The first question, you need to answer is whether you expect any time-based variations in what you are trying to predict?

Case 1: Time based output

Ispn most of the cases, the answer would be yes for such a long duration. It is hard to think of any thing, which you might be interested in predicting, but would not have any time-based changes. If the answer to the above question is yes, that it would always be a bad idea to split test and train using a time based (month / year) based variable. You will unnecessarily make your model blind to an important factor.

Let us understand this with an example:

* You are trying to predict demand / supply of a product based on several variables - Almost all products go through a cycle of early adaption, gaining momentum to dying off. If you keep your model blind to any of this period, you are making a gross error.
* stratified train and test split: sklearn.model\_selection.StratifiedKFold.html