**GROUP NUMBER:** 078

**GROUP MEMBERS:**  Siavash Azarhazin, Daniel Carlesso, Carmine Graniello

**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 | 0.9777 | 0.2940 | 0.3181 |
| [1,10000] | 10000 | 0.8606 | 0.3870 | 0.4040 |
| [1,5000] | 5000 | 0.6038 | 0.4909 | 0.5005 |
| [1,1000] | 1000 | 0.1762 | 0.5080 | 0.5126 |

**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 | 8.1224 | 15.8166 | 10.0405 |
| 50 | 11.9594 | 17.5077 | 21.2059 |
| 20 | 173.69 | 80.0573 | 248.4546 |
| 15 | 45.4852 | 66.1060 | 115.7460 |

**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 | 0.2338 | 0.3144 | 0.2780 |
| 50 | 24.2809 | 38.7437 | 20.2886 |
| 100 | 30.2727 | 122.6711 | 62.3959 |
| 200 | 784.2925 | 375.7488 | 413.1189 |