

MODUL
ANALISIS DATA
DENGAN



DISUSUN OLEH:
ANI SILVIA, S.TP, M.S.M.

JURUSAN MANAJEMEN
FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS MUHAMMADIYAH PROF. DR. HAMKA

I. PENDAHULUAN

1.1. Pengenalan STATA

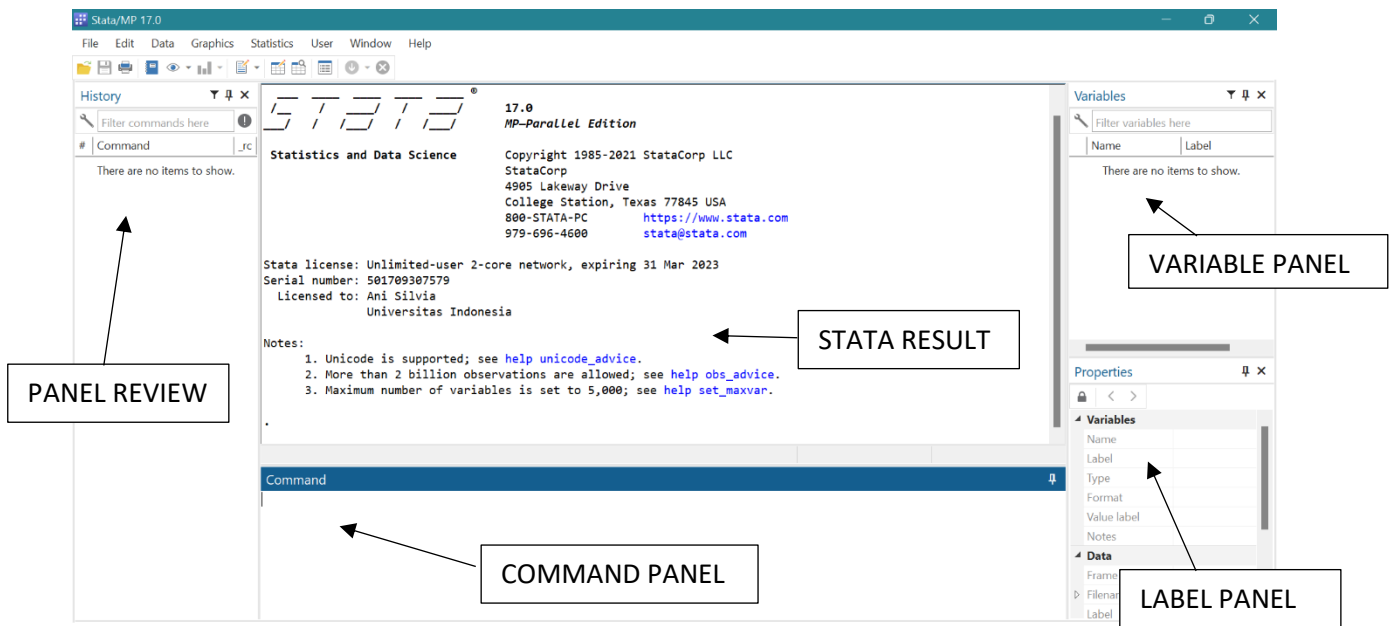
STATA merupakan salah satu software komputer untuk mengolah dan menganalisis data. Saat ini, STATA menjadi program statistik yang populer karena memiliki fungsi statistik yang lengkap dan kemampuan analisis yang akurat. Selain itu, STATA juga memiliki kelebihan lain diantaranya:

1. Dapat digunakan untuk berbagai bidang ilmu,
2. Mudah untuk digunakan (*user-friendly*),
3. Terdapat menu *help* untuk mencari informasi tentang perintah (command) yang dibutuhkan dalam analisis ekonometrika tanpa memerlukan koneksi internet,
4. Tersedianya banyak website-website yang berdiskusi tentang STATA,
5. Kemampuan untuk menerima fungsi-fungsi tambahan yang dibuat oleh pengguna STATA (*user-written*),
6. Adanya Do-file editor untuk memudahkan pengguna,
7. Terdapat online update untuk fungsi-fungsi statistika dan ekonometrika yang terbaru tanpa harus update software (untuk *original software*).

1.2. Tampilan STATA

Tampilan depan STATA memiliki lima window sekaligus yang memiliki fungsi yang berbeda, yaitu:

1. **Command Panel** adalah tempat untuk menuliskan syntax command (jika sudah hafal commandnya. Catatan: Kita tidak perlu menghafal command karena sudah ada Do-file editor)
2. **Panel Review (History)** menunjukkan seluruh perintah syntax yang pernah diketik pada Command Panel.
3. **STATA Results** menunjukkan hasil-hasil dari perintah yang diketikkan pada Command Panel (untuk mendapatkan hasil juga bisa menggunakan menu Statistic pada Toolbar di bagian atas)
4. **Variables Panel** menampilkan daftar singkatan nama-nama variabel dari data yang sedang digunakan serta nama lengkapnya di sisi kirinya. Singkatan nama-nama variabel inilah yang akan diketikkan pada Command Panel.
5. **Label Panel** digunakan jika kita ingin menambahkan nama lengkap (keterangan) pada variabel-variabel yang digunakan.



Gambar 1. Tampilan STATA

Pada bagian atas terdapat toolbar yang terdiri dari:

File, Edit, Data, Graphics, Statistics, Users, Window, dan help. Pilihan Statistics adalah pilihan utama untuk analisis statistik dan ekonometrika.

1.3. Langkah awal penggunaan STATA

Secara umum, analisis data dengan STATA meliputi beberapa proses yaitu:

1. **Menyiapkan data** dalam program Excel dengan formal .xls yang siap diolah dengan menggunakan program STATA
2. **Memasukan data tersebut ke dalam STATA** sehingga berubah menjadi file data dengan format .dta
3. **Mengolah data sesuai kebutuhan** dengan seperangkat perintah (Descriptive statistics, analisis regresi berganda, uji asumsi klasik dll)
4. **Menyimpan output hasil analisis**

Bagian pendahuluan hanya akan menjelaskan langkah (1) dan (2) yaitu untuk menyiapkan data dan memasukkan data ke dalam program STATA. Untuk langkah (3) dan (4) akan disampaikan pada bagian II, III dan IV dalam modul ini secara bersamaan.

Menyiapkan Data

Setelah data dikumpulkan, sebelum diolah menggunakan program STATA, data harus disiapkan sedemikian rupa. Kenali data kita apakah struktur data berupa data time series, data cross section, atau panel data.

Data pada Excel sebaiknya sudah dalam bentuk variabel penelitian. Jika struktur data berupa *time series*, maka susunan kolom pada Excel adalah sebagai berikut:

Tahun – Variabel Y – Variabel X1 – Variabel X2 – Variabel X3

Jika struktur data berupa *cross section*, maka susunannya:

Individu (ID) - Variabel Y – Variabel X1 – Variabel X2 – Variabel X3

Jika struktur data berupa *panel data*, maka susunannya:

Individu (ID) – Tahun – Variabel Y – Variabel X1 – Variabel X2 – Variabel X3

Dalam modul ini akan didemonstrasikan analisis data dengan menggunakan struktur data panel. Tabel 1 menunjukkan cara penyusunan data dengan struktur panel.

Tabel 1. Penyusunan Panel Data pada Excel

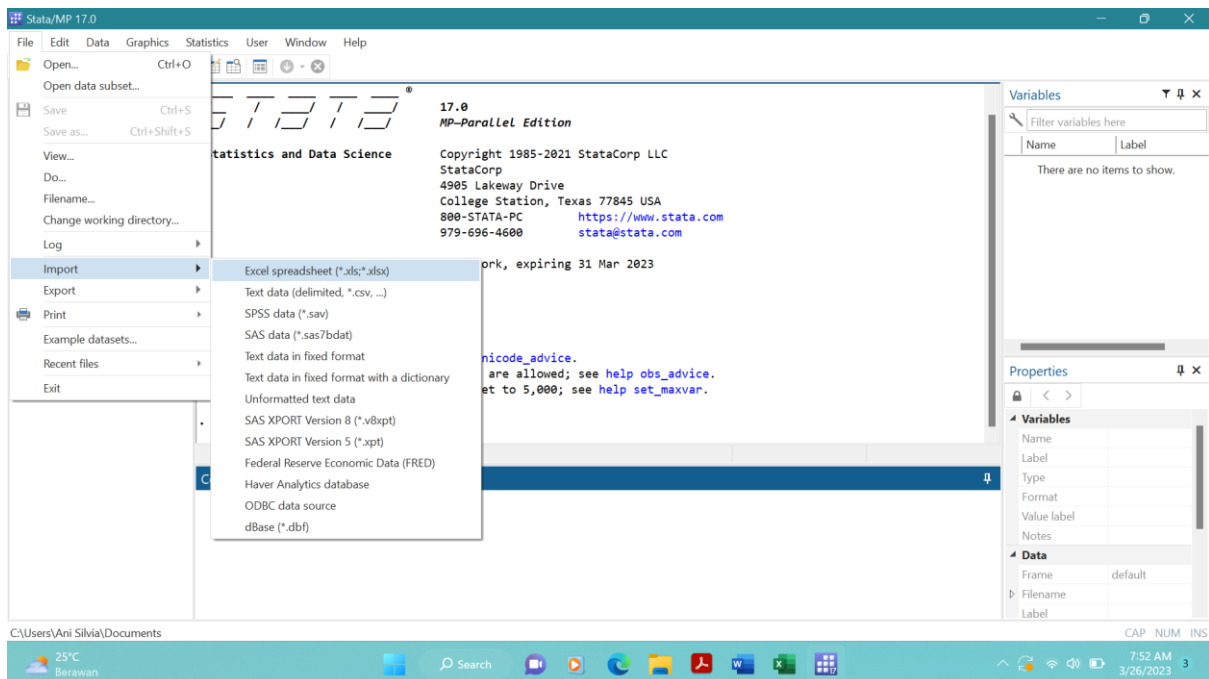
ID	Year	Y	X1	X2	X3	X4
1	2018	2.99	22.79	94.95	2.05	70.85
1	2019	2.95	24.59	98.85	2.15	82.71
1	2020	0.87	25.59	83.96	2.98	88.87
1	2021	1.02	26.38	84.56	2.84	86.62
2	2018	0.78	19.44	90.08	4.36	93.36
2	2019	1.3	19.89	86.32	2.77	87.04
2	2020	0.97	35.68	78.69	2.9	88.76
2	2021	0.73	34.94	68.97	3.2	90.07
3	2018	4.01	23.39	81.58	1.41	58.24
3	2019	4.02	23.8	80.47	1.34	59.09
3	2020	3.32	25.83	65.77	1.79	63.45
3	2021	3.41	25.66	61.96	2.16	54.15
4	2018	1.48	19.09	96.46	2.66	83.85
4	2019	1.09	21.42	94.13	3.62	89.28
4	2020	0.82	24.25	79.25	4.13	88.98
4	2021	1	26.38	76.28	4.04	84.94
5	2018	2.25	23.49	104.15	2.97	75.54
5	2019	2.09	24.07	107.92	3.02	77.04
5	2020	2.08	29.55	83.26	2.93	76.5

Memasukan data ke dalam STATA

Ada beberapa cara untuk memasukan data Excel menjadi data di STATA dengan format .dta. Pada modul ini akan dilakukan input data dengan cara yang paling mudah.

Langkah-langkahnya:

1. Buka program STATA
2. Masuk ke toolbar, klik file → Import → Excel spreadsheet seperti pada Gambar 2 berikut.



Gambar 2. Import Data dari Excel

3. Masukkan data Excel yang sudah disiapkan¹
4. Beritahu STATA jenis data yang dimasukkan yaitu data panel dengan menuliskan command sebagai berikut pada command panel.
tsset ID Year, yearly

¹ Catatan: seringkali setelah import data akan menghasilkan tampilan warna angka berwarna coklat pada Data Editor. Ini menunjukkan ketidaksinkronan antara sistem desimal pada program excel di komputer/laptop menggunakan tanda koma (“,”) bukan titik (“.”). STATA menggunakan sistem desimal dengan tanda titik mengikuti aturan internasional, sehingga format di Excel pada komputer/laptop harus diubah dulu menjadi format titik untuk desimal.

II. DESCRIPTIVE STATISTICS DENGAN STATA

Untuk menampilkan summary data secara keseluruhan atau menampilkan statistik secara deskriptif, maka dapat menuliskan command berikut:

summarize Y X1 X2 X3 X4

maka akan muncul output berikut:

Variable	Obs	Mean	Std. dev.	Min	Max
Y	112	1.604375	1.05829	-.73	4.02
X1	112	22.44313	6.902298	.18	67.15
X2	112	81.47857	18.63621	.82	113.5
X3	112	2.809375	1.217832	0	5.99
X4	112	81.63277	15.10004	.67	111.09

Output tersebut menampilkan nama variabel, jumlah observasi, nilai rata-rata, standar deviasi, nilai minimum dan nilai maksimum dari data secara keseluruhan (overall).

Jika ingin menampilkan data dengan analisis tambahan per individu (between) dan per tahun (within), maka command yang ditulis adalah sebagai berikut:

xtsum Y X1 X2 X3 X4

Variable		Mean	Std. dev.	Min	Max	Observations
Y	overall	1.604375	1.05829	-.73	4.02	N = 112
	between		.9785398	-.1625	3.69	n = 28
	within		.4339555	.351875	2.636875	T = 4
X1	overall	22.44313	6.902298	.18	67.15	N = 112
	between		5.227181	13.4575	41.885	n = 28
	within		4.588765	9.165625	47.70813	T = 4
X2	overall	81.47857	18.63621	.82	113.5	N = 112
	between		12.73185	41.0475	100.76	n = 28
	within		13.76915	13.08857	113.5736	T = 4
X3	overall	2.809375	1.217832	0	5.99	N = 112
	between		1.012204	.765	4.9175	n = 28
	within		.6973161	.801875	5.089375	T = 4
X4	overall	81.63277	15.10004	.67	111.09	N = 112
	between		10.11666	58.7325	97.6525	n = 28
	within		11.33272	7.187768	112.8878	T = 4

Untuk menampilkan analisis korelasi bisa menggunakan command Spearman correlation berikut:

spearman Y X1 X2 X3 X4

	Y	X1	X2	X3	X4
Y	1.0000				
X1	0.2522	1.0000			
X2	0.2844	-0.1426	1.0000		
X3	-0.4065	0.0101	-0.1433	1.0000	
X4	-0.8802	-0.2131	-0.1202	0.3486	1.0000

Arah hubungan dilihat dari tanda koefisien korelasi. Bila sign negatif, berarti hubungannya negatif atau berlawanan arah. Sebaliknya, bila sign positif berarti ada hubungan positif atau searah.

III. REGRESI LINEAR BERGANDA

Analisis korelasi bertujuan untuk mempelajari hubungan antara variabel independen X dengan variabel dependen Y. Korelasi tidak menjelaskan hubungan sebab-akibat atau causal relationship antara variabel X terhadap Y.

Berbeda dengan analisis korelasi, analisis regresi ditujukan untuk mempelajari hubungan sebab-akibat antara variabel independen X terhadap variabel dependen Y. Dalam regresi, variabel independen juga disebut sebagai variabel prediktor karena nilai variabel y dapat diprediksi dari nilai variabel X.

Model ekonometrika untuk analisis regresi dengan data panel adalah sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{1,it} + \beta_2 X_{2,it} + \beta_3 X_{3,it} + \beta_4 X_{4,it} + \varepsilon_{it}$$

Regresi pada data panel berbeda dengan data time series dan data cross-section. Regresi data panel bisa menggunakan model Common Effect (Pooled Least Square), Fixed Effect (FE) atau Random Effect (RE). Regresi model Common Effect dapat dilakukan dengan command metode OLS biasa sebagai berikut:

regress Y X1 X2 X3 X4

Source	SS	df	MS	Number of obs	=	112
Model	65.1152752	4	16.2788188	F(4, 107)	=	29.42
Residual	59.2022811	107	.553292347	Prob > F	=	0.0000
Total	124.317556	111	1.11997798	R-squared	=	0.5238
				Adj R-squared	=	0.5060
				Root MSE	=	.74384

Y	Coefficient	Std. err.	t	P> t	[95% conf. interval]
X1	.030415	.0105877	2.87	0.005	.0094261 .0514039
X2	.0218357	.0038634	5.65	0.000	.014177 .0294945
X3	-.2630486	.0598443	-4.40	0.000	-.381683 -.1444142
X4	-.0381999	.0048913	-7.81	0.000	-.0478963 -.0285034
_cons	2.999987	.5219947	5.75	0.000	1.965194 4.034781

Untuk menyimpan output hasil regresi OLS digunakan command berikut:

estimates store ols

Secara default, regresi data panel menggunakan command untuk model Random Effect dan untuk menyimpan output adalah sebagai berikut:

xtreg Y X1 X2 X3 X4

estimates store re

```

Random-effects GLS regression              Number of obs   =       112
Group variable: ID                        Number of groups =        28

R-squared:                                Obs per group:
  Within = 0.3681                          min =           4
  Between = 0.5643                         avg =           4.0
  Overall = 0.4797                          max =           4

corr(u_i, X) = 0 (assumed)                Wald chi2(4)    =       70.45
                                           Prob > chi2     =       0.0000
  
```

Y	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
X1	.0259424	.0090426	2.87	0.004	.0082192	.0436657
X2	.0187766	.0030994	6.06	0.000	.0127019	.0248513
X3	-.2422585	.0549902	-4.41	0.000	-.3500373	-.1344797
X4	-.0181979	.0037884	-4.80	0.000	-.025623	-.0107727
_cons	1.658392	.4235221	3.92	0.000	.8283044	2.48848
sigma_u	.46684699					
sigma_e	.40180409					
rho	.57446056	(fraction of variance due to u_i)				

Untuk merubahnya menjadi model Fixed Effect, command default harus ditambahkan “fe” di belakangnya sebagai berikut:

xtreg Y X1 X2 X3 X4, fe

estimates store fe

Fixed-effects (within) regression
 Group variable: ID

Number of obs = 112
 Number of groups = 28

R-squared:
 Within = 0.3821
 Between = 0.4822
 Overall = 0.4271

Obs per group:
 min = 4
 avg = 4.0
 max = 4

corr(u_i, Xb) = 0.3383
 F(4, 80) = 12.37
 Prob > F = 0.0000

Y	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
X1	.0233649	.008747	2.67	0.009	.0059579	.0407719
X2	.0168724	.0029323	5.75	0.000	.0110369	.0227079
X3	-.2083782	.0553593	-3.76	0.000	-.3185467	-.0982096
X4	-.0112418	.0035452	-3.17	0.002	-.0182969	-.0041867
_cons	1.208369	.3891913	3.10	0.003	.4338538	1.982885
sigma_u	.78078022					
sigma_e	.40180409					
rho	.79061859	(fraction of variance due to u_i)				

F test that all u_i=0: F(27, 80) = 10.62
 Prob > F = 0.0000

Untuk menampilkan tabel hasil output ketiga model dan bisa membandingkan koefisien serta signifikansi pengaruh variabel independen terhadap variabel dependen digunakan command berikut:

estimates table ols re fe, b(%4.3f) p(%4.3f) stats(N r2)

Variable	ols	re	fe
X1	0.030 0.005	0.026 0.004	0.023 0.009
X2	0.022 0.000	0.019 0.000	0.017 0.000
X3	-0.263 0.000	-0.242 0.000	-0.208 0.000
X4	-0.038 0.000	-0.018 0.000	-0.011 0.002
_cons	3.000 0.000	1.658 0.000	1.208 0.003
N	112	112	112
r2	0.524		0.382

Atau, untuk menampilkan variabel mana saja yang signifikan melalui tanda bintang, bisa dilakukan dengan command berikut:

estimates table ols re fe, b(%4.3f) star stats(N r2)

Variable	ols	re	fe
X1	0.030**	0.026**	0.023**
X2	0.022***	0.019***	0.017***
X3	-0.263***	-0.242***	-0.208***
X4	-0.038***	-0.018***	-0.011**
_cons	3.000***	1.658***	1.208**
N	112	112	112
r2	0.524		0.382

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Untuk menentukan model mana yang terbaik sesuai dengan data yang dimiliki, maka diperlukan uji statistik Chow test, Hausman test, dan Lagrange Multiplier test.

1. Chow Test

Uji ini digunakan untuk memilih model terbaik antara OLS dan Fixed Effect. Output Fixed Effect berguna untuk melihat apakah model Fixed Effect lebih baik daripada model OLS. Hipotesis pada Chow test adalah sebagai berikut:

Ho: Model OLS

H1: Model Fixed Effect

Jika F-statistik > dari F-table, maka Ho ditolak yang artinya model Fixed Effect adalah model terbaik. Probability dari F-Statistics adalah 0.0000, artinya tolak Ho sehingga berdasarkan Chow test terpilih model Fixed Effect.

```

Fixed-effects (within) regression
Group variable: ID
Number of obs   =   112
Number of groups =   28

R-squared:
  Within = 0.3821
  Between = 0.4822
  Overall = 0.4271

Obs per group:
  min = 4
  avg = 4.0
  max = 4

corr(u_i, Xb) = 0.3383
F(4, 80) = 12.37
Prob > F = 0.0000

```

Y	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
X1	.0233649	.008747	2.67	0.009	.0059579	.0407719
X2	.0168724	.0029323	5.75	0.000	.0110369	.0227079
X3	-.2083782	.0553593	-3.76	0.000	-.3185467	-.0982096
X4	-.0112418	.0035452	-3.17	0.002	-.0182969	-.0041867
_cons	1.208369	.3891913	3.10	0.003	.4338538	1.982885
sigma_u	.78078022					
sigma_e	.40180409					
rho	.79061859	(fraction of variance due to u_i)				

F test that all u_i=0: F(27, 80) = 10.62

Prob > F = 0.0000

2. Hausman Test

Uji ini dilakukan pada saat uji Chow memilih model Fixed Effect sehingga perlu dilakukan uji lanjutan untuk memilih model terbaik antara Fixed Effect dan Random Effect. Hipotesisnya adalah sebagai berikut:

Ho: Model Random Effect

H1: Model Fixed Effect

Uji hipotesis dalam uji Hausman mengikuti distribusi statistik Chi Square dengan degree of freedom sebanyak k variabel independen. Jika nilai statistik Chi Square lebih besar dari nilai kritisnya maka keputusannya adalah menolak Ho dan model terbaik adalah Fixed Effect (Gujarati, 2012).

Command dalam STATA untuk uji Hausman adalah sebagai berikut:

quietly xtreg Y X1 X2 X3 X4, fe

estimate store fe

quietly xtreg Y X1 X2 X3 X4

estimate store re

hausman fe re, sigmamore

	Coefficients		(b-B) Difference	sqrt(diag(V_b-V_B)) Std. err.
	(b) fe	(B) re		
X1	.0233649	.0259424	-.0025775	.0041486
X2	.0168724	.0187766	-.0019042	.0012319
X3	-.2083782	-.2422585	.0338803	.0306728
X4	-.0112418	-.0181979	.006956	.001381

b = Consistent under H0 and Ha; obtained from xtreg.
 B = Inconsistent under Ha, efficient under H0; obtained from xtreg.

Test of H0: Difference in coefficients not systematic

```
chi2(4) = (b-B)'[(V_b-V_B)^(-1)](b-B)
        = 27.38
Prob > chi2 = 0.0000
```

Berdasarkan output di atas, Probability dari Chi Square adalah 0.0000 yang artinya tolak H0 sehingga model yang terbaik tetap model Fixed Effect.

3. Lagrange Multiplier test

Jika hasil Chow test memilih model OLS dan Hausman test memilih model Random Effect, maka untuk memilih antara OLS dan Random Effect digunakan uji Lagrange Multiplier. Karena model terbaik dari hasil uji sebelumnya tetap memilih Fixed Effect sebagai model terbaik, maka uji Lagrange Multiplier tidak dibutuhkan.

Jika dibutuhkan untuk uji dengan data lain, uji Lagrange Multiplier dapat dilakukan dengan menggunakan command sebagai berikut:

quietly xtreg Y X1 X2 X3 X4

xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

$$Y_{[ID,t]} = Xb + u[ID] + e[ID,t]$$

Estimated results:

	Var	SD = sqrt(Var)
Y	1.119978	1.05829
e	.1614465	.4018041
u	.2179461	.466847

Test: Var(u) = 0

```
chibar2(01) = 39.94
Prob > chibar2 = 0.0000
```

Hipotesis statistik untuk uji ini adalah sebagai berikut:

Ho: Model OLS

H1: Model Random Effect

Berdasarkan output di atas, Probability dari Chi Square adalah 0.0000 yang artinya tolak Ho sehingga model yang terbaik menurut uji Lagrange Multiplier adalah model Random Effect.

IV. UJI ASUMSI KLASIK

Data panel merupakan data yang menggabungkan antara data time series dan cross section. Ada beberapa kelebihan penggunaan data panel, yaitu:

1. Meningkatkan jumlah observasi
2. Variasi antar unit berbeda antar ruang dan antar waktu
3. Sedikit terjadi kolinearitas antar variabel sehingga sangat kecil kemungkinan terjadi pelanggaran asumsi multikolinearitas (Gujarati, 2012)

Kelebihan ini membuat regresi pada data panel hanya memerlukan uji asumsi klasik berupa uji autokorelasi dan uji heteroskedastisitas saja.

Namun demikian, modul ini akan membahas semua pelanggaran asumsi klasik dimungkinkan dengan command pada STATA.

1. Uji Multikolinearitas

Uji ini ditujukan untuk mengetahui apakah terdapat korelasi antar independen variabel. Deteksi multikolinearitas dapat dilakukan melalui Variance Inflation Factor (VIF). Jika nilai $VIF > 10$, maka berarti terdapat pelanggaran asumsi klasik multikolinearitas. Pelanggaran ini menyebabkan standard error koefisien regresi yang semakin besar sehingga confidence interval untuk pendugaan parameter semakin lebar. Akibatnya, sangat besar kemungkinan terjadi kekeliruan dalam menyimpulkan hasil regresi. Cara lain untuk mendeteksi masalah multikolinearitas adalah melalui nilai korelasi berpasangan untuk variabel independen. Jika nilai korelasi $> \text{plus minus } 0.8$, kemungkinan besar terjadi pelanggaran multikolinearitas.

Command yang bisa dituliskan adalah sebagai berikut:

**quietly xtreg Y X1 X2 X3 X4, fe
vif, uncentered**

Variable	VIF	1/VIF
X4	23.26	0.042989
X2	14.09	0.070977
X1	10.54	0.094908
X3	6.72	0.148883
Mean VIF	13.65	

Konklusi: terjadi pelanggaran asumsi multikolinearitas.

Untuk mengatasi permasalahan multikolinearitas, beberapa variabel akan melalui proses ortogonalisasi menggunakan prosedur Gram-Schmidt yang akan menghilangkan secara parsial common variance dan merubahnya menjadi tidak berkorelasi antara satu variabel dengan variabel yang lain (Saville & Wood, 2012).

Kemudian, seluruh model diestimasi dengan menggunakan variabel yang sudah ditransformasi. Cara melakukan transformasi data dengan command berikut:

orthog X4 X2 X1, generate(oX4 oX2 oX1)

kemudian estimasi ulang model Fixed effect sebagai berikut:

**quietly xtreg Y oX1 oX2 X3 oX4, fe
vif, uncentered**

Variable	VIF	1/VIF
X3	1.01	0.990337
oX4	1.01	0.991905
oX1	1.00	0.998431
oX2	1.00	0.999975
Mean VIF	1.00	

Konklusi: sudah tidak terjadi pelanggaran asumsi multikolinearitas.

2. Uji Heteroskedastisitas

Pelanggaran heteroskedastisitas muncul apabila residual model regresi tidak memiliki variance yang konstan atau bisa dikatakan masih bervariasi antar waktu atau antar individu. Akibatnya, reliabilitas observasi akan berbeda-beda jika ada perubahan kondisi yang tidak tertangkap di dalam model ekonometrika (Kuncoro, 2011). Biasanya, gejala ini lebih sering muncul pada regresi dengan jenis data cross section dan juga data panel. Salah satu cara mendeteksi permasalahan heteroskedastisitas adalah menggunakan modified Wald test. Hipotesis dari tes ini adalah sebagai berikut:

Ho: Homoskedastisitas

H1: Heteroskedastisitas

Ada package tambahan agar bisa menggunakan command uji heteroskedastisitas pada model regresi data panel dengan model Fixed Effect. Install terlebih dulu package **st0004_2**, atau langsung **xttest3**
ssc install xttest3

Untuk menguji gejala heteroskedastisitas dalam regresi model Fixed Effect digunakan command berikut:

**quietly xtreg Y oX1 oX2 X3 oX4, fe
xttest3**

Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: $\sigma(i)^2 = \sigma^2$ for all i

chi2 (28) = 46964.95
Prob>chi2 = 0.0000

Kesimpulan: terdapat pelanggaran asumsi klasik heteroskedastisitas.

3. Uji Autokorelasi

Permasalahan autokorelasi terjadi karena residual model saling berkorelasi antar satu observasi ke observasi lainnya (Kuncoro, 2011). Permasalahan ini biasa terjadi pada data dengan struktur time series. Autokorelasi pada data panel dapat dideteksi dengan Wooldridge test. Hipotesis dari tes ini adalah sebagai berikut:

Ho: Tidak terdapat autokorelasi

H1: Terdapat autokorelasi

commandnya sebagai berikut:

xtserial Y oX1 oX2 X3 oX4

Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
F(1, 27) = 4.432
Prob > F = 0.0447

Kesimpulan: terdapat pelanggaran asumsi klasik autokorelasi

Permasalahan heteroskedastisitas dan autokorelasi pada data panel dengan metode Fixed Effect dapat diatasi dengan menggunakan robust standar error. Robust standard error pada dasarnya hanya mengoreksi perhitungan standard error tanpa mengubah hasil estimasi koefisien regresi. Meski hasil estimasi koefisien regresi tetap tidak efisien, inferensia statistik tetap bisa dilakukan.

Caranya dengan mengetikan command berikut:

xtreg Y oX1 oX2 X3 oX4, fe robust

```

Fixed-effects (within) regression      Number of obs   =   112
Group variable: ID                    Number of groups =    28

R-squared:                             Obs per group:
  Within = 0.3821                       min =         4
  Between = 0.4822                       avg =        4.0
  Overall = 0.4271                       max =         4

corr(u_i, Xb) = 0.3383                  F(4,27)        =    4.21
                                          Prob > F        =    0.0089

```

(Std. err. adjusted for 28 clusters in ID)

Y	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
oX1	.1559235	.0683976	2.28	0.031	.0155832	.2962639
oX2	.2852215	.0943162	3.02	0.005	.0917006	.4787423
X3	-.2083782	.0698262	-2.98	0.006	-.3516497	-.0651066
oX4	-.109001	.0818474	-1.33	0.194	-.2769379	.0589359
_cons	2.189787	.196168	11.16	0.000	1.787284	2.592291
sigma_u	.78078022					
sigma_e	.40180409					
rho	.79061859	(fraction of variance due to u_i)				

4. Uji normalitas

Uji Normalitas digunakan untuk menguji apakah residual atau error term dari model regresi panel terdistribusi normal atau tidak. Salah satu uji statistik yang untuk mendeteksi normalitas data adalah dengan Jarque-Berra test. Hipotesis dari tes ini adalah sebagai berikut:

Ho: Data terdistribusi normal

H1: Data tidak terdistribusi normal

Untuk dapat melakukan Jarque-Berra test, perlu membuat series baru yang berisi nilai dari residual model. Commandnya adalah sebagai berikut:

predict resid, residuals
histogram resid, normal
sktest resid

Variable	Obs	Pr(skewness)	Pr(kurtosis)	Joint test	
				Adj chi2(2)	Prob>chi2
resid	112	0.3456	0.2075	2.54	0.2814

Kesimpulan: residual model terdistribusi normal sehingga tidak terjadi pelanggaran asumsi klasik normalitas.

Command untuk data time series dan cross section

```
import data
*Seting data cross section tidak perlu dilakukan
tsset mdate, monthly // setting data time series
summarize Y X1 X2 X3 //statistic descriptive
spearman Y X1 X2 X3 // Correlation
regress Y X1 X2 X3 //OLS regression
estimate store ori //saving original regression
estat vif //calculate VIF
estat dwatson // Uji Durbin-Watson (DW Test)
estat bgodfrey // Uji Lagrange Multiplier (LM Test) atau (Breusch – Godfrey)
newey Y X1 X2 X3, lag(1) //Cara mengobati otokorelasi dengan newey - west
estimate store newey // saving modified regression
estat hettest // Deteksi Heteroskedastisitas
predict resid, residuals //Generate Residual series
histogram resid, normal //histogram graph for residual distribution
sktest resid //Skewness and kurtosis (Jarque Bera test for normality)
regress Y X1 X2 X3, robust //using robust standar error
estimate store robust // saving robust regression
estimates table ori newey robust, b(%4.3f) p(%4.3f) stats(N r2) //comparing
models
```

Referensi:

1. Gujarati, N. D. (2012). Dasar-Dasar Ekonometrika. Jakarta: Salemba.
2. Kuncoro, Mudrajat. 2011. Metode Kuantitatif. Yogyakarta: Sekolah Tinggi Ilmu Manajemen: YKPN.
3. Saville, D. J., & Wood, G. R. (2012). *Statistical methods: The geometric approach*. Springer Science & Business Media.