cos similarity histogram – layer 23 (bert-large-uncased\_wikitext\_\(Qx,Ky\)\_trained-params\_dense-off\_24-layers)

