

String Variables in Stata - Numeric to String, and Concatenating¹

The following document provides an example of how to create string variables from numeric variables, and then concatenate string variables into one. An example of when one might need to do this is if they needed to append several variable sets together, and they needed a unique identifier the included information about the source data set. The Stata commands are preceded with a period. Stata's responses are not. The particularly relevant Stata commands are in bold. Comments are preceded with an asterisk.

```
-----
      log:  C:\dl\chartest.log
      log type:  text
      opened on:  9 Jan 2003, 15:02:54

. cd c:\dl
C:\dl

. insheet using chartest.csv
(3 vars, 16 obs)

. list

      id      country      year
1.      1          AA          94
2.      2          AA          94
3.      3          AA          94
4.      4          AA          94
5.      5          AA          94
6.      6          AA          94
7.      7          AA          94
8.      8          AA          94
9.      1          BB          94
10.     2          BB          94
11.     3          BB          94
12.     4          BB          94
13.     5          BB          94
14.     6          BB          94
15.     7          BB          94
16.     8          BB          94

. * id ranges from 1 to 8.
. * In case ID values are up to four digits,
. * I will add 1000 .
. replace id = id + 1000
id was byte now int
(16 real changes made)

. * Create a string variable by specifying "str(length)"
. * the "string" function can be used to create a string from a number

. gen      str4 id2 = string(id)

. gen      str4 year2 = string(year)
```

¹Prepared by Patty Glynn, University of Washington. 1/9/03 C:\allhelp\helpnew\string_stata.wpd

```
. * String variables can be concatenated by using the + operator
```

```
. gen str8 newid = year2 + country + id2
```

```
. desc
```

```
Contains data
```

```
obs:      16
vars:      6
size:     400 (100.0% of memory free)
```

```
-----
```

variable name	storage type	display format	value label	variable label
id	int	%8.0g		
country	str2	%9s		
year	byte	%8.0g		
id2	str4	%9s		
year2	str4	%9s		
newid	str8	%9s		

```
-----
```

```
. list
```

	id	country	year	id	id2	year2	newid
1.	1001	AA	94	1001	1001	94	94AA1001
2.	1002	AA	94	1002	1002	94	94AA1002
3.	1003	AA	94	1003	1003	94	94AA1003
4.	1004	AA	94	1004	1004	94	94AA1004
5.	1005	AA	94	1005	1005	94	94AA1005
6.	1006	AA	94	1006	1006	94	94AA1006
7.	1007	AA	94	1007	1007	94	94AA1007
8.	1008	AA	94	1008	1008	94	94AA1008
9.	1001	BB	94	1001	1001	94	94BB1001
10.	1002	BB	94	1002	1002	94	94BB1002
11.	1003	BB	94	1003	1003	94	94BB1003
12.	1004	BB	94	1004	1004	94	94BB1004
13.	1005	BB	94	1005	1005	94	94BB1005
14.	1006	BB	94	1006	1006	94	94BB1006
15.	1007	BB	94	1007	1007	94	94BB1007
16.	1008	BB	94	1008	1008	94	94BB1008

```
. log close
```

```
log: C:\dl\chartest.log
log type: text
closed on: 9 Jan 2003, 15:56:49
```

```
-----
```