



The R Language

CS102

Winter 2019

R

Programming language and software environment for statistical computing, graphics, and (more recently) data analysis

- Dynamically typed and interpreted, like Python
- A primary data type is “data frame”, similar to relational table (and Python pandas dataframe)
- Many R features now also available in Python packages

R Versus Python

R

- Easier for experienced programmers
- Used by academics, researchers, hard-core data scientists
- Shorter code for complex analyses, statistics, graphics
- Extremely slow!

Python

- Good for beginners or experienced
- Used by software engineers of all types
- Better integrated with general-purpose coding
- Not especially fast

Agenda

- Data manipulation
- Plotting
- Machine Learning: Regression, Classification, Clustering