

#swRk

Introduction to R Programming for Social Workers



Gina Griffin. MSW, LCSW, DSWc

What is R Programming?

The language:

“R is a language and environment for statistical computing and graphics.”

“R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is highly extensible.”

“One of R’s strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed.”

(source: <https://www.r-project.org/about.html>)

What is R Programming?

The language (contd):

“R is available as Free Software under the terms of the [Free Software Foundation](#)’s [GNU General Public License](#) in source code form.”

(source: <https://www.r-project.org/about.html>)

What is R Programming?

The R Environment (contd):

“R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes

- an effective data handling and storage facility,
- a suite of operators for calculations on arrays, in particular matrices,
- a large, coherent, integrated collection of intermediate tools for data analysis,
- graphical facilities for data analysis and display either on-screen or on hardcopy, and
- a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

(source: <https://www.r-project.org/about.html>)

What is R Programming?

The R Environment (contd):

“R, like S, is designed around a true computer language, and it allows users to add additional functionality by defining new functions.”

“Many users think of R as a statistics system. We prefer to think of it as an environment within which statistical techniques are implemented. R can be extended (easily) via *packages*.”

“There are about eight packages supplied with the R distribution and many more are available through the CRAN family of Internet sites covering a very wide range of modern statistics.”

(source: <https://www.r-project.org/about.html>)

What are R Packages?

You install them from CRAN with `install.packages("x")`

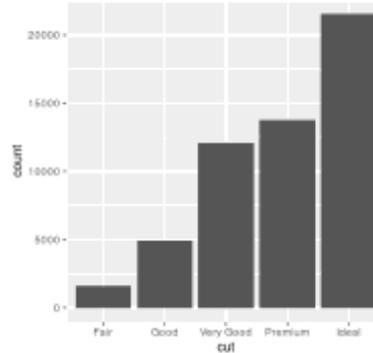
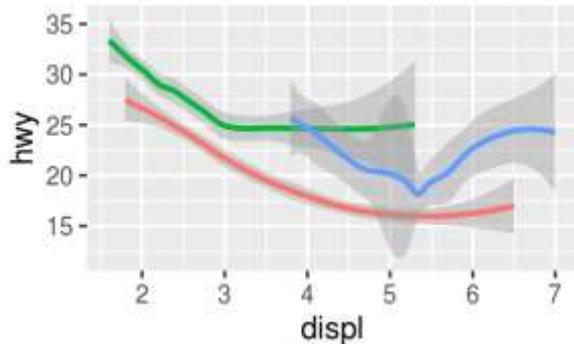
You use them in R with `library("x")`

You get help on them with `packages?x` and `help(package = "x")`

(source: <http://r-pkgs.had.co.nz/intro.html>)

How can I use R?

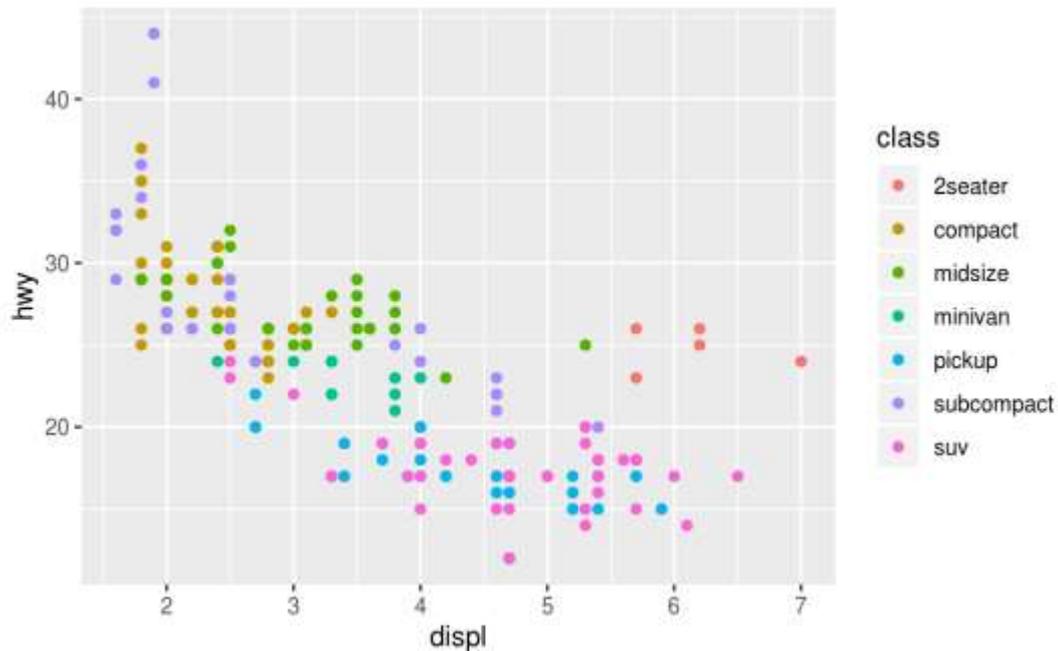
- Straight up statistical data analysis
- Data visualization
- Reproducible research
- Reports, papers, webpages, presentations and other outputs



How can I use R?

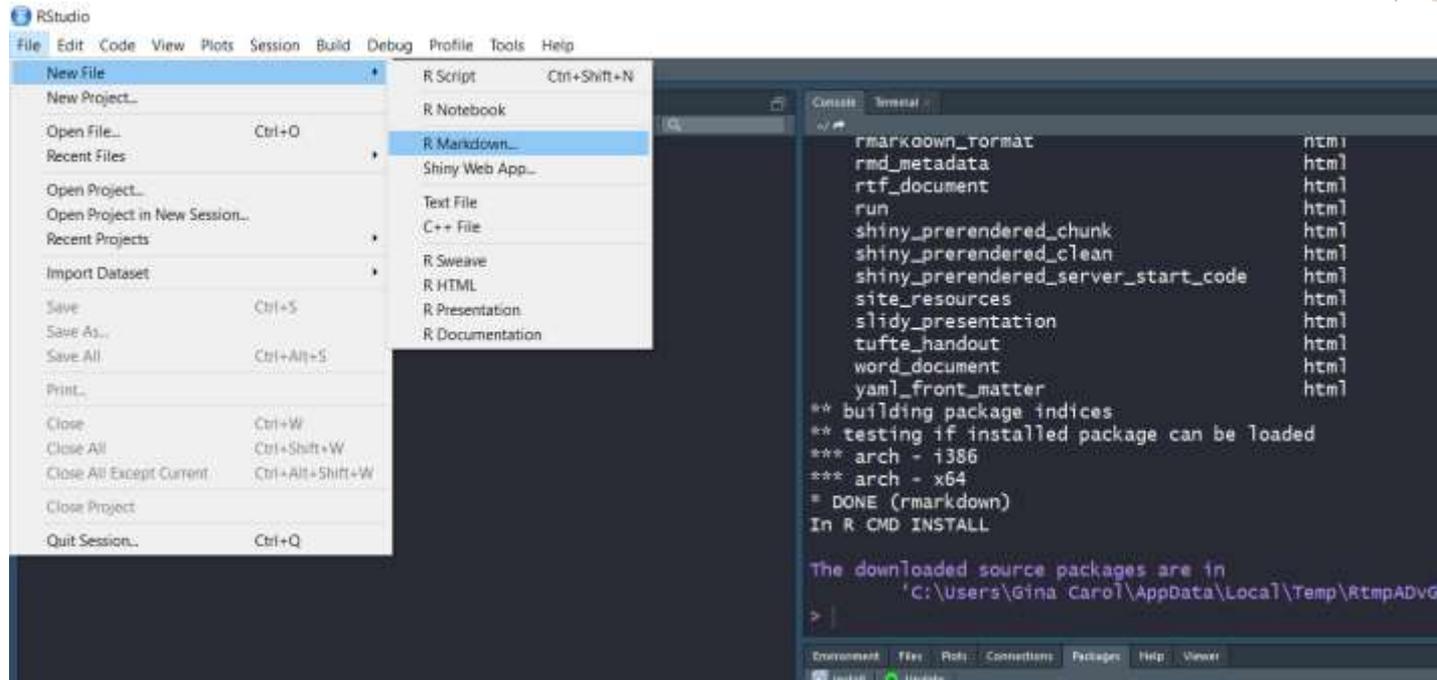
This is a plot with the colors of the points mapped to the **class** variable.

```
ggplot(data = mpg) +  
  geom_point(mapping =  
    aes(x = displ, y = hwy,  
        color = class))
```



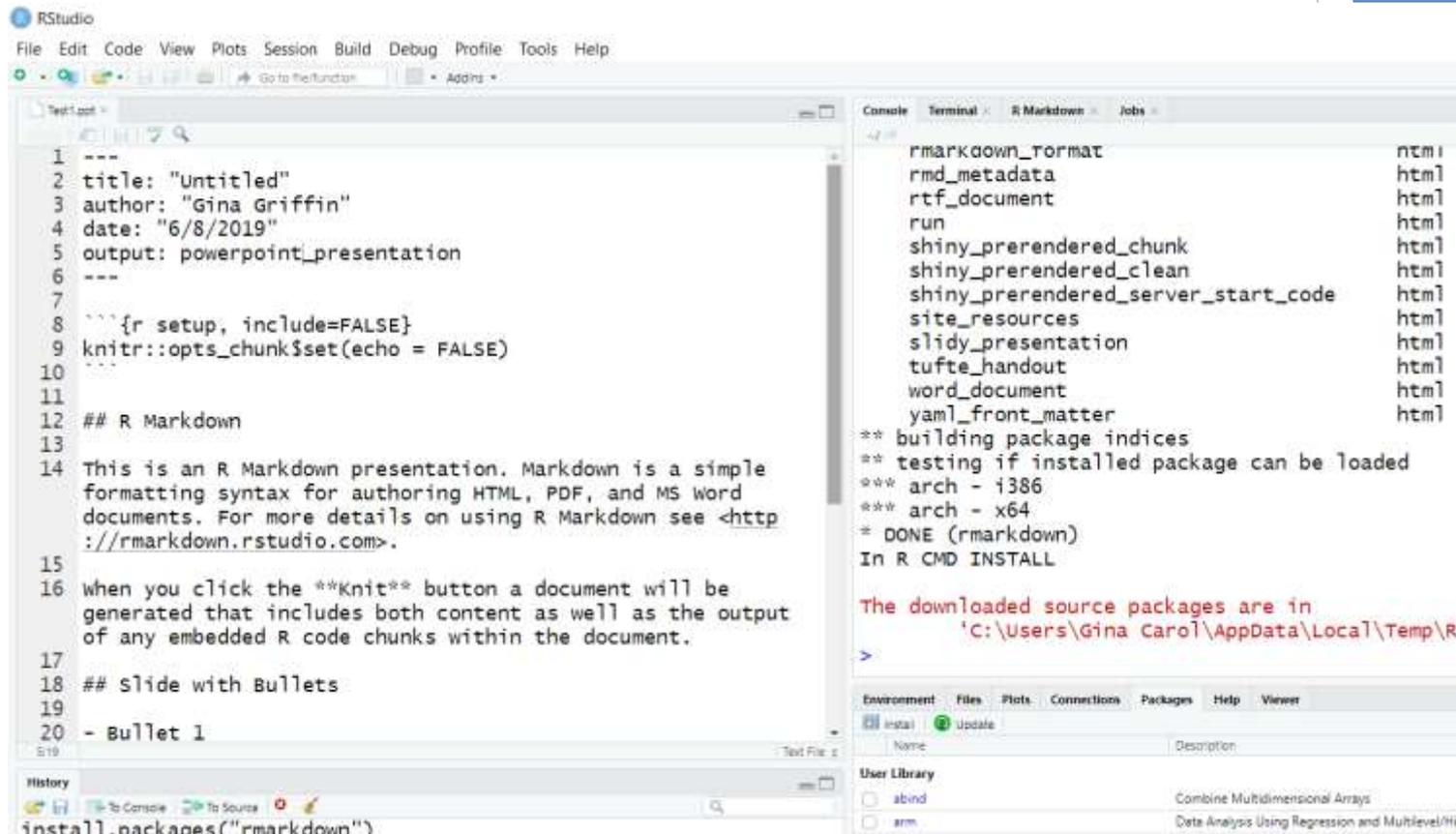
How can I use R?

R Markdown is a file format for making dynamic documents with R. An R Markdown document is written in markdown (an easy-to-write plain text format) and contains chunks of embedded code.



How can I use R?

R Markdown



The screenshot displays the RStudio interface with an R Markdown document open in the editor and the terminal window showing the output of the `install.packages("rmarkdown")` command.

R Markdown Document Content:

```
1 ---
2 title: "Untitled"
3 author: "Gina Griffin"
4 date: "6/8/2019"
5 output: powerpoint_presentation
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = FALSE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown presentation. Markdown is a simple
15 formatting syntax for authoring HTML, PDF, and MS word
16 documents. For more details on using R Markdown see <http
17 ://rmarkdown.rstudio.com>.
18
19 When you click the knit button a document will be
20 generated that includes both content as well as the output
21 of any embedded R code chunks within the document.
22
23 ## slide with Bullets
24
25 - Bullet 1
```

Terminal Output:

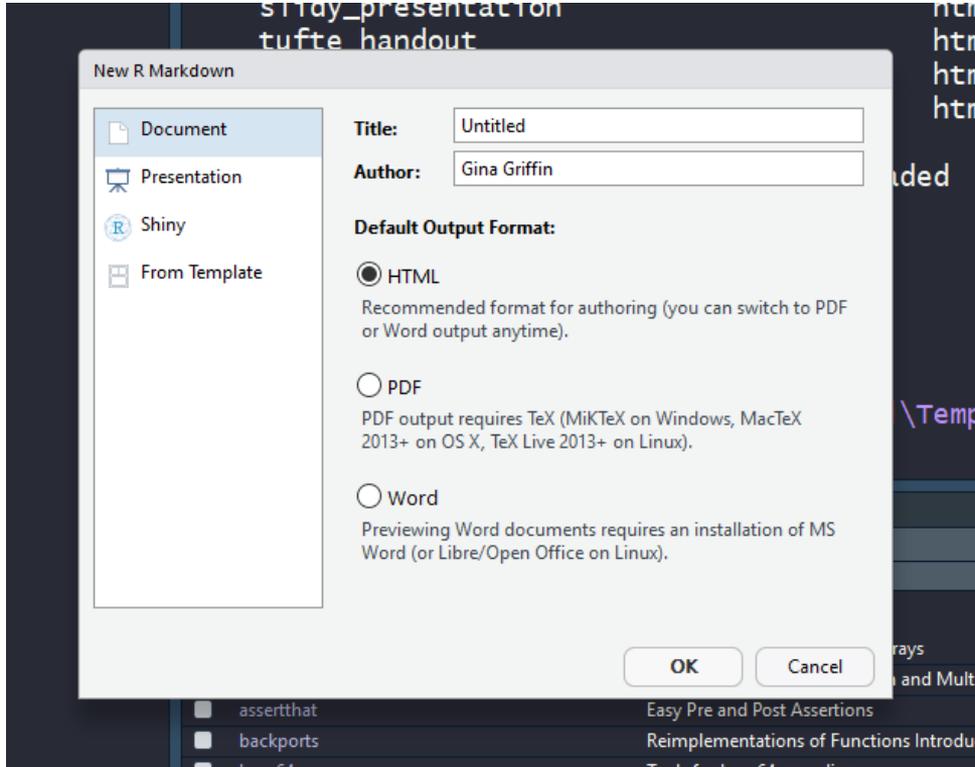
```
rmarkdown_format          html
rmd_metadata              html
rtf_document              html
run                       html
shiny_prerendered_chunk  html
shiny_prerendered_clean  html
shiny_prerendered_server_start_code  html
site_resources            html
slidy_presentation       html
tufte_handout            html
word_document            html
yaml_front_matter       html
** building package indices
** testing if installed package can be loaded
*** arch - i386
*** arch - x64
* DONE (rmarkdown)
In R CMD INSTALL

The downloaded source packages are in
  'C:\Users\Gina Carol\AppData\Local\Temp\R'
```

Environment:

Name	Description
abind	Combine Multidimensional Arrays
arm	Data Analysis Using Regression and Multilevel/Hierarchical Models

How can I use R?



Types of document outputs include HTML (the default), PDF, Word...

How can I use R?

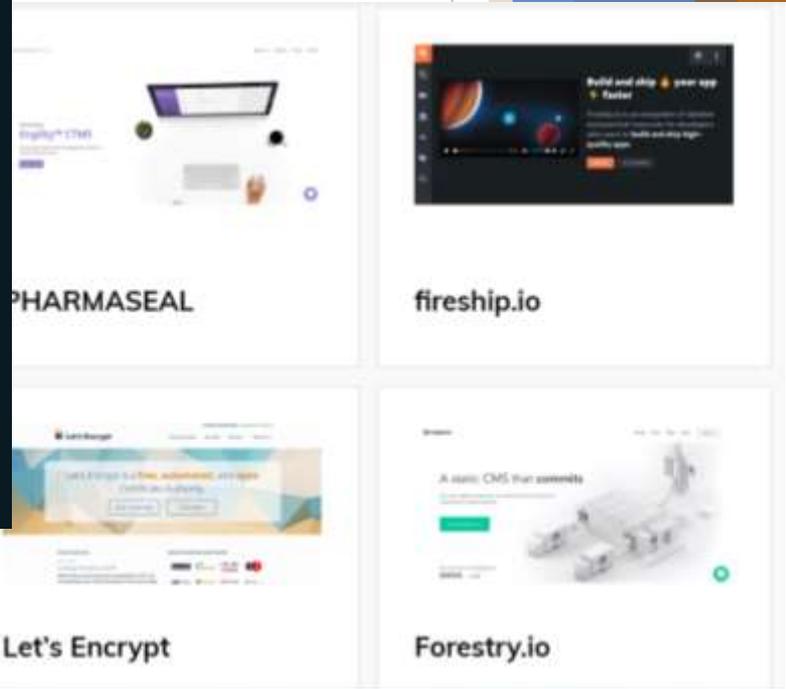
Create Websites



The world's fastest
framework for building
websites

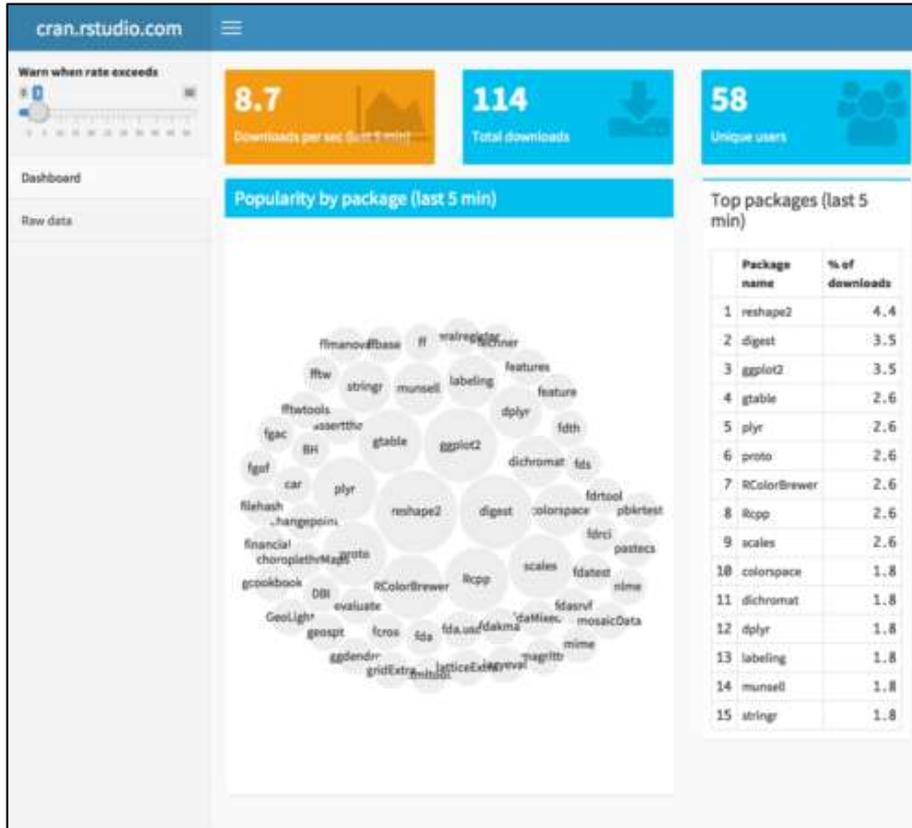
Hugo is one of the most popular open-source static site generators. With its amazing speed and flexibility, Hugo makes building websites fun again.

<https://gohugo.io/>



How can I use R?

Shiny Dashboards



A **data dashboard** is an information management tool that visually tracks, analyzes and displays **key performance indicators (KPI)**, metrics and key data points to monitor the health of a business, department or specific process. They are customizable to meet the specific needs of a department and company. Behind the scenes, a dashboard connects to your files, attachments, services and API's, but on the surface displays all this data in the form of tables, line charts, bar charts and gauges. A data dashboard is the most efficient way to track multiple data sources because it provides a central location for businesses to monitor and analyze performance. Real-time monitoring reduces the hours of analyzing and long line of communication that previously challenged businesses.

(Source:

<https://www.klipfolio.com/resources/articles/what-is-data-dashboard>)

Installing R Software

- You will need to install R and R Studio.
- You can use R independently. But if you want to use R Studio, you will also need to install R.

To install R:

<https://www.r-project.org/>

To install R Studio:

<https://rstudio.com/products/rstudio/download/>

RStudio Desktop <small>Open Source License</small> Free <small>Download</small> DOWNLOAD	RStudio Desktop <small>Commercial License</small> \$995 <small>Buy</small> BUY	RStudio Server <small>Open Source License</small> Free <small>Download</small> DOWNLOAD	RStudio Server Pro <small>Commercial License</small> \$4,975 <small>Buy</small> BUY
---	---	--	--

Now what? Shout-out to the tidyverse!

What is the tidyverse?

“The tidyverse is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures.”

(source: <https://www.tidyverse.org/>)

You can install them all at once with the following command:

```
install.packages("tidyverse")
```

Now what? Shout-out to the tidyverse!

Core tidyverse

“The core tidyverse includes the packages that you’re likely to use in everyday data analyses.”

(source: <https://www.tidyverse.org/packages/>)

Now what? Shout-out to the tidyverse!



Core tidyverse packages:

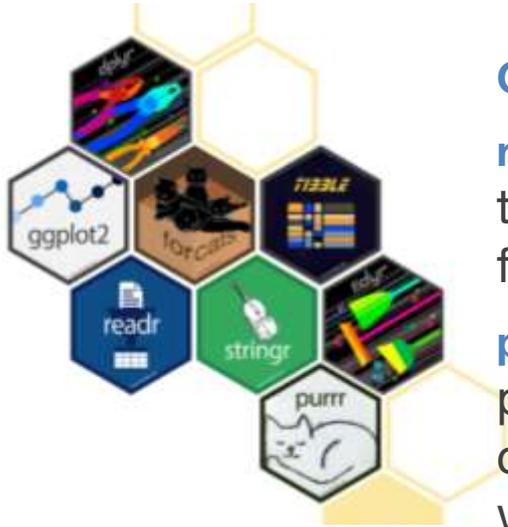
ggplot2: ggplot2 is a system for declaratively creating graphics.

dplyr: dplyr provides a grammar of data manipulation.

tidyr: tidyr provides a set of functions that help you get to tidy data. Tidy data is data with a consistent form: in brief, every variable goes in a column, and every column is a variable.

(source: <https://www.tidyverse.org/packages/>)

Now what? Shout-out to the tidyverse!



Core tidyverse packages (contd):

readr: readr provides a fast and friendly way to read rectangular data (like csv, tsv, and fwf).

purrr: purrr enhances R's functional programming (FP) toolkit by providing a complete and consistent set of tools for working with functions and vectors.

tibble: Tibbles are data.frames that are lazy and surly: they do less and complain more forcing you to confront problems earlier

(source: <https://www.tidyverse.org/packages/>)

Now what? Shout-out to the tidyverse!



Core tidyverse packages (contd):

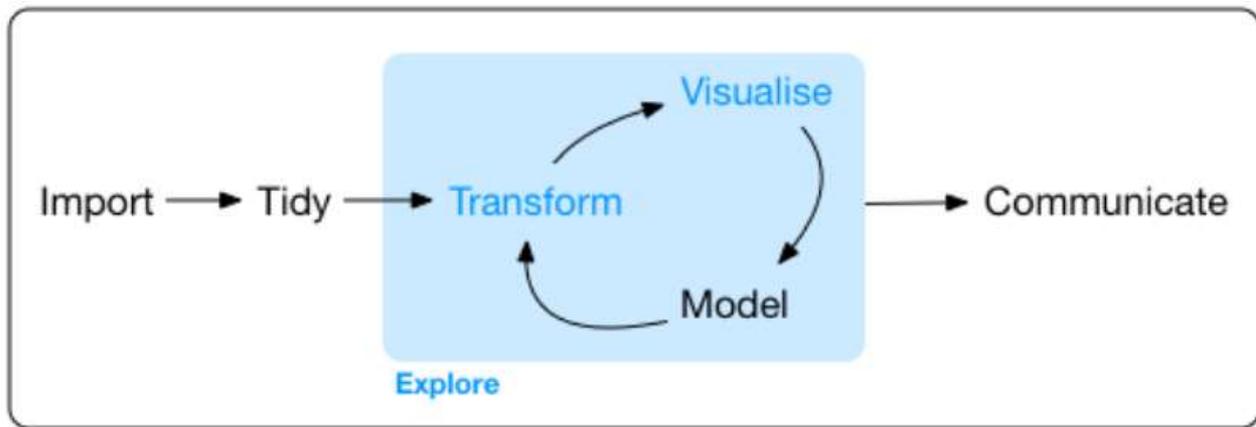
stringr: stringr provides a cohesive set of functions designed to make working with strings as easy as possible.

forcats: forcats provides a suite of useful tools that solve common problems with factors. R uses factors to handle categorical variables, variables that have a fixed and known set of possible values.

(source: <https://www.tidyverse.org/packages/>)

Now what? Shout-out to the tidyverse!

tidyverse data exploration model:



Program

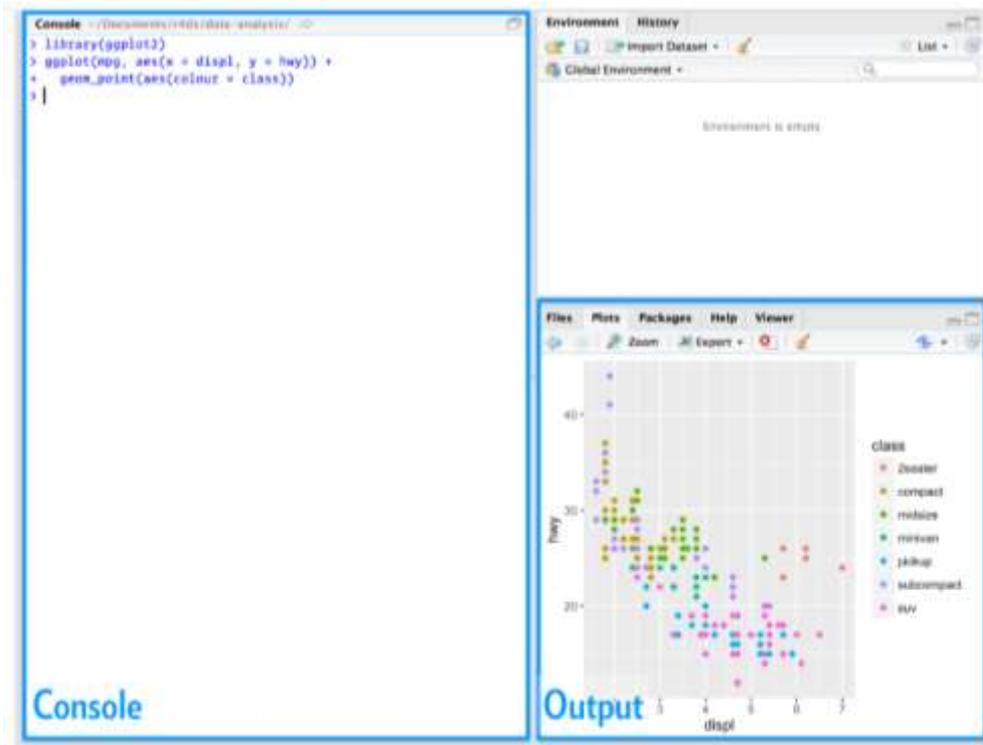
(source: <https://r4ds.had.co.nz/explore-intro.html>)

Putting it to Work

The background features a series of overlapping, semi-transparent geometric shapes in shades of blue and brown. These shapes are primarily located on the right side of the frame, creating a dynamic, layered effect. The colors range from a light tan to a deep, dark brown, with various shades of blue interspersed. The overall composition is clean and modern, with a white background on the left.

Exploring R Studio

The R Studio Environment



(source: <https://r4ds.had.co.nz/introduction.html>)

Exploring R Studio

Installing Packages:

```
install.packages("tidyverse")
```

```
library(tidyverse)
```

Packages in the tidyverse change fairly frequently. You can see if updates are available, and optionally install them, by running `tidyverse_update()`

Exploring R Studio

Installing Packages:

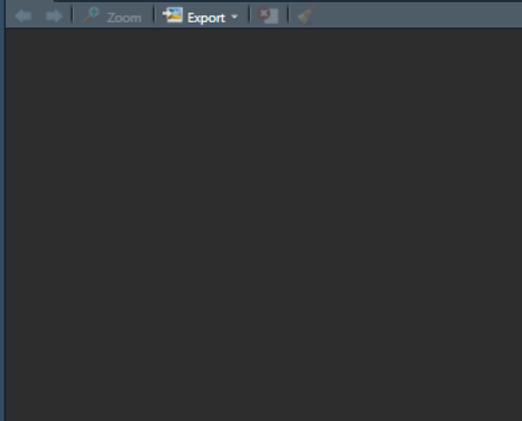
R C:/Users/Gina Carol/Desktop/Capstone Files and Documents/Shiny Stuff - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

```
Console Terminal Jobs
C:/Users/Gina Carol/Desktop/User Materials/
> myData$CupsOfCoffee
Error in myData$CupsOfCoffee : $ operator is invalid for atomic
vectors
>
> ggplot(data = myData) +
+ geom_point(mapping=aes(x = CupOfCoffee, y = WorkProductivity))
Error in ggplot(data = myData) : could not find function "ggplo
t"
> library(tidyverse)
-- Attaching packages ----- tidyvers
idyverse 1.3.0 --
v ggplot2 3.3.0    v purrr   0.3.4
v tibble  3.0.1    v dplyr   1.0.0
v tidyr   1.1.0    v stringr 1.4.0
v readr   1.3.1    v forcats 0.5.0
-- conflicts ----- tidyvers
e_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
> |
```

Plots



Environment History Files Connections Packages Help Viewer

Import Dataset

Global Environment

Data

CoffeePlo... 5 obs. of 3 variables

Values

myData "CoffePlot_csv"

Exploring R Studio

Coding Basics:

You can use R as a calculator:

```
>  
> 1/200 * 30  
[1] 0.15  
> (59 + 73 + 2 ) /3  
[1] 44.66667  
> |
```

(source: <https://r4ds.had.co.nz/workflow-basics.html>)

Exploring R Studio

Coding Basics:

You can create new objects with <-

```
x <- 3 * 4
```

All R statements where you create objects, **assignment** statements, have the same form:

```
object_name <- value
```

When reading that code say “object name gets value” in your head.

(source: <https://r4ds.had.co.nz/workflow-basics.html>)

Exploring R Studio

Coding Basics - Names:

Object names must start with a letter, and can only contain letters, numbers, `_` and `.`

You want your object names to be descriptive, so you'll need a convention for multiple words.

`i_use_snake_case` (separate lowercase words with `_`)

`otherPeopleUseCamelCase`

`some.people.use.periods`

`And_aFew.People_RENOUNCEconvention`

Exploring R Studio

Coding Basics - Names:

Object names must start with a letter, and can only contain letters, numbers, `_` and `.`

You want your object names to be descriptive, so you'll need a convention for multiple words.

`i_use_snake_case` (separate lowercase words with `_`)

`otherPeopleUseCamelCase`

`some.people.use.periods`

`And_aFew.People_RENOUNCEconvention`

Exploring R Studio

Coding Basics - Names:

Hash marks indicate notes in the code! It ignores this text.

```
> view(CoffeePlot2)
>
> # I always forget how to do this
> |
```

Exploring R Studio

Coding Basics - Names:

Quotation marks and parentheses must always come in a pair. RStudio does its best to help you, but it's still possible to mess up and end up with a mismatch. If this happens, R will show you the continuation character +

The + tells you that R is waiting for more input; it doesn't think you're done yet. Usually that means you've forgotten either a “ or a).

Either add the missing pair, or press ESCAPE to abort the expression and try again.

(source: <https://r4ds.had.co.nz/workflow-basics.html>)

Resources

All of the major R Studio Books are free, and online. You can buy a hardcopy, but they're there for you to use.

R for Data Science: <http://r4ds.had.co.nz/>

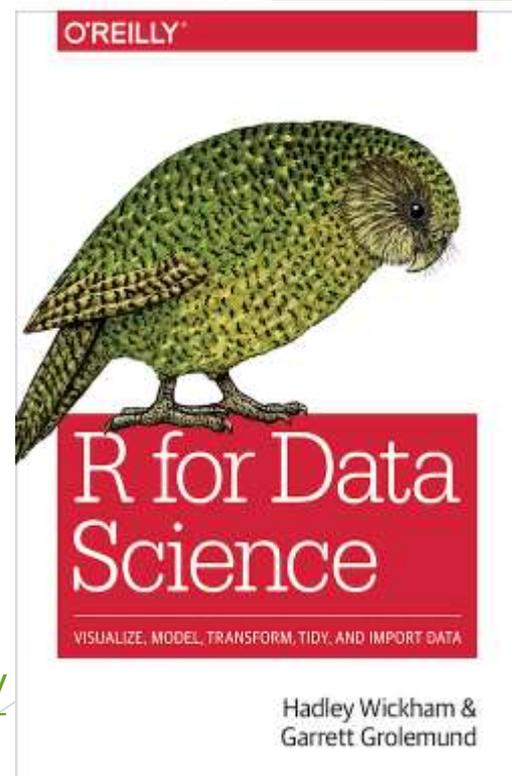
Introduction to Data Science:
<https://leanpub.com/datasciencebook>

Advanced R: <http://adv-r.had.co.nz/>

ggplot2: <https://github.com/hadley/ggplot2-book>

R Markdown – The Definitive Guide:
<https://bookdown.org/yihui/rmarkdown/>

Text Mining With R: <https://www.tidytextmining.com/>



Resources

R Conferences



useR! 2020
The R User Conference
July 7-10, 2020
St. Louis, MO, USA



Resources

Ways to Learn R



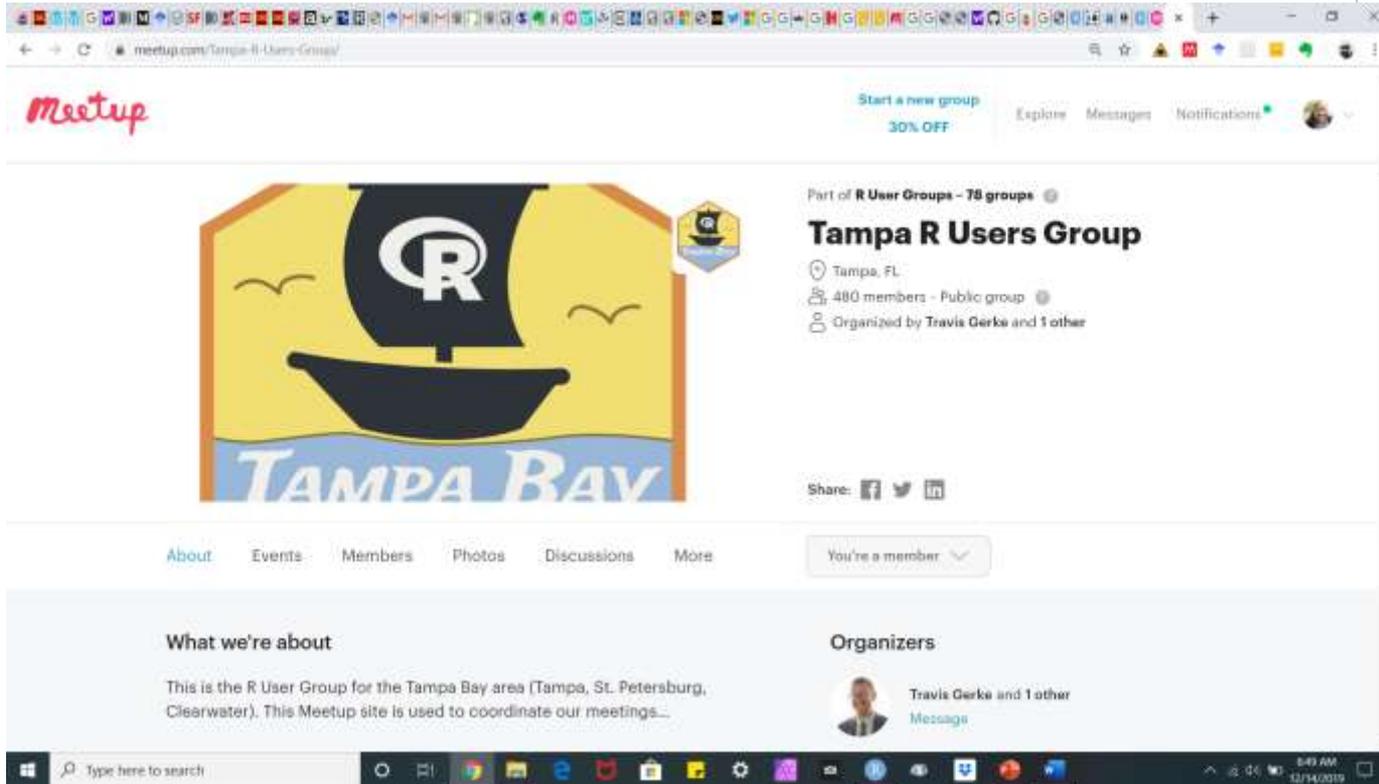
<https://www.rfordatasci.com/>



<https://rladies.org/>

Resources

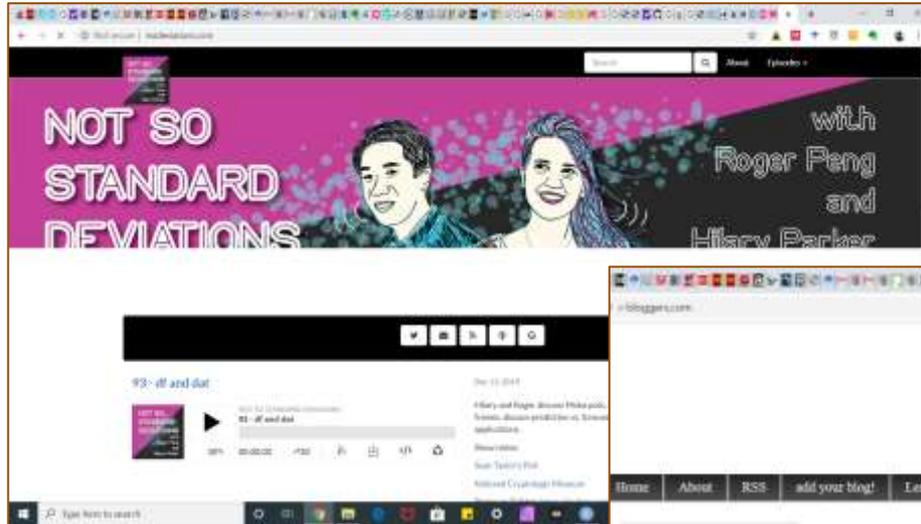
Ways to Learn R



The screenshot shows a web browser window displaying the Meetup.com page for the Tampa R Users Group. The browser's address bar shows the URL `meetup.com/Tampa-R-Users-Group/`. The Meetup logo is in the top left corner. The main content area features a large graphic of a sailboat with a white 'R' on its black sail, set against a yellow background with a blue banner at the bottom that reads 'TAMPA BAY'. To the right of the graphic, the group name 'Tampa R Users Group' is displayed, along with details: 'Part of R User Groups - 78 groups', 'Tampa, FL', '480 members - Public group', and 'Organized by Travis Gerke and 1 other'. Below the graphic, there are navigation tabs for 'About', 'Events', 'Members', 'Photos', 'Discussions', and 'More'. A 'What we're about' section contains the text: 'This is the R User Group for the Tampa Bay area (Tampa, St. Petersburg, Clearwater). This Meetup site is used to coordinate our meetings...'. An 'Organizers' section lists 'Travis Gerke and 1 other' with a 'Message' button. The Windows taskbar at the bottom shows the search bar and various application icons, with the system clock indicating 9:49 AM on 12/14/2019.

Resources

<http://nssdeviations.com/>



<https://www.r-bloggers.com/>

Resources

Data Transformation Cheat Sheet
dplyr provides a grammar for manipulating tables in R. This cheat sheet will guide you through the grammar, reminding you how to select, filter, arrange, mutate, summarise, group, and join data frames and tibbles. (Previous version) Updated January 17.

DOWNLOAD

Sparklyr Cheat Sheet
Sparklyr provides an R interface to Apache Spark, a fast and general engine for processing Big Data. With sparklyr, you can connect to a local or remote Spark session, use dplyr to manipulate data in Spark, and run Spark's built-in machine learning algorithms. Updated January 17.

DOWNLOAD

Functions Cheat Sheet
purrr makes it easy to work with lists and functions. This cheat sheet will remind you how to manipulate lists with purrr as well as how to apply functions iteratively to each element of a list or vector. The back of the book explains how to work with list-columns. With list columns, you can use a simple data frame to organize any collection of objects in R. Updated September 17.

DOWNLOAD

Import Cheat Sheet
tidyimport cheat sheet reminds you how to read in flat files with readr (tidyverse.org), work with the results as tibbles, and reshape messy data with tidyr. Use tidyr to reshape your tables into tidy data, the data format that works the most seamlessly with R and the tidyverse. Updated January 17.

DOWNLOAD

<https://rstudio.com/resources/cheatsheets/>

Resources

UseR! Videos

The screenshot shows the YouTube channel page for the R Consortium. The channel name is "R Consortium" with 4.83K subscribers. The page displays a list of videos under the "Uploads" tab, specifically "Lightning Talks" from the UseR! 2019 Toulouse conference. The videos are:

- Lightning Talk 3 - Spatial & ... (367 views, 3 months ago)
- Lightning Talk 4 - Forecasting - Eran Raviv (646 views, 3 months ago)
- Lightning Talk 7 - Methods ... (303 views, 3 months ago)
- Lightning Talk 8 - Models & ... (342 views, 3 months ago)
- Lightning Talk 6 - Shiny & ... (310 views, 3 months ago)

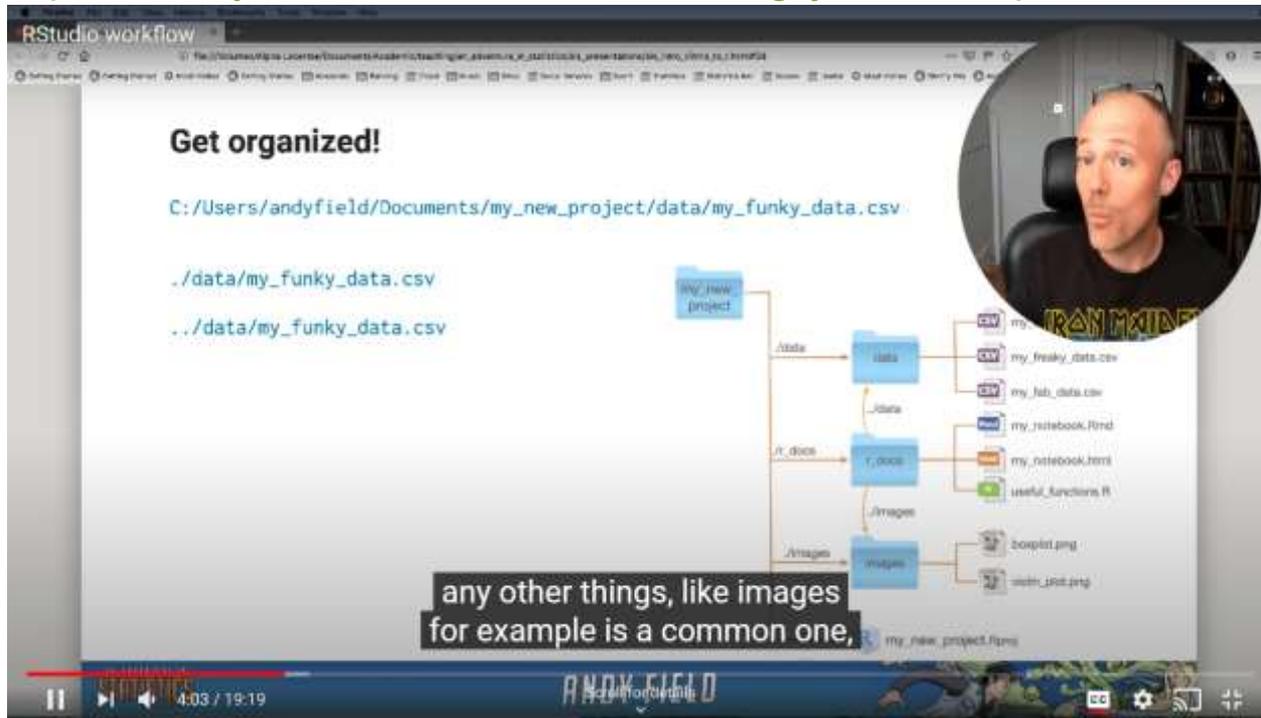
The page also features a navigation menu on the left with options like Home, Trending, Subscriptions, Library, and History. The top of the page includes a search bar with the text "user 2019 conference" and a "SIGN IN" button. The R Consortium logo and a "SUBSCRIBE" button are prominently displayed.

Resources

Andy Field

YouTube Channel:

<https://www.youtube.com/channel/UCakigkjm3vBzEHpFzECDXQQ>



The screenshot shows an RStudio workflow window with the title "Get organized!". The main content area displays the file path `C:/Users/andyfield/Documents/my_new_project/data/my_funky_data.csv` and three terminal commands: `./data/my_funky_data.csv`, `../data/my_funky_data.csv`, and `my_new_project`. To the right, a circular video feed shows a man with a beard and glasses, wearing a black t-shirt with "IRON MIA" on it. Below the main content, a file tree diagram illustrates the directory structure:

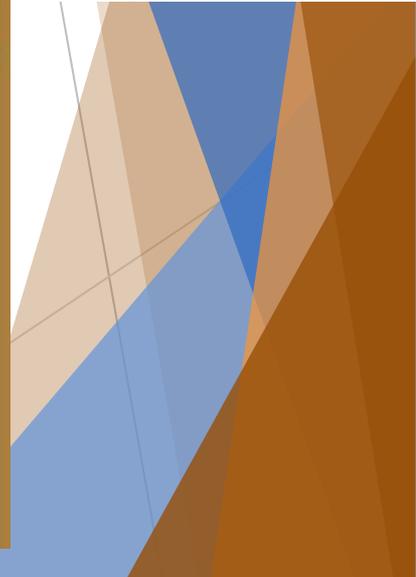
- my_new_project
 - data
 - my_funky_data.csv
 - my_funky_data.csv
 - my_fab_data.csv
 - it_does
 - my_notebook.Rmd
 - my_notebook.html
 - useful_functions.R
 - images
 - boxplot.png
 - with_plot.png

A text box at the bottom of the screenshot reads: "any other things, like images for example is a common one,". The video player interface at the bottom shows a progress bar at 4:03 / 19:19 and the channel name "ANDY FIELD".

Resources

Adventures in Social Work Research!

<https://aiswr78645859.wpcostaging.com/>



Thank you!

Contact:

aiswr-swrk@aiswr-swresearch.org

twitter: @TheGinaGi

