

Label Counts by Language — Qwen

Label	Counter-Speech	Neutral	Refusal	Hate Speech
	271	16	2	11
	299	0	1	0
	293	3	4	0
Language				
Albanian				
Arabic				
Chinese				
English				