**GROUP NUMBER: 046**

**GROUP MEMBERS:** Lorenzon Nicola, Manfè Alessandro, Mickel Mauro

**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 | 0.0062 | 0.4296 | 0.4298 |
| [1,10000] | 10000 | 0.0106 | 0.5020 | 0.5020 |
| [1,5000] | 5000 | 0.0185 | 0.4292 | 0.4289 |
| [1,1000] | 1000 | 0.0179 | 0.4731 | 0.4725 |

**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 | 0.0191 | 0.0462 | 0.0184 |
| 50 | 0.0567 | 0.0643 | 0.0926 |
| 20 | 0.4971 | 0.6434 | 0.7451 |
| 15 | 0.9236 | 0.8726 | 0.6437 |

**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 | 0.0041 | 0.0032 | 0.0048 |
| 50 | 0.0765 | 0.0836 | 0.1241 |
| 100 | 0.6545 | 0.4091 | 0.2503 |
| 200 | 1.4479 | 1.5005 | 1.2548 |