Breakdown of the variables used in tools_project till now

Variable

1.Libraries:

- **tweepy**: Used for accesing the twitter API
- **csv**: Handel CSV file operation
- **time**: Manges time-related tasks (e.g delays)
- **pandas**: Data manipulation and analysis
- **re**: Regular expression for text processing
- **MongoClient**: Interacts with MogoDB.
- **TextBlob**: performs sentiment analysis
- **SentimentIntesityAnalyzer**: Analyes sentiments using VADER
- **matplotlip.pyplot**: creates visualization
- **DateFornatter**: formats dates in plots

---

2. Data Extractions Variables:

- **ai_keywords**: a list of keywords related to AI
- **bearer_token**: Twitter API bearer token for authentication
- **client**: tweepy client object configured and their creation dates
- **file**: file object for writing tweets to ai_tweets.csv
- **writer**: CSV writer object to store tweets and their creation dates
- **tweets**: stores tweets fetched from twitter usign client.search_recent_tweets()

---

3. Sentiment Analysis Variables:

- **vader**: Initialized SentimentIntensityAnalyzer object for VADER sentiment analysis
- **analyze_vader()**: function to compute VADER sentiment score (returns compound score)
- **analyze_textblob()**: function to compute TextBlob sentiment polarity
- **vader_score**: column in df storing VADER sentiment score

- **textblob_score**: column in df storing TextBlob sentiment polarity.
- **label_sentiment()**:a function identify if sentiment score whether('Positive', 'Neutral', 'Negative')
- **Sentiment**: column in df storing sentiment labels

---

4. Comparsion between NLP, CV

- **nlp_keywords**: keywords related to NLP
- **cv_keywords**: keywords related to Computer Vision
- **nlp_tweets**: subset of df containing tweets with NLP keywords
- **cv_tweets**: subset of df containing tweets with CV keywords
- **avg_nlp_sentiment, avg_cv_sentiment**: average sentiment labels for NLP and CV tweets
- **common_nlp_sentiment, common_cv_sentiment**: most common sentiment labels for labels for NLp and CV tweets

---

5. Event analysis Variables

- **events**: dictionary definig events (e.g ChatGpt Homework surge) with date range
- **analyze_events()**: function to compute sentiment metrics for specific events
- **result**: list storing event analysis results

---

6. Temporal Analysis Variables

- **filtered_df**: subset of df with tweets created after 2025-02-15
- **daily_tweets**:aggregates tweet count and average sentiment by day
- **weekly_tweets**: aggregates tweet count and average sentiment by week

---

Visualization suggestion

برحتك طبعا الي تعمليه بس انا بقترح عشان انا الي عامله الانلسيس و كمان نقاشات التيم

1. Sentiment distribution:

Plot: Bar chart/ pie chart

Purpose: Show the proportion of Positive, negative and neutral

2. Sentiment Over time:

Plot: line chart

Purpose: Track sentiment changes daily/ weekly

3. NLP vs computer vision sentiment comparison:

Plot: grouped bar chart

Purpose: compare average senitment

4. sentiment during key events:

Plot vertical line annotation + line plot

Purpose: Highlight how sentiment shifted during events

5. most frequent AI key word (Optionally)

Plot: word cloud

Purpose: visualize which AI key words appeared most often in tweets

6. tweets volume over time:

Plot: area plot

Purpose: show tweet volume trends (daily/weekly)

7. sentiment by keyword

Plot: Heatmap

Purpose: compare sentiment scores across different AI keywords

8. Box plot of sentiment score

Plot: box plt

Purpose: compare distributions of sentiment scores for NLP vs CV