

NTUEE Algorithm Programming Assignment #1

B10901151 電機三 林祐群

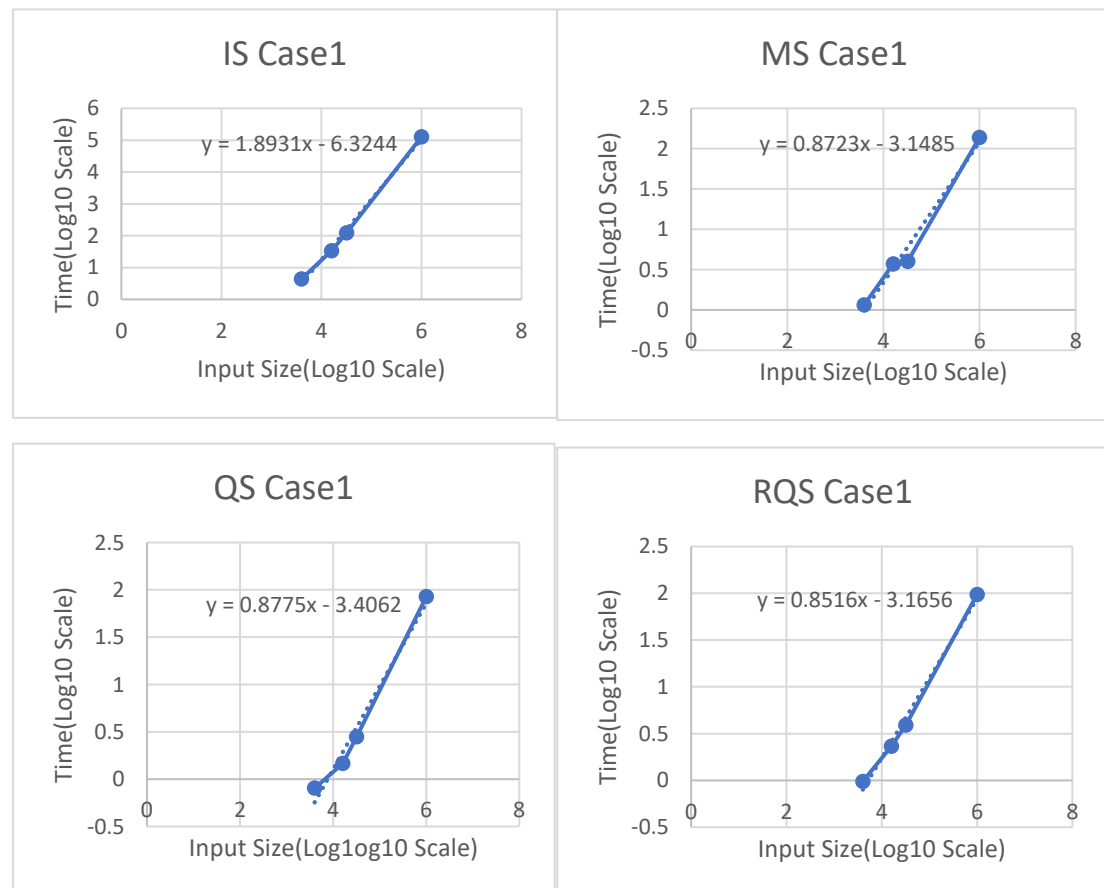
I. Running Time Table(EDA Union port 40060) :

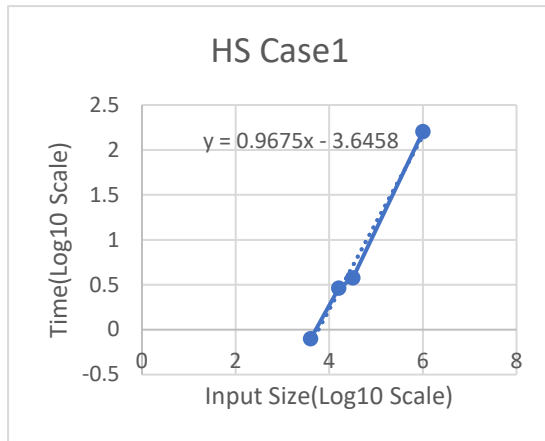
Input size	IS		MS		QS		RQS		HS	
	CPU time(ms)	Memory(KB)	CPU time(ms)	Memory(KB)	CPU time(ms)	Memory(KB)	CPU time(ms)	Memory(KB)	CPU time(ms)	Memory(KB)
4000.case2	0.115	5904	0.528	5904	17.107	6028	0.62	5904	0.654	5904
4000.case3	7.313	5904	0.523	5904	14.395	5904	0.663	5904	0.681	5904
4000.case1	4.438	5904	1.161	5904	0.804	5904	0.976	5904	0.788	5904
16000.case2	0.174	6056	0.96	6056	158.711	6928	1.44	6056	1.483	6056
16000.case3	63.494	6056	1.87	6056	125.164	6432	1.542	6056	1.453	6056
16000.case1	33.703	6056	3.725	6056	1.459	6056	2.312	6056	2.885	6056
32000.case2	0.121	6188	1.97	6188	665.593	8000	1.538	6188	1.86	6188
32000.case3	248.521	6188	1.798	6188	502.91	6948	2.299	6188	2.749	6188
32000.case1	122.595	6188	4.02	6188	2.802	6188	3.899	6188	3.755	6188
1000000.case2	1.332	12144	45.083	13872	597946	72468	49.222	12144	83.169	12144
1000000.case3	258412	12144	48.81	13876	325877	33016	54.463	12144	82.236	12144
1000000.case1	127342	12144	137.687	13880	84.282	12144	96.956	12144	159.064	12144

II. Trendline for five sorting algorithms in three cases :

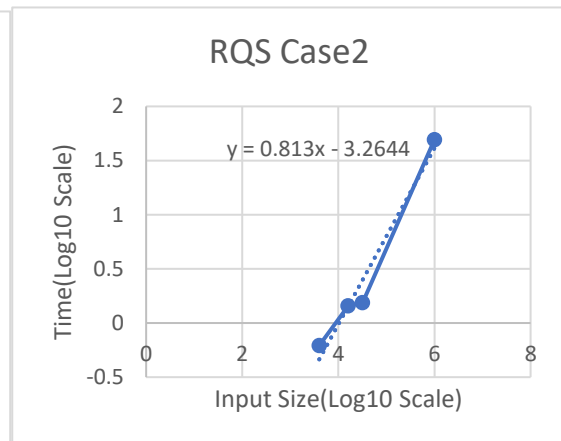
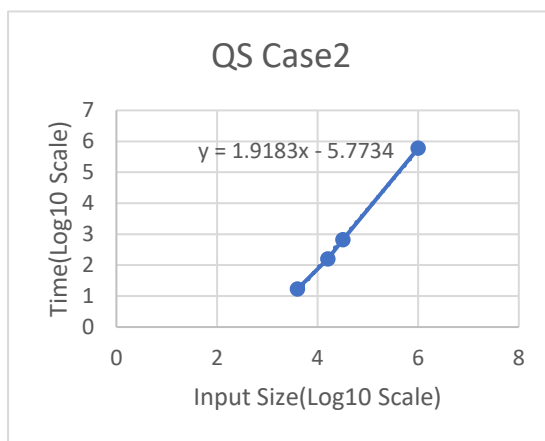
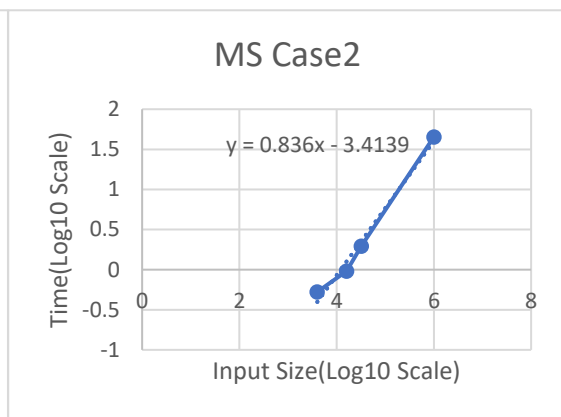
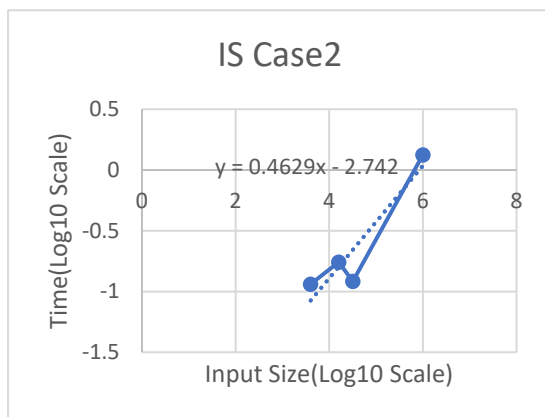
趨勢線之方程式標記於圖上，並以線性為基準。X 及 Y 座標皆為 Log10 轉換後之座標。IS: Insertion Sort ; MS: Merge Sort ; QS: Quick Sort ; RQS: Randomized Quick Sort ; HS: Heap Sort 。

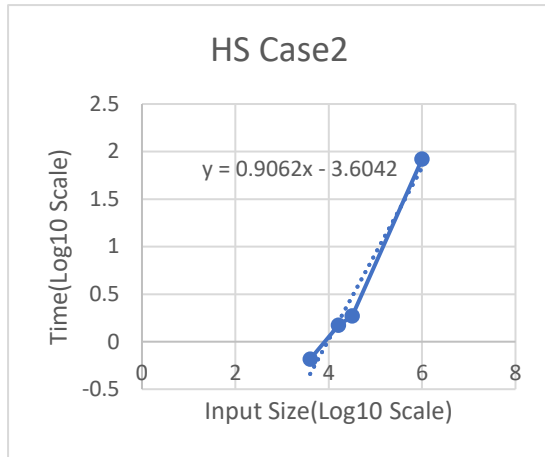
1. Case1(Average Case) :



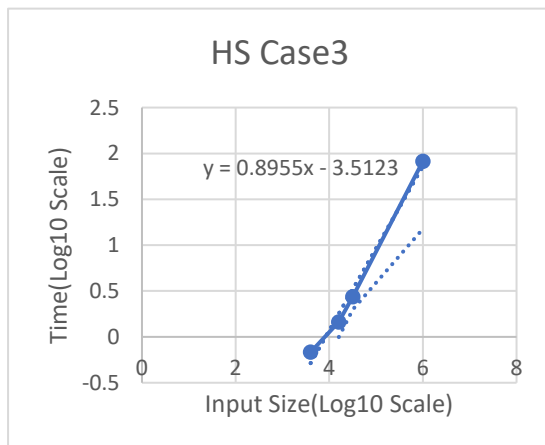
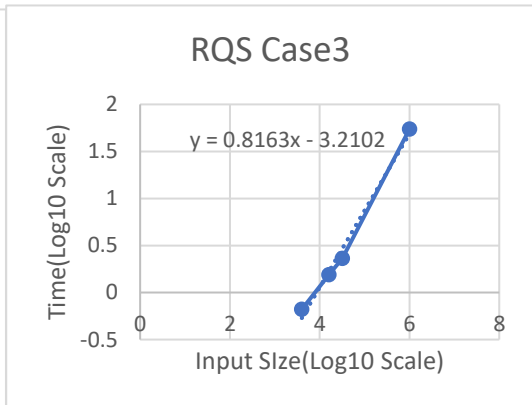
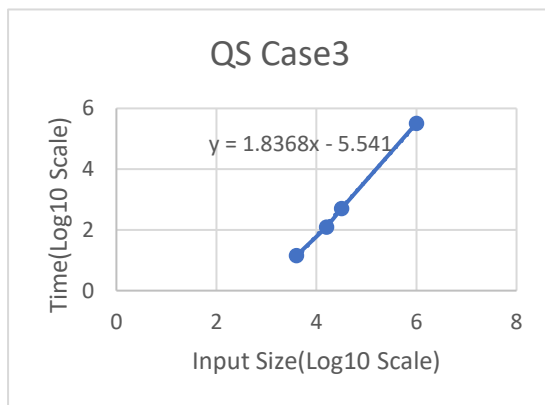
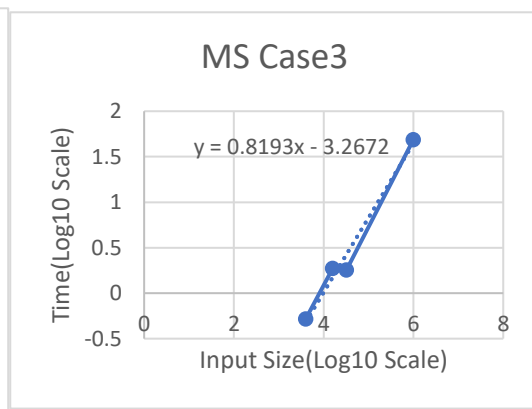
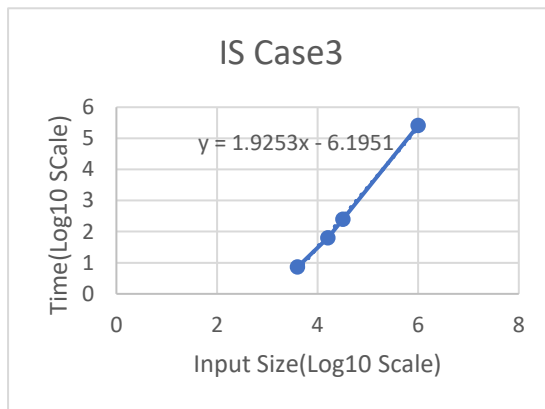


2. Case2(Best Case) :



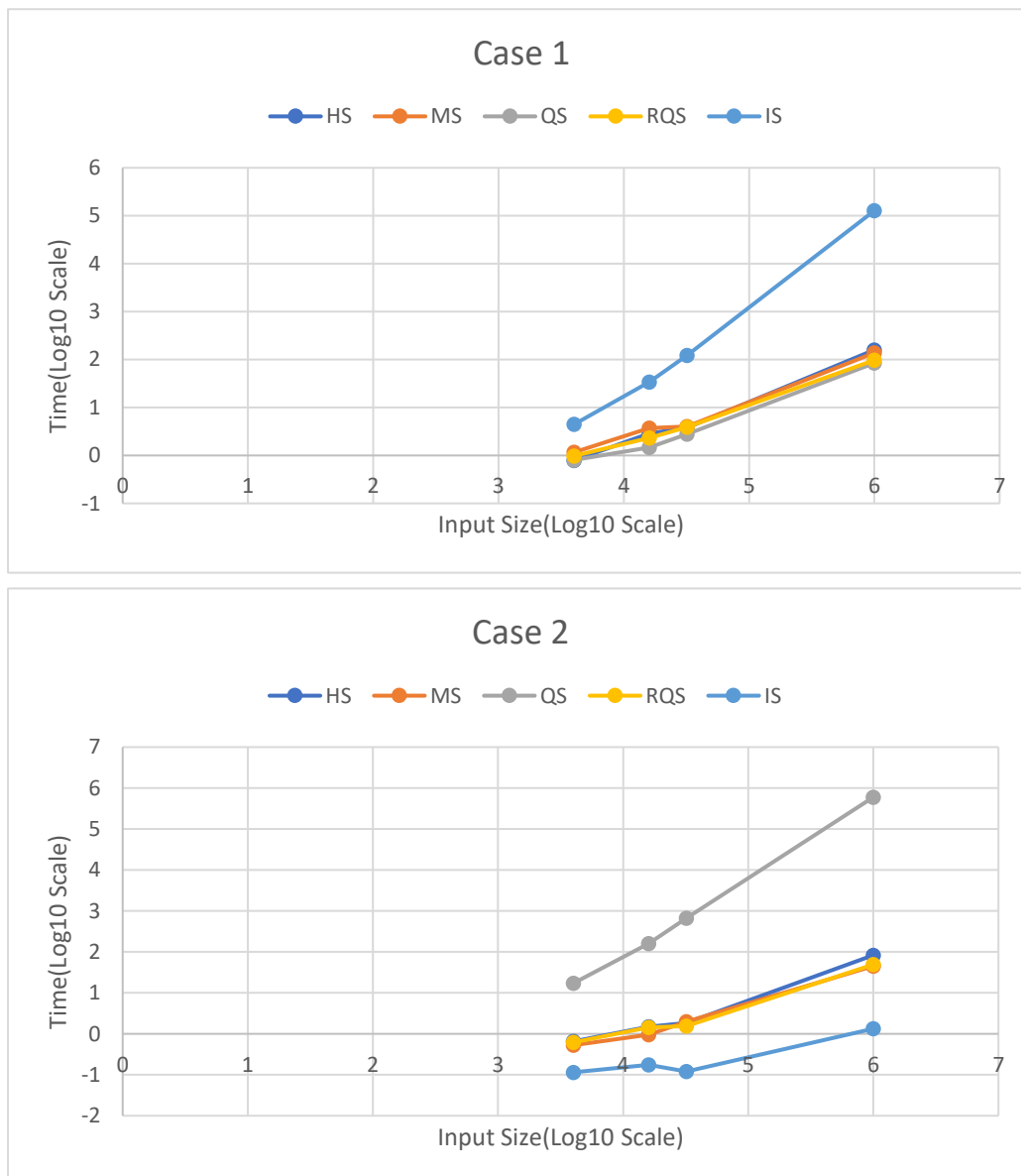


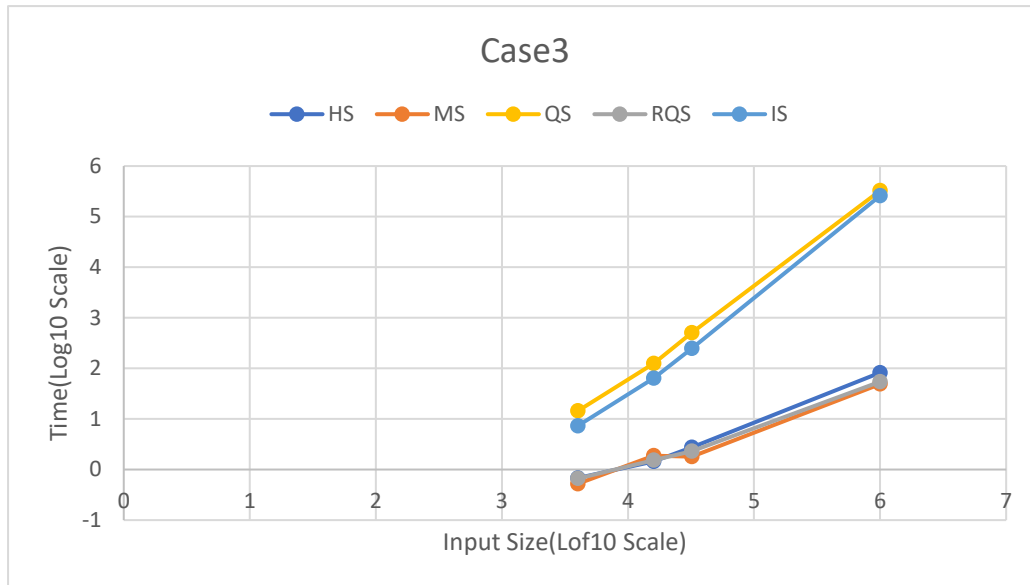
3. Case3(Worst Case) :



III. Graph Analysis :

理論上分析可以得知，IS 和 QS 的 **worst case time complexity** 是 $\Theta(n^2)$ ，而 IS 的 **average time complexity** 也是 $\Theta(n^2)$ ，因為是簡單指數函數，所以理論斜率在雙對數圖的表現為斜直線，且斜率為 2。而 IS 的 **best case running time** 是 $\Theta(n)$ ，所以在雙對數圖上為斜率為 1 之斜直線，其餘之時間複雜度為 $\Theta(n \lg(n))$ ，因此在雙對數圖型上之表現可以簡單表示為 $y' = x' + \log_{10}(x')$ 。以下將五種 sorting algorithm 在相同的 case 下的趨勢圖形化出來：





由 case1(average case)的分析可以發現，只有 IS 的趨勢和其他 sorting algorithms 有明顯的差別，而這和 IS 的 average time complexity 是 $\Theta(n^2)$ 有關。在 case2(best case)中首先可以發現 IS 斜率低於其他演算法，也就呼應了其時間複雜度為 $\Theta(n)$ 低於 $\Theta(n \lg(n))$ 的理論判斷。和理論不同的是 QS 的線性趨勢，以理論來說應該和其他演算法較為接近，因為時間複雜度皆為 $\Theta(n \lg(n))$ ！為了確認這件事情，我試了兩個 port(40055 和 40060)，結果卻是相同，這邊推測是因為一些 $\Theta(1)$ 部分的運算在 EDA Union 上執行會明顯可見，導致整體 Running Time 較其他演算法多上許多。在 Case3(worst case)中則可以發現 QS 和 IS 比其他演算法明顯需要更多時間，而這也呼應他們 $\Theta(n^2)$ 的 worst case time complexity。而綜合觀察，比較 QS 和 RQS 的差別則可以發現，在 case1(average case)中，兩者的圖形近乎重疊，也就是說加入 random selected pivot 這樣的改變之目的並非著重於改善 average case 的時間複雜度，而是希望可以大幅優化 pivot 的選擇，避免遇到 worst case，也因此我們在 case3(worst case)中可以明顯發現 QS 和 RQS 不再重合，也就呼應我們前面理論的分析。

另外，本次 inputs 中並未有所有資料皆為同一個數字的情形，然而如果是 QS 或 RQS 遇到這種狀況，因為 QS 每次會選擇最右邊的數字做為 Pivot，但因為所有數字都相同，所以每次僅能切割成 1 和 n (原先資料大小)-1 的方式一路遞迴下去，但這就如同 worst case 一樣，會得到很不好的 running time。同理，因為所有數字相同，所以 RQS 會表現得如 QS 一樣，選哪個數字做為 pivot 在這個特例之下變得沒有意義也沒有時間複雜度上的助益。