midterm.R

Dipro

2021-04-01

```
library(quantmod)
```

```
## Warning: package 'quantmod' was built under R version 4.0.4
```

```
## Loading required package: xts
```

```
## Warning: package 'xts' was built under R version 4.0.4
```

```
## Loading required package: zoo
```

```
## Warning: package 'zoo' was built under R version 4.0.4
```

```
##
## Attaching package: 'zoo'
```

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

```
## Loading required package: TTR
```

```
## Warning: package 'TTR' was built under R version 4.0.4
```

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

```
# Q 1.1

getSymbols("JPM")
```

```
## 'getSymbols' currently uses auto.assign=TRUE by default, but will
## use auto.assign=FALSE in 0.5-0. You will still be able to use
## 'loadSymbols' to automatically load data. getOption("getSymbols.env")
## and getOption("getSymbols.auto.assign") will still be checked for
## alternate defaults.
##
## This message is shown once per session and may be disabled by setting
## options("getSymbols.warning4.0"=FALSE). See ?getSymbols for details.
```

```
## [1] "JPM"
```

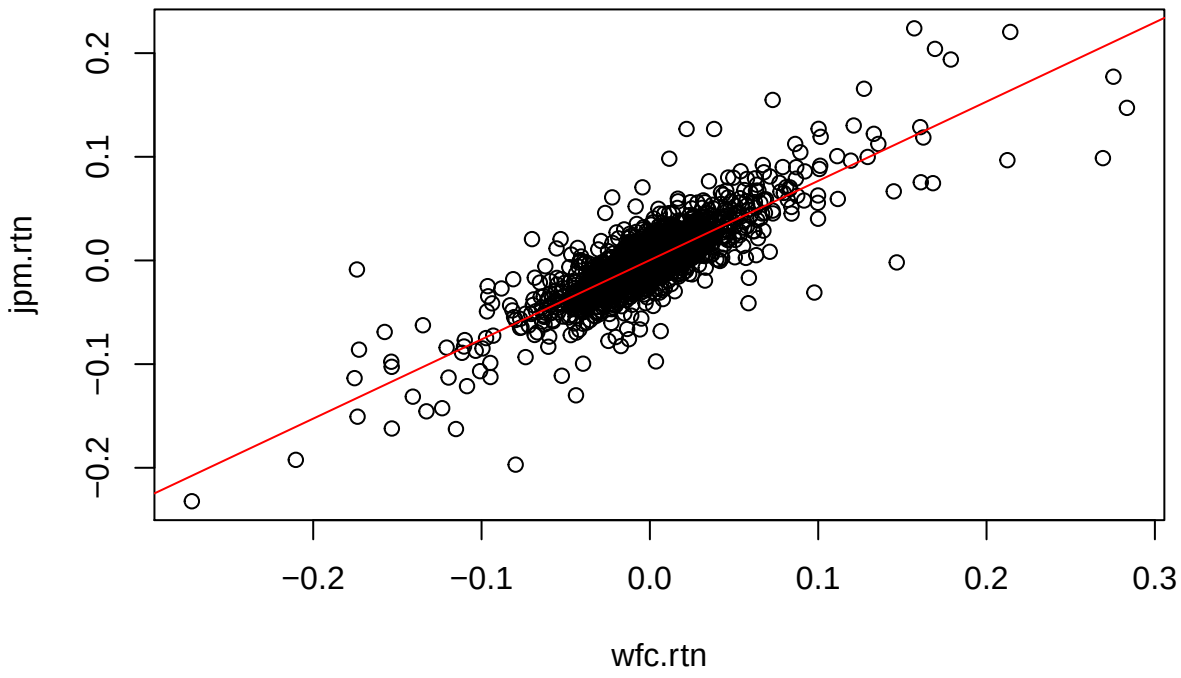
```
jpm <- data.frame(JPM)
#head(jpm)

getSymbols("WFC")
```

```
## [1] "WFC"
```

```r
wfc <- data.frame(WFC)
#head(wfc)
#tail(wfc)

jpm.price <- jpm$JPM.Adjusted
wfc.price <- wfc$WFC.Adjusted

jpm.rtn <- diff(log(jpm.price))
wfc.rtn <- diff(log(wfc.price))


#Q 1.2

lin.reg <- lm(jpm.rtn ~ wfc.rtn)
summary(lin.reg)
```

```
##
## Call:
## lm(formula = jpm.rtn ~ wfc.rtn)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.136326 -0.005699 -0.000127  0.005790  0.124016
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0003170  0.0002355   1.346    0.178
## wfc.rtn     0.7645407  0.0084388  90.598   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0141 on 3583 degrees of freedom
## Multiple R-squared:  0.6961, Adjusted R-squared:  0.696
## F-statistic:  8208 on 1 and 3583 DF,  p-value: < 2.2e-16
```

```r
#Q 1.3

plot(jpm.rtn ~ wfc.rtn)
abline(lin.reg , col="red")
```




```
#Q 1.4

cheese <- read.csv("cheese.csv")
#head(cheese)

cheese.model <- lm(taste ~ acetic + h2s + lactic, data = cheese) # full model
null.model <- lm(taste ~ 1, data = cheese) # model with no factor

summary(cheese.model)
```

```
##
## Call:
## lm(formula = taste ~ acetic + h2s + lactic, data = cheese)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.390  -6.612  -1.009   4.908  25.449
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -28.8768    19.7354  -1.463  0.15540
## acetic        0.3277     4.4598   0.073  0.94198
## h2s           3.9118     1.2484   3.133  0.00425 **
## lactic       19.6705     8.6291   2.280  0.03108 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 10.13 on 26 degrees of freedom
## Multiple R-squared:  0.6518, Adjusted R-squared:  0.6116
## F-statistic: 16.22 on 3 and 26 DF,  p-value: 3.81e-06
```

```
summary(null.model)
```

```
##
## Call:
## lm(formula = taste ~ 1, data = cheese)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -23.833 -10.983  -3.583  12.167  32.667
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   24.533      2.968   8.266  4.1e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.26 on 29 degrees of freedom
```

```
# Q 1.5

full.model.formula <- taste ~ acetic + h2s + lactic

# forward selection
step(null.model, full.model.formula, direction = "forward")
```

```
## Start:  AIC=168.29
## taste ~ 1
##
##          Df Sum of Sq    RSS    AIC
## + h2s     1    4376.7 3286.1 144.89
## + lactic  1    3800.4 3862.5 149.74
## + acetic  1    2314.1 5348.7 159.50
## <none>                7662.9 168.29
##
## Step:  AIC=144.89
## taste ~ h2s
##
##          Df Sum of Sq    RSS    AIC
## + lactic  1    617.18 2669.0 140.65
## <none>                3286.1 144.89
## + acetic  1     84.41 3201.7 146.11
##
## Step:  AIC=140.65
## taste ~ h2s + lactic
##
##          Df Sum of Sq    RSS    AIC
## <none>                2669.0 140.65
## + acetic  1   0.55427 2668.4 142.64
```

```
##
## Call:
```

```
## lm(formula = taste ~ h2s + lactic, data = cheese)
##
## Coefficients:
## (Intercept)          h2s       lactic
##     -27.592        3.946       19.887
```

```
#backward selection
step(cheese.model, full.model.formula, direction = "backward")
```

```
## Start:  AIC=142.64
## taste ~ acetic + h2s + lactic
##
##          Df Sum of Sq    RSS    AIC
## - acetic  1      0.55 2669.0 140.65
## <none>                2668.4 142.64
## - lactic  1    533.32 3201.7 146.11
## - h2s     1   1007.66 3676.1 150.25
##
## Step:  AIC=140.65
## taste ~ h2s + lactic
##
##          Df Sum of Sq    RSS    AIC
## <none>                2669.0 140.65
## - lactic  1    617.18 3286.1 144.89
## - h2s     1   1193.52 3862.5 149.74
```

```
##
## Call:
## lm(formula = taste ~ h2s + lactic, data = cheese)
##
## Coefficients:
## (Intercept)          h2s       lactic
##     -27.592        3.946       19.887
```

```
# both direction
step(null.model, full.model.formula, direction = "both")
```

```
## Start:  AIC=168.29
## taste ~ 1
##
##          Df Sum of Sq    RSS    AIC
## + h2s     1    4376.7 3286.1 144.89
## + lactic  1    3800.4 3862.5 149.74
## + acetic  1    2314.1 5348.7 159.50
## <none>                7662.9 168.29
##
## Step:  AIC=144.89
## taste ~ h2s
##
##          Df Sum of Sq    RSS    AIC
## + lactic  1     617.2 2669.0 140.65
## <none>                3286.1 144.89
## + acetic  1      84.4 3201.7 146.11
## - h2s     1    4376.7 7662.9 168.29
##
## Step:  AIC=140.65
```

```
## taste ~ h2s + lactic
##
##          Df Sum of Sq    RSS    AIC
## <none>                2669.0 140.65
## + acetic  1      0.55 2668.4 142.64
## - lactic  1    617.18 3286.1 144.89
## - h2s     1   1193.52 3862.5 149.74
```

```
##
## Call:
## lm(formula = taste ~ h2s + lactic, data = cheese)
##
## Coefficients:
## (Intercept)          h2s       lactic
##     -27.592        3.946       19.887
```

```
# All three selection methods give the same model: taste = c1 + c2*h2s + c3*lactic + e
# corresponding regression line is: taste = -27.592 + 3.946*h2s + 19.887 * lactic
```