IPUMS Training and Development: Requesting Terra Data as Area-Level

IPUMS TERRA

Area-Level Exercise

Objective: Use IPUMS Terra to obtain a customized dataset that can be used to answer research questions. This exercise uses area-level and raster data to explore temporal and spatial changes in Brazilian population and land cover.

IPUMS Terra: Area-level Data Extract Overview

RESEARCH QUESTIONS

Question 1

Examine the interoperable (raster and area-level) datasets for Brazil over three decennial censuses. What trends exist over time in education, employment, agriculture, urbanization, and deforestation?

Question 2

Examine the drivers of natural resource use in Brazil. How does population growth relate to natural resource consumption?

OBJECTIVES

- Create an IPUMS Terra account
- Create and download a IPUMS Terra area-level data extract
- Use IPUMS Terra to generate raster summarizations for geographic areas

TERRAPOP VARIABLES

Area-level variables

POPTOTAL:	Total population for tabulated census areas
UNEMP:	Percent of population unemployed in each census area
PROF:	Percent of workers in professional, technical or managerial occupations in each census area
EDATTAIN:	Percent of population with a specific educational attainment in each census area
AGRIC:	Percent of employed persons in each census area working in agriculture, fishing, or forestry sectors
URBANIZ:	Percent of population living in urban areas

Raster variables

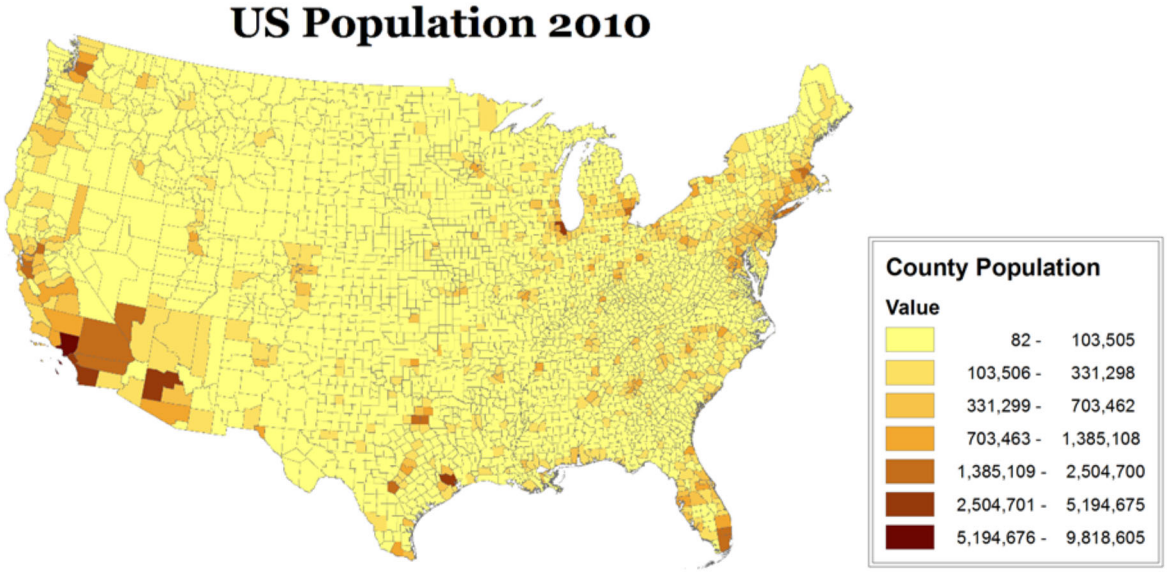
EVRGRNBRDLF:	Area of evergreen broadleaf forest
URBAN:	Urban area
CROPLAND:	Area devoted to crops

IPUMS Terra: Area-level Data Extract Overview

Data Type Descriptions

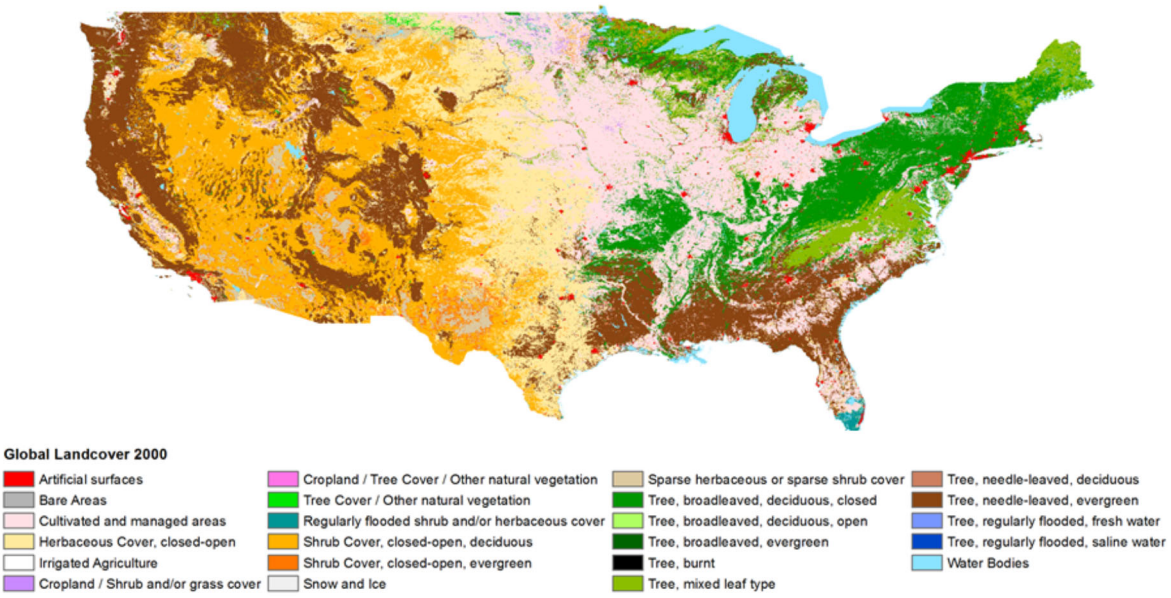
Area-level

Area-level data describe geographic units defined by boundaries. Geographic 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



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.



Step 1
Sign up

IPUMS uses a **common user management system** covering all IPUMS products. If you have an existing account with any IPUMS product, you will use the same account for IPUMS Terra.

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

Step 2
Email confirmation and log in

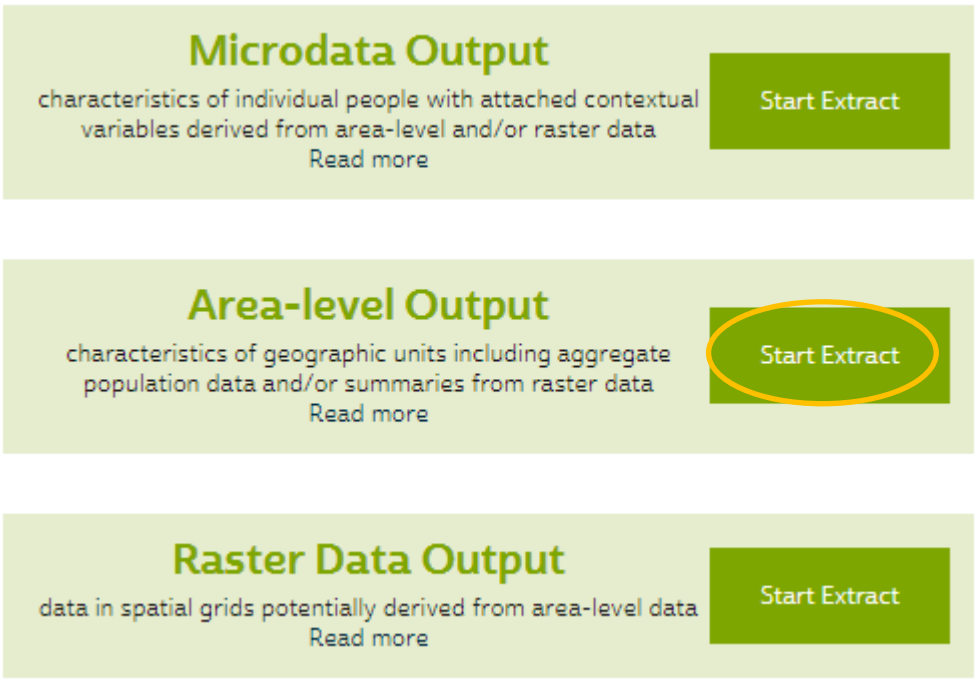
After you have registered with IPUMS Terra, an e-mail will be sent to your account notifying you of approval.
Note: Please be sure to check your trash/spam folders

- Open the e-mail and click on the confirmation link. You will then be logged into IPUMS Terra

Step 3
Start an area-level extract

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 be working with

- Click on the *Start Extract* button for Area-level Output



Step 4

Examine the IPUMS Terra interface

The IPUMS Terra interface for the first step of the workflow consists of the following elements:

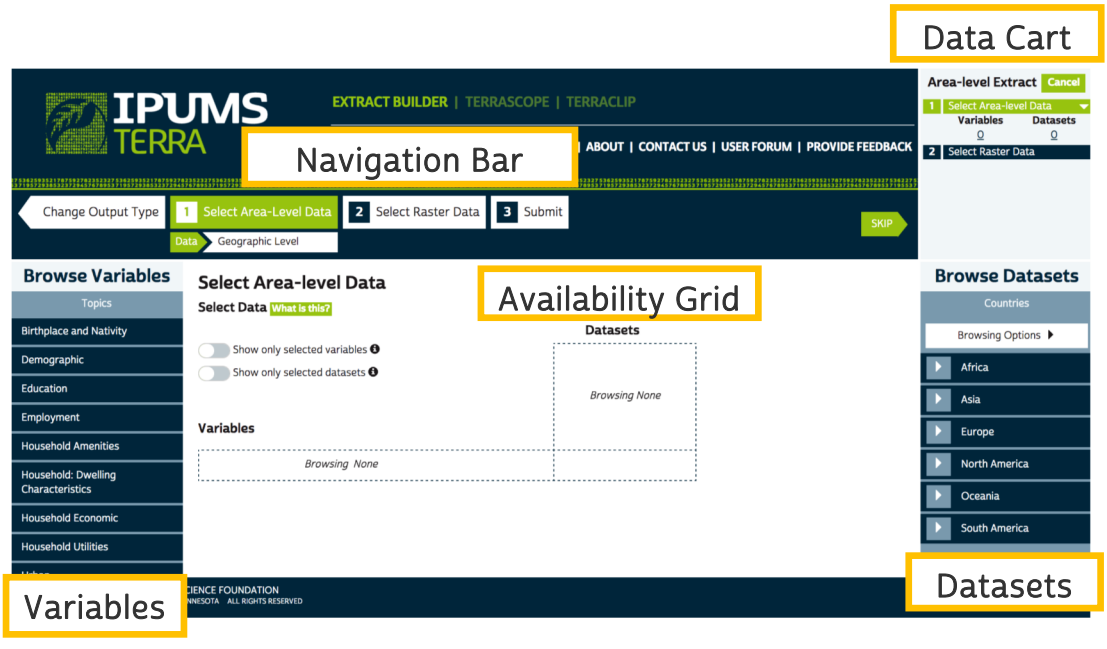
Navigation Bar: Shows the major steps in the workflow, the sub-steps of the current step, and your progress through the workflow. The navigation bar steps will reflect the workflow you select. The step and sub-step you are currently on are highlighted in green.

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.

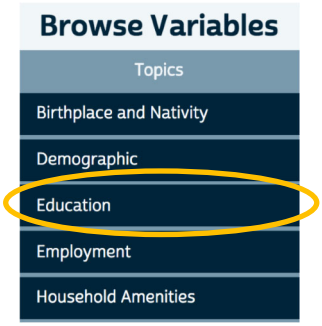
Availability Grid: Shows the availability of variables by dataset and enables selection of variables and datasets.

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 7
Select area-level variables

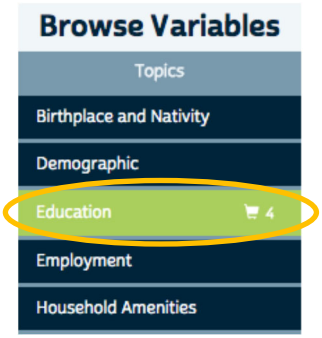


To see available variables, choose a topic in the Browse Variables panel.

- 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.

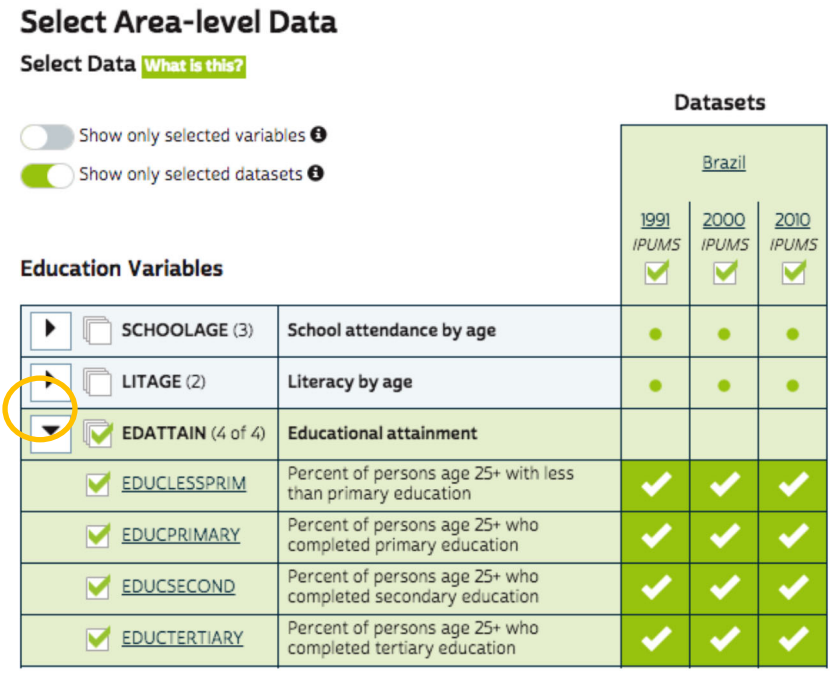


EDATTAIN is available for Brazil in the years 1991, 2000, and 2010. The variables in the EDATTAIN group will be added to your Data Cart.

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

Note: EDATTAIN contains 4 variables for different levels of education

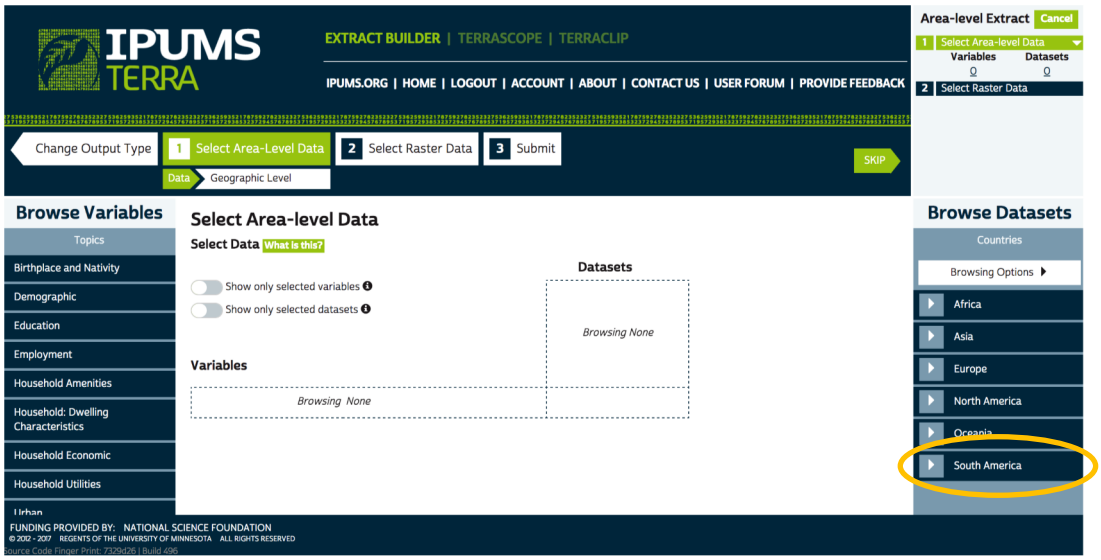
Note: You can hide unselected datasets by clicking the "Show only selected datasets" toggle.



Continue Adding variables to your extract by selecting the following area-level variables:

- Demographic → *POPTOTAL*
- Employment → *UNEMP*
- Employment → *PROF*
- Employment → *AGRIC*
- Urban → *URBANIZ*

Step 5
Browse datasets for Brazil



- Click on *South America* in the Browse Datasets panel.

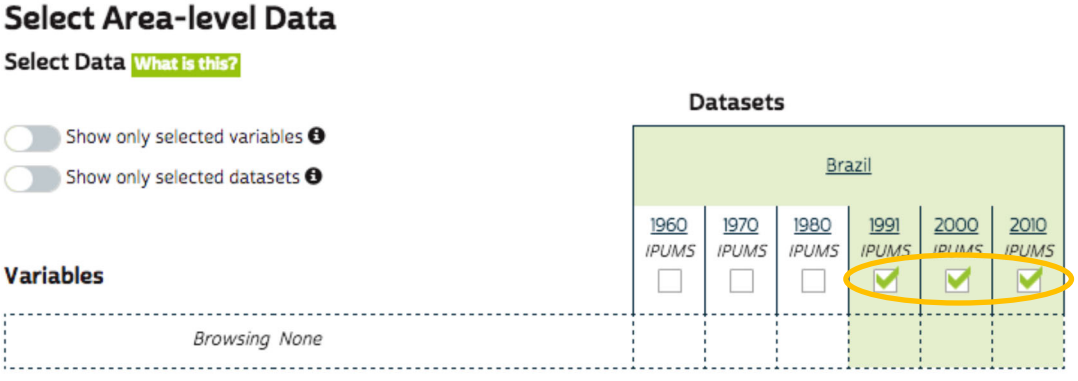
South American countries are listed with numbers in parentheses indicating how many years of data are available for the country.

Step 6
Select datasets

- Click on *Brazil*.

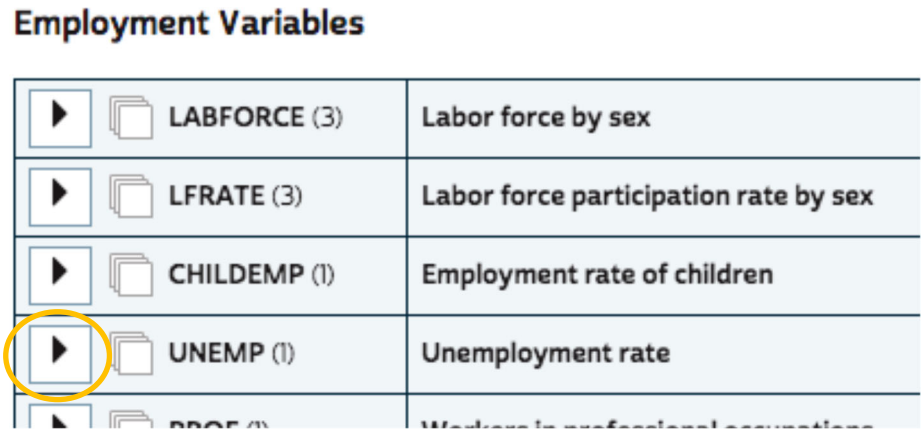
The available datasets for Brazil appear as columns in the availability grid.

- Check the boxes to select years: *1991, 2000, 2010*.

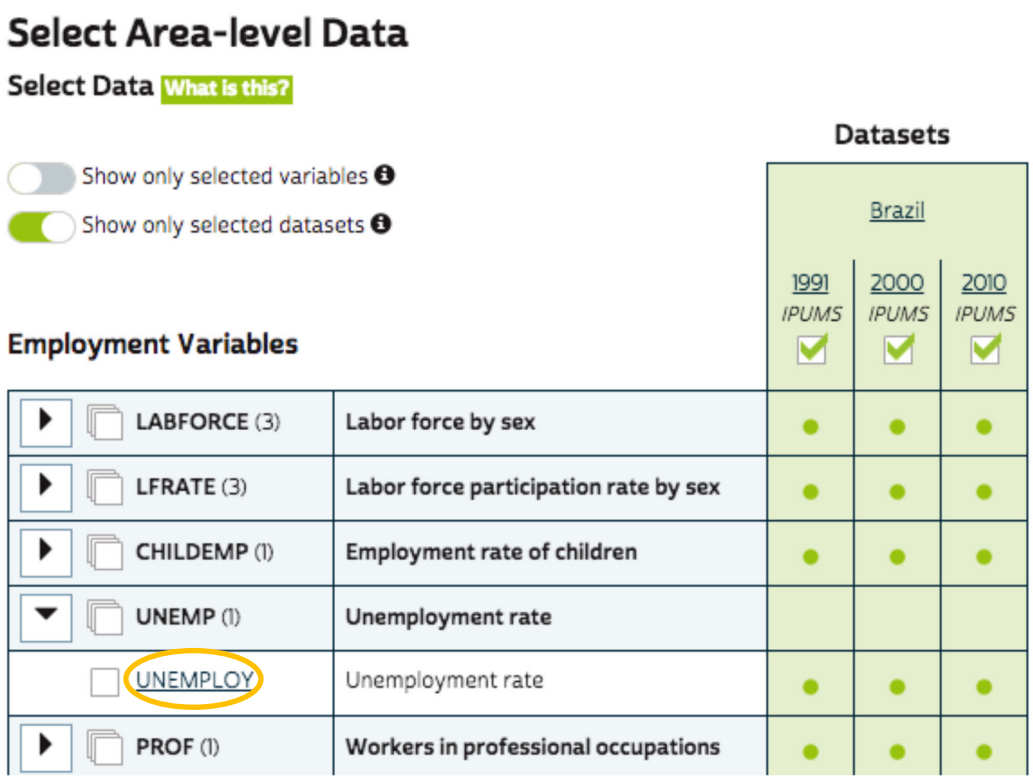


Step 8
View variable metadata

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

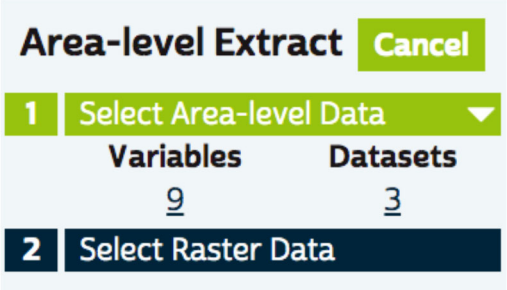


- Once the variable group is open, click on the individual variable name *UNEMPLOY* to get additional metadata (documentation about the variable, description, availability, and source).



Step 9
Select a harmonized geographic level

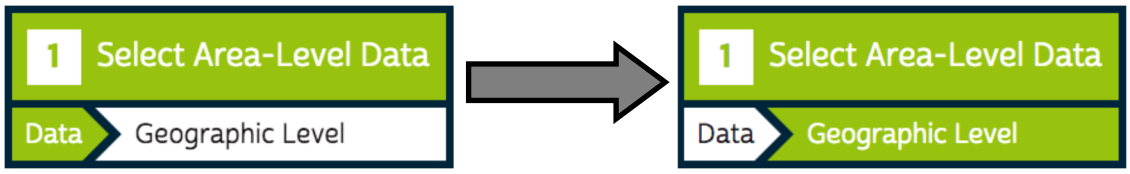
Before moving on to the next step, verify that your Data Cart has the correct number of variables and datasets.



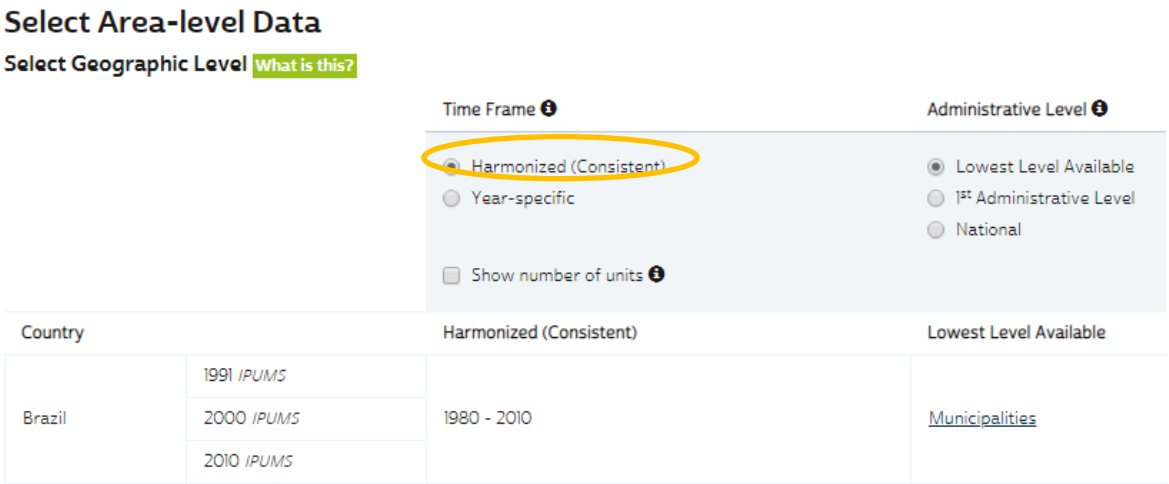
When you have selected both area-level variables and area-level datasets, the NEXT button will become active and turn green.

The Navigation Bar indicates that the next step will be to select geographic levels.

- Click ***NEXT*** to move to the geographic level selection screen.



Our research question involves examining how deforestation has changed between 1991 and 2010. In order to study change over time, it is important to use geographic units with boundaries that are consistent over the time frame under study. In IPUMS Terra, such units are referred to as harmonized. If harmonized units are not used, apparent changes may be due to changing boundaries rather than actual change in deforestation.



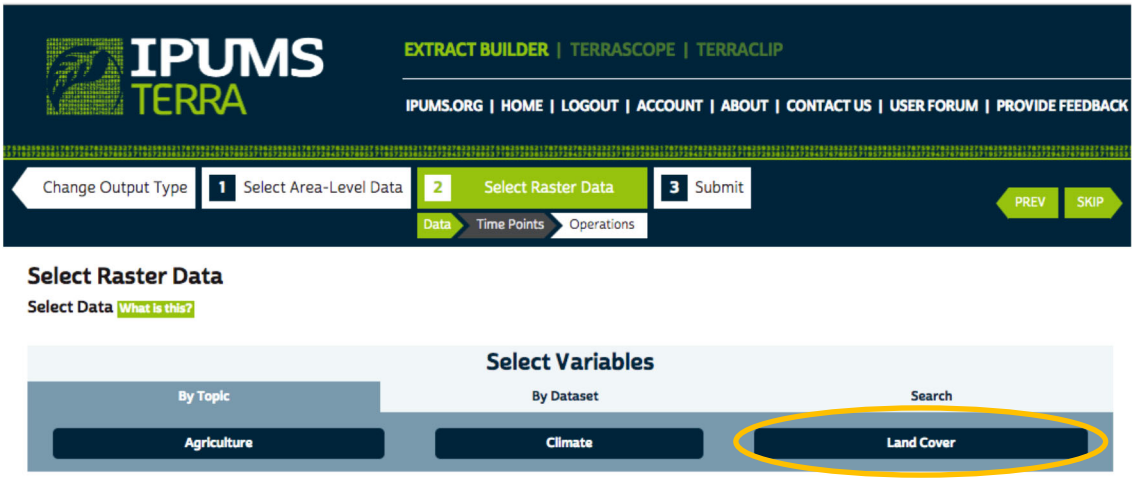
- Leave ***Harmonized (consistent)*** selected
- Click ***NEXT*** to move to raster data selection.

Step 10
Select raster variables

You will need variables from the Land Cover topic

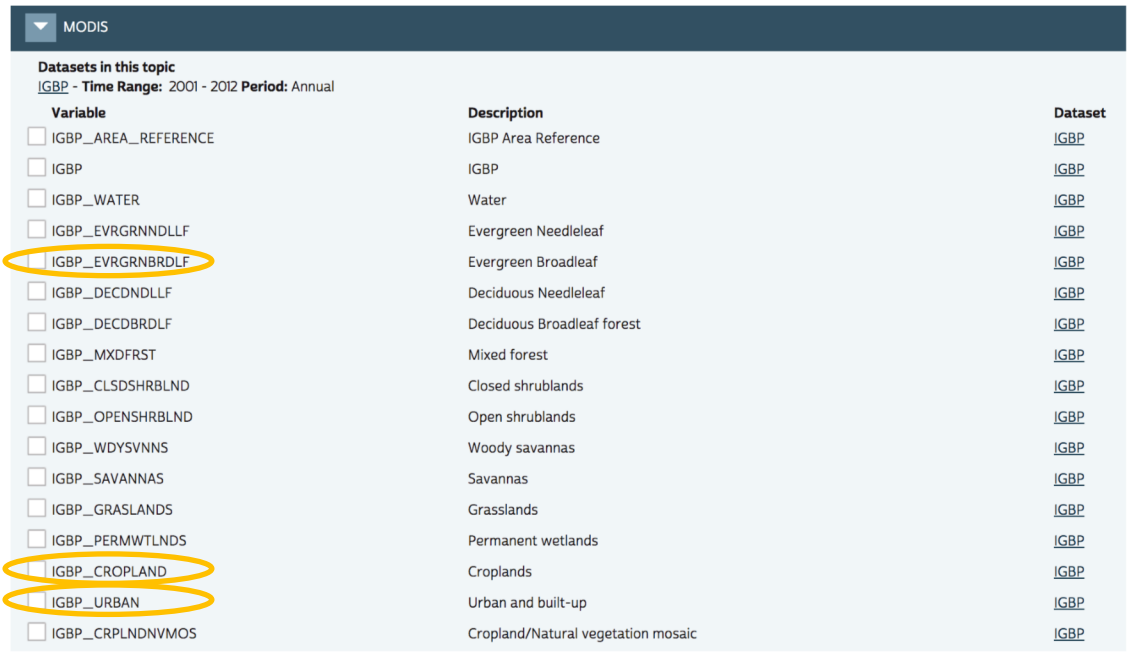
- Click on the ***Land Cover*** topic to list variable categories.
- Click on the ***MODIS*** variable category.

The MODIS variables will be listed.



- Select three variables: ***Evergreen Broadleaf***, ***Urban and built-up***, and ***Croplands*** to add them to your cart.

(The variables may be in a different order than shown below.)



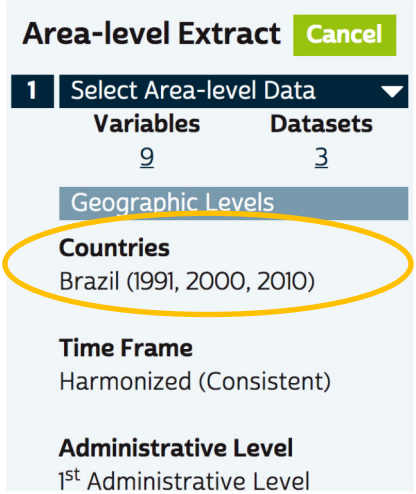
- Click ***NEXT*** when you are finished adding variables.

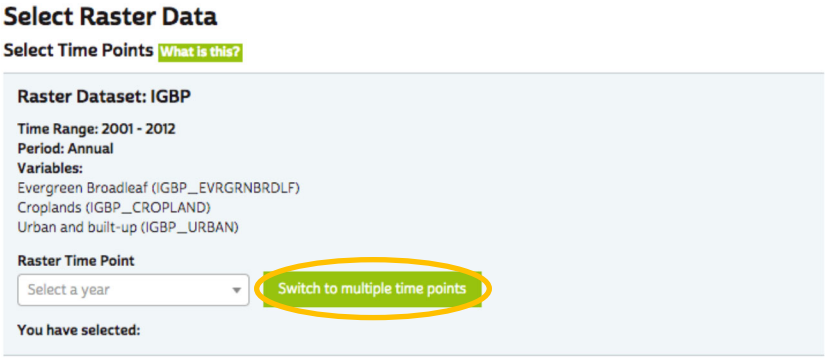
Step 11

Select time points for raster variables

MODIS data are available at annual time steps for 2001-2012. You must select which years to include in your extract. Ideally, the MODIS data years will match the census years of your area-level data.

Note: The Geographic Levels section of your cart shows the years of your area-level data, along with the Countries.





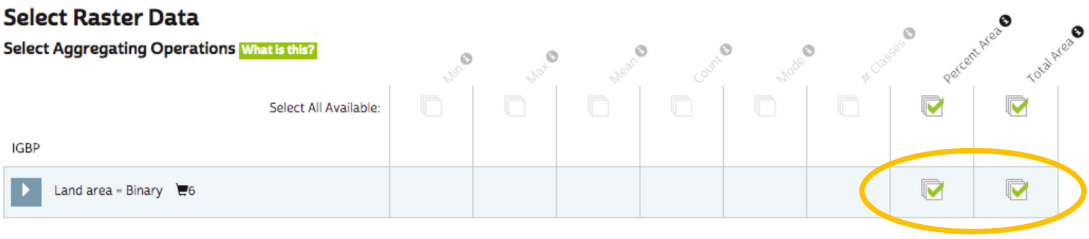
The time points closest to the area-level years are 2001 and 2010.

- Click on years *2001* and *2010*
- Click *NEXT*

Step 12
Select summarization operations for raster variables

Because we are creating an extract for output as area-level data, we must summarize the raster variables over each geographic unit. Depending on the type of raster variable, there are several possible ways to perform the summarization. The raster variables we have chosen are all "Binary" type.
In the original MODIS IGBP data, each cell has a value indicating the type of land cover at that location, such as Urban or Cropland. In the binary variables, the cells with a given land cover type, such as Urban, are given a value of 1 and all other cells are given a value of 0.

To summarize binary variables over geographic units, the options are Percent Area, which will calculate the percent of each unit's area that is covered by the variable's land cover type, and Total Area, which will calculate the total area (in square meters) of the variable's land cover in each unit.

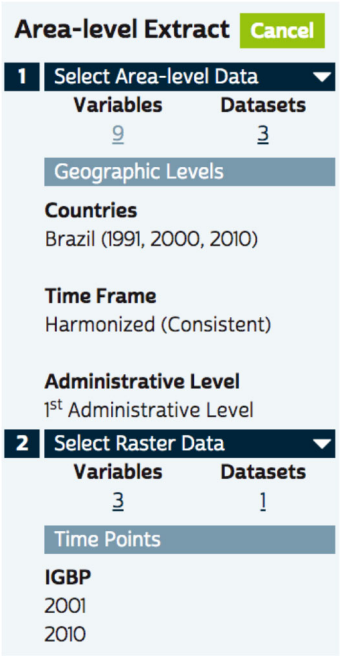


- Check *Percent Area* and *Total Area* on the Land area - Binary line of the table to select these operations for all of your selected variables.
- Click *NEXT* to go to the Submit step.

Note: You can see the individual variables by expanding the Land area - Binary section of the table.

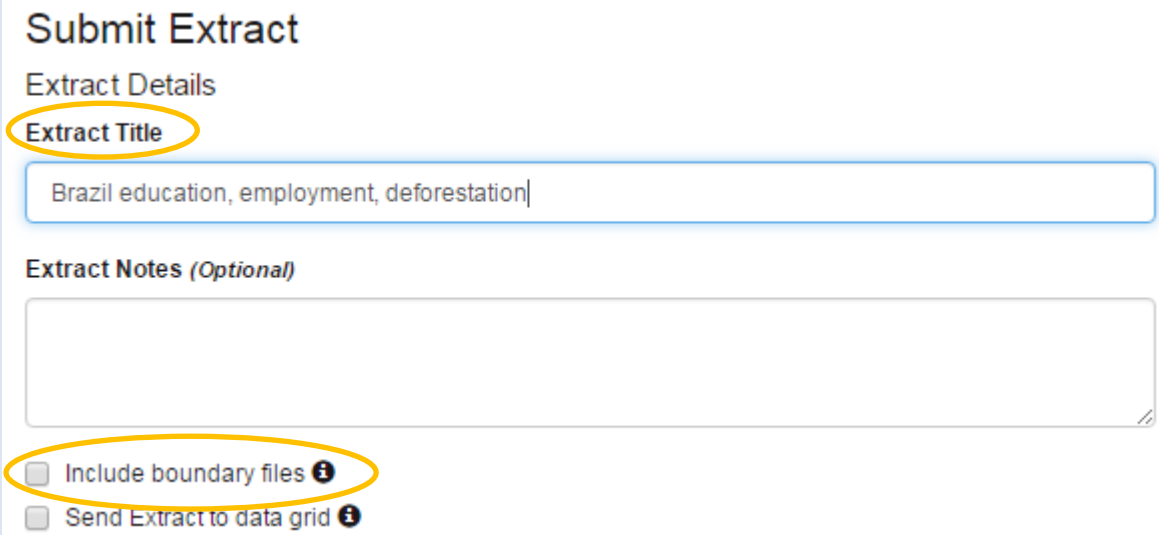
Step 13
Check data cart

Review your cart in the right panel



Step 14

Submit extract



- Give your extract a short, descriptive ***Extract Title***, maybe, "Brazil education, employment, deforestation". The Extract Title will appear in your Extract History.
- If you want the GIS shapefiles for the geographic levels used in your extract, check the box to ***Include boundary files.*** (Not required for this exercise)
- Click ***Submit Extract***

You will receive an email when the extract is ready

ipums@umn.edu
to me

Your IPUMS Terra extract 'Brazil education, employment, deforestation' is ready.

To retrieve your data, codebook, and command files, for Extract #10, go to the link below.

https://demo.terrapop.org/user/extract_requests

Thank you for your support.

Sincerely,
The IPUMS Terra Team

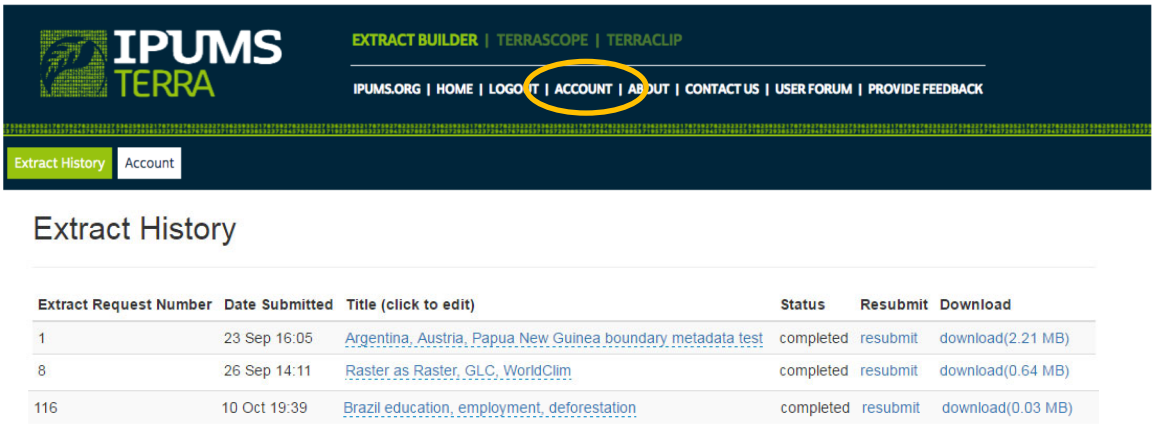
Step 15
Download

your extract

- To download the data, follow the link in the e-mail, which will take you to your account's "Extract History" page as shown below.

This page can also be accessed by clicking *ACCOUNT* in the IPUMS Terra page header

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



Extract History

Extract Request Number	Date Submitted	Title (click to edit)	Status	Resubmit	Download
1	23 Sep 16:05	Argentina, Austria, Papua New Guinea boundary metadata test	completed	resubmit	download(2.21 MB)
8	26 Sep 14:11	Raster as Raster, GLC, WorldClim	completed	resubmit	download(0.64 MB)
116	10 Oct 19:39	Brazil education, employment, deforestation	completed	resubmit	download(0.03 MB)