PDS ASSIGNMENT 3

SAI PRIYATHAM GUMMADI

16331093

In this assignment we were given a dataset and we were asked to perform these below mentioned tasks on it.

* Convert the text corpus into tokens.
* Perform stop word removal.
* Count Word frequencies
* Create word clouds.

For performing these tasks, we have given a code.

* In the code , we first load the dataset using Pandas. We then tokenize the text corpus by looping through each tweet, tokenizing it using NLTK's word\_tokenize function, and adding the tokens to a list.
* Next, we remove stop words from the token list using NLTK's stopwords corpus. We then count the frequency of each word in the token list using Python's Counter class.
* Finally, we use the WordCloud class from the WordCloud library to create a word cloud from the word frequency data. We set the word cloud's size and background color, and provide it with the stop words and minimum font size. We then show the word cloud using Matplotlib.

SNAPSHOTS OF THE CODE:

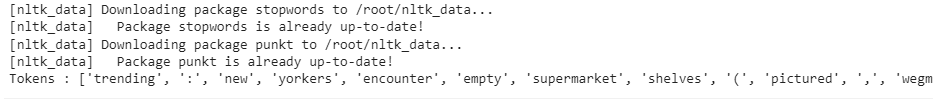
Graphical user interface, text, application, chat or text message

Description automatically generated

Graphical user interface, application

Description automatically generated

OUTPUTS:





A picture containing text

Description automatically generated