```
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

```