**GROUP NUMBER:** fill in your group number

**GROUP MEMBERS:** fill in your group members

**TEST 1:** The goal is to assess the accuracy of the count-sketch estimations as the number of distinct items (regulated by the interval [left,right]) varies. The values D, W, and K are fixed. You must fill in the following table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ACCURACY WITH RESPECT TO NUMBER OF DISTINCT ITEMS, USING D=9, W=30, K=10**  Use 4 decimal digits for floating points | | | | |
| **[left,right]** | **Number of received distinct items in [left,right]** | **Average relative error for items with top-K frequencies** | **True normalized F2** | **Approximate normalized F2** |
| [1,15000] | 15000 | 0.0323 | 0.3465 | 0.3455 |
| [1,10000] | 10000 | 0.0092 | 0.4453 | 0.4451 |
| [1,5000] | 5000 | 0.0212 | 0.3729 | 0.3712 |
| [1,1000] | 1000 | 0.0062 | 0.5639 | 0.5638 |

S

**TEST 2:** The goal is to assess the accuracy of the count-sketch estimations as the number W of columns of the sketch varies. The values D, K and the interval are fixed. Repeat each experiment 3 times. You must fill in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **ACCURACY WITH RESPECT TO NUMBER OF COLUMNS W, USING D=9, K=30, [left,right]=[1,10000]**  Use 4 decimal digits for floating points and report averages over 3 runs | | | |
| **W** | **Average relative error for items with top-K frequencies. RUN 1** | **Average relative error for items with top-K frequencies. RUN 2** | **Average relative error for items with top-K frequencies. RUN 3** |
| 100 | 0.0305 | 0.0300 | 0.0145 |
| 50 | 0.0834 | 0.0357 | 0.2824 |
| 20 | 1.1876 | 0.6913 | 2.2184 |
| 15 | 3.0718 | 0.8666 | 0.4274 |

**TEST 3:** The goal is to assess the accuracy of the count-sketch estimations as K varies. The values D, W and the interval are fixed. Repeat each experiment 3 times. You must fill in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **ACCURACY WITH RESPECT TO K, USING D=9, W=100, [left,right]=[1,10000]**  Use 4 decimal digits for floating points and report averages over 3 runs | | | |
| **K** | **Average relative error for items with top-K frequencies. RUN 1** | **Average relative error for items with top-K frequencies. RUN 2** | **Average relative error for items with top-K frequencies. RUN 3** |
| 10 | 0.0024 | 0.0057 | 0.0035 |
| 50 | 0.1145 | 0.0572 | 0.2115 |
| 100 | 0.2820 | 0.8279 | 0.3608 |
| 200 | 32.2223 | 1.2258 | 1.1745 |