

---

## Statistics 251: Lab 1 Exercises

### Part 1

Create a matrix with 3 rows and 5 columns such that it contains the numbers 1, 5, 9, 13, 17, ..., 57. Make sure the numbers are increasing column-wise; ie, 5 should be in the first column. Use the `seq()` function to generate the numbers. Don't type them out by hand!

### Part 2

In R, 'nlme' is a library or package containing many useful functions and datasets. We will focus on using one dataset inside this library, called Earthquake. In the R console, type:

```
library(nlme)
Earthquake
str(Earthquake)
?Earthquake
```

#### Questions:

1. How many variables are contained in this dataset?
2. Calculate the mean of *distance*. Calculate the sample variance of *distance* manually by translating the below equation of variance into R code,

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Note that the `length()` function can give you the length of a vector.

Confirm that you performed the calculation of variance correctly by comparing your result with `var(x)`.

3. Generate a histogram of *accel* and describe the shape of the histogram. Calculate the five-number summary for *accel*.
4. Generate side-by-side boxplots, comparing *Richter* to *soil*. Does it seem that the *Richter* intensity differs by *soil* type?
5. Generate a plot that illustrates the relationship between *distance* and *accel*. Describe the relationship between these two variables.