Introduction: Why use R?

Hannah Waddel

1/26/2022

What is R?

- Programming language specifically developed for *statistical computing*
- It is good at statistical tasks:
  - Manipulating data
  - Statistical tests
  - Drawing graphs
  - Fitting statistical models
  - etc.

Why use R?

- Popular
- Flexible
- FREE!

R is popular

- New statistical research is often done using R
- Out of 50 newest statistics papers on arXiv (preprint server)
  - 30 identified programming languages used
  - 14 used R
  - 8 used Python
  - 0 used SAS
  - 8 used another (Matlab, Julia, etc.)
- If you want to use a new statistical method, there's a good chance it's implemented in R

R is flexible

- Add packages to increase functionality
  - New statistical methods
  - Graphs and presentations
  - Reports
- R updates often
- R is available for Windows, Mac, and Linux
- You can use other programming languages (for example, Python), within R code
- I made these slides in R

R is free

- Open source software
- Anyone can contribute
  - Still managed by R core team

Caveats to Using R

- R is free
- R is less “verbose”

R is free

- SAS has a dedicated help team and statisticians
  - Quality control
  - Maintenance
  - Consistency
  - Help
- R is managed by the core team
  - R repository (CRAN) is generally high quality
  - R updates may change things, or (rarely) break code
  - Volunteers may stop maintaining packages
  - Many forums, websites, etc. for R help

SAS is verbose

- SAS gives a lot of output
- R gives less output
  - You must know what you want

What is RStudio?

- Integrated development environment (IDE) for R
- R is the programming language, RStudio gives you an easy way to use R
  - You don't need RStudio to use R
- RStudio is free
  - RStudio company creates RStudio and makes money by selling business licenses

Next Videos

- Install R and RStudio (Windows or Mac)
- Bios 500 Analyses
  - Two-sample T-test
  - χ^2 test