Subject: Aggregating household level data
Posted by Lukresha on Mon, 25 Jul 2016 10:16:05 GMT
View Forum Message <> Reply to Message

Hello,

I would like to aggregate the household level data mv717 (occupation grouped) to cluster level data.

What is the recommended way of going about it as I have used this command:
collapse mv717, by(hv001)

and from the output (shown below) I do not know whether I should only focus on the whole numbers that are similar to the initial mv717 categories (shown below) or all the values in the output table are important.
tab mv717

(mean) mv717	Freq.	Percent	Cum.
0	40	4.29	4.29
.25	1	0.11	4.39
.2857143	1	0.11	4.50
.3	1	0.11	4.61
.3333333	2	0.21	4.82
.4444444	1	0.11	4.93
.5	3	0.32	5.25
.5714286	1	0.11	5.36
.6	1	0.11	5.47
.6666667	2	0.21	5.68
.7142857	1	0.11	5.79
.75	1	0.11	5.89
.8	2	0.21	6.11
.8571429	1	0.11	6.22
.8888889	1	0.11	6.32
.9230769	1	0.11	6.43
1	18	1.93	8.36
1.142857	4	0.43	8.79
1.222222	1	0.11	8.90
1.230769	1	0.11	9.00
1.25	2	0.21	9.22
1.272727	1	0.11	9.32
1.285714	1	0.11	9.43
1.307692	1	0.11	9.54
1.333333	7	0.75	10.29
1.428571	1	0.11	10.40
1.454545	1	0.11	10.50
1.5	3	0.32	10.83

1.555556	1	0.11	10.93
1.6	5	0.54	11.47
1.625	1	0.11	11.58
1.636364	1	0.11	11.68
1.684211	1	0.11	11.79
1.714286	1	0.11	11.90
1.75	1	0.11	12.00
1.777778	1	0.11	12.11
1.8	3	0.32	12.43
1.818182	2	0.21	12.65
1.833333	1	0.11	12.75
1.846154	2	0.21	12.97
1.857143	2	0.21	13.18
1.875	1	0.11	13.29
1.913043	1	0.11	13.40
1.928571	1	0.11	13.50
2	26	2.79	16.29
2.083333	1	0.11	16.40
2.086957	1	0.11	16.51
2.111111	1	0.11	16.61
2.117647	1	0.11	16.72
2.125	2	0.21	16.93
2.133333	1	0.11	17.04
2.142857	1	0.11	17.15
2.166667	2	0.21	17.36
2.181818	1	0.11	17.47
2.214286	1	0.11	17.58
2.222222	5	0.54	18.11
2.25	8	0.86	18.97
2.272727	2	0.21	19.19
2.285714	4	0.43	19.61
2.3	1	0.11	19.72
2.307692	1	0.11	19.83
2.333333	5	0.54	20.36
2.384615	1	0.11	20.47
2.4	6	0.64	21.11
2.409091	1	0.11	21.22
2.428571	4	0.43	21.65
2.444444	2	0.21	21.86
2.454545	1	0.11	21.97
2.466667	1	0.11	22.08
2.5	1	0.11	22.19
2.555556	1	0.11	22.29
2.571429	4	0.43	22.72
2.588235	1	0.11	22.83
2.6	2	0.21	23.04
2.615385	1	0.11	23.15
2.625	1	0.11	23.26

2.666667	13	1.39	24.65
2.714286	4	0.43	25.08
2.727273	1	0.11	25.19
2.75	1	0.11	25.29
2.777778	1	0.11	25.40
2.8	4	0.43	25.83
2.833333	1	0.11	25.94
2.857143	2	0.21	26.15
2.875	3	0.32	26.47
2.888889	2	0.21	26.69
2.9	1	0.11	26.80
2.909091	1	0.11	26.90
2.92	1	0.11	27.01
3	25	2.68	29.69
3.058824	1	0.11	29.80
3.076923	3	0.32	30.12
3.1	1	0.11	30.23
3.111111	1	0.11	30.33
3.125	1	0.11	30.44
3.142857	2	0.21	30.65
3.166667	2	0.21	30.87
3.2	4	0.43	31.30
3.222222	1	0.11	31.40
3.230769	3	0.32	31.73
3.25	2	0.21	31.94
3.272727	1	0.11	32.05
3.294118	1	0.11	32.15
3.3	1	0.11	32.26
3.315789	1	0.11	32.37
3.333333	8	0.86	33.23
3.375	2	0.21	33.44
3.4	4	0.43	33.87
3.428571	4	0.43	34.30
3.4375	2	0.21	34.51
3.444444	3	0.32	34.83
3.466667	1	0.11	34.94
3.473684	1	0.11	35.05
3.5	17	1.82	36.87
3.538461	1	0.11	36.98
3.545455	1	0.11	37.08
3.555556	5	0.54	37.62
3.571429	4	0.43	38.05
3.583333	2	0.21	38.26
3.6	3	0.32	38.59
3.611111	1	0.11	38.69
3.625	6	0.64	39.34
3.636364	1	0.11	39.44
3.666667	1	0.11	39.55

3.692308	2	0.21	39.76
3.7	1	0.11	39.87
3.705882	1	0.11	39.98
3.714286	1	0.11	40.09
3.722222	1	0.11	40.19
3.75	5	0.54	40.73
3.777778	3	0.32	41.05
3.785714	1	0.11	41.16
3.789474	1	0.11	41.26
3.8	2	0.21	41.48
3.833333	1	0.11	41.59
3.846154	1	0.11	41.69
3.857143	2	0.21	41.91
3.875	1	0.11	42.02
3.9	1	0.11	42.12
3.909091	2	0.21	42.34
3.928571	1	0.11	42.44
3.947368	1	0.11	42.55
3.954545	1	0.11	42.66
4	103	11.04	53.70
4.111111	1	0.11	53.80
4.125	2	0.21	54.02
4.133333	1	0.11	54.13
4.153846	1	0.11	54.23
4.166667	3	0.32	54.56
4.181818	2	0.21	54.77
4.2	3	0.32	55.09
4.25	1	0.11	55.20
4.266667	1	0.11	55.31
4.277778	1	0.11	55.41
4.285714	5	0.54	55.95
4.307693	1	0.11	56.06
4.333333	7	0.75	56.81
4.357143	1	0.11	56.91
4.363636	3	0.32	57.23
4.375	1	0.11	57.34
4.384615	1	0.11	57.45
4.388889	1	0.11	57.56
4.4	3	0.32	57.88
4.428571	2	0.21	58.09
4.444445	1	0.11	58.20
4.454545	2	0.21	58.41
4.5	11	1.18	59.59
4.538462	2	0.21	59.81
4.555555	1	0.11	59.91
4.5625	1	0.11	60.02
4.571429	3	0.32	60.34
4.615385	2	0.21	60.56

4.625	1	0.11	60.66
4.666667	6	0.64	61.31
4.692307	1	0.11	61.41
4.7	1	0.11	61.52
4.714286	1	0.11	61.63
4.727273	1	0.11	61.74
4.75	6	0.64	62.38
4.769231	1	0.11	62.49
4.777778	1	0.11	62.59
4.8	11	1.18	63.77
4.8125	1	0.11	63.88
4.818182	1	0.11	63.99
4.833333	2	0.21	64.20
4.846154	1	0.11	64.31
4.857143	1	0.11	64.42
4.875	1	0.11	64.52
4.909091	2	0.21	64.74
4.9375	2	0.21	64.95
5	22	2.36	67.31
5.052631	1	0.11	67.42
5.055555	1	0.11	67.52
5.0625	1	0.11	67.63
5.076923	1	0.11	67.74
5.090909	4	0.43	68.17
5.1	1	0.11	68.27
5.142857	5	0.54	68.81
5.166667	3	0.32	69.13
5.181818	1	0.11	69.24
5.2	8	0.86	70.10
5.235294	1	0.11	70.20
5.25	3	0.32	70.53
5.333333	8	0.86	71.38
5.384615	2	0.21	71.60
5.4	3	0.32	71.92
5.428571	2	0.21	72.13
5.444445	1	0.11	72.24
5.454545	1	0.11	72.35
5.5	13	1.39	73.74
5.545455	1	0.11	73.85
5.555555	2	0.21	74.06
5.6	3	0.32	74.38
5.625	2	0.21	74.60
5.666667	10	1.07	75.67
5.7	1	0.11	75.78
5.714286	3	0.32	76.10
5.727273	1	0.11	76.21
5.733333	1	0.11	76.31
5.75	1	0.11	76.42

5.785714	1	0.11	76.53
5.8	1	0.11	76.63
5.833333	1	0.11	76.74
5.857143	1	0.11	76.85
5.888889	2	0.21	77.06
5.909091	1	0.11	77.17
6	33	3.54	80.71
6.111111	3	0.32	81.03
6.125	1	0.11	81.14
6.142857	1	0.11	81.24
6.153846	1	0.11	81.35
6.166667	1	0.11	81.46
6.2	3	0.32	81.78
6.222222	1	0.11	81.89
6.3	1	0.11	81.99
6.307693	1	0.11	82.10
6.333333	5	0.54	82.64
6.363636	2	0.21	82.85
6.375	1	0.11	82.96
6.4	2	0.21	83.17
6.421052	1	0.11	83.28
6.428571	3	0.32	83.60
6.5	14	1.50	85.10
6.545455	2	0.21	85.32
6.583333	1	0.11	85.42
6.6	2	0.21	85.64
6.642857	1	0.11	85.74
6.666667	1	0.11	85.85
6.692307	2	0.21	86.07
6.714286	4	0.43	86.50
6.75	3	0.32	86.82
6.8	2	0.21	87.03
6.857143	2	0.21	87.25
6.888889	2	0.21	87.46
6.909091	1	0.11	87.57
7	10	1.07	88.64
7.076923	1	0.11	88.75
7.090909	2	0.21	88.96
7.111111	4	0.43	89.39
7.142857	1	0.11	89.50
7.181818	1	0.11	89.60
7.2	2	0.21	89.82
7.214286	1	0.11	89.92
7.230769	1	0.11	90.03
7.25	1	0.11	90.14
7.307693	1	0.11	90.25
7.333333	8	0.86	91.10
7.4	1	0.11	91.21

```
  7.416667 |          1       0.11       91.32
       7.5 |          3       0.32       91.64
  7.538462 |          1       0.11       91.75
  7.555555 |          1       0.11       91.85
  7.571429 |          1       0.11       91.96
  7.583333 |          1       0.11       92.07
  7.636364 |          1       0.11       92.18
  7.666667 |          1       0.11       92.28
  7.714286 |          3       0.32       92.60
      7.75 |          1       0.11       92.71
  7.857143 |          2       0.21       92.93
  7.888889 |          1       0.11       93.03
         8 |         16       1.71       94.75
  8.142858 |          1       0.11       94.86
  8.166667 |          3       0.32       95.18
      8.25 |          2       0.21       95.39
  8.285714 |          1       0.11       95.50
       8.4 |          2       0.21       95.71
       8.6 |          2       0.21       95.93
  8.608696 |          1       0.11       96.03
  8.785714 |          1       0.11       96.14
  8.888889 |          1       0.11       96.25
         9 |         35       3.75      100.00
------------+-----------------------------------
     Total |        933     100.00
```

```
tab mv717, nol

 occupation |
  (grouped) |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      1,865       24.72       24.72
          1 |        515        6.83       31.55
          4 |      2,507       33.23       64.78
          6 |        978       12.96       77.74
          8 |        527        6.99       84.73
          9 |      1,152       15.27      100.00
------------+-----------------------------------
      Total |      7,544      100.00
```

Subject: Re: Aggregating household level data
Posted by Reduced-For(u)m on Mon, 25 Jul 2016 21:59:31 GMT
View Forum Message <> Reply to Message

It looks to me like you are telling Stata to treat your categorical variable as a continuous variable. If you want to keep each individual occupation, you'll need to create a separate dummy variable for each occupation, and then when you collapse those you will have a proportion of the population with that occupation in that cluster. If there are too many different groups, you'll probably need to combine them in some way.

But right now I think you are just getting a meaningless number because you are treating the variable value of, say, "4" as the number 4 and not as "occupation group 4." So those numbers below are meaningless.

Subject: Re: Aggregating household level data
Posted by Khanjila on Fri, 12 Aug 2016 06:59:04 GMT
View Forum Message <> Reply to Message

Hello,

I am also interested in doing the exact same thing: collapse the women's grouped occupation (v717) to the cluster level. I want to get the proportion of women who are in either of those occupation categories at the cluster level.

How do I go about it, which commands can I use in stata?

Subject: Re: Aggregating household level data
Posted by Bridgette-DHS on Mon, 15 Aug 2016 12:33:17 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

The following Stata lines will calculate the proportions of men in each cluster who are in each category of mv717. The crucial lines are the one that begins with "xi" and the one that begins with "collapse".

```
use e:\DHS\DHS_data\MR_files\KEMR70FL.dta, clear
keep mv001 mv005 mv717
tab mv717,m

* Give the NA cases a code of 99
replace mv717=99 if mv717==.

xi i.mv717, noomit
gen n=1
collapse (mean) _I* (sum) n (first) mv005, by(mv001)
```

```
rename _I* *
save e:\DHS\DHS_data\scratch\KEMRtemp.dta, replace

* This cluster-level file has the number of cases, weight, and the proportions
*  of men in each category of mv717.
* You will need to attach variable names to the mv717* variables

* Proceed to the following steps if you want to merge back with the MR file
drop mv005
sort mv001

use e:\DHS\DHS_data\MR_files\KEMR70FL.dta, clear

* I would reduce to just the variables you actually need in the MR file
* example:
keep mv0*

sort mv001 mv002 mv003
merge mv001 using e:\DHS\DHS_data\scratch\KEMRtemp.dta
drop _merge

* save this file with another name
```

Subject: Re: Aggregating household level data
Posted by Hassen on Wed, 02 May 2018 19:59:20 GMT
View Forum Message <> Reply to Message

Dear Bridgette,One of the independent variable in my MPH thesis,Multiple child deprivation index (children having low birth weght,not breast feed,etc) at the community level by Aggregating the individual variables to create the community index. Thus How can I create it using SPSS? Thank you for assistance and advice!!

Subject: Re: Aggregating household level data
Posted by Bridgette-DHS on Thu, 03 May 2018 12:59:52 GMT
View Forum Message <> Reply to Message

Following is another response from Senior DHS Stata Specialist, Tom Pullum:

The analysis team at DHS does not use SPSS, but we believe the relevant commands are aggregate and compute. (In Stata you could use "collapse" or some of the "egen" commands.) You might just try a google search, or perhaps another user can help

---