HOMEWORK 1  
SP6015 Analisis Kuantitatif Untuk Kebijakan Publik

Maghfira Ramadhani (20021140)

Institut Teknologi Bandung

Daftar Isi

[HOMEWORK 1 SP6015 Analisis Kuantitatif Untuk Kebijakan Publik 4](#_Toc95942949)

[Nomor 1 4](#_Toc95942950)

[Nomor 2 4](#_Toc95942951)

[Nomor 3 4](#_Toc95942952)

[Nomor 4 4](#_Toc95942953)

[Nomor 5 4](#_Toc95942954)

[Nomor 6 5](#_Toc95942955)

[Nomor 7 5](#_Toc95942956)

[Nomor 8 5](#_Toc95942957)

[Nomor 9 5](#_Toc95942958)

[Nomor 10 5](#_Toc95942959)

[Nomor 11 5](#_Toc95942960)

[Nomor 12 6](#_Toc95942961)

[Nomor 13 6](#_Toc95942962)

[Nomor 14 6](#_Toc95942963)

[Nomor 15 8](#_Toc95942964)

[Nomor 16 8](#_Toc95942965)

[Nomor 17 9](#_Toc95942966)

[Nomor 18 9](#_Toc95942967)

[Nomor 19 10](#_Toc95942968)

[Nomor 20 11](#_Toc95942969)

[Nomor 21 12](#_Toc95942970)

[Nomor 22 14](#_Toc95942971)

[Nomor 23 15](#_Toc95942972)

[Nomor 24 16](#_Toc95942973)

[Lampiran Screenshot Direktori, Do-File, dan Log-File 17](#_Toc95942974)

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) |
| VARIABLES | N | mean | sd | min | 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

Chart, histogram

Description automatically generated

figure\_1\_histogram\_logwage.png

Chart, histogram

Description automatically generated

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

Chart, scatter chart

Description automatically generated

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| VARIABLES | Model 1 | Model 2 | Model 3 | 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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| VARIABLES | Model 1 | Model 2 | Model 3 | 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

Chart, scatter chart

Description automatically generated

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

A picture containing text, screenshot, receipt

Description automatically generated

figure\_5\_correlation\_matrix\_graph.png

Diagram

Description automatically generated

Log file

A screenshot of a computer

Description automatically generated with low confidence

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

Text

Description automatically generated

figure\_6\_residual\_fitted\_plot.png

Chart, scatter chart

Description automatically generated

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

Text

Description automatically generated

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

Chart, histogram

Description automatically generated

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

Table

Description automatically generated