

Test accuracy time mean vs sd of test accuracy within cluster

