**GROUP NUMBER:** 047

**GROUP MEMBERS:** Alessandro Lucchiari, Francesco Pio Monaco, Lorenzo Riccò

**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 | 0.6682 | 0.3182 | 0.3309 |
| [1,10000] | 10000 | 0.3500 | 0.3278 | 0.3376 |
| [1,5000] | 5000 | 0.2500 | 0.4645 | 0.4689 |
| [1,1000] | 1000 | 0.0928 | 0.6066 | 0.6078 |

**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 |  |  |  |
| 50 |  |  |  |
| 20 |  |  |  |
| 15 |  |  |  |

**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 | 0.0883 | 0.1790 | 0.1314 |
| 50 | 3.8697 | 3.3423 | 5.5775 |
| 100 | 10.2576 | 8.0386 | 9.4603 |
| 200 | 27.5391 | 51.6479 | 52.2960 |