FIN3080 Project2 Question1

119020002 Cai Hongyu, 119020071 Zhai Haotian, 119020084 Zou Yipeng

2021.04.10

```
library(readxl)
library(ggplot2)
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
```

```
library(zoo)
```

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

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

Problem 1

```
#import data and data organization
BM_ratio<- read_excel("   110433826/AF_Actual.xlsx")
BM_ratio<-BM_ratio[-(1:2),]
BM_ratio<-na.omit(BM_ratio)
BM_ratio$BM<-as.numeric(BM_ratio$BM)
BM_ratio$Ddate<-as.Date(BM_ratio$Ddate,format="%Y-%m-%d")
BM_ratio<-BM_ratio[-(31351:31484),]
```

```
#import market value and monthly return
CNwhole<-read_excel("  +   .xlsx")
CNwhole<-CNwhole[-(1:2),]
```

```
#data cleaning
CNwhole$Msmvosd<-as.numeric(CNwhole$Msmvosd)
CNwhole$Mretwd<-as.numeric(CNwhole$Mretwd)
```

```
#import three-factor value factor
FAMA<-read_excel("    + .xlsx")
FAMA<-with(FAMA,FAMA[FAMA$MarkettypeID=='P9709',])
FAMA<-FAMA[,-4]
FAMA$HML<-0
FAMA$HML1<-as.numeric(FAMA$HML1)
```

```
#function of Percentile Separation
ZTT<-function(x) {
  x[(1:ceiling((length(x$BM)/10)*3)),]
}
TTS<-function(x) {
  x[(ceiling((length(x$BM)/10)*3)+1):ceiling((length(x$BM)/10)*7),]
}
STE<-function(x) {
  x[(ceiling((length(x$BM)/10)*7)+1):length(x$BM),]
}
```

```
Weight<-function(x) {
  x/sum(x)
}
```

```
#select from 2006.1~2006.6 to L, M, H
S1<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2004-12-31",])
S1<-arrange(S1,S1$BM)
S1_L<-ZTT(S1)
S1_M<-TTS(S1)
S1_H<-STE(S1)
```

```
for ( i in c("2006-01","2006-02","2006-03","2006-04","2006-05","2006-06") ) {

S1_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S1_L_M<-merge(S1_L,S1_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S1_H_M<-merge(S1_H,S1_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S1_H_M$Mretwd,Weight(S1_H_M$Msmvosd))-weighted.mean(S1_L_

}
```

```
#select from 2006.7~2007.6 to L, M, H
S2<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2005-12-31",])
S2<-arrange(S2,S2$BM)
S2_L<-ZTT(S2)
S2_M<-TTS(S2)
S2_H<-STE(S2)
```

```
for ( i in c("2006-07","2006-08","2006-09","2006-10","2006-11","2006-12","2007-01","2007-02","2007-03",'

S2_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S2_L_M<-merge(S2_L,S2_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S2_H_M<-merge(S2_H,S2_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S2_H_M$Mretwd,Weight(S2_H_M$Msmvosd))-weighted.mean(S2_L_M

}

#select from 2007.7~2008.6 to L, M, H
S3<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2006-12-31",])
S3<-arrange(S3,S3$BM)
S3_L<-ZTT(S3)
S3_M<-TTS(S3)
S3_H<-STE(S3)

for ( i in c("2007-07","2007-08","2007-09","2007-10","2007-11","2007-12","2008-01","2008-02","2008-03",'

S3_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S3_L_M<-merge(S3_L,S3_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S3_H_M<-merge(S3_H,S3_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S3_H_M$Mretwd,Weight(S3_H_M$Msmvosd))-weighted.mean(S3_L_M

}

#select from 2008.7~2009.6 to L, M, H
S4<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2007-12-31",])
S4<-arrange(S4,S4$BM)
S4_L<-ZTT(S4)
S4_M<-TTS(S4)
S4_H<-STE(S4)

for ( i in c("2008-07","2008-08","2008-09","2008-10","2008-11","2008-12","2009-01","2009-02","2009-03",'

S4_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S4_L_M<-merge(S4_L,S4_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S4_H_M<-merge(S4_H,S4_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S4_H_M$Mretwd,Weight(S4_H_M$Msmvosd))-weighted.mean(S4_L_M
```

```
}

#select from 2009.7~2010.6 to L, M, H
S5<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2008-12-31",])
S5<-arrange(S5,S5$BM)
S5_L<-ZTT(S5)
S5_M<-TTS(S5)
S5_H<-STE(S5)

for ( i in c("2009-07","2009-08","2009-09","2009-10","2009-11","2009-12","2010-01","2010-02","2010-03",

S5_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S5_L_M<-merge(S5_L,S5_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S5_H_M<-merge(S5_H,S5_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S5_H_M$Mretwd,Weight(S5_H_M$Msmvosd))-weighted.mean(S5_L_

}

#select from 2010.7~2011.6 to L, M, H
S6<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2009-12-31",])
S6<-arrange(S6,S6$BM)
S6_L<-ZTT(S6)
S6_M<-TTS(S6)
S6_H<-STE(S6)

for ( i in c("2010-07","2010-08","2010-09","2010-10","2010-11","2010-12","2011-01","2011-02","2011-03",

S6_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S6_L_M<-merge(S6_L,S6_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S6_H_M<-merge(S6_H,S6_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S6_H_M$Mretwd,Weight(S6_H_M$Msmvosd))-weighted.mean(S6_L_

}

#select from 2011.7~2012.6 to L, M, H
S7<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2010-12-31",])
S7<-arrange(S7,S7$BM)
S7_L<-ZTT(S7)
S7_M<-TTS(S7)
S7_H<-STE(S7)
```

```
for ( i in c("2011-07","2011-08","2011-09","2011-10","2011-11","2011-12","2012-01","2012-02","2012-03",'

S7_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S7_L_M<-merge(S7_L,S7_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S7_H_M<-merge(S7_H,S7_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S7_H_M$Mretwd,Weight(S7_H_M$Msmvosd))-weighted.mean(S7_L_

}
```

```
#select from 2012.7~2013.6 to L, M, H
S8<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2011-12-31",])
S8<-arrange(S8,S8$BM)
S8_L<-ZTT(S8)
S8_M<-TTS(S8)
S8_H<-STE(S8)
```

```
for ( i in c("2012-07","2012-08","2012-09","2012-10","2012-11","2012-12","2013-01","2013-02","2013-03",'

S8_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S8_L_M<-merge(S8_L,S8_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S8_H_M<-merge(S8_H,S8_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S8_H_M$Mretwd,Weight(S8_H_M$Msmvosd))-weighted.mean(S8_L_

}
```

```
#select from 2013.7~2014.6 to L, M, H
S9<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2012-12-31",])
S9<-arrange(S9,S9$BM)
S9_L<-ZTT(S9)
S9_M<-TTS(S9)
S9_H<-STE(S9)
```

```
for ( i in c("2013-07","2013-08","2013-09","2013-10","2013-11","2013-12","2014-01","2014-02","2014-03",'

S9_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S9_L_M<-merge(S9_L,S9_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S9_H_M<-merge(S9_H,S9_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S9_H_M$Mretwd,Weight(S9_H_M$Msmvosd))-weighted.mean(S9_L_
```

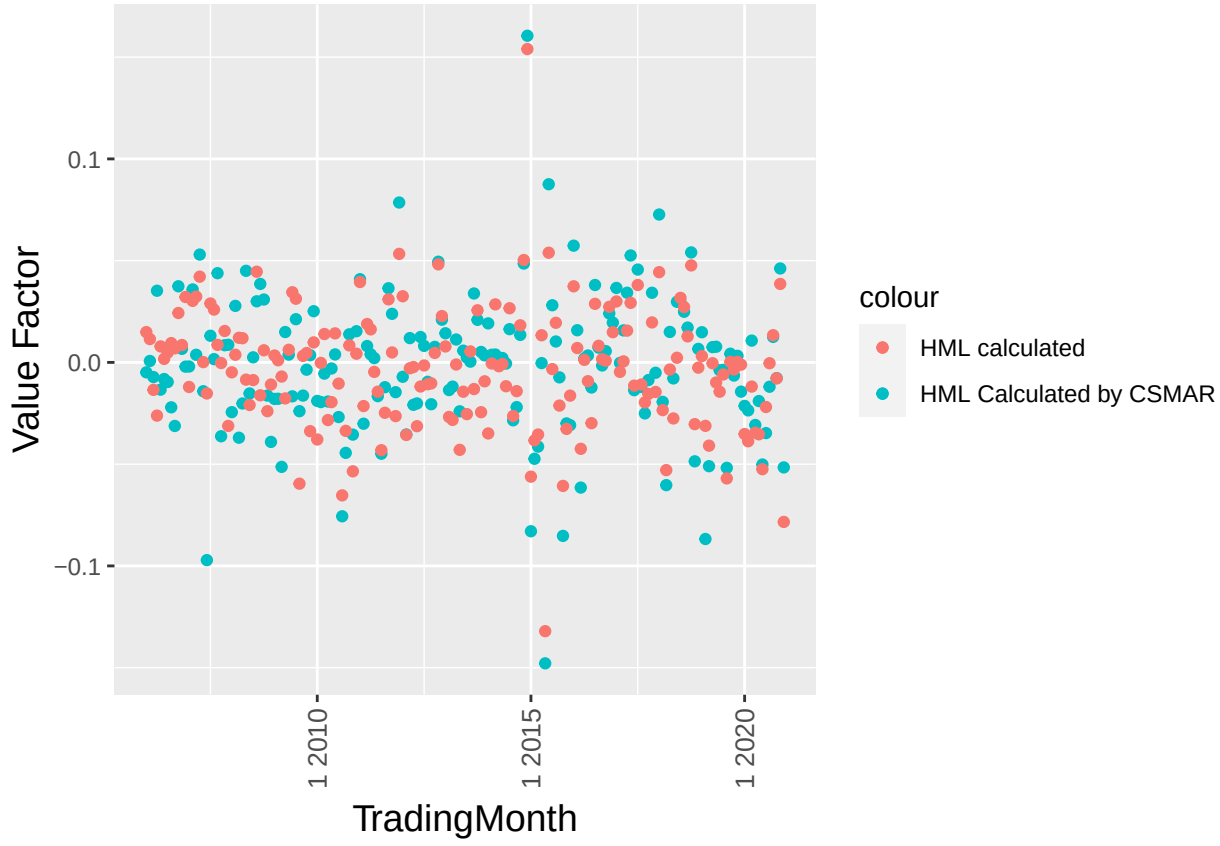
```
}

#select from 2014.7~2015.6 to L, M, H
S10<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2013-12-31",])
S10<-arrange(S10,S10$BM)
S10_L<-ZTT(S10)
S10_M<-TTS(S10)
S10_H<-STE(S10)

for ( i in c("2014-07","2014-08","2014-09","2014-10","2014-11","2014-12","2015-01","2015-02","2015-03",

S10_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S10_L_M<-merge(S10_L,S10_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S10_H_M<-merge(S10_H,S10_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S10_H_M$Mretwd,Weight(S10_H_M$Msmvosd))-weighted.mean(S10_

}

#select from 2015.7~2016.6 to L, M, H
S11<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2014-12-31",])
S11<-arrange(S11,S11$BM)
S11_L<-ZTT(S11)
S11_M<-TTS(S11)
S11_H<-STE(S11)

for ( i in c("2015-07","2015-08","2015-09","2015-10","2015-11","2015-12","2016-01","2016-02","2016-03",

S11_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S11_L_M<-merge(S11_L,S11_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S11_H_M<-merge(S11_H,S11_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S11_H_M$Mretwd,Weight(S11_H_M$Msmvosd))-weighted.mean(S11_

}

#select from 2016.7~2017.6 to L, M, H
S12<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2015-12-31",])
S12<-arrange(S12,S12$BM)
S12_L<-ZTT(S12)
S12_M<-TTS(S12)
S12_H<-STE(S12)
```

```
for ( i in c("2016-07","2016-08","2016-09","2016-10","2016-11","2016-12","2017-01","2017-02","2017-03",'

S12_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S12_L_M<-merge(S12_L,S12_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S12_H_M<-merge(S12_H,S12_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S12_H_M$Mretwd,Weight(S12_H_M$Msmvosd))-weighted.mean(S12_

}

#select from 2017.7~2018.6 to L, M, H
S13<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2016-12-31",])
S13<-arrange(S13,S13$BM)
S13_L<-ZTT(S13)
S13_M<-TTS(S13)
S13_H<-STE(S13)

for ( i in c("2017-07","2017-08","2017-09","2017-10","2017-11","2017-12","2018-01","2018-02","2018-03",'

S13_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S13_L_M<-merge(S13_L,S13_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S13_H_M<-merge(S13_H,S13_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S13_H_M$Mretwd,Weight(S13_H_M$Msmvosd))-weighted.mean(S13_

}

#select from 2018.7~2019.6 to L, M, H
S14<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2017-12-31",])
S14<-arrange(S14,S14$BM)
S14_L<-ZTT(S14)
S14_M<-TTS(S14)
S14_H<-STE(S14)

for ( i in c("2018-07","2018-08","2018-09","2018-10","2018-11","2018-12","2019-01","2019-02","2019-03",'

S14_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S14_L_M<-merge(S14_L,S14_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S14_H_M<-merge(S14_H,S14_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S14_H_M$Mretwd,Weight(S14_H_M$Msmvosd))-weighted.mean(S14_
```

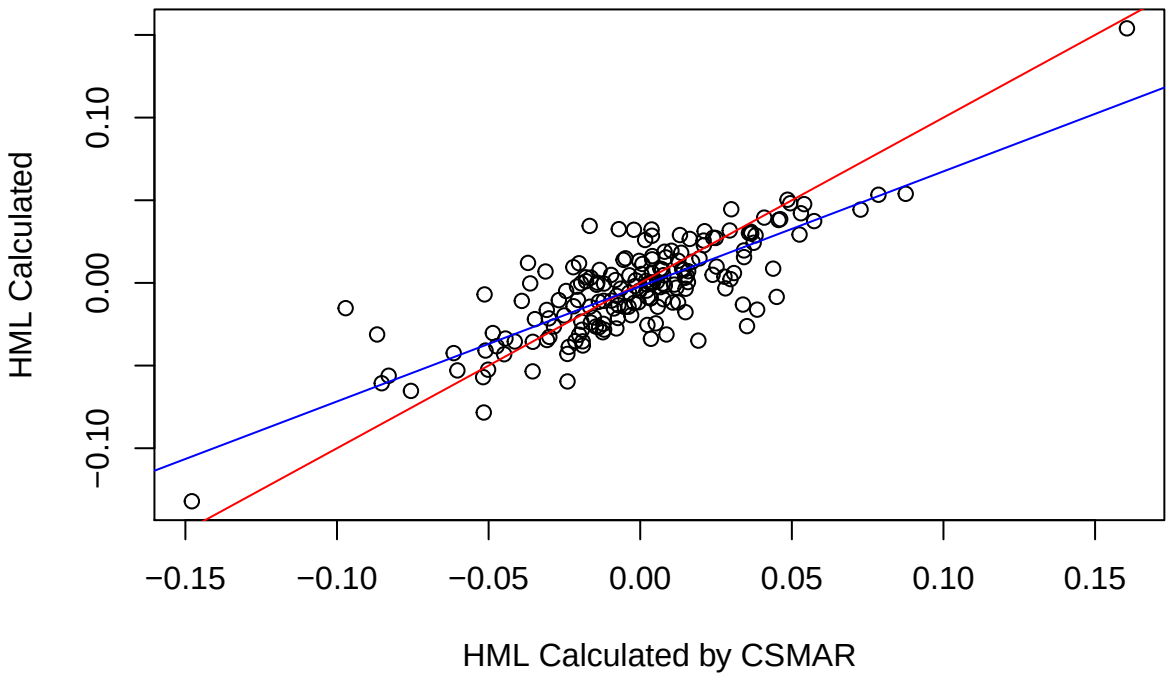
```
}

#select from 2019.7~2020.6 to L, M, H
S15<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2018-12-31",])
S15<-arrange(S15,S15$BM)
S15_L<-ZTT(S15)
S15_M<-TTS(S15)
S15_H<-STE(S15)

for ( i in c("2019-07","2019-08","2019-09","2019-10","2019-11","2019-12","2020-01","2020-02","2020-03",'

S15_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S15_L_M<-merge(S15_L,S15_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S15_H_M<-merge(S15_H,S15_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S15_H_M$Mretwd,Weight(S15_H_M$Msmvosd))-weighted.mean(S15_

}

#select from 2020.7~2020.12 to L, M, H
S16<-with(BM_ratio,BM_ratio[BM_ratio$Ddate=="2019-12-31",])
S16<-arrange(S16,S16$BM)
S16_L<-ZTT(S16)
S16_M<-TTS(S16)
S16_H<-STE(S16)

for ( i in c("2020-07","2020-08","2020-09","2020-10","2020-11","2020-12") ) {

S16_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
S16_L_M<-merge(S16_L,S16_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
S16_H_M<-merge(S16_H,S16_M,by.x='Stkcd',by.y = 'Stkcd')

FAMA[FAMA$TradingMonth==i,]$HML=weighted.mean(S16_H_M$Mretwd,Weight(S16_H_M$Msmvosd))-weighted.mean(S16_

}

FAMA$TradingMonth<-as.yearmon(FAMA$TradingMonth)

FAMA$HML<-FAMA$HML/2
cor.test(FAMA$HML1,FAMA$HML)


##
##  Pearson's product-moment correlation
```

```
##
## data:  FAMA$HML1 and FAMA$HML
## t = 17.984, df = 178, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7443005 0.8495943
## sample estimates:
##       cor
## 0.8031318
```

```
ggplot(data=FAMA,mapping=aes(x=TradingMonth,y=HML1))+
  geom_point(aes(x=TradingMonth,y=HML1,colour="HML Calculated by CSMAR"))+
  geom_point(data=FAMA,mapping=aes(x=TradingMonth,y=HML,colour="HML calculated"))+
  theme(axis.title.x=element_text(size=14),axis.title.y=element_text(size=14),axis.text.x = element_tex
  ylab("Value Factor")
```




```
reg<-lm(FAMA$HML~FAMA$HML1)
plot(x=FAMA$HML1,y=FAMA$HML,xlab="HML Calculated by CSMAR",ylab="HML Calculated")
abline(a=0,b=1,col="red")
abline(reg,col="blue")
```




```
FRENCH<-read_excel("    + .xlsx")
FRENCH<-with(FRENCH,FRENCH[FRENCH$MarkettypeID=='P9706',])
FRENCH<-FRENCH[,-4]
FRENCH$HML<-0
FRENCH$HML1<-as.numeric(FRENCH$HML1)
FRENCH<-na.omit(FRENCH)
```

```
#Select Stock not from second board
BM_ratio_1<-BM_ratio[-(13546:18200),]
```

```
#Select form 2010.7~2011.6
N1<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2009-12-31",])
N1<-arrange(N1,N1$BM)
N1_L<-ZTT(N1)
N1_M<-TTS(N1)
N1_H<-STE(N1)
```

```
for ( i in c("2010-07","2010-08","2010-09","2010-10","2010-11","2010-12","2011-01","2011-02","2011-03",

N1_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N1_L_M<-merge(N1_L,N1_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
```

```
N1_H_M<-merge(N1_H,N1_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N1_H_M$Mretwd,Weight(N1_H_M$Msmvosd))-weighted.mean(N

}
```

```
#Select form 2011.7~2012.6
N2<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2010-12-31",])
N2<-arrange(N2,N2$BM)
N2_L<-ZTT(N2)
N2_M<-TTS(N2)
N2_H<-STE(N2)
```

```
for ( i in c("2011-07","2011-08","2011-09","2011-10","2011-11","2011-12","2012-01","2012-02","2012-03",

N2_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N2_L_M<-merge(N2_L,N2_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N2_H_M<-merge(N2_H,N2_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N2_H_M$Mretwd,Weight(N2_H_M$Msmvosd))-weighted.mean(N

}
```

```
#Select form 2012.7~2013.6
N3<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2011-12-31",])
N3<-arrange(N3,N3$BM)
N3_L<-ZTT(N3)
N3_M<-TTS(N3)
N3_H<-STE(N3)
```

```
for ( i in c("2012-07","2012-08","2012-09","2012-10","2012-11","2012-12","2013-01","2013-02","2013-03",

N3_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N3_L_M<-merge(N3_L,N3_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N3_H_M<-merge(N3_H,N3_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N3_H_M$Mretwd,Weight(N3_H_M$Msmvosd))-weighted.mean(N

}
```

```
#Select form 2013.7~2014.6
N4<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2012-12-31",])
N4<-arrange(N4,N4$BM)
N4_L<-ZTT(N4)
```

```
N4_M<-TTS(N4)
N4_H<-STE(N4)
```

```
for ( i in c("2013-07","2013-08","2013-09","2013-10","2013-11","2013-12","2014-01","2014-02","2014-03",

N4_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N4_L_M<-merge(N4_L,N4_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N4_H_M<-merge(N4_H,N4_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N4_H_M$Mretwd,Weight(N4_H_M$Msmvosd))-weighted.mean(N4

}
```

```
#Select form 2014.7~2015.6
N5<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2013-12-31",])
N5<-arrange(N5,N5$BM)
N5_L<-ZTT(N5)
N5_M<-TTS(N5)
N5_H<-STE(N5)
```

```
for ( i in c("2014-07","2014-08","2014-09","2014-10","2014-11","2014-12","2015-01","2015-02","2015-03",

N5_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N5_L_M<-merge(N5_L,N5_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N5_H_M<-merge(N5_H,N5_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N5_H_M$Mretwd,Weight(N5_H_M$Msmvosd))-weighted.mean(N5

}
```

```
#Select form 2015.7~2016.6
N6<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2014-12-31",])
N6<-arrange(N6,N6$BM)
N6_L<-ZTT(N6)
N6_M<-TTS(N6)
N6_H<-STE(N6)
```

```
for ( i in c("2015-07","2015-08","2015-09","2015-10","2015-11","2015-12","2016-01","2016-02","2016-03",

N6_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N6_L_M<-merge(N6_L,N6_M,by.x='Stkcd',by.y = 'Stkcd')
```

```
#calculate weighted return for H
N6_H_M<-merge(N6_H,N6_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N6_H_M$Mretwd,Weight(N6_H_M$Msmvosd))-weighted.mean(N

}
```

```
#Select form 2016.7~2017.6
N7<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2015-12-31",])
N7<-arrange(N7,N7$BM)
N7_L<-ZTT(N7)
N7_M<-TTS(N7)
N7_H<-STE(N7)
```

```
for ( i in c("2016-07","2016-08","2016-09","2016-10","2016-11","2016-12","2017-01","2017-02","2017-03",

N7_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N7_L_M<-merge(N7_L,N7_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N7_H_M<-merge(N7_H,N7_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N7_H_M$Mretwd,Weight(N7_H_M$Msmvosd))-weighted.mean(N

}
```

```
#Select form 2017.7~2018.6
N8<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2016-12-31",])
N8<-arrange(N8,N8$BM)
N8_L<-ZTT(N8)
N8_M<-TTS(N8)
N8_H<-STE(N8)
```

```
for ( i in c("2017-07","2017-08","2017-09","2017-10","2017-11","2017-12","2018-01","2018-02","2018-03",

N8_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N8_L_M<-merge(N8_L,N8_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N8_H_M<-merge(N8_H,N8_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N8_H_M$Mretwd,Weight(N8_H_M$Msmvosd))-weighted.mean(N

}
```

```
#Select form 2018.7~2019.6
N9<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2017-12-31",])
N9<-arrange(N9,N9$BM)
```

```
N9_L<-ZTT(N9)
N9_M<-TTS(N9)
N9_H<-STE(N9)
```

```
for ( i in c("2018-07","2018-08","2018-09","2018-10","2018-11","2018-12","2019-01","2019-02","2019-03",'

N9_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N9_L_M<-merge(N9_L,N9_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N9_H_M<-merge(N9_H,N9_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N9_H_M$Mretwd,Weight(N9_H_M$Msmvosd))-weighted.mean(N

}
```

```
#Select form 2019.7~2020.6
N10<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2018-12-31",])
N10<-arrange(N10,N10$BM)
N10_L<-ZTT(N10)
N10_M<-TTS(N10)
N10_H<-STE(N10)
```

```
for ( i in c("2019-07","2019-08","2019-09","2019-10","2019-11","2019-12","2020-01","2020-02","2020-03",'

N10_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N10_L_M<-merge(N10_L,N10_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N10_H_M<-merge(N10_H,N10_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N10_H_M$Mretwd,Weight(N10_H_M$Msmvosd))-weighted.mean

}
```

```
#Select form 2020.7~2020.12
N11<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2019-12-31",])
N11<-arrange(N11,N11$BM)
N11_L<-ZTT(N11)
N11_M<-TTS(N11)
N11_H<-STE(N11)
```

```
for ( i in c("2020-07","2020-08","2020-09","2020-10","2020-11","2020-12") ) {

N11_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N11_L_M<-merge(N11_L,N11_M,by.x='Stkcd',by.y = 'Stkcd')
```

```
#calculate weighted return for H
N11_H_M<-merge(N11_H,N11_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N11_H_M$Mretwd,Weight(N11_H_M$Msmvosd))-weighted.mean

}
```

```
#Select form 2006.1~2006.6
N12<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2004-12-31",])
N12<-arrange(N12,N12$BM)
N12_L<-ZTT(N12)
N12_M<-TTS(N12)
N12_H<-STE(N12)
```

```
for ( i in c("2006-01","2006-02","2006-03","2006-04","2006-05","2006-06") ) {

N12_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N12_L_M<-merge(N12_L,N12_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N12_H_M<-merge(N12_H,N12_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N12_H_M$Mretwd,Weight(N12_H_M$Msmvosd))-weighted.mean

}
```

```
#Select form 2006.7~2007.6
N13<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2005-12-31",])
N13<-arrange(N13,N13$BM)
N13_L<-ZTT(N13)
N13_M<-TTS(N13)
N13_H<-STE(N13)
```

```
for ( i in c("2006-07","2006-08","2006-09","2006-10","2006-11","2006-12","2007-01","2007-02","2007-03",

N13_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N13_L_M<-merge(N13_L,N13_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N13_H_M<-merge(N13_H,N13_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N13_H_M$Mretwd,Weight(N13_H_M$Msmvosd))-weighted.mean

}
```

```
#Select form 2007.7~2008.6
N14<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2006-12-31",])
N14<-arrange(N14,N14$BM)
```

```
N14_L<-ZTT(N14)
N14_M<-TTS(N14)
N14_H<-STE(N14)
```

```
for ( i in c("2007-07","2007-08","2007-09","2007-10","2007-11","2007-12","2008-01","2008-02","2008-03",'

N14_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N14_L_M<-merge(N14_L,N14_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N14_H_M<-merge(N14_H,N14_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N14_H_M$Mretwd,Weight(N14_H_M$Msmvosd))-weighted.mean

}
```

```
#Select form 2008.7~2009.6
N15<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2007-12-31",])
N15<-arrange(N15,N15$BM)
N15_L<-ZTT(N15)
N15_M<-TTS(N15)
N15_H<-STE(N15)
```

```
for ( i in c("2008-07","2008-08","2008-09","2008-10","2008-11","2008-12","2009-01","2009-02","2009-03",'

N15_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N15_L_M<-merge(N15_L,N15_M,by.x='Stkcd',by.y = 'Stkcd')

#calculate weighted return for H
N15_H_M<-merge(N15_H,N15_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N15_H_M$Mretwd,Weight(N15_H_M$Msmvosd))-weighted.mean

}
```

```
#Select form 2009.7~2010.6
N16<-with(BM_ratio_1,BM_ratio_1[BM_ratio_1$Ddate=="2008-12-31",])
N16<-arrange(N16,N16$BM)
N16_L<-ZTT(N16)
N16_M<-TTS(N16)
N16_H<-STE(N16)
```

```
for ( i in c("2009-07","2009-08","2009-09","2009-10","2009-11","2009-12","2010-01","2010-02","2010-03",'

N16_M<-with(CNwhole,CNwhole[CNwhole$Trdmnt==i,])

#calculate weighted return for L
N16_L_M<-merge(N16_L,N16_M,by.x='Stkcd',by.y = 'Stkcd')
```

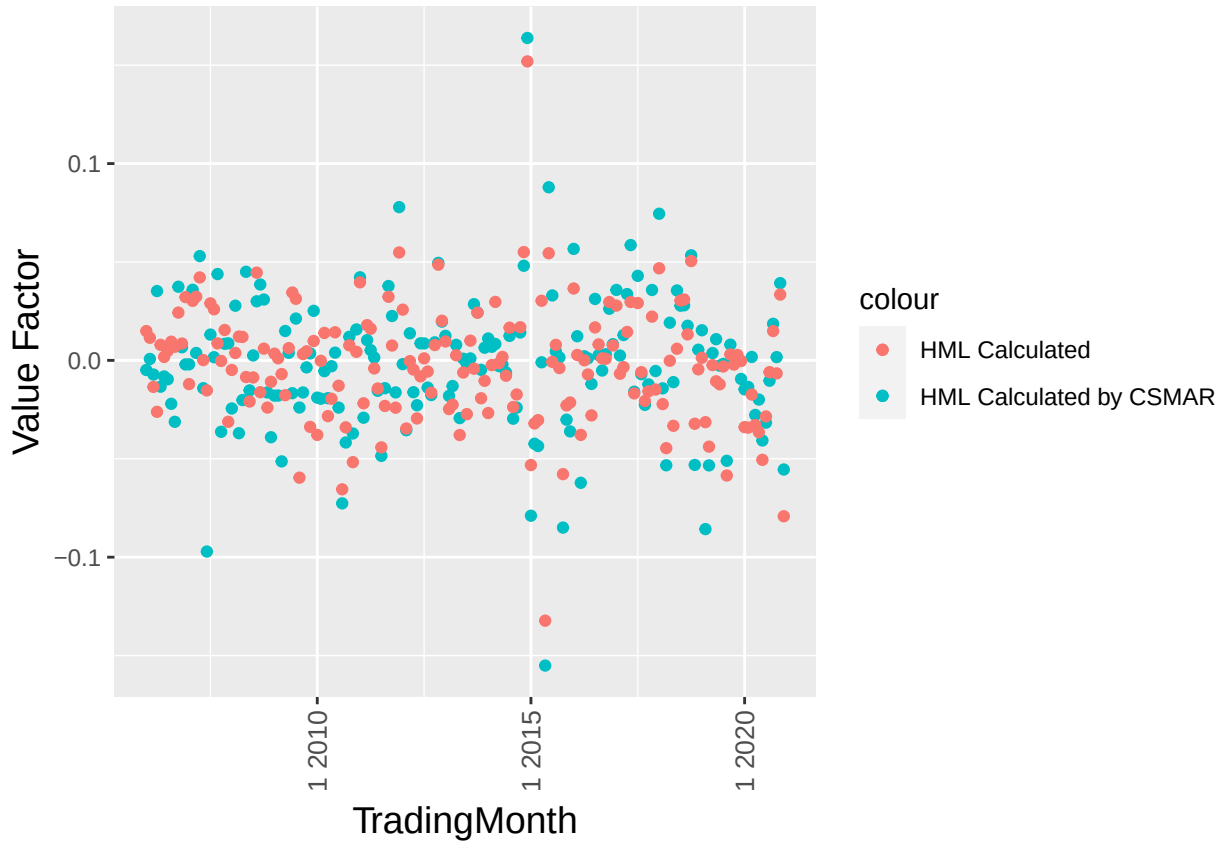
```
#calculate weighted return for H
N16_H_M<-merge(N16_H,N16_M,by.x='Stkcd',by.y = 'Stkcd')

FRENCH[FRENCH$TradingMonth==i,]$HML=weighted.mean(N16_H_M$Mretwd,Weight(N16_H_M$Msmvosd))-weighted.mean

}
```

```
FRENCH$TradingMonth<-as.yearmon(FRENCH$TradingMonth)
```

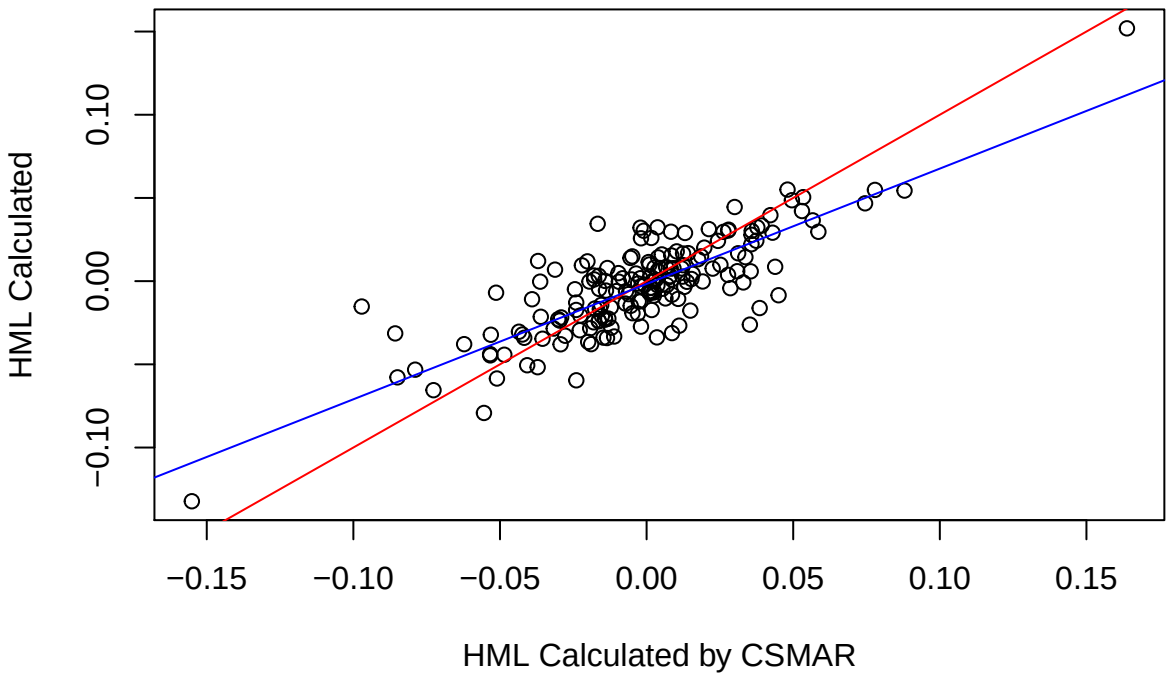
```
FRENCH$HML<-FRENCH$HML/2
cor.test(FRENCH$HML1,FRENCH$HML)
```

```
##
##  Pearson's product-moment correlation
##
## data:  FRENCH$HML1 and FRENCH$HML
## t = 18.714, df = 178, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7583122 0.8583049
## sample estimates:
##      cor
## 0.814263
```

```
ggplot(data=FRENCH,mapping=aes(x=TradingMonth,y=HML1))+
  geom_point(aes(x=TradingMonth,y=HML1,colour="HML Calculated by CSMAR"))+
  geom_point(mapping=aes(x=TradingMonth,y=HML,colour="HML Calculated"))+
  theme(axis.title.x=element_text(size=14),axis.title.y=element_text(size=14),axis.text.x = element_tex
  ylab("Value Factor")
```




```
reg<-lm(FRENCH$HML~FRENCH$HML1)
plot(x=FRENCH$HML1,y=FRENCH$HML,xlab="HML Calculated by CSMAR",ylab="HML Calculated")
abline(a=0,b=1,col="red")
abline(reg,col="blue")
```


HML Calculated
0.10
0.00
−0.10
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
HML Calculated by CSMAR