



Rania Hashem Abu_Dhaim

Fatima Shawkat Al_Awamreh



Stata



- Introduction about Stata
- Installation and Editions of Stata
- Adv./DisAdv of Stata.
- Component of Stata Interface
- Use, clear command.
- Save, replace command
- Log file
- Do file
- cd and dir command
- If condition
- describe, list, count commands
- lookfor, browse commands
- label command
- generate , replace command



- Practical Exercises



هي حزمة برامج إحصائية للأغراض العامة تم إنشاؤها في عام ١٩٨٥ من قبل **StataCorp**.

يُعَمَّلُ مُعَظَّمُ مُسْتَخَدِمِيهَا فِي البحوث، وَلَا سِيمَّا فِي مَجَالَاتِ الْإِقْتِصَادِ وَالْإِجْتِمَاعِ وَالْعِلُومِ السِّيَاسِيَّةِ وَالْطَّبِيعِيَّةِ وَالْحَيْوِيَّةِ وَالْعِلْمِ الْأَوْبَيَّةِ. وَلَسْتَاتَا قَدْرَاتٌ تَشْكُلُ إِدَارَةَ الْبَيَانَاتِ، وَالْتَّحْلِيلَ الإِحْصَائِيَّ، وَالرَّسُومَاتِ، وَالْمُحاَكَاهِ، وَتَحْلِيلَ الْانْهَارِ.

▪ **Stata**: هي اختصار لكلمة الاحصاء والبيانات.

Statistical

Data



✓ من مميزاتها:

- يمكن التعامل مع ملفات البيانات الكبيرة
- يمكن اجراء عمليات التحليل اما بكتابة الاوامر او من خلال القوائم
- يعتبر برنامج لديه قوة في التحليل الاحصائي

✓ من سلبياته:

- تكلفة الرخصة: تطلب رخصة مدفوعة وهذا يمكن أن يكون مكلفاً بالنسبة للأفراد والمؤسسات، خاصة إذا كانت هناك حاجة إلى عدد كبير من التراخيص.
- صعوبة تعلمها: تحتاج إلى تعلم لغة البرمجة التي تعتمد عليها، وهذا قد يكون تحدياً للأشخاص الذين ليس لديهم خلفية في البرمجة.



Stata/MP

Stata/SE

Stata/BE

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

File Edit Data Graphics Statistics User Window Help

Review T X

Command rc

There are no items to show.

Statistics/Data Analysis (R) 13.0 Copyright 1985-2013
Special Edition
StataCorp
4905 Lakeway Drive
College Station, Tex
800-STATA-PC
979-696-4600
979-696-4601 (fax)

3-user Stata network perpetual license:
Serial number: 501306208483
Licensed to: Beatriz Godoy
Systems Integrales

Notes:
1. (/v# option or -set maxvar-) 5000 maximum variab

Results (4)

Command

Properties CAP NUM OVR

Variables Variable Label

There are no items to show.

2 المراجعة (Review)

1 المتغيرات (Variables)

3 الأوامر (Command)

C:\Users\beatr\Documents



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 "F:\Stata_Training\hh.dta", clear
```



تخزين ملفات بصيغة Stata

✓ يتم تخزين ملف البيانات :

• يتم تخزين ملف جديد عن طريق الامر :

```
save "F:\Stata_Training\hh.dta"
```

• يتم تخزين ملف مخزن مسبقا عن طريق الامر :

```
save "F:\Stata_Training\hh.dta", replace
```



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 begin ومن ثم تظهر شاشة لتحديد المكان المراد تخزينه فيه او يمكن فتحه من خلال النقر على الأيقونة الموجودة في شريط الأدوات.



do file

- هذا الملف هو شبيه تماماً ب syntax spss في برنامج syntax ، ويستخدم لتخزين جميع الأوامر التي تم تنفيذها والتي ستنفذ لاحقاً ويمكن استخدامه بدلاً من نافذة الأوامر الموجودة في برنامج ستاتا (.do).
- يمكن فتح do file من خلال النقر على الأيقونة الموجودة في شريط الأدوات او من قائمة Window ثم نختار doedit او كتابة الامر new do file او كتابة الامر do file editor.



تثبيت مسار العمل ***cd command***

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

```
cd "F:\Stata_Training"
```

```
use "individual.dta"
```

```
save "ind.dta"
```

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

```
dir
```



تثبيت مسار العمل change directory

Exercise1

- قم بتنصيب مسار العمل على مجلد **Stata_Training**
- استعرض محتويات المجلد
- قم بفتح ملف **individual.dta**
- خزن الملف باسم **individual1.dta** على نفس المسار ومن ثم قم بفتحه وتأكد من انه تم تخزينه على نفس المسار



تثبيت مسار العمل change directory

solution

```
cd E:\Stata_Training
```

```
dir
```

```
use individual.dta
```

```
save individual1.dta,replace
```

```
use individual1.dta,clear
```

```
dir
```



If condition

==	مساواة
>	اكبر من
<	اقل من
>=	اكبر او يساوي
<=	اقل او يساوي
!=	لا يساوي
&	and
	or

{command} if var1 == var2



describe command

- وهو أمر بسيط لوصف المتغيرات داخل الملف من حيث نوعها و عددها و حجمها ...
- هنالك نوعان :

✓ وصف بيانات للملف مفتوح :

```
describe  
des  
d
```

✓ وصف بيانات لملف غير مفتوح:

```
describe using "hh"
```



list command

- هذا الأمر يقوم باستعراض البيانات الموجودة داخل الملف على شاشة النتائج بحيث يمكن استعراض جميع البيانات أو جزء من البيانات حسب الطلب.

```
list
```

```
list q201 gov
```

```
list q201 if gov==11
```

```
list in 1/10
```

```
list gov in 1/10
```



count command

- يستخدم لمعرفة عدد مشاهدات الملف المفتوح

```
count  
count if gov==11
```



describe,list,count commands

Exercise2

► قم بفتح ملف hh.dta ووصفه

```
use hh.dta  
desc
```

► قم بوصف متغير ال gov فقط

```
des governorate
```

► انشئ قائمة تحتوي على q201 فقط لمحافظة 31

```
list q201 if governorate==31
```



describe, list, count commands

Exercise2

► عدد المشاهدات في محافظة البلقاء والزرقاء (البلقاء= ١٢ و الزرقاء= ١٣)

```
count if governorate==12 | governorate==13
```

► وصف wage في ملف month1 من دون فتحه

```
des month1 using "wage"
```



lookfor command

- تستخدم للبحث عن اسم متغير او وصف متغير داخل الملف المفتوح.

```
lookfor x
```

```
lookfor 201
```

```
lookfor gov
```



browse command

- يستخدم هذا الامر لعرض البيانات لملف المفتوح.

```
browse
```

```
br gov
```

```
br q201 if gov==11
```

```
br, nolabel
```



lookfor,browse commands

Exercise3

► قم بفتح ملف hh ثم البحث عن house

```
use "hh.dta"
```

```
lookfor house
```

► عرض بيانات محافظة ٣١ اذ كان نوع المسكن = ٢ حيث يكون عرض لقيم المتغيرات بدون ظهور وصف لها

```
br if gov==31 & q201==2, nolabel
```



label command

✓ **label data:(label dataset):**

● لاعطاء ملف البيانات وصف عن بيانات هذا الملف .

label data "information of household"

● لمشاهدة هذا الوصف عن طريق امر

desc



label command

✓ **label variable:**

لإعطاء متغيرات ملف البيانات وصف يساعد في فهم هذه المتغيرات

```
label variable name of variable "label of variable"  
label variable governorate "governorate"
```

✓ **define value label:**

تعريف قيم المتغير واعطائه وصف

```
label define gv 11 amman 12 balqa 13 zarqa
```



label command

- ✓ assign value label to variables:

اسناد القيم للمتغير

```
label value gov gv
```

- ✓ list names of value labels:

```
label dir
```

- ✓ list names and contents of value labels:

```
label list
```

- ✓ drop labels:

```
label drop gv
```



label commands

Exercise4

► قم بفتح ملف hh ثم انشئ وصف لملف Householday1

```
use "hh.dta",clear
```

```
label data "Householday1"
```

► قم بإنشاء وصف للمتغير Type of Dwelling <- q201

```
label variable q201 "Type of Dwelling"
```



label commands

Exercise4

► قم بإنشاء وصف للمتغيرات السؤال q201 كالتالي:
dar ٢ villa ١

```
label define lb_q201 1 villa 2 dar
```

```
label value q201 lb_q201
```

► قم بتعديل وصف متغيرات السؤال q201 كالتالي: ٣
hut ٤ apartment ٣

```
label define lb_q201 3 apartment 4 hut, modify
```



label commands

Exercise4

► قم بإنشاء وصف لقيم المتغيرات q308 q309 q310 q312 كالتالي:

No ٢ Yes ١

```
label define lb_yn 1 yes 2 no
```

```
label value q308 q309 q310 q312 lb_yn
```

► استعرض label مع وصف قيمها

```
label list
```

► اشطب وصف المتغير Q201

```
label drop lb_q201
```



generate command

- ✓ يستخدم هذا الامر لانشاء متغير جديد ويمكن ان يكون انشاءه معتمد على متغير اخر في ملف البيانات.
- ✓ يستخدم لمعرفة التكرارات على المستوى الذي نختاره في البيانات و خاصة عند تدقيق البيانات.

```
gen x =1
```

```
gen r =1 if gov==11
```

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

```
gen var5= _n
```



replace command

✓ هذا الأمر هو 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
```



generate, replace commands

Exercise5

► قم بفتح ملف hh ثم انشئ متغير جديد باسم group

```
gen group =.
```

► استبدل قيم المتغير الجديد group كالتالي:

- محافظات الوسط ونوع السكن دار = ١
- محافظات الشمال ونوع السكن دار = ٢
- محافظات الجنوب ونوع السكن دار = ٣



generate, replace commands

Exercise5

```
replace group=1 if q201==2 & governorate<15
```

```
replace group=2 if q201==2 & governorate>14 & governorate<25
```

```
replace group=3 if q201==2 & governorate>24 & governorate<35
```

► قم بفتح ملف الافراد واعد تسلسل الافراد داخل كل اسرة:

```
use " individual.dta"
```

```
bys hhid: gen num = _n
```

Stata



- egen command
- recode command
- keep,drop commands
- rename command
- order,sort commands
- Tabulation
- Summarize command
- Append command
- Merge command

Day2

- Practical Exercises



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)



egen command

- ✓ عند استخدام امر egen مع rowtotal و كان جميع قيمه تساوي missing عذها تكون اجابة صفر، في حال اردنا ان تبقى نتائج missing نضع مع الامر m

```
egen total_wage1= rowtotal (month1 month2 month3),m
```



egen commands

Exercise6

- قم بفتح ملف hh ثم انشاء متغير جديد يحتوي على عدد الافراد في كل محافظة حيث يظهر على مستوى كل سطر

```
bys governorate: egen sum_family= total (hh_size)
```

- انشاء متغير جديد يحتوي على أقل اجابة للاسئلة : q308 q309 q310

```
egen min308_310= rowmin(q308 q309 q310 )
```



recode command

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

```
recode sex (1=0) (2=1)
```

```
recode sex (1=0) (2=1), gen (gender)
```

```
recode q522 (1 2 = 1) (3 = 2) (4/9= 3), gen(cat_edu)
```

```
recode q522 (1 2 = 1 "Illiterate") (3/6 = 2 "diploma and less") (7/12 = 3 "Bachelor and above"), gen(cat_edu1)
```



recode command

Exercise 7

► قم بفتح ملف individual ثم انشاء متغير جديد من متغير الجنس باسم sex_new بحيث تكون قيمة ٠ للذكر و ١ للانثى مع انشاء وصف للقيمة الجديدة

```
recode sex (2=1 "female") (1=0 "male"), gen(sex_new)
```



keep and drop commands

- إذا أردنا إبقاء عدد من المتغيرات أو عدد من المشاهدات نستخدم الأمر **keep**
- أما إذا أردنا حذف بعض المتغيرات والمشاهدات نستخدم الأمر **drop**.

```
keep gov cluster
```

```
keep if gov==11
```

```
keep in 1/20
```

```
drop gov q201
```

```
drop if gov==22
```

```
drop in 1/5
```

```
drop q*
```



rename command

► يستخدم هذا الامر لاعادة تسمية المتغيرات.

rename governorate gov لمتغير واحد

rename (q201 hhid) (tybe_of_department hosehold_id) لعدة متغيرات



keep,drop,rename commands

Exercise8

► قم بفتح ملف individual واحفظ أول ٥ حالات من البيانات

```
use individual
```

```
keep in 1/5
```

► احذف البيانات اذا كانت المحافظة عمان الرمز ١١

```
drop if gov==11
```

► اعد تسمية سؤال age بسؤال q111

```
rename age q111
```



order Command

► يستخدم أمر ال **order** لنقل المتغيرات واعادة ترتيبها حسب الطلب.

order cluster

order hhid cluster sex q506-q529

order sex , after (cluster)

order sex, before (hhid)



sort Command

➢ يستخدم أمر ال **sort** لترتيب البيانات (القيم) تصاعدياً أو تنازلياً

--Ascending order

sort id

sort id gov

--Descending order

gsort -id

gsort -gov



order,sort commands

Exercise9

► قم بفتح ملف **individual**

► قم بترتيب متغير **idcode** بعد **hhid**

```
use individual
```

```
order idcode, after(hhid)
```



order,sort commands

Exercise9

► قم بترتيب متغير **cluster** حيث يكون اخر متغير في الملف

```
order cluster, last
```

► قم بترتيب البيانات تنازلياً حسب سؤال **cluster**

```
gsort- cluster
```



Tabulation

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

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

tab1 q522

لظهور الحالات المفقودة (.) tab1 q522, m

لظهور خيارات السؤال بدون وصف قيم السؤال tab1 q522, nolabel

tab1 q522 if gov==11 بوجود شرط معين

bys sex: tab1 q502

tab1 q522 sex q502 q506



Tabulation

Exercise 10

► قم بفتح ملف individual

► انشئ جدول تكراري لسؤال q502 لمحافظة اربد

```
use individual
```

```
tab1 q502 if gov==21
```

► انشئ جدول تكراري لسؤال q502 مع ظهور حالات القيم المفقودة ان وجدت وبدون ظهور تعريف للقيم فقط قيم

```
tab1 q502 ,m nolab
```



Tabulation

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

tab q522 sex

ظهور الحالات المفقودة

tab q522 sex , m

ظهور خيارات السؤال بدون تعريف (ارقام)

tab q522 sex if gov==21

بوجود شرط معين

tab q522 sex, row noref

ظهور النسب افقي

tab q522 sex, col noref

ظهور النسب عمودي

bys gov: **tab** q502 sex

عرض التكرارات بشكل تفصيلي مجزأ



Tabulation

Exercise 11

- قم بفتح ملف individual
- انشاء جدول يربط بين سؤال q515 مع سؤال q502 فقط للذكور

```
use individual
```

```
tab q515 q502 if sex==1
```



Tabulation

Exercise12

- انشاء جدول يربط بين سؤال q515 مع سؤال q502 فقط للذكور مع ظهور النتيجة مجزأة حسب كل جنسية على حدة

```
bys q506: tab q515 q502 if sex==1
```



Tabulation

Exercise12

➢ انشاء جدول يربط بين سؤال q515 مع سؤال q502 فقط للذكور مع ظهور النتيجة مجزأه حسب كل جنسية على حدة مع ظهور القيم المفقودة وبدون تعريف لقيم السؤال
(الجدول بنسب فقط بدون قيم وبشكل افقي)

```
bys q506: tab q515 q502 if sex==1 , m row nofreq nolabel
```



summarize command

- يستخدم الأمر **summarize** لحساب وعرض بعض الاحصاءات الوصفية لمتغير ما مثل: المتوسط ، الانحراف المعياري، الحد الأدنى والحد الأقصى لكل متغير أو أكثر.
- إذا لم يتم تحديد/ تعين أي متغير، فإنه يتم حساب الاحصاءات لجميع متغيرات الملف.

```
use "hh.dta", clear  
  
sum hh_size  
  
sum hh_size if gov==11  
  
sum hh_size, detail  
  
bys gov: sum hh_size
```



Tabulate, Summarize

► وهو 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 q522 sex , sum (age)
```

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



Tabulate, Summarize

Exercise13

- قم بفتح ملف individual
- استخراج متوسط العمر (age) فقط اذا كان الجنس ذكر

```
use individual
```

```
sum age if sex==1
```



Tabulate, Summarize

Exercise13

► متوسط العمر (age) والقيمة الدنيا والعليا حسب الجنس مجزأ

```
bys sex: sum age
```

► جدول يبين الانحراف المعياري للعمر بين سؤالين الجنس و q513 بدون ظهور label وبدون ظهور عدد المشاهدات

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



Tabulation

✓ وهو أمر يستخدم لمعرفة التقابلات بين أكثر من متغيرين

(cross tabulation)

table q201 q204 q308

table q201 q204 q308, by(urban)

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

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 وافقى

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)



Table command

Exercise14

► قم بفتح ملف **wagem** و استخراج متوسط و وسيط قيم الدخل للشهر الثالث حسب ال group مع وجود المجموع افقي و

► قم بفتح ملف **wagem** و استخراج متوسط و وسيط قيم الدخل للشهر الثالث حسب ال group مع وجود المجموع افقي و

format 2 decimals

```
use wagem
```

```
table group , contents(mean month3 median month3) row
```

```
format(%9.2f)
```



Tabulation

✓ وهو امر يستخدم لعرض ملخص لعدة متغيرات رقمية (Continuous) متواصل في

جدول واحد ويستخدم لعرض نتيجة متغيرات (Continuous) حسب متغيرات مصنفة

(categories)

`tabstat month3`

يعطي متوسط القيم مباشرة

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

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



Tabulation

Exercise15

- قم بفتح ملف **wagem**
- استخراج متوسط ووسيط قيم الدخل للشهر الأول **month1** والشهر الثاني **month2** حسب متغير الـ **group**

```
use wagem
```

```
tabstat month1 month2, 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 متقطع لمتغيرين نستخدم الامر **tab**

```
tab q502 sex
```



. tab q502 sex

q502	Gender		Total
	male	female	
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 متقطع لاكثر من متغيرين نستخدم الامر ***table***

```
table q502 sex urban
```



. table q502 sex urban

q502	urban and Gender			
	1		2	
	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 q201 q204, contents(mean hh_size median hh_size)
```



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

q502	Gender	
	male	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		82.800003 80

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

append command



يتم اولا فتح احد الملفات ودمجه على الملف الآخر 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

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

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

append command



Exercise16

► قم بدمج الاسر الموجودة في ملف hh الى الاسر الموجودة في ملف hh_new

```
use hh.dta, clear
```

```
append using "hh_new.dta"
```

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]

	hhid[1]	n_head	101
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]

	hhid[1]	n_4y	101
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	



merge command



- Merge 1:1 --→ merge one to one

Data Editor (Browse) - [ind1.dta]

File Edit View Data Tools

hhid[1] 101

	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



First file

```
use ind1.dta, clear
```

```
merge 1:1 hhid using ind2.dta
```

One-to-one

merge

Key
variable

Second file

Exercise17



- فتح file do جديد وتخزينه باسم merge1_1
- فتح ملف ind1
- عمل merge مع ملف ind2
- تخزين الملف المدمج باسم merge_all
- ملاحظة: ملف ind1 وملف ind2 ملفين على مستوى الاسر وان key هو hhid

merge command



```
cd D:\Stata_Training_poverty\Data
```

```
*****1
```

```
use ind1,clear
```

```
br
```

```
*****2
```

```
merge 1:1 hhid using ind2
```

```
br
```

```
*****3
```

```
save merge_all,replace
```

```
*****
```

merge command



Result	Number of obs
Not matched	2
from master	1 (_merge==1)
from using	1 (_merge==2)
Matched	18 (_merge==3)

merge command



```
cd D:\Stata_Training_poverty\Data  
*****  
use ind1,clear  
br  
*****  
merge 1:1 hhid using ind2  
br  
*****  
keep if _merge==3  
drop _merge  
save merge_all,replace  
*****
```

merge command



- Merge m:1 --> merge many to one

Name of file: : indiv_m

Type of file:many

	hhid	idcode	q502	sex	q504_m	q504_y	age	q506	q513	q515	q522	q524	q529	q530	incom	exp	food
1	101	1	0	1	3	1967	50	3	2	2	7	2	1	1	450	.	.
2	101	3	2	1	9	1997	19	3	2	2	5	1	1	1	450	.	.
3	101	4	2	1	9	1997	19	3	2	2	5	1	2	1	450	.	.
4	101	5	2	1	5	2001	16	3	2	1	.	1	2	1	450	.	.
5	101	6	2	2	7	2009	8	3	2	1	.	.	.	2	1500	500	.
6	102	1	0	1	6	1964	53	1	2	2	9	3	2	1	1000	1000	400
7	103	1	0	1	3	1940	77	1	1	2	9	1	1	1	450	.	.
8	108	1	0	1	7	1950	67	1	1	2	3	2	2	1	450	.	.
9	108	2	1	2	1	1953	64	1	1	2	3	2	2	2	450	.	.
10	110	1	0	1	12	1949	67	1	1	2	10	2	2	1	450	.	.
11	110	4	2	2	4	1982	35	1	1	2	9	1	1	2	450	.	.
12	202	1	0	2	4	1960	57	1	2	2	7	2	2	2	450	.	.
13	202	3	2	2	2	1982	35	1	2	2	5	1	2	2	450	.	.
14	204	1	0	2	3	1971	46	1	2	2	9	1	1	2	450	.	.
15	207	1	0	1	11	1963	54	1	1	2	4	2	1	1	450	.	.
16	207	5	2	1	9	2008	9	1	1	1	.	.	.	1	450	.	.
17	207	7	2	2	9	2012	5	1	1	1	.	.	.	2	450	.	.
18	301	1	0	1	4	1969	49	1	2	2	9	2	1	1	450	.	.
19	301	5	2	1	11	1997	20	1	2	1	.	1	2	1	450	.	.
20	301	6	2	2	1	2006	12	1	2	1	.	.	.	2	450	.	.
21	301	9	2	1	11	2012	5	1	1	1	.	.	.	1	450	.	.
22	303	1	0	1	2	1962	56	4	1	2	9	2	2	1	450	.	.
23	303	2	1	2	6	1962	55	4	1	2	9	2	2	2	450	.	.
24	303	3	2	2	6	1989	28	4	1	2	9	1	2	2	450	.	.
25	303	4	2	2	6	1988	29	4	1	2	9	1	2	2	450	.	.

merge command



- Merge m:1 --> merge many to one

Name of file: hh_m

Type of file:one

	hhid	q201	q202	q203	q204	q308	q309	q310
1	101	3	80	2	3	2	1	2
2	102	3	90	1	1	2	2	2
3	103	3	170	1	3	1	2	2
4	108	3	529	3	1	1	2	1
5	110	3	130	1	1	1	1	2
6	202	3	200	2	1	2	1	2
7	204	3	55	1	3	2	1	2
8	207	3	105	1	1	1	1	1
9	301	3	100	2	3	1	1	2
10	303	3	50	1	3	2	1	1
11	304	3	85	1	3	2	1	2
12	305	3	95	3	3	1	1	1
13	404	3	145	1	1	2	2	2
14	406	3	110	1	1	2	1	2
15	407	3	130	1	1	2	1	1
16	412	3	180	1	1	2	1	2
17	501	3	200	1	1	2	1	2
18	506	3	100	1	3	2	1	2
19	507	3	270	1	1	2	1	2
20	602	3	100	1	3	2	1	2
21	604	3	120	1	1	2	1	1
22	605	3	120	1	1	2	1	2
23	606	3	160	1	1	2	1	2
24	608	3	100	1	3	2	1	2
25	702	3	225	2	1	1	1	2

merge command



- Merge m:1 --> merge many to one

	hhid	idcode	q502	sex	q504_m	q504_y	age	q506	q513	q515	q522	q524	q529	q530	incom	exp	food	nonfood	q201	q202	q203
1	101	1	0	1	3	1967	50	3	2	2	7	2	1	1	450	.	.	.	3	80	2
2	101	3	2	1	9	1997	19	3	2	2	5	1	1	1	450	.	.	.	3	80	2
3	101	4	2	1	9	1997	19	3	2	2	5	1	2	1	450	.	.	.	3	80	2
4	101	5	2	1	5	2001	16	3	2	1	.	1	2	1	450	.	.	.	3	80	2
5	101	6	2	2	7	2009	8	3	2	1	.	.	.	2	1500	500	.	350	3	80	2
6	102	1	0	1	6	1964	53	1	2	2	9	3	2	1	1000	1000	400	600	3	90	1
7	103	1	0	1	3	1940	77	1	1	2	9	1	1	1	450	.	.	.	3	170	1
8	108	1	0	1	7	1950	67	1	1	2	3	2	2	1	450	.	.	.	3	529	3
9	108	2	1	2	1	1953	64	1	1	2	3	2	2	2	450	.	.	.	3	529	3
10	110	1	0	1	12	1949	67	1	1	2	10	2	2	1	450	.	.	.	3	130	1
11	110	4	2	2	4	1982	35	1	1	2	9	1	1	2	450	.	.	.	3	130	1
12	202	1	0	2	4	1960	57	1	2	2	7	2	2	2	450	.	.	.	3	200	2
13	202	3	2	2	2	1982	35	1	2	2	5	1	2	2	450	.	.	.	3	200	2
14	204	1	0	2	3	1971	46	1	2	2	9	1	1	2	450	.	.	.	3	55	1
15	207	1	0	1	11	1963	54	1	1	2	4	2	1	1	450	.	.	.	3	105	1
16	207	5	2	1	9	2008	9	1	1	1	.	.	.	1	450	.	.	.	3	105	1
17	207	7	2	2	9	2012	5	1	1	1	.	.	.	2	450	.	.	.	3	105	1
18	301	1	0	1	4	1969	49	1	2	2	9	2	1	1	450	.	.	.	3	100	2
19	301	5	2	1	11	1997	20	1	2	1	.	1	2	1	450	.	.	.	3	100	2
20	301	6	2	2	1	2006	12	1	2	1	.	.	.	2	450	.	.	.	3	100	2
21	301	9	2	1	11	2012	5	1	1	1	.	.	.	1	450	.	.	.	3	100	2
22	303	1	0	1	2	1962	56	4	1	2	9	2	2	1	450	.	.	.	3	50	1
23	303	2	1	2	6	1962	55	4	1	2	9	2	2	2	450	.	.	.	3	50	1
24	303	3	2	2	6	1989	28	4	1	2	9	1	2	2	450	.	.	.	3	50	1
25	303	4	2	2	6	1988	29	4	1	2	9	1	2	2	450	.	.	.	3	50	1

Exercise18



- فتح `do file` جديد وتخزينه باسم `mergem_1`
- فتح ملف `indiv_m` وهو ملف على مستوى افراد
- عمل `merge` مع ملف ال `hh_m` وهو ملف على مستوى أسر
- تخزين الملف المدمج باسم `merge_all1`

merge command



```
cd D:\Stata_Training_poverty\Data
```

```
*****1*****
use indiv_m,clear
br
*****2*****
merge m:1 hhid using hh_m
br,nol
*****3*****
keep if _merge==3
drop _merge
save merge_all1,replace
*****
```

Stata



- Collapse command
- Reshape command
- Do_file (Full Exercise)

Day 3

- Practical Exercises

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)
```

collapse command



Exercise19

► قم بفتح ملف **individual**

► انشئ ملف بيانات على مستوى الأسرة (**hhid**) يتضمن متغيرين :

متغير **female** يتضمن مجموع الاناث

يتضمن مجموع الذكور **male** متغير

```
use individual
```

```
gen male =1 if sex==1
```

```
gen female =1 if sex==2
```

```
collapse (sum) male female, by(hhid)
```

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)`

i(*i*) → Id variable(s)
j(*j*) → variable(s)

2) long reshape: from wide to long

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

i(*i*) → Id variable
j(*j*) → new variable

wide reshape command

(from long to wide)



long

Data Editor (Browse) - [hh_exp.dta]

File Edit View Data Tools

Saved

	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)



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)

Exercise20



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

```
cd D:\Stata_Training_poverty\Data
*****Reshape from long to wide
use income_yearly,clear
br
*****
reshape long faminc ,i(famid) j(year)
br
*****
```

Exercise21



► قم بفتح ملف inc_sav_yearly.dta وحول البيانات بشكل عرضي ومن ثم استعرض البيانات

```
*****Reshape from wide to long
```

```
use inc_sav_yearly,clear
```

```
br
```

```
reshape wide inc sav,i(hhid) j(year)
```

```
br
```

```
*****
```

Full_Excersies



١. تثبيت المسار على F:\Stata_Training\Full_Ex1 وعرض جميع محتوياته
٢. افتح ملف البيانات .mother.dta.
٣. انشئ وصف للملف “mothers file” (label).
٤. اظهر وصف كامل للملف.
٥. انشئ label للمتغير q100 وقيمه

Full_Excersies



١. اوجد عدد مشاهدات داخل الملف
٢. ادمج بيانات الملف مع بيانات ملف "mother2.dta"
٣. حزن الملف بعد الدمج "mother_all.dta"
٤. افتح ملف البيانات .child.dta
٥. استعرض label value الموجودة في الملف

Full_Excersies



١. رتب المتغير child_id بحيث يكون بعد المتغير q201
٢. اجمع متغير q205 و q206 بمتغير جديد باسم st_no

age_group	q202
1 "0-5"	0-5
2 "6-10"	6-10
3 "11-15"	11-15
4 "16-20"	16-20
5 "21+"	21-99

٣. اظهر اعلى قيمة لعمر الابناء
٤. انشئ متغير باسم age_group من متغير q202 كالتالي:

Full_Excersies



١. استخرج جدول تكراري للمتغيرات q205 و q206 مع اظهار القيم المفقودة
٢. انشئ ملف جديد لابناء الذين اعمارهم ١٥ سنة فاكثر باسم child_15more
٣. اعد فتح ملف الابناء ومن ثم انشئ ملف جديد يحتوي على عدد الابناء لكل ام وتخزينه باسم mother_nochild
٤. ادمج ملف mother_nochild مع ملف mother
٥. افتح ملف الابناء واستخرج جدول يبين متوسط عمر الابناء حسب المستوى التعليمي
٦. افتح ملف الامهات واستخرج جدول يبين التوزيع النسبي بين جنسية الام والمحافظة
٧. قم بفتح ملف الاطفال child1.dta و حول البيانات من شكلها الطولي الى العرضي

Stata



- Import/export data
- Bar /plot charts
- Map
- Practical Exercises

Day4



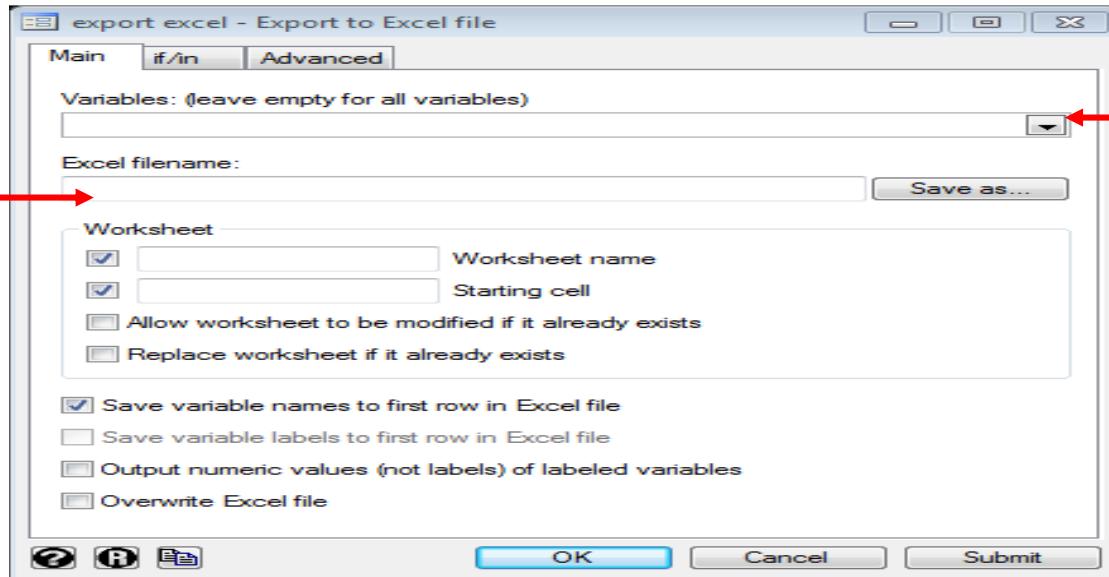
تصدير ملفات ستاتا الى ملف اكسل

✓ للتصدير ملف ستاتا الى اكسل (export) :

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



تصدير ملفات ستاتا الى ملف إكسل



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 لازريد جميع
المتغيرات.

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

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



تصدير ملفات ستاتا الى ملف إكسل

Exercise22

► قم بتصدير ملف wage.dta من مجلد stata_training الى wage_new.xlsx

```
use "F:\stata training\ wage.dta ", clear
```

```
export excel using "F:\Stata_Training\wage_new.xlsx", firstrow(variables) replace
```



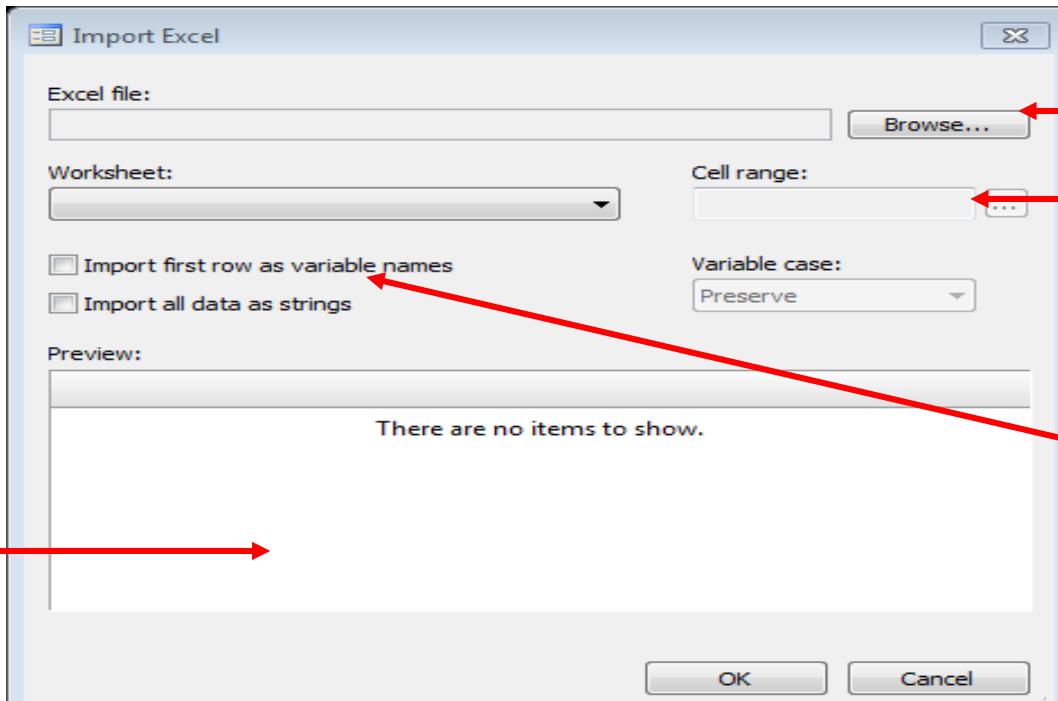
فتح او استيراد الملفات بصيغة الإكسل

✓ لاستيراد ملف اكسل الى Stata : (Import)

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



فتح او استيراد الملفات بصيغة الإكسل



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



فتح او استيراد الملفات بصيغة الاكسل

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



فتح او استيراد الملفات بصيغة الاكسل

Exercise23

► قم باستيراد ال sheet5 من ملف Wage من مجلد stata_training

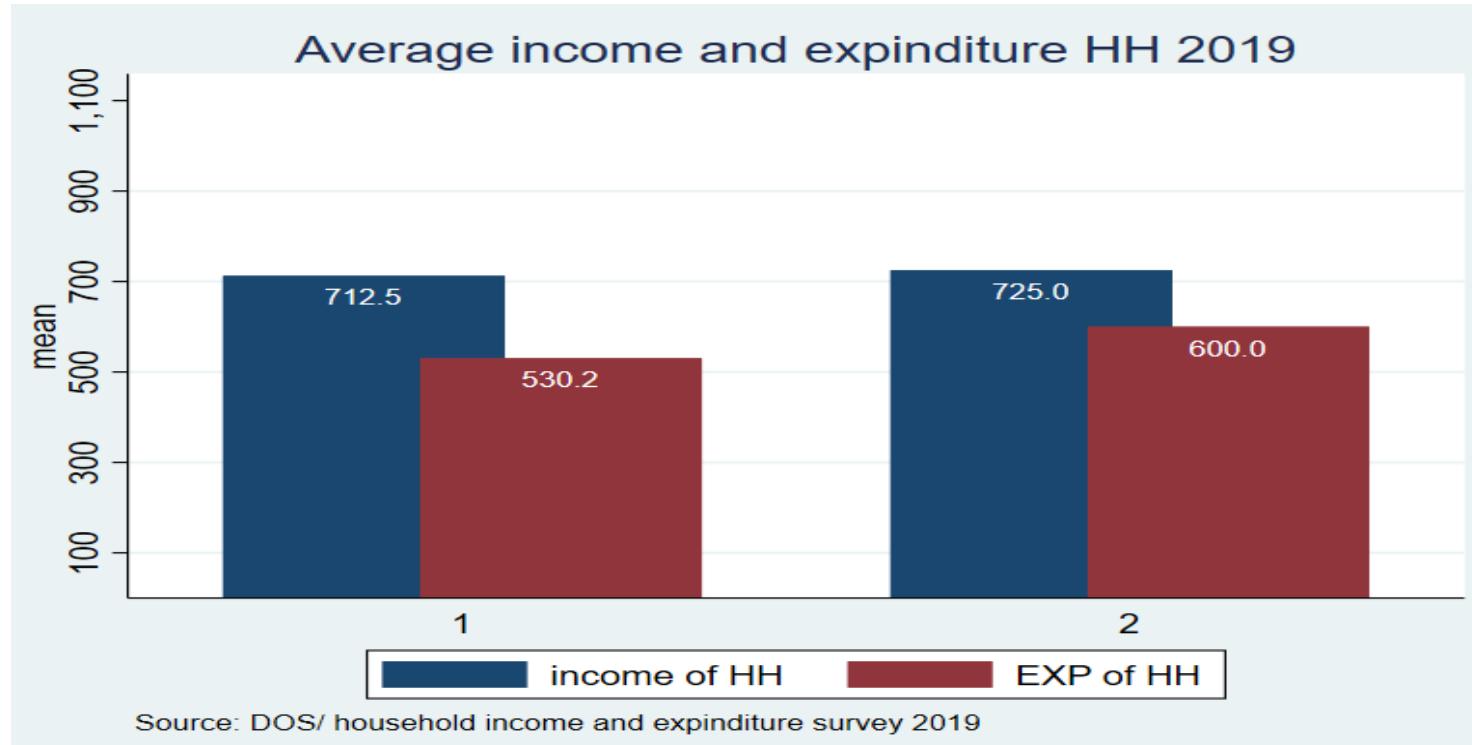
```
import excel "E:\stata_training\wage.xls", sheet("Sheet5") firstrow
```

Graph



- ✓ تستخدم الرسوم البيانية للتعبير عن العلاقة بين المتغيرات المختلفة أو لتوزيع مشاهدات العينة وفقاً لمتغير معين ويتم فتح نافذة مستقلة للرسم البياني.

Bar Chart



Bar Chart



```
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 expenditure HH 2019")  
note("Source: DOS/ household income and expenditure survey 2019")  
blabel(bar, position(inside) format(%9.1f) color(white))  
ylabel(100(200)1100);
```

Pie Chart



Expenditures, XYZ Comp.

2002



2003



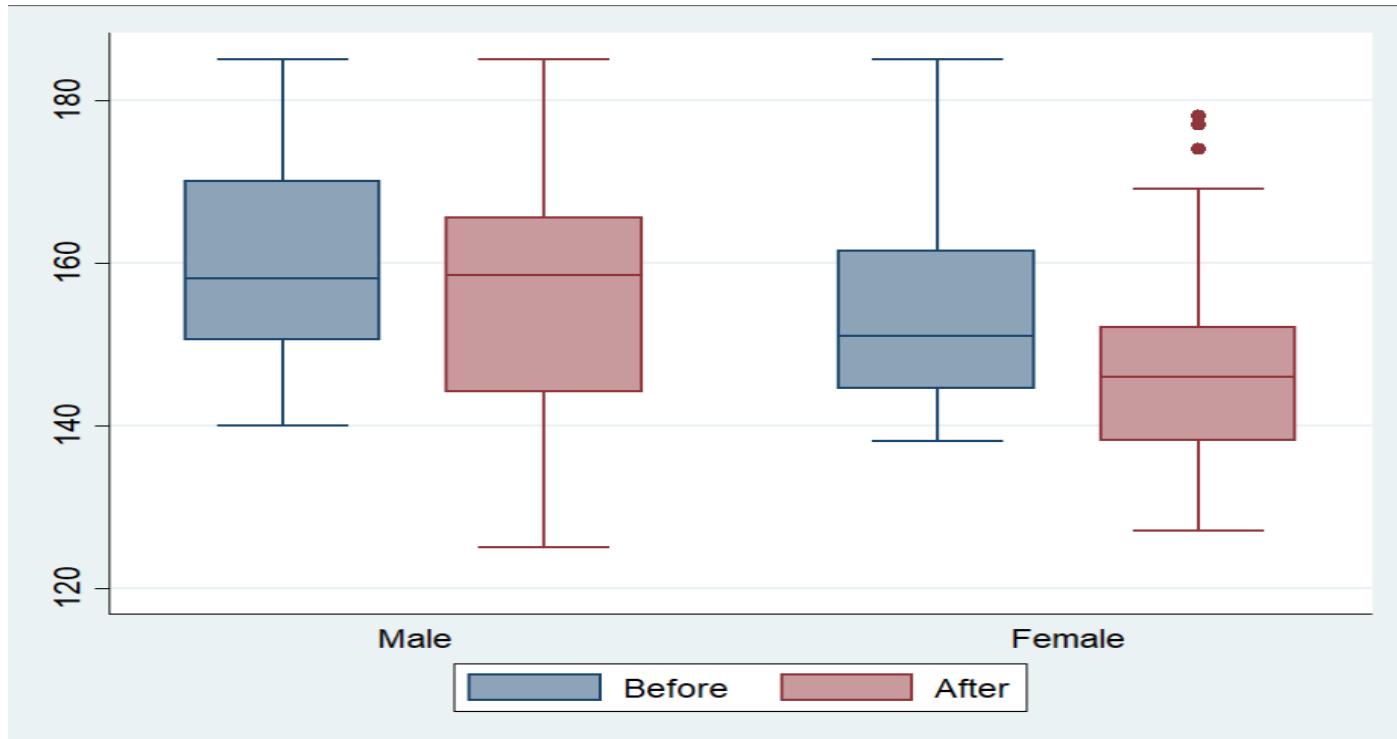
Source: 2002 Financial Report (sample data)

Pie Chart



```
use "company_balance.dta",clear  
br,nol  
#delim ;  
graph pie sales marketing research development,  
plabel(_all name, size(*1.5) color(white))  
by(year, legend(off) title("Expenditures, XYZ Comp.")  
note("Source: 2002 Financial Report ( sample data)") );
```

Box Chart

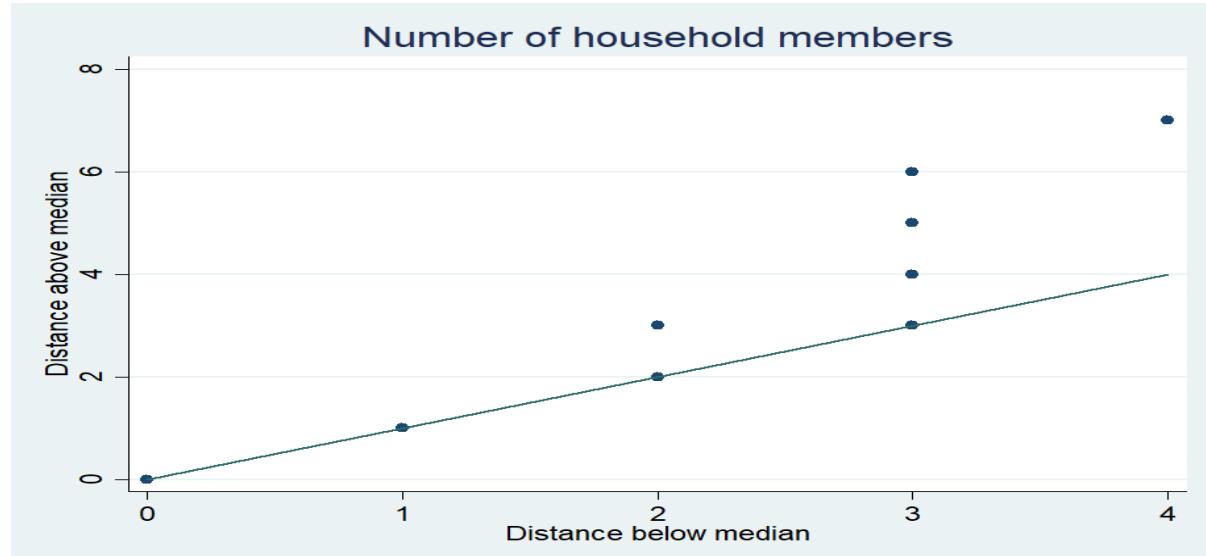


Box Chart



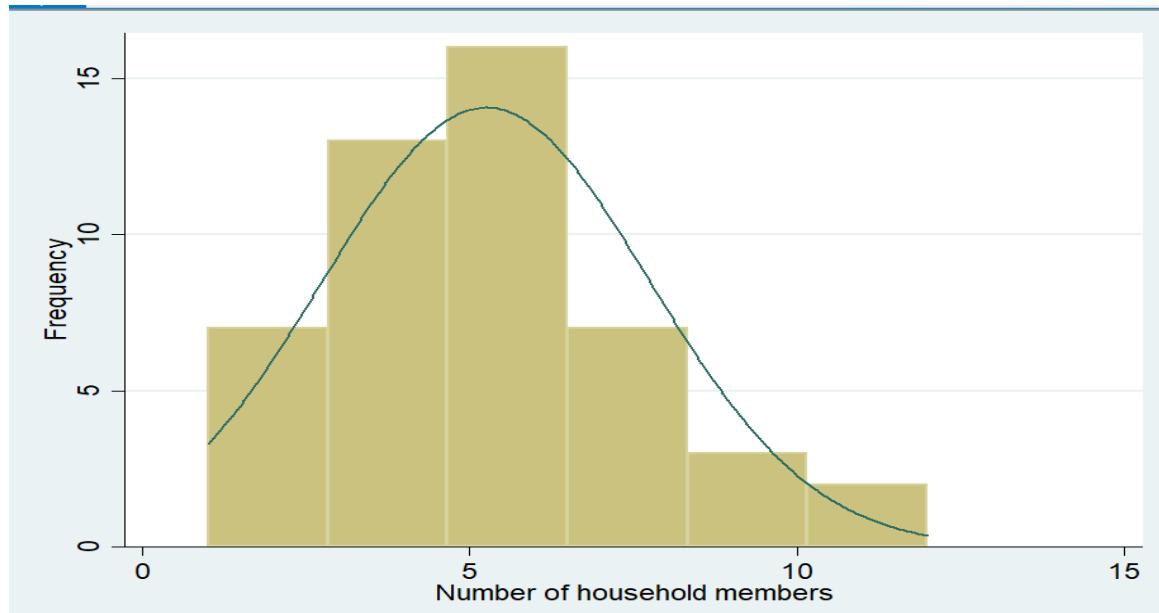
```
use "blood.dta",clear  
br,nol  
graph box bp_before bp_after, over(sex)
```

Plot Chart



```
use "hh.dta",clear  
symplot hh_size
```

Histogram Chart



```
use "hh.dta",clear  
histogram hh_size, freq normal
```



MAP

MAP



✓ خطوات تجهيز الخريطة:

DIVA-GIS

DATA

Data

١) يتم تحميل مجلد الخرائط الخاص بأي دولة عن طريق المسار التالي:

<http://www.diva-gis.org/gdata>

Select and download free geographic (GIS) data for any country in the world

Country

Jordan

Subject

Administrative areas

Download **JOR_adm.zip**

MAP



✓ خطوات تجهيز الخريطة:

- نقوم باختيار اسم الدولة من القائمة **country** ونختار نوع التوزيع من قائمة **(Administrative Areas) subject**
- نقوم بتنزيل الفولد **JOR_adm.zip** الموجود في نهاية الصفحة
- سيتم تنزيل الفولدر على **download** الخاص بجهازك ،فك الضغط للفولد وانسخ الملفات الموجودة بداخله على

F:\stata_training\Map

MAP



٢) يتم تزيل الامرین : **shp2dta,spmap**

```
ssc install shp2dta  
ssc install spmap
```

▪ يستخدم هذا الأمر لتحويل stata file الى shape file :**shp2dta**
▪ يستخدم هذا الأمر لرسم الخريطة :**spmap**

```
cd F:\stata_training\Map
```

٣) تثبيت ال directory على المجلد

MAP



٤) نقوم بتحويل ال **stata file** الى ملفين بصيغة **shape file**

```
shp2dta using JOR_adm1, data(JOR_gov_shp) coor(JOR_coorshp)
```

: اسم ملف ال **shape file** **JOR_adm1**

: اسم ملف البيانات بصيغة **Stata** والذي تم تحويله من ملف **shp_gov**

: اسم الملف الخاص بالحدود **X,Y** بصيغة **Stata** والذي تم تحويله من ملف **jor_coorshp**

MAP



٥) يتم تجميع الملف الخاص بحدود الخريطة وتجميعه على مستوى id و تغيير اسم id فيه

```
use JOR_coor_shp.dta,clear  
collapse (mean) _X _Y ,by (_ID)  
rename _ID id_1  
save JOR_coor_shp_id,replace
```

MAP



٦) نقوم بدمج ملف الحدود مع ملف البيانات:

```
merge 1:1 id_1 using JOR_gov_shp.dta  
drop _merge  
save JOR_gov_coor_shp.dta,replace
```

MAP



٧) يتم تجهيز ملف البيانات المراد عرضها على الخريطة على مستوى المحافظة

```
use data_show_map.dta , clear
```

٨) نقوم بدمج الملف اعلاه مع ملف stata shape file بصيغة ال

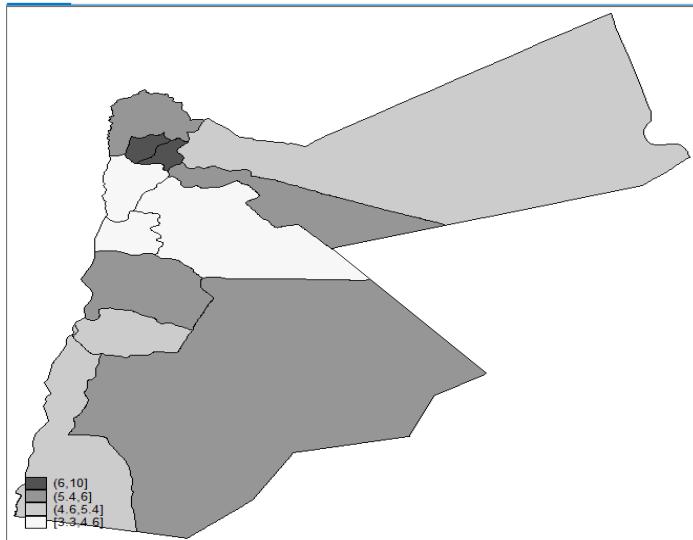
```
merge 1:1 id_1 using JOR_gov_coor_shp.dta  
drop _merge  
save "JOR_shp_gov_coor_var.dta",replace
```

MAP



٨) نقوم برسم الخريطة حسب المؤشر المراد عرض البيانات على اساسه

```
spmap hh_size using jor_coor_shap.dta, id(id_1)
```

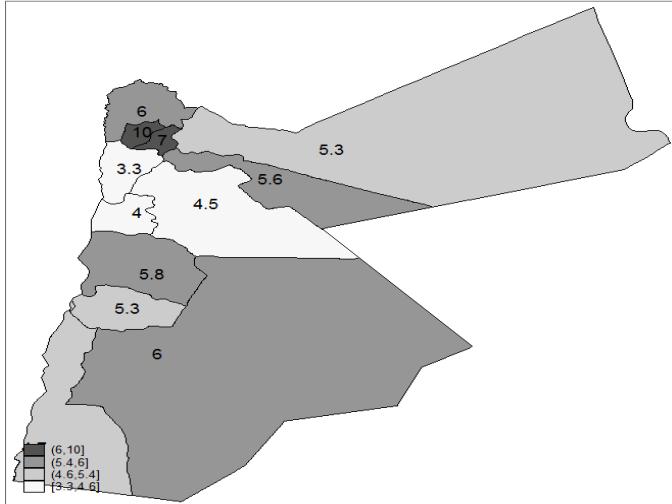


MAP



٩) عرض الوصف او قيم المؤشر موزعة حسب المحافظة على الخريطة

```
spmap hh_size using "jor_coor_shap.dta", id(id_1) label(xcoord(_X) ycoord(_Y) label(hh_size))
```



MAP



✓ لتغيير الالوان وبعض الخصائص على الخريطة :

نقوم باختيار زر Graph Editor الموجود في اعلى صفحة الخريطة وتعديل الالوان او وضع عنوان او اي خصائص
نريدها

