Work Instructions

Number: **029**

Name: **R Setup**

Title: Setting up of the software application R for manual preparation of a dataset and reporting

Status: Approved

Version: 1.0

Date: 2023-06-30

Authors: Gilles San Martin, Michael Rubinigg

This document is intended for the regulation of processes within the EU Pollinator Hub. It is published under the CC BY 4.0 licence. You may distribute, transform, adapt and build upon the material in any medium or format, including for commercial purposes, as long as you give credit to the author.

Table of contents

1 Document history 3

1.1 Release 3

1.2 Change history 3

2 Abbreviations 4

3 Introduction 5

4 Validity 6

4.1 Area 6

4.2 Time 6

4.3 Related Documents 6

5 Procedure 7

5.1 Installation of the software 7

5.2 Installation of required R packages 8

6 References 12

7 Annexes 13

1 Document history

1.1 Release

1.1.1 Drafted

Submitted on 2023-01-26 by Michael Rubinigg, quality manager.

1.1.2 Reviewed

Reviewed on 2023-06-30 by Noa Simòn Delso, general manager.

1.1.3 Approved

Approved on 2023-06-30 by Michael Rubinigg, quality manager.

1.2 Change history

n.a.

2 Abbreviations

IDE	Integrated Development Environment
n.a.	not available
WI	Work Instructions

3 Introduction

The document describes the process of setting up the software application R for manual preparation of datasets.

4 Validity

4.1 Area

The validity of this document is limited to the process of setting up the software application R for manual preparation of datasets.

4.2 Time

The period of validity of this WI begins on 2023-06-30 and ends with its revocation.

4.3 Related Documents

n.a.

5 Procedure

5.1 Installation of the software

5.1.1 R Language

5.1.1.1 Access the software product from operating system Windows

- Access the official R website (CRAN, https://cran.r-project.org/bin/windows/base/).

5.1.1.2 Access the software product from UNIX based operating systems (*e.g.* Linux, MacOS)

- The language R is installed through any suitable package manager.

5.1.1.3 Download the required version

- Download an R Language version ≥ R 4.0

5.1.1.4 Installation on operating system Windows

- Follow the instructions in the setup wizard.

5.1.2 RStudio Desktop

5.1.2.1 Access the software product

- It is recommended, but not mandatory to use the IDE RStudio Desktop for working with the R Language. This application has been specifically designed for data scientists using the language R.
- Access the IDE RStudio Desktop from the posit website (https://posit.co/download/rstudio-desktop/).
- Other IDE such as Visual Studio Code, Emacs text editor or Eclipse StatET may be used as well, but some of the functions used on the EU Pollinator Hub (*e.g.* the compilation of reports) might not be available.

5.1.2.2 Download the software product

- Download the desired version for the required software product

5.1.2.3 Installation on operating system Windows

- Follow the instructions in the setup wizard.

5.2 Installation of required R packages

5.2.1 Instructions for installation of additional software

5.2.1.1 General instructions for installation of packages

- R packages are extensions of the R language written by various contributors. They must be installed once on a computer.
- Open the IDE RStudio Desktop.
- Access the console pane through View > Show All Panes (Alt+Ctrl+Shift+0) and subsequently select the console pane through View > Panes > Zoom Console (Ctrl+Shift+2).
- Install
  - a single package defined by a `[package name]` by typing `install.packages("[package name]")` at the active command prompt (>) in the console pane;
  - multiple packages defined by `[name package 1]`, `[name package 2]`, `[name package 3]`, etc. with by typing : `install.packages(c('[name package 1]', '[name package 2]', '[name package 3]'))` at the active command prompt (>) in the console pane (note that the this can take some time).
- Wait until the installation is successfully completed:
  - the following messages appear:
    - `package '[package name]' successfully unpacked and MD5 sums checked`
  - the active command prompt appears.
- Check if the package has been correctly installed by typing `library([package name])` at the active command prompt in the console pane. If the package has been correctly installed no error message will be issued.

5.2.1.2 External dependencies

- If R packages depend on non-R-software that needs to be installed on your computer, the `install.package` command returns an error message and indicates, which software should be installed on the computer before trying to install the R package. Install the required software.

5.2.2 Installation of standard R packages

5.2.2.1 Installation of required Standard R packages

- Install the following packages based on the instructions given in section 5.2.1:

- `dplyr` (A Grammar of Data Manipulation. A fast, consistent tool for working with data frame like objects, both in memory and out of memory. More information at cran.r-project.org/web/packages/dplyr).
- `tidyr` (Tidy Messy Data; Tools to help to create tidy data, where each column is a variable, each row is an observation, and each cell contains a single value. 'tidyr' contains tools for changing the shape (pivoting) and hierarchy (nesting and 'unnesting') of a dataset, turning deeply nested lists into rectangular data frames ('rectangling'), and extracting values out of string columns. It also includes tools for working with missing values (both implicit and explicit. More information at cran.r-project.org/web/packages/tidyr).
- `ggplot2` (Create Elegant Data Visualisations Using the Grammar of Graphics. A system for 'declaratively' creating graphics, based on "The Grammar of Graphics". You provide the data, tell 'ggplot2' how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details. More information at cran.r-project.org/web/packages/ggplot2).
- `sf` (Simple Features for R. Support for simple features, a standardized way to encode spatial vector data. Binds to 'GDAL' for reading and writing data, to 'GEOS' for geometrical operations, and to 'PROJ' for projection conversions and datum transformations. Uses by default the 's2' package for spherical geometry operations on ellipsoidal (long/lat) coordinates. More information at cran.r-project.org/web/packages/sf).
- `pander` (An R 'Pandoc' Writer. Contains some functions catching all messages, 'stdout' and other useful information while evaluating R code and other helpers to return user specified text elements (like: header, paragraph, table, image, lists etc.) in 'pandoc' markdown or several type of R objects similarly automatically transformed to markdown format. Also capable of exporting/converting (the resulting) complex 'pandoc' documents to e.g. HTML, 'PDF', 'docx' or 'odt'. This latter reporting feature is supported in brew syntax or with a custom reference class with a smarty caching 'backend'. More information at cran.r-project.org/web/packages/pander).
- `bookdown` (Authoring Books and Technical Documents with R Markdown. Output formats and utilities for authoring books and technical documents with R Markdown. More information at cran.r-project.org/web/packages/bookdown).

5.2.2.2 Installation of frequently used Standard R packages

- Install the following packages based on the instructions given in section 5.2.1:
  - `lme4` (Linear Mixed-Effects Models using 'Eigen' and S4. Fit linear and generalized linear mixed-effects models. The models and their components are represented using S4 classes and methods. The core computational algorithms are implemented using the 'Eigen' C++ library

for numerical linear algebra and 'RcppEigen' "glue". More information at cran.r-project.org/web/packages/lme4).

- `car` (Companion to Applied Regression. Functions to Accompany J. Fox and S. Weisberg, An R Companion to Applied Regression, Third Edition, Sage, 2019. More information at cran.r-project.org/web/packages/car).
- `multcomp` (Simultaneous Inference in General Parametric Models. Simultaneous tests and confidence intervals for general linear hypotheses in parametric models, including linear, generalized linear, linear mixed effects, and survival models. The package includes demos reproducing analyzes presented in the book "Multiple Comparisons Using R" (Bretz, Hothorn, Westfall, 2010, CRC Press). More information at cran.r-project.org/web/packages/multcomp).
- `readr` (Read Rectangular Text Data. The goal of 'readr' is to provide a fast and friendly way to read rectangular data (like 'csv', 'tsv', and 'fwf'). It is designed to flexibly parse many types of data found in the wild, while still cleanly failing when data unexpectedly changes. More information at cran.r-project.org/web/packages/readr).
- `ggrepel` (Automatically Position Non-Overlapping Text Labels with 'ggplot2'. Provides text and label geoms for 'ggplot2' that help to avoid overlapping text labels. Labels repel away from each other and away from the data points. More information at cran.r-project.org/web/packages/ggrepel).
- `ggspatial` (Spatial Data Framework for ggplot2. Spatial data plus the power of the ggplot2 framework means easier mapping when input data are already in the form of spatial objects. More information at cran.r-project.org/web/packages/ggspatial/).
- `vegan` (Community Ecology Package. Ordination methods, diversity analysis and other functions for community and vegetation ecologists. More information at cran.r-project.org/web/packages/vegan).

5.2.2.3 Installation of interactive HTML reports

- Install the following packages based on the instructions given in section 5.2.1:
  - `mapview` (Interactive Viewing of Spatial Data in R. Quickly and conveniently create interactive visualisations of spatial data with or without background maps. Attributes of displayed features are fully queryable via pop-up windows. Additional functionality includes methods to visualise true- and false-color raster images and bounding boxes. More information at cran.r-project.org/web/packages/mapview).
  - `plotly` (Create Interactive Web Graphics via 'plotly.js'. Create interactive web graphics from 'ggplot2' graphs and/or a custom interface to the (MIT-licensed) JavaScript library 'plotly.js'

inspired by the grammar of graphics. More information at cran.r-project.org/web/packages/plotly).

- `echarts4r` (Create Interactive Graphs with 'Echarts JavaScript' Version 5. Easily create interactive charts by leveraging the 'Echarts Javascript' library which includes 36 chart types, themes, 'Shiny' proxies and animations. More information at cran.r-project.org/web/packages/echarts4r).

5.2.3 Installation of non-standard R packages

5.2.3.1 Complex Heatmap

- If required, install the package `ComplexHeatmap` based on the following instructions:
  - Type `install.packages("devtools")` at the active command prompt (>) in the console pane;
  - Wait until the installation is successfully completed;
  - Type `library(devtools)` at the active command prompt (>) in the console pane;
  - Type `install_github("jokergoo/ComplexHeatmap")` at the active command prompt (>) in the console pane;
  - Wait until the installation is successfully completed;

5.2.3.2 LaTeX support

- If reports should be created in PDF-format, a LaTeX distribution must be installed. It is recommended to install the `TinyTeX` package for this purpose (More information at yihui.org/tinytex):
  - Type `install.packages('tinytex')` at the active command prompt (>) in the console pane;
  - Wait until the installation is successfully completed;
  - Type `tinytex::install_tinytex()` at the active command prompt (>) in the console pane;
  - Wait until the installation is successfully completed;

6 References

There are no sources in the current document.

7 Annexes

There are no annexes in the current document.