

**fillin** — Rectangularize dataset[Description](#)[Remarks and examples](#)[Quick start](#)[References](#)[Menu](#)[Also see](#)[Syntax](#)

## Description

`fillin` adds observations with missing data so that all interactions of *varlist* exist, thus making a complete rectangularization of *varlist*. `fillin` also adds the variable `_fillin` to the dataset. `_fillin` is 1 for observations created by using `fillin` and 0 for previously existing observations.

*varlist* may not contain `strL`s.

## Quick start

Add observations so that all possible interactions of `v1` and `v2` exist and flag new observations with `_fillin = 1`

```
fillin v1 v2
```

As above, but also include interactions with `v3`

```
fillin v1 v2 v3
```

## Menu

Data > Create or change data > Other variable-transformation commands > Rectangularize dataset

## Syntax

```
fillin varlist
```

## Remarks and examples

### ► Example 1

We have data on something by sex, race, and age group. We suspect that some of the combinations of sex, race, and age do not exist, but if so, we want them to exist with whatever remaining variables there are in the dataset set to missing. For example, rather than having a missing observation for black females aged 20–24, we want to create an observation that contains missing values:

```
. use http://www.stata-press.com/data/r14/fillin1
. list
```

|    | sex    | race  | age_gr~p | x1    | x2   |
|----|--------|-------|----------|-------|------|
| 1. | female | white | 20-24    | 20393 | 14.5 |
| 2. | male   | white | 25-29    | 32750 | 12.7 |
| 3. | female | black | 30-34    | 39399 | 14.2 |

```
. fillin sex race age_group
. list, sepby(sex)
```

|     | sex    | race  | age_gr~p | x1    | x2   | _fillin |
|-----|--------|-------|----------|-------|------|---------|
| 1.  | female | white | 20-24    | 20393 | 14.5 | 0       |
| 2.  | female | white | 25-29    | .     | .    | 1       |
| 3.  | female | white | 30-34    | .     | .    | 1       |
| 4.  | female | black | 20-24    | .     | .    | 1       |
| 5.  | female | black | 25-29    | .     | .    | 1       |
| 6.  | female | black | 30-34    | 39399 | 14.2 | 0       |
| 7.  | male   | white | 20-24    | .     | .    | 1       |
| 8.  | male   | white | 25-29    | 32750 | 12.7 | 0       |
| 9.  | male   | white | 30-34    | .     | .    | 1       |
| 10. | male   | black | 20-24    | .     | .    | 1       |
| 11. | male   | black | 25-29    | .     | .    | 1       |
| 12. | male   | black | 30-34    | .     | .    | 1       |

◀

## References

- Baum, C. F. 2016. *An Introduction to Stata Programming*. 2nd ed. College Station, TX: Stata Press.
- Cox, N. J. 2005. *Stata tip 17: Filling in the gaps*. *Stata Journal* 5: 135–136.

## Also see

- [D] **cross** — Form every pairwise combination of two datasets
- [D] **expand** — Duplicate observations
- [D] **joinby** — Form all pairwise combinations within groups
- [D] **save** — Save Stata dataset