**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 | 0.0179 | 0.3876 | 0.3876 |
| [1,10000] | 10000 | 0.0275 | 0.3481 | 0.3476 |
| [1,5000] | 5000 | 0.0233 | 0.5861 | 0.5861 |
| [1,1000] | 1000 | 0.0210 | 0.5954 | 0.5954 |

**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 | 0.0601 | 0.0767 | 0.0654 |
| 50 | 0.0623 | 0.0842 | 0.0923 |
| 20 | 0.7185 | 40.9242 | 1.1808 |
| 15 | 33.1314 | 1.4176 | 9.0903 |

**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 | 0.0246 | 0.0037 | 0.0149 |
| 50 | 0.0675 | 0.0737 | 0.0847 |
| 100 | 28.1091 | 10.4413 | 14.0851 |
| 200 | 130.7262 | 129.2420 | 42.6413 |