Hi All,  
  
For Jan 31  
  
1) R does not have a built in function for calculating the standard error  
of the mean (SEM).  Create and demonstrate use of a function to calculate  
the SEM. (Send your code)  
  
HINTS: The formula for the SEM is the standard deviation divided by the  
square root of the number of samples.  You can determine the number of  
items by using length().  i.e.  
  
> myData <- c(93,45,67,23)  
> length(myData)  
[1] 4  
  
  
BONUS: have your function deal with missing values (NAs) correctly.  Do not  
solve this problem by googling and finding someone else's solution.  
  
ALTERNATIVE: If you are stuck on the above, write a function to return the  
product of two numbers.  
  
2) Read and work the examples in Chapters 1 ("Introduction") and 2  
("Getting started with qplot") of the ggplot2 book.  (Hadley Wickham, \*  
ggplot2\*, Springer: New York). This book can be accessed for free via the  
UC Davis library website.  
  
   - Go the UC Davis library website  
   - if you are off campus, click on the VPN link  
   - search the catalog by title for "ggplot2"  
   - click on the entry  
   - click on the online link.  
   - note that once you are on the Springer website you download PDFs of  
   each book chapter.  You can also purchase the book from this site at more  
   than 50% off.  
  
Send a transcript of your R session(s) (ie the console output)  
  
If you have problems installing ggplot2, ask me, Susan, or Mike for help.

Chp1:

Vocab:

The data that you want to visualise and a set of aesthetic mappings

describing how variables in the data are mapped to aesthetic attributes

that you can perceive.

• **Geometric objects, geoms** for short, represent what you actually see on

the plot: points, lines, polygons, etc.

• **Statistical transformations, stats** for short, summarise data in many useful

ways. For example, binning and counting observations to create a histogram,

or summarising a 2d relationship with a linear model. Stats are optional,

but very useful.

• The **scales map values** in the data space to values in an aesthetic space,

whether it be colour, or size, or shape. Scales draw a legend or axes, which

provide an inverse mapping to make it possible to read the original data

values from the graph.

• A **coordinate system, coord** for short, describes how data coordinates are

mapped to the plane of the graphic. It also provides axes and gridlines to

make it possible to read the graph. We normally use a Cartesian coordinate

system, but a number of others are available, including polar coordinates

and map projections.

• A **faceting** speciﬁcation describes how to break up the data into subsets

and how to display those subsets as small multiples. This is also known as

conditioning or latticing/trellising.

3) Do the following chapter two exercises  
  
#2.1a From the tomato data set (available on SmartSite under Resources) use  
qplot to plot totleng as a function of altitude.  
  
#2.1b Add a linear regression line to this plot  
  
#2.1c Add a locally smoothed line instead.  
  
#2.1d color the dots by light condition ("trt")  
  
#2.1e make the plot symbol correspond to species  
  
Send your code.  
  
Please include your NAME in the filename of your homework.