

# *Introduction to R for Data Management and Analysis*

Marcel Ramos, MPH

Tuesday, June 6, 2016

## *Notes on last week's lecture*

- Using `c()` with mixed data types
- ... ellipses
  - additional arguments to “lower level” functions (e.g., `par()` in `plot()`)
- Any questions?

- Basic features of the language
  - interactive and *interpreted*
  - case sensitive, ignores spaces except between objects and functions
  - an object is any type of variable stored in R (i.e., `data.frame`, `numeric` vector, `function`, etc.)
  - Be familiar with the parts of a function
  - know your help pages (use `?functionname` or `help("functionname")`)
  - Finding help and troubleshooting are critically important
  - Help pages can be intimidating but very useful

## Today's lecture overview

- Importing and Exporting Data
- Classes and Data types in R
  - `data.frame`
  - data type coercion
- Subsetting
  - using brackets and dollar signs (`[`, `$`)
  - using vector operands
- More troubleshooting

# Importing and Exporting Data

- Recognizing file types
  - File extensions (e.g., .tsv, .csv, .xlsx, .txt, .sav, .sasb7dat)
- Show file extensions on Windows
  - Windows Start button > Control panel > (View by: Large Icons) > Folder options > View tab > Advanced Settings: [**Deselect**] Hide extensions for known file types
- Downloading files from the internet

## *Packages for reading foreign data*

- readr, readxl, haven

# Exporting Data

- `write.table()`
- `write.csv()` / `write_csv()`
- `write_delim()`

## *Classes in R*

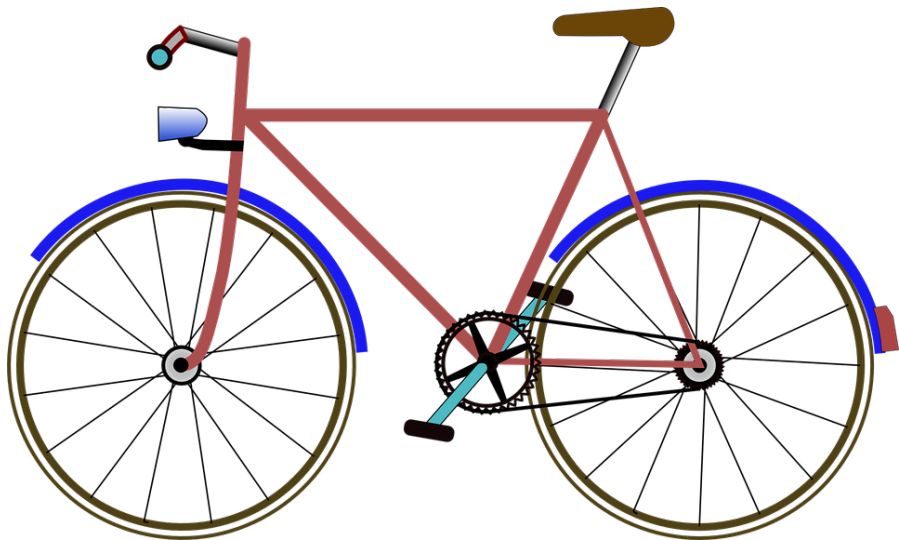


Figure 1: Bicycle



# *What are classes?*

- R objects have class attributes
- Define what functions/operations can be performed
- `class()` function
- Examples

# Vectors, classes

- Vectors of class (may be named):
  - character, integer, numeric, logical, complex, raw (bytes)
  - factor: discrete levels
  - missing (NA)
- Tabular classes:
  - `data.frame`
  - `matrix`
- Non-tabular:
  - `function`
  - `list`
  - `array`
  - custom classes