**SDMX STATISTICAL GUIDELINES**

**SDMX Cross-Domain Code Lists**

## CL\_AREA

**Name**: Code list for providing values to concepts such as “Reference area” [[1]](#footnote-1) and/or “Counterpart reference area” [[2]](#footnote-2).

**Description**: This code list provides code values for geographical areas, defined as areas included within the borders of a country, region, group of countries, etc.

**Input standards:**

* International Standard ISO 3166-1, Codes for the representation of names of countries and their subdivisions--Part 1: Country codes, ISO 3166-1: 2006 (E/F), International Organization on Standardization (Geneva, 2006). The latest version is available online at <http://www.iso.org/iso/home/standards/country_codes.htm>.
* “Standard country or area codes for statistical use” Series M, No. 49 (M49), United Nations, Statistics Division, New York website: <https://unstats.un.org/unsd/methodology/m49/> and a snapshot of the standard is prepared each year in the UN Statistical Yearbook, Annex I, for example (2018) https://unstats.un.org/unsd/publications/statistical-yearbook/files/syb61/SYB\_z\_Annex%20I.pdf.

**Version**: 2.0.

**Date**: March 2019.

|  |  |
| --- | --- |
| **Recommended code values** | **Recommended code description** |
| 001 | WORLD |
| 002 | AFRICA |
| 015 | NORTHERN AFRICA |
| 012 | ALGERIA |
| 818 | EGYPT |
| … | … |
| 896 | AREAS NOT ELSEWHERE SPECIFIED |
| … | … |
| DZ | ALGERIA |
| EG | EGYPT |

**Notes**

This version of CL\_Area supersedes an earlier version which presented solely the codes from the ISO 3166-1 alpha-2 list. This version extends the code list to include codes from the “Standard country or area codes for statistical use” Series M, No. 49 (M49). The difference between these code lists are as follows:

* The ISO 3166-1 alpha-2 are maintained by the International Organization for Standardization (ISO) for each country or area. Their purpose is to represent the “name” identity of the country or territory.
* The numerical codes are maintained by the United Nations Statistics Division as part of the Standard country or area codes for statistical use (M49), the codes are also presented as part of the ISO 3166-1 standard. Their purpose is to represent the “statistical” identity.

The Appendix provides further details and examples of the differences between the alpha-2 and numerical codes as well as additional numerical codes maintained under M49.

Since many statistical organisations prefer to use ISO 3166 alpha-2 as the codes are more easily identifiable with a country (i.e. United Kingdom = GB or UK = numerical code 826, United States of America = US = numerical code 840), alpha-2 and numerical codes are combined into a single code list for CL\_AREA to serve different preferences and application scenarios. One suggested approach is to adopt two hierarchical code lists in addition to the code list that contains both code lists, one that has only M49, and one that has only ISO 3166-1 alpha-2.

These codes can be supplemented by additional aggregate or subdivisions of these countries and territories. For example:

* + For coding areas within Europe, use of the Eurostat "Nomenclature of Territorial Units for Statistics" (NUTS) is recommended: <http://ec.europa.eu/eurostat/ramon/nuts/codelist_en.cfm?list=nuts>.
  + For coding internationally defined geographical, political and economic groups, M49 is recommended: <https://unstats.un.org/unsd/publications/catalogue?selectID=109>.

Depending on the context and domain, there may be differences in the territory definitions. Usually, guidance on this is provided in documents and guides maintained by the experts of particular domains. Information on political geography can also be explained in the footnotes attribute where utilised.

To address user requirements, additional extensions or user defined codes can be used to cover entities (usually, broader groupings) that are not included in the ISO or M49 standards.

The code list for this cross-domain concept is available from the global registry along with a structure set describing the mapping of the codes between the ISO and M49 standards.

# APPENDIX: Supplementary INFORMATION ON the ISO AND M49 CODE LISTs

## History

In 1974, **ISO 3166** was first published by the International Organization for Standardization. The standard originated from the distinguishing signs for vehicles under the Conventions on Road Traffic of 1949 and 1968. Updates were published in 1981, 1988, and 1993; and in 1996, it was announced that the standard would be divided into three parts: ISO 3166-1 would continue to provide two-letter, three-letter, and three-digit codes for countries and areas; ISO 3166-2 would contain codes for primary administrative subdivisions of countries and areas; and ISO 3166-3 would provide codes for obsolete country and area names, since 1974. The first publication of ISO 3166-1 was dated 1997.

In 1969, **Series M, No. 49** (known as M49) was first developed by the UN Statistical Office, now the United Nations Statistics Division, to facilitate information transmission and the standardization of material for data processing. It started as a list of **country or areas codes** as were reporting data to the UN Statistical Office at the time. Each country or area is represented by a unique three-digit numerical code covering a range from 000 to 899. Over four subsequent revisions in 1975, 1982, 1996, and finally in 1999; it was expanded to include many different types of codes for statistical reporting (see section on technical differences for more information).

## Common governance, differentiated maintenance and application

Despite the two standards being maintained by two organizations completely independent of each other, the governance arrangements and contents of the two standards are similar. Both code lists present the same two-**letter**, three-**letter**, and three-**digit** for countries and areas and apply the same country names as stipulated by terminology issued by the UN Secretariat[[3]](#footnote-3). However, the three-digit codes are maintained by the United Nations Statistics Division and the two-letter and three-letter codes are maintained by the International Organization for Standardization under different maintenance arrangements. As suggested by their individual origins, the main difference between the maintenance arrangements are M49 was created and is maintained for statistical reporting whereas the ISO 3166 is a country identifier for use in passports, internet domains and vehicle registrations etc.

### M49/ISO 3166 changes to country names and codes

Both M49 and ISO 3166 will only allocate a country code when a country has attained the stature and sovereignty required by the United Nations and the international community. Any change depends on a General Assembly resolution pertaining to an official announcement/ proclamation of a new sovereign nation or an official change of name or area nomenclature. Country terminology is codified by the United Nations Terminology Unit, considering country views and decisions of inter-governmental bodies. For other countries or areas, including sub-divisions, terminology is determined in consultation with Member States that might have an interest in the matter through the UN Secretariat's Office of Legal Affairs. The names of countries or areas refer to their short form used in day-to-day operations of the United Nations and not necessarily to their official name as used in formal documents.

Regular updates and maintenance to the country names, the three-digit, two- and three-letter codes in M49/ISO 3166 are made only when the following situations occur:

1. An official change in the name of a country or area while its political geography remains the same; this change is not usually accompanied by a change of the country’s three-digit code, only by a change in the two- and/or three-letter codes like Burkina Faso whose previous name was Upper Volta.
2. A major change in the political geography of a country or area; this is generally accompanied by a change in three-digit code used for statistical purposes and often in the name of the country. Possible scenarios:

* Two countries merge into one sovereign state as in the case of Germany from the former countries of the German Democratic Republic and the Federal Republic of Germany.
* A country may separate or dissolve like the Socialist Federal Republic of Yugoslavia where its states declared independence to become six different countries, i.e. Serbia, Croatia, etc.
* A country may secede, achieve independence or autonomy from another as the situation of Eritrea which was formerly part of Ethiopia.
* A sovereign country agrees and allows its autonomous entities or territories to provide their own data directly to the United Nations Statistical Division like the Åland Islands, which is a province of Finland.

## Technical differences between M49 and ISO 3166

The earlier section on *M49/ISO 3166 changes to country names and codes*, described how the three-digit code for countries and areas differed in their maintenance from the corresponding two- and three-letter codes. This section goes into more detail and provides further examples of how changes to the political geography or name of a country or area affect the maintenance of these codes. The section then considers the classes and types of three-digit codes which are unique to M49 or additional to the three-digit codes presented by ISO 3166.

### Political geography changes

As described in an earlier section, the three-digit code from M49 usually represents the “statistical” identity name of country or area, and the two-letter and three-letter codes from ISO 3166 represent the “name” identity of a country or area. Therefore, if the political geography of a country or area changes but not the name, the three-digit code will change but the two- and three- letter codes will stay the same. Since 1982, this situation has occurred in 6 instances, changes are highlighted in yellow, as shown in the table below:

| **M49 Code** | **ISO-3 code** | **ISO-2 code** | **Name (see M49 website for explanation of changes)** |
| --- | --- | --- | --- |
| 532 | ANT | AN | Netherlands Antilles [former, pre-1986] |
| 530 | ANT | AN | Netherlands Antilles [former, 1986 - 2010] |
| 533 | ABW | AW | Aruba |
|  |  |  |  |
| 278 | DDR | DD | German Democratic Republic [former] |
| 280 | DEU | DE | Germany, Federal Republic of [former] |
| 276 | DEU | DE | Germany |
|  |  |  |  |
| 230 | ETH | ET | Ethiopia [former] |
| 231 | ETH | ET | Ethiopia |
| 232 | ERI | ER | Eritrea |
|  |  |  |  |
| 736 | SDN | SD | Sudan [former] |
| 729 | SDN | SD | Sudan |
| 728 | SSD | SS | South Sudan |
|  |  |  |  |
| 720 | YMD | YD | Democratic Yemen [former] |  |  |
| 886 | YEM | YE | Yemen [former] |
| 887 | YEM | YE | Yemen |
|  |  |  |  |
| 890 | YUG | YU | Yugoslavia [former] |
| 891 | YUG | YU | Serbia and Montenegro [former] |
|  |  |  | + 5 other republics created out of “Yugoslavia” |

This makes it very difficult to track political geography changes which will affect a statistical series by using the two- and three- letter code codes and maintain consistent historical time series.

### Name changes

In contrast to the previous section, if the political geography of a country or area stays the same but the name changes, the three-digit code will stay the same but the two- and/or three- letter codes will change. Since 1982, this situation has occurred in 6 instances, changes are highlighted in yellow, as shown in the table below:

| **M49 Code** | **ISO-3 code** | **ISO-2 code** | **Name (*until year changed*)** |
| --- | --- | --- | --- |
| 104 | BUR | BU | Burma *(until 1989)* |
| 104 | MMR | MM | Myanmar |
| 112 | BLR | BY | Byelorussian SSR *(until 1992)* |
| 112 | BLR | BY | Belarus |
| 180 | ZAR | ZR | Zaire *(until 1997)* |
| 180 | COD | CD | Democratic Republic of the Congo |
| 626 | TMP | TP | East Timor *(until 2002)* |
| 626 | TLS | TL | Timor-Leste |
| 642 | ROM | RO | Romania *(no change)* |
| 642 | ROU | RO | Romania |
| 854 | HOV | HV | Upper Volta *(until 1984)* |
| 854 | BFA | BF | Burkina Faso |

By using the two- and three-letter codes, this makes it very difficult to a) maintain consistency in the codes since a name or label change does not affect the underlying statistical series, and b) maintain consistent historical time series.

### M49: Combinations of countries or areas

M49, separate from ISO 3166, provides codes for combinations of countries or areas which may represent more than one country or area identity. The most common example shown in the table below is the Channel Islands which is the combination of two areas, Guernsey and Jersey. This area is more often reported in statistics disseminated by international agencies than the individual areas.

|  |  |  |  |
| --- | --- | --- | --- |
| **M49 Code** | **ISO-3 code** | **ISO-2 code** | **Name** |
| 830 | <NA> | <NA> | Channel Islands |
| 831 | GGY | GG | Guernsey |
| 832 | JEY | JE | *Jersey* |

In previous revisions of the M49, many similar combinations were created as listed in the table below. Some of these codes are still used by databases within the UN Statistics Division like trade and energy. However, they are currently considered inactive and the current recommendation is in these cases to attach the data to the ‘larger’ area and footnote that it includes the other areas.

|  |  |  |  |
| --- | --- | --- | --- |
| **M49 Code** | **ISO-3 code** | **ISO-2 code** | **Name** |
| 58 | <NA> | <NA> | Belgium-Luxembourg |
| 251 | <NA> | <NA> | France-Monaco |
| 381 | <NA> | <NA> | Italy-San Marino-Holy See |
| 579 | <NA> | <NA> | Norway incl. Svalbard and Jan Mayen Is. |
| 757 | <NA> | <NA> | Switzerland-Liechtenstein |
| 841 | <NA> | <NA> | United States of America incl. Puerto Rico |
| 842 | <NA> | <NA> | USA incl. Puerto Rico, US Virgin Is. |

### M49: Sub-divisions of countries or areas

M49, again separate from ISO 3166, also provides codes for sub-divisions of countries or areas. Some active examples are given in the table below:

|  |  |  |  |
| --- | --- | --- | --- |
| **M49 Code** | **ISO-3 code** | **ISO-2 code** | **Name** |
| 655 | <NA> | <NA> | Ascension |
| 656 | <NA> | <NA> | Tristan da Cunha |
| 667 | <NA> | <NA> | Saba |
| 668 | <NA> | <NA> | Sint Eustatius |
| 669 | <NA> | <NA> | Bonaire |
| 836 | <NA> | <NA> | United Rep. of Tanzania (Zanzibar) |

### ISO 3166: Regions and other groupings

* Several organisations add extensions to the ISO country codes for representing various country groupings (e.g. North America, Central Africa, OPEC countries, rest of the world excluding the euro area, etc);
* Several organisations, especially in Europe, use some codes that may deviate from the ISO code list (e.g. “UK” for the United Kingdom while “GB” is the official ISO code). Political issues or country naming uncertainties may also exist for some other codes included in the ISO code list: such issues may lead to additional deviations in the actual implementations;
* For properly maintaining historical data in databases, there is often a need to keep using codes that ISO may suppress or switch them to inactive codes. Such codes (e.g. code for Eastern Germany) may be “deviations”, but they may need to continue being in use by organisations.
* The UNECE currently uses 6 groupings: (i) ECE - 52 of the 56 ECE member countries (i.e. excludes Andorra, Liechtenstein, Monaco and San Marino), (ii) NA - North America (this includes the USA and Canada), (iii) E2A - European Union (the definition of this code changes as necessary to reflect current membership, so for now it covers the 28 EU Member States), (iv) EAA - Euro Area (current membership), (v) CIS - the 12 current members of the CIS (includes Turkmenistan (associate member) and Georgia (notice to leave CIS becomes effective in August 2009)), (vi) ST7 - six Western Balkans countries (Albania, Bosnia and Herzegovina, Croatia, Montenegro, Serbia, the Former Yugoslav Republic of Macedonia);
* Concerning “subdivisions” of areas:

Some organisations may use the ISO code list for also coding subdivisions in Europe. (In the list of “recommended” codes, the Eurostat NUTS code list is recommended)

The UNECE uses the following code list: <http://www.unece.org/cefact/locode/service/sublocat.htm>

### M49: Regions and other groupings

The M49, separate from ISO 3166, also provides a classification of countries and areas by geographical regions and sub-regions. To ensure consistency in statistics and for convenience, all regions are mutually exclusive and therefore each country or area is shown in one region only. The macro geographical regions are arranged to the extent possible according to 6 continents; these continental regions, except Antarctica, are further subdivided into 22 sub-regions and 2 intermediary regions (*Sub-Saharan Africa* and *Latin America and the Caribbean*) that are drawn as to obtain greater homogeneity in sizes of population, demographic circumstances and accuracy of demographic statistics, see [M49](https://unstats.un.org/unsd/methodology/m49/)website for the full classification. The nomenclature, if not the codes, are used widely in statistics disseminated by international agencies. As discussed in an earlier section; from these sub-regions and intermediary regions, the Sustainable Development Goals (SDGs) regions were created that will be presented in the SDG “glossy” reports; defined in the following table:

| **M49 Code** | **ISO-3 code** | **ISO-2 code** | **Name** |
| --- | --- | --- | --- |
| 747 | <NA> | <NA> | Northern Africa and Western Asia |
| 202 | <NA> | <NA> | Sub-Saharan Africa |
| 513 | <NA> | <NA> | Northern America and Europe |
| 419 | <NA> | <NA> | Latin America and the Caribbean |
| 753 | <NA> | <NA> | Eastern and South-Eastern Asia |
| 62 | <NA> | <NA> | South-central Asia |
| 53 | <NA> | <NA> | Australia and New Zealand |
| 543 | <NA> | <NA> | Oceania (exc. Australia and New Zealand) |

For statistical and analytical purposes, the UN Statistics Division and Population Division have also been using a classification based on a combination of geographical criteria and level of development. The basic regions and sub-regions defined according to the M49 are further aggregated based on whether the regions are mostly composed of developing or developed countries[[4]](#footnote-4). Although these designations are not intended to express a judgment about the actual stage of development that a country or area has reached and are not based on the performance of any specific economic indicators, they are a convenient way to present the analysis and the assessment of progress separately for those parts of the world where development is generally still a significant concern.

Additionally, the M49 contains codes for other groupings of important socio-economic compositions of countries and areas, including a developed and developing regions classification which was formalized in 1996. Otherwise many of these codes represent trade associations and supranational organizations (*European Union*), but for the SDG reporting the main additional other groupings requested are as follows;

* the Least Developed Countries[[5]](#footnote-5), a group of countries, identified as being particularly vulnerable, with a low-income level and a low index of human assets, based on nutrition, health and education;
* the Landlocked Developing Countries[[6]](#footnote-6), a group of countries constrained in their development by the lack of territorial access to the sea, remoteness and isolation from world markets and high transit costs; and
* the Small Islands Developing States[[7]](#footnote-7), a group of small islands developing states, with a narrow resource base, a small domestic market, high costs for energy, infrastructure, transportation, communication and servicing, long distances from export markets and import resources, low and irregular international traffic volumes, little resilience to natural disasters and fragile natural environments.

|  |  |  |  |
| --- | --- | --- | --- |
| **Code** | **ISO-3 code** | **ISO-2 code** | **Name** |
| 199 | <NA> | <NA> | Least Developed Countries (LDCs) |
| 432 | <NA> | <NA> | Landlocked developing countries (LLDCs) |
| 722 | <NA> | <NA> | Small island developing States (SIDS) |

Unlike the country and areas three-digit codes, there is a maintenance policy among these regions and other groupings not to change the code when there is a change to underlying composition or membership of the group (i.e. Asia-Pacific Economic Cooperation (APEC), Central American Common Market (CACM), Economic Community of West African States (ECOWAS), European Free Trade Association (EFTA), EU/EC/EEC, Least Developed Countries (LDCs), Organization for Economic Co-operation and Development (OECD) and Organization of the Petroleum Exporting Countries (OPEC). The main exception is between the 1982 and 1996 M49 revisions, many code changes were made to the regions and sub-regions of Asia, Europe and Oceania to distinguish the new compositions of countries and areas under them from the previous classification. This was necessary due the dissolution of the former Soviet Union, where it was a separate region in the 1982 revision; and the break-up of the Pacific Islands (Trust Territory) into separate independent islands. However, with the recent break-up of Sudan a decision was made to put South Sudan in Eastern Africa rather than Northern Africa as it was when part of Sudan but this resulted in no code changes to these sub-regions.

1. Country or geographical area to which the measured statistical phenomenon relates (Source: [SDMX Glossary, Version 1.0](https://sdmx.org/wp-content/uploads/SDMX_Glossary_Version_1_0_February_2016.docx)). [↑](#footnote-ref-1)
2. Secondary area, as opposed to reference area, to which the measured data are in relation (Source: [SDMX Glossary, Version 1.0](https://sdmx.org/wp-content/uploads/SDMX_Glossary_Version_1_0_February_2016.docx)). [↑](#footnote-ref-2)
3. The listing of countries or areas in the six official languages of the United Nations can be found on the UNTERM website at <http://unterm.un.org>, and was previously available as Terminology Bulletin No. 347/Rev.1, prepared by the Department of General Assembly Affairs and Conference Services of the United Nations Secretariat. [↑](#footnote-ref-3)
4. Developed regions defined as Northern America (numerical code 021), Europe (150), Japan (392) and Australia and New Zealand (053). The other regions are defined as “developing”. [↑](#footnote-ref-4)
5. See website of the Development Policy & Analysis Division of the Department of Economic and Social Affairs: <https://www.un.org/development/desa/dpad/least-developed-country-category.html> [↑](#footnote-ref-5)
6. See website of the United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and the Small Island Developing States (UN-OHRLLS): <http://unohrlls.org/about-lldcs/> [↑](#footnote-ref-6)
7. See website of the Department of Sustainable Development of the Department of Economic and Social Affairs: <https://sustainabledevelopment.un.org/index.php?menu=1520> [↑](#footnote-ref-7)