

# BSDS 100: Intro to Data Science with R

## Assignment 7

by Abbie M Popa (University of San Francisco)

**Directions:** For all questions in this assignment, write complete sentences and fully answer any question that is asked, and use R to answer each question. Provide all R code and solutions by *knitting* your final RStudio file into a single file. Late assignments will automatically have 10 points deducted. Assignments turned in after the answer key is posted (one week after the due date) will not receive any points.

**Assignment:** For this assignment, please do the following

1. (8 pts) We have now discussed several libraries from the tidyverse, including `tibble`, `ggplot2`, `tidyr`, and `dplyr`. Select two and describe what they do, why you might use them, how to use one function from each, and why you might use that function.
2. (4 pts) Review the notes on different types of plots. Select two types of plots and describe what they are useful for. Then describe which functions you would use to make each.
3. (10 pts) Load the `ggplot2` library. Read **Sections 3.1** through **3.5** in the *R for Data Science* text on Data Visualization here <http://r4ds.had.co.nz/data-visualisation.html>. Complete Exercises 3.2.4 numbers 1-5. You will need to include code and sentence descriptions for each of the five questions, except question 4, which only requires code.
4. (8 pts) Load the `tibble` library and the `tidyr` library. Tidy the tibble below. Do you need to spread or gather it? What are the variables before tidying? After tidying? You must include code and sentence descriptions to receive full credit.

```
animals <- tribble(  
  ~species, ~weight_when_admitted, ~weight_when_adopted, ~supplement,  
  "cat",    8,    15,    "turkey",  
  "cat",    9,    11,    "kibbles",  
  "dog",   18,   27,    "turkey"  
)
```