ARAB INSTITUTE
FOR TRAINING
& RESEARCH IN STATISTICS

المعهد العربي
للتدريب والبحوث الإحصائية

ورشة التحليل الاحصائي باستخدام البرامج الجاهزة

١٨-١٩ اب / اغسطس ٢٠٢١

STATA

رانيا أبودهيم
الاحصاءات العامة الاردنية

اليوم الاول ٢٠٢١/٨/١٨

- مقدمة عامة عن برنامج الستاتا /مميزاته/سلبياته/اصداراته/مكونات الشاشة الرئيسية للبرنامج
- بعض الاوامر المستخدمة في تنظيف البيانات ومعالجتها
- بعض الاوامر المستخدمة في عرض النتائج

اليوم الثاني ٢٠٢١/٨/١٩

- مراجعة سريعة لليوم الاول
- كيفية استيراد وتصدير البيانات
  ماهو ال Do file
- تطبيق عملي شامل للمادة

- هي حزمة برامج إحصائية للأغراض العامة تم إنشاؤها في عام ١٩٨٥ من قبل **StataCorp**.
- يعمل معظم مستخدميها في البحوث، ولا سيما في مجالات الاقتصاد وعلم الاجتماع والعلوم السياسية والطب الحيوي وعلم الأوبئة. ولستاتا قدرات تشمل إدارة البيانات ، والتحليل الإحصائي ، والرسومات ، والمحاكاة ، وتحليل الانحدار .
- **Stata** هي اختصار لكلمتي الاحصاءات والبيانات.

من مميزاتها:

يمكن التعامل مع ملفات البيانات الكبيرة

يمكن اجراء عمليات التحليل اما بكتابة الاوامر او من خلال القوائم

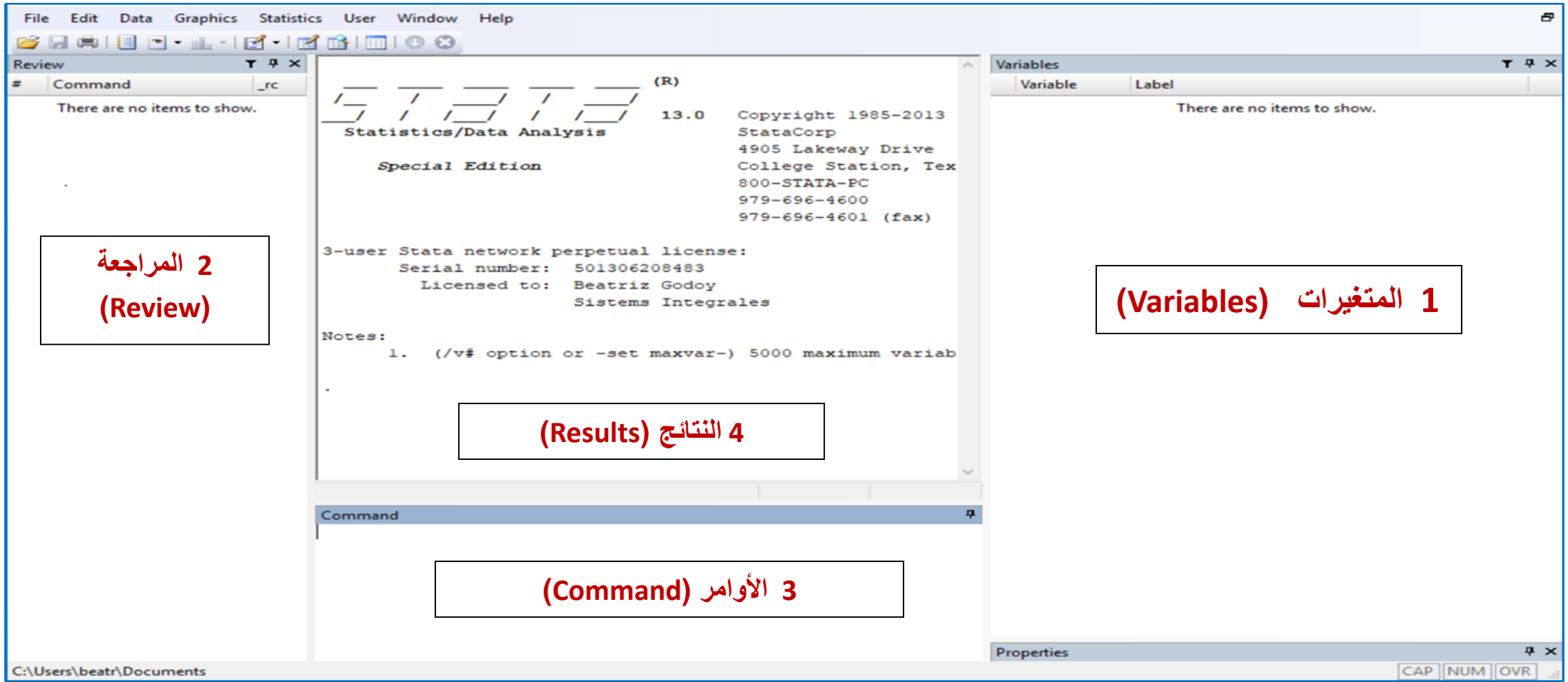
يعتبر برنامج مميز ومتفوق في اجراء تحليل الانحدار المتقدم

من سلبياته:

يعتبر برنامج صعب في التعلم مقارنة مع البرامج الاخرى

	Stata/MP	Stata/SE	Stata/IC
Max.Number of Variables	120,000	32767	2048
Max.Number of Independent Variables	10998	10998	798
Memory Requirements	4GB	2GB	1GB

مكونات الشاشة الرئيسية لبرنامج Stata



فتح ملفات بصيغة Stata
.dta

✓ يتم فتح ملف البيانات :

- من قائمة file نختار open ثم ابحث عن الملف المراد فتحه .
- النقر على اسم الملف مرتين .
- كتابة أمر **use** في شاشة كتابة الأوامر.

```
use   "E:\Stata_Training\hh.dta",   clear
```

تخزين ملفات بصيغة Stata

✓ **يتم تخزين ملف البيانات :**

- يتم تخزين ملف جديد عن طريق الامر :

```
save   "E:\Stata_Training\hh.dta”
```

- يتم تخزين ملف مخزن مسبقا عن طريق الامر :

```
save   "E:\Stata_Training\hh.dta”, replace
```

تثبيت مسار العمل change directory

cd command

- نستطيع من خلال هذا الامر قراءة او فتح الملفات بطريقة مختصرة من خلال تثبيت المكان المراد فتح او تخزين الملفات عليه،بمعنى أخر يستطيع المستخدم فتح الملف من المكان دون كتابة العنوان كاملا في كل مرة وانما يكتفي باسم الملف.

```
cd   "E:\Stata_Training“

use " individual.dta “

save " ind.dta "
```

- لعرض محتويات الملفات داخل المجلد عن طريق الامر **dir**

```
dir
```

التعرف على البيانات وخصائصها

browse command

```
. br hhid governorate region
.
```

Data Editor (Browse) - [RANIA]

File Edit View Data Tools

hhid[1] 101

	hhid	governorate	region
1	101	Amman	الوسط
2	102	Amman	الوسط
3	103	Amman	الوسط
4	105	Amman	الوسط
5	108	Amman	الوسط
6	109	Amman	الوسط
7	110	Amman	الوسط
8	201	Amman	الوسط
9	202	Amman	الوسط
10	203	Amman	الوسط
11	204	Amman	الوسط
12	205	Amman	الوسط
13	206	Amman	الوسط
14	207	Amman	الوسط
15	208	Amman	الوسط
16	301	Amman	الوسط
17	302	Amman	الوسط
18	303	Amman	الوسط
19	304	Amman	الوسط

التعرف على البيانات وخصائصها

browse command

```
. br hhid governorate region,nol
.
```

Data Editor (Browse) - [RANIA]

File Edit View Data Tools

hhid[1] 101

	hhid	governorate	region
1	101	11	1
2	102	11	1
3	103	11	1
4	105	11	1
5	108	11	1
6	109	11	1
7	110	11	1
8	201	11	1
9	202	11	1
10	203	11	1
11	204	11	1
12	205	11	1
13	206	11	1
14	207	11	1
15	208	11	1
16	301	11	1
17	302	11	1
18	303	11	1
19	304	11	1

التعرف على البيانات وخصائصها

describe command

```
. des governorat q506 sex age

              storage   display    value
variable name   type    format     label      variable label
------------------------------------------------------------------------------
governorate     byte    %8.0g      lq101      Governorate
q506            byte    %22.0g     lq506      Nationality
sex             byte    %10.0g     lq503      Gender
age             int     %10.0g                Age in years

.
```

التعرف على البيانات وخصائصها

sum command

```
. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
hhid	19,216	124457.1	71451.43	101	250008
governorate	19,216	19.13999	8.172717	11	34
region	19,216	1.708004	.7878405	1	3
district	19,216	2.265872	1.739302	1	9
subdistrict	19,216	1.2028	.6361759	1	5
urban	19,216	1.150812	.3578747	1	2
cluster	19,216	1244.52	714.5134	1	2500
hhn	19,216	5.076082	2.745296	1	12
q114	19,216	2017.59	.4918217	2017	2018
period	19,216	18.65794	10.35116	1	36
no_quarter	19,216	2.516393	1.114166	1	4
q108	19,216	19.59097	22.77178	1	161
hh_size	19,216	4.966746	2.250229	1	29
cat_hhsize	19,216	2.728924	1.142215	1	8
male	19,216	2.49771	1.466814	0	18
female	19,216	2.469036	1.45183	0	13
no_child	19,216	2.072856	1.886777	0	21
cat_no_child	19,216	2.226998	1.029865	1	6

التعرف على البيانات وخصائصها

codebook command

```
. codebook sex q522

-------------------------------------------------------------------------------
sex                                                                    Gender
-------------------------------------------------------------------------------

                  type:  numeric (byte)
                 label:  lq503

                 range:  [1,2]                        units:  1
        unique values:  2                         missing .:  0/19,216

            tabulation:  Freq.   Numeric  Label
                        16,606         1  male
                         2,610         2  female

-------------------------------------------------------------------------------
q522                                                   Highest level successfully completed
-------------------------------------------------------------------------------

                  type:  numeric (byte)
                 label:  lq522

                 range:  [1,12]                       units:  1
        unique values:  12                        missing .:  106/19,216

              examples:  3      Elementary
                         4      Preparatory
                         5      Basic
```

STATA®

تحضير البيانات

keep command

```
. keep hhid gov sex

. save file2.dta,replace
file file2.dta saved

. br,nol

.
```

Data Editor (Browse) - [file2]

File Edit View Data Tools

hhid[1] 101

	hhid	governorate	sex
1	101	11	1
2	102	11	1
3	103	11	1
4	105	11	1
5	108	11	1
6	109	11	1
7	110	11	1
8	201	11	1
9	202	11	2
10	203	11	1
11	204	11	2
12	205	11	1
13	206	11	1
14	207	11	1
15	208	11	1
16	301	11	1
17	302	11	1
18	303	11	1
19	304	11	1

تحضير البيانات

keep command

```
. keep if governorate==11
(13,893 observations deleted)

. save file3.dta,replace
(note: file file3.dta not found)
file file3.dta saved

. br,nol

. br,nol

.
```

Data Editor (Browse) - [file3]

File Edit View Data Tools

hhid[1] 101

	hhid	governorate	region	district	subdistrict	urban	cluster	hhn
1	101	11	1	1	1	1	1	
2	102	11	1	1	1	1	1	
3	103	11	1	1	1	1	1	
4	105	11	1	1	1	1	1	
5	108	11	1	1	1	1	1	
6	109	11	1	1	1	1	1	
7	110	11	1	1	1	1	1	
8	201	11	1	1	1	1	2	
9	202	11	1	1	1	1	2	
10	203	11	1	1	1	1	2	
11	204	11	1	1	1	1	2	
12	205	11	1	1	1	1	2	
13	206	11	1	1	1	1	2	
14	207	11	1	1	1	1	2	
15	208	11	1	1	1	1	2	
16	301	11	1	1	1	1	3	
17	302	11	1	1	1	1	3	
18	303	11	1	1	1	1	3	
19	304	11	1	1	1	1	3	

keep command

```
. keep in 1/5
(19,211 observations deleted)

. save file4.dta,replace
file file4.dta saved

. br,nol

.
```

Data Editor (Browse) - [file4]

File Edit View Data Tools

hhid[1] 101

	hhid	governorate	region	district	subdistrict	urban	cluster	hhn
1	101	11	1	1	1	1	1	1
2	102	11	1	1	1	1	1	2
3	103	11	1	1	1	1	1	3
4	105	11	1	1	1	1	1	5
5	108	11	1	1	1	1	1	8

generate command

✓ يستخدم هذا الامر لانشاء متغير جديد ويمكن ان يكون انشاءه معتمد على متغير اخر في ملف البيانات.

```
gen x =1

gen r =1 if gov==11

gen var1 = q201 if gov==11
```

replace command

- ✓ هذا الأمرهو لاستبدال قيم البيانات سواء كانت متغيرات جديدة أو متغيرات موجودة من قبل وهو مرتبط إلى حدٍ كبير مع الأمر **generate**.

```
gen x=.

replace x =1  if idcode<3

replace x =2  if  idcode>3 & idcode<8

replace x =3 if idcode>8
```

egen command

✓ يستخدم هذا الأمر لانشاء متغير جديد يعتمد على احصاء معين لمتغير آخر موجود.

```
egen mean_month1= mean (month1)
egen max_month= max (month1)
egen min_month= min(month1)
egen min_wage= rowmin (month1 month2 month3)
egen total_wage= rowtotal (month1 month2 month3)
egen mean_wage= rowmean (month1 month2 month3)
```

✓ عند استخدام امر egen مع rowtotal وكان row جميع قيمه تساوي missing عندها تكون اجابة صفر،في حال اردنا ان تبقى نتيجة missing نضع مع الامر m,

```
egen total_wage= rowtotal (month1 month2 month3),m
```

تحضير البيانات

recode command

- ✓ يعمل هذا الامرعلى انشاء متغير جديد او نفس المتغير يحتوي على مجموعات أو فئات بالاعتماد على نفس المتغير اوعلى متغير اخر.

```
recode sex (1=0) (2=1)
recode sex (1=0) (2=1), gen gender
recode educ_stat (1 2 = 1) (3 = 2) (4/9= 3), gen(cat_edu)
recode educ_stat (1 2 = 1 "Illiterate") (3/6 = 2 "diploma and less") (7/9 = 3 "Bachelor and above"), gen(cat_edu)
```

collapse command

- هذا الأمر شبيه تماماً بالأمر aggregate في برنامج spss وهو عملية تجميع الملف الواحد على مستوى مجمّع

```
collapse (sum)  month3, by(group)
collapse (count) idcode, by(hhid)
```

append command

- يستخدم هذا الامر لدمج الحالات (cases) من ملفات مختلفة داخل ملف واحد.

north.dta

gov	Tot of male
21	40000
22	15000
23	8000
24	3000

Middle.dta

gov	Tot of male
11	50000
12	20000
13	10000
14	5000

يتم اولا فتح احد الملفات ودمجه على الملف الاخر المغلق.

ولنفرض ان ملف ال middle هو المفتوح

append using north

Gov	Tot of male
11	50000
12	20000
13	10000
14	5000
21	40000
22	15000
23	8000
24	3000

merge command

- **يستخدم هذا الامر لدمج الملفات على مستويات معينة وينقسم الى الانواع التالية:**

- Merge 1:1 →merge one to one

وهو دمج ملف على مستوى واحد مع ملف اخر من نفس المستوى اي ملف اسر مع ملف اسر او ملف افراد مع ملف افراد.

- Merge 1:m→ merge one to many

- **Merge m:1→ merge many to one**

merge command

- Merge 1:1 --→merge one to one

Data Editor (Browse) - [ind1.dta]

File Edit View Data Tools

hhid[1] 101

Snapshots

	hhid	n_head
1	101	1
2	102	1
3	103	1
4	104	1
5	105	1
6	108	1
7	109	1
8	110	1
9	201	1
10	202	1
11	203	1
12	204	1
13	205	1
14	206	1
15	207	1
16	208	1
17	301	1
18	302	1
19	303	1
20	304	1
21	305	1
22	306	1
23	307	1
24	308	1
25	501	1
26	502	1

Data Editor (Browse) - [ind2.dta]

File Edit View Data Tools

hhid[1] 101

Snapshots

	hhid	n_4y
1	101	6
2	102	1
3	103	1
4	104	5
5	105	7
6	108	2
7	109	5
8	110	4
9	201	4
10	202	3
11	203	3
12	204	1
13	205	6
14	206	4
15	207	7
16	208	3
17	301	9
18	302	6
19	303	5
20	304	4
21	305	6
22	306	8
23	307	4
24	308	4
25	501	1
26	502	4

Data Editor (Browse) - [ind1.dta]

File Edit View Data Tools

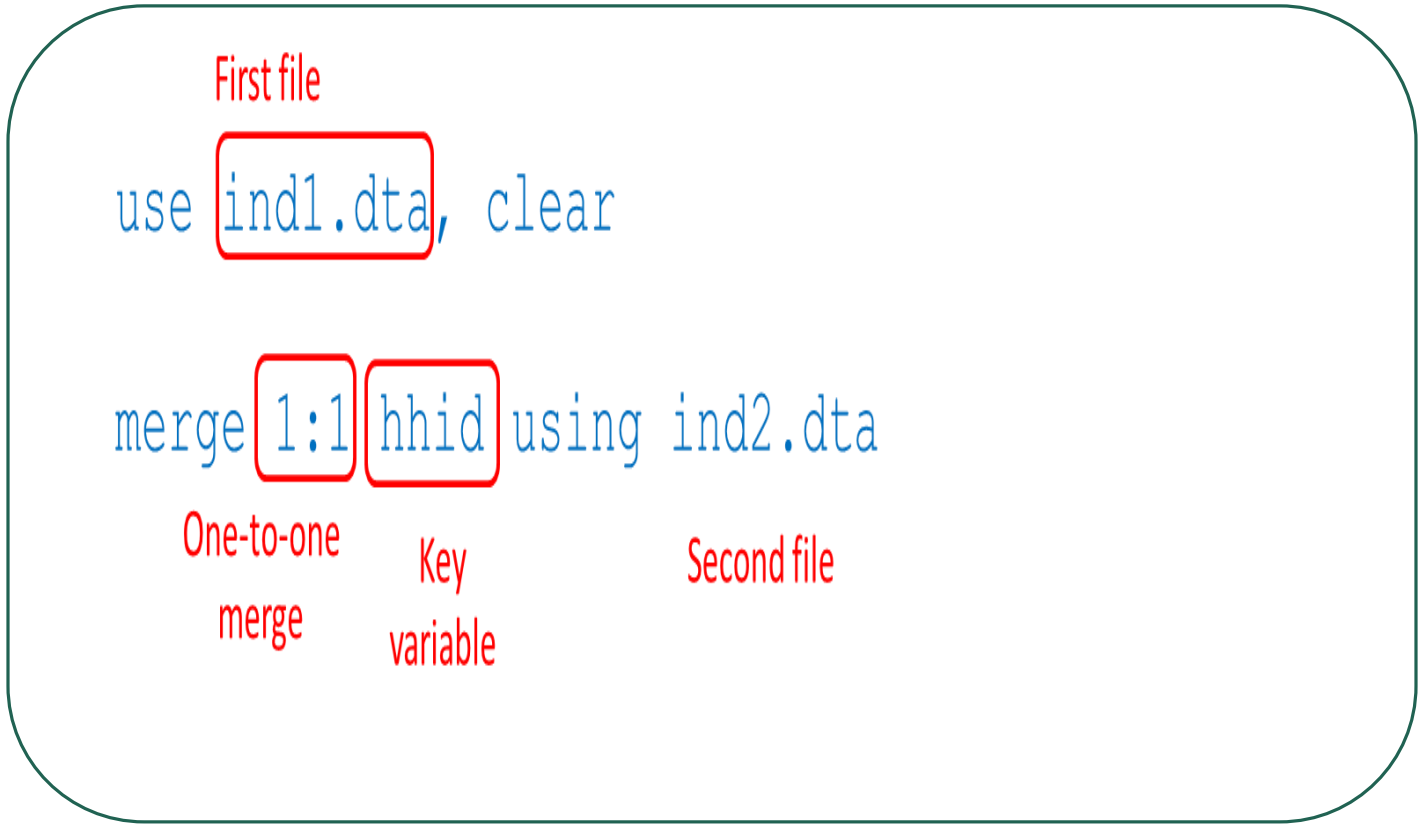
hhid[1] 101

Snapshots

	hhid	n_head	n_4y
1	101	1	6
2	102	1	1
3	103	1	1
4	104	1	5
5	105	1	7
6	108	1	2
7	109	1	5
8	110	1	4
9	201	1	4
10	202	1	3
11	203	1	3
12	204	1	1
13	205	1	6
14	206	1	4
15	207	1	7
16	208	1	3
17	301	1	9
18	302	1	6
19	303	1	5
20	304	1	4
21	305	1	6
22	306	1	8
23	307	1	4
24	308	1	4
25	501	1	1

merge command

- Merge 1:1 --→merge one to one

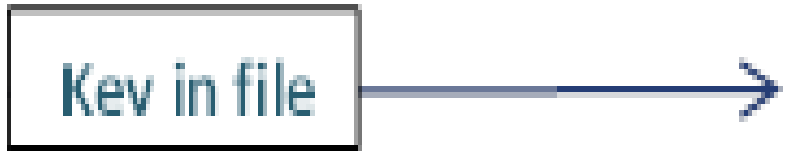


merge command

- Merge 1:m --→merge one to many

Name of file: hh

Type of file:one



	hhid	governorate	q201	q204	q215	q308	q309	q310
1	9306	11	3	1	1	2	1	2
2	19105	11	3	3	1	1	1	2
3	21310	11	3	1	1	2	1	2
4	22108	11	3	1	1	2	1	2
5	22303	11	3	3	1	1	1	2
6	23301	11	3	1	1	2	1	2
7	31502	11	3	1	1	1	1	1
8	34205	11	3	1	1	1	1	1
9	42510	11	3	1	1	2	1	2
10	46904	11	3	3	1	2	1	2
11	53001	11	3	1	1	1	1	2
12	68301	11	3	1	1	1	1	2
13	68502	11	2	1	1	2	1	2
14	73509	12	3	1	1	2	1	2
15	82304	12	2	1	1	2	1	2
16	84602	12	3	1	1	2	1	2
17	88706	13	2	2	1	1	1	2

merge command

- **Merge 1:m --→merge one to many**

Name of file: indv_m

Type of file:many

Key in file

	hhid	cluster	hhn	urban	period	hh_size	idcode	q502	sex	q504_m	q504_y	age
1	9306	93	6	1	5	3	3	Son\ Daughter	male	12	1993	23
2	9306	93	6	1	5	3	1	Head	male	9	1945	72
3	9306	93	6	1	5	3	2	Husband\ Wife	female	1	1948	69
4	19105	191	5	1	20	4	3	Son\ Daughter	male	9	2013	4
5	19105	191	5	1	20	4	2	Husband\ Wife	female	12	1991	26
6	19105	191	5	1	20	4	1	Head	male	8	1981	36
7	19105	191	5	1	20	4	4	Son\ Daughter	male	1	2015	3
8	21310	213	10	1	6	1	1	Head	female	1	1951	66
9	22108	221	8	1	6	4	3	Son\ Daughter	male	12	1976	40
10	22108	221	8	1	6	4	1	Head	male	1	1934	83
11	22108	221	8	1	6	4	4	Son\ Daughter	female	7	1981	36
12	22108	221	8	1	6	4	2	Husband\ Wife	female	1	1943	74
13	22303	223	3	1	26	8	6	Son\ Daughter	female	8	2005	12
14	22303	223	3	1	26	8	2	Husband\ Wife	female	6	1981	36
15	22303	223	3	1	26	8	5	Son\ Daughter	female	7	2003	14
16	22303	223	3	1	26	8	1	Head	male	8	1973	44
17	22303	223	3	1	26	8	4	Son\ Daughter	female	7	2001	16
18	22303	223	3	1	26	8	8	Son\ Daughter	male	3	2011	7
19	22303	223	3	1	26	8	3	Son\ Daughter	female	8	2000	17
20	22303	223	3	1	26	8	7	Son\ Daughter	female	10	2007	10

merge command

- Merge 1:m --→merge one to many

	hhid	governorate	q201	q204	q215	q308	q309	q310	q311	q312	q322	q329	cluster
1	9306	11	3	1	1	2	1	2	1	1	1	1	93
2	9306	11	3	1	1	2	1	2	1	1	1	1	93
3	9306	11	3	1	1	2	1	2	1	1	1	1	93
4	19105	11	3	3	1	1	1	2	1	2	2	1	191
5	19105	11	3	3	1	1	1	2	1	2	2	1	191
6	19105	11	3	3	1	1	1	2	1	2	2	1	191
7	19105	11	3	3	1	1	1	2	1	2	2	1	191
8	21310	11	3	1	1	2	1	2	1	1	1	2	213
9	22108	11	3	1	1	2	1	2	1	2	1	1	221
10	22108	11	3	1	1	2	1	2	1	2	1	1	221
11	22108	11	3	1	1	2	1	2	1	2	1	1	221
12	22108	11	3	1	1	2	1	2	1	2	1	1	221

merge command

- Merge 1:m --→merge one to many

```
use hh.dta
merge 1:m hhid using indv_m
save indv_hh.dta,replace
```

reshape command

✓ يستخدم هذا الامر لتحويل البيانات من افقي الى عمودي او العكس وهو نوعان:

1) wide reshape: from long to wide

```
reshape wide {list of variables to reshape} , i(i) j(j)
```

Id variable(s)

variable(s)

1) long reshape: from wide to long

```
reshape long {variable without id of item} , i(i) j(j)
```

Id variable

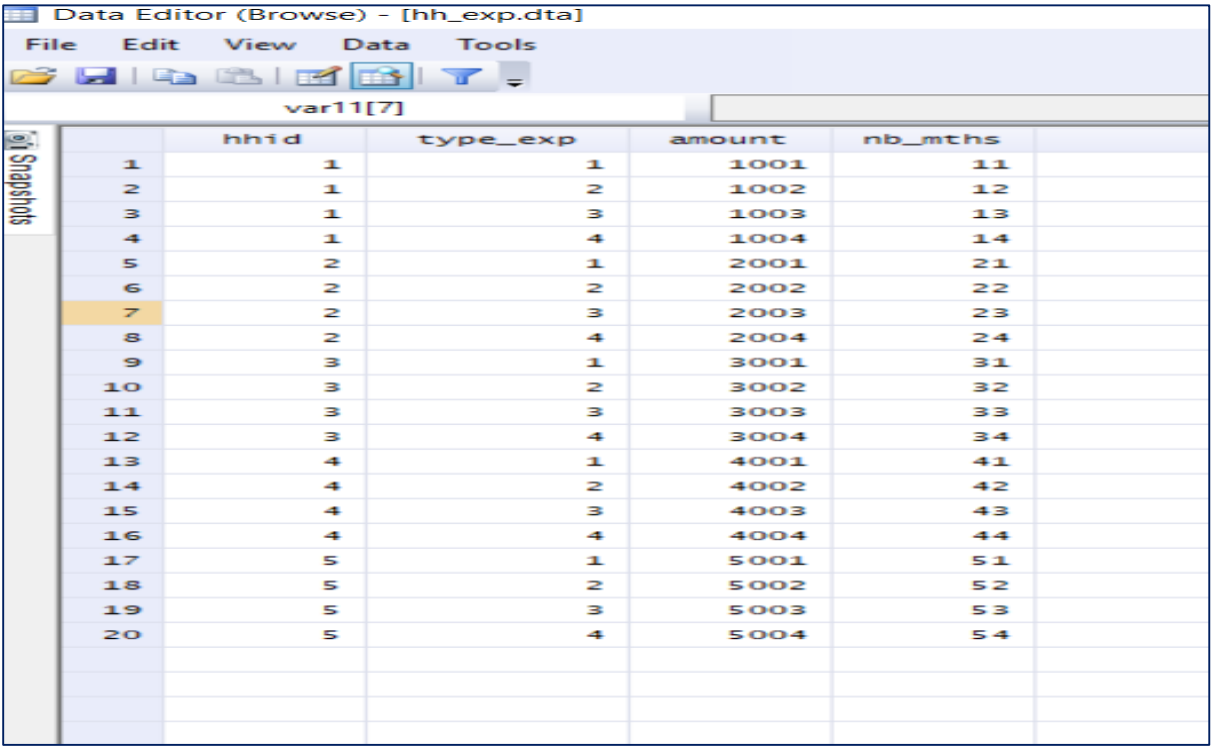
new variable

Wide reshape command

(from long to wide)

```
reshape wide {list of variables to reshape} , i(i) j(j)
```

long



	hhid	type_exp	amount	nb_mths
1	1	1	1001	11
2	1	2	1002	12
3	1	3	1003	13
4	1	4	1004	14
5	2	1	2001	21
6	2	2	2002	22
7	2	3	2003	23
8	2	4	2004	24
9	3	1	3001	31
10	3	2	3002	32
11	3	3	3003	33
12	3	4	3004	34
13	4	1	4001	41
14	4	2	4002	42
15	4	3	4003	43
16	4	4	4004	44
17	5	1	5001	51
18	5	2	5002	52
19	5	3	5003	53
20	5	4	5004	54

wide

hhid	amount1	nb_mths1	amount2	nb_mths2	amount3	nb_mths3	amount4	nb_mths4
1	1001	11	1002	12	1003	13	1004	14
2	2001	21	2002	22	2003	23	2004	24
3	3001	31	3002	32	3003	33	3004	34
4	4001	41	4002	42	4003	43	4004	44
5	5001	51	5002	52	5003	53	5004	54

{list of variables}? **amount** and **nb_mths**

{id variable}? **hhid**

{var}? **Type_exp**

```
reshape wide amount nb_mths, i( hhid ) j( type_exp )
```

Long reshape command

(from wide to long)

```
reshape long {variable without id of item} , i(i) j(j)
```

Id variable(s)

new variable

wide

Data Editor (Browse) - [cons_food.dta]

File Edit View Data Tools

var7[13]

Snapshots

	hhid	qty_1	qty_2	qty_3	qty_4
1	1	1001	1002	1003	1004
2	2	2001	2002	2003	2004
3	3	3001	3002	3003	3004
4	4	4001	4002	4003	4004
5	5	5001	5002	5003	5004

long

hhid	item	qty
1	1	1001
1	2	1002
1	3	1003
1	4	1004
2	1	2001
2	2	2002
2	3	2003
2	4	2004
3	1	3001
3	2	3002
3	3	3003
3	4	3004
4	1	4001
4	2	4002
4	3	4003
4	4	4004
5	1	5001
5	2	5002
5	3	5003
5	4	5004

{variable used to create the long form} **qty_**

{id variable}? **hhid**

{new_var}? **item**

```
reshape long qty_ , i(hhid) j(item)
```

Tabulation

- **Tab1**:هو أمرمهم لعمل الجداول التكرارية، حيث يظهر لنا عدد التكرارات الخاصة بهذا المتغير وهو شبيه تماماً بالأمر frequency في برنامج spss.

(لعمل جدول تكراري لمتغير واحد او أكثر نستخدم هذا الامر)

```
tab1 q522
tab1 q522, m                لظهور الحالات المفقودة (.)
tab1 q522, nolabel          لظهور خيارات السؤال بدون وصف قيم السؤال
tab1 q522 if gov==11        بوجود شرط معين
tab1 sex, gen(new_sex)
bys sex: tab1 q502
tab1 q522 sex q502 q506
```

Tabulation

- **Tab**:يستخدم لمعرفة التقاطعات بين متغيرين فقط (cross tabulation)

```
tab   q522   sex
tab q522 sex , m                          ظهور الحالات المفقودة
tab q522 q502, nolabel      ظهور خيارات السؤال  بدون تعريف (ارقام)
tab q522 sex  if gov==32                  بوجود شرط معين
tab q522 sex, nofreq                      ظهور النسب فقط بدون عدد الحالات
tab q522 sex, row nofreq                  ظهور النسب افقي
tab q522 sex, col nofreq                  ظهور النسب عمودي
bys sex: tab q502 sex                     عرض التكرارات بشكل تفصيلي مجزأ
```

Tabulation

- ***Tabulate, Summarize*** هو أمر يدمج بين **frequency and summarize** فينتج جدول يحتوي على وصف احصائي وتكرارات المتغيرات المختارة وهو يستخدم لمتغير واحد أو متغيرين فقط من ناحية الجداول التكرارية ويستخدم المتغير الثالث في انتاج وصف احصائي لمتغير مختلف.

```
tab sex q506, sum (incom)

tab sex q513  , sum ( age )  nolabel nofreq
```

Tabulation

✓ **table**:وهو أمر يستخدم لمعرفة التقاطعات بين اكثرمن متغيرين

(cross tabulation)

table q201 q204 q308

table q201 q204 q308, by(urban)

✓ لمعرفة المتوسط او الوسيط .. لمتغير متواصل (Continuous) مع متغيرات مصنفة (categories) يتم ذلك باستخدام الامر المساعد **contents**

table q201, contents(mean hh_size)

table q201 q204, contents(mean hh_size median hh_size)

Tabulation

يمكن استخدام (sub categories)

```
table q201 q204, contents(mean hh_size median hh_size) by(urban)
```

يمكن اضافة المجموع عمودي وافقي

```
table q201 q204, contents(mean hh_size median hh_size) by(urban) row col
```

يمكن وضع ال Format مناسب للنتيجة

```
table q201 q204, contents(mean hh_size median hh_size) by(urban) row col  format(%9.2f)
```

النتائج

Tabulation

- ✓ tabstat:وهو امر يستخدم لعرض ملخص لعدة متغيرات رقمية (Continuous) متواصل في جدول واحد ويستخدم لعرض نتيجة متغيرات (Continuous) حسب متغيرات مصنغة (categories)

tabstat month3 يعطي متوسط القيم مباشرة

tabstat month3, by(group) اضافة **subgroup** اذا اردنا التحليل حسب متغير مصنف

tabstat month1 month2 month3, by(group) stat(mean median)

Tabulation

- لعمل جدول تكراري لمتغير او اكثر سواء بشرط او لا نستخدم الامر ***tab1***

```
tab1 sex  gov q529
```

```
. tab1 sex gov q529

-> tabulation of sex

     Gender |      Freq.     Percent        Cum.
------------+-----------------------------------
       male |     47,996       50.29       50.29
     female |     47,445       49.71      100.00
------------+-----------------------------------
      Total |     95,441      100.00

-> tabulation of gov

        gov |      Freq.     Percent        Cum.
------------+-----------------------------------
         11 |     60,291       63.17       63.17
         12 |        505        0.53       63.70
         13 |     24,983       26.18       89.88
         14 |        478        0.50       90.38
         21 |      1,977        2.07       92.45
         31 |      7,207        7.55      100.00
------------+-----------------------------------
      Total |     95,441      100.00

-> tabulation of q529

       q529 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |     22,257       36.13       36.13
          2 |     39,338       63.87      100.00
------------+-----------------------------------
```

Tabulation

- لعمل جدول متقاطع لمتغيرين نستخدم الامر ***tab***

```
tab q502 sex
```

```
. tab    q502   sex
```

q502	Gender male	female	Total
Head	16,606	2,610	19,216
Husband\ Wife	15	15,675	15,690
Son\ Daughter	29,474	25,739	55,213
Stepson\ Stepdaughter	82	82	164
Father\ Mother	82	723	805
Parent-in-law	5	63	68
Grandson\ Granddaught	854	783	1,637
Brother\ Sister	420	556	976
Grandfather\ Grandmot	2	5	7
Sons wife\ Daughters	20	567	587
Other relatives	235	272	507
Servant\ Maid	4	363	367
No relation	197	7	204
Total	47,996	47,445	95,441

Tabulation

- لعمل جدول متقاطع لاكثر من متغيرين نستخدم الامر ***table***

```
table q502 sex urban
```

```
. table  q502 sex urban
```

	urban and Gender			
	1		2	
q502	male	female	male	female
Head	16,388	1,409	218	1,201
Husband\ Wife	15	15,675		
Son\ Daughter	17,099	13,445	12,375	12,294
Stepson\ Stepdaughter	55	48	27	34
Father\ Mother	66	490	16	233
Parent-in-law	4	47	1	16
Grandson\ Granddaughter	440	413	414	370
Brother\ Sister	390	482	30	74
Grandfather\ Grandmother	1	3	1	2
Sons wife\ Daughters husband	20	557		10
Other relatives	204	208	31	64
Servant\ Maid	4	340		23
No relation	194	6	3	1

.

Tabulation

- لعمل جدول يحتوي على الوسيط والمتوسط و ... نستخدم الامر ***table***

```
table q502 sex, contents(mean age median age)
```

```
. table  q502 sex , contents(mean age median age)
```

q502	Gender male	Gender female
Head	46.615318 45	57.860153 58
Husband\ Wife	57.133335 59	41.073822 40
Son\ Daughter	13.903502 13	13.34534 12
Stepson\ Stepdaughter	16.463415 14.5	17.597561 14
Father\ Mother	76.109756 78	68.914246 69
Parent-in-law	83.199997 83	69.428574 71
Grandson\ Granddaughter	6.5444965 5	6.2911878 5
Brother\ Sister	26.711905 25	37.884892 38
Grandfather\ Grandmother	84 84	82.800003 80

Tabulation

- جدول يوضح متوسط انفاق الاسرة مع جنس رب الاسرة.

```
. use individual.dta,clear
(dataset of person 2018)

. keep if q502==0
(18,965 observations deleted)

. table    sex, contents(mean exp)
```

Gender	mean(exp)
male	763.2135
female	500

Tabulation

- جدول يوضح نسبة الاسر التي تستخدم مصدر للتدفئة حسب المحافظة.

```
. use hh.dta,clear
(dataset of person 2018)

. tab gov q210 ,row nofreq
```

	Main source of heating in the dwelling?									
gov	Kerosene	Gas heate	Central h	Air-condi	Electrica	Firewood\	No heatin	No need f	Other	Total
11	30.73	46.02	2.04	1.38	7.23	9.06	1.33	2.15	0.05	100.00
14	32.08	43.40	3.77	0.00	11.32	1.89	0.94	6.60	0.00	100.00
21	31.64	49.60	2.86	0.64	6.52	4.77	1.27	2.70	0.00	100.00
31	29.83	37.94	0.23	0.57	6.74	17.14	4.46	2.97	0.11	100.00
Total	30.72	45.50	1.95	1.24	7.19	9.43	1.61	2.31	0.05	100.00

Tabulation

- جدول يوضح عدد الاسر التي تمتلك سيارة خاصة حسب المحافظة.

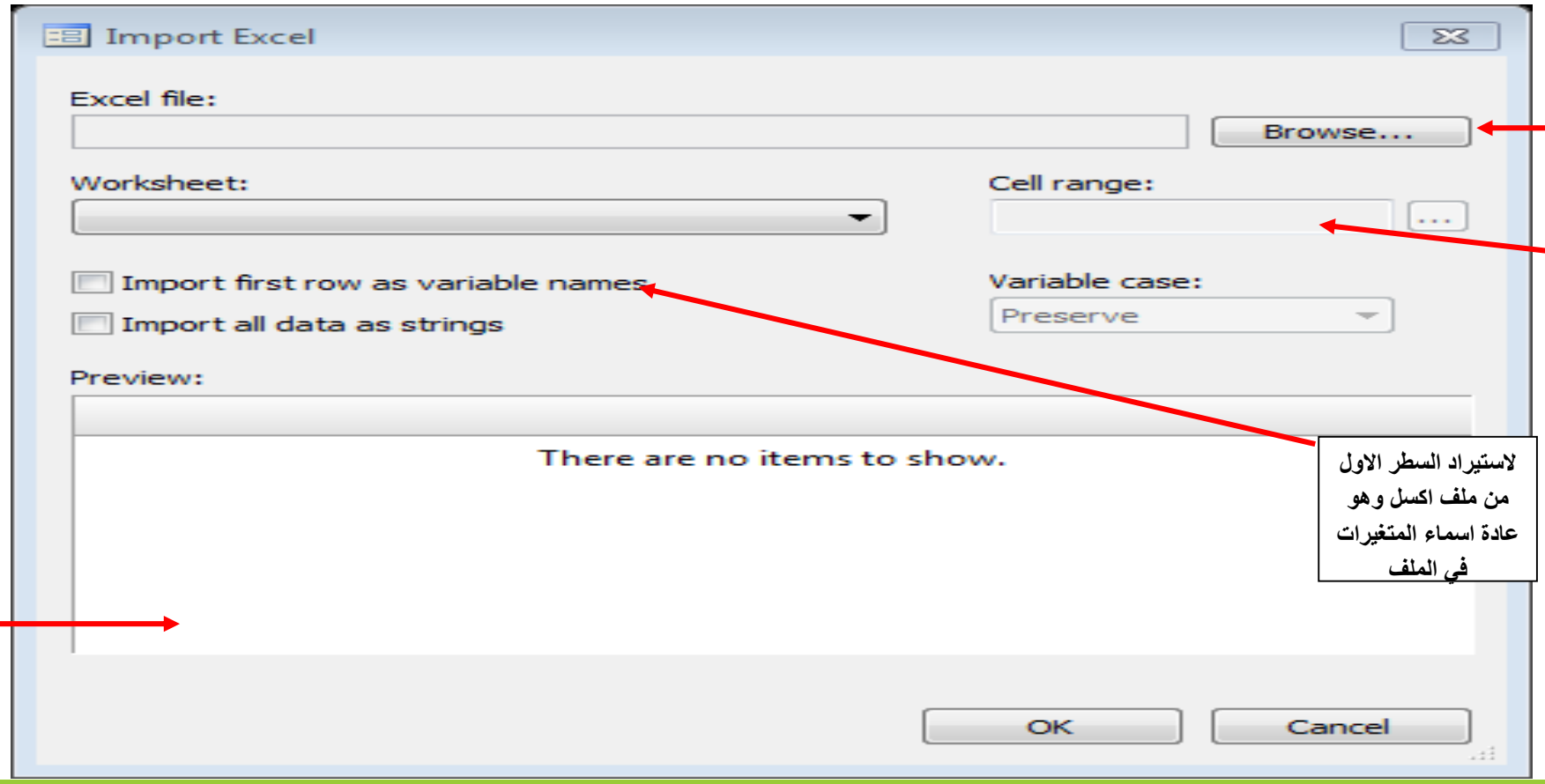
```
. use hh.dta,clear
(dataset of person 2018)

. tab gov car
```

gov	own car? yes	no	Total
11	4,179	3,856	8,035
14	32	74	106
21	192	437	629
31	156	719	875
Total	4,559	5,086	9,645

استيراد البيانات وتصديرها وكيفية التعامل معها

✓ لاستيراد ملف اكسل الى Stata (Import) :

- من قائمة file نختار import ثم نختار excel spreadsheet ثم تظهر الشاشة الاتية :



يتم اولا تحديد موقع ملف الاكسل المراد تحويله الى ستاتا

في حال عدم استيراد الملف كاملا يتم تحديد مدى البيانات المراد استيرادها

لاستيراد السطر الاول من ملف اكسل وهو عادة اسماء المتغيرات في الملف

يتم عرض ملف اكسل في هذه المنطقة

```
import excel "E:\Stata_Training\Train1.xls", sheet("Sheet1") firstrow clear
```

استيراد البيانات وتصديرها وكيفية التعامل معها

✓ **لاستيراد ملف spss الى ستاتا :**

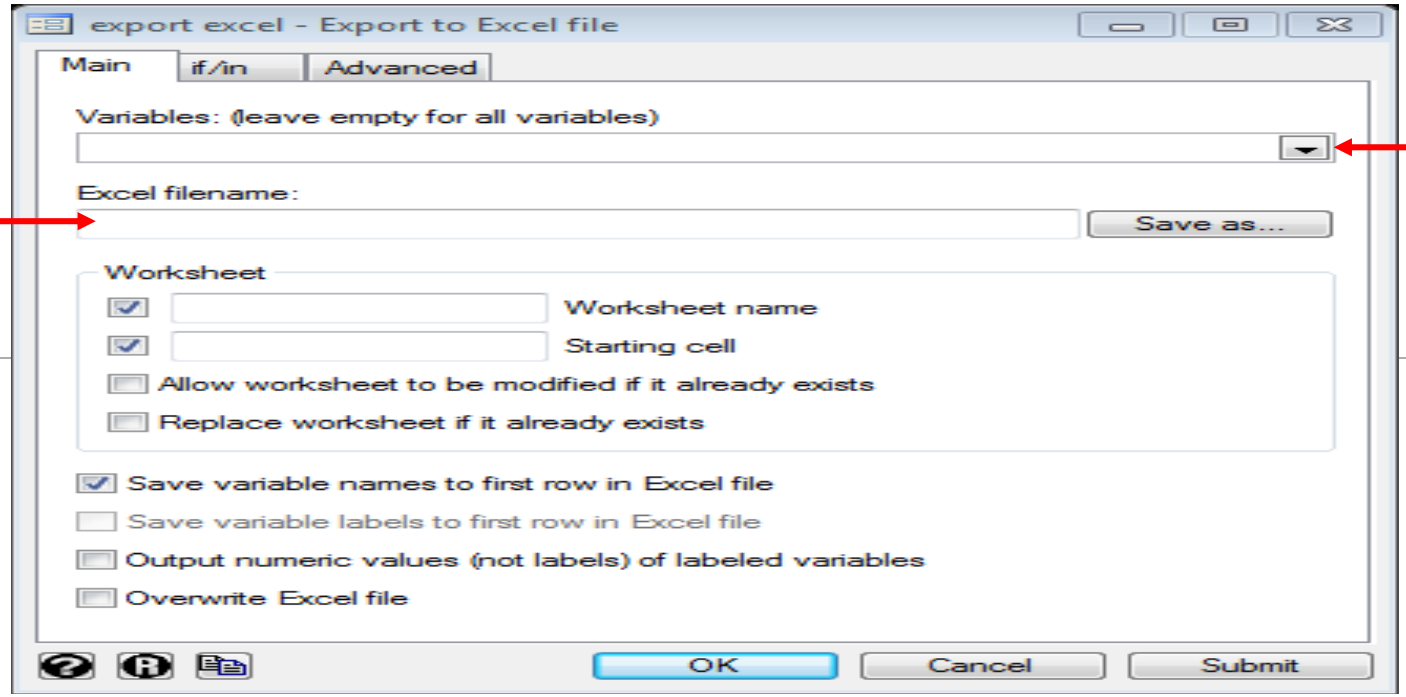
- يتم فتح الملف spss. على برنامج spss و من قائمة file نختار export ثم نختار Stata و يتم تخزين الملف باسم معين ومن ثم فتح الملف على ستاتا.

تصدير ملفات ستاتا الى ملف إكسل

✓ للتصدير ملف Stata الى اكسل (export) :

- يتم فتح الملف dta. المراد تحويله الى اكسل
- من قائمة file نختار export ثم نختار excel spreadsheet ثم تظهر الشاشة الاتية :

يتم كتابة اسم الملف المراد تخزينه بصيغة اكسل



تترك فارغة في حال تصدير جميع المتغيرات.

يتم اختيار المتغيرات المراد تصديرها من خلال الضغط على السهم الموجود على اليمين في حال لانريد جميع المتغيرات.

في حال تصدير الملف ضمن شرط معين يتم تحديده في تاب if

```
export excel using "E:\Stata_Training\hh_excel.xls", firstrow(variables) replace
```

تصدير ملف ستاتا الى ملف spss

✓ **للتصدير ملف Stata الى spss :**

- يتم فتح الملف dta. المراد تحويله الى spss

```
use "E:\Stata_Training\f1.dta"

savespss  "E:\Stata_Training\f1.sav"
```

- يمكن لبرنامج Stata رؤية خصائص ملف spss من دون فتح الملف.

```
desspss  using "E:\Stata_Training\f1.sav"
```

Log file

- تكمن أهمية هذه الملف في أنه يخزن جميع الأوامر والنتائج والجداول الاحصائية التي تظهر على نافذة النتائج ومن خصائص هذا الملف عند فتحه أنه يخزن اوتوماتيكياً بدون أن نعمل حفظ في كل مرة (.smcl)

- ويمكن فتح log من قائمة file ثم نختار log ثم نختار begin ومن ثم تظهر شاشة لتحديد المكان المراد تخزينه فيه او يمكن فتحه من خلال النقر على الأيقونة الموجودة في شريط الادوات.

do file

- هذا الملف هو شبيه تماماً ب syntax في برنامج spss ، وهذه النافذة تستخدم لتخزين جميع الأوامر التي تم تنفيذها والتي ستنفذ لاحقاً ويمكن استخدامه بدلا من نافذة الأوامر الموجودة في برنامج ستاتا (.do)

- يمكن فتح do file من خلال النقر على الأيقونة الموجودة في شريط الادوات او من قائمة Window ثم نختار do file editor ثم نختار new do file او كتابة الامر **doedit**.

انشاء ملف do file

١. تثبيت directory المراد فتح او تخزين الملفات عليه (حسب جهازك) ثم عرض محتوياته

٢. فتح ملف بيانات الافراد individual.dta واظهارها.

٣. اظهر وصف لكافة المتغيرات الموجود في ملف الافراد على شاشة النتائج.

٤. اظهر وصف للحقول: العمر والجنس

٥. اوجد عدد الافراد في الملف

٦. اوجد عدد الافراد الاردنيين الاناث

انشاء ملف do file

١. انشيء متغير جديد باسم q506_new بحيث تكون قيمه هي ١ اذا كانت قيمة سؤال q506 هي ١ وتكون قيمته هي ٢ اذا كانت قيمة سؤال q506 اكبر من ١.

٢. انشيء متغير max_age يحتوي على اكبر عمر فرد في كل اسرة .

٣. انشيء متغير hh_size يحتوي على عدد افراد الاسرة لكل اسرة.

٤. انشئ متغير جديد fam_type يحتوي على اسر صغيرة (hh_size 1-3) واسر متوسطة (hh_size 4-7) واسر كبيرة من (hh_size>7)

٥. اعد ترميز سؤال Q530 0→1 1→2

انشاء ملف do file

١. استخراج جدول تكراري للمتغير العلاقة برب الاسرة q502.

٢. استخرج عدد تكرارت الذكور والاناث حسب الجنسيات.

٣. استخرج جدول متقاطع بين متغير الالتحاق بالتعليم ومتغير الجنس.

٤. استخرج جدول تكراري من الحقل q502 الى الحقل exp للاردنيين فقط.

٥. استخرج الوسيط الحسابي والانحراف المعياري واعلى واقل قيمة للمتغير q506.

١. استخرج المتوسط الحسابي للمتغير العمر حسب الجنس.

انشاء ملف do file

١. استخرج جدول يبين نسبة الذكور والاناث حسب الجنسية

٢. استخرج المتوسط والوسيط للمتغيرات العمر ودخل الفرد حسب المحافظة

٣. استخرج متوسط العمر للذكور والاناث في كل محافظة.

٤. ادمج البيانات الموجودة في ملف q1 مع بيانات ملفين q2,q3 ومن ثم تخزين الملف باسم all

٥. من الملف all السابق قم بانشاء ملف مجمع لاعداد السكان على مستوى كل المحافظة ومن ثم تخزينه باسم all_agg

q1.dta + q2.dta + q3.dta = **all.dta**

	qno	gov	pop_tot
1	1	11	20000
2	1	12	10000
3	1	13	15000
4	1	14	5000
5	1	21	18000
6	1	22	2000
7	1	23	2500
8	1	24	8000
9	1	31	1500
10	1	32	2500
11	1	33	3500
12	1	34	3000

	qno	gov	pop_tot
1	2	11	21000
2	2	12	12000
3	2	13	16000
4	2	14	6000
5	2	21	20000
6	2	22	2200
7	2	23	2800
8	2	24	8500
9	2	31	1700
10	2	32	2600
11	2	33	3800
12	2	34	3200

	qno	gov	pop_tot
1	3	11	22000
2	3	12	12500
3	3	13	17000
4	3	14	6500
5	3	21	20500
6	3	22	3000
7	3	23	3200
8	3	24	8800
9	3	31	1900
10	3	32	2800
11	3	33	3600
12	3	34	3500

	qno	gov	pop_tot
1	1	11	20000
2	2	11	21000
3	3	11	22000
4	1	12	10000
5	2	12	12000
6	3	12	12500
7	1	13	15000
8	2	13	16000
9	3	13	17000
10	1	14	5000
11	2	14	6000
12	3	14	6500

انشاء ملف do file

١. من ملف الافراد قم بانشاء ملف لخصائص رب الاسرة (كل اسرة لها رب اسرة واحد) ودمجه مع ملف الاسر hh وتخزينه باسم hh_head

٢. ادمج حقل المحافظة وحقل الحضر والريف urbanمع ملف الافراد individual

٣. احفظ بيانات الاسئلة hhid و q308-q329 بملف جديد ind_long ثم قم بتحويل بيانات كل اسرة بشكل طولي ثم قم بعرض البيانات.

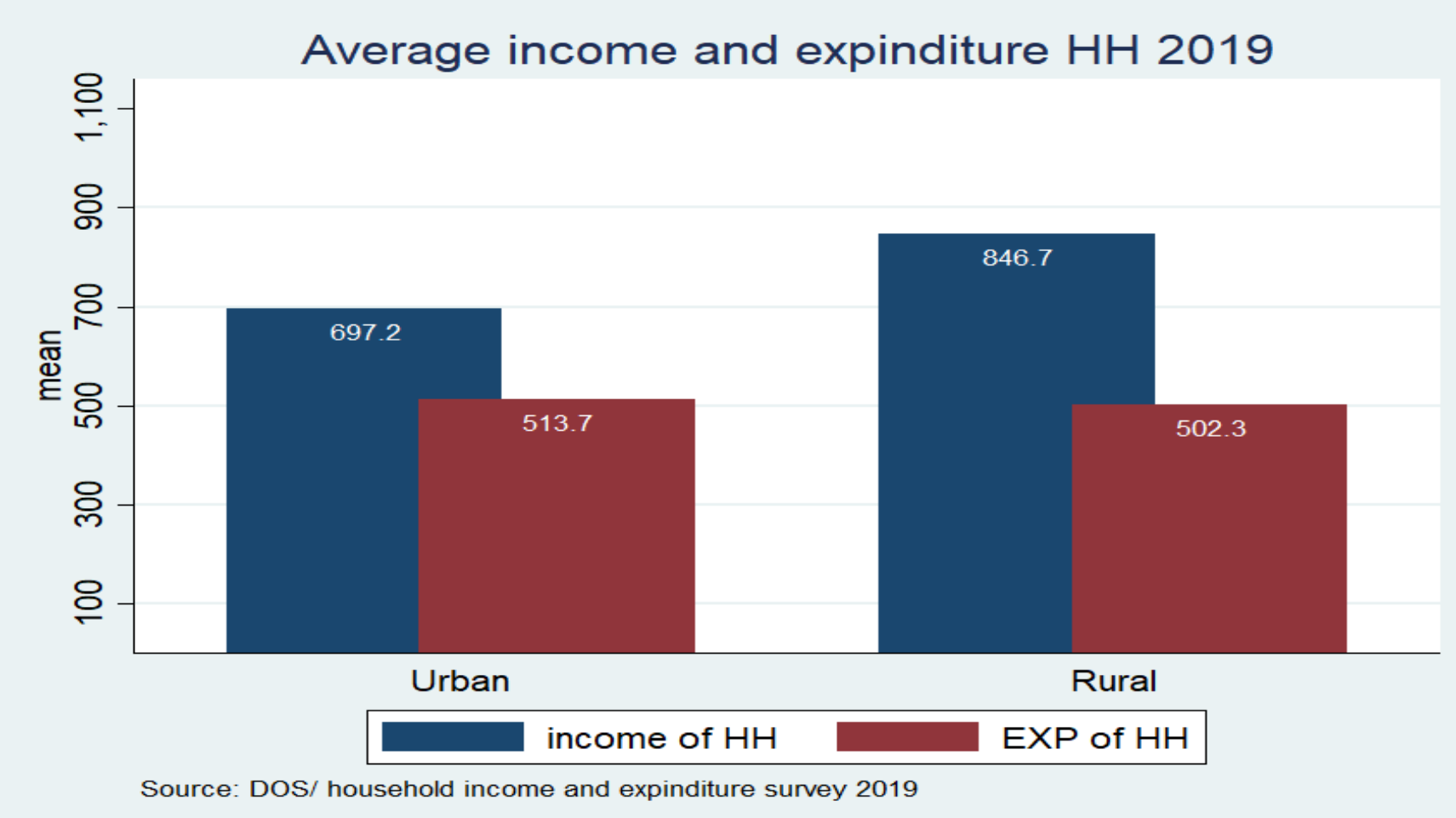
٤. قم بفتح ملف income_yearly.dta وحول البيانات بشكل طولي ومن ثم استعرض البيانات

Graph

- اعرض رسم بياني يمثل متوسط دخل وانفاق الفرد حسب الحضر والريف

```
use individual,clear
 #delim ;
graph bar exp incom ,over(urban)
legend(label(1 "income of HH") label(2 "EXP of HH"))
bargap(-30)
ytitle("mean")
title("Average income and expinditure HH 2019")
note("Source: DOS/ household income and expinditure survey 2019")
blabel(bar, position(inside) format(%9.1f) color(white))
ylabel(100(200)1100);
```

Graph

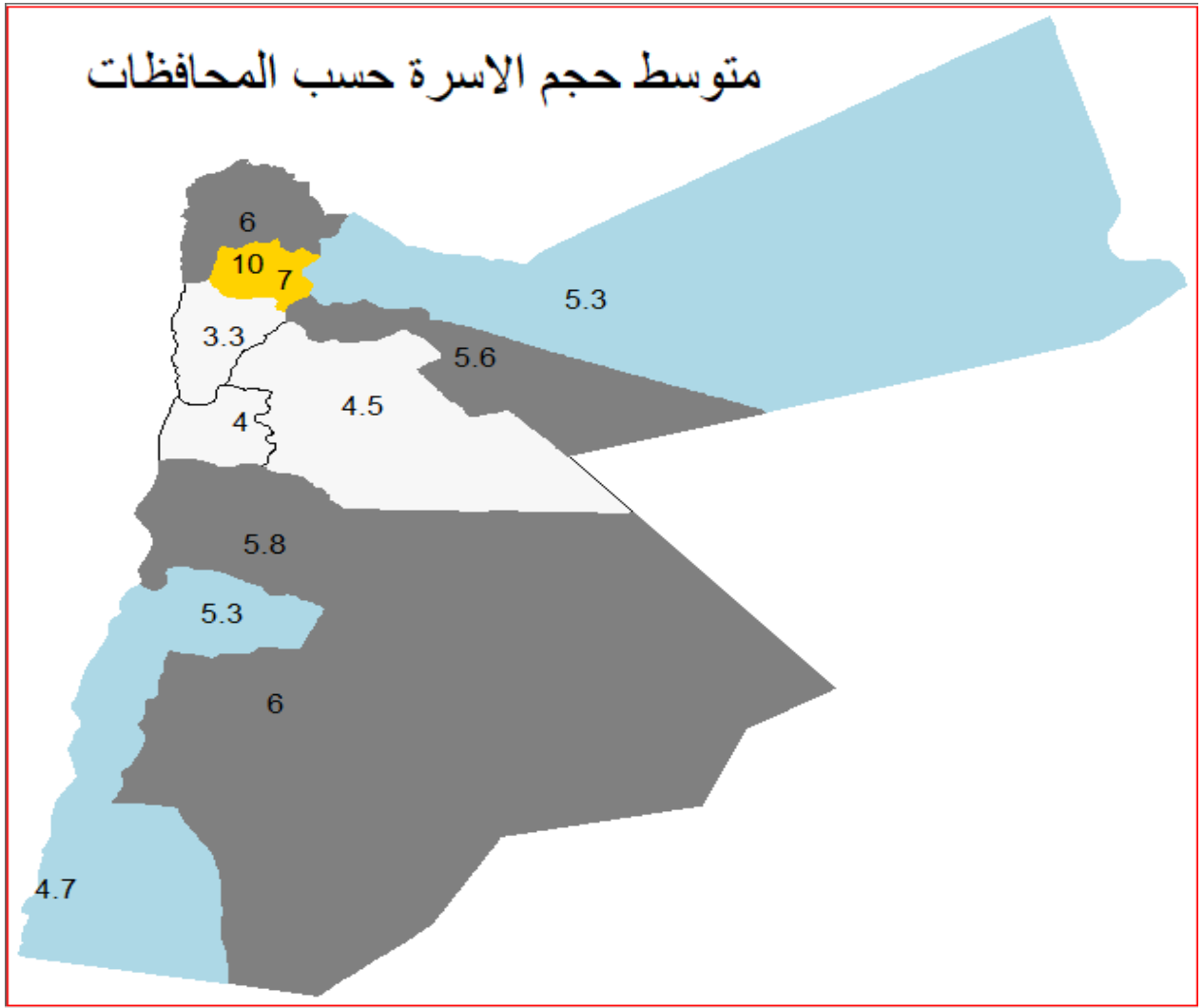

Average income and expinditure HH 2019
mean
100
300
500
700
900
1,100
697.2
513.7
846.7
502.3
Urban
Rural
income of HH
EXP of HH
Source: DOS/ household income and expinditure survey 2019

Map

- اعرض خريطة تمثل نسبة حجم الاسر في المحافظات

```
spmap  hh_size using jor_coor_shap.dta, id(id)

spmap hh_size using "jor_coor_shap.dta", id(id) label(xcoord(_X) ycoord(_Y)
label(hh_size))
```

Map



Thank
you