



POLS 7387 Global Governance
Dr. 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:** virtual mapping that uses data, especially latitude and longitude coordinates, to map information

Datasets

- <https://data.humdata.org> > Clean data with different files types that are mostly prepared to be mapped
- <https://unstats.un.org/sdgs/indicators/database> > messier data that needs to be processed before geospatial mapping

File Types

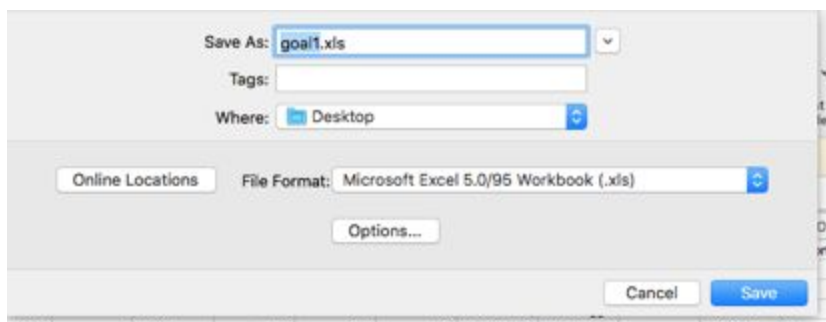
- .xlsx = Excel file, and can usually only be opened in Excel
- .csv = comma-separated values, or tabular data. Can be read by Excel & mapping software programs like Google My Maps and ArcGIS
- .kml = mapping file that can be fed into mapping software such as Google My Maps

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

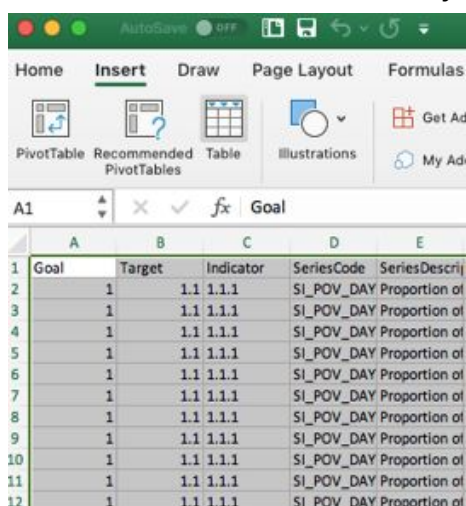
1. Go on the UN Sustainable Goals Website where the data is held:
<https://unstats.un.org/sdgs/indicators/database>
2. From here, look through to determine which goal(s) you would like to focus on. Remember to explore the Metadata repository of the goal you chose so you understand what is in the data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

- **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 poverty lines
 - ☒ **INDICATOR 1.2.1** Proportion of population living below the national poverty line, by sex and age
 - ☒ Proportion of population living below the national poverty line (%) SI_POV_NAHC
 - ☐ **TARGET 1.3** Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial increases in the number of people covered by social protection systems and measures
 - ☐ **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 financial services, markets, and assets
 - ☐ **TARGET 1.5** By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related events and disasters
 - ☐ **TARGET 1.a** Ensure significant mobilization of resources from a variety of sources, including through enhanced development cooperation, in order to implement measures to eradicate poverty in accordance with the Sustainable Development Goals

4. Download the data by clicking the “download” button on the top right corner of the webpage
5. 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)
6. 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. Then name your file something you will remember and select .xlsx file where it says File Format







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



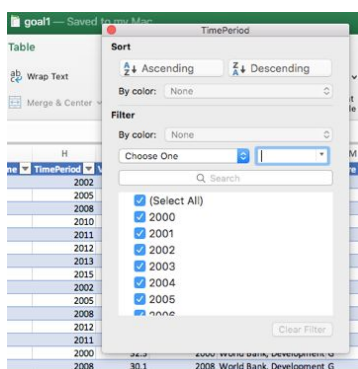
	A	B	C	D	E
1	Goal	Target	Indicator	SeriesCode	SeriesDescription
2	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
3	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
4	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
5	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
6	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
7	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
8	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
9	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
10	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
11	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day
12	1	1.1	1.1.1	SI_POV_DAY	Proportion of population living on less than \$1.25 a day

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

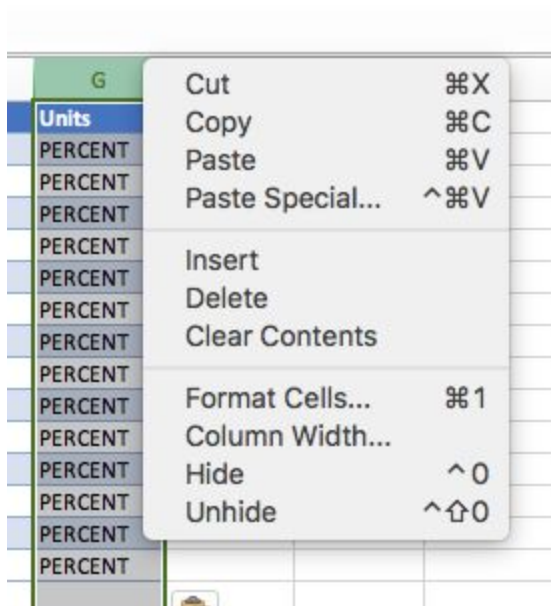
C12     1.1.1

	A	B	C	D	E	F
	Goal	Target	Indicator	SeriesCode	SeriesDescription	GeoAreaCode
1						
2	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat	1
3	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat	1
4	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat	1
5	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat	1
6	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat	1
7	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat	1
8	1	1.1	1.1.1	SI_POV_DAY1	Proportion of populat	1

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.



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
- After copying the data [control + c] → open a new blank Excel file (File → New)
 - Paste the data [control + v] into the new file and
 - 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 click → delete



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 comma separated .csv under file format → Save)

Things to Note

- Remember to always be saving your work as you go [command +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

Goal 1, Target 1.1, indicator 1.1.1	Goal 16, Target 16.2, indicator 16.2.2
<ol style="list-style-type: none"> 1. Go to https://unstats.un.org/sdgs/indicators/datab ase 2. Select Goal 1, Target 1.1, indicator 1.1.1, called “Proportion of population below international poverty line” 	<ol style="list-style-type: none"> 1. Go to https://unstats.un.org/sdgs/indicators/datab ase 2. Select Goal 16, Target 16.2, indicator 16.2.2, called “Age and sex distribution of detected victims of trafficking in persons



<ol style="list-style-type: none">Download this fileOpen in Excel → Save as Excel document (.xlsx)Select all cases (control + a/command + a)Insert → TableClean the data, select:<ol style="list-style-type: none">Year 2010Countries: Bolivia, Chile, Colombia, Costa Rica, Ecuador, Panama, and other specific countries (not regions like “West Asia” or “World”)Delete all columns except: value, country name, country code, year indicatorIf there are formatting issues with the columns → select all again (control + a/command + a) and “wrap text”Copy and paste your data into a new spreadsheet and save it as a .csv	<p>(%)”</p> <ol style="list-style-type: none">Download this fileOpen in Excel → Save as Excel document (.xlsx)Select all cases (control + a/command + a)Insert → TableClean the data, select:<ol style="list-style-type: none">Females onlyOnly those over the age of 15Delete columns so that it is only the value, country name, country code, year indicatorRename “value” to Female_Over15 to understand that’s what it entailsIf there are formatting issues with the columns → select all again (control + a/command + a) and “wrap text”Copy and paste your data into a new spreadsheet and save it as a .csv
---	---