Creating crowd sourced sentiment analysis for the dataset would need to be relatively structured. We would not have freely typed analysis, and instead would have set sentiments to decide on. If we were able to add to the given dataset slightly, we would add some additional options than simply “boredom, enthusiasm and anger”, at the very least being able to select an ‘other’ option or indicate if the message appeared to be positive, negative or neutral. Along with this, it may be a useful option to be able to select multiple options. Regardless of the choice, preset sentiments would need to be chosen to prevent multiple users from using different words to describe the same thing or deviating too much from the desired result set.

With a dataset such as this, we would aim to have a wide range of contributors and be unbiased to who annotates what. To accomplish this, a ‘push’ methodology would be given, one in which the workers would be given data to work with. This dataset is too large to expect one person to annotate the entire thing, and as such a random set of ~50 tweets would be given instead. This set could be randomly chosen or targeted towards the data with the lowest number of notations. This would allow us to gather information which is unbiased, as it is randomly distributed and sourced from multiple different people, giving results which would represent a more realistic distribution of how each tweet was perceived. If we allowed users to pick which tweets to annotate themselves, there could be a large discrepancy in the number of annotations each tweet has, along with potential for certain controversial tweets being ‘vandalised’ or otherwise explicitly targeted.

In order to incentivise getting information, it may be possible to provide a small monetary output to users who fill out the data, but it may also result in users who do not provide helpful annotations and who only join for the money. In order to gather a set of contributors who would truly represent a realistic depiction of what the general sentiment would be, we would want relatively less screening as that could introduce bias, but we also need to ensure that the quality remains high. Ensuring that contributors can only provide one set of annotations and gathering data from large amounts of contributors should reduce the impact of some poor contributors.