Forecasting Lyme Disease

Exploring Google Trends

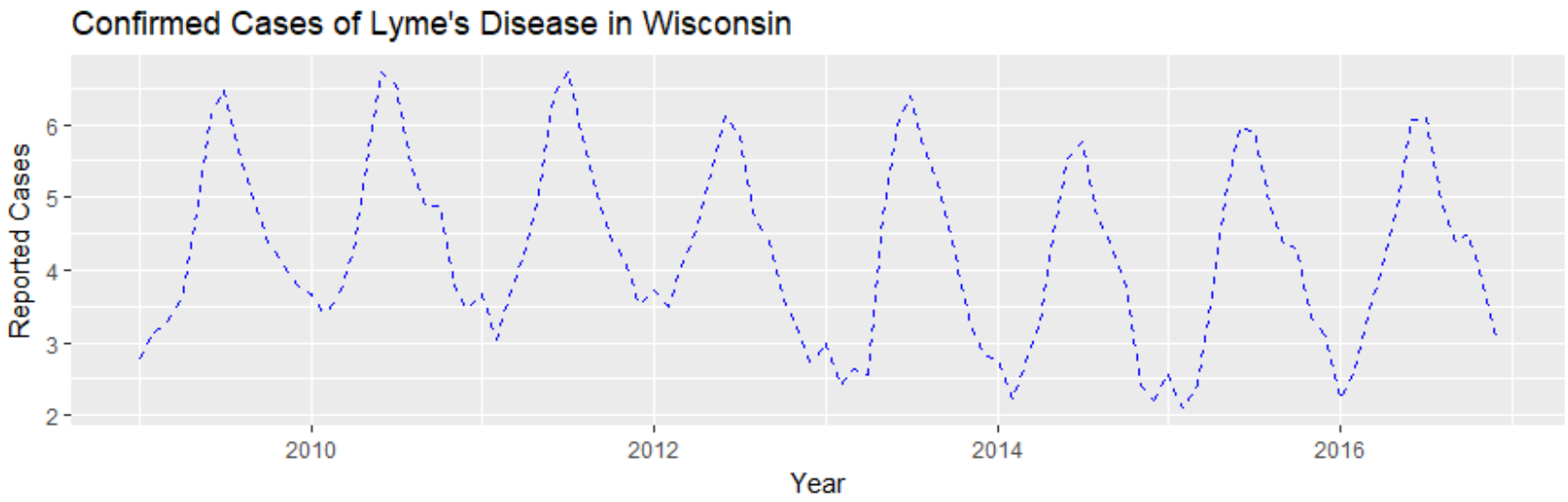
Jeremy Ling Alexis Rivera

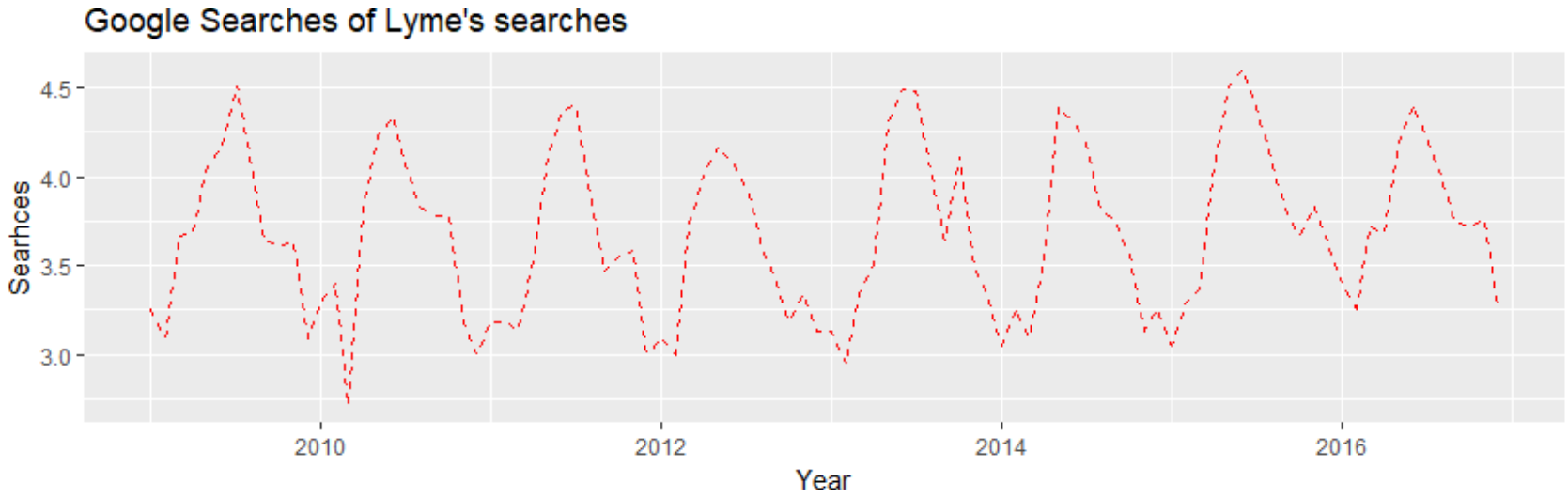
San Francisco State University

Spring 2018

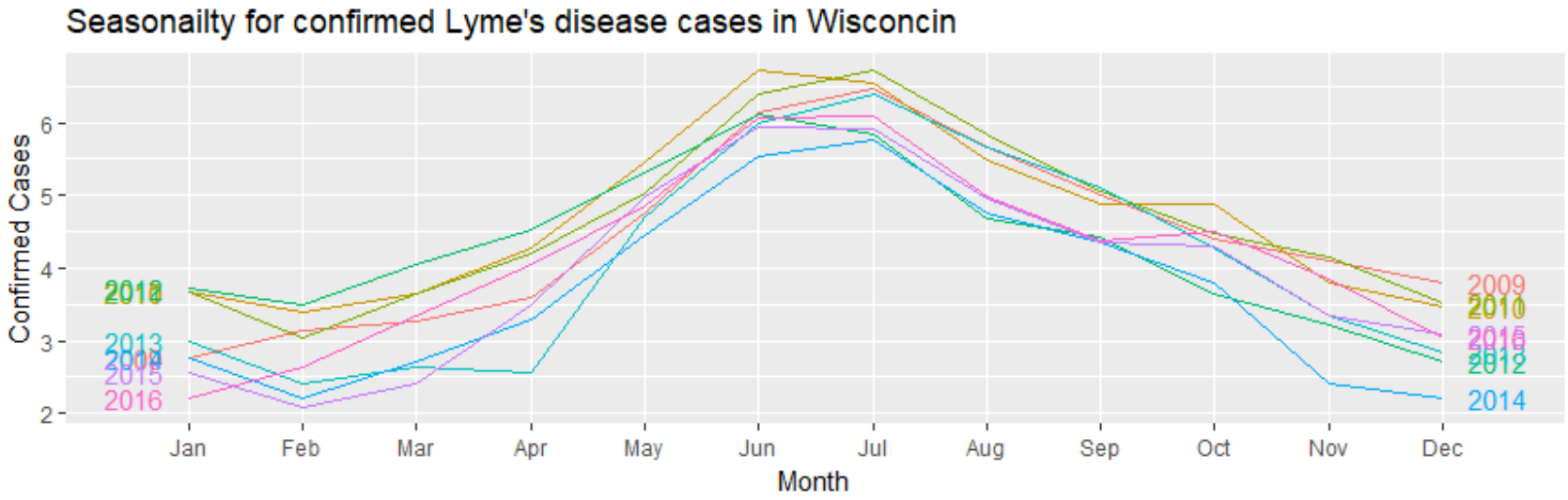
Table of Contents

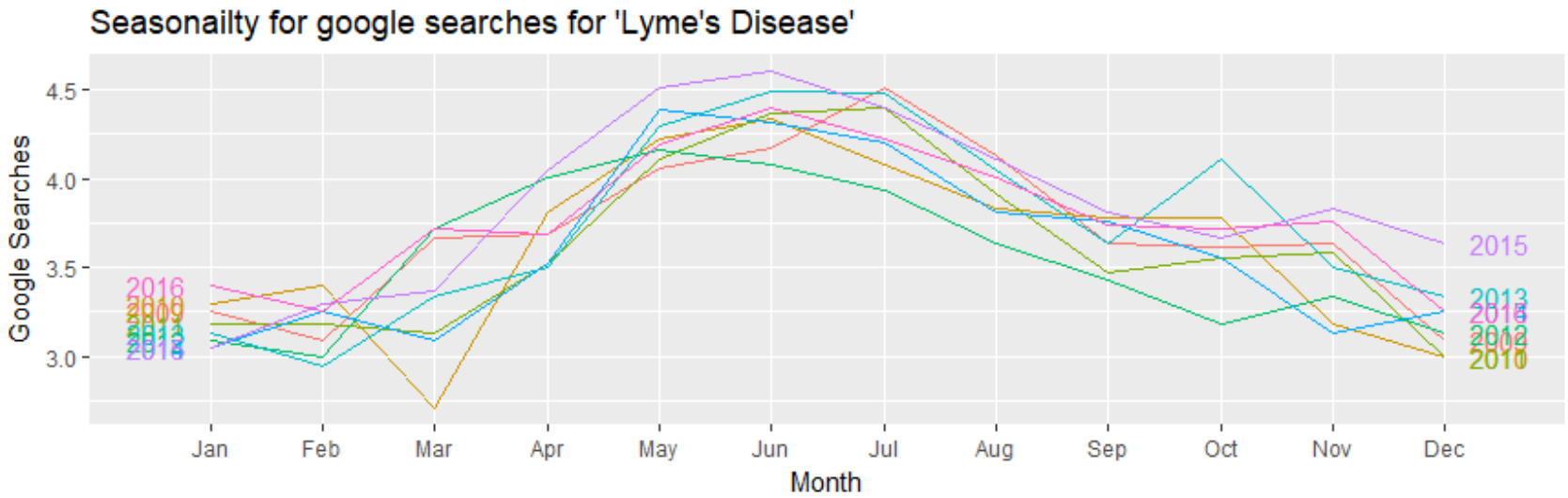
Initial Plots



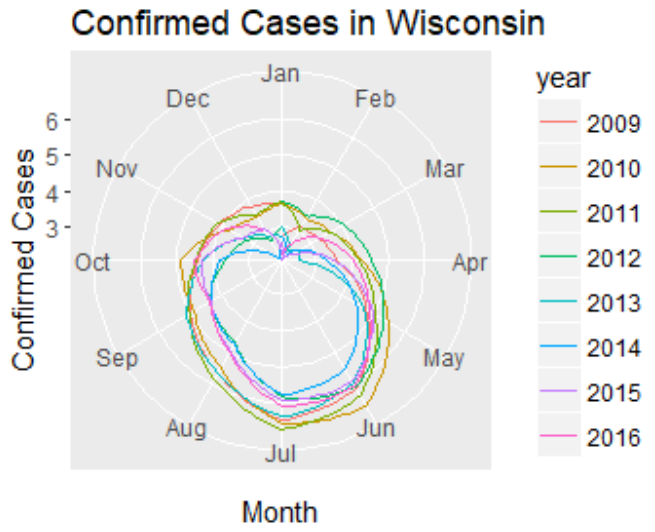


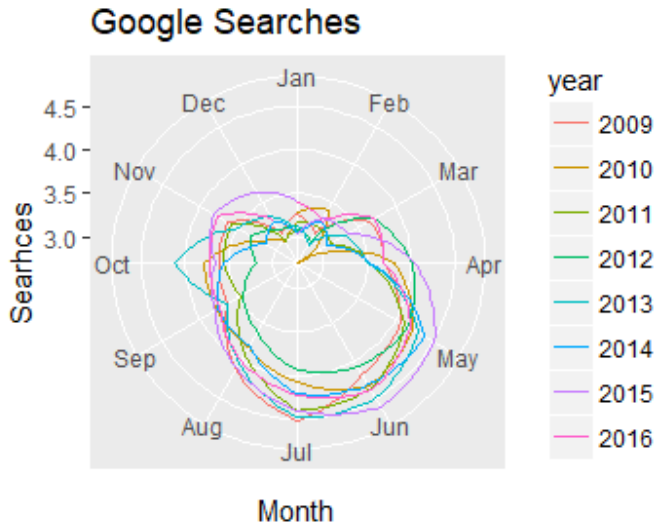
Seasonal Plots



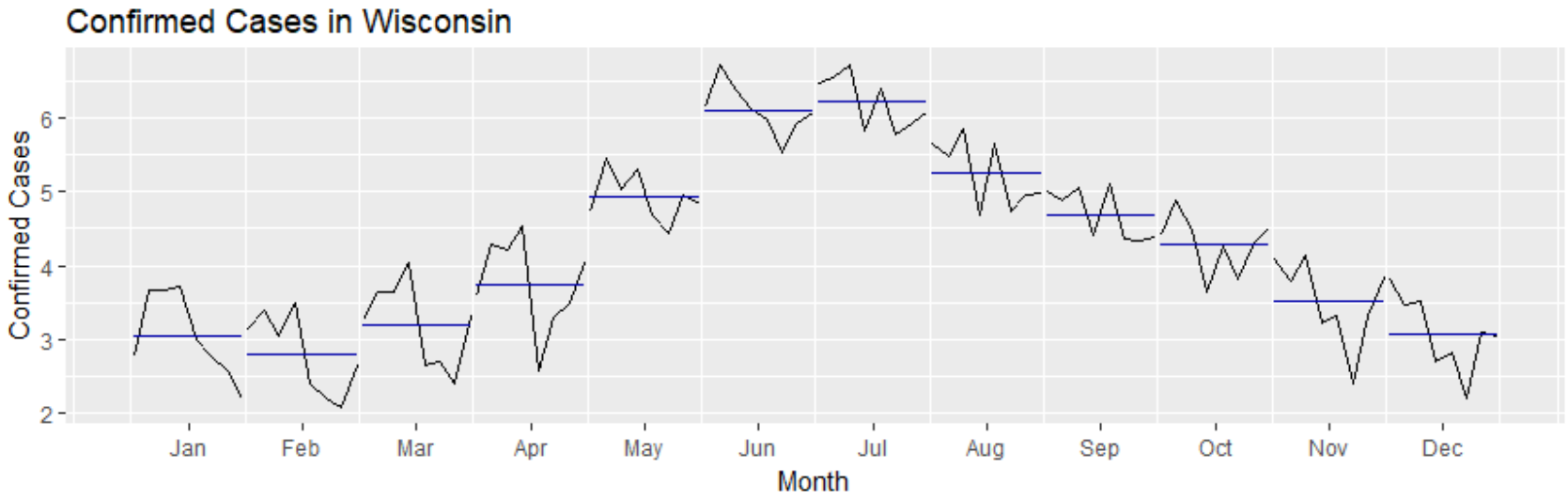


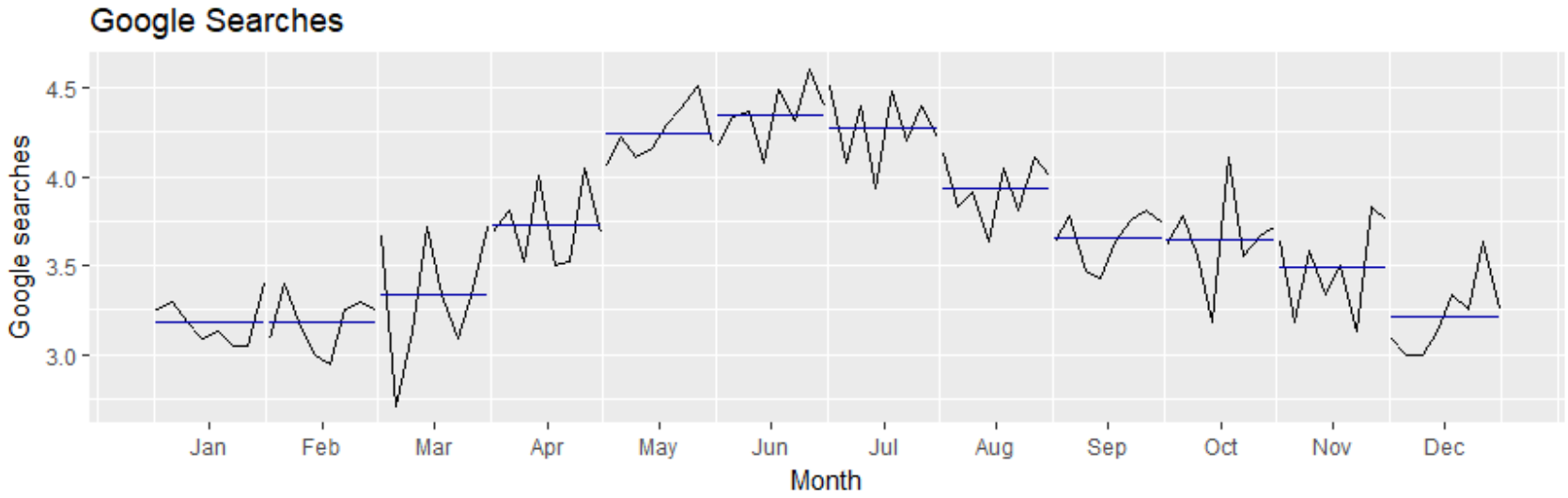
Polar Plots


Confirmed Cases in Wisconsin
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Confirmed Cases
3
4
5
6
Month
year
2009
2010
2011
2012
2013
2014
2015
2016

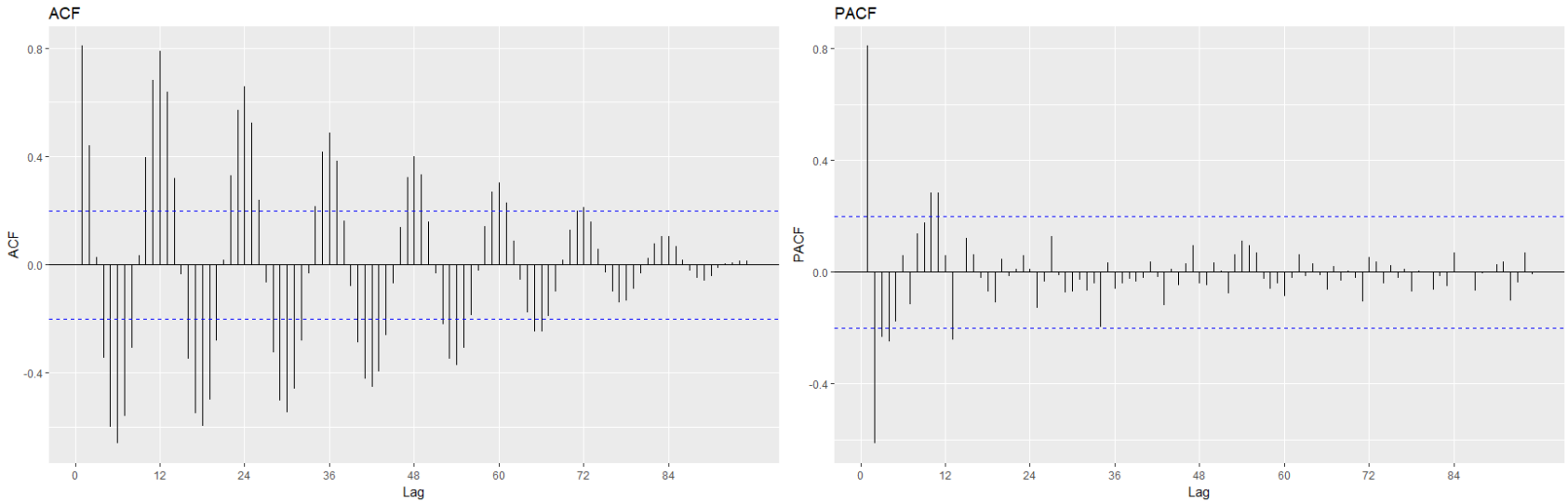

Google Searches
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Searhces
3.0
3.5
4.0
4.5
Month
year
2009
2010
2011
2012
2013
2014
2015
2016

Subseries Plots





ACF and PACF


ACF
0.8
0.4
0.0
-0.4
ACF
0
12
24
36
48
60
72
84
Lag
PACF
0.8
0.4
0.0
-0.4
PACF
0
12
24
36
48
60
72
84
Lag

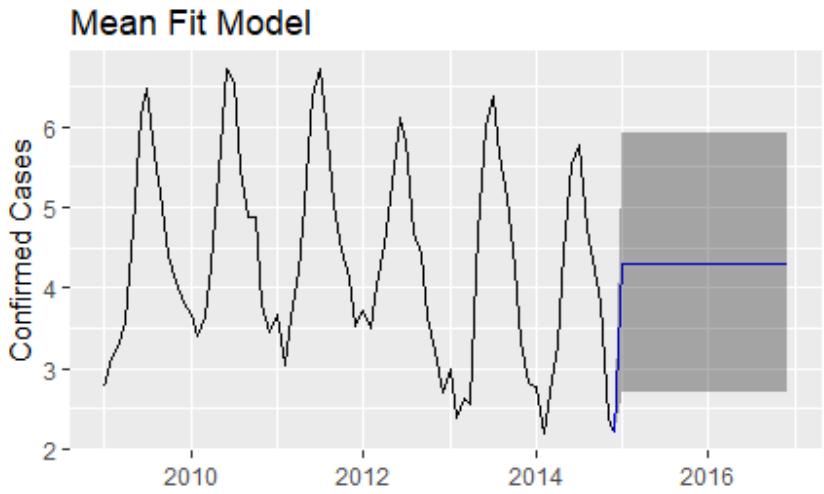
Reports Stationarity

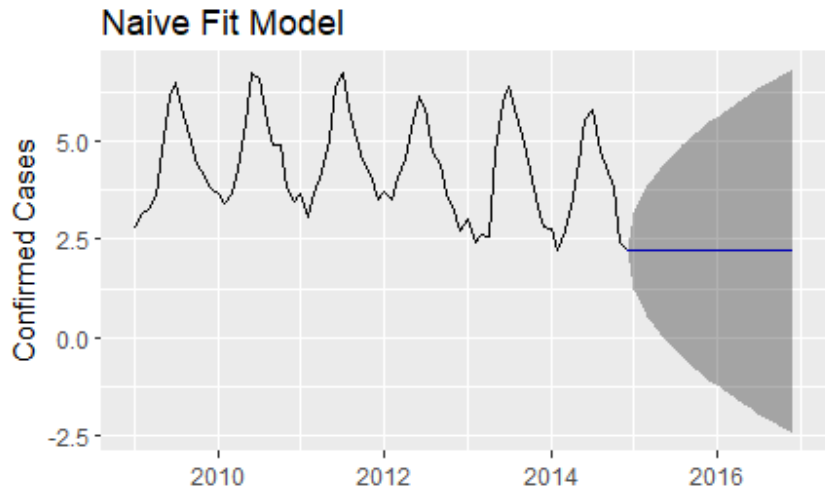
```
> adf.test(x = lyme.ts)

Augmented Dickey-Fuller Test

data:  lyme.ts
Dickey-Fuller = -7.8019, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(x = lyme.ts) : p-value smaller than printed p-value
> ndiffs(lyme.ts)
[1] 0
```

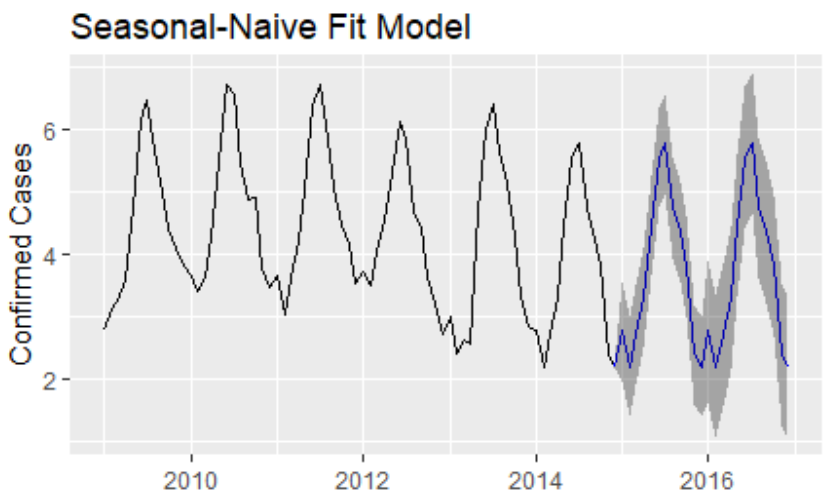
Google Trends Stationarity

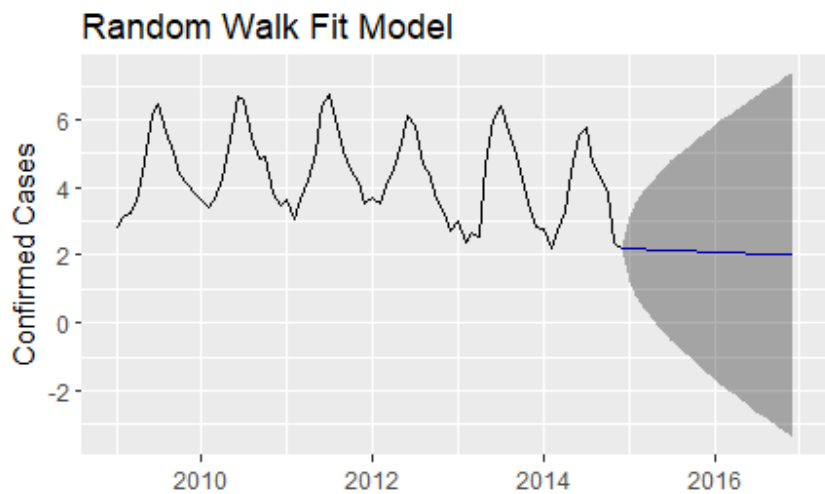
```
> adf.test(x = Searches.ts)

Augmented Dickey-Fuller Test

data:  Searches.ts
Dickey-Fuller = -7.0788, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(x = Searches.ts) : p-value smaller than printed p-value
> ndiffs(Searches.ts)
[1] 0
```

Fitting Simple Methods

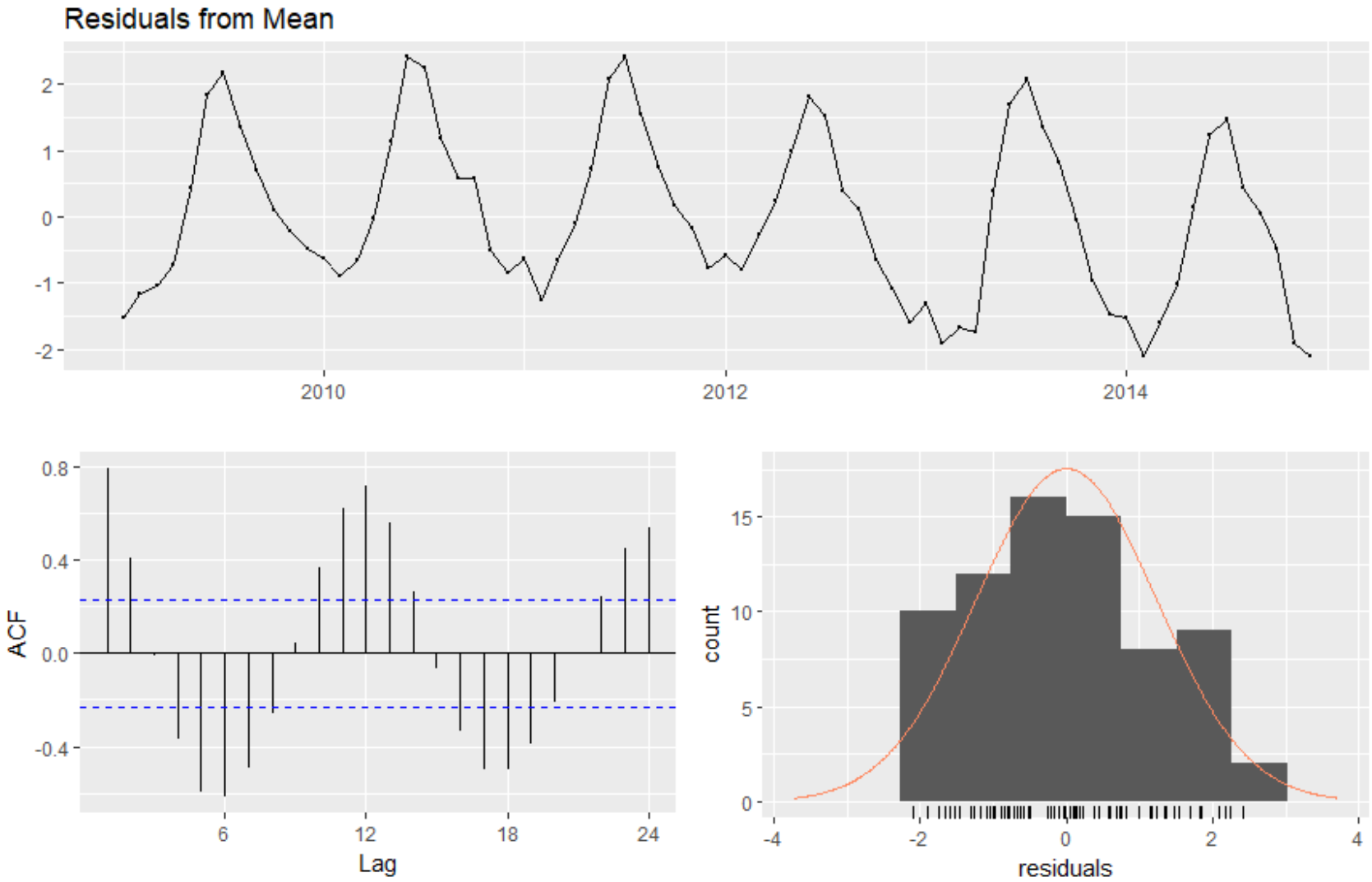




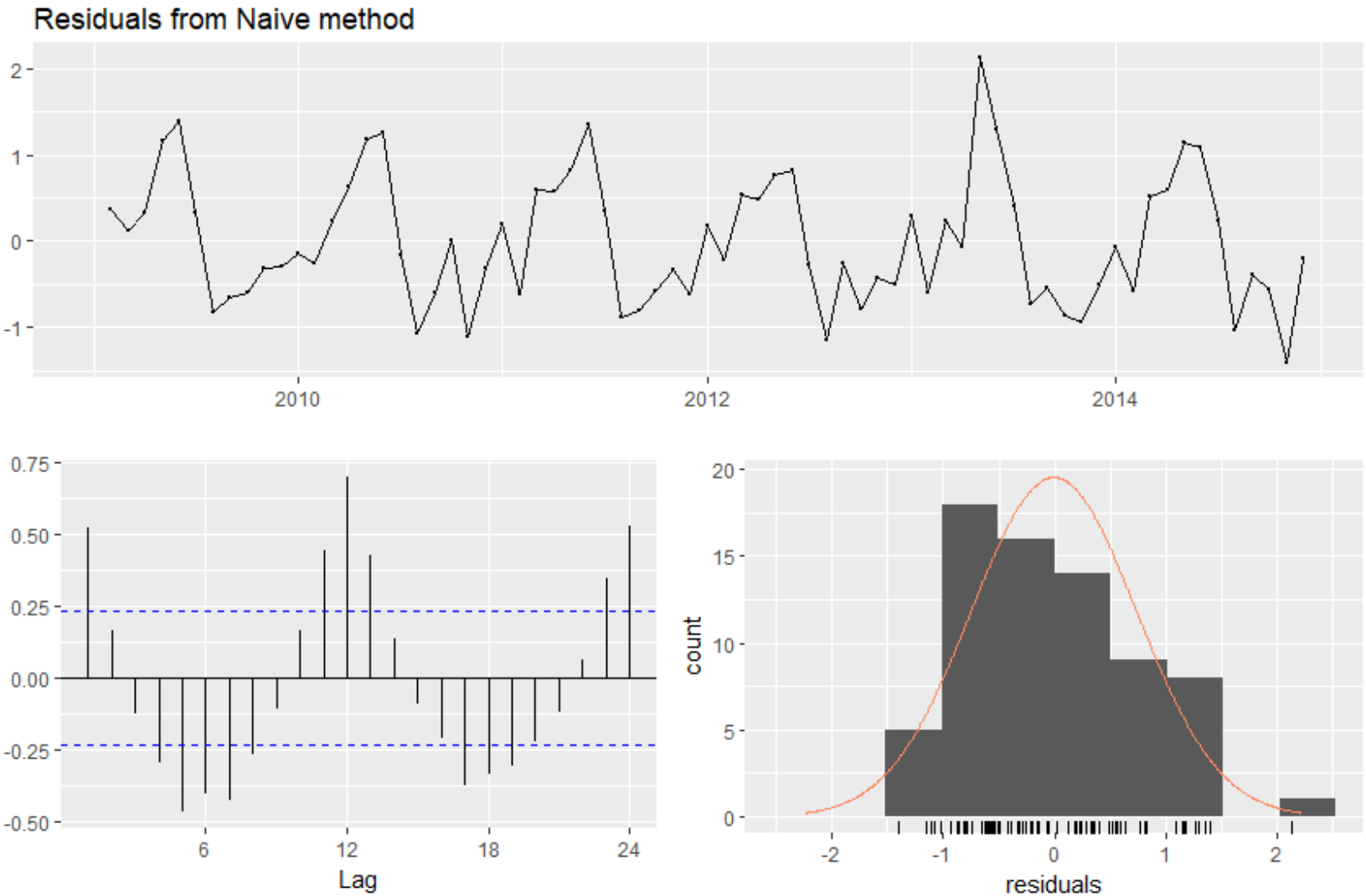




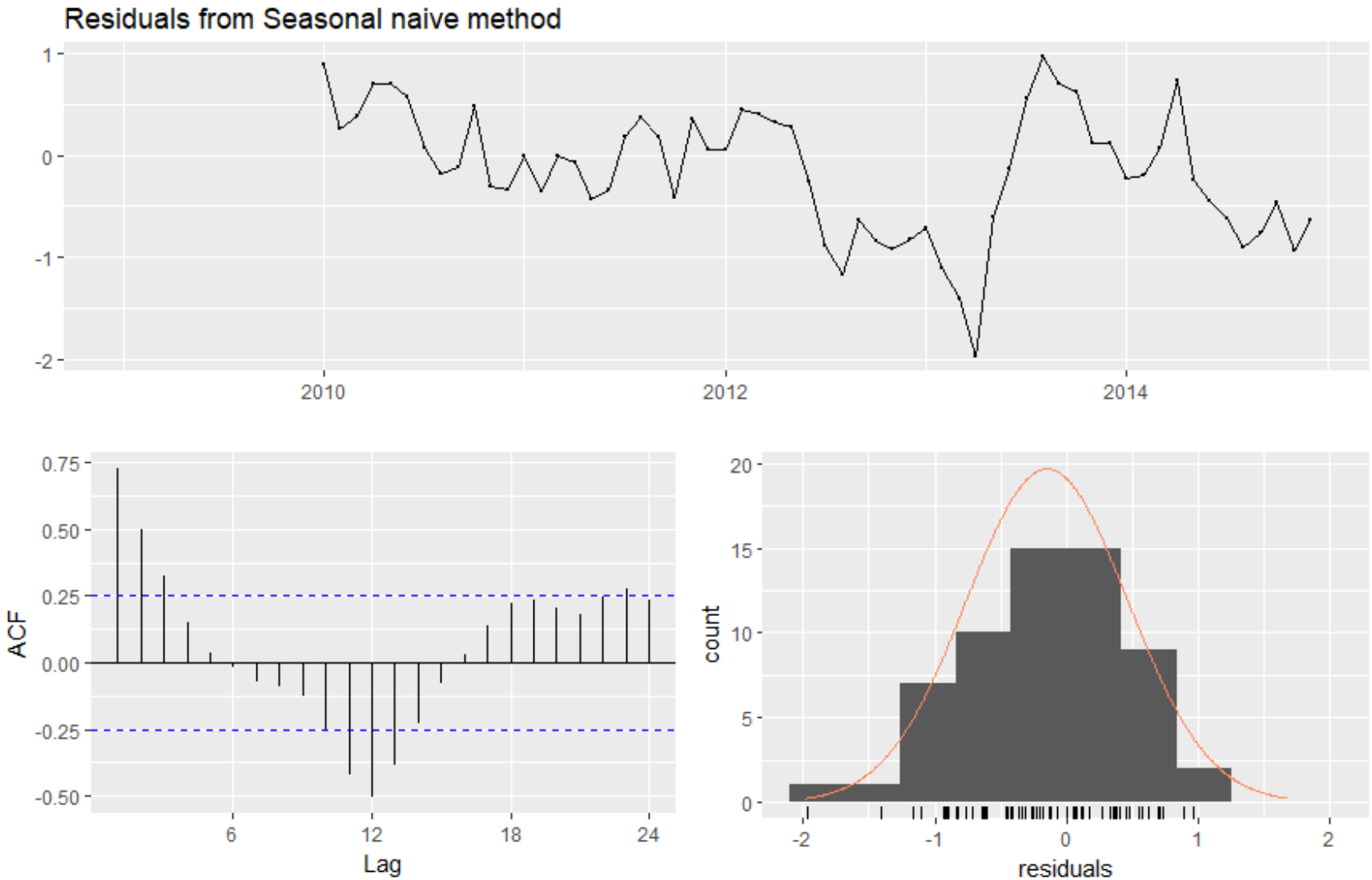
Mean Residuals



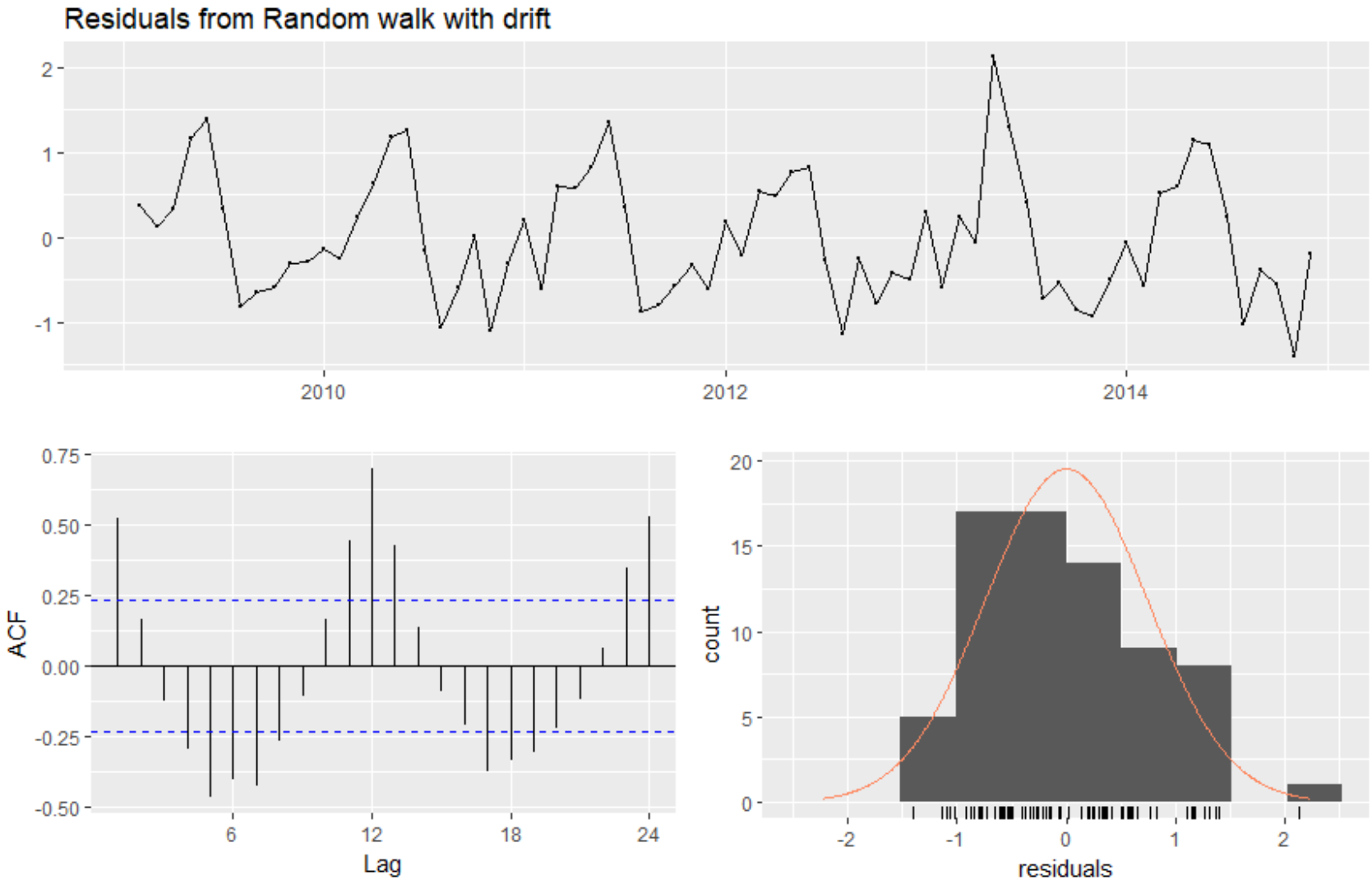
Naive Residuals



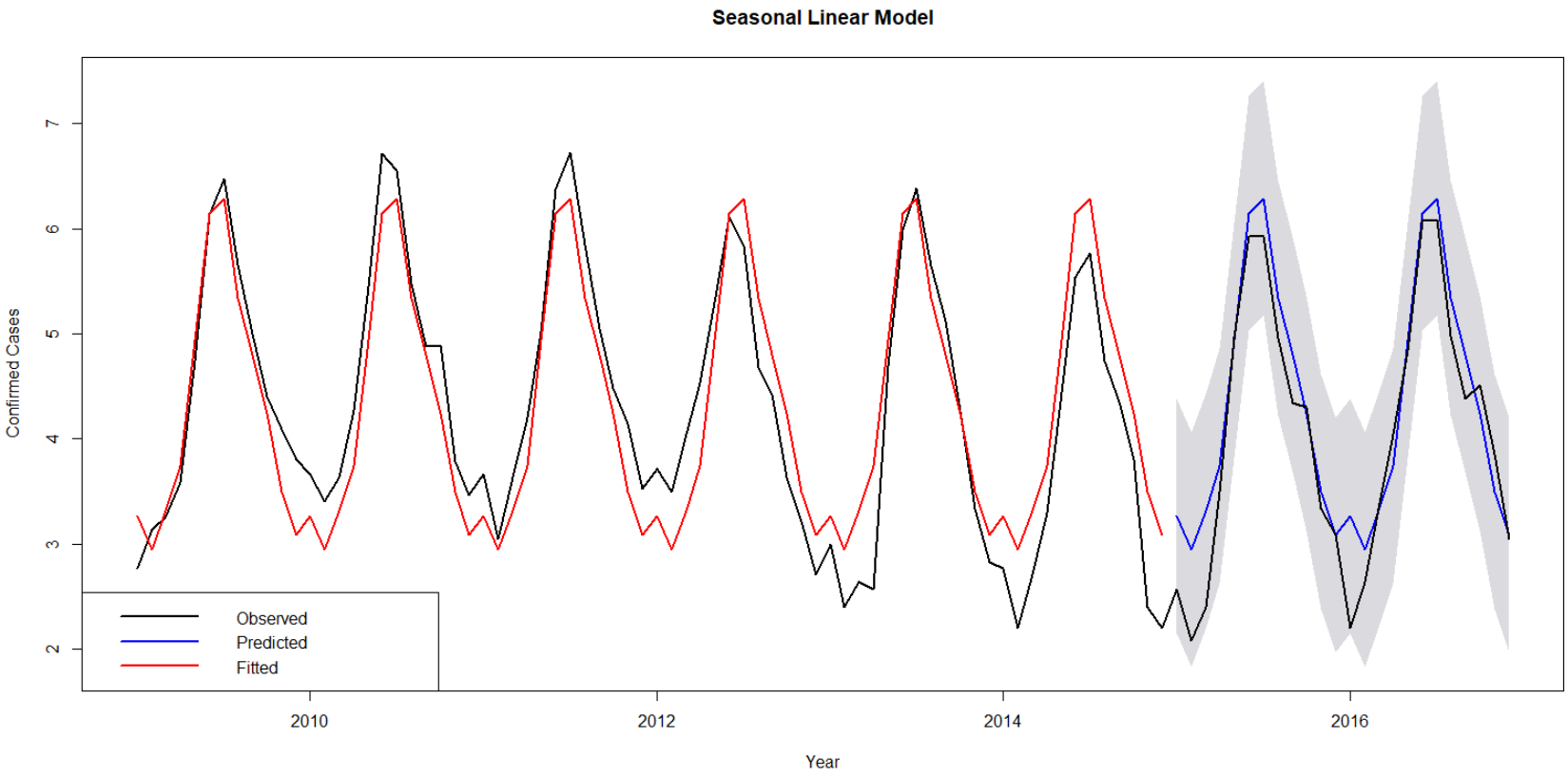
Seasonal Naive Residuals



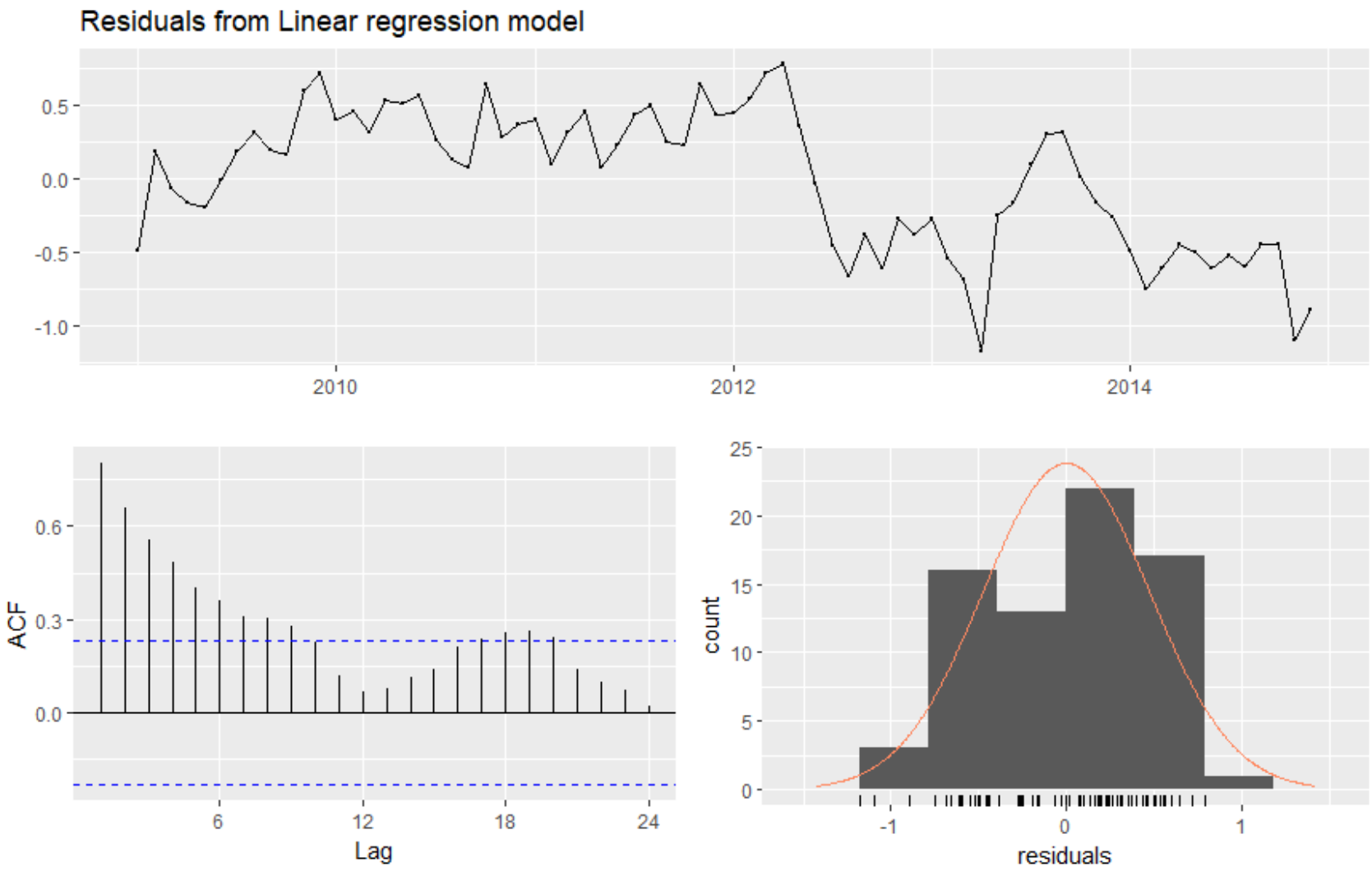
Random Walk Residuals



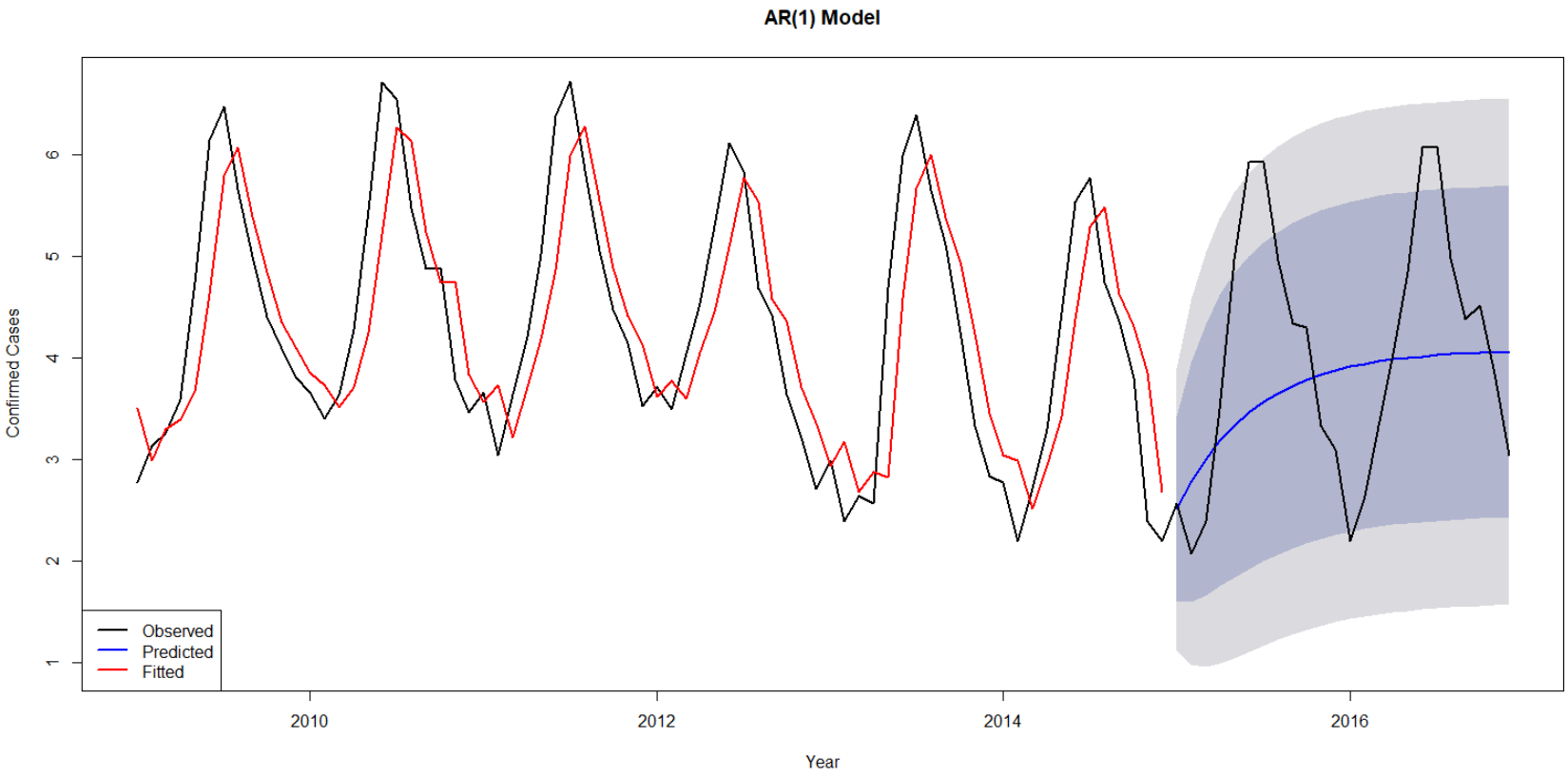
Fitting Seasonal Linear Model



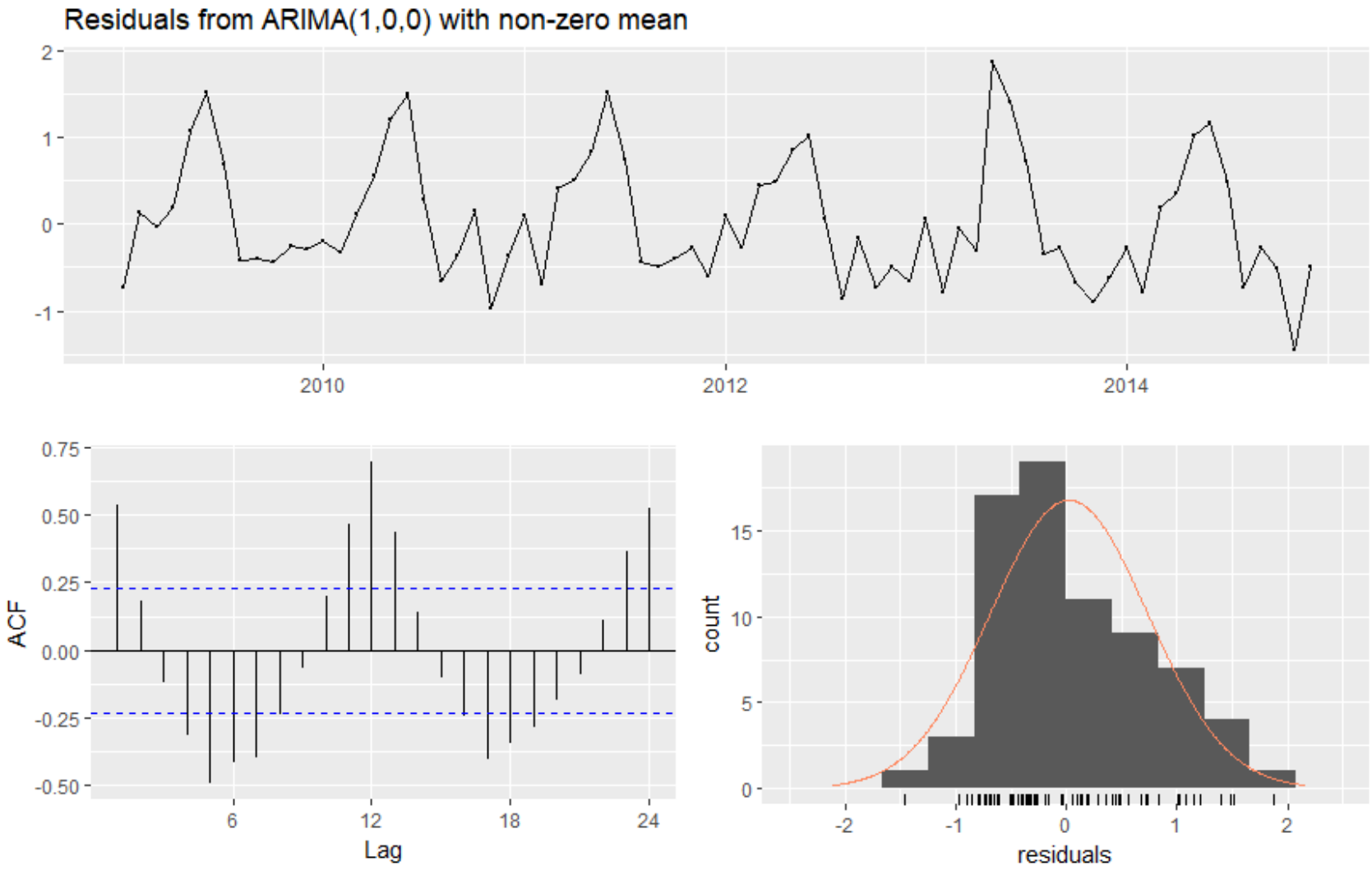
Seasonal Linear Model Residuals



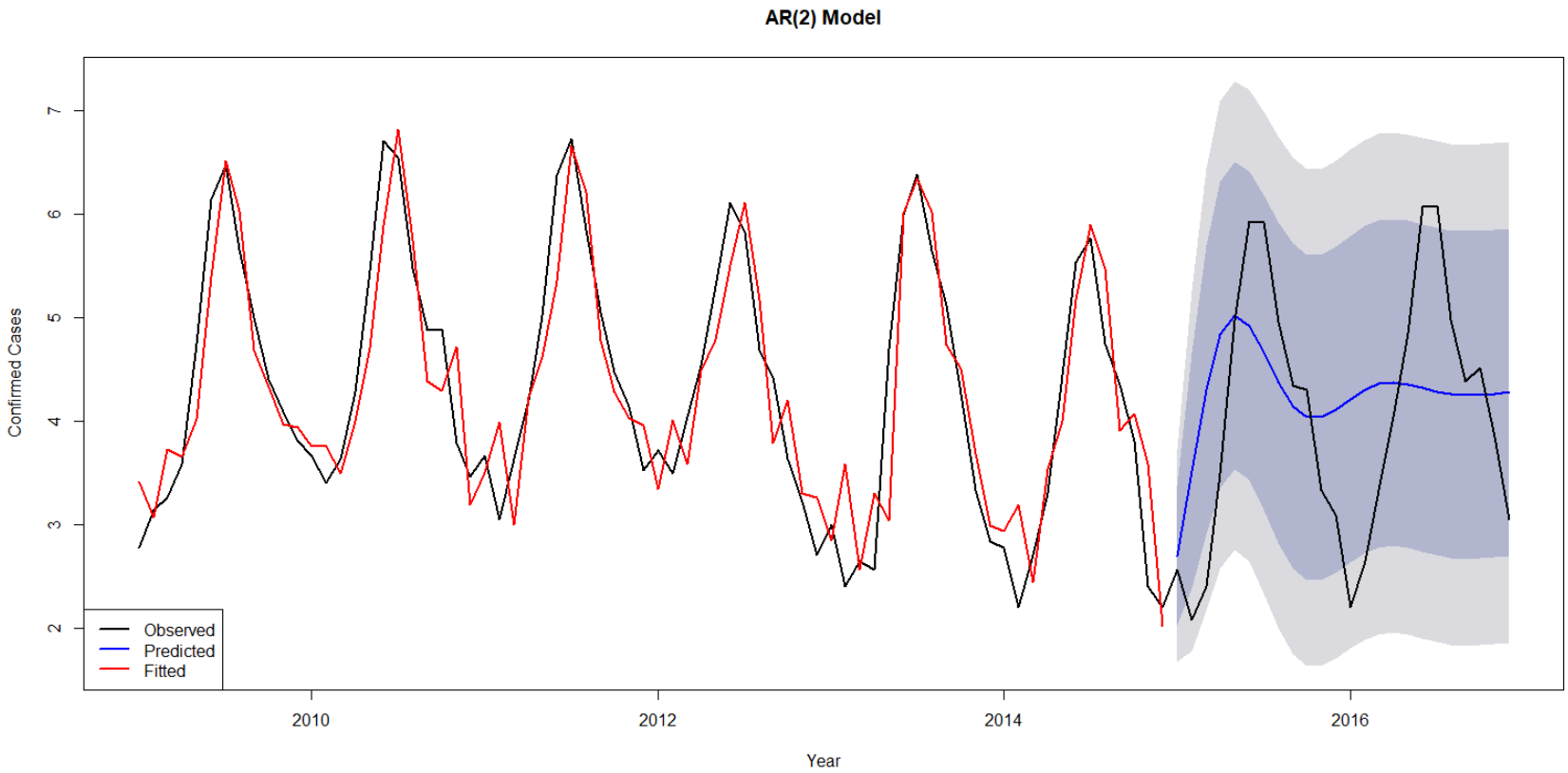
Fitting AR(1)


AR(1) Model
Confirmed Cases
1
2
3
4
5
6
2010
2012
2014
2016
Year
Observed
Predicted
Fitted

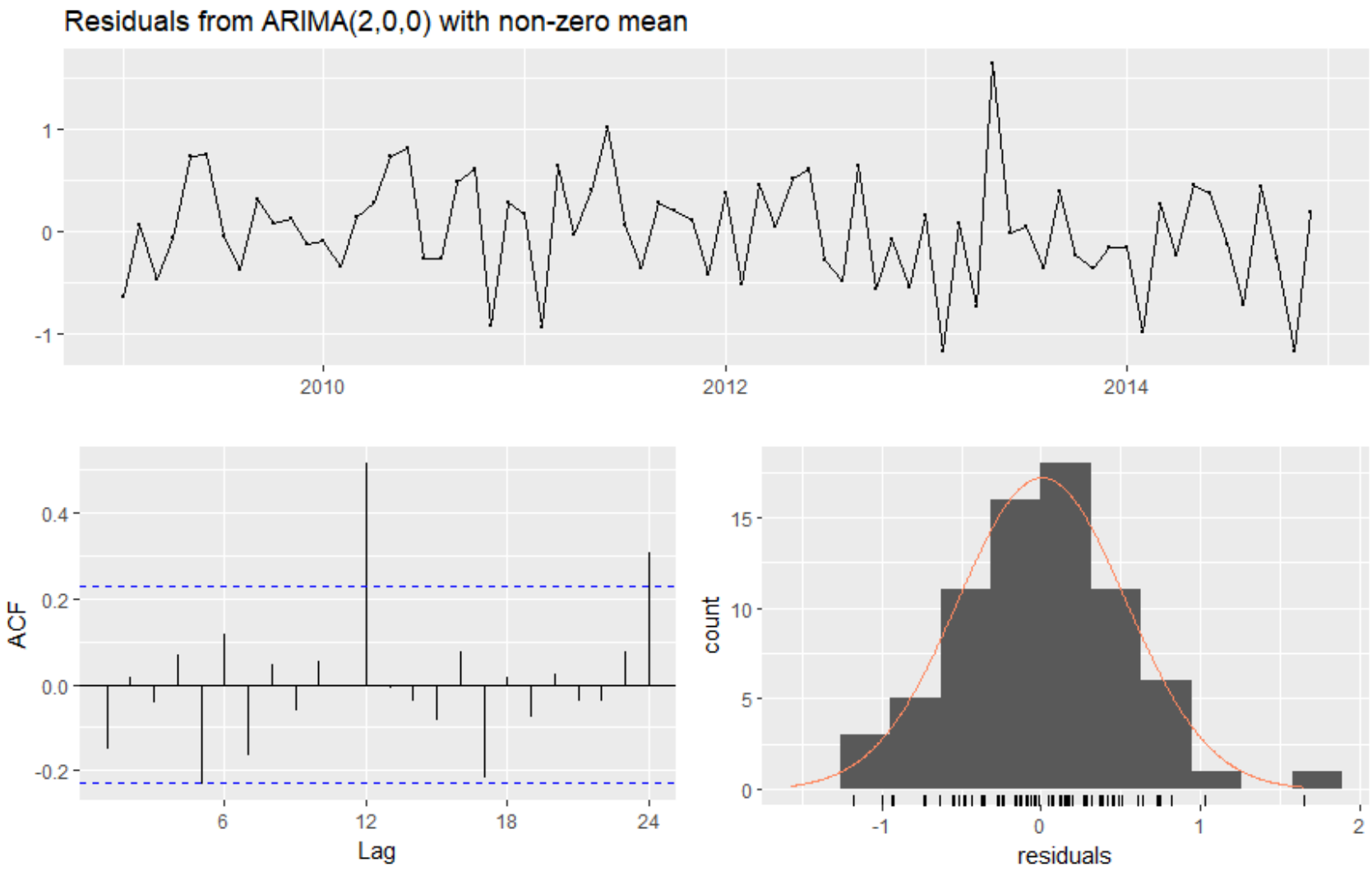
AR(1) Residuals



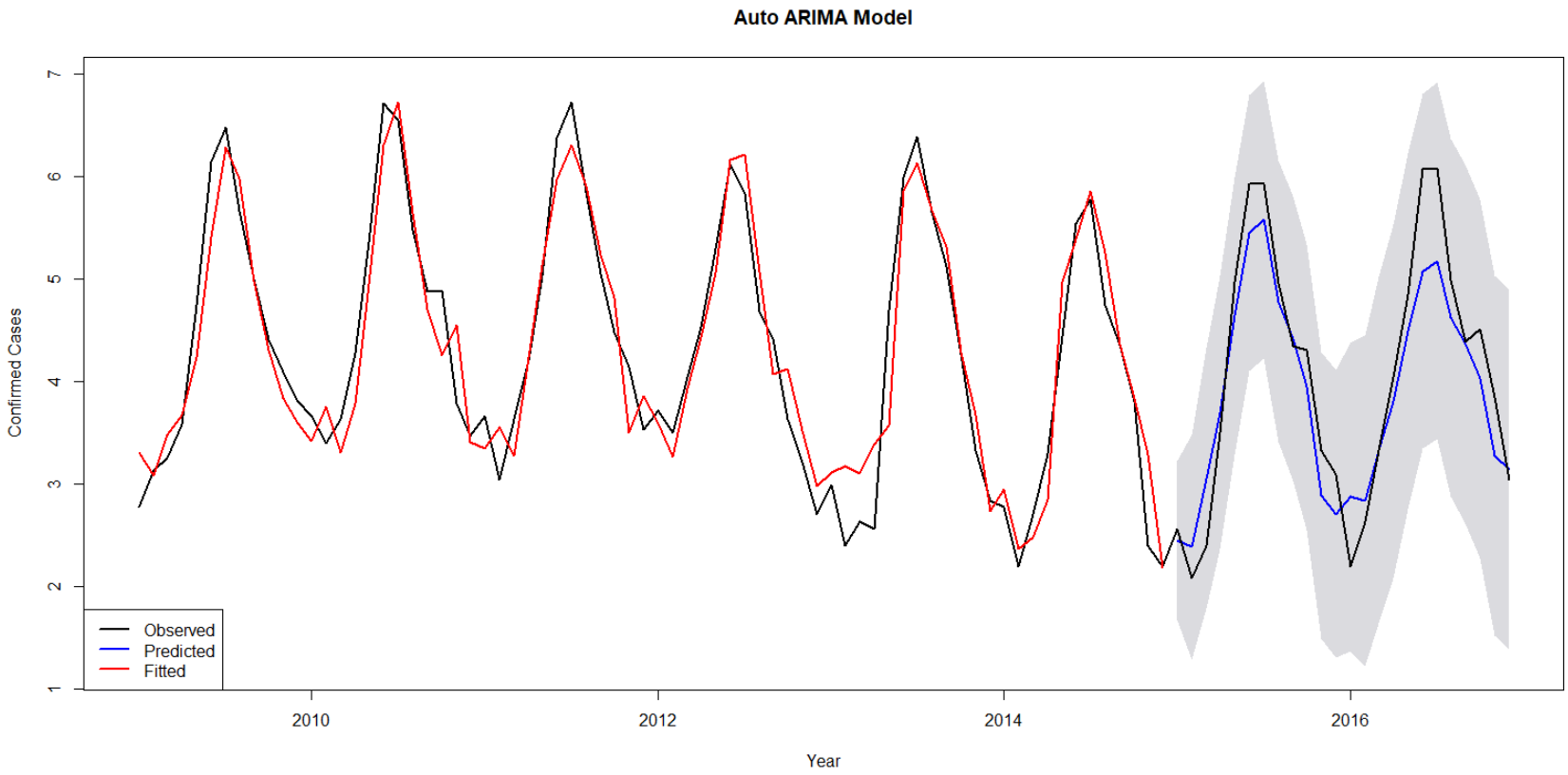
Fitting AR(2)


AR(2) Model
Confirmed Cases
2
3
4
5
6
7
2010
2012
2014
2016
Year
Observed
Predicted
Fitted

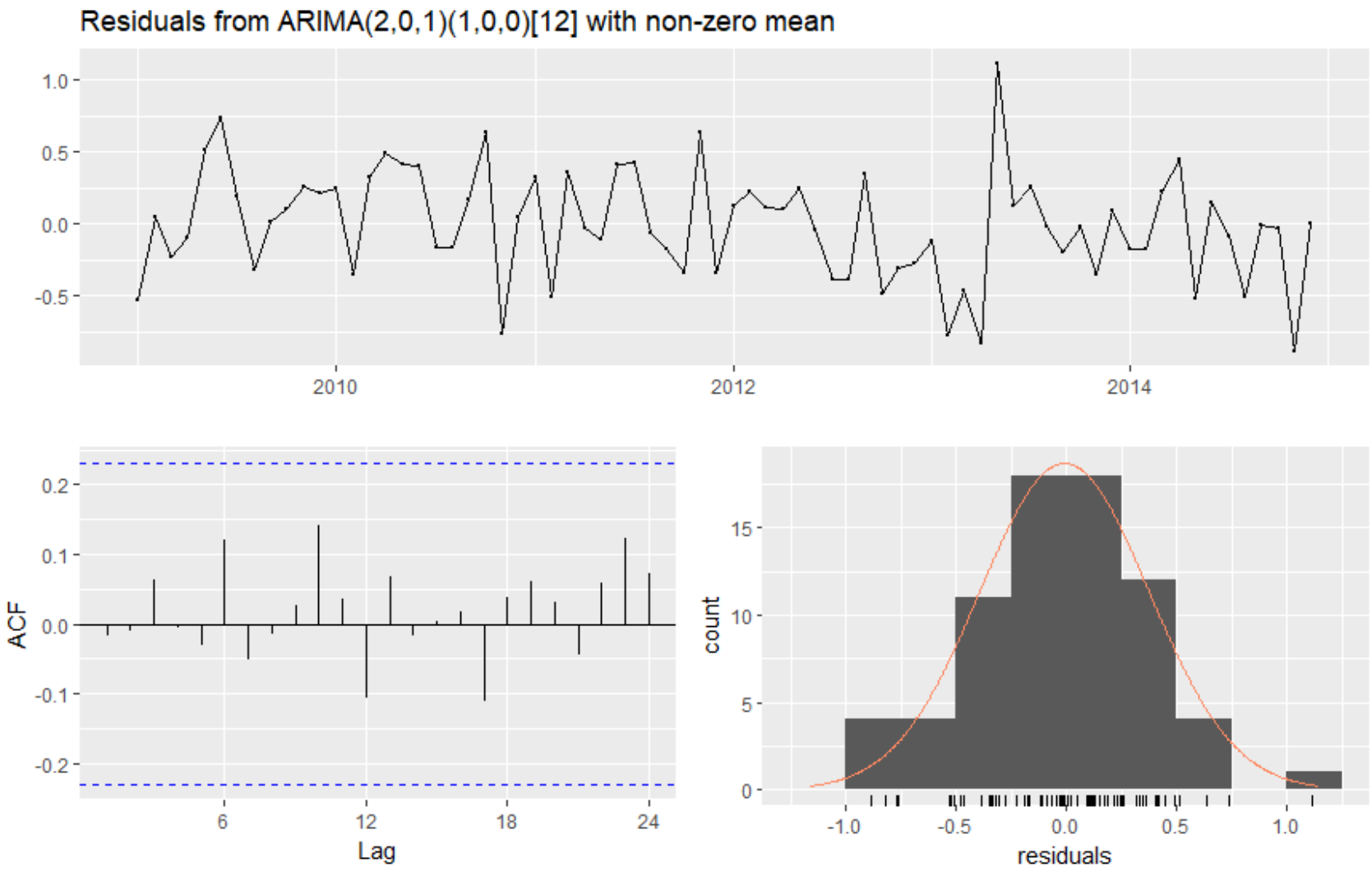
AR(2) Residuals



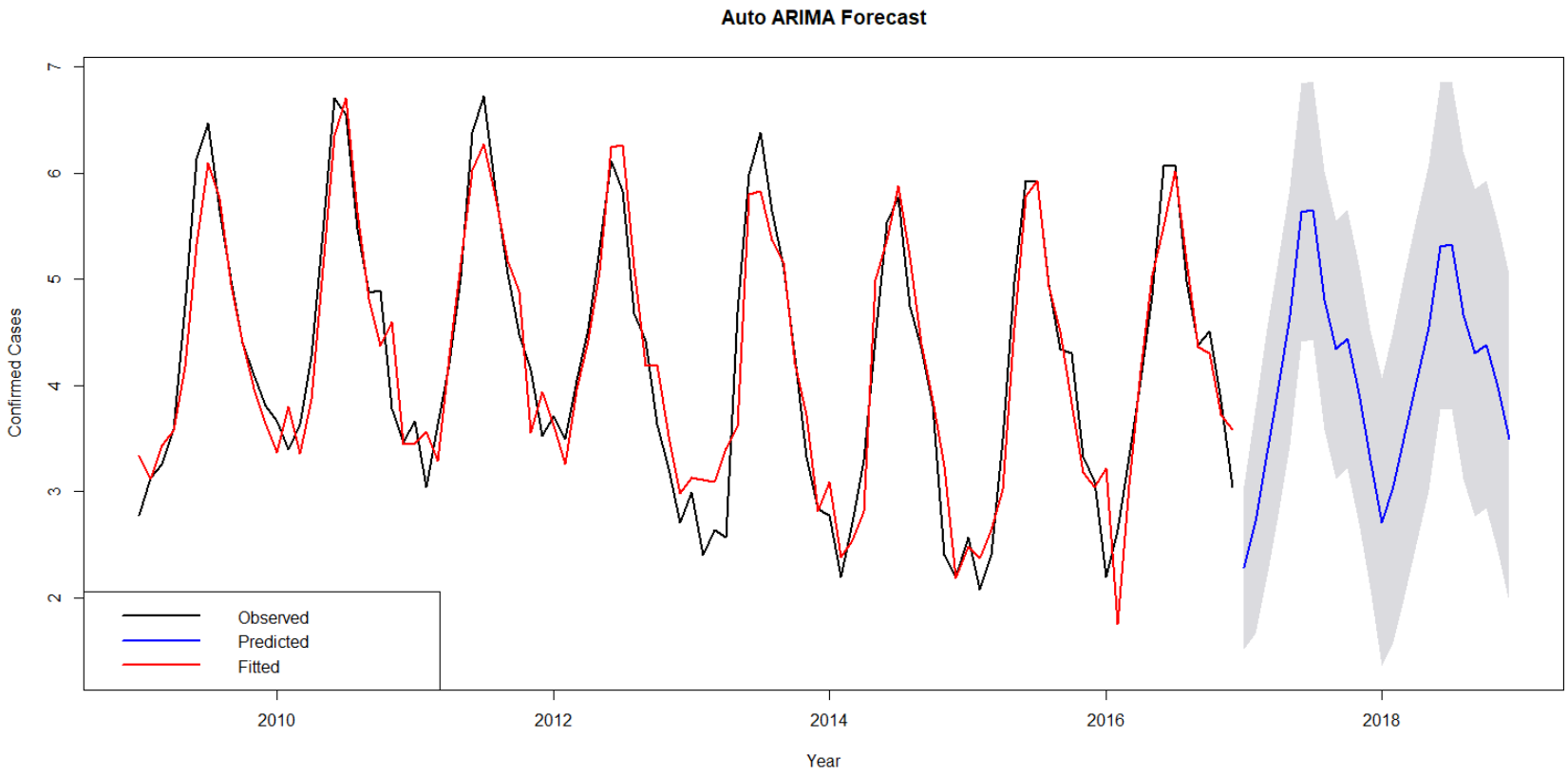
Fitting Auto ARIMA



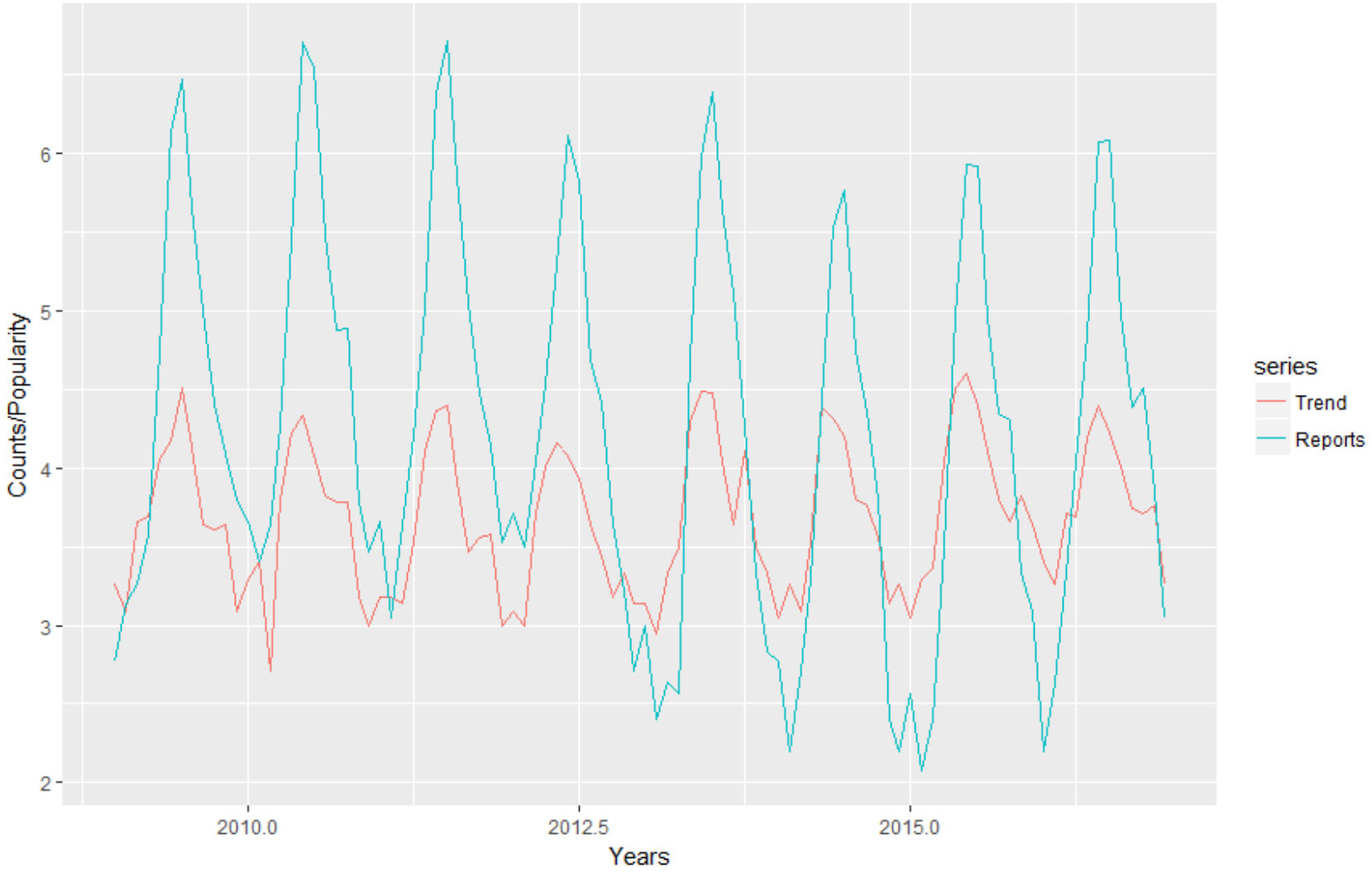
Auto ARIMA Residuals



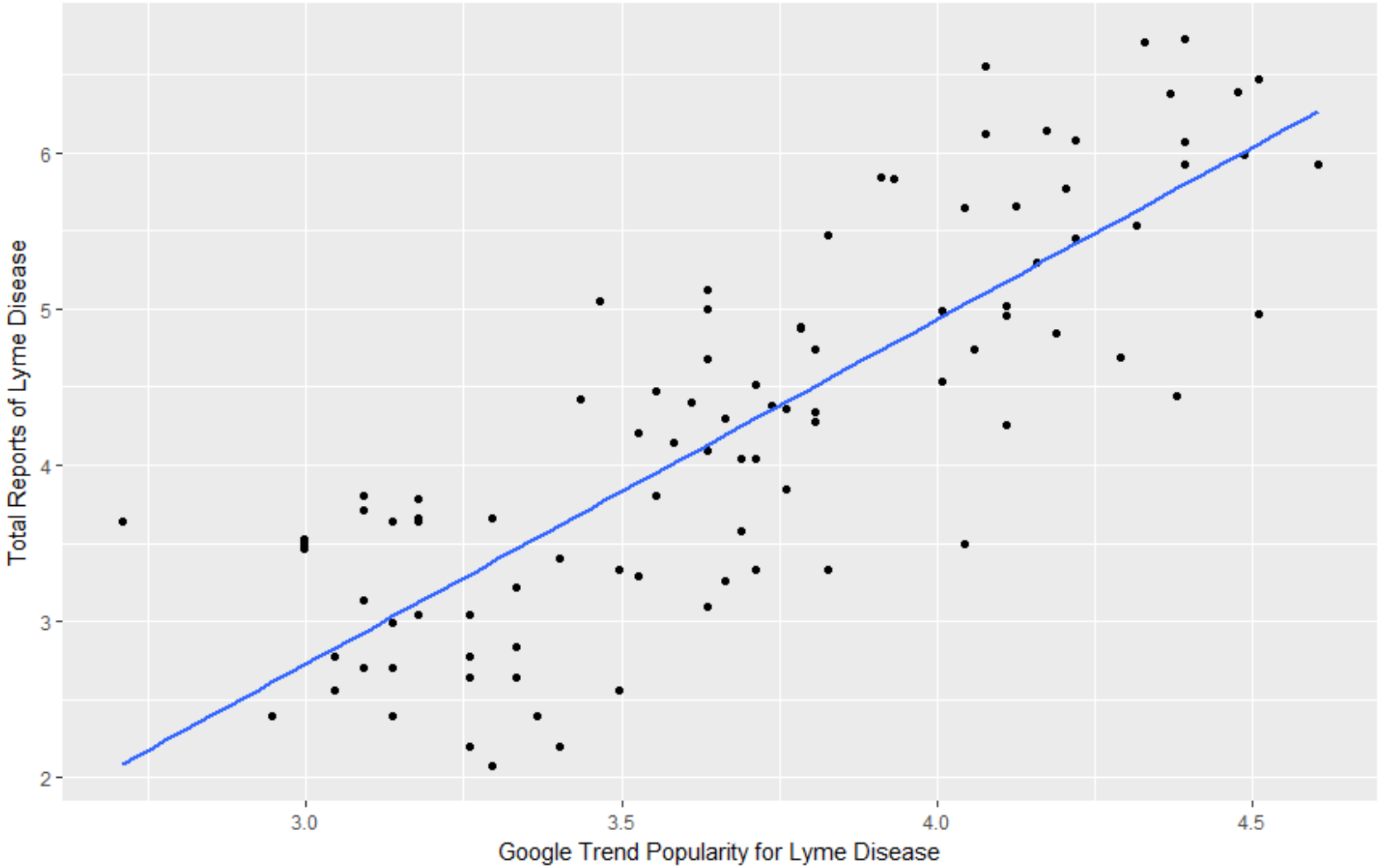
Auto ARIMA Forecast



Data Overlay



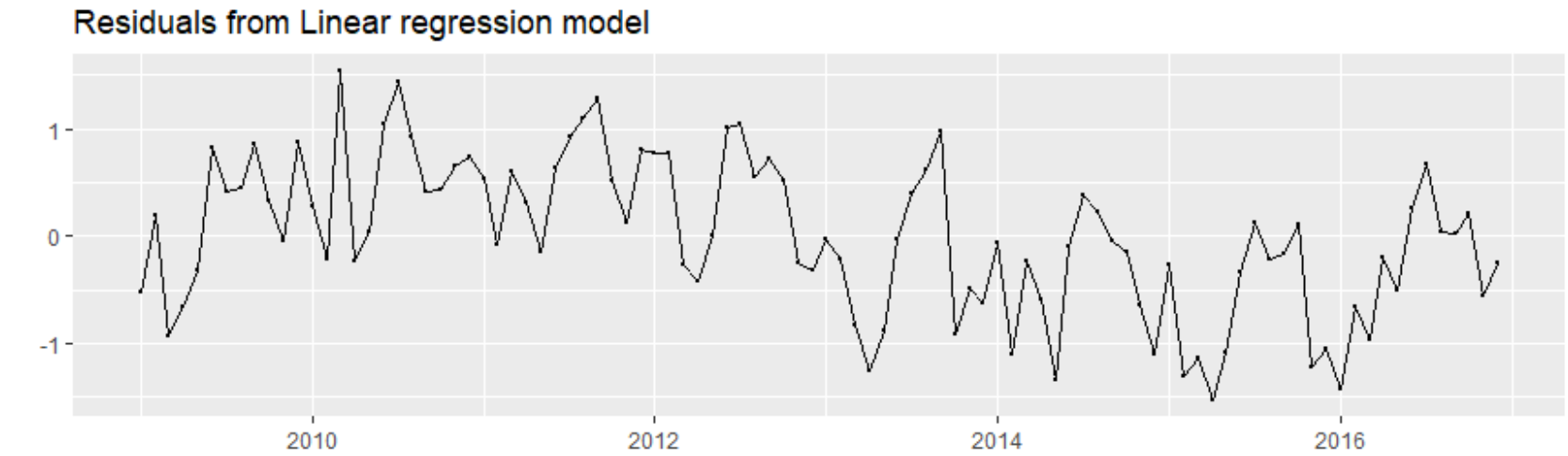
Linear Plot

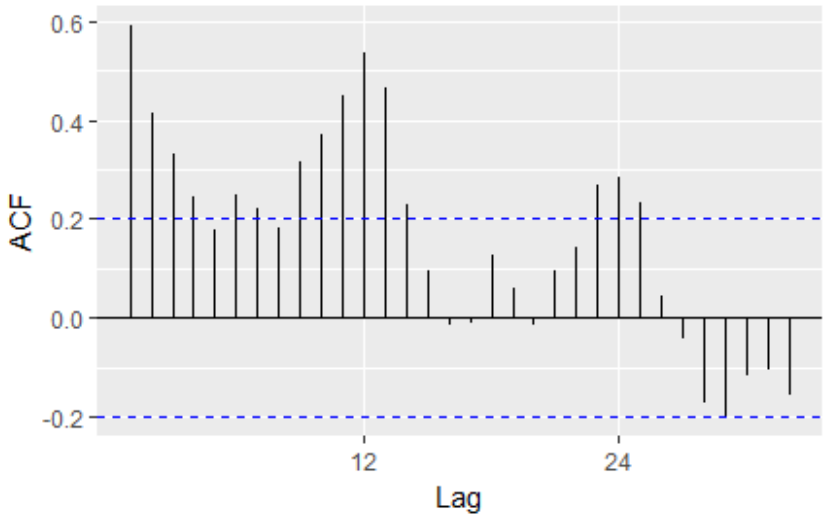

Total Reports of Lyme Disease
6
5
4
3
2
3.0
3.5
4.0
4.5
Google Trend Popularity for Lyme Disease

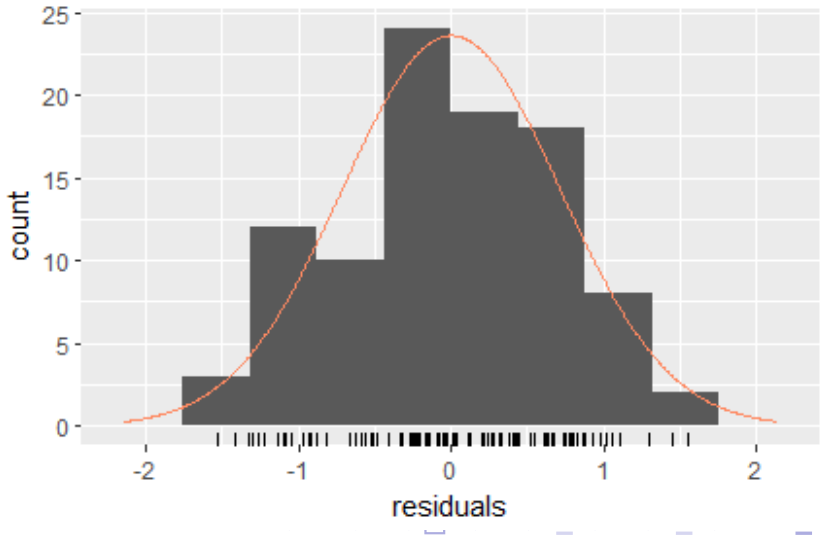
Summary Output

```
Call:
tslm(formula = Reports ~ Trend, data = master.ts)

Residuals:
     Min       1Q   Median       3Q      Max
-1.53066 -0.49533 -0.02761  0.54224  1.54912

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -3.8727     0.5944  -6.516 3.56e-09 ***
Trend         2.2013     0.1600  13.755  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7178 on 94 degrees of freedom
Multiple R-squared:  0.6681,Adjusted R-squared:  0.6646
F-statistic: 189.2 on 1 and 94 DF,  p-value: < 2.2e-16
```

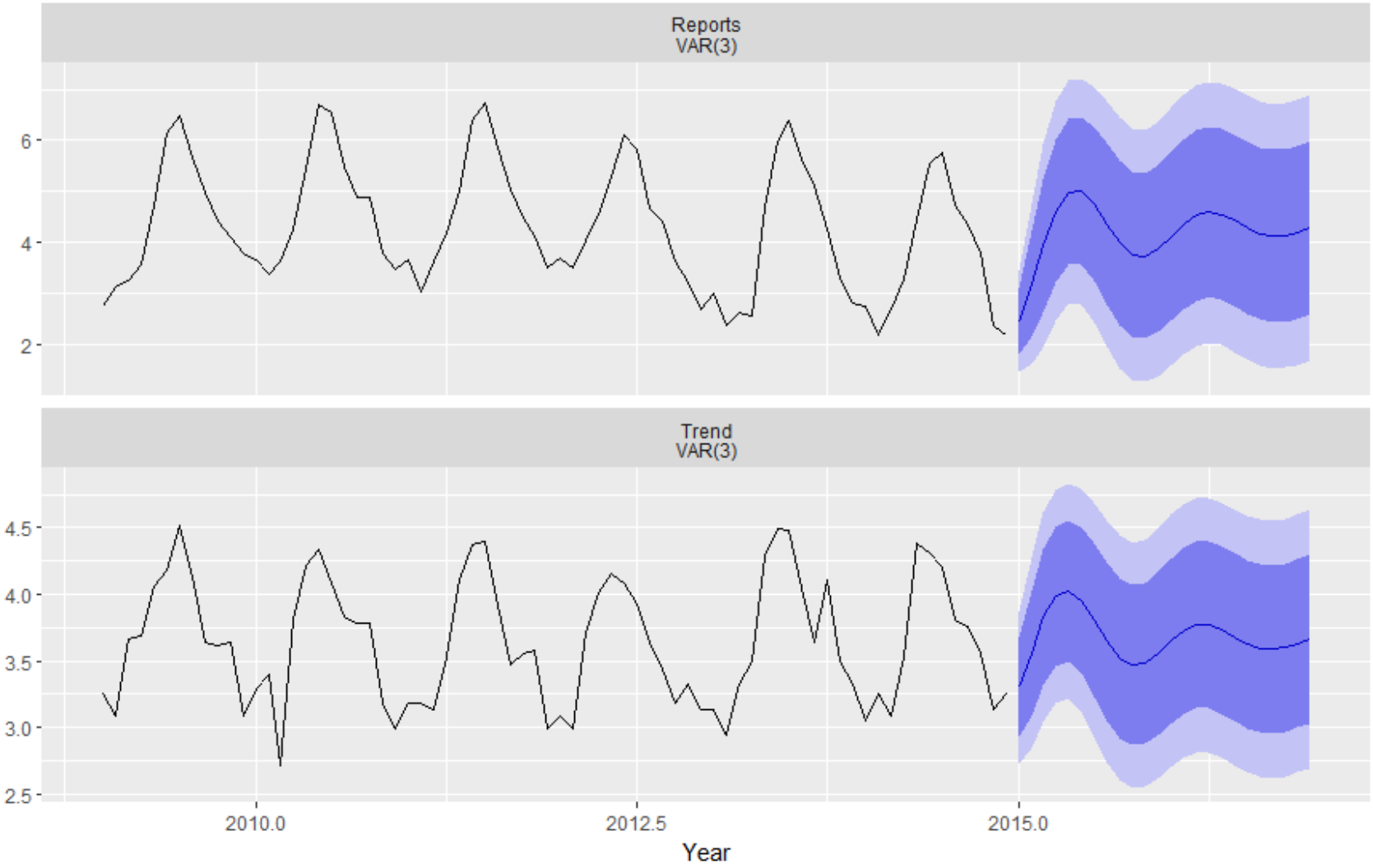
Linear Residuals



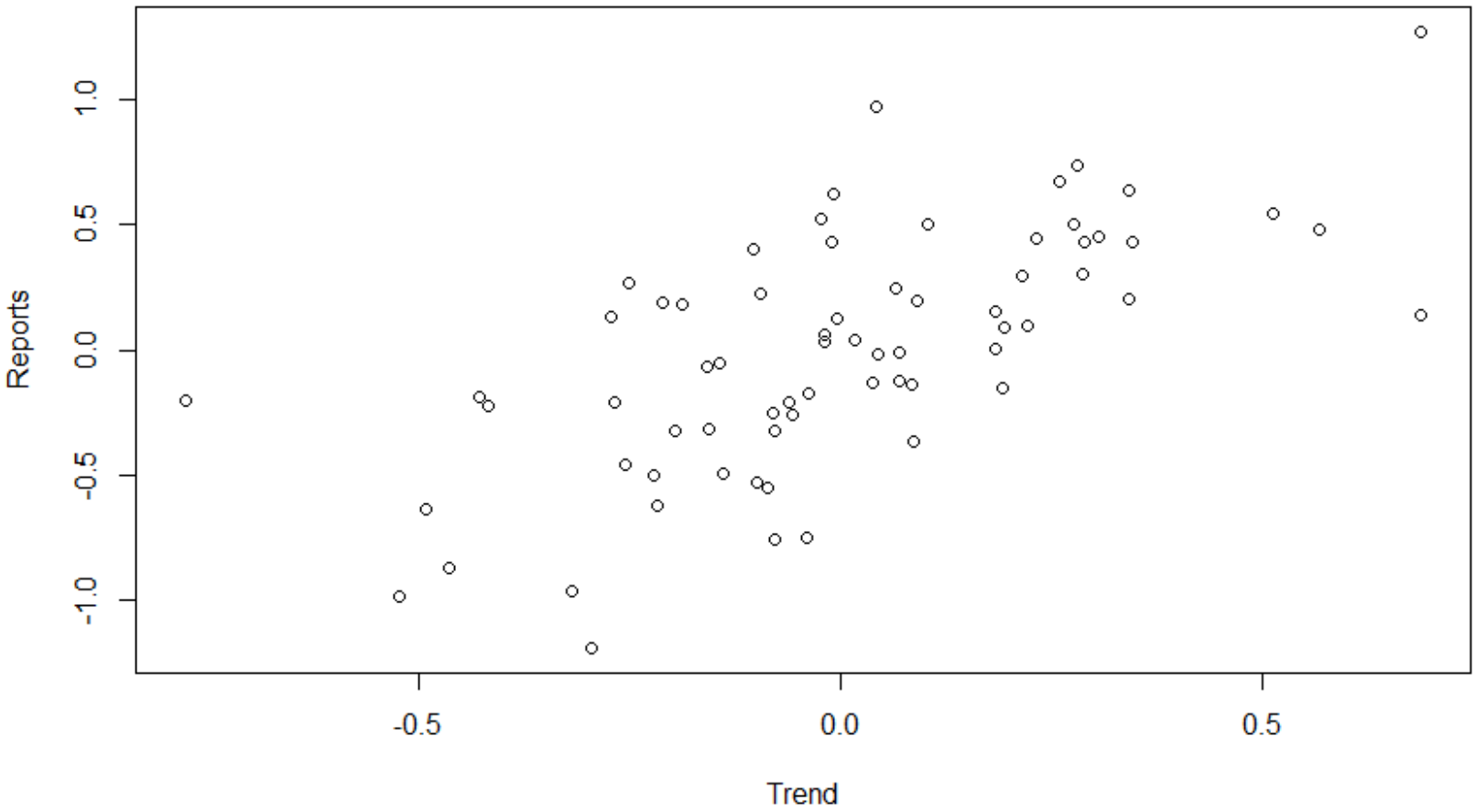




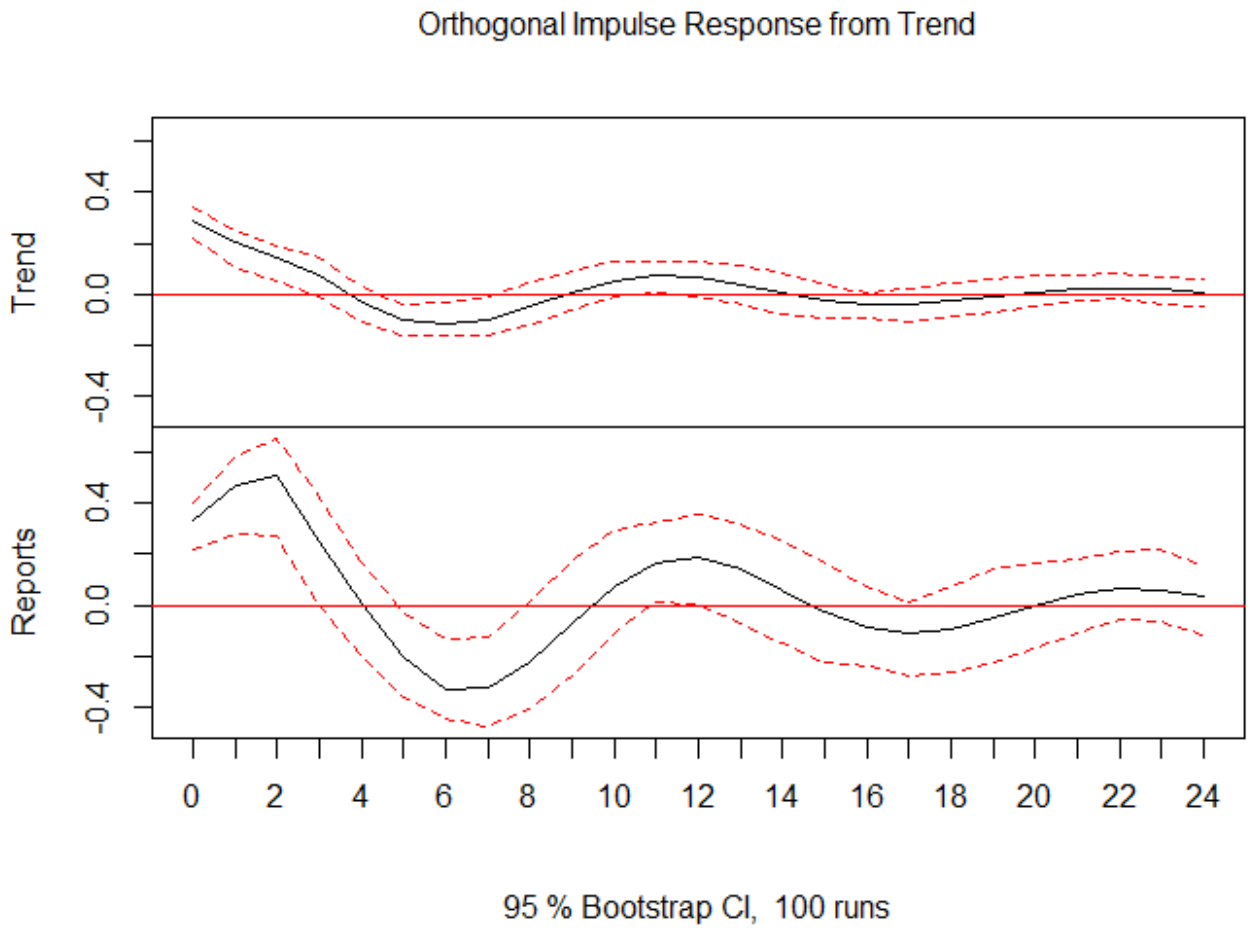
Fitting VAR


Reports
VAR(3)
6
4
2
Trend
VAR(3)
4.5
4.0
3.5
3.0
2.5
2010.0
2012.5
2015.0
Year

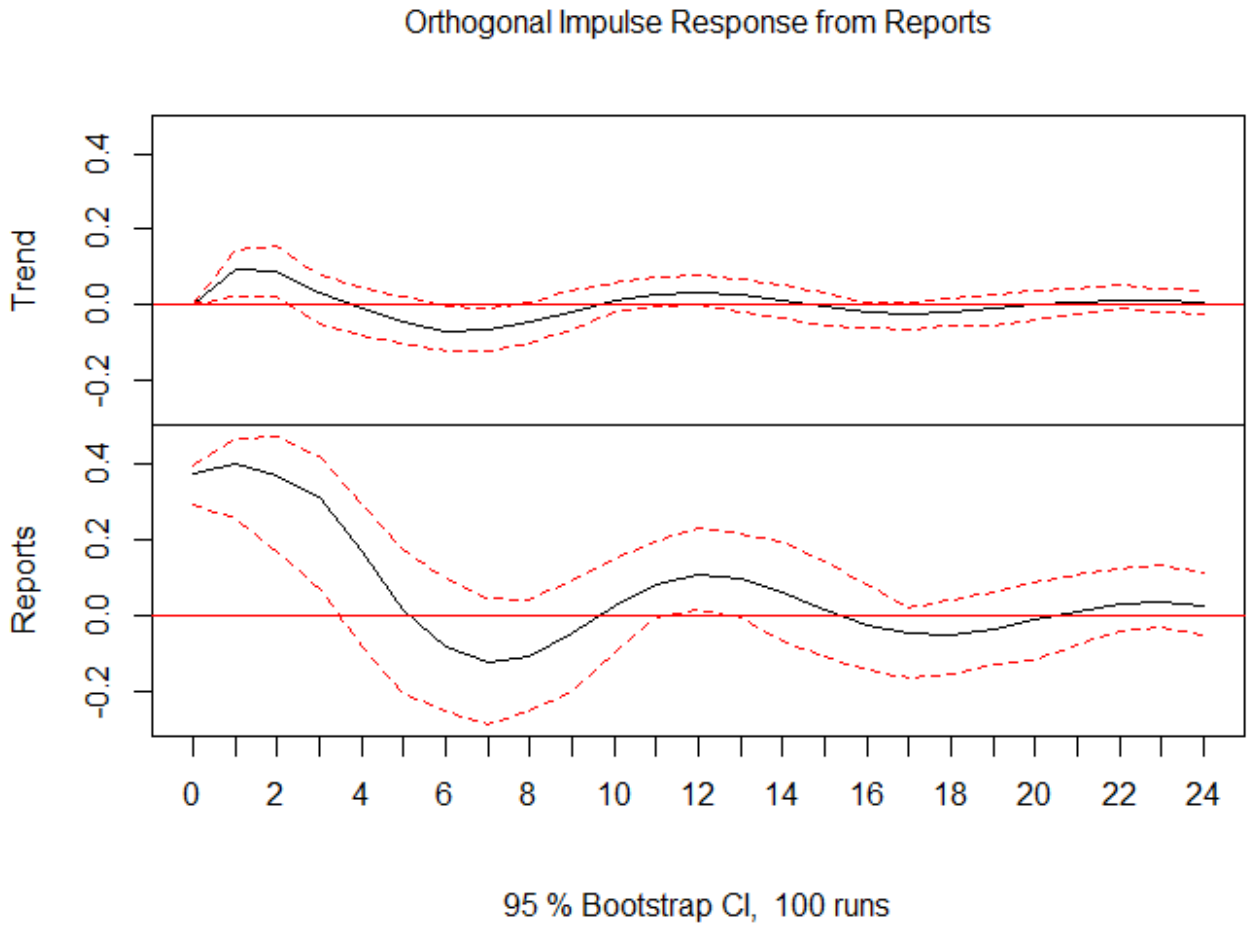
VAR Residuals



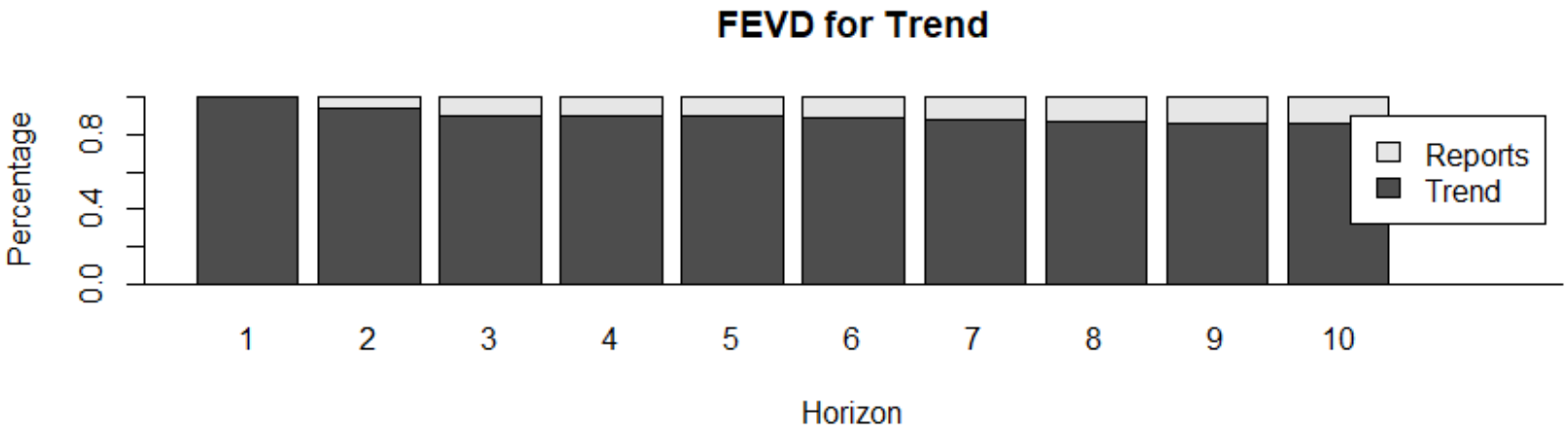
IRF Trends

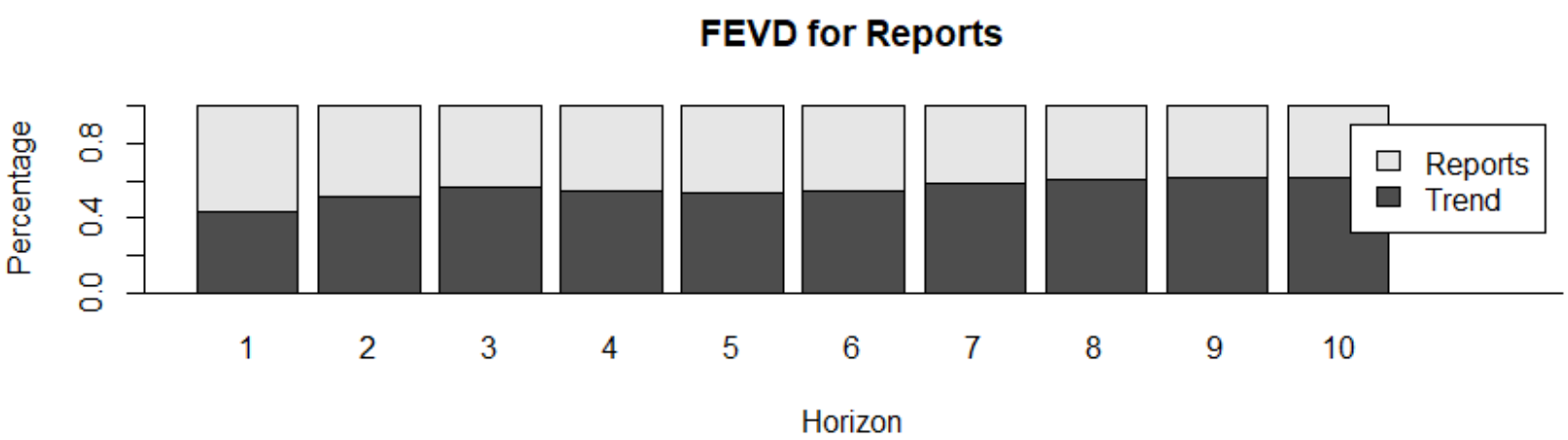

Orthogonal Impulse Response from Trend
Trend
Reports
0.4
0.0
-0.4
0
2
4
6
8
10
12
14
16
18
20
22
24
95 % Bootstrap CI, 100 runs

IRF Reports


Orthogonal Impulse Response from Reports
Trend
Reports
-0.2
0.0
0.2
0.4
0
2
4
6
8
10
12
14
16
18
20
22
24
95 % Bootstrap CI, 100 runs

FEVD


FEVD for Trend
Percentage
0.0
0.4
0.8
1
2
3
4
5
6
7
8
9
10
Horizon
Reports
Trend


FEVD for Reports
Percentage
0.0
0.4
0.8
1
2
3
4
5
6
7
8
9
10
Horizon
Reports
Trend