**GROUP NUMBER: group 061**

**GROUP MEMBERS: CANEL ALESSANDRO, GIACHINTA GABRIELLA, ZOLGHADR SHARARE**

**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 THE 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]** | **The average relative error for items with top-K frequencies** | **True normalized F2** | **Approximate normalized F2** |
| [1,15000] | 15000 | 0.4586 | 0.2889 | 0.3003 |
| [1,10000] | 10000 | 0.3775 | 0.3457 | 0.3552 |
| [1,5000] | 5000 | 0.1950 | 0.3775 | 0.3829 |
| [1,1000] | 1000 | 0.1200 | 0.6503 | 0.6515 |

**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** | **The average relative error for items with top-K frequencies. RUN 1** | **The average relative error for items with top-K frequencies. RUN 2** | **The average relative error for items with top-K frequencies. RUN 3** |
| 100 | 2.7971 | 0.7666 | 1.9246 |
| 50 | 4.9171 | 2.9678 | 4.4318 |
| 20 | 16.0224 | 17.2898 | 11.9175 |
| 15 | 1.3732 | 1.6778 | 1.0000 |

**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** | **The average relative error for items with top-K frequencies. RUN 1** | **The average relative error for items with top-K frequencies. RUN 2** | **The average relative error for items with top-K frequencies. RUN 3** |
| 10 | 0.2060 | 0.4408 | 0.3340 |
| 50 | 5.2038 | 2.1466 | 7.5037 |
| 100 | 11.37907 | 42.1283 | 16.5438 |
| 200 | 41.9468 | 65.9804 | 59.2655 |