The R Language

CS102
Spring 2020

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	Python
R	**Python**
▪ Easier for experienced programmers	▪ Good for beginners or experienced
▪ Used by academics, researchers, hard-core data scientists	▪ Used by software engineers of all types
▪ Shorter code for complex analyses, statistics, graphics	▪ Better integrated with general-purpose coding
▪ Extremely slow!	▪ Not especially fast

Agenda

- **Data manipulation**

- Plotting

- Machine Learning: Regression, Classification, Clustering