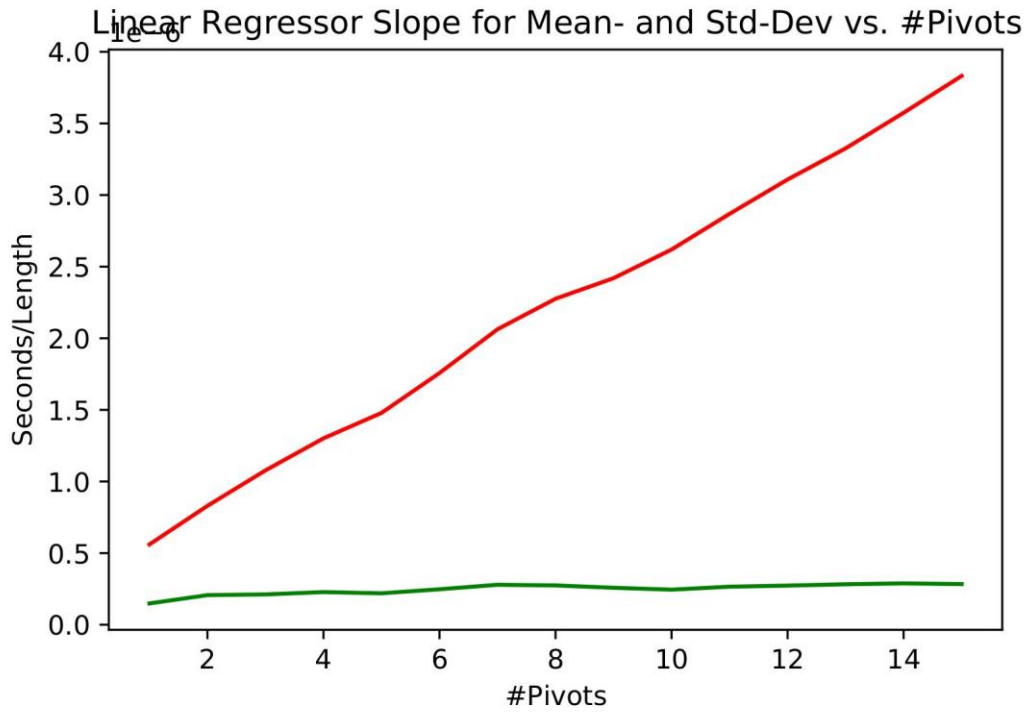
Programming Assignment 3

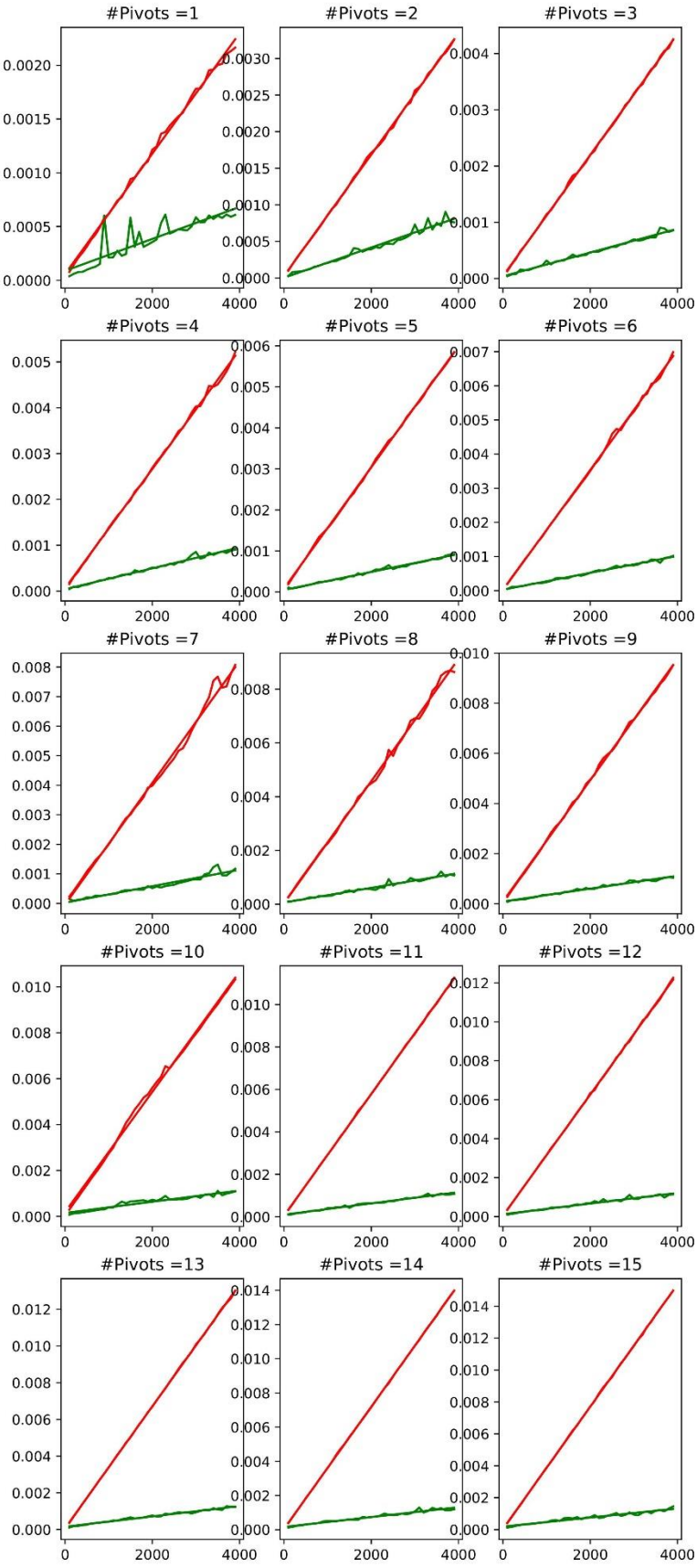
From Homework 2, we theoretically showed that the average running-time of Randomized-Selection Algorithm with m pivots is:

$$A_m(n) \leq \frac{2(m+1)}{m} cn$$

This suggests that the average-running time decreases as m increases. But a larger value of m means more comparisons in the process of selecting proper subarray, which increases the computational cost. Thus, we could suggest that there is no practical benefit of using multiple pivots. At least, there will be a point beyond which using more pivots will be more computationally time-consuming over the single-pivot case. In this programming assignment we trued to empirically find an optimal value of m.

The experiments are conducted on my personal laptop, which has other confounding factors, thus, the results might be different from the ones presented in the Assignment instructions. Since the running-time plot (slopes of the best-fit linear regressor vs number of pivots) does not have a "U-shape" there is no optimal value of m, and since plot has an upward trend, there is no computational advantage of using multiple pivots.




#Pivots =1
#Pivots =2
#Pivots =3
#Pivots =4
#Pivots =5
#Pivots =6
#Pivots =7
#Pivots =8
#Pivots =9
#Pivots =10
#Pivots =11
#Pivots =12
#Pivots =13
#Pivots =14
#Pivots =15