**GROUP NUMBER:** 058

**GROUP MEMBERS:** Mehran Farajinegarestan, Omid Airom, Nahid Jahanianarange

**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 | 0.4120 | 0.3353 | 0.3355 |
| [1,10000] | 10000 | 0.4087 | 0.3397 | 0.3400 |
| [1,5000] | 5000 | 0.4599 | 0.4451 | 0.4443 |
| [1,1000] | 1000 | 0.5177 | 0.4975 | 0.4978 |

**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 | 0.6106 | 0.6395 | 0.5610 |
| 50 | 0.7007 | 0.5230 | 0.6369 |
| 20 | 1.1207 | 1.4607 | 2.4321 |
| 15 | 2.1875 | 1.7327 | 1.1171 |

**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 | 0.3948 | 0.6080 | 0.4590 |
| 50 | 0.5641 | 0.6262 | 1.1454 |
| 100 | 1.5928 | 2.0548 | 1.2157 |
| 200 | 2.8404 | 1.8122 | 3.0722 |