rbtl - Research Beyond the Lab

Version control and more tools

Lars Schöbitz

2022-04-28

Today

1. Assignment grading and conversion
2. What is Version Control?
3. RStudio Cloud & RStudio IDE

Assignments & Grades

Due dates: Assignments 1, 2, and 3

Full points for assignment 1, 2, and 3 if submitted by:

15th March

Submission complete? Full points?

- If you haven't heard from me by Wednesday, then you have successfully submitted your assignment.

Grade conversion

grade	points
6.00	100
5.75	95
5.50	90
5.25	85
5.00	80
4.75	75
4.50	70
4.25	60
4.00	50
3.50	40
3.00	30
2.50	20
2.00	10
1.00	0

Grading scheme and conversion details in syllabus on course website

What is Version Control?

Version Control

A way to share files with others, so they can:

- download
- re-use
- contribute

You can view the history of files, and jump back in time to any point.

Why is it useful?

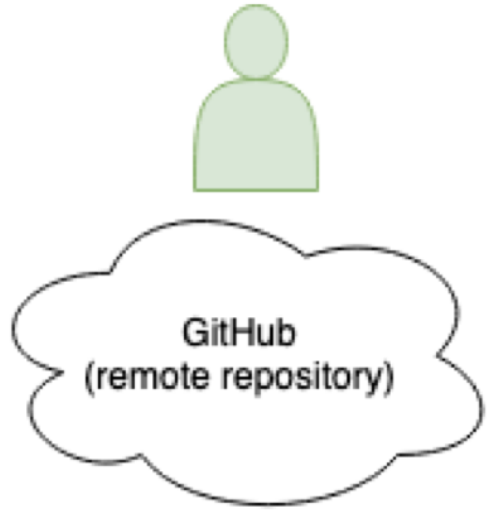
Git and GitHub

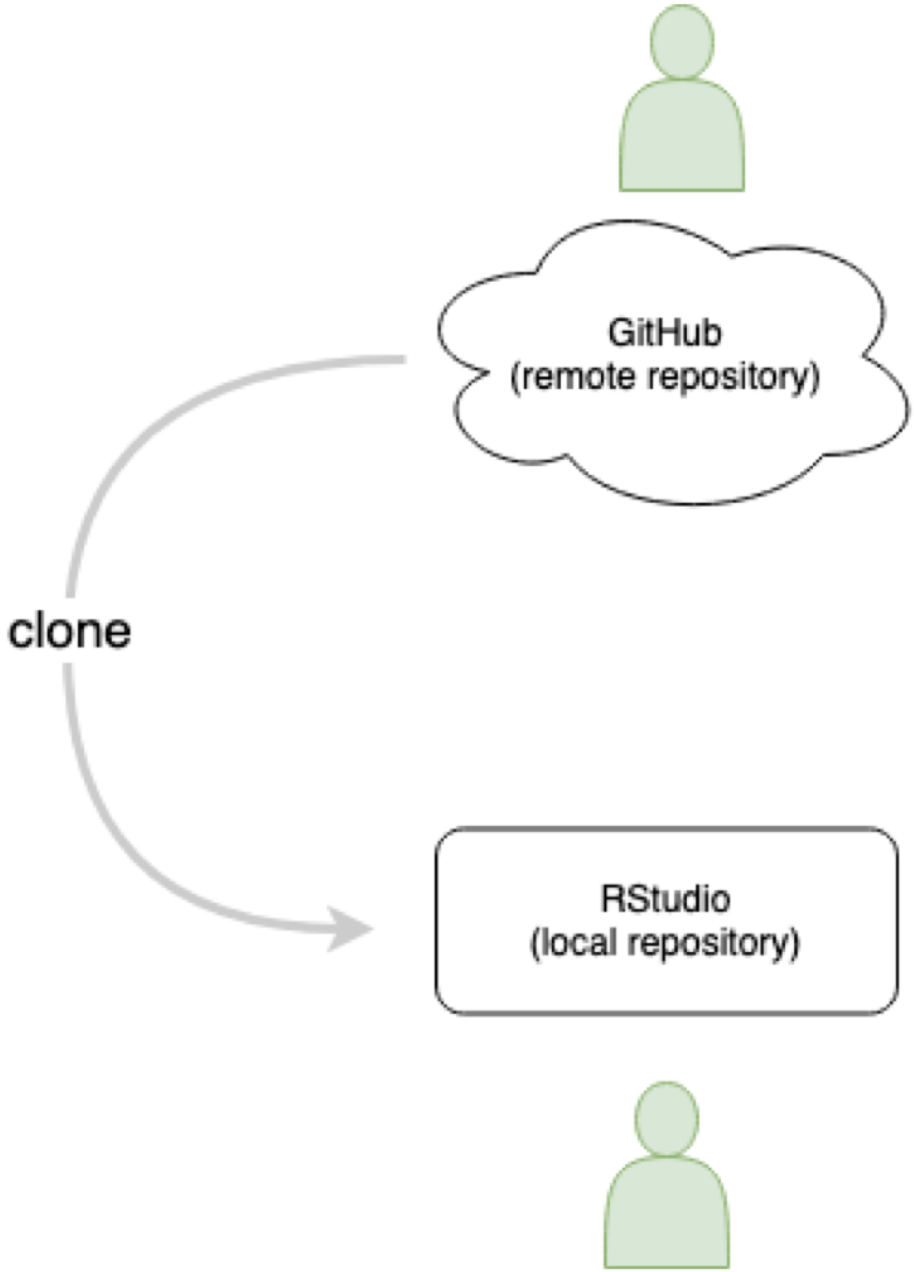
- Git is a software for version control
- Released in 2005
- Popular among programmers collaboratively developing code
- Tracks changes in a set of files (directory/folder/repository)

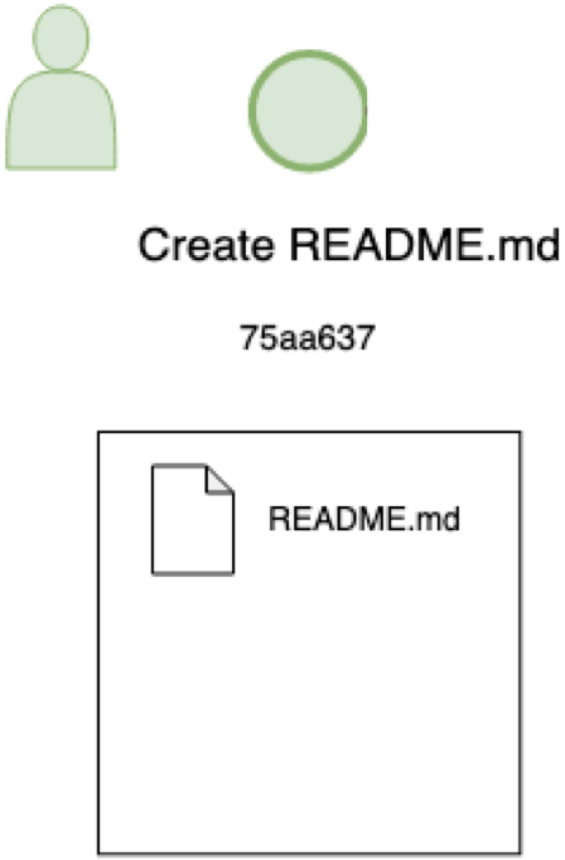
- GitHub is a hosting platform for version control using Git
- Launched in 2008, aquired by Microsoft in in 2018, Microsoft for US$ 7.5 billion
- 73 million Users (February, 2022)
- Facebook for Software Developers

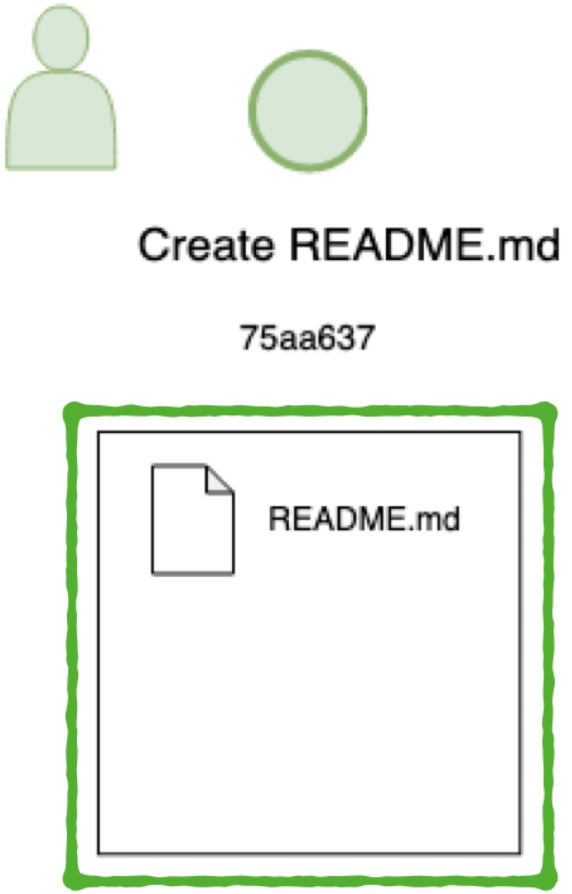
Homework Assignment 2

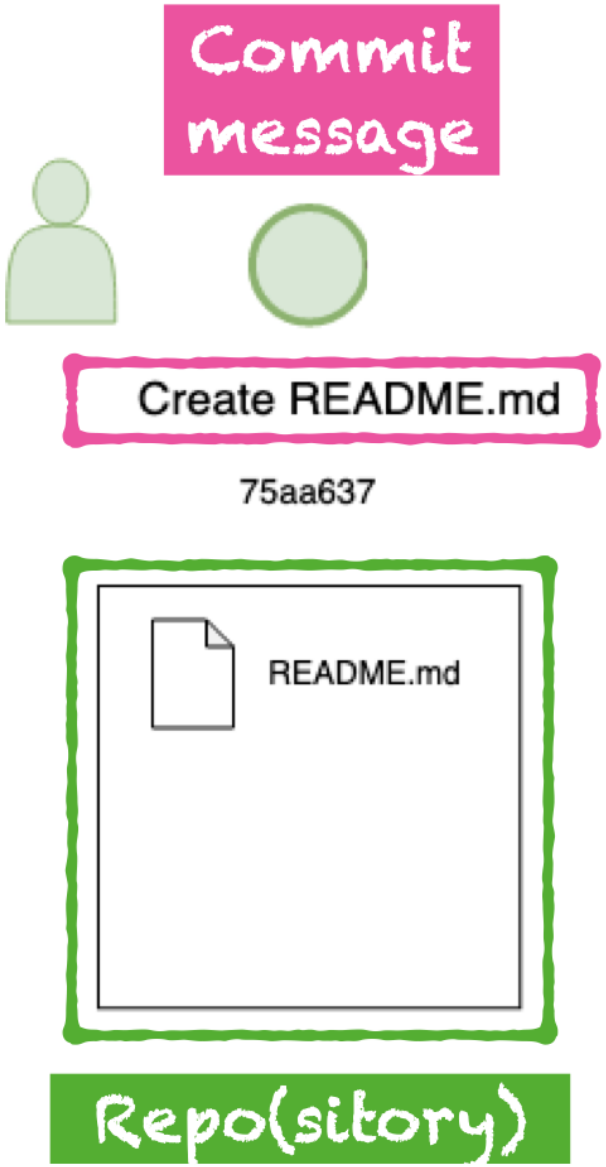
Version Control - Terminology

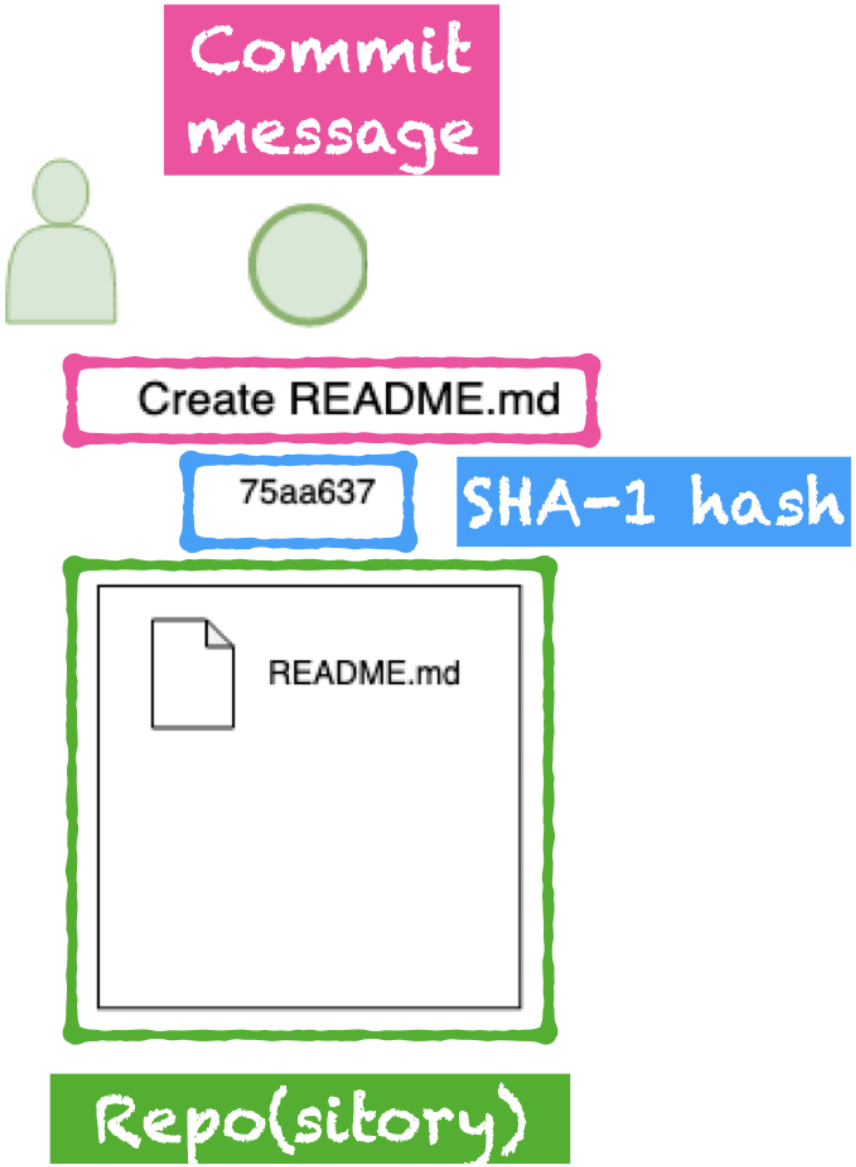

GitHub
(remote repository)

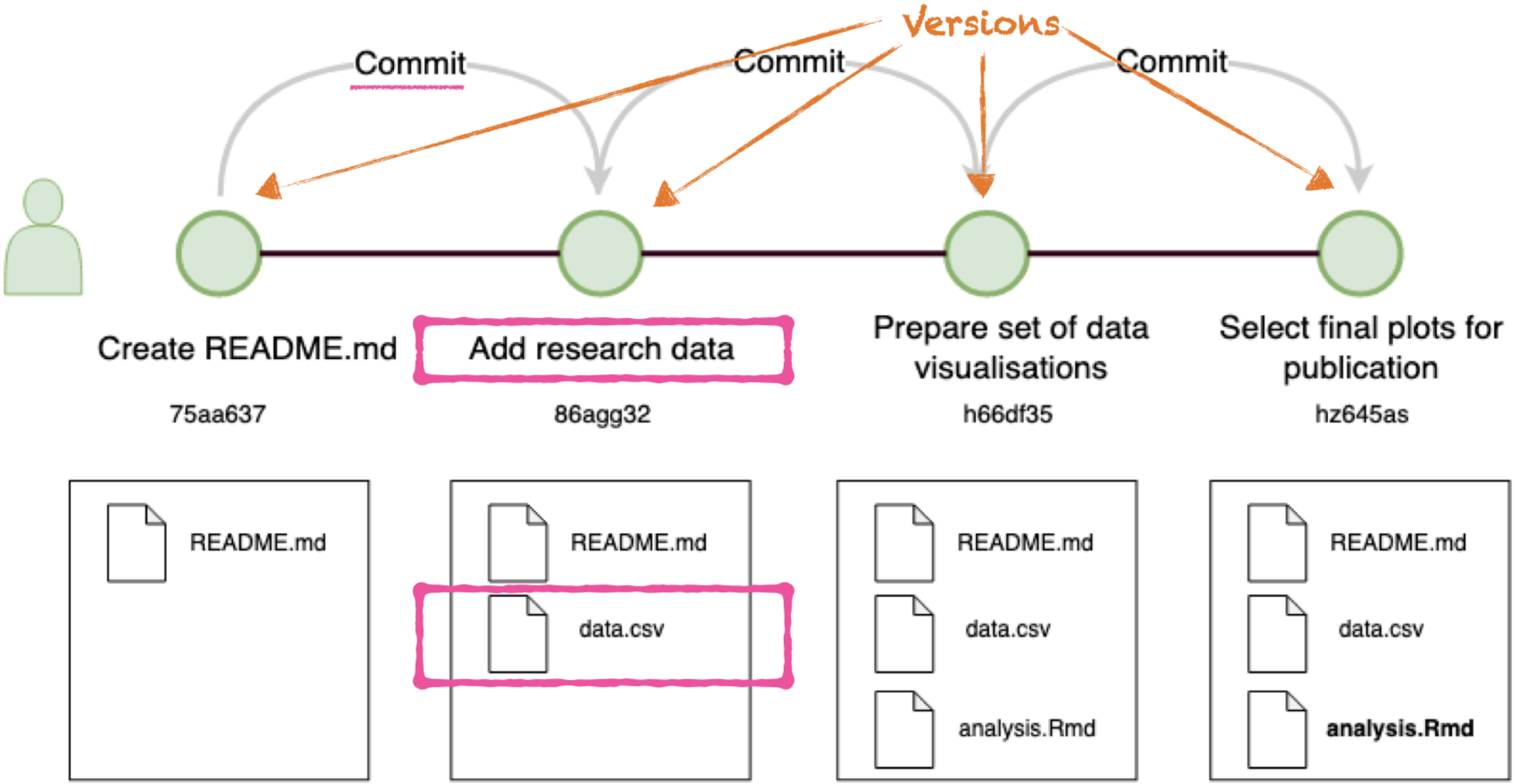

GitHub
(remote repository)
clone
RStudio
(local repository)

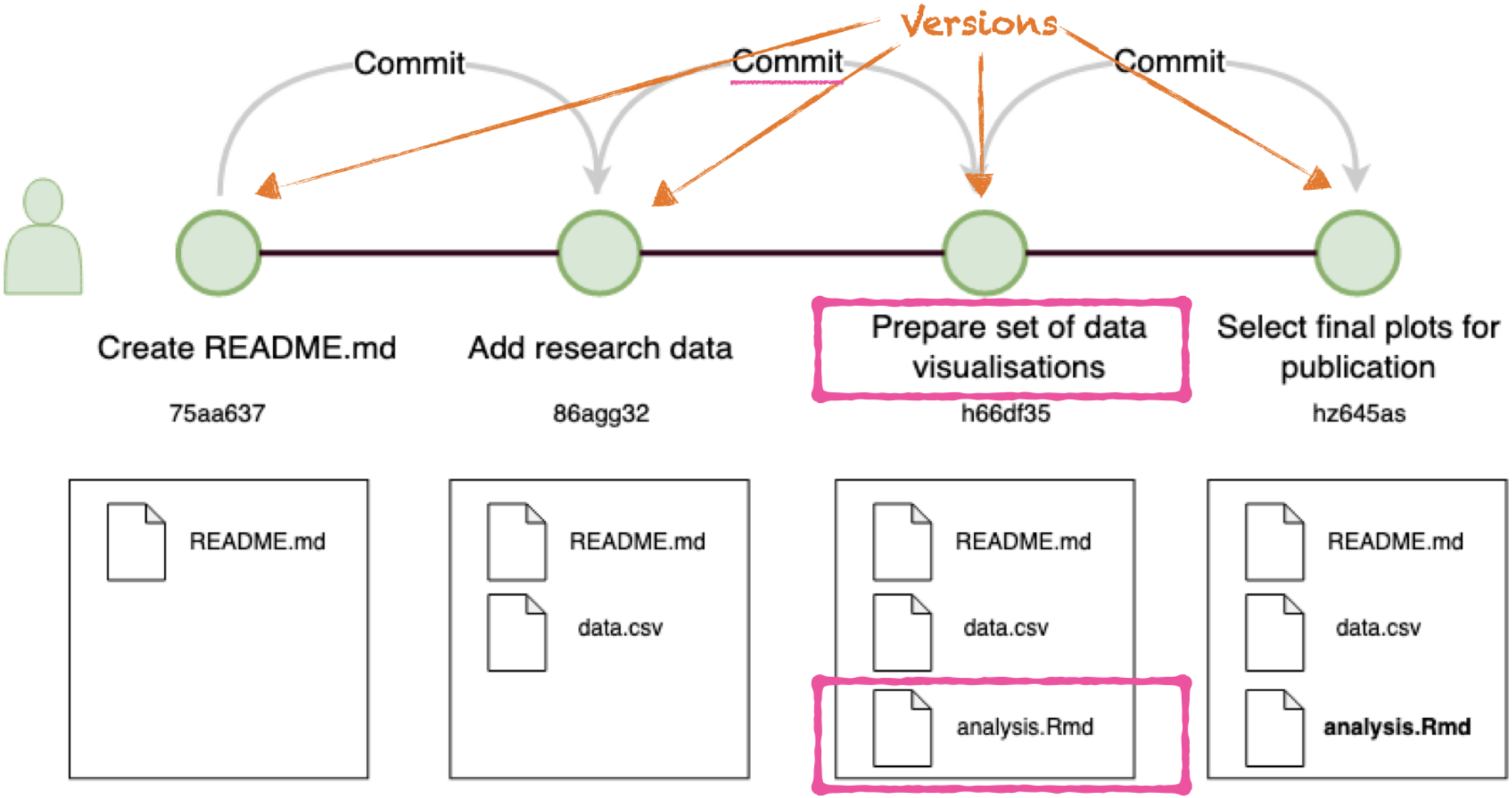

Create README.md
75aa637
README.md

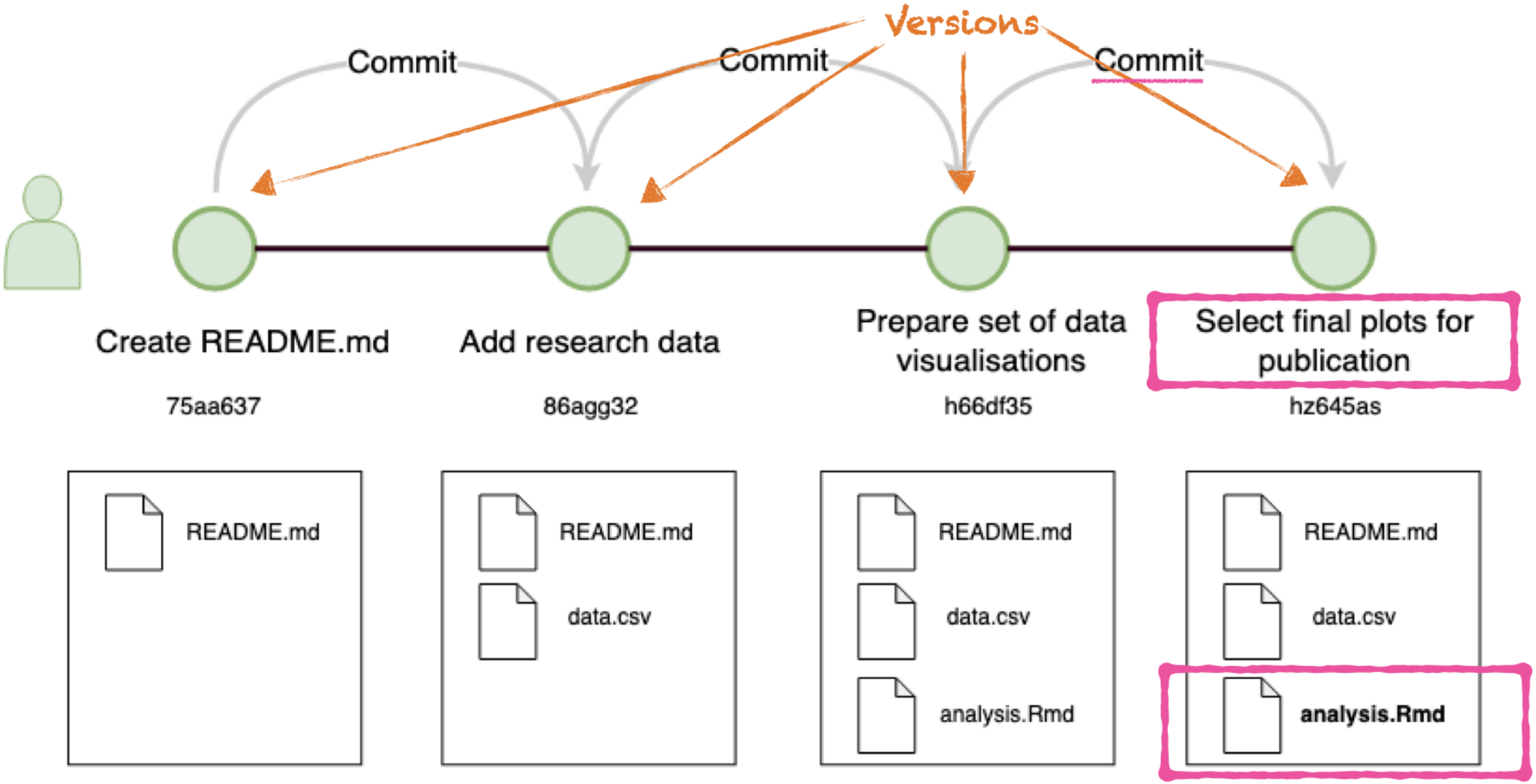

Create README.md
75aa637
README.md
Repo(sitory)

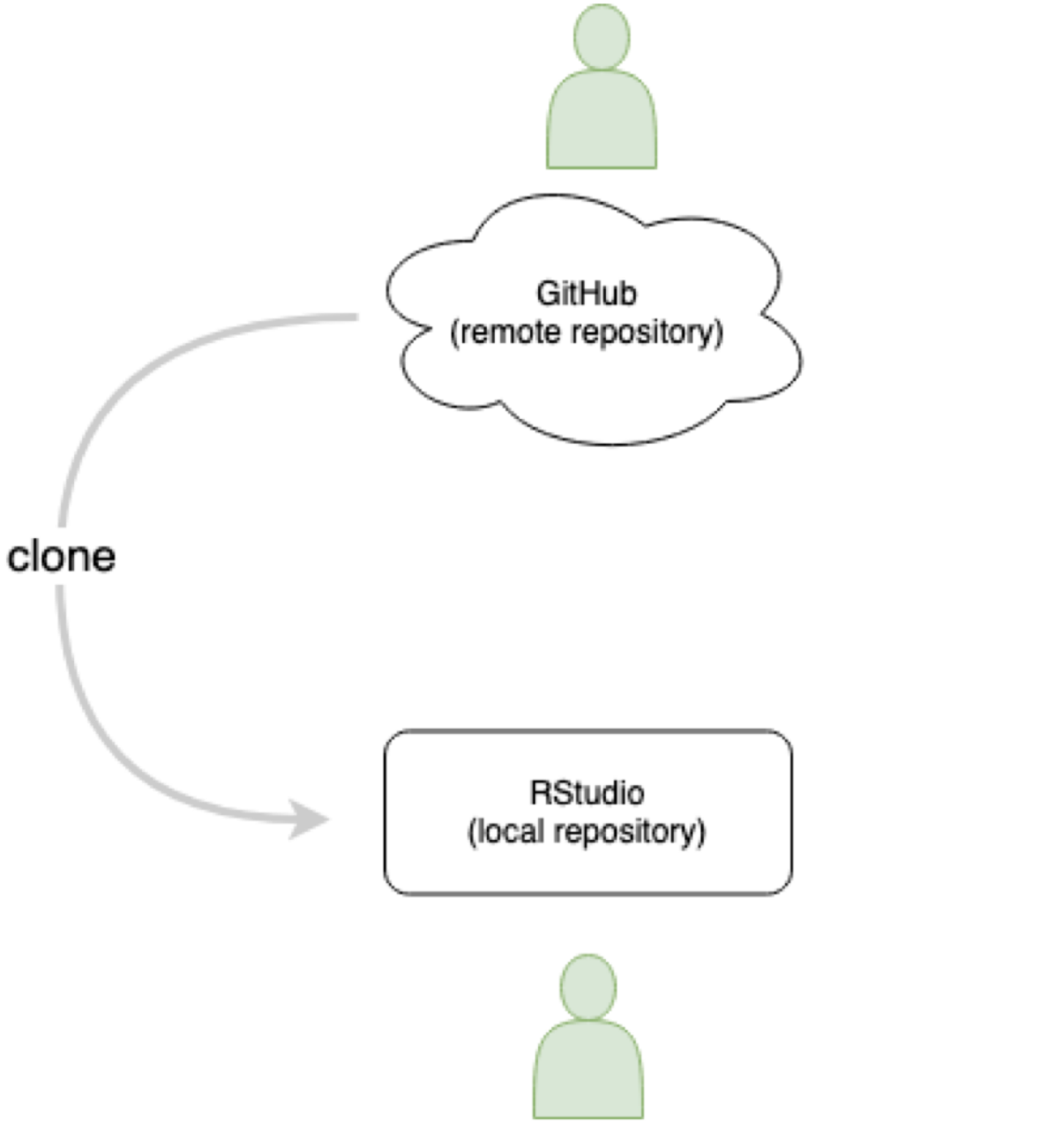

Commit message
Create README.md
75aa637
README.md
Repo(sitory)

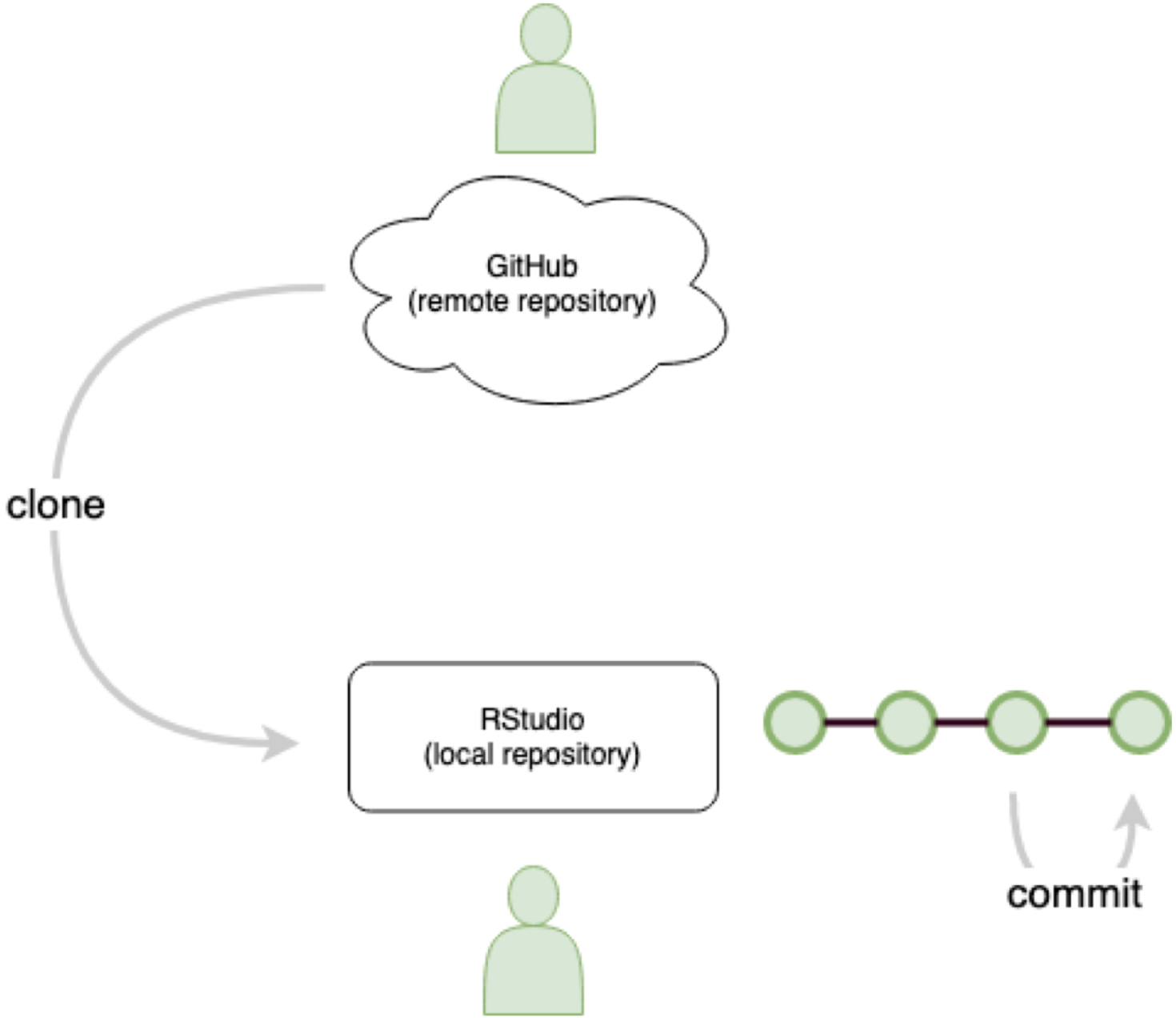

Commit message
Create README.md
75aa637
SHA-1 hash
README.md
Repo(sitory)

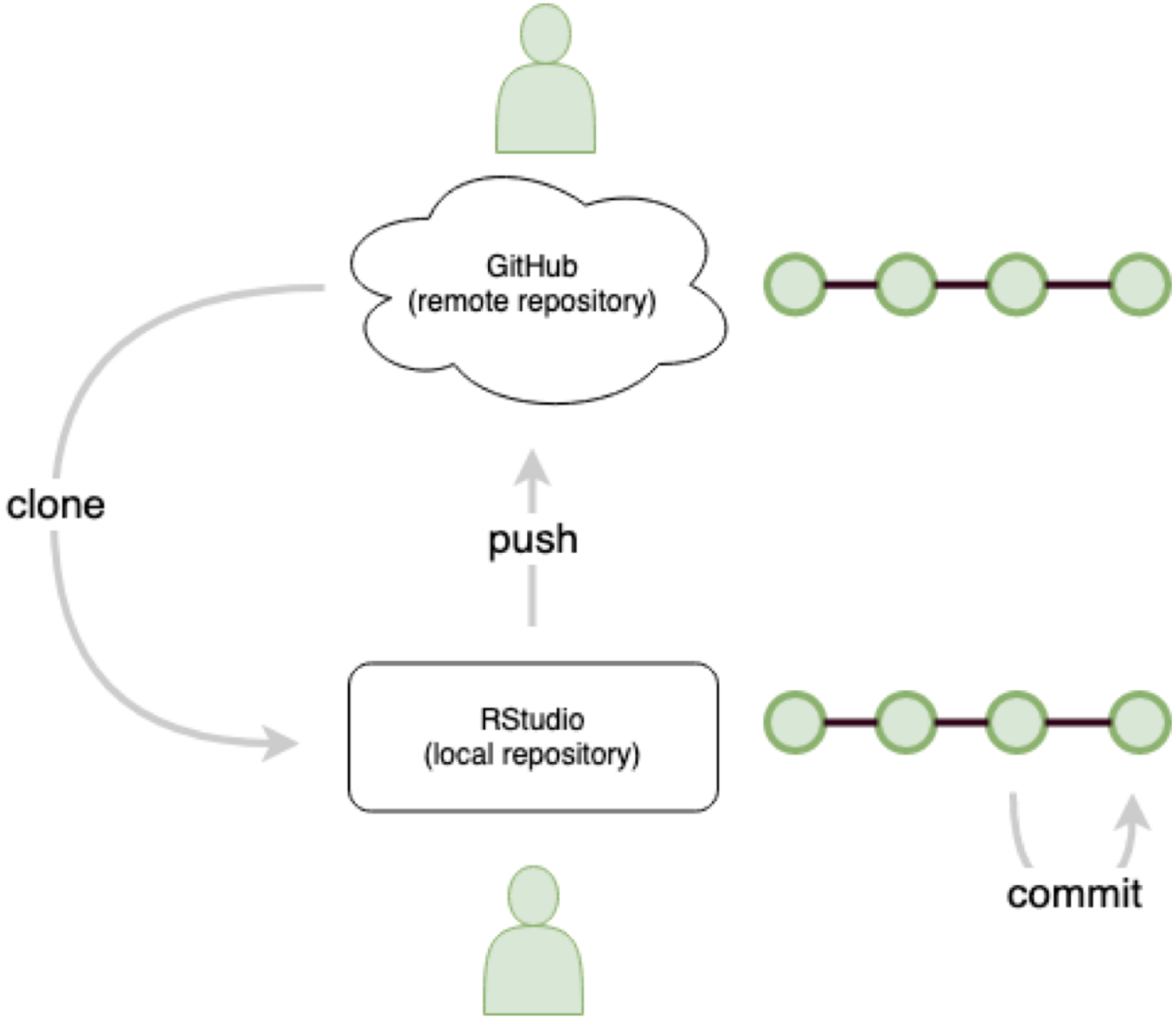

Versions
Commit
Commit
Commit
Create README.md
75aa637
Add research data
86agg32
Prepare set of data visualisations
h66df35
Select final plots for publication
hz645as
README.md
README.md
data.csv
README.md
data.csv
analysis.Rmd
README.md
data.csv
analysis.Rmd


Versions
Commit
Commit
Commit
Create README.md
75aa637
Add research data
86agg32
Prepare set of data visualisations
h66df35
Select final plots for publication
hz645as
README.md
README.md
data.csv
README.md
data.csv
analysis.Rmd
README.md
data.csv
analysis.Rmd


Versions
Commit
Commit
Commit
Create README.md
75aa637
Add research data
86agg32
Prepare set of data visualisations
h66df35
Select final plots for publication
hz645as
README.md
README.md
data.csv
README.md
data.csv
analysis.Rmd
README.md
data.csv
analysis.Rmd

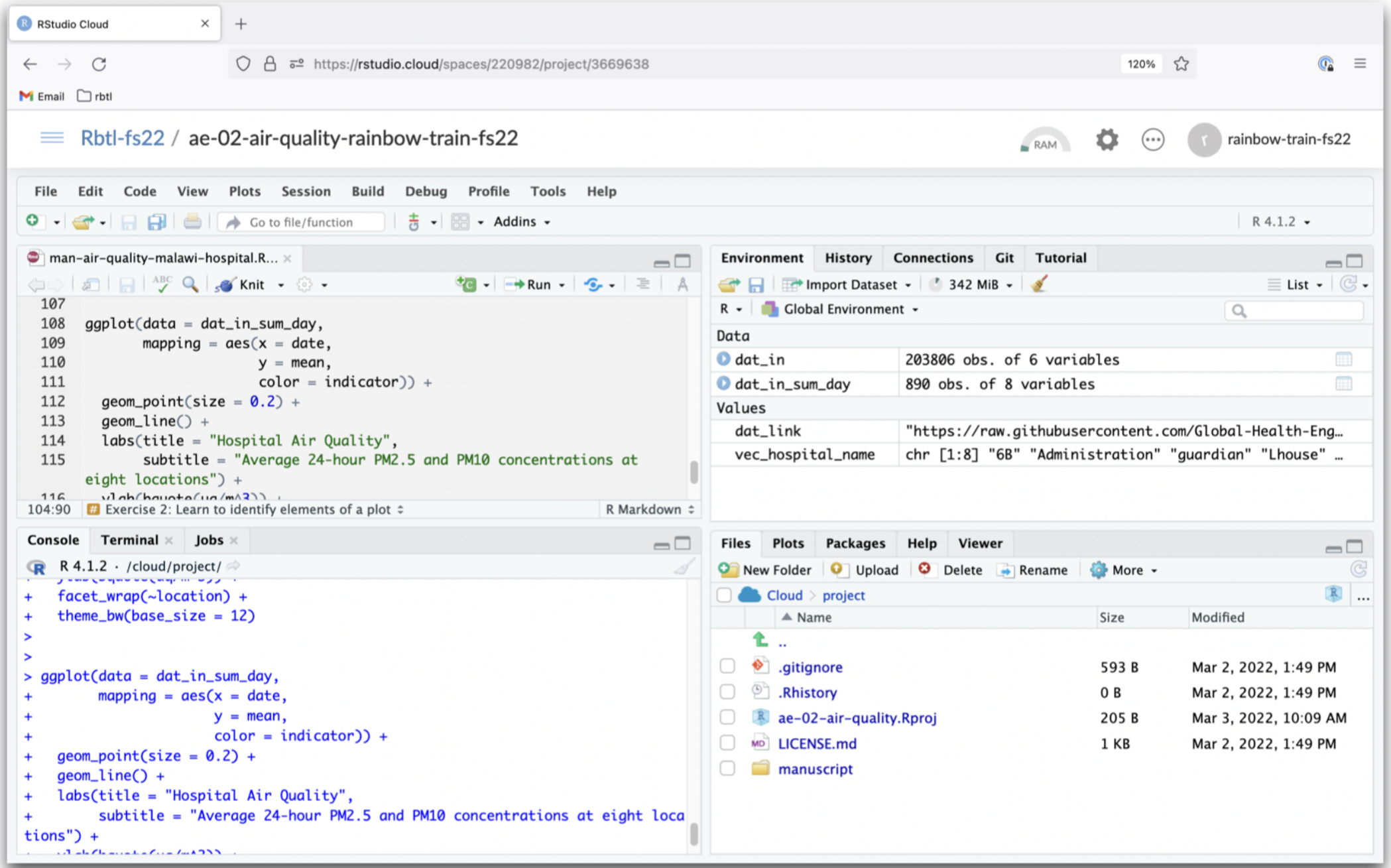

GitHub
(remote repository)
clone
RStudio
(local repository)

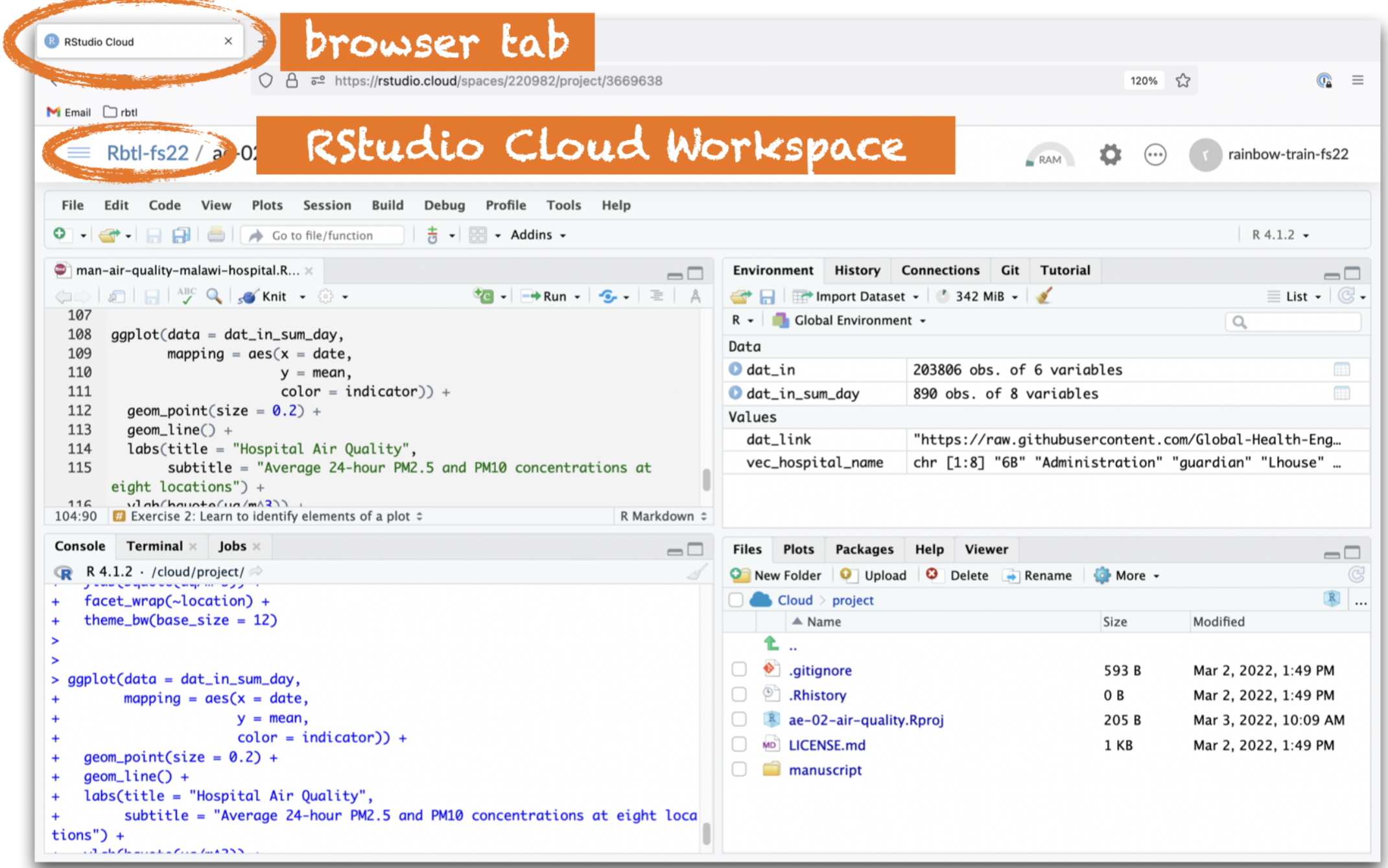

GitHub
(remote repository)
clone
RStudio
(local repository)
commit

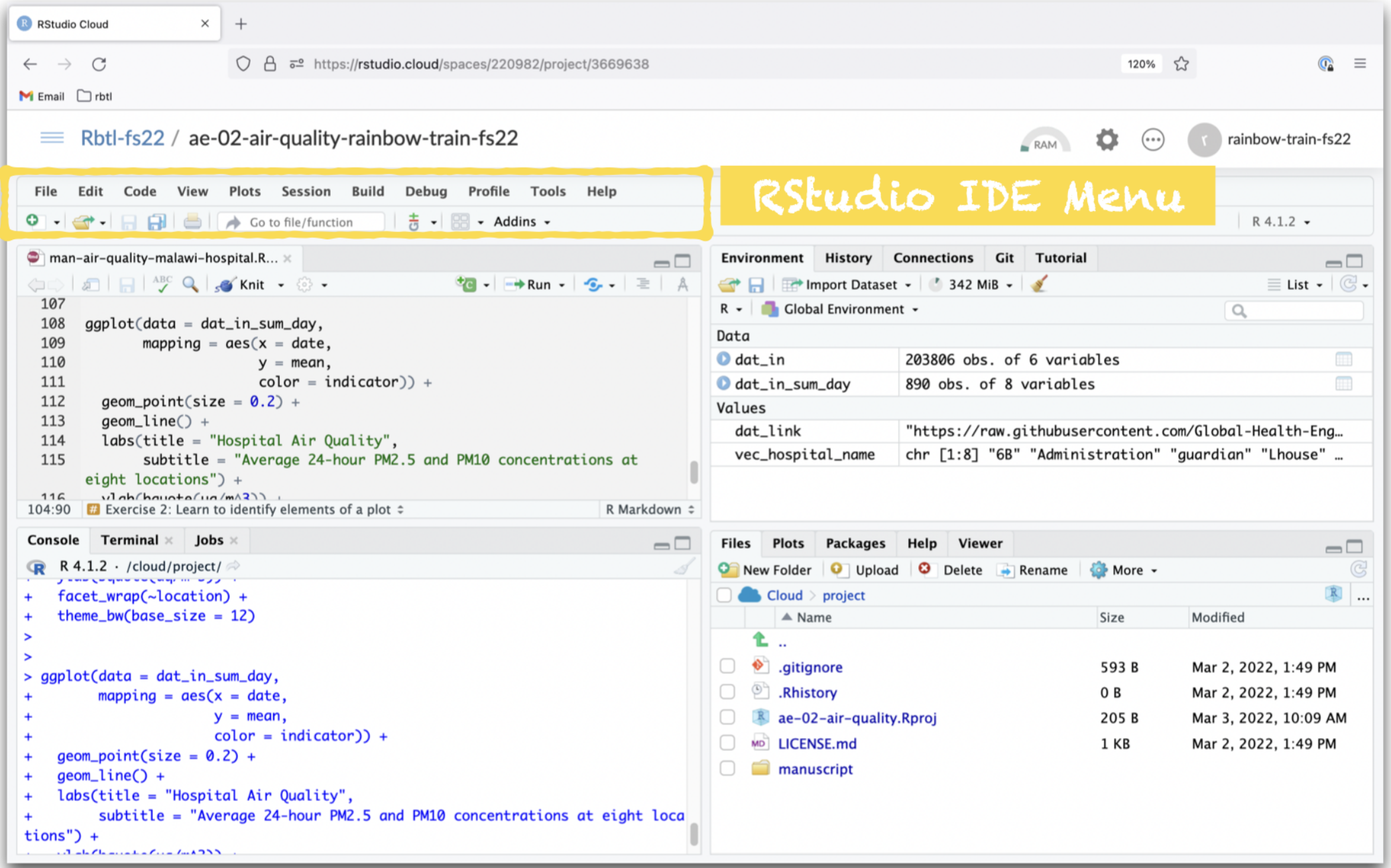

GitHub
(remote repository)
clone
push
RStudio
(local repository)
commit

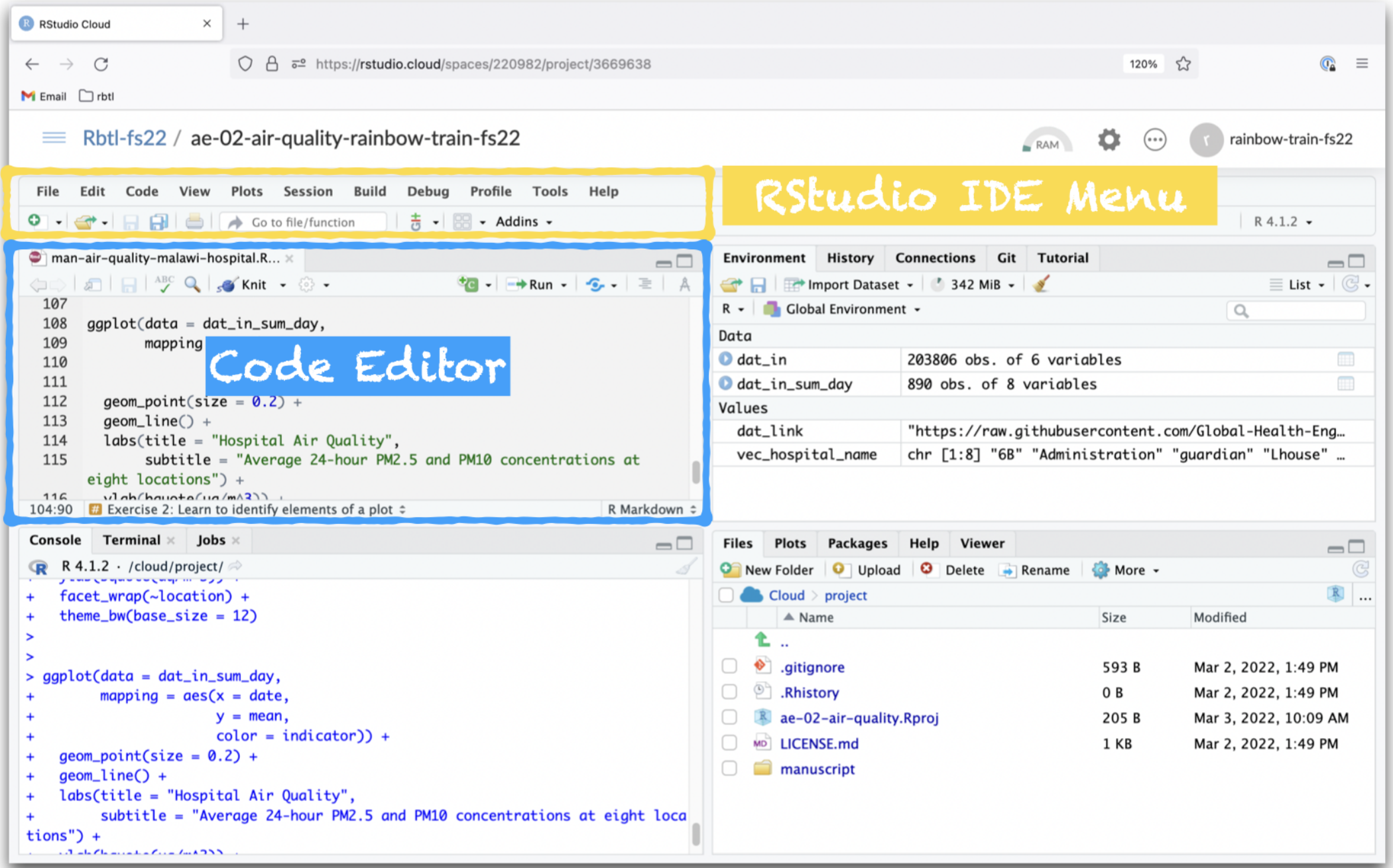
RStudio

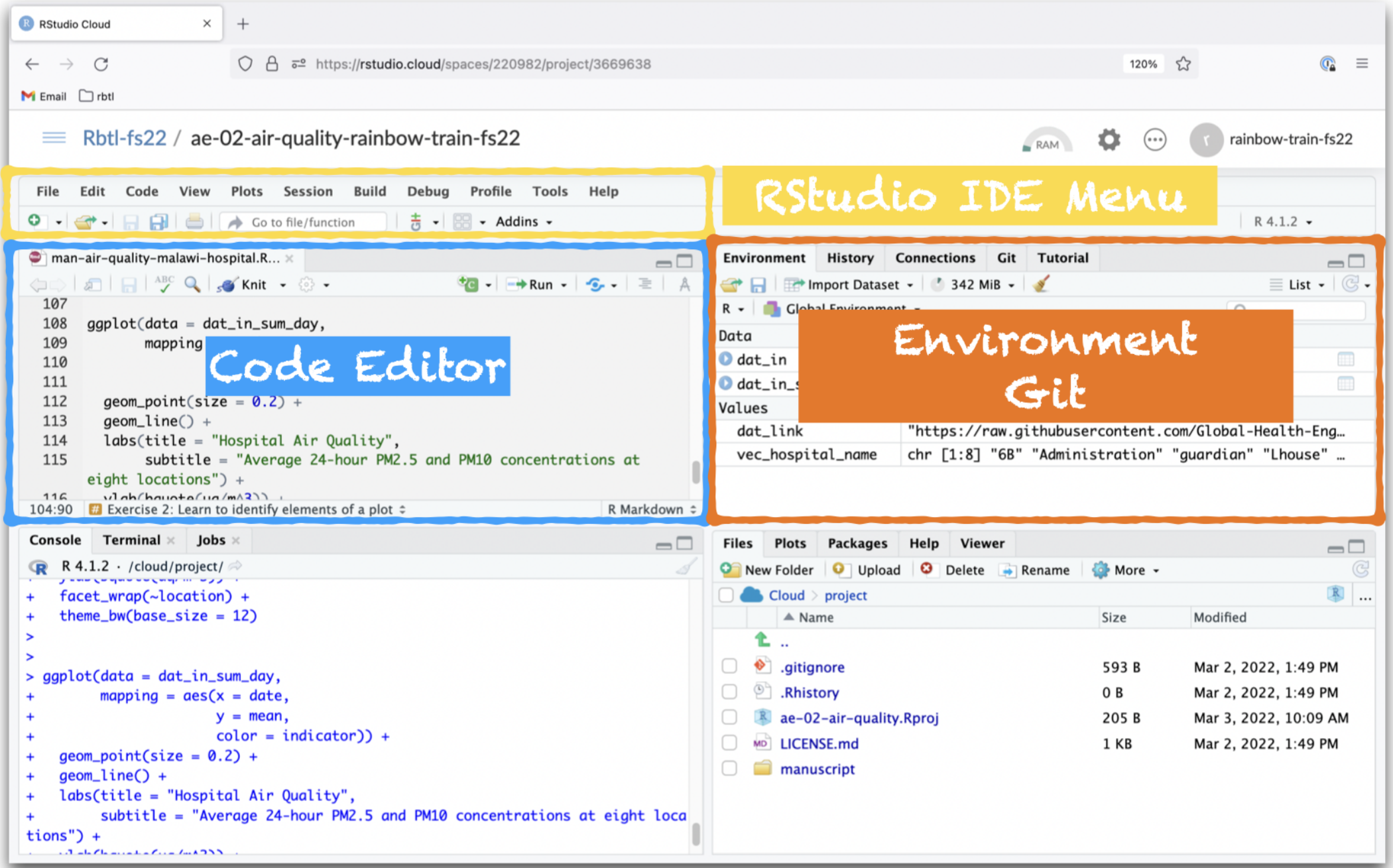
RStudio Cloud & RStudio IDE

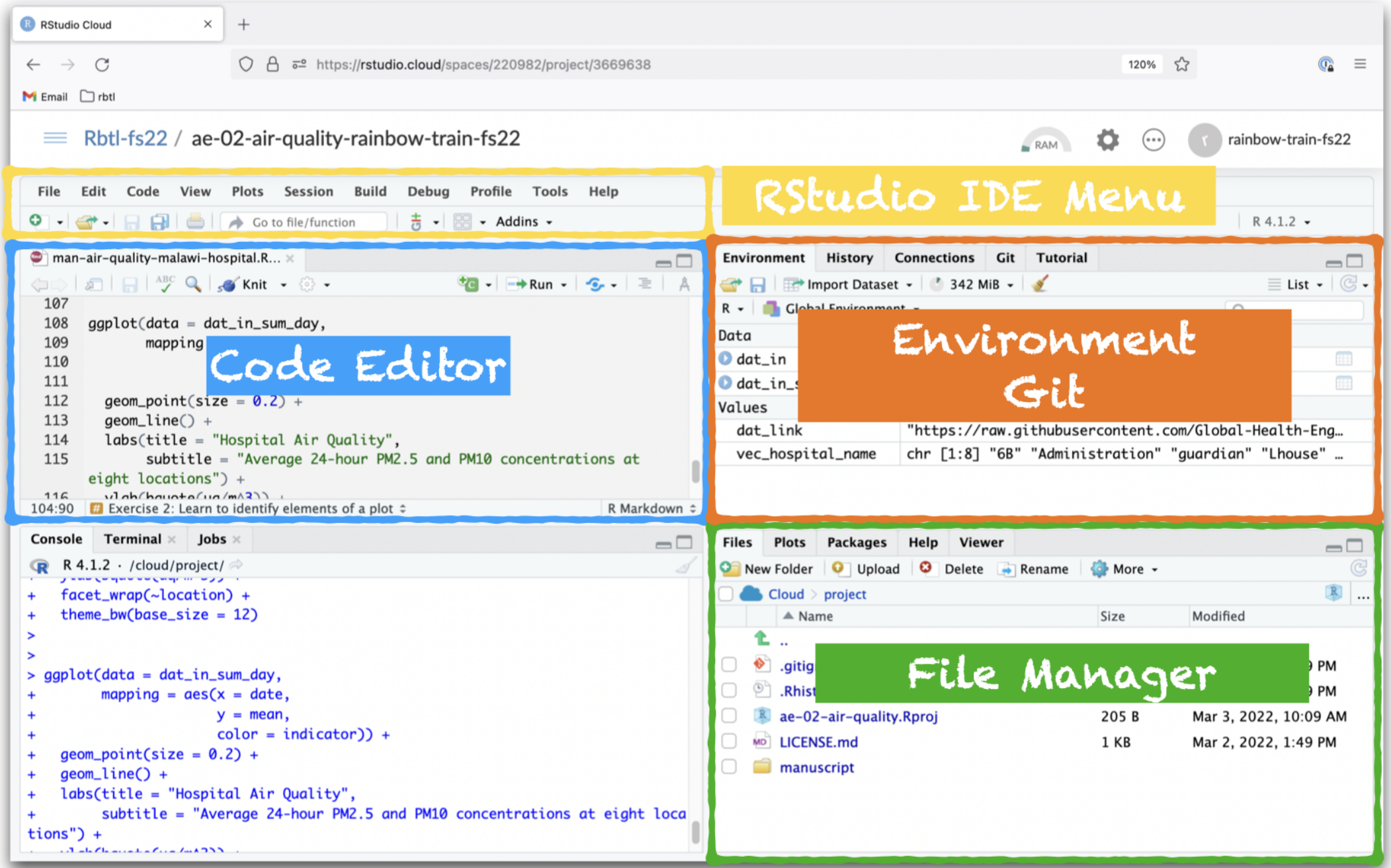

RStudio Cloud
https://rstudio.cloud/spaces/220982/project/3669638
120%
Email
rbtl
Rbtl-fs22 / ae-02-air-quality-rainbow-train-fs22
RAM
rainbow-train-fs22
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 4.1.2
man-air-quality-malawi-hospital.R...
Knit
Run
107
108 ggplot(data = dat_in_sum_day,
109 mapping = aes(x = date,
110 y = mean,
111 color = indicator)) +
112 geom_point(size = 0.2) +
113 geom_line() +
114 labs(title = "Hospital Air Quality",
115 subtitle = "Average 24-hour PM2.5 and PM10 concentrations at
eight locations") +
104:90
Exercise 2: Learn to identify elements of a plot
R Markdown
Environment History Connections Git Tutorial
Import Dataset
342 MiB
List
R
Global Environment
Data
dat_in 203806 obs. of 6 variables
dat_in_sum_day 890 obs. of 8 variables
Values
dat_link "https://raw.githubusercontent.com/Global-Health-Eng…
vec_hospital_name chr [1:8] "6B" "Administration" "guardian" "Lhouse" …
Console Terminal Jobs
R 4.1.2 · /cloud/project/
+ facet_wrap(~location) +
+ theme_bw(base_size = 12)
>
>
> ggplot(data = dat_in_sum_day,
+ mapping = aes(x = date,
+ y = mean,
+ color = indicator)) +
+ geom_point(size = 0.2) +
+ geom_line() +
+ labs(title = "Hospital Air Quality",
+ subtitle = "Average 24-hour PM2.5 and PM10 concentrations at eight loca
tions") +
Files Plots Packages Help Viewer
New Folder Upload Delete Rename More
Cloud project
Name Size Modified
..
.gitignore 593 B Mar 2, 2022, 1:49 PM
.Rhistory 0 B Mar 2, 2022, 1:49 PM
ae-02-air-quality.Rproj 205 B Mar 3, 2022, 10:09 AM
LICENSE.md 1 KB Mar 2, 2022, 1:49 PM
manuscript


RStudio Cloud
browser tab
https://rstudio.cloud/spaces/220982/project/3669638
120%
Email
rbtl
Rbtl-fs22 / a
RStudio Cloud Workspace
RAM
rainbow-train-fs22
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 4.1.2
man-air-quality-malawi-hospital.R...
Knit
Run
107
108 ggplot(data = dat_in_sum_day,
109 mapping = aes(x = date,
110 y = mean,
111 color = indicator)) +
112 geom_point(size = 0.2) +
113 geom_line() +
114 labs(title = "Hospital Air Quality",
115 subtitle = "Average 24-hour PM2.5 and PM10 concentrations at
eight locations") +
104:90
Exercise 2: Learn to identify elements of a plot
R Markdown
Environment History Connections Git Tutorial
Import Dataset
342 MiB
List
R
Global Environment
Data
dat_in 203806 obs. of 6 variables
dat_in_sum_day 890 obs. of 8 variables
Values
dat_link "https://raw.githubusercontent.com/Global-Health-Eng…
vec_hospital_name chr [1:8] "6B" "Administration" "guardian" "Lhouse" …
Console Terminal Jobs
R 4.1.2 · /cloud/project/
+ facet_wrap(~location) +
+ theme_bw(base_size = 12)
>
>
> ggplot(data = dat_in_sum_day,
+ mapping = aes(x = date,
+ y = mean,
+ color = indicator)) +
+ geom_point(size = 0.2) +
+ geom_line() +
+ labs(title = "Hospital Air Quality",
+ subtitle = "Average 24-hour PM2.5 and PM10 concentrations at eight loca
tions") +
Files Plots Packages Help Viewer
New Folder Upload Delete Rename More
Cloud project
Name Size Modified
..
.gitignore 593 B Mar 2, 2022, 1:49 PM
.Rhistory 0 B Mar 2, 2022, 1:49 PM
ae-02-air-quality.Rproj 205 B Mar 3, 2022, 10:09 AM
LICENSE.md 1 KB Mar 2, 2022, 1:49 PM
manuscript


RStudio Cloud
https://rstudio.cloud/spaces/220982/project/3669638
120%
Email
rbtl
Rbtl-fs22 / ae-02-air-quality-rainbow-train-fs22
RAM
rainbow-train-fs22
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
RStudio IDE Menu
R 4.1.2
man-air-quality-malawi-hospital.R...
Knit
Run
107
108 ggplot(data = dat_in_sum_day,
109 mapping = aes(x = date,
110 y = mean,
111 color = indicator)) +
112 geom_point(size = 0.2) +
113 geom_line() +
114 labs(title = "Hospital Air Quality",
115 subtitle = "Average 24-hour PM2.5 and PM10 concentrations at
eight locations") +
104:90
Exercise 2: Learn to identify elements of a plot
R Markdown
Environment History Connections Git Tutorial
Import Dataset
342 MiB
List
R
Global Environment
Data
dat_in 203806 obs. of 6 variables
dat_in_sum_day 890 obs. of 8 variables
Values
dat_link "https://raw.githubusercontent.com/Global-Health-Eng…
vec_hospital_name chr [1:8] "6B" "Administration" "guardian" "Lhouse" …
Console Terminal Jobs
R 4.1.2 · /cloud/project/
+ facet_wrap(~location) +
+ theme_bw(base_size = 12)
>
>
> ggplot(data = dat_in_sum_day,
+ mapping = aes(x = date,
+ y = mean,
+ color = indicator)) +
+ geom_point(size = 0.2) +
+ geom_line() +
+ labs(title = "Hospital Air Quality",
+ subtitle = "Average 24-hour PM2.5 and PM10 concentrations at eight loca
tions") +
Files Plots Packages Help Viewer
New Folder Upload Delete Rename More
Cloud project
Name Size Modified
..
.gitignore 593 B Mar 2, 2022, 1:49 PM
.Rhistory 0 B Mar 2, 2022, 1:49 PM
ae-02-air-quality.Rproj 205 B Mar 3, 2022, 10:09 AM
LICENSE.md 1 KB Mar 2, 2022, 1:49 PM
manuscript


RStudio Cloud
https://rstudio.cloud/spaces/220982/project/3669638
120%
Email
rbtl
Rbtl-fs22 / ae-02-air-quality-rainbow-train-fs22
RAM
rainbow-train-fs22
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
RStudio IDE Menu
R 4.1.2
man-air-quality-malawi-hospital.R...
Knit
Run
107
108 ggplot(data = dat_in_sum_day,
109 mapping
110
111
112 geom_point(size = 0.2) +
113 geom_line() +
114 labs(title = "Hospital Air Quality",
115 subtitle = "Average 24-hour PM2.5 and PM10 concentrations at
eight locations") +
Code Editor
104:90
Exercise 2: Learn to identify elements of a plot
R Markdown
Environment History Connections Git Tutorial
Import Dataset
342 MiB
List
R
Global Environment
Data
dat_in 203806 obs. of 6 variables
dat_in_sum_day 890 obs. of 8 variables
Values
dat_link "https://raw.githubusercontent.com/Global-Health-Eng…
vec_hospital_name chr [1:8] "6B" "Administration" "guardian" "Lhouse" …
Console Terminal Jobs
R 4.1.2 · /cloud/project/
+ facet_wrap(~location) +
+ theme_bw(base_size = 12)
>
>
> ggplot(data = dat_in_sum_day,
+ mapping = aes(x = date,
+ y = mean,
+ color = indicator)) +
+ geom_point(size = 0.2) +
+ geom_line() +
+ labs(title = "Hospital Air Quality",
+ subtitle = "Average 24-hour PM2.5 and PM10 concentrations at eight loca
tions") +
Files Plots Packages Help Viewer
New Folder
Upload
Delete
Rename
More
Cloud project
Name Size Modified
..
.gitignore 593 B Mar 2, 2022, 1:49 PM
.Rhistory 0 B Mar 2, 2022, 1:49 PM
ae-02-air-quality.Rproj 205 B Mar 3, 2022, 10:09 AM
LICENSE.md 1 KB Mar 2, 2022, 1:49 PM
manuscript


RStudio Cloud
https://rstudio.cloud/spaces/220982/project/3669638
120%
Email
rbtl
Rbtl-fs22 / ae-02-air-quality-rainbow-train-fs22
RAM
rainbow-train-fs22
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
RStudio IDE Menu
R 4.1.2
man-air-quality-malawi-hospital.R...
Knit
Run
107
108 ggplot(data = dat_in_sum_day,
109 mapping
110
111
112 geom_point(size = 0.2) +
113 geom_line() +
114 labs(title = "Hospital Air Quality",
115 subtitle = "Average 24-hour PM2.5 and PM10 concentrations at
eight locations") +
Code Editor
104:90
Exercise 2: Learn to identify elements of a plot
R Markdown
Environment
History
Connections
Git
Tutorial
Import Dataset
342 MiB
List
R
Global Environment
Data
dat_in
dat_in_s
Values
dat_link "https://raw.githubusercontent.com/Global-Health-Eng…
vec_hospital_name chr [1:8] "6B" "Administration" "guardian" "Lhouse" …
Environment
Git
Console
Terminal
Jobs
R 4.1.2 · /cloud/project/
+ facet_wrap(~location) +
+ theme_bw(base_size = 12)
>
>
> ggplot(data = dat_in_sum_day,
+ mapping = aes(x = date,
+ y = mean,
+ color = indicator)) +
+ geom_point(size = 0.2) +
+ geom_line() +
+ labs(title = "Hospital Air Quality",
+ subtitle = "Average 24-hour PM2.5 and PM10 concentrations at eight loca
tions") +
Files
Plots
Packages
Help
Viewer
New Folder
Upload
Delete
Rename
More
Cloud
project
Name
Size
Modified
..
.gitignore 593 B Mar 2, 2022, 1:49 PM
.Rhistory 0 B Mar 2, 2022, 1:49 PM
ae-02-air-quality.Rproj 205 B Mar 3, 2022, 10:09 AM
LICENSE.md 1 KB Mar 2, 2022, 1:49 PM
manuscript


RStudio Cloud
https://rstudio.cloud/spaces/220982/project/3669638
120%
Email
rbtl
Rbtl-fs22 / ae-02-air-quality-rainbow-train-fs22
RAM
rainbow-train-fs22
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
RStudio IDE Menu
R 4.1.2
man-air-quality-malawi-hospital.R...
Knit
Run
107
108 ggplot(data = dat_in_sum_day,
109 mapping
110
111
112 geom_point(size = 0.2) +
113 geom_line() +
114 labs(title = "Hospital Air Quality",
115 subtitle = "Average 24-hour PM2.5 and PM10 concentrations at
eight locations") +
Code Editor
104:90
Exercise 2: Learn to identify elements of a plot
R Markdown
Environment History Connections Git Tutorial
Import Dataset
342 MiB
List
R
Global Environment
Data
dat_in
dat_in_s
Values
dat_link "https://raw.githubusercontent.com/Global-Health-Eng…
vec_hospital_name chr [1:8] "6B" "Administration" "guardian" "Lhouse" …
Environment
Git
Console Terminal Jobs
R 4.1.2 · /cloud/project/
+ facet_wrap(~location) +
+ theme_bw(base_size = 12)
>
>
> ggplot(data = dat_in_sum_day,
+ mapping = aes(x = date,
+ y = mean,
+ color = indicator)) +
+ geom_point(size = 0.2) +
+ geom_line() +
+ labs(title = "Hospital Air Quality",
+ subtitle = "Average 24-hour PM2.5 and PM10 concentrations at eight loca
tions") +
Files Plots Packages Help Viewer
New Folder
Upload
Delete
Rename
More
Cloud > project
Name
Size
Modified
..
.gitig
.Rhist
ae-02-air-quality.Rproj 205 B Mar 3, 2022, 10:09 AM
LICENSE.md 1 KB Mar 2, 2022, 1:49 PM
manuscript
File Manager

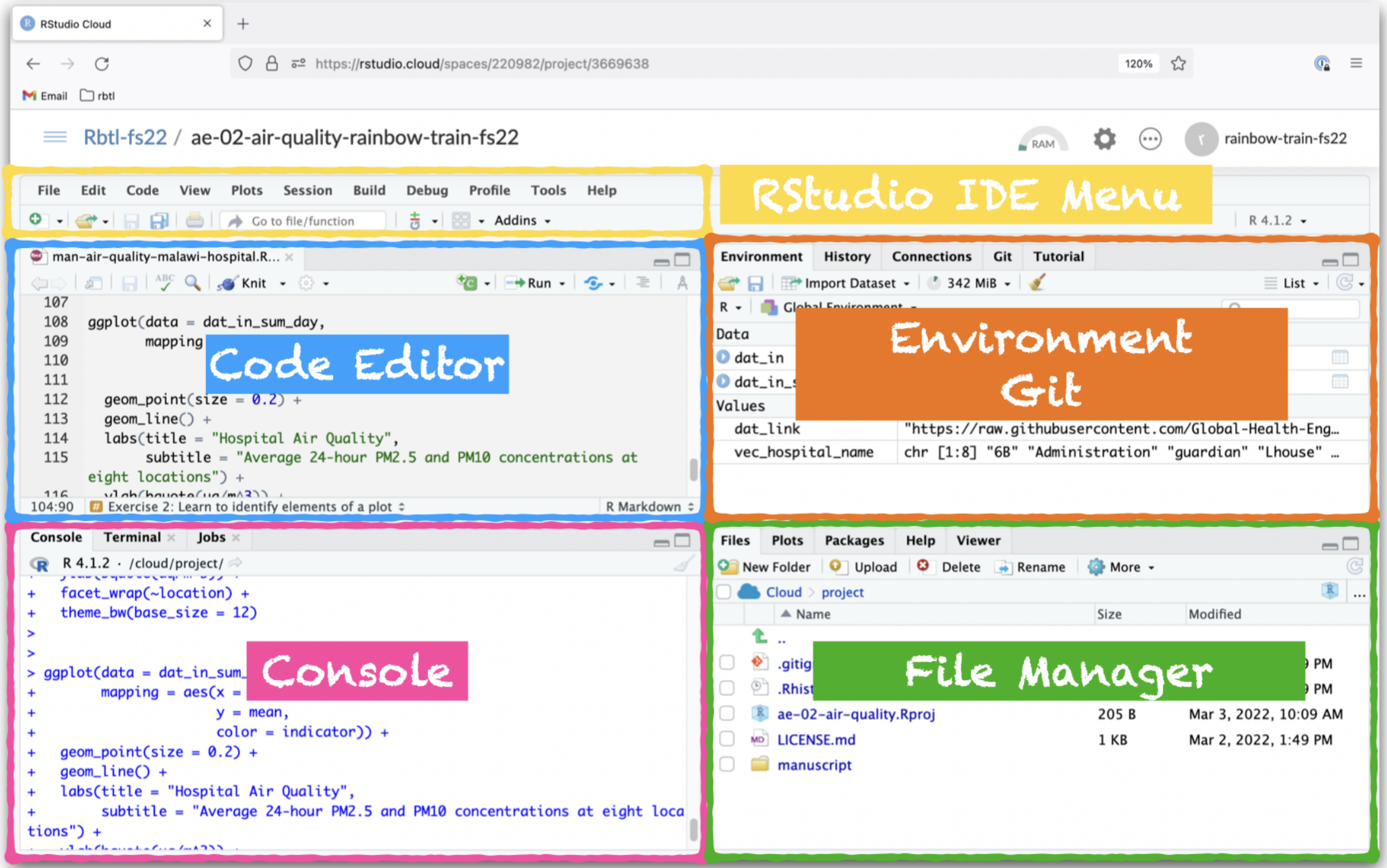

RStudio Cloud
https://rstudio.cloud/spaces/220982/project/3669638
120%
Email
rbtl
Rbtl-fs22 / ae-02-air-quality-rainbow-train-fs22
RAM
rainbow-train-fs22
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R 4.1.2
RStudio IDE Menu
man-air-quality-malawi-hospital.R...
Knit
Run
107
108 ggplot(data = dat_in_sum_day,
109 mapping
110
111
112 geom_point(size = 0.2) +
113 geom_line() +
114 labs(title = "Hospital Air Quality",
115 subtitle = "Average 24-hour PM2.5 and PM10 concentrations at
eight locations") +
Code Editor
104:90
Exercise 2: Learn to identify elements of a plot
R Markdown
Environment
History
Connections
Git
Tutorial
Import Dataset
342 MiB
List
R
Global Environment
Data
dat_in
dat_in_s
Values
dat_link "https://raw.githubusercontent.com/Global-Health-Eng…
vec_hospital_name chr [1:8] "6B" "Administration" "guardian" "Lhouse" …
Environment
Git
Console
Terminal
Jobs
R 4.1.2 · /cloud/project/
+ facet_wrap(~location) +
+ theme_bw(base_size = 12)
>
>
> ggplot(data = dat_in_sum_
+ mapping = aes(x =
+ y = mean,
+ color = indicator)) +
+ geom_point(size = 0.2) +
+ geom_line() +
+ labs(title = "Hospital Air Quality",
+ subtitle = "Average 24-hour PM2.5 and PM10 concentrations at eight loca
tions") +
Console
Files
Plots
Packages
Help
Viewer
New Folder
Upload
Delete
Rename
More
Cloud
project
Name
Size
Modified
..
.gitig
.Rhist
ae-02-air-quality.Rproj 205 B Mar 3, 2022, 10:09 AM
LICENSE.md 1 KB Mar 2, 2022, 1:49 PM
manuscript
File Manager

Photo by: Matt Walsh

Thanks! 🌻

Slides created via the R packages:

xaringan
gadenbuie/xaringanthemer

The chakra comes from remark.js, **knitr**, and R Markdown.

Access slides as PDF on GitHub

All material is licensed under Creative Commons Attribution Share Alike 4.0 International.