HDBSCAN cluster sizes of tokens x_i
(randomtext_albert-xlarge-v2_params-trained_dense-on-layers-48)