Lecture 1

2024-07-01

Lecture 1: Regular Typing

These are some notes. Teaching is confusing me. I don't understand what is happening.

제가 공부하고있어요.

- This is bullet 1
- This is bullet 2

1. This is item 1
2. This is item 2

This is **bold** text. This is *italic* text.

Lecture 1: R coding

Three backticks: `

```
2+2
```

```
## [1] 4
```

```
5*40
```

```
## [1] 200
```

```
whateveryouwant <- 5*40
```

Some more notes…why do we care about "object assignment operators" or saving things to "objects"?

Because, we can then refer to them and use them **later**.

```
whateveryouwant
```

```
## [1] 200
```

More sophisticated functions: - average - sum

(Aside: Fast commands) 1. PC: ctrl + alt + i 2. Mac: ctrl + option + i

Examples of averaging and summing

- First create 3 numbers and store them in an object

```
threeNumbs <- c(10,-100,50)
threeNumbs
```

```
## [1]   10 -100   50
```

- Second take the average with the `mean()` function

```
mean(threeNumbs)
```

```
## [1] -13.33333
```

```
mean(c(10,-100,50))
```

```
## [1] -13.33333
```

- Or take the median with the `median()` function

```
median(threeNumbs)
```

```
## [1] 10
```

- Or take the sum with the `sum()` function

```
sum(threeNumbs)
```

```
## [1] -40
```

Create your own function

```
customMean <- function(x) {
   sum(x) / length(x)
}
```

Packages

Packages are collections of functions written by someone else.

`tidyverse`

1. Step 1: `install.packages("tidyverse")`

- **DO THIS FROM THE CONSOLE…NOT IN A CODE CHUNK**

2. Step 2: `require(tidyverse)`

```
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## ── Attaching core tidyverse packages ──────────────────── tidyverse 2.0.0 ──
## ✓ dplyr     1.1.4     ✓ readr     2.1.5
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ ggplot2   3.5.1     ✓ tibble    3.2.1
## ✓ lubridate 1.9.3     ✓ tidyr     1.3.1
## ✓ purrr     1.0.2
## ── Conflicts ─────────────────────────────────── tidyverse_conflicts() ──
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts t
o become errors
```

Loading data

We use the `read_rds()` function to load data

```
debt <- read_rds("https://github.com/jbisbee1/ISP_Data_Science_2024/raw/main/data/sc_deb
t.Rds")

debt
```

```
## # A tibble: 2,546 × 16
##    unitid instnm   stabbr grad_debt_mdn control region preddeg openadmp adm_rate
##     <int> <chr>    <chr>          <int> <chr>   <chr>  <chr>      <int>    <dbl>
##  1 100654 Alabama… AL             33375 Public  South… Bachel…        2    0.918
##  2 100663 Univers… AL             22500 Public  South… Bachel…        2    0.737
##  3 100690 Amridge… AL             27334 Private South… Associ…        1   NA
##  4 100706 Univers… AL             21607 Public  South… Bachel…        2    0.826
##  5 100724 Alabama… AL             32000 Public  South… Bachel…        2    0.969
##  6 100751 The Uni… AL             23250 Public  South… Bachel…        2    0.827
##  7 100760 Central… AL             12500 Public  South… Associ…        1   NA
##  8 100812 Athens … AL             19500 Public  South… Bachel…       NA   NA
##  9 100830 Auburn … AL             24826 Public  South… Bachel…        2    0.904
## 10 100858 Auburn … AL             21281 Public  South… Bachel…        2    0.807
## # ℹ 2,536 more rows
## # ℹ 7 more variables: ccbasic <int>, sat_avg <int>, md_earn_wne_p6 <int>,
## #   ugds <int>, costt4_a <int>, selective <dbl>, research_u <dbl>
```