Distribution of Toxicity Scores in Filtered Data