

mkcorr and fre in Stata

SOC 561

Programming for the Social Sciences

Hyungjun Suh

Feb. 17. 2016



mkcorr

1) What is it and what does it do?

- ▶ It produces a publication-ready correlation table and a descriptive statistics table.
- ▶ Thus, it reduces working time and errors.
- ▶ It enables researchers to figure out the variance-covariance structure of variables.

mkcorr

2) How to install

```
. ssc install mkcorr
checking mkcorr consistency and verifying not already installed...
installing into c:\ado\plus\...
installation complete.
```

3

mkcorr

3) Basic Syntax

- ▶ **mkcorr** varlist [**if** exp] [**in** range], log(filename)
[replace means nocorr sig lab num cdec(#)
mdec(#) casewise]

4

mkcorr

4) Options and Examples

- ▶ sysuse nlsw88
- ▶ mkcorr age grade wage tenure, log(mkcorr) replace

	A	B	C	D	E
1		age	grade	wage	tenure
2	age	1			
3	grade	-0.03	1		
4	wage	-0.04	0.33	1	
5	tenure	0.07	0.12	0.18	1

5

mkcorr

4) Options and Examples

- ▶ mkcorr age grade wage tenure, log(mkcorr1) replace num label

	A	B	C	D	E	F
1			-1	-2	-3	-4
2	-1	age in current year	1			
3	-2	current grade completed	-0.03	1		
4	-3	hourly wage	-0.04	0.33	1	
5	-4	job tenure (years)	0.07	0.12	0.18	1

6

mkcorr

4) Options and Examples

- ▶ `mkcorr age grade wage tenure, log(mkcorr2) replace means mdec(2) cdec(3)`

	A	B	C	D	E	F	G	H	I
1		Mean	S.D.	Min	Max	age	grade	wage	tenure
2	age	39.15	3.06	34	46	1			
3	grade	13.1	2.52	0	18	-0.033	1		
4	wage	7.77	5.76	1	40.75	-0.036	0.325	1	
5	tenure	5.98	5.51	0	25.92	0.074	0.123	0.178	1

7

mkcorr

4) Options and Examples

- ▶ `mkcorr age grade wage tenure, log(mkcorr3) replace means nocorr`

	A	B	C	D	E
1		Mean	S.D.	Min	Max
2	age	39.153	3.06	34	46
3	grade	13.099	2.521	0	18
4	wage	7.767	5.756	1.005	40.747
5	tenure	5.978	5.51	0	25.917

8

mkcorr

4) Options and Examples

- ▶ `mkcorr age grade wage tenure, log(mkcorr4) replace sig`

	A	B	C	D	E
1		age	grade	wage	tenure
2	age	1			
3					
4	grade	-0.03	1		
5		-0.12			
6	wage	-0.04	0.33	1	
7		-0.09	0		
8	tenure	0.07	0.12	0.18	1
9		0	0	0	

9

mkcorr

4) Options and Examples

- ▶ `mkcorr age grade wage tenure, log(mkcorr5) replace casewise`

	A	B	C	D	E
1		age	grade	wage	tenure
2	age	1			
3	grade	-0.03	1		
4	wage	-0.04	0.33	1	
5	tenure	0.07	0.12	0.18	1

10

mkcorr

4) Options and Examples

- ▶ `mkcorr age grade wage tenure if south==1, log(mkcorr6) replace`

	A	B	C	D	E
1		age	grade	wage	tenure
2	age	1			
3	grade	-0.03	1		
4	wage	-0.04	0.33	1	
5	tenure	0.07	0.12	0.18	1

overall

	A	B	C	D	E
1		age	grade	wage	tenure
2	age	1			
3	grade	-0.05	1		
4	wage	-0.03	0.37	1	
5	tenure	0.08	0.15	0.13	1

South==1

11

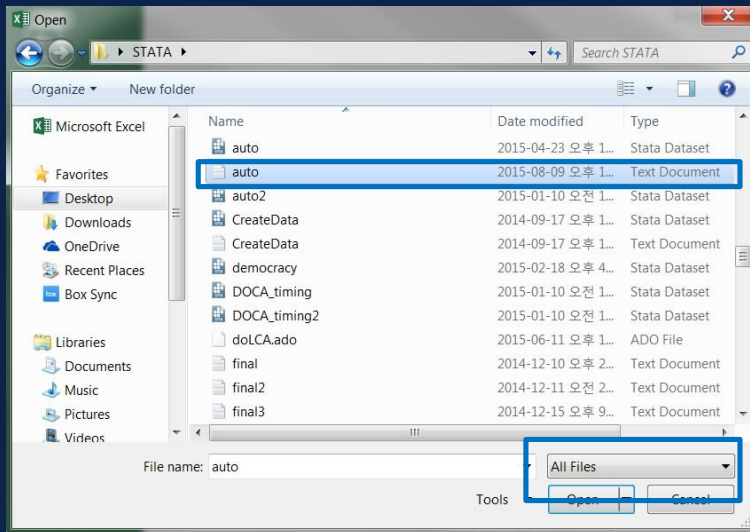
mkcorr

Caveat: Commands that do not work

- ▶ `bysort south: mkcorr age grade wage tenure, log(mkcorr7) replace`
- ▶ `mkcorr age grade wage tenure, log(mkcorr8) replace by(south)`
- ▶ `mkcorr age grade wage tenure if south==0, log(mkcorr6) append`

12

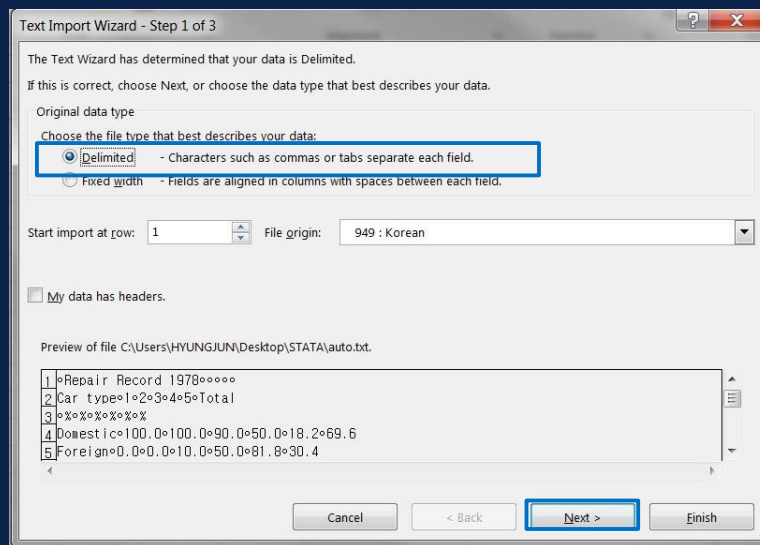
- Export Result Files to MS Excel



Choose 'all files' for the file type and open the result txt file.

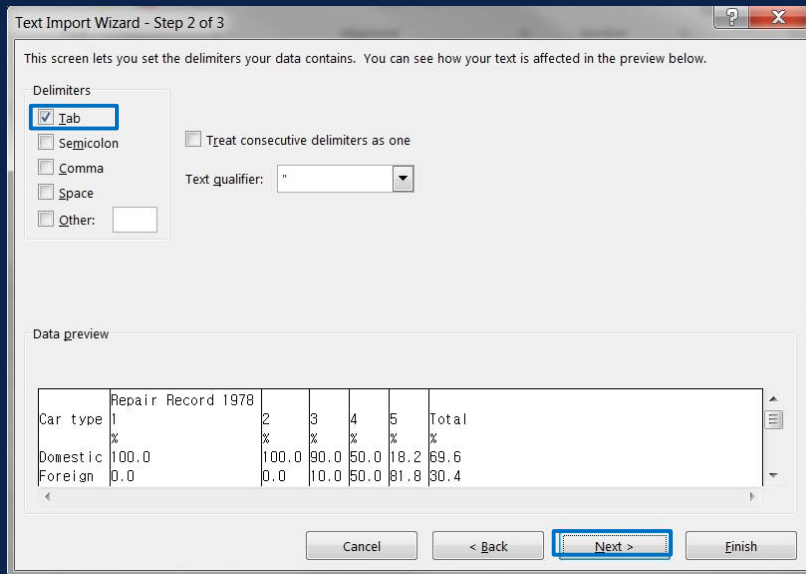
13

- Export Result Files to MS Excel (cont.)



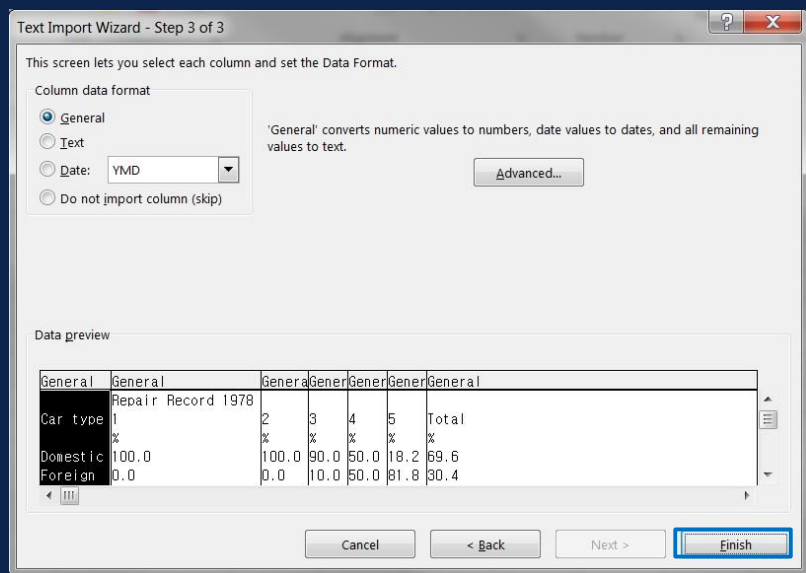
14

- Export Result Files to MS Excel (cont.)



15

- Export Result Files to MS Excel (cont.)



16

- Export Result Files to MS Excel (cont.)

The screenshot shows an Excel spreadsheet with the following data:

Repair Record 1978						
Car type	1	2	3	4	5	Total
%	%	%	%	%	%	%
Domestic	100	100	90	50	18.2	69.6
Foreign	0	0	10	50	81.8	30.4
Total	100	100	100	100	100	100

Car type	Mean price	Mean mpg
Domestic	6072.42	19.83
Foreign	6384.68	24.77
Total	6165.26	21.3

Car type	Mean headroom	Mean trunk
Domestic	3.15	14.75
Foreign	2.61	11.41
Total	2.99	13.76

17

fre

1) What is it and what does it do?

- ▶ It enhances `tab1` command, so that one can reduce errors.
- ▶ It produces a publication-ready frequency table (tab-delimited or LaTeX-formatted).

18

Problems of tab1

```
. tab1 jobstat3
```

```
-> tabulation of jobstat3
```

Job Status (three cat)	Freq.	Percent	Cum.
Employed	4,380	45.89	45.89
Unemployed	1,094	11.46	57.35
others	4,071	42.65	100.00
Total	9,545	100.00	

Valid percentage only

No missing values unless specified

Values are not shown

19

fre

```
. fre jobstat3
```

```
jobstat3 — Job Status (three cat)
```

	Freq.	Percent	Valid	Cum.
Valid 1 Employed	4380	45.57	45.89	45.89
2 Unemployed	1094	11.38	11.46	57.35
3 others	4071	42.36	42.65	100.00
Total	9545	99.31	100.00	
Missing .	66	0.69		
Total	9611	100.00		

20

fre

```
. fre stabili
```

stabili — the period of living in current housing

		Freq.	Percent	Valid	Cum.
Valid	0	128	1.33	1.37	1.37
	1	661	6.88	7.05	8.42
	2	485	5.05	5.17	13.59
	3	374	3.89	3.99	17.58
	4	417	4.34	4.45	22.02
	5	496	5.16	5.29	27.31
	6	263	2.74	2.81	30.12
	7	280	2.91	2.99	33.11
	8	244	2.54	2.60	35.71
	9	235	2.45	2.51	38.21
	10	291	3.03	3.10	41.32
	11	184	1.91	1.96	43.28
	12	189	1.97	2.02	45.30
	13	172	1.79	1.83	47.13
	14	183	1.90	1.95	49.08
	15	356	3.70	3.80	52.88
	16	127	1.32	1.35	54.23
	17	182	1.89	1.94	56.18
	18	241	2.51	2.57	58.75
	19	223	2.32	2.38	61.12
	:	:	:	:	:
	37	74	0.77	0.79	92.24

21

fre

2) How to install

```
. ssc install fre
checking fre consistency and verifying not already installed...
```

22

fre

3) Basic Syntax

- ▶ **fre** varlist [**if** exp] [**in** exp] [weight], options
- ▶ **fre** varlist using filename [**if** exp] [**in** exp] [weight], options export_options

23

fre

4) Options and Examples

```
. fre age
age — age in current year
```

		Freq.	Percent	Valid	Cum.
Valid	34	53	2.36	2.36	2.36
	35	260	11.58	11.58	13.94
	36	257	11.44	11.44	25.38
	37	225	10.02	10.02	35.40
	38	219	9.75	9.75	45.15
	39	234	10.42	10.42	55.57
	40	208	9.26	9.26	64.83
	41	222	9.88	9.88	74.71
	42	160	7.12	7.12	81.83
	43	165	7.35	7.35	89.18
	44	163	7.26	7.26	96.44
	45	78	3.47	3.47	99.91
	46	2	0.09	0.09	100.00
	Total	2246	100.00	100.00	

24

fre

4) Options and Examples

```
. fre age, subpop(south)
age — age in current year
```

		Freq.	Percent	Valid	Cum.
Valid	34	25	2.65	2.65	2.65
	35	116	12.31	12.31	14.97
	36	105	11.15	11.15	26.11
	37	95	10.08	10.08	36.20
	38	83	8.81	8.81	45.01
	39	90	9.55	9.55	54.56
	40	80	8.49	8.49	63.06
	41	96	10.19	10.19	73.25
	42	75	7.96	7.96	81.21
	43	67	7.11	7.11	88.32
	44	80	8.49	8.49	96.82
	45	30	3.18	3.18	100.00
	46	0	0.00	0.00	100.00
	Total	942	100.00	100.00	

Same with the command "fre age if south==1"

25

fre

4) Options and Examples

```
. label define south 1 "This long label will be wrapped", modify
. label value south south
. fre south, width(15) include (.a .b .c)
south — lives in south
```

		Freq.	Percent	Valid	Cum.
Valid	0	1304	58.06	58.06	58.06
	1 This long label will be wrapped	942	41.94	41.94	100.00
	Total	2246	100.00	100.00	
Missing	.a	0	0.00		
	.b	0	0.00		
	.c	0	0.00		
	Total	0	0.00		
Total		2246	100.00		

26

fre

4) Options and Examples

```
. fre south using south.tex
(output written to south.tex)

. type south.tex

south --- lives in south

\begin{tabular}{lllrrrr}\hline
& & & & Freq. & Percent & Valid & Cum. \\ \hline
Valid & 0 & & & 1304 & 58.06 & 58.06 & 58.06 \\
& 1 & This long label will be wrapped & & 942 & 41.94 & 41.94 & 100.00 \\
& Total & & & 2246 & 100.00 & 100.00 & \\ \hline\end{tabular}
```

27

fre

4) Options and Examples

```
. fre south using south.txt
(output written to south.txt)

. type south.txt

south  lives in south

Valid  0          Freq.  Percent Valid  Cum.
       1          This long label will be wrapped 942  41.94  41.94  100.00
Total  2246       100.00  100.00
```

28

fre

4) Options and Examples – ‘by’ option

```
. bysort south: fre age
```

```
-> south = 0
```

age — age in current year

		Freq.	Percent	Valid	Cum.
Valid	34	28	2.15	2.15	2.15
	35	144	11.04	11.04	13.19
	36	152	11.66	11.66	24.85
	37	130	9.97	9.97	34.82
	38	136	10.43	10.43	45.25
	39	144	11.04	11.04	56.29
	40	128	9.82	9.82	66.10
	41	126	9.66	9.66	75.77
	42	85	6.52	6.52	82.29
	43	98	7.52	7.52	89.80
	44	83	6.37	6.37	96.17
	45	48	3.68	3.68	99.85
	46	2	0.15	0.15	100.00
Total		1304	100.00	100.00	

```
-> south = This long label will be wrapped
```

age — age in current year

		Freq.	Percent	Valid	Cum.
Valid	34	25	2.65	2.65	2.65
	35	116	12.31	12.31	14.97
	36	105	11.15	11.15	26.11
	37	95	10.08	10.08	36.20
	38	83	8.81	8.81	45.01
	39	90	9.55	9.55	54.56
	40	80	8.49	8.49	63.06
	41	96	10.19	10.19	73.25
	42	75	7.96	7.96	81.21
	43	67	7.11	7.11	88.32
	44	80	8.49	8.49	96.82
	45	30	3.18	3.18	100.00
Total		942	100.00	100.00	

29

fre

4) Options and Examples

- ▶ fre south using south.txt, replace

	A	B	C	D	E	F	G
1							
2	south	lives in south					
3							
4				Freq.	Percent	Valid	Cum.
5	Valid	0 Non-South	1304	58.06	58.06	58.06	
6		1 South	942	41.94	41.94	100	
7		Total	2246	100	100		

30

fre

4) Options and Examples

- ▶ fre south using south1.txt, bodyonly

	A	B	C	D	E	F	G
1							
2	Valid	0	Non-South	1304	58.06	58.06	58.06
3		1	South	942	41.94	41.94	100
4		Total		2246	100	100	

31

fre

4) Options and Examples

- ▶ fre south using south2.txt, combine

	A	B	C	D	E	F
1						
2	south	lives in south				
3						
4			Freq.	Percent	Valid	Cum.
5	Valid	0	Non-South	1304	58.06	58.06
6		1	South	942	41.94	100
7		Total		2246	100	100

32

fre

4) Options and Examples

- ▶ fre south using south3.txt, pre(descriptive statistics)

	A	B	C	D	E	F	G
1	descriptive	statistics					
2	south	lives in south					
3							
4				Freq.	Percent	Valid	Cum.
5	Valid	0 Non-South		1304	58.06	58.06	58.06
6		1 South		942	41.94	41.94	100
7		Total		2246	100	100	

33

fre

4) Options and Examples

- ▶ fre married using south.txt, append
- ▶ Using `order` and `append`, one can make as many number of tables as one wants!

	A	B	C	D	E	F	G
1							
2	south	lives in south					
3							
4				Freq.	Percent	Valid	Cum.
5	Valid	0 Non-South		1304	58.06	58.06	58.06
6		1 South		942	41.94	41.94	100
7		Total		2246	100	100	
8							
9							
10							
11	married	married					
12							
13				Freq.	Percent	Valid	Cum.
14	Valid	0 single		804	35.8	35.8	35.8
15		1 married		1442	64.2	64.2	100
16		Total		2246	100	100	

34

fre

4) Options and Examples

- ▶ order_all, alphabetic
- ▶ fre age-wage using nlsw.txt

35

תודה
 Dankie Gracias
 Спасибо
 شكرًا
 Köszönjük
 Merci Takk
 Grazie Dziękujemy Terima kasih
 Ďakujeme Vielen Dank Paldies
 Kiitos Täname teid 谢谢
Thank You Tak
 感谢您 Obrigado Teşekkür Ederiz
 Σας Ευχαριστούμ 감사합니다
 Bedankt Děkujeme vám
 ありがとうございます
 Tack