Stephanie Mecham
EPID 674 Section 2

Exercise 2A

Problem 1.

A. Dimensions are 4784 observations x 22 variables.
B. Dimensions of the reimported bpa data from CSV format is now 4784 observations x 23 variables.

Problem 2.

a. Age
   i. Mean: 49.49
   ii. Standard deviation: 18.52
   iii. Minimum: 20.00
   iv. 25th percentile: 33.00
   v. Median: 48.00
   vi. 75th percentile: 65.00
   vii. Maximum: 85.00

b. Urinary bisphenol-A
   i. Mean: 3.712
   ii. Standard deviation: 5.77
   iii. Minimum: 0.280
   iv. 25th percentile: 1.00
   v. Median: 2.100
   vi. 75th percentile: 4.100
   vii. Maximum: 79.800

CODE:

```
>setwd("\\Users\smecham\Desktop\Hw2") #setting working directory
>
>install(sas7bdat) #installing sas7bdat package
>library(sas7bdat) #calling sas7bdat package
>
>options(max.print=999999) #setting parameters for file input
>read.sas7bdat("bpa.sas7bdat") #reading in SAS file
>
> bpa <- read.sas7bdat("bpa.sas7bdat") #setting file to a variable
> save(bpa, file="bpa.rda") #saving for future use
>
> write.csv(bpa, file="bpa.csv") #Converting to CSV
```

```
>new_bpa <- read.csv("bpa.csv") #Setting file to a new variable
> save(new_bpa, file="new_bpa.rda") #saving for future use
>
> age_data <- c("age") #Subsetting dataset to isolate age variable
> age_summary <- bpa[age_data] #Subsetting dataset to isolate age variable
>
> summary(age_summary) #Obtaining five-number summary
      age
 Min.   :20.00
 1st Qu.:33.00
 Median :48.00
 Mean   :49.49
 3rd Qu.:65.00
 Max.   :85.00
>> sd(bpa$age, na.rm= TRUE) #Obtaining standard deviation
[1] 18.51889
>
> ubpa_data <- c("ubpa") #Subsetting dataset to isolate urinary bpa variable
> ubpa_summary <- bpa[ubpa_data] #Subsetting dataset to isolate urinary bpa variable
>
> summary(ubpa_summary) #Obtaining five-number summary
      ubpa
 Min.   : 0.280
 1st Qu.: 1.000
 Median : 2.100
 Mean   : 3.712
 3rd Qu.: 4.100
 Max.   :79.800
>
> sd(bpa$ubpa, na.rm=TRUE) #Obtaining standard deviation
[1] 5.77308
>
```