

When dealing with time-series data, it can be tempting to simply disregard the timing structure and simply treat it as the appropriate form of categorical or numerical data.

One important concern, however, is that if you are building a predictive project looking at forecasting future data points. In this case, it is important **NOT** to use the future as a source of information! Since "hindsight is 20/20" and retrodictions are much easier than predictions, in predictive tasks it's generally a good idea to use a training set made up of data from before a certain point, a validation set of data from some dates beyond that, and testing data leading up to the present. This way your algorithm won't overfit by learning future trends.