Heap Sort

Dharmik Kurlawala CSU ID - 2886995

Introduction:

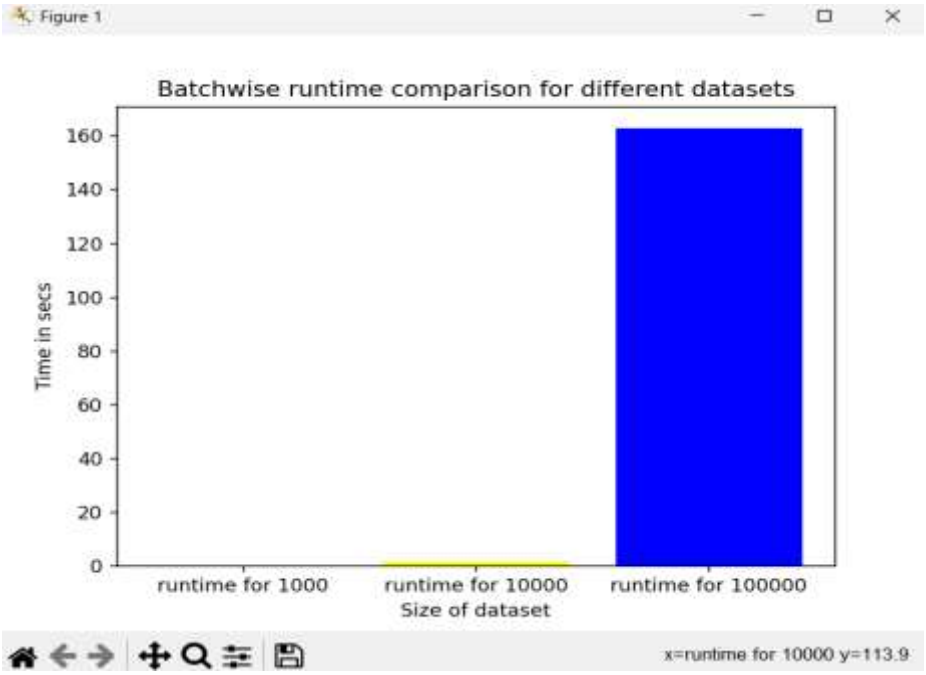
- Aim of this assignment was to implement a heap sorting algorithm to facilitate the selection of movies with the highest ratings for movie nights. This involved constructing a max heap from a given array of movie ratings, implementing heap operations such as insert and remove, and sorting the movies based on their ratings. Additionally, the assignment required testing the implemented heap sorting methods with different dataset sizes and comparing their performance with built-in sorting methods in Python.

Results and Analysis:

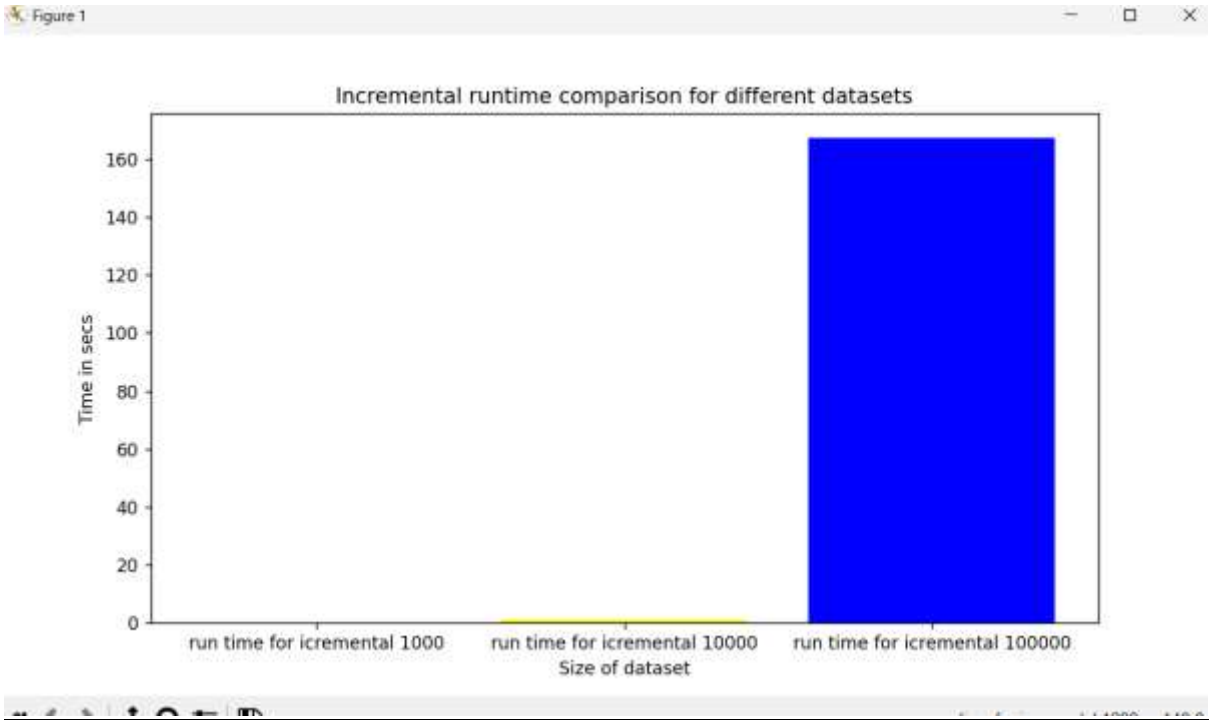
- Observed that the runtime of the heap sorting methods increased with the size of the dataset, as expected. Batch sorting showed better performance compared to incremental sorting due to its direct construction of the max heap from the entire dataset.
- Comparison with Built-in Sorting: Compared the performance of the implemented heap sorting methods with built-in sorting methods in Python that is sorted() function. Found that the built-in sorting methods outperformed the implemented heap sorting methods, especially for larger dataset sizes.
- The sorting for inbuilt sorted() function was done with in a seconds where as compared to batch and incremental sorting.

Following is the visual representation of dataset comparison:

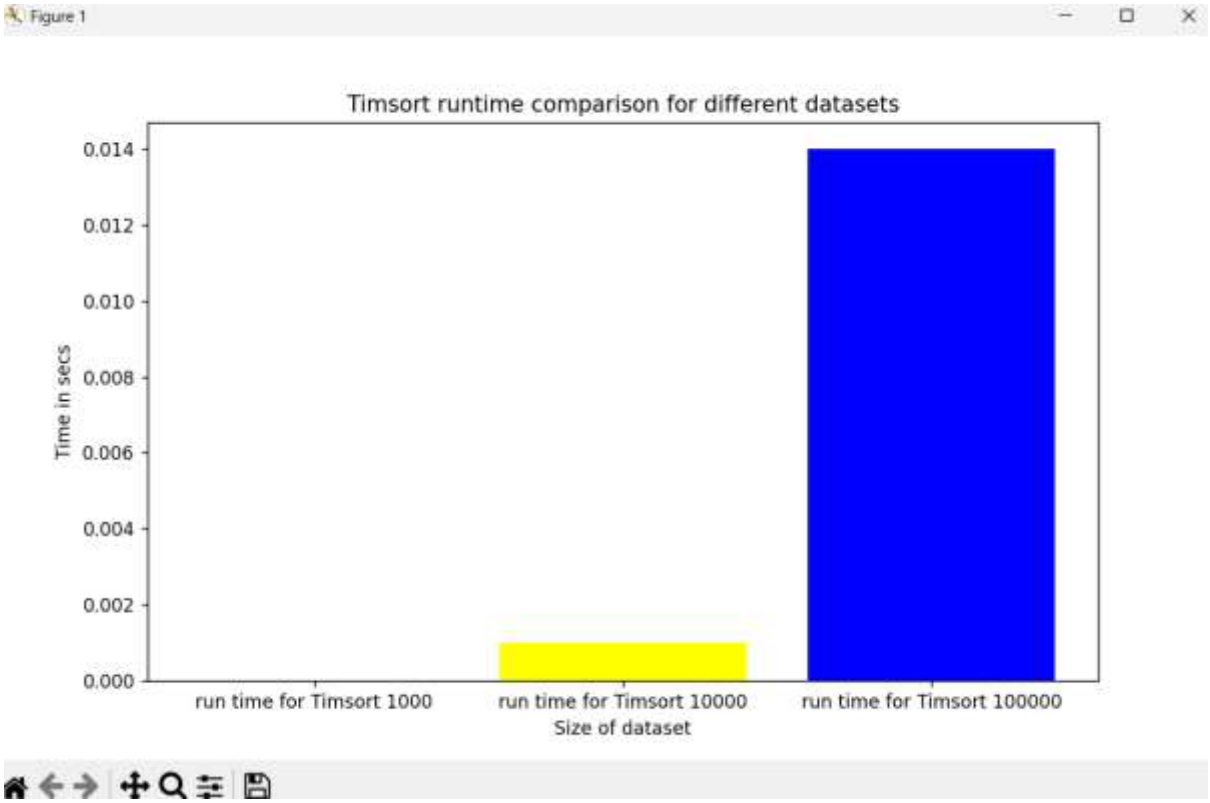
Batchwise comparison for 1000,10000 and 100000 data sets:



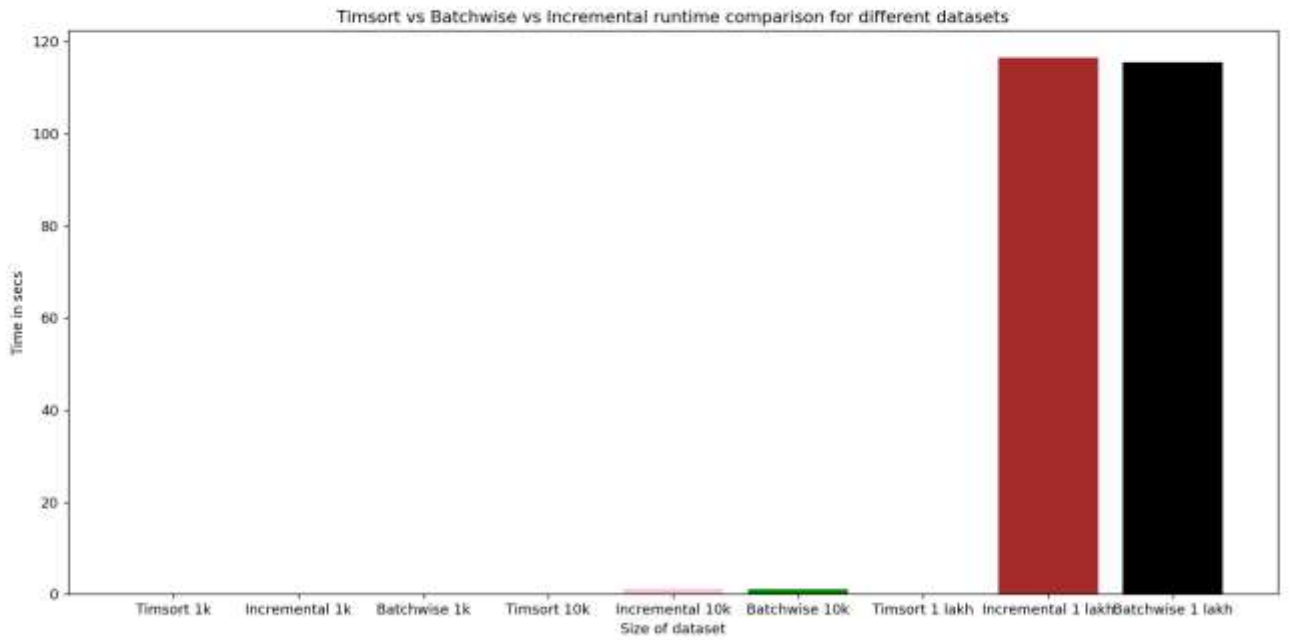
- Incremental comparison for 1000,10000 and 100000 data sets:



- Inbuilt sort(Timsort) comparison for 1000,10000 and 100000 data sets:



- Overall comparison of all the sorting techniques with dataset of 1000, 10000 and 100000:



Conclusion:

- In conclusion, the implementation of heap sorting methods proved effective in sorting movies based on their ratings. However, as the data set increases the sorting time increases. The performance of these methods was not as efficient as built-in sorting methods in Python. Nevertheless, the heap sorting methods provided a useful alternative for sorting tasks, especially for smaller dataset sizes.