Week-2: Code-along

NM2207: Computational Media Literacy

2023-08-17

Welcome! Go through the steps described below, *carefully*. It is totally fine to get stuck - **ASK FOR HELP**; reach out to your friends, TAs, or the discussion forum on Canvas.

Here is what you have to do,

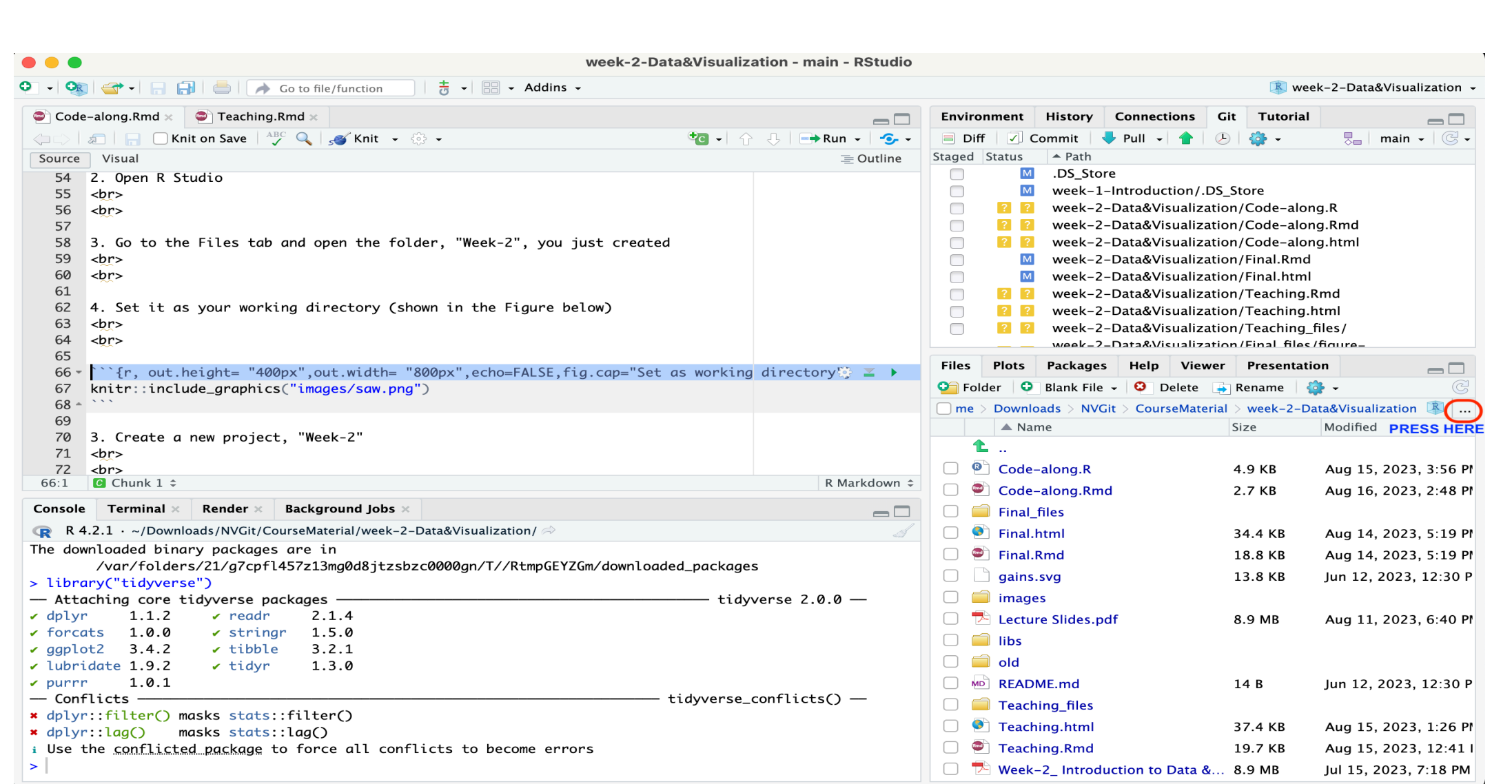
1. **Listen** to the video lectures, and while doing so,
2. **Follow** the instructions in this file

I. Preliminaries

1. Create a new folder, "Week-2", inside "NM2207" folder you created last week

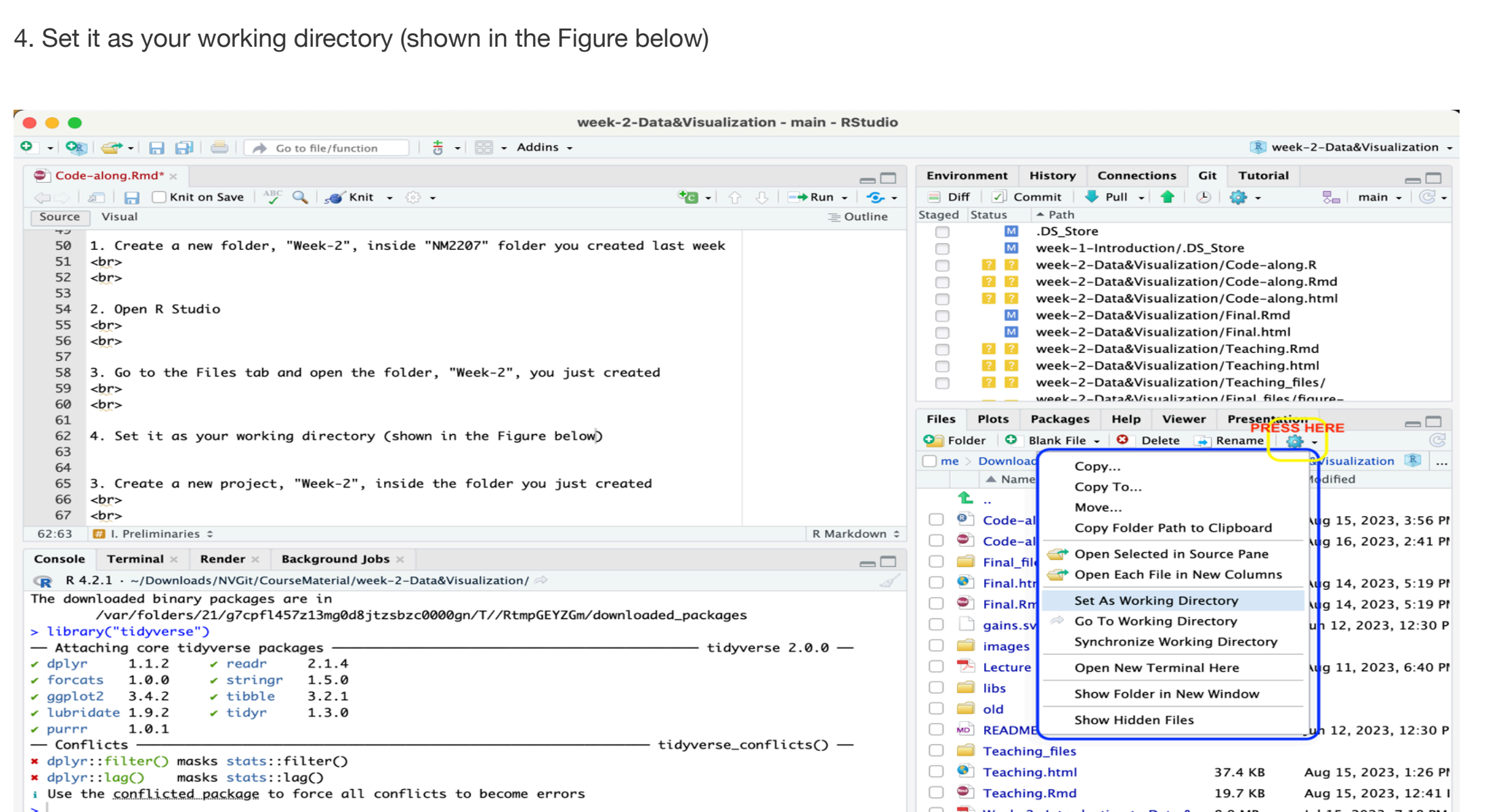
2. Open R Studio

3. Go to the Files tab and open the folder, "Week-2", you just created
   - Press the three horizontal dots highlighted in the Figure below
   - Browse and select "Week-2" folder inside "NM2207"



Navigating folders

4. Set it as your working directory (shown in the Figure below)



Set as working directory

5. Now, create a new project and name it "Week-2"

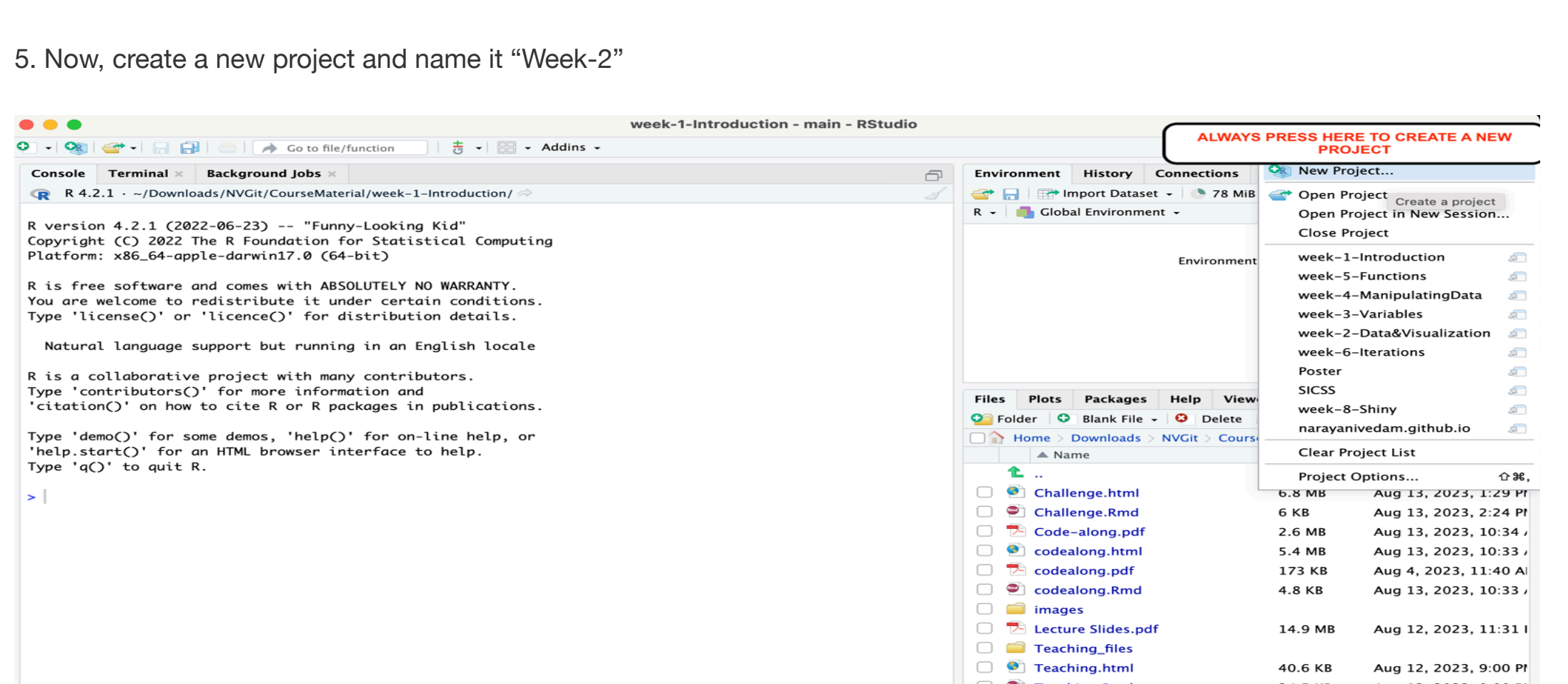
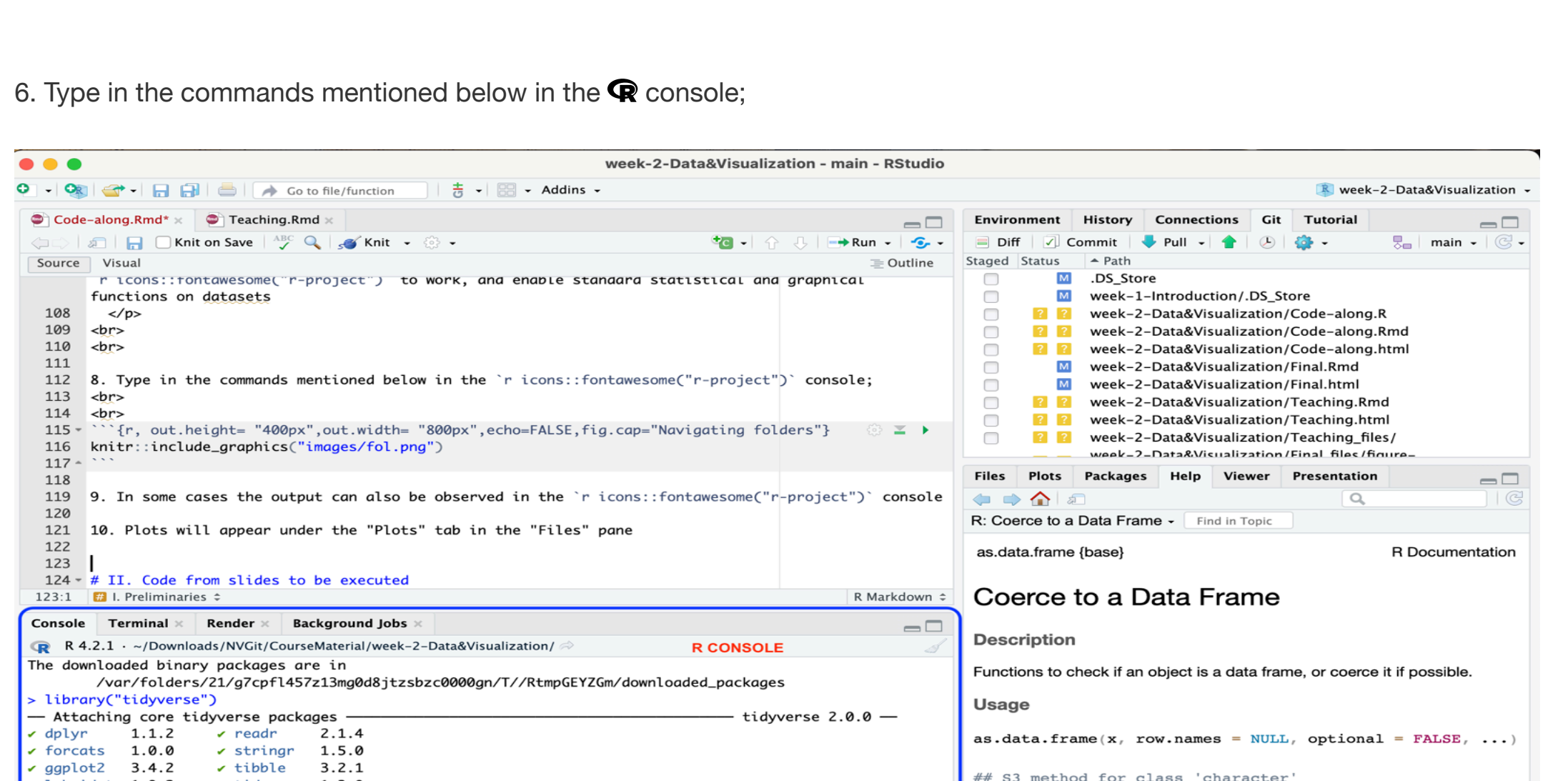


Figure: Creating a new project

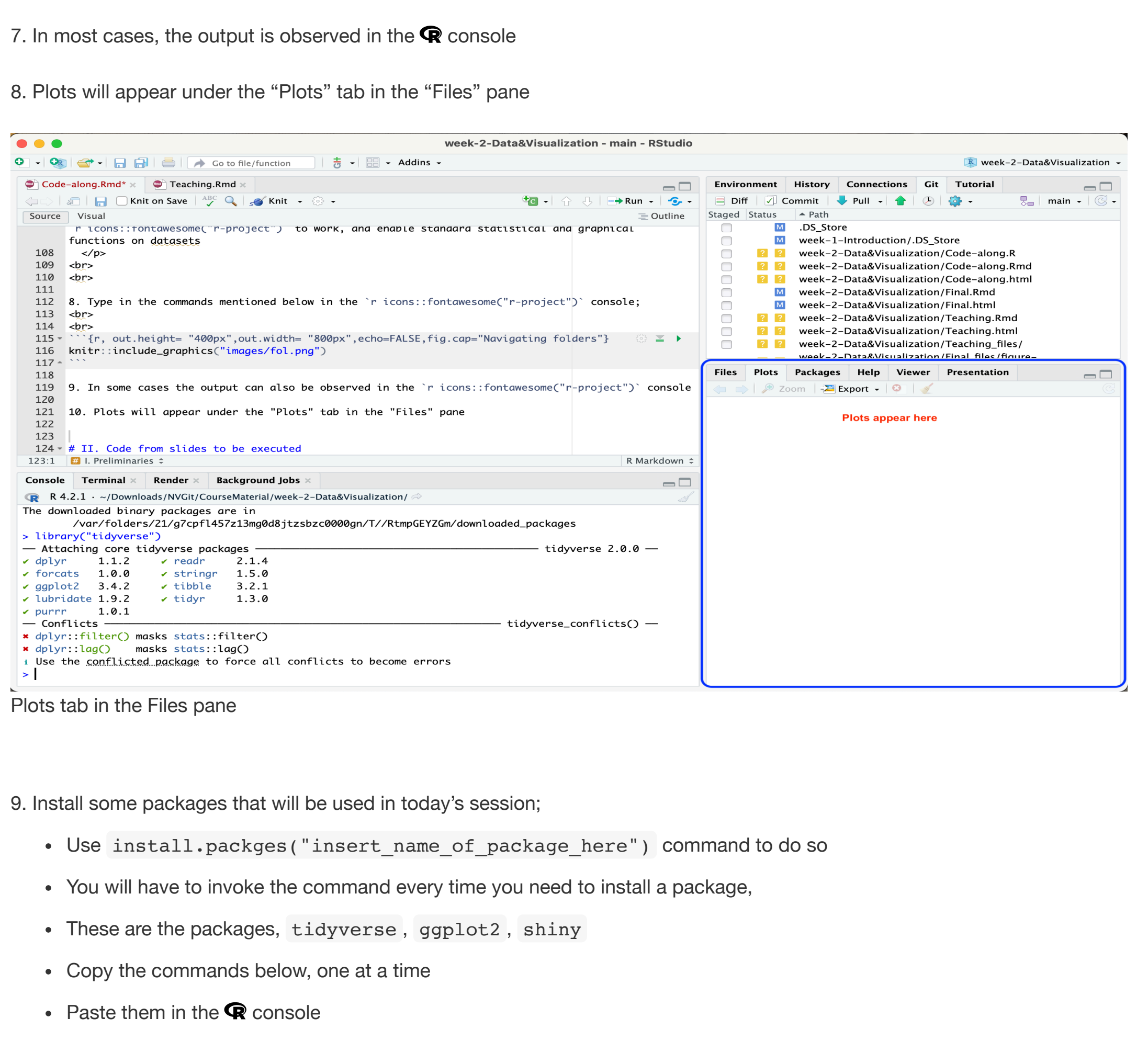
6. Type in the commands mentioned below in the R console;



R Console

7. In most cases, the output is observed in the R console

8. Plots will appear under the "Plots" tab in the "Files" pane



Plots tab in the Files pane

9. Install some packages that will be used in today's session;
   - Use `install.packges("insert_name_of_package_here")` command to do so
   - You will have to invoke the command every time you need to install a package,
   - These are the packages, `tidyverse`, `ggplot2`, `shiny`
   - Copy the commands below, one at a time
   - Paste them in the R console
   - Press enter
   - Messages will be printed on your screen about the progress
   - Here are the commands to install R packages
     - `install.packages("tidyverse")`
     - `install.packages("ggplot2")`
     - `install.packages("shiny")`

> *What are **packages**?*
> *Packages are collections of R functions, data, and compiled code in a well-defined format, created to add specific functionality. There are 10,000+ user contributed packages and growing.*
> *There are a set of standard (or base) packages which are considered part of the R source code and automatically available as part of your R installation. Base packages contain the basic functions that allow R to work, and enable standard statistical and graphical functions on datasets*

10. Once you complete the installation of packages, load them;
    - Use `library(insert_name_here)` command to do so
    - Copy the commands below, one at a time
    - Paste them in the R console
    - Press enter
    - Here are the commands,
      - `library(tidyverse)`
      - `library(ggplot2)`
      - `library(shiny)`

II. Code from slides to be executed

1. Copy the commands inside the shaded box
2. Paste them in the R console, one at a time
3. Press enter and observe the output
4. Pay attention to what each line of code manages to execute

Slide Number 5

```
# Load R packages for data science
library(tidyverse)
# Data in starwars data-set
starwars
```

Slide Number 7

```
# Salient features of the data-set
?starwars
```

Slide Number 8

```
# Catch a glimpse starwars data-set
glimpse(starwars)
```

Slide Number 10

```
# Access column "height"
starwars$height
```

Slide Number 11

```
# Access column "mass"
starwars$mass
```

Slide Number 12

```
# Access column "gender"
starwars$gender
```

Slide Number 13

```
# Access column "gender"
starwars$films[1:3]
```

Slide Number 14

```
# Rows of interest
filter_rows <- c("Luke Skywalker","R2-D2","Darth Vader") # Extract row corresponding to Luke Skywalker
starwars %>% filter(name%in%filter_rows)
```

Slide Number 15

```
# Rows of interest
filter_rows <- c("Luke Skywalker","R2-D2")
# Extract row corresponding to Luke Skywalker
starwars %>% filter(name%in%filter_rows) %>% select(name,height,mass,homeworld,films)
```

Slide Number 16

```
# Rows of interest
filter_rows <- c("Luke Skywalker","R2-D2")
# Extract rows in `rows`
starwars %>% filter(name%in%filter_rows) %>% pull(films)
```

pull extracts 'film' columns for filtered rows

pull() only pulls out 1 column

Slide Number 17

```
# Number of rows in the data-set
nrow(starwars)
ncol(starwars)
dim(starwars)
```

no. of rows first then column

Slide Number 23

```
# Invoke the library
library(Tmisc)
# Filter data-set I in quartet
quartet %>% filter(set=="I")
```

```
# Invoke the library
library(Tmisc)
# Filter data-set I in quartet
quartet %>% filter(set=="II")
```

Slide Number 24

```
# Invoke the library
library(Tmisc)
# Filter data-set I in quartet
quartet %>% filter(set=="III")
```

```
# Invoke the library
library(Tmisc)
# Filter data-set I in quartet
quartet %>% filter(set=="IV")
```

Slide Number 25

```
# Obtain the needed statistics
grouped_quartet %>%
  summarise(
    mean_x = mean(x),
    mean_y = mean(y),
    sd_x = sd(x),
    sd_y = sd(y),
r = cor(x, y) )
```

Slide Number 29

```
# Plot the data
ggplot(data=starwars)
```

Slide Number 30

```
# Plot height along x-axis
ggplot(data=starwars,mapping=aes(x=height))
```

Slide Number 31

```
# Plot mass along y-axis
ggplot(data=starwars,mapping=aes(x=height,y=mass))
```

Slide Number 32

```
ggplot(data=starwars,mapping=aes(x=height,y=mass)) +
geom_point()
```

Slide Number 33

```
ggplot(data=starwars,mapping=aes(x=height,y=mass)) + geom_point() +
labs(x="Height (cm)",y="Weight (Kg)")
```

Slide Number 34

```
ggplot(data=starwars,mapping=aes(x=height,y=mass)) +
  geom_point() +
  labs(x="Height (cm)",y="Weight (Kg)",
title="Mass versus Height")
```

Slide Number 35

```
ggplot(data=starwars,mapping=aes(x=height,y=mass)) +
  geom_point() +
  labs(x="Height (cm)",y="Weight (Kg)",
title="Mass versus Height",
caption="Source: tidyverse/ starwars dataset")
```

Slide Number 40

```
# Install package
install.packages("shiny") # Invoke the package
library(shiny)
# Run an example from the library
runExample("01_hello")
```

Slide Number 42

```
# Install package
install.packages("shiny") # Invoke the package
library(shiny)
# Run an example from the library
runExample("06_tabsets")
```