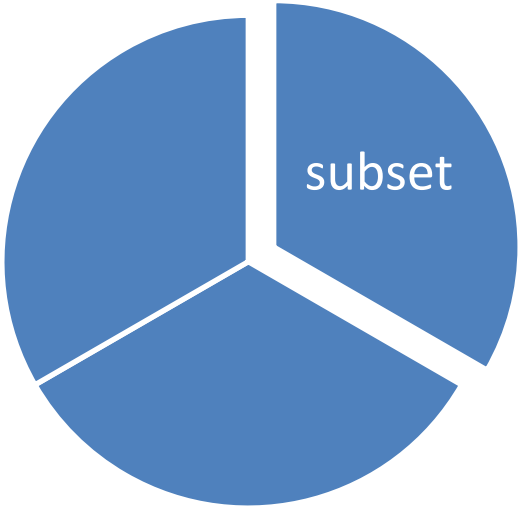
Subsets in SPSS



Subsetting options

- Select by
  - Variables
  - Observations
    - Conditional
    - Random
    - Range
- Permanent versus temporary/filtered

Selecting variables

- Delete variables
- Get file

  /keep=

  /drop=
- Save outfile

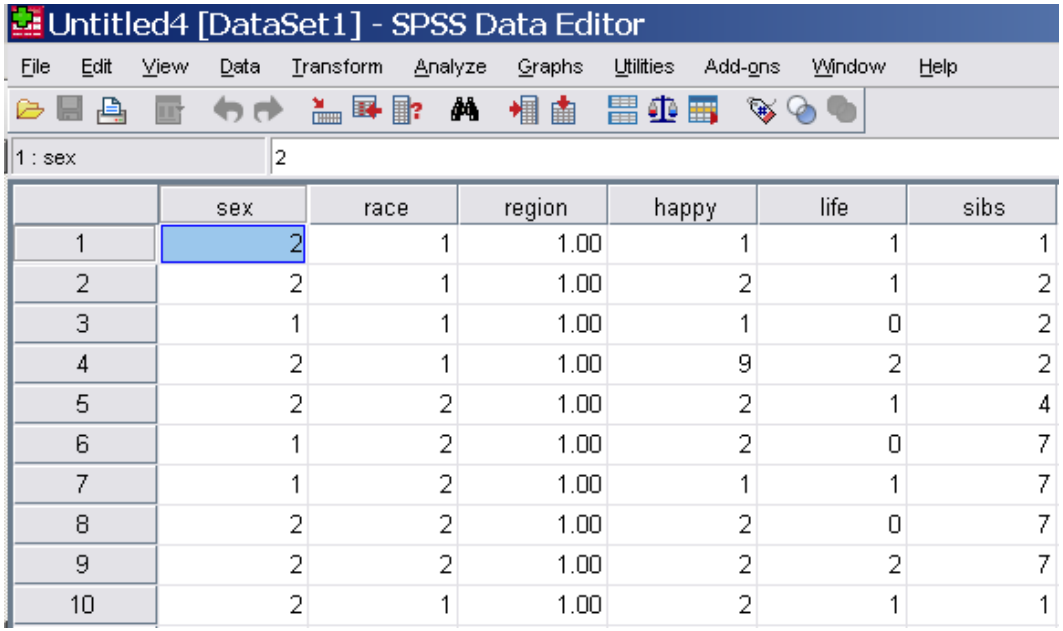
  /keep=

  /drop=

Delete variables

get file='Y:\SPSS\Data\1991 U.S. General Social Survey.sav'.

delete variables race sex.



	sex	race	region	happy	life	sibs
1	2	1	1.00	1	1	1
2	2	1	1.00	2	1	2
3	1	1	1.00	1	0	2
4	2	1	1.00	9	2	2
5	2	2	1.00	2	1	4
6	1	2	1.00	2	0	7
7	1	2	1.00	1	1	7
8	2	2	1.00	2	0	7
9	2	2	1.00	2	2	7
10	2	1	1.00	2	1	1

Delete variables illustrated

Untitled4 [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : sex 2

	sex	race	region	happy	life	sibs
1	2	1	1.00	1	1	1
2	2	1	1.00	2	1	2
3	1	1	1.00	1	0	2
4	2	1	1.00	9	2	2
5	2	2	1.00	2	1	4
6	1	2	1.00			
7	1	2	1.00			
8	2	2	1.00			
9	2	2	1.00			
10	2	1	1.00			

*1991 U.S. General Social Survey.sav [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : region 1

	region	happy	life	sibs	childs	age
1	1.00	1	1	1	2	61
2	1.00	2	1	2	1	32
3	1.00	1	0	2	1	35
4	1.00	9	2	2	0	26
5	1.00	2	1	4	0	25
6	1.00	2	0	7	5	59
7	1.00	1	1	7	3	46
8	1.00	2	0	7	4	99
9	1.00	2	2	7	3	57
10	1.00	2	1	1	2	64

Delete variables nuances

- Ignored if transformations are pending

```
get file='Y:\SPSS\Data\1991 U.S. General Social Survey.sav'.
/*error, DELETE cannot be processed while transformations are pending */
compute black=(race=2).
delete variables race.
frequencies variables=black.

/* This works because EXECUTE clears pending transformations */
compute black=(race=2).
execute.
delete variables race sex.
frequencies variables=black.
```

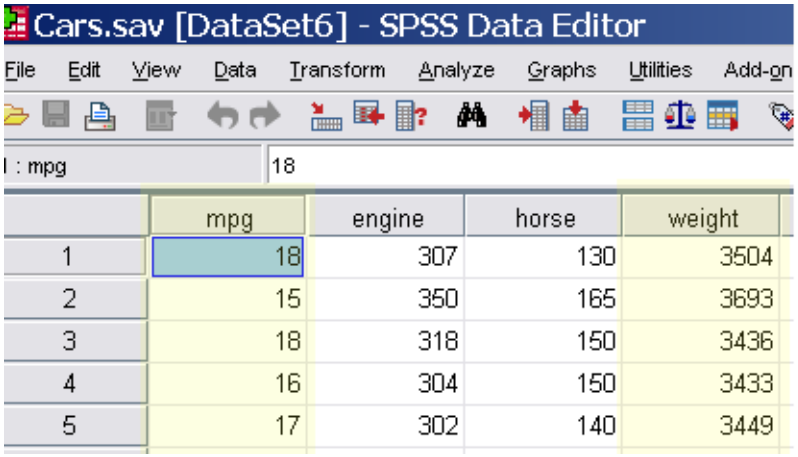
continued

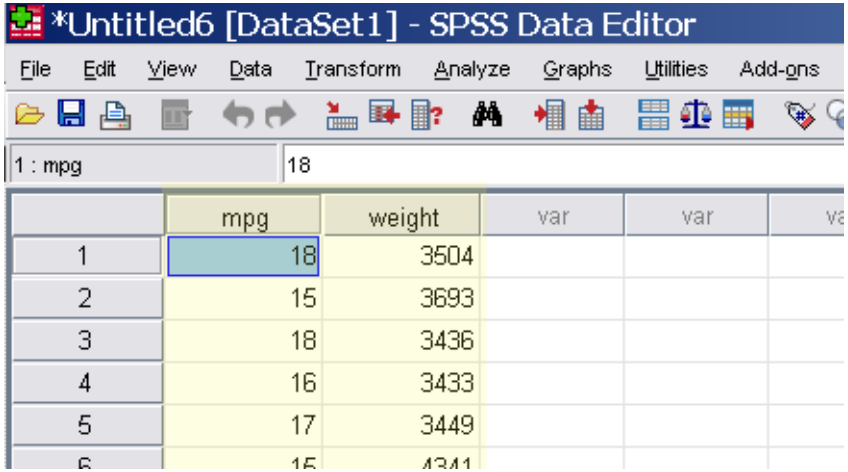
```
/* This works because FREQUENCIES clears pending transformations */
get file='Y:\SPSS\Data\1991 U.S. General Social Survey.sav'.
compute black=(race=2).
frequencies variables=black.
delete variables race.

/* This works because DATASET ACTIVATE clears pending transformations */
get file='Y:\SPSS\Data\1991 U.S. General Social Survey.sav'.
compute black=(race=2).
dataset activate DataSet1.
delete variables race sex.
frequencies variables=black.
```

Get file

- /keep=

```
get file='Y:\SPSS\Data\Cars.sav'
  /keep=mpg weight.
```

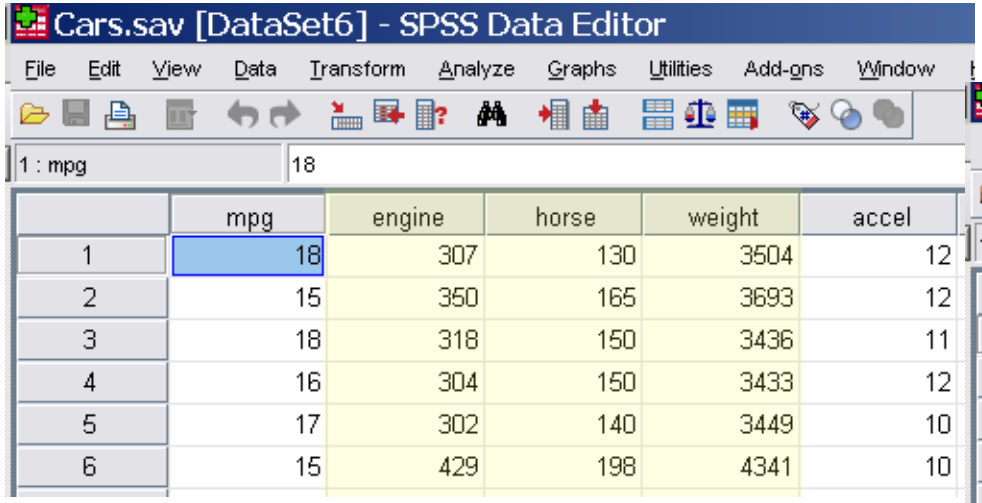




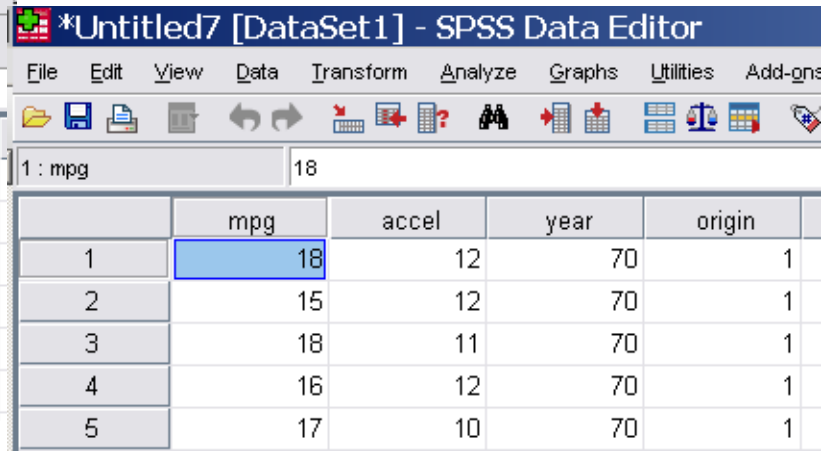
Get file

- /drop=

```
get file='Y:\SPSS\Data\Cars.sav'
  /drop=engine horse weight.
```



	mpg	engine	horse	weight	accel
1	18	307	130	3504	12
2	15	350	165	3693	12
3	18	318	150	3436	11
4	16	304	150	3433	12
5	17	302	140	3449	10
6	15	429	198	4341	10



	mpg	accel	year	origin
1	18	12	70	1
2	15	12	70	1
3	18	11	70	1
4	16	12	70	1
5	17	10	70	1

Save file

- Similar to get file

```
get file='Y:\SPSS\Data\Employee data.sav'.

save outfile="u:\employees1.sav"
  /keep=gender salary salbegin.

/* note the working data set is not changed */
save outfile="u:\employees2.sav"
  /drop=bdate salary jobcat minority.
```

Selecting observations

- Conditional on a data value
- Random
- Range of observations

- Temporary vs. permanent

Temporary, filtered, permanent

- “Permanent” – removes data from the working data set
- “Temporary” – skips data for ***one*** pass through the data set
- “Filtered” – skips data until filtering is turned off

Conditional selection

- Select if

```
get file='Y:\SPSS\Data\1991 U.S. General Social Survey.sav'.
descriptives variables=age.

select if region=1.
descriptives variables=age.
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age of Respondent	1514	18	89	45.63	17.808
Valid N (listwise)	1514				

Descriptives

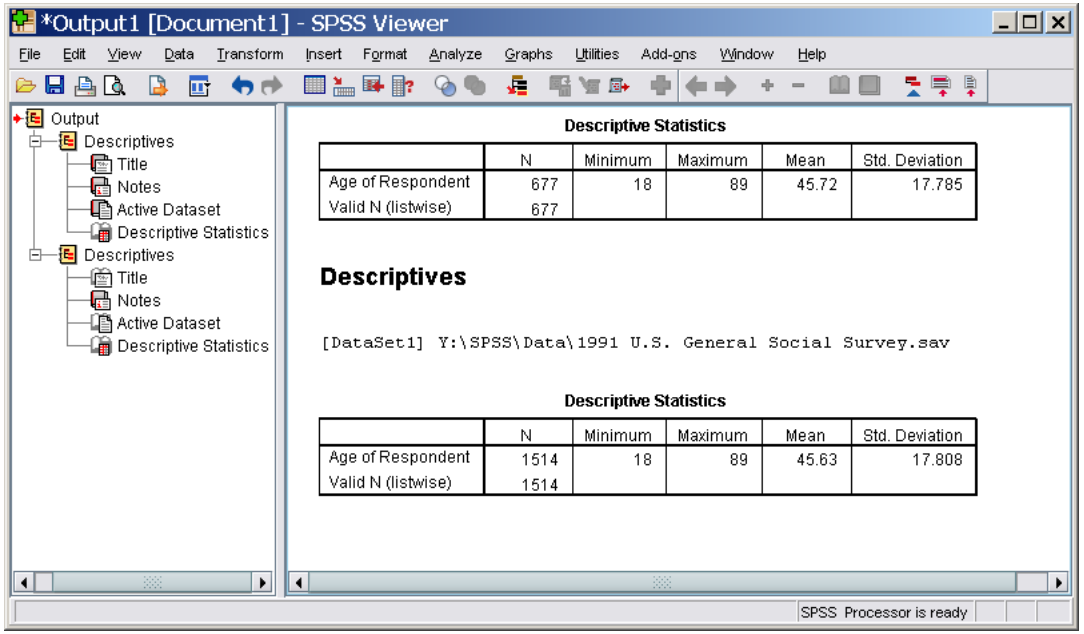
[DataSet1]

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age of Respondent	677	18	89	45.72	17.785
Valid N (listwise)	677				

Temporary select if

```
get file='Y:\SPSS\Data\1991 U.S. General Social Survey.sav'.

temporary.
select if region=1.
descriptives variables=age.
descriptives variables=age.
```

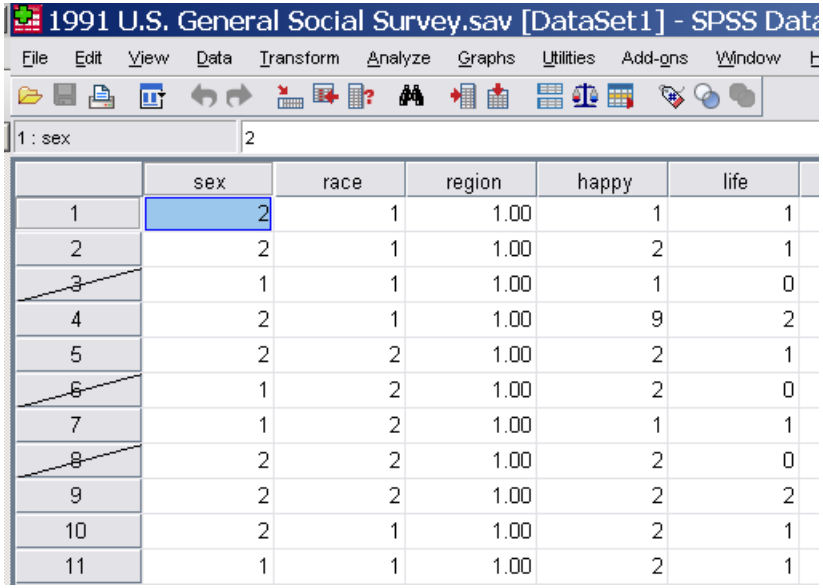


continued

```
temporary.
select if sex = 1.
save outfile="U:\men.sav".
temporary.
select if sex = 2.
save outfile="U:\women.sav".
```

Filtered selection

- Filter by
  - Usually used with a dummy variable
  - Filters cases coded 0 or missing (system & user)
  - Visible in the data editor



	sex	race	region	happy	life
1	2	1	1.00	1	1
2	2	1	1.00	2	1
3	1	1	1.00	1	0
4	2	1	1.00	9	2
5	2	2	1.00	2	1
6	1	2	1.00	2	0
7	1	2	1.00	1	1
8	2	2	1.00	2	0
9	2	2	1.00	2	2
10	2	1	1.00	2	1
11	1	1	1.00	2	1

Random Samples

- Simple random sampling
  - Proportion (approximate)
  - Number of cases (exact)

Ranges

- Sampling by observation range