

Annex: Data Sources and Technical Notes for the Preliminary 2012 Subnational Human Development Indices

The Human Development Index (HDI) is a summary measure of human development. It measures the average achievement in a country in three basic dimensions of human development: longevity or a long and healthy life, access to knowledge, and a decent standard of living. These dimensions are measured by a set of indicators that are aggregated into indices.

Indicators for each of the dimension are sourced from available secondary data collected by government agencies (see Table below).

Dimension	Indicator	Source
Longevity	Life expectancy at birth	Flieger and Cabigon [1999] Cabigon [2009]
Knowledge	Mean years of schooling Expected years of schooling	Labor Force Survey (LFS), Philippine Statistics Authority (PSA)
Standard of living	Per capita income	Family Income and Expenditures Survey (FIES), PSA Regional Consumer Price Index (RCPI), PSA 2012 Provincial Poverty Threshold, PSA

The life expectancy at birth for 2012 were generated using estimated life expectancy at birth from Cabigon (2009) and population figures based on actual values from Census of Population and linear projections using the 1995 and 2000 figures.

As for education, the mean years of schooling is the average years of prior schooling for adults aged 25 and older in a province while the expected years of schooling is the sum of the enrolment rate per age level from ages 6 to 24.

To compute for the mean years of schooling an equivalent number of years were assigned to each response under the “*highest grade completed*” variable (i.e., 1 year for “*Grade 1*”, and so forth, to 16 years for “*Graduate in Masters/PhD*”) in the LFS. Weighted average of the number of years of schooling among adults aged 25 and older is then generated for each province.

Expected years of schooling is calculated using the weighted sum of the complete number of years of a school-age child of a particular age (which is equal to 1), where the weights are the probability of being enrolled in school at age i (6 to 24), which is assumed equal to the current enrolment ratio at age i (6 to 24) in the LFS. [Klugman, J. et. al., 2012]

The real per capita income is the weighted mean total family income divided by family size converted in NCR 2012 pesos. This was computed using the trimmed¹ 2012 FIES data, converted to 2012 prices using RCPI (for FIES years other than 2012) and adjusted by the ratio of 2012 poverty thresholds of the province with Metro Manila, which serve as base.

Before the HDI itself is calculated, an index needs to be created first for each of the three dimensions. To calculate these dimension indices, minimum and maximum values (goalposts) are chosen for each

¹Symmetric trimming was done, i.e., 0.5% of upper and lower values.

underlying indicator. The index for each dimension is then expressed as a value between 0 and 1 by applying the general formula:

$$\text{Dimension Index} = \frac{\text{actual value} - \text{minimum value}}{\text{maximum value} - \text{minimum value}}$$

Global goalposts are adopted except in the case of education and income indices, specifically, for the maximum for the combined education index and maximum and minimum for the income index. For the education index, the actual maximum value observed from 1997 to 2012 is used. For the income index, the minimum is set at 90 percent of the actual minimum value observed while the maximum is set at 110 percent of the actual maximum value observed. The latter is to avoid an undefined estimate when the general formula is applied. The goalposts used are as follows:

Indicator	Maximum	Minimum
Life expectancy at birth, years	85.0 (Global)	20.0 (Global)
Mean Years of Schooling	11.52 (Batanes, 2008)	0.0
Expected Years of Schooling	14.6 (Benguet, 2012)	0.0
Combined Education Index	0.930 (Benguet, 2012)	0.0
Real per capita income, 2012 NCR Pesos	105,761* (based on Metro Manila, 1997)	17,255* (based on Lanao del Sur, 2012)

*Maximum is 110 percent of actual maximum; minimum is 90 percent of actual minimum

Indices for both subcomponents of education are first calculated before computing their geometric mean. The combined education index is then computed by applying the general formula with minimum equal to 0 and maximum equal to the highest geometric mean observed among provinces (i.e., Benguet in 2012).

The data for Metro Manila is used as an illustration.

Dimension Index	Estimate (2012)
Life expectancy (I_{Life})	$(73.8 - 20)/(85.0 - 20) = 0.827$
Education	
▪ Mean years of schooling	$(11.02 - 0)/(11.52 - 0) = 0.957$
▪ Expected years of schooling	$(12.60 - 0)/(14.60 - 0) = 0.866$
▪ Combined education ($I_{\text{Education}}$)	$[(0.957 \times 0.866)^{1/2} - 0]/(0.930 - 0) = 0.979$
Income (I_{Income})	$(79,488 - 17,255)/(105,761 - 17,255) = 0.703$

Once the dimension indices have been calculated, the HDI is determined by computing the geometric mean of the three dimension indices:

$$\begin{aligned} \text{HDI}_{\text{MM}} &= (\text{I}_{\text{Life}} \times \text{I}_{\text{Education}} \times \text{I}_{\text{Income}})^{\frac{1}{3}} \\ \text{HDI}_{\text{MM}} &= (0.827 \times 0.979 \times 0.703)^{\frac{1}{3}} \\ \text{HDI}_{\text{MM}} &= \mathbf{0.829} \end{aligned}$$

As it is for every new set of sub-national HDIs - due to the updating of goalposts - these **2012 estimates are not comparable to previously published provincial HDIs**, i.e. in past volumes of the Philippine Human Development Report. Back-computations to 1997 (and for the FIES survey years of 2000, 2003, 2006, 2009), for purposes of comparability over time, will be featured in the next volume of the Philippine Human Development Report (PHDR).

The above computations adopt the interim methodology agreed upon by the former National Statistical Coordination Board (now part of PSA) and HDN (in February 2012) and previously applied to the 2009 subnational HDI. The methodology is still subject for review by the relevant technical/interagency committees being coordinated by the PSA and for approval by the PSA Board.

Data Limitations

HDI estimates are subject to the same limitations as all survey-based province-level estimates released by the PSA (e.g. poverty statistics).

In particular, the LFS and FIES dataset, which were used in the estimation of the indicators of 2 of the 3 subcomponents of HDI, are based on the 2003 Master Sample which has as its sampling domain, the Region and not the Province. Hence, estimates for some provinces have high standard errors and/or high coefficients of variation.

For instance, 7 out of the 79 provinces have sample sizes in the 2012 FIES (and LFS) which are less than 100. These provinces are Aurora, Apayao, Batanes, Camiguin, Catanduanes, Guimaras, and Siquijor. Small sample sizes affect the precision and reliability of resulting estimates.

In order to better guide users of the data therefore, measures of *dispersion* (i.e. scatter about an average value) of the HDI component data, where applicable, are presented along with the HDI estimates. These measures are the standard errors, coefficients of variation, and confidence intervals (refer to Tables 2 and 3). The measures of dispersion of the indices shall be released in the next volume of the PHDR.

The **standard error**(SE) is a measure of absolute dispersion and is used as a measure of variability of one data set.²If there is a small amount of variation in the data set, then data values will be close to the mean and standard errors will be small. That is, the lower the standard error, the more precise the estimate.

The **coefficient of variation** (CV) is a measure of relative dispersion, used to compare two or more data sets with different means and/or different units of measurement. As a measure, it utilizes the mean and

²See also 2012 FIES Technical Notes at <https://psa.gov.ph/content/filipino-families-poorest-decile-earn-six-thousand-pesos-monthly-average-2012-results-2012>

standard error and is expressed as a percentage; it is unitless. At present, there is no internationally agreed cut-off point for what are 'acceptable' levels of the CV.³ What is important however is for CVs to be fully disclosed so that users may be guided in their interpretation of estimates. In the computation of mean per capita income in PPPNCR 2012 pesos for instance, 16 out of the 80 provinces have CVs greater than 10 percent but less than 15 percent (see Table 2); 3 more have CVs between 15 percent and 20 percent (i.e. Ifugao, Apayao, and Guimaras); and 2 provinces have CVs between 30 percent and 33 percent (i.e. Siquijor and Aurora). If a 10 percent cut-off is used, then the estimated values for 21 provinces would be considered as 'unreliable'. If a 15 percent cut-off is used, then the estimate for 5 provinces would be considered as 'unreliable.'

Reporting the standard error provides some idea of the accuracy of an estimate (of say, mean per capita income) from a particular sample, but does not really tell us how far that sample estimate is likely to be from the population value (or 'true' mean). For this purpose, an *interval estimate* of the true mean is appropriate which the confidence interval (CI) provides. The CI is computed using the sample mean and standard error; the larger the 95 percent confidence interval, the more caution is required when using the estimated sample mean. The technical interpretation of a confidence interval is as follows: Say the constructed confidence interval for the true mean per capita income of Catanduanes is [P27,634, P41,142]. This means that for 95% of random samples (drawn from the population), the constructed interval [27,634, 41,142] will contain the true mean.

A number of provinces have overlapping CIs of their mean per capita income (in PPP NCR 2012 pesos), mean years of schooling and expected years of schooling. Thus, ranking of provinces on these bases should always be undertaken with caution.

³Some developed countries like Australia and the United States have released to the public, statistics with CVs higher than ten percent. To quote from the Australian Bureau of Statistics' publication on Crime and Safety, Australia, for instance, "For tables in the publication, only estimates with relative standard errors (RSEs) of 25% or less, and percentages based on such estimates, are considered sufficiently reliable for most purposes".