

Occurrences data tutorial

Main recommendations:

- Use clear names for your species name. We suggest use underline to avoid any error in R or your preferred modeling platform
- Never truncate coordinates!
- Please use G or H to describe germplasm and other accessions source respectively
- Curate your geographical data. Please check coordinates always!
- Please keep the names of the columns as it is shown in this presentation. Never modify the names to avoid problems at running the analysis!
- Please save your files in CSV format!

SDM occurrences data tutorial

Wallace example

Prepare your occurrence data with scientific_name, latitude and longitude

- Please format your occurrences to this format:

scientific_name	latitude	longitude
Bactris_gasipaes	3.976222	-76.184472
Bactris_gasipaes	-0.956542	-77.860912
Bactris_gasipaes	0.183	-78.918167
Bactris_gasipaes	0.123425	-78.620453

Columns:

- Scientific name (species name using underline or without underline. (We suggest use underline to avoid any error in R or your preferred modeling platform),
 - latitude in decimal coordinate format
 - longitude in decimal coordinate format

CWRGA occurrence data tutorial

Prepare your occurrence data to fit the table format

- Please open the file [CWRGA_input_format.csv](#) to obtain a template

Species	latitude	longitude	type
<i>Cucurbita_cordata</i>	23.375	-109.775	H
<i>Cucurbita_cordata</i>	23.4513848	-109.9279362	H
<i>Cucurbita_palmata</i>	34.1233	-112.95191	G

Columns:

- species (species name using underline to separate genus and species),
 - latitude in decimal coordinate format
 - longitude in decimal coordinate format
- type: It has the values H for occurrences not found in the germplasm collection and G to represent the germplasm occurrences for Crop Wild Relatives

LGA occurrence data tutorial

Prepare your occurrence data to fit the table format

- Please open the file [Triticum turgidum example](#) to obtain a template

crop_name	Latitude	Longitude	status	source_db	database_id
triticum_turgidum	31	51G		GENESYS	922
triticum_turgidum	31	51G		GENESYS	935
triticum_turgidum	34.716667	32.533333G		GENESYS	937

Columns:

- crop_name (species name using underline to separate genus and species),
 - Latitude in decimal coordinate format
 - Longitude in decimal coordinate format
- Status: It has the values H for occurrences not found in the germplasm collection and G to represent the germplasm occurrences for landraces
- source_db: source where the accession was extracted
- database_id represents the database id for the record