

Data Visualization with R

Instructor: Arvind Iyer



www.arvindkiyer.com



[@Arvind K Iyer](#)



ayalurarvind@gmail.com

Learning Objectives

- **Able to do data formatting and cleaning.**
- **Learn basics of data visualization with ggplot2.**
 - Introduction to grammar of graphics
 - Univariate Graphs (bar plot, pie chart etc)
 - Bivariate Graphs (box plot, violin plot etc)
 - Multivariate Graphs (scatter plot etc)
 - Other Graphs (heatmap etc)
 - Using of Themes
- **A hands on workshop with a project at the end. (~70-80% doing)**

Pre-requestives

- Knowing the basic of R programming
- Have R and Rstudio installed (Rstudio is an IDE, any other IDE will also work)

Installation of R and Rstudio

https://cran.rstudio.com

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux \(Debian, Fedora/Redhat, Ubuntu\)](#)
- [Download R for macOS](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

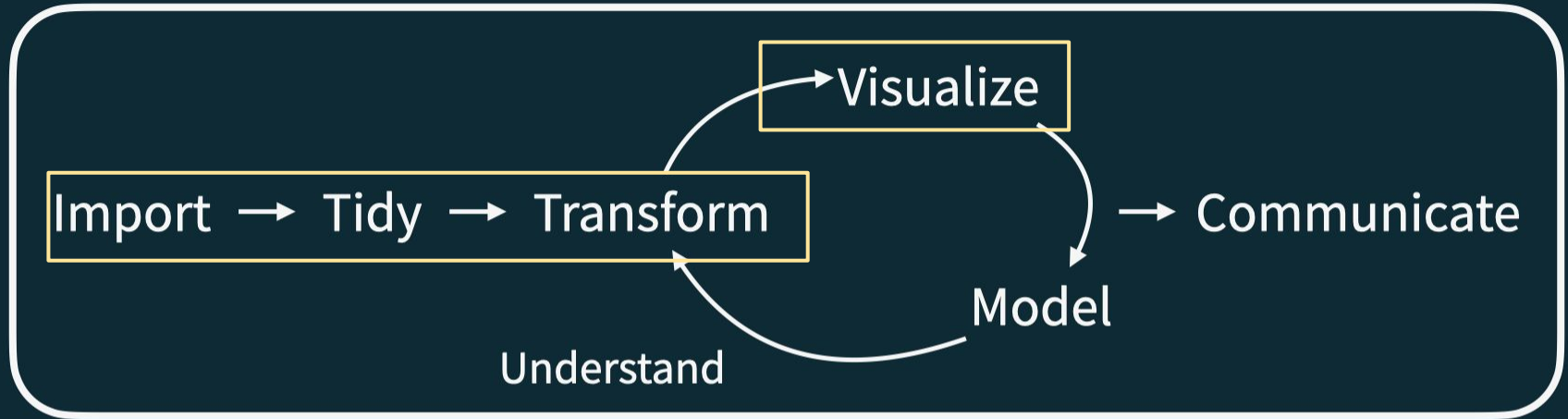
- The latest release (2024-06-14, Race for Your Life) [R-4.4.1.tar.gz](#), read [what's new](#) in the latest version.
- The CRAN directory [src/base-prerelease](#) contains R alpha, beta, and rc releases as daily snapshots in time periods before a planned release.
- Between releases, the same directory [src/base-prerelease](#) contains snapshots of current patched and development versions. Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Alternatively, daily snapshots are [available here](#).
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#).

<https://cran.rstudio.com>

OS	Download	Size	SHA-256
Windows 10/11	RSTUDIO-2024.09.0-375.EXE ±	265.55 MB	513216FE
macOS 12+	RSTUDIO-2024.09.0-375.DMG ±	621.00 MB	54D722FD
Ubuntu 20/Debian 11	RSTUDIO-2024.09.0-375-AMD64.DEB ±	203.93 MB	DB096050
Ubuntu 22/Debian 12	RSTUDIO-2024.09.0-375-AMD64.DEB ±	203.92 MB	111C64DB
Ubuntu 24	RSTUDIO-2024.09.0-375-AMD64.DEB ±	203.92 MB	111C64DB

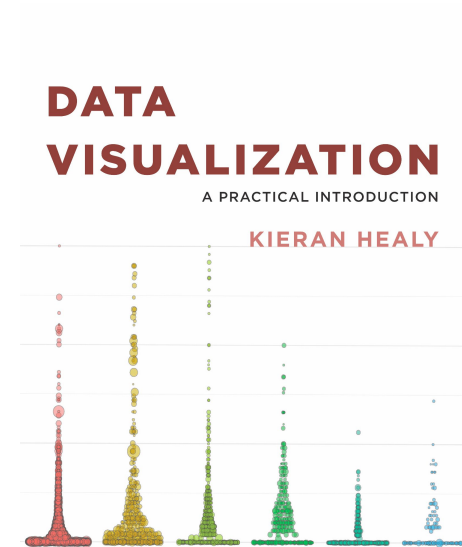
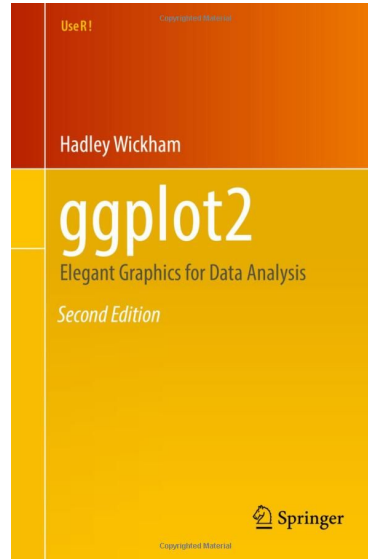
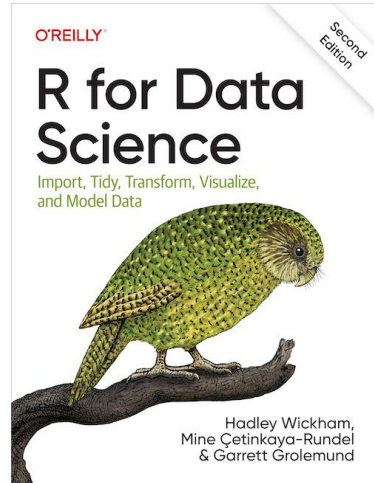
<https://posit.co/download/rstudio-desktop/>

Data science life cycle



Reading Materials

- Content of this tutorial is taken from multiple sources (Reference will be mentioned)
- Major reading materials and books (what I have used)



Things to Do

Try out in Rstudio these data structures. Be familiar with it. Complete week-0 and week-1

<https://bioinformatics-core-shared-training.github.io/Bitesize-R/>

More R tutorial over here.

<http://venus.ifca.unican.es/Rintro/dataStruct.html>