The Shocking Truth about History

Clayton Glenn

27 April, 2018

Abstract

This project uses pplications of SLR to Electric Prices using R, RMD, Word, Beamer, PDF, and HTML. The project clearly defines SLR in terms of theory, words, and graphics. Upon analysis of the Electric Prices throughout 1960-2011, the data tells a story of history of policies changing hands and being redistributed. Various graphs show the bulk of the project highlighted with text to clarify any uninvolved areas.

Contents

Figure 1: Clayton Glenn

Introduction

Throughout history, Energy Costs have been have been steadily increasing due to Inflation of Currency, Scarcity of Resources, and Green Laws Barring Steady Production. With the Supplimentation of Energy into our Everyday living Trending Further toward Green and Efficient Methods, we have seen Influxes of Costs when New Alternatives are Introduced.

This Project will Breakdown the Average Retail Prices of Electricity from 1960-2011 (Cents per Kilowatthour, Including Taxes) to Attempt to Find the Years that New and Improved Ways have been Implemented to Produce Energy.

Data has been Collected by the US Energy Information Association. (see "U.S. Energy Information Administration - Eia - Independent Statistics and Analysis" 2012)

Data of Electric Prices from 1960 - 2011

```
ep.df <- read.csv("ep.csv")
head(ep.df)
```

YEAR	IND	RES	CAR	COM	OTH	TOT
1960	1.1	2.6	0	2.4	1.9	1.8
1961	1.1	2.6	0	2.4	1.8	1.8
1962	1.1	2.6	0	2.4	1.9	1.8
1963	1.0	2.5	0	2.3	1.8	1.8
1964	1.0	2.5	0	2.2	1.8	1.7

YEAR	IND	RES	CAR	COM	OTH	TOT
1965	1.0	2.4	0	2.2	1.8	1.7

[1]

Variables of the data file

```
names(ep.df)
```

```
## [1] "YEAR" "IND"  "RES"  "CAR"  "COM"  "OTH"  "TOT"
```

For the sake of linear regression, we will only focus on The 2 Categorical variables Total Electric Prices (Discrete) corresponding to each Year (Continuous), and throughout the theory, the year will be a range of 0:51 to find accurate coefficients.[1]

Plot of Electric Prices from 1960 - 2011

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
x = 0:(length(ep.df$YEAR)-1)
ep.lm <- lm(TOT ~ x, data = ep.df)
g = ggplot(ep.lm, aes(x = x, y = TOT, color = x)) + geom_point()
g = g + ylab("Electric Prices (Cents per Kilowatthour)") + xlab("Year") + ggtitle("Electric Price ~ Yea
g = g + geom_smooth(method = "loess")
g
```

Data Collection and Why Data was Collected

“According to EIA, the Data Collection Began in 2003. The Category”Other" has been Replaced by “Transportation,” and the Categories “Commercial” and “Industrial” have been Redefined. The Data Represents Revenue from Electricity Retail Sales Divided by Electricity Retail Sales. Prices include State and Local Taxes, Energy or Demand Charges, Customer service charges, Environmental Surcharges, Franchise Fees, Fuel Adjustments, and Other Miscellaneous Charges Applied to End-Use Customers during Normal Billing Operations. Prices do not Include Deferred charges, Credits, or other Adjustments, such as Fuel or Revenue from Purchased Power, from previous Reporting Periods. Through 1979, The Data is for Classes A and B Privately Owned Electric Utilities only. (Class A Utilities are Those with Operating Revenues of 2.5 Million Dollars or More; Class B Utilities are Those with between 1 Million Dollars and 2.5 Million Dollars.) For 1980 - 1982, the Data is for Selected Class A Utilities Whose Electric Operating Revenues were 100 Million Dollars or More during the Previous Year. For 1983, the Data is for a Selected Sample of Electric Utilities. Beginning in 1984, the Data is for a Census of Electric Utilities. Beginning in 1996, the Data also Includes Energy Service Providers Selling to Retail Customers." -Per EIA

Figure 2: Alternative Energy Resources (http://energyfive.net/2018/02/05/what-is-hydroelectric-power-plant/)

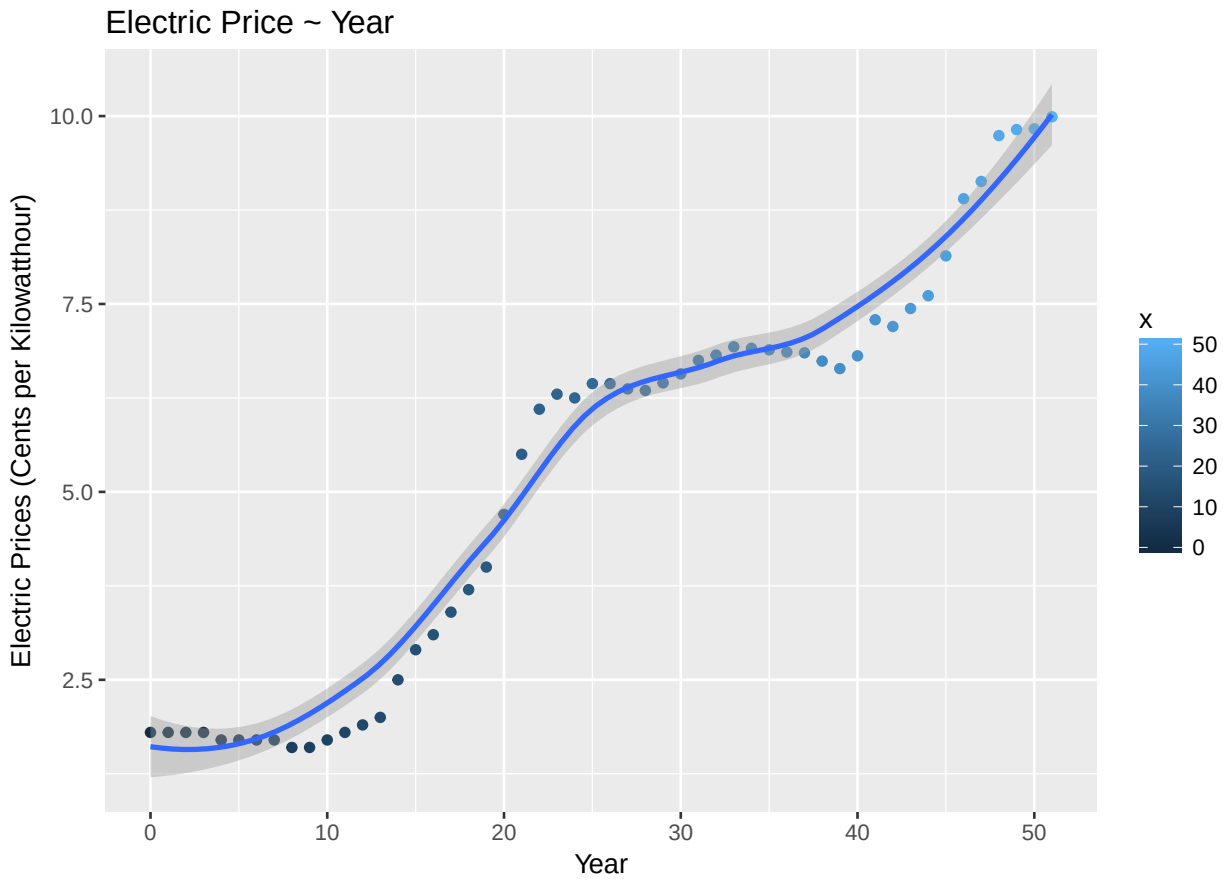


Figure 3: Electricity Prices

[1]

The Data was Initially Collected to Analyze Energy Prices. In Hindsight, the Data has been used to Generate Forecasts and Projections for Energy Prices in the Future.

Why so Interesting

I have a Monumental Interest with the Data mostly due to our Carbon Footprint. Hopefully I will be able to Research the Located Spots in the Historical Timeline that led to better Efficiency of Energy Production and Consumption.

SLR Theory

Simple Linear Regression can be used to create a formula to predict points of interest, find outliers, and show variation of the points about the mean and the regressive line itself. The formula I produced outputs a list of useful equations that output my form of SLR gold as you could call it. Each formula will be broken down to show its very meaning behind the data.

```
mysummary = function(x=x,y=y){
  ssxx  = sum((x - mean(x)) ^ 2)
  ssxy  = sum((x - mean(x)) * (y - mean(y)))
  ssyy  = sum((y - mean(y)) ^ 2)
  b1hat = ssxy / ssxx
  b0hat = mean(y) - b1hat * mean(x)
  yhat  = b0hat + b1hat * x
  rss   = sum((y - yhat) ^ 2)
  mss   = sum((yhat - mean(y)) ^ 2)
  tss   = sum((y - mean(y)) ^ 2)
  yerr  = y - yhat
  rerr  = y - b0hat - b1hat*x
  sdhat = sd(y) / ssxx
  r     = ssxy / sqrt(ssxx*ssyy)

  return(list(ssxx = ssxx, ssxy = ssxy, ssyy = ssyy, b1hat = b1hat, b0hat = b0hat, yhat = yhat, rss = rs
```

```
}
ep.summary = mysummary(x, ep.df$TOT)
```

(see "Assumptions of Linear Regression" 2018) ## Sum of Squares of X

$$SS_{xx} = \sum_{i=1}^{n}(x - \bar{x})^2 = S_x^2(n-1)$$

Sum of Squares of X shows the variability of the Independent Variables or the Covariance of x.

```
ep.summary$ssxx
```

```
## [1] 11713
```

We can check SSxx by taking the variance of the TOTAL electric prices across the US and multiplying by the array of years minus 1.

```
var(x) * (length(x) - 1)
```

```
## [1] 11713
```

Sum of Squares of XY

$$SS_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = S_{xy}^2(n-1)$$

Sum of Squares of X and Y measures how X and Y varies together or the Covariance of x,y.

```
ep.summary$ssxy
```

```
## [1] 2009.71
```

SSxy can be checked by taking the covariance of x and y and multiplying it by the number of years data was taken minus 1.

```
cov(x,ep.df$TOT) * (length(x) - 1)
```

```
## [1] 2009.71
```

Slope of Linear Model

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$$

By taking the sum of squares of xy and dividing by sum of squares of x, we obtain Beta1hat. The predicted value of Beta 1 shows the slope of our linear model we are going to apply to the data.

```
ep.summary$b1hat
```

```
## [1] 0.1715794
```

Intercept of Linear Model

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

After finding the predicted value of Beta1, we can now find the predicted value for Beta-not. Beta-not is the y-intercept in our linear model formula and can be found by taking difference of the mean of y and the predicted value of Beta1 times the mean of x.

```
ep.summary$b0hat
```

```
## [1] 0.9124165
```

Residual of Xi, Error

$$Residual_i = \hat{\epsilon}_i = Y_i - \hat{Y}_i$$

The residual can also be denoted as the prediction of the error difference between the actual y value and its prediction corresponding the the x value. Each y value has an error to the linear model.

```
ep.summary$yerr
```

```
##  [1]  0.88758345  0.71600401  0.54442457  0.37284513  0.10126569
##  [6] -0.07031375 -0.24189320 -0.41347264 -0.68505208 -0.85663152
## [11] -0.92821096 -0.99979040 -1.07136985 -1.14294929 -0.81452873
## [16] -0.58610817 -0.55768761 -0.42926705 -0.30084650 -0.17242594
## [21]  0.35599462  0.98441518  1.41283574  1.44125630  1.21967685
## [26]  1.23809741  1.06651797  0.82493853  0.63335909  0.56177965
## [31]  0.51020020  0.51862076  0.41704132  0.35546188  0.16388244
## [36] -0.02769700 -0.22927644 -0.41085589 -0.69243533 -0.96401477
## [41] -0.96559421 -0.65717365 -0.91875309 -0.85033254 -0.85191198
## [46] -0.49349142  0.09492914  0.15334970  0.59177026  0.50019081
## [51]  0.33861137  0.32703193
```

Residual Sum of Squares or Error Sum of Squares

$$RSS = \sum_{i=1}^{n}(y_i - \hat{y})^2 = \sum_{i=1}^{n}(\epsilon_i)^2 = s_r^2(n-2)$$

The residual sum of squares measures the difference of the linear model as opposed to the real data. As the Residual sum of squares tends smaller, the linear model better fits the data.

```
ep.summary$rss
```

```
## [1] 27.0884
```

To check the Residual sum of squares, we can take the sum of the errors of each x value squared.

```
sum(ep.summary$yerr^2)
```

```
## [1] 27.0884
```

Model Sum of Squares

$$MSS = \sum_{i=1}^{n}(\hat{y} - \bar{y})^2$$

The model sum of squares is the variation of the predicted value of y to the mean of y.

```
ep.summary$mss
```

```
## [1] 344.8249
```

Total Sum of Squares or Sum of Squares of Y

$$TSS = SS_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = S_y^2(n-1) = RSS + MSS$$

The total sum of squares is the squared difference of the actual y value as opposed to the mean of y. A steeper linear model shows a higher total sum of squares and vice versa.

```
ep.summary$tss
```

```
## [1] 371.9133
```

We can check the validity of the total sum of squares multiplying the variance of y by total years minus 1

```
var(ep.df$TOT) * (length(ep.df$TOT) - 1)
```

```
## [1] 371.9133
```

Another way to check the validity of the Total sum of squares is by pathagorean theorem. A^2+B^2=C^2 where c^2 is the TSS, b^2 is RSS, and a^2 is MSS.

```
ep.summary$mss + ep.summary$rss
```

```
## [1] 371.9133
```

Linear Model vs Actual Data

Equation of Linear Model

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

After finding the bulk of our equation, we can now find yhat. Each x value has its own unique predicted value in the model and can be found by the equation above. Yhat can be used to find the Residual and Model sum of squares and the error for each x value to the linear model.

```
ep.summary$yhat
```

```
##  [1] 0.9124165 1.0839960 1.2555754 1.4271549 1.5987343 1.7703138 1.9418932
##  [8] 2.1134726 2.2850521 2.4566315 2.6282110 2.7997904 2.9713698 3.1429493
## [15] 3.3145287 3.4861082 3.6576876 3.8292671 4.0008465 4.1724259 4.3440054
## [22] 4.5155848 4.6871643 4.8587437 5.0303231 5.2019026 5.3734820 5.5450615
## [29] 5.7166409 5.8882204 6.0597998 6.2313792 6.4029587 6.5745381 6.7461176
## [36] 6.9176970 7.0892764 7.2608559 7.4324353 7.6040148 7.7755942 7.9471737
## [43] 8.1187531 8.2903325 8.4619120 8.6334914 8.8050709 8.9766503 9.1482297
## [50] 9.3198092 9.4913886 9.6629681
```

Equation of Data

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 x + \epsilon_i$$

Now that we have found the predicted values of beta1 and beta2, we can find the actual position of each data point from the linear model regarding to the data file of Total Electric Prices.

```
ep.summary$b0hat+ep.summary$b1hat*x+ep.summary$rerr
```

```
##  [1] 1.80 1.80 1.80 1.80 1.70 1.70 1.70 1.70 1.60 1.60 1.70 1.80 1.90 2.00
## [15] 2.50 2.90 3.10 3.40 3.70 4.00 4.70 5.50 6.10 6.30 6.25 6.44 6.44 6.37
## [29] 6.35 6.45 6.57 6.75 6.82 6.93 6.91 6.89 6.86 6.85 6.74 6.64 6.81 7.29
## [43] 7.20 7.44 7.61 8.14 8.90 9.13 9.74 9.82 9.83 9.99
```

We can check the validity of our theoretical model of the actual data by just comparing the formula to the actual data itself

```
ep.df$TOT
```

```
##  [1] 1.80 1.80 1.80 1.80 1.70 1.70 1.70 1.70 1.60 1.60 1.70 1.80 1.90 2.00
## [15] 2.50 2.90 3.10 3.40 3.70 4.00 4.70 5.50 6.10 6.30 6.25 6.44 6.44 6.37
## [29] 6.35 6.45 6.57 6.75 6.82 6.93 6.91 6.89 6.86 6.85 6.74 6.64 6.81 7.29
## [43] 7.20 7.44 7.61 8.14 8.90 9.13 9.74 9.82 9.83 9.99
```

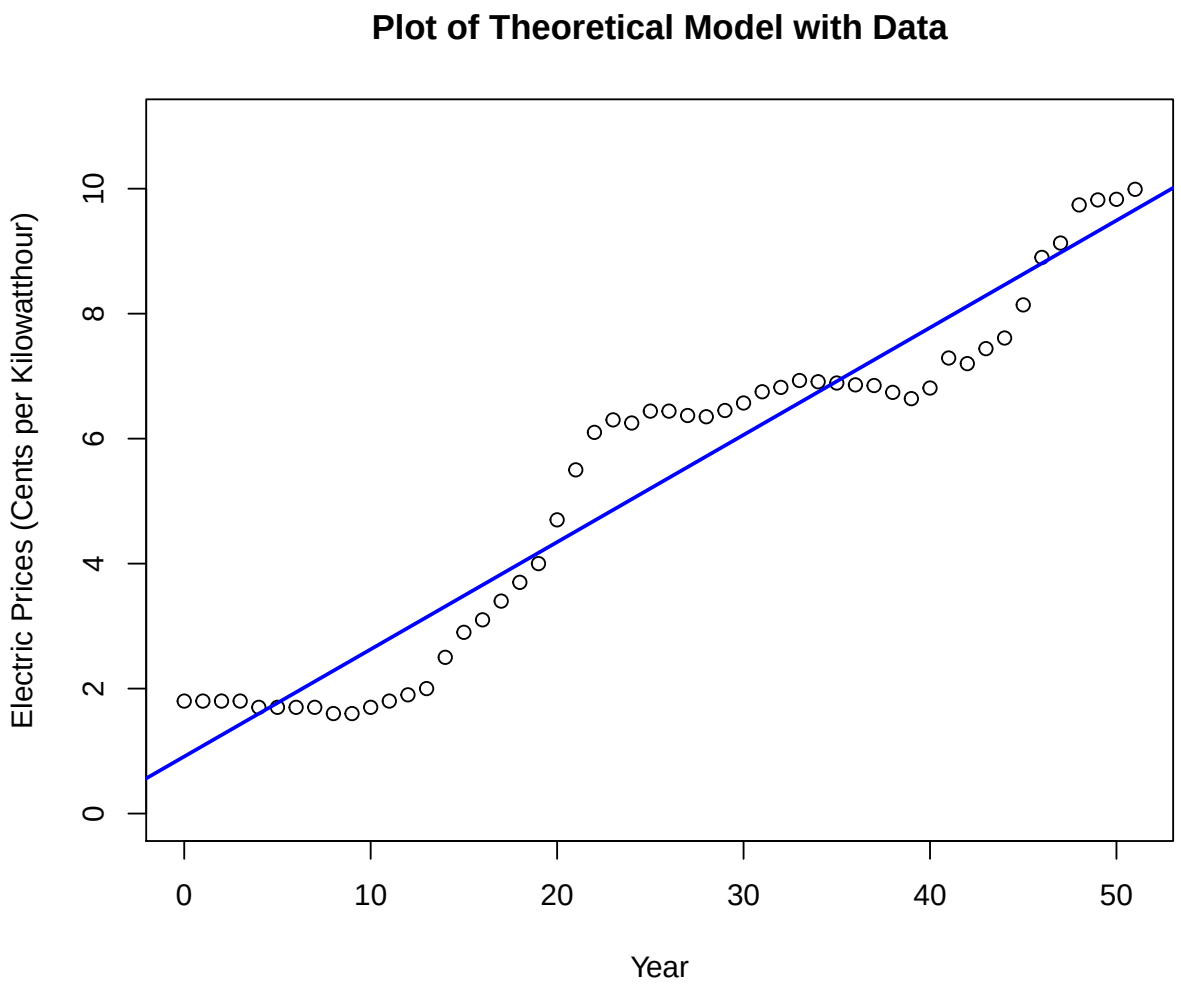
Electric Price Linear Model

We can now use

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

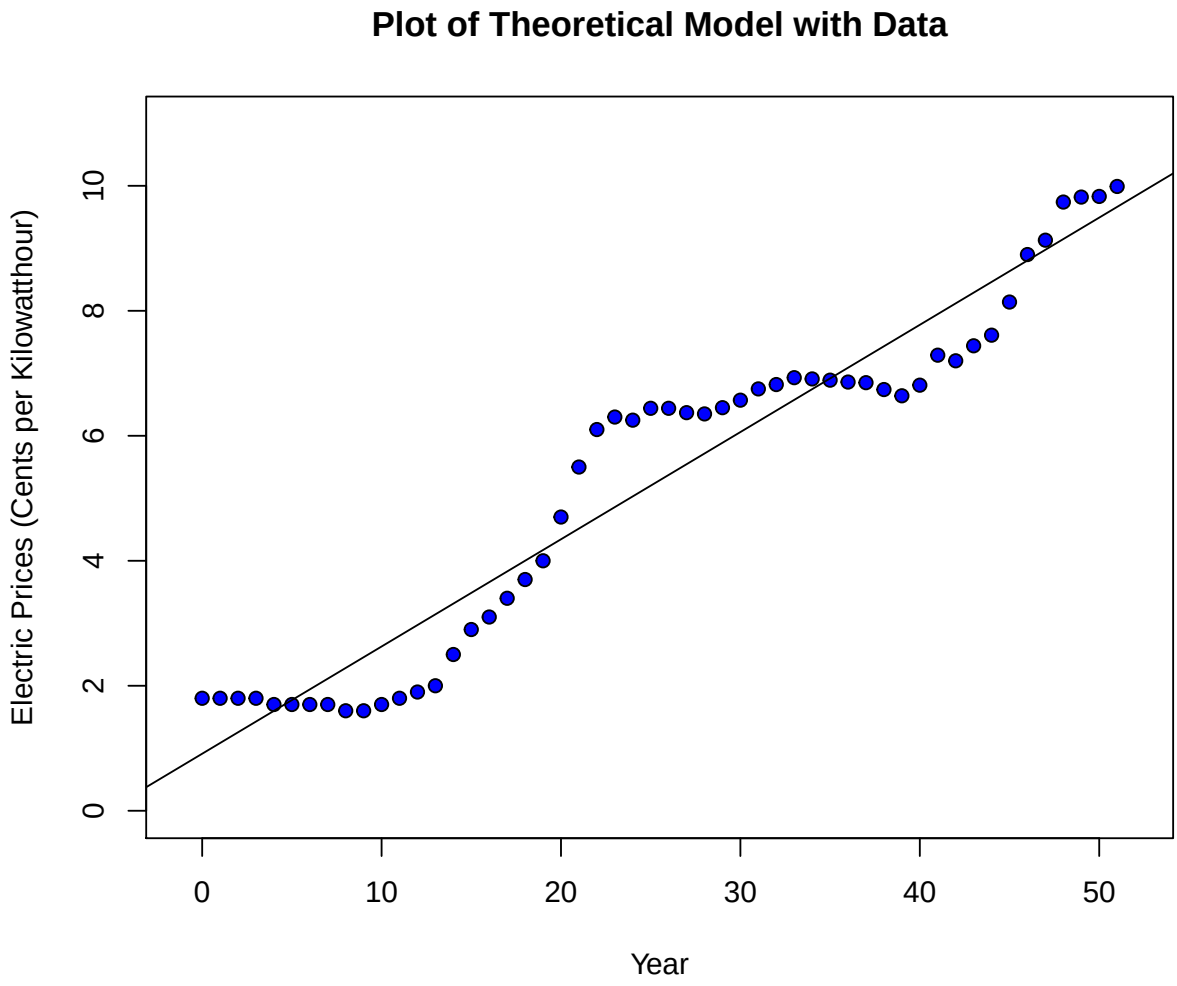
as a predictor of each datum according to each independant variable. The line can be applied to the actual data to show the linear model.

```
plot(ep.df$TOT~x, ylim = c(0, max(ep.df$TOT) + 1), xlim = c(0, max(x)), ylab = "Electric Prices (Cents p
abline(ep.summary$b0hat, ep.summary$b1hat, lwd = 2, col = "Blue")
```



Using the linear model we produced with code, we can check the validity of our theoretical linear model.

```
with(ep.df, {
  plot(TOT~x,bg="Blue",pch=21,ylim=c(0,1.1*max(TOT)),xlim=c(min(x)-1,max(x)+1), ylab = "Electric Prices
  abline(ep.lm)
})
```

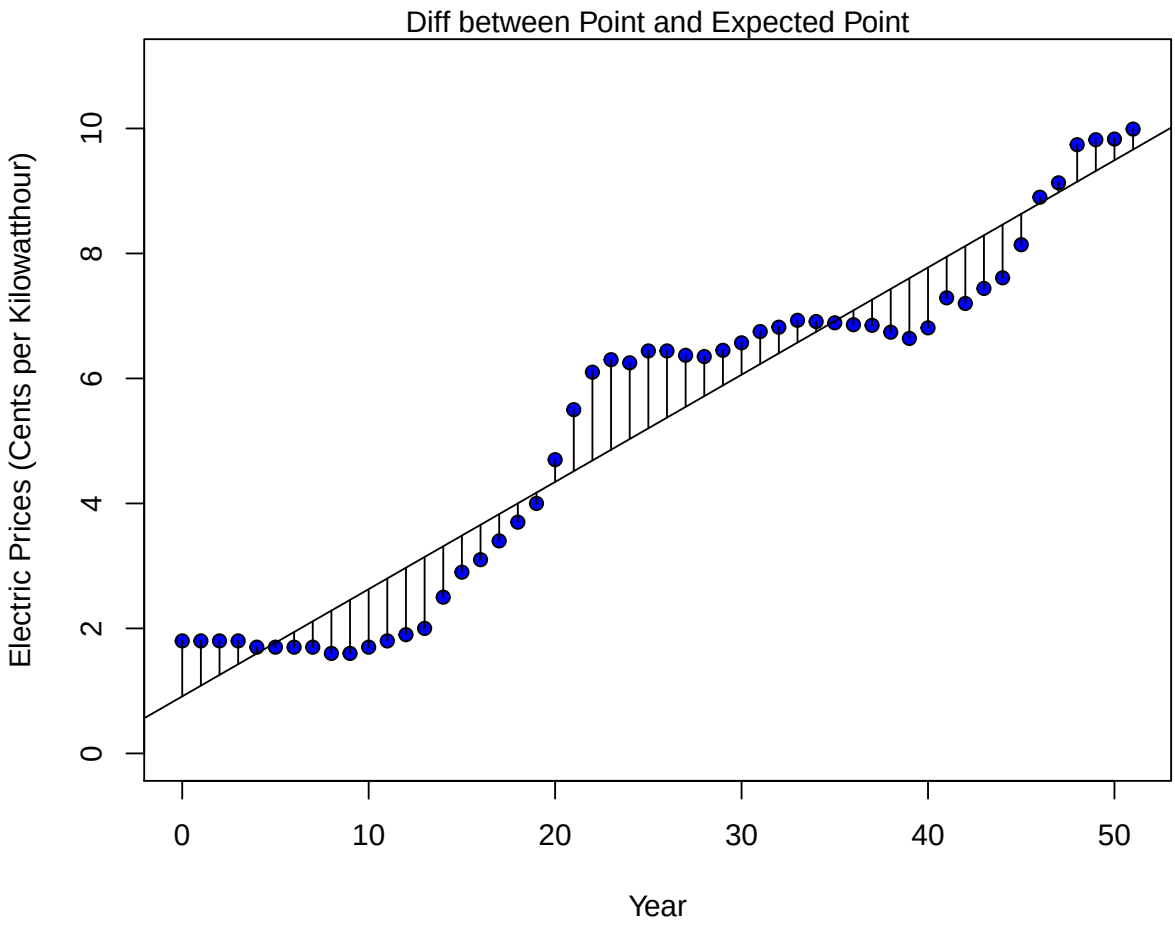


According to the linear models, we have a perfect match, so our theory is valid.

Error Plot from Linear Model

With the data and the linear model, we can show how much each datum varies from the theoretical linear model we have produced.

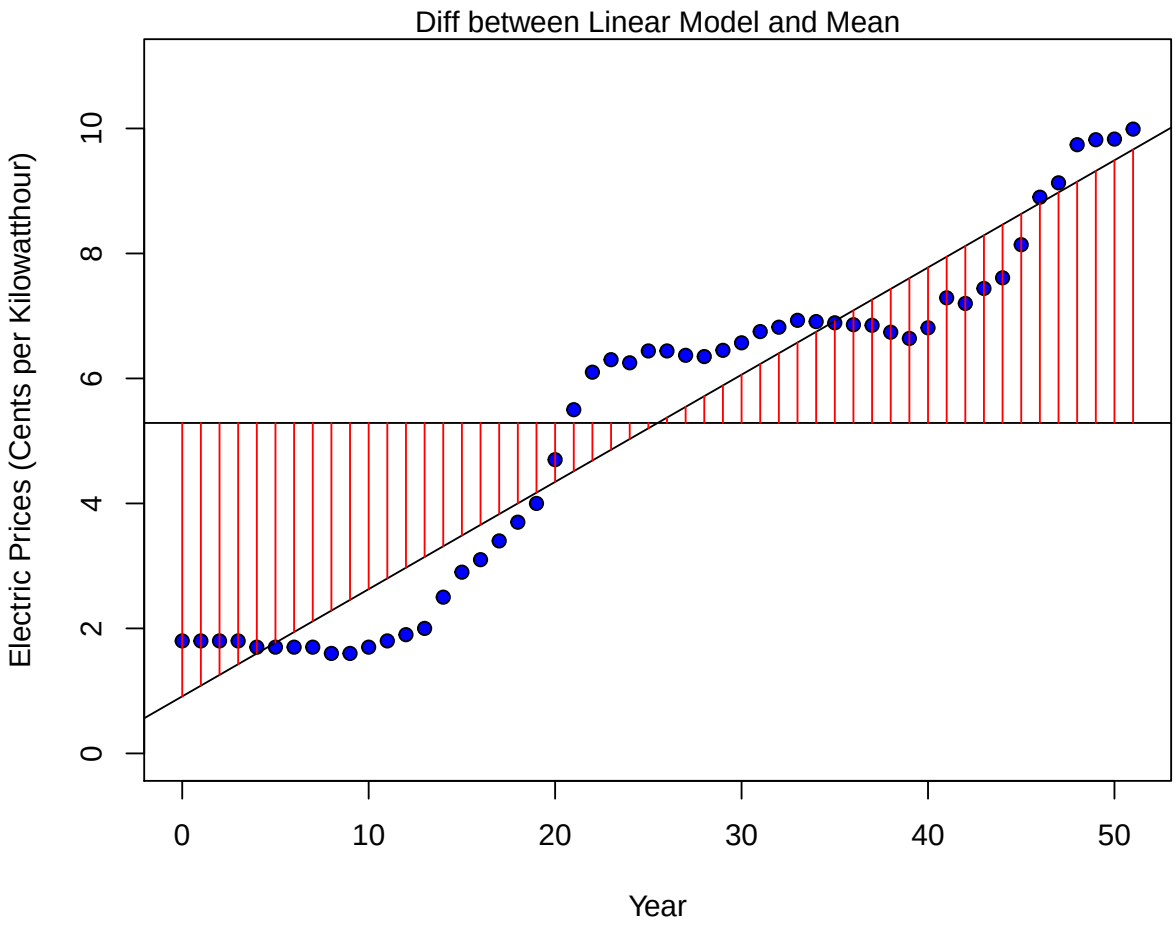
```
with(ep.df, {
  plot(TOT~x,bg="Blue",pch=21,ylim=c(0,1.1*max(TOT)),xlim=c(min(x),max(x)), ylab = "Electric Prices (Ce
  segments(x,TOT,x,ep.summary$yhat)
  abline(ep.lm)
  mtext("Diff between Point and Expected Point")
})
```



Plot of the Difference between our Theoretical Model and the Mean of the Electric Prices

We can also show the difference between the mean and the theoretical linear model.

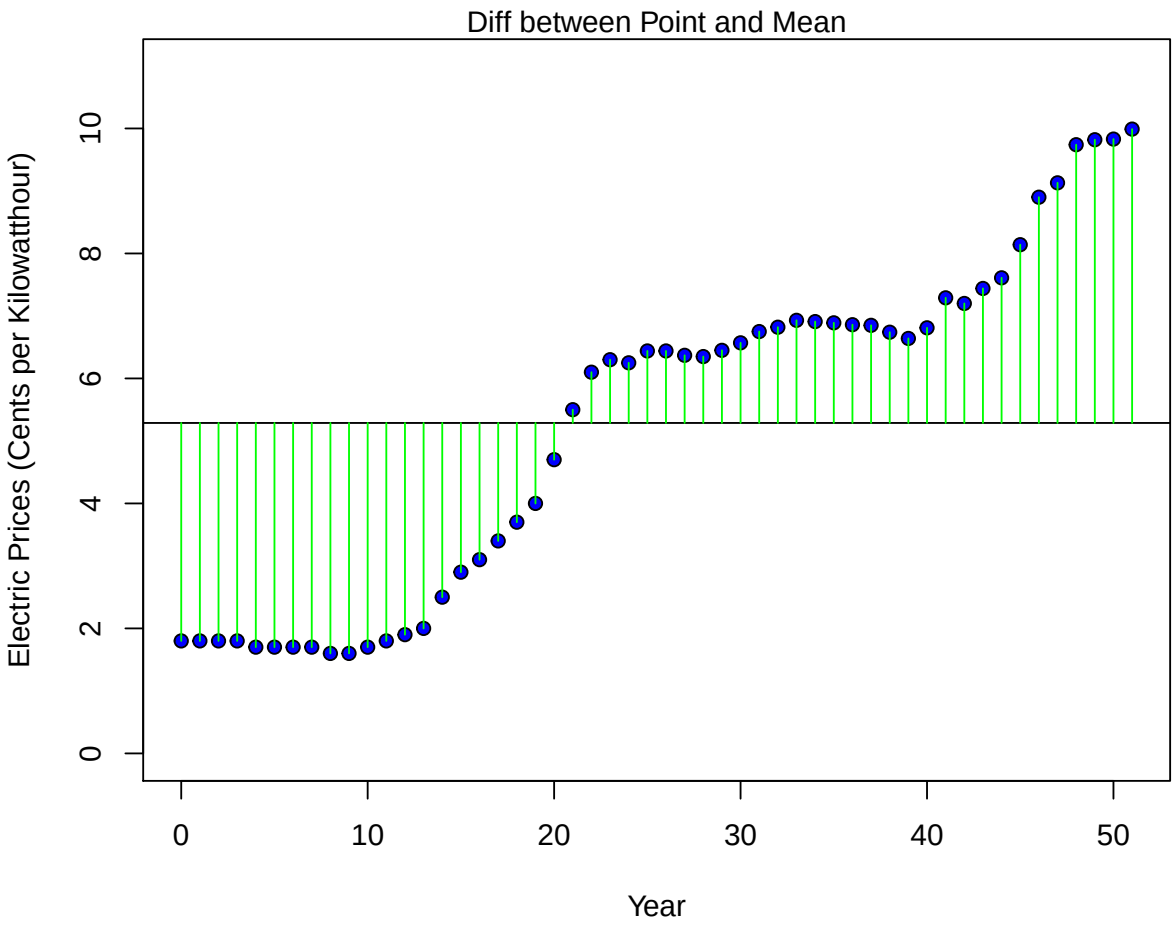
```
with(ep.df, {
  plot(TOT~x,bg="Blue",pch=21,ylim=c(0,1.1*max(TOT)),xlim=c(min(x),max(x)), ylab = "Electric Prices (Ce
  abline(ep.lm)
  abline(h=mean(TOT))
  segments(x,mean(TOT),x,ep.summary$yhat,col="Red")
  mtext("Diff between Linear Model and Mean")
})
```



Plot of the Difference between each Datum and the Mean of Electric Prices

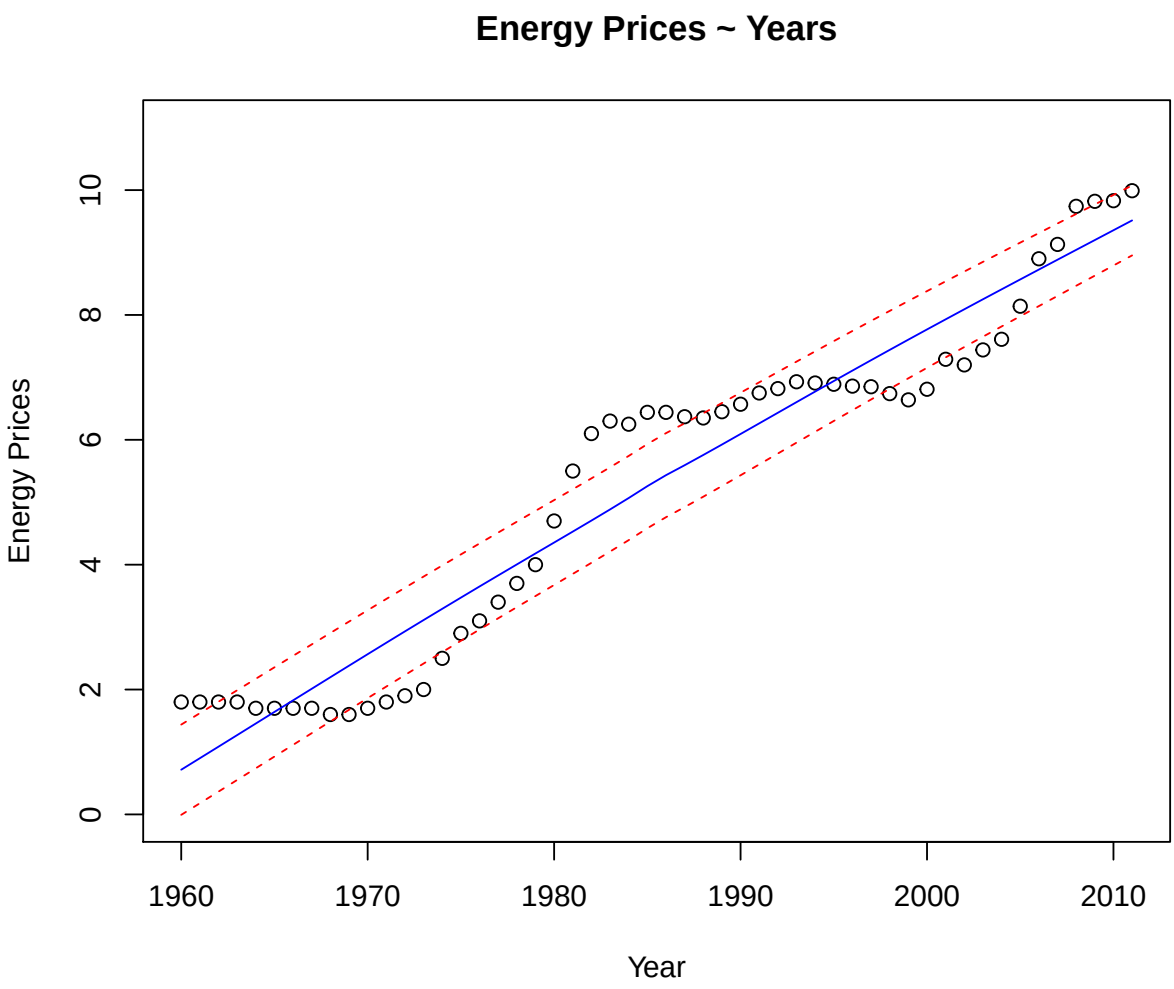
With the mean, we can show the difference between each datum and the mean about the electric prices.

```
with(ep.df, {
  plot(TOT~x,bg="Blue",pch=21,ylim=c(0,1.1*max(TOT)),xlim=c(min(x),max(x)), ylab = "Electric Prices (Ce
  abline(h=mean(TOT))
  segments(x,TOT,x,mean(TOT),col="Green")
  mtext("Diff between Point and Mean")
})
```



Trendscatter of Electric Prices

```
library(s20x)
trendscatter(ep.df$TOT~ep.df$YEAR, f = 1, ylim = c(0,11), main = "Energy Prices ~ Years", xlab = "Year"
```



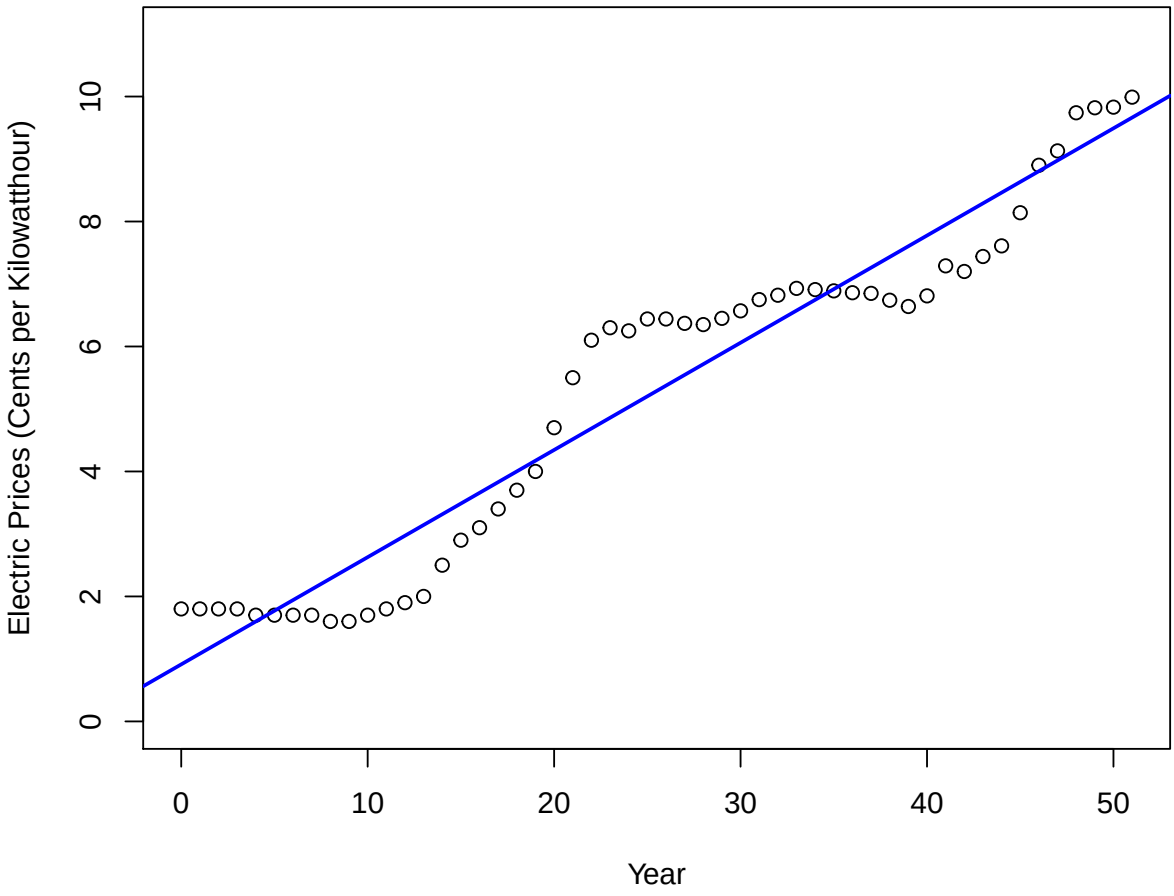
Using a trend scatter plot, we can show the standard deviation of error on each side of the linear model with shows the most outlying points of time when electric prices increased or decreased dramatically. This gains momentum to our best guesses on which are the best times to look into history.

Assumption Check of SLR

Linearity of Data

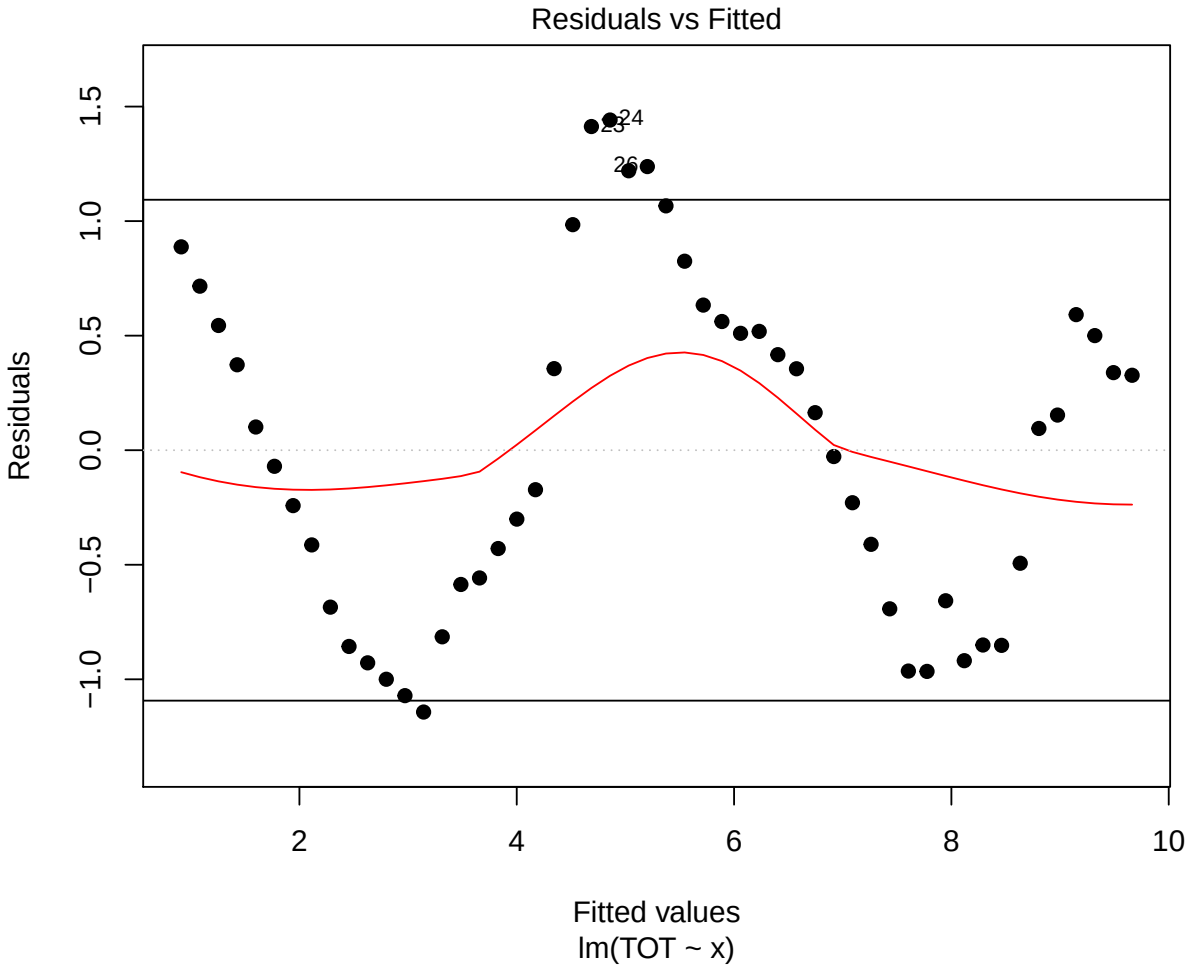
Plotted Data with Linear Model

```
plot(ep.df$TOT~x, ylim = c(0, max(ep.df$TOT) + 1), xlim = c(0, max(x)), ylab = "Electric Prices (Cents p
abline(ep.summary$b0hat, ep.summary$b1hat, lwd = 2, col = "Blue")
```



Residual vs Fitted Values

```
plot(ep.lm, which =1, pch = 19)
abline(mean(ep.lm$residuals)+1.5*sd(ep.lm$residuals),0)
abline(mean(ep.lm$residuals)-1.5*sd(ep.lm$residuals),0)
```



From the two graphs above, the plotted data and the plotted residuals~fittedvalues, we can see linearity in the data.

Data is Multivariate Normal

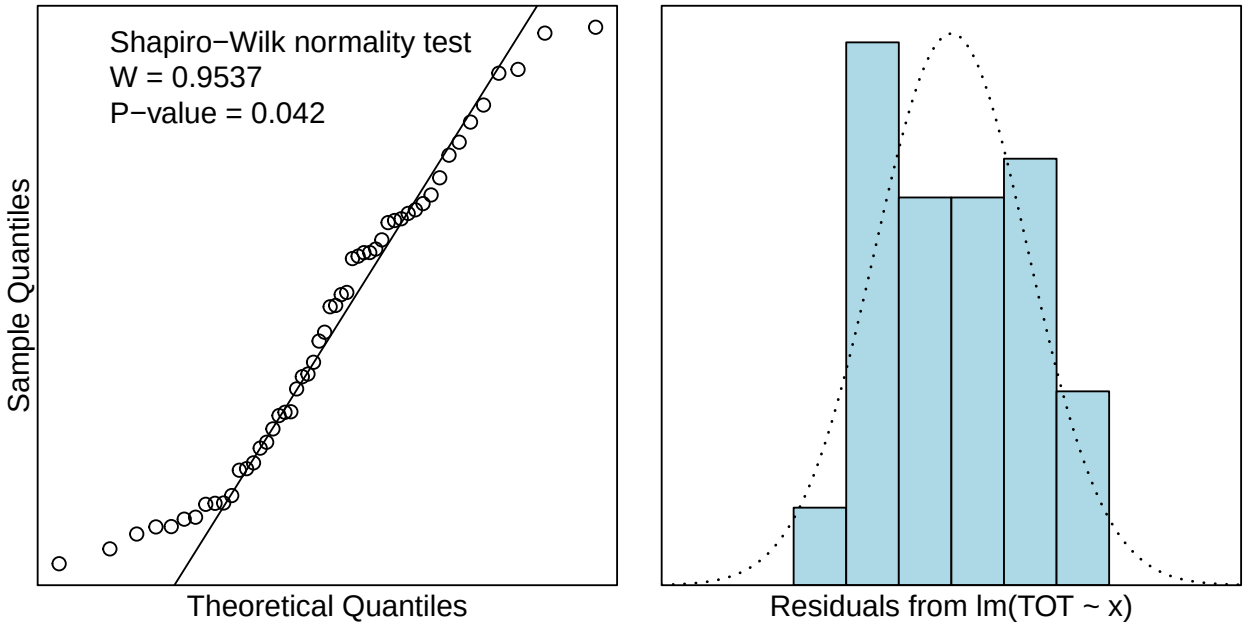
For a linear model to be a linear regression, we must have a normal distribution with the residuals.

$$\epsilon_i \sim N(0, \sigma^2)$$

Shapiro-wilk

We can check normality of the residuals by using a Shapiro Wilk graph of the linear model.

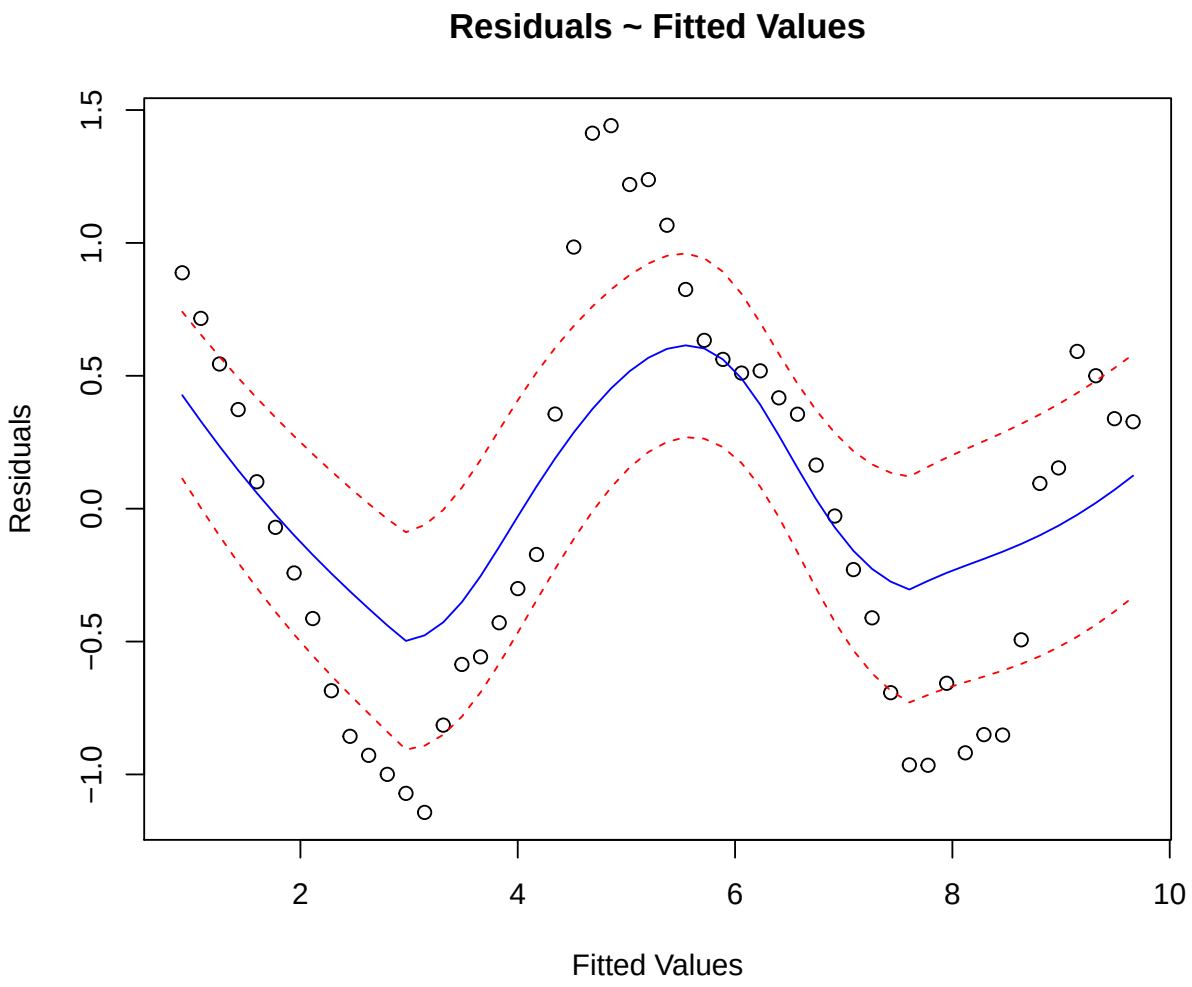
```
library(s20x)
normcheck(ep.lm, shapiro.wilk = TRUE)
```



As you can see, the residuals are distributed somewhat normally throughout the past electric prices, but we do not have sufficient evidence to reject the evidence since the data is larger than >30

trendscatter on Residual Vs Fitted

```
trendscatter(ep.lm$residuals~ep.lm$fitted.values, f = 0.5, main = "Residuals ~ Fitted Values", ylab = "
```



The data's residuals correspond with the fitted values and follow a general pattern, so constant variance cannot be claimed with the model.

Zero mean value of ϵ

```
mean(ep.summary$yerr)
```

```
## [1] -3.843465e-16
```

This shows the mean of the residuals are approximately zero i.e. $\bar{\epsilon} \approx 0$. Since $\bar{\epsilon}$ is approximately zero, we do not have evidence to reject the mean of error being zero.

```
t.test(ep.summary$yerr, mu=0, conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  ep.summary$yerr
## t = -3.8029e-15, df = 51, p-value = 1
## alternative hypothesis: true mean is not equal to 0
```

```
## 95 percent confidence interval:
##  -0.2028984  0.2028984
## sample estimates:
##     mean of x
## -3.843465e-16
```

This can also be shown by the pvalue of $\epsilon_i > 0.05$.

Independence of data

We can check independence of data through Tolerance or $T = 1 - R^2$

```
1-.9272^2
```

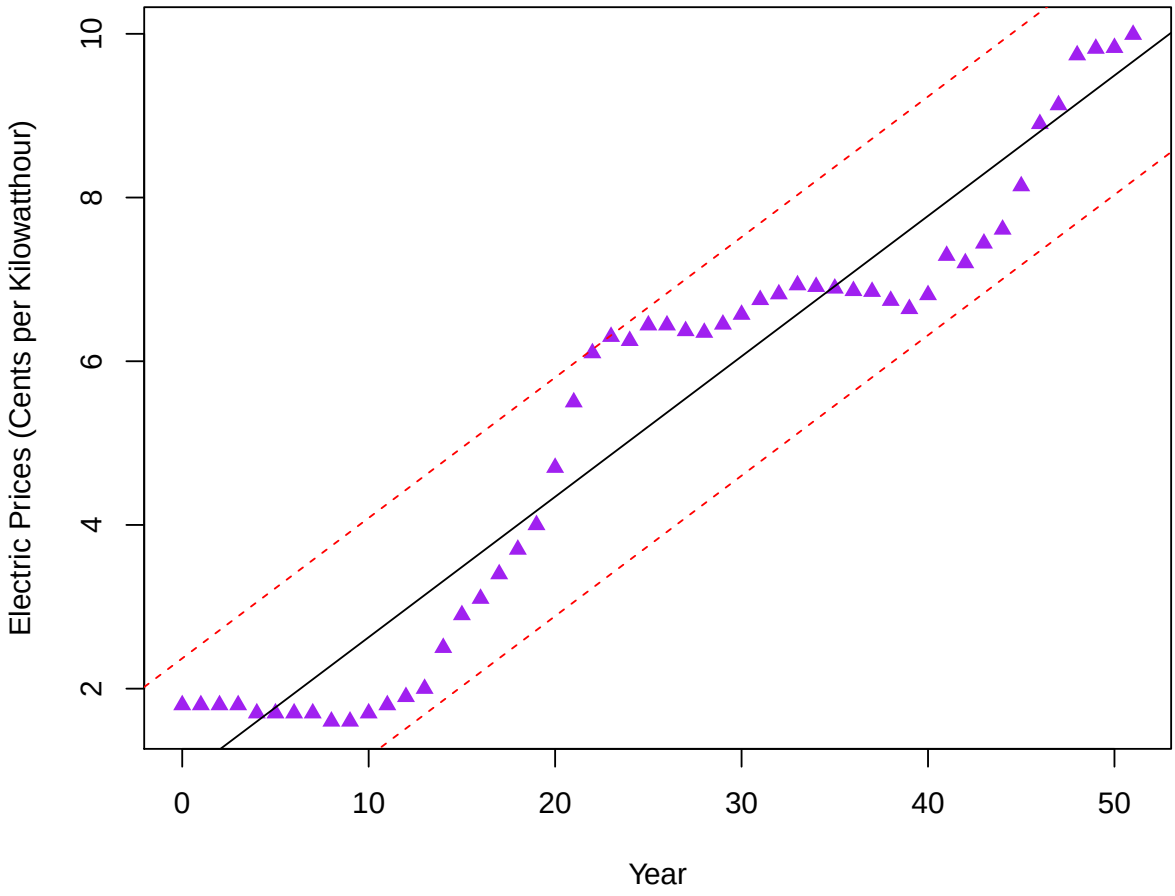
```
## [1] 0.1403002
```

Since tolerance is >.1, we cannot reject that the data is Independent.

Homoscedasticity

We can check if the data is identically distributed through the plotting of the data with lines of same slope showing variation

```
plot(ep.df$TOT~x, pch = 17, col = "Purple", ylab = "Electric Prices (Cents per Kilowatthour)", xlab = "
abline(ep.summary$b0hat, ep.summary$b1hat)
abline(ep.summary$b0hat+2*sd(ep.lm$residuals), ep.summary$b1hat, h = 0, lty = 2, col = "Red")
abline(ep.summary$b0hat-2*sd(ep.lm$residuals), ep.summary$b1hat, h = 0, lty = 2, col = "Red")
```



As we can see, the residuals of the Electric Prices lie within 2 standard deviations from the linear model which now, we can accept that the data is identically distributed.

Analysis Summary

Summary lm object

```
summary(ep.lm)
```

```
##
## Call:
## lm(formula = TOT ~ x, data = ep.df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14295 -0.66414  0.03362  0.52507  1.44126
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.912417   0.201234   4.534 3.63e-05 ***
## x           0.171579   0.006801  25.229  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.736 on 50 degrees of freedom
## Multiple R-squared:  0.9272, Adjusted R-squared:  0.9257
## F-statistic: 636.5 on 1 and 50 DF,  p-value: < 2.2e-16
```

With the summary of the linear model placed on Electric Prices, we can state that the linear model follows $y_i = 0.9124 + 0.1716x_i$ for prediction of any Electric Price given a year. Also, we can see that the residual standard error is 0.736, which shows the data fits the line rather well, but is not perfect. Multiple R Squared also shows a good data fit to the linear model. As multiple R squared tends to 100%, the data is near perfect. With Adjusted R Squared, we can see R squared adjusted to better fit the ammount of predictors we have in the linear model, but Adjusted R Squared is still higher than 90%, which shows the model fits the data very well.

Calculate cis for β parameter estimates

With analysis of the linear model summary, we can find the standard error of beta1 and beta2. With this we can find 95% confidence intervals of both betas.

```
ci=c()
t=abs(qt(0.05/2,length(x)-1))
ci[1]=ep.summary$b1hat-t*(0.006801)
ci[2]=ep.summary$b1hat+t*(0.006801)
ci
```

```
## [1] 0.1579259 0.1852330
```

```
t=abs(qt(0.05/2,length(x)-1))
ci[1]=ep.summary$b0hat-t*(0.201234)
ci[2]=ep.summary$b0hat+t*(0.201234)
ci
```

```
## [1] 0.5084224 1.3164107
```

ci's for Beta1 and Beta0

```
library(s20x)
ciReg(ep.lm)
```

```
##             95 % C.I.lower    95 % C.I.upper
## (Intercept)        0.50823           1.31661
## x                  0.15792           0.18524
```

As you can see, we can generate the 95% confidence intervals for β_1 and β_0 theorically and computer generated, and they are very close, if not spot on.

Predictions

We can now, with the linear model gain access to predictions and with this function, we can produce any approximate prediction between 1960-2011.

```
mypred = function(y, ep.lm){
  predict(ep.lm)[y - 1960]
}
mypred(1961, ep.lm)
```

```
##         1
## 0.9124165
```

```
mypred(1991, ep.lm)
```

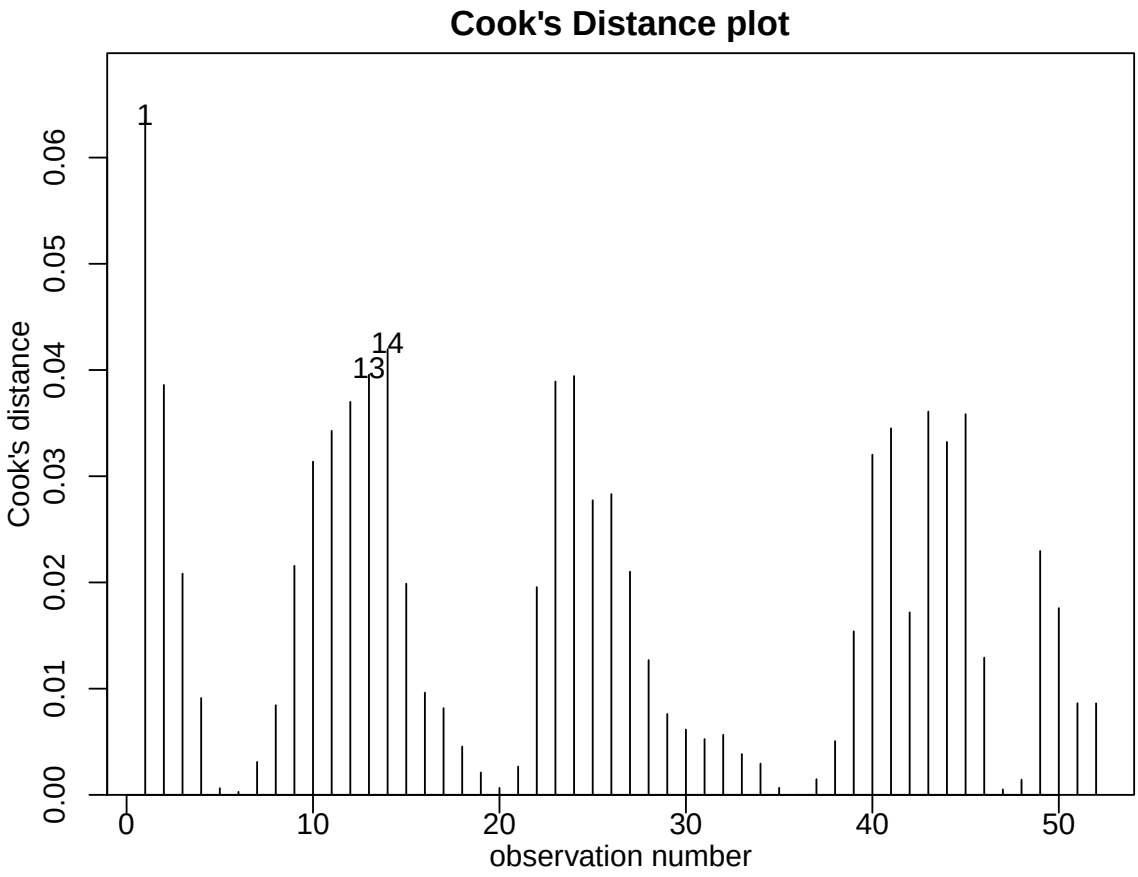
```
##     31
## 6.0598
```

```
mypred(2011, ep.lm)
```

```
##       51
## 9.491389
```

Outliers using cooks plots

```
cooks20x(ep.lm)
```

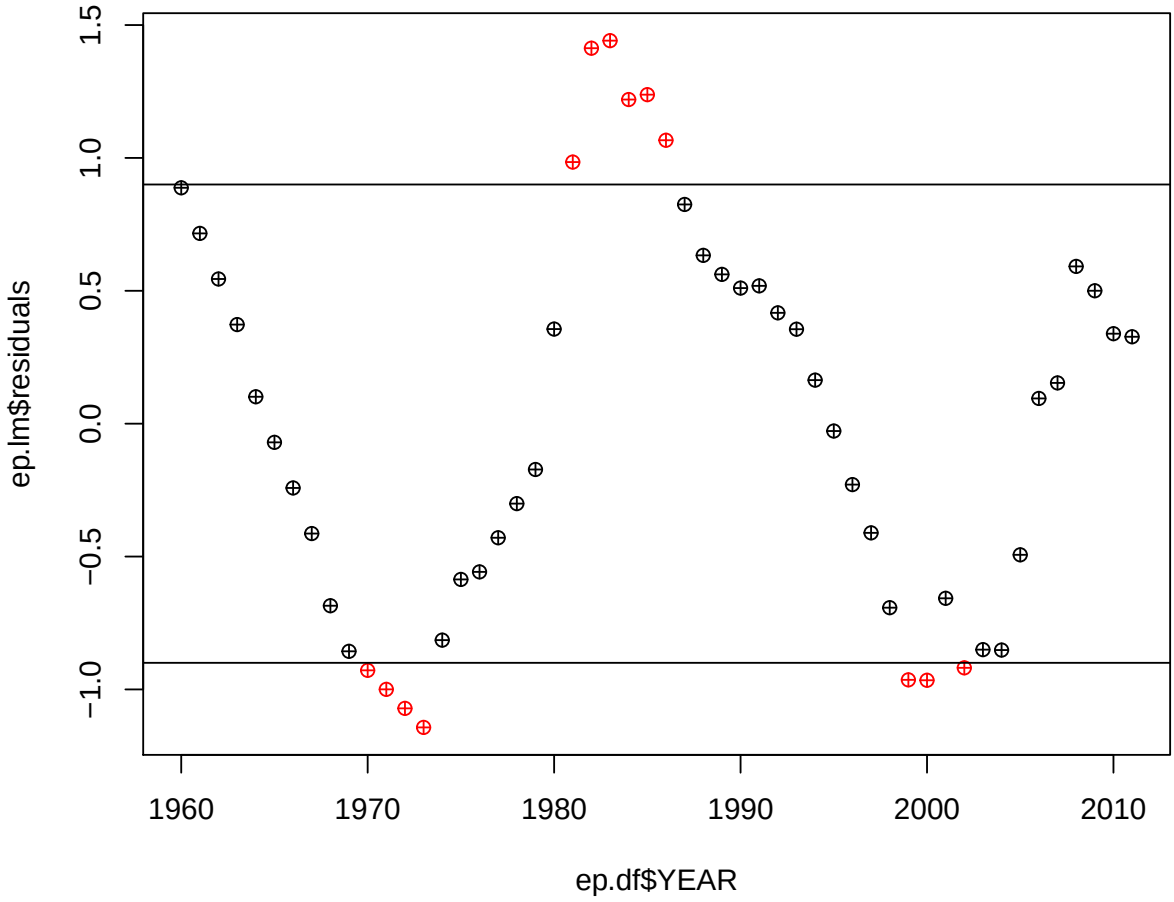


This cooks distance plot, shows each observation and the residual corresponding with the observation. The Higher the line, the more of an error we acheive. In the Electric Prices, we can see that the cooks plot cooked up 3 outliers, and if we add 1960 to each of these outliers, we have 1961, 1973, and 1974. These will be some places we will look for events that have dramatically changed electric prices in the past.

Plot of residuals

Before we come to our conclusion and the time points we want to investigate, lets tighten the trend that was loosely placed on the cooks plot. We can do this using a plot of the Y_{error} vs. Year.

```
plot(y=ep.lm$residuals,x=ep.df$YEAR, col= ifelse(ep.lm$residuals >= 0.9, "Red", ifelse(ep.lm$residuals <
abline(0.9,0)
abline(-0.9,0)
```



Looking at the graph, we can definitely see 3 data groups of outliers that we can use as a timeline for history. Lets gather these data.

```
outliers = c()
ctr = 1
for(i in 1:length(ep.lm$residuals)){
  if(ep.lm$residuals[i] >= 0.9){
    outliers[ctr] = ep.df$YEAR[i]
    ctr = ctr + 1
  }
  if(ep.lm$residuals[i] <= -0.9){
    outliers[ctr] = ep.df$YEAR[i]
    ctr = ctr + 1
  }
}
outliers
```

```
##  [1] 1970 1971 1972 1973 1981 1982 1983 1984 1985 1986 1999 2000 2002
```

Now we absolutely have the years for suspicion.

Conclusion

Looking at the data group we found above, we have found historical events that correspond with influx of Electricity Prices barring inflation.

1970's

"The Oil Crisis (October 1973 - March 1974) caused by oil embargo of Organization of Arab Petroleum Countries has significant impact on oil consumption and efficiency of the electric plants. The crisis showed great dependence of Unites States and other western countries on oil price and led to greater interest interest in renewable energy and induced research in solar and wind power. Petroleum consumption for electricity at 1970 was roughly 15%." [3] The Oil Crisis alone created a spike in Electricity Prices, and companies were pushed toward renewable electricity. (see Iskhakov, n.d.)

1980's

The United States however adopted the Fuel Use Act in 1980 that prohibited new electric generators from using petroleum because it is a nonrenewable resource. This led to cleaner and more efficient means of generating electricity and prices stabilized for a decade.(see Iskhakov, n.d.)

2000's

In the new millenium, the Clean Air Act of 1992 took affect and the sight of new improvements to electrity production was in full swing. Electricity Prices rose to back the corperations that generate power to fund these innovative plans. After a decade, the Electricity economy finally stabilized and prices stayed constant.(see Iskhakov, n.d.)

Result

In Conclusion, the major influxes of Electricity Prices can be explained. After visualization of the influxes, recently, it seems that prices inflate as new acts are imposed on electric companies and their attempts to work with these acts. However, the influxes from earlier in the 1900's was due to scarcity of coal resources. Any deflations in price is usually shown by electric companies finally catching the trend and stabilizing the electric sector of the economy.

Improvements

Adding another factor such as inflation of the United States dollar bill would show error accuracy better than this model.

References

"Assumptions of Linear Regression." 2018. *Statistics Solutions*. http://www.statisticssolutions.com/assumptions-of-linear-regression/.

Iskhakov, Ruslan. n.d. "History of Electricity in the United States." *History of Electricity in the United States*. http://large.stanford.edu/courses/2013/ph240/iskhakov2/.

"U.S. Energy Information Administration - Eia - Independent Statistics and Analysis." 2012. *Table 8.10*

Average Retail Prices of Electricity, 1960-2011 (Cents Per Kilowatthour, Including Taxes). http://www.eia.gov/totalenergy/data/annual/showtext.php?t=ptb0810.