

STAT-612 Final Project: Investigating Trends in Violent Conflict

Erik Lovece, Peter Matarrese, and Alif Zabidi

Introduction

This project focuses on violent conflict throughout the world. Within the last 30 years, we've seen a number of major conflicts (e.g. the Gulf War, the Rwandan Genocide) as well as sustained asymmetric violence (e.g. a rise in international terrorism, the War on Drugs and the collateral damage from that). However, with our general awareness of this violence comes some assumptions; mainly, that the world seems to be a much more violent place. We decided to ask some questions and make some initial hypotheses to test these assumptions. Do note that while we had already found our dataset, we had brainstormed our questions and hypotheses before getting a detailed look at our data.

To that end, we chose to investigate data collected by the Uppsala Conflict Data Program (UCDP), run out of the University of Uppsala in Sweden. While they have multiple datasets that span all the way back to 1946, we chose to utilize their Georeferenced Event Dataset (GED) version 19.1 (as well as integrating their Syria-specific dataset). This dataset provides records sourced from news-wire and other local news agencies around the world that include the perpetrator and victim(s) of violent acts, locational data (including latitude and longitude of the nearest reliably referenced location where the violence occurred), the start and end date of the event, the source article and publisher, and low, high, and best estimates for the number of deaths from the event (among other variables provided, 50 in total). From this dataset, we hope to look at trends in various regions of the world for violence perpetrated by and against different groups of peoples.

For a second level of analysis we downloaded some datasets from the World Bank Databank pertaining to cellphone and internet users in a given country, as well as national population estimates. These datasets were merged with our main UCDP data using `dplyr` joins by the `year` and `country_id` variables (there was admittedly some data loss as resulting from the join, but most of the failures for converting country code types accounted for small island nations in the Caribbean and the south Pacific, which don't appear often in our UCDP data). One major case of data loss (which is discussed later) stems from the lack of recent data from the World Bank concerning Syria. As stated above, our team proceeded to come up with questions and hypotheses we thought we could use this data to answer.

Initial hypotheses

Question 1: “What kinds of conflicts have become more prevalent worldwide in the past three decades?”

As stated above, we believe that there is a general sense that today's world is a more violent place. We're not sure if this is just a learned assumption conditioned into us by sensationalized news stories, or if it's true. Our first question asks what kinds of conflicts have become more prevalent in the last three decades. Specifically, we

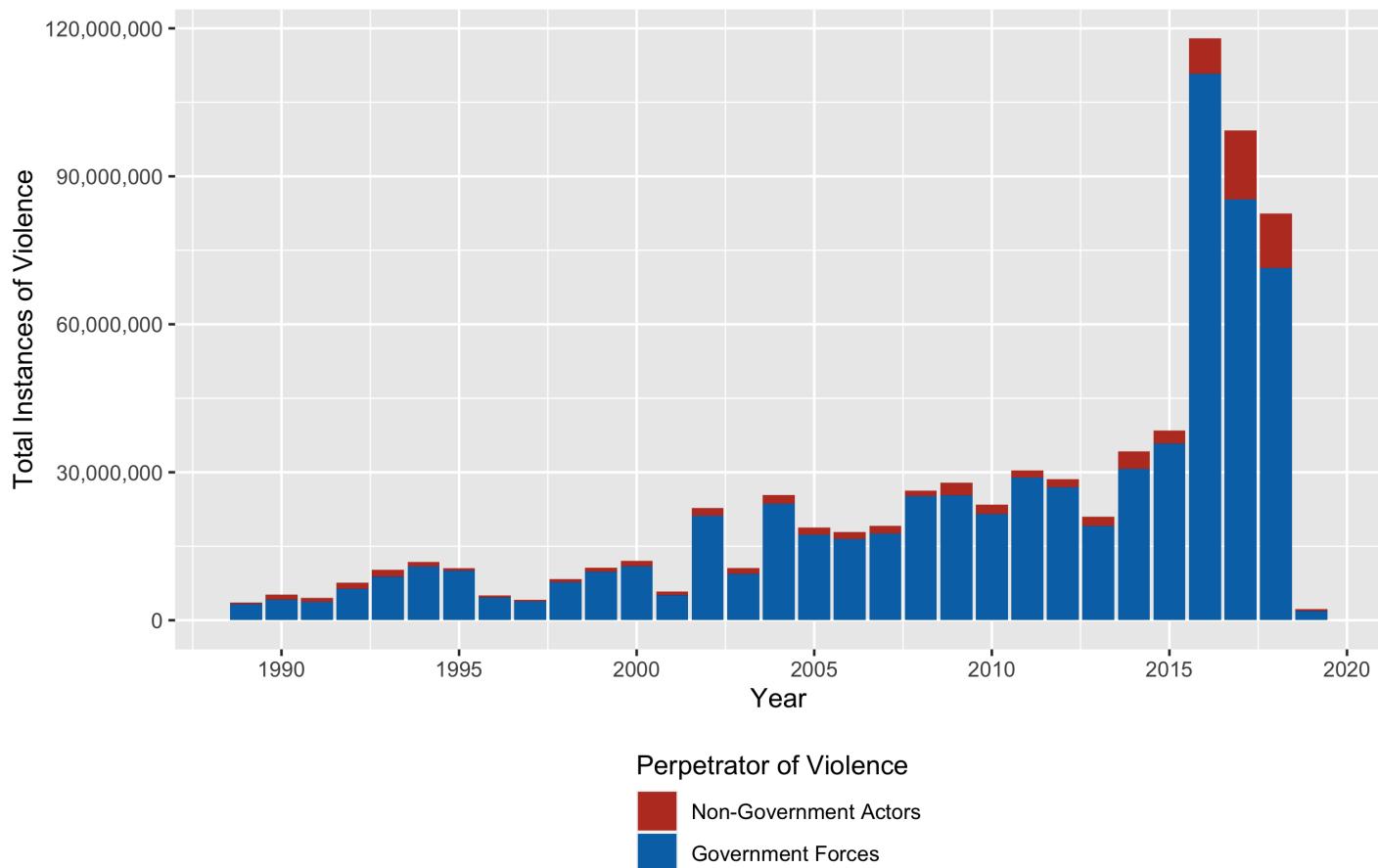
wanted to differentiate between violence perpetrated by government forces (whether against other governments or non-state actors), and violence conducted by non-state actors. Below is our initial hypothesis.

Hypothesis 1: “There has been a reduction in violence perpetrated by government forces, but an increase in violent conflicts by and between non-state actors in recent years.”

To investigate this, we needed to classify the events in our data by who the perpetrator was. We then wanted to count the total number of reports in the dataset for each year, and visualize who the aggressor was. We created a binary variable which separates the data into two groups: those events where the aggressor was government forces, and those which the aggressor was specifically non-government (it doesn’t matter if they were working in conjunction with government forces over the course of the whole conflict, since we are looking at the data at the level of individual instances of violence).

Who is perpetrating violence in the world?

The number of total instances of violence seems to have increased over the last 30 years.



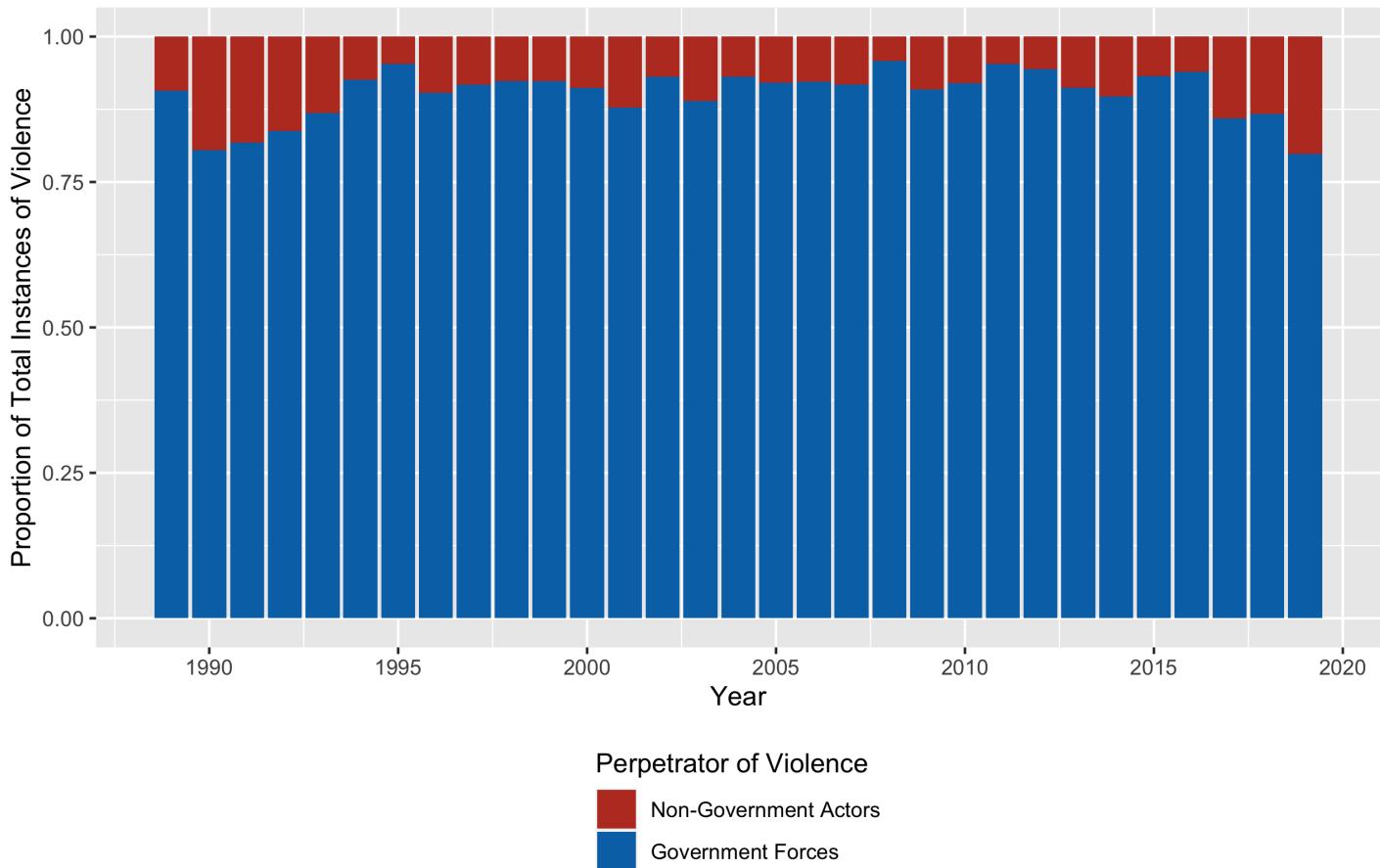
Source: UCDP dataset

This graph gives us a startling picture. Over the last thirty-one years, there has been a rise in reported instances of violence, with an incredible spike in the last five years (the drop-off in 2018 onward is discussed later). This immediate picture shows that the first half of our hypothesis is unlikely to be true; government-based violence has been on the rise, not a decline.

Another noticeable trend has been the increase in non-state violence in the last few years as well. To see this trend in a different way, we plotted a proportion graph of the same binary variable.

Who is perpetrating violence in the world?

In the last decade, it seems like there is a higher proportion of non-state actors perpetrating violent acts, though the majority of violent acts are still carried out by government forces.



Source: UCDP dataset

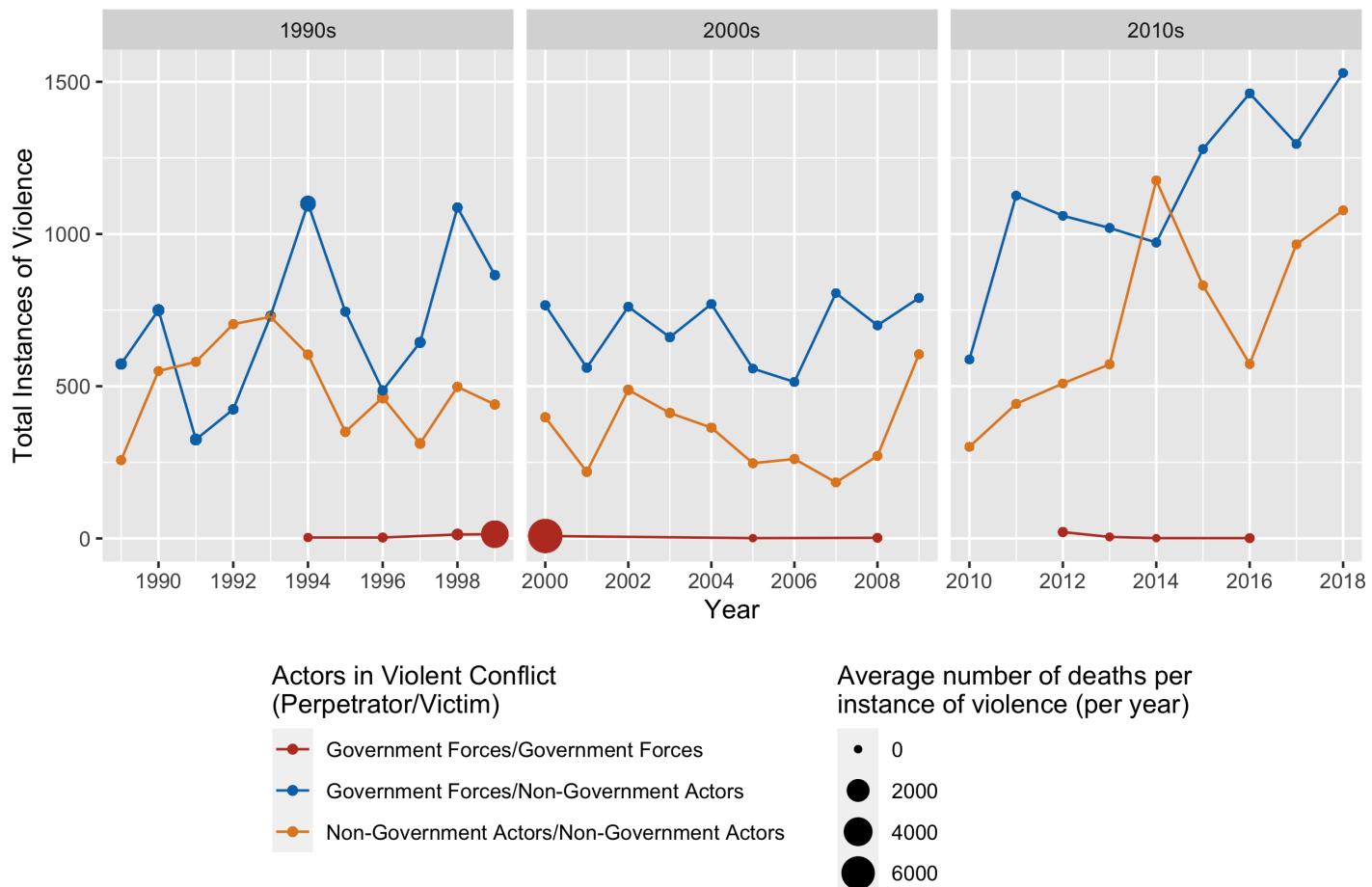
As this plot shows, there has been a shift to a higher proportion of non-state actors accounting for violent acts in the last five years. This also give us a clear visual that governments still perpetrate the vast majority of violent acts. Now that we have an idea of a worldwide trend, let's see if any interesting trends pop up when we look at the various regions of the world.

For this section, rather than looking at just government vice non-state actors as the perpetrator, we wanted to further classify based on both aggressor and victim. This can provide a visual context for inter-state conflicts (i.e. government forces attacking other government forces) in different parts of the world. We `mutate` a new variable that codes for three levels: government vs government forces, government vs non-government forces, and non-state based violence.

Africa

Violence in Africa

The number of violent acts has been on the rise in the last decade, though the average number of deaths per action is not as high as it was during the 1990s.

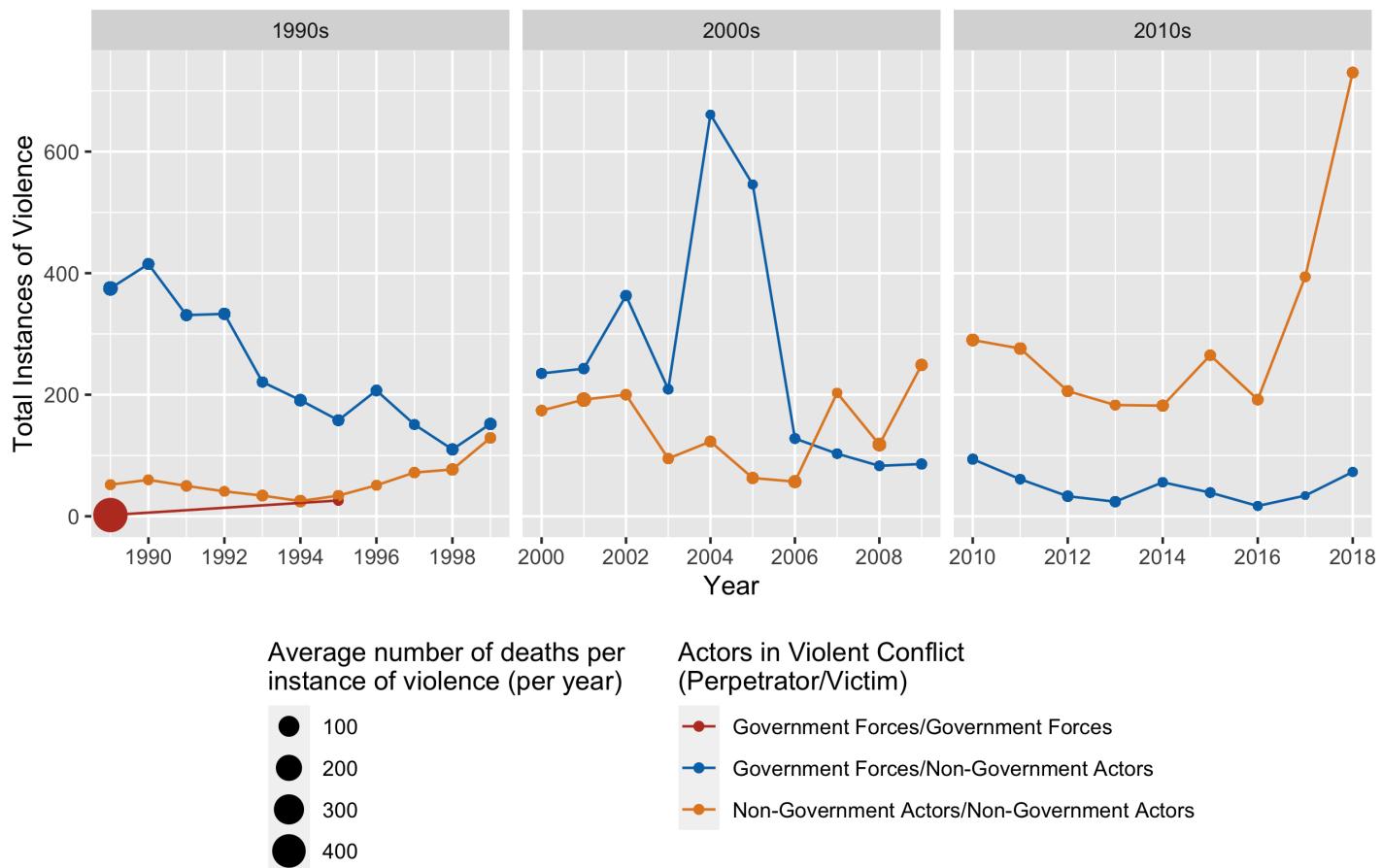


It seems that Africa had a tumultuous period in the 1990s. This period included one of the most famous conflict events: the Rwandan Genocide (how this affected our data will be discussed later). In the 2000s, it seems as if violence, while still present, seemed to almost stabilize; meaning, there was not as many peaks and valleys across the different aggressor/victim types. Unfortunately, the 2010s do see sharp increases in both government vs non-government violence, and non-state actor based violence. Across the whole dataset , government vs government violence looks oddly low. However, if you think about it, a rebel group is considered a non-state actor during a conflict; if the group succeeds in a coup, they only become “government forces” after the fact. This is demonstrated by the AFDL and the creation of the Democratic Republic of the Congo in the 1990s (see hypothesis 2).

The Americas

Violence in the Americas

While the Americas has seen a steady decrease in government-sponsored violence, there has been a steady increase in non-state violence, with a severe spike in recent years.

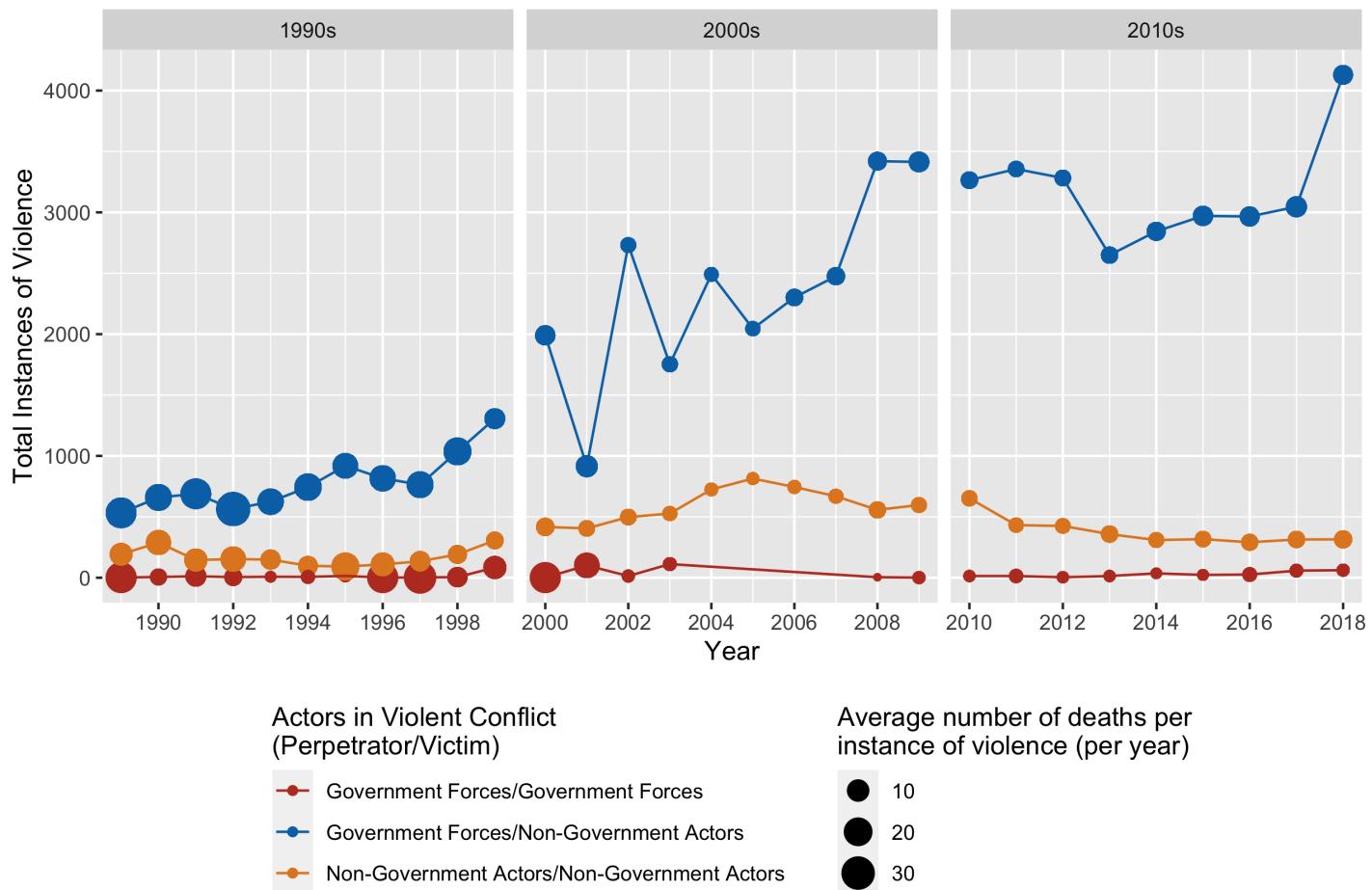


Two trends that are quite clear across the decades in the Americas is a decline in government based violence and an increase in non-state actor violence. The spike in government action in the 2000s represents a period of terrible conflict between the Government of Colombia and the FARC rebels. The upward trend in non-state actor violence is discussed further in our second question. The profile of violence in the Americas does seem to match our initial hypothesis: a decline in state-sponsored violence with an increase in non-state actor based violence.

Asia

Violence in Asia

Government-sponsored violence in Asia has been on a steady rise, and government v government violence has been regularly occurring over the last 30 years.

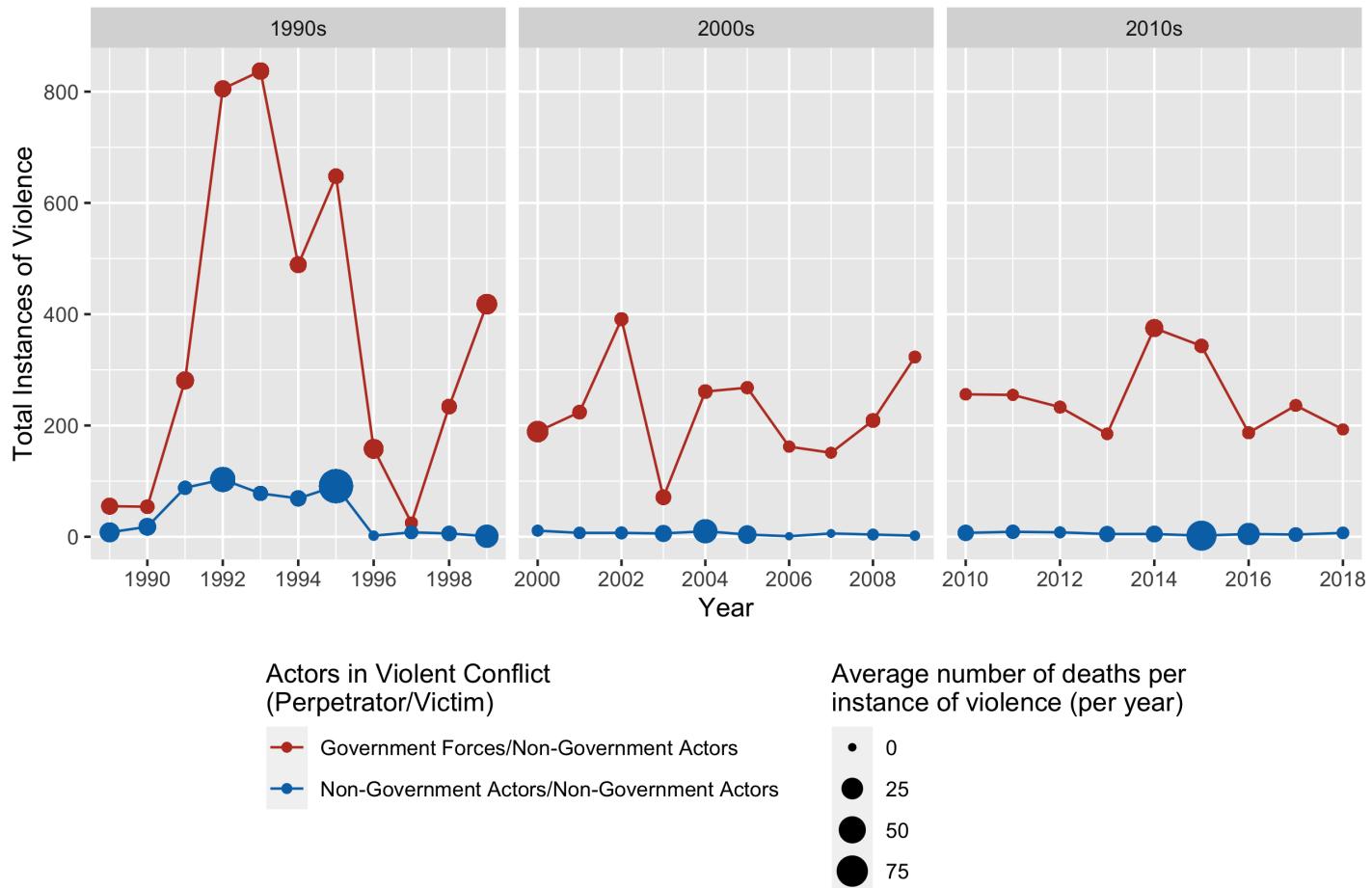


Two notable points from this graph is the definite increase in government action over the last three decades, and the continual presence of government vs government action. One thing to note about our methodology is that government action within a region does not have to be perpetrated by a government of a country within that region. For example, increased government vs non-state violence in Asia does include U.S. forces fighting against Taliban and al-Qaeda forces. The War on Terror could certainly account for the increase in reported instances in Asia. Asia, unlike in the Americas, seems to see a trend directly opposite from our hypothesis: state-sponsored violence has been on a steady rise, while non-state violence has declined in the last decade after a peak in the 2000s.

Europe

Violence in Europe

Europe has not seen government v government violence in the last 30 years, and overall the number of violent acts has seen a decline in recent decades.

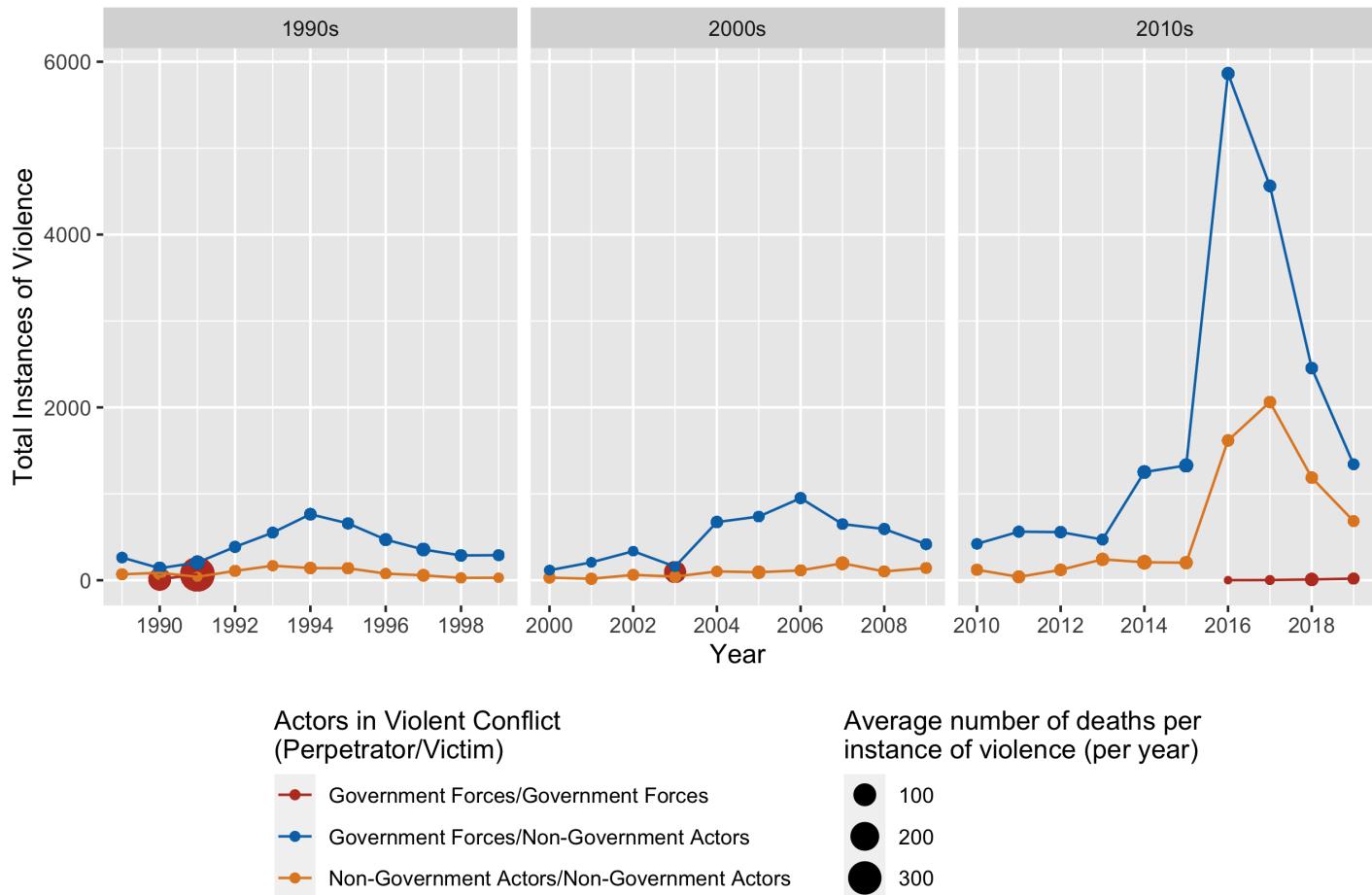


Note that this graph does not depict any instances of government vs government violence. Europe in the 1990s saw a number of violent conflicts, including the First Chechen War from 1994 to 1996. Aside from that spike, there's been a relative decline in instances of violence, and the numbers are much smaller in general than the other regions. One other item of interest is the islands of relative increase in average fatalities. They don't seem to follow any sort of trend, possibly indicating that they are localized on particular violent conflicts.

The Middle East

Violence in the Middle East

Relative to the previous two decades, the 2010s has seen a huge spike in violence, particularly in the last five years.



This graph is quite striking. The sheer magnitude of the spike in the 2010s is stark in itself. Of particular note is the spike in non-state actor violence. This may account for the jump in recent years in our proportion plot. Looking at the average fatalities in the government vs government data, we can most likely attribute these spikes to the major conflicts in the region within the time frame of our dataset, namely the Gulf War and the invasion of Iraq after 9/11.

To further explore the role of non-state actors in our data, we decided to drill down on the most violent players in each of the decades.

Drill Down: Who are the deadliest non-state actors in each decade?

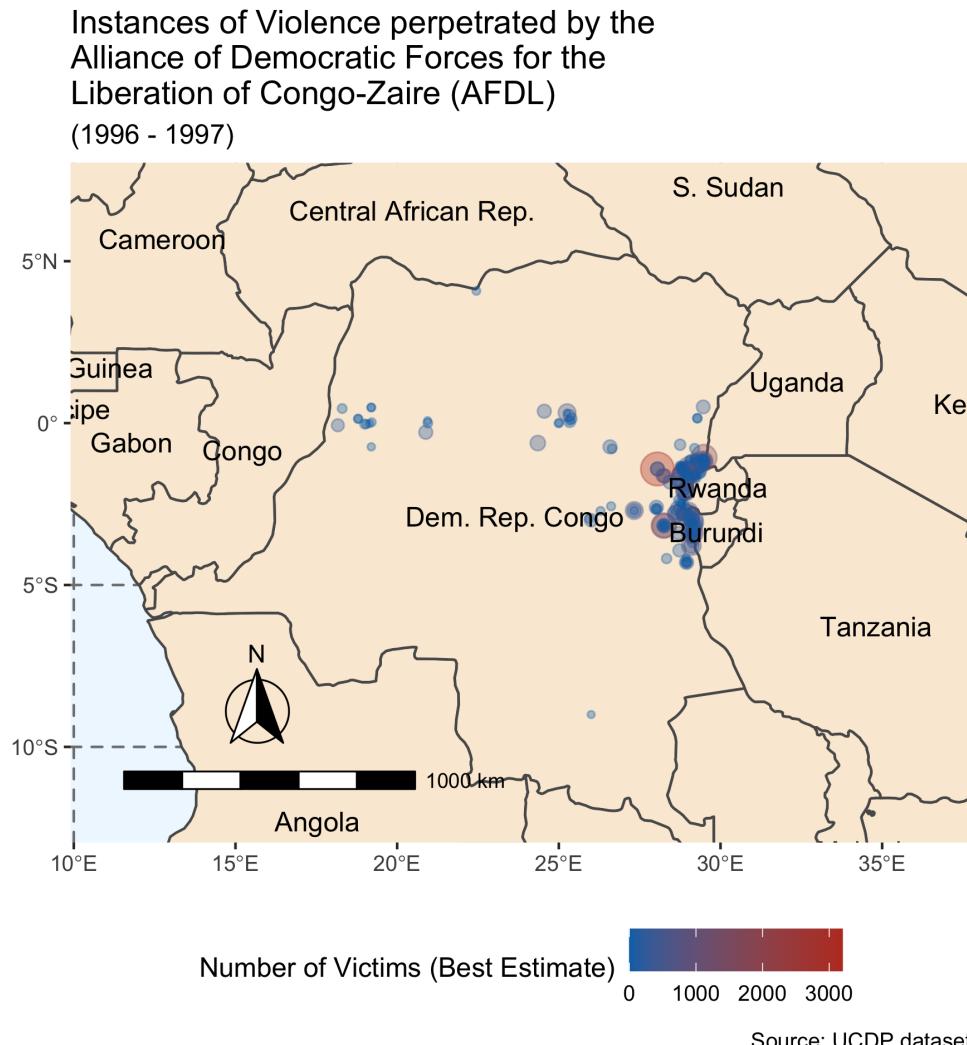
From our original dataset, we can find out which non-state group claimed the highest number of victims in each decade, which would yield the following table:

The Deadliest Non-State Actors by Decade

Decade	Non-State Actor	Total number of Victims (Best Estimate)
1990s	AFDL	35287
2000s	LRA	5474
2010s	IS	42316

While we didn't group by region for this finding, it is quite telling that for both the 1990s and the 2000s, the deadliest groups in the world were both active in central Africa. Let's take a look at each one.

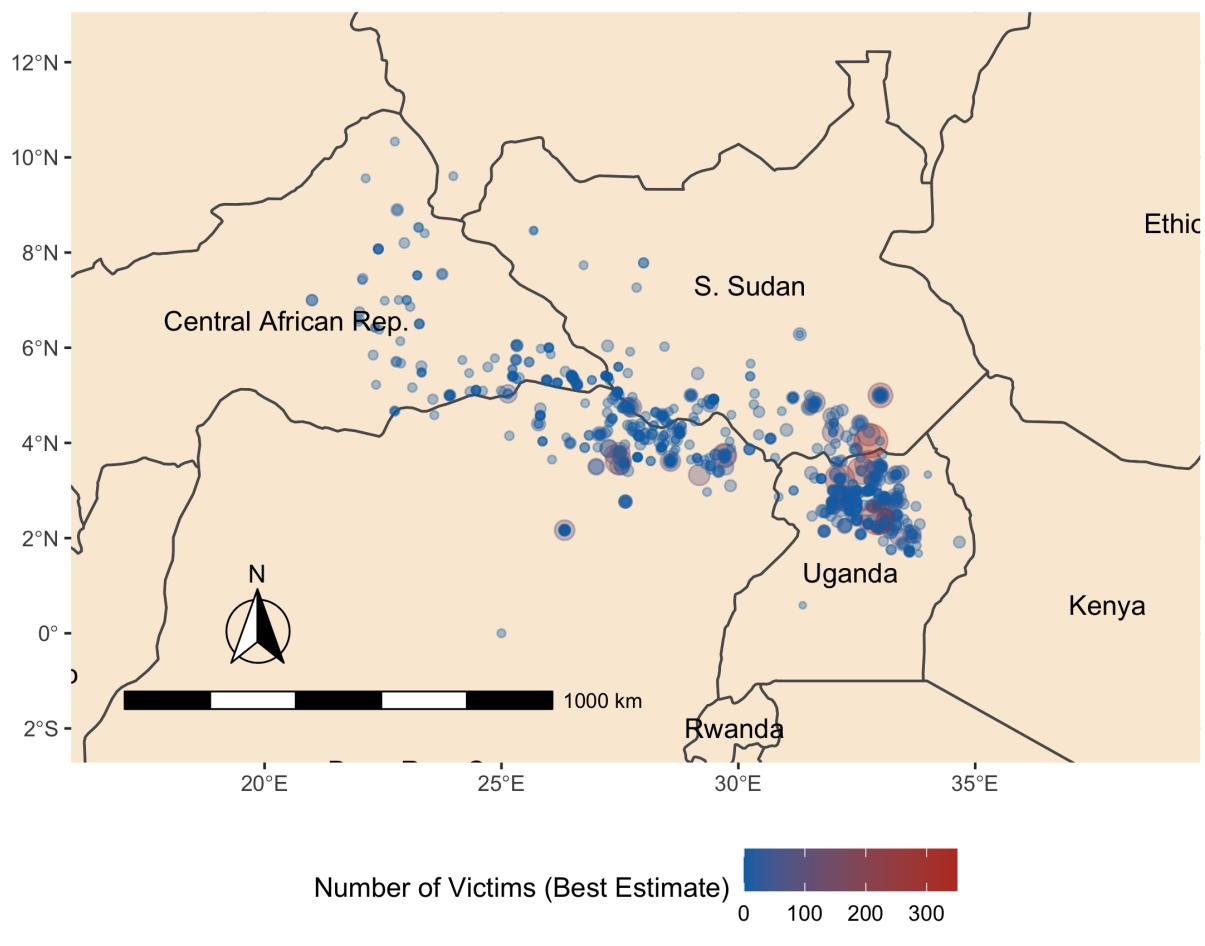
The Alliance of Democratic Forces for the Liberation of Congo-Zaire (AFDL)



The deadliest non-state group in the 1990s was The Alliance of Democratic Forces for the Liberation of Congo-Zaire, or the AFDL. The Rwandan Genocide had left the eastern part of then- Zaire destabilized, and the AFDL successfully defeated the government to form the Democratic Republic of the Congo after the First Congo War. Much of the conflict along the borders of Uganda, Rwanda, and Burundi reflect the violence in this period.

The Lord's Resistance Army (LRA)

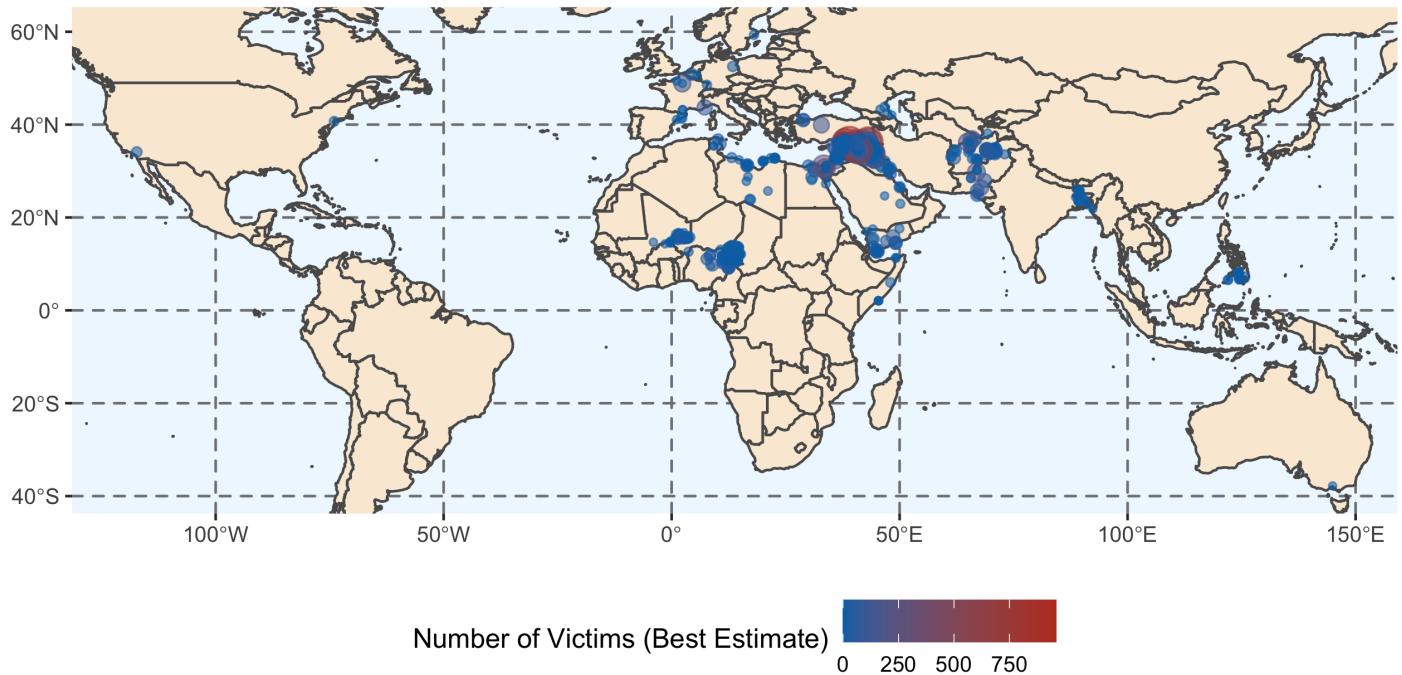
Instances of Violence perpetrated by the Lord's Resistance Army (LRA) (1989 - 2018)



In the 2000s, the Lord's Resistance Army (LRA) had massive amounts of conflict in central Africa. Led by Joseph Kony, they had not only abducted an estimated 25,000 children to act as child soldiers, but are responsible for thousands of deaths in southern Sudan, northern Uganda, and the DRC, as well as the displacement of hundreds of thousands of Congolese.

Islamic State (IS)

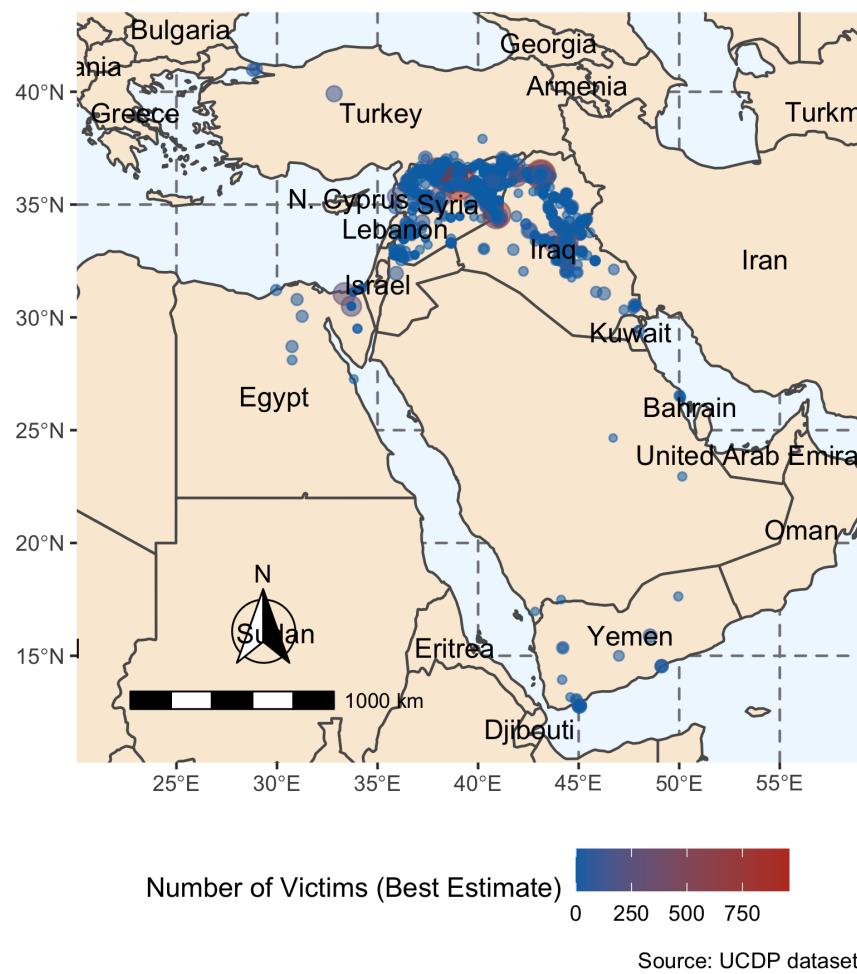
Instances of Violence perpetrated by the Islamic State (IS) (2004 - 2019)



Source: UCDP dataset

In the 2010s, the Islamic State (IS) has clearly taken over as the most violent group. What is distinct about IS's actions is their global reach. As opposed to the first two non-state actors, IS is shown to have a much further reach. It is not isolated to a particular country or region. The difficulty in combating international terrorist organizations in particular partially comes from their lack of specific geo-political goals. The ADFL and LRA were tied to geopolitical and ethnic feuds, and thus localized to where the affected populations exist. IS's jihad against western normative culture allows them free reign to target any particular place or peoples that draw their ire. Additionally, splintering and factioning within IS, as well as cross-group support, can create multiple "fronts", as can be seen in the worldwide map. Let's take a look at their actions specifically in the Middle East.

Instances of Violence perpetrated by IS in the Middle East (2004 - 2019)



As we can see, there is a heavy concentration in Syria and northern Iraq. With the Syrian Civil War ongoing, the instability in the country made it a perfect environment for extremist ideologies to gain a foothold. In the future it would be interesting to graph a fine detail time series on IS's violent acts, but for now we can see by the sheer number of instances that IS has dominated the last decade as the non-state group responsible for the most instances of violence.

In summary, our initial hypothesis was mostly incorrect. The perception that state-based violence has gone down is probably due to the lack of "traditional" wars in recent years. There are always hotspots that crop up: India-Pakistan border disputes, North Korea/South Korea shelling across the DMZ. But state-sponsored violence has been on the rise. Another explanation is the bias against one's own country. An American may not view U.S. military action against terrorists as instances of violence, but that doesn't mean that those events shouldn't be counted. The second half of the initial hypothesis does have some merit. Relative to previous decades, the 2010s did see a surge in instances of violence by non-state actors (both in sheer number of instances and in proportion of total instances). The rise of international terrorism adds a particularly worrying dimension to it by making it borderless, or rather delocalized.

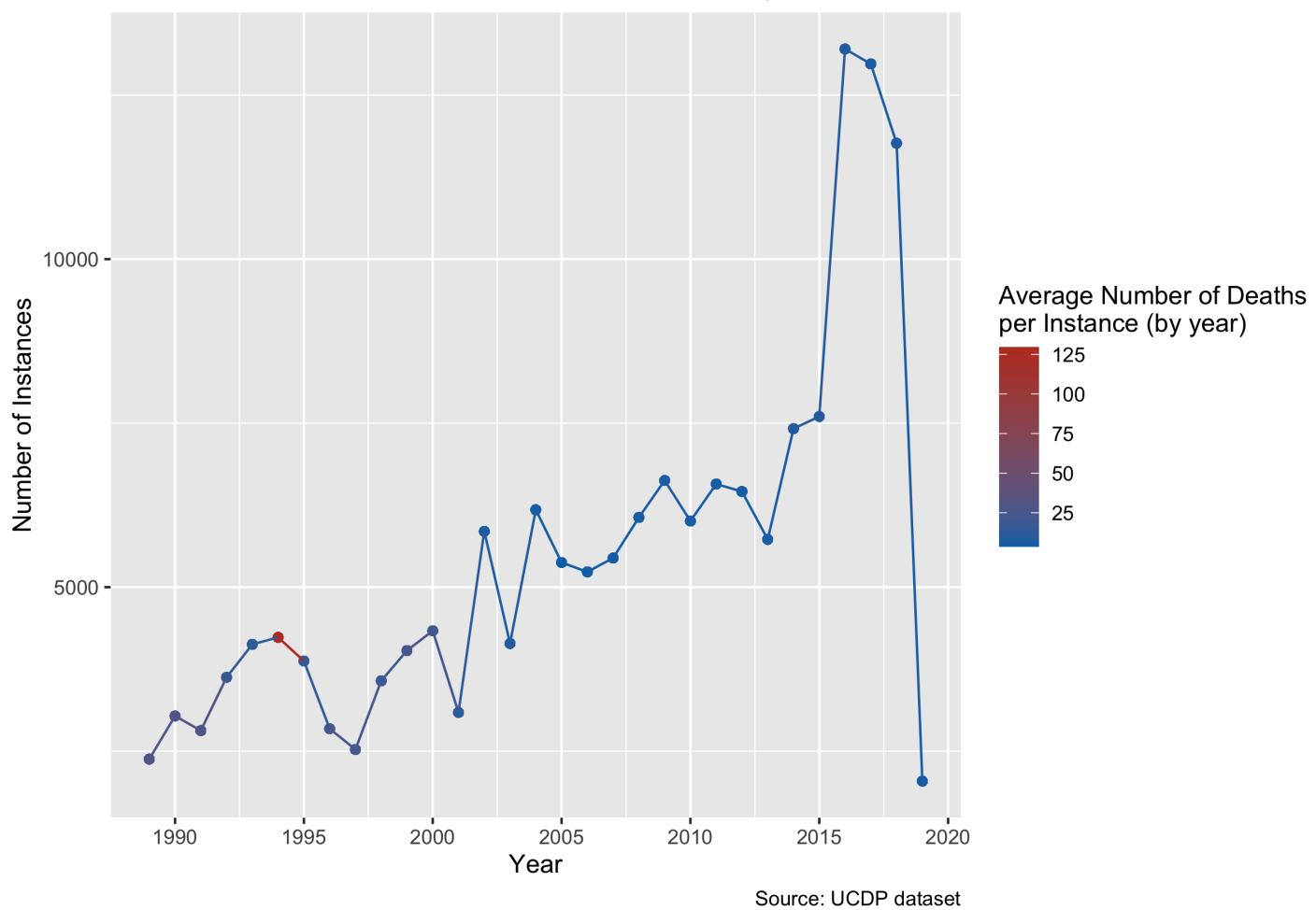
Hypothesis 2: "Violent conflict has risen in frequency over the last decade, but overall is lower than in previous decades"

Our second hypothesis is posed as a sort of alternative of our first; it's a prediction that the last decade has seen a rise in instances of violence, **but** has less conflict than in previous decades. Again, we are looking at instances of violence (i.e. number of observations in our dataset) by year. As always, we want to start by looking at a world

overview.

Reported Instances of Violence Worldwide (1989 - 2019)

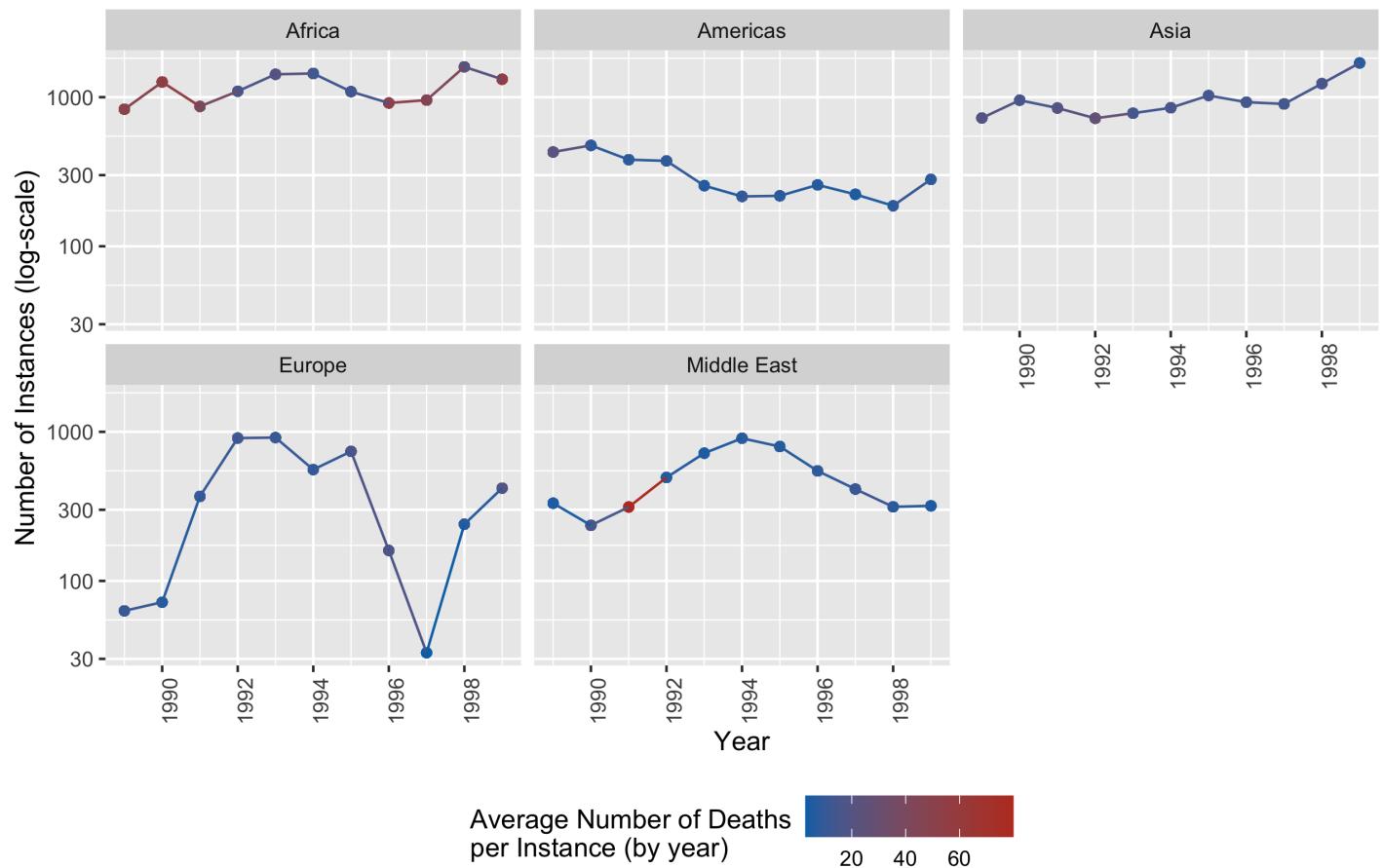
How much of the upward trend is related to better reporting?



Already, we can see that our hypothesis is incorrect. We were right in thinking that violence has increased in the last decade, but it is far and away higher than in previous decades. This plot does have some interesting implications though. Two initial thoughts are raised about validity. First, at the far right of the graph we see the tremendous drop for 2019 (which follows smaller drops after a peak in 2016). The drops from 2016 to 2017 and 2017 to 2018 are themselves not too incredible to be believed, but 2019 clearly is not accurate. Before looking at this further, it's best to investigate each decade in detail.

Instances of Violence by Region (1990s)

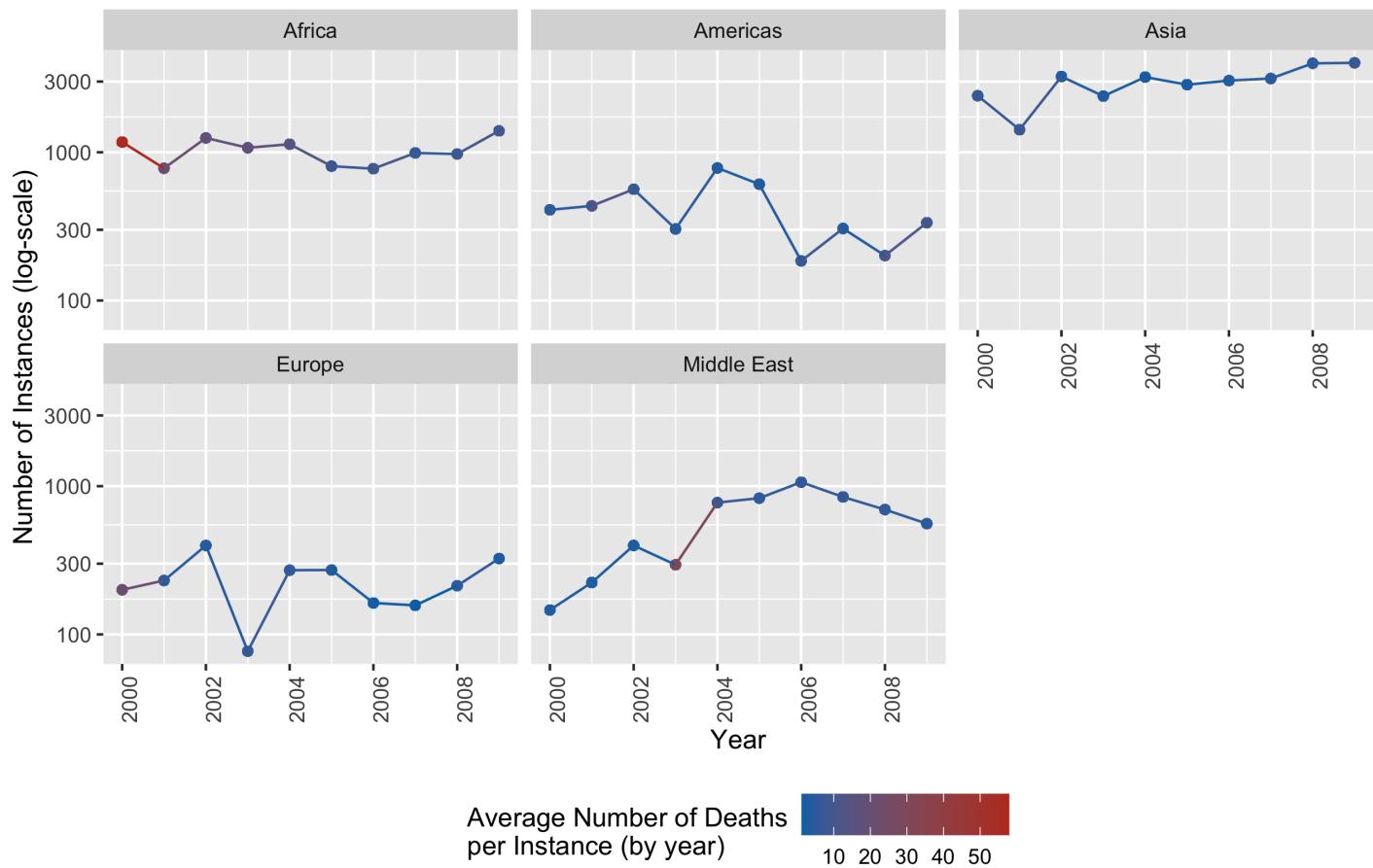
Numbers were calculated after removing data of instances in Rwanda between 1990 and 1996, to avoid heavily skewed graphs



Source: UCDP dataset

Instances of Violence by Region (2000s)

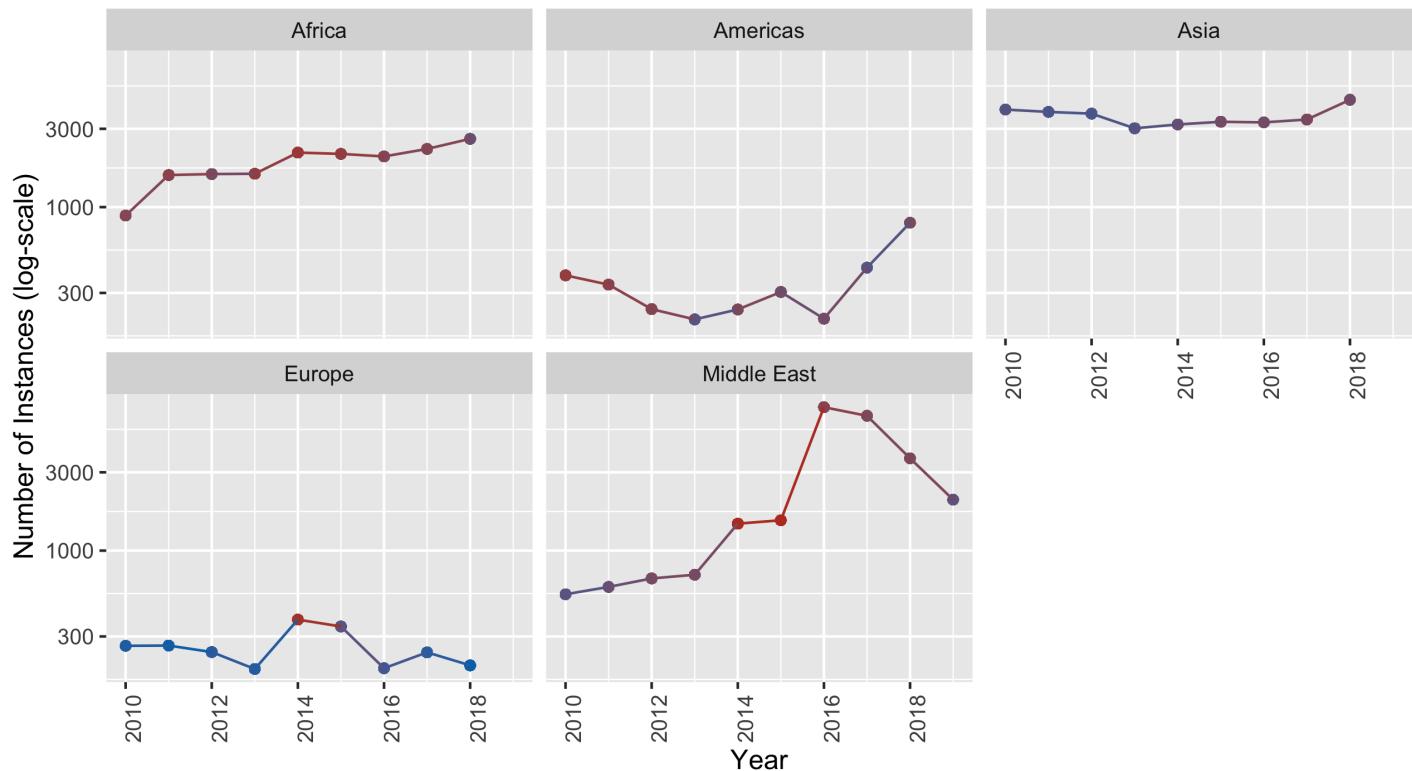
Numbers were calculated after removing data of instances in Rwanda between 1990 and 1996, to avoid heavily skewed graphs

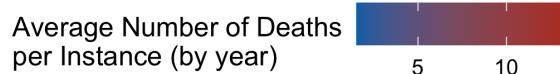


Source: UCDP dataset

Instances of Violence by Region (2010s)

Numbers were calculated after removing data of instances in Rwanda between 1990 and 1996, to avoid heavily skewed graphs



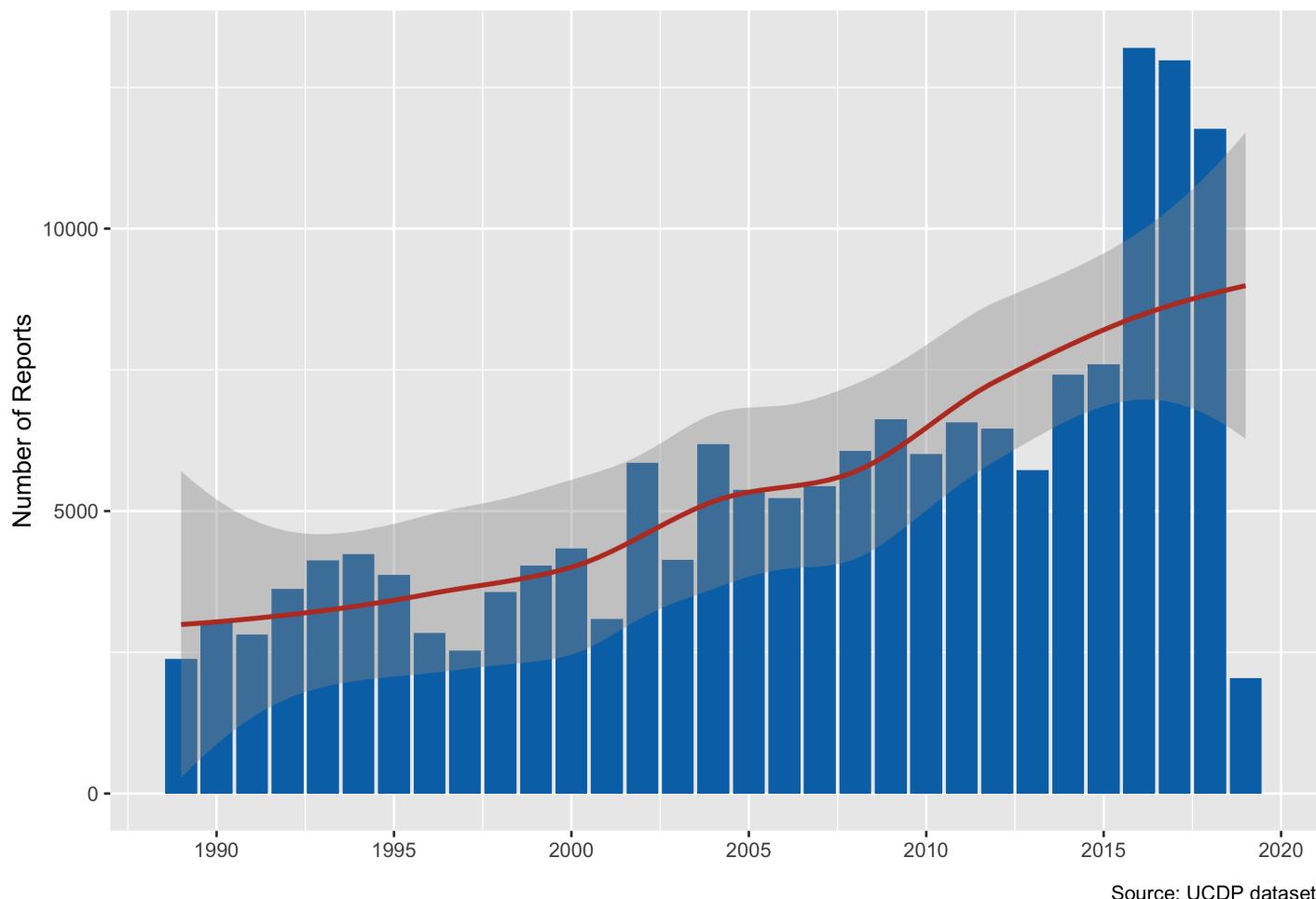


Source: UCDP dataset

These plots show that Europe and the Americas have seen an overall decrease across decades, while all other regions have seen a steady increase across our dataset. Of particular interest is the spike in the Middle East. Again, we want to take a look at the validity of our dataset. If you look closely, one notices that for the 2010s, only one region (the Middle East) has a data point in 2019. All other regions' data ends in 2018. This in itself is enough to tell us that the data for 2019 is incomplete. Let's take a look at another graph to check our suspicion:

Number of Observations in UCDP Dataset

We should expect many more reported instances for 2019



Source: UCDP dataset

With the added loess (locally estimated scatter plot smoothing) line, we see that our 2019 data is certainly outside what we would expect it to be.

Looking back at the codebook provided by UCDP, we note that their process for collecting data has multiple steps, included automated collection, human-driven data validation, as well as secondary human-driven data collection from local news sources; this last step also includes language translation, which is a time- and labor-intensive process. It only makes sense that there would be a lag in reporting, as there is for a lot of datasets that are built on collection from disparate sources. Future iterations of this dataset are sure to include more data.

To our second question about validity of this methodology (looking at trends in number of instances), we may ask the question: how much of the observed upward trends is due to an actual increase in violent acts, and how much could be accounted for by better reporting? The world of telecommunication today is a far cry from what it

was in 1989. The spread of internet and mobile telecomm (i.e. cellphones) may play a large role in how conflict data is being reported on and collected. Let's bring in some World Bank Development Index data on cellphone and internet usage.

Drill-down: Is there a correlation between telecommunication development and reported violence?

For a closer look at the data, we wanted to identify the countries with the highest number of reported cases in a given year. The following table shows the top twenty country-years for reported instances of violence.

The Top 20 Country-Years of Reported Violence in the UCDP dataset

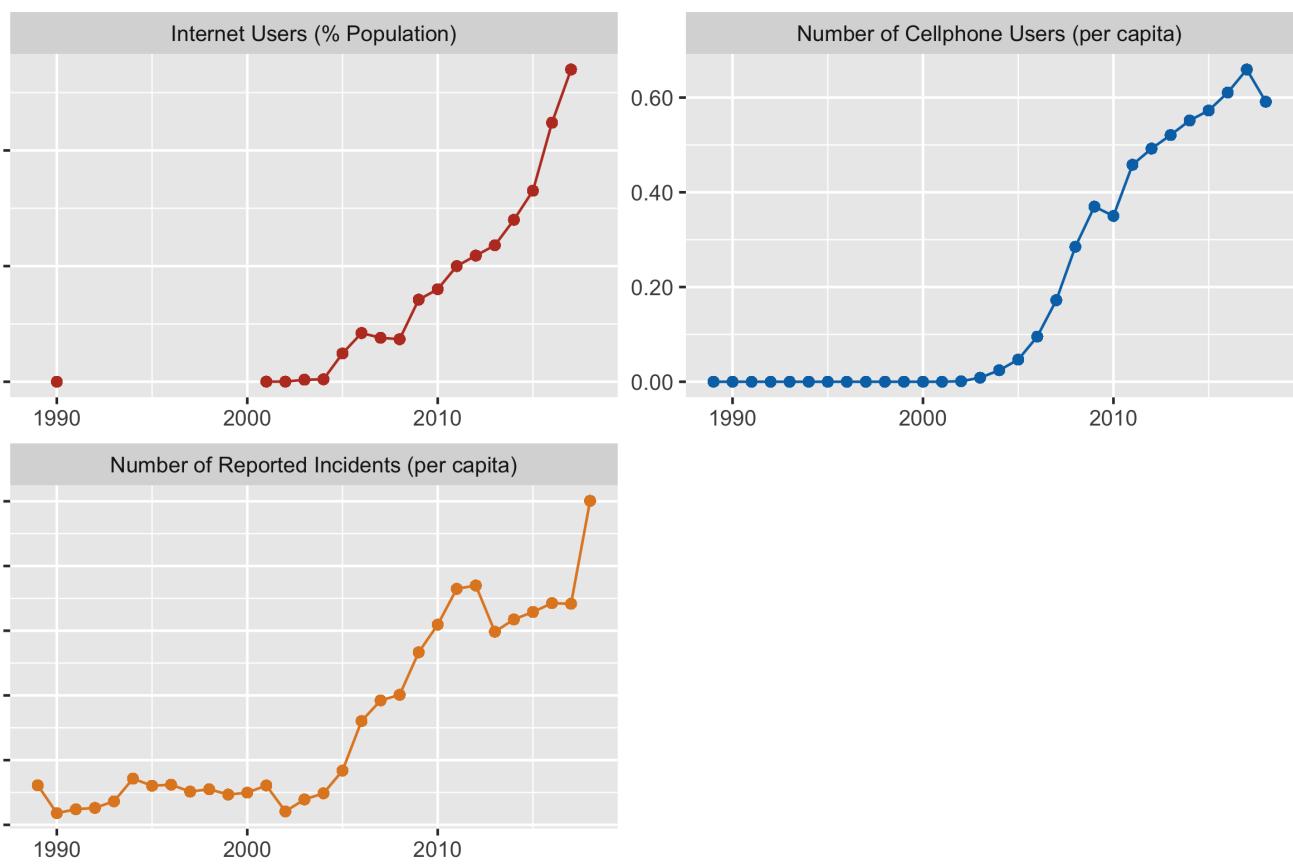
Year	Country	Total Number of Reports of Violence
2016	Syria	6170
2017	Syria	5587
2018	Afghanistan	3722
2018	Syria	2765
2017	Afghanistan	2480
2016	Afghanistan	2424
2012	Afghanistan	2305
2015	Afghanistan	2264
2011	Afghanistan	2197
2014	Afghanistan	2119
2019	Syria	2044
2013	Afghanistan	1926
2010	Afghanistan	1806
2002	Nepal	1531
2009	Afghanistan	1514
2004	Nepal	1360
2000	India	1286
2008	Afghanistan	1114
2007	Afghanistan	1042
2008	Sri Lanka	1016

Within this table there are five countries across the top twenty entries: Syria, Afghanistan, Nepal, India, and Sri Lanka. Unfortunately, the World Bank didn't have updated data on Syria for these indices, so we had to look at the other four countries. The following are visualizations for a countries number of reported instances of violence (per capita, calculated with World Bank population estimates), its cell phone users by year (also per capita), and its number of internet subscribers (as a percentage of population).

Afghanistan

Does the number of reported incidents per capita increase with the introduction of telecommunication infrastructure?

The number of incidents per capita in Afghanistan seem to correlate with the development of telecommunications infrastructure.



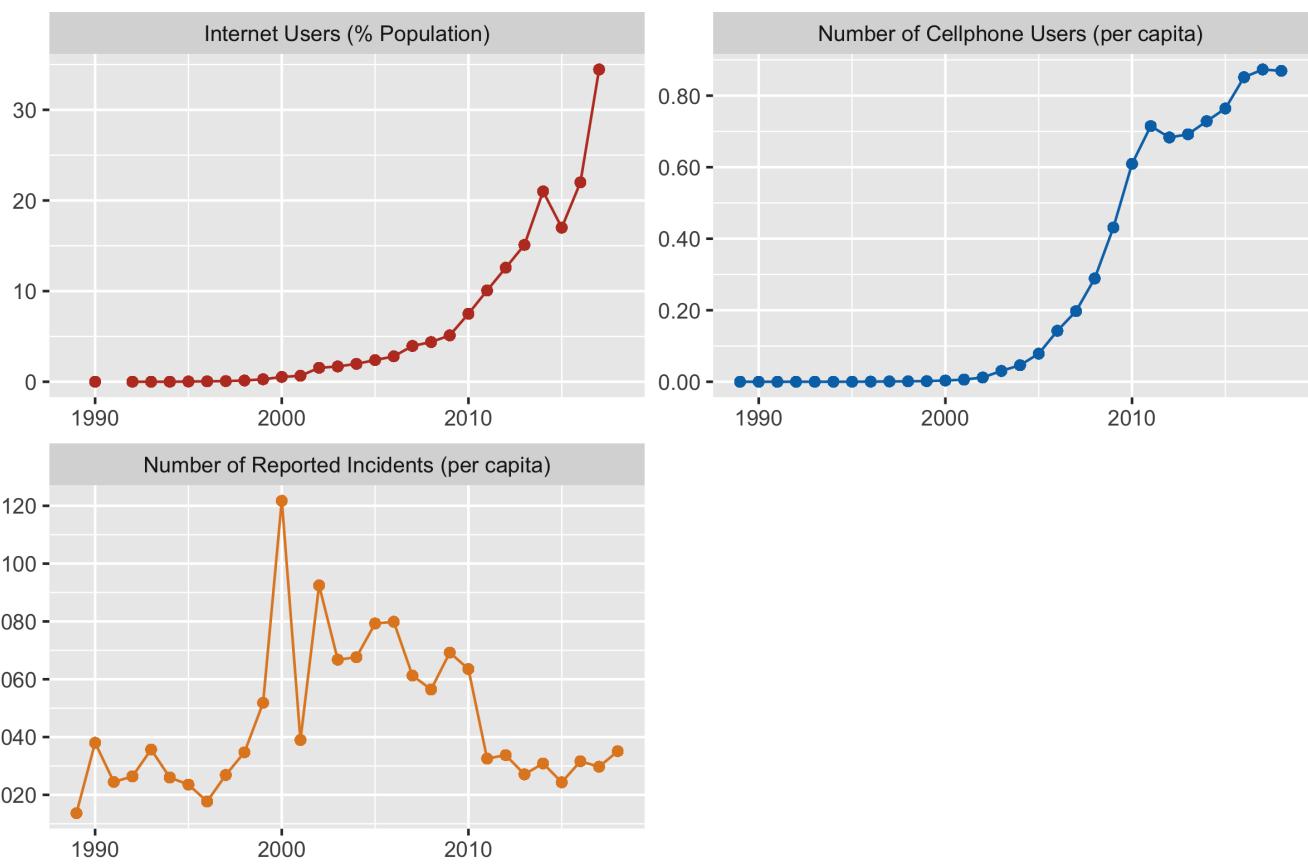
Source: UCDP dataset and World Bank dataset

This graph seems to show a correlation between number of reports and the increased use of telecomm in the country. However, these might be have a confounding variable: the War on Terror. An increase in violence in Afghanistan can be attributed to the post-9/11 surge in military action by coalition forces. Those coalition forces also initiated nation-building efforts, which I'm sure involved developing better telecommunication infrastructure throughout the country. So these indices could be related but not in a causal fashion.

India

Does the number of reported incidents per capita increase with the introduction of telecommunication infrastructure?

India's numbers for reported instances of violence sees a rise that predates an expanded use of cellphones and internet.



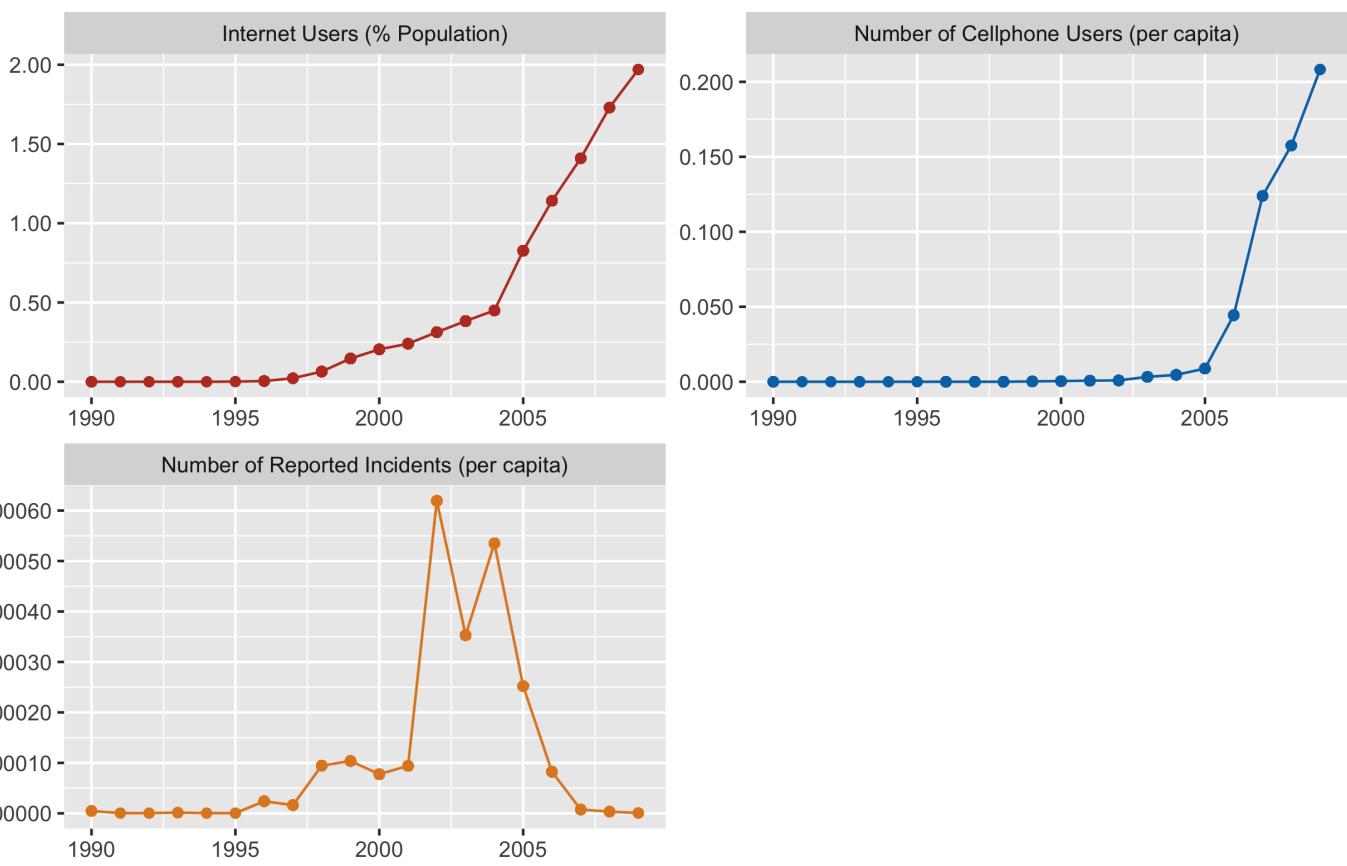
Source: UCDP dataset and World Bank dataset

India's violence graph, conversely, doesn't see as strong correlation between violence and telecomm. One contributing factor in this, I believe, is the sheer size of India, both in area and in population. India's history also lends itself to bucking the trend among the five countries we're looking at; as a former British colony, there's a good chance that in more developed parts of the country there already exists communication infrastructure necessary to report violence by more traditional means (i.e. local news reports and news-wire for major events). Additionally, the size of the population affects the scale of per capita graphs by an order of magnitude. In this way, dramatic shifts in the graph are an over-representation of movement that is in reality smaller than shown. A further drill down in granularity may help clarify (e.g. looking at state-level data for India, instead of country-wide).

Nepal

Does the number of reported incidents per capita increase with the introduction of telecommunication infrastructure?

Nepal's number of incidents per capita track an increase with the introduction of the internet.



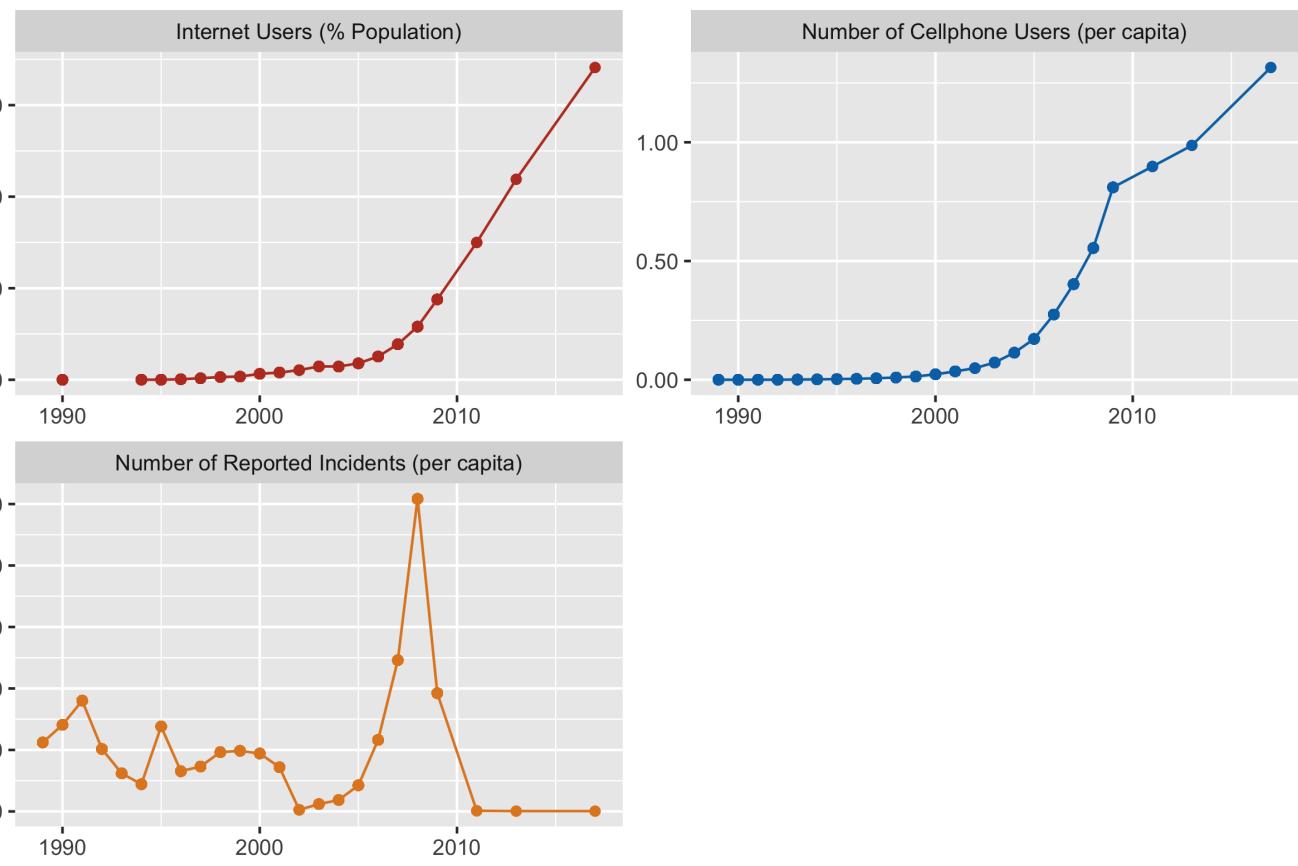
Source: UCDP dataset and World Bank dataset

This graph is quite striking. There again seems to be a correlation between telecomm and reporting of violence. The spike in the violence curve is representative of the collapse of peace talks during the Nepalese Civil War in 2002, with increased violence between the Government of Nepal and the Maoist rebels. However, the mountainous nature of the country could still lend credence to the correlation; would some of the reports of violence by the Maoists in rural parts of the country have been reported otherwise? Further research into this topic would certainly help.

Sri Lanka

Does the number of reported incidents per capita increase with the introduction of telecommunication infrastructure?

Sri Lanka's incidents do see a large spike around the same time that internet and cellphone usage begin to grow throughout the country.



Source: UCDP dataset and World Bank dataset

This graph is somewhere between India's graph and Nepal's graph. There are significant counts of violence before the increases in telecommunication, but there also exists the spike in violence between 2005 and 2009. This time frame is reflective of the fourth phase of violence between the Sri Lankan government and the Liberation Tigers of Tamil Eelam (LTTE), during the Sri Lankan Civil War, also known as the Sixth Eelam War.

Does this mean that our drill down question is invalid? Each of the spikes in violence in these graphs (that have them; India excluded) happens to represent major conflicts, either civil wars or the war on terror. Do these just happen to line up with the increase in telecomm usage in the countries? Our first look at this subject has yielded some interesting results, but does leave additional questions to be answered upon further investigation.

Question 2: “Is violent conflict becoming more deadly or less deadly over the years? Are there any exceptions to these trends?”

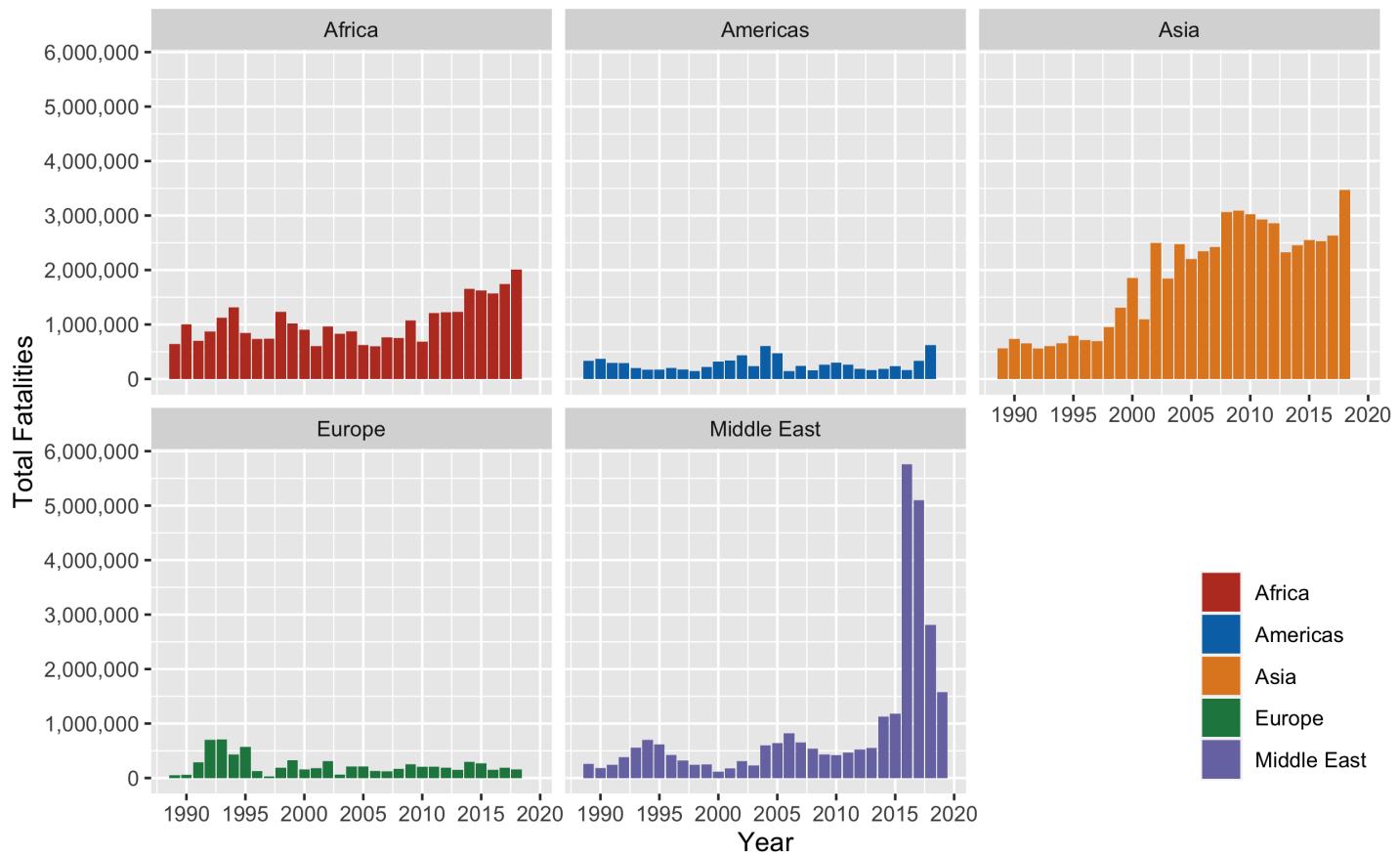
From our first question, we have seen that instances of violence have been increasing worldwide (whether as a function of better reporting or not is unclear). However, has that violence been more deadly as well as more frequent? Our initial hypothesis is as follows:

Hypothesis: “Although violent conflict has been on the rise globally in the past decade, average fatalities are likely to be lower than conflicts in previous decades.”

First, let's look at the total number of fatalities (from the best estimate variable in our dataset) per year and region.

Total fatalities by year and by region (1989-2019)

Log transformation will be required to better understand the data, as we have massive peaks in the Middle East and Africa that dwarf the scale of fatality counts in other regions.



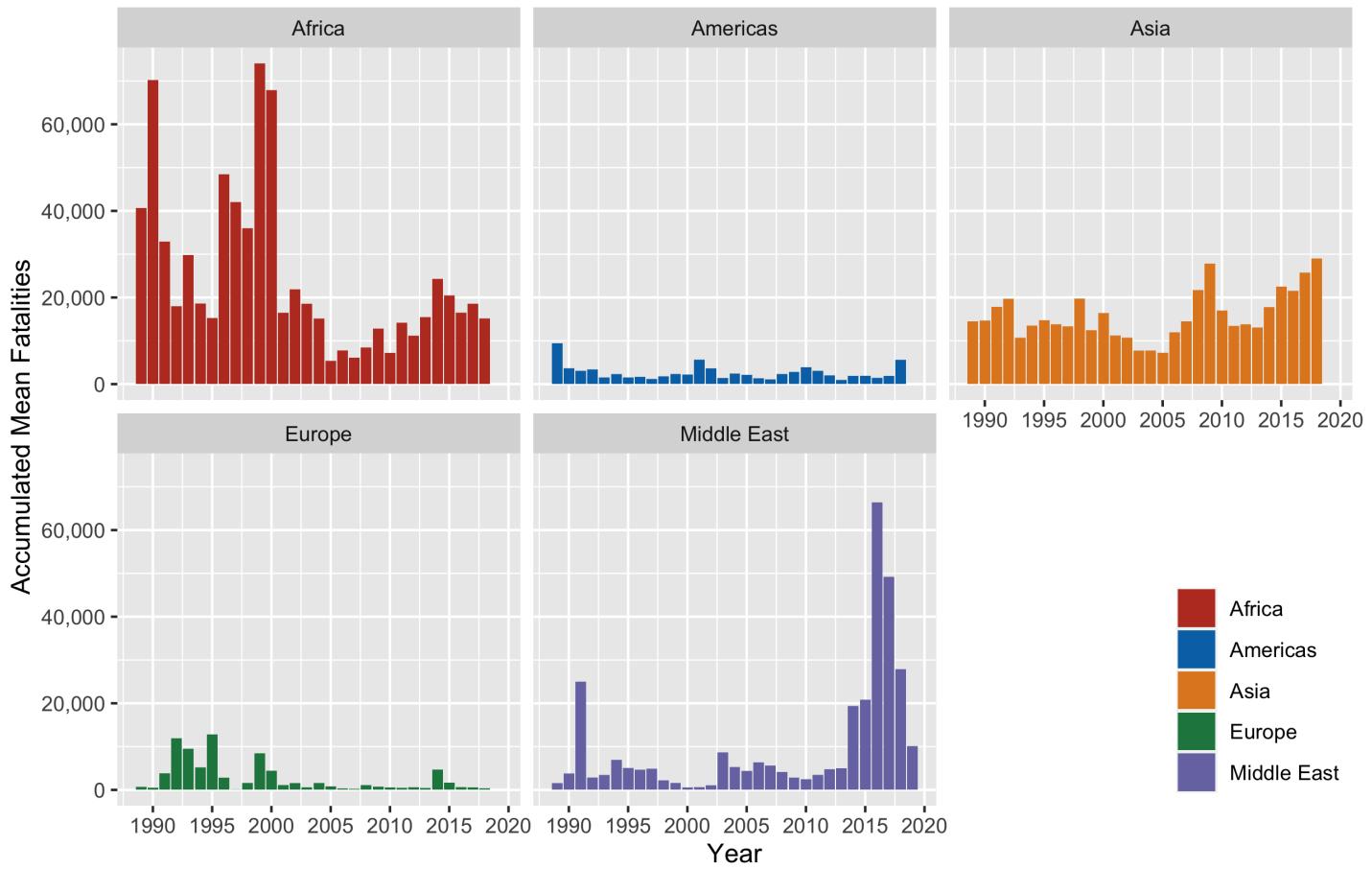
Source: UCDP dataset

In four of the five regions, we see an increase in yearly number of deaths. Only Europe has seen a downward trend in the last decade. One item of note is the spike in the Americas in 2018. Could the region be trending up again, and what could be causing this?

Next, we'll look at the average fatalities per instance per year.

Accumulated mean fatalities by year and by region (1989-2019)

Mean fatalities have spiked in Africa and the Middle East in the last decade, but are falling over the 31 year time period. In Asia however, mean fatalities are rising.



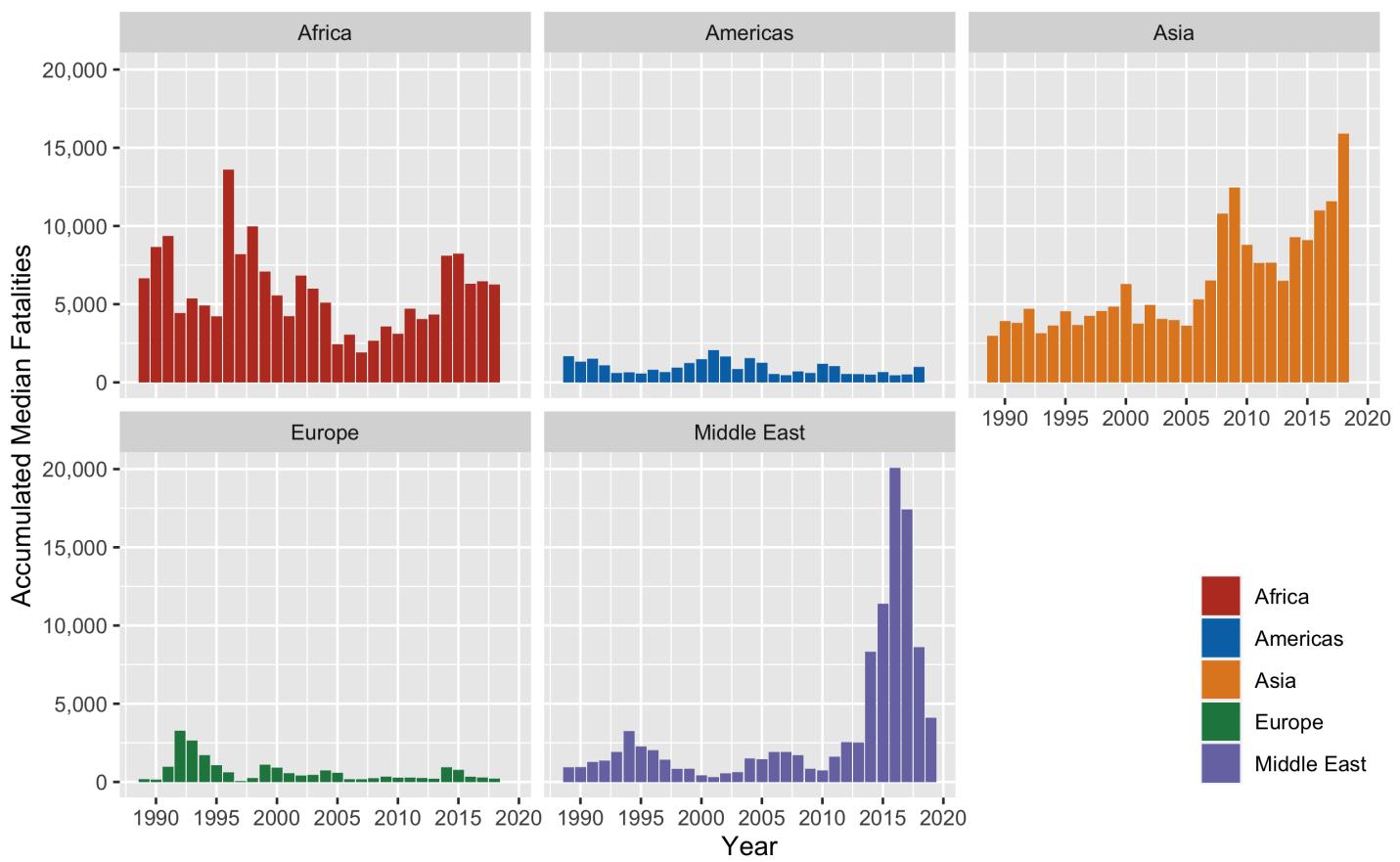
Source: UCDP dataset

Europe and the Americas are quite low relative to the other regions, while the Middle East and Asia have both seen an increase in mean fatalities in the last decade. Interestingly, Africa actually saw a decline from the peaks seen in the late 1990s and early 2000s. One obvious possible reason for this is the extremely high numbers from the Rwandan Genocide skewing the average numbers up. In the Middle East, there is a clear dramatic spike in 2016, while Asia has seen a more steady upward trend over the last ten years.

To try to account for skewing by outliers, let's look at the same data by median deaths rather than average fatalities:

Accumulated median fatalities by year and by region (1989-2019)

Median fatalities follow a similar trend to mean fatalities, in that they are decreasing in most regions, with the exception of a minor rise in the Americas and a much higher growth rate in Asia.

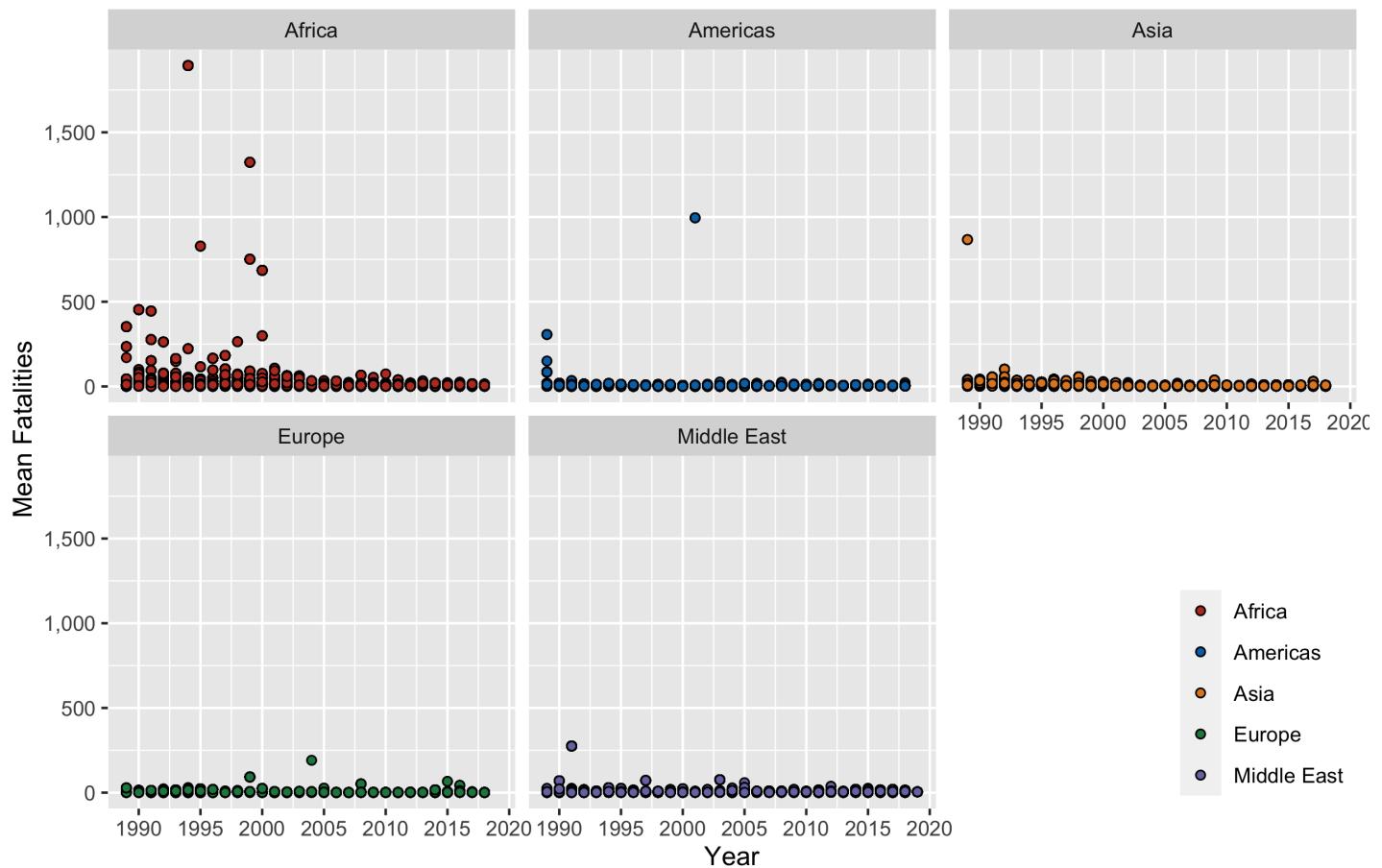


Source: UCDP dataset

The median plot in Africa doesn't have as dramatic of peaks in the 1990s, but they are still visible. Additionally, the median deaths in Africa have crept higher in the last decade than the mean fatalities. In Asia, the median plot is even more striking. Aggregating both metrics, the deadliness of conflicts have gone up at a regional scale. To look at the national level, we have to take a look at the distribution of national mean and median deaths over the years by region (as well as log-scale graphs for outlier normalization).

Distribution of mean fatalities by year and region (1989-2019)

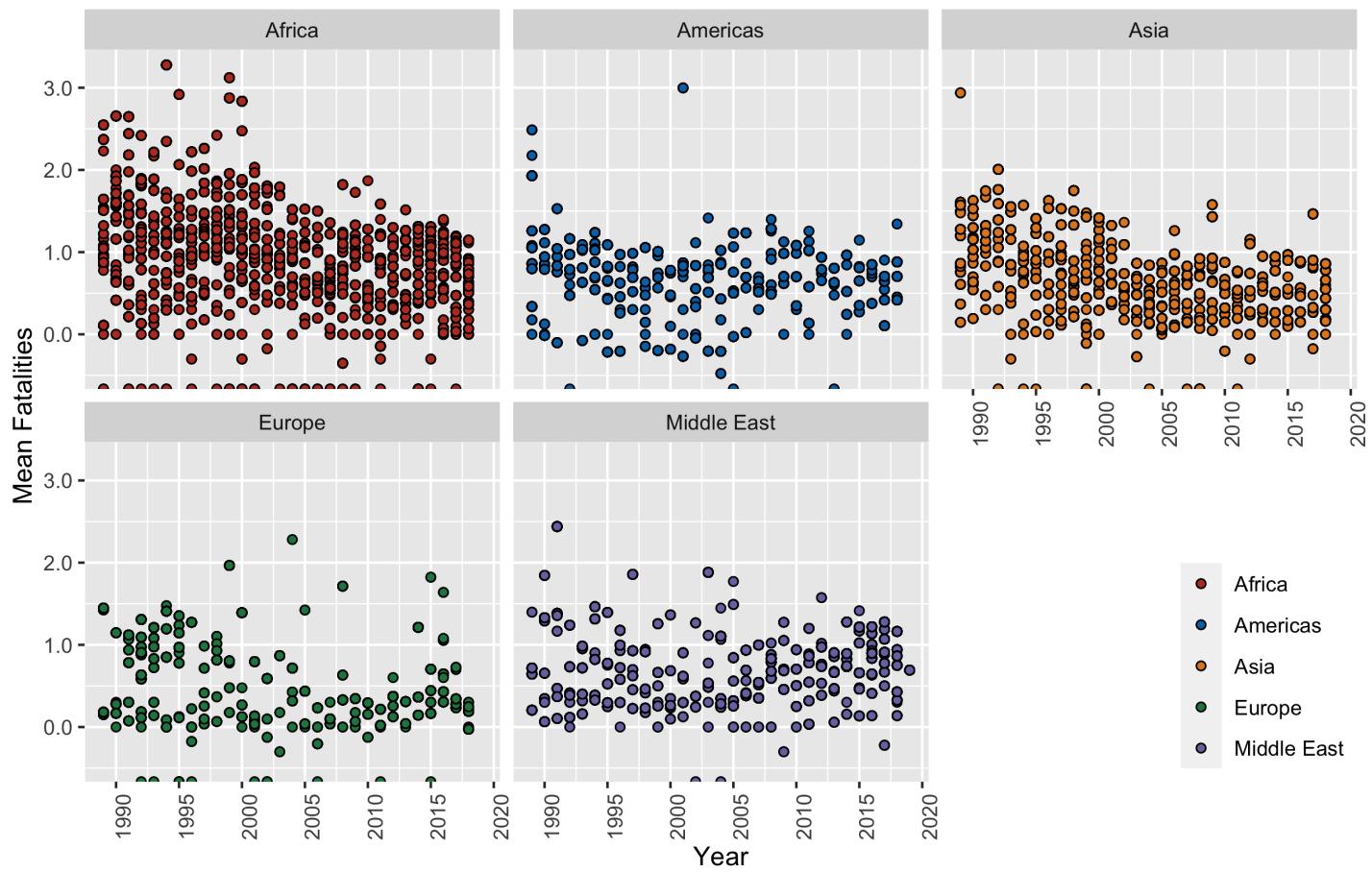
The distribution of mean yearly fatalities shows that fatalities for countries on year to year basis have fallen over the past three decades



Source: UCDP dataset

Distribution of mean fatalities by year and region, Log10 Scale (1989-2019)

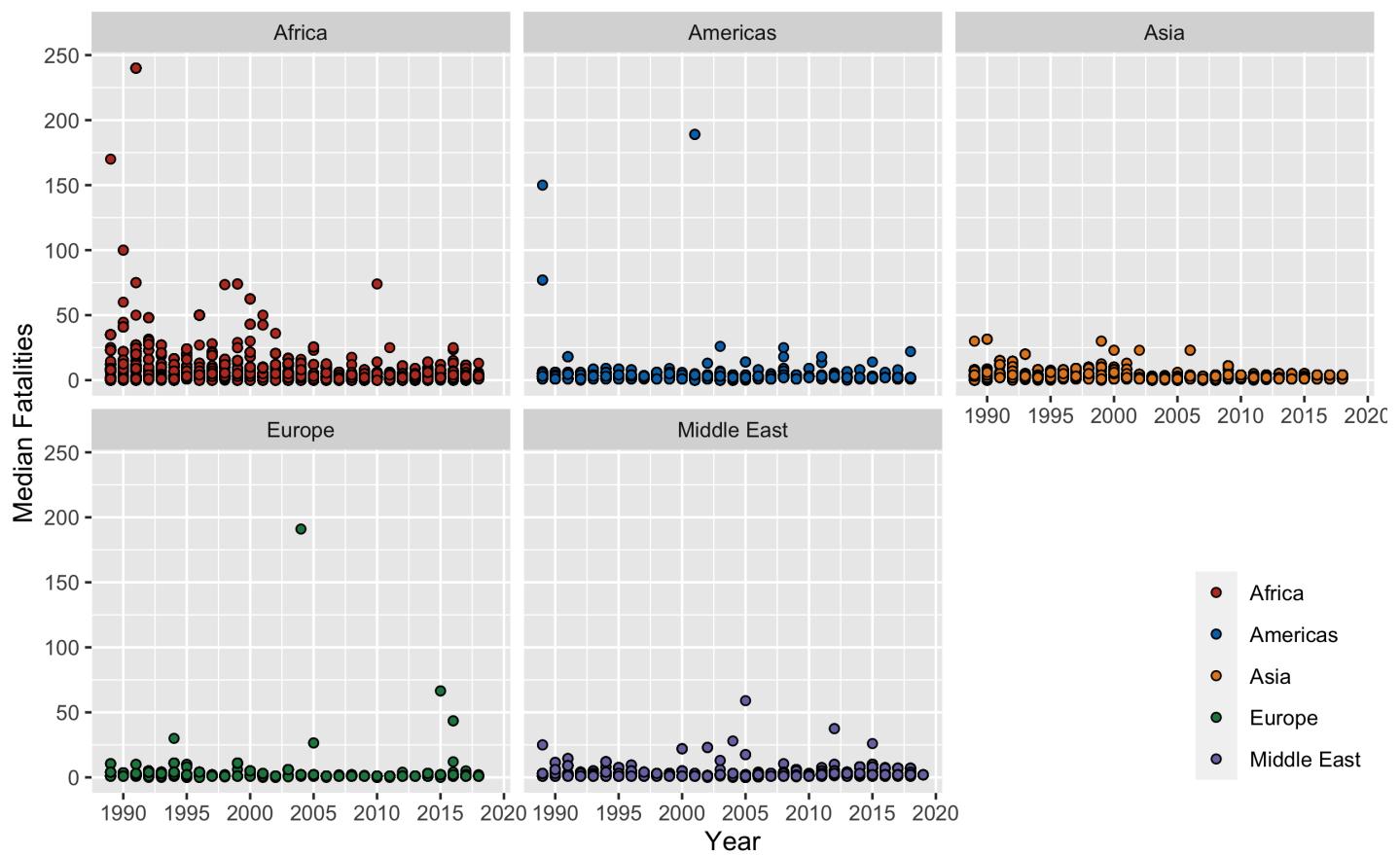
The distribution of mean yearly fatalities shows that fatalities for countries on year to year basis have fallen over the past three decades



Source: UCDP dataset

Distribution of median fatalities by year and region (1989-2019)

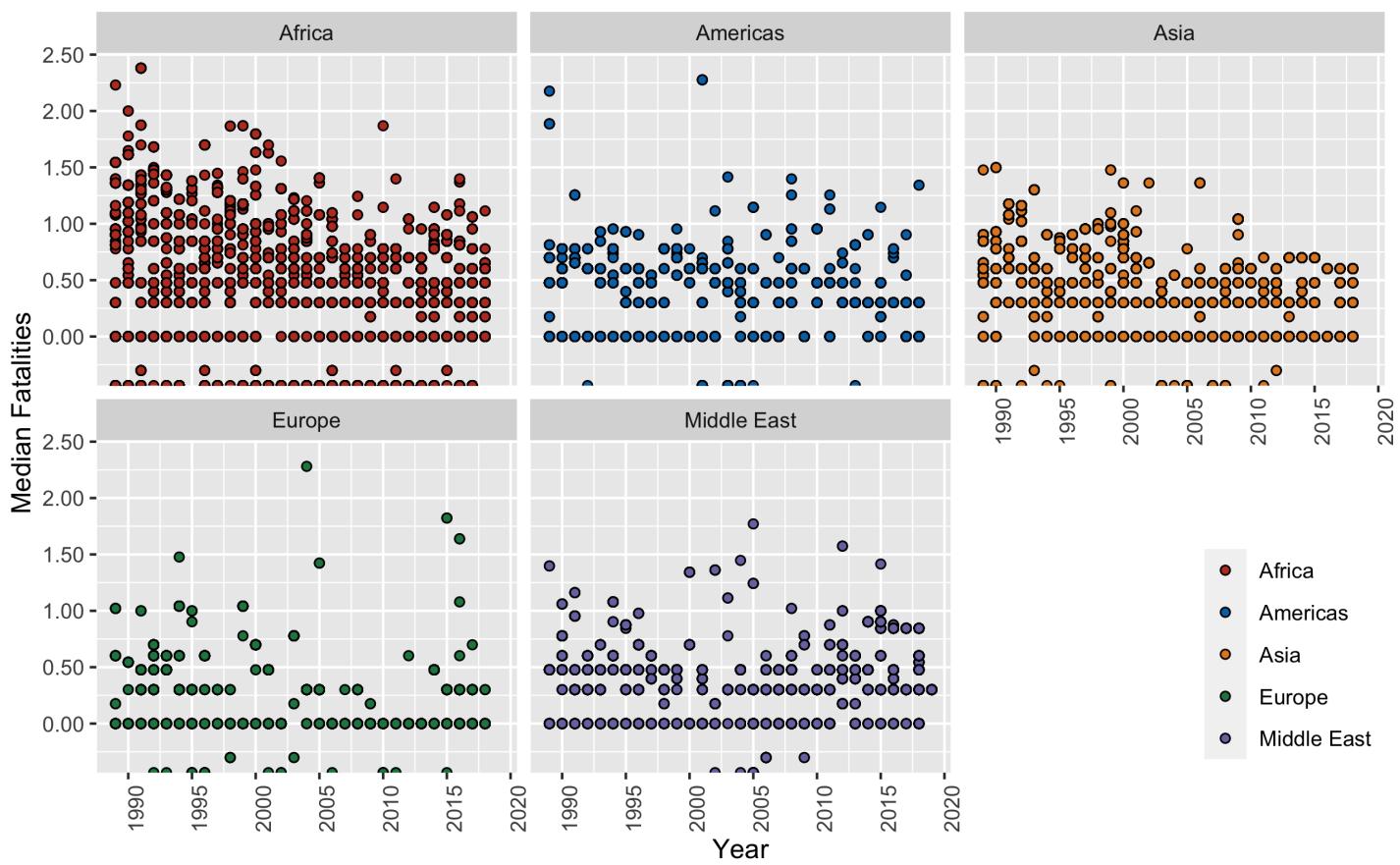
The distribution of median yearly fatalities shows a more uneven trend over time across regions, with growth in the Middle East and Europe. However, it seems that yearly medians are falling for Asia and Africa.



Source: UCDP dataset

Distribution of median fatalities by year and region (1989-2019)

The distribution of median yearly fatalities shows a more uneven trend over time across regions, with growth in the Middle East and Europe. However, it seems that yearly medians are falling for Asia and Africa.



Source: UCDP dataset

In general we see downward trends in the distribution of national means and medians. When we consider this alongside our observation of a rise in total number of fatalities in most of the regions, we have to conclude that there must be significant conflicts that are producing large numbers of deaths to account for the rise in totals. The upward trend in mean and median deaths in Asia draws our attention. In the same way as above, we wanted to take a closer look at the national distribution within Asia for mean and median deaths to see if we can extract validation for our analysis.

Distribution of average fatalities from conflict by year in Asia (1989-2019)

The distribution of both mean (in blue) and median (in red) fatalities of Asian conflicts are decreasing in number of fatalities but as we will see in later plots, this does not mean total yearly fatalities are falling.



Source: UCDP dataset

This graph shows us that, although our initial graphs showed that fatalities are on the rise, the average deaths per event (aggregated at a national level) has dropped, lending credence to our hypothesis. The spread in Asia in this distribution is much narrower in the 2010s than it was in the 1990s, (mostly) averaging fewer than 20 deaths per event. Still, we have to account for the rise in total, mean, and median deaths in Asia. Let's drill down.

Drill-down: Who is responsible for the rise in total, mean, and median deaths in Asia in the 2010s?

In order to do this, we wanted to identify the events that produced the highest death (estimate) counts to help us identify the dyads that are possibly responsible for the rise in fatalities.

The Top 10 Deadliest Instances of Violence by Year, with mean fatalities and type of violence (2010s in Asia)

Dyad Name	Country	Year	Best Estimate of Deaths	Total Fatalities per Year	Mean Fatalities	Type of Violence
Government of Afghanistan - Taliban	Afghanistan	2014	1899	12549	5.922133	1
Government of Afghanistan - Taliban	Afghanistan	2015	1576	17987	7.944788	1
Government of Myanmar (Burma) - Civilians	Myanmar (Burma)	2017	750	2331	29.137500	3
Government of Afghanistan - Taliban	Afghanistan	2018	576	26888	7.224073	1
Government of Afghanistan - Taliban	Afghanistan	2014	507	12549	5.922133	1
Government of Afghanistan - Taliban	Afghanistan	2014	374	12549	5.922133	1
Government of Afghanistan - Taliban	Afghanistan	2015	367	17987	7.944788	1
Government of Myanmar (Burma) - Civilians	Myanmar (Burma)	2017	358	2331	29.137500	3
Government of Afghanistan - Taliban	Afghanistan	2018	316	26888	7.224073	1
Government of Afghanistan - Taliban	Afghanistan	2013	297	8131	4.221703	1

As might be predicted, violence in Afghanistan has been rather deadly, with the deadliest event estimated to have nearly 1,900 victims. Violence connected to the Taliban accounts for over 25,000 entries in our dataset of about 170,000 entries (this is nearly 15%). Below are some of the top deadliest events involving the Taliban (nearly identical to the table from all of Asia).

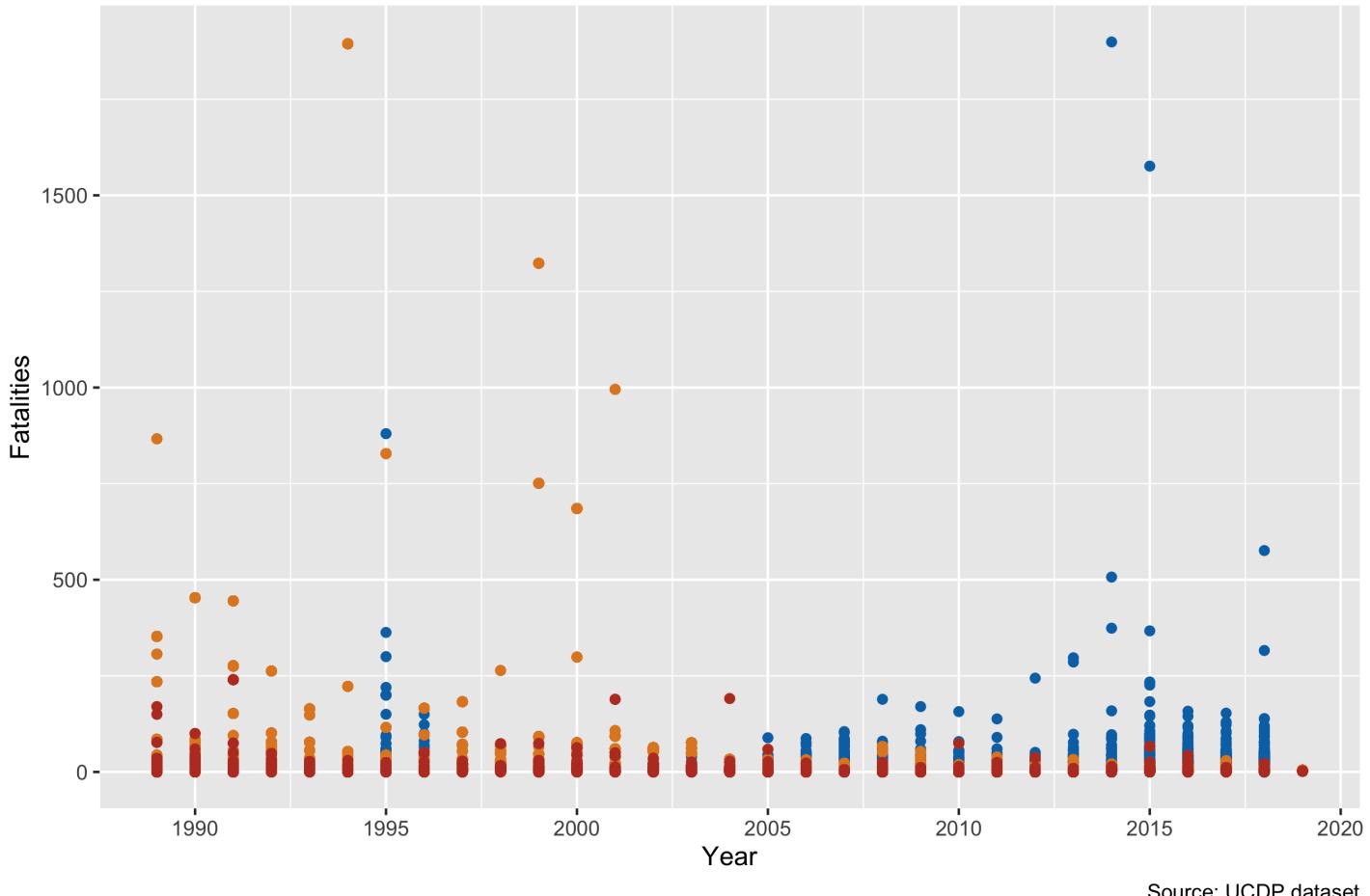
Instances of Violence Involving the Taliban

Dyad Name	Country	Year	Best Estimate of Deaths	Total Fatalities per Year	Mean Fatalities	Type of Violence
Government of Afghanistan - Taliban	Afghanistan	2014	1899	12549	5.922133	1
Government of Afghanistan - Taliban	Afghanistan	2015	1576	17987	7.944788	1
Government of Afghanistan - Taliban	Afghanistan	1995	880	5618	25.652968	1
Government of Afghanistan - Taliban	Afghanistan	2018	576	26888	7.224073	1
Government of Afghanistan - Taliban	Afghanistan	2014	507	12549	5.922133	1
Government of Afghanistan - Taliban	Afghanistan	2014	374	12549	5.922133	1
Government of Afghanistan - Taliban	Afghanistan	2015	367	17987	7.944788	1
Government of Afghanistan - Taliban	Afghanistan	1995	363	5618	25.652968	1
Government of Afghanistan - Taliban	Afghanistan	2018	316	26888	7.224073	1
Government of Afghanistan - Taliban	Afghanistan	1995	300	5618	25.652968	1

To cast a comparison, let's look at a similar density plot for numbers of deaths per entry across the years, looking at total deaths for Taliban related events against the means and medians for events in Asia:

Comparison of Taliban related conflict fatalities to Asian averages (1989-2019)

Violent conflict events in Afghanistan (in blue) have a much higher fatality count than the Asian mean (in orange) or median (in red)



Source: UCDP dataset

Since we can see that the Taliban related events have increased in deadliness (as indicated by the upward spread in the densities) vice the means and medians in Asia, we can see that these Taliban related violent acts have contributed significantly to the rise that we see in the regional totals in Asia over the last decade. The tables show us that the Taliban related events are not only some of the deadliest in Asia (with two over 1,500 victims), but constitute the outliers that would bring the total death counts much higher than the density plots would have us otherwise believe.

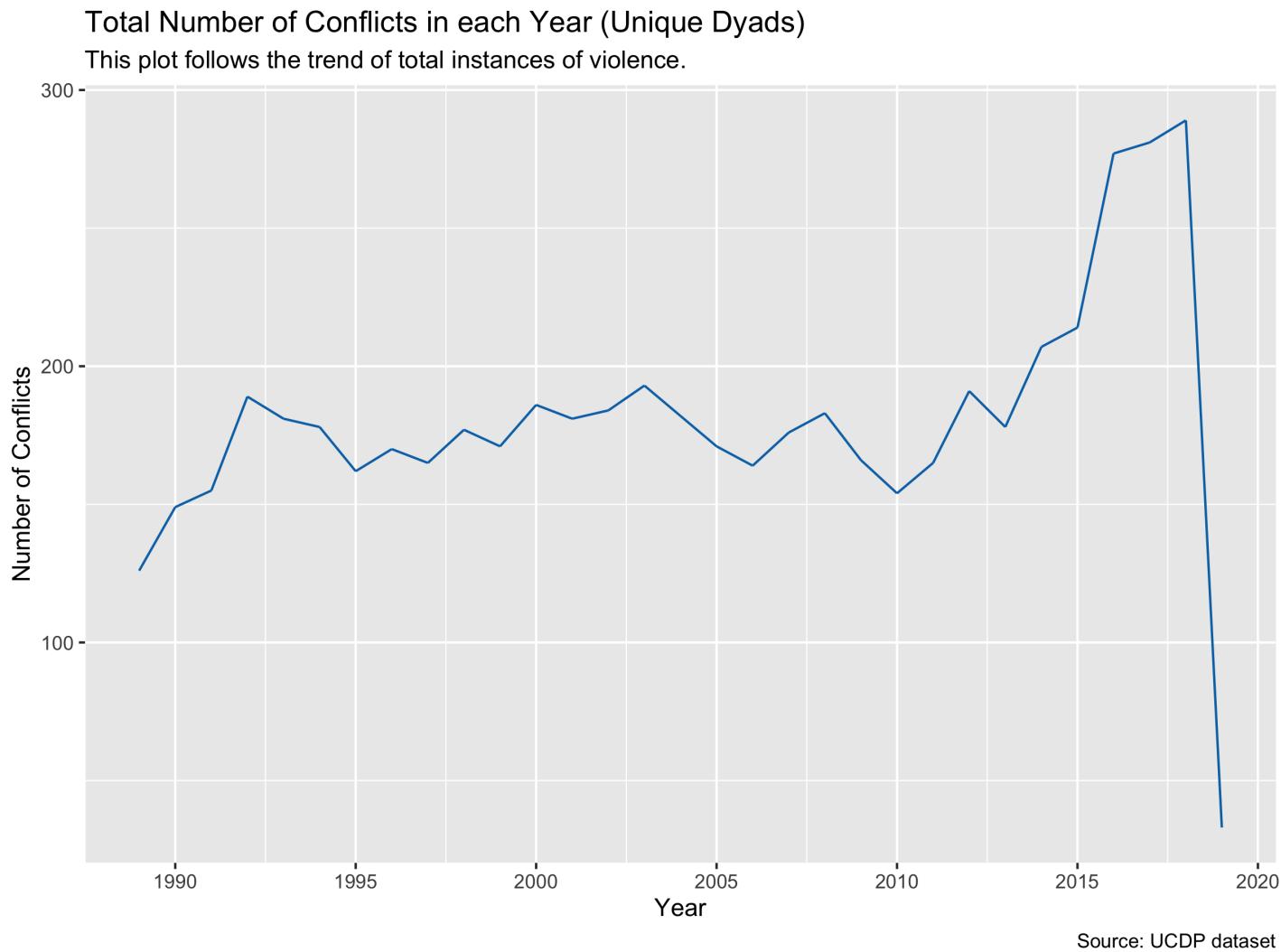
In this way, our hypothesis has some merit, as long as we qualify our response to include specific major conflicts that can bring regional totals up, even as national means and medians within each region is trending down.

Question 3: “Which regions have seen the highest growth in number of violent conflicts in recent years? Have any regions seen a reduction in conflict?”

In our previous two questions we looked at numbers of instances of violence (i.e. observations in our dataset), and number of fatalities (i.e. the `bषt` variable from our dataset). In this question, we specifically wanted to look at total conflicts, meaning numbers of pairings of identifiable groups (i.e. the `conflict_new_id` in our dataset). We want to look across the regions to see which way each is trending. Will this match up with our previous analyses?

Hypothesis: “There is likely to be a noticeable increase in number of conflicts in the Middle East and Africa. Regions in the Americas and Asia may have seen a relative reduction in conflict in the past decade.”

Again, our hypothesis was formed mostly from general perceptions and assumptions about global conflict. With so much news coverage in the past decade of the War on Terror, we assume the Middle East and Africa will have an increase in conflicts, while other regions may have a relative reduction in number of conflicts. As always, we started with an overview of global numbers:



This plot mirrors our total instances and fatalities, though most closely in the spike in the last decade. The number of conflicts shows a much flatter trend between the 1990s and 2000s. From 2010 onward, we have only seen one dip, otherwise the trend has been up. 2019, as accounted for already, is probably incomplete at the time of publishing of this iteration of the dataset. What will the regional data look like for conflict numbers?

Number of Conflicts per Year

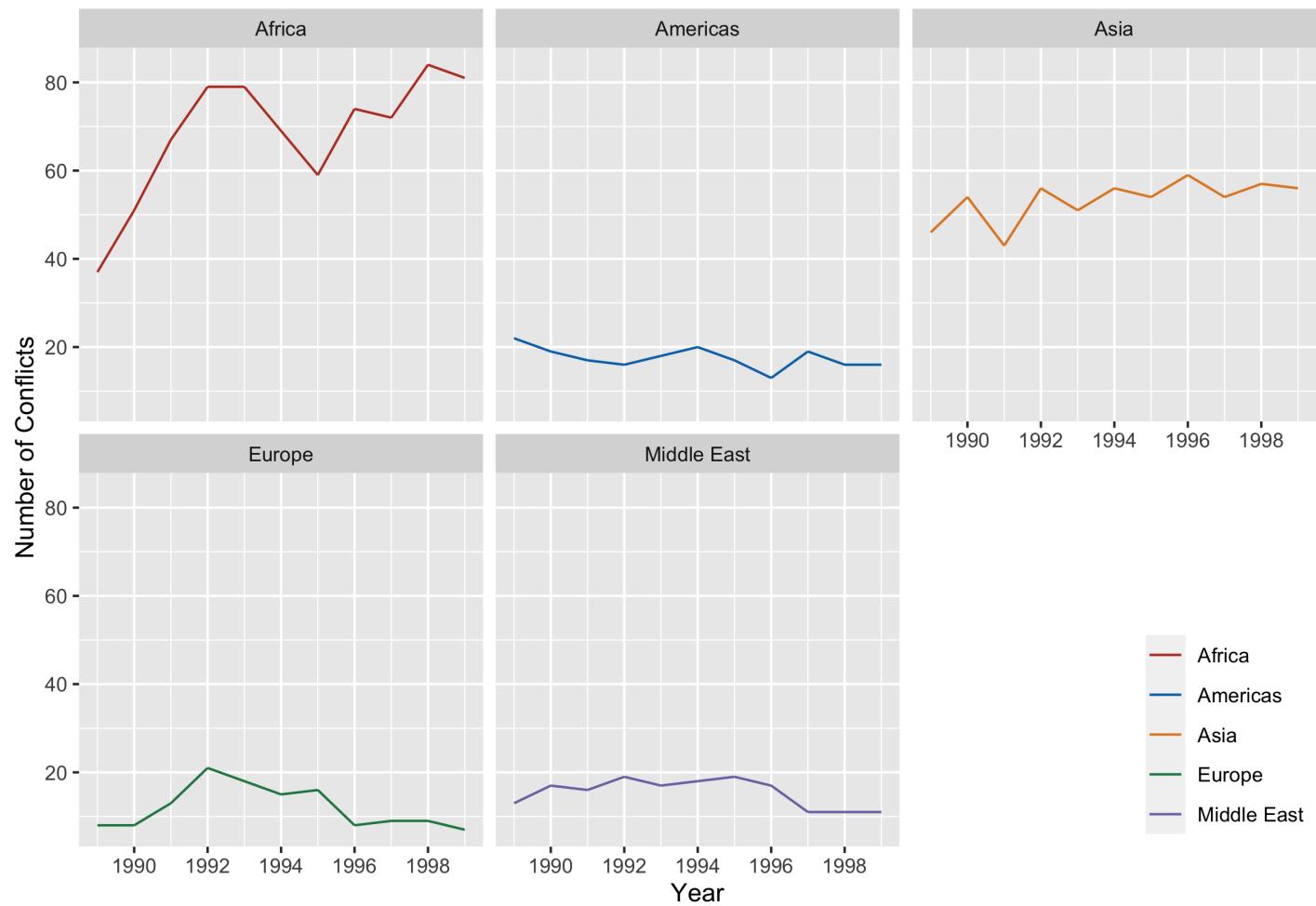
Given in Log Base-10 Form and Separated by Region



Source: UCDP dataset

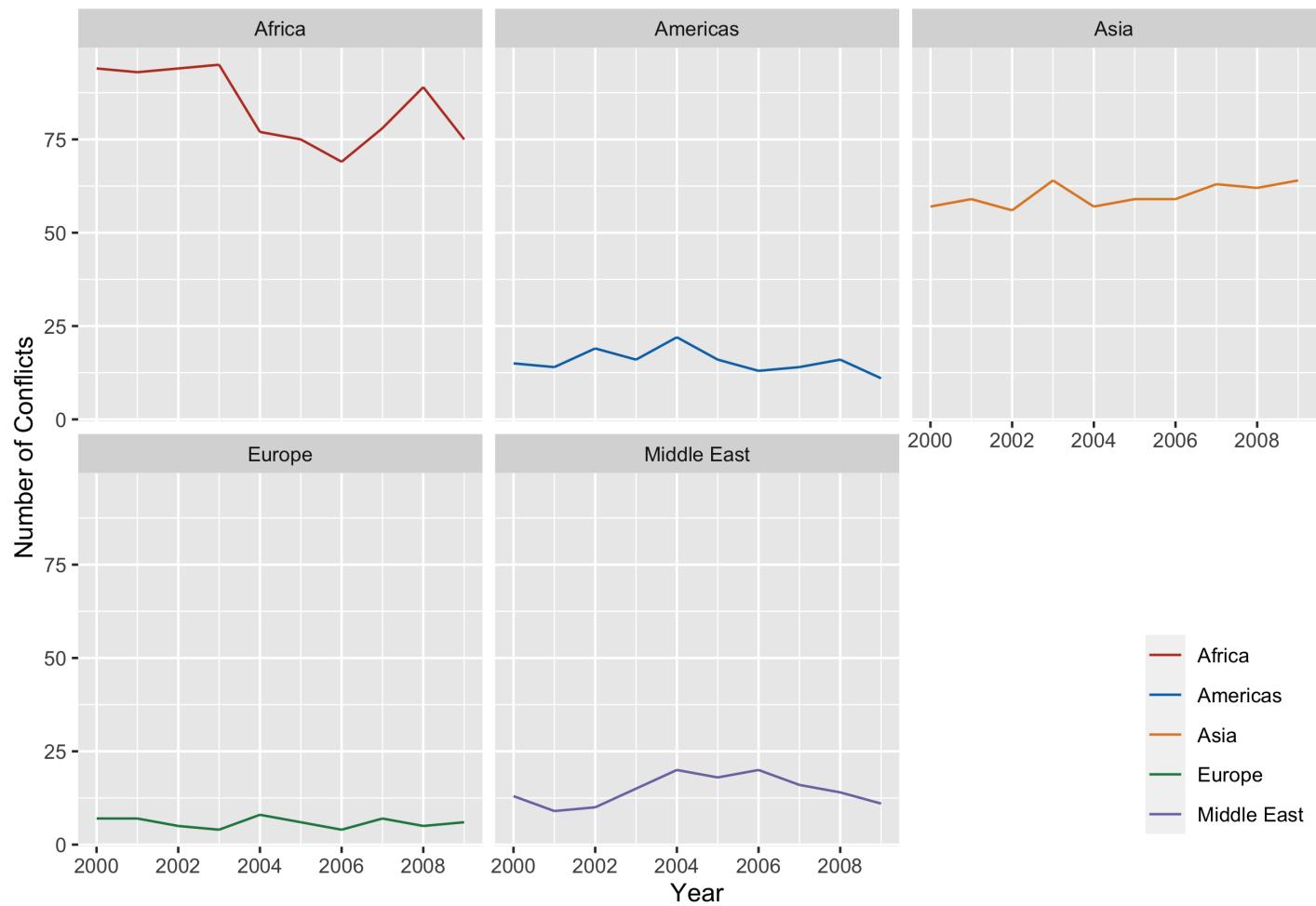
As predicted by our hypothesis, the Middle East and Africa have observed an upward trend in number of conflicts, and Asia and Europe have seen a slight decrease in the last ten years. In our hypothesis we referenced Latin America. If we take this to mean countries where Spanish, Portuguese, and French are primary languages, we can conclude that our datasets Americas region is analogous for Latin America (Canada and the United States account for a *total* of 40 entries in the dataset of about 170,000). That being said, the last ten years have seen a slight increase in the number of conflicts seen in the Americas (further discussion will take place in our last question).

Total Number of Conflicts in each Year 1990s

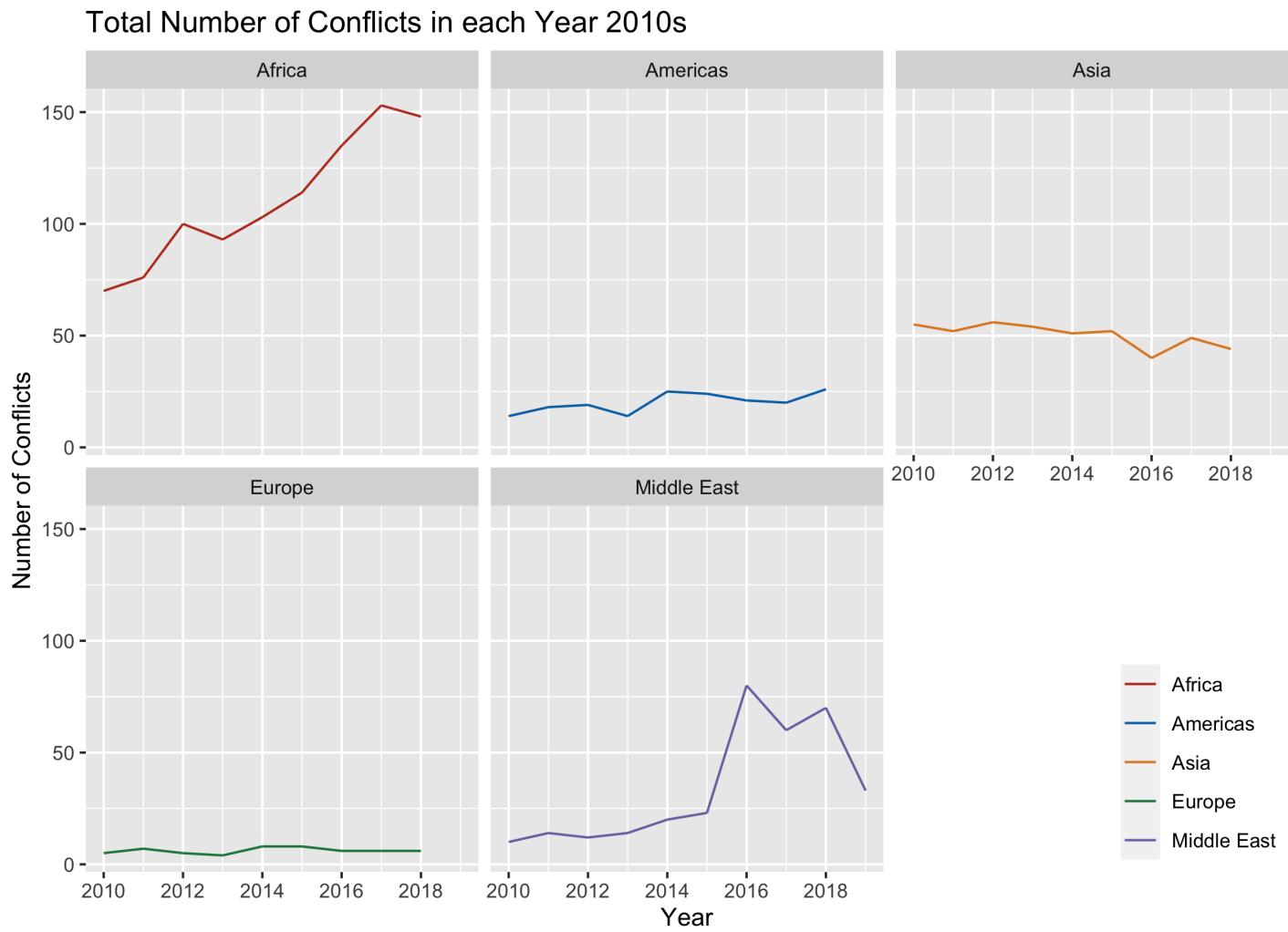


Source: UCDP dataset

Total Number of Conflicts in each Year 2000s



Source: UCDP dataset



Source: UCDP dataset

Further investigation by region and decade doesn't yield additional insight other than the apparent rise in the initial region graph is a result of the log-scale applied to the Y-axis. Our hypothesis remains intact, though further analysis is warranted to see if the trends for number of conflicts and numbers of instances or deaths diverge at a more granular level.

Question 4: “In the regions with the highest growth rate in violent conflict, which conflict type has seen the greatest increase in the past decade?”

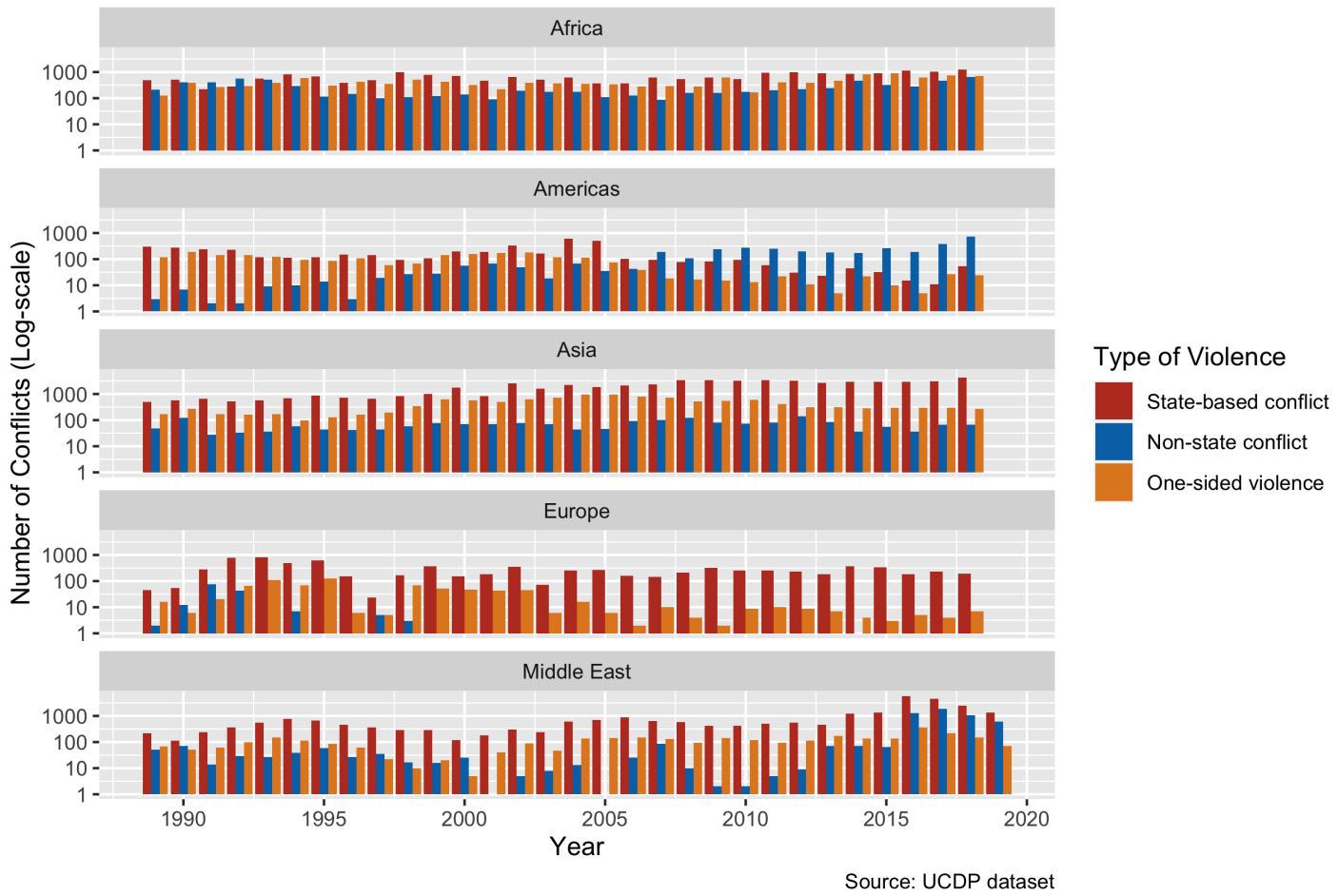
In our final question, we wanted to investigate type of violence. Our dataset defines three types of violence: state-based, non-state based, and one-sided violence. We have previously discussed state vs non-state violence when it comes to the aggressor, and the codebook defines one-sided violence as any action against non-combatants (labeled as “Civilians” in the `side_b` variable). In forming our hypothesis, we are primarily interested in non-state violence.

Hypothesis: “It is likely that non-state actor conflict is likely to be a major contributor to overall growth rates in violent conflicts.”

To test this hypothesis, we want to see if there are any trends in the different types of conflicts:

Number of conflicts by region and type of violence (1989-2019)

State-based conflict remains the modal type of conflict in Africa, Asia and the Middle East.



Source: UCDP dataset

Having discussed trends in Africa and the Middle East, let's take a closer look at the Americas. This region stands out in this graph because the non-state conflicts show a steady rise (evident in the blue bars). The increases in the Middle East and Africa are also apparent, but in those regions, instances of violence as a whole are on the rise. Line graphs help differentiate the trends for each type of violence:

Number of conflicts by region and type of violence (1989-2019)

Trends in conflicts by type seem somewhat stable in Africa and Asia, while there has been a dip in the Middle East over the past 5 years. However, we can see a clear upsurge in non-state conflict in the Americas



Source: UCDP dataset

Taking a look at the Americas graph, we can definitely see a major shift in type of violence, where in Africa and the Middle East, the rise in each type of violence seems to be proportional (even if to a great magnitude as in the Middle East). This plot confirms our assessment. Isolating the non-state conflict data yields the following graph:

Non-State Violence by Region (1989 - 2019)

The rise in non-state violence in the Americas is of particular interest

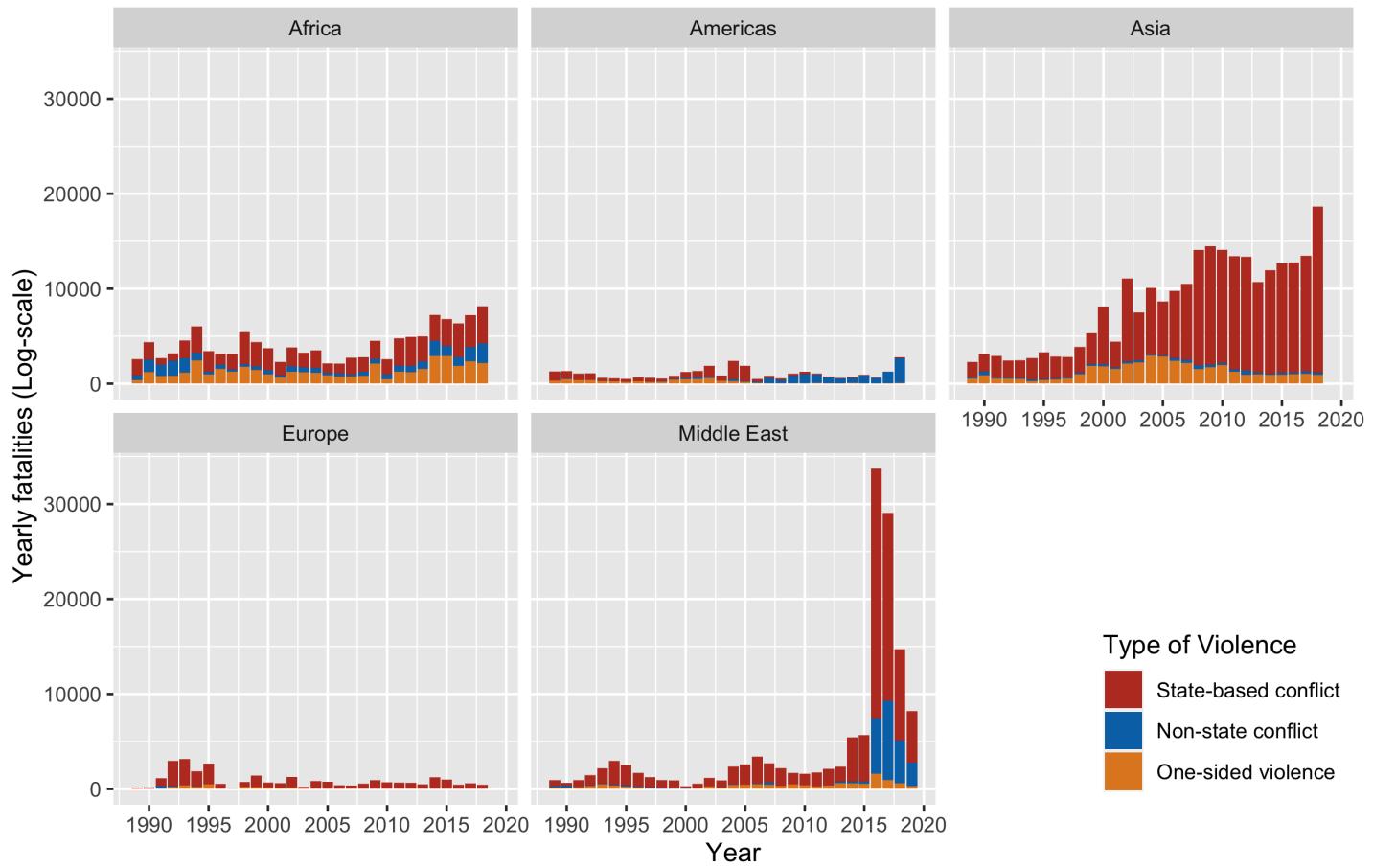


Source: UCDP dataset

Again, the counts of instances of violence have increased for non-state based violence in the Americas, Africa, and the Middle East. To see how this trends for fatalities, we can graph total fatalities by region and type of conflict:

Fatalities by region and type of violence (1989-2019)

While we can see an upsurge in non-state conflicts in three world regions, it has become the most prevalent type of conflict in the Americas.

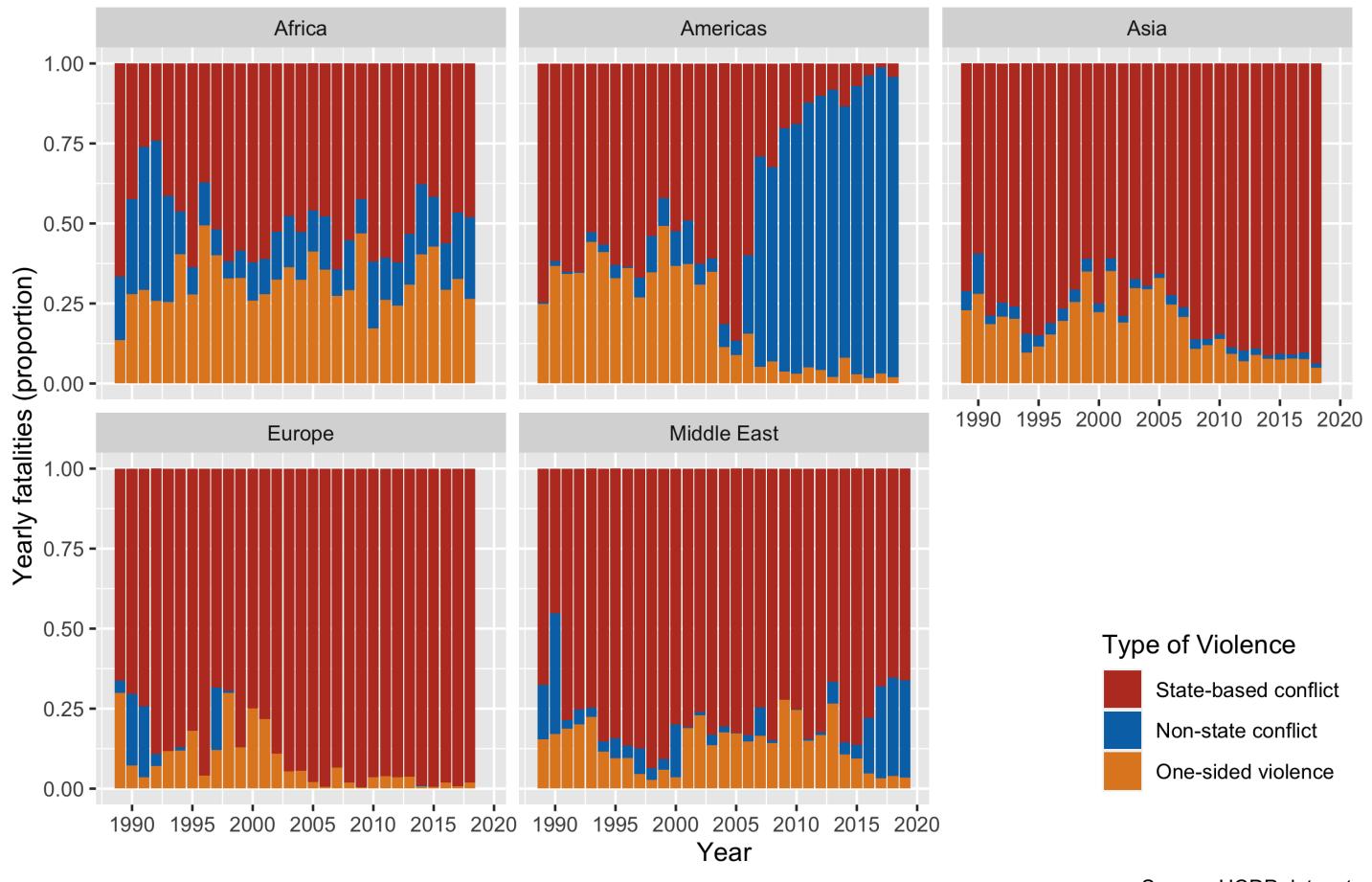


Source: UCDP dataset

The Americas is distinct from the other two regions in that since 2007, non-state based violence accounts for nearly all of the deaths in the region:

Fatalities by region and type of violence (1989-2019)

While we can see an upsurge in non-state conflicts in three world regions, it has become the most prevalent type of conflict in the Americas.



Outside of the Americas and the Middle East, non-state based violence has not increased in its proportion accounting for deaths globally. So while our initial hypothesis could be correct, it would be wrong to say that non-state action was a **major** contributor to the increase in total deaths. As stated above, the Americas take a drastic departure from other regions in its proportion of deaths from non-state action. Let's drill down on it.

Drill-down: What has caused such an increase in non-state based violence and deaths in the Americas in the 2010s?

In order to find out what is responsible for the increase, we wanted to look at how many fatalities per year can be attributed to which dyad pairs. When we look at the following table, we can see there is a large number of drug cartels involved in the deadliest years.

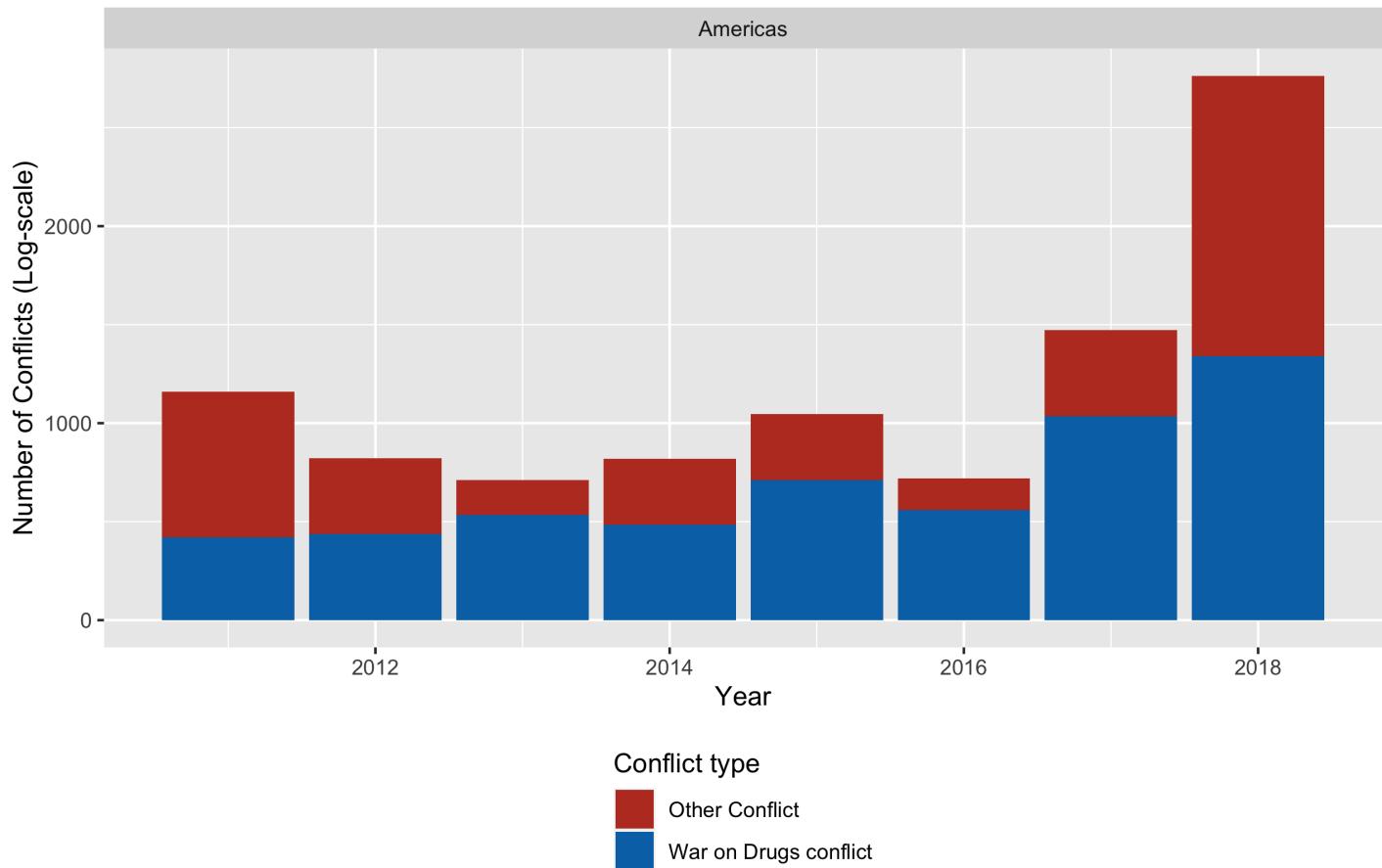
The Top 15 Deadliest Years by Dyad Pairing (2010s in the Americas)

Year	Dyad Name	Country	Total Fatalities per Year
2018	Jalisco Cartel New Generation - Santa Rosa de Lima Cartel	Mexico	4903
2018	Jalisco Cartel New Generation - Sinaloa Cartel	Mexico	4903
2018	Jalisco Cartel New Generation - Santa Rosa de Lima Cartel	Mexico	4903
2018	Jalisco Cartel New Generation - Sinaloa Cartel	Mexico	4903
2018	Sinaloa Cartel - Tijuana Cartel	Mexico	4903
2018	Jalisco Cartel New Generation - Santa Rosa de Lima Cartel	Mexico	4903
2018	Los Aztecas - La Linea	Mexico	4903
2018	Jalisco Cartel New Generation - Nueva Plaza Cartel	Mexico	4903
2018	Jalisco Cartel New Generation - La Nueva Familia	Mexico	4903
2018	Los Aztecas - La Linea	Mexico	4903
2018	La Union de Tepito - Fuerza Anti Unión	Mexico	4903
2018	Juarez Cartel - Sinaloa Cartel	Mexico	4903
2018	Los Aztecas - La Linea	Mexico	4903
2018	Jalisco Cartel New Generation - Sinaloa Cartel	Mexico	4903
2018	Jalisco Cartel New Generation - Los Zetas	Mexico	4903

In order to get a closer look at the relevant data, we wanted to create a subset of data that attempts to extract observations specifically related to drugs. Using regular expressions and `str_detect`, we drilled down and produced about 2,700 cartel related observations (in the Americas in 2010s), with about 1,600 that are specifically drug related (while the regex was crafted as carefully as possible, it is a) not comprehensive and may have missed some cartel related cases and b) any attempts to widen the regex had diminishing returns by capturing non-cartel related murders, insurgencies, etc that fell outside of what we wanted to look at).

Number of War on Drugs conflicts compared to other conflicts (2010-2018)

Although drug conflicts are not the dominant form of conflict in 2018, they have clearly been on the rise in the past decade, as is the number of conflicts in general.

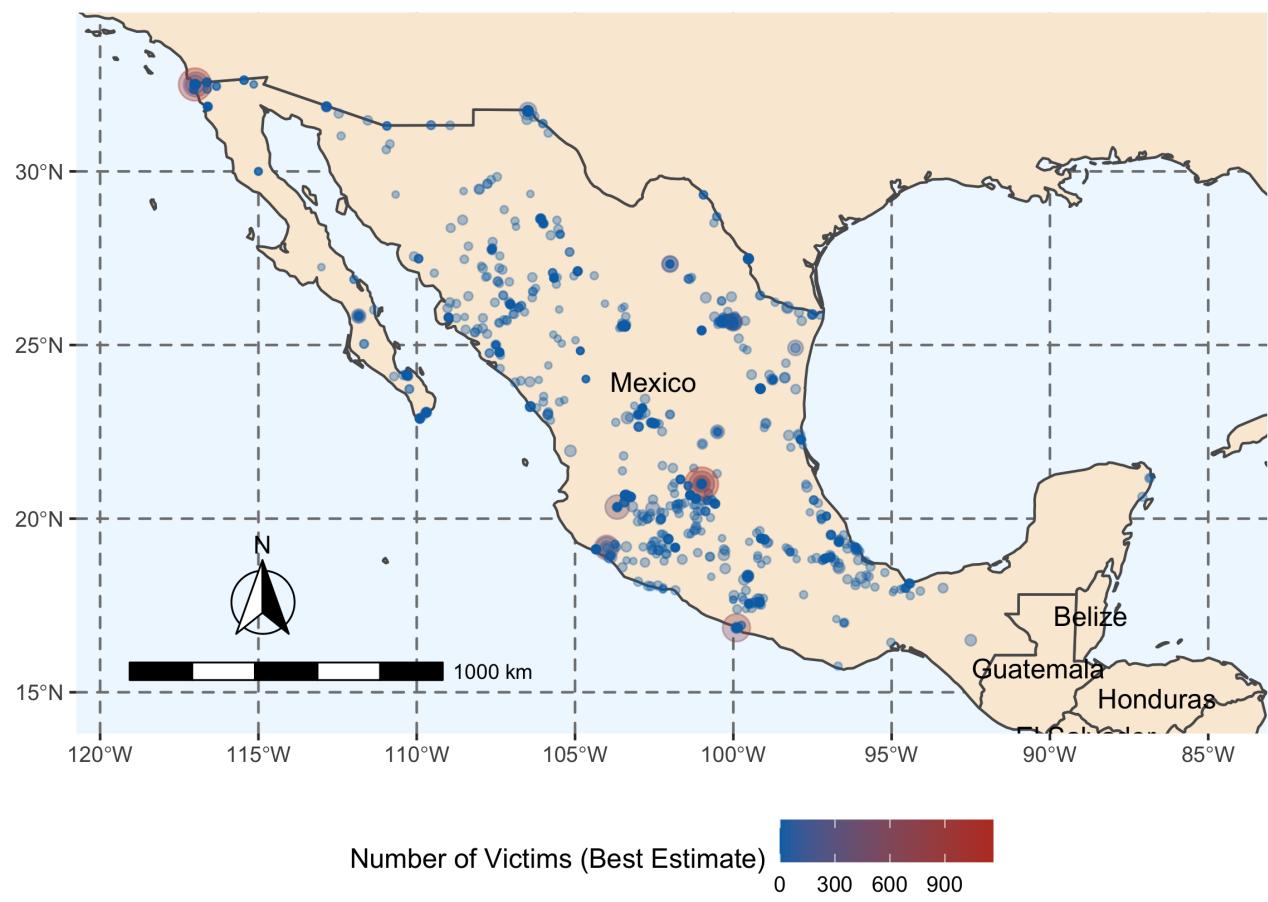


Source: UCDP dataset

This graph shows that from within those 2,700 observations, there has been an upward trend in the last ten years with drug related violence. Looking at our subset of data, almost all of the violence takes place in Mexico.

The War on Drugs (Mexico in the 2010s)

Cartels often perpetrate violence along the U.S.-Mexico Border to gain control of the illicit drug smuggling channels.



Source: UCDP dataset

While the violence is quite dispersed throughout the country, we can see that there is still considerable violence that occurs along the U.S.-Mexico border. We can hypothesize that cartels might fight each other over control of these illicit smuggling channels. They may also have no issue with perpetrating violence against any local authorities that try to stand in their way.

Again we wanted to highlight that our regex methodology is rather basic, and we can't say that the graphs are conclusive. But they did provide a potential glimpse as the state of violence in the Americas for the last decade.

In all, our initial hypothesis has to be moderated; worldwide, non-state violence hasn't greatly increased the fatalities in a given year. We can say this is true specifically about the Americas, mainly due to cartel violence.

Final Thoughts

Conflict studies is incredibly complex. There are so many factors to consider when making any sort of assumptions or hypotheses about specific topics or regions. The UCDP dataset is remarkably robust, though it falls a bit flat for near-term analysis, given the lag in reporting. It also allows us to check our own assumptions about the state of violence in the world. Media has a tendency to focus on certain topics for a time, but the public attention span seems to fall off quickly as news coverage decreases. With this constant cycling, it's easy to imagine that the general population may have some incorrect perceptions of conflicts around the world.

References

Uppsala Conflict Data Program (UCDP): <https://ucdp.uu.se/> (<https://ucdp.uu.se/>)

Sundberg, Ralph, and Erik Melander, 2013, "Introducing the UCDP Georeferenced Event Dataset", Journal of Peace Research, vol.50, no.4, 523-532

Högbladh Stina, 2019, "UCDP GED Codebook version 19.1", Department of Peace and Conflict Research, Uppsala University

World Bank Databank:

Individuals using the Internet (% of population) (<https://data.worldbank.org/indicator/IT.NET.USER.ZS?view=chart>)

Mobile Cellular Subscriptions (<https://data.worldbank.org/indicator/IT.CEL.SETS?view=chart>)

Population, total (<https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>)

GINI index (World Bank Estimates) (<https://data.worldbank.org/indicator/SI.POV.GINI?view=chart>)

GPD per capita (constant 2010 US\$) (<https://data.worldbank.org/indicator/NY.GDP.PCAP.KD?view=chart>)

Coding references:

Mapping Data (<https://www.r-spatial.org/r/2018/10/25/ggplot2-sf.html>)

Graphics Design (<https://www.r-graph-gallery.com/index.html>)

Matching WB and UCDP country codes (https://stats.andrewheiss.com/donors-ngo-restrictions/01_get-merge-data.html#conflict), coding (<https://cran.r-project.org/web/packages/countrycode/countrycode.pdf>)

Session Info

```
sessionInfo()
```

```
## R version 3.6.3 (2020-02-29)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Mojave 10.14.6
##
## Matrix products: default
## BLAS:    /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRblas.0.dylib
## LAPACK:  /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics   grDevices  utils       datasets   methods    base
##
## other attached packages:
##  [1] magick_2.3           ggspatial_1.1.1      countrycode_1.1.2
##  [4] tinytex_0.22          ggsci_2.9            scales_1.1.0
##  [7] rnaturalearthdata_0.1.0 rnaturalearth_0.1.0 sf_0.9-2
## [10] data.table_1.12.8     magrittr_1.5         kableExtra_1.1.0
## [13] forcats_0.5.0         stringr_1.4.0        dplyr_0.8.5
## [16] purrrr_0.3.4          readr_1.3.1         tidyverse_1.0.2
## [19] tibble_3.0.1          ggplot2_3.3.0        tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
##  [1] httr_1.4.1           splines_3.6.3        jsonlite_1.6.1      viridisLite_0.3.0
##  [5] modelr_0.1.6          assertthat_0.2.1     sp_1.4-1           highr_0.8
##  [9] cellranger_1.1.0      yaml_2.2.1           pillar_1.4.3       backports_1.1.6
## [13] lattice_0.20-41       glue_1.4.0           digest_0.6.25      rvest_0.3.5
## [17] colorspace_1.4-1      Matrix_1.2-18        htmltools_0.4.0    pkgconfig_2.0.3
## [21] broom_0.5.6          haven_2.2.0          webshot_0.5.2     processx_3.4.2
## [25] mgcv_1.8-31          generics_0.0.2       farver_2.0.3       ellipsis_0.3.0
## [29] withr_2.2.0          cli_2.0.2            crayon_1.3.4      readxl_1.3.1
## [33] ps_1.3.2             evaluate_0.14       fs_1.4.1           fansi_0.4.1
## [37] nlme_3.1-147         xml2_1.3.2          class_7.3-17      tools_3.6.3
## [41] hms_0.5.3            lifecycle_0.2.0     munsell_0.5.0     reprex_0.3.0
## [45] callr_3.4.3          compiler_3.6.3      e1071_1.7-3       rlang_0.4.5
## [49] classInt_0.4-3        units_0.6-6          grid_3.6.3         rstudioapi_0.11
## [53] labeling_0.3          rmarkdown_2.1        gtable_0.3.0      DBI_1.1.0
## [57] R6_2.4.1              lubridate_1.7.8      knitr_1.28         rgeos_0.5-2
## [61] KernSmooth_2.23-17    stringi_1.4.6       Rcpp_1.0.4.6      vctrs_0.2.4
## [65] dbplyr_1.4.3          tidyselect_1.0.0     xfun_0.13
```