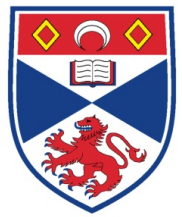
A primer on R
for the
Intermediate Distance Sampling workshop

CREEM
Centre for Research into Ecological and Environmental Modelling



University of St Andrews

Learning goals

- Obtain a working knowledge of the R language and environment, to be able to implement analysis for the Intermediate Distance Sampling workshop

- Learn how to access R via the RStudio software

The learning session will be almost entirely hands-on.
Thanks to Tiago Marques and Danielle Harris for tutorial material.

About R

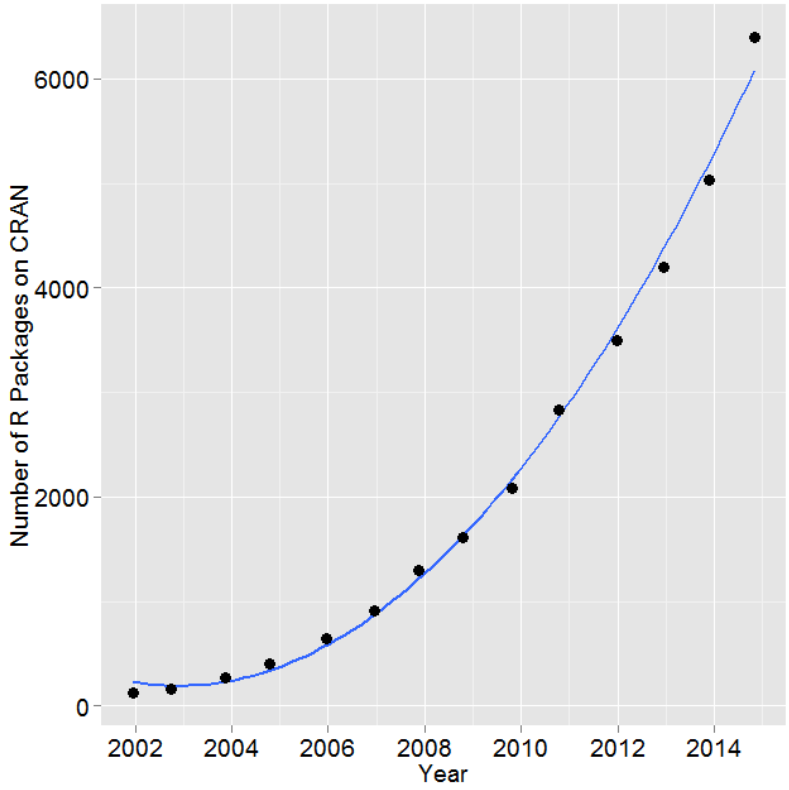
- R is both a programming language and an environment for statistical computation
- R is free open source software
- Created in 1995 by Ross Ihaka and Robert Gentleman, (University of Auckland, Auckland)
- Based on previously established language, S
- Now managed by the “R Core Development Team”
- Software and resources available at

http://www.r-project.org

Latest version was **R version 3.4.1**

The R base + packages

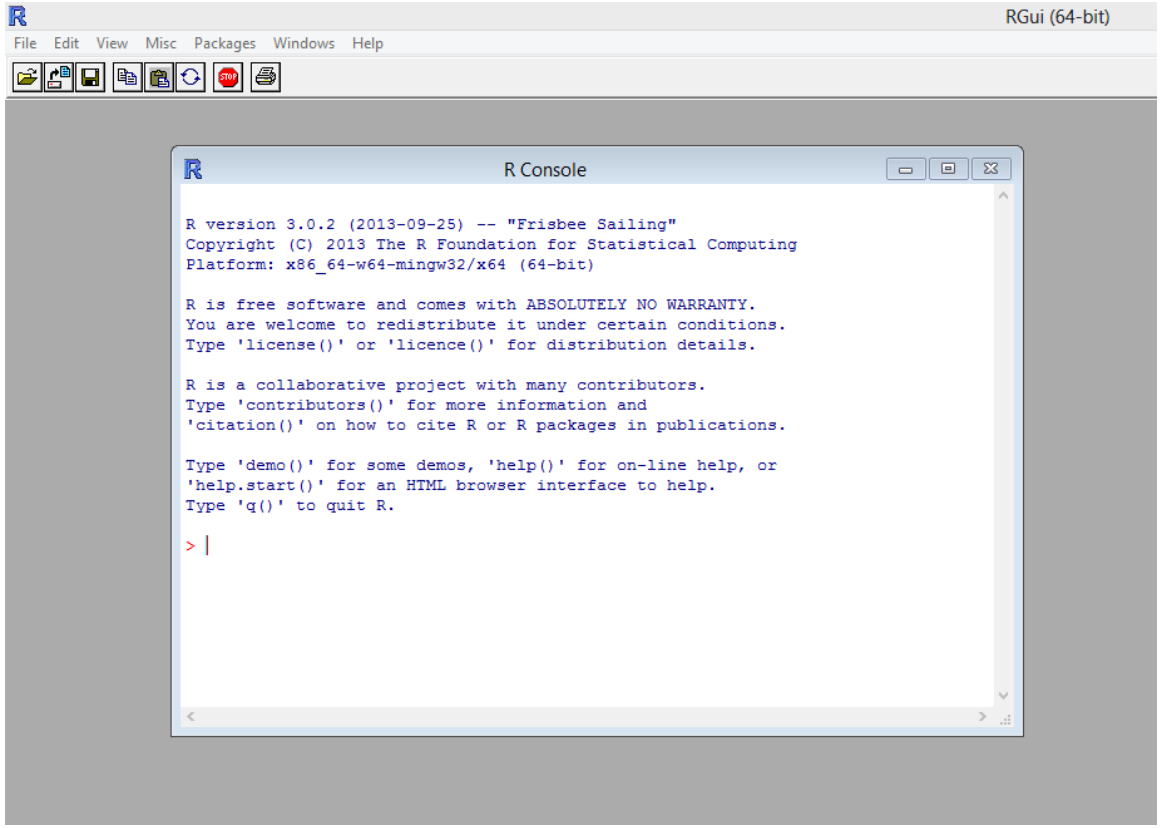
- R base installation comes with some packages that provide access to a large number of commonly used features (e.g. fancy plotting, survival analysis, nonlinear modelling, etc)
- Many more packages are available online, contributed by users



Source: http://r4stats.com/articles/popularity/

The R environment

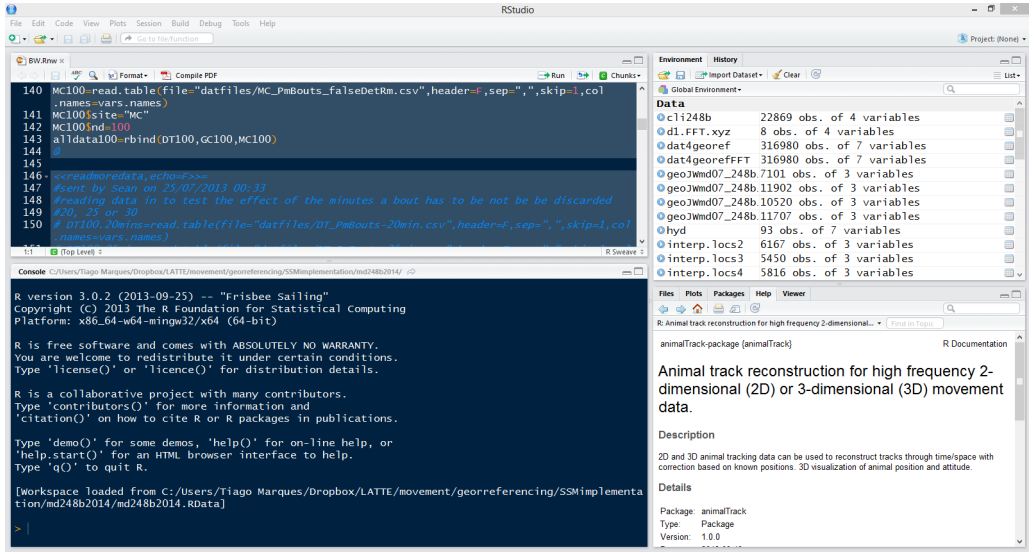
- R comes with a simple interface, allowing you to
  - type/paste commands
  - view plots



- ... but RStudio is much easier to use

RStudio

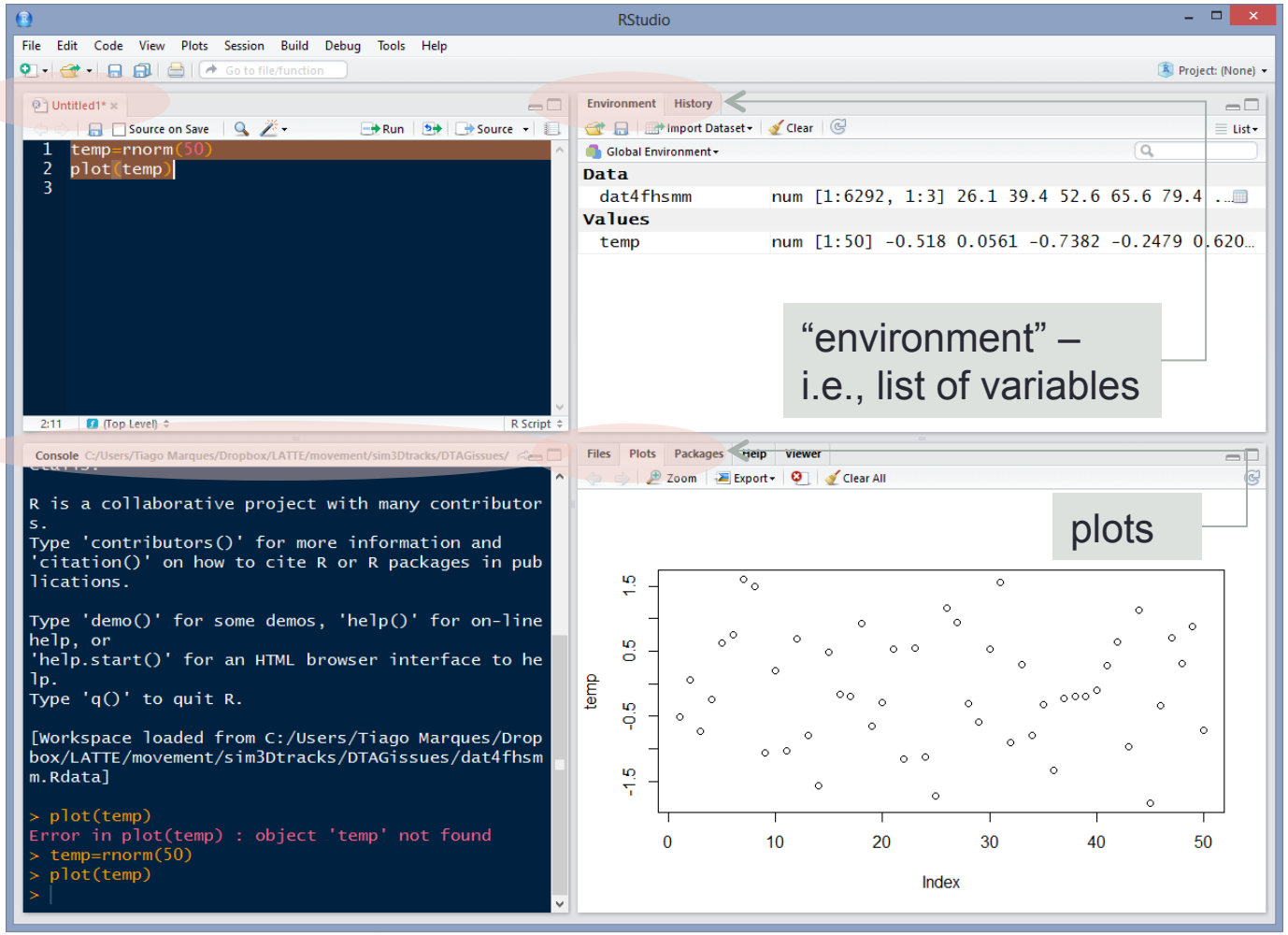
- An feature-rich “integrated development environment” for data analysis
- Simplifies the user interaction with R
- In a single application you have access to a number of windows containing your code files, an execution console, your plots, help, a file explorer, etc.
- (as R) Free and open source



4 windows; 4 ways to interact with R

code file (“script”)

active console

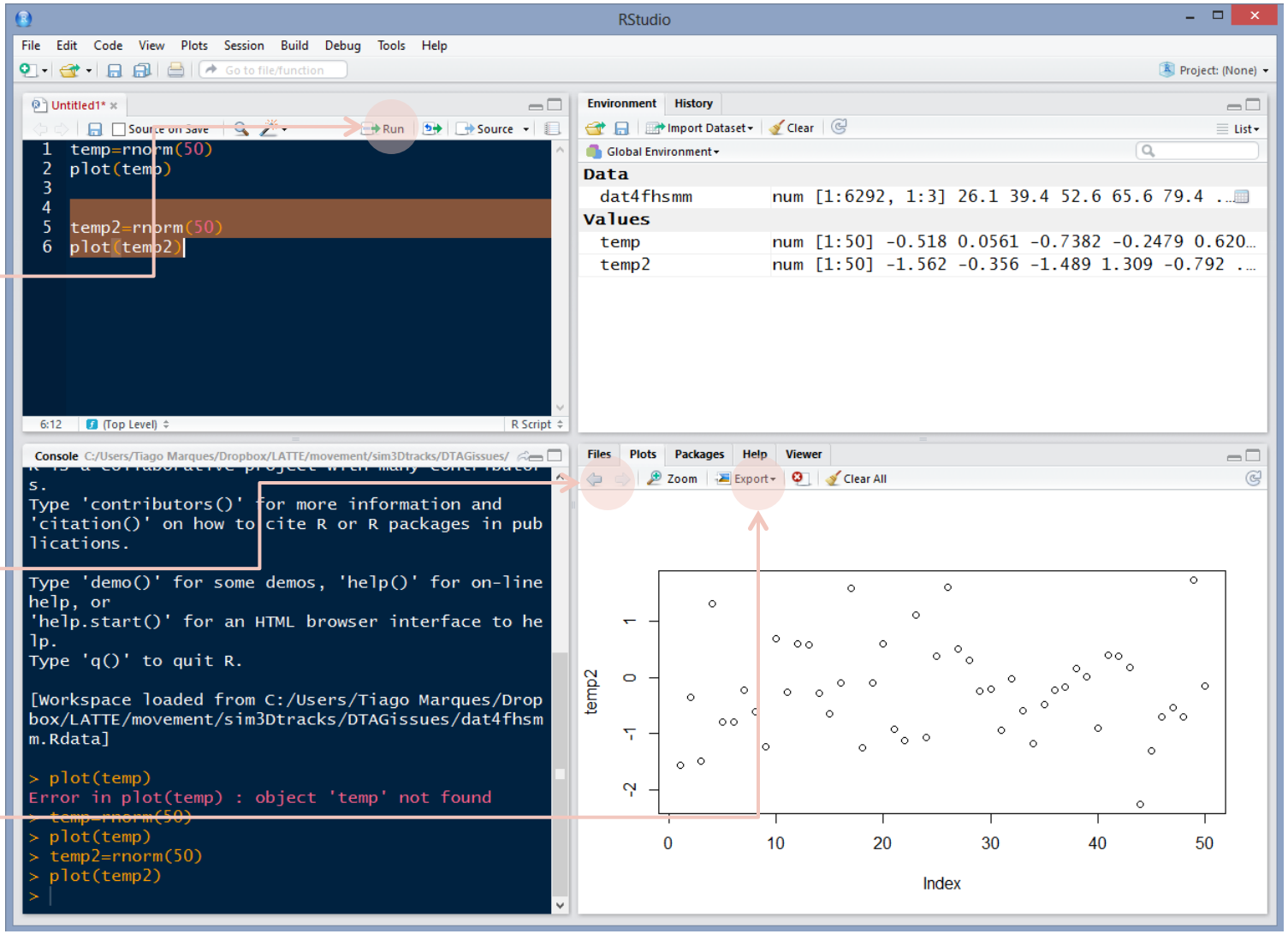


The code and plots tabs

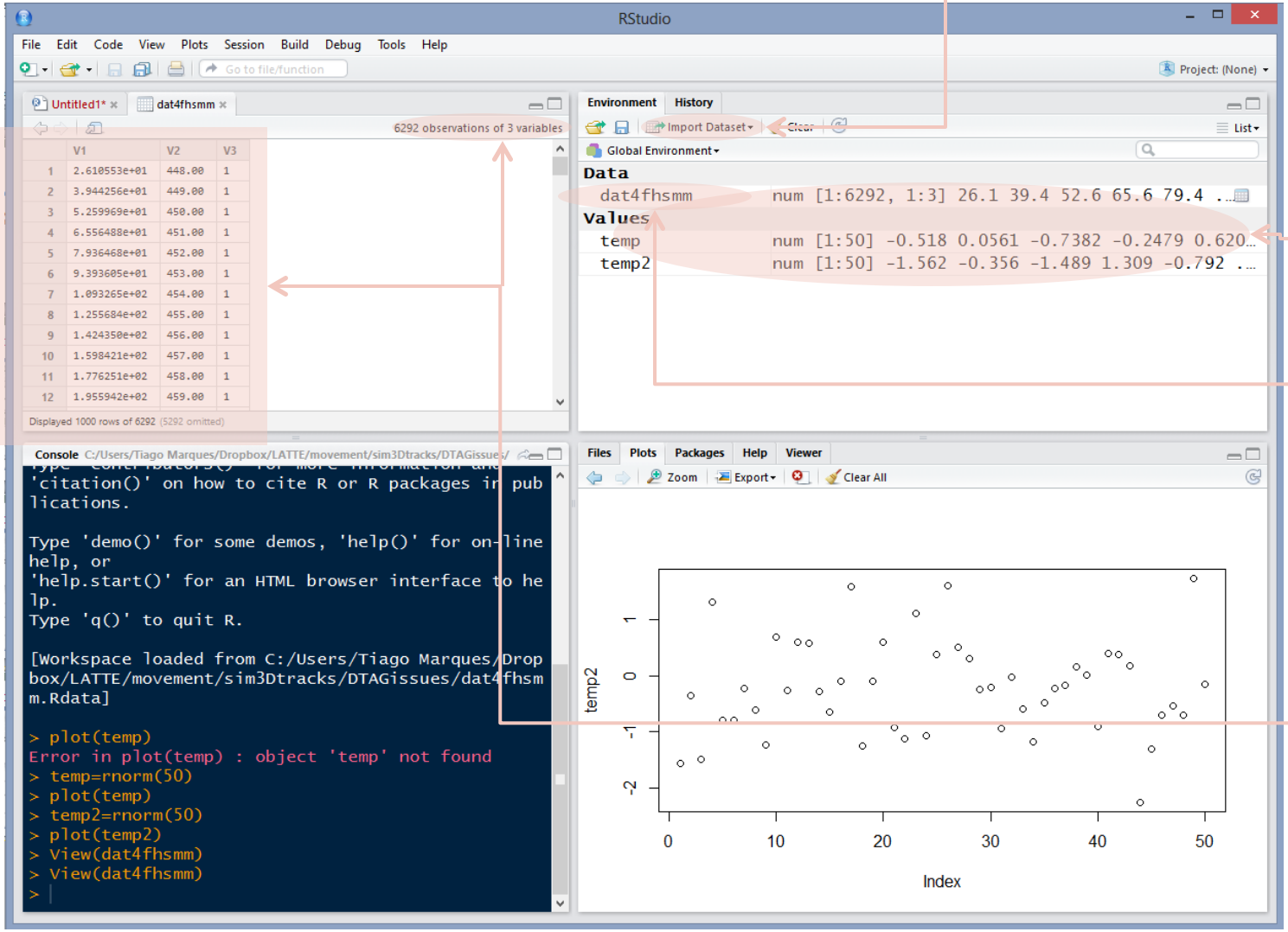
Sends and executes current line or selected text to console

Move back and forth in a figure log

Export figures as pdf or jpeg



The environment tab

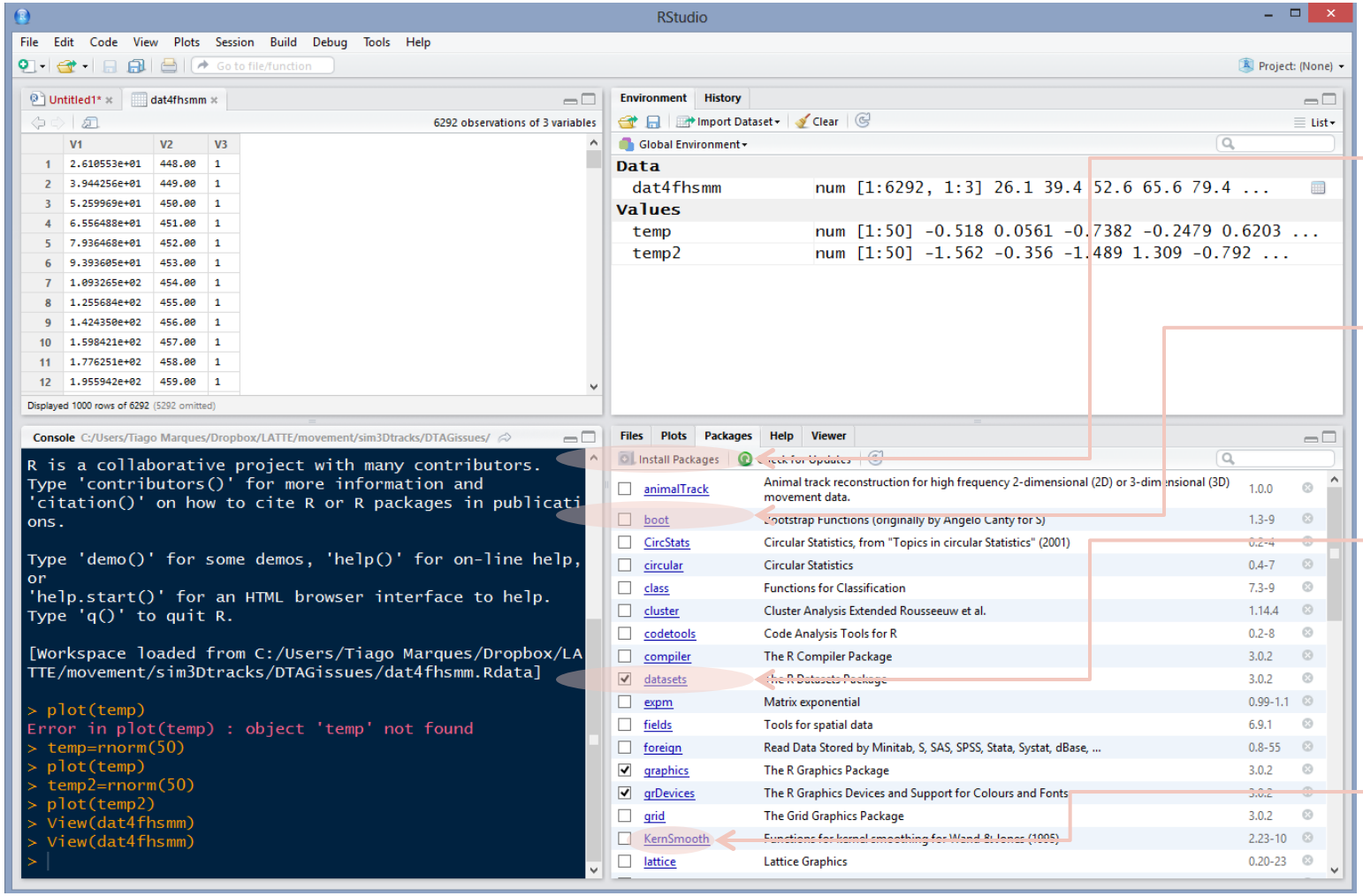


Shortcut for data import

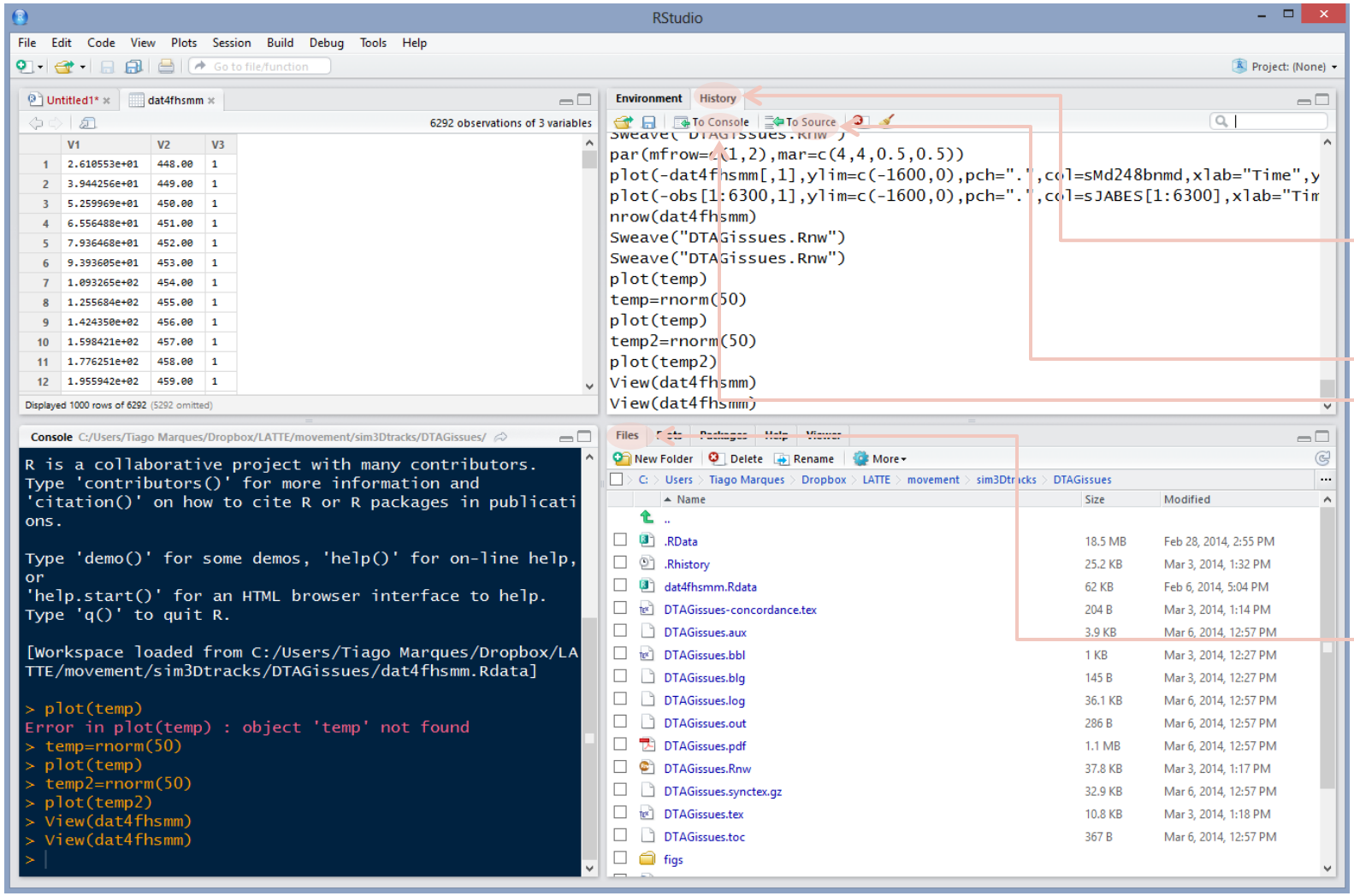
Objects available in workspace (+ details)

Pressing most objects names gives you a preview of what they are

The packages tab



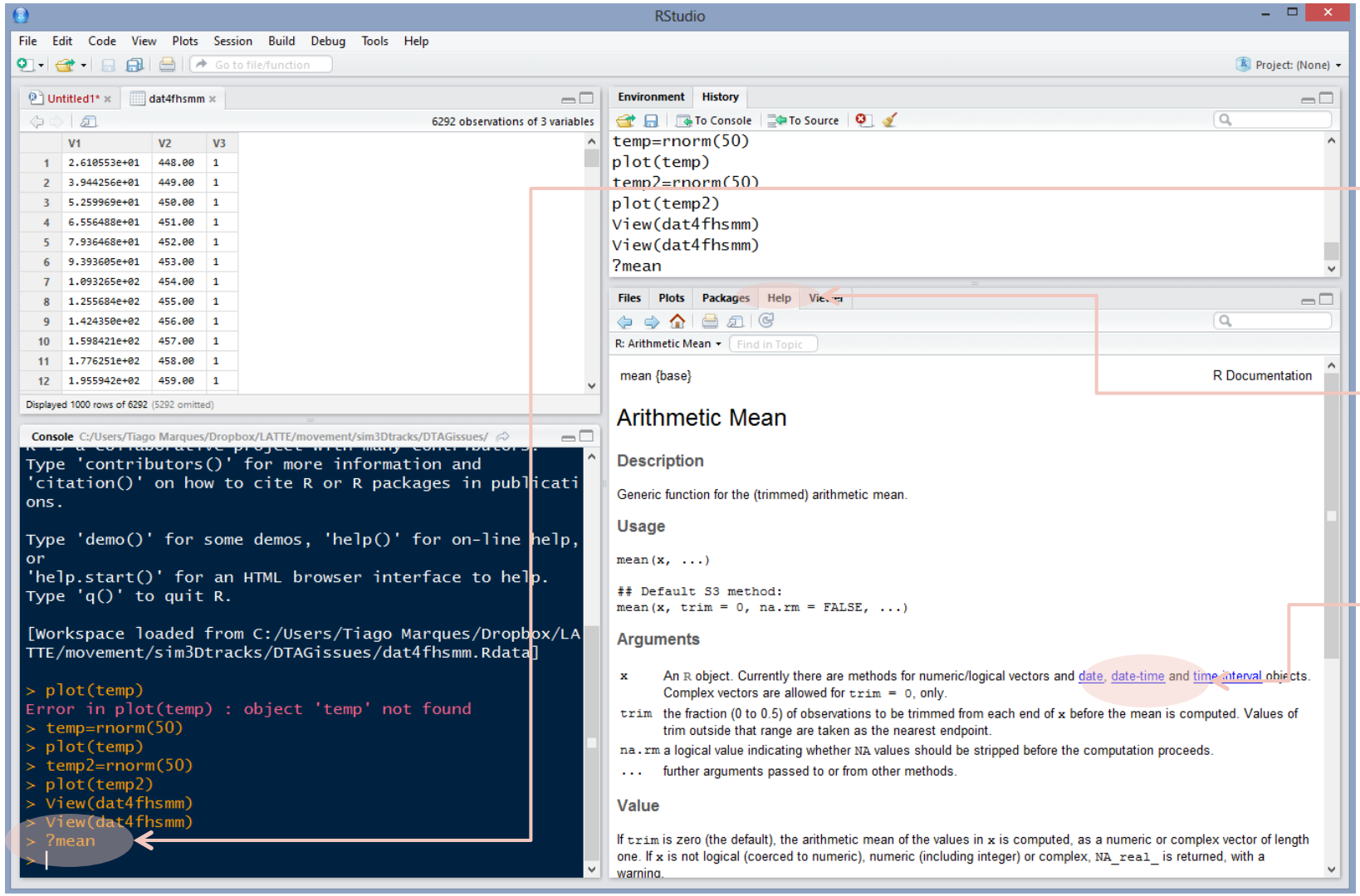
The history and files tabs



The command history (with shortcuts to send commands to source or console)

To see the files in your working directory (select a file to load or open it)

The help tab



If you call for help...

This is where it will show up

With hyperlinks for easy navigation

Now... open the hands on tutorial

A hands on introduction to R tutorial

> prepared by Tiago A. Marques, Danielle Harris & Len Thomas

Introduction

This tutorial was created as a gentle introduction to the R environment. It does not assume any basic knowledge about R, but some basic programming notions would be desirable.

There is an extensive community revolving around R, and abundant courses, tutorials, books, blogs, list servers, etc, freely available online. We provide here a small list of some of these:

- R webpage - the main R webpage, including links to downloading R, manuals, tutorials, dedicated search engines, etc.
- R video tutorials - video how to's in R
- Online tutorial - a course with interactive exercises
- Online course - DataCamp commercial site
- Reference card - A very handy list of useful R functions
- Short reference card - A longer reference card with most commonly used R functions

To facilitate the interaction with R we leverage on RStudio, a piece of software which allows users to have at a click's distance many useful features in R. In the following sections of the tutorial you will be guided through a first session of R via RStudio.

The tutorial is intended to follow a brief presentation about R and RStudio, their interaction and capabilities. It assumes that R and RStudio have been previously installed in the computer you are using. The latest version of both software packages is recommended. Both are free and open source.

Introduction to RStudio

Working with Rmarkdown files

- Having downloaded and opened *01-Rtutorial.Rmd*
  - You will see a mixture of readable text along with R code
  - The bits of R code are called “chunks” and those chunks usually have labels
  - We are likely to refer to these chunk names when describing computation steps
  - Lines 62-66 is a chunk named ‘calc’
  - Green box indicates icons that can be clicked to send code within chunk to console for processing
  - Orange ellipse indicates icon that can be used to render entire document