$\label{eq:hdbscan} \begin{aligned} &\text{HDBSCAN cluster sizes of tokens } x_i \\ &(\texttt{wikitext\_albert-xlarge-v2\_params-random\_dense-on\_layers-48}) \end{aligned}$ 

