

---

Subject: Re: Difference DTA and SAV

Posted by [Bridgette-DHS](#) on Tue, 03 Jan 2023 13:34:54 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS staff members, Tom Pullum and Trevor Croft:

In CPro, the package in which the data files are originally constructed, we have two different codes:

1. Blank this is the not applicable code for use when a question is not applicable.
2. 9 (or 99, 999, etc.) this is the missing value code for use when a question is applicable, but a response was not given.

9 (etc.) is used very rarely nowadays with CAPI, but was needed for paper questionnaires, and was used by data entry staff when a particular question was not NA but the interviewer forgot to enter a response.

Usually in SPSS, blank is the system-missing value, and 9 is the user-missing value. In Stata, usually both cases are converted to missing.

We are surprised that the 9 shows up in the Stata file. Perhaps the DTA file for this survey was not created in our usual way, which would convert the CPro 'missing' to a dot in Stata.

You are reading the Stata and SPSS files with R, and we believe the conversion to R is being handled in different ways. You will get agreement if, when you read the Stata file with R, you add a line to change the 9 to a dot.