

CPSC 331 Empirical Exercise III

Ali Akbari

TOTAL POINTS

3 / 3

QUESTION 1

1 Report **3 / 3**

✓ **+ 3 pts** Good:

Clear and consist data collection.

Observations properly explained and connected to theory.

Plot shows expected trends, clear explanations included.

Report generally well-written.

- **0.5 pts** Inaccurate mean partitioning ration or not mentioned for some k (it is about 0.75)

+ **2 pts** Fair:

Some evidence of data, incorrect data etc.

Non-technical explanations, incorrect analysis/interpretation of data etc.

Plot included but not explained, unclear trend etc.

Report written well for the most part but has grammar / style issues.

+ **1 pts** Deficient:

No evidence of data, little to no effort made to collect data.

Data not explained/connected to theory.

Plot lacking, poorly written report.

+ **0 pts** Not done:

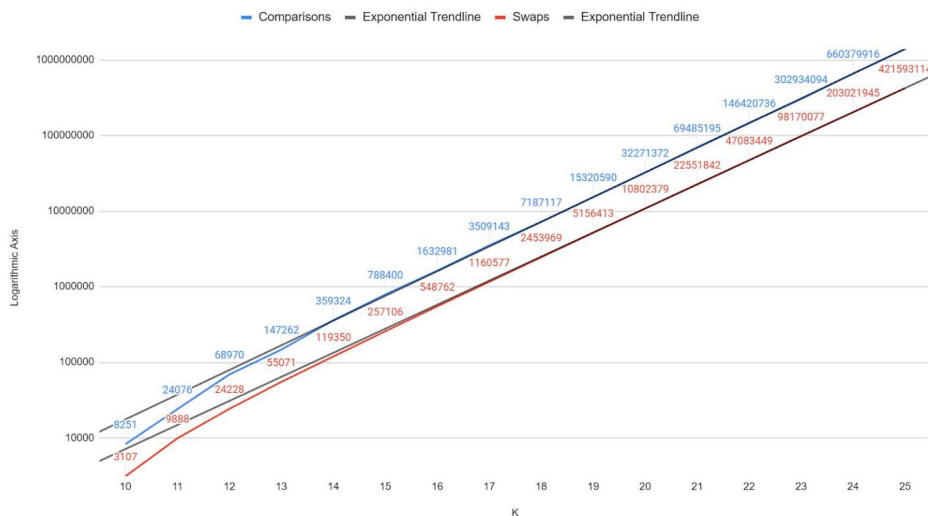
Nothing submitted.

- **0.5 pts** minor mistake or missing

Data Table /Chart

k:	#comps	#swaps	Ratio Min	Ratio Max	Ratio-mean
10	8251	3107	0.5	1	0.7517254335506212
11	24076	9888	0.5	1	0.7557678214777293
12	68970	24228	0.5	1	0.7584961154552092
13	147262	55071	0.5	1	0.7517175389084219
14	359324	119350	0.5	1	0.7542070839022786
15	788400	257106	0.5	1	0.753824157577661
16	1632981	548762	0.5	1	0.7550041709706712
17	3509143	1160577	0.5	1	0.7548572926334607
18	7187117	2453969	0.5	1	0.7554173029557153
19	15320590	5156413	0.5	1	0.7551763105187846
20	32271372	10802379	0.5	1	0.7551341867477577
21	69485195	22551842	0.5	1	0.7552929530042592
22	146420736	47083449	0.5	1	0.755214946223507
23	302934094	98170077	0.5	1	0.7552703563287357
24	660379916	203021945	0.5	1	0.7553595021534296
25	1387949424	421593114	0.5	1	0.7554399057942583

Number of Comparison & Swaps VS K (power)



Methodology:

I had multiple counters, one to count the number of swaps, one to count the number of comparisons, and one to count the number of partitions made. Swap counter incremented every time in the swaps function. I.e increments whenever the function is called. Whenever compareTo function is used to compare the comparison counter increments. Every time the partitions are made in the two if statements the partition counter is incremented. Done for both the left and right partitions. Also compared both left and right partitions to find larger partitions. In the end, I printed the number of comparisons and swaps using the TA provided print format. It also printed the total number of the larger partition divided by the partition counter.

Analysis:

The size of our array exponentially grows in each test, as well as both the number of comparisons and swaps. So when we change the y-axis to show it on a log scale that negates the exponential growth and makes it seem like it is linear. That is expected and can be seen following a similar pattern of an exponential trend line. Any slight variation could be due to the random array favoring the best/ worst case of quicksort. From the table, we see that the mean partitioning ratio averages to around 0.7543.... Which falls between our expected range of $0.5 \leq r \leq 1$. 0.5 being our best case where the array is split evenly at half (50%). For our result, 0.7543 means that on average the sub-array/partition is 75.43% split of larger array/partition. This value is expected as the ratio average of the ratio is $(0.5(\text{minimum}) + 1(\text{maximum}))/2 = 0.75$. However, it is skewed a little to the higher end, and this could be due to the fact that Java's own random algorithm is not precisely random and also can favor the best or worst case of quicksort. One key takes away is that at lower k values (10,11,12) the number of swaps and comparisons does not follow the trend line, and the higher k value almost matches the trend line. Which is true as for smaller values quicksort is quite faster but not as fast as insertion sort for smaller sized inputs.

1 Report 3 / 3

✓ + 3 pts Good:

Clear and consist data collection.

Observations properly explained and connected to theory.

Plot shows expected trends, clear explanations included.

Report generally well-written.

- 0.5 pts Inaccurate mean partitioning ration or not mentioned for some k (it is about 0.75)

+ 2 pts Fair:

Some evidence of data, incorrect data etc.

Non-technical explanations, incorrect analysis/interpretation of data etc.

Plot included but not explained, unclear trend etc.

Report written well for the most part but has grammar / style issues.

+ 1 pts Deficient:

No evidence of data, little to no effort made to collect data.

Data not explained/connected to theory.

Plot lacking, poorly written report.

+ 0 pts Not done:

Nothing submitted.

- 0.5 pts minor mistake or missing