## 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

week-2-Data&Visualization -

week-2-Data&Visualization -

Environment History Connections Git Tutorial

week-1-Introduction/.DS\_Store week-2-Data&Visualization/Code-along.R

□ Diff
✓ Commit
→ Pull
→
♠
✓ L
♠
♠
✓ -

.DS\_Store

**Environment History Connections** 

■ Diff 🗸 Commit 🕨 Pull 🗸 👚 🕗 🥸 🕶

 week−1-Introduction/.DS\_Store week-2-Data&Visualization/Code-along.R

week-2-Data&Visualization/Code-along.Rmd week-2-Data&Visualization/Code-along.html

week-2-Data&Visualization/Final.Rmd

week-2-Data&Visualization/Final.html

week-2-Data&Visualization/Final.html
week-2-Data&Visualization/Teaching.Rmd
week-2-Data&Visualization/Teaching.html
week-2-Data&Visualization/Teaching\_files/

.DS\_Store

Staged Status - Path

Staged Status A Path

.DS\_Store

week-1-Introduction/.DS\_Store week-2-Data&Visualization/Code-along.R 🔞 🔞 week-2-Data&Visualization/Code-along.Rmd

😢 😢 week-2-Data&Visualization/Code-along.html week-2-Data&Visualization/Final.Rmd

week-2-Data&Visualization/Final.html week-2-Data&Visualization/Teaching.Rmd

week\_2\_Data&Visualization/Final\_files/figu

week-2-Data&Visualization/Teaching.html

week-2-Data&Visualization/Teaching\_files/

Outline

☐ Import Dataset → ↑ 78 MiB

R → ☐ Global Environment → Open Project in New Session.

Close Project

week-1-Introduction week-5-Functions

week-3-Variables week-2-Data&Visualization

week-4-ManipulatingData

🚚 main 🗸 🥝 🗸

Staged Status

Outline

## I. Preliminaries

2. Open R Studio

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

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"

 

 Image: Comparit of the control of ≣ Outline Staged Status ▲ Path Source Visual .DS\_Store 54 2. Open R Studio ■ week-1-Introduction/.DS\_Store week-2-Data&Visualization/Code-along.R week-2-Data&Visualization/Code-along.Rmd 3. Go to the Files tab and open the folder, "Week-2", you just created week-2-Data&Visualization/Code-along.html week-2-Data&Visualization/Final.Rmd
 week-2-Data&Visualization/Final.html week-2-Data&Visualization/Teaching.Rmd 62 4. Set it as your working directory (shown in the Figure below) week-2-Data&Visualization/Teaching.html week-2-Data&Visualization/Teaching\_files/ 64 <br> week-2-Data&Visualization/Final files/figure-66 \* ```{r, out.height= "400px",out.width= "800px",echo=FALSE,fig.cap="Set as working directory > Files Plots Packages Help Viewer Presentation Folder Delete Rename knitr::include\_graphics("images/saw.png") 3. Create a new project, "Week-2" 71 <<u>br</u>> 4.9 KB Aug 15, 2023, 3:56 Pi Code-along.Rmd Aug 16, 2023, 2:48 Pf Render × Background Jobs ☐ ☐ Final\_files R 4.2.1 · ~/Downloads/NVGit/CourseMaterial/week-2-Data&Visualization/ ☐ Final.html Aug 14, 2023, 5:19 Pf 34.4 KB The downloaded binary packages are in ☐ **Pinal.Rmd** 18.8 KB Aug 14, 2023, 5:19 Pf /var/folders/21/g7cpfl457z13mg0d8jtzsbzc0000gn/T//RtmpGEYZGm/downloaded\_packages 13.8 KB Jun 12, 2023, 12:30 P gains.svg > library("tidyverse") — Attaching core tidyverse packages images — tidyverse 2.0.0 — ✓ dplyr readr 2.1.4 1.1.2 Aug 11, 2023, 6:40 Pf ✓ forcats 1.0.0 ✓ stringr 1.5.0 ✓ tibble 3.2.1 ✓ aaplot2 3.4.2 ✓ lubridate 1.9.2 🗸 tidyr 1.3.0 README.md ✓ purrr 1.0.1 Jun 12, 2023, 12:30 P - Conflicts tidyverse\_conflicts() — Teaching\_files \* dplyr::filter() masks stats::filter() Teaching.html 37.4 KB Aug 15, 2023, 1:26 PM \* dplyr::lag() masks stats::lag() Teaching.Rmd Aug 15, 2023, 12:41 I i Use the conflicted package to force all conflicts to become errors 19.7 KB ☐ Week-2\_ Introduction to Data &... 8.9 MB Jul 15, 2023, 7:18 PM Navigating folders

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

🔁 - | ↑ 🕕 | → Run - | 🦫 -Source Visual 1. Create a new folder, "Week-2", inside "NM2207" folder you created last week

52 <br> week-2-Data&Visualization/Code-along.Rmd 53 week-2-Data&Visualization/Code-along.html 54 2. Open R Studio week-2-Data&Visualization/Final.Rmd 55 <br> week-2-Data&Visualization/Final.html 56 <br> week-2-Data&Visualization/Teaching.Rmd week-2-Data&Visualization/Teaching.html 58 3. Go to the Files tab and open the folder, "Week-2", you just created week-2-Data&Visualization/Teaching\_files/ week\_2\_Data&Visualization/Final\_files/figure Packages Help Viewer Presentation PRESS HERE 62 4. Set it as your working directory (shown in the Figure below) Folder Blank File - Delete 🙀 Rename 🙋 -Copy... 65 3. Create a new project, "Week-2", inside the folder you just created Copy To... **1** .. 66 <br> 67 <br> g 15, 2023, 3:56 Pl Copy Folder Path to Clipboard 62:63 # I. Preliminaries \$ ☐ **②** Code−al g 16, 2023, 2:41 Pi TOPEN Selected in Source Pane Console Terminal × Render × Background Jobs × ☐ **i** Final\_file TOPEN Each File in New Columns R 4.2.1 · ~/Downloads/NVGit/CourseMaterial/week-2-Data&Visualization/ ☐ **●** Final.htr g 14, 2023, 5:19 Pi The downloaded binary packages are in Set As Working Directory g 14, 2023, 5:19 Př /var/folders/21/g7cpfl457z13mg0d8jtzsbzc0000gn/T//RtmpGEYZGm/downloaded\_packages Go To Working Directory n 12, 2023, 12:30 P > library("tidyverse") images Synchronize Working Directory — Attaching core tidyverse packages Lecture g 11, 2023, 6:40 Pi Open New Terminal Here ✓ forcats 1.0.0 libs Show Folder in New Window ✓ tibble 3.2.1 old ✓ lubridate 1.9.2 🗸 tidyr Show Hidden Files README 1.0.1 n 12, 2023, 12:30 P — Conflicts — - tidyverse\_conflicts() — Teaching\_files \* dplyr::filter() masks stats::filter() Teaching.html Aug 15, 2023, 1:26 Pf i Use the <u>conflicted package</u> to force all conflicts to become errors Teaching.Rmd 19.7 KB Aug 15, 2023, 12:41 I ☐ № Week-2\_ Introduction to Data &... 8.9 MB Jul 15, 2023, 7:18 PM Set as working directory 5. Now, create a new project and name it "Week-2" week-1-Introduction - main - RStudio ALWAYS PRESS HERE TO CREATE A NEW 

week-2-Data&Visualization - main - RStudio

Copyright (C) 2022 The R Foundation for Statistical Computing Platform: x86\_64-apple-darwin17.0 (64-bit) R is free software and comes with ABSOLUTELY NO WARRANTY. You are welcome to redistribute it under certain conditions. Type 'license()' or 'licence()' for distribution details.

R 4.2.1 · ~/Downloads/NVGit/CourseMaterial/week-1-Introduction/

R version 4.2.1 (2022-06-23) -- "Funny-Looking Kid"

Console Terminal  $\times$  Background Jobs  $\times$ 

Natural language support but running in an English locale Poster R is a collaborative project with many contributors. Type 'contributors()' for more information and SICSS Files Plots Packages Help View 'citation()' on how to cite R or R packages in publications. week-8-Shiny Folder □ Blank File → □ Delete narayanivedam.github.io Type 'demo()' for some demos, 'help()' for on-line help, or □ 
 ↑ Home > Downloads > NVGit > Course 'help.start()' for an HTML browser interface to help. Clear Project List Type 'q()' to quit R. Challenge.html 6.8 MB Aug 13, 2023, 1:29 Pf Challenge.Rmd Aug 13, 2023, 2:24 Pf Code-along.pdf 2.6 MB Aug 13, 2023, 10:34 codealong.pdf 173 KB Aug 4, 2023, 11:40 A codealong.Rmd 4.8 KB Aug 13, 2023, 10:33 / images Lecture Slides.pdf 14.9 MB Aug 12, 2023, 11:31 I Teaching.html 40.6 KB Aug 12, 2023, 9:00 Pl Teaching.Rmd 24.5 KB Aug 12, 2023, 9:00 Př Untitled files Untitled.html 35.7 KB Aug 12, 2023, 11:30 I Untitled.Rmd 22.9 KB Aug 12, 2023, 11:30 I Figure: Creating a new project 6. Type in the commands mentioned below in the  $\mathbf{Q}$  console; week-2-Data&Visualization - main - RStudio ○ - On of the function Code-along.Rmd\* × Teaching.Rmd × Environment History Connections Git Tutorial

**\***□ • | ↑ ↓ | → Run • | • • •

114 <br> 115 - ```{r, out.height= "400px",out.width= "800px",echo=FALSE,fig.cap="Navigating folders"} 116 knitr::include\_graphics("images/fol.png")

 $\verb"ricons::tontawesome("r-project")" to work, and enable standard statistical and graphical \\$ 

112 8. Type in the commands mentioned below in the `r icons::fontawesome("r-project")` console;

functions on datasets

109 <br>

110 <br>

113 <br>

Source Visual

110 <br>

114 <br>

111

week-2-Data&Visualization/Final files/figure-Files Plots Packages Help Viewer Presentation 9. In some cases the output can also be observed in the `r icons::fontawesome("r-project")` console R: Coerce to a Data Frame - Find in Topic 121 10. Plots will appear under the "Plots" tab in the "Files" pane as.data.frame {base} R Documentation 123 | 124 - # II. Code from slides to be executed R Markdown \$ Coerce to a Data Frame 123:1 # I. Preliminaries \$ Console Terminal × Render × Background Jobs  $ightharpoonup R 4.2.1 \cdot \sim /Downloads/NVGit/CourseMaterial/week-2-Data&Visualization/ <math>\approx$ The downloaded binary packages are in Functions to check if an object is a data frame, or coerce it if possible. /var/folders/21/g7cpfl457z13mg0d8jtzsbzc0000gn/T//RtmpGEYZGm/downloaded\_packages > library("tidyverse") — tidyverse 2.0.0 — Attaching core tidyverse packages ✓ dplyr 1.1.2 ✓ readr 2.1.4 ✓ forcats 1.0.0 ✓ stringr 1.5.0 dplyr as.data.frame(x, row.names = NULL, optional = FALSE, ...) ✓ ggplot2 3.4.2 ✓ tibble 3.2.1 ✓ lubridate 1.9.2 tidyr as.data.frame(x, ..., ✓ purrr 1.0.1 stringsAsFactors = FALSE) tidyverse\_conflicts() — \* dplyr::filter() masks stats::filter() \* dplyr::lag() masks stats::lag() ## S3 method for class 'list' i Use the conflicted package to force all conflicts to become errors as.data.frame(x, row.names = NULL, optional = FALSE, ..., cut.names = FALSE, col.names = names(x), fix R Console 7. In most cases, the output is observed in the  $\mathbf{Q}$  console 8. Plots will appear under the "Plots" tab in the "Files" pane week-2-Data&Visualization - main - RStudio ◆ Go to file/function 🗦 🗸 🔡 - Addins week-2-Data&Visualization Code-along.Rmd\* x
□ Teaching.Rmd > Environment History Connections Git Tutorial 🗸 📗 🗌 🗀 Knit on Save | 💯 🔍 | 🦋 Knit 🔻 🔅 🔻 👣 - | ↑ 🕕 | 🗪 Run - | 🦫 - | 🚚 | main 🗸 | 🕝 🗸

115 - ```{r, out.height= "400px",out.width= "800px",echo=FALSE,fig.cap="Navigating folders"} 116 knitr::include\_graphics("images/fol.png") 117 -

r icons::tontawesome("r-project") to work, and enable standard statistical and graphical

112 8. Type in the commands mentioned below in the `r icons::fontawesome("r-project")` console;

Files Plots Packages Help Viewer Presentation 9. In some cases the output can also be observed in the `r icons::fontawesome("r-project")` console 🇼 | 🔑 Zoom | -- Export 🗸 🗯 121 10. Plots will appear under the "Plots" tab in the "Files" pane Plots appear here 124 - # II. Code from slides to be executed 123:1 # I. Preliminaries \$ R Markdown Console Terminal × Render × Background Jobs × R 4.2.1 · ~/Downloads/NVGit/CourseMaterial/week-2-Data&Visualization/ The downloaded binary packages are in /var/folders/21/g7cpfl457z13mg0d8jtzsbzc0000gn/T//RtmpGEYZGm/downloaded\_packages > library("tidyverse") — Attaching core tidyverse packages – 1.1.2 🗸 readr 2.1.4 ✓ forcats 1.0.0 ✓ stringr 1.5.0 ✓ ggplot2 3.4.2 ✓ tibble 3.2.1 ✓ lubridate 1.9.2 🗸 tidyr 1.0.1 - Conflicts -· tidyverse\_conflicts() — \* dplyr::filter() masks stats::filter() \* dplyr::lag() masks stats::lag() i Use the conflicted package to force all conflicts to become errors 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

 Press enter Messages will be printed on your screen about the progress

datasets

Press enter

• Paste them in the **Q** console

o install.packages("tidyverse")

o install.packages("shiny")

- o install.packages("ggplot2")
- What are packages? Packages are collections of **Q** functions, data, and compiled code in a well-defined format, created to add specific functionality. There are 10,000+ user contributed packages and growing.

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 **Q** console

· Here are the commands, o library(tidyverse) o library(ggplot2) o library(shiny)

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 to work, and enable standard statistical and graphical functions on

3. Press enter and observe the output 4. Pay attention to what each line of code manages to execute Slide Number 5

1. Copy the commands inside the shaded box

2. Paste them in the **Q** console, one at a time

# Load R packages for data science

# Catch a glimpse starwars data-set

# Data in starwars data-set

library(tidyverse)

Slide Number 7

Slide Number 8

glimpse(starwars)

Slide Number 11

starwars\$mass

# Access column "mass"

Slide Number 13

starwars\$films[1:3]

Slide Number 14

# Rows of interest

# Rows of interest

Slide Number 16

# Rows of interest

# Access column "gender"

starwars

II. Code from slides to be executed

# Salient features of the data-set ?starwars

Slide Number 10 # Access column "height" starwars\$height

Slide Number 12 # Access column "gender" starwars\$gender

filter\_rows <- c("Luke Skywalker","R2-D2")</pre>

filter\_rows <- c("Luke Skywalker","R2-D2")</pre>

starwars %>% filter(name%in%filter\_rows) %>% pull(films)

# Extract row corresponding to Luke Skywalker

starwars %>% filter(name%in%filter\_rows) Slide Number 15

starwars %>% filter(name%in%filter\_rows) %>% select(name,height,mass,homeworld,films)

filter\_rows <- c("Luke Skywalker", "R2-D2", "Darth Vader") # Extract row corresponding to Luke Skywalker

Slide Number 17 # Number of rows in the data-set

nrow(starwars)

ncol(starwars) dim(starwars)

library(Tmisc)

Slide Number 23

# Invoke the library

Slide Number 24

# Invoke the library

library(Tmisc)

summarise(

r = cor(x, y)

mean x = mean(x),

 $mean_y = mean(y)$ ,

 $sd_x = sd(x)$ , sd y = sd(y),

Slide Number 31

Slide Number 33

Slide Number 34

Slide Number 35

runExample("06\_tabsets")

labs(x="Height (cm)",y="Weight (Kg)")

# Plot mass along y-axis

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

# Filter data-set I in quartet

# Filter data-set I in quartet

# Filter data-set I in quartet

quartet %>% filter(set=="III")

quartet %>% filter(set=="II")

# Extract rows in 'rows'

quartet %>% filter(set=="I") # Invoke the library library(Tmisc)

# Invoke the library library(Tmisc) # Filter data-set I in quartet quartet %>% filter(set=="IV") Slide Number 25 # Obtain the needed statistics grouped\_quartet %>%

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 32 ggplot(data=starwars,mapping=aes(x=height,y=mass)) + geom\_point()

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

ggplot(data=starwars,mapping=aes(x=height,y=mass)) + geom\_point() +

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