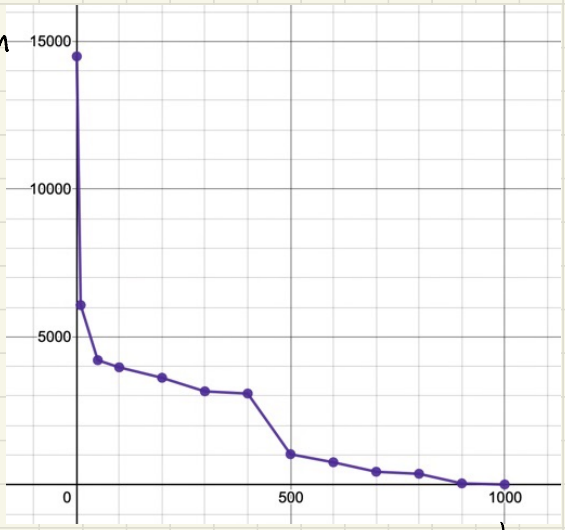
1. minimum: 1

   maximum: n

2. distortion



numbers of clusters

numbers of clusters

numbers of clusters	distortion
1	14492.3280552
10	6069.469572
50	4206.799990
100	3964.969487
200	3610.835770
300	3151.958677
400	3077.380405
500	1023.573223
600	750.568263
700	430.226641
800	360.810977
900	35.153152
1000	0

3. The lowest possible distortion is 0, it happens when numbers of clusters = 1000 (n, maximum).

The reason for it is: the number of data is equal to the number of clusters, so every data belongs to one cluster center which is itself. Thus, total distortion is equal to 0.

4. We could use Gap statistic method.

1. Cluster the observed data, varying the number of clusters from k = 1, ..., k_{max}, and compute the corresponding total within intra-cluster variation W_k.
2. Generate B reference data sets with a random uniform distribution. Cluster each of these reference data sets with varying number of clusters k = 1, ..., k_{max}, and compute the corresponding total within intra-cluster variation W_{kb}.
3. Compute the estimated gap statistic as the deviation of the observed W_k value from its expected value W_{kb} under the null hypothesis:
   $Gap(k) = \frac{1}{B}\sum_{b=1}^{B} log(W_{kb}^*) - log(W_k)$. Compute also the standard deviation of the statistics.
4. Choose the number of clusters as the smallest value of k such that the gap statistic is within one standard deviation of the gap at k+1: $Gap(k) \geq Gap(k+1) - s_{k+1}$.

Reference: https://www.datanovia.com/en/lessons/determining-the-optimal-number-of-clusters-3-must-know-methods/