# Sources and corrections

Data Pluralism is fundamental to the construction of the dataset.

To understand what it means, consider the simplest case in which the data from a provider is modified to clean it or extend it. Thus, the dataset provides, for convenience, a calculation of the US dollar value of GDP of all countries covered by the Penn World tables national accounts data. This figure is not present in the original data but can be reconstructed, because the Penn data includes an exchange rate that can be used to convert Local currency estimates into US dollar estimates. If this estimate is attributed to the Penn dataset itself, the implication is that it has been provided by the compilers of the dataset, which is not the case. Moreover, there is more than one possible such estimate because, for example, the exchange rates supplied by IFS differ in some cases from those supplied by Penn.

Another example is ‘de-indexing’. Both Jorda et al. and Barro-Ursua, supply ‘indexed real GDP’ figures, in the first case indexed to 2005=100 and in the second to 2006=100. But neither compiler supplies the base GDP from which the original GDP may be inferred. This baseline has to be supplied from another source (in the dataset, this is the UN real GDP estimate) for real GDP in the base year, in order to reconstruct a constant GDP series. Whilst not critical in calculating growth (since the growth of an indexed variable is the same as the growth of the original) it is useful to be able to compare the estimate of GDP from one source, with the estimate from another, to see if it affects any qualitative conclusions.

Yet a third example is splicing; researchers often wish to amalgamate data from two different sources, in order to create a continuous series.

In such cases the intention is that the Macrohistory dataset should *both* the original data and the derived data. Until 2020 this was accomplished by means of a ‘Definition’ dimension that specifies whether the data series is the original, or is modified. This proved too cumbersome for everyday use and has been withdrawn. The intention now us to modify the ‘Source’ dimension to provide multiple definitions based on the same source data.

# Germany: how difficult problems may be overcome, if authorities think it merits attention, and why we should not take the result for granted

In 1990, Volcker Ruhe famously announced that ‘We have to stop treating Germany as a state on wheels’. This followed the *Widervereinigung* (re-unification) of Eastern and Western Germany following a referendum securing consent for the creation of a single state incorporating the territories of both sovereign states. At this point, as the German right began reviving romantic memories of lost Polish and Czech territories, maps of Germany began to appear dating back to the Holy Roman Empire, including Mediaeval Germany’s historic Mediterranean seaboard. Ruhe’s offhand remark served equally to rebuke the romantic expansionism, and to put a lid on any residual East German separatism by declaring a final boundary of Germany.

To put this in context, the statistician must recognise just how frequently and how much Germany’s boundaries have moved, and recognise the issues of measurement and classification which arise. If a state’s boundaries have extended or contracted over a period under consideration, how do we know whether, for example, any economic changes are due to increased or decreased industriousness, or simply to the state being bigger or smaller?

To limit the extent of this historical-statistical essay, we simply leave out of our account the pre-war incorporation into Germany of significant parts of what is now the Czech Republic or the state of Poland and the Rhineland including Luxemburg (leading to a general strike in that country in 1942, surely one of the most heroic in history), not to mention the *Anschluss*.

We also set aside the general tendency of Germany’s statistical authorities, a tendency shared with France and England but restrained by the outcome of World War II, to incorporate the economic activities of overseas territories into those given by the traditional boundaries of their country. We note in passing that to this day the currency of French Guiana is the Euro and that, being part of the European Union, its economic product is included in the output of that august body, the very same which so recently angered the British by defining Gibraltar as a colony.

The Saarland [(en.wikipedia.org/wiki/Saarland](https://en.wikipedia.org/wiki/Saarland)) was established as a joint protectorate of the British and French in 1920 under the name ‘Saar Basin’. It was ceded to Germany under the 1935 Saar Referendum and its statistics were then reported under the provincial statistics of the state of Germany. However, in 1947, the Allied Occupied Germany organised it as the Saar protectorate of France ([en.wikipedia.org/wiki/List\_of\_French\_possessions\_and\_colonies](https://en.wikipedia.org/wiki/List_of_French_possessions_and_colonies)) , joining many other such illustrious entities. Until 1956 the Saar Basin was recognised as a member of the Council of Europe, issued its own postage stamps and, most significantly for statistical purposes, issued its own currency, the Saar Franc.

In 1955, the French administration organised a referendum which would make the territory independent under the auspices of a European [Commissioner](https://en.wikipedia.org/wiki/Commissioner) to be appointed by the Council of Ministers of the [Western European Union](https://en.wikipedia.org/wiki/Western_European_Union), while remaining in the [economic union](https://en.wikipedia.org/wiki/Economic_union) with [France](https://en.wikipedia.org/wiki/France). They rejected this by 424,423 votes to 201,973 with 15,725 blank or invalid votes, and in 1957 under the Saar Treaty joined the state of Germany.

At this point, German statisticians reorganised the reporting system to include Saarland, and retrospectively incorporate the statistics of the Saar Basin into those of West Germany.

The political jurisdiction of West Berlin provides a further complication which it is by no means simple to resolve. Prior to re-unification, West German statistics reported Berlin as a separate entity from the Federal Republic, adding its GDP to that of the rest of the territory only from 1958 onwards. I have in my possession the *Redivierte Volkswirtschaftliche Gesamtrechnungen* (Revised national accounts) of Germany for 1981 in which the distinction is still clear in the historical record.

With reunification, a further issue emerged: how should the statistical systems of East and West Germany be reconciled, consistently, with each other? East Germany reported on the basis of the ‘Material Balance system’ of Eastern Europe and the USSR, which involves many calculations contested until this day. When such issues of measurement and conversion arise in relation to entire countries such as Russia, China, or the Czech Republic, it is resolved on the basis of discussions between the statistical agencies of these sovereign countries. Germany is the only case in which two separate sets of GDP measures, calculated on a different basis, had to be aggregated.

For many years after the reunification, separate statistical estimates were published, and could be obtained, for each of the entities we have listed: Saar Basin, West Berlin, Federal Germany, German Democratic Republic. Now, a single statistical series is published by most authorities, not least the Germans, covering at least 1950 to the present day. This is considered authoritative for the purpose of international comparisons.

The difficulty this presents for historical comparisons is by no means easily overcome. We treat the current boundaries of Germany as its historical boundaries, in order to get a long-run comparison, but in so doing, we eliminate the actual historical record, and cannot (for example) easily establish how much of its growth from 1945 to 1989 was caused by capitalism, and how much by communism. Moreover, there is a symmetric problem: when a previously sovereign state is broken into component parts, as was the USSR and Yugoslavia, not to mention Czechoslovakia, Indonesia, Sudan, or Ethiopia, should we write history backwards in the same way, and wipe out the history of that state? Does this add to, or detract from, our understanding of the real course of economic history?

To allow the greatest possible flexibility in answering such questions, our dataset introduces the concept of a Reporting Unit.

# Reporting Units

A Reporting Unit, such as ‘Czechoslovakia (Former)’ or ‘USSR’ is a geopolitical entity, normally a sovereign state which has either broken up, or united, over the period covered by the data.

Statistical agencies usually – but not always – report data in such cases as a ‘broken’ series with some overlap. Thus the UN reports the GDP of Czechoslovakia as a single figure from 1970 until 1990, and from 1990 onwards as two separate figures for the Czech Republic and Slovakia. In 1990 there is an overlap.

Since the purpose of the dataset is to cover long-term economic trends, the data is ‘cleaned’ to provide a continuous series in such cases, but the data on individual parts is preserved. In the geographical hierarchy, therefore, the reporting unit provides continuous data for a single entity, but can be drilled down to give data on components, when available.

For the years in which there is an overlap, the duplicated data has to be removed. Thus if, in 1990, the UN data on both Czechoslovakia and its components are retained, then for that year, the GDP would appear to be doubled.

The modifications made to the ‘-EXTENDED’ definitions are given in a worksheet in ‘Meta.xlsx’ entitled *Extended Corrections*

## Semi-sovereign entities, colonies, departments

(eg Virgin Islands, Micronesia, Puerto Rico, Falklands, etc)

This section to be developed.

Westphalia has not yet quite reached Europe’s colonies. Note that this is almost entirely a European issue.

BUT See <http://theconversation.com/why-puerto-rico-doesnt-count-to-the-us-government-86262>

Virgin Islands, Guam, American Samoa and the Mariana Islands.

# AGGREGATES

(Eg ‘advanced economies’)

This section first to deal with the approach of the main authorities, then with the classification system in the dataset.

Key point is that the classification is not fixed but can be adapted by users. Users thus create the definitions appropriate for the object of study and also discuss, and attempt to reach consensus, on which of a range of alternatives is preferable.

## World Economic Outlook

2004 changed classification of countries: now ‘Advanced’ and the rest

2002 is the last year when WEO data were published from start year 1970

(<https://www.imf.org/external/pubs/ft/weo/2002/02/data/>)

Country groups 2018: (<https://www.imf.org/external/pubs/ft/weo/2018/02/weodata/weoselagr.aspx>)

|  |  |  |
| --- | --- | --- |
| **Country Groups Information** | |  |
|  | **World** Composed of 194 countries: Afghanistan, Albania, Algeria, Angola, Antigua and Barbuda, Argentina, Armenia, Aruba, Australia, Austria, Azerbaijan, The Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Burkina Faso, Burundi, Cabo Verde, Cambodia, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Comoros, Democratic Republic of the Congo, Republic of Congo, Costa Rica, Côte d'Ivoire, Croatia, Cyprus, Czech Republic, Denmark, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Eswatini, Ethiopia, Fiji, Finland, France, Gabon, The Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong SAR, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kiribati, Korea, Kosovo, Kuwait, Kyrgyz Republic, Lao P.D.R., Latvia, Lebanon, Lesotho, Liberia, Libya, Lithuania, Luxembourg, Macao SAR, FYR Macedonia, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Marshall Islands, Mauritania, Mauritius, Mexico, Micronesia, Moldova, Mongolia, Montenegro, Morocco, Mozambique, Myanmar, Namibia, Nauru, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Palau, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Puerto Rico, Qatar, Romania, Russia, Rwanda, Samoa, San Marino, São Tomé and Príncipe, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Singapore, Slovak Republic, Slovenia, Solomon Islands, Somalia, South Africa, South Sudan, Spain, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Sudan, Suriname, Sweden, Switzerland, Syria, Taiwan Province of China, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Tuvalu, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, Yemen, Zambia, and Zimbabwe.  **Advanced economies** Composed of 39 countries: Australia, Austria, Belgium, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong SAR, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Macao SAR, Malta, Netherlands, New Zealand, Norway, Portugal, Puerto Rico, San Marino, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Taiwan Province of China, United Kingdom, and United States.  **Euro area**  Composed of 19 countries: Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Portugal, Slovak Republic, Slovenia, and Spain.  **Major advanced economies (G7)** Composed of 7 countries: Canada, France, Germany, Italy, Japan, United Kingdom, and United States.  **Other advanced economies (Advanced economies excluding G7 and euro area)** Composed of 16 countries: Australia, Czech Republic, Denmark, Hong Kong SAR, Iceland, Israel, Korea, Macao SAR, New Zealand, Norway, Puerto Rico, San Marino, Singapore, Sweden, Switzerland, and Taiwan Province of China.  **European Union** Composed of 28 countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Romania, and United Kingdom.  **Emerging market and developing economies** Composed of 155 countries: Afghanistan, Albania, Algeria, Angola, Antigua and Barbuda, Argentina, Armenia, Aruba, Azerbaijan, The Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Burkina Faso, Burundi, Cabo Verde, Cambodia, Cameroon, Central African Republic, Chad, Chile, China, Colombia, Comoros, Democratic Republic of the Congo, Republic of Congo, Costa Rica, Côte d'Ivoire, Croatia, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Eswatini, Ethiopia, Fiji, Gabon, The Gambia, Georgia, Ghana, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hungary, India, Indonesia, Iran, Iraq, Jamaica, Jordan, Kazakhstan, Kenya, Kiribati, Kosovo, Kuwait, Kyrgyz Republic, Lao P.D.R., Lebanon, Lesotho, Liberia, Libya, FYR Macedonia, Madagascar, Malawi, Malaysia, Maldives, Mali, Marshall Islands, Mauritania, Mauritius, Mexico, Micronesia, Moldova, Mongolia, Montenegro, Morocco, Mozambique, Myanmar, Namibia, Nauru, Nepal, Nicaragua, Niger, Nigeria, Oman, Pakistan, Palau, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Qatar, Romania, Russia, Rwanda, Samoa, São Tomé and Príncipe, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Solomon Islands, Somalia, South Africa, South Sudan, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Sudan, Suriname, Syria, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Tuvalu, Uganda, Ukraine, United Arab Emirates, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, Yemen, Zambia, and Zimbabwe.  **Commonwealth of Independent States** Composed of 12 countries: Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyz Republic, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan. Georgia, which is not a member of the Commonwealth of Independent States, is included in this group for reasons of geography and similarities in economic structure.  **Emerging and developing Asia** Composed of 30 countries: Bangladesh, Bhutan, Brunei Darussalam, Cambodia, China, Fiji, India, Indonesia, Kiribati, Lao P.D.R., Malaysia, Maldives, Marshall Islands, Micronesia, Mongolia, Myanmar, Nauru, Nepal, Palau, Papua New Guinea, Philippines, Samoa, Solomon Islands, Sri Lanka, Thailand, Timor-Leste, Tonga, Tuvalu, Vanuatu, and Vietnam.  **ASEAN-5** Composed of 5 countries: Indonesia, Malaysia, Philippines, Thailand, and Vietnam.  **Emerging and developing Europe** Composed of 12 countries: Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Hungary, Kosovo, FYR Macedonia, Montenegro, Poland, Romania, Serbia, and Turkey.  **Latin America and the Caribbean** Composed of 33 countries: Antigua and Barbuda, Argentina, Aruba, The Bahamas, Barbados, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominica, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Suriname, Trinidad and Tobago, Uruguay, and Venezuela.  **Middle East, North Africa, Afghanistan, and Pakistan** Composed of 23 countries: Afghanistan, Algeria, Bahrain, Djibouti, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, Pakistan, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates, and Yemen.  **Middle East and North Africa** Composed of 21 countries: Algeria, Bahrain, Djibouti, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates, and Yemen.  **Sub-Saharan Africa** Composed of 45 countries: Angola, Benin, Botswana, Burkina Faso, Burundi, Cabo Verde, Cameroon, Central African Republic, Chad, Comoros, Democratic Republic of the Congo, Republic of Congo, Côte d'Ivoire, Equatorial Guinea, Eritrea, Eswatini, Ethiopia, Gabon, The Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, São Tomé and Príncipe, Senegal, Seychelles, Sierra Leone, South Africa, South Sudan, Tanzania, Togo, Uganda, Zambia, and Zimbabwe. | |