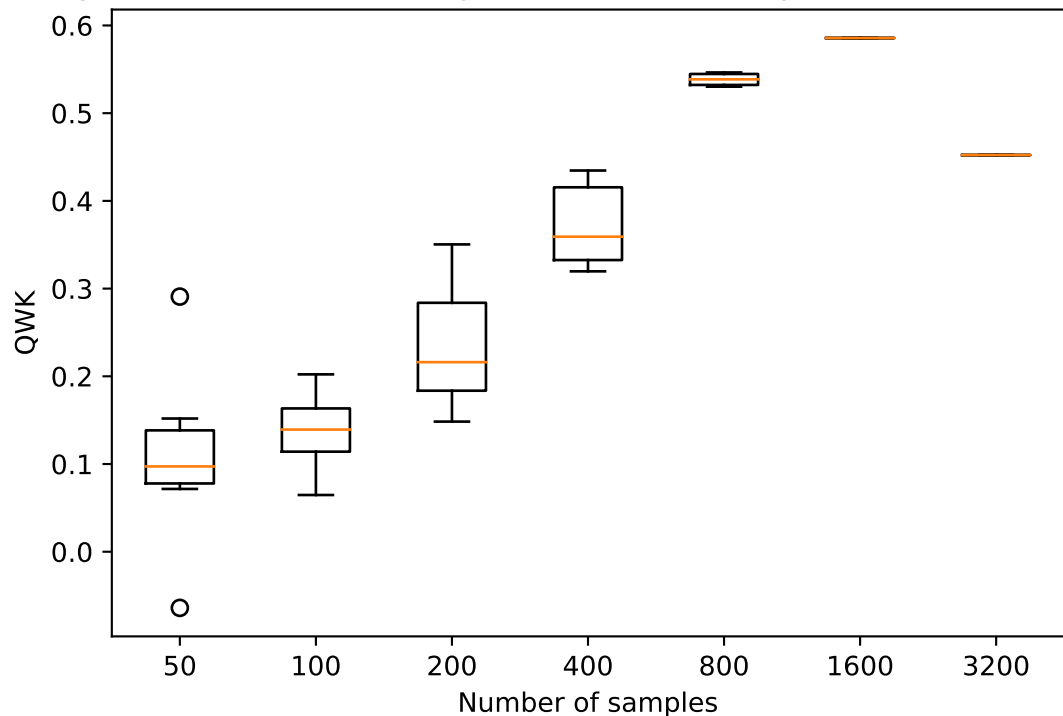XGB Model trained with samples created by text-davinci-003 rating answers created by text-davinci-003 annotated by gpt-3.5-turbo used for testing

Each box represents exactly 6 kappa values