Introduction to R

What is R?

An environment that combines:

- effective handling of big data

- collection of integrated tools

- graphical facilities

- simple and effective programming language

Visualization   Open source

Data science   Platform agnostic

Computational
statistics

R packages

- collections of R functions, data, and compiled code in a well-defined format

- base packages come with R installation

- additional packages for specific functionality

- 7,000+ user contributed packages and growing

CRAN (Comprehensive R Archive Network)

- a repository for packages; source available for download or install directly within R

- An alternative package repository; "..provides tools for the analysis and comprehension of *high-throughput genomic data*."

- Includes (but is not limited to) tools for:

  ✦ performing statistical analysis

  ✦ accessing public datasets

- Open source and open development

- Free

www.bioconductor.org

Integrated Development Environment (IDE)

- Graphical user interface, not just a command prompt

- Great learning tool
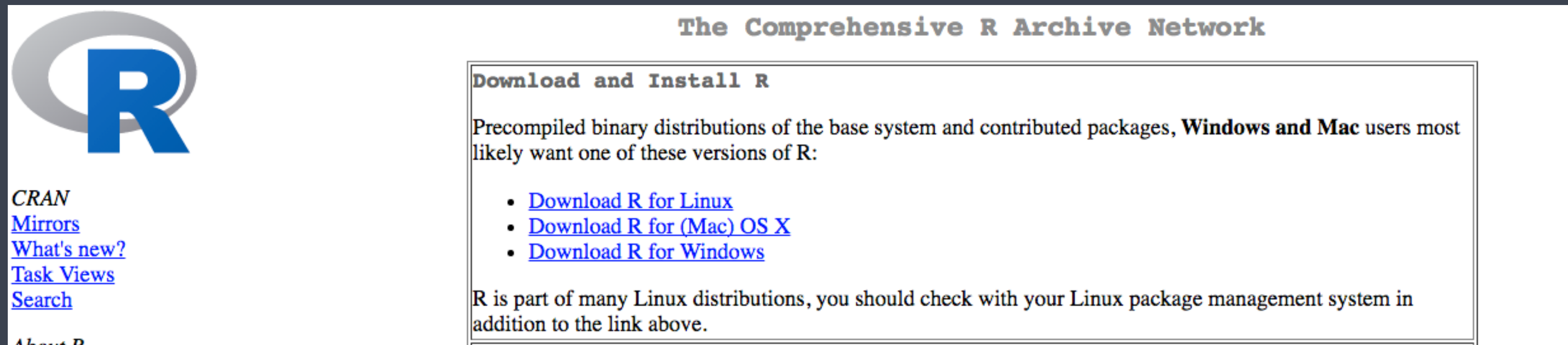
- Free for academic use

- Program agnostic

- Open source

www.rstudio.com

Learning objectives

- How to interact with R using RStudio

- Become familiar with R syntax

- Understand data structures in R

- Inspect and manipulate data structures

- Install packages and use functions in R

- Visualize data using simple and complex plotting methods

Installations

- R: https://cran.r-project.org/

Installations

- RStudio: https://www.rstudio.com/products/rstudio/download/