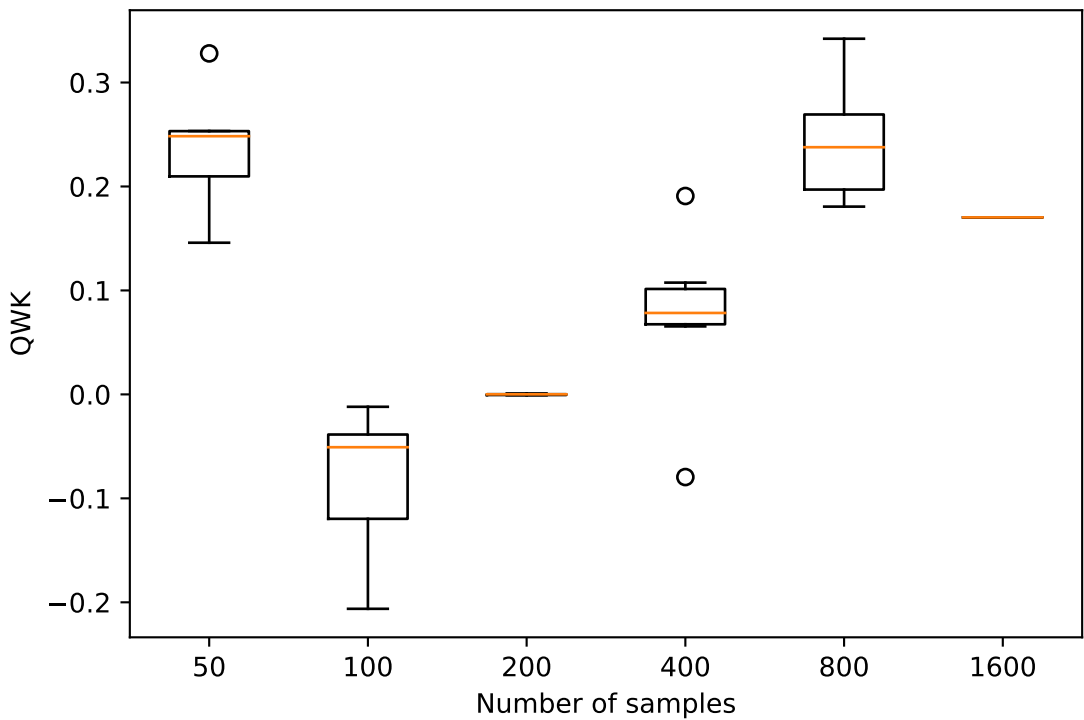


QWK of score\_1 assigned to samples created by davinci-testing by two BERT Models  
One trained with samples created by experts one trained with samples created by davinci.



Each box represents exactly 6 kappa values