a.

The aggregator argument determines the final prediction from the labels of the nearest neighbors. The mode, typically used in Classification problems, is necessary to determine the most common label among the nearest neighbors, in order to make the choice for a data point given the majority of a particular label that surrounds it. The mean, typically used in Regression problems, is necessary for obtaining a nuanced estimate of the target value by essentially taking in the summary of the distribution of values for each of the nearest neighbors surrounding a point. The difference between using the mode vs. the mean is that the mode provides the most common label whereas the mean provides the average value of the labels.

b.

Perhaps a third option could be using the median of the nearest neighbors to a data point. This would provide the middle-most value of the labels of the nearest neighbors. This can be best applied to situations that involve a plethora of outliers or extreme cases of values. This can provide more of an estimate than the mean for example, since the median is not affected *(as much)* by the outliers as a mean aggregator would be. Though, a drawback would not represent complex relationships well and can be more computational expensive for searching for that median value among the nearest neighbors.