

```

* Exercise 1 & 2

* how to set mem in STATA?
set mem 100m

* setting working directory in STATA
cd"C:\Users\dtrifunovic\Desktop\STATA\Dragana_vezbe"

*uploading data from csv in STATA
insheet using "C:\Users\dtrifunovic\Desktop\STATA\Data\data.csv"

*****
*****
* Generating practice in STATA
gen var1=1

* change var1 in zero if person is selfemployed.
replace var1=0 if selfemp==2

* labeling variables in STATA for graphs and tables
label var var1 "Pass Final Exam"

* labeling variables in Indian data
label var educat "Education Level"
label var totwage "Wage per week"
label var exper "Experience"
label var marst "Marital Status"

* value names for numbers 0 "Fail" and 1 "Pass"
label def dummy 1 "Pass" 0 "Fail"
label val var1 dummy

* how to drop variables in STATA?
drop var1

* introduction to data / summarizing data
sum totwage exper educat sex marst

* getting to know data a bit better
sum totwage exper educat sex marst, detail

* getting to know data even better :)
tab totwage
tab educat
tab marst

* Graphs in STATA
tway histogram totwage, bin(15)
tway histogram totwage, bin(55)
tway scatter totwage educat

* graphs with labeled variables
tway scatter var1 educat

```

```

* graphs of wage with independent variables

* 1. wage and education
tway scatter totwage educat

* 2. wage and experience
tway scatter totwage exper

* 3. wage and sex
tway scatter totwage sex

* 4. wage and marital status
tway scatter totwage marst

* 5. age and experience
tway scatter age exper

**** Exercise 3

* Dropping extremes

* Wage per week
drop if totwage>8000

* inspecting variable experience
tway histogram exper, bin(55)

* dropping extreme values of experience
drop if exper>65

* examining correlations
** wage and experience
corr totwage exper

** wage and education
pccorr totwage educat, sig

* tab with two variables
tab totwage
tab marst

* tab with two variables
tab marst educat
tab marst sex
tab sex

***** Exercise 4

*** When starting stata always set the memory, working directory and
**** withdraw the data (from beginning of the file)

* repeating the graphs
* 2. wage and experience
tway scatter totwage exper

```

```

** repeating exclusions of extremes
drop if totwage>8000

** repeat tab command
tab marst
tab marst sex

** repeat correlation between variables
pwcorr totwage educat, sig

** when we need more explanation about the command we use help
help drop

*** how to do correlation in STATA?
help correlation

*** Wage model

** choosing variables
*** totwage educat exper marst sex age

* for start we summarize
sum totwage educat exper marst sex age

** econometric model
*** regression analysis I
reg totwage educat exper marst sex age

*** corrections of the wage model

* 1) change the wage to log wage
gen logtotwage=log(totwage)
sum logtotwage

* 2) inspect age and experience
* 5. age and experience
twoway scatter age exper
*** as a result of perfect correlation we exclude age from the model.

*** regression analysis II
reg logtotwage educat exper marst sex

* quadratic variables
*** how experience affects wage
twoway scatter totwage exper

*** regression analysis III
**** including experience square - concave shape

* creating exper square
gen exper2=exper^2

* regression with new variable exper2
reg logtotwage educat exper exper2 marst sex

```

```

***** Exercise 5

*** after set memory, directory and uploading data, we gen logs and drop
extremes
gen logtotwage=log(totwage)
gen exper2=exper^2
drop if totwage>8000

*** working with dummy variables

**** generate - rename : male dummy
rename sex male

*** generate female dummy
gen female = male == 0

*** generate married dummy
gen married = marst == 1

*** regression analysis IV
reg logtotwage educat exper exper2 male married
reg logtotwage educat exper exper2 female married

*** standard errors and concept of homoscedasticity
reg logtotwage educat exper exper2 male married, vce (robust)

reg logtotwage educat exper exper2 male married, cluster (state)

*** comparing results for different groups
rename male sex
**** two tables in STATA by gender
sort sex
by sex: reg logtotwage educat exper exper2 married, cluster (state)

*** storing tables and printing tables

**** command quietly runs regression without output
quietly reg logtotwage educat exper exper2 married if sex==1
*** saving the results
estimates store men

quietly reg logtotwage educat exper exper2 married if sex==0
*** saving the results
estimates store women

*** put the results in tables
estimates table men women

*** put the results in tables with standard errors
estimates table men women, b(%5.3f) se(%5.3f) t(%5.3f)

*** how to copy tables in word
**** copy paste text from STATA to Word using Courier 8

*** saving the list of independent variables
local controls `educat exper exper2 married'

```

```

reg logtotwage `controls'

*** PREDICTED VALUES
quietly reg logtotwage educat exper exper2 married sex, cluster (state)
predict yhat
predict resid, residuals

**** plot predicted values and residuals
twoway scatter yhat resid

*** Exercise 6
* probability of being married

** linear probability model
reg married educat exper exper2 totwage

*** saving the results
estimates store LPM

** Probability model
probit married educat exper exper2 totwage
*** saving the results
estimates store PROBIT

** Logit model
logit married educat exper exper2 totwage
*** saving the results
estimates store LOGIT

estimates table LPM PROBIT LOGIT, b(%5.3f) se(%5.3f)

*** marginal effects are not given in LPM PROBIT or LOGIT
**** interpretation of parameters beta can not be as for OLS.

**** to get marginal effects we use a special command, first the model
probit married educat exper exper2 totwage
**** second, command for marginal effects
mfx

*** to check the marginal values in OLS -- using LPM (Linear model!)
reg married educat exper exper2 totwage
mfx

*** How to capture state specific effects on getting married? (by using
cluster)
probit married educat exper exper2 totwage, cluster (state)
mfx
**** the effect of education does not differ across states (our marginal
effects
**** do not differ)

*** How to estimate prob of married by gender using Probit?
**** two tables in STATA by gender (by male since we renamed the var.sex)
rename sex male
sort male
by male: probit married educat exper exper2 totwage, cluster (state)

```

```
*** two comparable tables by gender
```

```
quietly probit married educat exper exper2 totwage if male==1  
estimates store MEN
```

```
quietly probit married educat exper exper2 totwage if male==0  
estimates store WOMEN
```

```
estimates table MEN WOMEN, b(%5.3f) se(%5.3f)
```