Exercise A: Understanding random sampling and the law of large numbers

In this exercise, we will visually explore random samples of different sizes from a given population. In particular, we will try to demonstrate that larger sample sizes tend to be more reflective of the underlying population.

1) Open the file "ExA_SmpDst.xlsm".
2) If prompted, select "Enable Macros".
3) Navigate to the "Randomize" worksheet, which allows you to choose a random sample of size "Sample Size" from the data contained in the "control" worksheet.
4) Enter "10" for "Sample Size and click the "Randomize" button. Observe the distribution of the various characteristics between Treatment, Control and Expected. With a sample size this small, the percentage difference from the expected average is quite high for reading scores. Click "Randomize" multiple times and observe how the distribution changes.
5) Now, try "50" for the sample size. What happens to the distributions? Randomize a few times and observe the percentage difference for the reading scores.
6) Increase the sample size to "500", "2000" and "10000", and repeat the observations from step 5. What can we say about larger sample sizes? How do they affect our Treatment and Control samples? Should the percentage difference between Treatment, Control and Expected always go down as we increase sample size?

Resource: Abdul Latif Jameel Poverty Action Lab Executive Training: Evaluating Social Programs
Dr. Rachel Glennerster, Prof. Abhijit Banerjee, Prof. Esther Duflo