

Documentation for
**GDP Per Capita by
Purchasing Power Parities**
for countries and territories

Gapminder Documentation constitutes work in stepwise progress.
We welcome all sorts of comments, corrections and suggestions through e-mail to the author.

Gapminder Documentation 001

Version 3

Uploaded: 2008-10-08

Published by: The Gapminder Foundation,
Sweden, Stockholm, 2008

Author: Mattias Lindgren

E-mail: [mattias.lindgren \(at\) Gapminder.org](mailto:mattias.lindgren@gapminder.org)

Gapminder is a non-profit foundation
promoting sustainable global development
by increased use and understanding of statistics.
www.gapminder.org



1. Introduction

This is the documentation for the Gapminder compilation of GDP per capita by PPP. For more information please go to:

<http://www.gapminder.org/downloads/documentation/#gd001>

The ambition has been to include as many countries and territories as possible. A guiding principle has been to use the most transparent and well-documented source whenever possible, however, preference was also given to sources with as wide coverage of countries and territories as possible.

The main purpose of the data is to produce graphical presentations that display the magnitude of income disparities in the world over time. Therefore, we have also included rough estimates for countries and territories for which reliable data was not available. These estimates can only be taken as an indication of the order of magnitude for the indicator. Furthermore, we have not been able to make sure that every single observation is based on the best estimates available. Hence we discourage the use of this data set for statistical analysis and advise those who require more exact data to investigate the available data more carefully and look for additional sources, when appropriate.

For a discussion on what countries and territories we try to cover, and how we try to handle border changes and the like, see the document “Countries and Territories in Gapminder World”. To be succinct, we will hereafter jointly refer to all countries and other types of geographical entities and territories as “areas”, irrespectively of their statehood status. The inclusion of any area in this data set does, in no way, imply a stated opinion of Gapminder on the legal status of the area.

It is not always clear to what extent certain semi-autonomous or disputed territories also are included in the data for their “mother country”, so there may be some “double counting” in these instances.

2. Sources used

Six major sources have been used:

1. The main source has been the data available at the homepage of Angus Maddison:

www.ggdc.net/maddison/

The data can be found under “Historical Statistics: [World Population, GDP and Per Capita GDP, 1-2003 AD](#) (Updated: August 2007)” and the data can be downloaded as an excel-file from:

www.ggdc.net/maddison/Historical_Statistics/horizontal-file_03-2007.xls

The data is described in Maddison (2003, X, X and X). We used the indicator called “Per Capita GDP (1990 International Geary-Khamis dollars)”. This source will be referred to as “Maddison online” below.

2. The World Development Indicators 2007 (WDI) has an indicator called “GDP per capita, PPP (constant 2000 international \$)”. The data is available online for a fee at <http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:21298138~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html> The site also contain some documentation for the data. This data has been used whenever there was no data in the above source. This source will be referred to as “WDI” below.

Since the data is expressed in the international prices of 2000, rather than of 1990 as in Maddison, the data has to be recalculated. For countries for which there was data for both Maddison and the WDI the data from the WDI was simply chained to Maddison (at the earliest year of Maddison if filling

backward and at the latest year of Maddison when filling forward).

For several countries, however, there was no data at all in Maddison. For these the ratio between WDI and Maddison for USA 1990 was used to deflate the WDI data.

3. The Penn World Tables 6.2 (PWT) has an indicator called “Real GDP per capita (Constant Prices: Laspeyres)” or RGDPL. The data is available online:
http://pwt.econ.upenn.edu/php_site/pwt_index.php The site also contain the documentation for the data. This data has been used whenever there was no data in the two source as above.

Since the data is expressed in the international prices of 2000, rather than of 1990 as in Maddison, the data has to be recalculated. For countries for which there was data for both Maddison and the PWT the data from the PWT was simply chained to Maddison (at the earliest year of Maddison if filling backward and at the latest year of Maddison when filling forward).

For several countries, however, there was no data at all in Maddison. For these the ratio between PWT and Maddison for USA 1990 was used to deflate the WDI data. The discussion for this is equivalent to the one for WDI. This source will be referred to as “PWT” below.

4. In Maddison (2003) there are additional data in the tables 1-4, 4-5 and 5-4 that, for some reasons, are not available in the on-line data set. This data only cover a few scattered years. This data has occasionally been used when no other data was available. The source will be referred to as “Maddison (2003)” below. This should not be confused with “Maddison online” above.
5. The CIA world fact book, available at

www.cia.gov/library/publications/the-world-factbook/index.html

(accessed 2008-01-14). One of their indicators is “GDP per capita (PPP)”. Only one year was available for each country and the year for which the data applies differ from country to country (even though most are for 2005 or 2006). There is no clear definition of the indicator, and no detailed documentation.

If we take the ratio between the CIA data and the relevant data from all the countries for which we have created data from all the other sources they seem to be concentrated somewhere between the ratios of 1,25 to 1,5. This is incidentally where we find the ratio of USA (1,43). Hence we use this ratio to deflate all the CIA data. Note, however, that the data for USA used as the denominator is for 2005, while the numerator (CIA) is for 2006. Furthermore, we are not certain whether the data is expressed in international dollars for a specific year, and what year that supposedly should be. Hence, the data used from this source is particularly shaky and should only be taken as a very rough indicator of the order of magnitude of the GDP per capita.

6. The World Development Indicators 2008 (WDI 2008) has an indicator called “GDP per capita, PPP (constant 2000 international \$)”. The real growth rate between 2003 and 2006 (for 173 areas) or between 2003 and 2005 (for 7 areas) was linked to the data for each country at 2003. If there were existing data for 2004 or 2005 this was replaced with this new data.

3. The different methods used

There have been various combinations of sources for the included areas, and the methods of linking sources together vary somewhat between the areas. The different methods have been grouped into a number of groups which are described below. The updates with WDI 2008 is not described below since it was applied to a majority of the areas.

1. *For 18 areas* Maddison online was the only source. For Qatar we used the average for 'Total Arabia' taken from Maddison (2003), table 5.6, for the years 1820, 1870, 1913.
2. *For 137 areas* Maddison online was the main source, updated with the WDI for 2004 and onwards (WDI was linked to Maddison at 2003). The four French overseas territories Guadeloupe, Martinique, Reunion and French Guyana are, in all likelihood, included in the data for France.¹ Since these four countries are included in the data set it is likely that at least in this case there is some double-counting going on.

For Bahrain, Kuwait, Oman, Saudi Arabia, United Arab Emirates and Yemen we used the average for 'Total Arabia' taken from Maddison (2003), table 5.6 for the years 1820, 1870, 1913.

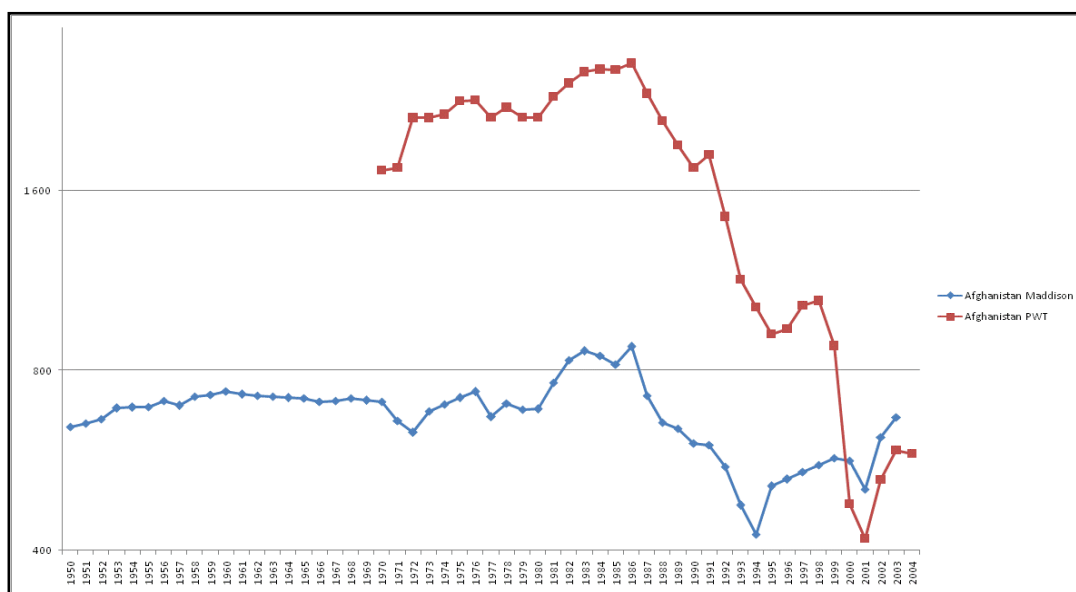
3. *For 5 areas* Maddison online was the main source, updated with the WDI for 2004 and onwards (WDI was linked to Maddison at 2003). Furthermore, the data was updated backwards from 1990 with the WDI (the WDI was linked to Maddison at 1990). The WDI had data back to 1987 (Uzbekistan), 1980 (Estonia and Moldavia), and 1975 (Georgia and Latvia). Overlapping years match well.

For 1973 data from Maddison online was used. The observations for 1973 seem to be in line with earliest available year from the WDI (that has been linked to Maddison). This goes for all five countries, with the possible exception of Latvia.

4. *For the Slovak republic* Maddison online was the main source, updated with the WDI for 2004 and onwards (the WDI was linked to Maddison at 2003) and 1989 and backwards (the WDI was linked to Maddison at 2003). Overlapping years fit quite well.
5. *For French Guiana, Guadeloupe and Martinique* data calculated from the text of Maddison (2003), table 4.5, was used (for four years only).
6. *For Taiwan and Afghanistan* Maddison online was the main source, updated with the PWT for 2004 (PWT was linked to Maddison at 2003).

For Taiwan the overlapping years for the Maddison and the PWT fit well. For Afghanistan, however, there was a sharp drop between 1999 and 2000 in the PWT not shown in Maddison (see figure X below). The years 2000-2003 were quite similar though, so the (linked) PWT was used for 2004.

1 The World Bank data source for GNI per capita, available at XXXX, explicitly states that the four areas are included in France. We have not found any explicit statement of this in the sources we actually use for France, but we cannot see any evidence that these data are not included in the estimate for France.



It should be noted that the PWT also had data for 2004 for Somalia, but here the inconsistencies between the PWT and Maddison was felt to be too large. Hence, Somalia was not updated with the PWT, but is included in the first category above (where Maddison is the only source).

7. For Eritrea and Guyana the WDI was used. To recalculate the WDI to 1990 international dollars all the WDI observations were deflated with the ratio between WDI and Maddison for USA 1990. For Guyana there was data for 1950 in Maddison (2003), but these were not used since it seemed to be too inconsistent with the WDI.
8. For 11 areas the WDI was used, updated backwards with the PWT. To recalculate the WDI to 1990 international dollars all the WDI observations were deflated with the ratio between WDI and Maddison for USA 1990. The PWT was linked to the earliest year available in the WDI.

The earliest year for which there was data in the WDI differed between the countries: Solomon Islands, Iceland, Luxembourg, Suriname and Kiribati 1975; Vanuatu 1979; Ethiopia and Tonga 1982; Samoa and Macao 1983; Micronesia 1987.

For Suriname and Macao there was data for 1950 in the text of Maddison (2003), but these were not used since it seemed to be too inconsistent with the WDI.

9. For 9 areas the WDI was used, updated backwards with the PWT in addition to using data from Maddison (2003) for 1950. To recalculate the WDI to 1990 international dollars all the WDI observations were deflated with the ratio between WDI and Maddison for USA 1990. The PWT was linked to the earliest year available in the WDI. For Malta and Cyprus the data for 1950 was taken from the text of Maddison (2003), table 1-4. For the other areas the data for 1950 was calculated from the text of Maddison (2003) table 4-5.

The earliest year for which there was data in the WDI differed between the countries: Malta, Cyprus, St. Vincent & the Grenadines and Belize 1975; St. Lucia 1980, the other five areas 1977.

10. For the Bahamas the WDI was used, updated backward and forward with the PWT for the years 1970-74 and 2003. To recalculate the WDI to 1990 international dollars all the WDI observations were deflated with the ratio between WDI and Maddison for USA 1990. The PWT was linked to the earliest the WDI at 1975 and 2002 respectively.

11. *For Papua New Guinea* the WDI was used, in addition to using data from the text of Maddison (2003) for 1950. To recalculate the WDI to 1990 international dollars all the WDI observations were deflated with the ratio between WDI and Maddison for USA 1990. The data for 1950 was calculated from Maddison, table 5-4.

There was data for 1970-74 in the PWT, but this was not used since it seemed to be too inconsistent with the WDI. In particular the PWT makes big upward jump between 1975 and 1976 not seen in the WDI.

12. *For Fiji* the WDI was used, updated backwards to 1970-76 with the PWT in addition to using data from the text of Maddison (2003) for 1950. To recalculate the WDI to 1990 international dollars all the WDI observations were deflated with the ratio between WDI and Maddison for USA 1990. The PWT was linked to the WDI at 1977. The data for 1950 was calculated from the text of Maddison (2003), table 5-4. However, since the few years in Maddison (2003) seemed to be consistently higher the data has been deflated with the ratio between the data for 1973 in table 5-4 and the data in the PWT for the same year.

13. *For six areas* the PWT was used. To recalculate the PWT to 1990 international dollars all the PWT observations were deflated with the ratio between PWT and Maddison for USA 1990. For all countries except Palau there was also data for 1950 in Maddison (2003). However, these data seemed to be too inconsistent with the PWT so it was decided not to use them.

14. *For Maldives* the PWT was used in addition to using data from Maddison (2003) for 1950. To recalculate the PWT to 1990 international dollars all the PWT observations were deflated with the ratio between PWT and Maddison for USA 1990. The data for 1950 was calculated from the text of Maddison (2003), table 5-4.

15. *For 37 areas*, for which no other data was available, we used the data from CIA. To make the data comparable to the other observations, expressed in 1990 international prices, we deflated all the observations with the ratio between the CIA data for USA for 2006 and our data for the USA for 2005. However, we do not know how the CIA data originally was calculated, so we are not certain whether it make sense to do such a deflation.

Only one year per area was available from this source (which year depends on the area). In order to get coverage for at least 2002 and 2003 it has been assumed that the GDP for those two years were the same as for 2005.

As noted above, this data should be considered particularly shaky.

UPDATES VERSION 2 TO VERSION 3:

The title and the front-page has been changed. One graph has been removed.

UPDATES VERSION 1 TO VERSION 2:

We have updated the previous dataset with new data up to 2006 for 180 countries and territories. This data was based on the real growth rate taken from World Development Indicators 2008. For all these countries and territories the previous data for 2004-2006 was replaced with the new data.

For most of these 180 areas this has not caused any major changes in the data. The exceptions are Equatorial Guinea, Montenegro, Dominica and Serbia for which there seem to be discrepancies between the old and new data of up to 20%. See below for a description of how the update was made.

References

Forthcoming

