STAT 478 Project

Logan Rosentreter & Zachary Kern

Mar 19, 2023

Things to include

- plot, acf, variogram, (decompose)
- qqplot, aug dick full,
- white noise test : box-pierce
- normality test : anderson-darling
- auotcorrelation & time series regression : dwt
- auotcorrelation & regression : cochrane orcutt
- auotcorrelation : Ljung-Box

general approach

1. plot

- determine basic features
  - trend, season, outliers

2. elimin trend, seas

- diff
- apply appropriate model

3. develop forecast model for residuals
4. validate performance

- split-sample
- cross-validation

5. find diff b/n orig and forecast / smoothed
6. find prediction intervals of forecast
7. develop procedure for detecting deterioration in forecast, quickly

also want to use prophet

evaluation

- ME, MAD, MSE, MPE, MAPE

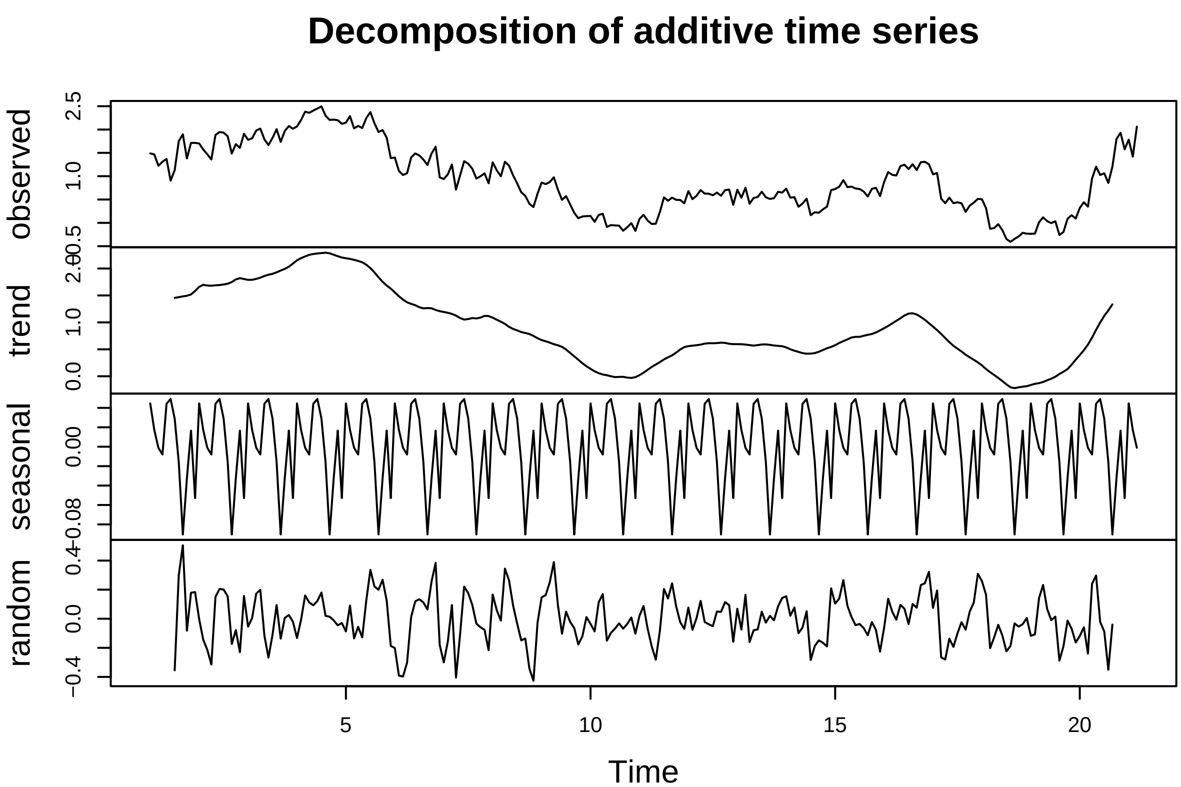
1. Problem definition / Introduction

The Fred STL dataset tracks the 10-Year Real Interest Rate in the United States ("10-Year Real Interest Rate" 2023). The 10-Year Real Interest Rate provides valuable insights into the state of the economy and the financial market, as it provides a measure of the real cost of borrowing and the expected return on investment. When the 10-Year Real Interest Rate is low, it can stimulate economic growth by making borrowing cheaper and encouraging investment. When the 10-Year Real Interest Rate is high, it can restrict economic growth by increasing the cost of borrowing and reducing investment.

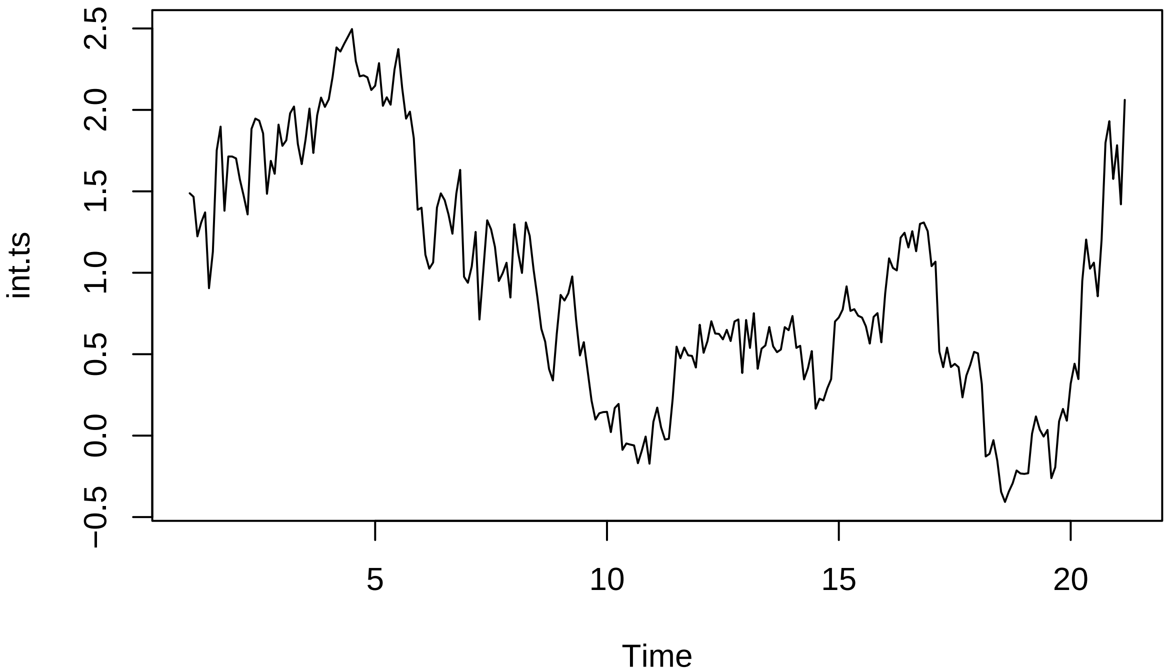
```
# read in data set
int.rate <- read.csv("RIRA10Y.csv")

# rename columns for easier data manipulation
int.rate <- int.rate %>%
    rename(INT_RATE_10Y = REAINTRATREARAT10Y)

# convert date to type(date)
int.rate <- int.rate %>%
    mutate(DATE = as.Date(DATE, sep = "-", "%Y-%m-%d"))
str(int.rate)
```

```
## 'data.frame':    243 obs. of  2 variables:
##  $ DATE        : Date, format: "2003-01-01" "2003-02-01" "2003-03-01" ...
##  $ INT_RATE_10Y: num  1.49 1.47 1.22 1.31 1.37 ...
```

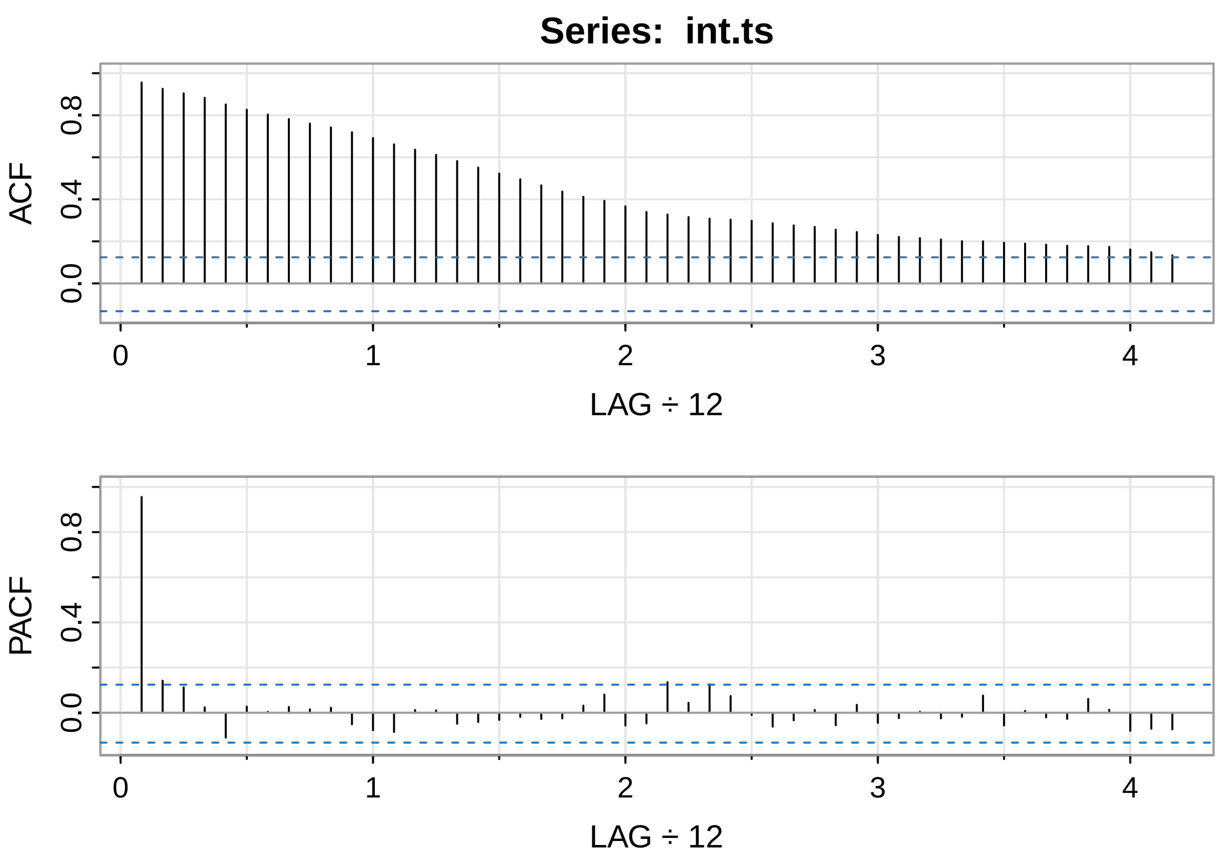
```
int.ts <- ts(int.rate$INT_RATE_10Y, freq = 12)
plot(decompose(int.ts))
```



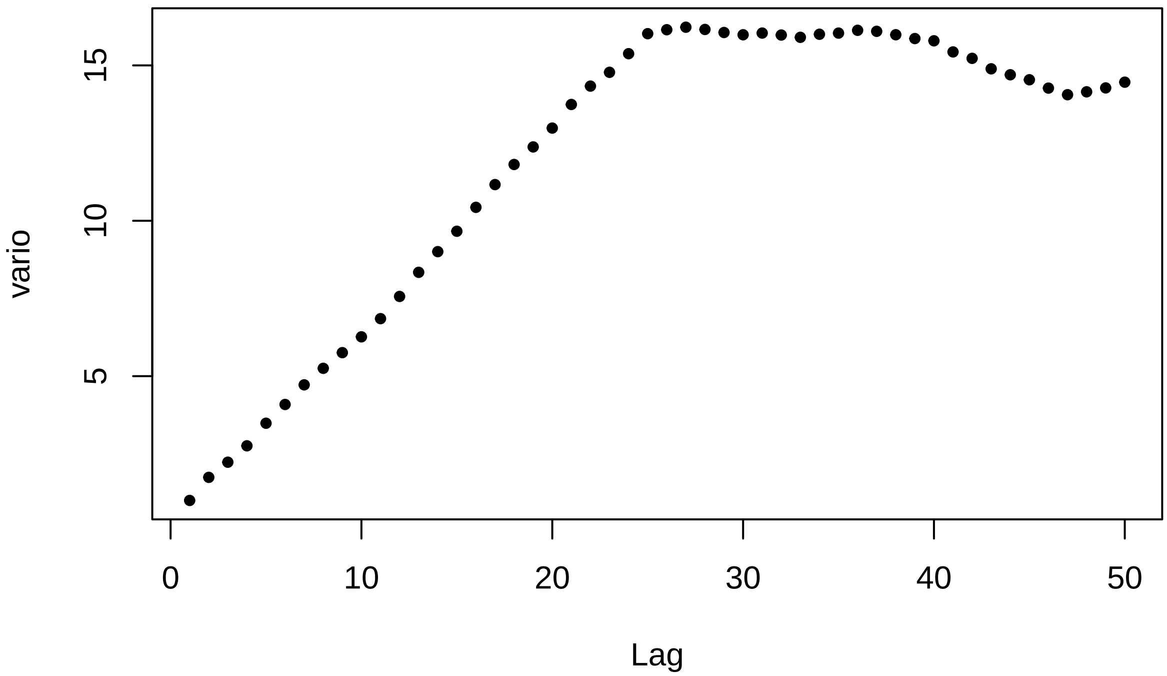
```r
plot(int.ts)
```




```r
int_acf2 <- acf2(int.ts, max.lag = 50)
```




```r
plot(variogram(int.ts, 50), pch = 19, cex = 0.65)
```



notes - orig

- data shows down trend, no seas
- noncons mean, cons variance
- acf -> nonstation
- variogram -> monotonically increasing for long period of time -> nonstat

```
int.lm <- lm(int.ts ~ int.rate$DATE)
summary(int.lm)
```

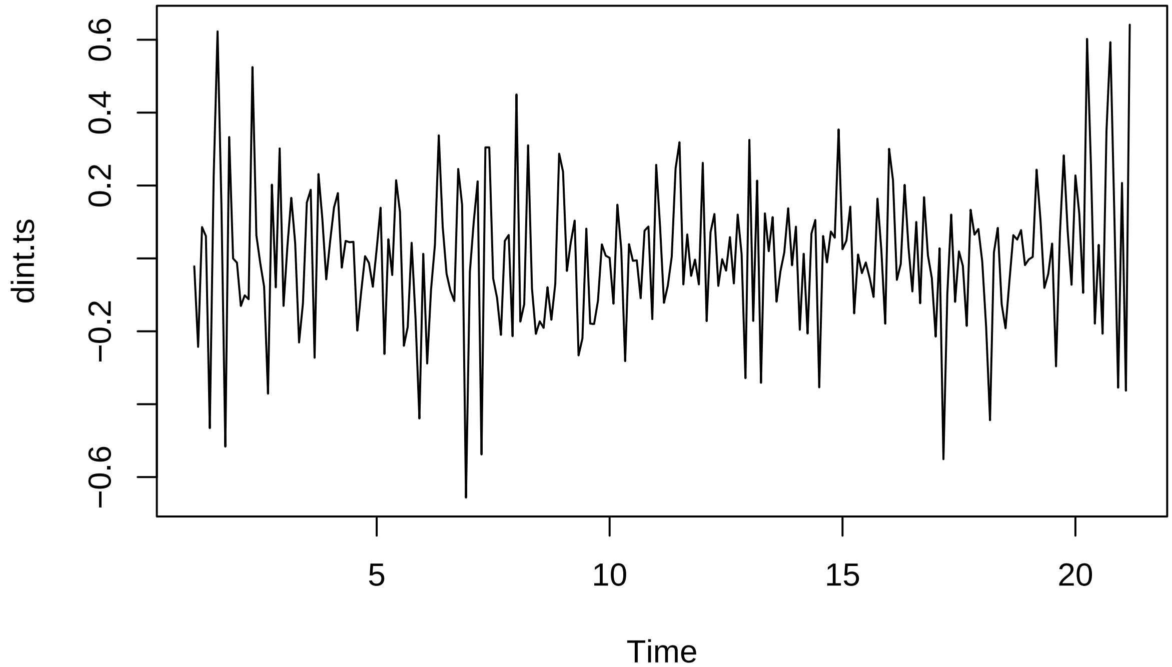
```
##
## Call:
## lm(formula = int.ts ~ int.rate$DATE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12750 -0.34482 -0.07004  0.29486  1.90524
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.219e+00  2.695e-01   15.65   <2e-16 ***
## int.rate$DATE -2.092e-04  1.697e-05  -12.33   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5649 on 241 degrees of freedom
## Multiple R-squared:  0.3867, Adjusted R-squared:  0.3842
## F-statistic:   152 on 1 and 241 DF,  p-value: < 2.2e-16
```

```
anova(int.lm)
```

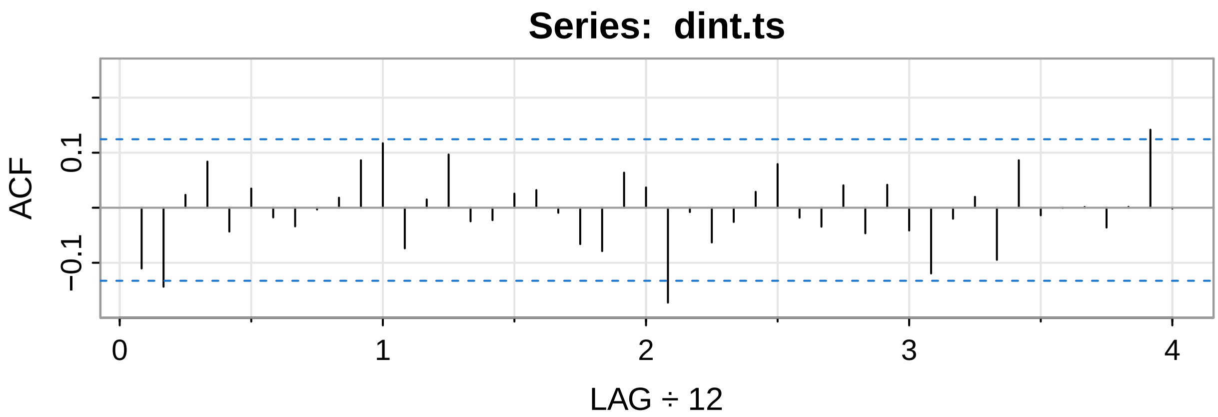
```
## Analysis of Variance Table
##
## Response: int.ts
```

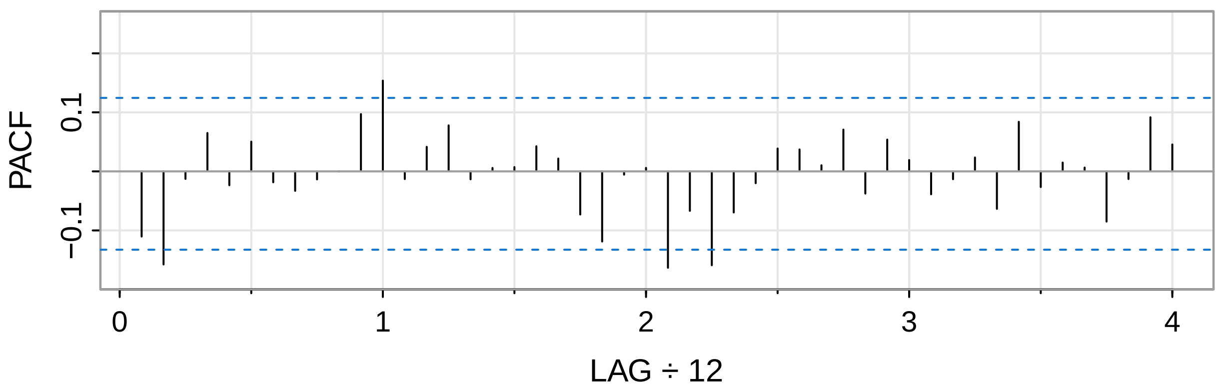
```
##                  Df Sum Sq Mean Sq F value    Pr(>F)
## int.rate$DATE   1 48.494  48.494  151.99 < 2.2e-16 ***
## Residuals     241 76.896   0.319
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# differenced data
dint.ts <- diff(int.ts)
plot(dint.ts)
```

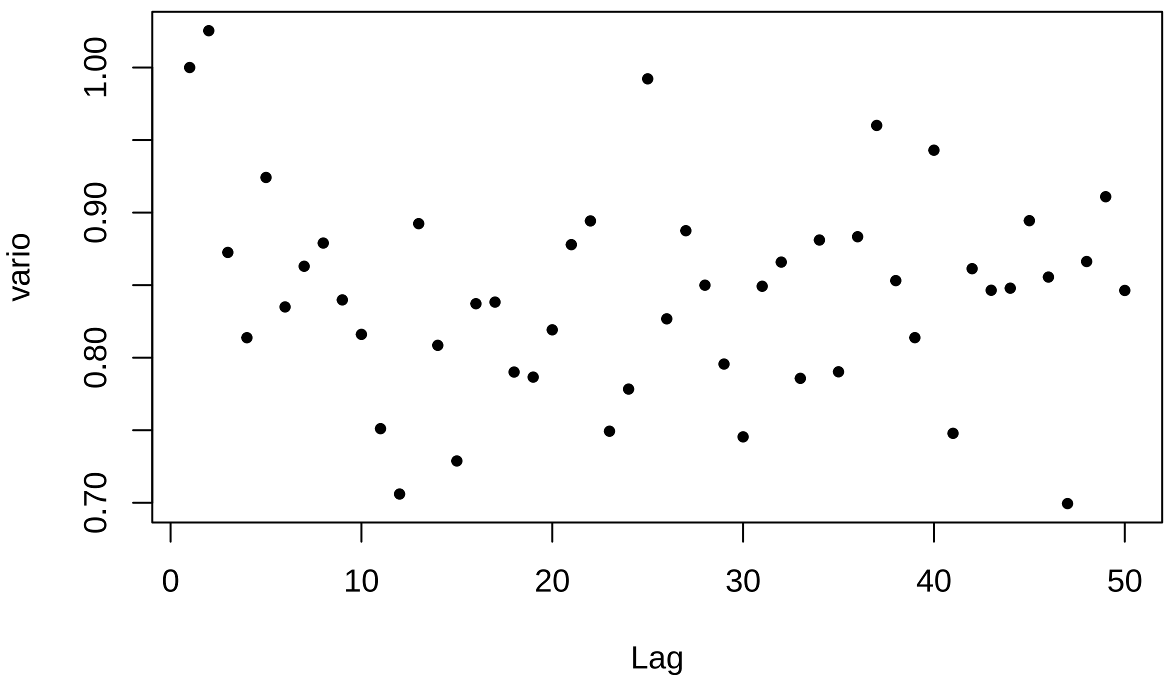



```
dint.acf2 <- acf2(dint.ts)
```






```
plot(variogram(dint.ts, 50), pch = 19, cex = 0.65)
```



notes - diff

- data shows no trend, no seas, rand scatter

- cons mean, cons variance
- acf -> station
- variogram -> random scatter -> stat

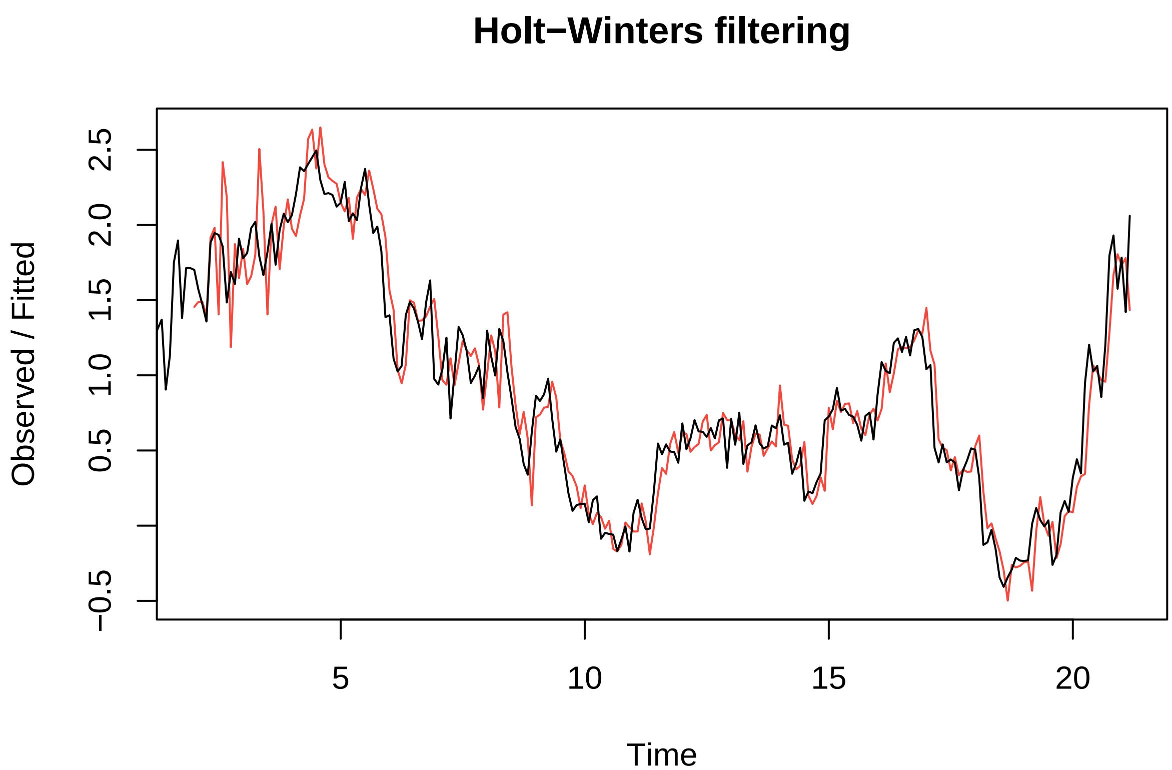
```
dint.lm <- lm(dint.ts ~ int.rate$DATE[-1])
summary(dint.lm)
```

```
##
## Call:
## lm(formula = dint.ts ~ int.rate$DATE[-1])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.64981 -0.11172  0.00324  0.10350  0.64063
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -8.795e-02  9.531e-02  -0.923    0.357
## int.rate$DATE[-1]  5.734e-06  5.997e-06   0.956    0.340
##
## Residual standard error: 0.1984 on 240 degrees of freedom
## Multiple R-squared:  0.003795,   Adjusted R-squared:  -0.0003556
## F-statistic: 0.9143 on 1 and 240 DF,  p-value: 0.3399
```

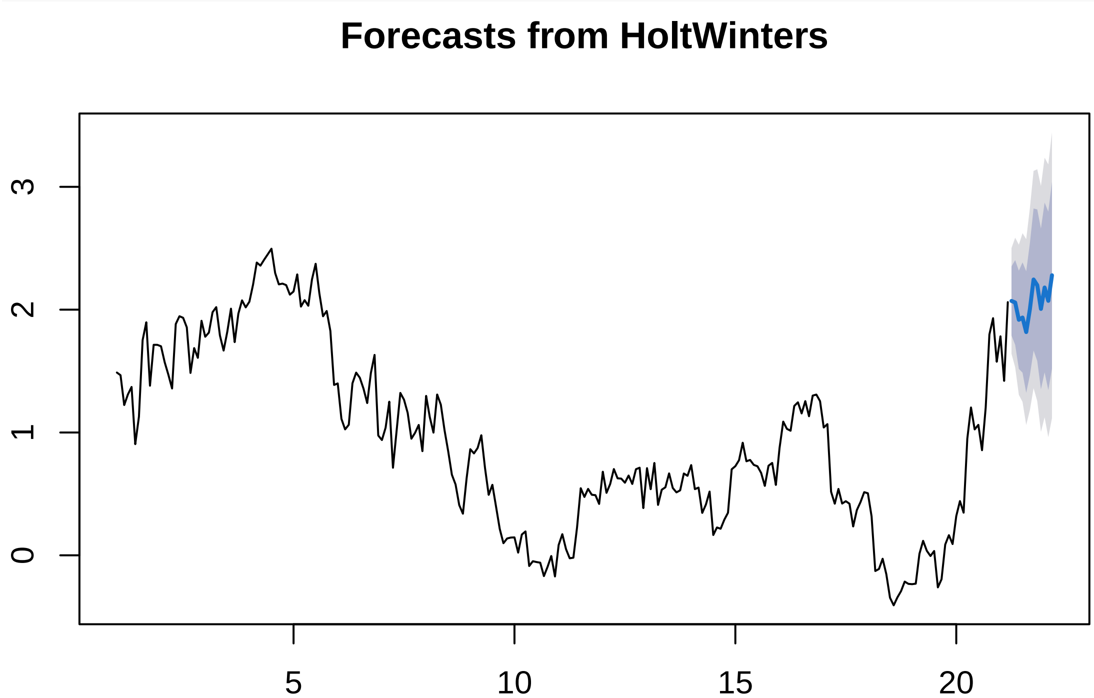
```
anova(dint.lm)
```

```
## Analysis of Variance Table
##
## Response: dint.ts
##                    Df Sum Sq  Mean Sq F value Pr(>F)
## int.rate$DATE[-1]   1 0.0360 0.035979  0.9143 0.3399
## Residuals         240 9.4439 0.039350
```

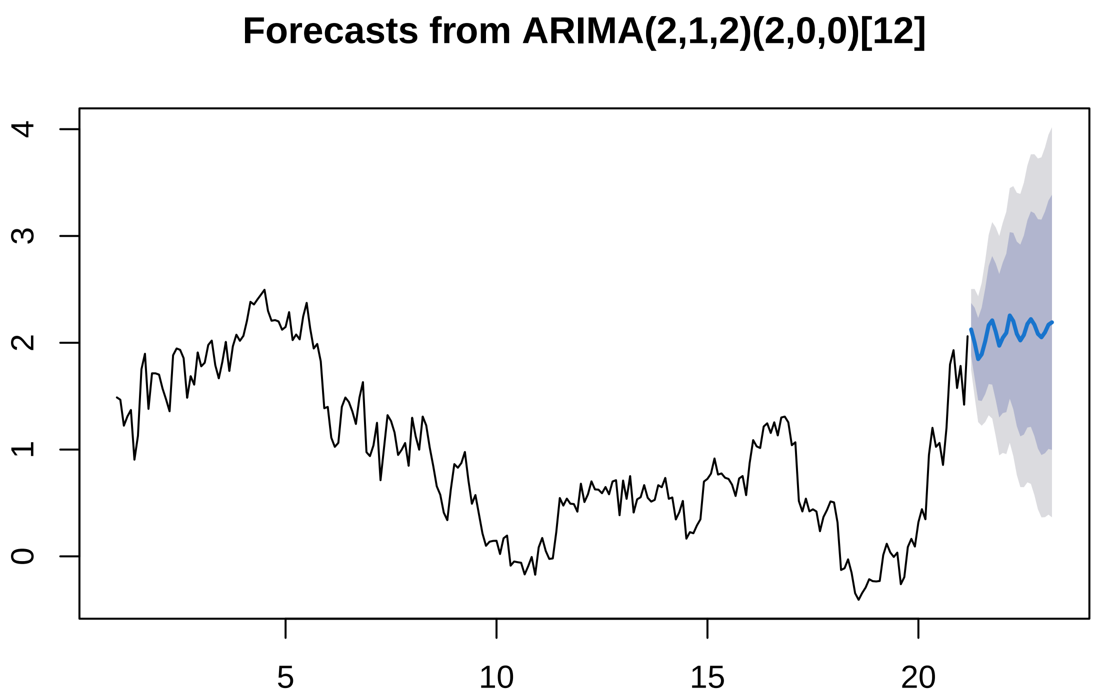
```
int.ses <- HoltWinters(int.ts)
plot(int.ses)
```




```
int.ses.for <- forecast(int.ses, h = 12)
plot(int.ses.for)
```




```
int.arima <- auto.arima(int.ts)
plot(forecast(int.arima))
```

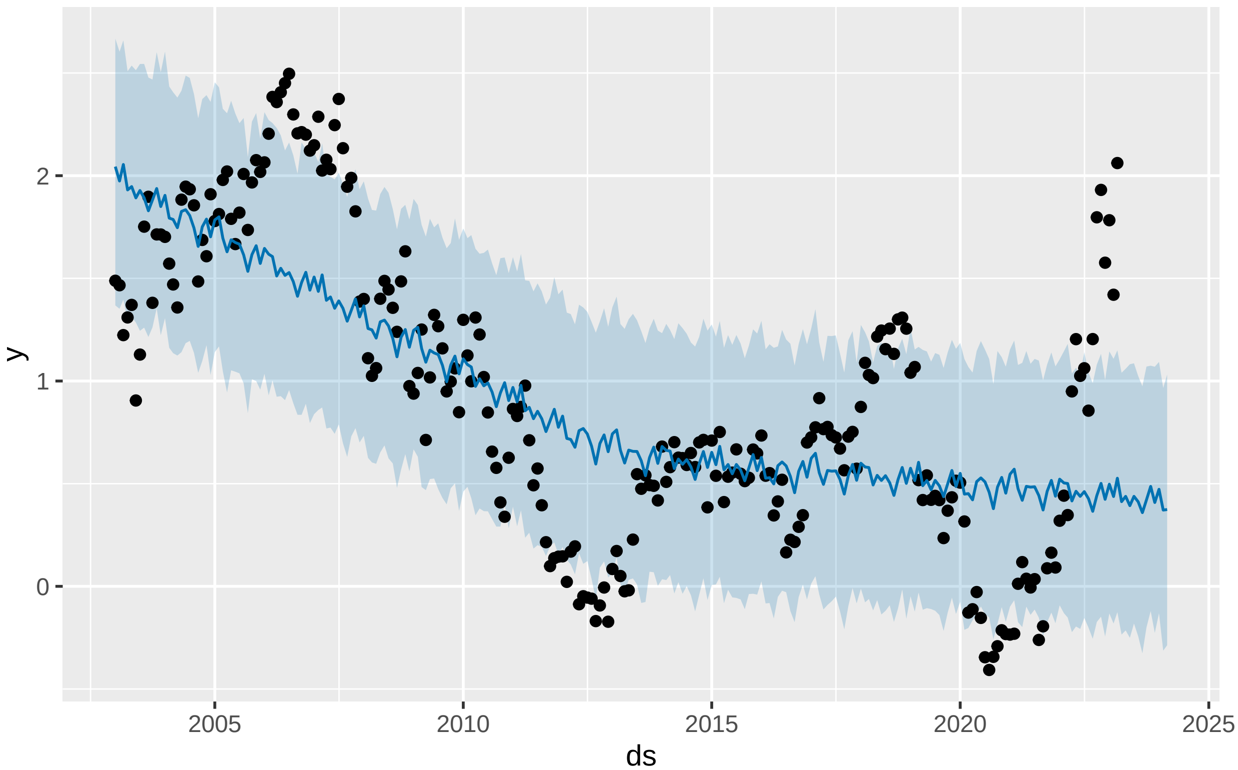



```
int.rate.prophet <- int.rate %>%
    rename(ds = DATE, y = INT_RATE_10Y)

# implementing prophet
m <- prophet(int.rate.prophet)
```

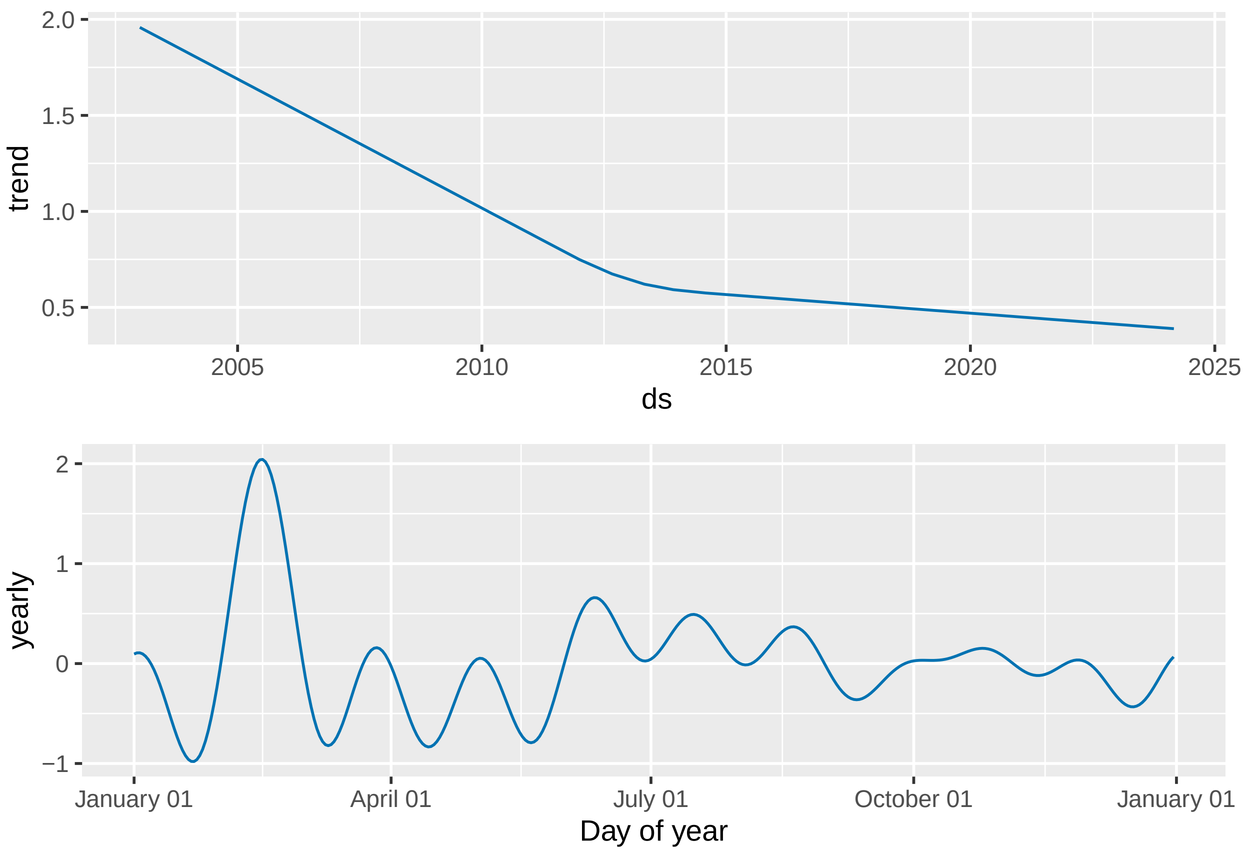
```
## Disabling weekly seasonality. Run prophet with weekly.seasonality=TRUE to override this.
```

```
## Disabling daily seasonality. Run prophet with daily.seasonality=TRUE to override this.
```

```
future <- make_future_dataframe(m, freq = "month", periods = 12)
forecast <- predict(m, future)
plot(m, forecast)
```




```r
prophet_plot_components(m, forecast)
```



notes - log trans

- cant use logarithm, because of negatives

2. Data description

3. Data Analysis

4. Model specification and fitting

5. Model validation and diagnostics

6. Forecasting

Conclusion

"10-Year Real Interest Rate." 2023. *FRED*. FRED. https://fred.stlouisfed.org/series/REAINTRATREARAT10Y.