

Using NLP to Classify Mood Disorders

Overview

This project analyzes an integrated dataset from Kaggle: a CSV file of compiled and cleaned data from 9 different sources pulling and classifying social media posts into 7 mental health statuses.

The data can be accessed here: [Sentiment Analysis for Mental Health](#)

Business Problem

According to the *World Health Organization*, in 2019 970 million people were living with a mental disorder. COVID-19 exacerbated this, with a 26% and 28% increase in anxiety and major depressive disorders, respectively. With the surge in technological advancements and use of social media, paired with a greater shift toward remote and virtual communications as a result of the pandemic...accessible, online resources for people seeking mental health support are essential.

- My goal is to build a Natural Language Processing Model which can analyze a person's posts on social media, and accurately categorize whether they have a mood disorder, as well as which disorder they have if relevant
- By identifying key words, engineering additional important features, and fine tuning the best model, we can hope to use this information to develop a chatbot that can direct people to further relevant resources (such as therapists specializing in the particular mood disorder a person needs help with)

Data Understanding

Due to size constraints of these data sources, I am not able to store the raw data in the online repository where this analysis resides. If the reader wishes to explore those data sets or run the code within this file, it will be necessary to find the sources online and download them. These sources include:

- [3k Conversations Dataset for Chatbot](#)
- [Depression Reddit Cleaned](#)
- [Human Stress Prediction](#)
- [Predicting Anxiety in Mental Health Data](#)
- [Mental Health Dataset Bipolar](#)

- Reddit Mental Health Data
- Students Anxiety and Depression Dataset
- Suicidal Mental Health Dataset
- Suicidal Tweet Detection Dataset

Data Preparation

The initial dataset was over 51,000 rows with 7 different classes to predict in our target variable. Therefore, one of my main tasks was balancing detail-oriented fine tuning with reducing computational expense. A huge part of data preparation was trying to cut off as much of the “dead weight” as possible, to reduce noise and be able to parse through which features are truly important and unique to predicting our different target classes. To cut the “dead weight” I dropped stop words, punctuation and special characters (such as emojis), combined classes where possible, and ended up dropping our ‘Normal’ status rows altogether. To be meticulous and detail-oriented with regards to identifying and engineering rare, predictive features, I used lemmatization rather than stemming on tokenized words, created new numeric features (such as Percentage of Upper-Case Words, and Exclamation Point Count) that might aide my model and allow me to drop more unnecessary tokens, and combed through bigrams and tokens common to multiple classes to relabel rows where necessary.

To prepare data for all models I used TF-IDF Vectorization.

Modeling

The scikit-learn package was primarily used to run Multinomial Naive Bayes, Multinomial Logistic Regression, and Random Forest. I quickly realized our dataset was quite messy, large, and overly complex, and at first Multinomial Naive Bayes was the only model that would successfully run without worker timeouts, errors regarding max_iterations, etc. A large part of my tuning was figuring out the best way to select features (SelectKBest with chi-squared ended up being our most helpful feature selection tool when possible), as well as figuring out how many maximum features, and/or maximum iterations a model needed to run without running a high risk of overfitting.

Other hyper-parameters were tested and tuned, but balancing a model’s need for many features and/or iterations to run, with protecting against overfitting was my main struggle.

The classes were heavily imbalanced, so adjusting class weights (class_prior for Naive Bayes) was also essential.

Evaluation

Random Forest was the most successful model. My main metric was F1-score (weighted to account for imbalance): false positives and false negatives run equal measures of risk...for example when misdiagnosis of a mood disorder results in incorrect prescribed medications that severely alter brain chemistry, or suicidal threats are mislabeled leading to a lack of proper immediate assistance and intervention. Furthermore F1 Score is helpful when classes are imbalanced.

I used train-test-split and k-fold cross-validation.

Initial scores on the first Naive Bayes model were (as I said, Random Forest would not even successfully run):

Baseline Best F1 Score: 0.4305

Baseline Test Set Score: 0.4659

There were continued discrepancies between a model's Best F1 Score and Test Set Score at the start of this project, as I continued to run models trying to get them to work. After much careful tuning and feature engineering, my final models yielded nearly identical scores between Best F1 and Test Set.

Final Random Forest mode scores were:

Final Best F1 Score: 0.7643

Final Test Set Score: 0.7668

Given the complexity and messiness of this dataset, I'm very happy with the proportional magnitude at which my final scores had improved from my baseline model, even if there's still room for further improvement.

(XGBoost was also tested at the very end, but it yielded similar results to our final Random Forest model. I've picked random forest because it is less computational expensive.)

```
In [46]: import pandas as pd
import numpy as np

import seaborn as sns
import geopandas as gpd
import matplotlib.pyplot as plt
from matplotlib.colors import ListedColormap
import plotly.express as px
from wordcloud import WordCloud
import os

from statsmodels.stats.outliers_influence import variance_inflation_factor
```

```
from statsmodels.tools.tools import add_constant

from sklearn.preprocessing import OneHotEncoder, StandardScaler, LabelEncoder
from sklearn.model_selection import train_test_split, GridSearchCV, cross_val_score
from sklearn.linear_model import LogisticRegression, Lasso
from sklearn.metrics import classification_report, silhouette_score, confusion_matrix
from sklearn.tree import DecisionTreeClassifier, plot_tree
from sklearn.ensemble import RandomForestClassifier
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
from sklearn.impute import SimpleImputer
from sklearn.neighbors import KNeighborsClassifier
from sklearn.feature_extraction.text import TfidfVectorizer, CountVectorizer
from sklearn.naive_bayes import MultinomialNB, ComplementNB
from sklearn.cluster import KMeans
from sklearn.base import BaseEstimator, TransformerMixin
from sklearn.feature_selection import SelectKBest, chi2, SelectFromModel

from imblearn.pipeline import make_pipeline
from imblearn.over_sampling import SMOTE

from scipy.stats import chi2_contingency, multinomial
from scipy.sparse import hstack, csr_matrix

import nltk
import re
nltk.download('punkt')
from nltk.tokenize import word_tokenize, RegexpTokenizer, regexp_tokenize, sent_tokenize
nltk.download('stopwords')
from nltk.corpus import stopwords, wordnet
from nltk.stem import SnowballStemmer, LancasterStemmer, PorterStemmer
from nltk.probability import FreqDist
from nltk import WordNetLemmatizer, pos_tag
from nltk.util import bigrams
nltk.download('averaged_perceptron_tagger')
nltk.download('wordnet')
nltk.download('omw-1.4')
#nltk.download('all')
from nltk.collocations import BigramCollocationFinder, BigramAssocMeasures
from nltk.metrics import BigramAssocMeasures

import itertools
import string
import contractions

import os
import sys
module_path = os.path.abspath(os.path.join(os.pardir, os.pardir))
if module_path not in sys.path:
    sys.path.append(module_path)

from copy import deepcopy

from xgboost import XGBClassifier

from joblib import parallel_backend
```

```
from memory_profiler import profile

from collections import Counter

[nltk_data] Downloading package punkt to
[nltk_data]      /Users/emmascotson/nltk_data...
[nltk_data]      Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data]      /Users/emmascotson/nltk_data...
[nltk_data]      Package stopwords is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]      /Users/emmascotson/nltk_data...
[nltk_data]      Package averaged_perceptron_tagger is already up-to-
[nltk_data]      date!
[nltk_data] Downloading package wordnet to
[nltk_data]      /Users/emmascotson/nltk_data...
[nltk_data]      Package wordnet is already up-to-date!
[nltk_data] Downloading package omw-1.4 to
[nltk_data]      /Users/emmascotson/nltk_data...
[nltk_data]      Package omw-1.4 is already up-to-date!
```

Loading the Data

Let's load our dataset and take a look at it.

```
In [47]: data = pd.read_csv('data/Combined Data.csv')
```

```
In [48]: data.head()
```

```
Out[48]:   Unnamed: 0                      statement  status
0            0                           oh my gosh  Anxiety
1            1  trouble sleeping, confused mind, restless hear...  Anxiety
2            2    All wrong, back off dear, forward doubt. Stay ...  Anxiety
3            3    I've shifted my focus to something else but I'...  Anxiety
4            4    I'm restless and restless, it's been a month n...  Anxiety
```

```
In [49]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 53043 entries, 0 to 53042
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Unnamed: 0   53043 non-null   int64  
 1   statement    52681 non-null   object  
 2   status       53043 non-null   object  
dtypes: int64(1), object(2)
memory usage: 1.2+ MB
```

Looks like we have some NaNs. Let's take a closer look.

In [50]: `data[pd.isna(data['statement'])]`

Out[50]:

	Unnamed: 0	statement	status
293	293	NaN	Anxiety
572	572	NaN	Anxiety
595	595	NaN	Anxiety
1539	1539	NaN	Normal
2448	2448	NaN	Normal
...
52838	52838	NaN	Anxiety
52870	52870	NaN	Anxiety
52936	52936	NaN	Anxiety
53010	53010	NaN	Anxiety
53031	53031	NaN	Anxiety

362 rows × 3 columns

There are so few NaN's compared to the number of rows in our entire dataframe. Furthermore, rows with no 'statement' value are useless to us. Let's drop them.

In [51]: `data = data.dropna(subset=['statement'])`

Now let's check if there are duplicate rows.

In [52]: `data['statement'].nunique()`

Out[52]: 51073

It looks like there are!

In [53]: `data[data['statement'].duplicated()]`

Out[53]:

	Unnamed: 0		statement	status
97	97	"No regrets or grudges/angry at things that ha...	Anxiety	
138	138	but my heart is still restless even though my ...	Anxiety	
167	167	I want to exhale the restlessness in my chest ...	Anxiety	
228	228	Do not compare yourself to others. Envy only m...	Anxiety	
244	244	people seem calm, happy like there's no proble...	Anxiety	
...
53038	53038	Nobody takes me seriously I've (24M) dealt wit...	Anxiety	
53039	53039	selfishness "I don't feel very good, it's lik...	Anxiety	
53040	53040	Is there any way to sleep better? I can't slee...	Anxiety	
53041	53041	Public speaking tips? Hi, all. I have to give ...	Anxiety	
53042	53042	I have really bad door anxiety! It's not about...	Anxiety	

1608 rows × 3 columns

These statements all look pretty specific and personalized. AKA it does seem like our duplicate values are true duplicates of the same social media statements made by the same singular user. Furthermore, we have over 50,000 rows in our dataset and can afford to lose these rows.

Let's drop duplicates as well.

```
In [54]: # Removing 'statement' duplicates, keeping first instance
data = data.drop_duplicates(subset=['statement'], keep='first')
```

```
In [55]: data[data['statement'].duplicated()]
```

```
Out[55]: Unnamed: 0 statement status
```

```
In [56]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 51073 entries, 0 to 52840
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Unnamed: 0   51073 non-null   int64  
 1   statement    51073 non-null   object  
 2   status       51073 non-null   object  
dtypes: int64(1), object(2)
memory usage: 1.6+ MB
```

Furthermore, our 'Unnamed: 0' column appears to just be a duplicate of our index. Let's explore.

```
In [57]: data['Unnamed: 0'].nunique()
```

```
Out[57]: 51073
```

```
In [58]: # Checking if 'Unnamed: 0' equals our index  
is_identical = data['Unnamed: 0'].equals(data.index)  
is_identical
```

```
Out[58]: False
```

Hm...We can safely assume this column doesn't provide information about multiple 'statement' values being generated by the same singular user...because our nunique() matches the total number of rows in our dataframe, which eliminates the possibility of duplicates.

This column is most likely just an old index from previous datasets, which is no longer useful to use. We'll drop it for now. We can always go back and edit the code if we decide we need it later.

```
In [59]: data.drop(columns=['Unnamed: 0'], inplace=True)
```

Let's look at the value_counts() for our target variable 'status'.

```
In [60]: data['status'].value_counts()
```

```
Out[60]: status  
Normal           16039  
Depression       15087  
Suicidal         10641  
Anxiety          3617  
Bipolar           2501  
Stress            2293  
Personality disorder    895  
Name: count, dtype: int64
```

'Personality disorder'

'Personality disorder' is an ambiguous label compared to the others. It might be a version of a 'placeholder' value for not 'Normal' users where further categorization of disorder (ex. 'Depression', 'Suicidal', 'Anxiety') was inconclusive. Alternatively, it could be indication of other, entirely different personality disorders not listed above.

We'll keep this in mind and explore as we go. Depending on what we find, it might be helpful to drop these values when building a predictive model.

Preprocessing & EDA

Prior to Cleaning...

We'll want to clean the text in our data by performing operations that remove punctuation and special characters, lowercase text, remove newline '\n' characters, etc.

However, **before** we do this...let's look at whether or not we can use any of our 'pre-cleaned' text to create features that might be useful down the road.

Percentage Upper-Case

Let's think about whether or not any unique punctuation and/or text characteristics might be indicative of a personality disorder.

Upper-case text might help identify whether or not a person is in some kind of distress. If user posts 'A MESSAGE ENTIRELY IN UPPER-CASE LIKE SO', that's an unusual behavior that we should try and quantify.

Let's create a new column 'perc_upper' that calculates the percentage of upper-case letters to the total number of letters in each 'statement'.

```
In [61]: # Define function to calculate percentage of upper-case letters
def perc_upper(text):
    if not isinstance(text, str):
        return 0
    letters = re.findall(r'[a-zA-Z]', text) # Using regex to isolate letters
    if not letters:
        return 0
    upper_count = len(re.findall(r'[A-Z]', text))
    return (upper_count / len(letters)) * 100

data['perc_upper'] = data['statement'].apply(perc_upper)
```

```
In [62]: upper_sorted = data.sort_values(by='perc_upper', ascending=False)
upper_sorted.head()
```

Out [62]:

		statement	status	perc_upper
3546	[HELP RT] WE FANBASE SHAKE RP! JOIN? FOLLOW FI...		Normal	100.0
2436		ONAKA GA SUITA	Normal	100.0
6787		TODAY NO CLASS YAAYYY	Normal	100.0
5303	6 HALF HOURS AGAINIII ULULU I WANT TO SLEEP ...		Normal	100.0
2478	I DON'T HAVE A HOLIDAY AS WELL AS EVIL		Normal	100.0

METRIC CHANGE - Percentage by Upper-Case Words

On second thought, we should find a different way to quantify a notable amount of upper-case in a string of text. With our current 'perc_upper' value...a statement such as

'Hi' will have a 50% perc_upper. However this would be due to normal grammatical capitalization techniques that are of no note.

Let's alter our metric slightly, to calculate the **percentage of upper-case words compared to the total number of words in a statement**.

We'll stick with regex for now to tokenize our words and calculate this metric.

```
In [63]: # Dropping perc_upper column before to reduce computation time
data.drop(columns=['perc_upper'], inplace=True)

# Updating perc_upper to calculate based on number of words
def perc_upper_words(text):
    if not isinstance(text, str):
        return 0
    words = re.findall(r'\b[a-zA-Z\-\']+\b', text) # Including apostrophes are included in the word
    words = [word for word in words if not re.search(r'\d', word)] # Filter out numbers
    if not words:
        return 0
    uppercase_words = [word for word in words if word.isupper()]
    return len(uppercase_words) / len(words) * 100

data['perc_upper_words'] = data['statement'].apply(perc_upper_words)
```

```
In [64]: upper_sorted = data.sort_values(by='perc_upper_words', ascending=False)
upper_sorted.head()
```

		statement	status	perc_upper_words
11210	I KEEP MESSING THINGS UP ALL DAY EVERY DAY THE...	Suicidal	100.0	
5917	#PECAT WHOSE NAME IS SI ALI MOCHTAR NYEBELIN #	Normal	100.0	
4173	OH MY GOD FEAR CANCEL	Normal	100.0	
1581	HIS BD CARD KANON IS CUTE HSHSHSHS SAD NO FREE...	Normal	100.0	
2811	YAALLAH SO NATION OF SM DREAMIES	Normal	100.0	

```
In [65]: # Sampling rows from the middle
filtered_rows = upper_sorted[(upper_sorted['perc_upper_words'] >= 50.0) & (upper_sorted['perc_upper_words'] <= 100.0)]
filtered_rows.head(20)
```

Out[65]:

		statement	status	perc_upper_words
10953	NOBODY WANTS TO TALK TO ME	Going to kill myself	Depression	60.000000
498	the habit of chatting for a day, gadicchat imm...	Anxiety	60.000000	
17851	I feel like I cannot even explain myself becau...	Suicidal	60.000000	
4525	NOT FUNNY MORNING â€“ Runny nose	Normal	60.000000	
24841	I try my best but I just want to die. I hate m...	Depression	59.420290	
6357	HAHAAAA GET THE ASSIGNMENT OF MAKING A SCIENTI...	Normal	57.894737	
1113	WANT TO SIN BUT Afraid to Laugh	Normal	57.142857	
36232	Everything is temporary. EVIL EYES OFF SHIVRIT	Normal	57.142857	
830	MORNING AGAIN I've made a AMENDE SPACE	Normal	57.142857	
48673	AAAAAAAAAAAAAAAAAA aaAAAAAAAAA FIDBWJSKLA oaa...	Stress	57.142857	
4557	if this passes, DROP YOUR LAST COPY !	Normal	57.142857	
7490	This is what the government wants poor people ...	Suicidal	56.862745	
19137	I am here to talk if you need help HAHAHAHAA A...	Depression	56.756757	
22163	I MIGHT KILL MSELF BECAUSE I am A RETARDED USE...	Suicidal	56.521739	
23744	i just want to fit in. i do not want to be mad...	Depression	56.074766	
8319	Bitch why the fuck did you just tell me to che...	Depression	56.020942	
15768	Welp fuck it everything I do gets put on the b...	Depression	55.882353	
25969	My humiliation and mental torture of sleeping ...	Suicidal	55.882353	
5121	type IMB space DENITAKOPLOK send to 14045, TON...	Normal	55.555556	
8227	I really regret talkin or sayin anything. I am...	Suicidal	54.255319	

Let's see if there are any trends regarding the distribution of these numbers within each of our 'status' values.

In [66]:

```
upper_stats = data.groupby('status')['perc_upper_words'].agg(['mean', 'min', 'upper_stats'])
```

Out[66]:

		mean	min	max
	status			
	Anxiety	5.511061	0.0	100.000000
	Bipolar	5.978278	0.0	25.000000
	Depression	5.667380	0.0	100.000000
	Normal	3.903474	0.0	100.000000
	Personality disorder	5.424887	0.0	27.118644
	Stress	5.664143	0.0	57.142857
	Suicidal	7.922907	0.0	100.000000

Interesting. 'Normal' has the lowest average perc_upper_words with a mean of 3.90, whereas 'Suicidal' has the highest with a mean of 7.92.

Let's see if there are any noticeable trends for the same metric applied to lower-case words. A higher percentage of lower-case words might be indicative of disorders like 'Depression'.

In [67]:

```
# Updating perc_upper to calculate based on number of words
def perc_lower_words(text):
    if not isinstance(text, str):
        return 0
    words = re.findall(r'\b[a-zA-Z\-\']+\\b', text) # Including apostrophes are included in the word
    words = [word for word in words if not re.search(r'^\\d', word)] # Filter out numbers
    if not words:
        return 0
    lowercase_words = [word for word in words if word.islower()]
    return len(lowercase_words) / len(words) * 100

data['perc_lower_words'] = data['statement'].apply(perc_lower_words)
```

In [68]:

```
lower_stats = data.groupby('status')['perc_lower_words'].agg(['mean', 'min', 'lower_stats'])
```

Out[68]:

		mean	min	max
	status			
	Anxiety	87.107075	0.000000	100.0
	Bipolar	86.984639	25.000000	100.0
	Depression	90.887056	0.000000	100.0
	Normal	89.098266	0.000000	100.0
	Personality disorder	87.738038	45.454545	100.0
	Stress	87.853044	14.285714	100.0
	Suicidal	87.430360	0.000000	100.0

'Depression' does indeed have the highest mean compared with our other 'status' values.

However, typing in all lower-case is far more socially "standard" than typing in caps-lock. People often remove automatic capitalization as a setting on their phone or computer, which sets their default text to all lower-case.

Given that lower-case typing is a societal 'norm', and a personal preference people often make regardless of psychological state...the 'lower_stats' numbers above do not vary drastically between 'status' groups enough for us to be able to use this as a predictive metric.

We'll drop it and just keep perc_upper_words.

```
In [69]: data.drop(columns=['perc_lower_words'], inplace=True)
```

Special Characters

Let's explore if it will be useful to apply similar logic to special characters and punctuation. Let's examine all the different special characters in our entire dataframe, excluding numbers.

```
In [70]: # Searching for special characters, excluding numbers
def extract_special(text):
    if not isinstance(text, str):
        return ''
    return re.findall(r'[^w\s\d]', text)

special_chars = data['statement'].apply(extract_special).explode().unique()
```

```
In [71]: special_chars
```

Emojis

Let's try and categorize the emojis used. There are few enough that we can manually create lists identifying emojis we deem to carry clear 'positive' or 'negative' sentiments. We'll ignore any "neutral" emojis: we're about to get into more detailed and layered analysis of the actual text in a bit, so we don't need to waste time getting too intricate with these additional features!

If we have to choose one or the other, we'll probably choose **negative** emojis as our main predictive metric (with regards to emoji sentiments)...since positive emojis could be used sarcastically.

The sarcastic use of positive "sentiments" to indicate negativity is far more common than the sarcastic use of negative sentiments to indicate positivity.

Varying Cultural Expression

There can be differences in a person's interpretation and use of emojis when conveying sentiments, based on a variety of cultural factors. We don't know enough about the

people who created this dataset, nor do we have demographic information on the users in the dataset itself, to make those determinations...We'll categorize emoji sentiments through our own cultural lens, and see how well our version of this metric helps or hurts our predictive model.

Binary Classification

As we stated, we don't need to get too intricate. Let's create a binary classifier to simply determine whether a 'statement' has any number of positive or negative emojis in it's text.

```
In [73]: # 1 if any positive emojis in 'statement', else 0
def has_pos_emoji(text):
    if not isinstance(text, str):
        return 0
    return 1 if any(emoji in text for emoji in pos_emojis) else 0

data['pos_emoji'] = data['statement'].apply(has_pos_emoji)
```

```
In [74]: # Same for negative emojis
def has_neg_emoji(text):
    if not isinstance(text, str):
        return 0
    return 1 if any(emoji in text for emoji in neg_emojis) else 0

data['neg_emoji'] = data['statement'].apply(has_neg_emoji)
```

```
In [75]: # Define a function to count rows with certain values
def count_rows(df, pos_val, neg_val):
    filtered_df = df[(df['pos_emoji'] == pos_val) & (df['neg_emoji'] == neg_val)]
    return filtered_df.groupby('status').size()

print("Number of Rows with POSITIVE Emojis:")
print(count_rows(data, 1, 0))
print("")
print("Number of Rows with NEGATIVE Emojis:")
print(count_rows(data, 0, 1))
print("")
print("Number of Rows with BOTH:")
print(count_rows(data, 1, 1))
```

```
Number of Rows with POSITIVE Emojis:
```

```
status
Anxiety           39
Bipolar          17
Depression       10
Normal            15
Personality disorder  22
Stress             7
dtype: int64
```

```
Number of Rows with NEGATIVE Emojis:
```

```
status
Anxiety           40
Bipolar          12
Depression        8
Personality disorder  14
Stress             13
dtype: int64
```

```
Number of Rows with BOTH:
```

```
status
Anxiety           2
Depression        2
Personality disorder  4
dtype: int64
```

We were correct in assuming that the use of negative emojis might be a more telling metric than the use of positive emojis. There are no 'Normal' status users that have used any negative emojis whatsoever. The results for positive emojis are more ambiguous. Let's drop that column.

(There are a surprisingly few amount of rows in our dataset that contain emojis. Our neg_emoji column is probably useless as well, but we'll keep it for right now).

```
In [76]: data.drop(columns=['pos_emoji'], inplace=True)
```

! and ?

The last notable special characters we might want to document prior to text cleaning are exclamation points ('!') and question marks ('?'). These are frequently used to convey emotional sentiment, especially distress when used repeatedly (ex. '!!!!!', '????!', '!?!?!?!?').

Let's add columns counting the number of times a person uses these in a 'statement'.

Normalization might not be useful here. Using special characters such as '!' and '?' more than once is already noteworthy, especially in brief statements made on social media. Identifying an excess count, regardless of statement length, is probably our best metric.

We'll try normalizing first, to see what the numbers look like.

```
In [77]: # Function to count characters excluding whitespace and calculate proportion
def calculate_char_ratio(text, char):
    if not isinstance(text, str):
        return 0

    # Remove whitespace and count non-whitespace characters
    non_whitespace_text = re.sub(r'\s+', '', text)
    total_characters = len(non_whitespace_text)

    # Count occurrences of 'char'
    char_count = text.count(char)

    # Calculate proportion
    if total_characters == 0:
        return 0
    ratio = char_count / total_characters

    return ratio

# Apply the function to calculate proportions for '!' and '?'
data['exclamation_ratio'] = data['statement'].apply(lambda x: calculate_char_ratio(x, '!'))
data['question_ratio'] = data['statement'].apply(lambda x: calculate_char_ratio(x, '?'))
```

```
In [78]: # Printing stats by 'status' group
exclamation_stats = data.groupby('status')['exclamation_ratio'].agg(['mean', 'min', 'max'])
print("Ratio of Exclamation Points by Status")
exclamation_stats
```

Ratio of Exclamation Points by Status

```
Out[78]:
```

	mean	min	max
status			
Anxiety	0.000642	0.0	0.157895
Bipolar	0.000562	0.0	0.056604
Depression	0.000236	0.0	0.138298
Normal	0.002042	0.0	0.740741
Personality disorder	0.000318	0.0	0.026157
Stress	0.000396	0.0	0.031008
Suicidal	0.000417	0.0	0.246154

```
In [79]: question_stats = data.groupby('status')['question_ratio'].agg(['mean', 'min', 'max'])
print("Ratio of Question Marks by Status")
question_stats
```

Ratio of Question Marks by Status

Out[79]:

mean min max

status

Anxiety	0.002883	0.0	0.100000
Bipolar	0.003458	0.0	0.090909
Depression	0.001452	0.0	0.166667
Normal	0.005444	0.0	0.466667
Personality disorder	0.003138	0.0	0.062500
Stress	0.001802	0.0	0.210811
Suicidal	0.001741	0.0	0.111111

Interesting. Just like we said, normalizing to account or statement length might not be ideal here. Our non-normalized counts might be easier to interpret.

Let's run the same for merely the count of each character, not the ratio proportional to the total number of characters in the statement.

In [80]:

```
# Function to count occurrences of '!'
def count_exclamation(text):
    if not isinstance(text, str):
        return 0
    return text.count('!')


# Add 'exclamtion_count' column to the DataFrame
data['exclamation_count'] = data['statement'].apply(count_exclamation)

# Printing stats by 'status' group
exccount_stats = data.groupby('status')['exclamation_count'].agg(['mean', 'min', 'max'])
print("Exclamation Point Count by Status")
exccount_stats
```

Exclamation Point Count by Status

Out[80]:

mean min max

status

Anxiety	0.301355	0	27
Bipolar	0.301479	0	9
Depression	0.110824	0	39
Normal	0.098635	0	20
Personality disorder	0.178771	0	26
Stress	0.176625	0	14
Suicidal	0.116906	0	58

```
In [81]: # Function to count occurrences of '!'
def count_question(text):
    if not isinstance(text, str):
        return 0
    return text.count('?')

# Add 'exc_count' column to the DataFrame
data['question_count'] = data['statement'].apply(count_question)

# Printing stats by 'status' group
questcount_stats = data.groupby('status')['question_count'].agg(['mean', 'min', 'max'])
print("Question Mark Count by Status")
questcount_stats
```

Question Mark Count by Status

Out[81]:

		mean	min	max
	status			
	Anxiety	1.056953	0	15
	Bipolar	1.565374	0	20
	Depression	0.652151	0	47
	Normal	0.170896	0	8
	Personality disorder	1.264804	0	22
	Stress	0.648059	0	39
	Suicidal	0.616389	0	34

Yes!! These numbers are far more useful, and give us interesting insight that aligns with what we know about personality disorders. It makes a lot of sense that 'Suicidal' would have the highest maximum value of exclamation points...since an excess of exclamation points typically indicates high levels of distress or excitement. It also makes sense that 'Normal' has the lowest maximum value with regards to question mark counts...for similar reasons.

We won't want to jump to concrete conclusions based on personal perception and prior knowledge...we'll try running our model and see how well these columns actually correlate to and are predictive of 'status'!

We'll keep these two new count features for now and drop our ratio columns.

In [82]:

```
data.drop(columns=['exclamation_ratio', 'question_ratio'], inplace=True)
```

In [83]:

```
# Checking value_counts to see how useful new columns are
data['question_count'].value_counts()
```

```
Out[83]: question_count
```

```
0      35919  
1      8260  
2      3689  
3      1627  
4      680  
5      354  
6      212  
7      116  
8      66  
9      39  
10     31  
11     24  
12     12  
14     9  
16     7  
13     7  
15     3  
17     2  
18     2  
21     2  
19     2  
22     2  
23     2  
20     2  
31     1  
47     1  
34     1  
39     1  
Name: count, dtype: int64
```

```
In [84]: data['exclamation_count'].value_counts()
```

```
Out[84]: exclamation_count
0      47503
1      2237
2      666
3      279
4      164
5      72
6      46
8      27
7      19
9      15
14     7
10     7
12     7
13     5
16     4
11     2
24     2
15     2
20     2
17     1
18     1
58     1
19     1
39     1
26     1
27     1
Name: count, dtype: int64
```

Text Cleaning

Now let's clean our text by making everything lower-case, removing special characters, etc.

Contractions

We could include a step in our cleaning that prepares our text to analyze contractions (ex. "don't", "aren't", "they're") as *bigrams*. However there aren't any contraction words that are particularly indicative of sentiment, especially with regards to personality disorders. We want to be detailed, without wasting time - so we will expand our contractions, but won't worry about providing any further analysis on them.

```
In [85]: # Expanding contractions
data['statement'] = data['statement'].apply(lambda x: contractions.fix(x))
```

```
In [86]: # Define function to remove emojis
def remove_emojis(text):
    # Regex pattern to match emojis (covers a wide range of emojis)
    emoji_pattern = re.compile(
        '['
        '\U0001F600-\U0001F64F'  # emoticons
        '\U0001F300-\U0001F5FF'  # symbols & pictographs
    )
```

```
'\U0001F680-\U0001F6FF' # transport & map symbols
'\U0001F700-\U0001F77F' # alchemical symbols
'\U0001F780-\U0001F7FF' # Geometric Shapes Extended
'\U0001F800-\U0001F8FF' # Supplemental Arrows-C
'\U0001F900-\U0001F9FF' # Supplemental Symbols and Pictographs
'\U0001FA00-\U0001FA6F' # Chess Symbols
'\U0001FA70-\U0001FAFF' # Symbols and Pictographs Extended-A
'\U00002700-\U000027BF' # Dingbats
'\U000024C2-\U0001F251'
']+',
flags=re.UNICODE
)
return emoji_pattern.sub(r'', text)

# Define function to preprocess text
def clean_text(text):
    if not isinstance(text, str):
        return ""
    text = text.lower()
    text = re.sub(r'https?://\S+|www\.\S+', '', text) # Remove links
    text = re.sub(r'[^w\s]', '', text) # Remove special char's & punct
    text = re.sub(r'\[|\]', '', text) # Remove square brackets
    text = re.sub(r'\d+', '', text) # Remove numbers
    text = re.sub(r'<.*?>', '', text) # Remove HTML tags
    text = re.sub(r'\n', '', text) # Remove newlines
    text = remove_emojis(text) # Apply emoji function from above
    return text

data['cleaned_statement'] = data['statement'].apply(lambda x: clean_text(x))
```

In [87]: `data.head()`

	statement	status	perc_upper_words	neg_emoji	exclamation_count	question_count
0	oh my gosh	Anxiety	0.0	0	0	
1	trouble sleeping, confused mind, restless hear...	Anxiety	0.0	0	0	
2	All wrong, back off dear, forward doubt. Stay ...	Anxiety	0.0	0	0	
3	I have shifted my focus to something else but ...	Anxiety	0.0	0	0	
4	I am restless and restless, it is been a month...	Anxiety	0.0	0	0	

Tokenization, Stop Words, Lemmatization

Let's tokenize our data, remove stopwords, then use lemmatization.

'Not'

Once we went ahead with tokenization, removal of stopwords, lemmatization, and printing a frequency distribution...some of the most-frequent tokens were words such as 'like', 'want' and 'know'...the meaning and interpretation of these words can be drastically altered depending on whether or not the word 'not' comes before them.

Let's alter our code to make sure '**not**' is excluded from the stop words we remove. Then we can create **bigrams** such as ('not', 'like') and ('not', 'want') that might help us contextualize our tokens more specifically. We'll wait to revisit our bigrams later on.

```
In [88]: # Creating list of stop words
stopwords_list = stopwords.words('english')

# Excluding 'not'
important_words = ['not']
```

```
# Initialize lemmatizer
lemmatizer = WordNetLemmatizer()

# Define function to tokenize, remove stopwords, and lemmatize
def tokenize_stopwords_lemmatize(text):
    tokens = word_tokenize(text)
    cleaned_tokens = [w for w in tokens if w not in stopwords_list or w in i]
    lemmatized_tokens = [lemmatizer.lemmatize(w) for w in cleaned_tokens]
    return lemmatized_tokens

# Tokenize
data['tokens'] = data['cleaned_statement'].apply(tokenize_stopwords_lemmatiz
```

In [89]: `data.head()`

	statement	status	perc_upper_words	neg_emoji	exclamation_count	question_count
0	oh my gosh	Anxiety	0.0	0	0	0
1	trouble sleeping, confused mind, restless hear...	Anxiety	0.0	0	0	0
2	All wrong, back off dear, forward doubt. Stay ...	Anxiety	0.0	0	0	0
3	I have shifted my focus to something else but ...	Anxiety	0.0	0	0	0
4	I am restless and restless, it is been a month...	Anxiety	0.0	0	0	0

Frequency Distribution by 'Status' Group

Let's plot our frequency distributions for each 'status' group, to see which words are most commonly used for each personality disorder.

We'll want to make use of common bigrams, to see how they compare to our single-word tokens with regards to frequency. We tried combining our tokens with common bigrams

and plotting the frequency of both for each status...however our bigrams didn't seem to have as common frequency distribution, and therefore didn't appear on our graphs.

We'll plot them separately for now, and keep the idea of combining their distributions in the back of our minds.

COMPUTATIONALLY EXPENSIVE: LEAVING AS MARKDOWN

Group by 'status' and aggregate tokens

```
status_groups = data.groupby('status')['tokens'].sum()
```

Function to get the most common words for each group

```
def get_most_common_words(tokens, num_common=10): freqdist = FreqDist(tokens)
return freqdist.most_common(num_common)
```

Apply frequency distribution calculation for each status

```
most_common_words_by_status = status_groups.apply(lambda x:
get_most_common_words(x))
```

Normalize Frequency Distribution Counts

Because our dataframe is so large, the code takes quite a while to run. Therefore we decided it's better to test things out, then go back and edit our previous code...to decrease the size and computational runtime of our notebook.

We tried normalizing our frequency distributions prior to plotting, since we saw in one of the first iterations of this code that they cover a wide range of numbers. We used *log normalization*, since our values cover a wide range and will also be susceptible to outliers.

However, our graphs with log normalization were not very useful to interpret. It might actually be better to get a sense of *how* varied the distribution counts are among words first, *then* normalize if that's helpful.

We will **not** normalize our counts for now.

```
In [90]: # Plotting
# Set up figure and axes
fig, axes = plt.subplots(nrows=7, figsize=(12, 12))

# Empty dict to hold words that have already been plotted and their colors
plotted_words_and_colors = {}

# Establish color palette to pull from
color_palette = sns.color_palette('cividis', n_colors=38)

# Creating a plot for each unique status
status_groups = data.groupby('status')
```

```
for idx, (status, status_df) in enumerate(status_groups):
    # Combine all tokens in the current status group
    all_tokens_in_status = status_df['tokens'].explode()

    # Calculate frequency distribution
    freq_dist = FreqDist(all_tokens_in_status)

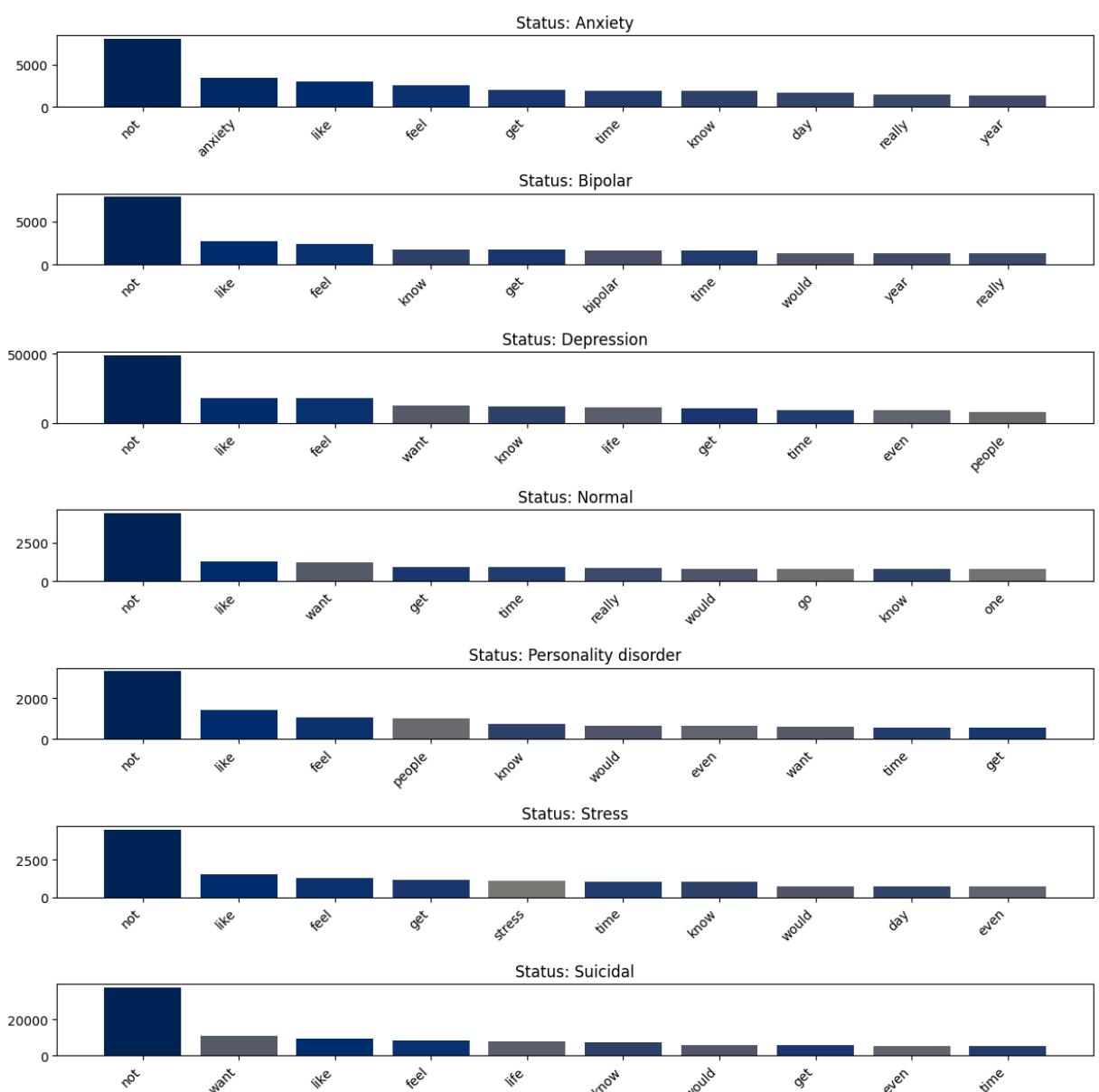
    # Get the top 10 tokens and their frequencies
    top_10 = freq_dist.most_common(10)
    tokens, counts = zip(*top_10)

    # Select appropriate colors, reusing colors if tokens repeat
    colors = []
    for token in tokens:
        if token not in plotted_words_and_colors:
            new_color = color_palette.pop(0)
            plotted_words_and_colors[token] = new_color
        colors.append(plotted_words_and_colors[token])

    # Select axes, plot data, set title
    ax = axes[idx]
    bars = ax.bar(tokens, counts, color=colors)
    ax.set_title(f"Status: {status}")

    # Rotate x-axis labels
    ax.set_xticks(tokens)
    ax.set_xticklabels(tokens, rotation=45, ha='right')

fig.tight_layout()
plt.show()
```



In [95]: #SAVING FOR PRESENTATION

```

import matplotlib.pyplot as plt
import seaborn as sns
from nltk.probability import FreqDist

# Define custom colors for specific status groups
status_colors = {
    'Bipolar': '#8dd8d3',      # Turquoise
    'Depression': '#69c6ef',   # Light Blue
    'Suicidal': '#d7e6a3',     # Light Green
    'Anxiety': '#df6fba',      # Pink
    'Normal': '#e4a959',       # Orange
    'Personality disorder': '#ed8975', # Peach
    'Stress': '#4a4ae7'        # Blue
}

# Set up figure and axes
fig, axes = plt.subplots(nrows=7, figsize=(12, 12))

```

```
# Create a plot for each unique status
status_groups = data.groupby('status')
for idx, (status, status_df) in enumerate(status_groups):
    # Combine all tokens in the current status group
    all_tokens_in_status = status_df['tokens'].explode()

    # Calculate frequency distribution
    freq_dist = FreqDist(all_tokens_in_status)

    # Get the top 10 tokens and their frequencies
    top_10 = freq_dist.most_common(10)
    tokens, counts = zip(*top_10)

    # Set color for all bars in this subplot based on the status
    color = status_colors.get(status, '#000000') # Default to black if stat

    # Plot data and set title
    ax = axes[idx]
    bars = ax.bar(tokens, counts, color=color)
    ax.set_title(f"Status: {status}", color='white')

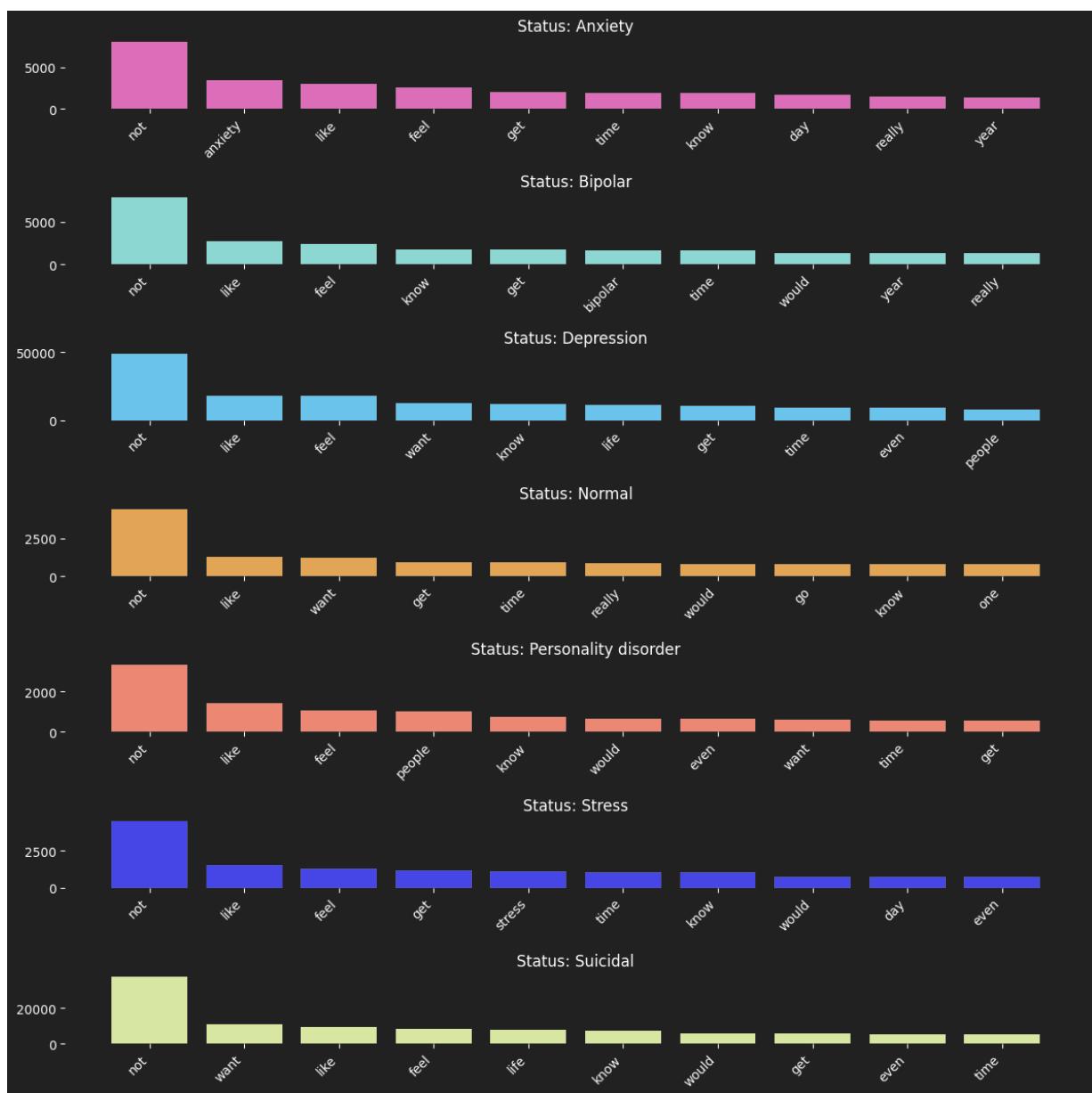
    # Rotate x-axis labels and set color
    ax.set_xticks(range(len(tokens)))
    ax.set_xticklabels(tokens, rotation=45, ha='right', color='white')
    ax.yaxis.label.set_color('white') # Set y-axis label color to white

    # Set transparent background for the axes and plot area
    ax.set_facecolor('none') # Transparent background for the plot area
    for spine in ax.spines.values():
        spine.set_visible(False) # Hide the spines (borders around the plot)
    ax.tick_params(axis='both', colors='white') # Set the ticks color to white

    # Set figure background color to transparent
    fig.patch.set_alpha(0.0)
    fig.patch.set_facecolor('none')

    fig.tight_layout()

    # Save the figure with a transparent background
    file_path = 'freq_words_status.png'
    plt.savefig(file_path, bbox_inches='tight', transparent=True)
    plt.show()
```



Dropping Common Tokens

We need to filter our tokens further. We have a lot of overlap of frequent tokens among our status groups which makes sense...tokens such as 'like' are often used as colloquial "filler-words" with no substantial meaning, tokens such as 'get', 'would', and 'know' could be useful...but they're also common words used in every-day speech that have no meaning out of context, especially when they also appear across all of our status groups.

We need to find a way to identify common tokens specific to a singular status group, that will help us better predict features unique to that status group alone.

Let's define a **threshold** the filters out tokens appearing a certain number of times across all 'statement's in our dataframe. We'll start with 70%, and go from there.

```
In [96]: # Combine all tokens across the DataFrame
all_tokens = [token for token_list in data['tokens'] for token in token_list]

# Calculate frequency distribution of all tokens
freq_dist = FreqDist(all_tokens)

# Set a frequency threshold of 70%
threshold = 0.7 * len(data)

# Identify common tokens
common_tokens = {token for token, count in freq_dist.items() if count > threshold}

# Function to remove common tokens
def remove_common_tokens(tokens):
    return [token for token in tokens if token not in common_tokens]

# Apply filtering to the tokens column
data['filtered_tokens'] = data['tokens'].apply(remove_common_tokens)
```

```
In [97]: # Group by 'status' and aggregate filtered tokens
status_groups_filtered = data.groupby('status')['filtered_tokens'].sum()

# Function to get the most common words for each group
def get_most_common_words(tokens, num_common=10):
    freqdist = FreqDist(tokens)
    return freqdist.most_common(num_common)

# Apply frequency distribution calculation for each status
most_common_words_by_status_filtered = status_groups_filtered.apply(lambda x: get_most_common_words(x))

# Set up figure and axes
fig, axes = plt.subplots(nrows=7, figsize=(12, 12))

# Empty dict to hold words that have already been plotted and their colors
plotted_words_and_colors = {}
# Establish color palette to pull from
color_palette = sns.color_palette('cividis', n_colors=38)

# Creating a plot for each unique status
for idx, (status, words_list) in enumerate(most_common_words_by_status_filtered):
    # Get the top words and their frequencies
    top_words = words_list
    words, counts = zip(*top_words)

    # Select appropriate colors, reusing colors if words repeat
    colors = []
    for word in words:
        if word not in plotted_words_and_colors:
            new_color = color_palette.pop(0)
            plotted_words_and_colors[word] = new_color
        colors.append(plotted_words_and_colors[word])

    # Select axes, plot data, set title
    ax = axes[idx]
    ax.bar(words, counts, color=colors)
```

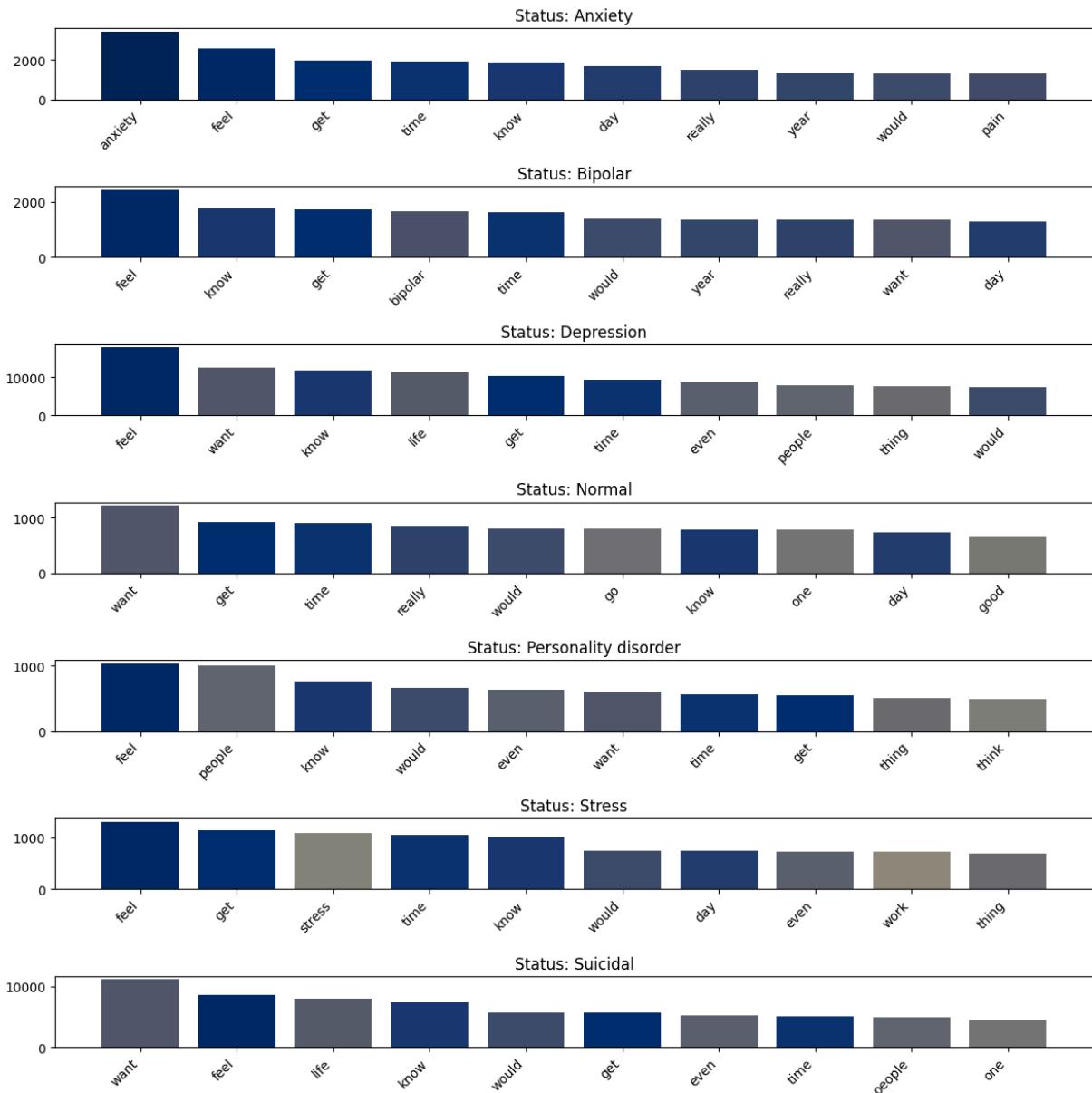
```

ax.set_title(f"Status: {status}")

# Rotate x-axis labels
ax.set_xticks(range(len(words)))
ax.set_xticklabels(words, rotation=45, ha='right')

fig.tight_layout()
plt.show()

```



This did seem to help a bit. For example, 'pain' was added to our top 10 most-common words for users with 'Anxiety', which is far more useful than the word 'like'.

However we still have a lot of overlap. Let's adjust our threshold further and try 60%. We can always go back and adjust.

```
In [98]: # Set a frequency threshold of 60%
threshold = 0.6 * len(data)
```

```
# Identify common tokens
common_tokens = {token for token, count in freq_dist.items() if count > threshold}

# Function to remove common tokens
def remove_common_tokens(tokens):
    return [token for token in tokens if token not in common_tokens]

# Apply filtering to the tokens column
data['filtered_tokens'] = data['tokens'].apply(remove_common_tokens)
```

```
In [99]: # Group by 'status' and aggregate filtered tokens
status_groups_filtered = data.groupby('status')['filtered_tokens'].sum()

# Function to get the most common words for each group
def get_most_common_words(tokens, num_common=10):
    freqdist = FreqDist(tokens)
    return freqdist.most_common(num_common)

# Apply frequency distribution calculation for each status
most_common_words_by_status_filtered = status_groups_filtered.apply(lambda x:
    get_most_common_words(x, num_common))

# Set up figure and axes
fig, axes = plt.subplots(nrows=7, figsize=(12, 12))

# Empty dict to hold words that have already been plotted and their colors
plotted_words_and_colors = {}

# Establish color palette to pull from
color_palette = sns.color_palette('cividis', n_colors=38)

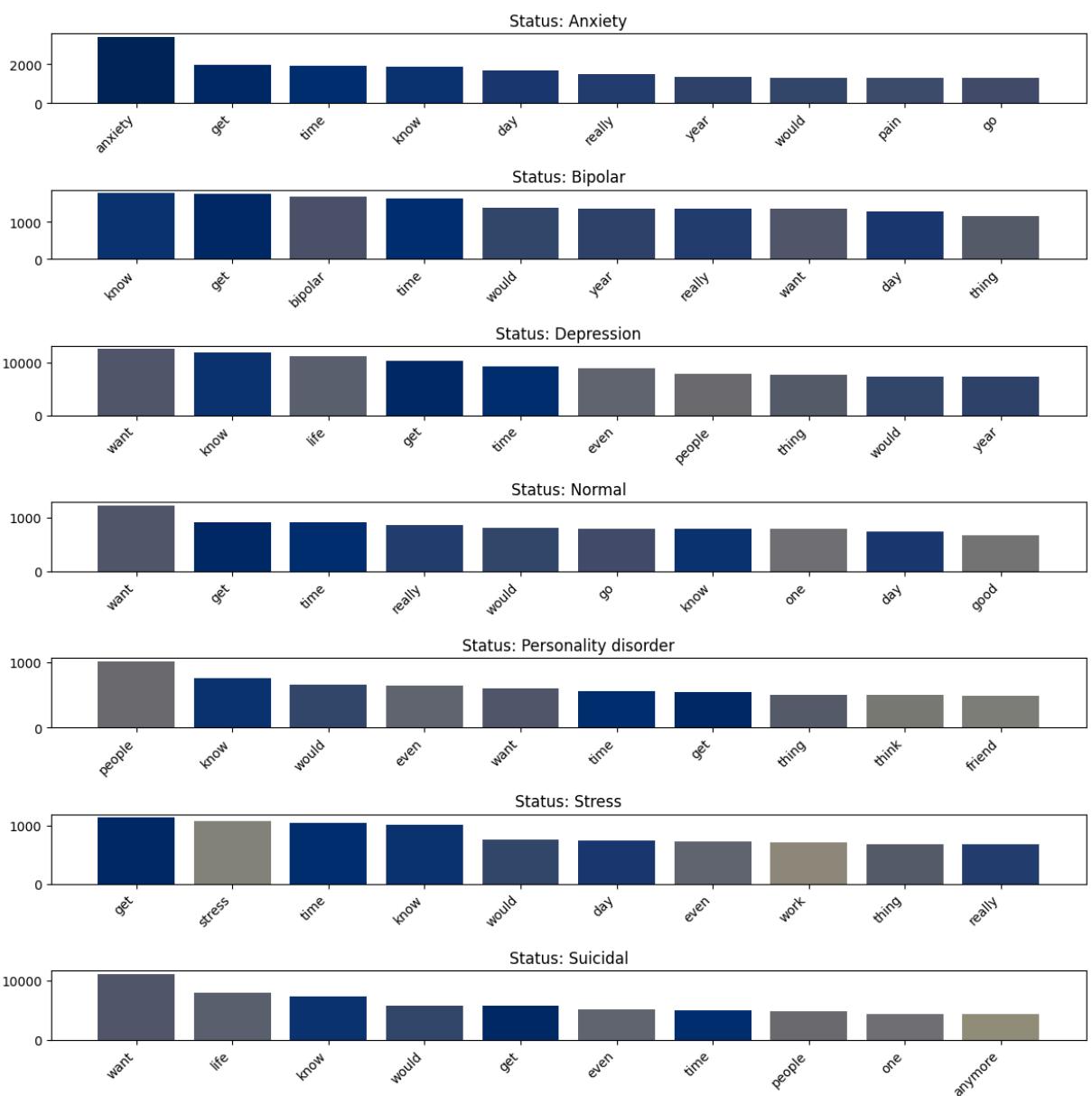
# Creating a plot for each unique status
for idx, (status, words_list) in enumerate(most_common_words_by_status_filtered):
    # Get the top words and their frequencies
    top_words = words_list
    words, counts = zip(*top_words)

    # Select appropriate colors, reusing colors if words repeat
    colors = []
    for word in words:
        if word not in plotted_words_and_colors:
            new_color = color_palette.pop(0)
            plotted_words_and_colors[word] = new_color
        colors.append(plotted_words_and_colors[word])

    # Select axes, plot data, set title
    ax = axes[idx]
    ax.bar(words, counts, color=colors)
    ax.set_title(f"Status: {status}")

    # Rotate x-axis labels
    ax.set_xticks(range(len(words)))
    ax.set_xticklabels(words, rotation=45, ha='right')

fig.tight_layout()
plt.show()
```



Super helpful! Notice how 'anymore' was added to 'Suicidal', and 'friend' was added to 'Personality disorder'.

Let's try it one more time, and go quite extreme. We tried filtering again to 50%, and it didn't make much of a difference. Let's set our threshold at 35% as an experiment and see what happens.

```
In [100]: # Set a frequency threshold of 35%
threshold = 0.35 * len(data)

# Identify common tokens
common_tokens = {token for token, count in freq_dist.items() if count > threshold}

# Function to remove common tokens
def remove_common_tokens(tokens):
    return [token for token in tokens if token not in common_tokens]
```

```
# Apply filtering to the tokens column
data['filtered_tokens'] = data['tokens'].apply(remove_common_tokens)
```

In [101...]

```
# Group by 'status' and aggregate filtered tokens
status_groups_filtered = data.groupby('status')['filtered_tokens'].sum()

# Function to get the most common words for each group
def get_most_common_words(tokens, num_common=10):
    freqdist = FreqDist(tokens)
    return freqdist.most_common(num_common)

# Apply frequency distribution calculation for each status
most_common_words_by_status_filtered = status_groups_filtered.apply(lambda x:
    get_most_common_words(x, num_common=10))

# Set up figure and axes
fig, axes = plt.subplots(nrows=7, figsize=(12, 12))

# Empty dict to hold words that have already been plotted and their colors
plotted_words_and_colors = {}

# Establish color palette to pull from
color_palette = sns.color_palette('cividis', n_colors=38)

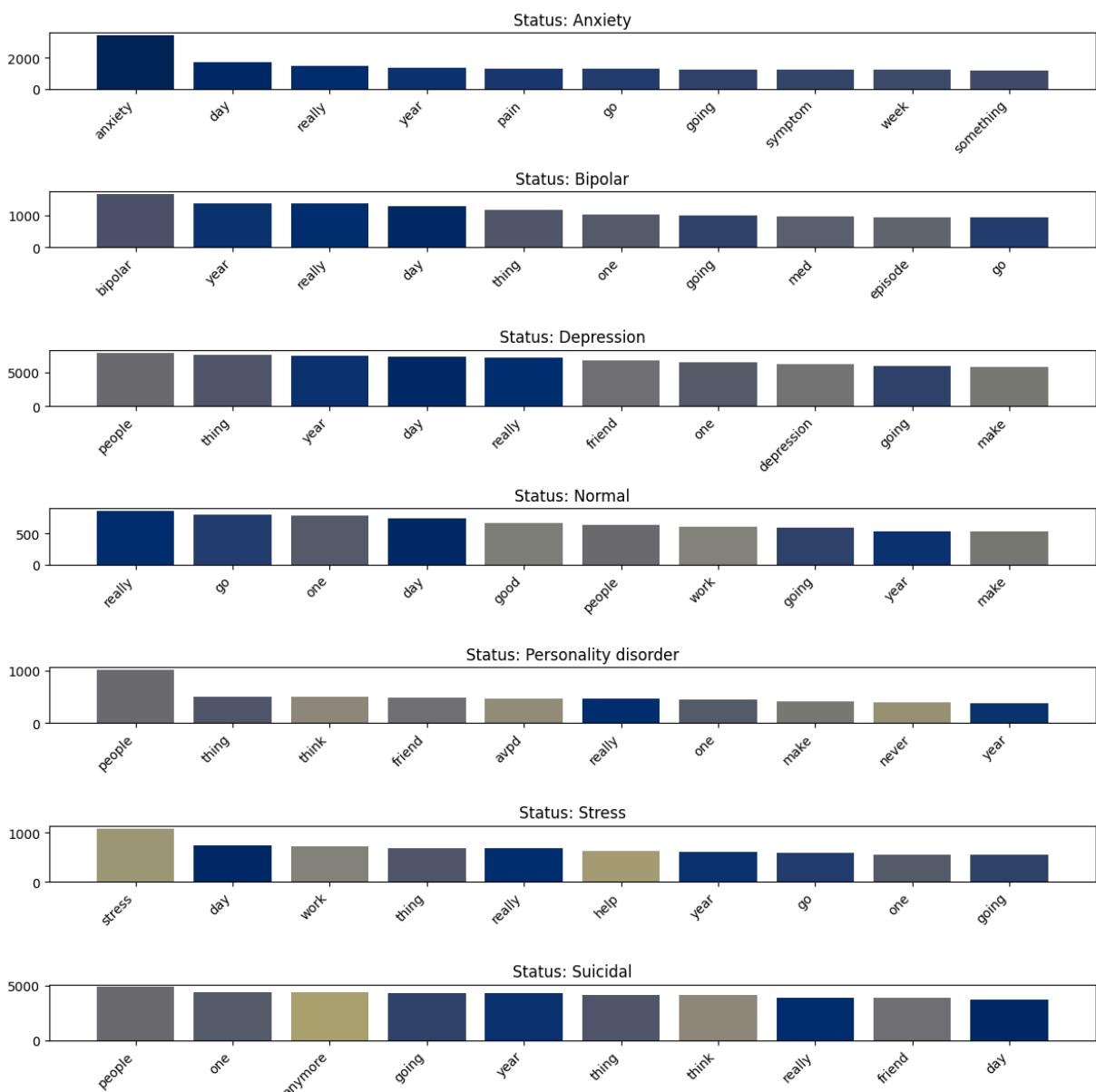
# Creating a plot for each unique status
for idx, (status, words_list) in enumerate(most_common_words_by_status_filtered):
    # Get the top words and their frequencies
    top_words = words_list
    words, counts = zip(*top_words)

    # Select appropriate colors, reusing colors if words repeat
    colors = []
    for word in words:
        if word not in plotted_words_and_colors:
            new_color = color_palette.pop(0)
            plotted_words_and_colors[word] = new_color
        colors.append(plotted_words_and_colors[word])

    # Select axes, plot data, set title
    ax = axes[idx]
    ax.bar(words, counts, color=colors)
    ax.set_title(f"Status: {status}")

    # Rotate x-axis labels
    ax.set_xticks(range(len(words)))
    ax.set_xticklabels(words, rotation=45, ha='right')

fig.tight_layout()
plt.show()
```



In [102]:

```

import matplotlib.pyplot as plt
import seaborn as sns
from nltk.probability import FreqDist

# Define custom colors for specific status groups
status_colors = {
    'Bipolar': '#8dd8d3',      # Turquoise
    'Depression': '#69c6ef',   # Light Blue
    'Suicidal': '#d7e6a3',     # Light Green
    'Anxiety': '#df6fba',      # Pink
    'Normal': '#e4a959',       # Orange
    'Personality disorder': '#ed8975', # Peach
    'Stress': '#4a4ae7'        # Blue
}

# Group by 'status' and aggregate filtered tokens
status_groups_filtered = data.groupby('status')['filtered_tokens'].sum()

# Function to get the most common words for each group

```

```
def get_most_common_words(tokens, num_common=10):
    freqdist = FreqDist(tokens)
    return freqdist.most_common(num_common)

# Apply frequency distribution calculation for each status
most_common_words_by_status_filtered = status_groups_filtered.apply(lambda x:
    # Set up figure and axes
    fig, axes = plt.subplots(nrows=7, figsize=(12, 12))

    # Creating a plot for each unique status
    for idx, (status, words_list) in enumerate(most_common_words_by_status_filtered):
        # Get the top words and their frequencies
        top_words = words_list
        words, counts = zip(*top_words)

        # Set color for all bars in this subplot based on the status
        color = status_colors.get(status, '#000000') # Default to black if status
            # Plot data and set title
            ax = axes[idx]
            bars = ax.bar(words, counts, color=color)
            ax.set_title(f"Status: {status}", color='white')

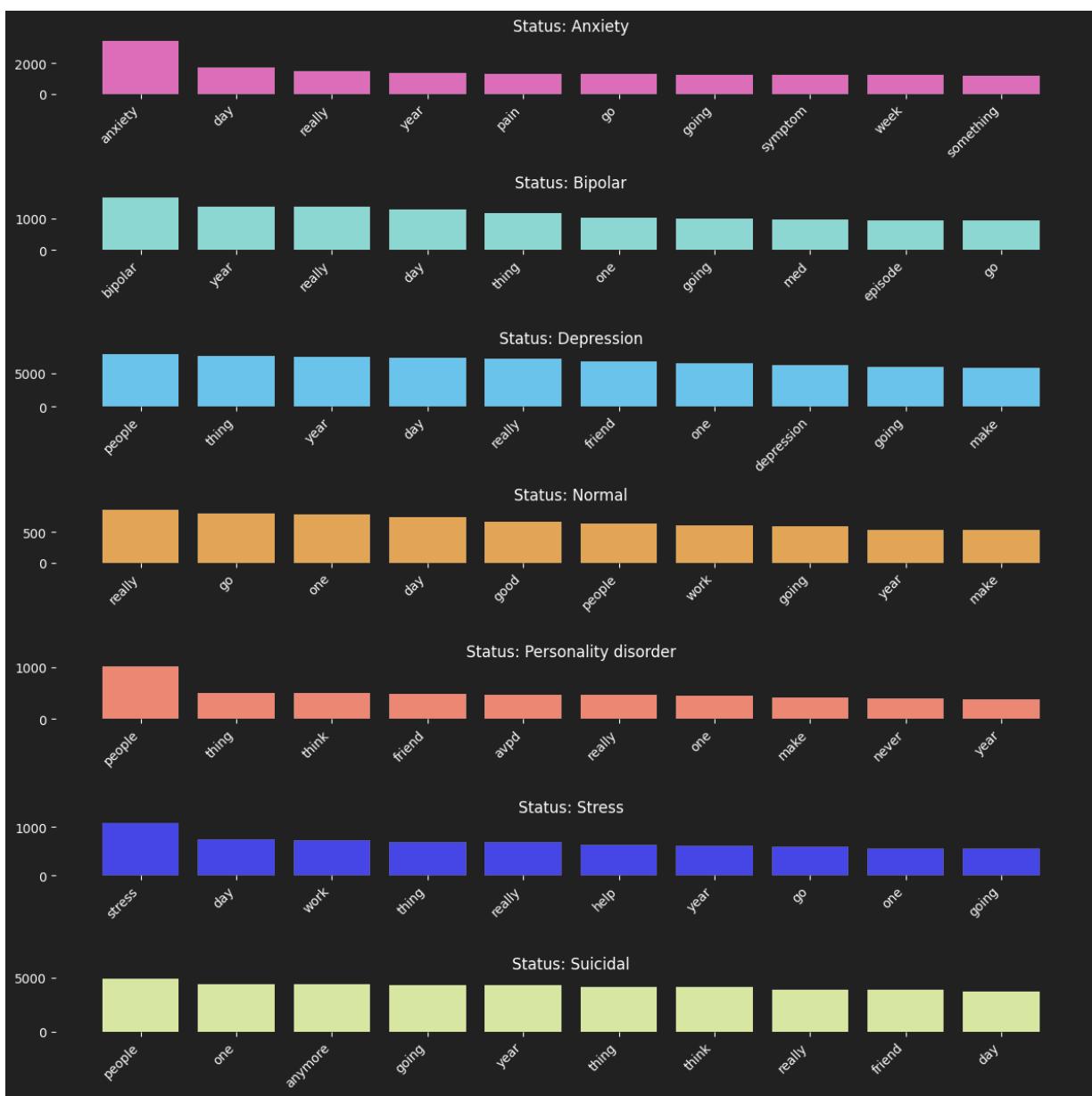
            # Rotate x-axis labels and set color
            ax.set_xticks(range(len(words)))
            ax.set_xticklabels(words, rotation=45, ha='right', color='white')
            ax.yaxis.label.set_color('white') # Set y-axis label color to white

            # Set transparent background for the axes and plot area
            ax.set_facecolor('none') # Transparent background for the plot area
            for spine in ax.spines.values():
                spine.set_visible(False) # Hide the spines (borders around the plot
            ax.tick_params(axis='both', colors='white') # Set the ticks color to white

            # Set figure background color to transparent
            fig.patch.set_alpha(0.0)
            fig.patch.set_facecolor('none')

            fig.tight_layout()

    # Save the figure with a transparent background
    file_path = 'freq_words_thresh_35.png'
    plt.savefig(file_path, bbox_inches='tight', transparent=True)
    plt.show()
```



Super interesting and potentially very useful!

Bigrams

Let's revisit bigrams. We kept the stop word 'not' in our tokens, in order to see whether it's particularly useful in defining the sentiment of a following word in a bigram.

We can furthermore calculate the **Raw Frequency** of bigrams with other tokens in our dataset as well as the **Pointwise Mutual Information Score** between bigram pairs in our dataset. Mutual Information Score essentially tells us the mutual dependence between two words.

We can then determine whether any other bigram pairs are important.

We'll perform this on our original 'tokens' column, rather than the lists of tokens filtered with thresholds. There may be some very common words that become unique to a particular status, when paired with another token in a bigram!

Filtering out Noise

It might be helpful to filter out pairs of words that occur frequently across our entire dataset, regardless of 'status' group. The same way we eliminate stop words from our tokens prior to drawing any meaningful conclusions.

By first filtering out the most frequent bigrams in our dataset, we can help eliminate unnecessary noise.

To improve computational efficiency, we ran some code already and have some interesting takeaways that we can apply here prior to re-running...which will make our results more useful and improve run-times.

We ran code filtering through the raw frequencies and PMI scores for the top 50 bigrams in the entire dataset, as well as by status group

'Not' and 'Like'

Our main takeaways are regarding the words **not** and **like**.

Like

As we mentioned earlier, **like** is a pretty useless filler word that's only important if it's conveying some type of sentiment in its verb form ("to like"). Once we looked at the most common bigrams, we found this was *not* the case in our dataset

Ex. If the bigrams had included phrases such as ('i', 'like') or ('not', 'like') or ('you', 'like') in relation to specific 'status' groups, they might have been useful in predicting a person's mental state). These pairings did **not** appear in our results, so we can safely *drop the token 'like' from our entire dataframe*.

```
In [103]: # Remove 'like' from tokens, now that we've made bigrams
def remove_like(tokens):
    return [token for token in tokens if token != 'like']

# Apply function
data['tokens'] = data['tokens'].apply(remove_like)
```

Not

Our results related to **not** were more nuanced. Some bigrams containing 'not' definitely *do* convey interesting and informative sentiments...however they appear so frequently

across our entire dataframe, *and* all or most of our 'status' groups...that they really will not be helpful predicting specific personality disorders.

We carefully filtered through the raw frequencies and PMI scores for the top 50 bigrams in the entire dataset, as well as by status group...and we retrieved certain 'not' bigrams that we think will be useful to include. Some were found frequently across multiple 'status' groups in our dataset, that might still be informative when breaking down their frequency counts by status in a further filtered dataframe...others were meaningful bigrams found frequently in individual 'status' groups alone.

There's a list of our selects below.

NOT BIGRAMS: LEAVING AS MARKDOWN

We ran this and it was super unhelpful. We ended up dropping 'not', because it only added noise. We'll leave this code below as markdowns for reference, and we will remove 'not' from the tokens in our dataframe.

List of bigrams to keep

```
bigrams_to_keep = [('could', 'not'), ('not', 'stop'), ('not', 'able'), ('not', 'help'), ('not', 'take'), ('not', 'seem'), ('not', 'work'), ('life', 'not'), ('not', 'see'), ('people', 'not'), ('not', 'anything'), ('not', 'understand'), ('not', 'good'), ('not', 'going'), ('not', 'sleep'), ('not', 'afford'), ('anymore', 'not'), ('die', 'not'), ('not', 'anymore'), ('would', 'not')]
```

We're running into problems with accidentally creating or removing duplicate tokens, which might be a problem when we get to count vectorization shortly.

We'll create additional lists below, which will help us determine whether to drop 'not' as a token from our 'tokens' list...while preserving the structure of all other tokens in the row.

```
In [104]: # Remove 'not' from tokens, now that we've made bigrams
def remove_not(tokens):
    return [token for token in tokens if token != 'not']

# Apply function
data['tokens'] = data['tokens'].apply(remove_not)
```

words where 'not' is first word in bigram selects

```
not_first = ['stop', 'able', 'help', 'take', 'seem', 'work', 'see', 'anything', 'understand',
'good', 'going', 'sleep', 'afford', 'anymore']
```

import nltk from nltk.util import bigrams

Define function to remove 'not' tokens based on conditions

```

def remove_not(tokens, not_first): # Convert tokens to bigrams
    bigram_list = list(bigrams(tokens))

    # Determine indices of 'not' to remove
    remove_not_indices = set()

    # Identify indices where 'not' should be removed
    for i, bigram in enumerate(bigram_list):
        if bigram[0] == 'not' and bigram[1] not in not_first:
            # If 'not' is the first word and the second word is
            # NOT in 'not_first', mark 'not' for removal
            remove_not_indices.add(i)
        elif bigram[0] == 'not' and bigram[1] in not_first:
            # If 'not' is the first word and the second word is
            # in 'not_first', keep 'not'
            # Remove index from removal set if it exists
            remove_not_indices.discard(i)

    # Create a new list of tokens excluding the 'not' tokens that
    # are to be removed
    filtered_tokens = [token for i, token in enumerate(tokens) if
not (token == 'not' and (i-1) in remove_not_indices)]

    return filtered_tokens

```

Apply the function to the DataFrame

```
data['tokens'] = data['tokens'].apply(lambda x: remove_not(x, not_first))
```

Verify the changes

```
data.head()
```

Combining tokens and bigrams

```

def combine_tokens_bigrams(row): tokens = row['tokens'] bigrams =
row['bigrams_following_not'] return tokens + bigrams

```

```
data['combined_tokens_bigrams'] = data.apply(combine_tokens_bigrams, axis=1)
```

Group by status and aggregate

```
status_groups = data.groupby('status')['combined_tokens_bigrams'].sum()
```

Now we run our code calculating the frequency and PMI scores for bigrams by 'status', with our filtered data.

Raw Frequency by Status

In [105...]

```

# Define a function to score bigrams
def score_bigrams(tokens):
    bigram_finder = BigramCollocationFinder.from_words(tokens)

```

```
bigrams_scored = bigram_finder.score_ngrams(BigramAssocMeasures.raw_freq)
return bigrams_scored[:50] # Return top 30 scored bigrams

# Group by 'status' and apply the scoring function
bigrams_by_status = data.groupby('status')['tokens'].apply(lambda x: score_bigrams(x))

# Print or access the results for each status
for status, bigrams in bigrams_by_status.items():
    print(f"Status: {status}")
    for idx, (bigram, score) in enumerate(bigrams, 1):
        print(f"{idx}. {bigram}: {score}")
    print()
```

Status: Anxiety

1. ('health', 'anxiety'): 0.0028422468886646485
2. ('panic', 'attack'): 0.0018710057181298351
3. ('anyone', 'else'): 0.0014043054154053145
4. ('year', 'old'): 0.000853514968045745
5. ('year', 'ago'): 0.0008030608812647158
6. ('month', 'ago'): 0.000786242852337706
7. ('even', 'though'): 0.0007820383451059536
8. ('heart', 'attack'): 0.0007736293306424487
9. ('came', 'back'): 0.0007399932727884292
10. ('lymph', 'node'): 0.0007063572149344097
11. ('heart', 'rate'): 0.0007021527077026573
12. ('week', 'ago'): 0.0007021527077026573
13. ('chest', 'pain'): 0.0006643121426168853
14. ('go', 'away'): 0.0006264715775311134
15. ('every', 'time'): 0.0005760174907500841
16. ('blood', 'pressure'): 0.0005676084762865792
17. ('blood', 'test'): 0.0005591994618230744
18. ('last', 'year'): 0.000546585940127817
19. ('first', 'time'): 0.0005339724184325597
20. ('every', 'day'): 0.0005171543895055499
21. ('left', 'side'): 0.0004877228388832829
22. ('last', 'week'): 0.000479313824419778
23. ('really', 'bad'): 0.000479313824419778
24. ('something', 'wrong'): 0.0004582912882610158
25. ('going', 'die'): 0.0004540867810292634
26. ('last', 'night'): 0.0004330642448705012
27. ('day', 'ago'): 0.0004246552304069963
28. ('brain', 'tumor'): 0.0004036326942482341
29. ('side', 'effect'): 0.0004036326942482341
30. ('go', 'doctor'): 0.0003994281870164817
31. ('went', 'doctor'): 0.0003994281870164817
32. ('ever', 'since'): 0.0003952236797847292
33. ('come', 'back'): 0.000374201143625967
34. ('past', 'month'): 0.000374201143625967
35. ('long', 'time'): 0.0003699966363942146
36. ('two', 'week'): 0.0003699966363942146
37. ('go', 'back'): 0.00035738311469895727
38. ('doctor', 'said'): 0.00034476959300369994
39. ('make', 'feel'): 0.0003405650857719475
40. ('blood', 'work'): 0.00033215607130844267
41. ('last', 'month'): 0.00033215607130844267
42. ('past', 'week'): 0.00033215607130844267
43. ('right', 'side'): 0.00031954254961318534
44. ('ct', 'scan'): 0.0003153380423814329
45. ('anxiety', 'attack'): 0.0003113353514968043
46. ('past', 'year'): 0.0003113353514968043
47. ('pretty', 'much'): 0.000306929027917928
48. ('back', 'normal'): 0.00029852001345442316
49. ('back', 'pain'): 0.00029852001345442316
50. ('something', 'else'): 0.0002943155062226707

Status: Bipolar

1. ('anyone', 'else'): 0.0014429296447407702
2. ('manic', 'episode'): 0.0013334660165190567
3. ('bipolar', 'disorder'): 0.0011991242909742262

4. ('diagnosed', 'bipolar'): 0.0009901482734600458
5. ('mental', 'health'): 0.0009155139814906956
6. ('side', 'effect'): 0.0008956115036322022
7. ('depressive', 'episode'): 0.0008806846452383321
8. ('year', 'ago'): 0.0007314160612996319
9. ('first', 'time'): 0.0006020499552194248
10. ('even', 'though'): 0.000577171857896308
11. ('last', 'year'): 0.0005672206189670614
12. ('month', 'ago'): 0.0005423425216439446
13. ('mental', 'illness'): 0.000517464424320828
14. ('year', 'old'): 0.0005075131853915812
15. ('make', 'feel'): 0.00048761070753308787
16. ('get', 'better'): 0.00045278137128072445
17. ('mood', 'swing'): 0.0004129764155637377
18. ('every', 'day'): 0.000403025176634491
19. ('last', 'week'): 0.0003980495571698676
20. ('mixed', 'episode'): 0.0003980495571698676
21. ('long', 'time'): 0.00038809831824062095
22. ('last', 'night'): 0.00038312269877599763
23. ('two', 'week'): 0.00037814707931137426
24. ('panic', 'attack'): 0.00035326898198825753
25. ('make', 'sense'): 0.00034829336252363417
26. ('go', 'back'): 0.00034331774305901085
27. ('want', 'go'): 0.00033336650412976417
28. ('want', 'get'): 0.0003283908846651408
29. ('really', 'want'): 0.0003234152652005175
30. ('week', 'ago'): 0.0003234152652005175
31. ('one', 'day'): 0.0003035127873420241
32. ('recently', 'diagnosed'): 0.0003035127873420241
33. ('bipolar', 'ii'): 0.00029853716787740076
34. ('feel', 'better'): 0.0002935615484127774
35. ('mood', 'stabilizer'): 0.000288585928948154
36. ('pretty', 'much'): 0.000288585928948154
37. ('really', 'bad'): 0.000288585928948154
38. ('two', 'year'): 0.0002836103094835307
39. ('taking', 'med'): 0.00027863469001890734
40. ('go', 'away'): 0.00027365907055428403
41. ('started', 'taking'): 0.00026868345108966066
42. ('hypomanic', 'episode'): 0.0002637078316250373
43. ('day', 'ago'): 0.000258732212160414
44. ('every', 'time'): 0.0002537565926957906
45. ('got', 'diagnosed'): 0.0002487809732311673
46. ('hard', 'time'): 0.0002487809732311673
47. ('last', 'month'): 0.0002487809732311673
48. ('take', 'med'): 0.0002487809732311673
49. ('want', 'know'): 0.0002388297343019206
50. ('rapid', 'cycling'): 0.00023385411483729725

Status: Depression

1. ('get', 'better'): 0.0008258946040631805
2. ('mental', 'health'): 0.0008037970895731957
3. ('make', 'feel'): 0.0008019556300323637
4. ('even', 'though'): 0.0007642057094453063
5. ('every', 'day'): 0.0007421081949553216
6. ('year', 'old'): 0.000722772869776585
7. ('year', 'ago'): 0.0006214925950308215

8. ('want', 'die'): 0.0005911085126070924
9. ('last', 'year'): 0.0005395476454637946
10. ('anyone', 'else'): 0.0005174501309738098
11. ('wish', 'could'): 0.0004824623996980006
12. ('even', 'know'): 0.00045115758750385556
13. ('first', 'time'): 0.0004502368577334395
14. ('high', 'school'): 0.0004502368577334395
15. ('every', 'time'): 0.0004474746684221914
16. ('really', 'want'): 0.0004327429920955349
17. ('take', 'anymore'): 0.00041616985622804634
18. ('long', 'time'): 0.00041248693714638225
19. ('one', 'day'): 0.00041064547760555017
20. ('go', 'back'): 0.0004005174501309738
21. ('best', 'friend'): 0.00039959672036055777
22. ('panic', 'attack'): 0.0003977552608197257
23. ('month', 'ago'): 0.0003848650440339013
24. ('depression', 'anxiety'): 0.0003802613951818211
25. ('suicidal', 'thought'): 0.0003710540974776608
26. ('feel', 'better'): 0.0003701333677072448
27. ('want', 'go'): 0.00036921263793682874
28. ('need', 'help'): 0.00035632242115100427
29. ('want', 'live'): 0.0003397492852835157
30. ('know', 'anymore'): 0.0003250176089568592
31. ('every', 'single'): 0.00032133468987519505
32. ('want', 'get'): 0.00030844447308937063
33. ('know', 'feel'): 0.00029371279676271414
34. ('feel', 'bad'): 0.00029002987768105
35. ('past', 'year'): 0.00029002987768105
36. ('want', 'feel'): 0.00028910914791063395
37. ('anymore', 'take'): 0.0002845054990585538
38. ('feel', 'way'): 0.0002835847692881378
39. ('life', 'feel'): 0.00027898112043605764
40. ('keep', 'going'): 0.00027069455250231336
41. ('friend', 'family'): 0.0002697738227318973
42. ('really', 'know'): 0.00026609090365023317
43. ('pretty', 'much'): 0.00026148725479815303
44. ('really', 'bad'): 0.0002587250654869049
45. ('ever', 'since'): 0.0002578043357164889
46. ('go', 'away'): 0.0002578043357164889
47. ('still', 'feel'): 0.0002559628761756568
48. ('video', 'game'): 0.0002513592273235767
49. ('take', 'care'): 0.00024215192961941635
50. ('feel', 'alone'): 0.00024031047007858427

Status: Normal

1. ('let', 'u'): 0.0012639668335102886
2. ('even', 'though'): 0.0008883881172672315
3. ('want', 'go'): 0.0006572627534253501
4. ('last', 'night'): 0.000628372082945115
5. ('good', 'morning'): 0.0004983640657840566
6. ('every', 'day'): 0.0004550280600637039
7. ('year', 'old'): 0.00044780539244364514
8. ('first', 'time'): 0.00044058272482358637
9. ('oh', 'god'): 0.00044058272482358637
10. ('long', 'time'): 0.0004044693867232924
11. ('really', 'want'): 0.0003900240514831748

12. ('year', 'ago'): 0.00035391071338288085
13. ('gon', 'na'): 0.000346688045762822
14. ('go', 'back'): 0.00033946537814276325
15. ('wan', 'na'): 0.00033946537814276325
16. ('come', 'back'): 0.00031057470766252806
17. ('mental', 'health'): 0.00029612937242241047
18. ('high', 'school'): 0.0002816840371822929
19. ('best', 'friend'): 0.00026723870194217533
20. ('let', 'know'): 0.0002600160343221165
21. ('really', 'good'): 0.0002600160343221165
22. ('http', 'twitpic'): 0.00024557069908199896
23. ('take', 'care'): 0.00024557069908199896
24. ('twitpic', 'com'): 0.00024557069908199896
25. ('â', 'â'): 0.00024557069908199896
26. ('last', 'year'): 0.00023112536384188137
27. ('month', 'ago'): 0.00023112536384188137
28. ('every', 'time'): 0.00022390269622182257
29. ('want', 'see'): 0.00022390269622182257
30. ('anyone', 'else'): 0.00021668002860176377
31. ('go', 'home'): 0.00021668002860176377
32. ('someone', 'else'): 0.00020945736098170497
33. ('want', 'get'): 0.00020945736098170497
34. ('last', 'week'): 0.0002022346933616462
35. ('next', 'week'): 0.0002022346933616462
36. ('want', 'buy'): 0.0002022346933616462
37. ('would', 'love'): 0.0002022346933616462
38. ('ó', 'ó'): 0.0002022346933616462
39. ('get', 'back'): 0.0001950120257415874
40. ('one', 'day'): 0.0001950120257415874
41. ('looking', 'forward'): 0.0001877893581215286
42. ('wish', 'could'): 0.0001877893581215286
43. ('good', 'luck'): 0.000173344022881411
44. ('last', 'time'): 0.000173344022881411
45. ('sorry', 'hear'): 0.000173344022881411
46. ('tinyurl', 'com'): 0.000173344022881411
47. ('u', 'mutual'): 0.000173344022881411
48. ('anyone', 'want'): 0.00016612135526135224
49. ('go', 'school'): 0.00016612135526135224
50. ('help', 'rtlike'): 0.00016612135526135224

Status: Personality disorder

1. ('anyone', 'else'): 0.001561045470819907
2. ('make', 'feel'): 0.0010025062656641604
3. ('social', 'anxiety'): 0.0009881847475832438
4. ('even', 'though'): 0.0009595417114214107
5. ('mental', 'health'): 0.0007590404582885786
6. ('year', 'old'): 0.0005871822413175796
7. ('even', 'know'): 0.0004869316147511636
8. ('personality', 'disorder'): 0.0004869316147511636
9. ('social', 'medium'): 0.00047261009667024705
10. ('year', 'ago'): 0.00047261009667024705
11. ('everyone', 'else'): 0.00045828857858933047
12. ('wish', 'could'): 0.00045828857858933047
13. ('social', 'interaction'): 0.0004296455424274973
14. ('every', 'time'): 0.00041532402434658075
15. ('first', 'time'): 0.00041532402434658075

16. ('last', 'year'): 0.00041532402434658075
17. ('never', 'really'): 0.00041532402434658075
18. ('really', 'want'): 0.00041532402434658075
19. ('go', 'back'): 0.0004010025062656642
20. ('high', 'school'): 0.0003866809881847476
21. ('make', 'friend'): 0.0003866809881847476
22. ('social', 'situation'): 0.00037235947010383103
23. ('get', 'better'): 0.00034371643394199783
24. ('pretty', 'much'): 0.00034371643394199783
25. ('talk', 'people'): 0.00034371643394199783
26. ('diagnosed', 'avpd'): 0.00032939491586108125
27. ('else', 'feel'): 0.00032939491586108125
28. ('know', 'people'): 0.00032939491586108125
29. ('one', 'thing'): 0.00032939491586108125
30. ('people', 'know'): 0.00032939491586108125
31. ('avoidant', 'personality'): 0.0003150733977801647
32. ('many', 'people'): 0.0003150733977801647
33. ('something', 'wrong'): 0.0003150733977801647
34. ('want', 'get'): 0.0003150733977801647
35. ('feel', 'bad'): 0.0003007518796992481
36. ('good', 'enough'): 0.0003007518796992481
37. ('leave', 'house'): 0.0003007518796992481
38. ('social', 'skill'): 0.0003007518796992481
39. ('long', 'time'): 0.00028643036161833153
40. ('people', 'feel'): 0.00028643036161833153
41. ('whole', 'life'): 0.00028643036161833153
42. ('best', 'friend'): 0.00027210884353741496
43. ('close', 'friend'): 0.00027210884353741496
44. ('every', 'day'): 0.00027210884353741496
45. ('many', 'time'): 0.00027210884353741496
46. ('one', 'day'): 0.00027210884353741496
47. ('people', 'avpd'): 0.00027210884353741496
48. ('people', 'get'): 0.00027210884353741496
49. ('time', 'feel'): 0.00027210884353741496
50. ('around', 'people'): 0.0002577873254564984

Status: Stress

1. ('panic', 'attack'): 0.0008889984244483368
2. ('anyone', 'else'): 0.0007305630616753659
3. ('even', 'though'): 0.0007041571678798708
4. ('year', 'old'): 0.0006689493094858772
5. ('month', 'ago'): 0.0005369198405084014
6. ('make', 'feel'): 0.000501711982114408
7. ('every', 'day'): 0.0004929100175159096
8. ('need', 'help'): 0.000457702159121916
9. ('mental', 'health'): 0.0004489001945234176
10. ('year', 'ago'): 0.00044009822992491926
11. ('last', 'year'): 0.00037848447773543054
12. ('every', 'time'): 0.00036088054853843375
13. ('first', 'time'): 0.00036088054853843375
14. ('go', 'back'): 0.00036088054853843375
15. ('new', 'job'): 0.0003520785839399354
16. ('stress', 'anxiety'): 0.000343276619341437
17. ('high', 'school'): 0.00032567269014444023
18. ('feel', 'bad'): 0.00031687072554594187
19. ('get', 'better'): 0.0002992667963489451

20. ('really', 'bad'): 0.0002992667963489451
21. ('feel', 'stressed'): 0.0002816628671519483
22. ('last', 'night'): 0.0002816628671519483
23. ('even', 'know'): 0.00027286090255344993
24. ('long', 'time'): 0.00027286090255344993
25. ('really', 'know'): 0.00027286090255344993
26. ('stress', 'management'): 0.00027286090255344993
27. ('take', 'care'): 0.00027286090255344993
28. ('last', 'week'): 0.00026405893795495157
29. ('really', 'want'): 0.00026405893795495157
30. ('one', 'day'): 0.00025525697335645315
31. ('best', 'friend'): 0.0002464550087579548
32. ('get', 'back'): 0.0002464550087579548
33. ('get', 'job'): 0.0002464550087579548
34. ('manage', 'stress'): 0.0002464550087579548
35. ('pretty', 'much'): 0.0002464550087579548
36. ('health', 'issue'): 0.0002376530441594564
37. ('still', 'feel'): 0.0002376530441594564
38. ('know', 'feel'): 0.000228851079560958
39. ('next', 'day'): 0.000228851079560958
40. ('past', 'month'): 0.000228851079560958
41. ('deal', 'stress'): 0.00022004911496245963
42. ('depression', 'anxiety'): 0.00022004911496245963
43. ('family', 'member'): 0.00022004911496245963
44. ('many', 'thing'): 0.00022004911496245963
45. ('stress', 'level'): 0.00022004911496245963
46. ('two', 'week'): 0.00022004911496245963
47. ('anxiety', 'disorder'): 0.00021124715036396124
48. ('came', 'back'): 0.00021124715036396124
49. ('day', 'day'): 0.00021124715036396124
50. ('know', 'going'): 0.00021124715036396124

Status: Suicidal

1. ('want', 'die'): 0.0022462648432937512
2. ('take', 'anymore'): 0.0021854316721822077
3. ('anymore', 'take'): 0.0019360156706248782
4. ('get', 'better'): 0.001044809713840763
5. ('every', 'day'): 0.0008303727856725715
6. ('suicidal', 'thought'): 0.0007680187852832392
7. ('year', 'old'): 0.0007102272727272727
8. ('want', 'live'): 0.0006934981506715982
9. ('wish', 'could'): 0.0006889356628382324
10. ('die', 'want'): 0.0006539565894490948
11. ('make', 'feel'): 0.0006539565894490948
12. ('want', 'kill'): 0.0006524357601713063
13. ('even', 'though'): 0.000649394101615729
14. ('mental', 'health'): 0.0006402691259489975
15. ('really', 'want'): 0.0006113733696710142
16. ('year', 'ago'): 0.0006037692232820712
17. ('want', 'end'): 0.0005459777107261047
18. ('want', 'go'): 0.0004744987346700409
19. ('even', 'know'): 0.0004653737590033093
20. ('every', 'time'): 0.000459290441892155
21. ('one', 'day'): 0.00042431136850301735
22. ('best', 'friend'): 0.0004091030757251314
23. ('end', 'life'): 0.0004060614171695542

```

24. ('high', 'school'): 0.0003847698072805139
25. ('every', 'single'): 0.00038324897800272534
26. ('last', 'year'): 0.00037868649016935957
27. ('know', 'anymore'): 0.00036956151450262797
28. ('long', 'time'): 0.0003680406852248394
29. ('need', 'help'): 0.0003604365388358964
30. ('first', 'time'): 0.00035587405100253065
31. ('anymore', 'want'): 0.00035435322172474204
32. ('keep', 'going'): 0.00034370741678022194
33. ('want', 'feel'): 0.00033762409966906756
34. ('commit', 'suicide'): 0.00033458244111349034
35. ('go', 'back'): 0.0003269782947245474
36. ('going', 'kill'): 0.00031785331905781586
37. ('know', 'want'): 0.00031633248978002725
38. ('feel', 'better'): 0.00031481166050223864
39. ('live', 'anymore'): 0.0003087283433910843
40. ('want', 'anymore'): 0.0003087283433910843
41. ('month', 'ago'): 0.0003072075141132957
42. ('perdedorsoy', 'un'): 0.00030264502627992993
43. ('un', 'perdedorsoy'): 0.00030264502627992993
44. ('anyone', 'else'): 0.00029047839205762117
45. ('get', 'help'): 0.00028895756277983256
46. ('friend', 'family'): 0.000287436733502044
47. ('past', 'year'): 0.00028287424566867823
48. ('self', 'harm'): 0.0002813534163908896
49. ('ever', 'since'): 0.000279832587113101
50. ('want', 'get'): 0.00027831175783531246

```

Mutual Information Score by Status

```
In [106]: bigram_measures = BigramAssocMeasures()

# Define a function to score bigrams using PMI
def score_bigrams(tokens):
    bigram_finder = BigramCollocationFinder.from_words(tokens)
    bigram_finder.apply_freq_filter(20) # Higher filter for larger dataframes
    bigrams_scored = bigram_finder.score_ngrams(bigram_measures.pmi)
    return bigrams_scored

# Group by 'status' and apply the scoring function
bigrams_by_status = data.groupby('status')['tokens'].apply(lambda x: score_tokens(x))

# Print or access the results for each status
for status, bigrams in bigrams_by_status.items():
    print(f"Status: {status}")
    for idx, (bigram, score) in enumerate(bigrams[:50], 1): # Print top 30
        print(f"{idx}. {bigram}: {score}")
    print()
```

Status: Anxiety

1. ('bowel', 'movement'): 11.581182384574525
2. ('acid', 'reflux'): 11.317567300511563
3. ('greatly', 'appreciated'): 11.004021752130182
4. ('ct', 'scan'): 10.397579523998575
5. ('shortness', 'breath'): 10.228421709514455
6. ('lymph', 'node'): 10.200825311554311
7. ('birth', 'control'): 10.167754138157338
8. ('urgent', 'care'): 10.140535238618813
9. ('full', 'blown'): 10.098080610350527
10. ('falling', 'asleep'): 10.040614003211822
11. ('fall', 'asleep'): 9.83293972517001
12. ('oh', 'god'): 9.80582934762625
13. ('yo', 'male'): 9.481336986883115
14. ('family', 'member'): 9.36030587704876
15. ('fast', 'forward'): 9.352348705895725
16. ('laying', 'bed'): 9.024405257291612
17. ('swollen', 'lymph'): 8.967848139576697
18. ('weight', 'loss'): 8.574923394102843
19. ('brain', 'fog'): 8.530897872104557
20. ('deep', 'breath'): 8.417517720375784
21. ('story', 'short'): 8.387144071332262
22. ('old', 'male'): 8.355806104799255
23. ('side', 'effect'): 8.323546554829242
24. ('burning', 'sensation'): 8.303648409045616
25. ('tension', 'headache'): 8.301124159936487
26. ('blood', 'clot'): 8.247458234388754
27. ('family', 'history'): 8.245966798996474
28. ('brain', 'tumour'): 8.18986095426949
29. ('let', 'u'): 8.137165818323917
30. ('brain', 'tumor'): 8.130967265215919
31. ('wanted', 'share'): 8.110829103092515
32. ('sore', 'throat'): 8.069427109977463
33. ('colon', 'cancer'): 8.065688681246275
34. ('intrusive', 'thought'): 8.06557705642198
35. ('old', 'female'): 8.04585065157797
36. ('long', 'term'): 8.032024496710955
37. ('hey', 'guy'): 7.9822431244705125
38. ('driving', 'crazy'): 7.874857382748166
39. ('ruining', 'life'): 7.784420131446765
40. ('chest', 'xray'): 7.746435810017676
41. ('made', 'mistake'): 7.73034882585004
42. ('panic', 'attack'): 7.722945961225726
43. ('every', 'single'): 7.588016575662834
44. ('heart', 'rate'): 7.58701602537608
45. ('high', 'school'): 7.538944096997277
46. ('blood', 'pressure'): 7.472970961371104
47. ('night', 'sweat'): 7.4565345812630035
48. ('hi', 'everyone'): 7.324356466174201
49. ('anybody', 'else'): 7.229778512419257
50. ('similar', 'experience'): 7.221985187402641

Status: Bipolar

1. ('ups', 'down'): 12.111783167798182
2. ('coping', 'mechanism'): 11.478071535428999
3. ('credit', 'card'): 11.410026868695972

4. ('rapid', 'cycling'): 10.696054490888672
5. ('panic', 'attack'): 10.269086642980934
6. ('full', 'blown'): 10.07975143109509
7. ('fall', 'asleep'): 9.825994838849306
8. ('suicide', 'attempt'): 9.63369884225899
9. ('suicidal', 'ideation'): 9.518270363713095
10. ('psych', 'ward'): 9.509213770326685
11. ('weight', 'gain'): 9.496022530313043
12. ('birth', 'control'): 9.310445944298209
13. ('self', 'harm'): 9.217106993710503
14. ('side', 'effect'): 8.99342055933467
15. ('sex', 'drive'): 8.90949764773684
16. ('story', 'short'): 8.809337494895699
17. ('family', 'member'): 8.777751151837863
18. ('mood', 'stabilizer'): 8.66540770198633
19. ('mental', 'health'): 8.454478986670253
20. ('mood', 'swing'): 8.220268170918066
21. ('top', 'world'): 8.154190144356038
22. ('every', 'single'): 8.028914900625093
23. ('mental', 'illness'): 7.999260272442079
24. ('let', 'u'): 7.907177756246586
25. ('intrusive', 'thought'): 7.740298258170059
26. ('wanted', 'share'): 7.643114622571343
27. ('trying', 'figure'): 7.350262317283287
28. ('health', 'insurance'): 7.263689465306337
29. ('long', 'term'): 7.183106975802806
30. ('racing', 'thought'): 7.0979440935605655
31. ('major', 'depressive'): 7.091483065716501
32. ('leave', 'house'): 7.061144645306237
33. ('mixed', 'state'): 7.000143573174313
34. ('anyone', 'else'): 6.97399333451898
35. ('depressive', 'episode'): 6.964607036998348
36. ('low', 'dose'): 6.869959046538275
37. ('make', 'sense'): 6.83753590002669
38. ('hi', 'everyone'): 6.813058020573816
39. ('mixed', 'episode'): 6.767482816384039
40. ('stopped', 'taking'): 6.71922507523141
41. ('social', 'anxiety'): 6.703138091063771
42. ('high', 'school'): 6.66185997759727
43. ('long', 'story'): 6.604570744241084
44. ('little', 'bit'): 6.594879593368466
45. ('quit', 'job'): 6.582105055291768
46. ('high', 'low'): 6.435126009075592
47. ('even', 'though'): 6.406063055301598
48. ('year', 'old'): 6.351927873843611
49. ('bipolar', 'ii'): 6.324370784151263
50. ('anyone', 'relate'): 6.27887663039648

Status: Depression

1. ('ups', 'down'): 13.472451021404577
2. ('anti', 'depressant'): 13.217828855386983
3. ('non', 'existent'): 13.191553881710274
4. ('blah', 'blah'): 12.902588991834817
5. ('sexually', 'assaulted'): 12.899617742401066
6. ('minimum', 'wage'): 12.877070782066696
7. ('bare', 'minimum'): 12.560988945115021

8. ('c', 'est'): 12.516627485644802
9. ('psych', 'ward'): 12.403372876954236
10. ('coping', 'mechanism'): 12.35730231317677
11. ('downward', 'spiral'): 12.153209510237515
12. ('cold', 'turkey'): 12.131855632277132
13. ('rough', 'patch'): 12.105567698767858
14. ('credit', 'card'): 12.091586292762068
15. ('http', 'www'): 12.02835105652327
16. ('brushing', 'teeth'): 11.915722201412366
17. ('rabbit', 'hole'): 11.763006490002276
18. ('rock', 'bottom'): 11.705117945679996
19. ('brush', 'teeth'): 11.68585465951377
20. ('http', 'co'): 11.662801740063543
21. ('existential', 'crisis'): 11.603259892580505
22. ('customer', 'service'): 11.510629587569204
23. ('grocery', 'store'): 11.500549277631228
24. ('greatly', 'appreciated'): 11.423445563972603
25. ('gon', 'na'): 11.21777261106266
26. ('wan', 'na'): 11.196425391777904
27. ('vicious', 'cycle'): 11.179861005187963
28. ('bank', 'account'): 11.081932096728114
29. ('trigger', 'warning'): 11.042850626218385
30. ('x', 'b'): 10.9312589350924
31. ('bachelor', 'degree'): 10.898203546369043
32. ('mood', 'swing'): 10.849984925444263
33. ('socially', 'awkward'): 10.84504384316021
34. ('daily', 'basis'): 10.748025552677248
35. ('dating', 'apps'): 10.425966922747719
36. ('burst', 'tear'): 10.41015955219044
37. ('smoke', 'weed'): 10.404921896302568
38. ('master', 'degree'): 10.37367689950986
39. ('que', 'eu'): 10.369640215820802
40. ('co', 'worker'): 10.290613658117238
41. ('substance', 'abuse'): 10.28887860674649
42. ('panic', 'attack'): 10.21923567660123
43. ('fall', 'asleep'): 10.174136891318618
44. ('smoking', 'weed'): 10.146426370069932
45. ('deeper', 'deeper'): 10.134764818992073
46. ('side', 'effect'): 9.986138019439089
47. ('th', 'july'): 9.94314511794199
48. ('depressive', 'episode'): 9.932135741263966
49. ('falling', 'apart'): 9.894583832027632
50. ('black', 'hole'): 9.758769497115805

Status: Normal

1. ('ó', 'ó'): 11.711466028357531
2. ('social', 'medium'): 10.793634569938359
3. ('ð', 'ð'): 10.750069235453424
4. ('mental', 'health'): 10.328064336640521
5. ('tinyurl', 'com'): 10.2846209224505
6. ('twitpic', 'com'): 10.2846209224505
7. ('http', 'twitpic'): 10.10175686530069
8. ('http', 'tinyurl'): 10.040356320636548
9. ('gon', 'na'): 10.023754353299417
10. ('wan', 'na'): 9.934487015202329
11. ('ya', 'allah'): 9.138478480426693

12. ('â', 'âï'): 8.994228400996246
13. ('help', 'rtlike'): 8.097159323275127
14. ('high', 'school'): 8.088314602388618
15. ('looking', 'forward'): 7.985466861920727
16. ('let', 'u'): 7.516794364579535
17. ('got', 'ta'): 7.472763781834038
18. ('oh', 'god'): 7.284301353274115
19. ('sorry', 'hear'): 7.236686445386798
20. ('â', 'â'): 7.166807855884535
21. ('anyone', 'else'): 7.077347120396393
22. ('last', 'night'): 7.0214806999540755
23. ('even', 'though'): 6.920314225292586
24. ('u', 'mutual'): 6.872938174804808
25. ('someone', 'else'): 6.821161302013103
26. ('month', 'ago'): 6.73783072236713
27. ('good', 'luck'): 6.659120489206369
28. ('little', 'bit'): 6.6227270191313945
29. ('take', 'care'): 6.551876046588539
30. ('year', 'ago'): 6.537722827196113
31. ('thank', 'god'): 6.317787786385701
32. ('year', 'old'): 6.317029893676509
33. ('next', 'week'): 6.310941098895819
34. ('wish', 'could'): 6.135789392899053
35. ('feel', 'free'): 5.9941982093014285
36. ('wish', 'wa'): 5.964969141635342
37. ('every', 'day'): 5.747035772174073
38. ('best', 'friend'): 5.713157730768039
39. ('last', 'week'): 5.66499339683882
40. ('good', 'morning'): 5.305539193048919
41. ('thank', 'much'): 4.9456193142041975
42. ('long', 'time'): 4.912010732940267
43. ('come', 'back'): 4.883306938658016
44. ('last', 'year'): 4.782640326716354
45. ('first', 'time'): 4.737107895428707
46. ('many', 'people'): 4.571242148601911
47. ('really', 'hard'): 4.558596681711677
48. ('every', 'time'): 4.435326039324497
49. ('want', 'buy'): 4.196734913402754
50. ('next', 'time'): 4.180418972503482

Status: Personality disorder

1. ('mental', 'health'): 9.249520894286926
2. ('leave', 'house'): 8.431205420029281
3. ('avoidant', 'personality'): 7.932043364925319
4. ('social', 'medium'): 7.806053829192427
5. ('personality', 'disorder'): 7.8037436685052235
6. ('high', 'school'): 7.644219726387684
7. ('social', 'interaction'): 7.310096334530016
8. ('social', 'skill'): 6.947526255145306
9. ('anyone', 'else'): 6.617545540705368
10. ('year', 'old'): 6.59755179563942
11. ('social', 'anxiety'): 6.578157557438642
12. ('wish', 'could'): 6.413075982945557
13. ('year', 'ago'): 6.143531374539938
14. ('social', 'situation'): 5.914611632672573
15. ('everyone', 'else'): 5.862261704636246

16. ('even', 'though'): 5.810996878865254
17. ('pretty', 'much'): 5.6981996763063005
18. ('last', 'year'): 5.655462550447956
19. ('good', 'enough'): 5.553682978135585
20. ('diagnosed', 'avpd'): 5.440872536351179
21. ('something', 'wrong'): 5.425232045251494
22. ('whole', 'life'): 5.332899685874468
23. ('go', 'back'): 5.113792806264414
24. ('first', 'time'): 4.637670992558519
25. ('every', 'time'): 4.588761392077572
26. ('get', 'better'): 4.4441100554571875
27. ('long', 'time'): 4.315742897671157
28. ('many', 'people'): 3.590614168497744
29. ('make', 'feel'): 3.5062788765436146
30. ('never', 'really'): 3.4638278276557237
31. ('feel', 'bad'): 3.382290159268159
32. ('make', 'friend'): 3.206324617762359
33. ('really', 'want'): 2.8349219793469693
34. ('one', 'thing'): 2.821821145050741
35. ('else', 'feel'): 2.7555074828523765
36. ('talk', 'people'): 2.704740287309356
37. ('even', 'know'): 2.2848827343553246
38. ('want', 'get'): 2.2293118352762384
39. ('know', 'people'): 1.062168821086427
40. ('people', 'know'): 1.062168821086427
41. ('people', 'feel'): 0.41637009793279134

Status: Stress

1. ('panic', 'attack'): 9.193977424287317
2. ('mental', 'health'): 8.289216073734716
3. ('family', 'member'): 8.048505667969753
4. ('anyone', 'else'): 7.067381196117321
5. ('high', 'school'): 6.905987222178114
6. ('anxiety', 'disorder'): 6.837003855040422
7. ('year', 'old'): 6.763656717424929
8. ('month', 'ago'): 6.703413617499113
9. ('health', 'issue'): 6.4972091160105485
10. ('take', 'care'): 6.3626498425323526
11. ('even', 'though'): 6.228962350002774
12. ('stress', 'management'): 5.997363160780232
13. ('last', 'night'): 5.995321827870546
14. ('ever', 'since'): 5.887568204575567
15. ('pretty', 'much'): 5.860921890964105
16. ('came', 'back'): 5.772676279936226
17. ('year', 'ago'): 5.7527684012821965
18. ('stress', 'relief'): 5.715592192364852
19. ('feel', 'guilty'): 5.659841858575415
20. ('best', 'friend'): 5.595670739566128
21. ('depression', 'anxiety'): 5.529769645296094
22. ('two', 'week'): 5.473845122945038
23. ('new', 'job'): 5.4084198239236265
24. ('manage', 'stress'): 5.256160573727556
25. ('chronic', 'stress'): 5.225266565781768
26. ('past', 'month'): 5.162775877465169
27. ('every', 'day'): 5.075837120120013
28. ('week', 'ago'): 5.075745347572269

29. ('stress', 'level'): 5.074046163277327
30. ('past', 'week'): 5.028439632793916
31. ('please', 'help'): 5.013408291975608
32. ('last', 'week'): 4.9929778874022475
33. ('full', 'time'): 4.992310201310108
34. ('let', 'go'): 4.965285348140313
35. ('something', 'else'): 4.893127729917467
36. ('last', 'year'): 4.86744700531839
37. ('next', 'day'): 4.8247993458927
38. ('anxiety', 'attack'): 4.762041797359199
39. ('many', 'thing'): 4.6468179902372135
40. ('last', 'month'): 4.628508275102057
41. ('go', 'away'): 4.558027905337603
42. ('friend', 'family'): 4.387330001995
43. ('past', 'year'): 4.383534591616478
44. ('first', 'time'): 4.377782291385632
45. ('need', 'help'): 4.3001569365453705
46. ('get', 'better'): 4.275842076892026
47. ('really', 'bad'): 4.230266304331256
48. ('long', 'time'): 4.1767347724475385
49. ('know', 'anymore'): 4.1352069150196655
50. ('deal', 'stress'): 4.130629691643694

Status: Suicidal

1. ('carbon', 'monoxide'): 14.076615008220664
2. ('anti', 'depressant'): 13.721495433599353
3. ('blah', 'blah'): 13.32671035834696
4. ('minimum', 'wage'): 12.713227817615753
5. ('coping', 'mechanism'): 12.649825368184763
6. ('survival', 'instinct'): 12.5807559809535
7. ('rock', 'bottom'): 12.16715433028819
8. ('mood', 'swing'): 12.145271368887606
9. ('sexually', 'assaulted'): 12.025093833488956
10. ('psych', 'ward'): 11.78693469078472
11. ('perdedorsoy', 'un'): 11.4379671094487
12. ('un', 'perdedorsoy'): 11.4379671094487
13. ('slit', 'wrist'): 11.395973020784075
14. ('depressive', 'episode'): 11.32295822348585
15. ('soy', 'un'): 11.253542538311274
16. ('panic', 'attack'): 11.170045368128683
17. ('un', 'soy'): 11.11603901456134
18. ('borderline', 'personality'): 11.087623332397254
19. ('daily', 'basis'): 10.900445603644862
20. ('bank', 'account'): 10.876677437711914
21. ('pull', 'trigger'): 10.623960479518665
22. ('fall', 'asleep'): 10.334594330966423
23. ('falling', 'apart'): 10.323639704386993
24. ('permanent', 'solution'): 10.309054672667827
25. ('video', 'game'): 10.159292212515222
26. ('cease', 'exist'): 10.157161107666218
27. ('social', 'medium'): 10.105221209649569
28. ('waste', 'space'): 9.868518373318537
29. ('personality', 'disorder'): 9.761608279986644
30. ('closer', 'closer'): 9.647500958728017
31. ('drug', 'addict'): 9.612464840680838
32. ('self', 'esteem'): 9.586929748573699

```
33. ('attention', 'seeking'): 9.473140881306533
34. ('laying', 'bed'): 9.470284829721429
35. ('bipolar', 'disorder'): 9.460903456186799
36. ('side', 'effect'): 9.428985888130516
37. ('front', 'train'): 9.413678498526773
38. ('eating', 'disorder'): 9.402065893299506
39. ('th', 'grade'): 9.346756300548716
40. ('sexually', 'abused'): 9.328791636892419
41. ('lay', 'bed'): 9.279807588461745
42. ('jump', 'bridge'): 9.205059113230055
43. ('self', 'harm'): 9.079392219797711
44. ('playing', 'video'): 9.03867485129728
45. ('self', 'harming'): 9.011815033808892
46. ('self', 'harmed'): 9.001967247852543
47. ('thanks', 'reading'): 8.945866032472123
48. ('blow', 'brain'): 8.905246589908682
49. ('sexual', 'abuse'): 8.892677916405628
50. ('fast', 'forward'): 8.856803152804726
```

There is a lot of really fascinating and informative findings in our exploration of bigrams.

For example, 'Anxiety' users appear to speak about physical health ailments and concerns (such as brain tumors, lymph nodes, heart attacks), which makes a lot of sense. 'Bipolar' users speak in 'up & down', 'high & low', 'mood swings'...which also makes sense. 'Depression' users conjure "downward" imagery of 'rock bottom', 'rabbit hole', etc.

This is all super helpful!

Vectorization

Let's move ahead with our train test split and begin vectorization of a model.

We'll use **TF-IDF** vectorization, and we'll compare token frequencies not only to their total frequencies across documents/'statement' values...but across entire 'status' groups as well.

Let's make a new dataframe 'filtered_df', that extracts only the necessary features we need for our models.

```
In [107]: data.head()
```

Out [107...]

	statement	status	perc_upper_words	neg_emoji	exclamation_count	question_count
0	oh my gosh	Anxiety	0.0	0	0	
1	trouble sleeping, confused mind, restless hear...	Anxiety	0.0	0	0	
2	All wrong, back off dear, forward doubt. Stay ...	Anxiety	0.0	0	0	
3	I have shifted my focus to something else but ...	Anxiety	0.0	0	0	
4	I am restless and restless, it is been a month...	Anxiety	0.0	0	0	

In [108...]

```
columns_to_keep = ['status', 'perc_upper_words', 'neg_emoji', 'exclamation_count', 'question_count']
filtered_df = data.loc[:, columns_to_keep]
filtered_df.head()
```

Out [108...]

	status	perc_upper_words	neg_emoji	exclamation_count	question_count	tokens
0	Anxiety	0.0	0	0	0	[oh, go...]
1	Anxiety	0.0	0	0	0	[trou...]
2	Anxiety	0.0	0	0	0	[wro...]
3	Anxiety	0.0	0	0	0	[shift...]
4	Anxiety	0.0	0	0	1	[restle...]

In [109...]

filtered_df.info()

```

<class 'pandas.core.frame.DataFrame'>
Index: 51073 entries, 0 to 52840
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   status            51073 non-null   object 
 1   perc_upper_words  51073 non-null   float64
 2   neg_emoji          51073 non-null   int64  
 3   exclamation_count 51073 non-null   int64  
 4   question_count    51073 non-null   int64  
 5   tokens             51073 non-null   object 
dtypes: float64(1), int64(3), object(2)
memory usage: 2.7+ MB

```

Filtering for Empty Tokens

After all of our preprocessing and cleaning...we might have some rows that are essentially "empty", with 0 tokens. This could happen if a person's statement included only stop words, emoticons, punctuation, etc.

We'll filter for these and set them to NaN. If there are relatively few, compared to the 51,000+ rows we currently have in our dataframe, we can feel good about simply dropping them.

In [110...]

```
# Set rows with empty 'tokens' lists to NaN
filtered_df.loc[filtered_df['tokens'].apply(lambda x: len(x) == 0), 'tokens']
```

```
# Optional: Print the DataFrame to verify
filtered_df.info()

<class 'pandas.core.frame.DataFrame'>
Index: 51073 entries, 0 to 52840
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   status            51073 non-null   object  
 1   perc_upper_words  51073 non-null   float64 
 2   neg_emoji         51073 non-null   int64  
 3   exclamation_count 51073 non-null   int64  
 4   question_count    51073 non-null   int64  
 5   tokens             50955 non-null   object  
dtypes: float64(1), int64(3), object(2)
memory usage: 2.7+ MB
```

In [111...]: # Dropping NaNs
`filtered_df = filtered_df.dropna(subset=['tokens'])`

Multinomial Naive Bayes Classifier

For our first model, we'll try a multinomial naive bayes classifier. This type of model deals with text data alone, so we'll leave the other numeric features we added out of this train test split.

Train Test Split and Vectorize

In [112...]: # Convert token lists to strings
`filtered_df['joined_tokens'] = filtered_df['tokens'].apply(lambda x: ' '.join(x))`

Define features and target
`X = filtered_df['joined_tokens']
y = filtered_df['status'] # 'status' is the target variable`

Train-test split
`X_train, X_test, y_train, y_test = train_test_split(
 X, y, test_size=0.3, random_state=42
)`

GridSearchCV

We can either vectorize our singular tokens as unigrams, or include bigrams in our analysis.

Furthermore, our dataset is larger so a higher number of max_features might be better.

We can test the best parameters for our model using GridSearchCV

Check for Class Imbalance

Before we fit our model, we'll need to handle class imbalance.

```
In [113...]: # Assuming `y_train` contains the target labels
class_distribution = y_train.value_counts(normalize=True)
class_distribution
```

```
Out[113...]: status
Normal           0.307895
Depression       0.300998
Suicidal         0.208843
Anxiety          0.071885
Bipolar           0.048419
Stress            0.044438
Personality disorder  0.017523
Name: proportion, dtype: float64
```

Large Dataset - Class Imbalance, Feature Importance

Because our dataset is larger, we'll try to account for class imbalance by adjusting the class weights with the MultinomialNB 'class_prior'.

We'll also want to optimize computational runtimes, by narrowing our feature selection if possible...We tried using **SelectKBest** with a **chi-squared** test. This would evaluate the statistical importance of each feature, by measuring the independence between a feature and target variable.

SelectKBest would be the most suitable option for a model like Multinomial Naive Bayes. We'll look at other tree-based models and options for feature selection in a bit.

However, our scores were even worse, so we'll remove SelectKbest from our Pipeline for now.

```
In [114...]: # Set class prior probabilities based on class distribution
class_prior = [0.307895, 0.300998, 0.208843, 0.071885, 0.048419, 0.044438, 0.017523]
```



```
In [115...]: # Define the parameter grid for GridSearchCV
param_grid = {
    'tfidf_max_features': [10000, 20000], # Number of features to use in the vectorizer
    'tfidf_ngram_range': [(1, 1), (1, 2)], # Use unigrams and/or bigrams
    'feature_selection_k': [1000, 5000, 'all'],
    'clf_alpha': [0.1, 0.5, 1.0], # Smoothing parameter for Naive Bayes
    'clf_fit_prior': [True, False] # Whether to learn class prior probabilities
}

# Create a pipeline that first transforms data using TF-IDF and then applies the classifier
pipeline = Pipeline([
    ('tfidf', TfidfVectorizer()), # Placeholder for TF-IDF Vectorizer
    ('feature_selection', SelectKBest(score_func=chi2)),
    ('clf', MultinomialNB(class_prior=class_prior)) # Naive Bayes with class prior
```

```
])

# Initialize GridSearchCV with timeout for joblib workers
timeout_in_seconds = 300 # Timeout set to 5 minutes (300 seconds)
with parallel_backend('loky'):
    grid_search = GridSearchCV(
        pipeline,
        param_grid,
        cv=5, # Number of folds for cross-validation
        scoring='f1_weighted', # F1 score, weighted for imbalanced classes
        verbose=1, # Print progress
        n_jobs=-1 # Use all available cores
)

# Fit GridSearchCV
grid_search.fit(X_train, y_train)

# After fitting, now you can access best parameters and scores
print(f"Best parameters: {grid_search.best_params_}")
print(f"Best F1 score: {grid_search.best_score_}")

# Evaluate on the test set
best_model = grid_search.best_estimator_
test_score = best_model.score(X_test, y_test)
print(f"Test set score: {test_score}")
```

Fitting 5 folds for each of 72 candidates, totalling 360 fits
 Best parameters: {'clf_alpha': 0.1, 'clf_fit_prior': True, 'feature_selection_k': 'all', 'tfidf_max_features': 20000, 'tfidf_ngram_range': (1, 2)}
 Best F1 score: 0.4304857142867046
 Test set score: 0.46588604696801206

In [116]: # Assuming best_model is your best estimator from GridSearchCV
 tfidf_vectorizer = best_model.named_steps['tfidf']
 X_train_tfidf = tfidf_vectorizer.transform(X_train)

 # Inspect the shape of the transformed feature matrix
 print(f"Shape of X_train_tfidf: {X_train_tfidf.shape}")

Shape of X_train_tfidf: (35668, 20000)

In [117]: # Generate predictions on the test set
y_pred = best_model.predict(X_test)

Print classification report
report = classification_report(y_test, y_pred, target_names=best_model.classes_)
print("Classification Report:\n", report)

Optionally, print confusion matrix
cm = confusion_matrix(y_test, y_pred, labels=best_model.classes_)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=best_model.classes_, yticklabels=best_model.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()

Classification Report:

	precision	recall	f1-score	support	
Personality disorder	Anxiety	0.33	0.85	0.47	1053
	Bipolar	0.23	0.78	0.36	774
	Depression	0.46	0.80	0.58	4350
	Normal	0.94	0.38	0.54	4944
	Stress	0.82	0.27	0.41	270
	Suicidal	0.70	0.12	0.20	708
	accuracy			0.47	15287
	macro avg	0.62	0.46	0.37	15287
	weighted avg	0.70	0.47	0.42	15287

Confusion Matrix



These scores are not very good. Which is okay! We shall continue tuning!

Stress and Suicidal cases seem particularly hard to accurately predict, with lower f1-scores. We will keep this in mind as we tune and test other models...in case this seems to be a trend.

We can come back to this model and tune it to yield better results...but first let's test some other models and see if they perform better as a baseline.

Before we move on, let's see whether there's any informative results on feature importance or irrelevancy, that we can highlight and compare with other models...We moved ahead and tried running other models like *XGBoost*, however they unsurprisingly didn't reach convergence due to extraneous computational runtimes. This makes sense, because our dataset is currently so large and complex with 7 target variables. We need to figure out how to narrow things down if possible.

We can try making lists of most and least important features across models...then we can cross-reference using these lists and figure out which tokens to drop and which to keep as we hone in on feature selection.

Feature Importance - MultinomialNB

Multinomial Bayes doesn't extract feature importance like some of our other predictive models...but we can still get some good idea of feature relevancy using log probabilities and extracting feature names. This will tell us the log of the probability that a particular word ends up in a feature class.

We'll have to extract from SelectKBest

```
In [118]: # Accessing the TfIdfVectorizer and SelectKBest from the pipeline
tfidf_vectorizer = best_model.named_steps['tfidf']
select_k_best = best_model.named_steps['feature_selection'] # Assuming 'fea

# Getting feature names (words) from TfIdfVectorizer
feature_names = tfidf_vectorizer.get_feature_names_out()

# Get indices of selected features from SelectKBest
selected_feature_indices = select_k_best.get_support(indices=True)

# Accessing MultinomialNB classifier from the pipeline
clf = best_model.named_steps['clf']

# Getting the log probabilities of features given each class
log_probabilities = clf.feature_log_prob_

# Determine top and bottom features for all classes combined and store in co
top_n = 50 # Number of top and bottom features to consider

multinomialnb_top_features = set() # Using set to ensure uniqueness
multinomialnb_bottom_features = set()

for i, class_label in enumerate(clf.classes_):
    # Filter log_probabilities and feature_names with selected_feature_indices
    class_log_probabilities = log_probabilities[i, selected_feature_indices]
    class_feature_names = [feature_names[idx] for idx in selected_feature_in

    # Sort indices based on log probabilities
    top_features_idx = np.argsort(class_log_probabilities)[-1:-top_n]
    bottom_features_idx = np.argsort(class_log_probabilities)[:top_n]
```

```
# Retrieve top and bottom features
top_features = [class_feature_names[idx] for idx in top_features_idx]
bottom_features = [class_feature_names[idx] for idx in bottom_features_i

# Extend sets with top and bottom features
multinomialnb_top_features.update(top_features)
multinomialnb_bottom_features.update(bottom_features)

# Convert sets back to lists
multinomialnb_top_features = list(multinomialnb_top_features)
multinomialnb_bottom_features = list(multinomialnb_bottom_features)
```

```
multinomialnb_top_features
```

```
multinomialnb_bottom_features
```

In [119...]

```
# Print top 20 and bottom 20 features for each class individually
for i, class_label in enumerate(clf.classes_):
    # Filter log_probabilities and feature_names with selected_feature_indices
    class_log_probabilities = log_probabilities[i, selected_feature_indices]
    class_feature_names = [feature_names[idx] for idx in selected_feature_ir

    # Sort indices based on log probabilities
    top_features_idx = np.argsort(class_log_probabilities)[::-1][:20]
    bottom_features_idx = np.argsort(class_log_probabilities)[:20]

    # Retrieve top and bottom features
    top_features = [class_feature_names[idx] for idx in top_features_idx]
    bottom_features = [class_feature_names[idx] for idx in bottom_features_i

    print(f"Top 20 features for class '{class_label}':")
    print(top_features)
    print()

    print(f"Bottom 20 features for class '{class_label}':")
    print(bottom_features)
    print()
```

Top 20 features for class 'Anxiety':

```
['anxiety', 'restless', 'feel', 'heart', 'worried', 'symptom', 'get', 'anxious', 'day', 'time', 'know', 'nervous', 'pain', 'health', 'doctor', 'really', 'feeling', 'week', 'something', 'cancer']
```

Bottom 20 features for class 'Anxiety':

```
['aa', 'know wanted', 'know wa', 'know use', 'know tried', 'know told', 'know talk', 'know suicidal', 'know since', 'know side', 'know shit', 'know selfish', 'know pretty', 'know place', 'know parent', 'know wish', 'know nobody', 'know may', 'know man', 'know loved']
```

Top 20 features for class 'Bipolar':

```
['bipolar', 'feel', 'med', 'episode', 'manic', 'get', 'know', 'time', 'really', 'anyone', 'would', 'year', 'day', 'want', 'mg', 'mania', 'thing', 'medication', 'diagnosed', 'work']
```

Bottom 20 features for class 'Bipolar':

```
['öy', 'told love', 'told loved', 'ibuprofen', 'sidewalk', 'sided', 'told mom', 'side neck', 'id get', 'told mother', 'id never', 'id say', 'destroys', 'told needed', 'told nothing', 'destroy everything', 'idc', 'destiny', 'devastate', 'destined']
```

Top 20 features for class 'Depression':

```
['feel', 'depression', 'want', 'know', 'life', 'get', 'even', 'time', 'people', 'day', 'thing', 'friend', 'really', 'year', 'would', 'one', 'make', 'going', 'think', 'help']
```

Bottom 20 features for class 'Depression':

```
['öy', 'cbd oil', 'uti', 'jyp', 'vaguely', 'valentine', 'kill somebody', 'similar symptom', 'side neck', 'side face', 'know bipolar', 'chronic stress', 'iâm tired', 'die keep', 'view poll', 'seller', 'anxiety hi', 'lebaran', 'lecturer', 'left arm']
```

Top 20 features for class 'Normal':

```
['want', 'good', 'really', 'go', 'morning', 'yes', 'oh', 'day', 'get', 'today', 'love', 'time', 'work', 'going', 'one', 'know', 'still', 'think', 'come', 'people']
```

Bottom 20 features for class 'Normal':

```
['literally never', 'please anyone', 'frequently', 'playing videogames', 'playing video', 'play video', 'atitude', 'please forgive', 'planning suicide', 'planning end', 'plan work', 'plan kill', 'plan future', 'plagued', 'placebo', 'planning kill', 'friend amp', 'please know', 'please try', 'point done']
```

Top 20 features for class 'Personality disorder':

```
['avpd', 'people', 'feel', 'know', 'would', 'even', 'think', 'want', 'get', 'friend', 'time', 'thing', 'anyone', 'social', 'make', 'life', 'really', 'one', 'poll', 'someone']
```

Bottom 20 features for class 'Personality disorder':

```
['aa', 'mg zoloft', 'michael', 'middle class', 'middle nowhere', 'midnight', 'midst', 'might better', 'might bipolar', 'might end', 'might find', 'might get', 'might go', 'might happen', 'might hurt', 'might lose', 'might need', 'might seem', 'might something', 'might take']
```

```
Top 20 features for class 'Stress':
['stress', 'feel', 'get', 'know', 'time', 'help', 'work', 'would', 'really',
'day', 'thing', 'anxiety', 'even', 'stressed', 'year', 'job', 'go', 'want',
'need', 'one']
```

```
Bottom 20 features for class 'Stress':
['ðÿ', 'someone happy', 'feel running', 'one found', 'feel restless', 'feel
relief', 'someone hug', 'someone kill', 'life ever', 'life every', 'life eve
ryday', 'feel purpose', 'one first', 'feel powerless', 'feel pointless', 'li
fe failed', 'feel pleasure', 'feel place', 'feel piece', 'someone need']
```

```
Top 20 features for class 'Suicidal':
['want', 'life', 'feel', 'know', 'die', 'would', 'get', 'people', 'kill', 'g
oing', 'even', 'anymore', 'fucking', 'one', 'think', 'time', 'hate', 'frien
d', 'suicide', 'really']
```

```
Bottom 20 features for class 'Suicidal':
['aa', 'fun time', 'fuckfuckfuck fuck', 'fuckfuckfuck', 'fuck fuckfuckfuck',
'friend wa', 'friend depression', 'friend back', 'gabapentin', 'frantic', 'f
mla', 'floater', 'fixate', 'fitbit', 'first got', 'first episode', 'finding
hard', 'found nothing', 'gallbladder', 'geodon']
```

Dropping - 'know', 'told', 'might', 'feel'

Already from these, we can gain some valuable insight into features we can drop. Although these selections were made by log probabilities, which don't directly tell us a certain correlation or impact a feature might have on predicting a given class...we can highlight some repetitive features that add nothing but noise to our data.

'know' and **'told'** are two that might be helpful to drop. We'll keep these for now, since there's a *chance* they might implicate a unique mental state (ex. 'know' can help us determine whether or not a person feels grounded in reality 'i know', or lost 'i don't know'...'told' could help implicate whether someone feels unheard 'i told', or attacked 'they told me!'...it's unlikely we'll need them but we'll run some more models before coming to any concrete conclusion).

'feel' can be dropped. It's been appearing as one of the most frequently used tokens in all of our outputs regarding feature/token importance...which makes sense given that this dataset is exploring internal emotional states of users. Furthermore, 'feel' doesn't give us any further insight...because whatever word follows it (ex 'i feel **stressed**', 'i feel **happy**', 'i feel **worried**') is the actually informative word, *and* will implicate the verb "to feel".

'might' can also be dropped. It's a filler word that doesn't seem to be operating in bigrams in a way that helpfully implicates emotional uncertainty. There will be other more unique ways to determine uncertainty of mental state by class if necessary.

```
In [120]: # Remove 'like' from tokens, now that we've made bigrams
def remove_feel(tokens):
    return [token for token in tokens if token != 'feel']
```

```

def remove_might(tokens):
    return [token for token in tokens if token != 'might']

# Apply function
data['tokens'] = data['tokens'].apply(remove_feel)
filtered_df['tokens'] = filtered_df['tokens'].apply(remove_feel)

data['tokens'] = data['tokens'].apply(remove_might)
filtered_df['tokens'] = filtered_df['tokens'].apply(remove_might)

# Set rows with empty 'tokens' lists to NaN
filtered_df.loc[filtered_df['tokens'].apply(lambda x: len(x) == 0), 'tokens'] = np.nan

# Optional: Print the DataFrame to verify
filtered_df.info()

<class 'pandas.core.frame.DataFrame'>
Index: 50955 entries, 0 to 52840
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype  
---  -- 
 0   status            50955 non-null   object 
 1   perc_upper_words  50955 non-null   float64
 2   neg_emoji          50955 non-null   int64  
 3   exclamation_count 50955 non-null   int64  
 4   question_count     50955 non-null   int64  
 5   tokens             50953 non-null   object 
 6   joined_tokens      50955 non-null   object 
dtypes: float64(1), int64(3), object(3)
memory usage: 3.1+ MB

```

In [121...]

```

# Dropping NaNs
filtered_df = filtered_df.dropna(subset=['tokens'])

# Convert token lists to strings
filtered_df['joined_tokens'] = filtered_df['tokens'].apply(lambda x: ' '.join(x))

```

Feature Selection

We need to continue to hone and select our most important features, in order to hopefully build models with better scores that are less computationally expensive.

Multinomial Logistic Regression: LEAVING AS MARKDOWN

The code below for this logistic regression model did not run and ran into many errors, due to further tuning and engineering that was needed of our hyperparameters, our dataset as a whole, and our features.

To reduce computational expense, we're going to put it as markdowns, so it does not run everytime we run our notebook.

We tried many iterations of Random Forest and Logistic Regression models that aren't all evidenced here (as we deleted their code cells) -- but we want to leave some evidence of our train of thought and how we got from point A to point B. So we'll keep this particular code in.

CODE FOR MULTINOMIAL LOGISTIC REGRESSION ATTEMPT #1:

Let's see how multinomial logistic regression performs as a model, and extract relevant features like we did for Multinomial Naive Bayes.

We'll deal with **class imbalance** by adjusting our **class weight hyperparameter**, rather than using a technique like SMOTE.

Including Numeric Features

We'll be able to include the numeric features we calculated earlier on, in addition to our vectorized text data.

LabelEncoder()

We'll need to use labelencoder() to encode our target variable classes, so they are integer values that the model can interpret.

```
X = filtered_df.drop(columns=['status', 'tokens']) y = filtered_df['status']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

encoder = LabelEncoder() y_train_encoded = encoder.fit_transform(y_train)
y_test_encoded = encoder.transform(y_test) y_train = pd.Series(y_train_encoded) y_test
= pd.Series(y_test_encoded)
```

Check encoded labels

```
print("Encoded y_train:") print(y_train.value_counts())

print("Encoded y_test:") print(y_test.value_counts())
```

Mapping from encoded labels to original labels

```
label_mapping = dict(zip(encoder.classes_, encoder.transform(encoder.classes_)))
print("Label Mapping:") print(label_mapping)

vectorizer = TfidfVectorizer(max_features=20000, ngram_range=(1, 2))

X_train_text = vectorizer.fit_transform(X_train['joined_tokens']) X_test_text =
vectorizer.transform(X_test['joined_tokens'])
```

Scale numeric data

Initialize StandardScaler

```
scaler = StandardScaler()
```

Fit and transform StandardScaler on numeric data

```
X_train_numeric_scaled = scaler.fit_transform(X_train[['perc_upper_words', 'neg_emoji', 'exclamation_count', 'question_count']]) X_test_numeric_scaled = scaler.transform(X_test[['perc_upper_words', 'neg_emoji', 'exclamation_count', 'question_count']])
```

```
X_train_combined = hstack([X_train_text, X_train_numeric_scaled]) X_test_combined = hstack([X_test_text, X_test_numeric_scaled])
```

Feature Importance - SelectFromModel, Random Forest

To minimize computational expense, we need to hone our feature selection in our model once again. Instead of using SelectKBest this time with a chi-squared test, let's use **SelectFromModel** paired with **RandomForestClassifier()** to select predictive features.

We'll use Random Forest as opposed to Gradient Boosting, since our dataset is quite large and Gradient Boosting might be more computationally expensive.

Example param_grid for GridSearchCV with SelectFromModel

```
param_grid = { 'select__threshold': ['mean', 'median', '2.*mean'], 'clf__C': [0.1, 1.0, 10.0], 'clf__solver': ['lbfgs', 'saga'], 'clf__max_iter': [2000], # Balance computational cost with convergence needs for max_iter }
```

Example pipeline with SelectFromModel and Logistic Regression

```
pipeline = Pipeline([ ('select', SelectFromModel(estimator=RandomForestClassifier())), ('clf', LogisticRegression(class_weight='balanced', random_state=42)), # Removed multi_class to account for warning, multinomial is default ])
```

Initialize GridSearchCV

```
grid_search = GridSearchCV( pipeline, param_grid, cv=5, scoring='f1_weighted', verbose=0, n_jobs=-1, )
```

Fit GridSearchCV

```
grid_search.fit(X_train_combined, y_train)
```

Evaluate on the test set

```
test_score = grid_search.score(X_test_combined, y_test) print(f"Test set score: {test_score}")
```

```
print("Best parameters:", grid_search.best_params_) print("Best F1 score:",  
grid_search.best_score_)
```

Convergence Errors

We've run our logistic regression model many times now, and continue to get errors regarding our max_iterations. We've tried multiple values, yet it seems this model might not be working. Our dataset is pretty large and complex, so logistic regression might simply not be the best model.

We'll leave the code in here as markdowns, in case it's helpful to refer back to at some point. There's no use having it run unnecessarily everytime we run our notebook.

Modifying our Dataset

We've continue to run into some pretty serious issues with this dataset, as we've attempted to run different models. To eliminate computational cost, we have not included all of the iterations of the models we ran. After carefully and thoroughly tuning hyperparameters, adjusting feature selection, etc. for Multinomial Naive Bayes, Multinomial Logistic Regression, Random Forest, and XGBoost...it's clear that our dataset is currently formatted in a way that will either yield poor model results or result in errors and model failure.

Our **dataset is currently very large and complex**, with 7 different subcategories of our main target variable that we are trying to predict from 50,000 rows of vectorized text data.

If we had more time, we could invest in fine tuning these complex models further and relying on more machine learning resources and programs that would allow our models to run with these complex structures. *However*, we have a deadline to complete for this particular project, and furthermore upon reviewing our dataset it's current formatting might not be the most optimal anyway.

Reassessing our Target Status Groups

Let's examine our 7 different 'status' groups. This is a lot of groups to predict, which creates the need more complex feature engineering on top of greater computational expense. Most importantly, we noticed there are some overlaps among the group 'statement' values, that might make it hard to accurately predict all 7 groups. AKA...it seems like our data might not be categorized specifically enough to yield great models as is.

Take a look at these selections 50 **statement values from *Anxiety* and from *Stress***...

```
In [122]: # Filter rows where 'status' is 'Anxiety' and get the 'statement' values
anxiety_statements = data.loc[data['status'] == 'Anxiety', 'statement'].head(50)

# Print the first 50 statements
for idx, statement in enumerate(anxiety_statements):
    print(f"Statement {idx + 1}: {statement}")
```

Statement 1: oh my gosh
Statement 2: trouble sleeping, confused mind, restless heart. All out of tune
Statement 3: All wrong, back off dear, forward doubt. Stay in a restless and restless place
Statement 4: I have shifted my focus to something else but I am still worried
Statement 5: I am restless and restless, it is been a month now, boy. What do you mean?
Statement 6: every break, you must be nervous, like something is wrong, but what the heck
Statement 7: I feel scared, anxious, what can I do? And may my family or us be protected :)
Statement 8: Have you ever felt nervous but did not know why?
Statement 9: I have not slept well for 2 days, it is like I am restless. why huh :([].
Statement 10: I am really worried, I want to cry.
Statement 11: always restless every night, even though I do not know why, what is wrong. strange.
Statement 12: I am confused, I am not feeling good lately. Every time I want to sleep, I always feel restless
Statement 13: sometimes what is needed when there is a problem is to laugh until you forget that there is a problem, when you remember it, you feel restless like that well, it turns out that I still have a sad burden blah blah b lah
Statement 14: Because this worry is you.
Statement 15: Sometimes it is your own thoughts that make you anxious and afraid to close your eyes until you do not sleep
Statement 16: Every time I wake up, I am definitely nervous and excited, until when are you going to try ^",
Statement 17: I can only hope, even though I am worried if it starts like this....
Statement 18: restive = restless, restless
Statement 19: "No regrets or grudges/angry at things that have passed, and not worrying too much about the future, that is true serenity."
Statement 20: If the guy is sick, the girl actually feels sick indirectly. He must be feeling restless and restless. :')
Statement 21: I have been the most anxious person before I became the most sincere person, I have also been the most worried person before it all ended
Statement 22: Successfully online at 9am. Even sleeping in a state of restlessness what is wrong with me
Statement 23: Cannot sleep. Just restless by nature
Statement 24: Already worried about picking up the phone
Statement 25: a little restless. a little restless. always feel guilty. sorry if i cannot be the best as i have been :'
Statement 26: I have been so restless since this morning
Statement 27: it makes you feel restless
Statement 28: Sleepy but restless when you sleep
Statement 29: Why is this cave... I am already in a bad mood and then my heart seems to be beating really fast... I am really nervous. Is there something wrong???

Statement 30: He is very busy, He is restless, geez, what is going on?
Statement 31: I am broken and fragile, I am hurt, I cannot stand the anxiety anymore, I cannot stand the love. Until now I was ignored, not seen, even left far away. I am not like what he wants, so I will just step back.
Statement 32: # Suggest a song that makes you calm, do not know why suddenly

this feeling of anger / restlessness appears. Thanks!

Statement 33: Teroosss nervous huuuu

Statement 34: why today's heart does not feel good..jealous?.anxious?.disappoointed?.feeling failed?

Statement 35: It is a habit from elementary school if I think about something I will get worried I cannot sleep

Statement 36: Since yesterday yesterday, I have not been in the mood, very sensitive, feeling restless, always surprised and many surprises every day that almost make my brain explode, everyone is the target of ups and downs in mood. But today I feel very grateful that this lecture drama is over.

Statement 37: Tonight he continues to chatter about money, I am so confused what to do. Once he said do not worry about money, but he kept babbling about money.

Statement 38: Feeling so restless lately.

Statement 39: Oh god I am restless

Statement 40: but my heart is still restless even though my mouth is talking

Statement 41: waiting for uncertainty, please be anxious

Statement 42: been the most anxious person, before becoming the most sincere person

Statement 43: Did you ever read it when you read it to RL, I mean the atmosphere, the taste. It is because I am like this. when I read au, the topic was about mental illness " my days became restless....and now I am scared.

Statement 44: Hi, I want to tell you... Lately I have been feeling restless, have trouble sleeping, I searched on google it says it is a mild symptom of depression, I used to tell my mom a psychologist friend "do not think too much, it is not important you get depressed easily" then I frequent irregular breathing.

Statement 45: I have been feeling restless for the past few days, now I am nauseous, like someone...

Statement 46: Oh no I am restless. The time will come. I walk with Jesus. Then my heart is calm.. Every time I sing a song I do not know, tomorrow in this section I always want to cry . Because I sing it but I do not do it . Sorry God I worry too much

Statement 47: Controlling anxiety is not as easy as it sounds.

Statement 48: Since the pandemic, I have not opened my laptop for a day, I feel restless, I have a very heavy life load

Statement 49: People who are restless and afraid can be seen from the frequent touching of the face and ringing the fingers.

Statement 50: I do not know if my mind breaks down, but I do not want to think about it :(, why do I feel restless?

```
In [123]: # Filter rows where 'status' is 'Anxiety' and get the 'statement' values
stress_statements = data.loc[data['status'] == 'Stress', 'statement'].head(5)

# Print the first 50 statements
for idx, statement in enumerate(stress_statements):
    print(f"Statement {idx + 1}: {statement}")
```

Statement 1: He said he had not felt that way before, suggested I go rest and so ..TRIGGER AHEAD IF YOU'RE A HYPOCONDRIAC LIKE ME: i decide to look up "feelings of doom" in hopes of maybe getting sucked into some rabbit hole of ludicrous conspiracy, a stupid "are you psychic" test or new age b.s., something I could even laugh at down the road. No, I ended up reading that this sense of doom can be indicative of various health ailments; one of which I am prone to.. So on top of my "doom" to my gloom..I am now f'n worried about my heart. I do happen to have a physical in 48 hours.

Statement 2: My mom then hit me with the newspaper and it shocked me that she would do this, she knows I do not like play hitting, smacking, striking, hitting or violence of any sort on my person. Do I send out this vibe asking for it from the universe? Then yesterday I decided to take my friend to go help another "friend" move to a new place. While we were driving the friend we are moving strikes me on my shoulder. And I address it immediately because this is the 4th time I have told him not to do these things, then my other friend who is driving nearly gets into an collision with another car i think because he was high on marijuana and the friend we are moving in the backseat is like "you have to understand I was just trying to get your attention" you know the thing 5 year olds do to get peoples attention by smacking them, this guy is in his 60's.

Statement 3: until i met my new boyfriend, he is amazing, he is kind, he is sweet, he is a good student, he likes the same things as me, my family likes him, and so on... but i do not feel that passion that rush i felt with my ex, the truth is that when i started going out with my boyfriend i secretly saw my ex a few times to see if i really did not feel nothing for him, but it was disgusting, i did not even want him to touch me, i feel bad with myself i did not want him, but still, i was there. then i kind of realized i felt nothing love related for him and it was ok, HE was HURT when he knew i was dating this boy and he even begged me to stay but of course not. but now the problem is that when i am with my boyfriend i do not feel like i love him, like that thing you kind of have to feel with a new love, i just feel "ok" with him, and i catch myself thinking about my ex from time to time, remembering all the good things we had and it drives me crazy because i know that if i see him again i will not feel that way, that "love" that my mind makes me think stills there. and recently i found out that he has a girl and he is actually enjoying the experience and i got so mad and so hurt (i know i do not have any right to feel that way) i felt betrayed and I STILL feel that way, i gross myself out.

Statement 4: October is Domestic Violence Awareness Month and I am a domestic violence survivor who is still struggling, even after over four years. Lately I have been feeling very angry. Angry that my abusive ex received no real consequences for his actions. This man abused me in all manners: physically, sexually, emotionally, verbally, financially, etc. I was granted a restraining order against him (and it was renewed a year later) but I was unable to press criminal charges against him because I did not have enough evidence to have a case.

Statement 5: I think he does not want to put in the effort for the relationship to work (and we are both so difficult that we have to work on our relationships, does not matter with whom) but he cannot be without me either. What should I do? I am afraid this is going to happen over and over again, because I am always forgiving him at some point. Am I being strung along? TL;DR: Boyfriend [28,M] broke up with me [23,F] after on-off for 1.5 years, I thought we just got it together and am devastated...do not know what to do, want to keep fighting but should I?

Statement 6: It cleared up and I was okay but. On Monday I was thinking about humans and how the brain works and it tripped me out I got worried that

because I was thinking about how the brain works that I would lose sleep and I did. That night was bad just like last time. Also yesterday my sleep was bad I woke up like every hour of the night just like last time. I got kind of scared like I did last time but this time I think that this is fake life which is absurd but I just think about it then get really scared then I think rationally then calm down.

Statement 7: I actually give an assistant half my emergency fee to come help me sometimes. I feel unsafe driving to "house calls" in the middle of the night by myself. I frequently bring my husband along as a makeshift "tech" or when I feel unsafe, which is obviously unfair to him. Unhappy Husband: My husband is miserable both with his 60 minute commute and with our isolation out here in the country. We feel like we have no social life out here.

Statement 8: I just feel like the street life has fucked my head up. There is so much I do not even know how to talk about anymore, I just hold that shit. Only person I can really chat with is a pal I know at the bar. He has PTSD and shit from the military bad, hard up alcoholic nowadays after killing people.. We talk once every few weeks and we are open and it is cool. But normal people?

Statement 9: Next week I will be flying for our family vacation. The flights will not be very long (just MSY to LAS then LAX to MSY on the way home), but flying really triggers my anxiety. Mostly I just like having control over myself and my surroundings, so the idea of being in a metal tube 30,000 feet in the air is not ideal for me. I also have a lot of fears about terrorist attacks/mass shootings (movie theaters and other crowded public places are also a problem for me). I was wondering if anyone has any tips for flying anxiety/fear.

Statement 10: Everything sets me off and I am almost having a panic attack. I am going to my dad's soon and that also makes me anxious because I am afraid that I am not in the mental state to go there and I want to stay at my mom's but he really is not happy when I do that. Anyway when she complains or is feeling bad I comfort her and I have never raised my voice at her even when frustrated. When she thinks I disagree with her she raises her voice, I have to talk about what she wants to in general right now because she gets impatient easily. I feel like vomiting because she also has so many redeeming factors but I do not want to lose her as a friend.

Statement 11: I have been taking 12.5 mgs zoloft for about 6 months (yeah, super small dose .. but it was taking the edge off, i think). I decided to stop taking it for a few days to see if it was really helping. Now, about 40 hours since my last dose, I am suddenly kind of shaky and also flushed (my cheeks are burning for no reason. And THOSE side effects are about to freak me out - do you think it is from my missed dose? Is it possible to have such a fast reaction?

Statement 12: No place in my city has shelter space for us, and I will not put my baby on the literal street. (I have been there before. Fine for new, but I do not want to have to go through that with a baby). What cities have good shelter programs for homeless mothers and children? I would love to be able to make some calls and get on a greyhound between the 7th and 16th (when I will likely be able to get the money to go).

Statement 13: These past couple of months have been the worst. My anxiety has gotten so bad it is effecting my sleep and relationship. I have become so paranoid about my health as well. I do not feel like me anymore and I just feel scared all the time now over every little thing. I do not have money to see a therapist either...

Statement 14: I asked him three times what happened. And after the third time I cried and went home. Month later he still do not talk to me and he and my mother started fighting for the first time in the relationship. Me and my mo

ther honestly do not know what to do, he just ignores me. I even told him that I do not date that guy(even though I date him) and he did not even react to it.

Statement 15: The sensation lasted hours. It seemed like a circulation problem, and I panicked and of course ended up in the ER again. This time a doctor came to see me immediately. He wanted to talk about my anxiety. He said he could run some more tests, but he did not think it would help.

Statement 16: It did hurt but did not leave any marks. When I lied down I kept coughing and today it feels kind of constricted like there is a lump in my throat. I have googled and realised how serious it can be and now I am terrified but scared to go doctor. What do I do I am not calling police and I am not going to any hospital. I will go to doctor if they are sensitive about it

Statement 17: he still holds me back in life. he still finds ways to get to me. YET I STILL TALK TO HIM. and i always feel worse after i do. i have tried blocking him on social media, but he still finds ways to get to me.

Statement 18: Through work I have been in some dodgy situations abroad. A number of times my life has been at risk, and I have had to defend myself a few times. I am a civilian and have only had basic weapon training. ​ We have always had an ex-military security contractor with us, but still had to perform aggressive roles when approached.

Statement 19: If anything goes wrong, sexually, I want to cease to exist, and a lot of the time, I am not able to do it out of sheer anxiety. I used to be a very soft person, and I hope I am still kind, but now I am quicker to anger, and I do not want to become an abuser. Basically... was I abused? Does that sound like something an abused person would say? I feel like I have the symptoms of PTSD.

Statement 20: I really do not know what to do about this, and I am wondering if I am even able to get over this fear. A traumatic experience compounded it in me as a kid, and I have been experiencing it for so long that it feels like second nature. It is likely so ingrained with me that therapy will not do much good. I need some help. If anyone can offer some, feel free.

Statement 21: November has been one of the worst months in a long time. My mom was in a car accident and has no idea how she is getting to work. A few days later my grandmother died and I did not even have enough money to attend the funeral. I drove home to be with my family, and to drive my mom 5 hours to the airport. I spent a ton on gas and now money is tight.

Statement 22: Also the headaches. LOADS of headaches all the time. I am so done. I hate this almost as bad as my brain constantly telling me I am a POS. Anxiety is fun :)

Statement 23: I am really feeling like there are no good men. They are all just horrible to women. I think I would rather be alone than deal with any man again. Has anyone else felt like this? Did your feelings ever change?

Statement 24: You can see the detail on my [update 2 right <url> and here is a link to the [project github <url> So I am begging you here, literally begging you. If there is anyone out there who would be kind enough to give me a second chance to fix my fuck up and finish my degree, would you please consider giving me some financial aid via [my goFundMe <url> I am literally in dire need of help right here and I do not know where else to turn other than the people of the internet :(***Kind regards,

Statement 25: Did I mention my parents are religious? I do not know if this is normal for religious people to treat. Whenever I tell them I am terrified of being homeless they tell me I am a "acting like a baby" and "get over it" my parents parents did not treat them this way. They are basically mad because they(I guess 30 or 40 years ago were different when they were my age? Because they said they both lived on their own at 17 and that they find it cree

py I am 16+ and they find it creepy being around me).

Statement 26: But it is been 2 months already this time. We did not speak for Christmas or new year. I am lonely, sad, angry at the situation (not angry at him!) and the worst part is not being able to talk or even know what is going on. We did not fight before this so he is not angry at me.

Statement 27: Sorry this going to be super long, I have a lot to say... Tl;dr... see title. It was constant up and down; one minute we were happy and he was being more affectionate than he has in years, the next he is screaming at me and calling me a whore. I get it, I broke him. I feel so disgusted and ashamed with myself but I truly want to spend my life with no one else.

Statement 28: And it makes me more anxious which causes me to slip up *more *. It is getting WORSE though. I thought I could hold it down, this has been an ongoing issue for about a year or so and it has never been this bad but lately it is getting to the point where I cannot even talk to my own MOTHER without freaking out over this shit. She came back in town from a very long half year vacation recently and I had to go to dinner with her because that is what you are obligated to do and I had to sit there and try and make small talk with my brother the whole damn night looking anxious as fuck... I know they noticed.

Statement 29: Our dog Jett has been diagnosed with diabetes and is now in the hospital to stabilize his blood sugar. Luckily, he seems to be doing well and he will be home with us soon. Unfortunately, his bill is large enough that we just cannot cover it on our own (especially with our poor financial situation). We are being evicted from our home soon and trying to find a place with this bill is just too much for us by ourselves. [To help us with the bill, we set up a GoFundMe <url>

Statement 30: I did not do anything during spring break, except maybe go to work. And even then, I missed a day because I just could not deal with the stress. Most adults get excited to tell you they did absolutely nothing during their break. And there was a time when I enjoyed taking a break too. But now it seems like every time I have a break or a day off from work, it makes the stress worse.

Statement 31: I feel like I will never trust my own judgement again. 3) I fucking HATE being this person. I hate being a victim. I would never judge anyone else in my situation, but I cannot stand having to see myself as someone who was abused. My grandfather abused my grandmother her whole marriage and I always felt like, because of that, I would always be on the lookout.

Statement 32: Hey everyone. I need a place to vent and place my feelings because I do not know what to do anymore. This evening while playing a boardgame with my BF and some friends, I found out my BF lied big time. Short: last year we decided that I would join him this year on his skiing trip with his friends (his idea). This was not possible in the end (or so he told me) because his friends did not want to rent a cabin but go on a cheap student skiing trip.

Statement 33: I have been leaving lights on at night both in the main rooms and my bedroom. The fear is both physical and emotional. I am so jumpy at work and at home. I feel like I am seeing things everywhere and little waves of panic wash through me all day. I am really just finding this incredibly difficult right now and needed to write it out I guess.

Statement 34: Why would you say that to me" and gets extremely offended. He continues on by giving me the silent treatment for 50 minutes and then comes in the room and tells me "You do not understand me. No one understands me. I give up. On everything."

Statement 35: She was already against smoking weed after college at first but now she seems to have moved to a stance where she is okay with me smoking weed as long as she is there when it happens, as in I could smoke with just

her or when we are both hanging out with the same friends. I have tried asking her about it and genuinely understanding why she feels this way about it but she cannot seem to describe it rationally, saying that it is a gut feeling that she gets that legitimately bothers her and gives her anxiety. She seemed very distressed talking about it and was almost at the point of tears saying that she wishes she knew why it bothers her so much but it does and she cannot help it no matter how much she thinks about it or how much I try to explain it to her. I do not know how to explain this to her. I do not want to give up the option of smoking weed with friends but I do not want to break up with her over something like this because I love her.

Statement 36: I am worried I have a blood clot or something that gets aggravated when I wear them. I just want to be okay and to have a good time on this trip, but I have been so out of it, and I am at my wit's end. Right now I am lying down and I feel blood rushing all through my head, and bulging of blood vessels around my nose. I am extremely sleep deprived and woozy but I am scared to go to sleep and am in pain. I am so scared.

Statement 37: I have PTSD from my sexual abuse, and this week has been really hard. Really, since Dr. Ford came out against Kavanaugh and the fallout of it. Today, though, was really hard. I come home to see my father watching fox news and agreeing that Dr. Ford is a lying whore. His words.

Statement 38: And I do not ask for the nightmares to drain me of all energy and make me feel worse throughout the day the harder I try to be normal. I was planning on seeking counselling without my parent's knowledge for this condition once I get to university and get settled, but now I am wondering if I am going to be told the same thing by a therapist there. I am not looking for a diagnosis, just people who share my experience or know what I am talking about. I am sorry if this seems kind of ranty or rambling, my thoughts are a bit of a mess and I am not sure what to think. Thank you for reading all this, anyhow.

Statement 39: i knew gay was humiliation. i tried bringing up to my dad, and he said he would not care if his son was a fudge packer, nor would he care bout that son anymore. i then had to speak to the bishop of my church and tell him about me possibly being gay. my dad was in the back of the room and another man i had never met was sitting next to him. He asked me if i had ever had any gay encounters.

Statement 40: I know every time I have panic attacks to remind myself it is just in my head, but it sucks when little things just make you feel like shit! Does anyone else get triggered by stupid things like this? ? I am even just pissed at myself for now feeling mad at my mom, like do not you know not to tell someone with health anxiety that "they do not look right"?? Ugh!

Statement 41: In class, i am always on edge, i cannot focus on one thing for too long and i constantly scan my environment and I think they think i am looking at them which i am not. this also made it more awkward to talk to the m. I been there for a month but i am still very nervous around my peers and teachers. I know my teacher and the tutor there get nervous/uncomfortable around me, lately they have not been talking to me much or coming up to me especially the tutor to ask if I need help which they did before, my anxiety has gotten worse and I probably look more visibly uncomfortable than ever so that why they have barely asked if I needed help or said much to me ha! yet seem more at ease with other people.

Statement 42: I know this may seem like a small thing, but to someone with these problems, it is a huge deal and to be at a store where you do not feel "safe" around your managers and you cannot talk with them and grow as an employee, what is the point of continuing to work there? So I luckily got them to put me on 3rd shift. Minimal contact and I still get to work for the company. As I did mention, I have been struggling lately with some things: being

slow with completing tasks on the work list is the biggest issue I believe. I am not sure why I am slow, in the past a supervisor has said "Yes you are slow, but you are thorough."

Statement 43: I was consuming quite a lot of caffeine (up to 500mg per day) during the summer and I am medicated with 18mg methylphenidate for ADHD. (I was rarely taking caffeine with methylphenidate together as there are a lot of side effects after doing that, one of them is anxiety and increased heart rate). I do not take methylphenidate now because of the high bp and my chest pains. I have an appointment to the cardiologist but until then I will have all my trial exams. I am only 18 years old, is it really possible to have a chest angina at this age?

Statement 44: But she says she cannot stand hearing it because she is got enough on her plate as she actually suffers with depression herself. I do read other people's posts and think maybe I have actually had it easy. Maybe I am just over exaggerating thing, but i feel like I need to get it off my chest. Anyway thanks for reading. (sorry the post is so long).

Statement 45: I am 24 and have massive amounts of stress going on with job hunting. I live with parents who are constantly on my back and getting angry with me for not having a job. They keep trying to help me by saying "oh this place is hiring and that place is hiring again." I have tried all places where I live and every time I do not get an interview let alone hired. Even if they have a sign up that they are hiring, they tell me "sorry we are not hiring right now."

Statement 46: When I was a kid, I suffered a kind of unlikely injury that basically ruined me mentally, and it seems like I just CAN NOT get away from it, because it is such cheap gory shock value that everyone uses it when they need cheap gory shock value. And if I am not mentally prepared to see it I have a panic attack or just start crying. A fucking jokey homebrew cursed item on r/DND did that to me today, and I hate that my trauma is so bad that I have to stay guarded even when I am just reading a funny thread about my favorite hobby. I am still thinking about that fucking item. I am so tired.

Statement 47: I do not know. Was this okay? Should I hate him? Or was it just something new? I really do not know what to make of the situation.

Statement 48: Any tips for making more friends? Tl;dr: I have had very few friends throughout my life. Recently I have tried reaching out to a lot of people but almost all of them will decline my offer to hang out at some point. I guess maybe this has to do with how I am socially awkward in the past but I have also been told that I am trying too hard. Any tips for making more friends?

Statement 49: Hello guys. This is my first post here. Ever since I cannot fall asleep on a random night (few months ago) , I started getting anxiety at night and before heading to bed but everyday is different. Usually, when I get on bed n off the lights, my body started to heat up and i would sweat a little which makes me really uncomfortable. My heart would pound fast to which makes me anxious.

Statement 50: I am completely devastated. My cat was just hit and killed by a car. I do not think I am allowed to bury him in the garden (we rent) and I do not know what to do. I have searched online and found loads of pet cremation services but they all cost a fortune. I do not have a penny to my name as I am unemployed.

Take a look at some of these examples from the Stress statements printed above...

Statement 15: The sensation lasted hours. It seemed like a circulation problem, and I panicked and of course ended up in the ER again. This time a doctor came to see me

immediately. He wanted to talk about my anxiety. He said he could run some more tests, but he did not think it would help.

Statement 10: Everything sets me off and I am almost having a panic attack. I am going to my dads soon and that also makes me anxious because I am afraid that I am not in the mental state to go there...

Statement 11: I have been taking 12.5 mgs zoloft for about 6 months (yeah, super small dose .. but it was taking the edge off, i think). I decided to stop taking it for a few days to see if it was really helping. Now, about 40 hours since my last dose, I am suddenly kind of shaky and also flushed...

Furthermore, a lot of the Stress statements seem to be talking about physical illness, ailments, emergencies, etc. If we remember back to our most important **bigrams in Anxiety**, they seemed to point to very similar discussions of physical health.

Stress VS. Anxiety

According to the American Psychological Association, stress and anxiety can present with quite similar symptoms...however stress is a temporary state due to external circumstances, whereas anxiety lasts even after external stressors resolve.

Furthermore, **stress can be a symptom of anxiety**.

In other contexts, it might be useful to make the distinctions between the two. However for our agenda...if a user is in a state of emotional disturbance and stress severe enough that they turn to a chatbot for mental health resources...it will be reasonable (at least for these initial notebook algorithms) to refer them to the underlying mood disorder most commonly associated with stress...anxiety.

We'll combine these under *Anxiety*.

(There are other mood-disorders and psychological conditions that can present as symptoms of stress...such as the manic episodes associated with Bi-Polar, or symptoms of PTSD. Given our examination of the 'statement' values categorized as stress, we feel comfortable assuming the 'stress' identified in this dataset is more closely related to anxiety...However we want to acknowledge that **before any next steps are taken to develop software from these models, and before any actual medical or psychological advice is given to real human subjects, these metrics should be thoroughly re-evaluated**.

Depression VS. Suicidal

Let's examine some statement selects from *Depression* and *Suicidal*. These are more noticeably distinct compared with Stress and Anxiety -- as *Suicidal* implies a very specific threat of action that neither Depression, Anxiety, nor Stress do...However we'll double check to make sure this specificity of distinction is relatively clear in our statements.

In [124...]

```
# Filter rows where 'status' is 'Anxiety' and get the 'statement' values
depression_statements = data.loc[data['status'] == 'Depression', 'statement']

# Print the first 50 statements
for idx, statement in enumerate(depression_statements):
    print(f"Statement {idx + 1}: {statement}")
```

Statement 1: I recently went through a breakup and she said she still wants to be friends so I said I can try doing that but when she talks to me about things it always hurts. I just want to lose feelings so all this pain can stop it hurts so much and I cannot even cry about it. I do not want to hurt her because she said she does not want to never speak to me again but I do not know what to do here. When we were together she always hurt me so I do not know why I still love her. I wish we never met it would be much less painful How do I lose feelings?

Statement 2: I do not know how to navigate these feelings, not that its a new feeling by any stretch. I just do not understand how I can go on from being so pleasantly fine to suicidal and self loathing within in minutes and not be able to step out of it. I have been suffering from poor mental health for over 10 years and I am about to turn 26 next week and I feel so ashamed for how serious my negative thoughts are. Covid was a hard transition especially having moved to a new city a few months before everything started locking down, but now I am here and settled and I have a loving partner, a great connection to their family, a really fun job and I finished my degree. But when I stop and look around I feel like nothing will ever be good enough for me and I will never be good enough, I will always have skin problems and bad teeth and acne. I will never be the artist I used to be (I have a Bachelor in fine art) because I do not have the same drive or passionate that I used to have for it. I will never have the career I built my entire plan around because I am losing my interest in art. I will never have the relationship with my family that I want because I could not stand to stay in my hometown. All I can feel is this tornado of my failures swirling around me all the damn time. It just seems that no matter how many positives come my way they do not last long enough to out weight the negatives. I do not really know what I am looking for I guess I just wanted to say this out loud, or whatever. 0 to 100

Statement 3: So I have been with my bf for 5 months , and he already told me he was depressed. To this week nothing particular happened, but i can now feel he is bothered by it. He does not ask me to sleep together, do things and etcI know I am not the problem I am aware of that, I just wanted to ask: how can I be there without suffocating him? I am really clingy, a thing which he usually loves about me but right now I kind of have a feeling that I am doing too much.Any suggestion will be gladly accepted, be frank and tell me how I can not make it worse: its that that bothers me. My bf (21) is having a bad time due to his depression

Statement 4: I have been severely bullied since i was 5 till 15, this resulted in me being a depressed misanthrope with ptss and multiple anxiety disorders. I am now 18 have no friends,or social contacts. my study will be online so i will not meet people there, i do not sport since i have poor health because i was sick when i was a child. I do not drink so going to a bar to meet new people is not it, i do not like party's or festivals since its too crowded and filthy, dating sites are shit since most people that are on it are superficial, and even with the rare occasion i get a match and message them i get no response, I am just done. I have wonderfull close family whom i love a lot but, i just do not want to deal with this constant loneliness, does anyone have some recommandation of which activatys would bring me more into contact with people(i do not have any hobbies or special intrest besides my obsessive facination about war and the reasons etc) How long am i suposed to go on until the "better" days come

Statement 5: My mom made me go to a camp that she knows I hate. Now I hate most days the only good time is at midnight where I can think to myself, but I do not want to call CPS because she does small things but they build up and I do not want to be taken away from my dog. So does it get better? Does it get better

Statement 6: Help me for ideas simple healthy meals to make when you feel depressed Simple meals when feeling depressed

Statement 7: Been stuck in a loop of ruminating/ overthinking. I still can't seem to connect with another person I find attractive even though that is the least of my problems right now. Just overwhelmed with what is in front of me and trying to stay calm but imposter syndrome is ringing in my ears Low

Statement 8: Since I began seeing a therapist 5 months ago I have realised the thought of being happy genuinely terrifies me, and I do not even really know why. I think its probably caused by multiple things. If I had to guess, I think I just do not feel like I deserve to be happy. I have hated myself for so long that maybe I have found comfort in my misery, like its my punishment or something. And in a related way, I think I am scared of my happiness being selfish. Having been miserable for so long, I think Id feel like a selfish asshole if I suddenly got over my issues and starting liking life when I know there is countless people in the world still feeling how I did. And furthermore, its only in the last while have I really begun to take my problems seriously and have been trying to be more sympathetic to myself, but even now with all the shit I am going through its like I can still barely take care of myself. And its like I am afraid of losing the various anxiety issues and what not that make me miserable, because if I lose only them and not my self-hating mindset, Ill have even less reason to sympathise with myself. Ill be some privileged selfish asshole and be even harder on myself than I have ever been. Lastly, I think I am just afraid of the thought of living in general. Death just seems soo simple and peaceful, and I think I have become fixated on the idea of dying. Having been miserable for so many years I think the idea of any more life terrifies me. because even if I did make an amazing recovery, Ill still get new problems, which could be just as bad or worse, and I could also relapse and redevelop all my anxiety and insecurity issues. In fact I have before, in 2018 I made huge strides and then in 2019 it all turned to shit again. And a lot of it was because I had one panic attack in front of an entire class in school, in front of most of the new friends Id made in 2018 who did not know I was an insecure, anxiety ridden loser with no friends before I met them. It pretty much shattered any progress I had made, and thereafter I had multiple panic attacks in school, making it even worse. Those are my interpretations anyway. Whenever I feel like I am making any progress on myself it sends waves of panic through me. I hate it so much, and I wish to fuck I knew what was causing it, because these are all good guesses but I still do not know if that is all there is to it. I am genuinely terrified of being happy and I do not even really know why

Statement 9: Did your depression start after a time of serious stress/anxiety? If so, this may be why you feel so disconnected. After times of intense stress, where we do not give ourselves any leeway or breaks, our body naturally will downregulate hormones to essentially force us to stop what were doing. So we may have super high levels of glutamate and norepinephrine keeping us in a stress response, but low levels of dopamine keeping us in a depression, and the high amount of anxiety neurotransmitters because us to feel both depressed and depersonalized (freeze response) If you feel disconnected from reality this may be why

Statement 10: I am going to make this short because i hate thinking about it, but long story short, I am getting better after months of medication, therapy, etc. My anxiety is at an all-time low and my antidepressants ""do not allow"" my mind to wander too much into suicidal thoughts. But i hate it, i hate that I am getting better, i feel embarrassed whenever i catch myself doing something not depressing, like taking a walk, laughing, or doing stuff by my own will. And i feel SO SO selfish, i was given immense support from everyone, my mum has spent so so much time and money on therapists, psychiatri

sts, pills, etc; she is suffered more than anyone else, even more than me since i seem to enjoy being depressed. she is the happiest she is been in months since she can tell I am in a better mood and she is the only reason i have not "tried harder" to get worse, but i do not know how long i can keep it going. Something in me just wants me to maybe stop taking my meds without letting her know, get a new self destructive habit, and go back to the old me that did not leave his bed in days and could not keep a conversation for more than 2 minutes, i really miss it. I am going back to school in a couple of months (i dropped out), i started driving lessons yesterday, I am eating better, sleeping better, I am starting to live and behave like a regular person but i cannot handle it, it makes me feel bad, gross, EMBARRASSED, very embarrassed, and its just a cycle of: I am ok this makes me depressed I am ok again this makes me depressed etc. I do not know what to do, I am not even sure why i am even writing this since I will probably end up ignoring whatever help or advice i get, but i guess i needed to vent, thanks for reading and as always sorry for my english how can i be happy about getting better??

Statement 11: I have been talking to someone who I really liked. They were a great person and we really clicked. I told her that I could not do this with her because I am dealing with so much shit in my head that it is not fair to her. She said she wanted to be there for me and I shut her down. I am scared to be depressed around other people now. My poor mental health ruined one relationship and I do not want it to ruin another. I know I am much better now than I was in my previous relationship but I am scared that I might be a horrible person again. I am scared to let anyone be close. I feel like I am just getting by daily and cannot make any real connections anymore. The ability to get close to people is non-existent because I am scared that when they see the real me, they will leave. I am trying to improve but what if I cannot, what if I fall apart again? What if I am the piece of shit suicidal asshole that I was before? I am scared to put anyone through that. I kept making up excuses for why this was not going to work out. Distance, what we wanted, culture, but truthfully I am just fucking scared to be close to someone again. I am scared to be vulnerable. I wish I knew how to not be depressed. It controls my life. I hate it I am pushing people away.

Statement 12: Not because of death, but because she wanted to get away. I have been depressed for over 10 years and maybe this is the lowest point I have ever been to. The last few days it seems like there is been an invisible force squeezing my chest all the time. I can barely eat. I spent 48 hours with only 1 sandwich in my stomach and after that I have forced myself to eat just to be able to get drunk after I get home from work just so the time goes by quickly and I do not feel this damn pain. She was my last friend. I am completely alone now. Losing people is always a disgrace to me, but ok. I have known this friend for 6 years, and even though we had a few periods without talking, we were doing well for almost 2 years, we never even had a fight until last week. Turns out I was pretty drunk and I ended up insisting on telling her that she did not really care about me and how I was a nobody to her. I really insisted on it, because I was drunk and could not even think straight, although she said clearly that she did care. Turns out she was hurt and does not want to talk to me anymore. She was the only person who really mattered and it made me want to keep living, I did not want to bring her pain. I tried to apologize, tried to ask him to understand and forgive me, but to no avail. The worst thing is that I did not even want to attack her the way I did, it is very easy for me to think that someone does not like me, after all I myself think I am rubbish and in no way worthy of any love. I did not even sleep the night that happened. I spent the whole night thinking about how I could kill myself. what is the fucking point on living if my life is a piece of trash, and I feel bad absolutely all the time? I suck at everything and have

no friends to have fun with. The ones I had I ruined it. How to deal with the pain of losing someone when you are depressed and that person was one of the pillars of your life?

Statement 13: To me, its all I think about when I think about what I lost, its just a matter of when and how at this point. I have never felt that I want to do it this much before

Statement 14: I have felt this way for so long I have always assumed it was normal- not so much that I was necessarily sad, but more often just numb? Like I never really feel joy or excitement, although occasionally (I distinctly remember the occasions) I have noticeable sense of peace that cuts through my general sense of heaviness. I am not sure if this depression, and I have not wanted to face it or label myself as such and be forced to deal with the reality of it if it is. I think I am having a hard time differentiating being depressed due to an internal chemical imbalance that needs to be reconciled, or just having a natural reaction to the state of the world or events (previous abuse, climate change, corruption, realizing you likely will not grow up to be what you thought you would be, the passage of time and loss of loved ones etc), or if that is a distinction that even matters? Depressed?

Statement 15: My ex and I broke up after 3 years a few months back. I tried desperately to get her back after realizing i really did still love her. It failed and she is done and she has our cat we got together. I still love them dearly and I know ill probably never see them again. I finally got the courage to go through my phone and try to erase the photos of the cat. I just cannot do it. I have spent an hour going through the gallery and snap chat. I have just set them to hidden. It hurts far too much. I have been suffering through depression after realizing what I had let go, what I had lost. My little family. These photos are almost the last moments and belongings I have left of them. I am trying so hard to heal, but it hurts thinking ill need to forget our kitten. Does anyone else have any of advice for this? It really reopened a wound that was starting to heal. Maybe I just needed to vent for the day I had to go through all the pictures of my ex and kitten today

Statement 16: Hello guys , i m 21 with high eye pressure and currently on lumigan . I recently get a few physical conditons and probably partly because of my mental diaorders (doctors says that mental health affects esp. Some parts of our body , i have depression , anxiety and ocd by the way) my doctor wanted to prescribe me cybalex but i know these drugs triggers glaucoma , so is there any of you who has been on psychological pills and do not get any side effects/? worsen you are pressure etc.? If so pls help me , is there any psychological medicine that does not worsen high eye pressure ? Having mental illnesses and glaucoma

Statement 17: Not a day goes by that I do not break down in tears at some point. 2021 has been the worst. Lost my job, my Aunt, and my stepfather. Now I am anxiously awaiting the results from a chest x-ray. My nerves are shot, I am stressed, I am struggling to pay my bills, keep my car on the road, still trying to figure out how I am going to afford the new tires I need while not missing a payment. My credit card has been living about \$100 away from maxed for the last few months. I am starting to think everything would just be better if I did have lung cancer, then when I am gone I can stop worrying about money, and happiness, and I will not be such a burden on my family. My wife and I have talked about what is going on, but I have kept a lot of it from her, she sees me crying and does her best to ignore me. I am sure she resents me for being such a useless waste of space. When I am gone she will be able to find someone she can actually be happy with, guilt free. I just, I do not know, I guess I just needed to say it, not expecting advice, or help, just wanted to get some of these ramblings out of my head. Feels like I am all alone in this

Statement 18: I feel like I am at the end, nothing I do is ever right, I am stupid and worthless. I just do not see the point in living anymore I do not know if I can keep going...

Statement 19: I have not seen my 7 year old daughter in a couple of months because she moved across the country with her mom, and it has literally been killing me. My mind was telling me our relationship was damaged forever because of this. She is the only bright spot in my entire life. I have just been alternating between drinking and crying, mostly. Today we connected in an online game and had a great time, we are hopefully going to be playing regularly from now on and I will get to talk to her and bond with her through this. It is a pretty basic online game for kids but the first thing she did when she met my avatar was to shrink her own avatar into a little baby and jump on my shoulders like she did when she was a toddler in real life. I literally saw tears. I have also agreed with her mom to fly over there and spend a week or so with them soon. My life has been just a black hole for many weeks now and it is so incredibly refreshing to poke your head above the water and see at least a glimmer of hope and sunshine. She does not deserve to lose her dad and I now have new motivation to push on and try to get better. If a terminal case like me can do it, so can you, hopefully. My daughter just pulled me out of a black pit of despair.

Statement 20: I cannot seem to go a couple of months without self-sabotaging myself. I do not know what comes first the depression or the self-sabotage but they work hand in hand to make me absolutely miserable. For the first time in two years I am considering going back to therapy because I just cannot live like this anymore. I cannot keep repeatedly, just making things unreasonably hard for myself. I deserve happiness and a chance at success just as much as anyone else. I just have to figure out a way to actually believe that, to put my words into thought and action. Self-sabotage

Statement 21: I cannot fucking feel a single fucking thing man . I bottle up every feeling and I am so far away from everyone in my life . I have no friends or anyone I can talk to , I feel like shit right this now , I do not want anything in life , I do not want to grow old , travel , make friends or whatever the fuck everyone says we should live for . I just want to end it , I wish I died in my sleep , everyday I wake up I feel shock and sadness . I did not choose to be born so why do I have to suffer for people who do not give a fuck about me . oh god !! just kill me nowMusic used to make me feel better (it was the only thing) , but now even music does nothing . I cannot tell anyone about how I feel , I feel stuck I cannot take this anymore . My parents think that I am very lazy and stupid so anything I tell them would be because I am lazy . Like I told my mom n dad about how I do not find anything interesting while talking about career and they said it is because I am lazy and stupid that I think too highly of myself and that I am overconfident . They also said that I am useless and would not accomplish anything in life . when actually I have low self-esteem and zero self-confidence . how am I supposed to tell them about my depression that I have got because of neglect and abuse from them in childhood . Like I tell them that I do not feel like doing anything , then they would just say it is because I am lazy , you do not want to do this then you are fucking lazy and stupid "I just want to die in my sleep , that is the only thing I want . I feel so alone , I wish I had someone to support me emotionally The worst thing about all this is being alone

Statement 22: I am not excited, I am not happy for him, i just want to get this over with and get on to the next day My brother is getting married today and i feel absolutely nothing

Statement 23: cut onions so that i could get some tears out, since i cannot seem to cry and she would a single tear even though i feel like I am on the

verge of crying for hours each day. it worked though, for the first few tears. but then nothing :(hit a new low recently

Statement 24: I am only 21. But everyone around me just tells me that I always act so serious and sometimes even dress too conservatively for my age. It.. honestly hurts my feelings. I grew up in an environment where I had so much pressure put on to me to take abuse and not complain about it, and learn to do everything on my own. I am not claiming to be more mature than people my own age but I am just tired of people making fun of me for not being more "fun" i do not know, it seems something dumb to complain about but it just makes me feel sad today. I get told I act and look like an old person

Statement 25: Yeah, this is a question I have been chewing on for a long time. I have been in the depths before, as I know you all have, too, but even in those deepest depressive periods, I still wonder how I know for certain that what I am experiencing is depression. It being of the subjective mind, it seems like depression does not have strict dimensions or even definitions. it is hard to pin down and say for sure, one way or another, "I am depressed." This has been an obstacle for me. Trying to take it seriously. I doubt that what I have is depression. I doubt my own ability to objectively judge this. I doubt my problems are problems at all and this just makes me feel selfish and kind of like a loser that has to fake a mood disorder to get along in the world. The ambiguous nature of depression, other mood disorders, and the feeling mind alone is enough reasonable doubt to because even more suffering in me, like depression that is accompanied by a feeling of guilt and self-deprecation. Is this normal? Part of depression's deception? Or are there ways to draw a line near-ish the middle between depressed and not depressed? We can do it to a very fine degree with physiological illnesses. Surely we could do the same for the mind, but are we there yet? How do you know for certain that depression is "depression"? I invite you to share your information and opinions.

Statement 26: If you could understand, maybe you would see why I am so angry, so hurt, so destroyed. I know from your perspective, maybe you see someone who is hurting, maybe you see someone who needs help. But if you could look beyond my eyes, behind my face. The constant struggle to just be alright and it feels like you are completely alone because you actually are. Alone in this desolate wasteland with crumpled buildings, decimated roadways and highways. Twisted metal signs, relics of the past. With old directions guiding to better places than here. The ghosts of yesterday yelling out from the corridors of collapsed buildings as the wind whistle through the burned out city. The sun glaring down as dark clouds approach, gusts of wind kicking up large plumes of dust in the distance, heading for the city as lightning begins to strike and bombs begin to drop. Within a moment you are surrounded by flashes of memories all over, all around you and in the winds you can hear the faint voices, some screaming, some crying, some are my own voice filled with regret and sorrow and loss. The flashes are so intense that the images of memories begin to bleed together as you see my life unfold in bits and pieces and for a moment, just a brief second, you see me for who I really am but I whisp away into the surrounding debris just as quick as you can blink. As though it were just a ghost, just a projection of all the things I wished I could ever be. Strewn to the wind and shattered into a million pieces as you look onward trying to make sure you were not just imagining things. The storm still flickering lightning and gunfire sounds in the fractured buildings echo around you as all sunlight fades and just as quickly as it came, the storm has passed. you are left still wondering if everything you saw was real or not. That image of me, the forgotten dreams, the screams. How much was real? How much was just your mind trying to put the pieces together? All these

questions and doubts, like the particles of dust now drifting further away. Like a candy wrapper caught in the wind, the idea of ever knowing dies along side the winds as they begin to settle and the sun blares bright again. You put your hand above your eyes to shade them from the sun as you look to the sky in some hope if seeing more storm clouds. They say every cloud has a silver lining but none of these clouds seem to be anything but black, darkness, harsh. Until you realize that you do not have to be here and your sight zooms out, slowly revealing my eyes and the pain they hide as you travel back behind your own eyes and see this figure in front of you, torn by war and so much pain. Not quite human but the figure tries. Just as quickly as that image popped up during the storm then dissipated, the notion of who this creature is in front of you begins to blur. A hollowed out shell of all the grand things that could have been. It is then, I think, you would begin to understand that there is nothing left here in front of you worth salvaging and so it would just be easier, simpler, to walk away. My bad place

Statement 27: I literally feel so worthless I have written and deleted three posts in the last hour and a half. None of them were worth posting and I actually still believe that. Anyway, none of you guys better off them selves! That goes for the ladies too! . And all the freaks like me. A stupid reason to post

Statement 28: Today is the last day of being 27. To Be Honest I always thought I will die way before 28. Life is not great right now. I still have A LOT of issues with my depression and taking different types of pills, most of the days I do not have the tiniest bit of focus and feel mad and upset for no reason. Still, I managed to keep my job and my alcohol consumption under control. It sucks but I also have a big exam tomorrow right on my birthday. Wish me luck Going to be 28 tomorrow, never though I will make it this far

Statement 29: I have a close friend who has suffered from periods of depression for years. she is been doing well for the last year, but for the past month has clearly been struggling. she is lonely and does not have anyone else in her life with whom to talk. When we do talk, she has been dropping obvious signs that she is aware that she is become depressed, but when I try to push her to go to counseling, she gets annoyed and wants to change the subject. She admits that she should go, but has no motivation to physically go. And she also has doubts that it would actually help based on past experiences. My question is, how hard should I push her to get help? Should I keep bringing it up even when she says she does not want to talk about it? Should I tell her that I will not talk to her anymore if she does not go? Or should I give in, ignore her depression, and only talk about light topics.. maybe at least providing her social stimulus and relief from her own thoughts? How much pressure to apply to a friend to get help?

Statement 30: So i do not take it for BP, i take it for PTSD, OCD and migraines and so far it appears to be lifting me out of the depression that comes as a symptom of those conditions. Instead of wanting to just lay in my bed all day and shut myself away, I am actually wanting to do stuff and lay with my cats in the living room and interact. I have like 1 to 2 tiny depressive mood drops on it each day but they last literally like 5 min and then it brings me back up Lamictal is pulling me out of this

Statement 31: I am 19 male. Both of my parents where immigrants my dad died in a car accident and my mom crazy and killed herself shortly after when I was 8. My only family left was my grandma. She raised me and I fucking love her so much, but she is old and has Alzheimers. Its so fucking hard seeing her drift away. I should be grateful that I was given such a loving and awesome grandma for the time she was there. I just do not know what Ill do when she forgets me. My depression has really gotten bad I do not want to take paxi

l because I am scared of the suicidal thoughts. Any tips? What do I even do? Statement 32: I have been reaching out for help. I do EVERYTHING that I am told to help me feel better. Well today I called a place about the sexual abuse I went through as a kid. A charity that gives therapy to people who have gone through that. Well. Because I was not raped, it is not deemed bad enough. I was simply molested for 2 years by a neighbor as a kid so they do not think I can get help from them. I was referred to this place by a NHS mental health team who keep referring to these places. One place will tell me I am too severe to help. And now, this place says what happened was not bad enough. I was not even abused the right way. Even my abuse was done wrong. Which is the SECOND time I have heard it this week, about TWO different types of abuse. And then people have the cheek to ask me why I am so miserable. I am literally asking for help and doing all the rubbish I am told to do, and yet here we are. Damn. Just damn. (please no "it gets better"). That makes me feel worse EVERY TIME) I was just told my abuse was not severe enough

Statement 33: It always feels so peaceful the first 5 seconds after waking up then it all goes to shit, anyone else experience that? Last night I had someone over to help me take my mind off shit and it should have worked but it did not. I should be on top of the world but it does not feel that way. I just got pussy why am I so fckin empty? I am so fckin down. And now I feel like I have used someone and they do not deserve that. I miss my girl so fckin much. My mind keeps going to this picture of her laying with me eyes closed and I am just running my hand through her hair admiring how beautiful she is. Fck I miss how it felt when I kissed her. What do I do with this energy. I cannot get her back and I am left with this feeling that clearly I cannot put into someone else and be happy with it. It just sits with me waiting till I am alone so it can Fck me up. I cannot get through my head that it was not going to work with her its like I am going through the grieving process again and I am at denial. This weighing too much on me. Pray for me you all its one hell of a morning Still empty

Statement 34: So I am 22f, I have a great and supportive family and a boyfriend whom I met in University. I am in my last year of university and the pressure to figure out what I should do after getting graduated is getting to me, since the pandemic started we are having classes from home so attending classes is not a big deal. The problem is I do not feel like doing anything I just want to lay in bed and do nothing. So I do not do anything which is making me guilty. I have a lot to do to be honest searching for an internship. focusing on my subjects, deciding if I want to go abroad for further studying if I decide to abroad the admission process and exams I need to give, there is a lot going on in my head. When I sit down and try to do anything I feel anxious like I am missing out on something and thinking about the things which I could be finishing. and I end up not doing nothing and just browsing youtube for hours. I just do not know what to do I feel lost and anxious. Does anyone has any advice? I feel guilty

Statement 35: I no longer look forward to ANYTHING. I have a beautiful girlfriend, a caring family, job opportunities, talent, looks, personality, but inside I am miserable and have lost my love for life. I spent about 3 years off my life going through all the popular antidepressants, benzodiazepines, etc. but none of them worked. I no longer drink alcohol because I abuse it and I no longer smoke weed because when I do, I smoke non-stop and get even more depressed. I am at the end of my rope and I no longer know what to do. I am not going to commit suicide but I just wish I could find some hope. Any suggestions would be appreciated I am 27 (m) & Everything in my life is picture perfect from the outside, but my depression is eating me up inside day by day.

Statement 36: I have just been sitting on my couch and I am just feeling los

t about what to do about my dog because he was ran over. Like when I think about him i start crying to the point I cannot even breathe right and i miss him so much I do not want to do anything after he died ion really have much people to talk to about this and I talked to a friend and I feel better about it but I am still crying every single time I think about him or look at his equally sad sister. I have never been the type to be sad and gloomy but this broke my heart too much and I do not know how to handle it on my own. I am just confused as to how it happened why it happened and I do not even know how to feel about this. I am lost and confused

Statement 37: 19 years old, male, from the Philippines. The lockdown here has been going on for 2 years and more now. It completely sucks here. Not only the government, but also the parenting style and community. My parents are giant narcissist, gaslighting bastards, apathetic, scapegoating, victim blaming, and abusive people. I grew up as a kid being parentified by my mom, while experiencing how my dad openly cheated on my mom while wasting the money we have. I did not feel like I grew with a father because he never tried to form an emotional connection with me. When he does, it will be him giving something while saying he worked his ass off for it so we should be grateful for it. He has forever glorified himself having a job, and he will use it in any arguments we had; "I am the one who has the job, who feeds you, and who gives you everything. Be fucking grateful!" in his words. I would be beaten up by him with a belt as a kid so it has left me traumatized by flinching at any thing that is loud, surprising, or both. He has not changed from my childhood until now, except it has only gotten worse as time goes on. He has endlessly made me hate him with no possible resolutions. Calling me a failure, useless, good-for-nothing, and disgrace of a son does not help at all. He is the reason why I despise cheating, abusive, and manipulative people with a passion. His god complex shows to everyone else. It does not help that he is a cop so his temper is extremely bad. He has a fragile ego that will automatically collapse when told he is wrong or disagreed with. He also thinks he is doing our pets a favour by caging and leashing them their entire lives. he is truly an abusive piece of shit that does not deserve a family of his own. he is a manchild that accidentally had children. As for my mother, she is the epitome of gaslighting, victim blaming, and narcissism. You literally cannot talk to her without her taking it personally. Like my horrible father, she also thinks the world revolves around her. This is the person who made me go through parentification as a fucking 6 year old. Until now, I have been her ownpour of traumas, pain, bad memories, and rants. I am literally exhausted from everything because of this. And I HAVE to listen to her or else everything turns bad; she manipulates my father to teach me a lesson. She also believes she cannot be wrong. She strongly believes feng shui, the supernatural, superstition, and the occult. She has made our entire lives revolve around superstition. I have been gaslighted by her for who knows how much; I do not even know if my feelings are valid anymore. She would try to have an emotional connection with me, but she would just make it about herself after. Literally goes and minimizes my experiences just because she thinks she had it worse. If I have not said it before, these two are very religious. They would rather pray and complain instead of doing something. They are fanatics of the horrible president we have, and they think everyone should respect them. As their child, I am riddled with face and body acne with no help at all to relieve it; father had extreme face acne, mom has extreme body acne but no help from them. Instead, I am called out for being careless about my body and face by both of them for years. They also told me that it will just go away anyway "like theirs", so no medicine is needed. Due to it, I experienced a lot of bullying in my highschool years because of my acne. I could not form proper connections with others and often dressed weirdly to hide my body. I was

called the "Pimple King" by my classmates; I wanted to kill myself so much from it because I had NO ONE to talk to about it. Fast forward to now, I have fought and kind of got some freedom, but here I am again wanting to fucking die. I had already wanted to die from my highschool days about 6 years continuously. I had attempts but none worked. I am doing everything in the house, yet I am called a fucking failure yet again by my piece of shit father for something I did not do; our cat accidentally unplugged the doorbell ringer, which my father uses to call us when he needs us for something, in our room. When he realized there was not any doorbell noises, he checked it out and saw it was unplugged. He called me a failure for it because "I do not care about anything and do not have a sense for anything". These days, I cannot find anything enjoyable, interesting, or fun anymore. I just want to die. I am trying to find something to do, but my brain just rings out "what is the point?" and I stop. I am just in bed right now wanting to fucking die. Someone take my life please. Please. [LONG POST] Someone kill me please

Statement 38: Boyfriend deals with depression. Is not wanting to be alive. Affecting our relationship. I have tried to understand but I cannot see his position. Trying to be there. Please any tips I want him to be happy and be the best partner I can be to him but it seems like its lost. help

Statement 39: I have a good life if you were to look at me from the outside, but I am rarely ever happy. I have noticed lately, that when I am on social media and see something not ideal or bad happen to someone (unplanned pregnancy, breakups, etc.) I get a good feeling. Of course, there are some instances where I am happy for family members and friends who have positive things happen to them, but if its someone who has either ignored me (especially if the opposite sex), I get an especiallyyyy good feeling when I see bad shit happen to them. I really want to work on this, but it is not something that I have dealt with much before and it makes me feel like just a bad person at the end of the day. Any thoughts or similar experiences dealing with this? Thanks. Is this Normal?

Statement 40: I finally got my depression well under control at the end of this last year. I was managing an exercise routine, cooking lots of healthy and fun meals instead of junk food, spending more quality time with my family, enjoying hobbies again, thriving in school, and managing a self care and hygiene routine (which was always one of my big struggles). I felt like I finally got things under control but I made a big mistake that is sent everything crashing down. I took on extra responsibility and I put too much on my plate. I would rather not elaborate on what it was, but now I feel like I have no time for myself so all of my routine went out the window. I had to do a late withdraw from my last 2 college courses when I had maintained a 3.9 prior. I do not shower or brush my teeth for days at a time, I find myself binge eating and falling back into restrictive eating disorder habits, I only really look forward to eating and getting in bed every night. The Fall semester is set to begin and I do not even think I will be able to do it. If I am lucky I might slide by and pass but I am going to suffer even more with that on my plate too. Why the fuck cannot I be a functioning person? Sometimes I want to move away from my family and just let myself self destruct to my hearts content. I feel like I live to eat and sleep again

Statement 41: It always gets worse I have no friends, nobody can stand me, I have horrible acne and I look absolutely disgusting Terrible social anxiety, it is only been getting worse It's not getting better and I wish I had the energy to kill myself I used to try a lot, Meditate and do yoga and eat healthy but it never did anything. I am just really, really tired now Been depressed for 7 years

Statement 42: There is almost 100% chance that my situation is due to all the toxic and stupid people around me. I am cutting everyone off and disappear

ing to start a new life for myself. Wish me luck I decided to cut off everyone I know from my life

Statement 43: A week ago my overseas boyfriend of 6 years went to prison. This means no contact for 18 months for us. I Do Not Know how I have coped honestly. There is been days where I do not feel like living and are never ending. I am young I am 17 and he is 19, but this is the man I want to spend the rest of my life with. We are not able to call due to me not being 18 yet and I cannot go over seas due to covid and financial situations atm. I knew it would be hard but I never thought it would be this hard. I have started sleeping all day, not eating, falling behind on class work and constantly crying/panic attacks. I miss him so fucking much every single day, and dread the 17 months ahead of us. There is a possibility he could get out in 12 but I do not want to get my hopes up. I am so grateful it did not hit the 2 year mark or was not a sentence like 5 years but its still so hard. It constantly feels like I do not want to live anymore but I am trying hard to hold on for him. I have not felt this pain in a while. Heart ache more like heart attack.

Statement 44: Quick background; I suffered social anxiety for most of my 20's and went in and out of depression throughout the years. I am now 31 with no friends, just a boyfriend and all my family who I barely get along with anyway live in another state. Over the last few years I have learnt to be ok with feeling lonely and being by myself. I have learnt to enjoy my own company and that is great and all but I am still friendless. I also went through a horrible year in 2019 where I went numb to something that was happening in my relationship so I would not "feel" so much kind of like depersonalization/derealization. Things have gotten better but the numbness has remained. I am not sure how to explain it properly but I do not have anything in my life that is thrilling or exciting to me nor do I care to seek this out. Most people have a vice like shopping, gambling, some sort of sport or activity that gives them a rush. I used to want to travel, I thought maybe that could fill my time or get me excited but even though I have the money and could travel (once out of lockdown), I just do not care for it anymore. It is like, I can not take the memories when I die so what is the point? I could sit in my lounge room and street view places in Canada for example so once again, what is the point? All I do these days is work, come home, click around on some sites and do it all again the next day. I barely talk to anyone online. My relationship is boring and stale and even though we could spice it up, sometimes I just do not care to. I love my boyfriend but I want someone else to talk to and hang out with. I almost have no reaction to things. If I won a million dollars or even 5, I think my reaction honestly would be "meh". Everyone my age is either having kids, married or have successful careers. I have never chased success so I really do not care to climb the career ladder or chase money. None of that matters to me. I am not sure what does. I cannot help but feel as if I am wasting my life away. Sometimes I wish I could have a crazy life. Sometimes I get the urge to do something so out of character to shock myself. To get that thrill. To feel excitement. God, I am so utterly bored with my life. I feel numb about most things. I feel very little excitement or thrill in life.

Statement 45: I grew up being a very bad kid, I smoked weed from age 9 first time I got drunk I was 13, since childhood we moved from place to place, never had my friends as we always moved, since I was 18 I moved away from my parents to another country and 3 years later, still cannot make friends because people are so fucking antisocial here, I lost my first son after 3 days of birth due to placenta separation, and now my job is pissing me off, I want to break up with my gf but I love her, money is a big issue and recently I have been having suicidal thoughts and other thoughts, is there anyone that has any tips? Tired of living, weird thoughts

Statement 46: I do not know what happened. Over the course of one night I went from optimistic to questioning my reality. I am a 27M and it seemed like I got under eye wrinkles within a night. I do not know if I did not realise it before but I certainly did the morning after and the revelation hit me like a bus. I look old. Older than a 27 year old. I am single, I have no career, no degree, no money and have nothing going on in my life. But the odd thing was before that night I was still hopeful about things turning around in my life. But as soon as I saw my wrinkles that hope just seemed to vanish into the creases under my eyes. I realised I was no longer young and the thought of getting old alone is scary. Finding a partner was hard enough but this will only make it harder. The thought of aging is depressing me.

Statement 47: tw talk about eating disorder further downHi, sorry for long post and bad formatting. I do not know who to ask this as I am on a long waiting list for a psychologist and have no one else to talk to.I have always had problems with breakfast, mostly because I have trouble waking up and am so tired that the thought of food makes me nauseous. Before work (I am currently on sick leave) I used to force down a sandwich only because I know I faint easily. Work also helped me maintain a somewhat proper meal schedule as I had designated lunch and snack breaks, and I always had an appetite right after work.Now when I have been home several months, eating is a massive problem. Getting up earlier than 2pm is a huge but rare success but even if I do I just do not feel like eating for hours, despite feeling so hungry my stomach hurts and my head spins. I just do not crave anything and I try to drown the hunger with water and cigarettes and occasionally save myself from fainting by eating dextrose.This results in me lying around like a zombie for hours, not having the energy to do anything, until I feel too weak to make even a sandwich without having to sit down. Basically, during daytime, I HATE eating. I wish humans did not HAVE to eat and could survive off water.But then in the evening, I often find the cravings wash over me. Not only do I tend to eat unhealthy to begin with because I do not have the energy to cook (for example I often just have microwave food or frozen pizza), but at night while watching shows or playing games I feel some kind of NEED to have something. Sometimes cigarettes and tea works, but often the cravings are too strong. It can be anything I have lying around, like crisps, chocolate, nuts, anything. Sometimes it feels like a proper craving but sometimes it feels like I just have to occupy my hands and mouth. Not giving into this often results in nail biting instead, or even worse pulling at my toenails until they bled (gross i know, but I physically cannot stop and have not been able to show my feet to anyone throughout my life). Giving into the cravings on the other hand makes me feel disgusting, like a grotesque emotional meat machine that keeps eating just for the sake of eating. I feel disgusted by the warm feeling I got after eating those tasty donuts, like how I sometimes feel disgusted after spending money on some useless expensive shit I do not really need but that gave me happiness for one small moment. It makes me feel materialistic but like, with food, if that makes any sense.Now I know the lack of food during the day probably is a big reason for the nocturnal cravings. But it is so hard forcing myself to eat when my appetite is not there, let alone cook proper meals. I am also very conscious about my body and feel like shit knowing that I have put on some weight in recent years, even though I KNOW deep inside that no one cares about that and that weight should not define my attractiveness. But despite this, I have never had an ED. And despite having somewhat ED-like thoughts about food, this feels like mainly a problem with appetite related to my depression and apathetic tendencies. I know that if I could just pull myself together, I could have somewhat healthy eating habits.My past therapist kept nagging me about meal planning but even when I am grocery shopping I cannot for the love of god come up with any ingredients to

buy, despite being kind of good at cooking. Hell, planning ANYTHING nowadays feels like climbing a mountain. Does anyone have similar problems, like lack of appetite? How do you refrain from giving into cravings/comfort eating? How do you manage your daily food intake when depression takes away most energy and enthusiasm to do anything at all? How do you deal with apathy, the kind that makes it hard to even walk from the couch to the bed? How do you keep afloat? Because I am on the verge of drowning. No appetite during day, cravings at night

Statement 48: I am so damn exhausted of my mind screwing everything up. I am about to turn 30 in a few weeks and although I have been fortunate enough to have had many women interested in me throughout the years, I just cannot seem to hold a stable relationship for longer than a few months. When I was younger it was not a huge deal, but now that I am getting older I am beginning to feel I may never be able to develop a long-lasting meaningful romantic connection with anyone and it makes me feel so deeply alone. The thing is, I am able to clearly see how it is me pushing my partners away. Once the initial excitement wears off, I go back to being aloof, insecure and self-centered and I can sense my masculine essence eroding. I can see how my partners go from being attracted to me to feeling sorry and completely turned off. I do not blame them. I do not want them to carry any of my burdens. I want to own up to them and keep fighting to get better. But with each passing relationship, I just become more afraid of opening up about my mental struggles and having them realize that the confident, fun version of me is kind of a part-time deal. On top of it all, I seem to be developing some kind of body dysmorphia. Depression sure seems to have done a number on me physically and its causing me to age a lot more quickly than my peers. I am rapidly losing my hair, about a third of my beard and body hair is already gray, I have bad eye bags and dark circles and, although I have tried, I just do not look good completely bald and I have been told as much by friends. I legitimately hate looking in the mirror most days. I even started dating older women in hopes that it would even things out, but the same thing happened. My last partner recently left me after I became insecure and started acting weird and it really hurt me because I actually did care about her. But somehow, just a week later, the universe has already put two new women in my life who are showing interest. The thing is, I am seriously afraid of trying this again and screwing it up once more. I do not think my heart can take much more disappointment and I am tired of superficial connections. I know I need to develop self-love and find something to get excited about every time I wake up before I try to find someone again, but the thing is I am also very lost and lack direction in life. The only thing I am truly passionate about is music, and that is just not going to pay the bills or provide me with the financial stability that a relationship requires. Besides that, I am working on finishing a degree in biochemistry (two semesters left), but I have no desire to go to grad school or Med school, and the job prospects with a B.S. in that field look bleak. Are there any guys out there who have managed to be in successful or satisfying relationships in spite of all your struggles? If so, how did you manage? I do not think my depression will ever go away at this point, and I do not want to end up alone or hurting myself and others even further. Thank you for reading. Dating as a Man with Depression

Statement 49: Hello everyone I am sorry for this question but I honestly feel curious on why people get depressed, what caused it, and last but not least why cannot some people treat it? I know it is a bit too rude to ask this but if I never ask I would never learn so again sorry for this post. What caused you to be depressed?

Statement 50: After this pandemic is over and i can finally go out, i will overdose myself will sleeping pills. to be honest i do not want to die but I

am just too tired to live. the last time i felt true happiness was in the 6th grade i mean my family is ok and they did not do anything wrong but for the past 6 years I have never been truly happy i decided after the pandemic is over I am going to kill myself

```
In [125]: # Filter rows where 'status' is 'Anxiety' and get the 'statement' values
suicidal_statements = data.loc[data['status'] == 'Suicidal', 'statement'].head(50)

# Print the first 50 statements
for idx, statement in enumerate(suicidal_statements):
    print(f"Statement {idx + 1}: {statement}")
```

Statement 1: I am so exhausted of this. Just when I think I can finally rest, just when I think maybe things are starting to settle, another hurdle comes flying at me. This month alone we found out my mum could be dying, my girlfriend left me, my parents revealed that they wanted a divorce, my grandad was hospitalised again and just now my little sister's been rushed to A&E with possible brain damage. If there is a god up there they must fucking hate me. It is like life is trying to get me to kill myself and honestly I think I would be better off dead. I attempted when I was 12 but I was stupid and there was no way I could cut deep enough. Now I am 15 and everything is so much worse than it ever has been and I just cannot hold on much longer -- it is going to take a miracle to get me through this. I feel so alone. I feel like the world hates me and I have no idea what I did wrong to deserve this. I thought I was getting better. I was doing so well and now everything's just come crashing down again. I was thinking about carbon monoxide poisoning but I highly doubt there is any way for me to get that. It is looking like I will overdose but that is unreliable as fuck and surviving sounds like hell so I do not know. Just got to stick around long enough to make sure my sister's alright. I have already contacted one of my ex's friends to make sure they are looked after when I am gone. I do not think there is long left for me. Sorry. I have posted on this subreddit too many times in the span of a week and that is probably really annoying I just have absolutely no idea what to do anymore. I feel sick. If my sister dies then I really cannot do this anymore. It all hurts. Life is really fucking trying me.

Statement 2: I am 20 year old with some good friends but I am just tired and had a problem with bullying when I was little and that hit me hard but the thing that annoys me the most is like all my friends succeeding in life without even trying like the relationship part without even wanting one and I just feel fucking jealous.. I am just fucking ugly and all I want is to find a person to love.. I feel disgusted by myself that I am jealous of my friends and I just want to end it all but I just see my family's faces and I do not want to put them through this.. I feel I am playing my life on hard mode while everyone is on easy mode.. I just cannot do this anymore I am just tired.

Statement 3: It is looming around the corner again. It always comes back. Like a wave, like the moon. I can hear the silence, feel the breeze, see the cold light. Death. I try and try and try. I choose and I survive. Then it hits again and everything goes dark. Suicide. Singing to me peace, quiet, end of everything. Resisting is difficult, so very difficult. Just a bright white piece of string and it would be over. Peace, quiet, end. I live with death, it is always there. It is looming around the corner.

Statement 4: There is..... food and other things I will be judged for and for having weird views because maybe I am delusional I want to go to jail. Life would be better there I am not joking. Maybe suicide would not even cross my mind in jail.

Statement 5: I am on zoloft and focalin and it is changed my life over the past four years for the better. But this week I have been feeling like I am regressing a lot. I do not have any interest in any activities I normally like. My brain feels like there is a fog around it and I love my new job but I do not have the energy to perform as well. My fiancee went out of town on Tuesday as well which just made things worse. I feel useless and tired and just like shit in general. I think it is hitting extra hard because I have been doing so well for so long. I would talk to my fiancee but she is been through so much with me that I cannot bear to put this on her. I am just lost... What am I doing wrong? I just want to get back in my routine but I do not know shit sucks. It is getting worse again...

Statement 6: I have really bad feelings, but always when I try to look up something related to that on a forum or other website, they quickly go to a go

d. I din't want to hear about that stupid god, but I always find a perspn wh o does that. Because of that, I always quit the website and do sonething els e stupid instead of searching for some comforting text. Why do so many peopl e go to religion when they try to help someone with depression etc.

Statement 7: Everyone in my schooling system thinks I am autistic but I real ly faked it so my school work would be easier, I stopped going to school bec ause I do not fit in there with anyone and nobody likes how different I am, I have problems about my body I have a dent on my chest and it makes me so s elf conscious I cannot do anything that requires me to not wear a shirt or p eople laugh at me and call me an alien, I have a gf but as of recently she i s been loosing connection to me since she has her bestfriend back she does n ot want to be my bestfriend anymore, I have some friends they are really nic e but I do not think they understand me, I think that my Dad hates me becaus e I am not the son that he wanted, I feel like a disappointment to him and m y family name because I am "autistic" and I have not gone to school for 3 ye ars and I feel weak because I was molested and used when I was 10. I am 17 y ears old. I do not have a job and I used to be a drug dealer 1 year ago just because I hated going to school, when I was dealing I was addicted to MDMA a nd Xanax, they have both helped heavily with destroying my views and how I s ee myself, I have lost all of my happy emotions the only feelings that I fee l is sadness and hatred to who I am. I cry almost every day of every hour ju st because living in my body hurts my soul so much. I do not belong on this planet or in this body. I hope nobody feels the ways that I do it is a shit feeling, I just needed to get this off my chest. I am planning on trying DMT as I have experienced LSD over 30 times but I would just like to try it to s ee if it changes anything about myself and if it does not change me I do not want to live here anymore. Sorry for huge wall of text Ik a lot of people on this subreddit probably do not actually care about anyone venting you all ju st care about punctuation and karma but I am doing it anyway I do not feel c omfortable on this Earth anymore and I want to die

Statement 8: I have so many stressors in my life, all major things that tear me down. Every day I fight a battle not just with my mental health but with the world. Last night I finally broke. After 33 years of a shit show of a li fe, I am done. What ultimately triggered this is that my son (7m) with ADHD has been expelled from holiday club. I quit my job Jan 2020 because he was n ot coping in school and managed to find him a specialist educational provide r as well as get him on medication. He did brilliantly! I however, hated bei ng a SAHM and a carer. The loneliness is overwhelming and the fact that I on ly live to provide food and a clean house to 3 children. So, I started to th ink about going back to work in my area of specialty which is law. I got a j ob with a police department. Something for me and something where I will mee t other adults without being attached to my children. I got all 3 of my chil dren into a childcare provider and settled them in over months. I have offic ially been working for less than 2 weeks and my son has regressed back to hi s bad behaviour. He has been kicked out and now I have to quit my job and go back to being his full time carer. I cannot do this for the next 11 years of his life and possibly beyond that. I do not want to live anymore and there i s no help out there for my situation. I can go back on antidepressants, I ca n go to therapy but none of these things fix my problem that I no longer hav e any autonomy over my life. I am not a person any more, I am merely a mum w ith no identity other than that of my children. I find no pleasure in anythi ng and exhaust myself pretending to be happy for the sake of my family. I ca nnot work, and now as it turns out I cannot even die. I cannot leave behind my children, I am their only parent. I have no control in this life over any thing. I now think I am too far gone to ever feel happy again. I want to di e. I cannot go back to Groundhog Day

Statement 9: Imagine waking up only hanging on by a thread and then you get a call from your bank that you lost \$9000 to fraud from a possible job posting. Not to be dramatic but at this point I am convinced that life enjoys torturing me. And I am not sure how much longer I can be it is joke I am life's rag doll

Statement 10: Anyone know of a way that has no chance of failure, is decently quick and is not a gun? It would be much appreciated Painless way except gun?

Statement 11: i feel like there is nothing left for me. i cannot go to school again i cannot handle that again. my entire life is about to change in a couple weeks and i just want it all to stop i want to die so bad

Statement 12: It sounds quite stupid, I know. For the last couple of days, I have prepared everything I need (I will not bring the details, but it is a very elaborate plan) except one: a note. I started brainstorming and drafting what is useless and what is not, decided whether to post or write by hand, and checked for grammar. But I felt like even my last piece of writing, however short, was not good enough. Nothing I do is ever good enough. that is how it has always been. I want to perfect something in my life and end myself in a grand finale. Yet, I do not know. Perhaps I can give it another few days. Today is a new one. I am planning to end it all, but as a writer, I keep editing my suicide note

Statement 13: If I find out I have another illness on top of my epilepsy I will end my life. I am done. I am fucking done. If I find out I have something else

Statement 14: I am f21. i know it is pathetic to ask for help but i just want to talk and at least have a bit of hope that the future will be different. The last days have been though and i do not know how it got to this, now i want to kill myself. i could talk to my friends but i do not want to worry them without actually trying it, they have their lives and my problems feel just too stupid to put that burden on someone else's mind. i do not want to bother anyone. i do not know if this kind of posts is allowed here but all i see are hotlines from other countries and not mine (Venezuela) so I do not know what else to do. i just want to feel better and talk to someone that can offer some kind advice, that would be nice :) i do not really want to kill myself but i feel it is the only way of getting rid of all my problems i just want to talk to somebody

Statement 15: If I fail, there is no more hope left for me, in this shitty life. it is going to be my 3rd attempt this year, so I am seriously messed up and ready to go. Previously I gave up because it was too cold and I want to die in the wilderness. This time I am better prepared. I do not have much more strength to mask my emotions. 1 week then I gtfo I give myself 1 week to find a new job

Statement 16: My life has always been a mess but now I am at a point where I do not have much energy left to go on any further. Every second is tormenting me. I have made several mistakes that threw what is been left of me and my life completely off the rails and now I am devastated. I do not have any perspectives whatsoever and self hate eats me. What can I do? Nothing is fun anymore and I have no one. It gets worse everyday. People are trying to help me and they really do see what I am saying but they do not know how to help me either and they are professionals. I am in mental hospitals most of the time but it does not help. I do not know what to do anymore I feel horrible

Statement 17: I have given up on life. I hate my emotionally abusive parents. I wish they just fucking divorced. they are always arguing over literally everything, they are always fucking insulting me every fucking day. they are also overcontrolling af, treating me like a fucking object and making me literally feel like I am physically and mentally suffocating every day. I feel

empty, useless and hopeless every fucking day. I wish i was aborted. I am in uni, but i have no fucking clue what I want to do with my life, and I do not have anywhere near enough money to move out. Fucking sucks having to pretend to friends that I am fine, when in reality I am constantly bawling my eyes out everyday because I think about killing myself everyday. I wake up everyday wishing I died in my sleep. I have tried so fucking hard to fight my depression. But it literally has destroyed me. I have got no fight left in me. Life just is not meant to be for me. I have had enough

Statement 18: Yes. It is necessary. However it is not a cure and does not fix everything. It does not fix trauma or trust issues when they are deeply ingrained because nearly everyone in your life abused and betrayed you. It does not fix depression. Coping mechanisms only go so far. You go to therapy, you talk about it, they give you coping mechanisms that trauma, depression and anxiety supersedes. And you have to trick yourself into having a positive mindset to the point where it is toxic or invalidating or bypassing very real issues sometimes. I hate it. If you do not have love, support or an environment that is not triggering then sometimes therapy is not enough. A recent K-pop star attempted suicide and all I see in the comments are: Get therapy. She needs professional help. The presence of a therapist is only temporary. Even if they are a good one. What about being home alone with your thoughts, trust issues and depression? What about feeling like you need to be held all the time? there is no medium for that. For Christ's sakes. I am so sick of people telling people to get therapy.

Statement 19: Than having to live my life. The negative outweigh the positives. The highs are not worth it. it is been like this for years. I have not become stronger, I am still the same. And still, nobody gives a fuck, because everyone is got problems. Why? Death seems a better option

Statement 20: I have horrible ocd (pureo) and it constantly bothers me. I cannot be happy for more than 10 mins a day. I am miserable. I cannot even kill myself because people would miss me. I wish I could kill myself

Statement 21: had a lot of little half assed attempts i did not really start or go far enough through, but i did last night. the other time was when i OD'd at 16was suffocating for an hour but i could not pass out just crying a lot at one point and writing out my thoughts. my neck and shoulders hurt a lot and i keep feeling tingly and dizzy at points. i really wanted to go to the ER but it is scary, you know? i have therapy later but i do not know if I will tell him. a lot of awful stuff has been going on. do not know if i can talk about any of it though. it is not like he can help me. he would probably have me shipped off somewhere where I will lose all my rights and probably be mistreated at best, but likely molested and abused at worst. and then i will not be able to get top surgery anymore because I am too cわazywish i could get like. some kind of medication to get me through these shit things but i just get treated like a druggie for wanting something more intensive than shit SSRIs texts someone sent me today are just sending me even more over the edge. nobody understands attempt #2

Statement 22: I have suicidal thoughts, anger and I just do not care anymore. On edge from transitioning and transitioning genders, getting angry at people attacking them verbally even physically, if it got to that stage and just want to end it all. As a coping mechanism, I have posted to NSFW subreddits on an another account and cannot stop. Is this what my life has come to? Sometimes in my life wish I were invisible.

Statement 23: I am over eating, my sleep schedule is erratic, my room is a mess, I am binging, I am moving slow, I am hurting myself, my head is blank, I have deep dark eye circles, I am over spending. I should have ended my life at April but my father is so selfish for getting the knife in my roomI need help. Help me stop please. I am desperate. I do not know if talking can he

lp, but I do not think it can. It never does. I am alive but not living Statement 24: // general narcissist venting, feel free to comment if you want The thought of dying without being able to see if anyone cares or not just fucks with me too much to actually commit suicide. I mean of course they would care. People even care when strangers die. The initial shock of someone you knew dying and knowing that you could not do anything about it is very traumatizing. But I think what would affect them is the idea of a family member dying, and not that they would miss things about me or reminisce over memories or anything like that. The only memories they have of me is me being a dumbass or saying "what?" five times in a row because my hearing is going. But the thought of this makes me want to go out and be something that I know I cannot be. I want to make something of myself but I just, I cannot. I have nothing to put on my resume, I cannot do taxes, there is so many obvious things that other people know that I just do not! And nobody thinks to tell me because I should already know! Why do not I already know basic crap! Why did I spend my childhood locked in my room instead of trying to learn and develop like a normal child! Why cannot I take care of myself without burning out? If I brush my teeth, shower, get dressed, and do my hair routine all in the same day then I am burnt out! that is it! I cannot do anything else for the day, I am too tired! there is no way I could work or get a job like this! I would burn out after one task and get fired! I hate that my huge sense of self importance has manifested into me always focusing on how I am the worst instead of atleast making me delusional enough to think that I am the best. I hate that my self esteem is based on other people's approval, but unlike other narcs I do not get approval because I am literally shit at everything so I am just an unmotivated sack of shit who cannot do anything. I hate that I even told anyone that I am a narc. Of course they would not accept me!! What the hell was I thinking!! What kind of idiot would expect anything different!! Things are not going to get better for the NPD community, not in my life time atleast! Well, maybe that will be another reason to stay alive. Imagine being an old man and seeing 20 something year olds openly talk about their NPD recovery process that they felt no shame entering because there is no more stigma, experiencing something I could not. Maybe. When I am in my 80s. And my mom, my poor mom. She has to pay for everything because I cannot do shit on my own. Physical therapy, regular therapy, medication, my food, I hate this!! You should not have to take care of me!! You should be relaxing somewhere without worrying about my next diagnosis or my next medication or whether or not any treatment is working!! You should be doing your hobbies and I should be the one buying you things!! Maybe I will get to the point of recovery where I am not so self-centered, and then I can finally let go. And they can all finally live their lives. I wish I was not such a narc so I could just find peace in death

Statement 25: Started volunteering at a local animal shelter. I feel better, not happier. I still really hate myself, but I think I can push forward. I want to start painting and drawing to express myself, I know I'll start eventually. Days just blend in

Statement 26: I have been so disconnected with the world, my family, and myself for the past few years. I do not know what my face looks like. I see it in a mirror but cannot picture that is actually me. My dad talked about me as a child and I cannot picture that happened. I was never a child. I never existed. Nothing does. I Do Not Know how to talk to my family friends and even little sister who I was so close with. I Do Not Know what to do Disconnected

Statement 27: I feel like nothing everyday, I wake up and ask why am I alive and why did not I die in my sleep? I mean almost nothing to most members but of society so if I offed myself people would forget about me after using my n

ame for publicity after 3 days. When you die people only care then use your name for clout. Tired of living a life I did not ask for, I have no goals, no motivation. I live in a bum ass town encompassed in a shitty county with no opportunities. People tell me no excuses but, I do not care I am not as strong as everybody. And being a man they expect you to suck it up and hold it in. But in reality I just rather off myself or die in my sleep. On the edge and probably about to write my final note.

Statement 28: I have called the Samaritans, I used to contact childline and I cannot go counselling or to my family or to my school. they will contact safeguarding n Ill be placed in a mental hospital. I want to commit but Ill feel guilty but then I do not want to continue. When I die I will not feel guilt, I will not have memories or any worries. So peaceful. I am going to do it

Statement 29: I am so sick and tired of not being able to do anything the right way, first my ex broke up with me, then alot of my friends started ignoring me, and now I just do not have the energy to do much and the stuff I want to do, I cannot do correctly, and its a stupidly easy thing id rather not mention, I am really really tired of trying anymore. I do not have a clue what I am doing anymore.

Statement 30: Content warning: Please do not read this post if you are suicidal/prone to suicidal thoughts. I do not want to trigger a negative spiral for anyone else. I am 19M from Turkey. My family history from the maternal side is filled with mental health issues, I have a depressed mother and a suicidal sister who is older than me. I have a very sensitive temperament, am very neurotic and introverted. I am somewhat bright academically (not a genius, but significantly above average), have no hobbies or social life. Ever since the age of 13, I have been severely depressed and isolated. I know many people have it much harder than me, and I am somewhat in a privileged demographic but I really need to vent without holding back. Living in Turkey is very hard nowadays with the islamist and nationalistic fascist regime, and even if I want to migrate, I will be met with some degree of prejudice because I am a middle easterner. I do not want to leave my country ideally too, but there is no choice other than to strive for an escape. I am not religious, believe life is meaningless. Those are not the reasons behind my depression, though. I do not give a shit about life being meaningless, because I know the brain only seeks happiness and meaning is irrelevant in this case. The problem is, I cannot fulfill my needs and it is destroying me. My lack of social life especially hurts in my relationships with girls. I have gone through many rejections, only to deduce that (a) my need for approval is making me unattractive (b) I am not good/attractive/sociable enough for women to take notice of me, want me. Of course nobody tells those to my face, they are rather the constant messages I come across reading dating advice blogs, psychological studies on attraction etc. It is even more depressing when dating, intimacy and relationships are portrayed through a lens of competition. And even more depressing that I have no satisfying counter arguments against that. I, too, can see that this is all just a competition; but some of us (me included) are not comfortable admitting. I remember listening to the talks of evolutionary psychologists, how they talk about "everyone is trying to get the best deal they can" and "you may be miserable as an 8, since you lust after 10s but 6s lust after you." Dating is a market; and I just cannot see how you can be happy or fulfilled unless you are in the 1% as a male. I know suicide, if done right, would be the same as falling asleep, just not to wake up. I know painless ways, and losing hope in ever attaining a fulfilling intimacy just makes me lose interest in living. I have a very painful daily lived experience, and I just wish for the pain to go away but it persists. There is always an unfulfilled need for love, and it is very hard to fulfill that need. And I do not have

the skillset to have it fulfilled (and mind you; I have been trying very hard to be social. My temperament is just too apparent) I am going to wait for my therapy sessions. I just want this painful feeling of isolation and rejection to go away. And it will not. I may decide on killing myself a year later if I feel the same way, and things did not change I am considering suicide Statement 31: I have been through hell and back just to be myself, lost my (extremely abusive) birth family in the process (who never accepted that they had a daughter instead of a son anyway), accidentally found out about a ton of memories from my childhood, and I just do not want to be in all this pain. I cannot work because my stupid brain does not function properly, especially around other people, and I am running out of food. I started therapy, but I do not know how long it will take to get better and I am just hurting so much. I should not have to deal with all this pain just because I was born in the wrong body and to the wrong family, it is just not fair. It does not feel like things will get better for me

Statement 32: I have not been suicidal for awhile, a good 3-4 months! I am still living with my parents, and my mom keeps all pills (& other things I could use to hurt myself) in her room. She fills up a pill container for a few days, no more than 4 at a time. And will not let me take more Atarax which I need throughout the day, so I have to ask her continuously. She also does not think I should take much, even though the safe amount is way more than I take everyday. Basically, I do not get why I am trying if nothing is changing. I attempted my worst attempt about a year ago, I am about 5 months clean from self harm and am not suicidal! When does the trust come back? I understand she is scared, that is valid. But I will not grow at all if I am kept under lock & key forever. Why am I trying

Statement 33: A week ago I was fine. Now all my friends hate my guts, and I cannot go outside because I have fuckhbgb Covid. If the virus does not kill me then I sure as hell will do its job for it. My life has gone downhill so fast

Statement 34: Why am I so stupid? I fell for someone who cannot love me back. Loving him hurts so much but I do not know how to stop. I know he will not be in my life forever and when he starts college I will not have much chance to see him anymore. I really do not want us to just slowly drift apart and stop talking. I do not want to be alone. I already have nothing to live for and if he's gone as well there will not be anything worth staying alive for. I hate myself for loving someone

Statement 35: Throwaway account because God forbid anyone who actually knows me sees this. I am just tired and it does not even feel worth it anymore. It just feels so futile and hopeless to stay alive. I was diagnosed with ADHD, depression, and binge eating disorder in February, I am on medication and yet I am still fucking up a lot. I am still late for work on occasion, I am still having a hard time getting things done around the apartment, I keep telling myself that I will do good things like cook healthy food and wake up and take walks and wash my face "from this point forward" but it never happens or I do it for a day or 2 and cannot keep it up. Before, I thought what my family said was true and I am just lazy but I cannot function without medication. And what is the point of living a life where I need medication and I am still having trouble getting it together anyway? I would rather be dead than wasting air and people's time. I Wish I Was Dead

Statement 36: I have been going through a lot lately and all I can think of is taking my own life. All I feel is pain and depression and I am honestly not sure if I really want help or just for everything to end I need help

Statement 37: I do not know what my purpose is on this planet. No idea what I want to do in terms of a job. I feel no passions or have any kind of dream I wish to chase. It makes my life and existents almost feel pointlessWhat am

I actually here for. what is the point in living 80 or more years unhappy. I just want to not exist but that is not possible without serious repercussions. Just wanted to rant and see if I was the only one. Struggling to understand a purpose

Statement 38: I have posted about it several times, but here goes. 12 months ago I joined a dating website and met a Thai woman. We got on EXTREMELY well. I thought she was the one for me. It later transpired that she had 2 children which she did not tell me about. Odd, but ok. I decided to go with the flow and see whether we could make it work. Shortly after, we got into a relationship. One night I video-called her and she did not reply. She said she was feeling unwell and she would see me tomorrow. When I turned up the next day, she was standoff-ish. I saw her phone on the side so I grabbed it and took a look at her photo album, looking for the photos we would take a week earlier. I was horrified to see a picture of her in bed with another man. She would cheat on me the night I called her. Even worse, after I confronted her, she told me it was for money. Yes, he would pay my girlfriend to sleep with him. Turns out she was on some kind of dating app and had arranged that behind my back. There were a few other guys she would arrange to see but she did not in the end. She assured me it would not happen again and she had not done it before. Turns out it was all a big fat lie. I found out she was a hooker. By this point we would already given it another go. I would get attached to her and her two children. I would cook for them, bought them things - I felt like a part of their family. Anyway, I thought it was all going well and that was behind us, when she cheated. Again. I saw a guy online reviewing her. Yes, REVIEWING her service(s) and he described what happened and how much she charged. I was devastated. I cried buckets of tears. Fast forward a few months and we had given it another go and we were getting along well - when all of a sudden she dumped me. She said she was going back to that lifestyle and it left me a broken man. I would love her, forgive her several times, I would been looking after her children. I did absolutely everything I could for them. I am just left with pain. I will not commit suicide, but the pain has been so bad I have thought things like: "when I am gone I will not be in this pain". She is blocking me on everything. I called her phone from another number and she hung up when she heard my voice. She is treated me like a piece of trash. At an all time low

Statement 39: I have been in therapy and I feel so much better. It is been a year since the last time I thought about ending my life, I have done so much work and I still have to, but I feel really proud of myself. It gets better, I promise!! I no longer want to kill myself

Statement 40: I do not know why I even had hope that it would get better. Things only get worse and I hate when people say it will get better. My mind is telling me to give up everyday I am trying to fight but I do not think I was made for this world. I love my parents and they are very caring but I do not think that is enough reason for me to keep going. I have nothing going for me and I hate myself. I remember when I was a kid and was actually happy with myself but those days are long gone. Now all I can think about is how worthless I am and how I do not belong. I think this might be my last few days here. I hope everyone has a good day and keeps fighting. It does not get better

Statement 41: Does anyone have experience with brain damage after a suicide attempt? Brain damage

Statement 42: I am not sure if I am incredibly mentally unwell but scrolling through gore posts and seeing failed suicides make me incredibly scared that if I fail, I would be horrifically deformed and the pain would get worse. Gore subs and website scare me into not committing suicide.

Statement 43: I cannot do it. I cannot do it. I cannot do it. I want to die

and its all they care about

Statement 44: I tried it a few times but it never worked except for the last time but I got scared and called the ambulance. But now I am sad again and I think about killing myself again. My medication does not help me enough and everyone seems to hate me that is why I cannot find a reason to stay alive. I do not want to live anymore

Statement 45: All it takes is any small thing to trigger me into suicidal mode. Can I please talk to someone, I need a friend. I hate how I go from 100 feeling amazing, to wanting to blow my brains out in seconds

Statement 46: Have to tie up a few things. Final blow came today. I am done. I survived an attempt over a year ago and since then things have got considerably worse. I cannot manage myself in this world. I cannot do the things other people do. I do not understand the tricks people do to get along. it is not mental illness. it is this place. I hate it. I always did. I love you all. Over and out. Checking out

Statement 47: Cannot take this domestic bs anymore and do not think it will ever end. did not ask for this crap, and did not ask to be born either. I hope my piece of crap father burns in hell for putting me in this situation which has not been helped by some of my decisions either. cannot wait to end it. Will end it soon now

Statement 48: Today I was on guard duty in my base. I went back to my room and received a call from my commander, saying that I was spotted going to my room with my cartridge inside my m16. I told him its not true and that I removed the magazine myself. Apparently, you are not allowed to do that and now I am going to be trialled. I am so scared. I do not want to be trialed. I want to go home, I want to finish it right now. I do not know if I should pull the trigger.... I have a gun in my room, I think about it a lot, but I do not know if I should.

Statement 49: I am sad and I am scaredGuys i do not know where else to turn to with my feelings right now. I am just so frustrated. I have been in my antidepressants for 6 weeks now and I am not doing okay this past week. I Do Not Know if its the circumstances of my life right now or if my medication is not working.. what started this all was me kicking my sister out of my house for disrespect and overstepping boundaries, then basically my mom was supposed to go to rehab so I dedicated 2 days to try and help her get ready.. while I was at her house getting her ready some drama happened and My sister (the one I kicked out) was trying to fight my mom, so I just stepped between them and then my dad comes between me and my sister and yells at me like I was the one instigating. He backed me into a corner almost yelling and cussing at me. He never touched me, but he kept walking towards me so of course I stepped back. And it just brought back bad memories of my childhood but not a specific one if that makes sense anyway I ended up bringing her to the airport the second day. Well something happened at the airport that was neither of our fault and she said she did not want to go anymore so that upset me and I cried the entire way to her house which was an hour. Then my cat has to have his leg amputated and its going to cost us \$1000 or more after everything is said and done with his er visit, preop exam, and surgery My sleep schedule has been messed up and it gets worse everyday. I sleep till 4 in the afternoon and I am ashamed of myself. I feel disgusted by myself honestly. I start school august 18 and part of me feels like I cannot make it because of the way I am dealing with stress right now. I am overwhelmed and now I am discouraged and that makes me feel even worse. I hope someone can understand what I mean. Its almost like I am sad BECAUSE I am sad in a way. I am in therapy but I did not get to make it to my appointment yesterday because of my sleep schedule and also staying home with my cat. I have no friends and I feel like I annoy everyone so that puts me into antisocial mode. I just want it all to

stop. I was doing so good the first 4 weeks of my meds and slowly it got worse and I feel I am right where I was before and I am scared that I am gonna have to be depressed forever because medication is not working . I am sorry if this sounds dramatic but its literally how I feel. I just feel hopeless, my antidepressants have not seemed to be working this past week AT ALL.. I just wish I had a friend who has been through this to kind of help me because I am lost.

Statement 50: Hey um, I do not like talking to strangers but I feel like I need to do something before it goes too far down the rabbit hole. I am 30 with a 1 yr old and another coming in December been married to 2 yrs (do not know for how much longer though). I write songs to cope(I will post them if you want to read them) but to be honest I will not lie I want to throw my hands up so bad..... I am angry with a lot different things. I miss the me before I got corrupted with life and it is experiences. I cannot be an effective father and husband like this. Truthfully ik my family would be better off without me. No more worrying about me. They could focus on themselves. Idek why I am posting this. Maybe crying through reddit. Looking for God through. Hell....i do not know.... I have been diagnosed with Depression and Anxiety disorder by the military and civilian psychiatrists.... Anything to avoid it

It does seem like there is a much clearer distinction between 'Suicidal' and 'Depression' than there was between Stress and Anxiety.

Let's quickly examine whether or not any of our 'Depression' rows might be mislabeled. We don't know how these datasets were initially integrated and classified.

```
In [126...]: # Example tokens to check for
tokens = ['suicide', 'suicidal', 'kill']

# Filter rows where 'status' is 'Depression' and 'statement' contains any of the tokens
filtered_rows = data[(data['status'] == 'Depression') &
                     (data['statement'].str.contains('|'.join(tokens))]
```

```
In [127...]: filtered_rows
```

Out[127...]

	statement	status	perc_upper_words	neg_emoji	exclamation_count	que
6983	I do not know how to navigate these feelings, ...	Depression	9.539474	0	0	
6996	I am going to make this short because i hate t...	Depression	4.249292	0	0	
6999	I have been talking to someone who I really li...	Depression	11.627907	0	0	
7004	Not because of death, but because she wanted t...	Depression	6.852792	0	0	
7015	I have not seen my 7 year old daughter in a co...	Depression	2.380952	0	0	
...						
48038	It is getting worse 19F, I have been struggling...	Depression	2.958580	0	0	
48045	Tried committing suicide twice today failed mis...	Depression	3.658537	0	0	
48051	I feel like I was not meant to be born (litera...	Depression	1.575931	0	0	

	statement	status	perc_upper_words	neg_emoji	exclamation_count	que
48058	I hate this phrase People always jump to say, ...	Depression	9.433962	0	0	
48059	Everytime I think of my "future" I just think,...	Depression	7.500000	0	0	

3113 rows × 9 columns

Holy moly! That's a lot! Let's remove 'kill' (our intention was to isolate rows with anything similar to 'kill myself' in the statement, however with the amount of rows that were returned we'll have to get more specific) and try again...

```
In [128]: # Example tokens to check for
tokens = ['suicide', 'suicidal']

# Filter rows where 'status' is 'Depression' and 'statement' contains any to
filtered_rows = data[(data['status'] == 'Depression') &
                     (data['statement'].str.contains('|'.join(tokens))]

filtered_rows
```

Out [128...]

	statement	status	perc_upper_words	neg_emoji	exclamation_count	que...
6983	I do not know how to navigate these feelings, ...	Depression	9.539474	0	0	
6996	I am going to make this short because i hate t...	Depression	4.249292	0	0	
6999	I have been talking to someone who I really li...	Depression	11.627907	0	0	
7040	I am 19 male. Both of my parents where immigr...	Depression	7.500000	0	0	
7044	I no longer look forward to ANYTHING. I have a...	Depression	11.347518	0	0	
48030
48030	I have been struggling with depression for aro...	Depression	9.743590	0	0	
48037	another episode is suddenly hitting me and i d...	Depression	8.673469	1	0	
48045	Tried committing suicide twice today failed mis...	Depression	3.658537	0	0	
48051	I feel like I was not meant to be born (litera...	Depression	1.575931	0	0	

	statement	status	perc_upper_words	neg_emoji	exclamation_count	que
48058	I hate this phrase People always jump to say, ...	Depression	9.433962	0	0	

1790 rows × 9 columns

In [129...]

```
# Print the first 50 statements
for idx, statement in enumerate(filtered_rows['statement'].head(50), start=1)
    print(f"Statement {idx}: {statement}")
```

Statement 1: I do not know how to navigate these feelings, not that its a new feeling by any stretch. I just do not understand how I can go on from being so pleasantly fine to suicidal and self loathing within in minutes and not be able to step out of it. I have been suffering from poor mental health for over 10 years and I am about to turn 26 next week and I feel so ashamed for how serious my negative thoughts are. Covid was a hard transition especially having moved to a new city a few months before everything started locking down, but now I am here and settled and I have a loving partner, a great connection to their family, a really fun job and I finished my degree. But when I stop and look around I feel like nothing will ever be good enough for me and I will never be good enough, I will always have skin problems and bad teeth and acne. I will never be the artist I used to be (I have a Bachelor in fine art) because I do not have the same drive or passionate that I used to have for it. I will never have the career I built my entire plan around because I am losing my interest in art. I will never have the relationship with my family that I want because I could not stand to stay in my hometown. All I can feel is this tornado of my failures swirling around me all the damn time. It just seems that no matter how many positives come my way they do not last long enough to out weight the negatives. I do not really know what I am looking for I guess I just wanted to say this out loud, or whatever. 0 to 100

Statement 2: I am going to make this short because i hate thinking about it, but long story short, I am getting better after months of medication, therapy, etc. My anxiety is at an all-time low and my antidepressants ""do not allow"" my mind to wander too much into suicidal thoughts. But i hate it, i hate that I am getting better, i feel embarrassed whenever i catch myself doing something not depressing, like taking a walk, laughing, or doing stuff by my own will. And i feel SO SO selfish, i was given immense support from everyone, my mum has spent so so much time and money on therapists, psychiatrists, pills, etc; she is suffered more than anyone else, even more than me since i seem to enjoy being depressed. she is the happiest she is been in months since she can tell I am in a better mood and she is the only reason i have not "tried harder" to get worse, but i do not know how long i can keep it going. Something in me just wants me to maybe stop taking my meds without letting her know, get a new self destructive habit, and go back to the old me that did not leave his bed in days and could not keep a conversation for more than 2 minutes, i really miss it.I am going back to school in a couple of months (i dropped out), i started driving lessons yesterday, I am eating better, sleeping better, I am starting to live and behave like a regular person but i cannot handle it, it makes me feel bad, gross, EMBARRASSED, very embarrassed, and its just a cycle of:I am ok this makes me depressed I am ok again this makes me depressed etc.I do not know what to do, I am not even sure why i am even writing this since I will probably end up ignoring whatever help or advice i get, but i guess i needed to vent, thanks for reading and as always sorry for my english how can i be happy about getting better??

Statement 3: I have been talking to someone who I really liked. They were a great person and we really clicked. I told her that I could not do this with her because I am dealing with so much shit in my head that it is not fair to her. She said she wanted to be there for me and I shut her down. I am scared to be depressed around other people now. My poor mental health ruined one relationship and I do not want it to ruin another. I know I am much better now than I was in my previous relationship but I am scared that I might be a horrible person again.I am scared to let anyone be close. I feel like I am just getting by daily and cannot make any real connections anymore. The ability to get close to people is non-existent because I am scared that when they see the real me, they will leave. I am trying to improve but what if I cannot, what if I fall apart again? What if I am the piece of shit suicidal asshole t

hat I was before? I am scared to put anyone through that. I kept making up excuses for why this was not going to work out. Distance, what we wanted, culture, but truthfully I am just fucking scared to be close to someone again. I am scared to be vulnerable. I wish I knew how to not be depressed. It controls my life. I hate it I am pushing people away

Statement 4: I am 19 male. Both of my parents were immigrants my dad died in a car accident and my mom crazy and killed herself shortly after when I was 8. My only family left was my grandma. She raised me and I fucking love her so much, but she is old and has Alzheimers. Its so fucking hard seeing her drift away. I should be grateful that I was given such a loving and awesome grandma for the time she was there. I just do not know what Ill do when she forgets me. My depression has really gotten bad I do not want to take paxil because I am scared of the suicidal thoughts. Any tips? What do I even do?

Statement 5: I no longer look forward to ANYTHING. I have a beautiful girlfriend, a caring family, job opportunities, talent, looks, personality, but inside I am miserable and have lost my love for life. I spent about 3 years of my life going through all the popular antidepressants, benzodiazepines, etc. but none of them worked. I no longer drink alcohol because I abuse it and I no longer smoke weed because when I do, I smoke non-stop and get even more depressed. I am at the end of my rope and I no longer know what to do. I am not going to commit suicide but I just wish I could find some hope. Any suggestions would be appreciated I am 27 (m) & Everything in my life is picture perfect from the outside, but my depression is eating me up inside day by day.

Statement 6: I grew up being a very bad kid, i smoked weed from age 9 first time i got drunk i was 13, since child hood we moved from place to place, never had my friends as we always moved, since i was 18 i moved away from my parents to another country and 3 years later, still cannot make friends because people are so fucking antisocial here, i lost my first son after 3 days of birth due to placenta seperation, and now my job is pissing me off, i want to break up with my gf but i love her, money is a big issue and recently I have been having suicidal thoughts and other thoughts, is there anyone that has any tips? Tired of living, weird thoughts

Statement 7: I am okay with dying or living but I will not fight for either. Like if I get a terminal illness, Ill probably just let it be. I cannot imagine myself reaching 30, I expect that somewhere between my college and grad school years, Ill just die of some illness or probably suicide but I do not really want to commit suicide because of my family will blame each other and stuff like that and also my mom, I do not want to put her in a situation like that. I am okay with getting killed through natural causes where no one gets blamed. Right now, I think I am okay with this; nothing really matters anymore or something. Basically I do not have a will to live or a reason to fight for my life should I be in a fatal situation. I know some people feels like this too but i do not know what this is

Statement 8: I am tired of putting up a fake smile, fake laugh and shit. I think I am depressed, like I have a lot of self-hatred, suicidal thoughts almost everyday. it is so fucking hard to concentrate on studies. My phone is my only escape. I do not know what the fuck to do now. I do not know if I am depressed or am I just acting like an idiot. I do not know what the fuck I am going to do now.

Statement 9: I have had more downs in my life than ups to be honest but the main core of this situation was and still is because of my parents' we were taught to respect them and love them because they were the ones who raised us and "put a shelter over ours heads". For me these last two years were absolute hell and it was not because of the pandemic but there are several reasons like:1. I am pretty smart and got into med university by the age of 17 an

d because of the amount of stress and trauma I had to deal with I failed2. After failing my parents gave me a second chance to reenter the university and I took it but I was so unhappy that I chose not to do it and went against my parents simply because I was severely depressed and hospitalized due to psychological pains like constant chest pains and I could not breathe3. My parents were not supportive of me. They were supportive when they were showing off to their peers or friends and my dad told me " You are nothing but a disappointment and you embarrass us wherever we go" I was highly discouraged 4. I still get yelled at and pretty much am verbally insulted and was told I cannot do shit in life because I am a complete failure. I really hated when my dad used to tell me "You are going to be a doctor so you should know this". The only thing I know, is that I really really hate you and this hate adds up even more daily.5. I have this new opportunity where I technically can go to another country and go to university there and my parents will not be able to come. I feel so restrained and so closed in like in a box and I suffocate so I feel like I should say fuck you to my parents and leave but I do not know how well I can deal with myself there too6. Technically they killed all my dreams, what I could have became and they took my dreams and crushed it and are the MOST negative people I have ever seen in my life and I do not even want to look at them. I know it is unethical to hate your parents but do not blame the ones they do because some parents clearly suck and I even told them " You should have never became parents" or "I wish I was never born I did not ask to you to be my parents" My parents makes me want to kill myself and have suicidal thoughts and I cannot take it anymore. They do not know how to communicate at all and they have zero empathy and funnily enough my parents hate each other and themselves as well; weird how they have to reflect that on me too. I was an innocent kid with big dreams. I was not born to be insulted like this, they are the fakest people I have ever met too. Acting like they care about me in front of others all the fake hugs and such and ask me why I do not ever let them to hug me; it is because I hate you and I do not want to see you. I do not know what to do anymore...should I leave to an other country and do not see them again? It seems pretty huge but I cannot deal with this kind of toxicity anymore.. It feels like it keeps getting worse

Statement 10: I feel awful right now, and almost every night. I cannot go to bed anymore without realizing how much no one loves me, or how much I hate college! Its only worse that I have been thinking a little more about the idea of suicide. I swear, I am considering it after graduation. One day, I am feeling great, the next day just okay. Right now, I feel like absolute shit

Statement 11: I am so fucking suicidal, but cannot bring myself to do it, i pull the knife to my throat but cannot convince myself, everybody in my family loves me, but myself. Why do i hate myself, there is no reason, am i just being a spoiled, ugly, brat? I feel like every part of me is too tired for anything, but i try, if i could i would sleep my life away. I try to convince myself to tell my family and reach out, but younger me has little devil ears and clicked the button of Life Ruiner. Is their anyway to kill myself, but painless and quick? Kill me

Statement 12: My life is very hard. I am in pain all the time. I have been working on healing for years and there is been some improvement but overall I am still in bad shape. I can usually only sleep during the day. I play chess but my pain really affects my game and its really disappointing. I play guitar but have really low quality gear because I have no money. I live with my parents. I cannot work at all and my disability claim was denied. I sent it in an appeal, hoping for the best. Its always hard being around my father. It seems like hes so disappointed in me and it creates a very tense atmosphere. Its easier to be around my mom but its still hard. My older brother blew

up at me a few weeks ago telling me how lazy I am. He does not believe I have a disability. He was incredibly insulting. I blocked him but he apologized so I am talking to him again. Hopefully it will stay civil. My younger brother is nice but does not really understand my situation. I spend all my time either at home or at the park a few feet from my house. I tried to get gas the other day but was too anxious. I drove by the gas station and went back home. My mental health has never been decent. In 2013 my physical health began doing downhill. One issue after another. No help from any of the many doctors I tried. I smoked weed every day from 12 to 18 to try and cope with my issues. When I got sober a lot of the people I thought were my friends wanted little to nothing to do with me. College was terrible. I dropped out my senior year. Every year got harder and harder since my health was declining. Then I moved back in with my parents and it was doctor after doctor, protocol after protocol. None of it helped. It generally just made things worse and I voice this to my parents and it caused more tension. I was kicked out of the house countless times and often slept at motel 6 with the money I had at the time. I have had such a bizarrely difficult life. I have been expressing suicidal ideation since 3rd grade. I keep hoping things will pick up for me and that I'll finally reach a place where I feel stable. I have hopes and dreams. Well see what happens. Thanks for reading. All the best to you Tanner A difficult life

Statement 13: i do not know if its depression last year i used to love studying and did it for around 8 hours on top of school because i just enjoyed it . maybe its burnout but now i cannot even study for 10 minutes without feeling like its not worth it. this is the course i wanted to study my entire life. i wake up in the morning full of anxiety and have to calm myself down because the thought of university is too much. maybe its the fact that i was anticipating life to be good at this stage and for me to be happy that now that I am not, it is heartbreaking. i will never drop out of school but i have considered suicide because staying does not seem worth it. I have a bf and a loving family but seeing them no longer brings me joy either. but i want to know if i should visit a doctor and get on medication or should i just wait it out for a few more months. not worth it

Statement 14: i have been depressed most of my life, i am now 27 , and every time i seem to get somewhere with being happy my mind just forces itself to become depressed again, back to the comfort zone.. i was reading that depression swells the brain over the years or something.. is it possible to be depressed beyond return? i have no wants in life at all anymore i have long just been waiting to die / contemplating suicide but the trail of destruction i would leave behind for my partner, parents, siblings ect is something i cannot do, i do not want to disperse my depression among those who have tried their best to help me. depressed beyond the point of return

Statement 15: With what medications can I almost commit suicide? I just want to almost die and go to the hospital, I am running out of options to get my ex's attention, she did not care that i cut myself, I just want to talk to her. What meds or pills can i take?

Statement 16: Hi my name is Pierre I am a 22 year old guy from Houston and for about 7 years I have been really depressed. A lot of people close to me really do not know because I have always been the funny guy that makes people laugh. I have had a tuff coming up me and my mother had a very bad relationship coming up I am not really close to my family but I love them though. I try to keep a smile on everyone's face if I can but at the end of the day I am somewhere questioning why am I alive I have thought about suicide multiple times but I have a niece and nephew that I could not do that to my relationship with my gf is so weird I really like her but i know deep down she does not care about me its shows but I stay with her I am currently trying to g

et a car so I can move around without having to ask someone who dislikes me I am actually crying right now my heart is so heavy I can never keep anyone happy I understand men are not supposed to cry but man I just want to be happy and make others happy but it seems I cannot I pray but not as much as I should know I will not get a reply I am not to much worried but I needed to let something is off my chest it hurts sleeping in the bed with someone who is obviously cheating on you and is not interested but before I stayed with her I was homeless and just graduated she knows I do not have anywhere else to live so she treats me like shit because she know I do not have anywhere to go I am so tired of being tired no matter what I do lately the tears have been uncontrollable I need help I am tired I just want to be happy man with peace of mind please God I am so ready to end it all I cannot continue like this Help me

Statement 17: I have had suicidal thought i do not know if that is a sign of sadness or depression ? Can someone help Am I depressed or just sad?

Statement 18: Posting here mostly as a last resort since I feel like I bother others when I talk about my depression. I am a 23 years old girl and feel like a loser. I have suffered from depression for as long as I can remember, which was made worse when I was a teenager due to drug use/excessive drinking, being sexually assaulted, and briefly working as a sex worker although in a pretty sketchy way. I feel like no one cares. I feel like my friends do not really give a shit about me or would not care if I was not around. I feel like my band sucks even though I put so much of myself into it. I feel like people think I am stupid. I fear I may actually just be stupid. I feel unremarkable, worthless. I have not attempted suicide in many years but sometimes I feel I am getting there again. I feel unhappy in my relationships and I think the problem is and always will be me. I try DBT and am on antidepressants. Any advice is appreciated. 23 yo really depressed and feel lost.

Statement 19: All I have to do is browse the reddit comments. "You look like mom's neck beard basement dweller", "I would date your second chin". And I think that some of these "roasts" could apply to me. I am a woman. I have been told I am attractive. I think I am. I only weigh 110 lbs and I am 5'6". But I still have excess weight I gained after losing a pregnancy and being put on birth control. So the double chin comment really gets to me. Everytime I see a negative comment, I think about how my closest friends could say or think the same thing about me. Especially since one of my closer guy friends is a guy I crushed on, recently told and got rejected by in the nicest way possible. I have trust issues. I sometimes think I am not as great as I think I am in the moment. That my friends are internally roasting me in their heads while I look like an egotistical, arrogant, a-hole. Thank you for coming to my ted talk. Today's been rough. I have been struggling with suicidal ideation ever since I flirted with my guy friend again and made things awkward. It seemed like he was flirting at first but I think I misperceived the intent in his speech and screwed up again. there is more to this story but I do not want to get into it. I cannot tell if I am overthinking, paranoid and waiting for betrayal or if my intuition is on to some things. So I deactivate everything. I am staying away. I personalize insults aimed at other people then want to die.

Statement 20: At first I thought it was normal to not feel anything to hate myself ect then when I was about 6-9 I became suicidal which was when I got the idea of it so I did not tell anyone Am I the only one here? I have been depressed for all my life

Statement 21: Before I really wanted to be happy, i know it is not just something you can get but i was tired of being depressed all the time. At this point, I do not even want to be happy anymore. It seems like too much of a mess and too much effort for something that will not even last. I feel like i

have already tried so many times to be happy, and none of them worked. Even when I have people that love me, friends, hobbies, life still seemed so shit. Is this just it? At some point i just stopped attempting and accepted that I will never be happy with my life. The only way i see an end to this is suicide, but I am just too scared to do that either. I also know that i will hurt some people if i do die, and in the end, i just feel like I am trapped here with nowhere to go. It makes it all the more frustrating that i hate being here but i cannot just kill myself . I feel like i do not even want to be happy anymore. I just want to die instead of doing anything else.

Statement 22: that is it. I feel helpless, desperate and hopeless. My hair is damaged & weak because I am stressing out, I am stressing about not being stressed enough, I want to be busy. I do not want to feel like this ever again. I cannot live anymore. My hard work will never pay off, so I am giving up. Fuck this shit, why only me? From now onwards, I will never do anything. I will never work hard, I will never practice, I will not give a fuck about my career and future. The toll that all of this has taken on me is that now I feel happiness only when I think about suicide. No pain, no sufferings, No worries, nothing. I give up.

Statement 23: I am 25 years old and I have never had a girlfriend, had sex, been on a date, kissed, held hands, been touched by, been asked out, etc. I graduated college last May. At my on campus job I became friends with a girl who I eventually fell head over heels in love with. I have been in love with her since 2018. Was rejected and ghosted in 2020. She texted me last year to say hi, I answered and she never answered me back. I cry constantly and have been grieving every day for 3 years. The pain from never being able to be with her has made me suicidal and I have already attempted but failed to hang myself a few times. Its something I have to cope with every day. Its a sickness that never goes away. Its a pain that is not just emotional but physical. Again, I have been dealing with this every single day since fall of 2018. Something this is rooted in is my crush wrote poetry a lot and I bought a copy of a collection she published while we were still friends in 2019. In the collection there was a poem about her losing her virginity. Reading that absolutely destroyed me. It kept me awake for days. I could not stand the thought of someone else having sex with someone I loved so dearly and longed to care for so much. One day, she confided in me that her recent ex had been physically abusing her. More pain. Id fantasize about having sex with her but it would not help. I just imagined what it would be like if she actually loved me back and was attracted to me. Its getting to the point where everything related to having sex and relationships in general is painful to me. I even am starting to experience sexual dysfunction: I do not get aroused as easily and I can no longer watch pornography because it upsets me to see people being intimate because it makes me jealous and depressed that this is something I have never experienced before. I hear people talk about their gfs and wives and I get extremely sad. I found out my 20 year old sister lost her virginity months ago and when I did I cried for days. I felt indescribably disgusting and pathetic. I am 25. I should have lost my virginity first. I should have a gf or wife by now. Everything related to sex and relationships makes me violently depressed. It ruins my day. I am so fucking sick of feeling like this. Every time I hear about sex i just think Yep, another person experiencing what I'll never get to experience. Everything and anything related to sex makes me want to kill myself.

Statement 24: I am 17, been on Effexor for about a year. I was suicidal before starting it and now I feel mostly fine. I am going off of it because I feel like it has affected my ability to maintain relationships - both girlfriends I have had while I have been on it, I feel like I have not been able to actually connect with them emotionally, like I cannot fall in love with the

m. I have heard other people share similar stories. Also, I am sexually active with my girlfriend and I cannot cum – we have had sex probably like twice a week, for the last 2 months, and I have finished exactly 1 time. Its horrible. For a while I could not even get it up. But anyways, that is why I am going off it. What should I do to try to prevent my depression from coming back once I am off it? I am currently in the process of weaning off of Effexor. I am at like half of my dose of 75 mg, planning on cutting that to 0 in 2 weeks or so, once I run out of pills. How do I make sure my depression does not come back after going off of antidepressants?

Statement 25: I have been hiding my tears all day and my boss let me leave early. I thought spending close to \$1k would help me cope, but it did not. Now I am just home and feeling worse. I want to attempt suicide once more, last time I really almost did it, until my brother found me. It just brought shame and made me want to move so my family will not find my body. My only real friends are all online and I feel really lonely. At this point I am thinking of making a plan, find my cat a home, give away my belonging and then end it. I am done trying, nothing ever works out for me. I want to be free already

Statement 26: let us see.... where to begin. I feel like I am stuck in a rut. I have 2 people I can truly call a friend, I bury myself in work (80+ hours a week), and on my spare time I sit at home alone. I hate myself 90% of the time, my depression/anxiety is a bitch, I feel like I would be better off dead, or at least alone. I crave attention, but hate when it is given. My thoughts are dark, sometimes scaring myself, its like I am in a race with myself, and I am losing. I am single, not attractive, early 30s male. I have a job, car, apt., motorcycle, etc. but something is missing. I do not actively seek out a relationship anymore, to me it seems pointless. Any ideas on how to cope? Oh, I am in no way suicidal. I just feel like I would be better off alone, falling off the radar, or dead. How to cope?

Statement 27: I went to residential for 2 weeks and it was amazingly helpful. It did so much to change my outlook on things. While I was gone, my wife realized she can not keep doing this. I spent all of my momentum coming out of residential on convincing her to at least give us a chance. I was able to get a job, only one place even reached out after I applied. It is physically wrecking my body. I am in pain every day. The pay is crap. Plus I am sleeping in the guest room which has a shitty mattress. I do not know if I can make it through this. I got home from work tonight actually feeling pretty good. My wife was out so I just lounged around for a bit. I felt like I was getting some of my confidence back. Then it came crashing down around me. I am still a fat ass, unattractive. My wife told me before I even went to residential that we needed to stop with the physical in part because I need to lose weight. Before she got home I moved my shotgun out of our, sorry *her* room and put it in my closet. Just in case. I promised her months ago that suicide was off the table, that I was all in. What good did that do me? I applied to yet another job. If I do not get it, I think that will be the end of it all. It will be the first time I fire my shotgun. With my luck it will not work. I can not even work up enough to care. Falling Again

Statement 28: As much supportive as my bf is with my depression, its so fucking scary to tell someone that would never do this that if it was possible to kill myself painlessly, I would. I think I got to the conclusion today that my life is not going to get better if I do not admit this. However it feels like such a selfish thing to do because I do not him to worry all the time and this to mess with his mood. I feel like I am selfish with everyone for not caring because I am so depressed, and then I also feel like its selfish if I tell someone that I would like to die because of the burden I am putting on them. Please help me by saying how you told someone you felt this way and

how they took it. Note: I am aware that I need a psychiatrist and probably taking meds. How to explain to someone your suicidal thoughts

Statement 29: I am just fed up of everything. Everything and everyone pisses me off and I have tried to be happy but that just does not work. I have made plans of suicide for sometime after I pass my driving test and get a car. It's the only thing I can think of, I have tried self harm but I am too scared of pain, thought about jumping off a building like my brother tried but I am a coward and anything else is too painful or I am scared to do it. I know I have not passed my GCSEs because I did not try so I probably will not be able to get into college, I am so negative to everyone around me I am scared I am losing friends. Everything about me is just a fucking waste, anyone else could have been born in my place and done better, children are dying in impoverished countries who could have had great potential but people like me were born who are selfish and useless. I do not know what to title this

Statement 30: I have been reflecting on my life a lot lately, and I always find myself going back and thinking about the bullying I have dealt with throughout my life. Hopefully sharing my story here can get some outside perspectives and give me an idea on how to move on. So a little about myself, I am an introverted guy whose hobbies fall in Video Games, anime and music. I have been playing video games since I was 3 and they have been my major escape for most of my life. I grew up in a small rural town for most of my life (From when I was born to my senior year, am currently 25) and was actually a pretty optimistic kid in my youth, free-spirited and just looking to have fun and make people smile. My parents have been pretty decent, but they were never around to support me so I mostly just stayed in my room playing games. It was around the age of 4 that I first experienced any sort of bullying. A couple of my older cousins would pick on me for some of the games I played. (I was a PlayStation kid growing up and they were Nintendo kids and around 8-9 at the time this was happening) Unfortunately nobody really stepped in to stop them but things never went beyond verbal insults. This went on till I was about 8-9 myself and they grew up enough to stop it and we got along on the rare occasions I see them. Elementary school was pretty uneventful as I mostly just kept to myself. Middle school was where things took a turn for the worse. In my 5th grade gym class we had to start changing into gym clothes and I never got the memo that boxers were even a thing and was still wearing briefs at this point. We were all getting changed and I realized what was up, got really nervous because I noticed everyone else was wearing clothes that I did not even know existed and went into the one stall in the locker room to change, needless to say the locker room was full of boys making fun of the fact that I was wearing briefs. It was at this point most of the male classmates in my class and the grade above me started just verbally making fun of me, and the female students would either just do nothing or tell me to not make eye contact with them. By the time I got to my 6th grade, I only had a few friends and most of them were not actually friends with me and ended doing some things to me. I was getting no support from family or teachers and was starting to feel isolated, so I tried to fake being cool, which did not really do anything in the long run. It was also this year where I saw another student in my grade wearing sweatpants to school, and since I had some figured I could wear some. I remember walking into class and a few classmates giggled a bit amongst themselves. I went to use the bathroom at some point in the day and on my way back was stopped by a female bully, who proceeded to call me gay and saying wearing sweatpants is "for girls" and other verbal insults. This was the only time I ever cried in school for something like this. I feel like I had to grow up really quickly emotionally in school due to a mixture of all this bullying and teachers would also make me hang out with special needs students, even if I did not want to. In 7th grade I had a crush on another girl

in my class (She was one of the only people in my class, let alone the school that actually treated me like an equal person) and some of my "friends" physically dragged me over to her after I let them know that I had a crush on her and made me grope her breast, all the while I was begging them not to. (Thankfully she did not hate me for this) 8th grade was the first time I can remember wanting to die, not exactly wanting to commit suicide but wishing I could just cease to exist. I started to become really jaded, I could not trust anyone save for a few friends that I barely saw because I was too busy still trying to live this "cool" facade externally but just feeling so angry inside. Being exposed to all these people hating and belittling me made me start loathing myself too. Near the end of the year, the same group of "friends" from earlier pressured me into smoking some cigarettes that they managed to find. (I do not even remember where) I took one whiff, and in the back of my head I just said "that is it, I am done with these guys, this is not what I want to be" I left that school with one friend that I had just made that year, social anxiety and self-hatred from all the harassment. I remember that summer feeling really lonely, like I was just existing but I lost a lot of my drive and happiness. High school was better at least, there was still bullying, nothing compared to my old school but I think some people could sense my insecurities and would pressure me into saying or doing things, I tried to own this "one man wolf pack" loner sort of shtick but deep down I just wanted to be accepted. Things ended completely by Junior year but by that point the damage was done. Nowadays I have a small group of friends who seem to like me for who I am. I still deal with the insecurities and fears from all those years ago and feel really emotionally stunted, like there is still this scared kid trapped in the body of this cold, logical body. For those who took the time, thank you so much for reading all of this, any advice on what I can do to move past my fears and insecurities are appreciated. Hopefully I did not come off as too whiny. Looking to share my story of being bullied.

Statement 31: My depression is fucking everything. I have had to quit 2 different jobs (mechanic and lawncare). Both should not seem stressful but each caused my suicidal thoughts to sky rocket. I just need something low stress that likely will not because this what is a low stress job?

Statement 32: Life is fcking crazy, I want everyone to realize how sad and weird life can get. I have been struggling deeply for the longest. To be honest I do not know why I am still going but I am hoping one day I can look back and be glad that I am still breathing. Everyday feels the same. I am constantly thinking about suicide but Ill probably never do it. Its a really sad feeling when you have no one to turn to. I am so scared of the future and where Ill be. Am I destined to be alone forever?? I have always thought it was cringe how people just rant on Reddit but here I am on Reddit. I hope I can look back on this post and remember these times. If anyone ever reads this.. know that we are looking at the same stars and moon and that you are not struggling alone. Everyday feels the same.

Statement 33: Sorry if this sort of thing is not allowed, I just need to tell someone. I tried to kill myself when I was six, and once again a year later. I have never stopped being suicidal, I had to go to the hospital when I was 14 because I felt like I would kill myself. Since I was 12-13, I have been planning on killing myself when I am 19, the only reason why I did not kill myself when I was 14 was because I wanted to give myself a few more years for my life to get better. I have 2 years until I am 19, my life has only gotten worse. I have been saving up to get a (mentally) life saving surgery when I am 18. My family struggles with money and my dad does not pay child support, I want to kill myself so the money for the surgery can go to my family. Paying for one funeral will be a lot cheaper than spending money on me for the rest of my life when I will never even be happy. depressed

Statement 34: I am tired of feeling tired of being depressed. My life is a bucket full of shit in which I stick my head into every morning. Most people are shallow and act only for their self interest but cover that under a veneer of false benevolence and think we cannot see through their bs. We live in a sick world where we only think about ourselves, use nice people ditch them then repeat the process. The idea of suicide seems more and more attractive to me as days go by. Life is made of pain.

Statement 35: I was just on my bed, feeling suicidal as usual, and then I decided to try browse r/nihilism. I know it sounds odd but hear me out. The realisation that nothing is going to matter kind of helped me a bit. The ability to separate myself from my feelings and thoughts. Nothing will matter in the end, but that does not mean that you will never be able to feel validity in happiness. It is ok to just wander through life enjoying it until we die. I just got out of a suicidal spiral in 5 minutes and I am oddly proud of myself.

Statement 36: Is anyone having a particularly bad time right now? I really need to get outside my head, and maybe gain some perspective through the problems of others. Last summer I lost my job. I have been working towards improving my symptoms, and I do see a steady improvement. A few months ago I could not sleep for more than 3 hours. Falling asleep was almost impossible. My anxiety shot through the roof, as I was having daily panic attacks (a new thing for me). I lost my job, went back to school and got my degree. I am worried all the time. I ruminate all the time. I was bedridden for a few months and I lost about 50 pounds. My muscles have atrophied. My teeth rotted as I was not caring for myself. I am in a lot of pain physically and mentally. I cry multiple times a day sometimes uncontrollably. I will be 30 in less than a year and I feel like I am already dying. All I can think of is the negative, even though I have some positives. I am relatively healthy, no serious physical diseases. I have an amazing and beautiful girlfriend. I love her more than anything and she always returns the love tenfold. She has been trying to support me during this very dark period and I feel horrible for dragging her down. When the depression and anxiety were at their peak I was considering suicide every minute of the day. I had to call the suicide hotline and the lady I talked with used a tough love approach which actually helped put things in perspective. Anyway, even though the symptoms have been improving things keep happening in the middle of my recovery that trigger the anxiety and depression. Just yesterday when I was feeling quite good about my situation, I became the victim of identity theft. Here we go again. One step forward and 10 steps back. I obsess about things, and it has been driving me crazy. There are compounding problems contributing to this hole I have dug besides my mental health. I need a kind stranger to tell me what they are going through as well so I do not feel so alone. I am hurting really bad. My heart and spirit feel broken. Anyone else? Anyone Else Having The Worst Year of Their Lives? I desperately want to feel less alone and reading about the what others are going through helps.

Statement 37: My partner broke up with me. We dated for 8 months and were extremely co dependant. It was not a healthy relationship because she did not want to get better and I wanted her to. We met up and she is with someone else as soon as a month after, an extreme alcoholic drinking 750mls every night and I am still in love with her. The breakup is what originally sent me into the depression, but it was not just that. Where I live was going above and beyond with covid restrictions even though people were getting vaccinated and you would be fined for even seeing your family. My parents were being insane, my brother was not doing well, and it took a huge toll on me too. More than I know, I think. Then I made the impulsive decision to get an apartment right after the breakup. The move in date is September 1st and as much as I s

ould be happy, that makes me even more depressed for some reason. I have no friends with depression, I have no one that talks to me constantly (my ex did so I am feeling extremely lonely after just getting cut off from that), and I just feel so exhausted and alone. I want to go out, but I want to be alone. I feel nothing – literally nothing. I am just a numb husk that gets stuck in depressive thought loops over and over and wakes up to panic attacks. I have been in bed for over a week now, my room is beyond disgusting, I cannot eat and look like skin and bones, and my Dad basically has to take care of me because I cannot. I have gone out once or twice but I feel nothing the whole time. I cannot live like this. I have clinical depression and have extremely intense depressive episodes multiple times a year but this is the worst one I have had since I was in fucking jr high. (I am in college now). I think I am suicidal but I am so numb and empty I cannot even tell. The healthcare system in general, let alone the mental healthcare, where I live is abysmal but I am genuinely considering calling the crisis line and getting myself admitted to the mental health ward. The only thing stopping me is how much it would affect my family. I just feel so lost and SO alone and I do not know what to do. I think I came here to just talk to more people who are struggling like me, but if you have any advice please let me know. I am so tired. I am in one of the worst depressive episodes of my life.

Statement 38: I have been walking around being productive, I have a boyfriend, I have emancipation papers (not filled out), I have this big plan on how I am going to go out and live life and pursue college. I have told all my friends that I am going to leave my family (which I feel like is emotionally and mentally abusive but I am not sure) and go get an apartment and finish high school and go straight to college and then Med school. Problem is, I have not actually done any of that. This past year I got pulled out of school twice for suicide watch, my grades plummeted, I quit my sports team and I cut off most of my friends, I stopped showing up to work and I just tried to not to kill myself. Recently its been getting worse at home. My father has become more controlling, grounding me for stupid things, taking away my car, not letting me see my boyfriend or friends, gaslighting me for not spending time with him, etc. My mother has gotten better she has not manipulated me in a while so I think were fine. Her husband on the other hand, says that I am the reason my mothers health has deteriorated, both of my stepparents have said that I am the reason our family is dysfunctional. Sometimes I like to think I am stronger than others, that I have it worse than my friends, but that is not true at all. I am weaker than all of them. My life is so easy. So why should I want to stop existing? there is been awful shit that is happened when I was a kid Ill admit, but as of late lifes not as bad as it could be. I Do Not Know I just feel so stupid. Living the vida depressy

Statement 39: I have been waiting for my death like looking forward to it , expecting death to happen today while traveling and wanted myself to fall from the vehicle . Is this a symptoms or it is just that I am sad now Am i suicidal person ?

Statement 40: It has been 7 years since my brother died by suicide and honestly, life has never been the same but not better. He killed himself using a gun owned by my father and left a suicide note written in japanese (he self -studied the language). We never saw the whole note since it was surrendered together with the gun to the police and my parents never wanted to read it. After his death, they burned all of his things (laptop, books, etc.) and never talked about the reason of his death ever again. There was always a time once a year when my curiosity got the better of me (I asked them if they could get a copy of it or if the police can translate it for us) and it always ended with disastrous fights with me against my parents. I always feel like I will never have my peace of mind if I will not be able to read about the l

etter. In the end, I always feel like the note could be some sort of closure for our family. Today, I feel like I just want to follow my brothers footsteps. My parents became more emotionally absent. They always feel like they are right and their ideas must always be followed. I also just feel like a burden to this family. I do not even see myself alive after 5 years. I just want to die. My brother died by suicide

Statement 41: I have been trying to hold out on suicide for 4 years now, and last night, I remembered something horrible I did and broke down and lost it. The thought of it scars me and the people I hurt do not believe me when I genuinely say I am sorry either. Piled with the suicidal thoughts I have daily and family conflict was just too much. The only reason I am still here is because of my family and the friends who have not ghosted me out of concern for them. They cried begging me not to go and just hearing them like that is the only reason I have not tried again yet, they do not know how much it hurts hearing them say that. I really do not think it will get better and wished I was never born so everyone could be so much happier. I do not understand how people feel I deserve life when all I am is a waste of time and resources. I wish I never existed

Statement 42: I am going through a tough mental week and I feel like I would be the person who "randomly" commits suicide. I seem happy af when I talk to friends but deep down I feel like I am dying and no one really knows or fully understands how deep it goes. Just needed to get my thoughts out of my head so thank you if you read it and I am sorry for anyone who feels the same way as I do. Horrible intrusive thoughts/a need to vent

Statement 43: 21M. So. I have been feeling depressed, very anxious and very low the past while. I grew up in a abusive household until I was 18, I moved in with my sister for 3 years and that turned toxic and it did not work out (she was angry, hatred, thinks she still hates me), I moved into student accommodation yesterday and I am still overwhelmed and depressed. I have been like this I was a little kid. but what is happened over the past few days just made me feel so low. My family were basically just my brother and sister. My sister now thinks I am backstabbing her for leaving. I tried to explain but she got over sensitive and angry, it is for the best. We were living in a 1 bedroom small flat. She was annoyed and unhappy over time. So I decided to leave, but she was not having any of it. So I got the bus an hour away from home to student accommodation. I got off the bus, ran to the bathroom and took a panic attack. I was there for 45 mins. Left walked round town and checked in to the accomodation. did not eat anything just spent most of the day annoyed and crying. So I was sitting for hours last night and this morning just anxiously looking to click that 'call' button on the support line. I explained my situation. They asked questions, asked about my GP and I agreed to have another assessment. They said it was a suicide line, but I have dark feelings, I feel hopeless alot, pretty much every day since I was no age. But I could not bring myself to end it, I just want the pain to end. It felt ok taking the first step. I know I needed help for years. I do not know what to do with the sister situation. I am still very low. But if anyone else is feeling low or suicidal. If this quiet, socially anxious weirdo like me can do it, anyone of you can I rung a support group

Statement 44: well, i do not even know how to start this. I am so confused and lonely. just as info I am a 17 y/o female. i have no idea if i even have depression but reading all of your stories, i recognize myself in most of them so i guess I am one of you. I have been experiencing extreme self-hatred for the past 5 years, triggered by a tragic event in my life and then developing rapidly. its like my thoughts are my biggest enemy and I am constantly bringing myself down for the things i do and the way i look. generally, i do not share personal thoughts with people, its actually the first time i am spe

aking about this. i just ignored it but lately its been getting so troubleso me i cannot sleep and I am always in a bad mood. suicidal thoughts have been coming up as well which scares me a lot.furthermore, i just feel apathy towa rds everything and everyone. yesterday i got my first job offer and i just f elt numb. i know I am not worthy and i do not deserve it so i just did not t ell anyone about it even if it was an achievement. same thing with good grad es, i just do not care anymore. guilt has also been my best friend throughou t these years. something happens and then i just put the guilt on myself. it s like i know it straightaway even before considering the situation.its just getting worse after every day and i am really lost. thanks for reading and s orry i bothered you with my thoughts. just another depressed buddy here

Statement 45: Last year I got into music production and i quote it as the ha ppiest year of my life.I was working on music 6 hours a day (at least). I on ce managed to pull off working for 12 hours.I really loved what I was doing , but this year..its different I do not know how.I am just lost , I can bare ly work for 2 hours and worst thing is games are back into my life. In 2020 everything that could possibly happen that is good to me has happend. I lost all my weight and got on some serious muscles all by working from home. I ha d such a mentality for work that year , it was ahead of everything.Back then I loved music so much I managed to not play games at all , I once did not bo ot up anything for 2 weeks straight. Now , I play games almost everyday (for 2 hours or so) and the worst thing is this fucking game Hearts of Iron 4 got me hooked onto it but it did not made me political. I just remember the days when I was a fat piece of shit and was playing that game in 2016 and I get a nxious about all of that coming back right at me.I even uninstalled the game few days ago in hopes of my work hours going up but it did no such thing.Wha t I am trying to say is that I lost the working spirit and I do not know how , I am genuinely not a lazy person.There is a person in my life that is stre ssing me a lot , I think about it everyday and it does bother me.I do not wa nt to talk about it but I do want to meantion it , I was mad at the same per son last year but I was still managing to work very hard.I am not suicidal n or do I think about it , just trapped in an endless loop of anger towards my self.Feel like I am in a minefield , that I have to look at every single ste p I take.There is something to meantion though , I have quit drinking coffee and any unhealthy drinks (energy drinks , coca cola , pepsi and alike) for 6 months now.Those drinks gave me the ability to stay up at night and make mus ic , which was perfect because every producer encourages that , its just lik e magic when you work on music while your brain is half dead.But now I canno t make it past 10 , what is interesting is that I can play games untill 1 in the morning but not make music and that is 1 of the many reasons I hate myse lf. What is wrong with me

Statement 46: The memories I suppressed for so long suddenly triggered toda y. I remembered how I used to feel like I was trapped in my own world and ho w i was suicidal in junior school, at around age 10. I remember how my mum u sed to go into rage at mistakes i would make or me being annoying. I remembe red how mum would tell me to put my arm out for her to hit me, and if i pull ed my arm back in fear, she would threaten to hit me even harder. I was real ly fucked up back then. Every other night Id think about running away, never having to face anyone who would hurt me or berate me or make fun of me like my classmates and teachers at school would or like my mum would at home. I t hink today was the day my old self that i had abandoned to heal finally came back to haunt me. Just needed to write this down. I just needed to write thi s down

Statement 47: So introduction I guess.. my name is Michael and I am 20 yo. I live in Slovakia though I would be surprised if anybody knew where that is l ocated. And I am struggling with..I do not know maybe deppresion?The problem

s started like long time ago back when I was still a child. My mum hated the fuck out of my dad and for some reason my dad hated her too. My dad is like the typical stereotype guy who thinks that woman should be in kitchen taking care of her husband. He was and still is to be honest addicted to alcohol. I remember when I was a child and my mum had to go away for a couple of days that everytime my dad stepped through front door I prayed to god because I was always afraid that he arrives drunk again. I was so genuinely happy when he did not and I always hid in my closet when he did. In school I was not very popular. I am the kind of guy who thinks that if you have the means, if you have the option and ability to help somebody you should by all means do that. Of course people tend to take advantage of that. Everytime they needed something I would do it eventhough if it meant that I had to spent 4 hours doing their homework or writing an essay. Of course I was bullied too when they did not need me because why the fuck not right I deserved that. I did not have a single true friend everybody I considered to be my friends they insulted me and humiliated me most of the time and I did not like it. I was so alone that I just decided that I had to have a friend. And that is when my schizophrenia probably started.I made myself a friend. He is called Martin Lade I even created him a facebook account so I could chat with him because I had always wanted to have friend that I could chat with on facebook. Everytime I was alone I talked to him. My imaginary best friend. It all started like nothing too serious every conversation I had with him always played in my head like I only imagined that I was talking to him because I had known that he was not real. But then I started talking out loud to him.. It just went like this I do not really know how to explain it properly. I just sat on the ground and started crying and telling him how I was feeling but after I was done talking I got up smiled at the spot where I had been crying and tried to reassure the spot you know like if I was Martin cheering up Michael. After some time I got better but then I started hearing voices it was around 8th or 9th grade. They were telling me how pitiful I was, that I should kill myself that everybody hated me. I started to feel really depressed and it was not helping when the actual people were telling me how pitiful I am and how they hated me. So at the end of elementary school my mind was shattered like glass into so many parts. Part that wants to kill me, other part who argues with the first part and telling me to live, Martin and myself.Then the high school came and I found new friends..like true friends. I liked them so much and they seemed to like me too. There was this particular guy that I enjoyed spending my time with we had so much in common. He loved the same type of music I did, movies, TV series and so on. Of course I had fallen in love with him. I did not even realized that I was gay but when I did I just went along with it. But I loved and still I love him so much..I guess I was really obsessed with him. I was jealous everytime he went out with somebody else (of course i did not tell him I am not dumb I knew it would ruin everything) I was thinking of him everytime I went to bed and everytime I woke up. He had a best friend, our classmate, and the three of us started hanging out a lot. I t got to a point when we had like really deep shit conversation and I remember that the other guy (not the one I love) told me (because he could see that I had been struggling) : ,,"The stupidest thing that a person can do when he has some problems is that he never tells anyone about those problems. Talk about it." and then the guy I loved said : ,,"We are telling everything to each other. Tell us." And I told them because I believed in them. I told them about the voices and about Martin and everything. It went fine it seemed they believed me but of course.. shit had to get worse. (for the sake of this the guy I love is : T and his best friend is : S)At one point I do not know what happened but S accused me of lying. He said that nothing that happened to me never really happened to me and that I like the attention I am ge

tting because of that and I just want people to pity me. He told other guys in my class and so I became the guy who is constantly lying, exaggerating and depressed. I never stopped talking to them though. I understood that this shit is really hard to believe. T was still my number 1 friend because he seemed that he still believed me. I started to love him even more. I have always had money I did not even know how. From insurance, when my grandma and aunt died and what they left me. So always when I needed or wanted something I had it. And he did not have the luck he had like no money. Like he was not poor or anything but when he wanted Converse for example he could not afford. And I felt selfish because I had the money and he did not so I felt like it would be really selfish of me not to buy things for him. I bought him those. And trousers. And jackets. And games on ps4 (which by the way I bought just so I could play with him). I remember that once he told me how much he dreamed of getting a new computer but his mum did not let him have it. So I literally started building my own new pc for 600 even though the one I currently had was completely fine. The old one I had cost me around 500. I told I am that I had built myself a new pc so I can give him my old one. For free. And I did. I literally spent on him around 600 because I gave him my monitor because he did not have one and I had to buy myself a new one. And I still was so happy when I bought those things for him and gave those things to him. I loved him so much like this is probably the true love you know. Then last year came and it all went to shit. It was my last year on my school and we had a party the entire class after Stukova. (Stukova is like the Ball you have in last year of high school. After Stukova you have Dozvuky where all students just drink a lot). S dated my classmate but they broke up before Dozvuky. I was really good friend with S' ex at that time. I am not going to explain what happened exactly this post is long enough even without those details but basically S "fooled around" with two girls. Two girls his ex hated the most. She was really sad and I was talking to her the entire time. After Dozvuky and everything I met him and he told me that I really fucked up. He told me that I disgust him and if I feel hurt that I deserve it. He blamed me for something I did not do. He did not want me to explain to him and he stopped talking to me. In the same week every guy from our class stopped talking to me including T and I realized that he told them his version and they all feel same as him about me. It could not come in worse moment. I was feeling during this time depressed a lot. I had a friend (M) She was like my best friend after S and T. She knew that I was gay and she knew that I loved T. Still she fooled around with him when we had parties. When I was there. 3 times. Everytime she did it she told me that she was sorry and promised me she would never do it again. And she did it two more times. One time i was in the same bed as T and M and they started making out next to me. It was the worst feeling ever because I loved him so much and she knew and still she did it. I was heartbroken and defeated and I started cutting myself. And then S dumped me because of the misunderstanding and every guy dumped me and I was alone again. Fortunately I am not that stupid and I realized that I couldn't continue like this. I am proud of myself that I realized that I need professional help because most people cannot do it. So I started looking for a psychologist at first. Someone to talk to. I tried 9 different ones. And they all told me the same thing. Only time can help. I did the Rorschach test numerous times and still they were not able to find anything wrong with me even though my every goddamn answer was really depressing, negative and not positive at all. I cut myself. I hear voices telling me that I should kill myself or the ones that hurt me. I created an imaginary friend. For fuck sake even I as a patient was able to realize that something is a bit fucking fishy about me. They did not help. One psychologist even tried to touch my thigh. So that much for professional help. So I started looking for a psychiatrist . I t

old him the same story and he gave me pills. Arisppa. It did not help. The best part is that I told him that I had told S about my voices and stuff and how I hurt felt when he did not believe me and accused me of being a drama queen and guess what did the fucking psychiatrist asked? " Michael, are you telling me the truth? Do you really hear the voices or have you made this all up just to skip school or be different? To let people pity you?" Is this even possible in modern world? in which fucking dimension would sane psychiatrist ask such a dumb question? I snapped and I started laughing hysterically after that. I was laughing at my place for good 3 hours and then I started crying. And laughing after that. Some people just insult those who kill themselves and say that it is selfish decision. I totally understand suicide now. I have problems I accept that. I wanted help I desperately needed help but everything is against me. In this world if you are not selfish bastard you cannot be happy. This world is not made for people like me. it is really easy to say that " get some help" because apparently nobody can help. Like I do not even know if anybody is going to read this or not. But if some lost soul read this entire novel to the end just wanted to let you know that do not worry I promise I WILL NOT kill myself. I tried sure but I am not strong enough to do that. Besides my paranoid me thinks that it would make other guys happy and I do not want to give them this satisfaction. Honestly I do not know what I expected from this. I do not know if I want your help because I know you cannot help me. But I tried everything. Psychologists, psychiatrists, telling friends, group therapies, fucking taiwan mantras to calm my mind, tropical teas and herbs to soothe my mind. Nothing worked so I just thought fuck it let us give it a shot. Maybe miracle will happen. do not know what else to do so I just leave it up here

Statement 48: I am a 15M. I smoke weed as a way to cope with my anxiety and depression that gets worse at night. I also use it as a way to fall asleep easier. My dad found my weed last night and told me he was disappointed, he said it is not a good way to deal with things. He said he cannot force me what to do or what not to do, but he trusted me enough and handed it back to me. I do not know what to feel about that. Right now, my dad is also having troubles with my mom, and sadly my little sister got stuck in the crossfire. My dad is barely talking to anyone like he used to. He does this very often. When he gets angry, he shuts himself off and does not talk to anyone. He probably has a mental illness of some sort, he is a very complex person. My mom once told him to go see a psychiatrist but he refused to because that is just how he is. He does not have a good relationship with his own father, which might be a because or something. I also recently got diagnosed with mild depression and anxiety and got prescribed with Lexapro. I have been taking them for a week now. My mother is the only person I can always talk to but I cannot tell her about the weed thing because I got caught before in school and she knew. I want to explain to my dad about everything but knowing that he is the person that he is, a part of me also does not want to open up to him. he is a sensitive person, something so slight can trigger him to the point of shutting himself off from his family. Smoking weed has always been that barrier between me appreciating life and jumping off a cliff. I do not know if I should continue smoking or not, I feel so guilty if I smoke because I know my dad is probably stressed about this situation also. I am pretty well known in my city, I know a lot of people but that does not help either. I cannot trust anyone at all. I have only 2 real friends, my best friend and his girlfriend. I am so young, I do not know why I am feeling this way, it is so overwhelming. I just want to sleep all day and never wake up. I want to kill myself but in a way that no one would be sad about. I know there are people that care about me. My mom does. My dad does. My friends do. But when you are suicidal, nothing really even matters anymore. I feel like I cannot take this. it

is so overwhelming. I just want someone to read this and help me understand myself better because I barely know myself anymore. I hope someone reads this to the very end

Statement 49: I have always hated myself through my 17 years being here. but i never really know if I am actually sad as fuck all the time or doing it for attention... hearing everyone talking about their past and traumas makes me feel like i honestly have no right to be suicidal and sad. I am honestly going crazy trying not to end it all here. all of my pillars of support are in different schools and i barely see them as much as i used to.. now i spend my days in my own school barely talking to anyone and leaving straightaway every time my teacher dismisses us. yea sure i have a boyfriend and i know that i can always talk to him but i still feel so alone. am i bad person? am i faking it..?

Statement 50: "you have been procrastinating from more than a year you piece of shit! And now that you are responsibilities r butt fucking you, you r tryn a blame lockdown, covid and isolation for this huh? How big of a coward you are, you deserve to die. Nah wait, you thinking about suicide from a few months? you coward tryna exit you are problems like that? Ever thought how you are family will feel? Oh you r such a selfish coward, you definitely deserve to die. I think you were born evil, an absolute madman, go die you fucking piece of trash, you will die regardless. And now you posting this paragraph on Reddit to get some sympathy huh? 17 year olds have nothing to be depressed about, it is all in you are head. you r responsible for every bad thing that is happening, yes I said you! Been isolating yourself from a year, who needs friends huh? Just go lock yourself in the bedroom again and pretend to be good around others. you told you are parents but they do not care, no one cares for you, you ugly fat motherfucker, no girl will ever like to be with you. you deserve to die. Just fuckin' die and end this shit once and for all. "I honestly do not want any sympathy at this point. Ik I am responsible for everything, I have been responsible for everything that has went wrong with me. guilt and shame r going to kill me someday, I guess that day is coming soon. I am sorry that I made you read such a long paragraph that does not even matter(does not even make sense, I am just insane. No, I am just a dramatic piece of shit) . I really hope you all have a nice day ahead. To all the people suffering with any type of mental illness, I hope God gives you the strength to survive through it, I believe you r strong enough. Once again, I am extremely sorry for my rant. I have just gone insane maybe(nah tryna push my responsibilities once again). Just downvote this post. Peace Just Fuckin' Die

Wow! It seems like these are all mislabeled! Jeez Louise!

Let's run the value counts on our data to see how much this will alter our class imbalance.

```
In [130]: filtered_df['status'].value_counts()
```

```
Out[130]: status
Normal           15924
Depression      15086
Suicidal        10637
Anxiety          3617
Bipolar          2501
Stress           2293
Personality disorder  895
Name: count, dtype: int64
```

We will change all of these to 'Suicidal'.

Word Count -- Feature Selection

Upon closer examination of these statement values, it's become clear that certain posts have a far higher word count than others. Furthermore, higher word counts seem to be associated with long tangents of vented distress.

Maybe we should add a word_count column that counts the number of words in each statement, and see whether or not there's any clear difference in the distribution of numbers by status that might help us better predict our target variables.

In [131...]	filtered_df.head()					
Out[131...]	status	perc_upper_words	neg_emoji	exclamation_count	question_count	tokens
	0 Anxiety	0.0	0	0	0	[oh, go
	1 Anxiety	0.0	0	0	0	[trou
	2 Anxiety	0.0	0	0	0	sleepi
	3 Anxiety	0.0	0	0	0	confus
	4 Anxiety	0.0	0	0	1	mi
						restless
						[wro
						back, de
						forwa
						dou
						stay, res
						[shift
						foc
						somethi
						else, s
						worri
						[restle
						restle
						mor
						boy, me

```
# Function to calculate word count
def count_words(text):
    return len(text.split())

# Apply word count function to 'statement' column
data['word_count'] = data['statement'].apply(count_words)

# Group by 'status' and calculate mean, max, min word count
grouped_data = data.groupby('status')['word_count'].agg(['mean', 'max', 'min'])

print("Mean, Max, Min Word Count by Status:")
print(grouped_data)
```

status	mean	max	min
Anxiety	148.093724	1673	1
Bipolar	184.722111	4850	4
Depression	168.955326	4245	1
Normal	17.914646	264	1
Personality disorder	184.788827	5580	5
Stress	115.350632	1648	3
Suicidal	146.726811	6300	2

Wow! **Huge** difference with regards to 'Normal' users VS. users with a mood disorder.

Let's add this to our filtered_df.

```
In [133...]: filtered_df['word_count'] = data['word_count']
```

Dropping 'Personality Disorder'

Lastly, 'Personality Disorder' is an entirely unhelpful label. We have no idea what this means or how to harness domain knowledge to engineer feature selects that might help us predict 'Personality Disorder' as an outcome. There's a strong possibility it overlaps with our other status groups.

We kept this group in our dataframe, in case our exploration of relevant tokens/bigrams, and running subsequent models, helped us understand what kind of condition it's meant to label...however we've found no new useful information and our models are having trouble running. Since this is such a small portion of our dataset, we can drop it entirely.

Furthermore, combining certain status groups, and dropping others that are improperly labeled and in the minority class **might help with the class imbalances which are potentially affecting the success of our model attempts.**

Applying Changes

```
In [134...]: # Drop rows where 'status' is 'Personality disorder'
filtered_df = filtered_df[filtered_df['status'] != 'Personality disorder']

In [135...]: # Relabeling stress as anxiety
filtered_df.loc[filtered_df['status'] == 'Stress', 'status'] = 'Anxiety'

In [136...]: # Relabeling depression rows that should be suicidal
# Step 1: Filter rows where status is 'Depression' and tokens contain 'suici
filtered_rows = filtered_df[(filtered_df['status'] == 'Depression') &
                             (filtered_df['tokens'].apply(lambda x: 'suicide' in x))]

# Step 2: Update status to 'Suicidal' for these filtered rows
filtered_df.loc[filtered_rows.index, 'status'] = 'Suicidal'
```

```
In [137...]: filtered_df['status'].value_counts(normalize=True)
```

```
Out[137]: status
Normal      0.318111
Depression  0.266451
Suicidal    0.247413
Anxiety     0.118063
Bipolar     0.049962
Name: proportion, dtype: float64
```

This seriously helped improve the class imbalance in our model!

Let's try our models that had problems running, and see if things go more smoothly now.

Numeric Feature Evaluation - Random Forest

Our models are still not running properly -- it's a lot of data to deal with!

We have those numeric features that we've engineered...Let's use random on our numeric features only to hone in on feature selection and drop irrelevant numeric columns. Then we can limit the number of features in our more complex models once we combine our numeric and vectorized text data.

```
In [138...]: X = filtered_df.drop(columns=['status', 'tokens'])
y = filtered_df['status']

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

X_train_numeric = X_train[['perc_upper_words', 'neg_emoji', 'exclamation_count']]
X_test_numeric = X_test[['perc_upper_words', 'neg_emoji', 'exclamation_count']]

param_grid = {
    'clf_n_estimators': [100, 200, 300],
    'clf_max_depth': [None, 10, 20],
    'clf_min_samples_split': [2, 5, 10],
    'clf_min_samples_leaf': [1, 2, 4],
}

pipeline = Pipeline([
    ('clf', RandomForestClassifier(random_state=42, class_weight='balanced'))
])

grid_search = GridSearchCV(
    pipeline,
    param_grid,
    cv=5,
    scoring='f1_weighted',
    verbose=0,
```

```

    n_jobs=-1,
)

# Fit GridSearchCV
grid_search.fit(X_train_numeric, y_train)

# Evaluate on the test set
test_score = grid_search.score(X_test_numeric, y_test)
print(f"Test set score: {test_score}")

print("Best parameters:", grid_search.best_params_)
print("Best F1 score:", grid_search.best_score_)

```

/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/joblib/externals/loky/process_executor.py:752: UserWarning: A worker stopped while some jobs were given to the executor. This can be caused by a too short worker timeout or by a memory leak.

```
warnings.warn(
```

```
Test set score: 0.5322223156636798
```

```
Best parameters: {'clf__max_depth': None, 'clf__min_samples_leaf': 2, 'clf__min_samples_split': 5, 'clf__n_estimators': 200}
```

```
Best F1 score: 0.5272141199535815
```

Visualizing Numeric Features by Importance

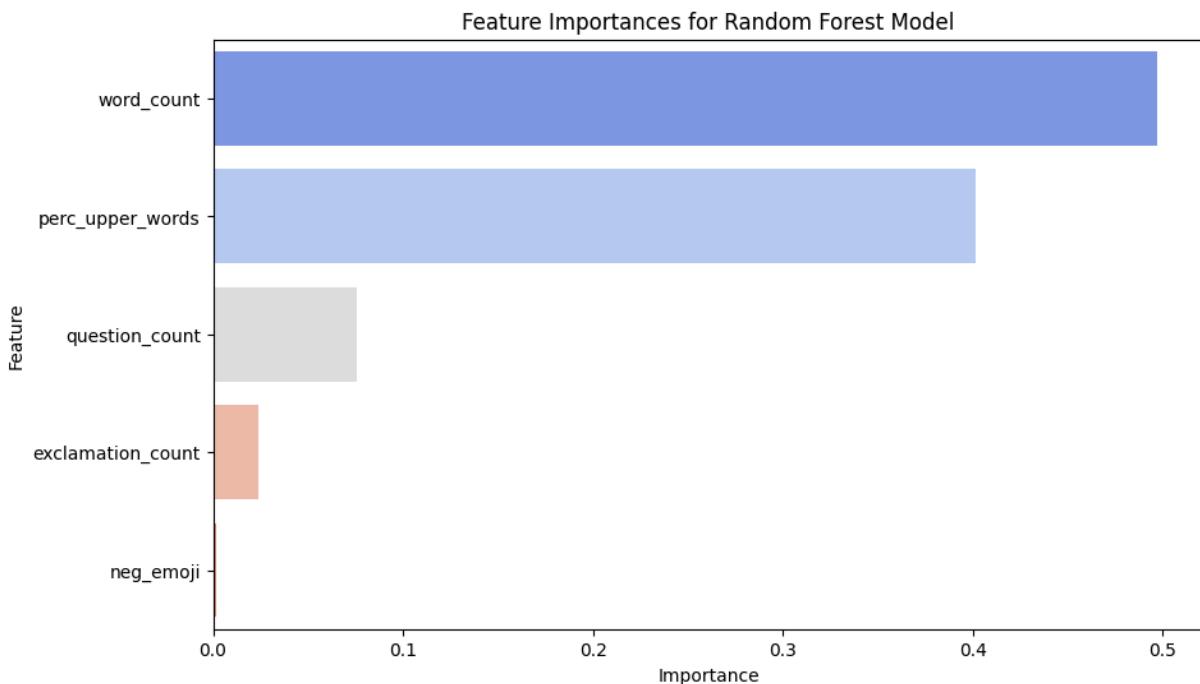
```
In [139]: # Access the best estimator (fitted RandomForestClassifier) from grid_search
best_rf_model = grid_search.best_estimator_.named_steps['clf']

# Get feature importances
importances = best_rf_model.feature_importances_

# Create a DataFrame for easier plotting
feature_importance_df = pd.DataFrame({
    'Feature': ['perc_upper_words', 'neg_emoji', 'exclamation_count', 'question_marks'],
    'Importance': importances
})

# Sort features by importance (optional)
feature_importance_df = feature_importance_df.sort_values(by='Importance', ascending=False)

# Plotting feature importances
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df, hue='Feature')
plt.title('Feature Importances for Random Forest Model')
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



Further Numeric Feature Evaluation - Correlation

Our added word_count feature was clearly a success compared to our other features! Let's compare these results with correlation coefficients between our numeric features and each class of our target variable, using Multinomial Logistic Regression.

```
In [140]: # Prepare X and y
X = filtered_df.drop(columns=['status', 'tokens'])
y = filtered_df['status']

# Initialize LabelEncoder
label_encoder = LabelEncoder()

# Fit LabelEncoder on y and transform it
y_encoded = label_encoder.fit_transform(y)

# Now 'y_encoded' contains numerical labels for your categories
# Print the mapping of original labels to encoded labels
label_mapping = dict(zip(label_encoder.classes_, label_encoder.transform(label)))
print("Label mapping:")
print(label_mapping)
```

Label mapping:

```
{'Anxiety': np.int64(0), 'Bipolar': np.int64(1), 'Depression': np.int64(2),
'Normal': np.int64(3), 'Suicidal': np.int64(4)}
```

```
In [142]: # Assign encoded target variable to y
y = y_encoded

# Split data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)
```

```
# Select numeric features
numeric_features = ['perc_upper_words', 'neg_emoji', 'exclamation_count', 'c
X_train_numeric = X_train[numeric_features]

# Convert y_train to a Pandas Series for correlation calculation
y_train_series = pd.Series(y_train).reset_index(drop=True)

# Calculate correlation coefficients with each category in y_train_series
correlation_coefficients = {}
for category in np.unique(y_train_series):
    category_mask = (y_train_series == category)
    if np.sum(category_mask) > 0: # Check if there are samples for this cat
        correlations = X_train_numeric.corrwith(pd.Series(y_train_series[cat
        correlation_coefficients[category] = correlations
    else:
        print(f"No samples found for category {category}. Skipping correlati

# Print correlation coefficients for each category
for category, correlations in correlation_coefficients.items():
    category_name = label_encoder.inverse_transform([category])[0] # Conver
    print(f"Correlation coefficients for category '{category_name}':")
    print(correlations)
    print("\n")
```

```
Correlation coefficients for category 'Anxiety':
```

```
perc_upper_words      NaN
neg_emoji              NaN
exclamation_count     NaN
question_count         NaN
word_count             NaN
dtype: float64
```

```
Correlation coefficients for category 'Bipolar':
```

```
perc_upper_words      NaN
neg_emoji              NaN
exclamation_count     NaN
question_count         NaN
word_count             NaN
dtype: float64
```

```
Correlation coefficients for category 'Depression':
```

```
perc_upper_words      NaN
neg_emoji              NaN
exclamation_count     NaN
question_count         NaN
word_count             NaN
dtype: float64
```

```
Correlation coefficients for category 'Normal':
```

```
perc_upper_words      NaN
neg_emoji              NaN
exclamation_count     NaN
question_count         NaN
word_count             NaN
dtype: float64
```

```
Correlation coefficients for category 'Suicidal':
```

```
perc_upper_words      NaN
neg_emoji              NaN
exclamation_count     NaN
question_count         NaN
word_count             NaN
dtype: float64
```

```
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2922: RuntimeWarning: invalid value encountered in divide
    c /= stddev[:, None]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2923: RuntimeWarning: invalid value encountered in divide
    c /= stddev[None, :]
```

```
In [143]: # Check for missing values in X_train_numeric
missing_values = X_train_numeric.isnull().sum()
```

```
print("Missing values in X_train_numeric:\n", missing_values)
```

```
Missing values in X_train_numeric:  
perc_upper_words      0  
neg_emoji              0  
exclamation_count     0  
question_count         0  
word_count             0  
dtype: int64
```

```
In [144]: # Check for constant columns in X_train_numeric  
constant_columns = X_train_numeric.columns[X_train_numeric.nunique() == 1]  
print("Constant columns in X_train_numeric:\n", constant_columns)
```

```
Constant columns in X_train_numeric:  
Index([], dtype='object')
```

```
In [145]: # Check variance of numeric features  
feature_variances = X_train_numeric.var()  
print("Variance of numeric features:\n", feature_variances)
```

```
Variance of numeric features:  
perc_upper_words      97.963622  
neg_emoji              0.001766  
exclamation_count     0.540314  
question_count         1.675379  
word_count             27657.184830  
dtype: float64
```

```
In [146]: # Check data types and alignment  
print("X_train_numeric shape:", X_train_numeric.shape)  
print("y_train shape:", y_train.shape)
```

```
X_train_numeric shape: (35040, 5)  
y_train shape: (35040,)
```

```
In [147]: # Convert y_train from numpy array to pandas Series  
y_train_series = pd.Series(y_train)  
  
# Now you can check the distribution of y_train categories  
print("Distribution of y_train categories:")  
print(y_train_series.value_counts())
```

```
Distribution of y_train categories:  
3    11057  
2     9344  
4     8766  
0     4152  
1     1721  
Name: count, dtype: int64
```

```
In [148]: # Tried Kendall and Spearman to try and debug, problem splitting y_train int  
# Trying Pearson again after troubleshooting without splitting y_train  
  
# Select numeric features  
numeric_features = ['perc_upper_words', 'neg_emoji', 'exclamation_count', 'c  
X_train_numeric = X_train[numeric_features]
```

```
# Convert y_train to a pandas Series
y_train_series = pd.Series(y_train, index=X_train.index)

# Calculate correlation coefficients with the target variable y_train
correlation_coefficients = X_train_numeric.corrwith(y_train_series)

# Print correlation coefficients
print("Correlation coefficients with y_train:\n")
print(correlation_coefficients)
```

Correlation coefficients with y_train:

```
perc_upper_words      0.038906
neg_emoji             -0.067907
exclamation_count    -0.061985
question_count        -0.131667
word_count            -0.056776
dtype: float64
```

In [149]:

```
# Tried Kendall and Spearman, problem splitting y_train into classes
# Trying Pearson again after troubleshooting
# Perform train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Initialize LabelEncoder for target variable
label_encoder = LabelEncoder()
y_train_encoded = label_encoder.fit_transform(y_train)

# Convert y_train_encoded back to a Series with original index
y_train_series = pd.Series(y_train_encoded, index=X_train.index)

# Select numeric features
numeric_features = ['perc_upper_words', 'neg_emoji', 'exclamation_count', 'category']
X_train_numeric = X_train[numeric_features]

# Category labels (e.g., 0, 1, 2, 3, 4)
categories = [0, 1, 2, 3, 4]

# Calculate and print Pearson correlation coefficients for each category
for category in categories:
    category_mask = (y_train_series == category)
    correlations_pearson = X_train_numeric[category_mask].corrwith(y_train_series)

    # Convert category to its name or label (if applicable)
    category_name = label_encoder.inverse_transform([category])[0] # Adjusted from [0] to [category]
    print(f"Pearson correlation coefficients for category '{category_name}':")
    print(correlations_pearson)
    print("\n")
```

```
Pearson correlation coefficients for category '0':
```

```
perc_upper_words      NaN  
neg_emoji             NaN  
exclamation_count    NaN  
question_count        NaN  
word_count            NaN  
dtype: float64
```

```
Pearson correlation coefficients for category '1':
```

```
perc_upper_words      NaN  
neg_emoji             NaN  
exclamation_count    NaN  
question_count        NaN  
word_count            NaN  
dtype: float64
```

```
Pearson correlation coefficients for category '2':
```

```
perc_upper_words      NaN  
neg_emoji             NaN  
exclamation_count    NaN  
question_count        NaN  
word_count            NaN  
dtype: float64
```

```
Pearson correlation coefficients for category '3':
```

```
perc_upper_words      NaN  
neg_emoji             NaN  
exclamation_count    NaN  
question_count        NaN  
word_count            NaN  
dtype: float64
```

```
Pearson correlation coefficients for category '4':
```

```
perc_upper_words      NaN  
neg_emoji             NaN  
exclamation_count    NaN  
question_count        NaN  
word_count            NaN  
dtype: float64
```

```
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2922: RuntimeWarning: invalid value encountered in divide
    c /= stddev[:, None]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2923: RuntimeWarning: invalid value encountered in divide
    c /= stddev[None, :]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2922: RuntimeWarning: invalid value encountered in divide
    c /= stddev[:, None]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2923: RuntimeWarning: invalid value encountered in divide
    c /= stddev[None, :]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2922: RuntimeWarning: invalid value encountered in divide
    c /= stddev[:, None]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2923: RuntimeWarning: invalid value encountered in divide
    c /= stddev[None, :]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2922: RuntimeWarning: invalid value encountered in divide
    c /= stddev[:, None]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2923: RuntimeWarning: invalid value encountered in divide
    c /= stddev[None, :]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2922: RuntimeWarning: invalid value encountered in divide
    c /= stddev[:, None]
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/numpy/lib/_function_base_impl.py:2923: RuntimeWarning: invalid value encountered in divide
    c /= stddev[None, :]
c /= stddev[None, :]
```

Oh well. Not helpful. We need to move forward rather than wasting more time.

Numeric Feature Evaluation - Multinomial Logistic Regression

Let's run logistic regression on our numeric features only, while we're still evaluating them. In case comparing these scores happens to be helpful to us later on.

In [150...]

```
# Prepare X and y
X = filtered_df.drop(columns=['status', 'tokens'])
y = filtered_df['status']
```

```
# Split data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

# Select numeric features
numeric_features = ['perc_upper_words', 'neg_emoji', 'exclamation_count', 'c']
X_train_numeric = X_train[numeric_features]
X_test_numeric = X_test[numeric_features]

# Initialize StandardScaler and fit on training data
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train_numeric)
X_test_scaled = scaler.transform(X_test_numeric)

# Define parameter grid for Logistic Regression
param_grid = {
    'clf__C': [0.1, 1.0, 10.0, 100.0], # Regularization parameter
    'clf__solver': ['lbfgs', 'saga'], # Multinomial Logistic Regression sol
    'clf__max_iter': [500, 1000, 1500], # Maximum number of iterations
}

# Setup the pipeline with Logistic Regression
pipeline = Pipeline([
    ('clf', LogisticRegression(class_weight='balanced', random_state=42)),
])

# Perform Grid Search CV with f1_weighted as scoring metric
grid_search = GridSearchCV(
    pipeline,
    param_grid,
    cv=5,
    scoring='f1_weighted', # Use f1_weighted for grid search scoring
    verbose=0,
    n_jobs=-1,
)

# Fit GridSearchCV
grid_search.fit(X_train_scaled, y_train) # Use X_train_scaled here

# Evaluate on the test set
test_score = grid_search.score(X_test_scaled, y_test) # Use X_test_scaled h
print(f"Test set score (F1 weighted): {test_score}")

# Print best parameters and best F1 weighted score
print("Best parameters:", grid_search.best_params_)
print("Best F1 weighted score:", grid_search.best_score_)

# Optionally, print classification report
y_pred = grid_search.predict(X_test_scaled)
print("\nClassification Report:")
print(classification_report(y_test, y_pred))
```

```
Test set score (F1 weighted): 0.4744715327476619
Best parameters: {'clf__C': 100.0, 'clf__max_iter': 500, 'clf__solver': 'lbfgs'}
Best F1 weighted score: 0.4749906955089175
```

Classification Report:

	precision	recall	f1-score	support
Anxiety	0.28	0.16	0.20	1758
Bipolar	0.18	0.43	0.25	780
Depression	0.43	0.28	0.34	3994
Normal	0.68	0.89	0.77	4867
Suicidal	0.43	0.38	0.40	3619
accuracy			0.50	15018
macro avg	0.40	0.43	0.39	15018
weighted avg	0.48	0.50	0.47	15018

In [151...]

```
# Printing stats by 'status' group
exccount_stats = filtered_df.groupby('status')[['exclamation_count']].agg(['mean', 'min', 'max'])
print("Exclamation Point Count by Status")
exccount_stats
```

Exclamation Point Count by Status

Out[151...]

mean min max

status			
Anxiety	0.252961	0	27
Bipolar	0.301479	0	9
Depression	0.106613	0	39
Normal	0.099347	0	20
Suicidal	0.120630	0	58

In [152...]

```
wordcount_stats = filtered_df.groupby('status')[['word_count']].agg(['mean', 'min', 'max'])
wordcount_stats
```

Out[152...]

mean min max

status			
Anxiety	135.389848	1	1673
Bipolar	184.722111	4	4850
Depression	153.843230	1	4245
Normal	18.015825	1	264
Suicidal	166.198950	2	6300

Text Feature Evaluation - Narrowing Down

Before we try to run anymore models again, we clearly need to cut down on the amount of text data in the form of tokens that we're running through our models as features.

We've spent hours and hours and hours trying to troubleshoot these models to get them to work prior to making any big cuts to our dataset...but it's clearly not working. It's time to get bold and vicious! Make some serious cuts just to see what happens. We're going to take the liberate to experiment. We can always go back and edit our code later on and undo what we did.

Multinomial Naive Bayes

Our only model that properly ran with our text data was Multinomial Naive Bayes. Let's run it again and use it cut out a ton of words that have very low probabilities of fitting into any of our 5 target categories.

First, let's see how many tokens we actually have. So we can get a sense of how many we should drop.

```
In [153]: # Step 1: Flatten the lists in 'tokens' column
flat_list = [item for sublist in filtered_df['tokens'] for item in sublist]

# Step 2: Count unique strings
unique_strings_count = len(set(flat_list))

print(f"Number of unique strings across the dataset: {unique_strings_count}")
```

Number of unique strings across the dataset: 74412

That is a freaking ton! No wonder our models are getting overwhelmed.

To jog our memories...here were our scores from the first time we ran Multinomial Naive Bayes as a baseline model.

Baseline Scores

Fitting 5 folds for each of 72 candidates, totalling 360 fits

Best parameters: {'clf__alpha': 0.1, 'clf__fit_prior': True, 'feature_selection__k': 'all', 'tfidf__max_features': 20000, 'tfidf__ngram_range': (1, 2)}

Best F1 score: 0.4304857142867046

Test set score: 0.46588604696801206

Shape of X_train_tfidf: (35668, 20000)

Classification Report:

support		precision	recall	f1-score	
	Anxiety	0.33	0.85	0.47	1053
	Bipolar	0.23	0.78	0.36	774
	Depression	0.46	0.80	0.58	4350
	Normal	0.94	0.38	0.54	4944
Personality disorder		0.82	0.27	0.41	270
	Stress	0.70	0.12	0.20	708
	Suicidal	0.88	0.03	0.06	3188
	accuracy			0.47	15287
	macro avg	0.62	0.46	0.37	15287
	weighted avg	0.70	0.47	0.42	15287

Class Imbalance - Class Prior

If we remember back to our original model, we had to account for class imbalance using `class_prior`. However, now that we've tuned our dataset, our class imbalances are different. We'll need to rerun those numbers.

```
In [154...]: # Define features and target
X = filtered_df['joined_tokens']
y = filtered_df['status'] # 'status' is the target variable

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

# Assuming `y_train` contains the target labels
class_distribution = y_train.value_counts(normalize=True)
class_distribution
```

```
Out[154...]: status
Normal      0.315554
Depression  0.266667
Suicidal   0.250171
Anxiety     0.118493
Bipolar     0.049115
Name: proportion, dtype: float64
```

Select Features

Now's the time to be bold! We'll set limits on our Vectorizer, as well as our SelectKBest.

```
In [155...]: # TF-IDF Vectorization
tfidf_vectorizer = TfidfVectorizer(max_features=20000) # Limit vocabulary size
X_train_tfidf = tfidf_vectorizer.fit_transform(X_train)
```

```
# Feature Selection using SelectKBest with chi2 score
k_best = 10000 # Select top 5000 features
selector = SelectKBest(score_func=chi2, k=k_best)
X_train_selected = selector.fit_transform(X_train_tfidf, y_train)
```

```
In [156...]: # Define class priors
class_priors = [0.315554, 0.266667, 0.250171, 0.118493, 0.049115]

# Define the parameter grid for GridSearchCV
param_grid = {
    'tfidf_ngram_range': [(1, 1), (1, 2)], # Use unigrams and/or bigrams
    'clf_alpha': [0.1, 0.5, 1.0], # Smoothing parameter for Naive Bayes
    'clf_fit_prior': [True] # Whether to learn class prior probabilities or not
}

# Create a pipeline that first transforms data using TF-IDF and then applies the classifier
pipeline = Pipeline([
    ('tfidf', TfidfVectorizer(max_features=20000)), # Placeholder for TF-IDF
    ('feature_selection', SelectKBest(score_func=chi2, k=1000)),
    ('clf', MultinomialNB(class_prior=class_priors)) # Naive Bayes with class priors
])

# Initialize GridSearchCV with timeout for joblib workers
timeout_in_seconds = 300 # Timeout set to 5 minutes (300 seconds)
with parallel_backend('loky'):
    grid_search = GridSearchCV(
        pipeline,
        param_grid,
        cv=5, # Number of folds for cross-validation
        scoring='f1_weighted', # F1 score, weighted for imbalanced classes
        verbose=1, # Print progress
        n_jobs=-1 # Use all available cores
    )

    # Fit GridSearchCV
    grid_search.fit(X_train, y_train)

    # After fitting, now you can access best parameters and scores
    print(f"Best parameters: {grid_search.best_params_}")
    print(f"Best F1 score: {grid_search.best_score_}")

    # Evaluate on the test set
    best_model = grid_search.best_estimator_
    test_score = best_model.score(X_test, y_test)
    print(f"Test set score: {test_score}")

    # Assuming best_model is your best estimator from GridSearchCV
    tfidf_vectorizer = best_model.named_steps['tfidf']
    X_train_tfidf = tfidf_vectorizer.transform(X_train)

    # Inspect the shape of the transformed feature matrix
    print(f"Shape of X_train_tfidf: {X_train_tfidf.shape}")

    # Generate predictions on the test set
    y_pred = best_model.predict(X_test)
```

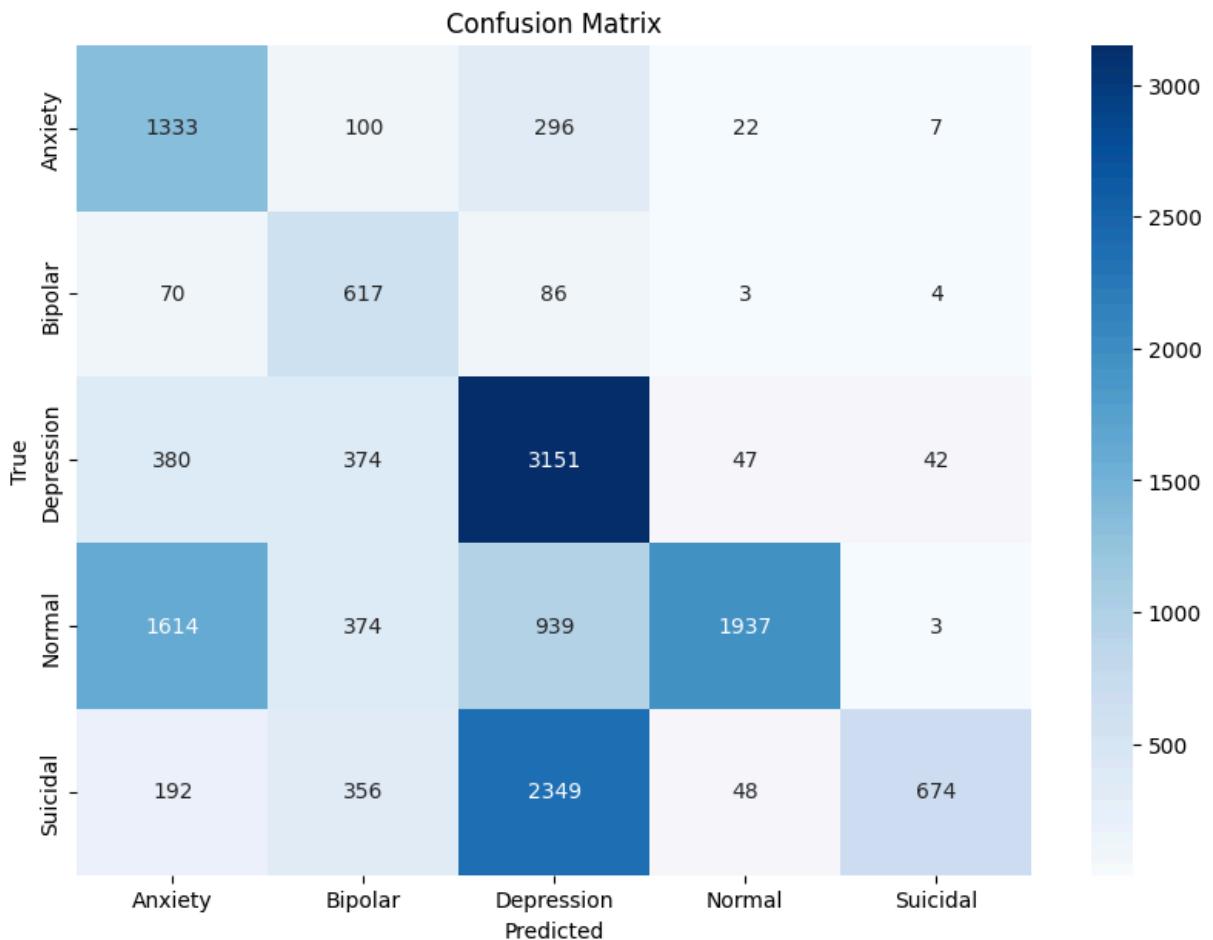
```
# Print classification report
report = classification_report(y_test, y_pred, target_names=best_model.classes_)
print("Classification Report:\n", report)

Fitting 5 folds for each of 6 candidates, totalling 30 fits
Best parameters: {'clf_alpha': 0.1, 'clf_fit_prior': True, 'tfidf_ngram_range': (1, 2)}
Best F1 score: 0.4932794230060484
Test set score: 0.5135171127979757
Shape of X_train_tfidf: (35040, 20000)
Classification Report:
precision    recall   f1-score   support
Anxiety       0.37      0.76      0.50      1758
Bipolar        0.34      0.79      0.47       780
Depression     0.46      0.79      0.58      3994
Normal         0.94      0.40      0.56      4867
Suicidal       0.92      0.19      0.31      3619
accuracy          0.51      0.51      0.51      15018
macro avg       0.61      0.58      0.49      15018
weighted avg    0.71      0.51      0.49      15018
```

```
In [157...]: # Optionally, print confusion matrix
cm = confusion_matrix(y_test, y_pred, labels=best_model.classes_)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=best_model.classes_, yticklabels=best_model.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')

# Save the plot with transparent background
plt.savefig('multi_naive_bays_confusion_matrix.png', transparent=True)

plt.show()
```



These scores are marginally better, but any improvement is hopeful!

This confusion matrix is super, super interesting...There are still a lot of 'Depression' cases being incorrectly predicted as 'Suicidal'.

There are also a lot of 'Anxiety' cases being incorrectly predicted as 'Normal', and quite a bit of 'Depression' cases being incorrectly predicted as normal as well.

Let's explore the overlap in tokens and bigrams among these categories.

Depression VS. Suicidal

```
In [158]: filtered_df.head()
```

Out [158...]

	status	perc_upper_words	neg_emoji	exclamation_count	question_count	tokens
0	Anxiety	0.0	0	0	0	[oh, go...]
1	Anxiety	0.0	0	0	0	[trou...]
2	Anxiety	0.0	0	0	0	[wro...]
3	Anxiety	0.0	0	0	0	[shift...]
4	Anxiety	0.0	0	0	1	[restle...]

In [159...]

```
# Function to tokenize and extract n-grams
def extract_ngrams(text, ngram_range=(1, 1)):
    vectorizer = CountVectorizer(ngram_range=ngram_range, analyzer='word')
    X = vectorizer.fit_transform(text)
    feature_names = vectorizer.get_feature_names_out()
    return feature_names

# Filter data to include only 'Depression' and 'Suicidal' statuses
dep_suic = filtered_df[filtered_df['status'].isin(['Depression', 'Suicidal'])]

# Extract unigrams (tokens) for Depression
tokens_depression_unigrams = extract_ngrams(dep_suic[dep_suic['status'] == 'Depression'])

# Extract bigrams for Depression
bigrams_depression = extract_ngrams(dep_suic[dep_suic['status'] == 'Depression'])

# Extract unigrams (tokens) for Suicidal
tokens_suicidal_unigrams = extract_ngrams(dep_suic[dep_suic['status'] == 'Suicidal'])

# Extract bigrams for Suicidal
bigrams_suicidal = extract_ngrams(dep_suic[dep_suic['status'] == 'Suicidal'])

# Find common tokens and bigrams
common_tokens = set(tokens_depression_unigrams).intersection(tokens_suicidal_unigrams)
common_bigrams = set(bigrams_depression).intersection(bigrams_suicidal)

# Print the number of common tokens and bigrams
print(f"Number of common tokens (unigrams) between 'Depression' and 'Suicidal':", len(common_tokens))
print("Common tokens (unigrams):", common_tokens)
```

```
print(f"Number of common bigrams between 'Depression' and 'Suicidal': {len(common_bigrams)}")
print("Common bigrams:", common_bigrams)
```

Number of common tokens (unigrams) between 'Depression' and 'Suicidal': 1570

4

Common tokens (unigrams): {'freeload', 'inconsistent', 'fooling', 'socialising', 'compulsively', 'crew', 'hunting', 'gibberish', 'mentality', 'dried', 'inept', 'paved', 'tatted', 'clingy', 'argumentative', 'possessed', 'mistakenly', 'agonizingly', 'trainer', 'shop', 'offer', 'operator', 'permanent', 'guilty', 'thoroughly', 'imploding', 'equipment', 'front', 'listing', 'secretive', 'destruction', 'poco', 'redecorate', 'obvious', 'phenomenon', 'piloting', 'dit', 'motorcycle', 'benefit', 'alls', 'soft', 'overthinking', 'participates', 'switched', 'silly', 'starter', 'custody', 'anything', 'drive', 'clicking', 'pharma', 'bothersome', 'cashier', 'smashing', 'goodnight', 'engagement', 'been', 'mewell', 'truck', 'unblocking', 'pressure', 'rotten', 'disapproval', 'eats', 'jacking', 'three', 'preferable', 'chicken', 'prozac', 'fighter', 'du', 'subside', 'blinded', 'knowand', 'painmy', 'agree', 'phpio', 'srsly', 'lurker', 'cam', 'mayora', 'flew', 'naltrexone', 'wither', 'asf', 'flirty', 'in', 'agreed', 'humanly', 'longterm', 'consists', 'controlled', 'themso', 'margin', 'bitter', 'usafter', 'trusting', 'volatile', 'linked in', 'carolina', 'shepherd', 'thingit', 'colitis', 'tying', 'embarrasses', 'chief', 'fcking', 'dj', 'paying', 'retail', 'excrutiating', 'toil', 'exploration', 'mouse', 'toture', 'widow', 'profitable', 'geography', 'vodka', 'negate', 'dictated', 'humiliated', 'demolishing', 'towni', 'individualistic', 'penis', 'unclear', 'administrator', 'tidy', 'ami', 'burry', 'destruct', 'association', 'harassing', 'despair', 'accept', 'lifeall', 'season', 'tipsy', 'trophy', 'impersonal', 'sens', 'largest', 'typed', 'nowbut', 'politician', 'marble', 'headjust', 'peach', 'talkative', 'immortality', 'psychiatristtherapist', 'wat', 'basicallyi', 'dobut', 'capitalistic', 'host', 'ultimately', 'engage', 'insulting', 'rocd', 'theory', 'overpass', 'phoned', 'tip', 'noticed', 'education', 'ugliness', 'safer', 'utilizing', 'layer', 'cruelty', 'iceberg', 'nightly', 'undiagnosed', 'coronavirus', 'uncommon', 'unchanged', 'planted', 'issueso', 'moping', 'evangelical', 'anythingthe', 'nearest', 'ponder', 'crossing', 'valueless', 'assembly', 'coca', 'ignored', 'ist', 'frighten', 'exhausting', 'stagger', 'anchored', 'significant', 'midst', 'misdiagnosed', 'cater', 'sweetened', 'familial', 'plastered', 'sunshine', 'kissed', 'communication', 'protest', 'deductible', 'stance', 'foster', 'selfhate', 'celebrating', 'progressed', 'save', 'thereby', 'dissociation', 'grope', 'everytime', 'billion', 'stagnate', 'dependency', 'snippet', 'blurting', 'solicitor', 'tapping', 'itnow', 'handler', 'helmet', 'muchbut', 'contradicting', 'msc', 'walkin', 'suburban', 'icing', 'mit', 'pup', 'extraordinary', 'rag', 'grouping', 'scenario', 'dvd', 'wheelchair', 'unfuck', 'achieve', 'mdd', 'succesful', 'reality', 'insomniac', 'detained', 'tug', 'molehill', 'pedestal', 'diarrhea', 'gross', 'gf', 'innocent', 'overlap', 'calmness', 'photographer', 'discussing', 'biochemistry', 'boil', 'bone', 'invading', 'developmenti', 'painfuli', 'reputation', 'immediately', 'respecting', 'whit', 'hey', 'mandated', 'secondary', 'relathionship', 'default', 'judgmentally', 'stock', 'lego', 'tandem', 'myselfsorry', 'pass', 'bullshiti', 'lefti', 'genetics', 'mindfulness', 'killin', 'port', 'bi', 'circling', 'dosen', 'ket', 'entrance', 'recognising', 'clinic', 'herculean', 'shard', 'intoxicated', 'sober', 'aid', 'cream', 'honesti', 'alignment', 'extinction', 'regained', 'nan', 'roleplayiing', 'disguise', 'horizontal', 'melast', 'companion', 'malaise', 'worship', 'symptom', 'precipice', 'individual', 'peel', 'occurring', 'inform', 'mehr', 'posture', 'hw', 'analog', 'island', 'wit', 'expense', 'milky', 'liveeveryone', 'mf', 'bloom', 'disappearance', 'stepping', 'deluding', 'unpopular', 'vaped', 'blind', 'blissful', 'upvotes', 'reminding', 'possiblebut', 'gentle', 'asian', 'housework', 'compensation', 'marking', 'meditative', 'problematic', 'reveal', 'laze', 'cheering', 'meteor', 'abuse', 'china', 'blackhole', 'sighti', 'mood', 'bombing', 'suicidalness', 'videogame', 'postponed', 'ange'}

ri', 'bud', 'dance', 'liver', 'childlike', 'shiti', 'sugary', 'mistakesi', 'profound', 'coaching', 'discard', 'unreachable', 'homeschooled', 'horseshi t', 'stupidass', 'incriminate', 'bombarded', 'refilled', 'approached', 'latc hing', 'englisch', 'nasal', 'abyss', 'volume', 'haywire', 'itthankfully', 's peak', 'boss', 'mebut', 'pc', 'worth', 'bestfriends', 'nothingness', 'happin essi', 'redeeming', 'beated', 'cotton', 'assulted', 'field', 'overbearing', 'awkwardness', 'rumour', 'plaster', 'ithow', 'traumai', 'xo', 'hahaha', 'met his', 'catfish', 'skyrocketed', 'visiting', 'meme', 'hpv', 'concealer', 'adj usts', 'casino', 'outcasted', 'responsibly', 'dessert', 'ded', 'todaywell', 'interface', 'facing', 'effortlessly', 'recently', 'shark', 'whiny', 'spil l', 'mortality', 'reguardless', 'def', 'based', 'gorgeous', 'lamp', 'literat ure', 'rehab', 'puedo', 'emailed', 'enclosed', 'landline', 'rd', 'musical', 'postcovid', 'talk', 'bit', 'esta', 'thrash', 'cheery', 'toilet', 'thoughtle ss', 'ware', 'cuticle', 'magnetic', 'backmy', 'fps', 'cricket', 'contentmen t', 'dwell', 'passerby', 'emasculate', 'epic', 'methese', 'faulty', 'dolla r', 'coworker', 'heading', 'martyr', 'panick', 'available', 'lifeampxbthe', 'eventually', 'cheapest', 'process', 'bum', 'exploitation', 'raw', 'fo', 'sc reamed', 'lostnothing', 'cosmic', 'learnt', 'frustrationi', 'speaks', 'nativ e', 'upis', 'conform', 'seizure', 'tboned', 'devaluing', 'hiding', 'lifestyl e', 'worsethat', 'skip', 'bonding', 'imploded', 'nonethe', 'appartment', 'emptied', 'exiled', 'gardening', 'normali', 'call', 'twig', 'canyon', 'dominat ion', 'cuz', 'fated', 'committed', 'rewind', 'backyard', 'underwater', 'suppl ement', 'trow', 'obituary', 'muster', 'mizzzidc', 'realize', 'caregiver', 'l ightheaded', 'moody', 'vous', 'pause', 'misused', 'brand', 'ghosted', 'laur a', 'vengeful', 'disintegrating', 'unwell', 'rem', 'oppressive', 'quitter', 'postsecondary', 'seeing', 'youshe', 'strenght', 'nail', 'embarrassingly', 'helping', 'posting', 'sure', 'lifeso', 'homosexuality', 'neatly', 'physiqu e', 'pillar', 'betteri', 'pressing', 'smallest', 'timeat', 'reflected', 'sch oolit', 'gaslighted', 'repugnant', 'meani', 'addictioni', 'caused', 'sewag e', 'noticeable', 'star', 'ranty', 'mental', 'carpal', 'essen', 'ehich', 'sa p', 'multivitamin', 'seized', 'oregon', 'godfather', 'flirt', 'picked', 'vap ing', 'regularlyi', 'kind', 'hahahaha', 'dishonest', 'occurred', 'pitty', 'rum', 'peoplelit', 'squashed', 'swan', 'tightened', 'pantry', 'purposely', 'ath letic', 'drifted', 'fertile', 'swapping', 'evidence', 'anythingafter', 'glo w', 'lifehe', 'whipping', 'mundane', 'beacon', 'crumble', 'frequent', 'forma lity', 'unreal', 'lab', 'wimp', 'realy', 'payoff', 'bereavement', 'passpor t', 'pharmacist', 'hydrated', 'miraculous', 'obeying', 'prize', 'misinterpre ted', 'pft', 'coldly', 'alittle', 'strangled', 'retake', 'iron', 'flat', 'ea ger', 'prom', 'tap', 'egocentric', 'standard', 'puked', 'consulting', 'beb ut', 'awe', 'teddy', 'conversing', 'capitalize', 'ting', 'betterthis', 'years all', 'perfect', 'pleasantly', 'feelingis', 'firefighter', 'assume', 'entrap ped', 'detach', 'itching', 'scope', 'suburb', 'whether', 'boost', 'smth', 'd ispatch', 'hectic', 'pmdd', 'silverware', 'trywhen', 'hyperbole', 'sense', 'supermarket', 'inevitability', 'subsidized', 'treading', 'coop', 'apologize', 'sadly', 'flare', 'thoughti', 'car', 'triple', 'lettuce', 'jean', 'oce a', 'compulsory', 'plagiarism', 'bitch', 'sport', 'daring', 'filler', 'weekd ay', 'blur', 'dancing', 'logging', 'reached', 'thinkin', 'algorithm', 'matri x', 'unconditionally', 'tit', 'se', 'versus', 'fabrication', 'depressionsuicide', 'contradictory', 'disgusted', 'pinning', 'pleasant', 'immediatelythi s', 'themand', 'transformed', 'zoey', 'rod', 'beforethe', 'esteemi', 'cpul d', 'indie', 'plea', 'harm', 'infatuated', 'bothering', 'earns', 'thro', 'terrifies', 'familyjust', 'bend', 'exists', 'injected', 'jd', 'creativity', 'massive', 'ahold', 'heart', 'depressive', 'hooking', 'incoherent', 'rped', 'har mi', 'woe', 'advice', 'preteen', 'shitless', 'viagra', 'hygiene', 'foreverbut', 'parlour', 'portray', 'cynical', 'diary', 'luck', 'start', 'stolen', 'elusive', 'melatonin', 'metric', 'resorting', 'caregiving', 'anybody', 'spe

d', 'enduring', 'ace', 'screenshots', 'diluted', 'predicament', 'primary', 'sucksmy', 'meaningless', 'expiration', 'airbnb', 'doe', 'ramen', 'convince', 'rounded', 'handcuff', 'hoe', 'miscommunication', 'yup', 'screwed', 'wayyy', 'giggle', 'despairing', 'join', 'goose', 'intervention', 'mediocrity', 'literally', 'schizo', 'sends', 'unworthiness', 'blackmailed', 'diewhat', 'weakyoure', 'freinds', 'capsule', 'lower', 'haunted', 'pm', 'inadequate', 'no ah', 'insect', 'snuck', 'whatsapp', 'swears', 'qualification', 'devolved', 'noninvasive', 'plumbing', 'mebeing', 'dummy', 'habit', 'perspective', 'precisely', 'dr', 'mathematics', 'resurfaced', 'pmi', 'gravitated', 'read', 'ein em', 'alcoholism', 'doomedi', 'boatload', 'minimumwage', 'spurred', 'pussy', 'repair', 'hypothyroidism', 'footprint', 'agony', 'occur', 'scarf', 'rainforest', 'install', 'center', 'baklava', 'rocky', 'reflex', 'thai', 'back', 'pes simist', 'pleasing', 'stats', 'amongst', 'fortunately', 'resurface', 'dust', 'proceeded', 'productivity', 'healing', 'ungodly', 'ish', 'sketch', 'spontaneous', 'believe', 'bagging', 'paracetamol', 'daughter', 'ew', 'spur', 'bonfire', 'oxygen', 'staffed', 'belonging', 'thorough', 'wtf', 'questionnaire', 'specific', 'supernatural', 'exist', 'puddle', 'dehumanizing', 'receding', 'gallery', 'comical', 'operate', 'brink', 'extreme', 'malnutrition', 'pinch', 'depot', 'search', 'vividly', 'hateful', 'immigrant', 'insignificant', 'unconsciously', 'streamer', 'assisted', 'ati', 'myselfwhy', 'wellmeaning', 'stil', 'zealous', 'shenanigan', 'reasoni', 'joyful', 'workday', 'cheated', 'bath', 'integrity', 'communicating', 'foolproof', 'peanut', 'intubation', 'hinge', 'bawling', 'filming', 'releasing', 'sertaline', 'snack', 'unloveable', 'mimicking', 'lawabiding', 'stressor', 'toby', 'atrophied', 'terminal', 'unbreakable', 'giggled', 'glaucoma', 'mainstream', 'particulary', 'experienced', 'reinforcement', 'arise', 'insurance', 'prioritized', 'bonus', 'pop', 'sorta', 'videocall', 'catastrophe', 'ample', 'hater', 'rectangle', 'settling', 'impediment', 'invite', 'fracturing', 'hand', 'chemistry', 'beginningi', 'unique', 'easei', 'judgemental', 'relaxer', 'skiing', 'paradox', 'roommate', 'lurked', 'herkill', 'loose', 'ser', 'unfolds', 'reflecting', 'toss', 'assusage', 'qualify', 'replacement', 'grandmother', 'weirdness', 'tad', 'cremate', 'friendless', 'uncovered', 'andni', 'gall', 'mindless', 'vm', 'helpwhen', 'childhoodteenagehood', 'intercourse', 'benifits', 'crysies', 'friendly', 'dogshit', 'manipulative', 'infallible', 'murderous', 'wealth', 'metabolism', 'french', 'windshield', 'affair', 'simultaneously', 'mon', 'hah', 'nj', 'truei', 'financial', 'mess', 'tape', 'facethe', 'jobno', 'lemonade', 'achining', 'itjust', 'heatwave', 'jason', 'trudge', 'coerce', 'enter', 'modicum', 'triedbut', 'booted', 'comparing', 'dipshit', 'surgery', 'condemned', 'infuriates', 'gummies', 'minimise', 'movemy', 'organize', 'overcame', 'ranked', 'laptop', 'bangladesh', 'housekeeping', 'desperately', 'viral', 'affective', 'felling', 'thishow', 'mewhat', 'amphetamine', 'chillin', 'downloaded', 'persistently', 'overweight', 'disposed', 'dissappear', 'heartless', 'shear', 'y k', 'sabotage', 'doubted', 'reliant', 'slim', 'newsometimes', 'awnsers', 'skating', 'mourn', 'ostomy', 'impressive', 'attorney', 'deadbeat', 'prideful', 'idid', 'surgeon', 'baggy', 'broken', 'clonazepam', 'remeber', 'wounded', 'defiant', 'todayi', 'burnout', 'piercings', 'sorysometimes', 'normally', 'sweeden', 'naive', 'contracted', 'paralysed', 'lige', 'dryer', 'download', 'howeveri', 'ugly', 'bragged', 'materialistic', 'heightened', 'amso', 'pilot', 'caked', 'flexing', 'thanked', 'worseits', 'judt', 'blister', 'scapegoat', 'stimulant', 'maybe', 'stocking', 'everday', 'stool', 'sleepover', 'compassionate', 'embrace', 'unsafe', 'ampxbso', 'pinpointed', 'friendfamily', 'taided', 'piecesi', 'controli', 'indulges', 'retaliate', 'revert', 'luckily', 'violence', 'birthdayi', 'raven', 'youtuber', 'divine', 'vocalist', 'ultrasound', 'percocets', 'obscure', 'aiming', 'pressed', 'refill', 'lead', 'temperature', 'mcat', 'foreseeable', 'nanobots', 'ampxbthe', 'dim', 'myasthenia', 'offend', 'problem', 'stationary', 'title', 'othersi', 'preschool', 'reputabl

e', 'overwhelmingly', 'humiliating', 'defined', 'entanglement', 'werid', 'eging', 'er', 'nightstand', 'jsut', 'imagining', 'mount', 'politics', 'selfpit y', 'garden', 'measure', 'tortured', 'visualize', 'scapegoating', 'atleast', 'sputtering', 'inexperience', 'analyse', 'havehad', 'simpson', 'flooded', 'ifs', 'awesome', 'speaki', 'inducing', 'havei', 'meditated', 'flicked', 'gu m', 'kiddos', 'spoke', 'vertebra', 'freelance', 'hoped', 'hadi', 'stockhol m', 'hilft', 'skinny', 'mantra', 'acknowledgement', 'atmosphere', 'diei', 'admitting', 'selfish', 'sleeve', 'frigging', 'selfharmits', 'debilitation', 'icky', 'various', 'vulgar', 'cure', 'tw', 'arduous', 'manifest', 'sedentary', 'smalltalk', 'terrorized', 'electrical', 'floated', 'tugging', 'abundance', 'rather', 'returning', 'violate', 'existing', 'ingenuity', 'maturity', 'offensive', 'survivable', 'ahora', 'toying', 'trickling', 'speakingi', 'comp', 'complimenting', 'overreacting', 'tendon', 'apparantly', 'midas', 'tasted', 'tend', 'hanging', 'rethinking', 'bare', 'happywhen', 'botch', 'alarme d', 'ignorantly', 'sanctuary', 'scoff', 'repressed', 'ride', 'organise', 'rpe', 'snot', 'australia', 'fogged', 'tbis', 'tic', 'metaphorical', 'bind', 'carers', 'refused', 'iran', 'studio', 'tale', 'wholeheartedly', 'overwhelmin g', 'satisfying', 'annoy', 'resting', 'relax', 'formula', 'youtubers', 'berating', 'scarcity', 'fairy', 'somewhere', 'unavailable', 'pegged', 'requirin g', 'viable', 'selfhelp', 'spread', 'steady', 'hen', 'sinister', 'multi', 'leetting', 'name', 'vocally', 'uselessi', 'respective', 'recommending', 'insisted', 'dwelling', 'bark', 'idea', 'deprecating', 'significantly', 'usi', 'replied', 'mustered', 'executioner', 'klonopin', 'appreciatedi', 'character', 'keto', 'drastically', 'rainbow', 'dissipate', 'ugandan', 'endi', 'superviso r', 'infuriating', 'suport', 'organised', 'privately', 'reallife', 'anything s', 'entirety', 'dare', 'meh', 'mortified', 'tylenol', 'classi', 'aboutwha t', 'energize', 'nevertheless', 'strangling', 'unpaid', 'preachy', 'splash', 'inferior', 'easyi', 'sweeping', 'aspies', 'basically', 'suddenly', 'conflic t', 'november', 'anticipated', 'juice', 'againwhy', 'stronger', 'whatnot', 'go', 'tier', 'withdrawal', 'awful', 'ee', 'quitting', 'switzerland', 'prestigious', 'offended', 'seeming', 'gun', 'humbly', 'steam', 'vast', 'referenc e', 'inability', 'unforgivable', 'monday', 'expanding', 'unrecognizable', 'bait', 'occupied', 'condo', 'locker', 'haunt', 'digital', 'fuckjng', 'came', 'drag', 'betray', 'elated', 'falleni', 'judgmental', 'kt', 'associated', 'enjoyi', 'pill', 'endless', 'indescribable', 'vehicle', 'volcano', 'asexual', 'hellhole', 'evaluated', 'community', 'latest', 'depressionso', 'mission', 'happythe', 'massacre', 'branch', 'blocked', 'intended', 'filipino', 'happiness', 'ampxbi', 'breach', 'infinity', 'vaccination', 'bos', 'node', 'greif', 'bulk', 'undergraduate', 'disadvantage', 'othersthat', 'lash', 'dresser', 'hyperacusis', 'chalk', 'conquered', 'dehumanized', 'invalidation', 'irreversible', 'loneliness', 'madi', 'morsel', 'officer', 'personality', 'shaving', 'underfunded', 'abilify', 'instance', 'lenient', 'acid', 'dissociated', 'unprofessional', 'inflicted', 'moving', 'substitute', 'ai', 'undeserving', 'statement', 'many', 'behaving', 'shattered', 'tramadol', 'likable', 'radius', 'hot', 'himmy', 'listener', 'dioxide', 'yadayadayada', 'question', 'serioumsy', 'memory', 'reminiscent', 'calender', 'andor', 'freed', 'kin', 'nowat', 'goat', 'teenager', 'skydiving', 'stormed', 'tothe', 'unfounded', 'achy', 'banning', 'weaker', 'institute', 'originality', 'austin', 'finding', 'hypochondriac', 'ball', 'intelligence', 'detail', 'dampd', 'lectured', 'validate', 'funniest', 'chnage', 'memo', 'quit', 'sash', 'forgetting', 'blob', 'tribe', 'enticing', 'trying', 'extra', 'rare', 'tightness', 'skipping', 'concussion', 'wantedit', 'damaging', 'chest', 'yeh', 'realization', 'decay', 'daythe', 'recreational', 'bot', 'fella', 'lot', 'therefor', 'collegeuniversity', 'shivering', 'unsurprisingly', 'unimportant', 'piecing', 'beatit', 'middlehigh', 'initiating', 'blanche', 'pokemon', 'sidewalk', 'includes', 'slowed', 'witnessed', 'thrown', 'want', 'wrestling', 'exceeds', 'stepmother', 'start

i', 'partum', 'againthey', 'disposing', 'drown', 'improvement', 'hunk', 'copping', 'painlessly', 'review', 'slay', 'spending', 'squad', 'su', 'weirded', 'danish', 'hardi', 'crystal', 'regress', 'clearest', 'awkwardly', 'destitute', 'twin', 'identity', 'mattersi', 'password', 'tyia', 'goi', 'traumatized', 'audition', 'brewing', 'forma', 'unconventional', 'circus', 'realism', 'von', 'conscious', 'canot', 'trimester', 'txts', 'colorless', 'internet', 'stumbled', 'paramedic', 'tipping', 'proved', 'pursuing', 'nervousness', 'et hnicity', 'condominium', 'reporting', 'besti', 'prepare', 'fooled', 'ke', 'relation', 'rowing', 'band', 'pedophile', 'withdrawls', 'ledge', 'polished', 'anymore', 'mold', 'leaf', 'via', 'guitar', 'grownass', 'immortal', 'smartest', 'tanked', 'nowhere', 'physician', 'vertically', 'cringe', 'bean', 'drag on', 'attached', 'shooter', 'dan', 'roomi', 'reverting', 'disowning', 'wron g', 'warmer', 'unfunny', 'hating', 'brew', 'employee', 'artwork', 'buisnes s', 'inflamed', 'lonley', 'passable', 'sucde', 'thrill', 'encouraging', 'symbolic', 'well', 'ashwaganda', 'blushing', 'provocation', 'accumulated', 'institutionalized', 'itright', 'problemsi', 'await', 'disordered', 'japan', 'eurer', 'sameit', 'positivei', 'upshe', 'obtained', 'guardian', 'acknowledge', 'cancerous', 'sext', 'golden', 'legislation', 'offends', 'subpar', 'suggests', 'torso', 'hardworking', 'misuse', 'clad', 'endso', 'clashing', 'faster', 'orientation', 'hardware', 'strain', 'assaulted', 'unaffected', 'oily', 'prioritizing', 'contemplates', 'bedridden', 'divorce', 'burnt', 'adoptive', 'citing', 'pizza', 'alternatively', 'manageable', 'daysit', 'impacting', 'floor why', 'ordinary', 'intact', 'flowed', 'met', 'plannedi', 'trim', 'grounded', 'melted', 'saidhi', 'bout', 'cancer', 'try', 'goodnow', 'constrained', 'evict', 'strand', 'suicidaldepressed', 'percieved', 'tummy', 'uncontrollably', 'unpacking', 'violating', 'fizzle', 'dependable', 'considered', 'unacceptin g', 'became', 'statewhen', 'swamp', 'swirl', 'numbing', 'accepting', 'min', 'senseso', 'half', 'anymore', 'act', 'cheerful', 'savior', 'force', 'defending', 'knob', 'avoidance', 'confidently', 'creation', 'breakfast', 'prev', 'stale', 'careless', 'healty', 'temporary', 'aspergers', 'minister', 'tangible', 'input', 'optimistic', 'stage', 'visitor', 'ground', 'firstly', 'chor e', 'anatomy', 'grandfather', 'menopause', 'clone', 'rotate', 'eventual', 'reaming', 'disposable', 'inflict', 'oki', 'problemmand', 'shitshow', 'smoke', 'prioritizes', 'terminology', 'confronts', 'unresponsive', 'defiance', 'saver', 'unhappyampxbthe', 'shyness', 'prescribe', 'continuing', 'fade', 'episodic', 'dyscalculia', 'stop', 'depends', 'beast', 'elseit', 'slump', 'overdue', 'biden', 'degrades', 'snapped', 'surrounding', 'bombard', 'airy', 'jobif', 'mumble', 'problemsnobody', 'creative', 'span', 'anesthesia', 'attempt', 'stimulus', 'unapologetic', 'criticizes', 'ranged', 'irl', 'plus', 'thisand', 'electronics', 'watchi', 'uselessthe', 'burner', 'eas', 'enticed', 'majoring', 'overwhelm', 'count', 'deactivate', 'foreveri', 'solution', 'could', 'completed', 'craziness', 'wronged', 'commit', 'spectrum', 'absolute', 'ليل', 'monthsi', 'needwant', 'load', 'fromthe', 'pepper', 'soda', 'cannotso', 'minimised', 'trapped', 'cryptocurrency', 'faithful', 'sorrow', 'usmy', 'yeari', 'holed', 'messedyup', 'unintelligent', 'kid', 'peek', 'network', 'compassion', 'slumber', 'machine', 'communicator', 'accurate', 'amends', 'memorial', 'destined', 'creeped', 'deception', 'scde', 'abort', 'chipped', 'lifting', 'supplier', 'fluctuation', 'integrating', 'taper', 'bcuz', 'triumph', 'conditon', 'info', 'practitioner', 'chosen', 'discontinued', 'haben', 'benadryl', 'responding', 'wailing', 'idiotic', 'history', 'resilience', 'jeff', 'thankfully', 'disheartened', 'itwhy', 'earbuds', 'needlessly', 'pump', 'delved', 'behavioral', 'lethal', 'offing', 'evokes', 'trump', 'carry', 'dedicate', 'entered', 'situationi', 'reflects', 'uh', 'stabbing', 'formed', 'istg', 'recognize', 'sleepless', 'brush', 'measurement', 'frequently', 'doubled', 'defeatist', 'owe', 'hapoy', 'criticize', 'leading', 'bore', 'dbt', 'gruesome', 'tail', 'compelling', 'scumbags', 'catchup', 'fraternity', 'anyways', 'c

e', 'study want', 'year found', 'manly', 'put dog', 'telling one', 'already talked', 'impure', 'posted', 'uncertainty future', 'waiting whenever', 'sent ir', 'month old', 'people take', 'anyone stay', 'constantly call', 'matter o thers', 'difficult also', 'date friend', 'equally worthless', 'shared feelin g', 'mom understand', 'much forever', 'thought guilt', 'lived without', 'sim ilar boat', 'lasted month', 'able save', 'late bloomer', 'empty stuck', 'dep ressed friend', 'relationship mental', 'disorder people', 'spending hundre d', 'freaking right', 'talking wa', 'stopped though', 'whenever write', 'ali ve everyone', 'life thisi', 'becomes', 'vastly', 'looking tried', 'plenty re ason', 'place new', 'fact extremely', 'treat depression', 'friendsto', 'take n point', 'prescribed lexapro', 'life truly', 'contact completely', 'lie los t', 'friend monday', 'bullshit need', 'come poor', 'lost right', 'without do ubt', 'doubt much', 'matter find', 'ever call', 'could held', 'good idea', 'give goal', 'tried feeling', 'decision right', 'happen day', 'unfortunate ly', 'fix would', 'happy continue', 'separate occasion', 'overtaken', 'one tu rn', 'happened taken', 'certain people', 'deserve around', 'saying thing', 'take leave', 'something completely', 'started sobbing', 'excitement', 'wa l east', 'fuck place', 'stayed week', 'nobody believe', 'relate connect', 'int eraction someone', 'people generally', 'fat ill', 'time ago', 'wanted want', 'someone love', 'someone hoping', 'fuck yeah', 'reason number', 'wanted liv e', 'feeling lethargic', 'fine never', 'ignored tried', 'sister home', 'some one calm', 'nothing improve', 'therapist time', 'fighting way', 'bin ich', 'year tell', 'constant pressure', 'thr', 'general know', 'get major', 'frien d city', 'gotten interview', 'stress right', 'told truth', 'sea weight', 'wa nt single', 'apartment today', 'increase', 'dining hall', 'know painful', 'p articipate self', 'thing pour', 'maybe good', 'started cut', 'completely wre cked', 'scared get', 'hurting almost', 'needed tried', 'pain mom', 'instead whole', 'pain alive', 'short getting', 'mask everyone', 'closely', 'help tru e', 'even license', 'suffocate feeling', 'hurt far', 'least accepted', 'stic k certain', 'rude people', 'year accomplish', 'feeling relatively', 'therapy taking', 'solution worked', 'format', 'home take', 'resist urge', 'glad mad e', 'go maybe', 'head alot', 'sleep live', 'one talk', 'fuck man', 'compared everything', 'okay going', 'going rough', 'issue everyone', 'today stop', 's tress whole', 'amount help', 'tunnel vision', 'somewhere really', 'sign thin g', 'college still', 'brain even', 'myselfit', 'girl voice', 'think nothin g', 'cycle wish', 'normal front', 'dayevery', 'send pic', 'first long', 'ram bling know', 'told suck', 'everyone behind', 'handling well', 'terrible mont h', 'spread positive', 'away health', 'hour break', 'working week', 'hope mo ve', 'long night', 'everyday', 'pay somehow', 'tried told', 'one type', 'pat hological', 'care treat', 'family sometimes', 'tired even', 'helping hand', 'meafter', 'either post', 'change much', 'could person', 'everyday tried', 'morning right', 'helped either', 'fault everyone', 'care title', 'place suc cessful', 'weird could', 'please remove', 'friend often', 'blame truly', 'mi litaristic rhetoric', 'cycle end', 'attack day'}

Too many to interpret! Let's set a threshold. And let's focus on only unigrams for now.

```
In [160]: # Function to extract most frequent ngrams (unigrams in this case)
def extract_top_ngrams(text, ngram_range=(1, 1), top_n=300):
    vectorizer = CountVectorizer(ngram_range=ngram_range, analyzer='word')
    X = vectorizer.fit_transform(text)
    feature_names = vectorizer.get_feature_names_out()
    frequencies = X.sum(axis=0).A1 # Sum frequencies of each ngram
    ngram_freq = dict(zip(feature_names, frequencies))
    sorted_ngrams = dict(Counter(ngram_freq).most_common(top_n)) # Get top
    return sorted_ngrams
```

```

# Filter data to include only 'Depression' and 'Suicidal' statuses
dep_suic = filtered_df[filtered_df['status'].isin(['Depression', 'Suicidal'])

# Extract top 100 unigrams for 'Depression'
tokens_depression = dep_suic[dep_suic['status'] == 'Depression']['joined_tokens']
top_depression_unigrams = extract_top_ngrams(tokens_depression, ngram_range=(1, 1))

# Extract top 100 unigrams for 'Suicidal'
tokens_suicidal = dep_suic[dep_suic['status'] == 'Suicidal']['joined_tokens']
top_suicidal_unigrams = extract_top_ngrams(tokens_suicidal, ngram_range=(1, 1))

# Find common top unigrams
common_top_unigrams = set(top_depression_unigrams.keys()).intersection(top_suicidal_unigrams)

# Print the number of common top unigrams
print(f"Number of common top unigrams between 'Depression' and 'Suicidal': {len(common_top_unigrams)}")
print("Common top unigrams:", common_top_unigrams)

```

Number of common top unigrams between 'Depression' and 'Suicidal': 269
 Common top unigrams: {'mind', 'depressed', 'matter', 'literally', 'people', 'though', 'hope', 'le', 'great', 'wa', 'finally', 'moment', 'issue', 'kid', 'home', 'come', 'guy', 'body', 'health', 'wanted', 'depression', 'therapy', 'friend', 'sure', 'day', 'parent', 'head', 'read', 'put', 'able', 'death', 'away', 'room', 'another', 'girlfriend', 'anymore', 'stuff', 'happen', 'tol', 'd', 'seem', 'maybe', 'go', 'make', 'back', 'better', 'pain', 'new', 'end', 'see', 'wrong', 'least', 'house', 'felt', 'gone', 'mental', 'high', 'norma', 'l', 'pretty', 'may', 'found', 'fact', 'bed', 'stuck', 'else', 'believe', 'do', 'ne', 'ever', 'kind', 'without', 'post', 'around', 'therapist', 'since', 'tod', 'ay', 'telling', 'hate', 'well', 'last', 'college', 'world', 'face', 'every', 'wish', 'ago', 'problem', 'sad', 'support', 'together', 'looking', 'real', 'nobody', 'please', 'care', 'live', 'made', 'first', 'came', 'okay', 'get', 'either', 'thought', 'love', 'best', 'right', 'also', 'thing', 'left', 'hel', 'l', 'man', 'job', 'way', 'mother', 'experience', 'doctor', 'soon', 'stay', 'went', 'happened', 'different', 'struggling', 'everything', 'guess', 'tim', 'e', 'good', 'couple', 'die', 'girl', 'took', 'point', 'honestly', 'making', 'work', 'night', 'due', 'money', 'anyone', 'would', 'going', 'try', 'takin', 'g', 'single', 'lost', 'give', 'worth', 'scared', 'sick', 'lonely', 'brain', 'deal', 'started', 'whole', 'kill', 'self', 'shit', 'never', 'many', 'clos', 'e', 'alone', 'already', 'think', 'one', 'place', 'family', 'really', 'wors', 'e', 'actually', 'sometimes', 'sorry', 'probably', 'recently', 'long', 'cry', 'longer', 'mean', 'break', 'thinking', 'others', 'something', 'living', 'two', 'hour', 'loved', 'week', 'enough', 'still', 'old', 'next', 'understand', 'remember', 'month', 'move', 'talk', 'saying', 'bit', 'someone', 'know', 'pa', 'rt', 'yet', 'say', 'always', 'alive', 'ask', 'trying', 'fuck', 'sleep', 'lit', 'tle', 'person', 'stop', 'relationship', 'feeling', 'anxiety', 'fucking', 'ad', 'vice', 'stupid', 'past', 'everyone', 'happy', 'take', 'lot', 'hurt', 'med', 'social', 'tried', 'dad', 'start', 'become', 'much', 'anything', 'tell', 'se', 'ems', 'want', 'keep', 'mom', 'help', 'completely', 'almost', 'got', 'let', 'said', 'change', 'look', 'ill', 'future', 'everyday', 'find', 'constantly', 'reason', 'tired', 'year', 'bad', 'used', 'call', 'even', 'dream', 'wake', 'leave', 'school', 'nothing', 'life', 'hard', 'getting', 'talking', 'could', 'idea', 'situation', 'working', 'worst', 'need', 'far'}

'Kill' and 'Die'

Let's reexamine if we can relabel some more rows. Let's search for rows in depression with bigrams ('kill', 'myself') and ('want', 'die')

```
In [161]: # Function to check if the text contains specific bigrams
def has_specific_bigrams(text, bigrams):
    for bigram in bigrams:
        if all(word in text.split() for word in bigram):
            return True
    return False

# Filter rows with status 'Depression'
depression_df = filtered_df[filtered_df['status'] == 'Depression']

# Define the bigrams to search for
target_bigrams = [('kill', 'myself'), ('want', 'die')]

# Find rows containing the specific bigrams
filtered_rows = depression_df[depression_df['joined_tokens'].apply(lambda x: has_specific_bigrams(x, target_bigrams))]

# Print the filtered rows
print("Rows with status 'Depression' containing bigrams ('kill', 'myself') or ('want', 'die'):")
filtered_rows.head(50)
```

Rows with status 'Depression' containing bigrams ('kill', 'myself') or ('want', 'die'):

Out[161...]

		status	perc_upper_words	neg_emoji	exclamation_count	question_count
7022	Depression		11.538462	0	2	0
7048	Depression		4.252874	0	1	1
7061	Depression		8.407080	0	0	2
7071	Depression		3.614458	0	0	0
7078	Depression		7.173913	0	0	0
7080	Depression		6.368330	0	0	0
7094	Depression		7.440476	0	0	0
7126	Depression		7.266436	0	0	1
7163	Depression		7.812500	0	0	0
7198	Depression		13.636364	0	0	0
7226	Depression		9.405941	0	0	0

		status	perc_upper_words	neg_emoji	exclamation_count	question_count
7261	Depression		5.078125	0	0	3
7282	Depression		5.577689	0	0	1
7298	Depression		12.621359	0	0	0
7311	Depression		1.923077	0	0	0
7318	Depression		7.973422	0	0	0
7319	Depression		5.913978	0	0	0
7323	Depression		6.417112	0	0	0
7328	Depression		20.000000	0	0	3
7377	Depression		10.204082	0	0	0
7381	Depression		7.836991	0	1	1
7522	Depression		5.154639	0	0	0

		status	perc_upper_words	neg_emoji	exclamation_count	question_count
7527	Depression		9.900990	0	0	0
7544	Depression		7.419355	0	1	1
7558	Depression		10.471204	0	0	2
7576	Depression		4.232804	0	0	0
7631	Depression		11.320755	0	0	2
7688	Depression		9.836066	0	0	0
7701	Depression		15.789474	0	0	0
7711	Depression		11.811024	0	1	0
7770	Depression		1.565558	0	0	5
7787	Depression		11.278195	0	0	0

		status	perc_upper_words	neg_emoji	exclamation_count	question_count
7793	Depression		4.535147	0	0	3
7816	Depression		8.080808	0	0	3
7843	Depression		2.952030	0	0	0
7856	Depression		5.555556	0	0	4
7883	Depression		9.565217	0	0	0
7891	Depression		7.734807	0	0	0
7920	Depression		11.949686	0	0	0
7937	Depression		6.197183	0	0	0
8138	Depression		11.538462	0	0	0
8154	Depression		4.651163	0	0	0
8215	Depression		8.787879	0	1	1

		status	perc_upper_words	neg_emoji	exclamation_count	question_count
8365	Depression		11.261261	0	0	0
8449	Depression		16.000000	0	0	0
8462	Depression		10.714286	0	0	0
8479	Depression		11.627907	0	0	0
8520	Depression		10.382514	0	0	0
8570	Depression		0.549451	0	0	7
8598	Depression		7.462687	0	0	0

In [162]: `filtered_rows.info()`

```
<class 'pandas.core.frame.DataFrame'>
Index: 788 entries, 7022 to 48059
Data columns (total 8 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   status            788 non-null    object  
 1   perc_upper_words  788 non-null    float64 
 2   neg_emoji          788 non-null    int64  
 3   exclamation_count 788 non-null    int64  
 4   question_count    788 non-null    int64  
 5   tokens             788 non-null    object  
 6   joined_tokens     788 non-null    object  
 7   word_count         788 non-null    int64  
dtypes: float64(1), int64(4), object(3)
memory usage: 55.4+ KB
```

```
In [163]: # Assuming 'filtered_df' has the indices you want to match with 'data'
filtered_indices = filtered_rows.index

# Print 'statement' values from data with the same index as filtered_rows
for idx in filtered_indices:
    statement = data.loc[idx, 'statement']
    print(f"Statement {idx}: {statement}")
```

Statement 7022: I cannot fucking feel a single fucking thing man . I bottle up every feeling and I am so far away from everyone in my life . I have no friends or anyone I can talk to , I feel like shit right this now , I do not want anything in life , I do not want to grow old , travel , make friends or whatever the fuck everyone says we should live for . I just want to end it , I wish I died in my sleep , everyday I wake up I feel shock and sadness . I did not choose to be born so why do I have to suffer for people who do not give a fuck about me . oh god !! just kill me nowMusic used to make me feel better (it was the only thing) , but now even music does nothing . I cannot tell anyone about how I feel , I feel stuck I cannot take this anymore . My parents think that I am very lazy and stupid so anything I tell them would be because I am m lazy . Like I told my mom n dad about how I do not find any thing interest while talking about career and they said it is because I am l azy and stupid that I think too highly of myself and that I am overconfident .They also said that I am useless and would not accomplish anything in life . when actually I have low self-esteem and zero self-confidence . how am I supposed to tell them about my depression that I have got because of neglect and abuse from them in childhood . Like I tell them that I do not feel like doing anything , then they would just say it is because I am lazy , you do not want to do this then you are fucking lazy and stupid "I just want to die in my sleep , that is the only thing I want . I feel so alone , I wish I had someone to support me emotionally The worst thing about all this is being alone

Statement 7048: 19 years old, male, from the Philippines. The lockdown here has been going on for 2 years and more now. It completely sucks here. Not only the government, but also the parenting style and community. My parents are giant narcissist, gaslighting bastards, apathetic, scapegoating, victim blaming, and abusive people. I grew up as a kid being parentified by my mom, while experiencing how my dad openly cheated on my mom while wasting the money we have. I did not feel like I grew with a father because he never tried to form an emotional connection with me. When he does, it will be him giving something while saying he worked his ass off for it so we should be grateful for it. He has forever glorified himself having a job, and he will use it in any arguments we had; "I am the one who has the job, who feeds you, and who gives you everything. Be fucking grateful!" in his words. I would be beaten up by him with a belt as a kid so it has left me traumatized by flinching at anything that is loud, surprising, or both. He has not changed from my childhood until now, except it has only gotten worse as time goes on. He has endlessly made me hate him with no possible resolutions. Calling me a failure, useless, good-for-nothing, and disgrace of a son does not help at all. He is the reason why I despise cheating, abusive, and manipulative people with a passion. His god complex shows to everyone else. It does not help that he is a cop so his temper is extremely bad. He has a fragile ego that will automatically collapse when told he is wrong or disagreed with. He also thinks he is doing our pets a favour by caging and leashing them their entire lives. he is truly an abusive piece of shit that does not deserve a family of his own. he is a manchild that accidentally had children.As for my mother, she is the epitome of gaslighting, victim blaming, and narcissism. You literally cannot talk to her without her taking it personally. Like my horrible father, she also thinks the world revolves around her. This is the person who made me go through parentification as a fucking 6 year old. Until now, I have been her downpour of traumas, pain, bad memories, and rants. I am literally exhausted from everything because of this. And I HAVE to listen to her or else everything turns bad; she manipulates my father to teach me a lesson. She also believes she cannot be wrong. She strongly believes feng shui, the supernatural, superstition, and the occult. She has made our entire lives revolve around s

e stories they both tell. Why is that not the case anymore? Or is it just me?

Again I ask, if the world wants me to live so much, than why make it so goddamn hard to actually live???

I am so fucking tired of living. Nothing here is worth it. I hate this fucking world. I hate every single person on this planet. I hate my stupid job. Hate my stupid boss. I hate how everything revolves around money. I hate the society humans created. I hate myself for not just being born a normal person.

Fck I wish someone would kill me. People die everyday, why will not someone just kill me?

Check out some excerpts from the statements above that are currently labeled as status 'Depression':

Statement 7071: After this pandemic is over and i can finally go out, i will overdose myself will sleeping pills. to be honest i do not want to die but I am just too tired to live.

Statement 7126:...I do not want to do anything. And those horrible thoughts of ending myself. I just want to die. If I could just leave this world and sleep forever. The only thing holding me above the surface are my kids...

Statement 7198: I want to die. I am about to mark my wrist with my nail help please I am from Lima peru Depresin

Let's label all of these rows as 'Depression' and see how our models scores change. We only had 42 'Depression' rows mislabeled as 'Suicidal'...compared to the 2349 'Suicidal' mislabeled as 'Depression'. So clearly the problem is in the current way our 'Depression' rows have been categorized and are therefore being tagged, and not vice versa.

```
In [164]: filtered_df['status'].value_counts(normalize=True)
```

```
Out[164]: status
Normal      0.318111
Depression  0.266451
Suicidal    0.247413
Anxiety     0.118063
Bipolar     0.049962
Name: proportion, dtype: float64
```

```
In [165]: # Assuming filtered_rows is a subset of filtered_df and has the same indices
          filtered_df.loc[filtered_rows.index, 'status'] = 'Suicidal'
```

```
In [166]: filtered_df['status'].value_counts(normalize=True)
```

```
Out[166... status
Normal      0.318111
Suicidal    0.263155
Depression   0.250709
Anxiety     0.118063
Bipolar      0.049962
Name: proportion, dtype: float64
```

Unigrams: 'Done', 'Leave', 'End'

Bigrams: ('want', 'live'), ('take',)

This only relabeled around 800 rows, when we had over 2,000 that were being mislabeled.

Let's examine a few more shared words that could be indicative of 'Suicidal' status rather than depression.

'done', 'leave', and 'end' can all implicate a persons wish to end their life, in addition to bigram ('want', 'live').

Bigrams with the first letter of ('take') could also help identify people wanting to take substances in order to overdose. Because this is our most unique identifier, we'll focus on this first.

```
In [167... # Filter rows with status 'Depression'
depression_df = filtered_df[filtered_df['status'] == 'Depression']

# Function to extract bigrams where the first word is 'take'
def extract_take_bigrams(text):
    vectorizer = CountVectorizer(ngram_range=(2, 2), analyzer='word')
    X = vectorizer.fit_transform(text)
    feature_names = vectorizer.get_feature_names_out()

    # Filter for bigrams where the first word is 'take'
    take_bigrams = [bigram for bigram in feature_names if bigram.startswith('take')]

    return take_bigrams

# Filter and print bigrams where the first word is 'take'
depression_statements = depression_df[depression_df['status'] == 'Depression']
take_bigrams = extract_take_bigrams(depression_statements)

print("Bigrams in 'Depression' statements where the first word is 'take':")
for idx, bigram in enumerate(take_bigrams, start=1):
    print(f"{idx}: {bigram}")
```

Bigrams in 'Depression' statements where the first word is 'take':

1: take abandon
2: take abilify
3: take ability
4: take able
5: take abroad
6: take absolute
7: take absolutely
8: take abuse
9: take abusive
10: take accept
11: take account
12: take accountability
13: take action
14: take actionmy
15: take active
16: take ad
17: take adderall
18: take admission
19: take advantage
20: take advantagei
21: take advice
22: take afford
23: take age
24: take agency
25: take ailment
26: take air
27: take alone
28: take alot
29: take already
30: take also
31: take always
32: take ambien
33: take amount
34: take amp
35: take anger
36: take annoyed
37: take another
38: take anti
39: take antidepressant
40: take anxiety
41: take anymore
42: take anymoreampxbi
43: take anymorefuck
44: take anymorei
45: take anymoremy
46: take anymorepeople
47: take anymorethe
48: take anyone
49: take anything
50: take anywhere
51: take ap
52: take appointment
53: take around
54: take aspirin
55: take associate

56: take ativan
57: take attention
58: take avoiding
59: take away
60: take awhile
61: take baby
62: take babysitter
63: take back
64: take backpack
65: take bad
66: take badly
67: take basically
68: take bath
69: take beating
70: take become
71: take bed
72: take believe
73: take benadryl
74: take benzos
75: take best
76: take beta
77: take better
78: take bicycle
79: take big
80: take bit
81: take bite
82: take blade
83: take blame
84: take blamed
85: take blow
86: take blue
87: take body
88: take bottle
89: take bp
90: take brain
91: take brake
92: take break
93: take breath
94: take breather
95: take broken
96: take bunch
97: take bus
98: take busit
99: take buspirone
100: take calling
101: take care
102: take career
103: take case
104: take cat
105: take cbt
106: take challenge
107: take challenging
108: take chance
109: take change
110: take charge
111: take chest

112: take chin
113: take chunk
114: take citalopram
115: take city
116: take class
117: take classesand
118: take cold
119: take college
120: take come
121: take comfort
122: take commission
123: take company
124: take complex
125: take concern
126: take concerta
127: take concrete
128: take confidence
129: take connect
130: take consistently
131: take control
132: take coping
133: take counseling
134: take counselling
135: take counter
136: take couple
137: take course
138: take court
139: take coz
140: take custody
141: take dad
142: take daily
143: take dark
144: take date
145: take dating
146: take day
147: take deal
148: take decision
149: take deep
150: take depression
151: take desk
152: take desvenlafaxine
153: take different
154: take dinner
155: take dirt
156: take disappointed
157: take disassociating
158: take discus
159: take distorts
160: take doctor
161: take dog
162: take dose
163: take double
164: take downhill
165: take drag
166: take drastic
167: take drive

168: take drug
169: take dull
170: take early
171: take ease
172: take east
173: take easy
174: take eat
175: take edge
176: take effect
177: take effort
178: take either
179: take embracing
180: take emergency
181: take emotion
182: take empty
183: take end
184: take endless
185: take enemy
186: take energy
187: take enjoyment
188: take enough
189: take entire
190: take entirely
191: take escape
192: take especially
193: take etc
194: take even
195: take ever
196: take every
197: take everyone
198: take everything
199: take everywhere
200: take exactly
201: take exam
202: take excruciating
203: take exercise
204: take exercising
205: take expect
206: take extra
207: take eye
208: take fair
209: take fake
210: take family
211: take far
212: take fck
213: take fcking
214: take feeling
215: take felt
216: take field
217: take final
218: take financially
219: take find
220: take finding
221: take first
222: take five
223: take fixed

224: take fluoxetine
225: take fmla
226: take foam
227: take food
228: take forever
229: take forward
230: take foster
231: take free
232: take friend
233: take friendship
234: take frustration
235: take fuck
236: take full
237: take gap
238: take garbage
239: take generally
240: take get
241: take given
242: take glass
243: take glue
244: take go
245: take god
246: take going
247: take good
248: take goodbye
249: take got
250: take grain
251: take granted
252: take great
253: take grief
254: take groundbreaking
255: take group
256: take guess
257: take gummies
258: take gummy
259: take hair
260: take half
261: take hand
262: take handincluding
263: take handle
264: take handling
265: take happens
266: take happy
267: take hard
268: take headset
269: take heartbeat
270: take heed
271: take hell
272: take help
273: take herbal
274: take high
275: take higher
276: take hit
277: take hobby
278: take hold
279: take holistic

280: take home
281: take honest
282: take honor
283: take hormone
284: take hospital
285: take hour
286: take house
287: take huge
288: take hurt
289: take hypersensitive
290: take image
291: take immediately
292: take immoralyou
293: take imodium
294: take important
295: take improvementdoes
296: take increased
297: take information
298: take initiative
299: take inside
300: take instead
301: take insult
302: take insurance
303: take interest
304: take internship
305: take interpret
306: take interview
307: take invalid
308: take investment
309: take issue
310: take italian
311: take italso
312: take iteither
313: take ithowever
314: take iti
315: take itnothing
316: take job
317: take joke
318: take jump
319: take keep
320: take kept
321: take kid
322: take kind
323: take klonopin
324: take knife
325: take knock
326: take knoe
327: take know
328: take large
329: take last
330: take latter
331: take le
332: take leap
333: take learn
334: take lease
335: take leash

336: take least
337: take leave
338: take lebothyroxine
339: take left
340: take letter
341: take level
342: take lexapro
343: take lie
344: take life
345: take lifebut
346: take lightly
347: take likely
348: take lithium
349: take little
350: take live
351: take living
352: take load
353: take loan
354: take loneliness
355: take lonely
356: take long
357: take longer
358: take look
359: take lost
360: take lot
361: take lotion
362: take love
363: take low
364: take lowest
365: take lyft
366: take macchiato
367: take made
368: take magnitude
369: take maintenance
370: take make
371: take many
372: take mask
373: take massive
374: take matter
375: take max
376: take mean
377: take meant
378: take med
379: take medication
380: take medicationspm
381: take medicine
382: take mefocused
383: take mei
384: take melatonin
385: take men
386: take mental
387: take mentalhealth
388: take message
389: take mg
390: take middle
391: take million

392: take min
393: take mind
394: take mine
395: take minute
396: take miracle
397: take mirtazapine
398: take miserable
399: take moderately
400: take mom
401: take moment
402: take money
403: take month
404: take morning
405: take mother
406: take motivation
407: take move
408: take much
409: take multiple
410: take name
411: take nap
412: take near
413: take nearly
414: take need
415: take needed
416: take negatively
417: take never
418: take new
419: take next
420: take nice
421: take night
422: take nose
423: take note
424: take notice
425: take noting
426: take nowhere
427: take nsaid
428: take ocd
429: take od
430: take offence
431: take offense
432: take offensive
433: take offer
434: take offi
435: take often
436: take oh
437: take one
438: take online
439: take opportunity
440: take order
441: take others
442: take otherwise
443: take outside
444: take overcome
445: take overdose
446: take overwhy
447: take pain

448: take panic
449: take parent
450: take park
451: take part
452: take past
453: take pathetic
454: take patience
455: take patient
456: take paxil
457: take pay
458: take paycut
459: take pc
460: take pediatrician
461: take people
462: take perc
463: take perfect
464: take personally
465: take pet
466: take pharmacy
467: take phone
468: take photo
469: take photography
470: take picture
471: take piece
472: take pill
473: take piss
474: take pity
475: take place
476: take plan
477: take please
478: take plus
479: take pm
480: take point
481: take possibility
482: take post
483: take praise
484: take prerequisite
485: take prescribed
486: take prescription
487: take pressure
488: take pride
489: take privately
490: take prob
491: take probably
492: take problem
493: take process
494: take professionally
495: take program
496: take project
497: take propranolol
498: take proscribed
499: take protect
500: take prozac
501: take psychiatrist
502: take ptsd
503: take public

504: take pulse
505: take punch
506: take push
507: take pushed
508: take put
509: take quite
510: take random
511: take rapid
512: take real
513: take reality
514: take really
515: take reasonable
516: take recently
517: take reddit
518: take refuse
519: take regardless
520: take regret
521: take regular
522: take rein
523: take rejection
524: take religion
525: take remembered
526: take research
527: take response
528: take responsibility
529: take responsibilty
530: take responsibity
531: take responsibly
532: take rest
533: take revenge
534: take reward
535: take rewrite
536: take right
537: take risk
538: take road
539: take rocket
540: take root
541: take sadness
542: take safety
543: take said
544: take scare
545: take scared
546: take scenic
547: take school
548: take scissors
549: take scoop
550: take seatbelt
551: take secluded
552: take second
553: take sedative
554: take see
555: take selfie
556: take semester
557: take sequence
558: take serious
559: take seriously

560: take seriouslyhow
561: take seriouslymy
562: take seroquel
563: take sertraline
564: take session
565: take severely
566: take shape
567: take share
568: take shelter
569: take shit
570: take shitload
571: take show
572: take shower
573: take sibling
574: take sick
575: take side
576: take sign
577: take since
578: take single
579: take sip
580: take sister
581: take sits
582: take sleep
583: take sleepaid
584: take sleeping
585: take slept
586: take slow
587: take slowly
588: take small
589: take smallest
590: take smoking
591: take sober
592: take software
593: take solace
594: take something
595: take somewhere
596: take sorry
597: take sort
598: take soul
599: take sound
600: take space
601: take specialist
602: take specific
603: take spend
604: take spent
605: take sporadically
606: take ssri
607: take stand
608: take starting
609: take state
610: take step
611: take still
612: take stimulant
613: take store
614: take stress
615: take strong

616: take stuck
617: take stupid
618: take stutter
619: take sublease
620: take substantial
621: take sudden
622: take suddenly
623: take suffer
624: take suggestion
625: take summer
626: take supplement
627: take supplementation
628: take survive
629: take tablet
630: take take
631: take taking
632: take talent
633: take teach
634: take telling
635: take test
636: take text
637: take thanks
638: take therapist
639: take therapy
640: take thereidk
641: take thing
642: take think
643: take thinking
644: take thought
645: take threw
646: take thyroid
647: take tia
648: take time
649: take tiniest
650: take tiny
651: take today
652: take together
653: take told
654: take toll
655: take tomorrow
656: take took
657: take top
658: take topamax
659: take totally
660: take train
661: take training
662: take trapped
663: take trash
664: take treatment
665: take tried
666: take trip
667: take truck
668: take trucker
669: take try
670: take trying
671: take tum

672: take turn
673: take turning
674: take twice
675: take two
676: take type
677: take understand
678: take ungrateful
679: take unit
680: take vacation
681: take valuable
682: take vast
683: take vet
684: take viibryd
685: take virginity
686: take vitamin
687: take vote
688: take wa
689: take waitlists
690: take walk
691: take walkbefore
692: take walking
693: take want
694: take way
695: take week
696: take wellbutrin
697: take whatever
698: take whenever
699: take wherever
700: take whilefuck
701: take whilst
702: take whole
703: take wild
704: take win
705: take winning
706: take wish
707: take withdrawal
708: take within
709: take wondering
710: take word
711: take work
712: take world
713: take worldweeeeelllllll
714: take worse
715: take worthless
716: take would
717: take wrong
718: take xanax
719: take year
720: take younger
721: take zero
722: take zolpidem

There are quite a lot of these that *could* be used in conversations of suicide. However, we don't have the time to comb through too many lines of data. We have to be sparing.

There are lots of prescription drugs that could be taken to overdose. We can't include every single prescription drug name, because that will be too many to look through. Furthermore, they're probably helpful indicators of mood disorders, and there will be other words that will help distinguish suicidal tendencies specifically.

We will however include any ambiguous references to taking a drug. In these cases, it's more likely someone is discussing using a drug for unhealthy purposes (ex. "I'm going to take a pill"), rather than discussing the use of prescription drugs to healthily treat a psychological ailment (ex. "I'm going to take an ambien"). We'll include these, along with any items that could be used as weapons for self harm, and vocabulary terms that implicate wanting to end a person's life (ex. "I can't take anymore").

Let's start with bigrams that are highly likely to be really clear indicators of suicidal tendencies. We have a few more selects that we can add to our list, depending on how useful our initial results are. We want to be meticulous, without wasting time!

In [168...]

```
# Function to check if the text contains specific bigrams
def has_specific_bigrams(text, bigrams):
    for bigram in bigrams:
        if all(word in text.split() for word in bigram):
            return True
    return False

# Filter rows with status 'Depression'
depression_df = filtered_df[filtered_df['status'] == 'Depression']

# Define the bigrams to search for
target_bigrams = [('take', 'anymore'), ('take', 'dose'), ('take', 'drug'),
                   ('take', 'overdose'), ('take', 'knife'), ('take', 'perc')]

# Find rows containing the specific bigrams
filtered_rows = depression_df[depression_df['joined_tokens'].apply(lambda x:
    any(bigram in x for bigram in target_bigrams))]

# Print the filtered rows
print("Rows with status 'Depression' containing 'take' bigram selects:")
# Assuming 'filtered_df' has the indices you want to match with 'data'
filtered_indices = filtered_rows.index

# Print 'statement' values from data with the same index as filtered_rows
for idx in filtered_indices:
    statement = data.loc[idx, 'statement']
    print(f"Statement {idx}: {statement}")
```

Rows with status 'Depression' containing 'take' bigram selects:

Statement 6993: Since I began seeing a therapist 5 months ago I have realised the thought of being happy genuinely terrifies me, and I do not even really know why. I think its probably caused by multiple things. If I had to guess, I think I just do not feel like I deserve to be happy. I have hated myself for so long that maybe I have found comfort in my misery, like its my punishment or something. And in a related way, I think I am scared of my happiness being selfish. Having been miserable for so long, I think Id feel like a selfish asshole if I suddenly got over my issues and starting liking life when I know there is countless people in the world still feeling how I did. And furthermore, its only in the last while have I really begun to take my problems seriously and have been trying to be more sympathetic to myself, but even now with all the shit I am going through its like I can still barely take care of myself. And its like I am afraid of losing the various anxiety issues and what not that make me miserable, because if I lose only them and not my self-hating mindset, Ill have even less reason to sympathise with myself. Ill be some privileged selfish asshole and be even harder on myself than I have ever been. Lastly, I think I am just afraid of the thought of living in general. Death just seems soo simple and peaceful, and I think I have become fixated on the idea of dying. Having been miserable for so many years I think the idea of any more life terrifies me. because even if I did make an amazing recovery, Ill still get new problems, which could be just as bad or worse, and I could also relapse and redevelop all my anxiety and insecurity issues. In fact I have before, in 2018 I made huge strides and then in 2019 it all turned to shit again. And a lot of it was because I had one panic attack in front of an entire class in school, in front of most of the new friends I'd made in 2018 who did not know I was an insecure, anxiety ridden loser with no friends before I met them. It pretty much shattered any progress I had made, and thereafter I had multiple panic attacks in school, making it even worse. Those are my interpretations anyway. Whenever I feel like I am making any progress on myself it sends waves of panic through me. I hate it so much, and I wish to fuck I knew what was causing it, because these are all good guesses but I still do not know if that is all there is to it. I am genuinely terrified of being happy and I do not even really know why.

Statement 7064: tw talk about eating disorder further downHi, sorry for long post and bad formatting. I do not know who to ask this as I am on a long waiting list for a psychologist and have no one else to talk to. I have always had problems with breakfast, mostly because I have trouble waking up and am so tired that the thought of food makes me nauseous. Before work (I am currently on sick leave) I used to force down a sandwich only because I know I faint easily. Work also helped me maintain a somewhat proper meal schedule as I had designated lunch and snack breaks, and I always had an appetite right after work. Now when I have been home several months, eating is a massive problem. Getting up earlier than 2pm is a huge but rare success but even if I do I just do not feel like eating for hours, despite feeling so hungry my stomach hurts and my head spins. I just do not crave anything and I try to drown the hunger with water and cigarettes and occasionally save myself from fainting by eating dextrose. This results in me lying around like a zombie for hours, not having the energy to do anything, until I feel too weak to make even a sandwich without having to sit down. Basically, during daytime, I HATE eating. I wish humans did not HAVE to eat and could survive off water. But then in the evening, I often find the cravings wash over me. Not only do I tend to eat unhealthy to begin with because I do not have the energy to cook (for example I often just have microwave food or frozen pizza), but at night while watching shows or playing games I feel some kind of NEED to have something. Sometimes cigarettes and tea works, but often the cravings are too str

go do whatever he wants and so will I, i am tired of being depressed, if I do not change I will k.o myself eventually. I love my children and him more than I hate myself to do that to them.

So I guess it is time to start learning how to live for myself for once.

Statement 48048: Our control, our fear, our anxiety, our stress... Why does it always have to lead to more of them? A plain and simple hi to begin

Most of this discussion will be off the top of my head and I will be adding my own experiences.

To start, there are situations in life that we genuinely have absolutely no control over. I myself suffer from MDD and struggle with it on a daily basis. I take medication, I try and get out more, I see a therapist and I am trying to find a job that would interest me. I am constantly worried about my future and whether I want to continue or not. One thing that happens quite frequently is that I create non-existent situations in my head which because even more stress and anxiety. Once this cycle begins, it does not matter what it is, whether I have control over it or not, it is a nightmare to deal with this constant stress of anything that may happen. Why does this occur? Why are we as humans so unequipped to handle these challenges? One could say the society we are surrounded by and what is expected of us causes a drastic change in the way we look, feel and think. To an extent this is true, but why is it so terribly exhausting and anxiety stricken to be able to even begin coping with what is going on around us. It is an obstacle I am incapable of clearing, I worry about anything and everything that involves me and those close to me. I am consistently tired, in pain, anxious, worried and have had this heavy weight on my chest each and every minute of everyday for over 5 years. Nothing seems to change, I keep telling myself tomorrow will be different, I will do this and that and it will all be better. But when the time comes.... I fall short, I let myself down, I let others down and most of all I continue the same cycle over and over again. Is this a situation or cycle I have made up to comfort myself as all I know is this feeling of dread and doubt and fear and anxiety, or does it truly exist and hinder me? Am I really my own worst enemy? Why am I constantly at war with myself? It feels like a never ending battle, I fight back, but I do not fight hard enough. Am I even capable of doing so?

Just what exactly is it that I am waiting for... An idea, a person, an event or situation in my life? These consist of things I can and cannot control... I feel as though I have no control over the majority of these things, which further pushes me to feel as though I have control over nothing. Why do we have to suffer the way we do? Why are our minds and body working against us? Is it not the job of our brain to keep us alive and well? It is supposed to keep us alive for as long as possible, and we even have built in ways of surviving dangerous situations, but when it comes down to the core of it, our brains are what control us. Our thoughts, our actions and feelings.... Yet it's perfectly happy with the idea of us ending it...

Too many, and too unclear. Let's narrow it further.

```
In [169]: # Define the bigrams to search for  
target_bigrams = [('take', 'soul'), ('take', 'life'), ('take', 'overdose'),  
  
# Find rows containing the specific bigrams  
filtered_rows = depression_df[depression_df['joined_tokens'].apply(lambda x:  
  
# Print the filtered rows
```

```
print("Rows with status 'Depression' containing 'take' bigram selects:")
# Assuming 'filtered_df' has the indices you want to match with 'data'
filtered_indices = filtered_rows.index

# Print 'statement' values from data with the same index as filtered_rows
for idx in filtered_indices:
    statement = data.loc[idx, 'statement']
    print(f"Statement {idx}: {statement}")
```

Rows with status 'Depression' containing 'take' bigram selects:

Statement 6993: Since I began seeing a therapist 5 months ago I have realised the thought of being happy genuinely terrifies me, and I do not even really know why. I think its probably caused by multiple things. If I had to guess, I think I just do not feel like I deserve to be happy. I have hated myself for so long that maybe I have found comfort in my misery, like its my punishment or something. And in a related way, I think I am scared of my happiness being selfish. Having been miserable for so long, I think Id feel like a selfish asshole if I suddenly got over my issues and starting liking life when I know there is countless people in the world still feeling how I did. And furthermore, its only in the last while have I really begun to take my problems seriously and have been trying to be more sympathetic to myself, but even now with all the shit I am going through its like I can still barely take care of myself. And its like I am afraid of losing the various anxiety issues and what not that make me miserable, because if I lose only them and not my self-hating mindset, Ill have even less reason to sympathise with myself. Ill be some privileged selfish asshole and be even harder on myself than I have ever been. Lastly, I think I am just afraid of the thought of living in general. Death just seems soo simple and peaceful, and I think I have become fixated on the idea of dying. Having been miserable for so many years I think the idea of any more life terrifies me. because even if I did make an amazing recovery, Ill still get new problems, which could be just as bad or worse, and I could also relapse and redevelop all my anxiety and insecurity issues. In fact I have before, in 2018 I made huge strides and then in 2019 it all turned to shit again. And a lot of it was because I had one panic attack in front of an entire class in school, in front of most of the new friends I'd made in 2018 who did not know I was an insecure, anxiety ridden loser with no friends before I met them. It pretty much shattered any progress I had made, and thereafter I had multiple panic attacks in school, making it even worse. Those are my interpretations anyway. Whenever I feel like I am making any progress on myself it sends waves of panic through me. I hate it so much, and I wish to fuck I knew what was causing it, because these are all good guesses but I still do not know if that is all there is to it. I am genuinely terrified of being happy and I do not even really know why.

Statement 7064: tw talk about eating disorder further downHi, sorry for long post and bad formatting. I do not know who to ask this as I am on a long waiting list for a psychologist and have no one else to talk to. I have always had problems with breakfast, mostly because I have trouble waking up and am so tired that the thought of food makes me nauseous. Before work (I am currently on sick leave) I used to force down a sandwich only because I know I faint easily. Work also helped me maintain a somewhat proper meal schedule as I had designated lunch and snack breaks, and I always had an appetite right after work. Now when I have been home several months, eating is a massive problem. Getting up earlier than 2pm is a huge but rare success but even if I do I just do not feel like eating for hours, despite feeling so hungry my stomach hurts and my head spins. I just do not crave anything and I try to drown the hunger with water and cigarettes and occasionally save myself from fainting by eating dextrose. This results in me lying around like a zombie for hours, not having the energy to do anything, until I feel too weak to make even a sandwich without having to sit down. Basically, during daytime, I HATE eating. I wish humans did not HAVE to eat and could survive off water. But then in the evening, I often find the cravings wash over me. Not only do I tend to eat unhealthy to begin with because I do not have the energy to cook (for example I often just have microwave food or frozen pizza), but at night while watching shows or playing games I feel some kind of NEED to have something. Sometimes cigarettes and tea works, but often the cravings are too str

ave now, I feel like my depression is more debilitating than ever. I am kind of scared that soon I will not be able to manage anything and I will let my life slip by. :(

Tldr; I probably average 1~2 hours of productivity (work and hobbies) a day, autopilot is not working since i have physical symptoms that make it hard to move and leave me dead tired, i feel like i am making excuses, and my depression feels inscreasingly debilitating.

Just wondering how you guys manage to get things done or if there is others that may be going through the same struggle.

Statement 47950: I am having such a hard timw with sleeping Like my sleeping habits have started to make my life feel worse. Usually I cannot sleep until 3am, most of the time it takes longer. I do like 1-2 all nighters a week. Wh enever I do one, i start to hear this whistling. A melody. Every morning I have an alarm and I automatically turn it off while sleeping. I cannot get a hold on my sleeping habits. Melatonin does not work

Statement 47958: It is not fcking fair It is not fair in any way. I only did what I had every right too he ruined my entire life and broke it I had every right on this stupid planet to tell him to fck off and just leave me be with his manipulation and bullshit. But not he had to go around crying to people how he has such a terrible daughter and is the victim in al, of this. Destroying the entire relationship I had with my grandmother to just end up in her telling me I cannot come visit anymore. I could not even tell her my side of the story I just broke down. It is not fair that he always goes around telling shit about everyone and how he is the fcking victim letting one suffer fr om all that by destroying everything one build. I hate him I hate him so muc h. I am happy to go infront of court and take away his rights about me. Let Us see how you lie your shit out if this I am sure the judges will decide quick when I tell them how many times he fcking hit me and caused most of my p roblems in my life.

Statement 48048: Our control, our fear, our anxiety, our stress... Why does it always have to lead to more of them? A plain and simple hi to begin

Most of this discussion will be off the top of my head and I will be adding my own experiences.

To start, there are situations in life that we genuinely have absolutely no control over. I myself suffer from MDD and struggle with it on a daily basi s. I take medication, I try and get out more, I see a therapist and I am trying to find a job that would interest me. I am constantly worried about my f uture and whether I want to continue or not. One thing that happens quite fr equently is that I create non existent situations in my head which because e ven more stress and anxiety. Once this cycle begins, it does not matter what it is, whether I have control over it or not, it is a nightmare to deal with this constant stress of anything that may happen. Why does this occur? Why a re we as humans so unequipped to handle these challange? One could say the s ociety we are surrounded by and what is expected of us causes a drastic chan ge in the way we look, feel and think. To an extent this is true, but why is it so terribly exhausting and anxiety stricken to be able to even begin copi ng with what is going on around us. It is an obstacle I am incapable of clea ring, I worry about anything and everything that involves me and those close to me. I am consistently tired, in pain, anxious, worried and have had this heavy weight on my chest each and every minute of everyday for over 5 years.

Nothing seems to change, I keep telling myself tomorrow will be different, I will do this and that and it will all be better. But when the time comes.... I fall short, I let myself down, I let others down and most of all I continue the same cycle over and over again. Is this a situation or cycle I have made up to comfort myself as all I know is this feeling of dread and doubt and fear and anxiety, or does it truly exist and hinder me? Am I really my own worst enemy? Why am I constantly at war with myself? It feels like a never ending battle, I fight back, but I do not fight hard enough. Am I even capable of doing so?

Just what exactly is it that I am waiting for... An idea, a person, an event or situation in my life? These consist of things I can and cannot control... I feel as though I have no control over the majority of these things, which further pushes me to feel as though I have control over nothing. Why do we have to suffer the way we do? Why are our minds and body working against us? Is it not the job of our brain to keep us alive and well? It is supposed to keep us alive for as long as possible, and we even have built in ways of surviving dangerous situations, but when it comes down to the core of it, our brains are what control us. Our thoughts, our actions and feelings.... Yet it's perfectly happy with the idea of us ending it...

Nope. Can't reasonably do this with enough certainty that we are correctly relabeling.

I have a feeling the same is going to be true for our unigrams 'done', 'end' and 'leave'. These words can be used in so many different contexts, other than suicidal speech.

Dropping Common Words

There were a lot of words in our list of frequent tokens shared by 'Depression' and 'Suicidal', that are incredibly common in everyday speech. Let's filter again through our dataframe and drop the top 50 tokens that occur most frequently across all status groups.

This will affect the usefulness of some of our important bigrams. But like we said - now's the time to be bold and experiment!

```
In [170]: # Combine all tokens into a single list
all_tokens = [token for token_list in filtered_df['tokens'] for token in token_list]

# Count frequencies of each token
token_counts = Counter(all_tokens)

# Get the top 50 most frequent tokens
top_50_tokens = token_counts.most_common(50)

# Print the top 50 tokens
for token, frequency in top_50_tokens:
    print(f'{token}: {frequency}'')
```

want: 27815
know: 24724
life: 21927
get: 21838
time: 19817
even: 17418
would: 17408
year: 15551
day: 15458
people: 15390
really: 15367
thing: 15131
one: 14465
going: 13571
go: 12852
think: 12748
friend: 12316
make: 11731
never: 11037
much: 11024
help: 10684
could: 10014
feeling: 9196
thought: 9156
anymore: 9096
work: 9030
back: 8842
take: 8528
way: 8500
anxiety: 8489
still: 8415
anything: 8376
depression: 8285
got: 8026
something: 7985
good: 7886
always: 7803
everything: 7733
need: 7706
better: 7421
anyone: 7420
see: 7264
every: 7255
nothing: 7047
month: 6840
right: 6727
someone: 6715
family: 6713
bad: 6684
hate: 6613

Some of these might still be useful, and our classes are imbalanced. We want to balance dropping common words that corrupt our data, with preserving words that are common simply because they're highly predictive of one of our unique majority classes.

Let's print the frequencies for these words across our unique status groups and compare them.

```
In [171]: from collections import defaultdict

# Initialize a defaultdict to store token frequencies by status
token_freq_by_status = defaultdict(Counter)

# Iterate over each row in filtered_df
for index, row in filtered_df.iterrows():
    status = row['status']
    tokens = row['tokens']

    # Count frequencies of tokens in this row
    token_counts = Counter(tokens)

    # Update token frequencies for this status group
    token_freq_by_status[status].update(token_counts)

# Combine token frequencies across all status groups
overall_token_counts = Counter()
for status, token_counts in token_freq_by_status.items():
    overall_token_counts.update(token_counts)

# Get the top 50 tokens sorted by frequency
top_50_tokens = overall_token_counts.most_common(50)

# Print frequencies for each status group
for status, token_counts in token_freq_by_status.items():
    print(f"Status: {status}")
    sorted_token_counts = sorted(token_counts.items(), key=lambda item: item[1])
    for token, frequency in top_50_tokens:
        print(f"\t{token}: {token_counts[token]}")
    print()
```

Status: Anxiety
want: 1569
know: 2898
life: 1322
get: 3101
time: 2958
even: 1837
would: 2069
year: 1971
day: 2426
people: 1118
really: 2160
thing: 1827
one: 1693
going: 1790
go: 1864
think: 1535
friend: 773
make: 1202
never: 1043
much: 1236
help: 1508
could: 1493
feeling: 1617
thought: 1105
anymore: 311
work: 1350
back: 1666
take: 944
way: 940
anxiety: 3952
still: 1166
anything: 923
depression: 282
got: 1380
something: 1565
good: 733
always: 977
everything: 688
need: 1030
better: 596
anyone: 1270
see: 834
every: 959
nothing: 614
month: 1342
right: 1141
someone: 572
family: 585
bad: 1079
hate: 356

Status: Normal
want: 1222
know: 795
life: 347

get: 917
time: 910
even: 475
would: 808
year: 533
day: 745
people: 644
really: 856
thing: 423
one: 791
going: 587
go: 797
think: 523
friend: 465
make: 529
never: 314
much: 425
help: 401
could: 358
feeling: 206
thought: 230
anymore: 102
work: 604
back: 492
take: 383
way: 290
anxiety: 107
still: 526
anything: 201
depression: 37
got: 496
something: 260
good: 671
always: 262
everything: 156
need: 455
better: 182
anyone: 233
see: 437
every: 218
nothing: 134
month: 303
right: 324
someone: 269
family: 189
bad: 241
hate: 153

Status: Depression
want: 8297
know: 8712
life: 7871
get: 7783
time: 6805
even: 6585
would: 5179

year: 5238
day: 5443
people: 5838
really: 5273
thing: 5734
one: 4806
going: 4261
go: 4169
think: 4133
friend: 4926
make: 4332
never: 3745
much: 3727
help: 3686
could: 3048
feeling: 3565
thought: 2459
anymore: 2892
work: 3256
back: 2946
take: 2481
way: 2887
anxiety: 2656
still: 2696
anything: 3283
depression: 5053
got: 2624
something: 2777
good: 2833
always: 2827
everything: 2798
need: 2623
better: 2746
anyone: 2486
see: 2449
every: 2554
nothing: 2568
month: 2087
right: 2138
someone: 2528
family: 2114
bad: 2357
hate: 2594

Status: Suicidal
want: 15375
know: 10546
life: 11400
get: 8294
time: 7509
even: 7553
would: 7977
year: 6445
day: 5559
people: 6901
really: 5714

thing: 5987
one: 6141
going: 5941
go: 5088
think: 5663
friend: 5605
make: 4864
never: 5314
much: 4848
help: 4303
could: 4319
feeling: 3054
thought: 4604
anymore: 5519
work: 2956
back: 3019
take: 3982
way: 3799
anxiety: 1278
still: 3390
anything: 3420
depression: 2168
got: 2890
something: 2789
good: 3060
always: 3297
everything: 3629
need: 2941
better: 3415
anyone: 2529
see: 3014
every: 3119
nothing: 3445
month: 2293
right: 2510
someone: 2978
family: 3511
bad: 2498
hate: 3283

Status: Bipolar
want: 1352
know: 1773
life: 987
get: 1743
time: 1635
even: 968
would: 1375
year: 1364
day: 1285
people: 889
really: 1364
thing: 1160
one: 1034
going: 992
go: 934

```
think: 894
friend: 547
make: 804
never: 621
much: 788
help: 786
could: 796
feeling: 754
thought: 758
anymore: 272
work: 864
back: 719
take: 738
way: 584
anxiety: 496
still: 637
anything: 549
depression: 745
got: 636
something: 594
good: 589
always: 440
everything: 462
need: 657
better: 482
anyone: 902
see: 530
every: 405
nothing: 286
month: 815
right: 614
someone: 368
family: 314
bad: 509
hate: 227
```

```
In [172]: filtered_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 50058 entries, 0 to 52840
Data columns (total 8 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   status            50058 non-null   object 
 1   perc_upper_words  50058 non-null   float64
 2   neg_emoji         50058 non-null   int64  
 3   exclamation_count 50058 non-null   int64  
 4   question_count    50058 non-null   int64  
 5   tokens             50058 non-null   object 
 6   joined_tokens     50058 non-null   object 
 7   word_count         50058 non-null   int64  
dtypes: float64(1), int64(4), object(3)
memory usage: 4.4+ MB
```

```
In [173...]: dropped_words = ['want', 'know', 'get', 'time', 'even', 'would', 'year', 'da...
          'thing', 'one', 'go', 'think', 'make', 'much', 'could', 'th...
# Define a function to filter out the dropped words
def filter_tokens(tokens):
    return [token for token in tokens if token not in dropped_words]

# Apply the function to each row in the 'tokens' column
filtered_df['tokens'] = filtered_df['tokens'].apply(filter_tokens)

# Optionally, print the updated DataFrame to verify changes
filtered_df.info
```

```
Out[173...]: <bound method DataFrame.info of
exclamation_count \
0      Anxiety      0.000000      0      0
1      Anxiety      0.000000      0      0
2      Anxiety      0.000000      0      0
3      Anxiety      0.000000      0      0
4      Anxiety      0.000000      0      0
...
52835  Anxiety      0.000000      0      0
52836  Anxiety      3.174603      0      0
52837  Anxiety      8.805031      0      1
52839  Anxiety      0.000000      0      0
52840  Anxiety      2.739726      0      0

question_count                                     tokens \
0                  0                         [oh, gosh]
1                  0           [trouble, sleeping, confused, mind, restless, ...
2                  0           [wrong, back, dear, forward, doubt, stay, rest...
3                  0           [shifted, focus, something, else, still, worried]
4                  1           [restless, restless, month, boy, mean]
...
52835             3           [anxiety, faintness, standing, title, anxiety, ...
52836             4           [anxiety, heart, symptom, anyone, else, someth...
52837             2           [travel, anxiety, hi, long, anxiety, sufferer, ...
52839             3           [fomo, involved, anyone, else, recently, watch...
52840             1           [getting, anything, done, anxiety, keep, house...

joined_tokens  word_count
0            oh gosh        3
1  trouble sleeping confused mind restless heart ...      10
2  wrong back dear forward doubt stay restless re...      14
3  shifted focus something else still worried        13
4  restless restless month boy mean        16
...
52835  anxiety faintness standing title anxiety faint...      23
52836  anxiety heart symptom anyone else something si...      64
52837  travel anxiety hi long time anxiety sufferer f...     161
52839  fomo thing involved anyone else get recently w...     180
52840  getting day get anything done anxiety keep hou...      76

[50058 rows x 8 columns]>
```

```
In [174... # Define a function to set empty lists to NaN
def set_empty_list_to_nan(tokens):
    return np.nan if not tokens else tokens

# Apply the function to each row in the 'tokens' column
filtered_df['tokens'] = filtered_df['tokens'].apply(set_empty_list_to_nan)

# Optionally, print the updated DataFrame to verify changes
print(filtered_df.head())

      status  perc_upper_words  neg_emoji  exclamation_count  question_count
\\
0  Anxiety          0.0           0            0                  0
1  Anxiety          0.0           0            0                  0
2  Anxiety          0.0           0            0                  0
3  Anxiety          0.0           0            0                  0
4  Anxiety          0.0           0            1                  1

                                tokens  \
0                           [oh, gosh]
1  [trouble, sleeping, confused, mind, restless, ...
2  [wrong, back, dear, forward, doubt, stay, rest...
3  [shifted, focus, something, else, still, worried]
4                [restless, restless, month, boy, mean]

      joined_tokens  word_count
0             oh gosh            3
1  trouble sleeping confused mind restless heart ...
2  wrong back dear forward doubt stay restless re...
3        shifted focus something else still worried
4            restless restless month boy mean            16
```

```
In [175... filtered_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 50058 entries, 0 to 52840
Data columns (total 8 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   status            50058 non-null   object 
 1   perc_upper_words  50058 non-null   float64
 2   neg_emoji          50058 non-null   int64  
 3   exclamation_count 50058 non-null   int64  
 4   question_count    50058 non-null   int64  
 5   tokens             49958 non-null   object 
 6   joined_tokens     50058 non-null   object 
 7   word_count         50058 non-null   int64  
dtypes: float64(1), int64(4), object(3)
memory usage: 4.4+ MB
```

```
In [176... # Drop rows where 'tokens' column is NaN
filtered_df = filtered_df.dropna(subset=['tokens'])
```

Dropping neg_emoji

If we remember back to our random forest numeric feature selection...neg_emoji had barely any influence on our target features. Since nothing came of trying to run the correlation coefficients...we forgot to drop it! Let's drop that too.

```
In [177]: # Drop the 'neg_emoji' column
filtered_df = filtered_df.drop(columns=['neg_emoji'])
```

```
In [178]: # Convert token lists to strings
filtered_df['joined_tokens'] = filtered_df['tokens'].apply(lambda x: ' '.join(x))
```

Anxiety VS Normal

If we remember back to our confusion matrix, there were also 1614 'Normal' cases incorrectly predicted as 'Anxiety'. Let's examine any overlap and/or mislabeling between those as well.

```
# Function to extract most frequent ngrams (unigrams in this case)
def extract_top_ngrams(text, ngram_range=(1, 1), top_n=300):
    vectorizer = CountVectorizer(ngram_range=ngram_range, analyzer='word')
    X = vectorizer.fit_transform(text)
    feature_names = vectorizer.get_feature_names_out()
    frequencies = X.sum(axis=0).A1 # Sum frequencies of each ngram
    ngram_freq = dict(zip(feature_names, frequencies))
    sorted_ngrams = dict(Counter(ngram_freq).most_common(top_n)) # Get top
    return sorted_ngrams

# Filter data to include only Normal and Anxiety statuses
norm_anx = filtered_df[filtered_df['status'].isin(['Anxiety', 'Normal'])]

# Extract top 100 unigrams for Anxiety
tokens_anxiety = norm_anx[norm_anx['status'] == 'Anxiety']['joined_tokens']
top_anxiety_unigrams = extract_top_ngrams(tokens_anxiety, ngram_range=(1, 1))

# Extract top 100 unigrams for Normal
tokens_normal = norm_anx[norm_anx['status'] == 'Normal']['joined_tokens']
top_normal_unigrams = extract_top_ngrams(tokens_normal, ngram_range=(1, 1))

# Find common top unigrams
common_top_unigrams = set(top_anxiety_unigrams.keys()).intersection(top_normal_unigrams)

# Print the number of common top unigrams
print(f"Number of common top unigrams between 'Anxiety' and 'Normal': {len(common_top_unigrams)}")
print("Common top unigrams:", common_top_unigrams)
```

Number of common top unigrams between 'Anxiety' and 'Normal': 202
Common top unigrams: {'mind', 'great', 'though', 'hope', 'le', 'big', 'finally', 'issue', 'check', 'home', 'come', 'guy', 'wanted', 'friend', 'sure', 'parent', 'read', 'head', 'put', 'able', 'away', 'room', 'another', 'use', 'anymore', 'coming', 'stuff', 'told', 'maybe', 'new', 'back', 'better', 'house', 'morning', 'end', 'see', 'wrong', 'least', 'felt', 'gone', 'high', 'pretty', 'may', 'bed', 'else', 'believe', 'crazy', 'kind', 'ever', 'done', 'with out', 'post', 'around', 'hi', 'since', 'today', 'hate', 'well', 'last', 'face', 'college', 'every', 'ago', 'problem', 'looking', 'real', 'please', 'care', 'made', 'live', 'first', 'came', 'okay', 'love', 'best', 'right', 'also', 'left', 'thanks', 'job', 'way', 'experience', 'eat', 'doctor', 'soon', 'went', 'different', 'happened', 'everything', 'guess', 'good', 'couple', 'took', 'point', 'making', 'night', 'work', 'money', 'phone', 'full', 'anyone', 'going', 'try', 'asked', 'taking', 'give', 'sick', 'food', 'started', 'hand', 'never', 'sound', 'car', 'many', 'cold', 'eye', 'already', 'place', 'family', 'actually', 'question', 'sometimes', 'later', 'yesterday', 'probability', 'long', 'cry', 'mean', 'thinking', 'found', 'something', 'two', 'hour', 'week', 'old', 'still', 'enough', 'next', 'understand', 'remember', 'month', 'move', 'talk', 'bit', 'someone', 'part', 'turn', 'yet', 'say', 'always', 'ask', 'trying', 'story', 'sleep', 'heart', 'little', 'person', 'stop', 'relationship', 'feeling', 'anxiety', 'minute', 'everyone', 'take', 'lot', 'hurt', 'start', 'tried', 'dad', 'course', 'ha', 'anything', 'tell', 'afraid', 'keep', 'mom', 'help', 'almost', 'got', 'let', 'said', 'change', 'look', 'small', 'find', 'reason', 'tired', 'bad', 'used', 'call', 'wake', 'leave', 'school', 'hard', 'life', 'nothing', 'getting', 'talking', 'idea', 'working', 'need', 'etc'}

```
In [180]: # Define the words to check for
words_to_check = {'crazy', 'anxiety', 'afraid'}

# Define a function to check if any of the words are in the tokens list
def contains_words(tokens):
    return any(word in tokens for word in words_to_check)

# Filter the DataFrame
filtered_rows = filtered_df[(filtered_df['status'] == 'Normal') & (filtered_
```

```
In [181]: # Assuming 'filtered_df' has the indices you want to match with 'data'
filtered_indices = filtered_rows.index

# Print 'statement' values from data with the same index as filtered_rows
for idx in filtered_indices:
    statement = data.loc[idx, 'statement']
    print(f"Statement {idx}: {statement}")
```

Statement 918: Crazy!! I just woke up at 9.30, I already know I am late for work, I can still tweet first hahaha

Statement 1020: want to talk but afraid of not being allowed

Statement 1042: if there is following this account, then I am afraid to make a connection

Statement 1045: body comparison that does not make sense, I AM CRAZY Screaming

Statement 1047: Possessive right, afraid it will not work, boss?

Statement 1060: Got TS's desire to be bangchan. But afraid to be comfortable.

Statement 1113: WANT TO SIN BUT Afraid to Laugh

Statement 1288: crazy fast wednesday already takbiran

Statement 1295: Want to lick your own spit but afraid to regret it

Statement 1315: Honestly, I am really crazy, suspend even laugh

Statement 1433: Hurry up for sahur. where is my water? It is crazy how fast I cook, I want fried rice from the eye egg villageâ€ .. it is fun to play with these requests hahahah I am not a wife yet

Statement 1460: I really want to look like other people, the ones with the twt notifications are accommodated a lot... But I am a crazy person who does not wear it.. I am afraid there will be mutual reps and then I do not know and end up getting peanuts.. so in the end I cannot collect much, that is the note: ")

Statement 1567: It is Mother's Day but I am in pain from noon to night. I massaged at the new place, after that, I went to buy stuff and I got dizzy and dizzy for more than 2 hours. Until the house is brave, do not cook to open it. Pstu Hbs Isha, Tuam Ais' husband at 4 places who is crazy sick. skit relief. But still sick

Statement 1576: I want to make a fa.. but I am afraid I will not take care of it "

Statement 1589: Let us go crazy together when mv dream is released, zimzalabim calls all nctzen wayzenni reveluv smstann let us have a good time! Unflop pleaseee help rt+like yaa tsysm

Statement 1657: Seriously crazy.. looking at SG, old friends or new friends are jealous and insecure.. AAAAAAAAAAAAAA

Statement 1712: {mu} anxiety, stress, selfharm affect your face or not? I am afraid that skincare is already expensive, why is my face suddenly?

Statement 1739: Aduu this Lebaran cannot get together again, this year we also do not gather. Oh my goodness, where are you afraid that the virus is getting fiercer, wehh

Statement 1754: I swear I am afraid of baper with Twitter kids, I am a little doubtful about that

Statement 1928: No need to show off about the existing reality. Later, when the harsh reality hits, you will know the feeling. I am afraid no one will help. Because you were arrogant in the beginning and without realizing it, you hurt people you did not know who they were.

Statement 2098: I am afraid that the sweet potato is not gelso again, wkkssh

Statement 2364: it is so quiet it is crazy

Statement 2380: Me when looking in the mirror: "crazy, my dog is really handsome" .. Me when trying to take a selfie: "ouch, what the heck is this huek (those ticks... x). Why is it still so dirty??"

Statement 2442: Crazy long hours because the roadblock is near SG Besi. Ok, find another way, do not go here

Statement 2469: afraid

Statement 2611: I am a crazy fan, are not I.. I usually get a cute scene in the drama, instead I like to scream and shout incoherently.. it does not work

Statement 41785: i m afraid i had bad code
 Statement 41857: sara kate i am afraid too you are reply about uni from age ago
 Statement 41992: i m very glad britney isn t crazy anymore that wa one hell of a show now i m sad that it s over next up ap at cook county
 Statement 42132: listening to murd and 9th wonder just chillen out missing m y crazy sex life
 Statement 42170: justgelo it sad knowing how they found her body what a crazy world we live in no child deserves that ya know
 Statement 42581: is afraid that her g i note will not read themselves
 Statement 42676: it just make me happy over and over again i wish i wan t afraid to fly http tinyurl com skpp
 Statement 42687: it just make me happy over and over again i just wish i was n t afraid to fly http tinyurl com skpp
 Statement 42713: horror wench me tooooooo i feel like i ve been on the verge of an anxiety attack all day
 Statement 42946: need s help with this anxiety crap
 Statement 43062: shannamoakler i m afraid to fly too i straight up refuse to get on a plane it suck because it really limit where you can go
 Statement 43090: all these crazy pollen tree in washington state are making my sinus go crazy
 Statement 43900: wow nw00 00 are now operated with a 0 we can get ac power cool but ticket price is hyper crazy expensive
 Statement 44591: davidt 00 virgin again today i m afraid a i m pushed for time having said that i wa delayed 0 min at brum
 Statement 44802: finding it incredibly unfair that mom got more tan than i even though i spent more time in the sun miss him like crazy
 Statement 45068: sneffielynn i wish i knew what s going on with tb it s driving me crazy
 Statement 45359: ugh doe anyone know what i can do to stop anxiety attack pl eeeease i need help

Nope! Too general. Let's just print statements with 'Anxiety'

In [182...]

```

# Define the words to check for
words_to_check = {'anxiety'}

# Define a function to check if any of the words are in the tokens list
def contains_words(tokens):
    return any(word in tokens for word in words_to_check)

# Filter the DataFrame
filtered_rows = filtered_df[(filtered_df['status'] == 'Normal') & (filtered_df['tokens'].apply(lambda x: contains_words(x)))] 

# Assuming 'filtered_df' has the indices you want to match with 'data'
filtered_indices = filtered_rows.index

# Print 'statement' values from data with the same index as filtered_rows
for idx in filtered_indices:
    statement = data.loc[idx, 'statement']
    print(f"Statement {idx}: {statement}")
  
```

For now, I am just going to keep saving up money for the future, taking precautions and doing things to protect my own mental health. I am seeing a therapist for my anxiety, and the next time I have a session I am going to bring up everything I talked about here and maybe get some help in dealing with it. Again, thanks to everyone who took time out of their day to give me advice when I was struggling, I really appreciate it.

Statement 30102: I asked my friend and Psychological guru if he happened to know of a washing machine I could pop my brain into for a while. He said yes and I have a neat little machine (expensive at \$800+) which literally rinses my brain, I run it at the highest charge possible to electrodes attached to each ear lobe and I sleep for an hour, waking up with a rinsed brain. Electric pulses and the use of magnetic fields in the treatment of psychiatric issues is not new but we no longer have to hook you up to the transfer station, in fact despite big pharma trying to stop the research (I kid you not), it is moved forward exponentially with researchers in the University of Texas in Dallas amongst many pushing their findings into the stimulation of the vagus nerve to alleviate the symptoms of PTSD, anxiety, depression and over all mental f*ckery.

Statement 30121: Secondly; after 3 years of countless meds and as many strains of dank as you can think of, I finally discovered my cure to anxiety. I have heard handfuls of people say music helps but that is never been that case for me. Until tonight. I started listening to Kid Cudi during an anxiety attack, I have been listening to him for a long time but never during an episode, I discovered his music helps me significantly. More than any meds I have ever tried.

Statement 30122: Anxiety my ***, that **** gave ME anxiety, being surrounded like that. * ​ Skipping to college, at this point I really did not care about anything anymore, and it did not help that I met a girl who completely split me and my friend apart (that is a long story in of itself), and started skipping a lot of my classes. I had good grades in my computer programming class, because I was actually curious for a while, but that was about it.

Statement 30133: – Lastly, it may be perceived to imply that having a positive mind frame will solve problems. As I understand it, this is not the message, only that there are methods to help overcome negative thinking of which this is one, but there are no guarantees. To be clear, this is not intended to be a 'cure', nor do I know if it applies to everyone, understanding is an important to recovery/reducing anxiety. and for me it helped. Beneath I will post what I discovered and how this has helped me.

Statement 30157: I have been suffering from anxiety and ptsd since I was 13, actively working towards recovery for the past three years. It is been up and down, of course, but overall I have made really encouraging progress. For the past six months I have been dating a really sweet guy, someone I think is worthy of my trust and good for me. We are both interested in being together long-term. This is the first healthy relationship I have ever been in, and I have been working really hard not to let past experiences and my own issues negatively affect our partnership.

Statement 42713: horror wench me toooooo i feel like i've been on the verge of an anxiety attack all day

Statement 42946: need s help with this anxiety crap

Statement 45359: ugh doe anyone know what i can do to stop anxiety attack pl eeeease i need help

In [183]: [filtered_rows.info\(\)](#)

```
<class 'pandas.core.frame.DataFrame'>
Index: 77 entries, 1712 to 45359
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   status            77 non-null      object  
 1   perc_upper_words  77 non-null      float64 
 2   exclamation_count 77 non-null      int64   
 3   question_count    77 non-null      int64   
 4   tokens             77 non-null      object  
 5   joined_tokens     77 non-null      object  
 6   word_count         77 non-null      int64  
dtypes: float64(1), int64(3), object(3)
memory usage: 4.8+ KB
```

There's not nearly enough for this to even be a useful relabel. Furthermore, it doesn't seem like all of these statement values definitely indicate a mood disorder, rather than a general use of the word 'anxiety' to casually state some sort of concern or stress.

As far as bold choices are considered, **we might consider dropping 'Normal' altogether**, since it's not useful for our chatbot agenda of giving mental health help to people with mood disorders, and it might be corrupting our model's success rates in a way that is currently unsolvable with the timeline of this project.

Let's run our Multinomial Naive Bayes model again and see how our scores and confusion matrix due now that we've dropped more common words, and cleaned some status labeling between Depression and Suicidal.

Multinomial Naive Bayes, again

```
In [184...]: # Define features and target
X = filtered_df['joined_tokens']
y = filtered_df['status'] # 'status' is the target variable

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

# Assuming `y_train` contains the target labels
class_distribution = y_train.value_counts(normalize=True)
class_distribution
```

```
Out[184...]: status
Normal      0.314298
Suicidal    0.265227
Depression  0.252045
Anxiety     0.118902
Bipolar     0.049528
Name: proportion, dtype: float64
```

In [185...]

```
# TF-IDF Vectorization
tfidf_vectorizer = TfidfVectorizer(max_features=20000) # Limit vocabulary size
X_train_tfidf = tfidf_vectorizer.fit_transform(X_train)

# Feature Selection using SelectKBest with chi2 score
k_best = 10000 # Select top 5000 features
selector = SelectKBest(score_func=chi2, k=k_best)
X_train_selected = selector.fit_transform(X_train_tfidf, y_train)

# Define class priors
class_priors = [0.314298, 0.265227, 0.252045, 0.118902, 0.049528]

# Define the parameter grid for GridSearchCV
param_grid = {
    'tfidf_ngram_range': [(1, 1), (1, 2)], # Use unigrams and/or bigrams
    'clf_alpha': [0.1, 0.5, 1.0], # Smoothing parameter for Naive Bayes
    'clf_fit_prior': [True] # Whether to learn class prior probabilities or not
}

# Create a pipeline that first transforms data using TF-IDF and then applies feature selection and classification
pipeline = Pipeline([
    ('tfidf', TfidfVectorizer(max_features=20000)), # Placeholder for TF-IDF
    ('feature_selection', SelectKBest(score_func=chi2, k=10000)),
    ('clf', MultinomialNB(class_prior=class_priors)) # Naive Bayes with class priors
])

# Initialize GridSearchCV with timeout for joblib workers
timeout_in_seconds = 300 # Timeout set to 5 minutes (300 seconds)
with parallel_backend('loky'):
    grid_search = GridSearchCV(
        pipeline,
        param_grid,
        cv=5, # Number of folds for cross-validation
        scoring='f1_weighted', # F1 score, weighted for imbalanced classes
        verbose=1, # Print progress
        n_jobs=-1 # Use all available cores
    )

# Fit GridSearchCV
grid_search.fit(X_train, y_train)

# After fitting, now you can access best parameters and scores
print(f"Best parameters: {grid_search.best_params_}")
print(f"Best F1 score: {grid_search.best_score_}")

# Evaluate on the test set
best_model = grid_search.best_estimator_
test_score = best_model.score(X_test, y_test)
print(f"Test set score: {test_score}")

# Assuming best_model is your best estimator from GridSearchCV
tfidf_vectorizer = best_model.named_steps['tfidf']
X_train_tfidf = tfidf_vectorizer.transform(X_train)

# Inspect the shape of the transformed feature matrix
```

```
print(f"Shape of X_train_tfidf: {X_train_tfidf.shape}")

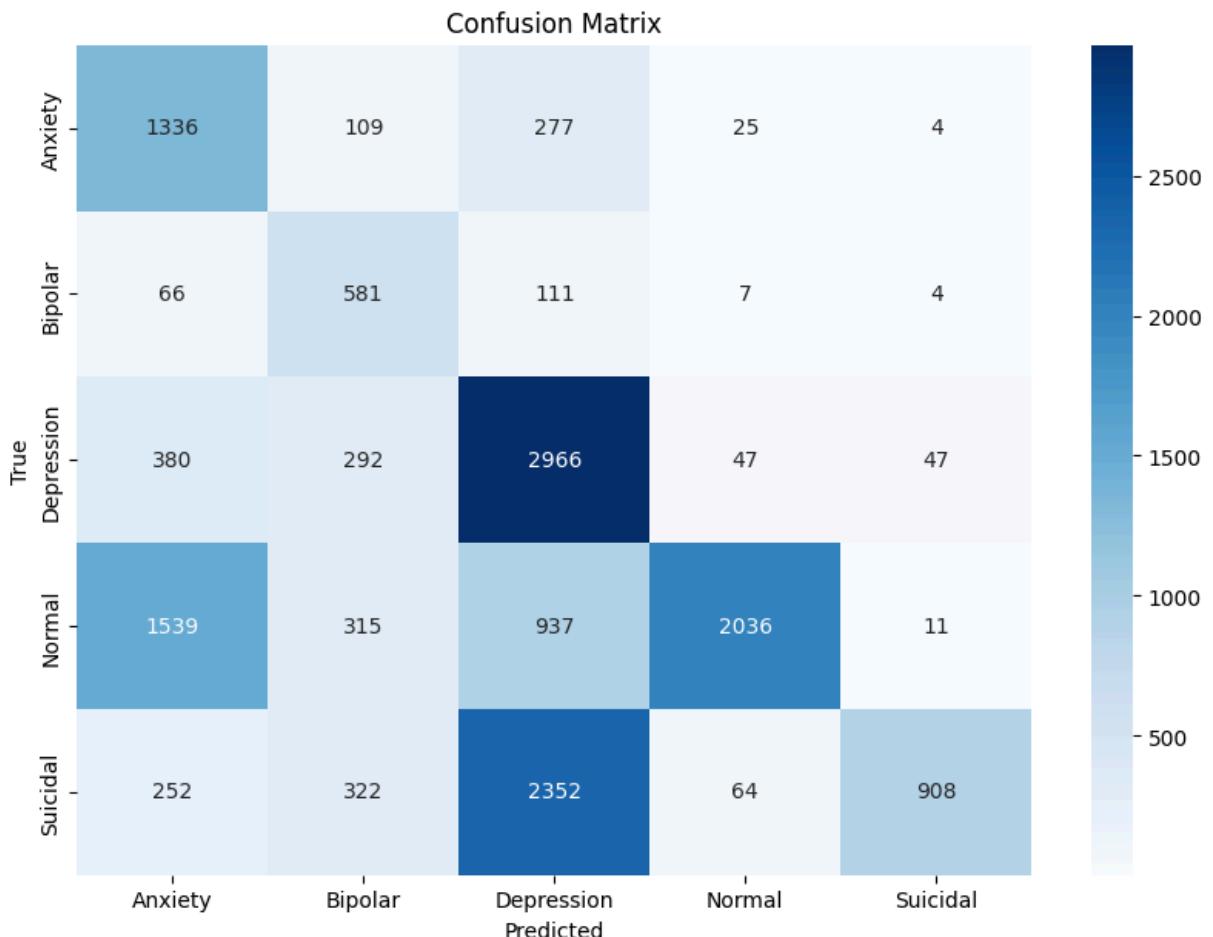
# Generate predictions on the test set
y_pred = best_model.predict(X_test)

# Print classification report
report = classification_report(y_test, y_pred, target_names=best_model.classes_)
print("Classification Report:\n", report)

# Optionally, print confusion matrix
cm = confusion_matrix(y_test, y_pred, labels=best_model.classes_)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=best_model.classes_, yticklabels=best_model.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()
```

Fitting 5 folds for each of 6 candidates, totalling 30 fits
Best parameters: {'clf_alpha': 0.1, 'clf_fit_prior': True, 'tfidf_ngram_range': (1, 2)}
Best F1 score: 0.5058909112949941
Test set score: 0.5222177742193755
Shape of X_train_tfidf: (34970, 20000)
Classification Report:

	precision	recall	f1-score	support
Anxiety	0.37	0.76	0.50	1751
Bipolar	0.36	0.76	0.49	769
Depression	0.45	0.79	0.57	3732
Normal	0.93	0.42	0.58	4838
Suicidal	0.93	0.23	0.37	3898
accuracy			0.52	14988
macro avg	0.61	0.59	0.50	14988
weighted avg	0.72	0.52	0.51	14988



Our F1-Scores improved slightly...but parts of our confusion matrix actually got worse!
How!?!? Why!!

Let's see what happens if we drop Normal altogether.

```
In [186]: filtered_df['status'].value_counts()
```

```
Out[186]: status
Normal      15829
Suicidal    13173
Depression  12546
Anxiety     5909
Bipolar     2501
Name: count, dtype: int64
```

Experiment: Dropping Normal

We'll test new numbers for SelectKBest and Max_Features now that we have a considerable amount fewer rows.

```
In [187...]: test_df = filtered_df

# Drop rows where 'status' is 'Normal'
test_df = test_df[test_df['status'] != 'Normal']
```

```
In [188...]: # Define features and target
X = test_df['joined_tokens']
y = test_df['status'] # 'status' is the target variable

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

# Assuming `y_train` contains the target labels
class_distribution = y_train.value_counts(normalize=True)
class_distribution
```

```
Out[188...]: status
Suicidal      0.384554
Depression    0.367560
Anxiety       0.176769
Bipolar        0.071118
Name: proportion, dtype: float64
```

```
In [189...]: # TF-IDF Vectorization
tfidf_vectorizer = TfidfVectorizer(max_features=20000) # Limit vocabulary size
X_train_tfidf = tfidf_vectorizer.fit_transform(X_train)

# Feature Selection using SelectKBest with chi2 score
k_best = 10000 # Select top 5000 features
selector = SelectKBest(score_func=chi2, k=k_best)
X_train_selected = selector.fit_transform(X_train_tfidf, y_train)

# Define class priors
class_priors = [0.384554, 0.367560, 0.176769, 0.071118]

param_grid = {
    'tfidf__ngram_range': [(1, 1), (1, 2)], # Use unigrams and/or bigrams
    'feature_selection__k': [2000, 5000], # Adjusted values for feature selection
    'clf__alpha': [0.1, 0.5, 1.0], # Smoothing parameter for Naive Bayes
    'clf__fit_prior': [True] # Whether to learn class prior probabilities or not
}

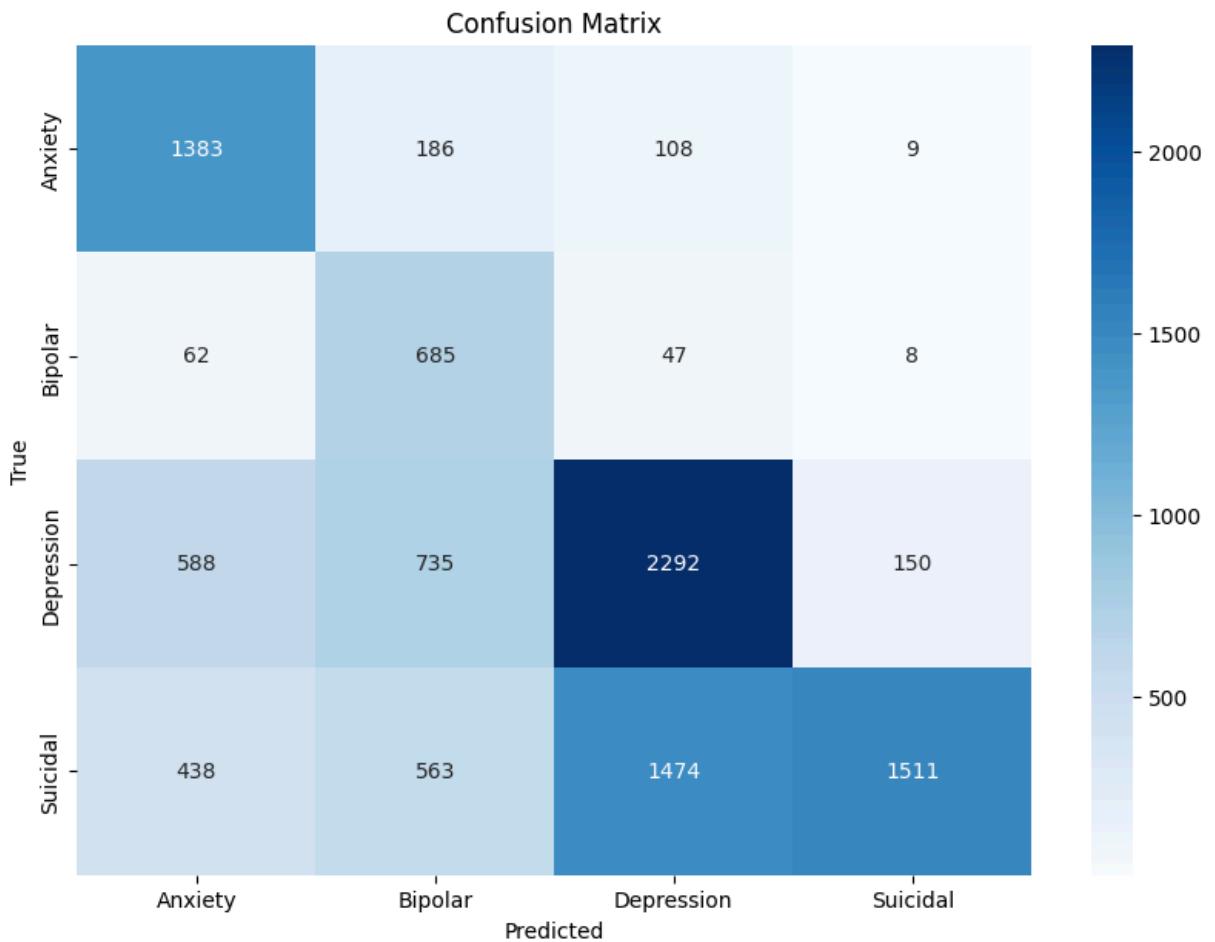
# Create a pipeline that first transforms data using TF-IDF and then applies the classifier
pipeline = Pipeline([
    ('tfidf', TfidfVectorizer(max_features=10000)), # Placeholder for TF-IDF
    ('feature_selection', SelectKBest(score_func=chi2, k=5000)),
    ('clf', MultinomialNB(class_prior=class_priors)) # Naive Bayes with class priors
])

# Initialize GridSearchCV with timeout for joblib workers
timeout_in_seconds = 300 # Timeout set to 5 minutes (300 seconds)
with parallel_backend('loky'):
```

```
grid_search = GridSearchCV(  
    pipeline,  
    param_grid,  
    cv=5, # Number of folds for cross-validation  
    scoring='f1_weighted', # F1 score, weighted for imbalanced classes  
    verbose=1, # Print progress  
    n_jobs=-1 # Use all available cores  
)  
  
# Fit GridSearchCV  
grid_search.fit(X_train, y_train)  
  
# After fitting, now you can access best parameters and scores  
print(f"Best parameters: {grid_search.best_params_}")  
print(f"Best F1 score: {grid_search.best_score_}")  
  
# Evaluate on the test set  
best_model = grid_search.best_estimator_  
test_score = best_model.score(X_test, y_test)  
print(f"Test set score: {test_score}")  
  
# Assuming best_model is your best estimator from GridSearchCV  
tfidf_vectorizer = best_model.named_steps['tfidf']  
X_train_tfidf = tfidf_vectorizer.transform(X_train)  
  
# Inspect the shape of the transformed feature matrix  
print(f"Shape of X_train_tfidf: {X_train_tfidf.shape}")  
  
# Generate predictions on the test set  
y_pred = best_model.predict(X_test)  
  
# Print classification report  
report = classification_report(y_test, y_pred, target_names=best_model.classes_)  
print("Classification Report:\n", report)  
  
# Optionally, print confusion matrix  
cm = confusion_matrix(y_test, y_pred, labels=best_model.classes_)  
plt.figure(figsize=(10, 7))  
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=best_model.classes_,  
            yticklabels=best_model.classes_)  
plt.xlabel('Predicted')  
plt.ylabel('True')  
plt.title('Confusion Matrix')  
plt.show()
```

Fitting 5 folds for each of 12 candidates, totalling 60 fits
 Best parameters: {'clf_alpha': 0.1, 'clf_fit_prior': True, 'feature_selection_k': 5000, 'tfidf_ngram_range': (1, 2)}
 Best F1 score: 0.5683273337747731
 Test set score: 0.5733958394374451
 Shape of X_train_tfidf: (23890, 10000)
 Classification Report:

	precision	recall	f1-score	support
Anxiety	0.56	0.82	0.67	1686
Bipolar	0.32	0.85	0.46	802
Depression	0.58	0.61	0.60	3765
Suicidal	0.90	0.38	0.53	3986
accuracy			0.57	10239
macro avg	0.59	0.67	0.56	10239
weighted avg	0.68	0.57	0.57	10239



Multinomial Logistic Regression

We tried running it, and it's still taking awhile. We're going to refrain from using SelectFromModel with Random Forest to select features...to try and cut down on runtime further. Especially now that we have honed some of our dataset.

```
In [190...]: X = filtered_df.drop(columns=['status', 'tokens'])

y = filtered_df['status']

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

encoder = LabelEncoder()
y_train_encoded = encoder.fit_transform(y_train)
y_test_encoded = encoder.transform(y_test)
y_train = pd.Series(y_train_encoded)
y_test = pd.Series(y_test_encoded)

# Check encoded labels
print("Encoded y_train:")
print(y_train.value_counts())

print("Encoded y_test:")
print(y_test.value_counts())

# Mapping from encoded labels to original labels
label_mapping = dict(zip(encoder.classes_, encoder.transform(encoder.classes)))
print("Label Mapping:")
print(label_mapping)
```

```
Encoded y_train:
3    10991
4     9275
2     8814
0     4158
1     1732
Name: count, dtype: int64
Encoded y_test:
3     4838
4     3898
2     3732
0     1751
1      769
Name: count, dtype: int64
Label Mapping:
{'Anxiety': np.int64(0), 'Bipolar': np.int64(1), 'Depression': np.int64(2),
'Normal': np.int64(3), 'Suicidal': np.int64(4)}
```

```
In [192...]: vectorizer = TfidfVectorizer(max_features=20000, ngram_range=(1, 2))

X_train_text = vectorizer.fit_transform(X_train['joined_tokens'])
X_test_text = vectorizer.transform(X_test['joined_tokens'])

# Scale numeric data
# Initialize StandardScaler
scaler = StandardScaler()

# Fit and transform StandardScaler on numeric data
# ADDING NEW WORD COUNT FEATURE
X_train_numeric_scaled = scaler.fit_transform(X_train[['perc_upper_words', 'exclam']])
X_test_numeric_scaled = scaler.transform(X_test[['perc_upper_words', 'exclam']])
```

```
X_train_combined = hstack([X_train_text, X_train_numeric_scaled])
X_test_combined = hstack([X_test_text, X_test_numeric_scaled])
```

```
In [193]: param_grid = {
    'clf_C': [0.1, 1.0, 10.0],
    'clf_solver': ['lbfgs', 'saga'],
    'clf_max_iter': [2000], # Balance computational cost with convergence needs
}

# Logistic Regression
pipeline = Pipeline([
    ('clf', LogisticRegression(class_weight='balanced', random_state=42)), #
])

# Initialize GridSearchCV
grid_search = GridSearchCV(
    pipeline,
    param_grid,
    cv=5,
    scoring='f1_weighted',
    verbose=0,
    n_jobs=-1,
)

# Fit GridSearchCV
grid_search.fit(X_train_combined, y_train)

# Evaluate on the test set
test_score = grid_search.score(X_test_combined, y_test)
print(f"Test set score: {test_score}")

print("Best parameters:", grid_search.best_params_)
print("Best F1 score:", grid_search.best_score_)
```

```
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linear_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which means the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linear_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which means the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linear_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which means the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linear_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which means the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linear_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which means the coef_ did not converge
    warnings.warn(
```

```
Test set score: 0.7830027185650968
Best parameters: {'clf__C': 10.0, 'clf__max_iter': 2000, 'clf__solver': 'sag
a'}
Best F1 score: 0.7783384727718577
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linea
r_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which mean
s the coef_ did not converge
warnings.warn(
```

Random Forest

Let's try building a Random Forest model, which is known to handle non-linear complex relationships in large datasets quite well. Yet is less computationally expensive than a model like Gradient Boosting.

We will not need to use a Label Encoder or Standard Scaler for Random Forest.

```
In [194...]: X = test_df.drop(columns=['status', 'tokens'])
y = test_df['status']

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)
```

```
In [195...]: vectorizer = TfidfVectorizer(max_features=10000, ngram_range=(1, 2))

X_train_text = vectorizer.fit_transform(X_train['joined_tokens'])
X_test_text = vectorizer.transform(X_test['joined_tokens'])

X_train_numeric = X_train[['perc_upper_words', 'exclamation_count', 'questio
X_test_numeric = X_test[['perc_upper_words', 'exclamation_count', 'question_'

X_train_combined = hstack([X_train_text, X_train_numeric])
X_test_combined = hstack([X_test_text, X_test_numeric])
```

```
In [196...]: param_grid = {
    'clf__n_estimators': [100, 200, 300],
    'clf__max_depth': [None, 10, 20],
    'clf__min_samples_split': [2, 5, 10],
    'clf__min_samples_leaf': [1, 2, 4],
}

pipeline = Pipeline([
    ('clf', RandomForestClassifier(random_state=42, class_weight='balanced'))
])
```

```
In [198...]: with parallel_backend('loky'):
    grid_search = GridSearchCV(
        pipeline,
        param_grid,
        cv=5,
        scoring='f1_weighted',
```

```

        verbose=0,
        n_jobs=-1,
    )

# Fit GridSearchCV
grid_search.fit(X_train_combined, y_train)

# Evaluate on the test set
test_score = grid_search.score(X_test_combined, y_test)
print(f"Test set score: {test_score}")

print("Best parameters:", grid_search.best_params_)
print("Best F1 score:", grid_search.best_score_)

```

Test set score: 0.7668400989946974
 Best parameters: {'clf__max_depth': None, 'clf__min_samples_leaf': 1, 'clf__min_samples_split': 10, 'clf__n_estimators': 300}
 Best F1 score: 0.7642999388055169

In []: # Assuming grid_search is your GridSearchCV instance and it has been fitted

```

# Get the best model from the grid search
best_model = grid_search.best_estimator_

# Generate predictions on the test set
y_pred = best_model.predict(X_test_combined)

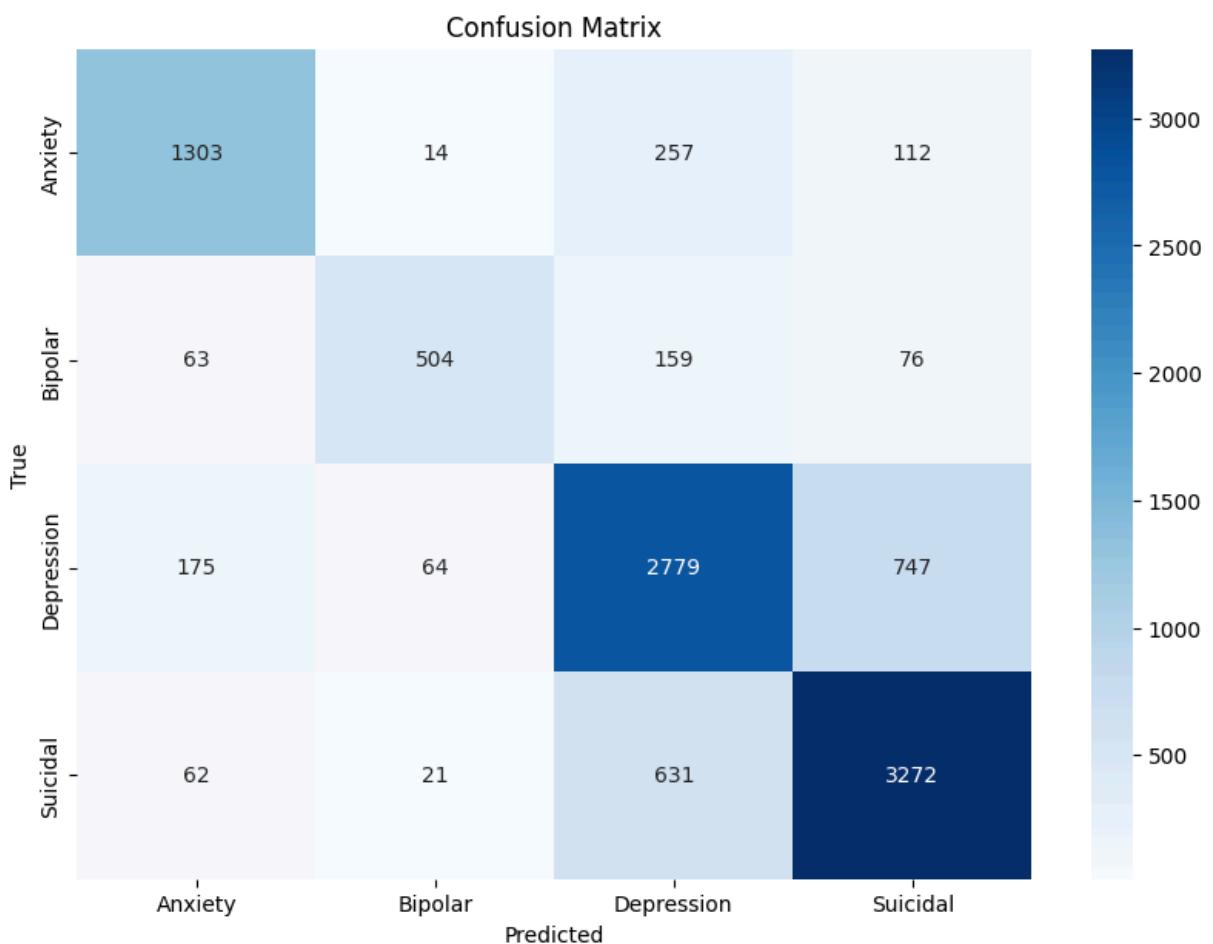
# Print classification report
report = classification_report(y_test, y_pred, target_names=best_model.classes_)
print("Classification Report:\n", report)

# Compute and print confusion matrix
cm = confusion_matrix(y_test, y_pred, labels=best_model.classes_)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=best_model.classes_, yticklabels=best_model.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()

```

Classification Report:

	precision	recall	f1-score	support
Anxiety	0.81	0.77	0.79	1686
Bipolar	0.84	0.63	0.72	802
Depression	0.73	0.74	0.73	3765
Suicidal	0.78	0.82	0.80	3986
accuracy			0.77	10239
macro avg	0.79	0.74	0.76	10239
weighted avg	0.77	0.77	0.77	10239



Woohoo!!! These scores are way better! That's awesome! Let's print our most and least important features.

In []:

```
# Extract feature importances
importances = best_model.named_steps['clf'].feature_importances_

# Feature names from vectorizer and numeric features
text_feature_names = vectorizer.get_feature_names_out()
numeric_feature_names = ['perc_upper_words', 'exclamation_count', 'quest']

# Combine feature names
all_feature_names = np.concatenate([text_feature_names, numeric_feature_names])

# Create a DataFrame for feature importances
feature_importance_df = pd.DataFrame({
    'Feature': all_feature_names,
    'Importance': importances
}).sort_values(by='Importance', ascending=False)

# Print top 100 most important features
top_100_features = feature_importance_df.head(100)
print("\nTop 100 Most Important Features:")
for idx, row in top_100_features.iterrows():
    print(f"{idx + 1}. {row['Feature']}: {row['Importance']:.4f}")

# Print 100 least important features (handling case where there are fewer than 100)
least_important_features = feature_importance_df.sort_values(by='Importance').head(100)
print("\nLeast Important Features:")
for idx, row in least_important_features.iterrows():
    print(f"{idx + 1}. {row['Feature']}: {row['Importance']:.4f}")
```

```
if len(feature_importance_df) > 100:  
    least_100_features = feature_importance_df.tail(100)  
    print("\n100 Least Important Features:")  
    for idx, row in least_100_features.iterrows():  
        print(f"{idx + 1}. {row['Feature']}: {row['Importance']:.4f}")  
else:  
    print("\nFeatures less than 100:")  
    for idx, row in feature_importance_df.iterrows():  
        print(f"{idx + 1}. {row['Feature']}: {row['Importance']:.4f}")
```

Top 100 Most Important Features:

1078. bipolar: 0.0225
2190. die: 0.0180
2076. depression: 0.0173
8547. suicide: 0.0164
10001. perc_upper_words: 0.0133
8536. suicidal: 0.0127
419. anxiety: 0.0105
10003. question_count: 0.0100
5550. manic: 0.0097
10004. word_count: 0.0096
2658. episode: 0.0091
4801. kill: 0.0087
5009. life: 0.0081
5649. med: 0.0077
7304. restless: 0.0069
8441. stress: 0.0063
5549. mania: 0.0058
478. anymore: 0.0051
472. anxious: 0.0047
2058. depressed: 0.0047
9897. worried: 0.0045
9473. wa: 0.0044
6025. nervous: 0.0040
5664. medication: 0.0038
4110. health anxiety: 0.0037
2567. end: 0.0037
8640. symptom: 0.0035
4425. hypomanic: 0.0033
5254. live: 0.0033
2178. diagnosed: 0.0033
3364. friend: 0.0032
4860. lamictal: 0.0032
531. anyone: 0.0032
5867. mood: 0.0030
10002. exclamation_count: 0.0030
2290. disorder: 0.0030
3496. fucking: 0.0030
2899. experience: 0.0030
8235. stable: 0.0029
5551. manic episode: 0.0029
5242. lithium: 0.0028
2319. doctor: 0.0028
5732. mg: 0.0027
4127. heart: 0.0025
4003. happy: 0.0025
2186. diagnosis: 0.0025
4055. hate: 0.0025
6935. psychiatrist: 0.0024
2181. diagnosed bipolar: 0.0024
1011. better: 0.0024
6200. nothing: 0.0024
9752. wish: 0.0023
1081. bipolar disorder: 0.0023
572. anything: 0.0022
4424. hypomania: 0.0021

9900. worry: 0.0021
6423. pain: 0.0021
2113. depressive: 0.0021
3673. going: 0.0021
8954. tired: 0.0021
9636. week: 0.0021
1352. care: 0.0020
8446. stressed: 0.0019
1197. bp: 0.0019
7773. shit: 0.0019
3059. feeling: 0.0019
8692. taking: 0.0019
4109. health: 0.0018
6033. never: 0.0018
4160. help: 0.0018
8048. someone: 0.0018
1130. blood: 0.0018
2796. everything: 0.0018
2773. everyone: 0.0018
1979. death: 0.0018
9883. world: 0.0018
8708. talk: 0.0018
3477. fuck: 0.0018
1337. cancer: 0.0017
282. also: 0.0017
9821. work: 0.0017
217. alive: 0.0017
9572. way: 0.0016
5964. need: 0.0016
5287. living: 0.0016
5422. love: 0.0016
2983. family: 0.0016
3775. good: 0.0016
8653. take: 0.0016
243. alone: 0.0016
540. anyone else: 0.0015
2728. every: 0.0015
7940. sleep: 0.0015
5838. month: 0.0015
2512. else: 0.0015
8260. started: 0.0015
2492. effect: 0.0015
800. back: 0.0014
6463. panic: 0.0014
8080. something: 0.0014

100 Least Important Features:
9952. wrote suicide: 0.0000
613. anything wanted: 0.0000
9884. world anymore: 0.0000
6658. piece paper: 0.0000
3602. german: 0.0000
5415. lot stuff: 0.0000
6722. pleaser: 0.0000
4707. job live: 0.0000
1196. boyfriend love: 0.0000

9130. tried tried: 0.0000
3618. getting life: 0.0000
1988. decent job: 0.0000
4389. human interaction: 0.0000
830. back right: 0.0000
1320. called police: 0.0000
3336. found new: 0.0000
6607. person made: 0.0000
7090. ready die: 0.0000
7994. smile face: 0.0000
9142. trough: 0.0000
3429. friend none: 0.0000
6203. nothing anything: 0.0000
7500. say oh: 0.0000
3355. freezing: 0.0000
3830. got bullied: 0.0000
8663. take everything: 0.0000
8858. themself: 0.0000
2591. end way: 0.0000
4008. happy content: 0.0000
6875. problem life: 0.0000
7082. read online: 0.0000
1061. bible: 0.0000
3376. friend barely: 0.0000
2528. else see: 0.0000
1098. bla: 0.0000
9526. wanted life: 0.0000
3435. friend problem: 0.0000
9475. wa always: 0.0000
6765. polite: 0.0000
9103. tremendous: 0.0000
7919. sit home: 0.0000
3788. good guy: 0.0000
7902. since still: 0.0000
5929. mutual friend: 0.0000
9819. word say: 0.0000
6152. night cry: 0.0000
8553. suicide lot: 0.0000
2651. entity: 0.0000
9807. wolf: 0.0000
1332. camp: 0.0000
288. also going: 0.0000
6336. old still: 0.0000
7547. school always: 0.0000
2675. escapism: 0.0000
2823. everything look: 0.0000
5012. life ahead: 0.0000
3462. friend wa: 0.0000
9379. used friend: 0.0000
8292. stating: 0.0000
2338. done bad: 0.0000
9860. work towards: 0.0000
5119. life moment: 0.0000
8156. soulmate: 0.0000
1614. committed: 0.0000
4813. kill somebody: 0.0000

```
3568. gallbladder: 0.0000
7828. shred: 0.0000
2484. economy: 0.0000
4338. hoping better: 0.0000
6790. posture: 0.0000
361. amazing friend: 0.0000
7161. red flag: 0.0000
5179. life turn: 0.0000
9543. wanting live: 0.0000
7233. remained: 0.0000
2853. ex girlfriend: 0.0000
6860. primary school: 0.0000
4292. holter: 0.0000
2859. examination: 0.0000
4318. hope dream: 0.0000
8687. taken advantage: 0.0000
1049. better tried: 0.0000
5536. making progress: 0.0000
3242. first love: 0.0000
7453. said needed: 0.0000
2661. equally: 0.0000
7344. right hand: 0.0000
9372. urine test: 0.0000
1427. change fact: 0.0000
2852. ex friend: 0.0000
9480. wa getting: 0.0000
9503. waking moment: 0.0000
8445. stress test: 0.0000
337. always knew: 0.0000
7034. radiation: 0.0000
4350. horror movie: 0.0000
3546. fundamental: 0.0000
3174. final exam: 0.0000
339. always made: 0.0000
3902. gray: 0.0000
```

Dropping all Numeric Features

Let's experiment with dropping all of our numeric features, and just focusing on text data, and see how that affects our model.

```
In [ ]: X = test_df.drop(columns=['status', 'tokens'])
y = test_df['status']

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

vectorizer = TfidfVectorizer(max_features=10000, ngram_range=(1, 2))

X_train_text = vectorizer.fit_transform(X_train['joined_tokens'])
X_test_text = vectorizer.transform(X_test['joined_tokens'])

param_grid = {
    'clf__n_estimators': [100, 200, 300],
```

```
'clf__max_depth': [None, 10, 20],
'clf__min_samples_split': [2, 5, 10],
'clf__min_samples_leaf': [1, 2, 4],
}

pipeline = Pipeline([
    ('clf', RandomForestClassifier(random_state=42, class_weight='balanced'))
])

with parallel_backend('loky'):
    grid_search = GridSearchCV(
        pipeline,
        param_grid,
        cv=5,
        scoring='f1_weighted',
        verbose=0,
        n_jobs=-1,
    )

    # Fit GridSearchCV
    grid_search.fit(X_train_text, y_train)

    # Evaluate on the test set
    test_score = grid_search.score(X_test_text, y_test)
    print(f"Test set score: {test_score}")

    print("Best parameters:", grid_search.best_params_)
    print("Best F1 score:", grid_search.best_score_)

# Assuming grid_search is your GridSearchCV instance and it has been fitted

# Get the best model from the grid search
best_model = grid_search.best_estimator_

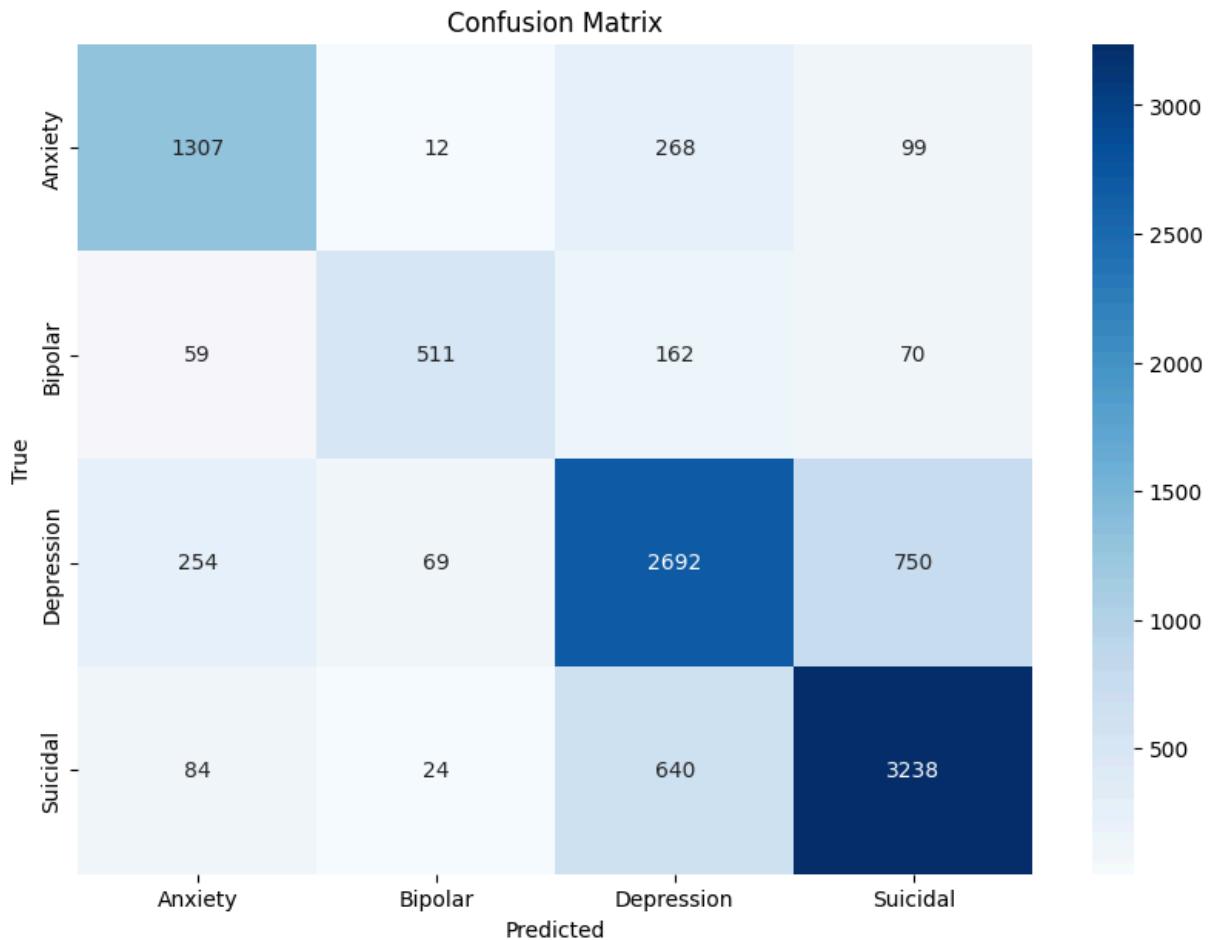
# Generate predictions on the test set
y_pred = best_model.predict(X_test_text)

# Print classification report
report = classification_report(y_test, y_pred, target_names=best_model.classes_)
print("Classification Report:\n", report)

# Compute and print confusion matrix
cm = confusion_matrix(y_test, y_pred, labels=best_model.classes_)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=best_model.classes_, yticklabels=best_model.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()
```

Test set score: 0.7560469299821828
 Best parameters: {'clf__max_depth': None, 'clf__min_samples_leaf': 1, 'clf__min_samples_split': 10, 'clf__n_estimators': 300}
 Best F1 score: 0.753766714699417
 Classification Report:

	precision	recall	f1-score	support
Anxiety	0.77	0.78	0.77	1686
Bipolar	0.83	0.64	0.72	802
Depression	0.72	0.72	0.72	3765
Suicidal	0.78	0.81	0.80	3986
accuracy			0.76	10239
macro avg	0.77	0.73	0.75	10239
weighted avg	0.76	0.76	0.76	10239



Nope! Our scores were worse. We'll keep our numeric features, and experiment with dropping them again if need be.

Now let's try Multinomial Logistic Regression and compare all our models.

Multinomial Logistic Regression

```
In [ ]: X = test_df.drop(columns=['status', 'tokens'])
y = test_df['status']

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)

encoder = LabelEncoder()
y_train_encoded = encoder.fit_transform(y_train)
y_test_encoded = encoder.transform(y_test)
y_train = pd.Series(y_train_encoded)
y_test = pd.Series(y_test_encoded)

# Check encoded labels
print("Encoded y_train:")
print(y_train.value_counts())

print("Encoded y_test:")
print(y_test.value_counts())

# Mapping from encoded labels to original labels
label_mapping = dict(zip(encoder.classes_, encoder.transform(encoder.classes)))
print("Label Mapping:")
print(label_mapping)

vectorizer = TfidfVectorizer(max_features=20000, ngram_range=(1, 2))

X_train_text = vectorizer.fit_transform(X_train['joined_tokens'])
X_test_text = vectorizer.transform(X_test['joined_tokens'])

# Scale numeric data
# Initialize StandardScaler
scaler = StandardScaler()

# Fit and transform StandardScaler on numeric data
X_train_numeric_scaled = scaler.fit_transform(X_train[['perc_upper_words', 'exclam']])
X_test_numeric_scaled = scaler.transform(X_test[['perc_upper_words', 'exclam']])

X_train_combined = hstack([X_train_text, X_train_numeric_scaled])
X_test_combined = hstack([X_test_text, X_test_numeric_scaled])

param_grid = {
    'clf__C': [0.1, 1.0, 10.0],
    'clf__solver': ['lbfgs', 'saga'],
    'clf__max_iter': [2000], # Balance computational cost with convergence needs
}

# Logistic Regression
pipeline = Pipeline([
    ('clf', LogisticRegression(class_weight='balanced', random_state=42)), #
])

# Initialize GridSearchCV
grid_search = GridSearchCV(
    pipeline,
```

```
param_grid,
cv=5,
scoring='f1_weighted',
verbose=0,
n_jobs=-1,
)

# Fit GridSearchCV
grid_search.fit(X_train_combined, y_train)

# Evaluate on the test set
test_score = grid_search.score(X_test_combined, y_test)
print(f"Test set score: {test_score}")

print("Best parameters:", grid_search.best_params_)
print("Best F1 score:", grid_search.best_score_)
```

Encoded y_train:

```
3    9187
2    8781
0    4223
1    1699
Name: count, dtype: int64
Encoded y_test:
3    3986
2    3765
0    1686
1     802
Name: count, dtype: int64
Label Mapping:
{'Anxiety': np.int64(0), 'Bipolar': np.int64(1), 'Depression': np.int64(2),
 'Suicidal': np.int64(3)}
```

```
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linea
r_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which mean
s the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linea
r_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which mean
s the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linea
r_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which mean
s the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linea
r_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which mean
s the coef_ did not converge
    warnings.warn(
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linea
r_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which mean
s the coef_ did not converge
    warnings.warn(
```

```
Test set score: 0.7487672292766013
Best parameters: {'clf__C': 10.0, 'clf__max_iter': 2000, 'clf__solver': 'sag
a'}
Best F1 score: 0.747527147065292
```

```
/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/sklearn/linea  
r_model/_sag.py:349: ConvergenceWarning: The max_iter was reached which mean  
s the coef_ did not converge  
    warnings.warn(
```

Woohoo! We got some warnings that our model didn't reach convergence, but at least our models are actually running now and printing out scores!!! That's a huge improvement!

Let's try increasing our max iterations. We can also experiment with dropping our numeric features.

```
In [ ]: X = test_df.drop(columns=['status', 'tokens'])  
y = test_df['status']  
  
X_train, X_test, y_train, y_test = train_test_split(  
    X, y, test_size=0.3, random_state=42  
)  
  
encoder = LabelEncoder()  
y_train_encoded = encoder.fit_transform(y_train)  
y_test_encoded = encoder.transform(y_test)  
y_train = pd.Series(y_train_encoded)  
y_test = pd.Series(y_test_encoded)  
  
# Check encoded labels  
print("Encoded y_train:")  
print(y_train.value_counts())  
  
print("Encoded y_test:")  
print(y_test.value_counts())  
  
# Mapping from encoded labels to original labels  
label_mapping = dict(zip(encoder.classes_, encoder.transform(encoder.classes)))  
print("Label Mapping:")  
print(label_mapping)  
  
vectorizer = TfidfVectorizer(max_features=20000, ngram_range=(1, 2))  
  
X_train_text = vectorizer.fit_transform(X_train['joined_tokens'])  
X_test_text = vectorizer.transform(X_test['joined_tokens'])  
  
# Scale numeric data  
# Initialize StandardScaler  
scaler = StandardScaler()  
  
# Fit and transform StandardScaler on numeric data  
X_train_numeric_scaled = scaler.fit_transform(X_train[['perc_upper_words', 'exclamation_marks', 'question_marks', 'length']] )  
X_test_numeric_scaled = scaler.transform(X_test[['perc_upper_words', 'exclamation_marks', 'question_marks', 'length']])  
  
X_train_combined = hstack([X_train_text, X_train_numeric_scaled])  
X_test_combined = hstack([X_test_text, X_test_numeric_scaled])  
  
param_grid = {
```

```

    'clf__C': [0.1, 1.0, 10.0],
    'clf__solver': ['lbfgs', 'saga'],
    'clf__max_iter': [5000, 10000], # Balance computational cost with coverage
}

# Logistic Regression
pipeline = Pipeline([
    ('clf', LogisticRegression(class_weight='balanced', random_state=42)), #
])

# Initialize GridSearchCV
grid_search = GridSearchCV(
    pipeline,
    param_grid,
    cv=5,
    scoring='f1_weighted',
    verbose=0,
    n_jobs=-1,
)

# Fit GridSearchCV
grid_search.fit(X_train_combined, y_train)

# Evaluate on the test set
test_score = grid_search.score(X_test_combined, y_test)
print(f"Test set score: {test_score}")

print("Best parameters:", grid_search.best_params_)
print("Best F1 score:", grid_search.best_score_)

```

Encoded y_train:

```

3      9187
2      8781
0      4223
1      1699
Name: count, dtype: int64

```

Encoded y_test:

```

3      3986
2      3765
0      1686
1      802
Name: count, dtype: int64

```

Label Mapping:

```
{'Anxiety': np.int64(0), 'Bipolar': np.int64(1), 'Depression': np.int64(2),
'Suicidal': np.int64(3)}
```

Test set score: 0.7512178561917288

Best parameters: {'clf__C': 10.0, 'clf__max_iter': 5000, 'clf__solver': 'saga'}

Best F1 score: 0.7486999850146129

```
In [ ]: # Get the best model from the grid search
best_model = grid_search.best_estimator_

# Generate predictions on the test set
y_pred = best_model.predict(X_test_combined)
```

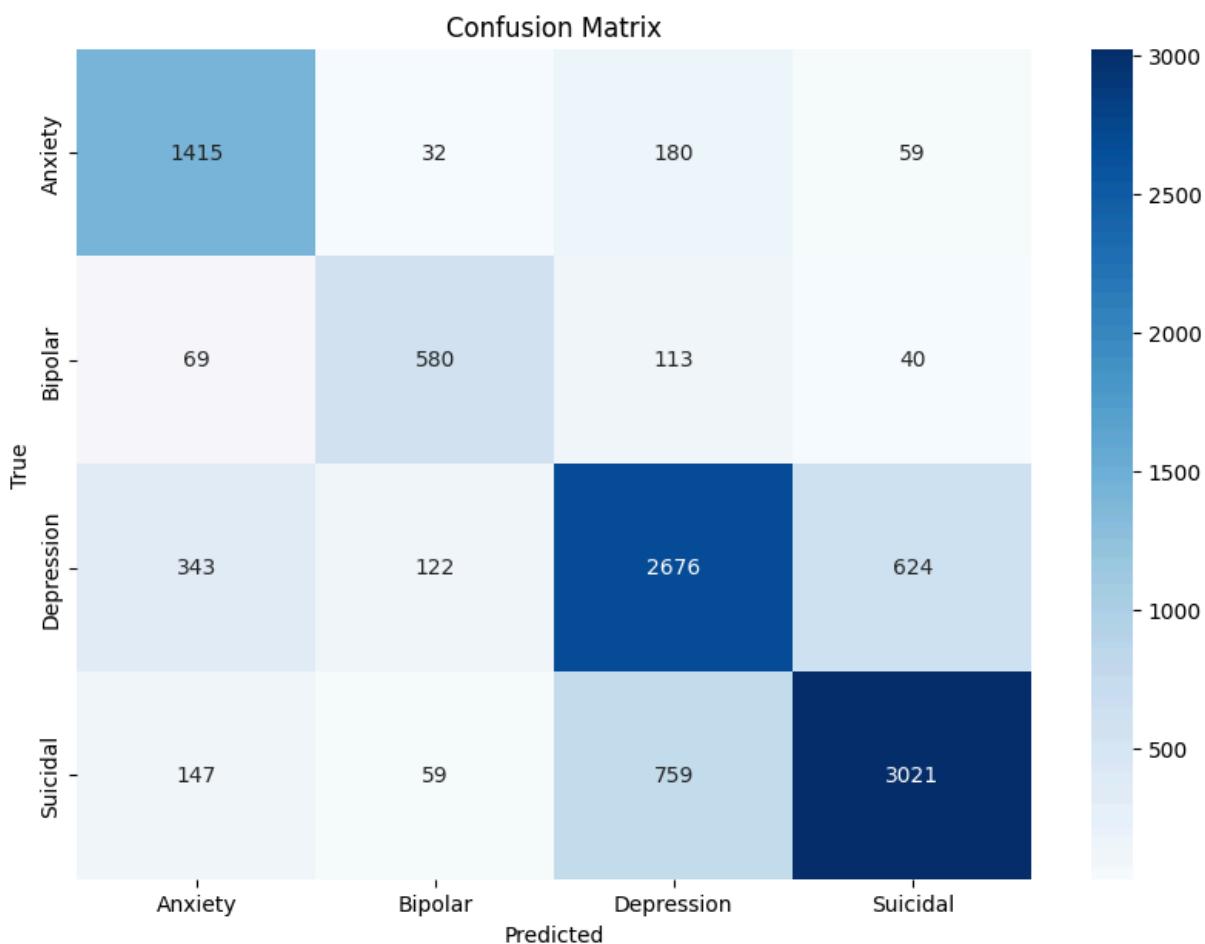
```
# Convert encoded predictions back to original labels
y_pred_original = encoder.inverse_transform(y_pred)
y_test_original = encoder.inverse_transform(y_test)

# Print classification report
report = classification_report(y_test_original, y_pred_original, target_names=target_names)
print("Classification Report:\n", report)

# Compute and print confusion matrix
cm = confusion_matrix(y_test_original, y_pred_original, labels=encoder.classes_)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=encoder.classes_, yticklabels=encoder.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()
```

Classification Report:

	precision	recall	f1-score	support
Anxiety	0.72	0.84	0.77	1686
Bipolar	0.73	0.72	0.73	802
Depression	0.72	0.71	0.71	3765
Suicidal	0.81	0.76	0.78	3986
accuracy			0.75	10239
macro avg	0.74	0.76	0.75	10239
weighted avg	0.75	0.75	0.75	10239



Woohoo!!! Our scores are so much better!!! I couldn't be happier!! This is awesome!

XGBoost: LEAVING AS MARKDOWN

We tried running XGBoost at the beginning of our process. It unsurprisingly took too long. For this particular project, we can't keep trying to run models that will take too much of the limited time we have. We'll leave the code we tried for XGBoost below as markdowns...in case we want to come revisit and edit this later on!

CODE FOR XGBOOST:

Let's try XGBoost, which is known to yield extraordinary results on large, complex datasets that require multi-categorical targets.

Let's train, test, split.

We can **include our numeric features** this time, in addition to our text data which will be vectorized.

In []:

```
##### Define features and target
X = test_df.drop(columns=['status', 'tokens'])
y = test_df['status'] # 'status' is the target variable
```

```
##### Train-test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42
)
```

LabelEncoder()

We need to use LabelEncoder() to make sure we have correctly formatted indices for XGBoost, starting at '0'.

```
In [ ]: y_train.value_counts().sort_index()
```

```
Out[ ]: status
Anxiety      4223
Bipolar      1699
Depression   8781
Suicidal     9187
Name: count, dtype: int64
```

```
In [ ]: ##### Instantiate the encoder
encoder = LabelEncoder()

##### Fit the encoder on y_train and transform it
y_train_encoded = encoder.fit_transform(y_train)

##### Transform y_test based on the encoder fitted on y_train
y_test_encoded = encoder.transform(y_test)

##### Convert back to pandas Series
y_train = pd.Series(y_train_encoded)
y_test = pd.Series(y_test_encoded)

##### Check encoded labels
print("Encoded y_train:")
print(y_train.value_counts())

print("Encoded y_test:")
print(y_test.value_counts())

##### Mapping from encoded labels to original labels
label_mapping = dict(zip(encoder.classes_, encoder.transform(encoder.classes)))

print("Label Mapping:")
label_mapping
```

```

Encoded y_train:
3    9187
2    8781
0    4223
1    1699
Name: count, dtype: int64
Encoded y_test:
3    3986
2    3765
0    1686
1     802
Name: count, dtype: int64
Label Mapping:
Out[ ]: {'Anxiety': np.int64(0),
          'Bipolar': np.int64(1),
          'Depression': np.int64(2),
          'Suicidal': np.int64(3)}

```

Vectorize - TF-IDF

We need to vectorize our 'joined_tokens' again.

```

In [ ]: # Step 1: Initialize TfidfVectorizer
vectorizer = TfidfVectorizer(max_features=10000, ngram_range=(1, 2))

# Step 2: Transform text data
X_train_text = vectorizer.fit_transform(X_train['joined_tokens'])
X_test_text = vectorizer.transform(X_test['joined_tokens'])

# Prepare numeric features
X_train_numeric = X_train[['perc_upper_words', 'exclamation_count', 'question_
X_test_numeric = X_test[['perc_upper_words', 'exclamation_count', 'question_'

# Convert numeric arrays to sparse matrices
X_train_numeric_sparse = csr_matrix(X_train_numeric)
X_test_numeric_sparse = csr_matrix(X_test_numeric)

# Step 3: Combine text and numeric features
X_train_combined = hstack([X_train_text, X_train_numeric_sparse])
X_test_combined = hstack([X_test_text, X_test_numeric_sparse])

```

Best Parameters - GridSearchCV

There are so many parameters we can test for...but we'll want to limit the number we run through GridSearchCV to cut down on computational cost.

```

In [ ]: ##### Define the parameter grid
param_grid = {
    'learning_rate': [0.1, 0.2],
    'max_depth': [3, 5],
    'min_child_weight': [1, 3],
    'subsample': [0.8],
}

```

```

    'gamma': [0, 0.1],
}

In [ ]: ##### Initialize XGBClassifier
xgb_model = XGBClassifier(objective='multi:softmax', num_class=7, random_state=42)

##### Initialize GridSearchCV
grid_search = GridSearchCV(
    estimator=xgb_model,
    param_grid=param_grid,
    scoring='f1_weighted', # Use F1 score to account for class imbalance
    cv=5, # Number of cross-validation folds
    verbose=0,
    n_jobs=-1 # Use all available cores
)

##### Fit GridSearchCV
grid_search.fit(X_train_combined, y_train)

##### Best parameters and score
print(f"Best parameters: {grid_search.best_params_}")
print(f"Best F1 score: {grid_search.best_score_}")

##### Evaluate on the test set
best_model = grid_search.best_estimator_
test_score = best_model.score(X_test_combined, y_test)
print(f"Test set score: {test_score}")

```

/opt/anaconda3/envs/learn-env-3.9/lib/python3.10/site-packages/joblib/externals/loky/process_executor.py:752: UserWarning: A worker stopped while some jobs were given to the executor. This can be caused by a too short worker timeout or by a memory leak.

```

warnings.warn(
Best parameters: {'gamma': 0, 'learning_rate': 0.2, 'max_depth': 5, 'min_child_weight': 1, 'subsample': 0.8}
Best F1 score: 0.7638513087005425
Test set score: 0.7691180779372986

```

Graphs For Presentation

```

In [ ]: # Calculate mean exclamation_count for each status group
mean_counts = filtered_df.groupby('status')['exclamation_count'].mean().reset_index()

# Assuming mean_counts is your DataFrame with mean exclamation counts for each status
mean_counts = mean_counts.sort_values(by='exclamation_count')

# Define custom colors for specific status groups
status_colors = {
    'Bipolar': '#8dd8d3', # Turquoise
    'Depression': '#69c6ef', # Light Blue
    'Suicidal': '#d7e6a3', # Light Green
    'Anxiety': '#df6fba', # Pink
    'Normal': '#e4a959' # Orange

```

```

}

# Plotting bar graph with custom colors
plt.figure(figsize=(10, 6))
sns.barplot(x='status', y='exclamation_count', data=mean_counts, palette=status_colors)
plt.title('Mean Exclamation Count by Status', color='white') # Set title color to white
plt.xlabel('Status', color='white') # Set x-axis label color to white
plt.ylabel('Mean Exclamation Count', color='white') # Set y-axis label color to white
plt.xticks(color='white') # Set x-axis tick label color to white
plt.yticks(color='white') # Set y-axis tick label color to white
plt.xticks(rotation=45) # Rotate x-axis labels for better readability if needed
plt.gca().spines['top'].set_color('none') # Remove top spine
plt.gca().spines['right'].set_color('none') # Remove right spine
plt.gca().spines['left'].set_color('white') # Set left spine color to white
plt.gca().spines['bottom'].set_color('white') # Set bottom spine color to white
plt.tight_layout()
plt.savefig('mean_exclamation_count_by_status.png', transparent=True) # Save plot
plt.show()

```

/var/folders/1k/x4fbmx7x5jq2rlv4qntp985r0000gn/T/ipykernel_87545/4048049819.py:19: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```
sns.barplot(x='status', y='exclamation_count', data=mean_counts, palette=status_colors.values())
```



In []: # Define custom colors for specific status groups
status_colors = {
 'Bipolar': '#8dd8d3', # Turquoise
 'Depression': '#69c6ef', # Light Blue
 'Suicidal': '#d7e6a3', # Light Green
 'Anxiety': '#df6fba', # Pink

```
'Normal': '#e4a959'      # Orange
}

# Calculate mean exclamation_count for each status group
mean_counts = filtered_df.groupby('status')['exclamation_count'].mean().reset_index()

# Assuming mean_counts is your DataFrame with mean exclamation counts for each status
mean_counts = mean_counts.sort_values(by='exclamation_count')

# Plotting bar graph with custom colors
plt.figure(figsize=(10, 6))
sns.barplot(x='status', y='exclamation_count', data=mean_counts,
            palette=status_colors) # Use palette argument to assign colors
plt.title('Mean Exclamation Count by Status', color='white') # Set title color to white
plt.xlabel('Status', color='white') # Set x-axis label color to white
plt.ylabel('Mean Exclamation Count', color='white') # Set y-axis label color to white
plt.xticks(color='white') # Set x-axis tick label color to white
plt.yticks(color='white') # Set y-axis tick label color to white
plt.xticks(rotation=45) # Rotate x-axis labels for better readability if needed
plt.gca().spines['top'].set_color('none') # Remove top spine
plt.gca().spines['right'].set_color('none') # Remove right spine
plt.gca().spines['left'].set_color('white') # Set left spine color to white
plt.gca().spines['bottom'].set_color('white') # Set bottom spine color to white
plt.tight_layout()
plt.savefig('mean_exc_count_by_status.png', transparent=True) # Save with transparent background
plt.show()
```

/var/folders/1k/x4fbmx7x5jq2rlv4qntp985r0000gn/T/ipykernel_87545/354745326.py:18: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```
sns.barplot(x='status', y='exclamation_count', data=mean_counts,
```



```
In [ ]: # Define custom colors for specific status groups
status_colors = {
    'Bipolar': '#8dd8d3',      # Turquoise
    'Depression': '#69c6ef',   # Light Blue
    'Suicidal': '#d7e6a3',     # Light Green
    'Anxiety': '#df6fba',      # Pink
    'Normal': '#e4a959'        # Orange
}

# Calculate mean exclamation_count for each status group
mean_counts = filtered_df.groupby('status')['word_count'].mean().reset_index

# Assuming mean_counts is your DataFrame with mean exclamation counts for each status group
mean_counts = mean_counts.sort_values(by='word_count')

# Plotting bar graph with custom colors
plt.figure(figsize=(10, 6))
sns.barplot(x='status', y='word_count', data=mean_counts,
            palette=status_colors) # Use palette argument to assign colors
plt.title('Mean Word Count by Status', color='white') # Set title color to white
plt.xlabel('Status', color='white') # Set x-axis label color to white
plt.ylabel('Mean Word Count', color='white') # Set y-axis label color to white
plt.xticks(color='white') # Set x-axis tick label color to white
plt.yticks(color='white') # Set y-axis tick label color to white
plt.xticks(rotation=45) # Rotate x-axis labels for better readability if needed
plt.gca().spines['top'].set_color('none') # Remove top spine
plt.gca().spines['right'].set_color('none') # Remove right spine
plt.gca().spines['left'].set_color('white') # Set left spine color to white
plt.gca().spines['bottom'].set_color('white') # Set bottom spine color to white
plt.tight_layout()
plt.savefig('mean_word_count_by_status.png', transparent=True) # Save with transparent background
plt.show()
```

```
/var/folders/1k/x4fbmx7x5jq2rlv4qntp985r0000gn/T/ipykernel_87545/547643034.py:18: FutureWarning:
```

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```
sns.barplot(x='status', y='word_count', data=mean_counts,
```



```
In [ ]: # Same graph for README

import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd

# Define custom colors for specific status groups
status_colors = {
    'Bipolar': '#8dd8d3',      # Turquoise
    'Depression': '#69c6ef',   # Light Blue
    'Suicidal': '#d7e6a3',     # Light Green
    'Anxiety': '#df6fba',      # Pink
    'Normal': '#e4a959'        # Orange
}

# Calculate mean word_count for each status group
mean_counts = filtered_df.groupby('status')['word_count'].mean().reset_index

# Assuming mean_counts is your DataFrame with mean word counts for each stat
mean_counts = mean_counts.sort_values(by='word_count')

# Plotting bar graph with custom colors
plt.figure(figsize=(10, 6))
sns.barplot(x='status', y='word_count', data=mean_counts,
            palette=status_colors) # Use palette argument to assign colors
```

```

# Update plot appearance
plt.title('Mean Word Count by Status', color='black') # Set title color to black
plt.xlabel('Status', color='black') # Set x-axis label color to black
plt.ylabel('Mean Word Count', color='black') # Set y-axis label color to black
plt.xticks(color='black') # Set x-axis tick label color to black
plt.yticks(color='black') # Set y-axis tick label color to black
plt.xticks(rotation=45) # Rotate x-axis labels for better readability if needed

# Configure gridlines to be black
plt.grid(True, linestyle='--', color='black', alpha=0.6) # Add gridlines with black dashed lines

# Set spine colors
plt.gca().spines['top'].set_color('black') # Set top spine color to black
plt.gca().spines['right'].set_color('black') # Set right spine color to black
plt.gca().spines['left'].set_color('black') # Set left spine color to black
plt.gca().spines['bottom'].set_color('black') # Set bottom spine color to black

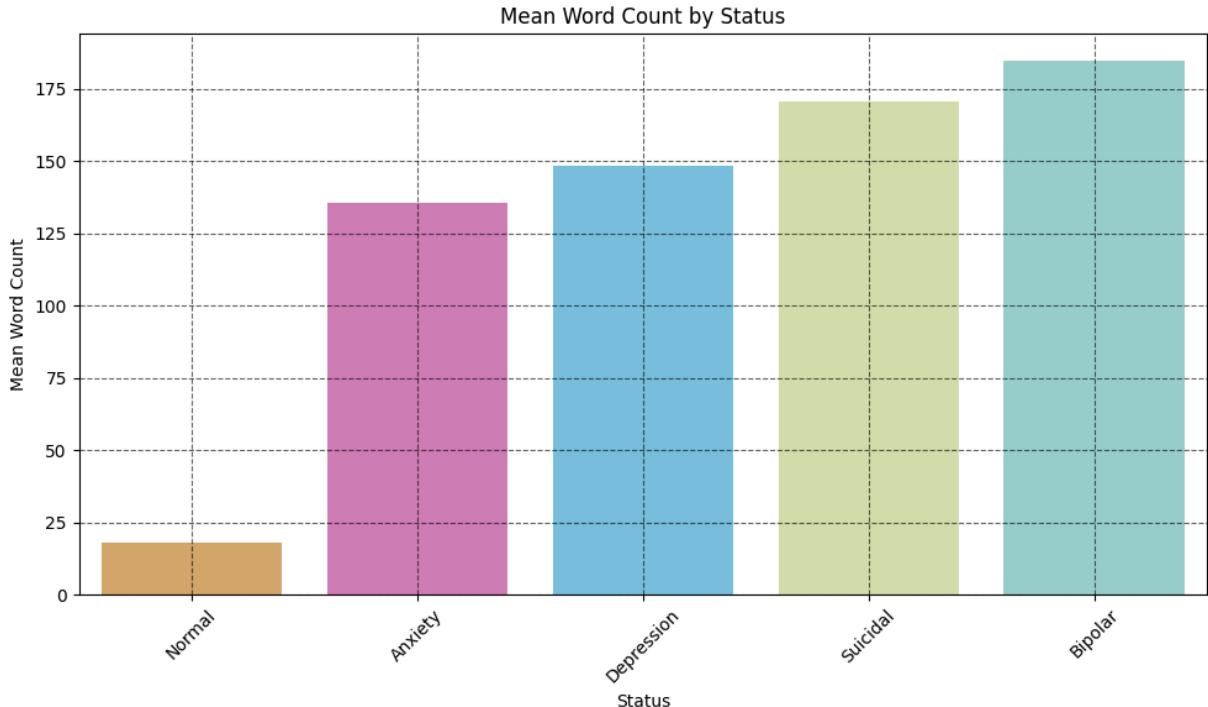
plt.tight_layout()
plt.savefig('README_wordcount_status.png', dpi=300, bbox_inches='tight') # Save the figure
plt.show()

```

/var/folders/1k/x4fbmx7x5jq2rlv4qntp985r0000gn/T/ipykernel_87545/1962860000.py:24: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```
sns.barplot(x='status', y='word_count', data=mean_counts,
```



In []: mean_counts

```
Out[ ]:    status  word_count
            3      Normal   18.094763
            0     Anxiety   135.412422
            2  Depression   148.468516
            4   Suicidal   170.623320
            1     Bipolar   184.722111
```

```
In [ ]: filtered_df['status'].value_counts()
```

```
Out[ ]: status
Normal      15829
Suicidal    13173
Depression  12546
Anxiety      5909
Bipolar      2501
Name: count, dtype: int64
```

```
In [ ]: # Create a legend with custom colors
legend_handles = [plt.Line2D([0, 1], [0, 0], color=color, linewidth=10, linecap='round')
                  for color in status_colors.values()]
legend_labels = list(status_colors.keys())

# Plot the legend
plt.figure(figsize=(8, 2))
legend = plt.legend(legend_handles, legend_labels, loc='center', frameon=False)
plt.axis('off')

# Set legend text color to white
for text in legend.get_texts():
    text.set_color('white')

# Save the legend with transparent background
plt.savefig('legend.png', transparent=True, bbox_inches='tight', pad_inches=0)
plt.show()
```



```
In [ ]: # Step 1: Prepare the TF-IDF vectorizer and fit_transform on the combined text
vectorizer = TfidfVectorizer(max_features=20000, ngram_range=(1, 2))
X_text = vectorizer.fit_transform(test_df['joined_tokens'])

# Step 2: Fit TF-IDF vectorizer for each class
class_labels = test_df['status'].unique()
```

```
# Create a dictionary to map class labels to their respective TF-IDF scores
class_tfidf_scores = {}

for label in class_labels:
    class_text = test_df[test_df['status'] == label]['joined_tokens']
    class_tfidf = vectorizer.transform(class_text)
    # Compute the mean TF-IDF score for each feature
    mean_tfidf_scores = class_tfidf.mean(axis=0).A1
    feature_names = vectorizer.get_feature_names_out()
    # Store the TF-IDF scores for the current class
    class_tfidf_scores[label] = dict(zip(feature_names, mean_tfidf_scores))

# Step 3: Generate and plot word clouds for each class
output_dir = '/Users/emmascotson/Documents/project_4/images/' # Updated path

# Ensure the output directory exists
if not os.path.exists(output_dir):
    os.makedirs(output_dir)

dpi = 300 # Higher DPI for better quality
figsize = (12, 6) # Increase figure size for better resolution

for i, label in enumerate(class_labels):
    plt.figure(figsize=figsize, dpi=dpi)

    # Get top words for the current class
    tfidf_scores = class_tfidf_scores[label]
    sorted_words = sorted(tfidf_scores.items(), key=lambda x: x[1], reverse=True)
    top_words = dict(sorted_words[:200]) # Adjust the number of top words if needed

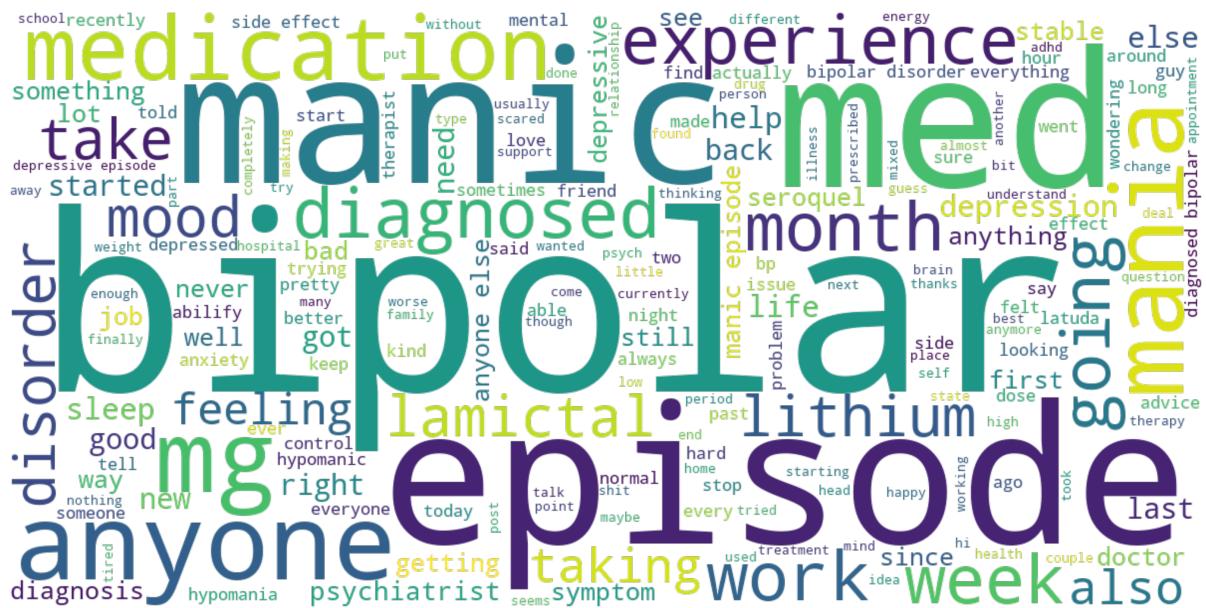
    # Generate word cloud with transparent background
    wordcloud = WordCloud(
        background_color=None, # Transparent background
        mode='RGBA', # Ensure transparency is maintained
        width=1200, # Width of the word cloud image
        height=600 # Height of the word cloud image
    ).generate_from_frequencies(top_words)

    # Plot word cloud
    plt.imshow(wordcloud, interpolation='bilinear')
    plt.axis('off') # No axis
    # Remove title of status
    plt.title('')

    # Save word cloud to file
    wordcloud_image_path = os.path.join(output_dir, f'{label}_wordcloud.png')
    plt.savefig(wordcloud_image_path, format='png', dpi=dpi, bbox_inches='tight')

plt.show()
```





Conclusions

Best Model: Random Forest

XGBoost yielded similar results, yet is more computationally expensive.

Next Steps: Tuning, Further EDA, Reclassifying

Next steps will be continuing to tune hyperparameters, and spending more time seriously examining how the rows in this dataframe are being labeled and whether there are more mislabeled disorders (like the 'Depression' rows that were really 'Suicidal').

For example, when combing through 'take' bigrams in 'Depression', I noticed medications that are prescribed for Bipolar. With more time, I can examine the statements associated with these values and determine what context in which they were used to possibly reclassify these rows.

I can also more thoroughly research how and why certain mood disorders are diagnosed, to examine the dataset's original labeling system and harness domain knowledge to improve my models.

More improvements on my models will be needed before any steps can be taken toward developing a resource-providing chatbot.