

# Quant II

## Lab 0: Intro to R Markdown

Junlong Aaron Zhou

February 04, 2021

# Introduce the TA

- Junlong Aaron Zhou, 4<sup>th</sup> year PhD.
- Working on political methodology and comparative politics
- My personal website: <http://www.zhoujunlong.com/>
- Email: **jlzhou@nyu.edu**

- Lab: Fridays, 12:30 - 14:30 EST
  - Recording will be posted after class.
  - Materials can be find at [github repo]
- Office hours: Mondays 16-18 EST; Tuesdays 9-10 EST
  - book my office hour [here]

- Lab: Fridays, 12:30 - 14:30 EST
  - Recording will be posted after class.
  - Materials can be find at [github repo]
- Office hours: Mondays 16-18 EST; Tuesdays 9-10 EST
  - book my office hour [here]
- All homework will be submitted in soft copy (pdf) to both Cyrus and me
- **Code should be well commented**
- Only accept high quality tables and plots

# Plan

- R Markdown
- Basic R
- Extra Basic R (time allows)

# Install R

- It depends on whether you are a Mac or Windows user
- For MAC, go to <http://cran.r-project.org/bin/macosx/>
- For Windows, go to <http://cran.r-project.org/bin/windows/base/>
- You should choose an editor and learn how it works
  - One common choice is RStudio, but there is vim, emacs, Notepad++, etc.

# When things break

- Documentation - Ex: ?lm
- Google
- CRAN (Reference manuals, vignettes, etc) - Ex:  
<http://cran.r-project.org/web/packages/AER/index.html>
- JSS - Ex: <http://www.jstatsoft.org/v27/i02>
- Stack Overflow - <http://stackoverflow.com/questions/tagged/r>
- Listservs - <http://www.r-project.org/mail.html>

# Resources

- The Art of R Programming - N. Matloff
- Modern Applied Statistics with S - W. Venables and B. Ripley
- Advanced R Programming - H. Wickham
- The R Inferno - P. Burns
- Rdataviz - a talk by P. Barberá on ggplot2
- Basic Intro to R - also by P. Barberá
- Jamie Monogan



# What is R Markdown?

R Markdown is a tool within RStudio that allows you to write documents, presentations, or webpages and combine written text with code. The text in the document can be fully formatted, and you can choose to make your code visible or not. Documents can be output as PDFs, Word Documents, HTML, and other formats.

# Why people use R Markdown?

- It has all the advantages of LaTeX
- Everything (analysis + texts + graphs + tables) is on one place.
- The document changes when the data gets updated (this is very important).
- Better bibliography system

# What if I prefer STATA?

- You can still use Markdown via Stata Markdown
  - Markdown is a language and R Markdown is just one realization of it.
  - Stata Markdown is similar.
- To export results, you can try `outreg` in STATA
- Nevertheless, we focus on R in this class.

# Getting started

It is very easy. You of course need to download R, RStudio. In newer versions you don't need to install anything else, however, you might need to install packages once (knitr, rmarkdown), and you may need some distribution of LaTeX if you want pdf output.

Once you have done all that,

- 1 Open RStudio
- 2 Select File > New File > R Markdown
- 3 Enter Title, Author, Output Format (which can easily be changed later)
- 4 You are good to go! You will see that there is some example text that you can delete.

## Markdown document

The documents typically have four different pieces: the YAML, the formatted text, code chunks, and inline code.

# YAML

At the beginning of any R Markdown script is a YAML header section enclosed by `---`. This includes a title, author, date and the type of output you want to produce. Many other options are available for different functions and formatting, like `beamer_presentation`, `pdf_document`, `html_document`, etc. Take a look [\[here\]](#) at all the available output formats. Rules in the header section will alter the whole document.

You will always insert something like this at the top of your new .Rmd script:

```
---  
title: "0. Intro to R Markdown"  
author: "Junlong Aaron Zhou"  
date: "February 04, 2021"  
output: html_document  
---
```

# Code chunks

Code chunks are the best part of R Markdown. The code you enter is executed and the results are displayed in the document.

- The syntax for code chunks looks like this:

```
'''{r}  
basic code  
'''
```

# Code chunks

Note that these are backticks, not quotes or apostrophes. Everything that goes between these backticks should have R syntax.

```
x <- runif(10)
y <- rnorm(10, x, 1)
df <- data.frame(x, y)
df
```

```
##           x           y
## 1  0.5147753  0.3329132
## 2  0.5611408  1.3061028
## 3  0.3493416  1.4063719
## 4  0.9383252  1.5653399
## 5  0.2144501 -1.7998002
## 6  0.5247760 -0.3212232
## 7  0.6252590  0.2608835
## 8  0.4347588 -1.1921694
```

# Code chunk options

```
' ' ' {r name, eval=FALSE, warning=FALSE, message=FALSE}  
CODE GOES HERE  
' ' '
```

- name - A label for your code chunk.
- echo - Whether to display the code chunk or just show the results. If you want the code embedded in your document but don't want the reader of the document to see it, you can set echo=FALSE
- eval - Whether to run the code in the code chunk.
- warning - Whether to display warning messages in the document.
- message - Whether to display code messages in the document.
- results - Whether and how to display the computation of the results



# Formatted text

It is very easy to format your text in Markdown.

- `*italics*` *italics*
- `**bold**` **bold**
- `~~strikethrough~~` ~~strikethrough~~
- `[nyu](https://www.nyu.edu/)` nyu
- `superscript^2^` superscript<sup>2</sup>

And many more you can find here, including headers, lists, etc.

# Equations

Equations use LaTeX notation and you can include them either between single dollar signs (inline equations) or double dollar signs (displayed equations).

We can use inline equations such as  $y_i = \alpha + \beta x_i + e_i$  which is displayed as  $y_i = \alpha + \beta x_i + e_i$ .

Or displayed formulas:

$$f(x) = \frac{e^{(-x-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

Which looks like this:

$$f(x) = \frac{e^{(-x-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$
$$y = x + 3$$

# Equations

For multiline equations, you need the `aligned` environment:

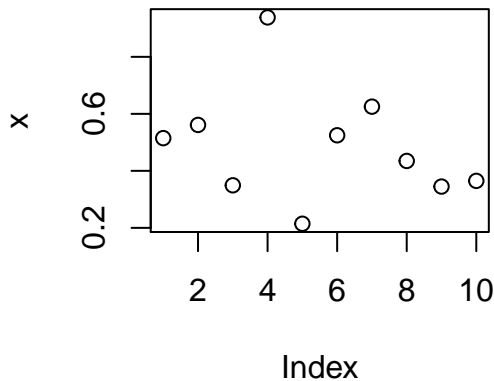
$$x = 2 * z + 3$$

$$y = 5 * z + 6$$

# Plot

We can also plot figures,

```
plot(x)
```



# Plot

Or attach a picture





The screenshot displays the RStudio environment. The left pane shows the source editor with an R Markdown document titled '0. Intro to R Markdown'. The document includes a title, author, date, output format (HTML), and a section titled 'Introduce the TA' with a list of bullet points. The right pane shows the rendered HTML output, which includes the title, author, date, and the 'Introduce the TA' section with the same bullet points. The bottom pane shows the 'Logistics' section.

```
1 * ---
2 title: "0. Intro to R Markdown"
3 author: "Junlong Aaron Zhou"
4 date: "`r format(Sys.time(), '%B %d, %Y')`"
5 output:
6   html_document:
7     df_print: paged
8   pdf_document: default
9   biblio-style: apsr
10 theme: Singapore
11 * ---
12
13 # Introduce the TA
14 - Junlong Aaron Zhou, 4th year PhD.
15 - Working on political methodology and comparative politics
16 - My personal website: [http://www.zhoujunlong.com/](http://www.zhoujunlong.com/)
17
```

## 0. Intro to R Markdown

Junlong Aaron Zhou  
February 04, 2021

### Introduce the TA

- Junlong Aaron Zhou, 4<sup>th</sup> year PhD.
- Working on political methodology and comparative politics
- My personal website: <http://www.zhoujunlong.com/>

## Logistics