

Introduction

Your task is to ingest, clean and visualize global cropland area data obtained from the Food and Agricultural Organization (FAO). This task should be done within the next 24 hours and should not take longer than 7-9 hours to finish. You can use either Python or R to complete this task.

The input data required to finish this task is in 2 excel files, namely:

FAO_Crops_area_harvested_Data.xlsx and *crop_codes_names.xlsx*

FAO_Crops_area_harvested_Data.xlsx contains the harvested crop area for multiple crops for each country from 1961 to 2013. Some key columns in this file are:

Country Code: Unique identifier for each country

Country: Name of each country

Item Code: Unique identifier for each crop

Item: Name of crop

Element Code: Identifier corresponding to reported activity (area harvested in this case)

Element: Reported activity name (area harvested)

Unit: Units of area harvested (hectares)

Y1961, Y1962 ... Y2013: Columns representing the crop harvested area for each year from 1961 ... 2013.

Y1961F, Y1962F ... Y2013F: Redundant columns, can be ignored

Crop_codes_names.xlsx is a lookup table that provides a mapping between the crops in *FAO_Crops_area_harvested_Data.xlsx* (identified by *Item Code*) and 5 functional crop types (C3annual, C3perennial, C4annual, C4perennial, N-fixing).

Tasks

Please perform the following steps and provide us with your documented code, data files from steps 4 and 7, and images from steps 5 and 6.

1. Read in input file and drop/delete all columns with column names ending in 'F' i.e. Y1961F, Y1962F ... Y2013F.
2. The file contains data for several continents and regions as well. Please drop/delete these rows. The continents/regions to be dropped can be identified as those rows with a 'Country Code' value exceeding 1000.
3. Read in data from *Crop_codes_names.xlsx*
4. Merge data from step 2 and step 3 on the *Item Code* column and provide us with the resulting data in a excel or csv file.
5. Create a stacked plot showing the cropland area in the top 10 countries by cropland area (averaged from 1961 - 2013). Please provide the result as a TIF or PNG image.

6. Create a stacked plot showing the area of each crop functional type in the top 10 countries by cropland area (averaged from 1961 - 2013). Your results should look similar to figure 1, but might have different units/values in case you report in hectares. Please provide the result as a TIF or PNG image.

7. As you can see in *FAO_Crops_area_harvested_Data.xlsx*, FAO data separates the countries Belgium and Luxembourg starting in year 2000, but before that it treats them as a single country called Belgium-Luxembourg. Your task is to compute separate crop area values for Belgium and Luxembourg. Please do this by computing the average fraction of crop area in each country (based on data from year 2000 onwards), and applying this fraction to the previous years data. Please provide the resulting data as an excel or csv file.

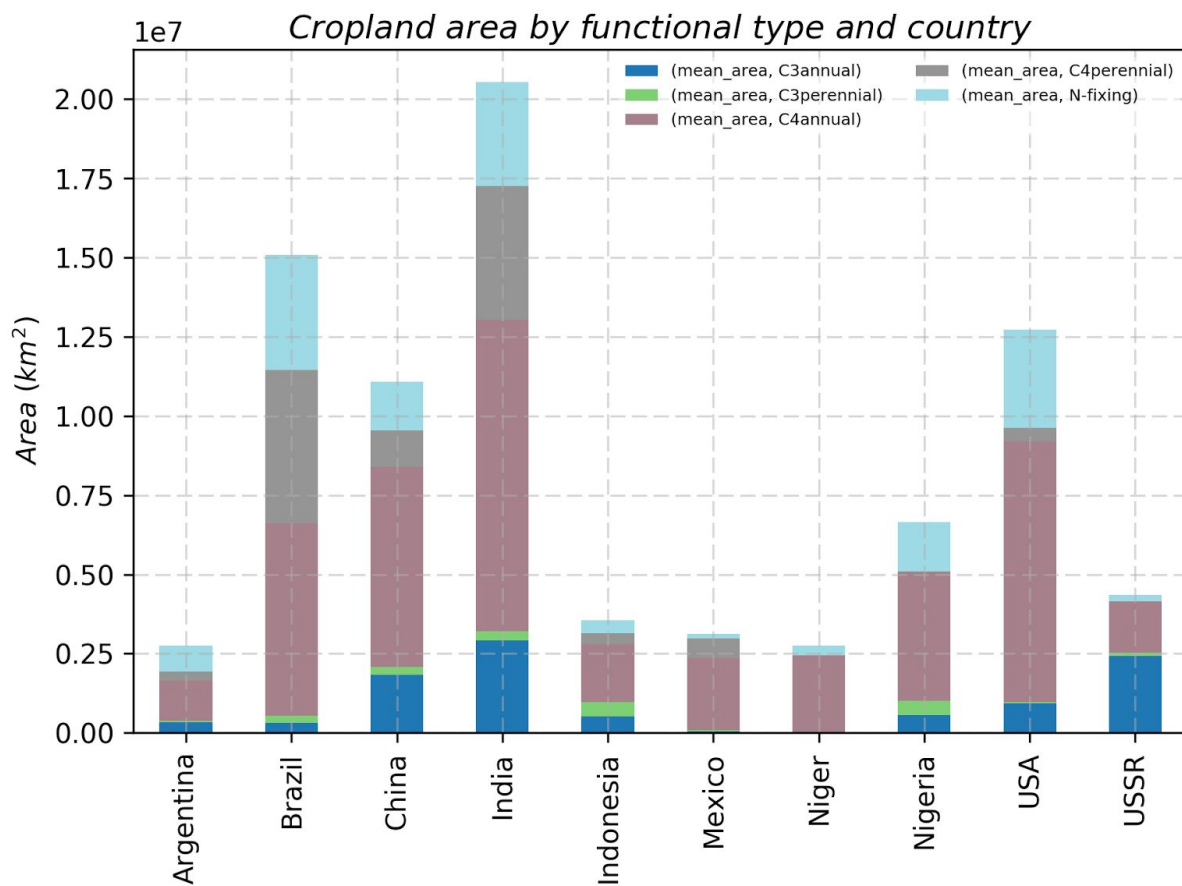


Figure 1. Global cropland area by crop functional type and country