# R Workshop - Day 1 Exercise Solutions

* Material borrowed from A. Culhane (Thank you!) *

## Read in Data

```
# Read in file from the web
myURL <- "http://bcb.dfci.harvard.edu/~aedin/courses/Bioconductor/Women.txt"
women <- read.table(myURL, sep = "\t", header = TRUE)
head(women)
```

```
##   height weight age
## 1     58    115  33
## 2     59    117  34
## 3     60    120  37
## 4     61    123  31
## 5     62    126  31
## 6     63    129  34
```

```
str(women)
```

```
## 'data.frame': 15 obs. of  3 variables:
##  $ height: int  58 59 60 61 62 63 64 65 66 67 ...
##  $ weight: int  115 117 120 123 126 129 132 135 139 142 ...
##  $ age   : int  33 34 37 31 31 34 31 39 35 34 ...
```

# Solutions

```
# 1. What are the column names of the imported data?
colnames(women)
```

```
## [1] "height" "weight" "age"
```

```
# 2. What is the class of this data set?
class(women)
```

```
## [1] "data.frame"
```

```
# 3. How many rows and columns are in the data?
dim(women)
```

```
## [1] 15  3
```

```
nrow(women)
```

```
## [1] 15
```

```r
ncol(women)
```

```
## [1] 3
```

```r
# 4. What is the mean height, weight and age of the women?
summary(women)
```

```
##      height         weight         age
##  Min.   :58.0   Min.   :115   Min.   :30.0
##  1st Qu.:61.5   1st Qu.:124   1st Qu.:32.0
##  Median :65.0   Median :135   Median :34.0
##  Mean   :65.0   Mean   :137   Mean   :33.9
##  3rd Qu.:68.5   3rd Qu.:148   3rd Qu.:35.5
##  Max.   :72.0   Max.   :164   Max.   :39.0
```

```r
# 5. Compare the above result to using the function colMeans().
colMeans(women)
```

```
## height weight    age
##  65.00 136.73  33.93
```

```r
# 6. How many women have a weight under 120?
sum(women$weight < 120)
```

```
## [1] 2
```

```r
# 7. What is the average height of women who weigh between 124 and 150
# pounds?
mean(subset(women, women$weight > 124 & women$weight < 150)$height)
```

```
## [1] 65
```

```r
# 8. Sort the data by age.

# BEFORE
head(women)
```

```
##   height weight age
## 1     58    115  33
## 2     59    117  34
## 3     60    120  37
## 4     61    123  31
## 5     62    126  31
## 6     63    129  34
```

```r
# AFTER
women <- women[order(women$age), ]  # this stores the change
head(women)
```

```
##    height weight age
```

```
## 14     71    159  30
## 4      61    123  31
## 5      62    126  31
## 7      64    132  31
## 1      58    115  33
## 13     70    154  33
```

```
# 9. Give the 5th row the row name Lucy.
rownames(women)[5] <- "Lucy"
head(women)
```

```
##      height weight age
## 14       71    159  30
## 4        61    123  31
## 5        62    126  31
## 7        64    132  31
## Lucy     58    115  33
## 13       70    154  33
```

```
# 10. Write out the data frame as a tab delimited file (or csv file) using
# write.table() (or write.csv()).

# As tab-delimited
write.table(women, file = "women_out.txt", sep = "\t")

# As csv
write.csv(women, file = "women_out.csv")
```