Tables

Classification Report:

	precision	recall	f1-score	support
Hate (0)	0.34	0.65	0.44	52
Offensive (1)	0.99	0.97	0.98	2662
accuracy			0.97	2714
macro avg	0.66	0.81	0.71	2714
weighted avg	0.98	0.97	0.97	2714

Figure 1. Classification Report — Hate vs Offensive

The report highlights the challenge of class imbalance. Offensive tweets achieve near-perfect precision (0.99), recall (0.97), and F1 (0.98). Hate tweets, however, exhibit low precision (0.34) and F1 score (0.44), indicating a high number of false positives and fewer correct identifications.

	neg				neu				\
	mean	median	std	count	mean	median	std	count	
class									
0	0.290893	0.281	0.231636	261	0.648471	0.637	0.230802	261	
1	0.266070	0.264	0.200650	13306	0.639633	0.631	0.207114	13306	
2	0.061550	0.000	0.112813	2753	0.830401	0.844	0.164431	2753	

	pos				compound			
	mean	median	std	count	mean	median	std	count
class								
0	0.060644	0.0	0.103726	261	-0.409708	-0.5719	0.453484	261
1	0.094296	0.0	0.127670	13306	-0.333923	-0.5267	0.462047	13306
2	0.108052	0.0	0.142384	2753	0.095741	0.0000	0.403498	2753

Figure 2. Sentiment Statistics by Class

This table summarizes VADER sentiment scores. Hate (0) and Offensive (1) tweets have higher negative sentiment means (0.29 and 0.27) and strongly negative compound scores (–0.41, –0.33). In contrast, Neutral tweets (Class 2) skew strongly towards neutrality (mean 0.83) with a near-zero compound score.