

Corruption Perceptions Index 2016: Short Methodology Note

The Corruption Perceptions Index aggregates data from a number of different sources that provide perceptions of business people and country experts of the level of corruption in the public sector.

The following steps are followed to calculate the CPI:

- 1. Select data sources:** Each data source that is used to construct the Corruption Perceptions Index must fulfil the following criteria to qualify as a valid source:
 - Quantifies perceptions of corruption in the public sector
 - Be based on a reliable and valid methodology, which scores and ranks multiple countries on the same scale
 - Performed by a credible institution and expected to be repeated regularly
 - Allow for sufficient variation of scores to distinguish between countries

The CPI 2016 is calculated using 13 different data sources from 12 different institutions that capture perceptions of corruption within the past two years. These sources are described in detail in the accompanying source description document.

- 2. Standardise data sources** to a scale of 0-100 where a 0 equals the highest level of perceived corruption and 100 equals the lowest level of perceived corruption. This is done by subtracting the mean of the data set and dividing by the standard deviation and results in z-scores, which are then adjusted to have a mean of approximately 45 and a standard deviation of approximately 20 so that the data set fits the CPI's 0-100 scale. The mean and standard deviation are taken from the 2012 scores, so that the rescaled scores can be compared over time against the baseline year.
- 3. Calculate the average:** For a country or territory to be included in the CPI, a minimum of three sources must assess that country. A country's CPI score is then calculated as the average of all standardised scores available for that country. Scores are rounded to whole numbers.
- 4. Report a measure of uncertainty:** The CPI is accompanied by a standard error and confidence interval associated with the score, which capture the variation in scores of the data sources available for that country/territory.