* Next, we'll randomize the data. It's important to have the labels well shuffled for the training and test distributions to match.
  + Why? Distribution doesn’t matter on order…