

```
-----
name: <unnamed>
log: d:\temp\data management.log
log type: text
opened on: 25 Jan 2021, 10:52:36
```

```
.
.
. *****
. * Test some Stata commands
. *****
. clear

. set maxvar 10000 // set the maxium number of variables to be read in Stata

.
.
. *set more on //Tell Stata to pause for --more-- messages
. set more off //Tell Stata not to pause for --more-- messages

.
. *****
. * Open the data file
. *****
. use d:\temp\auto.dta, clear
(1978 Automobile Data)

.
.
. *****
. * The list command
. *****
. list
```

	make	price	mpg	rep78	headroom	trunk	weight	length	turn	displa~t	gear_r~o	for
1.	AMC Concord	4,099	22	3	2.5	11	2,930	186	40	121	3.58	Dome
2.	AMC Pacer	4,749	17	3	3.0	11	3,350	173	40	258	2.53	Dome
3.	AMC Spirit	3,799	22	.	3.0	12	2,640	168	35	121	3.08	Dome
4.	Buick Century	4,816	20	3	4.5	16	3,250	196	40	196	2.93	Dome
5.	Buick Electra	7,827	15	4	4.0	20	4,080	222	43	350	2.41	Dome
6.	Buick LeSabre	5,788	18	3	4.0	21	3,670	218	43	231	2.73	Dome
7.	Buick Opel	4,453	26	.	3.0	10	2,230	170	34	304	2.87	Dome
8.	Buick Regal	5,189	20	3	2.0	16	3,280	200	42	196	2.93	Dome
9.	Buick Riviera	10,372	16	3	3.5	17	3,880	207	43	231	2.93	Dome
10.	Buick Skylark	4,082	19	3	3.5	13	3,400	200	42	231	3.08	Dome
11.	Cad. Deville	11,385	14	3	4.0	20	4,330	221	44	425	2.28	Dome
12.	Cad. Eldorado	14,500	14	2	3.5	16	3,900	204	43	350	2.19	Dome
13.	Cad. Seville	15,906	21	3	3.0	13	4,290	204	45	350	2.24	Dome
14.	Chev. Chevette	3,299	29	3	2.5	9	2,110	163	34	231	2.93	Dome
15.	Chev. Impala	5,705	16	4	4.0	20	3,690	212	43	250	2.56	Dome
16.	Chev. Malibu	4,504	22	3	3.5	17	3,180	193	31	200	2.73	Dome
17.	Chev. Monte Carlo	5,104	22	2	2.0	16	3,220	200	41	200	2.73	Dome
18.	Chev. Monza	3,667	24	2	2.0	7	2,750	179	40	151	2.73	Dome
19.	Chev. Nova	3,955	19	3	3.5	13	3,430	197	43	250	2.56	Dome
20.	Dodge Colt	3,984	30	5	2.0	8	2,120	163	35	98	3.54	Dome
21.	Dodge Diplomat	4,010	18	2	4.0	17	3,600	206	46	318	2.47	Dome
22.	Dodge Magnum	5,886	16	2	4.0	17	3,600	206	46	318	2.47	Dome
23.	Dodge St. Regis	6,342	17	2	4.5	21	3,740	220	46	225	2.94	Dome
24.	Ford Fiesta	4,389	28	4	1.5	9	1,800	147	33	98	3.15	Dome
25.	Ford Mustang	4,187	21	3	2.0	10	2,650	179	43	140	3.08	Dome
26.	Linc. Continental	11,497	12	3	3.5	22	4,840	233	51	400	2.47	Dome
27.	Linc. Mark V	13,594	12	3	2.5	18	4,720	230	48	400	2.47	Dome
28.	Linc. Versailles	13,466	14	3	3.5	15	3,830	201	41	302	2.47	Dome
29.	Merc. Bobcat	3,829	22	4	3.0	9	2,580	169	39	140	2.73	Dome
30.	Merc. Cougar	5,379	14	4	3.5	16	4,060	221	48	302	2.75	Dome
31.	Merc. Marquis	6,165	15	3	3.5	23	3,720	212	44	302	2.26	Dome
32.	Merc. Monarch	4,516	18	3	3.0	15	3,370	198	41	250	2.43	Dome
33.	Merc. XR-7	6,303	14	4	3.0	16	4,130	217	45	302	2.75	Dome
34.	Merc. Zephyr	3,291	20	3	3.5	17	2,830	195	43	140	3.08	Dome
35.	Olds 98	8,814	21	4	4.0	20	4,060	220	43	350	2.41	Dome
36.	Olds Cutl Supr	5,172	19	3	2.0	16	3,310	198	42	231	2.93	Dome
37.	Olds Cutlass	4,733	19	3	4.5	16	3,300	198	42	231	2.93	Dome
38.	Olds Delta 88	4,890	18	4	4.0	20	3,690	218	42	231	2.73	Dome

39.	Olds Omega	4,181	19	3	4.5	14	3,370	200	43	231	3.08	Dome
40.	Olds Starfire	4,195	24	1	2.0	10	2,730	180	40	151	2.73	Dome
41.	Olds Toronado	10,371	16	3	3.5	17	4,030	206	43	350	2.41	Dome
42.	Plym. Arrow	4,647	28	3	2.0	11	3,260	170	37	156	3.05	Dome
43.	Plym. Champ	4,425	34	5	2.5	11	1,800	157	37	86	2.97	Dome
44.	Plym. Horizon	4,482	25	3	4.0	17	2,200	165	36	105	3.37	Dome
45.	Plym. Sapporo	6,486	26	.	1.5	8	2,520	182	38	119	3.54	Dome
46.	Plym. Volare	4,060	18	2	5.0	16	3,330	201	44	225	3.23	Dome
47.	Pont. Catalina	5,798	18	4	4.0	20	3,700	214	42	231	2.73	Dome
48.	Pont. Firebird	4,934	18	1	1.5	7	3,470	198	42	231	3.08	Dome
49.	Pont. Grand Prix	5,222	19	3	2.0	16	3,210	201	45	231	2.93	Dome
50.	Pont. Le Mans	4,723	19	3	3.5	17	3,200	199	40	231	2.93	Dome
51.	Pont. Phoenix	4,424	19	.	3.5	13	3,420	203	43	231	3.08	Dome
52.	Pont. Sunbird	4,172	24	2	2.0	7	2,690	179	41	151	2.73	Dome
53.	Audi 5000	9,690	17	5	3.0	15	2,830	189	37	131	3.20	For
54.	Audi Fox	6,295	23	3	2.5	11	2,070	174	36	97	3.70	For
55.	BMW 320i	9,735	25	4	2.5	12	2,650	177	34	121	3.64	For
56.	Datsun 200	6,229	23	4	1.5	6	2,370	170	35	119	3.89	For
57.	Datsun 210	4,589	35	5	2.0	8	2,020	165	32	85	3.70	For
58.	Datsun 510	5,079	24	4	2.5	8	2,280	170	34	119	3.54	For
59.	Datsun 810	8,129	21	4	2.5	8	2,750	184	38	146	3.55	For
60.	Fiat Strada	4,296	21	3	2.5	16	2,130	161	36	105	3.37	For
61.	Honda Accord	5,799	25	5	3.0	10	2,240	172	36	107	3.05	For
62.	Honda Civic	4,499	28	4	2.5	5	1,760	149	34	91	3.30	For
63.	Mazda GLC	3,995	30	4	3.5	11	1,980	154	33	86	3.73	For
64.	Peugeot 604	12,990	14	.	3.5	14	3,420	192	38	163	3.58	For
65.	Renault Le Car	3,895	26	3	3.0	10	1,830	142	34	79	3.72	For
66.	Subaru	3,798	35	5	2.5	11	2,050	164	36	97	3.81	For
67.	Toyota Celica	5,899	18	5	2.5	14	2,410	174	36	134	3.06	For
68.	Toyota Corolla	3,748	31	5	3.0	9	2,200	165	35	97	3.21	For
69.	Toyota Corona	5,719	18	5	2.0	11	2,670	175	36	134	3.05	For
70.	VW Dasher	7,140	23	4	2.5	12	2,160	172	36	97	3.74	For
71.	VW Diesel	5,397	41	5	3.0	15	2,040	155	35	90	3.78	For
72.	VW Rabbit	4,697	25	4	3.0	15	1,930	155	35	89	3.78	For
73.	VW Scirocco	6,850	25	4	2.0	16	1,990	156	36	97	3.78	For
74.	Volvo 260	11,995	17	5	2.5	14	3,170	193	37	163	2.98	For

. list if foreign ==1

	make	price	mpg	rep78	headroom	trunk	weight	length	turn	displa~t	gear_r~o	foreign
53.	Audi 5000	9,690	17	5	3.0	15	2,830	189	37	131	3.20	Foreign
54.	Audi Fox	6,295	23	3	2.5	11	2,070	174	36	97	3.70	Foreign
55.	BMW 320i	9,735	25	4	2.5	12	2,650	177	34	121	3.64	Foreign
56.	Datsun 200	6,229	23	4	1.5	6	2,370	170	35	119	3.89	Foreign
57.	Datsun 210	4,589	35	5	2.0	8	2,020	165	32	85	3.70	Foreign
58.	Datsun 510	5,079	24	4	2.5	8	2,280	170	34	119	3.54	Foreign
59.	Datsun 810	8,129	21	4	2.5	8	2,750	184	38	146	3.55	Foreign
60.	Fiat Strada	4,296	21	3	2.5	16	2,130	161	36	105	3.37	Foreign
61.	Honda Accord	5,799	25	5	3.0	10	2,240	172	36	107	3.05	Foreign
62.	Honda Civic	4,499	28	4	2.5	5	1,760	149	34	91	3.30	Foreign
63.	Mazda GLC	3,995	30	4	3.5	11	1,980	154	33	86	3.73	Foreign
64.	Peugeot 604	12,990	14	.	3.5	14	3,420	192	38	163	3.58	Foreign
65.	Renault Le Car	3,895	26	3	3.0	10	1,830	142	34	79	3.72	Foreign
66.	Subaru	3,798	35	5	2.5	11	2,050	164	36	97	3.81	Foreign
67.	Toyota Celica	5,899	18	5	2.5	14	2,410	174	36	134	3.06	Foreign
68.	Toyota Corolla	3,748	31	5	3.0	9	2,200	165	35	97	3.21	Foreign
69.	Toyota Corona	5,719	18	5	2.0	11	2,670	175	36	134	3.05	Foreign
70.	VW Dasher	7,140	23	4	2.5	12	2,160	172	36	97	3.74	Foreign
71.	VW Diesel	5,397	41	5	3.0	15	2,040	155	35	90	3.78	Foreign
72.	VW Rabbit	4,697	25	4	3.0	15	1,930	155	35	89	3.78	Foreign
73.	VW Scirocco	6,850	25	4	2.0	16	1,990	156	36	97	3.78	Foreign
74.	Volvo 260	11,995	17	5	2.5	14	3,170	193	37	163	2.98	Foreign

. list if inlist(rep78,1,2,3)

	make	price	mpg	rep78	headroom	trunk	weight	length	turn	displa~t	gear_r~o	for
1.	AMC Concord	4,099	22	3	2.5	11	2,930	186	40	121	3.58	Dome

2.	AMC Pacer	4,749	17	3	3.0	11	3,350	173	40	258	2.53	Dome
4.	Buick Century	4,816	20	3	4.5	16	3,250	196	40	196	2.93	Dome
6.	Buick LeSabre	5,788	18	3	4.0	21	3,670	218	43	231	2.73	Dome
8.	Buick Regal	5,189	20	3	2.0	16	3,280	200	42	196	2.93	Dome
9.	Buick Riviera	10,372	16	3	3.5	17	3,880	207	43	231	2.93	Dome
10.	Buick Skylark	4,082	19	3	3.5	13	3,400	200	42	231	3.08	Dome
11.	Cad. Deville	11,385	14	3	4.0	20	4,330	221	44	425	2.28	Dome
12.	Cad. Eldorado	14,500	14	2	3.5	16	3,900	204	43	350	2.19	Dome
13.	Cad. Seville	15,906	21	3	3.0	13	4,290	204	45	350	2.24	Dome
14.	Chev. Chevette	3,299	29	3	2.5	9	2,110	163	34	231	2.93	Dome
16.	Chev. Malibu	4,504	22	3	3.5	17	3,180	193	31	200	2.73	Dome
17.	Chev. Monte Carlo	5,104	22	2	2.0	16	3,220	200	41	200	2.73	Dome
18.	Chev. Monza	3,667	24	2	2.0	7	2,750	179	40	151	2.73	Dome
19.	Chev. Nova	3,955	19	3	3.5	13	3,430	197	43	250	2.56	Dome
21.	Dodge Diplomat	4,010	18	2	4.0	17	3,600	206	46	318	2.47	Dome
22.	Dodge Magnum	5,886	16	2	4.0	17	3,600	206	46	318	2.47	Dome
23.	Dodge St. Regis	6,342	17	2	4.5	21	3,740	220	46	225	2.94	Dome
25.	Ford Mustang	4,187	21	3	2.0	10	2,650	179	43	140	3.08	Dome
26.	Linc. Continental	11,497	12	3	3.5	22	4,840	233	51	400	2.47	Dome
27.	Linc. Mark V	13,594	12	3	2.5	18	4,720	230	48	400	2.47	Dome
28.	Linc. Versailles	13,466	14	3	3.5	15	3,830	201	41	302	2.47	Dome
31.	Merc. Marquis	6,165	15	3	3.5	23	3,720	212	44	302	2.26	Dome
32.	Merc. Monarch	4,516	18	3	3.0	15	3,370	198	41	250	2.43	Dome
34.	Merc. Zephyr	3,291	20	3	3.5	17	2,830	195	43	140	3.08	Dome
36.	Olds Cutl Supr	5,172	19	3	2.0	16	3,310	198	42	231	2.93	Dome
37.	Olds Cutlass	4,733	19	3	4.5	16	3,300	198	42	231	2.93	Dome
39.	Olds Omega	4,181	19	3	4.5	14	3,370	200	43	231	3.08	Dome
40.	Olds Starfire	4,195	24	1	2.0	10	2,730	180	40	151	2.73	Dome
41.	Olds Toronado	10,371	16	3	3.5	17	4,030	206	43	350	2.41	Dome
42.	Plym. Arrow	4,647	28	3	2.0	11	3,260	170	37	156	3.05	Dome
44.	Plym. Horizon	4,482	25	3	4.0	17	2,200	165	36	105	3.37	Dome
46.	Plym. Volare	4,060	18	2	5.0	16	3,330	201	44	225	3.23	Dome
48.	Pont. Firebird	4,934	18	1	1.5	7	3,470	198	42	231	3.08	Dome
49.	Pont. Grand Prix	5,222	19	3	2.0	16	3,210	201	45	231	2.93	Dome
50.	Pont. Le Mans	4,723	19	3	3.5	17	3,200	199	40	231	2.93	Dome
52.	Pont. Sunbird	4,172	24	2	2.0	7	2,690	179	41	151	2.73	Dome
54.	Audi Fox	6,295	23	3	2.5	11	2,070	174	36	97	3.70	For
60.	Fiat Strada	4,296	21	3	2.5	16	2,130	161	36	105	3.37	For
65.	Renault Le Car	3,895	26	3	3.0	10	1,830	142	34	79	3.72	For

. list if inrange(rep78,1,3)

	make	price	mpg	rep78	headroom	trunk	weight	length	turn	displa~t	gear_r~o	for
1.	AMC Concord	4,099	22	3	2.5	11	2,930	186	40	121	3.58	Dome
2.	AMC Pacer	4,749	17	3	3.0	11	3,350	173	40	258	2.53	Dome
4.	Buick Century	4,816	20	3	4.5	16	3,250	196	40	196	2.93	Dome
6.	Buick LeSabre	5,788	18	3	4.0	21	3,670	218	43	231	2.73	Dome
8.	Buick Regal	5,189	20	3	2.0	16	3,280	200	42	196	2.93	Dome
9.	Buick Riviera	10,372	16	3	3.5	17	3,880	207	43	231	2.93	Dome
10.	Buick Skylark	4,082	19	3	3.5	13	3,400	200	42	231	3.08	Dome
11.	Cad. Deville	11,385	14	3	4.0	20	4,330	221	44	425	2.28	Dome
12.	Cad. Eldorado	14,500	14	2	3.5	16	3,900	204	43	350	2.19	Dome
13.	Cad. Seville	15,906	21	3	3.0	13	4,290	204	45	350	2.24	Dome
14.	Chev. Chevette	3,299	29	3	2.5	9	2,110	163	34	231	2.93	Dome
16.	Chev. Malibu	4,504	22	3	3.5	17	3,180	193	31	200	2.73	Dome
17.	Chev. Monte Carlo	5,104	22	2	2.0	16	3,220	200	41	200	2.73	Dome
18.	Chev. Monza	3,667	24	2	2.0	7	2,750	179	40	151	2.73	Dome
19.	Chev. Nova	3,955	19	3	3.5	13	3,430	197	43	250	2.56	Dome
21.	Dodge Diplomat	4,010	18	2	4.0	17	3,600	206	46	318	2.47	Dome
22.	Dodge Magnum	5,886	16	2	4.0	17	3,600	206	46	318	2.47	Dome
23.	Dodge St. Regis	6,342	17	2	4.5	21	3,740	220	46	225	2.94	Dome
25.	Ford Mustang	4,187	21	3	2.0	10	2,650	179	43	140	3.08	Dome
26.	Linc. Continental	11,497	12	3	3.5	22	4,840	233	51	400	2.47	Dome
27.	Linc. Mark V	13,594	12	3	2.5	18	4,720	230	48	400	2.47	Dome
28.	Linc. Versailles	13,466	14	3	3.5	15	3,830	201	41	302	2.47	Dome
31.	Merc. Marquis	6,165	15	3	3.5	23	3,720	212	44	302	2.26	Dome
32.	Merc. Monarch	4,516	18	3	3.0	15	3,370	198	41	250	2.43	Dome
34.	Merc. Zephyr	3,291	20	3	3.5	17	2,830	195	43	140	3.08	Dome
36.	Olds Cutl Supr	5,172	19	3	2.0	16	3,310	198	42	231	2.93	Dome

37.	Olds Cutlass	4,733	19	3	4.5	16	3,300	198	42	231	2.93	Dome
39.	Olds Omega	4,181	19	3	4.5	14	3,370	200	43	231	3.08	Dome
40.	Olds Starfire	4,195	24	1	2.0	10	2,730	180	40	151	2.73	Dome
41.	Olds Toronado	10,371	16	3	3.5	17	4,030	206	43	350	2.41	Dome
-----												
42.	Plym. Arrow	4,647	28	3	2.0	11	3,260	170	37	156	3.05	Dome
44.	Plym. Horizon	4,482	25	3	4.0	17	2,200	165	36	105	3.37	Dome
46.	Plym. Volare	4,060	18	2	5.0	16	3,330	201	44	225	3.23	Dome
48.	Pont. Firebird	4,934	18	1	1.5	7	3,470	198	42	231	3.08	Dome
49.	Pont. Grand Prix	5,222	19	3	2.0	16	3,210	201	45	231	2.93	Dome
-----												
50.	Pont. Le Mans	4,723	19	3	3.5	17	3,200	199	40	231	2.93	Dome
52.	Pont. Sunbird	4,172	24	2	2.0	7	2,690	179	41	151	2.73	Dome
54.	Audi Fox	6,295	23	3	2.5	11	2,070	174	36	97	3.70	For
60.	Fiat Strada	4,296	21	3	2.5	16	2,130	161	36	105	3.37	For
65.	Renault Le Car	3,895	26	3	3.0	10	1,830	142	34	79	3.72	For

```
.
.
. *****
. * Commands for managing dta sets
. *****
.
```

```
. use d:\temp\auto.dta, clear
(1978 Automobile Data)
```

```
. des
```

```
Contains data from d:\temp\auto.dta
```

```
obs:      74      1978 Automobile Data
vars:     12      24 Jan 2021 18:30
size:    3,182    (_dta has notes)
```

variable name	storage type	display format	value label	variable label
make	str18	%-18s		Make and Model
price	int	%8.0gc		Price
mpg	int	%8.0g		Mileage (mpg)
rep78	int	%8.0g		Repair Record 1978
headroom	float	%6.1f		Headroom (in.)
trunk	int	%8.0g		Trunk space (cu. ft.)
weight	int	%8.0gc		Weight (lbs.)
length	int	%8.0g		Length (in.)
turn	int	%8.0g		Turn Circle (ft.)
displacement	int	%8.0g		Displacement (cu. in.)
gear_ratio	float	%6.2f		Gear Ratio
foreign	byte	%8.0g	origin	Car type

```
Sorted by: foreign
```

```
. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
make	0				
price	74	6165.257	2949.496	3291	15906
mpg	74	21.2973	5.785503	12	41
rep78	69	3.405797	.9899323	1	5
headroom	74	2.993243	.8459948	1.5	5
trunk	74	13.75676	4.277404	5	23
weight	74	3019.459	777.1936	1760	4840
length	74	187.9324	22.26634	142	233
turn	74	39.64865	4.399354	31	51
displacement	74	197.2973	91.83722	79	425
gear_ratio	74	3.014865	.4562871	2.19	3.89
foreign	74	.2972973	.4601885	0	1

```
. *****
. * Check duplicate records
. *****
.
```

```
. duplicates report foreign rep78
```

```
Duplicates in terms of foreign rep78
```

copies	observations	surplus
+		

1	1	0
2	4	2
3	3	2
4	4	3
8	8	7
9	27	24
27	27	26

```
-----
. duplicates tag foreign rep78, generate(same)
```

```
Duplicates in terms of foreign rep78
```

```
. label variable same "respondents have the same values of foreign and rep78"
```

```
. by foreign rep78, sort: gen n = _n
```

```
. by foreign rep78, sort: gen N = _N
```

```
. list foreign rep78 same n N, sepby(foreign rep78)
```

	foreign	rep78	same	n	N
1.	Domestic	1	1	1	2
2.	Domestic	1	1	2	2
3.	Domestic	2	7	1	8
4.	Domestic	2	7	2	8
5.	Domestic	2	7	3	8
6.	Domestic	2	7	4	8
7.	Domestic	2	7	5	8
8.	Domestic	2	7	6	8
9.	Domestic	2	7	7	8
10.	Domestic	2	7	8	8
11.	Domestic	3	26	1	27
12.	Domestic	3	26	2	27
13.	Domestic	3	26	3	27
14.	Domestic	3	26	4	27
15.	Domestic	3	26	5	27
16.	Domestic	3	26	6	27
17.	Domestic	3	26	7	27
18.	Domestic	3	26	8	27
19.	Domestic	3	26	9	27
20.	Domestic	3	26	10	27
21.	Domestic	3	26	11	27
22.	Domestic	3	26	12	27
23.	Domestic	3	26	13	27
24.	Domestic	3	26	14	27
25.	Domestic	3	26	15	27
26.	Domestic	3	26	16	27
27.	Domestic	3	26	17	27
28.	Domestic	3	26	18	27
29.	Domestic	3	26	19	27
30.	Domestic	3	26	20	27
31.	Domestic	3	26	21	27
32.	Domestic	3	26	22	27
33.	Domestic	3	26	23	27
34.	Domestic	3	26	24	27
35.	Domestic	3	26	25	27
36.	Domestic	3	26	26	27
37.	Domestic	3	26	27	27
38.	Domestic	4	8	1	9
39.	Domestic	4	8	2	9
40.	Domestic	4	8	3	9
41.	Domestic	4	8	4	9
42.	Domestic	4	8	5	9
43.	Domestic	4	8	6	9
44.	Domestic	4	8	7	9
45.	Domestic	4	8	8	9
46.	Domestic	4	8	9	9
47.	Domestic	5	1	1	2
48.	Domestic	5	1	2	2
49.	Domestic	.	3	1	4
50.	Domestic	.	3	2	4
51.	Domestic	.	3	3	4

```

52. | Domestic      .      3      4      4 |
    |-----|
53. | Foreign       3      2      1      3 |
54. | Foreign       3      2      2      3 |
55. | Foreign       3      2      3      3 |
    |-----|
56. | Foreign       4      8      1      9 |
57. | Foreign       4      8      2      9 |
58. | Foreign       4      8      3      9 |
59. | Foreign       4      8      4      9 |
60. | Foreign       4      8      5      9 |
61. | Foreign       4      8      6      9 |
62. | Foreign       4      8      7      9 |
63. | Foreign       4      8      8      9 |
64. | Foreign       4      8      9      9 |
    |-----|
65. | Foreign       5      8      1      9 |
66. | Foreign       5      8      2      9 |
67. | Foreign       5      8      3      9 |
68. | Foreign       5      8      4      9 |
69. | Foreign       5      8      5      9 |
70. | Foreign       5      8      6      9 |
71. | Foreign       5      8      7      9 |
72. | Foreign       5      8      8      9 |
73. | Foreign       5      8      9      9 |
    |-----|
74. | Foreign       .      0      1      1 |
    +-----+
    
```

```

.
.
. *****
. * Merge data
. *****
.
. *****
. * 1:1 match merge
. *****
.
.
. webuse autosize, clear
(1978 Automobile Data)
    
```

```

. list

+-----+
| make          weight   length |
+-----+
1. | Toyota Celica    2,410    174 |
2. | BMW 320i        2,650    177 |
3. | Cad. Seville    4,290    204 |
4. | Pont. Grand Prix 3,210    201 |
5. | Datsun 210      2,020    165 |
    |-----|
6. | Plym. Arrow     3,260    170 |
    +-----+
    
```

```

. webuse autoexpense, clear
(1978 Automobile Data)
    
```

```

. list

+-----+
| make          price    mpg |
+-----+
1. | Toyota Celica    5,899    18 |
2. | BMW 320i        9,735    25 |
3. | Cad. Seville    15,906    21 |
4. | Pont. Grand Prix 5,222    19 |
5. | Datsun 210      4,589    35 |
    +-----+
    
```

```

. webuse autosize, clear
(1978 Automobile Data)
    
```

```

. merge 1:1 make using http://www.stata-press.com/data/r15/autoexpense
    
```

```

Result          # of obs.
-----
not matched                1
  from master              1  (_merge==1)
  from using                0  (_merge==2)
    
```

matched 5 (\_merge==3)

. list

```

+-----+
| make          weight  length  price  mpg      _merge |
+-----+
1. | BMW 320i      2,650   177   9,735   25      matched (3) |
2. | Cad. Seville  4,290   204  15,906   21      matched (3) |
3. | Datsun 210    2,020   165   4,589   35      matched (3) |
4. | Plym. Arrow   3,260   170   .        .      master only (1) |
5. | Pont. Grand Prix 3,210   201   5,222   19      matched (3) |
+-----+
6. | Toyota Celica 2,410   174   5,899   18      matched (3) |
+-----+
    
```

```

. *****
. * m:1 match merge
. *****
.
. webuse dollars, clear
. (Regional Sales & Costs)
    
```

. list

```

+-----+
| region      sales    cost |
+-----+
1. | N Cntrl    419,472  227,677 |
2. | NE         360,523  138,097 |
3. | South      532,399  330,499 |
4. | West       310,565  165,348 |
+-----+
    
```

```

. webuse sforce, clear
. (Sales Force)
    
```

. list

```

+-----+
| region      name |
+-----+
1. | N Cntrl    Krantz |
2. | N Cntrl    Phipps |
3. | N Cntrl    Willis |
4. | NE         Ecklund |
5. | NE         Franks |
+-----+
6. | South      Anderson |
7. | South      Dubnoff |
8. | South      Lee |
9. | South      McNeil |
10. | West       Charles |
+-----+
11. | West       Cobb |
12. | West       Grant |
+-----+
    
```

```

. merge m:1 region using http://www.stata-press.com/data/r15/dollars
. (label region already defined)
    
```

```

Result          # of obs.
-----
not matched          0
matched             12 (_merge==3)
    
```

. list

```

+-----+
| region      name    sales    cost      _merge |
+-----+
1. | N Cntrl    Krantz  419,472  227,677  matched (3) |
2. | N Cntrl    Phipps  419,472  227,677  matched (3) |
3. | N Cntrl    Willis  419,472  227,677  matched (3) |
4. | NE         Ecklund  360,523  138,097  matched (3) |
5. | NE         Franks  360,523  138,097  matched (3) |
+-----+
6. | South      Anderson  532,399  330,499  matched (3) |
7. | South      Dubnoff  532,399  330,499  matched (3) |
8. | South      Lee      532,399  330,499  matched (3) |
+-----+
    
```

```

9. | South      McNeil  532,399  330,499  matched (3) |
10. | West       Charles 310,565  165,348  matched (3) |
-----+-----
11. | West       Cobb    310,565  165,348  matched (3) |
12. | West       Grant   310,565  165,348  matched (3) |
-----+-----

```

```

.
.
. *****
. * 1:m match merge
. *****
.
.
. webuse sforce, clear
(Sales Force)
. list

```

```

-----+-----
| region      name |
-----+-----
1. | N Cntrl    Krantz |
2. | N Cntrl    Phipps |
3. | N Cntrl    Willis |
4. | NE         Ecklund |
5. | NE         Franks |
-----+-----
6. | South      Anderson |
7. | South      Dubnoff |
8. | South      Lee |
9. | South      McNeil |
10. | West       Charles |
-----+-----
11. | West       Cobb |
12. | West       Grant |
-----+-----

```

```

.
. webuse dollars, clear
(Regional Sales & Costs)
. list

```

```

-----+-----+
| region      sales      cost |
-----+-----+
1. | N Cntrl    419,472  227,677 |
2. | NE         360,523  138,097 |
3. | South      532,399  330,499 |
4. | West       310,565  165,348 |
-----+-----+

```

```

.
. merge 1:m region using http://www.stata-press.com/data/r15/sforce
(label region already defined)

```

```

Result                                     # of obs.
-----+-----
not matched                                0
matched                                    12  (_merge==3)
-----+-----

```

```

. list
-----+-----+
| region      sales      cost      name      _merge |
-----+-----+
1. | N Cntrl    419,472  227,677  Krantz    matched (3) |
2. | NE         360,523  138,097  Ecklund   matched (3) |
3. | South      532,399  330,499  Anderson  matched (3) |
4. | West       310,565  165,348  Charles   matched (3) |
5. | N Cntrl    419,472  227,677  Phipps    matched (3) |
-----+-----+
6. | N Cntrl    419,472  227,677  Willis    matched (3) |
7. | NE         360,523  138,097  Franks    matched (3) |
8. | South      532,399  330,499  Dubnoff   matched (3) |
9. | South      532,399  330,499  Lee       matched (3) |
10. | South      532,399  330,499  McNeil    matched (3) |
-----+-----+
11. | West       310,565  165,348  Cobb      matched (3) |
12. | West       310,565  165,348  Grant     matched (3) |
-----+-----+

```



```
.
.
.
. *****
. * Reshape the data
. *****
.
. webuse reshapel, clear
```

```
. list
```

	id	sex	inc80	inc81	inc82	ue80	ue81	ue82
1.	1	0	5000	5500	6000	0	1	0
2.	2	1	2000	2200	3300	1	0	0
3.	3	0	3000	2000	1000	0	0	1

```
. reshape long inc ue, i(id) j(year)
(note: j = 80 81 82)
```

```
Data
```

	wide	->	long
Number of obs.	3	->	9
Number of variables	8	->	5
j variable (3 values)		->	year
xij variables:			
	inc80 inc81 inc82	->	inc
	ue80 ue81 ue82	->	ue

```
. list, sepby(id)
```

	id	year	sex	inc	ue
1.	1	80	0	5000	0
2.	1	81	0	5500	1
3.	1	82	0	6000	0
4.	2	80	1	2000	1
5.	2	81	1	2200	0
6.	2	82	1	3300	0
7.	3	80	0	3000	0
8.	3	81	0	2000	0
9.	3	82	0	1000	1

```
. reshape wide inc ue, i(id) j(year)
(note: j = 80 81 82)
```

```
Data
```

	long	->	wide
Number of obs.	9	->	3
Number of variables	5	->	8
j variable (3 values)	year	->	(dropped)
xij variables:			
	inc	->	inc80 inc81 inc82
	ue	->	ue80 ue81 ue82

```
. list, sepby(id)
```

	id	inc80	ue80	inc81	ue81	inc82	ue82	sex
1.	1	5000	0	5500	1	6000	0	0
2.	2	2000	1	2200	0	3300	0	1
3.	3	3000	0	2000	0	1000	1	0

```
. *****
. * Tabulations of frequencies
. *****
.
. webuse auto.dta, clear
(1978 Automobile Data)
```

```
. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
make	0				
price	74	6165.257	2949.496	3291	15906
mpg	74	21.2973	5.785503	12	41
rep78	69	3.405797	.9899323	1	5
headroom	74	2.993243	.8459948	1.5	5
trunk	74	13.75676	4.277404	5	23
weight	74	3019.459	777.1936	1760	4840
length	74	187.9324	22.26634	142	233
turn	74	39.64865	4.399354	31	51
displacement	74	197.2973	91.83722	79	425
gear_ratio	74	3.014865	.4562871	2.19	3.89
foreign	74	.2972973	.4601885	0	1

```
. tab1 rep78
-> tabulation of rep78
```

Repair Record 1978	Freq.	Percent	Cum.
1	2	2.90	2.90
2	8	11.59	14.49
3	30	43.48	57.97
4	18	26.09	84.06
5	11	15.94	100.00
Total	69	100.00	

```
. tab1 rep78, mis
-> tabulation of rep78
```

Repair Record 1978	Freq.	Percent	Cum.
1	2	2.70	2.70
2	8	10.81	13.51
3	30	40.54	54.05
4	18	24.32	78.38
5	11	14.86	93.24
.	5	6.76	100.00
Total	74	100.00	

```
.
.
. *****
. * Commands for Modifying Existing Variables
. *****
.
. webuse auto.dta, clear
(1978 Automobile Data)
```

```
.
.
. *****
. * rename variable
. *****
. des
```

```
Contains data from http://www.stata-press.com/data/r15/auto.dta
obs:      74      1978 Automobile Data
vars:     12      13 Apr 2016 17:45
size:    3,182    (_dta has notes)
```

variable name	storage type	display format	value label	variable label
make	str18	%-18s		Make and Model
price	int	%8.0gc		Price

```
mpg          int      %8.0g          Mileage (mpg)
rep78       int      %8.0g          Repair Record 1978
headroom    float    %6.1f          Headroom (in.)
trunk       int      %8.0g          Trunk space (cu. ft.)
weight      int      %8.0gc         Weight (lbs.)
length      int      %8.0g          Length (in.)
turn        int      %8.0g          Turn Circle (ft.)
displacement int    %8.0g          Displacement (cu. in.)
gear_ratio  float    %6.2f          Gear Ratio
foreign     byte     %8.0g          origin      Car type
```

Sorted by: foreign

```
. rename *, upper
```

```
. des
```

```
Contains data from http://www.stata-press.com/data/r15/auto.dta
  obs:          74          1978 Automobile Data
  vars:         12          13 Apr 2016 17:45
  size:        3,182        (_dta has notes)
```

```
-----
variable name  storage  display  value  variable label
              type   format   label
-----
MAKE          str18   %-18s    Make and Model
PRICE         int     %8.0gc   Price
MPG           int     %8.0g    Mileage (mpg)
REP78        int     %8.0g    Repair Record 1978
HEADROOM     float   %6.1f    Headroom (in.)
TRUNK        int     %8.0g    Trunk space (cu. ft.)
WEIGHT       int     %8.0gc   Weight (lbs.)
LENGTH       int     %8.0g    Length (in.)
TURN         int     %8.0g    Turn Circle (ft.)
DISPLACEMENT int     %8.0g    Displacement (cu. in.)
GEAR_RATIO   float   %6.2f    Gear Ratio
FOREIGN      byte    %8.0g    origin    Car type
```

Sorted by: FOREIGN

Note: Dataset has changed since last saved.

```
. rename *, lower
```

```
. des
```

```
Contains data from http://www.stata-press.com/data/r15/auto.dta
  obs:          74          1978 Automobile Data
  vars:         12          13 Apr 2016 17:45
  size:        3,182        (_dta has notes)
```

```
-----
variable name  storage  display  value  variable label
              type   format   label
-----
make          str18   %-18s    Make and Model
price         int     %8.0gc   Price
mpg           int     %8.0g    Mileage (mpg)
rep78        int     %8.0g    Repair Record 1978
headroom     float   %6.1f    Headroom (in.)
trunk        int     %8.0g    Trunk space (cu. ft.)
weight       int     %8.0gc   Weight (lbs.)
length       int     %8.0g    Length (in.)
turn         int     %8.0g    Turn Circle (ft.)
displacement int     %8.0g    Displacement (cu. in.)
gear_ratio   float   %6.2f    Gear Ratio
foreign      byte    %8.0g    origin    Car type
```

Sorted by: foreign

Note: Dataset has changed since last saved.

```
. rename make make2
```

```
. des
```

```
Contains data from http://www.stata-press.com/data/r15/auto.dta
  obs:          74          1978 Automobile Data
  vars:         12          13 Apr 2016 17:45
  size:        3,182        (_dta has notes)
```

```
-----
variable name  storage  display  value  variable label
              type   format   label
```

```
make2      str18   %-18s           Make and Model
price      int      %8.0gc         Price
mpg        int      %8.0g          Mileage (mpg)
rep78      int      %8.0g          Repair Record 1978
headroom   float    %6.1f          Headroom (in.)
trunk      int      %8.0g          Trunk space (cu. ft.)
weight     int      %8.0gc         Weight (lbs.)
length     int      %8.0g          Length (in.)
turn       int      %8.0g          Turn Circle (ft.)
displacement int    %8.0g          Displacement (cu. in.)
gear_ratio float    %6.2f          Gear Ratio
foreign    byte     %8.0g          origin    Car type
```

Sorted by: foreign  
Note: Dataset has changed since last saved.

```
.
.
.
. *****
. * recast variables
. *****
.
. des make
```

variable name	storage type	display format	value label	variable label
make2	str18	%-18s		Make and Model

```
. recast str30 make
. des
```

Contains data from <http://www.stata-press.com/data/r15/auto.dta>  
obs: 74 1978 Automobile Data  
vars: 12 13 Apr 2016 17:45  
size: 4,070 (\_dta has notes)

variable name	storage type	display format	value label	variable label
make2	str30	%-30s		Make and Model
price	int	%8.0gc		Price
mpg	int	%8.0g		Mileage (mpg)
rep78	int	%8.0g		Repair Record 1978
headroom	float	%6.1f		Headroom (in.)
trunk	int	%8.0g		Trunk space (cu. ft.)
weight	int	%8.0gc		Weight (lbs.)
length	int	%8.0g		Length (in.)
turn	int	%8.0g		Turn Circle (ft.)
displacement	int	%8.0g		Displacement (cu. in.)
gear_ratio	float	%6.2f		Gear Ratio
foreign	byte	%8.0g	origin	Car type

Sorted by: foreign  
Note: Dataset has changed since last saved.

```
.
.
.
. exit
end of do-file
```

```
. do "D:\jason\workshop\data management\data management.do"
```

```
. /*****
> This command file was created for Stata commands used in CFDR workshop on "Data Management
> in Stata" on Jan. 25, 2021.
>
>
> The command file includes five groups of Stata commands, including
> (1) general commands;
> (2) commands for Managing data sets,
> (3) commands for checking variables,
> (4) commands for modifying variables,
> (5) Generating new variables.
>
> The final command and log files (i.e., data management.do and data management.log) are saved
> in the folder (c:\temp\workshop\data management).
>
```

```
> *****/
.
. *****/
. * The findit command
. *****/
.
. findit mim //the mim is a user-written ado file for analyzed imputed data
.
. findit esttab //the esttab command is for creating Publication-Quality Tables in Stata
.
. *****/
. * Create a log file
. *****/
.
. *log using "d:\temp\workshop\data management\data management.log", replace
.
. log using "d:\temp\data management.log", replace
log file already open
r(604);

end of do-file

r(604);

. log close
      name: <unnamed>
      log:  d:\temp\data management.log
      log type: text
      closed on: 25 Jan 2021, 10:53:14
-----
```