

Mental health text sentiment — concise summary

I ran VADER sentiment analysis on the statements in the dataset to understand emotional tone across mental-health labels. The dataset contains 53,043 rows of short, informal text classified into seven labels (Normal, Depression, Suicidal, Anxiety, Bipolar, Stress, Personality disorder). The analysis produces per-row sentiment scores (positive/neutral/negative components and a compound score from -1 to +1) and a categorical sentiment label (negative / neutral / positive).

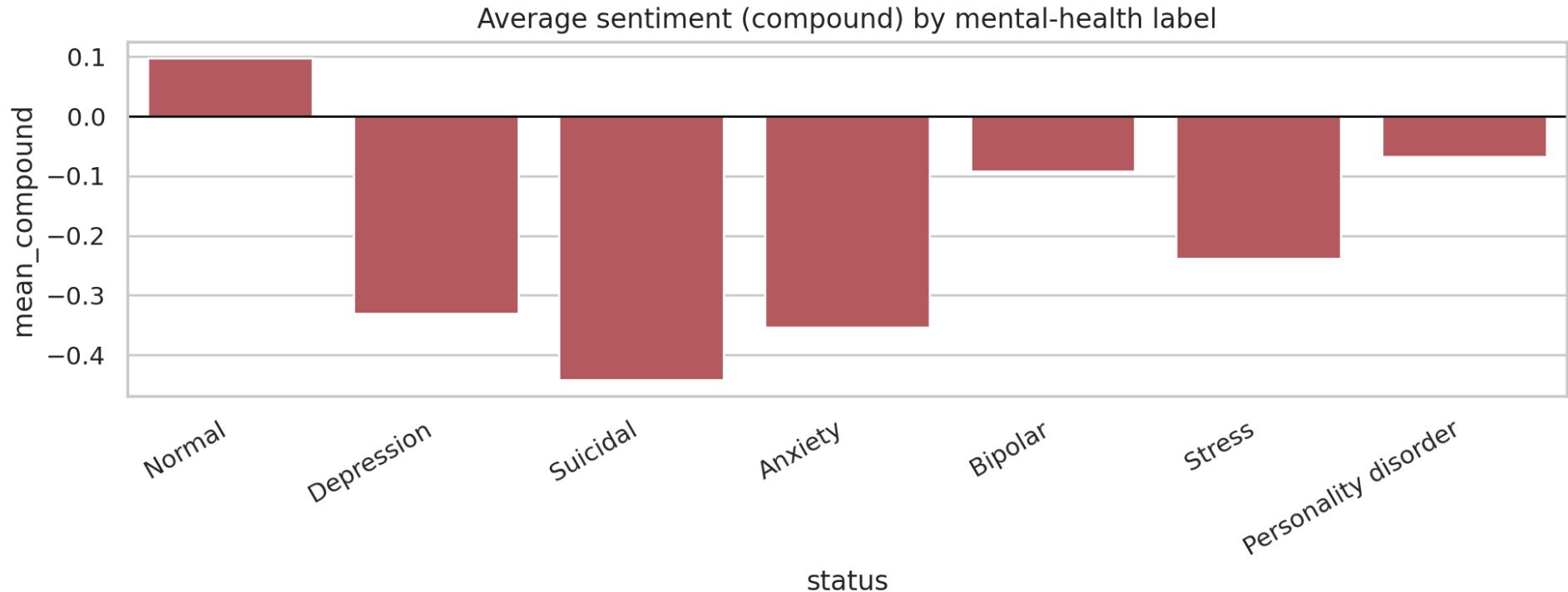
Top-line summary

- Overall the corpus skews negative: the mean compound score is about -0.199 and the median is -0.258.
- Depression and Suicidal posts are the most negative on average and by median — these categories show the strongest negative language.
- Normal posts are the most positive on average and have the largest share of positive statements.
- Anxiety and Stress skew negative too, but include a meaningful share of positive-scored statements (often indicating coping language or hopeful phrasing).
- Neutral-scored statements are relatively uncommon in clinical labels (Depression, Suicidal, Anxiety) — text there is usually emotionally charged.

Insight — Depression and Suicidal content shows the strongest negative tone

- Supporting data: By-label summary (count, mean/median compound, and percent negative/neutral/positive) shows:
 - Depression: n = 15,404, mean compound ≈ -0.330, median ≈ -0.7092, ~67.7% negative.

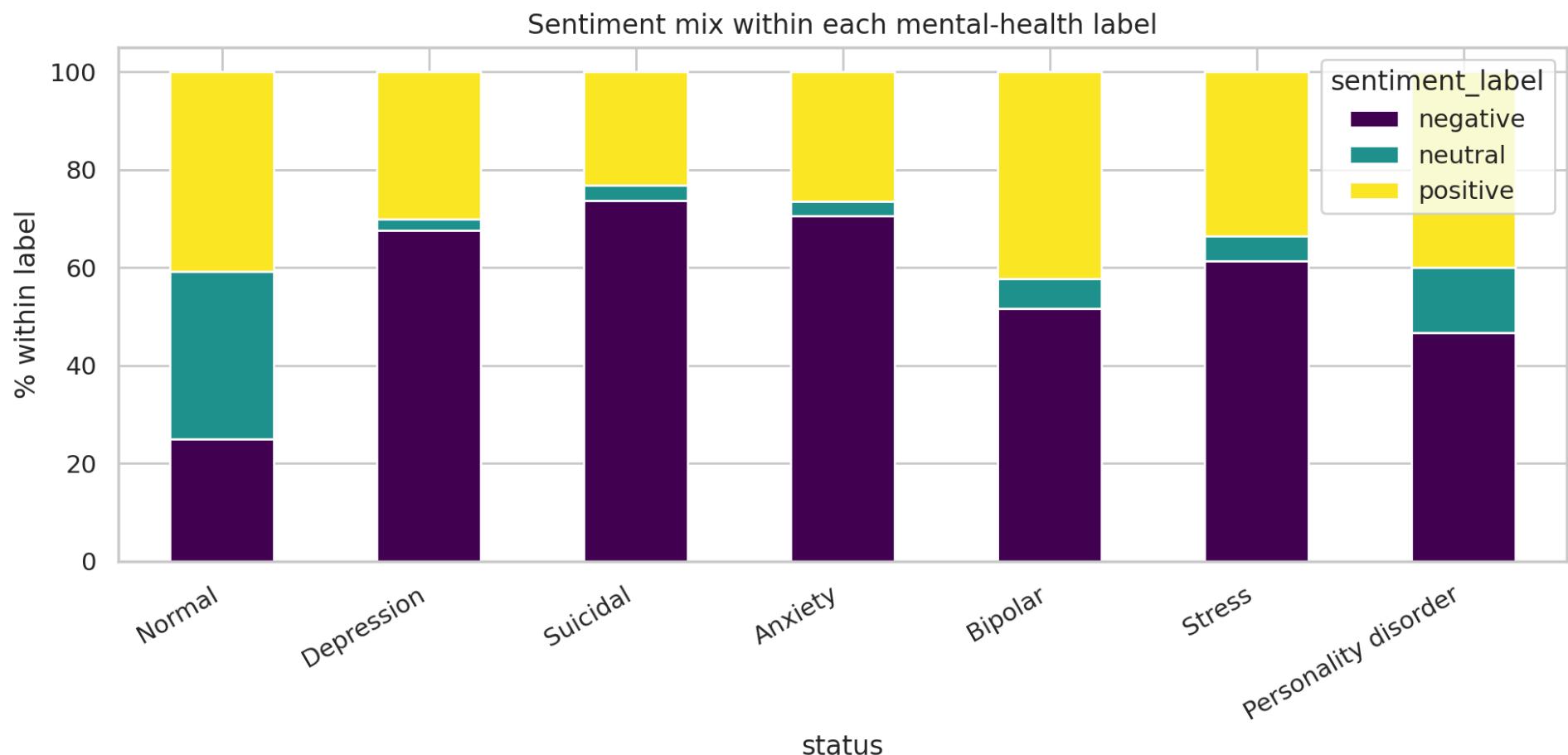
- Suicidal: n = 10,653, mean compound \approx -0.442, median \approx -0.8, ~73.6% negative.
- Visual reference: Average sentiment by label chart (see the PDF or image below for the bar plot that highlights these low mean compound scores).

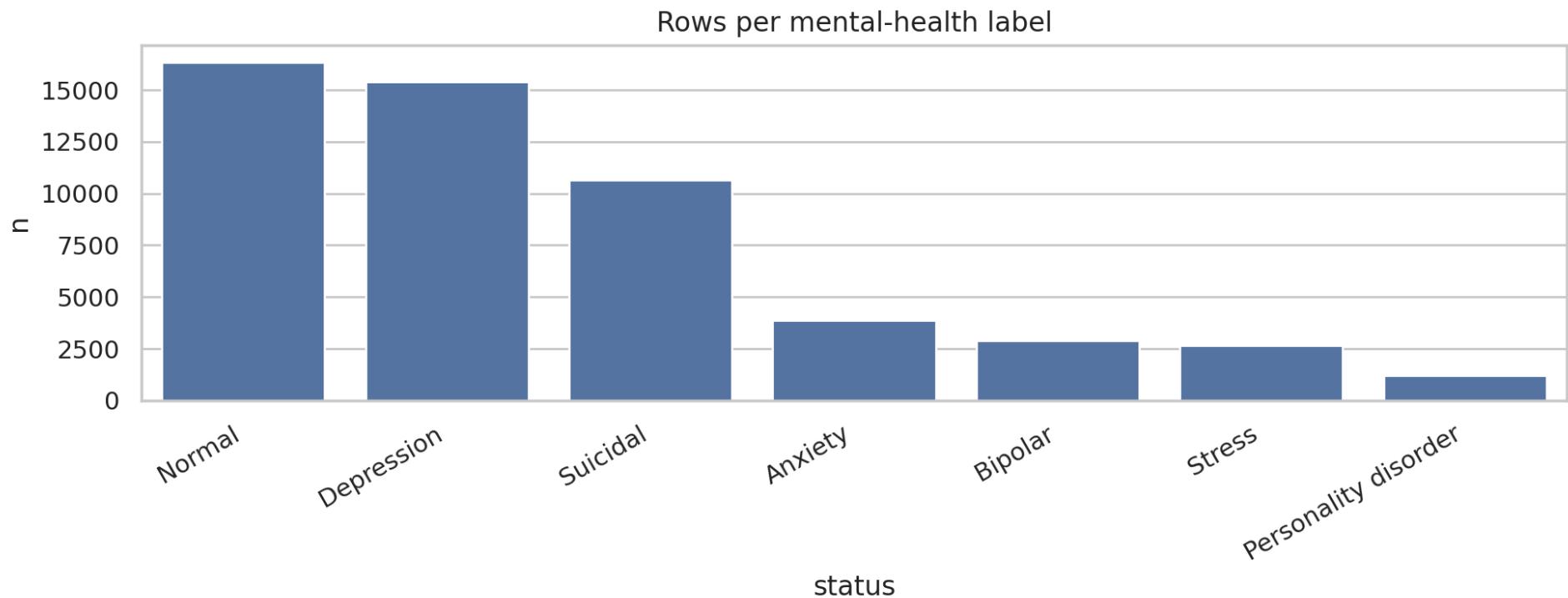
[!\[\]\(99f58673407353e96a019fbca558fd72_img.jpg\) Download](#)

Insight — Normal posts trend positive and are more balanced

- Supporting data: Normal: n = 16,351, mean compound \approx 0.098, median \approx 0.0, with ~40.7% positive and ~34.3% neutral.
- This indicates Normal-class text often uses neutral-to-positive language rather than strongly negative affect.

- Visual reference: Sentiment mix and counts charts show Normal's larger positive/neutral shares.

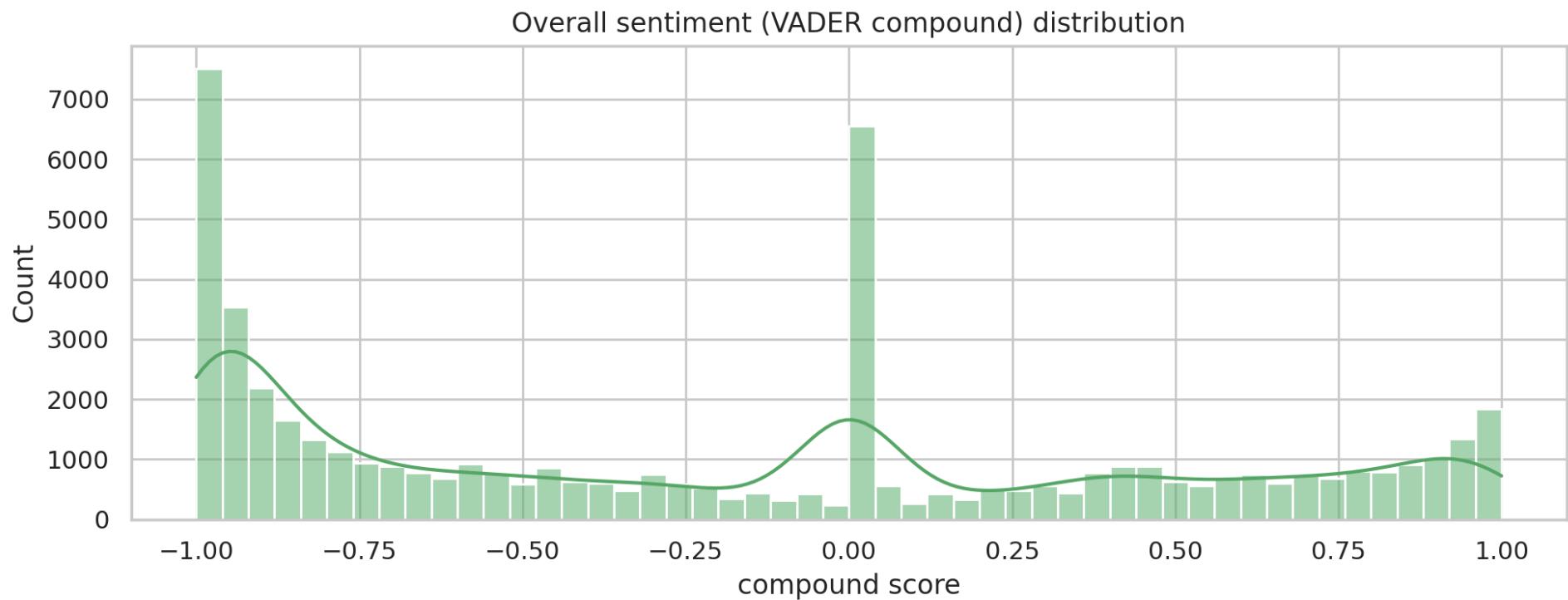
[!\[\]\(4729e517bc6a7cd81c8025b9646574fb_img.jpg\) Download](#)[!\[\]\(a03a7eb2f4046e1d3c76772003e549ea_img.jpg\) Download](#)



Insight — Anxiety and Stress are negative but contain meaningful positive language

- Supporting data:
 - Anxiety: n = 3,888, mean compound ≈ -0.354, ~70.6% negative, ~26.5% positive.
 - Stress: n = 2,669, mean compound ≈ -0.239, ~61.4% negative, ~33.5% positive.
- Interpretation: Posts about anxiety and stress frequently contain both negative descriptions and positive coping or reassurance language — hence the mixed sentiment profile.

- Visual reference: Sentiment distribution histogram and label mix chart show these mixed profiles.

[Download](#)


Compact table of per-label results

- Table shows n, mean and median compound, and percent negative/neutral/positive for each label.

status	n	mean_compound	median_compound	neg_pct	neu_pct	pos_pct
Normal	16351	0.0983	0.0	24.9709	34.3037	40.7253

status	n	mean_compound	median_compound	neg_pct	neu_pct	pos_pct
Depression	15404	-0.3302	-0.7092	67.6772	2.2397	30.0831
Suicidal	10653	-0.4417	-0.8	73.6037	3.1447	23.2517
Anxiety	3888	-0.3538	-0.6124	70.6019	2.8807	26.5175
Bipolar	2877	-0.0915	-0.1969	51.6858	6.0132	42.3010
Stress	2669	-0.2386	-0.4917	61.3713	5.1330	33.4957
Personality disorder	1201	-0.0672	0.0	46.7111	13.3222	39.9667

(Note: sentiment label rules used — negative if compound < -0.05; neutral if $-0.05 \leq \text{compound} \leq 0.05$; positive if compound > 0.05.)

Concluding summary: the most critical signals are strong negative language concentrated in Depression and Suicidal posts, while Normal posts are notably more positive and balanced. Anxiety and Stress sit between those extremes with substantial negative language but also visible positive/coping content.