```
from google.colab import drive
drive.mount('/content/drive')

import nltk
nltk.download('vader_lexicon')
```

```
Mounted at /content/drive
[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
True
```

```
import requests
import time
import json
import urllib.parse
import pandas as pd
import random
from datetime import datetime
from nltk.sentiment.vader import SentimentIntensityAnalyzer
```

```
analyzer = SentimentIntensityAnalyzer()
```

```
HEADERS = {
    "accept": "*/*",
    "accept-encoding": "gzip, deflate, br, zstd",
    "accept-language": "en-US,en;q=0.9",
    "authorization": "Bearer AuthToken", #removed the token for sharing notebook
    "content-type": "application/json",
    "cookie": ("cookie_value_of_the_session"), #removed the token for sharing notebook
    "priority": "u=1, i",
    "referer": "https://x.com/elonmusk",
    "sec-ch-ua": "\"Chromium\";v=\"134\", \"Not:A-Brand\";v=\"24\", \"Google Chrome\";v=\"134\"",
    "sec-ch-ua-mobile": "?0",
    "sec-ch-ua-platform": "\"Windows\"",
    "sec-fetch-dest": "empty",
    "sec-fetch-mode": "cors",
    "sec-fetch-site": "same-origin",
    "sec-gpc": "1",
    "user-agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/134.0.0.0 Safari/537.36",
    "x-client-transaction-id": "token-id", #removed the token for sharing notebook
    "x-csrf-token": "token-value", #removed the token for sharing notebook
    "x-twitter-active-user": "yes",
    "x-twitter-auth-type": "OAuth2Session",
    "x-twitter-client-language": "en"
}
```

```
# graphql endpoint - constant for all the tweets
BASE_ENDPOINT = "https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail"
CSV_FILENAME = "highlighted_elon_tweets_Jan1toMar12.csv"
csv_drive_path = "/content/drive/My Drive/" + CSV_FILENAME
csv_local_path = CSV_FILENAME
```

```
# building each url using the focaltweet and cursor
def build_tweet_detail_url(focalTweetId, cursor=""):
    variables = {
        "focalTweetId": focalTweetId,
        "with_rux_injections": False,
        "rankingMode": "Relevance",
        "includePromotedContent": True,
        "withCommunity": True,
        "withQuickPromoteEligibilityTweetFields": True,
        "withBirdwatchNotes": True,
        "withVoice": True,
        "cursor": cursor
    }
    variables_encoded = urllib.parse.quote(json.dumps(variables))
    # features extracted from the browser session, need for browser emulation
    features = ("&features=%7B%22rweb_video_screen_enabled%22%3Afalse%2C"
                "%22profile_label_improvements_pcf_label_in_post_enabled%22%3Atrue%2C"
                "%22rweb_tipjar_consumption_enabled%22%3Atrue%2C"
                "%22responsive_web_graphql_exclude_directive_enabled%22%3Atrue%2C"
                "%22verified_phone_label_enabled%22%3Afalse%2C"
                "%22creator_subscriptions_tweet_preview_api_enabled%22%3Atrue%2C"
                "%22responsive_web_graphql_timeline_navigation_enabled%22%3Atrue%2C"
                "%22responsive_web_graphql_skip_user_profile_image_extensions_enabled%22%3Afalse%2C"
                "%22premium_content_api_read_enabled%22%3Afalse%2C"
                "%22communities_web_enable_tweet_community_results_fetch%22%3Atrue%2C"
                "%22c9s_tweet_anatomy_moderator_badge_enabled%22%3Atrue%2C"
                "%22responsive_web_grok_analyze_button_fetch_trends_enabled%22%3Atrue%2C"
                "%22responsive_web_grok_analyze_post_followups_enabled%22%3Atrue%2C"
                "%22responsive_web_jetfuel_frame%22%3Afalse%2C"
                "%22responsive_web_grok_share_attachment_enabled%22%3Atrue%2C"
                "%22articles_preview_enabled%22%3Atrue%2C"
                "%22responsive_web_edit_tweet_api_enabled%22%3Atrue%2C"
                "%22graphql_is_translatable_rweb_tweet_is_translatable_enabled%22%3Atrue%2C"
                "%22view_counts_everywhere_api_enabled%22%3Atrue%2C"
                "%22longform_notetweets_consumption_enabled%22%3Atrue%2C"
                "%22responsive_web_twitter_article_tweet_consumption_enabled%22%3Atrue%2C"
                "%22tweet_awards_web_tipping_enabled%22%3Afalse%2C"
                "%22responsive_web_grok_analysis_button_from_backend%22%3Atrue%2C"
                "%22creator_subscriptions_quote_tweet_preview_enabled%22%3Afalse%2C"
                "%22freedom_of_speech_not_reach_fetch_enabled%22%3Atrue%2C"
                "%22standardized_nudges_misinfo%22%3Atrue%2C"
                "%22tweet_with_visibility_results_prefer_gql_limited_actions_policy_enabled%22%3Atrue%2C"
                "%22rweb_video_timestamps_enabled%22%3Atrue%2C"
                "%22longform_notetweets_rich_text_read_enabled%22%3Atrue%2C"
                "%22longform_notetweets_inline_media_enabled%22%3Atrue%2C"
                "%22responsive_web_grok_image_annotation_enabled%22%3Afalse%2C"
                "%22responsive_web_enhance_cards_enabled%22%3Afalse%7D")
    field_toggles = "&fieldToggles=%7B%22withArticlePlainText%22%3Afalse%7D"
    return f"{BASE_ENDPOINT}?variables={variables_encoded}{features}{field_toggles}"
```

```
def find_tweets(obj):
    tweets = []
    if isinstance(obj, dict):
        if "tweet_results" in obj and isinstance(obj["tweet_results"], dict) and "result" in obj["tweet_results"]:
            tweets.append(obj)
        for key, value in obj.items():
            tweets.extend(find_tweets(value))
    elif isinstance(obj, list):
        for item in obj:
            tweets.extend(find_tweets(item))
    return tweets
```

```
def extract_tweet_data(tweet_node, focal_id):
    result = tweet_node["tweet_results"]["result"]
    tweet_id = result.get("legacy", {}).get("id_str", "") or result.get("id_str", "")

    if "note_tweet" in result and result["note_tweet"]:
        note_res = result["note_tweet"].get("note_tweet_results", {}).get("result", {})
        tweet_text = note_res.get("text", "")
    else:
        tweet_text = result.get("legacy", {}).get("full_text", "")

    created_at = result.get("legacy", {}).get("created_at", "")
    retweet_count = result.get("legacy", {}).get("retweet_count", 0)
    reply_count = result.get("legacy", {}).get("reply_count", 0)
    user_legacy = result.get("core", {}).get("user_results", {}).get("result", {}).get("legacy", {})
    screen_name = user_legacy.get("screen_name", "")
    user_name = user_legacy.get("name", "")
    followers_count = user_legacy.get("followers_count", 0)

    sentiment = analyzer.polarity_scores(tweet_text)['compound']
    tweet_type = "Post" if tweet_id == focal_id else "Comment"

    # Building tweet URL
    tweet_url = f"https://x.com/{screen_name}/status/{tweet_id}" if tweet_id and screen_name else ""

    return {
        "Tweet ID": tweet_id,
        "Type": tweet_type,
        "Tweet Text": tweet_text,
        "Sentiment": sentiment,
        "Created At": created_at,
        "Screen Name": screen_name,
        "User Name": user_name,
        "Followers Count": followers_count,
        "Retweet Count": retweet_count,
        "Reply Count": reply_count,
        "Article URL": tweet_url
    }
```

```
def extract_cursor(json_data):
    def recurse(obj):
        if isinstance(obj, dict):
            if obj.get("itemType") == "TimelineTimelineCursor":
                return obj.get("value")
            for k, v in obj.items():
                result = recurse(v)
                if result:
                    return result
        elif isinstance(obj, list):
            for item in obj:
                result = recurse(item)
                if result:
                    return result
        return None
    return recurse(json_data)
```

```
def extract_tweet_url(tweet_node):
    result = tweet_node["tweet_results"]["result"]
    tweet_id = result.get("legacy", {}).get("id_str", "") or result.get("id_str", "")
    user_legacy = result.get("core", {}).get("user_results", {}).get("result", {}).get("legacy", {})
    screen_name = user_legacy.get("screen_name", "")
    created_str = result.get("legacy", {}).get("created_at", "")
    if tweet_id and screen_name:
        tweet_url = f"https://x.com/{screen_name}/status/{tweet_id}"
        return tweet_url, created_str
    return None, None
```

```
import requests
bearer_token = 'AAAAAAAAAAAAAAAAAAAAAOvizgEAAAAAqZYjhNkue0JoFODm87bfXsT5kMk%3DVVvD3hCDW9VxtPKk46W9lhsEsc3ijxmwv8BxNSvmeLCGRnGhIp'
user_id = '44196397'
url = f'https://api.twitter.com/users/{user_id}/tweets'

headers = {
    'Authorization': f'Bearer {bearer_token}'
}
response = requests.get(url, headers=headers)
if response.status_code == 200:
    tweets = response.json()
    for tweet in tweets['data']:
```

```
# Tweet IDs collected from Twitter's basic API plan and through scrapping - it is some chunk, we have more in different files
tweet_ids = [
  "1899983951306907662",
  "1899583456050598202",
  "1898369343399899218",
  "1898187378847592571",
  "1896995745280266333",
  "1896813119999209477",
  "1896158580690350303",
  "1895491688984182890",
  "1895113845447827802",
  "1894499799702843572",
  "1894173297786720718",
  "1893823851668091147",
  "1893227632906383424",
  "1892567553655865732",
  "1892413385804792307",
  "1891898474125529393",
  "1891401201993793677",
  "1890958798841389499",
  "1890612949368058118",
  "1890173376964313388",
  "1889896827513868734",
  "1889396492181426613",
  "1888989308708696177",
  "1888860423648145525",
  "1888314848477376744",
  "1887867644814020902",
  "1887535165603561505",
  "1887160798511804431",
  "1886840365329608708",
  "1886633448078475593",
  "1886129005759262964",
  "1885768824135307369",
  "1885293172458242296",
  "1885089037112791069",
  "1884470666100904015",
  "1883416606933680469",
  "1882732717139534034",
  "1881760015725523053",
  "1881547272556777647",
  "1881075557993165260",
  "1880536914690036015",
  "1880469184603431196",
  "1880159432308408468",
  "1879531470886465545",
  "1879423228978077761",
  "1878909785757880662",
  "1877402081340043512",
  "1877124097806565490",
  "1876891493014569426",
  "1876325206484038113",
  "1875931298143084755",
  "1875583572092104795",
  "1875150194909823085",
  "1874862847123165557",
  "1874625502323134570",
  "1899316339522609235",
  "1898953216852726090",
  "1897671125775138945",
  "1897538144653717996",
  "1895844665599148136",
  "1891359958333083924",
  "1878556378412798174",
  "1894744125938819481",
  "1884605023440515503",
  "1893101853090501112",
  "1882410078269329860",
  "1878030421713203225",
  "1882131458183336303",
  "1883771171155661236",
  "1883922457259876352",
  "1877800144596353068"
]
```

```
df_all = pd.DataFrame()

for tid in tweet_ids:
    print(f"Processing tweet ID: {tid}")
    current_vars = {
        "focalTweetId": tid,
        "with_rux_injections": False,
        "rankingMode": "Relevance",
        "includePromotedContent": True,
        "withCommunity": True,
        "withQuickPromoteEligibilityTweetFields": True,
        "withBirdwatchNotes": True,
        "withVoice": True,
        "cursor": ""
    }

    tweet_results = []
    while True:
        url = build_tweet_detail_url(tid, cursor=current_vars["cursor"])
        print("Fetching URL:", url)
        response = requests.get(url, headers=HEADERS)
        if response.status_code != 200:
            print(f"Failed to fetch data for tweet ID {tid}. Status code:", response.status_code)
            break

        json_data = response.json()
        tweet_nodes = find_tweets(json_data)
        print(f"Found {len(tweet_nodes)} tweet nodes on this page for tweet ID {tid}.")

        if tweet_nodes:
            for node in tweet_nodes:
                info = extract_tweet_data(node, tid)
                tweet_results.append(info)

        next_cursor = extract_cursor(json_data)
        print("Next cursor:", next_cursor)
        if not next_cursor or next_cursor == current_vars.get("cursor"):
            print(f"No new cursor found. Pagination finished for tweet ID {tid}.")
            break

        current_vars["cursor"] = next_cursor
        delay = random.randint(3, 10)
        print(f"Waiting for {delay} seconds before next fetch for tweet ID {tid}...")
        time.sleep(delay)

    if tweet_results:
        df_new = pd.DataFrame(tweet_results)
        df_all = pd.concat([df_all, df_new], ignore_index=True)
        print(f"Accumulated {len(df_new)} rows for tweet ID {tid}. Total rows so far: {len(df_all)}")
    else:
        print(f"No tweet details found for tweet ID {tid}.")

df_all.to_csv(csv_local_path, index=False)
df_all.to_csv(csv_drive_path, index=False)
print(f"Saved all tweet details to {csv_local_path} and {csv_drive_path}")
```

```
Processing tweet ID: 1899983951306907662
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 11 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: ZAAAAPBVHBmmpIe-_fri0OA0gMK9qdCfzuA01oLW2eLXruA0ioG65fjVw-A0koez9bqaw-E0pMO3ob2Zrd40tMKxjbbukN40yIe4gdKBl940vsC9mabnnN40
Waiting for 3 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 9 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: vwAAAPCwHBn2FLyD2-mNyJfeNOCBveHu2YHfNOKBsPndnPneNKSHvv364tDgNIDCvanQn87gNNaC1tni167gNIqBuuX41cPgNJKHs_W6msPhNMDFuPGcncng
Waiting for 10 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 9 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: GQEAAPD_CxwZ9h64hNuli9zJ3jTEhbulydjZ3jTyxrqZ6uS23zS8g9vpjciX3jTggb3h7tmB3zTigbD53Zz53jTwg7uJhvKM3jSkh779-uLQ4DSAwr2p0J_O
Waiting for 8 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 10 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: cwEAAPD_ZRwZ9ijmwLnZut_T3jT6h7mRpPaM3jS4hNuli9zJ3jTEhbulydjZ3jTyxrqZ6uS23zTmhryFg6264TS8g9vpjciX3jTggb3h7tmB3zTigbD53Zz5
Waiting for 5 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 10 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: zQEAAPD_vxwZ9jLqhdul373E3jTmwLnZut_T3jT6h7mRpPaM3jS4hNuli9zJ3jTEhbulydjZ3jT6hbfl2rvS4DTyxrqZ6uS23zTohrWN4euM3jTmhryFg626
Waiting for 5 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 8 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: JwIAAPH4HBn2POqF26XfvcTeNObAudm639PeNPqHuZGk9ozeNIqGsdWXgLDeNLiE26WL3MneNMSFu6XJ2NneNJqHurna5uHfNKSBvpHz6ozeNPqFt-Xau9Lg
Waiting for 5 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 9 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: gQIAAPH_JhwZ9kbqhdul373E3jTmwLnZut_T3jT6h7mRpPaM3jSKhrHVl4Cw3jS4hNuli9zJ3jTEhbulydjZ3jSah7q52ubh3zSkgb6R8-qM3jT6hbfl2rvS
Waiting for 3 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 10 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: 2wIAAPH_UxwZ9lDqhdul373E3jTmwLnZut_T3jT6h7mRpPaM3jSKhrHVl4Cw3jS4hNuli9zJ3jTEhbulydjZ3jSah7q52ubh3zSkgb6R8-qM3jT6hbfl2rvS
Waiting for 10 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 8 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: NQMAAPH_khwZ9lrqhdul373E3jTmwLnZut_T3jT6h7mRpPaM3jSKhrHVl4Cw3jS4hNuli9zJ3jTEhbulydjZ3jTMgNvZx7SN3jSah7q52ubh3zSkgb6R8-qM
Waiting for 8 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 10 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: jwMAAPH_thwZ9mTqhdul373E3jTmwLnZut_T3jTaxbXtoaaN3jT6h7mRpPaM3jSKhrHVl4Cw3jSqwLTZj9393jS4hNuli9zJ3jTEhbulydjZ3jSqx7rBlJeN
Waiting for 3 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 10 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: 6QMAAPD_HBwZ9m7qhdul373E3jTmwLnZut_T3jTaxbXtoaaN3jT6h7mRpPaM3jSKhrHVl4Cw3jSqwLTZj9393jS4hNuli9zJ3jTEhbulydjZ3jSqx7rBlJeN
Waiting for 7 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 9 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: QwQAAPD_LhwZ9njqhdul373E3jTmwLnZut_T3jTaxbXtoaaN3jSQxN3RpuWM3jTMhtvZgpmO3jT6h7mRpPaM3jSKhrHVl4Cw3jSqwLTZj9393jS4hNuli9zJ
Waiting for 7 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 10 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: ngQAAPD_ZRwZ9oIB6oXbpd-9xN405sC52brf09402sW17aGmjd40kMTd0abljN40zIbb2YKZjt40-oe5kaT2jN40ioax1ZeAsN40qsC02Y_d_d40uITbpYvc
Waiting for 10 seconds before next fetch for tweet ID 1899983951306907662...
Fetching URL: https://x.com/i/api/graphql/9ZmyKy7sj-7TWN39bvK8SA/TweetDetail?variables=%7B%22focalTweetId%22%3A%20%221899983951306907
Found 9 tweet nodes on this page for tweet ID 1899983951306907662.
Next cursor: -AQAAPD_ZRwZ9owB6oXbpd-9xN405sC52brf09402sW17aGmjd40kMTd0abljN40zIbb2YKZjt40-oe5kaT2jN40ioax1ZeAsN40qsC02Y_d_d40uITbpYvc
Waiting for 7 seconds before next fetch for tweet ID 1899983951306907662...
```