Heatmaps of attention logits $\frac{1}{\sqrt{d'}}\langle Qx_i, Kx_j\rangle$

