

- **Welcome to the Etherpad for the Software Carpentries Workshop at Rutgers University-Camden!**

## **When, Where, and What**

This Software Carpentry workshop will be held on Thursday and Friday, March 17-18, 2022, from 9am to 5pm ET.

The workshop will be held online and at Room V12, RWJMS Research Tower, Busch Campus, Rutgers University - Camden (675 Hoes Lane West, Piscataway).

On this Software Carpentry workshop we will cover:

- Automating Tasks with the Unix Shell: <https://swcarpentry.github.io/shell-novice/>,
- Version Control with Git: <https://swcarpentry.github.io/git-novice/>, and
- Programming with R: <http://swcarpentry.github.io/r-novice-inflammation/>.

## **Working collaboratively**

This is an Etherpad, a document that is synchronized as you type, so that everyone viewing this page sees the same text. This allows you to collaborate seamlessly during class.

We will use this Etherpad to share links and snippets of code, take notes, ask and answer questions, and whatever else comes to mind.

The Etherpad is displayed as a screen with three major parts:

- The left side holds all notes: please work on this area as we go along.
- The top right side shows the names of users who are logged in: please add your name. You can also change the color to one that you like or that best reflects your mood and personality.
- The bottom right is a real time chat window for asking questions of the instructor and your fellow learners.

## **Pre-workshop Survey**

Please fill out the pre-training survey at <https://carpentries.typeform.com/to/QVOarK#slug=2022-03-17-rutgers-online>

## **Code of Conduct & License**

Users are expected to follow our code of conduct: [https://docs.carpentries.org/topic\\_folders/policies/code-of-conduct.html](https://docs.carpentries.org/topic_folders/policies/code-of-conduct.html)

All content is publicly available under the Creative Commons Attribution License:  
<https://creativecommons.org/licenses/by/4.0/>

## Getting Started

To start each session, please sign in the attendee list so we can record your attendance.

**Sign in format is Name, Pronouns, Institution, Email (optional), Twitter (optional)**

- Instructor: Heather Ciallela, she/her, Rutgers University-Camden
  - Instructor: Luna L. Sanchez Reyes, she/her, University of California, Merced
  - Helper: Kate Douthat
  - Helper: Jiayi Liu, She/Her, Rutgers, Molecular Biosciences, jl2791@rutgers.edu, <https://github.com/jiayiliujiayi>
  - Helper: Kellen Xu
  - Student: Bipin Kumar Tripathi, he/him, Rutgers-Busch Campus
- 
- 

## Day 1

### Morning session - the Unix Shell

Follow the lesson here:

<https://swcarpentry.github.io/shell-novice/>

Setup > Download shell-lesson-data.zip > Move to Desktop and unzip for later

### Sign in

**Sign in format: Name, Pronouns, Institution, Email (optional), Twitter (optional)**

- Instructor: Heather Ciallela, she/her, Rutgers University-Camden
- Helper: Kate Douthat she/her, Rutgers New Brunswick, Ecology and Evolution, kdd56@scarletmail.rutgers.edu, @ThatOakDuet
- Helper: Jiayi Liu, She/Her, Rutgers, Molecular Biosciences, jl2791@rutgers.edu, <https://github.com/jiayiliujiayi>
- Helper: Kellen Xu
- Student: Abby Heller She/Her, Rutgers-Molecular Biosciences PhD
- Postdoc: Aniket Bhattacharya He/Him, Cell Biology and Neuroscience, a.bhattacharya@rutgers.edu
- Alumnus: Iman Amer, she/her
- Postdoc: Consuelo Ibar, she/her, Waksman Institute
- Keyaara Robinson (she/her), Pharmaceutical Sciences
- Student: Mary Lally, she/her, Rutgers-Molecular Biosciences PhD
- Student: Xia Qiu, she/her, Rutgers-Molecular Biosciences PhD
- Student: Chang Liu, She/her, NJMS, cyl44@njms.rutgers.edu
- Postdoc: Jihyun Kim, she/her, Rutgers NB, CBN
- Student: Michael Warkala, he/him, Rutgers NJMS, @mwarkala
- Omkaram Gangisetty, Research Associate, Rutgers university, og53@sebs.rutgers.edu

- Student: Kyla Johnson, she/her, Rutgers GSBS, Newark
- Student: Natalie Losada, she/her, Rutgers - CABM, Busch
- Postdoc: Jennifer Timm, she/her, Rutgers - CABM, NewBrunswick
- Postdoc: Dana Reid, she/her, Rutgers (Cook-SEBS); dr990@sebs.rutgers.edu
- Postdoc: Edward Martinez, he/him, Rutgers-New Brunswick
- Student: Lara Brindisi, she/her, Plant Biology PhD, lara.brindisi@rutgers.edu
- Student: Bignyi Liang, she/her, NJIT
- Postdoc: Prasanta Dey (he/him), Rutgers Cancer Institute of New Jersey, pd510@cinj.rutgers.edu
- Postdoc: Bipin Kumar Tripathi, he/him, Rutgers Busch Campus Waksman Institute of Microbiology
- Postdoc: Guoqiang Wang, he/him, Rutgers
- postdoc: Jeesebastian, He/him, Public Health Research Institute, Rutgers University, Newark, Nj
- Postdoc : Raja Khan, he/him :Rutgers - CABM, Piscataway
- Postdoc: Maria Gomez, Rutgers Cancer Institute, mlg196@cinj.rutgers.edu
- 
- 
- 
- 
- 

## **Morning Break Feedback**

Please provide feedback for this session through this google form before leaving for break:  
<https://forms.gle/qrh4aRvpy7DM7ZHR6>

## **Lunch Break Feedback**

Please provide some more feedback through this google form before leaving for lunch:  
<https://forms.gle/8wpNJJd4r9jb8ok46>

---

## **Afternoon session day 1 - Programming with R**

Follow the lesson here:  
<http://swcarpentry.github.io/r-novice-inflammation/>

## **Sign in**

**Sign in format: Name, Pronouns, Institution, Email (optional), Twitter (optional)**

- Instructor: Luna L. Sanchez Reyes, she/her, University of California, Merced
- Helper: Heather Ciallela, she/her, Rutgers University-Camden

- Helper: Kate Douthat she/her, Rutgers New Brunswick, PhD Candidate, Ecology and Evolution, kdd56@scarletmail.rutgers.edu, @ThatOakDuet
- Helper: Jiayi Liu, She/Her, Rutgers, Molecular Biosciences, jl2791@rutgers.edu, <https://github.com/jiayiliujiayi>
- Helper: Kellen Xu
- Postdoc: Guoqiang wang, he/him, Rutgers, wangg5@dls.rutgers.edu
- Student: Kyla Johnson, she/her, Rutgers GSBS- Newark
- Student: Michael Warkala, he/him, Rutgers NJMS, @mwarkala
- Keyaara Robinson (she/her), Pharmaceutical Sciences
- Postdoc: Jihyun Kim (she/her), Rutgers NB, CBN
- Student: Mary Lally, she/her, Rutgers-Molecular Biosciences PhD
- student: Xia Qiu, she/her, Rutgers-Molecular Biosciences PhD
- Postdoc: Consuelo Ibar, she/her, Waksman Institute
- Alumnus: Iman Amer, she/her
- Student: Chang Liu, she/her, NJMS, cyl44@njms.rutgers.edu
- Postdoc: Edward Martinez, he/him, Rutgers-New Brunswick
- Postdoc: Prasanta Dey (he/him), Rutgers Cancer Institute of New Jersey, pd510@cinj.rutgers.edu
- Postdoc: Dana Reid (she/her), Rutgers (Cook-SEBS) dr990@sebs.rutgers.edu
- Postdoc: Aniket Bhattacharya (he/him), CHI, a.bhattacharya@rutgers.edu
- Research Associate: Omkaram Gangisetty, Rutgers university, og53@sebs.rutgers.edu
- Student: Abby Heller She/Her, Rutgers-Molecular Biosciences PhD
- Postdoc: Jennifer Timm, she/her, Rutgers - CABM, Piscataway
- Postdoc: Bipin Kumar Tripathi, he/him, Rutgers Busch Campus Waksman Institute of Microbiology
- postdoc: Jeesebastian, He/him, Public Health Research Institute , Rutgers University, Newark, Nj
- student: Bingyi Liang, she/her, NJIT
- Postdoc : Raja Khan :Rutgers, he/him - CABM, Piscataway
- Student: Lara Brindisi, she/her, Plant Biology PhD, lara.brindisi@rutgers.edu
- Postdoc: Maria Gomez, Rutgers Cancer Institute, mlg196@cinj.rutgers.edu
- 
- 
- 

Questions about the first R session:

Can you explain the line "less data/inflammation-01.csv"?

Results of day20 Exercise:

```
my_data[,20]
day20patients <- my_data[,20]
day20patients
```

```
patients_day20 <- my_data[, 20]
patients_day20
```

```
day_20 <- my_data[,20]
```

## Afternoon Coffee Break Feedback

Please, provide some feedback for this session through this google form before leaving for coffee break:  
<https://forms.gle/MJReYcZpzmvcUcL96>

Second exercise code:

```
patient3_day10 <- my_data [3 , 10:15]
```

```
patient3day10 <- my_data[3, 10:15]  
patient3day10
```

```
patient3 <- c(3)  
days10_15 <- c(10:15)  
my_data[patient3, days10_15]
```

```
patient_3 <- 10:15  
patient3_day10 <- my_data[3,patient_3]
```

```
my_data[3,10:15]
```

Exercise 3 code:

```
t(patient3)  
transposedpatient3 <- t(patient3)  
summary(transposedpatient3)
```

```
patient3 <- sum(patient3_day10)  
patient3  
patient3_average <- "patient3"/ 6  
patient3_average
```

Exercise 4 Code:

```
average_date_inflamation <- apply(X= my_data, MARGIN = 2, FUN = mean)
```

```
average_day_inflammation <- apply(X=my_data, MARGIN=2, FUN=mean)  
average_day_inflammation
```

```
average_inflammmation <- apply(X= my_data, MARGIN = 1, FUN = mean)
```

```
average_day_inflammation <- apply(X = my_data, MARGIN = 2, FUN = mean)  
average_day_inflammation
```

```
average_day_inflammation <- apply(X= my_data, MARGIN=1, mean)
```

```
averagedayinflam <- apply(X=my_data, MARGIN = 2, FUN = mean)
```

```

plot(apply(my_data, 1, max), ylab="Maximum Patient Inflammation")
plot(apply(my_data, 2, max), ylab="Maximum Inflammation per day")

plot(apply(my_data, 1, max)) #max inflammation per patient

max_inflammation <- apply(my_data, 2, max)
max_inflammation
plot(max_inflammation, xlab = "day", ylab = "Max inflammation")

plot(apply(my_data,1,max),xlab="day",ylab="max patient inflammation")

plot(apply(my_data,2,max), ylab = "Maximum Day Inflammation")
plot(apply(my_data,1,max), ylab = "Maximum Patient Inflammation")

```

Exercise 6: Create a function

```

function_name <- function() {
  print("Natalie")
}

```

```

function_name()

```

```

function_name <- function() {
  print("Maria")
}
function_name()

```

```

function_name_2 <- function(){
  1 + 2
}
function_name_2()

```

```

analyze <- function(my_file){
  internal_data <- read.csv(file = my_file)
  average_day <- apply(my_data,2,mean)
  min_day <- apply(my_data,2,min)
  max_day <- apply(my_data,2,max)
  plot(average_day, ylab = "Average Inflammation per Day", xlab = "Day")
  plot(min_day, ylab = "Minimum Inflammation per Day", xlab = "Day")
  plot(max_day, ylab = "Maximum Inflammation per Day", xlab = "Day")
}
analyze(my_file = "data/inflammation-01.csv")

```

```
Analyze <- function(filename) {
  the_data <- read.csv(file = filename)
  average_day <- apply(X=the_data,2,mean)
  min_day <- apply(X=the_data,2,min)
  max_day <- apply(X=the_data,2,max)
  plot(average_day)
  plot(min_day)
  plot(max_day)
}
Analyze("data/inflammation-01.csv")
```

End of Day Feedback

Please let us know how we are doing through this google form before leaving:

<https://forms.gle/nc5MqywcGKh3eszf8>

---



---

## Day 2

### Morning session - Version Control with Git

Follow the lesson here:

<https://swcarpentry.github.io/git-novice/>

Heather's Git slides:

[https://docs.google.com/presentation/d/1\\_LgvVayKjzjaokIIeZ1Q\\_5hZKqtl4QYVnGOGqz1thA/edit?usp=sharing](https://docs.google.com/presentation/d/1_LgvVayKjzjaokIIeZ1Q_5hZKqtl4QYVnGOGqz1thA/edit?usp=sharing)

### Sign in

**Sign in format: Name, Pronouns, Institution, Email (optional), Twitter (optional)**

- Instructor: Heather Ciallella, she/her, Rutgers University-Camden, hlciallella@gmail.com
- Helper: Jiayi Liu, She/Her, Rutgers, Molecular Biosciences, jl2791@rutgers.edu, <https://github.com/jiayiliujiayi>
- Helper: Kellen Xu
- Student: Kyla Johnson, she/her, Rutgers GSBS
- Alumnus: Iman Amer, she/her
- Omkaram Gangisetty, he/him, Rutgers university
- Student: Abby Heller She/Her, Rutgers-Molecular Biosciences PhD
- Postdoc: Jihyun Kim, sge/her, Rutgers-NB, CBN

- Keyaara Robinson, Pharmaceutical Sciences
- Postdoc: Jennifer Timm, Rutgers - CABM, Piscataway
- Postdoc: Edward Martinez, he/him, Rutgers-New Brunswick
- Postdoc : Raja Khan :Rutgers, he/him - CABM, Piscataway
- Postdoc: Dana Reid, she/her- Rutgers (Cook-SEBS)
- Student: Chang Liu, she/her, NJMS, cyl44@njms.rutgers.edu
- Postdoc: Aniket Bhattacharya, he/him, CHI, a.bhattacharya@rutgers.edu
- Postdoc: Guoqiang, he/him, Rutgers-New Brunswick, wangg5@dls.rutgers.edu
- student: Xia Qiu, she/her, Rutgers-Molecular Biosciences PhD
- Student: Mary Lally, she/her, Rutgers-Molecular Biosciences PhD
- student: Bingyi Liang, she/her, NJIT
- Postdoc: Maria Gomez - Rutgers Cancer Institute
- Post Doc: Jeesebastian, PHRI, Rutgers, Newark
- Postdoc:Bipin Kumar Tripathi, he/him, Rutgers Busch Campus Waksman Institute of Microbiology
- student: Natalie Losada, she/her, CABM, Busch, Rutgers
- Student: Lara Brindisi, she/her, Plant Biology PhD, lara.brindisi@rutgers.edu
- git config --global user.name "
- git config --global user.email
- git config --global core.autocrlf input
- git config --global core.autocrlf true
- git config --global core.editor "nano -w"
- git config --global init.defaultBranch main # to change to by default called "main" instead of "master"
- git config -h # to get help
- git checkout -b main # Switched to a new branch 'main'
- git status
  - create some files
- git add <file>
- git commit -m "Some message you want to say about this cimmit." # git requires you to add a sticknote message
- git log
- git status
  - some editing
- git diff
- git commit -m
- git diff --staged # to see what's changed in the staging files
- git log -1
- git log -2
- git log --oneline
- git diff HEAD mars.txt
- git diff HEAD~2 mars.txt
- git show HEAD~2 mars.txt
- git checkout HEAD~2 mars.tx
- git checkout main
- git revert
- git revert --abort

## Morning Break Feedback

Please provide feedback for this session through this google form before leaving for morning break:  
<https://forms.gle/FDEDjnW9JqXZLN3p6>

## Lunch Break Feedback

Please provide some more feedback through this google form before leaving for lunch:  
<https://forms.gle/HZr1dKjxY1tFBYmY6>

---

## Afternoon session day 2 - Programming with R (Continued)

Follow the lesson here:  
<http://swcarpentry.github.io/r-novice-inflammation/>

## Sign in

**Sign in format: Name, Pronouns, Institution, Email (optional), Twitter (optional)**

- Instructor: Luna L. Sanchez Reyes, she/her, University of California, Merced
- Helper: Heather Ciallella, she/her, Rutgers University-Camden
- Helper: Jiayi Liu, She/Her, Rutgers, Molecular Biosciences, jl2791@rutgers.edu, <https://github.com/jiayiliujiayi>
- Helper: Kellen Xu
- Research Associate: Omkaram Gangisetty (he/his) Rutgers university, og53@sebs.rutgers.edu
- Postdoc: Jennifer Timm, Rutgers-CABM, Piscataway
- Postdoc: Jihyun Kim (she/her), Rutgers University-NB, CBN
- Postdoc: Jeesebastian, PHRI, Rutgers
- Postdoc: Consuelo Ibar, she/her, Waksman Institute
- Student: Abby Heller She/Her, Rutgers-Molecular Biosciences PhD
- Alumnus: Iman Amer, she/her
- Student: Kyla Johnson, she/her, Rutgers- GSBS
- Postdoc: Edward Martinez, he/him, Rutgers-New Brunswick
- Keyaara Robinson, Pharmaceutical Sciences
- Student: Chang Liu, she/her, njms, cyl44@njms.rutgers.edu
- Postdoc: Bipin Kumar Tripathi, he/him, Rutgers Busch Campus Waksman Institute of Microbiology
- Postdoc: Aniket Bhattacharya (he/him), CHI, a.bhattacharya@rutgers.edu
- Student: Mary Lally, she/her, Rutgers-Molecular Biosciences PhD
- student: Bingyi Liang, she/her, NJIT

- Postdoc: Dana Reid, she/her, Rutgers (Cook-SEBS)
- student: Natalie Losada, she/her, CABM, Busch, Rutgers
- Postdoc: Maria Gomez, CINJ
- Postdoc : Raja Khan :Rutgers, he/him - CABM, Piscataway
- postdoc: Guoqiang Wang, he/him, Rutgers-New Brunswick, wangg5@dls.rutgers.edu
- Student: Lara Brindisi, she/her, Plant Biology PhD, lara.brindisi@rutgers.edu
- 
- 
- 
- 
- 
- 

## Exercise 2.1

# Exercise: Get all values from day15 by creating an object called day15

```
> day15 <- my_data[,15]
```

```
> day15
```

```
[1] 4 4 8 3 9 4 10 10 13 13 5 6 13 13 5 9 12 8 11 10 9 10 4 12 12
```

```
[26] 3 14 7 3 14 14 13 12 5 3 6 5 9 4 13 8 8 6 8 4 8 13 4 5 13
```

```
[51] 3 14 6 13 3 6 14 5 12
```

```
dim(my_data)
```

```
day15 <- my_data[,15:40]
```

```
day15 <- my_data$X7.1
```

```
day15
```

```
#1
```

```
colnames(my_data)
```

```
Day_15 <- my_data$X7.1
```

```
Day_15
```

```
#2
```

```
Day_15_2 <- my_data[,15]
```

```
Day_15_2
```

```
#3
```

```
Day_15_3 <- my_data[, "X7.1"]
```

```
Day_15_3
```

```
day15 <- my_data[, 15]
```

```
day15
```

```
day15 <- my_data[,15]
```

Does this work:

```
summary(t(my_data))
```

Yes!! This is a great solution!

```
# summary of all days
apply(my_data, 2, summary)
summary(my_data) # also works and less typing Yes! good work!
```

```
# summary of all patients
summary(t(my_data))
apply(my_data, 1, summary)
```

## Exercise 2.2

```
avg_day_inflammation <- apply(my_data, 2, mean)
min_day_inflammation <- apply(my_data, 2, min)
max_day_inflammation <- apply(my_data, 2, max)
```

```
avg_pt_inflammation <- apply(my_data, 1, mean)
min_pt_inflammation <- apply(my_data, 1, min)
max_pt_inflammation <- apply(my_data, 1, max)
```

```
#day wise data
avge_day_infla <- apply(my_data, 2, mean)
min_day_inf <- apply(my_data, 2, min)
max_day_infla <- apply(my_data, 2, max)
```

```
#patient wise
```

```
avge_patient_infla <- apply(my_data, 1, mean)
min_patient_inf <- apply(my_data, 1, min)
max_patient_infla <- apply(my_data, 1, max)
```

```
min_day_inflammation <- apply(my_data, 2, min)
max_day_inflammation <- apply(my_data, 2, max)
```

```
patient_mean_day_inflammation <- apply(my_data, 1, mean)
patient_min_day_inflammation <- apply(my_data, 1, min)
patient_max_day_inflammation <- apply(my_data, 1, max)
```

```
# Exercise 2.2: Create two objects one with minimum inflammation per day and another maximum
inflammation per day
```

```
min_day_inflammation <- apply(X = my_data, MARGIN = 2, FUN = min)
min_day_inflammation
max_day_inflammation <- apply(X = my_data, MARGIN = 2, FUN = max)
max_day_inflammation
```

```
# Per patient
```

```
mean_per_patient <- apply(X = my_data, MARGIN = 1, FUN = mean)
mean_per_patient
```

```
min_per_patient <- apply (X = my_data, MARGIN = 1, FUN = min)
min_per_patient
max_per_patient <- apply (X = my_data, MARGIN = 1, FUN = max)
max_per_patient
```

Explore more on color palettes:

<https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html>

Afternoon Coffee Break Feedback

Please, provide some feedback for this session through this google form before leaving for coffee break:

<https://forms.gle/oFj8c64mE8TeSCkc7>

Getting error bars for your points

<https://stackoverflow.com/questions/13032777/scatter-plot-with-error-bars>

Resources for calculating error bars to histograms

# Solutions for exercise 1.7

```
filename = read.csv()
```

```
analyze <- function(filename) {
  # read the inflammation files
  internal_data <- read.csv(file = filename)
  head(internal_data)
}
```

```
analyze <- function(filename, output_name) {
  read_data <- read.csv(file = filename) # my_data
  mean_data <- apply(read_data, 2, mean)
  min_data <- apply(read_data, 2, min)
  max_data <- apply(read_data, 2, max)
  pdf(output_name)
  plot(mean_data,
        xlab = "Day",
        ylab = "Inflammation",
        ylim = c(0, 30),
        col = "#FF0000",
        pch = 1)
  par(new = TRUE)
  plot(min_data,
        xlab = "",
        ylab = "",
        ylim = c(0, 30),
        pch = 6,
        col = "#FF8000")
  par(new = TRUE)
  plot(max_data,
        xlab = "",
```

```

ylab = "",
col = "#FFFF00",
ylim = c(0, 30),
pch = 5)
# Adding a legend()
legend(x = 25.15,
      y = 30.12,
      legend = c("Minimum", "Average", "Maximum"),
      cex = 0.5,
      pch = c(6, 1, 5),
      col = c("#FF8000", "#FF0000", "#FFFF00"))
dev.off( )
}

```

```

analyze(filename = "data/inflammation-01.csv")
analyze(filename = "data/inflammation-04.csv")

```

## Post-workshop Survey

Please let us know how we did by completing the post-workshop survey at  
<https://carpentries.typeform.com/to/UgVdRQ#slug=2022-03-17-rutgers-online>

This is the last one! Well done!!