Introduction to $\mathcal{R}$

Sessions 2: Data management

Dag Tanneberg[1]

Potsdam Center for Quantitative Research
University of Potsdam, Germany
October 11/12, 2018

[1]Chair of Comparative Politics, UP dag.tanneberg@uni-potsdam.de

Introduction

Words of Warning

Data management in $\mathcal{R}$ is notoriously hard. Reasons include:

- Where's my spreadsheet?
- Barrage of new concepts
- Explosion of functions and notation
- Confusing diversity of options

Everybody struggles here.

Bear with me, and...

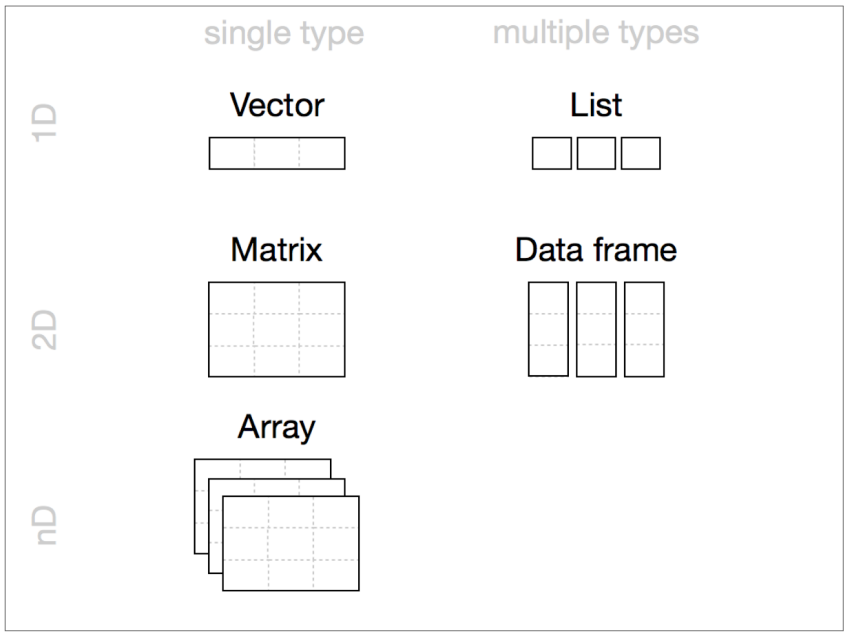
Outline

Data Structures

Organizing Principles[2]

1. **Dimensionality**
   How many qualities allowed?
2. **Homogeneity**
   All data single-typed?

Figure 1: Typical Data Structures



[2] Grolemund, G. 2014. Hands-on Programming with R. Sebastopol: O'Reilly, 62.

Vectors & Attributes

Matrices & Arrays

Data Frames & Lists

Data I/O

R Notation

Integers

Logical Values

Names

Dollar Signs & Double Brackets

Practical Data Management