A Spatial Analysis of the Target Demographics of Hate Crimes

Phoenix Nénar Williams

May 20, 2025

1 Introduction

Despite the ongoing efforts of Civil Rights, Women's Rights, Queer Liberation, and Black Lives Matter movements (among many others) to create a more just and equitable world where diversity is celebrated, people across our world and even within our city remain intolerant towards diversity, to the point of violence. From 2010 to 2022, nearly 600 Hate Crimes occurred in the City of Denver. In this project, we will examine patterns of spatial clustering within this dataset of Hate Crimes with cases (in our case-control point pattern analysis) delineated by the type of target group (the case status and the crime type).

To acknowledge the complexity of the issue of Hate Crimes, let us frame our analysis with the legal definition of a hate crime. Colorado law distinguishes Hate Crimes from other criminal activity with three criteria:

1. The accused person knowingly committed the crime.
2. The accused person intended to intimidate or harass the victim.
3. The accused person committed the criminal activity because of the victim's perceived or actual protected class or identity (i.e.: race, ethnicitiy, gender, sexual orientation, religion, disability, etc.).

How does one show the accused had the intention to intimidate or harass the victim? How do one tell if the victim was chosen randomly or specifically for their race? The complexity of proving a crime was a hate crime (in the legal sense at least) can be difficult, but understanding the requirements can help us understand the value and limits of our analysis.

The Denver City Open Data Catalog provided the dataset for this project, which consists of the following information about 567 distinct and relevant hate crimes that occurred from 2010 to 2022:

- Case Numbers
- Dates
- Deidentified location information
- Case Status
- Crime Type
- Target Groups

2 Methodology

We focused our analysis on the Target Group information, treating the data set as a case-control point pattern. To ensure a minimum degree of statistical validity, we grouped the 21 distinct target group values into the following 4 case groups:

1. Anti-LGBTQ+: This case group contains 177 data points labeled as Gay, Lesbian, Bisexual, Transgender, or Mixed group
2. Anti-Race/Ethnicity: This case group contains 293 data points labeled as American Indian/Alaskan Native, Arab, Asian, Black/African American, Hispanic/Latino, Multiple, Other, or White.
3. Anti-Religion: This case group contains 92 data points labeled as Catholic, Jewish, Muslim, Multiple, Other Christian, or Other.
4. Anti-Disability: This case group contains 5 data points labeled as Physical or Mental.

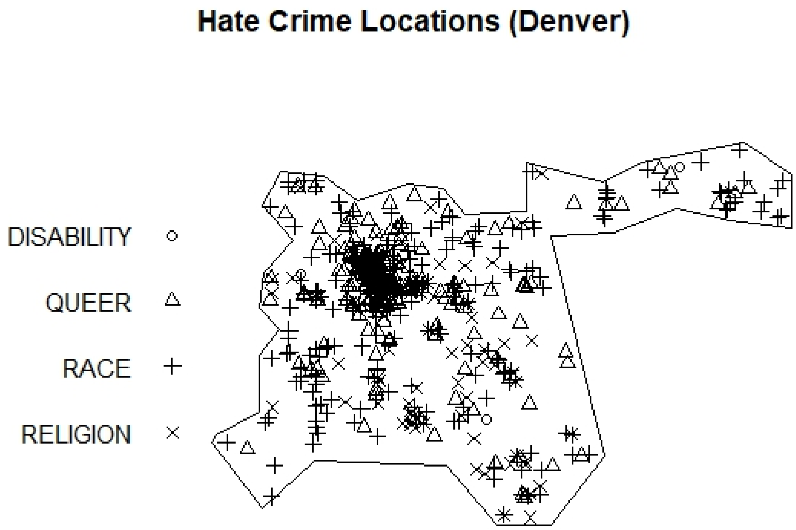


Figure 1: Preliminary Plot of Data Set by Case Group

We applied 4 different methods for detecting local clusters and global clustering in case-control point patterns, relative to the random labeling hypothesis. The random labeling hypothesis assumes the point pattern matches a heterogeneous Poisson Process where case and control events locations follow the same spatial point process and each location has a constant probability of being a case. By using the random labeling hypothesis, we can easily interpret the results of our hypothesis testing in the context of the data we are studying. The methods in the first category compare summaries of first-order or second-order properties for cases and controls, while the methods in the second category scan estimations of local rates of hate crimes. We iterated each of our methods over each case group serving as the cases in turn.

From the first category, we used the Log Relative Risk method, which compares the spatial density functions (a first-order property) of cases and controls with the test statistic:

$$\tilde{r_b}(s) = \log \frac{\tilde{f_b}(s)}{\tilde{g_b}(s)}$$

where $\tilde{f_b}(s)$ is the pointwise empirical estimate of the spatial density function of case locations and $\tilde{g_b}(s)$ is the pointwise empirical estimate of the spatial density function of control locations. Using Monte Carlo testing (simulating a large number of data sets under the random labeling hypothesis and calculating the proportion of total data sets with a test statistic at least as extreme as the observed data), we constructed 95% tolerance envelopes to observe local clusters of cases relative to controls or vice versa. Furthermore, by integrating the log relative risk test statistic over the study area (proposed by Kelsall and Diggle, 1995) and repeating Monte Carlo testing to find a p-value (with the standard significance bound of 0.05), we can identify if there is a pattern of global clustering in the sense of whether or not the spatial densities of cases and controls are equal for all locations in the study area. If the p-value indicates significance, we have evidence that the spatial densities of cases and controls differ for at least one location in the study area.

Also from the first category, we used the Difference Between K Functions methods which compares the spatial dependence or variance of cases and controls (a second-order property) to reveal global or local clustering relative to the size of the clusters. Using the edge-corrected K-function estimator:

$$\hat{K_{ec}}(h) = \hat{\lambda}^{-1}\frac{1}{N}\Sigma_{i=1}^{N}\Sigma_{j=1,j\neq i}^{N} w_{ij}^{-1}\delta(d(i,j) < h)$$

for the cases and controls separately and calculating the difference (denoted $KD(h)$). Using Monte Carlo testing on $KD(h)$ for a range of spatial lags h (from 0 to a maximum spatial lag h^*), we can create tolerance envelopes for that range of spatial lags. The spatial lag distances where the value of $KD(h)$ for our observed data exceeds the tolerance envelopes indicates clustering of cases relative to controls exists on the scale of that spatial lag (and vice versa for where $KD(h)$ falls below the tolerance envelopes). Furthermore, we can modify $KD(h)$ to test for global clustering with the test statistic suggested by Diggle and Chetwynd (1991):

$$KD_+ = \Sigma_{k=1}^{m}\frac{\hat{KD}(h_k)}{\sqrt{\mathrm{Var}[\hat{KD}(h_k)]}}$$

where h_k are the spatial lags used to estimate $KD(h)$. By applying Monte Carlo testing with this test statistic, we can check whether there is evidence of clustering of cases beyond controls at the spatial scales considered. If our Monte Carlo p-value indicates significance, we conclude there is at least one spatial lag scale for which we observe clustering of cases relative to controls.

In the second category, the spatial scan method compares local rates inside and outside a moving window to determine significant global clustering and significant most likely local clusters. For windows, we generate circles of variable radii (incrementing by increasing the size event by event until we reach the upper bound of half the width of the study area) around each observed event location. We denote:

- $N_{1,in}$ as the number of case locations inside the window,
- N_{in} as the number of people at risk inside the window,
- $N_{1,out}$ as the number of case locations outside the window,
- N_{out} as the number of people at risk outside the window,
- and w_i to index the windows considered.

Using the test statistic suggested by Kulldorff (1997),

$$T_{scan} = \max_{w_i}\left(\frac{N_{1,in}}{N_{in}}\right)^{N_{1,in}}\left(\frac{N_{1,out}}{N_{out}}\right)^{N_{1,out}} I\left(\frac{N_{1,in}}{N_{in}} > \frac{N_{1,out}}{N_{out}}\right),$$

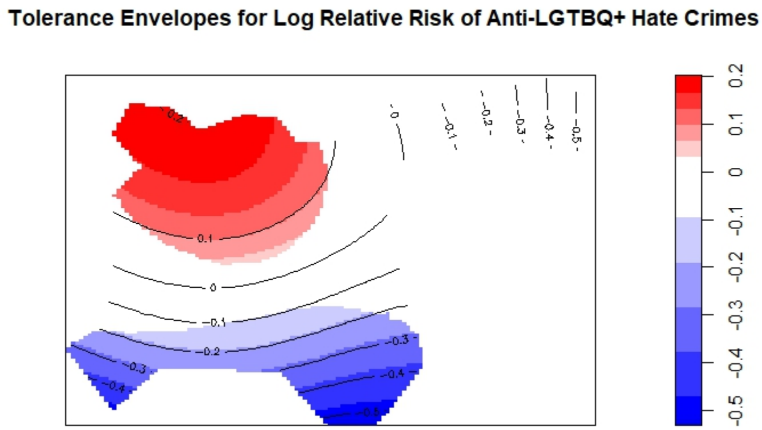
we apply Monte Carlo testing to determine if any window contains a most likely cluster more unusual than what is expected under the random labeling hypothesis. If our p-value indicates significance, then we have evidence of at least one cluster of cases in the study area, and we can determine the most likely cluster by the window which generated the largest test statistic. We can subsequently run Monte Carlo testing on all of the other windows to determine and hierarchically rank other most likely and significant clusters.

Finally, from the second category, we used the q nearest neighbors method, which checks for global clustering on specific scales of nearest neighbors and the contrasts between those scales. For a test statistic T_q, we set a particular value for q and count up the number of cases that occur within the q nearest events to other cases. We can apply Monte Carlo testing to find a p-value for each value of q we consider, indicating whether the number of cases within the q nearest neighbors of other cases is higher than what is expected under the random labeling hypothesis. Subsequently, we consider the differences in the count of T_q between different values of q and apply Monte Carlo testing again to determine whether clustering for larger values of q sources from clustering occurring for smaller values of q. If the p-value yielded by this second part of the q-nearest neighbors test does not indicate significance, we say the clustering for the larger value of q is caused by the clustering occurring at the smaller value of q. For each case group, we considered $q = 3, 5, 7, 9, 11, 13, 15$.

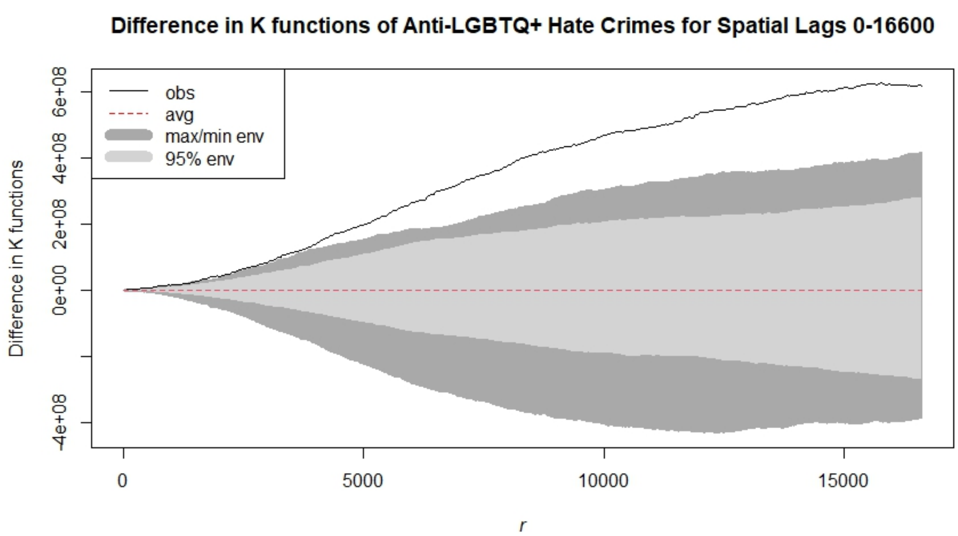
3 Results

3.1 Anti-LGBTQ+ Hate Crimes

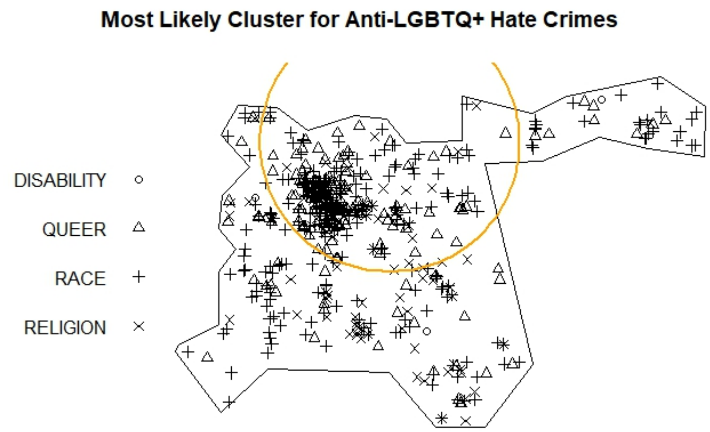
For Anti-LGBTQ+ Hate Crimes, all four methods provided evidence of local clustering and global clustering with respect to their particular test statistics.



For the local log relative risk test statistic, our Monte Carlo Tolerance Envelopes suggested a large cluster of Anti-LGBTQ hate crimes relative to Non-LGBTQ+ related hate crimes in the central north portion of Denver, as well as a large cluster of Non-LGBTQ+ related hate crimes relative to Anti-LGBTQ+ hate crimes along the southern border of the region (as seen in figure 3.1). With a p-value of 0.037 generated from 999 simulations, we rejected the null hypothesis for the global log relative risk test, concluding that the spatial density of Anti-LGBTQ+ hate crimes differs from the spatial density of Non-LGBTQ+ related hate crimes for at least one location in Denver.



For the local difference in K-functions test statistic, Monte Carlo tolerance envelopes indicated clear spatial clustering of Anti-LGBTQ+ hate crimes above any clustering observed in Non-LGBTQ+ related hate crimes for spatial lag distances between roughly 4500 and 16500, with borderline clustering of Anti-LGBTQ+ hate crimes above any clustering in Non-LGBTQ+ related hate crimes for distances between 2000 and 4500 (as seen in figure 3.1). And with a p-value of 0.002 from 499 simulations, we clearly reject the null hypothesis of the global difference in K-functions test statistic, providing evidence of clustering of Anti-LGBTQ+ hate crimes above and beyond Non-LGBTQ+ related hate crimes for at least one spatial scale between 0 and 16631.02 (i.e,.: $KD(h) > 0$ for at least one value of h between 0 and 16631.02).



With a p-value of 0.004, we rejected our null hypothesis for the spatial scan method, concluding there must be at least one cluster of Anti-LGBTQ+ hate crimes in Denver (more unusual than what is expected under the random labeling hypothesis). Furthermore, the spatial scan method only identified one significant most likely cluster of Anti-LGBTQ+ hate crimes in Denver, which covers nearly the same central northern region of Denver as the cluster of Anti-LGBTQ hate crimes relative to Non-LGBTQ+ related hate crimes that was identified by the log relative risk method (as seen in figure 3.1).

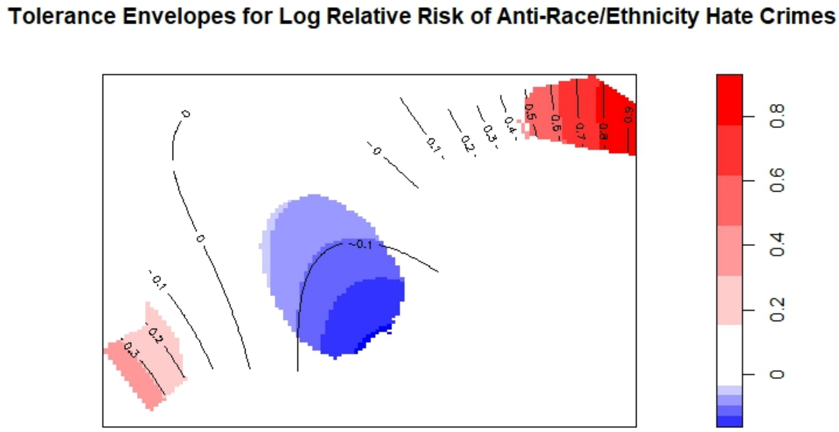
All values of q except $q = 5$ yielded a p-value of significance (between 0.004 and 0.024) for T_q, indicating that the number of Anti-LGBTQ+ hate crimes that occurred within the $3, 7, 9, 11, 13$ or 15 nearest neighbors of other Anit-LGBTQ+ hate crimes is greater than what is expected under the random labeling hypothesis. However, in considering the contrasts between test statistics, with a p-value of 0.048, the clustering occurring when $q = 15$ seems to be the only scale of nearest neighbors clustering that does not arise from clustering of the 3 nearest neighbors.

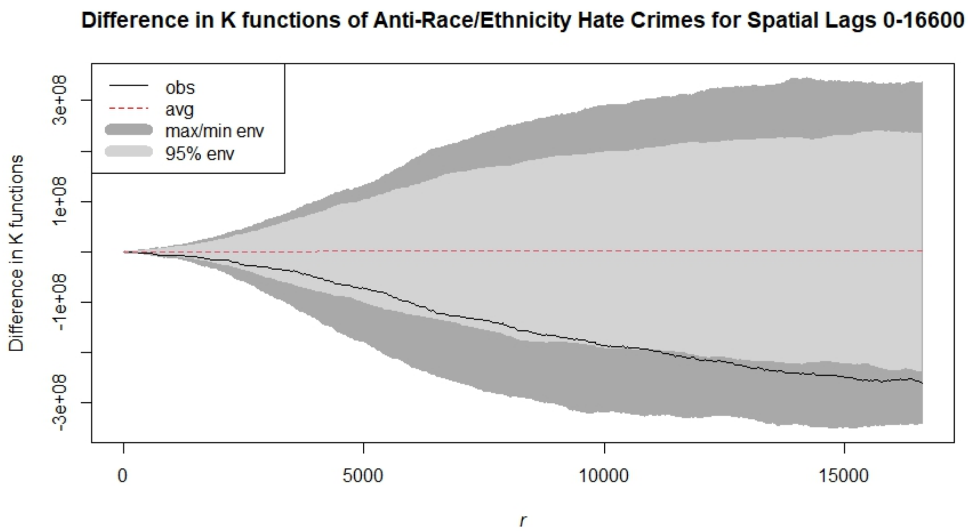
3.2 Anti-Race/Ethnicity Hate Crimes

For Anti-Race/Ethnicity Hate Crimes, our spatial scan and log relative risk tests identified local clusters and global clustering, but our difference in k-functions and q-nearest neighbors tests indicated no apparent significant scales of clustering.

For the local log relative risk test statistic, our Monte Carlo Tolerance Envelopes suggested two clusters of Anti-Race/Ethnicity hate crimes relative to hate crimes unrelated to race or ethnicity, one clear one of moderate size in the northeastern corner of Denver and another less certain and smaller cluster in the southwestern corner of Denver. Additionally, these tolerance envelopes indicated a moderately-sized cluster of hate crimes unrelated to race or ethnicity relative to Anti-Race/Ethnicity hate crimes in the southeastern corner of the region (as seen in figure 3.2). Furthermore, a p-value of 0.015 generated from 999 simulations caused us to reject the null hypothesis for the global log relative risk test, concluding that the spatial density of Anti-Race/Ethnicity hate crimes differs from the spatial density of hate crimes unrelated to race or ethnicity for at least one location in Denver.

For the local difference in K-functions test statistic, the difference in K functions never exceeded or fell below the max/min Monte Carlo Tolerance envelopes we found, indicating that there is no evidence of spatial clustering



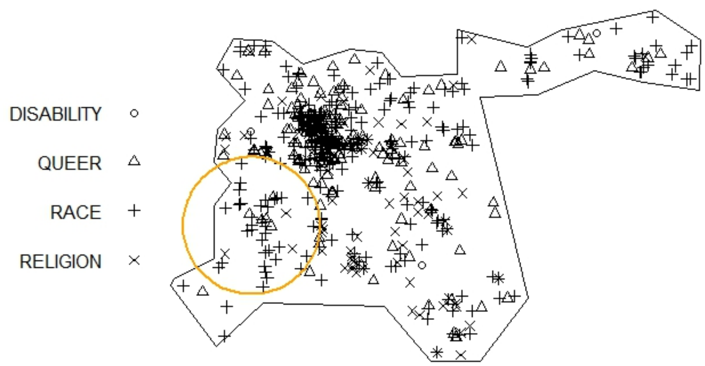


of Anti-Race/Ethnicity hate crimes relative to clustering of hate crimes unrelated to race or ethnicity for any spatial lag distances between 0 and roughly 16500 (as seen in figure 3.2). And with a p-value of 0.97 from 499 simulations, we failed to reject the null hypothesis of the global difference in K-functions test statistic, concluding that there is no evidence of clustering of Anti-Race/Ethnicity hate crimes above and beyond hate crimes unrelated to race or ethnicity for all spatial scales between 0 and 16631.02.

With a p-value of 0.031, we rejected our null hypothesis for the spatial scan method, concluding there must be at least one cluster of Anti-Race/Ethnicity hate crimes in Denver (that is more unusual than what is expected under the random labeling hypothesis). Furthermore, the spatial scan method only identified one significant most likely cluster of Anti-Race/Ethnicity hate crimes in Denver, which covers a southwestern region of Denver of moderate size just offset from the southwestern cluster of Anti-Race/Ethnicity hate crimes relative to hate crimes unrelated to race or ethnicity and the southeastern cluster of hate crimes unrelated to race or ethnicity relative to Anti-Race/Ethnicity hate crimes that were identified by the log relative risk method (as seen in figure 3.2).

Only $q = 9, 11, 13, 15$ yielded p-values of significance (0.032, 0.046, 0.036, and 0.016, respectively) for T_q, indicating that the number of Anti-Race/Ethnicity hate crimes that occurred within the $9, 11, 13$ or 15 nearest neighbors of other Anti-Race/Ethnicity hate crimes is greater than what is expected under the random labeling

Most Likely Cluster for Anti-Race/Ethnicity Hate Crimes (Spatial Scan)

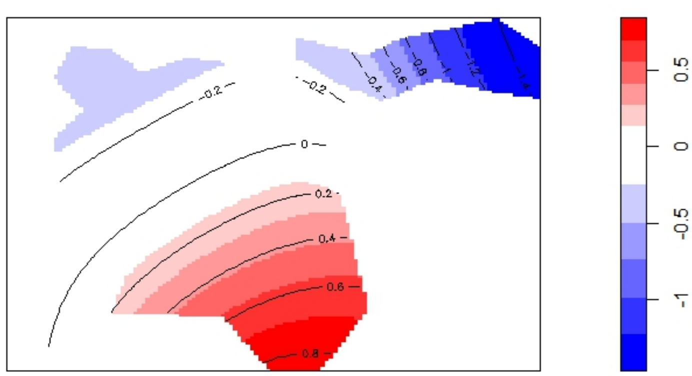


hypothesis. However, in considering the contrasts between test statistics, no contrasts $T_q's$ for these values of q yielded a significant Monte Carlo p-value. Thus, we concluded that the clustering occurring when $q = 11, 13, 15$ seems to be caused or sourced from the clustering occurring when $q = 9$.

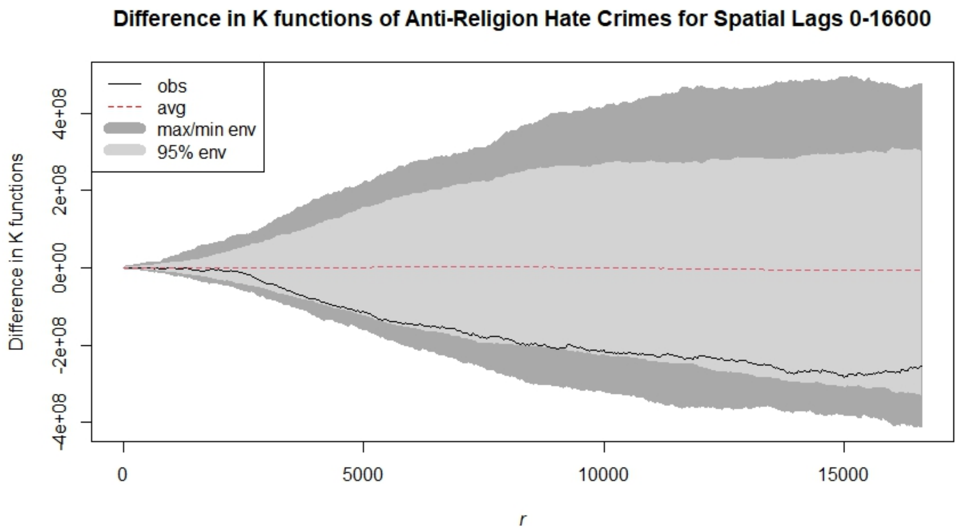
3.3 Anti-Religion Hate Crimes

Choosing Anti-Religion hate crimes as our case group, the log relative risk, spatial scan, and q nearest neighbors methods each provided evidence of statistically significant local clusters and global clustering, while the difference in K functions method did not provide any evidence of local or global clustering.

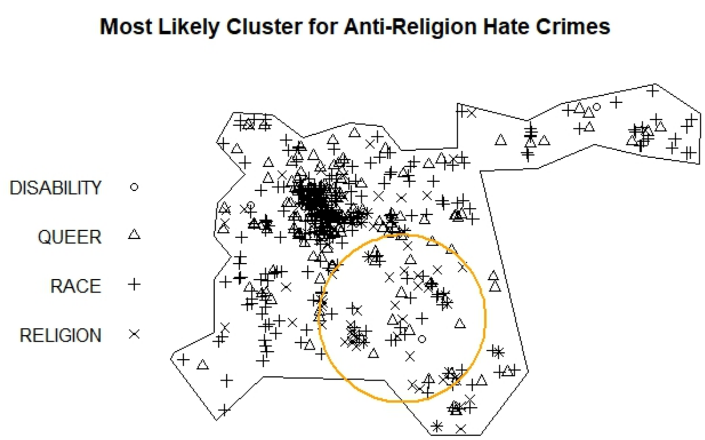
Tolerance Envelopes for Log Relative Risk of Anti-Religion Hate Crimes



For the local log relative risk test statistic, our Monte Carlo Tolerance Envelopes suggested a large cluster of Anti-Religion hate crimes relative to hate crimes unrelated to religion in the southeastern corner of Denver. Additionally, these tolerance envelopes indicated two moderately-sized clusters of hate crimes unrelated to religion relative to Anti-Religion hate crimes, one clear cluster in the northeastern corner of Denver and another less definitive cluster in the northwestern corner of Denver (as seen in figure 3.3). Furthermore, a p-value of 0.005 generated from 999 simulations caused us to reject the null hypothesis for the global log relative risk test, concluding that the spatial density of Anti-Religion hate crimes differs from the spatial density of hate crimes unrelated to religion for at least one location in Denver.



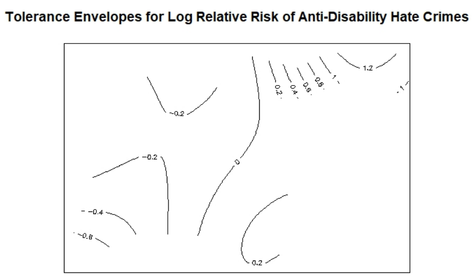
For the local difference in K-functions test statistic, the difference in K functions never exceeded or fell below the 95% Monte Carlo Tolerance envelopes we found, indicating that there is no evidence of spatial clustering of Anti-Religion hate crimes relative to clustering of hate crimes unrelated to religion for any spatial lag distances between 0 and roughly 16500 (as seen in figure 3.3). And with a p-value of 0.958 from 499 simulations, we failed to reject the null hypothesis of the global difference in K-functions test statistic, concluding that there is no evidence of clustering of Anti-Religion hate crimes above and beyond hate crimes unrelated to religion for all spatial scales between 0 and 16631.02.

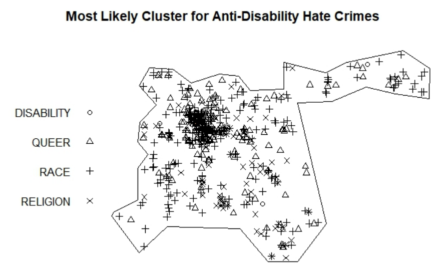


With a p-value of 0.001, we rejected our null hypothesis for the spatial scan method, concluding there must be at least one cluster of Anti-Religion hate crimes in Denver (that is more unusual than what is expected under the random labeling hypothesis). Furthermore, the spatial scan method identified one significant most likely cluster of Anti-Religion hate crimes in Denver, which covers a central southeastern region of Denver of moderate size that covers much of the same region as the southeastern cluster of Anti-Religion hate crimes relative to hate crimes unrelated to religion that was identified by the log relative risk method (as seen in figure 3.3).

All values of q yielded p-values of significance (between 0.002 and 0.008) for T_q, indicating that the number of Anti-Religion hate crimes that occurred within the 3, 5, 7, 9, 11, 13 or 15 nearest neighbors of other Anti-Religion hate crimes is greater than what is expected under the random labeling hypothesis. However, in considering the contrasts between test statistics, all contrasts except $T_{13} - T_{11}$ and $T_9 - T_7$ yielded a significant Monte Carlo p-value (0.082 and 0.244, respectively). Thus, we concluded that the clustering occurring when $q = 13$ seems to be caused or sourced from the clustering occurring when $q = 11$, and the clustering occurring when $q = 9$ seems to arise from the clustering occurring when $q = 7$, but the clustering occurring when $q = 3, 5, 7, 11$, or 15 are all independently significant.

3.4 Anti-Disability Hate Crimes





Lastly though, Anti-Disability Hate Crimes revealed no apparent clustering through any test. Plots for the results of the local log relative risk and spatial scan methods are included in 3.4 for partial verification of the insignificance. All Monte Carlo p-values were in the range of 0.9 to 1. However, because the number of points considered cases for this particular case group was so low, we considered these results inconclusive and have omitted full details of the test results for the Anti-Disability case group.

4 Conclusion

In conclusion, there is strong evidence for local clusters and global clustering of Anti-LGBTQ+ (as provided by all four methods) and Anti-Religion hate crimes (as provided by the log relative risk, spatial scan, and q nearest neighbors methods) and moderate evidence for clustering of Anti-Race/Ethnicity hate crimes (as provided by the log relative risk, spatial scan, and q nearest neighbors methods). The evidence for local clusters of Anti-Race/Ethnicity hate crimes is less certain due to conflicting results from the log relative risk and spatial scan methods as to where Anti-Race/Ethnicity hate crimes are clustered in Denver.

In considering the results of this spatial analysis and what policy recommendations to make from it, one of the most significant and justified hindrances to the relevance and significance of this analysis is the deidentification of location information within the dataset. Deidentifying location information related to hate crimes in public datasets is certainly an important choice to protect the victims of hate crimes. While we made effort to contact the people who maintain the dataset in the hopes of ethically obtaining location information that referenced back to Denver in a clear way, we were unsuccessful in finding a way to reference the coordinates in the dataset back to Denver. So while one could compare the study area to a map of Denver to get a rough estimate of where the coordinate locations are, there is no way to know exactly which neighborhoods in Denver are contained within the local clusters for different case groups. If the chance to reference these coordinates back to the city of Denver arises in the future, this analysis could provide valuable insight for people who identify as part of one of the target groups about which areas of the city are safer for them to live in.

Another limitation to this analysis is the generalization of the case groups. To policymakers or newcomers to Denver, it is likely more valuable to distinguish between clusters of hate crimes against people who identify as Black or African American and clusters of hate crimes against people who identify as White, or any other race/ethnicity. Unfortunately, to perform this kind of analysis would likely require a larger dataset, perhaps from the state or national level, so that we could have large enough counts of hate crimes targeted at each specific identity to perform analysis on smaller and more specific case groups.

This analysis could also be used in conjunction with analysis on other data sets or further analysis on this dataset. The same methods used here could be simply applied to case status or crime type information in this dataset. By comparing the local clusters of case statuses, crime types, and target groups, one could obtain a general understanding of the type of justice and equity that our justice system provides to people with different identities who have been affected by similar types of crimes, or which areas of the city of Denver receive the least justice. We could also search for a separate point pattern or geostatistical data set of the levels of policing in various parts of the city. By applying the same methods used in this analysis to a data set about policing in Denver and observing the levels of policing in the local clusters of target groups of hate crimes, we might find insight into how effectively police protect the most vulnerable members of our society. Whichever of the continuations explained above is pursued, expanding on this analysis can provide valuable support to efforts of Civil Rights, Women's Rights, Queer Liberation, and Black Lives Matter movements (among many others) to create a more just and equitable world where diversity is celebrated and reduce the incidents of violence and crime that arise from intolerance of diversity.