GTAPSSP: SSPs for GTAP Framework

Thiago Simonato

2025-01-23

Table of contents

1 Introduction

This tutorial demonstrates the utilization of the `gtapssp` package in R for data processing. It covers various steps such as reading, transforming, and analyzing data, making it suitable for both beginners and advanced users.

The package provides optimized and user-friendly functions to **download** SSP data, **interpolate** data using **spline** and **beers** methods. The *gtapssp* functions is accompanied by detailed manual, you can also access this manual by running `?gtapssp` in the R console or pressing `F1` on the function name in RStudio.

2 Installation

To use the *gtapssp* package, it's necessary to have *R* installed on your computer, which can be downloaded from here. Additionally, we recommend downloading *RStudio*, available at here, which provides a user-friendly interface to work with *R*.

R install details

R is a versatile programming language, with a focus on statistical computing. It is a big part of academic research in the social sciences. R is free and open-source and runs on Windows, Mac OS X, and Linux.

1. Download the R installer from the Comprehensive R Archive Network (CRAN)

- Choose the appropriate installer for your operating system and computer architecture (32-bit or 64-bit).

- If on Mac, you will need to know if you are using an Intel or Apple Silicon (M1) processor.

2. Run the installer and follow the instructions.
3. We recommend install **R Tools**. Many of the packages we will be using in this course require R Tools to be installed.

   If you are on Windows:

- Download the latest version of the software from the R Tools for Windows page.
- Run the installer and follow the instructions.

  If you are on Mac:

- First, install the Xcode Command Line Tools. Go to the R Tools for Mac page and follow the instructions. Note: the precise instructions will vary according to the version of macOS you are using.
- Install the gfortran compiler, as also indicated on the R Tools for Mac page.

IDE install details

We recommend RStudio as IDE for R.
RStudio is by far the most popular IDE used by R programmers. It is free and open-source and comes with a console and syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging, and workspace management.
Alternative: If you are already familiar with Visual Studio Code (VS Code), say because you have already used it for Python, you can also use it for R programming. You will need to do a bit of configuration to get it to work, though. If you choose to use VS Code, download and install the R Extension for Visual Studio Code and remember to follow the instructions on the *Getting Started* section of the extension page.

You can install the development version of *gtapssp* from GitHub with:

```
# If the devtools package is not already installed, please run the disabled line below.
# install.packages("devtools")
devtools::install_github("tsimonato/gtapssp")
```

3 Quick Start: Run the Entire Pipeline

If you'd like to execute the entire pipeline described in this tutorial with minimal effort, you can use the `gtapssp::iiasa_gtap()` function. This function runs all steps, including data

aggregation, interpolation, expansion, label merging, and optional file exports.

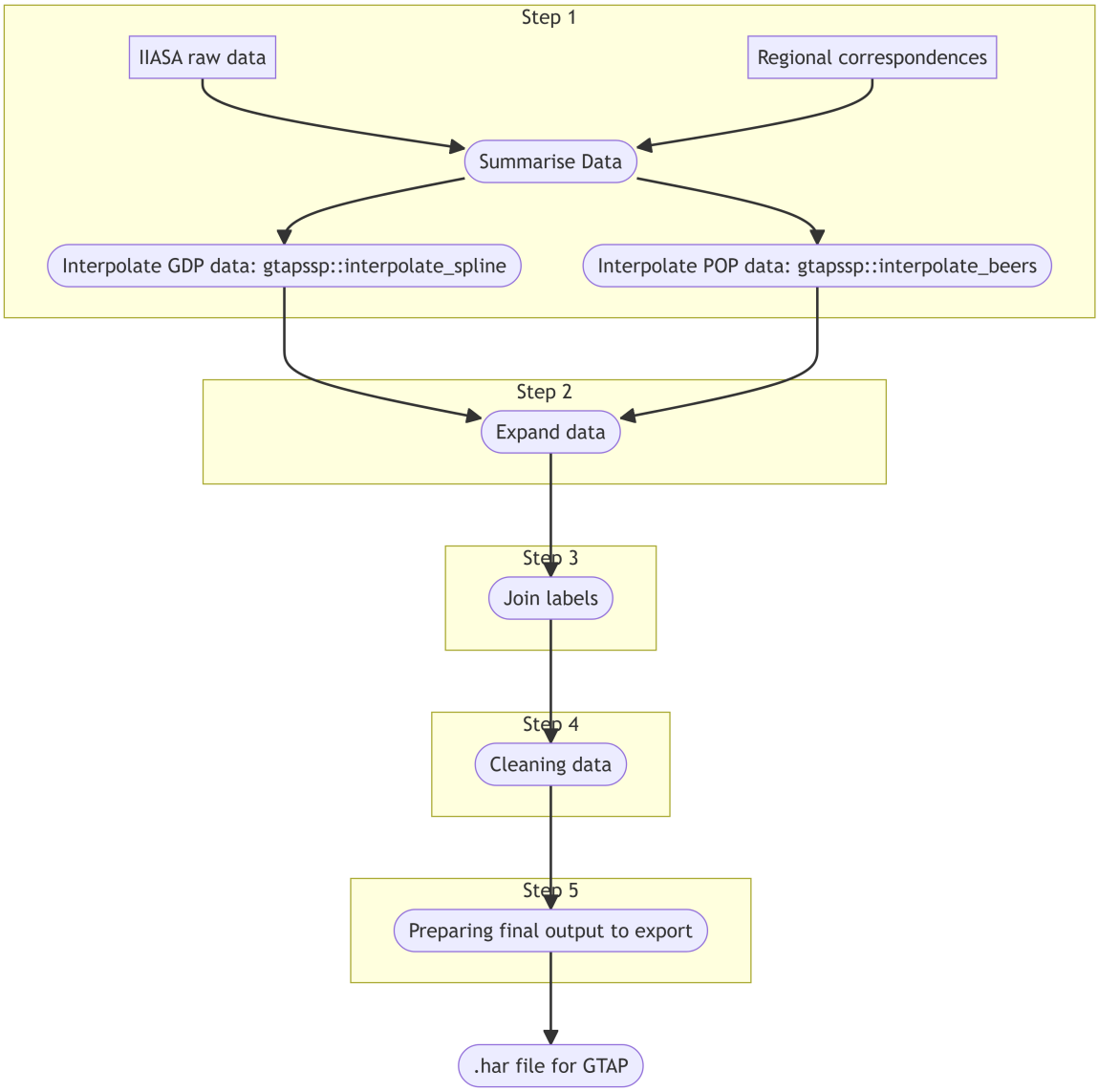
Here's an example:

```
# Run the entire pipeline and save the output as a .HAR file
gtapssp::iiasa_gtap(outFile = "gtap_ssp.har")

# Or save as a .CSV file
gtapssp::iiasa_gtap(outFile = "gtap_ssp.csv")

# To simply process the data without saving it
final_data <- gtapssp::iiasa_gtap()
```

This function provides flexibility to either save the output in .har or .csv formats or return the processed dataset for further use in R.

4 Workflow Overview

The pipeline involves the following steps:



5 Data Source

The pipeline uses projections from the Shared Socioeconomic Pathways (SSPs) developed by IIASA (version 3.0.1, March 2024). These projections include **GDP** and **Population** data and are publicly available under a license allowing reuse by other research communities. For more details, visit the official IIASA SSP database.

Below is a preview of the default IIASA dataset used in the `gtapssp` package. This dataset can be updated with the `gtapssp::updateData` function if newer data or custom updates are

required.

```
gtapssp::iiasa_raw$data
```

Show 10 entries Search:

	model	scenario	region	variable	unit	year	value
1	IIASA GDP 2023	SSP2	Africa (R10)	GDP|PPP	billion USD_2017/yr	2025	7739.894416053859
2	IIASA GDP 2023	SSP2	Africa (R10)	GDP|PPP	billion USD_2017/yr	2030	8869.621612801477
3	IIASA GDP 2023	SSP2	Africa (R10)	GDP|PPP	billion USD_2017/yr	2035	11028.55398825598
4	IIASA GDP 2023	SSP2	Africa (R10)	GDP|PPP	billion USD_2017/yr	2040	13678.58485495506
5	IIASA GDP 2023	SSP2	Africa (R10)	GDP|PPP	billion USD_2017/yr	2045	16798.78489408052
6	IIASA GDP 2023	SSP2	Africa (R10)	GDP|PPP	billion USD_2017/yr	2050	20480.93617311901
7	IIASA GDP 2023	SSP2	Africa (R10)	GDP|PPP	billion USD_2017/yr	2055	24750.77319241611

Showing 1 to 10 of 40 entries Previous 1 2 3 4 Next

6 Step 1: Preprocessing and Aggregation

Aggregate the raw data using `gtapssp::aggData`. This function groups the dataset by specified columns and sums the values within each group, ensuring compatibility with the GTAP regional structure.

```
# Define grouping columns
group_cols <- c("model", "scenario", "reg_iso3", "variable", "unit")

# Aggregate raw data
agg_iiasa <- gtapssp::aggData(
  iiasa_raw = gtapssp::iiasa_raw,
  group_cols = group_cols
)
```