Lab 11: Solutions

The folder `Lab11Data` contains several CSV data files.

```
dfiles <- dir("Lab11Data",full.names=TRUE)
dfiles
```

```
## [1] "Lab11Data/study1.csv" "Lab11Data/study2.csv" "Lab11Data/study3.csv"
## [4] "Lab11Data/study4.csv" "Lab11Data/study5.csv" "Lab11Data/study6.csv"
## [7] "Lab11Data/study7.csv" "Lab11Data/study8.csv" "Lab11Data/study9.csv"
```

1. Write R code to read in the first file. Print the tibble that you just read in. Use `names()` to change the column names of the tibble to `x` and `y`. Repeat for the second file. How many observations are in these first two files?

```
library(tidyverse)
f <- read_csv(dfiles[1])
f
```

```
## # A tibble: 100 x 2
##    x.study1 y.study1
##       <dbl>    <dbl>
##  1  -0.560    -1.77
##  2  -0.230    -1.95
##  3   1.56      1.17
##  4   0.0705   -1.41
##  5   0.129    -2.22
##  6   1.72      1.03
##  7   0.461     0.785
##  8  -1.27     -2.90
##  9  -0.687    -0.972
## 10  -0.446    -2.89
## # ... with 90 more rows
```

```
names(f) <- c("x","y")

f <- read_csv(dfiles[2])
f
```

```
## # A tibble: 150 x 2
##    x.study2 y.study2
##       <dbl>    <dbl>
##  1   0.543    0.591
##  2  -0.414    0.0422
##  3  -0.476   -2.40
##  4  -0.789    2.61
##  5  -0.595   -0.146
##  6   1.65     0.990
##  7  -0.0540   0.385
```

```
##  8   0.119    1.15
##  9   0.244    0.961
## 10   1.23     0.0638
## # ... with 140 more rows
```

```
names(f) <- c("x","y")
```

There are 100 and 150 observations in the files 1 and 2, respectively.

2. Use `vector()` to create an empty vector called `ff` that is of mode "list" and length 9. Now write a `for()` loop to loop over the 9 files in `dfiles` and for each (i) read the file in to a tibble, and change the column names to `x` and `y` as in part (1), and (ii) copy the tibble to an element of your list `ff`.

```
ff <- vector(mode="list",length=9)

for(i in seq_along(ff)) {
  f <- read_csv(dfiles[i])
  names(f) <- c("x","y")
  ff[[i]] <- f
}
```

3. Write a function called `read.study_data` that takes a vector of data file names (like `dfile`) as input, reads the data files into a list, ~~assigns class "study_data" to the list~~, and returns the list. Your function should use `length(dfile)` to determine the number of files.

```
read.study_data <- function(dfiles) {
  ff <- vector(mode="list",length=length(dfiles))
  for(i in seq_along(ff)) {
    f <- read_csv(dfiles[i])
    names(f) <- c("x","y")
    ff[[i]] <- f
  }
  class(ff) <- "study_data"
  ff
}
ss <- read.study_data(dfiles)
```