
L12: Central Limit Theorem Lab

Work in your small group to solve the following questions. Please complete the first page on pen and paper and then use R to confirm your results on the second page. The `college_data` dataset will be used from the `MSMU` library.

1. What does the Central Limit Theorem tell us? Be as specific as possible

2. The `Terminal` variable within the `college_data` dataset is highly skewed with a mean of 79.70 and a standard deviation of 14.72. Approximate the mean and standard deviation for a distribution of sample means relating to this variable. You may need to use a calculator to carry out the operations.

sample size	mean	standard deviation
n = 1		
n = 4		
n = 10		
n = 20		
n = 30		
n = 200		

-
3. Carry out a simulation in R for the distribution of sample means relating to the **Terminal** variable to confirm the results you got on the previous page. Also describe how your histogram looks in terms of skew/normal-ness.

sample size	mean	standard deviation	histogram
n = 1			
n = 4			
n = 10			
n = 20			
n = 30			
n = 200			

4. Based on the previous problems, what can we say about the Central Limit Theorem and why is it so important for us data scientists?

5. Turn in this paper filled-in and turn in your RMarkdown file running through your simulations to the Dropbox.