Philadelphia Crime

Forecasting Methods

Nick Morris

12/18/2016

Table of Contents

Data Description

This dataset comes with 14 columns and 2,140,485 rows to characterize over 10 years of Philadelphia crime, where each row corresponds to a police officer dispatch. The time span of the dataset is from 2006-01-01 to 2016-08-18. There are 3 external variables of interest in this study: Time, Location, and Offense. Time will be evaluated hourly, daily, weekly, and monthly to find seasonality. Location will be evaluated through the columns Dc_Dist, Psa, Police_Districts, Lon, and Lat to find subsets of seasonality. Offense will be evaluated through the columns UCR_General and Text_General_Code to find subsets of seasonality. The hypothesis for each of these 3 variables, is that each can be used to predict crime.

Table 1 below includes a description of 12 columns from the original dataset. There were 2 columns removed due to their redundancy, which is explained in the Data Cleansing section. The following link will lead you to the source of the data.

https://www.kaggle.com/mchirico/philadelphiacrimedata

Table 1: Column Description

Column	Description
Dc_Dist	A two character field that names the District boundary.
Psa	A single character field that names the Police Service Area boundary.
Dispatch_Date_Time	The date and time that the officer was dispatched to the scene.
Dispatch_Date	The date that the officer was dispatched to the scene.
Hour	The hour of the day that the officer was dispatched to the scene.
Location_Block	The location of crime generalized by street block.
UCR_General	The rounded crime code.
Text_General_Code	The generalized text for the crime code.
Police_Districts	A single character field that names the Police District boundary.
Month	The month that the officer was dispatched to the scene.
Lon	Longitude Number.
Lat	Latitude Number.

Table 2 below shows how the offense code in the column UCR_General maps to the offense description in the column Text_General_Code.

Table 2: Mapping UCR_General to Text_General_Code

UCR_General	Text_General_Code
100	HOMICIDE - CRIMINAL
100	HOMICIDE - JUSTIFIABLE
100	HOMICIDE - GROSS NEGLIGENCE
200	RAPE
300	ROBBERY FIREAM
300	ROBBERY NO FIREAM
400	AGGRAVATED ASSAULT FIREARM
400	AGGRAVATED ASSAULT NO FIREARM
500	BURGLARY RESIDENTIAL
500	BURGLARY NON-RESIDENTIAL
600	THEFTS-(EXCLUDING THEFT FROM VEHICLE)
600	THEFT FROM VEHICLE
700	MOTOR VEHICLE THEFT
700	RECOVERED STOLEN MOTOR VEHICLE
800	OTHER ASSAULTS
900	ARSON
1000	FORGERY AND COUNTERFEITING
1100	FRAUD
1200	EMBEZZLEMENT
1300	RECEIVING STOLEN PROPERTY
1400	VANDALISM / CRIMINAL MISCHIEF
1500	WEAPON VIOLATIONS
1600	PROSTITUTION AND COMMERCIALIZED VICE
1700	OTHER SEX OFFENSES Not Commercialized
1800	NARCOTIC / DRUG LAW VIOLATIONS
1900	GAMBLING VIOLATIONS
2000	OFFENSES AGAINST FAMILY AND CHILDREN
2100	DRIVING UNDER THE INFLUENCE, D.U.I
2200	LIQUOR LAW VIOLATIONS
2300	PUBLIC DRUNKENNESS
2400	DISORDERLY CONDUCT
2500	VAGRANCY / LOITERING
2600	ALL OTHER OFFENSES

The columns Lon and Lat were used to create a heat map of the entire dataset. Figure 1 shows where all of the crime is located, dark red indicating a lot of crime and light yellow indicating low amounts of crime. This figure is also representative of each year from 2006 to 2016, because the distribution of crime is similar year to year. The inner shell of the heat map is where most of the crimes occur, so the outside edge is going to be trimmed off. Furthermore, this inner shell will then be split into 3 groups corresponding to these locations on the map: Central, West, and North Philadelphia.

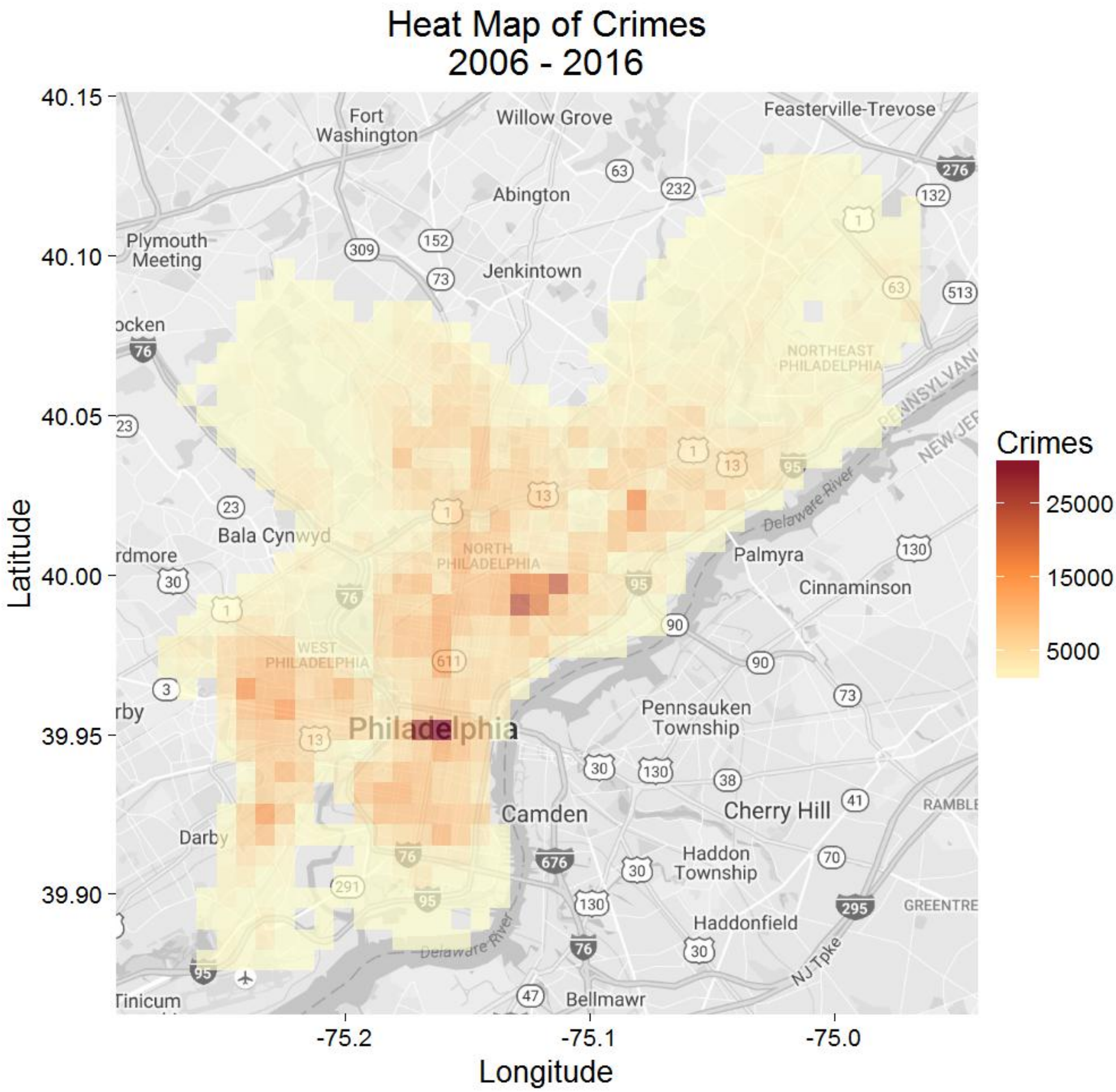


Figure 1: Heat Map of the Data

The Central Philadelphia subsection is shown below in Figure 2.

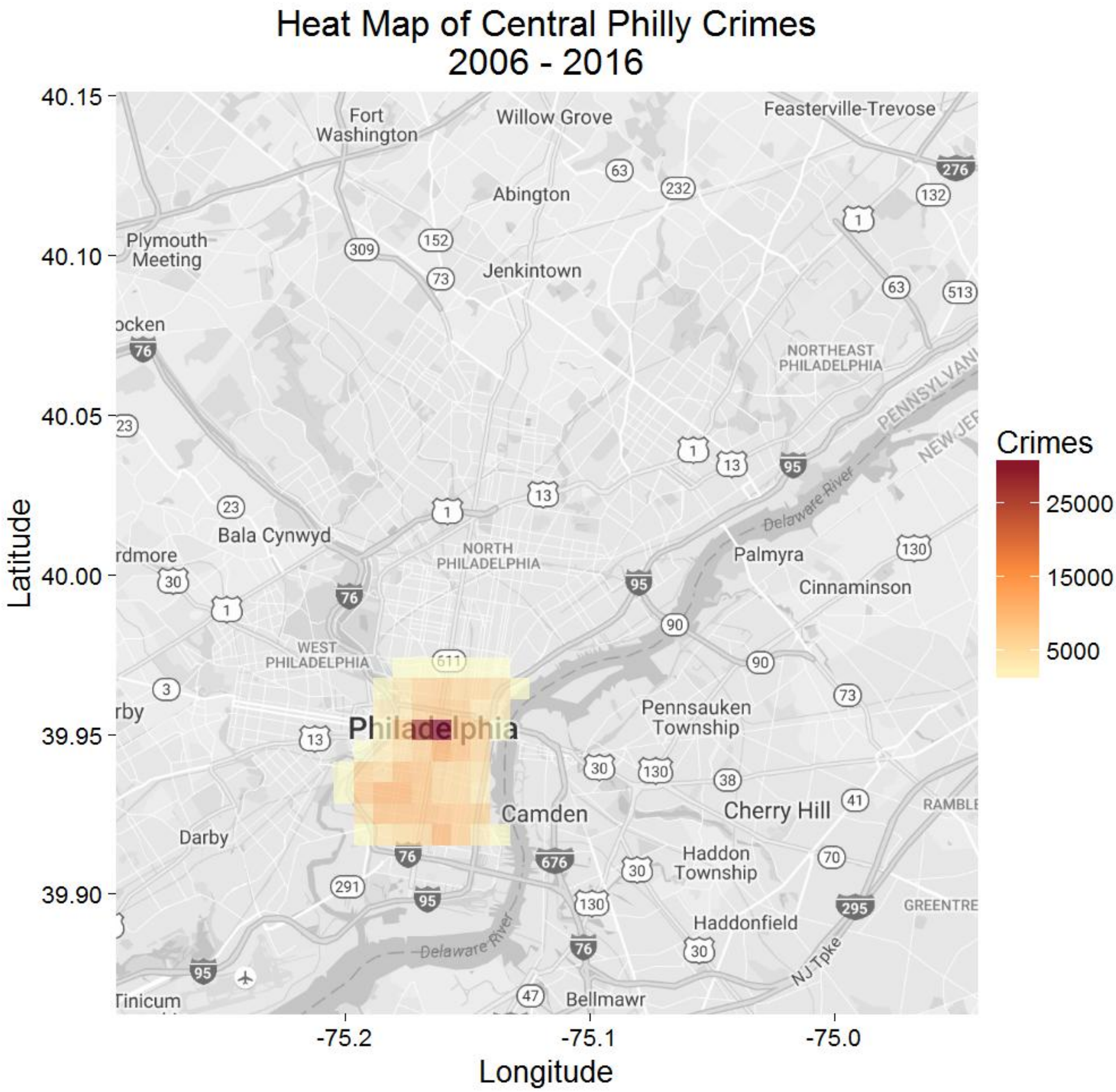


Figure 2: Heat Map of Central Philadelphia

The West Philadelphia subsection is shown below in Figure 3.

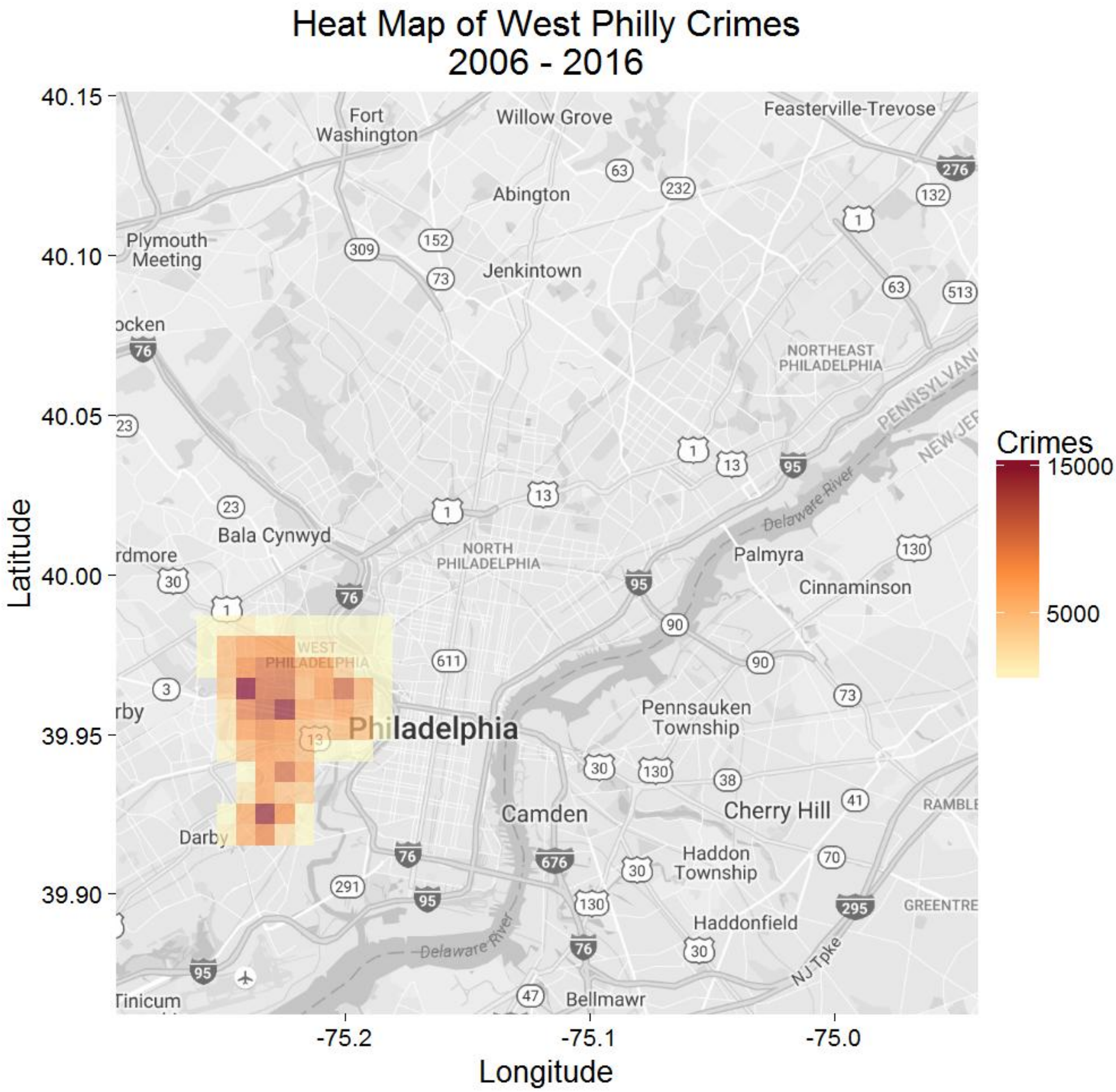


Figure 3: Heat Map of West Philadelphia

The North Philadelphia subsection is shown below in Figure 4.

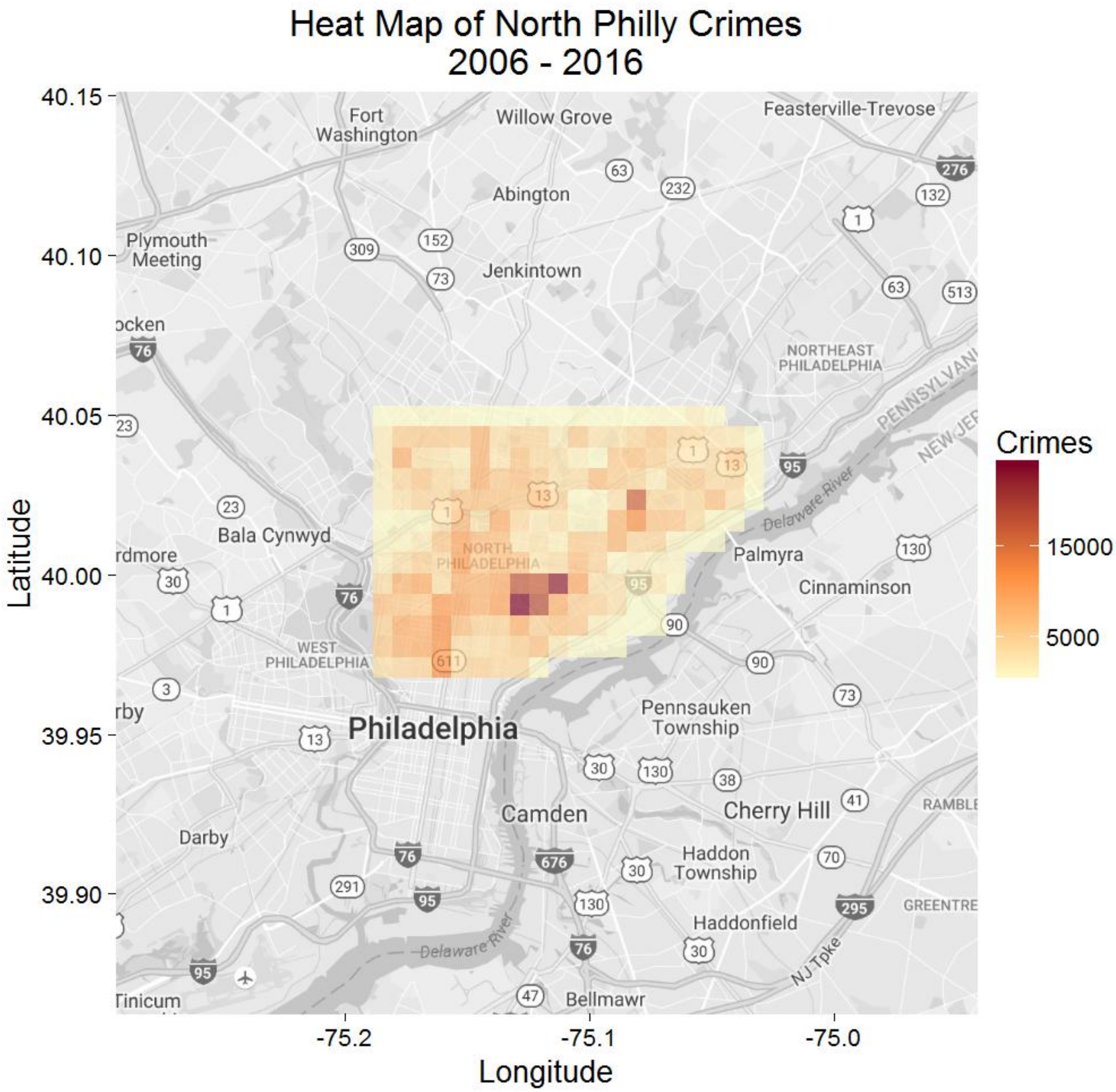


Figure 4: Heat Map of North Philadelphia

There were four more columns added to the dataset to facilitate the analysis of the external variables, Time and Location. These columns are described below in Table 3.

Table 3: Additional Columns

Column	Description
Day	The day of the week that the officer was dispatched to the scene.
Week	The week of the year that the officer was dispatched to the scene.
Year	The year that the officer was dispatched to the scene.
Area	The location of crime generalized by Central, West, and North Philly.

Data Cleansing

1. Removing variables.

There were two columns, Dispatch_Time and Dc_Key that were removed. Dispatch_Time was removed because the column Dispatch_Date_Time captures this information. The column Dc_Key was removed because this is a unique ID number based on the location, type, and time of crime which are represented by other columns.

2. Removing rows due to one or more NA's present.

Table 6 below shows the percentage of NA's within each column. There were very small amounts of NA's present in each column, so rows were removed if there was at least one NA present across all columns.

Table 4: Columns with NA's

Column	Percent NA
Dc_Dist	0%
Psa	0%
Dispatch_Date_Time	0%
Dispatch_Date	0%
Hour	0%
Location_Block	0%
UCR_General	0.01%
Text_General_Code	0%
Police_Districts	0.89%
Month	0%
Lon	0.78%
Lat	0.78%

3. Daylight Savings Time Violations

There were 15 instances of crime at 2am on days where 2am didn't occur due to daylight savings moving forward. These violations were handled by moving them up to 3am. There was no indication in the dataset for days when 1am occurred twice due to daylight savings moving backward. This was handled by taking all 1am data on corresponding dates, and splitting them in half, one half was placed under the first 1am time, and the second half was placed under the second 1am time. There were a total of 529 instances that had to be split.

4. Making sure there is a 1 hour increment.

There were 258 hours unaccounted for in this dataset. These were handled by replicating the information in the same hour from a week ago. The previous week was chosen to respect the day of the week that the hour comes from. For example, crime rates on Sundays may be different than Saturdays, so using a Saturday hour to represent a missing Sunday hour may not be appropriate.

5. Adjusting Anomalies

There was only one anomaly found when plotting hourly, daily, weekly, and monthly crimes across the entire dataset. The point circled in red within Figure 5 is the anomaly. The information in this data point from the column Block_Location shows many non-residential burglaries at a public storage center. This repeated information may be true, but it inflates the value of the data point such that it becomes an anomaly. This event occurred at 8am which has an average of 18 crimes across the 10 years, so these burglaries were removed until this data point reached 18 crimes.

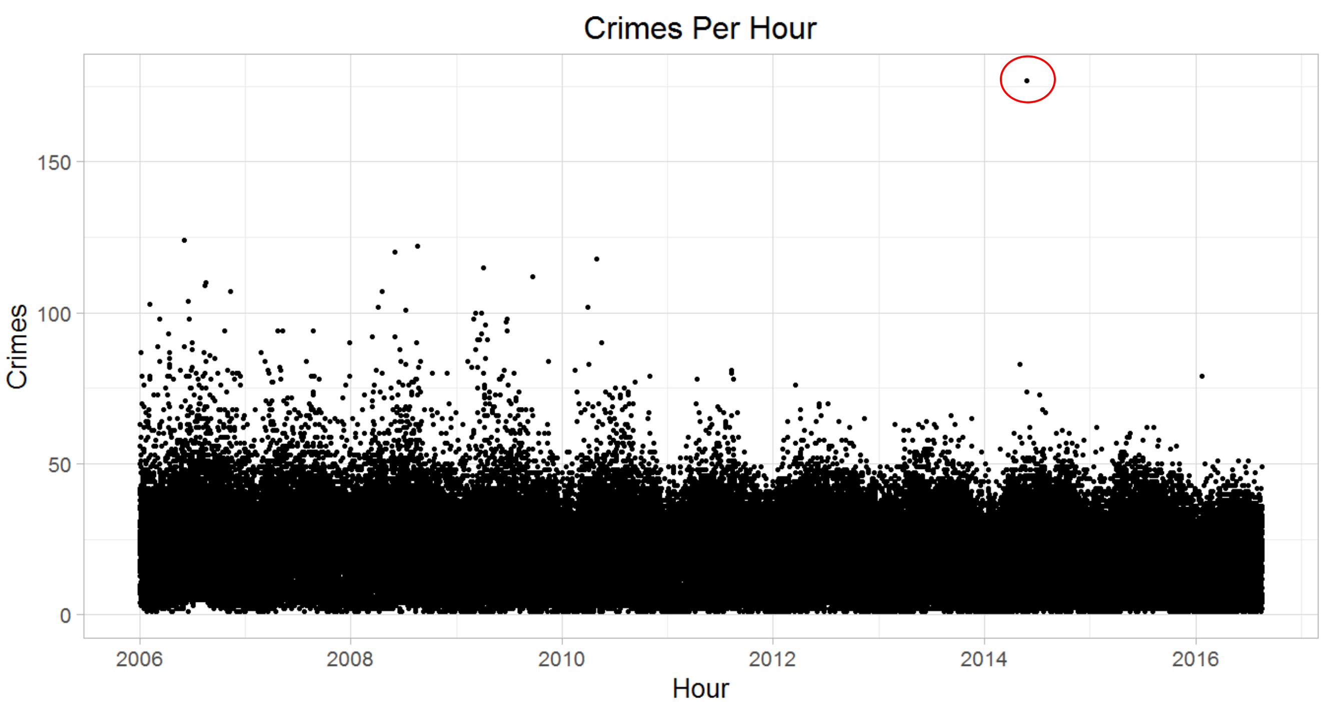


Figure 5: Hourly Crime Anomaly

Trend & Seasonality

The three external variables of interest will be shown in the following order: Time, Location, and Offense. The plot of 10 years of hourly crimes is shown below in Figure 6. There is annual seasonality present because of the periodic bumps at an annual frequency. Let's zoom in on the first week of this plot to see how hourly crimes behave at a more detailed level.

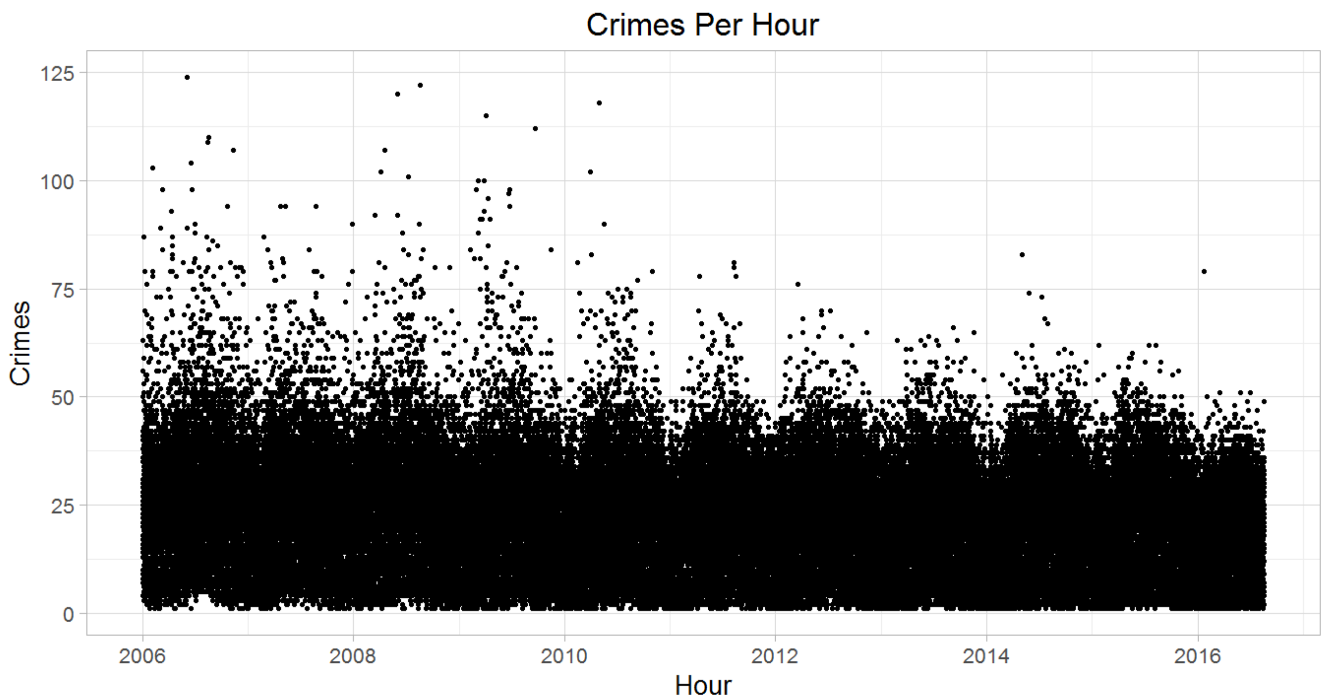


Figure 6: 10 Years of Hourly Crimes

The plot of 1 week of hourly crimes is shown below in Figure 7 with each hour given a color to highlight the behavior of hourly crimes. This plot shows seasonality on a daily basis with peaks during the afternoon and night, and valleys during the early morning. This makes since, as people are awake during the afternoon and night, and asleep during the early morning.

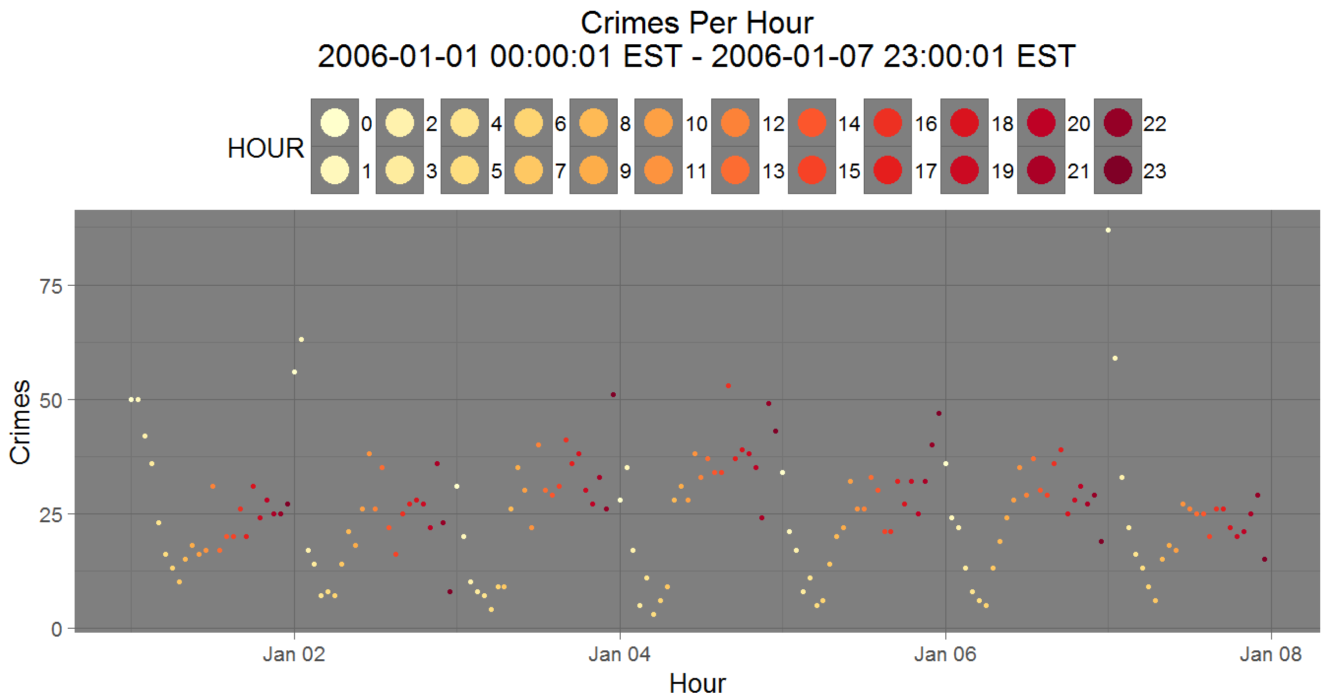


Figure 7: 1 Week of Hourly Crimes

The plot of 10 years of daily crimes is shown below in Figure 8 with each month given a color to highlight the behavior of daily crimes. This plot shows seasonality on an annual basis with peaks during warmer months and valleys during colder months. This makes since, as people are outside more often during the warmer months and the day is longer. Whereas during colder months, people spend more time inside, and the day is shorter. The first three months of this plot was zoomed in on to see how daily crimes behave at a more detailed level, and no seasonality was present. Let's see if daily crimes vary during the week.

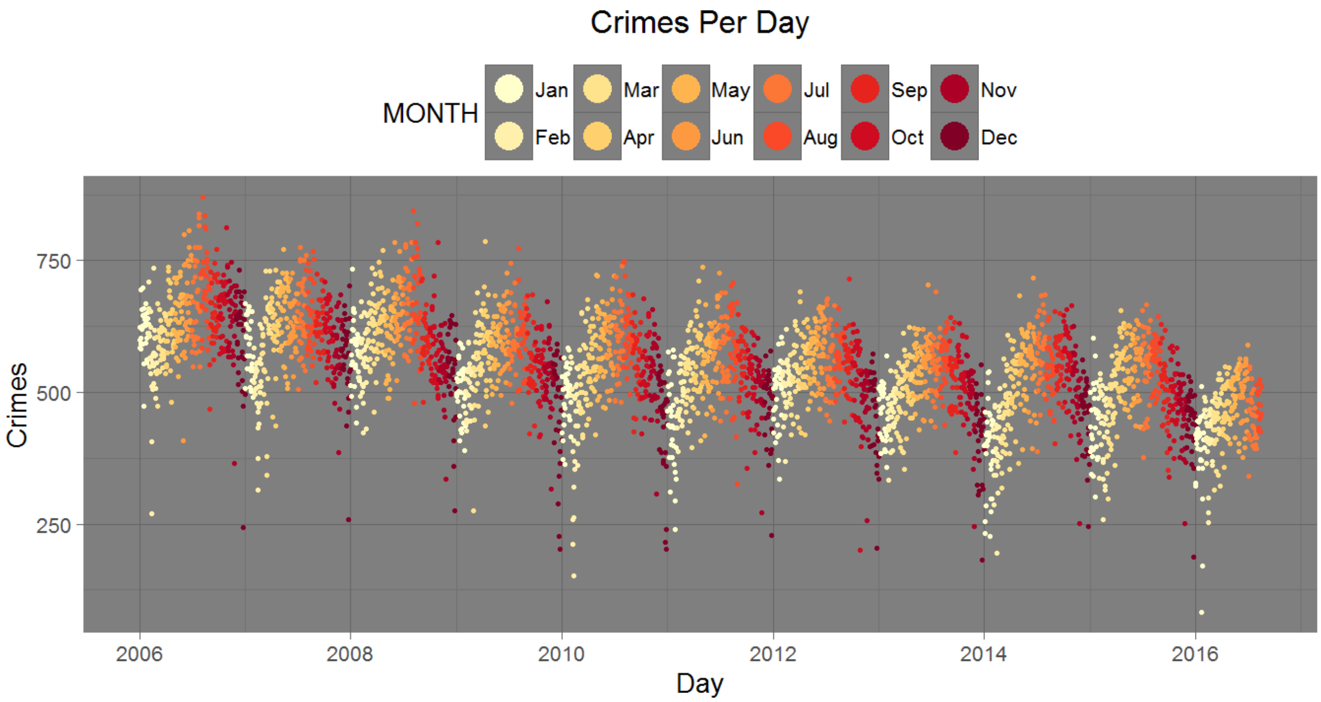


Figure 8: 10 Years of Daily Crimes

The behavior of daily crimes throughout the week is shown below in Figure 9. This shows that Sunday has less crimes but the other days have roughly the same amount of crimes, on average. Other performance measures such as the total sum, median, and quartiles show similar behavior as the average.

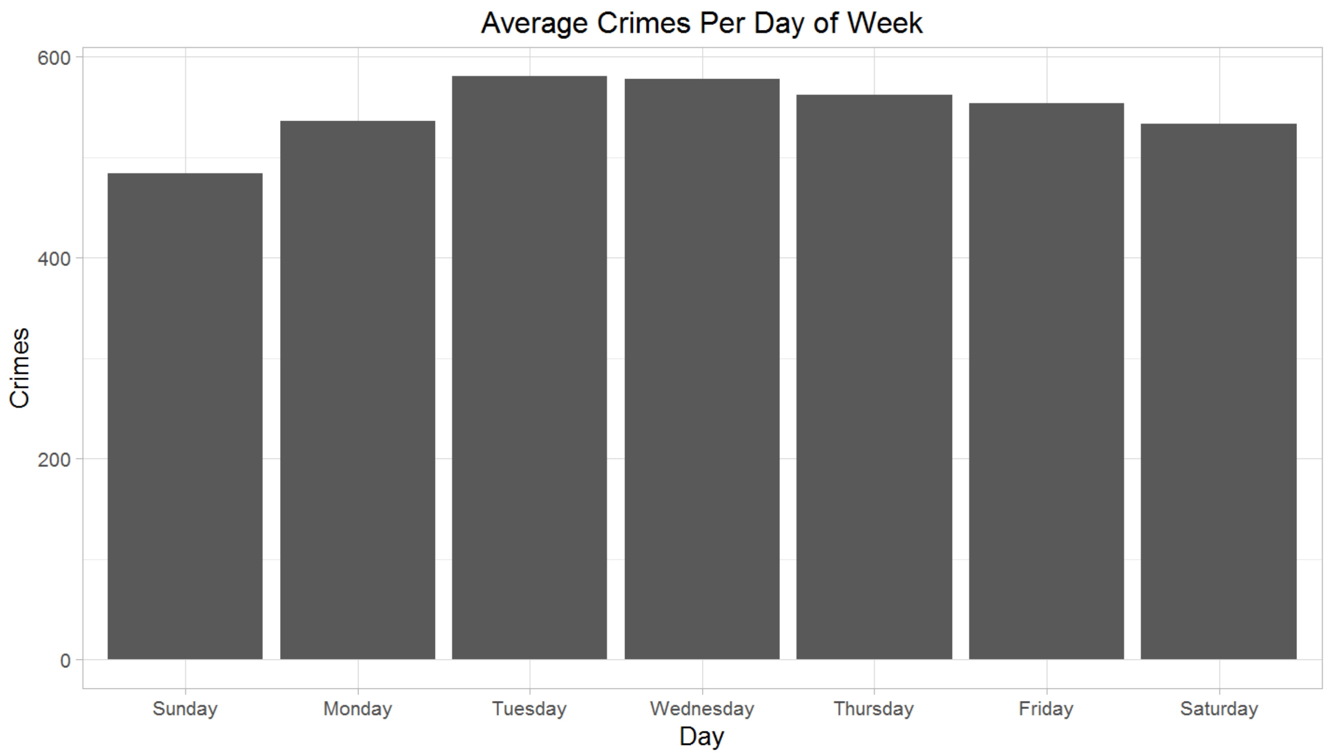


Figure 9: Average Crimes by Day of Week

The plot of 10 years of weekly crimes is shown below in Figure 10. There is annual seasonality present because of the periodic bumps at an annual frequency. There are very low values during the transition from year to year, and then another low value as the very last point. These low values can be explained by the number of days in each week.

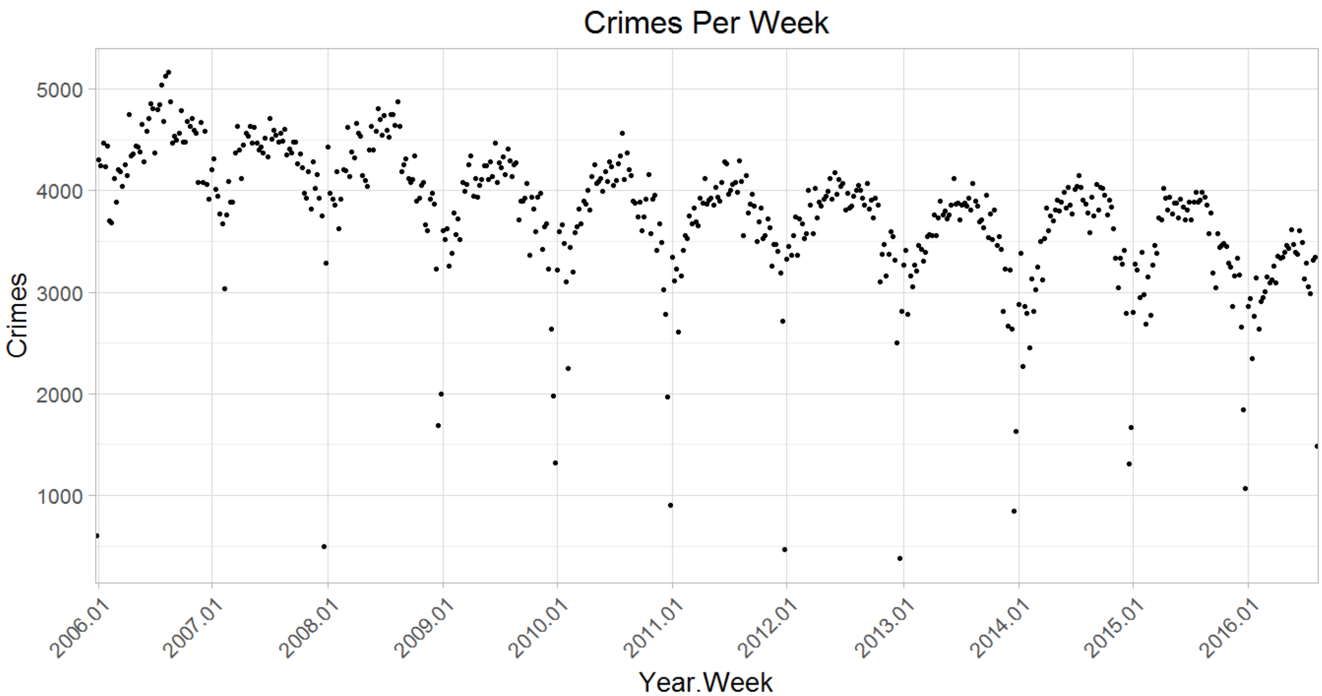


Figure 10: 10 Years of Weekly Crimes

This plot of weekly crimes in Figure 11 colors each data point based on the number of days it has, as to highlight the behavior of weekly crimes. The reason for why some weeks have less days is due to the way the calendar year shifts by one day every year, which often reduces the number of days for the last and first week of a year. Now the reason that the last data point is so low is because the last day in the dataset is in the middle of a week, hence an incomplete week.

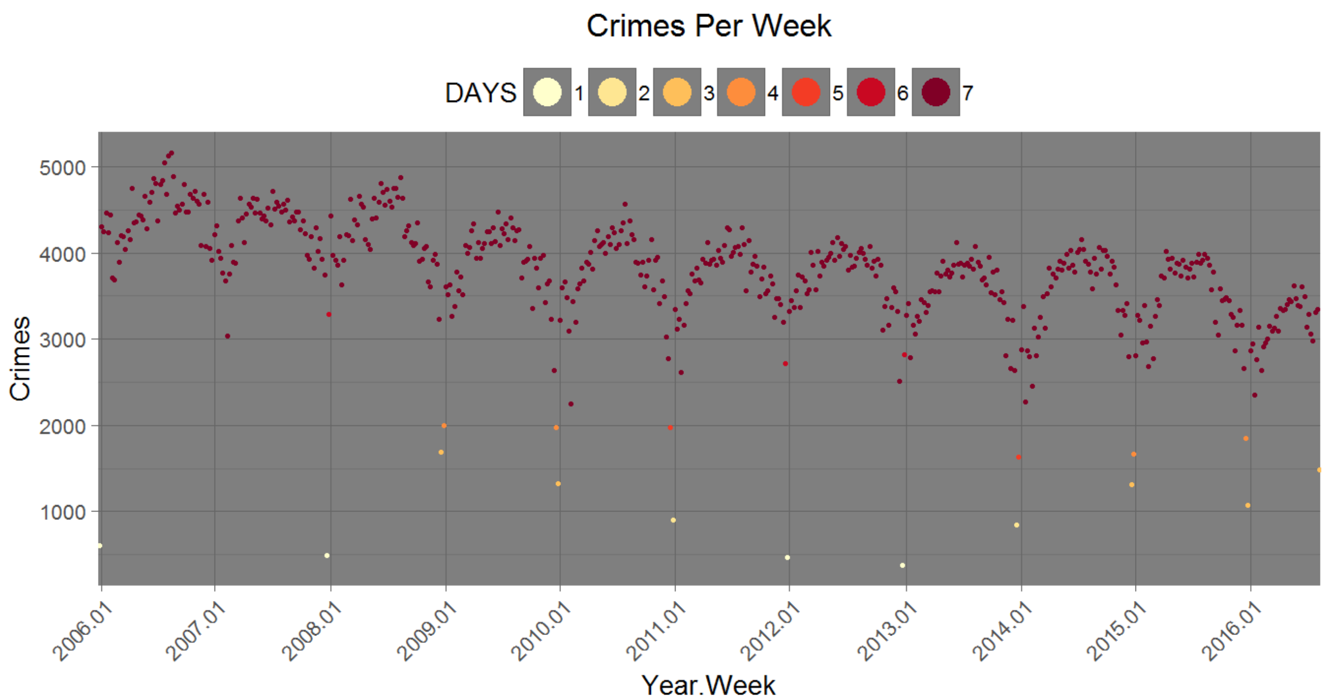


Figure 11: 10 Years of Weekly Crimes Colored by Days per Week

The plot of 10 years of monthly crimes is shown below in Figure 12. There is annual seasonality present because of the periodic bumps at an annual frequency. Overall, hourly, daily, weekly, and monthly crimes show evidence of seasonality. All of the plots on the 10 year scale have also shown a downward trend up until the start of 2012, no trend after the start of 2012, and then a shift downward at the start of 2016. This trend may be due to better policing and/or an increase in police staffing.

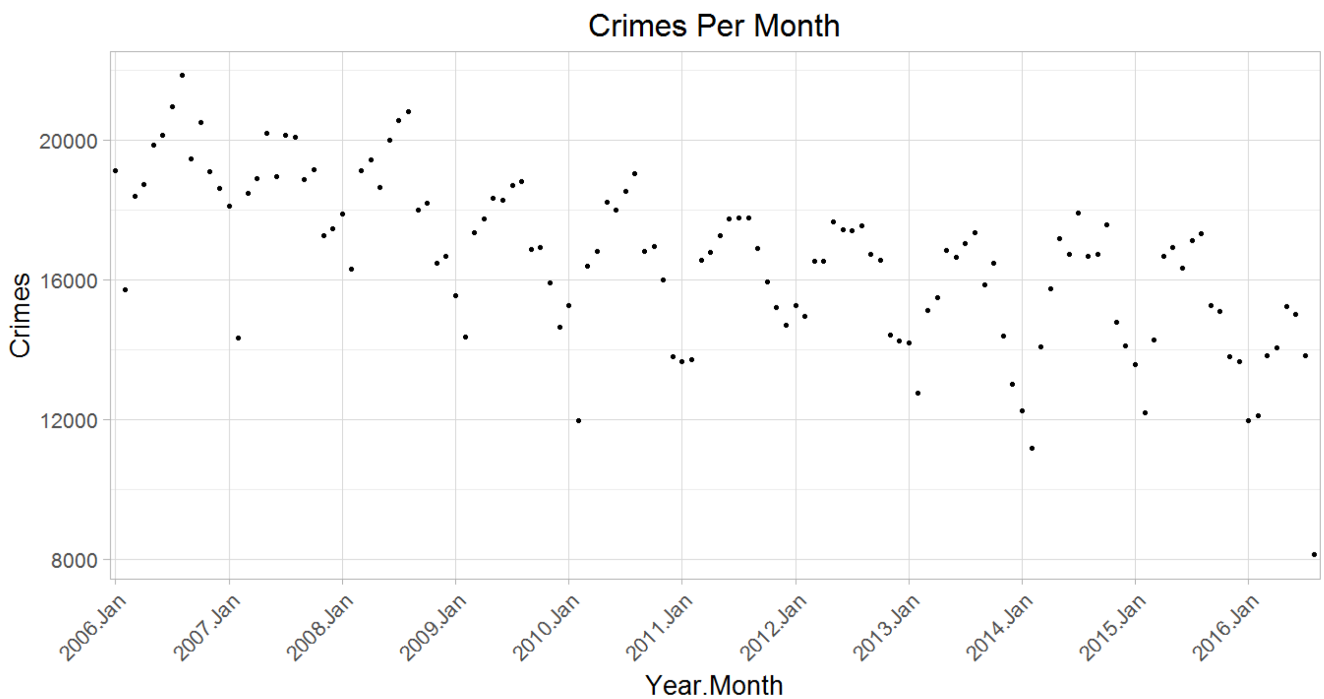


Figure 12: 10 Years of Monthly Crimes

The next external variable to evaluate is Location. The columns in the dataset corresponding to location are factor columns, therefore each level of each column can be plotted as its own subset. The number of plots that can be made for each of the location related columns is shown below in Table 5. The column Location_Block will be ignored due to its large amount of levels. A 10 year plot was made for each level of each column, for hourly, daily, weekly, and monthly crimes. Most of the plots have seasonality but only 3 subsets, Central, West, and North Philadelphia weekly crimes will be evaluated further.

Table 5: Levels in Location Columns

Column	Levels
Dc_Dist	25
Psa	30
Location_Block	114446
Police_Districts	22
Area	4

The following 3 plots are 10 years of weekly crimes for Central, West, and North Philadelphia. Central and North Philadelphia share the same behavior as the 10 years of weekly crimes for all of Philadelphia, whereas West Philadelphia is different in that it has no trend. This may be due to local improvements in policing for Central and North Philadelphia, and a lack of improvement for West Philadelphia.

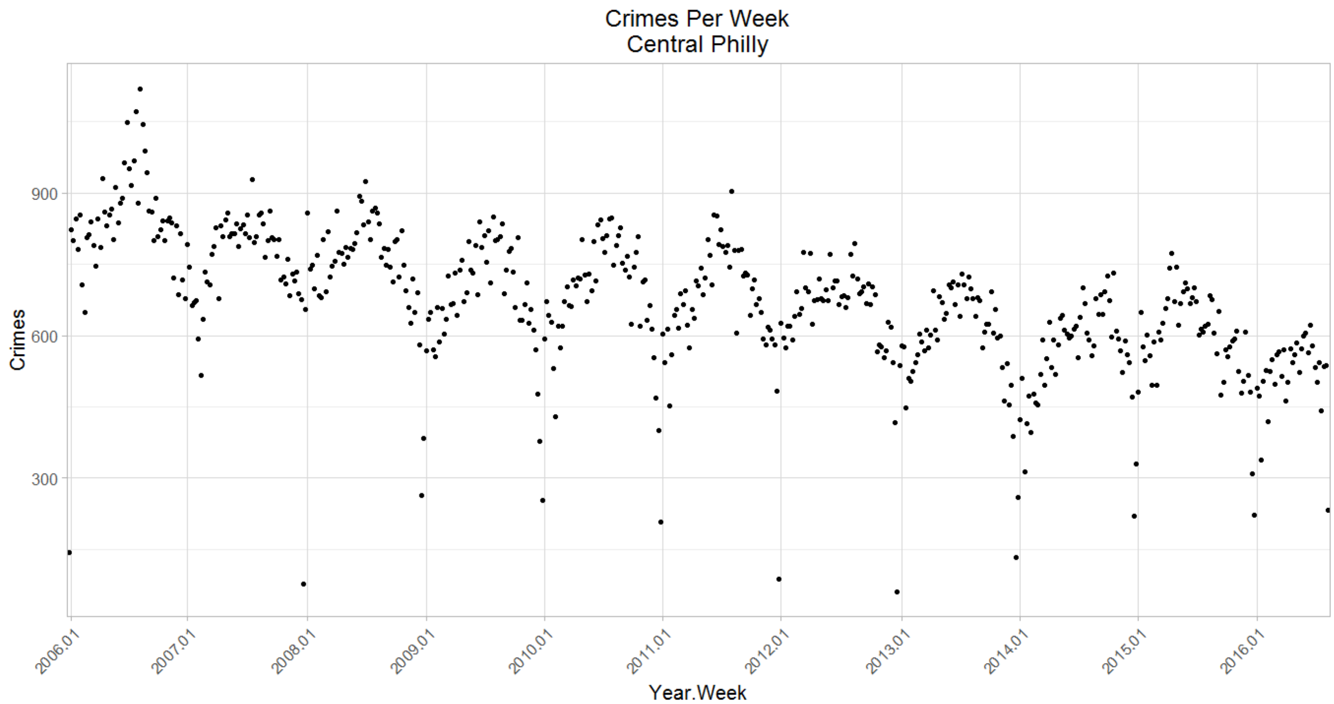


Figure 13: 10 Years of Weekly Crimes for Central Philadelphia

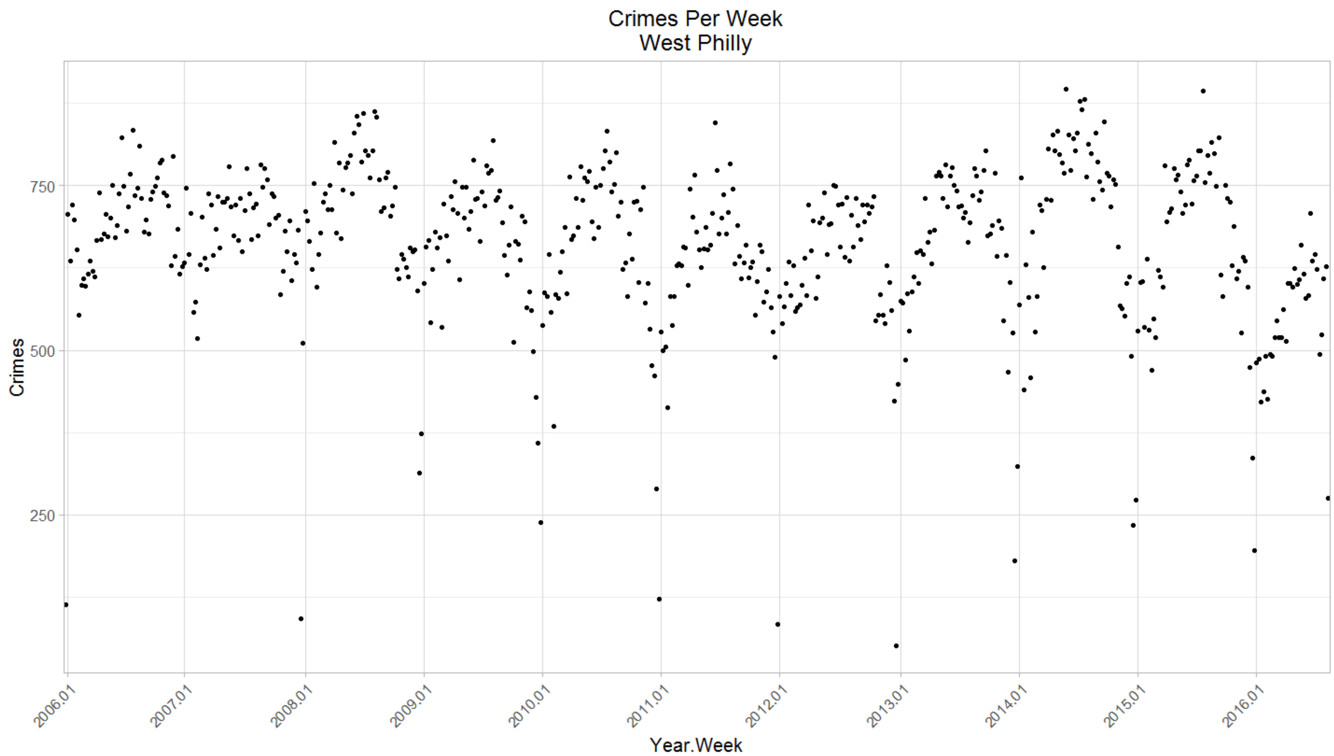


Figure 14: 10 Years of Weekly Crimes for West Philadelphia

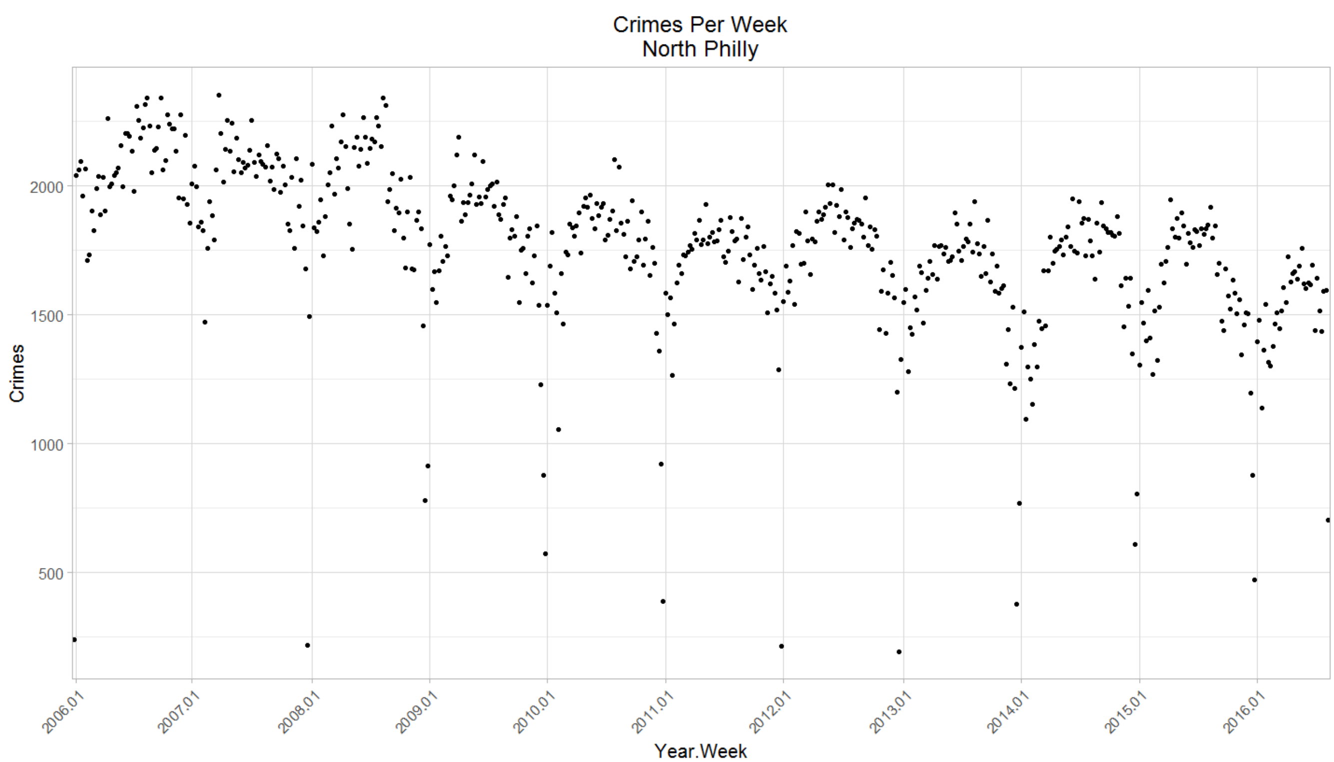


Figure 15: 10 Years of Weekly Crimes for North Philadelphia

The next external variable is Offense. The columns in the dataset corresponding to offense are factor columns, therefore each level of each column can be plotted as its own subset. The number of plots that can be made for each of the offense related columns is shown below in Table 6. A 10 year plot was made for each level of each column, for hourly, daily, weekly, and monthly crimes.

Table 6: Levels in Offense Columns

Column	Levels
UCR_General	26
Text_General_Code	33

Offense will no longer be evaluated. Hourly, daily, and monthly crimes will no longer be evaluated either. The scope for this analysis is shrinking due to the amount of work required to fully cover Time, Location, and Offense as external variables thus far. Therefore, this analysis is now interested in forecasting weekly crimes in Central Philadelphia, West Philadelphia, and North Philadelphia to show how time and location can be used to model Philadelphia crime.

Smoothing Methods

Three smoothing methods will be used to further highlight the behavior of weekly crimes for all of Philadelphia. The following 3 plots show 10 years of weekly crimes and the smoothers. The moving average smoother in Figure 16 has a span of 5 weeks which was chosen through trial and error until a desirable visual was achieved. Tukey's Running Median and LOWESS smoothers were run with the default parameter setting chosen by R. These smoothers all support the behavior of seasonality, a downward trend up until the start of 2012, no trend after the start of 2012, and then a shift downward at the start of 2016 for weekly crimes.

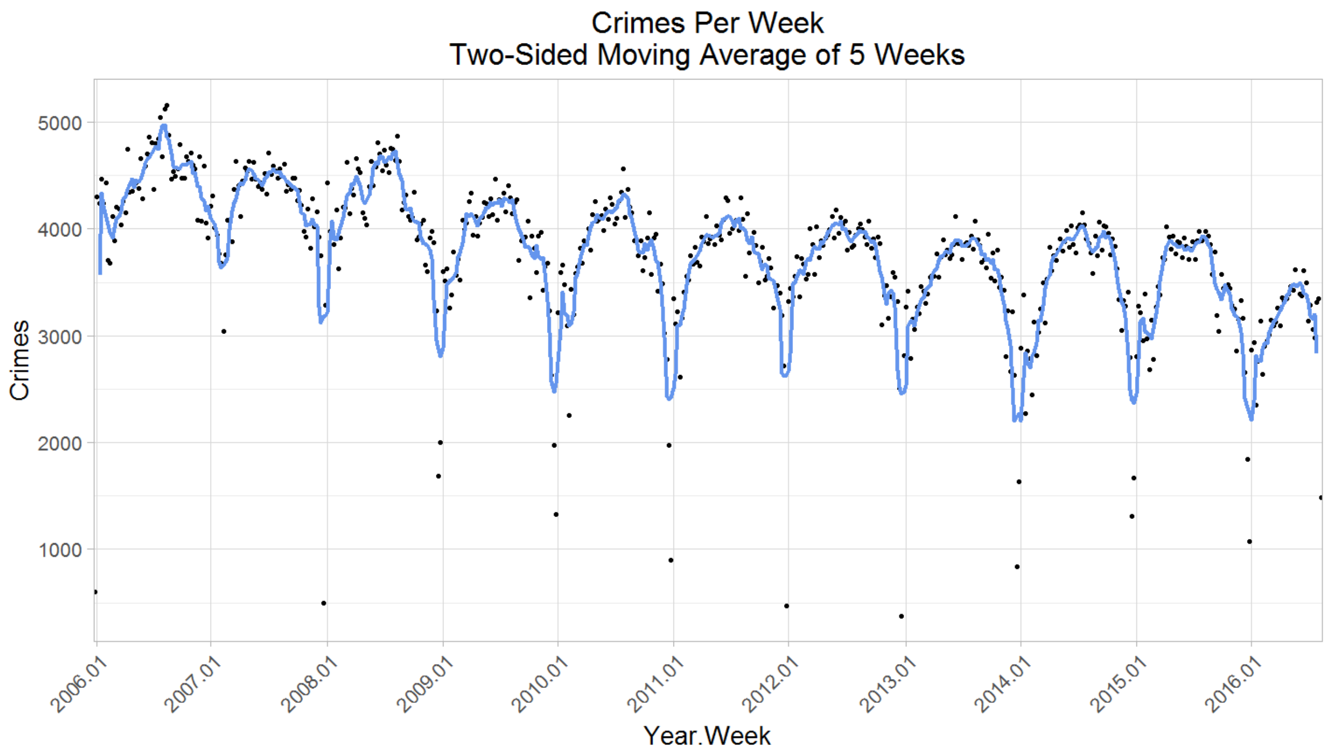


Figure 16: 10 Years of Weekly Crime & Moving Average

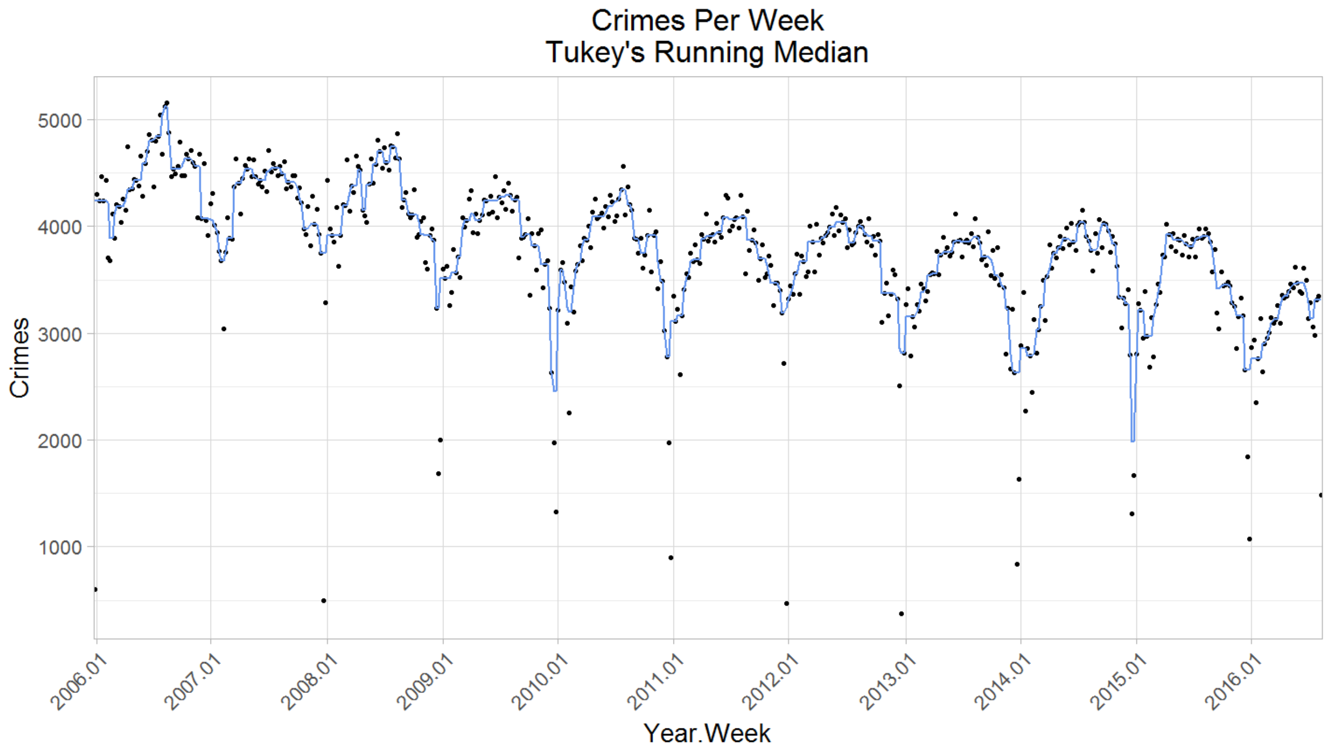


Figure 17: 10 Years of Weekly Crime & Tukey's Running Median

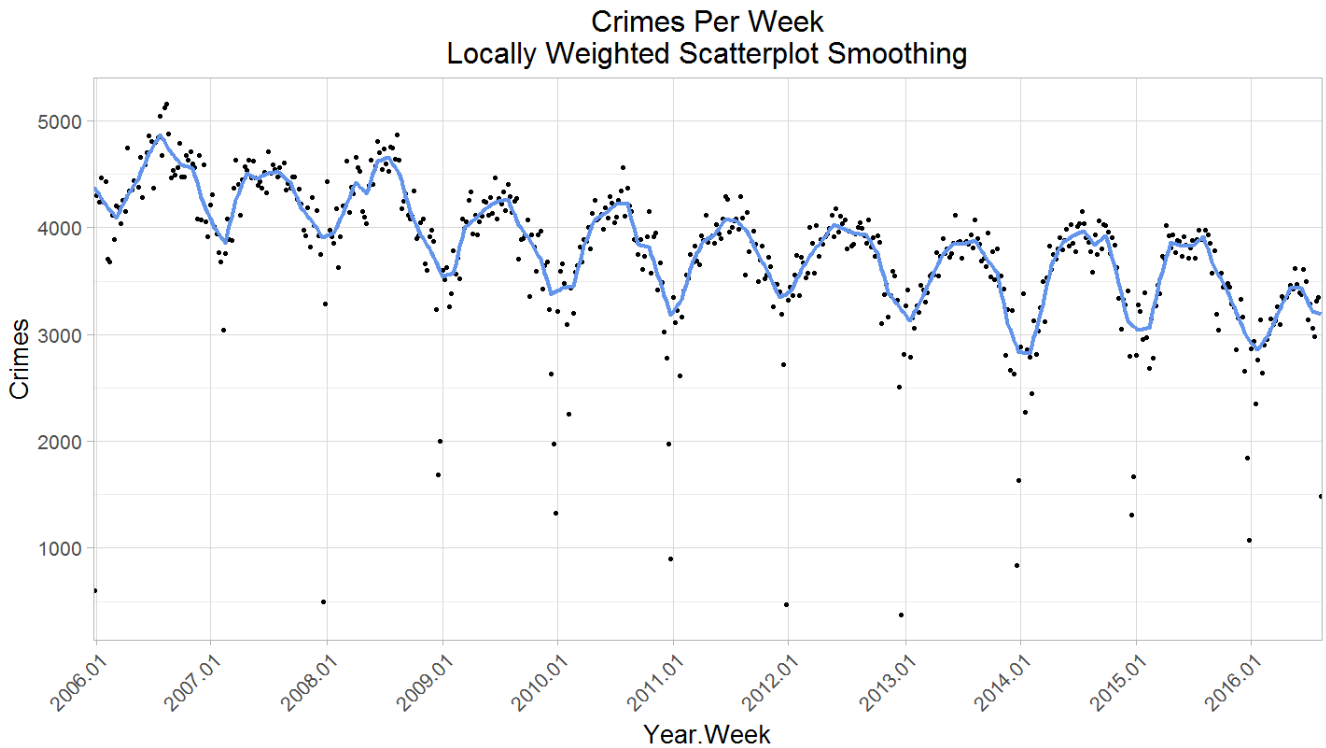


Figure 18: 10 Years of Weekly Crime & LOWESS

Autocovariance

The autocovariance of weekly crimes for all of Philadelphia will be evaluated through an ACF plot. The autocorrelation levels for weekly crimes is very high which supports the seasonality seen in the previous weekly time series plots. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below support the annual seasonality with a large gradual curve. The sample mean and variance for 10 years of weekly crime is 3758.796 and 451025.1 respectively.

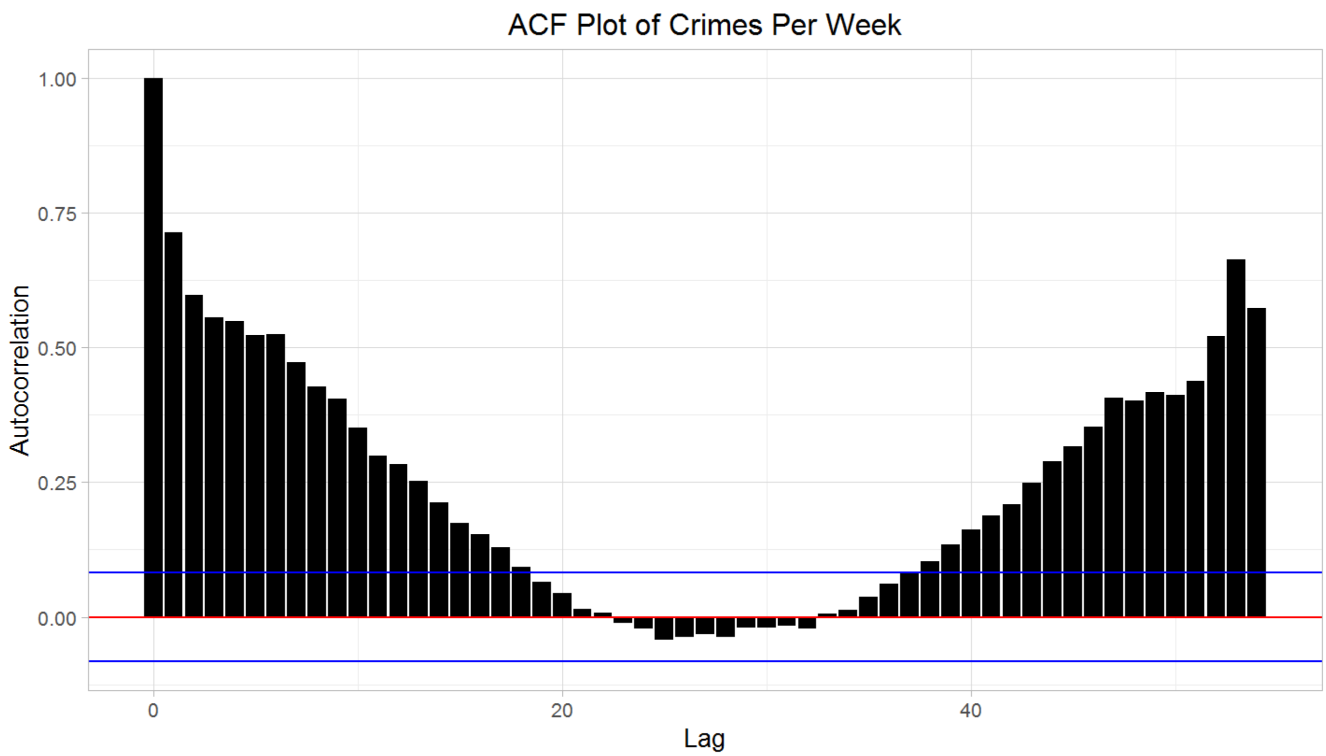


Figure 19: ACF Plot of 10 Years of Weekly Crime – Lag of 54 Weeks

Stationarity

The stationarity of weekly crimes for all of Philadelphia will be evaluated though a variogram. The variogram levels for weekly crimes has a distinct curve to it which supports the seasonality seen in the previous weekly time series plots. This figure compares the changing variance throughout one season, a lag up to 54 weeks. You can see that the variogram is at its highest during the middle of the season, which makes sense because this is the summer time where crime levels are at its highest annually. Then the variogram drops back down near the end of the season which is also consistent with the seasonal behavior of crime levels being at its lowest during the winter. The variogram shows that the seasonality of weekly crimes makes it non-stationary.

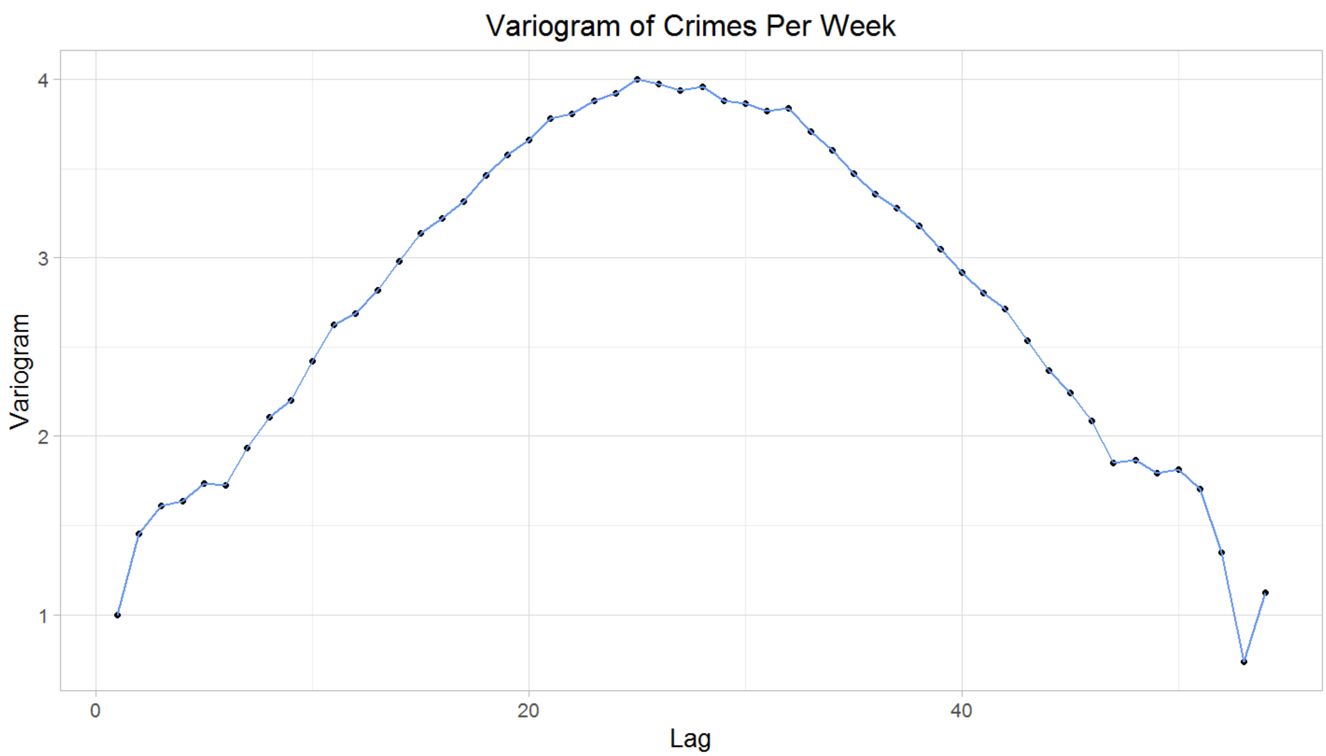


Figure 20: Variogram of 10 Years of Weekly Crimes – Lag of 54 Weeks

Removing Trend & Seasonality

The trend and seasonality of weekly crimes for all of Philadelphia will be removed through linear regression. The dataset created to remove trend and seasonality for 10 years of weekly crimes is shown below in Table 7. The SIN and COS variables use sin and cos functions to compute $\frac{2*\pi*T}{54}$ where T is the week of the year of the observation, for every observation. These SIN and COS variables capture the seasonal period of 54 weeks. The MA.54 variable is the moving average smoother with a span of 54 weeks to capture the direction of trend year to year. The DAY variable is a categorical variable that captures the low points of weekly crime as shown in Figure 11. The CRIMES variable is the response variable for weekly crimes.

Table 7: Dataset for Weekly Crime Linear Regression Model

	SIN	COS	MA.54	DAY	CRIMES
1:	0.11609	0.99324	4190.04	1	600
2:	0.23062	0.97304	4190.04	7	4304
3:	0.34202	0.93969	4190.04	7	4243
4:	0.4488	0.89363	4190.04	7	4465
5:	0.54951	0.83549	4190.04	7	4237

561:	0.64279	-0.766	3202.37	7	3057
562:	0.54951	-0.8355	3202.37	7	2984
563:	0.4488	-0.8936	3202.37	7	3312
564:	0.34202	-0.9397	3202.37	7	3345
565:	0.23062	-0.973	3202.37	3	1483

The fit of the linear regression model is shown below in Figure 21. The model captures those low drops between the transitions of years very well. The model also follows the center of the data very well.

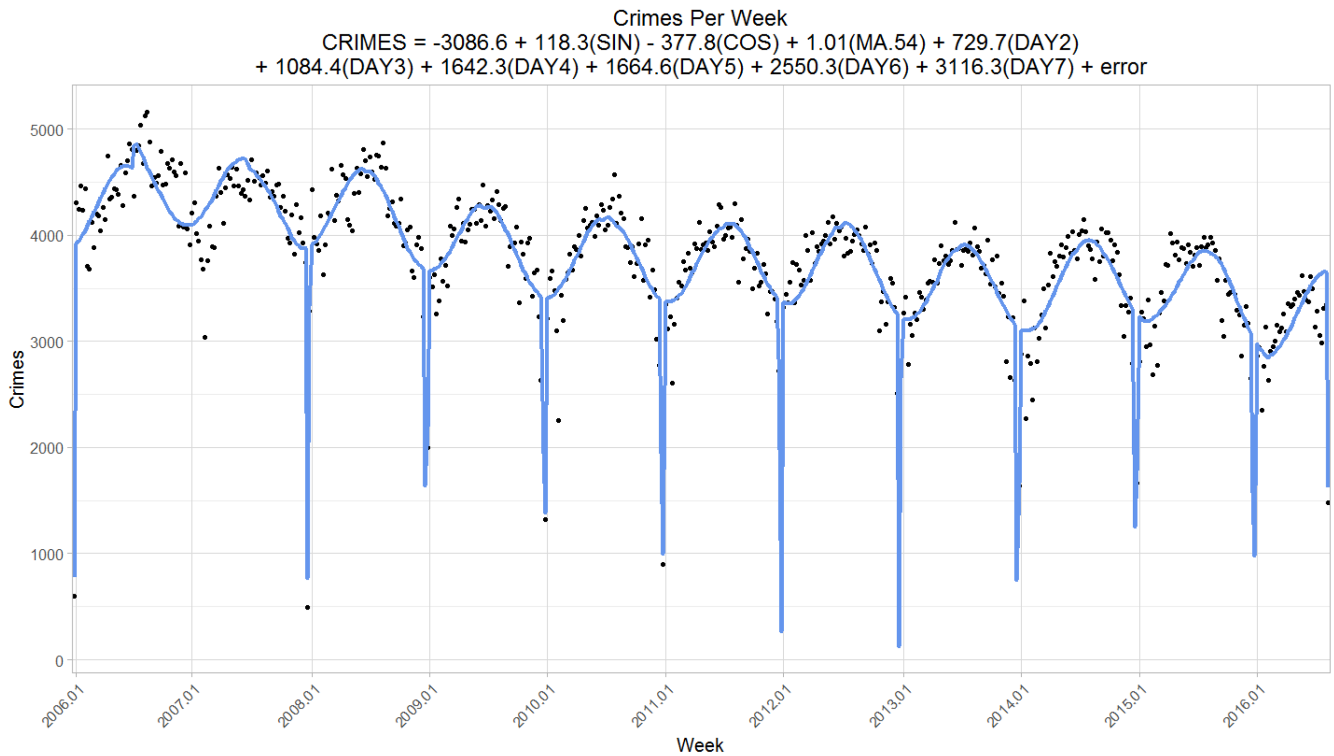


Figure 21: 10 Years of Weekly Crime & Linear Regression Fit

The residual time series of the linear regression model is the result of trying to remove trend and seasonality from weekly crimes, as shown below in Figure 22. This plot shows that annual seasonality was removed successfully. Let's verify the removal of trend and seasonality with the ACF plot of these residuals.

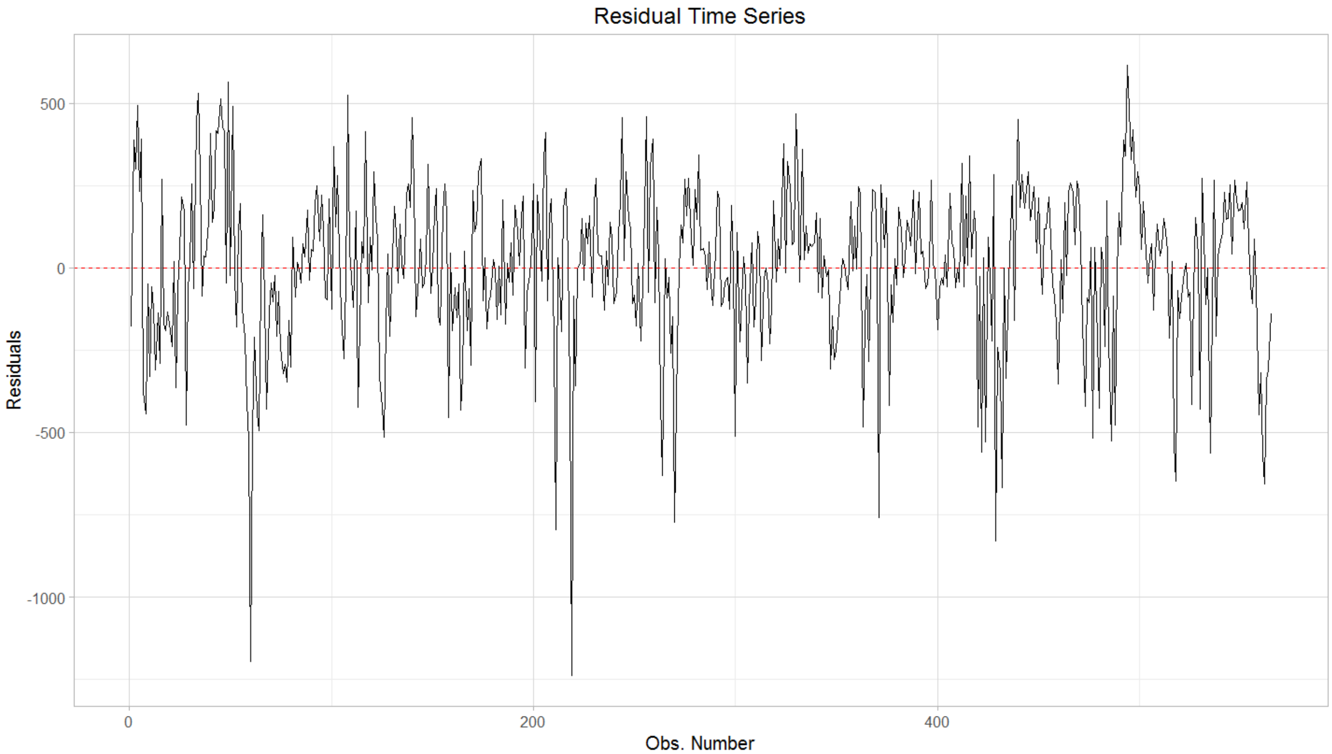


Figure 22: Residual v. Time for Weekly Crime Linear Regression Model

The autocorrelation levels of the residuals are not too high. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below show that there's room for improvement, but most of trend and seasonality was removed.

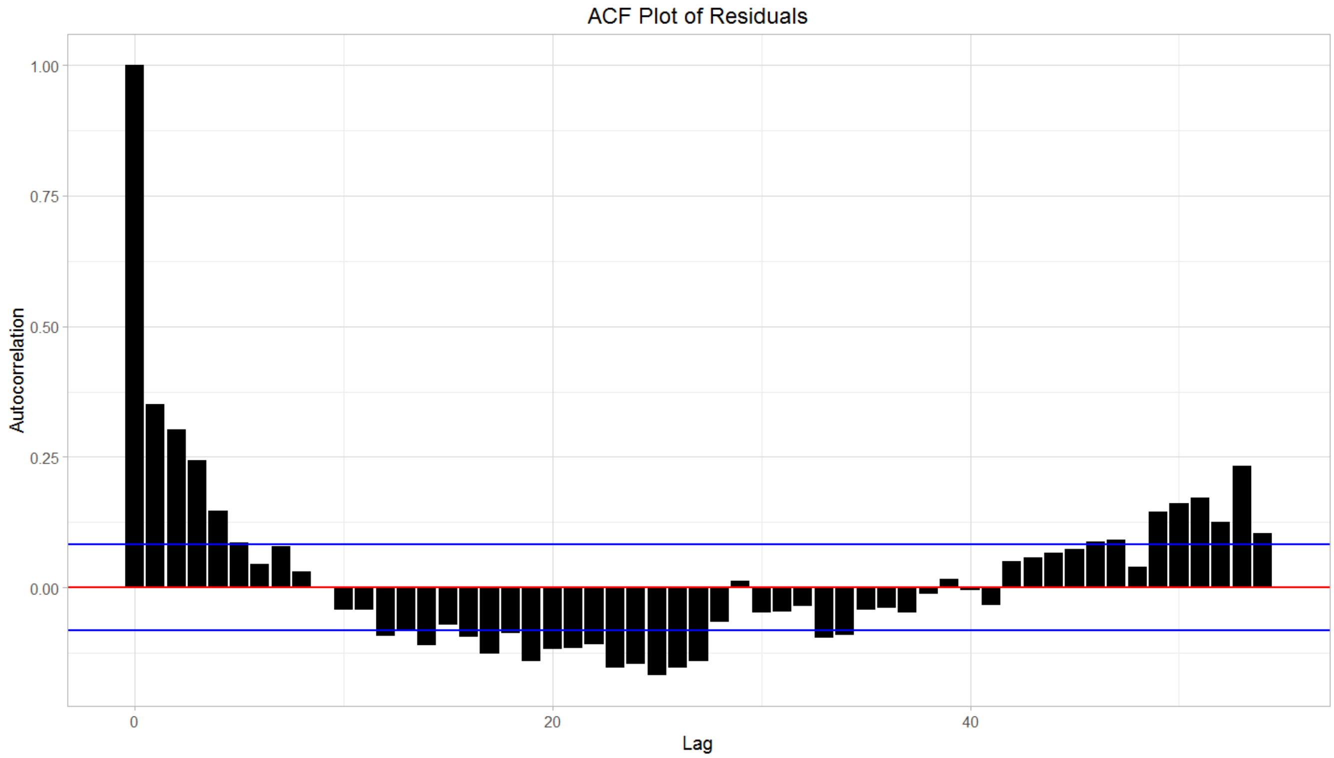


Figure 23: ACF Plot of Residuals for Weekly Crime Linear Regression Model

The residuals v. fitted plot show that a constant mean was achieved. Variance doesn't appear constant, but this is due to the distinct grouping of low values for incomplete weeks, and higher values for full weeks. Therefore this type of behavior for the variance makes sense, where variance is constant for weeks with the same days. This plot also shows that the residuals are centered at zero.

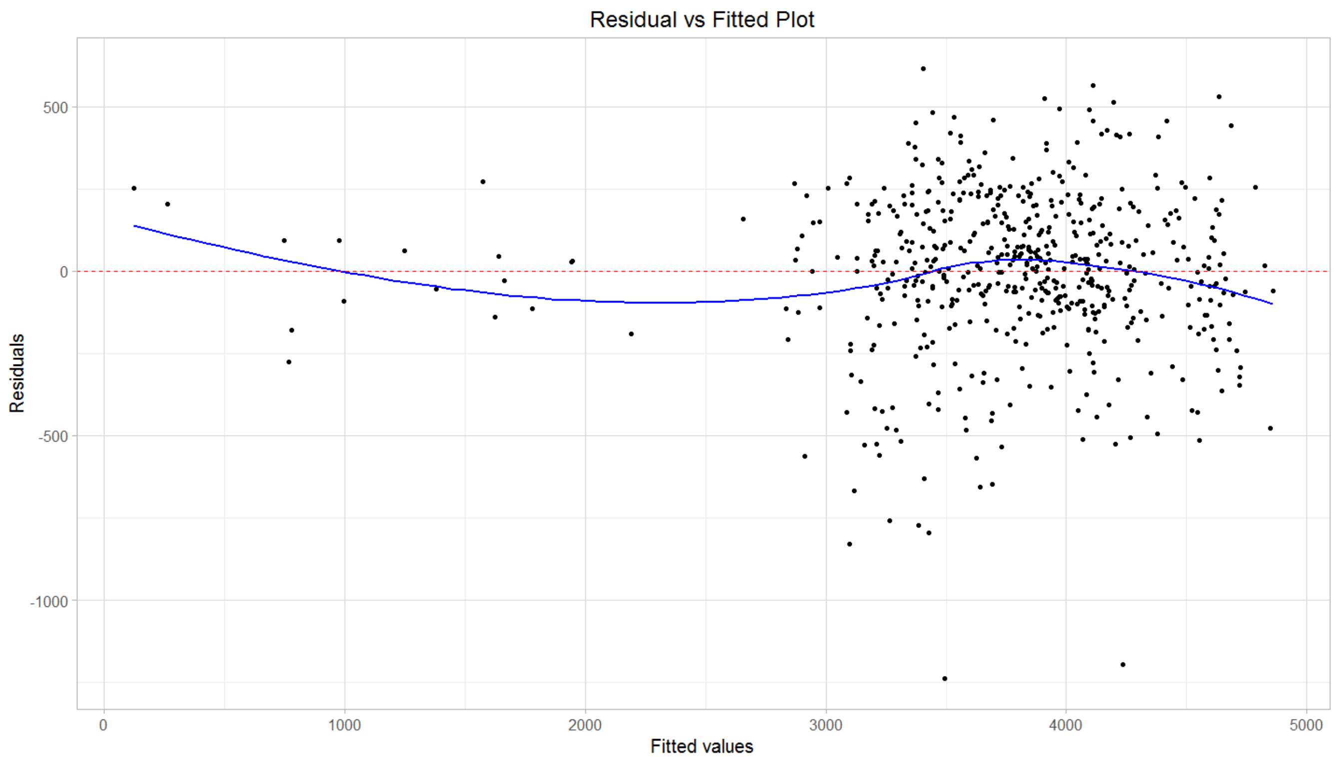


Figure 24: Residual v. Fitted for Weekly Crime Linear Regression Model

Power transformations on the weekly crime linear regression model would not improve its performance as proven by the relationships between the response variable and the continuous regressor variables in the pairs plot below. The lower triangle is the pairs plot section, the upper triangle is the correlation corresponding to each of the pairs plot, and the diagonal contains histograms of each individual regression variable. The most important plots to look at are the first three in the bottom row. These scatter plots indicate no visible relationship between the response and the regressor variables. This evidence shows that transforming the regression variables may not be necessary.

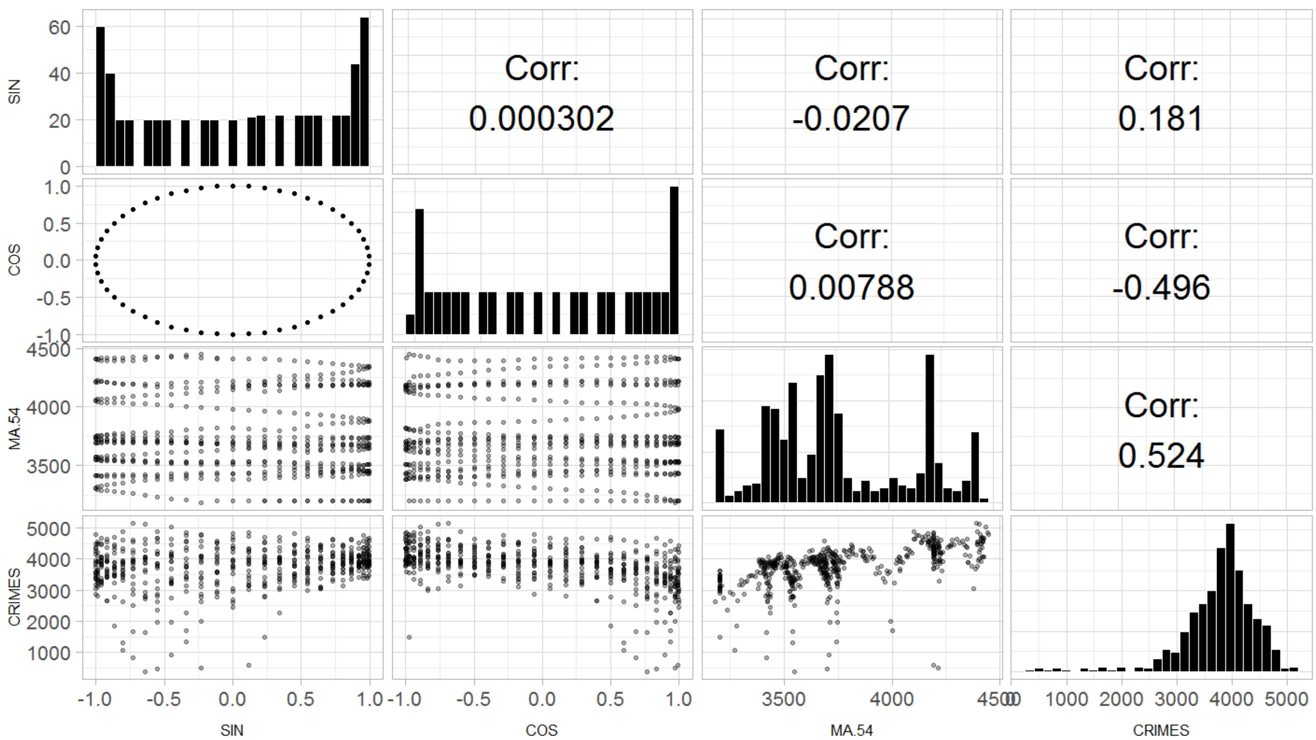


Figure 25: Pairs Plot of Regression Variables for Weekly Crime Linear Regression Model

Holt-Winters Forecasting

A decision matrix for determining how to choose the appropriate exponential smoothing method will be proposed. Then the proper exponential smoothing method will be used to fit and forecast weekly crimes for Central, West, and North Philadelphia.

The first row of the matrix below should be used if your data has no trend, whereas the second row should be used if your data has trend. The first column should be used if your data has no seasonality, the second column should be used if your data has additive seasonality, and the third column should be used if your data has multiplicative seasonality. Only one of these plots will apply to your data based on the criterion of trend and seasonality. This matrix shows that single exponential smoothing handles 1 instance, even if the data is non-stationary as shown in the plot. Double exponential smoothing handles 1 instance, and triple exponential smoothing handles 4 instances. The reason for these assignments is because single exponential smoothing estimates the mean, double exponential smoothing estimates the mean and trend, and triple exponential smoothing estimates the mean, trend, and seasonality. The row 2 – column 2 plot applies to Philadelphia crime due to additive seasonality with segments of downward trend seen in the time series plots.

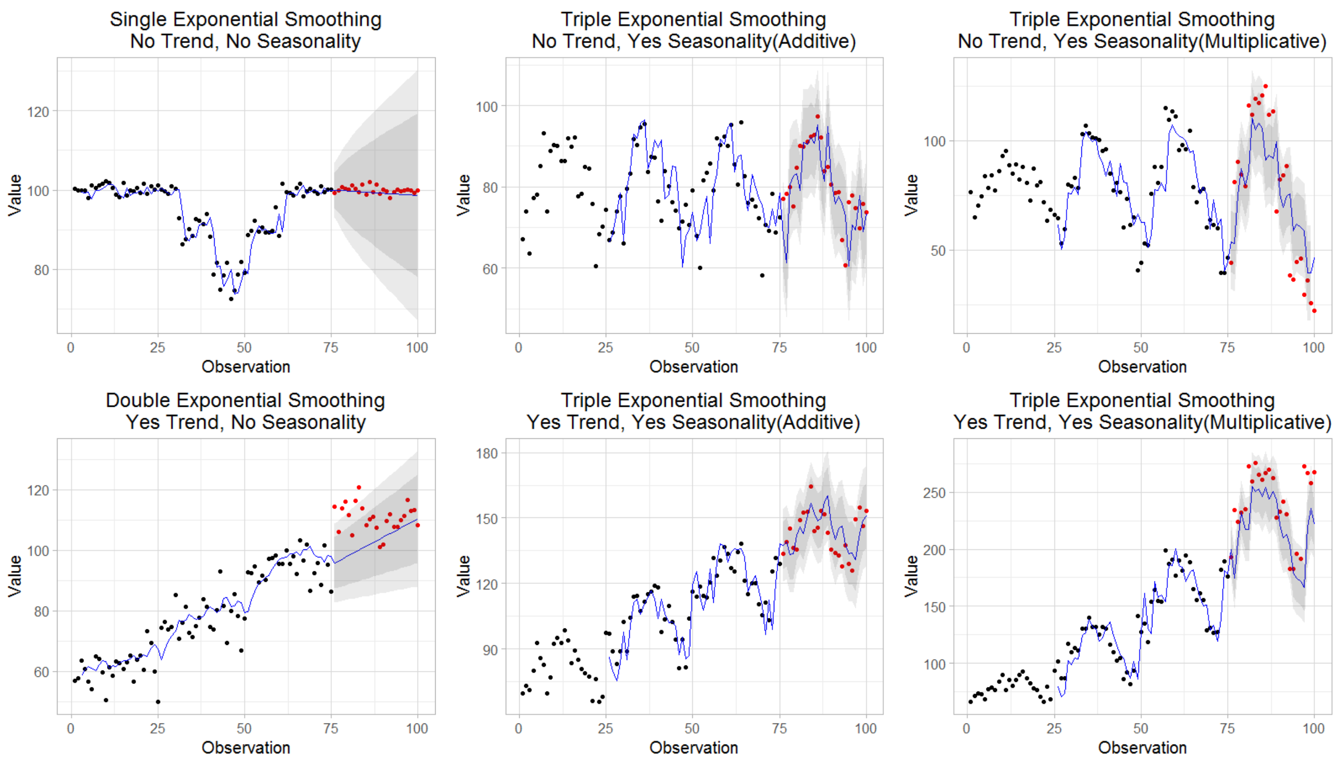


Figure 26: Exponential Smoothing Decision Matrix

The following 9 steps is the experimentation process for fitting and forecasting with exponential smoothing models. The forecasting will use the data from 2006 – 2014 as the training data to predict the data from 2015 – 2016.

1. Fit the automated model to all of the data and retrieve the estimated parameters: alpha, beta, gamma.
2. Fit the automated model to the training data and retrieve the estimated parameters.
3. Create a general factorial DOE based off the estimated parameters from Step 1 or Step 2.
    - Dependent Variables & their levels.
        - **alpha**: from = estimate – 0.25, to = estimate + 0.25, incremented by = 0.01
            - Plus the 2 alpha parameters from the models in Step 1 and 2.
        - **beta**: from = estimate – 0.25, to = estimate + 0.25, incremented by = 0.01
            - Plus the 2 beta parameters from the models in Step 1 and 2.
        - **gamma**: from = estimate – 0.25, to = estimate + 0.25, incremented by = 0.01
            - Plus the 2 gamma parameters from the models in Step 1 and 2.
        - A +- 0.25 range and an increment of 0.01 was chosen through trial and error to keep the computation time of all models around 5 minutes.
    - Build a model for every scenario of alpha, beta, and gamma.
        - These models use the training data only.
    - Independent Variables & their ideal values: These are computed on residuals of the test data.
        - **One-Sample t-test P-Value**: Values above 0.05.
            - H_0: Mean = 0
            - H_1: Mean $\neq$ 0
        - **Mean Error**: Values close to zero.
        - **Root Mean Square Error**: Values close to zero.
        - **Mean Absolute Error**: Values close to zero.
        - **Mean Percent Error**: Values close to zero.
        - **Mean Absolute Percent Error**: Values close to zero.
4. Plot the histograms of the Independent Variables to decide how to filter models.
5. Choose one model from the DOE.
6. Compare the values of the independent variables for the 3 final models.
    - The model from Step 1
    - The model from Step 2
    - The model from the DOE
7. Plot the fitted values on the actual values for the 3 final models.
8. Plot the residuals for the 3 final models.
    - Residual v. Fitted, Normal QQ Plot, Residual v. Time, Variogram, Histogram, ACF Plot
9. Choose the best model based on numeric and graphic comparison.

The DOE results for the final 2 fitted models, and then a plot of the best fitted model will be shown for Central, West, and North Philadelphia weekly crimes. The DOE results for the final 3 forecasting models, and then various plots of the best forecasting model will be shown for Central, West, and North Philadelphia weekly crimes.

Central Philadelphia

The numerical performance of the 2 Central Philadelphia fitted models is shown below in Table 8. Scenario 33517 showed much better performance in t.pval and ME for its values of alpha, beta, and gamma. The values for alpha, beta, and gamma between the 2 models are not that different but yield a differing results. Scenario 33517 is the model that was eventually chosen as the best exponential smoothing model for fitting Central Philadelphia weekly crimes for 2006 – 2016.

Table 8: DOE Results for Fitting Central Philadelphia Weekly Crimes 2006 - 2016

Scenario	alpha	beta	gamma	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0.13895	0.00146	0.5037	0.2684	4.32374	88.2359	60.5956	-4.013	13.8481
33517	0.15	0.01	0.58	0.90127	0.48752	88.7029	61.075	-4.5767	13.9393

The 1-step ahead performance for the best fitted Central Philadelphia Holt-Winters model is shown below. The model captures the center of the data and the peaks very well, but misses the lowest values often. The corresponding values of alpha, beta, and gamma show that the model relies on the most recent observations slightly, doesn't rely on the most recent trend, and relies on the most recent seasonality heavily. The Residual Time Series, ACF, and PACF plots of this model will be addressed in the Moving Average & Autoregressive Processes section.

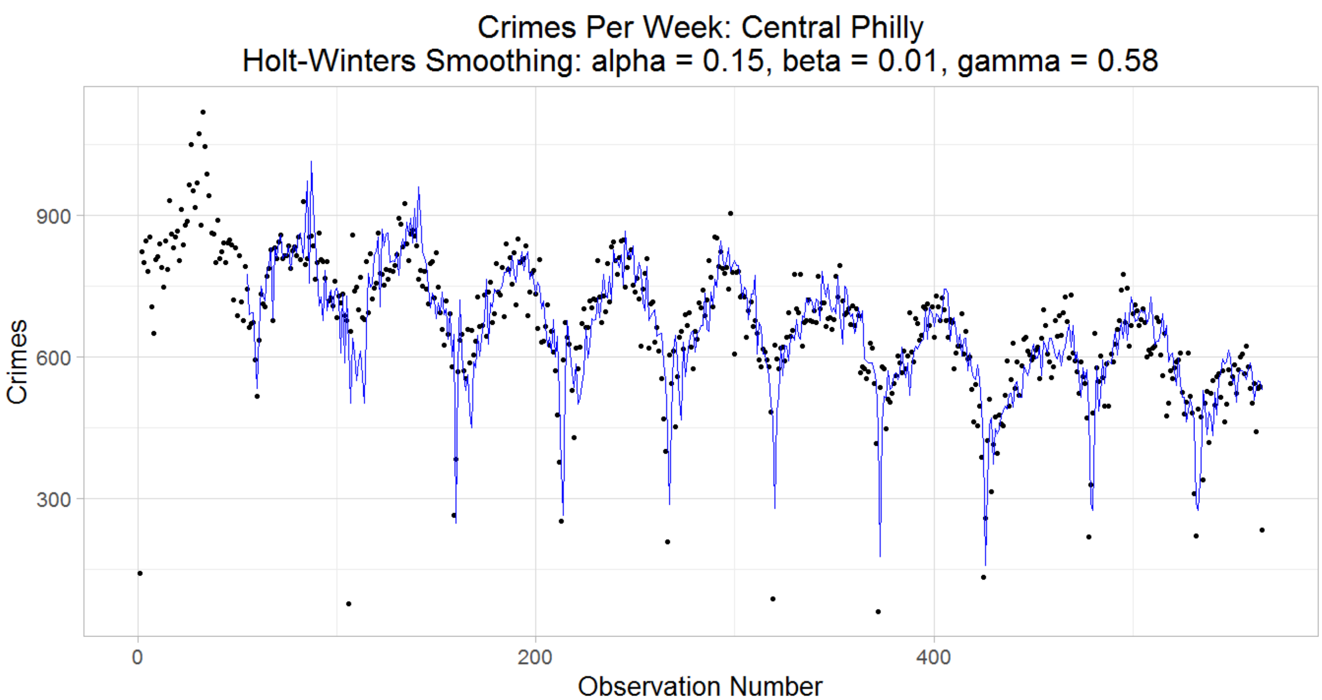


Figure 27: 10 Years of Central Philadelphia Weekly Crime & Holt-Winters Model Fit

The numerical performance of the 3 Central Philadelphia forecasting models is shown below in Table 9. Scenario 10352 had the best performance across all independent variables for its values of alpha, beta, and gamma. The values for alpha, beta, and gamma between the 3 models are different and yield distinctly different results. Scenario 10352 is the model that was eventually chosen as the best exponential smoothing model for forecasting Central Philadelphia weekly crimes for 2015 – 2016.

Table 9: DOE Results for Forecasting Central Philadelphia Weekly Crimes 2015 - 2016

Scenario	alpha	beta	gamma	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0.09852	0.00311	0.50784	1.07E-06	50.2407	101.946	78.819	6.95389	14.8738
2	0.13895	0.00146	0.5037	2.52E-14	80.4748	114.627	94.7632	12.4874	17.1357
10352	0.11	0.08	0.4	9.16E-01	0.85826	74.9167	55.7464	-2.1416	11.5526

The forecast performance for the best Central Philadelphia Holt-Winters model is shown below. The model follows the actual values well and the prediction intervals at 80% and 95% confidence levels, capture the actual values very well. The corresponding values of alpha, beta, and gamma show that the model relies on the most recent observations slightly, relies on the most recent trend slightly, and relies on the most recent seasonality heavily.

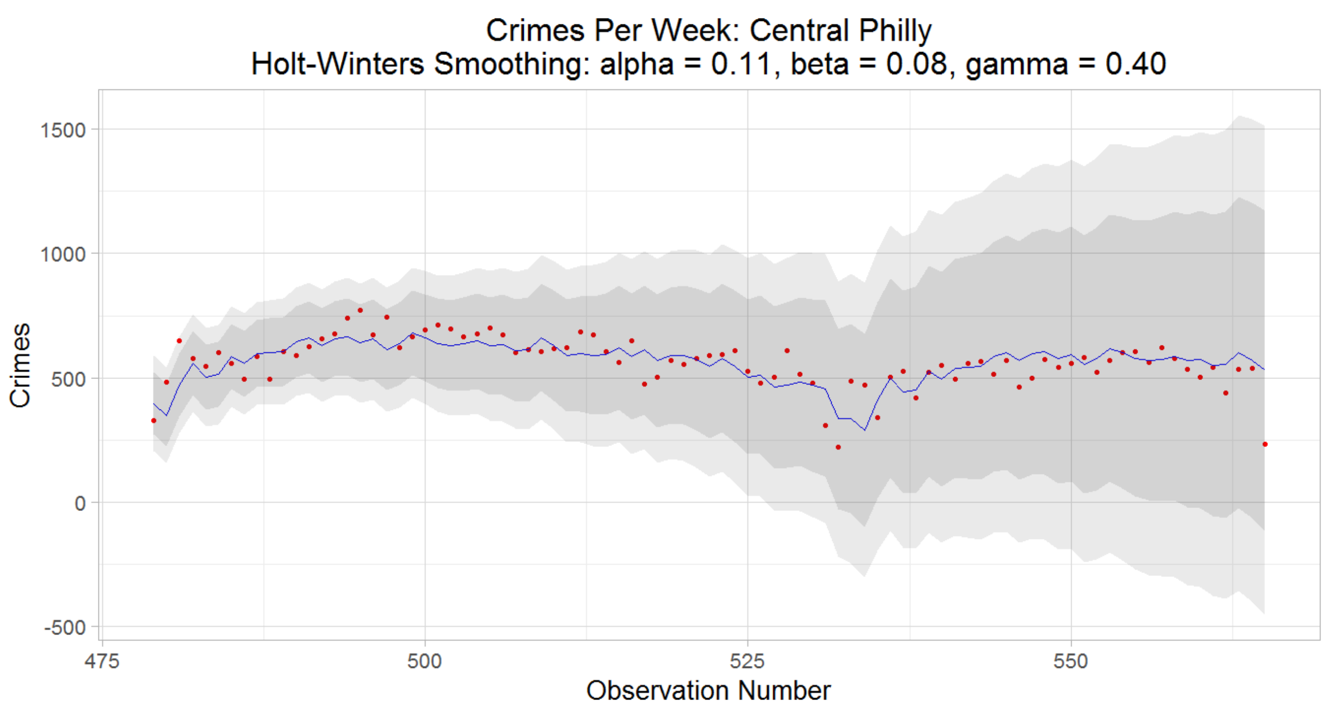


Figure 28: Prediction Plot for Central Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residual time series of the Central Philadelphia Holt-Winters forecasting model is the result of trying to remove trend and seasonality from weekly crimes, as shown below in Figure 29. This plot shows that annual seasonality was removed successfully. Let's verify the removal of trend and seasonality with the ACF plot of these residuals.

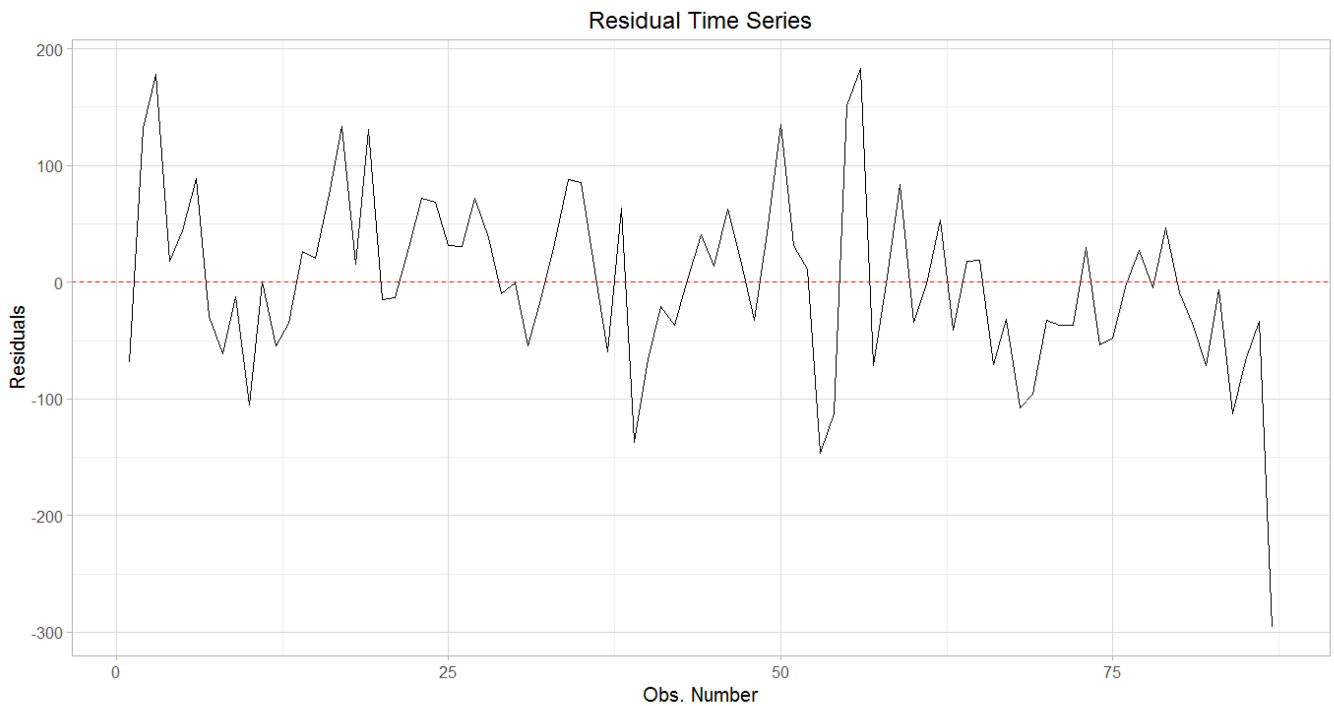


Figure 29: Residual v. Time for Central Philadelphia Weekly Crime Holt-Winters Forecasting Model

The autocorrelation levels of the residuals for the forecasting model look good. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below show that there is a first order moving average process present due to the lag of 1 violating the limits.

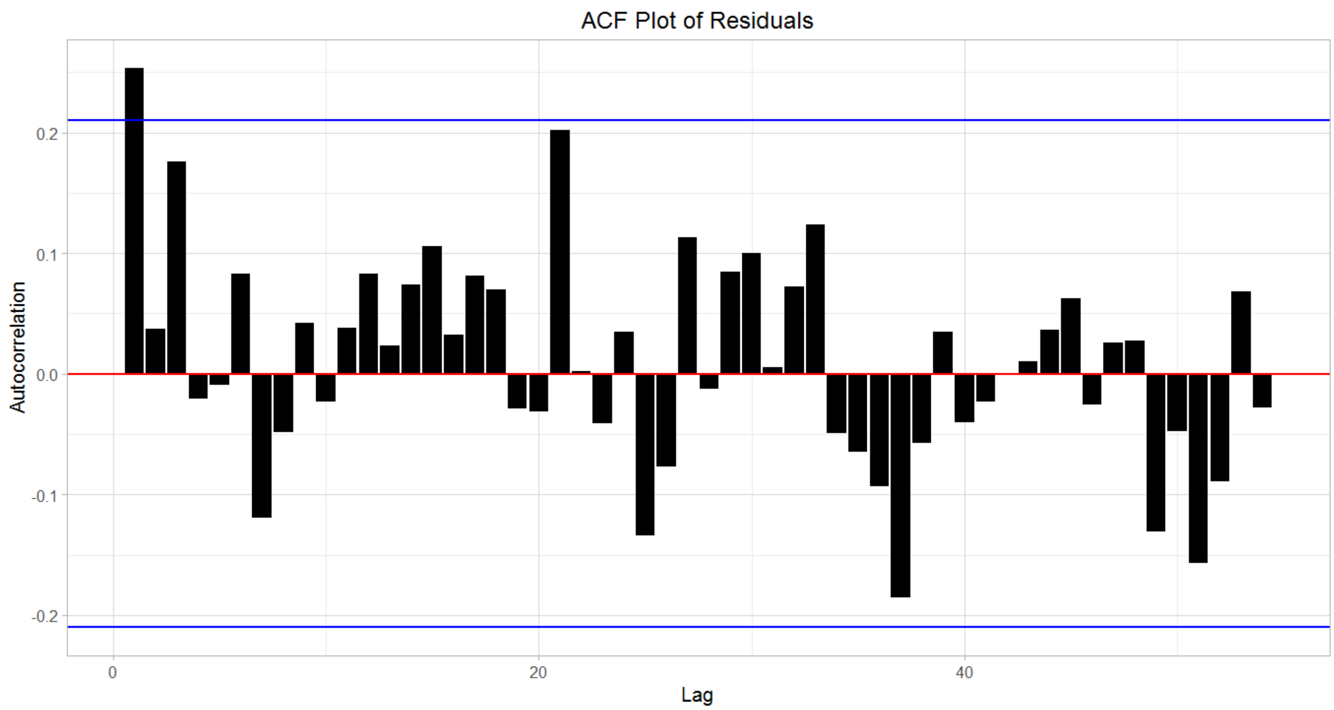


Figure 30: ACF Plot of Residuals for Central Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residuals v. fitted plot for the forecasting model show that a constant mean was achieved. Variance appears constant too. This plot also shows that the residuals are centered at zero.

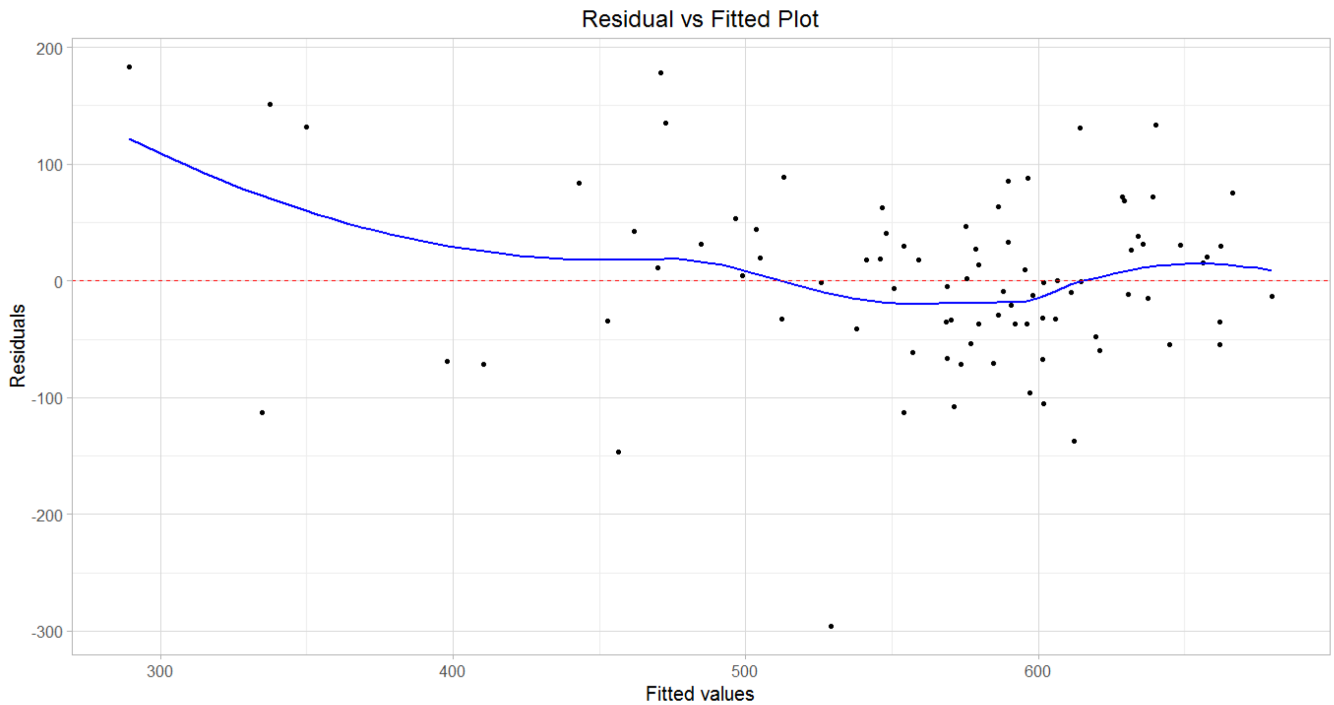


Figure 31: Residual v. Fitted for Central Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residuals for the forecasting model may follow a normal distribution, with most points following the line, and all points except for the upper tail and one lower tail point staying within the 95% confidence interval ribbon.

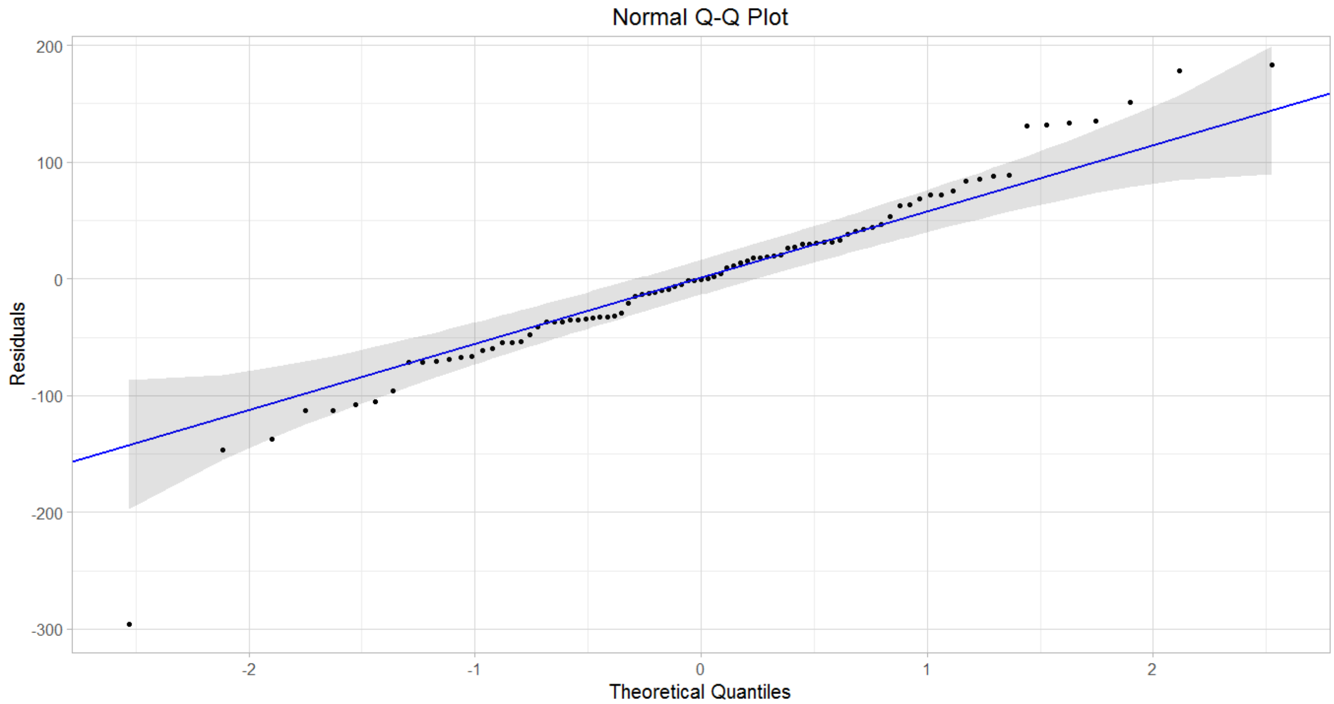


Figure 32: Residual Normal Q-Q Plot for Central Philadelphia Weekly Crime Holt-Winters Forecasting Model

West Philadelphia

The numerical performance of the 2 West Philadelphia fitted models is shown below in Table 10. Scenario 22157 had better performance in t.pval for its values of alpha, beta, and gamma. The values for alpha, beta, and gamma between the 2 models are different but yield very similar results except in t.pval. Scenario 22157 is the model that was eventually chosen as the best exponential smoothing model for fitting West Philadelphia weekly crimes for 2006 – 2016.

Table 10: DOE Results for Fitting West Philadelphia Weekly Crimes 2006 - 2016

Scenario	alpha	beta	gamma	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0.12106	0	0.52593	0.620888	-1.9552	89.2407	62.3286	-5.6628	14.4877
22157	0.3	0	0.51	0.851735	-0.7449	89.967	62.0272	-5.1612	14.3997

The 1-step ahead performance for the best fitted West Philadelphia Holt-Winters model is shown below. The model captures the center of the data very well and most of the peaks, but misses the lowest values often. The corresponding values of alpha, beta, and gamma show that the model relies on the most recent observations heavily, doesn't rely on the most recent trend, and relies on the most recent seasonality heavily. The Residual Time Series, ACF, and PACF plots of this model will be addressed in the Moving Average & Autoregressive Processes section.

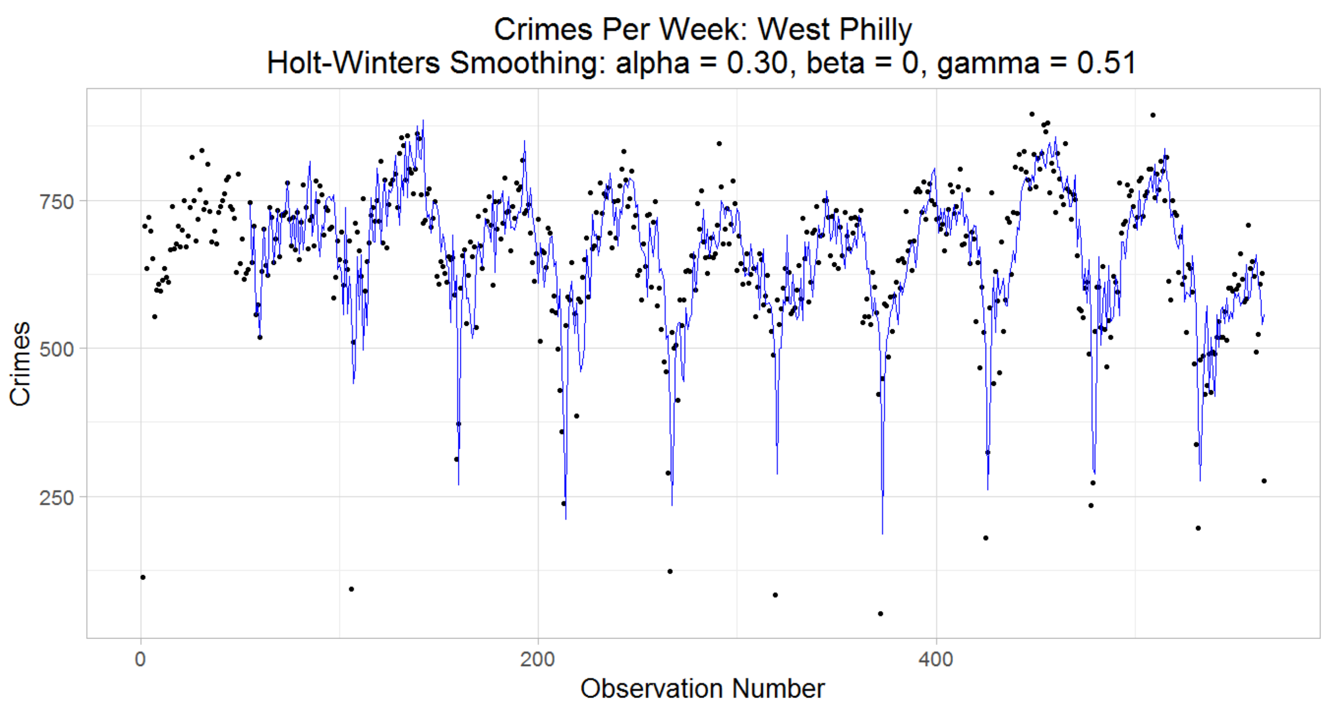


Figure 33: 10 Years of West Philadelphia Weekly Crime & Holt-Winters Model Fit

The numerical performance of the 3 West Philadelphia forecasting models is shown below in Table 11. Scenario 8827 had the best performance across all independent variables for its values of alpha, beta, and gamma. The values for alpha, beta, and gamma between the 3 models are different and yield different results. Scenario 8827 is the model that was eventually chosen as the best exponential smoothing model for forecasting West Philadelphia weekly crimes for 2015 – 2016.

Table 11: DOE Results for Forecasting West Philadelphia Weekly Crimes 2015 - 2016

Scenario	alpha	beta	gamma	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0.26415	0	0.48657	3.24E-12	96.7748	147.055	120.669	12.1713	19.3247
2	0.12106	0	0.52593	3.39E-01	-10.562	102.416	81.611	-4.8619	15.4318
8827	0.02	0.17	0.3	9.63E-01	-0.4019	80.5195	58.8806	-2.8648	11.8654

The forecast performance for the best West Philadelphia Holt-Winters model is shown below. The model captures 2015 fairly well and captures 2016 better. The prediction intervals at 80% and 95% confidence levels, capture all the actual values except the first. The corresponding values of alpha, beta, and gamma show that the model relies on the most recent observations very little, relies on the most recent trend, and relies on the most recent seasonality heavily.

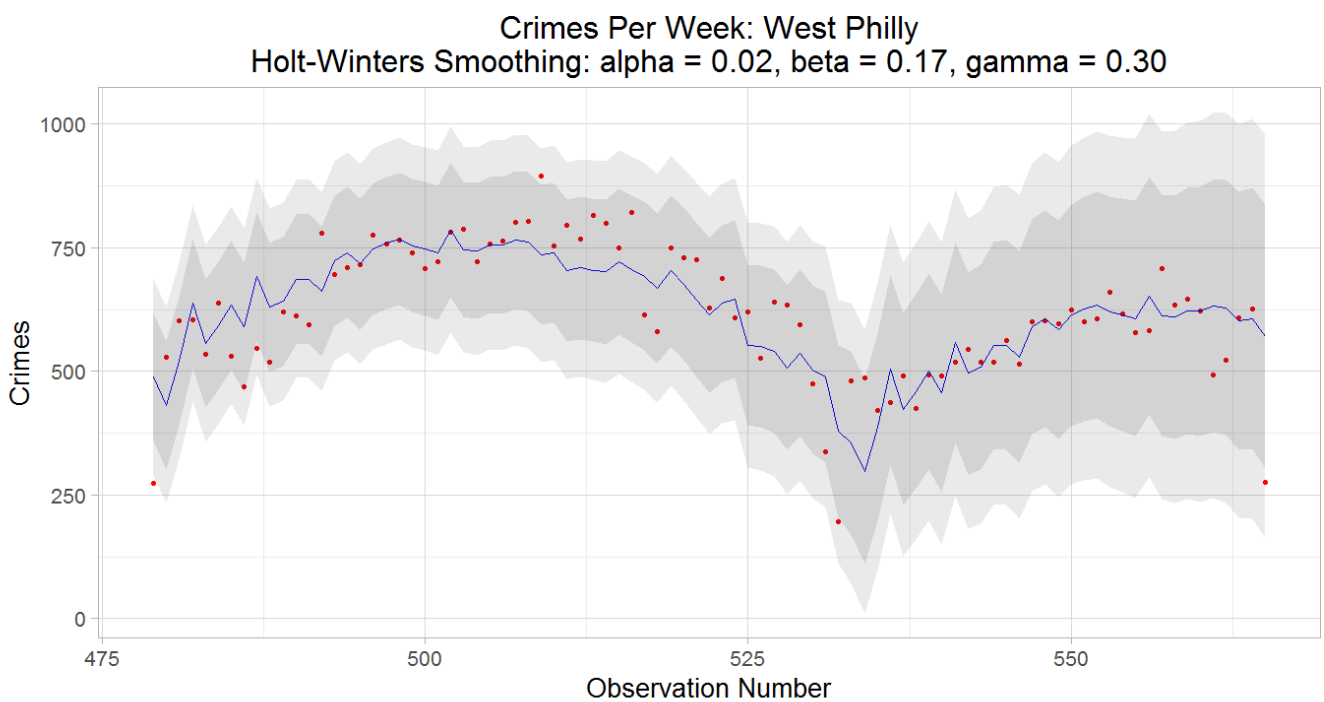


Figure 34: Prediction Plot for West Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residual time series of the West Philadelphia Holt-Winters forecasting model is the result of trying to remove trend and seasonality from weekly crimes, as shown below in Figure 35. This plot shows that annual seasonality has been removed. Let's verify the removal of trend and seasonality with the ACF plot of these residuals.

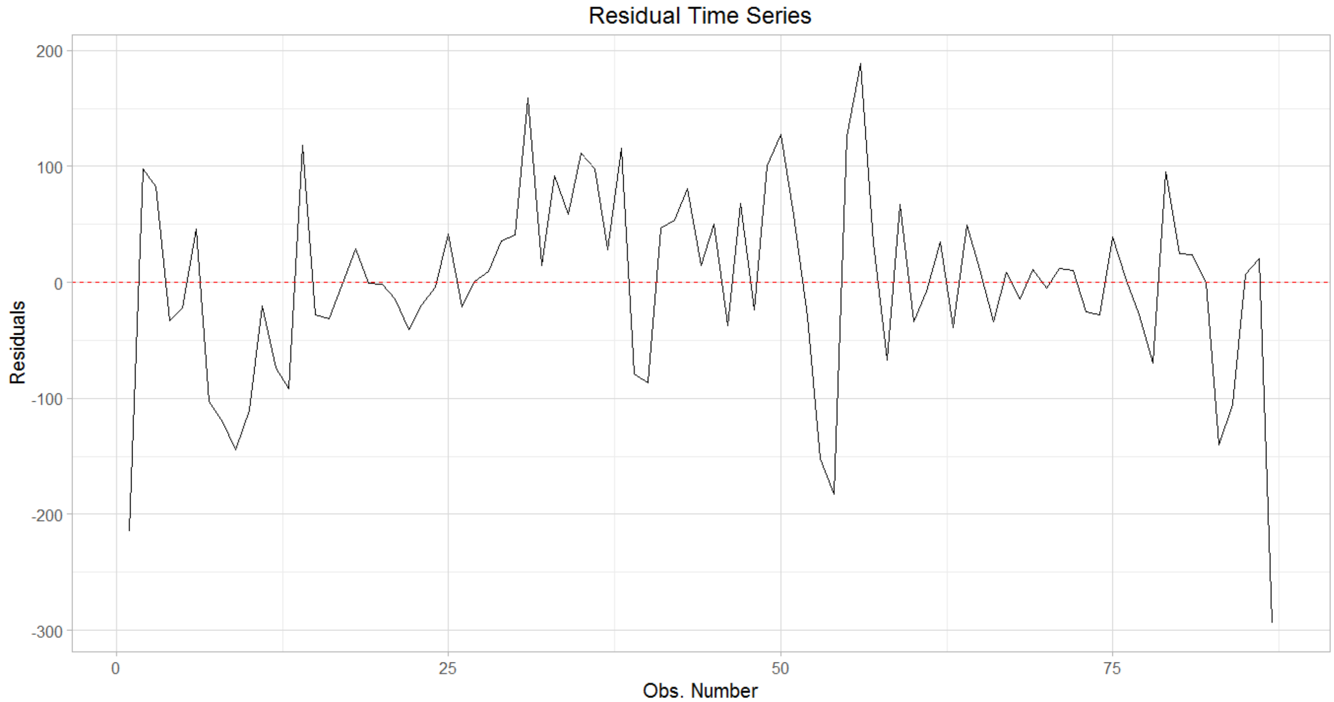


Figure 35: Residual v. Time for West Philadelphia Weekly Crime Holt-Winters Forecasting Model

The autocorrelation levels of the residuals for the forecasting model are not significant. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below show that autocorrelation has been completely removed.

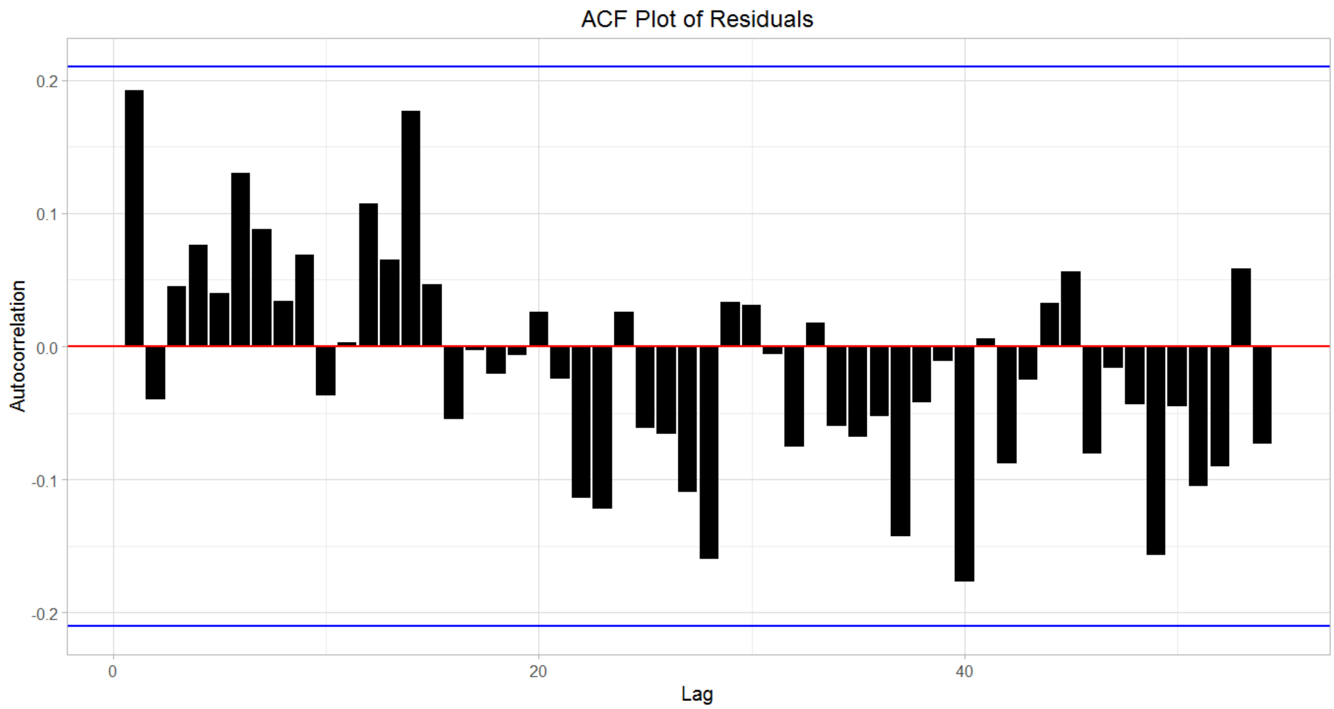


Figure 36: ACF Plot of Residuals for West Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residuals v. fitted plot for the forecasting model show that a constant mean was achieved. Variance appears constant too. This plot also shows that the residuals are centered at zero.

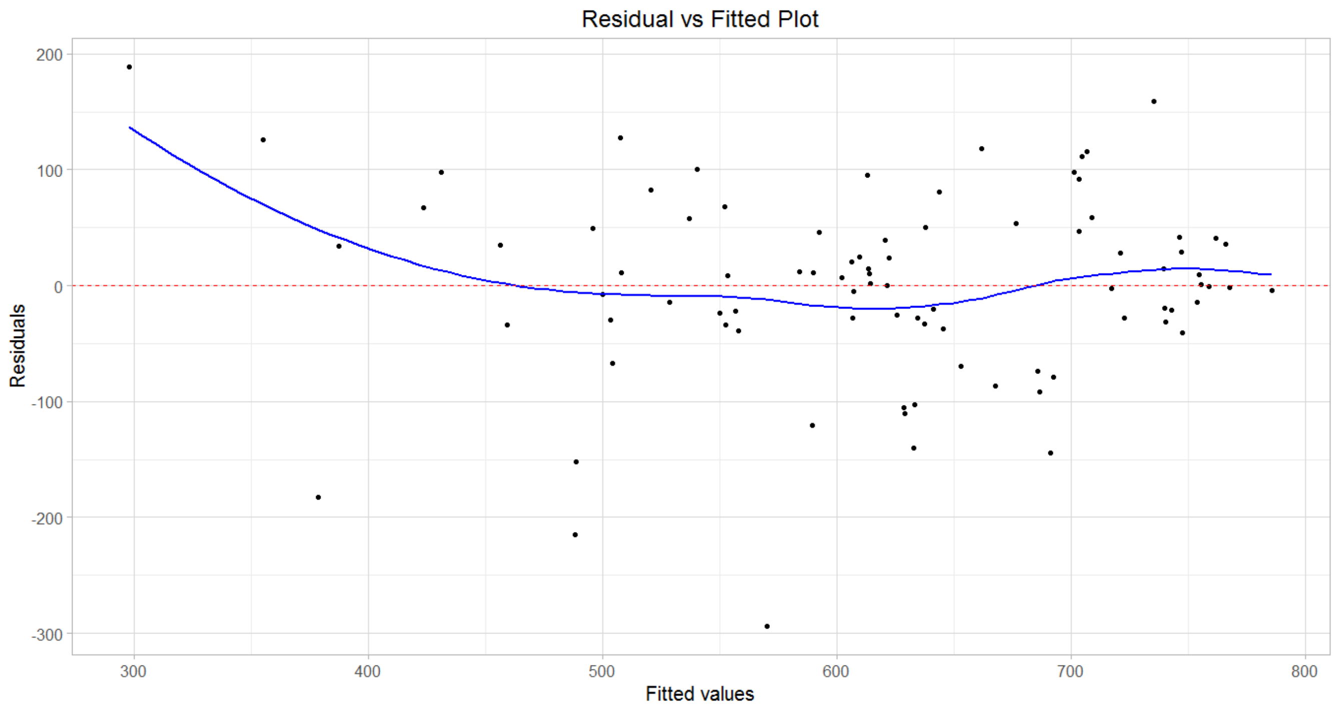


Figure 37: Residual v. Fitted for West Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residuals for the forecasting model do not appear to follow a normal distribution due to the left side trailing off the line and 95% confidence interval band too much.

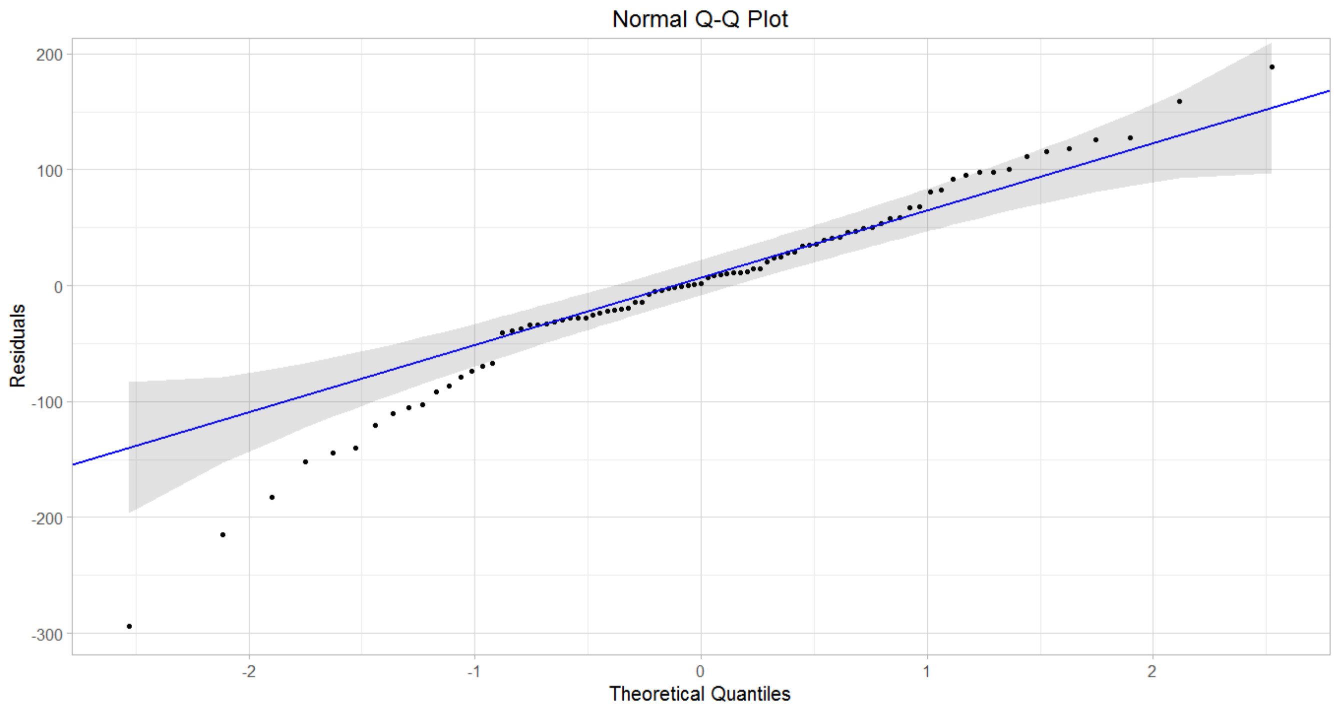


Figure 38: Residual Normal Q-Q Plot for West Philadelphia Weekly Crime Holt-Winters Forecasting Model

North Philadelphia

The numerical performance of the 2 North Philadelphia fitted models is shown below in Table 12. Both scenarios have nearly identical performance across all independent variables. This is due to the values for alpha, beta, and gamma between the 2 models are very similar. Scenario 25701, with a lower MAPE, is the model that was eventually chosen as the best exponential smoothing model for fitting North Philadelphia weekly crimes for 2006 – 2016.

Table 12: DOE Results for Fitting North Philadelphia Weekly Crimes 2006 - 2016

Scenario	alpha	beta	gamma	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0.13088	0	0.58501	0.797599	-2.4812	218.388	137.418	-4.8015	12.5817
25701	0.12	0	0.6	0.793611	-2.5317	218.431	137.381	-4.7704	12.5506

The 1-step ahead performance for the best fitted North Philadelphia Holt-Winters model is shown below. The corresponding values of alpha, beta, and gamma show that the model relies on the most recent observations slightly, doesn't rely on the most recent trend, and relies on the most recent seasonality heavily. The Residual Time Series, ACF, and PACF plots of this model will be addressed in the Moving Average & Autoregressive Processes section.

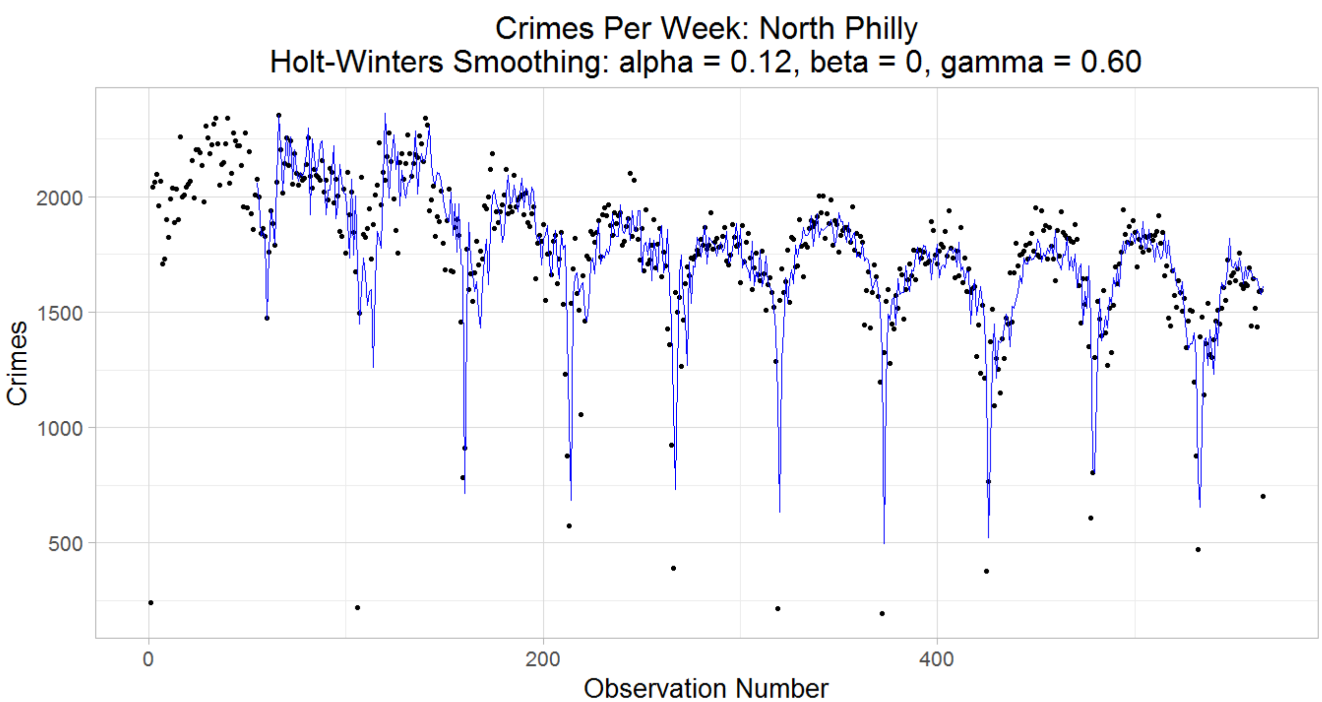


Figure 39: 10 Years of North Philadelphia Weekly Crime & Holt-Winters Model Fit

The numerical performance of the 3 North Philadelphia forecasting models is shown below in Table 13. Scenario 7257 had the best performance across all independent variables for its values of alpha, beta, and gamma. The values for alpha, beta, and gamma between the 3 models are different and yield different results. Scenario 7257 is the model that was eventually chosen as the best exponential smoothing model for forecasting North Philadelphia weekly crimes for 2015 – 2016.

Table 13: DOE Results for Forecasting North Philadelphia Weekly Crimes 2015 - 2016

Scenario	alpha	beta	gamma	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0.0894	0	0.58688	8.01E-01	5.66763	207.784	131.234	-1.6334	10.3079
2	0.13088	0	0.58501	2.29E-01	26.0071	200.676	128.825	-0.3499	10.0249
7257	0.13	0.05	0.42	8.86E-01	-2.8834	185.661	121.258	-2.4663	9.86163

The forecast performance for the best North Philadelphia Holt-Winters model is shown below. The model captures the actual values very well. The prediction intervals at 80% and 95% confidence levels, capture the actual values very well. The corresponding values of alpha, beta, and gamma show that the model relies on the most recent observations slightly, relies on the most recent trend very little, and relies on the most recent seasonality heavily.

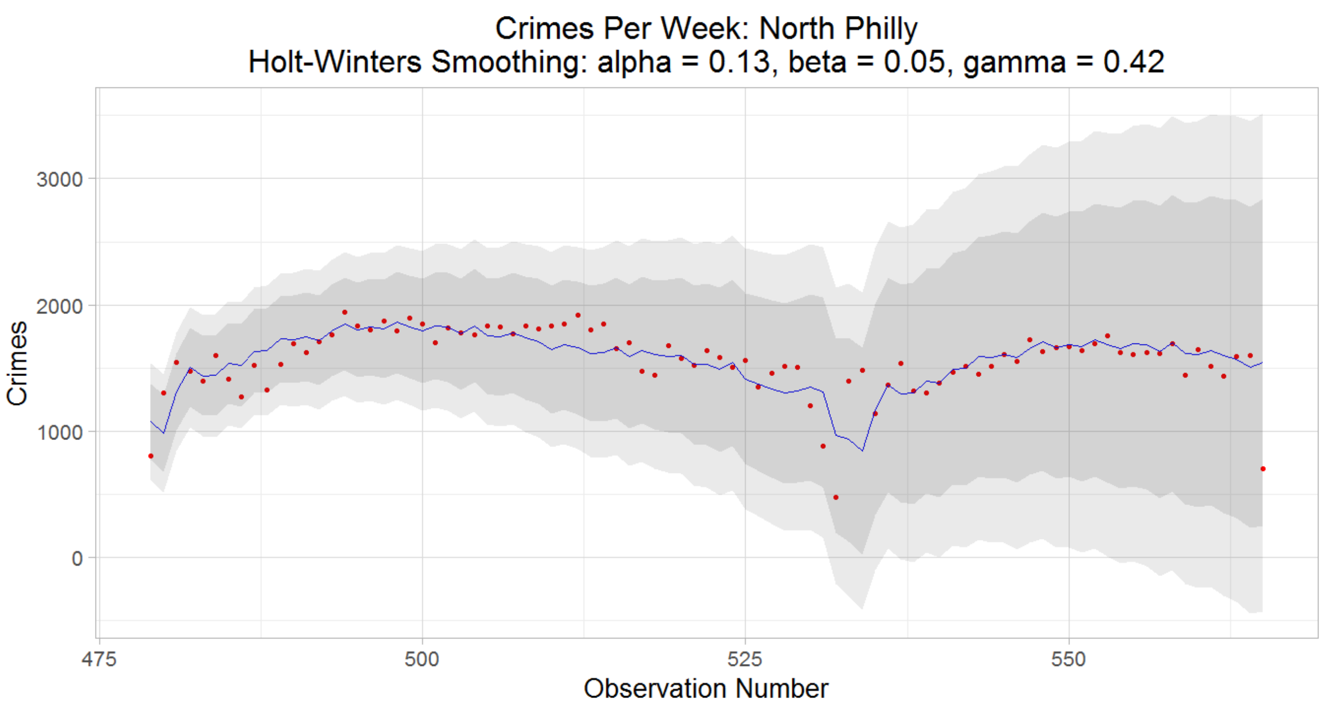


Figure 40: Prediction Plot for North Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residual time series of the North Philadelphia Holt-Winters forecasting model is the result of trying to remove trend and seasonality from weekly crimes, as shown below in Figure 41. This plot shows that annual seasonality has been removed. Let's verify the removal of trend and seasonality with the ACF plot of these residuals.

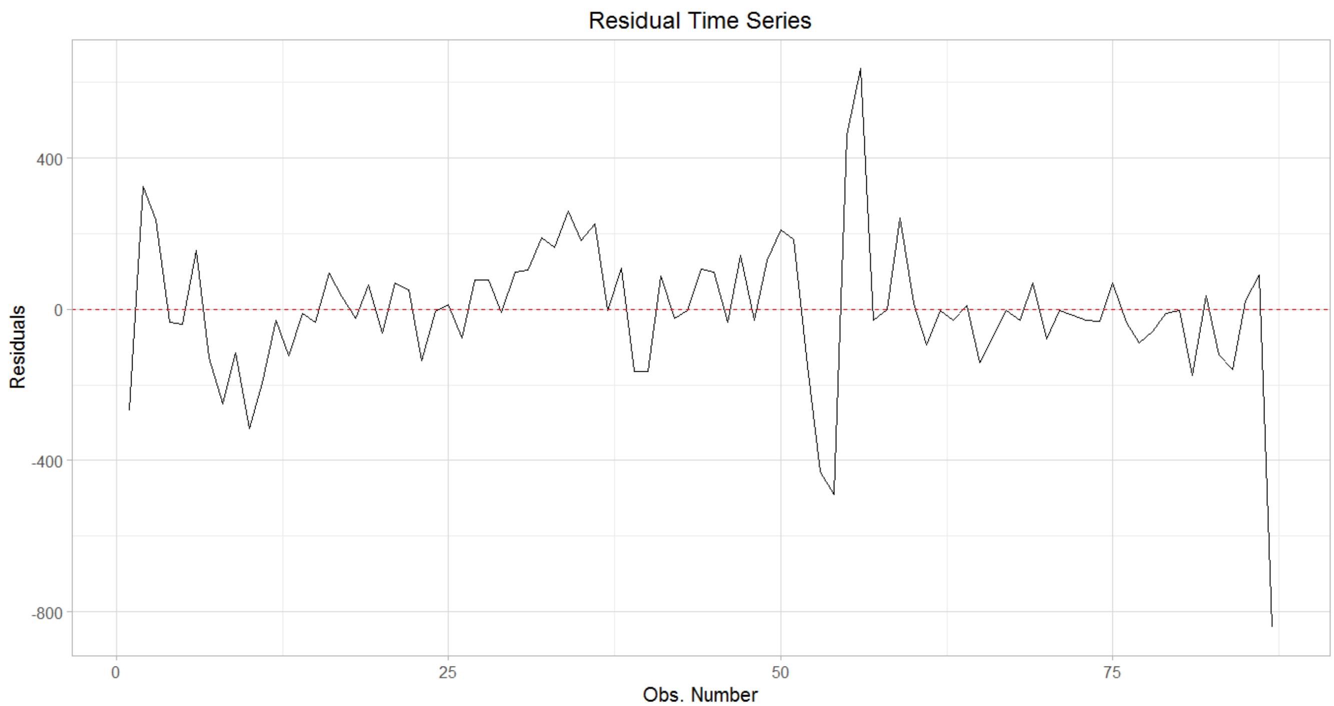


Figure 41: Residual v. Time for North Philadelphia Weekly Crime Holt-Winters Forecasting Model

The autocorrelation levels of the residuals for the forecasting model are not significant. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below show that autocorrelation has been completely removed.

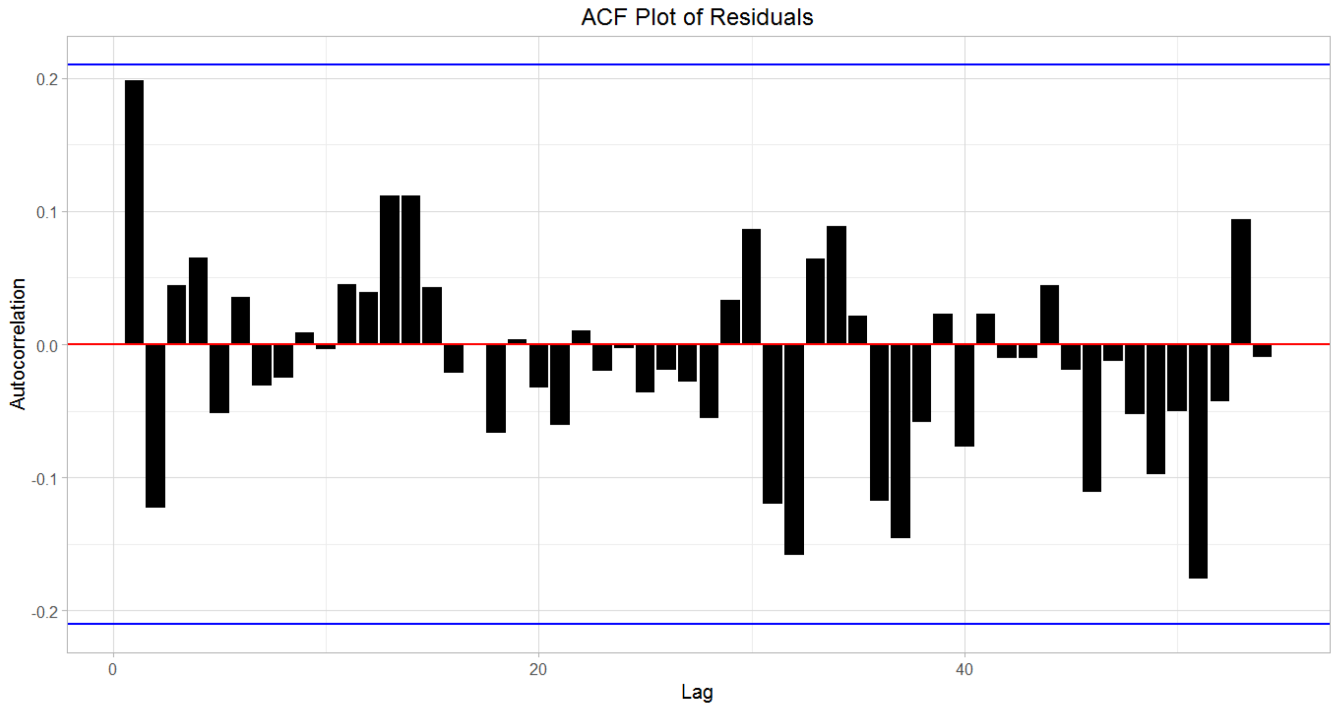


Figure 42: ACF Plot of Residuals for North Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residuals v. fitted plot for the forecasting model show that a constant mean was achieved. Variance appears constant too. This plot also shows that the residuals are centered at zero.

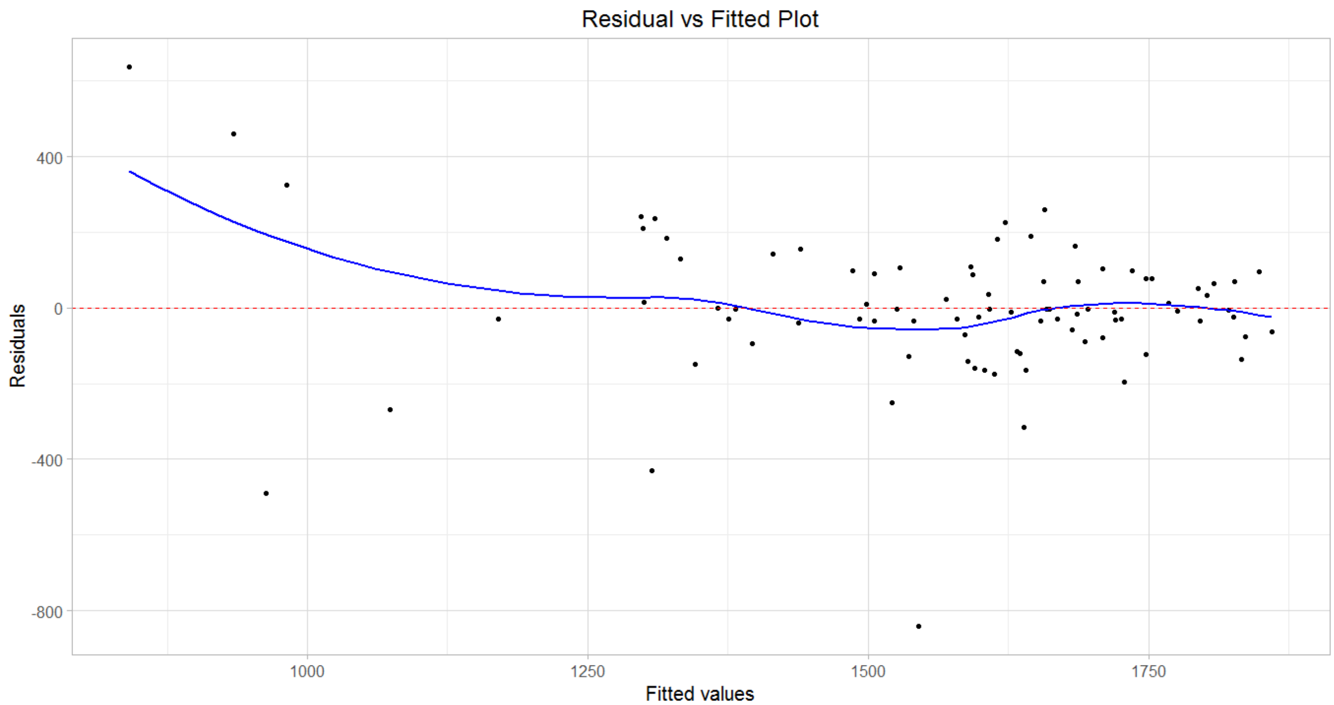


Figure 43: Residual v. Fitted for North Philadelphia Weekly Crime Holt-Winters Forecasting Model

The residuals for the forecasting model may follow a normal distribution. The tails are trailing off the line and 95% confidence interval band a little too much.

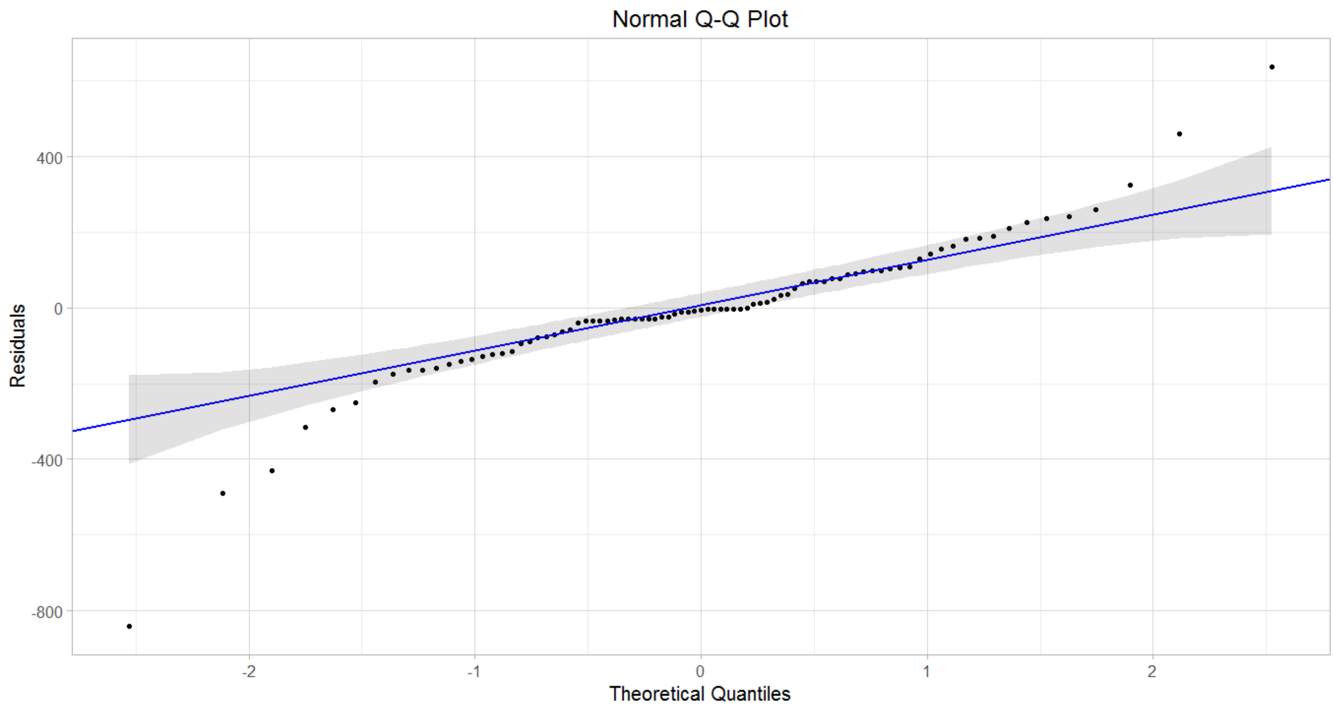


Figure 44: Residual Normal Q-Q Plot for North Philadelphia Weekly Crime Holt-Winters Forecasting Model

Moving Average & Autoregressive Processes

The residual time series of the fitted Central Philadelphia Holt-Winters smoothing model is the result of trying to remove trend and seasonality from weekly crimes, as shown below in Figure 45. This plot shows that annual seasonality may have been removed. Let's look at the ACF and PACF plots.

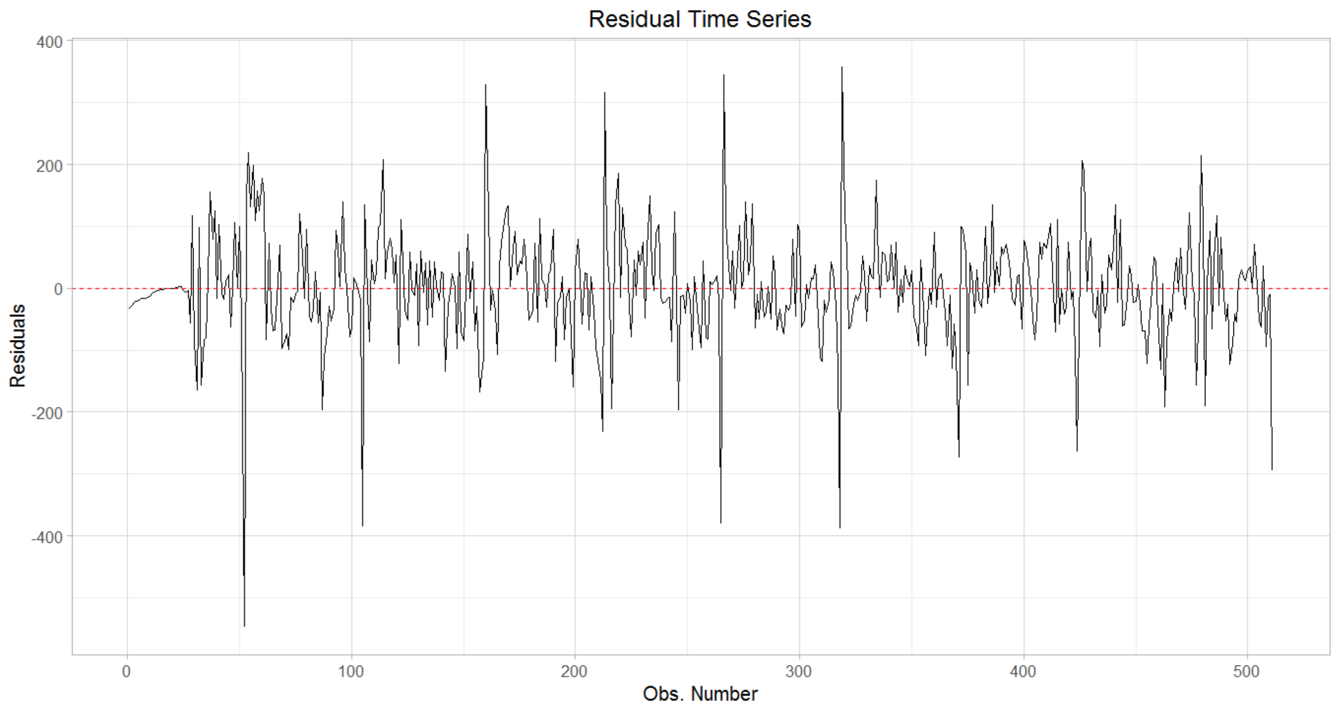


Figure 45: Residual v. Time for Central Philadelphia Weekly Crime Holt-Winters Smoothing Model

The bars in these charts are not perfect, especially at the end. This may be due to the fact that each year has 52 – 54 weeks technically. Overall the ACF and PACF plot appear to drop into the blue interval after a lag of 1. Therefore an ARMA(1, 1) process may be worth trying on these residuals. It seems that differencing might be worth exploring too, to create an ARIMA process for Central Philadelphia weekly crimes.

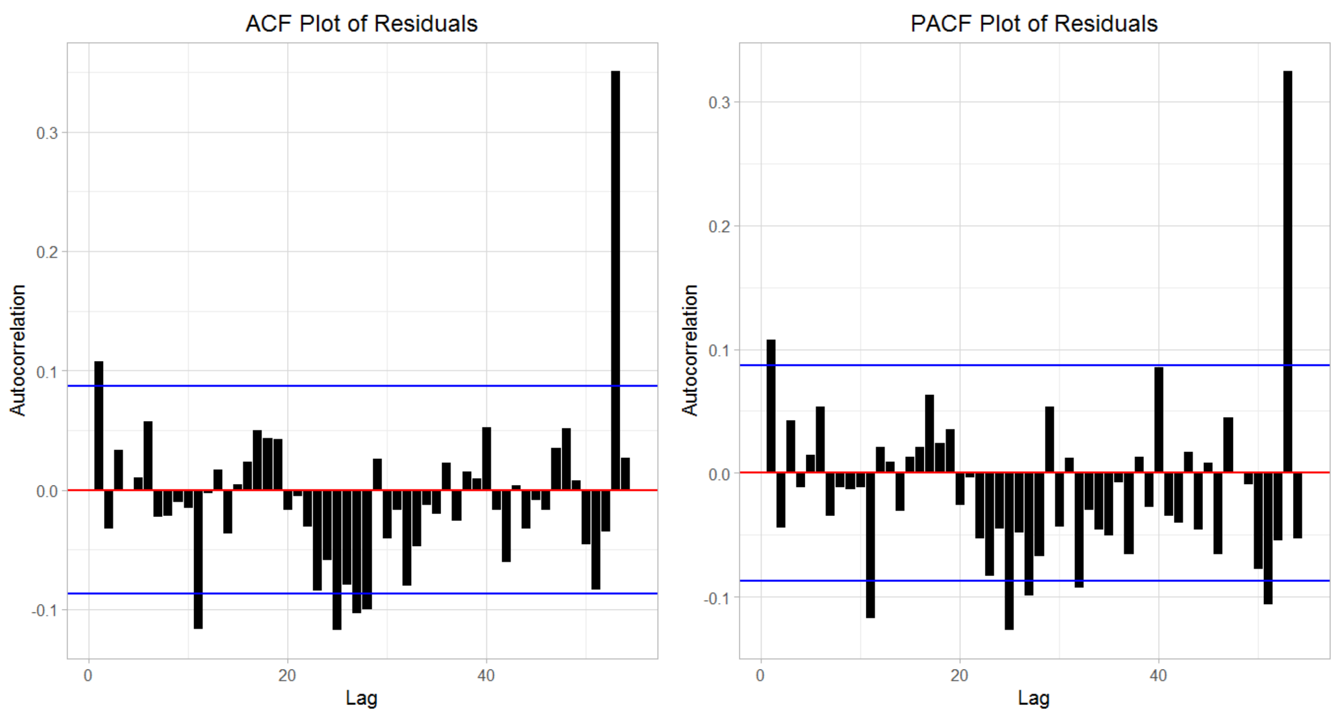


Figure 46: ACF & PACF Plots of Residuals for Central Philadelphia Weekly Crime Holt-Winters Smoothing Model

The residual time series of the West Philadelphia Holt-Winters smoothing model is the result of trying to remove trend and seasonality from weekly crimes, as shown below in Figure 47. This plot shows that annual seasonality may have been removed. Let's look at the ACF and PACF plots.

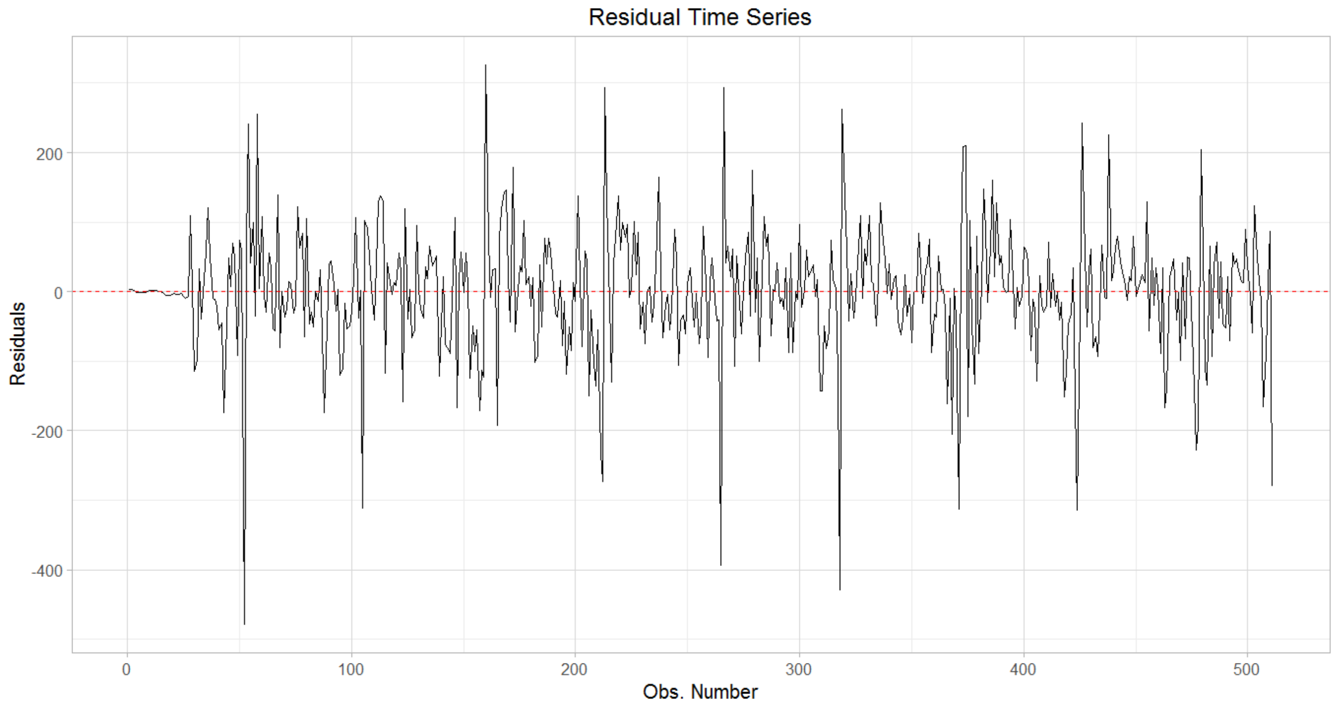


Figure 47: Residual v. Time for West Philadelphia Weekly Crime Holt-Winters Smoothing Model

The bars in these charts look good except at the end. This may be due to the fact that each year has 52 – 54 weeks technically. Overall the ACF and PACF plot appear to drop into the blue interval immediately. Therefore an ARMA process may not be worth trying. Differencing might be worth exploring too, to create an ARIMA process for West Philadelphia weekly crimes.

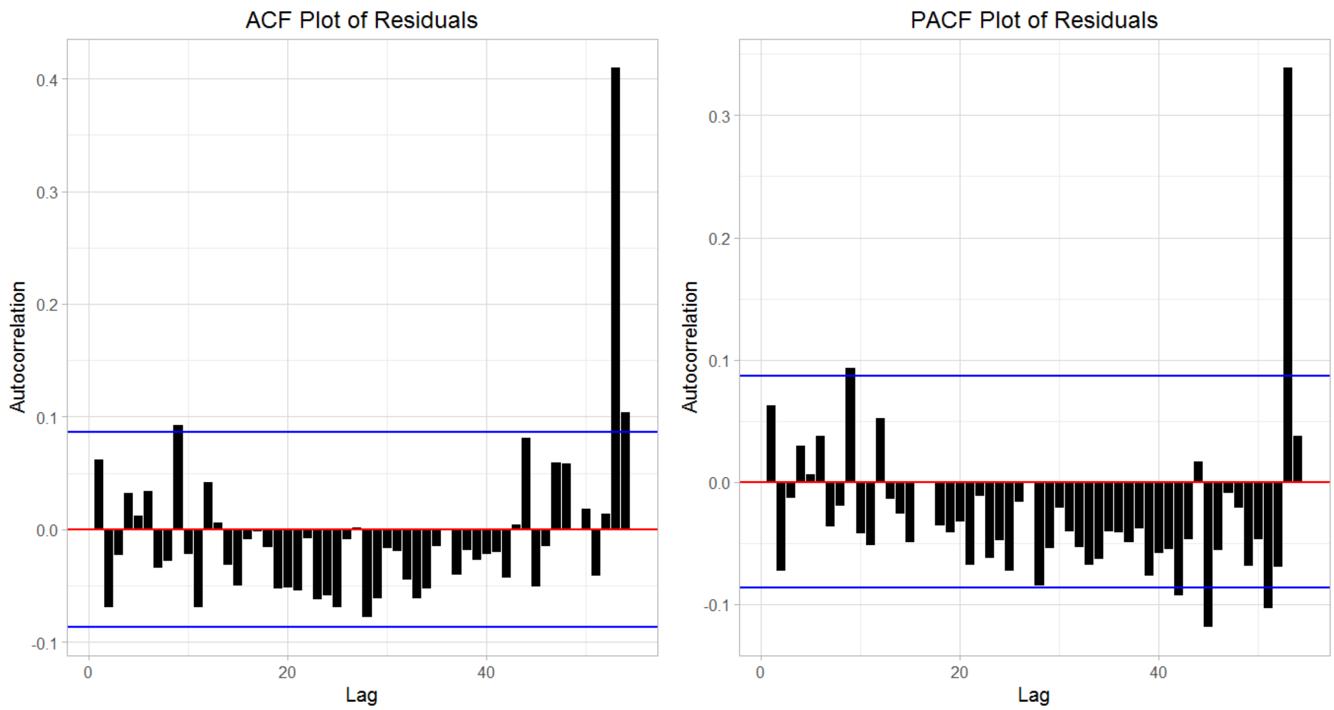


Figure 48: ACF & PACF Plots of Residuals for West Philadelphia Weekly Crime Holt-Winters Smoothing Model

The residual time series of the North Philadelphia Holt-Winters smoothing model is the result of trying to remove trend and seasonality from weekly crimes, as shown below in Figure 49. This plot shows that annual seasonality may have been removed. Let's look at the ACF and PACF plots.

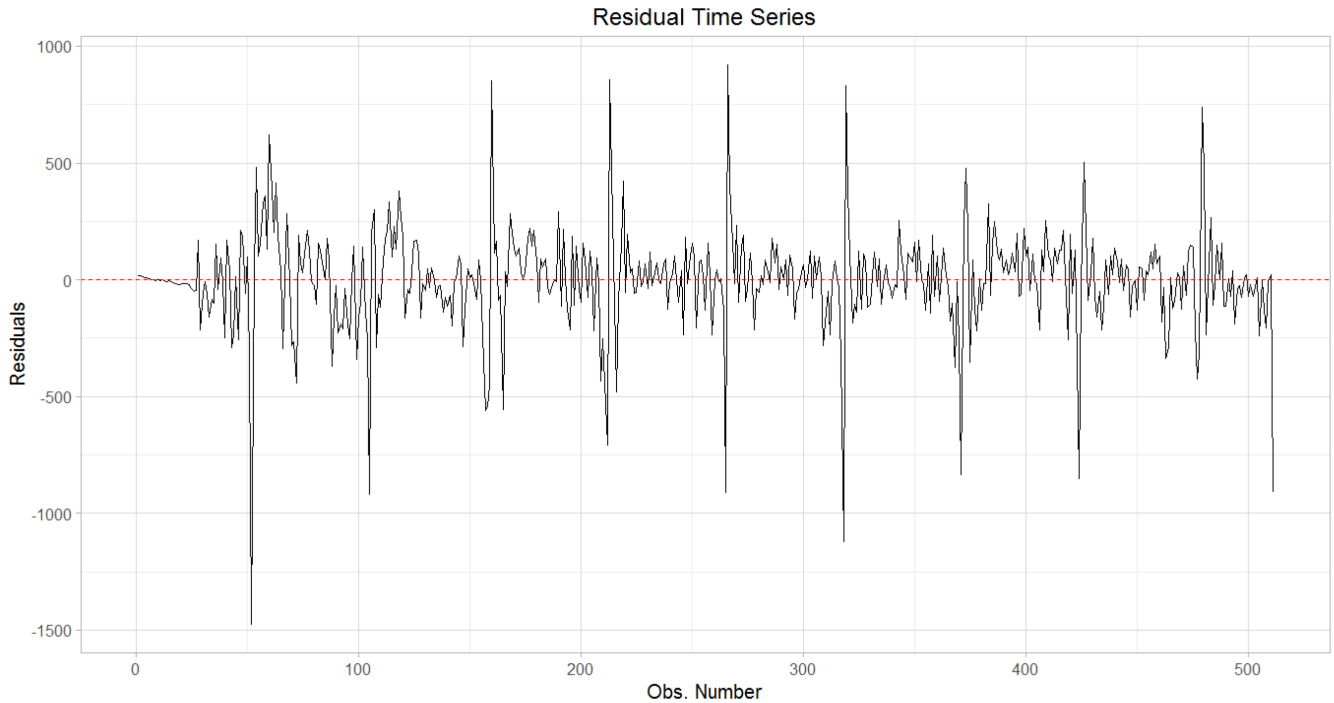


Figure 49: Residual v. Time for North Philadelphia Weekly Crime Holt-Winters Smoothing Model

The bars in these charts are not perfect, especially at the end. This may be due to the fact that each year has 52 – 54 weeks technically. Overall the ACF and PACF plot appear to drop into the blue interval after a lag of 2. Therefore an ARMA(2, 2) process may be worth trying on these residuals. It seems that differencing might be worth exploring too, to create an ARIMA process for North Philadelphia weekly crimes.

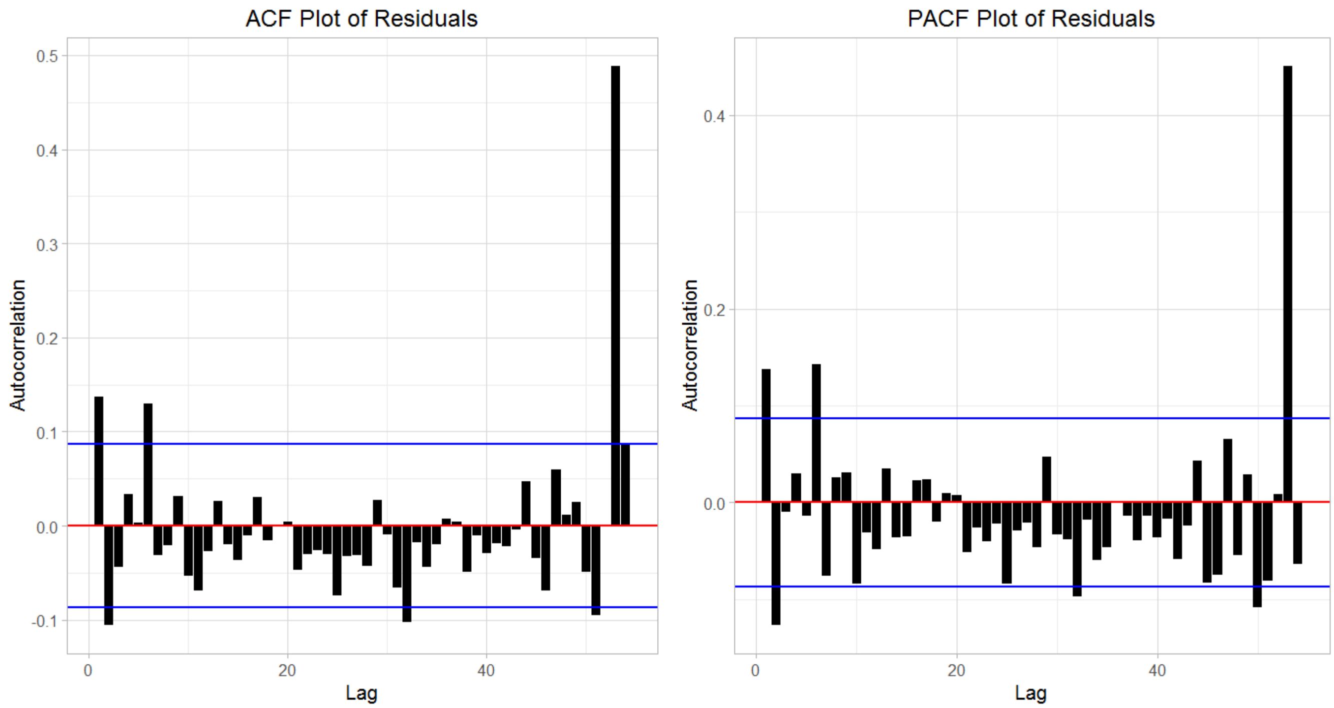


Figure 50: ACF & PACF Plots of Residuals for North Philadelphia Weekly Crime Holt-Winters Smoothing Model

ARIMA Forecasting

The process for choosing the best ARIMA fit and forecast models for Central, West, and North Philadelphia is very similar to the process in the Holt-Winters Forecasting section. The only difference is that ACF and PACF plots of differenced data were used to choose the ranges of ARIMA parameters (p, d, q) in the experiments. Whereas in Holt-Winters, the autofit was used to determine how to vary parameters. The ranges for the seasonal ARIMA parameters (P, D, Q) were only varied between 0 and 1, because the computation time is long.

Central Philadelphia

The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 2 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 48 models, one of which being the auto fit (Scenario 1), were tested.

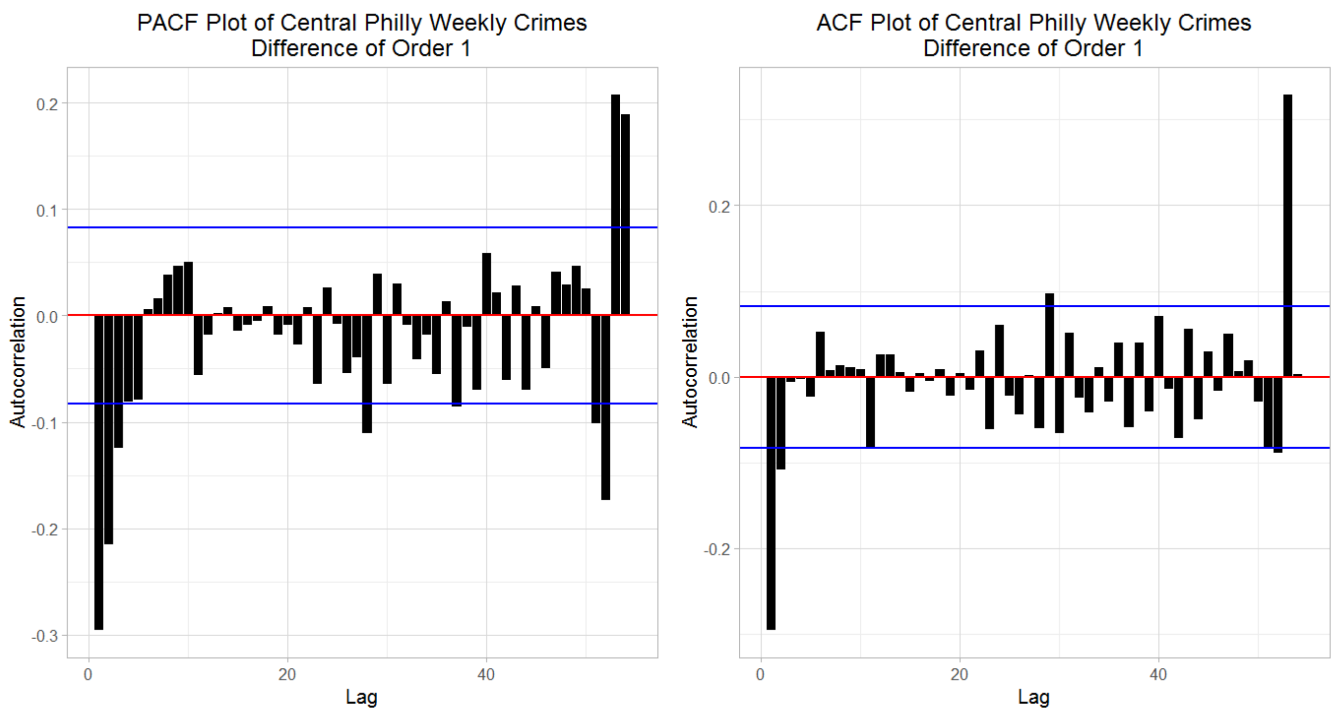


Figure 51: ACF & PACF Plots of Differenced Central Philadelphia Weekly Crimes

The numerical performance of the 2 Central Philadelphia fitted models is shown below. Scenario 17 had slightly better performance across all independent variables. Scenario 17 is the model that was eventually chosen as the best ARIMA model for fitting Central Philadelphia weekly crimes for 2006 – 2016.

Table 14: DOE Results for Fitting Central Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	1	1	1	0	0	1	3	0.59472	-1.9491	82.7099	55.5347	-5.6169	13.5795
17	2	1	1	0	0	1	4	0.60257	-1.9049	82.589	55.509	-5.6036	13.5742

The 1-step ahead performance for the best fitted Central Philadelphia ARIMA model is shown below. The model captures the center of the data and the peaks very well, but misses the lower values often.

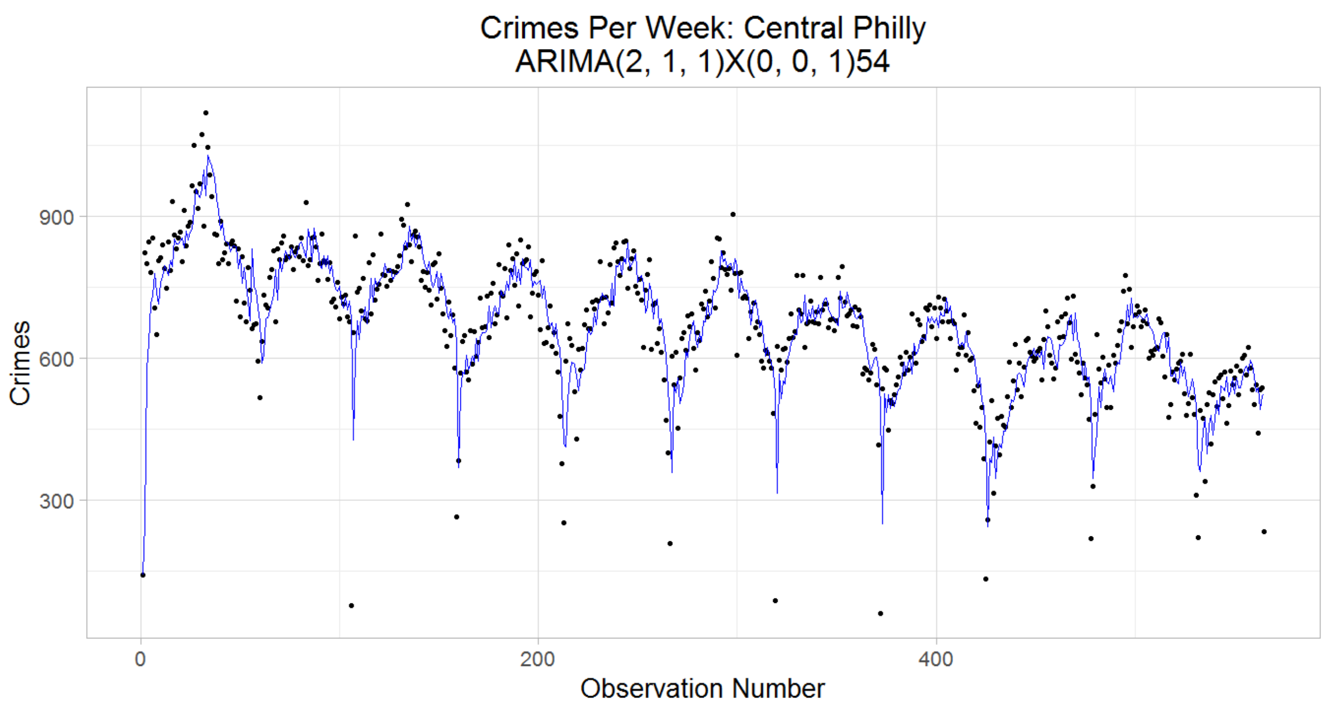


Figure 52: 10 Years of Central Philadelphia Weekly Crime & ARIMA Model Fit

The residual plots of the best fitted Central Philadelphia ARIMA model are shown below. The residuals v. fitted plot for the model show that a constant mean was achieved. Variance appears constant too. This plot also shows that the residuals are centered at zero. The residual time series of the model is the result of trying to remove trend and seasonality from weekly crimes. This plot shows that annual seasonality may have been removed. The residuals for the model don't follow a normal distribution. The tails in the QQ plot are trailing off the line and 95% confidence interval band too much. The histogram shows that most of the data is center about zero, and that zero is contained within the one sample t-test 95% confidence interval.

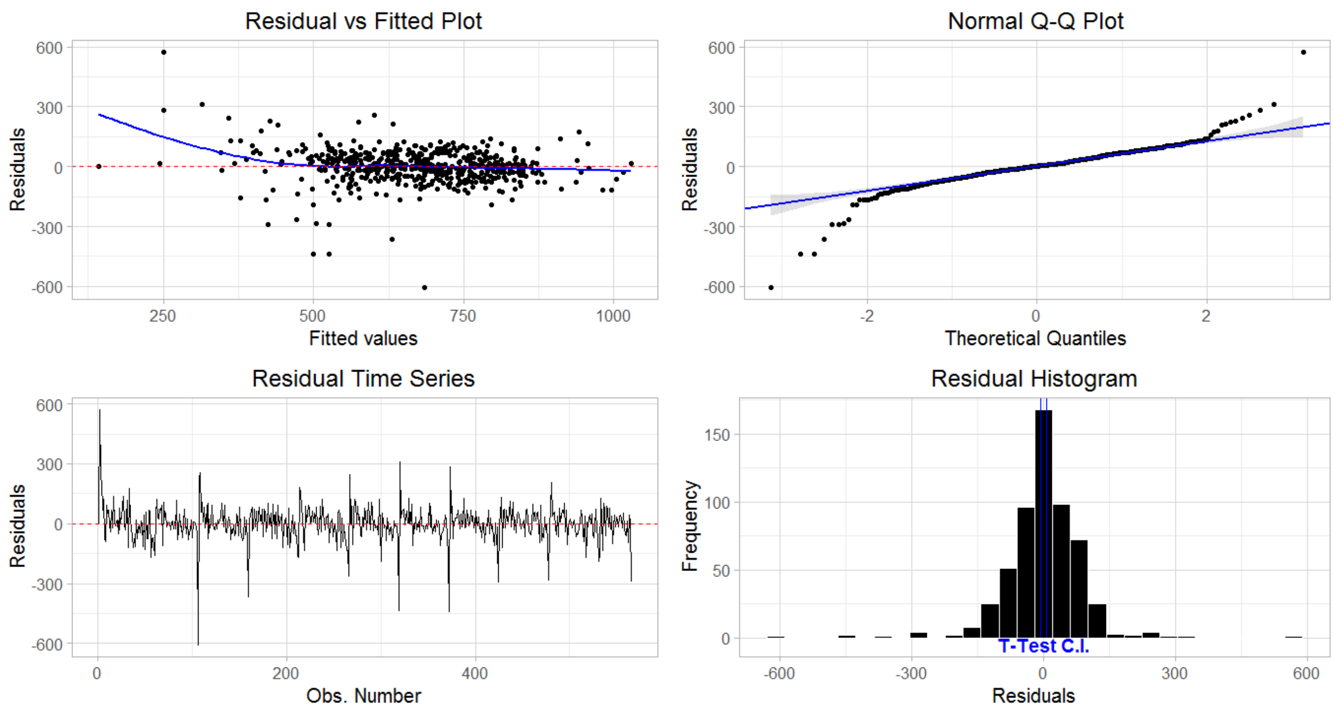


Figure 53: Residual Diagnostics of Central Philadelphia Weekly Crime ARIMA Model Fit

The autocorrelation levels of the residuals for the model are not significant, except for the end which is likely due to the varying seasonal length of 52 – 54 weeks. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below show that autocorrelation has been removed.

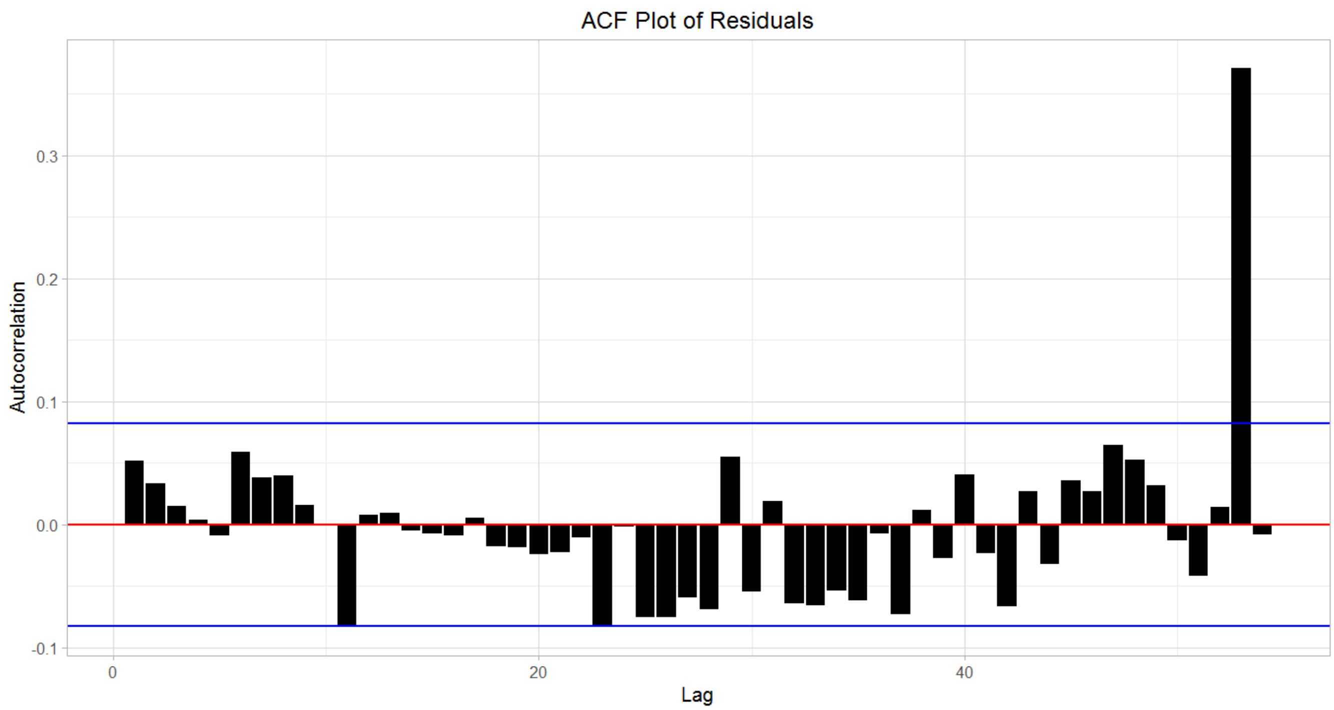


Figure 54: ACF Plot of Residuals for Central Philadelphia Weekly Crime ARIMA Model Fit

The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 24 models, one of which being the auto fit (Scenario 1), were tested.

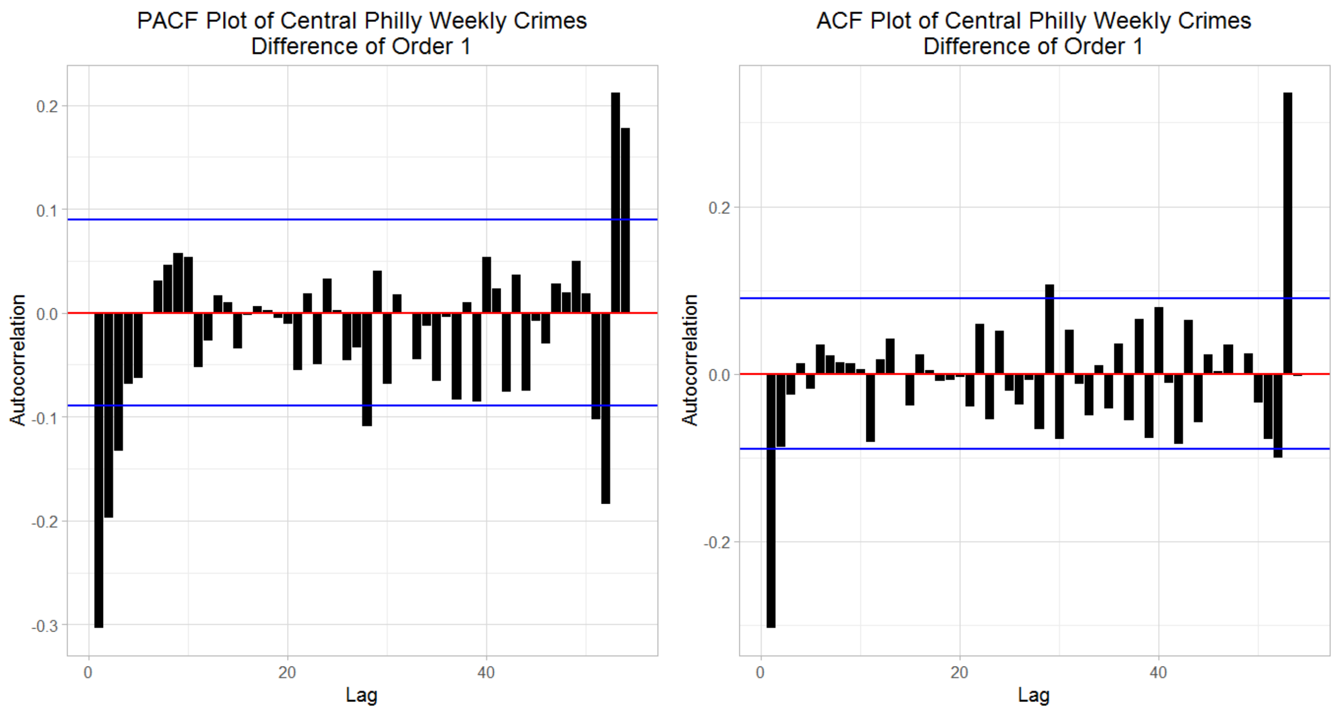


Figure 55: ACF & PACF Plots of Differenced Central Philadelphia Weekly Crimes - Training Data

The numerical performance of the 2 Central Philadelphia forecasting models is shown below. Scenario 13 had better performance across all independent variables. Scenario 13 is the model that was eventually chosen as the best ARIMA model for forecasting Central Philadelphia weekly crimes for 2015 – 2016.

Table 15: DOE Results for Forecasting Central Philadelphia Weekly Crimes 2015 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	1	1	1	0	0	1	3	5.18E-22	128.013	157.23	141.208	19.6141	24.7333
13	2	1	1	0	1	1	4	1.01E-06	57.4061	116.252	91.6005	7.80079	17.4541

The forecast performance for the best Central Philadelphia ARIMA model is shown below. The model under predicts the first half of each season, and then predicts the center of the second half of each very well. The prediction intervals at 80% and 95% confidence levels, capture most of the actual values.

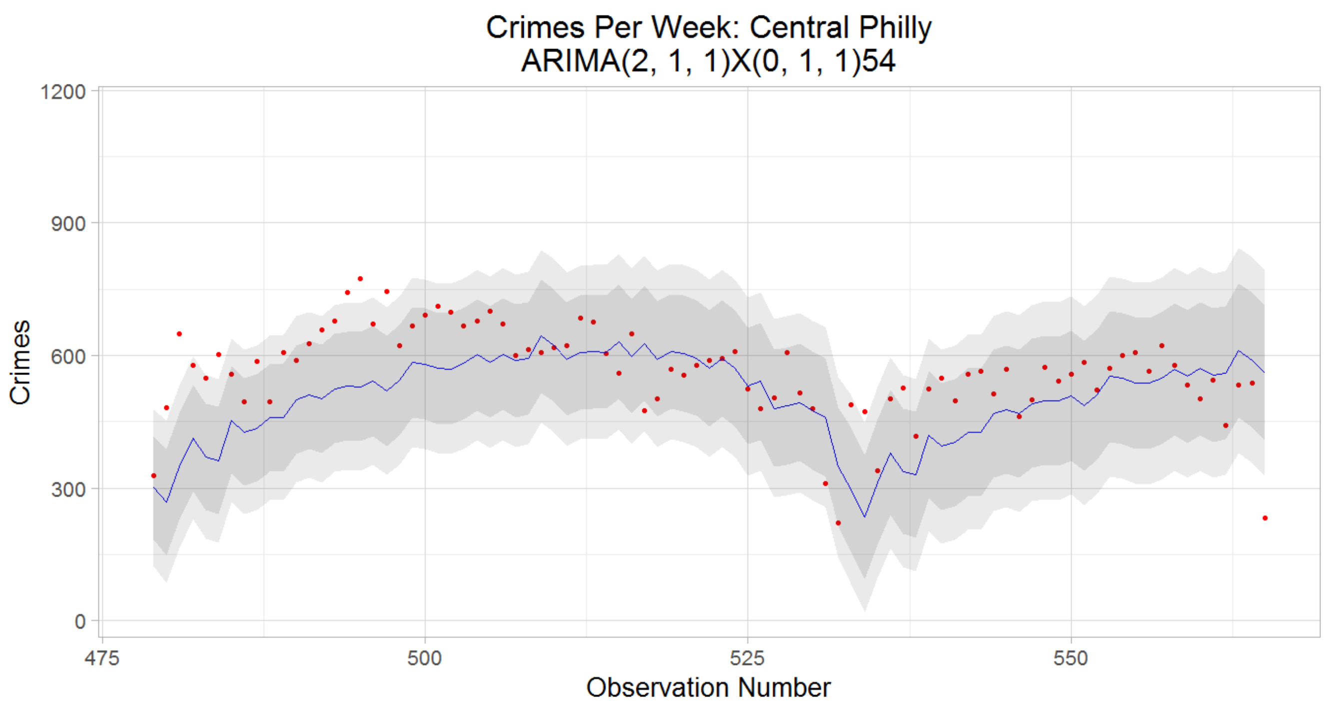


Figure 56: Prediction Plot for Central Philadelphia Weekly Crime ARIMA Forecasting Model

The forecasting comparison of the best Holt-Winter model and the best ARIMA model for Central Philadelphia 2015 – 2016 is shown below. The Holt-Winter model outperforms the ARIMA model across all independent variables. This may be due to the Holt-Winters function being able to compute many more models within the chosen 5 minute limit for model building. This efficiency allows for a larger sample of models to choose from, and results in the best model being found easier.

Table 16: Central Philadelphia Weekly Crime Forecasting Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	9.16E-01	0.85826	74.9167	55.7464	-2.1416	11.5526
ARIMA	1.01E-06	57.4061	116.252	91.6005	7.80079	17.4541

West Philadelphia

The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 2 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 49 models, one of which being the auto fit (Scenario 1), were tested.

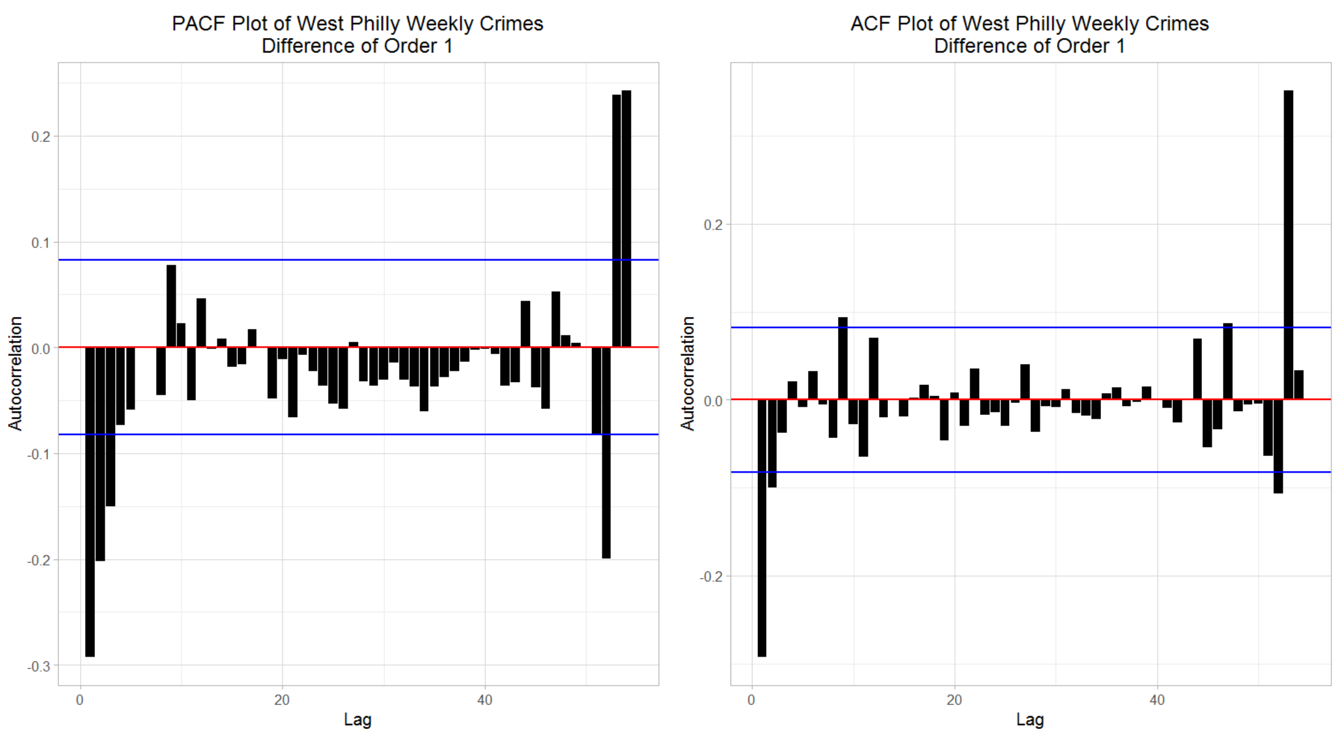


Figure 57: ACF & PACF Plots of Differenced West Philadelphia Weekly Crimes

The numerical performance of the 2 West Philadelphia fitted models is shown below. Scenario 1 had slightly better performance across most independent variables. Scenario 1 is the model that was eventually chosen as the best ARIMA model for fitting West Philadelphia weekly crimes for 2006 – 2016.

Table 17: DOE Results for Fitting West Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	1	0	1	0	0	1	3	9.40E-01	-0.2807	83.8639	57.1777	-5.8464	13.9571
8	3	1	1	1	0	0	5	7.07E-01	-1.4113	84.8766	58.517	-5.5576	14.0096

The 1-step ahead performance for the best fitted West Philadelphia ARIMA model is shown below. The model captures the center of the data well, but misses the peaks and lower values often.

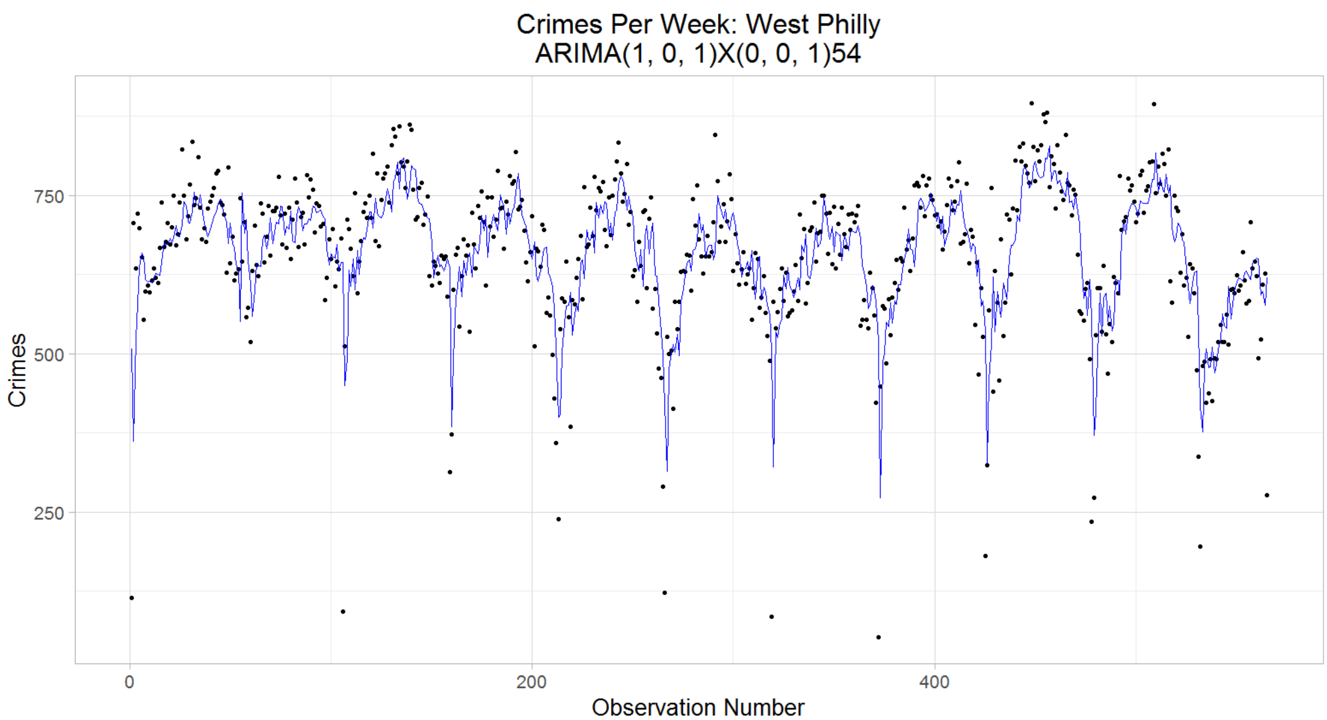


Figure 58: 10 Years of West Philadelphia Weekly Crime & ARIMA Model Fit

The residual plots of the best fitted West Philadelphia ARIMA model are shown below. The residuals v. fitted plot for the model show that a constant mean was achieved. Variance is not constant. This plot also shows that the residuals are centered at zero. The residual time series of the model is the result of trying to remove trend and seasonality from weekly crimes. This plot shows that annual seasonality may have been removed. The residuals for the model don't follow a normal distribution. The tails in the QQ plot are trailing off the line and 95% confidence interval band too much. The histogram shows that most of the data is center about zero, and that zero is contained within the one sample t-test 95% confidence interval.

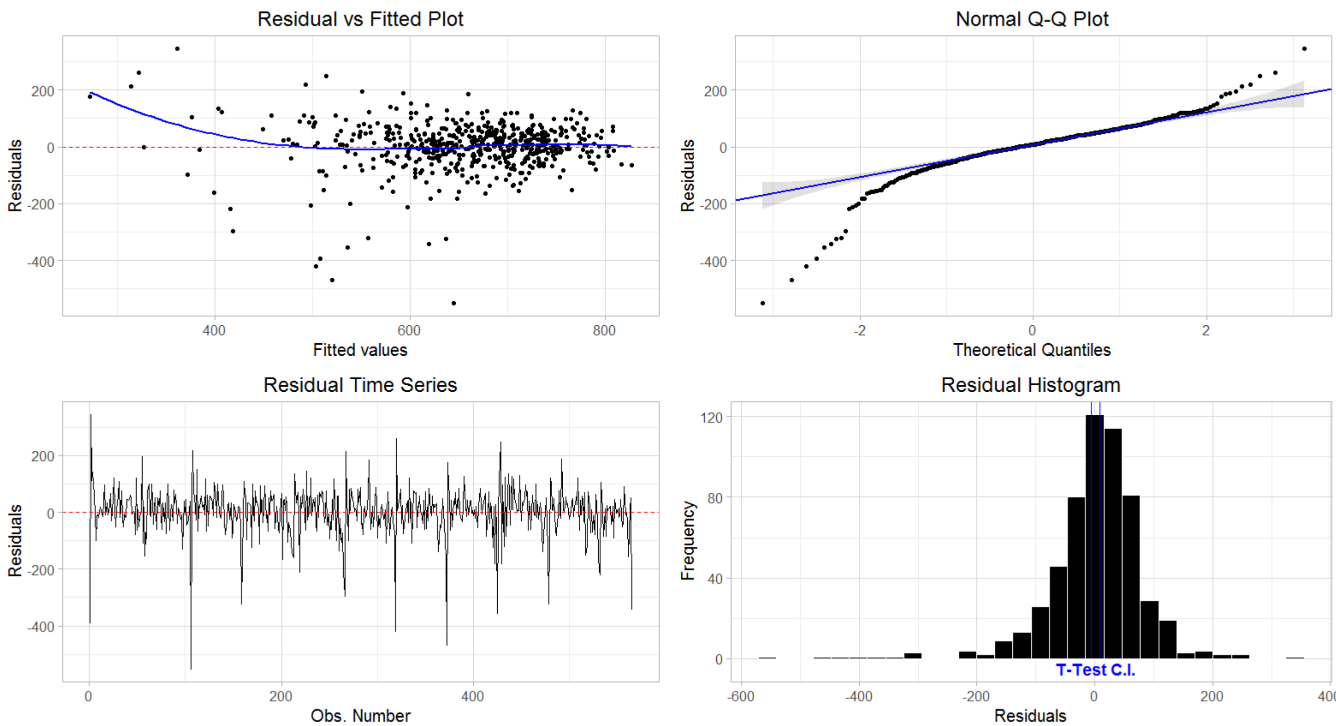


Figure 59: Residual Diagnostics of West Philadelphia Weekly Crime ARIMA Model Fit

The autocorrelation levels of the residuals for the model are not significant, except for the end which is likely due to the varying seasonal length of 52 – 54 weeks. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below show that autocorrelation has likely been removed.

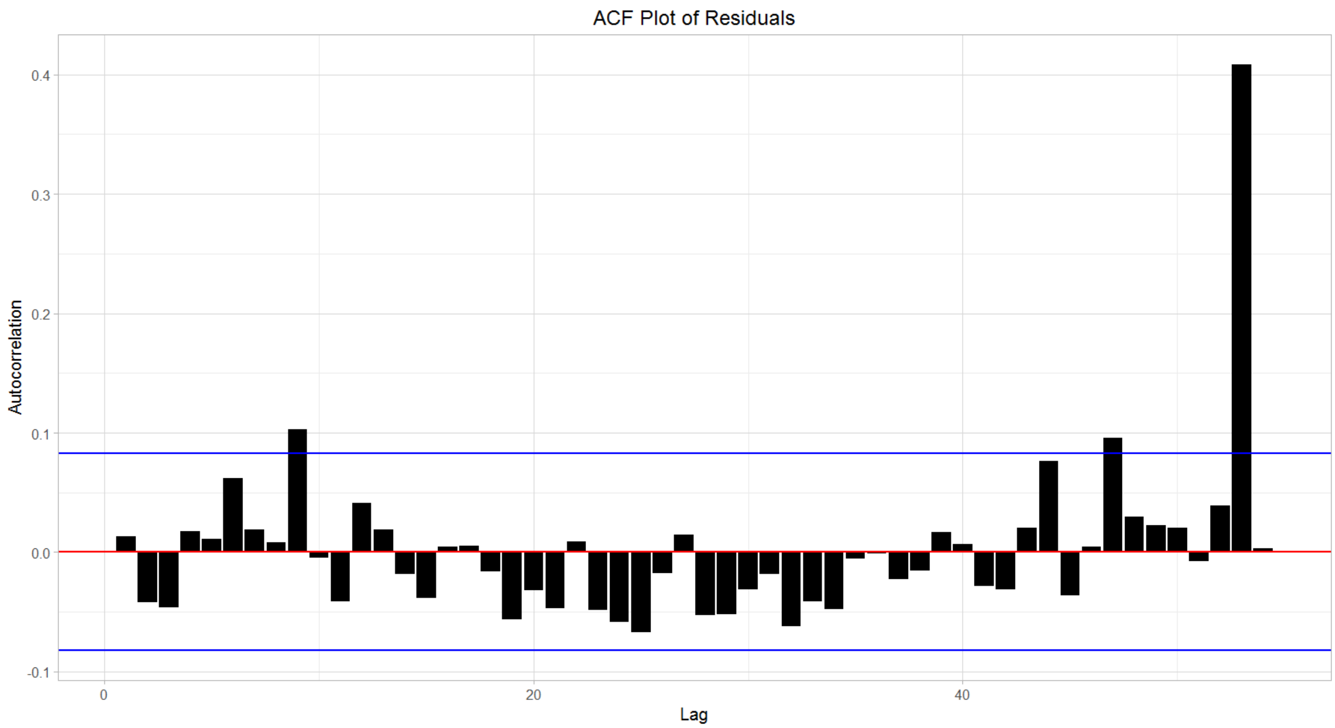


Figure 60: ACF Plot of Residuals for West Philadelphia Weekly Crime ARIMA Model Fit

The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 25 models, one of which being the auto fit (Scenario 1), were tested.

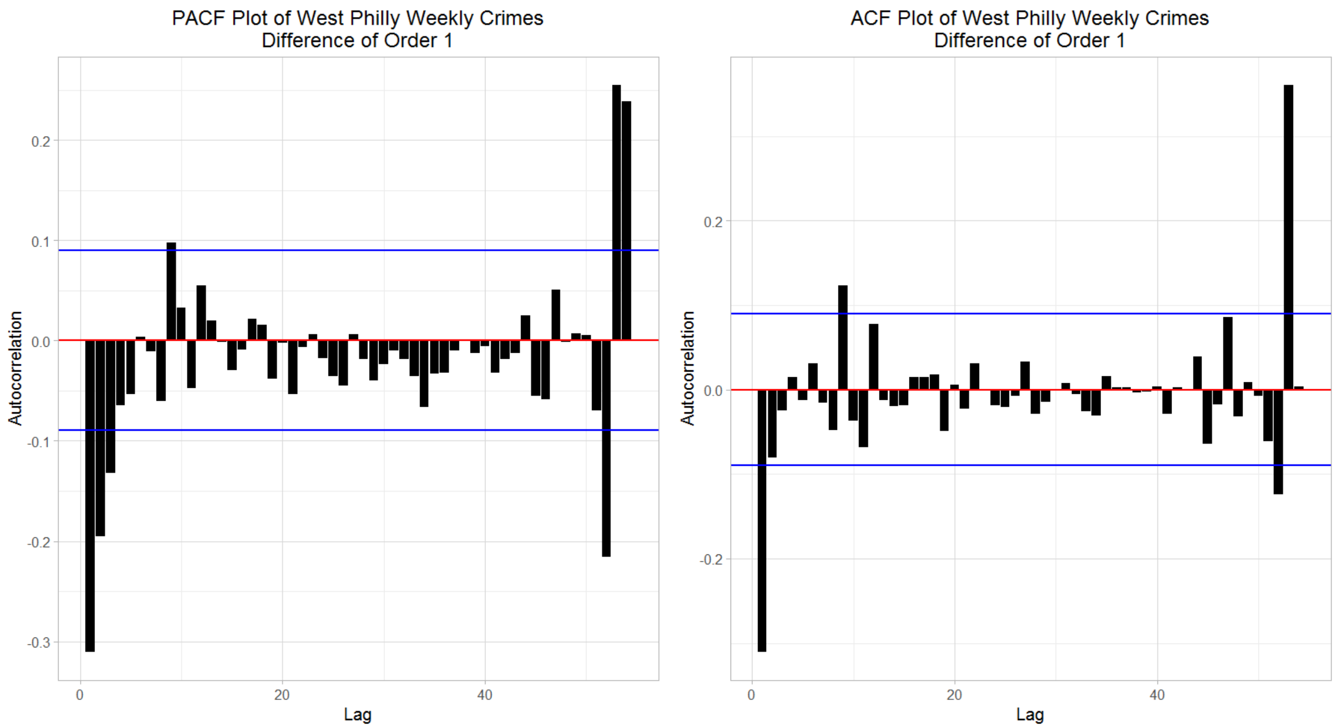


Figure 61: ACF & PACF Plots of Differenced West Philadelphia Weekly Crimes - Training Data

The numerical performance of the 2 West Philadelphia forecasting models is shown below. Scenario 16 had better performance across all independent variables. Scenario 16 is the model that was eventually chosen as the best ARIMA model for forecasting West Philadelphia weekly crimes for 2015 – 2016.

Table 18: DOE Results for Forecasting West Philadelphia Weekly Crimes 2015 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	1	0	1	0	0	1	3	7.37E-02	-22.255	116.156	92.6281	-9.1577	18.6198
16	3	1	1	0	1	1	5	6.01E-01	-5.6713	100.412	75.5358	-4.0598	14.6643

The forecast performance for the best West Philadelphia ARIMA model is shown below. The model misses the peak of 2015 and then over predicts the peak in 2016. The prediction intervals at 80% and 95% confidence levels, capture almost all of the actual values.

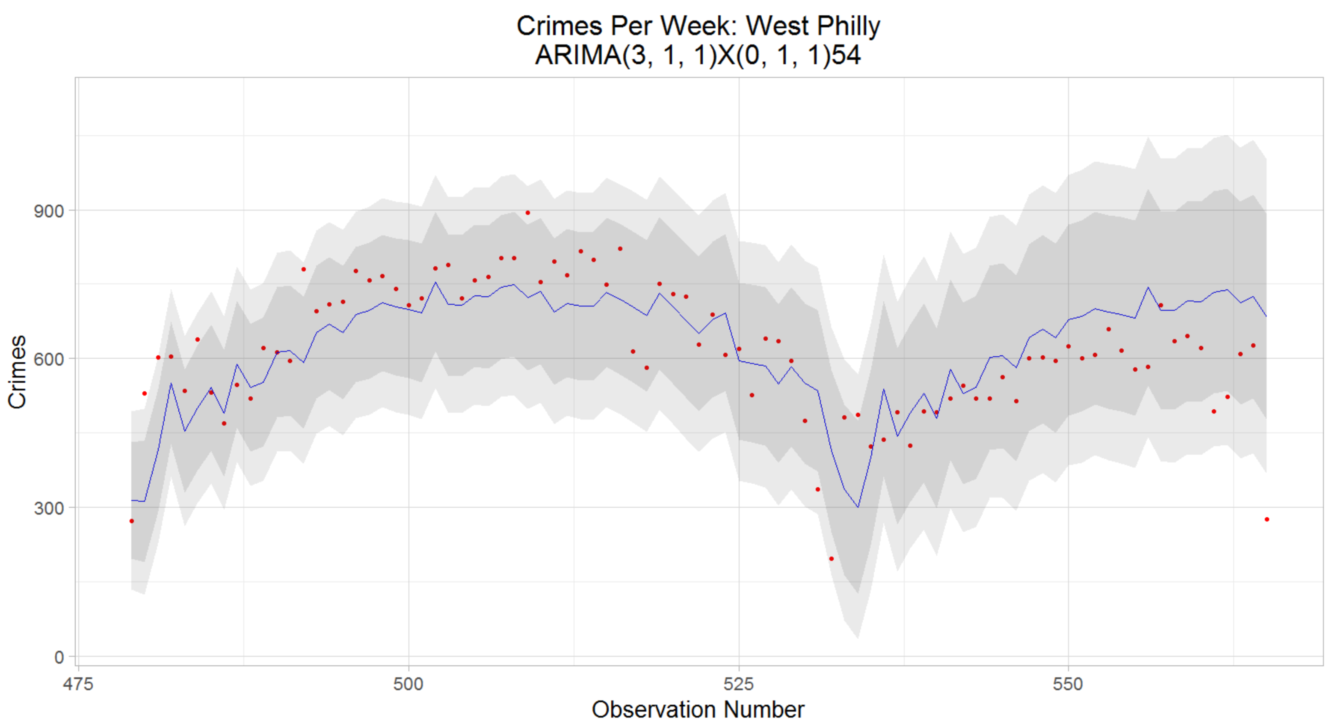


Figure 62: Prediction Plot for West Philadelphia Weekly Crime ARIMA Forecasting Model

The forecasting comparison of the best Holt-Winter model and the best ARIMA model for West Philadelphia 2015 – 2016 is shown below. The Holt-Winter model outperforms the ARIMA model across all independent variables. This may be due to the Holt-Winters function being able to compute many more models within the chosen 5 minute limit for model building. This efficiency allows for a larger sample of models to choose from, and results in the best model being found easier.

Table 19: West Philadelphia Weekly Crime Forecasting Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	9.63E-01	-0.4019	80.5195	58.8806	-2.8648	11.8654
ARIMA	6.01E-01	-5.6713	100.412	75.5358	-4.0598	14.6643

North Philadelphia

The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 2 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 49 models, one of which being the auto fit (Scenario 1), were tested.

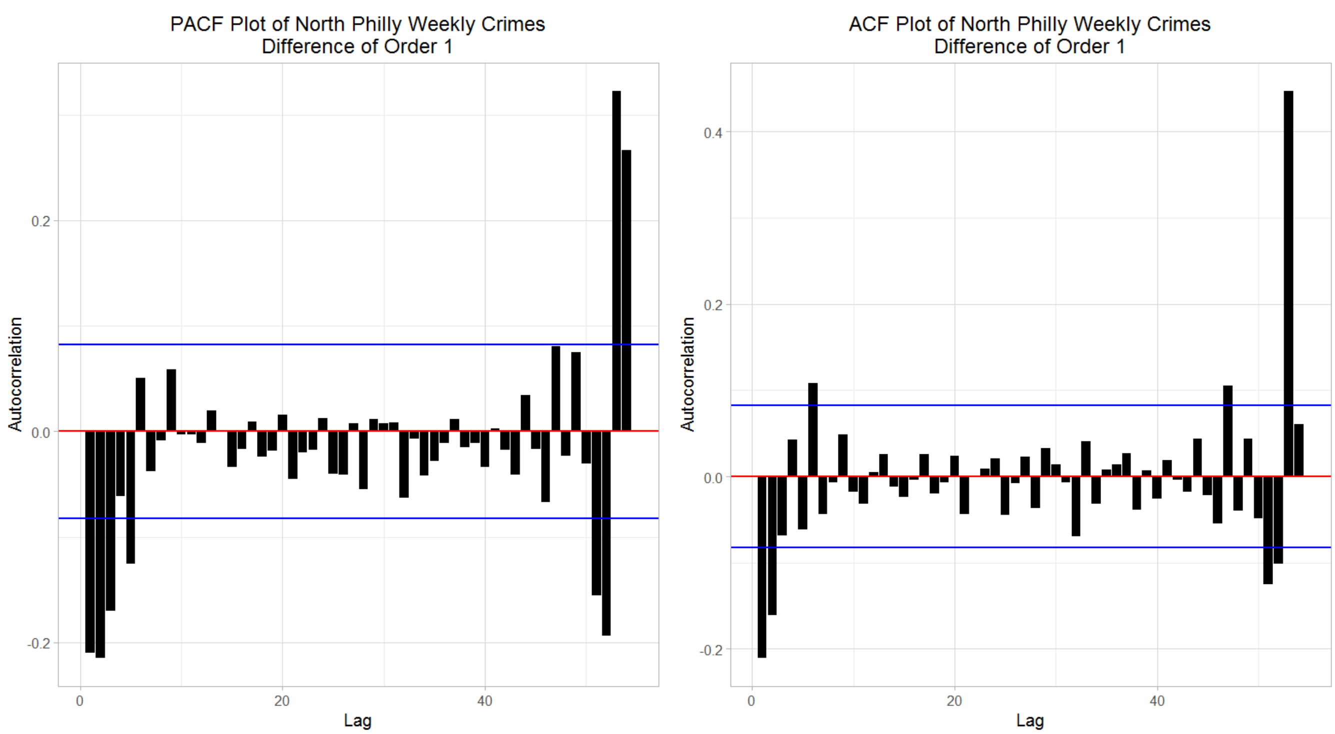


Figure 63: ACF & PACF Plots of Differenced North Philadelphia Weekly Crimes

The numerical performance of the 2 North Philadelphia fitted models is shown below. Scenario 10 had slightly better performance across all independent variables. Scenario 10 is the model that was eventually chosen as the best ARIMA model for fitting North Philadelphia weekly crimes for 2006 – 2016.

Table 20: DOE Results for Fitting North Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	1	1	2	0	0	2	5	5.29E-01	-5.7311	205.687	129.786	-5.2429	12.3078
10	2	1	2	1	0	0	5	5.74E-01	-5.1099	205.406	129.643	-5.1518	12.2856

The 1-step ahead performance for the best fitted North Philadelphia ARIMA model is shown below. The model captures the center of the data and the peaks very well, but misses the lower values often.

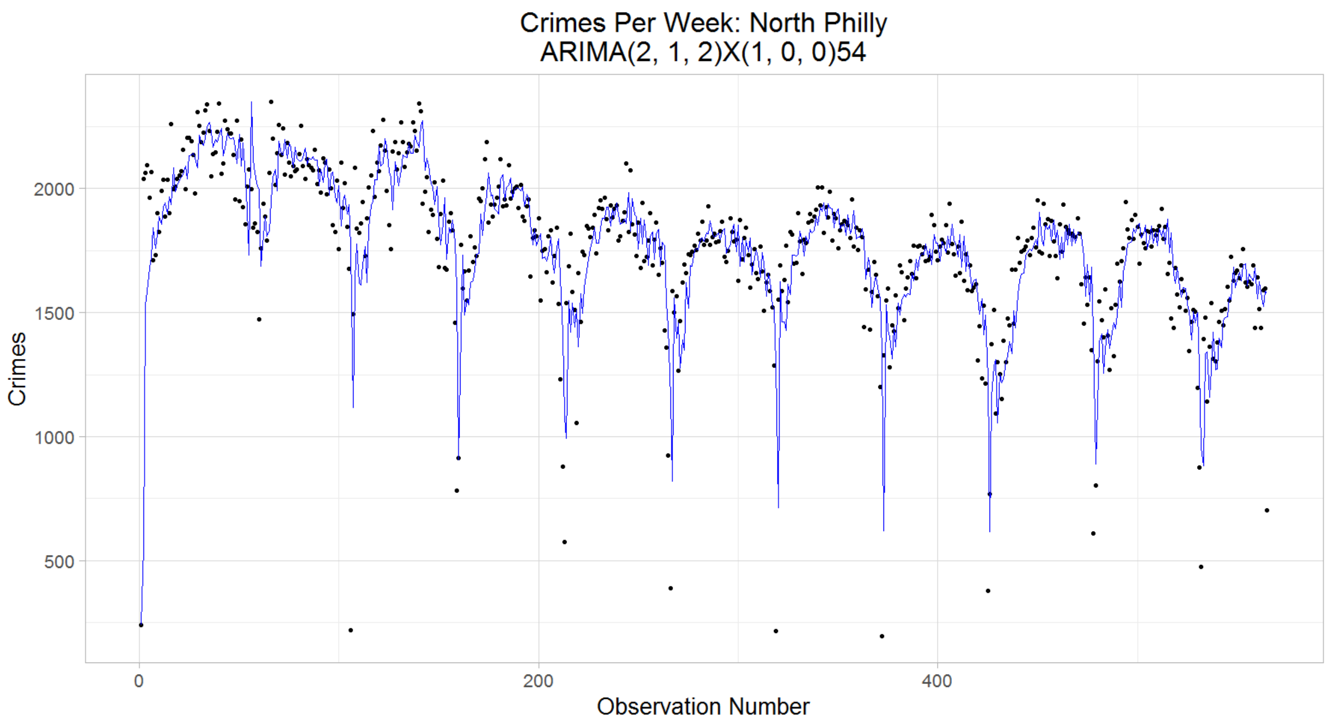


Figure 64: 10 Years of North Philadelphia Weekly Crime & ARIMA Model Fit

The residual plots of the best fitted North Philadelphia ARIMA model are shown below. The residuals v. fitted plot for the model show that a constant mean was achieved. Variance appears constant too. This plot also shows that the residuals are centered at zero. The residual time series of the model is the result of trying to remove trend and seasonality from weekly crimes. This plot shows that annual seasonality may have been removed. The residuals for the model don't follow a normal distribution. The tails in the QQ plot are trailing off the line and 95% confidence interval band too much. The histogram shows that most of the data is center about zero, and that zero is contained within the one sample t-test 95% confidence interval.

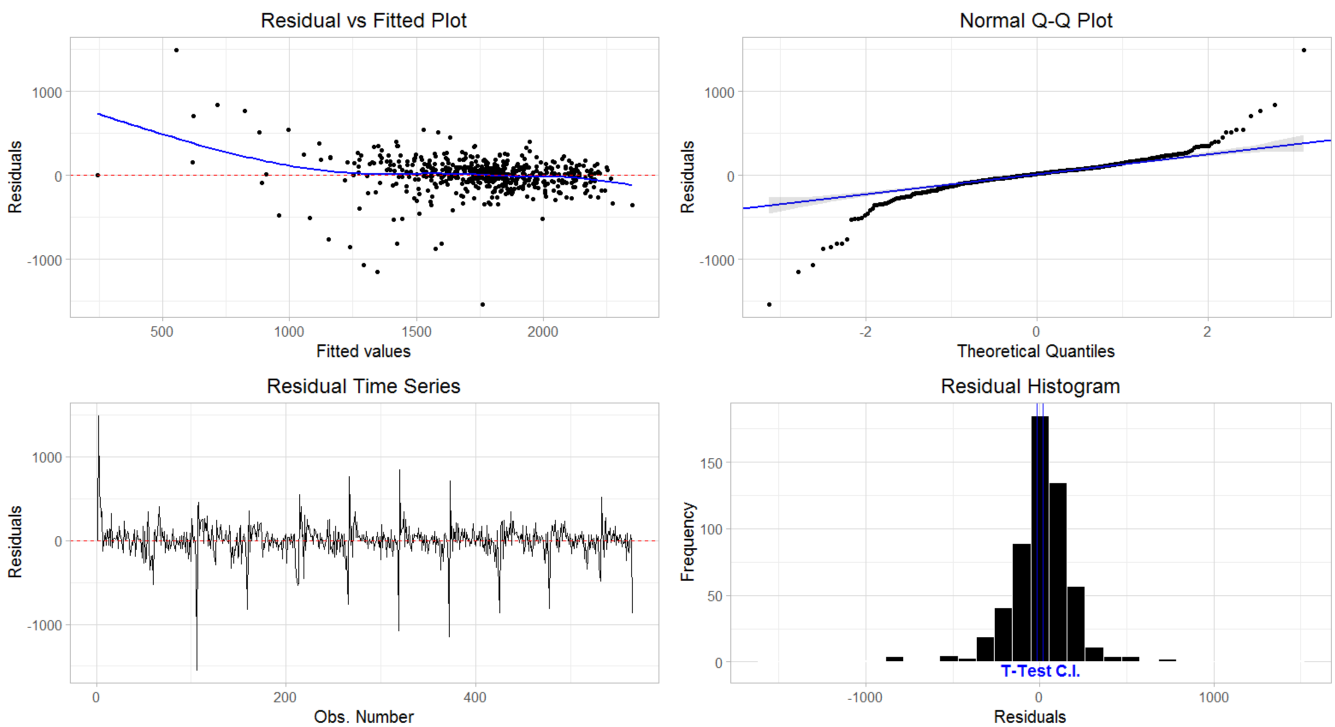


Figure 65: Residual Diagnostics of North Philadelphia Weekly Crime ARIMA Model Fit

The autocorrelation levels of the residuals for the model are not significant, except for the end which is likely due to the varying seasonal length of 52 – 54 weeks. This ACF plot goes out to a lag of 54 weeks to capture a full season of autocorrelation. The behavior of the bars in the ACF plot below show that autocorrelation has been removed.

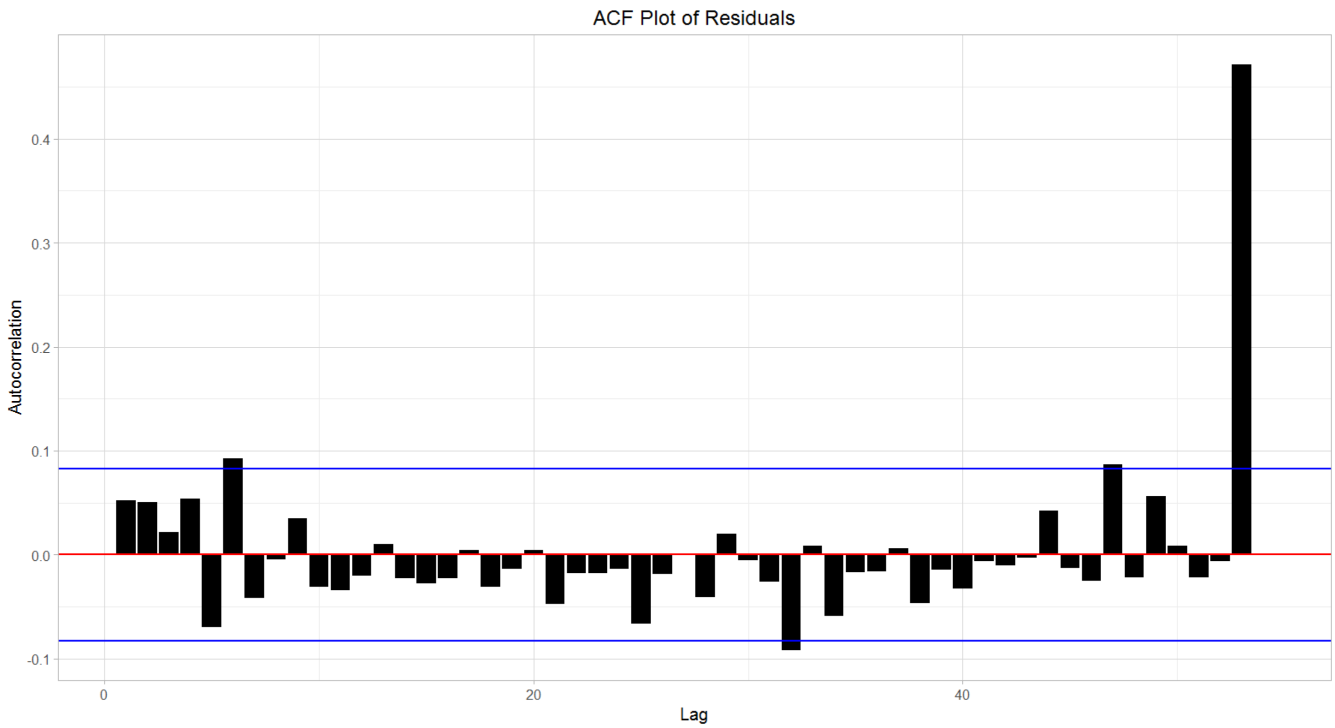


Figure 66: ACF Plot of Residuals for North Philadelphia Weekly Crime ARIMA Model Fit

The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 2 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 48 models, one of which being the auto fit (Scenario 1), were tested.

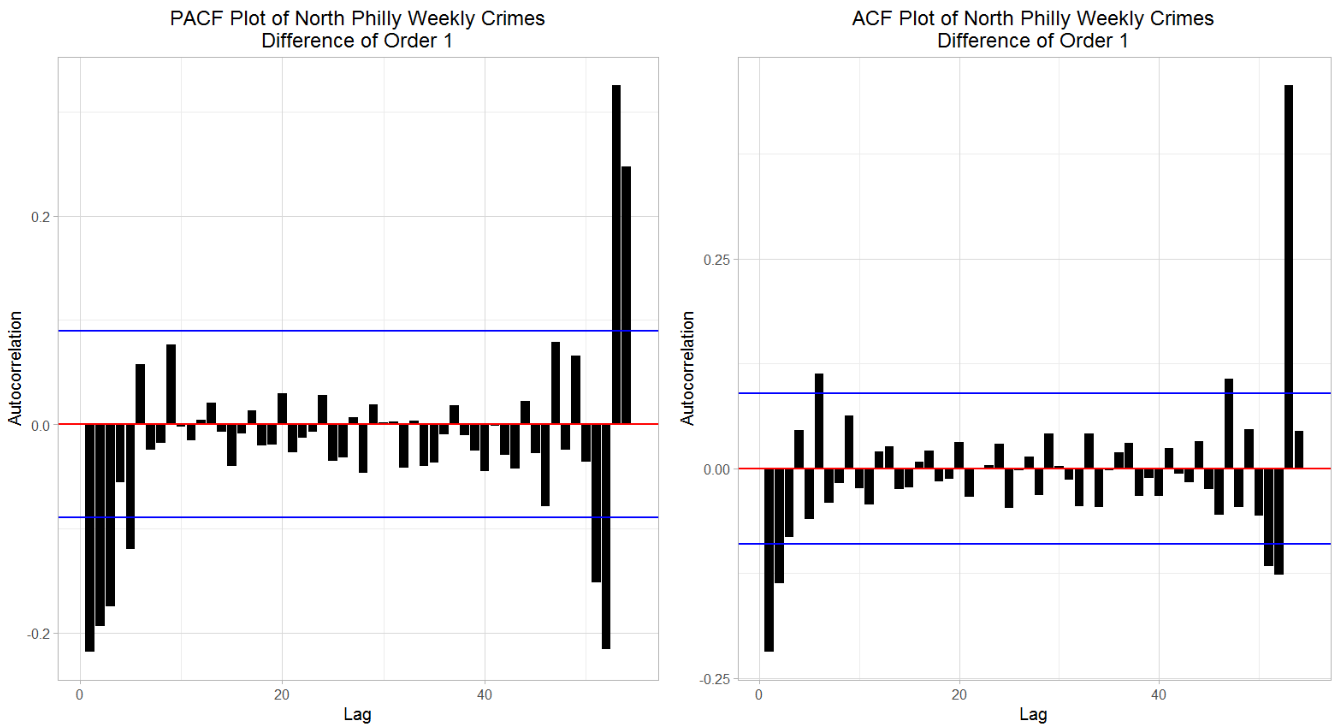


Figure 67: ACF & PACF Plots of Differenced North Philadelphia Weekly Crimes - Training Data

The numerical performance of the 2 North Philadelphia forecasting models is shown below. Scenario 31 had better performance across all independent variables. Scenario 31 is the model that was eventually chosen as the best ARIMA model for forecasting North Philadelphia weekly crimes for 2015 – 2016.

Table 21: DOE Results for Forecasting North Philadelphia Weekly Crimes 2015 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	1	1	2	0	0	1	4	6.97E-16	233.506	319.816	282.594	11.717	18.9551
31	2	1	1	0	1	1	4	2.22E-01	29.1673	221.682	152.684	-0.6937	12.105

The forecast performance for the best North Philadelphia ARIMA model is shown below. The model under predicts up through the peak for each season, and then over predicts the end of each season. The prediction intervals at 80% and 95% confidence levels, capture almost all of the actual values.

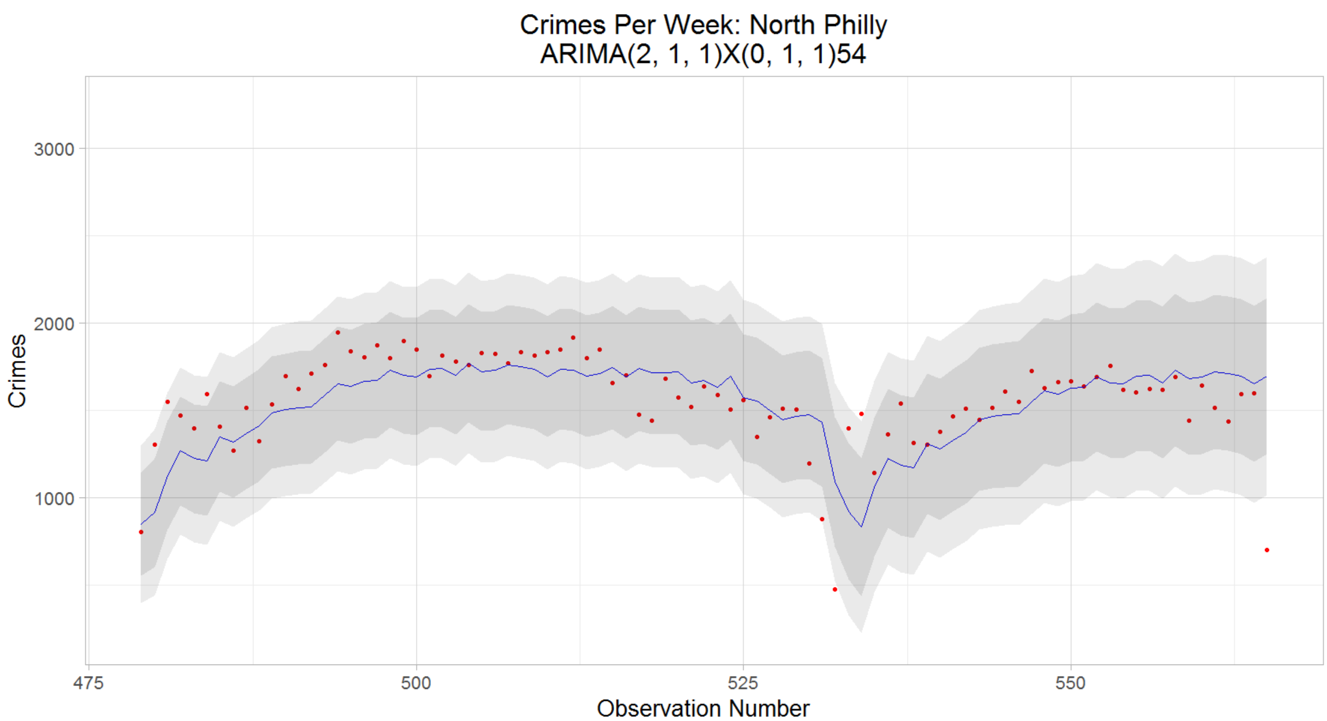


Figure 68: Prediction Plot for North Philadelphia Weekly Crime ARIMA Forecasting Model

The forecasting comparison of the best Holt-Winter model and the best ARIMA model for North Philadelphia 2015 – 2016 is shown below. The Holt-Winter model outperforms the ARIMA model across most independent variables. This may be due to the Holt-Winters function being able to compute many more models within the chosen 5 minute limit for model building. This efficiency allows for a larger sample of models to choose from, and results in the best model being found easier.

Table 22: North Philadelphia Weekly Crime Forecasting Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	8.86E-01	-2.8834	185.661	121.258	-2.4663	9.86163
ARIMA	2.22E-01	29.1673	221.682	152.684	-0.6937	12.105

Dynamic Regression Forecasting

The process for choosing the best dynamic regression fit and forecast models for Central, West, and North Philadelphia is very similar to the process in the ARIMA Forecasting Section. The only difference is that a linear regression model of the regressors was built and then ACF and PACF plots of the residuals from that linear regression model were used to choose the ranges of dynamic regression parameters in the experiments. The forecasting time period will be 2015 – 2016 to allow for comparison between all forecasting methods evaluated thus far. The time period of 2015 – 2016 is chosen also because it contains one season and the most recent data. This allows for each forecasting method to show how well it can capture an entire season, which is important because the data is seasonal. This also allows for the forecasting methods to capture the most recent crime behavior which is important because predicting crime in the near future will determine how useful these forecasting methods are for Philadelphia law enforcement.

The 2 regressors chosen for Central, West, and North Philadelphia are Days and Summer. Days represents the number of days in a given week. This was chosen as a regressor to better capture the low values of weekly crime. Summer represents whether or not a given week is fully or partially contained within the summer season: June Solstice to September Equinox. This was chosen as a regressor to better capture the high values of weekly crime. The plots below are violin plots for Central Philadelphia, which show the distribution and density of weekly crimes across the values of each regressor. These plots show that there is a positive linear relationship between weekly crime and each regressor. The violin plots for West and North Philadelphia show the same relationships. Days is modeled as a continuous regressor because computation time lengthens as more coefficients need to be estimated. Summer was modeled as a categorical variable because it only required one dummy variable. These regessors are computed based on the calendar year.

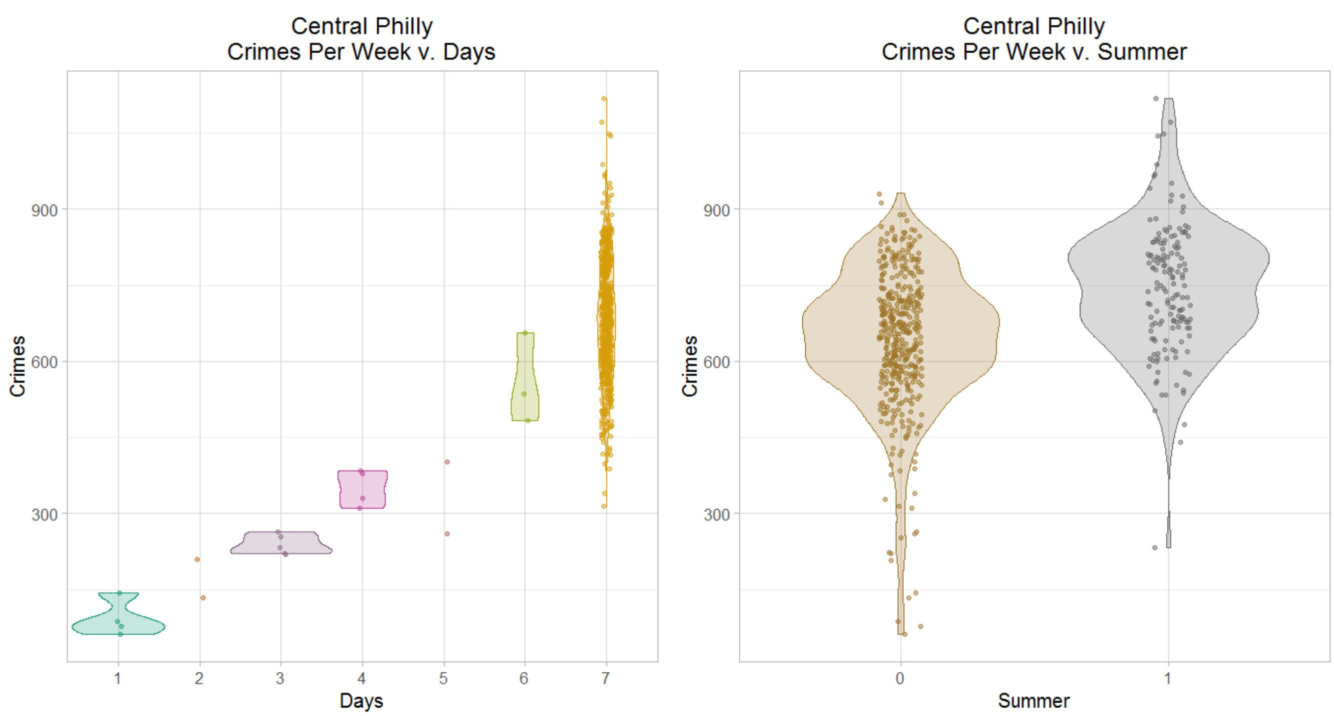


Figure 69: Central Philadelphia Dynamic Regression Regressors

Central Philadelphia

The need for these regressors was decided by creating dynamic regression autofits for all regressors and each individual regressor, and then comparing the fitness. The table below shows that the autofit using all regressors has better fitness. Therefore both regressors will be used in fitting and forecasting Central Philadelphia weekly crimes.

Table 23: Central Philadelphia Autofit Comparison Across Regressors

Variables	ME	RMSE	MAE	MPE	MAPE	AIC	AICc	BIC
SUMMER,DAYS	-0.7559	58.529	46.1823	-0.6807	7.80796	6200.42	6200.49	6217.76
SUMMER	0.15337	85.7864	57.1039	-4.9292	13.1194	6636.3	6636.41	6657.98
DAYS	-0.4454	58.6125	46.3762	-0.5478	7.84931	6204.06	6204.16	6225.73

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days and Summer as regressors and weekly crimes as the response variable. Evaluating the autocorrelation of these residuals will give insight onto what ARIMA parameters (p, d, q) are worth testing for the dynamic regression model. The bars in these charts indicate that an autoregressive order up to 2, and a moving average order up to 2 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 37 models, one of which being the auto fit (Scenario 1), were tested.

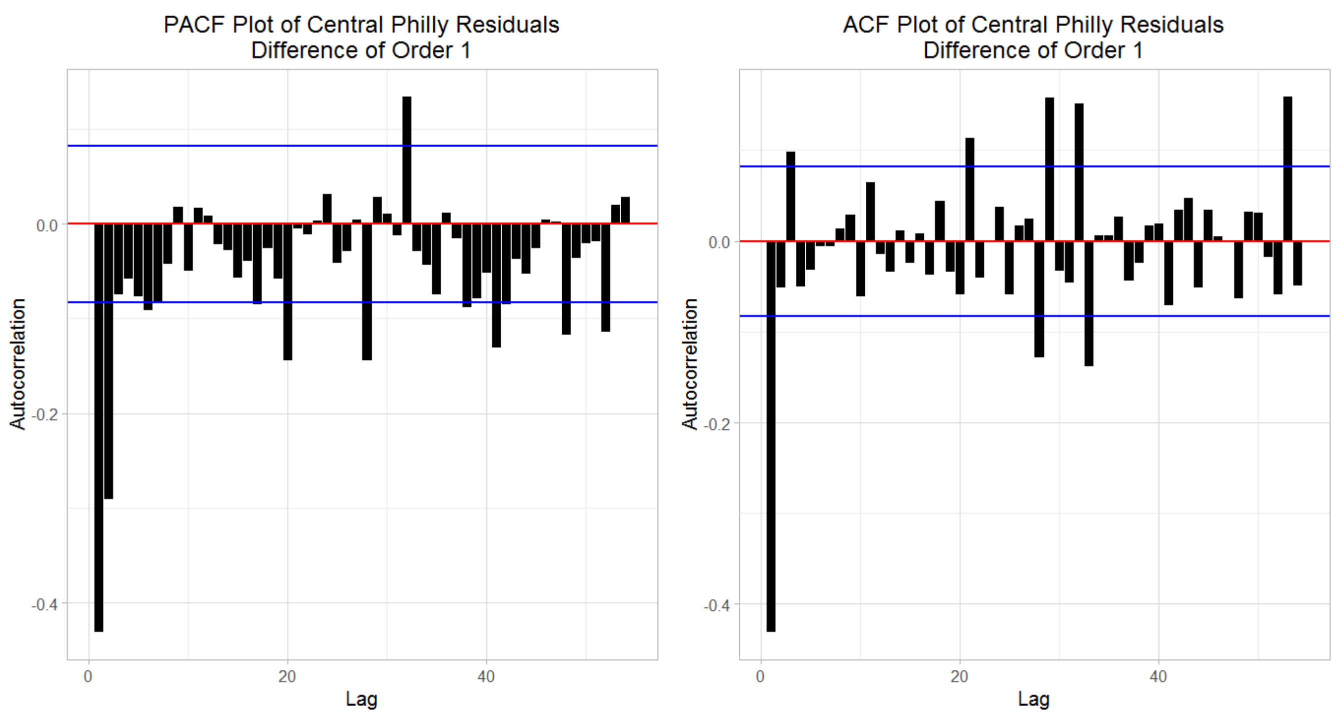


Figure 70: ACF & PACF Plot of Linear Regression Residuals for Central Philadelphia Dynamic Regression Fitting

The numerical performance of the 2 Central Philadelphia fitted models is shown below. Scenario 1 had slightly better performance across most independent variables. Scenario 1 is the model that was eventually chosen as the best dynamic regression model for fitting Central Philadelphia weekly crimes for 2006 – 2016.

Table 24: DOE Results for Fitting Central Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0	1	1	0	0	0	1	6.53E-01	-1.1316	56.8011	44.7697	-0.7319	7.88243
21	0	1	1	0	1	1	2	8.84E-01	-0.3871	59.835	47.415	-0.109	8.24459

The 1-step ahead performance for the best fitted Central Philadelphia dynamic regression model is shown below. The model captures the center of the data, the peaks, and the low values very well. The positive coefficients for each regressor agree with the positive linear relationship seen in the violin plots. The regressors are helping, especially Days which captures the low values that previous models failed to capture.

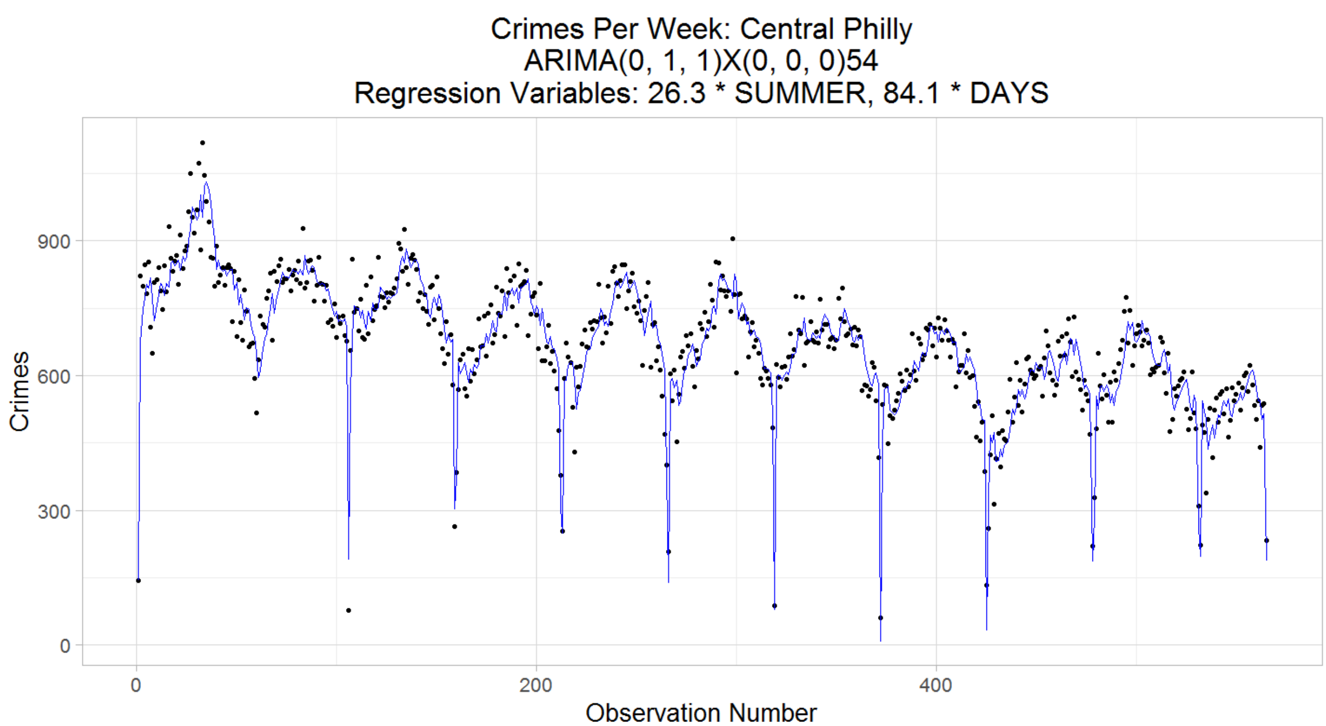


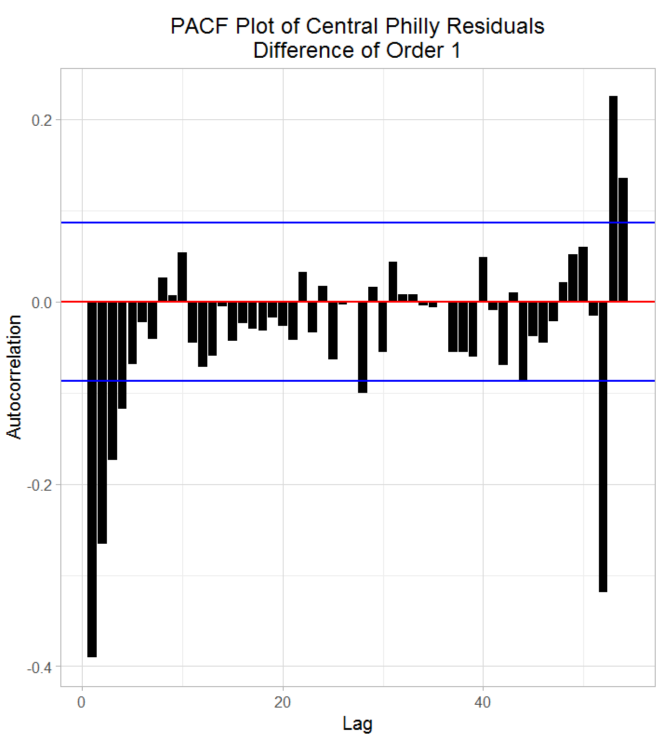
Figure 71: 10 Years of Central Philadelphia Weekly Crime & Dynamic Regression Model Fit

The choice of potential lagged regressors were based on lags of 1 week, 4 weeks, and 54 weeks. These lags were chosen based on the weekly, monthly, and annual seasonality of the data as shown in earlier time series. The chosen lagged regressors were based on trying 9 possible combinations shown in the table below. Each row corresponds to the fitness of a linear regression model. The table indicates that lagging Days and Summer by 54 weeks, scenario 9, is the better option.

Table 25: Results for Testing Lagged Regressors

Scenario	SUMMER	DAYS	r2pred	AIC	BIC
1	SUMMER1	DAYS1	0.126	6363.84	6380.78
2	SUMMER4	DAYS1	0.1057	6375.77	6392.72
3	SUMMER54	DAYS1	0.1263	6363.64	6380.58
4	SUMMER1	DAYS4	0.0998	6381.32	6398.27
5	SUMMER4	DAYS4	0.0787	6393.32	6410.26
6	SUMMER54	DAYS4	0.1001	6381.12	6398.07
7	SUMMER1	DAYS54	0.1448	6351.24	6368.18
8	SUMMER4	DAYS54	0.1249	6363.28	6380.23
9	SUMMER54	DAYS54	0.1452	6351.04	6367.98

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days54 and Summer54 as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 4, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 21 models, one of which being the auto fit (Scenario 1), were tested.



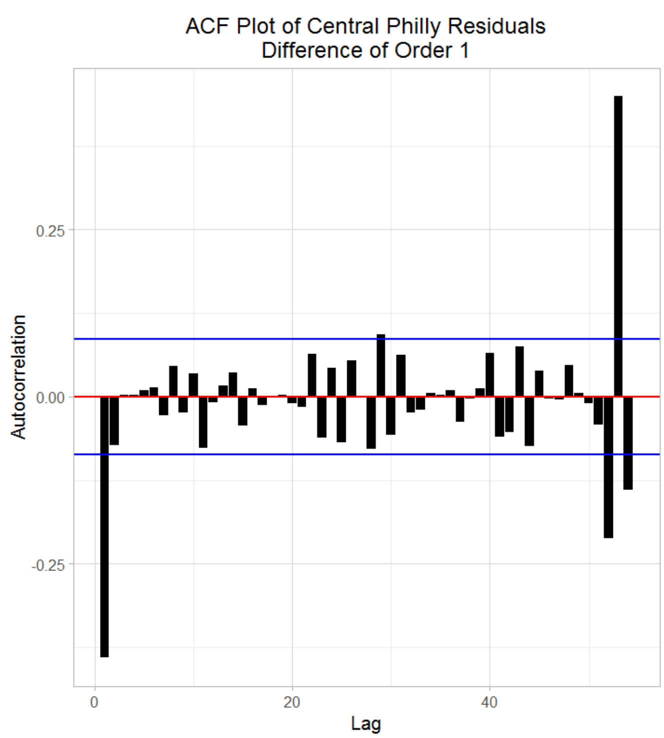


Figure 72: ACF & PACF Plot of Linear Regression Residuals for Central Philadelphia Dynamic Regression Fitting - Lagged

The numerical performance of the 2 Central Philadelphia fitted models is shown below. Scenario 1 had better performance across most independent variables. Scenario 1 is the model that was eventually chosen as the best lagged dynamic model for fitting Central Philadelphia weekly crimes for 2006 – 2016.

Table 26: DOE Results for Fitting Central Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0	1	1	1	0	0	2	6.69E-01	-1.5608	77.8596	54.3154	-4.4361	12.4707
13	2	1	0	1	1	1	4	7.26E-01	-1.4096	85.9765	58.6933	-3.2286	12.519

The 1-step ahead performance for the best fitted Central Philadelphia lagged dynamic regression model is shown below. The first season is left out because of the lagged regressors. The model captures the center of the data and the peaks very well, but misses the low values often. The lagged regressors are not helping compared to the no lagged dynamic regression model.

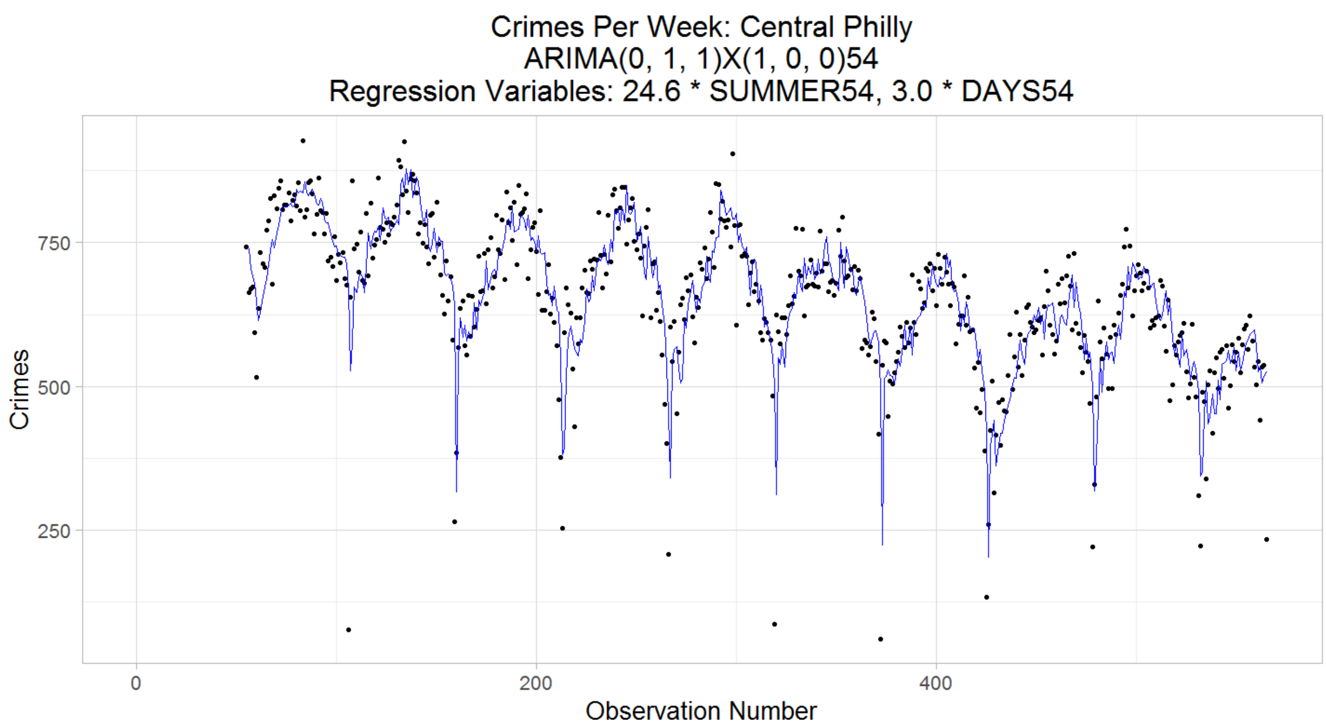


Figure 73: 10 Years of Central Philadelphia Weekly Crime & Dynamic Regression Model Fit - Lagged

The fit comparison of the best no lagged dynamic regression model and the best lagged dynamic regression model for Central Philadelphia is shown below. The no lagged model outperforms the lagged model across all independent variables. This may be due to the lagged regressors breaking the positive linear relationship with weekly crimes.

Table 27: Central Philadelphia Weekly Crime Fit Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
DR	6.53E-01	-1.1316	56.8011	44.7697	-0.7319	7.88243
DR.Lag	6.69E-01	-1.5608	77.8596	54.3154	-4.4361	12.4707

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days and Summer as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 21 models, one of which being the auto fit (Scenario 1), were tested.

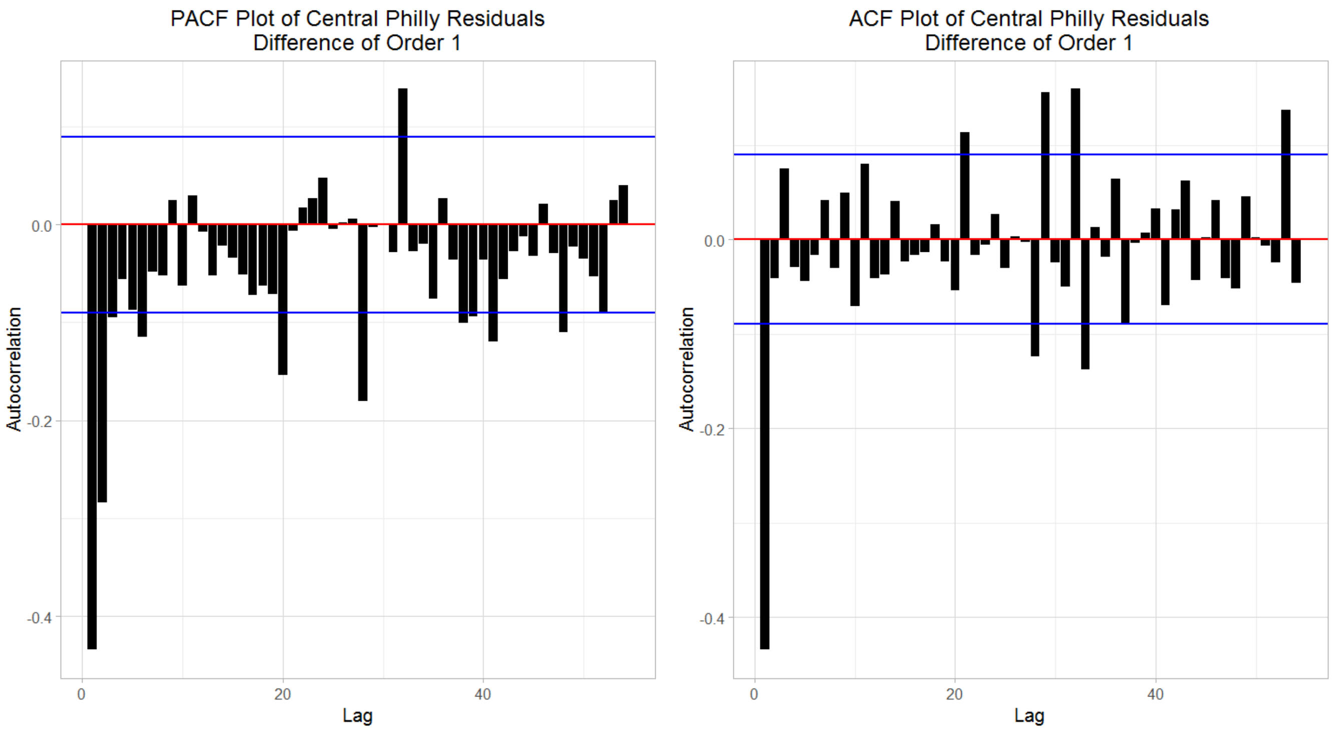


Figure 74: ACF & PACF Plot of Linear Regression Residuals for Central Philadelphia Dynamic Regression Forecasting

The numerical performance of the 2 Central Philadelphia forecasting models is shown below. Scenario 17 had better performance across most independent variables. Scenario 17 is the model that was eventually chosen as the best dynamic regression model for forecasting Central Philadelphia weekly crimes for 2015 – 2016.

Table 28: DOE Results for Forecasting Central Philadelphia Weekly Crimes 2015 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0	1	1	0	0	0	1	1.06E-03	27.8328	81.1016	63.6293	3.30395	11.1328
17	1	1	0	0	1	1	2	4.05E-01	7.33899	81.6045	64.4347	0.17805	11.8345

The forecast performance for the best Central Philadelphia dynamic regression model is shown below. The model under predicts the first half of 2015, over predicts the second half of 2015, and then predicts the center of 2016 well. The under and over predicting in 2015 is due to the 2015 peak starting earlier, compared to the 2014 season. The prediction intervals at 80% and 95% confidence levels, capture all of the actual values.

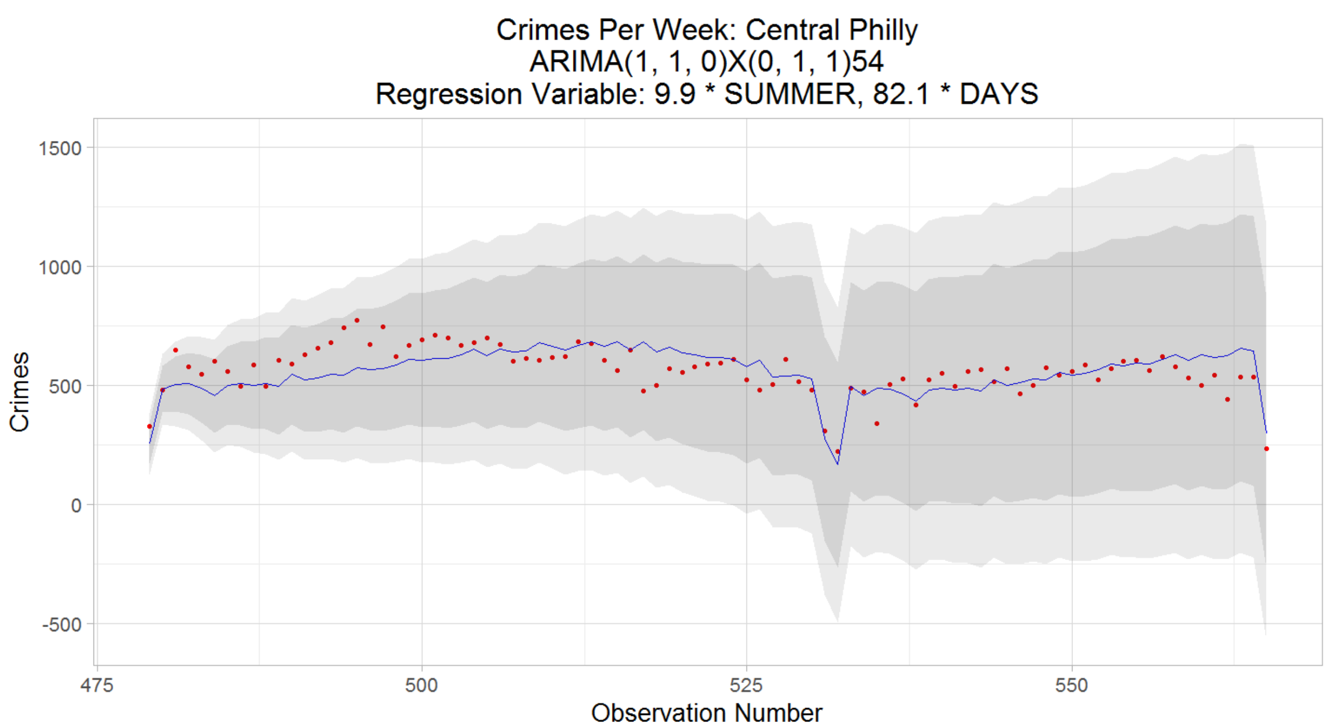


Figure 75: Prediction Plot for Central Philadelphia Weekly Crime Dynamic Regression Forecasting Model

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days54 and Summer54 as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 4, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 21 models, one of which being the auto fit (Scenario 1), were tested.

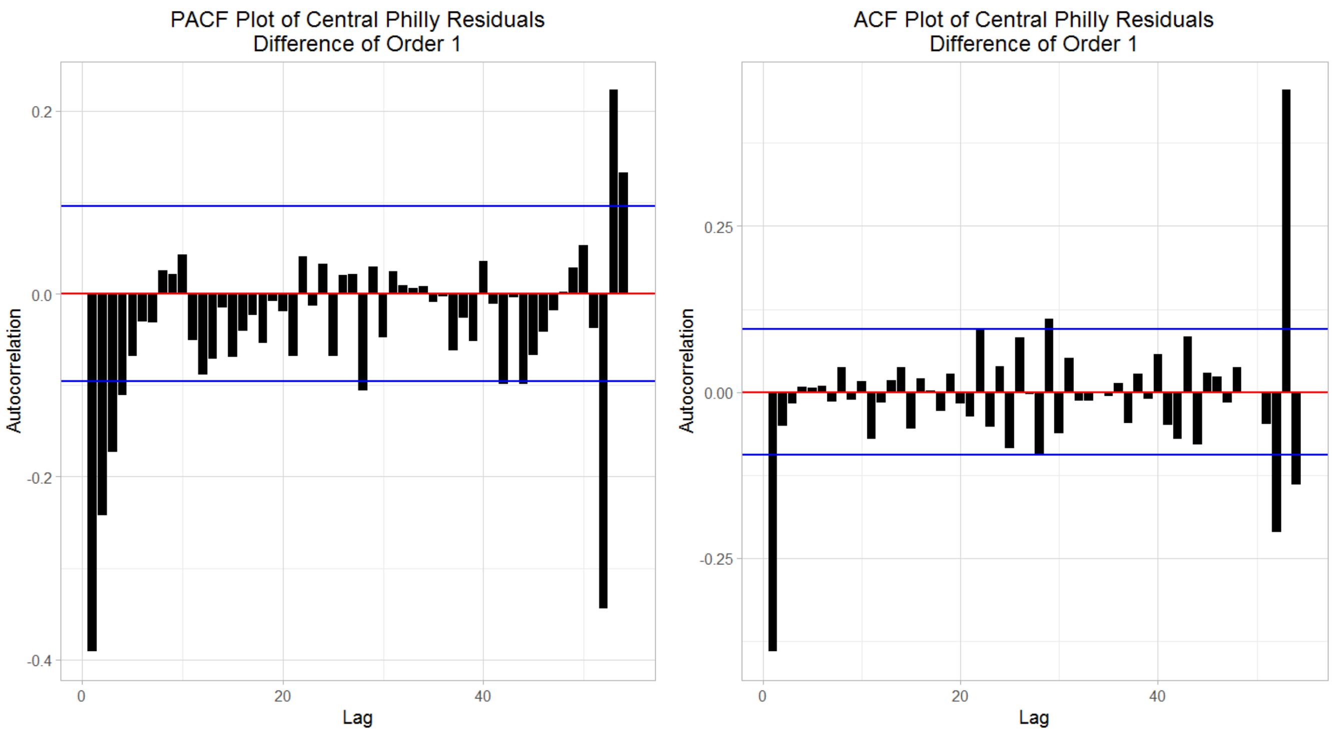


Figure 76: ACF & PACF Plot of Linear Regression Residuals for Central Philadelphia Dynamic Regression Forecasting - Lagged

The numerical performance of the 2 Central Philadelphia forecasting models is shown below. Scenario 15 had better performance across all independent variables. Scenario 15 is the model that was eventually chosen as the best lagged dynamic regression model for forecasting Central Philadelphia weekly crimes for 2015 – 2016.

Table 29: DOE Results for Forecasting Central Philadelphia Weekly Crimes 2015 – 2016 - Lagged

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	5	1	0	1	0	0	6	2.67E-21	120.007	148.866	133.429	18.4149	23.5205
15	0	1	1	1	1	1	3	7.23E-04	40.6128	114.814	89.5056	4.57841	17.2713

The forecast performance for the best Central Philadelphia lagged dynamic regression model is shown below. The model under predicts the first half of each season, and then predicts the center of the second half of each season well. The prediction intervals at 80% and 95% confidence levels, capture most of the actual values.

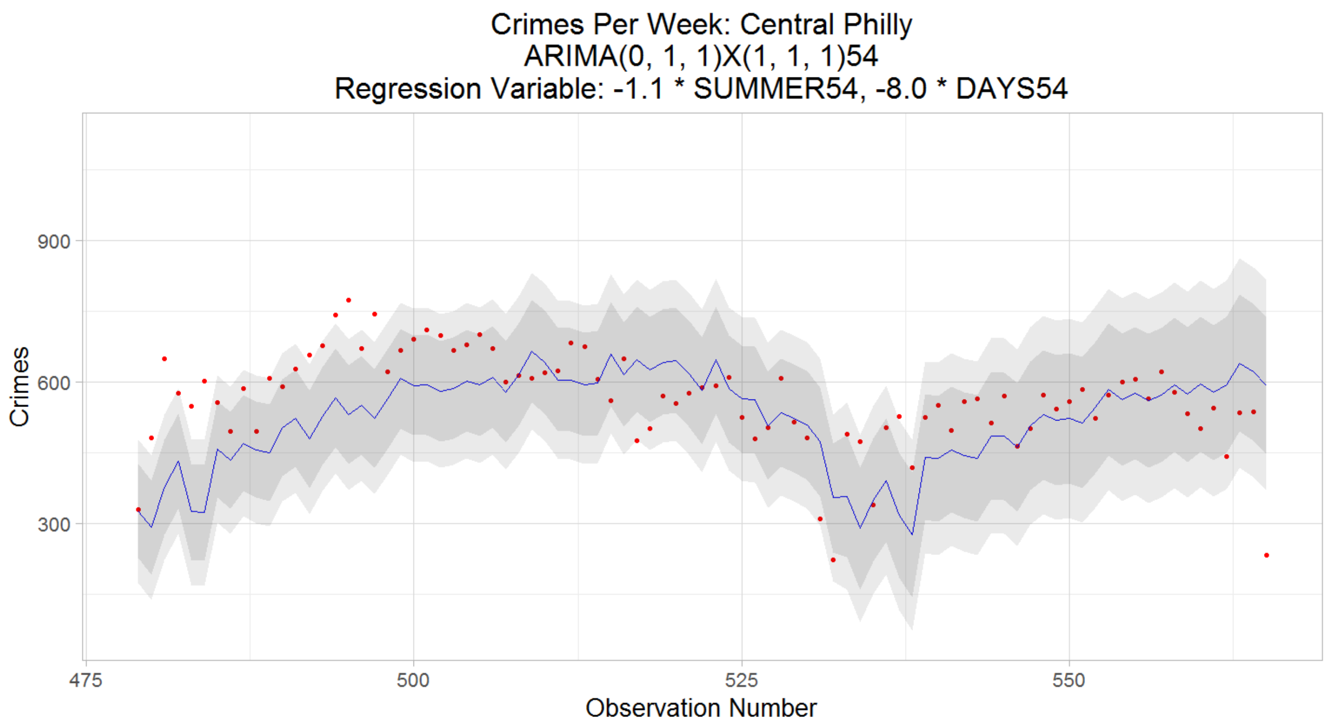


Figure 77: Prediction Plot for Central Philadelphia Weekly Crime Dynamic Regression Forecasting Model - Lagged

The forecasting comparison of the best forecasting model for Central Philadelphia 2015 – 2016 is shown below. The Holt-Winter model has the best performance across all independent variables. This may be due to the Holt-Winters function being able to compute many more models within the chosen 5 minute limit for model building. This efficiency allows for a larger sample of models to choose from, and results in the best model being found easier. The dynamic regression model captures the lowest values whereas the other models don't, but it misses the peaks of 2015 – 2016, resulting in worse performance.

Table 30: Central Philadelphia Weekly Crime Forecasting Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	9.16E-01	0.85826	74.9167	55.7464	-2.1416	11.5526
ARIMA	1.01E-06	57.4061	116.252	91.6005	7.80079	17.4541
DR	4.05E-01	7.33899	81.6045	64.4347	0.17805	11.8345
DR.Lag	7.23E-04	40.6128	114.814	89.5056	4.57841	17.2713

West Philadelphia

The need for these regressors was decided by creating dynamic regression autofits for all regressors and each individual regressor, and then comparing the fitness. The table below shows that the autofit using all regressors has better fitness. Therefore both regressors will be used in fitting and forecasting West Philadelphia weekly crimes.

Table 31: West Philadelphia Autofit Comparison Across Regressors

Variables	ME	RMSE	MAE	MPE	MAPE	AIC	AICc	BIC
SUMMER,DAYS	0.07073	59.1357	46.2776	-0.5979	7.66053	6218.88	6219.08	6249.22
SUMMER	0.72411	85.0743	57.9159	-4.7918	13.3908	6632.1	6632.31	6662.45
DAYS	-0.1864	59.2868	46.2724	-0.6187	7.66356	6219.68	6219.84	6245.7

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days and Summer as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 2, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 25 models, one of which being the auto fit (Scenario 1), were tested.

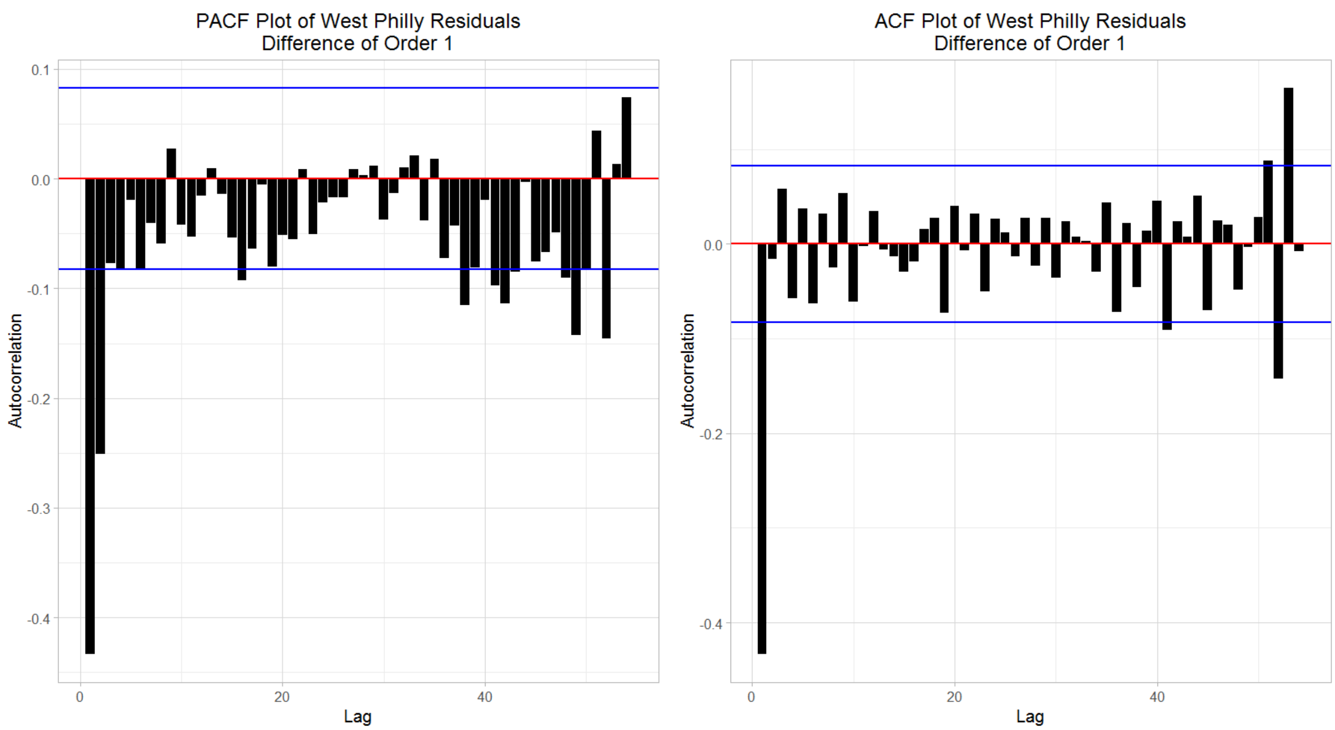


Figure 78: ACF & PACF Plot of Linear Regression Residuals for West Philadelphia Dynamic Regression Fitting

The numerical performance of the 2 West Philadelphia fitted models is shown below. Scenario 1 had slightly better performance across most independent variables. Scenario 16, with less parameters, is the model that was eventually chosen as the best dynamic regression model for fitting West Philadelphia weekly crimes for 2006 – 2016.

Table 32: DOE Results for Fitting West Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0	1	1	0	0	2	3	8.82E-01	-0.3943	59.9111	46.8186	-0.6931	7.84865
16	0	1	1	0	1	1	2	5.64E-01	-1.5515	60.783	47.0652	-0.4812	7.89742

The 1-step ahead performance for the best fitted West Philadelphia dynamic regression model is shown below. The model captures the center of the data, the peaks, and the low values very well. The positive coefficients for each regressor agree with the positive linear relationship seen in the violin plots. The regressors are helping, especially Days which captures the low values that previous models failed to capture.

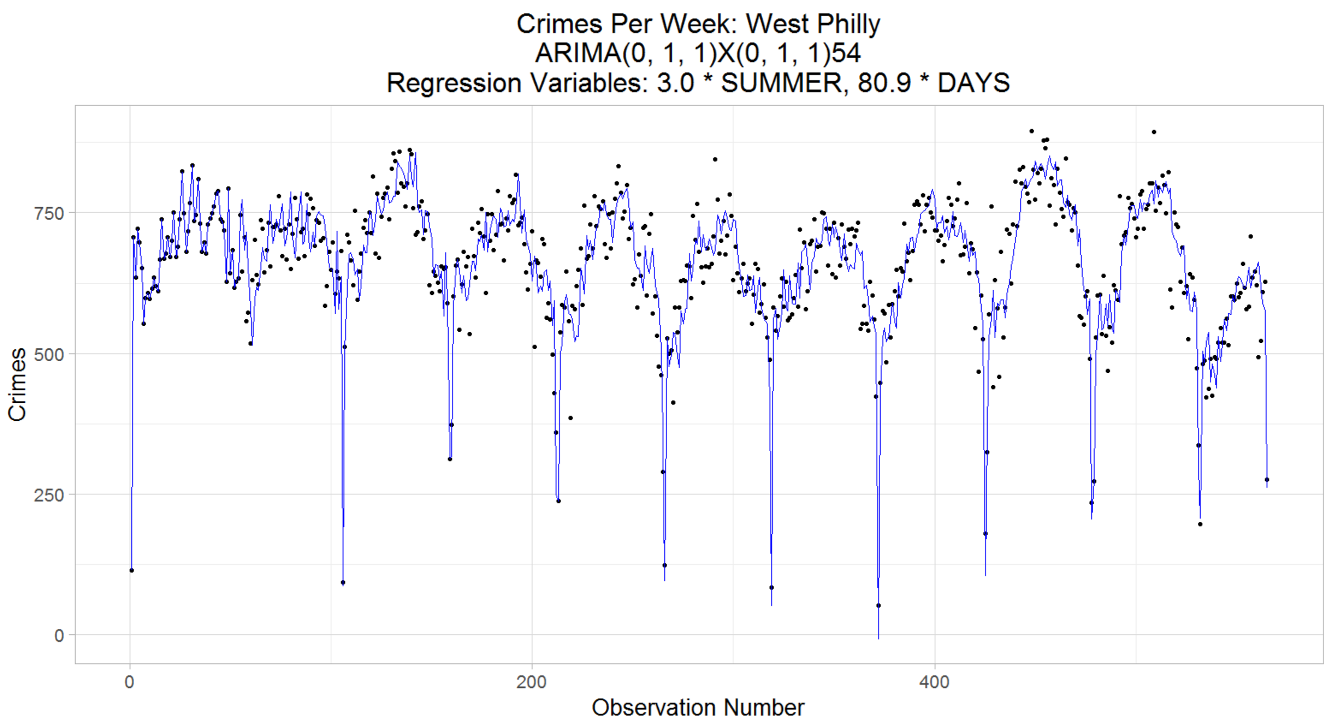


Figure 79: 10 Years of West Philadelphia Weekly Crime & Dynamic Regression Model Fit

The choice of potential lagged regressors were based on lags of 1 week, 4 weeks, and 54 weeks. These lags were chosen based on the weekly, monthly, and annual seasonality of the data as shown in earlier time series. The chosen lagged regressors were based on trying 9 possible combinations shown in the table below. Each row corresponds to the fitness of a linear regression model. The table indicates that lagging Days by 54 weeks and Summer by 1 week, scenario 7, is the better option.

Table 33: Results for Testing Lagged Regressors

Scenario	SUMMER	DAYS	r2pred	AIC	BIC
1	SUMMER1	DAYS1	0.1891	6251.14	6268.08
2	SUMMER4	DAYS1	0.1633	6267.22	6284.17
3	SUMMER54	DAYS1	0.187	6252.46	6269.41
4	SUMMER1	DAYS4	0.1321	6288.76	6305.71
5	SUMMER4	DAYS4	0.1049	6304.59	6321.54
6	SUMMER54	DAYS4	0.1299	6290.06	6307
7	SUMMER1	DAYS54	0.2014	6239.76	6256.71
8	SUMMER4	DAYS54	0.1758	6256.08	6273.03
9	SUMMER54	DAYS54	0.1993	6241.11	6258.05

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days54 and Summer1 as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 4, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 21 models, one of which being the auto fit (Scenario 1), were tested.

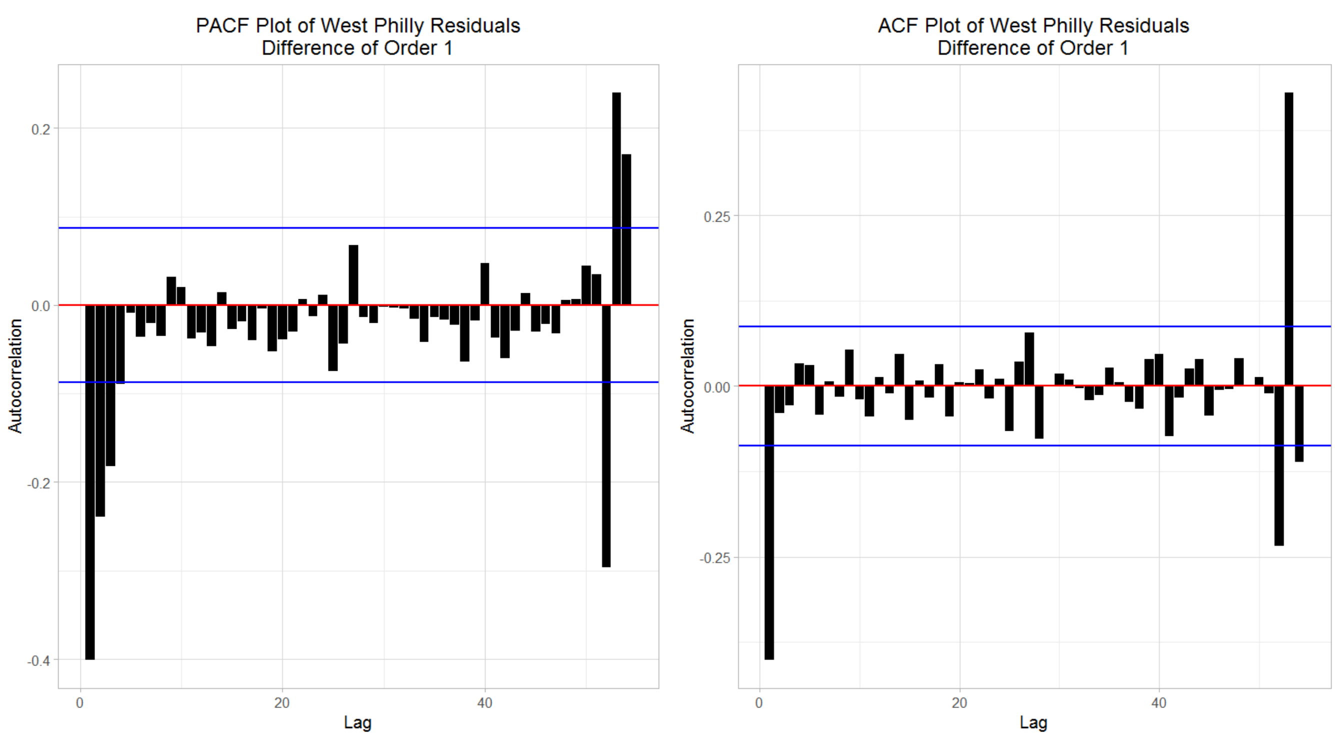


Figure 80: ACF & PACF Plot of Linear Regression Residuals for West Philadelphia Dynamic Regression Fitting - Lagged

The numerical performance of the 2 West Philadelphia fitted models is shown below. Scenario 1 had slightly better performance across most independent variables. Scenario 13, with less parameters, is the model that was eventually chosen as the best lagged dynamic model for fitting West Philadelphia weekly crimes for 2006 – 2016.

Table 34: DOE Results for Fitting West Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	5	1	0	1	0	0	6	8.58E-01	-0.6895	82.0394	57.8078	-4.6076	13.1999
13	3	1	0	1	1	1	5	8.37E-01	-0.8302	86.0306	59.8701	-3.6003	12.9665

The 1-step ahead performance for the best fitted West Philadelphia dynamic regression model is shown below. The first season is left out because of the lagged regressors. The model captures the center of the data, the peaks, and the low values well. The lagged regressors are not helping compared to the no lagged dynamic regression model.

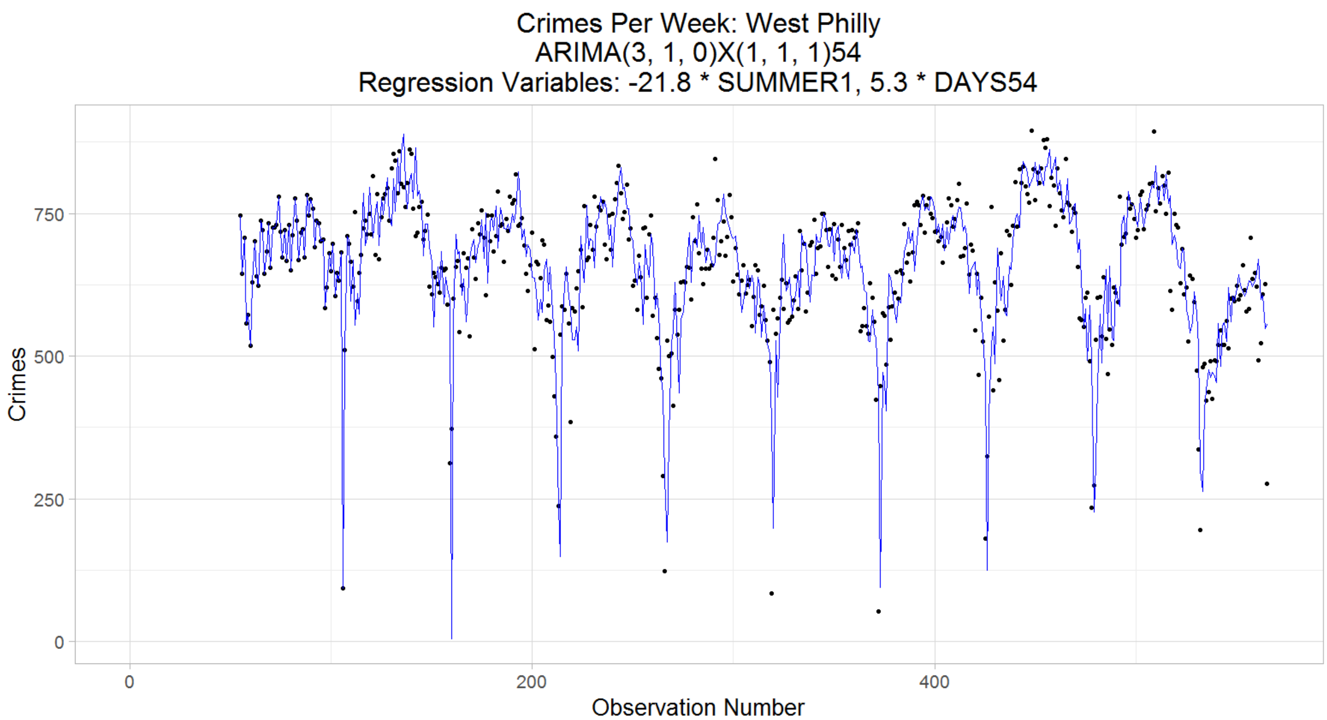


Figure 81: 10 Years of West Philadelphia Weekly Crime & Dynamic Regression Model Fit - Lagged

The fit comparison of the best no lagged dynamic regression model and the best lagged dynamic regression model for West Philadelphia is shown below. The no lagged model outperforms the lagged model across all independent variables. This may be due to the lagged regressors breaking the positive linear relationship with weekly crimes.

Table 35: West Philadelphia Weekly Crime Fit Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
DR	5.64E-01	-1.5515	60.783	47.0652	-0.4812	7.89742
DR.Lag	8.37E-01	-0.8302	86.0306	59.8701	-3.6003	12.9665

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days and Summer as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 2, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 25 models, one of which being the auto fit (Scenario 1), were tested.

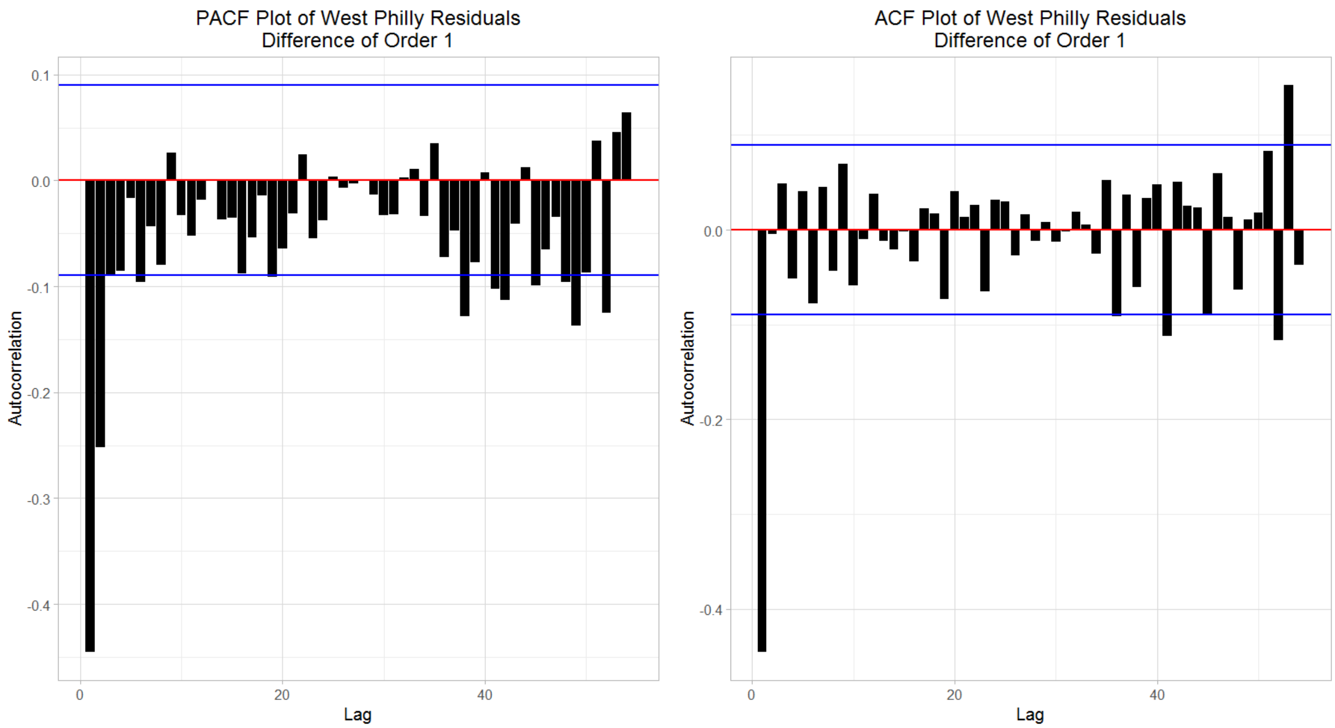


Figure 82: ACF & PACF Plot of Linear Regression Residuals for West Philadelphia Dynamic Regression Forecasting

The numerical performance of the 2 West Philadelphia forecasting models is shown below. Scenario 17 had better performance across all independent variables. Scenario 17 is the model that was eventually chosen as the best dynamic regression model for forecasting West Philadelphia weekly crimes for 2015 – 2016.

Table 36: DOE Results for Forecasting West Philadelphia Weekly Crimes 2015 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0	1	1	0	0	0	1	1.36E-07	65.2827	123.97	100.12	7.61945	15.2635
16	0	1	1	0	1	1	2	4.69E-01	5.96308	76.3293	61.7472	-0.7924	10.0733

The forecast performance for the best West Philadelphia dynamic regression model is shown below. The model misses the peak of 2015, and over predicts 2016. The prediction intervals at 80% and 95% confidence levels, capture almost all of the actual values.

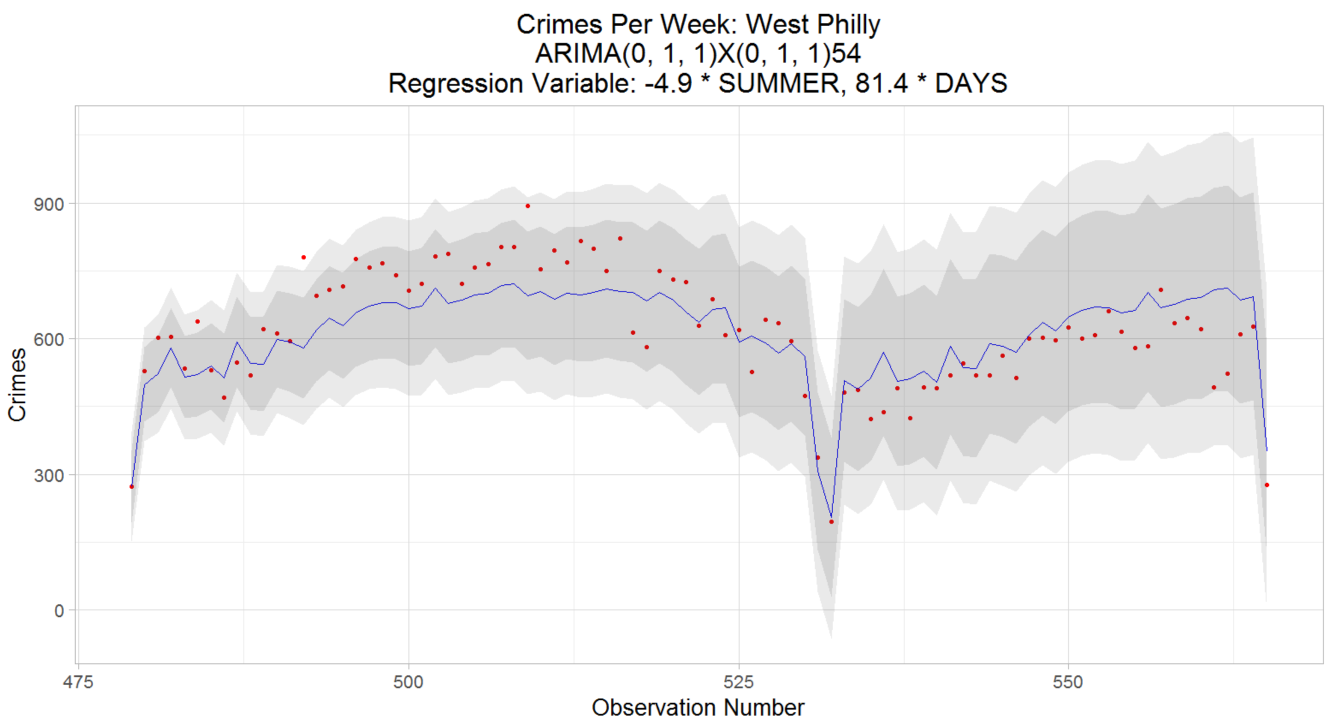


Figure 83: Prediction Plot for West Philadelphia Weekly Crime Dynamic Regression Forecasting Model

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days54 and Summer1 as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 17 models, one of which being the auto fit (Scenario 1), were tested.

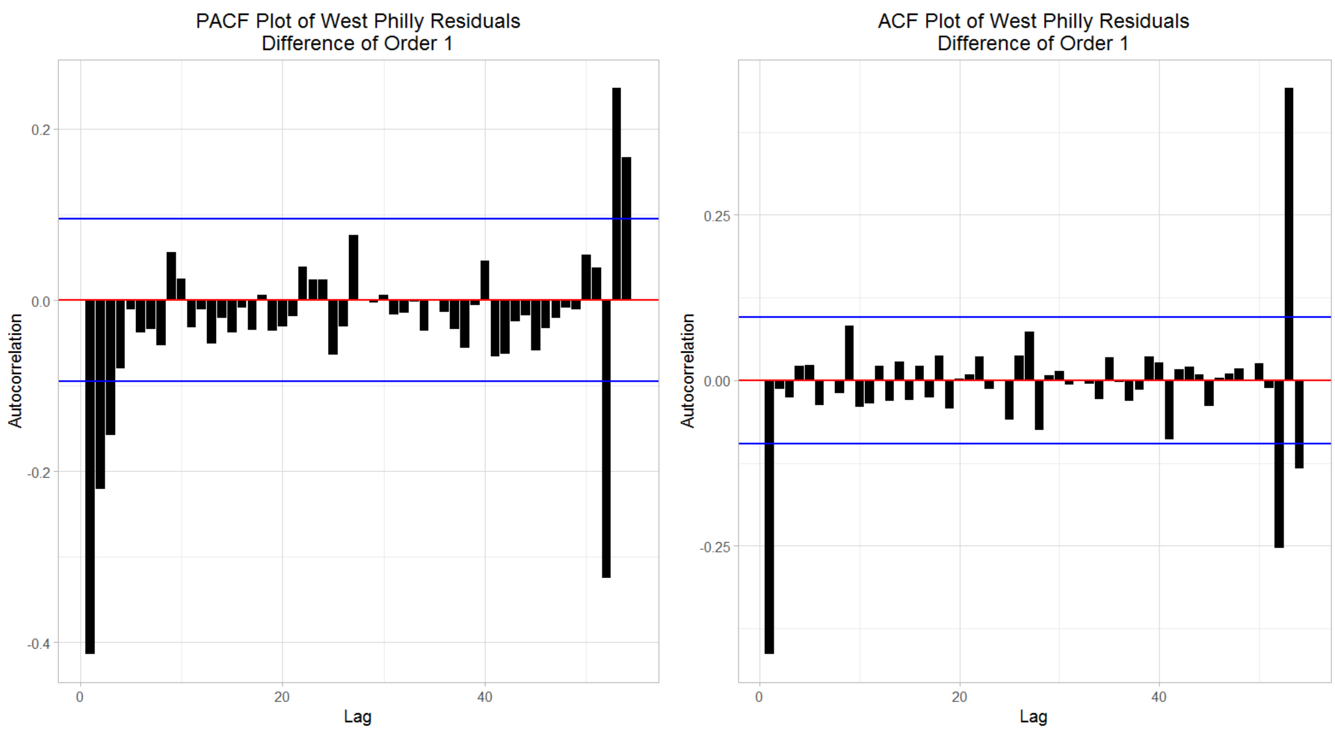


Figure 84: ACF & PACF Plot of Linear Regression Residuals for West Philadelphia Dynamic Regression Forecasting - Lagged

The numerical performance of the 2 West Philadelphia forecasting models is shown below. Scenario 13 had better performance across all independent variables. Scenario 13 is the model that was eventually chosen as the best lagged dynamic regression model for forecasting West Philadelphia weekly crimes for 2015 – 2016.

Table 37: DOE Results for Forecasting West Philadelphia Weekly Crimes 2015 – 2016 - Lagged

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	5	1	0	1	0	0	6	1.75E-27	168.423	194.717	177.548	24.2298	27.8853
13	0	1	1	1	1	1	3	1.79E-01	-13.639	94.3102	68.5543	-5.0163	13.4625

The forecast performance for the best West Philadelphia lagged dynamic regression model is shown below. The model predicts 2015 well, and then over predicts 2016. The prediction intervals at 80% and 95% confidence levels, capture almost all of the actual values.

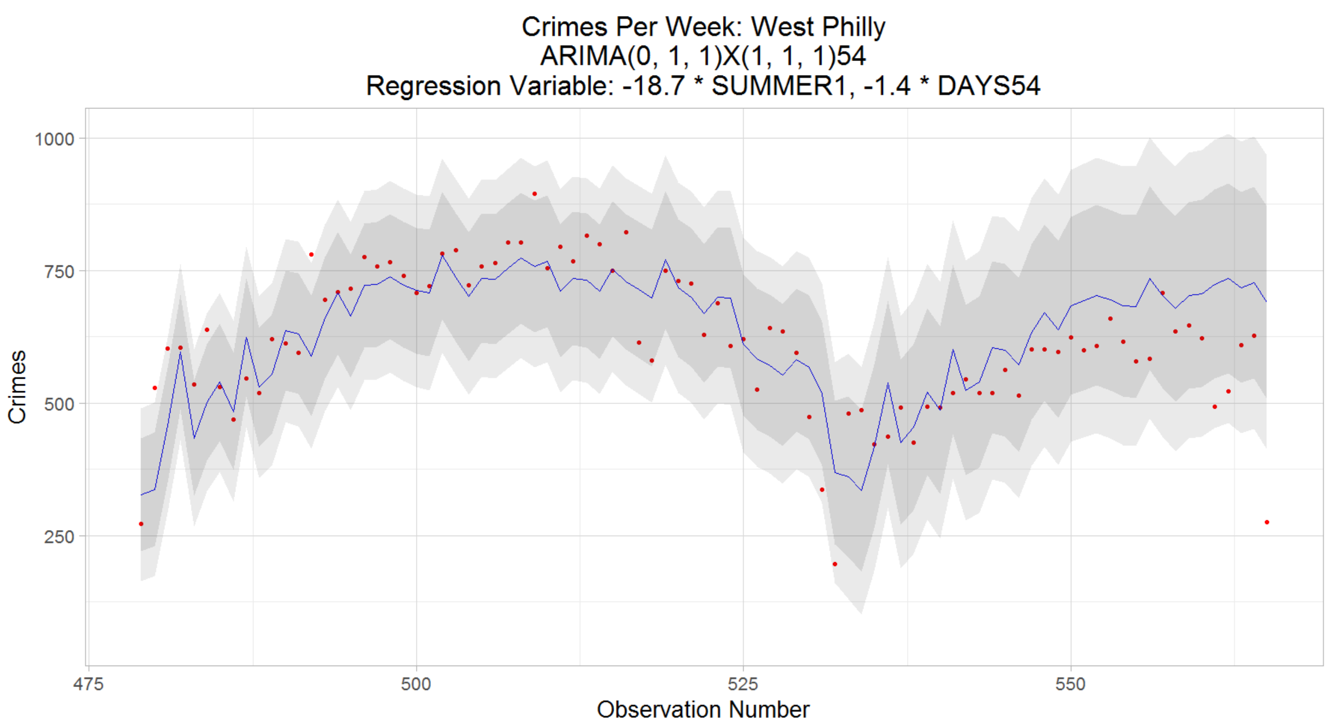


Figure 85: Prediction Plot for West Philadelphia Weekly Crime Dynamic Regression Forecasting Model - Lagged

The forecasting comparison of the best forecasting model for West Philadelphia 2015 – 2016 is shown below. The Holt-Winter model and dynamic regression model compete for the best performance across all independent variables. The dynamic regression model achieves better MAPE and RMSE due to the regressors reaching the lower values, and the Holt-Winter model achieves a better t.pval and MAE due to following the center of the data better. Overall, the dynamic regression model is better given that it won't over predict lower values as severe as Holt-Winter.

Table 38: West Philadelphia Weekly Crime Forecasting Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	9.63E-01	-0.4019	80.5195	58.8806	-2.8648	11.8654
ARIMA	6.01E-01	-5.6713	100.412	75.5358	-4.0598	14.6643
DR	4.69E-01	5.96308	76.3293	61.7472	-0.7924	10.0733
DR.Lag	1.79E-01	-13.639	94.3102	68.5543	-5.0163	13.4625

North Philadelphia

The need for these regressors was decided by creating dynamic regression autofits for all regressors and each individual regressor, and then comparing the fitness. The table below shows that the autofit using all regressors has better fitness. Therefore both regressors will be used in fitting and forecasting North Philadelphia weekly crimes.

Table 39: North Philadelphia Autofit Comparison Across Regressors

Variables	ME	RMSE	MAE	MPE	MAPE	AIC	AICc	BIC
SUMMER,DAYS	-0.7265	125.533	94.4181	-0.3446	6.07086	7071.83	7072.15	7110.84
SUMMER	1.70274	211.735	131.182	-4.4929	11.8424	7657.4	7657.5	7679.07
DAYS	-0.5863	125.709	94.6318	-0.3388	6.0841	7071.36	7071.62	7106.04

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days and Summer as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 33 models, one of which being the auto fit (Scenario 1), were tested.

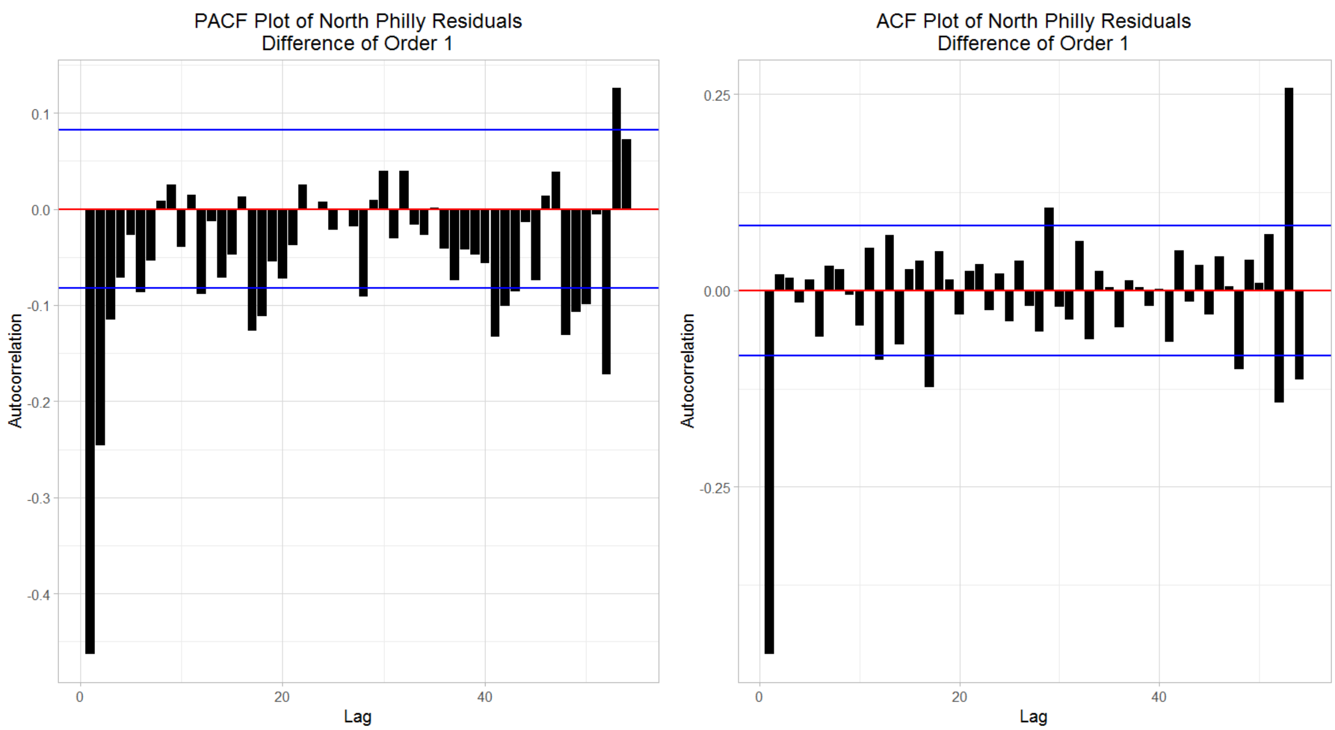


Figure 86: ACF & PACF Plot of Linear Regression Residuals for North Philadelphia Dynamic Regression Fitting

The numerical performance of the 2 North Philadelphia fitted models is shown below. Scenario 1 had slightly better performance across most independent variables. Scenario 17, with less parameters, is the model that was eventually chosen as the best dynamic regression model for fitting North Philadelphia weekly crimes for 2006 – 2016.

Table 40: DOE Results for Fitting North Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	2	1	3	1	0	0	6	7.33E-01	-1.879	124.45	93.22	-0.4021	6.16831
17	2	1	0	0	1	1	3	7.59E-01	-1.7354	127.733	98.7859	0.27499	6.6103

The 1-step ahead performance for the best fitted North Philadelphia dynamic regression model is shown below. The model captures the center of the data, the peaks, and the low values very well. The positive coefficients for each regressor agree with the positive linear relationship seen in the violin plots. The regressors are helping, especially Days which captures the low values that previous models failed to capture.

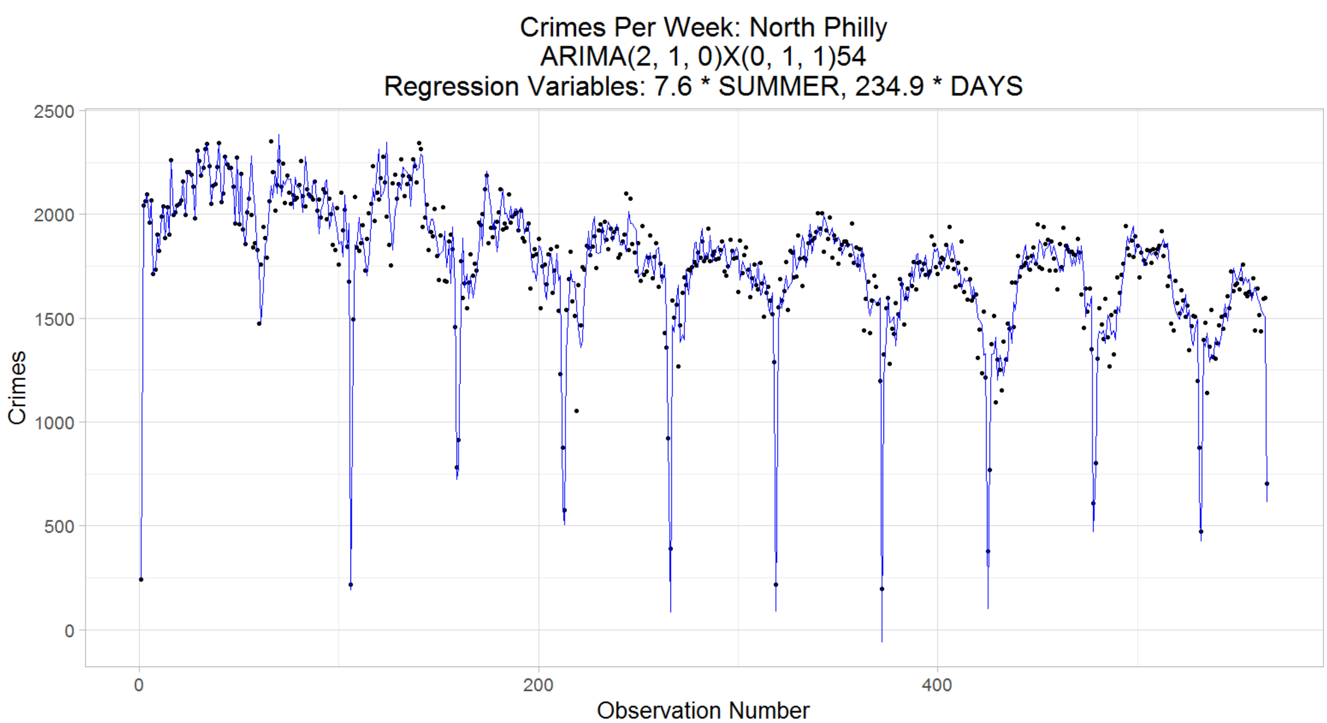


Figure 87: 10 Years of North Philadelphia Weekly Crime & Dynamic Regression Model Fit

The choice of potential lagged regressors were based on lags of 1 week, 4 weeks, and 54 weeks. These lags were chosen based on the weekly, monthly, and annual seasonality of the data as shown in earlier time series. The chosen lagged regressors were based on trying 9 possible combinations shown in the table below. Each row corresponds to the fitness of a linear regression model. The table indicates that lagging Days by 54 weeks and Summer by 1 week, scenario 7, is the better option.

Table 41: Results for Testing Lagged Regressors

Scenario	SUMMER	DAYS	r2pred	AIC	BIC
1	SUMMER1	DAYS1	0.1235	7223.5	7240.45
2	SUMMER4	DAYS1	0.1063	7233.54	7250.48
3	SUMMER54	DAYS1	0.1222	7224.25	7241.19
4	SUMMER1	DAYS4	0.0719	7256.97	7273.91
5	SUMMER4	DAYS4	0.0534	7267.13	7284.07
6	SUMMER54	DAYS4	0.0705	7257.71	7274.66
7	SUMMER1	DAYS54	0.1424	7208.41	7225.35
8	SUMMER4	DAYS54	0.1255	7218.57	7235.51
9	SUMMER54	DAYS54	0.1412	7209.16	7226.11

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days54 and Summer1 as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 5, and a moving average order up to 2 should be tested. Due to computation time, an autoregressive order up to 4, and a moving average order up to 1 was tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 21 models, one of which being the auto fit (Scenario 1), were tested.

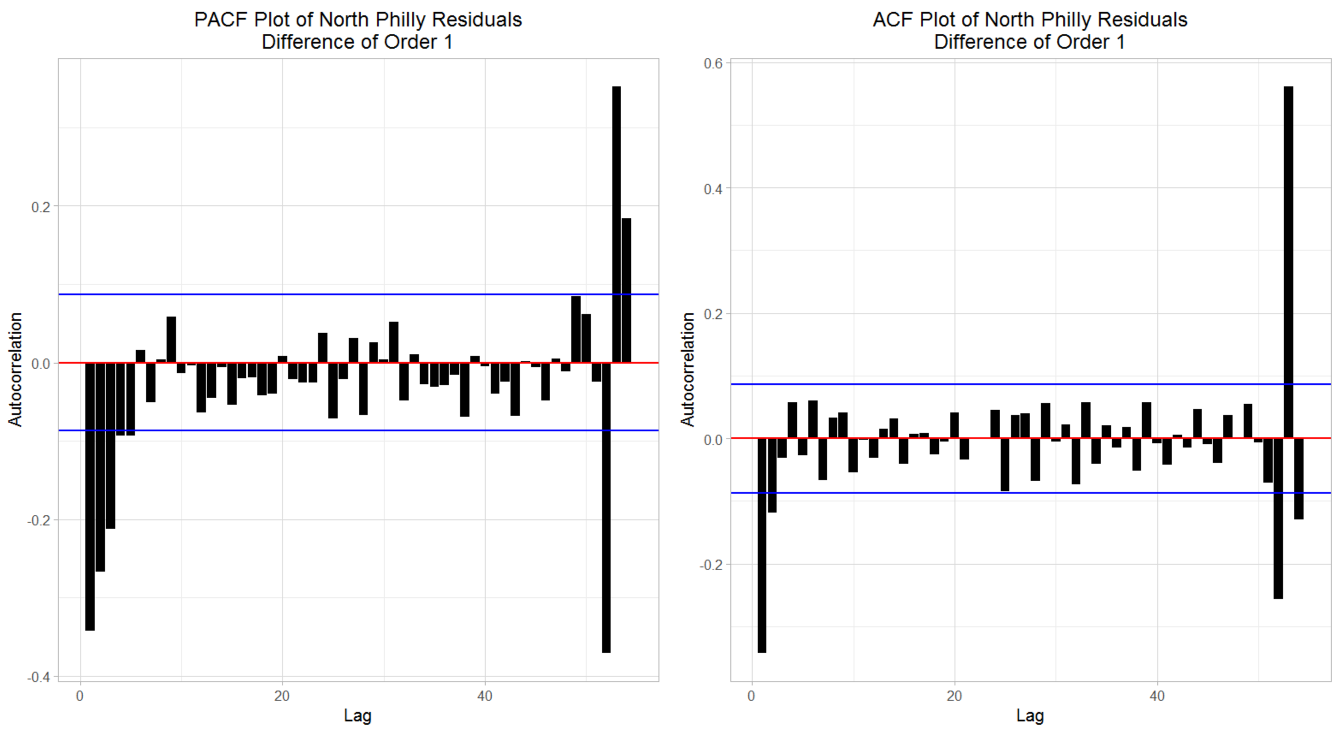


Figure 88: ACF & PACF Plot of Linear Regression Residuals for North Philadelphia Dynamic Regression Fitting - Lagged

The numerical performance of the 2 North Philadelphia fitted models is shown below. Scenario 1 had slightly better performance across most independent variables. Scenario 17, with less parameters, is the model that was eventually chosen as the best lagged dynamic model for fitting North Philadelphia weekly crimes for 2006 – 2016.

Table 42: DOE Results for Fitting North Philadelphia Weekly Crimes 2006 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	1	1	1	0	0	2	4	7.68E-01	-2.6868	194.291	125.336	-4.1021	11.2467
17	0	1	1	1	1	1	3	5.94E-01	-5.0096	200.339	130.914	-3.5023	11.1123

The 1-step ahead performance for the best fitted North Philadelphia dynamic regression model is shown below. The first season is left out because of the lagged regressors. The model captures the center of the data and the peaks very well, but misses the low values often. The lagged regressors are not helping compared to the no lagged dynamic regression model.

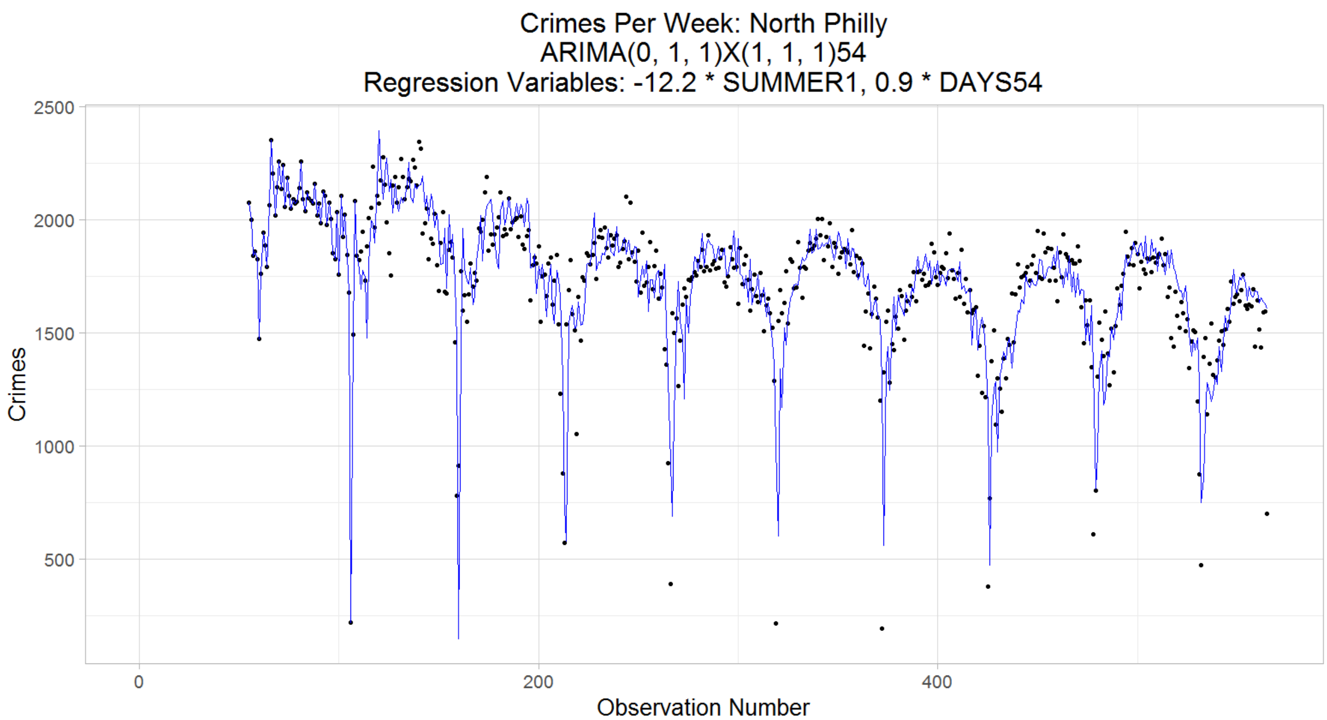


Figure 89: 10 Years of North Philadelphia Weekly Crime & Dynamic Regression Model Fit - Lagged

The fit comparison of the best no lagged dynamic regression model and the best lagged dynamic regression model for North Philadelphia is shown below. The no lagged model outperforms the lagged model across all independent variables. This may be due to the lagged regressors breaking the positive linear relationship with weekly crimes.

Table 43: North Philadelphia Weekly Crime Fit Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
DR	7.59E-01	-1.7354	127.733	98.7859	0.27499	6.6103
DR.Lag	5.94E-01	-5.0096	200.339	130.914	-3.5023	11.1123

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days and Summer as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 33 models, one of which being the auto fit (Scenario 1), were tested.

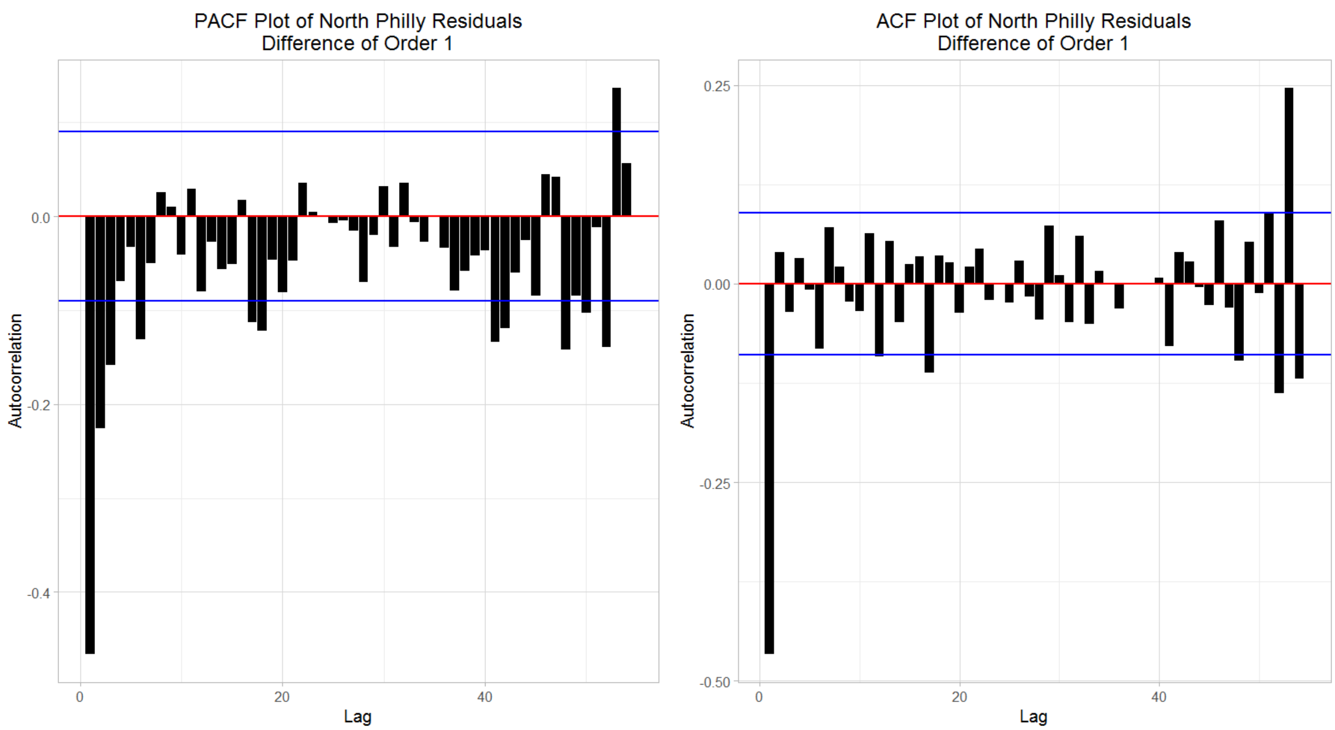


Figure 90: ACF & PACF Plot of Linear Regression Residuals for North Philadelphia Dynamic Regression Forecasting

The numerical performance of the 2 North Philadelphia forecasting models is shown below. Scenario 17 had better performance across all independent variables. Scenario 17 is the model that was eventually chosen as the best dynamic regression model for forecasting North Philadelphia weekly crimes for 2015 – 2016.

Table 44: DOE Results for Forecasting North Philadelphia Weekly Crimes 2015 - 2016

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0	1	1	1	0	0	2	3.50E-03	54.9814	178.466	145.86	2.21255	9.36633
20	1	1	1	0	1	1	3	7.63E-01	-4.3528	133.573	105.986	-1.096	7.02422

The forecast performance for the best North Philadelphia dynamic regression model is shown below. The model predicts 2015 – 2016 well, capturing the low values very well. The prediction intervals at 80% and 95% confidence levels, capture all of the actual values.

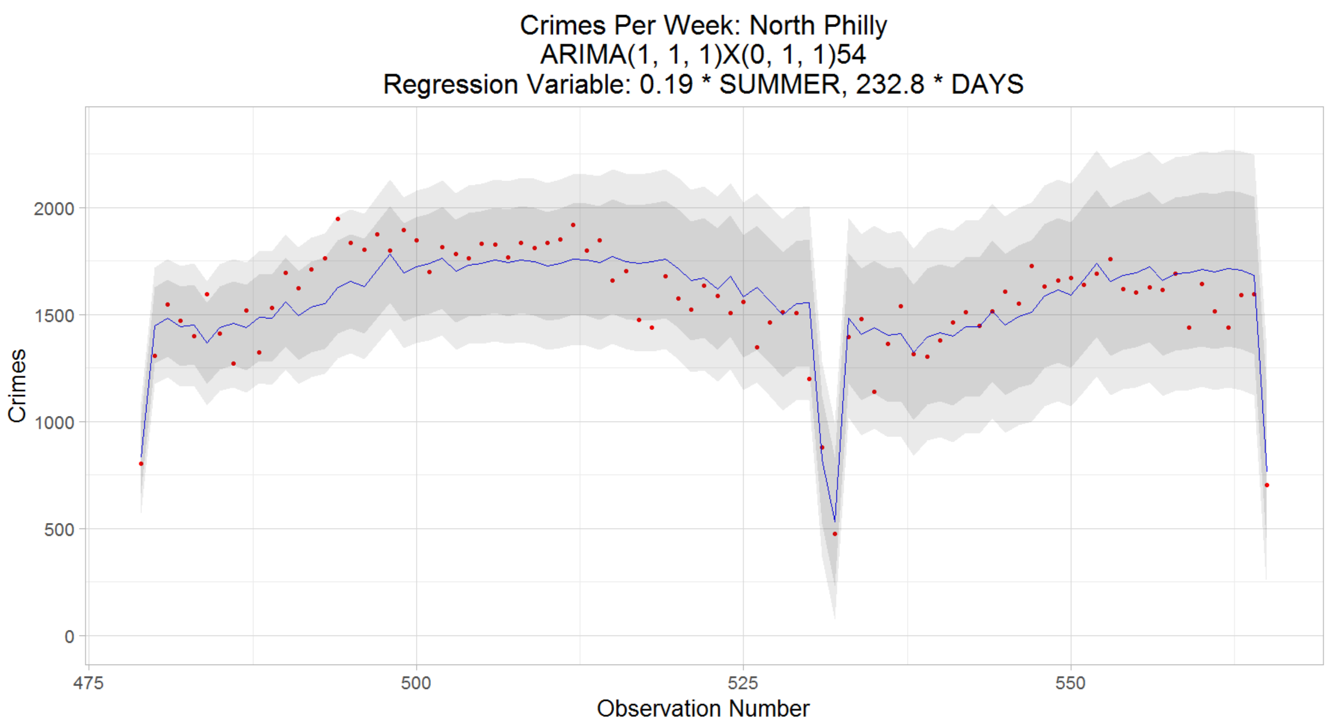


Figure 91: Prediction Plot for North Philadelphia Weekly Crime Dynamic Regression Forecasting Model

The following ACF and PACF plots are of the residuals from fitting a linear regression model of Days54 and Summer1 as regressors and weekly crimes as the response variable. The bars in these charts indicate that an autoregressive order up to 3, and a moving average order up to 1 will be tested. These ranges will be tested because the addition of seasonal parameters (P, D, Q) is only expected to reduce the (p, q) order, if at all. A total of 17 models, one of which being the auto fit (Scenario 1), were tested.

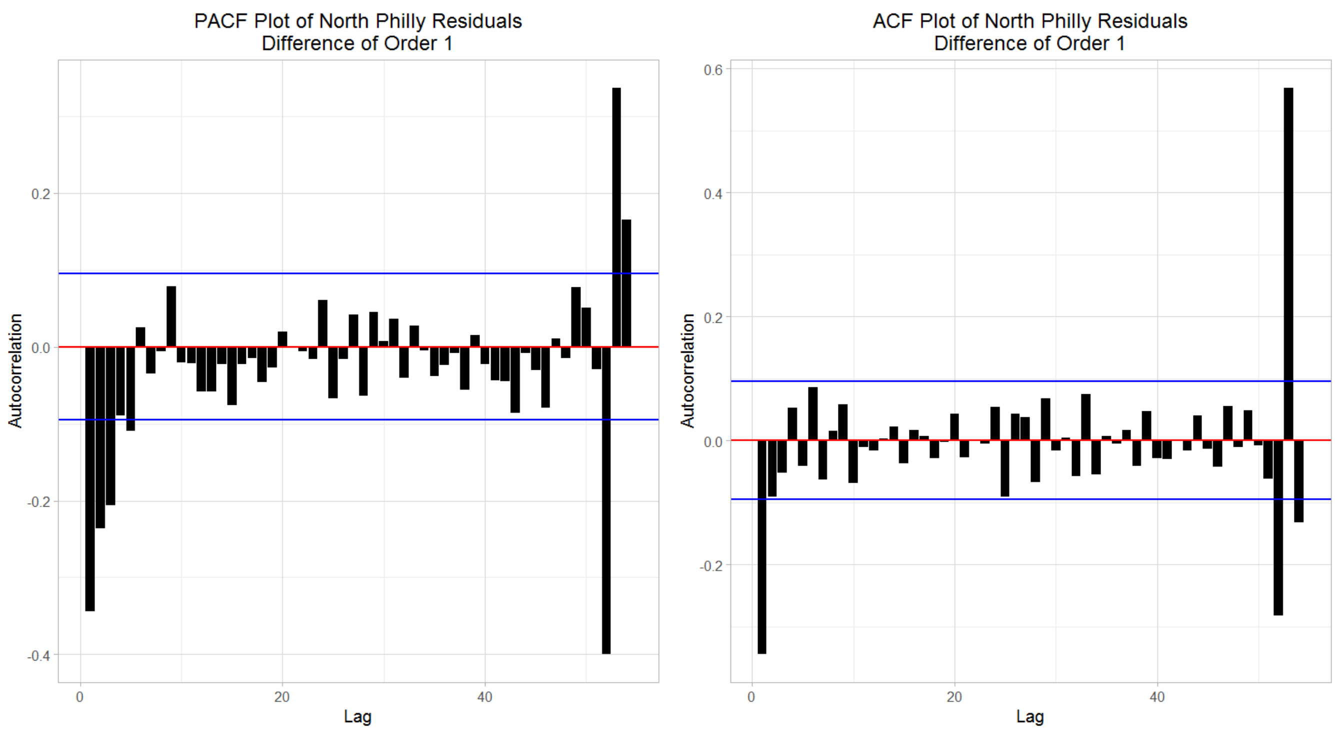


Figure 92: ACF & PACF Plot of Linear Regression Residuals for North Philadelphia Dynamic Regression Forecasting - Lagged

The numerical performance of the 2 North Philadelphia forecasting models is shown below. Scenario 13 had better performance across all independent variables. Scenario 13 is the model that was eventually chosen as the best lagged dynamic regression model for forecasting North Philadelphia weekly crimes for 2015 – 2016.

Table 45: DOE Results for Forecasting North Philadelphia Weekly Crimes 2015 – 2016 - Lagged

Scenario	p	d	q	P	D	Q	params	t.pval	ME	RMSE	MAE	MPE	MAPE
1	0	1	2	1	0	0	3	1.04E-09	148.645	250.139	210.63	6.51552	14.7749
13	0	1	1	1	1	1	3	3.60E-01	-22.613	228.777	160.81	-3.8451	12.6405

The forecast performance for the best North Philadelphia lagged dynamic regression model is shown below. The model predicts through the peak well for each season, and then over predicts the end of each season. The prediction intervals at 80% and 95% confidence levels, capture most of the actual values.

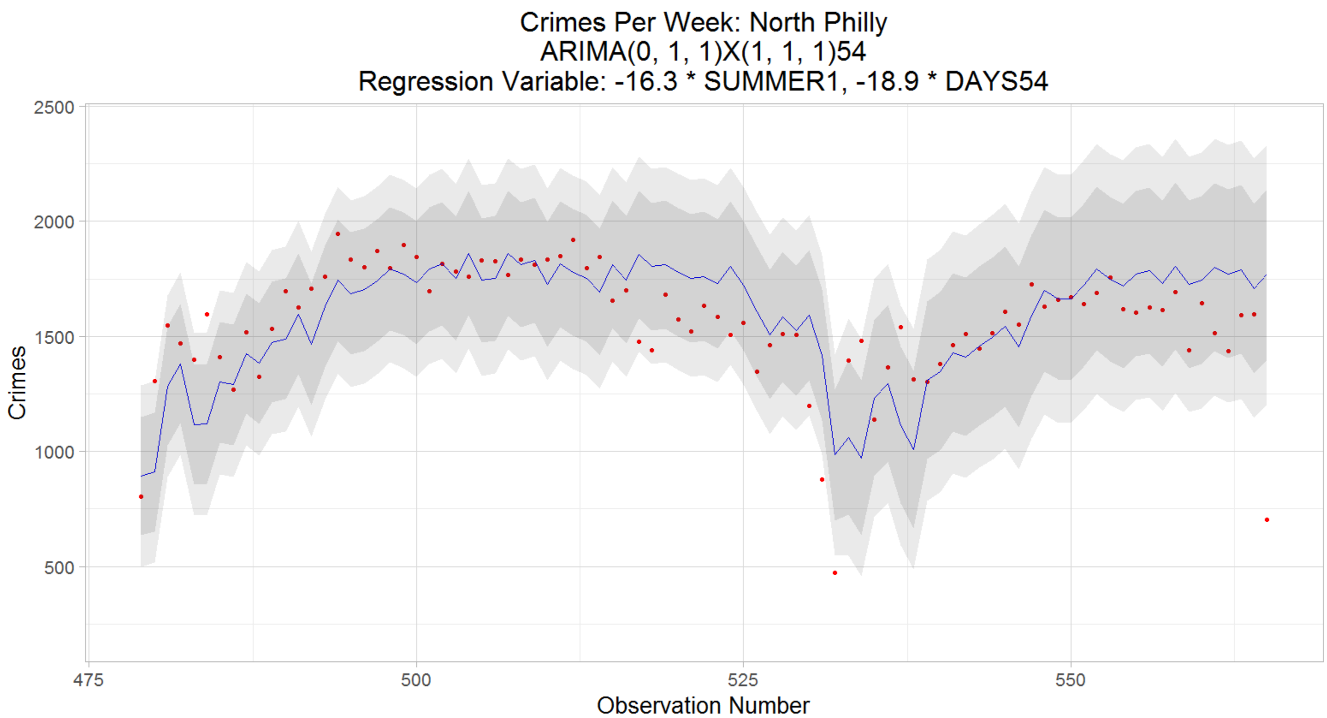


Figure 93: Prediction Plot for North Philadelphia Weekly Crime Dynamic Regression Forecasting Model - Lagged

The forecasting comparison of the best forecasting model for North Philadelphia 2015 – 2016 is shown below. The dynamic regression model has the best performance across all independent variables. The dynamic regression model achieves better performance due to the regressors reaching the lower values.

Table 46: North Philadelphia Weekly Crime Forecasting Comparison

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	8.86E-01	-2.8834	185.661	121.258	-2.4663	9.86163
ARIMA	2.22E-01	29.1673	221.682	152.684	-0.6937	12.105
DR	7.63E-01	-4.3528	133.573	105.986	-1.096	7.02422
DR.Lag	3.60E-01	-22.613	228.777	160.81	-3.8451	12.6405

The chosen regressors for the dynamic regression models enabled better predictions. This was mostly due to the regressor Days capturing the low values, while the regressor Summer didn't help as much with capturing the peaks. Other regressors regarding police staffing, police funding, Philadelphia economy, and Philadelphia prison population could be promising to further improve dynamic regression predictions.

Neural Network Forecasting

The process for choosing the best neural network fit and forecast models for Central, West, and North Philadelphia is similar to the processes used in the previous forecasting methods. There are three parameters of interest to vary: size (ie. the number of single layer hidden nodes), decay (ie. the penalty parameter for node weights), and maxit (ie. the maximum number of back propagation iterations to estimate node weights). Then the scenarios are evaluated graphically to filter out the best combination of size, decay, and maxit. The forecasting time period will be 2015 – 2016 to allow for comparison between all forecasting methods evaluated thus far. The time period of 2015 – 2016 is chosen also because it contains one season and the most recent data. This allows for each forecasting method to show how well it can capture an entire season, which is important because the data is seasonal. This also allows for the forecasting methods to capture the most recent crime behavior which is important because predicting crime in the near future will determine how useful these forecasting methods are for Philadelphia law enforcement. In this section, the best fits and forecasts from each forecasting method for Central, West, and North Philadelphia will be compared to ultimately determine the best descriptive model and the best predictive model for each of the three areas.

There were more regressors found for building a neural network model. Initially the regressors that were searched for included anything related to police staffing, police funding, Philadelphia economy, and Philadelphia prison population. The Philadelphia information that could be found publically within the time range of 2006 – 2016 was the number of unemployed people each month, the average electricity price per kWh each month, the average piped utility gas per therm each month, and the average gasoline price per gallon each month. This information was found on the official department of labor website. These regressors are evaluated below in Figure 94 for multicollinearity. The concerning pair of regressors are UNEMPLOYMENT and ELECTRICITY, with a decent positive linear relationship at a correlation value of 0.634.

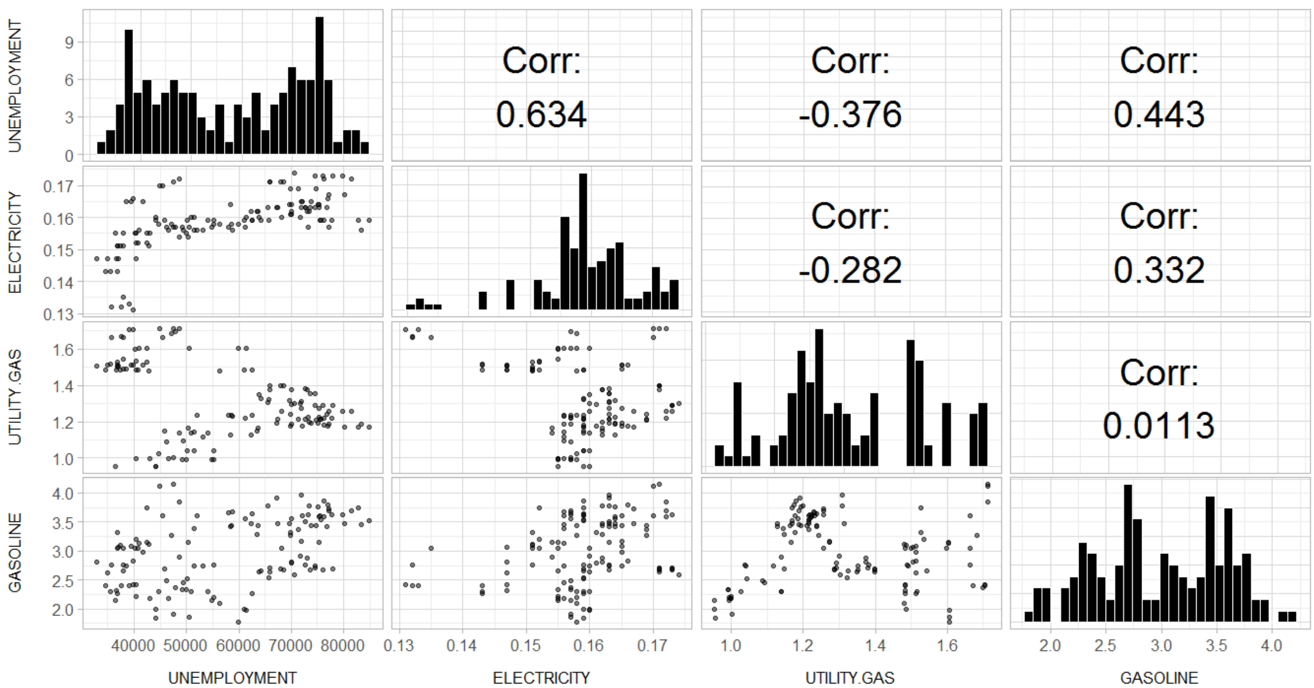


Figure 94: Pairs Plot of Potential Neural Network Regressors

There were two more regressors included based on their significance seen in previous models and plots. These regressors are the number of days in a week, and the month that a week is in. Given that a week can be in multiple months, a week is assigned the month that it first appeared in. Due to these two regressors being significant, they are included in the model, whereas the new four regressors were tested for significance. They were tested by building simple neural network models using parameter values (size, decay, maxit) = (5, 0, 1000). These models were built additively, and the procedure is as follows:

1. Build a base model of DAYS and MONTHS as regressors.
2. Build four models where each model includes DAYS, MONTHS, and one of the four new regressors.
3. Compare the four models to the base model and if a significant improvement in MAPE is shown by adding a particular regressor, then that regressor will be used, and that model becomes the new base model.
4. Repeat this procedure until adding one new regressor doesn't significantly improve MAPE.

The final set of regressors that were chosen for neural network modeling were: MONTH, DAYS, UNEMPLOYMENT, UTILITY.GAS.

Central Philadelphia

There were 125 fitted models built by varying size, decay, and maxit. The MAPE of each of those models is plotted below in Figure 95 where size is along the x-axis, MAPE is along the y-axis, maxit is the color of each data point, and decay is the value in the plot header. This figure indicates that to decrease MAPE, decay should equal 0, maxit should equal 2000, and size could equal anything between 20 and 50. The combination that was ultimately chosen was (size, decay, maxit) = (30, 0, 2000). The size value of 30 was chosen because it is not close to the behavior of a size value of 10, and adding up to 20 more nodes for a size value of 50 doesn't produce any significant improvement in MAPE.

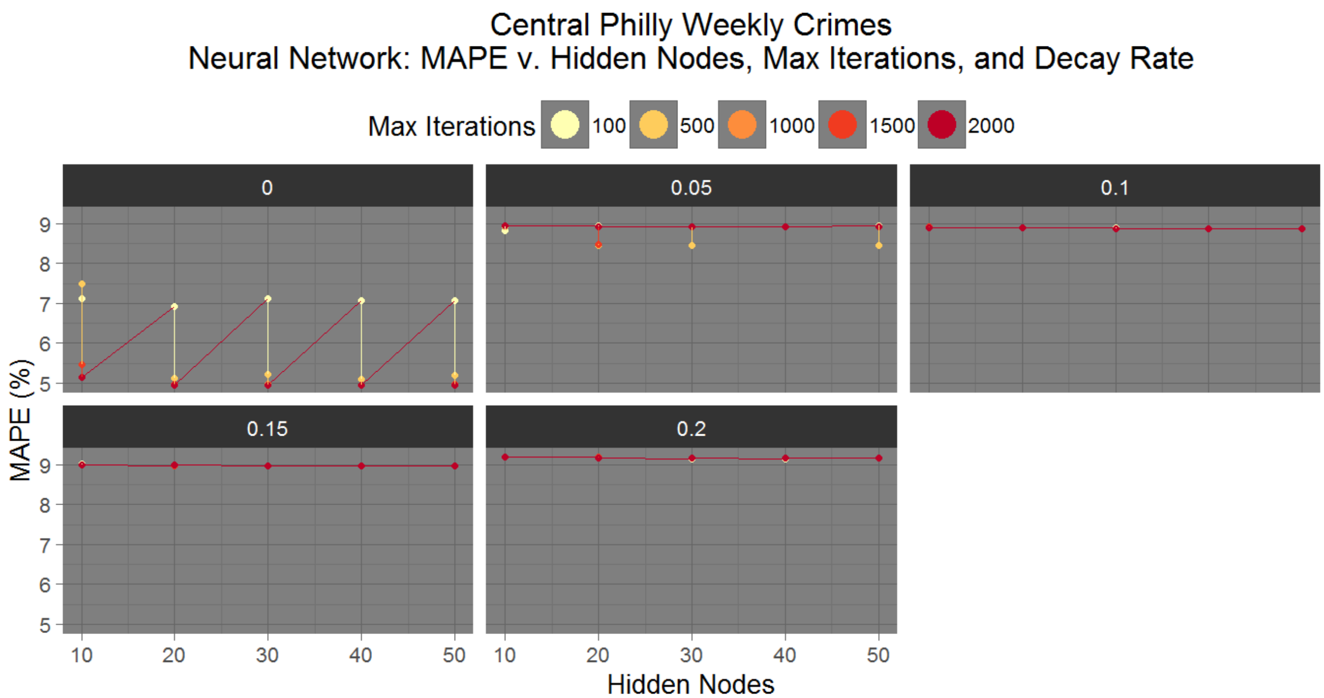


Figure 95: Neural Network Parameters Plot for Fitting Central Philadelphia Weekly Crimes

The performance metrics of the chosen model is shown below in Table 47. These values are significantly better than all previous fitted models, a complete comparison will be shown later in this section. The MAPE shows that this model does a good job at describing the behavior of the entire data set. The t-test p-value shows that we can be very confident that the residuals have an expected value of zero.

Table 47: DOE Result for Fitting Central Philadelphia Weekly Crimes 2006 - 2016

size	decay	maxit	t.pval	ME	RMSE	MAE	MPE	MAPE
30	0	2000	0.99996	8.7E-05	42.1312	32.2036	-0.4457	4.9575

The plot of fitted values on the actual values for the best fitted model is shown below in Figure 96, where the blue line is the fitted values, and the black points are the actual values. This plot shows that the model does a great job at capturing those low points, following the center of the data, and has distinct jumps to reach small clusters of data points. The stepwise behavior of the blue line is due to the regressors that follow a monthly scale, which are all of the regressors except DAYS.

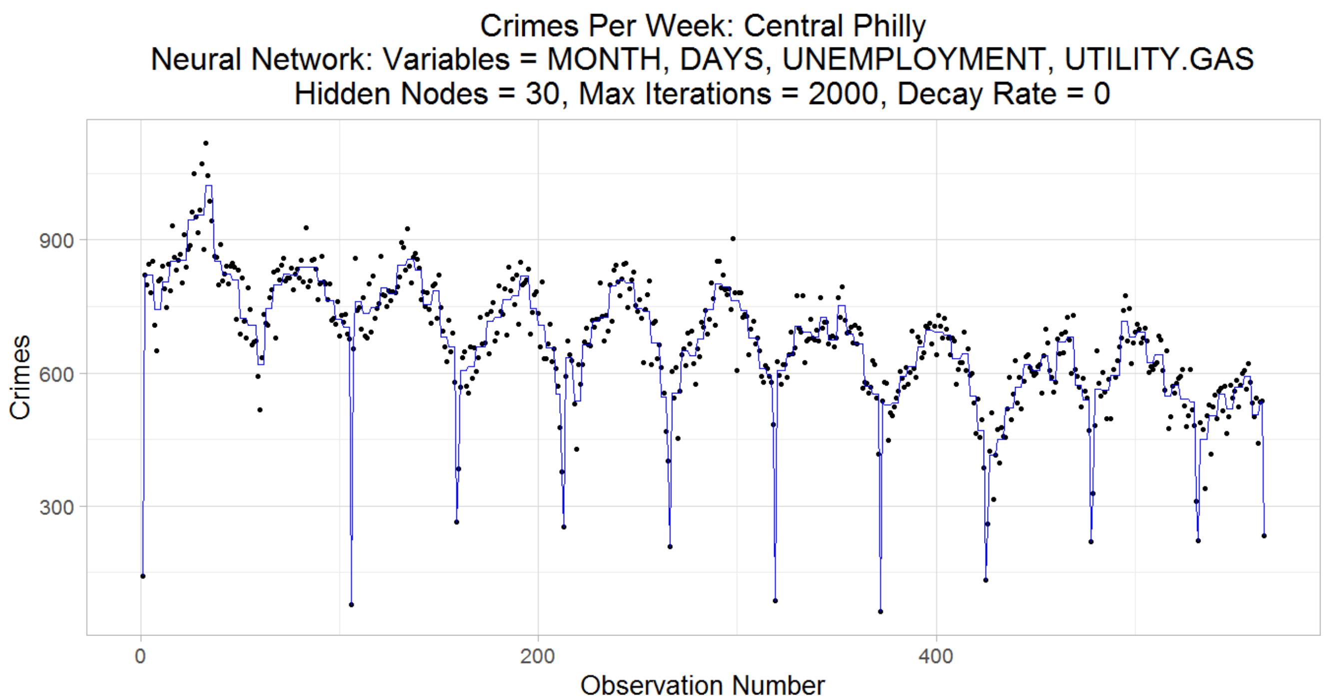


Figure 96: Prediction Plot for Central Philadelphia Weekly Crime Neural Network Fitted Model

The residual plots of the best fitted model are shown below. The residuals v. fitted plot for the model show that a constant mean was achieved. Variance appears constant too given that the lower values of crime don't vary as much as the higher values of crime do. This plot also shows that the residuals are centered at zero. The residual time series of the model shows that variance is constant over time and the trend and seasonality of the data was likely removed. The residuals for the model don't follow a normal distribution. The tails in the QQ plot are trailing off the line and 95% confidence interval band. The histogram shows that most of the data is center about zero, and that zero is contained within the one sample t-test 95% confidence interval.

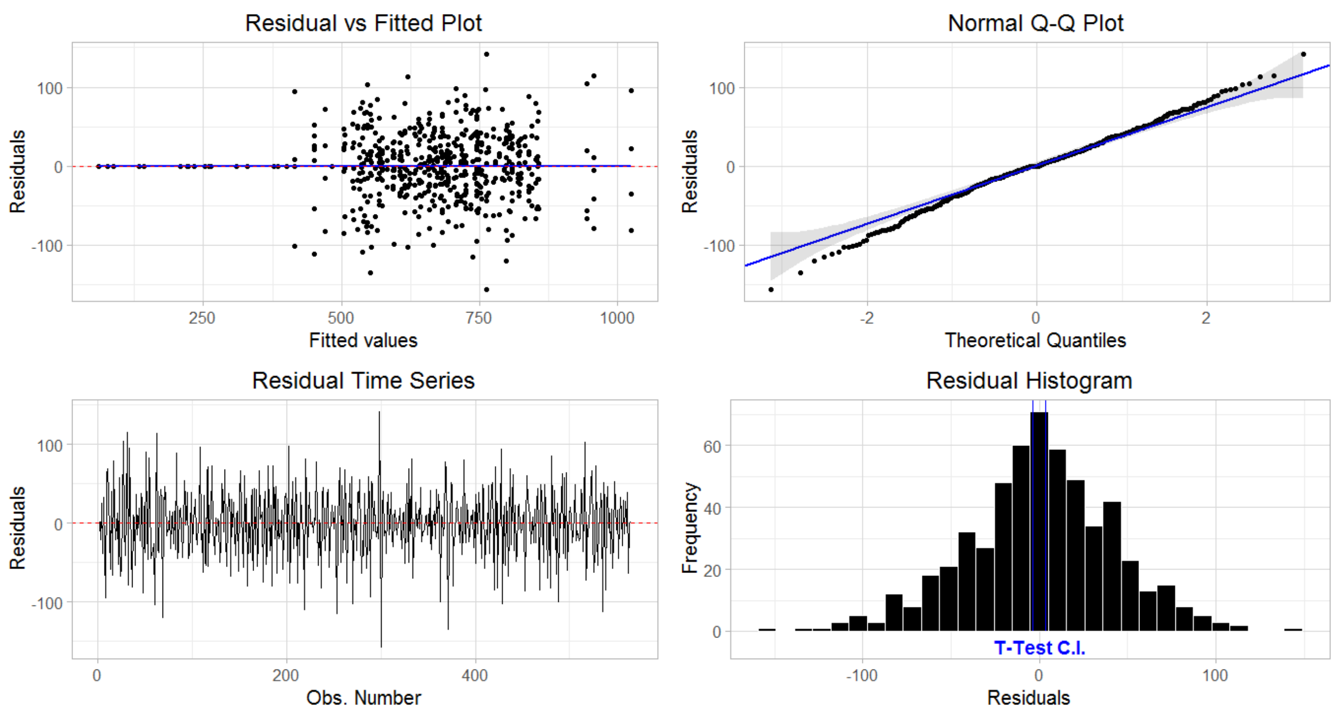


Figure 97: Residual Plots for Central Philadelphia Weekly Crime Neural Network Fitted Model

The performance metrics of the best fitted models from this analysis for Central Philadelphia are shown below. Across all performance metrics, the neural network model outperforms the other models significantly. This shows that the best descriptive model from this analysis for Central Philadelphia weekly crimes is the neural network model.

Table 48: Comparison of Central Philadelphia Weekly Crime Fitted Models

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	0.90127	0.48752	88.7029	61.075	-4.5767	13.9393
ARIMA	0.60257	-1.9049	82.589	55.509	-5.6036	13.5742
DR	0.65289	-1.1316	56.8011	44.7697	-0.7319	7.88243
NNET	0.99996	8.7E-05	42.1312	32.2036	-0.4457	4.9575

There were 125 forecasting models built by varying size, decay, and maxit. The MAPE of each of those models is plotted below in Figure 95 where size, MAPE, maxit, and decay are in the same plot dimension as the previous corresponding plot. This figure indicates that to decrease MAPE, decay should not equal 0, maxit could equal anything between 100 and 2000, and size could equal anything between 10 and 50. The combination that was ultimately chosen was (size, decay, maxit) = (40, 0.05, 2000). This combination was chosen by filtering the models based on a MAPE ≤ 12% and RMSE ≤ 73, and choosing from the remaining models which all had very similar performance metric values.

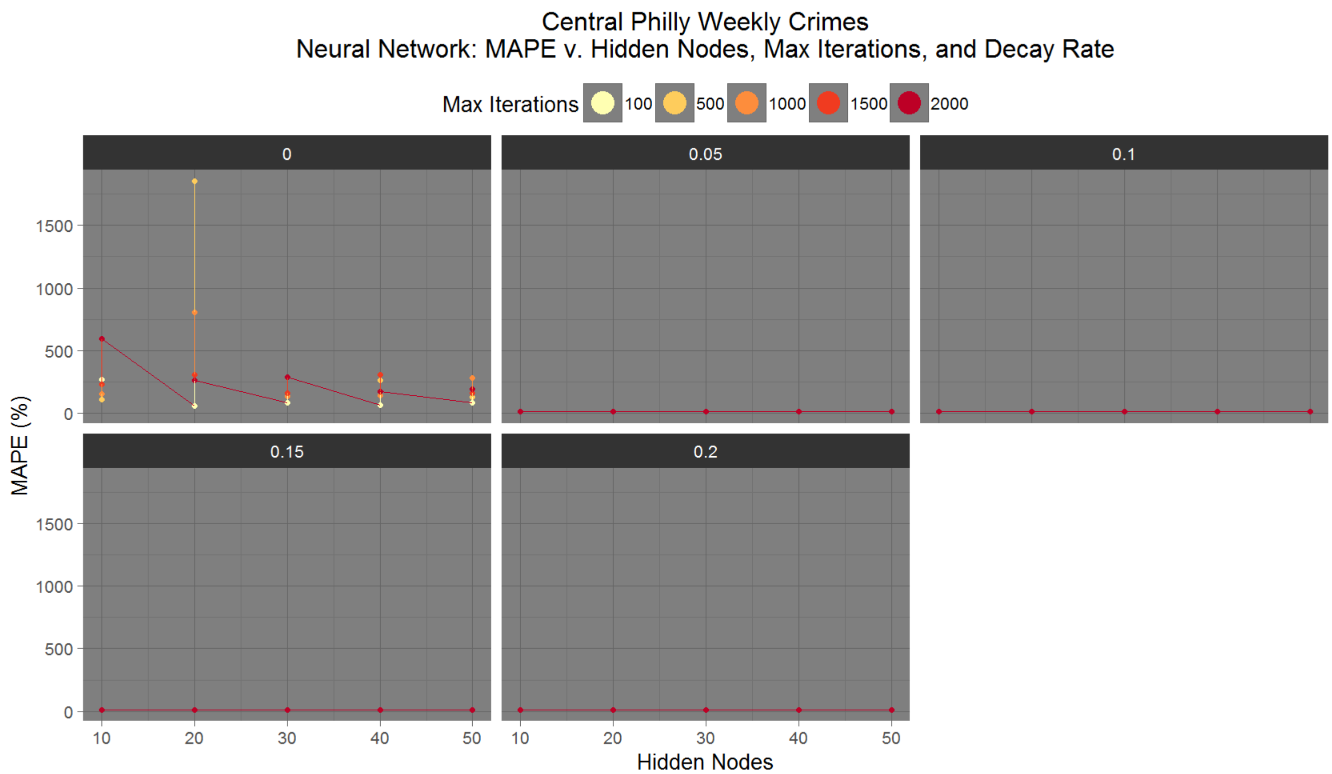


Figure 98: Neural Network Parameters Plot for Forecasting Central Philadelphia Weekly Crimes

The performance metrics of the chosen model is shown below in Table 47. These values are similar to a few previous forecasting models, a complete comparison will be shown later in this section. The MAPE shows that this model does a decent job at describing the behavior of the forecasting range. The t-test p-value shows that we can be confident that the residuals have an expected value of zero.

Table 49: DOE Result for Forecasting Central Philadelphia Weekly Crimes 2015 - 2016

size	decay	maxit	t.pval	ME	RMSE	MAE	MPE	MAPE
40	0.05	2000	0.11675	-12.17	72.2687	57.9407	-3.0798	11.1926

In an attempt to improve this forecast MAPE by 10%, two approaches were applied: Increasing the value of size, and replacing the month regressor with a week regressor that indicates what week of the year it is. Table 50 below shows that by increasing size, a small but steady improvement can be made. This table also shows that week appears to be a more promising regressor than month. Combining these two approaches where week is the new regressor, and the value of size is increased, an improvement in MAPE by 10% compared to Model 0 may be possible.

The issue with combining these approaches is benefit v. cost. The amount of time and resources required to build a model that uses a 53 regressors to represent week, instead of 11 regressors to represent month, and then uses more and more single layer hidden nodes, may be too much cost for the benefit of improving MAPE by 10%. This means that the current experiment process that builds 125 neural network models does a sufficient job at finding a high performing neural network model, and to test any more scenarios will likely result in insignificant improvement after spending a significant amount of time and computational resources.

Table 50: DOE Results for Improving the Forecasting of Central Philadelphia Weekly Crimes 2015 - 2016

Model	size	decay	maxit	t.pval	ME	RMSE	MAE	MPE	MAPE
0	40	0.05	2000	0.11675	-12.172	72.2687	57.9407	-3.0798	11.1926
1	60	0.05	2000	0.11495	-12.224	72.2287	57.905	-3.0905	11.1841
2	70	0.05	2000	0.1154	-12.205	72.2044	57.8855	-3.0876	11.1793
3	80	0.05	2000	0.11513	-12.213	72.1962	57.8785	-3.089	11.1777
4	90	0.05	2000	0.11486	-12.221	72.1893	57.8722	-3.0909	11.1761
5	100	0.05	2000	0.59846	-3.9932	70.1709	57.1254	-1.5854	10.9525
week	40	0.05	2000	0.2923	-8.0342	70.7735	56.2488	-2.2217	10.6227

The plot of fitted values on the actual values for the best forecasting model is shown below in Figure 96, where the blue line is the fitted values, and the red points are the actual values which the neural network model is tasked to predict. This plot shows similar behavior as the previous dynamic regression forecasting model without lagged regressors. The model under predicts the first half of 2015, over predicts the second half of 2015, and then predicts the center of 2016 well up until the end where it begins to over predict. The under and over predicting in 2015 is due to the 2015 peak starting earlier, compared to the 2014 season.

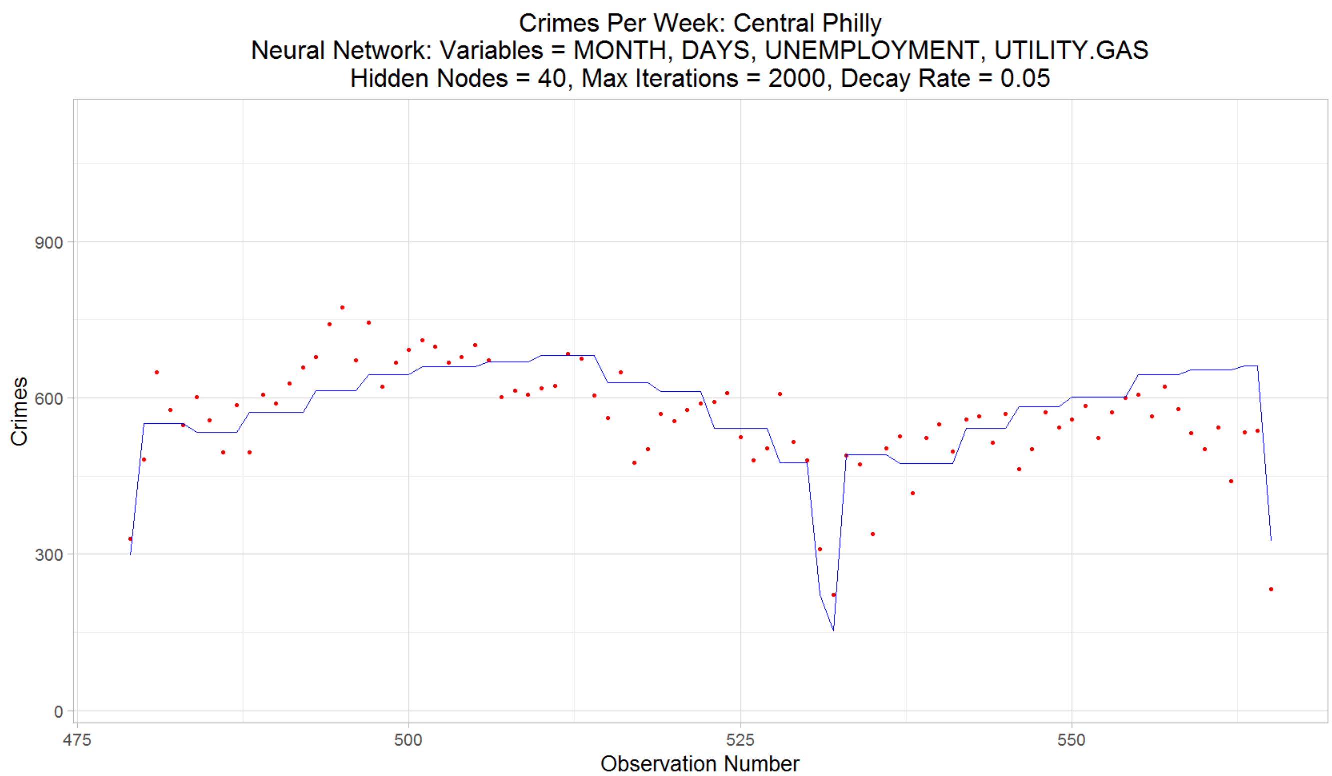


Figure 99: Prediction Plot for Central Philadelphia Weekly Crime Neural Network Forecasting Model

The residual plots of the best forecasting model are shown below. The residuals v. fitted plot for the model show that a constant mean was not achieved. Variance doesn't appears constant either. This plot also shows that the residuals may be centered at zero. The residual time series of the model shows that variance is not constant over time and the trend and seasonality of the data was not removed. The residuals for the model appear to follow a normal distribution. The points in the QQ plot are following the line and staying within the 95% confidence interval band. The histogram shows that the data is centered slightly to the left of zero, and that zero is contained within the one sample t-test 95% confidence interval.

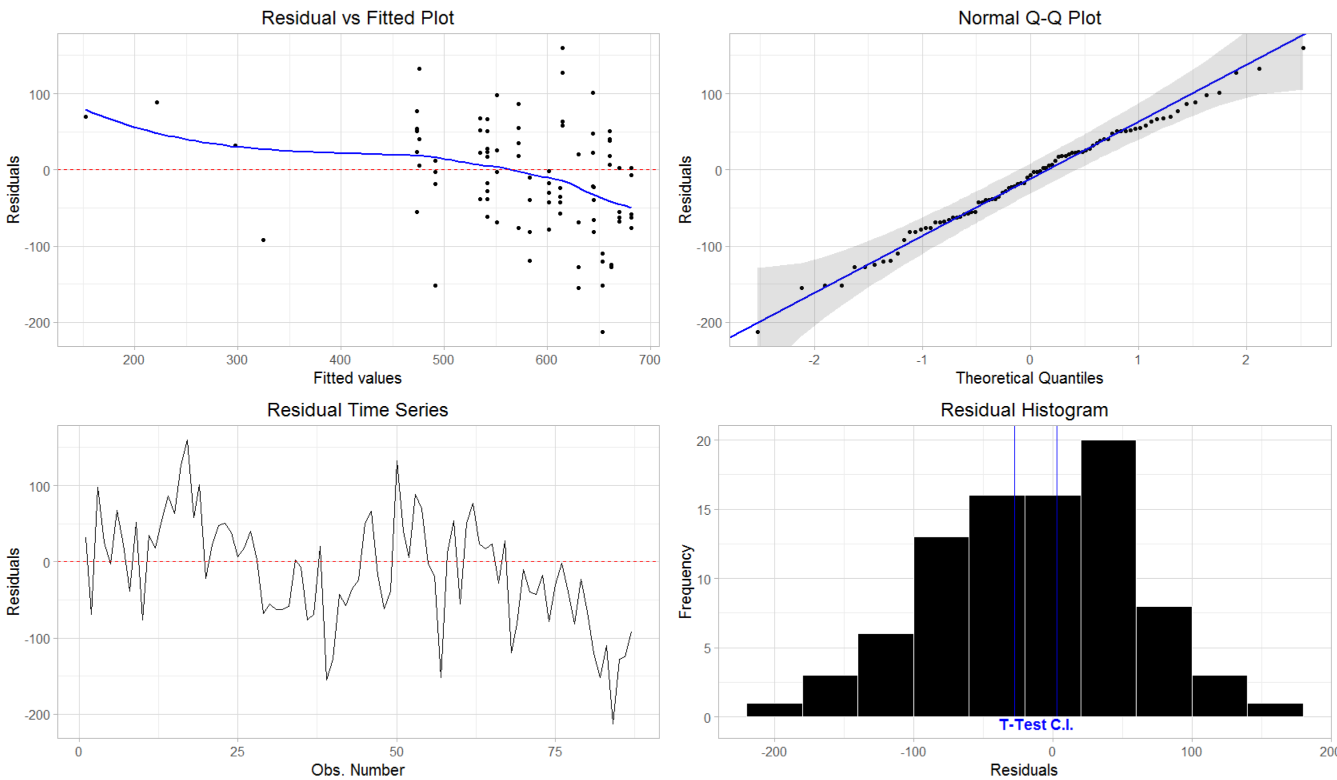


Figure 100: Residual Plots for Central Philadelphia Weekly Crime Neural Network Forecasting Model

The performance metrics of the best forecasting models from this analysis for Central Philadelphia are shown below. The neural network model has slightly better MAPE than the runner-ups, and it also has better RMSE than all other models. The concern with the neural network model is the clear bias shown by ME, which explains why the t-test p-value is low. Considering bias, residuals with an expected value of zero, and MAPE, the HoltWinters model is the best predictive model for Central Philadelphia weekly crimes in 2015 - 2016.

Table 51: Comparison of Central Philadelphia Weekly Crime Forecasting Models

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	0.91563	0.85826	74.9167	55.7464	-2.1416	11.5526
ARIMA	1E-06	57.4061	116.252	91.6005	7.80079	17.4541
DR	0.40469	7.33899	81.6045	64.4347	0.17805	11.8345
NNET	0.11675	-12.172	72.2687	57.9407	-3.0798	11.1926

West Philadelphia

There were 125 fitted models built by varying size, decay, and maxit. The MAPE of each of those models is plotted below in Figure 101 where size, MAPE, maxit, and decay are in the same plot dimension as the previous corresponding plot. This figure indicates that to decrease MAPE, decay should equal 0, maxit should equal 2000, and size could equal anything between 20 and 50. The combination that was ultimately chosen was (size, decay, maxit) = (30, 0, 2000). The size value of 30 was chosen because it is not close to the behavior of a size value of 10, and adding up to 20 more nodes for a size value of 50 doesn't produce any significant improvement in MAPE.

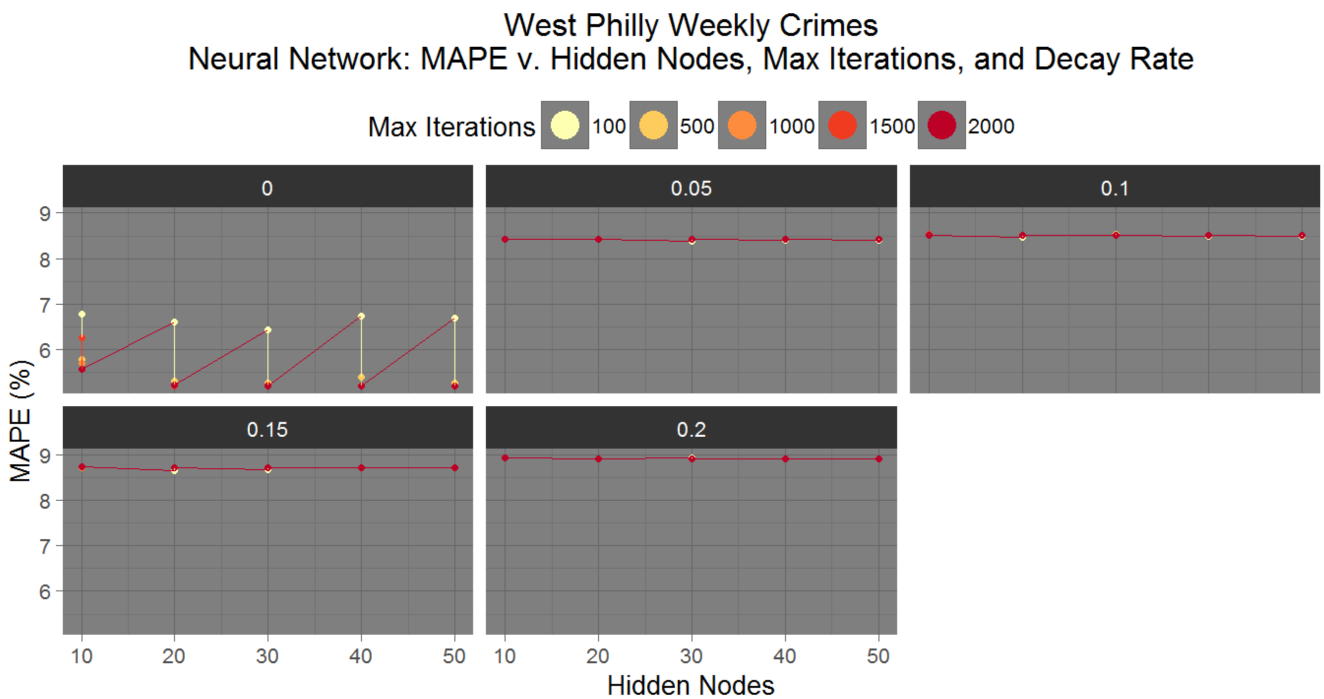


Figure 101: Neural Network Parameters Plot for Fitting West Philadelphia Weekly Crimes

The performance metrics of the chosen model is shown below in Table 52. These values are significantly better than all previous fitted models, a complete comparison will be shown later in this section. The MAPE shows that this model does a good job at describing the behavior of the entire data set. The t-test p-value shows that we can be very confident that the residuals have an expected value of zero.

Table 52: DOE Result for Fitting West Philadelphia Weekly Crimes 2006 - 2016

size	decay	maxit	t.pval	ME	RMSE	MAE	MPE	MAPE
30	0	2000	0.99995	-1E-04	44.2284	33.6294	-0.4932	5.21754

The plot of fitted values on the actual values for the best fitted model is shown below in Figure 102, where the blue line is the fitted values, and the black points are the actual values. This plot shows that the model does a great job at capturing those low points, following the center of the data, and has distinct jumps to reach small clusters of data points. The stepwise behavior of the blue line is due to the regressors that follow a monthly scale, which are all of the regressors except DAYS.

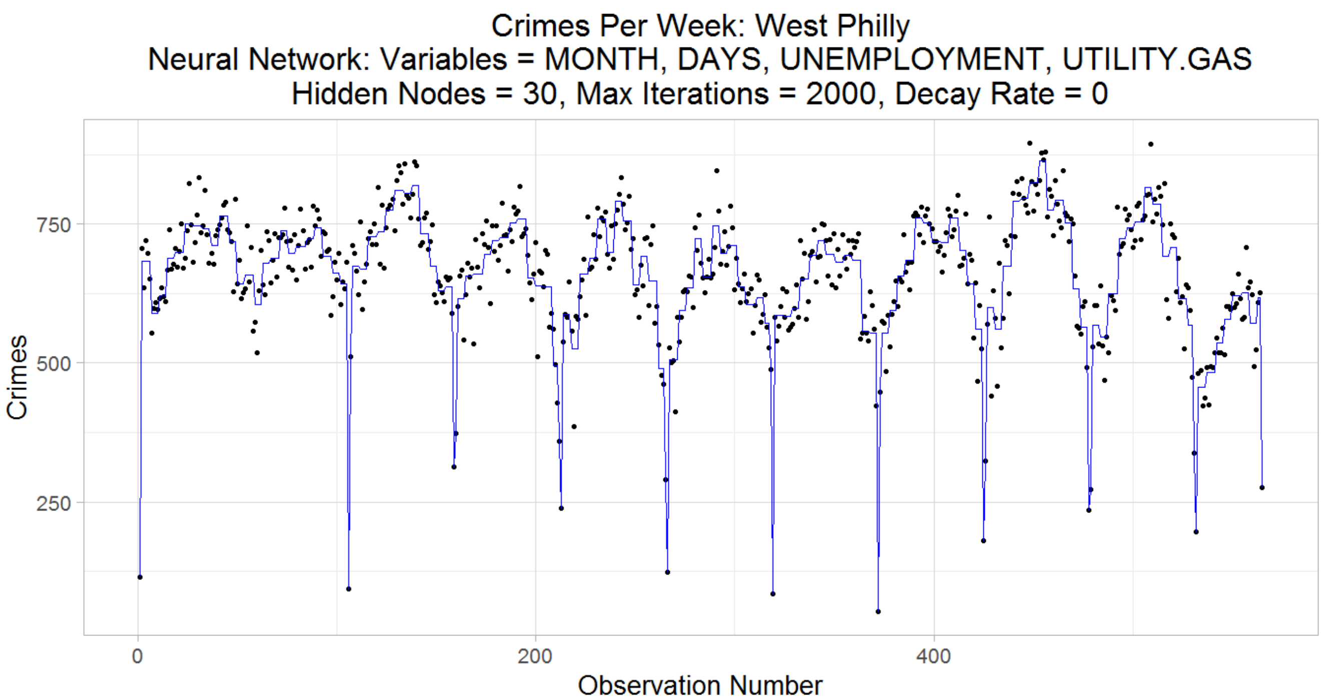


Figure 102: Prediction Plot for West Philadelphia Weekly Crime Neural Network Fitted Model

The residual plots of the best fitted model are shown below. The residuals v. fitted plot for the model show that a constant mean was achieved. Variance appears to have a bubble shape for the higher values of crime. This plot also shows that the residuals are centered at zero. The residual time series of the model shows that variance is constant over time up until the last two seasons where it increases. This plot also shows that the trend and seasonality of the data was likely removed. The residuals for the model don't follow a normal distribution. The tails in the QQ plot are trailing off the line and 95% confidence interval band. The histogram shows that most of the data is center about zero, and that zero is contained within the one sample t-test 95% confidence interval.

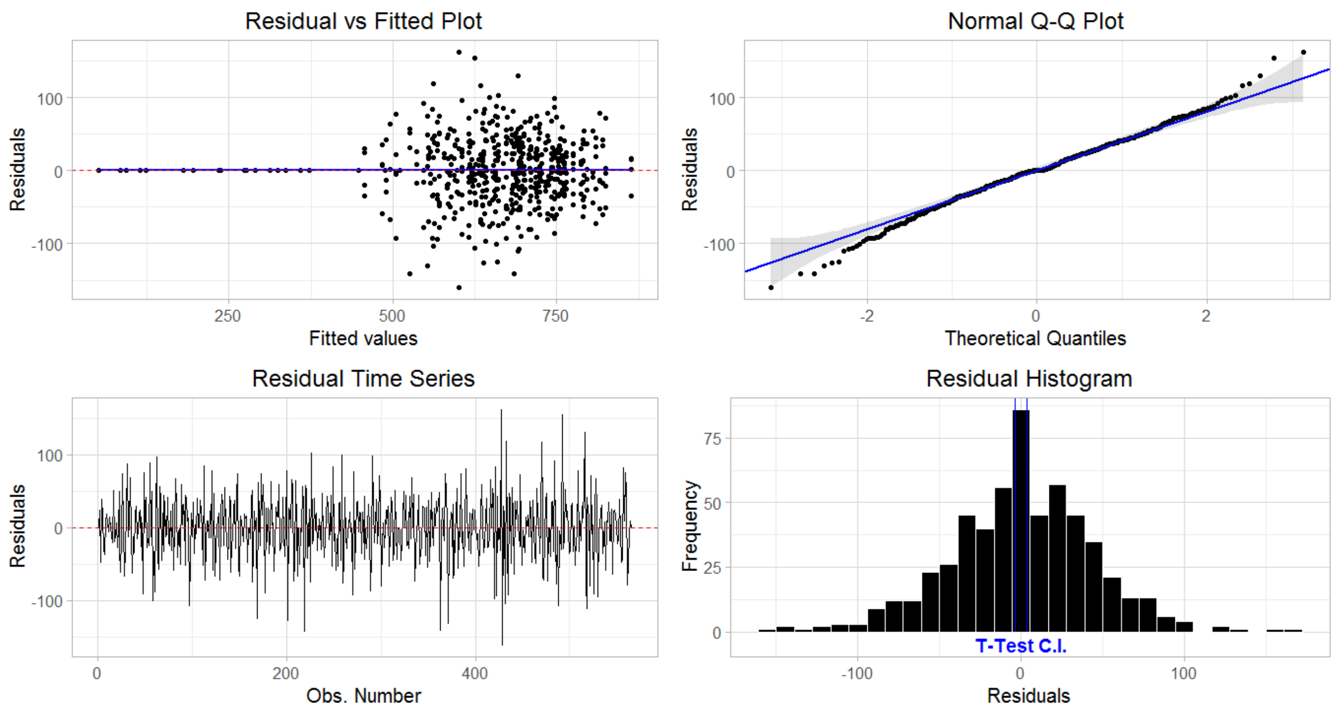


Figure 103: Residual Plots for West Philadelphia Weekly Crime Neural Network Fitted Model

The performance metrics of the best fitted models from this analysis for West Philadelphia are shown below. Across all performance metrics, the neural network model outperforms the other models. This shows that the best descriptive model from this analysis for West Philadelphia weekly crimes is the neural network model.

Table 53: Comparison of West Philadelphia Weekly Crime Fitted Models

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	0.85174	-0.7449	89.967	62.0272	-5.1612	14.3997
ARIMA	0.93978	-0.2807	83.8639	57.1777	-5.8464	13.9571
DR	0.56443	-1.5515	60.783	47.0652	-0.4812	7.89742
NNET	0.99995	-0.0001	44.2284	33.6294	-0.4932	5.21754

There were 125 forecasting models built by varying size, decay, and maxit. The MAPE of each of those models is plotted below in Figure 104 where size, MAPE, maxit, and decay are in the same plot dimension as the previous corresponding plot. This figure indicates that to decrease MAPE, decay should not equal 0, maxit could equal anything between 100 and 2000, and size could equal anything between 10 and 50. The combination that was ultimately chosen was (size, decay, maxit) = (50, 0.2, 2000). This combination was chosen by filtering the models based on a MAPE ≤ 20%, RMSE ≤ 120, and |ME| ≤ 80, and choosing from the remaining models which all had very similar performance metric values.

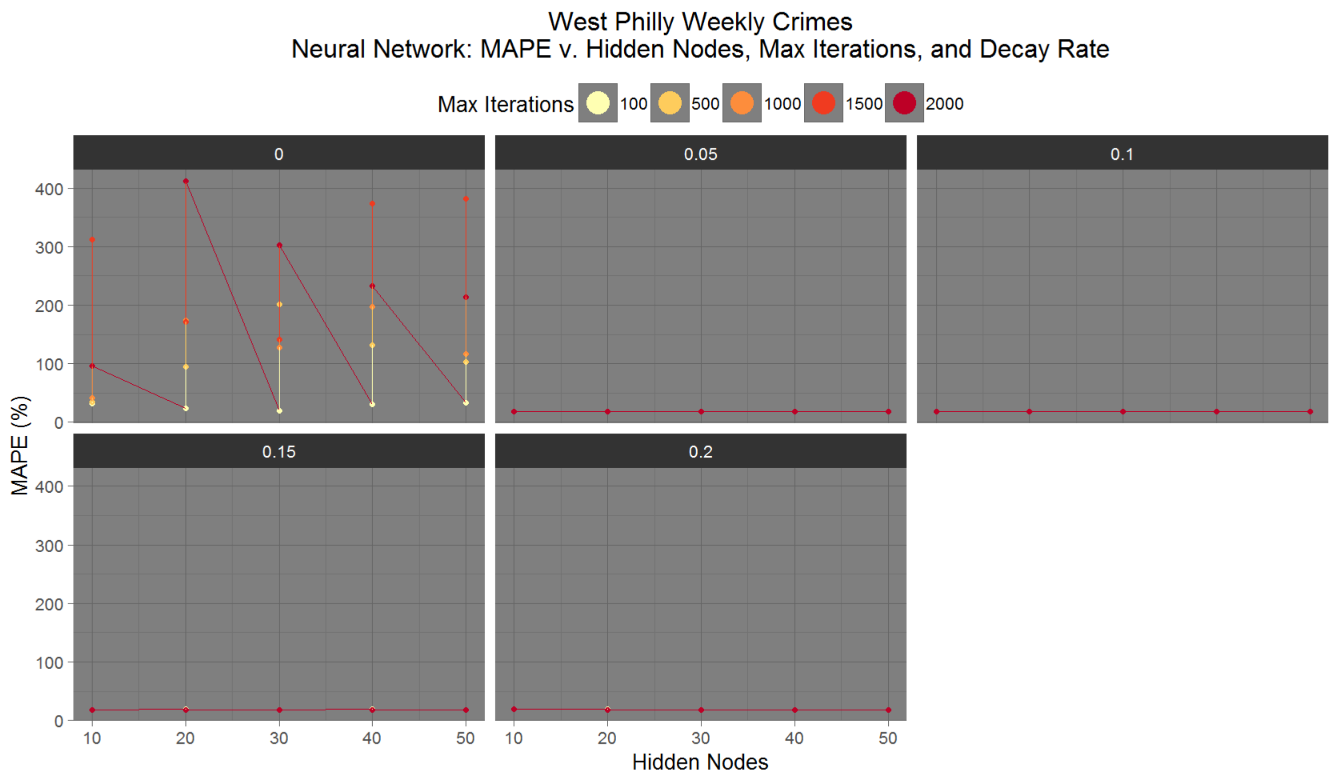


Figure 104: Neural Network Parameters Plot for Forecasting West Philadelphia Weekly Crimes

The performance metrics of the chosen model are shown below in Table 54. These values are worse than previous forecasting models, a complete comparison will be shown later in this section. The MAPE shows that this model does a poor job at describing the behavior of the forecasting range. The t-test p-value shows that the residuals don't have an expected value of zero, and the ME indicates that the model over predicts.

Table 54: DOE Result for Forecasting West Philadelphia Weekly Crimes 2015 - 2016

size	decay	maxit	t.pval	ME	RMSE	MAE	MPE	MAPE
50	0.2	2000	5.7E-13	-8E+01	117.319	95.4321	-16.446	18.5005

The plot of fitted values on the actual values for the best forecasting model is shown below in Figure 96, where the blue line is the fitted values, and the red points are the actual values which the neural network model is tasked to predict. The model over predicts the start and end of 2015, predicts the center of 2015 well, and over predicts all of 2016.

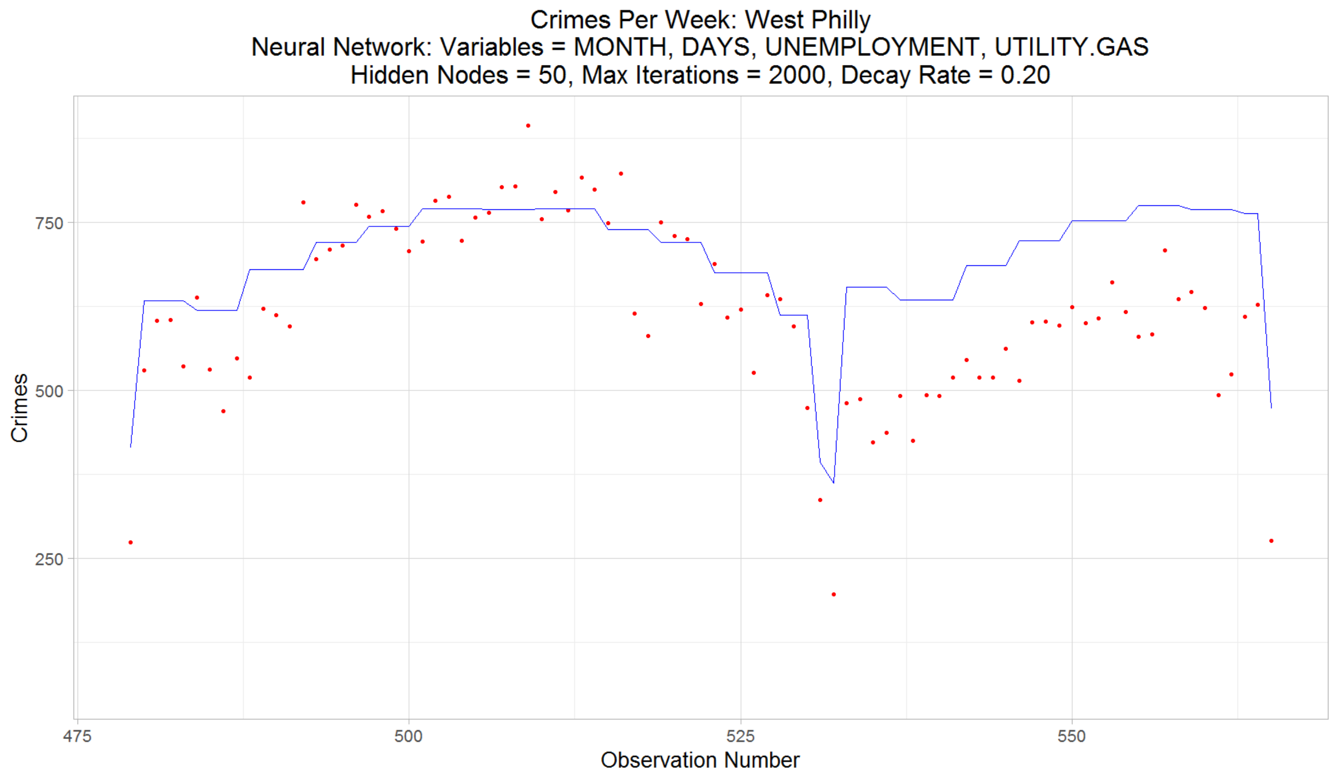


Figure 105: Prediction Plot for West Philadelphia Weekly Crime Neural Network Forecasting Model

The residual plots of the best forecasting model are shown below. The residuals v. fitted plot for the model show that a constant mean was not achieved. Variance doesn't appears constant either. This plot also shows that the residuals are not centered at zero. The residual time series of the model shows that variance is not constant over time and the trend and seasonality of the data was not removed. The residuals for the model don't follow a normal distribution. The points in the QQ plot don't follow the line and barely stay within the 95% confidence interval band. The histogram shows that the data is centered to the left of zero, zero is not contained within the one sample t-test 95% confidence interval, and the residuals are bimodal.

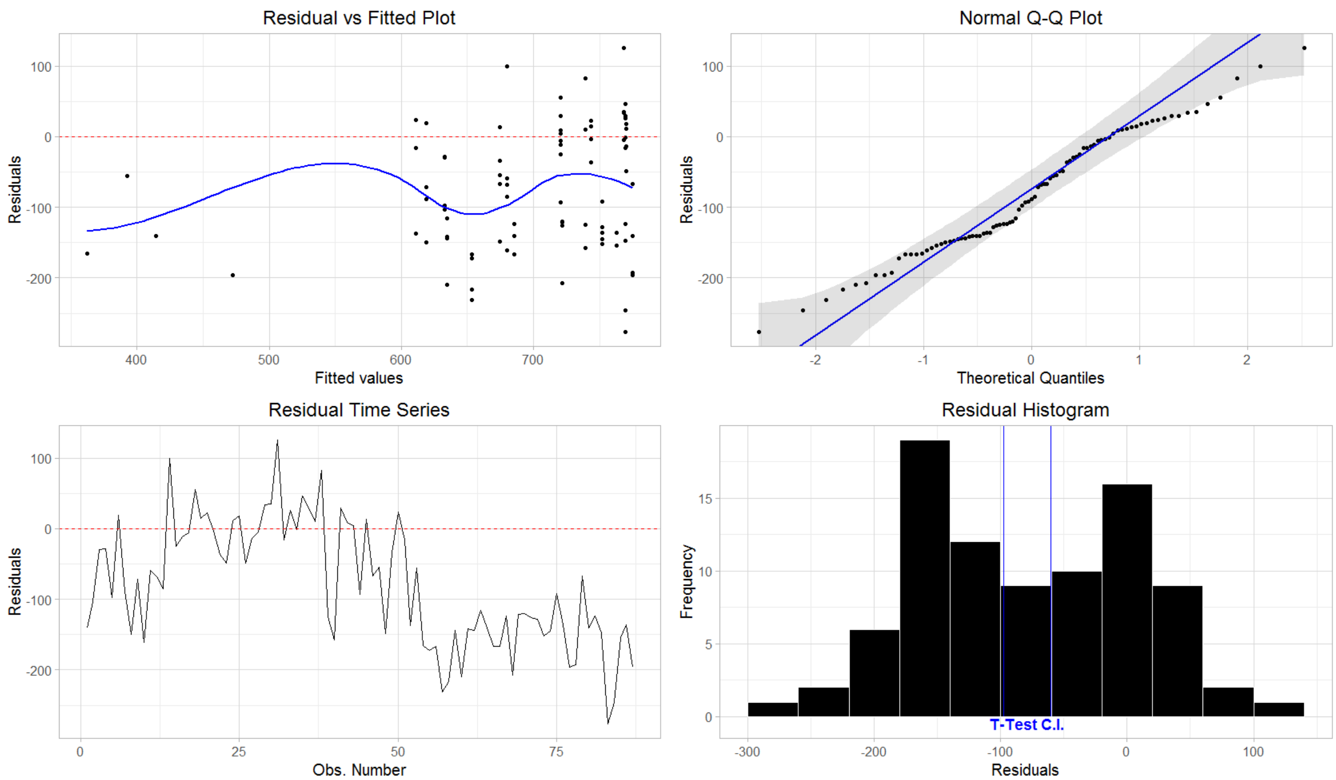


Figure 106: Residual Plots for West Philadelphia Weekly Crime Neural Network Forecasting Model

The performance metrics of the best forecasting models from this analysis for West Philadelphia are shown below. The neural network model is worse than all previous models. The dynamic regression model has a concerning ME but this doesn't appear too severe given the t-test p-value. Despite potential bias, and considering MAPE and RMSE, the dynamic regression model is the best predictive model for West Philadelphia weekly crimes in 2015 - 2016.

Table 55: Comparison of West Philadelphia Weekly Crime Forecasting Models

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	0.96319	-0.4019	80.5195	58.8806	-2.8648	11.8654
ARIMA	0.6012	-5.6713	100.412	75.5358	-4.0598	14.6643
DR	0.46938	5.96308	76.3293	61.7472	-0.7924	10.0733
NNET	5.7E-13	-79.163	117.319	95.4321	-16.446	18.5005

North Philadelphia

There were 125 fitted models built by varying size, decay, and maxit. The MAPE of each of those models is plotted below in Figure 107 where size, MAPE, maxit, and decay are in the same plot dimension as the previous corresponding plot. This figure indicates that to decrease MAPE, decay should equal 0, maxit should equal 2000, and size could equal anything between 20 and 50. The combination that was ultimately chosen was (size, decay, maxit) = (30, 0, 2000). The size value of 30 was chosen because it is not close to the behavior of a size value of 10, and adding up to 20 more nodes for a size value of 50 doesn't produce any significant improvement in MAPE.

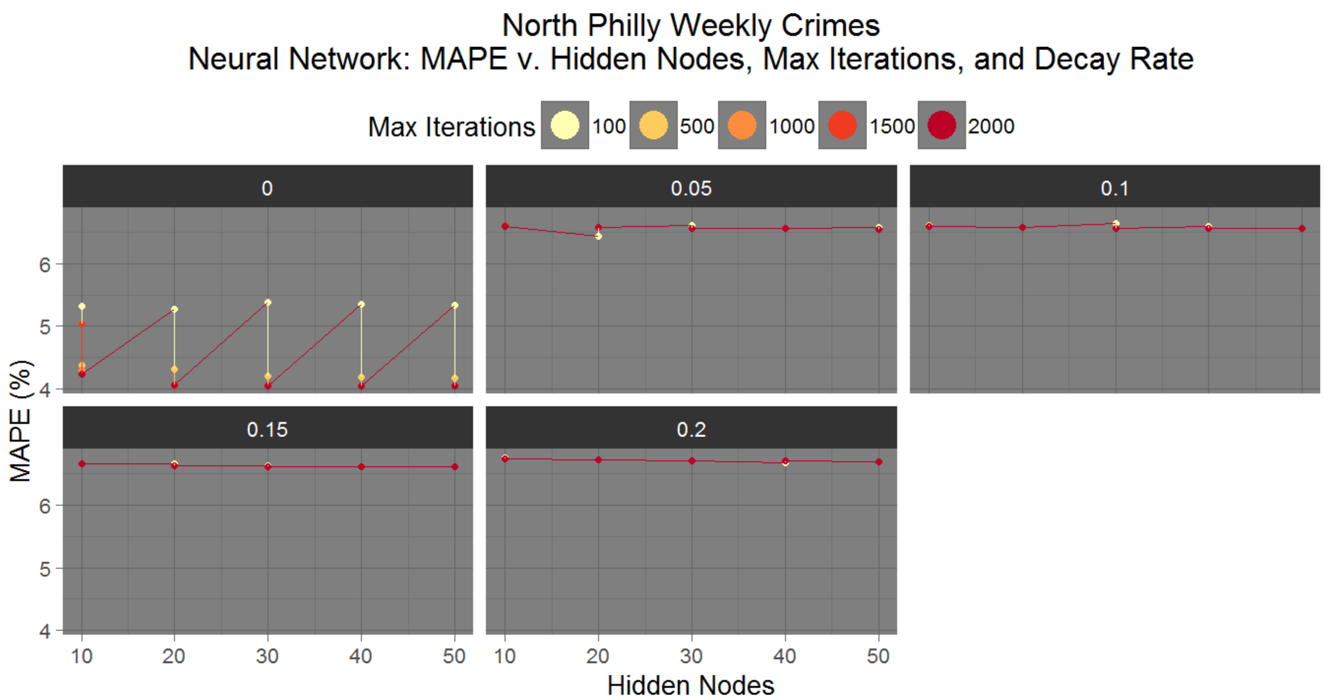


Figure 107: Neural Network Parameters Plot for Fitting North Philadelphia Weekly Crimes

The performance metrics of the chosen model is shown below in Table 56. These values are significantly better than all previous fitted models, a complete comparison will be shown later in this section. The MAPE shows that this model does a good job at describing the behavior of the entire data set. The t-test p-value shows that we can be very confident that the residuals have an expected value of zero.

Table 56: DOE Result for Fitting North Philadelphia Weekly Crimes 2006 - 2016

size	decay	maxit	t.pval	ME	RMSE	MAE	MPE	MAPE
30	0	2000	0.99999	6.1E-05	93.4906	69.5313	-0.3156	4.05409

The plot of fitted values on the actual values for the best fitted model is shown below in Figure 108, where the blue line is the fitted values, and the black points are the actual values. This plot shows that the model does a great job at capturing those low points, following the center of the data, and has distinct jumps to reach small clusters of data points. The stepwise behavior of the blue line is due to the regressors that follow a monthly scale, which are all of the regressors except DAYS.

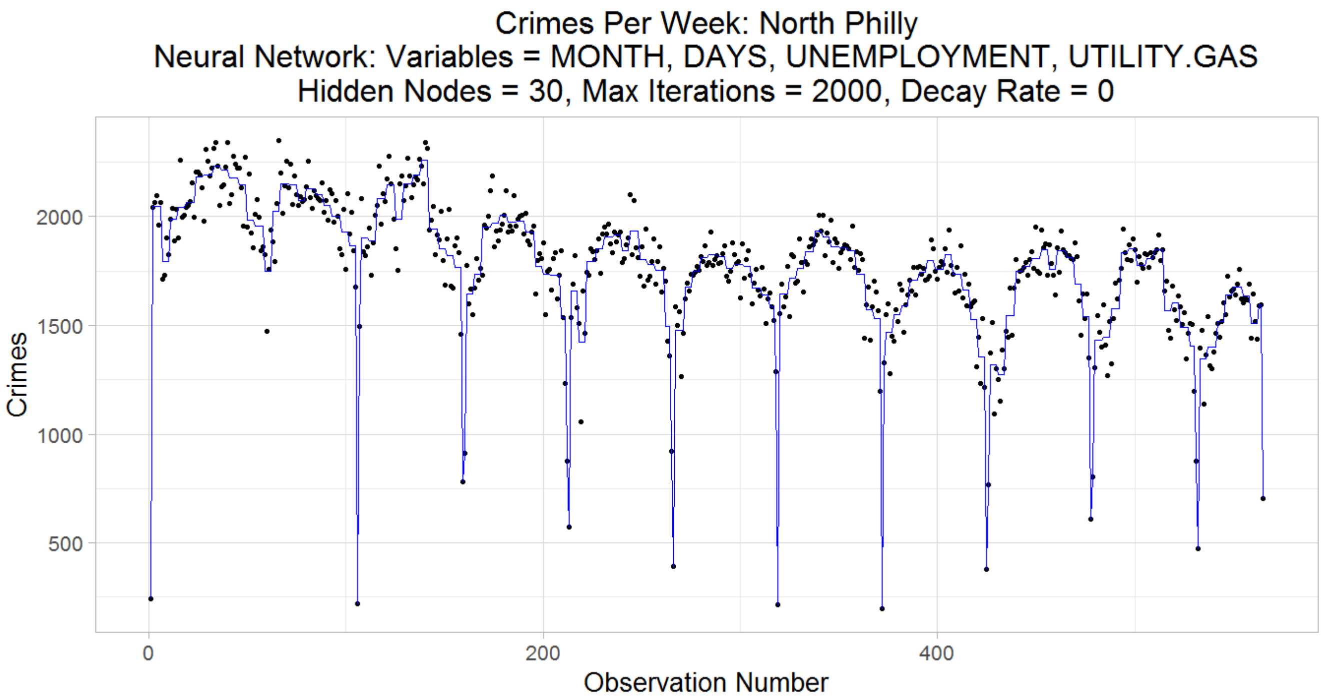


Figure 108: Prediction Plot for North Philadelphia Weekly Crime Neural Network Fitted Model

The residual plots of the best fitted model are shown below. The residuals v. fitted plot for the model show that a constant mean was achieved. Variance appears constant too given that the lower values of crime don't vary as much as the higher values of crime do. This plot also shows that the residuals are centered at zero. The residual time series of the model shows that variance is slightly higher during the first few seasons and then is constant for the remaining seasons. This plot also shows that the trend and seasonality of the data was likely removed. The residuals for the model don't follow a normal distribution. The tails in the QQ plot are trailing off the line and 95% confidence interval band. The histogram shows that most of the data is center about zero, and that zero is contained within the one sample t-test 95% confidence interval.

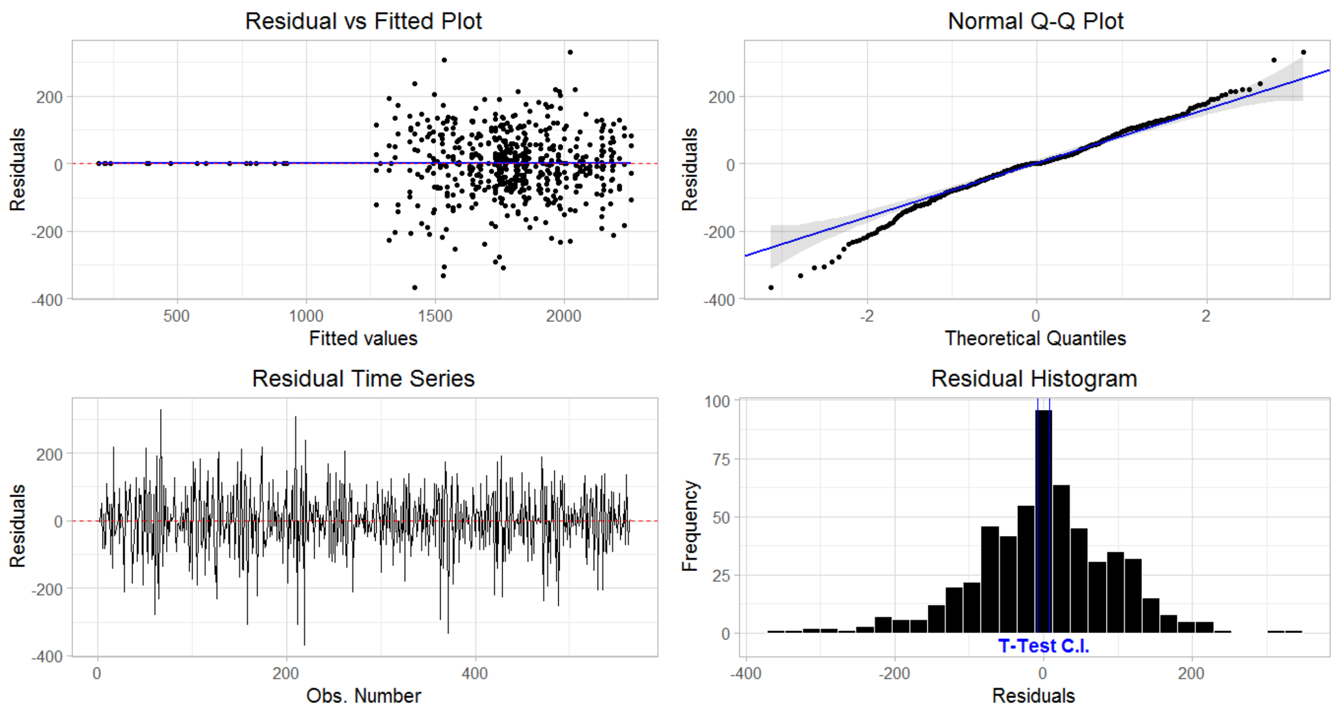


Figure 109: Residual Plots for North Philadelphia Weekly Crime Neural Network Fitted Model

The performance metrics of the best fitted models from this analysis for North Philadelphia are shown below. Across all performance metrics except MPE, the neural network model outperforms the other models. This shows that the best descriptive model from this analysis for North Philadelphia weekly crimes is the neural network model.

Table 57: Comparison of North Philadelphia Weekly Crime Fitted Models

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	0.79361	-2.5317	218.431	137.381	-4.7704	12.5506
ARIMA	0.57437	-5.1099	205.406	129.643	-5.1518	12.2856
DR	0.75908	-1.7354	127.733	98.7859	0.27499	6.6103
NNET	0.99999	6.1E-05	93.4906	69.5313	-0.3156	4.05409

There were 125 forecasting models built by varying size, decay, and maxit. The MAPE of each of those models is plotted below in Figure 110 where size, MAPE, maxit, and decay are in the same plot dimension as the previous corresponding plot. This figure indicates that to decrease MAPE, decay should not equal 0, maxit could equal anything between 100 and 2000, and size could equal anything between 10 and 50. The combination that was ultimately chosen was (size, decay, maxit) = (50, 0.05, 2000). This combination was chosen by filtering the models based on a MAPE ≤ 12% and RMSE ≤ 166, and choosing from the remaining models which all had very similar performance metric values.

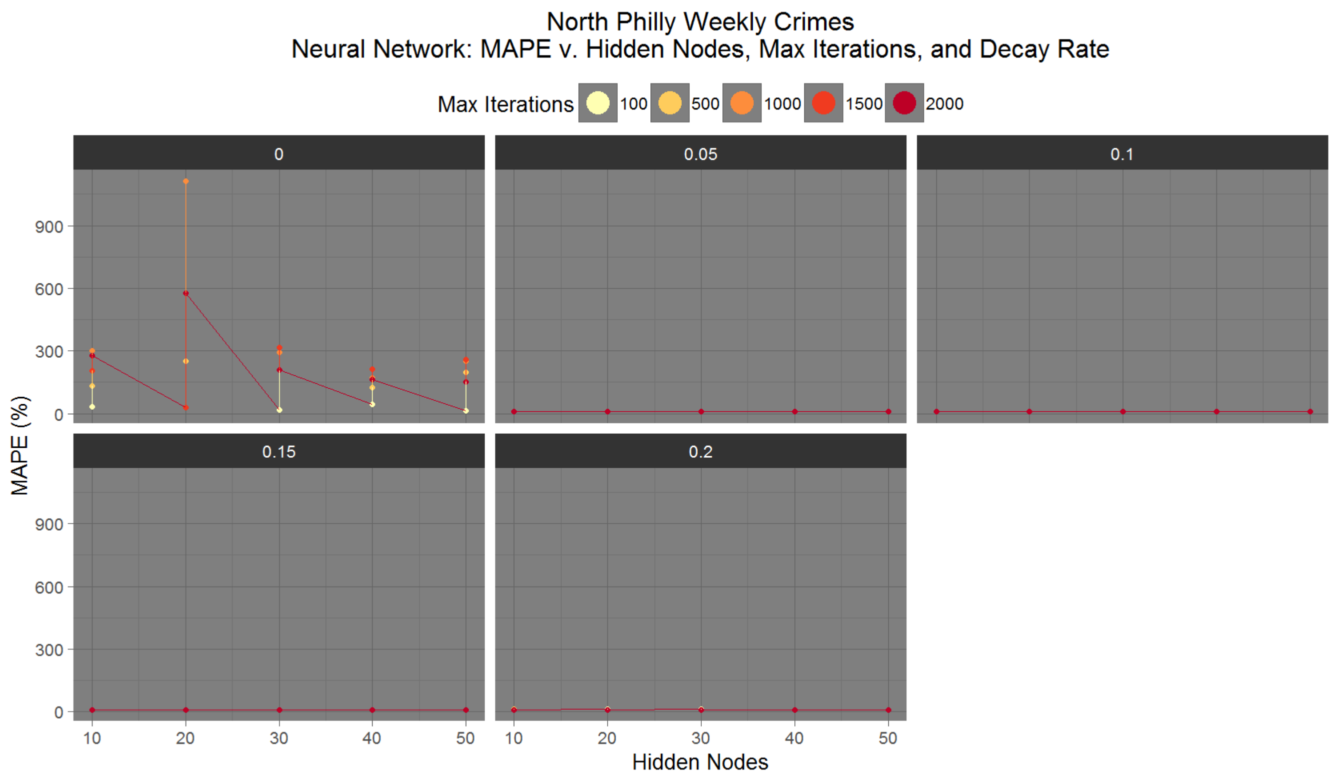


Figure 110: Neural Network Parameters Plot for Forecasting North Philadelphia Weekly Crimes

The performance metrics of the chosen model is shown below in Table 58. These values are similar to a few previous forecasting models, a complete comparison will be shown later in this section. The MAPE shows that this model does a decent job at describing the behavior of the forecasting range. The t-test p-value shows that the residuals don't have an expected value of zero, and the ME indicates that the model over predicts.

Table 58: DOE Result for Forecasting North Philadelphia Weekly Crimes 2015 - 2016

size	decay	maxit	t.pval	ME	RMSE	MAE	MPE	MAPE
50	0.05	2000	1.1E-14	-1E+02	165.758	129.488	-8.4212	9.15845

The plot of fitted values on the actual values for the best forecasting model is shown below in Figure 111, where the blue line is the fitted values, and the red points are the actual values which the neural network model is tasked to predict. The model over predicts the start and end of 2015, predicts the center of 2015 well, and over predicts all of 2016.

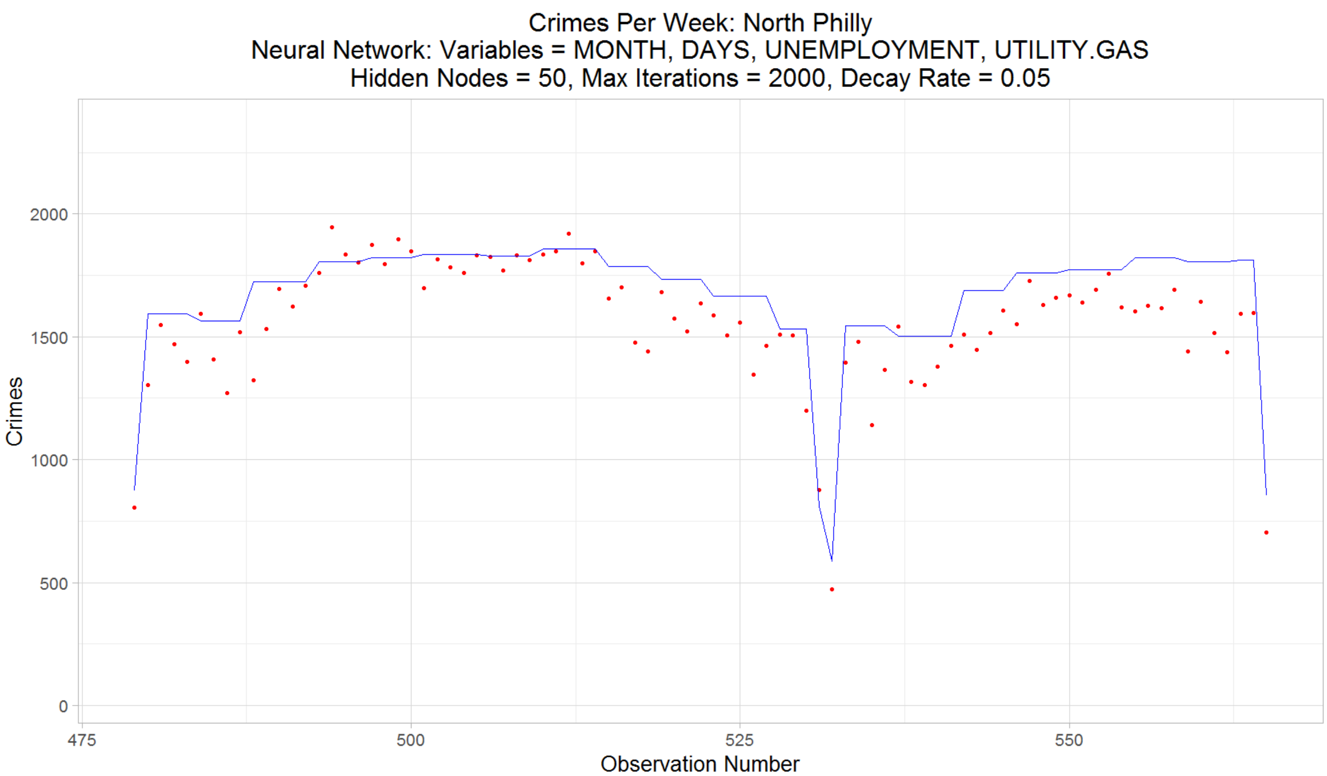


Figure 111: Prediction Plot for North Philadelphia Weekly Crime Neural Network Forecasting Model

The residual plots of the best forecasting model are shown below. The residuals v. fitted plot for the model show that a constant mean was not achieved. Variance doesn't appears constant either. This plot also shows that the residuals are not centered at zero. The residual time series of the model shows that variance is not constant over time and the trend and seasonality of the data was not removed. The residuals for the model don't follow a normal distribution. The tails in the QQ plot don't follow the line and barely stay within the 95% confidence interval band. The histogram shows that the data is centered to the left of zero, zero is not contained within the one sample t-test 95% confidence interval, and the residuals have a cluster of over predicted forecasts to the left.

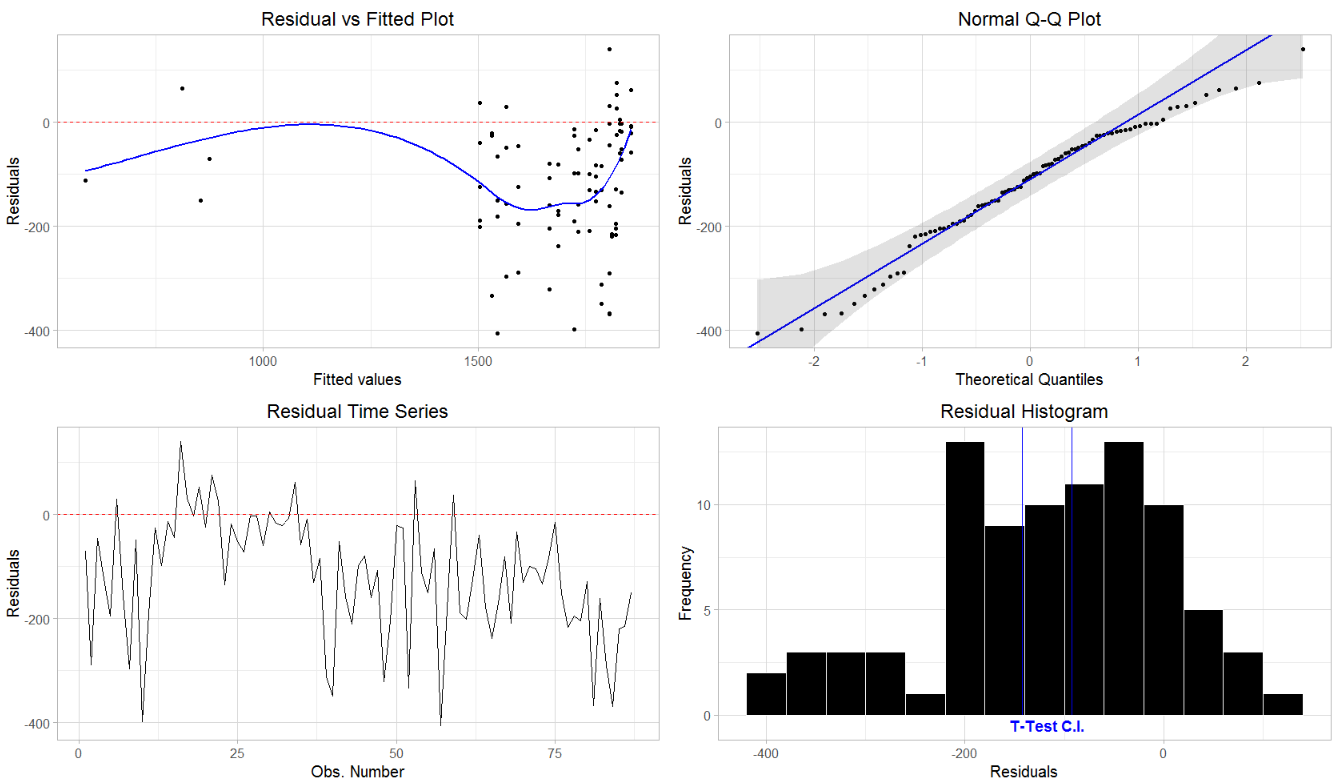


Figure 112: Residual Plots for North Philadelphia Weekly Crime Neural Network Forecasting Model

The performance metrics of the best forecasting models from this analysis for North Philadelphia are shown below. The neural network model is mediocre relative to all previous models. The dynamic regression model has the best overall performance metrics and is the best predictive model for North Philadelphia weekly crimes in 2015 - 2016.

Table 59: Comparison of North Philadelphia Weekly Crime Forecasting Models

Model	t.pval	ME	RMSE	MAE	MPE	MAPE
HW	0.8858	-2.8834	185.661	121.258	-2.4663	9.86163
ARIMA	0.22173	29.1673	221.682	152.684	-0.6937	12.105
DR	0.76311	-4.3528	133.573	105.986	-1.096	7.02422
NNET	1.1E-14	-117.53	165.758	129.488	-8.4212	9.15845

There are two consistent issues that all of these forecasting methods across all three areas of Philadelphia have yet to contain. The issues are that the weekly crimes data has shifting peaks from year to year, and any given year could be randomly higher or lower compared to the previous year. This means that when the season that is to be forecasted, has a similar shape and position as the previous season that is in the training set, a forecasting model can be expected to perform well, otherwise performance will degrade for that particular forecasted season. Perhaps, creating regressors that have information regarding the slope and intercept of first quarter of each season may allow the forecasting models to prepare for these two issues.