

# Generate Key PPC Indicators Script

Timothy Perez

2023-06-09

## PPC Data Extraction, Transformation, and Key Indicators Generation Script

Below are a series of 4 code sections. Each of these sections needs to be run in sequence. The sequence matters because the data from code 1 will be used in code 2, and the data from code 2 will be used in code 3, etc.

Before running each code, be sure to read the documentation and understand what the code does.

This code is a component of several files that you are likely to have downloaded from Conservation International's Center for Natural Climate Solutions (CNCS) Catalysts Teams' Github. All files should be within one folder called "PPC\_Data". Within the PPC\_Data folder there are 4 additional folders and 2 additional files:

- 1) Final\_Data- Are the Key Indicators that are produced after running all of the R scripts in this document. These data can be used later for further analysis
- 2) Old\_misc\_files- These are old miscellaneous files that I haven't deleted yet, but should probably delete
- 3) R\_scripts- This is the folder that contains all of the R scripts that are referenced in this script and are used to generate the Final\_Data.
- 4) Raw\_Data- This is the folder that contains all of the R data that is downloaded from Kobo, and is the repository for data in intermediate steps before producing the final data.
- 5) (this file) a .pdf called "Generate Key PPC Indicators Script"- sort of like Master R script. This file is used to run all of the R codes, which generate all of the intermediate data, and ultimately leads to the Final\_Data of Key Indicators.
- 6) a README file- A file describing the contents of the folder...
- 7) Tree Restoration Monitoring Framework - Field Test Edition 3.pdf - this file provides an overview of the methods used to monitor and measure the trees in the PPC project.

NOTE: All of the files should be present once downloaded from github, and are necessary for the script to run correctly. The first thing the user should do is enter the working directory, or, the file path to the PPC\_Data folder. To do this, copy and paste the code below into the R console or Script editor and hit enter. **You will be prompted to enter your working directory.**

*Copy and paste this code below:*

```
wd_prompt=readline("Enter the file path to the PPC ")
setwd(wd_prompt)
```

## Code 1: Extract data from Kobo

**Input:** Kobo username and password. User will be prompted for these inputs.

**Outputs:** 7 .csv file stored in a “Raw\_Data” folder

**Description:** This code is the 1st part of the analysis pipeline for key indicators for the Priceless Planet Coalition. This code allows an authorized user to access the Kobotools website and download the data that has been input by users. This code unnests and flattens the the data downloaded from Kobo. The outputs from this code are used for the “Taxonomic\_corrections\_script.R” and “Generate\_data\_for\_PPC\_indicators.R” R scripts.

In this code chunk, **you’re going to be prompted to enter your Kobo username and password.** This is needed to retrieve the PPC data from the Kobo website. Only users with access to the PPC data can download the data.

*Copy and paste this code below:*

```
file_path=getwd()
codePath_extract = paste(file_path, "R_scripts/Kobo_extraction_method.R", sep="/")
source(codePath_extract)
```

## Code 2: Taxonomic corrections

**Input:** The most recent “tree\_data\_PPC\_Data\_XXXX-XX\_XX.csv”, and the most recent Taxonomy\_updates\_XXXX-XX-XX.csv” if the latter exists.

**Outputs:** The final file is: tree\_data\_taxonomy\_corrections\_XXXX-XX-XX, and updated Taxonomy\_updates\_XXXX-XX-XX file is produced.

**Description:** This code is part 2 of the analysis pipeline for key indicators for the Priceless Planet Coalition. This code takes all of the species names entered in Kobo and attempts to resolve taxonomic information and adds family names. This script requires user input to manually correct names or add correct family names when the functions from other libraries fail to retrieve the data.

Note: This code can take a while to run and involves lots of user input. Follow the prompts to enter corrected names, then family names tht can’t be determine for certain species. (For entering unknown family names, I just do a quick google search).

*Copy and paste this code below:*

```
file_path=getwd()
codePath_tax = paste(file_path, "R_scripts/Taxonomic_corrections_script.R", sep="/")
source(codePath_tax)
```

## Code 3: Generate precursor for Key Indicator data

**Input:** The most recent “main\_data\_PPC\_Data\_XXXX-XX\_XX.csv”, and the most recent “tree\_data\_taxonomy\_corrections\_XXXX-XX-XX.csv”./ **Outputs:** The final file is: Data\_for\_PPC\_indicators\_XXXX-XX-XX.

**Description:** This code is part 3 of the analysis pipeline for key indicators for the Priceless Planet Coalition. This code takes takes plot size information and tree abundance for all plots and formats data from two data sources above to prep data to calculate 1) Number of trees restored 2) Number of of trees naturally regenerating, and 3) % survival of planted trees.

Note: As part of this script, users will be prompted to enter some data. A prompt will show what was entered by data collectors for planting dimensions. The user will then translate this into the dimension (in meters) of the plots with the following format: For example a 2 meter by 2 meter plot would be entered 2(\*)2. A 4 meter by 3 meter plot would be entered 4(\*)3. The entries are actually in the format of a multiplication operation. R will use multiply the users' input to get the plot area. This will be used to estimate

*Copy and paste this code below:*

```
file_path=getwd()
codePath_prep = paste(file_path, "R_scripts/Generate_data_for_PPC_indicators.R", sep="/")
source(codePath_prep)
```

## Code 4: Produce Key Indicator data

**Input:** The most recent "Data\_for\_PPC\_indicators\_XXXX-XX\_XX.csv" / **Outputs:** The final data are located in the Final\_Data Folder. There are 6 files: Regeneration, restored, and survival data for plot-levels and site-levels. Currentl

**Description:** This code is part 4 of the analysis pipeline for key indicators for the Priceless Planet Coalition. This code calculates regeneration, restored, and survival

*Copy and paste this code below:*

```
file_path=getwd()
codePath_prep = paste(file_path, "R_scripts/Key_Indicators_Script.R", sep="/")
source(codePath_prep)
```