

HW #1 SP6105 Analisis Kuantitatif Untuk Kebijakan Publik

Deadline: Rabu, 16 Februari 2022 Pukul 23.59

Asistensi: Jumat, 11 Februari 2022 pukul 13.30-14.30

1. Buatlah sebuah STATA do-file. Dengan menggunakan karakter untuk membuat komentar, buat identitas dari do-file ini yang berisikan nama mata kuliah, tugas, nama pembuat do-file, dan kapan do-file ini dibuat. Setelah ini untuk setiap Langkah-langkah tugas yang diberikan salin perintah ke dalam do-file sebelum menuliskan baris-baris perintah untuk menjalankan Langkah tugas yang diinstruksikan. Gunakan karakter komentar yang memastikan bahwa tulisan langkah tugas tersebut tidak dianggap sebagai baris perintah yang akan dieksekusi oleh STATA
2. Buat settingan environment software yang digunakan. Buat judul section yang sesuai dengan karakter komentar. Setting lingkungan perangkat lunak paling tidak memuat perintah sebagai berikut:
 - Gunakan kendali versi perangkat lunak (version control) untuk versi STATA 14
 - Perintahkan STATA untuk tidak berhenti (pause) pada setiap halaman ketika sedang mengeksekusi perintah
3. Unduh file `ps1_psid2003.dta` dan simpan di sebuah folder kerja (*working directory*). Perintahkan STATA untuk merubah *working directory* ke folder kerja dimana anda menyimpan file `.dta` yang baru saja anda unduh.
4. Buat sebuah *local macro* bernama `input_data`, yang menyimpan nama file `.dta` yang anda simpan pada Langkah 3. Buat sebuah *local macro* bernama `output_data` yang menyimpan nama file yang akan anda gunakan ketika menyimpan data final yang akan anda gunakan diakhir program ini. Jangan sertakan *path* file ke dalam local macro tersebut, cukup nama file `.dta` nya saja.
Hint: Untuk memahami apa itu dan menggunakan local macro anda dapat memperoleh bantuan dari STATA dengan memanggil perintah `help` diikuti dengan nama perintah baris yang ingin diketahui, dalam hal ini `local`.
5. Perintahkan STATA untuk membersihkan semua data yang sudah dimuat didalam memory. Dengan menggunakan macro yang anda buat pada Langkah sebelumnya, buka/muat data ke dalam STATA.
6. Buat sebuah variabel bernama `"totalhours"` dari hasil perkalian antara jumlah jam kerja per minggu (`hours`) dengan jumlah minggu bekerja (`weeks`)
7. Buat sebuah variable bernama `"wagerate"` dari variable `salary` dibagi oleh jumlah total waktu bekerja (`totalhours`)
8. Buat sebuah variable kategori bernama `"fulltime"` untuk setiap orang yang bekerja lebih dari 48 minggu dan total jam bekerja per minggunya lebih dari 35 jam. Variabel ini hanya memiliki nilai 1 untuk setiap orang yang memenuhi kriteria sebagai pekerja fulltime, dan nilai 0 untuk orang-orang yang tidak memenuhi kriteria diatas.
9. Buat sebuah variabel kategori bernama `"female"` dengan nilai 1 untuk setiap orang yang memiliki nilai variable `sex` bernilai 2 dan 0 untuk nilai variabel `sex` bernilai 1.

10. Buat sebuah variabel yang bernama “logwage” yang berisikan nilai logaritma dari variable `wagerate`

11. Beri label yang sesuai dengan variabel-variabel yang baru saja anda buat pada langkah 6-9.

12. Beri label pada nilai variabel `fulltime`. Untuk variable `fulltime` bernilai 1, beri label “Fulltime” sedangkan untuk nilai variabel `fulltime` 0, beri label “Parttime”

Kita akan melakukan analisis sederhana dan melaporkan hasilnya dalam format publikasi. Untuk itu kita akan menggunakan paket `outreg2` anda perlu memastikan bahwa paket ini telah terinstall dalam STATA dengan menggunakan perintah berikut:

```
ssc install outreg2
```

Ketik `help outreg2` untuk mempelajari dokumentasi dari paket ini. Anda juga dapat mempelajarinya dari dokumen ini: <https://www.princeton.edu/~otorres/Outreg2.pdf>

Pada HW ini kita akan menyajikan hasil deskriptif statistik dan beberapa model regresi kedalam 2 tabel.

13. Perintahkan STATA untuk melakukan perhitungan nilai statistika deskriptif untuk variabel-variabel dan simpan sebuah tabel kedalam file bernama `tabel1.doc` dengan perintah `outreg2`, yang akan menghasilkan tampilan seperti sebagai berikut:

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

14. Perintahkan STATA untuk melakukan histogram plot pada variabel “salary” dan “logwage”. Apakah anda melihat perubahan bentuk distribusi? Apa yang terjadi setelah kita melakukan transformasi logaritma pada data?

15. Perintahkan STATA untuk melakukan scatter plot pada variabel “age” pada sumbu X dan variabel “logwage” pada sumbu Y. Apakah anda melihat pola hubungan antara umur dan penghasilan?

16. Perintahkan STATA untuk membuat tabel tabulasi silang antara variabel “female” dan “fulltime”. Apakah anda melihat ada indikasi hubungan gender dan partisipasi dunia kerja? Kesimpulan apa yang anda dapat?

17. Simpan dataset yang anda buat dengan menggunakan macro `output_data` yang dibuat pada Langkah 4. Pastikan anda menggunakan opsi perintah yang memastikan bahwa dataset lama akan diganti dengan dataset terbaru setiap kali `do-file` dijalankan.

18. Perintahkan STATA untuk menjalankan Langkah-langkah perintah berikut ini:

- Buat sebuah local macro bernama `devar` yang berisikan nama variabel `logwage`

- Buat 3 (tiga) local macro indepvar1, indepvar2, dan indepvar3 yang berisikan nama-nama variabel berikut ini:
 - Macro indepvar1 berisikan nama variabel female
 - Macro indepvar2 berisikan nama variabel female, dan age
 - Macro indepvar3 berisikan nama variabel female, age, dan educ
 - Macro indepvar3 berisikan nama variabel female, age, dan educ
- Dengan menggunakan macro-macro diatas, buat perintah untuk menjalankan keempat buah model regresi secara berurutan antara dependent variabel logwage dengan independent variabel female; logwage dengan female dan age; logwage dengan female, age, dan educ; logwage dengan female, age, educ, dan fulltime. Kombinasikan dengan menggunakan perintah outreg2, sehingga akan dihasilkan satu buah file bernama table2.doc yang berisikan table hasil regresi dengan tampilan seperti sebagai berikut:

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

Gunakan robust standar error pada model regresi.

19. Ulangi Langkah ke -18 dengan menggunakan standar error konvensional.
20. Lakukan diagnose linearitas dengan menggunakan perintah acprplot. Periksa apakah hubungan antara logwage dan age tersebut linear atau tidak. Berikan interpretasi terhadap diagnosis anda.
21. Lakukan diagnosis multicollinearity pada variable logwage, female, age, educ, dan fulltime; dengan menggunakan (1) pairwise colleration matrix; (2) correlation matrix graph; dan (3) Variance Inflated Factor (VIF). Berikan interpretasi terhadap diagnosis anda.
22. Lakukan diagnosis homoskedasticity pada variable logwage, female, age, educ, dan fulltime ; dengan menggunakan (1) Breusch-pagan test; (2) residual-versus-fitted plot (rvfplot). Berikan interpretasi terhadap diagnosis anda.

23. Lakukan diagnosis omitted variable bias pada variable `logwage`, `female`, `age`, `educ`, dan `fulltime` ; dengan menggunakan the RESET (regression specification-error test) for omitted variables atau `ovtest`. Berikan interpretasi terhadap diagnosis anda.
24. Lakukan diagnosis normalitas pada error. Berikan interpretasi terhadap diagnosis anda.