

## **Refugees Welcome? A Dataset on Anti-refugee Violence in Germany**

### **Research Introduction and Motivations**

The purpose of the study is to interpret the drivers and consequences of right-wing violence due to the rise of xenophobic attacks against refugees in Germany. The number of refugees seeking asylum in Germany, as well as the number of xenophobic attacks against refugees has been on the rise. There is a need for systematic data collection and data processing in order to examine a broad range of research questions regarding xenophobic extremism in Germany. Research questions that we could answer are: *What geographic locations within Germany have a high number of attacks against refugees? What are the most frequent types of attacks? When are these attacks happening?* We can also uncover correlations or patterns in the data that may lead to a better understanding of this issue. We can ask: *Is there a correlation between the type and time/date of attack, such as events that precede or follow the attack? Is there any correlation between the type of attack and the location of attack?*

In 2015, several European Union states began to close their borders to refugees; however, German chancellor Angela Merkel stated that Germany's borders would remain open to all refugees. The increasingly stringent policies on refugee immigration in other EU states could be a potential reason as to why there was an increase in refugees who sought asylum in Germany. In 2015, 48% of the asylum applicants were Syrian refugees, a majority of whom were fleeing the government and Islamic State, while 21% of the arrivals were Afghan refugees who were fleeing because of the resurgence of Taliban control. A new right wing anti-immigration party (afD) developed in Germany in opposition to the current 'open-arms policy' and there was a noticeable increase in attacks against refugees. For example, on March 23, 2015, a group of attackers attempted to set fire to a school that housed refugees in Berlin-Kreuzberg. Another instance that occurred was in Hauseberg/Porta Westfalica in which people painted swastika symbols on a refugee shelter. Given the rise of the right wing anti-immigration party, the spread of anti-refugee sentiment in German politics, and the increase in attacks against refugees, there is a need for systematic data collection and data processing to understand this social and political phenomenon. This was the underlying motivation for the creation of the ARGIV dataset. This dataset categorizes four types of right-wing political violence and social unrest, and shows us the frequency, geographic location, and time of these events.

### **Stakeholders**

Stakeholders are individuals or entities who are interested in the research project and actively affect or affected by the outcomes of the research project. In this project there are seven key stakeholders: the Amadeu Antonio Foundation, PRO ASYL, MGRG, refugees seeking asylum in Germany, German political parties, German citizens, and other researchers.

The Antonio Amadeu Foundation is a civil society organization with the goal of strengthening German civil society activism against right-wing extremism, racism, and anti-semitism. This foundation is pairing up with PRO ASYL, another civil society organization supporting human-rights activism. Both of these organizations are part of the largest, most respected pro-immigration advocacy organizations. The Antonio Amadeu Foundation and the PRO ASYL are both key stakeholders, as they are responsible for collecting information and data for the research, which means that they will be affecting the outcomes. Their motivation is to combat political extremism and xenophobia as this aligns with each organization's values and mission. MGRG (*Mut gegen rechte Gewalt*) is an online chronicle that publicizes all instances of right-wing attacks directed at refugees, and their source of data is from the Antonio Amadeu foundation and PRO ASYL. Their motivation is to inform and educate German citizens about the attacks. Since MGRG has a pro-immigration stance, they may want to persuade people's attitudes to be accepting of refugees.

Refugees are also stakeholders as they are being targeted by extremist right-wing parties and are the victims of these political attacks. They may also have an interest in future research done using the dataset as it can potentially influence refugee and violence-related policies.

German citizens are stakeholders because they provide support to German political parties, and determine who is in power. They have an interest in this research as it informs them of the political climate. This may persuade them to oppose the anti-immigration party. On the other hand, it may have the opposite effect and induce fear in them. If they see that there is a rise in political violence and social unrest, there is may begin to favor the anti-immigration party.

German political parties are stakeholders as they foster the atmosphere of the political climate. Political parties may be motivated to use this research in order to support their stance on immigration policy. They can use this to leverage influence over German citizens and their support of certain policies. The more support a political party has, the more likely they are to be in power.

Other researchers are potential stakeholders, as they have this new dataset to further explore. This dataset was created to be "scientifically useable", which means that other researchers now have data to work with. These individuals affect the outcome of the research as they can shape the direction of future studies. These future studies can thereby inform policy and people's views regarding immigration.

## **Data Description**

The Antonio Amadeu Foundation and PRO ASYL collected data regarding the nature of the anti-refugee attack, as well as the federal state and location in which the attack occurred. The twelve-digit community identification number, as well as the exact time of the event, were taken from "official statistics". *Mut gegen rechte Gewalt*, a weekly magazine, publishes online chronicles using the data collected, which are then made available to the public. We formed partnerships with them to create the ARGIV dataset that includes a total of 1645 events of four different types of right-wing violence (2014-2015): xenophobic demonstrations, assault, arson

attacks, and miscellaneous attacks against refugee housing. The study focuses on perpetrators of right-wing violence in Germany by analyzing police data on biographical, socio-demographic characteristics, and public opinion polls. There were no issues gaining access to the data, since it is already publicly available.

Additionally, those of us who were involved in creating a scientifically useable dataset from the information in this chronicle did not face any serious collaboration issues. There were no issues in regards to sharing data. It was a highly successful collaboration as we were able to create a scientifically useable data set that all future parties can benefit from. A potential issue could be failing to share credit where it is due, but in the article, we clearly state where the data came from (MGRG), as well as how MGRG obtained it (Antonio Amadeu Foundation and the PRO ASYL). Another potential issue could be any issues that arise when making the data scientifically useable. These could be issues related to upholding the transparency and quality of the data. However, there is confidence in the quality and transparency of the data collection as these are credible organizations that belong to the largest and most respected immigration advocacy organization and work closely with international human-rights organizations. Most likely, there are not any confidentiality issues at play since this data is public information accessible by media. The data also does not contain any personal information on people. A potential problem that could arise is showing people where refugees are located since they report on attacks on refugee housing. This may lead to a rise in this type of attack, or increase discriminatory acts against these communities. IRB approval was not needed because the data comes from public attacks that happened, i.e. information such as the date, location, and category of attack is included. These variables do not include identifiable private information of human subjects.

### **Data/Policy Fit**

There are a number of disconnects and limitations in the dataset. The data is subject to high bias because it is pulled from a web scraped MGRG online chronicle where information on assault is only recorded if the victim has a *refugee status*. This means that assault on, for instance, left-wing and pro-refugee protesters, volunteers helping incoming refugees or journalists covering xenophobic rallies, are not recorded.

This also correlates to an underreporting of demonstrations and assaults. The MGRG chronicle points out that because anti-refugee demonstrations and rallies have been on the rise in recent years, it is impossible to collect information on every single one of them. Another reason for underreporting of results is that the Amadeu Antonio Foundation and PRO ASYL only report demonstrations that specifically disregard German law. This includes illegal demonstrations not registered with the authorities beforehand, demonstrations that included assaults against journalists and/or police or demonstrations in which hate speech was reported. The actual number of assaults, irrespective of the victim's status, is likely to be higher than is reported in the chronicle – a problem faced by many criminal statistics.

There is a small number of events in the dataset that are not categorized because they do not belong to any of the four basic event types. Examples of these include the distribution of

xenophobic leaflets or public banners with right-wing extremist slogans. Thus, when we are analyzing by type of attack (i.e. seeing which attack is the most common), we are unsure of the attacks that are uncategorized, which means we cannot conduct further analysis on them. The report and data reveals no indication of any information on these attacks.

The attacks are based on a variety of sources – public reporting in newspaper, press releases by the German police, parliamentary interpellations and publicly accessible reports by local and regional organizations offering consultation for victims for right-wing violence. By including events from the broad set of categories, the chronicle, as well as the dataset, cover a wide range of anti-refugee violence. This allows researchers to distinguish between the types of xenophobic attacks and study degrees of severity. Because of the systematic data collection and processing of anti-refugee events in Germany, we are able to conduct a thorough quantitative investigation of the phenomenon including the patterns, dynamics, drivers, and consequences. In this regard, the dataset can be used to answer research questions that deal with the escalation of xenophobic extremism.

Using a broad definition of what the MGRG describes as “right-wing violence” allows us to distinguish between and include different types of xenophobic attacks in terms of degrees of severity in our dataset. Because of the broad nature of the study, we can answer general questions on xenophobic extremism and the escalation of this phenomenon overtime in Germany as opposed to purely focusing on certain types of assaults. Despite the underreporting of attacks, the diverse range of attacks in the dataset is representative of the population we care about. We can explore patterns and correlations from a criminology, sociology, political science or economic point of view.

The study presented here ties into a broader literature on how immigration is linked to the rise of right-wing extremism and xenophobia in the Western world. Ideally, determinants of anti-refugee violence as a predictor for future attacks is a causal interpretation we would like to achieve out of this study. Theories such as (youth) unemployment, the success of right-wing political parties, and media discourses have each been identified as strong predictors of violent outbursts. However, this study cannot be hard evidence for causal related interpretations between these determinants and right-wing extremism. A continued study year by year of anti-refugee attacks may confirm theories and whether different predictors have a stronger explanatory power. This dataset would be useful for scholars trying to do further exploration with systematic case selection such as trying to compare municipalities with high levels of anti-refugee violence and unrest with municipalities with low levels of such violence.

Scholars can also easily misuse or misinterpret the results of the study. Since the ARVIG data includes all ranges of severity of xenophobic extremism, scholars or journalists may include the results of the study on their reports as a statistic of the total number of attacks. Usually, we think of attacks as demonstrations that involve casualties through mass physical assault, thereby having a big influence on political sentiment. However, the dataset also includes trivial events such as posters involving hate speech. It is important that scholars or journalists specify the type of attack and clearly define what it constitutes before a statistic about anti-refugee demonstrations is included in a certain report. The results, depending on the context of the report, may be subject to either underreporting or overreporting of demonstrations.

## Reproducibility

\_\_\_\_\_Attached as an .RMD file.

## Modeling

We would imagine that preliminary findings that did not reveal much meaningful information or extract patterns from the data would not be included in the published version of this paper. For example, graphs that did not show significant differences in anti-refugee attacks among specific groups would not be important for the purposes of our study, as we strive to examine what factors are associated with the occurrence and frequency of attacks. Nonetheless, in our case, since our study was rather exploratory to begin with, our code did document the majority of the analyses that were actually run. For example, our code, as well as our paper, showed basic summary statistics, such as the frequency and types of attacks.

For our additional analysis, we wanted to include a larger sample of attacks over a longer, possibly more recent, period of time. The dataset utilized to conduct the initial study is regularly updated to reflect the latest information regarding anti-refugee attacks. It turned out that the most up-to-date version contains information about attacks all the way until November 2017. Therefore, our dataset now contained information about 6698 anti-refugee attacks spanning the time period from January 2014 to November 2017.

## Additional Analyses

Category <fctr>	N <int>
arson	296
arson & miscellaneous attack	2
assault	1004
demonstration	541
demonstration & miscellaneous attack	2
miscellaneous attack	4825
miscellaneous attack & assault	12
other	25

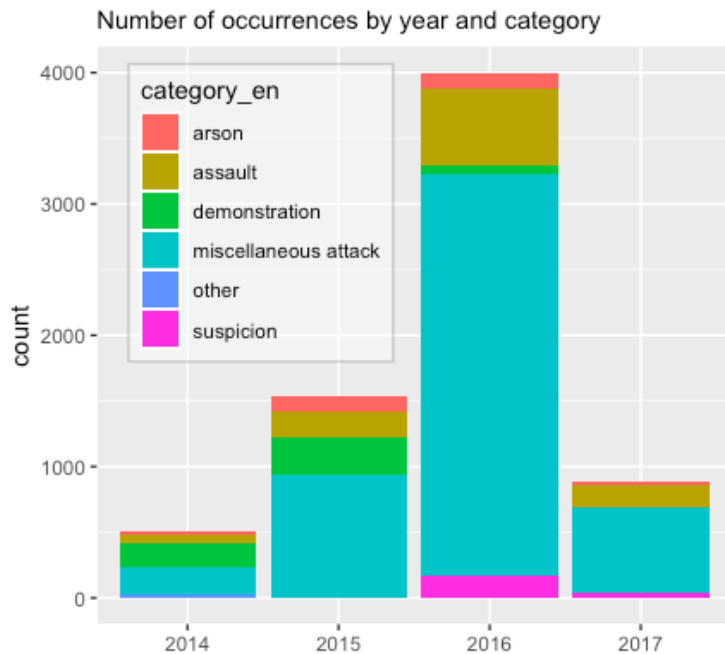
#1)

Category <chr>	N <int>
arson	157
arson & miscellaneous attack	8
assault	195
demonstration	443
demonstration & assault	8
demonstration & miscellaneous attack	16
demonstration & miscellaneous attack & assault	1
miscellaneous attack	763
miscellaneous attack & assault	29
other	25

The table on the left depicts the original analysis, while the table on the right depicts the original analysis with our supplementary analysis. Both tables represent the sheer amounts of the various types of anti-refugee attacks that occurred within their respective time periods. With

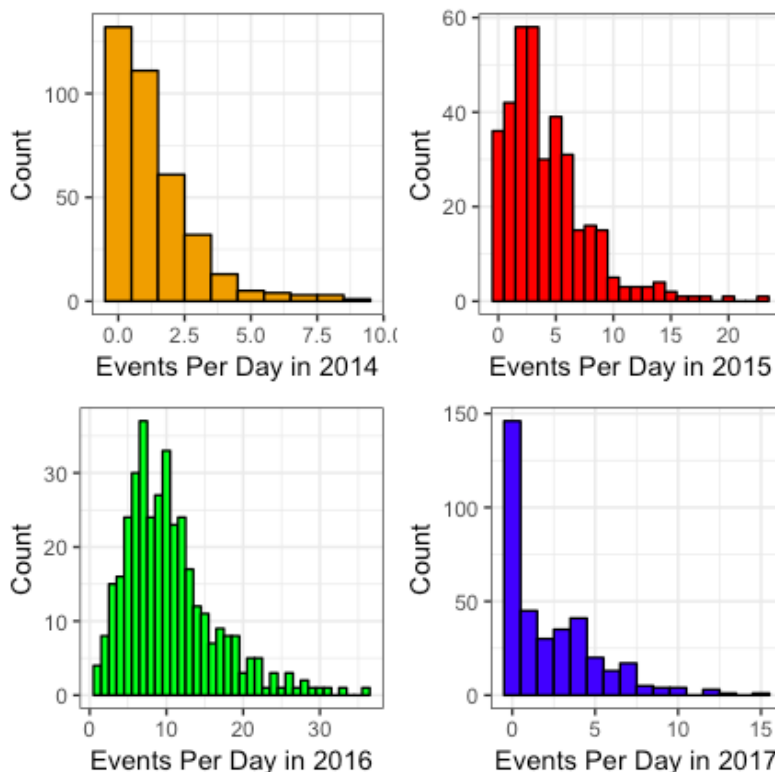
respect to the other categories, in both analyses, ‘miscellaneous attacks’ are the most frequent type of attack.

#2)



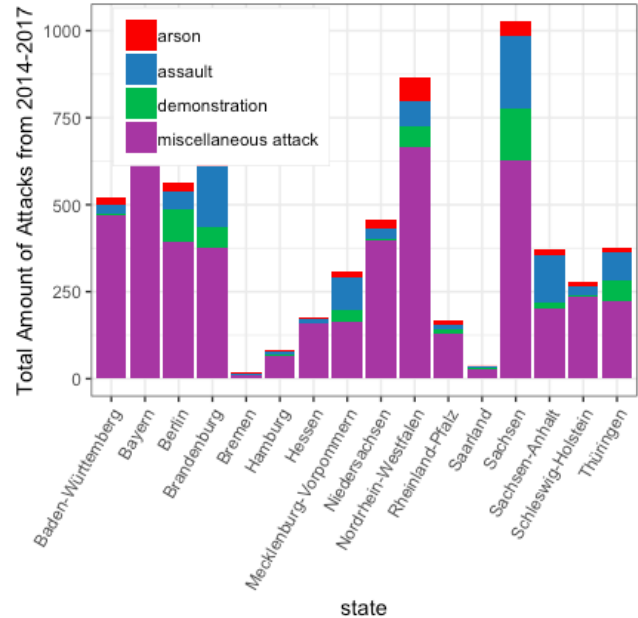
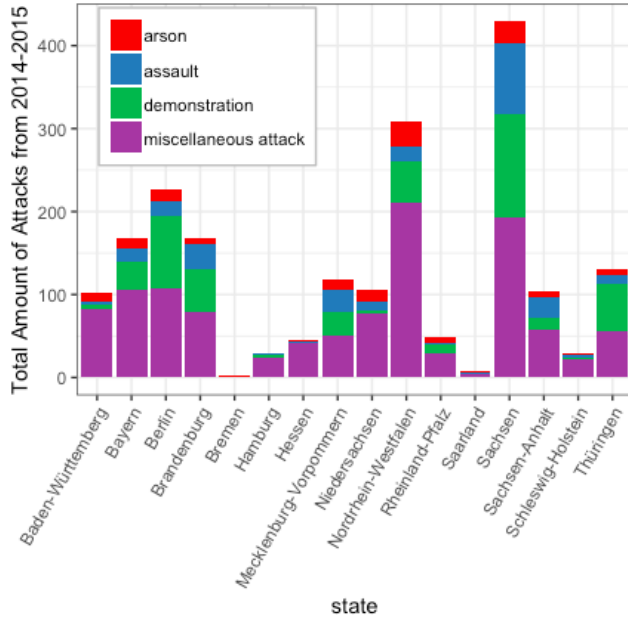
From this bar graph, we can see that violence and demonstrations against refugees soared in Germany during 2016. There is a sharp drop in the total number of attacks in 2017, which we can attribute to a drastic reduction in miscellaneous attacks. However, this additional analysis only includes data until November 2017, and there is a considerable time lag for when events are reported and included in the dataset. Therefore, as of right now, with the data we have, it would be inappropriate to make conclusions about 2017 in its entirety.

#3)



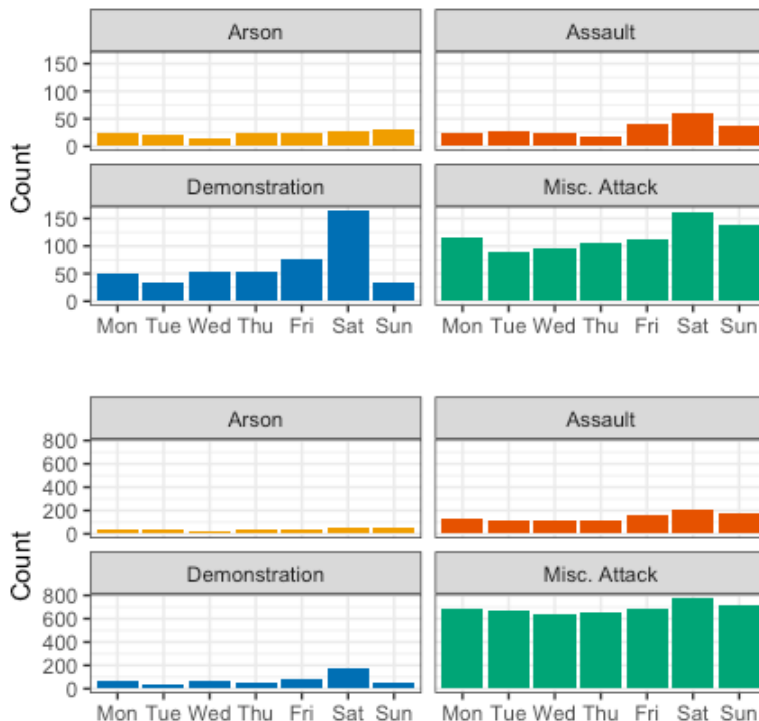
These four graphs show the daily frequency of anti-refugee attacks for every year in the dataset (2014-2017). These graphs depict the large surge in the mean amount of attacks that occurred in one single day between the years 2014 and 2016. The date for which the highest number of attacks was recorded is January 24, 2016. These attacks comprise of 25 miscellaneous attacks, 10 demonstrations, and 1 assault.

#4)



The bar graph on the left depicts the original analysis, while the bar graph on the right depicts the supplementary analysis. Both graphs represent the frequency of anti-refugee attacks across the sixteen German states. Some regions of Germany are more vulnerable to attacks than others.

This could be attributed to the fact that citizens' stance towards refugee differs across the nation of Germany. In both analyses, the German state of Sachsen has the greatest amount of attacks, while Bremen has the least amount of attacks.



#5)

The graph on the top depicts the original analysis, while the graph on the bottom depicts the supplementary analysis. Both depict the distribution of attacks by the day of the week. We observe that, generally, Saturday is the day in which attacks, regardless of type, occur most frequently.

## **Social Impact**

The results have a potential social impact on the political climate in Germany. The results from our analyses inform us about the frequency of attacks based on the time of year. Elections in Germany are held at certain times, so the political climate may shape attitudes towards refugees. Depending on the officials elected, their stance on immigration can determine the policies in place. Border policies and immigration laws implemented will directly affect refugees. The analyses also show us which communities have the highest frequencies of attacks. We also know the frequency of attacks in the four most populous German states. Nordrhein-Westfalen, one of the four most populous German states, has the highest number of attacks. This information can be turned into practice by increasing security measures in these communities where the most attacks happen, especially in Nordrhein-Westfalen where there are a lot of residents. Organizations who provide consultation to victims can also better prepare for how to help refugees with the trauma they may experience. If we know there are a lot of attacks in a certain area, organizations can make sure they are adequately staffed with appropriate personnel who can provide consultation. The information from our analysis can also help to predict the type of attack in each region. This can be turned into practice by creating security measures for certain types of attacks. It can also help us to understand what events may precede each type of attack which may aid in prevention measures. This can also inform policy. For example, if there are a high amount of arson attacks, policy regarding arson or gun sales can be put into place to reduce access.



## References

Benček, David, and Julia Strasheim. "Refugees Welcome? A Dataset on Anti-Refugee Violence in Germany." *Research & Politics*, vol. 3, no. 4, 2016, p. 205316801667959., doi:10.1177/2053168016679590.