Unit 4– Introduction to R
Homework 2 of 2
Solutions

Preliminaries

```
setwd("/Users/cbigelow/Desktop")

# Load datasets
load(file="intror1.Rdata")
load(file="intror2.Rdata")
```

Question #1 - Label Variables

```
library(Hmisc)
label(intror1$studyid) <- "studyid:  Study ID"
label(intror1$city) <- "city:  Woman Lives Metro"
label(intror1$dues) <- "dues:  Union Dues Paid"
```

Question #2 - Label (Factor) Variable Values

```
intror1$city <- factor(intror1$city,
                               levels = c(0,1),
                               labels = c("no", "yes"))
table(intror1$city)

##
##  no yes
## 296 704
```

Question #3 - Create new variable using natural logarithm function log( )

```
intror2$lnlead <- log(intror2$lead)
label(intror2$lnlead) <- "lnlead: ln(lead)"
summary(intror2$lnlead)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.6932  2.3979  2.6391  2.6150  2.9309  3.9120     286
```

Question #4 - Create a new variable that is a function of existing variables

```
intror2$bmi <- (100*100*intror2$weight )/(intror2$height^2 )
label(intror2$bmi) <- "bmi:  Body Mass Index"
summary(intror2$bmi)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   15.36   22.46   24.83   25.61   27.90   44.91
```

Question #5 - Create a grouped numerical variable.

```
intror2$bmi_group <- NA
intror2$bmi_group[intror2$bmi <= 18.49] <- 1
intror2$bmi_group[(intror2$bmi > 18.49) & (intror2$bmi <= 24.99)] <- 2
intror2$bmi_group[(intror2$bmi > 24.99) & (intror2$bmi <= 29.99)] <- 3
intror2$bmi_group[intror2$bmi > 29.99] <- 4
table(intror2$bmi_group)

##
##   1   2   3   4
##  17 241 172  70
```

Question #6 - Create Variable Label and Variable value labels

```
label(intror2$bmi_group) <- "Body Mass Index, Grouped"
intror2$bmi_group <- factor(intror2$bmi_group,
                            levels = c(1,2,3,4),
                            labels = c("Underweight (<= 18.5)",
                                       "Normal (18.5-24.99)",
                                       "Overweight (25.0-29.99)",
                                       "Obese (>= 30.0)"))
table(intror2$bmi_group)

##
##   Underweight (<= 18.5)     Normal (18.5-24.99) Overweight (25.0-29.99)
##                      17                     241                     172
##         Obese (>= 30.0)
##
```