

ETL and Sentiment Analysis

DASHBOARDS

- REDDIT STOCKS API
- TTM SQUEEZE STOCKS SCANNER API

Reddit APIs

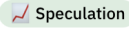
Get top 50 stocks discussed on Reddit subreddit - Wallstreetbets

```
curl -XGET 'https://tradestie.com/api/v1/apps/reddit'
```

You would get following JSON response...

```
[
  {
    "no_of_comments": 179,
    "sentiment": "Bullish",
    "sentiment_score": 0.13,
    "ticker": "GME"
  },
  {
    "no_of_comments": 37,
    "sentiment": "Bullish",
    "sentiment_score": 0.159,
    "ticker": "AMC"
  },
  {
    "no_of_comments": 17,
    "sentiment": "Bullish",
    "sentiment_score": 0.22,
    "ticker": "PLTR"
  },
  ...
]
```

↑ Posted by u/CaseFrequent8551 2 days ago




0  **Btc will drop back to 20s?**

↓ was waiting for BTC t go lower to buy but it only seems to be going up.

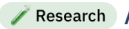
Anytime BTC or stocks go up really fast, it usually comes back down hard.

When can we expect the drop? I'm hoping blackrocks ETF gets rejected and reapplies so BTC will tank hard.

Or am I cometly wrong and the bull run has started?

 31 Comments  Share  Save ...

↑ Posted by u/Fabs_- 3 days ago

7  **An overview of a Bitcoin block (Personal summary by Fabs)**




↓ It's a too long to post here, nor fancy to read after the butchered formatting... Therefore I'll supply the link to the document.

<https://acrobat.adobe.com/id/urn:aaid:sc:EU:bf9f1598-db20-4fd9-a757-e02b326c756c>

I've (started) writing yet another (personal) summary about one of Bitcoin's many parts, and am pretty content about the summary so-far. A little bit of peer-reviewing and / or nitpicking can't hurt though.


I hope it manages to teach something new to someone!




*I'm new to this sub and hope to find more like-minded and tech-interested people here, [r/bitcoin](#) starts to become a circle-jerk...

 2 Comments  Share  Save ...

↑ Posted by u/IllusiveWebDesign 2 days ago

0 **BlackRock's Foray into Bitcoin Spot ETF - Potential Game-Changer for Crypto Miners and the Broader Market Dynamics - BTC Mox**

↓ btcmox.com/all-ey... 

 1 Comment  Share  Save ...

Assignment 3 - Social Speculation: Harnessing Reddit to Forecast Bitcoin Fluctuations

1. Get Reddit posts from cryptocurrency-related subreddits
2. Use sentiment analysis to determine the polarity and subjectivity of the posts
3. Use them to predict the price of Bitcoin
4. Create your own sentiment index using the Reddit posts

Assignment 3 Logistics

- Work in groups of 3-4 people during the class
- Submit individually
- Grading
 - ETL functions (80%): Functions must pass the testcases
 - Your sentiment index (20%): Points for how well your sentiment index is correlated with the price of Bitcoin. Your rationale for why your sentiment index is better along with evidence (e.g., data, chart) must be included in the notebook.

Data Source

- Reddit API (<https://www.reddit.com/dev/api/>)
- Alpha Vantage API (<https://www.alphavantage.co/documentation/>)

Package

- `TextBlob` for sentiment analysis (<https://textblob.readthedocs.io/en/dev/>)

Accessing Reddit API

1. Go to <https://www.reddit.com/prefs/apps>

- Sign up for a Reddit account if you don't have one (Copy **username** and **password**)

2. Click on "are you a developer? create an app..."

3. Fill in the form (randomly)

- name: random app name
- type: select "script"
- description: random description
- about url, redirect url: random url (e.g., <https://pyprogramming.net>)

4. Click on "create app"

5. Copy **client ID** and **client secret**

Accessing Reddit API

Four pieces of information are needed to generate authentication headers: `username`, `password`, `client ID`, `client secret`.

```
headers = get_authheaders(username, password, client_id, client_secret)
```

The authentication headers must be passed to the `requests.get()`

```
response = requests.get(url, headers=headers, params=params)
```


New post endpoint

- Endpoint: `https://oauth.reddit.com`
- Path: `/r/{subreddit}/new`
- Query parameters: [See here](#)

Sample API response

```
{
  "kind": "Listing",
  "data": {
    "after": "t3_16gufii",
    "dist": 25,
    "modhash": null,
    "geo_filter": null,
    "children": [
      {
        "kind": "t3",
        "data": {
          "approved_at_utc": null,
          "subreddit": "AAPL",
          "selftext": "...",
          "title": "let's take a look at $AAPL's gap fill data from 11/14/22 - 11/10/23",
          ...
        }
      },
      { ... },
      { ... },
      ...
    ],
    "before": null
  }
}
```

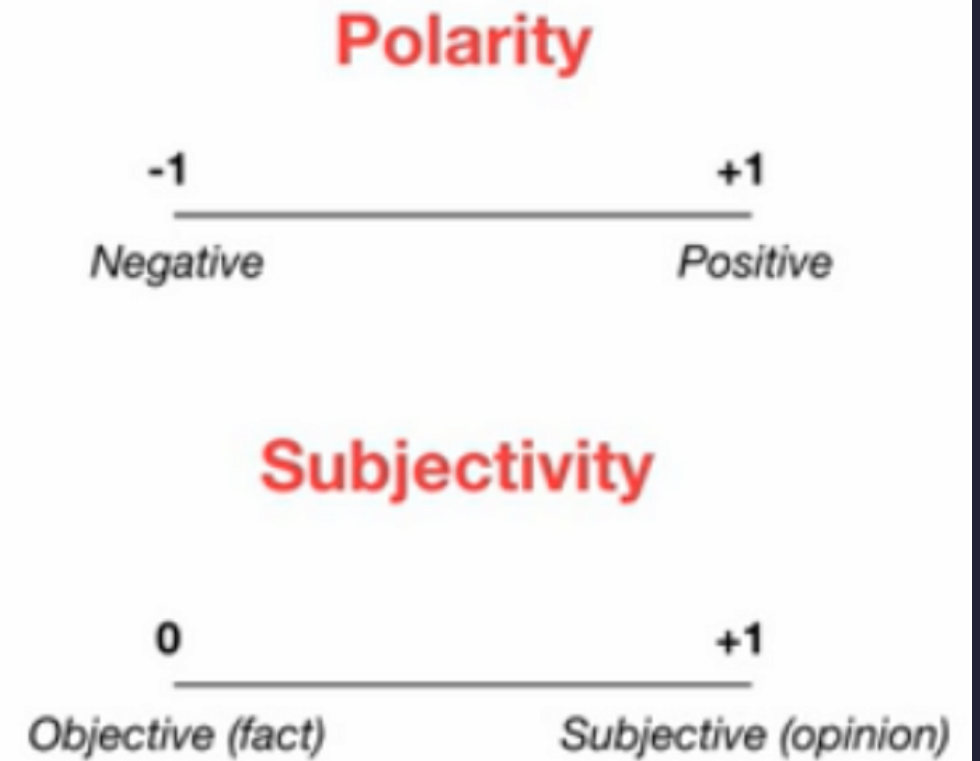
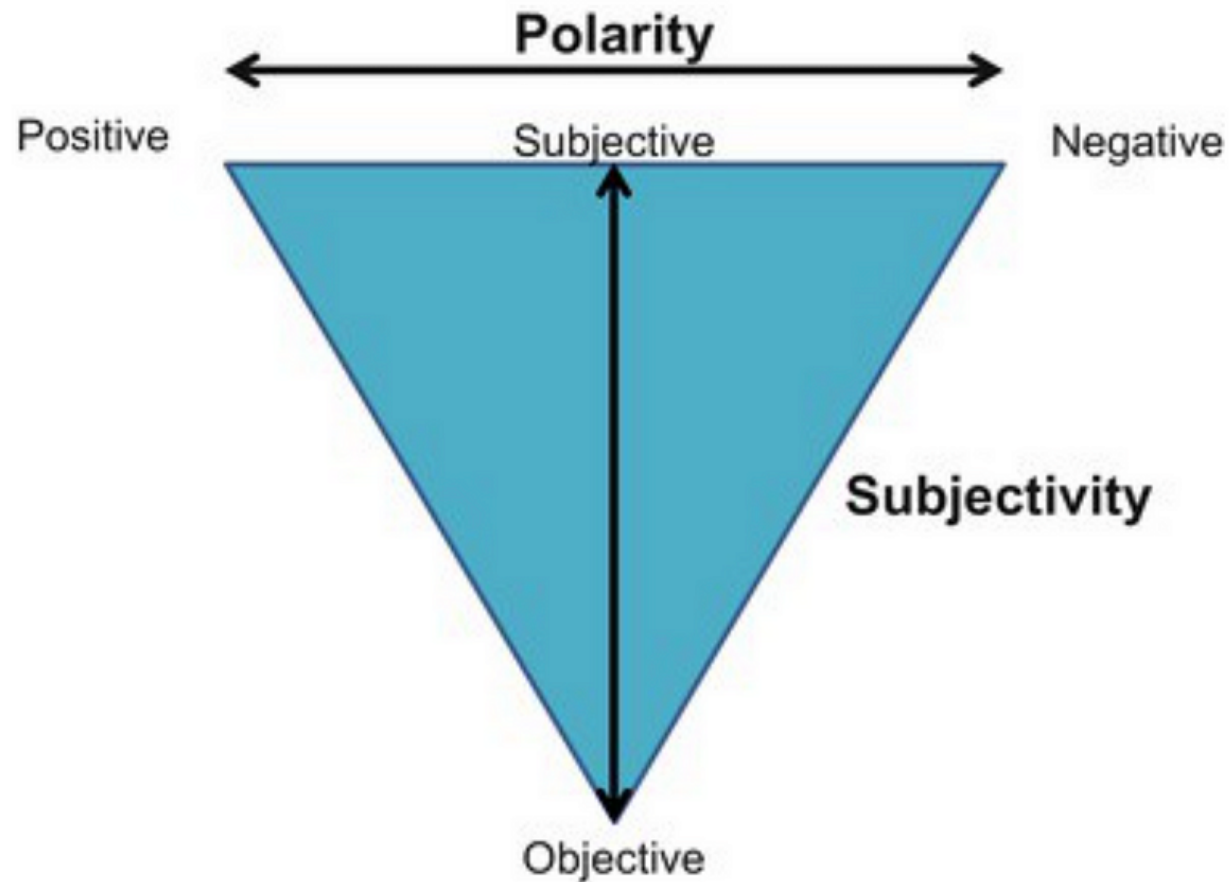
Tips for navigating through the response

- `response.json()` returns a dictionary
- `response.json().keys()` returns the keys of the dictionary
- `response.json()['data']` returns the value of the key `data`
- `response.json()['data'].keys()` returns the keys of the dictionary
- `response.json()['data']['children']` returns the value of the key `children`
- ...

Hint 1: Fetching more than limit

- Reddit API `limit` is the number of posts returned to 25 (default) or 100 (max)
- Pagination
 - `after` parameter: fullname of the last post in the current page
 - `before` parameter: fullname of the last post in the previous page

Hint 2: Sentiment analysis using TextBlob package



Hint 3: Better sentiment index

- different weights for different posts
- different subreddits
- different weights for different subreddits
- ...

Getting Started

1. Discuss with your team and decide on the subreddits you want to use. You can use the list of subreddits in the **Best Crypto Subreddits in 2023** page.
2. Plan your project, including pseudocode, database schema, and analysis and visualization ideas.
3. Divide the tasks among your team members.
4. Start coding!