Richard Qin

ECS36C

Professor Porquet

April 4 2020

HW1 city_safety Measurement Report

Purpose:

Using the program city_safety, I aim to show that linear search is more efficient in terms of run-time for smaller data sets, while binary search is generally more run-time efficient for common use and for larger data sets. In order to test this claim, I used the following benchmarking method:

Methodology:

The program was given three data sets to test, starting with 10 cities and 100 safety indices. Each subsequent data set is 10 times larger than the previous set, and every data set was tested three times to ensure more accurate results. This method of testing was applied to both the linear and binary search methods, and the results were averaged from each data set to provide six run-times which I will compare and use to derive a conclusion from.

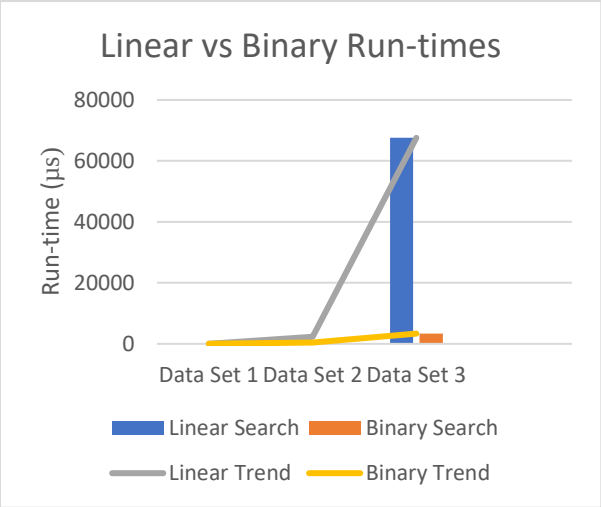
Data:

Linear Search: data size vs. run-time

	Runtime	Trial 1 (µs)	Trial 2 (µs)	Trial 3 (µs)	Average (µs)
Data Set	X	X	X	X	X
10 city, 100 safety	X	16.643	26.057	16.31	19.670
100 city, 1000 safety	X	2395.91	2548.29	1939.19	2294.463
1000 city, 10000 safety	X	64092.2	69856.7	68779	67575.967

Binary Search: data size vs. run-time

	Runtime	Trial 1 (µs)	Trial 2 (µs)	Trial 3 (µs)	Average (µs)
Data Set	X	X	X	X	X
10 city, 100 safety	X	49.259	34.134	21.801	35.065
100 city, 1000 safety	X	309.176	391.813	487.063	396.017
1000 city, 10000 safety	X	4588.64	2586.41	2911.34	3362.130



Taking the six average run-times, I graphed the data to more clearly visualize the trend which each search method displays as the input data set grows larger. As the data and the graph indicate, the binary search method yields only a relatively small increase in run-time even as the data set grows significantly. On the other hand, the run-times of the linear search method increase exponentially despite being supplied the same data sets as the binary method trials. Continuing this trend, I can conclude that the larger the data set, the more efficient that a binary search algorithm will be as compared to a linear search algorithm. However, while this is true for large data sets, small data sets lean in favor of linear searching in terms of run-time, as indicated in the data tables. For the smallest data set with 10 cities and 100 safety indices, the average run-time for linear search is 19.67µs as compared to binary search's run-time of 35.065µs. While the difference between the two run-times is relatively insignificant in the context of common use, it does show even if by a little that linear search is more run-time efficient when analyzing small data sets.

Conclusion:

While linear search is more run-time efficient for very small data sets, binary search is generally more efficient both in practical applications and for large data sets. This trend becomes more apparent as the data set grows larger, with the difference in run-times between the search algorithms growing exponentially.