Compute the total entropy in the dataset.

Illustrate splitting the data at feature Y (for yourselves on a sheet of paper) and compute the entropy after splitting at Y. Hint: Add up the entropy of each partition of the split. Compute the gain. Compute the gain for splitting at X.

It is ok to write down the calculations for just one case Y or X.

