1. Stepwise selection methods are preferred over full model subset selection as the number of features (p) in your model increases due to their relative lower computational burden. What are cases where stepwise selection methods would give you *dramatically* different "best" models than a full model subset method?
2. Information criterion, like AIC and BIC, are meant to adjust your goodness-of-fit measure to account for the complexity that arises with increased number of parameters in your model. Given how they are implemented they are not perfect adjustments. Is the bias of these sorts of measures in favor or against more complex models? Be specific as to why you think they may bias towards selection of models with more or fewer numbers of parameters.