Lab3

Subhashini R., Lei Yang, Ron Cordell

April 12, 2016

Lab 3

Part 1 - Modeling House Values

Exploratory Data Analysis

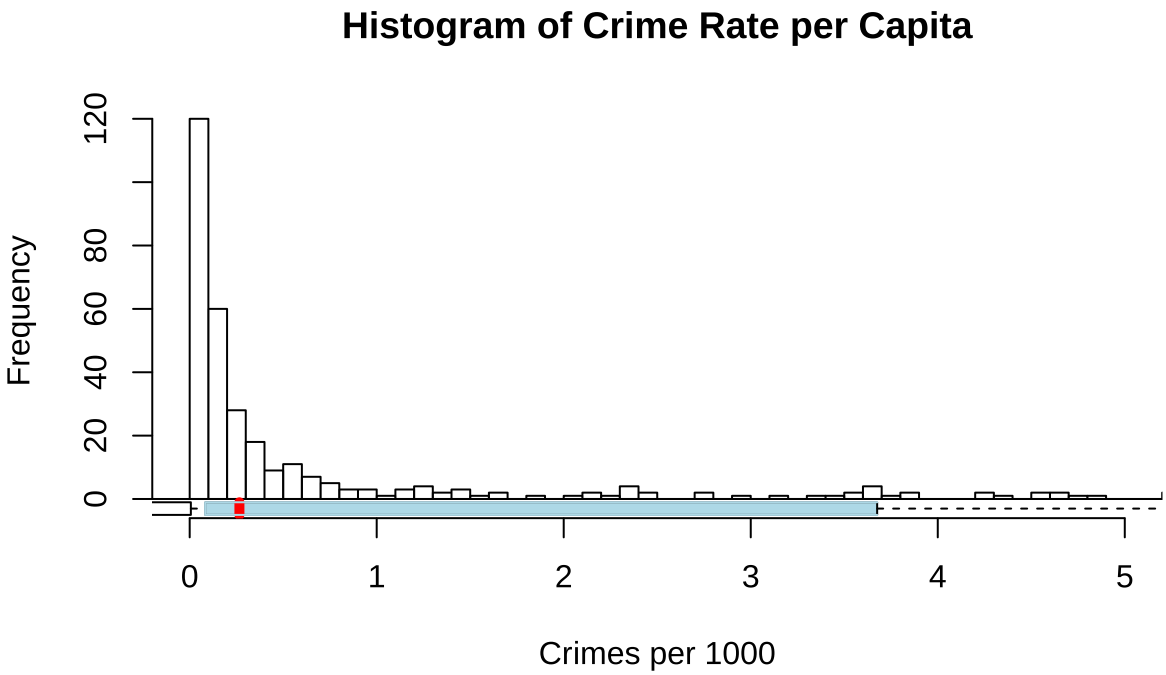
An examination of the provided data set reveals 11 variables of which *withWater* is binary and rest are continuous. There are no NA's in the data set, however the *distanceToHighway* variable appears to have a coding issue. We'll examine this variable in more detail a bit later, but first we summarize the variables in the following table.

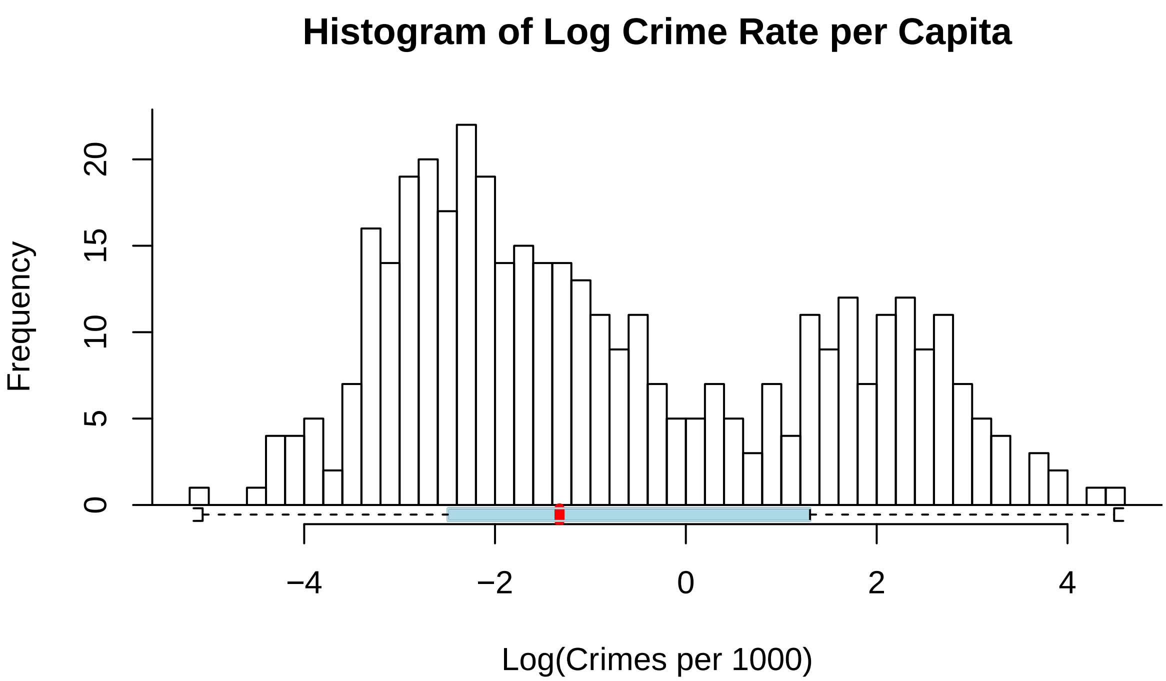
Table 1: Summary of Data

Statistic	N	Mean	St. Dev.	Min	Max
crimeRate_pc	400	3.763	8.872	0.006	88.976
nonRetailBusiness	400	0.112	0.070	0.007	0.277
withWater	400	0.068	0.251	0	1
ageHouse	400	68.932	27.977	2.900	100.000
distanceToCity	400	9.638	8.786	1.228	54.197
distanceToHighway	400	9.582	8.672	1	24
pupilTeacherRatio	400	21.391	2.168	15.600	25.000
pctLowIncome	400	15.795	9.341	2	49
homeValue	400	499,584.400	196,115.700	112,500	1,125,000
pollutionIndex	400	40.615	11.825	23.500	72.100
nBedRooms	400	4.266	0.719	1.561	6.780

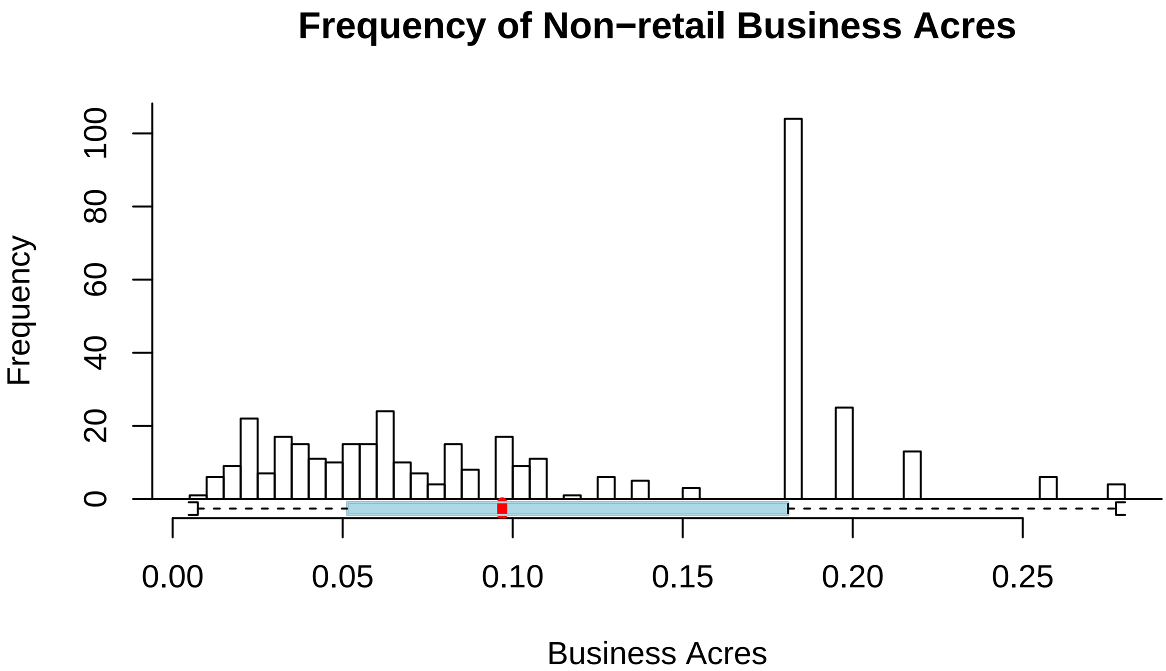
For the purposes of this analysis we categorize the variables *crimeRate_pc*, *nonRetailBusiness*, *withWater*, *distanceToCity*, *distanceToHighway*, *pollutionIndex*, *pupilTeacherRation* and *pctLowHousing* to be environmental variables. The variables *ageHouse* and *nBedRooms* are attributes of the house. The variable *homeValue* is the dependent variable we would like to explain in terms of primarily the environment variables but we will compare to explanations in terms of house attributes as well.

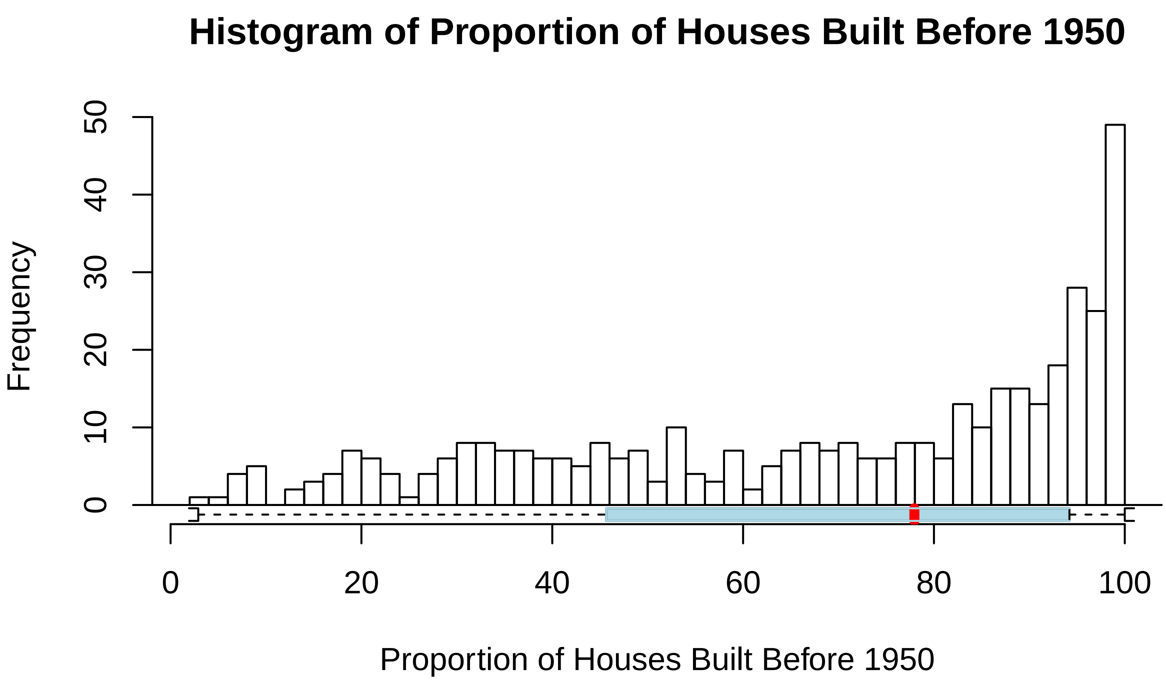
In the next several pages we examine the distribution of each of the variables and, where indicated, the distribution of the log(variable) as well.





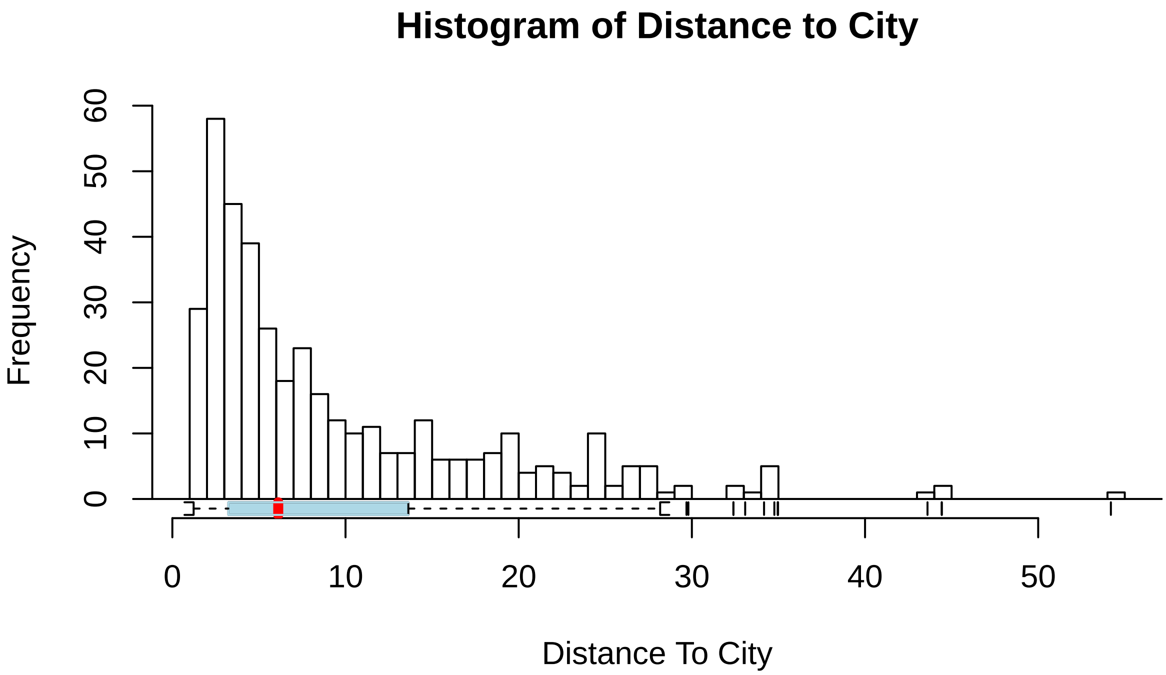
The distribution of *crimeRate_pc* is highly right-skewed with a very long tail. This makes sense as most neighborhoods are very low crime neighborhoods. The distribution of *log(crimeRate_pc)* appears almost bi-modal; however the analysis of skew and kurtosis show a significant improvement of each with a log transformation.

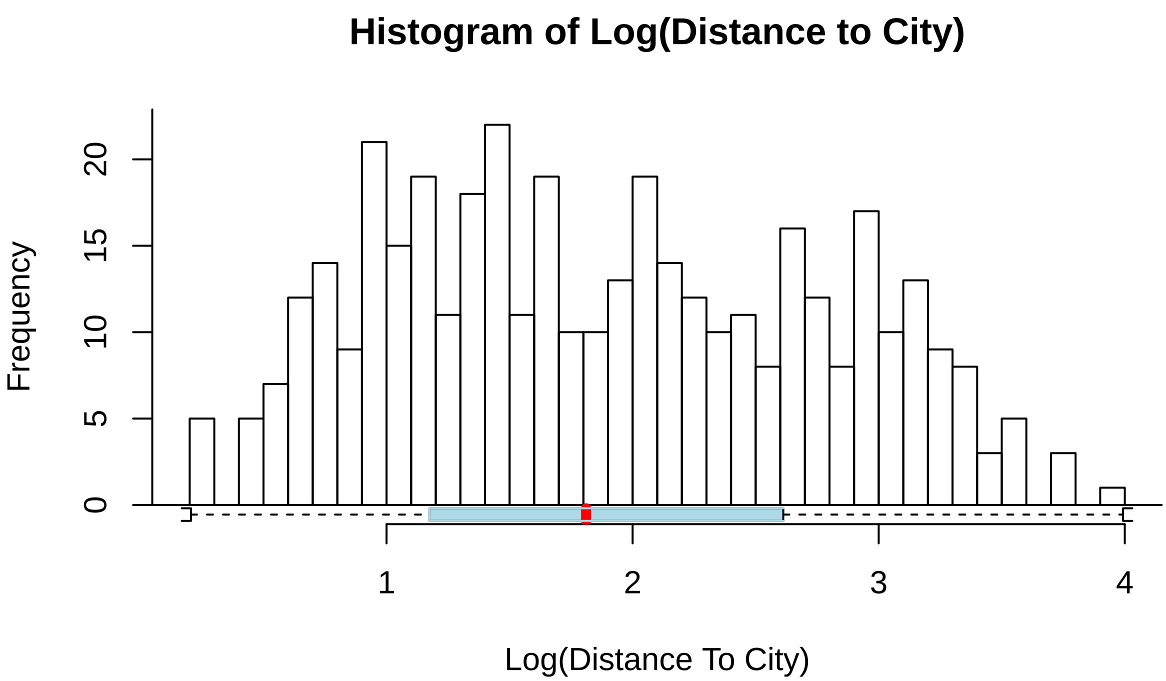




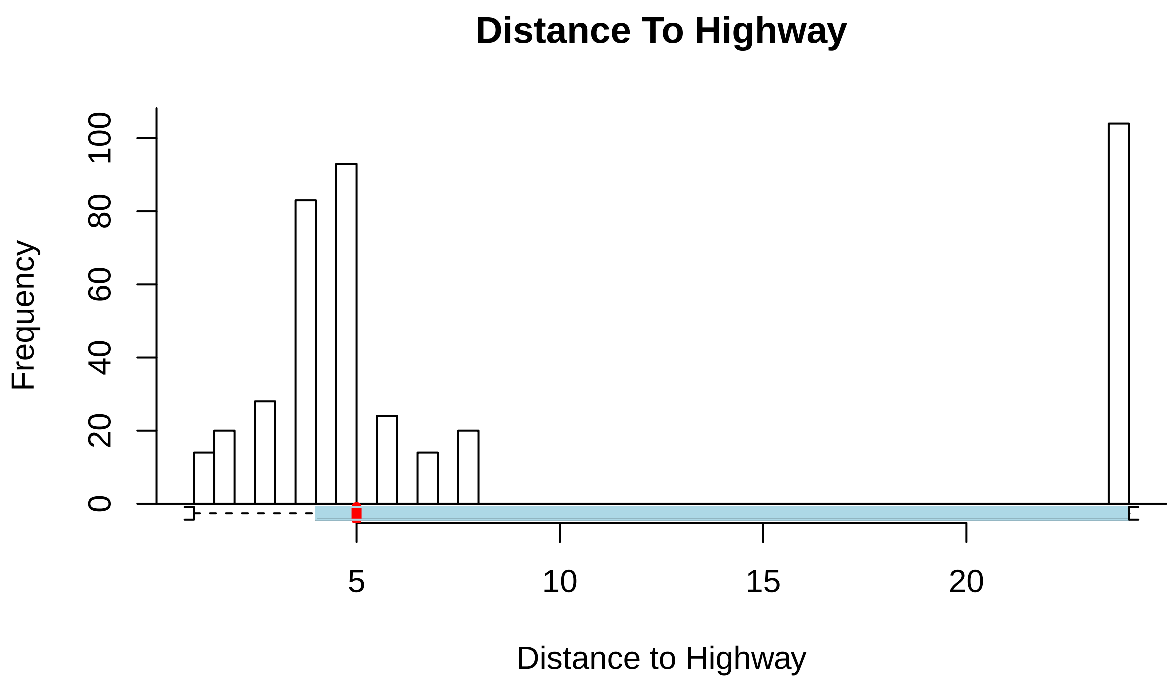
The variable *nonRetailBusiness* is a measure of the footprint of industy in a neaghborhood. This may range from light industrial to manufacturing but that information is not given. The distribution of *nonRetailBusiness* shows a spike at 0.18 business-acres but is otherwise somewhat uniform. There was no transform that improved skew or kurtosis for this variable.

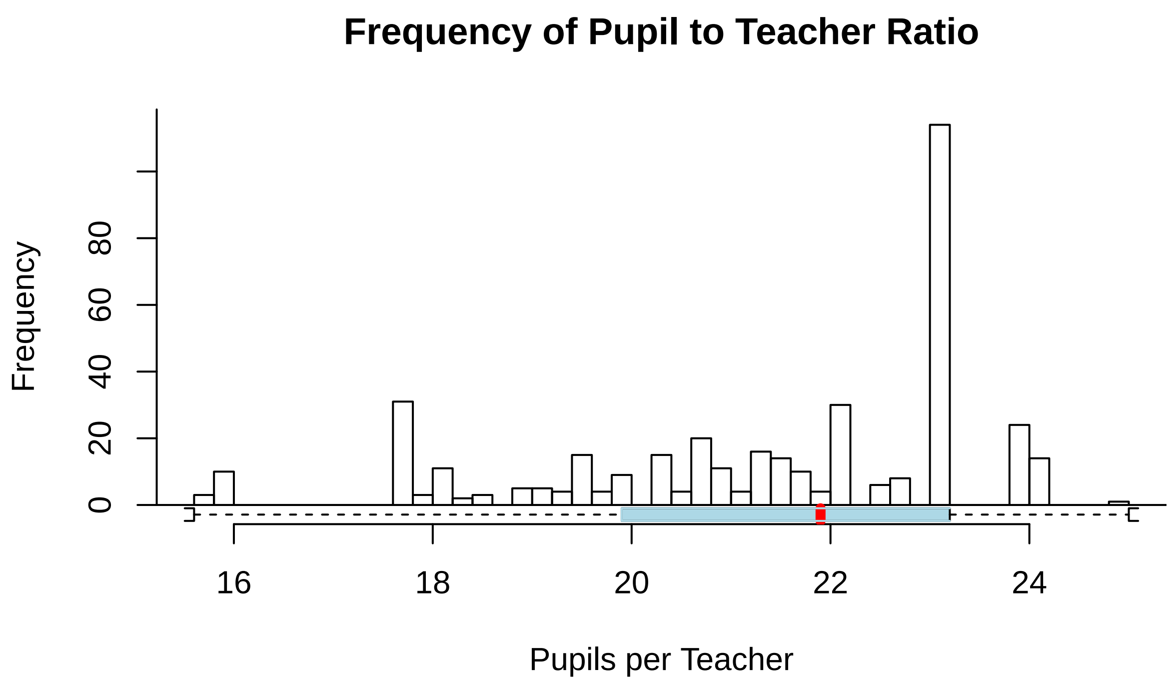
The variable *ageHouse* is the percentage of houses in a neighborhood built before 1950 and shows a significant left-skew with a long tail to the left. However no transformation was found that normalized the skew and kurtosis of this variable.





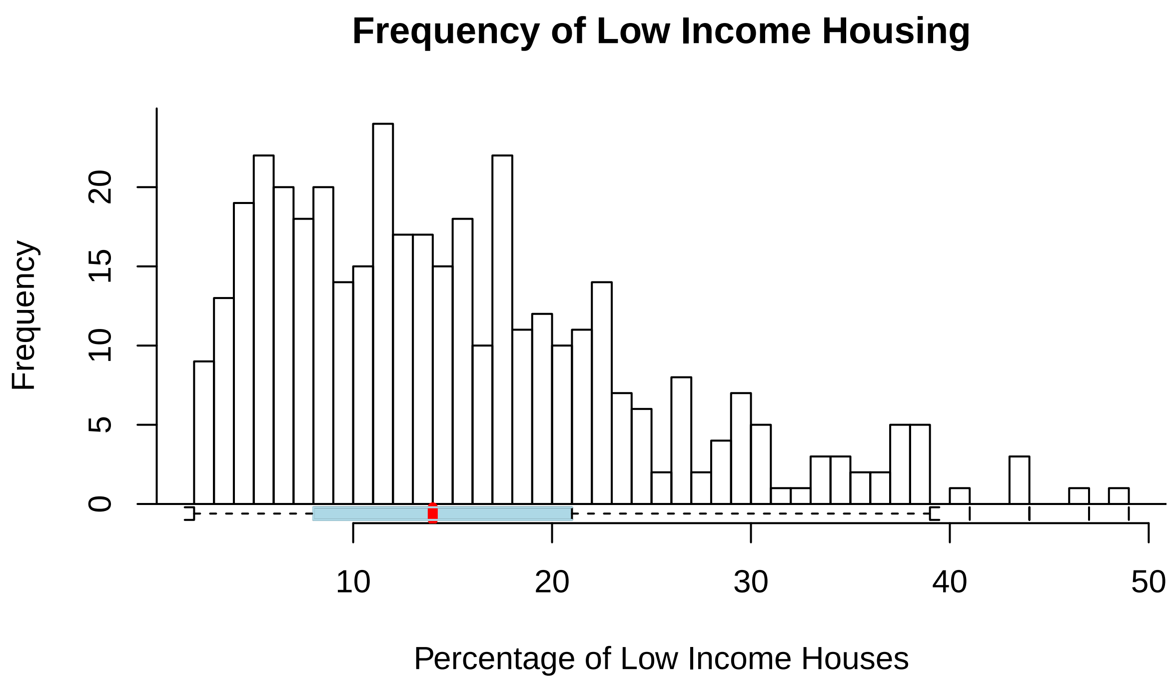
The *distanceToCity* variable shows a right-skewed distrubution that is much improved by a log transformation.

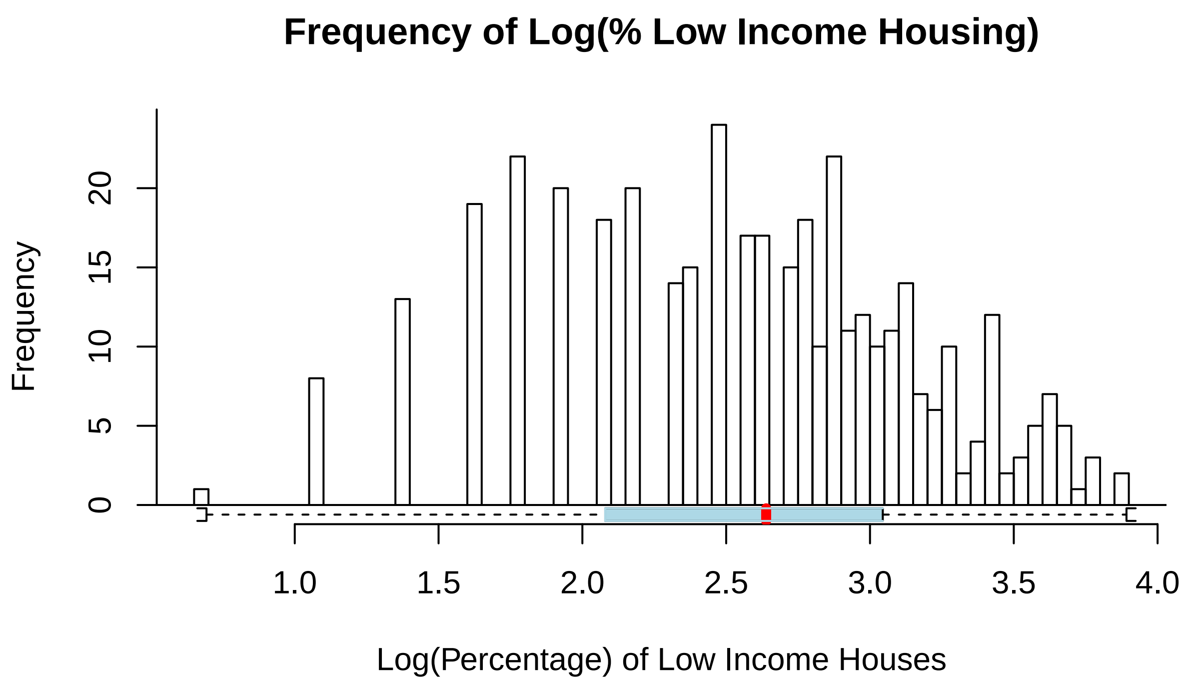




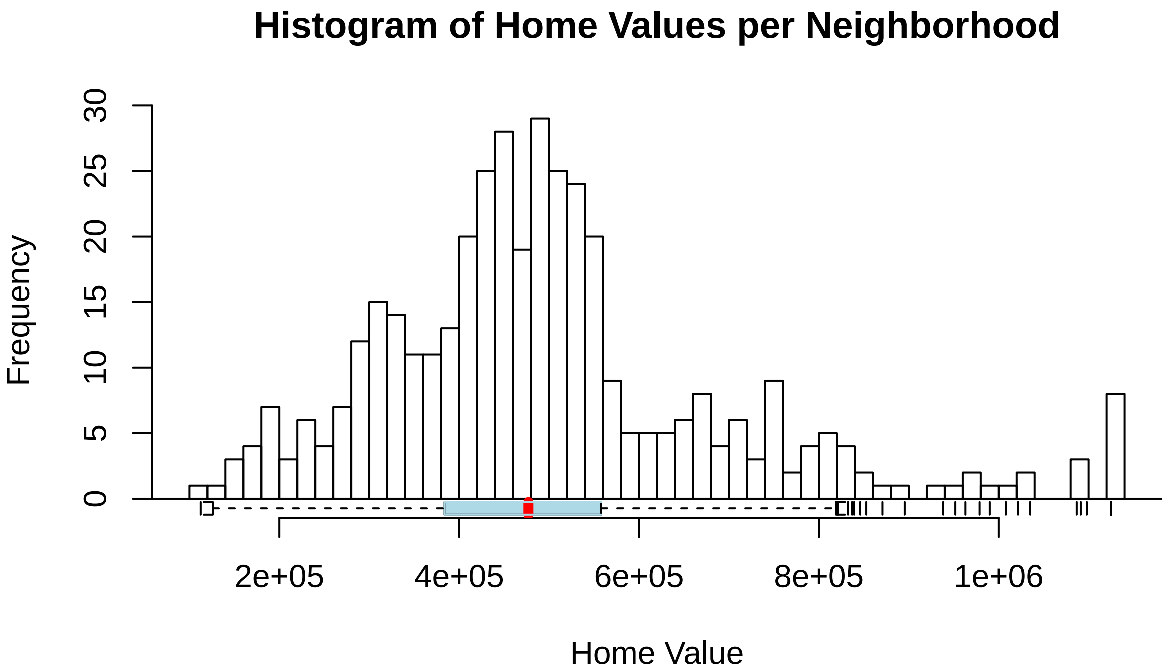
The *distanceToHighway* variable shows the concern with coding error in this histogram as there is a large occurance of the value *24*. About 25% of the dataset have this value, some of which may be correct but it seems unlikely that the *distanceToHighway* variable would be much greater than the *distanceToCity* variable.

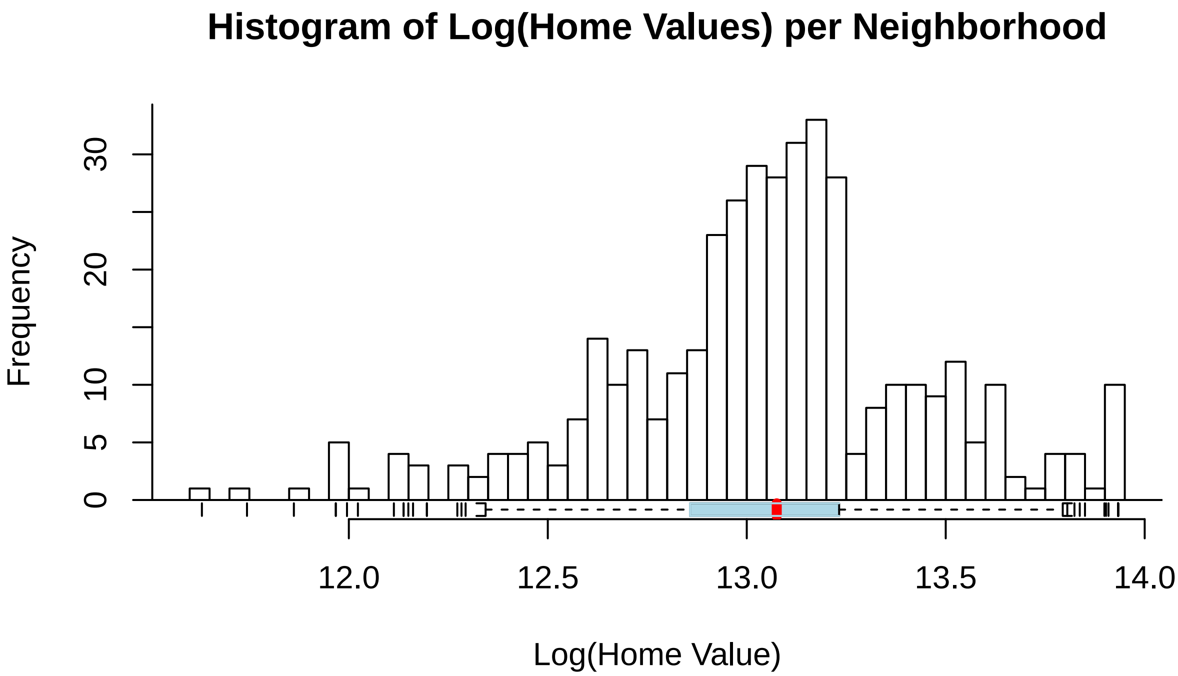
The *pupilTeacherRatio* variable shows a roughly uniform distribution except for a large number of occurances of the value *23*, which must be a more common classroom size.



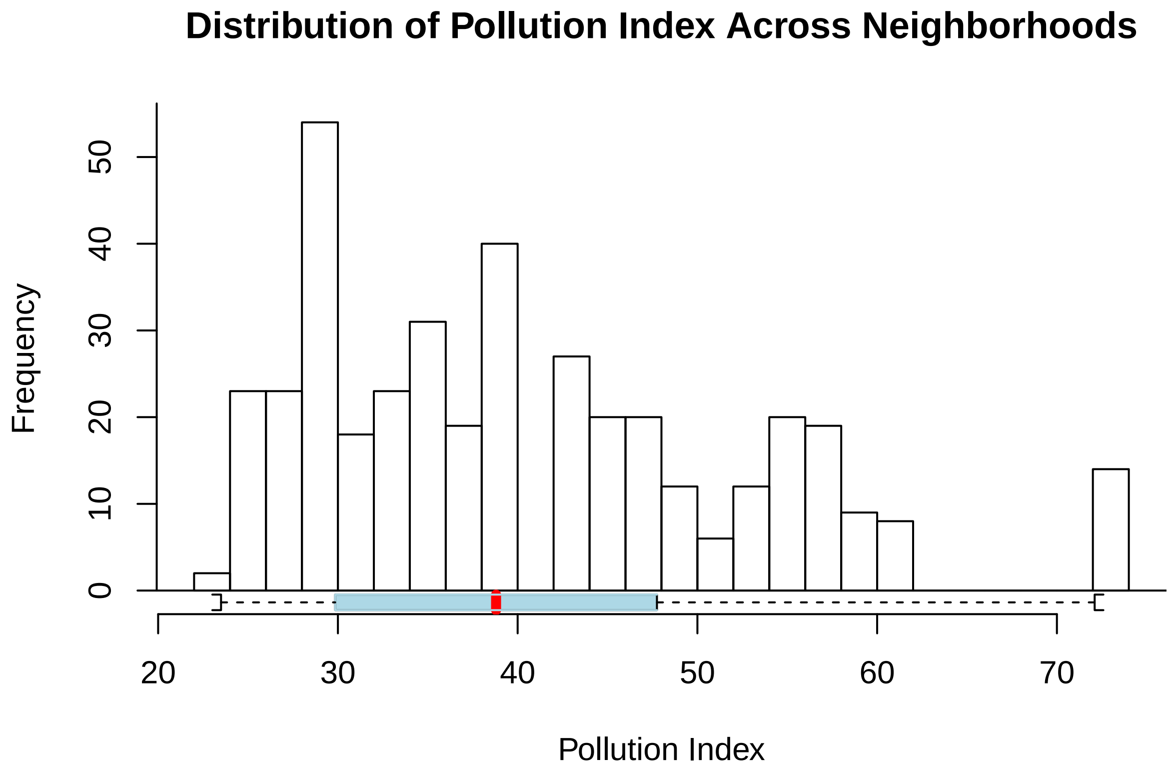


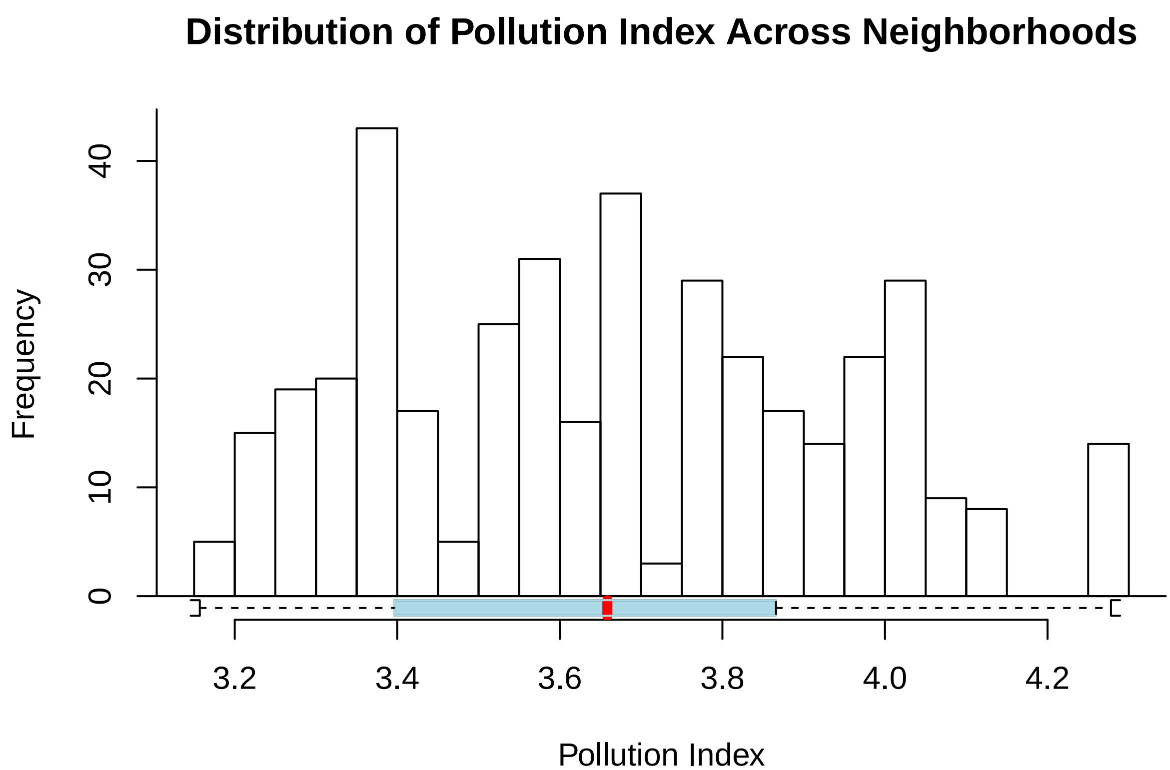
The *pctLowIncome* variable has a right-skewed distribution that tapers off to the right relatively quickly. A log transformation greatly improves the skew and kurtosis of the distribution.





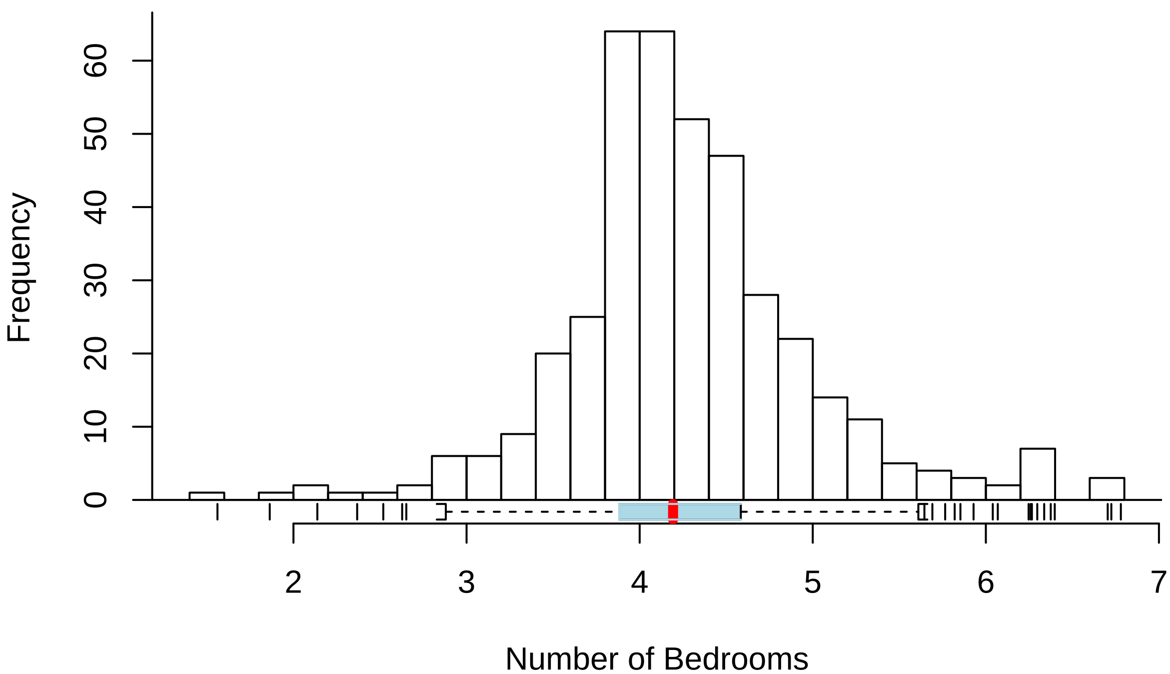
The *homeValue* variable shows a slight right-skew and a log transformation is used to help with this. It also allows discussion in terms of percentage change of home value when controlling for other variables.





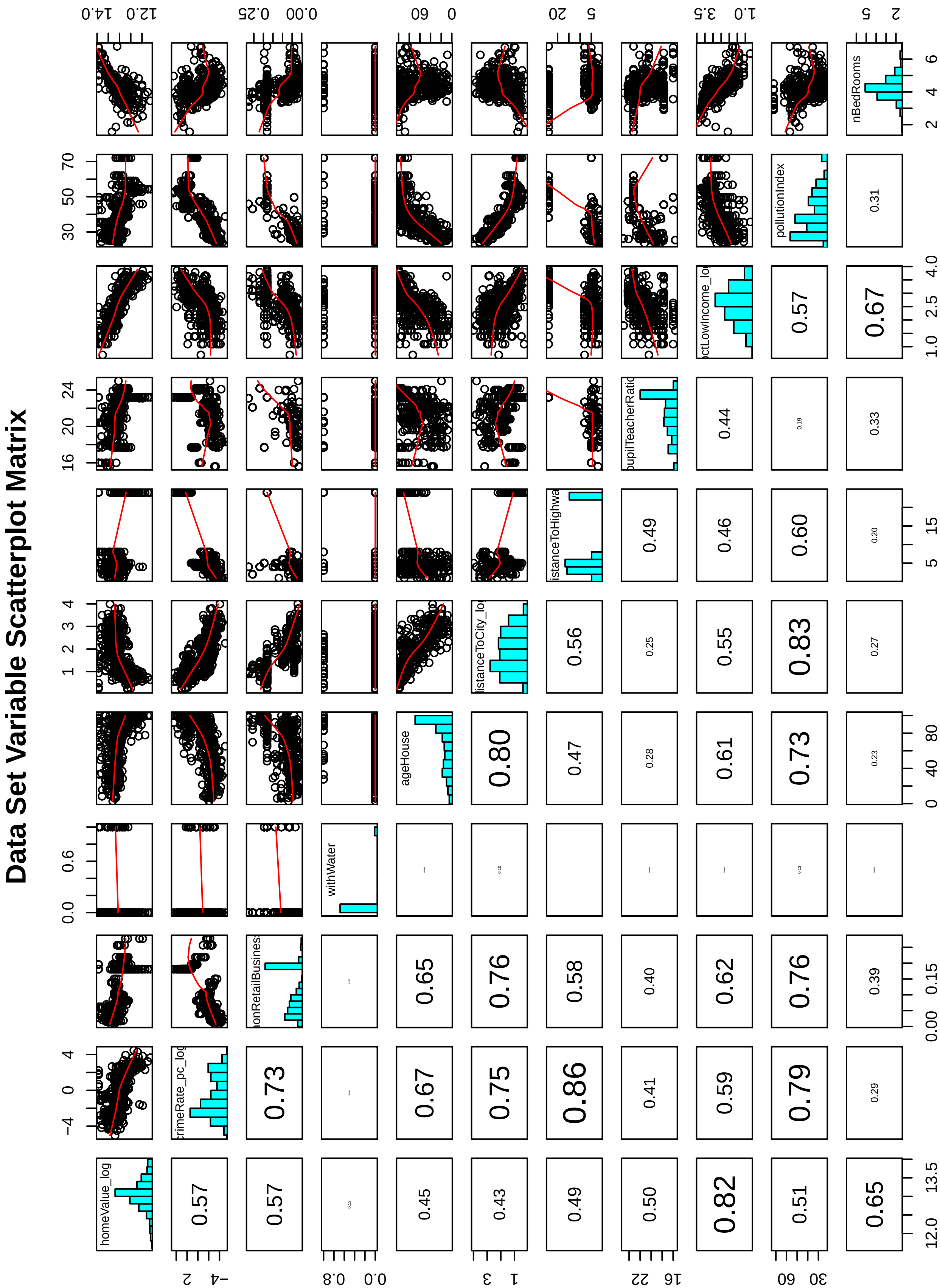
The *pollutionIndex* variable shows a right-skewed distribution upon which we perform a log transform.

Distribution of Number of Bedrooms



The *nBedRooms* variable appears amazingly normal-like in its distribution.

The next page brings all the variables into a single matrix for comparison and to get a first look at correlations to explore further.


Data Set Variable Scatterplot Matrix
homeValue_log
crimeRate_pc_log
withWater
ageHouse
pupilTeacherRatio
pollutionIndex
nBedRooms

DistanceToHighway Variable Detailed Examination

We saw previously that the *distanceToHighway* variable looked suspicious so in this section we look at how to address a possible coding issue. The number of rows in the dataset that have the *distanceToHighway* variable as 24 is 25% of the dataset. Removing these rows would remove a significant amount of data, reducing *N=400* to *N=296*. We examine two strategies and compare them to the row-removal option: replacing values of 24 with the mean of the filtered values or with the value of *distanceToCity*.

The following two tables compare the summaries of the filtered dataset with the summaries of the dataset with transformed values. Comparing the *distanceToHighway_meanMod* and *distanceToHighway_cityMod* shows that the latter is much closer to the values of the filtered dataset. This indicates that replacing the value of 24 with the value of *distanceToCity* is a reasonable transformation to deal with the coding issue. The idea is further substantiated by the proposition that the distance to a city is usually not greater than the distance to a highway as cities are generally located on highways.

The following page shows a set of comparison histograms for the *distanceToHighway* variable with the different transformations. A histogram of *distanceToCity* is included as a reference.

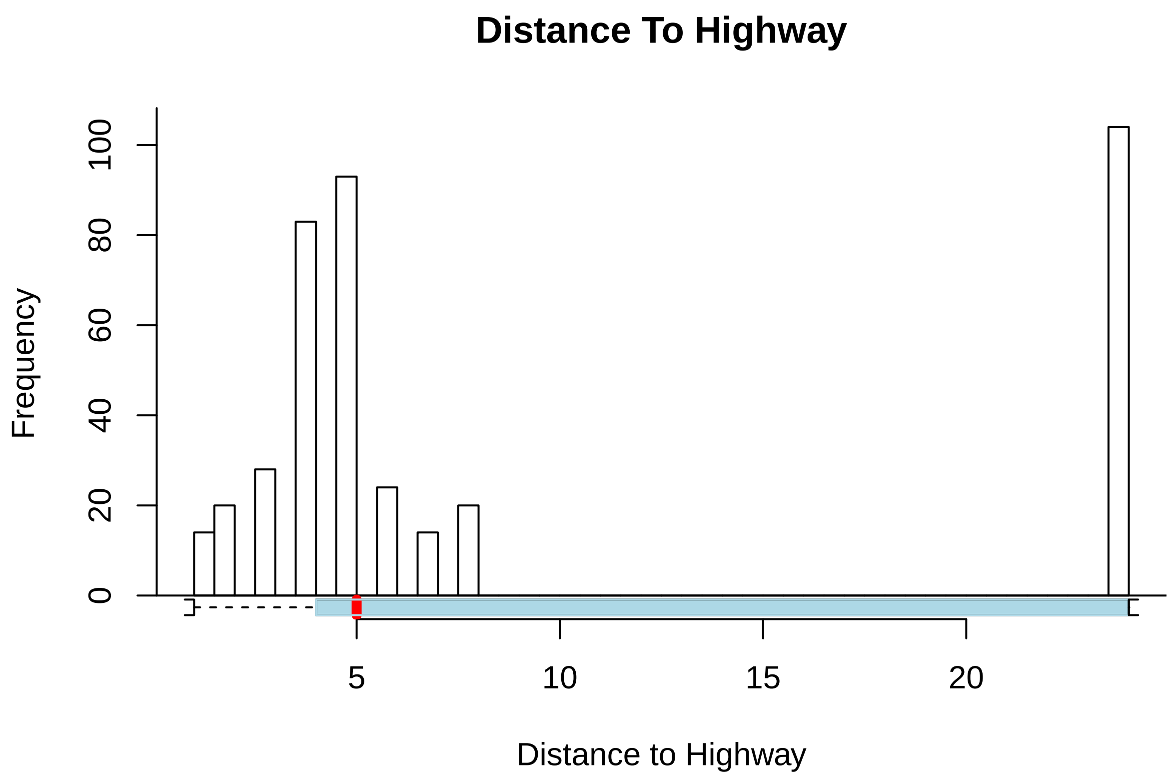
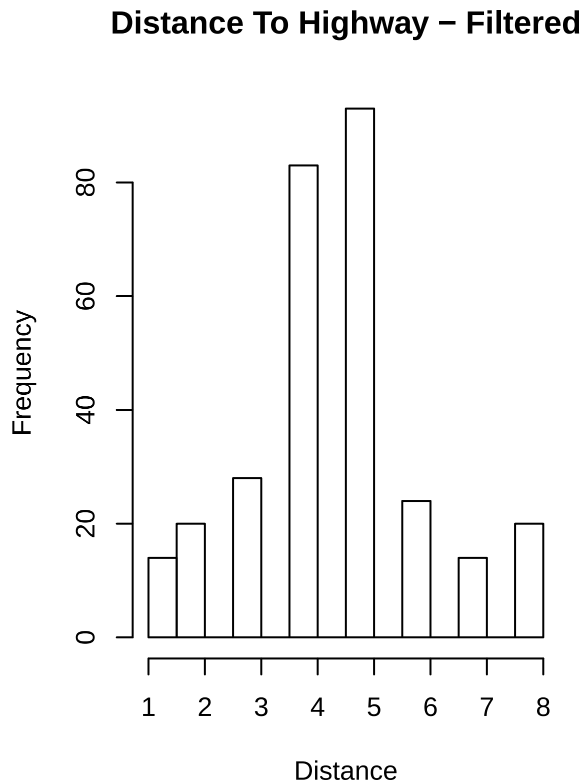


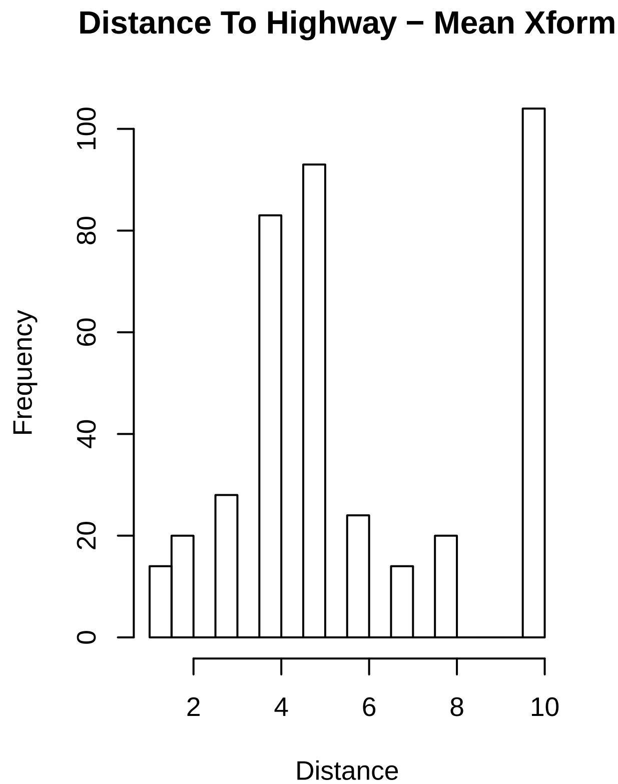
Table 2: Filtered Dataset

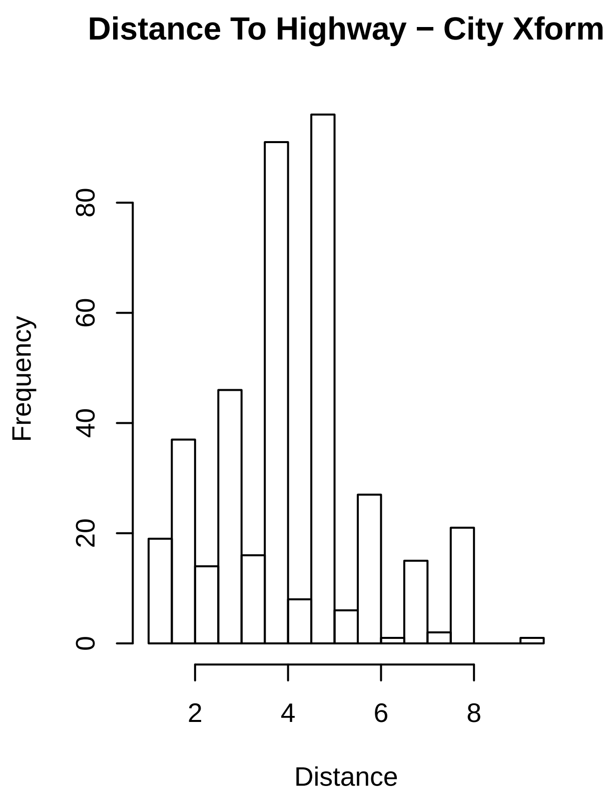
Statistic	N	Mean	St. Dev.	Min	Max
crimeRate_pc	296	0.381	0.610	0.006	3.535
nonRetailBusiness	296	0.087	0.065	0.007	0.277
withWater	296	0.068	0.251	0	1
ageHouse	296	61.367	28.205	2.900	100.000
distanceToCity	296	11.877	9.178	1.562	54.197
distanceToHighway	296	4.517	1.636	1	8
pupilTeacherRatio	296	20.756	2.191	15.600	25.000
pctLowIncome	296	13.037	7.690	2	44
homeValue	296	547,487.300	178,890.300	157,500	1,125,000
pollutionIndex	296	36.438	10.450	23.500	72.100
nBedRooms	296	4.356	0.676	2.903	6.725
crimeRate_pc_log	296	−1.840	1.297	−5.064	1.263
distanceToCity_log	296	2.175	0.802	0.446	3.993
pctLowIncome_log	296	2.403	0.587	0.693	3.784
ageHouse_log	296	3.959	0.643	1.065	4.605
homeValue_log	296	13.165	0.306	11.967	13.933
pollutionIndex_log	296	3.562	0.250	3.157	4.278

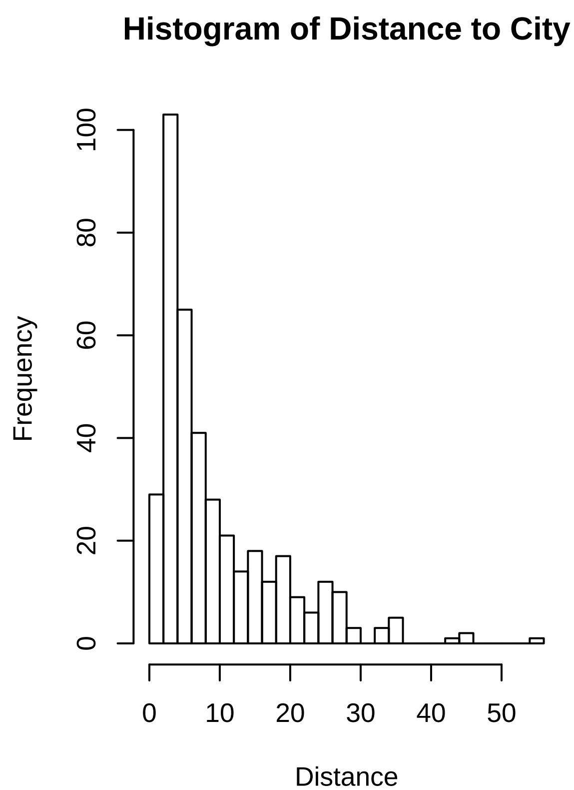
Table 3: Full Dataset

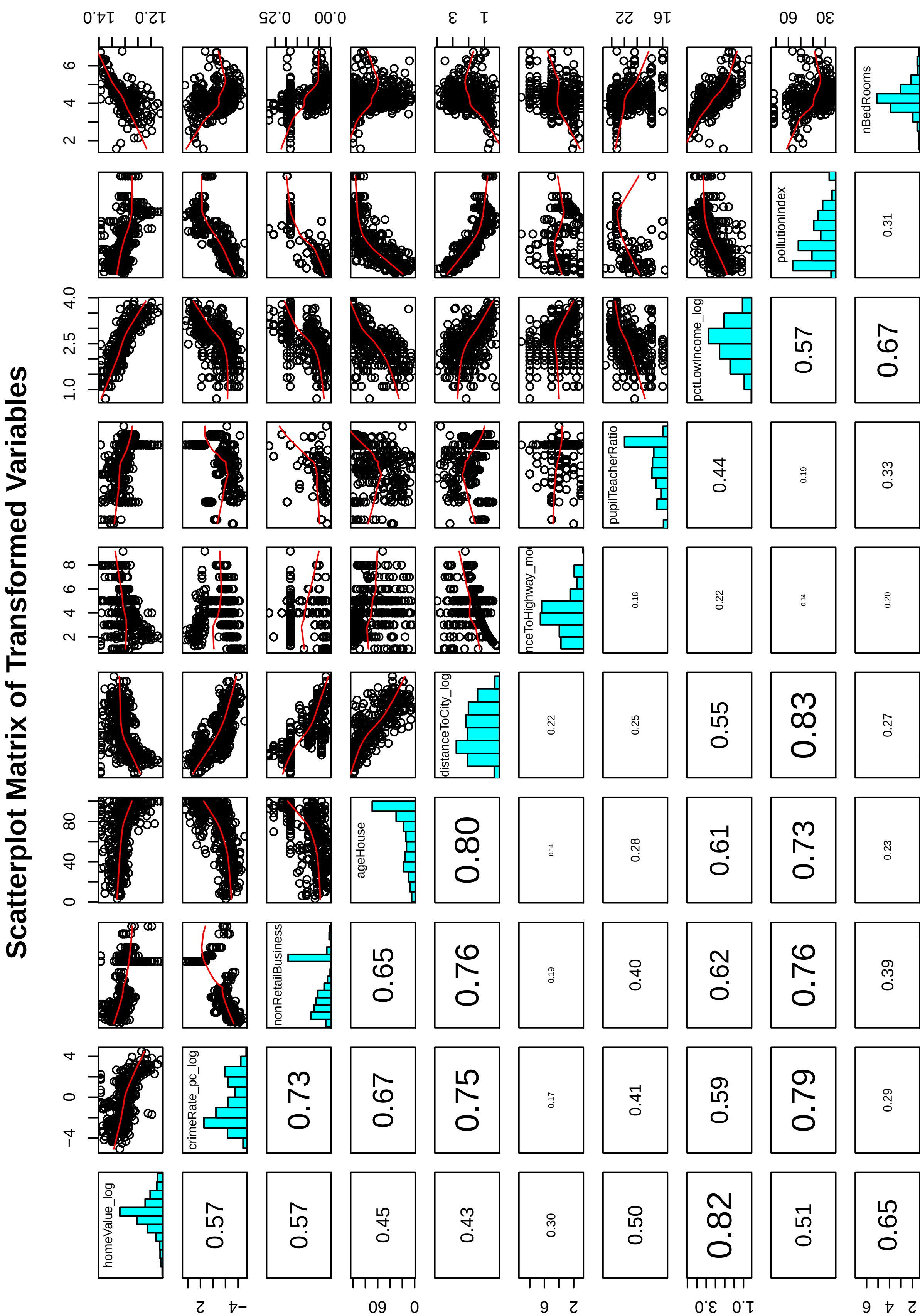
Statistic	N	Mean	St. Dev.	Min	Max
crimeRate_pc	400	3.763	8.872	0.006	88.976
nonRetailBusiness	400	0.112	0.070	0.007	0.277
withWater	400	0.068	0.251	0	1
ageHouse	400	68.932	27.977	2.900	100.000
distanceToCity	400	9.638	8.786	1.228	54.197
distanceToHighway	400	9.582	8.672	1	24
pupilTeacherRatio	400	21.391	2.168	15.600	25.000
pctLowIncome	400	15.795	9.341	2	49
homeValue	400	499,584.400	196,115.700	112,500	1,125,000
pollutionIndex	400	40.615	11.825	23.500	72.100
nBedRooms	400	4.266	0.719	1.561	6.780
crimeRate_pc_log	400	−0.763	2.164	−5.064	4.488
distanceToCity_log	400	1.892	0.868	0.205	3.993
pctLowIncome_log	400	2.577	0.631	0.693	3.892
ageHouse_log	400	4.099	0.605	1.065	4.605
homeValue_log	400	13.046	0.397	11.631	13.933
pollutionIndex_log	400	3.664	0.282	3.157	4.278
distanceToHighway_modMean	400	5.834	2.632	1.000	9.582
distanceToHighway_modCity	400	4.192	1.698	1.000	9.159


Distance To Highway − Filtered
Frequency
0
20
40
60
80
1
2
3
4
5
6
7
8
Distance


Distance To Highway − Mean Xform
Frequency
0
20
40
60
80
100
2
4
6
8
10
Distance

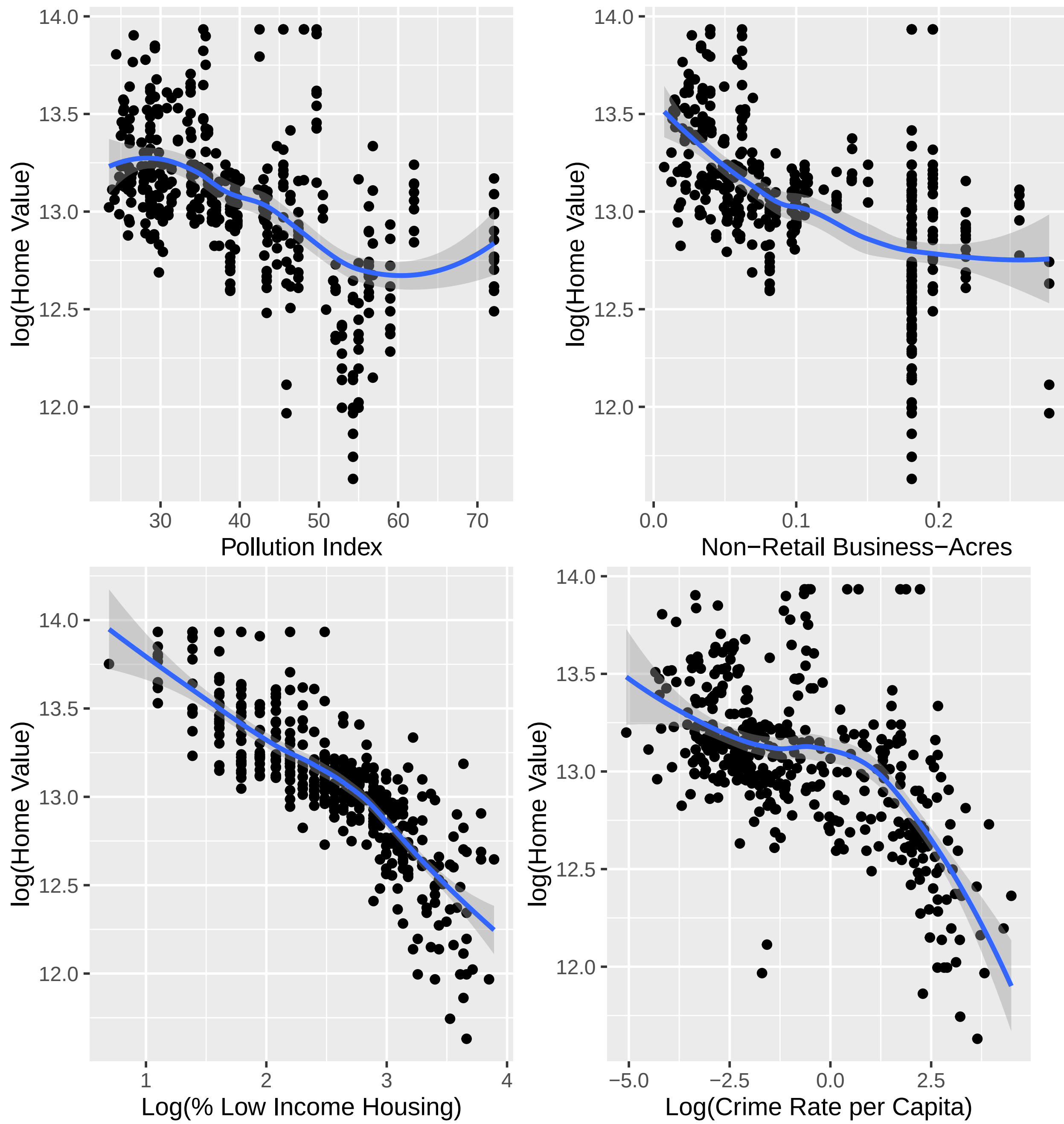

Distance To Highway − City Xform
Frequency
0
20
40
60
80
2
4
6
8
Distance

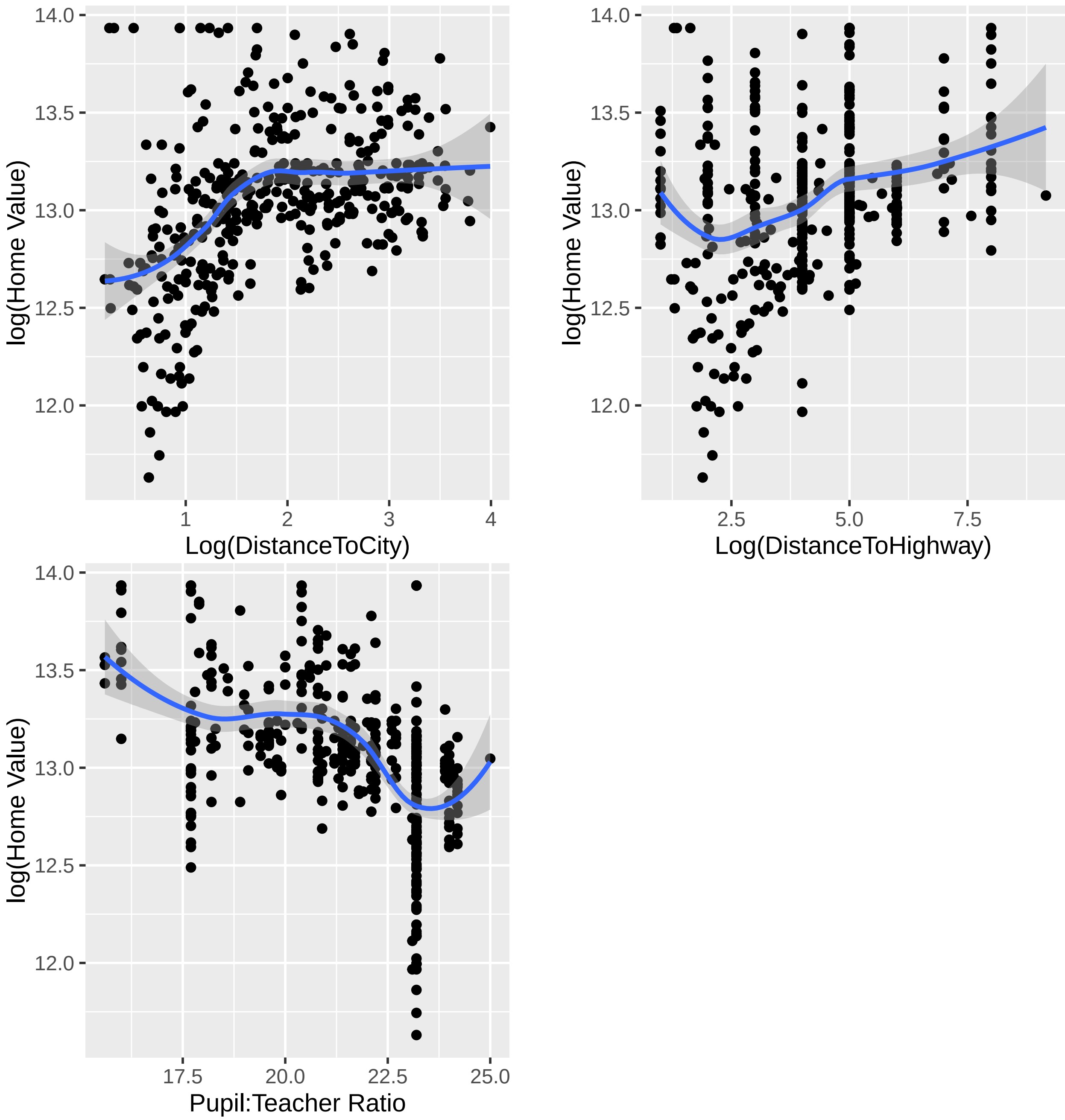

Histogram of Distance to City
Frequency
0
20
40
60
80
100
0
10
20
30
40
50
Distance


Scatterplot Matrix of Transformed Variables
homeValue_log
crimeRate_pc_log
nonRetailBusiness
ageHouse
distanceToCity_log
nceToHighway_mo
pupilTeacherRatio
pctLowIncome_log
pollutionIndex
nBedRooms

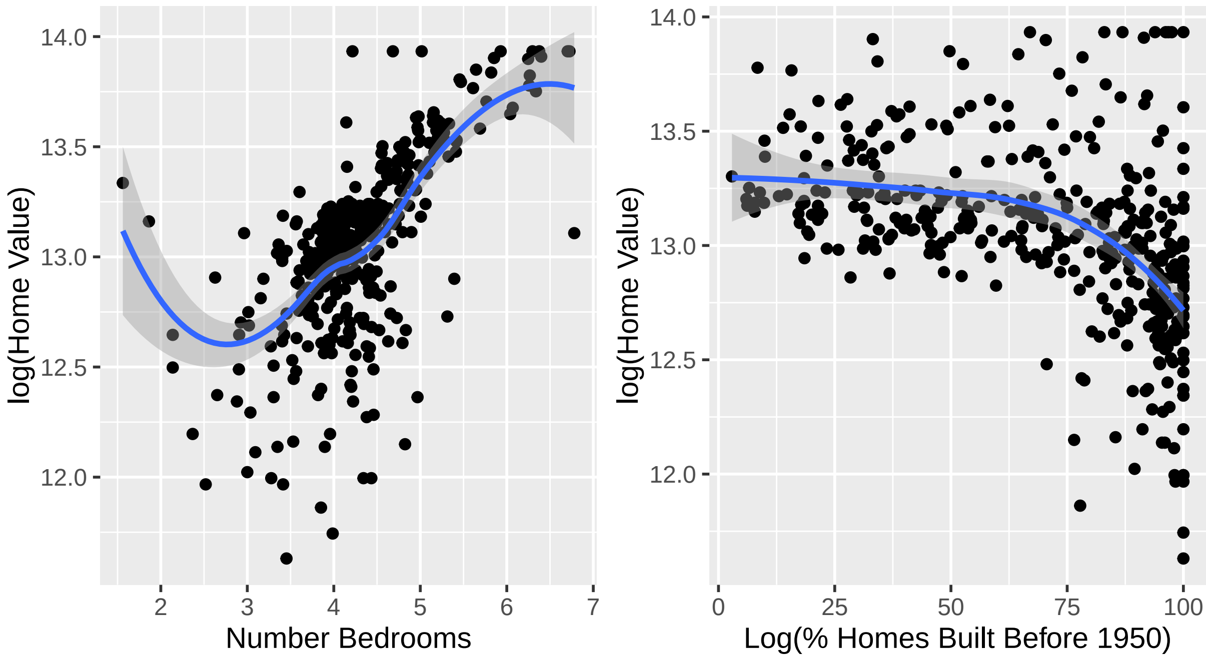
Multivariate Data Analysis

First we will examine the relationships of the environment variables on home values graphically. We can see that there are definitely relationships between most of the environmental variables and home values, as shown on the next two pages of graphs.




log(Home Value)
Log(DistanceToCity)
Log(DistanceToHighway)
Pupil:Teacher Ratio
14.0
13.5
13.0
12.5
12.0
1
2
3
4
2.5
5.0
7.5
17.5
20.0
22.5
25.0

These final two graphs show relationships between home attributes and home values.



Models Incorporating Environment Variables

Put something here about all the different models that were tried.

Include models with interaction terms

Discuss omitted variables and how they bias the models

What hypotheses can we test?

Choose the best model(s) and explain the parameters for them.

Table 4: Regression Model Comparison

	Dependent variable:				
	log(Home Values)				
	(1)	(2)	(3)	(4)	(5)
log(Low Income Housing)	-0.518^{***}	-0.469^{***}	-0.474^{***}	-0.461^{***}	-0.448^{***}
	(0.018)	(0.022)	(0.022)	(0.022)	(0.022)
log(Crime Rate)		-0.024^{***}	-0.030^{***}	-0.026^{***}	-0.024^{***}
		(0.006)	(0.009)	(0.006)	(0.006)
Pollution Index			0.002		
			(0.002)		
Close To Water					0.027^{***}
					(0.007)
Distance To Highway				0.144^{***}	0.137^{***}
				(0.044)	(0.043)
Constant	14.380^{***}	14.238^{***}	14.181^{***}	14.205^{***}	14.060^{***}
	(0.048)	(0.061)	(0.082)	(0.061)	(0.069)
Observations	400	400	400	400	400
R^2	0.677	0.688	0.689	0.696	0.709
Adjusted R^2	0.676	0.686	0.686	0.694	0.706
Residual Std. Error	0.226	0.222	0.222	0.220	0.215
F Statistic	834.337^{***}	437.399^{***}	291.964^{***}	302.291^{***}	240.388^{***}
Note:					$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Part 2 - Modeling and Forecasting a Real-World Macroeconomic Financial Time Series

Exploratory Data Analysis

In order to better understand the time series and analyze the possible underlying processes we must first observe and explore the time series. A summary of the time series is:

[1] "function"

Table 5: DXCM Series Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
DXCM Series	2,332	23.2	23.4	1.4	101.9

The first set of plots reveals:

- The series is non-stationary; it has a persistent upward trend, interrupted by shocks;
- There are shocks at approximately time periods 500, 1200, 1800 and 2200;
- There appears to be seasonality in the series;
- The autocorrelation shows a very slight decay over the entire correlagram;
- The partial autocorrelation shows barely significant results at lags 14 and 32;
- We do not know the frequency of the time series;
- The series is of closing prices of DXCM

To remove the trend from the series we take the first difference and replot to check the results.

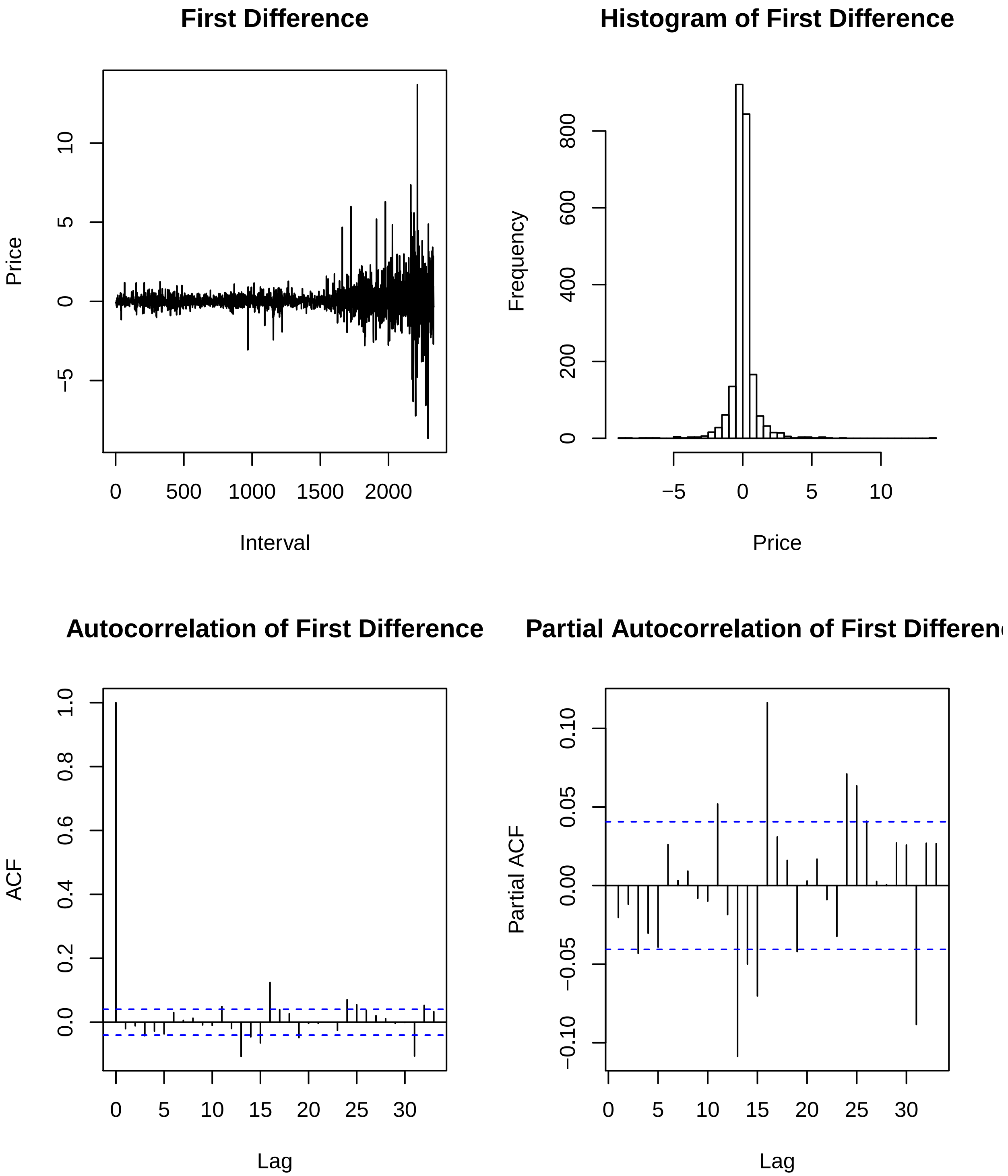
[1] "function"

Table 6: Differenced Series Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
DXCM Differenced Series	2,331	0.02	0.9	−8.6	13.7

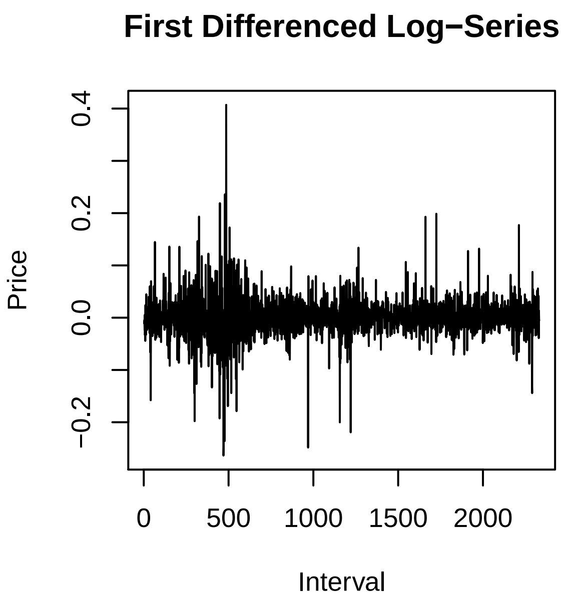
In the differenced series we observe:

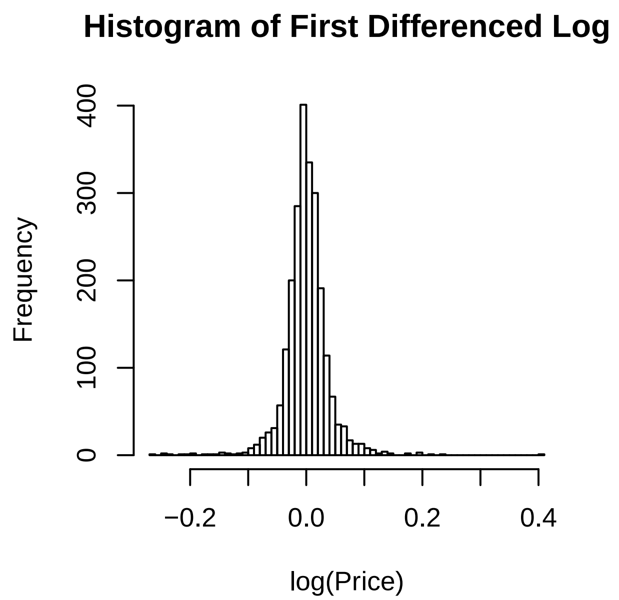
- The first difference series has a more or less white noise appearance until approximately time interval 1600 where the volatility of the series increases dramatically. This corresponds to the sudden, persistent upward trend in the original series.
- The autocorrelation shows marginally significant results at lags 13, 15, 16, 24, 31
- The partial autocorrelation shows a cyclic behavior that doesn't appear to decline, with significant results at lags 11, 13, 14, 15, 16, 24, 25, 31

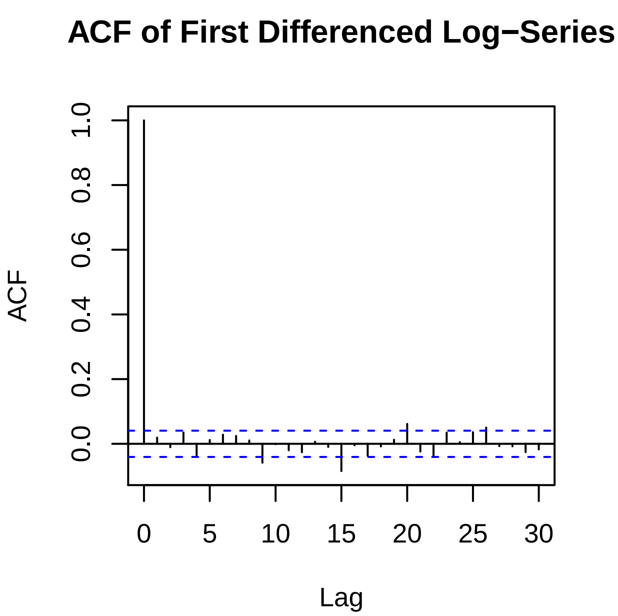

First Difference
Price
Interval
Histogram of First Difference
Frequency
Price
Autocorrelation of First Difference
ACF
Lag
Partial Autocorrelation of First Differen
Partial ACF
Lag

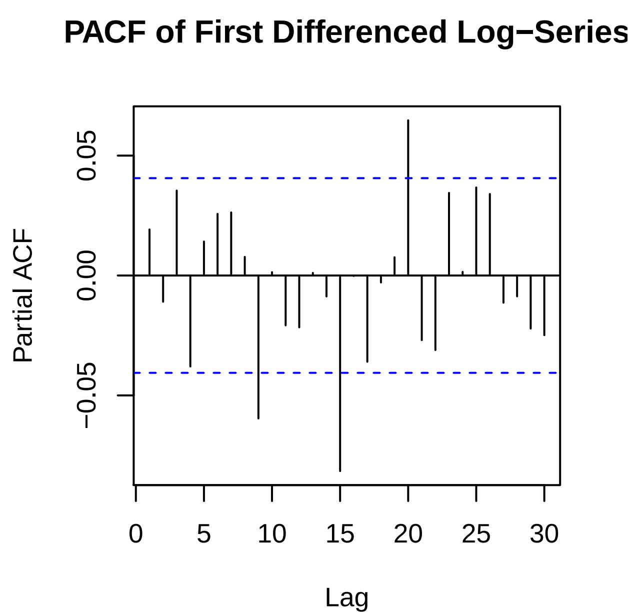
We also examine the first difference of the log of the series and replot to check results. In the differenced log series we observe:

- The volatility appears to be reversed such that it is around interval 300-500, and overall the volatility of the differenced log series is higher.
- The ACF shows only a small results at lag 15 and 20.
- The PACF shows a cyclic behavior with significant results at lags 9, 15, 20



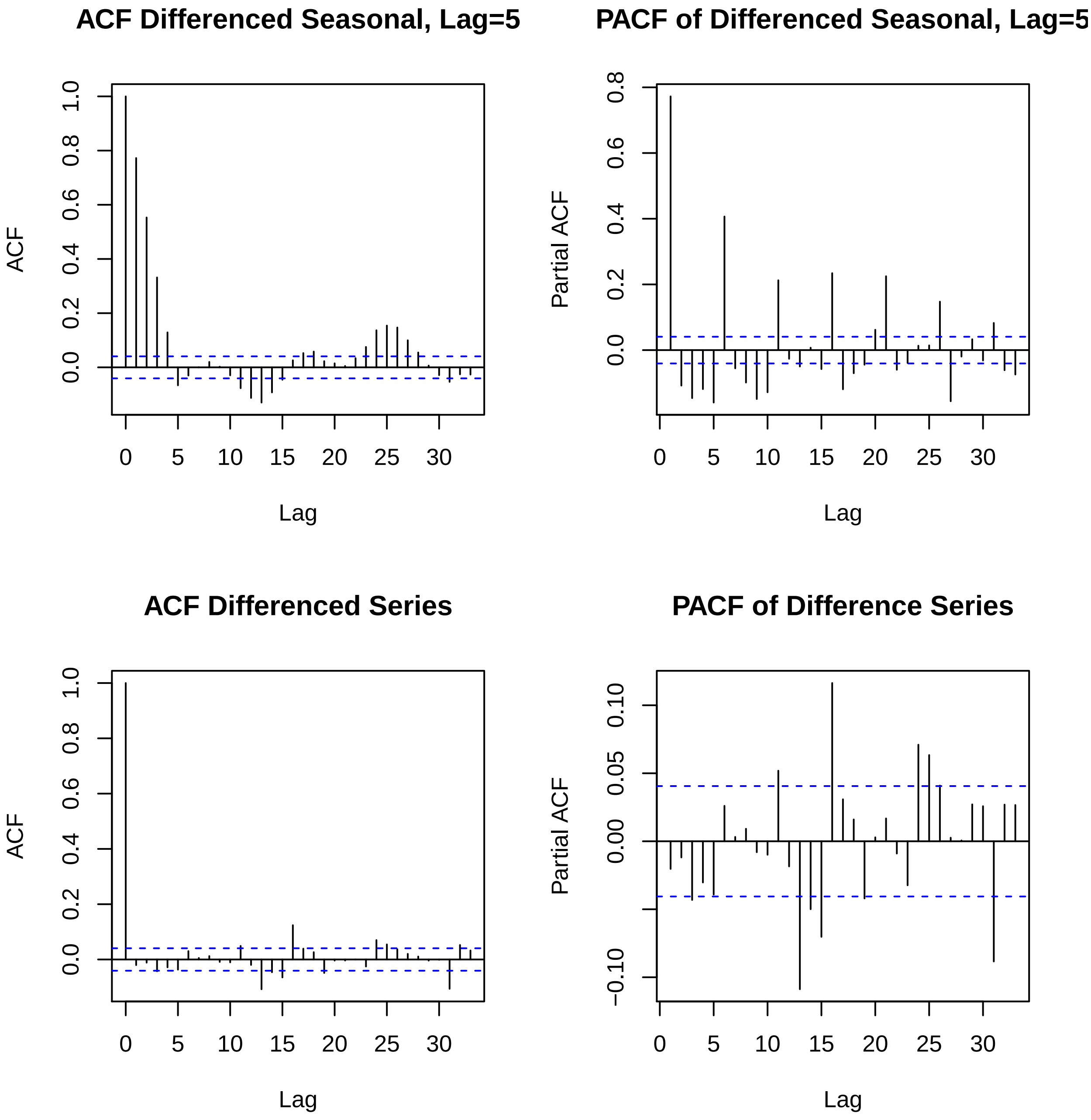






The seasonality of the series is not strong in the ACF of the differenced series. However, there are hints of a 5 day cycle that corresponds to the weekly frequency. There are stronger spikes that appear in lags 15 and 30 that support a multiple or harmonic of 5.

The ACF and PACF of the differenced seasonal series show evidence of an underlying MA(5) process



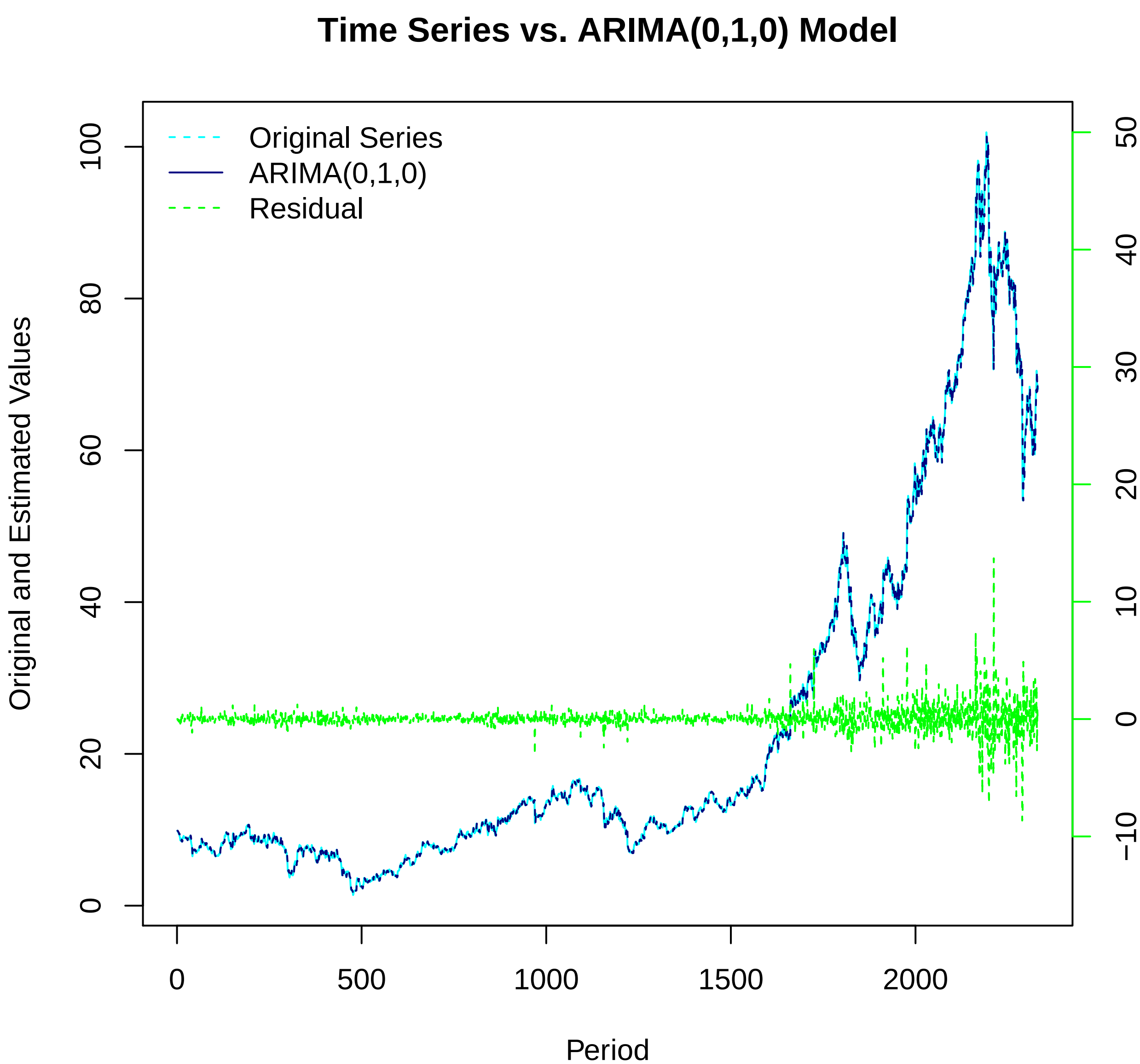
Model Selection

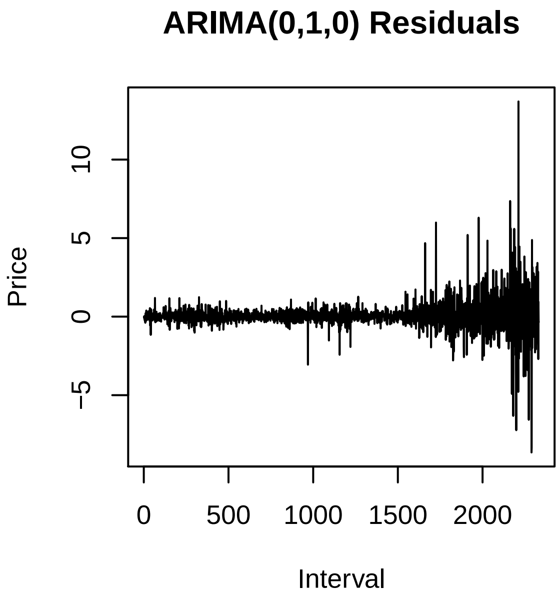
We use the `get.best.arima` function from the async materials and the Introductory Time Series with R book [Cowpertwait, Metcalfe - 2009] to perform a search for the best ARIMA model. The procedure results in an ARIMA(0,1,0) estimated model.

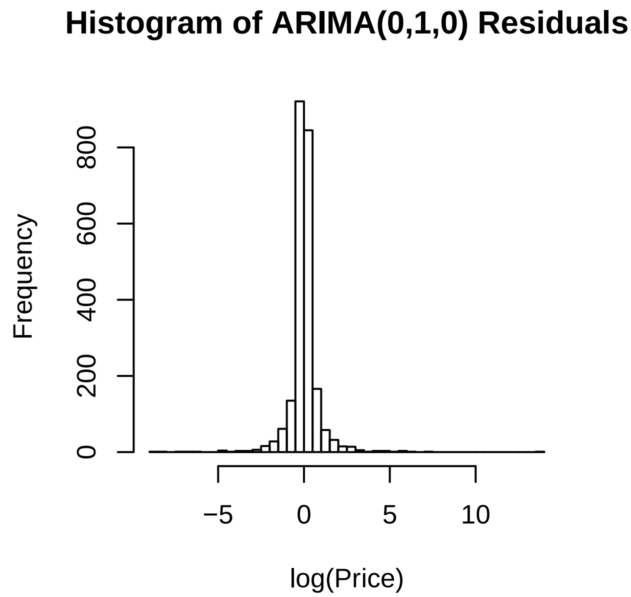
[1] "function"

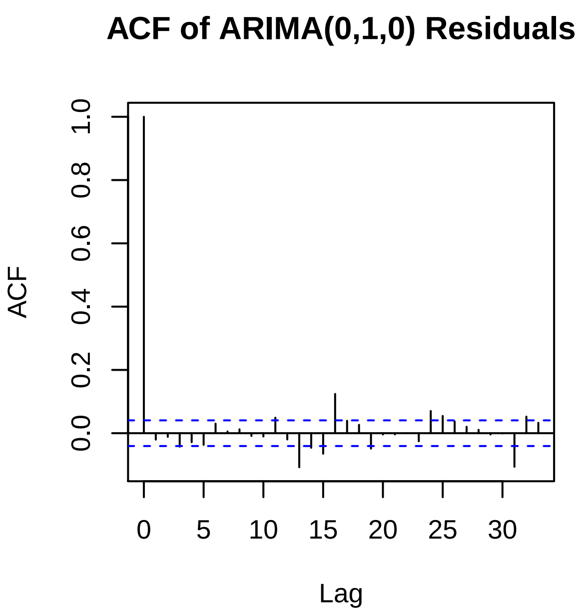
Table 7: Comparative Statistics

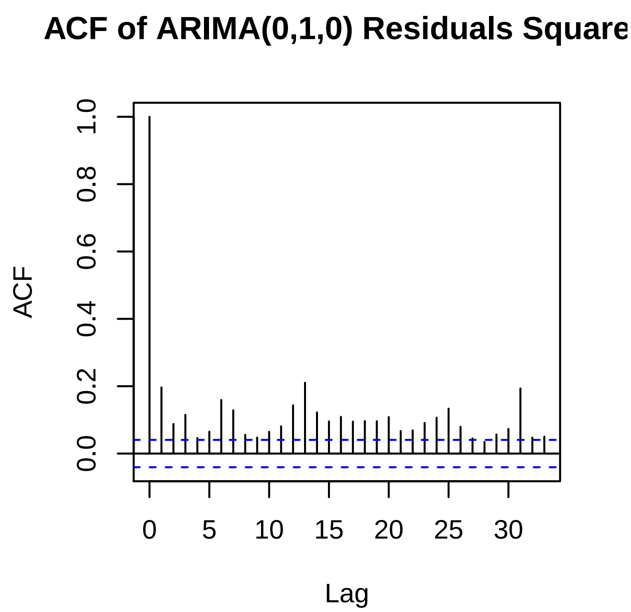
Statistic	N	Mean	St. Dev.	Min	Max
DXCS Series	2,332	23.2	23.4	1.4	101.9
ARIMA(0,1,0) Model	2,332	23.2	23.4	1.4	101.9
Residuals	2,332	0.02	0.9	−8.6	13.7







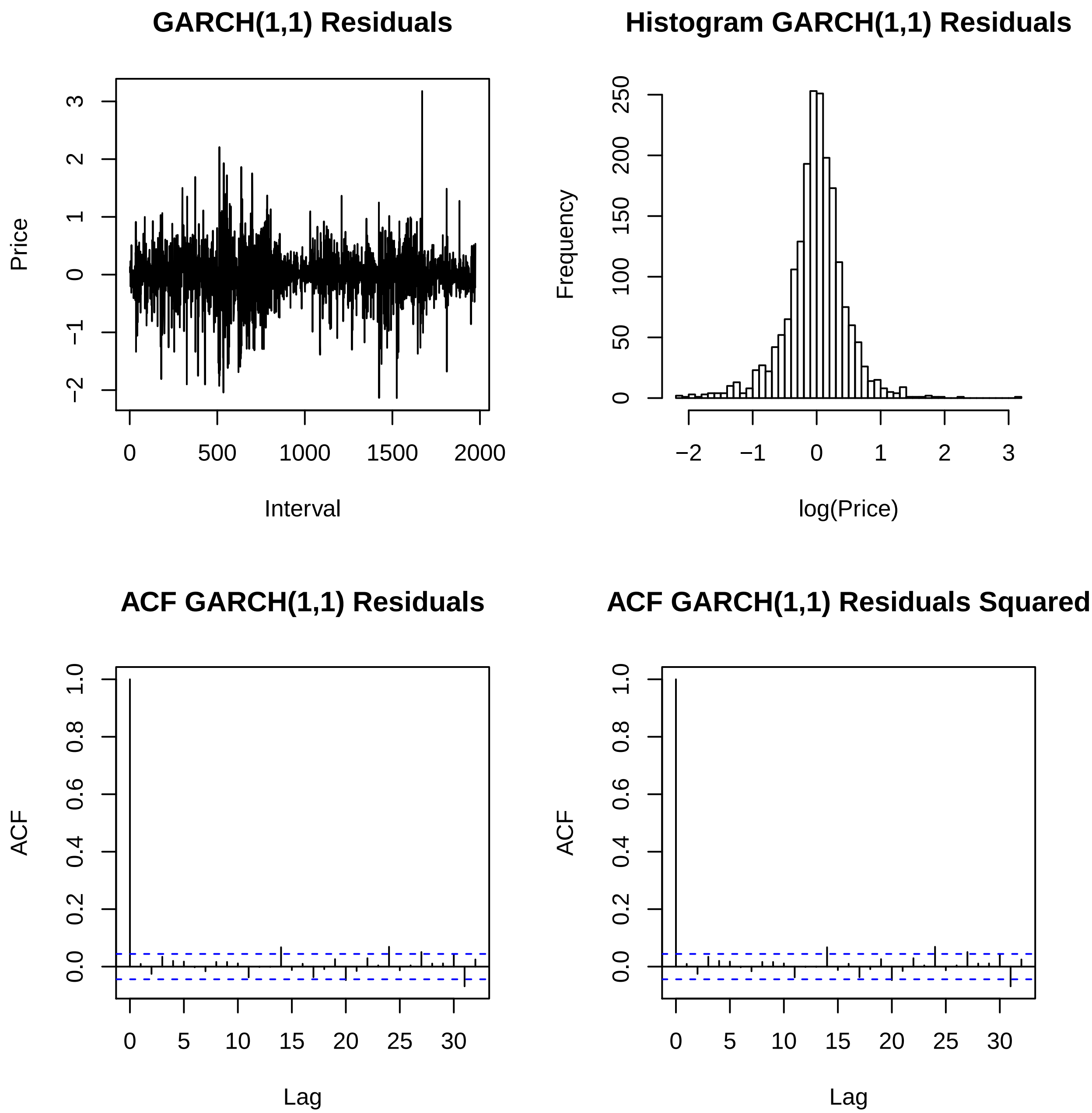




The residuals show an increased volatility at the right end of the residuals graph, and the PACF of the squared residuals shows definite autocorrelation of the residuals.

The ARIMA(0,1,0) model of the captures a good deal of the time series' behavior and it is a very parsimonious model.

Since the residuals of the ARIMA(0,1,0) model shows time-dependency we fit a GARCH model to the residual of the ARIMA model. The result is a GARCH(1,1) model where all parameters are significant. We fail to reject the hypothesis that the residuals are IID based on the results of the Ljung-Box test on the squared residuals. The summary of the GARCH model output is shown below.


GARCH(1,1) Residuals
Price
Interval
0
500
1000
1500
2000
−2
−1
0
1
2
3
Histogram GARCH(1,1) Residuals
Frequency
log(Price)
0
50
100
150
200
250
−2
−1
0
1
2
3
ACF GARCH(1,1) Residuals
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
ACF GARCH(1,1) Residuals Squared
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30

	GARCH(1,1)
mu	-0.006
	(0.008)
omega	0.011^{***}
	(0.003)
alpha1	0.153^{***}
	(0.026)
beta1	0.806^{***}
	(0.033)
Num. obs.	1974
AIC	1.125
Log Likelihood	1106.608

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

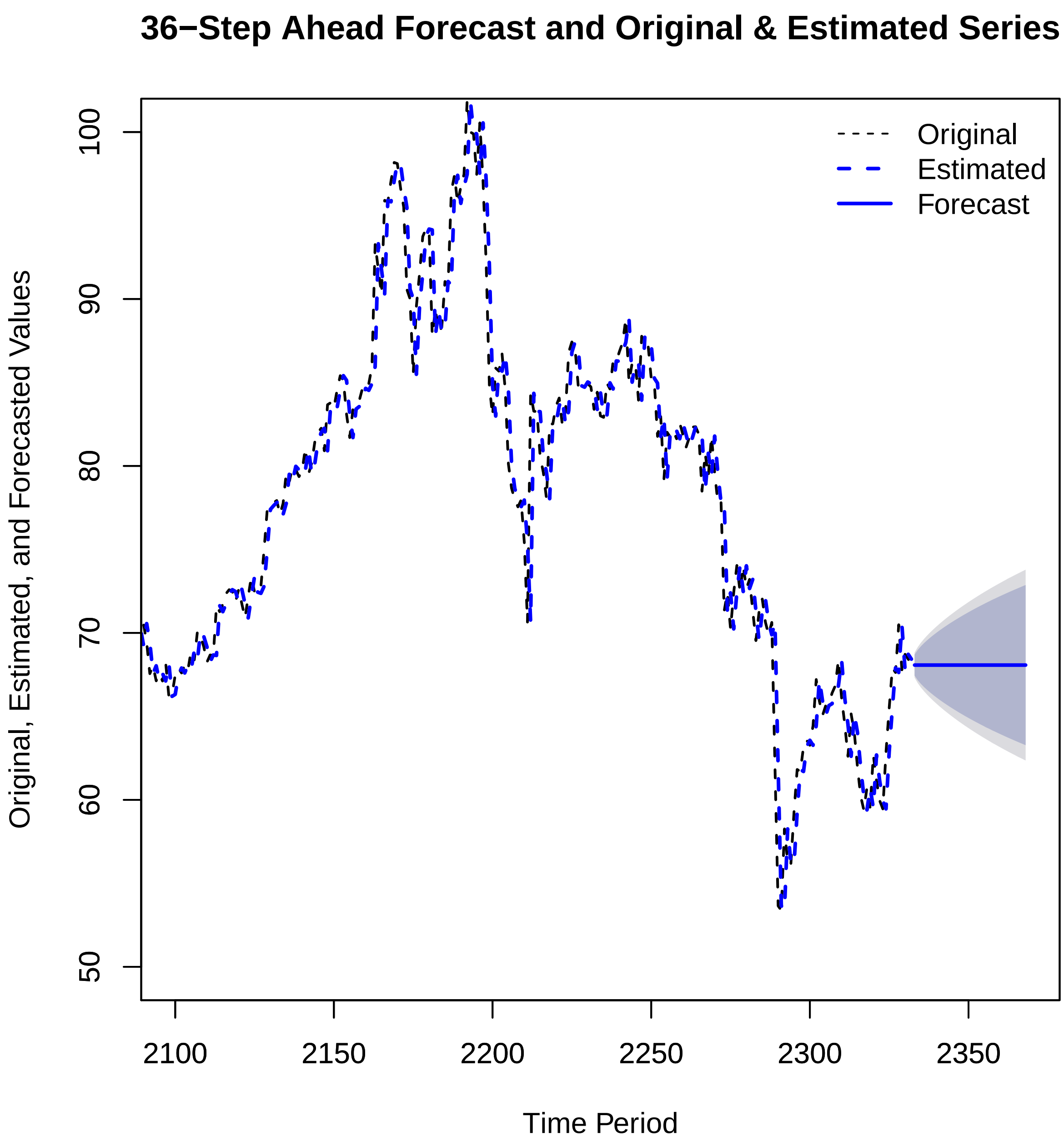
```
##
## Title:
##  GARCH Modelling
##
## Call:
##  garchFit(formula = ts1.fit1$residuals ~ garch(1, 1), trace = FALSE)
##
## Mean and Variance Equation:
##  data ~ garch(1, 1)
## <environment: 0x7f8d6da6db08>
##  [data = dem2gbp]
##
## Conditional Distribution:
##  norm
##
## Coefficient(s):
##         mu       omega      alpha1       beta1
## -0.0061903   0.0107614   0.1531341   0.8059737
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu     -0.006190    0.008462   -0.732 0.464447
## omega   0.010761    0.002838    3.793 0.000149 ***
## alpha1  0.153134    0.026422    5.796  6.8e-09 ***
## beta1   0.805974    0.033381   24.144  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  -1106.608    normalized:  -0.5605916
##
## Description:
##  Wed Apr 20 08:57:15 2016 by user:
##
##
```

```
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  1059.851  0
##  Shapiro-Wilk Test  R    W      0.9622848 0
##  Ljung-Box Test     R    Q(10)  10.12141  0.4299066
##  Ljung-Box Test     R    Q(15)  17.04349  0.3162711
##  Ljung-Box Test     R    Q(20)  19.29764  0.5025619
##  Ljung-Box Test     R^2  Q(10)  9.062556  0.5261773
##  Ljung-Box Test     R^2  Q(15)  16.07769  0.3769072
##  Ljung-Box Test     R^2  Q(20)  17.50715  0.6198388
##  LM Arch Test       R    TR^2   9.771217  0.6360238
##
## Information Criterion Statistics:
##      AIC      BIC      SIC     HQIC
## 1.125236 1.136559 1.125228 1.129396
```

The complete ARIMA(0,1,0)-GARCH(1,1) model looks like:

$$y_t - y_{t-1} = .010761 + .153134\epsilon_{t-1}^2 + .805974\hat{h}_{t-1}$$

All of these models exhibit time-dependent residuals so we fit a GARCH model to the residuals of the resulting model


36−Step Ahead Forecast and Original & Estimated Series
Original
Estimated
Forecast
Original, Estimated, and Forecasted Values
50
60
70
80
90
100
2100
2150
2200
2250
2300
2350
Time Period

Part 3 - Forecast Web Search Activity for "Global Warming"

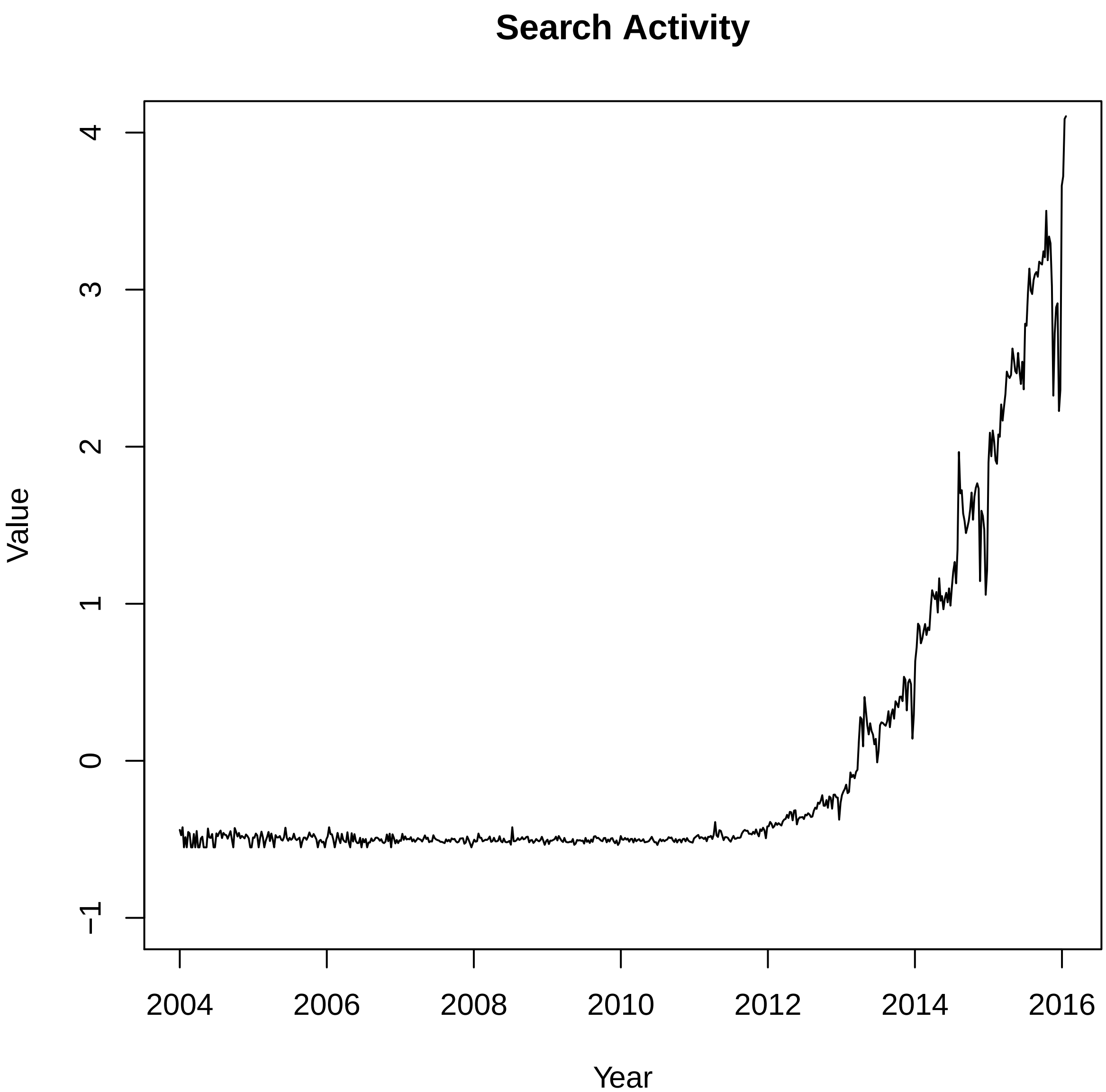
Exploratory Data Analysis

The time series is a weekly measurement of the frequency of a search phrase over a period from 1/4/2004 to 1/24/16 and we have no other information. The plot reveals that from 2004 to around 2012 there is very little activity; after 2012 the activity begins increasing persistently at a steep rate.

[1] "function"

Table 8: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Search Series	630	0.0000	1.00	−0.55	4.10



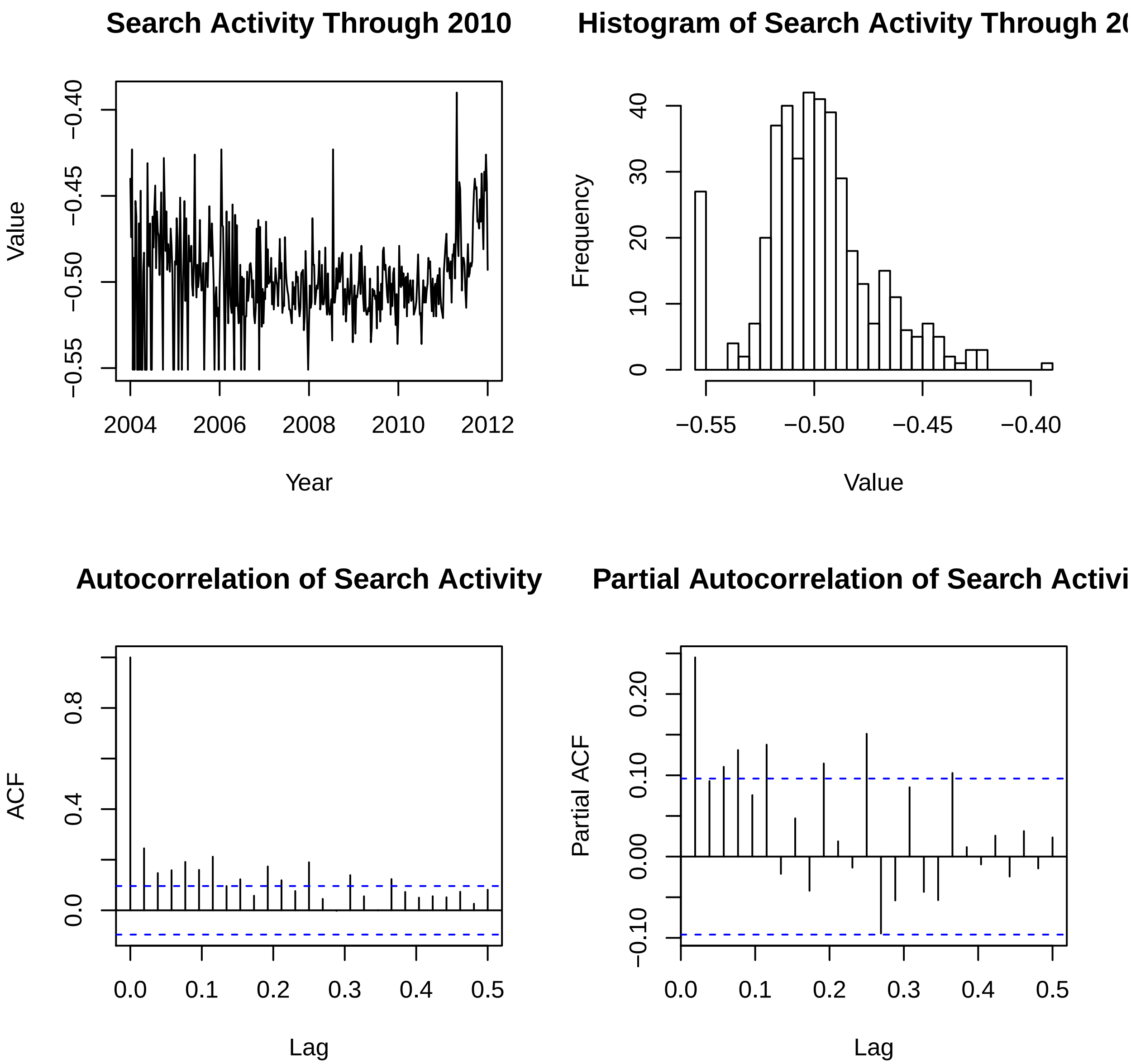
We can fit a model that does a better job of forecasting if we take that part of the series that contains the data since about 2012 on which to estimate a model. An analysis of the series indicates that the split point is around 2011. We split the series into 2004-2010 and 2011-2016 in order to capture the variation in the later part of the model. We measure this by observing the autocorrelation in each portion of the series to minimize the autocorrelation in the early series and maximize it in the later series.

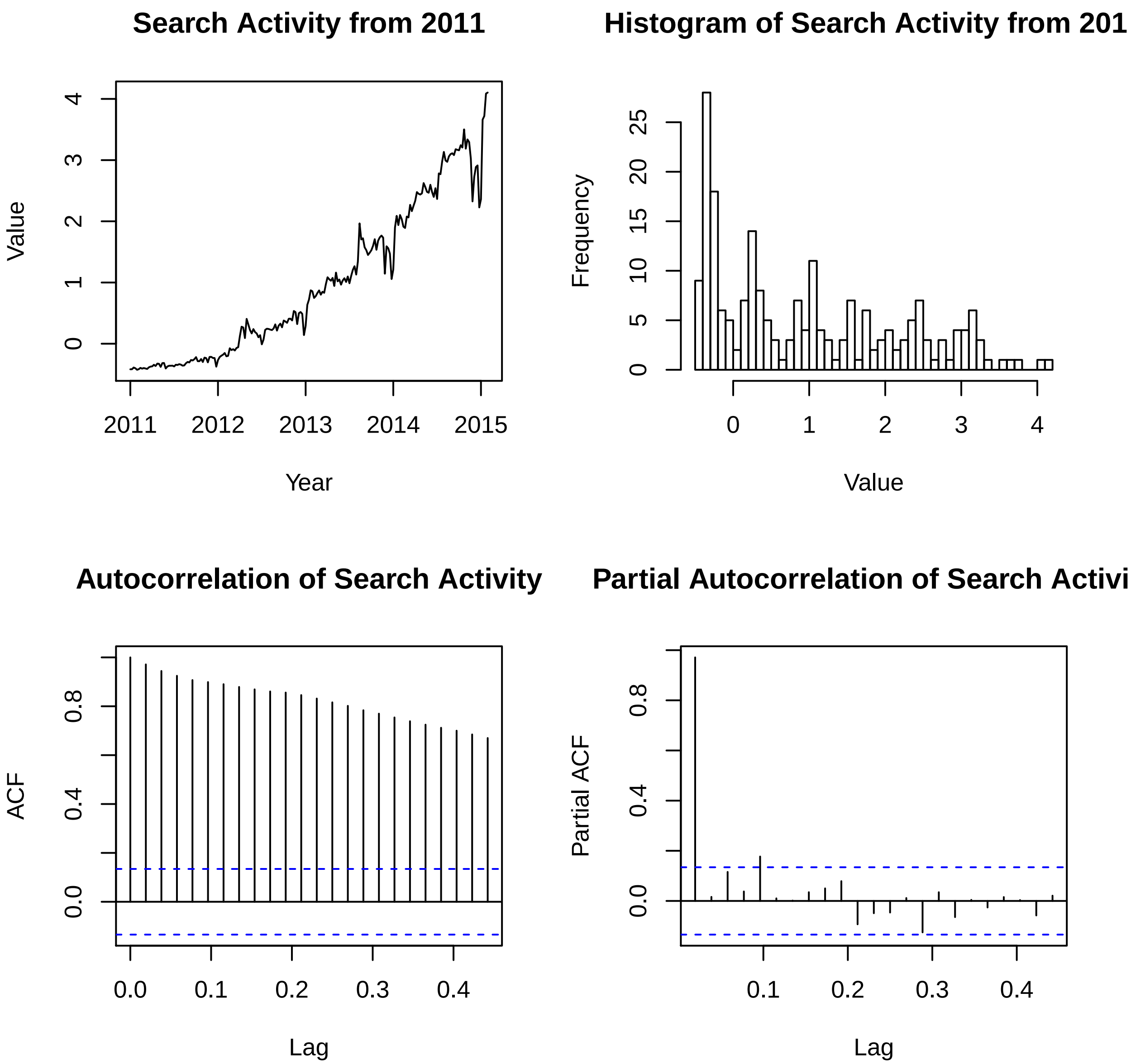
The following two pages show the comparative time series plots, ACF and PACF for the early and later time series.

[1] "function"

Table 9: Comparative Statistics

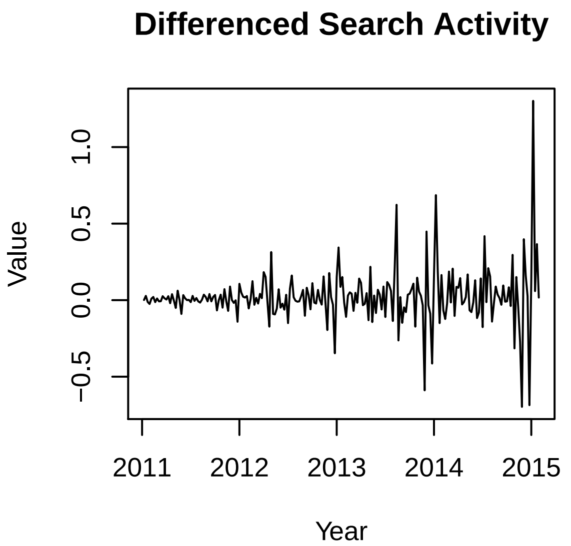
Statistic	N	Mean	St. Dev.	Min	Max
2004-2010	417	−0.50	0.03	−0.55	−0.39
2011-2016	213	0.98	1.23	−0.43	4.10

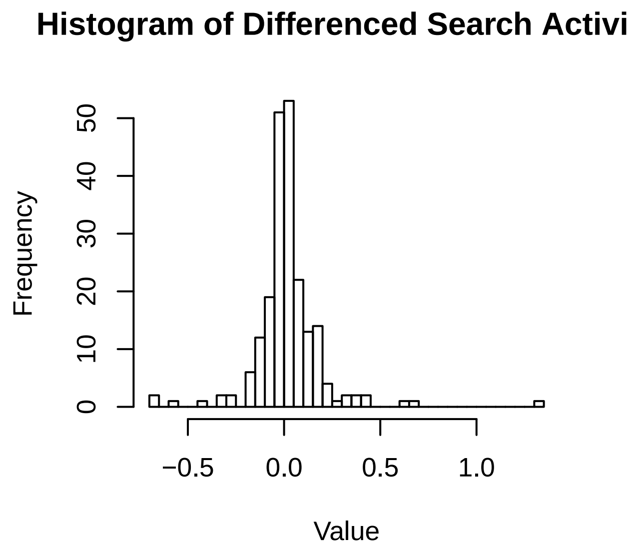

Search Activity Through 2010
Value
−0.55
−0.50
−0.45
−0.40
2004
2006
2008
2010
2012
Year
Histogram of Search Activity Through 20
Frequency
0
10
20
30
40
−0.55
−0.50
−0.45
−0.40
Value
Autocorrelation of Search Activity
ACF
0.0
0.4
0.8
0.0
0.1
0.2
0.3
0.4
0.5
Lag
Partial Autocorrelation of Search Activi
Partial ACF
−0.10
0.00
0.10
0.20
0.0
0.1
0.2
0.3
0.4
0.5
Lag

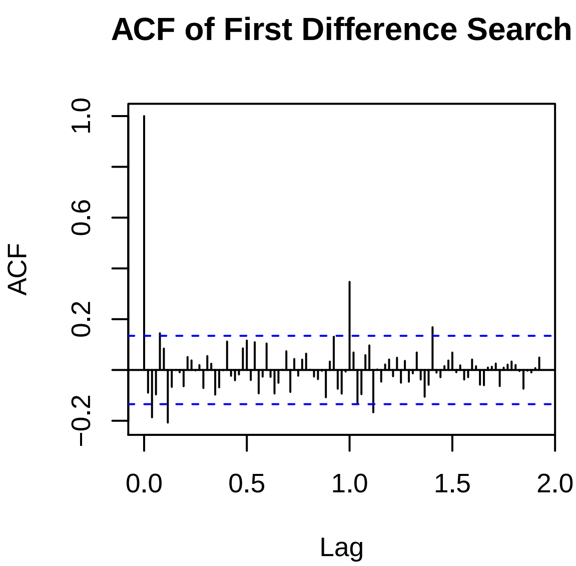

Search Activity from 2011
Value
Year
Histogram of Search Activity from 201
Frequency
Value
Autocorrelation of Search Activity
ACF
Lag
Partial Autocorrelation of Search Activi
Partial ACF
Lag

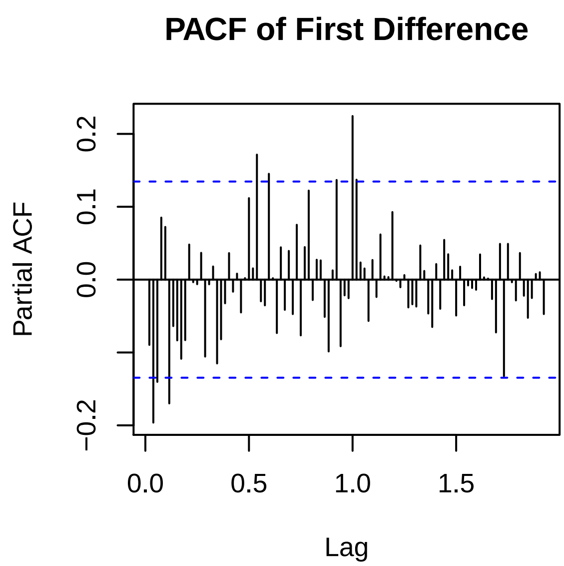
Time Series 2011-2015 Data Exploration

We can see from the plots on the previous page that the time series is not stationary. It has a persistent trend and exhibits seasonality with an increasing variance. We first examine the first difference of the series.



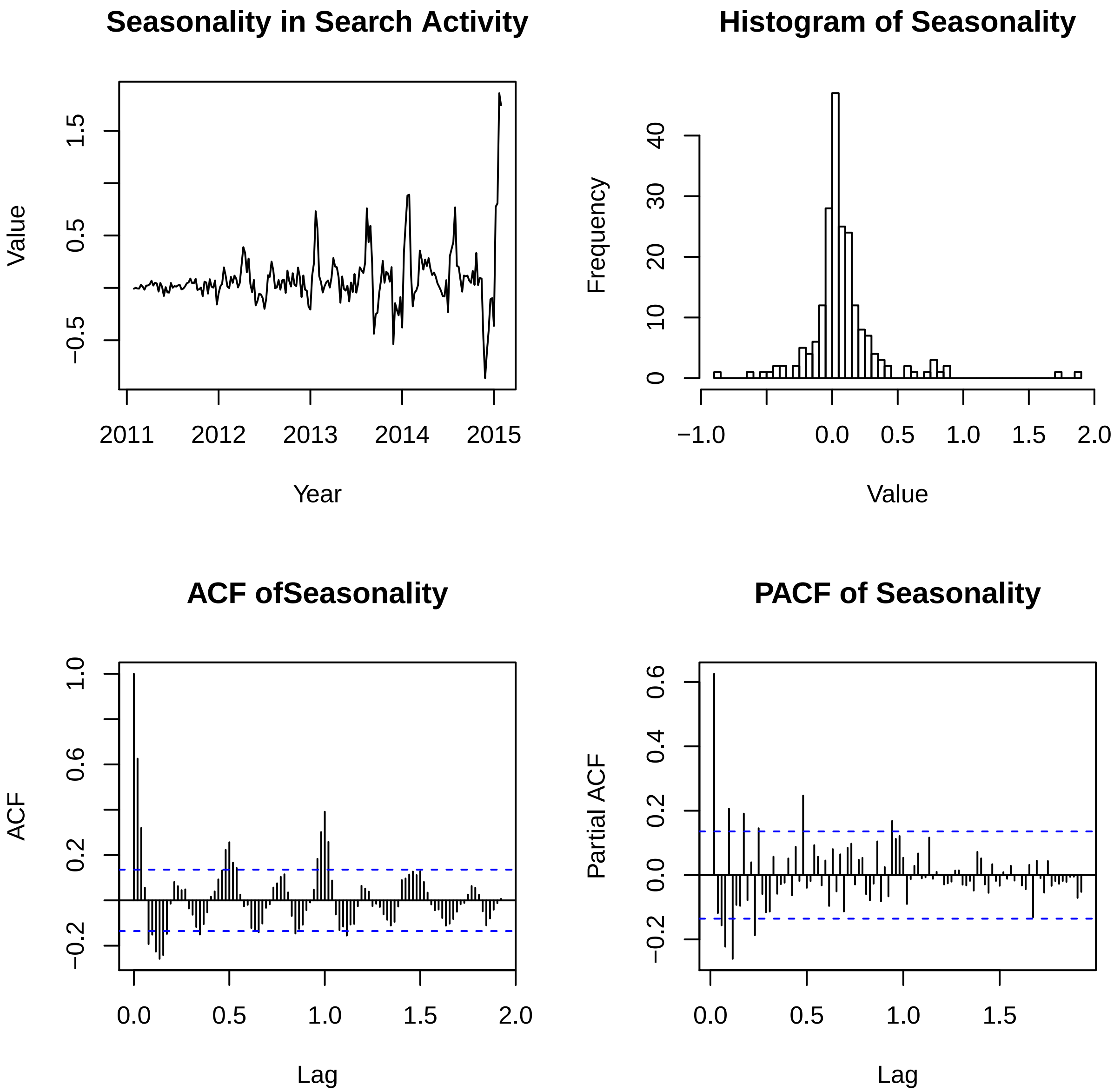






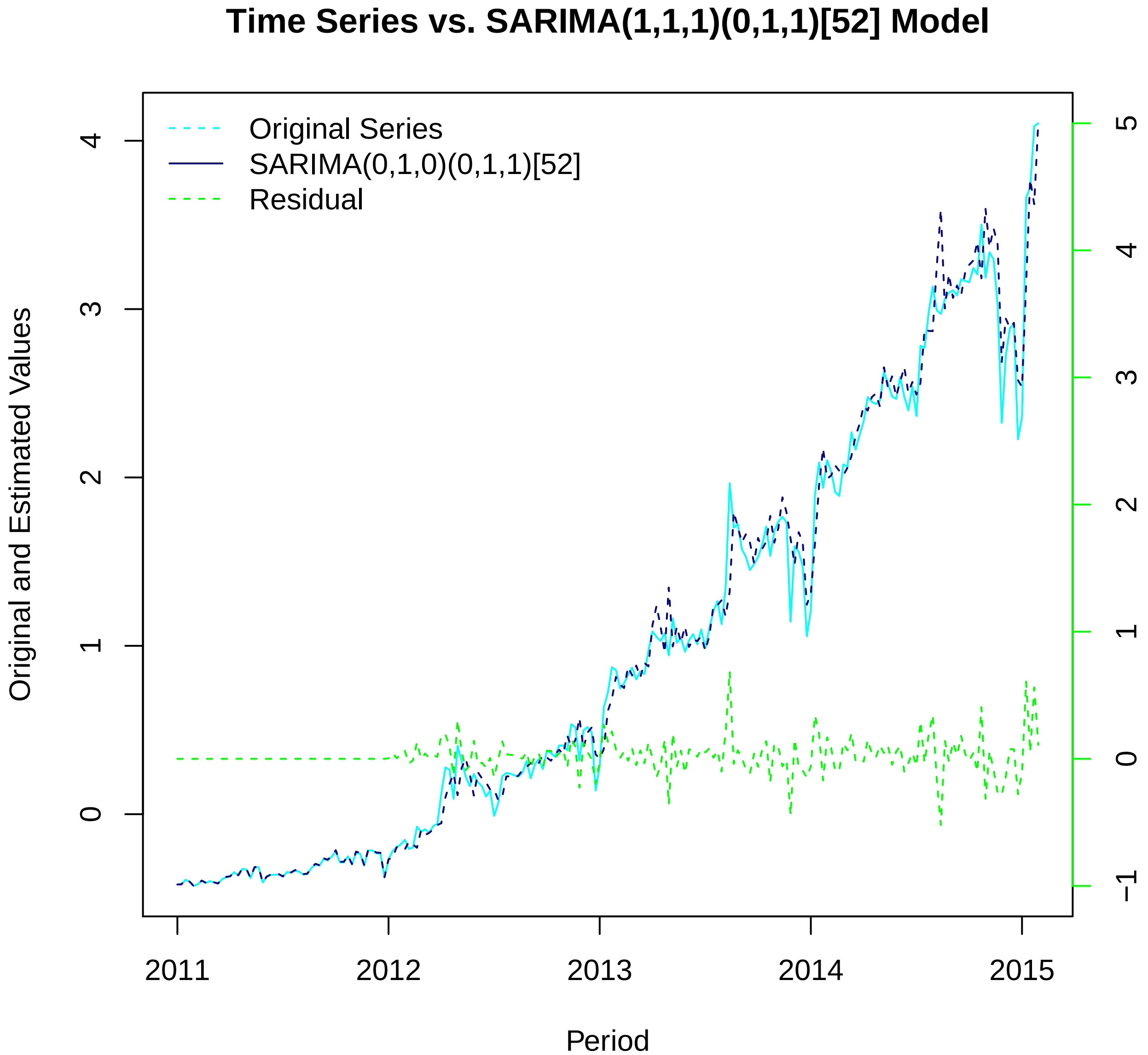
The resulting differenced series has increasing volatility over time. The ACF is indicative of an AR(2) process and the PACF indicates an MA process as well.

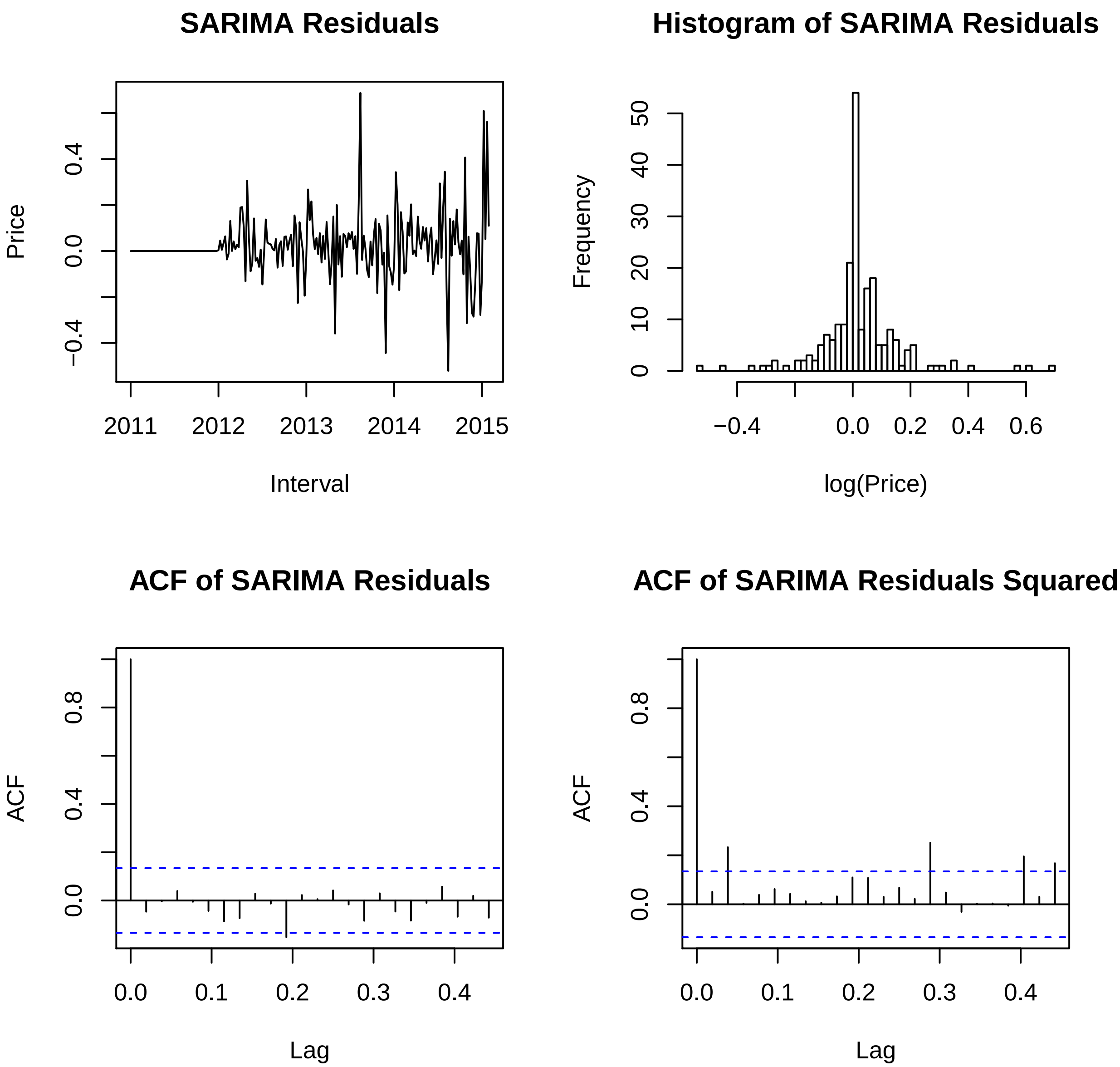
The seasonal component is present and seems to be in multiples of 4 and 13.



Modeling

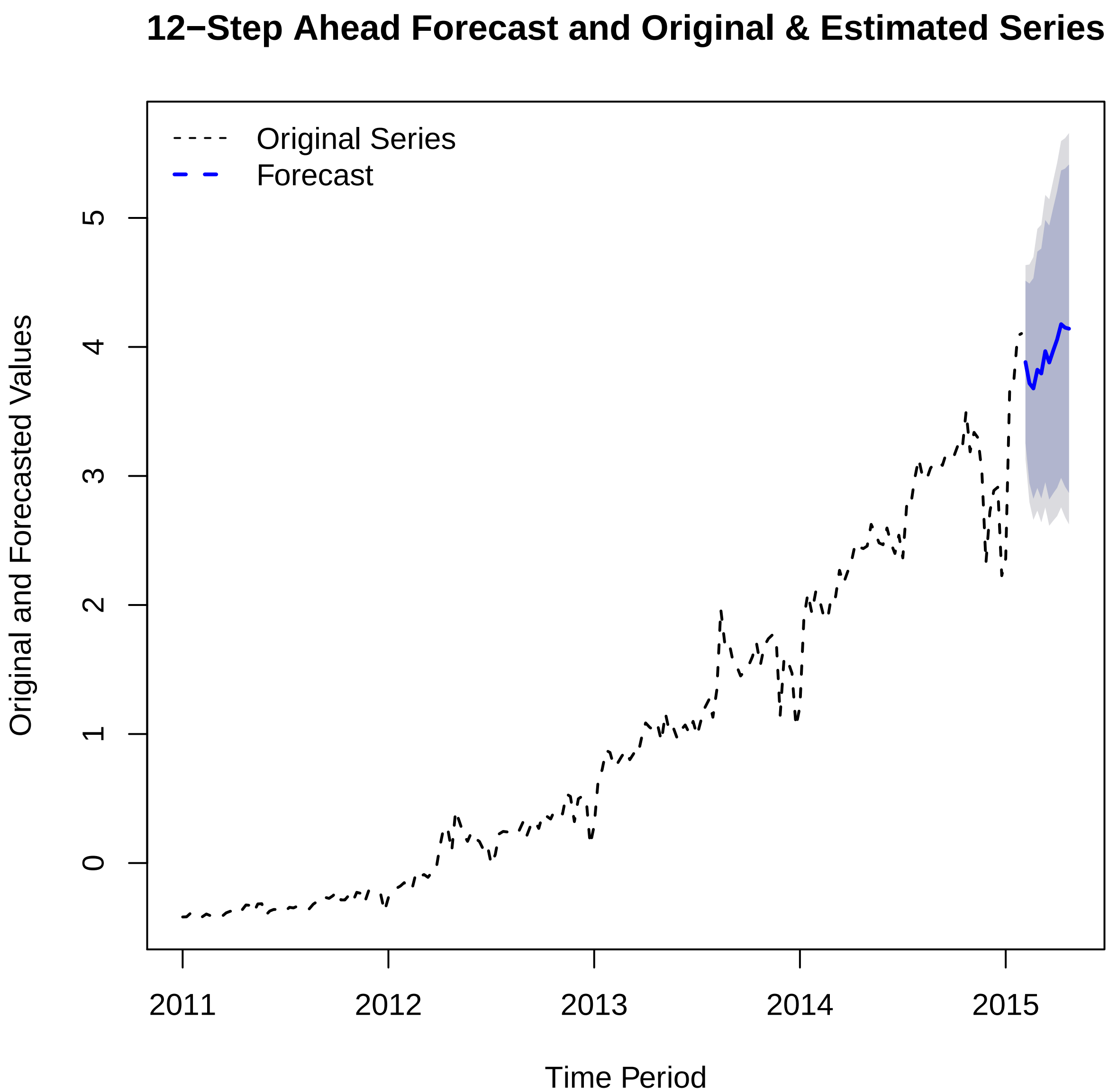
We use the `auto.arima` function to estimate the best SARIMA model and we obtain a SARIMA(1,1,1)(0,1,1)[52].




SARIMA Residuals
Price
−0.4
0.0
0.4
2011
2012
2013
2014
2015
Interval
Histogram of SARIMA Residuals
Frequency
0
10
20
30
40
50
−0.4
0.0
0.2
0.4
0.6
log(Price)
ACF of SARIMA Residuals
ACF
0.0
0.4
0.8
0.0
0.1
0.2
0.3
0.4
Lag
ACF of SARIMA Residuals Squared
ACF
0.0
0.4
0.8
0.0
0.1
0.2
0.3
0.4
Lag

Garch model

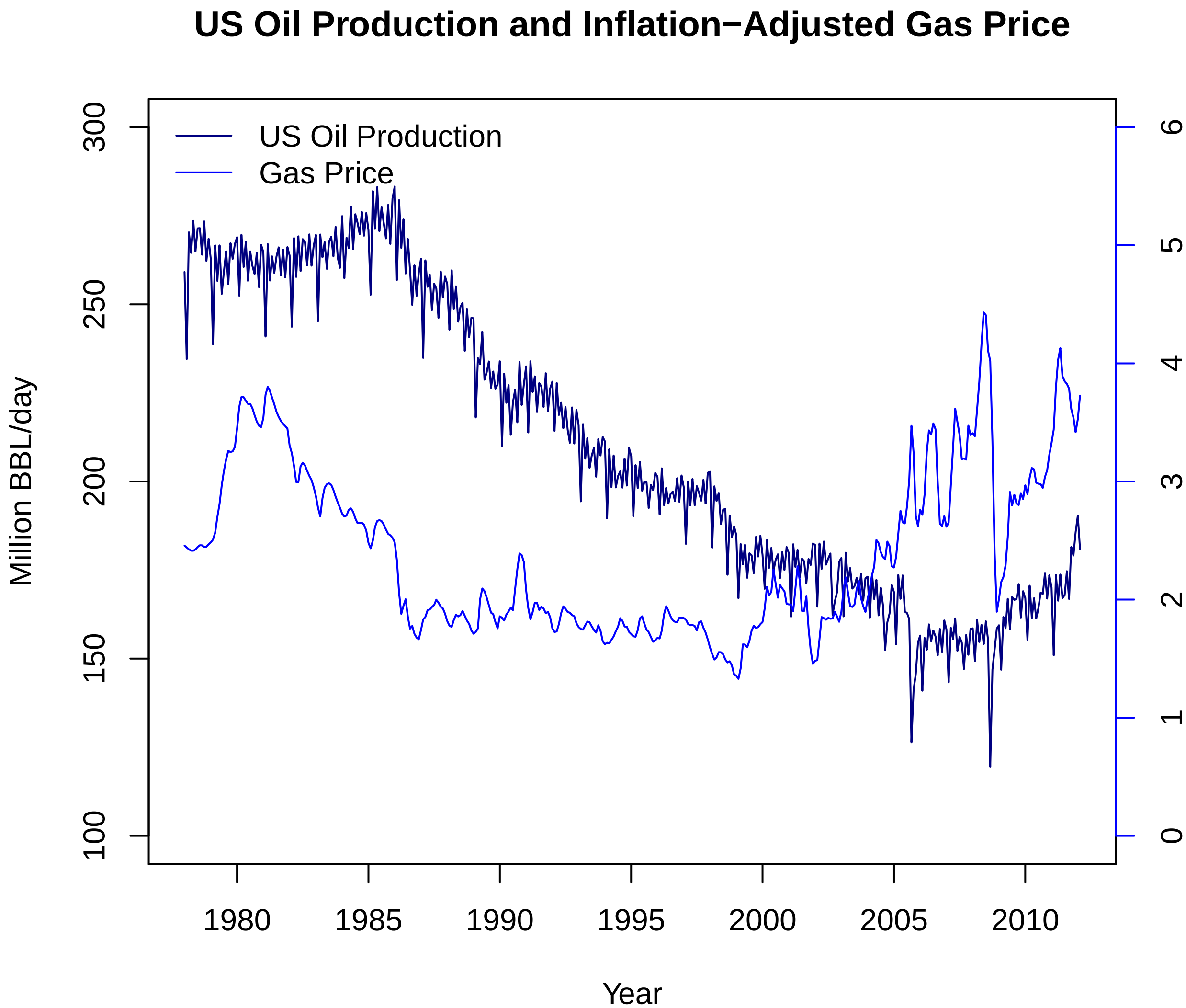
Forecast

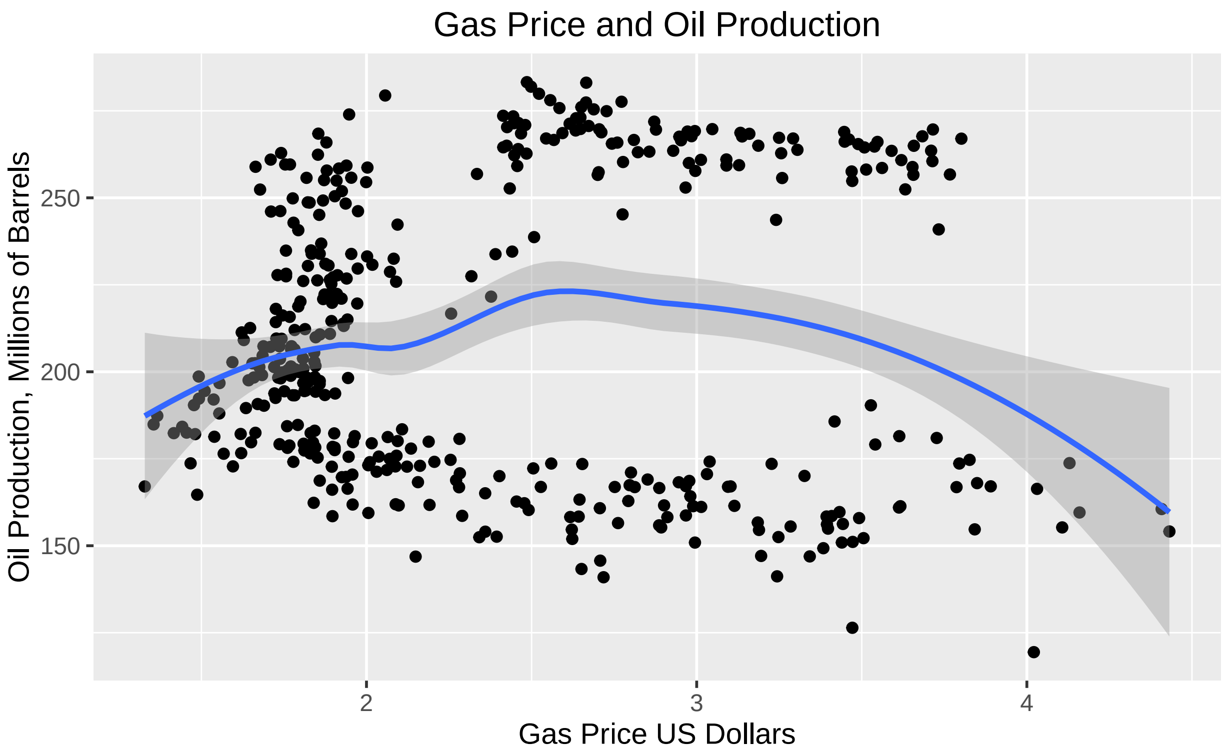


Part 4 - Forecast Inflation-Adjusted Gas Price

Exploratory Data Analysis

The *gasOil* series consists of 410 observations of two variables, *Production* and *Price*. As described the *Production* variable is in units of Million Barrels of Oil and *Price* is in inflation adjusted US Dollars. The observations are monthly beginning 1978/1/1 and ending 2012/2/1. The plots of the two variables indicate that both series appear similar to random walks: they have varying trends up and down over time. The *Production* series appears to have a seasonal component. The *Price* series may have a seasonal component as well but it isn't as prominent. A scatterplot of the two variables does not reveal any obvious correlation.



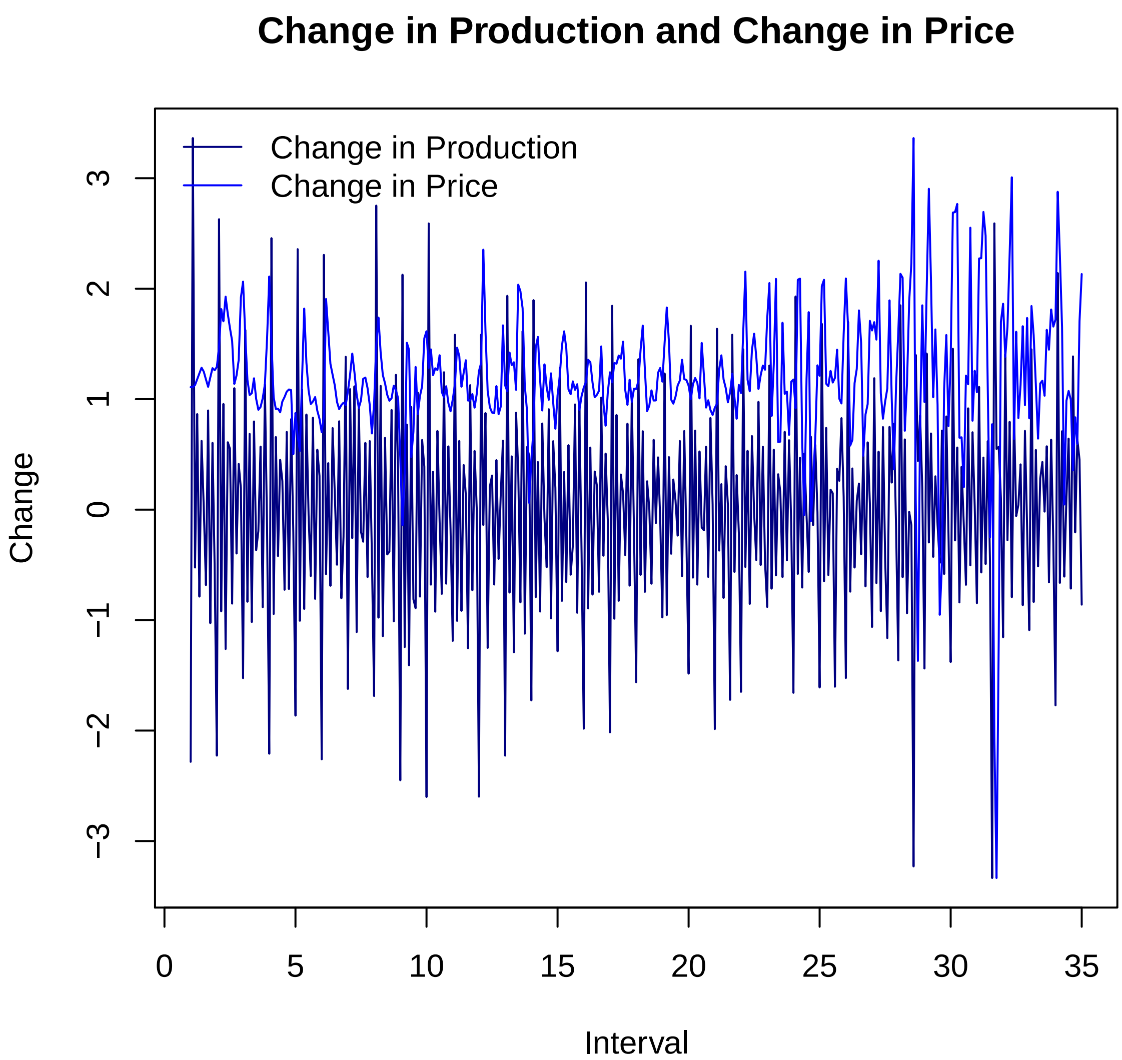


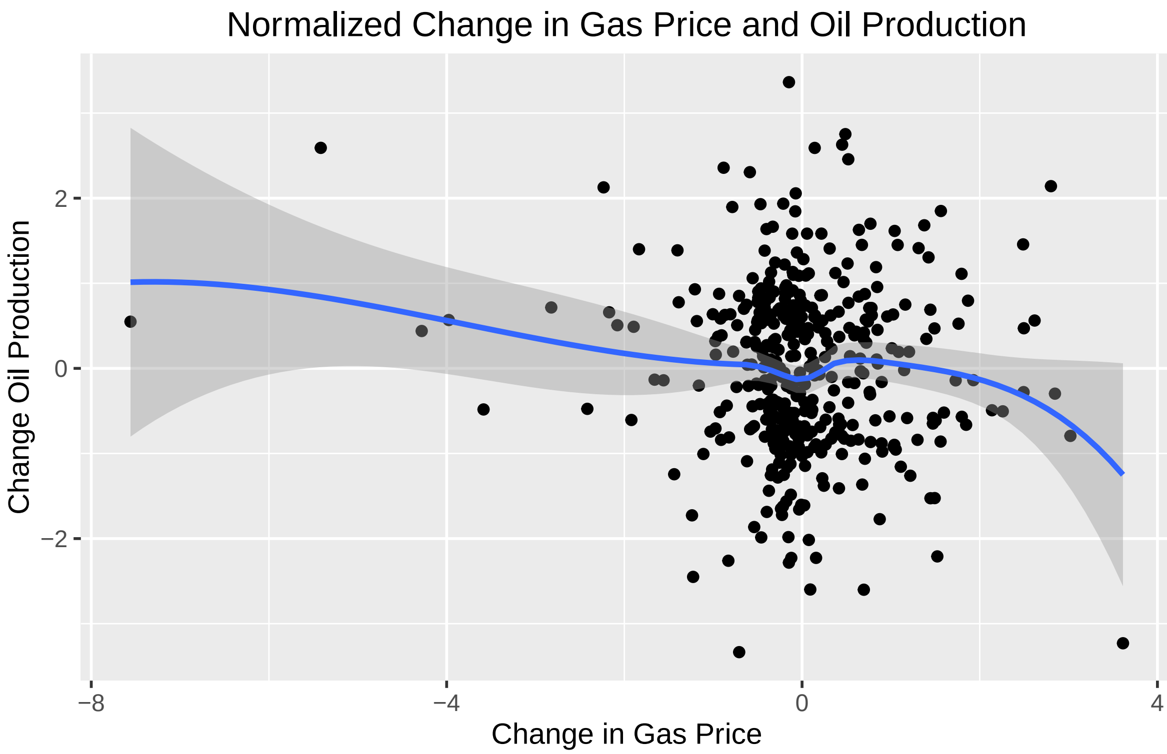
Change in Production and Change in Price

We take the difference of each series in order to compate the change in production to the change in price. We also scale the two variables so as to have zero mean and a standard deviation of 1.0 to make them more comparable. Now the scatterplot appears as a cluster about the origin of the graph.

Table 10: Descriptive Statistics of Differenced Series

Statistic	N	Mean	St. Dev.	Min	Max
Change in Production	409	−0.19	10.70	−35.85	35.78
Change in Price	409	0.003	0.13	−0.95	0.46
Scaled Change in Production	409	0.00	1.00	−3.33	3.36
Scaled Change in Price	409	0.00	1.00	−7.56	3.61


Change in Production and Change in Price
Change in Production
Change in Price
Change
3
2
1
0
−1
−2
−3
0
5
10
15
20
25
30
35
Interval



Model Recreation

The AP analysis states that there is no statistically significant correlation between oil production and gas prices. We can reconstruct a means of measuring a correlation and its significance using linear regression. We perform a regression on both the *Production* and *Price* variables as well as the *Change in Production* and *Change in Price*. The results of the regressions are summarized in the table, below.

The regression on the *change in production* and the *change in price* shows a marginally significant p-value but only at the $\alpha = 0.1$ level. The standard error shows that the estimate range crosses 0 so it is not distinguishable from 0. As expected all regression coefficients are not distinguishable from 0 and this reproduces the AP analysis conclusion that there is no statistically significate correlation between the two variables.

Critique

Comparing *Price* and *Production* in this way does not take into account the dependency on time. Each variable has it's own seasonal effects and its own volatility. A more careful analysis would compare the variables as time series and analyse the lagged relationships and seasonal factors.

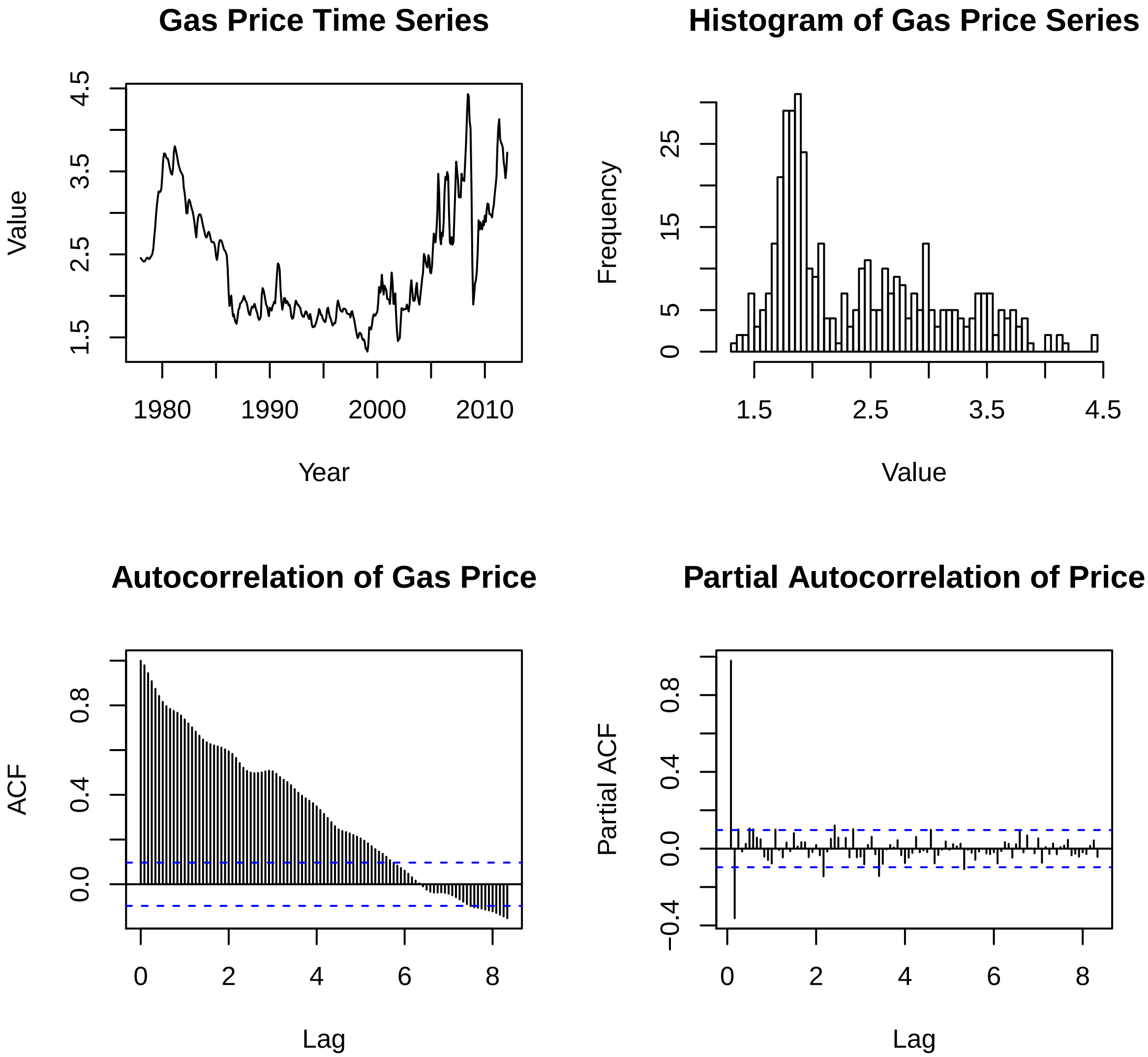
Table 11: Oil Production and Gas Price Model Summary

	Dependent variable:			
	Change in Production	Production	Change in Price	Price
	(1)	(2)	(3)	(4)
Change in Price	-0.08615^{*}			
	(0.04938)			
Price		1.66562		
		(2.96964)		
Change In Production			-0.08615^{*}	
			(0.04938)	
Production				0.00046
				(0.00082)
Constant	0.00000	206.02400^{***}	0.00000	2.29431^{***}
	(0.04932)	(7.39724)	(0.04932)	(0.17660)
Observations	409	410	409	410
R^2	0.00742	0.00077	0.00742	0.00077
Adjusted R^2	0.00498	−0.00168	0.00498	−0.00168
Residual Std. Error	0.99751 (df = 407)	41.91082 (df = 408)	0.99751 (df = 407)	0.69843 (df = 408)
F Statistic	3.04323^{*} (df = 1; 407)	0.31459 (df = 1; 408)	3.04323^{*} (df = 1; 407)	0.31459 (df = 1; 408)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Time Series Analysis

The *Price* series observations are...




```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.329   1.823   2.096   2.391   2.909   4.432
```