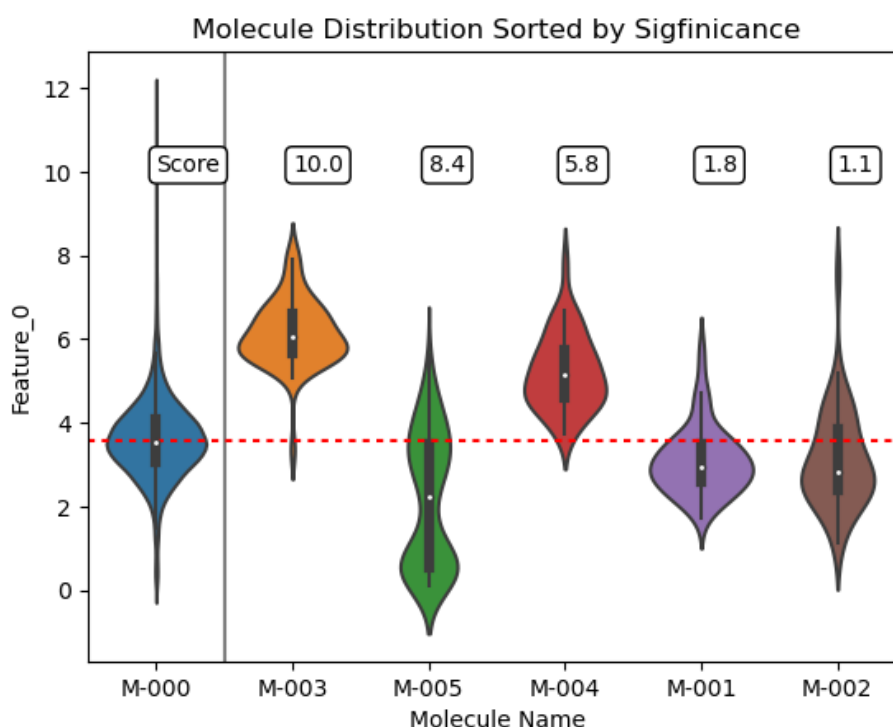


Scoring Metric:

To decide on a scoring metric, the first step is looking at the negative control distribution. According to the normality test, the negative control distribution is not normal therefore the scoring metric must take this into account. The Mann-Whitney U test if two distributions are derived from the same underlying source. In this application, this statistical test compares the distribution of *Feature_0* from the negative control and each molecule to evaluate if there is a significant difference. The score computed is an adjusted - log p-value computed by the Mann-Whitney U test. A larger - log p-value indicates that there is a larger difference in distribution between the negative control and the molecule, while a small - log p-value indicates little to no significant difference between the distributions. This score is scaled to have a maximum of 10 which is the molecule with the largest likelihood to be different than the negative control. It is also important to note that any molecule tested can have a difference either in the positive or negative direction. In other words, a positive direction indicates that the molecule mean is greater than the negative control while a negative direction indicates that the molecule mean is less than the negative control - both important types of molecules in drug discovery.

Plot Description:



This plot shows the distribution of each molecule. On the x-axis are the molecules, first the negative control followed by the other molecules sorted by the scoring metric. The y-axis is the response variable, *Feature_0*. The gray line separates the negative control and the other molecules being tested. The scoring metric above indicates that M-003 and M-005 are the most significant with different directions of effects. The red line is the mean of the negative control distribution to visually compare to the testing molecules direction of effect.