Useful R Packages

Heide Jackson
University of Maryland Population Research Center
heidej@umd.edu

September 2019

Contents

Installing and Using Packages 2
Version Control 2
Useful R Packages 3
- General Use Packages 3
- Version Control Packages 3
- Multiple Imputation Packages 3
- Survey Packages 3
- Statistical Methods Packages 4
- Data Visualization Packages 4

Disclaimers

R and R packages are user developed and user maintained. Some R packages may not be compatible with others and some R functions may not work on all versions of R.

Installing and Using Packages

Using packages within an R session requires two steps:

1. Installation
2. Loading

Below is a quick example of R code to install and load a package.

```
install.packages("Hmisc") #needed once
library("Hmisc") # needed every time you want to use the package in a session
```

Typically, installation will only need to be done once, but the library function will be needed every time you want to use a package's functions within your R session.

Version Control

R and R packages are dynamic. Sometimes if you want code to remain reproducible, it may be of interest to use an R package that maintains version control.

Useful R Packages

General Use Packages

- Tidyverse a collection of packages for cleaning, manipulating, and visualizing data. The grammar of functions in the tidyverse is very similar to Pandas in python.

  As mentioned, Tidyverse is a collection of packages. Installing Tidyverse will download all of these packages, but packages are best loaded individually. Below is a list of packages in the tidyverse I find particularly helpful:

  - haven Reads in and writes data from other statistical packages. My experience is this works better for Stata than SAS.
  - readr Reads in and writes rectangular data sets. Typically, these are csv or text files.
  - dplyr Manipulates data in various ways. Contains functions to keep, drop, reorganize, and modify data.
  - ggplot2 THE package for data visualization in R.

- stargazer Provides the ability to generate pretty tables from R output.
- Hmisc An assortment of miscellaneous packages that assist in data manipulation and analysis.

Version Control Packages

- Useful information about R version control and different options on how to go about it.
- packrat Version control package that I personally like.
- docker Another version control option that comes recommended.

Multiple Imputation Packages

- mice Offers different options for imputing data in R.
- mitools Offers tools for working with multiply imputed data in R.

Survey Packages

- survey Provides functions for working with survey data and incorporating survey weights into analysis.

Statistical Methods Packages

- nnet Functions for running multinomial logistic regression in R.
- survival Provides functions for fitting and evaluating single state survival models (event history analysis) in R.
- multistate Provides functions for fitting and evaluating multistate models in R.
- boot Bootstraps data in R.

Data Visualization Packages

- ggplot2 THE package for data visualization in R. A lot of people like and use this package. I like it mainly for difficult or complex graphics.
- lattice Offers function for making multi-panel graphs.
- RColorBrewer Allows for custom color creation in R.