Exercise 1 – R-Studio Skills Check

1. Install R and R-Studio. The URLs for these are:  
   <https://cran.r-project.org>   
   <https://www.rstudio.com/products/rstudio/download/#download>
2. Run R-Studio. At the console prompt (lower left pane), type **summary(Orange)** and press return.
3. Environment window (upper right): (Environment Tab) Load data from tinydata.Rdata (Hint, use the yellow icon of an opening folder). Import tinydata.csv using the Import Dataset button. (History Tab) review your code history so far. Copy the load( ) from the history window into the code window using the appropriate button.
4. Content window (lower right): (Packages tab) Review downloaded packages. Install the BayesFactor package. Load (library) the BayesFactor package into memory using the checkbox in the packages tab. (Help tab) At the console run **help("Control")** and examine the documentation.
5. Code/Data Window (upper left): Open a new code file. Type a comment in the code window (use the #) and type your name. Type this command in the same code window and run it: **View(Orange)**
6. Add **hist(Orange$circumference)**to your code and a comment describing the shape of the distribution. Also run this command:  
   **plot(Orange$age, Orange$circumference)**and describe what you see in a comment.

Exercise 2 – Descriptive Statistics

1. Here is R code to create a small data set: **smallData <- c(71,71,72,74)**  
   *Without using any functions except sum( ), length( ) and sqrt( )*, calculate the mean, the sum of squared deviations from the mean, the variance, and the standard deviation. Use the code window in order to save all of your code and comments. Why do people like to discuss the standard deviation of a variable as opposed to the variance?
2. The mean of a *very large data set* of temperature readings is 72. The median is also 72. What, if anything, can you say about the shape of the distribution? What, if anything, can you say about the mode? Add a comment with your answers.
3. Daily temperature readings collected from Alpha City for one year showed a mean of 72 (F) and a standard deviation of 6 degrees. Annual temperature readings collected from Omega City showed a mean of 72 (F) and a standard deviation of 18 degrees. In which city would you prefer to live and why? Add a comment with your answers.
4. Run this line of code in R-Studio: **myDist <- rbinom(1000,144,0.5)**  
   Calculate the mean, median, and standard deviation of myDist (you can use functions this time!). How many data points have values higher than the median? Create a histogram of myDist and note the shape. If you drew a vertical line at the 99th percentile of the histogram, how many data points would have values higher than that line? Add a comment with your answers.
5. In your new role as a traffic analyst, you set up your measuring station alongside a busy road where the speed limit is 30 MPH. Over the course of 24 hours, you collect measurements of the *speed* of each vehicle that passes, the number of new vehicles that pass your location each minute, and, through a cool new algorithm, the approximate *weight* of each vehicle. Use random number generators in R-Studio to create plausible distributions of observations for each of these three variables. Calculate the mean, median, and standard deviation of each distribution and create a histogram for each one, noting the shape in a comment.