VAE: loss ranges between 6.25 - 6.46
GAN: generator loss: 0.3-0.6, discriminator loss: 0.1

^ as shown in *_loss plots

gans: samples produced a lot of 1's, maybe because this does well against the discriminator.
    eventually produced more variance in numbers, but overall it doesnt stray far from 1.

vaes: samples produced fuzzy representations of numbers it got better with epochs
    they are still fuzzier than the gans but much better distribution

both were fine, gans has more intuitive set up I found. You have 2 models that are meant to optimize in
    step. Although both have very similar ideas.

vaes handled the training process a lot better, whereas gans required
more tedious training code.