Granger Causality Test

Katie Schilling - 501130072

07/02/2022

Required tools to be loaded

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
require(lmtest)
```

```
## Loading required package: lmtest

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

Import both Clean datasets

```
Covid_monthly <- read.csv("C:/Users/Katie Schilling/Downloads/covid_monthly_clean.csv")
Vital_Events <- read.csv("C:/Users/Katie Schilling/Downloads/vital_events_clean.csv")
```

Combine the vital events data with the Covid Monthly data

```
Final_dataset <- merge(x=Vital_Events, y=Covid_monthly, all = TRUE)
```

Check the data and ensure data merged properly

```
summary(Final_dataset)
```

```
##      Date               Births        Marriages         Deaths
##  Length:336         Min.   :10020   Min.   :  597   Min.   : 5926
##  Class :character   1st Qu.:11260   1st Qu.: 2596   1st Qu.: 6706
##  Mode  :character   Median :11818   Median : 3559   Median : 7326
##                     Mean   :11763   Mean   : 5085   Mean   : 7500
##                     3rd Qu.:12288   3rd Qu.: 7627   3rd Qu.: 8094
##                     Max.   :13398   Max.   :11532   Max.   :11390
##                     NA's   :6       NA's   :6       NA's   :6
##   Stillbirths         Covid
##  Min.   :  0.00   Min.   :    11
##  1st Qu.: 73.00   1st Qu.: 22889
##  Median : 90.50   Median : 75935
##  Mean   : 90.74   Mean   : 95144
##  3rd Qu.:114.00   3rd Qu.:134128
##  Max.   :156.00   Max.   :395815
##  NA's   :6        NA's   :313
```

Change the N/A in the Covid Positive Cases to 0 so that the data is not omitted from the predictions

```
Final_dataset$Covid[is.na(Final_dataset$Covid)] = 0
```

See if there are anymore NA's in the dataset

```
Final_dataset %>% filter_all(any_vars(is.na(.)))
```

```
##        Date Births Marriages Deaths Stillbirths  Covid
## 1 2021-07-1     NA        NA     NA          NA  15968
## 2 2021-08-1     NA        NA     NA          NA  67913
## 3 2021-09-1     NA        NA     NA          NA 125560
## 4 2021-10-1     NA        NA     NA          NA  91834
## 5 2021-11-1     NA        NA     NA          NA  75935
## 6 2021-12-1     NA        NA     NA          NA 395815
```

Remove rows with NA as they will skew the results

```
Final_dataset <- na.omit(Final_dataset)
```

Check for NA's to confirm all have been removed

```
Final_dataset %>% filter_all(any_vars(is.na(.)))
```

```
## [1] Date        Births      Marriages   Deaths      Stillbirths Covid
## <0 rows> (or 0-length row.names)
```

```
summary(Final_dataset)
```

```
##      Date               Births        Marriages         Deaths
##  Length:330         Min.   :10020   Min.   :  597   Min.   : 5926
##  Class :character   1st Qu.:11260   1st Qu.: 2596   1st Qu.: 6706
##  Mode  :character   Median :11818   Median : 3559   Median : 7326
##                     Mean   :11763   Mean   : 5085   Mean   : 7500
##                     3rd Qu.:12288   3rd Qu.: 7627   3rd Qu.: 8094
##                     Max.   :13398   Max.   :11532   Max.   :11390
##   Stillbirths         Covid
##  Min.   :  0.00   Min.   :     0
##  1st Qu.: 73.00   1st Qu.:     0
##  Median : 90.50   Median :     0
##  Mean   : 90.74   Mean   :  4289
##  3rd Qu.:114.00   3rd Qu.:     0
##  Max.   :156.00   Max.   :237308
```

```
Final_dataset$Date <- as.Date(Final_dataset$Date,"%Y-%m-%d")
```

```
glimpse(Final_dataset)
```

```
## Rows: 330
## Columns: 6
## $ Date        <date> 1994-01-01, 1994-02-01, 1994-03-01, 1994-04-01, 1994-05-0~
## $ Births      <int> 11631, 11254, 13003, 12576, 13240, 13072, 13045, 12982, 12~
## $ Marriages   <int> 2078, 2650, 2557, 3967, 6493, 7754, 9264, 9194, 8540, 7400~
## $ Deaths      <int> 8094, 6428, 6503, 6224, 6483, 6187, 6196, 5926, 6062, 6515~
## $ Stillbirths <int> 75, 62, 73, 74, 67, 66, 70, 79, 60, 59, 56, 43, 78, 84, 75~
## $ Covid       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
```

Normalize the Data

```
Final_dataset_standardized <- Final_dataset %>% mutate_each_(list(~scale(.) %>% as.vector),
                                  vars = c("Births","Marriages","Deaths","Stillbirths","Covid"))
```

```
## Warning: `mutate_each_()` was deprecated in dplyr 0.7.0.
## Please use `across()` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.
```

```
Final_ts <- ts(Final_dataset)
```

Running Granger Causality tests for each variable against each other variable one by one Previous Granger test showed one variable against all others did not show causality, so now determining if one variable against each other variable individually shows causality.

Covid vs. Deaths Causality

```
grangertest(Covid ~ Deaths, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Covid ~ Lags(Covid, 1:12) + Lags(Deaths, 1:12)
```

```
## Model 2: Covid ~ Lags(Covid, 1:12)
##   Res.Df  Df      F  Pr(>F)
## 1    293
## 2    305 -12 1.6482 0.07796 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Covid vs. Marriages Causality

```
grangertest(Covid ~ Marriages, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Covid ~ Lags(Covid, 1:12) + Lags(Marriages, 1:12)
## Model 2: Covid ~ Lags(Covid, 1:12)
##   Res.Df  Df     F  Pr(>F)
## 1    293
## 2    305 -12 1.803 0.04714 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Covid vs. Stillbirths Causality

```
grangertest(Covid ~ Stillbirths, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Covid ~ Lags(Covid, 1:12) + Lags(Stillbirths, 1:12)
## Model 2: Covid ~ Lags(Covid, 1:12)
##   Res.Df  Df      F  Pr(>F)
## 1    293
## 2    305 -12 1.6179 0.08576 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Covid vs. Births Causality

```
grangertest(Covid ~ Births, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Covid ~ Lags(Covid, 1:12) + Lags(Births, 1:12)
## Model 2: Covid ~ Lags(Covid, 1:12)
##   Res.Df  Df      F Pr(>F)
## 1    293
## 2    305 -12 0.7444 0.7073
```

Deaths vs. Marriages Causality

```
grangertest(Deaths ~ Marriages, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Deaths ~ Lags(Deaths, 1:12) + Lags(Marriages, 1:12)
## Model 2: Deaths ~ Lags(Deaths, 1:12)
##   Res.Df  Df      F   Pr(>F)
## 1    293
## 2    305 -12 6.7127 1.16e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Deaths vs. Stillbirths Causality

```
grangertest(Deaths ~ Stillbirths, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Deaths ~ Lags(Deaths, 1:12) + Lags(Stillbirths, 1:12)
## Model 2: Deaths ~ Lags(Deaths, 1:12)
##   Res.Df  Df      F  Pr(>F)
## 1    293
## 2    305 -12 2.1829 0.01257 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Deaths vs. Births Causality

```
grangertest(Deaths ~ Births, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Deaths ~ Lags(Deaths, 1:12) + Lags(Births, 1:12)
## Model 2: Deaths ~ Lags(Deaths, 1:12)
##   Res.Df  Df      F    Pr(>F)
## 1    293
## 2    305 -12 7.4423 5.739e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Births vs. Marriages Causality

```
grangertest(Births ~ Marriages, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Births ~ Lags(Births, 1:12) + Lags(Marriages, 1:12)
## Model 2: Births ~ Lags(Births, 1:12)
##   Res.Df  Df      F    Pr(>F)
## 1    293
## 2    305 -12 9.9338 2.624e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Births vs. Stillbirths Causality

```
grangertest(Births ~ Stillbirths, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Births ~ Lags(Births, 1:12) + Lags(Stillbirths, 1:12)
## Model 2: Births ~ Lags(Births, 1:12)
##   Res.Df  Df      F   Pr(>F)
## 1    293
## 2    305 -12 2.7585 0.001439 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Marriages vs. Stillbirths Causality

```
grangertest(Marriages ~ Stillbirths, order =12, data = Final_ts)
```

```
## Granger causality test
##
## Model 1: Marriages ~ Lags(Marriages, 1:12) + Lags(Stillbirths, 1:12)
## Model 2: Marriages ~ Lags(Marriages, 1:12)
##   Res.Df  Df      F   Pr(>F)
## 1    293
## 2    305 -12 2.9507 0.000677 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```