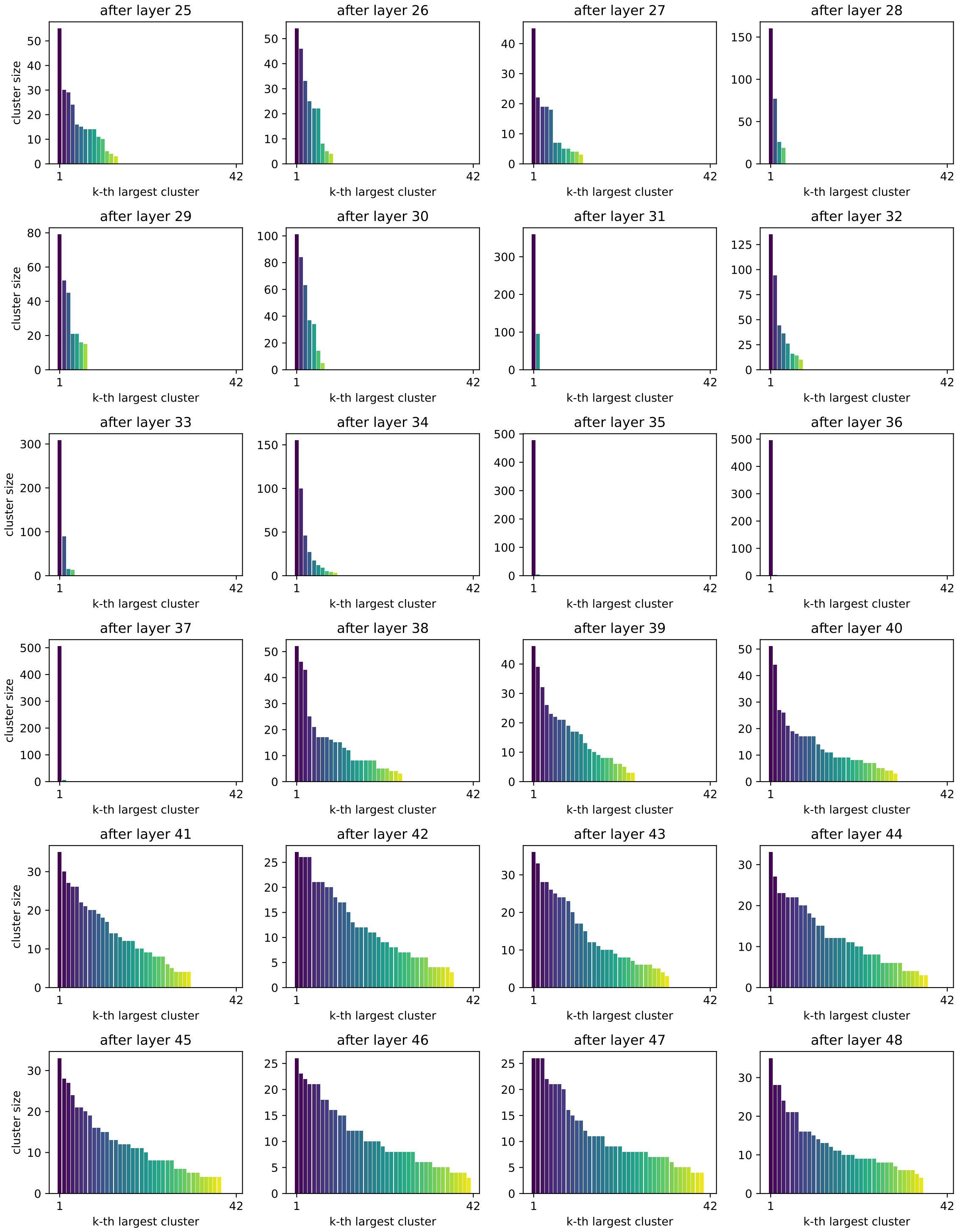HDBSCAN token cluster sizes
(randomtext_albert-large-v2_params-trained_dense-on_layers-48)