Required Libraries

praw: A Python library to interact with Reddit's API.

Creating the API (Client ID, Client Secret, and User Agent)

To interact with Reddit's API, you need an application key from Reddit:

1. **Sign in to Reddit**: Go to https://www.reddit.com/prefs/apps.

2. **Create an Application**:

   o Click on "Create App" or "Create Another App".

   o Fill out the form:

     ▪ **Name**: A name for your app.

     ▪ **App Type**: Choose "script" for personal use.

     ▪ **Redirect URL**: http://localhost:8080

   o Submit the form to generate:

     ▪ **Client ID**

     ▪ **Client Secret**

User Agent is a string that describes your app (Usually Reddit username is given here).

Explanation of the Code

1. **Authenticating Reddit API**:

reddit_read_only = praw.Reddit(client_id="...", client_secret="...", user_agent="...")

   o Allows you to fetch data from Reddit.

2. **Selecting a Subreddit**:

subreddit = reddit_read_only.subreddit("wallstreetbets")

   o Specifies the subreddit (e.g., "wallstreetbets") from which to scrape posts.

   o Modify this term to target a different subreddit.

3. **Fetching Top Posts**:

```
posts = subreddit.top(limit=None)
```

- o Retrieves top posts based on subreddit activity.

- o Change top to hot, new, or rising to fetch posts in different sorting orders.

4. **Creating a Dictionary for Data**:

```
posts_dict = {"Title": [], "Post Text": [], "ID": [], "Score": [], "Total Comments": [], "Post
URL": [], "Date": []}
```

- o Defines the structure for storing post details.

5. **Looping through Posts**:

```
for post in posts:

    posts_dict["Title"].append(post.title)

    ...
```

- o Iterates through the fetched posts and extracts details like title, text, ID, score, number of comments, URL, and creation date.

6. **Converting the Dictionary to a Dataframe**:

```
top_posts = pd.DataFrame(posts_dict)
```

- o Converts the dictionary into a Pandas Dataframe for easier manipulation and export.

7. **Saving the Dataframe as a CSV (Optional)**:

```
top_posts.to_csv("Top Posts.csv", index=True)
```

- o Saves the Dataframe to a CSV file named "Top Posts.csv".

Terms That Can Be Modified

1. **Subreddit Name**: Change subreddit("wallstreetbets") to another subreddit (e.g., "technology").

- o Effect: Retrieves data from the specified subreddit.

2. **Post Limit**: Adjust limit=None to a specific number (e.g., limit=100).

- o Effect: Limits the number of posts fetched.

3. **Sorting Method**: Replace top with hot, new, or rising.

   - o   Effect: Changes the type of posts fetched (e.g., most recent, trending).

Limitations

1. **API Limits**: Only retrieves up to 1,000 posts per query due to Reddit's API restrictions.

2. **Rate Limits**: Excessive requests may lead to temporary bans.