Distribution of evaluation metric scores