

# Chapter 1 Introduction to R and RStudio

Angelo LaCommare-Soto

2025-01-29

## Section 1 - anole

```
# set working directory for all chunks in this file (default working directory is wherever Rmd file is)  
getwd()
```

```
## [1] "C:/Users/Angelo L/Documents/GitHub/BIOL710/RCode710/RCode/working_directory"
```

## Sections 2 and 3 - anole

These sections were completed by working on a copy of the RMarkdown Lab Report template and continuously adding comments to my code.

## Section 4 - anole

```
#import data from .csv file  
#commands header=TRUE in order to treat the first row of the data frame as a header and stringsAsFactors=TRUE  
anole <- read.csv('anole.csv', header=TRUE, stringsAsFactors=TRUE)
```

## Section 5 - anole

```
#overview of the dataset anole  
str(anole)
```

```
## 'data.frame':    164 obs. of  21 variables:  
## $ ID             : Factor w/ 164 levels "537","539","540",...: 1 2 3 4 5 6 7 8 9 10 ...  
## $ Hurricane      : Factor w/ 2 levels "After","Before": 1 1 1 1 1 1 1 1 1 1 ...  
## $ Origin         : Factor w/ 2 levels "Pine Cay","Water Cay": 1 1 1 1 1 1 1 1 1 1 ...  
## $ Sex            : Factor w/ 2 levels "Female","Male": 2 1 2 1 1 1 2 2 1 2 ...  
## $ SVL            : num  48.7 40.3 58.3 43.1 45.5 ...  
## $ Femur           : num  10.39 8.66 12.87 8.55 10.26 ...  
## $ Tibia           : num  11.87 9.79 14.76 10.29 11.02 ...  
## $ Metatarsal      : num  7.52 6.18 9.45 6.6 6.89 6.85 7.9 8.24 6.88 7.77 ...  
## $ LongestToe       : num  7.43 6.2 9.58 6.26 7.02 7.18 8.23 8.02 6.7 7.7 ...  
## $ Humerus         : num  8.66 8.01 11.72 7.43 7.71 ...  
## $ Radius          : num  7.99 6.51 9.54 6.6 7.25 7.15 8.4 8.79 7.11 8.7 ...
```

```
## $ Metacarpal : num 2.22 2.38 3.54 2.79 2.52 2.39 3.15 3.18 2.82 3.05 ...
## $ LongestFinger: num 3.19 3.55 5.09 3.55 3.37 3.26 4.3 4.2 3.36 4.12 ...
## $ FingerCount : int 10 10 14 11 11 12 11 12 12 12 ...
## $ ToeCount : int 12 13 15 12 13 14 14 12 13 16 ...
## $ FingerArea1 : num 1.321 0.962 2.507 1.172 1.376 ...
## $ FingerArea2 : num 1.34 0.95 2.7 1.18 1.36 ...
## $ FingerArea3 : num 1.339 0.972 2.685 1.186 1.42 ...
## $ ToeArea1 : num 2.53 1.5 4.16 1.9 2.63 ...
## $ ToeArea2 : num 2.4 1.52 4.14 1.87 2.44 ...
## $ ToeArea3 : num 2.37 1.53 4 1.87 2.53 ...
```

## Question Answers

- The dataset “anole.csv” is of the class data frame
- “anole.csv” contains 164 total observations of 21 variables
- Within the dataset “anole.csv”, variables labeled ‘ID’, ‘Hurricane’, ‘Origin’, and ‘Sex’ are of the class factor, while variables labeled ‘SVL’, ‘Femur’, ‘Tibia’, ‘Metatarsal’, ‘LongestToe’, ‘Humerus’, ‘Radius’, ‘Metacarpal’, ‘LongestFinger’, ‘FingerArea1’, ‘FingerArea2’, ‘FingerArea3’, ‘ToeArea1’, ‘ToeArea2’, and ‘ToeArea3’ are of the class numerical, and variables labeled ‘FingerCount’ and ‘ToeCount’ are of the class integer.

```
# levels of the variable Sex, the $ sign means "within"
levels(anole$Sex)
```

```
## [1] "Female" "Male"
```

```
# summary of each variable
summary(anole)
```

```
##      ID      Hurricane      Origin      Sex      SVL
## 537 : 1 After :93 Pine Cay :79 Female:72 Min. :40.17
## 539 : 1 Before:71 Water Cay:85 Male :92 1st Qu.:43.48
## 540 : 1 Median :49.09
## 541 : 1 Mean :49.55
## 542 : 1 3rd Qu.:54.45
## 543 : 1 Max. :62.44
## (Other):158
##      Femur      Tibia      Metatarsal      LongestToe
## Min. : 7.900 Min. : 9.36 Min. :5.860 Min. :5.400
## 1st Qu.: 9.297 1st Qu.:10.30 1st Qu.:6.593 1st Qu.:6.367
## Median :10.715 Median :12.21 Median :7.740 Median :7.210
## Mean :10.919 Mean :12.03 Mean :7.544 Mean :7.274
## 3rd Qu.:12.377 3rd Qu.:13.32 3rd Qu.:8.383 3rd Qu.:8.037
## Max. :15.380 Max. :15.45 Max. :9.620 Max. :9.660
##
##      Humerus      Radius      Metacarpal      LongestFinger
## Min. : 6.770 Min. : 5.980 Min. :1.780 Min. :2.600
## 1st Qu.: 7.668 1st Qu.: 6.768 1st Qu.:2.500 1st Qu.:3.275
## Median : 9.125 Median : 7.885 Median :2.775 Median :3.825
## Mean : 8.996 Mean : 7.848 Mean :2.788 Mean :3.817
## 3rd Qu.:10.072 3rd Qu.: 8.745 3rd Qu.:3.080 3rd Qu.:4.230
## Max. :12.180 Max. :10.310 Max. :3.730 Max. :5.410
```

```
##
##   FingerCount      ToeCount      FingerArea1      FingerArea2
##   Min.       : 9.00    Min.       :11.00    Min.       :0.761    Min.       :0.414
##   1st Qu.:11.00    1st Qu.:13.00    1st Qu.:1.077    1st Qu.:1.067
##   Median :12.00    Median :14.00    Median :1.547    Median :1.535
##   Mean    :11.66    Mean    :13.98    Mean    :1.573    Mean    :1.566
##   3rd Qu.:12.00    3rd Qu.:15.00    3rd Qu.:1.996    3rd Qu.:1.995
##   Max.     :14.00    Max.     :16.00    Max.     :3.019    Max.     :3.010
##   NA's     :1      NA's     :1      NA's     :1      NA's     :1
##   FingerArea3      ToeArea1      ToeArea2      ToeArea3
##   Min.       :0.756    Min.       :1.122    Min.       :1.127    Min.       :1.116
##   1st Qu.:1.076    1st Qu.:1.702    1st Qu.:1.711    1st Qu.:1.689
##   Median :1.554    Median :2.469    Median :2.435    Median :2.468
##   Mean    :1.572    Mean    :2.526    Mean    :2.524    Mean    :2.512
##   3rd Qu.:2.001    3rd Qu.:3.260    3rd Qu.:3.240    3rd Qu.:3.220
##   Max.     :3.065    Max.     :4.798    Max.     :4.882    Max.     :4.839
##   NA's     :1      NA's     :1      NA's     :1      NA's     :1
```

```
# first rows of anole
head(anole)
```

```
##   ID Hurricane      Origin      Sex      SVL Femur Tibia Metatarsal LongestToe Humerus
## 1 537      After Pine Cay      Male 48.69 10.39 11.87         7.52         7.43      8.66
## 2 539      After Pine Cay     Female 40.31  8.66  9.79         6.18         6.20      8.01
## 3 540      After Pine Cay      Male 58.30 12.87 14.76         9.45         9.58     11.72
## 4 541      After Pine Cay     Female 43.15  8.55 10.29         6.60         6.26      7.43
## 5 542      After Pine Cay     Female 45.51 10.26 11.02         6.89         7.02      7.71
## 6 543      After Pine Cay     Female 46.97 10.02 10.78         6.85         7.18      8.45
##   Radius Metacarpal LongestFinger FingerCount ToeCount FingerArea1 FingerArea2
## 1    7.99         2.22           3.19          10         12         1.321         1.338
## 2    6.51         2.38           3.55          10         13         0.962         0.950
## 3    9.54         3.54           5.09          14         15         2.507         2.702
## 4    6.60         2.79           3.55          11         12         1.172         1.175
## 5    7.25         2.52           3.37          11         13         1.376         1.357
## 6    7.15         2.39           3.26          12         14         1.428         1.410
##   FingerArea3 ToeArea1 ToeArea2 ToeArea3
## 1         1.339     2.529     2.402     2.369
## 2         0.972     1.498     1.525     1.530
## 3         2.685     4.157     4.140     3.996
## 4         1.186     1.898     1.871     1.867
## 5         1.420     2.627     2.435     2.529
## 6         1.440     2.061     2.018     2.029
```

```
# viewing anole
View(anole)
```

## Question Answers

- The variable 'Sex' has two levels
- The mean femur length in "anole.csv" is 10.919 units
- The first three variables in "anole.csv" are 'ID', 'Hurricane', and 'Origin'

## Section 6 - anole

```
#loading the tidyverse package (previously downloaded and installed)  
library(tidyverse)
```

```
# selecting the column femur  
femurs <- select(anole, Femur)  
  
# checking the new object created  
#femurs
```

## Question Answers

- a. The object is in the Global Environment.

```
# filtering femurs by femur length > 10mm  
femurs_10mm <- filter(femurs, Femur>10)  
  
# checking the new object created  
#femurs_10mm  
  
# ranking femur of length > 10mm  
max(femurs_10mm)
```

```
## [1] 15.38
```

## Question Answers

- a. The longest femur length is 15.38mm  
b. There are 103 observations of femurs with length > 10mm

## Section 7 - anole

```
# mean femur length  
mean(anole$Femur)
```

```
## [1] 10.91866
```

```
# median femur length  
median(anole$Femur)
```

```
## [1] 10.715
```

```
# minimum femur length  
min(anole$Femur)
```

```
## [1] 7.9
```

```
# max femur length  
max(anole$Femur)
```

```
## [1] 15.38
```

```
# range of values for femur length  
range(anole$Femur)
```

```
## [1] 7.90 15.38
```

## Question Answers

- a. I did get the same results to populate as when I used the summary() function, only for the Femur variable.

## Stop, Think, Do:

### Section 1 - lizards

```
# set working directory for all chunks in this file (default working directory is wherever Rmd file is)  
getwd()
```

```
## [1] "C:/Users/Angelo L/Documents/GitHub/BIOL710/RCode710/RCode/working_directory"
```

### Sections 2 and 3 - lizards

These sections were completed by working on a copy of the RMarkdown Lab Report template and continuously adding comments to my code.

### Section 4 - lizards

```
#import data from .csv file  
#commands header=TRUE in order to treat the first row of the data frame as a header and stringsAsFactors=TRUE  
lizards <- read.csv('lizards.csv', header=TRUE, stringsAsFactors=TRUE)
```

### Section 5 - lizards

```
#overview of the dataset lizards  
str(lizards)
```

```
## 'data.frame': 104 obs. of 21 variables:  
## $ ID : Factor w/ 104 levels "572","574","576",...: 1 2 3 4 5 6 7 8 9 10 ...  
## $ Hurricane : Factor w/ 2 levels "After","Before": 1 1 1 1 1 1 1 1 1 ...  
## $ Origin : Factor w/ 2 levels "Pine Cay","Water Cay": 1 1 1 1 1 1 1 1 2 2 ...  
## $ Sex : Factor w/ 2 levels "Female","Male": 1 2 2 1 1 2 1 2 1 2 ...
```

```
## $ SVL      : num  43.5 56.7 59 45.9 42 ...
## $ Femur    : num  9.41 11.74 12.42 9.19 9.44 ...
## $ Tibia    : num  10 13.7 14.4 10.6 10.5 ...
## $ Metatarsal : num  6.54 8.75 9.12 6.99 6.77 7.61 6.48 8.54 5.98 7.9 ...
## $ LongestToe : num  6.84 8.49 9.19 6.26 6.34 7.32 5.73 7.67 5.72 7.57 ...
## $ Humerus   : num  7.64 11.24 10.94 8.28 7.58 ...
## $ Radius    : num  6.74 9.17 9.74 6.92 6.54 7.6 6.76 8.96 6.71 8.55 ...
## $ Metacarpal : num  2.58 3.23 3.34 2.74 2.93 2.65 2.37 2.89 1.94 2.98 ...
## $ LongestFinger: num  3.42 4.85 4.76 3.12 3.57 4 3.1 3.61 3.15 4.76 ...
## $ FingerCount : int  11 12 12 12 11 13 11 12 12 11 ...
## $ ToeCount    : int  12 15 14 12 14 15 13 14 14 14 ...
## $ FingerArea1 : num  1.12 2.05 2.44 1.39 1.08 ...
## $ FingerArea2 : num  1.11 2.06 2.55 1.3 1.04 ...
## $ FingerArea3 : num  1.12 2.05 2.45 1.26 1.09 ...
## $ ToeArea1    : num  1.25 3.67 4.33 1.95 1.7 ...
## $ ToeArea2    : num  1.23 3.73 4.42 1.97 1.73 ...
## $ ToeArea3    : num  1.25 3.7 4.35 1.9 1.72 ...
```

## Question Answers

- The dataset "lizards.csv" is of the class data frame
- "lizards.csv" contains 104 total observations of 21 variables
- Within the dataset "lizards.csv", variables labeled 'ID', 'Hurricane', 'Origin', and 'Sex' are of the class factor, while variables labeled 'SVL', 'Femur', 'Tibia', 'Metatarsal', 'LongestToe', 'Humerus', 'Radius', 'Metacarpal', 'LongestFinger', 'FingerArea1', 'FingerArea2', 'FingerArea3', 'ToeArea1', 'ToeArea2', and 'ToeArea3' are of the class numerical, and variables labeled 'FingerCount' and 'ToeCount' are of the class integer.

```
# levels of the variable Sex, the $ sign means "within"
levels(lizards$Sex)
```

```
## [1] "Female" "Male"
```

```
# summary of each variable
summary(lizards)
```

```
##      ID      Hurricane      Origin      Sex      SVL
## 572    : 1    After :54    Pine Cay :25    Female:48    Min.   :40.17
## 574    : 1    Before:50    Water Cay:79    Male   :56    1st Qu.:43.48
## 576    : 1                                     Median :48.95
## 577    : 1                                     Mean   :49.43
## 578    : 1                                     3rd Qu.:55.22
## 579    : 1                                     Max.   :62.10
## (Other):98
##      Femur      Tibia      Metatarsal      LongestToe
## Min.   : 7.900    Min.   : 9.55    Min.   :5.860    Min.   :5.520
## 1st Qu.: 9.188    1st Qu.:10.16    1st Qu.:6.540    1st Qu.:6.275
## Median :10.630    Median :12.18    Median :7.680    Median :7.155
## Mean   :10.859    Mean   :11.99    Mean   :7.537    Mean   :7.271
## 3rd Qu.:12.405    3rd Qu.:13.42    3rd Qu.:8.440    3rd Qu.:8.223
## Max.   :15.380    Max.   :14.88    Max.   :9.620    Max.   :9.660
##
```

```
##      Humerus      Radius      Metacarpal      LongestFinger
## Min.   : 6.910   Min.   : 5.980   Min.   :1.780   Min.   :2.600
## 1st Qu.: 7.660   1st Qu.: 6.740   1st Qu.:2.558   1st Qu.:3.225
## Median : 9.060   Median : 7.860   Median :2.810   Median :3.840
## Mean   : 8.988   Mean    : 7.859   Mean    :2.824   Mean    :3.823
## 3rd Qu.:10.082   3rd Qu.: 8.873   3rd Qu.:3.132   3rd Qu.:4.272
## Max.   :12.180   Max.    :10.130   Max.    :3.730   Max.    :5.410
##
##      FingerCount      ToeCount      FingerArea1      FingerArea2
## Min.   : 9.00   Min.   :11.00   Min.   :0.761   Min.   :0.414
## 1st Qu.:11.00   1st Qu.:13.00   1st Qu.:1.062   1st Qu.:1.052
## Median :12.00   Median :14.00   Median :1.585   Median :1.593
## Mean   :11.72   Mean    :14.03   Mean    :1.581   Mean    :1.570
## 3rd Qu.:13.00   3rd Qu.:15.00   3rd Qu.:2.038   3rd Qu.:2.052
## Max.   :14.00   Max.    :16.00   Max.    :3.019   Max.    :2.971
##
##      FingerArea3      ToeArea1      ToeArea2      ToeArea3
## Min.   :0.756   Min.   :1.122   Min.   :1.127   Min.   :1.116
## 1st Qu.:1.081   1st Qu.:1.660   1st Qu.:1.683   1st Qu.:1.652
## Median :1.585   Median :2.399   Median :2.402   Median :2.428
## Mean   :1.575   Mean    :2.485   Mean    :2.495   Mean    :2.479
## 3rd Qu.:2.018   3rd Qu.:3.292   3rd Qu.:3.329   3rd Qu.:3.332
## Max.   :3.065   Max.    :4.436   Max.    :4.425   Max.    :4.354
##
```

```
# first rows of lizards
head(lizards)
```

```
##      ID Hurricane      Origin      Sex      SVL Femur Tibia Metatarsal LongestToe Humerus
## 1 572      After Pine Cay Female 43.49  9.41 10.04      6.54      6.84      7.64
## 2 574      After Pine Cay  Male 56.69 11.74 13.69      8.75      8.49     11.24
## 3 576      After Pine Cay  Male 59.02 12.42 14.37      9.12      9.19     10.94
## 4 577      After Pine Cay Female 45.87  9.19 10.58      6.99      6.26      8.28
## 5 578      After Pine Cay Female 42.01  9.44 10.47      6.77      6.34      7.58
## 6 579      After Pine Cay  Male 48.96 10.25 12.05      7.61      7.32      8.87
##      Radius Metacarpal LongestFinger FingerCount ToeCount FingerArea1 FingerArea2
## 1   6.74      2.58      3.42      11      12      1.121      1.115
## 2   9.17      3.23      4.85      12      15      2.049      2.063
## 3   9.74      3.34      4.76      12      14      2.443      2.547
## 4   6.92      2.74      3.12      12      12      1.391      1.296
## 5   6.54      2.93      3.57      11      14      1.083      1.041
## 6   7.60      2.65      4.00      13      15      1.766      1.730
##      FingerArea3 ToeArea1 ToeArea2 ToeArea3
## 1      1.120      1.254      1.232      1.247
## 2      2.052      3.670      3.733      3.702
## 3      2.449      4.329      4.425      4.351
## 4      1.261      1.946      1.970      1.903
## 5      1.087      1.697      1.725      1.724
## 6      1.707      2.209      2.289      2.331
```

```
# viewing lizards
View(lizards)
```

## Question Answers

- The variable 'Sex' has two levels
- The mean femur length in "lizards.csv" is 10.859 units
- The first three variables in "lizards.csv" are 'ID', 'Hurricane', and 'Origin'

## Section 6 - lizards

```
#loading the tidyverse package (previously downloaded and installed)  
library(tidyverse)
```

```
# selecting the column femur  
lemurs <- select(lizards, Femur)  
  
# checking the new object created  
#lemurs
```

## Question Answers

- The object is in fact in the Global Environment.

```
# filtering femurs by femur length > 10mm  
lemurs_10mm <- filter(lemurs, Femur>10)  
  
# checking the new object created  
#lemurs_10mm  
  
# ranking femur of length > 10mm  
max(lemurs_10mm)
```

```
## [1] 15.38
```

## Question Answers

- The longest femur length is 15.38mm
- There are 61 observations of femurs with length > 10mm

## Section 7 - lizards

```
# mean femur length  
mean(lizards$Femur)
```

```
## [1] 10.85942
```

```
# median femur length  
median(lizards$Femur)
```

```
## [1] 10.63
```



```
# minimum femur length  
min(lizards$Femur)
```

```
## [1] 7.9
```

```
# max femur length  
max(lizards$Femur)
```

```
## [1] 15.38
```

```
# range of values for femur length  
range(lizards$Femur)
```

```
## [1] 7.90 15.38
```

### Question Answers

- a. I did get the same results to populate as when I used the summary() function, only for the Femur variable.

### Discussion Question Answers

- a. The layout in RStudio is comprised of a script window in which one can write live code, a console in which one can make quick commands and review script code that has been run, a global environment that keeps track of data frames, and a files/plots/packages that can serve to see active packages and set a working directory
- b. Annotating one's script can be beneficial for learning what certain functions and lines of code can do. It can also help with leaving a trace of one's thought process when they return to their code after a long period of time. Lastly, annotating is crucial for being able to share code with others so that they can more easily interpret your script and apply it to their needs.
- c. One can set a path between R and the location of one's files in the computer by clicking on the files tab of the files/plots/packages pane (principal component window), advancing to the folder in which one desires to set up their working directory and applying it after clicking on the gear icon.
- d. The str() function can be used to get a broad overview of an imported data file, including the structure of the data itself, as well as the total number of observations and variables with their respective variable categorizations and levels.