**GROUP NUMBER: 035**

**GROUP MEMBERS: AKBARI MOAFI MOHAMMADHOSSEIN, MIRZAZADEH BAHADOR, RAHGOOY ZAHRA**

**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] | 1335416 | 1.2432 | 0.4431 | 0.4526 |
| [1,10000] | 1167735 | 0.4391 | 0.4034 | 0.4105 |
| [1,5000] | 1187816 | 0.0420 | 0.4334 | 0.4343 |
| [1,1000] | 889438 | 0.0245 | 0.6666 | 0.6667 |

**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 | 0.6696 | 1.3124 | 0.1271 |
| 50 | 4.9618 | 3.0189 | 0.0580 |
| 20 | 1.3652 | 10.021 | 7.6391 |
| 15 | 0.5202 | 7.4653 | 0.8274 |

**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 | 0.0066 | 0.1858 | 0.1399 |
| 50 | 0.9177 | 8.0638 | 0.1571 |
| 100 | 3.5902 | 19.9039 | 8.8321 |
| 200 | 10.9819 | 10.4781 | 1.9129 |