$\begin{array}{c} \text{HDBSCAN cluster sizes of tokens } x_i \\ (\texttt{imdb\_albert-xlarge-v2\_params-random\_dense-on\_layers-48}) \end{array}$ 

