

Creating an R Data package

Anita Dhillon

2023-05-02

Creating a new repository locally as an R package

With RStudio: 1. Click option to create a new project 2. Create a New directory 3. Choose create an R package 4. Give a name to the package 5. Click on **create in new session**

Once we have a backbone for an R package, start customizing it - Modifying the DESCRIPTION file - This file contains the metadata of the package Metadata is “data about data”, it contains the information about a data set.

- License GPL (≥ 2) is the most common for data packages
- Say something about the particularities of a GPL license
- Our package will only depend on R itself for now, that is no other packages
 - It can be the latest version of R or any version that we are using

Add a README file

Use a new package that is called `usethis`

```
install.packages("usethis")
```

```
library(usethis)
```

The function `use_readme_md()` create a template for a README file in markdown

```
use_readme_md()
```

```
## v Setting active project to '/Users/anitadhillon/HelloR'
## * Update 'README.md' to include installation instructions.
```

Adding data to the package

We are going to use the function `'use_data_raw()'`, this creates a folder called `'data-raw/'` with an R script that will document the creation of the data object that will be stored in our package.

In this R script called `'DATASET.R'` we will document the steps needed to read the raw data, which is in the csv format, and save it in R format. This is useful so that the data set is always available when someone installs in R package.

Code to prepare `'taxonomy'` dataset goes here

```
usethis::use_data(DATASET, overwrite)
```