HDBSCAN attention_logit cluster sizes
(wikitext_bert-large-uncased_params-trained_dense-on_layers-24)