STAT 478 Project

Logan Rosentreter & Zachary Kern

Mar 19, 2023

```r
# Define a function variogram
variogram <- function(x, lag) {
    Lag <- NULL
    var_k <- NULL
    vario <- NULL
    for (k in 1:lag) {
        Lag[k] <- k
        var_k[k] = (sd(diff(x, k)))^2
        vario[k] = var_k[k]/var_k[1]
    }
    return(as.data.frame(cbind(Lag, vario)))
}
```

```r
# check if `tidyr` package is installed; otherwise, install
# and load if 'tidyr' is absent
if (!require("astsa")) install.packages("astsa", repos = "https://cloud.r-project.org")

# load libraries

# rename
library(dplyr)

# acf2
library(astsa)
```

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

evaluation

- ME, MAD, MSE, MPE, MAPE

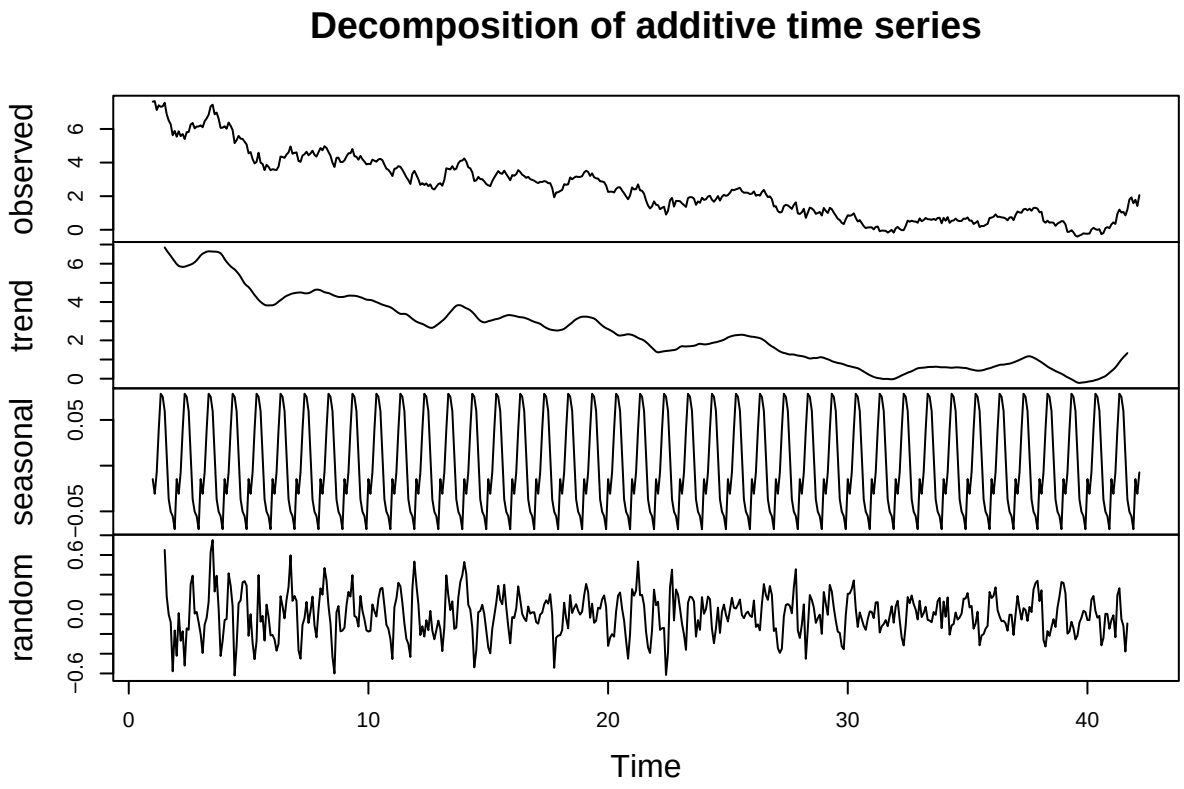
1. Problem definition / Introduction

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.

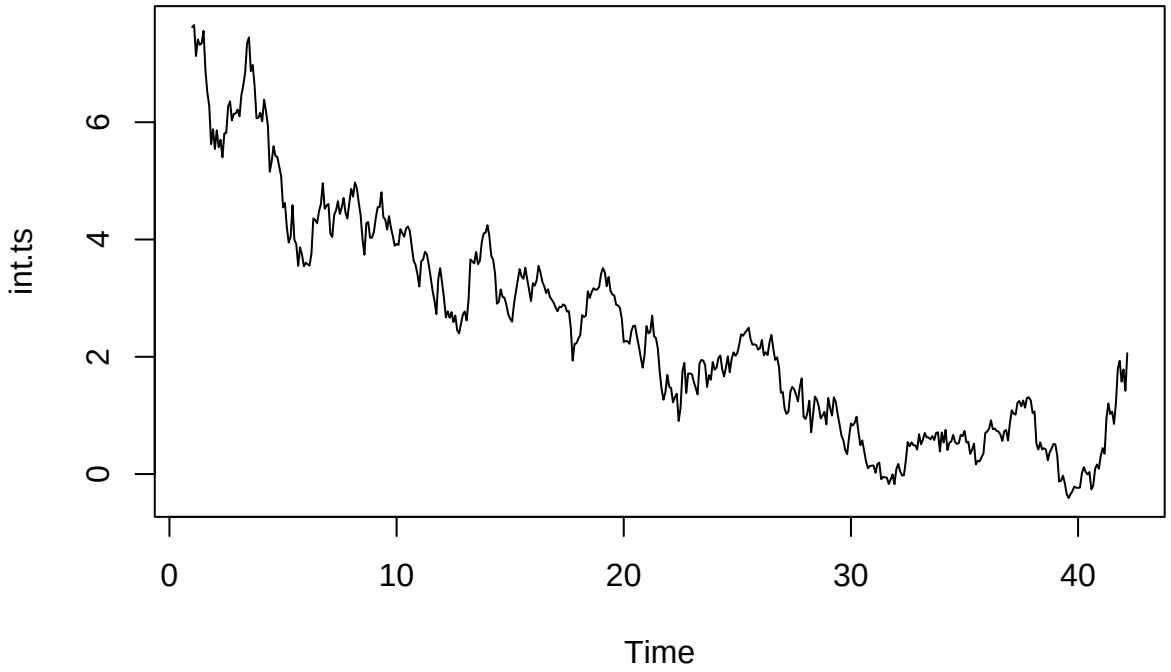
```
# read in data set
int.rate <- read.csv("RIRA10Y.csv")

# 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':    495 obs. of  2 variables:
##  $ DATE       : Date, format: "1982-01-01" "1982-02-01" "1982-03-01" ...
##  $ INT_RATE_10Y: num  7.62 7.66 7.13 7.41 7.32 ...
```

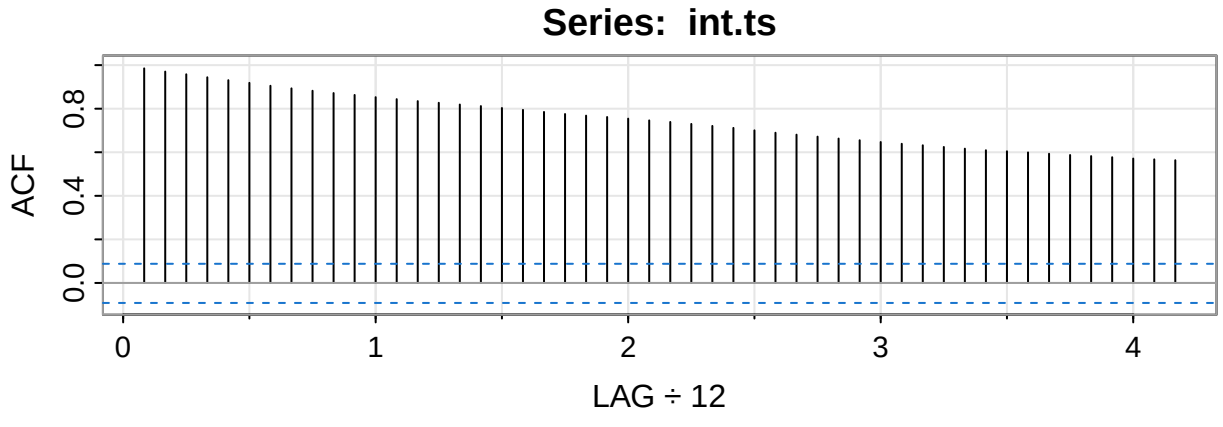
```
int.ts <- ts(int.rate$INT_RATE_10Y, freq = 12)
plot(decompose(int.ts))
```

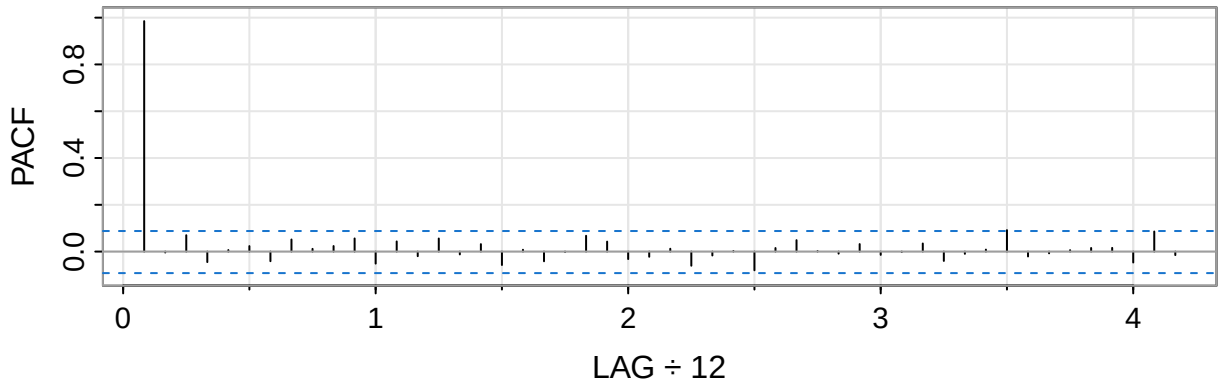



```
plot(int.ts)
```

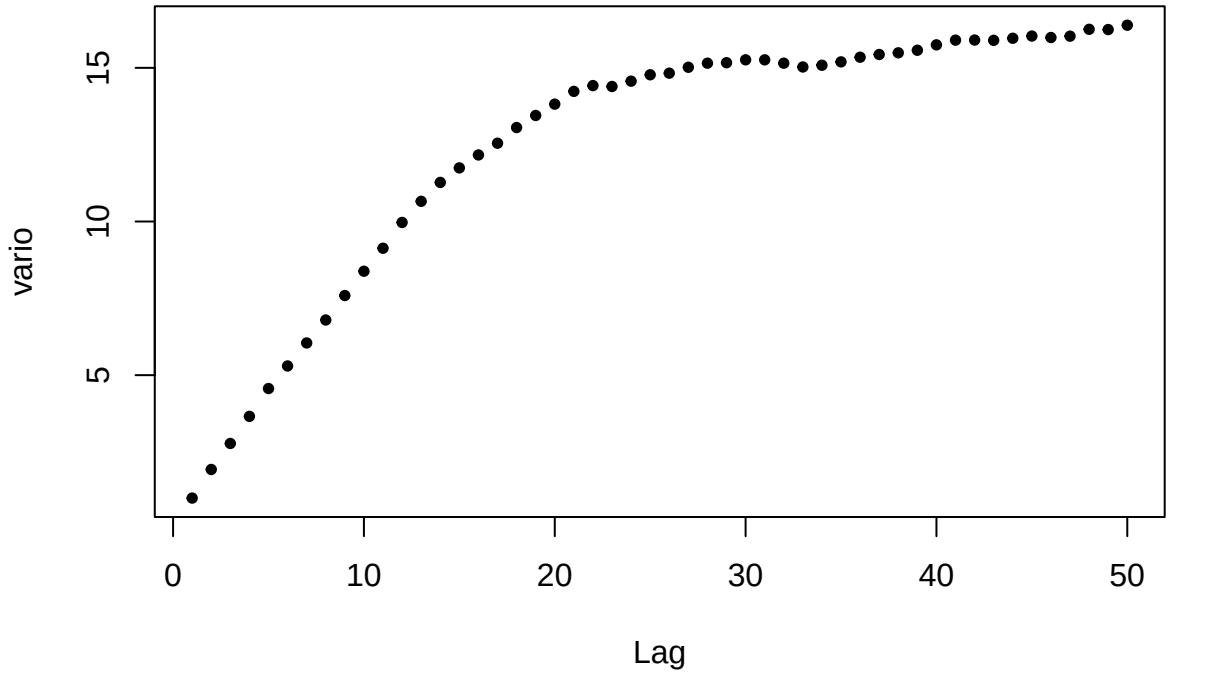



```
int_acf2 <- acf2(int.ts, max.lag = 50)
```






```
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

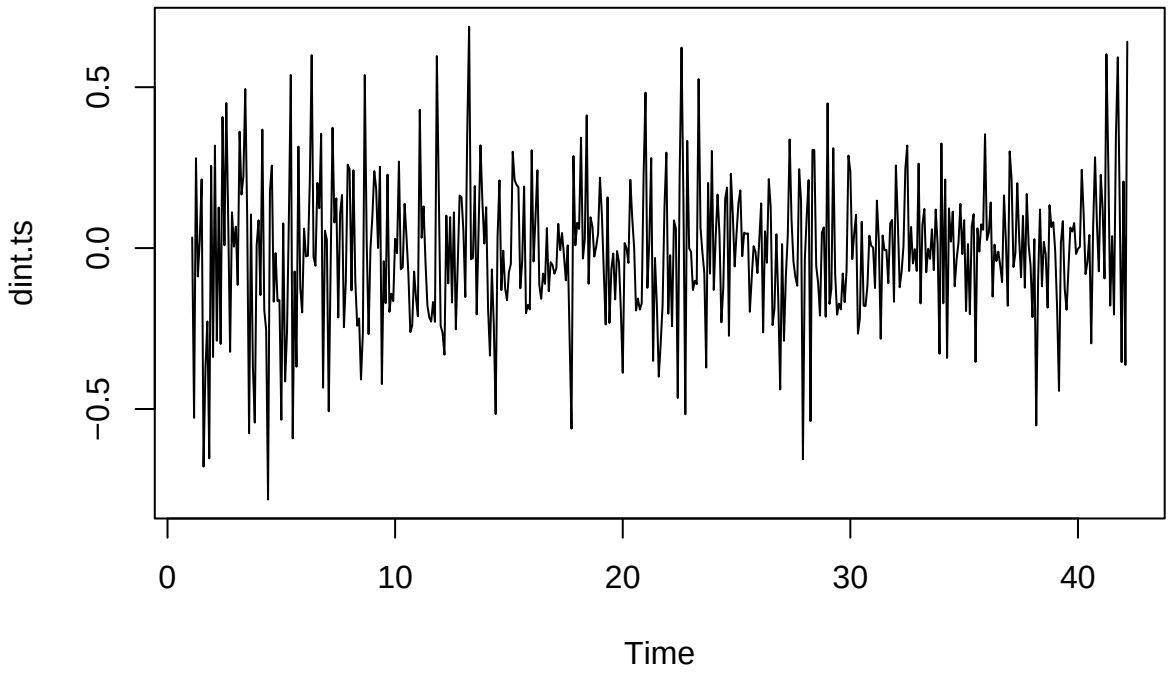
```
int.lm <- lm(int.ts ~ int.rate$DATE)
summary(int.lm)
```

```
##
## Call:
## lm(formula = int.ts ~ int.rate$DATE)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4336 -0.4273 -0.0615  0.3477  2.5861
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    7.183e+00  9.380e-02   76.58   <2e-16 ***
## int.rate$DATE -3.970e-04  7.404e-06  -53.62   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7164 on 493 degrees of freedom
## Multiple R-squared:  0.8536, Adjusted R-squared:  0.8533
## F-statistic:  2875 on 1 and 493 DF,  p-value: < 2.2e-16
```

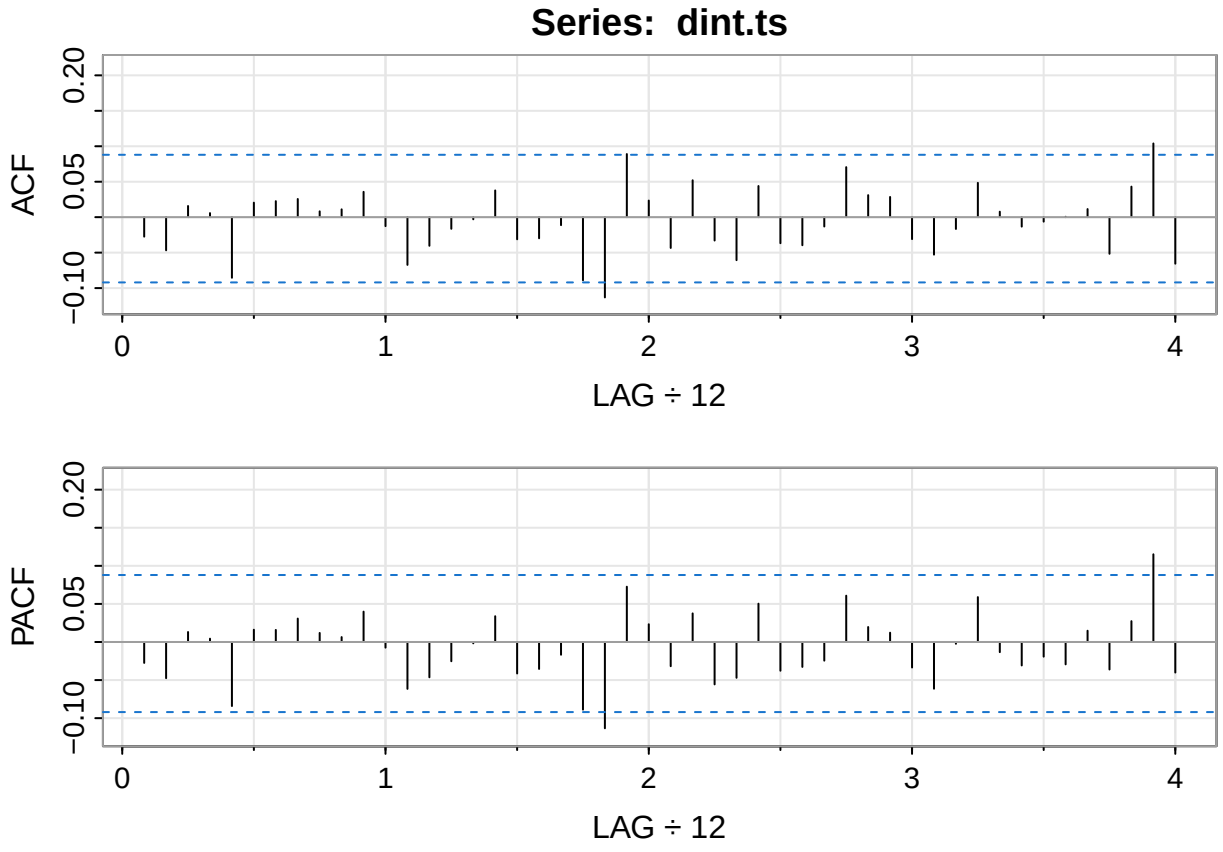
```
anova(int.lm)
```

```
## Analysis of Variance Table
##
## Response: int.ts
##                Df  Sum Sq Mean Sq F value    Pr(>F)
## int.rate$DATE   1 1475.48 1475.48  2874.7 < 2.2e-16 ***
## Residuals     493  253.04    0.51
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

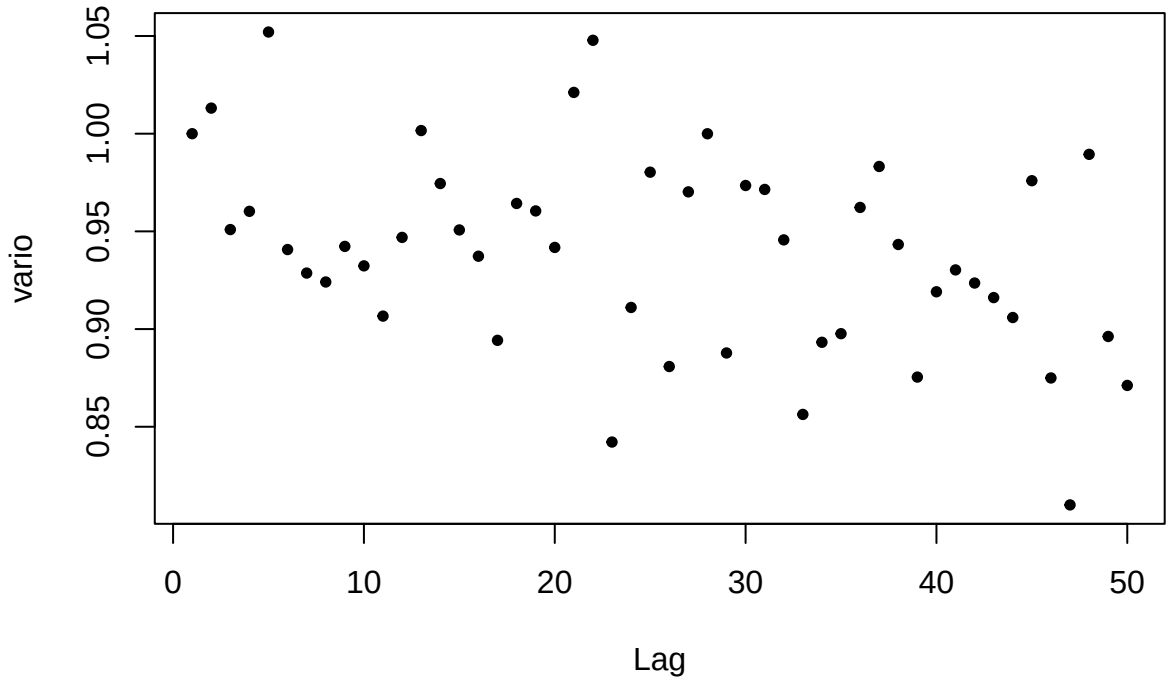
```
# differenced data
dint.ts <- diff(int.ts)
plot(dint.ts)
```




```
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$DATE[-1])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.74597 -0.13375  0.00034  0.12327  0.71104
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -5.723e-02  2.921e-02  -1.959   0.0507 .
## int.rate$DATE[-1]  3.858e-06  2.304e-06   1.675   0.0946 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2222 on 492 degrees of freedom
## Multiple R-squared:  0.005669,   Adjusted R-squared:  0.003648
## F-statistic: 2.805 on 1 and 492 DF,  p-value: 0.09462
```

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

```
## Analysis of Variance Table
##
```

```
## Response: dint.ts
##                     Df  Sum Sq  Mean Sq F value  Pr(>F)
## int.rate$DATE[-1]    1  0.1385 0.138527  2.8048 0.09462 .
## Residuals          492 24.2994 0.049389
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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.