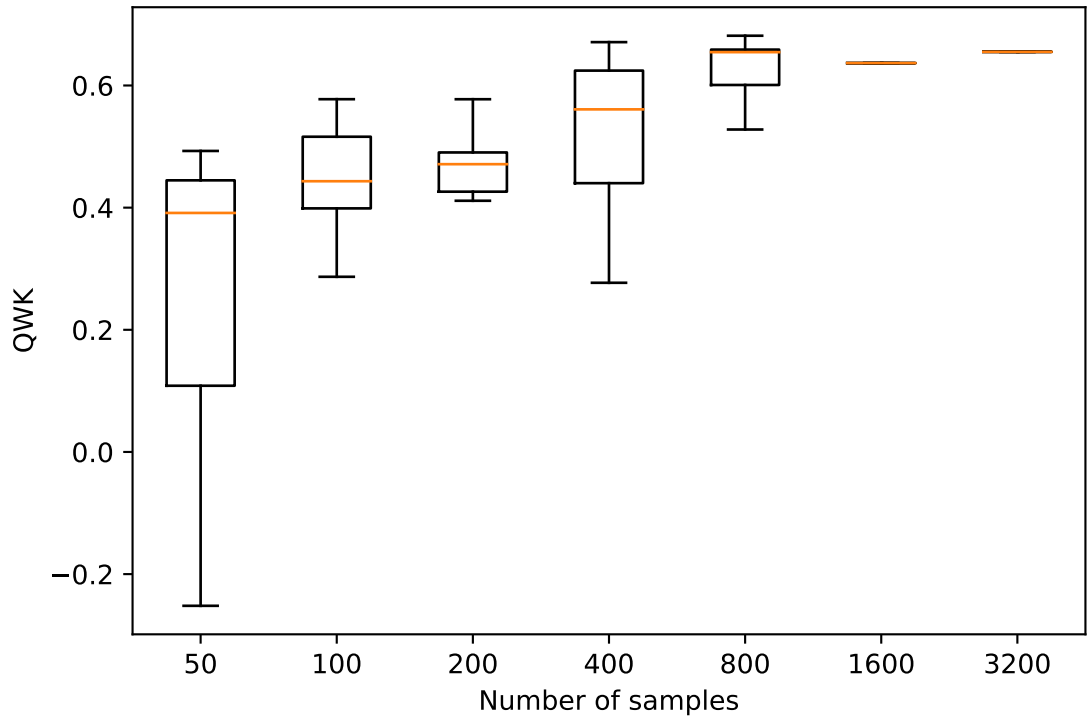


QWK of score\_1 assigned to samples created by experts-testing by two XGB Models  
One trained with samples created by davinci one trained with samples created by gpt4.



Each box represents exactly 6 kappa values