Allen Church Chapter 2 - Exercise 2 Accelerated Statistics for Public Policy 9/16/19

Load data and rename to data1. View data to look at dataset.

```
load("/Users/allenchurch/Ch2_Exercise2_Olympics.RData")
data1 <- dta
```

2a. Summarize medals, athletes, and GDP columns.

```
summary(data1$medals)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.000   0.000   1.751   0.000  37.000
```

```
summary(data1$athletes)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    0.00    2.00   18.17   13.00  230.00
```

```
summary(data1$GDP)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0110  0.1212  0.3849  1.1691  1.5127 14.5230     156
```

2b. List first 5 observations for country, year, medals, athletes, and GDP data.

```
##Create cols with 5 column names, subset data1 with cols, and print first 5 rows
cols <- c("country", "year", "medals", "athletes", "GDP")
subset <- data1[cols]
subset[1:5,]
```

```
##   country year medals athletes    GDP
## 1 Albania 1980      0        0     NA
## 2 Albania 1984      0        0 0.0641
## 3 Albania 1988      0        0 0.0637
## 4 Albania 1992      0        0 0.0206
## 5 Albania 1994      0        0 0.0587
```

2c. How many observations are there for each year?

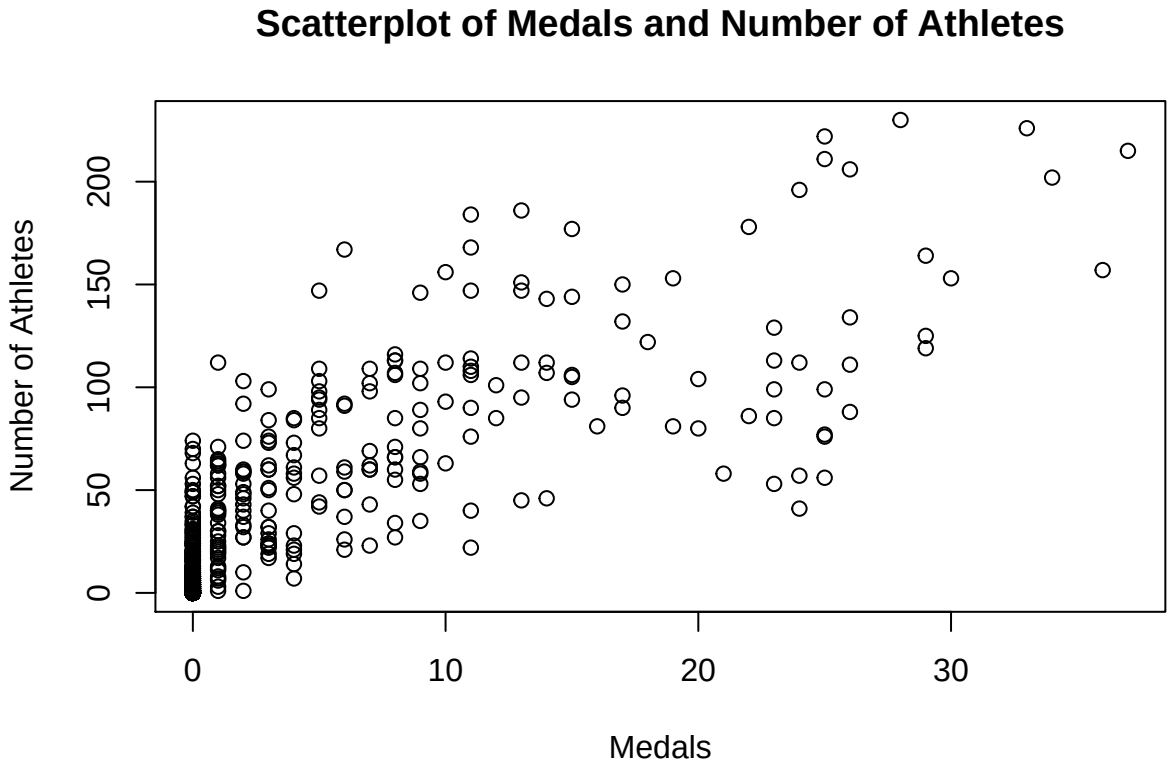
```
#Create table with year column in data1
table(data1$year)
```

```
##
## 1980 1984 1988 1992 1994 1998 2002 2006 2010 2014
##  117  117  117  113  110  110  110  110  109  109
```

2d. Produce scatterplot of medals and number of athletes. Describe relationship depicted.

The relationship depicted below shows high clustering in the number of athletes that received 0 medals. It also shows that a majority - or close to a majority - of athletes received under 10 medals. Moreover, the number of athletes that received 30 or more medals is incredibly low with a count of 5 athletes. It also shows a positive correlation between number of athletes and number of medals won.

```r
plot(data1$medals, data1$athletes, main="Scatterplot of Medals and Number of Athletes",
     xlab="Medals", ylab = "Number of Athletes")
```



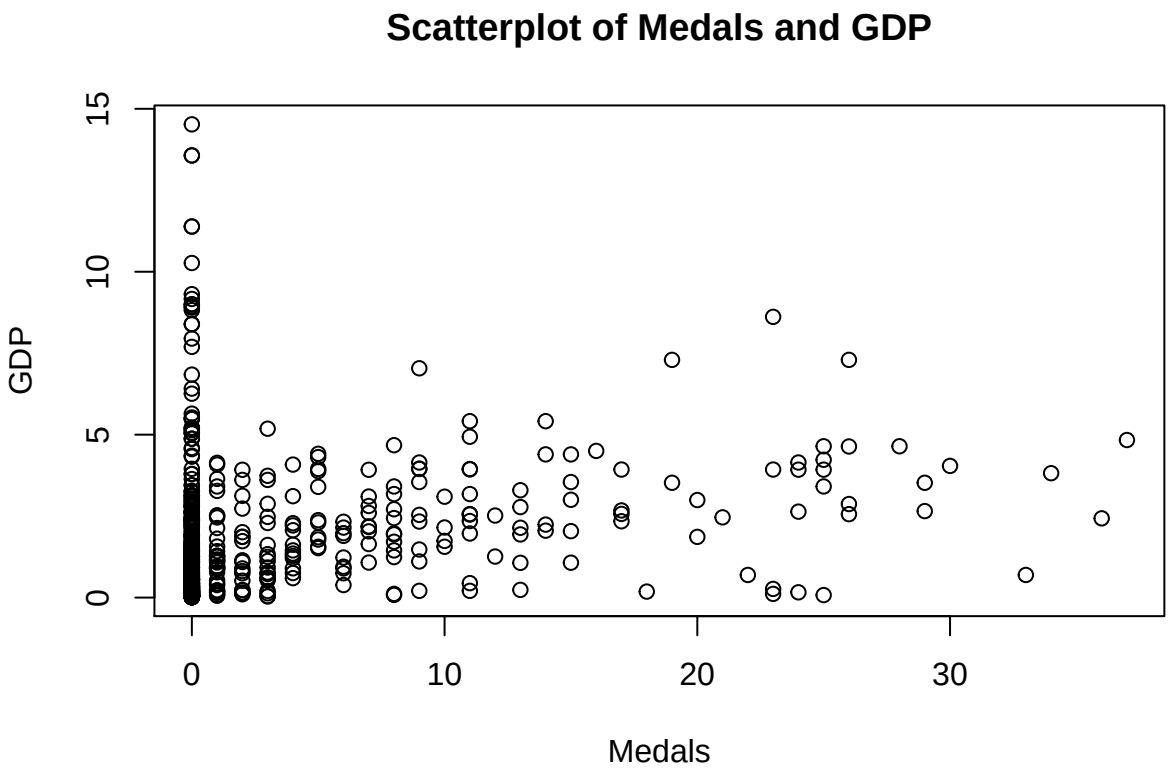
2e. Explain any suspicion you might have that other factors could explain the observed relationship between number of athletes and medals.

It seems there is a direct relationship between the number of athletes and the number of medals won by a country. This may cause the independent variable to be endogenous if changes in it are related to factors of the error term.

2f. Create scatterplot of medals and GDP. Briefly describe any clear patterns.

There does not seem to be a clear relationship between GDP and number of medals earned, meaning that GDP is not strongly correlated with number of medals. It also looks as the majority of countries won 15 or less medals, and there is also a high concentration of countries with 0 medals won.

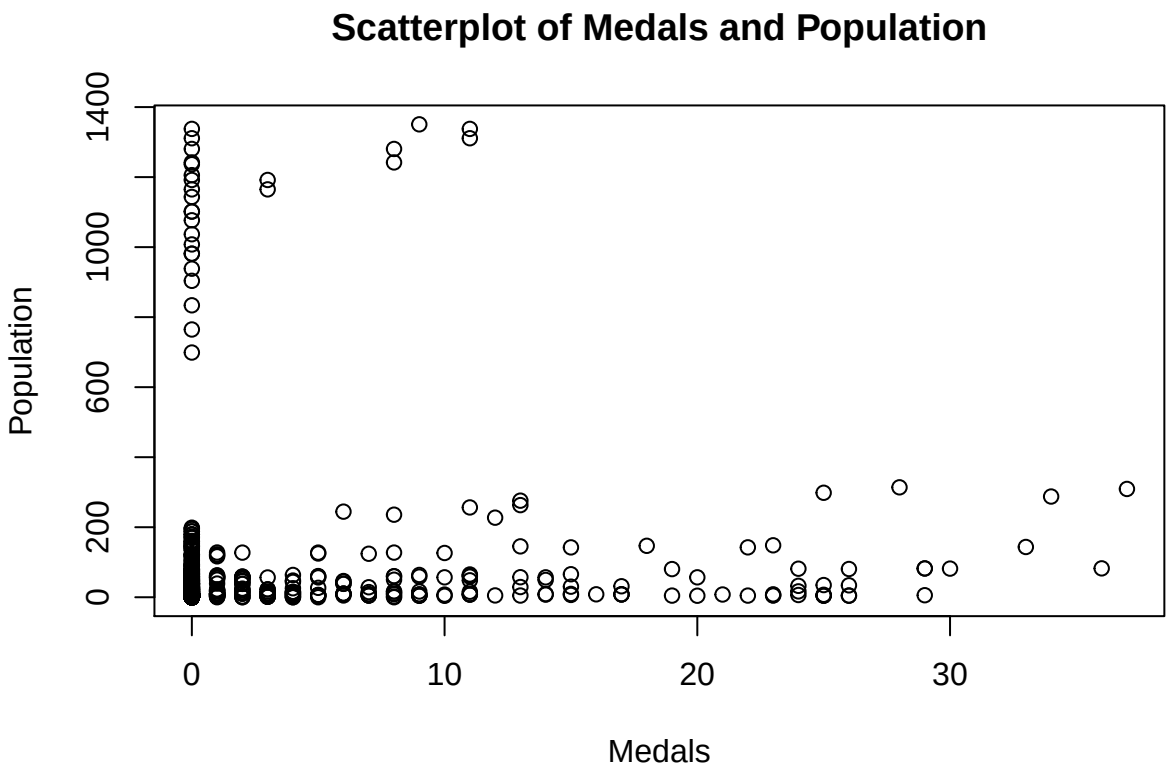
```r
plot(data1$medals, data1$GDP, main="Scatterplot of Medals and GDP",
     xlab="Medals", ylab = "GDP")
```



2g. Create scatterplot of medals and population. Briefly describe any clear patterns.

The relationship below shows that population size is not strongly correlated with number of medals won. As seen in the scatterplot below, countries with very large populations won the same, or less, medals than countries with smaller populations.

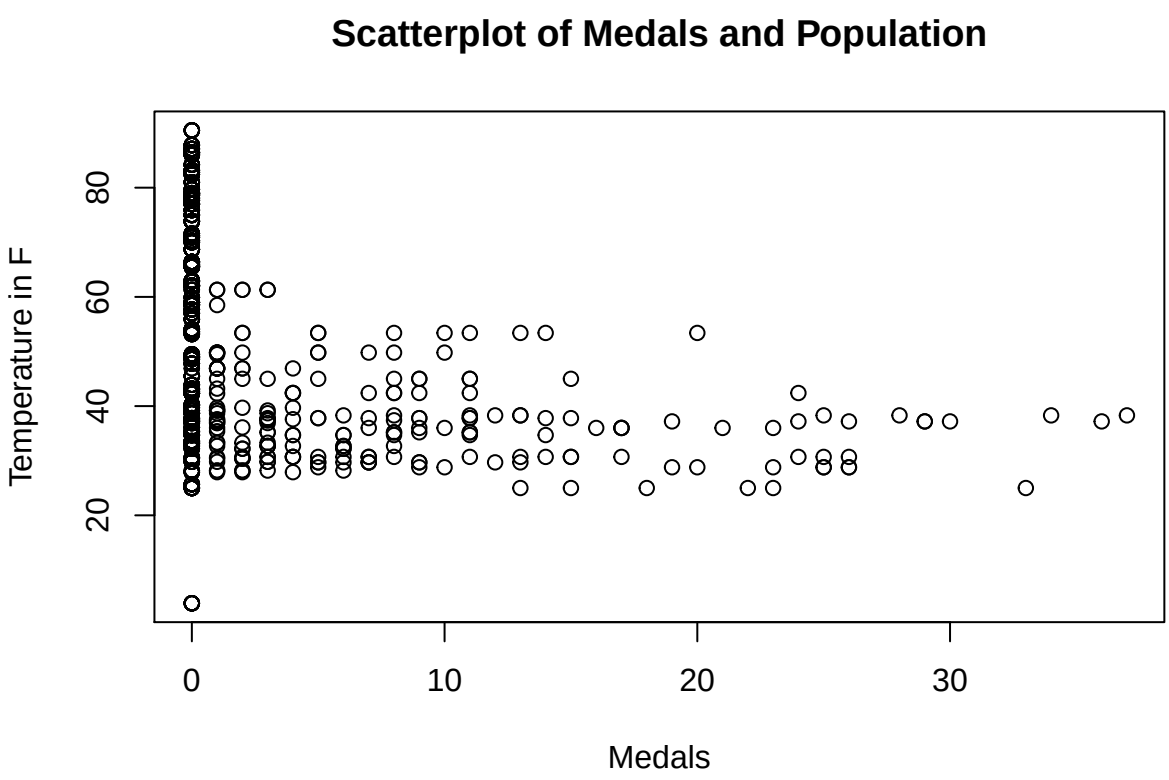
```
plot(data1$medals, data1$population, main="Scatterplot of Medals and Population",
     xlab="Medals", ylab = "Population")
```



2h. Create scatterplot of medals and temperature. Describe any clear patterns

The relationship below shows that countries with high medal counts - more than 20 - have an average temperature of approximately 40 degrees. Countries with extremely high temperatures - above 60 degrees F - mostly received 0 medals in the games. Since it was the Winter Olympics, this relationship is logical.

```
plot(data1$medals, data1$temp, main="Scatterplot of Medals and Population",
     xlab="Medals", ylab = "Temperature in F")
```


Scatterplot of Medals and Population
Temperature in F
20
40
60
80
0
10
20
30
Medals