

DAS
ONLINE



USER GUIDE

AUTHOR

Helen Jang
MPR Associates, Inc.

GRAPHIC DESIGNER

Leslie Retallick
MPR Associates, Inc.

June 2005

TABLE OF CONTENTS

Introduction	1
About the Guide	1
Overview	1
How Does the DAS Work?	1
What You Will Need	2
Starting the DAS	3
Major Components	3
The Main Window	4
The Variables Information Window	4
The Tag Options Window	5
The Tag-Specific Window	7
The TPF/CPF Window	7
The Confirm Input Window	8
The Output Window	8
DAS Functions and Utilities	9
Creating Tables	9
The Table Parameter File (TPF)	9
The Table	10
Creating Correlation Matrices	11
The Correlation Parameter File (CPF)	12
The Correlation Matrix	13
The Batch Processor	14
The Weight Converter	15
Using the DAS	15
Types of Variables	15
Categorical Variables	15
Continuous Variables	16
Viewing and Searching for Variables	17
Viewing Variables	17
Searching for Variables	17
Viewing Variable Descriptions/Percentages	17
Working with Parameters	18
Creating Parameters	18
Editing Parameters	18
Deleting Parameters	19

Working with Parameter Files—TPF/CPF	19
Creating Files	19
Importing Files	19
Editing Files	19
Saving Files	20
Running Files and Getting Output	20
Running Files	20
Printing Tables	20
Saving Output Files	20
Tags: Descriptions and Instructions	21
For Tables	21
Types of Variables and Tags	21
Tag Names and What They Mean	21
Column Tags	22
Average	23
Mean	24
Percent >	25
Sum	26
Centile	26
Centile>0	27
Column_Cat	28
Column_Cut	30
Column_Lump	32
Row Tags	34
Row_Cat	34
Row_Cut	35
Row_Lump	37
Span Tags	38
Span_Cat	38
Span_Cut	39
Span_Lump	41
By Tags	42
By_Cat	43
By_Cut	44
By_Lump	45
Filters	47
And_Filter	47
Or_Filter	48

Weight	50
Different Populations	50
Population Subsets	51
For Correlations	51
Continuous	51
Cut	52
Each	53
Lump	54
Frequently Asked Questions	56
Choosing Variables	56
Viewing Variable Information	56
Updating and Creating New Variables	57
Using Tags	57
For Correlation Matrices	58
DAS Resources	59
Tables Library	59
User Help Center	59
Getting Technical Assistance	59

DAS ONLINE USER GUIDE

INTRODUCTION ► The National Center for Education Statistics (NCES) is the primary federal agency reporting data related to education in the United States and other nations. One of its congressional mandates is to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States. The Data Analysis System (DAS) is a prime example of how NCES makes quality education data available to the public. Anyone can use this web application to generate tables using data from NCES surveys.

There is a separate DAS for each survey data set, but all applications have a consistent interface and command structure. The DAS can produce complicated tables and output files for use in multivariate analyses. There are also two utilities for the DAS, the Batch Processor and the Weight Converter. (See the Overview section.)

About the Guide Throughout the guide, variable labels are displayed in italics, and the variable names appear in parentheses, for example, *Age in 12/31/95* (AGE). The variables used in the examples may not be in the particular DAS you are using. The DAS Online application is called “the DAS” throughout the guide.

OVERVIEW ►

How Does the DAS Work? The DAS can create two types of output—Tables and Correlations. For tables, the DAS will generate table estimates, standard errors, and weighted sample sizes for the estimates. The standard errors are calculated by taking into account the complex sampling designs used in NCES surveys. For correlations, the DAS will create a correlation matrix that can be used as input for popular statistical software to conduct multivariate analyses. The design effects (DEFTs) for each variable are included in the output file. Because statistical procedures generally compute regression coefficients based on assumptions for a simple random sample, the standard errors must be adjusted with the design effect to take into account the stratified sampling method used in NCES surveys.

Creating a basic table involves four steps:

1. Specifying a column variable,
2. Defining a row variable,

3. Selecting a weight variable if the data set contains more than one weight option,¹ and
4. Running the Table Parameter File (TPF).

In the first three steps, you'll select column, row, and weight variables and define how they will appear in your table. This process is called "tagging." Many tags are available in the DAS and each is used for a different purpose. For a detailed discussion of tags, see the Tags section.

Creating a simple correlation matrix involves four steps:

1. Tagging the desired variables (most of the tag options are different than those used for generating tables).
2. Selecting a weight variable (again, only if there is more than one weight option for the data set).
3. Selecting an output format (SPSS or SAS)
4. Running the Correlation Parameter File (CPF)

Once a variable is tagged, it is called a "parameter." Parameters are stored in a parameter file— Table Parameter File (TPF) or a Correlation Parameter File (CPF). The DAS will run the TPF/CPF and create an output file (table/correlation matrix).

The DAS has two utilities—the Batch Processor and the Weight Converter. The Batch Processor will run multiple TPFs and CPFs from different data sets in one job. The Weight Converter replaces DAS Windows weight variables with corresponding DAS Online weight variables. You can submit individual TPFs/CPFs or compressed (.zip) files containing multiple files. Once the weight variables are changed, you can submit the new .zip file to the Batch Processor. For a detailed discussion of these utilities, see the Batch Processor and Weight Converter sections.

What You Will Need

To use the DAS, you will need a web browser. You can use Microsoft Internet Explorer (5.0 or higher) or Netscape (4.7 or higher).

¹Most data sets contain more than one weight variable (found under Surveys→Weights in the Variables List). If a data set contains only one weight, the DAS will automatically add it to your TPF.

Starting the DAS

To launch the application,

1. Go to the DAS web site (<http://nces.ed.gov/das>).
2. Click *DAS Online Application*.
3. From the left-hand-side menu, click *To Create Tables* or *To Create Correlations*. A list of surveys appears, organized by respondent group for each survey; for example, College Graduates.
4. Click the respondent group you wish to analyze. The NCES Data Usage Agreement appears.
5. Review the agreement and click *I agree to the terms above*. The DAS for that survey appears.

Major Components

This section shows the major components of the DAS.

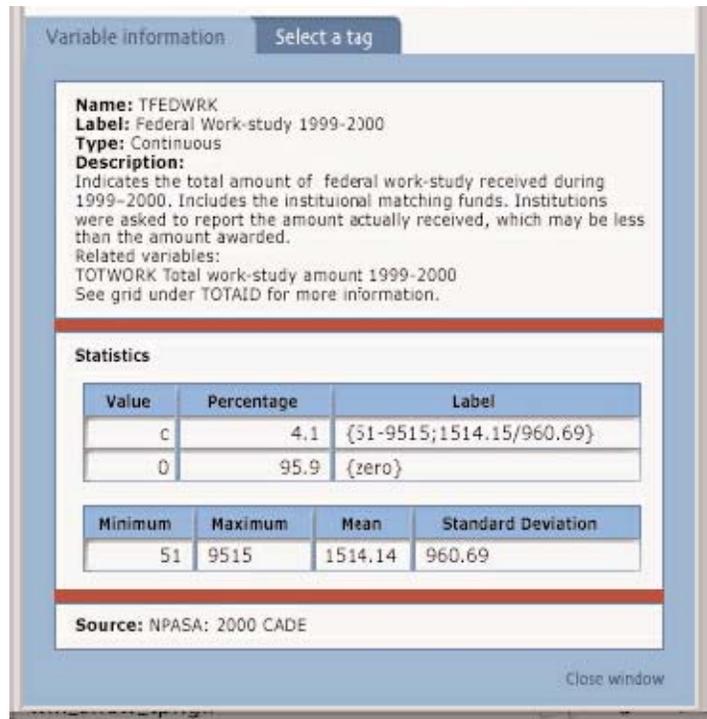
The Main Window

When you launch the DAS, the Main Window appears.



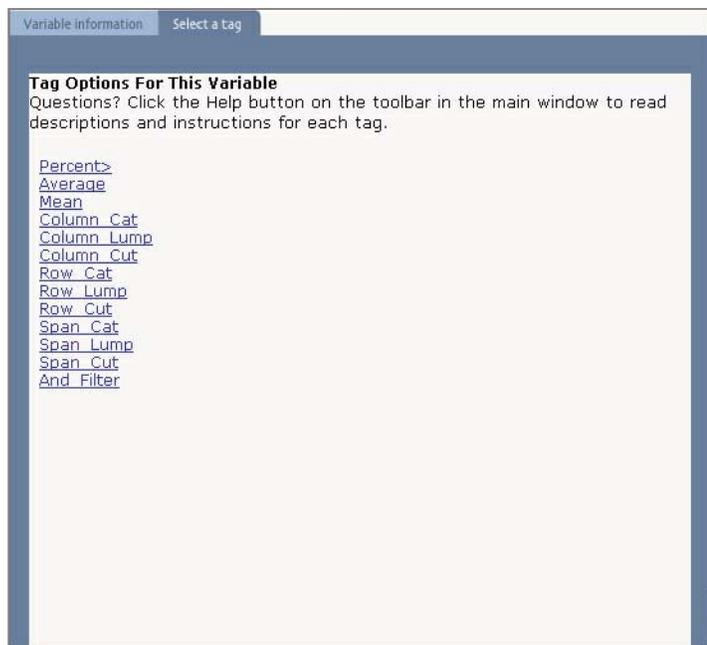
The Variable Information Window

When you select a variable in the Variable List, the Variable Information Window appears.



The Tag Options Window

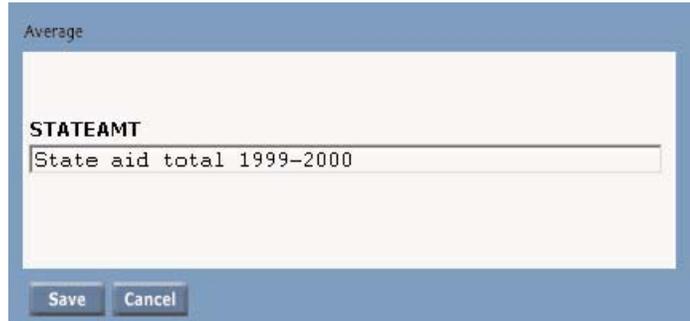
When you click the Select the Tag tab in the Variables Information Window, the Tag Options Window appears.



The Tag-Specific Window

When you click a tag from the list of the tag options, the Tag-Specific Window appears. The nine main types of tag windows are shown below.

The windows for Average, Mean, Centile, Centile>0, Continuous, and Weight tags look like this:



The window for the Percent> tag is similar, with the additional required specification of a cut-point value.

The windows for Cat tags (Column_Cat, Row_Cat, Span_Cat, and By_Cat) look like this:



The window for the Each tag (used only in correlations) is similar, with the additional required selection of a base category.

The windows for Cut tags (Column_Cut, Row_Cut, Span_Cut, and By_Cut for tables, and Cut tag for correlations) look like this:

Row_Cut

AGE
Age as of 12/31/03

Enter your cut points and customize labels below:

	Cut point	Label
1	-0.5	
2		
3		
4		
5		
6		
7		
8		
9		
10		

Save Cancel

The windows for Lump tags (Column_Lump, Row_Lump, Span_Lump, and By_Lump for tables, and Lump tag for correlations) look like this:

Column_Lump

RACE
Race-ethnicity (with multiple)

Select desired values to begin lumping process. At least two lumps are required.

- White
- Black or African American
- Hispanic or Latino
- Asian
- American Indian or Alaska Native
- Native Hawaiian / other Pacific Islander
- Other
- More than one race

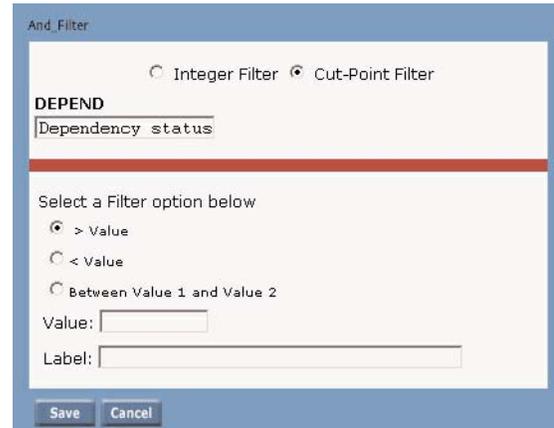
Lump title:

Add Lump Save Cancel

The windows for Filters (And_Filter and Or_Filter) look like this:
 For Integer Filters:

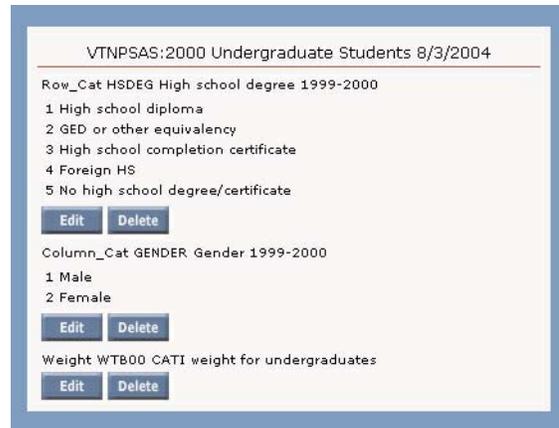


For Cut-Point Filters:



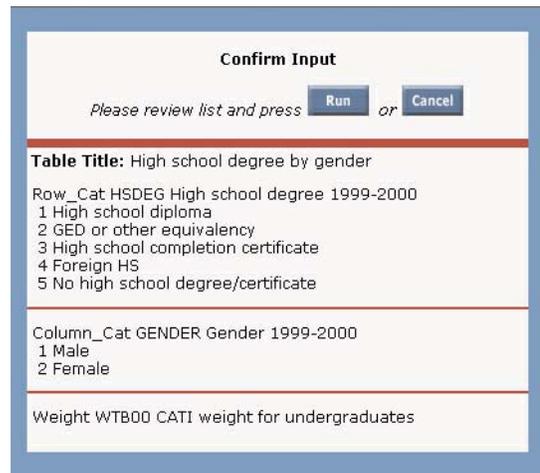
The TPF/CPF Window

You can view parameters at any time after tagging at least one variable by clicking *Show* on the Main Toolbar.



The Confirm Input Window

Before you can run a table, the DAS will ask you to confirm the TPF/CPF. This window will appear.



The Output Window

After you click *Run*, the DAS will process your table. When the table is ready, it will appear on your screen.

Computation by DAS-T Online Version 4.0 on 03/23/2005

High school degree by gender

Estimates	Gender 1999–2000	
	Male (%)	Female (%)
Total	43.6	56.4
High school degree 1999–2000		
High school diploma	43.9	56.1
GED or other equivalency	37.8	62.2
High school completion certificate	44.7	55.3
Foreign HS	46.3	53.7
No high school degree/certificate	42.4	57.6
Standard Errors (BRR)		
Total	0.09	0.09
High school degree 1999–2000		
High school diploma	0.12	0.12
GED or other equivalency	1.19	1.19
High school completion certificate	6.04	6.04
Foreign HS	7.57	7.57
No high school degree/certificate	4.96	4.96
Weighted sample sizes (n/1,000s)		
Total	16,538.4	
High school degree 1999–2000		
High school diploma	15,385.4	
GED or other equivalency	829.4	
High school completion certificate	56.1	
Foreign HS	150.4	
No high school degree/certificate	117.1	

Source: NCES, NPSAS:2000 Undergraduate Students 08/03/2004



VIEW
Printer-Friendly
Version



DOWNLOAD
Comma-Separated
Format



DOWNLOAD
Printer-Friendly
Web Page



RETURN
to
Specifications

Click the Back button on your browser when finished printing

DAS FEATURES AND UTILITIES

Creating Tables

One primary feature of the DAS is that it allows you to generate tables from NCES data sets based on your own specifications. The two main components of this feature—the TPF and the table—are discussed below.

The Table Parameter File (TPF)

As previously mentioned, once a variable is tagged, it is called a “parameter.” Parameters are stored in a parameter file—a Table Parameter File (TPF) or a Correlation Parameter File (CPF). The major components of a TPF are shown below. Information about each tag is provided in the Tags section.

1 Data Set

Indicates the selected data set. For this example, we use NPSAS 2000 Undergraduate Students.

2 Weight

Indicates the weight selected. For data sets that have multiple weights, you must tag a weight. For data sets that have only one weight, the DAS will automatically add it to your TPF.

3 Filters

Indicates the groups you have decided to include in the table. Based on these filters, the DAS will include estimates only for students who are U.S. citizens ages 17–49.

4 Column Tag

Indicates the column variable selected. In this case, the DAS will produce a column showing the percentage of students who received a Pell grant. It will be the last column in the table because this TPF also contains a By tag.

5 By Tag

Creates categories within the column variable. The DAS will create a column of estimates for students who received Pell grants within each category you specified. For example, the first column will be the percentage of students who received Pell grants within the traditional risk category.

6 Span Tag

Indicates the variable for which the DAS will create subtables. There will be a total of three subtables in this table, one for each lump you specified.

7 Row Tag

Indicates the variable for which the DAS will generate rows. Each category will be shown as a row.

8 Title

This is the title you entered for the table.

```

1  VTNPSAS:2000 Undergraduate Students 8/3/04
2  Weight WTB00 CAT1 weight for undergraduates
   Filter CITIZEN2 Student^s citizenship 1999-2000
   1 US citizen
3  Filter AGE Age as of 12/31/99
   16.5<X<49.5 Between 17 and 50
4  Percentage>0.5 PELLAMT Pell
   By_Cut RISKINDX Nontraditional status
   -0.5 Traditional
   0.5 Minimally
   1.5 Moderately
   3.5 Highly
5  Span_Lump AIDSECT Institution type
   1 3+4 Public 4-year
6  2 6+7 Private nfp 4-year
   3 2 Public 2-year
7  Row_Cat INCOME Total income by dependency
   1 Dependent: Less than $10,000
   2 Dependent: $10,000-$19,999
   3 Dependent: $20,000-$29,999
   4 Dependent: $30,000-$39,999
   5 Dependent: $40,000-$49,999
   6 Dependent: $50,000-$59,999
   7 Dependent: $60,000-$69,999
   8 Dependent: $70,000-$79,999
   9 Dependent: $80,000-$99,999
   10 Dependent: $100,000 or more
   11 Independent: Less than $5,000
   12 Independent: $5,000-$9,999
   13 Independent: $10,000-$19,999
   14 Independent: $20,000-$29,999
   15 Independent: $30,000-$49,999
   16 Independent: $50,000 or more
8  Title Percentage of US citizens between the ages of 17 and
   50, inclusive, who received Pell grants, by nontraditional
   status, institution type, and income: AY 1999-2000

```

The Table Using the TPF from the previous page, the DAS generated this table. The output file also contains the standard errors and weighted sample size. The numbers below correspond to those in the TPF explanation.

- 3** Filters
- 4** Column Tag
- 5** By Tab
- 6** Span Tag
- 7** Row Tag
- 8** Title

8 **Percentage of US citizens between the ages of 17 and 50, inclusive, who received Pell grants, by nontraditional status, institution type, and income: AY 1999-2000**

And Filters **3**
 Student's citizenship 1999-2000=US citizen
 Age as of 12/31/99= Between 16.5 AND 49.5

5
 Nontraditional status

	Traditional (%>0.5)	Minimally (%>0.5)	Moderately (%>0.5)	Highly (%>0.5)	Pell Total
6 Type of Institution = Public 4-year					
Estimates					
Total	19.58	23.59	26.96	28.91	23.33
Total income by dependency					
Dependent:					
Less than \$10,000	68.05	68.57	41.62	low n	65.34
\$10,000-\$19,999	76.42	71.96	42.92	low n	72.53
\$20,000-\$29,999	71.75	52.18	31.53	low n	63.04
\$30,000-\$39,999	39.64	32.67	12.6	low n	35.83
\$40,000-\$49,999	16.58	15.02	9.59	low n	15.57
\$50,000-\$59,999	3.83	0.61	1.82	low n	2.72
\$60,000-\$69,999	1.62	0.51	6.1	low n	1.62
\$70,000-\$79,999	0	0	0	low n	0
\$80,000-\$99,999	0	0	0	low n	0
\$100,000 or more	0	0	0	low n	0
Independent:					
Less than \$5,000	low n	70.1	69.69	66.23	68.92
\$5,000-\$9,999	low n	71.78	62.75	74.37	67.42
\$10,000-\$19,999	low n	34.4	36.97	51.79	41.35
\$20,000-\$29,999	low n	2.26	23.46	38.33	27.9
\$30,000-\$49,999	low n	0.32	14.54	19.07	16.02
\$50,000 or more	low n	0	0	0	0
6 Type of Institution = Public 2-year					
Estimates					
Total	22.61	16.69	13.29	20.6	17.75
Total income by dependency					
Dependent:					
Less than \$10,000	68.78	71.19	35.17	low n	55.51
\$10,000-\$19,999	73.01	48.79	26.64	low n	44.69
\$20,000-\$29,999	65.55	41.86	12.66	low n	37.81
\$30,000-\$39,999	34.07	25.37	15.73	19.53	24.17
\$40,000-\$49,999	13.53	4.93	5.98	low n	7.48
\$50,000-\$59,999	5.68	0	0	low n	1.9
\$60,000-\$69,999	0	2.69	0	low n	0.95
\$70,000-\$79,999	0	0	0	low n	0
\$80,000-\$99,999	0	0	0	low n	0
\$100,000 or more	0	0	0	low n	0
Independent:					
Less than \$5,000	low n	low n	40.1	51.11	47.46
\$5,000-\$9,999	low n	low n	50.09	58.92	55.32
\$10,000-\$19,999	low n	low n	20.67	46.6	35.71
\$20,000-\$29,999	low n	0	15.25	26.37	22.54
\$30,000-\$49,999	low n	0	9.19	9.43	9.2
\$50,000 or more	low n	0	0	0	0

Source: NCES, NPSAS:2000 Undergraduate Students 04/11/02.

Creating Correlation Matrices

The other main feature of the DAS is that it allows you to generate correlation matrices that can be used as input for linear regression models in SPSS or SAS. The design effects (DEFTs) for each variable are included in the output. Because statistical procedures generally compute regression coefficients based on a simple random sample assumption, the standard errors must be adjusted with the design effect to take into account the stratified sampling method used in NCES surveys.

The CPF allows you to create a two-category comparison in which the group specified is compared with all other cases. In most cases, the biggest group is not defined in the CPF and is used as the comparison group. You must tag each defined group so that it can be compared with the comparison group. In other words, the untagged group becomes the comparison group by default. You should define values that are not needed for comparison as a separate group so they will not be included in the comparison group. For examples, see the next page.

All correlations in the matrix are calculated using pairwise deletion of missing cases. Therefore, the definition of missing values is more flexible for correlations than for tables. For all data sets, -1 indicates useless data due to missing values, refusals, legitimate skips, and other factors. However, some variables have reserved codes of zero. In Cut specifications, the set of missing values may include -1 and 0.

It is helpful to have a basic knowledge of SPSS or SAS and a general understanding of multiple regressions before you use this DAS feature. For information about SPSS or SAS, visit <http://www.spss.com> or <http://www.sas.com>.

The two main components of this feature—the CPF and the output—are discussed below in the CPF section.



This feature is not available for the Integrated Postsecondary Education Data System (IPEDS) DAS.

The Correlation Parameter File (CPF)

The major components of a CPF are shown below. There is a different set of tags for creating correlation matrices. For explanations of the tags, see the Tags section.

1 Data Set

Indicates the selected data set. For this example, it is Baccalaureate and Beyond 2001.

2 Filters

Indicates the groups you have decided to include in the correlation. Based on these filters, the DAS will include estimates only for students receiving their first bachelor's degree at public 4-year or private not-for-profit 4-year institutions.

3 Cut Tag

Creates a two-category comparison in which the value range specified is compared with all other values. There are four variables in this CPF that are tagged as Cut. For example, *Time between postsecondary education and completion of a bachelor's degree* (PSE_BA). Students who took 4 years or less (1-48 months) to complete a bachelor's degree will be compared with the untagged group, students who required a longer period of time to do so.

4 Lump Tag

Creates a two-category comparison in which the category specified is compared with the untagged group. There are five variables in this CPF that are tagged as Lump. Each category for a variable is tagged separately in order to produce a two-category comparison. For example, the untagged group—White, non-Hispanic—for the variable Race/ethnicity (RACE1) will be compared with each of other race/ethnicity groups separately.

5 Each Tag

Compares a single category with each of the other variable groups. All categories are included in the correlation. In this case, the untagged group, Males, will be compared with the tagged group, Females.

6 Output File Selection

Indicates whether the output file will be in SPSS or SAS. We selected SPSS.

7 Title

This is the title you entered for the correlation matrix.

- 1 VCB&B:2001 07/08/03
- 2 F CBFSTBA First bachelor's degree yes/no
 - 1 First bachelor's degree
- 2 F I1SECT First PSE institution sector
 - 1 Public 4-year
 - 4 Private not-for-profit 4-year
- 3 X PSE_BA Time between PSE and BA completion
 - 1 0.5 48.5 4 years or less
- 4 L I1SECT First PSE institution sector
 - 1 4 Private not-for-profit 4-year
- 5 E GENDER Gender 1999-2000
 - 2 Female
- 4 L RACE1 Race/ethnicity (historical) 1999-2000
 - 1 2 Black, non-Hispanic
 - L RACE1 Race-ethnicity (historical) 1999-2000
 - 1 3 Hispanic or Latino
 - L RACE1 Race-ethnicity (historical) 1999-2000
 - 1 4+6 Asian/Pacific Islander
 - L RACE1 Race-ethnicity (historical) 1999-2000
 - 1 5 American Indian/Alaska Native
 - L RACE1 Race-ethnicity (historical) 1999-2000
 - 1 7 Other
- 3 X PCTALL2 Income percentile rank for all students 99-00
 - 1 0.5 25.5 Low quartile
 - 1 25.6 74.5 Middle quartiles
- 3 X RISKINDX Index of risk 1999-2000
 - 1 0.5 1.5 One risk factor
 - 1 1.6 9.5 Two or more
- 3 X GPA2 Cumulative undergraduate GPA
 - 1 -0.5 224.8 Less than 2.25
 - 1 224.9 325.1 2.25 to 3.25
- 4 L CBMULTPL Undergrad multiple enrollment
 - 1 1 Sometimes co-enrolled
- 4 L COMMCOLL Ever attended community/other 2 year
 - 1 1 Attended a 2-year college
- 4 L TRNPURP Transfer status
 - 1 1 Transferred
- 6 O SPSS_WIN
- 7 Title Time to degree regression

The Correlation Matrix

Using the CPF on the previous page, the DAS produced the following correlation matrix (the numbers below correspond to those in the CPF page):

1 Source: NCES, B&B:2001 08/15/03

7 Time to degree regression

2 Filters:
 First bachelor's degree yes/no=First bachelor's degree. (Value=1)
 First PSE institution sector=Public 4-year. (Value=1)
 =Private not-for-profit 4-year. (Value=4)
 Missing Values Excluded

6 -----SPSS for Windows -----.

```

MATRIX DATA VARIABLES=ROWTYPE_
PSE_BA1 I1SECT1 GENDER1 RACE11 RACE12 RACE13 RACE14
RACE15 PCTALL1 PCTALL2 RISKIN1 RISKIN2 GPA21 GPA22
CBMULT1 COMMCO1 TRNPUR1.
BEGIN DATA.
MEAN 0.4709 0.3505 0.5712 0.0831 0.0760 0.0634 0.0054 0.0141 0.2340 0.4511
      0.2240 0.3181 0.0377 0.5068 0.0776 0.3728 0.2555
STDDEV 0.5007 0.4798 0.4975 0.2789 0.2664 0.2441 0.0734 0.1179 0.4250 0.4985
        0.4184 0.4671 0.1917 0.5012 0.2679 0.4846 0.4376
CORR 1.0000
CORR 0.2587 1.0000
CORR 0.0778 -0.0136 1.0000
CORR -0.0778 -0.0388 0.0402 1.0000
CORR -0.0463 -0.0195 0.0349 -0.0874 1.0000
    
```

VAR LABELS

```

PSE_BA1 'Time between PSE and BA completion=4 years or less'/
I1SECT1 'First PSE institution sector=Private not-for-profit 4-year'/
GENDER1 'Gender 1999-2000=Female'/
RACE11 'Race-ethnicity (historical) 1999-2000=Black, non-Hispanic'/
RACE12 'Race-ethnicity (historical) 1999-2000=Hispanic or Latino'/
RACE13 'Race-ethnicity (historical) 1999-2000=Asian/PI'/
RACE14 'Race-ethnicity (historical) 1999-2000=American Indian/Alaska Native'/
RACE15 'Race-ethnicity (historical) 1999-2000=Other'/
    
```

	MEAN"	"S.E."	"DEFT"	"VAR"	"LABEL"
3	0.4709	0.0091	1.8584	"PSE_BA1"	"Time between PSE and BA completion=4 years or less"
2	0.3505	0.0079	2.0303	"I1SECT1"	"First PSE institution sector=Private not-for-profit 4-year"
5	0.5712	0.0078	2.1403	"GENDER1"	"Gender 1999-2000=Female"
4	0.0831	0.0056	1.6713	"RACE11"	"Race-ethnicity (historical) 1999-2000=Black, non-Hispanic"
	0.0760	0.0059	1.5098	"RACE12"	"Race-ethnicity (historical) 1999-2000=Hispanic or Latino"
	0.0634	0.0037	2.2296	"RACE13"	"Race-ethnicity (historical) 1999-2000=Asian/PI"
	0.0054	0.0000	0.0000	"RACE14"	"Race-ethnicity (historical) 1999-2000=American Indian/Alaska Native"
	0.0141	0.0000	0.0000	"RACE15"	"Race-ethnicity (historical) 1999-2000=Other"
3	0.2340	0.0066	2.1710	"PCTALL1"	"Income percentile rank for all students 99-00=Low quartile"
	0.4511	0.0076	2.1930	"PCTALL2"	"Income percentile rank for all students 99-00=Middle quartiles"
3	0.2240	0.0069	2.0318	"RISKIN1"	"Index of risk 1999-2000=One risk factor"
	0.3181	0.0087	1.7967	"RISKIN2"	"Index of risk 1999-2000=Two or more"
3	0.0377	0.0037	1.7339	"GPA21"	"Cumulative undergraduate GPA=Less than 2.25"
	0.5068	0.0082	2.0691	"GPA22"	"Cumulative undergraduate GPA=2.25 to 3.25"
4	0.0776	0.0046	1.9368	"CBMULT1"	"Undergrad multiple enrollment=Sometimes co-enrolled"
4	0.3728	0.0080	2.0222	"COMMCO1"	"Ever attended community/other 2 year=Attended a 2-year college"
4	0.2555	0.0073	2.0112	"TRNPUR1"	"Transfer status=Transferred"

The Batch Processor

Using the Batch Processor, you can submit many files, for both tables and correlation matrices from multiple data sets in one compressed file (.zip), to the DAS for processing. This utility requires a username and password.

To use the Batch Processor, go to the DAS website (<http://nces.ed.gov/das/>)

1. Click *DAS Online Application*.
2. From the left-hand-side menu, click *Batch Processor*.
3. Log in.
4. Upload your file.
The Batch Processor will assign a job code to your file. Be sure to note this number because you will need it to retrieve your processed file.
5. Retrieve your file.

Your output file will be in comma-separated format (CSV), which can be opened using Microsoft Excel. By default, the Batch Processor will also include printer-friendly web versions (in HTML format) of your tables. If you do not want these files, uncheck the box next to "Include Printer-Friendly HTML version" when you drop off your files for processing.

How to Compress Files

There are many ways to compress files; one way is to use an application called WinZip. To download an evaluation copy of this application, visit the WinZip website (<http://www.winzip.com/downwz.htm>) and then follow these steps:

1. Launch WinZip.
2. Click *New*.
The New Archive window appears.
3. Choose the directory containing the DAS files.
4. In the "File name" field, type a name.
5. Click *OK*.
The Add Window appears.
6. Choose your DAS files and click *Add* (hold down the CTRL key to select multiple files).
7. Click *Exit*.

The Weight Converter

The DAS was previously available only in a Windows format. The weight variables for the Windows and Online versions are different. Many analysis report tables were generated using the DAS for Windows, and the TPFs/CPFs for the tables are available for download on the DAS website. If you import a TPF/CPF containing a Windows weight variable into the DAS Online, the application will automatically replace the variable with the corresponding Online variable. However, to run these files using the Batch Processor, you must first convert the weights.

The Weight Converter will replace Windows weight variables with the correct Online variables. You can submit individual TPFs/CPFs or many files compressed into a zip file. Files must be in one of the following formats: .TPF, .CPF, or .ZIP. Once the weight variables are changed, you can submit the new zip file to the Batch Processor.

To use the Weight Converter, go to the DAS website (<http://nces.ed.gov/das/>)

1. Click *DAS Online Application*.
2. From the left-hand-side menu, click *Weight Converter*.
3. Upload your file.
4. Download your file.

USING THE DAS ►

Types of Variables

There are two types of variables in the DAS—Categorical and Continuous. In the Variables List, categorical variables are shown in green and continuous variables are displayed in blue. For both types, values of -1 usually indicate useless data due to missing values, refusals, legitimate skips, and other factors.

Categorical Variables

Categorical variables contain discrete groups and generally have nonadditive values—that is, the groups should usually not be lumped together. For example:

VARIABLE	CATEGORIES
GENDER	Male, Female

A few categorical variables may have special or restricted code values of zero. For example, for the variable *Type of first postsecondary institution* (PSE1ST), zero values indicate that the survey respondent did not enroll in a postsecondary institution.

The Statistics section of the Variable Information Window for a categorical variable looks like this:

Race/ethnicity (RACE)

Statistics

Value	Percentage	Label
1	70	White, non-Hispanic
2	12.2	Black, non-Hispanic
3	10.3	Hispanic
4	5.8	Asian/Pacific Islander
5	1	American Indian/Alaskan Native
6	0.7	Other

Continuous Variables

Continuous variables do not contain discrete groups, and the values can be added together. For example,

TO CREATE GROUP	ADD VALUES
Less than 18 years	1–18.5
19–24 years	18.5–23.5
25–29 years	24.5–29.5
30 years or more	29.5 and above

The Statistics section of the Variable Information Window for a continuous variable looks like this:

Age as of 12/31/99 (AGE)

Statistics

Value	Percentage	Label
C	100	{15-86;25.60/8.54}

Minimum	Maximum	Mean	Standard Deviation
15	86	25.6	9.57

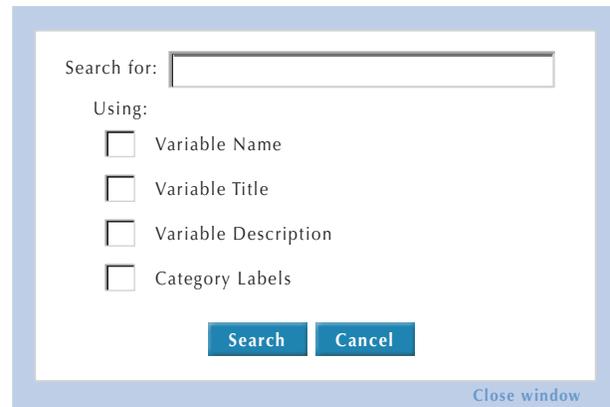
Viewing and Searching for Variables

Viewing Variables

In the Variable List, click  to expand the subject listing until the variables appear.

Searching for Variables

1. From the Toolbar, click *Search for Variable*. The Search Window appears.



2. In the “Search for” field, type the desired search term.
3. Select search method.
4. Click *Search*.

Viewing Variable Descriptions/ Percentages

1. From the Variable List, click the variable you wish to view. The Variable Information Window appears.
2. Scroll to the “Statistics” section to view the percentages.

Variable information Select a tag

Name: AGE
Label: Age as of 12/31/99
Type: Continuous
Description:
 Indicates student's age on 12/31/1999. Students who are 24 on or before this date are considered independent for financial aid purposes in the 1999-2000 academic year. Calculated from date of birth (DOB).
 0.5% of cases (unweighted) were statistically imputed.
 Related variable:
 ZAGE Source for student's age.

Statistics

Value	Percentage	Label
c	100	{ 15-86;25.60/8.54 }

Minimum	Maximum	Mean	Standard Deviation
15	86	25.6	8.54

Source: NPSAS:2000 Student CATI, NPSAS:2000 CADE, CPS:2000, Imputation.

[Close window](#)

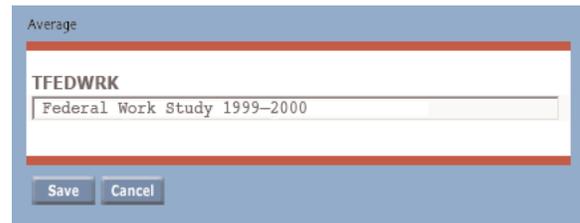
Working with Parameters

To generate a table of estimates, you must select the variables and specify how they should appear. This process is called “tagging,” and a tagged variable is called a “parameter.”

Creating parameters

In the Main Window,

1. Click the variable you wish to tag.
The Variable Information Window appears.
2. Click the *Select a Tag* Tab.
The Tag-Specific Window appears. For example:

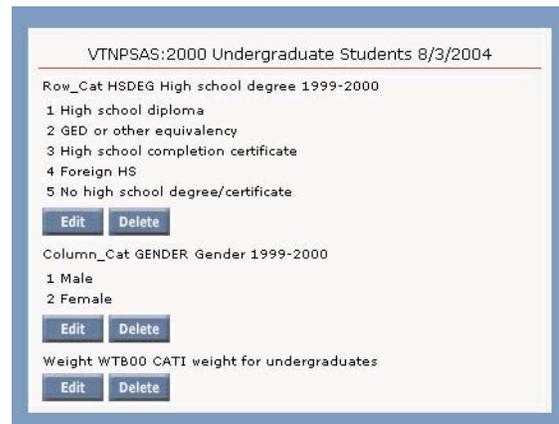


3. Edit specifications as desired. See the Tags section for more information.
4. Click *Save*.
The window closes.
5. A notation appears next to the variable you just tagged.
For example, **Used As: Average**

Editing parameters

From the Toolbar,

1. Click *Show*.
The Table Parameter File Window appears.



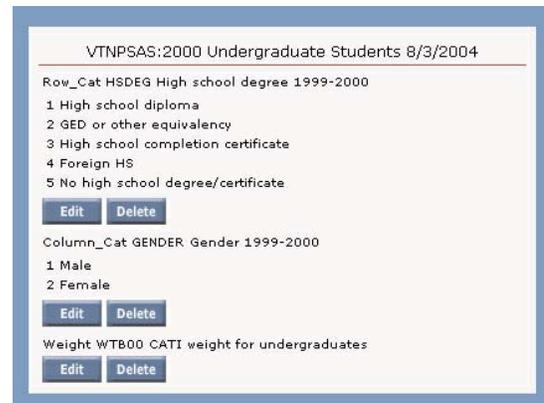
2. Click *Edit* for the parameter you wish to change.
The Tag-Specific Window appears.

3. Edit parameter as desired.
4. Click *Save*.
The TPF/CPF file is updated.

Deleting parameters

From the Toolbar,

1. Click *Show*.
The TPF/CPF Window appears.



2. Click *Delete* for the parameter you wish to remove.
The parameter is deleted.

Working with Parameter Files—TPF/CPF

Whenever you tag a variable, it appears in the TPF/CPF. You can view the TPF/CPF by clicking *Show* on the Main Toolbar.

Creating Files

From the Toolbar, click *New*.



If parameters are already specified, the DAS will ask if you wish to delete them before creating a new file. You can work on only one TPF/CPF at a time.

Importing Files

1. From the Toolbar, click *Import*.
The Import File Window appears.
2. Locate the file you wish to import.
3. Click *Import*.
A notation appears next to the variables you tagged.
For example, **Used as: Average**

Editing Files

1. From the Toolbar, click *Show*.
The TPF/CPF appears in a pop-up window.
2. Click *Edit* for the variable you wish to change.



This is the same process as the one for editing a parameter.

Saving Files

1. From the Toolbar, click *Save*.
The Title Window appears.
2. Type a title for your table in the "Title" field.
3. Click *Save*.

Running Files and Getting Output***Running Files***

1. From the Toolbar, click *Run*.
The Title Window appears.
2. Type a title for your table in the "Title" field.
3. Click *Run*.
The Confirm Input Window appears.
4. Review the parameters.
5. Click *Run*.
The "Processing..." message appears.

Tables will appear on the screen after they are processed. Correlations will not be displayed online; instead a note will appear informing you that the output file is ready for download.



When running correlation matrices, you must also choose either SPSS or SAS format for your output file.

Printing Tables

In the Table Window,

1. Click *Printer-Friendly Version*.
The printer-friendly version appears.
2. From your browser toolbar, click  , or File > Print.

Saving output files**Tables**

You can save the table in two formats: as a comma-separated file that can be opened in Excel or as a web page.

In the Table Window,

Click *Download Comma-Separated Format*.

Or

Click *Download Printer-Friendly Web Page*.

Correlations

In the Correlation Window, click *Download Output*.

Tags: Descriptions and Instructions ► Tags are used to specify how a variable functions in your table. There are two sets of tags: one for tables and the other for correlation matrices.

For Tables

The tags for creating tables are:

- Average, Mean, Percent>, Sum
- Centile, Centile>0
- Column_Cat, Column_Cut, Column_Lump,
- Row_Cat, Row_Cut, Row_Lump,
- Span_Cat, Span_Cut, Span_Lump,
- By_Cat, By_Cut, By_Lump,
- And_Filter, Or_Filter
- Weight

Types of variables and tags

As mentioned previously, there are two types of variables—Categorical and Continuous. When you select a variable and click *Select a Tag* in the Variable Information Window, a list of tag options appears. For categorical variables, you will see all the tags options. However, for continuous variable, you will not see Cat (Column_Cat, Row_Cat, Span_Cat, and By_Cat) tags or Lump (Column_Lump, Row_Lump, Span_Lump, and By_Lump) tags.

Some tags work in conjunction with others tags and will appear as an option only after you have tagged the required tags.

- By Tags: You must first create a parameter using only one of the following: Average, Mean, or Percent> tag.
- Or_Filter: You must first create a parameter using an And_Filter tag.

Tag names and what they mean

Many tag names have two parts. For example,

1 ————— 2
Column_Cat

The first part indicates where the estimates will appear in the table, whether estimates for a variable will appear as rows (Row), columns (Column), subtables (Span), or groups of columns (By) for an Average, Mean, or Percent> variable. The second part indicates what will be displayed. Cat tags (Column_Cat, Row_Cat, Span_Cat, and By_Cat) will include all the categories for a variable. Lump tags (By_Lump, Column_Lump, Row_Lump, and Span_Lump) will allow you to create new groups by combining existing categories.

Cut tags allow you to divide continuous variables and categories into customized categories.

Other tags function as indicated by their name: Average, Mean, Percent>, Sum, Centile, Centile>0, Filter, and Weight.

Each tag is explained in detail below.



The variables used in the examples may not be in the DAS you are using.

Column Tags

There are four types of column tags:

1. Tags that produce a single column of estimates (Average, Mean, Sum, and Percent>).
2. Tags that produce a set of columns showing a complete distribution (Column_Cat, Column_Cut, and Column_Lump).
3. Tags that produce columns showing percentile values at five (Centile) or six (Centile>0) distribution points.
3. One tag in the IPEDS DAS that produces a column of total counts (Sum).

Each tag is discussed below.

Average**Description**

The Average tag creates a column of average estimates. This tag will include positive values, but not zero and missing values. To include zeros, use a Mean tag.

Restrictions

When used with a By_Cut, By_Cat, or By_Lump tag, you can specify only one Average tag. Otherwise, you can define several Average tags in a single table. An Average tag can be included in a table with Percent> and Mean tags, but not with Column_Cut, Column_Cat, Column_Lump, Centile, or Centile>0 tags.

Example

Average variable: *Aid total amount 1999–2000 (TOTAID)*

Creates the middle column in this table:

Percentage of undergraduates who received aid and the average and mean amount received			
	Received aid	Average aid amount	Mean aid amount
Total	55	\$6,206	\$3,433
Attendance pattern			
Full-time, full-year	72	8,474	6,140
Part-time or part-year	45	3,902	1,741

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02



The variable labels have been edited.

Instructions

In the Average Window,

1. Edit the label if desired.
2. Click Save.



The instructions are the same for the Average, Mean, Sum, Centile, and Centile>0 tags.

Mean

Description

The Mean tag creates a column of mean estimates. This tag will include cases with values of zeros and exclude those with missing values (-1). To exclude zeros, use an Average tag.

Restrictions

When used with a By_Cut, By_Cat, or By_Lump tag, you can specify only one Mean tag. Otherwise, you can define several Mean tags in a single table. Mean tags can be included in a table with Percent> and Average tags, but not with Column_Cut, Column_Cat, Column_Lump, Centile, or Centile>0 tags.

Example

Mean variable: *Aid total amount 1999–2000* (TOTALID)
 Creates the third column in this table:

Percentage of undergraduates who received aid and the average and mean amount received			
	Received aid	Average aid amount	Mean aid amount
Total	55	\$6,206	\$3,433
Attendance pattern			
Full-time, full-year	72	8,474	6,140
Part-time of part-year	45	3,902	1,741

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02



The variable labels have been edited.

Instructions

In the Mean Window,

1. Edit the label if desired.
2. Click Save.



The instructions are the same for the Average, Mean, Sum, Centile, and Centile>0 tags.

Percent>**Description**

The Percent> tag creates a column of percentages in which the numerator is the weighted number of cases greater than the cut-point value), and the denominator is the weighted number of cases greater than -1. For example, a Percent> specification for *Income* (INCOME) with a cut-point value of 74999.5 will show the percentage of cases with incomes of \$75,000 or more. One of the most commonly used cut points for this tag is 0.5, which yields the percentage with nonzero values for student aid variables (the percentage receiving aid).

Restrictions

When used with a By_Cut, By_Cat, or By_Lump tag, you can specify only one Percent> tag. Otherwise, you can define several Percent> tags in a single table. Percent> tags can be included in a table with Average and Mean tags, but not with Column_Cut, Column_Cat, Column_Lump, Centile, or Centile>0 tags.

Example

Percent> variable: *Aid total amount 1999–2000* (TOTALID)

Greater than (>) value is 0.5.

Creates the first column in this table:

Percentage of undergraduates who received aid and the average and mean amount received			
	Received aid	Average aid amount	Mean aid amount
Total	55	\$6,206	\$3,433
Attendance pattern			
Full-time, full-year	72	8,474	6,140
Part-time of part-year	45	3,902	1,741

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02



The variable labels have been edited.

Instructions

In the Percent> Window,

1. Edit the label if desired.
2. Enter a cut-point value.
3. Click Save.

Sum

Description

The Sum tag creates a column of sum estimates. Positive values and zero values are included.



This tag is available in the IPEDS DAS only.

Restrictions

When used with a By_Cut, By_Cat, or By_Lump tag, you can specify only one Sum tag. Otherwise, you can define several Sum tags in a single table. A Sum tag can be included in a table with Average, Mean, and Percent> tags, but not with Column_Cut, Column_Cat, or Column_Lump tags.

Example

Sum variable: *Black, non-Hispanic total* (EFRACE18)

Creates:

The total number of Black, non-Hispanic men and women enrolled for credit	
	Black, non-Hispanic
Total	1,965,180
At least 2, but less than 4 years	844,378
4 or more years	1,050,694
Less than 2 years (below associate's degree)	70,108
Source: Postsecondary Institution Data (IPEDS:2002) 09/19/04	

Instructions

In the Sum Window,

1. Edit the label if desired.
2. Click Save.



The instructions are the same for the Average, Mean, Sum, Centile, and Centile>0 tags.

Centile

Description

The Centile tag generates percentile columns from continuous variables. It produces values for five percentiles: 10th, 25th, 50th (median), 75th, and 90th. The 10th percentile value represents the lowest decile, and the 90th percentile value represents the highest decile.

Restrictions

You cannot specify a subset of the percentiles. However, you can delete the unwanted columns in your output file (table).

The Centile tag can be used only as a column variable and cannot be used in combination with any other column tag. Minimum and maximum values are entered automatically.

Example

Centile variable: *Hours worked per week* (NDHOURS)

Creates:

Average hours worked per week at the 10th, 25th, 50th, 75th, and 90th centiles by unsubsidized federal loan amount					
	Centile				
	10th	25th	50th	75th	90th
Total	0.13	9.56	26.45	39.68	44.92
Unsubsidized federal loan recipient					
Did not receive	0.13	10.84	29.68	40.08	44.95
\$1 - \$1,499	0.18	8.49	22.73	35.64	43.28
\$1,500 - \$2,499	0.16	13.07	24.06	39.52	39.95
\$2,500 - \$3,999	0.18	5.56	20.53	30.88	41.57
\$4,000 or more	0.15	10.12	20.02	30.48	41.86

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Centile Window,

1. Edit the label if desired.
2. Click Save.



The instructions are the same for the Average, Mean, Sum, Centile, and Centile>0 tags.

Centile>0

Description

The Centile>0 tag generates percentile columns in your table from continuous variables. It filters out all zero values and starts with the first valid nonzero value. This tag produces values for five percentiles: 10th, 25th, 50th (median), 75th, and 90th. The 10th percentile value represents the lowest decile, and the 90th percentile value represents the highest decile. In addition, a sixth column will be generated showing the percentage of cases with a zero value. If your selected variable does not have zero values, Centile>0 will not generate values for this column. In such cases, the output for the Centile>0 tag will be identical to the output generated by a Centile tag.

Restrictions

You cannot specify a subset of the percentiles. However, you can delete the unwanted columns in your output file (table).

Centile>0 can be used only as a column variable and cannot be used in combination with any other column tag. Minimum and maximum values are entered automatically.

Example

Centile>0 variable: *Hours worked per week* (NDHOURS)

Creates:

Average hours worked per week at the 0th, 10th, 25th, 50th, 75th, and 90th centiles by unsubsidized federal loan amount						
	Centile					
	10th	25th	50th	75th	90th	Zero
Total	12.77	19.63	32.37	41.48	47.87	19.92
Unsubsidized federal loan recipient						
Did not receive	14.46	22.73	35.13	39.99	50.07	19.89
\$1 - \$1,499	10.88	19.53	29.26	39.49	41.07	21.75
\$1,500 - \$2,499	12.20	19.49	29.11	40.27	43.78	15.82
\$2,500 - \$3,999	9.57	14.87	24.60	36.24	42.98	22.28
\$4,000 or more	10.15	14.80	23.57	34.59	42.54	16.77

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Centile>0 Window,

1. Edit the label if desired.
2. Click Save.



The instructions are the same for the Average, Mean, Sum, Centile, and Centile>0 tags.

Column_Cat

Description

The Column_Cat tag creates a column of estimates for each variable category. To select or limit the categories shown as columns, use a Column_Lump tag instead.



Some categorical variables may have values of zero. For example, for the variable *Applied for financial aid* (AIDAPP), zero values indicate that the survey respondent did not apply for financial aid. If you use a Row_Cat, Column_Cat, By_Cat, or Span_Cat tag, the zero values will not be displayed. To include the zero values in your table, use a Row_Lump, Column_Lump, By_Lump, or Span_Lump tag instead.

To include the zero values in your table, use the following lump categories:

- 1 0 Did not apply
- 2 1 Applied for aid



The first numbers (1 and 2) are the lump numbers, and the adjacent numbers are the original category numbers.

Restrictions

This tag appears as an option for categorical variables only. You can specify only one Column_Cat variable per table.

Example

Column_Cat variable: *Gender* (GENDER)

- 1 Male
- 2 Female

Creates:

	Male	Female
Total	44	56
Race-ethnicity		
White, non-Hispanic	44	56
Black, non-Hispanic	37	63
Hispanic or Latino	44	56
Asian	49	51
American Indian/Alaska Native	39	61
Native Hawaiian/other Pacific Islander	46	54
Other	52	48
More than one race	47	53

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Column_Cat Window,

- 1. Edit the default labels if desired.
- 2. Click Save.



The instructions are the same for the Column_Cat, Row_Cat, Span_Cat, and By_Cat tags.

Column_Cut

Description

The Column_Cut tag allows you to divide a continuous variable into categories and creates a column of estimates for each category.



You can also use this tag to create customized categories for a categorical variable, but it is easier to use a Column_Lump tag instead.

Restrictions

A cut-point value must be a numeral with a decimal point (e.g., 999.5), and the first cut-point value must be the smallest value of the range. You can use only one Column_Cut tag per table.

How to create cut points

For *Aid total amount 1999–2000 (TOTAID)*, the cut points below will create two categories: those who did not receive aid and those who did.

CUT POINT	LABEL
-0.5	No aid
0.5	Received aid

To break up the range 0 through 2,000, use these cut points:

Zero is included.

CUT POINT	LABEL
-0.5	Below 400
399.5	400–699
699.5	700 or more

Missing values are included.

CUT POINT	LABEL
-1.5	Missing
-0.5	Below 400
399.5	400–999
999.5	1,000 or more

Missing and Zero values are included.

CUT POINT	LABEL
-1.5	Missing
-0.5	Zero (row includes only 0)
0.5	1–399 (includes >0 and <400)
399.5	400–999
999.5	1,000 or more

Example

Column_Cut variable: *Income quartiles* (PCTALL2)

Cut-point values and labels:

- 0.5 Low quartile
- 25.5 Middle quartiles
- 75.5 High quartile

Creates:

	Income quartiles		
	Low quartile	Middle quartiles	High quartile
Estimates			
Total	25	50	25
Race-ethnicity			
White, non-Hispanic	19	51	29
Black, non-Hispanic	38	48	15
Hispanic or Latino	36	47	17
Asian	33	46	21
American Indian/Alaska Native	25	55	20
Native Hawaiian/other Pacific Islander	31	46	23
Other	29	48	23
More than one race	31	50	19

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Column_Cut Window,

- Edit the variable label if desired.
- Enter the smallest cut-point value (i.e., -0.5) and the label for the first range of values (e.g., Low quartile).
- Enter the next cut-point value (e.g., 25.5) and the corresponding label (e.g., Middle quartile).

4. Continue entering values and labels until the last range is defined. You are not required to enter the upper cut-point value for the last range; a value of infinity is used by default.
5. Review the cut-point labels and edit as needed.
6. Click Save.



The instructions are the same for the Column_Cut, Row_Cut, Span_Cut, and By_Cut tags.

Column_Lump

Description

The Column_Lump tag allows you to create customized categories by grouping existing categories of a variable and displays a column of estimates for each new category.

Restrictions

This tag appears as an option for categorical variables only. You can specify only one Column_Lump variable per table.

Example

Column_Lump variable: *Parents' highest education level* (NPARED)

Existing categories:

- 1 Did not complete high school
- 2 High school diploma or equivalent
- 3 Vocational/technical training
- 4 Less than 2 years of college
- 5 2 or more years of college/associate's degree
- 6 Bachelor's degree
- 7 Master's degree or equivalent
- 8 MD, LLB, JD, or other advanced degree
- 9 PhD or equivalent

Customized categories:

- 1 1+2 High school or less
- 2 3+4+5 Some college
- 3 6+7+8+9 Bachelor's degree or higher



The first numbers (1–5) are the lump numbers, and the adjacent numbers are the original category numbers.

Creates:

	High school or less	Some college	BA or higher
Total	40	21	38
Race-ethnicity			
White, non-Hispanic	35	22	42
Black, non-Hispanic	51	24	25
Hispanic or Latino	61	17	22
Asian	36	16	49
American Indian/Alaska Native	47	24	29
Native Hawaiian/other Pacific Islander	33	16	51
Other	34	17	49
More than one race	36	27	37

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

The instructions are the same for the Column_Lump, Row_Lump, Span_Lump, and By_Lump tags.

In the Column_Lump Window,

1. Edit the variable label if desired.
2. Click the categories you wish to lump together.
3. Type a label for the new group.
4. Click *Add Lump*.
The new lump appears at the top of the window.
5. Repeat steps 2–4 until you have created all desired lumps.
6. Click *Save*.

Row Tags Rows define the denominator of the estimates in the table. You can specify rows using the Row_Cat, Row_Cut, and Row_Lump tags.

Row_Cat Row_Cat

Description

The Row_Cat tag creates a row of estimates for each variable category. That is, each category becomes a row defining the denominator of the estimates. To select or limit the categories shown as rows, use a Row_Lump tag instead.



Some categorical variables may have values of zero. For example, for the variable *Applied for financial aid (AIDAPP)*, zero values indicate that the survey respondent did not apply for financial aid. If you use a Row_Cat, Column_Cat, By_Cat, or Span_Cat tag, the zero values will not be displayed. To include the zero values in your table, use a Row_Lump, Column_Lump, By_Lump, or Span_Lump tag instead.

To include the zero values in your table, use the following lump categories.

- 1 0 Did not apply
- 2 1 Applied for aid



The first numbers (1 and 2) are the lump numbers, and the adjacent numbers are the original category numbers.

Restrictions

You can define many Row_Cat variables in a table, but this tag can be used only once for a given variable.

Example

Row_Cat variable: *Race/ethnicity (historical) 1999–2000 (RACE1)*

Categories:

- 1 White, non-Hispanic
- 2 Black, non-Hispanic
- 3 Hispanic or Latino
- 4 Asian
- 5 American Indian/Alaska Native
- 6 Native Hawaiian/Other Pacific Islander
- 7 Other

Creates:

Percentage distribution of undergraduates according to gender by race/ethnicity		
	Male	Female
Total	44	56
Race-ethnicity		
White, non-Hispanic	44	56
Black, non-Hispanic	37	63
Hispanic or Latino	44	56
Asian	49	51
American Indian/Alaska Native	39	61
Native Hawaiian/other Pacific Islander	46	54
Other	52	48
More than one race	47	53

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Row_Cat Window,

1. Edit the variable label if desired.
2. Click Save.



The instructions are the same for the Column_Cat, Row_Cat, Span_Cat, and By_Cat tags.

Row_Cut

Description

The Row_Cut tag allows you to divide a continuous variable into categories and creates a row of estimates for each category. See Column_Cut for instructions on how to create cut points.



You can use this tag to create customized categories for a categorical variable, but it is easier to use a Row_Lump tag instead.

Restrictions

A cut-point value must be a numeral with a decimal point (e.g., 999.5), and the first cut-point value must be the smallest value of the range. You can define many Row_Cut tags in a table.

Example

Row_Cut variable: *Age as of 12/31/99* (AGE)

Cut points and labels:

0.5 18 or younger

18.5 19-23

23.5 24-29

29.5 30-39

39.5 40 or older

Creates:

Percentage distribution of undergraduates according to gender by age		
	Male	Female
Total	43.72	56.28
Age as of 12/31/99		
18 or younger	41	59
19-23	45.84	54.16
24-29	46.53	53.47
30-39	40.03	59.97
40 or older	37.73	62.27

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Row_Cut Window,

1. Edit the variable label if desired.
2. Enter the smallest cut-point value (i.e., 0.5) and the label for the first range of values (e.g., 18 or younger).
3. Enter the next cut-point value (i.e., 18.5) and the corresponding label (e.g., 19–23).
4. Continue entering values and labels until the last range is defined. You are not required to enter the upper cut-point value for the last range; a value of infinity is used.
5. Review the cut-point labels and edit as needed.
6. Click Save.



The instructions are the same for the Column_Cut, Row_Cut, Span_Cut, and By_Cut tags.

Row_Lump**Description**

The Row_Lump tag allows you to create customized categories by grouping existing categories of a variable and displays a row of estimates for each of the new categories.

Restrictions

This tag appears as an option for categorical variables only.

Example

Row_Lump variable: *Parents' highest education level* (NPARED)

Existing categories:

- 1 Did not complete high school
- 2 High school diploma or equivalent
- 3 Vocational/technical training
- 4 Less than 2 years of college
- 5 2 or more years of college/associates's degree
- 6 Bachelor's degree
- 7 Master's degree or equivalent
- 8 MD, LLB, JD, or other advanced degree
- 9 PhD or equivalent

Customized categories:

- 1 1+2 High school or less
- 2 3+4+5 Some college
- 3 6+7+8+9 Bachelor's degree or higher



The first numbers (1–3) are the lump numbers, and the adjacent numbers are the original category numbers.

Creates:

	Male	Female
Estimates		
Total	44	56
Parents' highest education level		
High school or less	40	60
Some college	39	61
BA or higher	47	53

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Row_Lump Window,

1. Edit the variable label if desired.
2. Click the categories you wish to lump together.
3. Type a label for the new group.
4. Click *Add Lump*.
The new lump appears at the top of the window.
5. Repeat steps 2–4 until you have created all desired lumps.
6. Click *Save*.



The instructions are the same for the Column_Lump, Row_Lump, Span_Lump, and By_Lump tags.

Span Tags

Span tags create a subtable for each category of the Span variable. There are three Span tags: Span_Cat, Span_Cut, and Span_Lump.

Span_Cat**Description**

The Span_Cat tag creates a subtable of estimates for each variable category. Each category will be shown with its own set of rows. To select or limit the categories shown as subtables, use a Span_Lump tag instead.



Span variables may drastically increase the number of estimates in a table.

Restrictions

This tag appears as an option for categorical variables only. You can specify only one Span_Cat tag per table. The DAS will not produce a grand total row for a table containing a Span tag. (See the FAQ Section “How do I get grand totals when using a Span tag?”)

Example

Span_Cat variable: *Gender* (GENDER)

Categories:

- 1 Male
- 2 Female

Creates:

Percentage distribution of undergraduates according to type of institution by aid status and gender					
	Public 4-year	Private not-for-profit 4-year	Public 2-year	Private for-profit	More than one institution and other
MALE					
Total	33	14	42	4	7
Aid status					
No Aid	27	8	58	1	6
Received Aid	40	22	21	8	9
FEMALE					
Total	30	14	42	5	8
Aid status					
No Aid	26	9	58	2	6
Received Aid	35	19	27	9	10

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Span_Cat Window,

1. Edit the variable label if desired.
2. Click Save.



The instructions are the same for the Column_Cat, Row_Cat, Span_Cat, and By_Cat tags.

Span_Cut

Description

The Span_Cut tag allows you to divide a continuous variable into categories, and it creates subtables for each category.



You can use this tag to create customized categories for a categorical variable, but it is easier to use a Span_Lump instead.



Span variables may drastically increase the number of estimates in a table.

Restrictions

A cut-point value must be a numeral with a decimal point (e.g., 999.5), and the first cut-point value must be the smallest value of the range.

You can specify only one Span_Cut tag per table. The DAS will not produce a grand total row for a table containing a Span tag. (See the FAQ Section "How do I get grand totals when using a Span tag?")

Example

Span_Cut variable: *Total aid 1999–2000 (TOTAID)*

Cut-point values and labels:

-0.5 No aid

0.5 Received aid

Creates:

Percentage distribution of undergraduates according to type of institution by gender and aid status					
	Public 4-year	Private not-for-profit 4-year	Public 2-year	Private for-profit	More than one institution and other
NO AID					
Total	26	8	58	2	6
Gender					
Male	27	8	58	1	6
Female	26	9	58	2	6
RECEIVED AID					
Total	37	20	25	9	9
Gender					
Male	40	22	21	8	9
Female	35	19	27	9	10

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Span_Cut Window,

1. Edit the variable label if desired.
2. Enter the smallest cut-point value (i.e., -0.5) and the label for the first range of values (e.g., No aid).
3. Enter the next cut-point value (i.e., 0.5) and the corresponding label (e.g., Received aid).
4. Continue entering values and labels until the last range is defined. You are not required to enter the upper cut-point value for the last range; a value of infinity is used.
5. Review the cut-point labels and edit as needed.
6. Click Save.



The instructions are the same for the Column_Cut, Row_Cut, Span_Cut, and By_Cut tags.

Span_Lump**Description**

The Span_Lump tag allows you to create customized categories by grouping existing categories of a variable and it displays a subtable of estimates for each new category.



Span variables may drastically increase the number of estimates in a table.

Restrictions

This tag appears as an option for categorical variables only. You can specify only one Span_Lump tag per table. The DAS will not produce a grand total row for a table containing a Span tag. (See the FAQ section “How do I get grand totals when using a Span tag?”)

Example

Span_Lump variable: *Type of Institution* (SECTOR4)

Existing categories:

- 1 Public 4-year
- 2 Private not-for-profit 4-year
- 3 Public 2-year
- 4 Private for-profit
- 5 More than one institution and other

New categories:

- 1 1+2 4-year institutions
- 2 3 Public 2-year
- 3 4 Private for-profit



The first numbers (1–3) are the lump numbers, and the adjacent numbers are the original category numbers. the category “more than one institution and other” is not selected and will be omitted from the table.

Creates:

Percentage distribution of undergraduates according to gender by aid status and type of institution		
	Male	Female
ALL 4-YEAR INSTITUTIONS		
Total	45	55
Aid status		
No aid	47	53
Received aid	43	57
PUBLIC 2-YEAR		
Total	43	57
Aid status		
No aid	47	53
Received aid	34	66
PRIVATE FOR-PROFIT		
Total	40	60
Aid status		
No aid	43	57
Received aid	39	61

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the Span_Lump Window,

1. Edit the variable label if desired.
2. Click the categories you wish to lump together.
3. Type a label for the new group.
4. Click *Add Lump*.
The new lump appears at the top of the window.
5. Repeat steps 2–4 until you have created all desired lumps.
6. Click *Save*.



The instructions are the same for the Column_Lump, Row_Lump, Span_Lump, and By_Lump tags.

By Tags

By tags create a column or average, mean, total counts (IPEDS DAS only), or percent estimates for each category of the By variable.

By_Cat

Description

The By_Cat tag creates a column of averages, means, total counts (IPEDS DAS only), or percent estimates for each variable category. To select or limit the categories shown as columns, use a By_Lump tag instead.



The By_Cat tag option appears only after you have tagged a variable as Average, Mean, Sum, or Percent.



Use of a By tag can dramatically increase the number of columns in a table.

Restrictions

This tag appears as an option for categorical variables only. You can specify only one By_Cat tag per table, and it must be combined with an Average, Mean, Sum, or Percent tag.

Example

Average variable: *Aid total amount 1999–2000 (TOTAID)*

By_Cat variable: *Combinations of loans and other aid 1999–2000 (LOANAID2)*

Categories:

- 1 Aided, no loans
- 2 Student loan and other aid
- 3 Student loans only
- 4 No aid

Creates:

Average amount of aid received by loan and non-loan combinations				
	Combinations of loans and other aid			Aid total amount 1999-2000
	Aided, no loans	Student loan and other aid	Student loans only	
Total	\$2,913	\$10,625	\$5,156	\$6,206
Dependent 1998 income				
Low (under \$30,000)	3,527	10,682	4,719	7,005
Middle (\$30,000-80,000)	3,882	11,543	4,133	7,508
High (over \$80,000)	4,745	12,946	4,722	7,561
Independent 1998 Income				
Low (under \$12,000)	2,687	9,612	6,372	6,226
Middle (\$12,000-45,000)	2,014	9,012	6,088	4,753
High (over \$45,000)	1,512	9,131	6,537	3,051

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02



The “No aid” column was deleted in the table because the estimates were not applicable.

Instructions

In the By_Cat Window,

1. Edit the variable label if desired.
2. Click Save.



The instructions are the same for the Column_Cat, Row_Cat, Span_Cat, and By_Cat tags.

By_Cut

Description

The By_Cut tag allows you to divide a continuous variable into categories and creates a column of averages, means, total counts (IPEDS DAS only), or percent estimates for each variable category.



The By_Cut tag option appears only after you have tagged a variable as Average, Mean, Sum, or Percent.



Use of a By tag can dramatically increase the number of columns in a table.

Restrictions

You can specify only one By_Cat tag per table, and it must be combined with an Average, Mean, Sum, or Percent tag.

Example

Mean variable: *Student budget (adjusted) minus all aid 1999-2000* (NETCST1)
 By_Cut variable: *Income of independent students 1998* (INDEPINC)

Cut-point values and labels:

-0.5	Less than \$10,000
9999.5	\$10,000–19,999
19999.5	\$20,000–29,999
29999.5	\$30,000–49,999
49999.5	\$50,000 or more

Creates:

Mean net price for independent students by student income						
	Income of independent students 1998					Adjusted Student budget minus all aid
	Less than \$10,000	\$10,000-19,999	\$20,000-29,999	\$30,000-49,999	\$50,000 or more	
Total	\$4,744	\$5,086	\$4,669	\$4,490	\$4,329	\$4,644
Aid status						
No aid	6,335	5,919	4,692	4,373	4,254	4,821
Received aid	4,161	4,595	4,647	4,666	4,482	4,475
Grants status						
No grants	6,048	5,755	4,753	4,546	4,420	4,887
Received grants	4,163	4,490	4,545	4,358	4,042	4,312
Loan status						
No loans	5,687	5,527	4,672	4,346	4,182	4,758
Received loans	3,226	4,158	4,659	5,358	6,031	4,242

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the By_Cut Window,

1. Edit the variable label if desired.
2. Enter the smallest cut-point value (i.e., 0.5) and the label for the first range of values (e.g., Less than \$10,000).
3. Enter the next cut-point value (i.e., 9999.5) and the corresponding label (e.g., \$10,000–19,999).
4. Continue entering values and labels until the last range is defined. You are not required to enter the upper cut-point value for the last range; a value of infinity is used.
5. Review the cut-point labels and edit as needed.
6. Click Save.



The instructions are the same for the Column_Cut, Row_Cut, Span_Cut, and By_Cut tags.

By_Lump

Description

The By_Lump tag allows you to create customized categories by grouping existing categories of a variable and it will display a column of averages, means, total counts (IPEDS DAS only), or percent estimates for each variable category.



The By_Lump tag option appears only after you have tagged a variable as Average, Mean, Sum, or Percent.



Use of a By tag can dramatically increase the number of columns in a table.

Restrictions

This tag appears as an option for categorical variables only. You can specify only one By_Cat tag per table, and it must be combined with an Average, Mean, Sum, or Percent> tag.

Example

Average variable: *Stafford total* (STAFFAMT)

By_Lump variable: *Stafford loan combinations* (STAFTYPE)

Categories:

- 1 Subsidized loans only
- 2 Subsidized and unsubsidized
- 3 Unsubsidized only



By_Lump is used instead of By_Cat to exclude the 4th category: "No Stafford Loans" from the table.

Creates:

	Stafford loan combinations			Stafford total
	Subsidized loans only	Subsidized and unsubsidized	Unsubsidized only	
Total	3347.72	6209.27	4055.57	4492.08
Dependent 1998 income				
Low (under \$30,000)	3293.89	5928.21	3331.31	3759.49
Middle (\$30,000-80,000)	3397.29	4904.02	3405.72	3757.09
High (over \$80,000)	3735.96	4706.21	3837.5	3953.54
Independent 1998 Income				
Low (under \$12,000)	3274.83	6808.6	4610.49	5339.79
Middle (\$12,000-45,000)	3209.44	6684.95	5230.3	5473.29
High (over \$45,000)	3226.45	7212.81	6428.69	6277.79

Source: NCES, NPSAS:2000 Undergraduate Students 10/22/02

Instructions

In the By_Lump Window,

1. Edit the variable label if desired.
2. Click the categories you wish to lump together.
3. Type a label for the new group.
4. Click *Add Lump*.
The new lump appears at the top of the window.
5. Repeat steps 2–4 until you have created all desired lumps.
6. Click *Save*.



The instructions are the same for the Column_Lump, Row_Lump, Span_Lump, and By_Lump tags.

Filters Filters are used to subset the sample. For example, you can create a table of estimates that includes only students who attend college part time.

There are two types of filters—And_Filter and Or_Filter. Both operate using “select if” logic. Because Or_Filters work in conjunction with And_Filters, they appear as an option only after you have specified an And_Filter. Both types of filters are described below.

And_Filter

Description

There are two types of And_Filters—Integer and Cut Point. Integer filters are used for categorical variables, and cut-point filters are generally used for continuous variables.

Integer Filters

Each integer filter must be defined using whole numbers. When an integer filter has several values, the filter operates with “or” logic. For example,

Filter variable: *Race/ethnicity (historical 1999–2000)* (RACE1)

Categories:

Asian

Hispanic

Label: Asian or Hispanic

The DAS will create a table of estimates that includes only students who are Asian **or** Hispanic.

When multiple And_Filter tags are used, they operate with **and** logic. For example, if you add the following And_Filter to your TPE, the DAS will produce a table with estimates for only students who are Asian **or** Hispanic **and** who attended college full time, full year in 1999-2000.

Filter variable: *Attendance pattern 1999–2000* (ATTNSTAT)

Categories:

Full-time/full year, 1 institution

Full-time/full year, more than 1 institution.

Label: Attended full-time, full-year at one or more institutions

Cut-Point Filters

Cut-point values must be a numeral with a decimal point (i.e., 0.5). The same operation logic used for integer filters applies to cut-point filters.

Instructions

In the And_Filter Window, select the type of filter—Integer or Cut Point.

Integer Filter

1. Edit the variable label if desired.
2. For categorical variables, click the boxes for the categories you wish to include.
3. For continuous variables, enter the cut-point values and labels.
4. Click Save.

Cut-Point Filter

1. Edit the variable label if desired.
2. Select the type of filter: Greater than (> value), Less than (< value), or Between two values.
3. If you select “Greater than (> value)” or “Less than (< value)”, type the cut-point value in the Value field.

If you select “Between Value 1 and Value 2”, then two boxes appear. Type the first cut-point value in the “Value 1” field and the second cut-point value in the “Value 2” field.
4. Click Save.



The instructions are the same for the And_Filter and Or_Filter tags.

Or_Filter**Description**

The Or_Filter tag operates using “or” logic in relation to the subgroups defined by the And_Filters. There are two types of Or_Filters—Integer and Cut Point. Integer filters are used for categorical variables, and cut-point filters are generally used for continuous variables.

To be included in the table, the case must pass all of the And_Filters conditions **or** it must pass the Or_filter condition(s). This is true regardless of the order in which the filters are tagged.



The DAS can perform logic using these conditions: (A1 or A2) and (B1 or B2) or (C1 or C2), but not these (A and B) or (C and D).

Example

And_Filter variable: *Income of parents of dependent students* (DEPINC)

Select cut-point filter

Value < 19999.5

Label: Family income below \$20,000

And_Filter variable: *Attendance pattern 1999-2000* (ATTNSTAT)

Select integer filter

Categories:

Full-time/full-year, 1 institution

Full-time/full-year, more than 1 institution.

Label: Attended full-time, full-year at one or more institutions

Or_Filter variable: *Pell Grant total 1999–2000* (PELLAMT)

Select cut-point filter

Value > 0.5

Label: Received Pell grant

The DAS will create a table that includes estimates for students with family incomes below \$20,000 **and** who attended school full-time at one or more institutions **or** students who received Pell grants.

Instructions

In the Or_Filter Window, select the type of filter—Integer or Cut-Point.

Integer Filter

1. Edit the variable label if desired.
2. For categorical variables, click the boxes for the categories you wish to include. For continuous variables, enter the cut-point values and labels.
43. Click Save.

Cut-Point Filter

1. Edit the variable label if desired.
2. Select the type of filter: Greater than (> value), Less than (< value), or Between two values.
3. If you select “Greater than (> value)” or “Less than (< value)”, type the cut-point value in the Value field.
If you select “Between Value 1 and Value 2”, then two boxes appear. Type the first cut-point value in the “Value 1” field and the second cut-point value in the “Value 2” field.
4. Click Save.



The instructions are the same for the And_Filter and Or_Filter tags.

Weights**Description**

Weights can be used to restrict the data to a population subset or to ensure that the final output represents the population you wish to describe. If a data set contains more than one weight, the variables will be listed under Survey→Weights in the Variables List. They are shown in red. If you do not know which weight to use, read the variable descriptions or check the NCES methodology reports at <http://nces.ed.gov>. Each survey should have a corresponding report.

Some data sets have only one weight, which the DAS will automatically add to your TPF.

Different Populations

In some longitudinal surveys, weights are used to distinguish the different student populations for each survey year. For example, the Beginning Postsecondary Students (BPS) Longitudinal Study conducted in 1996, 1998, and 2001 has five analysis weight variables. They are the following:

1. *Weight for 1996 cross-sectional (WTA00)*
Represents respondents to the NPSAS:96 survey who were determined to belong to the BPS:96 cohort of first-time beginners. Use this weight for cross-sectional analysis of the first year of enrollment (1995–96).
2. *BPS 1996/1998 longitudinal weight (WTE00)*
Represents respondents to the BPS first follow-up survey in 1998 (BPS:96/98).
Use this weight for longitudinal comparisons of the responses of the BPS cohort in 1995-96 (NPSAS:96 survey) and in 1996/98 (BPS:96/98 survey).
3. *BPS 1996/2001 cross-sectional weight (WTB00)*
Represents respondents to the BPS survey in 2001.
Use this weight for cross-sectional analysis of respondents to the BPS survey in 2001.
4. *Longitudinal weight for 96, 98, 01 respondents (WTC00)*
Represents respondents to NPSAS:96, BPS:96/98, and BPS:1996/2001. Use this weight for longitudinal analysis of students who responded to all three surveys.
5. *Longitudinal weight for respondents in 96, 01 (WTD00)*
Represents respondents to NPSAS:96 and BPS:1996/2001. Use this weight for longitudinal analysis of students who responded to both surveys.

Population Subsets

In some cross-sectional surveys, there are subsets of the general survey population. For example, the data for the 1996 National Postsecondary Student Aid Survey (NPSAS) comes from two sources—(1) a Computer Assisted Telephone Interview (CATI), and (2) data gathered from the education institution (CADE), the Central Processing System (CPS), or the National Student Loan Data System (NSLDS). As a result, there two weight variables—*CATI weight for undergraduates* (WTB00) and *Full sample weight for undergraduates* (WTA00). The CATI respondents are a smaller group than the full sample, but the CATI weight provides the same weighted numbers as the full sample weight.

The general guideline is as follows:

- Use the CATI weight if you cross a CATI variable with another CATI variable
- Use the CATI weight if you cross a CATI variable with a full sample variable
- Use the full sample weight if you cross a full sample variable with another full sample variable

For Correlations

The following tags are used only for creating correlation parameter files. There are three tags that can be used for both tables and correlation matrices—`And_Filter`, `Or_Filter`, and `Weight` tags. These parameters label output and define the size of the sample used to calculate the pairwise correlations. The remaining parameters—`Continuous`, `Cut`, `Each`, and `Lump`—define the variables in the matrix. The code for each tag, as it appears in the CPF, is shown in parentheses next to the name. In addition, `And_Filter` is shown as *F*, `Or_Filter` as *R*, and `Weight` as *W*.

Continuous (C) Description

The `Continuous` tag adds a continuous variable to the correlation. The DAS will automatically assign the default code of `-1` for missing values.

Restrictions

This tag appears as an option for continuous variables only.



Due to string length limits in SAS and SPSS, you may need to shorten the DAS labels.

Cut (X) Description

The Cut tag allows you to define ranges for a continuous variable and to compare the new groups with a comparison group. Each defined range is a dummy variable. The unspecified range is the comparison group. Missing values must also be identified to exclude them from the comparison. The default missing values code is -1. See the example below.



Due to string length limits in SAS and SPSS, you may need to shorten the DAS labels.

Example

To create the following comparisons for the variable *Age as of 12/31/95 (AGE)*,

1. 24 years or younger vs. 25-29 years
2. 24 years or younger vs. 30 years or older

you would:

1. Enter 24.6 as the lower-cut value, 29.5 as the upper-cut value, and 25-29 years as the label in the first row.
2. Enter 29.6 as the lower-cut value, 99.5 as the upper-cut value, and 30 years or older as the label in the second row.

The range that is not specified, 13–24 years (24 years or younger), is the comparison group.

Instructions

In the Cut Window,

1. Edit the variable label if desired.
2. Enter the missing value code.
3. Enter the lower-cut value.
This value must be a fraction with a decimal point.
4. Enter the upper-cut value.
This value must be a fraction with a decimal point.
5. Enter the label for the defined range (i.e., 25-29 years old).
6. Repeat steps 3–4 until all dummy variables have been defined.
7. Click Save.

Each (E) Description

The Each tag compares a single category with each of the other categories of a categorical variable. All categories are included in the correlation matrix. You must enter the Base Category Code, which is the value of the comparison group. See the example below. The DAS will automatically assign the default code of -1 for missing values.



Some variables contain valid zero values (such as a variable indicating the number of months married, where zero indicates students who were never married). If you tag one of these variables with the Each tag, the DAS will treat the zero values as missing. To include them in your comparison, use a Lump tag instead.



To group categories together to create new categories or to exclude a category, use the Lump tag instead.



Due to string length limits in SAS and SPSS, you may need to shorten the DAS labels.

Restrictions

Use this tag with categorical variables only.

Example

The variable *Citizenship status 2001* (QFCITZN) contains these categories:

- 1 U.S. citizen/U.S. national
- 2 Resident alien
- 3 Student with visa

When you select U.S. citizen/U.S. national as the Base Category Code, it becomes the comparison group. The remaining categories become dummy variables. U.S. citizen/U.S. nationals will be compared first with resident aliens and second with students with visas. Resident aliens are not compared to students with visas.

Instructions

In the Each Window,

1. Edit the variable label if desired.
2. Select a Base Category Code.
3. View the Base Category pull-down menu for a list of valid values. Edit the value labels for the remaining categories if desired.
4. Click *Save*.

Lump (L) Description

The Lump tag allows you to define a new category by grouping existing categories together and to compare that category with the comparison group. Each defined category must be tagged separately. The untagged group(s) becomes the comparison group. As such, you must also tag values that are not needed for your analysis to exclude them from the comparison group. You must also identify missing values to exclude them from the comparison. The default missing values code is -1, which is shown in the “missing values” field. See the example below.

Restrictions

Use this tag with categorical variables only.



Due to string length limits in SAS and SPSS, you may need to shorten the DAS labels.

Example

The variable *Race/ethnicity (historical) 1999–2000 (RACE1)* has seven categories. They are:

1. White, non-Hispanic
2. Black, non-Hispanic
3. Hispanic or Latino
4. Asian
5. American Indian/Alaska Native
6. Native Hawaiian/Other Pacific Islander
7. Other

To make the following comparisons:

1. Black, non-Hispanic and Hispanic vs. White, non-Hispanic
2. Asian, Native Hawaiian/Other Pacific Islander and American Indian/Alaska Native vs. White, non-Hispanic

You would tag this variable three times:

- Lump 1 2+3 Black and Hispanic
 Lump 2 4+5+6 Asian, Native Hawaiian/Other Pacific Islander and American Indian/Alaska Native
 Lump 3 7 Other

Category 1 (White, non-Hispanic) is not tagged so it becomes the comparison group. Category 7 (Other) is tagged so it is not included in the comparison group. If you do not tag category 7, the comparison group will be 1 and 7 (White, non-Hispanic and other).

Instructions

In the Lump Window,

1. Enter a variable label if desired.
2. Select the Missing Value Code from the pull-down menu.
For many categorical variables with a reserved code of zero,
change the missing value to zero.
3. Check the boxes for the categories you wish to group.
4. Enter a label for your lump.
5. Click Save.

FREQUENTLY ASKED QUESTIONS

▶ The following is a list of frequently asked questions about using the DAS.

Choosing Variables



How do I choose between two similar variables?

Read the descriptions for both variables and select the one that is most appropriate for your table.



How do I pick the correct weight variable?

Read the descriptions for all the Weights and pick the best one for your dependent (column) variable. For data sets that have only one weight, the DAS will automatically add it to your TPF.

Viewing Variable Information



How do I view the frequency distribution for a variable?

In the Variable List,

1. Click the variable you want to look at.
The Variables Information Window appears.
2. Scroll to "Statistics."



How do I view the range and mean for a continuous variable?

Read the variable description in the "Statistics" section of the Variables Information Window.



How do I print a list of all variables?

Click *View/download list of variables* in the DAS and then *Print Variable List*.

Updating and Creating New Variables



Do variables get updated?

Yes. Variables may be updated after they have been released. To view a list of new/updated variables, go to the DAS Updates page (<http://nces.ed.gov/das/updates/index.asp>) and click the appropriate data set. You will see the variable name, status (type of change), and date of the change. New variables will be noted as such.



Do new variables get added?

Yes. New variables are created for analysis and added to the DAS after they have been approved by NCES. To view a list of new/updated variables, go to the DAS Updates page (<http://nces.ed.gov/das/updates/index.asp>) and click the appropriate dataset. You will see the variable name, status (type of change), and date of the change. New variables will be noted as such.



How do I create a new variable?

This option is not available. However, you can contact Aurora D'Amico of NCES at Aurora.D'Amico@ed.gov or (202) 502-7334 and request that the new variable be added. You must define the variable with SAS or SPSS code.

Using Tags



How do I get a median?

Use a Centile or Centile>0 tag.



How do I combine several categories for a variable into a new category?

Use a Column_Lump, Row_Lump, By_Lump, or Span_Lump tag.



How do I define a range from a continuous variable and have it appear as a row?

Use a Row_Cut tag.

**How do I define the subsample for estimates based on a range of a continuous variable?**

Use the And_Filter tag. Select the cut-point filter and the “Between Value 1 and Value 2” option.

For example, to define the subsample with incomes between \$20,000 and \$30,000 for the variable INCOME, use the following Filter parameters:

And_Filter INCOME Gross income

Value 1 19999.5

Value 2 30000.5

Label: Between \$20,000 and \$30,000

**How do I get grand totals when using a Span tag?**

Import the TPF, change the Span tag to a Row tag, and run the new table. The DAS will not provide grand totals for the TPF containing the Span tag.

For Correlation Matrices**How do I adjust *t*-tests in regressions for design effects?**

A: Divide the calculated *t* from SPSS or SAS by the design effect computed for the dependent variable. The design effects are the DEFTs displayed at the bottom of the correlations output file.

DAS RESOURCES ► NCES provides many resources to support DAS users. You can find these on the DAS website (<http://nces.ed.gov/das>). Each is discussed below.

Tables Library The Tables Library contains tables that focus on various topics in postsecondary education. These topics include student demographics, access to postsecondary education, persistence toward and attainment of a degree, financial aid, and institutional characteristics. The Library also allows you to view and download standard error tables, TPFs, and CPFs. You can edit and rerun TPFs and CPFs to produce customized tables.

Tables in reports published after August 2003 are posted in the Tables Library, while tables for reports released before this date are located in the Analysis Reports section of the DAS website. There are thousands of TPFs and CPFs available for download. To view a full list of analysis reports and table files, visit the DAS website (<http://nces.ed.gov/das>) and click *Tables Library* or *Analysis Reports*.

User Help Center The User Help Center contains web demonstrations of DAS features, comprehensive tutorials, a copy of the help files found in the application, and *t*-test tools. You can also find out how to get technical assistance. For a current selection of resources, visit the User Help Center at <http://nces.ed.gov/dasol/help/index.asp>,

Getting Technical Assistance If you have questions about the DAS, contact:

John Vavricka
MPR Associates, Inc.
2150 Shattuck Ave. Suite 800
Berkeley, CA 94704
(510) 849-4942
das-support@mprinc.com

or

Aurora D'Amico
NCES
1900 K Street, NW, Room 8115
Washington, DC 20006
(202) 502-7334
Aurora.D'Amico@ed.gov