Description of the Data

The primary data source for use in the analysis is from The Humanitarian Data Exchange (HDX) of The United Nations Office for the Coordination of Humanitarian Affairs (OCHA)[8]. It is described as follows:

"This dataset includes any reported incident (referred to as attack) that affected heatlhcare in Ukraine between 24 February 2022 and 04 March 2023."

The dataset consists of 1 x XLSX file:

- 2022-2023 Ukraine Attacks on Health Care Incident Data.xlsx

This is a detailed dataset from an official source, though it only contains data regarding attacks on health care workers and facilities. It covers the complete period from the beginning of the invasion on 24 February 2022 to 04 March 2023, encompassing all 4 campaign phases launched by the Russians thus far. It contains data on 785 discrete attacks. Included within the dataset is a description of the event in the Ukrainian language, coordinates for the location of the event, and the yield in terms of effect type and severity of the harmful act. Additionally there is attribution to an actor and also the type of weapon used. A description of the numeric variables follows:

	eventSindID	category HealthFacilitiesDamagedDestroyed	Infrastructure: Hospital	Health Transport	Infrastructure: Other	Health Building	Information	Everyday Activities	Health Facility	Number of Attacks on Health Facilities Reporting Destruction	Number of Attacks on Health Facilities Reporting Damaged	Health Workers Killed	Health Workers Kidnapped	Health Workers Arrested	Health Workers Injured	Health Workers Assaulted
count	785	785	785	785	785	785	785	785	785	785	785	785	785	785	785	785
mean	34662.56178	0.407643	0.365605	0.054777	0.243312	0.043312	0.008917	0.015287	0.035669	0.078981	0.365605	0.105732	0.078981	0.031847	0.086624	0.003822
std	1691.583912	0.49171	0.481906	0.22769	0.429356	0.203689	0.094069	0.122769	0.185581	0.317643	0.539358	0.599983	1.519055	0.721573	0.473951	0.107075
min	30995	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25%	33364	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50%	34717	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
75%	36069	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0
max	37668	1	1	1	1	1	1	1	1	5	5	12	42	20	7	3

Additional datasets regarding Russian personnel and equipment losses were gathered from Dr. Petro Ivanuik, a Lviv, Ukraine-based Data Scientist who shares data related to the conflict on Kaggle[6]. It is described as follows:

"This is a dataset that describes Equipment Losses, Death Toll, Military Wounded, and Prisoner of War of Russians in 2022 Ukraine Russia War."

The dataset used consists of 2 x CSV files:

- russia_losses_equipment.csv
- russia_losses_personnel.csv

These contained a cumulative daily total representing the physical costs of the war on the Russian MoD. For personnel losses, the dataset contains information from 25 February 2022 to 26 March 2023, representing a total of 170,550 Russians Killed in Action (KIA). A description of the numeric variables follows:

	day	personnel	POW	daily_increase
count	395	395	62	394
mean	199	66318.78734	386.387097	425.761421
std	114.170924	44805.05097	131.440363	305.628221
min	2	2800	0	0
25%	100.5	31000	389	200
50%	199	52250	421	350
75%	297.5	97985	474.5	620
max	396	170550	496	3160

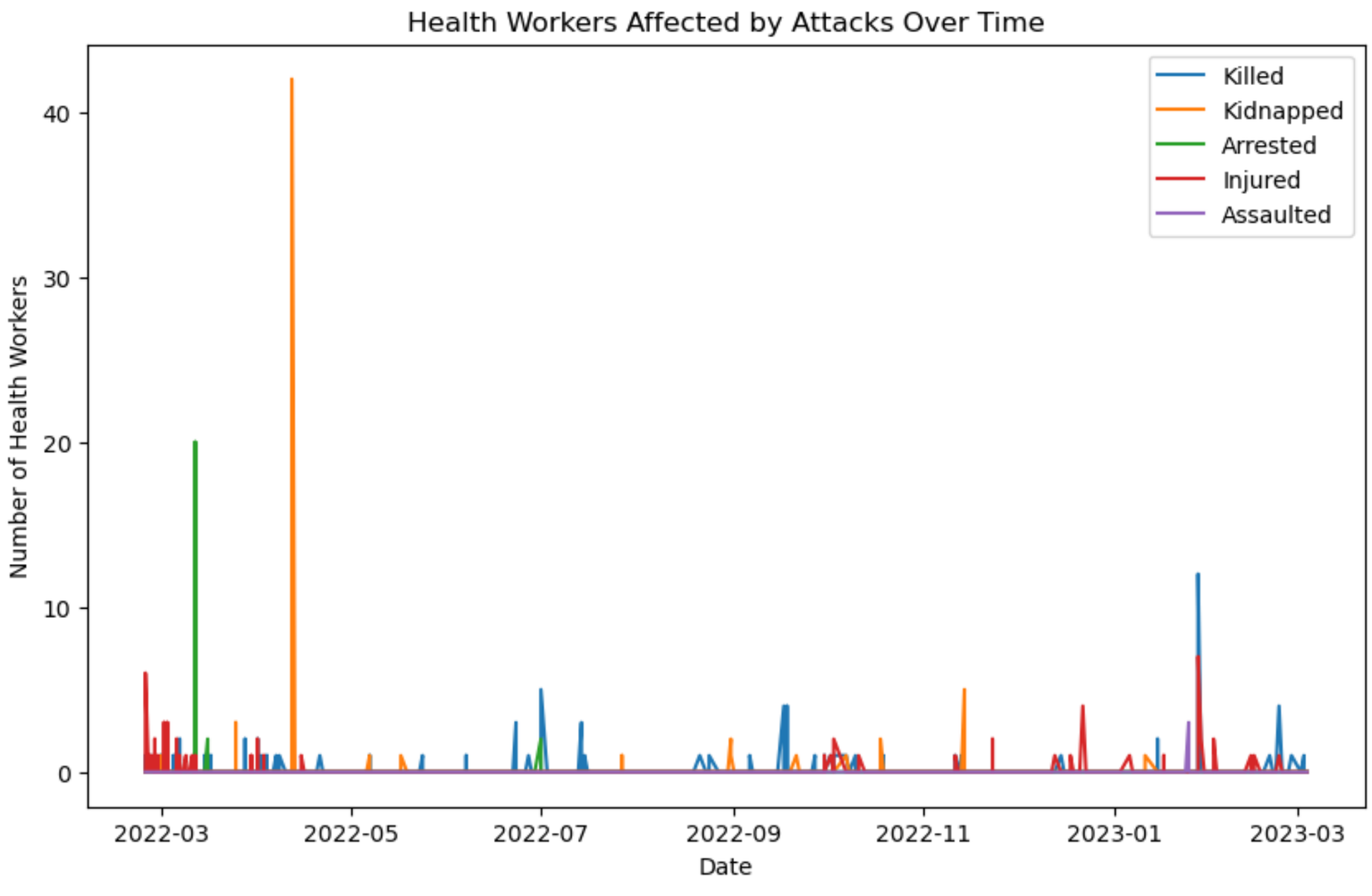
For the equipment losses, the dataset contains information from 25 February 2022 to 26 March 2023, representing various types of equipment lost or destroyed[6]. A description of the numeric variables follows:

	day	aircraft	helicopter	tank	APC	field artillery	MRL	military auto	fuel tank	drone	naval ship	anti-aircraft warfare	special equipment	mobile SRBM system
count	395	395	395	395	395	395	395	65	65	395	395	395	376	36
mean	199	232.921519	211.967089	2098.126582	4457.972152	1294.458228	301.726582	1047.507692	69.323077	1045.091139	13.840506	151.805063	121.255319	3.944444
std	114.170924	64.794571	61.818396	967.923865	1705.429199	743.553903	131.007173	466.16206	7.545917	682.146583	4.208691	69.647736	70.660833	0.333333
min	2	10	7	80	516	49	4	100	60	0	2	0	10	2
25%	100.5	210	175	1371.5	3372.5	677.5	207	600	60	537.5	13	95	55	4
50%	199	239	212	2136	4584	1259	311	1178	73	898	15	162	125	4
75%	297.5	281	264	2986	5960.5	1947.5	410	1437	76	1648.5	16	211	178	4
max	396	305	291	3595	6947	2631	522	1701	76	2216	18	277	282	4

Exploratory Analysis

As described in the methodology section of this paper, the creation and comparison of time plots is a valid and valuable technique for exploring data and identifying trends and patterns over time. This can help to uncover hidden relationships between variables and inform further analysis and modeling efforts. Visualization of single variables is provided for perpetrator of attack and weapon used in the form of a pie chart. An interactive map depicting the locations of the attacks along with date, description, and destructive yield (quantified effects on healthcare workers and infrastructure) was generated, and a sample view is provided.

Healthcare Workers

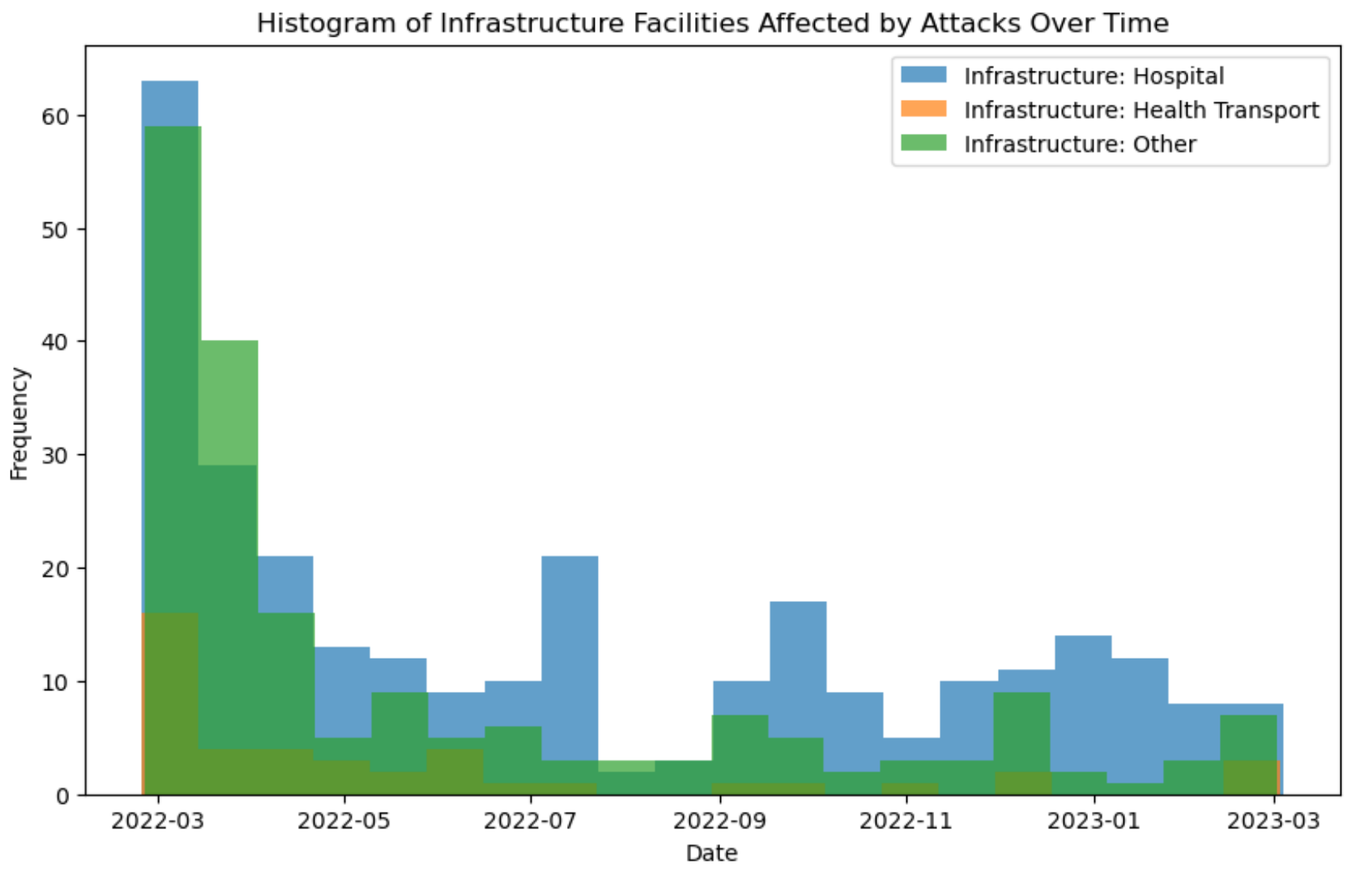


Plotting the crimes perpetrated against Healthcare workers over the time period analyzed paints a revealing picture. We see daily killing during the first four weeks following the invasion, followed by a lull in all types of harm. Category highs for kidnap and arrest also occur during this early period, though a review of the event description reveals that arbitrary categorization may have taken place regarding these two variables.

The 20 arrests on 12 March 2022 were the result of a regional hospital being taken-over by Russian Forces. Doctors and surrounding citizens were forced inside and prohibited to leave under threat of death. Russian Forces used the hospital as a military base and moved numerous weapons on to the grounds. The high for kidnapping was on 12 April 2022 when 42 doctors at a military hospital were held as prisoners of war by the Russians.

Surprisingly, the highest death toll waged against Healthcare workers during this period came at the hands of Ukrainian forces, on 28 January 2023. A hospital that had been taken over and used by Russian forces as a military hospital was hit by Ukrainian Forces rockets, killing 14 and injuring 24 hospital patients and medical staff. This was not the same location that was commandeered on 12 March 2022.

Healthcare Facilities

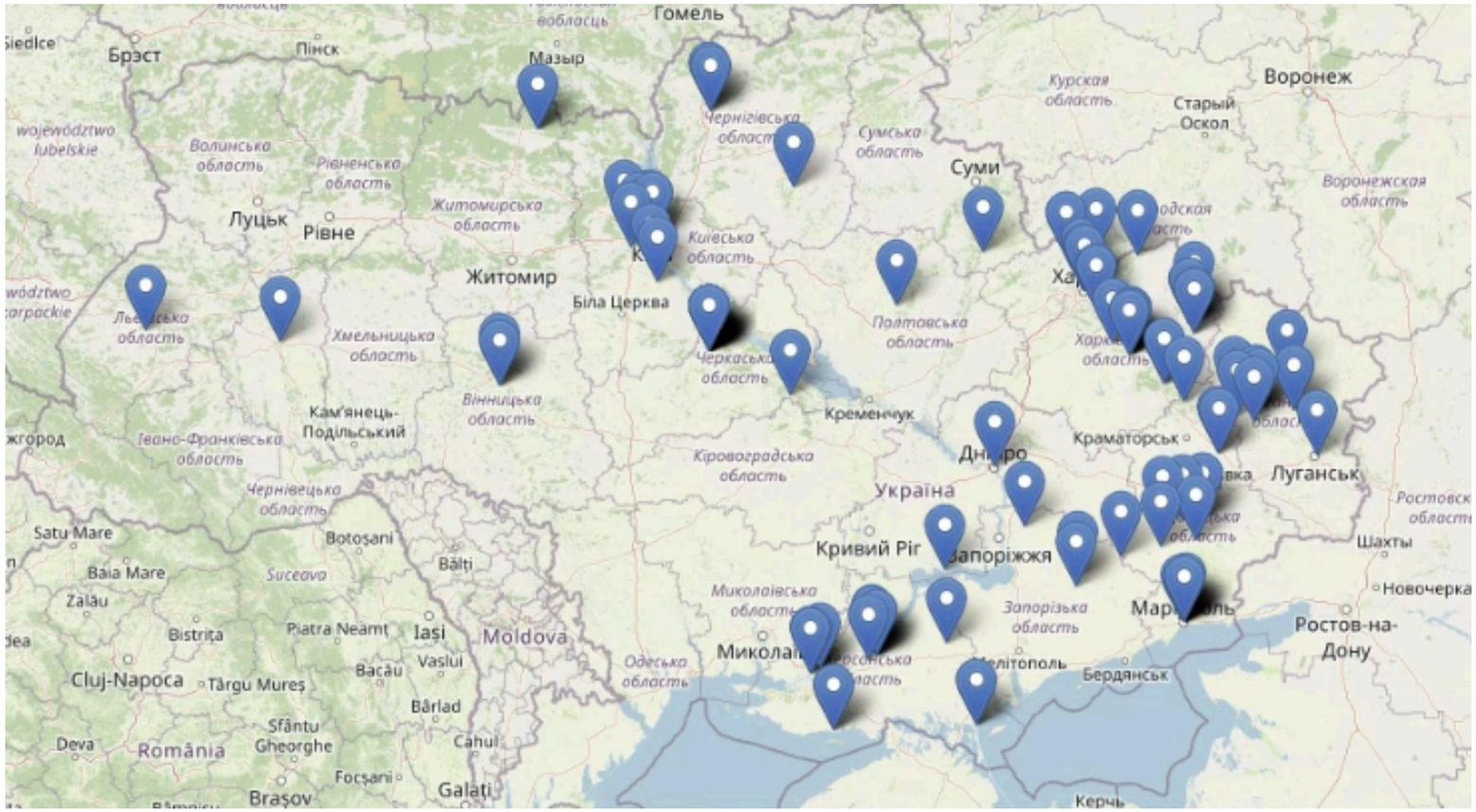


Constructing a histogram of healthcare facilities affected by attacks roughly mirrors the effects on workers, though with a few surprises.

While one may assume that attacks made against Healthcare workers on the road in the course of their duties, and in their villages may occur at higher frequency than attacks on infrastructure facilities, that is not true. In the early phase of the conflict especially, attacks on Healthcare infrastructure were very high. Had those facilities been fully staffed, the early war death toll would have been much higher for healthcare workers. Many of the healthcare facilities early in the war were destroyed after critical patients evacuated to western Ukraine, and have not been reopened to date.

While the attacks on workers tend to level off and then spike, it is also interesting to note that attacks on facilities maintain two week averages in the double digits throughout most of the conflict, with only a lull in attacks against Hospitals in August of 2022. While some of the effect is due to the daily line chart totals for workers vs. histogram visualization for facilities, there are only 3 events resulting in double digit effects for healthcare workers, each in a different category.

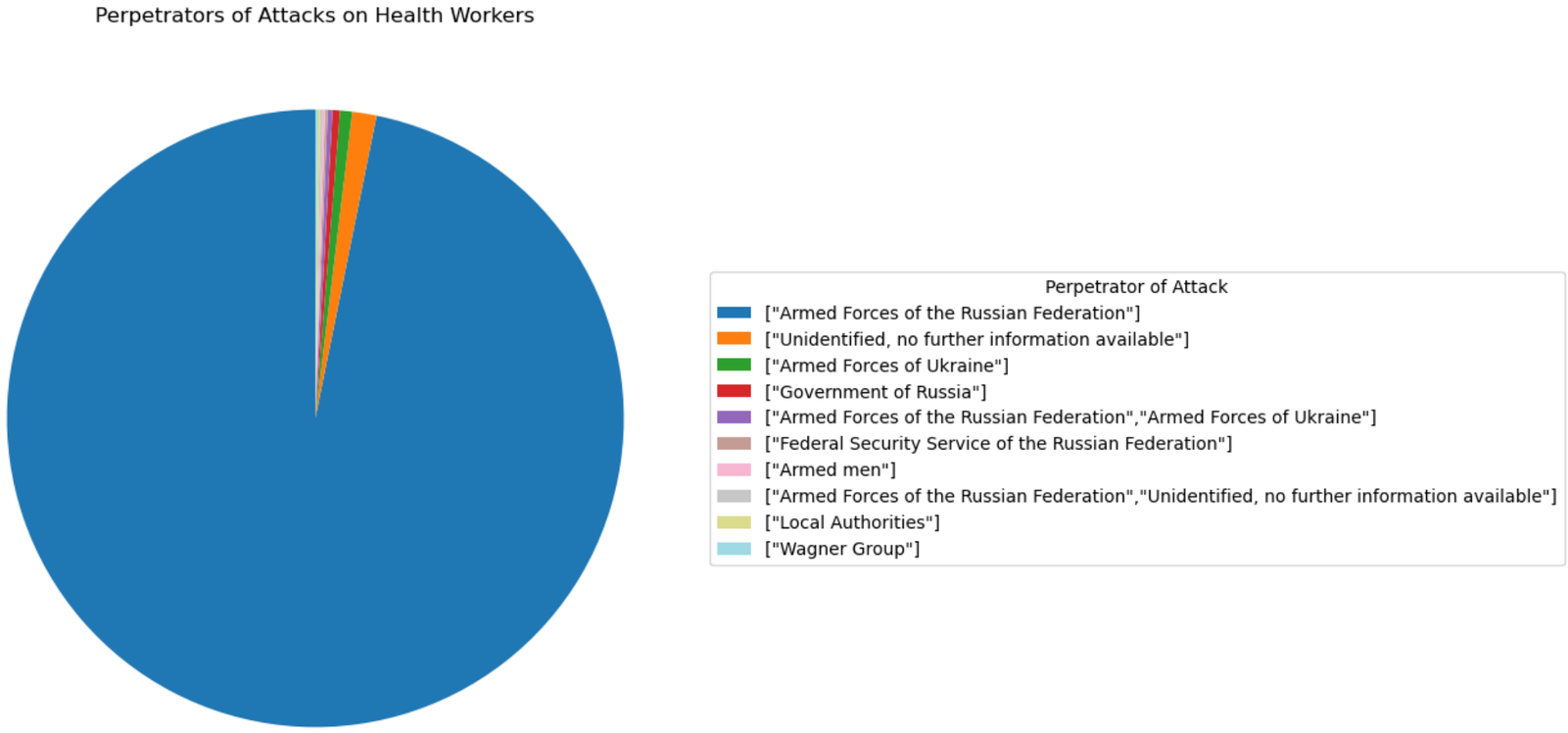
Locations of Attacks on Healthcare Workers and Infrastructure



A sample of the interactive visualization generated is depicted at left. Attacks on healthcare workers and infrastructure are plotted using their latitude and longitude coordinates contained within the dataset. As a tooltip (not pictured), mousing over a plot will reveal the date of the event, the translated, English language description of the event, and the yield in terms of quantitative effect (Example, 4 x Healthcare Workers Wounded, 2 x Healthcare Workers Wounded Killed).

This is a powerful tool for visualization, and shows the early events in the north and west eventually shifting to territories of the east in Donetsk and Lughansk Oblasts.

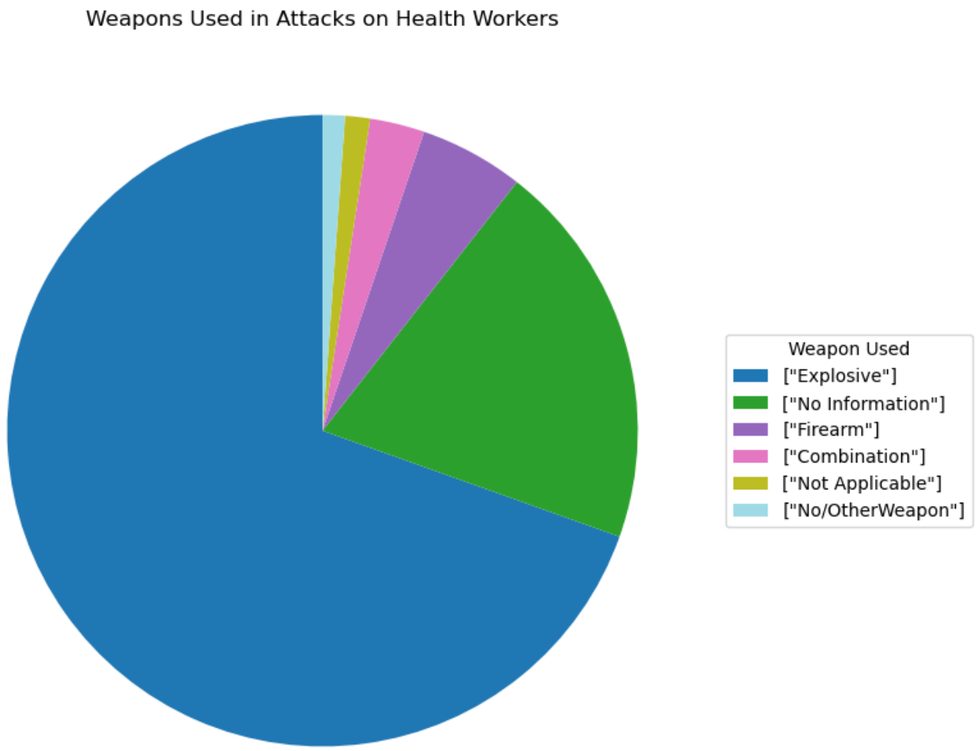
Perpetrators of Attacks



While the primary perpetrator, by far, of attacks against Healthcare workers is the Armed Forces of the Russian Federation and associated entities, it may come as a surprise to note that at total of 5 attacks, representing just over half a percentage of attacks on healthcare workers were perpetrated by the Armed Forces of Ukraine. The Private Military Corporation, Wagner group, known for it's brutality and lack of adherence to the laws of war, is attributed with only a single attack on healthcare workers.

Perpetrator of Attack	Total	Percentage
["Armed Forces of the Russian Federation"]	760	96.8%
["Unidentified, no further information available"]	10	1.3%
["Armed Forces of Ukraine"]	5	0.6%
["Government of Russia"]	3	0.4%
["Armed Forces of the Russian Federation","Armed Forces of Ukraine"]	2	0.3%
["Federal Security Service of the Russian Federation"]	1	0.1%
["Armed men"]	1	0.1%
["Armed Forces of the Russian Federation","Unidentified, nfia"]	1	0.1%
["Local Authorities"]	1	0.1%
["Wagner Group"]	1	0.1%

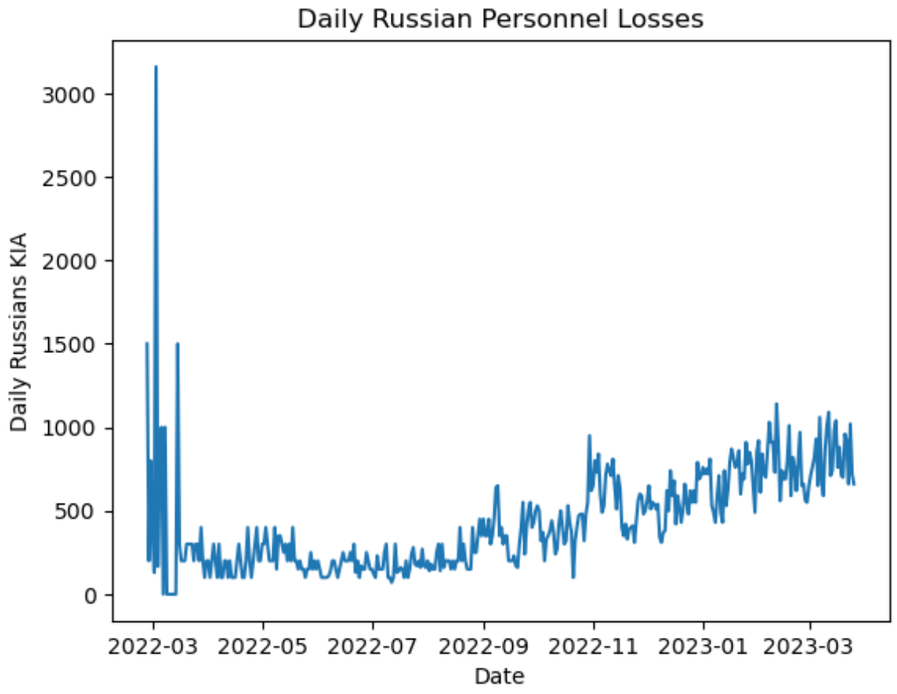
Weapons used in Attacks

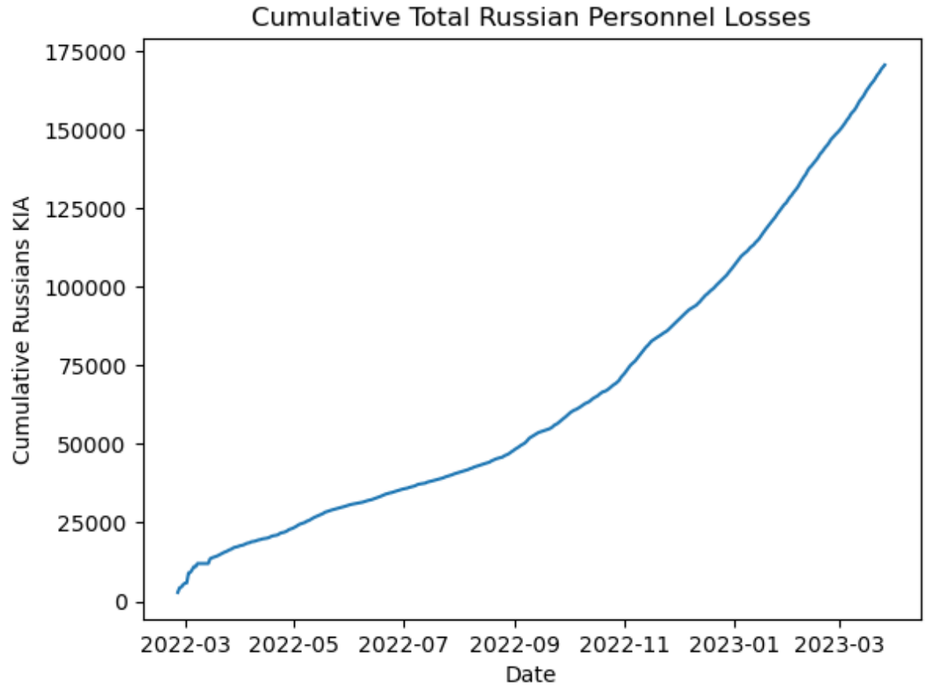


The vast majority of attacks on Health workers were carried out using explosives of some type- a category which includes missiles, rockets, artillery, and mortars, as well as land mines. This may be interpreted as a case of unintentional targeting, though in the early phase of the war Kaliber cruise missiles did specifically target major hospitals as well as other civilian infrastructure. In the case of firearms, it is highly likely that targeting was deliberate and intentional.

Weapon Used	Total	Percentage
["Explosive"]	546	69.6%
["No Information"]	156	19.9%
["Firearm"]	42	5.4%
["Combination"]	22	2.8%
["Not Applicable"]	10	1.3%
["No/Other Weapon"]	9	1.1%

Russian Personnel Losses

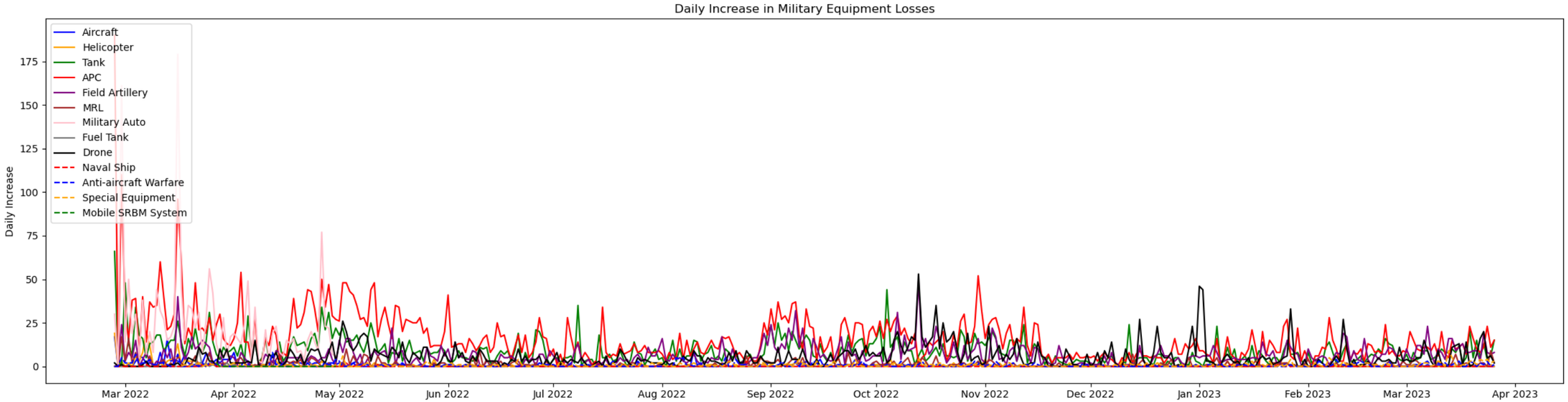




Transitioning to exploration of the datasets concerned with Russian battlefield losses, we can immediately see that the day-to-day trend of Russians Killed in Action (KIA) does not conform to the attacks on healthcare workers. After losing men in the thousands during the initial invasion, daily KIA totals have steadily climbed to reach an average of over 500 KIA per day for the past 90 days of the sampled date range. Viewing the cumulative total for Russian KIA, we can see a steep trend for the first month of the invasion that eventually becomes more gradual throughout the summer of 2022 during the Kherson Counteroffensive. Beginning in September of 2022 with the dramatically successful Kharkiv Counteroffensive, we see the curve become more steep as rate and scale improves for the Ukrainian Forces.

The hypothesis that attacks on Ukrainian healthcare workers and infrastructure is driven by Russian battlefield losses is therefore disproven. It even appears striking that increased operational tempo of the battlefield did not generate more Article 19 violations, considering that the Russian war machine currently relies on poorly trained conscripts and PMC Wagner forces, the vast majority of whom come from Russian jails and prisons.

Russian Equipment Losses



Another interesting aspect of the data revealed by this exploration exercise, specifically with respect to equipment losses, is that in the year 2023, equipment losses of all types have been relatively low despite the high Russian KIA rate. This is due to a lack of available equipment to field, combined with high KIA rates of the conscript soldiers. We see high losses during the initial invasion, as we have come to expect in examining the other datasets, but after that we see a large loss of Armored Personnel Carriers (APCs) before and during the Kharkiv Counteroffensive, and then relatively low rates of equipment loss. The one exception to this is the high rate of drone interception and destruction that accompanied the January drone attacks on the Ukrainian power grid.