

xtdescribe — Describe pattern of xt data

[Description
Options](#)[Quick start
Remarks and examples](#)[Menu
Reference](#)[Syntax
Also see](#)

Description

`xtdescribe` describes the participation pattern of cross-sectional time-series (xt) data.

Quick start

Describe the 9 most common participation patterns of `xtset` data

```
xtdescribe
```

Describe up to 15 of the most common participation patterns

```
xtdescribe, patterns(15)
```

As above, but list all participation patterns

```
xtdescribe, patterns(1000)
```

Describe patterns only for study subjects, denoted by binary variable `insample = 1`

```
xtdescribe if insample
```

Menu

[Statistics](#) > [Longitudinal/panel data](#) > [Setup and utilities](#) > [Describe pattern of xt data](#)

Syntax

```
xtdescribe [if] [in] [, options]
```

<i>options</i>	Description
----------------	-------------

Main

<code>patterns(#)</code>	maximum participation patterns; default is <code>patterns(9)</code>
<code>width(#)</code>	display # width of participation patterns; default is <code>width(100)</code>

A panel variable and a time variable must be specified; use `xtset`; see [\[XT\] xtset](#).
`by` is allowed; see [\[D\] by](#).

Options

Main

`patterns(#)` specifies the maximum number of participation patterns to be reported; `patterns(9)` is the default. Specifying `patterns(50)` would list up to 50 patterns. Specifying `patterns(1000)` is taken to mean `patterns(∞)`: all the patterns will be listed.

`width(#)` specifies the desired width of the participation patterns to be displayed; `width(100)` is the default. If the number of times is greater than `width()`, then each column in the participation pattern represents multiple periods as indicated in a footnote at the bottom of the table. The actual width may differ slightly from the requested width depending on the span of the time variable and the number of periods.

Remarks and examples

[stata.com](http://www.stata.com)

If you have not read [\[XT\] xt](#), please do so.

`xtdescribe` describes the cross-sectional and time-series aspects of the data in memory.

► Example 1

In [XT] `xt`, we introduced data based on a subsample of the NLSY data on young women aged 14–26 years in 1968. Here is a description of the data used in many of the [XT] `xt` examples:

```
. use http://www.stata-press.com/data/r14/nlswork
(National Longitudinal Survey. Young Women 14-26 years of age in 1968)

. xtdescribe
idcode: 1, 2, ..., 5159          n =          4711
year:   68, 69, ..., 88          T =           15
Delta(year) = 1 unit
Span(year) = 21 periods
(idcode*year uniquely identifies each observation)

Distribution of T_i:  min      5%      25%      50%      75%      95%      max
                   1         1         3         5         9        13        15

Freq.  Percent  Cum. | Pattern
-----|-----
  136    2.89    2.89 | 1.....
  114    2.42    5.31 | .....1
   89    1.89    7.20 | .....1.11
   87    1.85    9.04 | .....11
   86    1.83   10.87 | 111111.1.11.1.11.1.11
   61    1.29   12.16 | .....11.1.11
   56    1.19   13.35 | 11.....
   54    1.15   14.50 | .....1.1.11
   54    1.15   15.64 | .....1.11.1.11.1.11
 3974   84.36  100.00 | (other patterns)

-----|-----
 4711   100.00 | XXXXXX.X.XX.X.XX.X.XX
```

`xtdescribe` tells us that we have 4,711 women in our data and that the `idcode` that identifies each ranges from 1 to 5,159. We are also told that the maximum number of individual years over which we observe any woman is 15, though the `year` variable spans 21 years. The delta or periodicity of `year` is one unit, meaning that in principle we could observe each woman yearly. We are reassured that `idcode` and `year`, taken together, uniquely identify each observation in our data. We are also shown the distribution of T_i ; 50% of our women are observed 5 years or less. Only 5% of our women are observed for 13 years or more.

Finally, we are shown the participation pattern. A 1 in the pattern means one observation that year; a dot means no observation. The largest fraction of our women (still only 2.89%) was observed in the single year 1968 and not thereafter; the next largest fraction was observed in 1988 but not before; and the next largest fraction was observed in 1985, 1987, and 1988.

At the bottom is the sum of the participation patterns, including the patterns that were not shown. We can see that none of the women were observed in six of the years (there are six dots). (The survey was not administered in those six years.)

We could see more of the patterns by specifying the `patterns()` option, or we could see all the patterns by specifying `patterns(1000)`.



► Example 2

The strange participation patterns shown above have to do with our subsampling of the data, not with the administrators of the survey. Here are the data from which we drew the sample used in [XT] `xt`:

have an observation taken at either 4:00 PM or 5:00 PM on March 9, but we do not have observations for both times. There are three patients for whom we are missing both the 10:00 PM and 11:00 PM observations on March 10, and there are two patients for whom we are missing the 4:00 PM and 5:00 PM observations for March 9.



Reference

Cox, N. J. 2007. *Speaking Stata: Counting groups, especially panels*. *Stata Journal* 7: 571–581.

Also see

[XT] [xtsum](#) — Summarize xt data

[XT] [xttab](#) — Tabulate xt data