

RStudio Essentials

Managing Change 2: Version Control



git



Garrett Grolemond

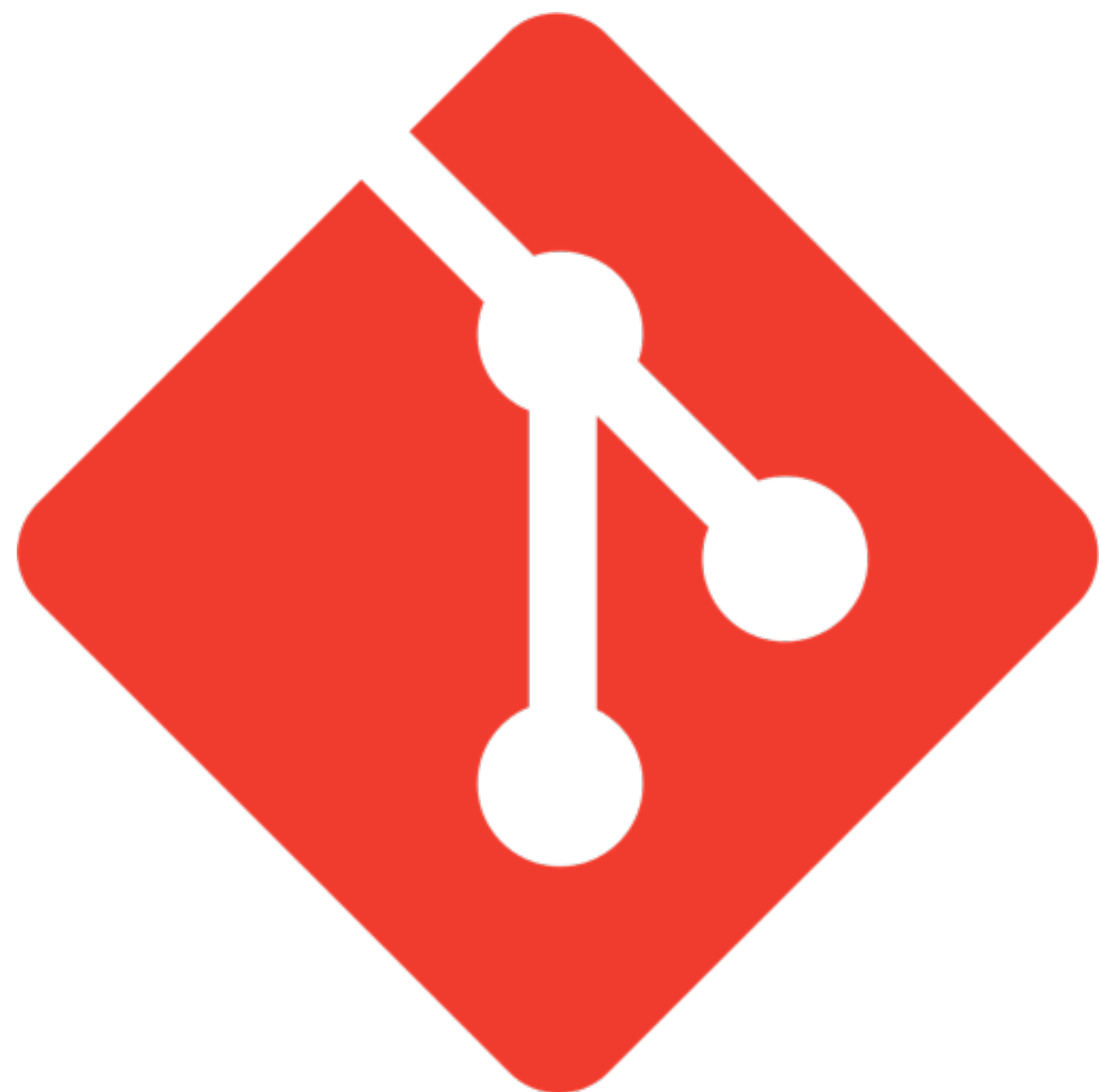
Data Scientist and Master Instructor
January 2016
Email: garrett@rstudio.com

Series

1. Writing Code (12/2)
2. Debugging Code (12/9)
3. Writing Packages (12/16)
4. Efficiency with Projects (1/6)
- 5. Version control, collaboration with Git (1/20)**
6. Reproducibility with Packrat (2/3)

[**www.rstudio.com/resources/webinars/**](http://www.rstudio.com/resources/webinars/)

**What is
version control
and why should
you use it?**



git

History

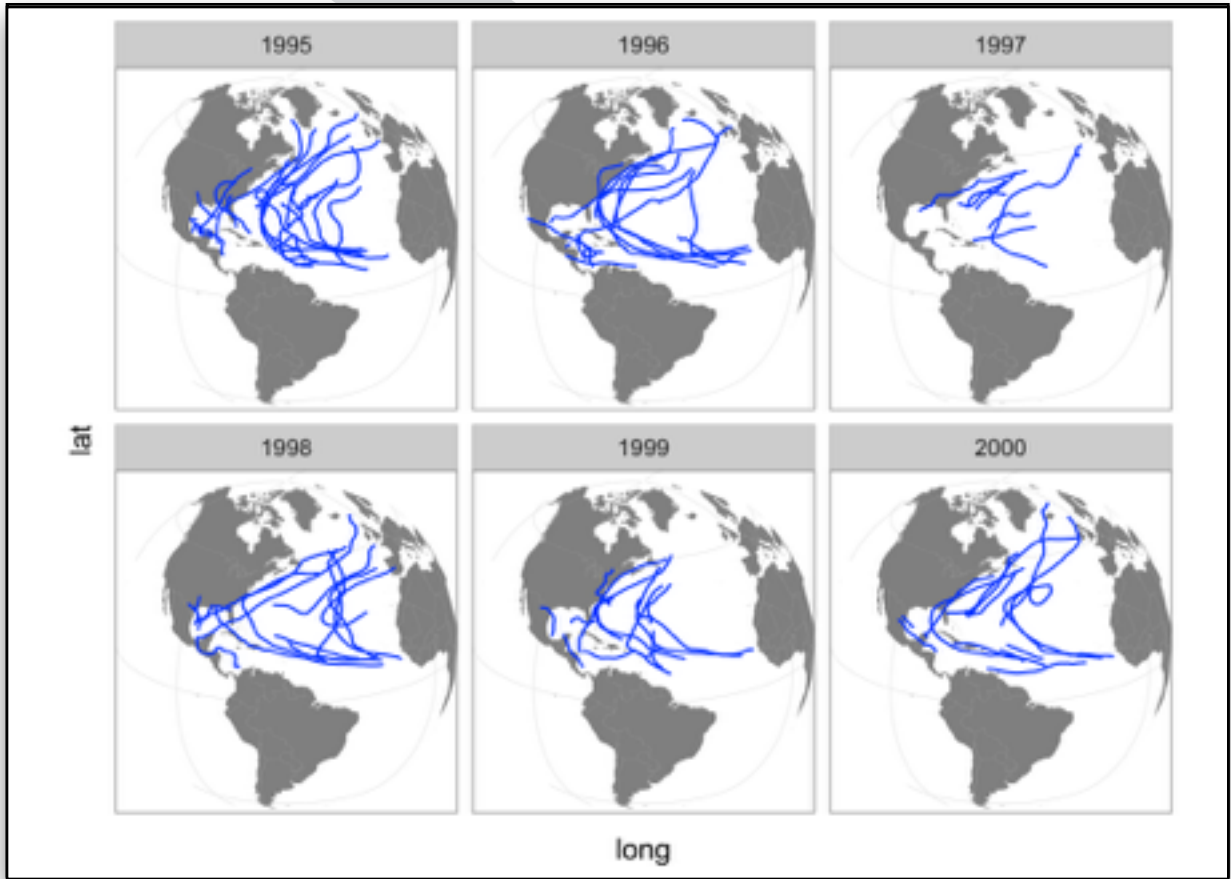
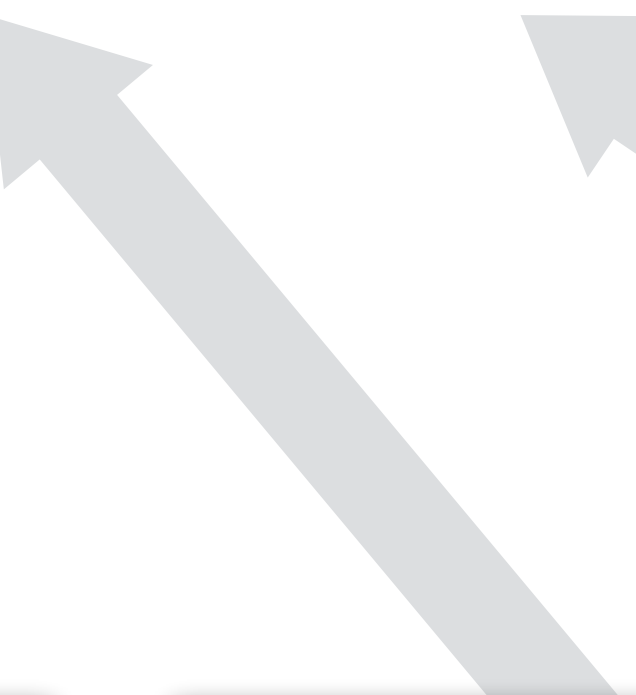
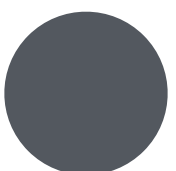
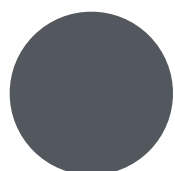
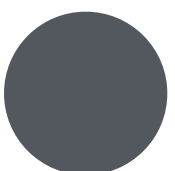
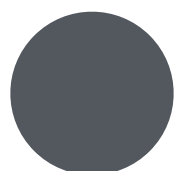
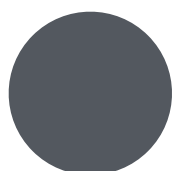
1st Commit

2nd Commit

3rd Commit

4th Commit

5th Commit



```
"name","year","time","lat","long"
"Allison",1995,1995-06-03 00:00:00,17.4,-84.3
"Allison",1995,1995-06-03 06:00:00,18.3,-84.9
"Allison",1995,1995-06-03 12:00:00,19.3,-85.7
"Allison",1995,1995-06-03 18:00:00,20.6,-85.8
"Allison",1995,1995-06-04 00:00:00,22,-86
"Allison",1995,1995-06-04 06:00:00,23.3,-86.3
"Allison",1995,1995-06-04 12:00:00,24.7,-86.2
"Allison",1995,1995-06-04 18:00:00,26.2,-86.2
"Allison",1995,1995-06-05 00:00:00,27.6,-86.1
"Allison",1995,1995-06-05 06:00:00,28.5,-85.6
"Allison",1995,1995-06-05 12:00:00,29.6,-84.7
"Allison",1995,1995-06-05 18:00:00,30.7,-83.8
"Allison",1995,1995-06-06 00:00:00,31.8,-82.8
"Allison",1995,1995-06-06 06:00:00,32.7,-81.5
"Allison",1995,1995-06-06 12:00:00,33.6,-80
"Allison",1995,1995-06-06 18:00:00,34.5,-78.1
"Allison",1995,1995-06-07 00:00:00,35.6,-75.9
"Allison",1995,1995-06-07 06:00:00,37.1,-73.6
"Allison",1995,1995-06-07 12:00:00,38.5,-71
"Allison",1995,1995-06-07 18:00:00,39.8,-69.2
"Allison",1995,1995-06-08 00:00:00,41,-67.7
"Allison",1995,1995-06-08 06:00:00,42.4,-66
"Allison",1995,1995-06-08 12:00:00,43.8,-63.7
```

```
# 0-Clean.R

library(dplyr)
library(lubridate)

storms <- read.csv("storms.csv")

storms <- storms %>%
  mutate(time = ymd_h(paste(year, month, day,
    hour))) %>%
  select(name, year, time, lat, long,
    pressure, wind, type)

write.csv(storms, file = "storms.csv",
  row.names = FALSE)
```

```
# 1-Plot.R

library(ggplot2)
library(dplyr)

map <- map_data("world") %>%
  filter(region != "USSR")

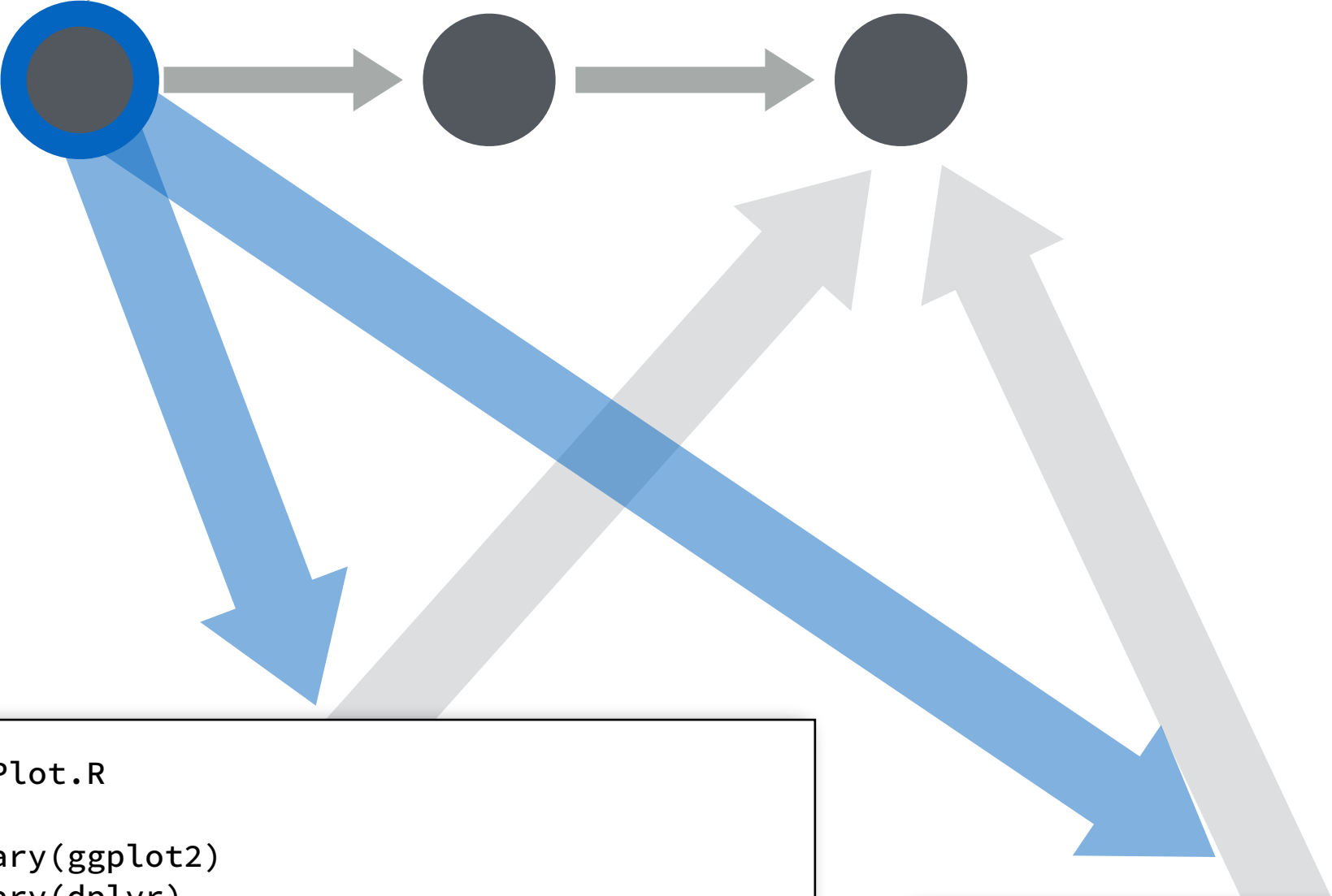
ggplot(storms, aes(x = long, y = lat)) +
  geom_polygon(aes(group = group),
    fill = "grey50", data = map) +
  geom_path(aes(group = name),
    color = "blue") +
  facet_wrap(~ year) +
  theme_bw() +
  coord_map(projection = "ortho",
    orientation = c(21, -60, 0))

ggsave("storms.png", width = 7, height = 5)
```

Project

History

1st Commit 2nd Commit 3rd Commit 4th Commit 5th Commit 6th Commit



```
"name","year","time","lat","long"
"Allison",1995,1995-06-03 00:00:00,17.4,-84.3
"Allison",1995,1995-06-03 06:00:00,18.3,-84.9
"Allison",1995,1995-06-03 12:00:00,19.3,-85.7
"Allison",1995,1995-06-03 18:00:00,20.6,-85.8
"Allison",1995,1995-06-04 00:00:00,22,-86
"Allison",1995,1995-06-04 06:00:00,23.3,-86.3
"Allison",1995,1995-06-04 12:00:00,24.7,-86.2
"Allison",1995,1995-06-04 18:00:00,26.2,-86.2
"Allison",1995,1995-06-05 00:00:00,27.6,-86.1
"Allison",1995,1995-06-05 06:00:00,28.5,-85.6
"Allison",1995,1995-06-05 12:00:00,29.6,-84.7
"Allison",1995,1995-06-05 18:00:00,30.7,-83.8
"Allison",1995,1995-06-06 00:00:00,31.8,-82.8
"Allison",1995,1995-06-06 06:00:00,32.7,-81.5
```

```
# 0-Clean.R

library(dplyr)
library(lubridate)

storms <- read.csv("storms.csv")

storms <- storms %>%
  mutate(time = ymd_h(paste(year, month, day,
                             hour))) %>%
  select(name, year, time, lat, long,
         pressure, wind, type)

write.csv(storms, file = "storms.csv",
          row.names = FALSE)
```

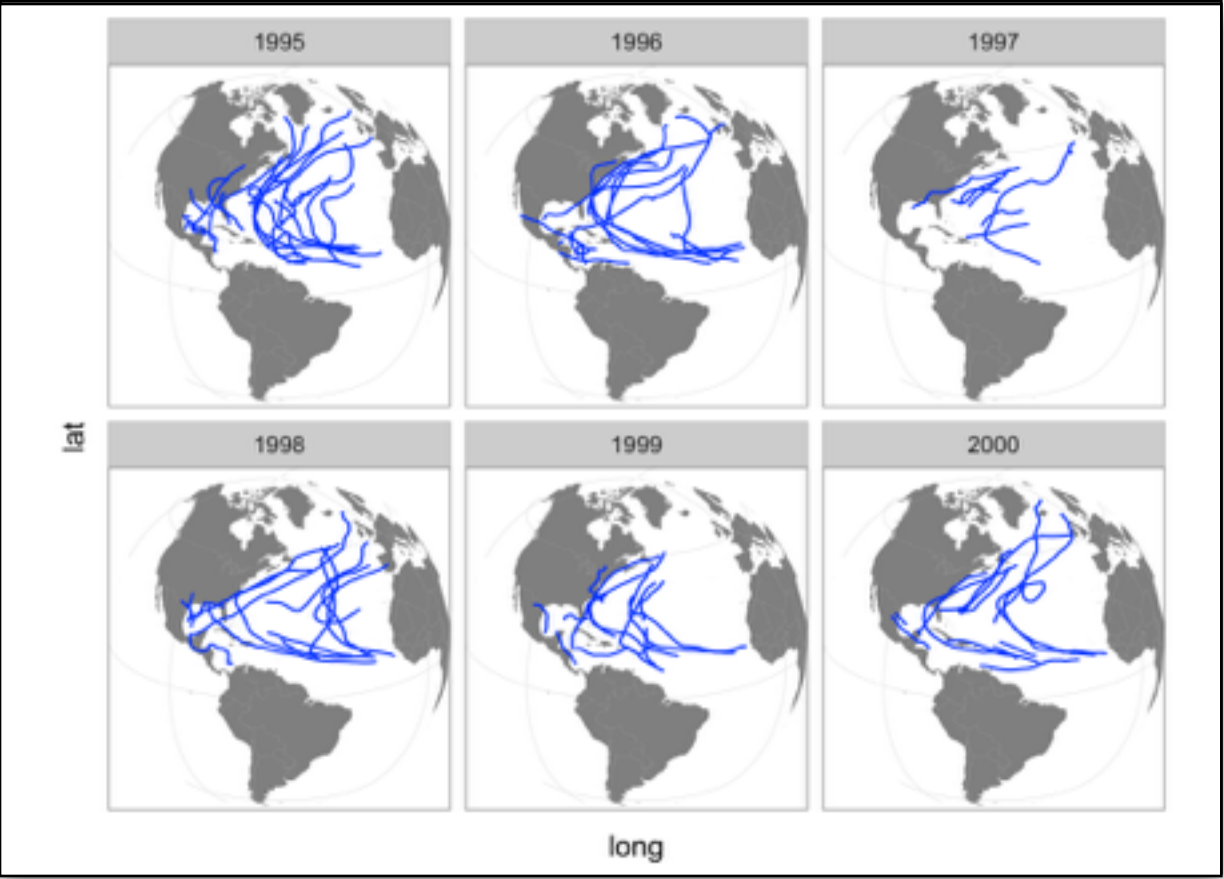
```
# 1-Plot.R

library(ggplot2)
library(dplyr)

map <- map_data("world") %>%
  filter(region != "USSR")

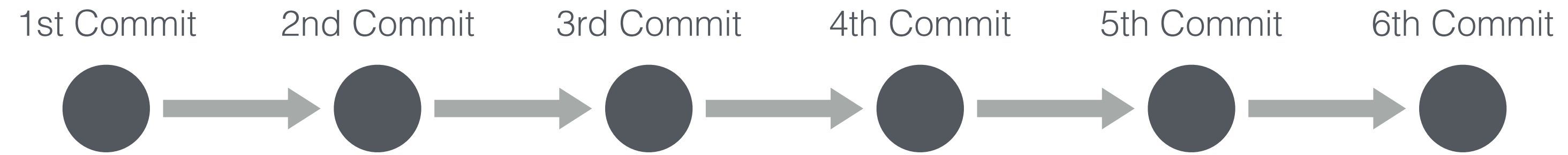
ggplot(storms, aes(x = long, y = lat)) +
  geom_polygon(aes(group = group),
              fill = "grey50", data = map) +
  geom_path(aes(group = name),
            color = "blue") +
  facet_wrap(~ year) +
  theme_bw() +
  coord_map(projection = "ortho",
            orientation = c(21, -60, 0))

ggsave("storms.png", width = 7, height = 5)
```



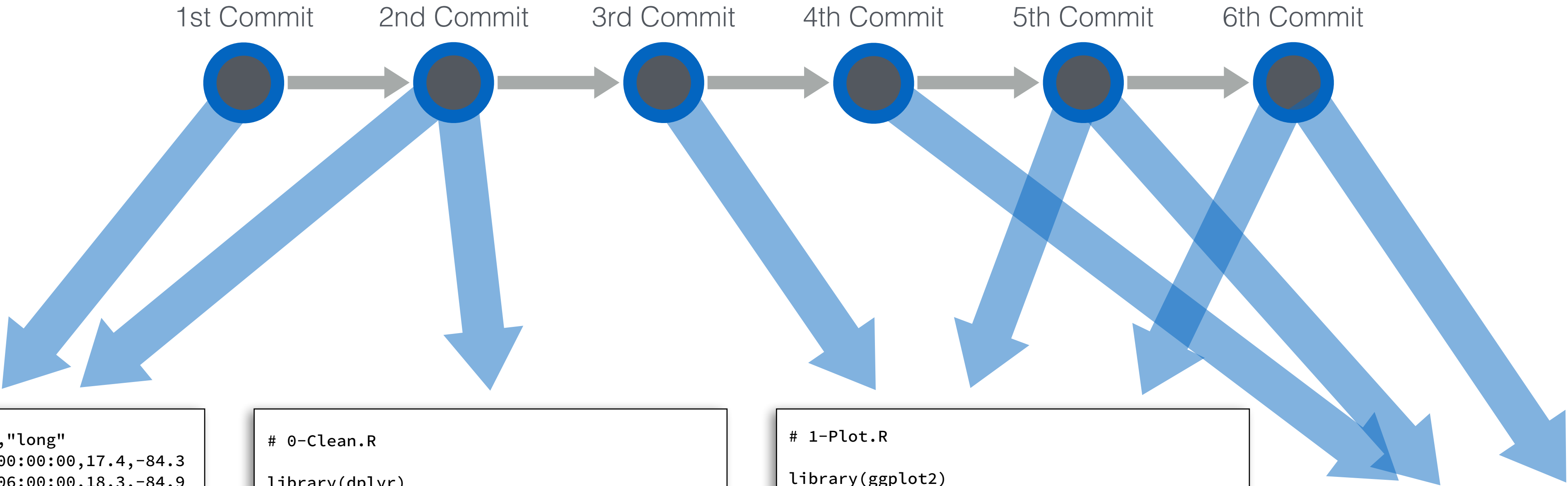
Project

History



.git

History



```
"name","year","time","lat","long"
"Allison",1995,1995-06-03 00:00:00,17.4,-84.3
"Allison",1995,1995-06-03 06:00:00,18.3,-84.9
"Allison",1995,1995-06-03 12:00:00,19.3,-85.7
"Allison",1995,1995-06-03 18:00:00,20.6,-85.8
"Allison",1995,1995-06-04 00:00:00,22,-86
"Allison",1995,1995-06-04 06:00:00,23.3,-86.3
"Allison",1995,1995-06-04 12:00:00,24.7,-86.2
"Allison",1995,1995-06-04 18:00:00,26.2,-86.2
"Allison",1995,1995-06-05 00:00:00,27.6,-86.1
"Allison",1995,1995-06-05 06:00:00,28.5,-85.6
"Allison",1995,1995-06-05 12:00:00,29.6,-84.7
"Allison",1995,1995-06-05 18:00:00,30.7,-83.8
"Allison",1995,1995-06-06 00:00:00,31.8,-82.8
"Allison",1995,1995-06-06 06:00:00,32.7,-81.5
"Allison",1995,1995-06-06 12:00:00,33.6,-80
"Allison",1995,1995-06-06 18:00:00,34.5,-78.1
"Allison",1995,1995-06-07 00:00:00,35.6,-75.9
"Allison",1995,1995-06-07 06:00:00,37.1,-73.6
"Allison",1995,1995-06-07 12:00:00,38.5,-71
"Allison",1995,1995-06-07 18:00:00,39.8,-69.2
"Allison",1995,1995-06-08 00:00:00,41,-67.7
"Allison",1995,1995-06-08 06:00:00,42.4,-66
"Allison",1995,1995-06-08 12:00:00,43.8,-63.7
```

```
# 0-Clean.R

library(dplyr)
library(lubridate)

storms <- read.csv("storms.csv")

storms <- storms %>%
  mutate(time = ymd_h(paste(year, month, day,
                             hour)))) %>%
  select(name, year, time, lat, long,
         pressure, wind, type)

write.csv(storms, file = "storms.csv",
          row.names = FALSE)
```

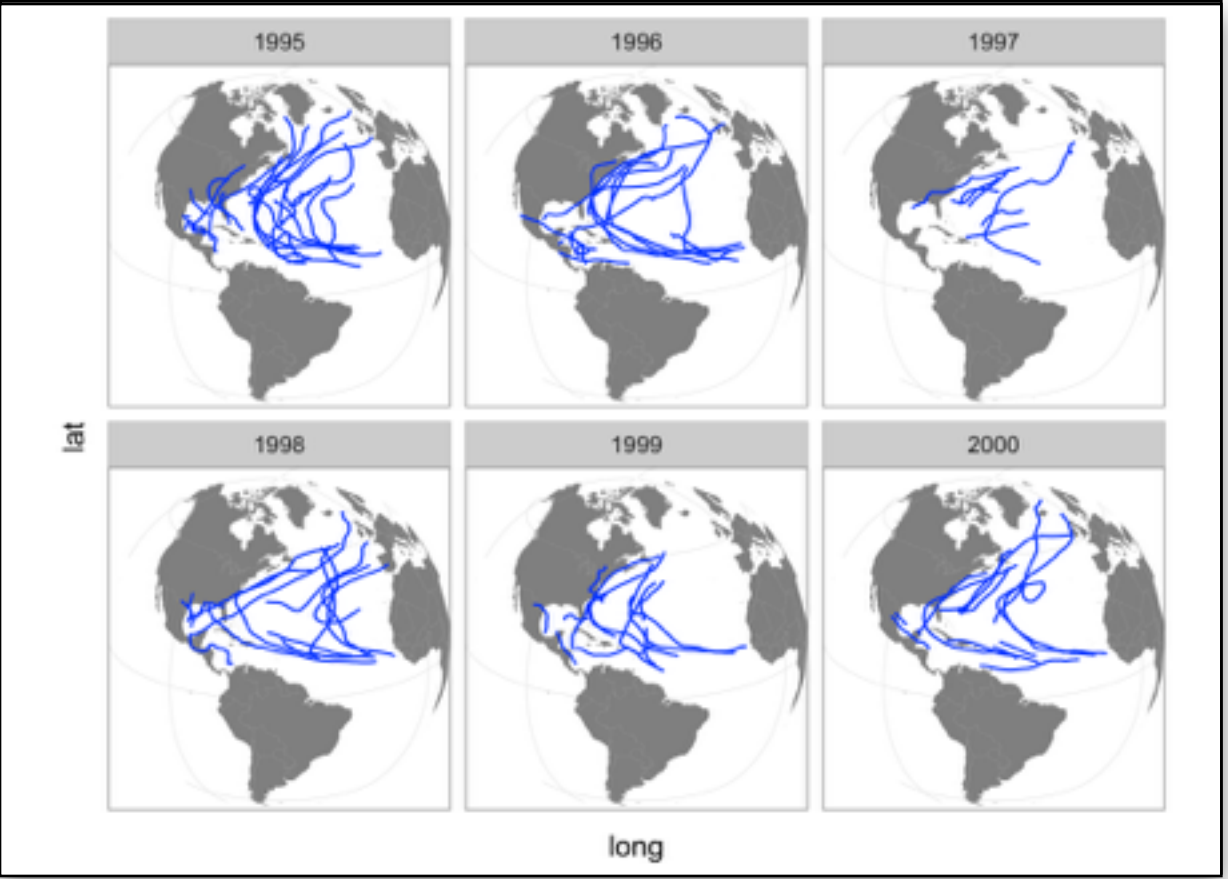
```
# 1-Plot.R

library(ggplot2)
library(dplyr)

map <- map_data("world") %>%
  filter(region != "USSR")

ggplot(storms, aes(x = long, y = lat)) +
  geom_polygon(aes(group = group),
              fill = "grey50", data = map) +
  geom_path(aes(group = name),
            color = "blue") +
  facet_wrap(~ year) +
  theme_bw() +
  coord_map(projection = "ortho",
            orientation = c(21, -60, 0))

ggsave("storms.png", width = 7, height = 5)
```



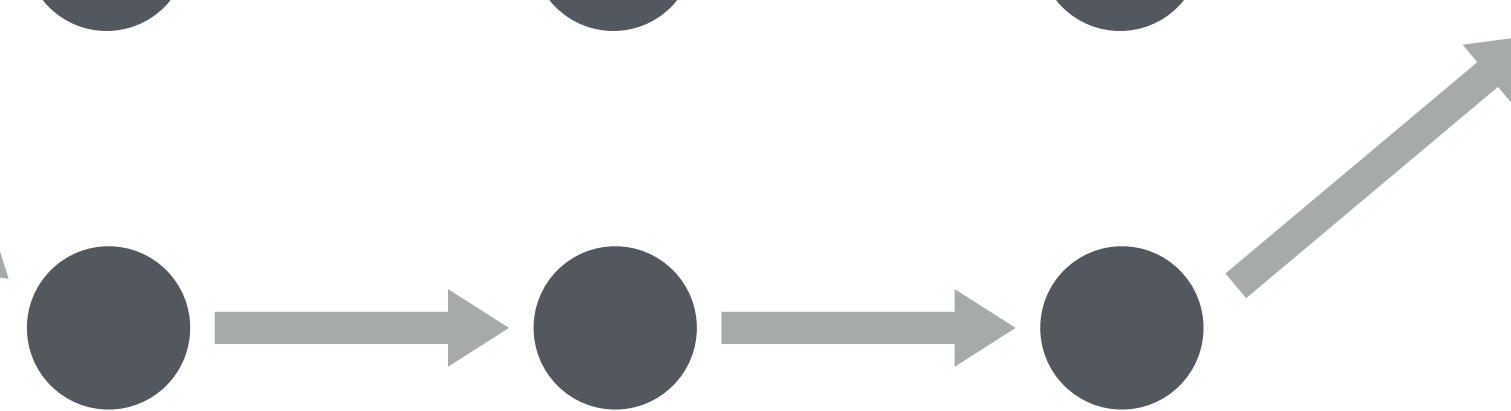
Project

History

Branch 1



Branch 2



"Official" Version

implied by commit history

```
"name","year","time","lat","long"
"Allison",1995,1995-06-03 00:00:00,17.4,-84.3
"Allison",1995,1995-06-03 06:00:00,18.3,-84.9
"Allison",1995,1995-06-03 12:00:00,19.3,-85.7
"Allison",1995,1995-06-03 18:00:00,20.6,-85.8
"Allison",1995,1995-06-04 00:00:00,22,-86
"Allison",1995,1995-06-04 06:00:00,23.3,-86.3
"Allison",1995,1995-06-04 12:00:00,24.7,-86.2
"Allison",1995,1995-06-04 18:00:00,26.2,-86.2
"Allison",1995,1995-06-05 00:00:00,27.6,-86.1
"Allison",1995,1995-06-05 06:00:00,28.5,-85.6
"Allison",1995,1995-06-05 12:00:00,29.6,-84.7
"Allison",1995,1995-06-05 18:00:00,30.7,-83.8
"Allison",1995,1995-06-06 00:00:00,31.8,-82.8
"Allison",1995,1995-06-06 06:00:00,32.7,-81.5
"Allison",1995,1995-06-06 12:00:00,33.6,-80
"Allison",1995,1995-06-06 18:00:00,34.5,-78.1
"Allison",1995,1995-06-07 00:00:00,35.6,-75.9
"Allison",1995,1995-06-07 06:00:00,37.1,-73.6
"Allison",1995,1995-06-07 12:00:00,38.5,-71
"Allison",1995,1995-06-07 18:00:00,39.8,-69.2
"Allison",1995,1995-06-08 00:00:00,41,-67.7
"Allison",1995,1995-06-08 06:00:00,42.4,-66
"Allison",1995,1995-06-08 12:00:00,43.8,-63.7
```

```
# 0-Clean.R

library(dplyr)
library(lubridate)

storms <- read.csv("storms.csv")

storms <- storms %>%
  mutate(time = ymd_h(paste(year,
month, day,
hour))) %>%
  select(name, year, time, lat,
long, pressure, wind, type)

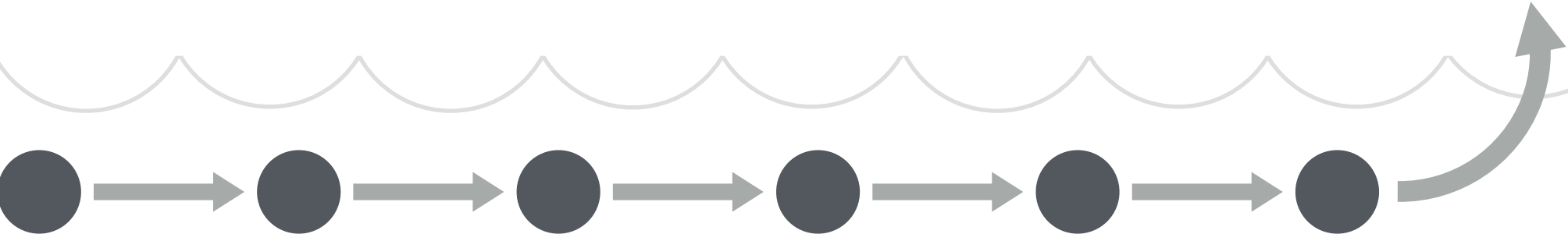
write.csv(storms, file =
"storms.csv",
row.names = FALSE)
```

```
# 1-Plot.R

library(ggplot2)
map <- map_data("world") %>%
  filter(region != "USSR")

ggplot(storms, aes(x = long, y =
lat)) +
  geom_polygon(aes(group = group),
fill = "grey50", data = map) +
  geom_path(aes(group = name),
color = "black") +
  facet_wrap(~ year) +
  theme_bw() +
  coord_map(projection = "ortho",
orientation = c(21, -60, 0))

ggsave("storms.png", width = 7,
height = 5)
```



Real Life Version

uncommitted changes
in blue

```
"name","year","time","lat","long"
"Allison",1995,1995-06-03 00:00:00,17.4,-84.3
"Allison",1995,1995-06-03 06:00:00,18.3,-84.9
"Allison",1995,1995-06-03 12:00:00,19.3,-85.7
"Allison",1995,1995-06-03 18:00:00,20.6,-85.8
"Allison",1995,1995-06-04 00:00:00,22,-86
"Allison",1995,1995-06-04 06:00:00,23.3,-86.3
"Allison",1995,1995-06-04 12:00:00,24.7,-86.2
"Allison",1995,1995-06-04 18:00:00,26.2,-86.2
"Allison",1995,1995-06-05 00:00:00,27.6,-86.1
"Allison",1995,1995-06-05 06:00:00,28.5,-85.6
"Allison",1995,1995-06-05 12:00:00,29.6,-84.7
"Allison",1995,1995-06-05 18:00:00,30.7,-83.8
"Allison",1995,1995-06-06 00:00:00,31.8,-82.8
"Allison",1995,1995-06-06 06:00:00,32.7,-81.5
>Allison",1995,1995-06-06 12:00:00,33.6,-80
>Allison",1995,1995-06-06 18:00:00,34.5,-78.1
>Allison",1995,1995-06-07 00:00:00,35.6,-75.9
>Allison",1995,1995-06-07 06:00:00,37.1,-73.6
>Allison",1995,1995-06-07 12:00:00,38.5,-71
>Allison",1995,1995-06-07 18:00:00,39.8,-69.2
>Allison",1995,1995-06-08 00:00:00,41,-67.7
>Allison",1995,1995-06-08 06:00:00,42.4,-66
>Allison",1995,1995-06-08 12:00:00,43.8,-63.7
```

```
# 0-Clean.R

library(dplyr)
library(lubridate)

storms <- read.csv("storms.csv")

storms <- storms %>%
  mutate(time = ymd_h(paste(year,
month, day,
hour))) %>%
  select(name, year, time, lat,
long, pressure, wind, type)

write.csv(storms,
file = "storms.csv",
row.names = FALSE)
```

```
# 1-Plot.R

library(ggplot2)
map <- map_data("world") %>%
  filter(region != "USSR")

ggplot(storms, aes(x = long, y =
lat)) +
  geom_polygon(aes(group = group),
fill = "grey50", data = map) +
  geom_path(aes(group = name),
color = "blue") +
  facet_wrap(~ year) +
  theme_bw() +
  coord_map(projection = "ortho",
orientation = c(21, -60, 0))

ggsave("storms.png", width = 7,
height = 5)
```



GitHub

www.github.com

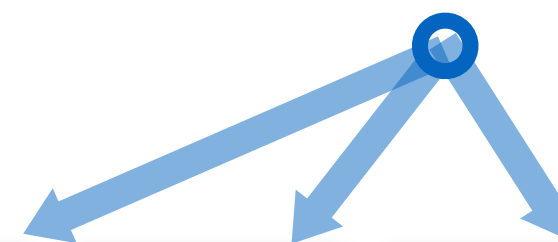


GitHub



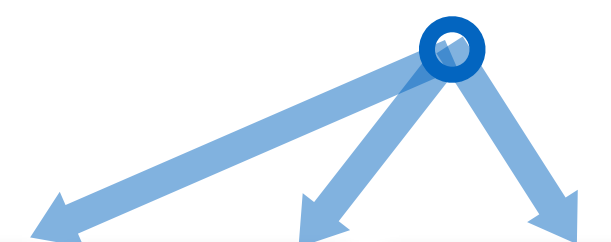
```
center <-  
function(x) {  
  x - mean(x)  
}  
  
scale <-  
function(x) {  
  x / sd(x)  
}  
  
standardize <-  
function(x) {  
  scale(center(x))  
}
```

You



```
center <-  
function(x) {  
  x - mean(x)  
}  
  
scale <-  
function(x) {  
  x / sd(x)  
}  
  
standardize <-  
function(x) {  
  scale(center(x))  
}
```

Collaborator 1



```
center <-  
function(x) {  
  x - mean(x)  
}  
  
scale <-  
function(x) {  
  x / sd(x)  
}  
  
standardize <-  
function(x) {  
  scale(center(x))  
}
```

Collaborator 2



GitHub



```
center <-  
function(x) {  
  x - mean(x)  
}
```

```
scale <-  
function(x) {  
  x / sd(x)  
}
```

```
standardize <-  
function(x) {  
  scale(center(x))  
}  
  
reverse <-  
function(x) {  
  rev(x)  
}
```

You



```
center <-  
function(x) {  
  x - mean(x)  
}
```

```
scale <-  
function(x) {  
  x / sd(x)  
}
```

```
standardize <-  
function(x) {  
  scale(center(x))  
}  
  
reverse <-  
function(x) {  
  rev(x)  
}
```

Collaborator 1



```
center <-  
function(x) {  
  x - mean(x)  
}
```

```
scale <-  
function(x) {  
  x / sd(x)  
}
```

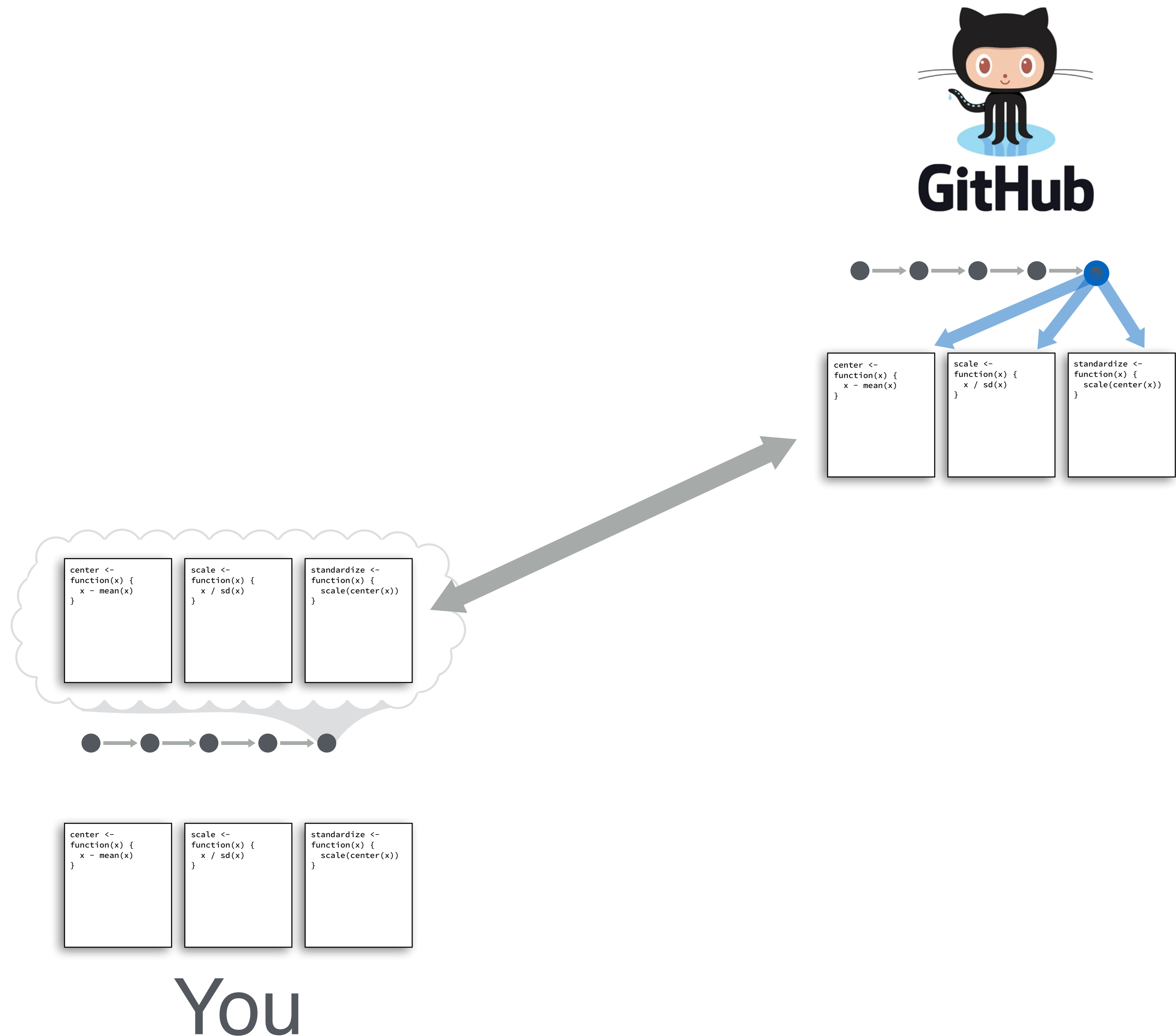
```
standardize <-  
function(x) {  
  scale(center(x))  
}  
  
reverse <-  
function(x) {  
  rev(x)  
}
```

Collaborator 2

1. Real Life Version
in your working directory

2. "Official" Version
implied by commit history

3. Github Version
available to public





3. Github Version

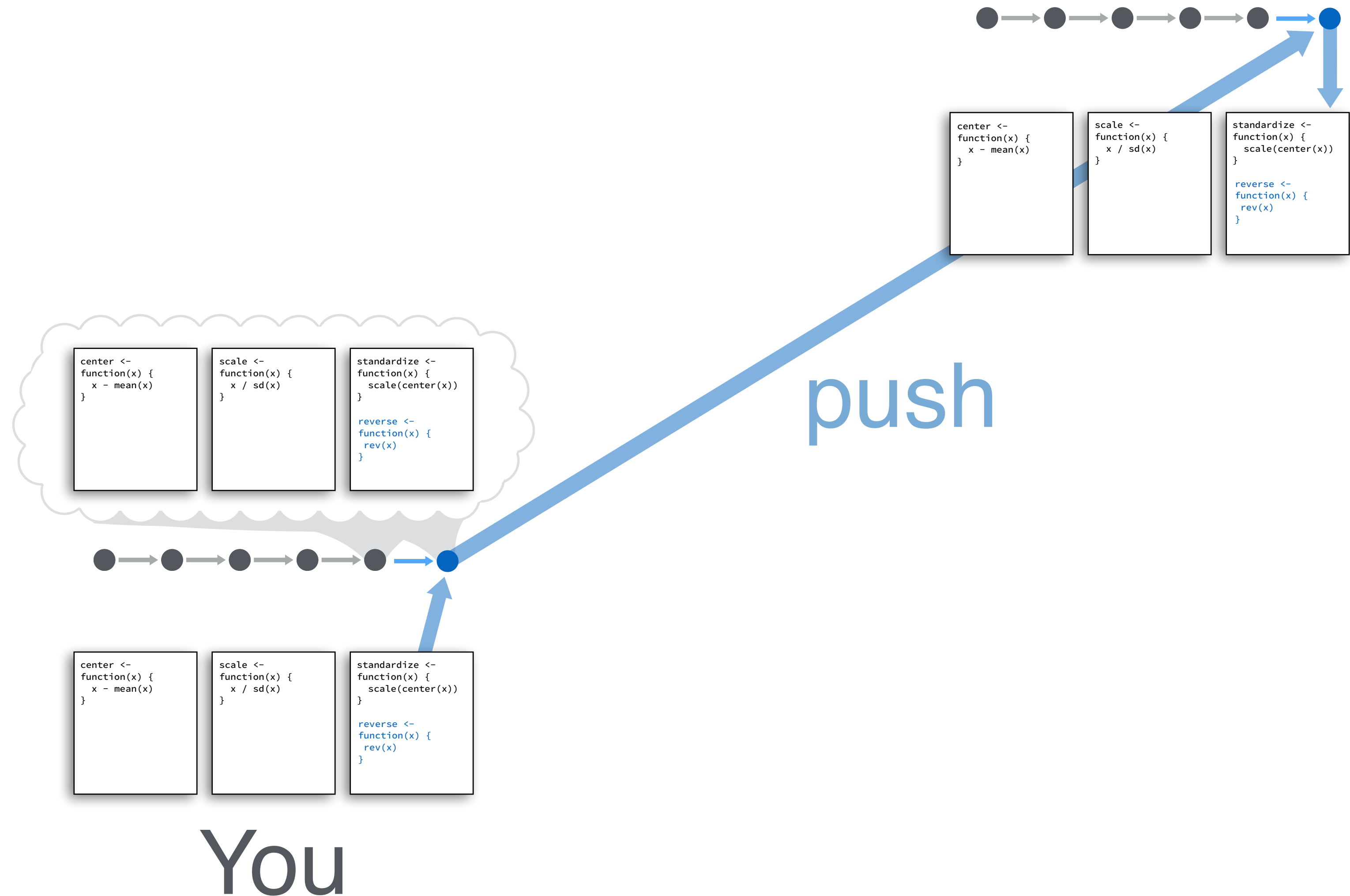
available to public

2. "Official" Version

implied by commit history

1. Real Life Version

in your working directory





3. Github Version

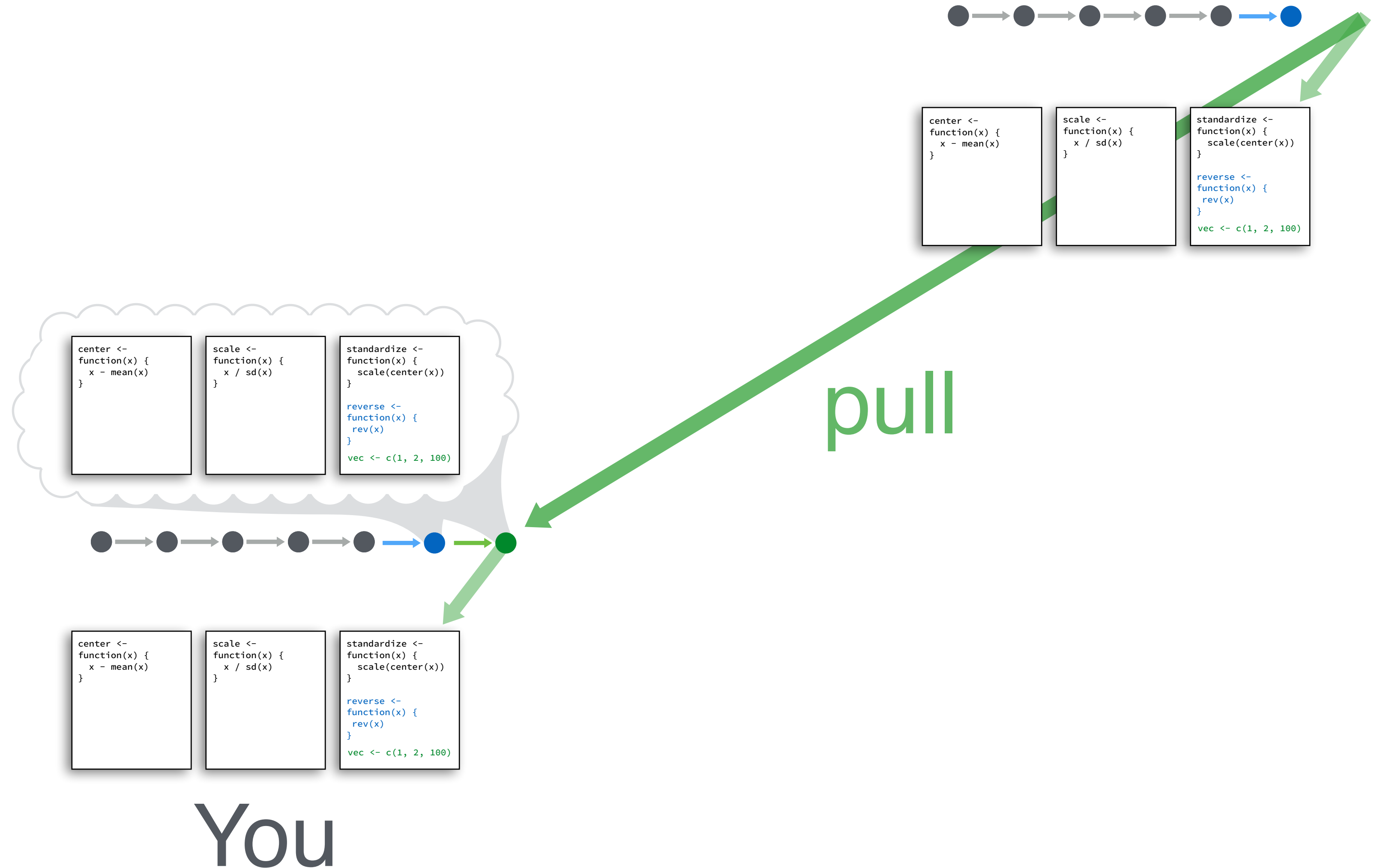
available to public

2. "Official" Version

implied by commit history

1. Real Life Version

in your working directory





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

GitUp

<http://gitup.co/>

RStudio IDE Cheatsheet


www.rstudio.com/resources/cheatsheets/

RStudio IDE Cheat Sheet

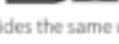
learn more at www.rstudio.com

RStudio


The RStudio IDE is an Integrated Development Environment in R that comes in three versions



Desktop IDE
A local version of the IDE for your desktop



Open Source Server
For larger compute resources and remote access



Professional Server
For teams that share large compute resources, large data, and uniform environments for collaboration

Download all at www.rstudio.com. Each provides the same useful interface:

Documents and Apps

- Check spelling
- Render output
- Choose output format
- Choose output location
- Open Shiny, R Markdown, knitr, Sweave, LaTeX, Rd files and more in Source Pane
- Jump to chunk
- Jump to next chunk
- Run selected lines
- Publish outline to server
- Show file outline
- Access Markdown guide at **Help > Markdown Quick Reference**
- Jump to chunk
- Set knitr all previous options
- Run this code chunk
- Run this code chunk

RStudio recognizes that files named **app.R**, **server.R**, **ui.R**, and **global.R** belong to a shiny app

Write Code

- Navigate tabs
- Open in new window
- Save
- Find and replace
- Compile as notebook
- Run selected code
- Multiple cursors/column selection with **Alt+mouse drag**
- Code diagnostics that appear in the margin
- Hover over diagnostic symbols for details
- Syntax highlighting based on your file's extension
- Tab completion to finish function names, file paths, arguments, and more
- Multi-language code snippets to quickly use common blocks of code
- Change file type
- Working Directory
- Press **Ctrl+H** to see command history
- Maximize, minimize panes
- Drag pane boundaries

R Support

- Import data file with wizard
- History of past commands to run/add to source
- Display RPR slideshows
- File > New File > R Presentation**
- Load workspace
- Save workspace
- Delete all saved objects
- Search inside environment
- Choose environment to display from list of parent environments
- Display objects as list or grid
- Displays saved objects by type with short description
- View in data viewer
- View function source code
- Create folder
- Upload file
- Delete file
- Rename file
- Path to displayed directory
- A File browser keyed to your working directory. Click on file or directory name to open.

Debug Mode

- Open with **debug()**, **browse()**, or a breakpoint. RStudio will open the debugger mode when it encounters a breakpoint while executing code.
- Launch debugger mode from origin of error
- Open traceback to examine the functions that R called before the error occurred
- Click next to line number to add/remove a breakpoint
- Highlighted line shows where execution has paused
- Run commands in environment where execution has paused
- Examine variables in executing environment
- Select function in traceback to debug
- Step through code one line at a time
- Step into and out of functions to run
- Resume execution mode
- Quit debug

Version Control with Git or SVN

- Turn on at **Tools > Project Options > Git/SVN**
- Stage files
- Show file diff
- Commit staged files
- Push/Pull staged files to remote
- View History
- Added
- Deleted
- Modified
- Renamed
- Untracked
- Open shell to type commands
- current branch

Package Writing

- File > New Project > New Directory > R Package**
- Turn project into package.
- Enable roxygen documentation with **Tools > Project Options > Build Tools**
- Roxygen guide at **Help > Roxygen Quick Reference**

RStudio Pro Features

- Share Project with Collaborators
- Active shared collaborators
- Start new R Session in current project
- Close R Session in project
- Select R Version
- Project System
- File > New Project**
- RStudio saves the call history, workspace, and working directory associated with a project. It reloads each when you re-open a project.

RStudio opens plots in a dedicated Plots pane

- Navigate
- Open in recent plots
- Export plot
- Delete plot
- Delete all plots

GUI Package manager lists every installed package

Install Update Create reproducible package Packages Packages library for your project

Click to load package with **library()**. Unlick to detach package with **detach()**

Delete installed package

RStudio opens documentation in a dedicated Help pane

- Home page of helpful links
- Search within help file
- Search for help file
- Viewer Pane displays HTML content, such as Shiny apps, RMarkdown reports, and interactive visualizations
- Stop Shiny app
- Publish to shinyapps.io, rpubs, RStudioConnect
- Refresh

View(<data>) opens spreadsheet like view of data set

Filter rows by value or value range

Sort by values

Search for value

1 LAYOUT

- Move focus to Source Editor
- Move focus to Console
- Move focus to Help
- Show History
- Show Files
- Show Plots
- Show Packages
- Show Environment
- Show Git/SVN
- Show Build

Windows/Linux	Mac
Ctrl+1	Ctrl+1
Ctrl+2	Ctrl+2
Ctrl+3	Ctrl+3
Ctrl+4	Ctrl+4
Ctrl+5	Ctrl+5
Ctrl+6	Ctrl+6
Ctrl+7	Ctrl+7
Ctrl+8	Ctrl+8
Ctrl+9	Ctrl+9
Ctrl+0	Ctrl+0

2 RUN CODE

- Search command history
- Navigate command history
- Move cursor to start of line
- Move cursor to end of line
- Change working directory
- Interrupt current command
- Clear console
- Quit Session (desktop only)
- Restart R Session
- Run current line/selection
- Run current (retain cursor)
- Run from current to end
- Run the current function
- Source a file
- Source the current file
- Source with echo

Windows/Linux	Mac
Ctrl+↑	Cmd+↑
Ctrl+↓	Cmd+↓
Home	Cmd+←
End	Cmd+→
Ctrl+Shift+H	Ctrl+Shift+H
Esc	Esc
Ctrl+L	Ctrl+L
Ctrl+Q	Cmd+Q
Ctrl+Shift+F10	Cmd+Shift+F10
Ctrl+Enter	Cmd+Enter
Alt+Enter	Option+Enter
Ctrl+Alt+E	Cmd+Option+E
Ctrl+Alt+F	Cmd+Option+F
Ctrl+Shift+O	Cmd+Shift+O
Ctrl+Shift+S	Cmd+Shift+S
Ctrl+Shift+Enter	Cmd+Shift+Enter

3 NAVIGATE CODE

- Go File/Function
- Fold Selected
- Unfold Selected
- Fold All
- Unfold All
- Go to line
- Jump to
- Switch to tab
- Previous tab
- Next tab
- First tab
- Last tab
- Navigate back
- Navigate forward
- Jump to Brace
- Select within Braces
- Use Selection for Find
- Find in Files
- Find Next
- Find Previous
- Jump to Word
- Jump to Start/End

Windows/Linux	Mac
Ctrl+.	Ctrl+.
Alt+L	Cmd+Option+L
Alt+O	Cmd+Option+O
Shift+Alt+O	Cmd+Shift+Option+O
Shift+Alt+G	Cmd+Shift+Option+G
Shift+Alt+J	Cmd+Shift+Option+J
Ctrl+Shift+.	Ctrl+Shift+.
Ctrl+F11	Ctrl+F11
Ctrl+F12	Ctrl+F12
Ctrl+Shift+F11	Ctrl+Shift+F11
Ctrl+Shift+F12	Ctrl+Shift+F12
Ctrl+F9	Cmd+F9
Ctrl+F10	Cmd+F10
Ctrl+P	Ctrl+P