How do sentiment and emotional tone in English and Arabic Twitter reactions to the Gaza War evolve between January and August 2022?

DISCLAIMER:

This Notebook only looks to explore this topic from an explorative standpoint without showing support or contempt for one side. Although I hold my personal views, this notebook looks to widen my understanding of this topic from an informative standpoint.

Background:

Between January and August 2022, the Gaza War understandably generated an intense global reaction on Twitter. Tweets in English and Arabic conveyed various opinions—from solidarity and calls for cease-fires to justification of military action from the IDF. Understanding how these sentiments and stances ("pro-Palestine," "neutral," and "pro-Israel") evolve over time and differ by language can inform researchers, journalists, and policymakers about public opinion dynamics during humanitarian crises(although the effects are varying).

Motivation:

First, there isn't a widely recognized, comprehensive sentiment tracker for this conflict. Second, my own exposure is nonetheless skewed. As a result, I've realized that my view of this conflict is filtered through a narrow lens. What I really wanted was to step outside of my personal "echo chamber" and see how people around the world—across different regions, languages, and political backgrounds—are feeling about the Gaza War. By analyzing a broader dataset of tweets, I can attempt to capture the spectrum of reactions, from solidarity and empathy to frustration and dissent, rather than just the perspectives I'm used to consuming.

Approach & Thesis.

I preprocessed and explored two parallel datasets (English, Arabic), labeled 10K tweets via zero-shot classification, then trained a fast classifier (sentence-embedding + logistic

regression). Then trained with a deep learning model (DistilBert) that was optimized for speed and accuracy.

Approach:

Data Description:

Source: Two Excel exports of tweets collected January–August 2022 containing keywords "Gaza" or "Israel."
Fields: Tweet text, timestamp, content, retweet and like counts, username.
Languages: English (≈75 K tweets) and Arabic (≈120K tweets).
Strengths: Temporal coverage across eight months, bi-lingual perspective, engagement metadata (likes/retweets)
Limitations: No native sentiment/emotion labels, potential noise/duplication (bots, retweets)

Preprocessing

Column pruning: Removed Tweet ID, source, username, and raw hashtags to focus on content.
Sampling: Drew 10,000 random tweets per language to balance scale and speed.
Cleaning: Early text-normalization methods decreased model performance and were omitted; raw text (minus URLs/mentions) was used for embedding.
Tokenization: Employed HuggingFace tokenizers matched to each model (AraBERT tokenizer for Arabic).

Methodology

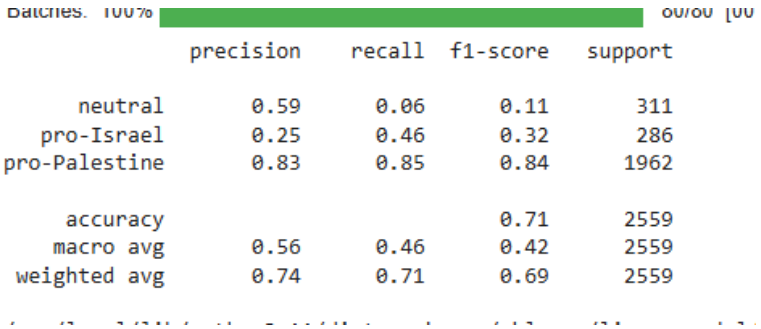
1. Zero-Shot Inference:

   Used a distilled NLI model (`typeform/distilbert-base-uncased-mnli`) to label samples of 10 K tweets per language with "pro-Palestine," "pro-Israel," or "neutral." **This technique proved to be the best chance at labeling sentiments across tweets. The other Steps were other ways I tried labeling data, but ultimately didn't employ them.**

2. Supervised Classification

- Generated 384-dim embeddings via a multilingual Sentence-Transformer (`paraphrase-multilingual-MiniLM-L12-v2`).
- Trained a Logistic Regression (weighted for class imbalance) on those embeddings.
- Tuned hyperparameters (L1 vs. L2 penalty, regularization strength, different class weights & solvers)

Final Model Performance



	precision	recall	f1-score	support
neutral	0.59	0.06	0.11	311
pro-Israel	0.25	0.46	0.32	286
pro-Palestine	0.83	0.85	0.84	1962
accuracy			0.71	2559
macro avg	0.56	0.46	0.42	2559
weighted avg	0.74	0.71	0.69	2559

(English Results)

3. Transformer Fine-Tuning
    - As a high-accuracy option, fine-tuned DistilBERT for 3 epochs on the same labeled subset, optimizing macro-F1.

Final DL Model Performance

	precision	recall	f1-score	support
neutral	0.10	0.50	0.17	2
pro-Israel	0.21	1.00	0.34	7
pro-Palestine	0.93	0.28	0.43	50
accuracy			0.37	59
macro avg	0.41	0.59	0.31	59
weighted avg	0.82	0.37	0.41	59

(Arabic Results, English were similar)

Evaluation Strategy

Primary focus: Pattern discovery—how stance counts change monthly—rather than maximizing held-out accuracy due to complexity of sentiment analysis.
Metrics: Precision, Recall, F-1 to assess balance across classes.
Visualization: Line plots of monthly stance counts, bar charts of overall distribution, and word clouds per stance.

Analysis & Results

Sentiment Over Time:

- Both languages experienced similar sentiment expression during the eight month period(1)
- Although “pro-Israel” sentiment among Arabic tweets remained consistent, there was a spike in th eight month for English Speakers (1)
- For English tweets, “pro-Israel” & “nuetral” had nearly identical behaivor during the eight month period(1)

- **Emotion Trends**:
  - Among English tweets, wordings were pretty similar, broadly speaking, for each sentiment(4)

- **Narrative Peaks**:
  - Both English and Arabic Speakers had proportionally unequal tweets heavily for “Pro-Palestine”(2)
  - Emotions & Tweet shifts tracked closely with on-ground events (April & August) (3)
  - There were quite a few mentions of God among Arabic tweets compared to English tweets(5)

Next Steps

Interpretation

1. Total tweets per month & tweet count for “Pro-Palestine” during this period followed similar trends, which is worth digging into. One Possibility is supporters for Palestine are more vocal. Another may be social pressure which affects both sides. Supporters of Israel may receive harsher backlash than vice versa. And supporters of Palestine receive positive support. This could work both ways. Regardless, this pattern is quite

interesting and I wish I had time to confidently explore the depth of this pattern alone.(1) (3)

2. One thing that immediately caught my attention was the spike in tweets in April & August. I looked online at events that happened during this time period and saw a correlation(ceasefire, missile strikes, humanitarian aid). I am not claiming causation, but was nonetheless interesting to see on ground events align with people's twitter behavior. (1)
3. **Pro-Israel**: Peaks (~20 %) immediately after cease-fire announcements, indicating reactive solidarity posts(August 2022).(1)
4. **Pro-Palestine**: Dominant (~60–70 %) throughout, with Arabic tweets showing slightly higher solidarity (~75 %) during humanitarian campaigns.(1)
5. One thing I wish I could also explore with confidence is the presence of certain themes between tweets of English & ARabic. Unfortunately the word cloud for Arabic tweets didn't work. But From what I saw there were persistent mentions of "God" or "Allah" among Arabic tweets.(5)

Limitations

1. No fine-tuning; possible misclassification especially in sarcasm or regional dialects
2. Bias toward publicly vocal users (vs. silent observers)
3. Language segmentation may have noise
4. ZeroShot Classifier was mostly accurate, Unfortunately I can't manually verify the stance of Arabic tweets as thoroughly as I did with English tweets. This may have caused the abnormally higher neutral tweets.

Conclusion

- Simple techniques are effective especially when you are constrained(i.e speed, class imbalances)
- Social media is not just reactive but emotionally and politically shaped by linguistic and cultural context
- Deep Learning Models are simple in implementation, but can be incredibly tricky when fine tuning, which is why I had to ditch using DL technique for classification

My biggest takeaway: A data science problem is an intensive, time consuming process and the amount of stories held within data feel infinite. This was a good reminder to maintain a consistent & manageable scope when handling a data science problem going into the future. Another thing, which I have heard often, domain knowledge is so important and means more, in my opinion for what it is worth, than your skill set of data science tools. Last thing is this was definitely one of the cooler projects I

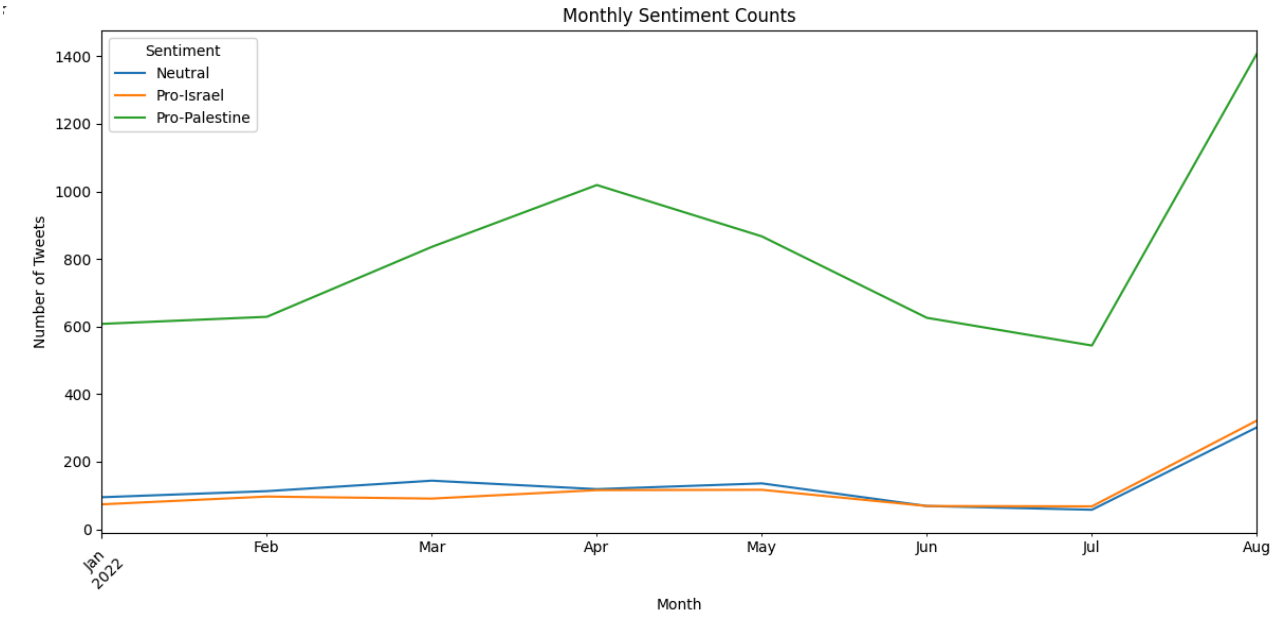
have done and learned quite a bit, and was genuinely invested and spent ~ 20 hours.

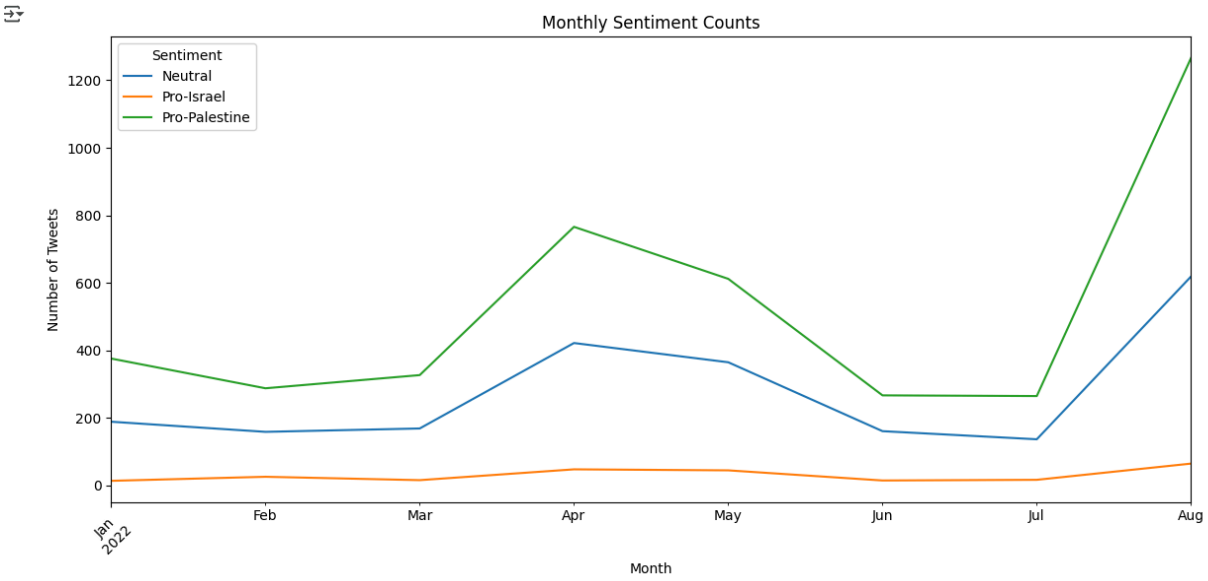
Next Steps

- Utilize location and map sentiment across locations & explore other geospatial data visualization
- Fine tune a transformer model and apply it to the expanded dataset to capture more information.
- Perform in depth data validation to supplement classifications with a deep learning model to ensure highest accuracy of tweet classes

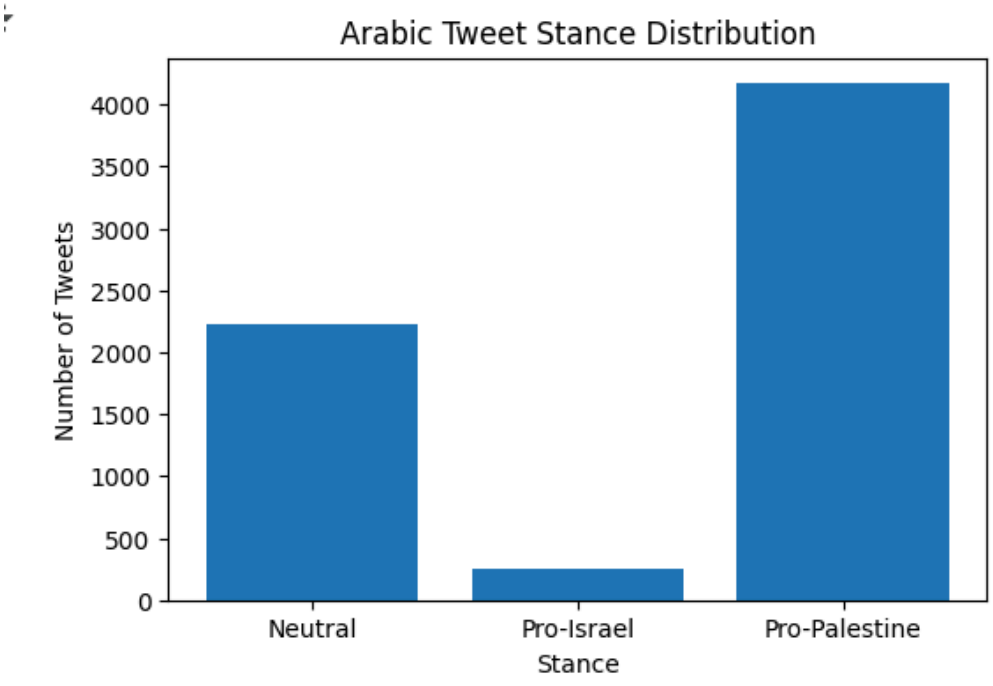
Appendix:

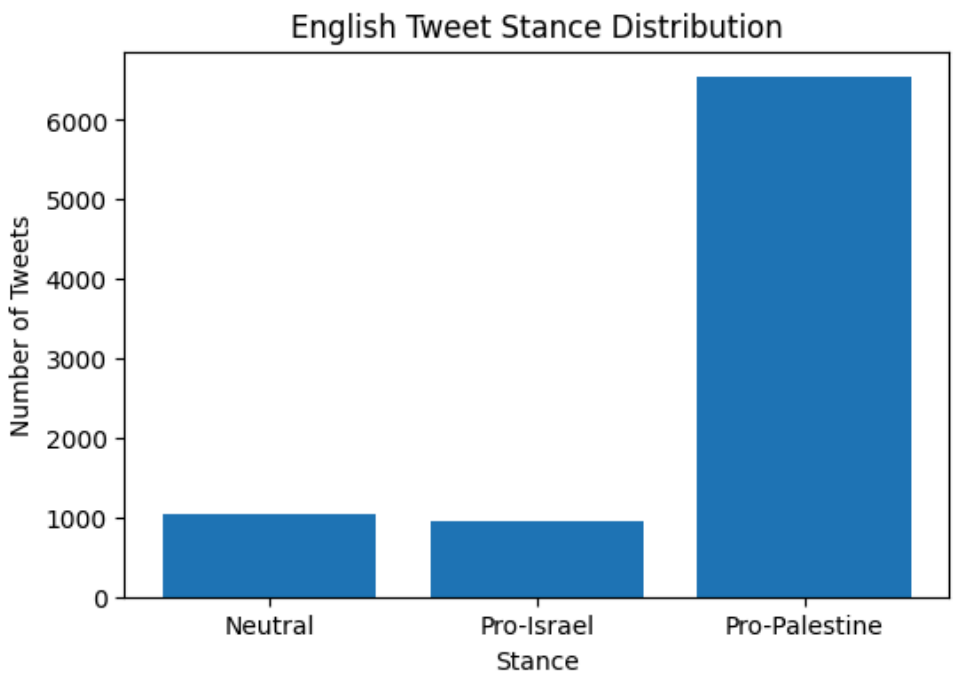
(1). Monthly Sentiment for English and Arabic, in respective order



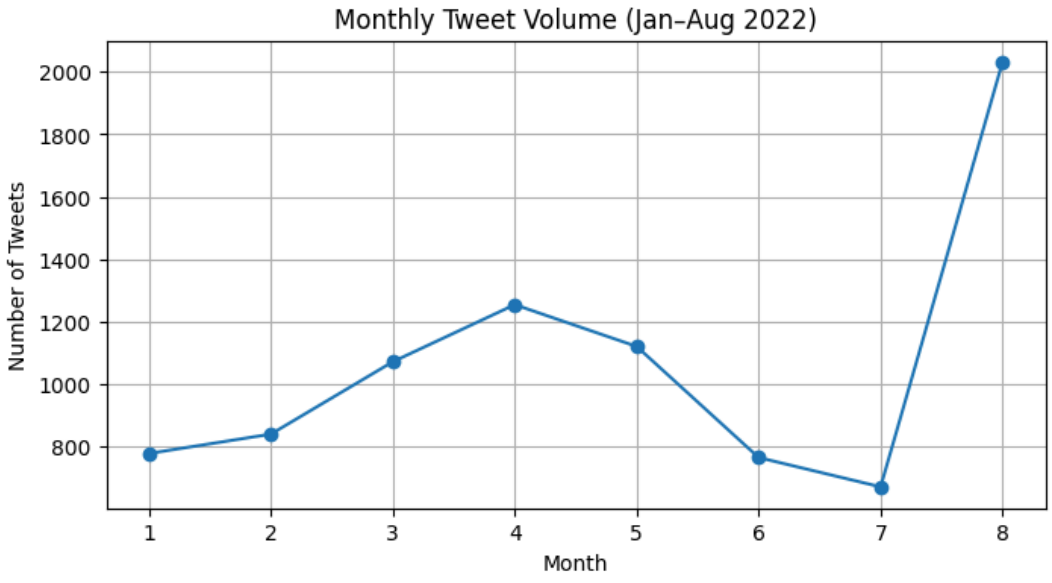


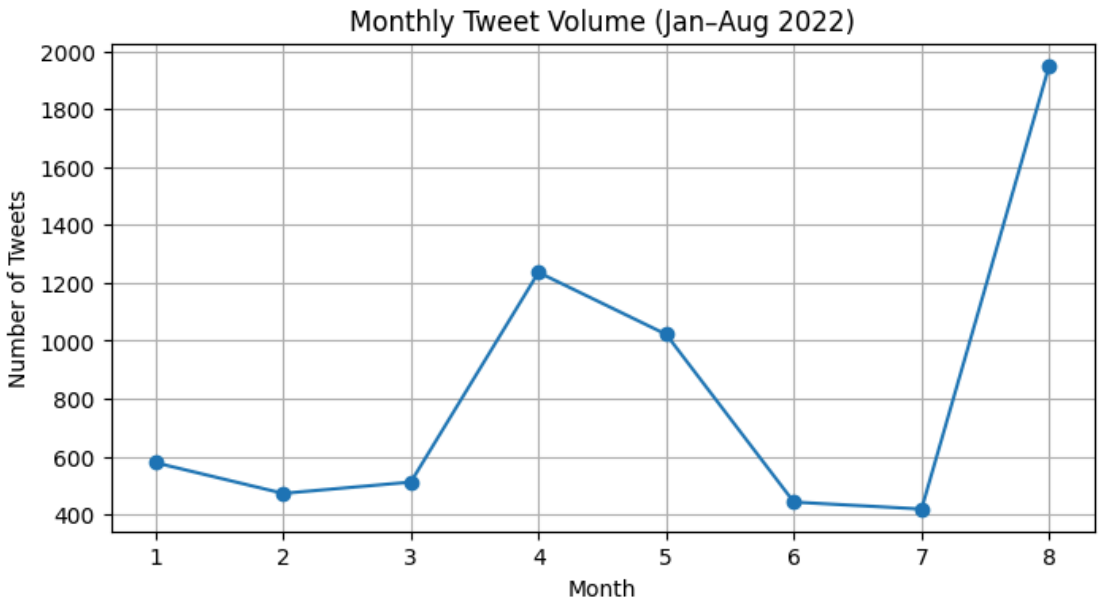
(2) Tweet Distribution



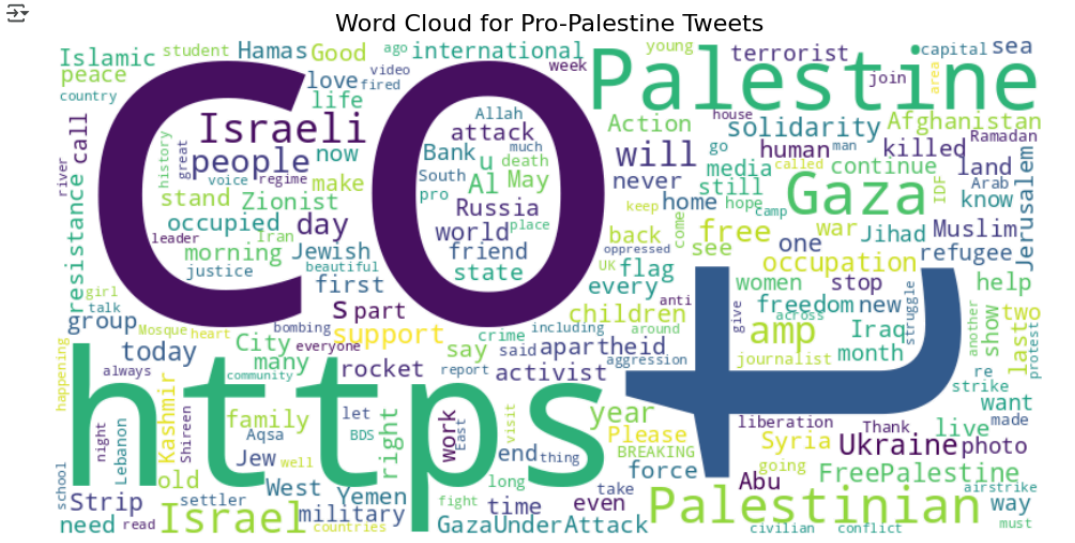


(3) Tweet Count per month (English, Arabic in that order)





(4) Word clouds for english tweets



(5) Most popular tweets Arabic vs English

English:

```
Bella Hadid has said that she is willing to lose jobs as a model to continue her advocacy for her homeland, Palestine.
Praying for the people of gaza. 🙏❤️
Gaza in pictures, 5th of August, 2022.
A Palestinian kid from Gaza during his father's funeral.
Dome intercepted 96% of all rockets fired into Israel from Gaza, saving thousands of lives of Israelis. God bless her. 🙏
```

Arabic:

Arabic ⇄ English

غزة تحت القصف الآن، اللهم احفظها بحفظك، واجعلها في ضمانك وأمانك، ولطفك ورعايتك

ghazat taht alqasf alan, allahuma ahfazha bihifzika, wajealha fi damanik wa'amanika, walutfik warieayatik

Gaza is under bombardment now, O God, protect it with your protection, and make it under your care and security, and your kindness and care

Open in Google Translate • Feedback

Arabic ⇄ English

ولندا وحدها استقبلت 2.5 مليون لاجئ من #أوكرانيا في 6 أسابيع، ولم نشاهد خيمة تغرق في المطر، ولا طفلا حافيا يغوص في الوحل، ولا امرأة تبحث في حاويات القمامة عما يسد رمق أولادها.\nتلك صور صارت حكرا على أمة العرب دون غيرها.\n#سوريا #فلسطين #العراق #اليمن

walanda wahdaha aistaqbalat 2.5 milyun laji min #'uwkrania fi 6 asabiea, walam nushahid khaymatan taghraq fi almutari, wala tiflan hafian yaghus fi alwahli, wala aimra'at tabhath fi hawiat alqumamat eamaa yasudu ramq 'awladiha.\ntilk sur sarat hikran ealaa 'umat alearab dun ghiriha.\n#suria #filastin #aleiraq #aliman

And the Netherlands alone received 2.5 million refugees from #Ukraine in 6 weeks, and we didn't see a tent flooding in the rain, a barefoot child drowning in the mud, or a woman searching through garbage containers for food to feed her children. These are images that have become the exclusive preserve of the Arab nation. #Syria #Palestine #Iraq #Yemen

Arabic ⇄ English

غزة تحت القصف
الآن، اللهم احفظها بحفظك،
واجعلها في ضمانك وأمانك،
ولطفك ورعايتك

ghazat taht alqasf alan, allahuma
ahfazha bihifzika, wajealha fi damanik
wa'amanika, walutfik warieayatik

Gaza is under bombardment now, O God, protect it with your protection, and make it under your care and security, and your kindness and care

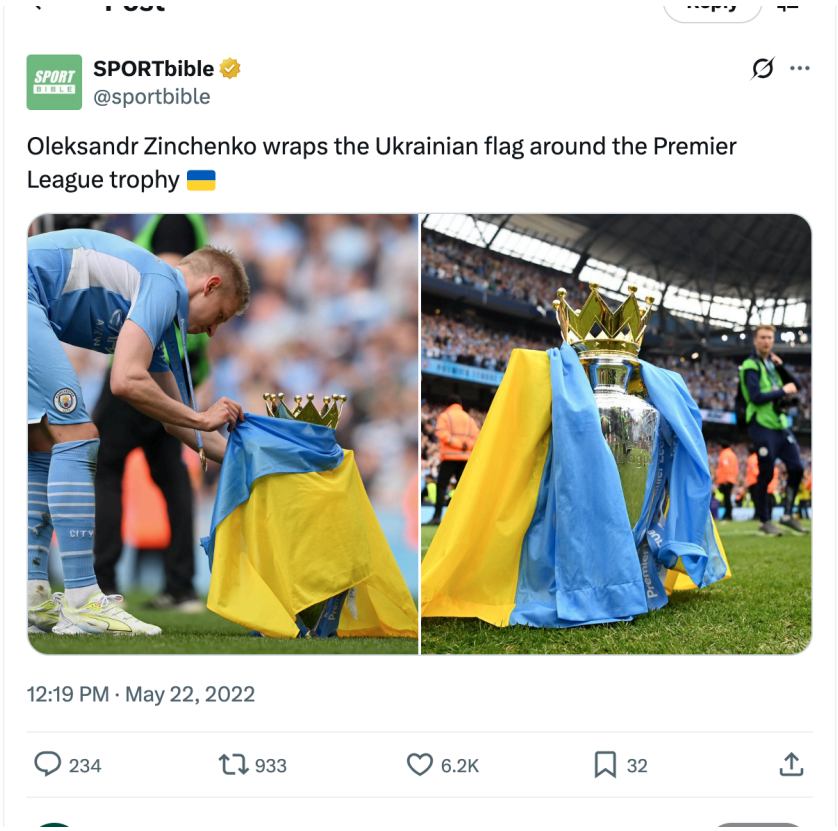
Open in Google Translate • Feedback

5533
هذا الطفل الجميل اسمه ماجد
أبو عودة، يبلغ من العمر 6
سنوات، أصبح اليوم أصغر
حافظ للقرآن الكريم في قطاع
غزة بل في فلسطين كلها
❤️ \nبمثل هذه النماذج
نفتخر، وبمثل هذه النماذج
ننتصر !!
https://t.co/2coDI4p
YrI

5533 hadha altifl aljamiil aismuh majid
abu eawdata, yablugh min aleumr 6
sanawati, 'asbah alyawm 'asghar
hafiz lilquran alkarim fi qitae ghazat
bal fi filastin kuliha ❤️ \nbmthl hadhih
alnamadhij naftakhru, wabimithl
hadhih alnamadhij nantasir !!
https://t.co/2coDI4pYrI

5533 This beautiful child's name is Majed Abu Odeh, he is 6 years old, today he became the youngest memorizer of the Holy Quran in the Gaza Strip, and in all of Palestine ❤️ We are proud of such models, and with such models we are victorious!!
https://t.co/2coDI4pYrI

Additional Findings:



The Most liked tweet(English) showed support for Palestine, and the photo was from the premier league which is the biggest soccer league in the world. This moment was captured after Man City won the Premier league. I remember watching this moment on TV, so it was surprising to stumble upon this tweet.