

Changes in DOSE in V2.9 (September 17 September 2024)

DOSE V2.9 is an update of DOSE V2. We made a few corrections and additions to the data where various users had identified gaps or inaccuracies.

1. We added missing output data for the following regions:

- Azerbaijan, all regions | 2016
- Brazil, all regions | 2010 (only aggregate GDP was missing)
- UK, all regions | 1994-1996
- Spain, regions ESP.8_1, ESP.9_1, ESP.10_1 | 1986 – 1992
- India, regions IND.1_1, IND.6_1, IND.10_1, IND.11_1, IND.13_1, IND.21_1, IND.23_1, IND.30_1, IND.33_1 | 2014

2. We corrected output values for the following observations where there were errors in the original calculation steps:

- IDN.12_1 (2013-2020)
- UZB, all regions, years 1995-1999 were corrected because of an error in the conversion from raw data in PPP USD into local currencies
- SWE.1_1 (1960, 1970, 1980), SWE.9_1 (1980, 1990), SWE.11_1 (1980), SWE.17_1 (1960, 1970, 1980, 1990), SWE.19_1 (1980,1990)
- BRA .14_1 (2010), BRA.16_1 (2010), BRA.25_1 (2010), BRA.26_1 (2010)
- UKR.12_1 and UKR.11_1 in 1995-2015
- KGZ.6_1/KGZ.7_1 2006-2015
- ARE.1_1 and ARE.2_1, all years
- KEN.22_1 and KEN.25_1 in 1997

3. We traced down a mistake in the agr_gdp_pc column for regions from India (after 2014), which was calculated at first using population figures that were one row off. We fixed this in the raw data and then re-added the correct Indian data to Dose.

4. We substantially increased the amount of data in the gdp_pc_lcu2015_usd (and sectoral versions) rows which were missing due to a mistake in the formula causing the column to be blank for any regions without a 2015 observation.

5. We noticed all observations from the year 2020 were missing the data for the columns `gdp_pc_usd` and `gpd_pc_usd_2015` (and their sectoral versions) due to our source for exchange rates, Penn World Tables 10, stopping in 2019. To address this, we added in exchange rates for 2020 from the World Bank (<https://data.worldbank.org/indicator/PA.NUS.FCRF>) and were then able to fill in all the USD columns for 2020 observations. We verified that the World Bank exchange rates closely match (and are in most cases identical) to the rates from PWT.
6. We added missing population values for 'Comunidad del Madrid' (ESP.8_1) from 1995-1999 from the same source as the other regions. The other columns were not affected by this change.
7. For MAR.10_1, the GID_1 identifier was corrected to MAR.22_1 in the period 2013 to 2019.
8. We noticed missing structural dummies for the start of a new datasource in Colombia (2000) and for changing district boundaries in Vietnam (2004-2008). These were added to the dataset.
9. Because of data irregularity that we couldn't verify, we removed the Belgian data before 2003.