

Aggregate - Calculating Values Across Cases¹

It is possible to get means, sums, and a variety of other measures across cases in SPSS by using "AGGREGATE". Following is a program that reads a few variables for four cases. It then uses Aggregate to calculate the mean and sum of annual income, by gender. The file created by the aggregate procedure is then merged with the original data file. At each stage, "list cases" is used to show the result.

```
title 'aggregate.sps' .
DATA LIST FREE / ID * name (A9) AGE (f3) hinch (f3) weight (f3)
              aninc (f5) gender (A) .
BEGIN DATA
1 George    22 72 203 35000 m
2 Frank    67 65 180 42000 m
3 Sally    27 62 120 65000 f
4 Michelle 33 66 145 43000 f
END DATA .
variable label  hinch 'Height in Inches'
                aninc 'Annual Income' .
sort cases by gender .
list var = all .
save outfile = 'c:\trash\aggl.sav' .
```

* The line "/ break = gender" tells SPSS that you want separate values by gender.

```
Aggregate outfile = *
  / break = gender
  / avginc = mean(aninc)
  / totinc = sum(aninc)
  / count = N .
variable label
  avginc = 'Mean Income for gender'
  totinc = 'Total Income for gender'.
list var = all .
match files
  table = * /
  file = 'c:\trash\aggl.sav' /
  by gender .
list var = all .
```

YOU WILL FIND OUTPUT FROM THIS JOB, AND MORE INFORMATION ON THE NEXT TWO PAGES.

¹Prepared by Patty Glynn, University of Washington, May 4, 2002.

```

title 'aggregate.sps' .

DATA LIST FREE / ID * name (A9) AGE (f3) hinch (f3) weight (f3)
  aninc (f5) gender (A) .
BEGIN DATA
1 George  22 72 203 35000 m
2 Frank   67 65 180 42000 m
3 Sally   27 62 120 65000 f
4 Michelle 33 66 145 43000 f
END DATA .
variable label      hinch 'Height in Inches'
                   aninc 'Annual Income' .

sort cases by gender .
list var = all .

```

List

ID	NAME	AGE	HINCH	WEIGHT	ANINC	GENDER
3.00	Sally	27	62	120	65000	f
4.00	Michelle	33	66	145	43000	f
1.00	George	22	72	203	35000	m
2.00	Frank	67	65	180	42000	m

Number of cases read: 4 Number of cases listed: 4

```

save outfile = 'c:\trash\agg1.sav' .
* The line "/ break = gender" tells SPSS that you want separate values by gender.
Aggregate outfile = *
  / break = gender
  / avginc = mean(aninc)
  / totinc = sum(aninc)
  / count = N .
variable label
  avginc = 'Mean Income for gender'
  totinc = 'Total Income for gender'.
list var = all .

```

List

GENDER	AVGINC	TOTINC	COUNT
f	54000.00	108000.0	2
m	38500.00	77000.00	2

Number of cases read: 2 Number of cases listed: 2

```

match files
  table = * /
  file = 'c:\trash\agg1.sav' /
  by gender .
list var = all .

```

List

GENDER	AVGINC	TOTINC	COUNT	ID	NAME	AGE	HINCH	WEIGHT	ANINC
f	54000.00	108000.0	2	3.00	Sally	27	62	120	65000
f	54000.00	108000.0	2	4.00	Michelle	33	66	145	43000
m	38500.00	77000.00	2	1.00	George	22	72	203	35000
m	38500.00	77000.00	2	2.00	Frank	67	65	180	42000

Number of cases read: 4 Number of cases listed: 4

Available functions in Aggregate:

SUM	. . .	Sum
MEAN	. . .	Mean
SD	Standard deviation
MAX	. . .	Maximum
MIN	. . .	Minimum
PGT	. . .	% of cases greater than value
PLT	. . .	% of cases less than value
PIN	. . .	% of cases between values
POUT	. . .	% of cases not in range
FGT	. . .	Fraction greater than value
FLT	. . .	Fraction less than value
FIN	. . .	Fraction between values
FOUT	. . .	Fraction not in range
N	Weighted number of cases
NU	Unweighted number of cases
NMISS	. .	Weighted number of missing cases
NUMISS	. .	Unweighted number of missing cases
FIRST	. .	First nonmissing value
LAST	. . .	Last nonmissing value