Homework 2: NRSG 741

Tommy Flynn*,[a]

[a]*Emory School of Nursing 1520 Clifton Rd Atlanta GA 30322*

Abstract

This assignment applies the dplyr and ggplot2 packages to work with the Davis dataset in the car package, which contains data on the measured and reported heights and weights of men and women engagedin regular exercise. The associated GitHub repository can be found at https://github.com/tommyflynn/N741_Homework/tree/master/Flynn_HW_02.

1. What kind of R object is the Davis dataset?

Answer: The Davis dataset is the "data.frame" class of R object.

2. How many observations are in the Davis dataset?

Answer: There are 200 observations.

```
# Use the is.na function to filter non-missing values from the repwt
# variable
repmissing <- filter(davis, is.na(repwt))
# count(repmissing)
```

3. For reported weight, how many observations have a missing value?

Answer: Reported weight has 17 missing values

```
# HINT: find the complete rows...
completeobs <- count(na.omit(davis))
```

4. How many observations have missing values?

Answer: The Davis dataset has 181 complete observations.

```
femaleset <- davis %>% subset(sex == "F")
```

5. Create a subset containing only females. How many females are in this subset?

Answer: A subset containing only females from the Davis data has 112 observations.

*Corresponding Author
Email address: tjflynn@emory.edu (Tommy Flynn)

```r
davisBMI <- davis %>% mutate(sqrmetHT = (height/100) * (height/100)) %>% mutate(BMI = weight/sqrmetHT) %>%
    mutate(BMIclass = if_else(BMI < 18.5, "Underweight", if_else(BMI < 25, "Normal",
        if_else(BMI < 30, "Overweight", "Obese", "Missing"), "Missing")))
# mean(davisBMI$BMI)
```

6. What is the average BMI for these individuals?

Answer: The mean BMI is 24.7009556.