Lesson 3: Dive into Tidy: Data Wrangling and Manipulation Part I - Homework Answers

Instructor: Caitlin Allen Akselrud (caitlin.allen_akselrud@noaa.gov)

January 29, 2021

Contents

Answers to Questions:

First, load your libraries!

```
library(tidyverse)
```

1 A read-in example. Use the error message to determine what the two columns should be correct the column types accordingly.

```
challenge <- read_csv(readr_example("challenge.csv"))
```

```
## Warning: 1000 parsing failures.
##  row col           expected     actual
## 1001   y 1/0/T/F/TRUE/FALSE 2015-01-16 'C:/Users/Emii/Documents/R/win-library/4.0/readr/extdata/chall
## 1002   y 1/0/T/F/TRUE/FALSE 2018-05-18 'C:/Users/Emii/Documents/R/win-library/4.0/readr/extdata/chall
## 1003   y 1/0/T/F/TRUE/FALSE 2015-09-05 'C:/Users/Emii/Documents/R/win-library/4.0/readr/extdata/chall
## 1004   y 1/0/T/F/TRUE/FALSE 2012-11-28 'C:/Users/Emii/Documents/R/win-library/4.0/readr/extdata/chall
## 1005   y 1/0/T/F/TRUE/FALSE 2020-01-13 'C:/Users/Emii/Documents/R/win-library/4.0/readr/extdata/chall
## .... ... .................. .......... ..................................................................
## See problems(...) for more details.
```

```
head(challenge)
```

x	y
404	NA
4172	NA
3004	NA
787	NA
37	NA
2332	NA

```
tail(challenge)
```

x	y
0.8052743	NA
0.1635163	NA
0.4719390	NA
0.7183186	NA
0.2698786	NA
0.6082372	NA

HINT: use View(challenge) or tail(challenge) to see more of column y

```
# No error!
challenge <- read_csv(
  readr_example("challenge.csv"),
  col_types = cols(
    x = col_double(),
    y = col_date()
  )
)
head(challenge)
```

x	y
404	NA
4172	NA
3004	NA
787	NA
37	NA
2332	NA

```
tail(challenge)
```

x	y
0.8052743	2019-11-21
0.1635163	2018-03-29
0.4719390	2014-08-04
0.7183186	2015-08-16
0.2698786	2020-02-04
0.6082372	2019-01-06

2 2) write:

```
fish_names <- c("Nemo", "Bubbles", "Jack", "Captain", "Finley",
                "Goldie", "Dory", "Ariel", "Angel", "Minnie")

name_rank <- rep(1:5, times = 2)
```

2.1 a) write your corrected challenge object to a .csv and a .txt file in the correct subfolder

```
write_csv(challenge, path = here("output", "challenge_correct.csv"))
# The warning (which you may or may not also get) here has to do with the version of the package you ar
```

2.2 b) write challenge, fish_names, and name_rank to a .txt file

```
save_stuff <- list("challenge" = challenge,
                   "names" = fish_names,
                   "rank" = name_rank)

save_dir <- here('output', "saving_stuff.txt")
sink(save_dir)
save_stuff
```

```
## $challenge
## # A tibble: 2,000 x 2
##        x y
##    <dbl> <date>
##  1   404 NA
##  2  4172 NA
##  3  3004 NA
##  4   787 NA
##  5    37 NA
##  6  2332 NA
##  7  2489 NA
##  8  1449 NA
##  9  3665 NA
## 10  3863 NA
## # ... with 1,990 more rows
##
## $names
##  [1] "Nemo"    "Bubbles" "Jack"    "Captain" "Finley"  "Goldie"  "Dory"    "Ariel"   "Angel"   "Minni
##
## $rank
##  [1] 1 2 3 4 5 1 2 3 4 5
```

```
sink()
```

3 Find a data set you've worked with and pull it into R

Fix any problems with column designation, and clean up the column names (Hint: remember lecture 2)

Note: We can help you find a dataset if you don't have one on hand. We're excited to see what you're working on next class!