Fighting Hunger Worldwide

VAM Guidance Paper

Creation of a Wealth Index

WFP
World Food Programme
wfp.org

Prepared by: Lisa Hjelm, Astrid Mathiassen, Darryl Miller, Amit Wadhwa.

Reviewed by: Rossella Bottone (June 2017)

Contact: Claudia AhPoe (claudia.ahpoe@wfp.org) or Kurt Burja (kurt.burja@wfp.org)

Contents

Background

Wealth is the value of all natural, physical and financial assets owned by a household, reduced by its liabilities. Household wealth is a measure commonly used in food security assessments. It gives an idea of households' ability to access food, the severity of food insecurity and gives information about the economic situation of the food insecure. It is used to differentiate between the poorer and the wealthier households in food security related indicators, such as food consumption, and thereby provides information on how to target the food insecure.

This guidance aims at complementing the Comprehensive Food Security & Vulnerability Analysis Guidelines (CFSVA 2009) with a practical step-by-step guidance on how to create one, while providing VAM officers and other food security analysts with a more general background of the wealth index (WI).

Measurements of wealth

There are several ways in which wealth, economic status of households and living standards can be measured. Income, expenditure and consumption are three common measurements.

However, there are challenges in collecting and measuring income and expenditure accurately. An alternative is to use data on asset ownership and housing characteristics and combine this information into a proxy indicator such as the wealth index, which is created using principal component analysis (PCA). Asset ownership gives an indication of the longer-term economic status of a household and is less dependent on short-term economic changes compared with other wealth or poverty measures.

The wealth index measures relative wealth and, unlike a poverty line, is not an absolute measure of poverty or wealth. When referring to the wealth of households based on the wealth index we can talk about poor*er* and wealthi*er* households but we cannot conclude who is absolutely poor and wealthy. The wealth index quintiles divide the whole population into five equally large groups, based on their wealth rank. For example, in an area where only 10% of households fall below the poverty line, 40% of households will still fall into the two poorest quintiles and therefore be classified as the poorest.

Current use of wealth index (WI)

The wealth index is commonly used in reports and analysis based on datasets from Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and WFP surveys (mainly CFSVAs[1] and EFSAs), and is used to rank households into quintiles. The value of using the wealth index is especially recognised in contexts where reliable income and expenditure data is absent.

The research questions related to the wealth index vary according to the different interests of the surveys. In DHS, the wealth index is chosen because of the major impact that wealth status has on household level health. It allows the researchers to identify the impact of wealth status on health outcomes[2]. For MICS, the wealth index serves a similar purpose in terms of understanding health outcomes. It is also used to target poverty alleviation programmes and projects[3].

Use of the wealth index in WFP's household surveys

For WFP-VAM food security analyses, the wealth index quintiles are useful for cross-tabulation with other variables in the dataset. For example, cross-tabulating with regions can show areas with higher proportions of poor households or cross-tabulating with food consumption groups could determine

[1] WFP, *Comprehensive Food Security & Vulnerability Analysis Guidelines,* 2009 page 211

[2] DHS wealth index. See at. http://dhsprogram.com/topics/Wealth-Index.cfm

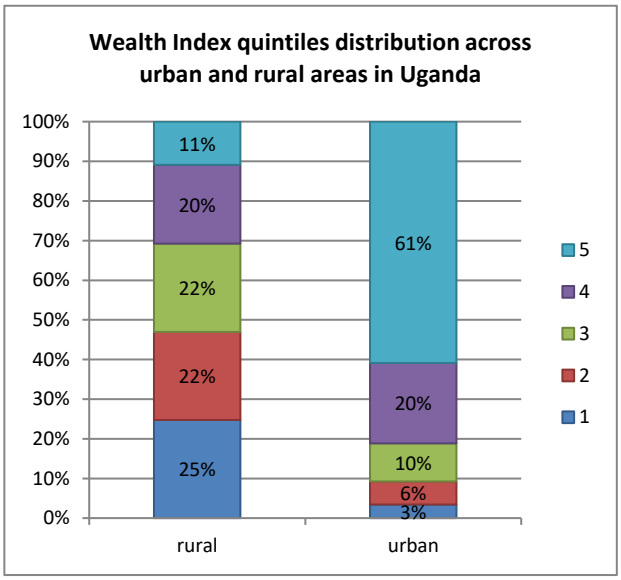
[3] UNICEF(2008). Measuring child poverty. See at http://www.southampton.ac.uk/ghp3/docs/unicef/presentation2.3.pdf

what proportion of households with poor food consumption are also in the poorer groups. In addition, the wealth index can be used as a proxy for food access..

Urban-rural considerations

One consideration that should be taken into account in relation to the wealth index is that wealth is characterized by ownership of different types of assets in urban areas compared with rural areas. Depending on the variables included in the index, the wealth measure can be biased towards urban or rural households. One solution is to include variables that are valid as proxies of wealth in both urban and rural areas. For example, if a high percentage of households live in urban areas and few households practise agriculture, we may consider excluding productive assets and livestock.

If the living conditions in urban and rural areas are very different, another approach can be to create separate indices for urban and rural areas. The following chart shows the distribution of rural and urban households in Uganda using the uniform national wealth index. We observe a remarkable urban-rural divergence in terms of relative wealth status. The analysts may need to assess if this distribution reflects the subsutantial urban-rural inequality or if it is a result of variable selection bias. If the latter situation is likely to be true, we may need to construct a separate wealth index for rural and urban households that take into account the differences in assets owned.



The question of when to create a separate urban-rural wealth index is an open one that calls for both researchers and analysts to examine the specific context. In this guidance, we focus on the methods of constructing a national wealth index. Similar methodologies can be adopted for index creation separately for urban and rural areas.

Steps for the creation of a wealth index (WI)

1. **Select variables**
2. **Explore variables**
    a. **Frequencies**
    b. **Missing values**
3. **Recode into binary variables**
4. **Principal components analysis (PCA)**
5. **Create wealth index quintiles**
6. **Graph the index**
7. **Select the final result and report the variables**

Note: Uganda LSMS 08/09 dataset is used to demonstrate the WI creation and SPSS (Statistical Package for the Social Sciences) procedures in this guidance.

STEP 1: Select variables

The first step in creating the wealth index is to select variables to be incorporated in the questionnaire. The questionnaires ask what the households own based on an extensive list of assets and other housing characteristics that reflect their economic status (see example in the table below).

A broad range of variables could be included in the analysis: a greater number can reduce the sampling bias and generate a better distribution of households. The final list should be country-specific, and simultaneously capture the differences in ownership among households (see more in Step 2).

Example of variables that can be included in household surveys:

Productive assets	Non-productive assets	Household utilities and other
Hand mill	Radio	Types of:
Sickle	Refrigerator	Water supply
Axe	TV	Toilet
Livestock	Bicycle	Flooring
Hoe	Motorbike	Walls/house
Tractor	Phone/cell phone	Roof
Plough	Chair	Light source
	Table	
	Bed	Persons sleeping per room
		Land ownership
		Livestock ownership

STEP 2: Explore variables

Once the possible variables have been identified in the data set, they require further investigation before being selected for the index.

2.a. Explore the variables by running descriptive analysis including a frequency of each variable.

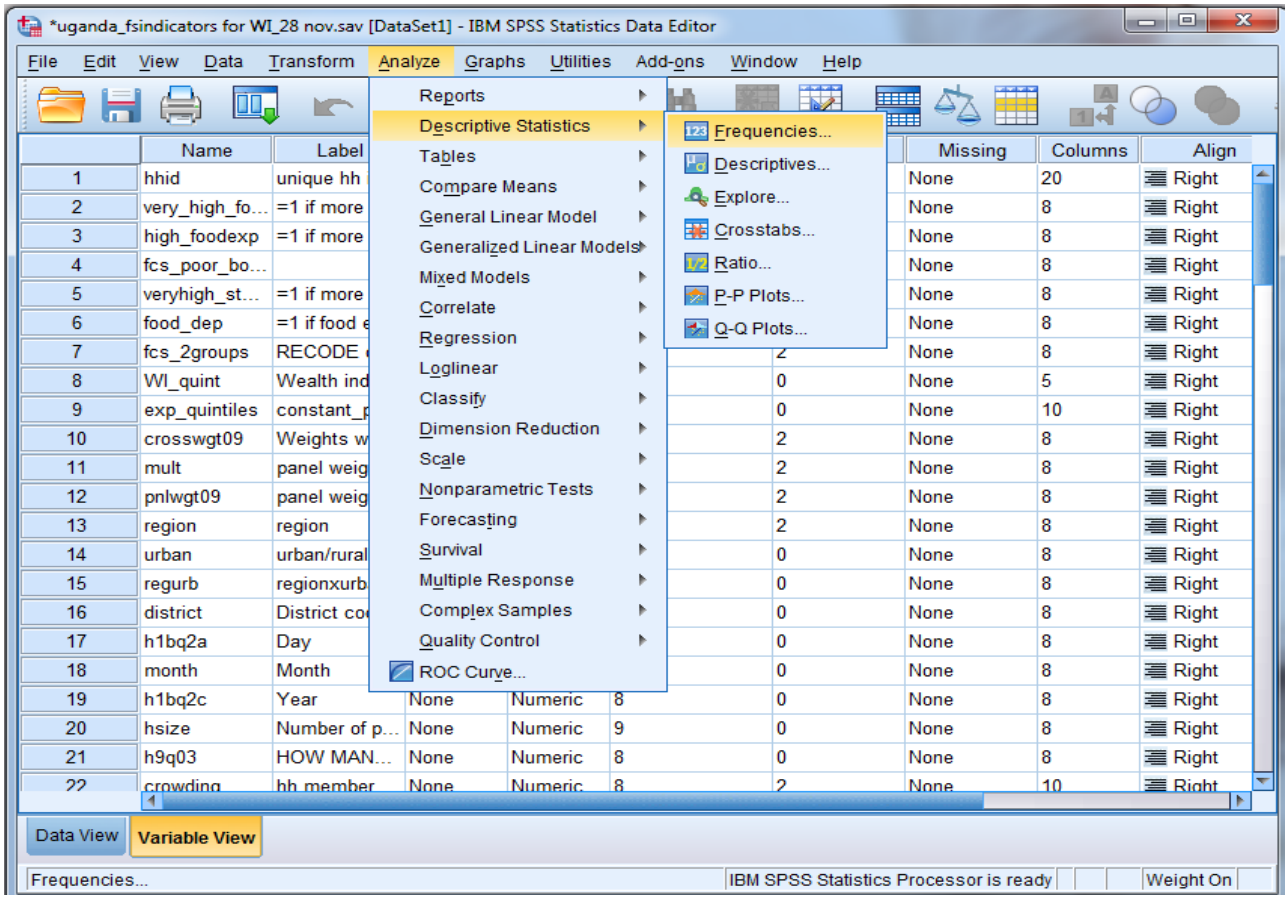
As a first step in exploring the variables to include in the index some basic cleaning of data may be needed. A household which has missing values for any of the assets will be excluded in the wealth index construction. If a substantial proportion of missing values is detected, the analysts should check the data quality again and if possible go back to the enumerators to ensure accurate data collection and entry.

We need to select the variables that are capable of distinguishing relatively "wealthy" households and relatively "poor" ones. The rule of thumb is that if a variable/asset is owned by more than 95% or less than 5% of the sample, it should be excluded from the analysis. For example, knowing that 99.2% of Ugandan households don't own a generator will not help the analyst to distinguish between richer and poorer households by this asset ownership (see table below). Thus, this variable will be excluded from the index.

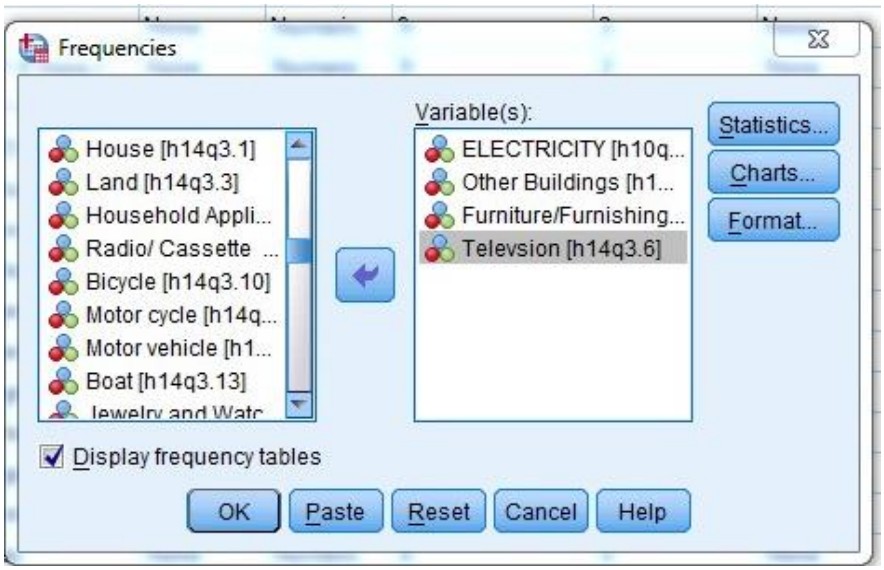
DOES THIS HOUSE HAVE A GENERATOR? (SPSS frequency table)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	no	5152193	98.6	99.2	99.2
	yes	44014	.8	.8	100.0
	Total	5196207	99.5	100.0	
Missing	System	28473	.5		
Total		5224680	100.0		

SPSS: Analyse → Descriptive statistics → Frequencies



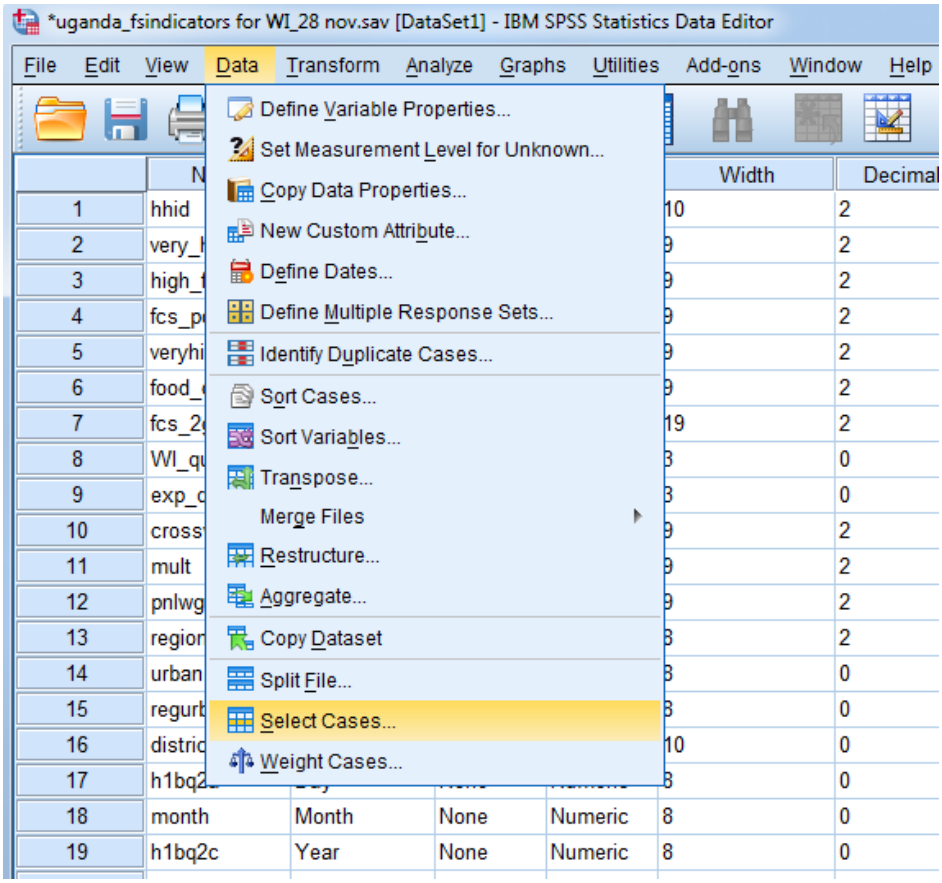
In the dialogue box, click the variables in the left field that we want to run frequencies for and click on the right arrow to move the variables into the right field called '**Variable(s)**'. The option '**Display frequency tables**' at the bottom of the dialogue box is checked as default.



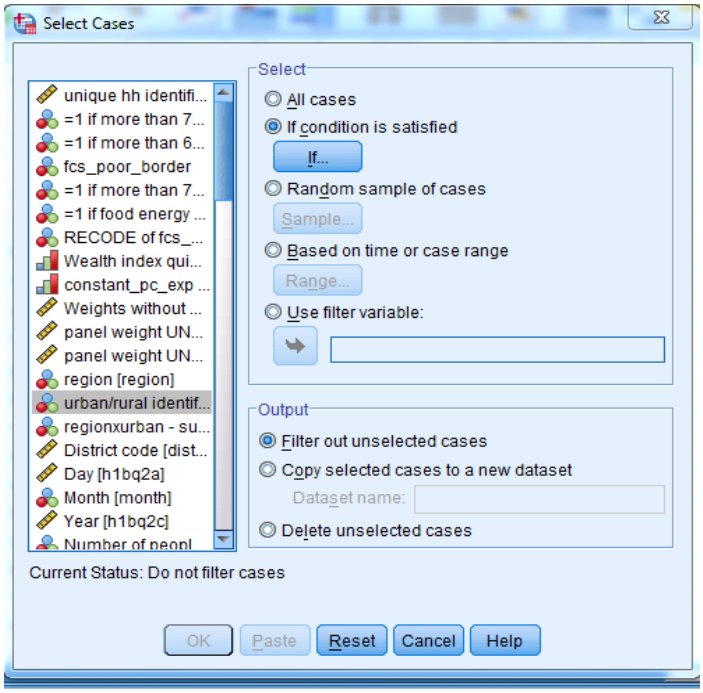
2 b. Run the frequencies for urban and rural areas separately

In a similar way to the rule of thumb discussed in 2.a, we run the frequencies for urban and rural areas separately to determine the variables to create a national wealth index. If there are certain assets owned by very few in either urban or rural areas, we will consider not including them because the national index needs to represent both urban and rural households. The inclusion of productive assets/livestock and land ownership should be reassessed if a high percentage of households do not practise agriculture or if many households are located in urban areas.

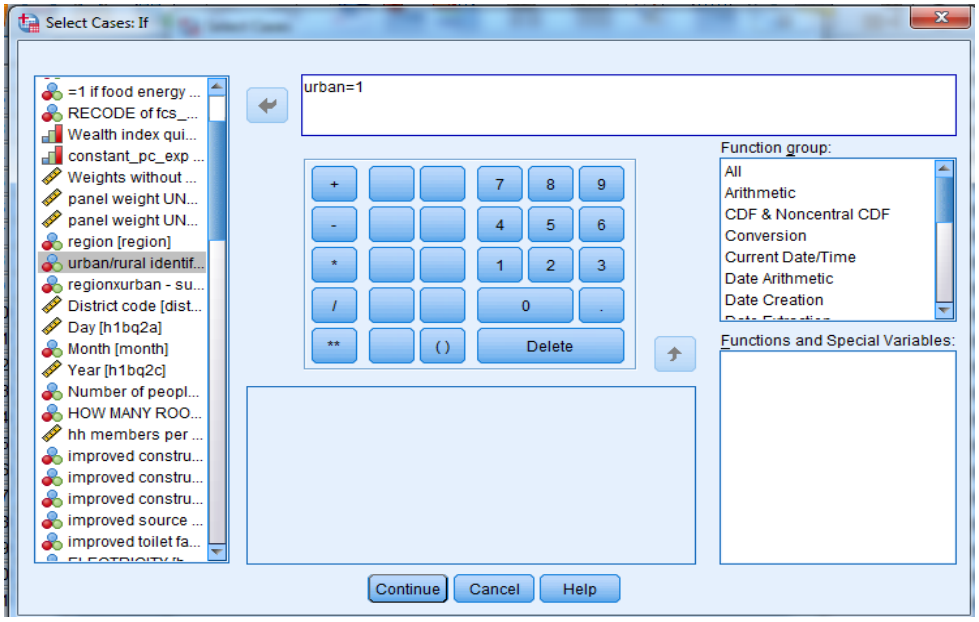
SPSS: Data→ Select cases



When the dialogue box pops up, select the urban/rural categorical variable ('urban/rural identifier' in this database) in the left field and select the option '**If condition is satisfied**' in the right field. Once selected, click the button **'If'** to customize specific conditions for the cases we want to analyse.



In this step, we want to look at data for urban and rural areas separately. Double click the selected categorical variable ('urban/rural identifier' in this database) and type '=1' (in this dataset, urban=1 and rural=0) in the blank column right of the arrow. Now we are narrowing the analysis scope down to urban households.



Click '**Continue**' and repeat the 'frequency' process in 2.a. to examine the asset ownership frequencies of urban households. To analyse rural households, change the 'if' conditions into 'urban/rural identifier=0' and follow the same procedures as above.

After steps above, remember to turn off the selection by going to **Data→ Select cases and click '**All cases**'.

Step 3: Recode into scale variables

Before creating the wealth index, all variables should be transformed into scale variables. Most often asset ownership questions are categorical variables. Yes/no variables should be recoded into binary variables (variables that takes only 0 and 1 to represent the categorical effect) and variables with more than 2 categories should be transformed into bivariate variables**.**

Questions regarding housing characteristics and access to services are commonly categorical variables with several options. When this is the case a decision has to be made on how to recode these variables so that there are only two categories. When doing this, identify the alternatives that are more likely to be found in wealthy households compared with poorer households. How the variables should be recoded depends on the context of the country and what is more likely to distinguish poorer households from wealthier households. For instance, if the light source has many options, such as "none," "wood fire," "oil lamp," "petrol light," "electricity," it might be appropriate to recode this into "none/primitive" for those households that answered "none" or "wood fire" and "purchased energy source," for the remaining light sources. In another country, the same options might be recoded differently: "no electric" versus "electric."

The choice is based on what is more likely to define wealth and also by looking at the prevalence of both categories: if the prevalence is between 30 and 70%, then the indicator will probably help categorize more households than if the prevalence is only 5%. Two similar variables can be combined if this will result in a summed prevalence of between 30 and 70%.

For sanitation facilities and source of water the UNICEF/WHO standards can be used (see table below)[4]. However, the recoding between improved/ not improved is just one possibility. Other classifications can be used if this is more likely to separate the poorer households from the richer in the country context. One example is the alternative classification of source of water: bottled water is regarded as an unimproved source since quantities are not usually large enough to supply a household, but in reality, those who can afford to buy bottled water, especially in less developed countries, are often wealthier. So we may consider including bottled water in the improved group.

[4] UNICEF&WHO. Progress on sanitation and drinking water.2013 Update. See at http://apps.who.int/iris/bitstream/10665/81245/1/9789241505390_eng.pdf

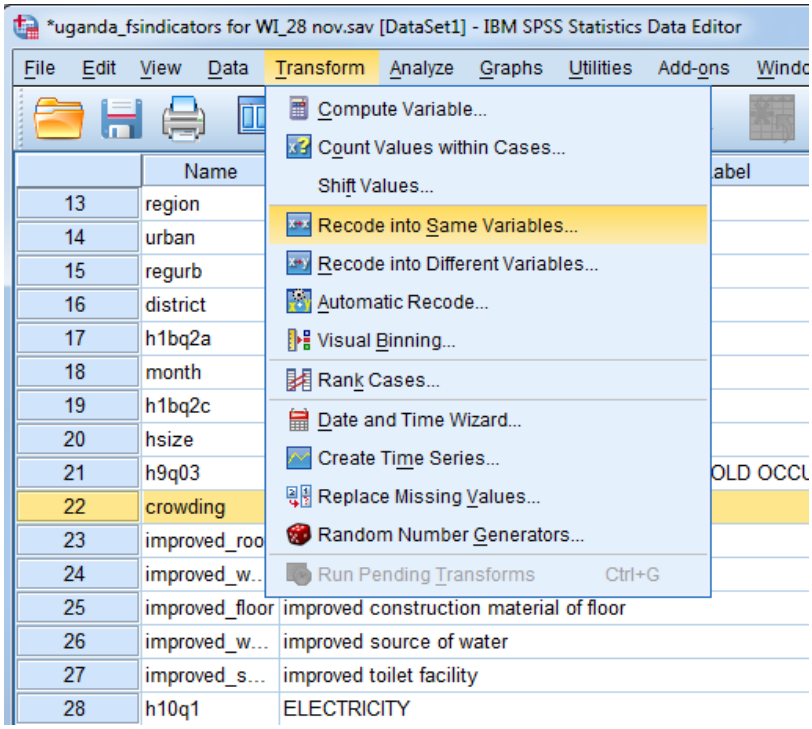
UNICEF/WHO Types of drinking water sources[5]	
IMPROVED	**UNIMPROVED**
Drinking water sources	**Drinking water sources**
Piped water on premises Public taps or standpipes Tube wells or boreholes Protected dug wells Protected springs Rainwater collection	Surface water + Unprotected dug well Unprotected spring Cart with small tank/drum Surface water Bottled water
Sanitation facilities	**Sanitation facilities**
Flush/pour flush to: Piped sewer system Specific tank Pit latrine Ventilated improved [VIP] pit Pit latrine with slab Composting toilet	Open defecation Shared sanitation facilities* Pit latrines without a slab or platform Hanging latrines Bucket latrines
+ Surface drinking water sources include river, dam, lake, pond, stream, canal, irrigation channels. * Sanitation facilities of an otherwise acceptable type shared between two or more households are shared sanitation facilities. Shared facilities include public toilets.	

When the variables have been reclassified, we assign categories values with 0 and 1. It is important to keep a record of how the variables have been recoded which the analyst can track and refer to during the analysis process. The record can also help the analyst adjust the recoding accordingly over time.

[5] Ibid.

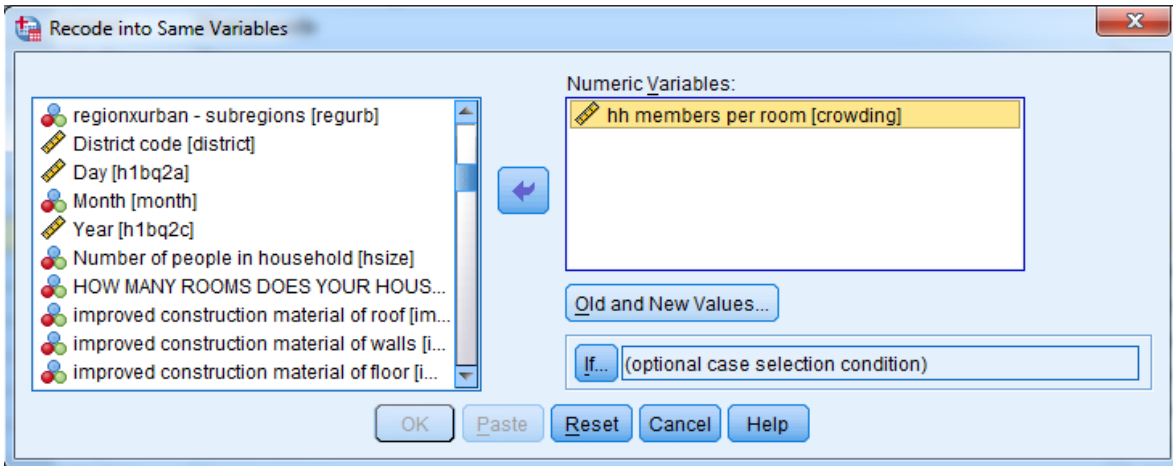
Name of the Variable	Wealthier	Poorer
Material of the house	1 = concrete or wood	0 = mud or thatch
Roof material	1= tiles or galvanized iron or concrete	0 = mud or thatch or plastic
Crowding	1 = 5 or fewer people per room	0 = 6 or more people per room
Type of lighting	1 = electricity or gas	0 = candle or wood
Source of water	1 = piped into dwelling or borehole with pump or protected dug well	0 = pond or unprotected well
Toilet facilities	1 = flush or ventilated improved latrine	0 = open pit or none (bush field)
Has a sewing machine	1 = yes	0 = no
Has a radio	1 = yes	0 = no
Has a TV	1 = yes	0 = no
Has a stove	1 = yes	0 = no
Has a fridge	1 = yes	0 = no
Has a mobile phone	1 = yes	0 = no
Has a bicycle	1 = yes	0 = no
Has a motorbike	1 = yes	0 = no
Has a car	1 = yes	0 = no

SPSS: Transform → Recode into same variables

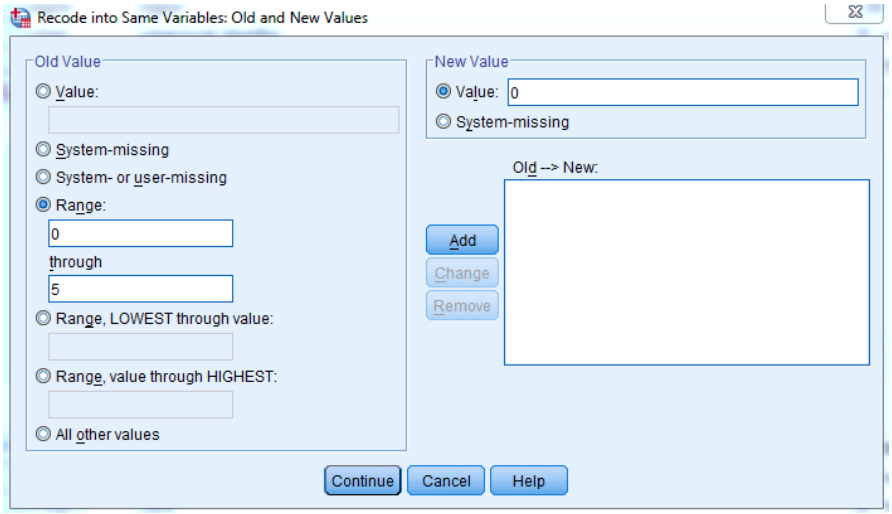


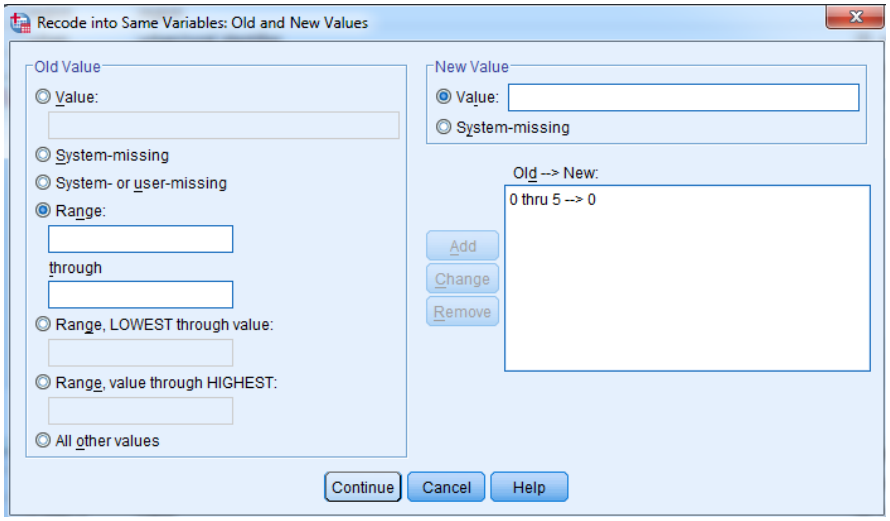
When the dialogue box pops up, select the variable ('crowding' which indicates the number of household members per room in this example) we want to recode from the left field and click the

arrow to move it into the right field. Then click the button '**Old and New Values**' to set the new values.

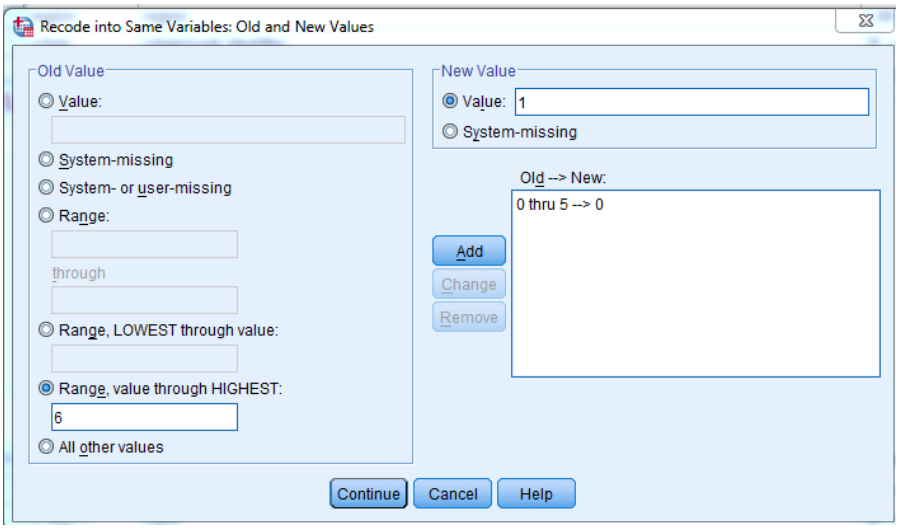


According to the binary classification for 'crowding' (see chart on P10), click '**Range**' in the '**Old Value**' field and enter '0' through '5'. Then click '**Value**' in the '**New Value**' field and enter '0'. Now we recode the 'number of household members per room ≤ 5' into '0'. Click the button '**Add**' to record this recoding.





In the same way, we recode the 'number of people per room ≥ 6' into '1'. Click '**Add**'. Recoding for this variable is now completed.



Click '**Continue**' to return to the dialogue box at the beginning of recoding. You can choose to recode other variables. Click '**OK**' to leave after all recoding is completed.

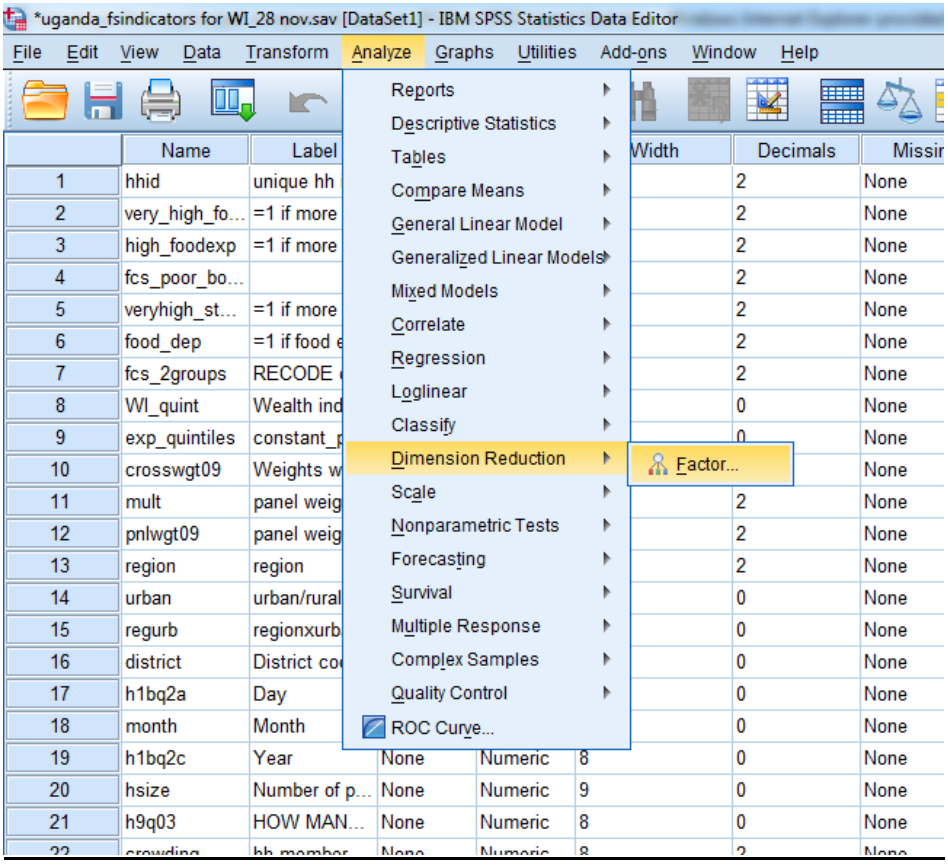
Step 4: Principal Component Analysis (PCA)

Once the variables have been selected, we need to run a principal component analysis to create the wealth index.

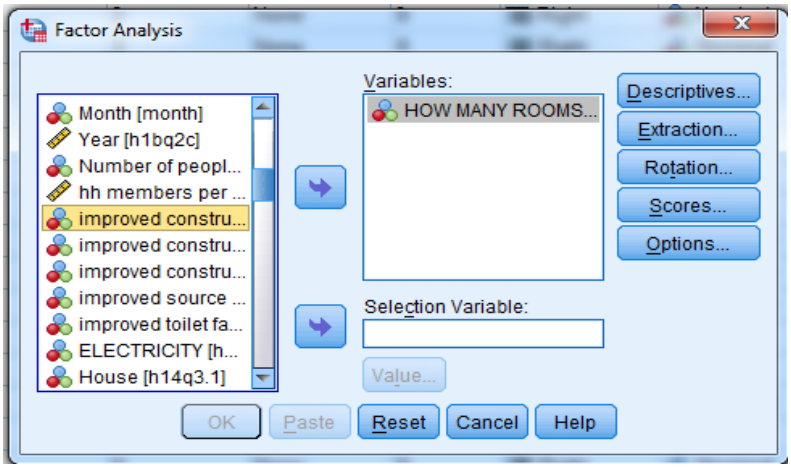
PCA is a 'data reduction' procedure. It involves replacing *many* correlated variables with a set of *principal* uncorrelated 'principal components' which can explain much of the variance and represent unobserved characteristics of the population. The objectives of a PCA are: i) to discover or reduce the dimensionality of the data set and ii) to identify new meaningful underlying variables. **The first principal component explains the largest proportion of the total variance and it is used as the wealth index to represent the household's wealth**.

In SPSS the factor analysis procedure is used to calculate the principal component. This procedure first standardizes the indicator variables by calculating the Z-scores. Then the factor coefficient scores which are also the factor loadings are generated. The indicator values are multiplied by the loadings and summed to the household wealth index. The wealth index as created is a continuous variable which can be used in correlations or regression models. The higher the score of the index, the wealthier the household.

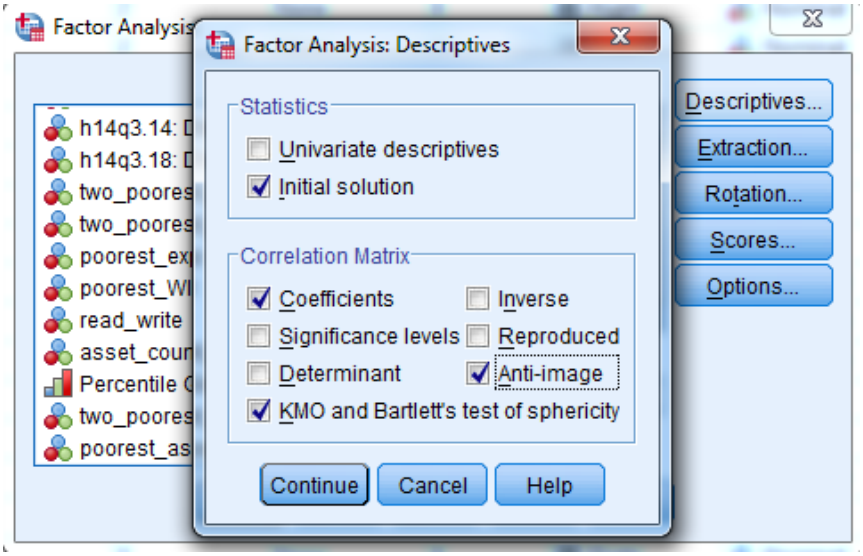
SPSS: Analyse > Dimension Reduction > Factor



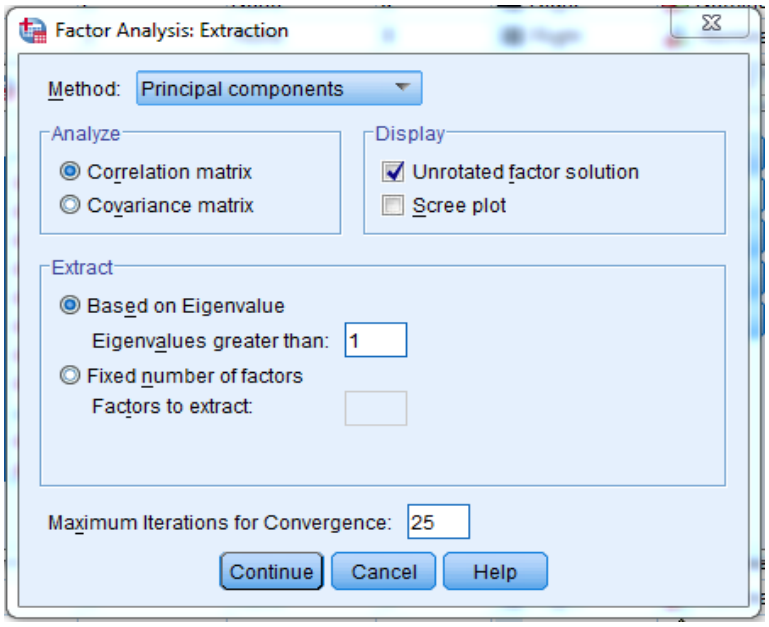
The '**Factor Analysis'** dialogue box will pop up. Click to select all the variables we want to include in the factor analysis (all asset variables in this example) and press the arrow to move them into the right '**Variables**' field.



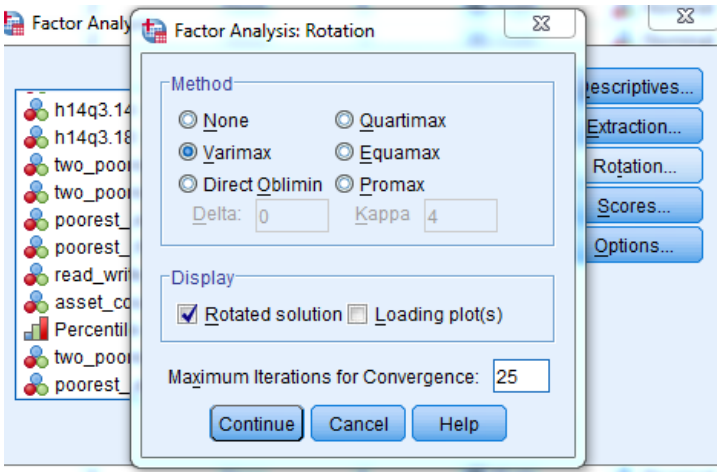
Click the '**Descriptives**' button to enter the **'Factor Analysis: Descriptives**' dialogue box. Check **Initial Solution, Coefficients, KMO and Bartlett's test of sphericity and Anti-image**. Click 'Continue' to return to the dialogue box.



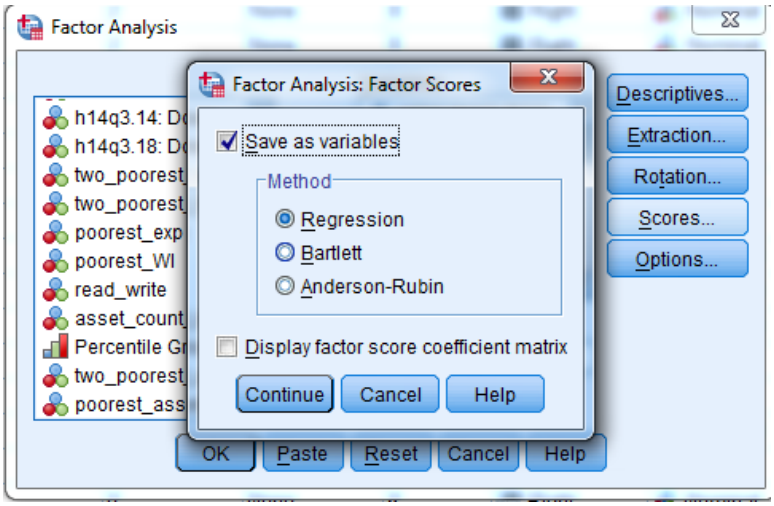
Click '**Extraction**' and check **Correlation Matrix, Unrotated Factor Solution, Scree Plot and Eigenvalues greater than 1**. Click 'Continue'.



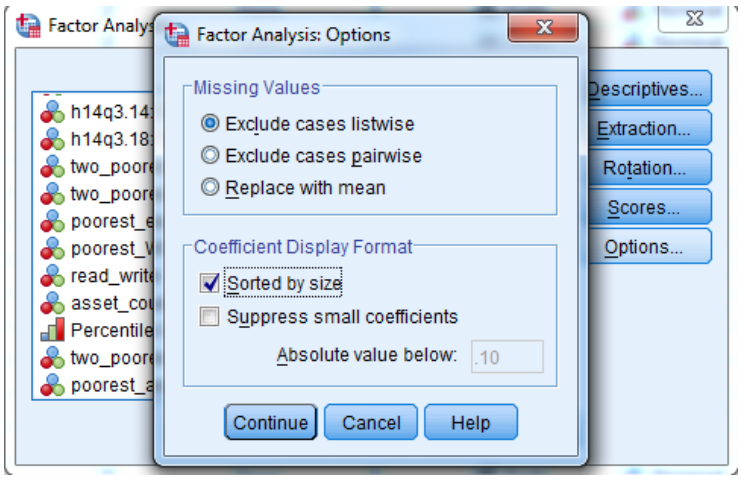
Click '**Rotation**'. Select **Varimax and Rotated Solution**. Click 'Continue'.



Click **'Scores'**. Select '**Save as variables**' in '**Regression**' method. Click 'Continue'.



Click '**Options**'. Check '**Exclude Cases Listwise'** and '**Sorted by Size'**. Click 'Continue'.



At this point, SPSS has completed the Principal Component Analysis.

In the SPSS output viewer[6], we are going to encounter the correlation matrix of the selected variables. If any of the correlations are too high (above 0.9), you may need to remove one of the variables, as the two variables are likely to be measuring the same thing. Another alternative would be to combine the two variables in some way (perhaps by taking the average). If the correlations are too low, (below 0.1), you may also remove it. If others do not behave as expected, one has to investigate this and maybe remediate the problem. For example, owning a radio turned out to be lower among the wealthiest in a country A, because these households often have radio-cassette players instead. The solution was to create a new variable: "radio or radio-cassette player."

In the correlation matrix of this example, we can see that the variable 'Boat' and 'Own_cattle' may be considered for removal.

Correlation Matrix

		HOW MANY ROOMS DOES YOUR HOUSE HOLD OCCUPY?03	improved construction material of roof	improved construction material of walls	improved construction material of floor	improved source of water	improved toilet facility	ELECTRICITY	House	Other Buildings	Land	Furniture/Furnishings
Correlation	HOW MANY ROOMS DOES YOUR HOUSE HOLD OCCUPY?03	1.000	.446	.103	.313	-.007	.365	.196	.179	.142	.070	.197
	improved construction material of roof	.446	1.000	.079	.382	-.063	.271	.197	.028	.008	.010	.266
	improved construction material of walls	.103	.079	1.000	.362	.201	-.006	.191	-.172	.147	-.008	.155
	improved construction material of floor	.313	.382	.362	1.000	.155	.133	.442	-.058	.057	-.089	.122
	improved source of water	-.007	-.063	.201	.155	1.000	-.071	.152	.008	-.045	-.024	-.072
	improved toilet facility	.365	.271	-.006	.133	-.071	1.000	.057	.170	.113	.065	.185
	ELECTRICITY	.196	.197	.191	.442	.152	.057	1.000	-.041	.004	-.059	.077
	House	.179	.028	-.172	-.058	.008	.170	-.041	1.000	-.032	.175	-.022
	Other Buildings	.142	.008	.147	.057	-.045	.113	.004	-.032	1.000	.164	.167
	Land	.070	.010	-.008	-.089	-.024	.065	-.059	.175	.164	1.000	.088
	Furniture/Furnishings	.197	.266	.155	.122	-.072	.185	.077	-.022	.167	.088	1.000
	Household Appliances e.g. Kettle, Flat iron, etc.	.214	.167	-.022	.217	-.004	.151	.339	.040	.013	-.049	.091
	Televsion	.206	.200	.182	.428	.113	.088	.682	.028	.036	-.019	.074
	Radio/ Cassette	.247	.247	.164	.243	.008	.180	.140	.033	.126	.065	.272
	Bicycle	.129	.009	.215	.083	.037	.153	-.027	.012	.175	.122	.235
	Motor cycle	.132	.123	.110	.178	.008	.097	.120	.064	.109	.010	.072
	Motor vehicle	.195	.112	.102	.256	.071	.055	.331	.033	.105	-.016	.039
	Boat	-.072	-.002	.031	-.064	-.048	-.083	-.032	.013	-.009	-.004	-.035
	Jewelry and Watches	.179	.154	.137	.278	.074	.035	.268	-.012	.074	-.047	.104
	Mobile phone	.288	.299	.180	.328	.029	.194	.268	.054	.133	.007	.226
	Computer	.131	.100	.077	.165	.055	.015	.305	.020	.019	.052	.029
	Other electronic equipment	.084	.098	.093	.208	.044	.019	.384	-.024	.023	.028	.046
	Other household assets e.g. lawn mowers, etc.	.106	.176	-.151	.041	-.166	.133	-.013	.066	.111	.127	.113
	own_cattle	.170	-.011	.113	.088	.007	.082	.053	.080	.157	-.006	.070

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy varies between 0 and 1. The values that are closer to 1 are better. A value of **0.6** is a suggested minimum acceptable value. In our example, we have a value of 0.803, which is satisfactory.

[6] See more interpretation of SPSS PCA outputs at http://statistics.ats.ucla.edu/stat/spss/output/principal_components.htm.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.803
Bartlett's Test of Sphericity	Approx. Chi-Square	14155065.780
	df	276
	Sig.	.000

The Component 1 is used as the wealth index as it accounts for the largest proportion of the variance. In the example dataset, the PCA generates a variable labelled as 'REGR factor score 1 for analysis 1' which is the wealth index.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.623	30.189	30.189	3.623	30.189	30.189
2	1.464	12.198	42.387	1.464	12.198	42.387
3	1.025	8.546	50.932	1.025	8.546	50.932
4	.938	7.814	58.746			
5	.887	7.395	66.141			
6	.840	7.001	73.142			
7	.725	6.045	79.188			
8	.679	5.656	84.844			
9	.593	4.945	89.789			
10	.545	4.545	94.334			
11	.387	3.224	97.559			
12	.293	2.441	100.000			

Extraction Method: Principal Component Analysis.

Initial Eigenvalues: 'Eigenvalues are the variances of the principal components. Because we conducted our principal components analysis on the correlation matrix, the variables are standardized, which means that each variable has a variance of 1, and the total variance is equal to the number of variables used in the analysis, in this case, 12'.

Total: 'This column contains the Eigenvalues. The first component will always account for the most variance (and hence have the highest Eigenvalue)'.

Extraction sums of squared loadings: 'The three columns of this half of the table exactly reproduce the values given on the same row on the left side of the table. The number of rows reproduced on the right side of the table is determined by the number of principal components whose Eigenvalues are 1 or greater'. (Sources: IDRE, UCLA)

In the output, we can also see the component matrix. This table contains the component loadings, which are the correlations between the variable and the component. Possible values range from -1 to +1.

Component Matrix[a]

	Component		
	1	2	3
HOW MANY ROOMS DOES YOUR HOUSE HOLD OCCUPY?03	.353	.553	-.271
improved construction material of roof	.543	.370	-.217
improved construction material of walls	.473	-.090	.618
improved construction material of floor	.757	-.151	.186
ELECTRICITY	.732	-.438	-.146
Furniture/Furnishings	.298	.524	-.104
Household Appliances e. g. Kettle, Flat iron, etc.	.570	-.275	-.451
Televsion	.722	-.404	-.115
Radio/ Cassette	.442	.480	.108
Motor cycle	.302	.132	.464
Jewelry and Watches	.484	-.004	.001
Mobile phone	.645	.192	.109

Extraction Method: Principal Component Analysis.

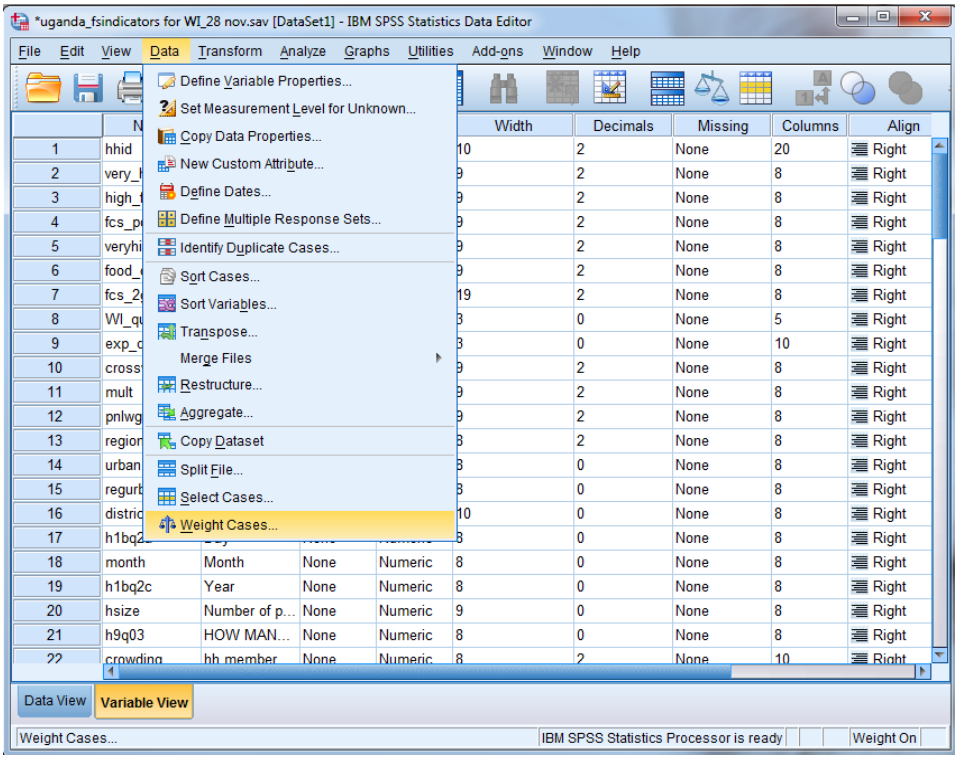
a. 3 components extracted.

** The construction of the wealth index is an **iterative process**. To obtain the best results, we usually need to conduct a few rounds of PCA including or excluding certain variables based on the factor coefficient scores we see in the PCA outputs.

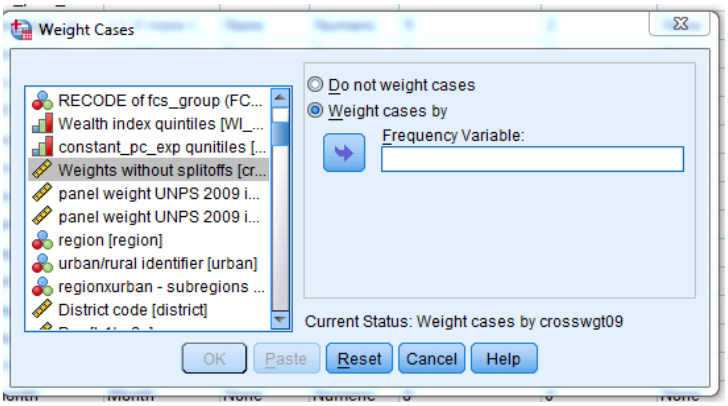
Step 5: Create wealth index quintiles

Before we rank the wealth index and create the quintiles, we need to turn weights on to correct over/under sampling of any group.

SPSS: Data→ Weight Cases

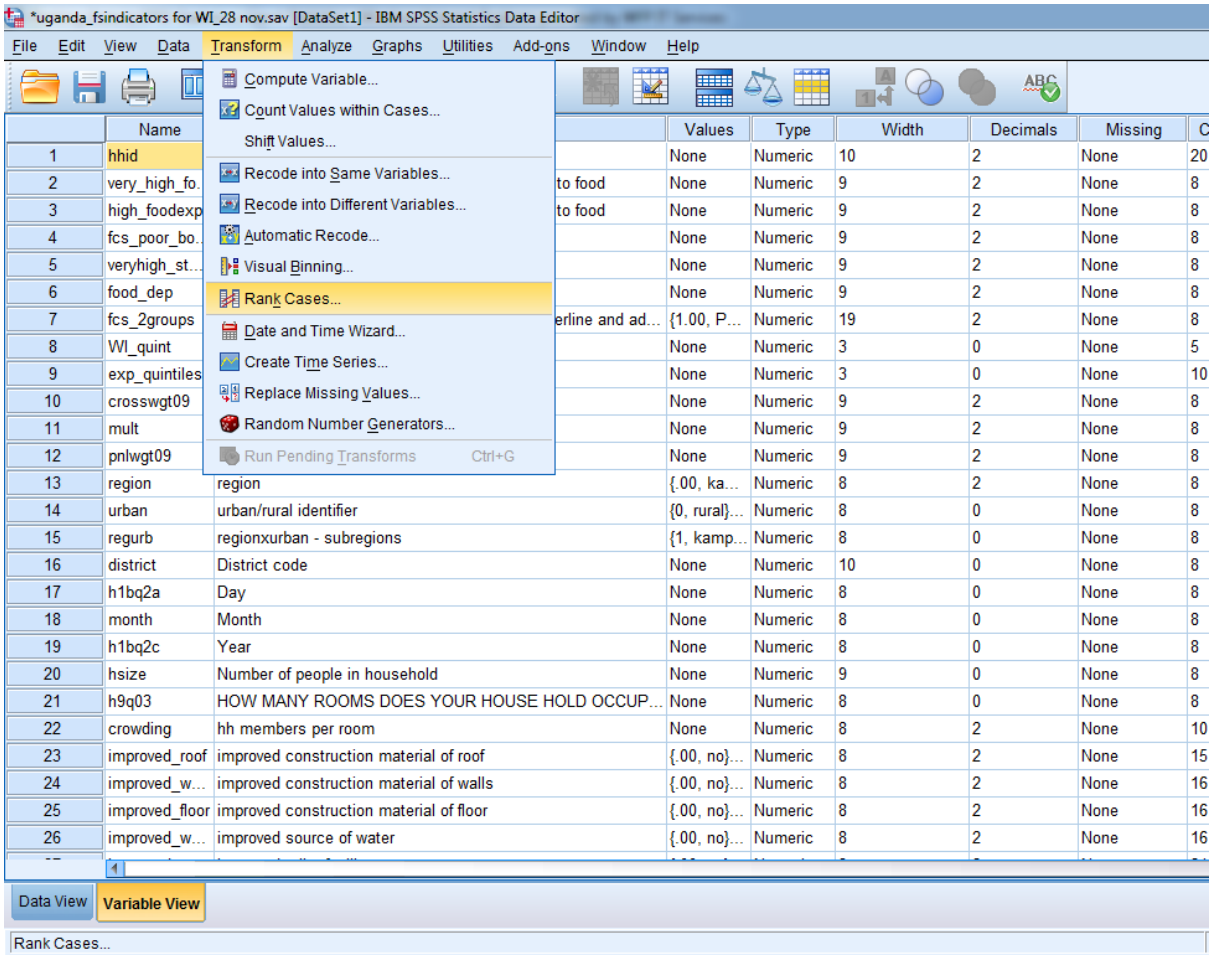


In the dialogue box that pops up, select the option '**Weight cases by'**. Click on the weighting variable in the left field and click on the right arrow to move this variable into the right field called **'Frequency Variable'**.

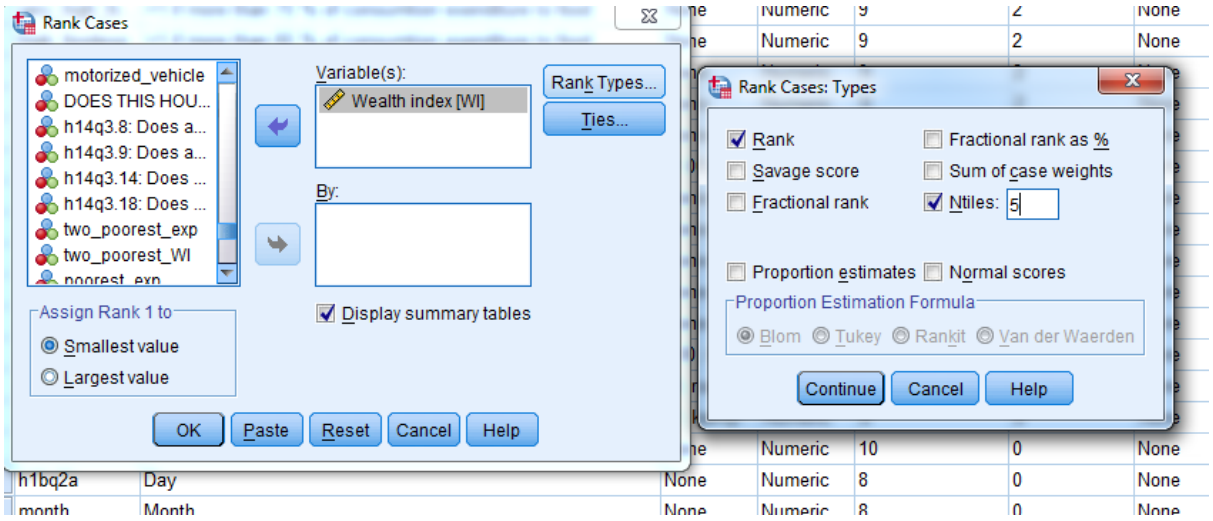


In order to better understand the wealth index, which is a continuous variable, it is useful to recode the index into a categorical variable. The best way to do it is to rank the WI (the first variable created from the PCA) into deciles or quintiles, dividing all households into five or ten equal groups. In the SPSS demonstration, we rank the WI into quintiles.

SPSS: Transform > Rank Cases



Select 'wealth index' in the left field and click the arrow to move it to the right '**Variable(s)**' field. Then click '**Rank Types**' to open the dialogue box. Check '**Rank**' and '**Ntiles**'. Set the number of Ntiles at **5** as we are ranking in **quintiles.**



Click 'Continue' and 'OK'. Two variables will be created and appear in the dataset. The variable called "percentile group of WI" is our wealth quintile variable.

Created variables[a]

Source variable	Function	New variable	Label
WI[b]	Rank	RAN001	Rank of WI
	Percentile group[c]	NTI001	Percentile group of WI

a. Mean rank of tied values is used for ties.

b. Ranks are in ascending order.

c. 5 groups are generated.

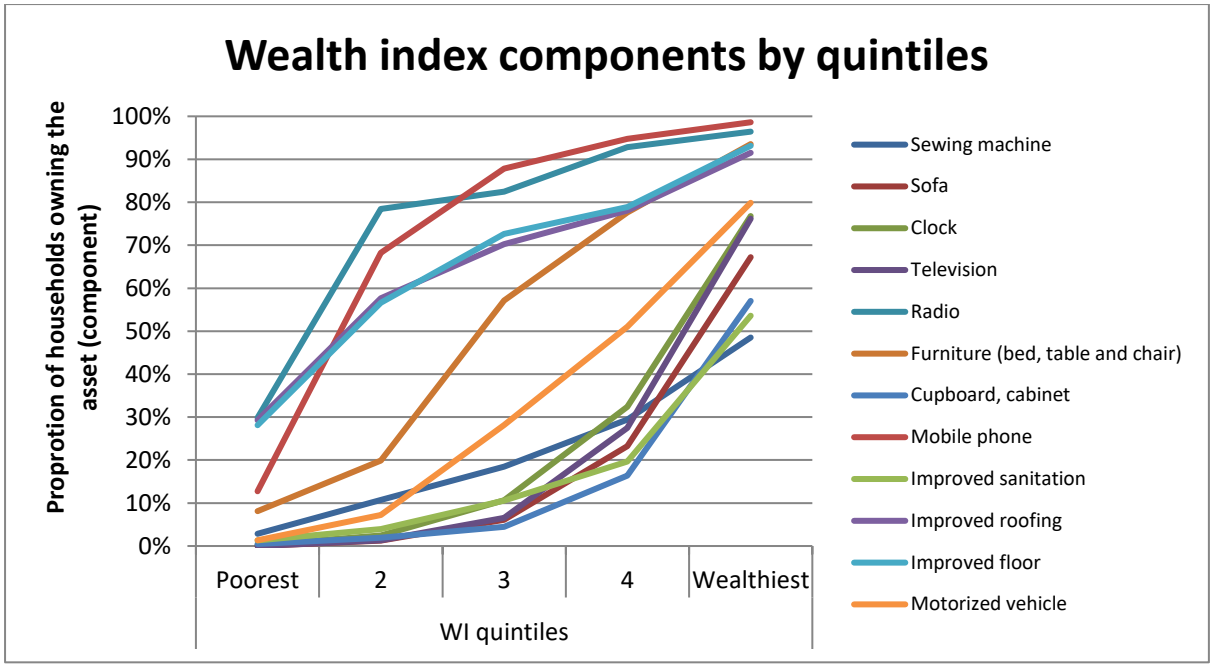
Step 6. Graph the wealth index

Graphing these quintiles or deciles by the variables included in the PCA will help the analyst to have an immediate idea about whether those variables are appropriate for the wealth index construction. Analysts can perform this step after running the PCA as a final check if the most desirable variables have been included in the wealth index.

To create the graph, run a cross-tabulation between the new categorical wealth index quintiles or deciles and the variables (assets and housing characteristics used in the analysis). This will show the prevalence of households that own the selected assets in each quintile. There should be a positive relationship between the independent and dependent variable. If there are variables included where this is not the case, they should be examined further and considered for removal.

In addition, the analysts may find it very useful to do this cross-tabulation by urban and rural areas separately as a double check on how this national index applies to both places of residence. If the variables in either rural or urban areas show insignificant or opposite patterns from what we expect, we may consider reconstructing the index or creating separate ones for both of them.

In the example graph below, ownership of every variable included in the PCA increases as the quintiles go from poorest to wealthiest. This indicates that the variables included for the PCA are appropriate.



Step 7. Select the final result and report the variables

After PCA procedures and wealth index graphing, we will have a good idea of the most appropriate wealth index. A rule of thumb to understand if the index created is appropriate is to run a correlation between the two latest first factors (of 2 different principal component analyses). If their correlation coefficient is close to 1 (0.998/0.999) then the two indicators are very similar and the wealth indices are also very similar.

When a satisfactory wealth index has been created this can be used for further analysis in relation to other indicators.

Additional Reading

Deaton and Zaidi, 2001: "Guidelines for Constructing Consumption Aggregates

For Welfare Analysis" Washington, DC: World Bank (2002)

Filmer and Prichett: "Estimating Wealth Effects without Expenditure Data -- or Tears:

An Application to Educational Enrollments in States of India" Demography. Feb;38(1):115-32.(2001)

Fimler and Scott, 2008: "Assessing Asset Indices" Policy Research Working Paper 4605, WB

Howe, L.D., Heargreave, J.R., Gabrysch S., and Huttly. S.R.A., 2009: "Is the Weath Index a Proxy for Consumption Expenditure? A Systematic Review" Journal of Epidemiol Community Health, 2009: 63, 871-880

Rutstein, 2008 : «The DHS Wealth Index: Approaches for Rural and Urban Areas" DHS Working Papers No. 60

McKenzie, D.: "Measuring Inequality with Asset Indicators" Journal of Population Economics 18(2): 229-260 (2005).