1. Introduction to R and Rstudio

Introduction to Social Network Analysis in R

Dr. Uma Ravat

University of California at Santa Barbara

Welcome

Introductions - Let's get to know one another

1. Introduction to R

2. Introduction to Rstudio

3. R Markdown documents

Welcome

Welcome

umaravat@ucsb.edu

Introduction to Social Network Analysis(SNA) in R

1. Introduction to R as we will use R language for SNA for the rest of the lecture series
2. Introduction to basic concepts in SNA and visualization of networks.
3. Metrics - Individual nodes.
4. Metrics - Whole network.
5. (Time permitting) Network models, algorithms and Inference.
6. SNA in Education, Surveys and Data Manipulation.
7. Ongoing Research Project with Keio University.

Introductions - Let's get to know one another

Introductions - Let's get to know one another

Go ahead and tell everybody

1. your name,
2. one of the following:
    - Your Hometown
    - Your favorite food or snack
    - Your hobbies and interests
    - An interesting fact about yourself
    - Two Truths and a Lie
    - Or anything else!
3. Why are you interested in this lecture series?

After you're done sharing, pick someone who hasn't shared yet to go next

1. Introduction to R

What *is* R?

- R is an open-source(free) statistical **programming language**
- R is also an environment for statistical computing and **graphics**
- It's easily extensible with *packages* (more on this later)
- R is based on the S language, which was developed by Bell laboratories in the 90's
- Home page: http://www.r-project.org

Why *use* `R`?

- Open source (free)
- Runs on just about any platform
- Great visualization capabilities (`ggplot2`)
- Read/write from/to various data sources
- Scripting language (interpreted)
- Massive library of data manipulation and statistical packages (including great network packages like `statnet` and `igraph`)

Why *use* R?

- several functions are vectorized in R
  - vectorized code saves time (any typing)
- There is an optimized engine —a basic linear algebra system (BLAS)—that is highly efficient at solving linear algebra problems
- A lot or R functions are written in C (or variants)

For more details see: https://www.noamross.net/archives/2014-04-16-vectorization-in-r-why/

R and RStudio

R: Engine	RStudio: Dashboard
	

- `R` is a programming language.
- `RStudio` is a convenient interface for R called an **IDE** (integrated development environment), e.g. *"I write R code in the RStudio IDE"*

2. Introduction to Rstudio

Rstudio

RStudio

Go to file/function | Addins | Project: (None)

L1_nw.Rmd | w1-d01-welcome.Rmd | unvotes.Rmd | L1.Rmd

Knit on Save | Knit | Run

Source | Visual | Outline

```
39  We can easily change which countries are being plotted by changing
    which countries the code above `filter`s for. Note that the
    country name should be spelled and capitalized exactly the same
    way as it appears in the data. See the [Appendix](#appendix) for a
    list of the countries in the data.
40
41  ```{r plot-yearly-yes-issue, fig.width=10, fig.height=6,
    message=FALSE}
42  unvotes %>%
43    filter(country %in% c("UK & NI", "US", "Turkey")) %>%
44    mutate(year = year(date)) %>%
45    group_by(country, year, issue) %>%
46    summarize(percent_yes = mean(vote == "yes")) %>%
47    ggplot(mapping = aes(x = year, y = percent_yes, color =
    country)) +
48    geom_point(alpha = 0.4) +
49    geom_smooth(method = "loess", se = FALSE) +
```

1:1 | UN Votes | R Markdown

Console | Render | Background Jobs

.../yt-01a-un-votes/unvotes.Rmd

```
 --toc-depth 3 --variable toc_float=1 --variable toc_selectors=h1,h2,h3
 --variable toc_collapsed=1 --variable toc_smooth_scroll=1 --variable to
c_print=1 --template /Library/Frameworks/R.framework/Versions/4.2/Resour
ces/library/rmarkdown/rmd/h/default.html --no-highlight --variable highl
ightjs=1 --variable theme=bootstrap --mathjax --variable 'mathjax-url=ht
tps://mathjax.rstudio.com/latest/MathJax.js?config=TeX-AMS-MML_HTMLorMM
L' --include-in-header /var/folders/wq/cjgk7v356dq1ljsj90802q9h0000gq/
T//Rtmpbfufuw/rmarkdown-str105d53f7a771d.html
output file: unvotes.knit.md


Output created: unvotes.html
```

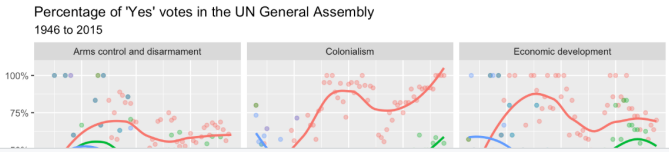
Environment | History | Connections | Tutorial

Files | Plots | Packages | Help | Viewer | Presentation

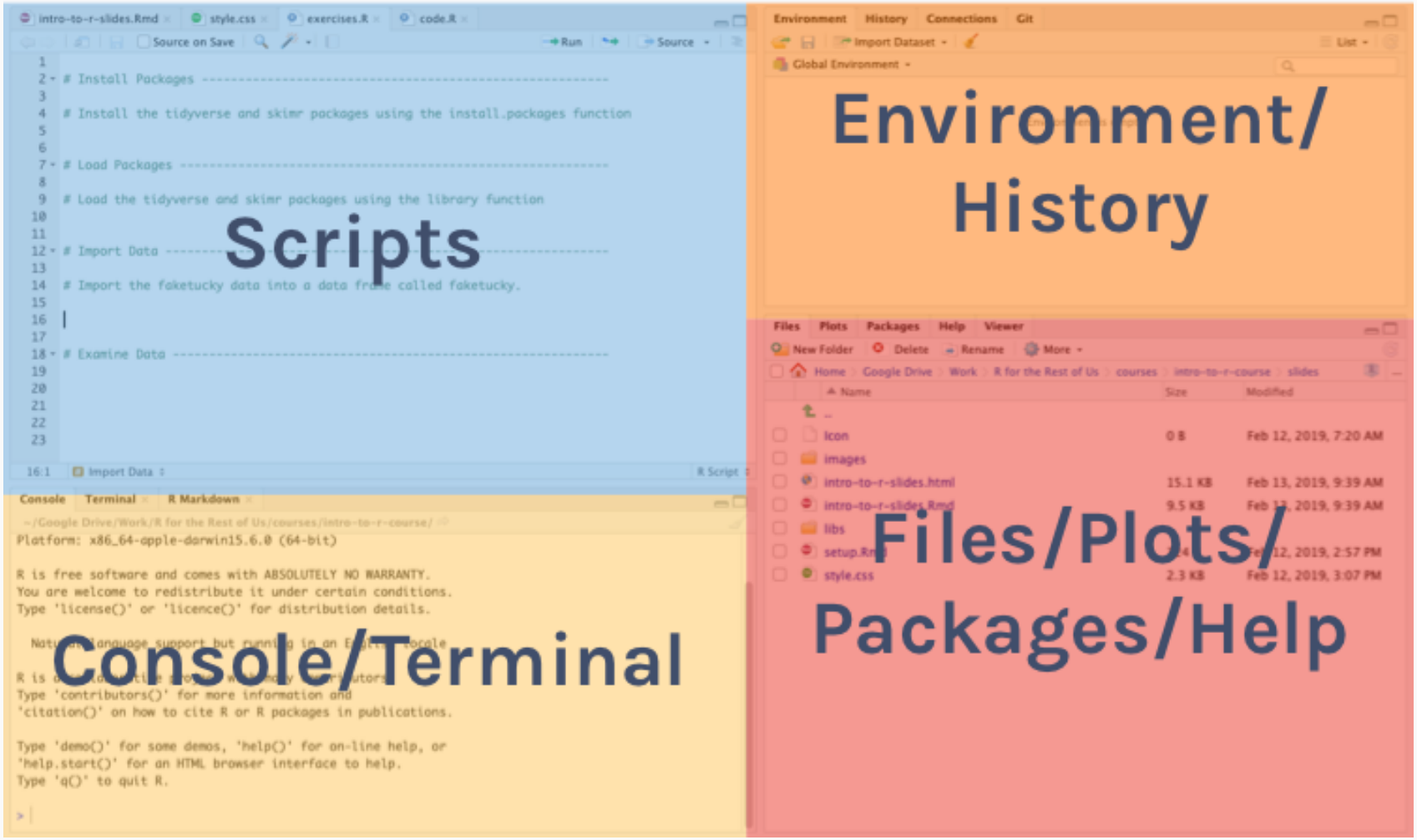
Publish

We can easily change which countries are being plotted by changing which countries the code above `filter`s for. Note that the country name should be spelled and capitalized exactly the same way as it appears in the data. See the Appendix for a list of the countries in the data.

```
unvotes %>%
  filter(country %in% c("UK & NI", "US", "Turkey")) %>%
  mutate(year = year(date)) %>%
  group_by(country, year, issue) %>%
  summarize(percent_yes = mean(vote == "yes")) %>%
  ggplot(mapping = aes(x = year, y = percent_yes, color = countr
y)) +
  geom_point(alpha = 0.4) +
  geom_smooth(method = "loess", se = FALSE) +
  facet_wrap(~issue) +
  scale_y_continuous(labels = percent) +
  labs(
    title = "Percentage of 'Yes' votes in the UN General Assembl
y",
    subtitle = "1946 to 2015",
    y = "% Yes",
    x = "Year",
    color = "Country"
  )
```



Tour of RStudio panes



R packages

- **Packages** are the fundamental units of reproducible R code. They include reusable R functions, the documentation that describes how to use them, and sample data
- There are over 18,000 R packages available on **CRAN** (the Comprehensive R Archive Network)1
- We will use various packages in this course such as `base` `R`,`graphics`, `igraph`, `igraphdata` etc.

1 Community contributed packages are stored at CRAN Comprehensive R Archive Network

Your Turn: 1. Tour of `Rstudio` panes

Download files from url provided.

Go to RStudio and open the file named `1_WA_IntroRstudioR.Rmd`

We will go over 1. Tour of four `Rstudio` panes

If you finish, react with a "thumbs up" in Zoom and help your class mates who may need help finishing.

A tour of R essentials

> *To understand computations in R, two slogans are helpful:*
> *Everything that exists is an object.*
> *Everything that happens is a function call.*
> *— John Chambers*

Even a *function* is an *object*.

R essentials (continued)

- **Functions** are (most often) verbs, followed by what they will be applied to in parentheses:

```
do_this(to_this)
```

Here `do_this` is the function and `to_this` is the **argument** to the function

```
do_that(to_this, to_that, with_those)
```

Here `do_that` is the function and `to_this, to_that, with_those` are the three **arguments** to the `do_that` function

- **Packages** are installed with the `install.packages` function and loaded with the `library` function, once per session:

```
install.packages("package_name")
library(package_name)
```

Tour of `R` essentials

We will use `Rmarkdown` documents to write our `R` code or programs.

Before we go on a tour of `R` essentials, let's look at `Rmarkdown` documents first

3. R Markdown documents

3. R Markdown documents

- with `.Rmd` extension
- **rmarkdown** is an `R` package
  - write code and prose in **reproducible** computational documents

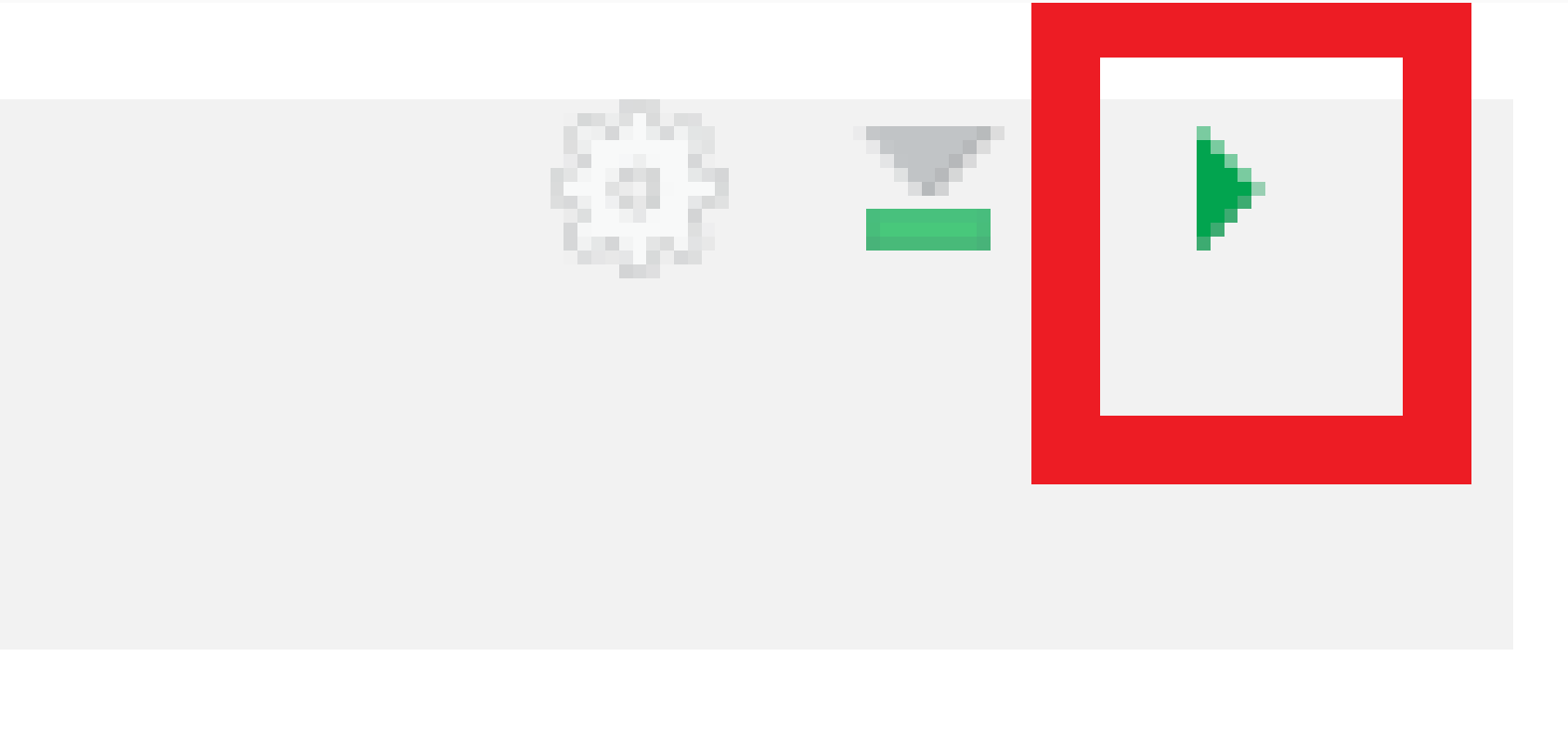
rmarkdown.rstudio.com

Take a look at the **Rmarkdown gallery**

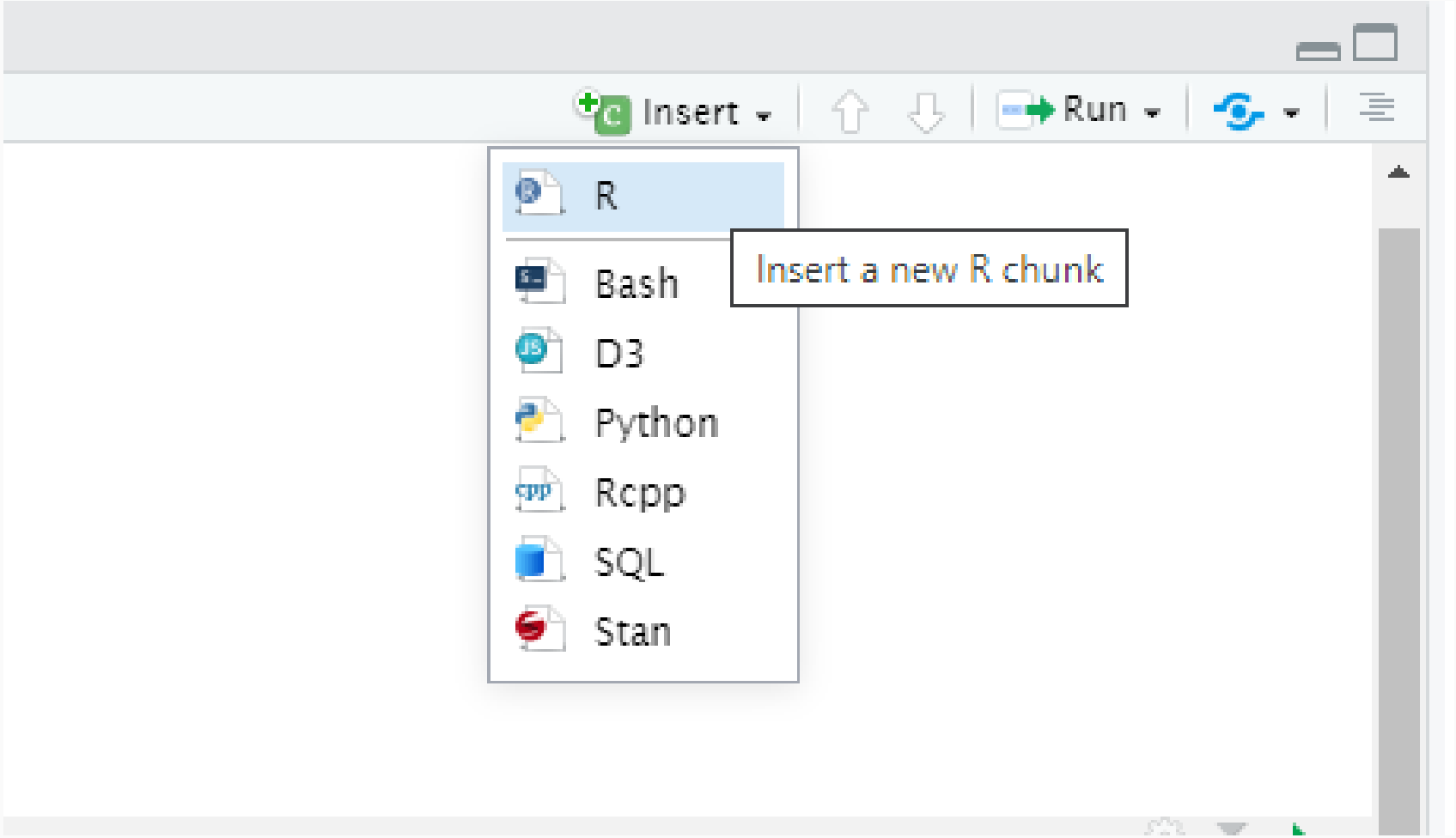
Anatomy of `RMarkdown` documents

- Code goes in code chunks, defined by three backticks
- narrative goes outside of chunks
- Simple markdown syntax for text

Demo - Run Code

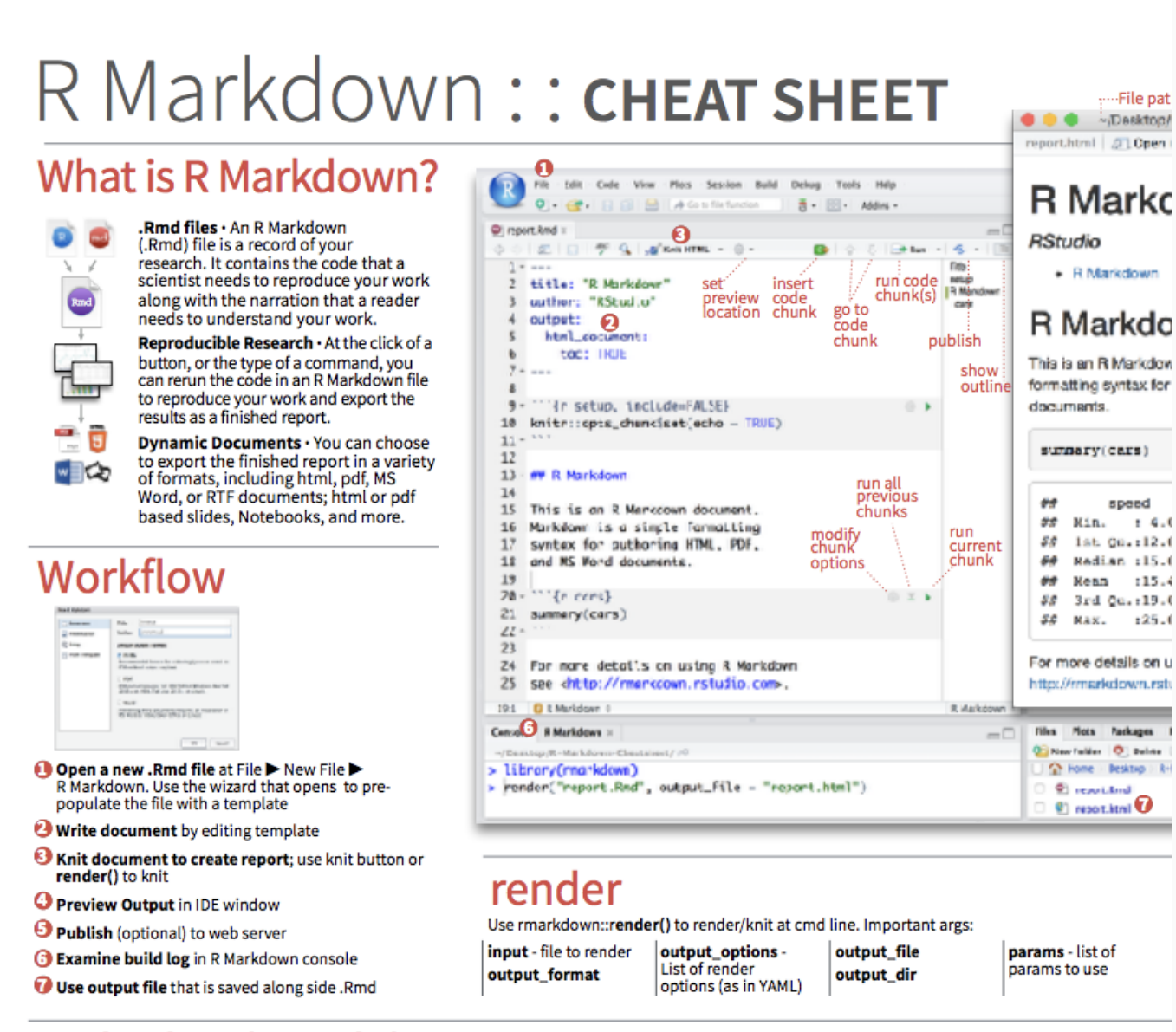


Demo - Adding Chunks



R Markdown help in RStudio

R Markdown Cheat Sheet `Help -> Cheatsheets`



Markdown Quick Reference

Help -> Markdown Quick Reference

Files | Plots | Packages | Help | Viewer

Markdown Quick Reference | Find in Topic

Markdown Quick Reference

R Markdown is an easy-to-write plain text format for creating dynamic documents and reports. See Using R Markdown to learn more.

Emphasis

```
*italic*   **bold**

_italic_   __bold__
```

Headers

```
# Header 1

## Header 2

### Header 3
```

Lists

Unordered List

```
* Item 1
* Item 2
    + Item 2a
    + Item 2b
```

Ordered List

```
1. Item 1
2. Item 2
3. Item 3
    + Item 3a
    + Item 3b
```

Manual Line Breaks

End a line with two or more spaces:

```
Roses are red,
Violets are blue.
```

(Optional) Your Turn: Rmarkdown basics

- We will not cover this here.
- You may do this to familiarize youself with `Rmarkdown` syntax and basics
- Open `2_WA_RmarkdownBasics.Rmd` and complete on your own
- Ask questions!

Tour of R essentials

Go to RStudio and we will nowgo over 2. Tour of R essentials.

Lecture 1 Summary: Introduction to R

- Rstudio
- R
- Rmarkdown essentials.

Next session:

2. Introduction to basic concepts in SNA and visualization of networks.

In preparataion of next session on Friday, you may take a look at network datasets available in `igraph` package and read the description of networks contained in this package at:

https://cran.r-project.org/web/packages/igraphdata/igraphdata.pdf