Time series analysis on racial hate crimes in the United States

October 21, 2022

Mandy Hong

Introduction

A crime where offenders' bias against race/ethnicity/ancestry acts as a primary motivation is defined as a *racial hate crime* (United States Department of Justice, n.d.). In 2020, hate crime in the US hit a historic high in 12 years (Carrega & Krishnakumar, 2021). Among 7,554 incidents, 61.9% were racial hate crimes, and among those racial hate crimes, two racial groups, Black or African American and Asian, had the highest increase (146% and 173% each) (Federal Bureau of Investigation, 2021). While most of the recent hate crimes involved a racial bias, past studies included all categories of hate crimes rather than focusing on racial hate crimes (Wang, 2021). Thus, this project aims to find a model that best predicts racial hate crimes. Moreover, we will subdivide racial hate crimes into two categories; anti-African American or Black and anti-Asian hate crimes. Thus, our research focuses on exploring the trend, finding the best predictive model by fitting the function of time model and the SARIMA model, and making the prediction of all racial, anti-African American or Black, and anti-Asian hate crimes.

Methods

1. Data

Data Acquisition and Biases

Our main datasets come from the Federal Bureau of Investigation Crime Data Explorer. Data is publicly available without needing permission to access in CSV format. FBI's hate crime data are generated based on the crime voluntarily reported to the FBI by individual law enforcement agencies. However, scholars argue that many hate crimes are often underreported compared to non-biased crimes since the motivation of the hate crime is based on the offender's bias (Lantz & Wenger, 2021). Since motivation is subjective, it is difficult to tell whether a crime is a hate crime (FBI, n.d.). Also, some police agencies often fail to accurately identify and report hate crimes due to the variability of the number of police trained to identify hate crimes on sight and in existing crime reporting programs (Farrell & Lockwood, 2022). Even though we have potential concerns, the FBI data was built based on data from multiple independent agencies.

Thus, if we consider our population to be hate crime cases that rise to the level of attention of the FBI, our data can be considered a random sample. As the awareness of racial hate crimes increases and police agencies develop more sophisticated methods to capture racial hate crimes, it will be possible to have more complete data that represents a larger population of hate crime victims so that we can make a better predictive model.

Data Preparation

The hate crime data from the FBI includes all categories of hate crimes. Thus, we first create a data frame that only has racial hate crimes. There are several incidents involving multiple biases, so we exclude those cases because we cannot tell which bias most influenced the crime. Next, we create three monthly time series data by counting the number of all racial, anti-African American, and anti-Asian hate crimes. Since our original data does not have a column that only shows the month the incident happened, we generate a new column to store monthly data. We extract the month from the incident date column. Then we aggregate the number of hate crimes by year and month. Thus, we have three time series data from 1991 to 2020 (monthly, 360 rows) with three columns; year, month, and the number of hate crimes in CSV format.

2. Analysis

Assumption and Validation

To do statistical analysis, we assume that each hate crime reported to the FBI is an independent case and motivated by a single bias, race. We implement the function of time model and SARIMA model to fit our time series data. To find a model that best predicts racial hate crimes, we first check the condition of each model (stationarity and independence). Then, we compare Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), and R-squared value. The model is better if it has a lower AIC, lower BIC, and higher R-squared value. If we have a model with stationary residuals, we add additional terms to see if the model can get better. We divide each data into two subsets and check the R-squared value of different models to validate our models. As we identify the best predictive model, we predict racial hate crimes for two years, 2021 and 2022.

The function of time model

If we identify a nonstationarity due to the trend in our time series, we can detrend it by fitting a different function of time model. We use a regression to discover linear, quadratic, cyclic and periodic trends. As we discover the trend, we can detrend our time series by fitting the appropriate function of time model and get residuals. If our residuals are linear with constant variance and no correlation, we achieve a stationarity.

SARIMA model

Often real-world time series are complex to explain solely by regression models (Shumway & Stoffer, 2019). To explain the dynamics of a time series that has seasonal behavior, we implement seasonal autoregressive integrated moving average (SARIMA) models . If the time series is nonstationary, we can first apply differencing to make it stationary. If there is a correlation between past time steps, we need autoregressive (AR) and moving average (MA) models. SARIMA$(p, d, q) \times (P, D, Q)_s$ model uses seven parameters; autoregressive operator (p,), order of differencing (d), moving average operator (q), seasonal autoregressive operator (P), order of seasonal differencing (D), seasonal moving average operator (Q), and seasonal period (s). If our residuals are stationary with no correlation after fitting the SARIMA model, we can do prediction using the time series data. If residuals are normally distributed, we can also trust forecast intervals.

We use *Python 3* to collect and clean data (Van Rossum & Drake, 2009). For descriptive and predictive analysis, we use *R v4.1.3* (R Core Team, 2022).

Results and Conclusions

1. Descriptive Analysis

From 1991 to 2020, the mean of all racial, anti-African American or Black, and anti-Asian hate crime cases is 389.1, 207.7, and 17.84. The time series plot in Figure 1. and 2. shows that all racial and anti-African American or Black hate crimes seem to follow similar trends. Also, we can observe seasonality within a 12 months period from both time series. Anti-Asian hate crimes do not show an apparent seasonal trend compared to the first two groups. All three time series have several spikes, and we use the interquartile range[1] to identify outliers (Figure 3). Racial hate crimes increased significantly during the last quarter of 2001 and after half of 2020. We identified a similar spike in the anti-African American of Black hate crimes. The first spikes could be explained by an increase in anti-Arab hate crimes after the 9/11 attacks in the US in 2001 (Hanes & Machin, 2014). The surge in the second half of 2020 might be due to the impacts of the Black Lives Matter movement after the death of George Floyd.

[1] Interquartile range (IQR) is a difference between the third and the first quartile (Q3-Q1). If values are outside the lower (Q1-(IQR*1.5)) and the upper (Q3-(IQR*1.5)) boundary, we consider them outliers (Vinutha et al., 2018).

However, to confirm this, further investigation is needed. The number of anti-Asian hate crimes was the highest in March 2020. After the COVID-19 outbreak in January 2020, the US declared a public health emergency in early February and states issued a lockdown order at the end of March (The American Journal of Managed Care, 2021). The rise of anti-Asian hate crimes might be due to anti-Asian sentiments after COVID-19, but further research is needed to confirm this relationship.

2. The Function of Time Models

For all three time series, we fitted five different function of time models, and the best model was the model with both linear and seasonal trends (12 months). The model has time (t=1:360) and 12 months as independent variables. All three model shows that all independent variables are significantly associated with our dependent variable ($p < 0.05$). Compared to all other models, these three models have the highest adjusted R-squared[2]. However, the residuals of these models are not linear and nonstationary (Figure 4). Therefore, it is not a good model for the prediction. Since all detrended time series are not stationary, we should consider differencing the data.

3. SARIMA Models

All racial hate crimes

The ACF shows seasonal patterns with a 12 months period that do not converge to zero quickly. This indicates the need for seasonal differencing to make time series stationary. The seasonal differenced time series passes four unit root tests (ADF, PP, KPSS, ERS), and the time series plot looks stationary. Thus, we proceed by fitting SARIMA models using information from the ACF and PACF. Among seven different models, the best model was $\mathrm{SARIMA}(2, 0, 1) \times (0, 1, 1)_{12}$ without the drift term (equation in Table 1). This means our model needs seasonal differencing, first-order autoregressive operator, second-order autoregressive operator, first-order moving average operator, and seasonal first-order moving average operator. All terms were significant to include ($p < 0.05$). The model suggests that the seasonal differenced racial hate crime of the current month t ($\Delta_{12}HC_t$=HC_t-HC_{t-12}) can be explained by the racial hate crime of the past month (HC_{t-1}), the racial hate crime of two months ago (HC_{t-2}), the error of the past month (ε_{t-1}), the error of the same month last year (ε_{t-12}), and the error in the current month (ε_t). The time series plot of the residuals of our final model looks stationary, passes four unit root tests, and has no significant autocorrelation identified from ACF and PACF (Figure 5). Also, this model has the best

2 Three best function of time models have adjusted R-squared values of 0.9417, 0.9388, 0.8687 meaning 94.17%, 93.88%, and 86.87% variability of each category of hate crime could be explained by our independent variables.

AIC, BIC, and R-squared values (4017.1, 4036.36, 0.579). Thus, we can use this model for prediction. However, it failed the normality test, which could be due to several outliers. Therefore, we can not trust forecast intervals.

Anti-African American or Black hate crimes

The ACF shows seasonal patterns with a 12 months period that slowly decay, which indicates the need for seasonal differencing. The seasonal differenced time series passes four unit root tests (ADF, PP, KPSS, ERS), and the time series plot looks stationary. We proceed by fitting SARIMA models using ACF and PACF. Among six different models, the best model was $\mathrm{SARIMA}(0, 1, 2) \times (0, 1, 1)_{12}$ without the drift term (equation in Table 1). This means our model needs first-order differencing, seasonal differencing, first-order moving average operator, second-order moving average operator, and seasonal first-order moving average operator. All terms were significant to include ($p < 0.05$). The model suggests that the first order and seasonal differenced anti-African American or Black hate crime of current month *t* ($\Delta_{12}\Delta B_t = (B_t - B_{t-1}) - (B_{t-12} - B_{t-13})$) can be explained by the error of the past month (ε_{t-1}), the error of two months ago (ε_{t-2}), the error of the same month last year (ε_{t-12}), and the error of the current month (ε_t). The time series plot of the residuals of our final model looks stationary, passes three unit root tests (ADF, PP, ERS), and has no significant autocorrelation identified from ACF and PACF (Figure 6). Also, this model has the best AIC, BIC, and R-squared values (3478.41, 3478.53, 0.688). Thus, we can use this model for prediction. However, it failed the normality test, which could be due to several outliers. Therefore, we can not trust forecast intervals.

Anti-Asian hate crimes

The ACF shows slow decay in lag which indicates the need for the first-order differencing to make time series stationary. The first-order differenced time series passes four unit root tests (ADF, PP, KPSS, ERS), and the time series plot looks stationary. Thus, we proceed by fitting SARIMA models using ACF and PACF. Among six different models, the best model was $\mathrm{SARIMA}(1, 1, 1) \times (1, 0, 1)_{12}$ without the drift term (Equation in Table 1). This means our model needs first-order differencing, first-order autoregressive operator, first-order moving average operator, seasonal first-order autoregressive operator, and seasonal first-order moving average operator. All terms were significant to include ($p < 0.05$). The model suggests that the first order differenced anti-Asian hate crime of the current month *t* ($\Delta A_t = A_t - A_{t-1}$) can be explained by the hate crime of the past month (A_{t-1}), the error of the past month (ε_{t-1}), the hate crime of the same month last year (A_{t-12}), the error of the same month last year (ε_{t-12}), and the error of the current month (ε_t). The time series plot of the residuals of our final model looks stationary, passes four unit root tests, and has no significant autocorrelation identified from ACF and PACF (Figure 7). Also, this model has the best AIC, BIC, and R-squared values (2285.7, 2305.12, 0.528). Thus, we can use this model for prediction. However, it failed the normality test, which could be due to several outliers. Thus, we can not trust forecast intervals.

Comparing function of time models and SARIMA models

We first compare the AIC and BIC of models. For all three time series, SARIMA models had better AIC and BIC (Table 2). Also, we calculate R-squared to compare models. To compare two models, we divide each time series data into two subsets; data from 1991-2018 and 2019. We fit both function of time models and SARIMA models using the first subset and make predictions on 2019 using the model. Then, we compared predicted values and the actual data for 2019. As a result, for all racial hate crimes and anti-African American or Black hate crimes, SARIMA models had better predictive performance when calculating the R-squared value (0.322, 0.310) compared to the function of time models (0.261, 0.209). However, for anti-Asian hate crimes, the function of time models had better performance (0.378) than the SARIMA model (0.009). However, by checking outliers of the original data, we know that 2020 had a historic high in anti-Asian hate crimes. Thus, we check how models perform if we create different subsets. Excluding 2020, we create two subsets; data from 1991-2017 and 2018-2019. Then we followed the same procedure from above. With these subsets, the predictive performance of the SARIMA model (0.165) was better than the function of the time model (0.054). However, all function of time models' residuals violate the stationarity condition. Thus, SARIMA models are better for predicting all racial, anti-African American or Black, and anti-Asian hate crimes (Figure 8).

Conclusion

This project was very meaningful in several ways to me. First, I could recall all concepts we had learned so far and reorganize them in my mind. As a result, I get a clearer picture of creating a statistical model with time series data. Second, I could practice planning all the steps needed to find predictive models using time series data by myself. Unlike labs with instructions on creating different models, I had to build a model from scratch. This helped me to practice the skills needed to make the best predictive models. Also, the data I used had several outliers and some concerns with passing the by-eye stationarity test. I get to think a lot about handling and interpreting those outliers in real-world settings. Finally, while writing a paper about this project, I read several time series analysis papers and learned how other fields use time series analysis and how researchers deliver their findings to audiences. In general, I figured out what concepts I did not understand clearly, which could help me go back to our class materials and catch up with those concepts.

Appendix

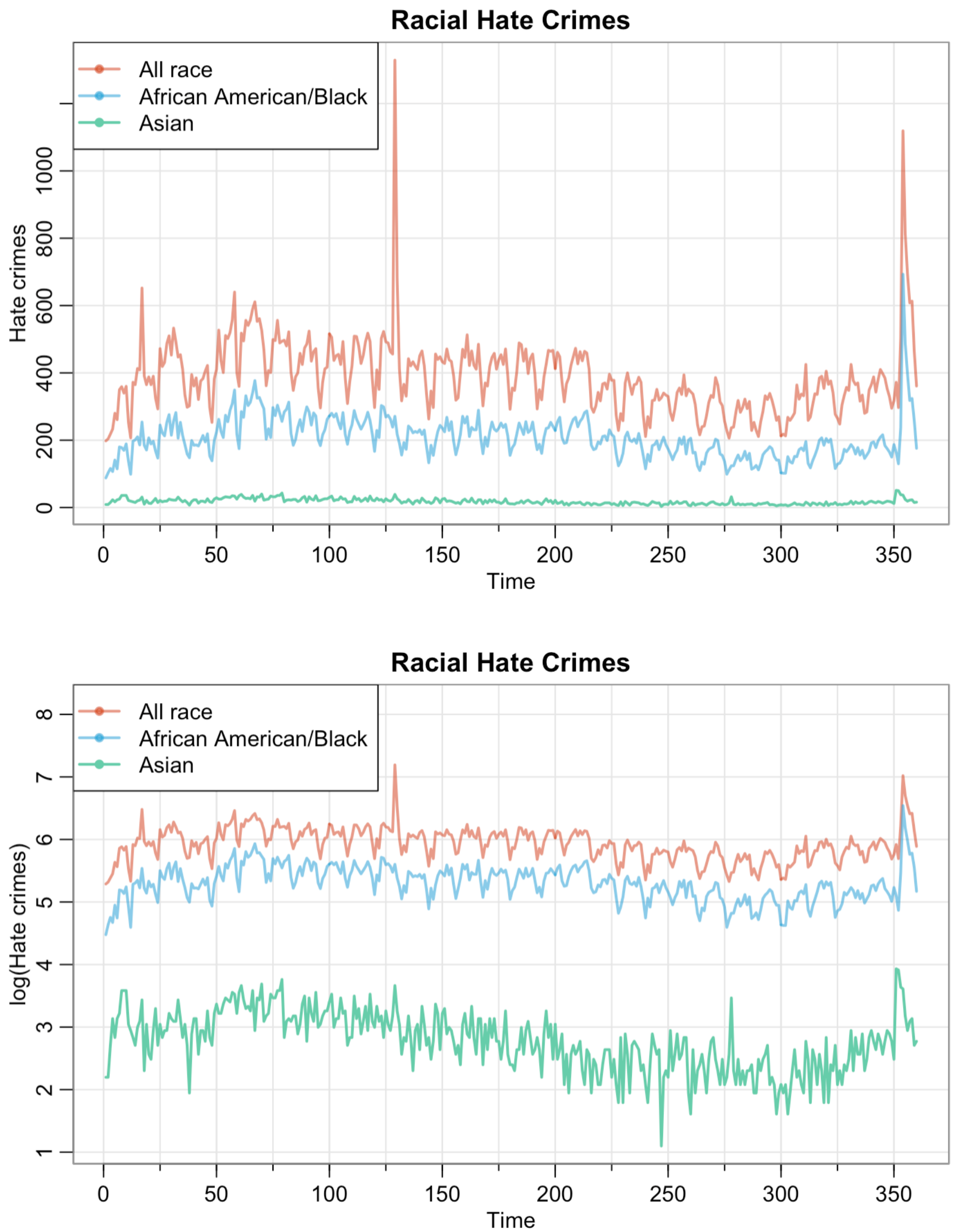

Figure 1. Time series plots of all racial, anti-African American, and anti-Asian hate crimes.

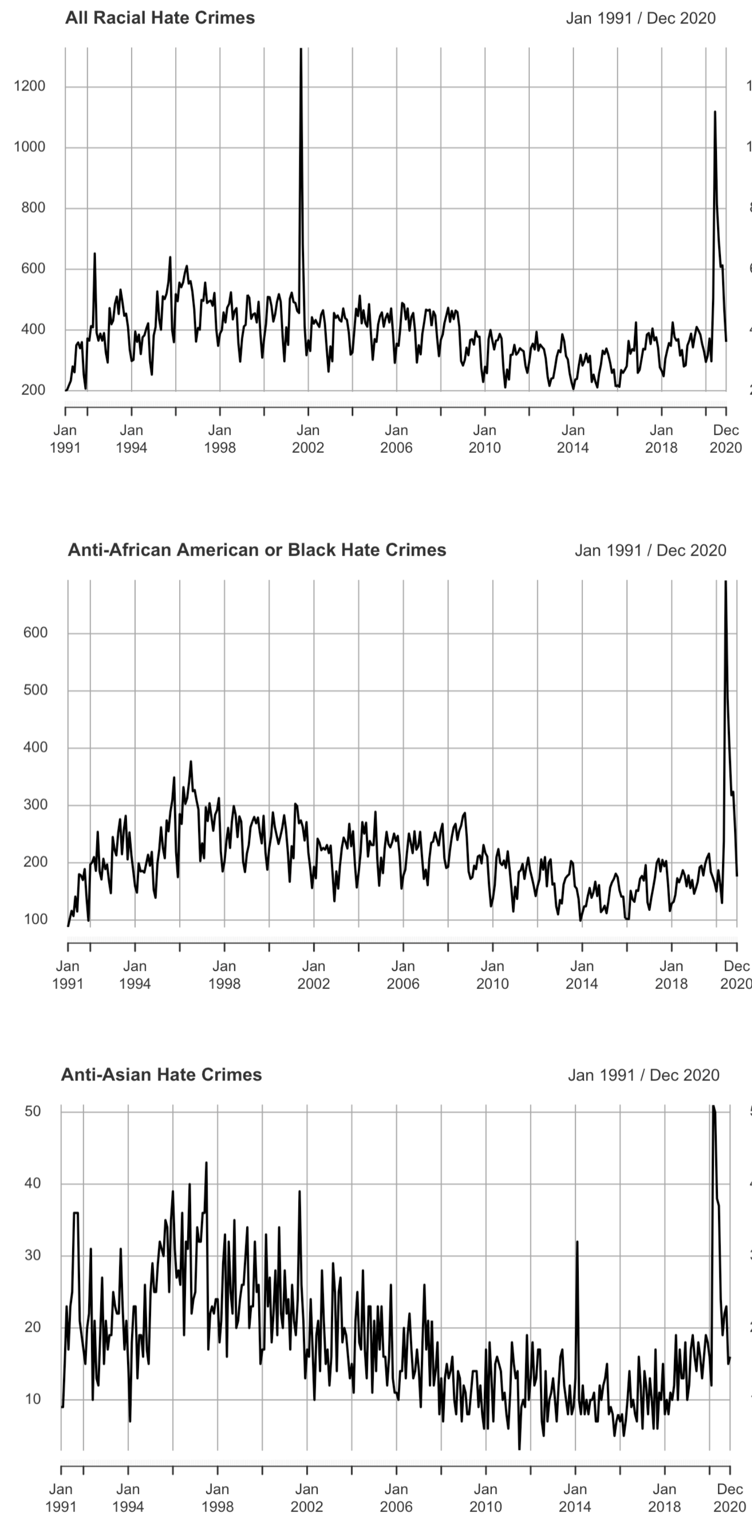

Figure 2. Time series plots of all racial, anti-African American, and anti-Asian hate crimes with time on the *x*-axis.

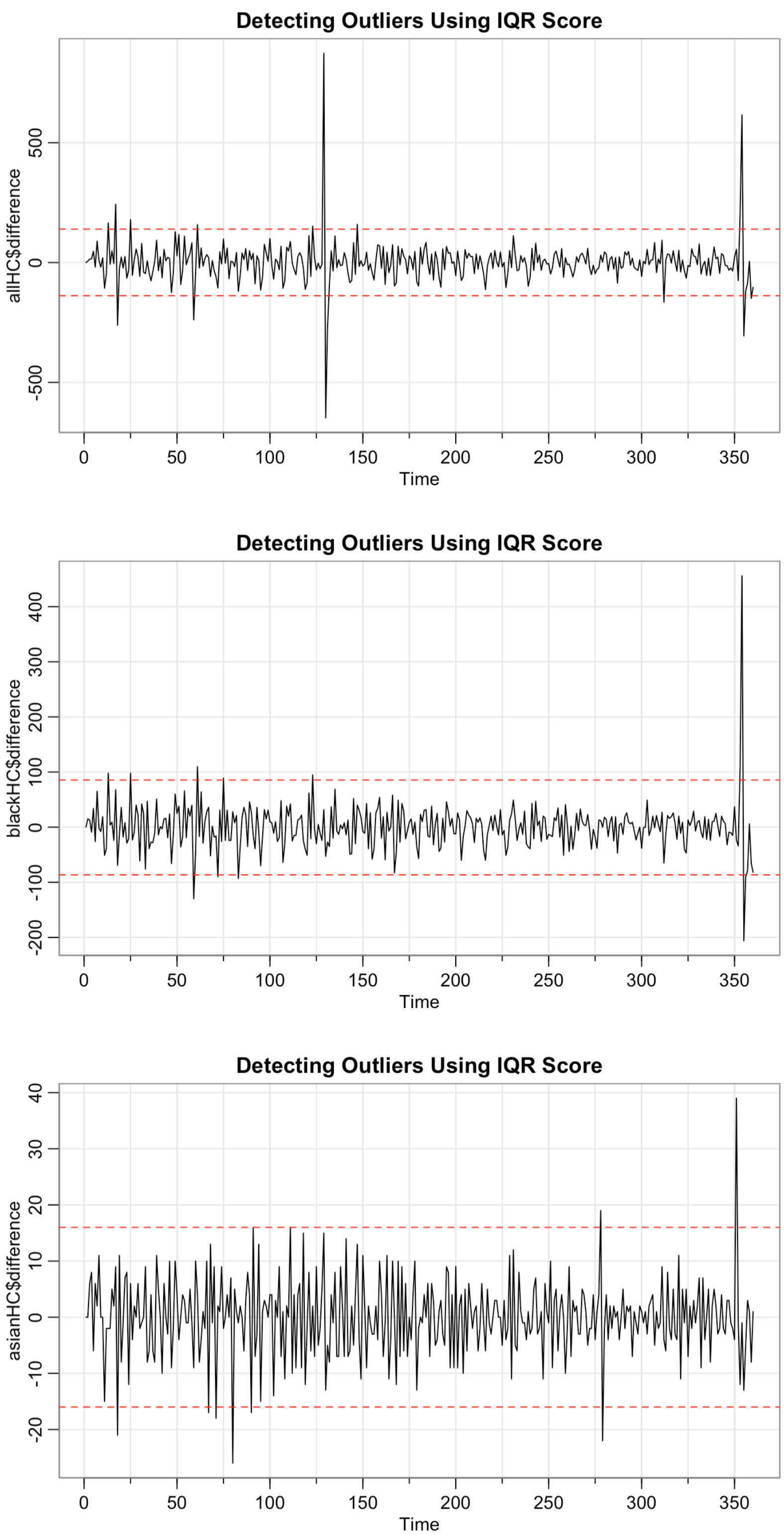

Figure 3. Detecting outliers using IQR scores.

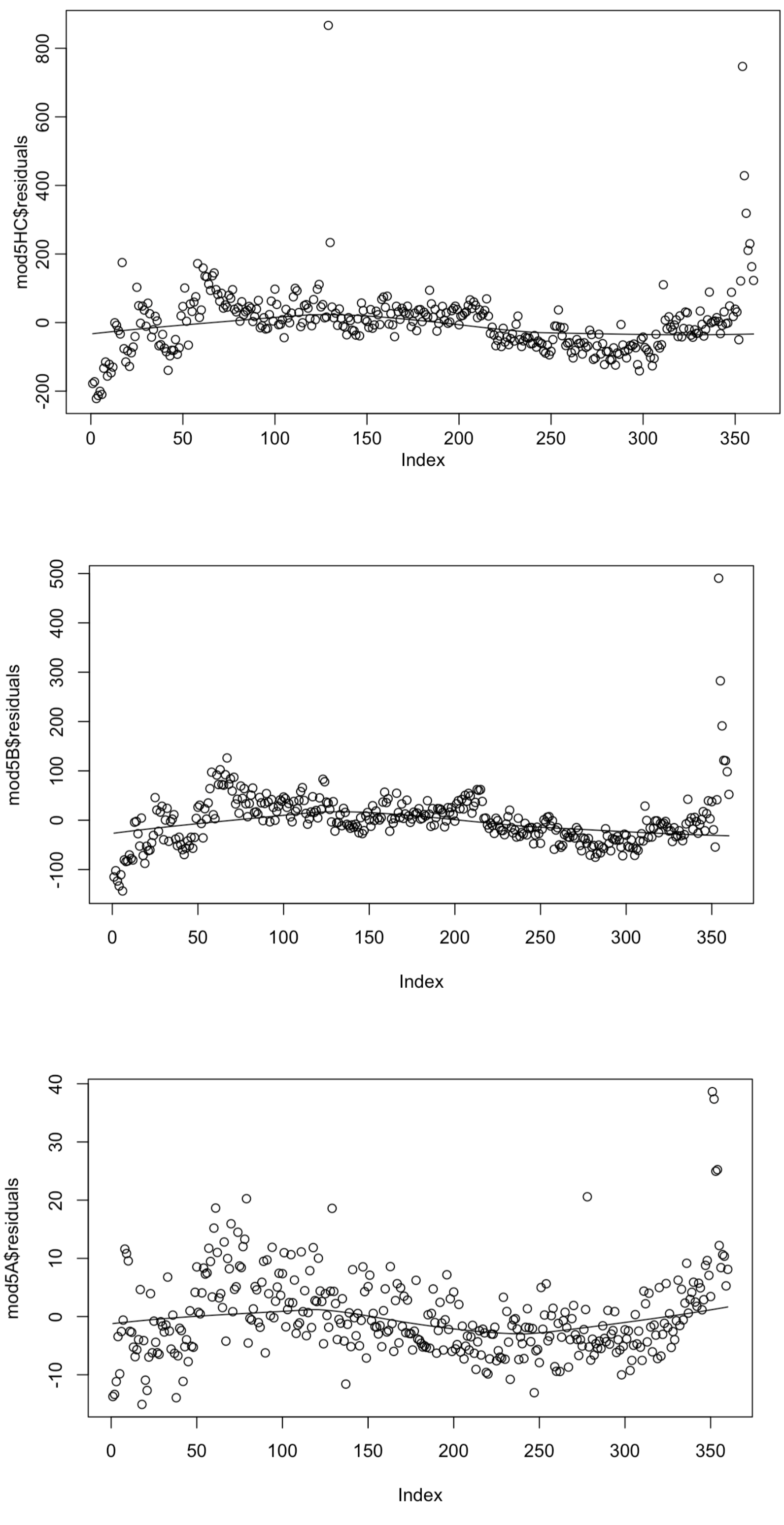

Figure 4. Nonlinear residuals of three best function of time models.

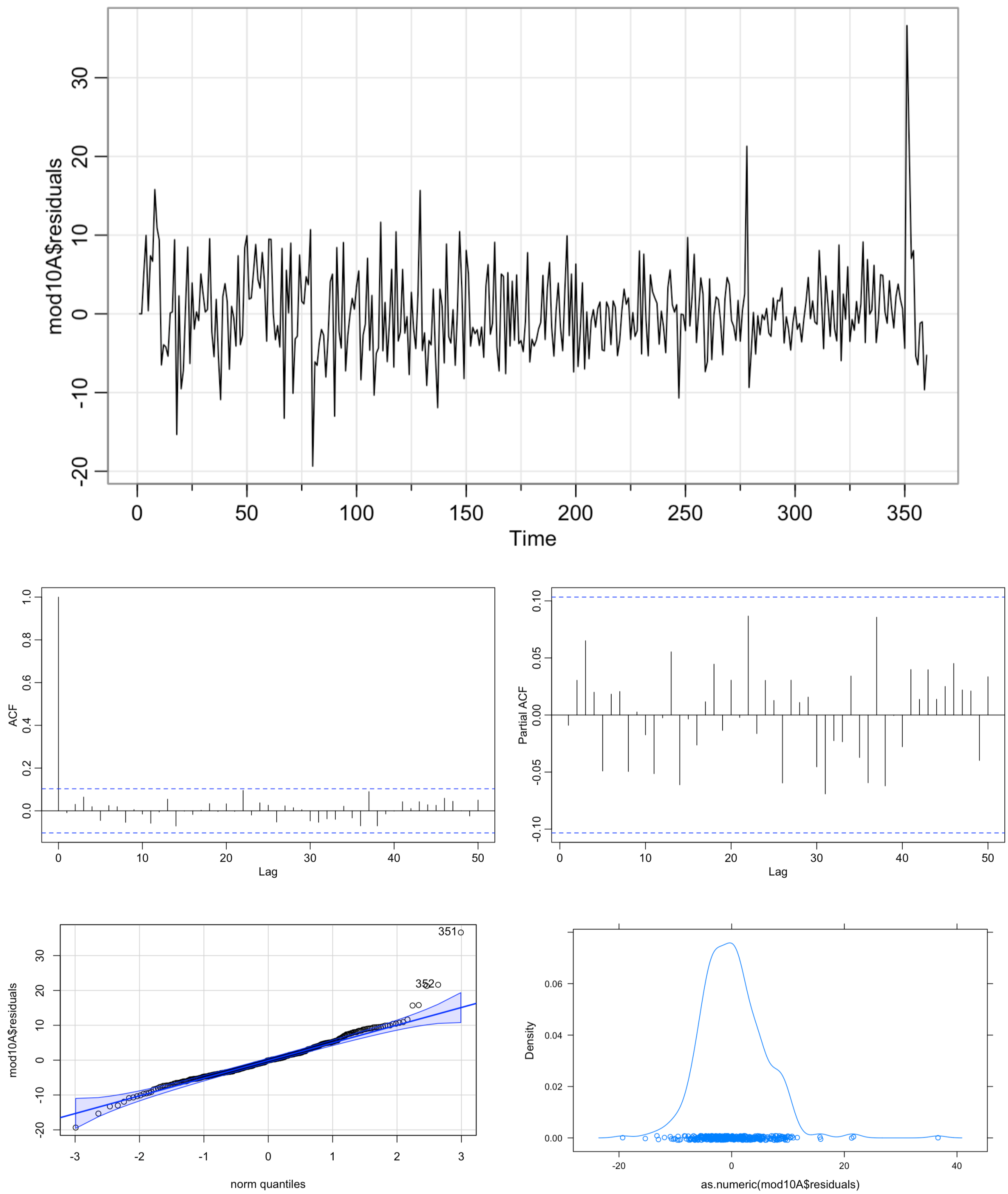

Figure 5. SARIMA$(2,0,1) \times (0,1,1)_{12}$ fitted on all racial hate crimes time series. Time series plots look stationary with some outliers. The ACF and PACF do not show significant lags. The residuals look normal, but the concern remains on outliers.

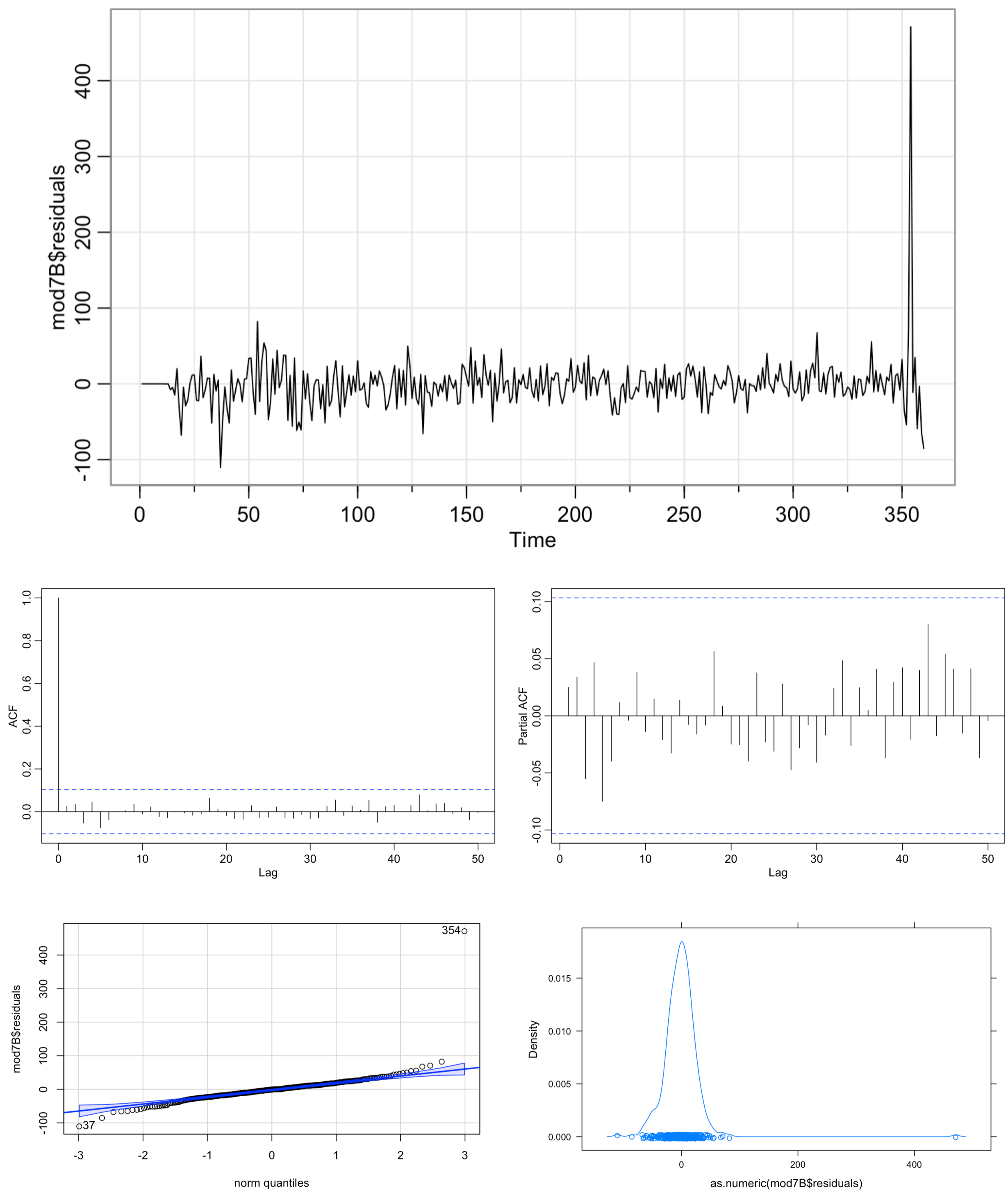

Figure 6. SARIMA$(0, 1, 2) \times (0, 1, 1)_{12}$ fitted on anti-African American or Black hate crimes time series. Time series plots look stationary, with outliers at the end of the data. The ACF and PACF do not show significant lags. The residuals look normal, but the concern remains on outliers.

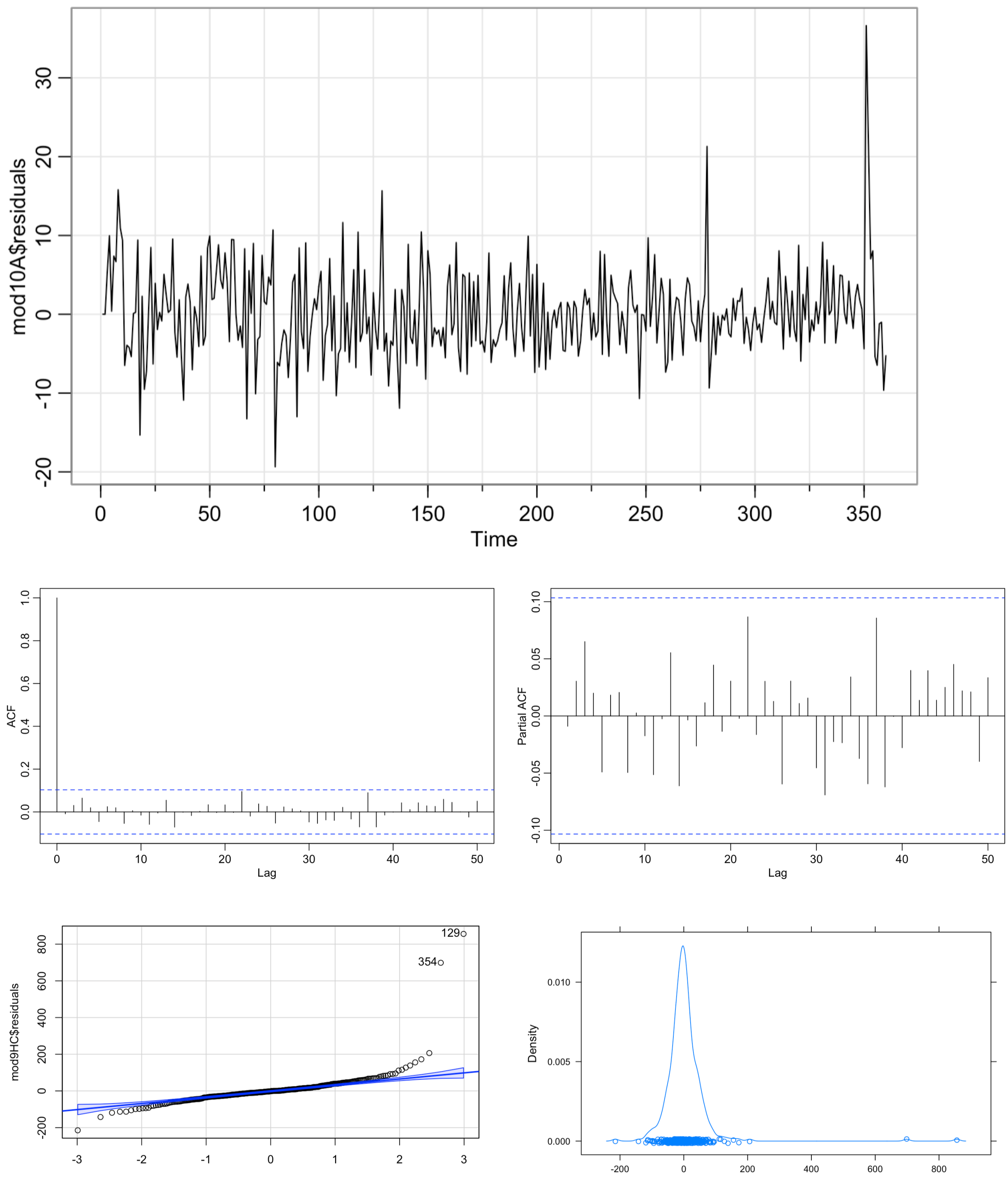

Figure 7. SARIMA$(1,1,1) \times (1,0,1)_{12}$ fitted on anti-Asian hate crimes time series. Time series plots look stationary with some outliers. The ACF and PACF do not show significant lags. The residuals look normal, but the concern remains on outliers.

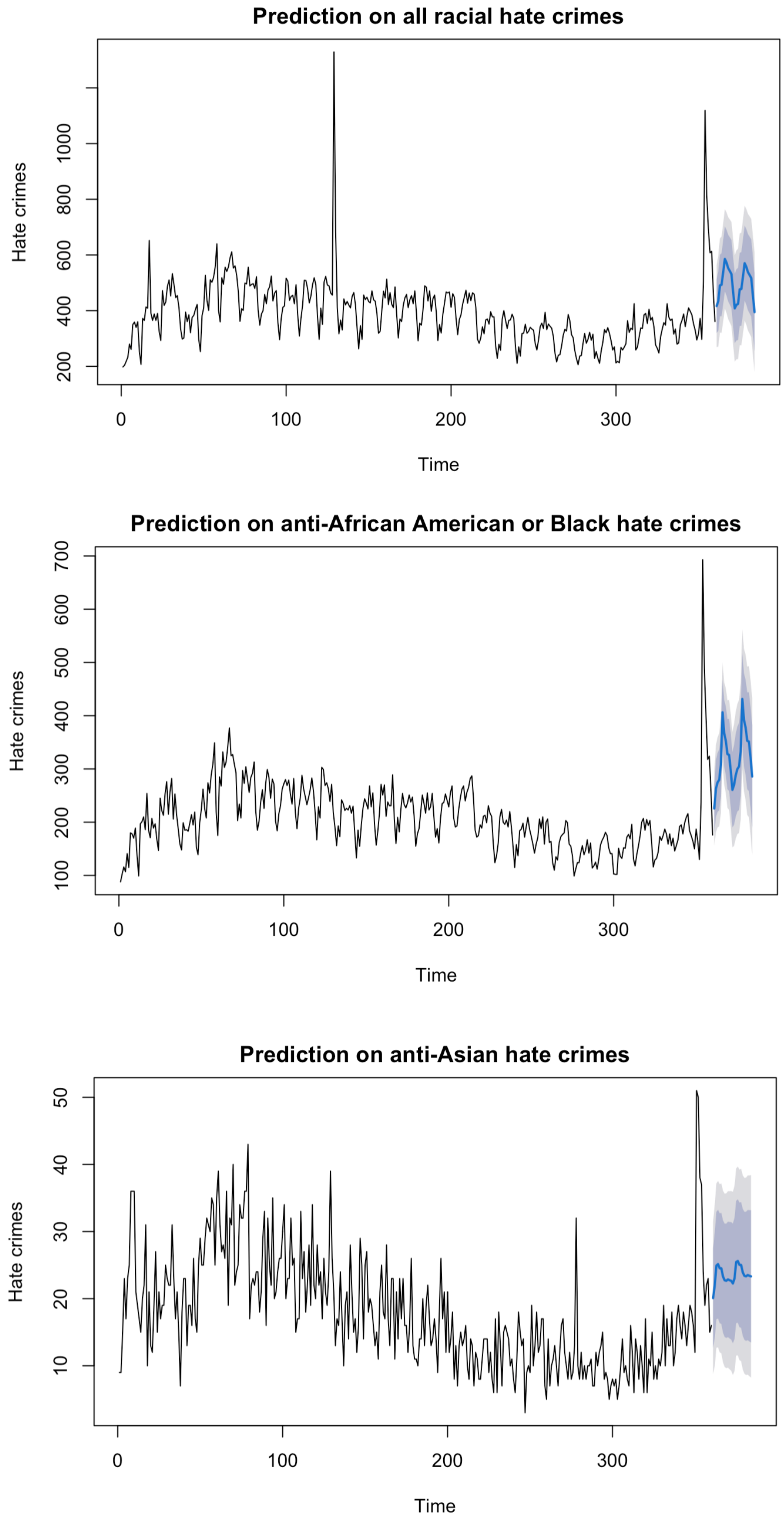

Figure 8. Prediction using SARIMA models. However, we cannot trust the forecast interval since the residuals are not normal for all three models.

Time Series	SARIMA Equation
All racial hate crimes $(2, 0, 1) \times (0, 1, 1)_{12}$	$\Delta_{12}HC_t =$ $1.31HC_{t-1} - 0.3239HC_{t-2} + \varepsilon_t - 0.8332\varepsilon_{t-1} - 0.8858\varepsilon_{t-12}$
Anti-African American or Black $(0, 1, 2) \times (0, 1, 1)_{12}$	$\Delta_{12}\Delta B_t =$ $\varepsilon_t - 0.3883\varepsilon_{t-1} - 0.2895\varepsilon_{t-2} - 0.7777\varepsilon_{t-12}$
Anti-Asian $(1, 1, 1) \times (1, 0, 1)_{12}$	$\Delta A_t =$ $0.2378A_{t-1} + 0.9106A_{t-12} + \varepsilon_t - 0.8797\varepsilon_{t-1} - 0.8371\varepsilon_{t-12}$

Table 1. Equations for SARIMA models.

Time Series	Model	**AIC**	**BIC**
All racial hate crimes	Function of time	4336.28	4390.685
	SARIMA	4017.1	4036.36
Anti-African American or Black	Function of time	3902.605	3957.01
	SARIMA	3478.41	3493.81
Anti-Asian	Function of time	2450.807	2505.212
	SARIMA	2285.7	2305.12

Table 2. Comparing AIC and BIC of the function of time model and the SARIMA model.

References

The American Journal of Managed Care. (2021, April 24). *A Timeline of COVID-19 Developments in 2020*. AJMC. Retrieved October 21, 2022, from https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020

Carrega, C., & Krishnakumar, P. (2021, October 26). *Hate crime reports in US surge to the highest level in 12 years, FBI says*. CNN. https://edition.cnn.com/2021/08/30/us/fbi-report-hate-crimes-rose-2020/index.html

Farrell, A., & Lockwood, S. (2022, September 2). Addressing Hate Crime in the 21st Century: Trends, Threats, and Opportunities for Intervention. *Annual Review of Criminology*, *6*(1). https://doi.org/10.1146/annurev-criminol-030920-091908

Federal Bureau of Investigation. (2021, August 30). *FBI Releases 2020 Hate Crime Statistics*. Retrieved September 16, 2022, from https://www.fbi.gov/news/press-releases/press-releases/press-releases/fbi-releases-2020-hate-crime-statistics

Federal Bureau of Investigation. (n.d.). *Hate Crime* [Dataset]. Federal Bureau of Investigation Crime Data Explorer. https://crime-data-explorer.fr.cloud.gov

Hanes, E., & Machin, S. (2014, June 6). Hate Crime in the Wake of Terror Attacks. *Journal of Contemporary Criminal Justice*, *30*(3), 247–267. https://doi.org/10.1177/1043986214536665

Hu, W., Tong, S., Mengersen, K., & Connell, D. (2007, September). Weather Variability and the Incidence of Cryptosporidiosis: Comparison of Time Series Poisson Regression and SARIMA Models. *Annals of Epidemiology*, *17*(9), 679–688. https://doi.org/10.1016/j.annepidem.2007.03.020

Lantz, B., & Wenger, M. R. (2021, August 20). Are Asian Victims Less Likely to Report Hate Crime Victimization to the Police? Implications for Research and Policy in the Wake of the COVID-19 Pandemic. *Crime &Amp; Delinquency*, *68*(8), 1292–1319. https://doi.org/10.1177/00111287211041521

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Shumway, R., & Stoffer, D. (2019, May 20). *Time Series (Chapman & Hall/CRC Texts in Statistical Science)* (1st ed.). Routledge.

The United States Department of Justice. (n.d.). *Learn About Hate Crimes*. Retrieved September 22, 2022, from https://www.justice.gov/hatecrimes/learn-about-hate-crimes

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: Create Space.

Vinutha, H. P., Poornima, B., & Sagar, B. M. (2018). Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset. *Advances in Intelligent Systems and Computing*, 511–518. https://doi.org/10.1007/978-981-10-7563-6_53