

# Vectors/Iterators Lab

David Gerard

2019-05-06

## Learning Objectives

- Practice manipulating vectors.
- Practice for loops.
- Practice `purrr`.

## Exercise 1: Simulations

1. Write a function that uses a for loop that, for each iteration, randomly draws 5 observations from an exponential distribution with rate parameter 1 (use `rexp()`) and calculates its mean. It should do this 10,000 times. Choose an appropriate plot to plot the distribution of means.
2. Repeat part 1 by using a `map_*()` function.
3. Repeat part 1 by using the `replicate()` function.
4. Use a another for loop that will print out plots for sample sizes of 5, 10, and 20 observations (instead of just 5).

## Exercise 2: Star Trek

For this exercise, load in the following dataset.

```
st <- list(  
  franchise = "Star Trek",  
  series = c("TOS", "TNG", "DS9", "VOY", "ENT"),  
  starship = tribble(~ID, ~class, ~name, ~length, ~launched,  
    ##-----/-----/-----/-----/  
    "NCC-1701", "constitution", "Enterprise", 288.646, 2245,  
    "NCC-1701D", "galaxy", "Enterprise", 642.5, 2361,  
    "NX-74205", "defiant", "Defiant", 170, 2370,  
    "NCC-74656", "intrepid", "Voyager", 343, 2371),  
  captain = tribble(~firstname, ~lastname, ~ship, ~drink,  
    ##-----/-----/-----/  
    "James", "Kirk", "NCC-1701", "brandy",  
    "Jean-luc", "Picard", "NCC-1701D", "tea",  
    "Benjamin", "Sisko", "NX-74205", "coffee",  
    "Katherine", "Janeway", "NCC-74656", "coffee"),  
  age = as.duration(today() - ymd(19660908))  
)
```

1. Use an appropriate function to see the components of `st` in a concise way.
2. Use R code to determine the class of each element of `st`.

3. It seems that `starship` and `captain` can be merged into one data frame. Do this now.

Your `st` list should now look like this:

```
## List of 4
## $ franchise: chr "Star Trek"
## $ series   : chr [1:5] "TOS" "TNG" "DS9" "VOY" ...
## $ starship :Classes 'tbl_df', 'tbl' and 'data.frame':  4 obs. of  8 variables:
## ..$ ID      : chr [1:4] "NCC-1701" "NCC-1701D" "NX-74205" "NCC-74656"
## ..$ class   : chr [1:4] "constitution" "galaxy" "defiant" "intrepid"
## ..$ name    : chr [1:4] "Enterprise" "Enterprise" "Defiant" "Voyager"
## ..$ length  : num [1:4] 289 642 170 343
## ..$ launched: num [1:4] 2245 2361 2370 2371
## ..$ firstname: chr [1:4] "James" "Jean-luc" "Benjamin" "Katherine"
## ..$ lastname : chr [1:4] "Kirk" "Picard" "Sisko" "Janeway"
## ..$ drink    : chr [1:4] "brandy" "tea" "coffee" "coffee"
## $ age        :Formal class 'Duration' [package "lubridate"] with 1 slot
## .. ..@ .Data: num 1.66e+09
```

## Exercise 3: Diamonds

1. Load in the `diamonds` data frame from the `ggplot2` package.
2. Calculate summary statistics (using `summary()`) for all quantitative variables.
3. To fit a linear model of `price` on `depth` and extract the  $p$ -value for the regression coefficient of `depth`, you would use the following R code:

```
lmout <- lm(price ~ depth, data = diamonds)
sumout <- summary(lmout)
sumout$coefficients["depth", "Pr(>|t|)"]
```

Fit a separate linear model for `price` on `depth` for each level of `color`. Extract the  $p$ -values for the regression coefficients.

4. Plot `depth` vs `price`, faceting by `color`, including an OLS line. Do the  $p$ -values in the previous question roughly match what you would expect based on the plots?