HOMEWORK 1

SP6015 Analisis Kuantitatif Untuk Kebijakan Publik

Maghfira Ramadhani (20021140)

Institut Teknologi Bandung

Daftar Isi

HOMEWORK 1

SP6015 Analisis Kuantitatif Untuk Kebijakan Publik

Nomor 1

Command:

```
*******************************************************************************
* COURSES: SP6015 Quantitative Method for Policy Analysis
* PROJECT: Homework 1
* SOURCE OF THE RAW DATA: ps1_psid2003.dta
* AUTHORS: Maghfira Ramadhani - 20021140
* DATE: February 2022
* STATA VERSION: Stata/SE 16.1 for Mac (Revision 19 Nov 2020)
*******************************************************************************

* 1 Create do-file
```

Nomor 2

Command:

```
* 2 Create version control and pause
version 14.2
set more off, permanently
capture log close
capture graph drop _all
```

Nomor 3

Command:

```
* 3 Set directory and use dataset (locate the directory of HW 1 folder)
cd "/Users/macbook/Documents/Work/SP6015/HW 1"
```

Nomor 4

Command:

```
* 4 Define local macro
local input_data "./data/ps1_psid2003.dta"
local output_data "./data/ps1_psid2003_edited.dta"
```

Nomor 5

Command:

```
* 5 Load data
use `input_data', clear
```

Nomor 6

Command:

```
* 6,7 Create new variable
generate totalhours = hours * weeks
generate wagerate = salary/totalhours
```

Nomor 7

Lihat nomor 6.

Nomor 8

Command:

```
* 8,9 Create categorical variable
gen fulltime=0
replace fulltime=1 if weeks>=48 & hours>=35

gen female=1
replace female=0 if sex==1
```

Nomor 9

Lihat nomor 8

Nomor 10

Command:

```
* 10 Create log wage
generate logwage = log(wagerate)
```

Nomor 11

Command:

```
* 11 Create label
label variable totalhours "hours worked per week multiplied by number of weeks worked"
label variable wagerate "salary divided by the total hours worked"
label variable logwage "the logarithmic value of the wage"
```

```
label variable fulltime "type of worker, fulltime or part time"
label define fulltime 0 "part time worker" 1 "fulltime worker"
label variable female "=1 if female worker"
label define female 0 "male" 1 "female"
```

Nomor 12

Command:

```
* 12 Create label values
label values fulltime fulltime
label values female female
```

Nomor 13

Command:

```
* 13 Produce summary table
outreg2 using "output/tables/table1.doc", replace sum(log) ///
keep(age educ weeks hours salary wagerate logwage)
```

Output:

table1.doc

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
age	510	27.14	11.29	18	88
educ	510	12.58	2.822	0	17
weeks	510	39.49	16.24	1	52
hours	510	31.70	12.59	3	72
salary	510	15,068	15,455	50	90,000
wagerate	510	13.43	26.24	0.0397	333.3
logwage	510	2.197	0.788	-3.227	5.809

Nomor 14

Command:

```
* 14 Produce histogram
histogram salary
graph export "output/figures/figure_1_histogram_salary.png", replace
histogram logwage
```

```
graph export "output/figures/figure_2_histogram_logwage.png", replace
```

Output:

figure_1_histogram_salary.png

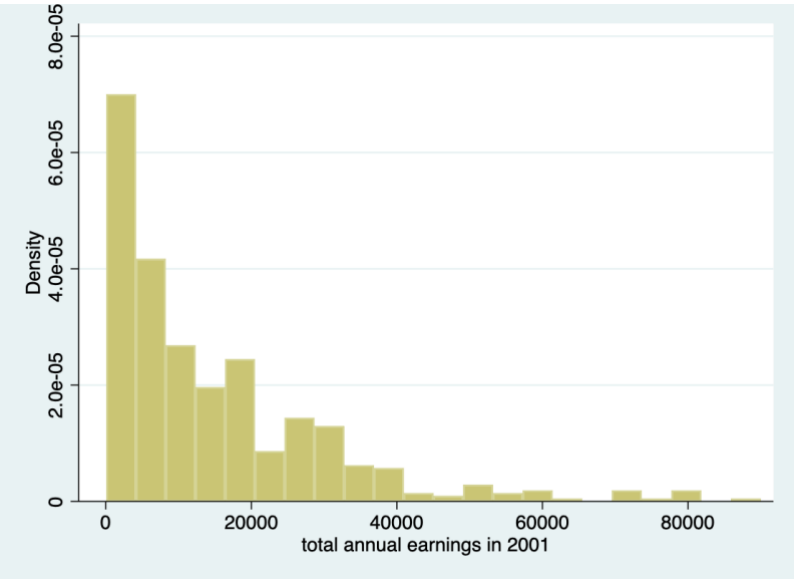
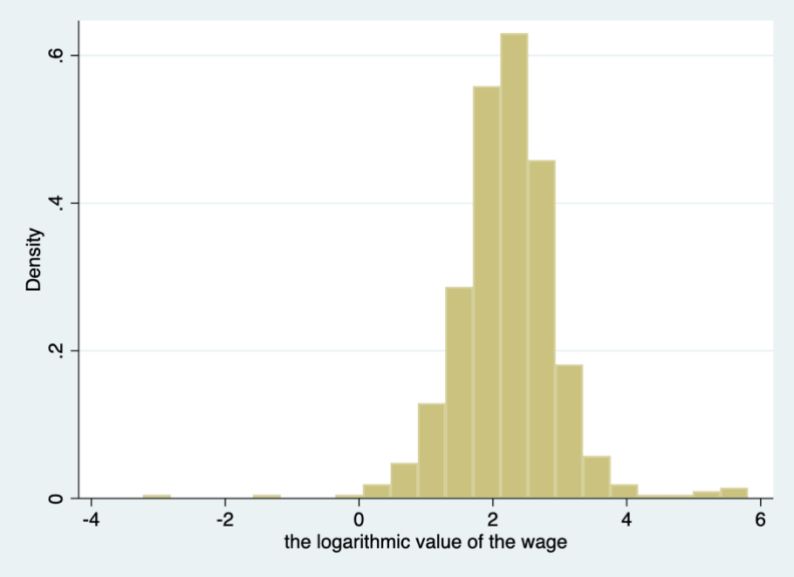


figure_1_histogram_logwage.png



Pada histogram plot variabel “salary” distribusi cenderung tidak mengikuti distribusi normal dengan kecenderungan density dari data semakin tinggi menuju nilai salary minimum.

Setelah dilakukan transformasi menjadi variable “logwage” terjadi perubahan bentuk distribusi menjadi mengikuti distribusi normal.

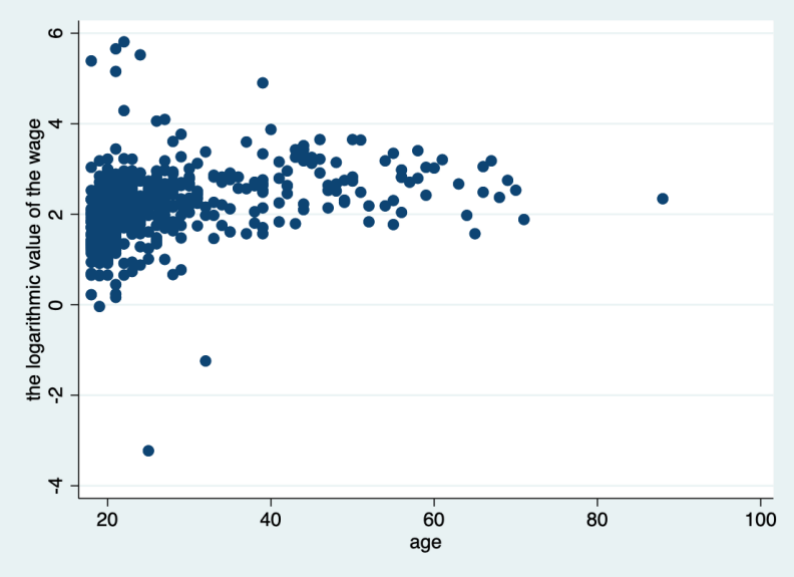
Nomor 15

Command:

```
* 15 Produce scatter plot
twoway (scatter logwage age, lcolor(emidblue) lwidth(medthick))
graph export "output/figures/figure_3_scatter_age_logwage.png", replace
```

Output:

figure_3_scatter_age_logwage.png



Pada scatter plot antara variabel “age” dan “logwage” secara sekilas teramati hubungan yang linear dengan data yang memiliki usia kurang dari 40 memiliki sebaran nilai logwage yang lebih besar.

Nomor 16

Command:

```
* 16 Produce twoway table
asdoc tabulate female fulltime, save(table2.doc)
copy table2.doc "output/tables/table2.doc", replace
erase "table2.doc"
```

Output:

table2.doc

Tabulation of female fulltime

=1 if female worker	type of worker, fulltime or part time		
	part time worker	fulltime worker	Total
male	151	117	268
female	154	88	242
Total	305	205	510

Berdasarkan hasil tabulasi silang secara umum pada sampel data, partisipasi gender pria dalam dunia kerja lebih besar dibandingkan wanita (selisih 26 orang). Jika dilihat lebih detail berdasarkan tipe pekerjaan, untuk tipe pekerjaan part time angka partisipasi kerja antar gender pria dan wanita relatif sama dengan angka partisipasi pria 151 dan wanita 154. Untuk jenis pekerjaan full time terdapat perbedaan yang cukup signifikan dimana sebanyak 117 pria bekerja secara fulltime sedangkan hanya sebanyak 88 wanita yang bekerja fulltime.

Nomor 17

Command:

```
* 17 Save edited data
save `output_data', replace
```

Output:

Nomor 18

Command:

```
* 18 Create regression local macro
local depvar logwage
local indepvar1 female
local indepvar2 female age
local indepvar3 female age educ
local indepvar4 female age educ fulltime

* 18 Produce regression table1 with robust
reg `depvar' `indepvar1', robust
outreg2 using "output/tables/table3.doc", replace ctitle(Model 1)
reg `depvar' `indepvar2', robust
outreg2 using "output/tables/table3.doc", append ctitle(Model 2)
reg `depvar' `indepvar3', robust
```

```
outreg2 using "output/tables/table3.doc", append ctitle(Model 3)
reg `depvar' `indepvar4', robust
outreg2 using "output/tables/table3.doc", append ctitle(Model 4)
```

Output:

table3.doc

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
female	-0.109	-0.170**	-0.176***	-0.162**
	(0.0697)	(0.0671)	(0.0668)	(0.0678)
age		0.0214***	0.0216***	0.0200***
		(0.00267)	(0.00252)	(0.00263)
educ			0.0354***	0.0346***
			(0.0115)	(0.0115)
fulltime				0.113*
				(0.0657)
Constant	2.249***	1.697***	1.249***	1.252***
	(0.0486)	(0.0848)	(0.168)	(0.169)
Observations	510	510	510	510
R-squared	0.005	0.097	0.113	0.118

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Nomor 19

Command:

```
* 19 Produce regression table3 with conventional standard error
reg `depvar' `indepvar1'
outreg2 using "output/tables/table4.doc", replace ctitle(Model 1)
reg `depvar' `indepvar2'
outreg2 using "output/tables/table4.doc", append ctitle(Model 2)
reg `depvar' `indepvar3'
outreg2 using "output/tables/table4.doc", append ctitle(Model 3)
reg `depvar' `indepvar4'
outreg2 using "output/tables/table4.doc", append ctitle(Model 4)
```

Output:

Table4.doc

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
female	-0.109	-0.170**	-0.176***	-0.162**

	(0.0698)	(0.0670)	(0.0665)	(0.0670)
age		0.0214***	0.0216***	0.0200***
		(0.00297)	(0.00295)	(0.00313)
educ			0.0354***	0.0346***
			(0.0117)	(0.0117)
fulltime				0.113
				(0.0716)
Constant	2.249***	1.697***	1.249***	1.252***
	(0.0481)	(0.0892)	(0.173)	(0.172)
Observations	510	510	510	510
R-squared	0.005	0.097	0.113	0.118

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

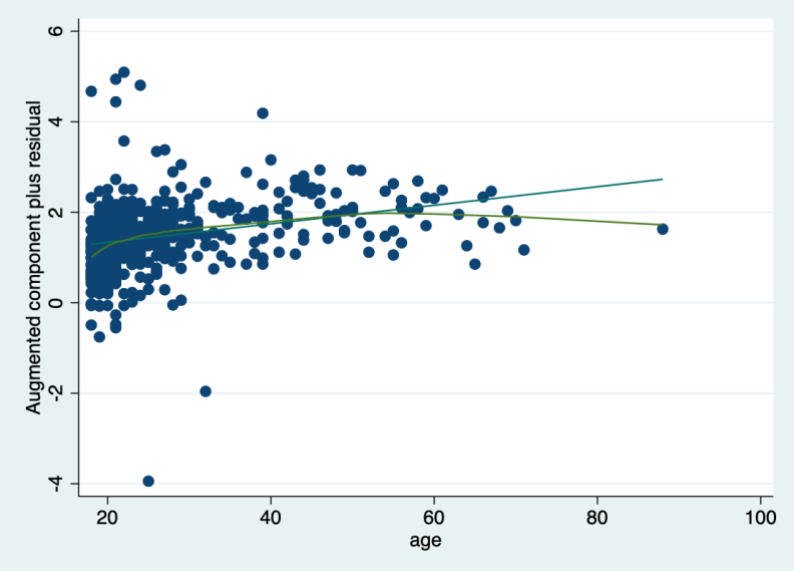
Nomor 20

Command:

```
* 20 Linearity diagnostics with acprplot on logwage and age
quietly reg logwage age
acprplot age, lowess
graph export "output/figures/figure_4_acprplot_age.png", replace
```

Output:

figure_4_acprplot_age.png



Pada kurva acprplot terlihat kedua garis antara garis linear dan garis hasil smoothing cukup berhimpit sehingga secara kualitatif dapat dikatakan tren variable “logwage” ini linear terhadap variable “age”.

Nomor 21

Command:

```
* 21 Multicolinearity diagnostics with
* (1) Pairwise correlation matrix
pwcorr `depvar' `indepvar4', star(0.05) sig

* (2) Correlation matrix graph
graph matrix `depvar' `indepvar4', half
graph export "output/figures/figure_5_correlation_matrix_graph.png", replace

* (3) Variance Inflation Factor (VIF)
quietly reg `depvar' `indepvar4'
vif
```

Output:

Log file

	logwage	female	age	educ	fulltime
logwage	1.0000				
female	-0.0692	1.0000			
	0.1188				
age	0.2933*	0.1262*	1.0000		
	0.0000	0.0043			
educ	0.1171*	0.0244	-0.0219	1.0000	
	0.0081	0.5830	0.6218		
fulltime	0.1754*	-0.0743	0.3275*	0.0281	1.0000
	0.0001	0.0938	0.0000	0.5263	

figure_5_correlation_matrix_graph.png

Log file

```
. * (3) Variance Inflation Factor (VIF)
. quietly reg `depvar' `indepvar4'

. vif

    Variable |       VIF       1/VIF
-------------+----------------------
         age |      1.15    0.868754
    fulltime |      1.14    0.877697
      female |      1.03    0.968112
        educ |      1.00    0.997089
-------------+----------------------
    Mean VIF |      1.08
```

Pada pairwise correlation matrix tidak ada koefisien korelasi yang memiliki nilai diatas 0.5 maka dapat dikatakan tidak ada variable yang memiliki hubungan kolinearitas secara signifikan.

Pada correlation matrix graph juga tidak teramati variable yang terlihat memiliki trend linear sempurna sehingga dapat dikatakan tidak terdapat multikolineritas.

Berdasarkan uji multikolinearitas menggunakan Variance Inflation Factor (VIF) menunjukkan bahwa keseluruhan variabel memiliki nilai VIF<10 sehingga dapat disimpulkan tidak terdapat masalah multikolinearitas

Nomor 22

Command:

```
* 22 Homoscedasticity diagnostics with
* (1) Breusch-Pagan test
quietly reg `depvar' `indepvar4'
estat hettest

* (2) Residual vs Fitted plot
rvfplot, yline(0)
graph export "output/figures/figure_6_residual_fitted_plot.png", replace
```

Output:

Log file

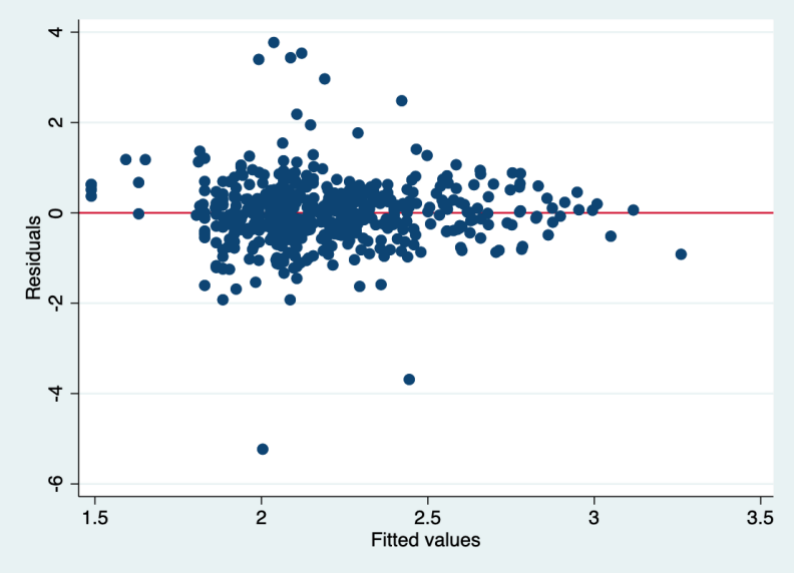
```
. * 22 Homoscedasticity diagnostics with
. * (1) Breusch-Pagan test
. quietly reg `depvar' `indepvar4'

. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of logwage

         chi2(1)      =    12.08
         Prob > chi2  =   0.0005
```

figure_6_residual_fitted_plot.png



Hasil dari Breush-Pagan test menunjukkan angka probability Chi-Square sebesar 0.0005. Dengan nilai p-value yang lebih kecil dari 0.01 (tingkat kepercayaan 99%), 0.05 (tingkat kepercayaan 95%), maupun 0.1 (tingkat kepercaaan 90%) maka null hypothesis ditolak, artinya terdapat cukup bukti untuk menerika alternative hypothesis bahwa variance error tidak konstan atau terdapat heteroskedasticity pada kasus ini

Berdasarkan residual versus fitted plot teramati bahwa variance error memiliki nilai yang tidak konstan dan cenderung lebih besar pada fitted values 2 hingga 2.5. Hal ini menandakan terdapat heteroskedasticity. Pada kasus ini asumsi OLS terkait homoskedasticity tidak terpenuhi.

Nomor 23

Command:

```
* 23 Omitted variable bias diagnostics with RESET
quietly reg `depvar' `indepvar4'
ovtest
```

Output:

Log file

```
. * 23 Omitted variable bias diagnostics with RESET
. quietly reg `depvar' `indepvar4'

. ovtest

Ramsey RESET test using powers of the fitted values of logwage
       Ho:  model has no omitted variables
                 F(3, 502) =      2.30
                  Prob > F =      0.0767
```

Dengan menggunakan Ramsey RESET test didapatkan p-value F-distribution sebesar 0.0767. Nilai P-value yang lebih besar dari 0.01 (tingkat kepercayaan 99%), 0.05 (tingkat kepercayaan 95%) maka null hypothesis tidak dapat ditolak pada tingkat kepercayaan tersebut. Artinya terdapat cukup bukti untuk menerima null hypothesis bahwa model tidak mengandung omittes variables.

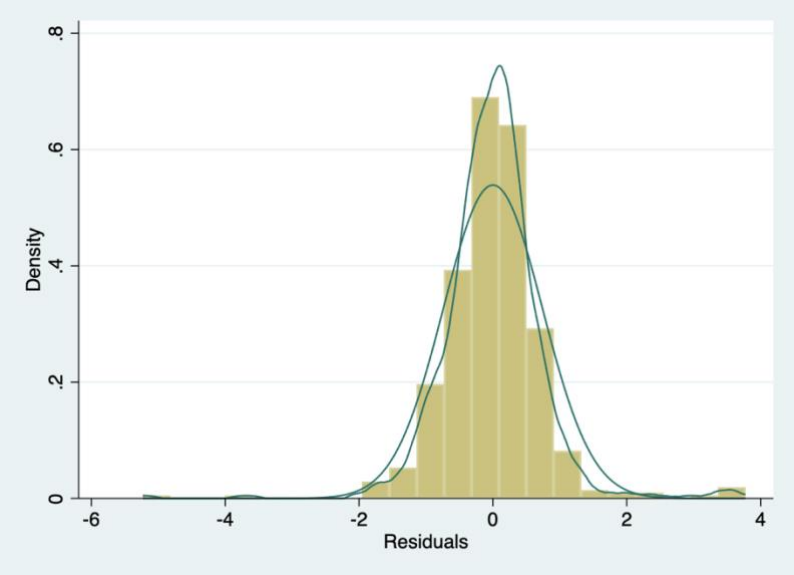
Nomor 24

Command:

```
* 24 Error normality diagnostics
quietly reg `depvar' `indepvar4', robust
predict e, resid
hist e, kdensity normal
graph export "output/figures/figure_7_error_normality_plot.png", replace
```

Output:

figure_7_error_normality_plot.png



Dengan melihat pada error normality plot dapat dilihat bahwa secara kualitatif error mengikuti karakteristik dari distribusi normal.

Lampiran Screenshot Direktori, Do-File, dan Log-File

code	Today at 10:27 PM	--	Folder
run_code	Today at 10:26 PM	5 KB	Stata Do-file
run_code	Today at 10:27 PM	59 KB	PDF Document
data	Today at 8:09 PM	--	Folder
ps1_psid2003_edited.dta	Today at 10:27 PM	27 KB	Stata Data File
ps1_psid2003.dta	Nov 25, 2021 at 12:37 PM	9 KB	Stata Data File
HW1_20021140_Maghfira_Ramadhani.docx	Today at 10:26 PM	869 KB	Micros...(.docx)
HW1-2022.pdf	Feb 5, 2022 at 7:47 AM	175 KB	PDF Document
output	Today at 10:27 PM	--	Folder
figures	Today at 10:27 PM	--	Folder
figure_1_histogram_salary.png	Today at 10:26 PM	71 KB	PNG image
figure_2_histogram_logwage.png	Today at 10:26 PM	55 KB	PNG image
figure_3_scatter_age_logwage.png	Today at 10:27 PM	114 KB	PNG image
figure_4_acprplot_age.png	Today at 10:27 PM	126 KB	PNG image
figure_5_correlation_matrix_graph.png	Today at 10:27 PM	139 KB	PNG image
figure_6_residual_fitted_plot.png	Today at 10:27 PM	65 KB	PNG image
figure_7_error_normality_plot.png	Today at 10:27 PM	54 KB	PNG image
logs	Today at 10:27 PM	--	Folder
HW1_16_Feb_2022_20_14_02.log	Today at 8:14 PM	17 KB	Log File
HW1_16_Feb_2022_20_56_58.log	Today at 8:57 PM	15 KB	Log File
HW1_16_Feb_2022_20_58_06.log	Today at 8:58 PM	17 KB	Log File
HW1_16_Feb_2022_21_51_12.log	Today at 9:51 PM	17 KB	Log File
HW1_16_Feb_2022_22_26_56.log	Today at 10:27 PM	17 KB	Log File
tables	Today at 8:58 PM	--	Folder
table1.doc	Today at 10:26 PM	2 KB	Micros...t (.doc)
table1.txt	Today at 10:26 PM	277 bytes	Plain Text
table2.doc	Today at 10:27 PM	2 KB	Micros...t (.doc)
table3.doc	Today at 10:27 PM	2 KB	Micros...t (.doc)
table3.txt	Today at 10:27 PM	515 bytes	Plain Text
table4.doc	Today at 10:27 PM	2 KB	Micros...t (.doc)
table4.txt	Today at 10:27 PM	507 bytes	Plain Text

```
1  ********************************************************************
   *****
2  * COURSES: SP6015 Quantitative Method for Policy Analysis
3  * PROJECT: Homework 1
4  * SOURCE OF THE RAW DATA: ps1_psid2003.dta
5  * AUTHORS: Maghfira Ramadhani - 20021140
6  * DATE: February 2022
7  * STATA VERSION: Stata/SE 16.1 for Mac (Revision 19 Nov 2020)
8  ********************************************************************
   *****
9
10 * 1 Create do-file
11
12 * 2 Create version control and pause
13 version 14.2
14 set more off, permanently
15 capture log close
16 capture graph drop _all
17
18 * Create log file
19 local c_time_date = "`c(current_date)'"+"_" +"`c(current_time)'"
20 local time_string = subinstr("`c_time_date'", ":", "_", .)
21 local time_string = subinstr("`time_string'", " ", "_", .)
22 log using "./output/logs/HW1_`time_string'.log", text
23
24 * 3 Set directory and use dataset (locate the directory of HW 1
   folder)
25 cd "/Users/macbook/Documents/Work/SP6015/HW 1"
26
27 * 4 Define local macro
28 local input_data "./data/ps1_psid2003.dta"
29 local output_data "./data/ps1_psid2003_edited.dta"
30
31 * 5 Load data
32 use `input_data', clear
33
34 * 6,7 Create new variable
35 generate totalhours = hours * weeks
36 generate wagerate = salary/totalhours
37
38 * 8,9 Create categorical variable
39 gen fulltime=0
40 replace fulltime=1 if weeks>=48 & hours>=35
41
42 gen female=1
```

```
gen female=1
replace female=0 if sex==1

* 10 Create log wage
generate logwage = log(wagerate)

* 11 Create label
label variable totalhours "hours worked per week multiplied by
number of weeks worked"
label variable wagerate "salary divided by the total hours worked"
label variable logwage "the logarithmic value of the wage"
label variable fulltime "type of worker, fulltime or part time"
label define fulltime 0 "part time worker" 1 "fulltime worker"
label variable female "=1 if female worker"
label define female 0 "male" 1 "female"

* 12 Create label values
label values fulltime fulltime
label values female female

* 13 Produce summary table1
outreg2 using "output/tables/table1.doc", replace sum(log) ///
keep(age educ weeks hours salary wagerate logwage)

* 14 Produce histogram
histogram salary
graph export "output/figures/figure_1_histogram_salary.png", replace
histogram logwage
graph export "output/figures/figure_2_histogram_logwage.png",
replace

* 15 Produce scatter plot
twoway (scatter logwage age, lcolor(emidblue) lwidth(medthick))
graph export "output/figures/figure_3_scatter_age_logwage.png",
replace

* 16 Produce twoway table2
asdoc tabulate female fulltime, save(table0.doc)
copy table0.doc "output/tables/table2.doc", replace
erase "table0.doc"

* 17 Save edited data
save `output_data', replace
```

```
save `output_data', replace

* 18 Create regression local macro
local depvar logwage
local indepvar1 female
local indepvar2 female age
local indepvar3 female age educ
local indepvar4 female age educ fulltime

* 18 Produce regression table3 with robust
reg `depvar' `indepvar1', robust
outreg2 using "output/tables/table3.doc", replace ctitle(Model 1)
reg `depvar' `indepvar2', robust
outreg2 using "output/tables/table3.doc", append ctitle(Model 2)
reg `depvar' `indepvar3', robust
outreg2 using "output/tables/table3.doc", append ctitle(Model 3)
reg `depvar' `indepvar4', robust
outreg2 using "output/tables/table3.doc", append ctitle(Model 4)

* 19 Produce regression table3 with conventional standard error
reg `depvar' `indepvar1'
outreg2 using "output/tables/table4.doc", replace ctitle(Model 1)
reg `depvar' `indepvar2'
outreg2 using "output/tables/table4.doc", append ctitle(Model 2)
reg `depvar' `indepvar3'
outreg2 using "output/tables/table4.doc", append ctitle(Model 3)
reg `depvar' `indepvar4'
outreg2 using "output/tables/table4.doc", append ctitle(Model 4)

* 20 Linearity diagnostics with acprplot on logwage and age
quietly reg logwage age
acprplot age, lowess
graph export "output/figures/figure_4_acprplot_age.png", replace

* 21 Multicolinearity diagnostics with
* (1) Pairwise correlation matrix
pwcorr `depvar' `indepvar4', star(0.05) sig
```

```
pwcorr `depvar' `indepvar4', star(0.05) sig

* (2) Correlation matrix graph
graph matrix `depvar' `indepvar4', half
graph export "output/figures/figure_5_correlation_matrix_graph.png"
, replace

* (3) Variance Inflation Factor (VIF)
quietly reg `depvar' `indepvar4'
vif

* 22 Homoscedasticity diagnostics with
* (1) Breusch-Pagan test
quietly reg `depvar' `indepvar4'
estat hettest

* (2) Residual vs Fitted plot
rvfplot, yline(0)
graph export "output/figures/figure_6_residual_fitted_plot.png",
replace

* 23 Omitted variable bias diagnostics with RESET
quietly reg `depvar' `indepvar4'
ovtest

* 24 Error normality diagnostics
quietly reg `depvar' `indepvar4', robust
predict e, resid
hist e, kdensity normal
graph export "output/figures/figure_7_error_normality_plot.png",
replace

log close

clear

exit

```

```
      ____________________________________________________________________________________
      > --------------
            name:  <unnamed>
             log:  /Users/macbook/Documents/Work/SP6015/HW 1/output/logs/HW1_16_Feb_2022_22_2
      > 6_56.log
        log type:  text
       opened on:  16 Feb 2022, 22:26:56

  1 .
  2 . * 3 Set directory and use dataset (locate the directory of HW 1 folder)
  3 . cd "/Users/macbook/Documents/Work/SP6015/HW 1"
      /Users/macbook/Documents/Work/SP6015/HW 1

  4 .
  5 . * 4 Define local macro
  6 . local input_data "./data/ps1_psid2003.dta"

  7 . local output_data "./data/ps1_psid2003_edited.dta"

  8 .
  9 . * 5 Load data
 10 . use `input_data', clear

 11 .
 12 . * 6,7 Create new variable
 13 . generate totalhours = hours * weeks

 14 . generate wagerate = salary/totalhours

 15 .
 16 . * 8,9 Create categorical variable
 17 . gen fulltime=0

 18 . replace fulltime=1 if weeks>=48 & hours>=35
      (205 real changes made)

 19 .
 20 . gen female=1

 21 . replace female=0 if sex==1
      (268 real changes made)

 22 .
 23 . * 10 Create log wage
 24 . generate logwage = log(wagerate)

 25 .
 26 . * 11 Create label
 27 . label variable totalhours "hours worked per week multiplied by number of weeks worked
      > "

 28 . label variable wagerate "salary divided by the total hours worked"

 29 . label variable logwage "the logarithmic value of the wage"
```

```
30 . label variable fulltime "type of worker, fulltime or part time"

31 . label define fulltime 0 "part time worker" 1 "fulltime worker"

32 . label variable female "=1 if female worker"

33 . label define female 0 "male" 1 "female"

34 .
35 . * 12 Create label values
36 . label values fulltime fulltime

37 . label values female female

38 .
39 . * 13 Produce summary table1
40 . outreg2 using "output/tables/table1.doc", replace sum(log) ///
   > keep(age educ weeks hours salary wagerate logwage)

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
     famid68 |        510    2888.275    1974.112         10       6864
         pid |        510    69.10588    69.26488          1        200
         sex |        510     1.47451    .4998401          1          2
         age |        510    27.13922    11.28842         18         88
        educ |        510    12.57647    2.821653          0         17
-------------+---------------------------------------------------------
       weeks |        510    39.48824     16.2374          1         52
       hours |        510    31.69804    12.59256          3         72
      salary |        510    15068.29    15454.56         50      90000
  totalhours |        510    1297.239    789.4712         12       3456
    wagerate |        510    13.42965     26.2417   .0396825   333.3333
-------------+---------------------------------------------------------
    fulltime |        510    .4019608    .4907755          0          1
      female |        510    .4745098    .4998401          0          1
     logwage |        510    2.197291    .7879197  -3.226844   5.809143


output/tables/table1.doc
dir : seeout

41 .
42 . * 14 Produce histogram
43 . histogram salary
(bin=22, start=50, width=4088.6364)

44 . graph export "output/figures/figure_1_histogram_salary.png", replace
(file /Users/macbook/Documents/Work/SP6015/HW 1/output/figures/figure_1_histogram_salar
> y.png written
> in PNG format)

45 . histogram logwage
(bin=22, start=-3.2268441, width=.41072669)
```

```
46 . graph export "output/figures/figure_2_histogram_logwage.png", replace
   (file /Users/macbook/Documents/Work/SP6015/HW 1/output/figures/figure_2_histogram_logwa
   > ge.png written
   >  in PNG format)

47 .
48 . * 15 Produce scatter plot
49 . twoway (scatter logwage age, lcolor(emidblue) lwidth(medthick))

50 . graph export "output/figures/figure_3_scatter_age_logwage.png", replace
   (file /Users/macbook/Documents/Work/SP6015/HW 1/output/figures/figure_3_scatter_age_log
   > wage.png writt
   > en in PNG format)

51 .
52 . * 16 Produce twoway table2
53 . asdoc tabulate female fulltime, save(table0.doc)

        =1 if |    type of worker,
       female | fulltime or part time
       worker | part time  fulltime  |     Total
   -----------+----------------------+----------
         male |       151        117 |       268
       female |       154         88 |       242
   -----------+----------------------+----------
        Total |       305        205 |       510
   Click to Open File:  table0.doc

54 . copy table0.doc "output/tables/table2.doc", replace

55 . erase "table0.doc"

56 .
57 . * 17 Save edited data
58 . save `output_data', replace
   file ./data/ps1_psid2003_edited.dta saved

59 .
60 . * 18 Create regression local macro
61 . local depvar logwage

62 . local indepvar1 female

63 . local indepvar2 female age

64 . local indepvar3 female age educ

65 . local indepvar4 female age educ fulltime

66 .
67 . * 18 Produce regression table3 with robust
68 . reg `depvar' `indepvar1', robust
```

```
Linear regression                               Number of obs     =        510
                                                F(1, 508)         =       2.45
                                                Prob > F          =     0.1184
                                                R-squared         =     0.0048
                                                Root MSE          =     .78681

------------------------------------------------------------------------------
             |               Robust
     logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1090065   .0696934    -1.56   0.118    -.2459293    .0279163
       _cons |   2.249015   .0485597    46.31   0.000     2.153613    2.344418
------------------------------------------------------------------------------

69 . outreg2 using "output/tables/table3.doc", replace ctitle(Model 1)
   output/tables/table3.doc
   dir : seeout

70 . reg `depvar' `indepvar2', robust

Linear regression                               Number of obs     =        510
                                                F(2, 507)         =      33.17
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0975
                                                Root MSE          =     .75002

------------------------------------------------------------------------------
             |               Robust
     logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1700327    .067121    -2.53   0.012    -.3019022   -.0381631
         age |   .0214198   .0026665     8.03   0.000     .0161811    .0266585
       _cons |   1.696655   .0847852    20.01   0.000     1.530082    1.863229
------------------------------------------------------------------------------

71 . outreg2 using "output/tables/table3.doc", append ctitle(Model 2)
   output/tables/table3.doc
   dir : seeout

72 . reg `depvar' `indepvar3', robust

Linear regression                               Number of obs     =        510
                                                F(3, 506)         =      27.17
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1135
                                                Root MSE          =     .74407

------------------------------------------------------------------------------
             |               Robust
     logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1755353   .0667826    -2.63   0.009    -.3067407     -.04433
         age |   .0216441   .0025202     8.59   0.000     .0166927    .0265954
        educ |   .0353564   .0114528     3.09   0.002     .0128556    .0578572
```

```
       _cons |   1.248523   .1684315     7.41   0.000     .9176114    1.579434
------------------------------------------------------------------------------

73 . outreg2 using "output/tables/table3.doc", append ctitle(Model 3)
   output/tables/table3.doc
   dir : seeout

74 . reg `depvar' `indepvar4', robust

   Linear regression                               Number of obs     =        510
                                                   F(4, 505)         =      22.75
                                                   Prob > F          =     0.0000
                                                   R-squared         =     0.1179
                                                   Root MSE          =     .74296

   ------------------------------------------------------------------------------
                |               Robust
        logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
   -------------+----------------------------------------------------------------
         female |  -.1623341   .0678176    -2.39   0.017    -.2955734   -.0290948
            age |   .0199511   .0026292     7.59   0.000     .0147856    .0251166
           educ |   .0345962   .0114684     3.02   0.003     .0120645    .0571279
       fulltime |   .1134394   .0656706     1.73   0.085    -.0155819    .2424607
          _cons |   1.252167   .1685941     7.43   0.000      .920935    1.583399
   ------------------------------------------------------------------------------

75 . outreg2 using "output/tables/table3.doc", append ctitle(Model 4)
   output/tables/table3.doc
   dir : seeout

76 .
77 . * 19 Produce regression table3 with conventional standard error
78 . reg `depvar' `indepvar1'

         Source |       SS           df       MS      Number of obs   =       510
   -------------+----------------------------------   F(1, 508)       =      2.44
          Model |   1.5110701         1   1.5110701   Prob > F        =    0.1188
       Residual |  314.485032       508  .619065023   R-squared       =    0.0048
   -------------+----------------------------------   Adj R-squared   =    0.0028
          Total |  315.996102       509  .620817489   Root MSE        =    .78681

   ------------------------------------------------------------------------------
        logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
   -------------+----------------------------------------------------------------
         female |  -.1090065   .0697715    -1.56   0.119    -.2460827    .0280698
          _cons |   2.249015   .0480619    46.79   0.000     2.154591     2.34344
   ------------------------------------------------------------------------------

79 . outreg2 using "output/tables/table4.doc", replace ctitle(Model 1)
   output/tables/table4.doc
   dir : seeout

80 . reg `depvar' `indepvar2'
```

```
      Source |       SS           df       MS      Number of obs   =       510
-------------+----------------------------------   F(2, 507)       =     27.37
       Model |  30.7963652         2  15.3981826   Prob > F        =    0.0000
    Residual |  285.199737       507  .562524135   R-squared       =    0.0975
-------------+----------------------------------   Adj R-squared   =    0.0939
       Total |  315.996102       509  .620817489   Root MSE        =    .75002

------------------------------------------------------------------------------
     logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1700327   .0670447    -2.54   0.012    -.3017522   -.0383131
         age |   .0214198   .0029687     7.22   0.000     .0155874    .0272522
       _cons |   1.696655    .089216    19.02   0.000     1.521377    1.871934
------------------------------------------------------------------------------

81 . outreg2 using "output/tables/table4.doc", append ctitle(Model 2)
   output/tables/table4.doc
   dir : seeout

82 . reg `depvar' `indepvar3'

      Source |       SS           df       MS      Number of obs   =       510
-------------+----------------------------------   F(3, 506)       =     21.59
       Model |  35.8560956         3  11.9520319   Prob > F        =    0.0000
    Residual |  280.140006       506  .553636376   R-squared       =    0.1135
-------------+----------------------------------   Adj R-squared   =    0.1082
       Total |  315.996102       509  .620817489   Root MSE        =    .74407

------------------------------------------------------------------------------
     logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1755353   .0665378    -2.64   0.009    -.3062597   -.0448109
         age |   .0216441   .0029461     7.35   0.000      .015856    .0274321
        educ |   .0353564   .0116954     3.02   0.003     .0123788     .058334
       _cons |   1.248523   .1726493     7.23   0.000     .9093249     1.58772
------------------------------------------------------------------------------

83 . outreg2 using "output/tables/table4.doc", append ctitle(Model 3)
   output/tables/table4.doc
   dir : seeout

84 . reg `depvar' `indepvar4'

      Source |       SS           df       MS      Number of obs   =       510
-------------+----------------------------------   F(4, 505)       =     16.87
       Model |  37.2407978         4  9.31019945   Prob > F        =    0.0000
    Residual |  278.755304       505  .551990701   R-squared       =    0.1179
-------------+----------------------------------   Adj R-squared   =    0.1109
       Total |  315.996102       509  .620817489   Root MSE        =    .74296

------------------------------------------------------------------------------
     logwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1623341   .0669596    -2.42   0.016    -.2938878   -.0307803
```

```
           age |   .0199511   .0031299     6.37   0.000     .0138019    .0261002
          educ |   .0345962   .0116879     2.96   0.003     .0116333    .0575591
      fulltime |   .1134394   .0716229     1.58   0.114    -.0272761    .2541549
         _cons |   1.252167   .1724079     7.26   0.000     .9134422    1.590892
------------------------------------------------------------------------------

85 . outreg2 using "output/tables/table4.doc", append ctitle(Model 4)
   output/tables/table4.doc
   dir : seeout

86 .
87 . * 20 Linearity diagnostics with acprplot on logwage and age
88 . quietly reg logwage age

89 . acprplot age, lowess

90 . graph export "output/figures/figure_4_acprplot_age.png", replace
   (file /Users/macbook/Documents/Work/SP6015/HW 1/output/figures/figure_4_acprplot_age.pn
   > g written in P
   > NG format)

91 .
92 . * 21 Multicolinearity diagnostics with
93 . * (1) Pairwise correlation matrix
94 . pwcorr `depvar' `indepvar4', star(0.05) sig

             |  logwage   female      age     educ fulltime
-------------+---------------------------------------------
     logwage |   1.0000
             |
             |
      female |  -0.0692   1.0000
             |   0.1188
             |
         age |   0.2933*  0.1262*  1.0000
             |   0.0000   0.0043
             |
        educ |   0.1171*  0.0244  -0.0219   1.0000
             |   0.0081   0.5830   0.6218
             |
    fulltime |   0.1754* -0.0743   0.3275*  0.0281   1.0000
             |   0.0001   0.0938   0.0000   0.5263
             |

95 .
96 . * (2) Correlation matrix graph
97 . graph matrix `depvar' `indepvar4', half

98 . graph export "output/figures/figure_5_correlation_matrix_graph.png", replace
   (file /Users/macbook/Documents/Work/SP6015/HW 1/output/figures/figure_5_correlation_mat
   > rix_graph.png
   > written in PNG format)

99 .
```

```
100 . * (3) Variance Inflation Factor (VIF)
101 . quietly reg `depvar' `indepvar4'

102 . vif

    Variable |       VIF       1/VIF
-------------+----------------------
         age |      1.15    0.868754
    fulltime |      1.14    0.877697
      female |      1.03    0.968112
        educ |      1.00    0.997089
-------------+----------------------
    Mean VIF |      1.08

103 .
104 . * 22 Homoscedasticity diagnostics with
105 . * (1) Breusch-Pagan test
106 . quietly reg `depvar' `indepvar4'

107 . estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of logwage

         chi2(1)      =    12.08
         Prob > chi2  =   0.0005

108 .
109 . * (2) Residual vs Fitted plot
110 . rvfplot, yline(0)

111 . graph export "output/figures/figure_6_residual_fitted_plot.png", replace
(file /Users/macbook/Documents/Work/SP6015/HW 1/output/figures/figure_6_residual_fitted
> _plot.png writ
> ten in PNG format)

112 .
113 . * 23 Omitted variable bias diagnostics with RESET
114 . quietly reg `depvar' `indepvar4'

115 . ovtest

Ramsey RESET test using powers of the fitted values of logwage
       Ho:  model has no omitted variables
                 F(3, 502) =      2.30
                  Prob > F =      0.0767

116 .
117 . * 24 Error normality diagnostics
118 . quietly reg `depvar' `indepvar4', robust

119 . predict e, resid
```

```
120 . hist e, kdensity normal
    (bin=22, start=-5.230608, width=.40921356)

121 . graph export "output/figures/figure_7_error_normality_plot.png", replace
    (file /Users/macbook/Documents/Work/SP6015/HW 1/output/figures/figure_7_error_normality
    > _plot.png writ
    > ten in PNG format)

122 .
123 . log close
          name:  <unnamed>
           log:  /Users/macbook/Documents/Work/SP6015/HW 1/output/logs/HW1_16_Feb_2022_22_2
    > 6_56.log
      log type:  text
     closed on:  16 Feb 2022, 22:27:12
    -------------------------------------------------------------------------------
    > --------------
```