Creating a Raster Extract

Exercise

Objective: Use the IPUMS Terra website to obtain customized datasets that can be used to answer research questions. This exercise uses raster datasets to explore education and agriculture in two economically different countries.
RESEARCH QUESTIONS

Question 1
Examine relationships between education and agriculture in Zambia and Switzerland. Is there a correlation between educational attainment levels and crop/pasture land use in each country? Is the relationship similar across the two countries?

Question 2
Examine the tree cover present in Zambia and Switzerland. How does population relate to the presence of broadleaved trees? Compare relationships in both countries, for evergreen and deciduous broadleaved trees where applicable.

OBJECTIVES

- Create an IPUMS Terra account
- Create and download an IPUMS Terra raster data extract
- Use IPUMS Terra to rasterize area-level data

IPUMS TERRA VARIABLES

Area-level variables

POPTOTAL: Total population for tabulated census areas
EDATTAIN: Percent of population with a specific level of educational attainment in each census area

Raster variables

CROPLAND2000: Area used as cropland
PASTURE2000: Area used as pasture
LCBRDEVGRN: Tree Cover, Broadleaved, Evergreen
LCDECIDCL: Tree Cover, Broadleaved, Deciduous, Closed
LCDECIDOP: Tree Cover, Broadleaved, Deciduous, Open
Data Type Descriptions

**Raster**

Raster data describe how the value of a variable varies over space. The data are structured as a grid of cells. Each cell is connected to a location, and contains the value of the variable at that location. For example, in a land cover raster, each cell indicates the type of land cover found at that location.

**Area-level**

Area-level data describe geographic units defined by boundaries. Units are grouped in sets, such as the counties of the United States or the states of Brazil. In IPUMS Terra, these sets of units are referred to as geographic levels. The data are structured as tables, with a row for each unit and a column for each variable. For example, you may have a table with a row for each county in the United States and columns containing the number of males and the number of females in each county.
Registering with IPUMS Terra

The Minnesota Population Center uses a common user management system for several data projects: IPUMS Terra, IPUMS International, IPUMS USA, IPUMS CPS, IPUMS Higher Ed, IPUMS NAPP, IPUMS NHGIS, IPUMS Health Surveys, and IPUMS Time Use. If you have an existing account with any of these systems, you will use the same account for IPUMS Terra.

- Go to https://data.terrapop.org/
- If you have an existing MPC account, click Login. After logging in, you will be directed to the registration page for IPUMS Terra.
- If you do not have an MPC account, click on Sign up to register for access.

Note: Microdata access is NOT required for this exercise. Access to international microdata requires application and approval by the IPUMS International project.

After you have registered with IPUMS Terra, an e-mail will be sent to your account notifying you of approval.

Note: Please be sure to check your trash/spam folders

- Open the e-mail and click on the confirmation link. You will then be logged into IPUMS Terra.
The extract builder website guides researchers through the workflows for building data extracts. Choose the type of data structure you would like to receive as output for further analysis. In this tutorial, we will get raster data output.

- Click on the **Start Extract** button for Raster Data Output
The IPUMS Terra interface for the first step of the workflow consists of the following elements:

**Navigation Bar:** Shows the major steps in the workflow, the sub-steps of the current step, and your progress through the workflow. The navigation bar steps will reflect the workflow you select. The step and sub-step you are currently on are highlighted in green.

**Data Cart:** Provides a summary of the data you have selected to include in your extract. The data cart is updated as you make selections throughout the workflow.

**Availability Grid:** Shows the availability of variables by dataset and enables selection of variables and datasets.

**Variables Panel:** Lists topics for which area-level variables are available. Clicking on a topic will populate the rows of the availability grid with the variables in that topic.

**Datasets Panel:** Lists countries in the IPUMS Terra system, and provides options to filter by time and hide countries without area-level data. Clicking on a continent will list the countries in the continent. Clicking on a country will populate the columns of the availability grid with the datasets available for that country. You may add all countries in a continent to the grid by clicking the “Browse All” line.
Step 5
Browse datasets for Zambia

- Click on **Africa** in the Browse Datasets Panel. Countries are listed alphabetically, with numbers in parentheses indicating how many years of data are available for the country.
- Click on **Zambia**. The available datasets for Zambia appear as columns in the availability grid.

Step 6
Browse datasets for Switzerland

- Click on **Europe** in the Browse Datasets Panel.
- Click on **Switzerland**. The available datasets for Switzerland will appear as columns in the availability grid.

Step 7
Select datasets

- Check the boxes to select year **2000** for both countries.
Step 8
Select Area-level variables

To see available variables, choose a topic in the Browse Variables panel.

- Choose the variable topic *Education*.

The availability grid will be updated with available education variables.

- Choose the variable group *EDATTAIN*, by checking the multi-select box.

**EDATTAIN** is available for Zambia and Switzerland in the year 2000. The variables in the EDATTAIN group will be added to your Data Cart.

You can expand the variable group to see the individual variables by clicking the arrow widget.

*Note: EDATTAIN contains 4 variables for different levels of education.*

*Note: To hide unchecked datasets in each country, click on “show only selected datasets”*

---

### Education Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Switzerland</th>
<th>Zambia</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCHOOLAGE (3)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>LITAGE (2)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>EDATTAIN (4 of 4)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>EDUCLESSPRIM</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>EDUCPRIMARY</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>EDUCSECOND</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>EDUCTERTIARY</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>EDYEARS (1)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Step 9
Add Area-level variables

- Continue Adding variables to your extract by selecting Demographic → POPTOTAL

IPUMS Terra also provides metadata about each variable. To access the metadata, you must first expose the individual variables within a variable group.

Step 10
View Variable Metadata

- Once the variable group is open, click on the individual variable name TOTPOP to get additional metadata (e.g., documentation about the variable, description, availability, and source).
When you have selected both area-level variables and datasets, the NEXT button will become active and turn green, allowing you to move on to the next step.

Note: Verify that your Data Cart has the correct number of variables and datasets.

The Navigation Bar indicates that the next step will be to select raster data.

- Click NEXT to move to the Raster Data selection screen.
Now we will begin adding raster data to our cart. You will need variables from the Agriculture and Land Cover topics.

- Click the **Agriculture** topic to list variable categories.

- Click on the **Crop and Pasture lands** variable category.

The Crop and Pasture lands variables will be listed.
• Select two variables, **CROPLAND2000** and **PASTURE2000**, to add them to your cart.

![Variable Selection](image)

Note: Clicking on variable labels will provide additional metadata

• Click on the **Land Cover** topic and then click on the **Global Land Cover 2000** variable category.

![Select Variables](image)

• After clicking **Global Land Cover 2000** select three variables, **LCBRDEVGRN**, **LCDECIDCL**, **LCDECIDOP**, to add them to your cart.

![Variable Selection](image)

• Click **NEXT** to go to the Submit step.
Step 13
Check data cart and Submit extract

Review your cart in the right panel
Your data cart should match the screenshot to the right

- Give your extract a short, descriptive **Extract Title**, maybe, “Zambia and Switzerland, education, land cover and use, population.” The Extract Title will appear in your Extract History.
- You may also provide more detailed **Extract Notes**, perhaps describing why you created the extract. These notes will appear on the extract details page. (The extract details page has not yet been implemented.)

**Submit Extract**

**Extract Details**

**Extract Title**
TerraPop Extract Zambia and Switzerland, education, land cover and use, population

**Extract Notes (Optional)**

- Click **Submit Extract**
Step 14
Download your extract

You will receive an email when the extract is ready

Your Terrapopulus extract 'Terrapop Extract_Zambia and Switzerland, education, land cover and use, population' is ready.

To retrieve your data, codebook, and command files, go to the link below.

https://demo.terrapop.org/extracts/57d3d800-0d745-0133-bf50-08508a37c4e/5_bundle.zip

- To download the data, follow the link in the e-mail. You can also access your extracts on your Account's “Extract History” page as shown below.

The data will be delivered in a compressed format, make sure you have software available to extract the files.

Raster extracts are provided in geoTIFF format, suitable for analysis in GIS or other software for handling spatial data.