Name: Aashir Shukla
SID: 917460238

Algorithm Analysis

Let the number of Cities in the dataset = N
Let the number of Safety indices in the dataset = M

Big O Complexity of the Binary Search Approach

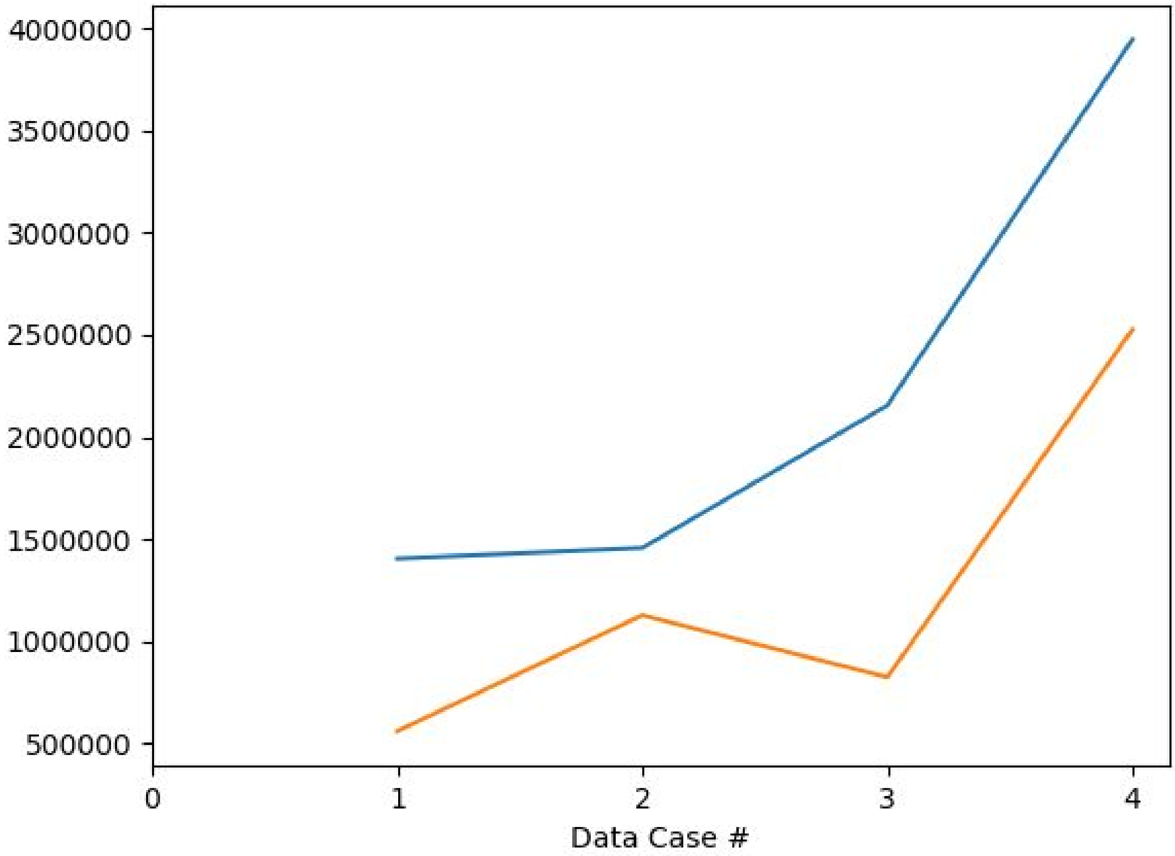
Sorting the cities using the std::sort algorithm will take $O(N * log(N))$ time.
So, the overall time complexity will be $O(N * log(N) + M * log(N))$. If $M \geq N$, the overall complexity would be $O(M * log(N))$. On the other hand, if $N \geq M$, the overall complexity would be $O(N * log(N))$. This is because it takes $log(N)$ time for each binary search, and that will be done M times - for each city.

Big O Complexity of the Linear Search Approach

The Big O Complexity of the Linear Search Approach would be $O(N * M)$. For each safety index, we will go through each of the cities.

Analysis of CPU Runtime

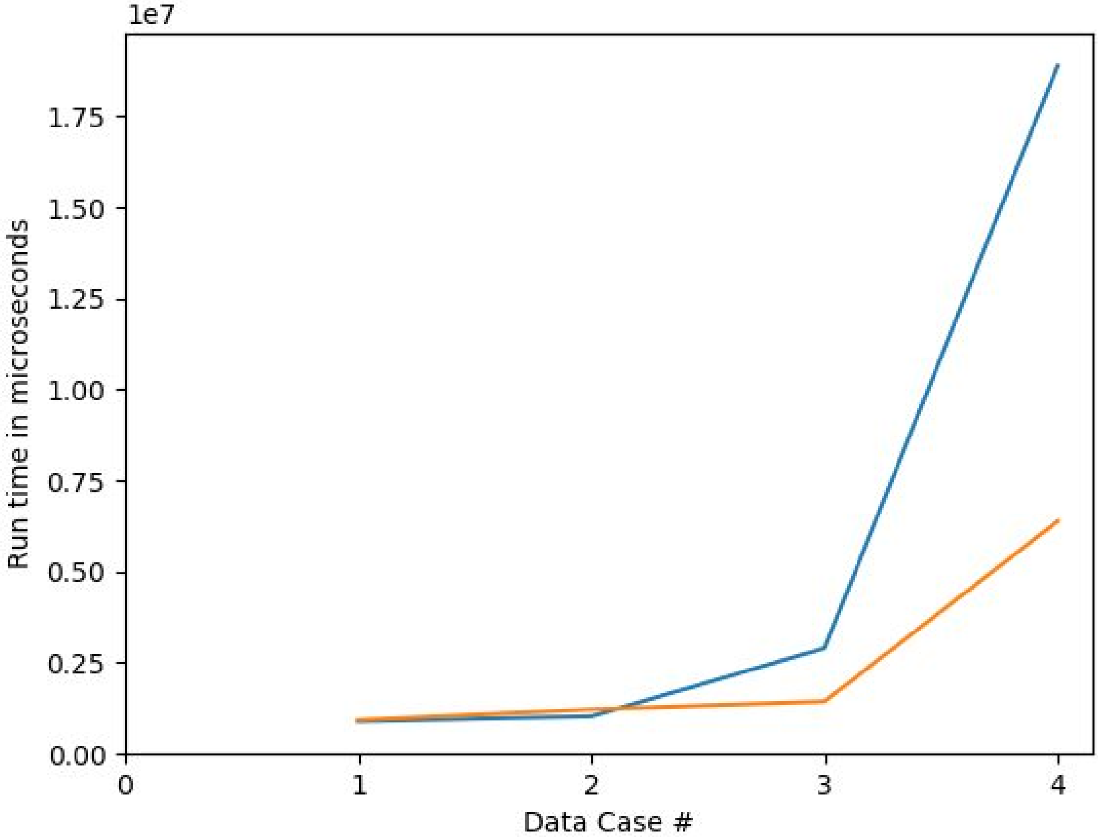
It is clear that the binary search algorithm and the linear search algorithm will increase at relatively the same rate as M increases but N remains constant. This result is displayed in the following graph:



Here, the y axis represents the time taken by the algorithm. For the data cases, N was taken constant at a 100, and M was increased from $1e4$ to $1e7$.

The yellow line represents the time taken with binary search, and the blue line represents the same with linear search. There isn't much of a difference between the two.

However, when N is varied and M is taken constant, binary search increases very slowly. This is because binary search only has a $log(N)$ factor, while runtime of linear search contains an N factor. We can see the sharp increase in linear search runtime below:



For the Data Cases, M was kept constant at 100. N was varied from $1e4$ to $1e7$. At $1e7$, the linear search algorithm took almost 10 seconds.

Result

Binary Search seems to be the best choice of algorithm in most data cases. At relatively small values of N, regardless of the value of M, both binary search and linear search will perform well, with binary search having a slight edge. As N increases, however, linear search will start to fail.