

KL divergence for 2dData; darker = larger training sets

