HDBSCAN cluster sizes of tokens  $x_i$  (imdb\_albert-xlarge-v2\_params-random\_dense-off\_layers-48)

