County\_Acadia.r

stareq

2021-01-21

library(urca)  
library(forecast)

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.4 v dplyr 1.0.2  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.0

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

#install.packages("vars")  
library(vars)

## Loading required package: MASS

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

## Loading required package: strucchange

## Loading required package: zoo

##   
## Attaching package: 'zoo'

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

## Loading required package: sandwich

##   
## Attaching package: 'strucchange'

## The following object is masked from 'package:stringr':  
##   
## boundary

## Loading required package: lmtest

library(lmtest)  
  
# Loading the Dataset  
  
#data\_frame = read\_csv("https://raw.githubusercontent.com/QinLab/COVID19\_transmission\_MS/main/county3RWM/Acadia%2C%20Louisiana%2C%20US.csv?token=APMEIZDYFM5QZG4AMN5DE4K72HT6O")  
  
data\_frame <- read\_csv("Data\_Covid.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## date = col\_character(),  
## dailyCases = col\_double(),  
## air\_temp = col\_double(),  
## dew\_point = col\_double(),  
## RH = col\_double(),  
## Rt = col\_double(),  
## AppleDriving = col\_double(),  
## GoogleWorkplace = col\_double(),  
## GoogleResidential = col\_double()  
## )

data\_frame[is.na(data\_frame)] = 0  
data\_frame

## # A tibble: 255 x 9  
## date dailyCases air\_temp dew\_point RH Rt AppleDriving GoogleWorkplace  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 15-0~ 0 8.33 1.87 65.3 0 144. 4  
## 2 16-0~ 0 15.0 12.0 83.0 0 114. -3  
## 3 17-0~ 0 15.1 14.2 94.6 0 113. 1  
## 4 18-0~ 0 20.8 19.1 90.4 0 121. 4  
## 5 19-0~ 0 13.1 11.9 92.5 0 137. 0  
## 6 20-0~ 0 9.24 8.09 92.6 0 135. 0  
## 7 21-0~ 0 5.84 -0.888 64.6 0 197. 1  
## 8 22-0~ 0 5.17 -1.64 65.5 0 164. 4  
## 9 23-0~ 0 10.6 6.39 77.1 0 207. -2  
## 10 24-0~ 0 15.4 12.9 85.8 0 154. -15  
## # ... with 245 more rows, and 1 more variable: GoogleResidential <dbl>

# DailyCases were converted into time series data. Since dates were not mentioned in the Date format (shown as string)  
# Thus general index were used (1 to 248)  
# Declaring time series objects  
dailyCases\_ts = ts(data\_frame$dailyCases, start = 1, end=248, frequency = 1)  
air\_temp\_ts = ts(data\_frame$air\_temp, start = 1, end=248, frequency = 1)  
RH\_ts = ts(data\_frame$RH, start = 1, end=248, frequency = 1)  
Rt\_ts = ts(data\_frame$Rt, start = 1, end=248, frequency = 1)  
apple\_driving\_ts = ts(data\_frame$AppleDriving, start = 1, end=248, frequency = 1)  
Google\_workplace\_ts = ts(data\_frame$GoogleWorkplace, start = 1, end=248, frequency = 1)  
Google\_residential\_ts = ts(data\_frame$GoogleResidential, start = 1, end=248, frequency = 1)  
  
  
# Bind into a system   
system1 = cbind(dailyCases\_ts, Rt\_ts, air\_temp\_ts)  
system2 = cbind(dailyCases\_ts, RH\_ts, apple\_driving\_ts)  
system3 = cbind(dailyCases\_ts, Google\_workplace\_ts, Google\_residential\_ts)  
  
# lag selection using VAR technique  
lagselect1 = VARselect(system1, lag.max = 7, type = "const")  
#lagselect1$selection  
K\_val1 =strtoi(names(sort(summary(as.factor(lagselect1$selection)), decreasing=T)[1]))-1  
  
if (K\_val1<2){  
 K\_val1=2  
}  
lagselect2 = VARselect(system2, lag.max = 7, type = "const")  
#lagselect2$selection  
K\_val2 =strtoi(names(sort(summary(as.factor(lagselect2$selection)), decreasing=T)[1]))-1  
  
if (K\_val2<2){  
 K\_val2=2  
}  
  
lagselect3 = VARselect(system3, lag.max = 7, type = "const")  
#lagselect3$selection  
K\_val3 =strtoi(names(sort(summary(as.factor(lagselect3$selection)), decreasing=T)[1]))-1  
  
if (K\_val3<2){  
 K\_val3=2  
}  
  
   
  
# Johansen testing(trace)  
  
test1 = ca.jo(system1, type='trace', ecdet = 'const', K=K\_val1)  
test2 = ca.jo(system2, type='trace', ecdet = 'const', K=K\_val2)  
test3 = ca.jo(system3, type='trace', ecdet = 'const', K=K\_val3)  
  
  
summary(test1)

##   
## ######################   
## # Johansen-Procedure #   
## ######################   
##   
## Test type: trace statistic , without linear trend and constant in cointegration   
##   
## Eigenvalues (lambda):  
## [1] 1.805645e-01 7.491011e-02 1.917411e-02 -2.543689e-308  
##   
## Values of teststatistic and critical values of test:  
##   
## test 10pct 5pct 1pct  
## r <= 2 | 4.76 7.52 9.24 12.97  
## r <= 1 | 23.92 17.85 19.96 24.60  
## r = 0 | 72.91 32.00 34.91 41.07  
##   
## Eigenvectors, normalised to first column:  
## (These are the cointegration relations)  
##   
## dailyCases\_ts.l2 Rt\_ts.l2 air\_temp\_ts.l2 constant  
## dailyCases\_ts.l2 1.000000 1.00000 1.000000 1.000000  
## Rt\_ts.l2 9.211152 -49.94739 -356.541761 828.080675  
## air\_temp\_ts.l2 -3.110456 13.14016 4.914035 -3.021227  
## constant 49.012173 -274.78631 282.399152 -164.033669  
##   
## Weights W:  
## (This is the loading matrix)  
##   
## dailyCases\_ts.l2 Rt\_ts.l2 air\_temp\_ts.l2 constant  
## dailyCases\_ts.d -0.3925821122 -3.068507e-02 -4.581181e-03 1.260548e-17  
## Rt\_ts.d -0.0002729301 -4.053662e-05 3.553009e-05 3.876761e-23  
## air\_temp\_ts.d 0.0267354806 -7.791316e-03 -4.606410e-04 -2.052369e-18

summary(test2)

##   
## ######################   
## # Johansen-Procedure #   
## ######################   
##   
## Test type: trace statistic , without linear trend and constant in cointegration   
##   
## Eigenvalues (lambda):  
## [1] 2.649299e-01 1.139255e-01 6.869784e-02 5.551115e-17  
##   
## Values of teststatistic and critical values of test:  
##   
## test 10pct 5pct 1pct  
## r <= 2 | 17.51 7.52 9.24 12.97  
## r <= 1 | 47.26 17.85 19.96 24.60  
## r = 0 | 122.98 32.00 34.91 41.07  
##   
## Eigenvectors, normalised to first column:  
## (These are the cointegration relations)  
##   
## dailyCases\_ts.l2 RH\_ts.l2 apple\_driving\_ts.l2  
## dailyCases\_ts.l2 1.0000000 1.00000000 1.000000  
## RH\_ts.l2 -8.9556211 0.18907898 6.123974  
## apple\_driving\_ts.l2 0.3847621 -0.09244073 3.778865  
## constant 621.6903882 -12.39140163 -1103.827747  
## constant  
## dailyCases\_ts.l2 1.000000  
## RH\_ts.l2 2.939054  
## apple\_driving\_ts.l2 1.394470  
## constant 2208.110923  
##   
## Weights W:  
## (This is the loading matrix)  
##   
## dailyCases\_ts.l2 RH\_ts.l2 apple\_driving\_ts.l2  
## dailyCases\_ts.d -0.03040052 -0.30292247 -0.006422801  
## RH\_ts.d 0.05456596 -0.01887693 -0.002105843  
## apple\_driving\_ts.d -0.09475738 0.24351208 -0.039664552  
## constant  
## dailyCases\_ts.d 6.893374e-19  
## RH\_ts.d 1.178182e-19  
## apple\_driving\_ts.d -7.215607e-19

summary(test3)

##   
## ######################   
## # Johansen-Procedure #   
## ######################   
##   
## Test type: trace statistic , without linear trend and constant in cointegration   
##   
## Eigenvalues (lambda):  
## [1] 9.792497e-02 4.007579e-02 2.756594e-02 4.163336e-17  
##   
## Values of teststatistic and critical values of test:  
##   
## test 10pct 5pct 1pct  
## r <= 2 | 6.76 7.52 9.24 12.97  
## r <= 1 | 16.66 17.85 19.96 24.60  
## r = 0 | 41.60 32.00 34.91 41.07  
##   
## Eigenvectors, normalised to first column:  
## (These are the cointegration relations)  
##   
## dailyCases\_ts.l6 Google\_workplace\_ts.l6  
## dailyCases\_ts.l6 1.0000 1.000000  
## Google\_workplace\_ts.l6 -118.3593 1.463141  
## Google\_residential\_ts.l6 -317.5302 -5.856418  
## constant -1110.0112 43.927921  
## Google\_residential\_ts.l6 constant  
## dailyCases\_ts.l6 1.000000 1.0000000  
## Google\_workplace\_ts.l6 -0.188060 -2.3164169  
## Google\_residential\_ts.l6 1.804158 -0.6564643  
## constant -25.694469 313.4928351  
##   
## Weights W:  
## (This is the loading matrix)  
##   
## dailyCases\_ts.l6 Google\_workplace\_ts.l6  
## dailyCases\_ts.d -0.0001196841 -0.05311015  
## Google\_workplace\_ts.d 0.0023961852 -0.03057509  
## Google\_residential\_ts.d 0.0003040770 0.02258053  
## Google\_residential\_ts.l6 constant  
## dailyCases\_ts.d -0.12600019 1.187801e-18  
## Google\_workplace\_ts.d 0.01638121 6.280028e-19  
## Google\_residential\_ts.d -0.01814482 -3.276985e-18

# Discussion  
# Null hypothesis (r = 0) is rejected as 5 pct value (19.96) < test value ( 39.87). Thus at least one co-integration relationship exist between daily cases and relative humidity. In the second hypothesis (r=1),value of 5 pct (9.24) > test value(6.48). Therefore there exist at most one integration relationship.  
  
# Johansen testing(MaxEigen)  
print("------Eigen value testing----")

## [1] "------Eigen value testing----"

test2ei = ca.jo(system1, type='eigen', ecdet = 'const', K=K\_val2)  
test3ei = ca.jo(system1, type='eigen', ecdet = 'const', K=K\_val3)  
test1ei = ca.jo(system1, type='eigen', ecdet = 'const', K=K\_val1)  
summary(test1ei)

##   
## ######################   
## # Johansen-Procedure #   
## ######################   
##   
## Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegration   
##   
## Eigenvalues (lambda):  
## [1] 1.805645e-01 7.491011e-02 1.917411e-02 -2.543689e-308  
##   
## Values of teststatistic and critical values of test:  
##   
## test 10pct 5pct 1pct  
## r <= 2 | 4.76 7.52 9.24 12.97  
## r <= 1 | 19.15 13.75 15.67 20.20  
## r = 0 | 48.99 19.77 22.00 26.81  
##   
## Eigenvectors, normalised to first column:  
## (These are the cointegration relations)  
##   
## dailyCases\_ts.l2 Rt\_ts.l2 air\_temp\_ts.l2 constant  
## dailyCases\_ts.l2 1.000000 1.00000 1.000000 1.000000  
## Rt\_ts.l2 9.211152 -49.94739 -356.541761 828.080675  
## air\_temp\_ts.l2 -3.110456 13.14016 4.914035 -3.021227  
## constant 49.012173 -274.78631 282.399152 -164.033669  
##   
## Weights W:  
## (This is the loading matrix)  
##   
## dailyCases\_ts.l2 Rt\_ts.l2 air\_temp\_ts.l2 constant  
## dailyCases\_ts.d -0.3925821122 -3.068507e-02 -4.581181e-03 1.260548e-17  
## Rt\_ts.d -0.0002729301 -4.053662e-05 3.553009e-05 3.876761e-23  
## air\_temp\_ts.d 0.0267354806 -7.791316e-03 -4.606410e-04 -2.052369e-18

summary(test2ei)

##   
## ######################   
## # Johansen-Procedure #   
## ######################   
##   
## Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegration   
##   
## Eigenvalues (lambda):  
## [1] 1.805645e-01 7.491011e-02 1.917411e-02 -2.543689e-308  
##   
## Values of teststatistic and critical values of test:  
##   
## test 10pct 5pct 1pct  
## r <= 2 | 4.76 7.52 9.24 12.97  
## r <= 1 | 19.15 13.75 15.67 20.20  
## r = 0 | 48.99 19.77 22.00 26.81  
##   
## Eigenvectors, normalised to first column:  
## (These are the cointegration relations)  
##   
## dailyCases\_ts.l2 Rt\_ts.l2 air\_temp\_ts.l2 constant  
## dailyCases\_ts.l2 1.000000 1.00000 1.000000 1.000000  
## Rt\_ts.l2 9.211152 -49.94739 -356.541761 828.080675  
## air\_temp\_ts.l2 -3.110456 13.14016 4.914035 -3.021227  
## constant 49.012173 -274.78631 282.399152 -164.033669  
##   
## Weights W:  
## (This is the loading matrix)  
##   
## dailyCases\_ts.l2 Rt\_ts.l2 air\_temp\_ts.l2 constant  
## dailyCases\_ts.d -0.3925821122 -3.068507e-02 -4.581181e-03 1.260548e-17  
## Rt\_ts.d -0.0002729301 -4.053662e-05 3.553009e-05 3.876761e-23  
## air\_temp\_ts.d 0.0267354806 -7.791316e-03 -4.606410e-04 -2.052369e-18

summary(test3ei)

##   
## ######################   
## # Johansen-Procedure #   
## ######################   
##   
## Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegration   
##   
## Eigenvalues (lambda):  
## [1] 7.269660e-02 3.811648e-02 2.127146e-02 2.775558e-17  
##   
## Values of teststatistic and critical values of test:  
##   
## test 10pct 5pct 1pct  
## r <= 2 | 5.20 7.52 9.24 12.97  
## r <= 1 | 9.40 13.75 15.67 20.20  
## r = 0 | 18.26 19.77 22.00 26.81  
##   
## Eigenvectors, normalised to first column:  
## (These are the cointegration relations)  
##   
## dailyCases\_ts.l6 Rt\_ts.l6 air\_temp\_ts.l6 constant  
## dailyCases\_ts.l6 1.00000 1.00000 1.000000 1.0000000  
## Rt\_ts.l6 12.52776 16.47995 -128.362767 103.3618012  
## air\_temp\_ts.l6 -3.17398 25.79796 5.537934 -0.7303494  
## constant 46.01867 -650.27059 -16.037272 21.4940411  
##   
## Weights W:  
## (This is the loading matrix)  
##   
## dailyCases\_ts.l6 Rt\_ts.l6 air\_temp\_ts.l6 constant  
## dailyCases\_ts.d -0.2445691953 -0.0049250474 -3.057831e-02 3.172907e-18  
## Rt\_ts.d -0.0005626546 -0.0000242899 9.943945e-05 1.222135e-19  
## air\_temp\_ts.d 0.0197182190 -0.0025844305 -1.630152e-03 4.222726e-18

# Discussion  
# Null hypothesis (r = 0) is rejected as 5 pct value (15.67) < test value ( 33.39). Thus at least one co-integration relationship exist between daily cases and relative humidity. In the second hypothesis (r=1),value of 5 pct (9.24) > test value(6.48). Therefore there exist at most one co-integration relationship.