Ahmet Çalışkan 31065 HW6 Report

In that homework, I use three class and implement it: BloomFilter, HyperLogLog, KMinVal.

Hash Functions are implemented by TA’s. I used these functions for classes. For example, I used ax+b hash functions for bloomFilter for each member of a and b. I stored a and b values in a array, I used all the numbers that given array for executing hash ax+b function. I used murmurhash as default, therefore I do not give any parameters for seed (it is already 0 for default.). Lastly I use a1 and b1 for ax+b hash function in KMinVal. I tried some of combinations but I was the best efficient situation as much as I try. Sum up, for some situations I use every a and b values, sometimes only a1 and b1, sometimes none of them (for murmurhash).

My Outputs:   
My outputs quite similar for HyperLogLog functions. There are some deviations in output for example if expected output is 1.8256, my output was roughly 2.12245 for HyperLogLog. For KMinVal there are some differences between my output and expected outputs. My MurmurHash outputs for KMinVal are exactly same with the expected. On the other hand My ax+b hash function results are very different from expected results, especially in low size arrays such as 2,4,8. But my outputs for high size arrays such as 8192, 4096 are exactly same with the expected. Lastly, my BloomFilter outputs are very similar with expected outputs. There are some little differences between my FP such as 22 and expected FP such as 23 or 24. Other than they estimate properly.