Performance Evaluation and Submission

As stated in the original writeup, the goal of the project is to perform an analysis on a number of different hash table algorithms. Below is a detailed list of all the different experiments that should be performed and described in the final deliverable.

- For Dataset A:
    - For hash function h(x) :
        - With Linear Probing:
            - Insert numbers till load factor becomes 0.1:
                - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert numbers till load factor becomes 0.2:
                - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert numbers till load factor becomes 0.5:
                - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert numbers till load factor becomes 0.7:
                - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert numbers till load factor becomes 0.9:
                - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
                - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.

        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 1:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
- With Chaining using a Linked List:
    - Insert numbers till load factor becomes 0.1:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.2:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.5:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.7:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.9:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.

        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 1:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
- With Chaining using a BST:
    - Insert numbers till load factor becomes 0.1:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.2:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.5:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.7:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 0.9:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.

            - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert numbers till load factor becomes 1:
            - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
- Repeat the above process for all collision resolution mechanisms with hash function h'(x)
- Using hash functions h(x) and h'(x) simultaneously:
    - Cuckoo Hashing:
        - Insert numbers till load factor becomes 0.1:
            - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert numbers till load factor becomes 0.2:
            - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert numbers till load factor becomes 0.5:
            - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert numbers till load factor becomes 0.7:
            - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
            - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert numbers till load factor becomes 0.9:

        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
    - Insert numbers till load factor becomes 1:
        - Delete 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Insert 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
        - Lookup 100 numbers. Measure time for these 100 operations. Find mean and standard deviation.
- Repeat the above process for Dataset B

Notes:

- From the above procedure, you will get a series of numbers for each collision resolution mechanism per hash function per dataset. You need to plot these vs their load factors.
- You can use a different sequence to time your operations as well. In the above illustration, we are timing delete, then insert and then lookup. You can use any sequence.
- All graphs should share a common scale.
- You might want to combine graphs since there are a lot of them. For example, for a particular dataset with a particular hash function, you can plot the times for all collision resolution mechanisms on a single graph. For example, you might have one graph for dataset A with hash function h(x) with 4 lines (linear probing, chaining + LL, chaining + BST, Cuckoo). Obviously, since cuckoo hashing uses both hash functions simultaneously, you will have the same line for cuckoo hashing on the graphs for both hash functions (h and h').
- You should plot the datasets as well so that you can visualize what the distribution of the data looks like. The behaviour of the hash table will change depending on the distribution of the data so this will help you in explaining your results. A simple way to do this is to use a histogram.
- You should explain why you think your implementation performs the way it does given a dataset.
- Code submission - it is up to you how you organize your code. You can have separate programs for the various collision resolutions, or have everything in a single program. You do not need a separate program showing each experiment you ran for the various load factors, as this would be identical code with only a single changed parameter. Whichever approach you choose, make sure your code is well organized, commented, indented, and legible.