-------------------------------------------------------------------------

**BAB 1 : STATA DATA MANAGEMENT**

-------------------------------------------------------------------------

Pada data management ini, kita akan melakukan beberapa hal, tapi fokusnya adalah melalui command bukan menggunakan toolbar.

1.1 EXPORTING DATA

--------------------------------------------------------------

Di stata, kita hanya bisa meload satu dataset dalam satu waktu.

Untuk itu, sebelum dataset lain di load, data yang tersimpan di memori harus dihapus dengan menggunakan perintah clear.

clear

1.1.1 Exporting Data dari Sumber Berbentuk CSV

--------------------------------------------------------------

Untuk membaca csv, saat perintah dijalankan, dia akan membaca baris pertama untuk mengidentifikasi separatornya.

Namun, separator sendiri bisa didefinisikan secara terpisah.

import delimited using hs0.csv, clear

describe

1.1.2 Exporting Data dari Sumber Berbentuk Excel

--------------------------------------------------------------

Untuk mengimport data dari excel, kita bisa menspesifikasi sheet mana yang akan digunakan dan menggunakan baris pertama sebagai nama tabel dengan -firstrow clear -.

import excel using hsbdemo.xlsx, sheet("hsbdemo") firstrow clear

1.1.3 Menyimpan Data ke Disk

--------------------------------------------------------------

Untuk menyimpan data ke disk, bisa dilakukan langsung dengan perintah save.

save dataku

Setelah itu, file akan tersimpan ke dalam lokal komputer dengan ekstensi .dta.

1.2 RETRIEVING DATA

--------------------------------------------------------------

1.2.1 Menampilkan isi dataset

--------------------------------------------------------------

Misalkan kita mengimpor "Chapter 2 Data" sebagai dataset:

Untuk menampilkan data dari dataset yang sudah di impor (use), bisa menggunakan -describe-

. use "Chapter 2 Data"

. describe

HASILNYA:

---------

Contains data from census2c.dta

obs: 21 1980 Census data for NE and NC states

vars: 7 9 Oct 2015 12:43

-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

storage display value

variable name type format label variable label

-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

state str13 %-13s State

region byte %-8.0g cenreg Census region

pop double %8.1f 1980 Population, '000

popurb double %8.1f 1980 Urban population, '000

medage float %9.2f Median age, years

marr double %8.1f Marriages, '000

divr double %8.1f Divorces, '000

-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

1.2.2 Menampilkan Isi Salah Satu Tabel di Dataset

--------------------------------------------------------------

Gunakan perintah -list- untuk menampilkan semua kolom dan baris dari satu tabel

. list

HASILNYA :

-----------

+---------------------------------------------------------------------+

| state region pop popurb medage marr divr |

|---------------------------------------------------------------------|

1. | Connecticut NE 3107.6 2449.8 32.00 26.0 13.5 |

2. | Illinois N Cntrl 11426.5 9518.0 29.90 109.8 51.0 |

3. | Indiana N Cntrl 5490.2 3525.3 29.20 57.9 40.0 |

4. | Iowa N Cntrl 2913.8 1708.2 30.00 27.5 11.9 |

5. | Kansas N Cntrl 2363.7 1575.9 30.10 24.8 13.4 |

6. | Maine NE 1124.7 534.1 30.40 12.0 6.2 |

7. | Massachusetts NE 5737.0 4808.3 31.20 46.3 17.9 |

8. | Michigan N Cntrl 9262.1 6551.6 28.80 86.9 45.0 |

9. | Minnesota N Cntrl 4076.0 2725.2 29.20 37.6 15.4 |

10. | Missouri N Cntrl 4916.7 3349.6 30.90 54.6 27.6 |

11. | Nebraska N Cntrl 1569.8 987.9 29.70 14.2 6.4 |

12. | New Hampshire NE 920.6 480.3 30.10 9.3 5.3 |

13. | New Jersey NE 7364.8 6557.4 32.20 55.8 27.8 |

14. | New York NE 17558.1 14858.1 31.90 144.5 62.0 |

15. | N. Dakota N Cntrl 652.7 318.3 28.30 6.1 2.1 |

16. | Ohio N Cntrl 10797.6 7918.3 29.90 99.8 58.8 |

17. | Pennsylvania NE 11863.9 8220.9 32.10 93.7 34.9 |

18. | Rhode Island NE 947.2 824.0 31.80 7.5 3.6 |

19. | S. Dakota N Cntrl 690.8 320.8 28.90 8.8 2.8 |

20. | Vermont NE 511.5 172.7 29.40 5.2 2.6 |

21. | Wisconsin N Cntrl 4705.8 3020.7 29.40 41.1 17.5 |

+---------------------------------------------------------------------+

Kemudian, apabila hanya ingin mengambil kolom state, pop, dan marr

. list region pop marr

HASILNYA :

-----------

+---------------------------+

| region pop marr |

|---------------------------|

1. | NE 3107.6 26.0 |

2. | N Cntrl 11426.5 109.8 |

3. | N Cntrl 5490.2 57.9 |

4. | N Cntrl 2913.8 27.5 |

5. | N Cntrl 2363.7 24.8 |

|---------------------------|

6. | NE 1124.7 12.0 |

7. | NE 5737.0 46.3 |

8. | N Cntrl 9262.1 86.9 |

9. | N Cntrl 4076.0 37.6 |

10. | N Cntrl 4916.7 54.6 |

|---------------------------|

11. | N Cntrl 1569.8 14.2 |

12. | NE 920.6 9.3 |

13. | NE 7364.8 55.8 |

14. | NE 17558.1 144.5 |

15. | N Cntrl 652.7 6.1 |

|---------------------------|

16. | N Cntrl 10797.6 99.8 |

17. | NE 11863.9 93.7 |

18. | NE 947.2 7.5 |

19. | N Cntrl 690.8 8.8 |

20. | NE 511.5 5.2 |

|---------------------------|

21. | N Cntrl 4705.8 41.1 |

+---------------------------+

Apabila ingin menampilkan urutan berdasarkan region.

. list idcode ccity hours uniondues married marriedyrs nevermarried in 1/3, abb(

> 20)

+-----------------------------------------------------------+

1. | idcode | ccity | hours | uniondues | married | marriedyrs |

| 5159 | 1 | 38 | 29 | 0 | 0 |

|-----------------------------------------------------------|

| nevermarried |

| 0 |

+-----------------------------------------------------------+

+-----------------------------------------------------------+

2. | idcode | ccity | hours | uniondues | married | marriedyrs |

| 5157 | 0 | 35 | 0 | 1 | 0 |

|-----------------------------------------------------------|

| nevermarried |

| 0 |

+-----------------------------------------------------------+

+-----------------------------------------------------------+

3. | idcode | ccity | hours | uniondues | married | marriedyrs |

| 5156 | 0 | 40 | 0 | 1 | 3 |

|-----------------------------------------------------------|

| nevermarried |

| 0 |

+-----------------------------------------------------------+

tanpa separator

. list state region pop marr, sep(0)

HASILNYA :

-----------

+-------------------------------------------+

| state region pop marr |

|-------------------------------------------|

1. | Connecticut NE 3107.6 26.0 |

2. | Illinois N Cntrl 11426.5 109.8 |

3. | Indiana N Cntrl 5490.2 57.9 |

4. | Iowa N Cntrl 2913.8 27.5 |

5. | Kansas N Cntrl 2363.7 24.8 |

6. | Maine NE 1124.7 12.0 |

7. | Massachusetts NE 5737.0 46.3 |

8. | Michigan N Cntrl 9262.1 86.9 |

9. | Minnesota N Cntrl 4076.0 37.6 |

10. | Missouri N Cntrl 4916.7 54.6 |

11. | Nebraska N Cntrl 1569.8 14.2 |

12. | New Hampshire NE 920.6 9.3 |

13. | New Jersey NE 7364.8 55.8 |

14. | New York NE 17558.1 144.5 |

15. | N. Dakota N Cntrl 652.7 6.1 |

16. | Ohio N Cntrl 10797.6 99.8 |

17. | Pennsylvania NE 11863.9 93.7 |

18. | Rhode Island NE 947.2 7.5 |

19. | S. Dakota N Cntrl 690.8 8.8 |

20. | Vermont NE 511.5 5.2 |

21. | Wisconsin N Cntrl 4705.8 41.1 |

+-------------------------------------------+

hanya menampilkan 1/5 data

. list state region pop marr in 1/5

HASILNYA :

-----------

+-----------------------------------------+

| state region pop marr |

|-----------------------------------------|

1. | Connecticut NE 3107.6 26.0 |

2. | Illinois N Cntrl 11426.5 109.8 |

3. | Indiana N Cntrl 5490.2 57.9 |

4. | Iowa N Cntrl 2913.8 27.5 |

5. | Kansas N Cntrl 2363.7 24.8 |

+-----------------------------------------+

dengan kondisional:

. list state region pop marr if pop>3000 in 1/5

HASILNYA :

-----------

+-----------------------------------------+

| state region pop marr |

|-----------------------------------------|

1. | Connecticut NE 3107.6 26.0 |

2. | Illinois N Cntrl 11426.5 109.8 |

3. | Indiana N Cntrl 5490.2 57.9 |

. list if substr(state,1,1)=="M"

HASILNYA :

-----------

+------------------------------------------------------------------+

| state region pop popurb medage marr divr |

|------------------------------------------------------------------|

6. | Maine NE 1124.7 534.1 30.40 12.0 6.2 |

7. | Massachusetts NE 5737.0 4808.3 31.20 46.3 17.9 |

8. | Michigan N Cntrl 9262.1 6551.6 28.80 86.9 45.0 |

9. | Minnesota N Cntrl 4076.0 2725.2 29.20 37.6 15.4 |

10. | Missouri N Cntrl 4916.7 3349.6 30.90 54.6 27.6 |

+------------------------------------------------------------------+

apabila header kepanjangan

. list idcode ccity hours uniondues married marriedyrs nevermarried in 1/3, noob> s

+-------------------------------------------------------------------+

| idcode ccity hours uniond~s married marrie~s neverm~d |

|-------------------------------------------------------------------|

| 5159 1 38 29 0 0 0 |

| 5157 0 35 0 1 0 0 |

| 5156 0 40 0 1 3 0 |

+-------------------------------------------------------------------+

spesifikasi agar tidak kepanjangan

. list idcode ccity hours uniondues married marriedyrs nevermarried in 1/3, abb(> 20) noobs

+--------------------------------------------------------------------------+

| idcode ccity hours uniondues married marriedyrs nevermarried |

|--------------------------------------------------------------------------|

| 5159 1 38 29 0 0 0 |

| 5157 0 35 0 1 0 0 |

| 5156 0 40 0 1 3 0 |

+--------------------------------------------------------------------------+

Summmarize

. summarize pop

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

pop | 21 5142.903 4675.152 511.456 17558.07

. summarize pop\*

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

pop | 21 5142.903 4675.152 511.456 17558.07

popurb | 21 3829.776 3851.458 172.735 14858.07