



POLS 7387 Global Governance
Professor Denise Garcia
Retrieving and Preparing Mapping Data in Excel

Important Vocabulary

- **Excel:** software that reads .xlsx and .csv files (tabular data)
- **Workbook:** the entire Excel file
- **Sheet:** individual sheets within an Excel workbook file
- **Columns:** vertical data (letters on top of Excel workbook)
- **Rows:** horizontal data (numbers on side of Excel workbook)
- **Metadata:** information about your data, such as what different column variables represent
- **Geospatial Mapping:** digital mapping that uses data, especially latitude and longitude coordinates, to map information

Dataset

- UN SDG Database: <https://unstats.un.org/sdgs/dataportal> > data that needs to be processed before geospatial mapping
- MetaData of UN SDG Database: <https://unstats.un.org/sdgs/metadata> > Metadata is essential in order to understand what the dataset does show, and what it does not show.

File Types

- .xlsx = Excel file; and can usually only be opened in Excel
- .csv = comma-separated values, an open format for delimited data organized into rows and columns. Can be read by Excel & mapping software programs like Google My Maps and ArcGIS
- .kml = mapping file that can be fed into mapping software such as Google My Maps

Data Cleaning in Excel for the Global Sustainable Development Goals Indicators Database

1. Go on the UN Sustainable Goals Website where the data is held:
<https://unstats.un.org/sdgs/dataportal>
2. From here, click on 'select indicators and country area' and look through to determine which goal(s) you would like to focus on. Remember to explore the Metadata repository of the goal you chose so you understand what is in the data
3. Pick the data you want to download – **make sure to be as specific as possible**



Indicators (Selected 2 of 669 series)
Country or Area (Selected 213 of 213)

☐ All

☒ **GOAL 1** End poverty in all its forms everywhere

- ☒ **TARGET 1.1** By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day
- ☐ **TARGET 1.2** By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions
- ☐ **TARGET 1.3** Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable
- ☐ **TARGET 1.4** By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including

Show Results

United Nations

Department of Economic and Social Affairs
Statistics • SDG Indicators Database

Home
SDG Indicators
Data
SDG Reports
HLG-PCCB
IAEG-SDG's
Events
Resources

Data Series (Selected 2 of 669)

+ Select

Geographic Areas (261 of 213)
All Groupings
Countries

By default Regional Groupings is selected. You can select a different geographic area

+ Select

Period
Range
Years

You can select single year or multiple years

1,602 observations

Important Information

Please select parameters from each of the sections on the left, and click on 'Show Results' to perform a search.

Database last updated on Friday, December 15, 2023. [See History](#)

If you need help using this site, [Read FAQs](#)

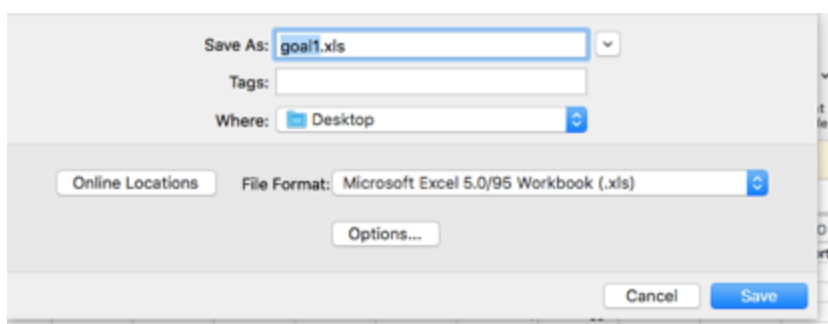
For latest reference metadata information for the Tier I and II indicators in the global indicator framework, explore the [Metadata Repository](#)

This new database interface has been launched on 15 September 2021. Please use the link below for questions and feedback. The old interface is still available for limited time at <https://unstats.un.org/sdgs/indicators/databaseLegacy>.

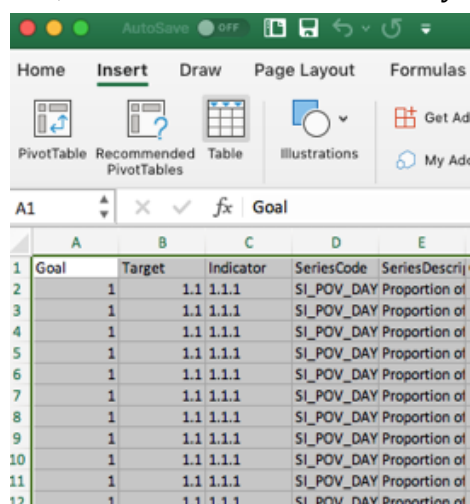
- Download the data by clicking the “download” button on the bottom of the webpage
- Once your data is downloaded, it should appear on the bottom of your web browser → click on it to open (it should then open in Excel as a workbook)
- The data needs to be saved as an Excel file, but right now it is saved as a CSV. To change the format, go to File → Save as. Name your file something you will remember and select .xlsx file where it says File Format

Questions? Contact Us!

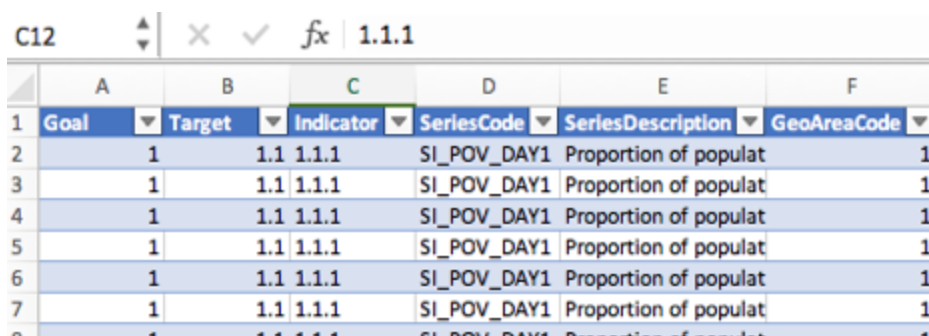
Schedule a meeting: <https://calendly.com/diti-nu>



7. Now we want the data to be in Table format so that we can clean it. First, select all the data [control + a/command + a]
8. Next, click Insert → Table → Okay



9. The file should then look similar to this in terms of format:



Goal	Target	Indicator	SeriesCode	SeriesDescription	GeoAreaCode
1	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
2	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
3	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
4	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
5	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
6	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
7	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
8	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
9	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
10	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
11	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat
12	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat

10. Now you can begin cleaning the data using the different columns. Filter the data to what you would like to include by clicking the column arrows next to each variable name. Then select the data you would like to keep in your file.

	R	S	T	U	V
	Age	Location	Nature	Reporting Type	Sex
25+			Sex		BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+					BOTHSEX
25+			M	G	BOTHSEX
25+			M	G	BOTHSEX

11. Once the data is filtered, the next step is to copy all of the data (use control + a/command + a again) to put in a new blank Excel file
 - a. After copying the data [control/command + c] → open a new blank Excel file (File → New)
 - b. Paste the data [control/command + v] into the new file and
 - c. Save your new file as something you will remember [File → Save as]
12. You may also want to delete variables that you do not need for mapping. To do so highlight the column that you want to delete → right/control click → delete

G		
Units	Cut	⌘X
PERCENT	Copy	⌘C
PERCENT	Paste	⌘V
PERCENT	Paste Special...	⇧⌘V
PERCENT	Insert	
PERCENT	Delete	
PERCENT	Clear Contents	
PERCENT	Format Cells...	⌘1
PERCENT	Column Width...	
PERCENT	Hide	⇧⌘0
PERCENT	Unhide	⇧⇧⌘0
PERCENT		
PERCENT		



13. Once you are happy with the way your file looks and it is ready for GIS, the last step is to save the file as a CSV (File → Save as → click "CSV UTF-8 (comma delimited) (.csv)" under file format → Save)

Things to Note

- Remember to always save your work as you go [command/control + s] or File → Save
- Relatedly, save your data using file names you will remember and store them in a folder you will remember later on
- When choosing how to filter your data, make sure that the data is complete
 - For example, if choosing two years (2001, 2002) make sure there is data for both dates for the countries you choose
- To realign your format – select all [control + a] → Wrap Text

Example Steps for Preparing Data - Class Demonstration

Goal 1, Target 1.1, indicator 1.1.1

1. Go to <https://unstats.un.org/sdgs/dataporta>
2. Select Data
 - a. Goal 1, Target 1.1, indicator 1.1.1, called “Employed population below international poverty line, by sex and age (%)”
 - b. All Countries
 - c. Year 2020
3. Download this file
4. Open in Excel → Save as Excel document (.xlsx)
5. Select all cases (control + a/command + a)
6. Insert → Table
7. Clean/Filter the data:
 - a. Select series code SI_POV_EMP1
 - b. Select both sexes
 - c. Select ages <15Y
8. Delete all columns except: value, geo area name, geo area code, time period, age
9. Copy and paste your data into a new spreadsheet and delete age column
10. If there are formatting issues with the columns → select all again (control + a/command + a) and “wrap text”
11. Save it as a .csv