# Statistical Inference Project Part 1 Comparing the exponential distribution in R and the Central Limit Theorem

***Lawrence Lau***

This is Part 1 of the course project from the Statistical Inference class in the Data Scientist Specialization course. The objective is to explore the exponential distribution in R and compare it with the Central Limit Theorem.

## Background

From Wikipedia (http://en.wikipedia.org/wiki/Central_limit_theorem):
The central limit theorem (CLT) states that, given certain conditions, the arithmetic mean of a sufficiently large number of iterates of independent random variables, each with a well-defined expected value and well-defined variance, will be approximately normally distributed, regardless of the underlying distribution.

## Overview

First, we simulate an exponential distribution of 40 exponentials. We expect this sample distribution to not be normally distributed, to have a mean of 1/lambda, a standard deviation of 1/lambda, and a variance of lower case sigma squared, or the standard deviation squared. Wikipedia (http://en.wikipedia.org/wiki/Exponential_distribution) provides further color on the derivation of the calculations. For this project, lambda = 0.2.

Secondly, we simulate the sampling distribution of the mean from the exponential distribution. We take the average of 40 exponentials 1,000 times. Given that we are taking the arithmetic mean 1,000 times, we expect the result to be normally distributed. Accordingly, we expect the mean to be the same as the exponential distribution, 1/lambda, the variance to be lower case sigma squared divided by n, where sigma = standard deviation and n = sample size used to compute the mean, and the standard deviation to equal the standard deviation of the sample population divided by the square root of n, again where n = sample size.

## Simulations