```stata
1
2
3    /*********************************************************************
     **************
4    Problem 2: CAPM and APT
5    *********************************************************************
     *************
6                          *2.1.1 Data pre-processing
7    *----------------------------------------------------------------
     ------------*/
8    import excel "data/DGS1.xls", sheet("FRED") firstrow clear
9
10   *create var fyear
11   gen fyear = yofd(observation_date)
12
13   *gen fyear = year(observation_date)
14
15
16   *create var r_f
17   gen r_f = DGS1/100
18
19   save "data/riskfree", replace
20
21   use "data/bvhpygvwvwympbr3.dta" if _n==1, clear
22
23   glo varlist GVKEY LINKPRIM fyear csho prcc_f bkvlps conml sic
24
25   use $varlist using "data/bvhpygvwvwympbr3.dta"
26
27   des $varlist
28   destring GVKEY, replace
29
30   keep if LINKPRIM == "P"
31   drop if fyear==.
32
33   merge m:1 fyear using "data/riskfree.dta"
34   drop if _merge !=3
35
36   /*********************************************************************
     **************
37           2.2 Variable construction and Market execess returns
38   *----------------------------------------------------------------
     ------------*/
39   * Market capitalization
40   gen mar_cap = prcc_f * csho
41   lab var mar_cap "firm market capitalization"
42
43   * Total market capitalization for every year
44   bysort fyear: egen tot_mar_cap= total(mar_cap)
45   lab var tot_mar_cap "total market capitalization for every year"
46
47   *Market porfolio weights for each stock
```

```stata
48    bysort conml : gen mar_por_wei = mar_cap/tot_mar_cap
49    lab var mar_por_wei "Market porfolio weights for each stock"
50
51    *One year ahead stock returns
52    bysort GVKEY : gen r = (prcc_f[_n+1] − prcc_f)/prcc_f
53    lab var r "stock return (annual)"
54
55    gen excess_r = r − r_f
56    lab var excess_r "excess return (annual)"
57
58    bysort fyear : egen market_r =total(r* mar_por_wei)
59    lab var market_r "market return (annual)"
60
61    gen excess_mar_r = market_r − r_f
62    lab var excess_r "excess market return (annual)"
63
64    *2.3 Scatter plot for Apple, Inc
65    list if conml=="Apple Inc"
66
67    tw (scatter excess_r excess_mar_r if conml=="Apple Inc",mc(navy)
      mlabel(fyear) mlabs(vsmall) mlabc(navy)) ///
68       (lfit excess_r excess_mar_r if conml=="Apple Inc", lc(red)
      lwidth(medthick)), ///
69       title("Apple Inc") legend(off) ///
70       xtitle("Market excess return (annual)") ///
71       ytitle("Firm excess stock return (annual)") ///
72       xlabel(, labsize(small)) ylabel(, nogrid labsize(small)) ///
73       plotregion(fcolor(white) lcolor(white)) ///
74       graphregion(fcolor(white) lcolor(white))
75    graph export "docs/Apple_Inc.pdf", replace
76
77    *2.4 Regression for different companies
78
79    reg excess_r excess_mar_r if conml== "Apple Inc"
80    outreg2 using "docs/capm_reg.xls", ctitle("Apple Inc") replace
81    reg excess_r excess_mar_r if conml== "The Kraft Heinz Co"
82    outreg2 using "docs/capm_reg.xls", ctitle("Kraft Heinz") append
83    reg excess_r excess_mar_r if conml==  "General Electric Co"
84    outreg2 using "docs/capm_reg.xls", ctitle( "General Electric")
      append
85
86    /*comments
87    the coefficients are not statistically significants with
      relatively high standard
88    error. The coef for apple and Kraft Heinz are negative while the
      ceof of General
89    Electric is positive*/
90
91    *2.5 Pooled CAPM
92    * sic
93    gen sic2= substr(sic,1,2)
94
```

```stata
95   *First way
96   tab sic, gen(sic_)
97   tab sic, gen(sic2_)
98
99   reg excess_r excess_mar_r
100  outreg2 using "docs/capm_reg.xls", ctitle(pooled regression)
     append
101  reg excess_r excess_mar_r sic2_*
102  outreg2 using "docs/capm_reg.xls", ctitle(first way, fixed effect
      1) append
103  reg excess_r excess_mar_r sic_*
104  outreg2 using "docs/capm_reg.xls", ctitle(first way, fixed effect
      2) append
105
106  *Second way
107  xi:reg excess_r excess_mar_r i.sic2
108  outreg2 using "docs/capm_reg.xls", ctitle(Second way, fixed
     effect 1) append
109
110  xi:reg excess_r excess_mar_r i.sic
111  outreg2 using "docs/capm_reg.xls", ctitle(Second way, fixed
     effect 2) append
112
113  ** Third way
114  areg excess_r excess_mar_r, absorb(sic2)
115  outreg2 using "docs/capm_reg.xls", ctitle(Third way, fixed effect
      1) append
116  areg excess_r excess_mar_r, absorb(sic)
117  outreg2 using "docs/capm_reg.xls", ctitle(Third way, fixed effect
      2) append
118
119  /*Regarless of the method used, the coefficients from the
     pooling regreesion
120  are quit similar*/
121
122  **2.6 Variable construction : Fama-French Factor "SML"
123
124  bysort fyear: egen cap_10pctile = pctile(mar_cap), p(10)
125  bysort fyear: egen cap_90pctile = pctile(mar_cap), p(90)
126
127  *percentile dummy
128  gen cap_lpctile = (mar_cap< cap_10pctile)
129  gen cap_hpctile = (mar_cap> cap_90pctile)
130
131
132  *Average stock return for 90th and 10th percentile
133  bysort fyear:egen average_ret90 = mean(r) if cap_hpctile==1
134  bysort fyear:egen average_ret10 = mean(r) if cap_lpctile==1
135
136  bysort fyear: egen r_l = min(average_ret90)
137  bysort fyear: egen r_s = min(average_ret10)
138
```

```
139    gen sml = r_s-r_l
140    sum sml
141
142    *2.7 variable construction:Fama-French Factor "HML"
143
144    gen BTM = bkvlps/prcc_f
145
146    *generate values of the 90th and 10 percentileof BTM
147
148    bysort fyear: egen btm_10pctile = pctile(BTM), p(10)
149    bysort fyear: egen btm_90pctile = pctile(BTM), p(90)
150
151    *percentile dummy
152    gen lpctile_btm = (BTM< btm_10pctile)
153    gen hpctile_btm = (BTM> btm_90pctile)
154
155    *Average stock return for 90th and 10th percentile
156    bysort fyear:egen average2_ret90 = mean(r) if hpctile_btm ==1
157    bysort fyear:egen average2_ret10 = mean(r) if lpctile_btm ==1
158
159    bysort fyear: egen r_low = min(average2_ret90)
160    bysort fyear: egen r_H = min(average2_ret10)
161
162    gen hml = r_H-r_low
163    sum hml
164
165    *2.8 Multivariate regression
166
167    reg excess_r excess_mar_r sml hml
168    outreg2 using "docs/multi_reg.xls", ctitle(True model) replace
169    *the excess market return coef becomes more .this change is due
       to the omitted
170    *variable sml and hml which was added.
171
172
173    *2.9 Partition regression
174    *setp1:
175    reg excess_r sml hml
176    predict xr_i_hat
177    *Residual
178    gen residual1 = excess_r - xr_i_hat
179
180    *Step 2
181    reg excess_mar_r sml hml
182    predict xr_m_hat
183    *residual2
184    gen residual2 = excess_mar_r - xr_m_hat
185
186    *Step 3
187    reg residual1 residual2
188    outreg2 using "docs/multi_reg.xls", ctitle(partition regression)
       append
```

```
189   *coef = .0539182 (.0140033 )
190   /*comparison of results with :
191   -coef of regression from problem 2.5 coef= .0428587 (.0139741 )
192
193   The regression from 2.5 underestimate the true effect
194
195   -coef of regression from problem 2.8
196   The regression results from 2.8 and 2.9 are the same which makes
      sense.
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
```