SARIMA Modeling for the M1 Supply

Pstat 274 Final Project

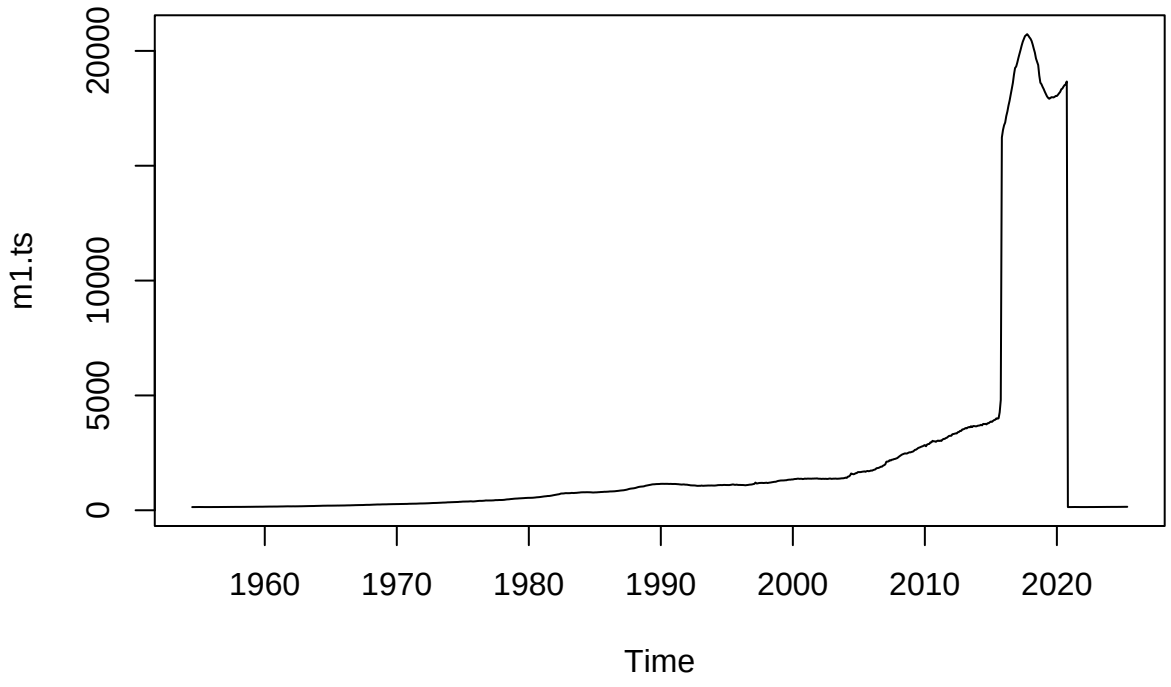
Jake Merry

06.06.2025

Setup Chuck:

```
knitr::opts_chunk$set(echo = TRUE)
library(MASS)
library(forecast)
```

```
m1 <- read.csv("data/m1Monthly.csv")
m1.ts <- ts(m1[ ,2], start=c(1954, 7), end=c(2025, 5), frequency=12)
plot.ts(m1.ts, type='l')
```



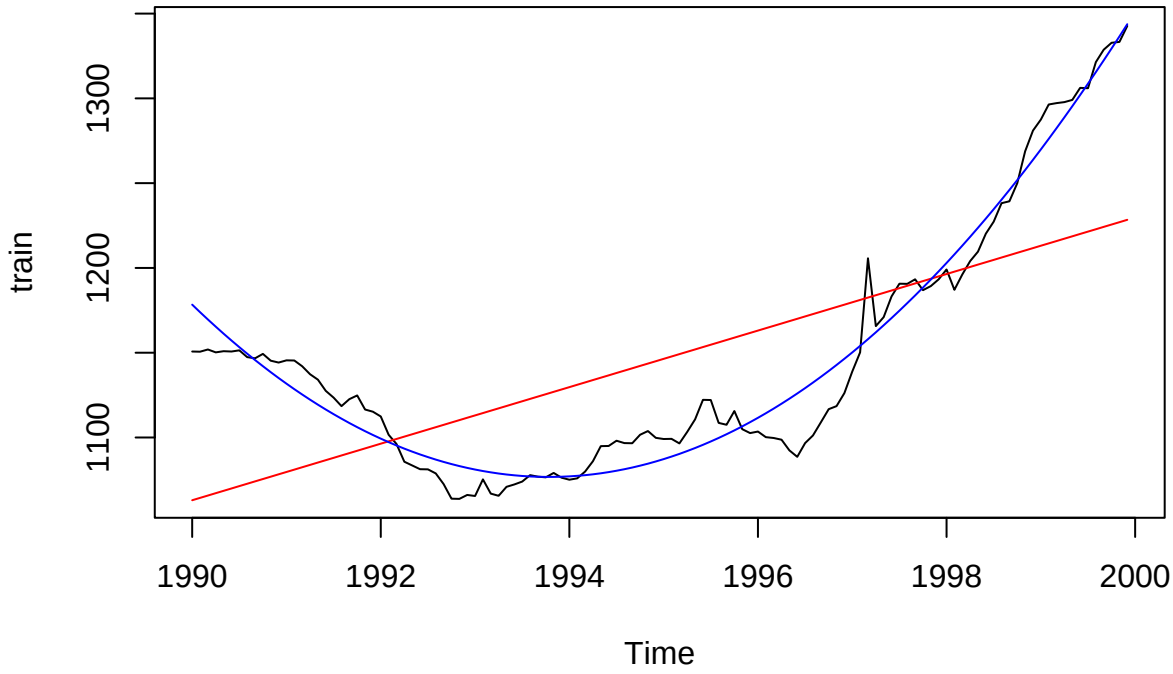
We will restrict to 1990-2000 for analysis since the data perturbs considerably.

```
train <- window(m1.ts, start=c(1990, 1), end=c(1999, 12), freq=12) # 120 obs
test <- window(m1.ts, start=c(2000, 1), end=c(2001, 12), freq=12) # 24 obs
```

Without Box-Cox Transformation

```
plot.ts(train, type='l')
lines(tslm(train ~ trend)$fitted, col="red") # Linear trend
```

```
lines(tslm(train ~ trend + I(trend^2))$fitted, col="blue") # Quadratic trend
```



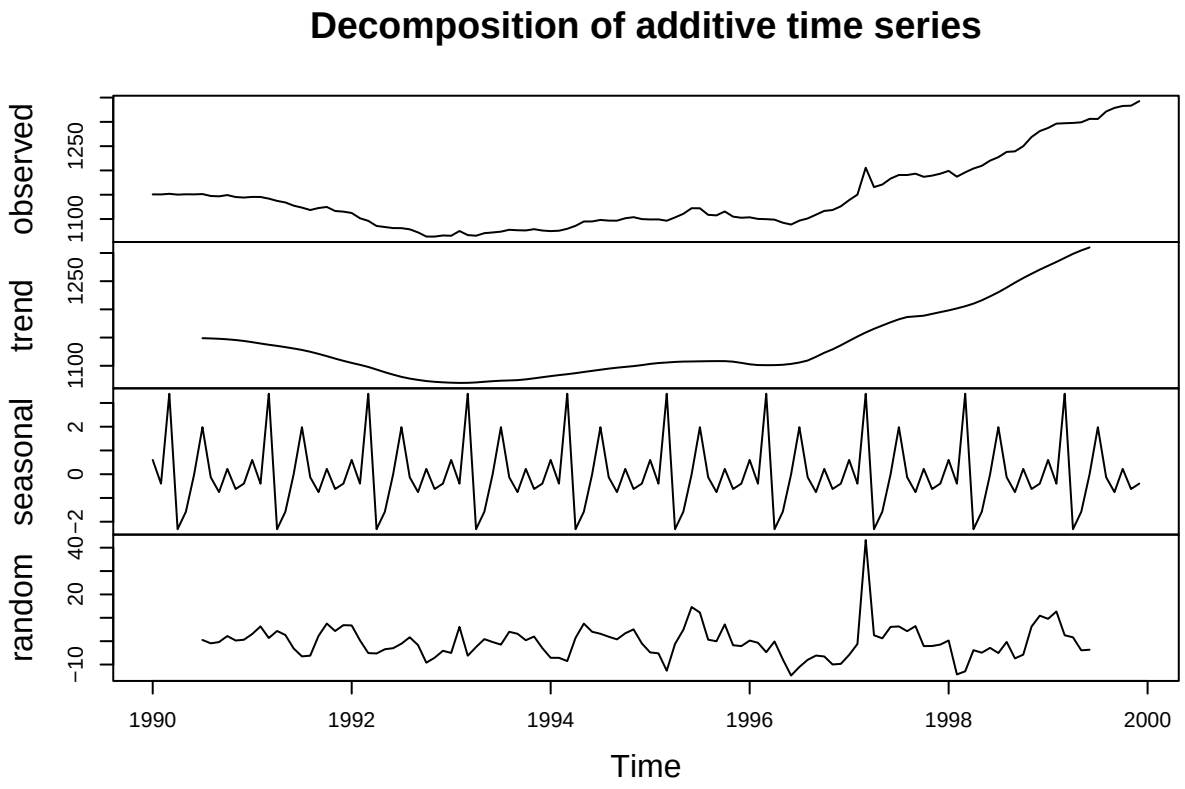

```
# Needs transformation to remove heteroskedasticity
# Difference to remove seasonality and trend
```

Clear quadratic trend.

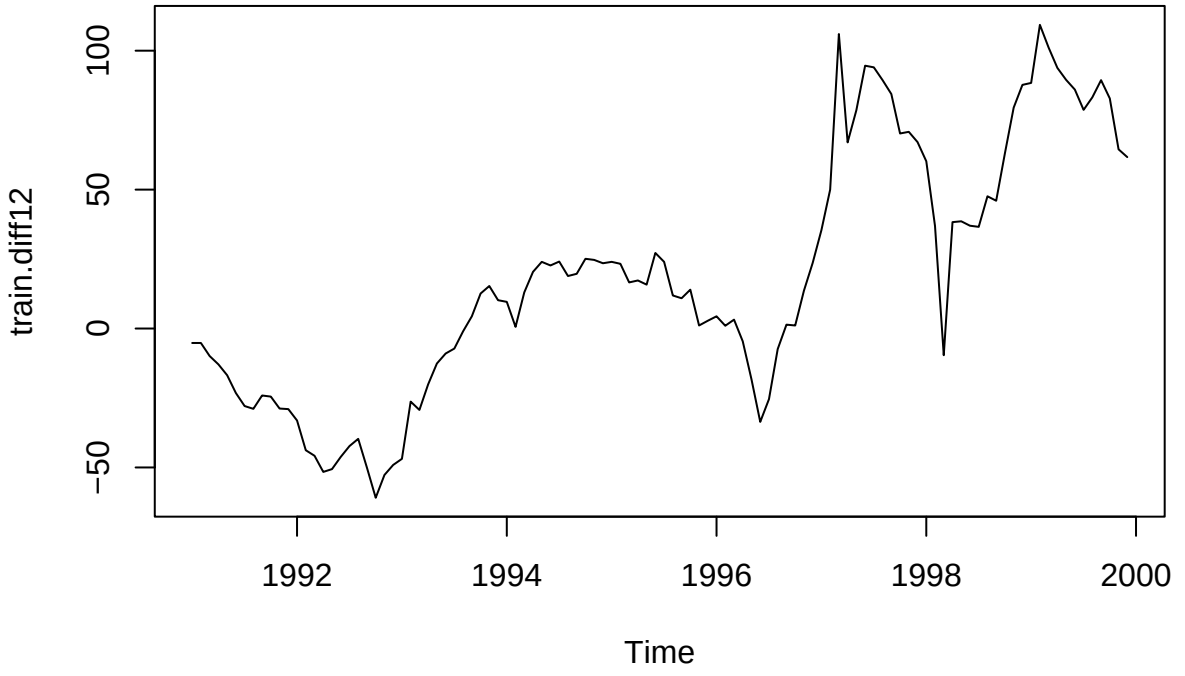
```
var(train)
```

```
## [1] 5369.962
```

```
plot(decompose(train))
```



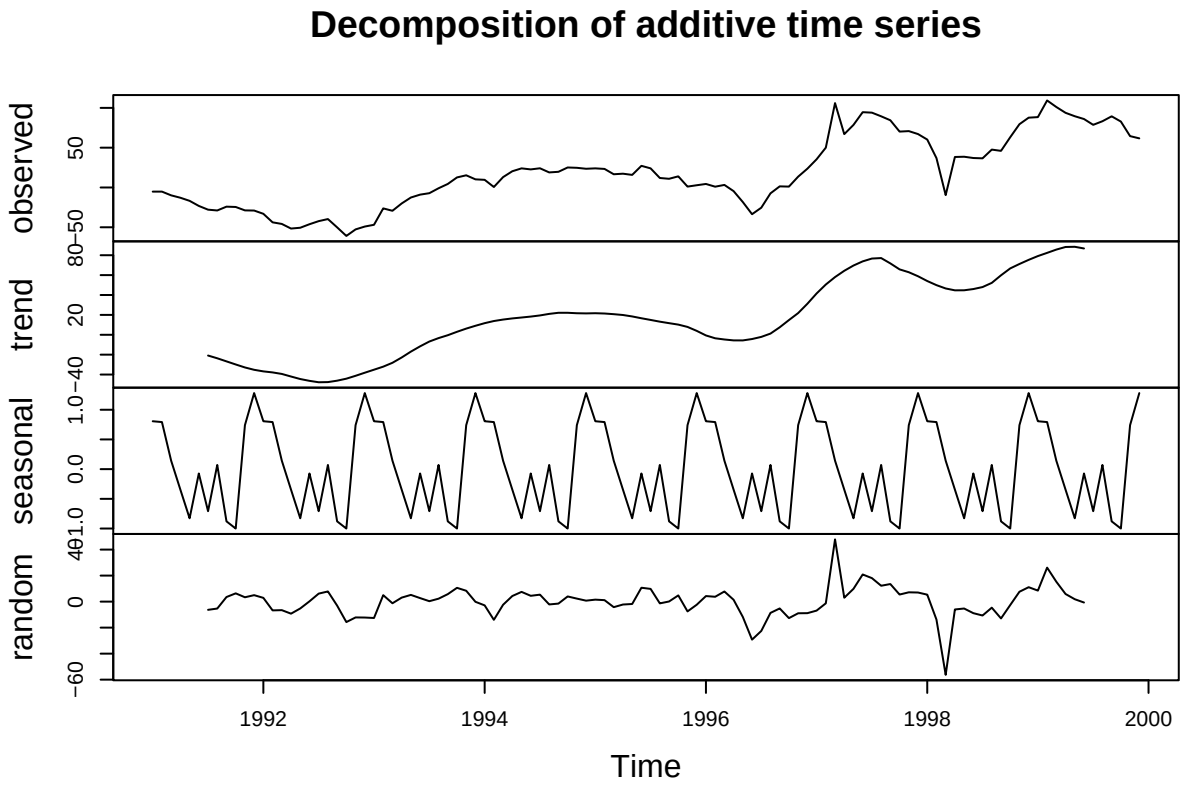

```
# Differencing at lag 12
train.diff12 <- diff(train, 12)
plot(train.diff12)
```



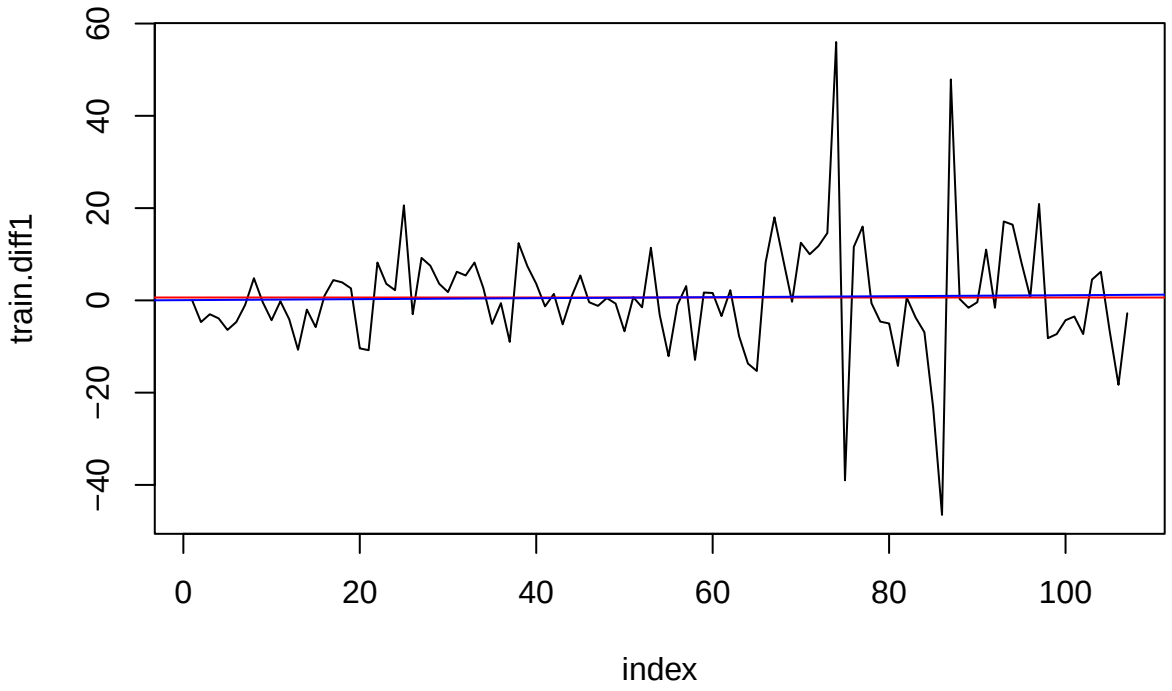

```
var(train.diff12)
```

```
## [1] 1963.123
```

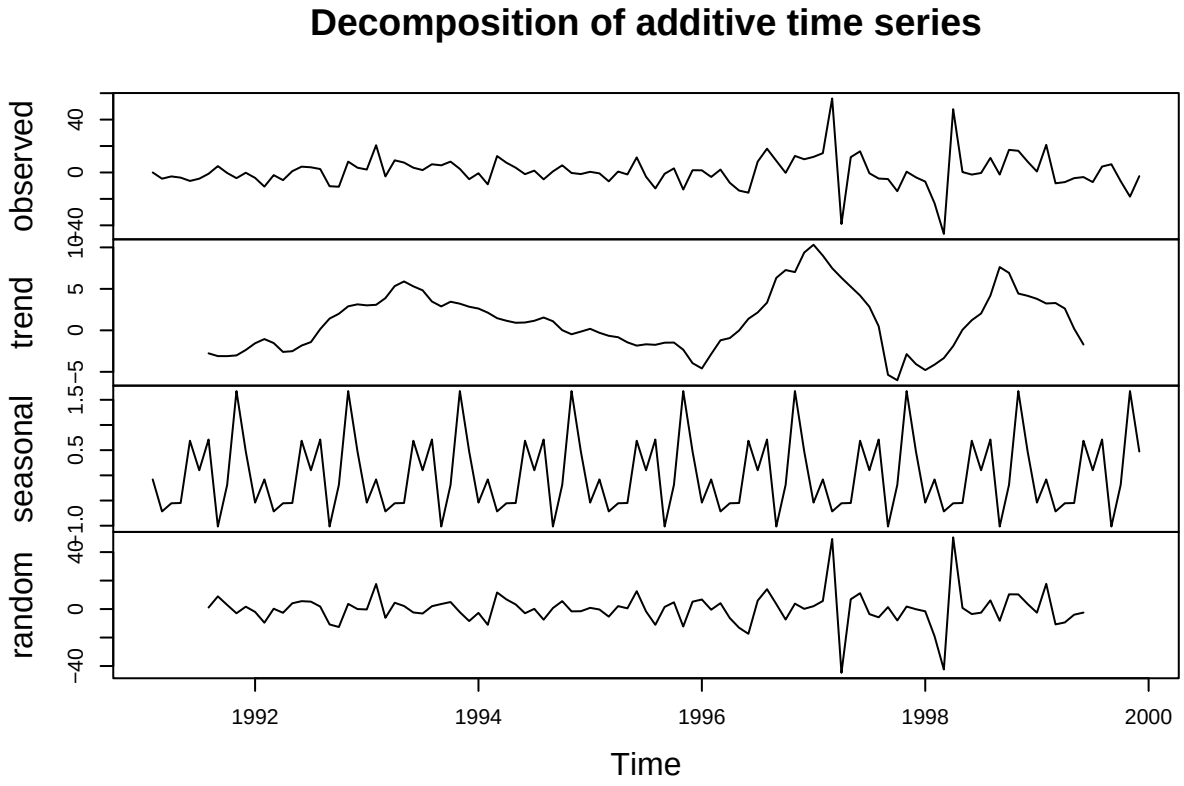
```r
plot(decompose(train.diff12))
```




```r
# Differencing at lag 1
train.diff1 <- diff(train.diff12, 1)
index <- 1:length(train.diff1)
plot(index, train.diff1, type='l')
abline(h=mean(train.diff1), col="red")
abline(lm(train.diff1 ~ index), col="blue")
```




```
plot(decompose(train.diff1))
```




```
mean(train.diff1)
```

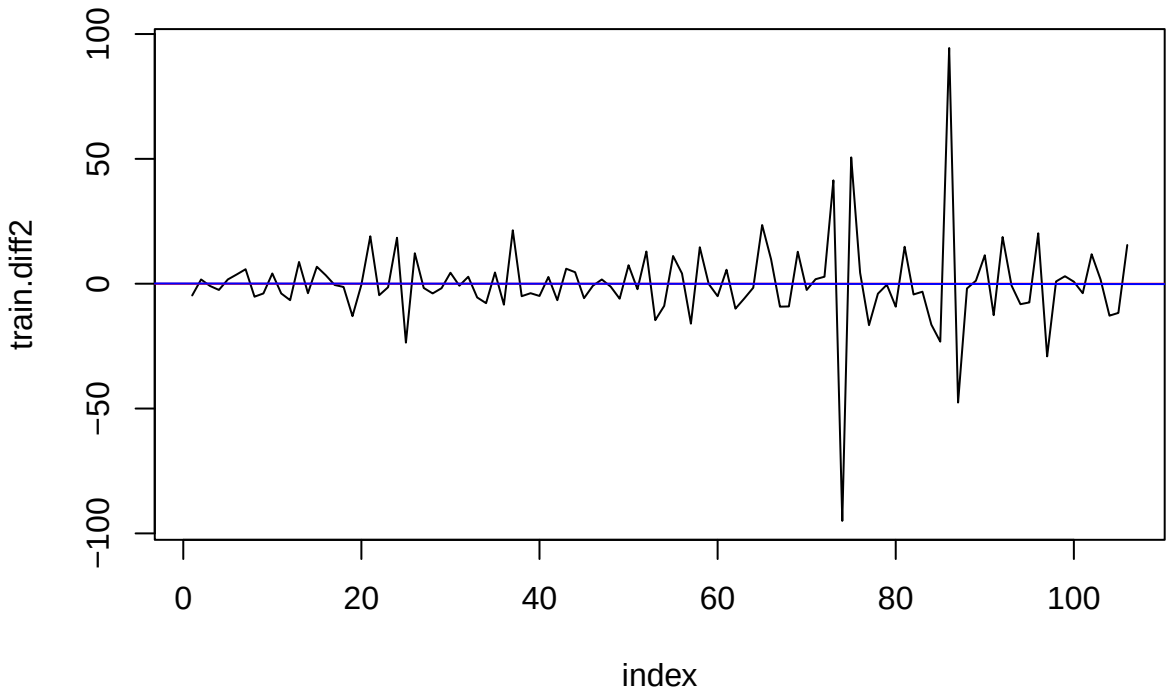
```
## [1] 0.6252336
```

```
var(train.diff1)
```

```
## [1] 151.3917
```

Variance still decreased so we accept both differences.

```
# Differencing at lag 2
train.diff2 <- diff(train.diff12, lag=1, differences=2)
index <- 1:length(train.diff2)
plot(index, train.diff2, type='l')
abline(h=mean(train.diff2), col="red")
abline(lm(train.diff2 ~ index), col="blue")
```




```
mean(train.diff2)
```

```
## [1] -0.02641509
```

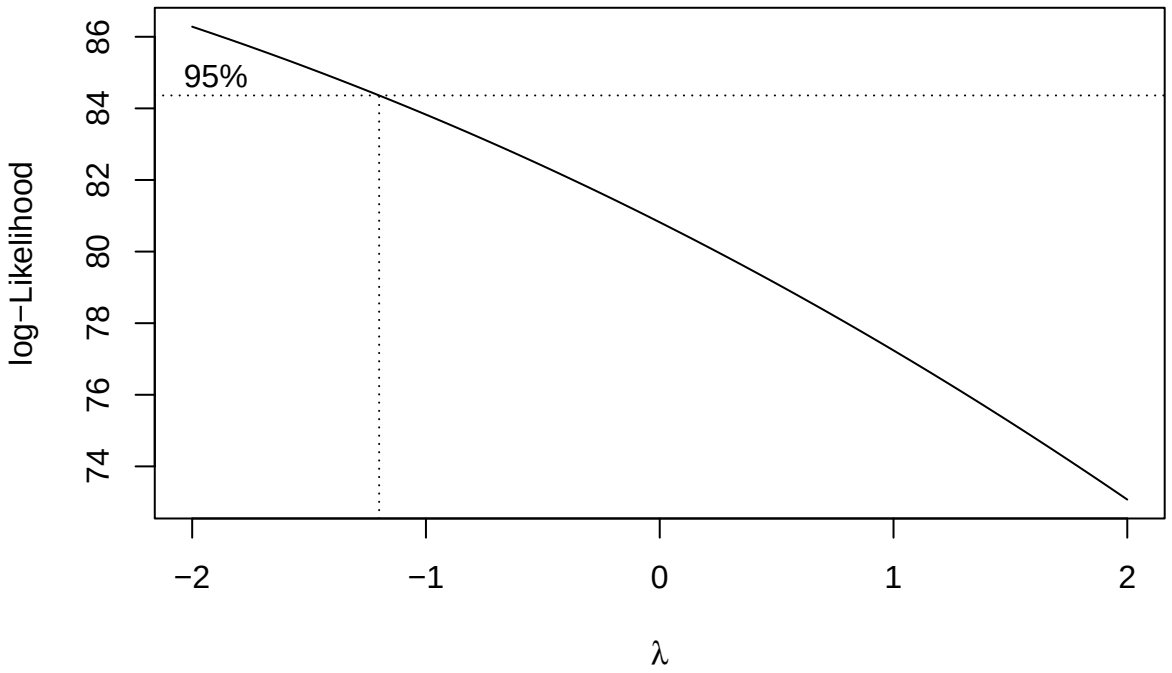
```
var(train.diff2)
```

```
## [1] 322.4597
```

Variance significantly higher than differncing only once at lag 1. Reject second differencing, keep at 1.

With Box-Cox Transformation

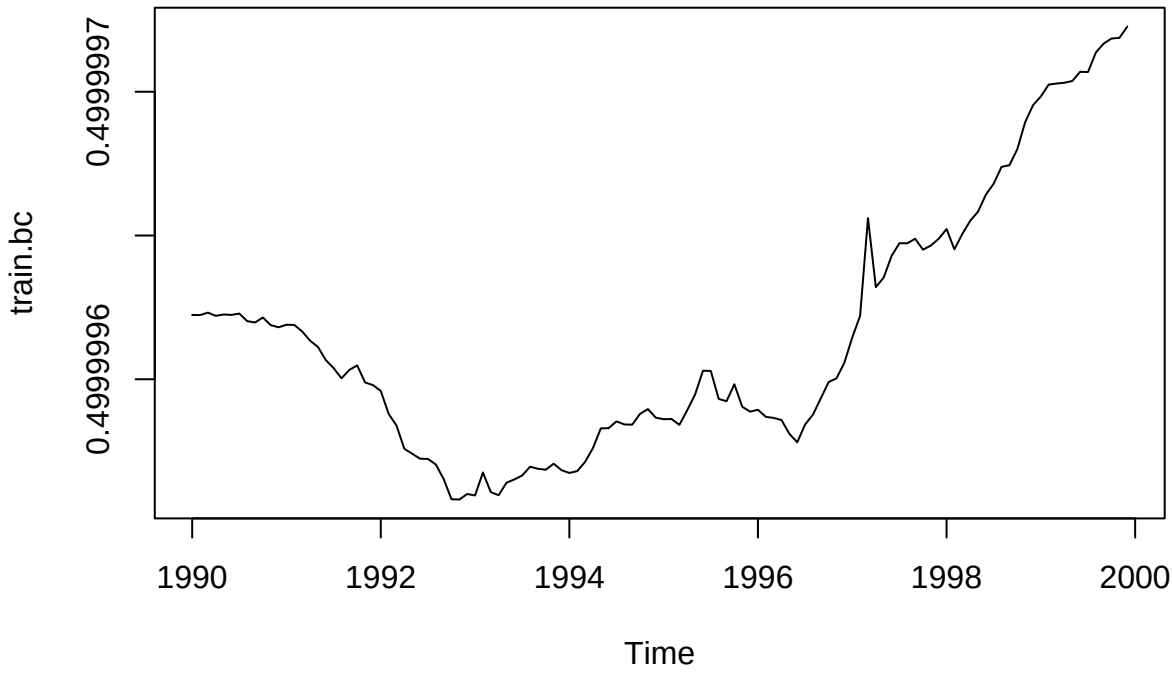
```
# Box-Cox transformation
index <- 1:length(train)
bcTransform <- boxcox(train ~ index, plotit=TRUE)
```




```
lambda <- bcTransform$x[which(bcTransform$y == max(bcTransform$y))]
lambda
```

```
## [1] -2
```

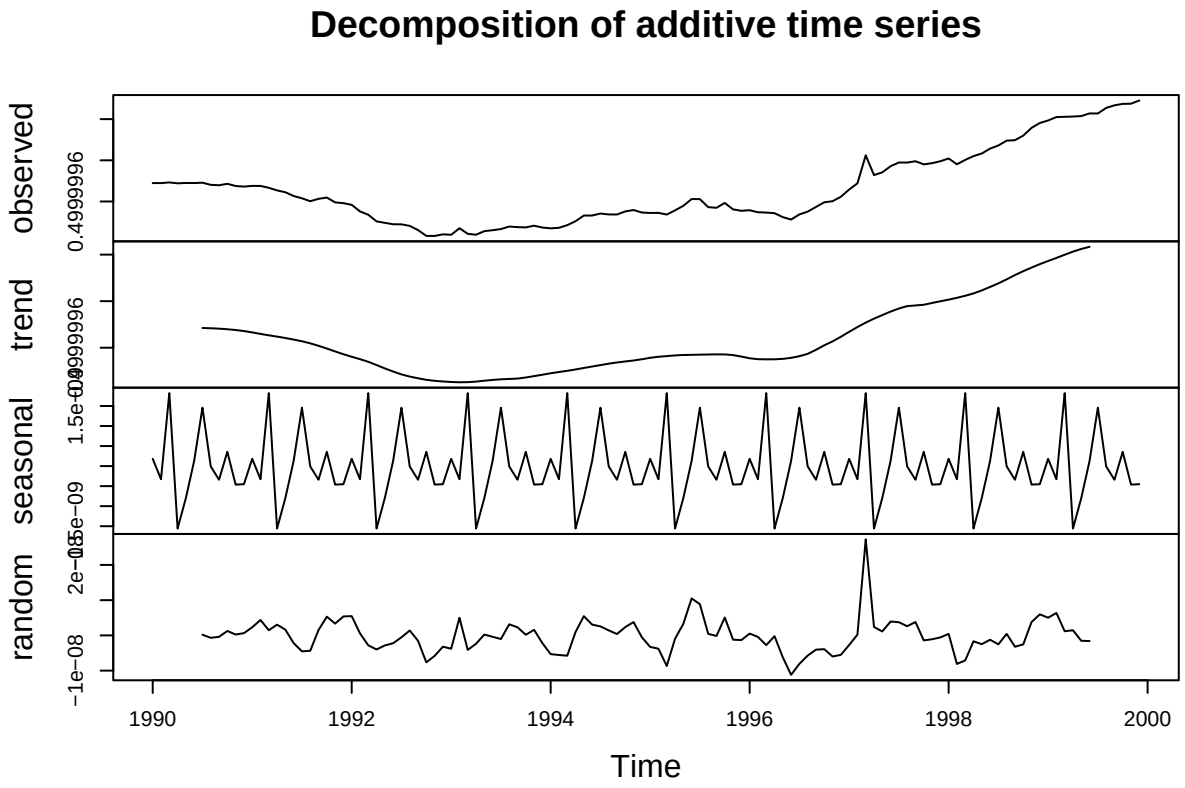
```
train.bc <- (1/lambda)*(train^lambda - 1)

plot.ts(train.bc)
```




```
var(train.bc)
```

```
## [1] 1.98852e-15
```

```
plot(decompose(train.bc))
```

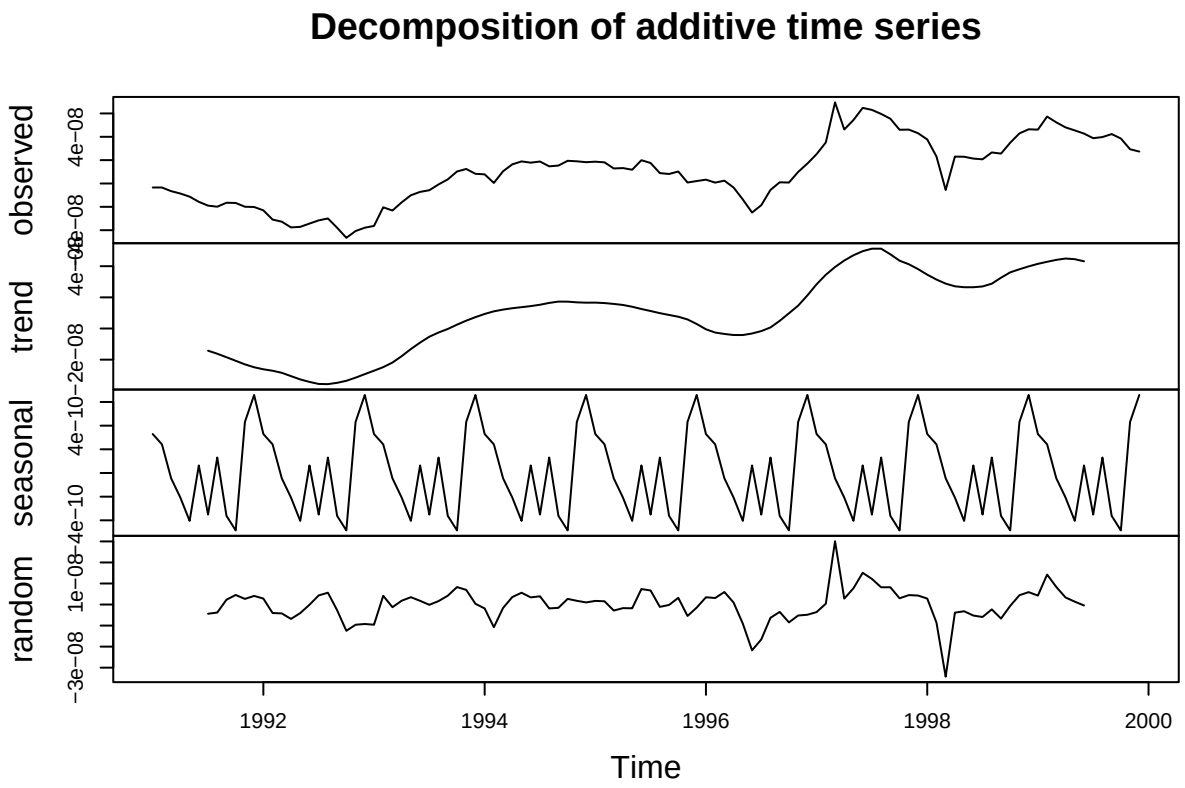



```
# Differencing at lag 12
train.bc12 <- diff(train.bc, 12)
plot(train.bc12)
```

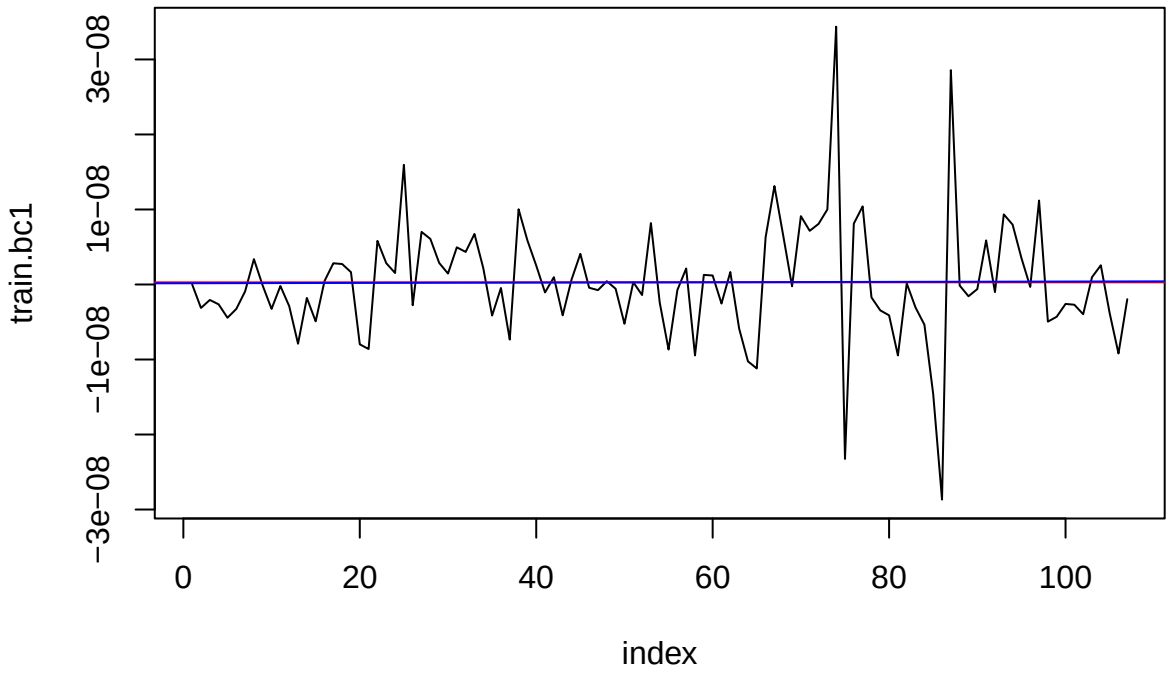

```
var(train.bc12)
```

```
## [1] 7.727189e-16
```

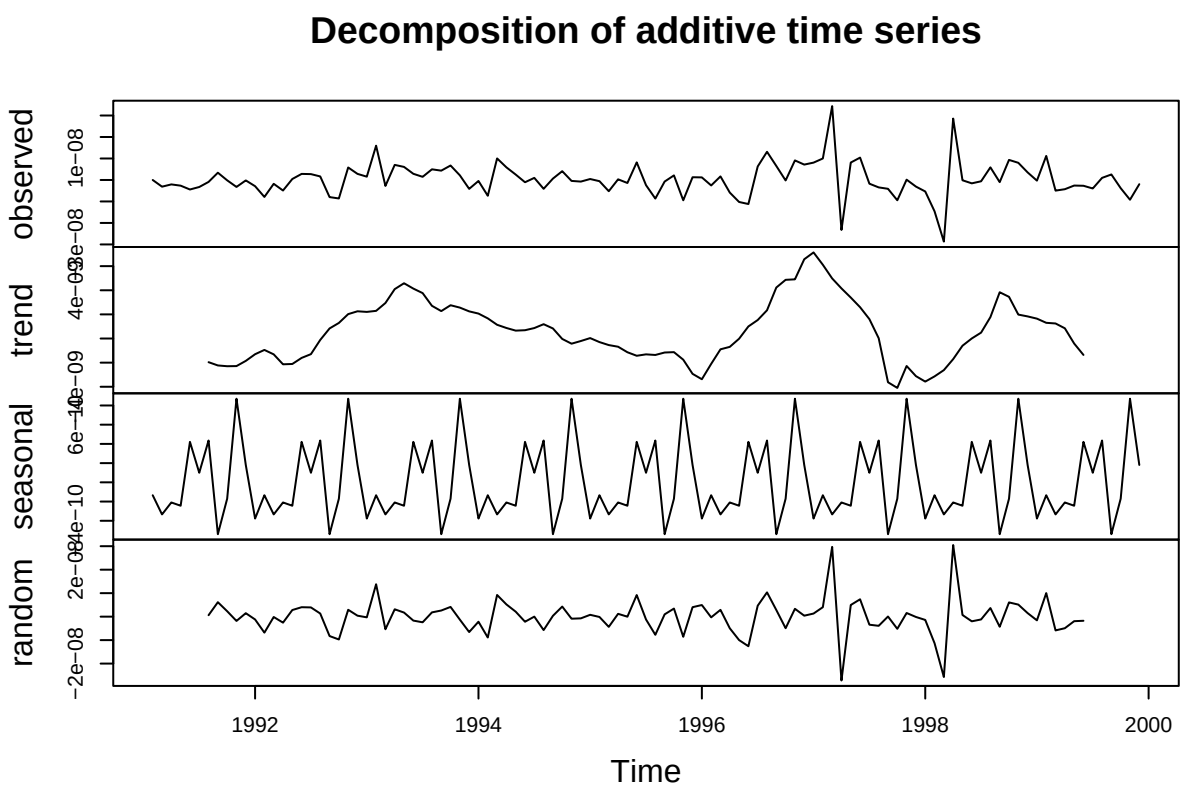
```
plot(decompose(train.bc12))
```




```
# Differencing at lag 1
train.bc1 <- diff(train.bc12, 1)
index <- 1:length(train.bc1)
plot(index, train.bc1, type='l')
abline(h=mean(train.bc1), col="red")
abline(lm(train.bc1 ~ index), col="blue")
```




```
plot(decompose(train.bc1))
```



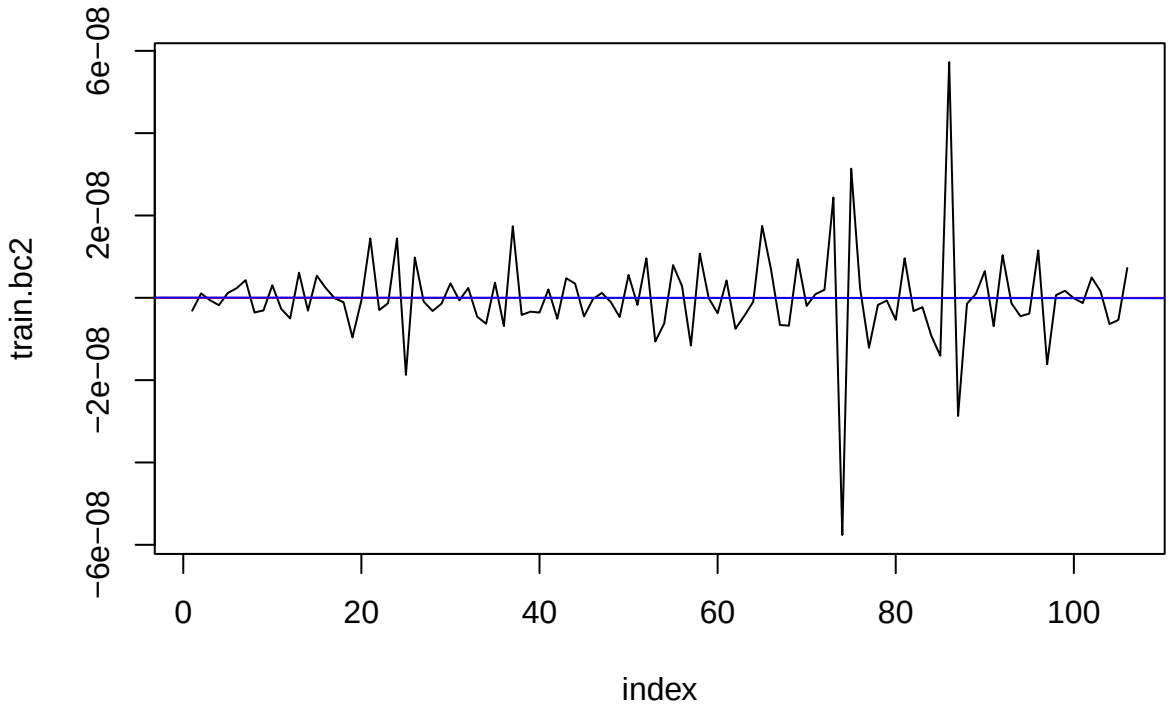

```
mean(train.bc1)
```

```
## [1] 2.878125e-10
```

```
var(train.bc1)
```

```
## [1] 6.214252e-17
```

```r
# Differencing at lag 2
train.bc2 <- diff(train.bc12, lag=1, differences=2)
index <- 1:length(train.bc2)
plot(index, train.bc2, type='l')
abline(h=mean(train.bc2), col="red")
abline(lm(train.bc2 ~ index), col="blue")
```




```r
mean(train.bc2)
```

```
## [1] -1.851421e-11
```

```r
var(train.bc2)
```

```
## [1] 1.27253e-16
```

Variance again higher than differencing only once at lag 1.

We choose to use the Box-Cox Transformation and difference once at lag 1 and once at lag 12.

```r
m1 <- train.bc1
```