**Project I: Comparison-based Sorting Algorithms**

Implement by: Sumedh Joglekar

Language: Python

Sorting Performed: Insertion Sort, Merge Sort, In-Place Quick Sort, Modified Quick Sort

1. **Insertion Sort** :-

* Algorithm:-

1. Defined an array n which contains the range of numbers for which sorting needs to be performed.
2. In order to measure the time required for the execution stating timer just after main for loop.
3. Array ‘arr’ will contain the all the randomly generated numbers.
4. Calling insertionSort(arr) function which will return the sorted array of sorted elements.
5. In the insertion sort function, setting key = arr[j] where j will be from 1 till the len(arr)
6. While ((i>=0) and (arr[i] > key)), swapping the array elements.
7. Updating the a[i+1] with the existing key.

* Time Complexity of Insertion Sort is O(n^2) in worst case.
* Best Case of insertion sort is sorted array and reversely sorted array and time complexity in best case is O(n).

1. **Merge Sort :-**

* Algorithm:-
  1. Merge Sort is recursive sort which divides given array into exactly half and send to the same function recursively until the length of the array is less than or equal to 1.
  2. In the code, merge\_sort function is dividing the input recursively into two parts, which keeps on creating internal stack till the last call.
  3. Once the last call to merge\_sort is made, program counter goes to merge(left,right), which combines the two part of sorted array.
  4. Sorting is done at the time of merging i.e. in merge function.
  5. Lastly insertion of remaining elements either from left array or right array is done depending upon the size of the array.
  6. At the end, timer is stopped and the difference between stop and start will give the execution time of the entire program.
* Time Complexity of Merge Sort is O(nlogn).
* There is no best or worst case for Merge Sort as we are dividing array in exact half regardless of the size and the position of the elements.

1. **Quick Sort** :-

* Algorithm:-
  1. Array ‘arr’ contains the randomly generated elements with varying length from 500 to 50000
  2. Function inPlaceQuickSort(arr) calls inPlacePartition which is a recursive function.
  3. Here, randomly number from the range is generated and the it is set as the index of the pivot element. Pivot element is chosen corresponding to the same index value.
  4. In the loop, scanning of i from left to right so long as (array[left] <pivot) and Scan j from right to left so long as(array[right]>pivot).
  5. If left<=right,we are swapping the elements.
  6. We will continue to do the same process until the left and right crosses.
  7. Once left and right cross, exchange pivot with a[right]
  8. Continue the same process recursively on the same array.
* Time complexity of Quick sort is O(nlogn) and worst case will be when all the element in the quick sort are sorted or reversely sorted and we continuously choose the lowest or highest element as the pivot. In such scenario time complexity will be O(n^2).

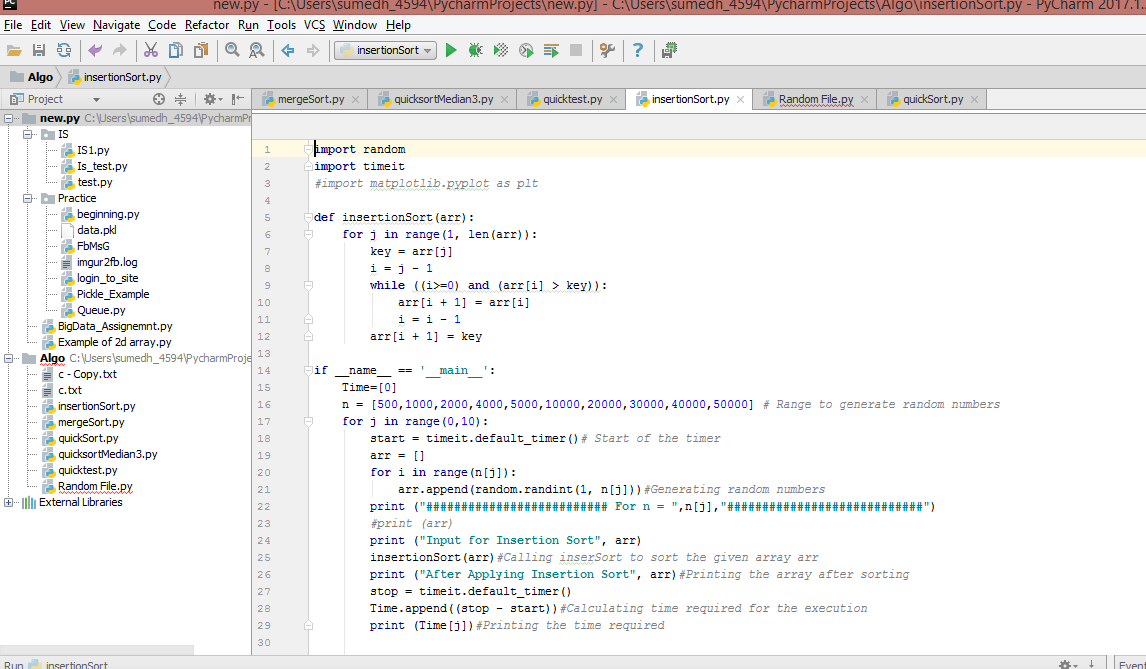
1. **Modified Quick Sort** :-

* Algorithm :-
  1. Pivot selection in the modified quick sort is done by taking the median of the first, last and middle element of the array.
  2. Before selecting the pivot the first, last and middle element is sorted according the value and then the array[center] is selected as the pivot. By this point we have already sorted 3 elements hence the probability of choosing bad pivot decreases.
  3. We will continue to perform the normal quick sort with the pivot as median of three elements until the length of the array becomes 10.
  4. Insertion sort is better option for the sorting of the elements with the length less than 10 and hence remaining sorting is performed with the help of insertion sort.

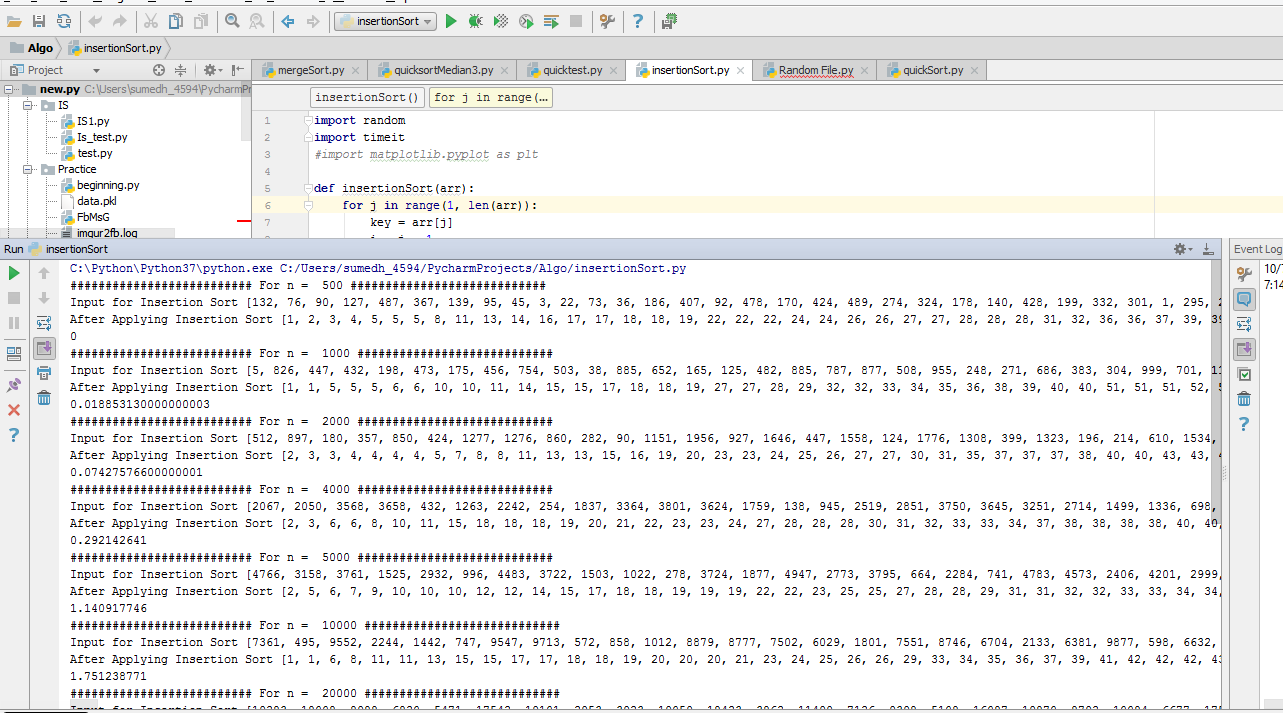
**Codes and Results** :-

1. **Insertion Sort** :-

**Code** :-

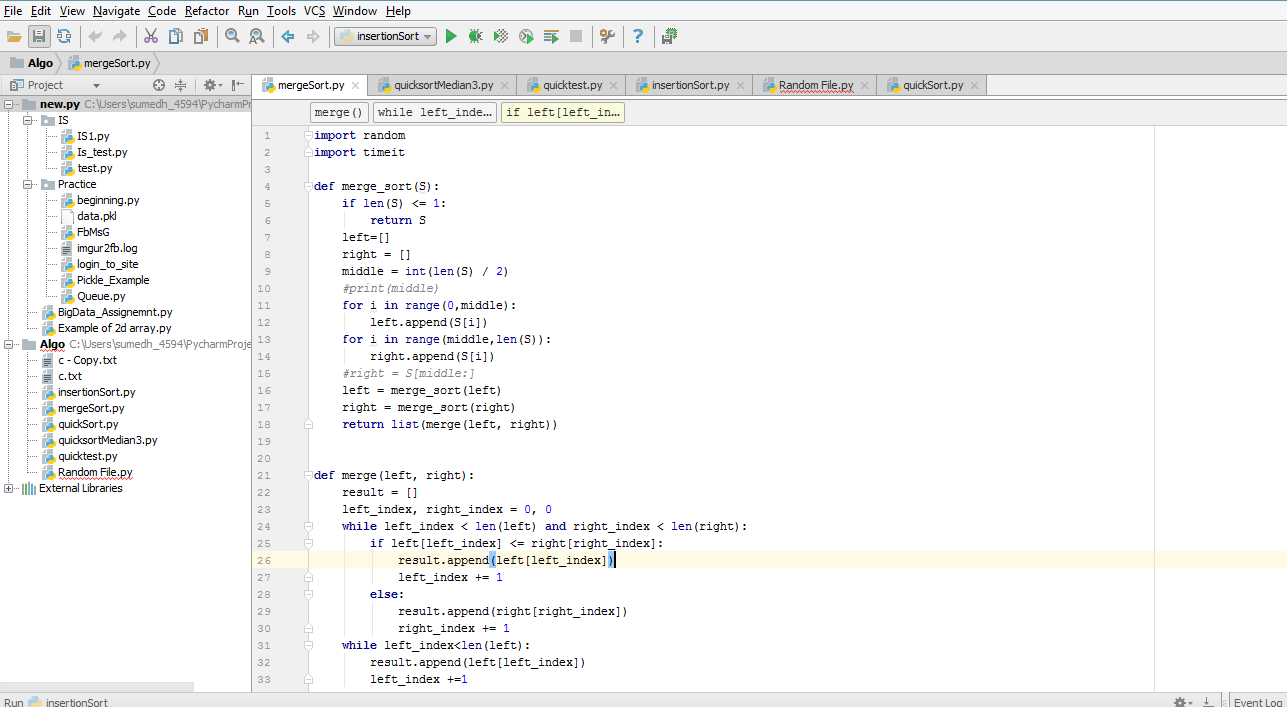


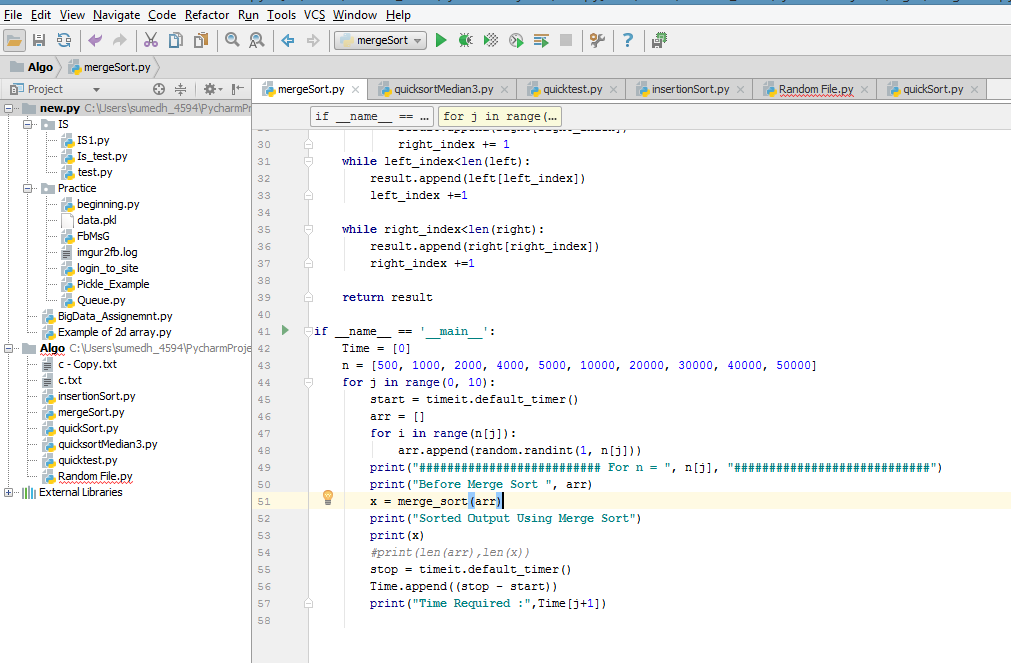
**Result** :-



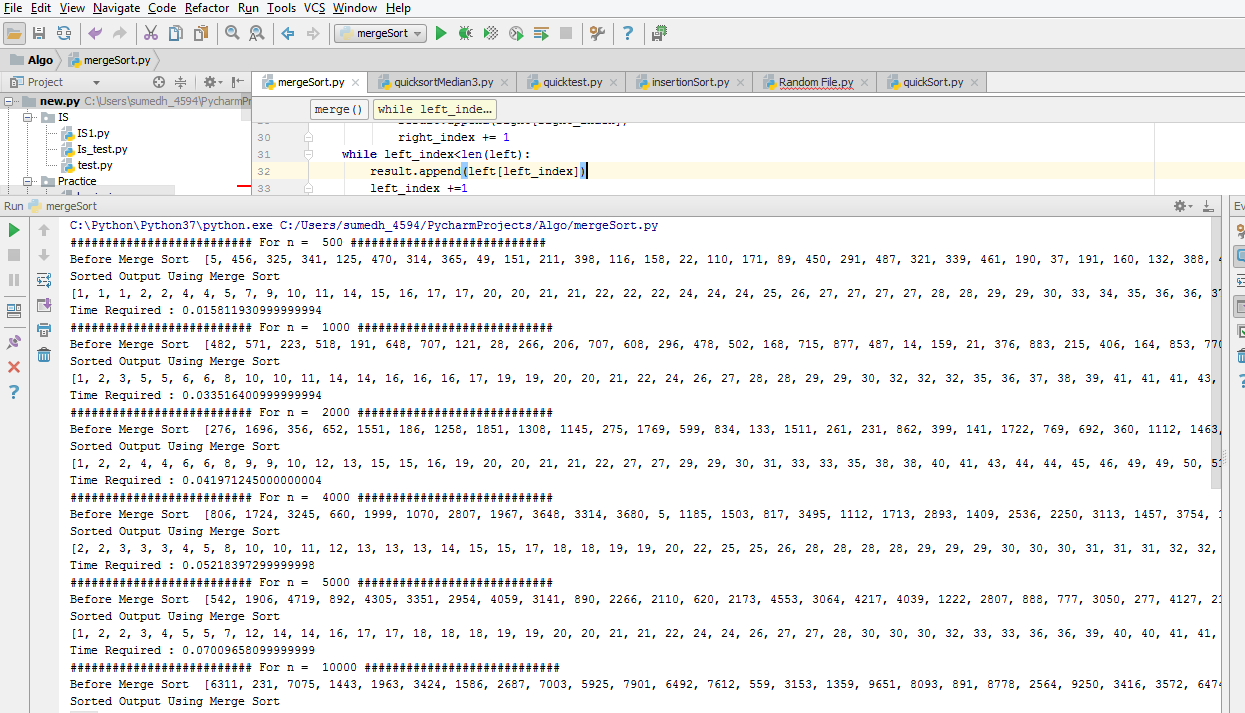
1. **Merge Sort**:-

**Code** :-



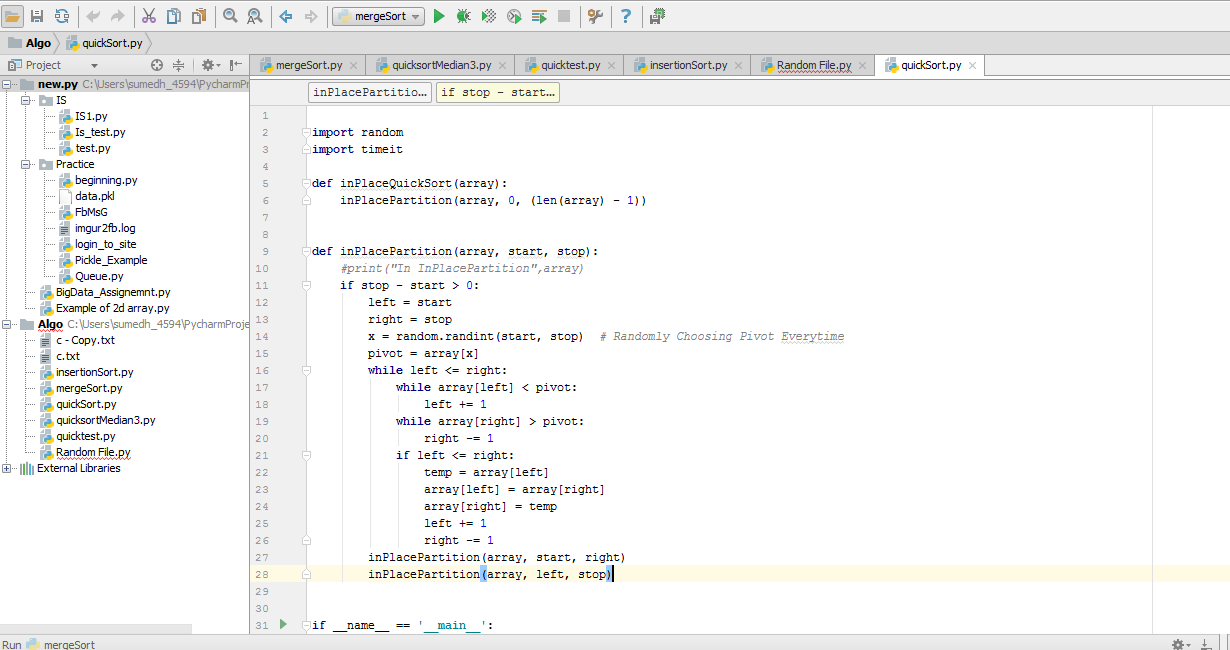


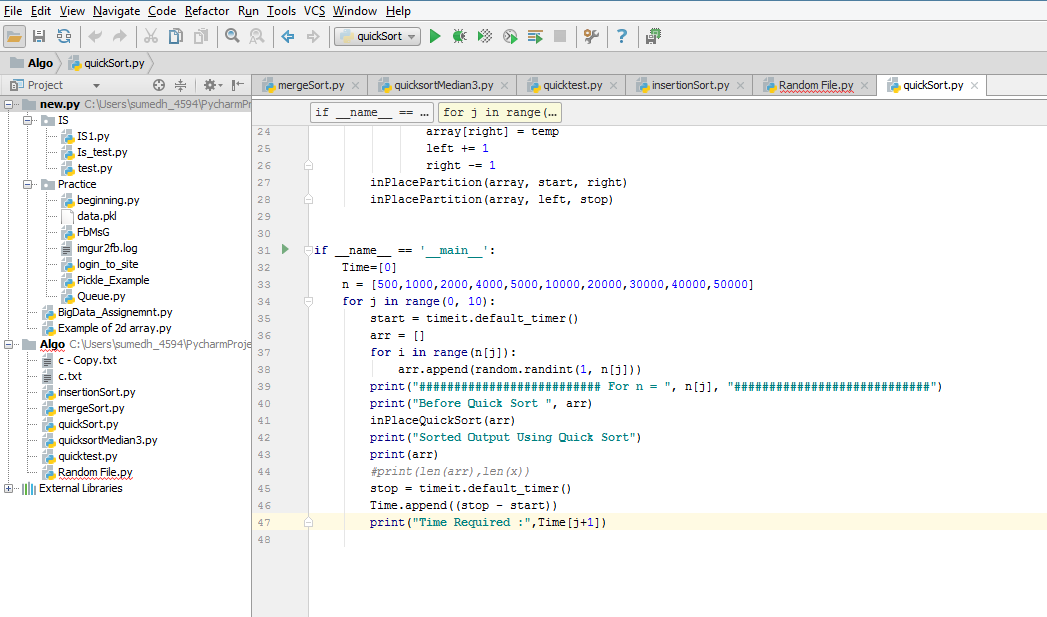
**Result** :-



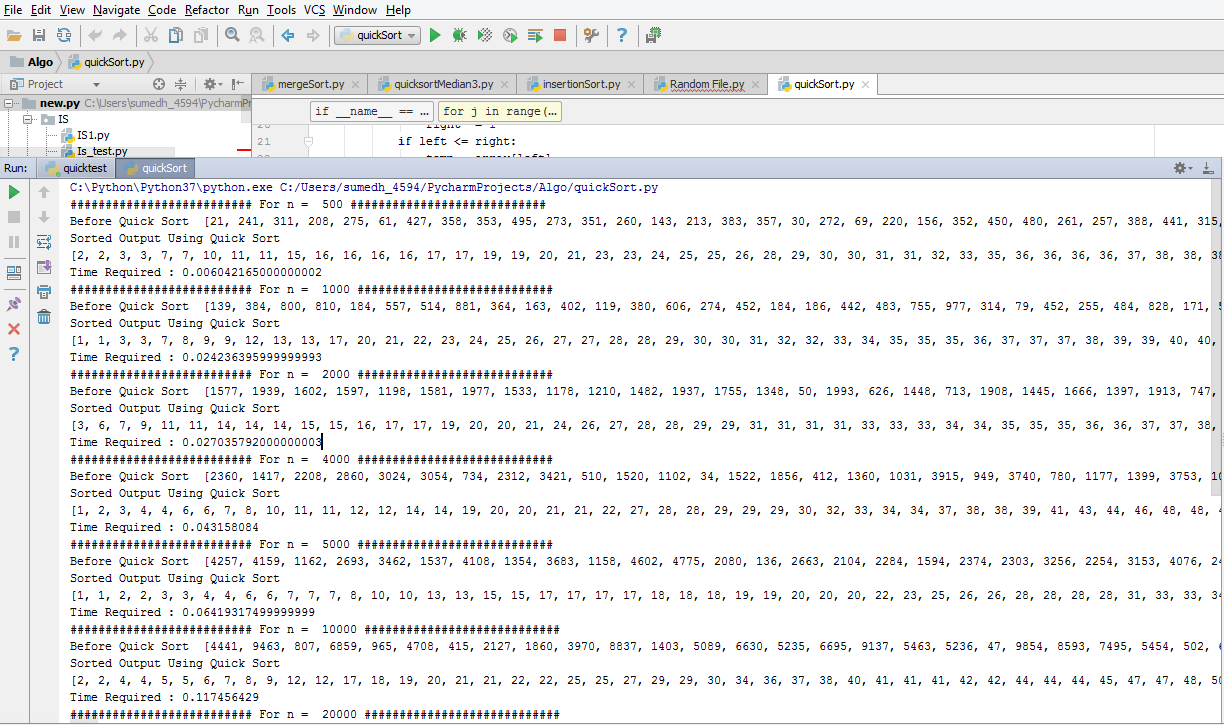
1. **In-Place Quick Sort** :-

**Code** :-



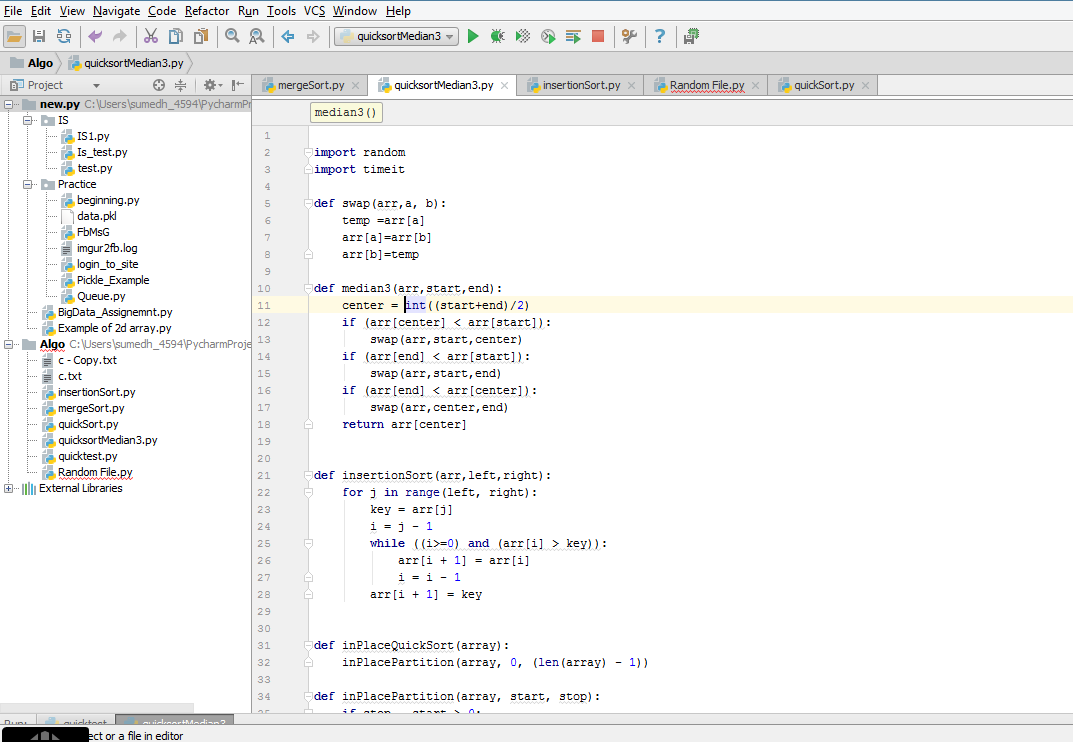


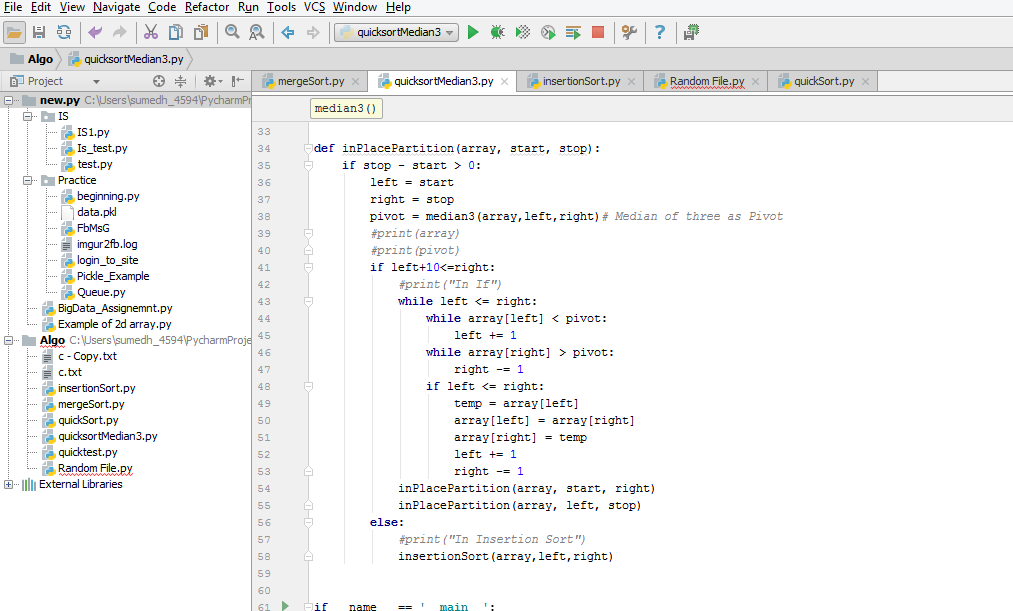
**Result** :-

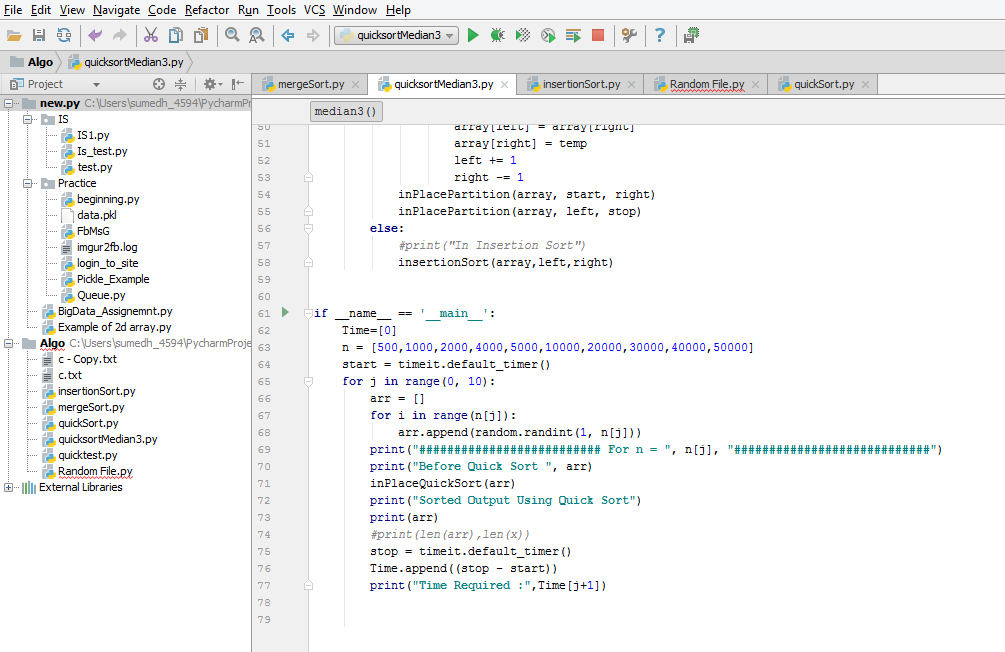


1. **Modified Quick Sort**:-

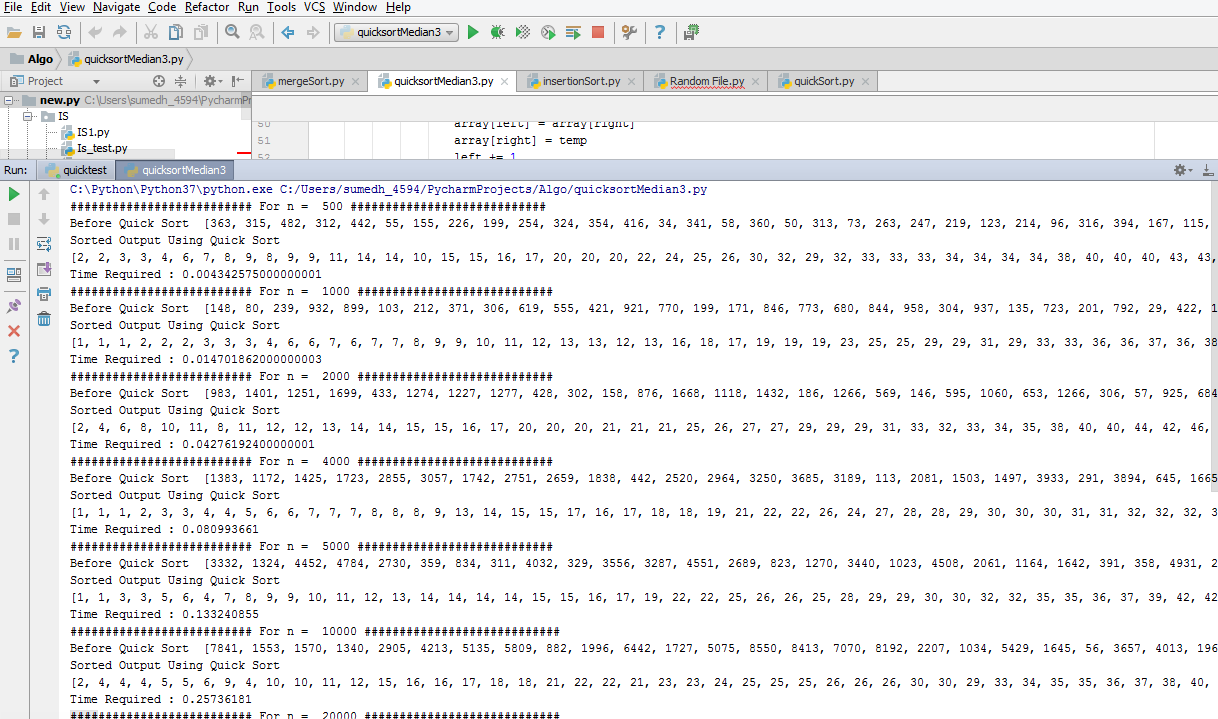
**Code** :-







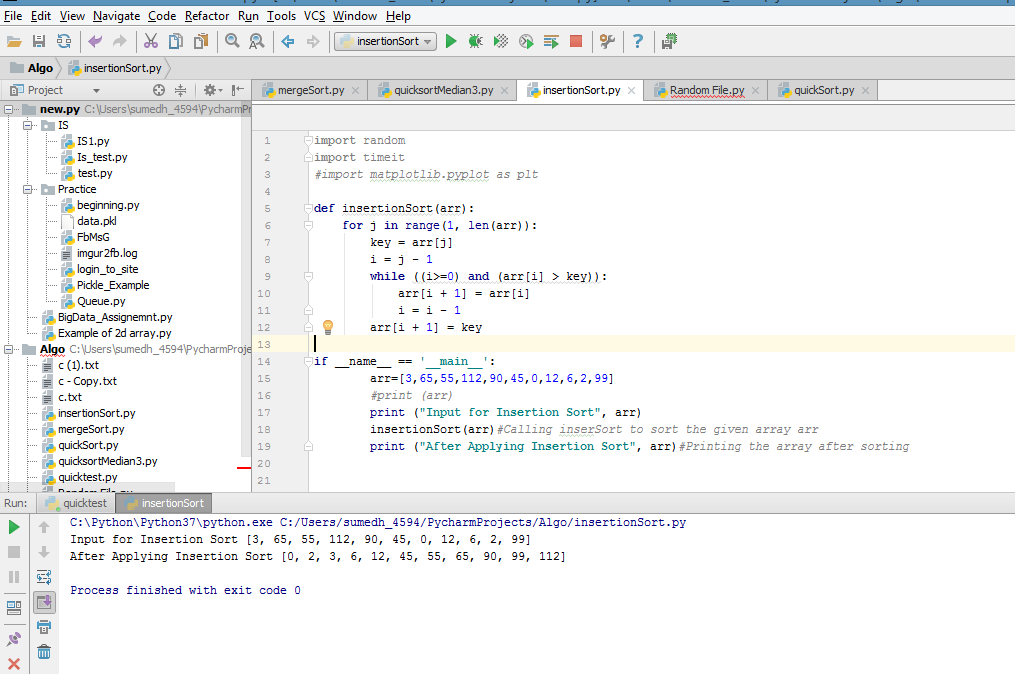
**Result**:-



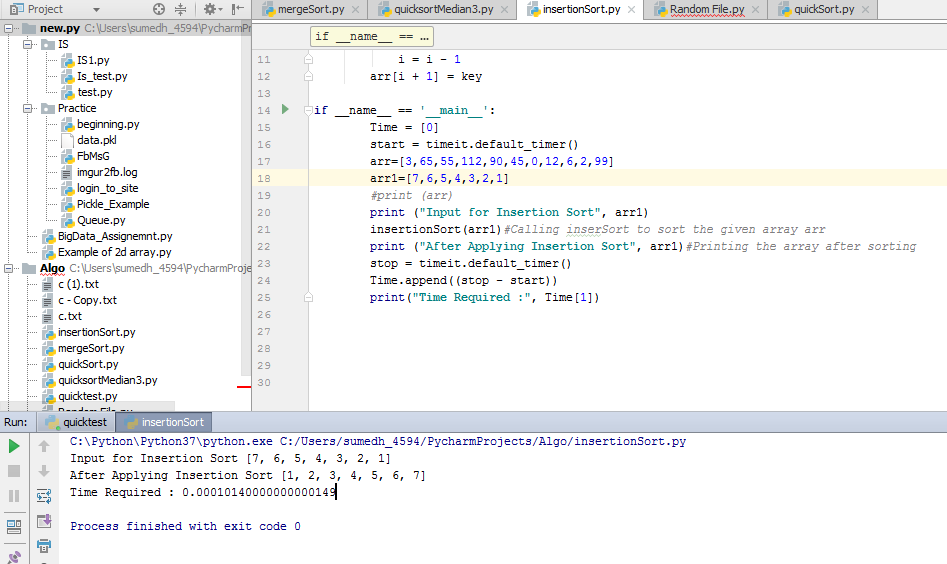
**Different Outputs :-**

Insertion Sort

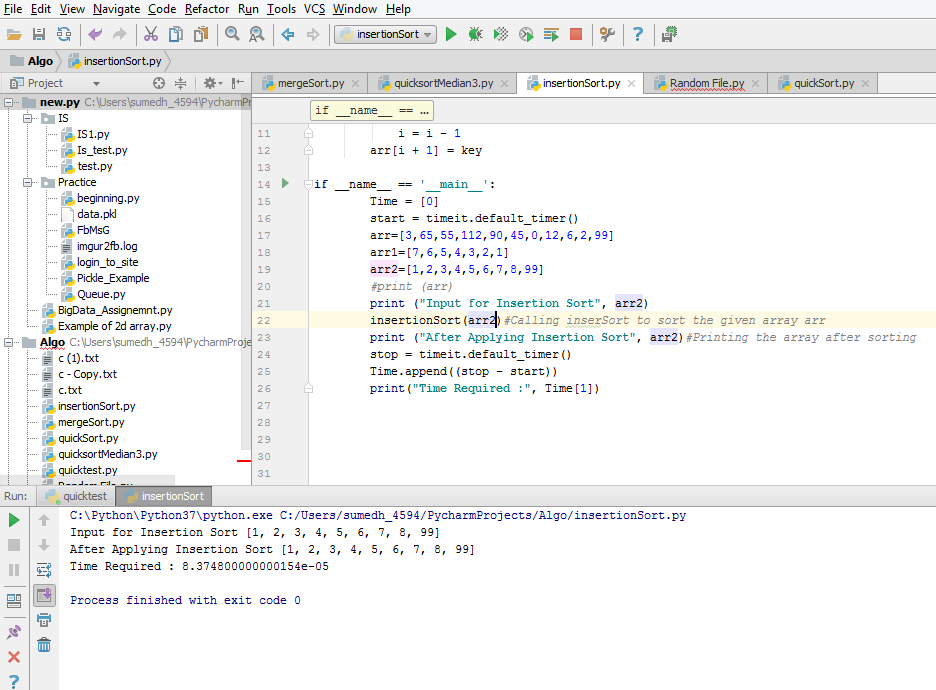
1. Random array :-



1. Reversely Sorted array :-



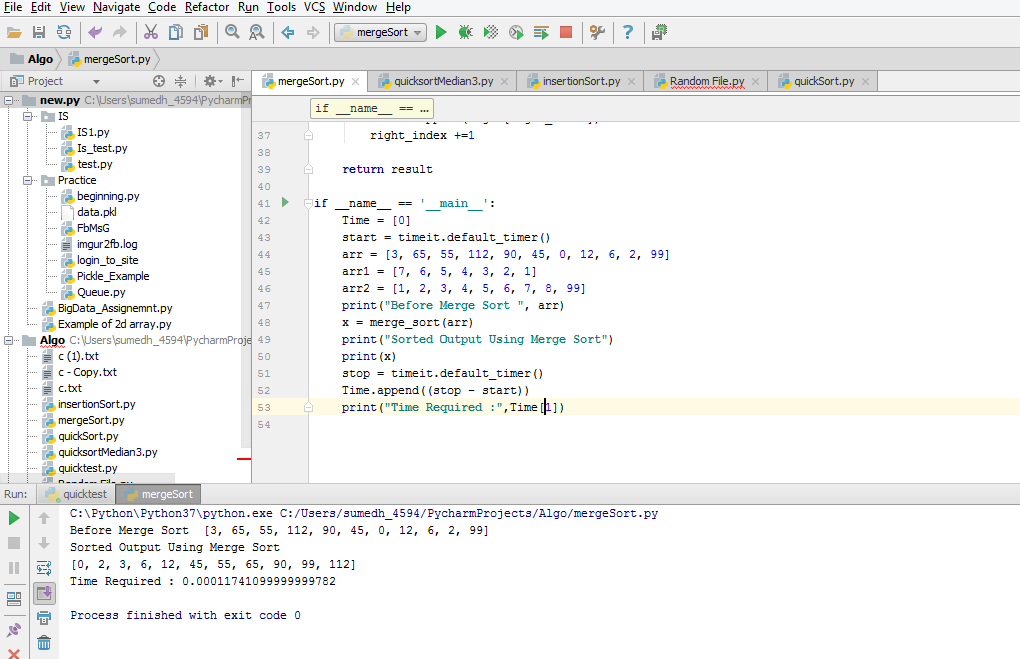
1. Sorted Array :



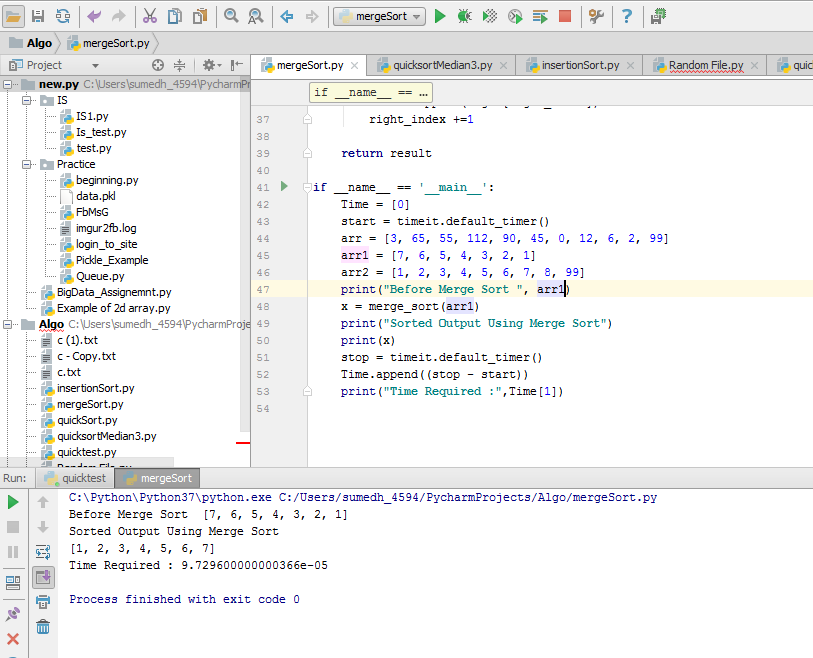
* Hence it can be conclude that the execution of the insertion sort is very fast for the sorted array as it does not go into the inner while loop of the code and as a result the time complexity of the algorithm reduces to O(n).

Merge Sort:-

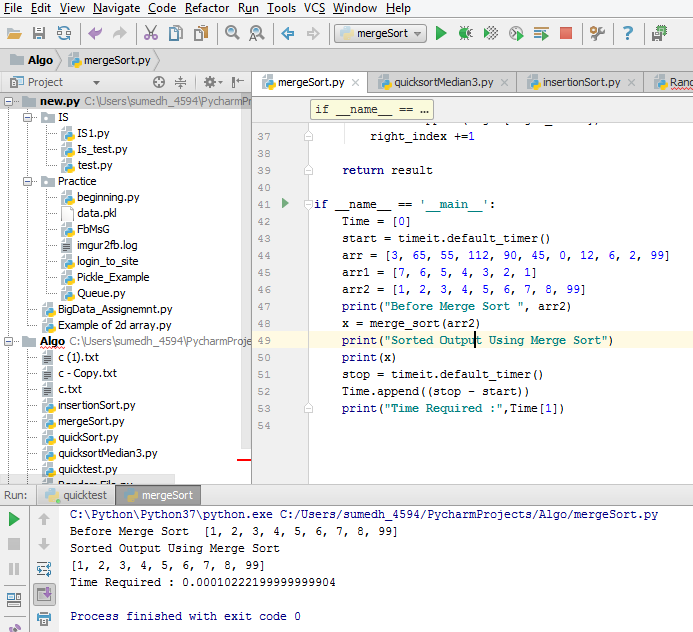
1. Random Input array:



1. Reversely Sorted Array:



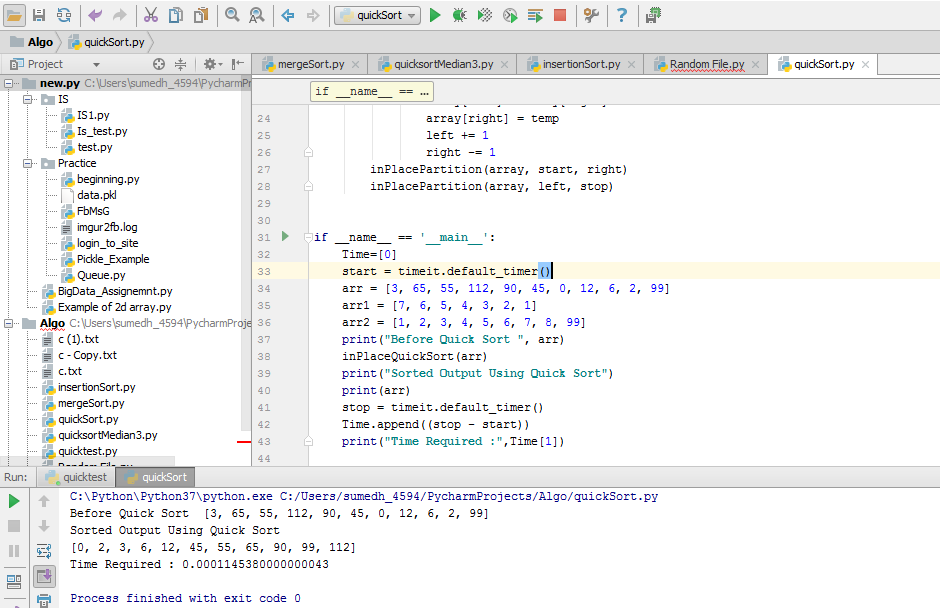
1. Sorted Array :



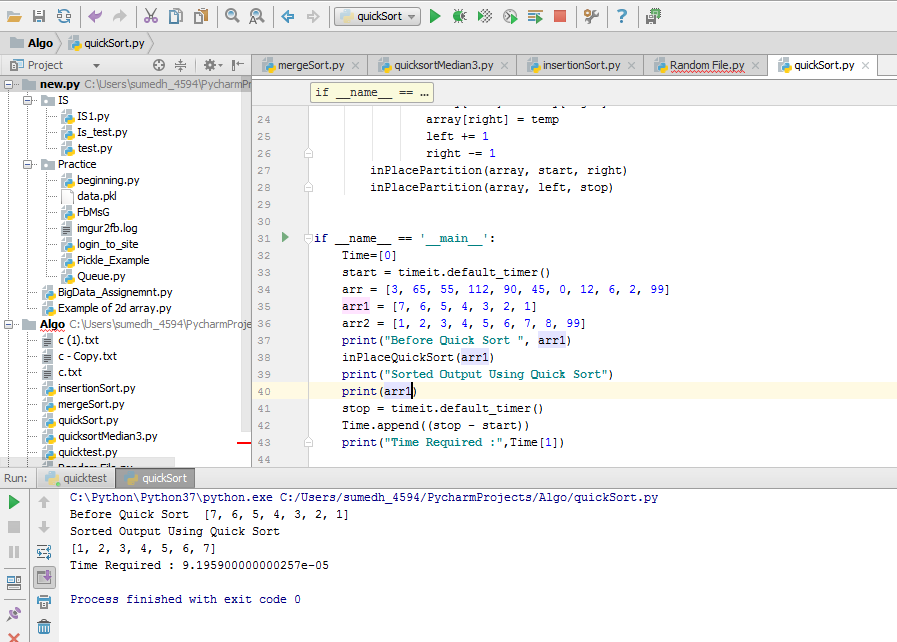
Merge sort is invariant of the value of the elements in an array. It will take perform same amount of work in order to sort the given array.

Quick Sort:

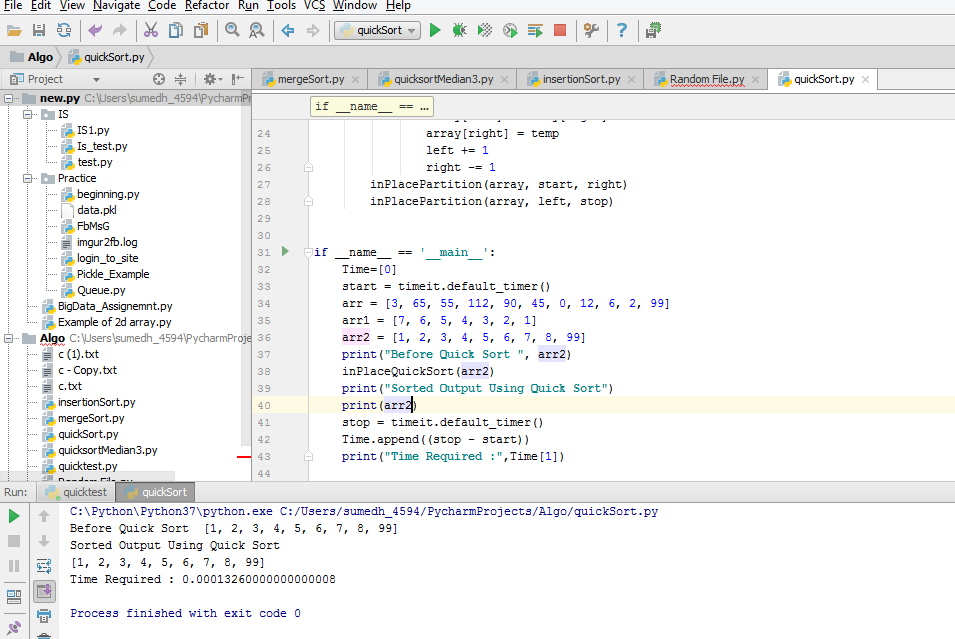
1. Random Input:-



1. Reversely Sorted Array :-



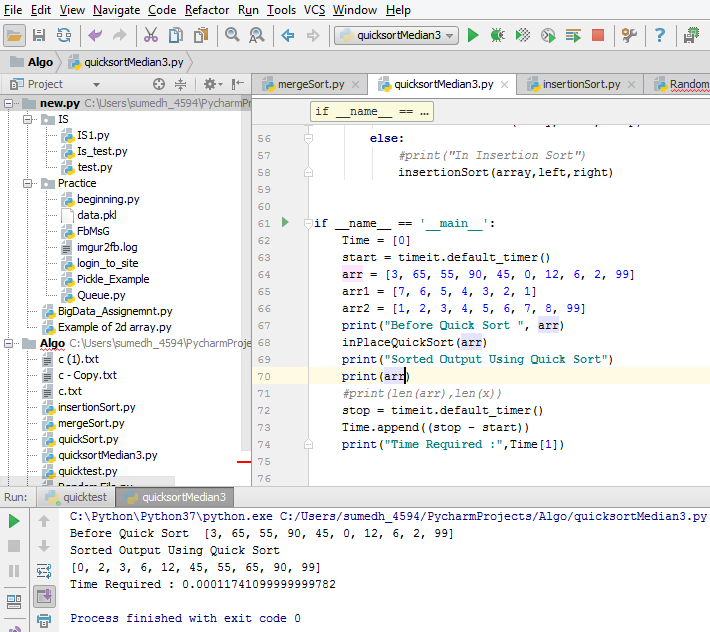
1. Sorted Array :-



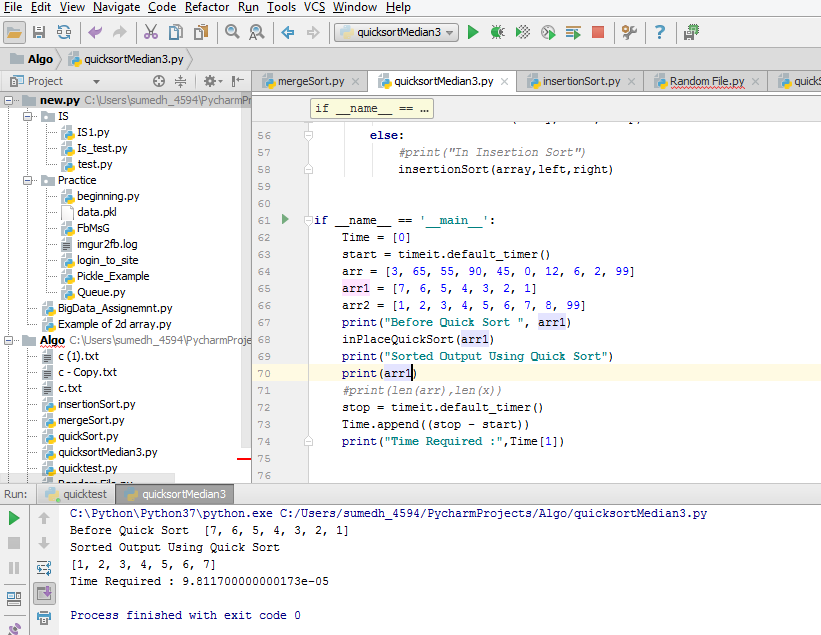
* Quick Sort performs worse for the sorted and reversely sorted array, if we choose pivot as minimum or maximum value. As it divides the array into 3 sub arrays with the length as (0, pivot, length(array)-1) or (length(array)-1, pivot, 0) repetitively and as a result time complexity reaches to O(n^2).

Modified Quick Sort :

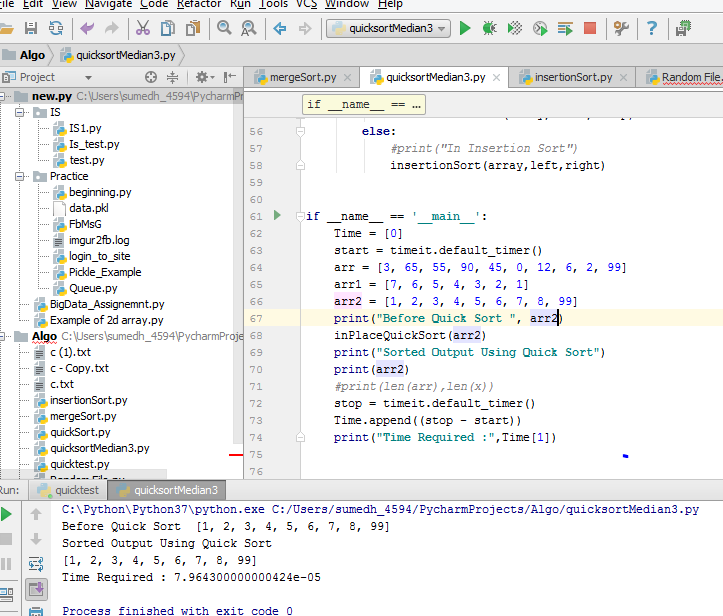
1. Random Input:



1. Reversely Sorted Input:



1. Sorted Input:



* **Output of the different sorting algorithms for large number of inputs.**

1. Code for this particular problem is present in the Random File.py.
2. Here I have imported all the files of individual sorting technique as a library files and have used random and time library from python to generate sequence of random numbers and to calculate time respectively.
3. Time\_Merge, Time\_Insertion, Time\_Quick, Time\_Quick\_Median correspond to the array where I am storing the time required for the n inputs in one iteration. Also, avg\_Merge, avg\_Insertion, avg\_Quick, avg\_Quick\_Median stores the average time required for particular algorithm for different input size.
4. Attached Result.txt contains the result of each iterations corresponding to n input size.
5. Time required is calculated just before the execution related to any sorting begins and it stops once the sorting is completed.
6. Below is the Final output, table and graph captured by the n number of inputs where n varies from 500 to 50,000.

**Output :-**

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 500)

('Insertion Sort :-', 0.018789546666666667)

('Merge Sort :-', 0.008804693333333334)

('Quick Sort :-', 0.002946559999999999)

('Quick With Median Sort :-', 0.0021009066666666625)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 1000)

('Insertion Sort :-', 0.07864447999999999)

('Merge Sort :-', 0.01576277333333333)

('Quick Sort :-', 0.0054677333333333356)

('Quick With Median Sort :-', 0.002996479999999996)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 2000)

('Insertion Sort :-', 0.2559488)

('Merge Sort :-', 0.04808362666666666)

('Quick Sort :-', 0.01817727999999999)

('Quick With Median Sort :-', 0.009412693333333333)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 4000)

('Insertion Sort :-', 1.0254890666666667)

('Merge Sort :-', 0.05067477333333331)

('Quick Sort :-', 0.013908906666666665)

('Quick With Median Sort :-', 0.022728533333333356)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 5000)

('Insertion Sort :-', 1.4276339200000001)

('Merge Sort :-', 0.10775765333333354)

('Quick Sort :-', 0.0294080000000001)

('Quick With Median Sort :-', 0.02731306666666633)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 10000)

('Insertion Sort :-', 4.958820266666667)

('Merge Sort :-', 0.15564629333333313)

('Quick Sort :-', 0.04750591999999987)

('Quick With Median Sort :-', 0.07373610666666686)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 20000)

('Insertion Sort :-', 20.423711146666665)

('Merge Sort :-', 0.4408512000000009)

('Quick Sort :-', 0.09950122666666594)

('Quick With Median Sort :-', 0.07246250666666754)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 30000)

('Insertion Sort :-', 44.106855253333336)

('Merge Sort :-', 0.4268753066666662)

('Quick Sort :-', 0.11471786666666617)

('Quick With Median Sort :-', 0.10657578666666723

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 40000)

('Insertion Sort :-', 67.41122048000001)

('Merge Sort :-', 0.5740659199999953)

('Quick Sort :-', 0.15427498666666395)

('Quick With Median Sort :-', 0.14127872000000252)

...........................Final Output...............................

('Avg Time for Different Sorting Techniques for n=', 50000)

('Insertion Sort :-', 111.26918570666666)

('Merge Sort :-', 0.7814920533333236)

('Quick Sort :-', 0.2346056533333467)

('Quick With Median Sort :-', 0.2257088000000067)

**Comparison Table 1**:-

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Length** | **Insertion Sort** | **Merge Sort** | **Quick Sort** | **Quick Sort(Median)** |
| 500 | 0.026698338 | 0.00789652 | 0.0042301 | 0.003987878 |
| 1000 | 0.09692793 | 0.02272154 | 0.0098892 | 0.005051559 |
| 2000 | 0.264120348 | 0.02934748 | 0.0177036 | 0.009878558 |
| 4000 | 1.190095826 | 0.05107431 | 0.0213951 | 0.030490804 |
| 5000 | 1.882631835 | 0.10584133 | 0.0403078 | 0.056903824 |
| 10000 | 6.800517431 | 0.14449838 | 0.0544977 | 0.056903824 |
| 20000 | 31.66799623 | 0.30603124 | 0.1194684 | 0.127404365 |
| 30000 | 62.43589262 | 0.5135413 | 0.1890847 | 0.282311294 |
| 40000 | 127.7116383 | 0.70684828 | 0.2449298 | 0.282311294 |
| 50000 | 206.3577304 | 1.05067364 | 0.367198 | 0.372183874 |

**Graph 1:-**

**Comparison Table 2**:-

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Length** | **Insertion Sort** | **Merge Sort** | **Quick Sort** | **Quick Sort(Median)** |
| 500 | 0.018789547 | 0.008804693 | 0.00294656 | 0.002100907 |
| 1000 | 0.07864448 | 0.015762773 | 0.005467733 | 0.00299648 |
| 2000 | 0.2559488 | 0.048083627 | 0.01817728 | 0.009412693 |
| 4000 | 1.025489067 | 0.050674773 | 0.013908907 | 0.022728533 |
| 5000 | 1.42763392 | 0.107757653 | 0.029408 | 0.027313067 |
| 10000 | 4.958820267 | 0.155646293 | 0.040307781 | 0.073736107 |
| 20000 | 20.42371115 | 0.4408512 | 0.099501227 | 0.072462507 |
| 30000 | 44.10685525 | 0.426875307 | 0.114717867 | 0.154274987 |
| 40000 | 67.41122048 | 0.57406592 | 0.154274987 | 0.14127872 |
| 50000 | 111.2691857 | 0.781492053 | 0.234605653 | 0.2257088 |

**Graph 2:-**

**Conclusion:-**

1. From the above graph it can be inferred that insertion sort is very much expensive for the large amount of inputs. Hence it is not advisable to use insertion sort for large amount of data.
2. Insertion sort is almost linear for very small amount of data and hence it is useful in case of modified quick search.
3. Although merge sort and quick sort have same time complexity – O(nlogn), quick sort performs well for large amount of input data.
4. Performance of the quick sort highly depends on the arrangement of an array (sorted/unsorted). Quick sort’s performance is of O(n^2) in the following cases:
5. If input array is sorted and selected pivot is either minimum or maximum
6. If input array is reversely sorted and selected pivot is either minimum or maximum
7. If input array has all the elements with same value
8. In order to avoid the limitation, quick sort can be used with pivot equal to median of three approach and with the use of insertion sort for the input data in an array less than 10.
9. One of the important advantage of merge sort over quick sort can be stated as, it takes run time in order of O(nlogn) regardless of the order(sorted/unsorted) of the elements in an array