# Stock Sentiment Analysis Code Explanation

## 1. Import Libraries

The script begins by importing necessary libraries:  
- `praw`: A Python wrapper for the Reddit API, used to fetch posts from subreddits.  
- `pandas`: A library for data manipulation and analysis, used to handle data in DataFrame format.  
- `matplotlib`: A library for creating static, animated, and interactive visualizations in Python.  
- `seaborn`: A visualization library based on matplotlib that provides a high-level interface for drawing attractive statistical graphics.  
- `textblob`: A library for processing textual data, providing a simple API for common natural language processing tasks.  
- `datetime`: A module to work with dates and times, used to convert timestamps from Reddit posts.

## 2. Reddit API Authentication

The script establishes a connection to the Reddit API using the `praw.Reddit` class. You need to provide your `client\_id`, `client\_secret`, and `user\_agent` to authenticate your app.

## 3. Fetch Reddit Posts

The `fetch\_reddit\_posts` function fetches posts from a specified subreddit. It takes two arguments:  
- `subreddit`: The name of the subreddit from which to fetch posts.  
- `limit`: The maximum number of posts to retrieve (default is 100).  
The function collects the post title, body, and creation date, and returns a DataFrame with this information.

## 4. Sentiment Analysis

The `analyze\_sentiment` function uses TextBlob to analyze the sentiment of the given text. It returns a polarity score between -1 (negative sentiment) and 1 (positive sentiment).

## 5. Data Preprocessing

The `preprocess\_data` function cleans and preprocesses the posts DataFrame:  
- It handles missing values by replacing them with an empty string.  
- It combines the title and body into a single text field.  
- It applies sentiment analysis to each text entry and labels the sentiment as positive, negative, or neutral based on the polarity score.

## 6. Data Visualization

The `visualize\_sentiment` function creates a count plot of the sentiment distribution using Seaborn. It visualizes how many posts fall into each sentiment category (positive, neutral, negative).

## 7. Main Function

The `main` function orchestrates the workflow:  
- It specifies the subreddit to scrape (e.g., r/wallstreetbets).  
- It calls the fetch, preprocess, and visualize functions in sequence.  
- Finally, it saves the resulting DataFrame to a CSV file for potential future analysis.

## 8. Running the Script

The script is designed to be executed as a standalone program. When run, it will scrape data, perform analysis, visualize the results, and save the data.