R functions

Math 241, Week 3

```
# it's good practice to check that all the packages required are loaded and installed
libs <- c('tidyverse','knitr','viridis','mdsr', 'macleish','babynames')
for(l in libs){
  if(!require(l,character.only = TRUE, quietly = TRUE)){
    message( sprintf('Did not have the required package << %s >> installed. Downloading now ... ',l))
    install.packages(l)
  }
  library(l, character.only = TRUE, quietly = TRUE)
}
```

Goals of this in-class activity:

- Practice creating functions in R.

Notes:

- When creating your graphs, consider context (i.e. axis labels, title, ...)!
- If I provide partially completed code, I will put `eval = FALSE` in the chunk. Make sure to change that to `eval = TRUE` once you have completed the code in the chunk.
- Be prepared to ask for help from me, Tory, and your classmates!

Problem 1 (Medium):

Write a function called `count_name` that, when given a name as an argument, returns the total number of births by year from the `babynames` data frame in the `babynames` package that match that name. The function should return one row per year that matches (and generate an error message if there are no matches). Run the function once with the argument Ezekiel and once with Ezze.

```
data(babynames) # this will explicitly ask R to load the babynames dataset to your environment
count_name <- function(name){
  # your code here
}
```

Problem 2 (Medium):

1. Write a function called `count_na` that, when given a vector as an argument, will count the number of NA's in that vector. Count the number of missing values in the `SEXRISK` variable in the `HELPfull` data frame in the `mosaicData` package.

```r
data(HELPfull) # this will explicitly ask R to load the HELPfull dataset to your environment
count_na <- function(x){
  # your code here
}
```

2. Apply `count_na` to the columns of the `Teams` data frame from the `Lahman` package. How many of the columns have missing data?

```r
data(Teams) # this will explicitly ask R to load the Teams dataset to your environment
```

Problem 3 (Medium):

Write a function called `prop_cancel` that takes as arguments a month number and destination airport and returns the proportion of flights missing arrival delay for each day to that destination. Apply this function to the `nycflights13` package for February and Atlanta airport ATL and again with an invalid month number.

```r
library(nycflights13)
prop_cancel <- function(monthnum, airport) {
  # your code here
}
```