

Missing in more ways than one

An introduction to the special missing values SAS has for numeric data

Missing values

- “.” (a period) for numeric
- “ ” (a blank) for character

Handling missing values

- Missing data can be explicitly handled in
 - IF-THEN-ELSE constructs;
 - recoded in PROC FORMATS and
 - in calculations

Special missing values

- `._`
- `.A .B .C .D .E .F .G .H .I .J .K .L .M`
- `.N .O .P .Q .R .S .T .U .V .W .X .Y .Z`

Some features

- .A is equivalent to .a
- displayed and printed as uppercase
- SAS prints only the letter or underscore

Setting missing values 1

- DATA step code
 - if age < 0 then age = .; or
 - if age < 0 then age = .a;

Setting missing values 2

- During input
 - Use the MISSING statement
 - Example from p102
- SAS® 9.1.3 Language Reference: Concepts**
Third Edition

Missing statement

```
data period_a;  
    missing X I;  
    input Id $4. Foodpr1 Foodpr2 Foodpr3 Coffeem1 Coffeem2;  
datalines;  
1001 115 45 65 I 78  
1002 86 27 55 72 86  
1004 93 52 X 76 88  
1015 73 35 43 112 108  
1027 101 127 39 76 79  
;
```

Output with Multiple Missing Values

Results of Test Period A

Obs	Id	Foodpr1	Foodpr2	Foodpr3	Coffeem1	Coffeem2
1	1001	115	45	65	I	78
2	1002	86	27	55	72	86
3	1004	93	52	X	76	88
4	1015	73	35	43	112	108
5	1027	101	127	39	76	79

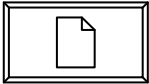
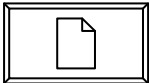
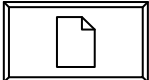
X indicates TESTER ABSENT

I indicates TEST WAS INCOMPLE

Order of missing values

- `._` Smallest
- `.`
- `.A ...`
- `.Z`
- Negative numbers
- Zero
- Positive numbers Largest

Example

- Report on work sites and wages for a company. 
- With data omitted or reported incorrectly 
- Enhanced report 

Data

Site	StartYr	EndYr	RptYr	WageAmt
Factory 2	2002	2004	2002	0.6
Factory 2	2002	2004	2004	1.2

Data expanded

Site	StartYr	EndYr	RptYr	WageAmt
Factory 2	2002	2004	2000	
Factory 2	2002	2004	2001	
Factory 2	2002	2004	2002	0.6
Factory 2	2002	2004	2003	
Factory 2	2002	2004	2004	1.2
Factory 2	2002	2004	2005	

Assign special missing values

```
if missing(WageAmt)
  then do;
    if RptYr < StartYr
      then WageAmt = .a;
    if RptYr > RptTo
      then WageAmt = .a;
    if RptYr >= StartYr and RptYr <= RptTo
      then WageAmt = .r;
  end;
```

Define a format

```
proc format;  
  value WagesFmt  
    .a      = ' '          /* Blank because  
                                not required */  
    .r      = 'Required' /* Not supplied  
                                contact company */  
  other = [10.1]  
;
```

