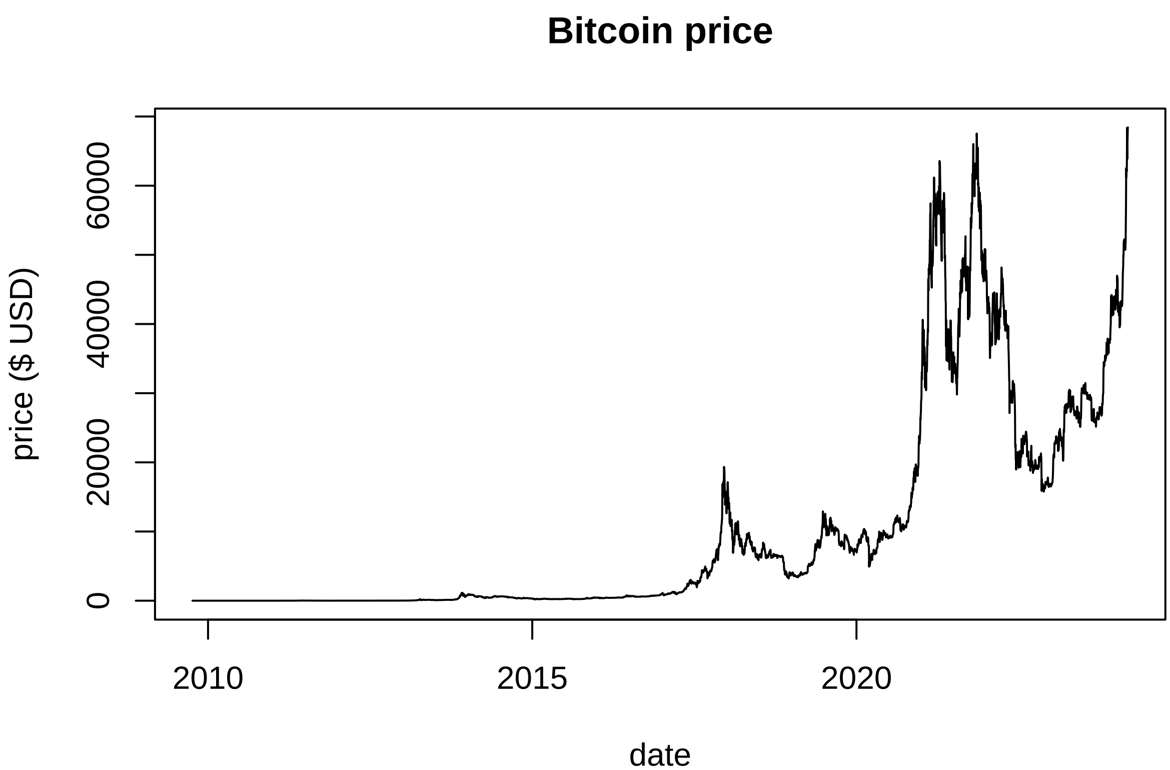
Exploratory Analysis

Alan Mathew

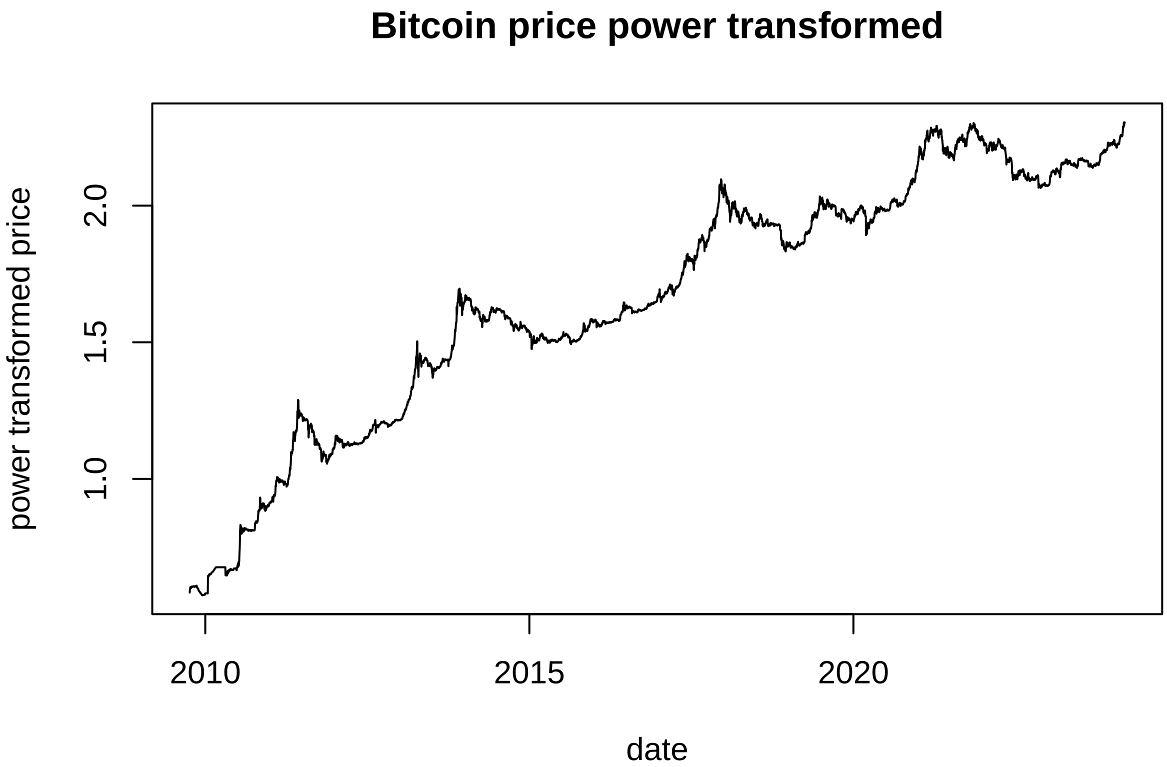
```
df = read.csv('/Users/pivaldhingra/Desktop/University courses/STAT 443 project /Data_Group24.csv')
head(df)
```

```
##         date    price
## 1 2009-10-05 0.000764
## 2 2009-10-06 0.000885
## 3 2009-10-07 0.001050
## 4 2009-10-08 0.001084
## 5 2009-10-09 0.001200
## 6 2009-10-10 0.001120
```

```
Y = ts(df$price)
date = as.Date(df$date)
plot(date, Y, main="Bitcoin price", ylab="price ($ USD)", ty='l')
```




```
plot(date, Y^0.075, main="Bitcoin price power transformed", ylab="power transformed price", ty='l')
```

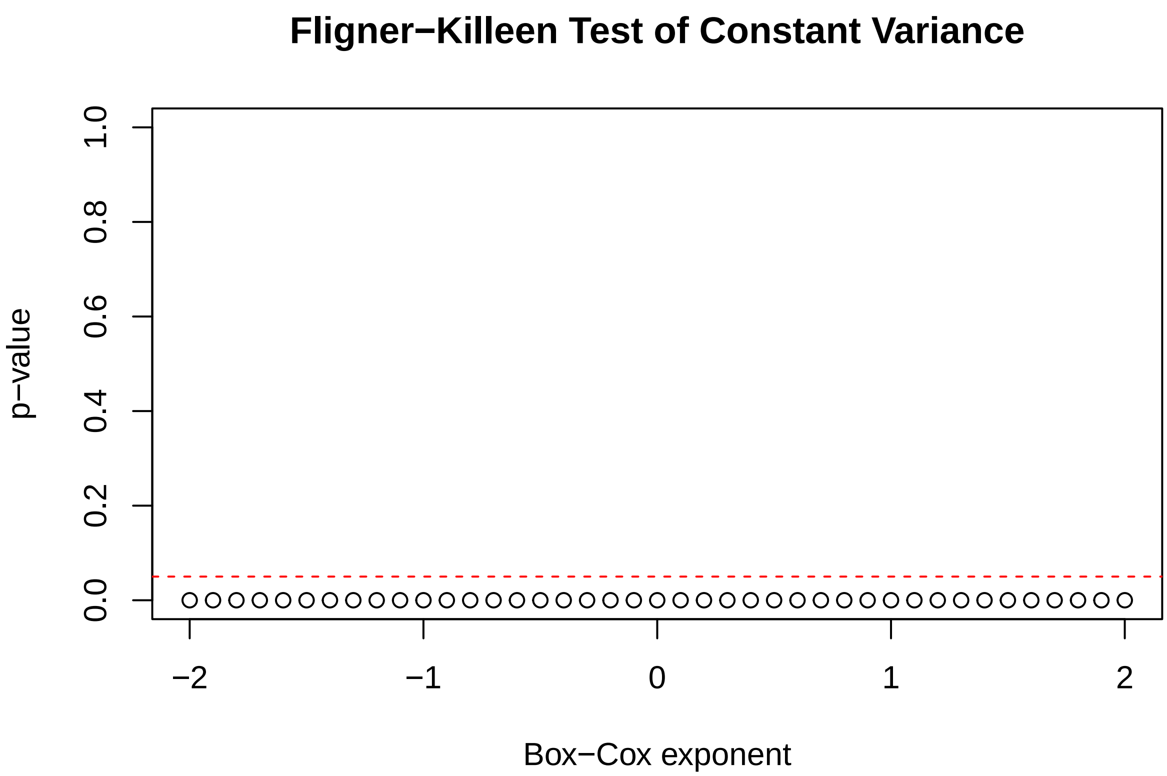


Observations:

- Trend/seasonality/non-constant variance
- Therefore, not stationary

Stabilizing Variance

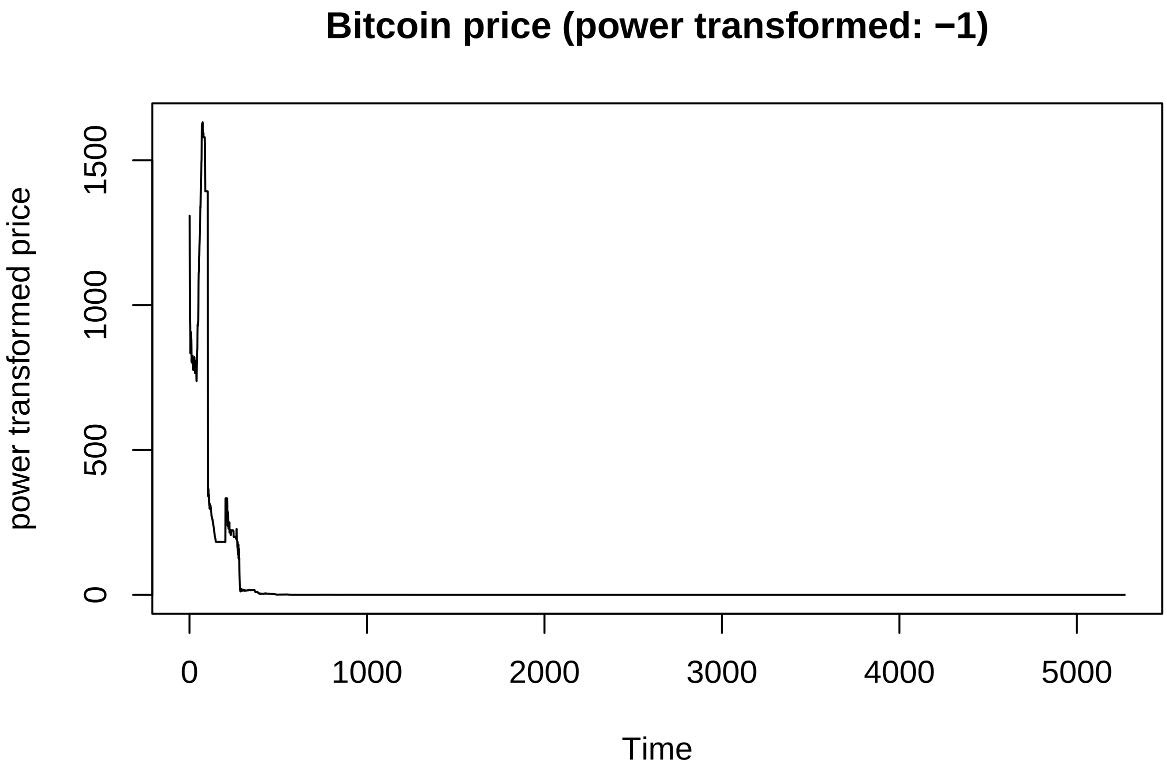
```
power.seq = seq(-2,2,by=0.1)
group = factor(rep(1:155,each=34))
filgner.p.value=c()
for(i in 1:length(power.seq)){
  if(power.seq[i]!=0){
    temp = Y^power.seq[i]
    filgner.p.value[i] = fligner.test(temp , group)$p.value
  } else {
    temp = log(Y)
    filgner.p.value[i] = fligner.test(temp , group)$p.value
  }
}
plot(power.seq, filgner.p.value, main="Fligner-Killeen Test of Constant Variance",
     xlab="Box-Cox exponent", ylab="p-value", ylim=c(0, 1))
abline(h=0.05 , col="red", lty=2)
```



- No power transformation results in a significant p-value for Fligner-Killeen
- Try a number of power transformations and examine which results in the most stable variance

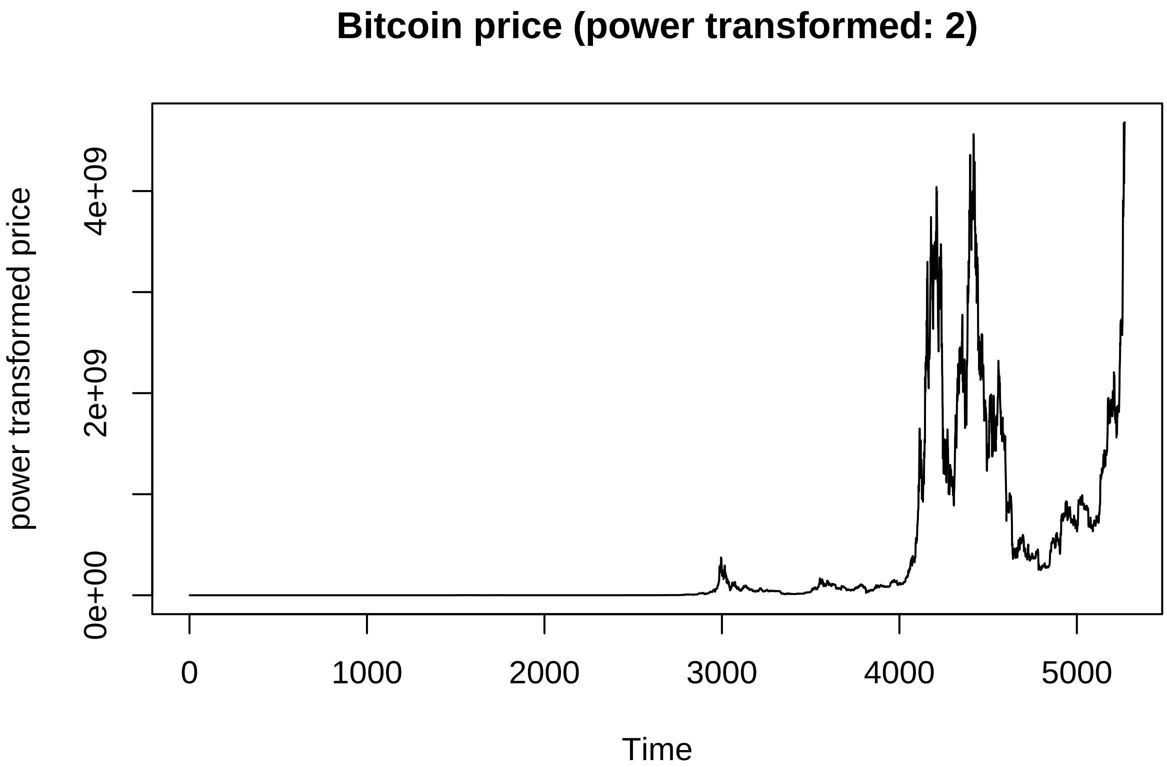
Y^-1

```
powerY = Y^-1
plot(powerY, main="Bitcoin price (power transformed: -1)", ylab="power transformed price")
```



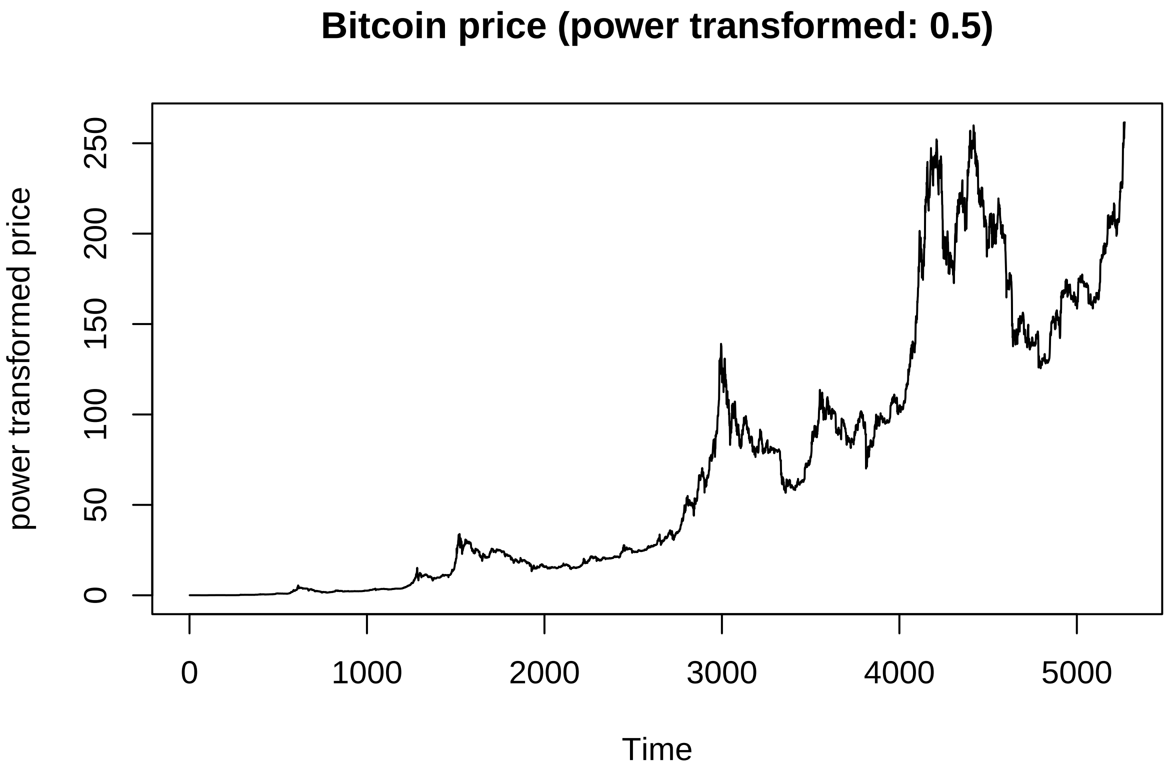
Y^2

```
powerY = Y^2
plot(powerY, main="Bitcoin price (power transformed: 2)", ylab="power transformed price")
```



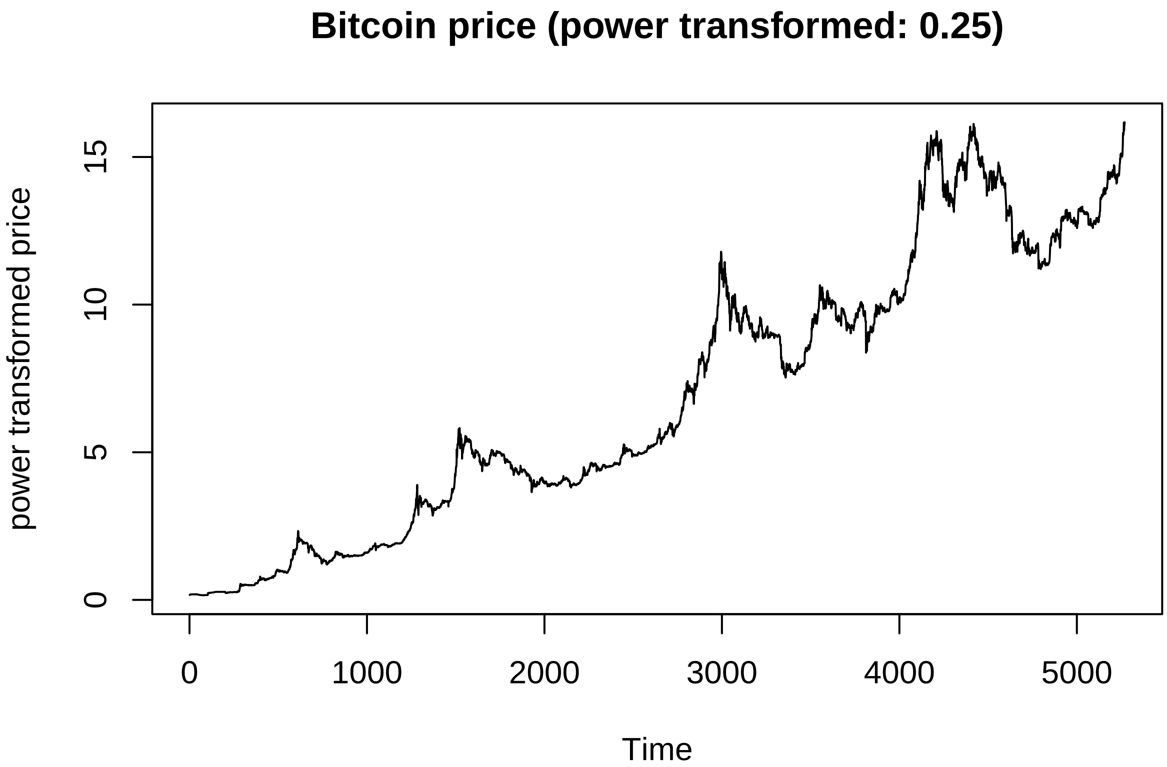
Y^0.5

```
powerY = Y^0.5
plot(powerY, main="Bitcoin price (power transformed: 0.5)", ylab="power transformed price")
```



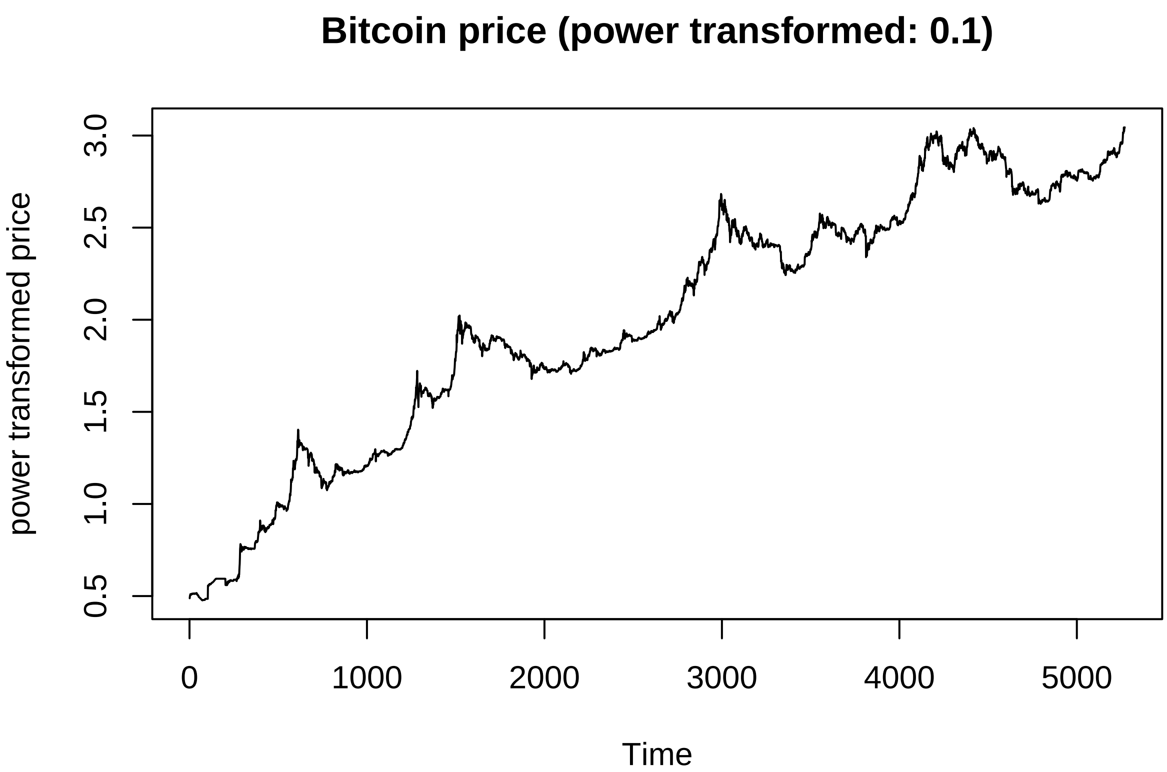
Y^0.25

```
powerY = Y^0.25
plot(powerY, main="Bitcoin price (power transformed: 0.25)", ylab="power transformed price")
```



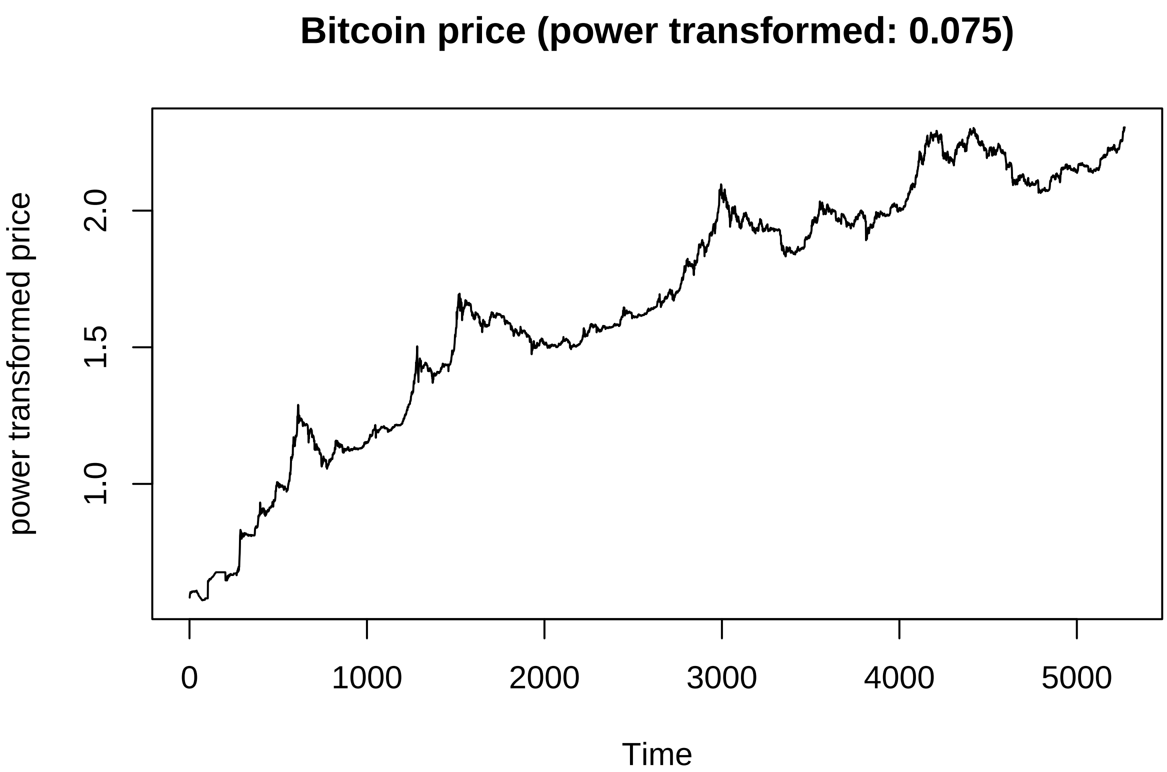
Y^0.1

```
powerY = Y^0.1
plot(powerY, main="Bitcoin price (power transformed: 0.1)", ylab="power transformed price")
```



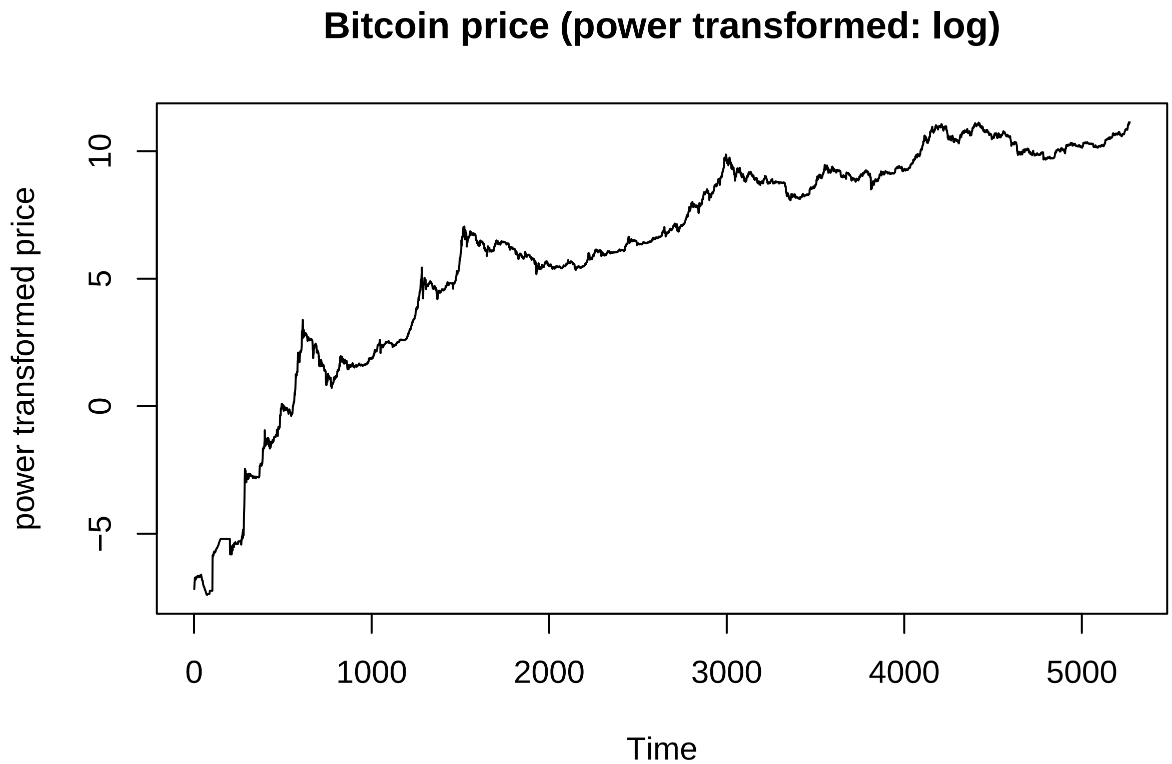
Y^0.075

```r
powerY = Y^0.075
plot(powerY, main="Bitcoin price (power transformed: 0.075)", ylab="power transformed price")
```



log (Y^0)

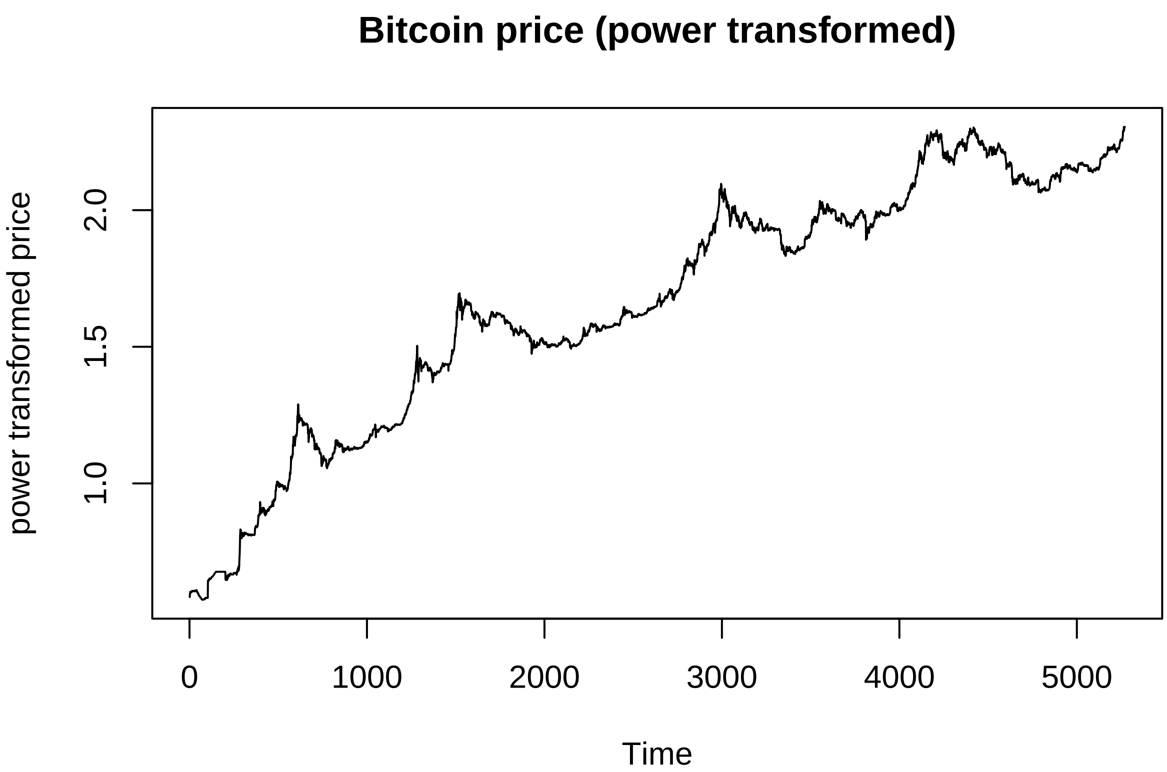
```
powerY = log(Y)
plot(powerY, main="Bitcoin price (power transformed: log)", ylab="power transformed price")
```


Bitcoin price (power transformed: log)
power transformed price
−5
0
5
10
0
1000
2000
3000
4000
5000
Time

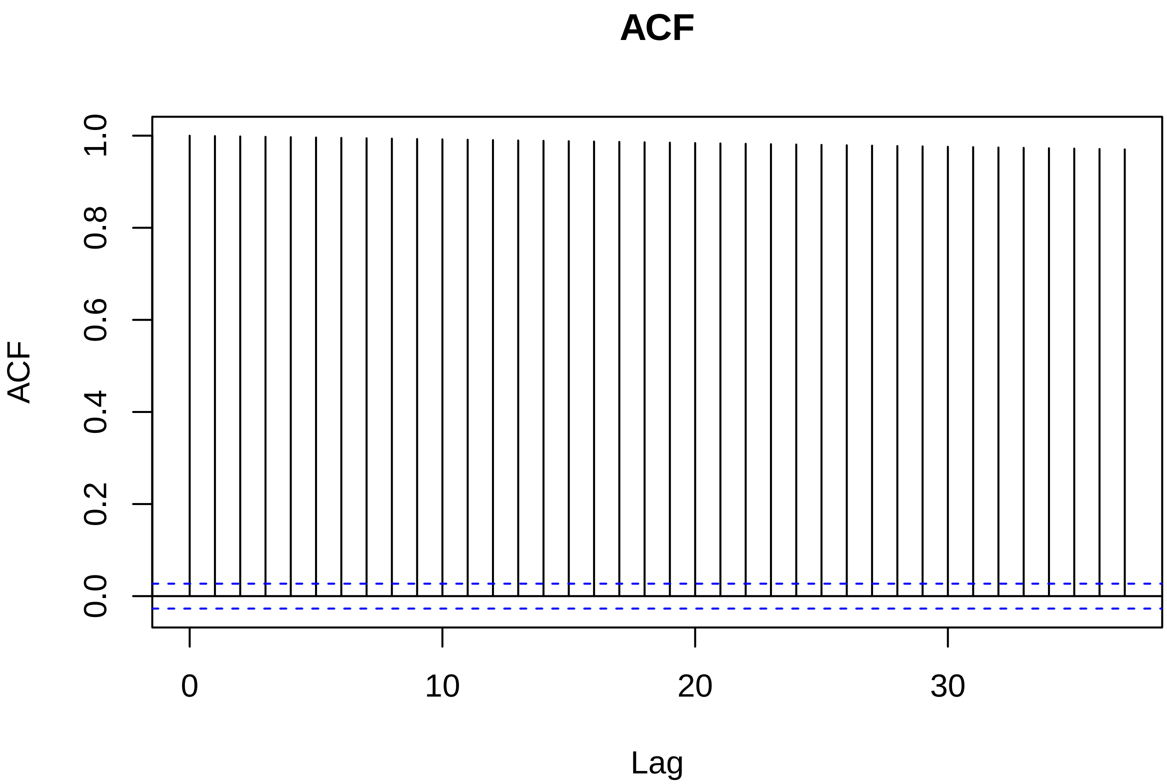
- While $Y^{0.25}$ seems to do the best job of making the trend linear, $Y^{0.075}$ creates the most stable variance

Trend & Seasonality

```
pwrY = Y^0.075
plot(pwrY, main="Bitcoin price (power transformed)", ylab="power transformed price")
```




```
acf(pwrY, main="ACF")
```

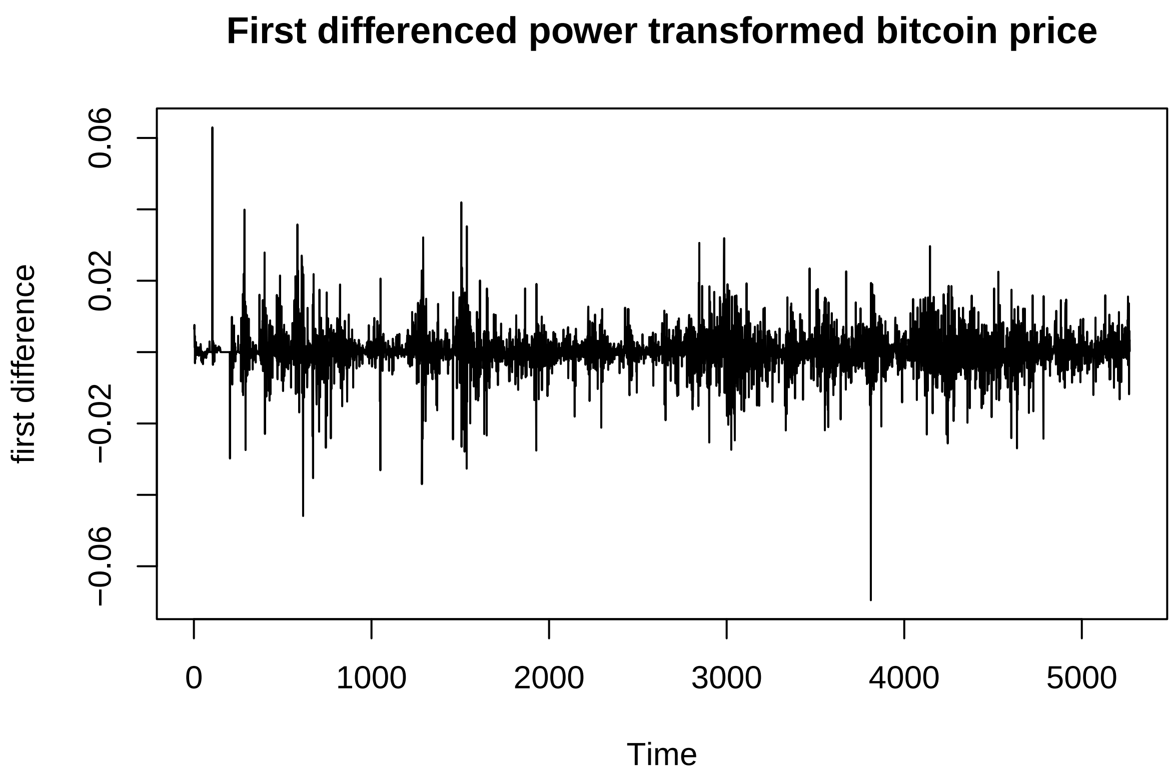


- There is a clear upward trend in the data
- Seasonality is also present, but season lengths seem to be increasing with time

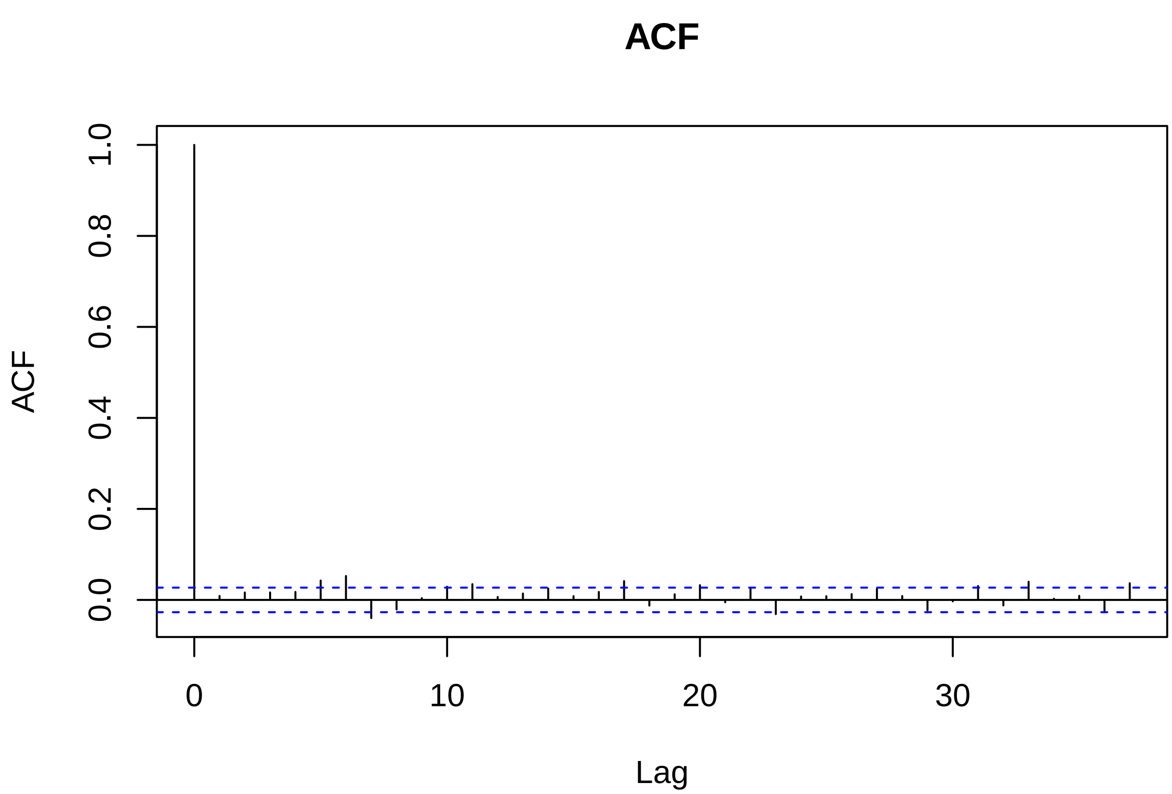
Differencing

- Differencing can be used to remove the trend

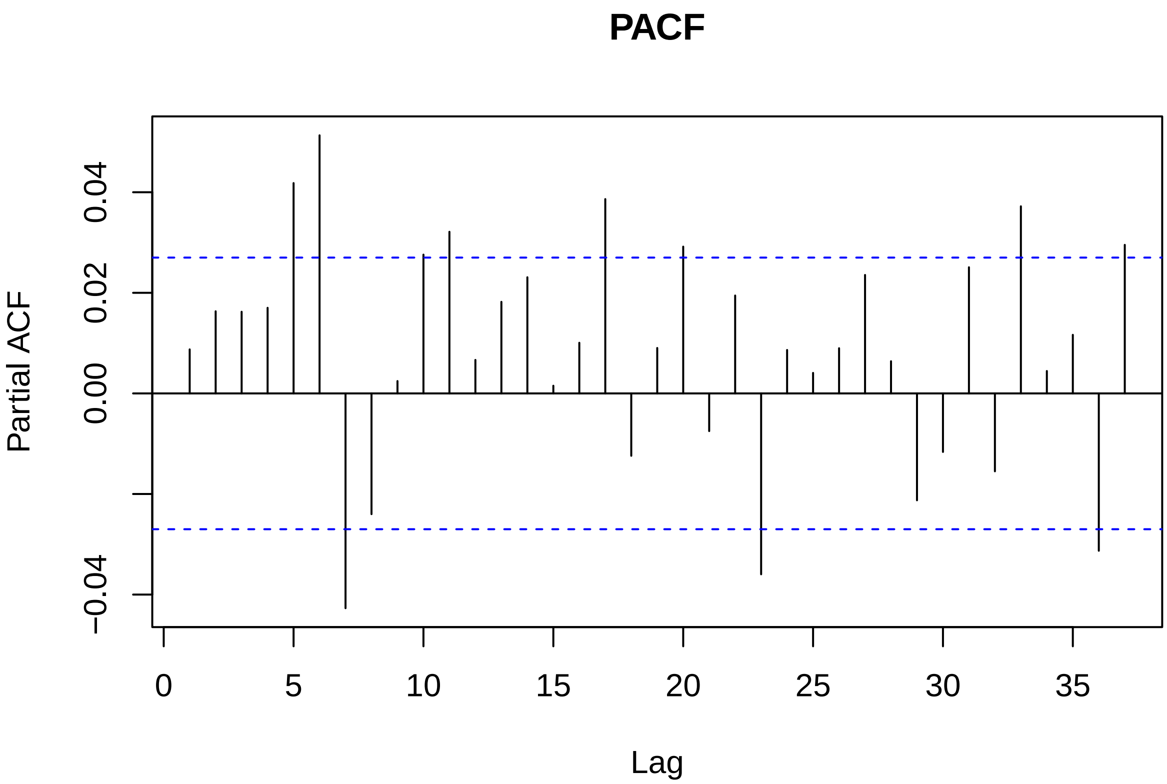
```
diffY = diff(pwrY)
plot(diffY, main="First differenced power transformed bitcoin price", ylab="first difference")
```




```r
acf(diffY, main="ACF")
```



```r
pacf(diffY, main="PACF")
```



Observations:

- Data seems to be stationary after first order differencing