

jackerparker:

- As in all Kaggle competitions (and all machine learning problems, for that matter), the most important first step is to get a validation set-up that matches the test set. There's no point in spending time on feature-engineering before your validation system is trustworthy.

[@jackerparker](#) what do you mean here by "get a validation set-up that matches the test set" ?