

# The R Language

## (and Using R for Machine Learning)

Professor Widom's Instructional Odyssey

[www.professorwidom.org](http://www.professorwidom.org)



# 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

1. Data manipulation
2. Plotting
3. Machine Learning: Regression, Classification, Clustering

Plenty of your turn!

For help while working with R:

**Tutorials and help pages**

(website Course Materials)

➤ **Web search**