

IMPERIAL

Exposomics Analytics - Pre-session Material

Contents

Getting started	2
Software Installation	2
Installing and Loading Packages	3
Check package informaton (version)	4

Getting started

Before we get started please make sure your R software is up to date.

1. Open a web browser and go to the R project website.
2. Download the latest version of R for macOS, Linux or Windows.
3. Open the downloaded .pkg file and follow the installation instructions.

Once you have installed the latest version of R, you can verify the installation by opening a terminal (or command prompt) and typing `R --version`. This command should display the version number of the installed R software

Software Installation

R is a programming language widely used for statistical analyses (see <https://www.r-project.org/about.html>).

R for Mac OS X can be downloaded from the following link: <https://cran.r-project.org/bin/macosx/> The link will take you to the download page on the CRAN. Please select the latest version by clicking on R-4.4.1.pkg, if this version can be supported on your machine. Otherwise, choose the latest supported version. If you have a laptop with the Apple chip, you can select the version, which has been specifically designed for that chip.

As indicated on the download page, you will also need to install XQuartz to be able to use some more advanced R functionalities. This can be done from this link: <https://www.xquartz.org>

R for Windows can be downloaded from the following link: <https://cran.r-project.org/bin/windows/base/> The link will take you to the download page on the CRAN. Please select the latest version by clicking on “Download R 4.4.1 for Windows”.

R Studio can be downloaded from <https://www.rstudio.com/products/rstudio/download/> This link will take you to the webpage. You will need to click on “Download”, which will take you to the version page. From there, you can select the installation file corresponding to your system. Once the file is downloaded, please follow the instructions for installation.

Once both R and R Studio are installed, you can open R Studio. You can then install the required packages (see next section).

Installing and Loading Packages

Installing packages in R is the process of downloading and setting up additional functionality and tools that extend the capabilities of the base R language. These packages contain functions, datasets, and other resources that users can utilize in their R scripts and analyses. In some cases, you may not be able to update the R software yourself, e.g. when using a remote server. In the case your R version does not support installation of the latest package version, e.g. Matrix version 1.7.0 required for lme4, you need to manually download the archived version of the packages that are compatible with your R version (e.g. $\geq 3.5.0$ and $< 4.4.0$) and run the following code to install the tar.gz files:

```
install.packages("~/Downloads/package_name.tar.gz", repos = NULL, type = "source")
```

Users can install packages from CRAN (the Comprehensive R Archive Network), GitHub, or other repositories using the `install.packages()` function in R.

Once installed, packages can be loaded into the R environment using the `library()` function, allowing users to access and utilize the functions and resources provided by the packages.

For convenience, we use a custom function that automatically loads all required packages from CRAN or Bioconductor and installs any missing ones. Bioconductor is an open source software repository for biological data analysis. Packages from Bioconductor are installed via the BiocManager using the `BiocManager::install()` function in R.

During the installation process, you may be prompted to update packages. If this occurs, type 'a' in the console to update all packages.

```
package_names = c("BiocManager", "lme4", "RColorBrewer", "pheatmap",
                  "tidyr", "dplyr", "stringr",
                  "glmnet", "ROCR", "ggplot2", "ggfortify",
                  "survival", "mixOmics", "utils", "corpcor",
                  "igraph", "glasso", "abind", "fake", "sharp",
                  "remotes", "OpenMx", "openxlsx", "FactoMineR",
                  "plotrix", "colorspace", "scales", "sgPLS",
                  "pcaMethods", "imputeLCMD", "impute", "devtools", "fpc",
                  "cluster", "mclust", "factoextra", "randomcoloR", "aricode", "mltools", "data")

require_packages <- function(...) {

  libs <- unlist(list(...))
```

```

# Load required libraries, see if any packages are missing
req <- unlist(lapply(libs, require, character.only = TRUE))
need <- libs[!req]
if (length(need) > 0) {
  for (pkg in need) {

    # First, try installing from CRAN
    tryCatch({
      utils::install.packages(pkg, dependencies = TRUE)
    }, warning = function(w) {

      # Make sure BiocManager is installed
      if (!requireNamespace("BiocManager", quietly = TRUE)) {
        install.packages("BiocManager")
        library(BiocManager)
      }

      # try Bioconductor
      message(paste("Attempting to install", pkg, "from Bioconductor..."))
      BiocManager::install(pkg)
    })

    # Load the package after installing
    require(pkg, character.only = TRUE)
  }
}

require_packages(package_names)

```

Check package informaton (version)

The following code efficiently retrieves and displays the version numbers of specified R packages, allowing users to quickly check the versions of the packages they are using in their R environment:

```

# Display package versions

for (pkg in package_names) {
  pkg_info <- installed.packages()[pkg, ]
}

```

```
cat(pkg, "version:", pkg_info[["Version"]], "\n")
}
```

```
## BiocManager version: 1.30.25
## lme4 version: 1.1-35.5
## RColorBrewer version: 1.1-3
## pheatmap version: 1.0.12
## tidyr version: 1.3.1
## dplyr version: 1.1.4
## stringr version: 1.5.1
## glmnet version: 4.1-8
## ROCR version: 1.0-11
## ggplot2 version: 3.5.1
## ggfortify version: 0.4.17
## survival version: 3.7-0
## mixOmics version: 6.28.0
## utils version: 4.4.1
## corpcor version: 1.6.10
## igraph version: 2.0.3
## glasso version: 1.11
## abind version: 1.4-8
## fake version: 1.4.0
## sharp version: 1.4.6
## remotes version: 2.5.0
## OpenMx version: 2.21.12
## openxlsx version: 4.2.7.1
## FactoMineR version: 2.11
## plotrix version: 3.8-4
## colorspace version: 2.1-1
## scales version: 1.3.0
## sgPLS version: 1.8
## pcaMethods version: 1.96.0
## imputeLCMD version: 2.1
## impute version: 1.78.0
## devtools version: 2.4.5
## fpc version: 2.2-13
## cluster version: 2.1.6
## mclust version: 6.1.1
## factoextra version: 1.0.7
```

```
## randomcoloR version: 1.1.0.1
## aricode version: 1.0.3
## mltools version: 0.3.5
## data.table version: 1.16.0
```