第一步：数据初步处理 # ProblemSet2 —

```
**************************ProblemSet2****************************************
/*a*/
/*     */
import delimited "TRD_Mnth.csv",encoding(UTF-16) clear
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth, replace
xtset yearmonth
save "return",replace                    /*   */

/*     cumulative return  monthly return*/

sort stkcd
bys stkcd: gen cumu= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*(
/*   cumulative return*/
drop if cumu==.                  /*   */
ssc install astile
bys trdmnt: astile portfolio = cumu, nq(10)
rename portfolio g
bysort trdmnt g:egen r=mean(mretwd)
duplicates drop  trdmnt g,force /*   */
sort g trdmnt
save portfolio                   /*      portfolio.dta */

/*   annualize return mean variance*/
bys g trdmnt: gen annual=r*12
by g:egen mean=mean(annual)
by g:egen variance=var(annual)
duplicates drop g,force          /*   */
keep g mean variance             /*    drop */
save annual_meanvar              /*  anuual_return mean variance   annual_return.dta */



/*b*/
/*     */
use "/Users/yinjie/Desktop/project2/Problemset2/portfolio.dta", clear
keep if g==1|g==10
sort trdmnt g
gen rwinner= r if g==10
gen rloser= r if g==1
bys trdmnt:gen rloser_sameline= rloser[_n-1] /* rwinner rloser     */
gen rstrategy= rwinner-rloser_sameline  /*generate strategy return*/
keep trdmnt rwinner rloser_sameline rstrategy
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
destring yearmonth,replace
drop year month
drop if rstrategy==.
save rstrategy,replace                /*    strategy return.dta */

import delimited "3factormodel.csv",encoding(UTF-8) clear
gen year=substr(date,1,4)
gen month=substr(date,6,2)
gen yearmonth= year+ month
destring yearmonth,replace
save 3factor                    /*3factor   */
```

```
/*   strategy return 3factor    alpha*/
use "3factor.dta"
merge 1:1 yearmonth using "rstrategy.dta"
drop _merge
merge 1:1 yearmonth using "rf.dta"
keep yearmonth rstrategy rf rp smb hml rwinner rloser_sameline
drop if rstrategy==.
gen Rstrategy=rstrategy-rf
reg Rstrategy rp smb hml
save momentum_alpha,replace

/*c*/
/*************************     ************************************/
use "return.dta"
bys stkcd: gen cumu3= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])-1
bys stkcd: gen cumu4= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])-1
bys stkcd: gen cumu5= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])-
bys stkcd: gen cumu6= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
bys stkcd: gen cumu7= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
bys stkcd: gen cumu8= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
bys stkcd: gen cumu9= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
bys stkcd: gen cumu10= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])
bys stkcd: gen cumu11= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])
bys stkcd: gen cumu12= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])
/*  3-12 cumulative return*/
bys trdmnt: astile g3 = cumu3, nq(10)
bys trdmnt: astile g4 = cumu4, nq(10)
bys trdmnt: astile g5 = cumu5, nq(10)
bys trdmnt: astile g6 = cumu6, nq(10)
bys trdmnt: astile g7 = cumu7, nq(10)
bys trdmnt: astile g8 = cumu8, nq(10)
bys trdmnt: astile g9 = cumu9, nq(10)
bys trdmnt: astile g10 = cumu10, nq(10)
bys trdmnt: astile g11 = cumu11, nq(10)
bys trdmnt: astile g12 = cumu12, nq(10)
/*  3-12 cumulative return */
bysort trdmnt g3:egen r3=mean(mretwd)
bysort trdmnt g4:egen r4=mean(mretwd)
bysort trdmnt g5:egen r5=mean(mretwd)
bysort trdmnt g6:egen r6=mean(mretwd)
bysort trdmnt g7:egen r7=mean(mretwd)
bysort trdmnt g8:egen r8=mean(mretwd)
bysort trdmnt g9:egen r9=mean(mretwd)
bysort trdmnt g10:egen r10=mean(mretwd)
bysort trdmnt g11:egen r11=mean(mretwd)
bysort trdmnt g12:egen r12=mean(mretwd)
/*     monthly return*/
save raw3_12,replace
/*     raw3_12.dta */

use "return.dta"
sort stkcd
bys stkcd: gen cumu3= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])-1
drop if cumu3==.
bys trdmnt: astile g3 = cumu3, nq(10)
bysort trdmnt g3:egen r3=mean(mretwd)
duplicates drop  trdmnt g3,force
sort g3 trdmnt
save portfolio3,replace
```

```
use "return.dta"
sort stkcd
bys stkcd: gen cumu4= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])-1
drop if cumu4==.
bys trdmnt: astile g4 = cumu4, nq(10)
bysort trdmnt g4:egen r4=mean(mretwd)
duplicates drop  trdmnt g4,force
sort g4 trdmnt
save portfolio4,replace

use "return.dta"
sort stkcd
bys stkcd: gen cumu5= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])-
drop if cumu5==.
bys trdmnt: astile g5 = cumu5, nq(10)
bysort trdmnt g5:egen r5=mean(mretwd)
duplicates drop  trdmnt g5,force
sort g5 trdmnt
save portfolio5,replace

use "return.dta"
sort stkcd
bys stkcd: gen cumu6= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
drop if cumu6==.
bys trdmnt: astile g6 = cumu6, nq(10)
bysort trdmnt g6:egen r6=mean(mretwd)
duplicates drop  trdmnt g6,force
sort g6 trdmnt
save portfolio6,replace

use "return.dta"
sort stkcd
bys stkcd: gen cumu7= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
drop if cumu7==.
bys trdmnt: astile g7 = cumu7, nq(10)
bysort trdmnt g7:egen r7=mean(mretwd)
duplicates drop  trdmnt g7,force
sort g7 trdmnt
save portfolio7,replace

use "return.dta"
sort stkcd
bys stkcd: gen cumu8= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
drop if cumu8==.
bys trdmnt: astile g8 = cumu8, nq(10)
bysort trdmnt g8:egen r8=mean(mretwd)
duplicates drop  trdmnt g8,force
sort g8 trdmnt
save portfolio8,replace

use "return.dta"
sort stkcd
bys stkcd: gen cumu9= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])*
drop if cumu9==.
bys trdmnt: astile g9 = cumu9, nq(10)
bysort trdmnt g9:egen r9=mean(mretwd)
duplicates drop  trdmnt g9,force
sort g9 trdmnt
save portfolio9,replace

use "return.dta"
```

```
sort stkcd
bys stkcd: gen cumu10= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])
drop if cumu10==.
bys trdmnt: astile g10 = cumu10, nq(10)
bysort trdmnt g10:egen r10=mean(mretwd)
duplicates drop  trdmnt g10,force
sort g10 trdmnt
save portfolio10,replace

use "return.dta"
sort stkcd
bys stkcd: gen cumu11= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])
drop if cumu11==.
bys trdmnt: astile g11 = cumu11, nq(10)
bysort trdmnt g11:egen r11=mean(mretwd)
duplicates drop  trdmnt g11,force
sort g11 trdmnt
save portfolio11,replace

use "return.dta"
sort stkcd
bys stkcd: gen cumu12= (1+ mretwd[_n-1])*(1+ mretwd[_n-2])*(1+ mretwd[_n-3])*(1+ mretwd[_n-4])*(1+ mretwd[_n-5])
drop if cumu12==.
bys trdmnt: astile g12 = cumu12, nq(10)
bysort trdmnt g12:egen r12=mean(mretwd)
duplicates drop  trdmnt g12,force
sort g12 trdmnt
save portfolio12,replace

import delimited "TRD_Nrrate.csv",encoding(UTF-16) clear
gen rf= nrrmtdt/100
gen year=substr(clsdt,1,4)
gen month=substr(clsdt,6,2)
gen yearmonth= year+ month
gen yearmonth= year+ month
duplicates drop yearmonth,force
keep rf yearmonth
save rf                              /* 10/6 19/6      "rf.dta" */


/***************     3-12  10 portfolio annualize return mean variance**************/
use "raw3_12.dta"
duplicates drop  trdmnt g3,force
keep g3 r3 trdmnt
bys g3 trdmnt: gen annual3=r3*12
by g3:egen mean=mean(annual3)
by g3:egen variance=var(annual3)
duplicates drop g3,force
drop trdmnt
drop if g3==.
list
save "annual3",replace

use "raw3_12.dta"
duplicates drop  trdmnt g4,force
keep g4 r4 trdmnt
bys g4 trdmnt: gen annual4=r4*12
by g4:egen mean=mean(annual4)
by g4:egen variance=var(annual4)
duplicates drop g4,force
drop trdmnt
```

```
drop if g4==.
list
save "annual4",replace

use "raw3_12.dta"
duplicates drop  trdmnt g5,force
keep g5 r5 trdmnt
bys g5 trdmnt: gen annual5=r5*12
by g5:egen mean=mean(annual5)
by g5:egen variance=var(annual5)
duplicates drop g5,force
drop trdmnt
drop if g5==.
list
save "annual5",replace

use "raw3_12.dta"
duplicates drop  trdmnt g6,force
keep g6 r6 trdmnt
bys g6 trdmnt: gen annual6=r6*12
by g:egen mean=mean(annual6)
by g:egen variance=var(annual6)
duplicates drop g6,force
drop trdmnt
drop if g6==.
list
save "annual6",replace

use "raw3_12.dta"
duplicates drop  trdmnt g7,force
keep g7 r7 trdmnt
bys g7 trdmnt: gen annual7=r7*12
by g:egen mean=mean(annual7)
by g:egen variance=var(annual7)
duplicates drop g7,force
drop trdmnt
drop if g7==.
list
save "annual7",replace

use "raw3_12.dta"
duplicates drop  trdmnt g8,force
keep g8 r8 trdmnt
bys g8 trdmnt: gen annual8=r8*12
by g:egen mean=mean(annual8)
by g:egen variance=var(annual8)
duplicates drop g8,force
drop trdmnt
drop if g8==.
list
save "annual8",replace

use "raw3_12.dta"
duplicates drop  trdmnt g9,force
keep g9 r9 trdmnt
bys g9 trdmnt: gen annual9=r9*12
by g:egen mean=mean(annual9)
by g:egen variance=var(annual9)
duplicates drop g9,force
drop trdmnt
drop if g9==.
```

```
list
save "annual9",replace

use "raw3_12.dta"
duplicates drop  trdmnt g10,force
keep g10 r10 trdmnt
bys g10 trdmnt: gen annual10=r10*12
by g:egen mean=mean(annual10)
by g:egen variance=var(annual10)
duplicates drop g10,force
drop trdmnt
drop if g10==.
list
save "annual10",replace

use "raw3_12.dta"
duplicates drop  trdmnt g11,force
keep g11 r11 trdmnt
bys g11 trdmnt: gen annual11=r11*12
by g:egen mean=mean(annual11)
by g:egen variance=var(annual11)
duplicates drop g11,force
drop trdmnt
drop if g11==.
list
save "annual11",replace

use "raw3_12.dta"
duplicates drop  trdmnt g12,force
keep g12 r12 trdmnt
bys g12 trdmnt: gen annual12=r12*12
by g:egen mean=mean(annual12)
by g:egen variance=var(annual12)
duplicates drop g12,force
drop trdmnt
drop if g12==.
list
save "annual12",replace
/*   3-12   annualized return mean variance*/


/***************     3-12 strategy return************************/
use "portfolio3.dta"
keep if g3==1|g3==10
sort trdmnt g3
gen rwinner= r3 if g3==10
gen rloser= r3 if g3==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy3= rwinner-rloser_sameline
keep trdmnt rwinner rloser rloser_sameline rstrategy
drop if rstrategy3==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy3,replace

use "portfolio4.dta"
duplicates drop g4 r4,force
```

```
keep if g4==1|g4==10
sort trdmnt g4
gen rwinner= r4 if g4==10
gen rloser= r4 if g4==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy4= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy4==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy4,replace

use "portfolio5.dta"
duplicates drop g5 r5,force
keep if g5==1|g5==10
sort trdmnt g5
gen rwinner= r5 if g5==10
gen rloser= r5 if g5==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy5= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy5==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy5,replace

use "portfolio6.dta"
duplicates drop g6 r6,force
keep if g6==1|g6==10
sort trdmnt g6
gen rwinner= r6 if g6==10
gen rloser= r6 if g6==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy6= rwinner-rloser_sameline
drop if rstrategy6==.
keep trdmnt rwinner rloser_sameline rstrategy
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy6,replace

use "portfolio7.dta"
duplicates drop g7 r7,force
keep if g7==1|g7==10
sort trdmnt g7
gen rwinner= r7 if g7==10
gen rloser= r7 if g7==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy7= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy7==.
```

```
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy7,replace

use "portfolio8.dta"
duplicates drop g8 r8,force
keep if g8==1|g8==10
sort trdmnt g8
gen rwinner= r8 if g8==10
gen rloser= r8 if g8==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy8= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy8==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy8,replace

use "portfolio9.dta"
duplicates drop g9 r9,force
keep if g9==1|g9==10
sort trdmnt g9
gen rwinner= r9 if g9==10
gen rloser= r9 if g9==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy9= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy9==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy9,replace

use "portfolio10.dta"
duplicates drop g10 r10,force
keep if g10==1|g10==10
sort trdmnt g10
gen rwinner= r10 if g10==10
gen rloser= r10 if g10==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy10= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy10==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy10,replace
```

```
use "portfolio11.dta"
duplicates drop g11 r11,force
keep if g11==1|g11==10
sort trdmnt g11
gen rwinner= r11 if g11==10
gen rloser= r11 if g11==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy11= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy11==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy11,replace

use "portfolio12.dta"
duplicates drop g12 r12,force
keep if g12==1|g12==10
sort trdmnt g12
gen rwinner= r12 if g12==10
gen rloser= r12 if g12==1
bys trdmnt:gen rloser_sameline= rloser[_n-1]
gen rstrategy12= rwinner-rloser_sameline
keep trdmnt rwinner rloser_sameline rstrategy
drop if rstrategy12==.
drop rloser
gen year=substr(trdmnt,1,4)
gen month=substr(trdmnt,-2,2)
gen yearmonth= year+ month
drop year month
destring yearmonth,replace
save rstrategy12,replace

use "rstrategy12.dta"
merge m:m yearmonth using "rstrategy3"
drop _merge
merge m:m yearmonth using "rstrategy4"
drop _merge
merge m:m yearmonth using "rstrategy5"
drop _merge
merge m:m yearmonth using "rstrategy6"
drop _merge
merge m:m yearmonth using "rstrategy7"
drop _merge
merge m:m yearmonth using "rstrategy8"
drop _merge
merge m:m yearmonth using "rstrategy9"
drop _merge
merge m:m yearmonth using "rstrategy10"
drop _merge
merge m:m yearmonth using "rstrategy11"
destring yearmonth,replace
drop _merge
merge m:m yearmonth using "3factor"
drop _merge
merge 1:1 yearmonth using "rf"
save strategy3_12_3factor,replace
```

```
gen Rstrategy3= rstrategy3-rf
gen Rstrategy4= rstrategy4-rf
gen Rstrategy5= rstrategy5-rf
gen Rstrategy6= rstrategy6-rf
gen Rstrategy7= rstrategy7-rf
gen Rstrategy8= rstrategy8-rf
gen Rstrategy9= rstrategy9-rf
gen Rstrategy10= rstrategy10-rf
gen Rstrategy11= rstrategy11-rf
gen Rstrategy12= rstrategy12-rf

reg Rstrategy3 rp smb hml
reg Rstrategy4 rp smb hml
reg Rstrategy5 rp smb hml
reg Rstrategy6 rp smb hml
reg Rstrategy7 rp smb hml
reg Rstrategy8 rp smb hml
reg Rstrategy9 rp smb hml
reg Rstrategy10 rp smb hml
reg Rstrategy11 rp smb hml
reg Rstrategy12 rp smb hml

save regression

gen annual3=rstrategy3*12
gen annual4=rstrategy4*12
gen annual5=rstrategy5*12
gen annual6=rstrategy6*12
gen annual7=rstrategy7*12
gen annual8=rstrategy8*12
gen annual9=rstrategy9*12
gen annual10=rstrategy10*12
gen annual11=rstrategy11*12
gen annual12=rstrategy12*12

sum annual3 annual4 annual5 annual6 annual7 annual8 annual9 annual10 annual11 annual12

/*d*/
/*   strategy return */
use "momentum_alpha.dta"
sort rstrategy
list
/*   market risk premium  */

/*      */
/*this part is conducted using Python*/
```