

# GEOG 5680

## Introduction to R

### 02: Variables in R

Simon Brewer

Geography Department  
University of Utah  
Salt Lake City, Utah 84112  
`simon.brewer@geog.utah.edu`

May 02, 2020

# The R Environment

- The working directory
  - The directory R is currently running in
  - Affects access to files, etc
  - When you start R, it will generally start in a root directory
  - Change R to directory where you store your files
  - 'Session' menu — 'Set working directory'
  - Can also be set from the 'Files' tab

# The R Environment

- The working directory
  - The directory R is currently running in
  - Affects access to files, etc
  - When you start R, it will generally start in a root directory
  - Change R to directory where you store your files
  - 'Session' menu — 'Set working directory'
  - Can also be set from the 'Files' tab
- The workspace
  - R's memory
  - Variables and data are held here during a session
  - Note that you are not altering the contents of a file as you work (c.f. Excel, etc)
  - Can cause some limitations on data size

# Variables in R

- All data is read in and stored in memory for analysis
- Variables are *objects* used to store data
- A variable consists of a name and a value *that can be changed*:
  - `x <- 5`
    - Creates a variable `x` with the value 5
  - `x <- 10`
    - Changes the value of `x` to 10
  - `x <- x + 10`
    - Changes the value of `x` to 20

# Variables in R

- Variables have different modes and classes
- Modes
  - Numeric: e.g. 5 or 6e-5
  - Character: e.g. "Green" or "Red"
  - Factor (categories): e.g. "male" "female"
  - Logic: True/False
- Classes
  - Scalar/Vector/Matrix/Array/Data frame/List

# Variables in R

- **Simple variable/scalar**
- No dimension, no index
- E.g. population of Salt Lake City



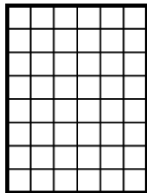
- **Vector**
- One dimension, one index  $[i]$
- Same mode
- E.g. series of monthly temperature



# Variables in R

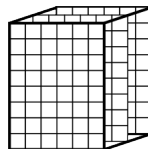
- **Matrix**

- Two dimensions, two indices  $[i, j]$
- Same mode
- E.g. gridded climate or raster image



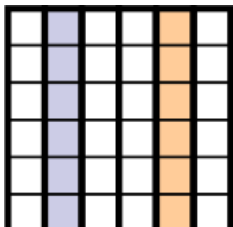
- **Array**

- $n$  dimension,  $n$  indices  $[i, j, \dots]$
- Same mode
- E.g. set of remote sensed images over time



# Variables in R

- **Data frame**
- 2 dimensions, 2 indices  $[i, j]$  or by name (\$)
- **Different** modes
- Most study data, e.g. census data (income class, household size),



- **List**
- Set of different objects, index by object name (\$)
- Frequently output from R functions — e.g. regression model may have coefficients,  $R^2$ , residuals

