Practice Problems #3

Alvin D. Jeffery, PhD, RN

Last Updated: 2018-05-02

The purpose of this problem set is to practice multiple aspects of wrangling and analyzing data. You will open up 3 separate datasets coming from the SUPPORT dataset, explore & manipulate them, and then do some analytical work.

Load and Review Data

Begin by reading in and exploring the data.

```
### read/load the data


### explore the data
```

Descriptive Statistics

Explore the following items:
1. Create a table of how many patients have each disease.
2. What is the mean heart rate?
3. Convert temperature from Celsius to Farenheit.
4. Set non-physiological respiratory rates to missing.

```
### disease table


### troubleshoot discovering the mean heart rate


### create new variable tempF for temperature in Farenheit


### if respiratory rate < 8, set to NA
```

Combine Datasets

Combine all 3 datasets into 1 big dataframe for conducting analyses. There should be 1 row per patient. Then, sort the large dataframe by patient's ages.

Challenge: Re-write all this code as one chain.

```
### combine demographics and vitals dataframes into an "mydata" object


### add the labs dataframe to the "mydata" object


### sort by age from oldest to youngest


### re-write above lines into 1 chain
```

Sex Differences in Heart Rate

Let's say you want to explore sex differences in the heart rate values. Perform the following procedures:
1. Create a boxplot figure of the heart rates, grouped by male vs. female
2. Perform a t-test for inferential analysis
3. Build a linear model for further inferential analysis

```
### boxplot (don't forget to add a title & consider rotating your y-axis labels)


### conduct a t-test


### build a linear model
```

Function for Boxplots

Your co-investigator really loved your beautiful boxplots. You want to create a function that makes it easier to create these. Write a function that takes 3 arguments (data, dependent variable, & grouping/independent variable) and produces a boxplot where the title is automatically populated according to the variable names.

Hint: You'll need to use a paste() function within a formula() function in order to pass it appropriately to the boxplot function().

```
### create function


### attempt some examples
```