Retrieving and Preparing Data for Mapping in Excel

POLS 7387 Global Governance
Professor Denise Garcia
Fall 2020

Northeastern University
NULab for Texts, Maps, and Networks

Feel free to ask questions at any point during the presentation!

Workshop Agenda

- Collecting and understanding data
- Collecting and preparing data from UN Sustainable Development Goals
- Preparing data for mapping using Excel

Slides, handouts, and data available at

https://bit.ly/diti-fall2020-garcia

Feel free to ask questions at any point during the presentation!

Workshop Objectives

- Know places where you can collect data
- Understand how to collect, store, and clean data in Excel
- Apply these methods to the UN Sustainable Development Goals data
- Prepare the UN SDG data for GIS mapping

Feel free to ask questions at any point during the presentation!

Collecting Data

- A dataset is a collection of several pieces of information called variables (usually arranged by columns).
- A variable can have one or several values (information for one or several cases).
- Qualitative (textual, interpretable) data vs quantitative (numerical) data

Country	Year	GDP
USA	2015	19.39

Feel free to ask questions at any point during the presentation!

Where Can You Get Data to Map?

https://unstats.un.org/sdgs/indicators/database > messier data that needs to be processed before geospatial mapping. These data are more globally focused.

https://unstats.un.org/sdgs/metadata/ > data about datasets. Unavoidable in order to understand what the dataset does show, and what it *does not show*

Feel free to ask questions at any point during the presentation!

Understanding Your Data

- Where does this data come from? Who collected it? For what purpose?
- Metadata for your data – use the metadata to understand what the particular variables represent.
    - For example, you might see unclear column names. Some data will have metadata attached, which may explain what the columns mean.
- Important data for mapping: *geo coordinates* (latitude and longitude) are necessary to produce any maps, just the names of countries and cities won't be sufficient.

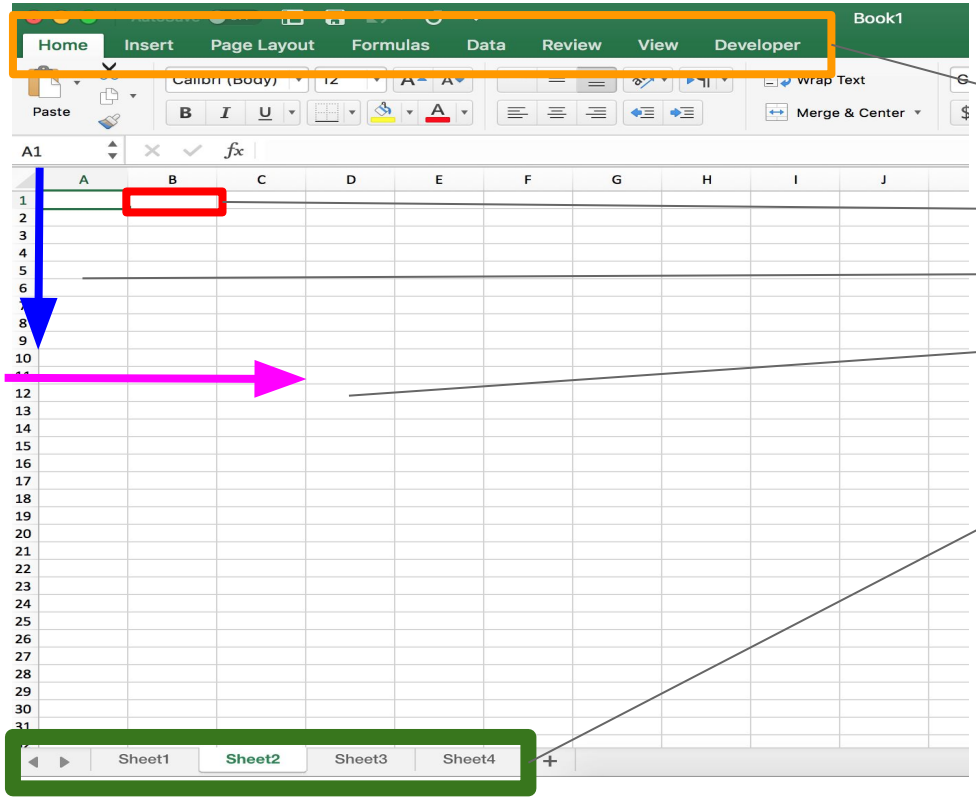
Feel free to ask questions at any point during the presentation!

Excel

Excel is a program that is used to create and edit tabular data (spreadsheets). In Excel, data are organized into rows and columns; data can be presented and analyzed using Excel's functions, such as pivot tables, charts, formulas, and more.

Feel free to ask questions at any point during the presentation!

Anatomy of Excel



TOOLBAR

CELL

COLUMN

ROW

SHEET

Feel free to ask questions at any point during the presentation!

Hands on activity: Download and prepare data

Follow the instructions from the handout when preparing your data

Example in the handout: Collect data from this goal:
Select **Goal 1, Target 1.1, indicator 1.1.1** "Employed population below international poverty line, by sex and age (%)"

Prepare this data and save it as a .CSV file:
Both sexes over 25 years of age in 2010

Feel free to ask questions at any point during the presentation!

Collecting & Preparing Data

Follow the instructions on handout-data_prep_excel:

- Go to https://unstats.un.org/sdgs/indicators/database
- Select the first goal: **Goal 1, Target 1.1, indicator 1.1.1** "Employed population below international poverty line, by sex and age (%)"
- Create a table from your Excel data
- Clean your data: choose the columns and variables to be mapped; make sure there are only **one** set of variables for each country
    - Select "2010" in year
    - Select only *countries* in the country variable (ignore "World," "North America", etc). We want *countries* because they have geocoordinates
    - Select both sexes
    - Select age 25+
- Copy and paste into a new Excel document. Save that document as .CSV

Feel free to ask questions at any point during the presentation!

In the next episode...

- How to add geo coordinates to your dataset
- How to make a map out of a dataset

Feel free to ask questions at any point during the presentation!

Thank you!

If you have any questions, contact us at:

Developed by Cara Messina
Digital Integration Teaching Initiative
DITI Research Fellow

Taught by Milan Skobic
Digital Integration Teaching Initiative
Assistant director

Slides, handouts, and data available at **https://bit.ly/diti-fall2020-garcia**

Feel free to ask questions at any point during the presentation!