1.00	0.23	0.21	-0.03	0.01	0.00	-0.10	-0.03	0.05	-0.05	-0.21	-0.10	0.09	0.04	-0.17	0.07	-0.10
17	pekarangan	0.10	0.00	0.05	-0.02	0.05	0.01	0.08	0.08	0.05	0.11	0.08	0.11	0.06	0.27	0.34	0.23	1.00	0.19	-0.04	-0.01	0.01	-0.13	-0.08	0.04	0.00	-0.31	-0.04	0.15	0.00	-0.06	0.05	-0.06
18	rumah	0.00	0.12	0.03	0.11	0.08	-0.12	0.10	0.09	0.04	0.06	-0.06	0.17	0.14	0.07	0.21	0.21	0.19	1.00	0.03	-0.14	0.09	-0.12	0.00	0.04	-0.03	-0.14	0.04	0.24	-0.01	-0.03	0.03	0.03
19	lantai_tanah	-0.22	-0.15	-0.37	-0.58	-0.41	-0.37	-0.20	-0.31	-0.60	-0.12	-0.30	-0.29	-0.60	0.03	0.00	-0.03	-0.04	0.03	1.00	-0.32	-0.49	0.29	-0.03	-0.24	0.24	0.07	-0.28	-0.31	-0.08	-0.02	0.10	-0.16
20	lantai_semen	-0.18	-0.31	-0.44	-0.30	-0.35	-0.27	-0.25	-0.39	-0.15	-0.08	-0.24	-0.49	-0.29	0.13	0.04	0.01	-0.01	-0.14	-0.32	1.00	-0.57	0.15	-0.23	-0.22	0.32	-0.04	-0.34	-0.26	-0.10	-0.03	0.14	-0.24
21	lantai_keramikubin	0.32	0.30	0.51	0.57	0.45	0.34	0.26	0.43	0.42	0.13	0.36	0.49	0.58	-0.11	-0.05	0.00	0.01	0.09	-0.49	-0.57	1.00	-0.38	0.18	0.33	-0.37	-0.01	0.38	0.34	0.11	0.05	-0.17	0.28
22	dinding_papan^u	-0.20	-0.14	-0.32	-0.13	-0.32	-0.30	-0.17	-0.22	-0.25	-0.01	-0.24	-0.23	-0.23	-0.06	-0.08	-0.10	-0.13	-0.12	0.29	0.15	-0.38	1.00	-0.56	-0.07	0.39	0.09	-0.11	-0.29	-0.04	0.03	0.12	-0.04
23	rangka_betonbaja	0.05	0.22	0.23	0.25	0.22	0.21	0.23	0.20	0.24	0.03	0.10	0.24	0.23	-0.11	-0.08	-0.03	-0.08	0.00	-0.03	-0.23	0.18	-0.56	1.00	-0.39	-0.47	0.33	0.25	0.21	0.19	0.11	-0.50	0.16
24	rangka_kayu	0.14	0.18	0.27	0.19	0.25	0.16	0.06	0.26	0.14	0.04	0.18	0.25	0.24	-0.07	0.02	0.05	0.04	0.04	-0.24	-0.22	0.33	-0.07	-0.39	1.00	-0.55	-0.03	0.22	0.24	0.11	0.05	-0.05	0.18
25	rangka_bambu	-0.13	-0.38	-0.36	-0.41	-0.34	-0.23	-0.24	-0.37	-0.31	-0.08	-0.17	-0.39	-0.41	0.12	-0.01	-0.05	0.00	-0.03	0.24	0.32	-0.37	0.39	-0.47	-0.55	1.00	-0.30	-0.35	-0.26	-0.15	-0.05	0.29	-0.24
26	atap_sengp^s	0.02	-0.08	0.01	0.11	0.01	0.01	-0.02	-0.02	0.05	-0.07	-0.06	0.00	-0.05	-0.25	-0.34	-0.21	-0.31	-0.14	0.07	-0.04	-0.01	0.09	0.33	-0.03	-0.30	1.00	0.14	-0.01	-0.03	0.12	-0.11	0.15
27	listrik_nonsubsidi	0.20	0.23	0.55	0.50	0.50	0.38	0.29	0.44	0.42	0.09	0.35	0.50	0.50	-0.17	-0.17	-0.10	-0.04	0.04	-0.28	-0.34	0.38	-0.11	0.25	0.22	-0.35	0.14	1.00	0.26	0.17	0.09	-0.28	0.30
28	fasilitas_bab	0.40	0.13	0.52	0.46	0.49	0.36	0.22	0.46	0.33	0.13	0.36	0.54	0.40	0.05	0.12	0.09	0.15	0.24	-0.31	-0.26	0.34	-0.29	0.21	0.24	-0.26	-0.01	0.26	1.00	0.38	0.12	-0.48	0.17
29	tpa_septictank	0.14	0.00	0.18	0.16	0.13	0.16	0.06	0.15	0.10	0.02	0.13	0.19	0.20	-0.02	-0.07	0.04	0.00	-0.01	-0.08	-0.10	0.11	-0.04	0.19	0.11	-0.15	-0.03	0.17	0.38	1.00	-0.50	-0.49	0.22
30	tpa_spal	-0.07	0.09	0.01	0.00	0.06	0.02	-0.06	0.05	0.01	-0.01	0.03	0.01	-0.04	-0.14	-0.10	-0.17	-0.06	-0.03	-0.02	-0.03	0.05	0.03	0.11	0.05	-0.05	0.12	0.09	0.12	-0.50	1.00	-0.48	0.00
31	tpa_kolamdll	-0.09	-0.11	-0.21	-0.19	-0.21	-0.20	-0.05	-0.21	-0.14	-0.02	-0.16	-0.24	-0.20	0.14	0.14	0.07	0.05	0.03	0.10	0.14	-0.17	0.12	-0.50	-0.05	0.29	-0.11	-0.28	-0.48	-0.49	-0.48	1.00	-0.24
32	air_layak	0.07	0.17	0.31	0.33	0.25	0.17	0.10	0.20	0.22	0.07	0.16	0.25	0.28	-0.10	-0.09	-0.10	-0.06	0.03	-0.16	-0.24	0.28	-0.04	0.16	0.18	-0.24	0.15	0.30	0.17	0.22	0.00	-0.24	1.00

Note:
Red colour signifies a weak correlation, whereas green colour signifies a strong correlation between two variables.
The blue colour indicates the correlation value of a variable with itself.
The numbers on the column header represent the variable's sequential order, referring to the variable list in the column.

Table 7. Correlation coefficient between potential variables in Wave 8

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	tv	1.00	-0.84	0.14	0.01	0.01	0.10	-0.10	0.04	-0.05	0.09	0.15	0.04	0.02	0.09	0.11	0.00	0.17	-0.09	-0.09	0.02	0.12	-0.13	-0.03	-0.04	0.09	0.03	-0.03	0.00	0.27	-0.03	-0.32
2	tvkabel	-0.84	1.00	0.30	0.28	0.28	0.26	0.31	0.17	0.17	0.05	0.21	0.21	0.25	-0.07	-0.06	0.04	-0.17	0.14	-0.26	-0.14	0.27	-0.02	0.10	0.19	-0.31	0.10	-0.10	0.22	-0.25	0.03	0.29
3	kulkas	0.14	0.30	1.00	0.49	0.70	0.64	0.44	0.52	0.36	0.25	0.60	0.59	0.47	-0.01	-0.02	-0.02	0.08	0.03	-0.35	-0.39	0.48	-0.28	0.11	0.20	-0.29	0.10	-0.10	0.39	0.05	0.01	-0.07
4	AC	0.01	0.28	0.49	1.00	0.49	0.41	0.69	0.48	0.40	0.19	0.22	0.55	0.64	-0.16	-0.11	0.01	0.02	0.16	-0.61	-0.30	0.53	-0.18	0.27	0.07	-0.36	-0.07	0.07	0.54	0.05	0.04	-0.10
5	mesin_cuci	0.01	0.28	0.70	0.49	1.00	0.60	0.43	0.49	0.36	0.19	0.53	0.56	0.36	-0.08	-0.03	0.02	0.09	0.07	-0.29	-0.35	0.40	-0.21	0.12	0.19	-0.30	0.08	-0.08	0.43	0.06	0.04	-0.11
6	telpon_seluler	0.10	0.26	0.64	0.41	0.60	1.00	0.22	0.64	0.40	0.23	0.83	0.55	0.31	-0.01	-0.14	-0.04	0.02	-0.22	-0.28	-0.29	0.37	-0.25	0.11	0.15	-0.24	0.08	-0.09	0.38	0.08	0.00	-0.07
7	pemanas_air	-0.10	0.31	0.44	0.69	0.43	0.22	1.00	0.42	0.31	0.09	0.05	0.52	0.60	-0.09	0.01	0.09	0.03	0.19	-0.61	-0.23	0.49	-0.10	0.18	0.12	-0.34	-0.03	0.04	0.44	0.02	0.06	-0.08
8	laptop	0.04	0.17	0.52	0.48	0.49	0.64	0.42	1.00	0.48	0.25	0.56	0.57	0.48	-0.07	0.09	0.09	0.18	0.07	-0.27	-0.36	0.38	-0.25	0.11	0.18	-0.29	0.06	-0.06	0.37	0.03	0.04	-0.09
9	pc	-0.05	0.17	0.36	0.40	0.36	0.40	0.31	0.48	1.00	0.16	0.36	0.45	0.44	-0.15	-0.03	0.08	0.10	0.10	-0.22	-0.29	0.29	-0.85	0.16	0.12	-0.27	-0.04	0.04	0.34	0.02	0.07	-0.08
10	sepeda	0.09	0.05	0.25	0.19	0.19	0.23	0.09	0.25	0.16	1.00	0.27	0.25	0.18	0.06	0.07	-0.05	0.11	0.05	-0.06	-0.11	0.09	-0.09	0.00	0.09	-0.11	0.06	-0.06	0.11	0.02	-0.01	-0.02
11	motor	0.15	0.21	0.60	0.22	0.53	0.83	0.05	0.56	0.36	0.27	1.00	0.41	0.14	0.11	0.00	0.07	0.09	-0.11	-0.21	-0.21	0.34	-0.27	0.06	0.16	-0.20	0.15	-0.15	0.29	0.09	-0.06	-0.03
12	mobil	0.04	0.21	0.59	0.55	0.56	0.55	0.52	0.57	0.45	0.25	0.41	1.00	0.57	-0.04	0.05	0.14	0.12	0.14	-0.27	-0.43	0.41	-0.28	0.17	0.16	-0.34	0.07	-0.07	0.48	0.04	0.07	-0.13
13	gas_nonsubsidi	0.02	0.25	0.47	0.64	0.36	0.31	0.60	0.48	0.44	0.18	0.14	0.57	1.00	-0.14	0.03	0.15	0.12	0.22	-0.60	-0.26	0.49	-0.24	0.15	0.15	-0.33	0.07	-0.07	0.47	0.08	0.04	-0.14
14	ternak	0.09	-0.07	-0.01	-0.16	-0.08	-0.01	-0.09	-0.07	-0.15	0.06	0.11	-0.04	-0.14	1.00	0.33	0.28	0.33	0.09	0.03	0.09	-0.07	-0.11	-0.12	0.00	0.09	0.17	-0.17	-0.15	-0.02	-0.06	0.08
15	sawah	0.11	-0.06	-0.02	-0.11	-0.03	-0.14	0.01	0.09	-0.03	0.07	0.00	0.05	0.03	0.33	1.00	0.33	0.30	0.24	0.02	0.02	-0.03	-0.03	-0.18	0.08	0.02	0.36	-0.36	-0.18	0.01	-0.08	0.04
16	kebun	0.00	0.04	-0.02	0.01	0.02	-0.04	0.09	0.09	0.08	-0.05	0.07	0.14	0.15	0.28	0.33	1.00	0.25	0.21	0.00	-0.01	-0.03	-0.03	-0.06	0.08	-0.06	0.12	-0.12	-0.04	0.10	-0.03	-0.10
17	pekarangan	0.17	-0.17	0.08	0.02	0.09	0.02	0.03	0.18	0.10	0.11	0.09	0.12	0.12	0.33	0.30	0.25	1.00	0.13	-0.01	0.03	-0.06	-0.11	-0.09	0.07	-0.02	0.23	-0.23	-0.05	0.06	0.01	-0.11
18	rumah	-0.09	0.14	0.03	0.16	0.07	-0.22	0.19	0.07	0.10	0.05	-0.11	0.14	0.22	0.09	0.24	0.21	0.13	1.00	-0.02	-0.13	0.08	-0.15	0.02	0.05	-0.09	0.18	-0.18	0.04	0.09	-0.09	-0.03
19	lantai_tanah	-0.09	-0.26	-0.35	-0.61	-0.29	-0.28	-0.61	-0.27	-0.22	-0.06	-0.21	-0.27	-0.60	0.03	0.02	0.00	-0.01	-0.02	1.00	-0.32	-0.49	0.15	-0.03	-0.17	0.23	-0.03	0.03	-0.30	-0.01	-0.01	0.01
20	lantai_semen	0.02	-0.14	-0.39	-0.30	-0.35	-0.29	-0.23	-0.36	-0.29	-0.11	-0.21	-0.43	-0.26	0.09	0.02	-0.01	0.03	-0.13	-0.32	1.00	-0.56	0.24	-0.16	-0.17	0.29	-0.06	0.06	-0.32	-0.05	-0.03	0.09
21	lantai_keramikubin	0.12	0.27	0.48	0.53	0.40	0.37	0.49	0.38	0.29	0.09	0.34	0.41	0.49	-0.07	-0.03	-0.03	-0.06	0.08	-0.49	-0.56	1.00	-0.34	0.14	0.25	-0.36	0.08	-0.08	0.36	0.05	0.04	-0.08
22	dinding_papan~u	-0.13	-0.02	-0.28	-0.18	-0.21	-0.25	-0.10	-0.25	-0.85	-0.09	-0.27	-0.28	-0.24	-0.11	-0.03	-0.03	-0.11	-0.15	0.15	0.24	-0.34	1.00	-0.14	-0.13	0.24	-0.17	0.17	-0.17	-0.06	-0.03	0.07
23	rangka_betonbaja	-0.03	0.10	0.11	0.27	0.12	0.11	0.18	0.11	0.16	0.00	0.06	0.17	0.15	-0.12	-0.18	-0.06	-0.09	0.02	-0.03	-0.16	0.14	-0.14	1.00	-0.49	-0.50	-0.36	0.36	0.23	0.03	0.05	-0.08
24	rangka_kayu	-0.04	0.19	0.20	0.07	0.19	0.15	0.12	0.18	0.12	0.09	0.16	0.16	0.15	0.00	0.08	0.08	0.07	0.05	-0.17	-0.17	0.25	-0.13	-0.49	1.00	-0.50	0.10	-0.10	0.16	0.02	0.00	0.00
25	rangka_bambu	0.09	-0.31	-0.29	-0.36	-0.30	-0.24	-0.34	-0.29	-0.27	-0.11	-0.20	-0.34	-0.33	0.09	0.02	-0.06	-0.02	-0.09	0.23	0.29	-0.36	0.24	-0.50	-0.50	1.00	0.24	-0.24	-0.35	-0.04	-0.03	0.06
26	atap_genteng	0.03	0.10	0.10	-0.07	0.08	0.08	-0.03	0.06	-0.04	0.06	0.15	0.07	0.07	0.17	0.36	0.12	0.23	0.18	-0.03	-0.06	0.08	-0.17	-0.36	0.10	0.24	1.00	-1.00	-0.12	0.01	0.00	-0.02
27	atap_sengp~s	-0.03	-0.10	-0.10	0.07	-0.08	-0.09	0.04	-0.06	0.04	-0.06	-0.15	-0.07	-0.07	-0.17	-0.36	-0.12	-0.23	-0.18	0.03	0.06	-0.08	0.17	0.36	-0.10	-0.24	-1.00	1.00	0.12	-0.01	0.00	0.02
28	listrik_nonsubsidi	0.00	0.22	0.39	0.54	0.43	0.38	0.44	0.37	0.34	0.11	0.29	0.48	0.47	-0.15	-0.18	-0.04	-0.05	0.04	-0.30	-0.32	0.36	-0.17	0.23	0.16	-0.35	-0.12	0.12	1.00	0.05	0.08	-0.13
29	tpa_septictank	0.27	-0.25	0.05	0.05	0.06	0.08	0.02	0.03	0.02	0.02	0.09	0.04	0.08	-0.02	0.01	0.10	0.06	0.09	-0.01	-0.05	0.05	-0.06	0.03	0.02	-0.04	0.01	-0.01	0.05	1.00	-0.50	-0.50
30	tpa_spal	-0.03	0.03	0.01	0.04	0.04	0.00	0.06	0.04	0.07	-0.01	-0.06	0.07	0.04	-0.06	-0.08	-0.03	0.01	-0.09	-0.01	-0.03	0.04	-0.03	0.05	0.00	-0.03	0.00	0.00	0.08	-0.50	1.00	-0.49
31	tpa_kolamdll	-0.32	0.29	-0.07	-0.10	-0.11	-0.07	-0.08	-0.09	-0.08	-0.02	-0.03	-0.13	-0.14	0.08	0.04	-0.10	-0.11	-0.03	0.01	0.09	-0.08	0.07	-0.08	0.00	0.06	-0.02	0.02	-0.13	-0.50	-0.49	1.00

Note:
Red colour signifies a weak correlation, whereas green colour signifies a strong correlation between two variables.
The blue colour indicates the correlation value of a variable with itself.
The numbers on the column header represent the variable's sequential order, referring to the variable list in the column.

3. Evaluating Variable Suitability

Table 8. Interpreting the KMO values

90 ≤	*Excellent*
80 ≤ KMO <90	*Meritorious*
70 ≤ KMO < 80	*Middling*
60 ≤ KMO < 70	Mediocre
50 ≤ KMO < 60	Miserable
<50	Unacceptable

There are two methods used to assess the suitability of variables for Principal Component Analysis (PCA): Bartlett's Test and Kaiser-Meyer-Olkin (KMO) measure. Bartlett's Test for sphericity examines whether the correlation matrix significantly differs from an identity matrix (matrix with all zeros except for the diagonal). The null hypothesis assumes that all correlation coefficients are zero, while the alternative hypothesis suggests that at least two variables have significant correlations.

On the other hand, the KMO test evaluates the sampling adequacy for each indicator. Its result ranges from 0 to 100, where a value of 0 indicates that the tested variable set is unsuitable for PCA, while a value of 100 indicates that the variable set will produce reliable factors (Kaiser, 1974).

The results of both tests are presented in **Table 9**. Bartlett's Test showed significant results for all three variable sets, indicating a significant departure from the null hypothesis. Furthermore, the KMO test yielded KMO values exceeding 80 for all three variable sets, suggesting that they are suitable for analysis using PCA.

Table 9. Results of variable suitability tests

Sample adequacy tests	Wave 1	Wave 4	Wave 8
Bartlett test of sphericity	0.000	0.000	0.000
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.879	0.857	0.833

4. Principal component analysis (PCA)

PCA was conducted to analyse the socio-economic status in three waves: Wave 1 with 18 variables, Wave 4 with 14 variables, and Wave 8 with 12 variables. Each wave produced principal components and total variance, matching the number of variables analysed. The eigenvalues (variances) for each principal component are shown in **Table 10**. In Wave 1, principal component 1 accounted for 49.3% (8.865/18*100) of the total variance (8.865). Similarly, principal component 1 in Wave 4 and Wave 8 explained 52% of the total variance each. Thus, the first principal component effectively captured the variation in socio-economic status within the sample. Additionally, according to McKenzie (2003), only the first principal component is required to calculate the wealth index.

Table 10. Results from Principal Component Analysis

Component	Wave 1 in 2015		Wave 4 in 2018		Wave 8 in 2022	
	Eigenvalue	Proportion	Eigenvalue	Proportion	Eigenvalue	Proportion
Comp1	8.865	0.493	7.286	0.520	6.238	0.520
Comp2	1.909	0.106	1.458	0.104	1.508	0.126
Comp3	1.268	0.070	0.973	0.070	0.860	0.072
Comp4	1.062	0.059	0.751	0.054	0.690	0.058
Comp5	0.991	0.055	0.671	0.048	0.563	0.047
Comp6	0.745	0.041	0.592	0.042	0.484	0.040
Comp7	0.592	0.033	0.471	0.034	0.444	0.037
Comp8	0.550	0.031	0.438	0.031	0.390	0.033
Comp9	0.418	0.023	0.399	0.029	0.330	0.028
Comp10	0.357	0.020	0.342	0.024	0.265	0.022
Comp11	0.307	0.017	0.280	0.020	0.228	0.019
Comp12	0.302	0.017	0.238	0.017	0.000	0.000
Comp13	0.240	0.013	0.102	0.007		
Comp14	0.220	0.012	0.000	0.000		
Comp15	0.173	0.010				
Comp16	0.000	0.000				
Comp17	0.000	0.000				
Comp18	0.000	0.000				

5. The socio-economic level scores

The socio-economic level scores of each household wave are then predicted based on principal component 1. The equation to calculate the socio-economic level score of a household (*fscore*) per wave is constructed from the eigenvectors of each indicator variable (see **equations 1-3**).

***fscore* Wave 1**= *0.236*tvkabel + 0.2762*kulkas + 0.2731*AC + 0.2653*mesin_cuci + 0.2314*telpon_seluler + 0.1792*pemanas_air + 0.257*laptop + 0.2381*pc + 0.2051*motor + 0.2652*mobil + 0.2593*gas_nonsubsidi - 0.2562*lantai_tanah + 0.2445*lantai_keramikubin - 0.1776*dinding_papan~u + 0.1925*rangka_betonbaja - 0.1962*rangka_bambu + 0.2328*listrik_nonsubsidi + 0.2198*fasilitas_bab*

(Equation 1)

***fscore* Wave 4**= *0.2997*kulkas + 0.2954*AC + 0.2926*mesin_cuci + 0.2702*telpon_seluler + 0.2835*laptop + 0.259*pc + 0.244*motor + 0.2875*mobil + 0.2827*gas_nonsubsid - 0.2763*lantai_tanah + 0.2617*lantai_keramikubin - 0.2058*rangka_bambu + 0.2374*listrik_nonsubsidi + 0.2265*fasilitas_bab*

(Equation 2)

***fscore* Wave 8**= *0.3037*kulkas + 0.328*AC + 0.2868*mesin_cuci + 0.2688*telpon_seluler + 0.2963*pemanas_air + 0.2851*laptop + 0.2384*pc + 0.303*mobil + 0.3074*gas_nonsubsidi - 0.3001*lantai_tanah + 0.2795*lantai_keramikubin + 0.2557*listrik_nonsubsidi*

(Equation 3)

The *fscore* reflects the socio-economic status of households. A higher *fscore* indicates a higher socio-economic status. Thus, indicators with positive eigenvector values suggest that these indicators are more commonly owned or accessible among households with a higher socio-economic status. For example, in Wave 1, households with high socio-economic status tend to have cable TV, a refrigerator, air conditioning, a washing machine, a mobile phone, a water heater, a laptop, a desktop computer, a motorcycle, a car, live in houses with ceramic/tile or concrete/steel flooring, have access to proper sanitation facilities, and use non-subsidised gas and electricity. Conversely, households with houses made of earthen floors, wooden walls, and bamboo structures are likely to have a low socio-economic status. Upon further examination, these indicators of low socio-economic status are negatively correlated with indicators of high socio-economic status. For instance, families living in houses with earthen floors are unlikely to have cable TV, air conditioning, and may use subsidised gas.

The statistical description of household *fscore* per Wave is presented in **Table 11**. The distribution of *fscores* across the three Waves is quite similar. The most noticeable difference is that the minimum score in Wave 1 is lower than the data from the other Waves.

Table 11. Descriptive statistics of the socio-economic level scores

	Mean	Standard deviation	Minimum	Maximum
fscore Wave 1	1.347	0.754	-0.630	3.575
fscore Wave 4	1.543	0.727	-0.482	3.240
fscore Wave 8	1.260	0.624	-0.300	3.153

6. Determining SES quintiles

Based on their *fscores*, households in Sleman HDSS were grouped into five socio-economic status quintiles per Wave. This grouping process involved sorting households based on their *fscores* from smallest to largest. Then, the households were divided into five roughly equal-sised groups. Thus, the higher the quintile group, the higher the household's socio-economic status: low (Q1), lower middle, middle, upper middle, and high (Q5). **Figure 2** presents the percentage distribution of socio-economic status quintiles for the entire population and by residential location for each data collection Wave.

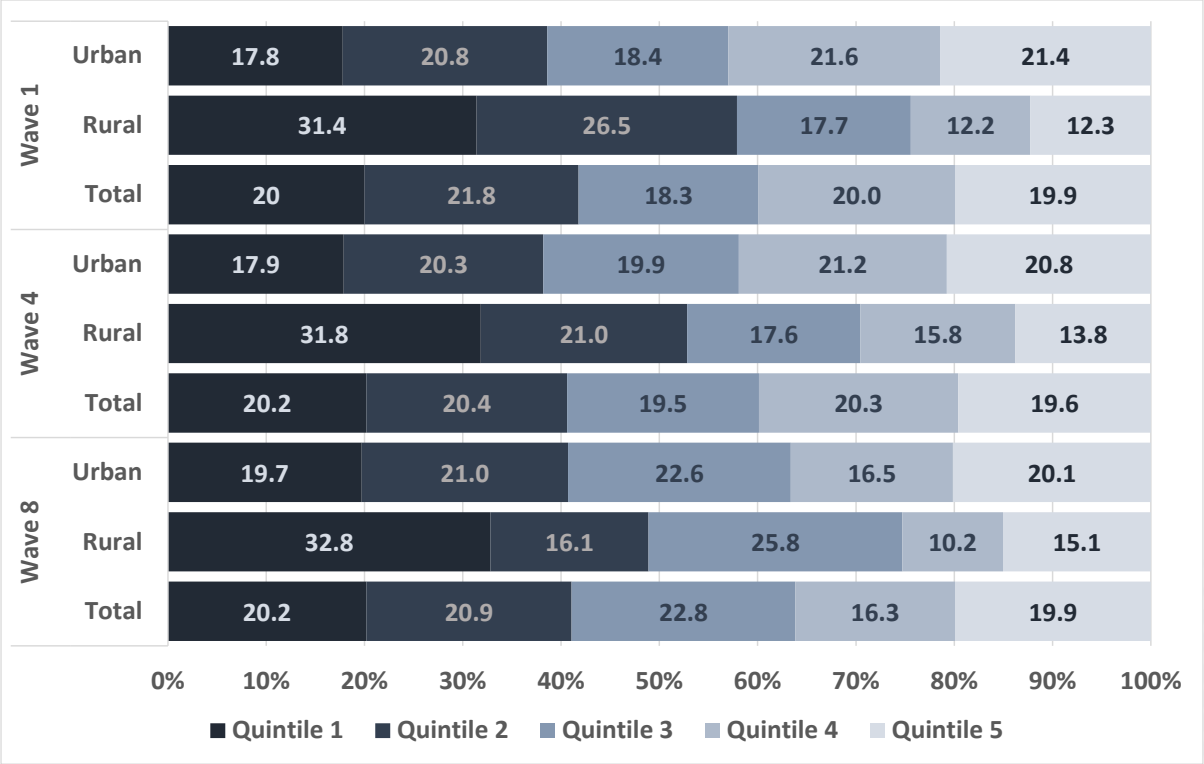

Figure 2. Households proportion by socio-economic quintile, residential area, and wave

Reference

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of* Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91-93. https://doi.org/https://doi.org/10.1016/j.tjem.2018.08.001

Ariawan, I. (2006). Indeks Sosio-ekonomi Menggunakan Principal Component Analysis. *2006*, 5. https://doi.org/10.21109/kesmas.v1i2.317

Depkes, R. (2013). Laporan Hasil Riset Kesehatan Dasar Indonesia Tahun 2013. *Badan Penelitian dan Pengembangan Kesehatan. Depkes RI. Jakarta*.

Howe, L. D., Galobardes, B., Matijasevich, A., Gordon, D., Johnston, D., Onwujekwe, O., . . . Hargreaves, J. R. (2012). Measuring socio-economic position for epidemiological studies in low- and middle-income countries: a methods of measurement in epidemiology paper. *International Journal of Epidemiology*, *41*(3), 871-886. https://doi.org/10.1093/ije/dys037

Jolliffe, I. T. (2002). *Principal component analysis for special types of data*. Springer.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. https://doi.org/10.1007/BF02291575

McKenzie, D. (2003). Measure Inequality With Asset Indicators (BREAD Working Paper No. 042). *Cambridge, MA: Bureau for Research and Economic Analysis of Development*.

Poirier, M. J. P., Grépin, K. A., & Grignon, M. (2020). Approaches and Alternatives to the Wealth Index to Measure Socio-economic Status Using Survey Data: A Critical Interpretive Synthesis. *Social Indicators Research*, *148*(1), 1-46. https://doi.org/10.1007/s11205-019-02187-9

Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning*, *21*(6), 459-468. https://doi.org/10.1093/heapol/czl029

Supplementary Files

The following link or QRcode provides the Stata Syntax for Computing Wealth Index as a Measure of Household Socio-Economic Status using Principal Component Analysis (PCA).

http://tinyurl.com/jmvwtfwk

UNIVERSITAS GADJAH MADA
FACULTY OF MEDICINE
PUBLIC HEALTH AND NURSING

SLEMAN HEALTH AND DEMOGRAPHIC
SURVEILLANCE SYSTEM
FMPHN UGM