

Doing your own analysis on single cell data.

Author: Teresa Attenborough (University of Glasgow)

First, decide where on your computer you would like to do your analysis, and save your plots, files etc. Location is up to you. I have a folder called R_projects in my home folder (ta13), and within that I have a folder for each R project.

Working in R

Next, install R and RStudio (if you don't already have them installed).

Windows install

Install R

To install R, you need to download the installer from the R website (<https://cran.r-project.org/>)

Click on either base or install R for the first time

Click on the download R for Windows link

Once downloaded, open the .exe file and follow the installation instructions on your computer

Install RStudio

To install RStudio we download it from the Posit website (<https://posit.co/download/rstudio-desktop/>)

Click on the Download RStudio Desktop link

Once downloaded, open the .exe file and follow the installation instructions on your computer

If you're having trouble installing R and/or RStudio, you can find guides online such as:

<https://www.dataquest.io/blog/installing-r-on-your-computer/>

Mac install

Install R

To install R on your Mac you need to know the type of processor your Mac uses. This is straightforward to find out:

On the top navigation bar on your Mac, click on the apple icon. From the drop down menu, select About This Mac. In Overview you will find the information about your Mac. If you have an Intel Mac, you will see the processor row, which has information that includes Intel. If you have an M1 or M2 Mac, you will see chip and M1/M2 in the Overview with something like Chip Apple M1.

M1 or M2 Mac

To install R, you need to download the installer from the R website (<https://cran.r-project.org/>)

If you have a M1 Mac you will need to click on the link that contains arm64 to download R. It will look something like R-4.3.1-arm64.pkg

Once downloaded, open the .pkg file and follow the installation instructions

Intel Mac

To install R, you need to download the installer from the R website (<https://cran.r-project.org/>)

If you have a Intel Mac you will need to click on the link that just contains the version of R. It will look something like R-4.3.1.pkg, and can be located a touch further down the page under the header Binaries for legacy macOS/OS X systems:

Once downloaded, open the .pkg file and follow the installation instructions

Install XQuartz

To run R on a Mac operating system, XQuartz is required. You can install it by following this link (<https://www.xquartz.org/>), downloading it and following the installation instructions.

Install RStudio

To install RStudio we download it from the Posit website (<https://posit.co/download/rstudio-desktop/>)

Click on the Download RStudio Desktop link

Once downloaded, open the .dmg file and follow the installation instructions on your computer

Help

If your installation for R and RStudio did not work, this is likely because your computer is running an older operating system. In these cases you will have to install an older version of the software.

Getting started - R projects

Now, open RStudio. I suggest you start a new **R project** to do this tutorial.

File -> New Project

Select the option 'New Directory' and then 'New Project'

You can choose the name of the new project, which will be the name of the folder you will be working in. For this, you might call it "singlecell_tutorial" or something similar. You can also set where this folder will be. For me, I will set it as ta13/R_projects, so it will be created in my folder of R projects. You can use the browse option to find the right folder for you.

More details on projects

Setting up a project in a new directory:

1) Select new project 2) Give your directory a name, e.g. "r-workshop" 3) Select browse and find the directory you want to start your project 4) Press Create Project 5) Your new project should open

Setting up a project in a existing directory:

1) Select browse and find the directory you want to start your project 2) Press Create Project 3) Your new project should open

To open and review your projects in future you will see them in a drop down menu in the top right corner of RStudio.

Data

Next, you need the data. This will be available from a google drive link. Download the sample1, sample2, and sample3 folders, .rds file (data) and .Rmd (script) file associated with this project, and place them in the R project folder you've made.

Get started

You can then open the Rmd file in RStudio and get started. In RStudio, select File -> Open File, then navigate to your chosen file. For this tutorial, the Rmd file is called `practice_analysis_2day_somules.Rmd`.

Introduction to packages

Packages are collections of functions, code, and sample data put together by the R community. Packages are one of the main benefits of R. As R is open source there can be lots of contributors who have made functions to do complex tasks, such as data cleaning, specific types of data analysis, or exciting data visualisation.

To install these packages onto your computer you have to download them from CRAN (The Comprehensive R Archive Network).

There are two ways of doing this, using code (recommended and easiest) or using the menus (`Tools > Install Packages`).

Using code involves using the `install.packages` function, which looks like:

`install.packages("package name")` . To install the package you would type something like: `install.packages("readr")` .

Once installed, you will not need to do this again unless you install a newer version of R.

Loading packages

Now you have installed the packages, you need to load them in order to use them. **Each time you load R you will need to re-load the packages you want to use.**

To load a package you need to use the `library()` function. For example, if I wanted to load the `readr` package I would type `library(readr)` .

We will need to install and load the following packages: 'tidyverse', 'Seurat', 'RColorBrewer', 'patchwork', 'clustree', 'biomaRt'. You can install these in advance (if possible), or we will install them at the beginning of the tutorial.

Potential problems/fixes

Some package dependencies may be an issue to install (e.g. RSpectra, uwot, and leiden). You can try installing them separately, and increasing time-out time to 300.

```
options(timeout=300)
```

Credit to github.com/andrewmoles2 for help with set-up guide.