Financial Data Analysis Tutorial

Contents

Setup

We load the required libraries

```
rm(list=ls())
require(tinytex) #LaTeX
require(ggplot2) #plots
require(AEC) #JP-Renne functions
require(forecast) #time series stuff
require(expm) #matrix exponents
require(here) #directory finder
require(stringr) # analysis of strings, important for the detection in tweets
require(dplyr) #data management
require(lubridate) #data dates management
require(zoo) #for lagging
require(jtools) #tables
require(huxtable) #tables
require(lmtest) #reg tests
require(vroom) #for loading data

getwd()
#setwd("...") -> set wd at base repo folder

#load helper functions
source(here("helperfunctions/data_loaders.R"))
source(here("helperfunctions/date_selector.R"))
source(here("helperfunctions/plotters.R"))
source(here("helperfunctions/quick_arma.R"))
source(here("helperfunctions/r.vol_calculators.R"))
```

Data

Note that this document uses intermediate data to show how we produce the financial part of the final data. It also shows various functions we developed to implement some quick analysis and plotting.

Raw Data

Use our custom data loading function as our financial data is fractured by month due to the downloading constraints.

```
#market prices (loads and names them automatically)

#S&P500
data_loader(symbol="SPY")

#STOXX50
data_loader(symbol="VGK")

#CSI 300 (China)
data_loader(symbol="ASHR")
```

Quick Analysis

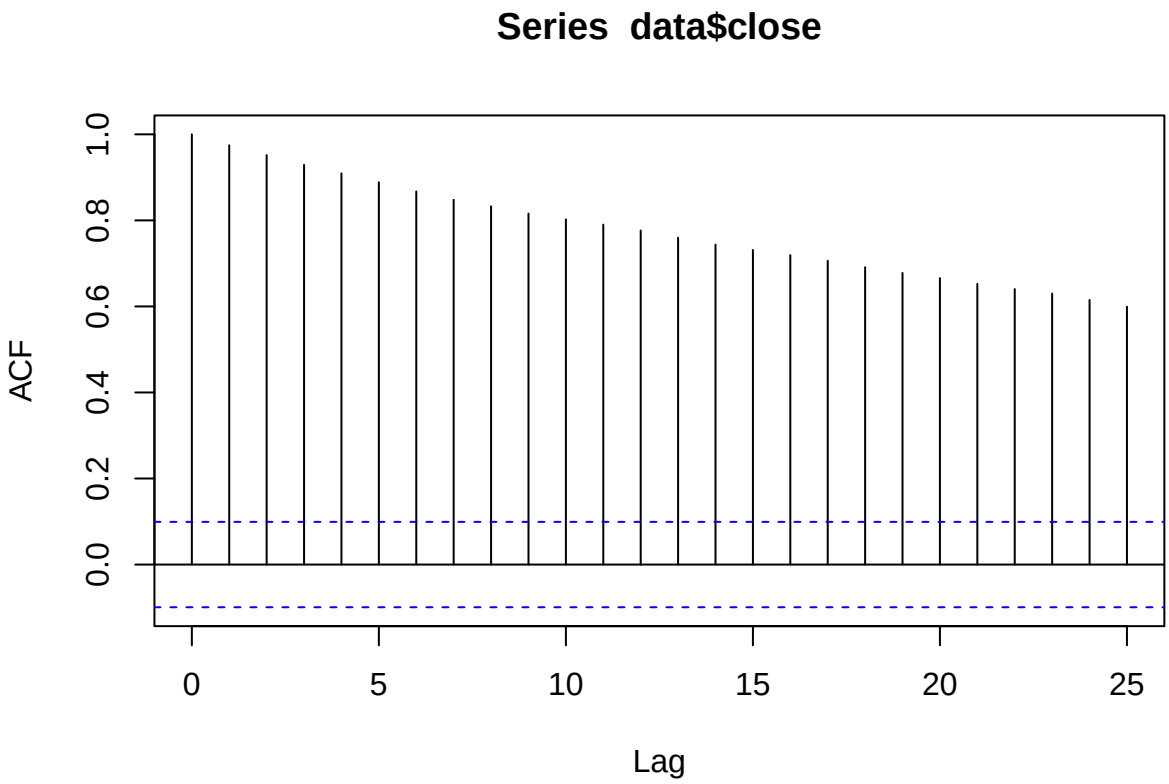
SPY April 2nd 2025

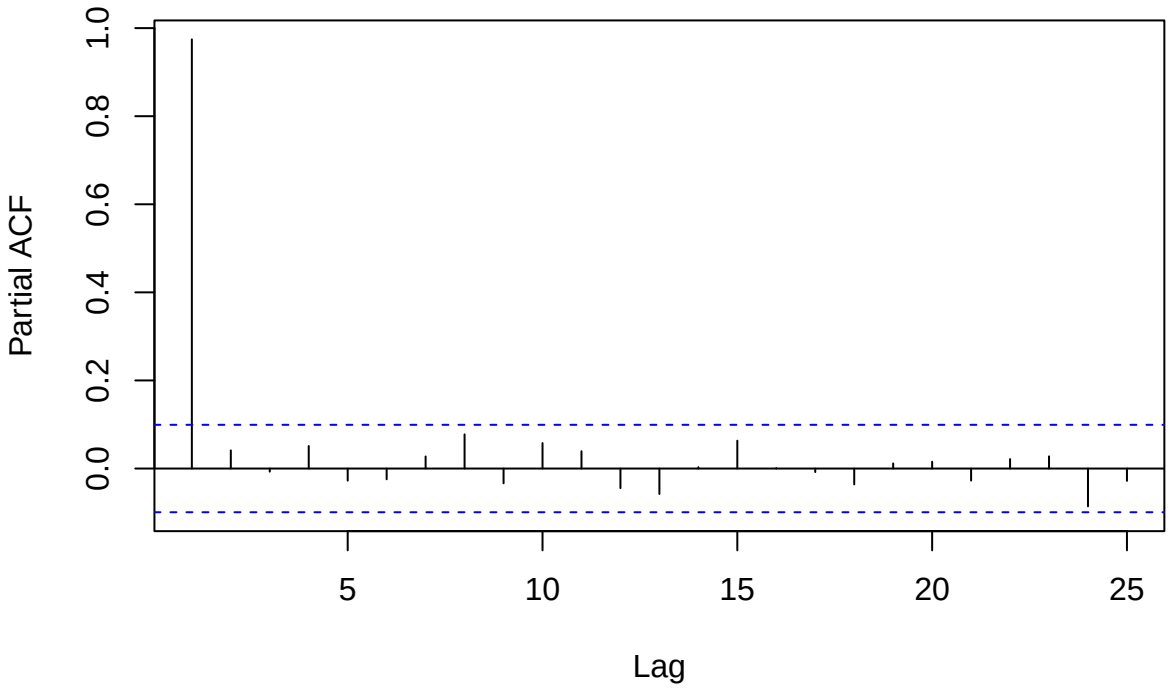
Here are some functions we developed in order to be able to easily select certain dates for our financial data. Comes in handy for our plotter functions as well as later in the calculation of the realised volatility.

```
#extract a particular day
SPY_25_04_02 = day_selector(raw_SPY,2025,04,02) #april 2nd 2025

#let's plot it
price_plotter_day(SPY_25_04_02,"SPY Price on April 2nd 2025")
```


```
#quickly test some ARMA specifications
quick_arma(SPY_25_04_02,1,0,0) #checking AR1,AR2,AR3
```


Series data$close
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
Lag




```
##                                AR Estimations
##
##                                  AR-1        AR-2         AR-3
##
##                  ar1            0.9975      0.9728       1.4609
##                                 (0.0030)    (0.0514)   (NaN)
##                  intercept    561.0971    561.3655     562.5635
##                                 (3.2897)    (3.4352)    (22.1897)
##                  ar2                         0.0249       0.0770
##                                              (0.0515)     (0.0013)
##                  ar3                                      -0.5386
##                                                           (0.0007)
##
##                  nobs         390         390          390
##                  sigma          0.2854      0.2853       0.3414
##                  logLik       -67.0847    -66.9808     -135.4359
##                  AIC          140.1693    141.9615     280.8718
##                  BIC          152.0678    157.8261     300.7025
##                  nobs.1       390.0000    390.0000     390.0000
##
##                  *** p < 0.001; ** p < 0.01; * p <
##                  0.05.
##
## Column names: names, AR-1, AR-2, AR-3
##                                Checking Residuals
##
```

```
##                         AR-1 Residuals   AR-2 Residuals   AR-3 Residuals
##
##        (Intercept)            0.0302 *         0.0291 *      -0.0051
##                              (0.0145)         (0.0145)       (0.0171)
##        REG1res_lagged        -0.0476
##                              (0.0510)
##        REG2res_lagged                         -0.0217
##                                               (0.0511)
##        REG3res_lagged                                        -0.1733 ***
##                                                              (0.0503)
##
##        N                    389              389            389
##        R2                     0.0022           0.0005         0.0297
##
##        *** p < 0.001; ** p < 0.01; * p < 0.05.
##
## Column names: names, AR-1 Residuals, AR-2 Residuals, AR-3 Residuals
```

```
#quick_arma(SPY_25_04_02,2,0,0) #checking AR2,AR3,AR4


#extract a particular month
SPY_24_09 = month_selector(raw_SPY,2024,09) #november 2024


#extract a particular year
SPY_24 = year_selector(raw_SPY,2024) #2024
```

Realised Volatility

Computations

We have developed several functions for the computation of realised volatility. The only areas where they differ is in the timeframe of the input dataset, and whether it calculates the daily average or the hourly average. Note that we use hourly data in our final analysis.
These functions can be found in the helperfunctions folder.

```
#avg per day for each month of any dataset
vol_SPY_daily = r.vol_daily(raw_SPY,merge=F)
head(vol_SPY_daily)
```

timestamp	r__vol__d
2013-01-02	0.00197
2013-01-03	0.00178
2013-01-04	0.0012
2013-01-07	0.000854
2013-01-08	0.00115
2013-01-09	0.000924

```
#can then filter out years, months, or days
vol_24d = year_selector(vol_SPY_daily,2024)
vol_24_08d = month_selector(vol_SPY_daily,2024,08)
vol_24_11_04d = day_selector(vol_SPY_daily,2024,11,04) #scalar


#avg per hour for each day of each month of any dataset
vol_SPY_hourly = r.vol_hourly(raw_SPY,merge=F)
head(vol_SPY_hourly)
```

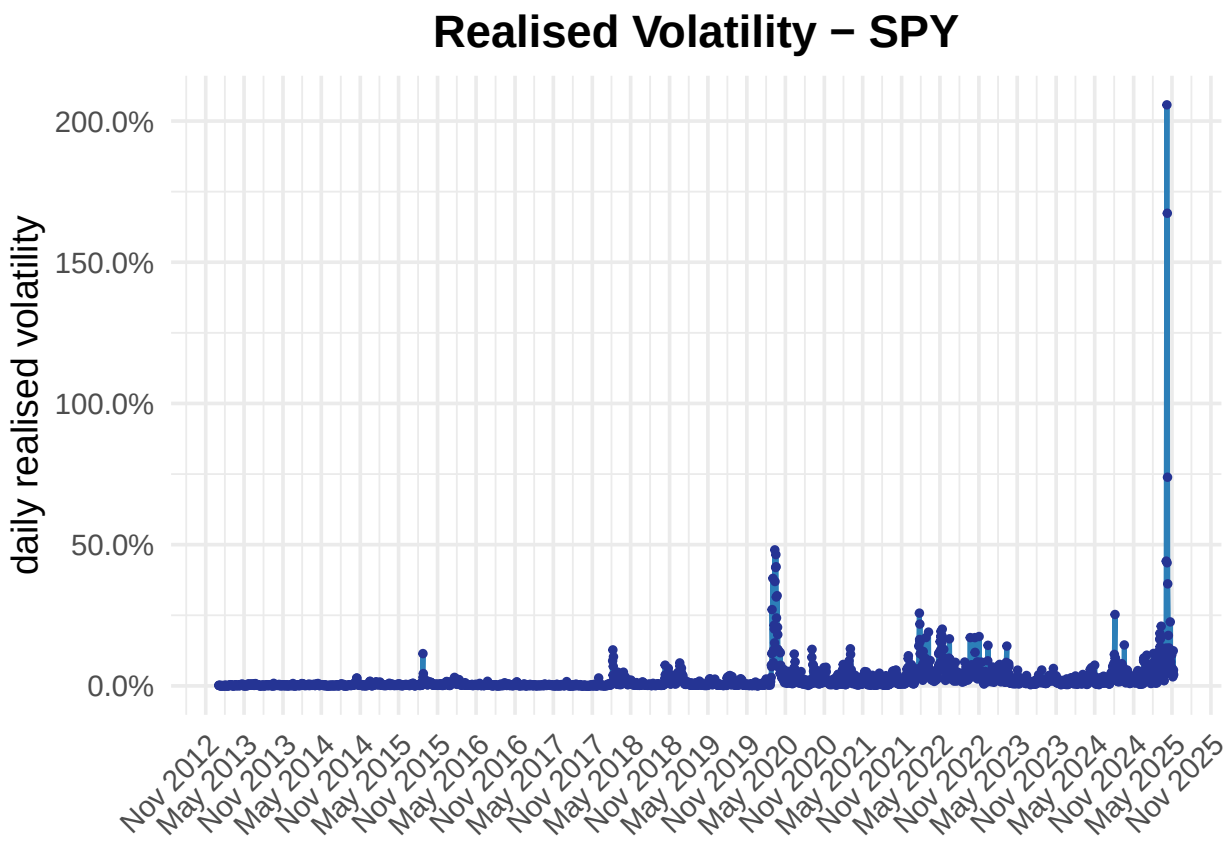
timestamp	r__vol__h
2013-01-02 09:00:00	0.00403
2013-01-02 10:00:00	0.00251
2013-01-02 11:00:00	0.00183
2013-01-02 12:00:00	0.00109
2013-01-02 13:00:00	0.001
2013-01-02 14:00:00	0.00149

```
#can then filter out years, months, or days
vol_24h = year_selector(vol_SPY_hourly,2024)
vol_24_08h = month_selector(vol_SPY_hourly,2024,08)
vol_24_11_04h = day_selector(vol_SPY_hourly,2024,11,04) #vector
```

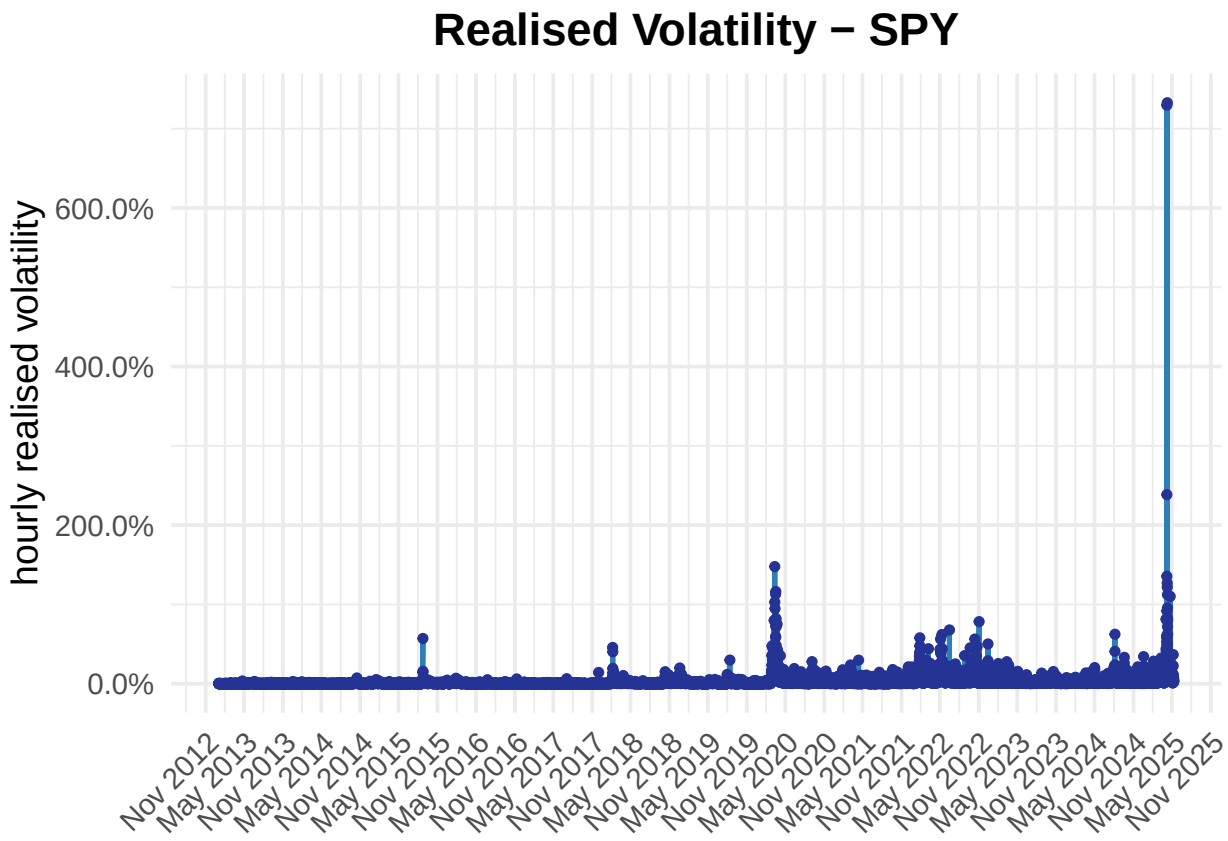
Plots

We have developed two functions in order to easily plot our volatility data in nice ggplots. One is for the daily average and the other for the hourly average.

```
#avg per day volatility all time
dvol_plotter(vol_SPY_daily,breaks="yearly",
             title="Realised Volatility - SPY")
```




```
#hourly volatility all time
hvol_plotter(vol_SPY_hourly,breaks="yearly",
             title="Realised Volatility - SPY")
```


Realised Volatility − SPY
hourly realised volatility
600.0%
400.0%
200.0%
0.0%
Nov 2012
May 2013
Nov 2013
May 2014
Nov 2014
May 2015
Nov 2015
May 2016
Nov 2016
May 2017
Nov 2017
May 2018
Nov 2018
May 2019
Nov 2019
May 2020
Nov 2020
May 2021
Nov 2021
May 2022
Nov 2022
May 2023
Nov 2023
May 2024
Nov 2024
May 2025
Nov 2025