RStudio

250 Northern Ave, Boston, MA 02210
Phone: 844-448-1212

Email: info@rstudio.com
Web: http://www.rstudio.com

RStudio Essentials

Managing Change 2: Version Control

Garrett Grolemund

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/

What is version control and why should you use it?


git

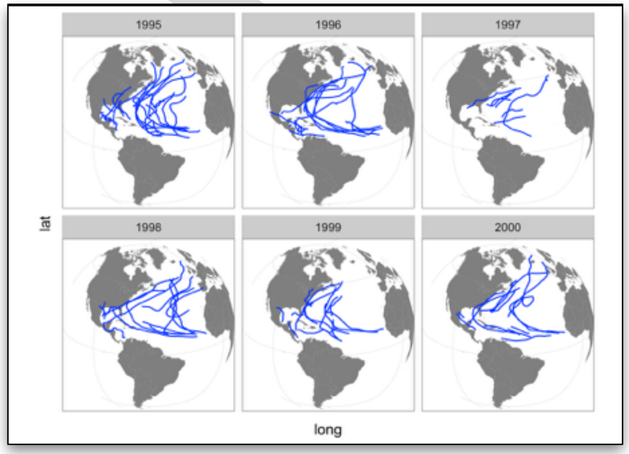
History

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

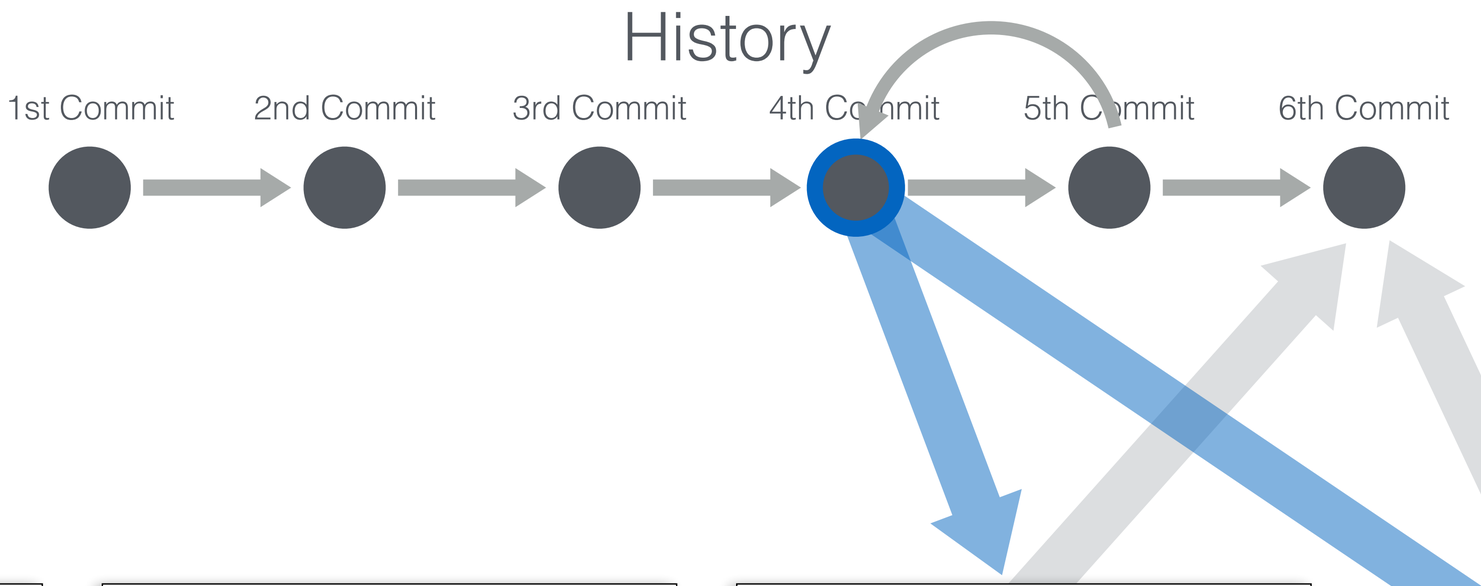
```
"name","year","time","lat","long"
"Allison",1995,1995-06-03 00:00:00,17.4,-84.3
"Allison",1995,1995-06-03 06:00:00,18.3,-84.9
"Allison",1995,1995-06-03 12:00:00,19.3,-85.7
"Allison",1995,1995-06-03 18:00:00,20.6,-85.8
"Allison",1995,1995-06-04 00:00:00,22,-86
"Allison",1995,1995-06-04 06:00:00,23.3,-86.3
"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)
```



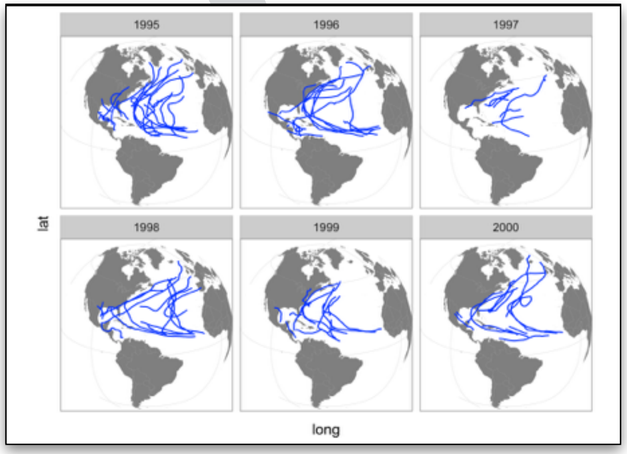
Project



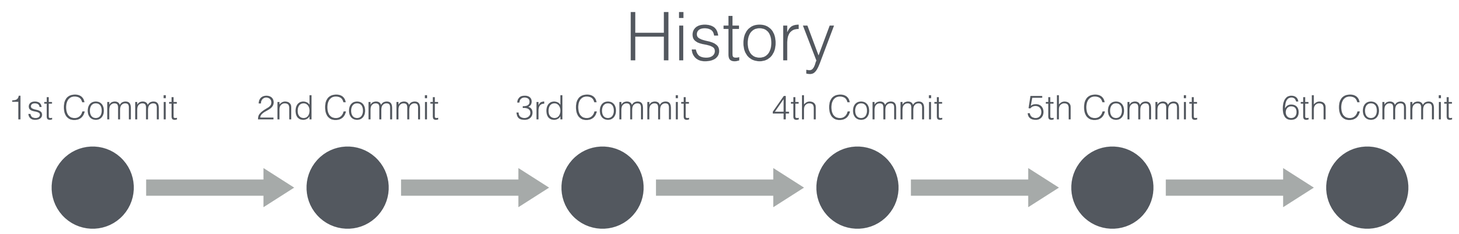

```
"name","year","time","lat","long"
"Allison",1995,1995-06-03 00:00:00,17.4,-84.3
"Allison",1995,1995-06-03 06:00:00,18.3,-84.9
"Allison",1995,1995-06-03 12:00:00,19.3,-85.7
"Allison",1995,1995-06-03 18:00:00,20.6,-85.8
"Allison",1995,1995-06-04 00:00:00,22,-86
"Allison",1995,1995-06-04 06:00:00,23.3,-86.3
"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)
```

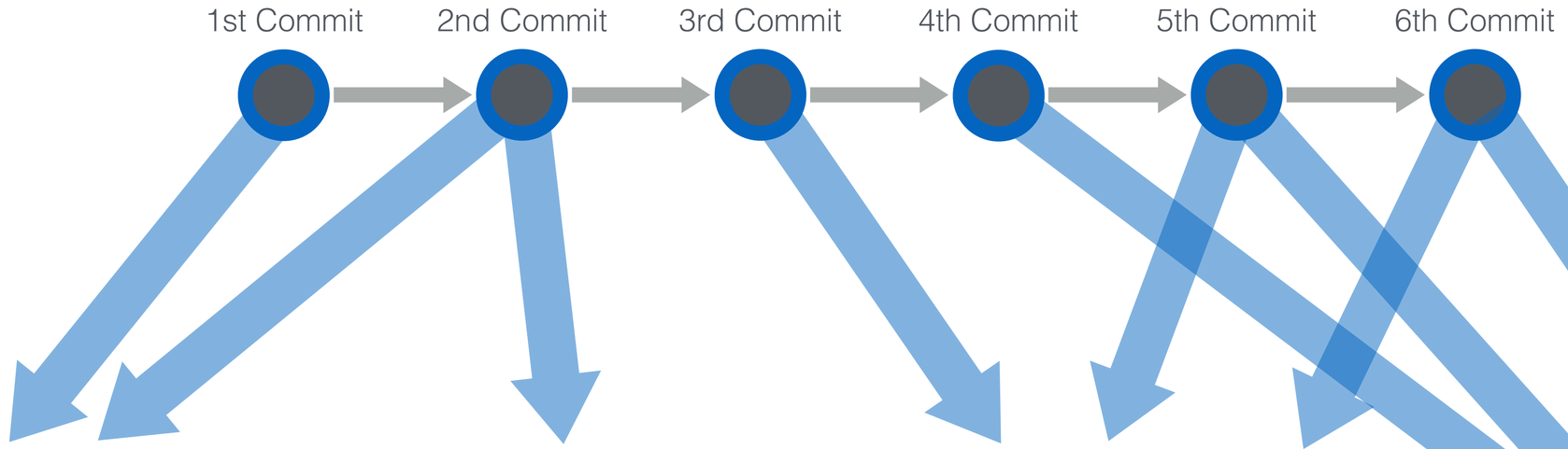


Project


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

.git

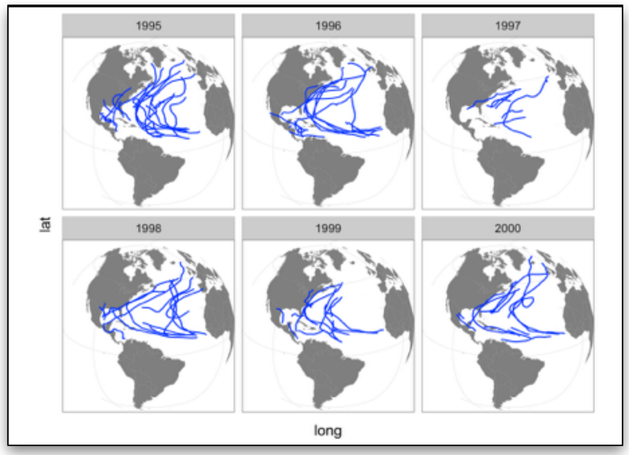
History




```
"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
```

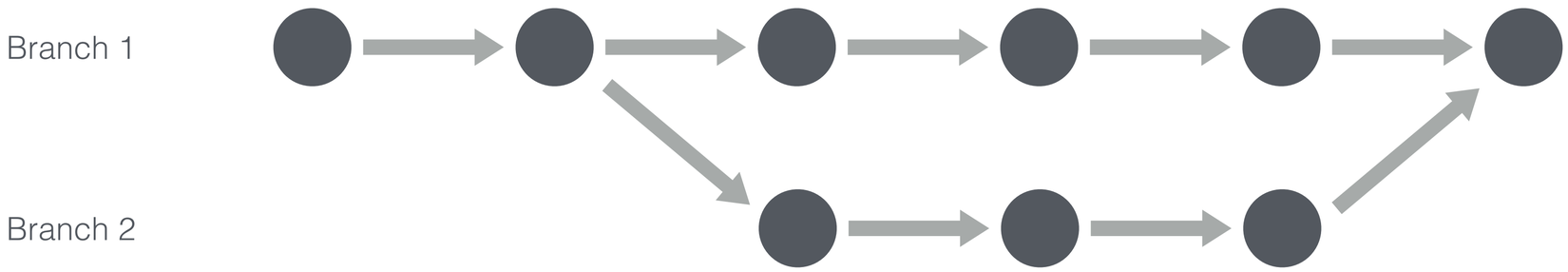
```
# 0-Clean.R

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



"Official" Version

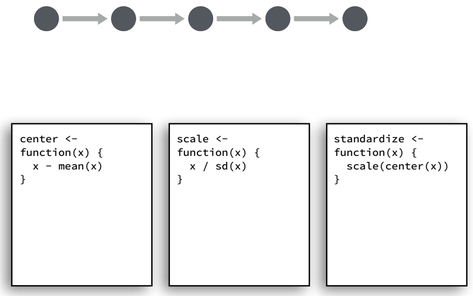
implied by commit history

Real Life Version

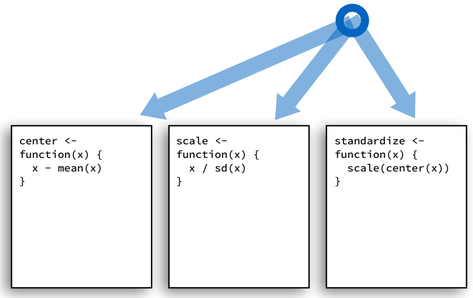
uncommitted changes in blue

GitHub

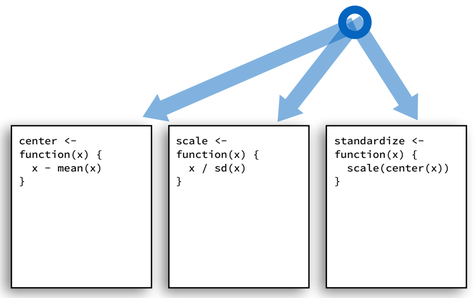
www.github.com



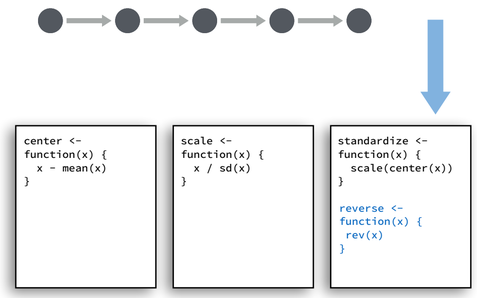
You



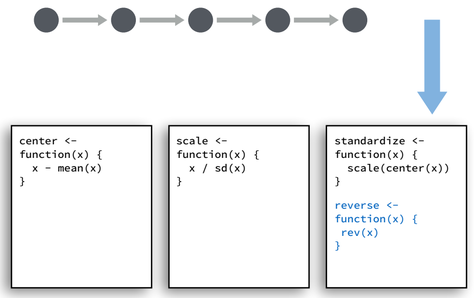
Collaborator 1



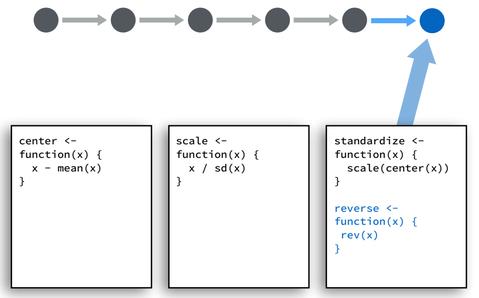
Collaborator 2



You



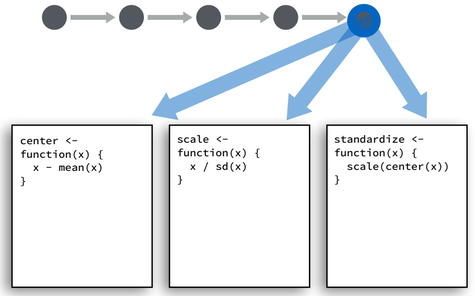
Collaborator 1



Collaborator 2

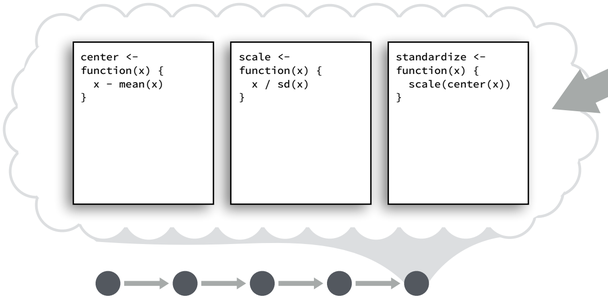
3. Github Version

available to public



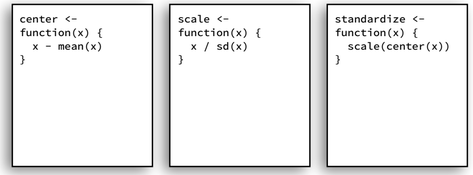
2."Official" Version

implied by commit history



1.Real Life Version

in your working directory



You

3. Github Version

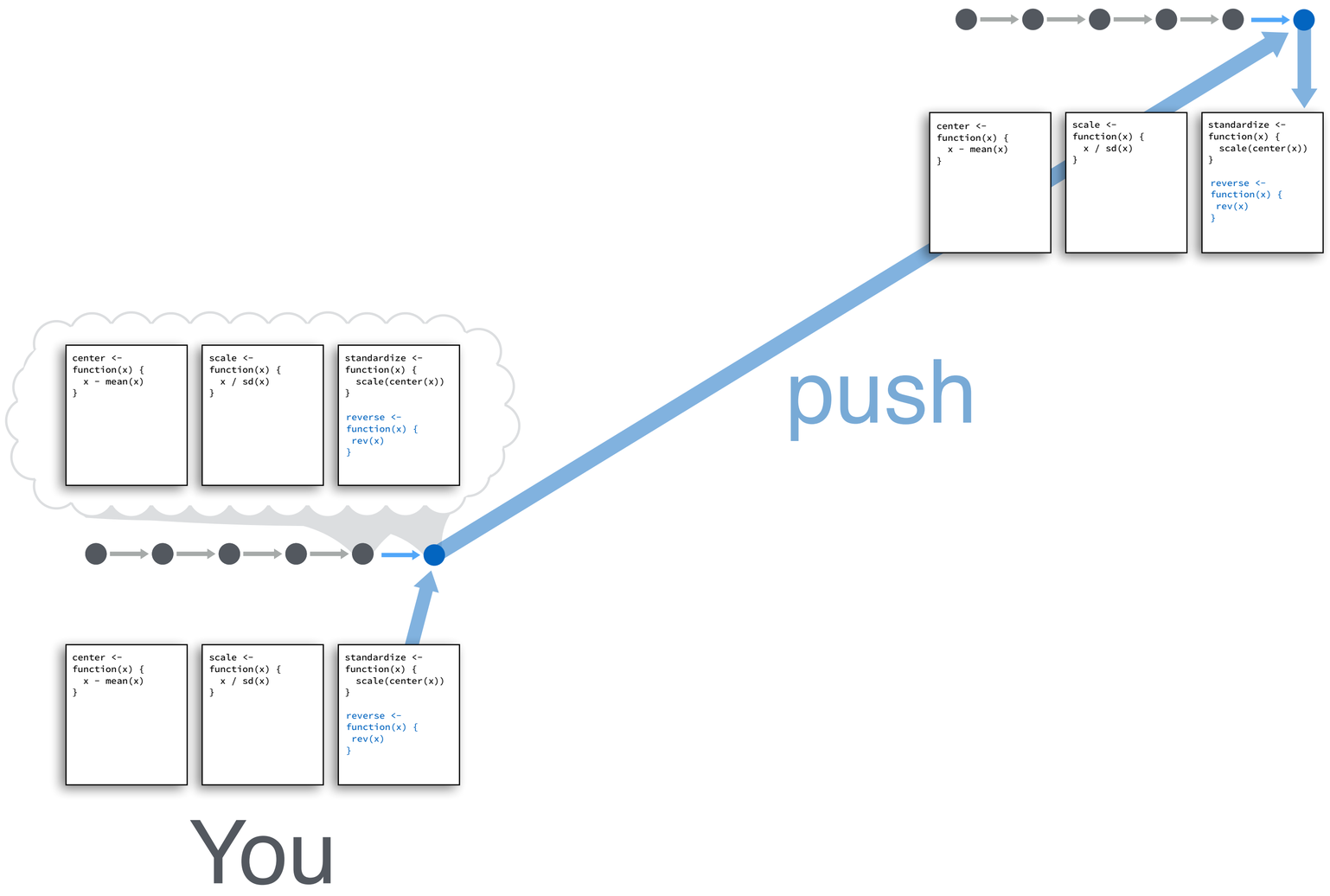
available to public

2."Official" Version

implied by commit history

1.Real Life Version

in your working directory



3. Github Version

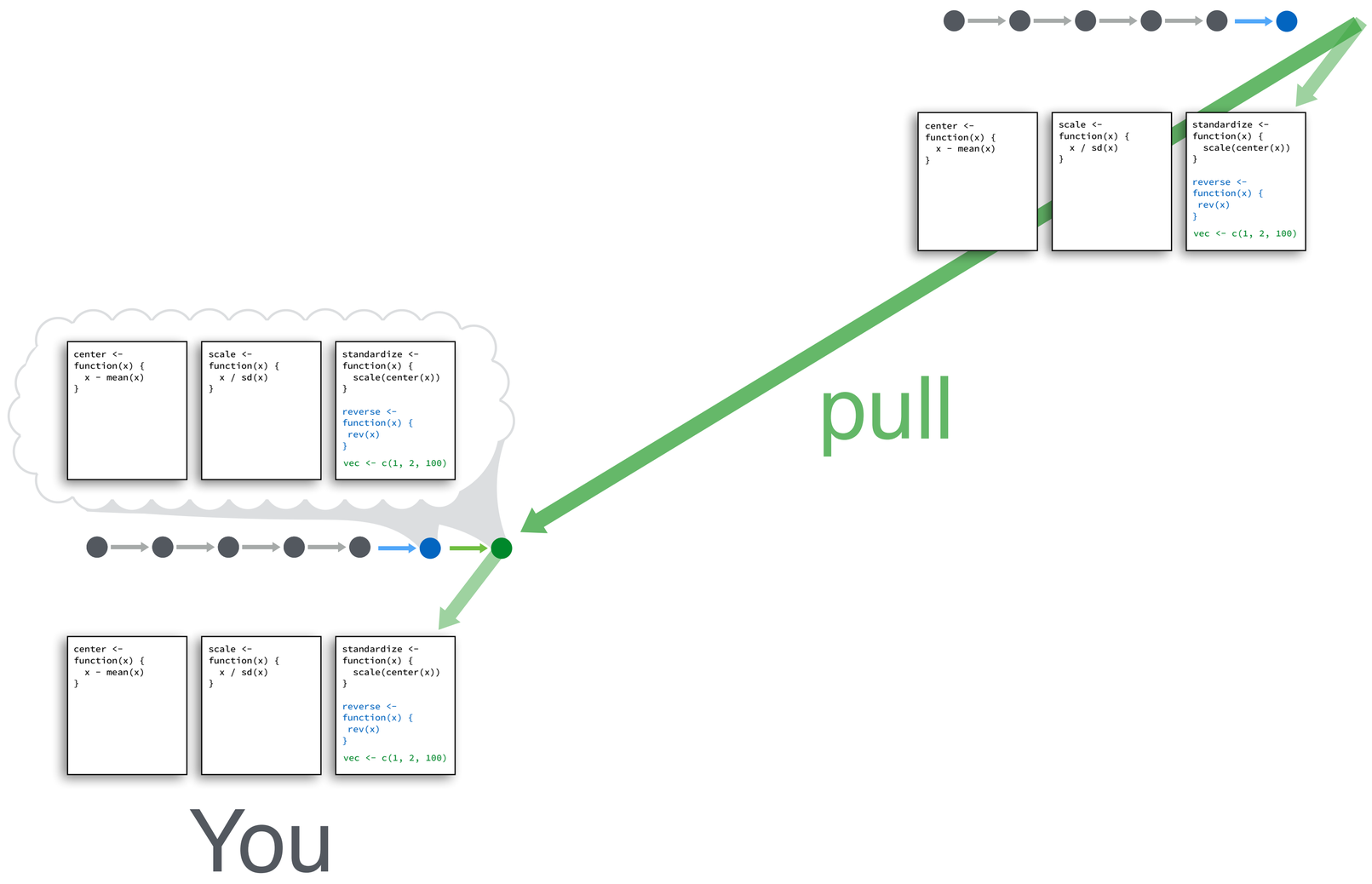
available to public

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/