Regional and country classification
Bert Lenaerts
09 November 2020

Suggested citation:
Lenaerts, Bert, 2020. <https://github.com/BertLenaerts/geodata.git>.

1   Overview

The first dataset (Names_global) converts country names written in full to the respective ISO3 codes. The second dataset (Labels) provides additional classifications and labels for each country included. Table 1 provides an overview of all variables.

Table 1 Definition of variables.

Variable name	Definition
ISO3	Official ISO3 country code
ISO2	Official ISO2 country code
ISO2_EU	Same as ISO2 with two exceptions for Greece (EL) and the United Kingdom (UK)
SIMPLE_NAMES	Simplified country names
FULL_NAMES	Official country names
EXTINCT	Indicator of countries no longer existing
CONTINENT	Simple regional classification
CONTINENT_WB	World Bank regional classification
CONTINENT_UN	United Nations regional classification
CONTINENT_UN2	United Nations subregional classification
CONTINENT_CGIAR	CGIAR regional classification
EU_map	Indicator of European and near-European countries that can be used for mapping purposes (that is, North Africa, the Middle East and Central Asia)
EU_extended	Extended Europe indicator, including Russia and Turkey.
LANDLOCKED	Indicator of landlockedness
SEALOCKED	Indicator of sealockedness
SOUTH	Southern hemisphere indicator.
DEPENDER	Sovereign state above dependent territories (if relevant)

2   ISO3, ISO2, ISO2_EU, SIMPLE_NAMES, FULL_NAMES

Some countries changed their ISO3 codes over time: Romania's switch from ROM to ROU, Burma (BUR) being renamed to Myanmar (MMR), East Timor (TMP) being renamed to Timor-Leste (TLS) and the Byelorussian SSR (BYS) being renamed to Belarus (BLR).

Kosovo (XKX) and Palestine (PSE)[1] are also included despite not having an official ISO3 code. The Channel islands Jersey (JEY) and Guernsey (GGY) are included separately but not as a whole[2].

Lastly, some codes differ between sources. The World Bank uses the code MON for Montenegro (MNE) and the code WSH for the Western Sahara (ESH). The European Union used the ISO2 code EL for Greece (GRC, GR) and the ISO2 code UK for United Kingdom (GBR, GB). The Seychelles (SYC) share the same ISO2 code with former Serbia and Montenegro (SCG): SC.

Full country names are derived from the European Union Style Guide (European Commission, 2020). See Table 2 for details.

3   EXTINCT

Table 2 Extinct countries

Extinct country	Constituting parts today
Included (separate from constituting parts today)	
Netherlands Antilles (ANT)	Aruba (ABW), Curaçao (CUW), Sint Maarten (Dutch part) (SXM) and the Caribbean Netherlands (Bonaire, Sint Eustatius and Saba) (BES)
Czechoslovakia (CSK)	Czechia (CZE) and Slovakia (SVK)
Yugoslavia (YUG)	Bosnia and Herzegovina (BIH), Croatia (HRV), North Macedonia (MKD), Serbia (SRB), Montenegro (MNE), Slovenia (SVN)
Serbia and Montenegro (SCG)	Serbia (SRB) and Montenegro (MNE)
Zaire (ZAR)	Democratic Republic of the Congo (Congo-Kinshasa; COD)
Excluded (considered part of constituting parts today)	
Saudi Arabian–Iraqi neutral zone (NTZ)	
Somaliland	Somalia
Northern Cyprus, Cyprus No Mans Area	Cyprus
People's Democratic Republic	Yemen

---

[1] Sometimes the code PAL is also used.
[2] Sometimes the code CHI is used for the Channel islands as a whole.

of Yemen (South)	
Republic of Vietnam	Vietnam
German Democratic Republic (East)	Germany
Federal Republic of Germany (West)	

4 CONTINENT

One approach to define supranational regions is the use of a continental classification, where a continent is understood as a large landmass. Given that continents are generally identified by convention rather than any strict criteria, we provide one simple classification suggested by Allroggen et al. (2015) (Figure 1). The following abbreviations are used: Africa, AFR. Oceania, OCE. Central America, South America, and Caribbean, CSC. East Asia and Central Asia, ECA. Europe, EU. Northern America, NAX. South Asia and South-East Asia, SSEA. West Asia / Middle East, MEA. Antarctica, ANT.
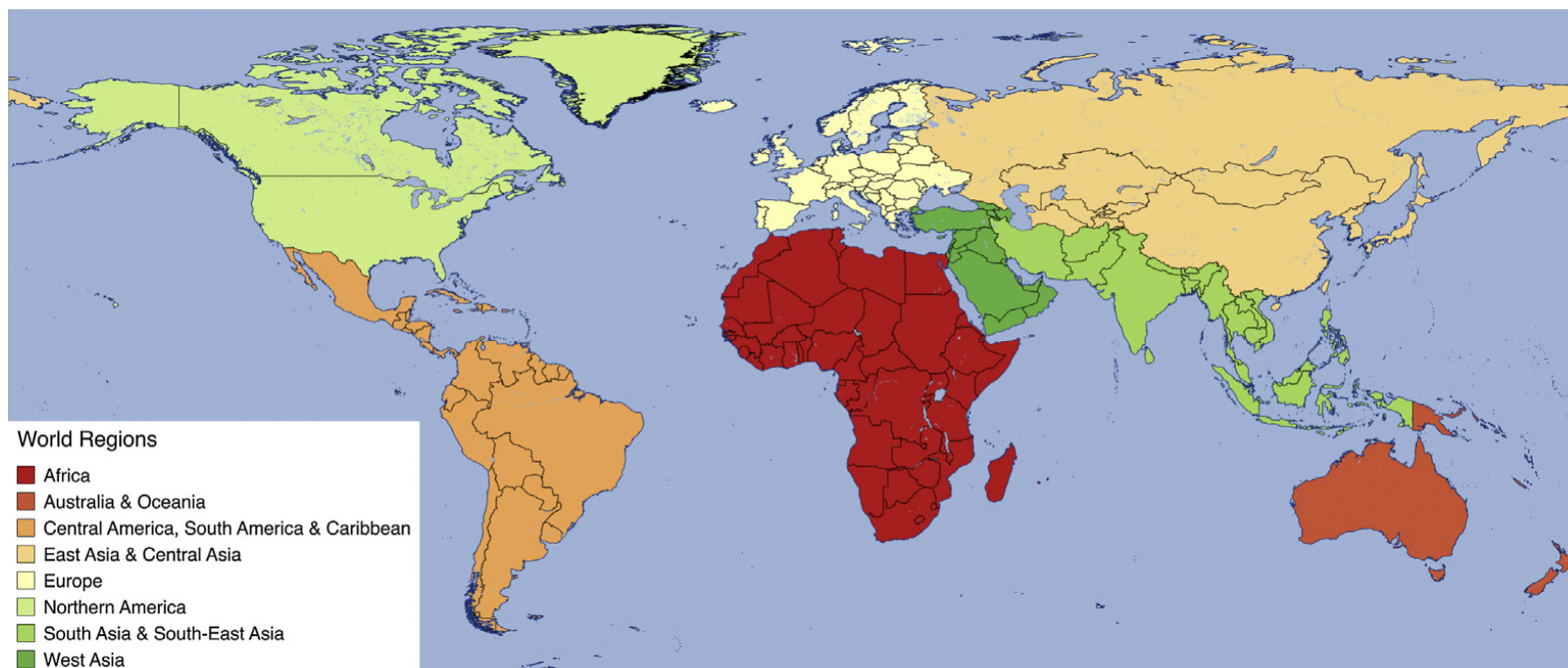
Figure 1 Simple regional classification
Source: Allroggen et al. (2015)

5   CONTINENT_WB

One common regional classification is the one developed by the World Bank (Figure 2) with the following abbreviations: (i) Sub-Saharan Africa (SSA), (ii) Middle East and North Africa (MNA), (iii) South Asia, (SA), (iv) Latin America and Caribbean (LCR), (v) East Asia and the Pacific (EAP), (vi) Europe and Central Asia (ECA), (vii) North America (NA) and (viii) Antarctica (ANT). Given that not all ISO3 codes are assigned a class (World Bank, 2018), any remaining ISO3 entities were assigned one based on best judgement (Table 3).

Table 3 List of countries unclassified by the World Bank

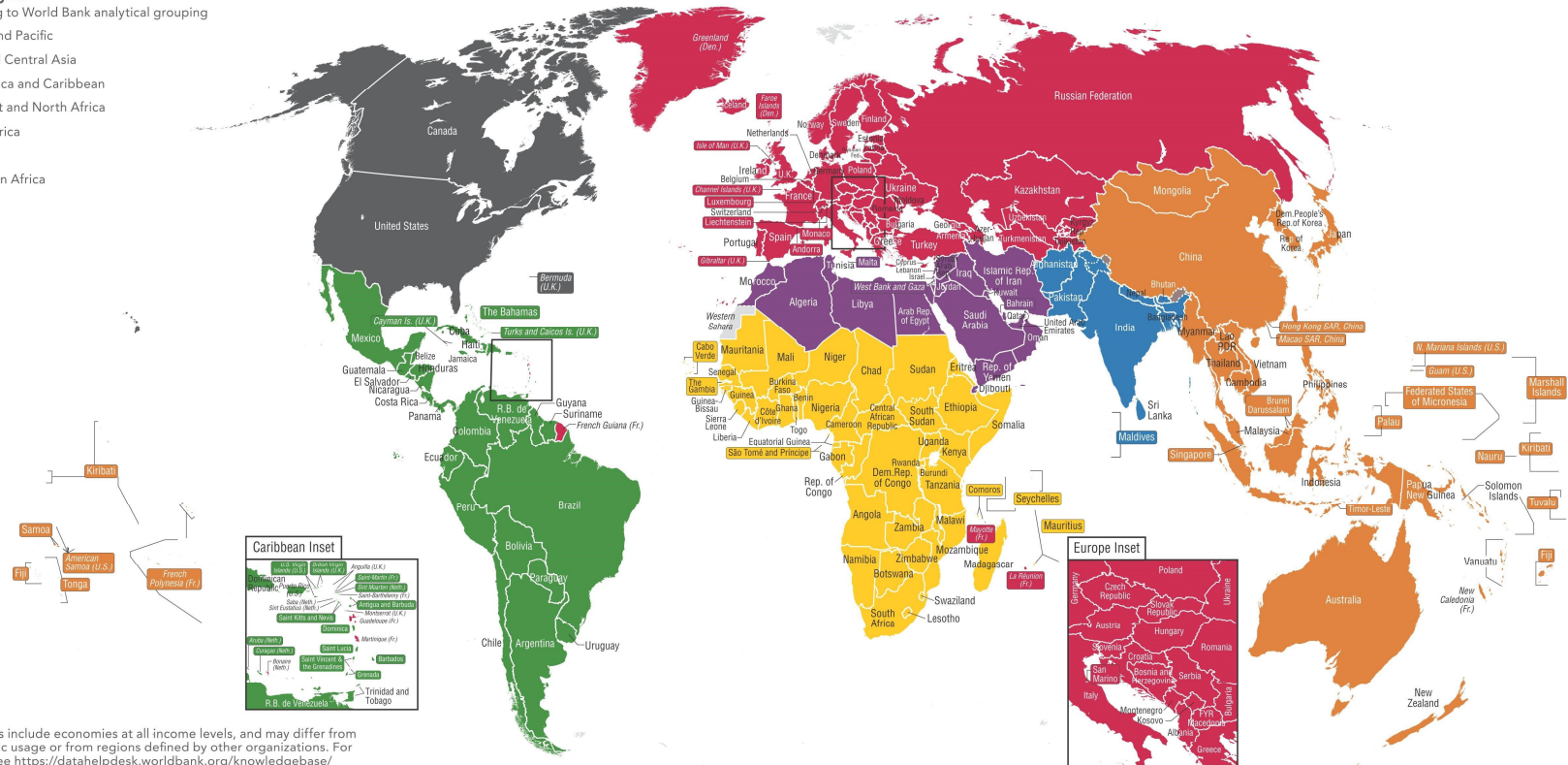| ISO3 code	Country name	Classification
ESH	Western Sahara	Middle East and North Africa
SHN	Saint Helena Islands	Europe and Central Asia
ATA	Antarctica	Antarctica
ATF	French Southern Territories	Antarctica
BVT	Bouvet Island	Antarctica
HMD	Heard and McDonald Islands	Antarctica
FLK	Falkland Islands	Europe and Central Asia
MSR	Montserrat	Europe and Central Asia
SGS	South Georgia Islands	Europe and Central Asia
TWN	Taiwan	East Asia and the Pacific
ALA	Aland Islands	Europe and Central Asia
SJM	Svalbard and Jan Mayen Islands	Europe and Central Asia
VAT	Vatican	Europe and Central Asia
SPM	Saint Pierre and Miquelon	North America
UMI	US Minor Outlying Islands	North America
CCK	Cocos Islands	East Asia and the Pacific
COK	Cook Islands	East Asia and the Pacific
CXR	Christmas Island	East Asia and the Pacific
NFK	Norfolk Island	East Asia and the Pacific
NIU	Niue	East Asia and the Pacific
PCN	Pitcairn Islands	East Asia and the Pacific
TKL	Tokelau	East Asia and the Pacific
WLF	Wallis and Futuna Islands	East Asia and the Pacific
IOT	British Indian Ocean Territory	Europe and Central Asia

Figure 2 World Bank regional classification

Source: World Bank (2018)

6 CONTINENT_UN, CONTINENT_UN2

Another common regional classification is the one developed by the United Nations (Figure 4), with the abbreviations for the regions and subregion defined in Table 4. Given that not all ISO3 codes are assigned a class (UN Statistics Division, 2020), any remaining ISO3 entities were assigned one based on best judgement (Table 5).

Table 4 United Nations regional classification

	Region	Abbr.	Subregion	Abbr.
1	Africa	AFR	Eastern Africa	EAF
			Northern Africa	NAF
			Southern Africa	SAF
			Western Africa	WAF
			Middle Africa	MAF
2	Asia	ASI	Western Asia	WAS
			Central Asia	CAS
			South-Eastern Asia	SEAS
			Eastern Asia	EAS
			Southern Asia	SAS
3	Europe	EU	Southern Europe	SEU
			Eastern Europe	EEU
			Northern Europe	NEU
			Western Europe	WEU
4	Latin America and the Caribbean	LCR	Caribbean	CAR
			South America	SAM
			Central America	CAM
5	Northern America	NAX	Northern America	NAX
6	Oceania	OCE	Polynesia	POL
			Melanesia	MEL
			Micronesia	MIC
			Australia/New Zealand	ANZ
7	Antarctica	ANT	Antarctica	ANT

Table 5 List of countries unclassified by the United Nations

ISO3 code	Country name	Classification
TWN	Taiwan	Eastern Asia
XKX	Kosovo	Southern Europe

Figure 3 United Nations regional classification
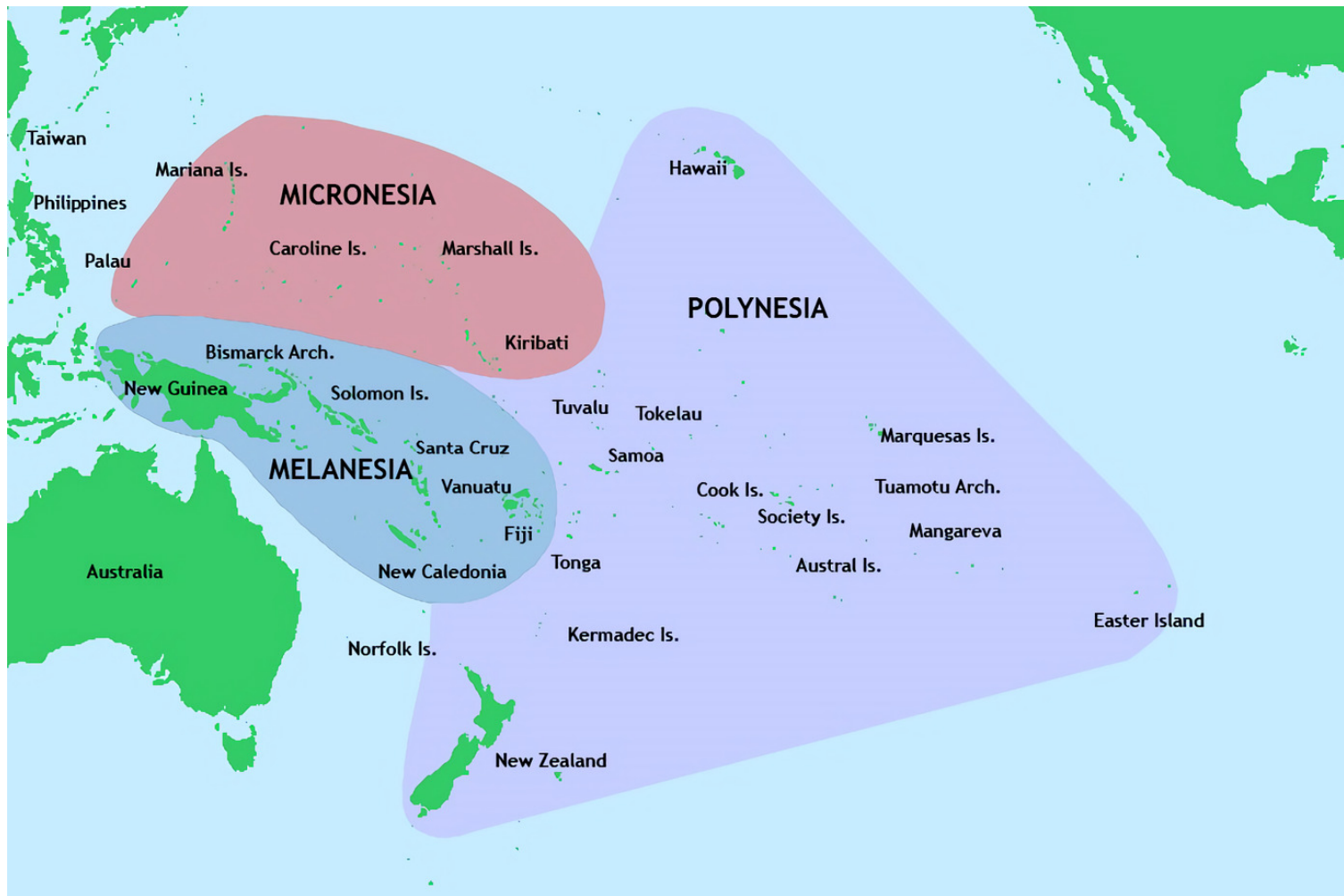Source: Wikipedia (2020a)

Figure 4 Pacific Culture Areas
Source: Wikipedia (2020b)

7   CONTINENT_CGIAR

The CGIAR has another regional classification (Figure 5): (i) West and Central Africa (WCA), (ii) East and Southern Africa (ESA), (iii) Central and West Asia and North Africa (CWANA), (iv) South Asia (SA), (v) Southeast Asia (SEA) and (vi) Latin America and Caribbean (LAC). The remaining countries are grouped here into: (vii) Europe and Offshoots (WST), (viii) Eastern Asia (EA) and (ix) Antarctica (ANT). The subregions are taken from the United Nations regional classification with the exception of Iran (here CWANA), Sudan (here ESA) and Russia (here EA). The pacific subregions Polynesia, Melanesia and Micronesia are considered part of Southeast Asia while the western Offshoots consist of the United States, Canada, Australia and New Zealand.
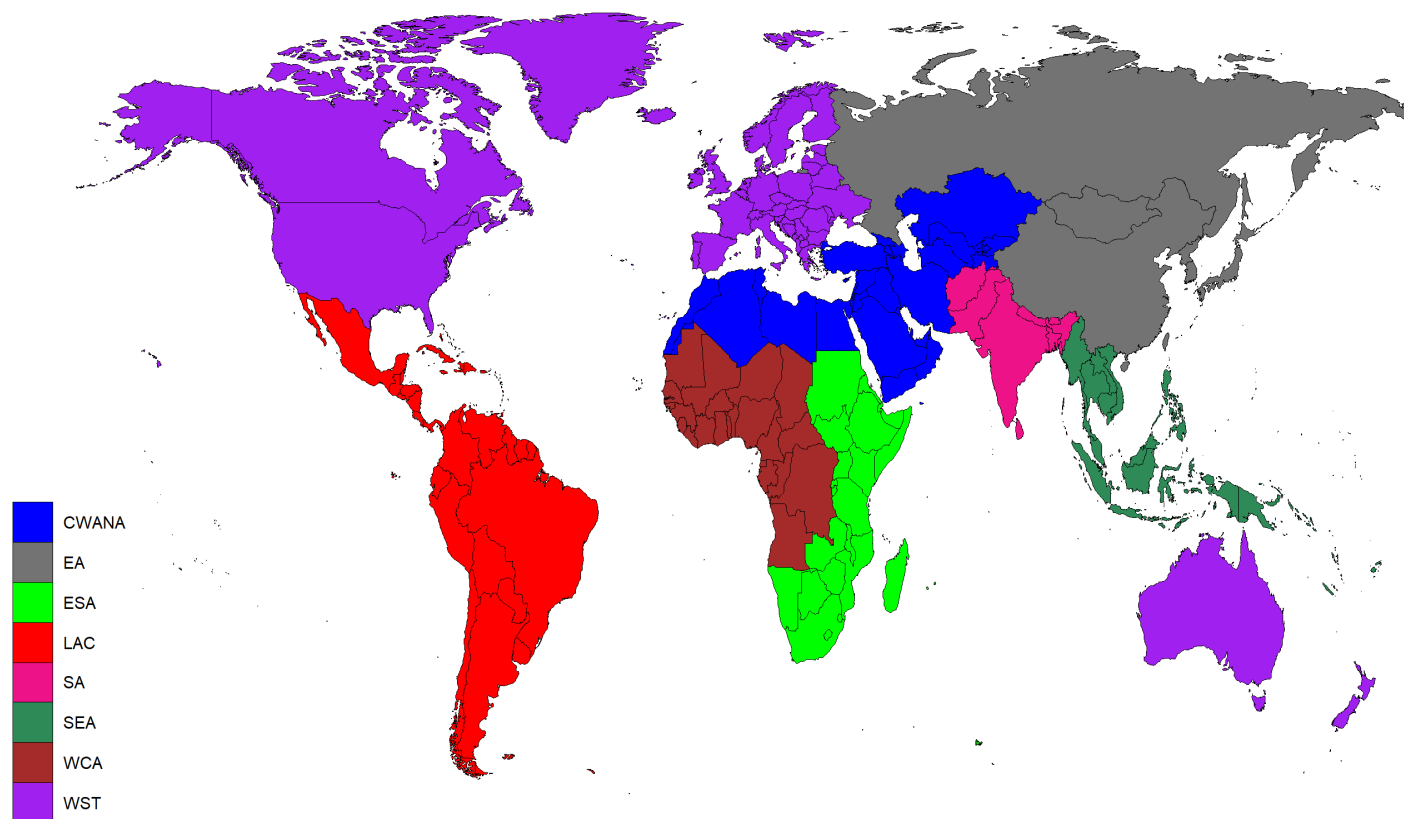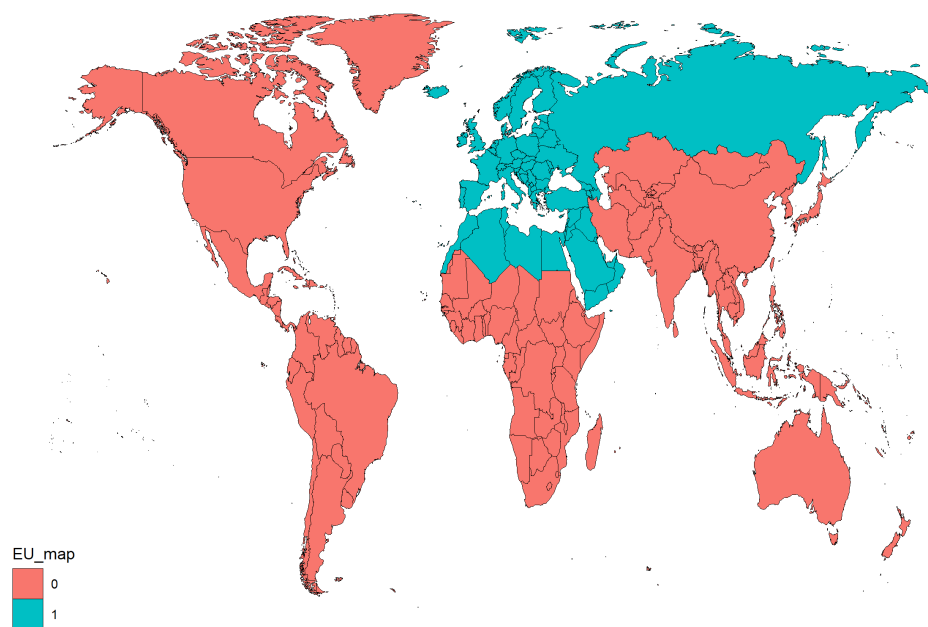
Figure 5 CGIAR regional classification

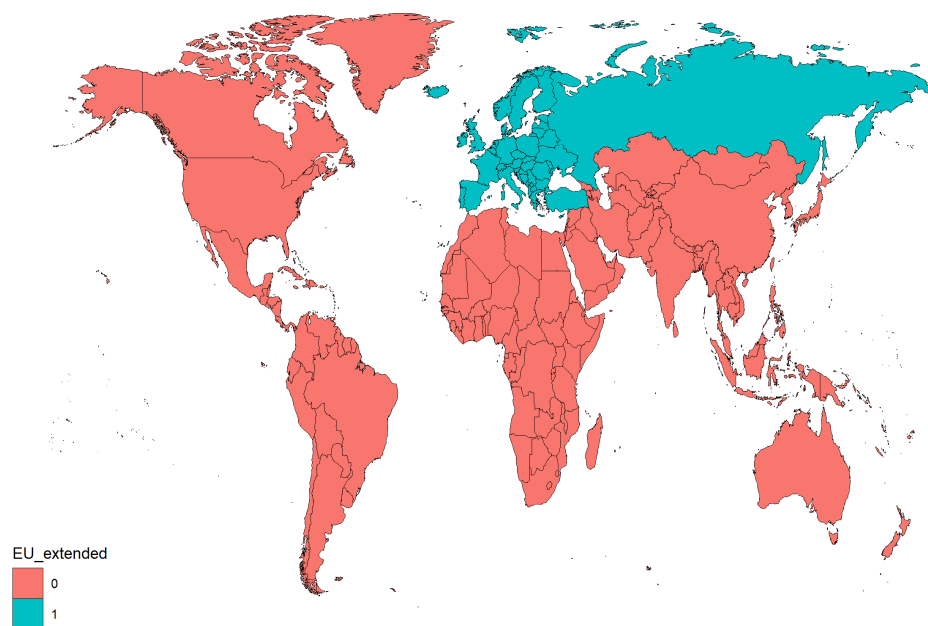# 8   EU_map, EU_extended

A

B

Figure 6 Map of countries included in (A) EU_map and (B) EE_extended.

9 LANDLOCKED, SEALOCKED, SOUTH

Sealocked entities consist of one or more islands or parts of islands (Wikipedia, 2020c) (Figure 5) , whereas landlocked entities are not connected to an ocean (Wikipedia, 2020d) (Figure 6).

Southern countries consist of countries wholly below the equator, as well as Brazil, Democratic Republic of Congo, Republic of Congo, Ecuador, Gabon, Indonesia, and Kenya, whose landmass lies predominantly below the equator.
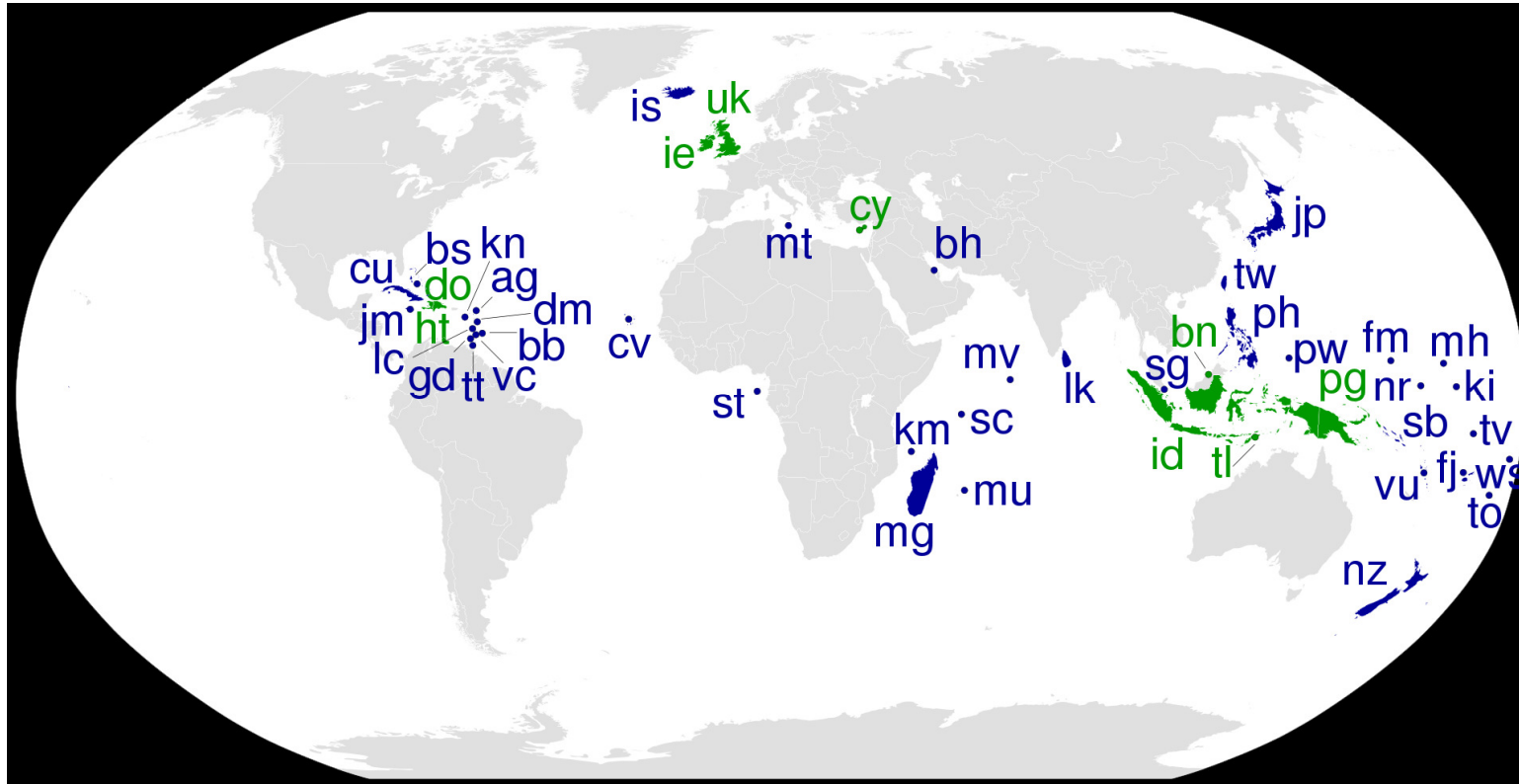
Figure 7 Sovereign states and states with limited recognition fully on islands (Australia is regarded a continent), those with land borders shaded green, and those without shaded dark blue
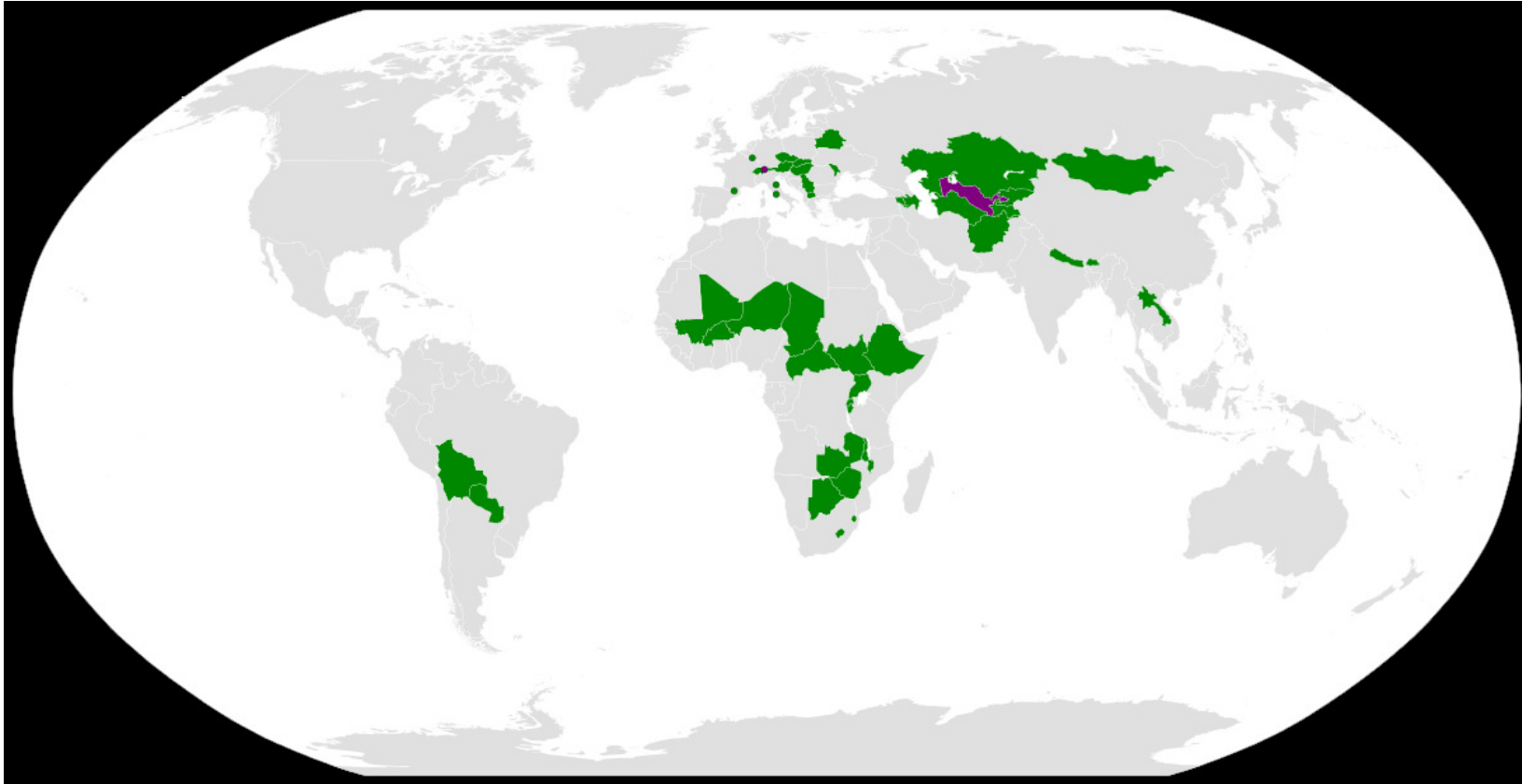
Source: Wikipedia (2020c)

Figure 8 42 landlocked (green) and 2 doubly landlocked (purple) countries
Source: Wikipedia (2020d)

References

Allroggen, F., Wittman, M.D., Malina, R., 2015. How air transport connects the world–A new metric of air connectivity and its evolution between 1990 and 2012. Transportation Research Part E: Logistics and Transportation Review 80, 184–201.

European Commission, 2020. Annex A5 - List of countries, territories and currencies [WWW Document]. URL https://publications.europa.eu/code/en/en-5000500.htm

UN Statistics Division, 2020. Standard country or area codes for statistical use (M49) [WWW Document]. URL https://unstats.un.org/unsd/methodology/m49/

Wikipedia, 2020a. File:United Nations geographical subregions.png [WWW Document]. URL https://commons.wikimedia.org/wiki/File:United_Nations_geographical_subregions.png

Wikipedia, 2020b. Pacific Culture Areas [WWW Document]. URL https://nl.m.wikipedia.org/wiki/Bestand:Pacific_Culture_Areas.png

Wikipedia, 2020c. List of island countries [WWW Document]. URL https://en.wikipedia.org/wiki/List_of_island_countries

Wikipedia, 2020d. Landlocked country [WWW Document]. URL https://en.wikipedia.org/wiki/Landlocked_country

World Bank, 2018. Atlas of Sustainable Development Goals 2018: World Development Indicators. World Bank, Washington, DC.