

A photograph showing a group of people, including several children, walking through a vast area of rubble and debris. In the background, partially destroyed buildings stand under a clear blue sky with scattered clouds. One person is carrying a large orange box. The scene depicts the severe impact of conflict on civilian life.

Text Analysis of the Conflict in Gaza Reveals the Civilian Impact

Kevin Linares

(https://github.com/klinares/airwars_scraping_project)

Israel-Hamas Conflict, 1 Year Later

- **Background:** 10/7/2023 Hamas surprise attack on Israel.
 - Estimated Palestinian dead ~44,000 (60% children & women).
- **Problem:** Hamas Ministry of Health (MoH) casualties, unconfirmed.
 - **Claim:** “Women & children disproportionately killed” [UN-OHCHR](#)
- **Research Question:** To what extent can open-source data be used to identify patterns in the targeting of Palestinian civilians in Gaza?



Methodology: Analyzing Civilian Casualties

Data Source: [Airwars](#) tracks casualties from conflicts.

1. Web Scrape: ~ 800 incidents (~9,000 deaths), & store in SQLite database.

2. Reverse Geo-coding: Query [Nominatim API](#) with coordinates, return site type (school, hospital).

3. Sentiment Analysis: Derive emotional tone from assessments.

- [DistilRoBERTa-base](#), classifies text into Ekman's 6 basic emotions.

4. Clustering Analysis: Cluster incidents geographically to discern civilian casualty patterns.

The screenshot shows a detailed report from Airwars. At the top, there's a navigation bar with links like 'Conflicts', 'Archive', 'Research', 'Investigations', 'News', 'About', and a language switcher. The main content area has a header 'Airwars Civilian Casualties' with a blurred background image. Below this, a section titled 'Airwars assessment' contains text about an incident on May 12, 2024, where at least 39 civilians were killed or injured. It mentions the Al-Zaytoun neighborhood in Gaza City and names several victims. A 'Summary' table provides key details: Strike status (Likely strike), Strike type (Airstrike), Civilian infrastructure (IDP or refugee camp, Residential building), Civilian harm reported (Yes), Civilians reported killed (39, including 4-5 children, 17 women, 21 men), Civilians reported injured (12-24), Cause of injury / death (Heavy weapons and explosive munitions), Airwars civilian harm grading (Fair), and Suspected belligerent (Israeli Military). Below the summary is a 'Sources (12)' section listing various social media accounts and their posts related to the incident.

Summary		
Strike status	Likely strike	
Strike type	Airstrike	
Civilian infrastructure	IDP or refugee camp, Residential building	
Civilian harm reported	Yes	
Civilians reported killed	39 (4-5 children, 17 women, 21 men)	
Civilians reported injured	12-24	
Cause of injury / death	Heavy weapons and explosive munitions	
Airwars civilian harm grading	Fair ⓘ	
Suspected belligerent	Israeli Military	

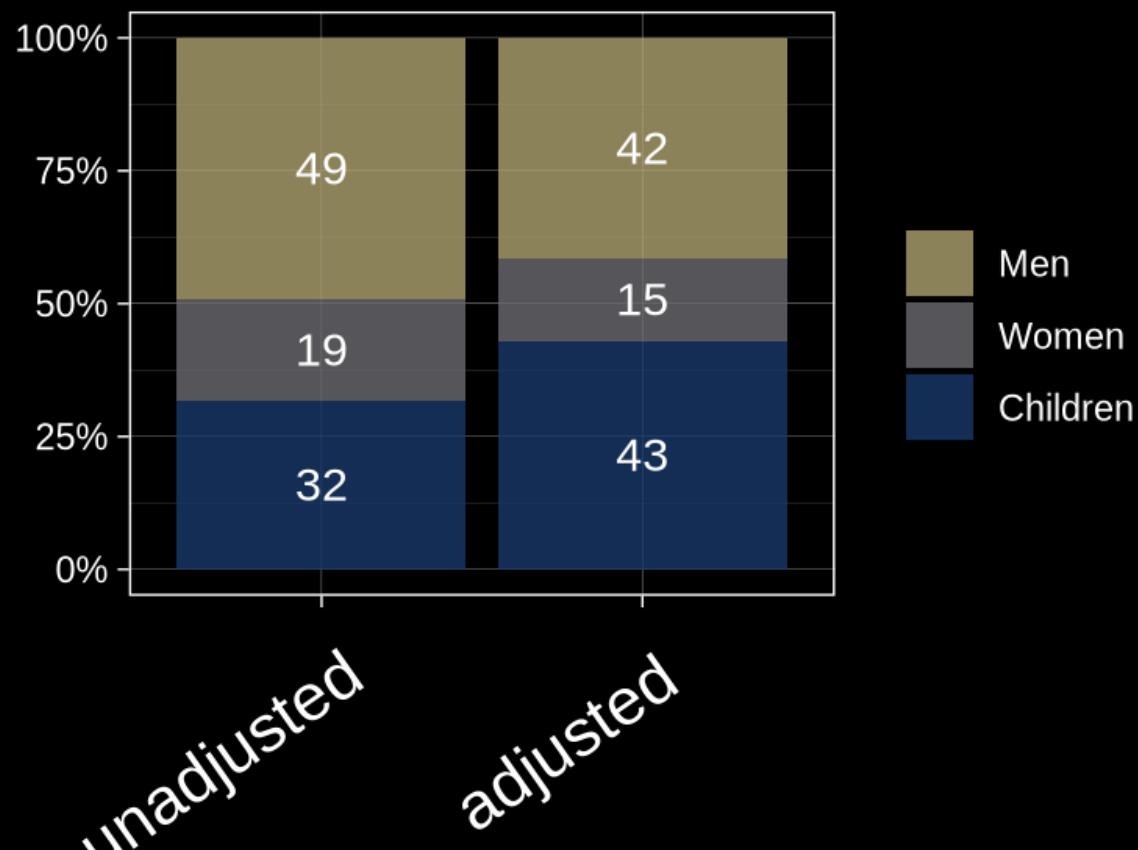
Sources (12) [↑ collapse]

- @alhawyahchannel TWITTER, ARABIC ⓘ Archive
- Tasneen.hussam FACEBOOK, ARABIC ⓘ Archive
- Samar Hanouna FACEBOOK, ARABIC ⓘ Archive
- Asaad Mohammad FACEBOOK, ARABIC ⓘ Archive
- Alzaytoon2017 FACEBOOK, ARABIC ⓘ Archive
- Rami A A Daloul FACEBOOK, ARABIC ⓘ Archive
- Samar Hanouna FACEBOOK, ARABIC ⓘ Archive
- Abu Jihad Daloul FACEBOOK, ARABIC ⓘ Archive
- Amly billah FACEBOOK, ARABIC ⓘ Archive
- Abu Jihad Daloul FACEBOOK, ARABIC ⓘ Archive
- Abdulla Daloul FACEBOOK, ARABIC ⓘ Archive
- Abu Jihad Daloul FACEBOOK, ARABIC ⓘ Archive

Results: Casualty Rate, Sentiment Analysis

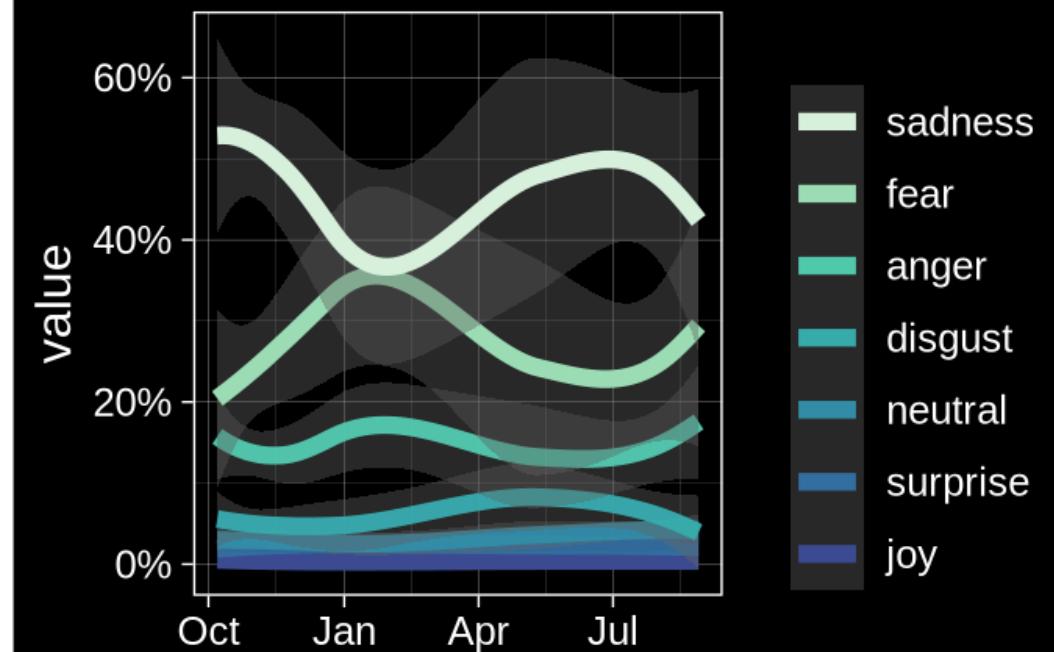
Children more likely to be killed, after adjusting for population size.

Current Casualty Rates



Negative emotional tones more prominent, vary across time.

Sentiment Analysis 2023-24



Results: Hierarchical Clustering

Three clusters were adequate for these data.
Cluster 3, incidents w/ highest child & women average deaths

Clusters 1 2 3

