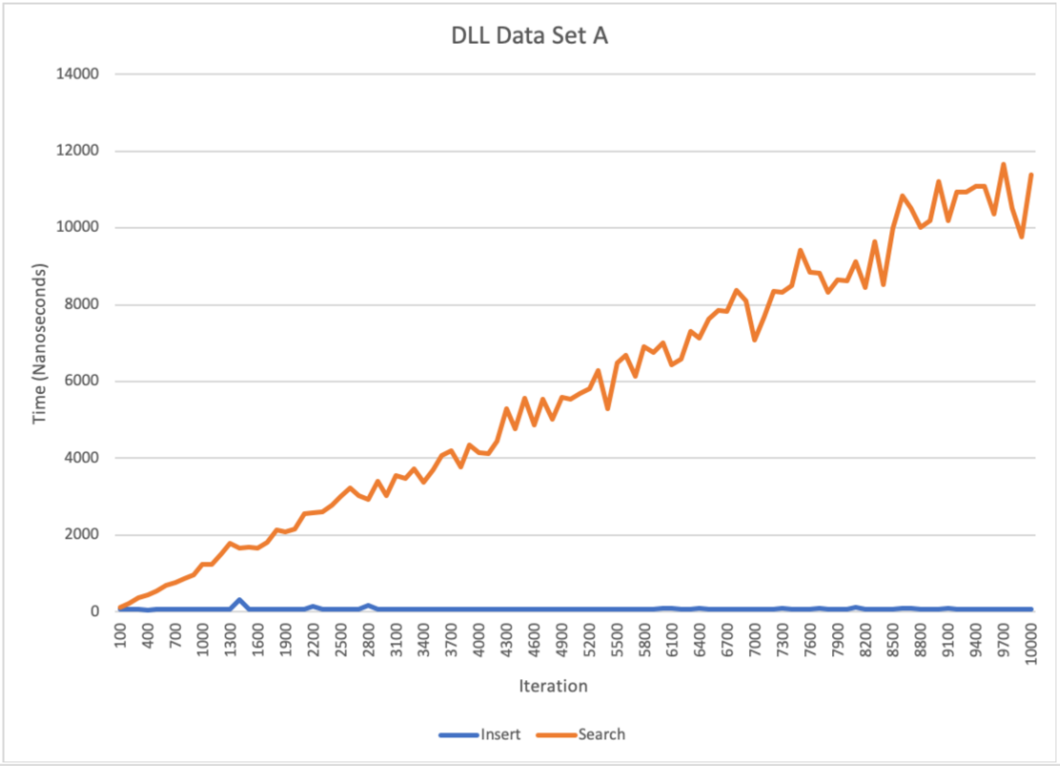
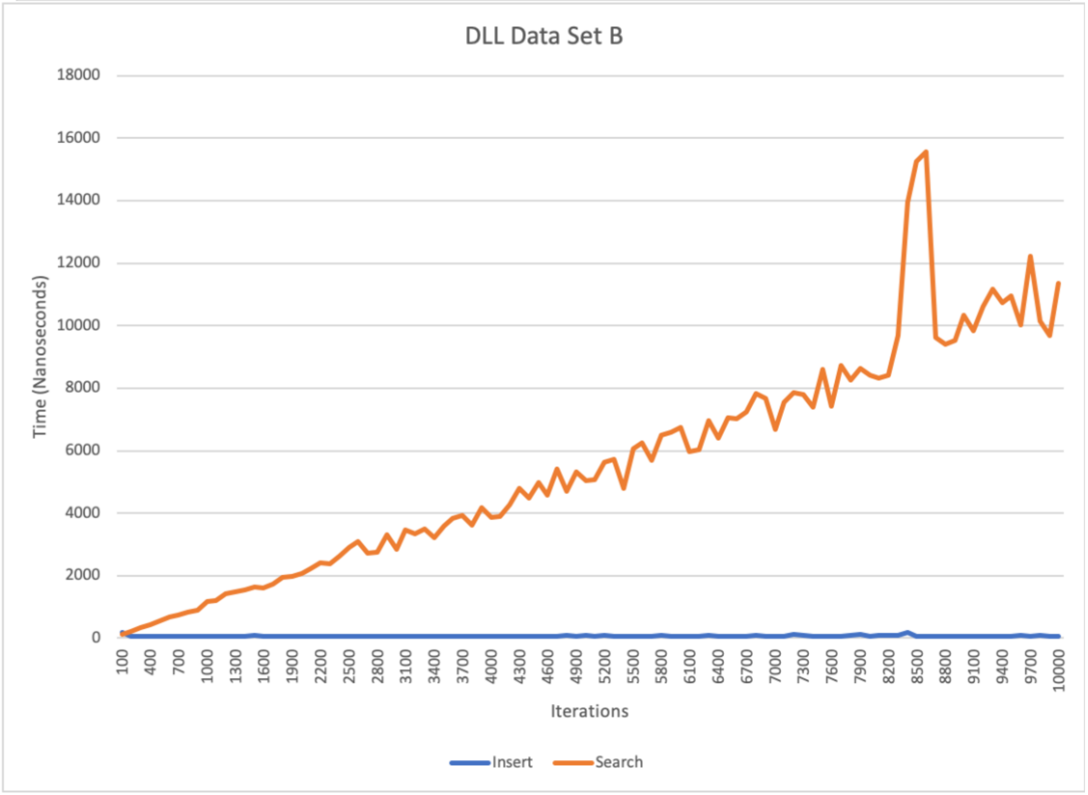
CSCI 2270: Data Structures
Final Project Writeup
Case Chrisbacher

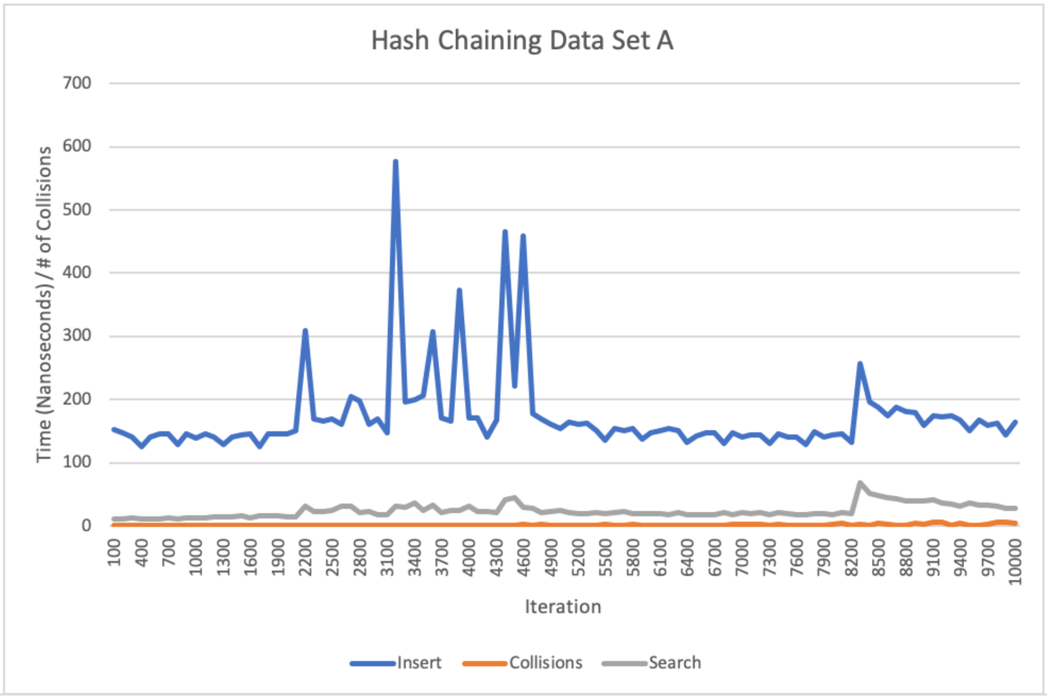
For the most efficient storage and retrieval of the patient ID's, my recommendation is hash chaining (figures 3 and 4). Chaining was extremely efficient at searching for the ID's, but was slightly slower at inserting, likely because of its need to create brand new nodes for the new data. Although some other structures could beat the insert time of chaining, the searching is more vital because the hospital will likely only insert once, and then search more frequently. Quadratic was consistently good throughout as well. All of the different structures struggled from spikes in time here and there, but the quadratic hashing graphs (figures 7 and 8) stayed fairly consistently low throughout with only a few spikes. Quadratic hashing would fit better if the hospital planned to input ID's just as frequently as they search. Linear hashing was a close third to quadratic, but it suffered from clustering in data set B, which really hurt the run time. Finally, doubly linked lists have a very good insert time because of the insert at the head, but an atrocious search time because of the need to peruse through the whole list (figures 1 and 2). The best insert and search times (figures 9 and 10) ended up being insert from linear hashing from data set A and search from chaining. The final graph really shows how efficient the search using chaining is compared to the doubly linked list.

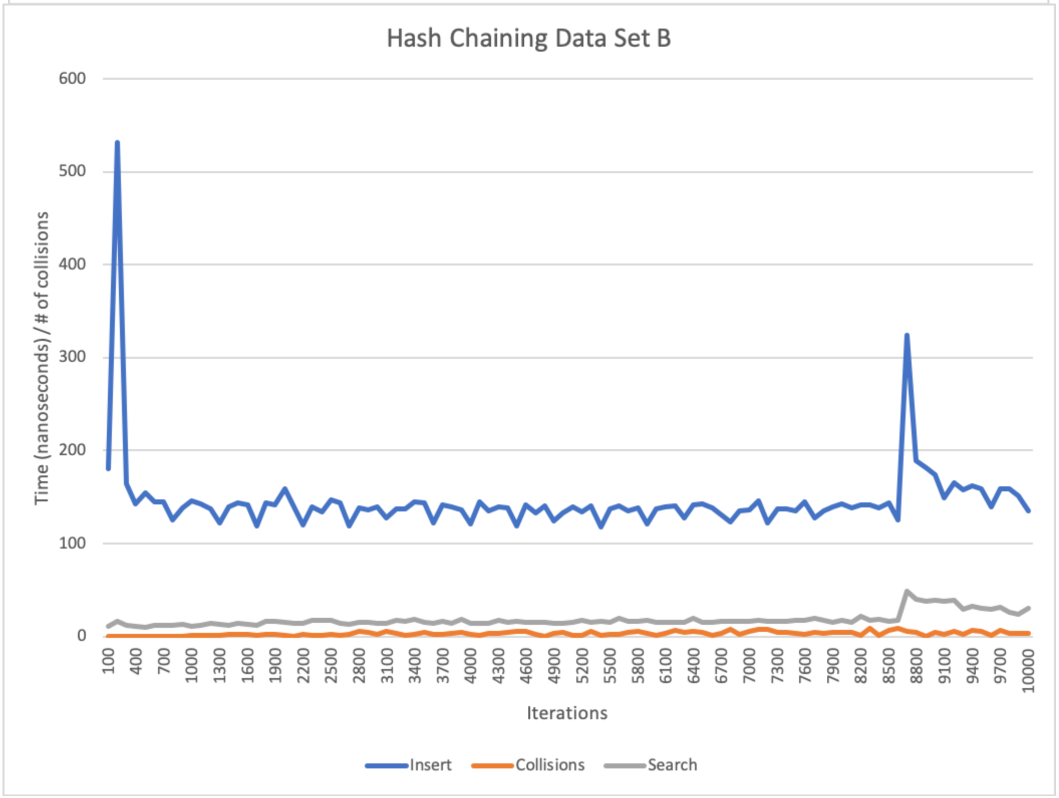
FIGURES BELOW



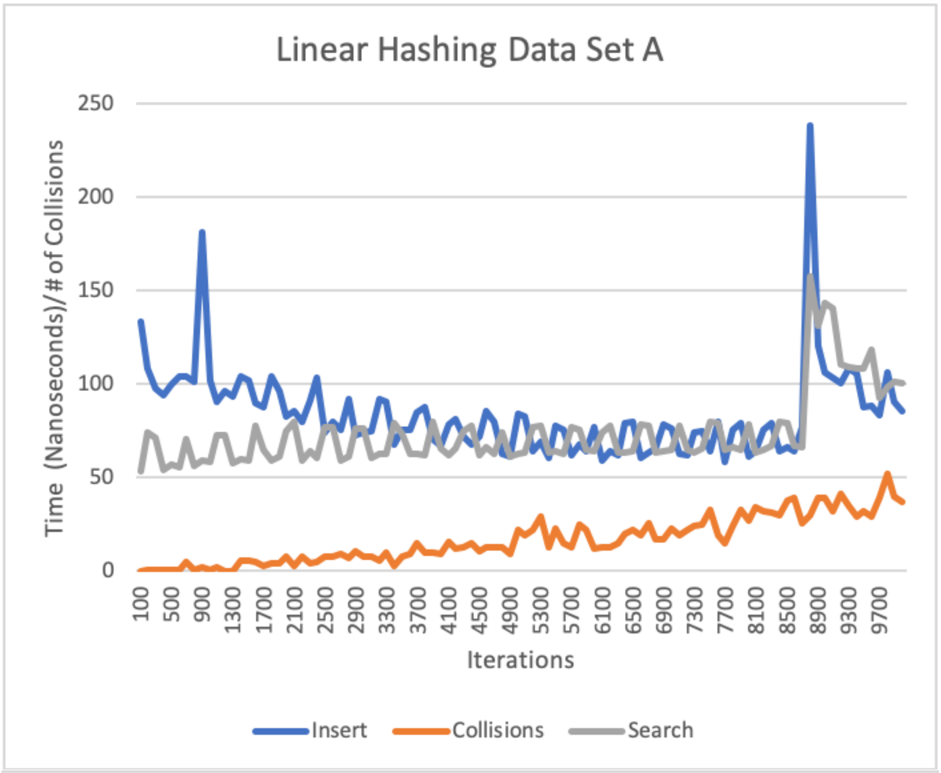


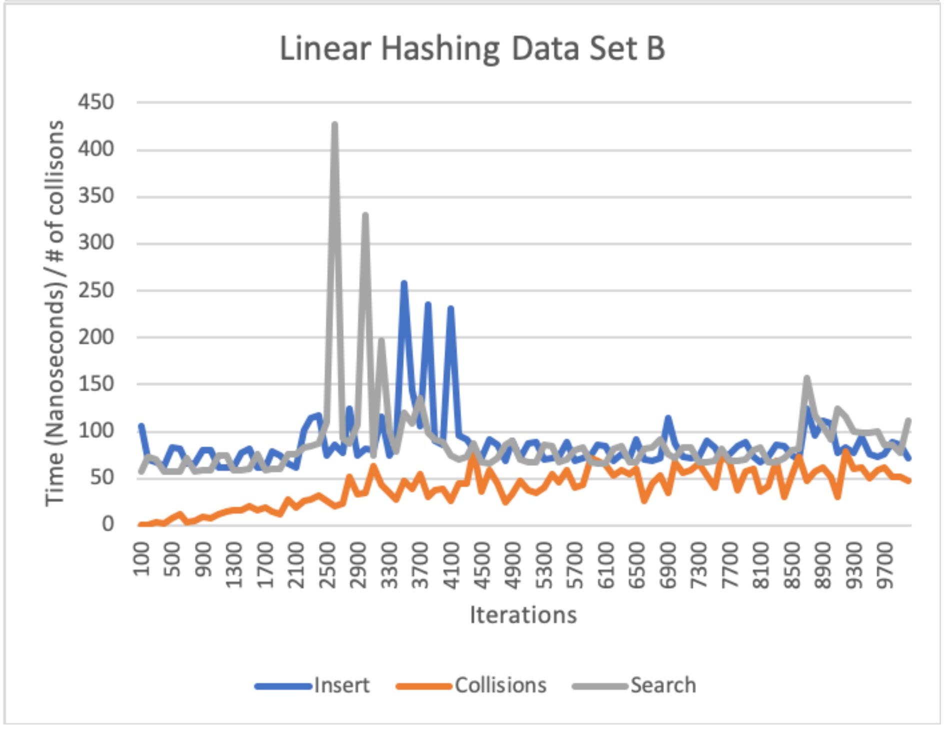
Figures 1 and 2: Average Insert and Search times for Doubly linked lists from set A (1) and B (2)



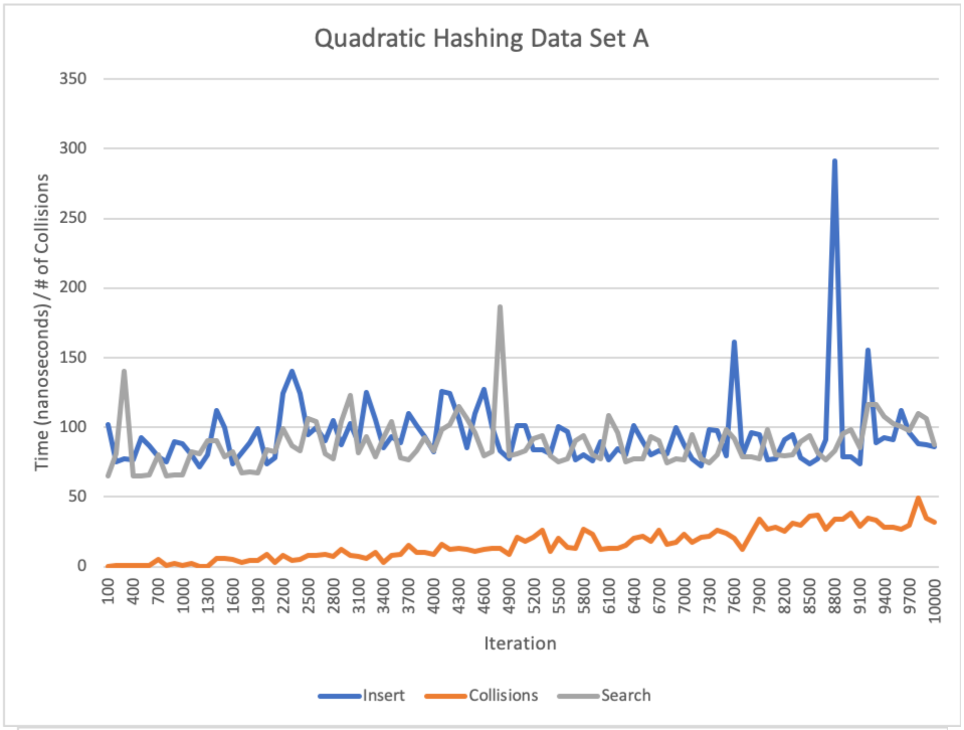


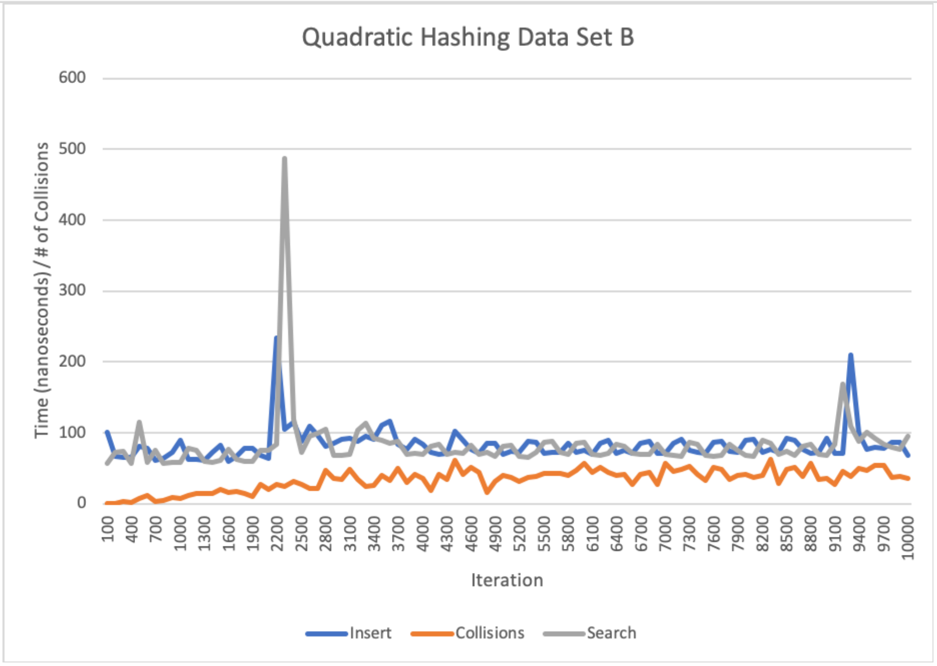
Figures 3 and 4: Average Insert and Search times for Hash Chaining from set A (1) and B (2)



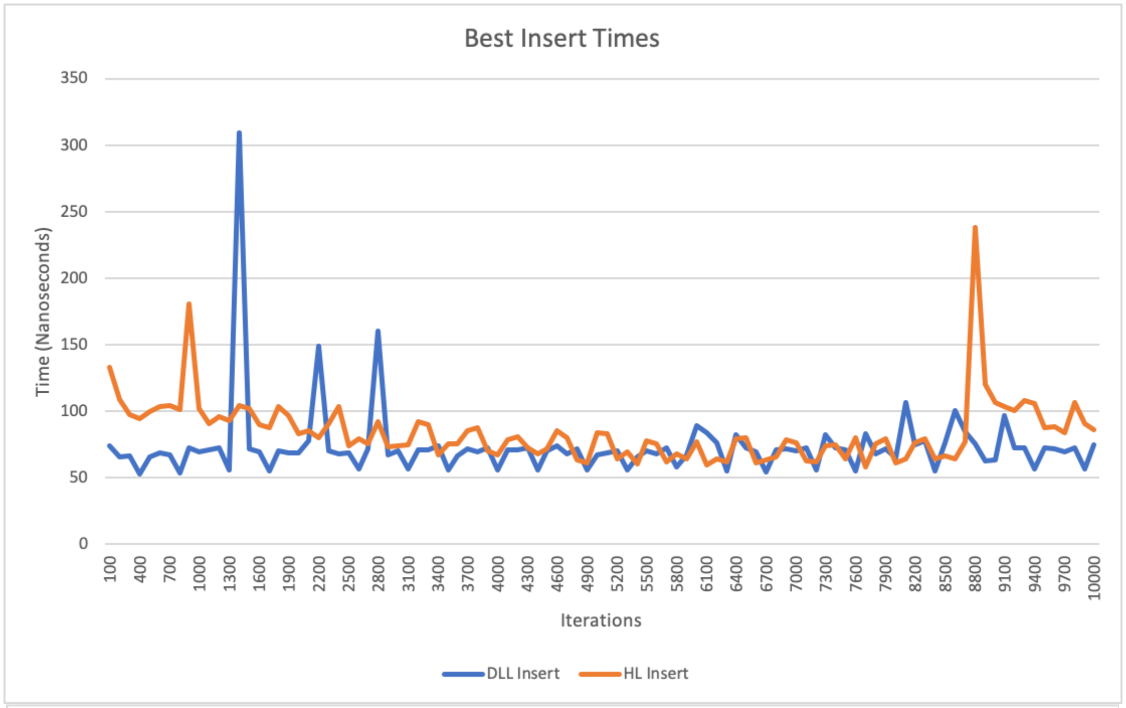


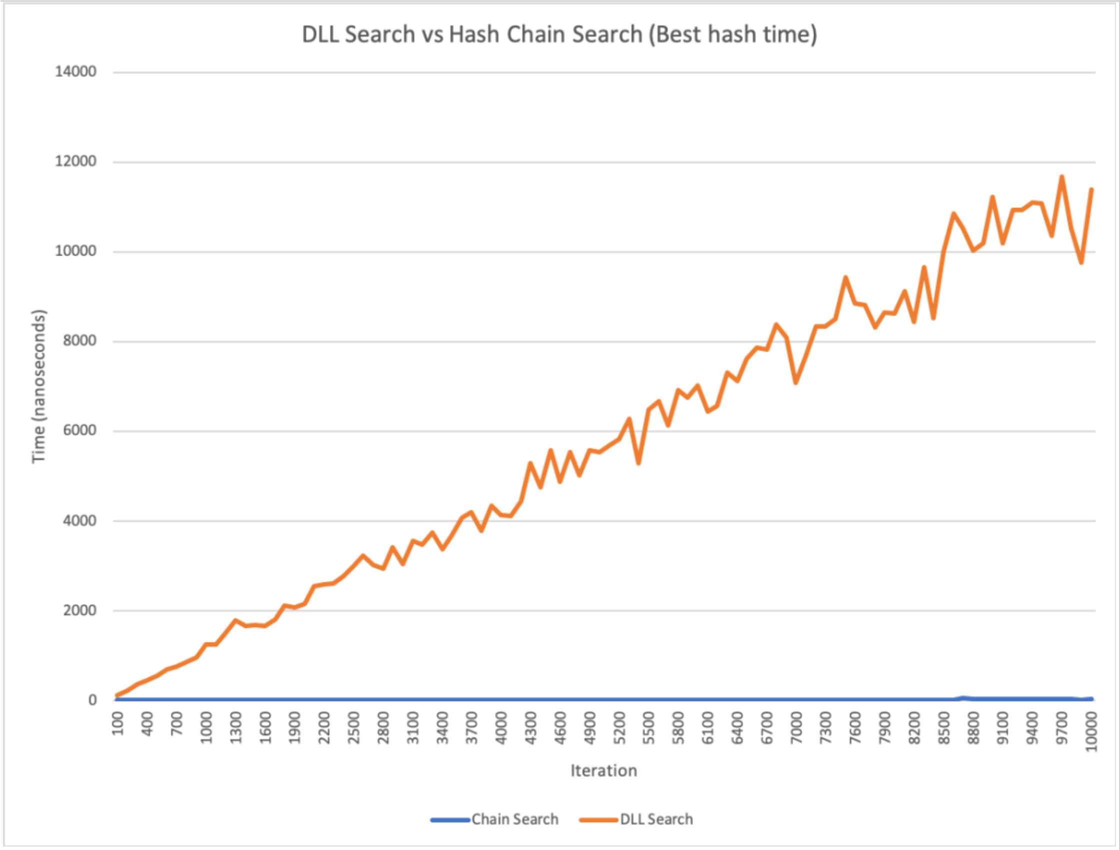
Figures 5 and 6: Average Insert and Search times for Linear Hashing from set A (1) and B (2)





Figures 7 and 8: Average Insert and Search times for Quadratic Hashing from set A (1) and B (2)





Figures 9 and 10: These graphs compare the average search times between the doubly linked list and the best average times for the graphs. The top graph uses the insert from linear hashing and the bottom uses the search time from chaining