the distribution of raw data