

Well, that's the definition of a validation set – a hold-out set that you “validate” on by checking the stats on it. If you train on it, of course it doesn't mean anything. The problem is that even if you don't train on it, it doesn't seem to mean much. What eras are “representative” of the larger population and indicate a good generalized model? Well, probably none of them. We need a validation set with years worth of data in it. Sounds like that's what we'll actually get soon if the rumblings about releasing the targets for the test set happen.