

# Statistika Kesehatan

## *Aplikasi Stata dan SPSS*

*Najmah, S.K.M., M.P.H*



**Penerbit  
Salemba Medika**

---

## Statistika Kesehatan: Aplikasi Stata dan SPSS

Najmah, S.K.M., M.P.H.

Manajer Penerbitan dan Produksi: Novietha Indra Sallama  
Supervisor Editor: Akliya Suslia  
Editor: Peni Puji Lestari  
Tata Letak: Leonardo Manggala Wardhana  
Desain Sampul: Ferdy Firnaldy



Hak Cipta © 2017 Penerbit Salemba Medika  
Jln. Raya Lenteng Agung No. 101  
Jagakarsa, Jakarta Selatan 12610  
Telp. : (021) 781 8616  
Faks. : (021) 781 8486  
Website: <http://www.penerbitsalemba.com>  
E-mail : [info@penerbitsalemba.com](mailto:info@penerbitsalemba.com)

**Hak cipta dilindungi undang-undang.** Dilarang memperbanyak sebagian atau seluruh isi buku ini dalam bentuk apa pun, baik secara elektronik maupun mekanis, termasuk tidak terbatas pada memfotokopi, merekam, atau dengan menggunakan sistem penyimpanan lainnya, tanpa izin tertulis dari Penerbit.

### UNDANG-UNDANG NOMOR 28 TAHUN 2014 TENTANG HAK CIPTA

1. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta yang meliputi penerjemahan dan pengadaptasian Ciptaan untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama **3 (tiga) tahun** dan/atau pidana denda paling banyak **Rp500.000.000,00 (lima ratus juta rupiah)**.
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta yang meliputi penerbitan, penggandaan dalam segala bentuknya, dan pendistribusian Ciptaan untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama **4 (empat) tahun** dan/atau pidana denda paling banyak **Rp1.000.000.000,00 (satu miliar rupiah)**.
3. Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada poin kedua di atas yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama **10 (sepuluh) tahun** dan/atau pidana denda paling banyak **Rp4.000.000.000,00 (empat miliar rupiah)**.

Pengetahuan medis senantiasa berubah. Oleh karena itu, standar tindakan pencegahan serta perubahan dalam perawatan dan terapi wajib diikuti seiring dengan penelitian dan pengalaman klinis baru yang memperluas pengetahuan. Pembaca disarankan untuk memeriksa informasi terbaru yang disediakan oleh produsen masing-masing obat (yang akan diberikan) untuk memverifikasi dosis, metode, dan interval pemberian yang direkomendasikan serta kontraindikasi. Merupakan tanggung jawab dari praktisi dengan memperhatikan pengalaman dan pengetahuan pasien untuk menentukan dosis dan perawatan terbaik bagi masing-masing pasien. Penerbit maupun penulis tidak bertanggung jawab atas kecelakaan dan/atau kerugian yang dialami seseorang atau sesuatu yang diakibatkan oleh penerbitan buku ini.

---

Najmah

Statistika Kesehatan: Aplikasi Stata dan SPSS/Najmah

—Jakarta: Salemba Medika, 2016  
1 jil., 214 hlm., 19 × 26 cm

ISBN: 978-602-6450-18-0

1. Kesehatan  
I. Judul

2. Statistika  
II. Najmah

*My Love*  
*Kusnan Sayuti,*  
*Queency Qoryra Himada,*  
*Maitreya Adilla Sultanah*





# Tentang Penulis



**Najmah, S.K.M., M.P.H.**  
(*PhD candidate*)

Lahir 1983 silam dan besar di kampung di bantaran Sungai Musi Palembang. Najmah adalah *PhD candidate* di School of Public Health di Auckland University of Technology, Selandia Baru (2015–2019) dengan sponsor *New Zealand Asean Scholarship*. Pendidikan Strata-2 diselesaikan di bidang Epidemiologi dan Biostatistika di School of Population Health di The University of Melbourne, Victoria (2008–2009) dengan sponsor *Australian Partnership Scholarship, Ausaid* dan Strata-1 di Prodi Kesehatan Masyarakat Fakultas Kedokteran Universitas Sriwijaya (2001–2005) dengan beasiswa BBM (Bantuan Belajar Akademik) Pemerintah Indonesia. Najmah bekerja sebagai dosen di Fakultas Kesehatan Masyarakat, Universitas Sriwijaya sejak 2006.

Tiga buku yang telah diterbitkan sebelumnya antara lain ***Manajemen dan Analisa Data Kesehatan, Kombinasi Teori dan Aplikasi SPSS*** (2011), ***Epidemiologi untuk Mahasiswa Kesehatan Masyarakat***-Edisi 2 (2015) dan ***Epidemiologi Penyakit Menular*** (2016) dan ini buku keempat bekerja sama dengan Salemba Medika,

Masih terus belajar dan belajar, karena ilmu kita yang masih sangat terbatas dan penuh kekurangan. Ke depan, ranah kualitatif akan menjadi target ranah buku kesehatan yang akan diterbitkan, sesuai dengan tantangan baru dalam proses Strata-3 yang sedang dijalaninya. Dunia Kesehatan penuh warna dengan mempelajari dan mengaplikasikan ranah kuantitatif dan kualitatif.

Saran, kritik silahkan email ke najem240783@yahoo.com/ <https://www.facebook.com/najmah.usman.7/>

Website interaktif yang telah dikembangkan:

<http://metopidfkmunsri.blogspot.co.id/>

<http://madfkmunsri.blogspot.co.id/>

<http://queencyhimada.blogspot.co.id/>

Auckland, 2017

Salam Cerdas, Tiada Batas untuk menjadi Cerdas

# Profil Editor Ahli Besar



Lahir di Kabupaten Agam Sumatera Barat 31 Januari 1972. Beliau menyelesaikan Sarjana Kesehatan Masyarakat di Universitas Indonesia Jakarta tahun 1994, dan Master of Science in Epidemiologi (Public Health) di University of the Philippines Manila tahun 2000, serta Doktor Kesehatan Masyarakat di Universitas Indonesia Jakarta tahun 2011.

Sejak lulus tahun 1994, bekerja sebagai peneliti di Pusat Penelitian Kesehatan Universitas Indonesia dan tahun 1997 diangkat menjadi pengajar tetap di Departemen Biostatistik dan Kependudukan, Fakultas Kesehatan Masyarakat Universitas Indonesia. Beliau mengajar mahasiswa S-1, S-2, dan S-3 di bidang Statistik Deskriptif dan Inferensial, termasuk Statistik Non-parametrik, Manajemen dan Analisis Data menggunakan Epi-Info/EpiData, SPSS, Stata, dan Minitab. Mengajar Metode Penelitian, Rancangan Sampel, dan Aplikasi Analisis Multivariabel pada berbagai jenis studi di bidang kesehatan masyarakat dan kedokteran.

Selain itu, beliau juga sering memberikan bantuan teknis dan konsultasi di bidang Biostatistik, Metode Penelitian, dan Teknik Analisis Data di berbagai universitas dan institusi kesehatan di Indonesia.





# Prakata

*Alhamdulillah*, puji syukur ke hadirat Allah Swt., karena atas *ridho*, rahmat, dan karunia-Nya penulis dapat menyelesaikan buku keempat yang berjudul *Statistika Kesehatan: Aplikasi Stata dan SPSS*. Buku ini disusun guna mempermudah mahasiswa dan praktisi khususnya di bidang kesehatan dalam mengolah data statistik menggunakan *software* Stata dan SPSS serta menginterpretasikan output yang didapatkan.

Buku ini memberikan informasi tentang aplikasi STATA dan SPSS dalam analisis deskriptif, uji hipotesis (*chi square*, *independent T-test*, *paired T-test*, dan anova), analisis regresi logistik sederhana, analisis data sekunder, serta aplikasi penghitungan epidemiologi, dan aplikasi regresi Cox. Aplikasi SPSS di buku ini menggunakan Stata versi 11 dan SPSS versi 20. Walaupun terdapat beberapa perbedaan menu data editor yang dimaksud dalam buku ini dengan program SPSS versi lainnya, diharapkan pembaca dapat menelaah menu data editor dalam buku ini. Pada setiap bab, penulis memberikan contoh dari beberapa penelitian penulis dan peneliti lainnya sehingga dapat meningkatkan pemahaman pembaca dalam mengaplikasikan Stata dan SPSS serta dapat dipelajari secara berkesinambungan.

Dalam proses penulisan buku ini, penulis tentunya mendapatkan bantuan dari semua pihak yang tulus dan ikhlas memberikan sumbangan berupa pikiran, bimbingan, dorongan, dan nasihat. Rasa terima kasih yang sebesar-besarnya penulis sampaikan kepada:

1. Rektor Universitas Sriwijaya, Prof. Dr. Ir. H. Anis Saggaff, MSCE; Dekan Fakultas Kesehatan Masyarakat Universitas Sriwijaya, Hamzah Hasyim, S.K.M., M.K.M. (Dekan periode 1); Iwan Stiabudi, S.K.M., M.Kes (Dekan periode 2); para Wakil Dekan FKM Unsri beserta Kaprodi IKM, S-2 dan Gizi FKM Unsri (Elvi Sunarsih, S.K.M., M.Kes, Dr. Misnaniarti, S.K.M., M.K.M., Dr. Novrikasari, S.K.M., M.Kes), serta Ketua Lembaga Penelitian Unsri (Prof. Drs. Tatang Suhery, M.A, Ph.D).
2. Pengayom pendidikan saya, Prof. Badia Perizade, Prof. Zarkasih Anwar, Sp.(A) dan dr. Husnil Farouk, M.P.H., serta Rekan kerja di lingkungan Fakultas Kesehatan Masyarakat Universitas Sriwijaya.
3. Kedua orang tua Ani Erosa (Alm.) dan Usman Nurdin, Sayuti dan Sumiati yang selalu mengutamakan pendidikan bagi anaknya dalam kondisi keterbatasan. Saudara-saudaraku: M. Reza Arsyadi & Ety Yulianti, A.Md., Rina Nur'ain & Iskandarian, A.Md., M. Faris Nurdiansyah, S.T., & Yunita Lestari, S.P., M. Nirwan Fauzan, S.T., & Widyawaiti, A.Md., Rumiati, A.Md., dan Sulaiman, S.H., dan Karmina, S.P., dan Halim, S.P., untuk tali persaudaraan yang tiada akhir.

4. Suami tercinta Kusnan Sayuti, S.E., dan anakku Queency Qoryra Himada dan Maitreya Adilla Sultanah, terima kasih atas cinta tak bersyarat yang telah diberikan.
5. Teman-teman yang telah membantu proses pengeditan buku ini, Yeni, S.K.M., M.K.M., Desy Indah Permatasari, S.K.M., Adelina Fitri, S.K.M., Nur Hidayah S.K.M., Eka Mujiati, S.K.M., dan teman-teman FKM Unsri lainnya yang tidak bisa disebutkan satu persatu.
6. Almamaterku tercinta, MI Azhariah Palembang, SMPN 35 Palembang, SMUN 8 Palembang, Prodi IKM FK Unsri, *School of Population Health-The University of Melbourne, Australia* dan tempat menempuh kuliahku sekarang, Auckland University of Technology, Selandia Baru.

Penulis menyadari bahwa masih banyak keterbatasan dalam buku ini dan jauh dari kesempurnaan karena penulis masih proses belajar dan akan terus belajar. Saran dan kritik yang membangun sangat dibutuhkan guna menyempurnakan buku ini, sehingga ke depan dapat menjadi lebih baik, *feel free to send me an email* (najem240783@gmail.com).

Auckland, Maret 2016

Penulis,

Najmah, S.K.M, M.P.H.



# Daftar Isi

TENTANG PENULIS	V
PROFIL EDITOR AHLI BESRAL	VII
PRAKATA	IX
DAFTAR ISI	XI
DAFTAR TABEL	XV
DAFTAR GAMBAR	XVII

## **BAB 1 PENGANTAR STATA 1**

Sekilas Tentang Stata	2
Stata Dasar	2
Tampilan Utama Stata	2
Menyimpan Output Stata di Log File	4
Syntax Stata atau Do file	4
Data Editor	5
Fungsi Syntax Sederhana dalam Stata	7

## **BAB 2 ANALISIS DESKRIPTIF PADA STATA 25**

Aplikasi Syntax pada Analisis Univariat	26
Review Mengenai Stata	26

## **BAB 3 APLIKASI STATA PADA PENGHITUNGAN EPIDEMIOLOGI (OR/RR) 33**

Pengenalan	34
Aplikasi Stata pada Penghitungan Risiko Rasio (Risk Ratio/Relative Risk)	35
Aplikasi Stata pada Penghitungan Odds Ratio	38
Aplikasi Stata pada Penghitungan Rasio Prevalensi (Prevalence Ratio)	40

**BAB 4 REGRESI LOGISTIK SEDERHANA DAN BERGANDA 43**

Aplikasi Stata pada Uji Regresi Logistik Sederhana 45

Aplikasi Stata pada Uji Regresi Logistik Berganda 49

**BAB 5 REGRESI COX (BIVARIAT-MULTIVARIAT) 53**

Analisis Regresi untuk Data Survival Regresi Cox Sederhana dan Berganda 54

Introduksi Konsep Analisis Survival dan Regresi Cox 54

Aplikasi Stata pada Regresi Cox 55

**BAB 6 PENGANTAR SPSS 67**

Icon penting SPSS 68

Memulai SPSS 68

Membuka Data 68

Menu Utama SPSS 69

Membuat Template 71

Entry Data 78

Data Editor 79

**BAB 7 ANALISIS DESKRIPTIF PADA SPSS 85**

Jenis Data dan Skala 86

Analisis Univariat 89

Analisis Deskriptif Data Kategorik 90

Analisis Deskriptif Data Numerik 95

Analisis Data Numerik dengan Kondisi Ada Missing Data Menggunakan Syntax 99

**BAB 8 UJI VALIDITAS DAN RELIABILITAS INSTRUMEN 107**

Tahap Pertama: Uji Validitas 108

Tahap Kedua: Uji Reliabilitas 108

Langkah-Langkah Uji Validitas dan Reliabilitas 110

Analisis 112

Analisis 1: Uji Validitas 112

Analisis 2: Uji Reliabilitas 113

**BAB 9 ANALISIS DATA SURVEI (DATA SEKUNDER) PADA SPSS 117**

Aplikasi Analisis Data Sekunder (SDKI 2012) 118

Proses Pengolahan Data SDKI 2012 121

Langkah 1: Membuat Normalisasi Bobot 121

Langkah 2: Membuat Plan untuk Analisis Pembobotan/Site Plan (Complex Sample) 122



Langkah 3: Analisis dengan Pembobotan 124  
Uji Regresi Linear Complex Sample 126

**BAB 10 REGRESI LOGISTIK SEDERHANA & GANDA DENGAN SPSS 133**

Perbedaan Regresi Logistik Sederhana dan Berganda 134  
Aplikasi Regresi Logistik Sederhana (Aplikasi SPSS) 135

**BAB 11 REGRESI LINEAR SEDERHANA & GANDA (SPSS & STATA) 141**

Regresi Linear Ganda 142  
    Fungsi Regresi Linear Ganda 142  
    Asumsi Regresi Linear 142  
Seleksi Variabel 147

**BAB 12 UJI CHI-SQUARE STATA & SPSS 153**

Uji Bivariat—Uji Hipotesis 154  
Uji Chi-square 155  
Uji Fisher Exact 156  
Aplikasi Uji Chi-square 157  
Chi-square dan Penggabungan Sel 163  
    Uji Hipotesis Tabel B - K Selain 2 - 2 dan 2 - K 163

**BAB 13 UJI INDEPENDENT STUDENT T-TEST Stata & SPSS 169**

Uji Independent Student T-Test 170  
    Uji Beda Rata-Rata Tidak Berpasangan (Uji T Independent/  
        Independent Sample T-Test) 170  
    Aplikasi Uji Student (T test) tidak berpasangan (Independent  
        T-Test) 170

**BAB 14 UJI ANOVA STATA & SPSS 177**

Uji Anova 178  
Aplikasi Uji Anova pada SPSS 178

**BAB 15 PAIRED T-TEST 183**

Uji Beda Rata-Rata Berpasangan/ Paired T-Test (Hipotesis Pre dan Post  
Test) 184

GLOSARIUM STATISTIKA KESEHATAN G-1

INDEKS STATISTIKA KESEHATAN I-1



# Daftar Tabel

TABEL 1.1	Command, Arti, dan Contoh Penulisan di Kolom Commands Stata	7
TABEL 1.2	Fungsi Paru-Paru pada 636 Anak yang Memiliki Keadaan Sosial-Ekonomi yang Kurang di Wilayah Lima, Peru	13
TABEL 1.3	Hasil Data Deskripsi Penelitian (Data Kategori dan Numerik)	16
TABEL 1.4	Ringkasan Variabel FEV <sub>1</sub>	19
TABEL 1.5	Ringkasan Variabel FEV <sub>1</sub>	19
TABEL 1.6	Prevalensi Gejala Penyakit Pernapasan	21
TABEL 1.7	Karakteristik Demografi, Tinggi Badan, FEV1 untuk Sampel Total dan Sampel Bertingkat Berdasarkan Ada-Tidaknya Gejala Penyakit Pernapasan	23
TABEL 2.1	Kode Variabel dan Keterangan Variabel	27
TABEL 2.2	Karakteristik Responden Berdasarkan Jenis Kelamin	29
TABEL 2.3	Karakteristik Responden Berdasarkan Tingkat Pendidikan	29
TABEL 2.4	Command Syntax	30
TABEL 2.5	Karakteristik Kelompok LJASS dan Non LJASS	32
TABEL 3.1	Penghitungan Asosiasi Sederhana	35
TABEL 3.2	Efektivitas Intervensi Terpadu terhadap Perilaku Merokok di Dalam Rumah	35
TABEL 3.3	Hubungan Antara Status Kebiasaan dan Kejadian Patah Tulang Pinggul pada Wanita Lansia di Geelong, Australia	38
TABEL 3.4	Kejadian Near Miss Berdasarkan Paritas Responden	40
TABEL 4.1	Contoh Command Aplikasi Stata pada Uji Regresi Sederhana dan Berganda	44
TABEL 4.2	Hubungan antara Karakteristik Penasun dan Upaya Akses LJASS	45
TABEL 4.3	Perbandingan Karakteristik Penasun Berstatus Memiliki Akses LJASS dan Tidak Memiliki Akses LJASS	49
TABEL 4.4	Hasil Multivariat Karakteristik Pengguna Napza Suntik dalam Mengakses Layanan Jarum dan Alat Suntik Steril (LJASS, 1=Ya, 0=Tidak)	50
TABEL 5.1	Crude Rates ILI (CI 95%) dalam Kelompok yang Mengikuti dan Tidak Mengikuti Penitipan Anak (Childcare)	56
TABEL 5.2	Rate Ratio ILI Membandingkan Kelompok Menggunakan dan Tidak Menggunakan Penitipan Anak	57
TABEL 5.3	Hasil Regresi Mantel-Cox dengan Anak Sebagai Model Variabel	58
TABEL 5.4	Hasil Regresi Mantel-Cox dengan Anak Sebagai Variabel dalam Model (Sebelum Penyesuaian Kovariat Lainnya)	60
TABEL 5.5	Hasil Regresi Cox dengan Status Childcare Anak Disesuaikan dengan Variabel Musim	62
TABEL 5.6	Hasil regresi Cox dengan Perawatan Anak dan Variabel bfcurent dalam Model	64
TABEL 6.1	Karakteristik Responden	78

TABEL 7.1	Contoh Jenis Data dan Skala	86
TABEL 7.2	Skala Pengukuran Tabel	87
TABEL 7.3	Gambaran Karakteristik Responden pada Desa yang Mendapatkan Intervensi dan Non-Intervensi Terpadu	89
TABEL 7.3	Gambaran Karakteristik Responden pada Desa yang Mendapatkan Intervensi dan Non-Intervensi Terpadu (lanjutan)	90
TABEL 7.4	Output Cross Tabulation Pendidikan Responden dan Desa yang Mendapatkan Intervensi dan Non-intervensi	92
TABEL 7.5	Output Cross Tabulation Status Pekerjaan Responden dan Desa yang Mendapatkan Intervensi dan Non-intervensi	93
TABEL 7.6	Gambaran Karakteristik Responden pada Desa yang Mendapatkan Intervensi dan Non-intervensi	94
TABEL 7.7	Uji Normalitas-Kolmogorov-Smirnov Data Numerik (Umur Bapak, Jumlah Anggota Keluarga dan Jumlah Balita)	96
TABEL 7.8	Gambaran Karakteristik Responden pada Desa yang Mendapat Intervensi dan Non-intervensi	98
TABEL 7.9	Gambaran Perilaku Merokok pada Desa yang Mendapat Intervensi dan Non-intervensi	99
TABEL 8.1	Kuesioner Pengetahuan Terkait Rokok	109
TABEL 8.2	Output Uji Validitas dan Reliabilitas tentang Pengetahuan Rokok	113
TABEL 8.3	Output Uji Validitas dan Reliabilitas tentang Pengetahuan Rokok	114
TABEL 9.1	Definisi Operasional Data SDKI 2012	119
TABEL 10.1	Model Parameter Tiap Odds Ratio	135
TABEL 10.2	Analisis Regresi Logistik Sederhana	136
TABEL 10.3	Karakteristik Responden pada Kelompok Patah Tulang dan Tidak Patah Tulang	140
TABEL 11.1	Hasil Uji Bivariabel antara Variabel Independen dan Variabel Dependen	148
TABEL 11.2	Tabel Perubahan Koefisien B Sebelum dan Sesudah Variabel Pengetahuan Dikeluarkan	149
TABEL 11.3	Tahap Pemodelan Multivariabel	150
TABEL 11.4	Laporan Hasil	150
TABEL 12.1	Tabel Uji Statistik pada Analisis Bivariat-Uji Hipotesis yang Sering Digunakan di Bidang Kesehatan	155
TABEL 12.3	Langkah-Langkah Penentuan Uji	157
TABEL 12.4	Output Cross Tabulation MPASI dalam Sehari dan Status Gizi Sampel Berdasarkan BB/TB	161
TABEL 12.5	Laporan Hasil MPASI	163
TABEL 12.6	Hubungan Tingkat Pengetahuan dan Intake Makanan	164
TABEL 13.1	Hubungan Berat Badan Anak dan Status Gizi Kurang	174
TABEL 13.2	Hubungan Usia Anak dan Status Gizi Kurang	175
TABEL 15.1	Laporan Hasil Uji T Dependen Intervensi Terpadu Pengurangan Dampak Buruk (Harm Reduction) Asap Rokok pada Ruangan Tertutup/Ber-AC di Lingkungan Universitas Sriwijaya	187

# Daftar Gambar

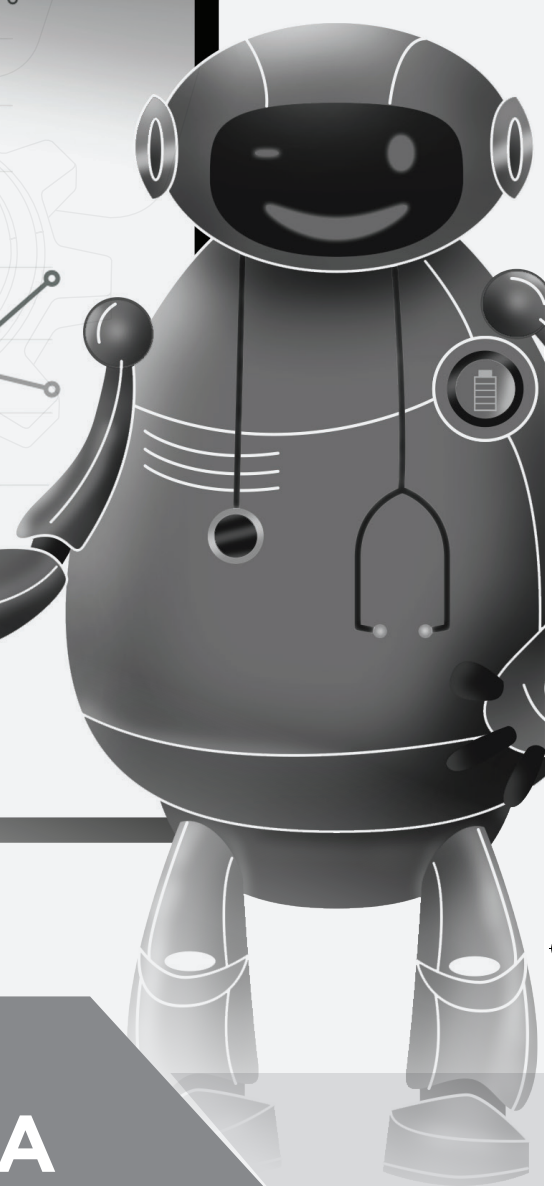
GAMBAR 1.1	Tampilan Utama Stata	3
GAMBAR 1.2	Tampilan Stata	3
GAMBAR 1.3	Membuat Log File	5
GAMBAR 1.4	Membuat Do-file baru	5
GAMBAR 1.5	Mengoperasikan Syntax di Do-file yang Sudah Ada	6
GAMBAR 1.6	Membuka Do-file yang Sudah Ada	6
GAMBAR 1.7	Editor di Stata	6
GAMBAR 5.1	Probabilitas Survival Kumulatif Kaplan-Meier pada Kelompok yang Menggunakan dan Tidak Menggunakan Penitipan Anak pada 234 Anak.	58
GAMBAR 5.2	Nelson-Aalen Perkiraan Cumulative Hazard untuk Variabel Penitipan Anak	59
GAMBAR 5.3	Nelson-Aalen Cumulative Hazard Estimates	61
GAMBAR 5.4	Nelson-Aalen Cumulative Hazard Estimates Bfcurrent	62
GAMBAR 6.1	Proses Memanggil Program SPSS <sup>(12)</sup>	68
GAMBAR 6.2	Tampilan 'Open Sample Files'	69
GAMBAR 6.3	Tampilan Data "template kuesioner_karakteristik responden. sav"	69
GAMBAR 6.4	Menu Utama IBM SPSS Statistics 20	70
GAMBAR 6.5	Tampilan Variabel 1	72
GAMBAR 6.6	Tampilan Variabel 2	73
GAMBAR 6.7	Tampilan Variabel 3	73
GAMBAR 6.8	Tampilan Variabel 4	74
GAMBAR 6.9	Tampilan Variabel 5	75
GAMBAR 6.10	Tampilan Variabel 6	75
GAMBAR 6.11	Tampilan Variabel 7	76
GAMBAR 6.12	Tampilan Variabel 8	77
GAMBAR 6.13	Tampilan Variabel 9	77
GAMBAR 6.14	Proses Membuka Data Kosong	79
GAMBAR 6.15	Import File Excell	80
GAMBAR 6.16	Window Konfirmasi	80
GAMBAR 6.17	Data akan Diganti	81
GAMBAR 6.18	Menyimpan File Data	81
GAMBAR 6.19	Menyimpan File Data	82
GAMBAR 6.20	Menghapus Data dengan Mengklik Kanan pada Mouse	82
GAMBAR 6.21	Copy Data	83
GAMBAR 6.22	Tampilan Fungsi "Edit"	83
GAMBAR 6.23	Output Data	84
GAMBAR 7.1	Langkah 1 "Descriptive Statistics >> Crosstabs"	91
GAMBAR 7.2	Langkah 2 "ROW dan COLOUMN"	91
GAMBAR 7.3	Langkah 3	91
GAMBAR 7.4	Langkah 1 Uji Normalitas	95

GAMBAR 7.5	Langkah 2 Uji Normalitas	96
GAMBAR 7.6	Langkah 3 Uji Normalitas	96
GAMBAR 7.7	Nilai Median, Minimum, dan Maksimum Variabel Numerik Dibagi Berdasarkan Status Intervensi	98
GAMBAR 7.8	Langkah 1 Seleksi Kasus yang Diperlukan	100
GAMBAR 7.9	Langkah 2 Seleksi Kasus	100
GAMBAR 7.10	Langkah 3 Seleksi Kasus	101
GAMBAR 7.11	Aplikasi Syntax pada SPSS	102
GAMBAR 8.1	Proses Uji Validitas dan Reliabilitas	111
GAMBAR 8.2	Kotak Dialog Items	111
GAMBAR 8.3	Kotak Dialog Option Statistic	112
GAMBAR 9.1	Open Data ADS	119
GAMBAR 9.2	Langkah Go to Variable ADS	120
GAMBAR 9.3	Langkah Copy Data pada Kode V005	120
GAMBAR 9.4	Panduan Normalisasi Bobot	121
GAMBAR 9.5	Langkah Compute Variabel	122
GAMBAR 9.6	Langkah Analyze Site Plan (Pembobotan)	122
GAMBAR 9.7	Langkah Create a New File untuk Plan	123
GAMBAR 9.8	Input Variabel Strata, Cluster, dan Bobot Normal	123
GAMBAR 9.9	Langkah Analyze Frequencies Complex Samples	124
GAMBAR 9.10	Input Plan yang sudah dibuat pada Analyze Frequencies	124
GAMBAR 9.11	Complex Sample Frequencies Analysis dan Statistik	125
GAMBAR 9.12	Output Complex Samples	125
GAMBAR 9.13	Complex Sample Descriptives	126
GAMBAR 9.14	Langkah Complex Samples General Linear Model	127
GAMBAR 9.15	Kotak Browse File Plan Complex Samples	127
GAMBAR 9.16	Kotak Input Variabel Dependen dan Independen CSGLM	128
GAMBAR 9.17	Langkah Complex Samples General Linear Model	130
GAMBAR 9.18	Kotak Browse File Plan Complex Samples	130
GAMBAR 9.19	Kotak Input Variabel Dependen dan Independen CSGLM	130
GAMBAR 9.20	Output Variabel Dependen dan Independen CSGLM	131
GAMBAR 10.1	Proses Analisis Regresi Logistik	136
GAMBAR 10.2	Kotak Dialog Regresi Logistik	137
GAMBAR 10.3	Kotak Dialog Options Regresi Logistik	137
GAMBAR 11.1	Langkah Analyze Regresi Linear Berganda	144
GAMBAR 11.2	Input Variabel Dependen dan Independen	144
GAMBAR 11.3	Regresi Linear: Statistics	145
GAMBAR 11.4	Regresi Linear: Plots	145
GAMBAR 11.5	Asumsi Homoscedasticity Regresi Linear Ganda	146
GAMBAR 11.6	Asumsi Linearitas Regresi Linear Ganda	146
GAMBAR 11.7	Asumsi Normalitas Data Regresi Linear Ganda	147
GAMBAR 11.8	Pengujian Kolinearitas Regresi Linear Berganda	147
GAMBAR 11.7	Asumsi Normalitas Data Regresi Linear Ganda	147
GAMBAR 11.9	Output Regresi Linear Ganda Full Model	148
GAMBAR 11.10a	Output Regresi Linear Ganda Model I	149
GAMBAR 11.10b	Output Regresi Linear Ganda Model I	151
GAMBAR 11.11	Output Regresi Linear Ganda Model II	152
GAMBAR 11.12	Output Regresi Linear Ganda Model III	152
GAMBAR 12.1	Jenis Penelitian secara Garis Besar	154
GAMBAR 12.2	Diagram Alur Uji Hipotesis Variabel Kategorikal Kelompok Tidak Berpasangan	157

GAMBAR 12.3	Proses Analisis <i>Chi-Square</i>	159
GAMBAR 12.4	Tampilan " <i>Crosstabs</i> "	160
GAMBAR 12.5	Tampilan Kolom Statistics pada " <i>Crosstabs</i> "	160
GAMBAR 12.6	Tampilan Kolom Cells pada " <i>Crosstabs</i> "	161
GAMBAR 12.7	Proses Pengkodean Variabel Baru	165
GAMBAR 12.8	Kotak Dialog "Recode into Different Variables"	165
GAMBAR 12.9	Kotak Dialog "Recode into Different Variables: Old and New Values"	166
GAMBAR 13.1	Proses Analisis T-Test	171
GAMBAR 13.2	Proses Pemilihan Independent-Sample T-Test	172
GAMBAR 13.3	Output Data Independent Samples T-Test BB Anak dan Status Gizi	172
GAMBAR 13.4	Proses Pemilihan Independent Samples T-Test	174
GAMBAR 13.5	Output Data Independent Samples T-Test Usia Awal Pemberian MPASI dan Status Gizi	174
GAMBAR 13.6	Uji Normalitas Kolmogorov-Smirnov Usia Awal Pemberian MPASI	175
GAMBAR 14.1	Proses Compare Means One Way Anova	179
GAMBAR 14.2	Kotak Dialog One Way Anova	180
GAMBAR 14.3	Kotak Dialog One Way Anova Bonferroni	180
GAMBAR 14.4	Kotak Dialog One Way Anova Options	181
GAMBAR 14.5	Output SPSS Uji Anova	181
GAMBAR 15.1	Langkah Analyze Paired-Sample T-Test	185
GAMBAR 15.2	Kotak Input Variabel Pengetahuan (Pre dan Post)	186
GAMBAR 15.3	Output SPSS Uji T Dependen Variabel Pengetahuan	186
GAMBAR 15.4	Output SPSS Uji T Dependen Variabel Sikap	186







# BAB 1

# PENGANTAR STATA

**Kompetensi Dasar**

Mampu menjelaskan aplikasi dasar statistik kesehatan dengan aplikasi Stata.

**Indikator Keberhasilan**

Mampu menjelaskan teori dasar Stata dan aplikasi dasar statistik kesehatan dengan *software* Stata.

**Materi Pembelajaran**

1. Sekilas Stata
2. Tampilan Stata dasar dan fungsi dasar Stata
3. Fungsi syntax sederhana dalam Stata
4. Aplikasi syntax sederhana dalam Stata

## SEKILAS TENTANG STATA

*Software* statistik Stata adalah paket *software* lengkap dan terintegrasi yang dapat memberikan apa pun yang kita butuhkan dalam menganalisis data, manajemen data, dan grafik. Stata merupakan perangkat lunak statistik yang berlisensi. Kita dapat membeli Stata melalui situs Web resminya, [www.stata.com](http://www.stata.com). Stata tidak dijual dalam bentuk modul terpisah, dan berarti kita mendapatkan semua itu dalam satu paket. Keunggulan Stata adalah cepat, akurat, dan mudah digunakan baik dengan aplikasi *tool bar* menu maupun *syntax*. Proses analisis data (*syntax* dan *output*) dapat didokumentasikan dalam bentuk *Do-file* dan *Log file* sehingga kita dapat melakukan analisis ulang pada data yang sama. Dokumentasi *syntax* pada *Do-file* dapat digunakan untuk analisis data lainnya dengan mengganti variabel pada data lainnya. Misal, pada tahun 2011 kita telah mengaplikasikan regresi sederhana pada data penelitian Pengguna Narkoba Suntik, dan kita telah menyimpan *syntax* pada *Do-file* di Stata. Kemudian, ketika kita ingin menganalisis data penelitian Narkoba Suntik pada tahun 2015 dengan analisis regresi sederhana, kita tinggal membuka data *Do-file* sebelumnya, dan mengganti variabel lama dengan variabel yang baru tanpa mengubah *syntax* yang ada.

Berbagai aplikasi statistik kesehatan yang dapat kita lakukan dengan Stata, antara lain: (1) analisis deskripsi sederhana, seperti menghitung rata-rata, nilai tengah, standar deviasi, derajat kepercayaan 95%, tabulasi; (2) analisis bivariat, seperti analisis *Chi-Square*, analisis *Student-T*, Anova/Manova, regresi linear, dan logistik sederhana; (3) aplikasi penghitungan epidemiologi seperti *risk ratio*, *odds ratio*; (4) penghitungan sampel, untuk beberapa uji proporsi, uji rata-rata, dan sebagainya hingga; (5) analisis lanjut seperti regresi linear dan logistik berganda, analisis survival, dan analisis berbagai teknik statistik lainnya.

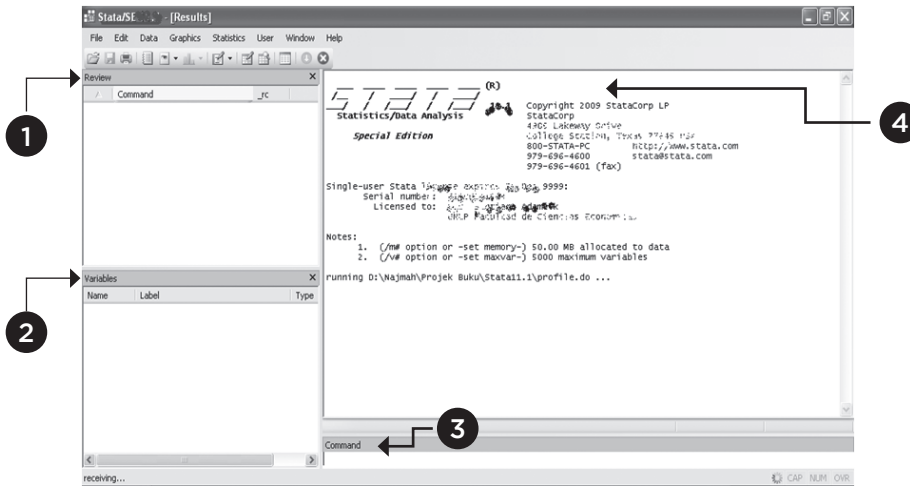
## STATA DASAR

### TAMPILAN UTAMA STATA

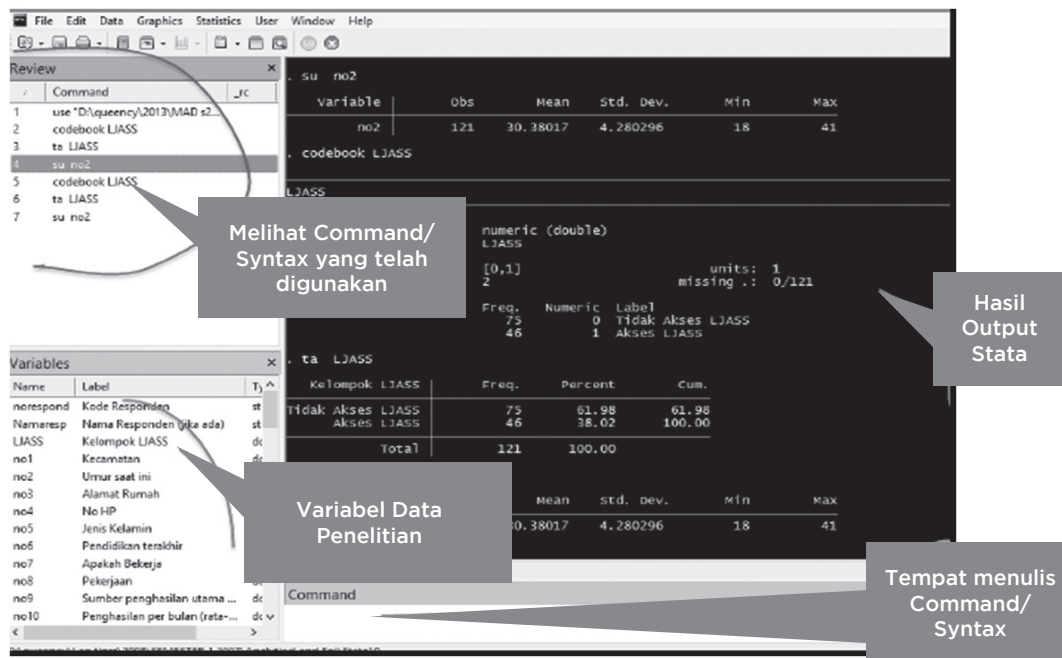
Berikut ini adalah tampilan utama ketika kita membuka aplikasi Stata. Ada *Drop Down Menu*, *File*, *Edit*, *Data*, *Graphics*, *Statistic*, *User*, *Window*, dan *Help*. Pada layar Stata ada beberapa tampilan jendela yang perlu kita ketahui seperti berikut ini.

- Jendela **Review**, untuk melihat kumpulan *command* yang berisi *syntax* yang sudah kita jalankan ketika mengolah data. Pada Gambar 1.1 ditandai dengan No. 1.
- Jendela **Variables**, yang memberi penjelasan data yang akan kita olah. Kolom ini terdiri atas nama variabel, label, atau penjelasan variabel tertentu, dan tipe data variabel. Pada Gambar 1.1 ditandai dengan No. 2.
- Jendela **Command**, tempat melakukan operasi Stata. Pada kolom ini, kita mengetik kode *syntax* yang akan kita operasikan, lalu tekan enter untuk melihat hasil operasi *syntax* pada layar Stata. Pada Gambar 1.1 ditandai dengan No. 3.

● **GAMBAR 1.1** Tampilan Utama Stata



● **GAMBAR 1.2** Tampilan Stata



■ **Jendela Stata**, pada layar ini kita melihat hasil *output* statistik dari operasi *syntax* yang kita lakukan, berupa tabel hasil olah statistik, baik analisis deskriptif maupun analitis. Pada Gambar 1.1 ditandai dengan No. 4.

## MENYIMPAN OUTPUT STATA DI LOG FILE

Ketika kita ingin menyimpan hasil analisis data atau *output* Stata, kita bisa menyimpannya dalam bentuk *log file*. Ada dua tipe *log file* yang bisa kita pilih, yaitu **Log** dan **Formatted Log**.

### 1. Log

Format *Log* ini lebih sering digunakan untuk menyimpan *output* Stata. Hasil *output* yang disimpan dalam *Log* adalah berupa *text file*, sehingga data *Log* bisa langsung dibuka dari program MsWord. Perintah *Log* dilakukan sebelum memulai analisis data, lalu menutup *Log* ketika analisis data telah selesai.

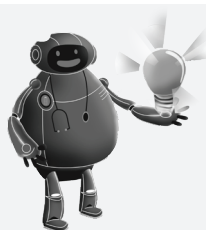
- Langkah Awal: klik **File** → **Log** → **Begin**, tulis nama file tempat kita menyimpan *output*, misal “Coba Coba”, dan pilih **Log** pada tipe data (*save a type*), lalu klik **Save**.
- Lakukan proses analisis data, *output* yang muncul di layar otomatis juga akan tersimpan ke file “Coba Coba”.
- Setelah selesai menganalisis data di Stata, kita tutup *log file*. Proses menutup log, klik **File** → **log** → **Close**.
- Setelah kita klik **Close**, kita dapat membuka hasil *output* Stata di folder tempat kita menyimpan. Kita bisa membukanya langsung dari folder tersebut, atau jika kita ingin data kita dalam aplikasi MsWords dan menginterpretasikannya langsung, kita bisa membuka hasil *output* di file “Coba Coba”, dari program MsWord pilih file → **Open**, dan **select all** pada tipe data agar data dalam bentuk log bisa terlihat, lalu pilih file “Coba Coba”.

### 2. Formatted Log (.smcl)

Format ini dipilih jika kita ingin membuka hasil *output* Stata dari program Stata langsung, kita menyimpan *output*-nya dalam *Formatted Log (.smcl)*. Prosesnya sama seperti penjelasan Log sebelumnya, hanya pada *save a type*, kita pilih **Formatted Log (\*.smcl)**. Hasil *output* yang tersimpan pada file smcl ini tidak dapat dibuka dengan MsWords, hanya dapat dibuka oleh program Stata.

## SYNTAX STATA ATAU DO-FILE

Menyimpan *syntax* sangat bermanfaat terutama ketika kita ingin menganalisis ulang data kita dengan variabel berbeda ataupun menganalisis data lainnya dengan analisis statistik yang sama. *Do-file* merupakan file untuk menyimpan *syntax* aplikasi Stata kita. Jika kita lupa kode *syntax* tertentu, dengan melihat kumpulan *syntax* di *Do-file* yang ada, kita akan bisa menelusuri kembali kode-kode tersebut.

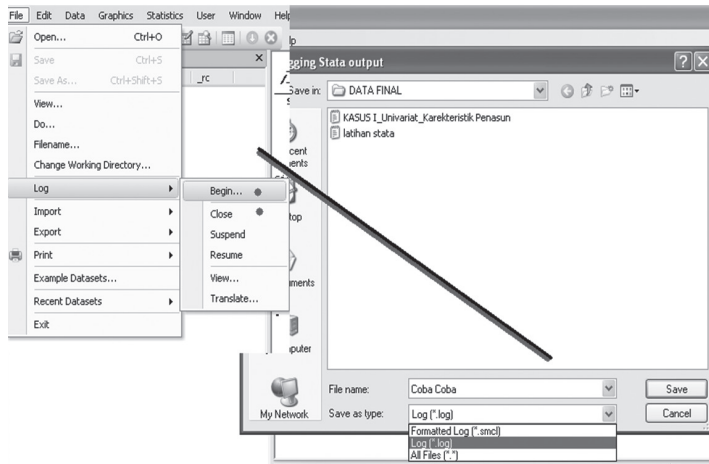


### CATATAN:

Kita juga dapat menyalin (*copy paste*) output stata ke MsWord atau ke MsExcel dengan mudah. *Block output* yang ingin disalin, kemudian klik kanan, pilih “*copy table*”, kemudian “*paste*” di MsWord atau di MsExcel.

Cara ini dikerjakan satu per satu per tabel, tidak bisa banyak tabel sekaligus dalam sekali *copy*.

● **GAMBAR 1.3** Membuat Log File



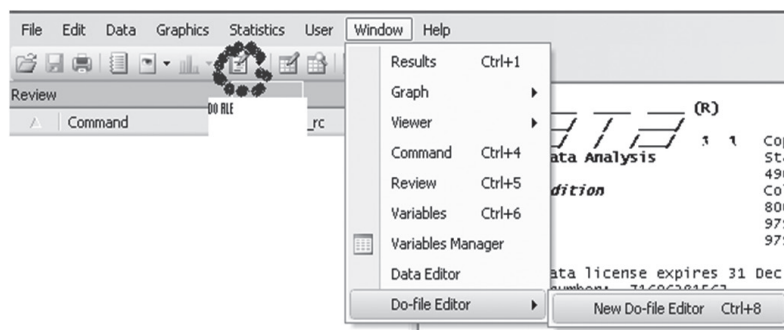
Tahap awal adalah membuat satu *Do-file*, selanjutnya kita simpan dengan nama tertentu, dalam bentuk *Do-file* (file dengan *ekstension do*), lalu kita bisa *copy* setiap operasi *syntax* yang kita jalankan pada jendela commands atau di layar Stata ke *Do-file* tersebut.

- Proses membuat *Do-file* baru: klik **Windows** → **Do-file Editor** → **New Do-file Editor** → **Save As** (ketik nama *file*) atau klik gambar *Do-file* di *menu bar*. Atau, pada tahap awal kita bisa menyetik semua *syntax* untuk proses statistik yang kita perlukan, dan tinggal kita klik **Execute(run)** pada *toolbars Do-file*. Perhatikan Gambar 1.4.
- Proses: sorot *syntax* yang telah kita ketik → lalu klik tombol **execute(run)**, lihat hasil di layar Stata. Perhatikan Gambar 1.5.
- Proses membuka *Do-file* yang telah ada: klik **File** → **Open** → klik *Do-file* pada **Files of Type** di bawah **File Name** → lalu pilih satu *Do-file* yang diinginkan. Perhatikan Gambar 1.6.

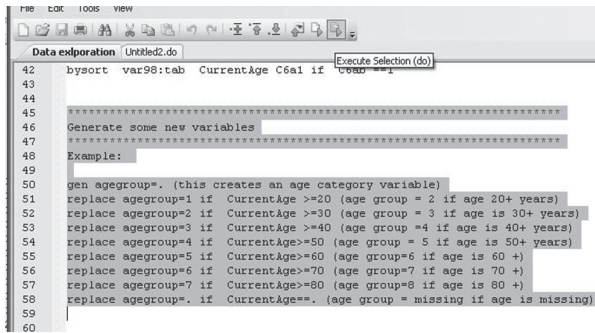
## DO-FILE EDITOR

Kita bisa melihat data dengan mengklik data editor pada *toolbars*. Lihat Gambar 1.7.

● **GAMBAR 1.4** Membuat Do-file baru



### ● GAMBAR 1.5 Mengoperasikan Syntax di Do-file yang Sudah Ada

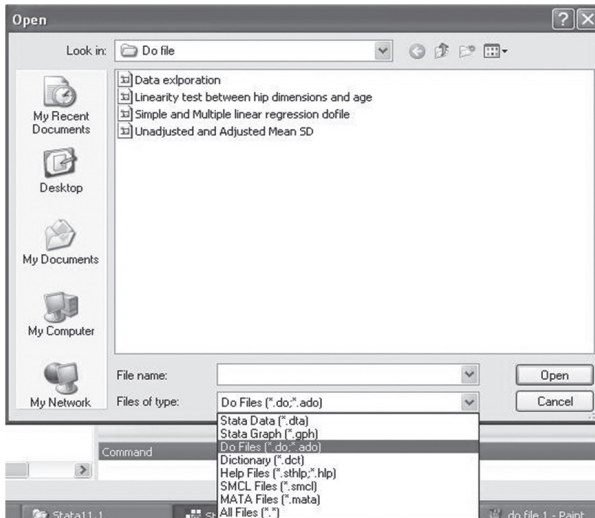


```

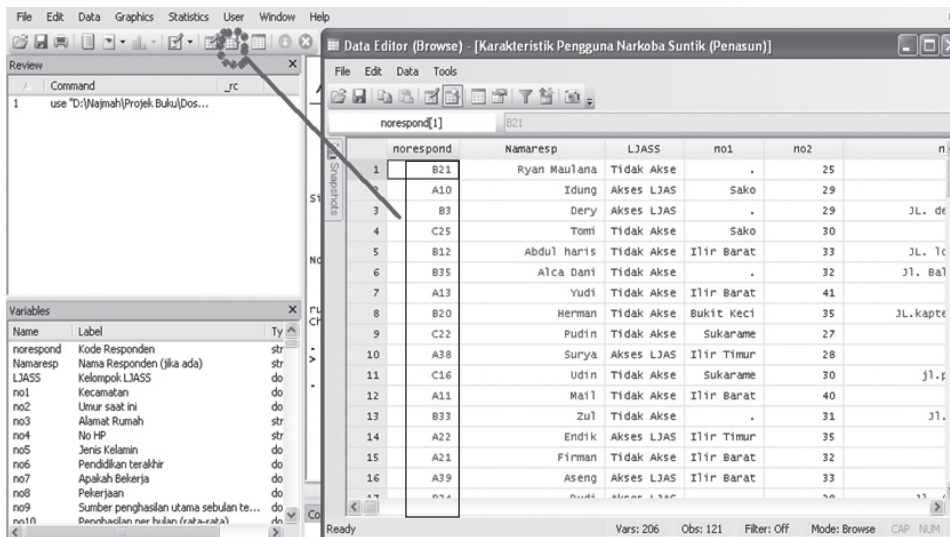
42  bsort var98:tab CurrentAge C6a1 if CurrentAge >= 1
43
44
45
46  Generate some new variables
47  .....
48  Example:
49
50  gen agegroup=. (this creates an age category variable)
51  replace agegroup=1 if CurrentAge >=20 (age group = 2 if age 20+ years)
52  replace agegroup=2 if CurrentAge >=30 (age group = 3 if age is 30+ years)
53  replace agegroup=3 if CurrentAge >=40 (age group = 4 if age is 40+ years)
54  replace agegroup=4 if CurrentAge >=50 (age group = 5 if age is 50+ years)
55  replace agegroup=5 if CurrentAge >=60 (age group=6 if age is 60+)
56  replace agegroup=6 if CurrentAge >=70 (age group=7 if age is 70+)
57  replace agegroup=7 if CurrentAge >=80 (age group=8 if age is 80+)
58  replace agegroup=. if CurrentAge==. (age group = missing if age is missing)
59
60

```

### ● GAMBAR 1.6 Membuka Do-file yang Sudah Ada



### ● GAMBAR 1.7 Editor di Stata



The Data Editor window shows the following dataset:

norespond	Namaresp	LJASS	no1	no2	n
1	B21	Ryan Maulana	Tidak Akse	.	25
2	A10	Idung	Akses LJAS	Sako	29
3	B3	Dery	Akses LJAS	.	29
4	C25	Toml	Tidak Akse	Sako	30
5	B12	Abdul haris	Tidak Akse	Iilir Barat	33
6	B35	Alca Dani	Tidak Akse	.	32
7	A13	Yudi	Tidak Akse	Iilir Barat	41
8	B20	Herman	Tidak Akse	Bukit Keci	35
9	C22	Pudin	Tidak Akse	Sukaname	27
10	A38	Surya	Akses LJAS	Iilir Timur	28
11	C16	Udin	Tidak Akse	Sukaname	30
12	A11	Mai1	Tidak Akse	Iilir Barat	40
13	B33	Zul	Tidak Akse	.	31
14	A22	Endik	Akses LJAS	Iilir Timur	35
15	A21	Firman	Tidak Akse	Iilir Barat	32
16	A39	ASeng	Akses LJAS	Iilir Barat	33

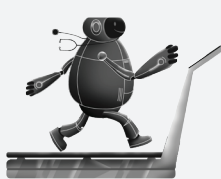
The 'Variables' list on the left includes: norespond (Kode Responden), Namaresp (Nama Responden (jika ada)), LJASS (Kelompok LJASS), no1 (Kecamatan), no2 (Umur saat ini), no3 (Alamat Rumah), no4 (No HP), no5 (Jenis Kelamin), no6 (Pendidikan terakhir), no7 (Apakah Bekerja), no8 (Pekerjaan), no9 (Sumber penghasilan utama sebulan te...), and no10 (Danharsian nar bahan (rak-arab)).




# FUNGSI SYNTAX SEDERHANA DALAM STATA

Berikut adalah beberapa aplikasi *syntax* dasar yang perlu dipahami dan diaplikasikan, mari kita perhatikan satu per satu, *command*, arti, dan contoh penulisan di kolom *Commands* di Stata.

## LATIHAN:



Buka  **Latihan Stata 1\_Penasun\_1.dta**, aplikasikan contoh pada Tabel 1.1 dengan data “**Latihan Stata 1\_Penasun\_1.dta**”, lalu buatlah *Do-file* untuk menyimpan *syntax* untuk operasi di bawah ini, silakan mengeksplorasi dengan menggunakan variabel yang berbeda.

● **TABEL 1.1** Command, Arti, dan Contoh Penulisan di Kolom Commands Stata

Command	Arti	Contoh dalam Kolom Command
<b>, missing</b>	Termasuk nilai yang hilang dalam data yang telah dihasilkan  Nilai hilang akan dilambangkan dengan titik “.” untuk data numerik dan “[spasi kosong]” untuk nilai string/huruf	tab LJASS, missing tab no9, missing  ta no 9, missing atau ta LJASS ta no9 (silakan bedakan hasilnya jika kita letakkan kata ‘missing’ dan tanpa ‘missing’)
<b>= =</b>	Sama dengan dua buah	tab [variabel yang akan dimasukkan] if [variabel yang digunakan untuk mengecek] = = Akses LJASS berdasarkan tingkat pendidikan(no6) tab no6 if LJASS==0 tab no6 if LJASS==1  Tabel frekuensi Akses LJASS berdasarkan jenis kelamin(no5) tab LJASS if kelamin==1 tab LJASS if kelamin==2
<b>By (Sort)</b>	Memberikan informasi mengenai dua variabel  Sering membutuhkan “sort” pada kolom command	by [variabel 1]: command [variabel 2]  sort sort [variabel 1] by [variabel 1]:command [variabel 2]  perilaku penasun (no44_1) disortir berdasarkan umur (no2)  sort no44_1 by no44_1 :sum no2

Command	Arti	Contoh dalam Kolom Command
<b>Clear</b>	Menghapus data dari file Stata	Clear [variabel 1]  Clear no2
<b>Close</b>	Tutup	Menutup perintah save <i>output</i>  Log close
<b>Codebook</b>	Memberikan informasi umum di mana setiap variabel dalam data yang telah dibuat  Type: tipe data Label: keterangan Range: kisaran data (nilai min-nilai max) Unique values: nilai Units: unit Missing values: nilai yang hilang Frequency: frekuensi Value labels: keterangan nilai	codebook [nama variabel]  codebook LJASS  codebook no44_1  codebook no2 if no44_1==0 codebook no2 if no44_1==1
<b>Decode</b>	Untuk mentransformasikan data numerik ke dalam bentuk teks berdasarkan variabel string.  Harus memiliki nilai pada keterangan.	decode variabel lama, gen(variabel baru)  decode LJASS, gen(perilaku_penasun)
<b>Describe</b>	Memberikan deskripsi atau gambaran tentang variabel  Variable name: nama variabel Storage type (numeric or string): Tipe penyimpanan (numerik atau string) Display format: format tampilan Value label: keterangan nilai Variable label: keterangan variabel	Describe describe [nama variabel]  describe LJASS describe no2
<b>Drop</b>	Digunakan untuk menghapus variabel atau hasil	drop [variabel]  atau  drop [variabel 1] [variabel 2]  drop no6
<b>Encode</b>	Digunakan untuk mengambil variabel dalam format teks dan mengubahnya ke dalam nilai numerik (kebalikan dari <i>Decode</i> )	encode [variabel lama], gen ([variabel baru])  encode perilaku_penasun, gen(LJASS_1)
<b>Exit</b>	Menutup atau keluar dari program Stata	Exit



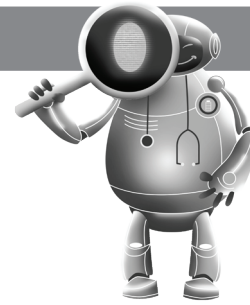
Command	Arti	Contoh dalam Kolom Command
<b>Generate</b>	Digunakan untuk membuat variabel baru.  Sering diikuti dengan perintah <i>'replace'</i> untuk mengategorikan variabel	generate [variabel baru] = [hasil yang diinginkan]  contoh 1: menghitung total pengetahuan dan mengelompokkan pengetahuan menjadi dua kelompok (mean 12)  generate SUM_TAHU= NO18ANEW + NO18BNEW + NO18CNEW +NO18DNEW + NO18ENEW + NO18FNEW+ NO18GNEW + NO18HNEW + NO18INEW + NO18JNEW + NO18KNEW + NO19ANEW + NO19BNEW + NO19CNEW + NO19DNEW + NO19ENEW + NO19FNEW+NO19GNEW  sum SUM_TAHU  generate tahukel=. replace tahukel=0 if SUM_TAHU<12 replace tahukel=1 if SUM_TAHU>=12
<b>Insheet</b>	Mengambil sebuah file TEXT (txt) (biasanya dikonversikan dari program excel)	Memperoleh sebuah data file teks Pada lembar kerja, gunakan "nama file"  Contoh "c:\my file.txt"
<b>Keep</b>	Digunakan untuk menyimpan variabel yang akan tetap digunakan dalam sebuah set data dan akan MENGHAPUS semua variabel lain/ hasil pengamatan selain variabel yang akan disimpan	keep [variabel 1] [variabel 2]  keep no6
<b>List</b>	Menghasilkan daftar keterangan nilai berdasarkan variabel yang sedang diamati Biasanya menggunakan sebuah pernyataan "if"	list [variabel]  list no responden dengan pendidikan SD (no6, SD=1)  list norespond if no6==1  list no responden dengan pendidikan SD (no6, SMP=3)  list norespond if no6==1
<b>Log</b>	Mengatur sebuah file yang isinya semua informasi yang baru saja diolah	Penggunaan log [nama file.log] Gunakan "close log" untuk menutup log Contoh: Log begin "hasil.txt" Log close.
<b>Mvdecode</b>	Digunakan untuk menggantikan nilai yang hilang (missing data) yang tidak memiliki kode.  Sebagai contoh keterangan "." berarti "9"	mvdecode [nama variabel], mv(nilai kode yang hilang)  misal :  mvdecode sex, mv(9)

Command	Arti	Contoh dalam Kolom Command
<b>Recode</b>	Digunakan untuk mengubah pengelompokan nilai numerik.  Sering digunakan dengan kode yang dihasilkan.  Yakni dari 0 & 1 untuk nilai 1 dan 2 untuk nilai 1  Biasanya dimulai dengan perintah GENERATE	<code>recode ([nomor lama] = [nomor baru]) ([nomor lama] = [nomor baru])</code>  misal: pendidikan (no6) 1= SD 3=SMP 4=SMU 5=D3 6=Univ 7=lainnya  diganti: 1=pendidikan rendah (SD&SMP) 2= pendidikan menengah (SMA) 3= pendidikan tinggi (D3, Univ)  <code>gen didikgroup=no6 recode didikgroup (1=1)(3=1)(4=2)(5=3) (6=3)(7=3)</code>
<b>Rename</b>	Digunakan untuk mengubah nama sebuah variabel  Sebagai contoh variabel 1 menjadi jenis kelamin	<code>rename [nama variabel lama] [nama variabel baru]</code>  <code>rename var1 jenis kelamin</code>
<b>Replace</b>	Digunakan untuk bentuk kategori dalam sebuah variabel.  Biasanya dimulai dengan perintah GENERATE Biasanya diikuti dengan perintah IF	<code>replace [nama variabel] = [nomor] if [nama grup][kondisi]</code>  Contoh: mengubah umur dari numerik menjadi kategorik <code>gen Agegrp= replace Agegrp=1 if CurrentAge &lt;70 replace Agegrp=2 if CurrentAge &gt;=70 replace Agegrp=3 if CurrentAge &gt;=75 replace Agegrp=4 if CurrentAge &gt;=80 replace Agegrp=5 if CurrentAge &gt;=85</code>
<b>Sort/by</b>	Digunakan untuk mensortir data yang akan diolah berdasarkan variabel lain.	<code>Sort [variabel yang akan disortir] By [variabel untuk penyortiran]: command [variabel hasil]</code>  <code>sort jeniskelamin by jeniskelamin: su umur</code>  <code>sort no44_1 by no44_1 :ta SIKAP</code>
<b>Summarise</b>	Memberikan ringkasan informasi statistik, biasanya kita gunakan untuk variabel numerik  Catatan: bisa meringkas lebih dari satu variabel dalam satu waktu  Obs: observasi; Mean: rata-rata; SD: standar deviasi; Min: nilai terkecil Max: nilai terbesar;	<code>summarize [nama variabel]</code>  <code>umur (no2) dan jumlah anak (no13) su no2 no13</code>

Command	Arti	Contoh dalam Kolom Command
<b>Tabulate</b>	<p>Memberikan data proporsi/ prevalensi</p> <p>Bisa membandingkan 2 variabel satu sama lain tapi hanya memberikan data frekuensi.</p> <p>Frequency: frekuensi Percentages: persentase Cumulative percentages: persentase kumulatif</p> <p>Penulisan dapat disingkat hanya 2 atau 3 huruf (TA atau TAB atau TAB1)</p>	<p>tabulate [nama variabel]</p> <p>ta didikgroup</p> <p>*JENIS KELAMIN</p> <p>tab no5 * AGAMA tab no11</p> <p>tab no5 LJASS</p> <p>Jika TA atau TAB hanya boleh satu variabel, sedangkan TAB1 boleh banyak variabel. Contoh: TAB didik TAB1 didik umur sosek dst</p>
<b>Value label</b>	<p>Hasil deksripsi yang diterapkan pada kelompok deskripsi untuk nilai numerik yang ekuivalen dalam bentuk data kategorik (kelompok lanjut)</p>	<p>la define [nama daftar] 1 "label 1" 2 "label 2" dst. label values [nama variabel] [nama untuk variabel]</p> <p>gunakan dalam urutan: la define daftarpendidikan 1"ya" 2"tidak" la values pendidikan daftar pendidikan</p> <p>Contoh Misal Variabel Pendidikan</p> <p>Kategori pendidikan dijadikan 3 kelompok saja 3 groups:(koding 3&lt;=SMP, koding 4 SMA, koding 5&gt;=D3)</p> <p>generate no6_1=no6 label define didik1 3"&lt;=SMP" 4"SMA" 5"&gt;=D3" label value no6_1 didik1 ta no61_1</p>

## STUDI KASUS 1.1

Latihan dasar stata deskriptif di bidang kesehatan. Data yang diberikan berikut ini termasuk data dari sebuah penelitian *cross sectional* mengenai fungsi paru-paru pada 636 anak yang memiliki keadaan sosial-ekonomi yang kurang di wilayah Lima, Peru. Variabel-variabel yang diikutkan dapat dilihat pada Tabel 1.2.



Buka Data: FEV data.dta

Sumber Data: Bennett, Cathrenie. 2008. *Kumpulan latihan Soal Stata, Epidemiology and Analytic Method I. School of Population Health The University of Melbourne.*

### PERTANYAAN DETEKTIF STATA



1. Buka file **FEV data.dta**,
2. Tentukan jumlah observasi, jumlah variabel, dan apa saja variabelnya!
3. Tambahkan “keterangan variabel” pada semua variabel yang ada sebagai berikut.
  - a. label var id “Subject ID Number”
  - b. label var fev1 “Forced Expiratory Volume”
  - c. label var age “Subject’s age”
  - d. label var height “Subject’s height”
  - e. label var sex “Subject’s gender”
  - f. label var respsymptoms “Respiratory symptoms last 12 months”
4. Buat *value label* untuk jenis kelamin dan *resymptom* sebagai berikut.
  - a. Perempuan=0 Laki-laki=1
  - b. *Respsymptoms* 0=no 1=yes
5. Bedakan tipe variabel (kategorik, diskrit, nominal, ordinal, atau lainnya), lalu ringkasan statistik (mean, median, frekuensi) dan pilih grafik (histogram, grafik batang, dll.) yang memungkinkan kita meringkas data dengan mudah.
6. Tentukan berapa banyak data hilang pada setiap variabel?
7. Buatlah daftar nomor ID responden yang memiliki nilai  $FEV_1$  lebih besar dari 2,4 liter!
8.
  - a. Hitung dan deskripsikan dengan singkat variabel  $FEV_1$  berdasarkan variabel jenis kelamin (nilai mean, median, standar deviasi, minimum, maksimum).
  - b. Apakah perkiraan ini dapat digunakan sebagai kesimpulan rata-rata statistik dari variabel  $FEV_1$ ?
9.
  - a. Buatlah grafik hubungan antara umur dengan fungsi paru!
  - b. Deskripsikan hubungan antara dua variabel ini. Bagaimana fungsi paru-paru berdasarkan umur?
  - c. Pada grafik tersebut, variabel mana yang bisa Anda deskripsikan sebagai variabel yang menggambarkan *outcome*/hasil dan mana yang menjelaskan paparan? Berikan alasan!

● **TABEL 1.2** Fungsi Paru-Paru pada 636 Anak yang Memiliki Keadaan Sosial-Ekonomi yang Kurang di Wilayah Lima, Peru

Nama Variabel	Keterangan	Nilai
<b>Id</b>	Nomor responden	n/a
<b>FEV<sub>1</sub></b>	Volume ekspirasi maksimum*	n/a
<b>Age</b>	Umur	n/a
<b>Height</b>	Tinggi badan	n/a
<b>Sex</b>	Jenis kelamin	0 = perempuan 1 = laki-laki
<b>Respsymptoms</b>	Masalah pernapasan dalam 12 bulan terakhir	0 = tidak 1 = ya

\* volume udara maksimal yang bisa dihembuskan dalam 1 detik  
n/a—nilai yang diisi berupa angka

10. Berapa prevalensi yang mengalami gejala penyakit pernapasan pada penelitian ini? Hitung juga prevalensi berdasarkan jenis kelamin responden.
11. Siapkan sebuah tabel karakteristik demografi, tinggi badan, dan FEV1 untuk total responden, dan juga berdasarkan ada tidaknya gejala penyakit pernapasan.

### JAWABAN DETEKTIF STATA



#### Pertanyaan No. 1:

- Buka  FEV data.dta.

#### Pertanyaan No. 2:

Menentukan angka observasi pada data. Jumlah observasi = 636

#### Syntax: describe

```
. describe
contains data
obs:      636
vars:      6
size:     12,720 (98.8% of memory free)
-----
variable name  storage  display  value  variable label
                type    format   label
-----
id             int      %8.0g
fev1           float   %9.0g
age            float   %9.0g
height        float   %9.0g
sex            byte    %8.0g
respsymptoms  byte    %8.0g
-----
Sorted by:
Note: dataset has changed since last saved
```

**Pertanyaan No. 3:**

Gunakan keterangan data yang ada pada tabel sebelumnya, tambahkan keterangan variabel pada keenam variabel yang ada pada data.

**Syntax:**

```
. label var id "Subject ID Number"
. label var fev1 "Forced Expiratory Volume"
. label var age "Subject's age"
. label var height "Subject's height"
. label var sex "Subject's gender"
. label var respsymptoms "Respiratory symptoms last 12 months"

. label var id "Subject ID Number"
. label var fev1 "Forced Expiratory Volume"
. label var age "Subject's age"
. label var height "Subject's height"
. label var sex "Subject's gender"
. label var respsymptoms "Respiratory symptoms last 12 months"
```

**Pertanyaan No. 4:**

Gunakan keterangan data yang berada pada tabel sebelumnya, tambahkan keterangan variabel jenis kelamin dan gejala.

**Syntax:**

```
. label define sex 0"female" 1"male"
. label define respsymptoms 0"no" 1"yes"
. label value sex sex
. label value respsymptoms respsymptoms
*****
. label define sex 0"female" 1"male"
. label define respsymptoms 0"no" 1"yes"
. label value sex sex
. label value respsymptoms respsymptoms
-----
```

**Pertanyaan No. 5:**

Bedakan tipe variabel (kategorik, diskrit, nominal, ordinal, atau lainnya), lalu ringkasan statistik (mean, median, frekuensi) dan pilih grafik (histogram, grafik batang, dll.) yang memungkinkan kita meringkas data dengan mudah.

**Syntax:**  
**Codebook**

-----  
id Subject ID Number  
-----

```

type: numeric (int)
range: [1,638]
unique values: 636 units: 1
missing : 0/636
mean: 319.761
std. dev: 184.488
percentiles: 10% 25% 50% 75% 90%
              64 160.5 319.5 479.5 575
    
```

-----  
Fev1 Forced Expiratory Volume  
-----

```

type: numeric (float)
range: [.64,2.69]
unique values: 144 units: .01
missing : 0/636
mean: 1.59465
std. dev: .304451
percentiles: 10% 25% 50% 75% 90%
              1.24 1.395 1.58 1.79 1.98
    
```

-----  
age Subject's age  
-----

```

type: numeric (float)
range: [7.116,10.44]
unique values: 443 units: .001
missing : 4/636
mean: 8.98147
std. dev: .719286
percentiles: 10% 25% 50% 75% 90%
              8.128 8.493 8.909 9.627 9.938
    
```

-----  
height Subject's height  
-----

```

type: numeric (float)
range: [105.6,149]
unique values: 215 units: .1
missing : 0/636
mean: 124.053
std. dev: 6.23791
percentiles: 10% 25% 50% 75% 90%
              116.2 119.9 124 128 131.8
    
```

-----  
sex Subject's gender  
-----

```

type: numeric (byte)
range: [0,1]
unique values: 2 units: 1
missing : 5/636
tabulation: Freq. Value
              333 0
              298 1
              5 .
    
```

-----  
respsymptoms Respiratory symptoms last 12 months  
-----

```

type: numeric (byte)
range: [0,1]
unique values: 2 units: 1
missing : 6/636
tabulation: Freq. Value
              486 0
              144 1
              6 .
    
```

Lalu lakukan operasi 'ta' untuk mengetahui persentase setiap unit pada data kategori.

**Syntax:**

**ta sex**

**ta respsymptoms**

```
. ta sex
```

Subject's gender	Freq.	Percent	Cum.
female	333	52.77	52.77
male	298	47.23	100.00
Total	631	100.00	

```
. ta respsymptoms
```

Respiratory symptoms last 12 months	Freq.	Percent	Cum.
no	486	77.14	77.14
yes	144	22.86	100.00
Total	630	100.00	

● **TABEL 1.3** Hasil Data Deskripsi Penelitian (Data Kategori dan Numerik)

Nama Variabel	Nama Variabel	Ringkasan Statistik	Pilihan Grafik
Id	Angka diskrit	Jumlah observasi: 636 Min: 1 Maks :638	-
FEV <sub>1</sub>	Numerik Berlanjut	Jumlah observasi: 636 Min: 0.64 Maks: 2.69 Mean: 1.59465 Median: 1.58 SD: 0.304451 Persentil 25%: 1.395 Persentil 75%: 1.79	Histogram Plot distribusi kumulatif <i>Dot/bubble plot</i> <i>Box plot</i>
Age/umur	Numerik berlanjut	Jumlah obsevasi: 632 Min: 7.116 Maks: 10.44 Mean: 8.98147 Median: 8.909 SD: 0.719286 Persentil 25%: 8.493 Persentil 75%: 9.627	Histogram Plot ditribusi kumulatif <i>Dot/bubble plot</i> <i>Box plot</i>
Height/Tinggi Badan	Numerik berlanjut	Jumlah observasi: 636 Min: 105.6 Maks: 149 Mean: 124.053 Median: 124 SD: 6.23791 Persentil 25%: 119.9 Persentil 75%: 128	Histogram Plot ditribusi kumulatif <i>Dot/bubble plot</i> <i>Box plot</i>
Sex/Jenis Kelamin	Nominal kategorik	Jumlah observasi: 631 Frek. wanita: 333 (52.77%) Frek. pria: 298 (47.23%)	Bagan batang dan lingkaran



● **TABEL 1.3** Hasil Data Deskripsi Penelitian (Data Kategori dan Numerik)

Nama Variabel	Nama Variabel	Ringkasan Statistik	Pilihan Grafik
Respsymptoms/ gejala penyakit pernapasan	Nominal kategorik	Jumlah observasi: 630 Frekuensi tidak: (77.14%) Frekuensi ya: 144 (22.86%)	Bagan batang dan lingkaran

**Pertanyaan No. 6:**

Menentukan berapa banyak nilai hilang pada setiap variabel.

**Syntax:**  
**Codebook**

```

-----
id                                     Subject ID Number
-----
      type: numeric (int)
      range: [1,638]
unique values: 636                      units: 1
                                         missing.: 0/636
      mean: 319.761
      std. dev: 184.488
percentiles:      10%      25%      50%      75%      90%
                  64      160.5    319.5    479.5    575
-----
fev1                                     Forced Expiratory Volume
-----
      type: numeric (float)
      range: [.64,2.69]
unique values: 144                      units: .01
                                         missing.: 0/636
      mean: 1.59465
      std. dev: .304451
percentiles:      10%      25%      50%      75%      90%
                  1.24     1.395    1.58     1.79     1.98
-----
age                                     Subject's age
-----
      type: numeric (float)
      range: [7.116,10.44]
unique values: 443                      units: .001
                                         missing.: 4/636
      mean: 8.98147
      std. dev: .719286
percentiles:      10%      25%      50%      75%      90%
                  8.128     8.493    8.909    9.627    9.938
-----
height                                 Subject's height
-----
      type: numeric (float)
      range: [105.6,149]
unique values: 215                      units: .1
                                         missing.: 0/636
      mean: 124.053
      std. dev: 6.23791
percentiles:      10%      25%      50%      75%      90%
                  116.2     119.9    124      128      131.8
-----
sex                                     Subject's gender
-----
      type: numeric (byte)
      range: [0,1]
unique values: 2                        units: 1
                                         missing.: 5/636
      tabulation: Freq. value
                  333 0
                  298 1
                   5  .
-----
respsymptoms                           Respiratory symptoms last 12 months
-----
      type: numeric (byte)
      range: [0,1]
unique values: 2                        units: 1
                                         missing.: 6/636
      tabulation: Freq. value
                  486 0
                  144 1
                   6  .
    
```

Variabel	Nilai yang hilang
Umur	4
Sex	5
Respsymptoms	6

**Pertanyaan No. 7:**

Membuat daftar nomor ID responden yang memiliki nilai FEV<sub>1</sub> lebih besar dari 2,4 liter.

**Syntax: list id if fev1 > 2.4**

```
. list id if fev1 > 2.4
```

	id
32.	32
290.	291
424.	426

The ID numbers of subjects: 32, 291, 426

**Pertanyaan No. 8:**

a. Menghitung nilai mean variabel FEV<sub>1</sub> berdasarkan variabel jenis kelamin.

**Syntax: bysort sex:su fev1**

```
. bysort sex:su fev1
```

-> sex = female					
variable	Obs	Mean	Std. Dev.	Min	Max
fev1	333	1.537868	.2906432	.68	2.43

-> sex = male					
variable	Obs	Mean	Std. Dev.	Min	Max
fev1	298	1.661409	.3050209	.64	2.69

-> sex = .					
variable	Obs	Mean	Std. Dev.	Min	Max
fev1	5	1.398	.382714	.93	1.91

Males = 1.661409 litres/sec

Females = 1.537868 litres/sec

\*\*\* Volume udara maksimal yang bisa dihembuskan wanita dalam 1 detik setelah menghirup udara adalah kurang lebih 1,537868 liter/detik. Namun volume udara yang dapat dihembuskan laki-laki dalam 1 detik setelah menghirup udara rata-rata 1,661409 liter/detik. Rata-rata udara ekspirasi maksimum dari 333 wanita lebih rendah daripada 298 laki-laki.

- b. Apakah perkiraan ini dapat digunakan sebagai kesimpulan rata-rata statistik dari variabel FEV<sub>1</sub>? Untuk menjawab pertanyaan tersebut, deskripsikan rata-rata dan ringkasan statistik yang kita pilih untuk menghitung variabel ini, dan juga gunakan grafik. Termasuk deskripsi tentang distribusi data.

**Syntax: su fev1**

```
. su fev1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
fev1	636	1.594654	.304451	.64	2.69

**Masukkan *syntax*: codebook fev1**

```
codebook fev1
```

```
-----+-----
fev1                                         Forced Expiratory Volume
-----+-----
      type: numeric (float)
      range: [.64,2.69]
unique values: 144                          units: .01
                                         missing.: 0/636
      mean: 1.59465
      std. dev: .304451]
percentiles:      10%      25%      50%      75%      90%
                  1.24    1.395    1.58    1.79    1.98
```

● **TABEL 1.4** Ringkasan Variabel FEV<sub>1</sub>

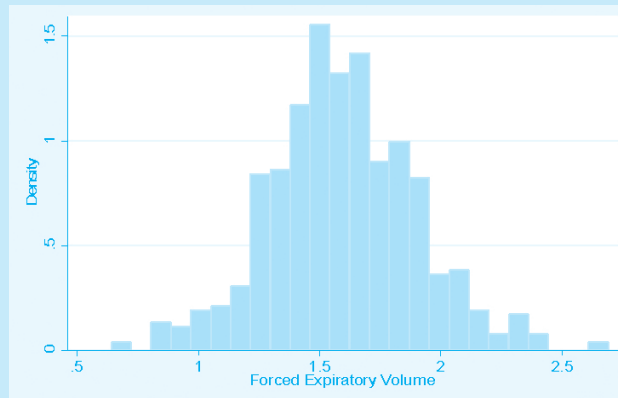
Variabel	Obs	Mean	Std. Dev.	Min	Max	Median
Laki-laki						
Perempuan						
Total	636	1.594654	0.304451	0.64	2.69	1.58

Ya, kurang lebih rata-rata tersebut dapat digunakan sebagai ringkasan statistik dari variabel FEV<sub>1</sub>. Distribusi data simetris (lihat grafik pada Gambar 1.8), sebagai hasil, kita bisa gunakan nilai rata-rata sampel variabel dan ringkasan statistik. Rata-rata mendekati median dengan 1,59465 dan 1,58 secara *respective*. Pada ringkasan data menambah informasi tentang jangkauan data (range), median, dan standar deviasi untuk mengetahui sebaran data. Sebaran data variabel FEV<sub>1</sub> antara 1,29 dan 1,89 (rata-rata ± SD adalah ± 1,59465 ± 0,304451) dengan jangkauan data (range) dari 0,64 hingga 2,69.

● **TABEL 1.5** Ringkasan Variabel FEV<sub>1</sub>

Variabel	Obs	Mean	Std. Dev.	Min	Max	Median
fev1	636	1.594654	0.304451	0.64	2.69	1.58

■ **GAMBAR 1.8** Grafik Distribusi Volume Udara Ekspirasi dalam 1 Detik



**Pertanyaan No. 9:**

- Grafik hubungan antara fungsi paru-paru dan umur responden.
- Deskripsikan hubungan antara dua variabel ini. Bagaimana fungsi paru-paru berdasarkan umur?

● **GAMBAR 1.9** Hubungan antara Fungsi Paru-Paru dan Umur Responden



Grafik pada Gambar 1.9 memperlihatkan hubungan yang linear dengan fungsi paru-paru dan umur reponden. Umur responden yang terdapat pada penelitian ini memiliki fungsi paru-paru yang lebih baik. Grafik tersebut menunjukkan kelompok umur di bawah 8 tahun memiliki nilai  $FEV_1$  lebih rendah dibanding kelompok lebih dari 8 tahun.

- c. Pada grafik tersebut, variabel mana yang bisa Anda deskripsikan sebagai variabel yang menggambarkan *outcome*/hasil dan mana yang menjelaskan paparan? Berikan alasan!

*Outcome*/hasil: fungsi paru-paru

Paparan: umur

Fungsi paru-paru sebagai variabel dependen dan umur sebagai variabel independen yang bertujuan untuk melihat dan mempelajari efek umur terhadap fungsi paru.

**Pertanyaan No.10:**

- a. Prevalensi gejala penyakit pernapasan pada sampel penelitian ini.

**Syntax: ta respsymptoms**

```
ta respsymptoms
*****
Respiratory symptoms last 12 months
```

	Freq.	Percent	Cum.
no	486	77.14	77.14
yes	144	22.86	100.00
Total	630	100.00	

Prevalensi gejala penyakit pernapasan dalam sampel penelitian ini adalah 22,86%.

- b. Menghitung prevalensi berdasarkan jenis kelamin responden. Siapkan tabel 2 × 2 untuk jenis kelamin dan gejala penyakit pernapasan (ingatlah jika variabel *outcome* (dependen) tempatkan di kolom dan paparan [independen] tempatkan di baris).

**Syntax: ta sex respsymptoms**

```
. ta sex respsymptoms
```

Subject's gender	Respiratory symptoms last 12 months		Total
	no	yes	
female	247	81	328
male	234	63	297
Total	481	144	625

● **TABEL 1.6** Prevalensi Gejala Penyakit Pernapasan

Jenis Kelamin Responden	Gejala Penyakit Pernapasan dalam 12 bulan Terakhir		
	Ya	Tidak	Total
Perempuan	81	247	328
Laki-laki	63	234	297
<b>Total</b>	144	481	625

- c. Menentukan jenis kelamin mana yang memiliki prevalensi penyakit gejala pernapasan yang lebih tinggi.

**Syntax:** `ta sex if respsymptoms==1`

```
ta sex if respsymptoms==1
```

Subject's gender	Freq.	Percent	Cum.
female	81	56.25	56.25
male	63	43.75	100.00
Total	144	100.00	

Perempuan memiliki prevalensi yang lebih tinggi 56,25 % dibanding laki-laki

### Pertanyaan No.11:

Siapkan sebuah tabel karakteristik demografi, tinggi badan dan FEV<sub>1</sub> untuk total responden, dan juga berdasarkan ada tidaknya gejala penyakit pernapasan (asumsi data berdistribusi normal).

**Syntax:**

```
su fev1 age height
```

```
ta sex
```

```
. su fev1 age height
```

Variable	Obs	Mean	Std. Dev.	Min	Max
fev1	636	1.594654	.304451	.64	2.69
age	632	8.98147	.7192857	7.116	10.44
height	636	124.053	6.237911	105.6	149

```
. ta sex
```

Subject's gender	Freq.	Percent	Cum.
female	333	52.77	52.77
male	298	47.23	100.00
Total	631	100.00	

- Kemudian lanjutkan *syntax*:

**bysort respsymptoms: su age height fev1**

```
. bysort respsymptoms : su age height fev1
```

```
-> respsymptoms = no
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	482	9.04779	.7336205	7.116	10.44
height	486	124.3934	6.157959	106	149
fev1	486	1.628971	.2884144	.81	2.69

```
-> respsymptoms = yes
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	144	8.755972	.626782	7.278	10.33
height	144	122.9264	6.408964	105.6	147.3
fev1	144	1.478819	.3336061	.64	2.39

```
-> respsymptoms = .
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	6	9.065667	.6204363	8.408	9.864
height	6	123.5167	6.370059	119.1	135.7
fev1	6	1.595	.1380942	1.49	1.87

■ Lanjutkan *syntax*:

```
bysort respsymptoms: ta sex
respsymptoms = no
respsymptoms = yes
respsymptoms = .
```

```
. bysort respsymptoms : ta sex
```

```
-> respsymptoms = no
```

Subject's gender	Freq.	Percent	Cum.
female	247	51.35	51.35
male	234	48.65	100.00
Total	481	100.00	

```
-> respsymptoms = yes
```

Subject's gender	Freq.	Percent	Cum.
female	81	56.25	56.25
male	63	43.75	100.00
Total	144	100.00	

```
-> respsymptoms = .
```

Subject's gender	Freq.	Percent	Cum.
female	5	83.33	83.33
male	1	16.67	100.00
Total	6	100.00	

● **TABEL 1.7** Karakteristik Demografi, Tinggi Badan, FEV<sub>1</sub> untuk Sampel Total dan Sampel Bertingkat Berdasarkan Ada-Tidaknya Gejala Penyakit Pernapasan

**Variabel Numerik**

Variabel	Sampel Total		Gejala Penyakit Pernapasan			
			Ya		Tidak	
	Mean	SD	Mean	SD	Mean	SD
FEV <sub>1</sub>	1.59	0.30	1.48	0.33	1.63	0.29
Umur	8.98	0.72	8.75	0.63	9.05	0.73
Tinggi badan	124.05	6.24	122.93	6.41	124.39	6.16

**Variabel Kategorik**

Jenis kelamin responden	Gejala Penyakit Pernapasan				Total	
	Ya		Tidak			
	n	%	N	%	n	%
Perempuan	81	56.25	247	51.35	333	52.77
Laki-laki	63	43.75	234	48.65	298	47.23







## BAB 2

# ANALISIS DESKRIPTIF PADA STATA

### **Kompetensi Dasar**

- Mampu menjelaskan univariat dengan aplikasi Stata.

### **Indikator Keberhasilan**

- Mampu melakukan proses Stata untuk menghasilkan tabel univariat pada laporan penelitian dan menginterpretasikan hasilnya.

### **Materi Pembelajaran**

1. Review Bab 1 Pengantar Stata.
2. Aplikasi syntax pada analisis univariat dan interpretasi data.

# APLIKASI SYNTAX PADA ANALISIS UNIVARIAT

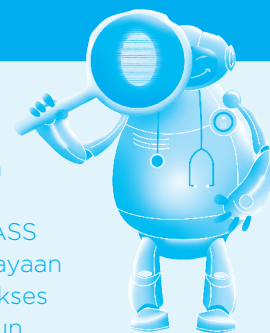
## REVIEW MENGENAI STATA

Berikut ini adalah *syntax* yang akan kita gunakan kembali pada Bab 1 Pengantar Stata. Bisakah Anda jelaskan dan berikan contoh untuk command *syntax* di bawah ini!

Command	Arti	Contoh yang biasa digunakan
By (sort)		
Codebook		
Generate		
Recode		
Tabulate		

### STUDI KASUS 2.1

Penelitian ini menggunakan studi desain kasus-kontrol (*case-control*). Responden diambil dengan menggunakan teknik bola salju (*snowball*) tahun 2010. Regresi logistik sederhana dan ganda dengan menggunakan program statistik (Stata versi 10) untuk mengidentifikasi perbedaan karakteristik Penasun yang Akses LJASS dan Penasun yang Tidak Akses LJASS. Odds ratio, derajat kepercayaan 95% dan nilai signifikansi dilaporkan. Terdapat 46 penasun yang akses LJASS dan 75 penasun yang tidak akses LJASS yang direkrut tahun 2010.



 Buka data: Latihan Stata 1\_Penasun.dta

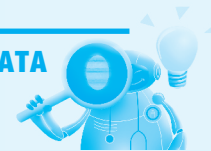
Sumber Data: Najmah. 2013. Characteristics Among Injecting Drug Users Accessing and Not Accessing Needle and Syringe Program in Palembang, South Sumatera, Jurnal Ilmu Kesehatan Masyarakat. Vol 4, 2013: FKM Unsi.

### PERTANYAAN DETEKTIF STATA



Aplikasikan Stata dalam analisis univariat Karakteristik Penasun yang Akses LJASS dan Tidak Akses LJASS (jenis kelamin, pendidikan, pendapatan, status pernikahan, perilaku menyuntik, pengetahuan HIV, dan umur, serta lama menggunakan narkoba suntik).

### JAWABAN DETEKTIF STATA



**Langkah I:** Analisis Data Deskriptif (Analisis Univariat)

SYNTAX : codebook  
SYNTAX :

<b>Data Kategorik</b>	sort (spasi) outcome By (outcome): ta(exposure)
<b>Data Numerik</b>	sort (spasi) outcome By (outcome): sum(exposure) Atau sort (spasi) outcome By (outcome): codebook (exposure)
<b>Codebook</b>	Codebook[variabel]

Aplikasi biostatistika Kasus 2.1:

Kode variabel dan keterangan variabel pada Kasus 2.1 tercantum pada Tabel 2.1.

● **TABEL 2.1** Kode Variabel dan Keterangan Variabel

Kode Variabel di Data	Keterangan Variabel
No10_1	Pendapatan (< Rp 1 juta, kategori referensi)
no2	Umur (tahun)
no37	Lama waktu penggunaan obat suntik (tahun)
no36a	Pernah berbagi jarum dan tabung suntik (0=pernah, 1=tidak pernah)
no45	Penyuntikkan obat satu minggu yang lalu (1=ya, 2=tidak)
tahukel1	Pengetahuan pencegahan HIV dan pengurangan dampak bahaya (0=kurang, 1=tinggi)
LJASS	Perilaku akses LJASS (Layanan Jarum dan Alat Suntik Steril (0=tidak akses, 1=akses)
SIKAP	Sikap terhadap pengurangan dampak buruk (1=negatif, 2=positif)

Aplikasi *syntax* ‘codebook’ untuk mengetahui pembagian kategori pada data kategori dan statistik dasar untuk data numerik, berikut beberapa contoh aplikasi *codebook* pada data di Kasus 2.1.

### 1. Aplikasi Deskripsi Variabel ‘Codebook’

→ Variabel LJASS (Perilaku Akses Layanan Jarum Suntik Steril)

```
. codebook LJASS
```

LJASS	Kelompok LJASS
type: numeric (double)	
label: LJASS	
range: [0,1]	units: 1
unique values: 2	missing .: 0/121
tabulation: Freq.	Numeric Label
75	0 Tidak Akses LJASS
46	1 Akses LJASS

→ Variabel Jenis Kelamin (no. 5)

```
. codebook no5
```

no5	Jenis Kelamin
type: numeric (double)	
label: no5	
range: [1,2]	units: 1
unique values: 2	missing .: 0/121
tabulation: Freq.	Numeric Label
120	1 Laki-laki
1	2 Perempuan

→ Variabel umur saat ini (no. 2)

```
. codebook no2
```

no2	Umur saat ini
type: numeric (double)	
range: [18,41]	units: 1
unique values: 21	missing .: 0/121
mean: 30.3802	
std. dev: 4.2803	
percentiles:	10% 25% 50% 75% 90%
	25 28 30 32 37

→ Variabel lama penggunaan narkoba suntik (no. 37)

Silakan eksplorasi variabel lainnya!

**2. Aplikasi ‘Tab Variabel Dependen Variabel Independen’ atau ‘By Sort’**

→ Untuk menampilkan persentase akses LJASS menurut jenis kelamin (no. 5)

Syntax:

Tab LJASS no 5

→ Untuk menampilkan persentase akses LJASS menurut tingkat pendidikan (no. 6)

Syntax:

Tab LJASS no 6

Tambahkan perintah ROW atau COL untuk menampilkan persen BARIS atau KOLOM

Tab LJASS no 5, COL

Tab LJASS no 5, ROW

Tab LJASS no 6, COL

Tab LJASS no 6, ROW

Coba perhatikan hasil output STATA di komputer Anda dan lakukan pelaporan seperti contoh di bawah ini, persentasi bisa Anda lakukan per kolom atau baris.

**Laporan Hasil**

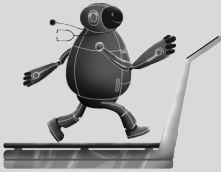
- Persentase akses LJASS menurut jenis kelamin (no. 5) lihat pada Tabel 2.2
- Persentase akses LJASS menurut tingkat pendidikan (no. 6)

● **TABEL 2.2** Karakteristik Responden Berdasarkan Jenis Kelamin

Karakteristik	Syntax	LJASS (n,%) n= 46	Non LJASS (n,%) n= 75
<b>Jenis kelamin</b>	Tab LJASS no 5		
<b>Laki-laki (n,%)</b>	atau  sort LJASS by LJASS :ta no5	45 (97,8%)	75 (100%)

● **TABEL 2.3** Karakteristik Responden Berdasarkan Tingkat Pendidikan

Karakteristik	Syntax	LJASS (n,%) n= 46	Non LJASS (n,%) n= 75
<b>Pendidikan</b>	Tab LJASS no6	1 (2,2%)	-
<b>Tidak pernah (n,%)</b>	atau	2 (4,3%)	5 (6,7%)
<b>SMP (n,%)</b>	sort LJASS	29 (63%)	49 (65,3%)
<b>SMA (n,%)</b>	by LJASS :ta	8 (17,4%)	10 (13,3%)
<b>Akademi (n,%)</b>	no6	6 (13%)	10 (13,3%)
<b>Universitas (n,%)</b>		-	1 (1,3%)
<b>Lain-lain (n,%)</b>			

**LATIHAN:**

Silakan lakukan proses Stata untuk karakteristik variabel kategorik lainnya, tuliskan command *syntax* dan hasilnya pada Tabel 2.4 di bawah ini!

■ **TABEL 2.4** Command Syntax

Variabel (kode variabel)	Syntax	LJASS	Non LJASS
Pendapatan/bulan (no10) Rp500.000-1 juta (n,%) Rp1 juta-1,5 juta (n,%) Rp1,5 juta-2 juta (n,%) Rp2 juta-2,5 juta (n,%) Rp2,5 juta-3 juta (n,%) Lebih dari Rp3 juta (n,%)			-
Status pernikahan (no12) Lajang (n,%) Menikah (n,%) Cerai Hidup (n,%) Cerai Mati (n,%)			
Perilaku Suntik			
Penyuntikkan 1 minggu yang lalu (no36a)			
Berbagi jarum dan tabung suntik (no44)			
Pengetahuan tentang HIV/AIDS dan pengurangan dampak bahaya (tahukel2)			

### 3. Aplikasi 'recode'

Aplikasi "recode" digunakan untuk membuat pengelompokan variabel, baik dari variabel numerik seperti umur maupun variabel kategorik yang memiliki banyak kategori (dibuat lebih sedikit kategorinya). Untuk beberapa variabel yang memiliki kategori banyak, kita kelompokkan kembali jumlah kategorinya menjadi lebih kecil. Misalnya: PENDIDIKAN (variabel no. 6), dikelompokkan menjadi 3 yaitu: SMP, SMA, D-3.

SYNTAX

```
Kategori pendidikan dijadikan 3 kelompok saja
3 groups: (koding 3<=SMP, koding 4 SMA, koding 5>=D3)

generate no6_1=no6

label define didik1 3" <=SMP" 4" SMA" 5" >=D3"

label value no6_1 didik1

codebook no6_1

sort LJASS
by LJASS: ta no6_1
```

**OUTPUT STATA:**

```
. sort LJASS
. by LJASS: ta no6_1
```

-----  
-> LJASS = TidakAkses

no6_1	Freq.	Percent	Cum.
<=SMP	6	8.00	8.00
SMA	49	65.33	73.33
>=D3	20	26.67	100.00
Total	75	100.00	

-----  
-> LJASS = Akses LJAS

no6_1	Freq.	Percent	Cum.
<=SMP	3	6.52	6.52
SMA	29	63.04	69.57
>=D3	14	30.43	100.00
Total	46	100.00	


● **TABEL 2.5** Karakteristik Kelompok LJASS dan Non LJASS

Karakteristik	LJASS (n,%) n= 46	Non LJASS (n,%) n= 75
<b>Jenis kelamin</b>		
Laki-laki (n,%)	45 (97,8%)	75 (100%)
<b>Pendidikan</b>		
Tidak pernah (n,%)	1 (2,2%)	-
SMP (n,%)	2 (4,3%)	5 (6,7%)
SMA (n,%)	29 (63%)	49 (65,3%)
Akademi (n,%)	8 (17,4%)	10 (13,3%)
Universitas (n,%)	6 (13%)	10 (13,3%)
Lain-lain (n,%)	-	1 (1,3%)
<b>Pendapatan/bulan</b>		
Rp500.000-1 juta (n,%)	7 (18,9%)	10 (20%)
Rp1 juta-1,5 juta (n,%)	13 (35,1%)	16 (32%)
Rp1,5 juta-2 juta (n,%)	7 (18,9%)	14 (28%)
Rp2 juta-2,5 juta (n,%)	8 (21,6%)	6 (12%)
Rp2,5 juta-3 juta (n,%)	1 (2,7%)	4 (8%)
Lebih dari Rp3 juta (n,%)	1 (2,7%)	-
<b>Status pernikahan</b>		
Lajang (n,%)	21 (45,7%)	47 (62,7%)
Menikah (n,%)	22 (47,8%)	26 (34,7%)
Cerai Hidup (n,%)	2 (4,3%)	2 (2,7%)
Cerai Mati (n,%)	1 (2,2%)	-
<b>Perilaku Suntik</b>		
Penyuntikkan 1 minggu yang lalu	41 (89,1%)	55 (73,3%)
Berbagi jarum dan tabung suntik	16 (34,8%)	25 (33,3%)
Pengetahuan tentang HIV/AIDS dan pengurangan dampak bahaya (%Baik)	43 (94%)	66 (88%)
	<b>Mean (SD)</b>	
Umur, Tahun	30 (4)	30,5 (4,5)
Lama waktu penggunaan obat suntikan (Tahun)	8,5 (4,6)	6,8 (4,2)

#### 4. Laporan dan Interpretasi

Berdasarkan Tabel 2.5 dapat diinterpretasikan untuk jenis kelamin semuanya laki-laki, mayoritas tidak memiliki akses LJASS dan minoritasnya memiliki akses LJASS. Berdasarkan pendidikan, baik yang memiliki akses dan yang tidak memiliki akses LJASS mayoritasnya berpendidikan SMA. Dari segi pendapatan, baik yang memiliki akses dan yang tidak memiliki akses LJASS mayoritas memiliki pendapatan antara 1 sampai 1,5 juta/bulan dan minoritas berpendapatan 2,5-3 juta/bulan (memiliki akses LJASS dan yang tidak) dan lebih dari 3 juta/bulan (yang memiliki akses LJASS). Dilihat dari status pernikahan, yang memiliki akses LJASS mayoritas berstatus telah menikah dan minoritasnya berstatus cerai mati. Pada kelompok yang tidak memiliki akses LJASS mayoritas berstatus lajang dan minoritasnya berstatus cerai hidup. Proporsi terbesar dilihat dari perilaku penyuntikkan, terdapat pada responden berstatus memiliki akses LJASS.





## BAB 3 APLIKASI STATA PADA PENGHITUNGAN EPIDEMIOLOGI (OR/RR)

### **Kompetensi Dasar**

Mampu menjelaskan aplikasi uji Stata pada studi epidemiologi.

### **Indikator Keberhasilan**

Mampu melakukan proses Stata penghitungan epidemiologi dan menginterpretasikan hasilnya.

### **Materi Pembelajaran**

1. Introduksi Konsep RR, OR, dan PR.
2. Aplikasi Stata pada penghitungan *risk ratio* serta interpretasi.
3. Aplikasi Stata pada penghitungan *odds ratio* serta interpretasi.
4. Aplikasi Stata pada penghitungan Prevalensi Rasio serta interpretasi.

## PENGENALAN

Pada studi epidemiologi, kita juga berhadapan dengan penghitungan data-data kesehatan. Tugas kita adalah bagaimana membuat data-data kesehatan yang ada menjadi lebih bermakna daripada hanya ditampilkan berupa grafik atau data deskripsi saja. Peneliti epidemiologi akan dihadapkan pertanyaan “**Mengapa** suatu masalah kesehatan terjadi?” “Siapa yang lebih berisiko terhadap suatu penyakit?” “Perilaku apa yang seharusnya dilakukan untuk pencegahan penyakit?” dan lain sebagainya. Oleh karena itu, ada penghitungan asosiasi sederhana untuk mengetahui risiko seseorang ataupun sekelompok, yaitu: *risk ratio* untuk studi eksperimental dan kohort, *odds ratio* untuk studi kasus kontrol, dan prevalensi rasio untuk studi potong lintang (*cross sectional*). Ada tiga kisaran nilai rasio sebagai berikut.

- Jika nilai rasio  $< 1$ ; ada asosiasi berupa menurunkan risiko terhadap suatu *outcome*, faktor protektif, faktor pelindung, faktor pencegah.
- Jika nilai rasio  $= 1$ ; tidak ada asosiasi.
- Jika nilai rasio  $> 1$ ; ada asosiasi berupa meningkatkan risiko terhadap suatu *outcome*/penyakit/kondisi kesehatan, faktor risiko.

**Berikut ini adalah beberapa contoh interpretasi nilai rasio** (catatan: pada Stata, *output/exposure* dengan nilai besar [misal 1] dibandingkan dengan *output/exposure* dengan nilai paling kecil [misal 0], pada SPSS sebaliknya):

### STUDI DESAIN KOHORT

- Risk Ratio (RR) = 3,4 (*exposure*: merokok [1=setiap hari, 0=kadang-kadang], *outcome*: kanker paru [0= Tidak, 1=Ya])  
Interpretasi: responden yang merokok setiap hari memiliki risiko menderita kanker paru 3,4 lebih tinggi dibandingkan responden yang tidak merokok setiap hari.

### STUDI DESAIN KASUS KONTROL

- Odds Ratio (OR) = 0,7 (*exposure*: konsumsi serat [1=setiap hari, 0=tidak setiap hari, *outcome*: kejadian sembelit]).  
Interpretasi: anak-anak yang mengonsumsi serat setiap hari *mengurangi* risiko sembelit sebesar 0,7 kali lebih rendah (30% lebih rendah) daripada anak-anak yang tidak mengonsumsi serat setiap hari.

### STUDI DESAIN POTONG LINTANG

- **Prevalence Ratio (PR)** = 1,0 (*exposure*: pengetahuan kesehatan reproduksi [1=baik, 0=kurang], sikap terhadap perilaku seks bebas [0: Tidak, 1: Ya]).  
Interpretasi: remaja yang memiliki pengetahuan kesehatan reproduksi baik memiliki sikap yang sama terhadap perilaku seks bebas dibandingkan remaja yang memiliki pengetahuan kesehatan reproduksi kurang baik.

● **TABEL 3.1** Penghitungan Asosiasi Sederhana

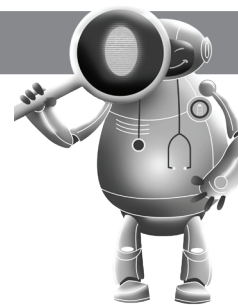
Ukuran	Rumus	Desain Studi
<b>Risk ratio</b>	$\frac{\text{Insidensi kelompok terpapar}}{\text{Insidensi kelompok tidak terpapar}}$	<ul style="list-style-type: none"> <li>■ Experimental</li> <li>■ Kohort</li> </ul>
<b>Odds ratio</b>	$\frac{\text{Odds Kasus}}{\text{Odds Kontrol}}$	<ul style="list-style-type: none"> <li>■ Kasus kontrol</li> <li>■ Potong lintang kejadian jarang</li> </ul>
<b>Prevalence ratio</b>	$\frac{\text{Prevalensi kelompok terpapar}}{\text{Prevalensi kelompok tidak terpapar}}$	<ul style="list-style-type: none"> <li>■ Potong lintang</li> </ul>

Keterangan:

- Insidensi/prevalensi kumulatif kelompok terpapar merupakan proporsi dari kasus baru/baru dan lama pada kelompok yang terpapar
- Insidensi/prevalensi kumulatif kelompok tidak terpapar merupakan proporsi dari kasus baru/baru dan lama pada kelompok yang tidak terpapar

## APLIKASI STATA PADA PENGHITUNGAN RISIKO RASIO (RISK RATIO/RELATIVE RISK)

### STUDI KASUS 3.1



Pada penelitian eksperimental Najmah, dkk., (2015) ingin menguji efektivitas intervensi pemodelan kawasan tanpa rokok pada tingkat rumah tangga di Ogan Ilir yang diadopsi dari *Centers for Disease Control and Prevention* (CDC). Salah satu variabel yang dilihat adalah perilaku merokok di dalam rumah (*outcome*) pada kelompok yang diberikan intervensi terpadu dan tidak (paparan).

Perhatikan data pada Tabel 3.2 berikut.

● **TABEL 3.2** Efektivitas Intervensi Terpadu terhadap Perilaku Merokok di Dalam Rumah

Status Pemberian Intervensi	Merokok dalam Rumah		Total
	Ya	Tidak	
Ya	56 (58%)	39 (42%)	95
Tidak	39 (70%)	30 (30%)	100
Total	126	69	195

Sumber: Najmah, Fenny Etrawati, Yeni, & Feranita Utama, 2015.

<b>Risk Ratio</b>	Insidensi kelompok terpapar	(56/95)
	Insidensi kelompok tidak terpapar	(79/100)

**SYNTAX RISK RATIO**

Jika ada dataset Stata

`cs [outcome] [exposure]`

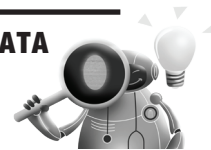
`cs [outcome] [exposure] {if expr}}, [by(varname) level(#)]`

Jika diketahui data 2x2

`csi [#outcome+ _exposure+] [#outcome+ _exposure+] [#outcome -exposure-] [#outcome-exposure-]`

**Keterangan**

Level (#) : bisa menspesifikkan derajat kepercayaan yang diinginkan (90, 95, atau 99%)  
 By [varname] : jika ingin distratifikasi oleh variabel lain  
 or : bisa menghasilkan odds ratio selain *risk ratio*

**JAWABAN DETEKTIF STATA**

Langkah-langkah penghitungan *risk ratio* untuk tabel 2 x 2 yang sudah diketahui adalah sebagai berikut.

1. Ketik *syntax risk ratio* di bawah ini pada kolom *command* di Stata, lalu klik Enter.

`csi 56 70 39 30`

(Ingat urutan *syntax*, `csi [#outcome+_exposure+] [#outcome+_exposure+] [#outcome -exposure-] [#outcome-exposure-]`)

2. Outcome Stata

→ `csi 56 70 39 30`

	Exposed	Unexposed	Total
Cases	56	70	126
Noncases	39	30	69
Total	95	100	195
Risk	.5894737	.7	.6461538
	Point estimate		[95% Conf. Interval]
Risk difference	-.1105263		-.2441394 .0230867
<b>Risk ratio</b>	<b>.8421053</b>		<b>.6817479 1.040181</b>
Prev. frac. ex.	.1578947		-.040181 .3182521
Prev. frac. pop	.0769231		

`chi2(1) = 2.60 Pr>chi2 = 0.1067`

→ `csi 56 70 39 30`, or, (jika ingin output Stata dengan *odds ratio*)

csi 56 70 39 30, or

	Exposed	Unexposed	Total
Cases	56	70	126
Noncases	39	30	69
Total	95	100	195
Risk	.5894737	.7	.6461538
	Point estimate		[95% Conf. Interval]
Risk difference	-.1105263		-.2441394 .0230867
Risk ratio	.8421053		.6817479 1.040181
Prev. frac. ex.	.1578947		-.040181 .3182521
Prev. frac. pop	.0769231		
Odds ratio	.6153846		.341515 1.108981 (Cornfield)

chi2(1) = 2.60 Pr>chi2 = 0.1067

→ csi 56 70 39 30, level(90), jika Confidence Interval kita ditentukan = 90%

**. csi 56 70 39 30, level(90)**

	Exposed	Unexposed	Total
Cases	56	70	126
Noncases	39	30	69
Total	95	100	195
Risk	.5894737	.7	.6461538
	Point estimate		[90% Conf. Interval]
Risk difference	-.1105263		-.2226579 .0016053
Risk ratio	.8421053		.7052995 1.005447
Prev. frac. ex.	.1578947		-.0054469 .2947005
Prev. frac. pop	.0769231		

chi2(1) = 2.60 Pr>chi2 = 0.1067

3. *Interpretasi:*

Risk ratio	.8421053	.6817479	1.040181
------------	----------	----------	----------

Berdasarkan penghitungan di atas, maka dengan RR sebesar 0.84 dapat diinterpretasikan sebagai risiko perilaku merokok di dalam rumah pada kelompok yang diberikan intervensi

terpadu 0.85 kali lebih rendah atau 15% mengurangi perilaku merokok di dalam rumah dibandingkan dengan kelompok non-intervensi. Dengan derajat kepercayaan 95%, intervensi bisa mengurangi perilaku merokok di dalam rumah 0.68 (menurunkan risiko 32%) atau meningkatkan perilaku sebesar 1.04 kali lebih tinggi dibandingkan kelompok non-intervensi. Signifikansi menunjukkan tidak ada hubungan antara intervensi terhadap perilaku merokok (p-value 0.1067).

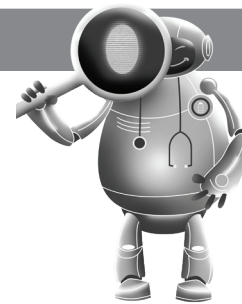
## APLIKASI STATA PADA PENGHITUNGAN ODDS RATIO

Asosiasi antara paparan (*exposure*) dan penyakit (*risk ratio*) dalam studi *kasus kontrol* diukur dengan menghitung *Odds Ratio* (OR), di mana OR merupakan rasio kemungkinan paparan pada kasus dan rasio kemungkinan paparan pada kontrol. Odds kasus artinya perbandingan jumlah kasus terpapar dengan kasus tidak terpapar, sedangkan odds kontrol artinya perbandingan jumlah kontrol terpapar dan kontrol tidak terpapar. Rumus *Odds Ratio* adalah:

$$\text{Odds Ratio (OR)} = \frac{\text{Odds Kasus} \left( \frac{a}{b} \right)}{\text{Odds Kontrol} \left( \frac{c}{d} \right)}$$

### STUDI KASUS 3.2

Berikut ini adalah contoh penghitungan OR. Suatu penelitian kasus kontrol ingin mengetahui hubungan antara kebiasaan minum alkohol dan kejadian patah tulang pinggul pada wanita lansia di Geelong, Australia, pada wanita lansia. Kasus adalah kelompok wanita patah tulang pinggul, dan kontrol adalah wanita lansia yang tidak patah tulang pinggul.



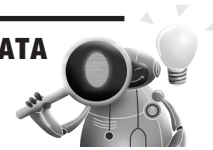
Buka Data: Buka data: kasus\_iii\_hip\_fracture1.dta

Sumber Data: Najmah, et al. Hip Structure Associated With Hip Fracture in Women: Data from the Geelong Osteoporosis Study (Gos) Data Analysis-Geelong, Australia. *International Journal of Public Health Research* 2011, 2011(Special Issue): p. 185-192.

● **TABEL 3.3** Hubungan Antara Status Kebiasaan dan Kejadian Patah Tulang Pinggul pada Wanita Lansia di Geelong, Australia

Kebiasaan Minum Alkohol	Status Patah Tulang Pinggul		TOTAL
	Ya (kasus)	Tidak (kontrol)	
Ya	35	10	45
tidak	9	444	453
Total	44	454	498

## JAWABAN DETEKTIF STATA



### 1. Kode Syntax

	Insidensi kelompok Kelompok kasus	$\frac{(35/4)}{(10/444)}$
<b>Odds ratio</b>	Insidensi kelompok kelompok kontrol	

#### SYNTAX RISK RATIO

Jika ada dataset Stata

cc [outcome] [exposure]

cc [outcome] [exposure] {if expr}, [by(varname) level(#)]

contoh: cc outcome exposure if sex==1

cc outcome exposure if sex==1

Jika diketahui data 2x2

cci [#case exposed] [#case unexposed] [#controlexposed] [#control unexposed]

Untuk tabel lebih dari 2x2

tabodds outcome exposure, base (#)

#-- reference group, missal 1=SD, 2=SMP 3. SMA

taboddskankerdidik, base(1)

### 2. Output Syntax

cci 35 9 10 444

. cci 35 9 10 444

	Exposed	Unexposed	Total	Proportion Exposed
Cases	35	9	44	0.7955
Controls	10	444	454	0.0220
Total	45	453	498	0.0904
	Point estimate		[95% Conf. Interval]	
Odds ratio	172.6667		60.29296	509.7318 (exact)
Attr. frac. ex.	.9942085		.9834143	.9980382 (exact)
Attr. frac. pop	.7908477			

chi2(1) = 291.92 Pr>chi2 = 0.0000

### 3. Interpretasi Hasil

Hasil penghitungan OR yang didapat adalah 172 kali. Artinya, risiko terjadinya patah tulang pinggul pada wanita lansia di Geelong pada wanita peminum alkohol adalah 172 kali lebih tinggi dibandingkan risiko terjadinya patah tulang pinggul pada wanita bukan peminum alkohol.

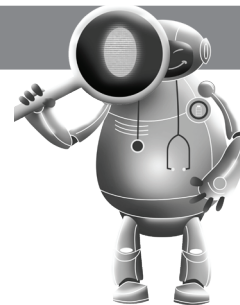
# APLIKASI STATA PADA PENGHITUNGAN RASIO PREVALENSI (PREVALENCE RATIO)

Penelitian *cross sectional* (potong-lintang) merupakan penelitian yang menggunakan data prevalensi. Rasio prevalensi dapat dihitung dengan menggunakan rumus *odds ratio* atau *risk ratio*, akan tetapi data yang digunakan bukan data insidensi kumulatif melainkan data prevalensi penyakit (Bonita R, 2006; Murti, 1997; Webb P. Bain C, 2005).

$$\text{Prevalensi Rasio} = \frac{\text{Angka Prevalensi Kelompok Terpapar} \left( \frac{a}{N1} \right)}{\text{Angka Prevalensi pada Kelompok Tidak Terpapar} \left( \frac{b}{N0} \right)}$$

## STUDI KASUS 3.3

Berikut ini merupakan contoh penelitian dengan judul *Prevalensi dan Faktor-Faktor yang Mempengaruhi Near Miss di Indonesia*, oleh Rini Mutahar, dkk (2014), Laporan Hasil Analisis lanjut Riset Kesehatan Dasar 2010. Peneliti ingin mengetahui hubungan antara paritas (jumlah anak) dan kejadian *near miss* (ibu hampir mati) pada proses persalinan di Indonesia.



Perhatikan data pada Tabel 3.4 berikut.

● **TABEL 3.4** Kejadian Near Miss Berdasarkan Paritas Responden

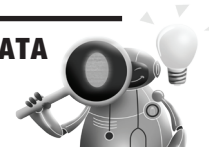
Faktor Paparan/ Risiko	Kejadian Nearmiss		Total	PR	
	Nearmiss	Bukan Nearmiss			
	N	n	n		
Paritas	Multipara	3534	3698	7232	0.89
	Primipara	2503	2070	4573	Ref
	<b>Total</b>	6037	5768	11805	

\*nilai signifikansi < 0.0001

Sumber: Mutahar, R., Najmah, and Yeni, *Prevalensi dan Faktor-Faktor yang Mempengaruhi Near Miss di Indonesia*, L.A.L.R. 2010, Editor. 2014, FKM Unsri-Litbangkes Depkes RI: Indralaya.



**JAWABAN DETEKTIF STATA**



SYNTAX

1. Prevalensi rasio dengan konsep *cross sectional*  
`csi3534 2503 3698 2070`

```
. csi 3534 2503 3698 2070
```

	Exposed	Unexposed	Total	
Cases	3534	2503	6037	
Noncases	3698	2070	5768	
Total	7232	4573	11805	
Risk	.4886615	.5473431	.5113935	
	Point estimate		[95% Conf. Interval]	
Risk difference	-.0586816		-.0771437	-.0402195
Risk ratio	.8927883		.8617684	.9249248
Prev. frac. ex.	.1072117		.0750752	.1382316
Prev. frac. pop	.0656802			

chi2(1) = 38.61 Pr>chi2 = 0.0000

2. Interpretasi: PR 0.89 (95% 0.86–0.92), p. < 0.0001

Berdasarkan hasil penghitungan prevalensi rasio dengan pendekatan *risk ratio* diperoleh, ibu yang memiliki anak 2 sampai 3 anak berisiko 0,89 kali lebih kecil atau 11% mengurangi risiko untuk mengalami kejadian *nearmiss* dibandingkan ibu yang memiliki satu orang anak (primipara). Dengan derajat kepercayaan 95%, di populasi ibu yang memiliki anak 2 hingga 3 mengurangi risiko untuk terjadinya *nearmiss* antara 0,86 hingga 0,92 (mengurangi risiko antara 8–14%) dibandingkan ibu dengan satu anak.





## BAB 4

# REGRESI LOGISTIK SEDERHANA DAN BERGANDA

### **Kompetensi Dasar**

- Mampu menjelaskan aplikasi uji logistik sederhana dan logistik berganda dengan aplikasi Stata

### **Indikator Keberhasilan**

- Mampu melakukan proses Stata untuk uji logistik sederhana dan logistik berganda serta menginterpretasikan hasilnya.

### **Materi Pembelajaran**

1. Fungsi syntax pada uji regresi logistik sederhana dan berganda.
2. Aplikasi syntax pada uji logistik sederhana.
3. Aplikasi syntax pada uji logistik berganda.

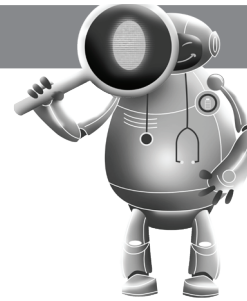
Kita akan menganalisis data Studi Kasus Penasun dengan mengaplikasikan uji regresi logistik sederhana dan berganda. Sebelum itu, mari perhatikan contoh *command* aplikasi Stata pada uji regresi sederhana dan berganda.

● **TABEL 4.1** Contoh Command Aplikasi Stata pada Uji Regresi Sederhana dan Berganda

Command	Arti	Contoh yang Biasa Digunakan
<b>Xi: logistic [var outcome] [var faktor risiko]</b>	Logistik sederhana untuk mengeluarkan nilai OR	<b>Xi:logistic BBLR i.merokok</b> <b>Xi:logistic Ca_paru merokok</b>
<b>Xi: logit [var outcome] [var faktor risiko]</b>	Logistik sederhana untuk mengeluarkan nilai koefisien regresi.	<b>Xi:logit BBLR i.merokok</b> <b>Xi:logit Ca_paru merokok</b>
<b>Xi: logistic [var outcome] [var fx risiko1] [var fx risiko2] [var fx risiko3]</b>	Logistik berganda (lebih dari satu risiko)	<b>Xi:logistic BBLR merokok gizi_ibu ekonomigr</b>
Atau <b>Xi: logistic [var outcome] [var fx risiko1] [var fx perancu1] [var fx perancu2] [var fx perancu3]</b>		<b>Xi:logistic Ca_paru merokok alkohol aktivitas fisik status_ekonomi</b>
<b>xi: logistic [outcome] [eksporur] est store A logistic [outcome] [eksporur] est store B lrtest A B</b>	Untuk mendapatkan nilai p value tunggal dari eksposur yang dengan kategori lebih dari 2 kategori	<b>xi: logistic LJASS no10_1 est store A</b> <b>xi: logistic LJASS i.no10_1 est store B</b> <b>lrtest A B</b>
<b>xi: logistic [outcome] [eksporur1] [eksporur2/perancu] est store A logistic [outcome] [eksporur2/perancu] est store B lrtest A B</b>	<b>Likelihood test untuk menguji hubungan antara outcome dan faktor risiko dikontrol oleh faktor perancu atau variabel interaksi</b>	

## APLIKASI STATA PADA UJI REGRESI LOGISTIK SEDERHANA

### STUDI KASUS 4.1



Lakukan uji statistik untuk mengetahui hubungan antara karakteristik responden dengan akses LJASS. Sajikan hasilnya dalam bentuk tabel yang penting sebagai panduan agar mudah membaca nilai OR (kolom berisi variabel akses LJASS, baris berisi variabel karakteristik)

Hitung nilai OR dan P-value (sajikan pada tabel yang sama) \*Tabel ini penting sebagai panduan agar mudah membaca nilai OR



Buka data: Latihan Stata 1\_Penasun.dta

Sumber Data: Najmah. 2013. Characteristics Among Injecting Drug Users Accessing and Not Accessing Needle and Syringe Program In Palembang, South Sumatera. Jurnal Ilmu Kesehatan Masyarakat. Vol 4. 2013: FKM Unsi

#### PERTANYAAN DETEKTIF STATA

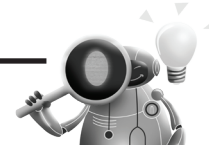
Apakah ada hubungan antara karakteristik penasun dan upaya akses LJASS?



#### JAWABAN DETEKTIF STATA

Langkah II: Analisis Data Regresi Logistik Sederhana

1. Aplikasi *syntax* 'xi: logistic outcome eksposur'



● **TABEL 4.2** Hubungan antara Karakteristik Penasun dan Upaya Akses LJASS

Variabel (Hasil, perilaku akses LJASS, 1=Ya, 0=Tidak)	Syntax	Odds Ratio/OR(95%CI)	P value
Pendapatan (< Rp 1 juta, kategori referensi)	xi: logistic LJASS no10_1 est store A xi: logistic LJASS i.no10_1 est store B lrtest A B	1.05(0.69-1.57)	0.82
Umur (tahun)	xi:logistic LJASS no2	0.98(0.90-1.07)	0.71
Lama penggunaan napza suntik (Tahun)	xi:logistic LJASS no37	1.09(1.002-1.19)	0.04
Pernah berbagi jarum dan tabung suntik (0=pernah, 1=tidak pernah)	xi:logistic LJASS no44	0.94(0.43-2.03)	0.87
Penyuntikkan satu minggu yang lalu (1=Ya, 2=tidak)	xi:logistic LJASS no36a	0.33(0.11-0.96)	0.043
Pengetahuan pencegahan HIV dan pengurangan dampak bahaya (0=kurang, 1=tinggi)	xi:logistic LJASS tahukel2	1.95(0.50-7.63)	0.33
Sikap terhadap pengurangan dampak bahaya (1=negatif, 2=positif)	xi:logistic LJASS SIKAP	2.76(1.30-5.89)	0.008

## 2. Output Stata

```
. *PENDAPATAN(p value gabungan, 0.45)
```

```
. xi: logistic LJASS no10_1
Logistic regression      Number of obs =    88
                        LR chi2(1)    =    0.05
Prob> chi2    =    0.8234
Log likelihood = -59.853656   Pseudo R2    =    0.0004

-----+-----
LJASS | Odds Ratio  Std. Err.   z    P>|z|   [95% Conf. Interval]
-----+-----
no10_1 |  1.047242   .2166035   0.22  0.823   .6982184   1.570736
-----+-----

.
. est store A
.
. xi: logistic LJASS i.no10_1
.
i.no10_1   _Ino10_1_3-6   (naturally coded; _Ino10_1_3 omitted)
Logistic regression      Number of obs =    88
                        LR chi2(3)    =    1.58
Prob> chi2    =    0.6632
Log likelihood = -59.086976   Pseudo R2    =    0.0132

-----+-----
LJASS | Odds Ratio  Std. Err.   z    P>|z|   [95% Conf. Interval]
-----+-----
_Ino10_1_4 |  1.160714   .7176559   0.24  0.810   .3454916   3.899538
_Ino10_1_5 |  .6666667   .4483951  -0.60  0.547   .1784008   2.491269
_Ino10_1_6 |  1.428571   .9506788   0.54  0.592   .3876578   5.264479
-----+-----

.
. est store B
.
. lrtest A B

Likelihood-ratio test      LR chi2(2) = 1.53
(Assumption: A nested in B) Prob> chi2 = 0.4646
.
```

. \*UMUR

. logistic LJASS no2

Logistic regression                      Number of obs = 121  
 LR chi2(1)                                = 0.14  
 Prob> chi2                                = 0.7089  
 Log likelihood = -80.291835            Pseudo R2                    = 0.0009

LJASS	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
no2	.9836824	.0434271	-0.37	0.709	.9021453 1.072589

. \*Lama penggunaan napza suntik (Tahun)

. logistic LJASS no37

Logistic regression                      Number of obs = 121  
 LR chi2(1)                                = 4.19  
 Prob> chi2 = 0.0405  
 Log likelihood = -78.264057            Pseudo R2                    = 0.0261

LJASS	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
no37	1.092465	.0479537	2.01	0.044	1.002407 1.190614

. \*Pernah berbagi jarum dan tabung suntik

. logistic LJASS no44

Logistic regression                      Number of obs = 121  
 LR chi2(1)                                = 0.03  
 Prob> chi2                                = 0.8702  
 Log likelihood = -80.348199            Pseudo R2                    = 0.0002

LJASS	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
no44	.9375	.370085	-0.16	0.870	.4324673 2.032307





### 3. Laporan Hasil

● **TABEL 4.3** Perbandingan Karakteristik Penasun Berstatus Memiliki Akses LJASS dan Tidak Memiliki Akses LJASS

Variabel (Hasil, perilaku akses LJASS, 1=Ya, 0=Tidak)	Odds ratio/OR (95%CI)	P value
Pendapatan (< Rp 1 juta, kategori referensi)	1.05(0.69-1.57)	0.82
Umur (tahun)	0.98(0.90-1.07)	0.71
Lama penggunaan napza suntik (tahun)	1.09(1.002-1.19)	0.04
Pernah berbagi jarum dan tabung suntik (0=pernah, 1=tidak pernah)	0.94(0.43-2.03)	0.87
Penyuntikkan satu minggu yang lalu (1=Ya, 2=tidak)	0.33(0.11-0.96)	0.043
Pengetahuan pencegahan HIV dan pengurangan dampak bahaya (0=kurang, 1=tinggi)	1.95(0.50-7.63)	0.33
Sikap terhadap pengurangan dampak bahaya (1=negatif, 2=positif)	2.76(1.30-5.89)	0.008

### 4. Interpretasi

Ada hubungan yang signifikan antara lama penggunaan napza suntik (OR 1.09, p. 0.71) aktivitas menyuntik napza satu minggu yang lalu (OR 0.33, p. 0.043) dan sikap terhadap pengurangan dampak buruk terhadap napza suntik (OR 2.7, p.0.008) terhadap perilaku akses LJASS. Penasun yang akses LJASS memiliki sikap positif terhadap pengurangan dampak buruk sebesar 2,76 kali lebih tinggi dibandingkan penasun yang tidak akses LJASS (p.0.008). Dengan derajat kepercayaan 95%, penasun yang akses LJASS di populasi memiliki sikap positif terhadap pengurangan dampak buruk napza suntik antara 1.30 hingga 5.89 lebih tinggi dibandingkan penasun yang tidak akses LJASS. Sementara itu, umur, tingkat pendapatan, perilaku berbagi jarum suntik tidak steril dan pengetahuan tentang HIV tidak berhubungan dengan perilaku akses LJASS (nilai P-value > 0,05 lihat Tabel 4.3)

## APLIKASI STATA PADA UJI REGRESI LOGISTIK BERGANDA

Langkah-langkah dalam Uji Regresi Logistik Berganda pada Studi Kasus 4.2 adalah sebagai berikut.

1. Aplikasi *syntax* 'xi: logistik outcome eksposur eksposur eksposur eksposur dst' (Asumsi semua variabel saling memengaruhi satu sama lain )

SYNTAX

xi:logistic LJASS tahukel2 no10\_1 no2 no37 no44 no36a SIKAP

## 2. Output Stata

```
. xi:logistic LJASS tahukel2 i.no10_1 no2 no37 no44 no36a SIKAP
i.no10_1 _Ino10_1_3-6 (naturally coded; _Ino10_1_3 omitted)
```

```
Logistic regression                Number of obs = 88
                                   LR chi2(9)       = 23.65

Prob> chi2          = 0.0049
Log likelihood      = -48.05381      Pseudo R2       = 0.1975
```

LJASS	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
tahukel2	1.495405	1.406861	0.43	0.669	.2365688	9.452794
_Ino10_1_4	1.036809	.7659802	0.05	0.961	.2436926	4.41118
_Ino10_1_5	.3739788	.3118172	-1.18	0.238	.0729675	1.916745
_Ino10_1_6	.8428233	.7041773	-0.20	0.838	.1638926	4.334249
no2	.9529861	.0807296	-0.57	0.570	.8071967	1.125107
no37	1.178731	.0844969	2.29	0.022	1.024228	1.356541
no44	.6667425	.4164947	-0.65	0.516	.1959915	2.268188
no36a	.0985447	.0728921	-3.13	0.002	.0231213	.4200053
SIKAP	3.82676	2.042489	2.51	0.012	1.34435	10.89307

## 3. Buatlah tabel laporan dan interpretasikan!

- **TABEL 4.4** Hasil Multivariat Karakteristik Pengguna Napza Suntik dalam Mengakses Layanan Jarum dan Alat Suntik Steril (LJASS, 1=Ya, 0=Tidak)

Variabel Independen	Adjusted OR (95%CI)	P value
Pengetahuan pencegahan HIV dan pengurangan dampak bahaya (0=kurang, 1=tinggi)	1.49 (0.23-9.45)	0.669
Pendapatan (< 1 juta sebagai kategori)		
Level 2	1.04 (0.24-4.41)	0.961
Level 3	0.37 (0.07-1.91)	0.238
Level 4	0.84 (0.16-4.33)	0.838
Umur (tahun)	0.95 (0.81-1.12)	0.57
Lama penggunaan Napza Suntik (tahun)	1.18 (1.02-1.35)	0.02
Pernah berbagi jarum dan tabung suntik (0=pernah, 1=tidak pernah)	0.67 (0.19-2.27)	0.19

Variabel Independen	Adjusted OR (95%CI)	P value
Penyuntikkan satu minggu yang lalu (1=Ya, 2=tidak)	0.09 (0.02-0.42)	0.002
Sikap terhadap pengurangan dampak bahaya (1=negatif, 2=positif)	3.82 (1.34-10.89)	0.012

\*Adjusted by all variables

#### CONTOH INTERPRETASI SINGKAT:

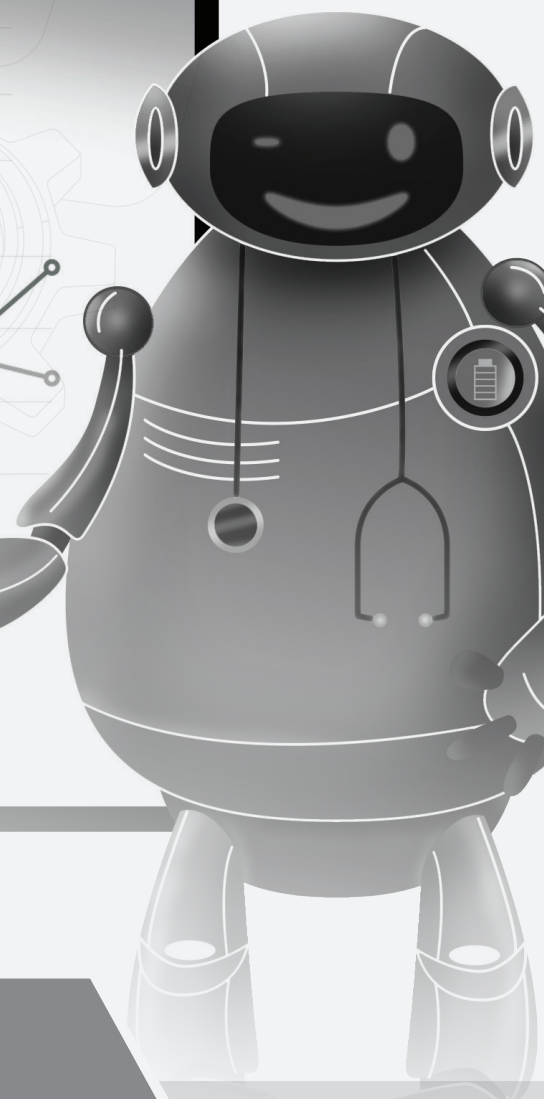
Berdasarkan analisis multivariat penasun yang memiliki sikap positif terhadap konsep *harm reduction* memiliki kecenderungan (*likelihood*) 3,8 kali lebih besar mengakses LJASS dari penasun yang memiliki sikap negatif (OR 3,8 [95% CI 1,4-10,9], p. 0.011), setelah disesuaikan pendidikan, pendapatan, usia, dan pengetahuan.

#### CONTOH INTERPRETASI PANJANG:

Berdasarkan analisis multivariat penasun yang memiliki sikap positif terhadap konsep *harm reduction* memiliki kecenderungan (*likelihood*) 3,8 kali lebih besar untuk mengakses LJASS dari penasun yang memiliki sikap negatif. Di populasi, dengan derajat kepercayaan 95%, sikap positif penasun yang akses LJASS berada pada rentang 1,4 hingga 10,9 lebih tinggi dibandingkan penasun yang tidak akses LJASS. Angka signifikansi menegaskan, adanya hubungan antara sikap penasun dan perilaku akses LJASS (OR 3,8 (95% CI 1,4-10,9), p. 0,012), setelah disesuaikan pendidikan, pendapatan, usia, dan pengetahuan.

Silakan interpretasi variabel lainnya!





# BAB 5

## REGRESI COX

### (BIVARIAT-MULTIVARIAT)

**Kompetensi Dasar**

Mampu menjelaskan aplikasi uji regresi Cox pada analisis survival

**Indikator Keberhasilan**

Mampu melakukan proses stata regresi Cox dan menginterpretasikan *hazard ratio*

**Materi Pembelajaran**

1. Introduksi Konsep Analisis Survival dan Regresi Cox
2. Aplikasi Stata pada Perhitungan Kaplar Meier dan *Proportional Hazard Ratio* serta Interpretasi
3. Aplikasi Stata pada Perhitungan Regresi Cox serta Interpretasi

# ANALISIS REGRESI UNTUK DATA SURVIVAL REGRESI COX SEDERHANA DAN BERGANDA

## INTRODUKSI KONSEP ANALISIS SURVIVAL DAN REGRESI COX

Ketika penelitian, kita sering kali menghadapi kenyataan bahwa banyak responden penelitian yang tidak terpantau hingga akhir penelitian (*loss of follow up*), terutama pada penelitian longitudinal seperti studi kohort prospektif atau eksperimental dalam jangka waktu tertentu, beberapa minggu, hari, bulan, bahkan bertahun-tahun. *Loss of follow up* ini dapat menghasilkan efek negatif terhadap hasil akhir penelitian jika jumlahnya cukup besar, melebihi 10–20% dari jumlah sampel semula. Akan tetapi, dengan analisis survival, kita bisa menganalisis responden yang *loss of follow up* (hilang atau meninggal) dengan responden yang mengikuti hingga akhir penelitian, dengan memperhitungkan waktu/*time frame* selama mereka mengikuti proses penelitian.

Kleinbaum (1997) dalam Besral (2012), menjelaskan bahwa analisis survival adalah kumpulan dari prosedur statistik untuk menganalisis data di mana variabel dependen (*outcome*) yang diteliti adalah waktu sampai suatu kejadian atau efek/penyakit, kondisi kesehatan (*event*) muncul. Kleinbaum (1997) dan Tabachnick (2001) dalam Besral (2012) menambahkan tujuan dari analisis survival antara lain: (1) mengestimasi/memperkirakan dan menginterpretasikan fungsi survivor atau hazard dari data survival, misalnya, kanker, meninggal, setelah operasi, (2) membandingkan fungsi survivor dan fungsi hazard pada dua kelompok atau lebih, dan (3) menilai hubungan variabel *explanatory* dari *survival time* atau menentukan faktor determinan ketahanan dari munculnya suatu kejadian (*event*).

Ada beberapa cara dalam melakukan analisis survival, antara lain metode Kaplan Meir, Metode tabel kehidupan (*life table*), dan regresi Cox. Pada bab ini, akan kita pelajari bagaimana aplikasi regresi Cox, sebuah fungsi eksponensial dari covariat dan menghitung Hazard Ratio. Regresi Cox atau regresi hazard proporsional (*proportional hazards regression*) berasumsi bahwa rasio hazard pada beberapa kelompok terpapar yang berbeda konstan setiap waktu. Regresi Cox sama seperti halnya *risk ratio* dengan membandingkan nilai subjek atau responden pada *exposure* dan *outcome*, tetapi pada regresi Cox, waktu dari setiap waktu *follow up* sangat dipertimbangkan, sehingga *exposure* pada subjek dengan *outcome* positif dapat dibandingkan dengan subjek dengan *outcome* yang masih diikuti dan *outcome* negatif (Kirkwood, 2003).

$$H(t) = H_0(t) \times \exp (B_1 \times 1 + B_2 \times 2+ \dots + B_p \times p)$$

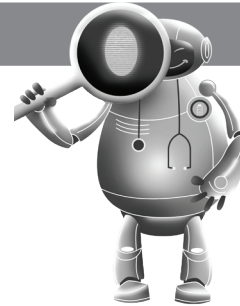
### Keterangan:

H(t) = hazard pada satu waktu  
 H<sub>0</sub> = hazard pada *baseline*  
 t = waktu  
 Exp = *exposure*

Sumber: Betty, R. K., & Jonathan, A. C. 2003. *Essential medical statistics*. Kirkwood and Jonathan AC Sterne: Blackwell Science Ltd, 414, 425.  
 Besral. 2012. *Regresi Cox Multivariat Analisis Survival Data Riset Kesehatan*. Departement Biostatistika, FKM, Universitas Indonesia.

# APLIKASI STATA PADA REGRESI COX

## STUDI KASUS 5.1



Tempat Penitipan Anak dan Gejala Influenza

Berikut adalah gambaran kondisi penyakit influenza atau disebut sebagai ILI (*an influenza-like illness*) di Rumah Sakit Anak Royal di Melbourne, Australia.

 Buka Data: Dataset revs\_LYLE.dta

*Sumber data: Lyle Gurrin, 2008, Assignment of Survival Analysis Course, School of Population Health, The University of Melbourne.*

**Dataset revs\_LYLE.dta** adalah dari **Respiratory Virus Study (ReVS)**, yang dilakukan di Melbourne oleh Vaccine Immunization Research Group (VIRGO), sebuah kolaborasi antara Jurusan Kesehatan Masyarakat dan Murdoch Children's Research Institute di Royal Children's Hospital. Penelitian ini adalah studi kohort prospektif dengan 234 anak yang terdaftar (kurang dari lima tahun pada saat pendaftaran) dan keluarga mereka terdaftar antara 17 Januari 2003 dan 31 Januari 2004. Banyak anak yang terdaftar sakit beberapa kali, episode ILI selama masa studi, dan setidaknya tiga hari sehat sebelum gejala influenza (ILI) terjadi lagi. Semua insiden setiap anak dicatat pada dataset, sehingga beberapa anak pada penelitian memberikan kontribusi lebih dari satu "*time-to-event*" observasi (ada 894 pengamatan total).

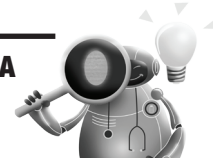
Berikut penjelasan data set

<b>kid</b>	Individual identifier (child identifier)
<b>id</b>	Individual identifier (individual ILI disease episode)
<b>enterdate</b>	Date the observation period started
<b>exitdate</b>	Date the observation period ended
<b>dob</b>	Child date of birth
<b>flu</b>	1 if observation period ends in ILI, 0 otherwise
<b>illtime</b>	Number of days after observation period started that an ILIevent occurred (if flu = 1) or time under observation to end of study (if flu = 0)
<b>season</b>	Season (1 = summer1 = summer 2003, 2 = autumn, 3 = winter, 4 = spring, 5 = summer2 = summer 2004)
<b>sex</b>	Sex (0 = male, 1 = female)
<b>bfccurrent</b>	0=never breastfed, 1=previously breastfed, 2=currenty breastfed
<b>childcare</b>	Child attends either formal or informal child care
<b>hhsiz</b>	Total number of people living in the household
<b>hhsmoke</b>	At least one adult in the household smokes

Hipotesis 0 = *Hazard ratio* ILI pada kelompok anak di *childcare* dan tidak di *childcare* adalah sama ( $h_0 = 1$ )

Hipotesis alternatif = *Hazard ratio* ILI pada kelompok anak di *childcare* dibandingkan pada kelompok anak tidak di *childcare* berbeda ( $h_a \neq 1$ )

### JAWABAN DETEKTIF STATA



Langkah-langkah:

1. Untuk melihat deskripsi data penelitian ini, kita lakukan CODEBOOK.

**Stata Command:**

**codebook kid id dob enterdate exitdate illtime flu season sex bcurrent childcare  
hsize hsmoke**

2. Sebelum melakukan analisis survival untuk data ini, kita perlu menetapkan/mendefinisikan data dan memerintahkan pada stata untuk mengingatnya. Pertama, waktu sakit (*time to illness*) dinyatakan sebagai waktu kegagalan (*time to failure*) berupa variabel "*illtime*". Variabel "*flu*" ditentukan sebagai status failure, dan *id* dinyatakan sebagai nomor identitas responden. Perintah: set *illtime*, failure (*flu*), *id* (*id*).

**Stata Command:**

**stset illtime, failure(flu) id(id)**

3. Menghitung risk atau rate ILI pada kelompok yang mengikuti *childcare* dan tidak mengikuti *childcare*, lalu menghitung *crude rate rasio* ILI pada kedua kelompok ini.

● **TABEL 5.1** Crude Rates ILI (CI 95%) dalam Kelompok yang Mengikuti dan Tidak Mengikuti Penitipan Anak (Childcare)

Penitipan Anak (Childcare)	Total ILI	Total pyar* (per 1000)	Rate	95% Confidence Interval
Tidak	189	15.713	12.03	10.43 to 13.87
Ya	492	30.550	16.11	14.74 to 17.59

*pyar\** person years at risk

**Stata Command:**

**strate childcare, per(1000)**

**Output Stata:**

childc~e	D	Y	Rate	Lower	Upper
0	189	15.7130	12.028	10.430	13.871
1	492	30.5500	16.105	14.743	17.593



**Interpretasi:**

*Crude rate ILI (Influenza Like Illness)* pada kelompok yang tidak menggunakan penitipan anak adalah 12,03 (95% CI 10,4-13,9) per 1.000 orang-tahun berisiko. Kejadian ILI pada kelompok yang tidak menghadiri penitipan anak antara 10,43 hingga 13,87 per 1.000 orang-tahun.

*Crude rate ILI* pada kelompok yang menggunakan penitipan anak adalah 16,11 (95% CI 14,74-17,59) per 1.000 orang-tahun. Kejadian ILI pada kelompok anak yang menghadiri penitipan anak antara 14,74 dan 17,59 per 1.000 orang-tahun.

● **TABEL 5.2** Rate Ratio ILI Membandingkan Kelompok Menggunakan dan Tidak Menggunakan Penitipan Anak

ILI	Hazard Ratio	P value	95% Confidence Interval
Penitipan Anak	1.34	0.0006	1.13 to 1.58

**Stata Command:**

```
stmh childcare
```

**Interpretasi:**

Rate rasio ILI pada anak-anak di *childcare* adalah 1,34 kali lebih tinggi dibandingkan dengan anak-anak yang tidak *childcare*. Dengan derajat kepercayaan 95% bahwa *rate ratio* ILI pada anak-anak di *childcare* antara 1,13 hingga 1,58 kali lebih tinggi dibandingkan anak-anak yang tidak mengikuti *childcare* di populasi. *P value* 0,0006 menunjukkan bukti kuat terhadap hipotesis nol ( $H_0$ ) bahwa ada risiko ILI lebih tinggi pada anak yang *childcare* dibanding yang tidak *childcare*.

4. Menghitung *Hazard ratio* dengan tiga pendekatan
  - a. **Kaplan-Meier**

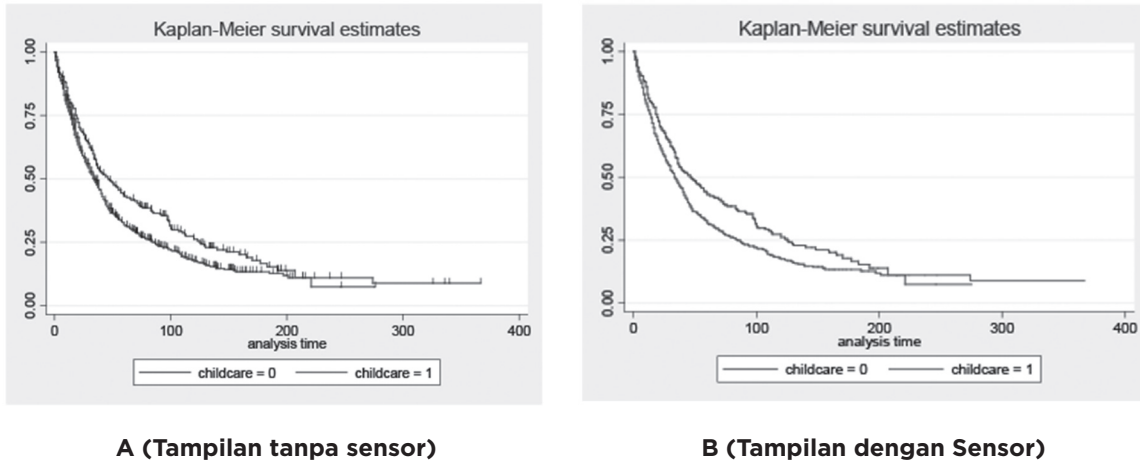
**Stata Command:**

```
sts graph, by(childcare) gwood
sts graph, by(childcare) na tmin(0.05) yscale(log) xscale(log)
sts graph, by(childcare) censored(multiple)
```

**Interpretasi:**

Probabilitas sehat tanpa menderita ILI di kelompok anak yang tidak mengikuti *childcare* lebih tinggi dibandingkan dengan kelompok anak di *childcare*. Pada saat 100 hari *follow-up*, kemungkinan bebas influenza pada kelompok yang menggunakan penitipan anak adalah

- **GAMBAR 5.1** Probabilitas Survival Kumulatif Kaplan-Meier pada Kelompok yang Menggunakan dan Tidak Menggunakan Penitipan Anak pada 234 Anak



sekitar 23% dan tidak menggunakan penitipan anak adalah sekitar 28%. Setelah 200 hari *follow-up*, probabilitas ILI untuk kedua kelompok cenderung saling tumpang tindih (*overlapping*). Hal ini bisa disebabkan oleh banyaknya peserta pada kedua kelompok yang disensor sebelum 200 hari *follow-up*, hanya beberapa anak yang bebas dari penyakit dan sejumlah kecil anak yang disebutkan ditindaklanjuti pada akhir penelitian.

b. **Hazard Rasio dengan Mantel-Cox (Regresi Cox)**

**Stata Command:**

**stcox childcare**

Hasil prosedur Mantel-Cox menunjukkan bahaya ILI di kelompok anak yang menggunakan *childcare* adalah hampir 1,29 kali tinggi dibandingkan kelompok anak yang tidak menggunakan *childcare*. Interval kepercayaan 95% menandakan bahwa risiko ILI di populasi dapat meningkatkan sekecil 1,09 atau setinggi 1,52 dalam kelompok anak di *childcare* dan tidak *childcare*. P value 0,003 menunjukkan bukti kuat untuk menolak  $H_0$  dari HR = 1, maka risiko ILI berbeda antara dua kelompok anak dalam populasi.

- **TABEL 5.3** Hasil Regresi Mantel-Cox dengan Anak Sebagai Model Variabel

ILI	Rate ratio	P value	95% Confidence Interval
Penitipan Anak	1.29	0.003	1.09 to 1.52

*Hazard ratio* dan rasio rata-rata kasar (*crude rate ratio*) membandingkan kelompok yang menggunakan dan tidak menggunakan penitipan anak masing-masing hampir sama dengan 1,29 (95% CI 1,09-1,52) dan 1,34 (95% CI 1,13-1,58). Rasio rata-rata tidak berubah terlalu banyak sebelum disesuaikan waktu dibandingkan dengan *hazard ratio* setelah disesuaikan waktu dalam regresi Cox. Hal ini dapat terjadi karena waktu bukan faktor perancu dalam hubungan antara anak dan risiko ILI berdasarkan asumsi tidak ada interaksi antara waktu dan ILI, saat periode penyakit pendek (maksimum 1 tahun), dan *baseline rate* adalah konstan.

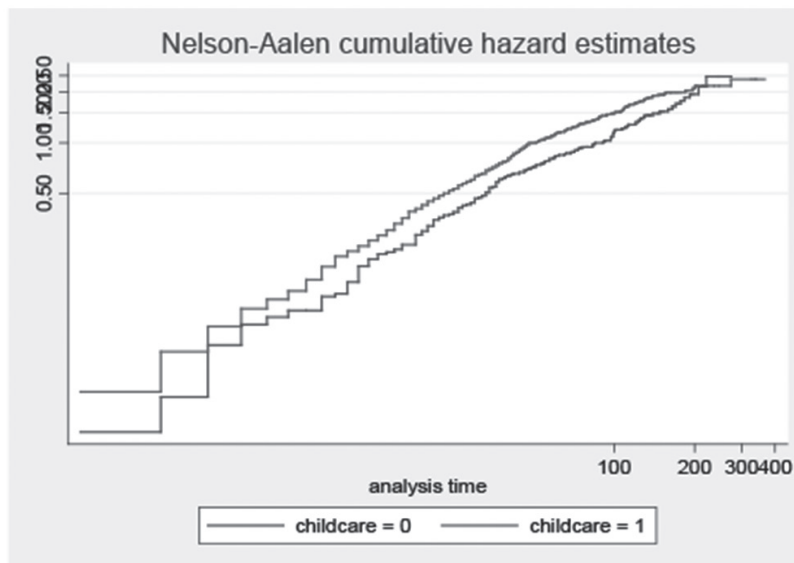
### c. Menilai Asumsi Proporsional Hazard (*Proportional Hazard Assumption*)

#### Stata Command:

```
sts graph,by(childcare) na yscale(log) xscale(log)
```

Asumsi bahaya proporsional terlihat dari kumulatif ILI di antara kelompok yang menggunakan penitipan anak dan tidak menggunakan penitipan anak terlihat paralel. Hasilnya, tidak ada pelanggaran utama asumsi PH dari fungsi *hazard ratio* pada kedua pengamatan, artinya HR tidak tergantung pada waktu atau waktu bukan faktor perancu pada kasus ini.

### ● GAMBAR 5.2 Nelson-Aalen Perkiraan Cumulative Hazard untuk Variabel Penitipan Anak



## 5. Menginvestigasi variabel perancu pada Regresi Cox

● **TABEL 5.4** Hasil Regresi Mantel-Cox dengan Anak Sebagai Variabel dalam Model (Sebelum Penyesuaian Kovariat Lainnya)

ILI	Rate ratio	P value	95% Confidence Interval
Penitipan Anak	1.29	0.003	1.09 to 1.52

Menggunakan perubahan 5% dalam estimasi rasio hazard sebagai ambang batas untuk memeriksa faktor perancu dari hubungan antara penitipan anak dan risiko ILI. Perkiraan bahaya kumulatif Nelson-Aalen dilakukan dengan penilaian visual informal (metode grafis) untuk memeriksa kesesuaian asumsi hazard proporsional.

a. **Identifikasi Musim sebagai faktor perancu dalam asosiasi risiko ILI dan penitipan anak**

Variabel musim terdiri atas lima kategori: musim panas 1, musim panas 2, musim dingin, musim semi, dan musim gugur. Rasio kemungkinan membandingkan model dengan menggabungkan musim panas 1 dan 2 dalam model dengan musim panas 1 dan 2 secara terpisah menunjukkan bukti kuat terhadap hipotesis nol dari rasio bahaya ILI pada penitipan anak adalah sama di musim panas 1 dan 2 digabung atau tidak. Oleh karena itu, musim panas 1 dan musim panas 2 harus dilakukan secara terpisah (LR Chi2 (1) = 112,61, Prob> Chi2 <0,0001).

**Stata Command:**

```
** pada data ada dua musim panas (summer)
generate season1=season
recode season1 5=1
tab season1
```

```
** kita harus memutuskan apakah kedua summer akan kita
gabungkan menjadi satu summer saja atau dua summer yang
berbeda (summer 1 dan summer 2) xi: stcox i.season1
est store seasoncat1
stcox i.season
est store seasoncat2
lrtest seasoncat1 seasoncat2
```

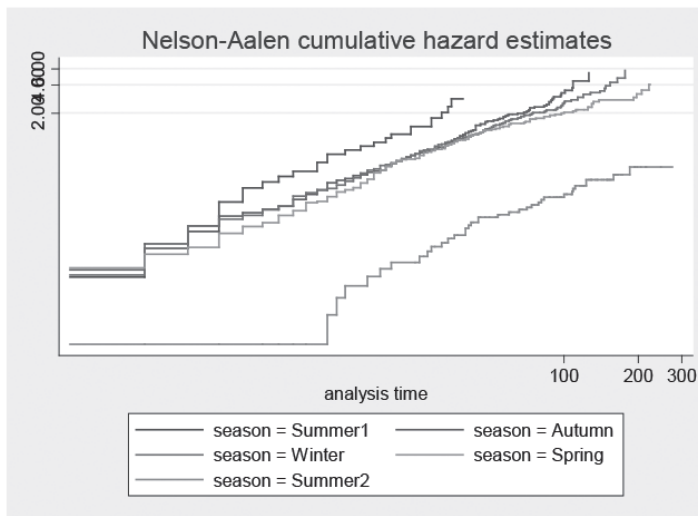
```
**Mengecek Proportional hazard assumption)
sts graph, by(season) na tmin(0.1) yscale(log) xscale(log)
```

```
stcox season, tvc(season) texp(ln(_t))
stcox season, schoenfeld(res)
estat phtest
```

```
** kemungkinan season berinteraksi dengan childcare, lalu
kita stratifikasi
stcox childcare, str(season)
```

**OUTPUT STATA:**

● **GAMBAR 5.3** Nelson-Aalen Cumulative Hazard Estimates



```
. stcox childcare, str(season)
      failure _d: flu
      analysis time _t: alltime
      id: id
```

```
Iteration 0: log likelihood = -2675.3122
Iteration 1: log likelihood = -2670.2483
Iteration 2: log likelihood = -2670.236
Iteration 3: log likelihood = -2670.236
Refining estimates:
Iteration 0: log likelihood = -2670.236
```

Stratified Cox regr. -- Breslow method for ties

```
No. of subjects =      779           Number of obs   =      779
No. of failures =      623
Time at risk    =    34149
Log likelihood  =   -2670.236
LR chi2(1)     =      10.15
Prob > chi2    =      0.0014
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
childcare	1.329595	.1212381	3.12	0.002	1.111996 1.589775

Stratified by season

■ **TABEL 5.5** Hasil Regresi Cox dengan Status Childcare Anak Disesuaikan dengan Variabel Musim

_t	Hazard ratio	P value	95% Confidence Interval
Penitipan Anak	1.33	0.002	1.11 to 1.59

Musim akan dijadikan sebagai variabel kategorik karena bukan angka numerik dan juga variabel kategori non-ordered.

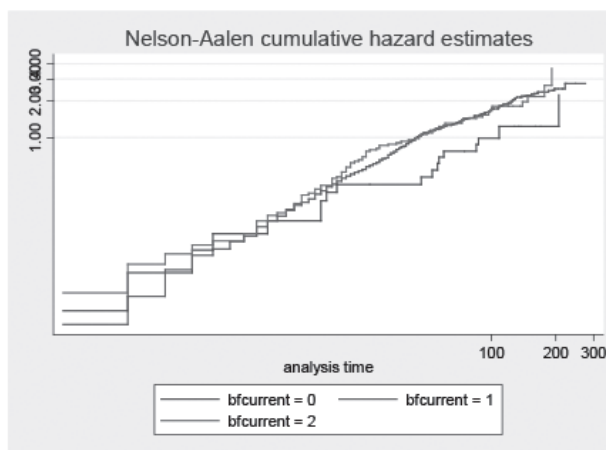
Perkiraan bahaya kumulatif Nelson-Aalen untuk variabel musim terlihat tidak paralel karena tumpang tindih dan melewati dari rasio bahaya dari musim panas 1, musim dingin, musim semi, dan musim gugur. Oleh karena itu, asumsi hazard proporsional tidak memenuhi syarat (pengaruh perubahan musim dengan waktu). Model regresi akan dikelompokkan berdasarkan nilai musim dengan asumsi non-PH.

Estimasi rasio bahaya ILI pada kelompok anak yang ditiip pada *Childcare* menggunakan 1.33 (95% CI 1,11-1,59) lebih tinggi dibandingkan dengan kelompok anak yang tidak ditiipkan pada *Childcare* setelah disesuaikan dengan musim. Perubahan persen dari perkiraan disesuaikan hanya 3%. Itu tidak mengubah lebih dari 5% dan karena musim tidak dianggap menjadi pembaur.

**b. Bfcurrent (status menyusui sekarang) sebagai kemungkinan perancu dalam hubungan risiko ILI dan penitipan anak**

Bfcurrent dipertahankan sebagai variabel kategoris karena bfcurrent tidak dapat ditandai dengan skor dan juga merupakan variabel kategori *non-ordered*. Selain itu, salah satu variabel kategoris (0 = tidak pernah disusui) yang terdapat angka yang rendah untuk meningkatkan kekuatan untuk analisis dalam setiap kelompok, sebagai hasilnya, hampir variabel biner jika kita menggabungkan dengan jumlah kelompok ASI lain.

● **GAMBAR 5.4** Nelson-Aalen Cumulative Hazard Estimates Bfcurrent



Perkiraan hazard kumulatif Nelson-Aalen untuk variabel `bfcurren` terlihat paralel. Hanya ada beberapa tumpang tindih dan melewati fungsi *hazard ratio*. Oleh karena itu, tidak ada pelanggaran utama dari asumsi hazard proporsional (lihat output Stata berikut).

**Stata Command:**

```
strate bfcurren, graph yscale (log)
**Mengecek Proportional hazard
stcox bfcurren, tvc(bfcurren) texp(ln(_t))

sts graph, by(bfcurren) na tmin(0.1) yscale(log) xscale(log)
```

**OUTPUT STATA:**

```
**Bfcurren sebagai possible confounder
stcox childcare bfcurren
xi:stcox childcare i.bfcurren
```

```
. stcox childcare bfcurren
      failure _d:  flu
analysis time _t: 111time
              id:  id

Iteration 0:  log likelihood = -3664.1644
Iteration 1:  log likelihood = -3655.0071
Iteration 2:  log likelihood = -3654.9672
Iteration 3:  log likelihood = -3654.9672
Refining estimates:
Iteration 0:  log likelihood = -3654.9672

Cox regression -- Breslow method for ties

No. of subjects =      779          Number of obs =      779
No. of failures =      623
Time at risk   =     34149
Log likelihood = -3654.9672          LR chi2(2) =      18.39
                                      Prob > chi2 =      0.0001
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
childcare	1.457837	.1415018	3.88	0.000	1.205283 1.763312
bfcurren	1.354473	.1409494	2.92	0.004	1.104568 1.660918

```
.
.
.
. xi:stcox childcare i.bfcurren
i.bfcurren      _ibfcurren_0-2 (naturally coded; _ibfcurren_0 omitted)
```

```
      failure _d:  flu
analysis time _t: 111time
              id:  id

Iteration 0:  log likelihood = -3664.1644
Iteration 1:  log likelihood = -3654.7276
Iteration 2:  log likelihood = -3654.6362
Iteration 3:  log likelihood = -3654.6361
Refining estimates:
Iteration 0:  log likelihood = -3654.6361

Cox regression -- Breslow method for ties

No. of subjects =      779          Number of obs =      779
No. of failures =      623
Time at risk   =     34149
Log likelihood = -3654.6361          LR chi2(3) =      19.06
                                      Prob > chi2 =      0.0003
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
childcare	1.437043	.1414947	3.88	0.000	1.184837 1.742935
_ibfcurren-1	1.595699	.3737293	2.00	0.046	1.008303 2.52529
_ibfcurren-2	2.049221	.5296223	2.78	0.006	1.234796 3.40081

● **TABEL 5.6** Hasil regresi Cox dengan Perawatan Anak dan Variabel *bfcurent* dalam Model

<b>_t</b>	<b>Hazard ratio</b>	<b>P value</b>	<b>95% Confidence Interval</b>
Childcare (0=tidak,1=ya)	1.40	<0.0001	1.17 to 1.68
Bfcurent (1= tidak pernah menyusui sebagai ref. group)			
Sebelumnya menyusui	1.59	0.046	1.01-2.52
Sekarang menyusui	2.05	0.006	1.23-3.40

Estimasi rasio bahaya ILI untuk kelompok anak menghadiri penitipan anak dan tidak menghadiri penitipan anak setelah disesuaikan dengan status ASI eksklusif telah berubah sebesar 8,5%. Perubahan rasio bahaya ILI sebelum dan sesudah disesuaikan faktor ASI eksklusif (*bfcurent*) dalam model regresi adalah lebih dari 5% sebagai ambang batas. Akibatnya, faktor ASI eksklusif mungkin sebagai perancu dalam hubungan antara risiko ILI dan penitipan anak dengan asumsi hazard proporsional.

### LATIHAN:



Masih menggunakan Dataset *revs\_Lyle.dta*. Coba Anda eksplorasi apakah jenis kelamin, jumlah anggota keluarga, dan keberadaan perokok dalam rumah anak merancu hubungan antara status childcare dan risiko ILI!

#### Age in a year

```
stset exitdate, failure(flu) enter(enterdate) scale(365.25) origin(dob) id(id)
```

#### Age in months

```
stset exitdate, failure(flu) enter(enterdate) scale(30.35) origin(dob) id(id)
```

**Langkah 2**, usia anak-anak yang dihasilkan dalam satu tahun atau satu bulan dan kemudian diringkas untuk mendapatkan usia minimum dan maksimum dari anak-anak.

#### Stata Command:

##### Age in year

```
gen age=enterdate-dob
replace age=age/364.25
su age
```

##### Age in month

```
gen agem=enterdate-dob
replace agem=agem/30.35
su agem
```



**Langkah 3**, usia dibagi menjadi 5 kelompok umur (berbagai usia satu tahun) atau ke-12 kelompok (kisaran usia 6 bulan), sehingga kita bisa membagi ILI ke dalam kelompok usia dan risiko tertular ILI. Command stata yang digunakan *stsplit*. Setiap kelompok usia dapat dihasilkan untuk menyelidiki rasio tingkat ILI pada setiap kelompok umur dan kelompok yang lebih rentan terhadap ILI (usia minimal anak adalah sekitar 6 bulan dan usia maksimal anak-anak adalah 5 tahun).

**Stata Command:**

Age in years

stsplit agegrp, at(0(1)5)

Age in months (range 6 months)

stsplit agegrp, at(0(6)60)



# BAB 6

## PENGANTAR SPSS

### Kompetensi Dasar

Mampu menjelaskan analisis deskriptif statistik kesehatan menggunakan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan *icon* penting dalam SPSS.
- Mampu menjelaskan cara membuat *template* dalam SPSS.
- Mampu menjelaskan cara *entry data* dalam SPSS.
- Mampu menjelaskan data editor dalam SPSS.

### Materi Pembelajaran

1. Icon Penting SPSS
2. Membuat Template
3. Entry Data
4. Data Editor

## ICON PENTING SPSS

SPSS adalah *software* pengolahan data yang penggunaannya sangat tergantung dari penguasaan materi statistik sekaligus pemahaman perintah-perintah atau menu-menu di dalamnya. Oleh karena itu berikut akan dibahas cara mengoperasikan SPSS.

### MEMULAI SPSS

Berikut adalah langkah-langkah memulai operasi SPSS.

1. Klik **START** → **Program** → **IBM SPSS Statistics** → **IBM SPSS Statistics**.
2. Anda bisa membuka data yang sudah tersedia “*open an existing data source*” atau membuka data kosong dengan memilih “*type in data*”

### MEMBUKA DATA

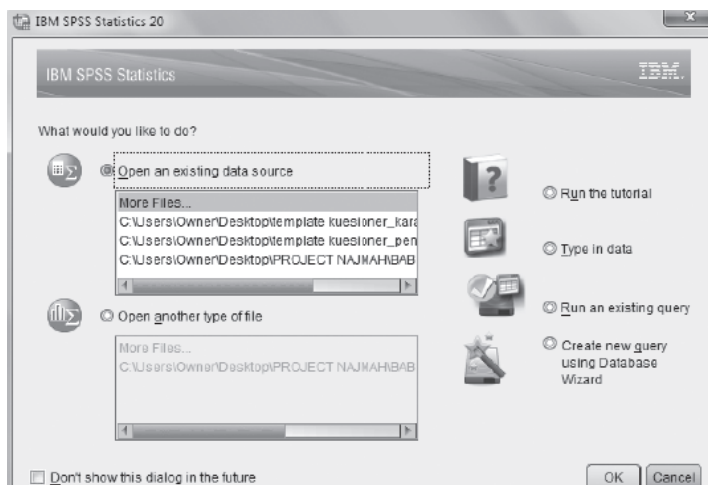
Dari menu “*file*” - **Open** → data atau bisa klik lambang berikut ini pada ‘*toolbar*’<sup>(9, 12)</sup>.



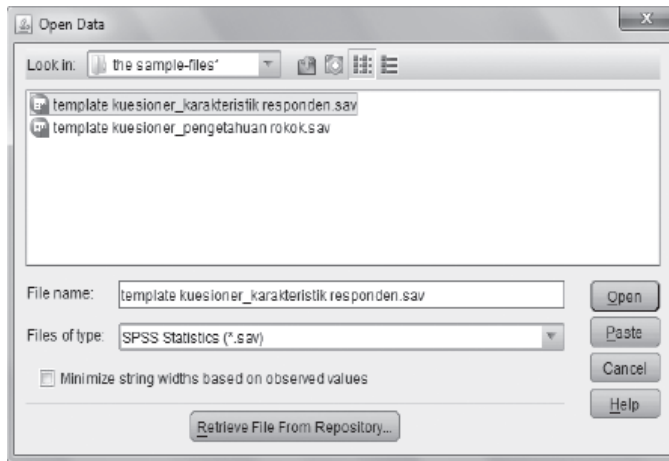
Untuk melihat contoh data dalam program SPSS, dapat dilakukan dengan:

- Klik 2 kali folder “tutorial”
- Klik 2 kali folder “the sample-files”
- Klik file dengan judul “template kuesioner\_karakteristik responden.sav” atau file lainnya
- Klik “Open” untuk membuka data SPSS

### ● GAMBAR 6.1 Proses Memanggil Program SPSS<sup>(12)</sup>



● **GAMBAR 6.2** Tampilan ‘Open Sample Files’



● **GAMBAR 6.3** Tampilan Data “template kuesioner\_karakteristik responden.sav”

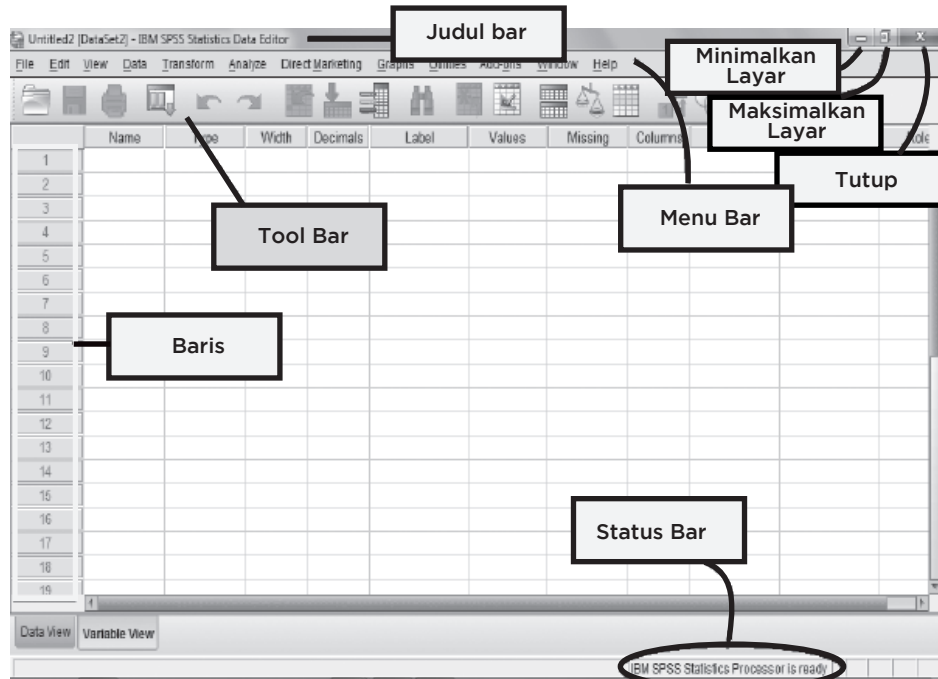
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	fakultas	Numeric	8	0	Fakultas Respo...	{1, FE}...	None	8	Right	Ordinal	Input
2	nama	String	20	0	Nama responden	None	None	8	Right	Nominal	Input
3	NIP_NIDN	Numeric	20	0	NIP/NIDN respo...	None	None	8	Right	Scale	Input
4	umur	Numeric	8	0	umur responden	None	None	8	Right	Scale	Input
5	didik	Numeric	8	0	pendidikan resp...	{1, SD}...	None	8	Right	Ordinal	Input
6	lamakerja	Numeric	8	0	lama kerja resp...	None	None	8	Right	Scale	Input
7	statuskawin	Numeric	8	0	status perkawin...	{1, lajang (li...	None	8	Right	Nominal	Input
8	kerjaistri	Numeric	8	0	pekerjaan istri	{1, tidak be...	None	8	Right	Nominal	Input
9	jumlahanak	Numeric	8	0	jumlah anak	None	None	8	Right	Nominal	Input
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											

## MENU UTAMA SPSS

Menu utama pada SPSS, di antaranya:

1. **FILE**  
Digunakan untuk membuat file data baru, membuka file yang tersimpan, atau membaca file dari program lain, menyimpan file, mencetak, dan lain-lain.
2. **EDIT**  
Digunakan untuk menyalin, menghapus, mencari, mengganti data, dan lain-lain.

● **GAMBAR 6.4** Menu Utama IBM SPSS Statistics 20



### 3. DATA

Digunakan untuk membuat/mendefinisikan variabel, mengambil/menganalisis sebagian data, menggabungkan data, menambah variabel, dan lain-lain.

### 4. TRANSFORM

Digunakan untuk transformasi/modifikasi data, seperti pengelompokan variabel, pembuatan variabel baru, dan lain-lain.

### 5. ANALYZE

Digunakan untuk melakukan/memilih berbagai analisis statistik, dari statistik deskriptif sampai statistik multivariat.

### 6. GRAPHS

Digunakan untuk membuat dan menampilkan grafik, meliputi grafik batang, pie, garis, histogram, scatter plot, dan lain-lain.

### 7. UTILITIES

Digunakan untuk menampilkan berbagai informasi tentang isi file.

### 8. WINDOW

Digunakan untuk berpindah-pindah antar-jendela/layar, misalnya dari jendela data editor ke *output*.

### 9. HELP

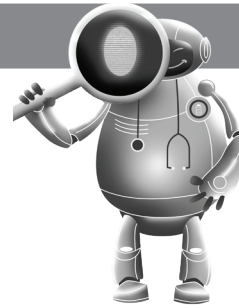
Memuat informasi bantuan bagaimana menggunakan berbagai fasilitas pada SPSS.

# MEMBUAT TEMPLATE

Mari kita aplikasikan studi kasus di bawah ini!

## STUDI KASUS 6.1

Tim Peneliti (Najmah, Fenny Etrawati, Feranita Utama, dan Yeni, 2015) akan melakukan penelitian terkait dengan intervensi terpadu pengurangan dampak buruk (*harm reduction*) asap rokok pada ruangan tertutup/ber-AC di lingkungan Universitas Sriwijaya. Tahap awal kuesioner, tim peneliti mengumpulkan karakteristik responden.



Buka: template kuesioner\_karakteristik responden.sav

### PERTANYAAN DETEKTIF SPSS

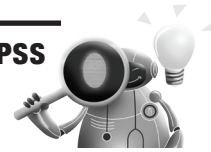
Buatlah template SPSS terkait dengan karakteristik responden sesuai dengan kuesioner di bawah ini.



**KUESIONER PENELITIAN HIBAH KOMPETITIF UNIVERSITAS SRIWIJAYA TAHUN 2015 INTERVENSI TERPADU PENGURANGAN DAMPAK BURUK (HARM REDUCTION) ASAP ROKOK PADA RUANGAN TERTUTUP/BER-AC DI LINGKUNGAN UNIVERSITAS SRIWIJAYA**

RAHASIA				
<b>I. PENGENALAN TEMPAT</b>				
1	Fakultas	1. FE 2. FH 3. FT 4. FK 5. FP	6. FKIP 7. FMIPA 8. FISIP 9. FASILKOM 10. FKM	££
<b>II. KARAKTERISTIK RESPONDEN</b>				
2	Nama :			
3	NIP/NIDN :			
4	Umur :			
5	Pendidikan terakhir : 1. SD 2. SMP 3. SMA /SMK/ sederajat 4.D1/D3 5.PT (a. S1 b. S2 c.S3)			
6	Lama kerja :            tahun			
7	Status perkawinan : 1. Lajang (Lanjut ke III) 2. Nikah 3. Cerai Hidup 4. Cerai Mati 5. Lainnya (.....)			
8	Pekerjaan Istri : 1. Tidak Bekerja 2. Bekerja			
9	Jumlah anak : .....orang			

## JAWABAN DETEKTIF SPSS



## Langkah-Langkah Pembuatan Template

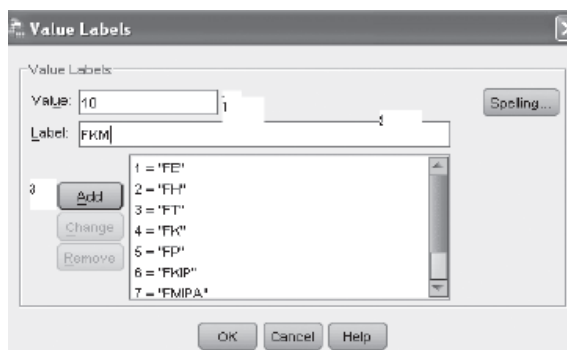
Klik 'Variable view' di bawah template SPSS



## 1. Variabel Fakultas

- Name* : ketik *fakultas*  
*Type* : *numeric*  
*Width* : secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai. Isi 8.  
*Decimal* : 0 (nol)  
*Label* : fakultas responden  
*Values* : 1. FE            5. FP            8. FISIP  
               2. FH            6. FKIP        9. FASILKOM  
               3. FT            7. FMIPA      10. FKM  
               4. FK  
*Missing* : tidak diisi  
*Columns* : 8  
*Align* : kanan (*right*)  
*Measure* : ordinal  
*Role* : *input*

## ● GAMBAR 6.5 Tampilan Variabel 1



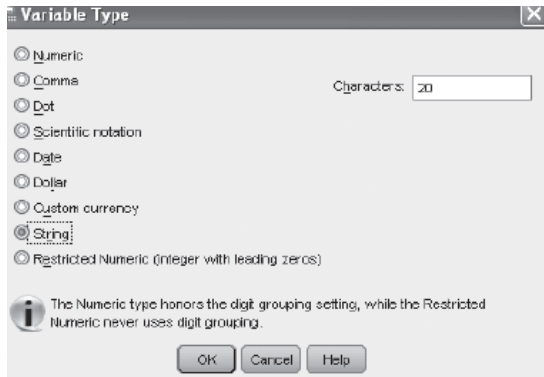
## 2. Variabel Nama

- Name* : ketik *nama*  
*Type* : *string*  
*Width* : secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai, isi 20.  
*Decimal* : 0 (nol)  
*Label* : nama responden  
*Values* : tidak diisi



*Missing* : tidak diisi  
*Columns* : 8  
*Align* : kanan (*right*)  
*Measure* : nominal  
*Role* : *input*

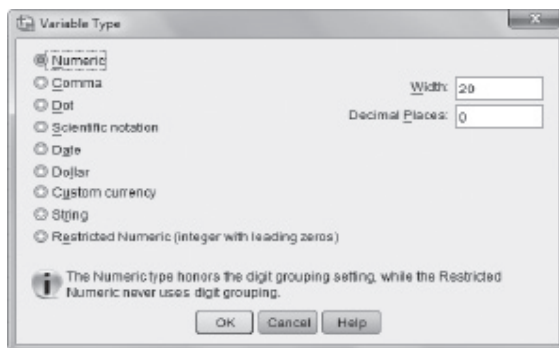
● **GAMBAR 6.6** Tampilan Variabel 2



3. **Variabel NIP/NIDN**

*Name* : ketik NIP/NIDN  
*Type* : *numeric*  
*Width* : secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai, isi 20.  
*Decimal* : 0 (nol)  
*Label* : NIP/NIDN responden  
*Values* : tidak diisi  
*Missing* : tidak diisi  
*Columns* : 8  
*Align* : kanan (*right*)  
*Measure* : scale  
*Role* : *input*

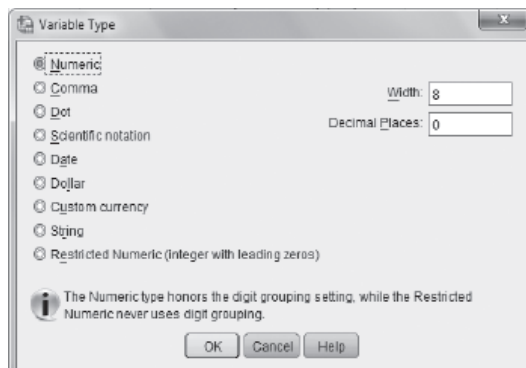
● **GAMBAR 6.7** Tampilan Variabel 3



#### 4. Variabel Umur

<i>Name</i>	: ketik <i>umur</i>
<i>Type</i>	: <i>numeric</i>
<i>Width</i>	: secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai, isi 8.
<i>Decimal</i>	: 0 (nol)
<i>Label</i>	: umur responden
<i>Values</i>	: tidak diisi
<i>Missing</i>	: tidak diisi
<i>Columns</i>	: 8
<i>Align</i>	: kanan ( <i>right</i> )
<i>Measure</i>	: scale
<i>Role</i>	: <i>input</i>

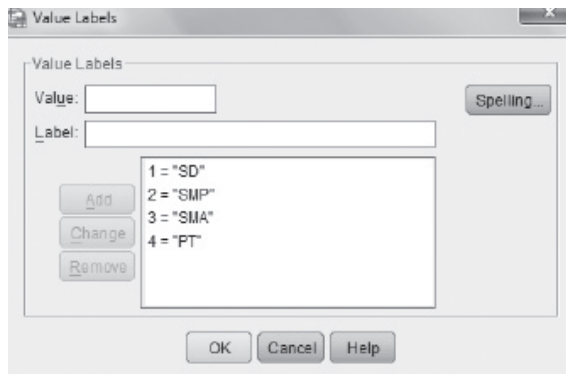
#### ● GAMBAR 6.8 Tampilan Variabel 4



#### 5. Variabel Pendidikan Terakhir

<i>Name</i>	: ketik <i>didik</i>
<i>Type</i>	: <i>numeric</i>
<i>Width</i>	: secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai, isi 8.
<i>Decimal</i>	: 0 (nol)
<i>Label</i>	: pendidikan responden
<i>Values</i>	: 1. SD 2. SMP 3. SMA 4. PT
<i>Align</i>	: kanan ( <i>right</i> )
<i>Measure</i>	: ordinal
<i>Role</i>	: <i>input</i>

● **GAMBAR 6.9** Tampilan Variabel 5



6. **Variabel Lama Bekerja**

*Name* : ketik *lamakerja*

*Type* : *numeric*

*Width* : secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai, isi 8.

*Decimal* : 0 (nol)

*Label* : lama kerja responden

*Values* : tidak diisi

*Missing* : tidak diisi

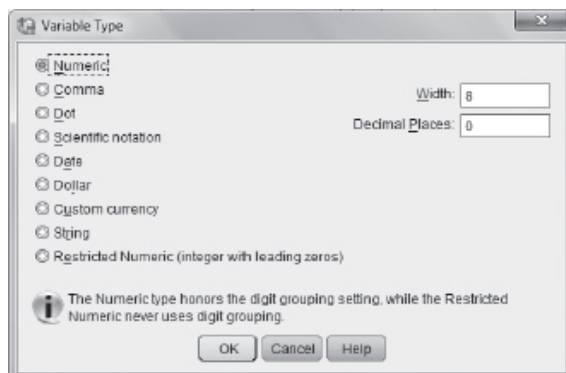
*Columns* : 8

*Align* : kanan (*right*)

*Measure* : scale

*Role* : *Input*

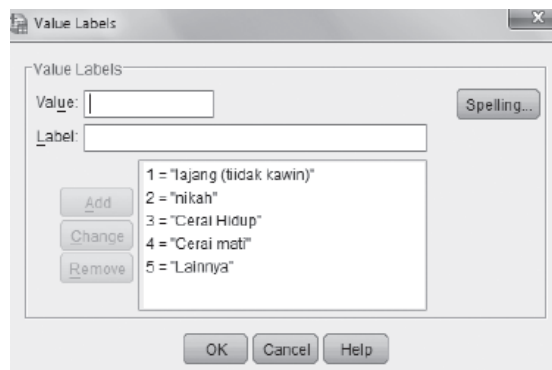
● **GAMBAR 6.10** Tampilan Variabel 6



## 7. Variabel Status Perkawinan

- Name* : ketik *statuskawin*  
*Type* : *numeric*  
*Width* : secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai. Isi 8.  
*Decimal* : 0 (nol)  
*Label* : status perkawinan  
*Values* : 1. Lajang/tidak kawin  
           2. Nikah  
           3. Cerai hidup  
           4. Cerai mati  
           5. Lainnya

● **GAMBAR 6.11** Tampilan Variabel 7

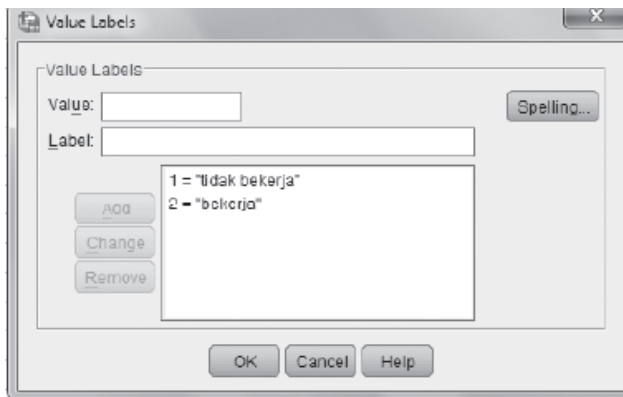


- Missing* : tidak diisi  
*Columns* : 8  
*Align* : kanan (*right*)  
*Measure* : nominal  
*Role* : *input*

## 8. Variabel Pekerjaan Istri

- Name* : ketik *kerjaistri*  
*Type* : *numeric*  
*Width* : secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai. Isi 8.  
*Decimal* : 0 (nol)  
*Label* : pekerjaan istri  
*Values* : 1. Tidak bekerja  
           2. Bekerja  
*Missing* : tidak diisi  
*Columns* : 8  
*Align* : kanan (*right*)  
*Measure* : nominal  
*Role* : *input*

● **GAMBAR 6.12** Tampilan Variabel 8



9. Variabel Jumlah Anak

*Name* : ketik *jumlahanak*

*Type* : *numeric*

*Width* : secara otomatis akan terisi pada saat mengisi tipe variabel, tapi jika tidak sesuai dapat diganti dengan angka yang sesuai. Isi 8.

*Decimal* : 0 (nol)

*Label* : jumlah anak

*Values* : tidak diisi

*Missing* : tidak diisi

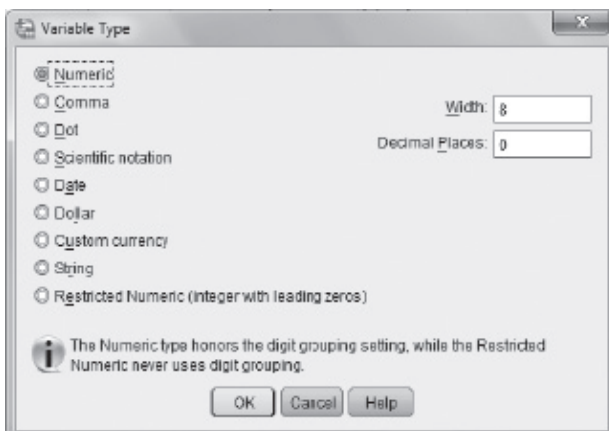
*Columns* : 8

*Align* : kanan (*right*)

*Measure* : nominal

*Role* : *Input*

● **GAMBAR 6.13** Tampilan Variabel 9



## ENTRY DATA

Masukkan data di bawah ini, klik “*data view*” untuk memasukkan data sesuai template yang telah kita buat tadi.

Data View Variable View

● **TABEL 6.1** Karakteristik Responden

No	Fakultas	Nama	NIDN	Umur	Didik	Lama kerja	Status kawin	Kerja istri	Jumlah anak
1	1	Ade	111110	58	5	26	2	1	5
2	2	Adi	111111	50	3	25	2	1	4
3	3	Mamad	111112	49	5	25	2	1	3
4	4	Mustofa	111113	38	5	15	1	88	88
5	5	Tantowi	111114	53	3	31	2	2	4
6	6	Yan	111115	40	5	15	2	1	2
7	7	Iwan	111116	23	3	2	1	88	88
8	8	Dudi	111117	50	4	24	2	1	3
9	9	Anto	111118	37	5	13	2	1	1
10	10	Dian	111119	35	4	7	2	2	1
11	5	Ichsan	111120	25	2	29	2	1	3
12	5	Yani	111121	45	3	20	2	1	3
13	5	Ahmad	111122	40	5	15	2	2	3
14	6	Baim	111123	44	2	25	2	1	3
15	6	Repila	111124	53	3	22	2	2	5
16	6	Suwarto	111125	35	3	5	2	1	2
17	7	Feris	111126	37	3	11	2	1	3
18	7	Udin	111127	50	5	26	2	2	3
19	7	Wawan	111128	38	5	10	2	1	3
20	8	Harun	111129	48	5	27	2	1	3
21	8	Zainal	111130	57	3	35	2	1	6
22	8	Dedi	111131	23	2	2	1	88	88
23	9	Jahri	111132	50	3	19	2	2	3
24	4	Bael	111133	51	5	25	2	1	3
25	3	Anto	111134	51	3	30	2	1	1
26	2	Rahmat	111135	53	5	29	2	1	5
27	1	Desman	111136	46	3	23	2	1	3
28	10	Taufik	111137	50	2	26	2	1	2
29	10	Kusnan	111138	30	3	8	2	1	2
30	3	Usman	111139	28	3	1	1	88	88

\*\*Bukan data sebenarnya, hanya contoh

# DATA EDITOR

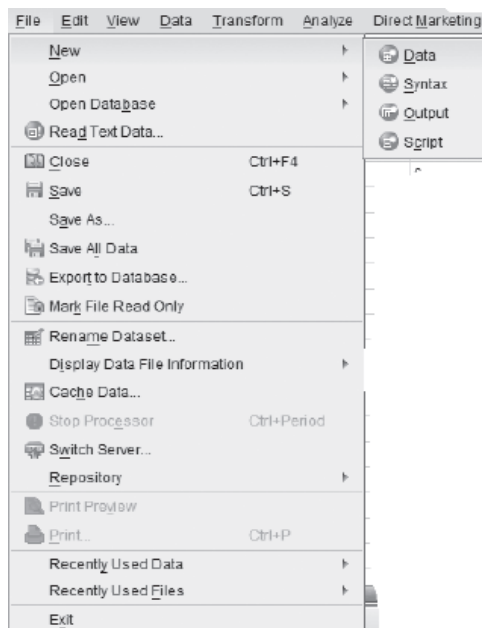
Data editor berkaitan erat dengan manajemen data atau pengelolaan data. Hal terpenting dalam manajemen data adalah menyimpan data, mengganti nilai data, menghapus, *me-copy* dan memindah sel, dan masih banyak lagi.<sup>(9, 12)</sup>

## 1. Membuka Data Kosong dalam Format SPSS

Langkah-langkah:

- Klik menu ***File*** → ***New*** → ***Data***

### ● GAMBAR 6.14 Proses Membuka Data Kosong



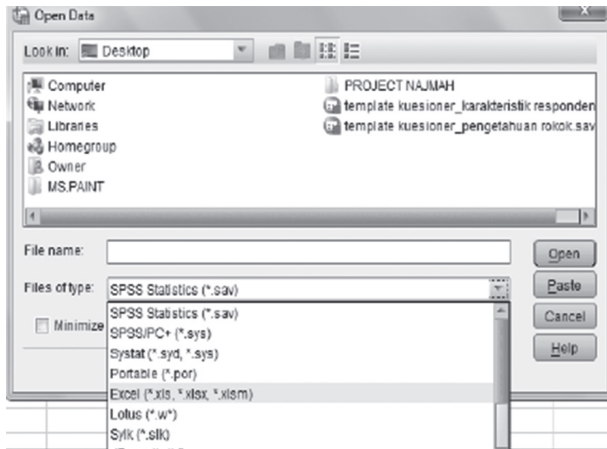
- Pilih file yang akan dibuka
- Klik *Open* jika yakin atau *Cancel* jika akan dibatalkan

## 2. Mengambil data yang tidak dalam format SPSS

Langkah-langkah:

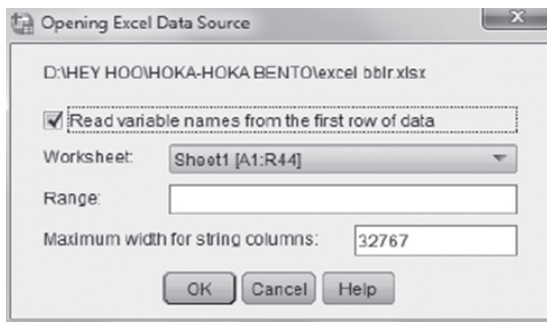
- Klik menu ***File*** → ***Open*** → ***Data***
- Pilih *File of Types*

### ● GAMBAR 6.15 Import File Excell



- Carilah ekstensi yang sesuai dengan *Microsoft Excell*
- Carilah direktori tempat file tersebut pada *look in*, lalu klik data yang mau dibuka “*Open*” lalu OK.

### ● GAMBAR 6.16 Window Konfirmasi



### 3. Mengganti Nilai Data

Langkah-langkah:

- Klik sel tempat data yang akan diganti
- Ketikkan data baru



● **GAMBAR 6.17** Data akan Diganti

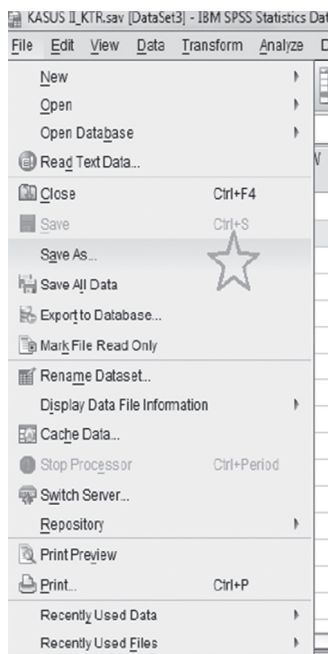
	ART1f_hari	ART1f_bulan	ART1f_tahun	ART1g
1	.00	.00	56,00	4
2	.00	.00	41,00	3
3	.00	.00	39,00	5
4	.00	.00	64,00	3
5	.00	.00	32,00	3

4. Menyimpan Data

Langkah-langkah:

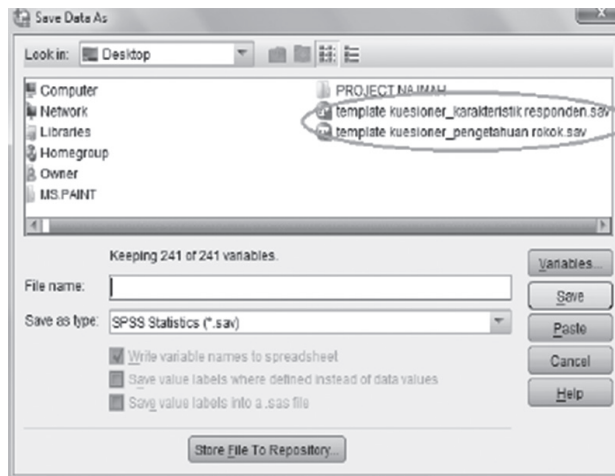
- Pilih Menu **File** → **Save**
- Ketikkan nama file pada *File Name*
- Klik **Save** jika ingin menyimpan dan klik **Cancel** jika ingin membatalkan

● **GAMBAR 6.18** Menyimpan File Data



- File data akan berekstensi **.sav** (nama file.sav) dan file *output* akan berekstensi.spo (nama file.spo).

● **GAMBAR 6.19** Menyimpan File Data



5. Menghapus Data

Langkah-langkah:

- menghapus isi sel dengan mengklik sel yang akan dihapus, lalu tekan tombol *delete*;
- menghapus sejumlah sel sekaligus dengan mengklik lalu tarik sehingga semua sel terblok, lalu tekan *delete*;
- menghapus isi sel satu kolom dengan mengklik heading kolom (nama variabel) yang akan dihapus, tekan *delete*;
- menghapus isi sel satu baris dengan mengklik baris (*nomor case*) yang akan dihapus, tekan *delete* atau dengan mengklik kanan pada mouse atau dengan mengklik menu *Edit*, pilih "*Clear*".

● **GAMBAR 6.20** Menghapus Data dengan Mengklik Kanan pada Mouse

	KAB	KEC
1	1	2
2	1	2
3		2
4		2
5		2
6		2
7		2
8		2

A context menu is overlaid on the table, showing options: Cut, Copy, Paste, Clear, and Insert Cases. The 'Clear' option is highlighted.

6. Menyalin (*copy*) data

Langkah-langkah:

- klik nilai data yang akan di-*copy*;
- klik menu *Edit*, pilih *Copy* atau Klik Ctrl C;
- arahkan pointer atau penunjuk sel ke lokasi tempat file tersebut akan di-*copy*;
- klik menu *Edit*, pilih *Paste* (Ctrl V).

● **GAMBAR 6.21** Copy Data

32,00	3
44,00	4
48,00	3
54,00	3
43,00	4
32,00	Cut
69,00	Copy
46,00	Paste
46,00	Clear
70,00	Grid_Font
50,00	Spelling...
37,00	

7. Memindahkan Nilai Sel

Langkah-langkah:

- klik data yang akan dipindah;
- klik edit, pilih *Cut* atau tekan tombol Ctrl + X;
- arahkan pointer ke lokasi baru tempat data akan dipindah;
- klik menu *Edit*, pilih *Paste* atau tekan tombol Ctrl + V.

● **GAMBAR 6.22** Tampilan Fungsi "Edit"

The image shows the 'Edit' menu in SPSS with several callout boxes explaining the functions of the menu items:

- Memotong/menghapus data (Ctrl X):** Points to the 'Cut' menu item.
- Meng-copy data (Ctrl C):** Points to the 'Copy' menu item.
- Menyisipkan variabel:** Points to the 'Insert Variable' menu item.
- Menyisipkan kolom dan sel:** Points to the 'Insert Cases' menu item.

The menu items and their keyboard shortcuts are as follows:

Menu Item	Keyboard Shortcut
Undo	Ctrl+Z
Redo	Ctrl+Y
Cut	Ctrl+X
Copy	Ctrl+C
Paste	Ctrl+V
Paste Variables...	
Clear	Delete
Insert Variable	
Insert Cases	
Find...	Ctrl+F
Find Next	F3
Replace...	Ctrl+H
Go to Case...	
Go to Variable...	
Go to Imputation...	
Options...	

## 8. Menyisipkan data

Langkah-langkah:

- pilih posisi baris/kolom yang akan disisipkan;
- klik menu Edit, pilih *Insert Case* untuk menyisipkan baris atau *Insert Variable* untuk menyisipkan Kolom (variabel);
- setelah penyisipan data baru dapat diinputkan;
- menyisipkan Kolom dengan meletakkan kursor pada kolom yang akan disisipi, klik Data, pilih *Insert Variable*;
- menyisipkan baris dengan meletakkan kursor pada baris yang akan disisipi, klik Data, pilih *Insert Case*.

## 9. Output

Window ini adalah berkaitan dengan hasil dari suatu analisis statistik. Misalkan *output* data deskripsi umur dan tingkat pendidikan ibu. Dari data tersebut hasil yang didapat seperti Gambar 6.23.

● **GAMBAR 6.23** Output Data

IBM SPSS Statistics Process

Descriptives

[DataSet3] C:\Users\Owner\Desktop\PROJECT NAJMA\BAB II\KASUS II\_KTR.sav

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Umur dalam Tahun	181	0	70	18,04	14,522	210,879
Status pendidikan tertinggi	148	0	7	3,60	1,644	2,708
Status Pekerjaan	154	0	9	2,80	2,519	6,345
Valid N (listwise)	149					

FREQUENCIES VARIABLES=ARTI1\_tahun ARTI1\_ARTI1  
/STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE  
/ORDER=ANALYSIS.

→ Frequencies

[DataSet3] C:\Users\Owner\Desktop\PROJECT NAJMA\BAB II\KASUS II\_KTR.sav

IBM SPSS Statistics Process

→ Frequencies

[DataSet3] C:\Users\Owner\Desktop\PROJECT NAJMA\BAB II\KASUS II\_KTR.sav

		Umur dalam Tahun	Status pendidikan tertinggi	Status Pekerjaan
N	Valid	195	195	185
	Missing	0	0	0
Mean		47,5487	3,23	5,27
Median		46,0000	3,00	6,00
Mode		46,00	3	6
Std. Deviation		12,46527	1,262	1,742
Variance		155,383	1,567	3,168
Minimum		23,00	0	1
Maximum		92,00	7	9

Frequency Table

Umur dalam Tahun



# BAB 7

## STATISTIK DESKRIPTIF PADA SPSS

### Kompetensi Dasar

Mampu menjelaskan analisis deskriptif statistik data kesehatan menggunakan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan jenis data dan skala.
- Mampu menjelaskan analisis univariat dengan SPSS.
- Mampu menjelaskan analisis deskriptif data kategorik dengan SPSS.
- Mampu menjelaskan analisis deskriptif data numerik dengan SPSS.
- Mampu menjelaskan analisis data numerik dengan kondisi missing data.

### Materi Pembelajaran

1. Jenis Data dan Skala
2. Analisis Univariat
3. Analisis Deskriptif Data Kategorik
4. Analisis Deskriptif Data Numerik
5. Analisis Data Numerik dengan Kondisi Ada Missing Data

## JENIS DATA DAN SKALA

Dalam analisis data, terlebih dahulu Anda harus membedakan atau mengklasifikasikan suatu kelompok data. Mari kita pahami definisi “data” dan variabel terlebih dahulu. Data adalah bentuk jamak (plural) dari “*datum*”. Definisi data adalah himpunan angka-angka yang merupakan nilai dari unit sampel kita sebagai hasil dari mengamati/mengukur. Sementara variabel adalah suatu sifat yang akan diukur atau diamati yang nilainya bervariasi antara satu objek ke objek lainnya. Variabel dalam suatu data pada umumnya dibedakan menjadi dua, kategori dan numerik. Pada variabel kategorik ada istilah nominal dan ordinal dan pada variabel numerik ada istilah rasio dan numerik, apakah perbedaannya, mari kita pelajari bersama.

Untuk mempelajarinya, coba pelajari klasifikasi data Tabel 7.1! Jika kita perhatikan, contoh yang termasuk variabel kategori dengan skala nominal yaitu pekerjaan, status kawin, dan jenis kelamin; sedangkan variabel kategori dengan skala ordinal adalah tingkat pendidikan. Semua data kategori baik nominal dan ordinal dilaporkan dalam bentuk frekuensi (jumlah) dan persentase dan juga bisa berupa gambar atau grafik. Apakah beda kedua skala nominal dan ordinal?

● **TABEL 7.1** Contoh Jenis Data dan Skala

VARIABEL	JENIS DATA
<b>Pekerjaan (n, %)</b> Tidak Kerja, TNI/POLRI, PNS/Pegawai, wiraswasta/Jasa/ Dagang, Petani, Buruh, Lainnya.	Kategori >>> Nominal
<b>Status kawin (n, %)</b> Belum Kawin, Kawin, Cerai Hidup, Cerai Mati	Kategori >>> Nominal
<b>Pendidikan (n, %)</b> Tidak Pernah Sekolah, Tidak Tamat SD/MI, Tamat SD/MI, Tamat SMP, Tamat SMA, Tamat D1/D2/D3, Tamat PT	Kategori >>> Ordinal
<b>Jenis kelamin (n, %)</b> Laki-laki, Perempuan	Kategorik >>> Nominal
<b>Umur Bapak (tahun)(rata-rata/SD)</b>	Numerik >>> Rasio
<b>Jumlah anggota keluarga (rata-rata/SD)</b>	Numerik >>> Rasio
<b>Jumlah balita (rata-rata/SD)</b>	Numerik >>> Rasio
<b>Suhu tubuh (°C) (rata-rata/SD)</b>	Numerik >>> Interval
<b>Usia berhenti merokok (tahun) (rata-rata/SD)</b>	Numerik >>> Rasio
<b>Membeli rokok perhari (jumlah rokok) (rata-rata/SD)</b>	Numerik >>> Rasio
<b>Tabungan perhari (rupiah) (rata-rata/SD)</b>	Numerik >>> Rasio
<b>Penghasilan perhari (rupiah) (rata-rata/SD)</b>	Numerik >>> Rasio
<b>Tingkat IQ Bapak (rata-rata/SD)</b>	Numerik >>> Interval

1. Skala Nominal

- Skala nominal adalah variabel yang bersifat kategorisasi dan tidak ada perjenjangan (*equal*).
- Variabel ini memiliki pengukuran paling lemah tingkatannya, terjadi apabila bilangan atau lambang-lambang lain digunakan untuk mengklasifikasikan objek pengamatan.

2. Skala Ordinal

- Skala variabel yang memiliki perjenjangan tetapi tidak sama/*equal*.
- Pengukuran ini tidak hanya membagi objek menjadi kelompok-kelompok yang tidak tumpang tindih, tetapi antara kelompok itu ada hubungan (*rangking*).
- Data skala ordinal mempunyai urutan kategori yang bermakna, tetapi tidak ada jarak yang terukur di antara kategori.

Coba sekarang perhatikan contoh variabel numerik dengan skala interval dan rasio dari Tabel 7.1. Perbedaan apa yang bisa Anda simpulkan contoh di atas? Yang termasuk contoh skala rasio adalah umur bapak, jumlah anggota keluarga, jumlah balita, usia berhenti merokok, jumlah rokok per hari, dan tabungan per hari serta jumlah penghasilan, sementara contoh skala interval adalah suhu tubuh dan tingkat IQ bapak. Semua variabel numerik pada tabel hasil penelitian akan melaporkan angka rata-rata dan standar deviasi atau standar *error*. Variabel numerik berkaitan dengan gambaran karakteristik satu set data dengan skala pengukuran numerik, dua parameter yang lazim. Jika data tidak normal, maka median dan angka range (min-maks) yang akan dilaporkan. Apa perbedaan skala interval dan rasio pada data numerik, perhatikan penjelasan di bawah ini.

3. Skala Interval

- Skala interval bersifat memiliki perjenjangan dan jenjang *equal*.
- Tidak memiliki nilai nol absolut.
- Membagi objek menjadi kelompok tertentu dan dapat diurutkan, juga dapat ditentukan jarak dari urutan kelompok tersebut.
- Misal: Suhu normal badan Andi biasanya 32°C. Ketika dia menderita demam, suhu tubuhnya menjadi 37°C. Berarti suhu Andi lebih panas 5°C daripada suhu normal.

● **TABEL 7.2** Skala Pengukuran Tabel

Sifat Skala	Nominal	Ordinal	Interval	Rasio
<b>a. Persamaan pengamatan (pengelompokan), klasifikasi pengamatan dapat dilakukan</b>	Ya	Ya	Ya	<b>Ya</b>
<b>b. Urutan tertentu, urutan pengamatan dapat dilakukan</b>	Tidak	Ya	Ya	<b>Ya</b>
<b>c. Jarak antara kelompok dapat ditentukan</b>	Tidak	Tidak	Ya	<b>Ya</b>
<b>d. Perbandingan antara kelompok</b>	<b>Tidak</b>	<b>Tidak</b>	<b>Tidak</b>	<b>Ya</b>

Nol derajat celsius bukan 0 absolut, artinya walaupun nilainya 0 bukan berarti suhu menjadi normal, tetapi tetap ada nilainya. Namun jika suhu tubuh dalam skala Kelvin ( $^{\circ}\text{K}$ ), termasuk dalam skala rasio karena memiliki 0 absolut/mutlak.

#### 4. Skala Rasio

- Skala rasio memiliki perjenjangan dan jenjang equal tetapi ada nilai nol absolut.
- Dengan skala rasio kita dapat mengelompokkan data, kelompok itu pun dapat diurutkan dan jarak antara urutan pun dapat ditentukan.
- Data dapat diperbandingkan (rasio) dan mempunyai titik “nol mutlak”.

### LATIHAN:



Coba Anda tuliskan contoh data lainnya berdasarkan skala nominal, ordinal, interval, dan rasio sebanyak-banyaknya, diskusikan dengan teman Anda!

#### SKALA PENGUKURAN

KATEGORIKAL/KUALITATIF/ DISKONTINU	NUMERIK/NON-KATEGORIKAL/ KUANTITATIF/ KONTINU
<b>Nominal</b>	<b>Rasio</b>
<b>Ordinal</b>	<b>Interval</b>

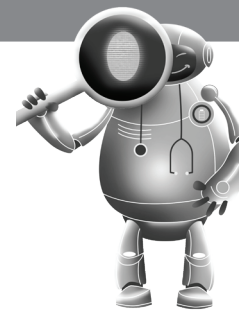



# ANALISIS UNIVARIAT

Pada tahap ini, kita akan mempelajari aplikasi analisis deskripsi untuk data numerik dan kategorik. Selain itu kita akan mempelajari bagaimana melaporkan data numerik dengan distribusi normal dan tidak normal dan menyeleksi data yang kita butuhkan pada proses analisis. Kita juga akan mempelajari bagaimana mengaplikasikan *syntax* seperti pengolahan data pada Stata pada aplikasi SPSS.

## STUDI KASUS 7.1

Kita ingin menghasilkan Tabel 1 karakteristik responden berdasarkan status intervensi untuk variabel numerik dan kategorik.



 Buka data: Kasus II\_KTR.sav

Sumber data: Najmah, Fenny Etrawati, Yeni, Feranita Utama. 2015. Studi Intervensi Klaster Kawasan Tanpa Rokok pada tingkat rumah tangga., 2015: Jurnal Ilmu Kesehatan Masyarakat Nasional: Universitas Indonesia, Vol 9 No 4 Mei 2015 (<http://jurnalkesmas.ui.ac.id/index.php/kesmas/article/view/752>.)

### PERTANYAAN DETEKTIF

- Aplikasikan SPSS dalam menghasilkan tabel karakteristik responden di bawah ini.

### JAWABAN DETEKTIF

- **TABEL 7.3** Gambaran Karakteristik Responden pada Desa yang Mendapatkan Intervensi dan Non-Intervensi Terpadu

Variabel		Status Intervensi	
		Intervensi (n = 95)	Non-intervensi (n = 100)
Pekerjaan	Tidak Kerja, n(%)	11 (11,6%)	10 (10%)
	TNI/POLRI, n(%)	1 (1%)	1 (1%)
	PNS/Pegawai, n(%)	1 (2%)	1 (1%)
	Wiraswasta/Jasa/Dagang, n(%)	19 (20%)	44 (44%)
	Petani, n(%)	58 (61,1%)	31 (31%)
	Buruh, n(%)	4 (4,2%)	9 (9%)
	Lainnya, n(%)	0 (0%)	4 (4%)

● **TABEL 7.3** Gambaran Karakteristik Responden pada Desa yang Mendapatkan Intervensi dan Non-Intervensi Terpadu (lanjutan)

Variabel		Status Intervensi	
		Intervensi (n = 95)	Non-intervensi (n = 100)
Status Kawin	Belum Kawin	0 (0%)	7 (7.5%)
	Kawin	93 (98.9%)	85 (91,4 %)
	Cerai Hidup	0 (0%)	1 (1.1 %)
	Cerai Mati	1 (1.1%)	0 (0%)
Pendidikan	Tidak Pernah Sekolah, n(%)	9 (9,5%)	3 (3%)
	Tidak Tamat SD/MI, n(%)	11 (11,6%)	17 (17%)
	Tamat SD/ MI, n(%)	53 (55,8%)	34 (34%)
	Tamat SMP, n(%)	11 (11,6%)	25 (25%)
	Tamat SMA, n(%)	11 (11,6%)	18 (18%)
	Tamat D1/D2/D3, n(%)	0 (0%)	1 (1%)
	Tamat PT	0 (0%)	2 (2%)
Umur Bapak (Median, Range)		46 (26-83)	44 (27-69)
Jumlah Anggota Keluarga (Median, Range)		4 (1-8)	4 (1-8)
Jumlah Balita (Median, Range)		0 (0-2)	1 (0-2)

Sumber: Najmah dkk, 2015, p. 377.

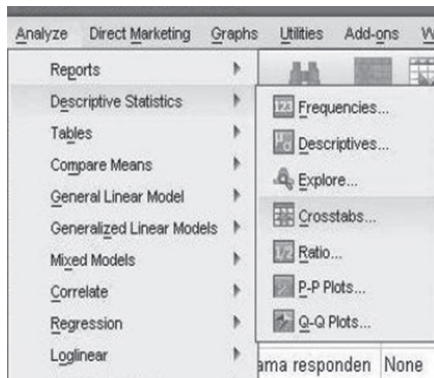
Hasil analisis karakteristik demografi responden pada Tabel 7.3 menunjukkan bahwa mayoritas responden pada kelompok intervensi bekerja sebagai petani (61%) sedangkan sebagian bekerja sebagai wiraswasta pada kelompok non-intervensi bekerja (44%). Tabel 7.3 menunjukkan bahwa sebanyak 55,8% responden pada kelompok intervensi tidak tamat SD/MI dan sebesar 34% pada kelompok non-intervensi. Rata-rata umur antara kelompok intervensi dan non intervensi 46 tahun dan 44 tahun. Sementara jumlah anggota keluarga berkisar 4 orang dan balita antara 0–2 balita per rumah pada kelompok intervensi dan non-intervensi

## ANALISIS DESKRIPTIF DATA KATEGORIK

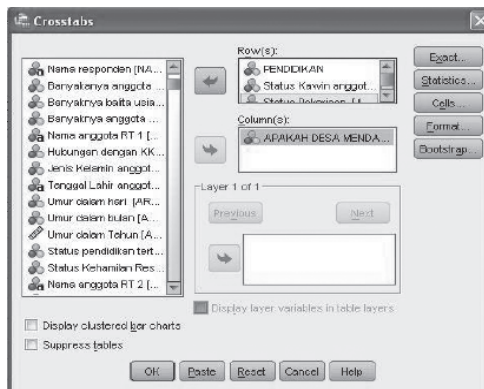
Untuk analisis deskripsi data kategorik, kita membutuhkan nilai observasi dan nilai persentase untuk Tabel 7.6 (Variabel Pekerjaan, Status Kawin, dan Tingkat Pendidikan).

1. **Langkah 1:** Untuk data kategorik, *analyze* → *descriptive statistics* → *crosstabs*
2. **Langkah 2:** Masukkan variabel Dependen (Intervensi) di **Column** dan variabel Independen (Pekerjaan, Pendidikan dan Status Kawin) di **ROW**.
3. **Langkah 3:** Klik *Cell Display*, untuk menampilkan persentase, kita bisa melakukan persentase total 100% per baris (*row*) atau per kolom ke bawah (*column*), kita klik keduanya.

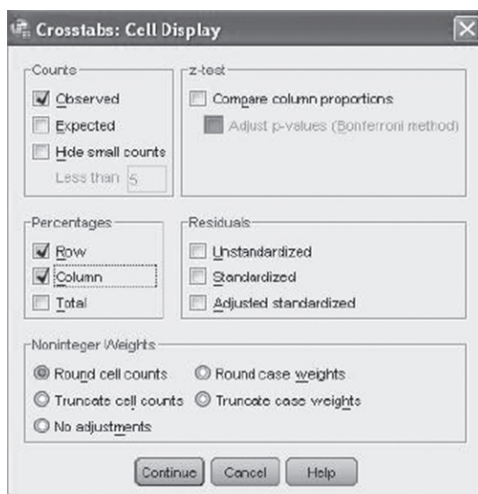
● **GAMBAR 7.1** Langkah 1 “Descriptive Statistics >> Crosstabs”



● **GAMBAR 7.2** Langkah 2 "ROW dan COLOUMN"



● **GAMBAR 7.3** Langkah 3



**OUTPUT SPSS:**

4. **Langkah 4:** Lihat *output* yang Anda hasilkan, dan silakan pindahkan ke Tabel 7.6, kita menggunakan persentase kolom ke bawah untuk total 100%. (Column)

● **TABEL 7.4** Output Cross Tabulation Pendidikan Responden dan Desa yang Mendapatkan Intervensi dan Non-intervensi

PENDIDIKAN \* APAKAH DESA MENDAPAT INTERVENSI CROSSTABULATION

			APAKAH DESA MENDAPAT INTERVENSI		TOTAL
			TIDAK	YA	
PENDIDIKAN	Tidak Pernah Sekolah	Count	3	9	12
		% within PENDIDIKAN	25,0%	75,0%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	3,0%	9,5%	6,2%
Tidak Tamat SD/MI	Tidak Tamat SD/MI	Count	17	11	28
		% within PENDIDIKAN	60,7%	39,3%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	17,0%	11,6%	14,4%
Tamat SD/MI	Tamat SD/MI	Count	34	53	87
		% within PENDIDIKAN	39,1%	60,9%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	34,0%	55,8%	44,6%
Tamat SMP	Tamat SMP	Count	25	11	36
		% within PENDIDIKAN	69,4%	30,6%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	25,0%	11,6%	18,5%
Tamat SMA	Tamat SMA	Count	18	11	29
		% within PENDIDIKAN	62,1%	37,9%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	18,0%	11,6%	14,9%
Tamat D1/D2/D3	Tamat D1/D2/D3	Count	1	0	1
		% within PENDIDIKAN	100,0%	0,0%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	1,0%	0,0%	0,5%
Tamat PT	Tamat PT	Count	2	0	2
		% within PENDIDIKAN	100,0%	0,0%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	2,0%	0,0%	1,0%
Total	Total	Count	100	95	195
		% within PENDIDIKAN	51,3%	48,7%	100,0%
		% within APAKAH DESA MENDAPAT INTERVENSI	100,0%	100,0%	100,0%

● **TABEL 7.5** Output Cross Tabulation Status Pekerjaan Responden dan Desa yang Mendapatkan Intervensi dan Non-intervensi

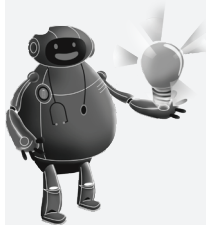
Status Pekerjaan \* Apakah Desa Mendapat Intervensi Crosstabulation

Status Pekerjaan		APAKAH DESA MENDAPAT INTERVENSI		TOTAL
		TIDAK	YA	
Tidak Kerja	Count	10	11	21
	% within Status Pekerjaan	47,6%	52,4%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	10,0%	11,6%	10,8%
TNI/ Polri	Count	1	1	2
	% within Status Pekerjaan	50,0%	50,0%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	1,0%	1,1%	1,0%
PNS/ Pegawai	Count	1	2	3
	% within Status Pekerjaan	33,3%	66,7%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	1,0%	2,1%	1,5%
Wiraswasta/ layanan jasa/ dagang	Count	44	19	63
	% within Status Pekerjaan	69,8%	30,2%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	44,0%	20,0%	32,3%
Petani	Count	31	58	89
	% within Status Pekerjaan	34,8%	65,2%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	31,0%	61,1%	45,6%
Buruh	Count	9	4	13
	% within Status Pekerjaan	69,2%	30,8%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	9,0%	4,2%	6,7%
Lainnya	Count	4	0	4
	% within Status Pekerjaan	100,0%	0,0%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	4,0%	0,0%	2,1%
Total	Count	100	95	195
	% within Status Pekerjaan	51,3%	48,7%	100,0%
	% within APAKAH DESA MENDAPAT INTERVENSI	100,0%	100,0%	100,0%

\*\* Silakan lakukan hal yang sama untuk variabel status kawin

● **TABEL 7.6** Gambaran Karakteristik Responden pada Desa yang Mendapatkan Intervensi dan Non-intervensi

	Status Intervensi	
	Intervensi (n = 95)	Non-intervensi (n = 100)
<b>Pekerjaan</b>	Tidak Kerja, n(%) TNI/POLRI, n(%) PNS/Pegawai, n(%) Wiraswasta/Jasa/Dagang, n(%) Petani, n(%) Buruh, n(%) Lainnya, n(%)	
<b>Status Kawin</b>	Belum Kawin Kawin Cerai Hidup Cerai Mati	
<b>Pendidikan</b>	Tidak Pernah Sekolah, n(%) Tidak Tamat SD/MI, n(%) Tamat SD/ MI, n(%) Tamat SMP, n(%) Tamat SMA, n(%) Tamat D1/D2/D3, n(%) Tamat PT	



#### CATATAN:

Masukkan data berdasarkan output SPSS Anda! Persentase bisa kita gunakan persentase *row* (100% pada kolom independen) atau persentase *column* (100% pada kolom dependen/outcome), dan hal yang perlu diperhatikan cara interpretasi data dan jenis studi desain penelitian.

Coba Anda interpretasikan data di atas menggunakan kalimat Anda sendiri (ingat, tidak semua data perlu diinterpretasikan, Anda bisa melihat nilai mayoritas atau minoritas, atau secara keseluruhan)!

# ANALISIS DESKRIPTIF DATA NUMERIK

Untuk analisis deskripsi pada data numerik, kita membutuhkan nilai *mean* (rata-rata) dan standar deviasi jika data numerik kita berdistribusi normal, dan kita membutuhkan nilai median (nilai tengah) dan nilai minimum-maksimum atau range jika data numerik kita berdistribusi tidak normal.

**Langkah 1:** Uji normalitas data numerik: umur bapak, jumlah anggota keluarga, dan jumlah balita

Hasil Output Uji Normalitas-Kolmogorov-Smirnov, perhatikan nilai p(sig) pada tabel di bawah ini!

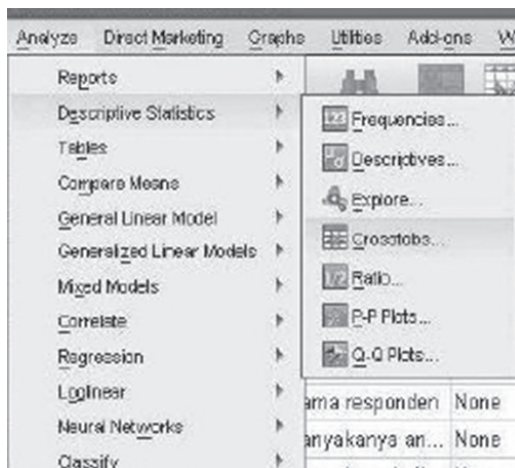
*Ho: variabel umur, jumlah anggota keluarga, jumlah balita berdistribusi normal*  
*Ha: variabel umur, jumlah anggota keluarga, jumlah balita tidak berdistribusi normal*

**Kesimpulan:** sig < 0.05, H0 ditolak, jadi distribusi tidak normal untuk ketiga data numerik (umur, p.0.005; jumlah anggota keluarga-p.<0.001, jumlah balita-p.<0.0001) yang kita punya. Jadi kita melaporkan median (nilai tengah) dan range (nilai min-maks).

Berikut adalah proses uji normalitas.

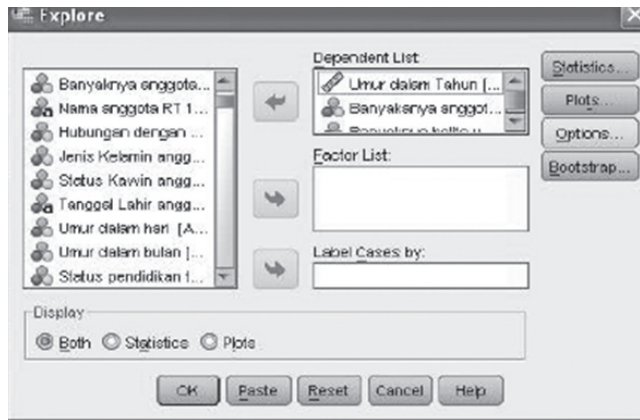
1. Klik **Analyze** → **Descriptive statistics** → **Explore**

● **GAMBAR 7.4** Langkah 1 Uji Normalitas



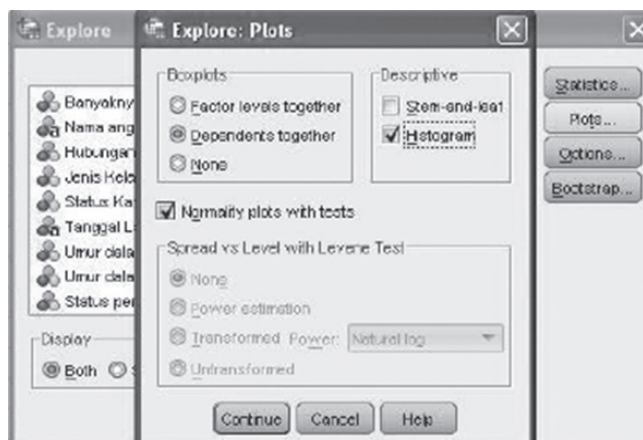
- Masukkan variabel numerik, umur, jumlah balita, jumlah anggota keluarga ke “**Dependent list**”.

● **GAMBAR 7.5** Langkah 2 Uji Normalitas



- Klik **PLOTS** pilih **Dependents together** dan **Normality plots with tests**

● **GAMBAR 7.6** Langkah 3 Uji Normalitas



**PERHATIKAN OUTPUT SPSS**

- **TABEL 7.7** Uji Normalitas-Kolmogorov-Smirnov Data Numerik (Umur Bapak, Jumlah Anggota Keluarga dan Jumlah Balita)

Hasil *output* yang dihasilkan dari langkah-langkah di atas adalah sebagai berikut!

**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Umur dalam Tahun	,092	143	,005	,963	143	,001
Banyaknya Anggota Rumah Tangga Selain Kepala Keluarga	,126	143	,000	,956	143	,000
Banyaknya Balita Usia 0–4 Tahun	,326	143	,000	,735	143	,000

<sup>a</sup> Lilliefors Significance Correction



**Langkah 2:** Melaporkan nilai mean dan SD untuk data yang berdistribusi normal, dan nilai median dan range untuk data yang berdistribusi tidak normal. Oleh karena variabel yang kita punya semua berdistribusi tidak normal, kita laporkan nilai MEDIAN dan RANGE. Perhatikan nilai di bawah ini dan masukkan ke dalam Tabel 7.9!

**Descriptives**

APAKAH DESA MENDAPAT INTERVENSI		Statistic	Std. Error	
Umur dalam Tahun	TIDAK	Mean	45,5882	1,33143
		<b>Median</b>	<b>44,0000</b>	
		Variance	120,544	
		Std. Deviation	10,97927	
		<b>Minimum</b>	<b>27,00</b>	
		<b>Maximum</b>	<b>69,00</b>	
		<b>Range</b>	<b>42,00</b>	
	Interquartile Range	18,50		
	YA	Mean	47,7333	1,44725
		5% Trimmed Mean	47,1519	
		<b>Median</b>	<b>46,0000</b>	
		Variance	157,090	
		Std. Deviation	12,53356	
		<b>Minimum</b>	<b>26,00</b>	
<b>Maximum</b>		<b>83,00</b>		
<b>Range</b>	<b>57,00</b>			
Interquartile Range	19,00			
Banyaknya anggota rumah tangga selain kepala keluarga	TIDAK	Mean	4,01	,209
		5% Trimmed Mean	4,00	
		<b>Median</b>	<b>4,00</b>	
		Variance	2,970	
		Std. Deviation	1,723	
		<b>Minimum</b>	<b>1</b>	
		<b>Maximum</b>	<b>8</b>	
	<b>Range</b>	<b>7</b>		
	Interquartile Range	2		
	YA	Mean	3,84	,195
		5% Trimmed Mean	3,81	
		<b>Median</b>	<b>4,00</b>	
		Variance	2,839	
		Std. Deviation	1,685	
<b>Minimum</b>		<b>1</b>		
<b>Maximum</b>		<b>8</b>		
<b>Range</b>	<b>7</b>			
Interquartile Range	2			
Banyaknya balita usia 0-4 tahun	TIDAK	Mean	,72	,083
		5% Trimmed Mean	,69	
		<b>Median</b>	<b>1,00</b>	
		Variance	,473	
		Std. Deviation	,688	
		<b>Minimum</b>	<b>0</b>	
		<b>Maximum</b>	<b>2</b>	
		<b>Range</b>	<b>2</b>	
		Interquartile Range	1	

APAKAH DESA MENDAPAT INTERVENSI		Statistic	Std. Error
YA	Mean	,41	,060
	5% Trimmed Mean	,39	
	<b>Median</b>	<b>,00</b>	
	Variance	,273	
	Std. Deviation	,522	
	<b>Minimum</b>	<b>0</b>	
	<b>Maximum</b>	<b>2</b>	
	<b>Range</b>	<b>2</b>	
	Interquartile Range	1	

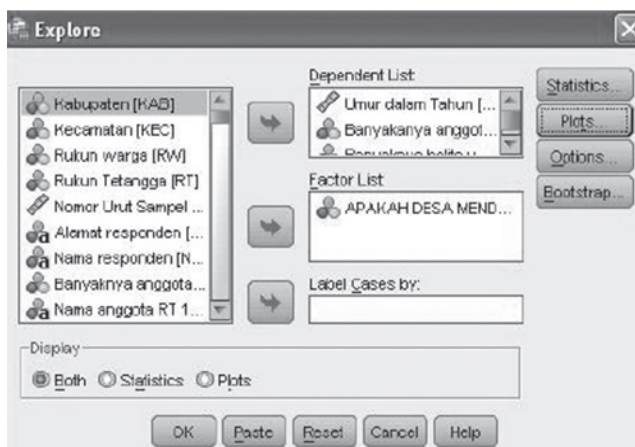
● **TABEL 7.8** Gambaran Karakteristik Responden pada Desa yang Mendapat Intervensi dan Non-intervensi

	Status Intervensi	
	Intervensi	Non-intervensi
Umur Bapak (Median, Range)		
Jumlah Anggota Keluarga (Median, Range)		
Jumlah Balita (Median, Range)		

#### Langkah-langkah:

1. *Descriptive statistics—Explore,*
2. Masukkan variabel umur, jumlah balita, dan jumlah anggota keluarga ke kolom “*Dependent list*” dan masukkan intervensi ke “*Factor list*” karena kita menampilkan tabel menjadi kelompok intervensi dan tidak intervensi,
3. Klik OK (tidak perlu mengklik plots, dan lain-lain).

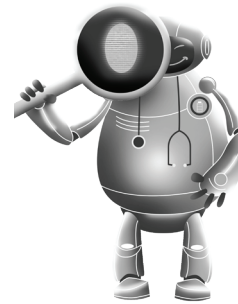
● **GAMBAR 7.7** Nilai Median, Minimum, dan Maksimum Variabel Numerik Dibagi Berdasarkan Status Intervensi




# ANALISIS DATA NUMERIK DENGAN KONDISI ADA MISSING DATA MENGGUNAKAN SYNTAX

## STUDI KASUS 7.2

Bagaimana kita bisa mendapatkan hasil pada Tabel 7.10 perilaku merokok pada desa yang mendapat intervensi dan non-intervensi berdasarkan status intervensi untuk variabel numerik dan ada banyak variabel yang hilang (*missing*) langkah apa yang harus kita lakukan?



 Buka data: Kasus II\_KTR.sav

**Notes: kode missing: 88**

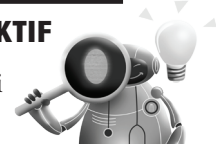
● **TABEL 7.9** Gambaran Perilaku Merokok pada Desa yang Mendapat Intervensi dan Non-intervensi

Variabel	Status Intervensi	
	Intervensi (n=95)	Non-intervensi (n=100)
Umur pertama kali merokok (Median, Range)	15.50( 5-53)	15 (6-35)
Missing	1	12
Umur pertama kali merokok tiap hari (Median, Range)	19 (7-60)	18(6-40)
Missing	6	20
Rata-rata batang rokok per hari (Median, Range)	12 (1-48)	12 (1-80)
Missing	0	0
Usia berhenti merokok (Mean, SD)	50 (12.2)	48(11.24)
Missing	42	60

Sumber: Najmah, dkk, 2015.

Berdasarkan Tabel 7.9 diketahui karakteristik perilaku merokok pada kelompok intervensi dan non-intervensi hampir sama. Rata-rata umur pertama kali merokok adalah 15 pada kedua kelompok dan rata-rata merokok tiap hari pada saat responden umur 19 tahun pada kelompok intervensi dan umur 18 tahun pada kelompok non-intervensi. Rata-rata jumlah rokok tiap hari mencapai 12 batang sedangkan usia berhenti merokok pada usia 50 pada kedua kelompok (lihat Tabel 7.9).

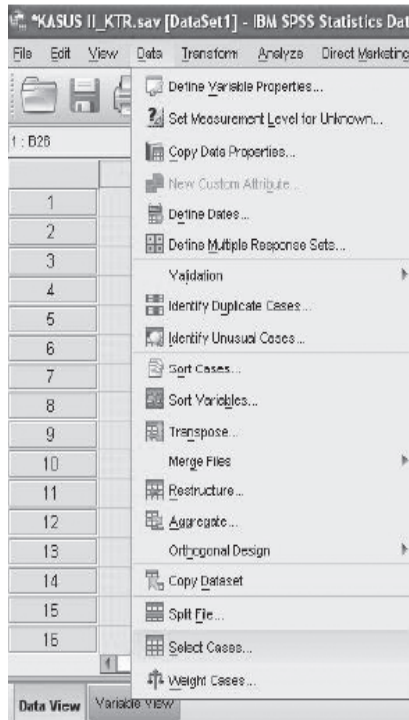
### JAWABAN DETEKTIF



Langkah-langkah menyeleksi kasus dengan “*Select Case*” dengan aplikasi **Drop Down Menu**.

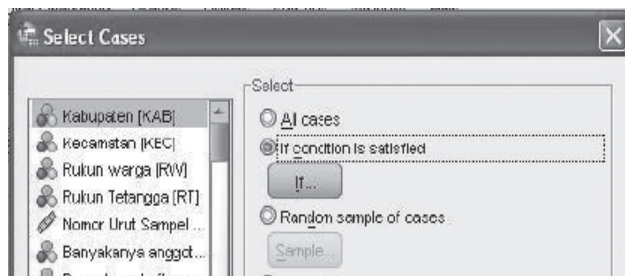
1. Pilih **Data** pada menu toolbar, pilih **Select Cases**.

#### ● GAMBAR 7.8 Langkah 1 Seleksi Kasus yang Diperlukan



2. Klik **if condition is satisfied**, pilih lalu klik button di bawahnya **if** (*filter out unselected cases*).

#### ● GAMBAR 7.9 Langkah 2 Seleksi Kasus



3. Kemudian pilih variabel B26 (Umur pertama kali merokok tiap hari) lalu masukkan variabel, kemudian bukan 88 lalu ( $B26 \neq 88$ ) klik **OK**.

● **GAMBAR 7.10** Langkah 3 Seleksi Kasus



4. Pada lembar data akan terlihat bahwa variabel yang telah dipilih sudah tereliminasi otomatis sesuai perintah yang dimasukkan.

***APLIKASI SYNTAX PADA SPSS PADA KASUS 7.2***

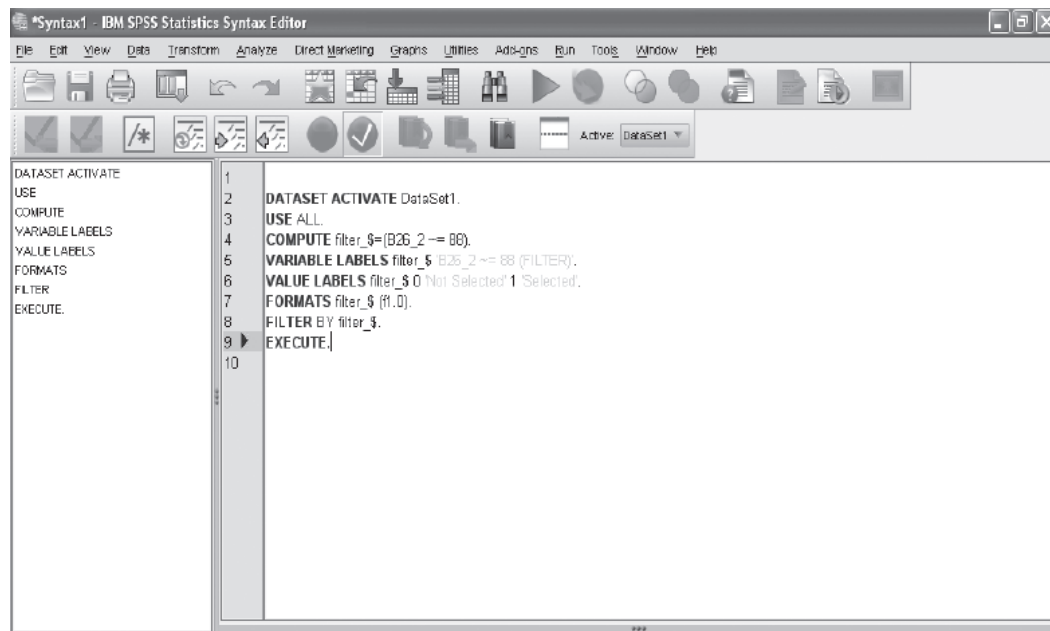
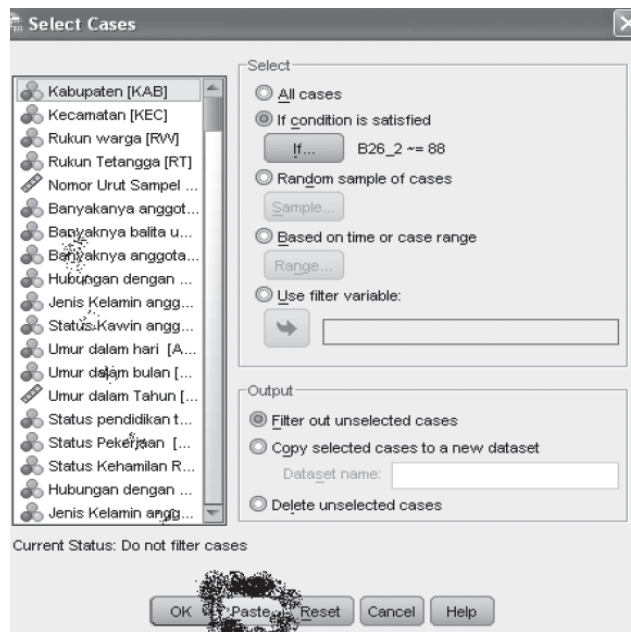
Jika Anda menggunakan aplikasi SPSS 20, maka pada *output* akan dihasilkan sekaligus *syntax* selain hasil olah data. Kita juga bisa meng-copy paste *syntax* proses kita olah data dengan menggunakan ‘drop down menu’ seperti langkah-langkah di atas, namun sebelum kita mengklik ‘OK’ pada proses akhir, kita klik ‘Paste’ terlebih dahulu (*button* sederetan dengan OK, lihat Gambar 7.11), maka, kode *syntax* akan otomatis keluar, dan akan tersimpan di file khusus SYNTAX.

Jadi, untuk variabel lainnya dengan proses yang sama, saya tinggal mengganti kode variabel pada *syntax* **B26 (umur merokok setiap hari)**, dengan kode variabel **umur pertama kali merokok (B24)**, **rata-rata batang rokok per hari (B27)**, dan **umur berhenti berokok (B33)**, lihat *syntax* di bawah ini, dan kita tinggal klik tombol **RUN (button hijau)**, untuk menghasilkan *output* SPSS, dengan menyorot kode *syntax* untuk variabel selanjutnya setelah diganti dengan kode variabel yang kita inginkan (📁 **Buka File: Syntax Perilaku Merokok.SPS**)

Kita akan mencoba mengaplikasikannya pada variabel lainnya, seperti umur pertama kali merokok tiap hari, rata-rata batang rokok perhari dan usia berhenti merokok dengan menggunakan *syntax*.

- 1) Umur pertama kali merokok tiap hari (Median, Range) `_B26Missing`

● **GAMBAR 7.11** Aplikasi Syntax pada SPSS



Select case dan explore : Umur Pertama Kali Merokok Tiap Hari, B26 (missing = 88 dikeluarkan)

\*\*\*\*\*

USE ALL.

COMPUTE filter\_\$(B26\_2 ~= 88).

VARIABLE LABEL filter\_\$ 'B26\_2 ~= 88 (FILTER)'.

```
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
```

```
EXAMINE
  VARIABLES=B26 BY INTERVENSI
  /PLOT BOXPLOT STEMLEAF NPLOT
  /COMPARE GROUP
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

```
FILTER OFF.
USE ALL.
EXECUTE .
```

## 2) Umur pertama kali merokok\_\_\_\_B24 Missing

Select case dan explore : Umur Pertama Kali Merokok, B24 (missing = 88 dikeluarkan)

```
*****
```

```
USE ALL.
COMPUTE filter_$=(B24 ~= 88).
VARIABLE LABEL filter_$ 'B24 ~= 88 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
```

```
EXAMINE
  VARIABLES=B24 BY INTERVENSI
  /PLOT BOXPLOT STEMLEAF NPLOT
  /COMPARE GROUP
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

```
FILTER OFF.
USE ALL.
```

```
EXECUTE .
```

### 3) Rata-rata batang rokok per hari (Median, Range) \_\_\_\_ B27 Missing

```
*****
```

```
USE ALL.
```

```
COMPUTE filter_$=(B27 ~= 88).
```

```
VARIABLE LABEL filter_$ 'B27 ~= 88 (FILTER)'.

```

```
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.

```

```
FORMAT filter_$ (f1.0).

```

```
FILTER BY filter_$.
```

```
EXECUTE .
```

```
EXAMINE
```

```
VARIABLES=B27 BY INTERVENSI
/PLOT BOXPLOT STEMLEAF NPLOT
/COMPARE GROUP
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

```
FILTER OFF.
```

```
USE ALL.
```

```
EXECUTE
```

### 4) Usia berhenti merokok(Mean, SD)\_\_ B33 Missing

Select case dan explore : Usia berhenti merokok , B33 (missing = 88 dikeluarkan)

```
*****
```

```
EXECUTE .
```

```
USE ALL.
```

```
COMPUTE filter_$=(B33 ~= 88 &B33 ~= 0).
```

```
VARIABLE LABEL filter_$ 'B33 ~= 88 &B33 ~= 0 (FILTER)'.

```

```
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.

```

```
FORMAT filter_$ (f1.0).

```

```
FILTER BY filter_$.
```

```
EXECUTE .
```

```
EXAMINE
```

```
VARIABLES=B33 BY INTERVENSI
/PLOT BOXPLOT STEMLEAF NPLOT
/COMPARE GROUP
```



```
/STATISTICS DESCRIPTIVES  
/CINTERVAL 95  
/MISSING LISTWISE  
/NOTOTAL.
```

```
FILTER OFF.  
USE ALL.  
EXECUTE .
```





## BAB 8

# UJI VALIDITAS & RELIABILITAS INSTRUMEN

### **Kompetensi Dasar**

Mampu menjelaskan analisis deskriptif statistik kesehatan menggunakan SPSS.

### **Indikator Keberhasilan**

Mampu menjelaskan uji reliabilitas dan validitas dalam SPSS.

### **Materi Pembelajaran**

Uji Reliabilitas dan Validitas

Salah satu bagian terpenting dalam penelitian adalah *tools* atau alat untuk mengukur penelitian kita. Dengan kata lain kita perlu *tools* untuk mengukur dan menilai hipotesis penelitian kita, salah satunya yaitu Kuesioner. Kumpulan pertanyaan pada kuesioner seperti pengetahuan terhadap suatu kondisi kesehatan, sikap, persepsi, harus mengukur apa yang kita ukur.

Sebagai peneliti, jika telah ada kuesioner baku dari penelitian sebelumnya, misalnya dari survei demografi dan kesehatan Indonesia (SDKI) ataupun riset kesehatan dasar nasional (Riskesdas) ataupun penelitian lain yang kuesionernya telah diuji validitas dan reliabilitasnya, kita tidak perlu melakukan uji coba kuesioner kepada minimal 30 responden. Namun tidak semua topik penelitian memiliki kuesioner baku ataupun yang telah diuji coba sebelumnya, sehingga peneliti perlu melakukan uji coba kuesioner penelitian yang akan dilakukan.

## TAHAP PERTAMA: UJI VALIDITAS

Uji Validitas digunakan untuk mengetahui sejauh mana ketepatan dan kecermatan suatu alat ukur dalam mengukur data. Misal kita ingin mengukur pengetahuan tentang rokok dan kawasan tanpa rokok di Universitas Sriwijaya. Untuk mengukur validitas pernyataan yang berkaitan dengan pengetahuan tersebut, dilakukan dengan cara melakukan korelasi antar-skor masing-masing pernyataan terhadap skor total.

Suatu pernyataan dikatakan valid bila skor pernyataan tersebut berkorelasi secara signifikan dengan skor totalnya. Keputusan uji, bila  $r$  hitung masing-masing pernyataan (dilihat pada *output data*) lebih besar dari  $r$  tabel maka  $H_0$  ditolak yang berarti valid dan jika  $r$  hitung lebih kecil dari  $r$  tabel maka  $H_0$  diterima yang berarti pernyataan tidak valid.


## TAHAP KEDUA: UJI RELIABILITAS

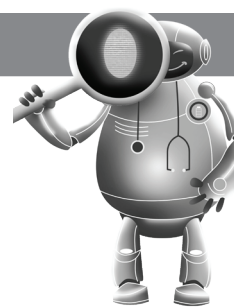
Uji reliabilitas adalah suatu konsistensi suatu hasil pengukuran. Dalam penelitian ini reliabilitas kuesioner diukur dengan cara satu waktu atau *one shot*. Di sini pengukurannya hanya sekali dan hasilnya dibandingkan dengan pernyataan lain. Setelah uji validitas dan dipastikan semua pertanyaan atau pernyataan yang tidak valid dikeluarkan dari uji coba, lalu uji reliabilitas dinilai. Pernyataan-pernyataan yang sudah valid kemudian baru secara bersama diukur reliabilitasnya. Untuk mengetahui reliabilitas pengetahuan tentang rokok dan kawasan tanpa rokok, maka kita membandingkan nilai  $r$  tabel dengan nilai  $r$  hasil (nilai ALPHA pada *output data*). Ketentuannya bila  $r$  Alpha lebih besar daripada  $r$  tabel maka pertanyaan tersebut *reliable* dan sebaliknya.

### Sumber:

Najmah. 2011. *Manajemen dan Analisis Data Kesehatan, Kombinasi Teori dan Aplikasi SPSS*. Nuha Medika: Yogyakarta.

## STUDI KASUS 8.2

 Buka Data: uji\_reliabilitas\_validitas.sav



Peneliti (Najmah, dkk., 2015) ingin mengetahui apakah pernyataan yang dibuat mengenai pengetahuan terkait rokok (merokok, bahaya rokok dan penyakit akibat rokok) benar-benar bisa digunakan untuk mengukur pengetahuan terkait rokok. Berikut pernyataan-pernyataan yang dibuat oleh tim peneliti. Uji coba kuesioner ini dilakukan pada 30 responden.

● **TABEL 8.1** Kuesioner Pengetahuan Terkait Rokok

IV. PENGETAHUAN TERKAIT ROKOK					
PETUNJUK: Isilah pernyataan berikut ini berdasarkan apa yang Bapak ketahui, pilihlah salah satu. Jawaban 1=Ya, 2=Tidak, atau 3=Tidak tahu					
PENGETAHUAN TENTANG ROKOK (Merokok, Bahaya Rokok, dan Penyakit Akibat Rokok)					
A01	Merokok merupakan hak asasi setiap manusia	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A02	Tidak merokok berarti menghargai kesehatan rekan kerja Anda	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A03	Merokok dapat merugikan kesehatan	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A04	Laki-laki/perempuan yang merokok terlihat lebih menarik	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A05	Merokok dapat menurunkan berat badan	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A06	Aman jika kita merokok selama 1 atau 2 tahun, asal kita berhenti merokok setelahnya	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A07	Seseorang yang tidak merokok namun menghirup asap rokok dari perokok yang ada di dekatnya adalah perokok pasif	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A08	Perokok aktif lebih berbahaya dibandingkan perokok pasif	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A09	Merokok di dalam ruangan tertutup dan ber-AC tidak berbahaya bagi diri sendiri dan orang lain	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A10	Bayi, balita dan perempuan tidak terganggu kesehatannya jika mengisap asap rokok	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A11	Kandungan zat kimia dalam rokok tidak berbahaya bagi kesehatan	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A12	Tar dan nikotin merupakan kandungan yang terdapat dalam rokok	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A13	Nikotin tidak menyebabkan ketagihan/kecanduan	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A14	Nikotin menjadi salah satu faktor timbulnya penyakit pembuluh darah seperti stroke	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A15	Merokok dapat mengganggu organ paru-paru, ginjal, pankreas, kandung kemih dan leher rahim	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A16	Kanker paru tidak disebabkan oleh perilaku merokok	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A17	Asap rokok mengandung ribuan macam zat kimia berbahaya bagi kesehatan	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A18	Rokok mengandung gas karbon monoksida (CO)	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>
A19	Rokok mengandung zat yang dapat mengikat oksigen dalam tubuh Anda	1. Ya	2. Tidak	3. Tidak Tahu	<input type="checkbox"/>

#### IV. PENGETAHUAN TERKAIT ROKOK

**PETUNJUK:** Isilah pernyataan berikut ini berdasarkan apa yang Bapak ketahui, pilihlah salah satu. Jawaban 1=Ya, 2=Tidak, atau 3=Tidak tahu

PENGETAHUAN TENTANG ROKOK (Merokok, Bahaya Rokok, dan Penyakit Akibat Rokok)			
A20	Rokok salah satu faktor penyebab Impotensi pada pria	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A21	Merokok bukan merupakan faktor risiko serangan stroke	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A22	Merokok dapat menyebabkan ISPA (Infeksi Saluran Pernapasan Atas)	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A23	Pemerintah mengatur Kawasan Tanpa Rokok melalui Peraturan Bersama Menteri Kesehatan dan Menteri Dalam Negeri No. 188/MENKES/PB/1/2011 No. 7 Tahun 2011	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A24	Kawasan tanpa rokok (KTR) adalah area yang bersih dari kegiatan produksi, penjualan, iklan, promosi ataupun penggunaan rokok	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A25	Pemerintah memberikan sanksi pada pelanggar peraturan Kawasan Tanpa Rokok	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A26	Perda Kawasan Tanpa Rokok (KTR) melarang warga untuk merokok pada tempat-tempat tertentu	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A27	Perda Kawasan Tanpa Rokok (KTR) hanya mengatur etika dalam merokok di tempat-tempat tertentu	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A28	Tempat belajar mengajar termasuk kawasan bebas rokok	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A29	Penerapan kawasan bebas rokok dapat menanggulangi masalah rokok	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A30	Mahasiswa, dosen, dan karyawan berhak menikmati lingkungan yang bebas asap rokok di tempat proses belajar mengajar	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A31	Pemberlakuan Kawasan Tanpa Rokok (KTR) melindungi generasi muda dari penyalahgunaan narkotika, psikotropika, dan zat adiktif	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>
A32	Menetapkan Kawasan Tanpa Rokok salah satu cara mengurangi pertambahan perokok anak-anak dan muda di masyarakat	1. Ya 2. Tidak 3. Tidak Tahu	<input type="checkbox"/>

Sumber: Najmah, Fenny Etrawati, Feranita Utama, Yeni, 2015 Intervensi terpadu pengurangan dampak buruk (harm reduction) asap rokok pada ruangan tertutup/ber-ac di lingkungan, Laporan Hibah Kompetitif, Lemlit Unsri, Indralaya.

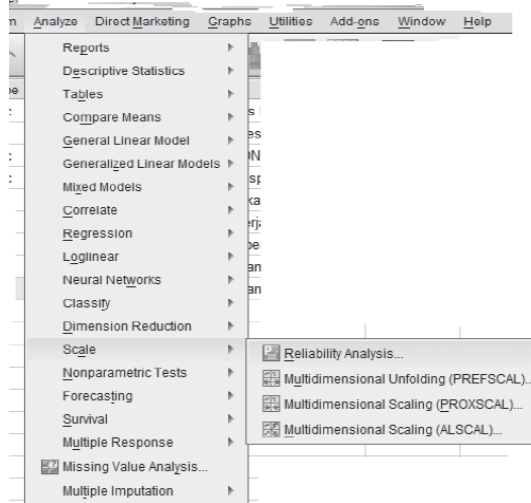
## LANGKAH-LANGKAH UJI VALIDITAS DAN RELIABILITAS

1. Pilih menu *Analyze* → *Scale* → *Reliabilityanalysis*
2. Masukkan variabel-variabel yang akan diuji ke dalam kotak items
3. Klik kotak **statistics** pada kotak '*Descriptive for*' pilih '*item, scale, scale if item deleted*'.
4. Pada Anova tabel pilih "*none*"
5. *Continue* → **OK**

Langkah-langkah uji validitas dan reliabilitas pada variabel pengetahuan:

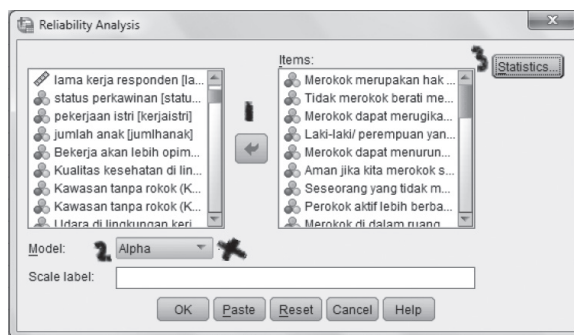
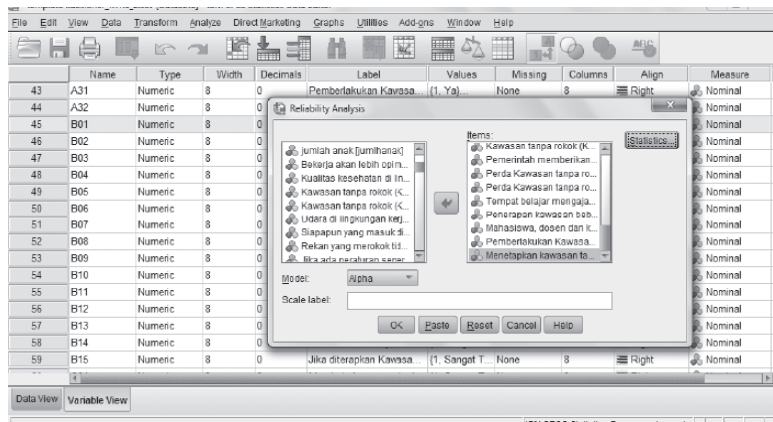
1. Klik *Analyze* → *Scale* → *Reliability Analysis*

● **GAMBAR 8.1** Proses Uji Validitas dan Reliabilitas



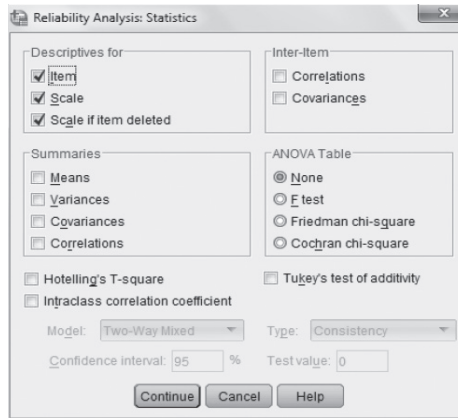
2. Masukkan semua variabel “Pengetahuan” yang akan diuji ke dalam kotak *items* (32 *items*).
3. Pada model biarkan pada *Alpha*.

● **GAMBAR 8.2** Kotak Dialog Items



- Klik options *Statistics* pada bagian *Descriptives*, klik: *Item, Scale, Scale if Item Deleted*.

● **GAMBAR 8.3** Kotak Dialog Option Statistic



## ANALISIS

Terdapat dua bagian dari hasil analisis reliabilitas dan validitas, yaitu:

- Bagian pertama menunjukkan hasil statistik deskriptif masing-masing variabel dalam bentuk mean, standar deviasi, varians, jumlah variabel.
- Bagian kedua memperlihatkan hasil dari proses validitas dan reliabilitas. Kaidah yang berlaku adalah dengan menguji validitas terlebih dahulu baru dilanjutkan uji reliabilitas (lihat nilai Alpha).

### ANALISIS 1: UJI VALIDITAS

**INGAT!** Bila  $r$  hasil  $>$   $r$  tabel, maka pertanyaan tersebut valid

- Nilai  $r$  tabel dilihat dengan tabel  $r$  menggunakan  $df = n - 2 = 30 - 2 = 28$ , pada tingkat kemaknaan 5% didapat angka  $r$  tabel = 0,351.
- Nilai  $r$  hitung dapat dilihat pada kolom "*corrected item-total correlation*".
- Keputusan: masing-masing pertanyaan variabel dibandingkan nilai  $r$  hasil dengan nilai tabel.
- Kesimpulan:  
Semua pertanyaan dinyatakan valid karena semua  $r$  hasil lebih besar dari  $r$  tabel, jika ada salah satu pertanyaan, nilai  $r$  hasil lebih kecil dari  $r$  tabel, maka lakukan uji selanjutnya dengan mengeluarkan pertanyaan tersebut.



## ANALISIS 2: UJI RELIABILITAS

**INGAT!** Bila  $r_{\alpha} > r_{\text{tabel}}$ , maka pertanyaan tersebut reliabel

Dari uji di atas ternyata nilai  $r_{\alpha}$  (0,899) > dibandingkan nilai  $r_{\text{tabel}}$ , maka pertanyaan di atas reliabel.

1. Output uji validitas dan reliabilitas tentang pengetahuan rokok:

● **TABEL 8.2** Output Uji Validitas dan Reliabilitas tentang Pengetahuan Rokok

Reliability Statistics				
Cronbach's Alpha	N of Items			
,899	32			

Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
A01	49,50	123,638	,159	,900
A02	49,63	119,413	,614	,894
A03	49,53	117,913	,609	,894
A04	48,90	120,438	,377	,897
A05	49,03	118,447	,430	,896
A06	48,80	115,890	,521	,895
A07	49,57	121,495	,335	,898
A08	49,10	120,507	,372	,897
A09	48,97	118,240	,502	,895
A10	49,03	123,413	,198	,900
A11	49,17	124,626	,100	,901
A12	49,67	121,885	,397	,897
A13	49,10	117,955	,560	,894
A14	49,00	115,172	,544	,894
A15	49,37	116,516	,506	,895
A16	48,83	117,799	,452	,896
A17	49,50	120,534	,302	,899
A18	49,03	117,206	,404	,897
A19	49,07	115,099	,532	,894
A20	49,17	113,523	,799	,890
A21	48,60	119,214	,403	,897

**Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
A22	49,20	115,614	,509	,895
A23	49,17	120,213	,238	,901
A24	49,30	116,838	,487	,895
A25	49,37	118,240	,406	,897
A26	49,60	116,524	,683	,892
A27	49,60	117,283	,625	,893
A28	49,60	124,386	,134	,900
A29	49,23	118,530	,487	,895
A30	49,63	121,482	,415	,897
A31	49,30	116,562	,574	,894
A32	49,27	115,995	,613	,893

**Reliability Coefficients**

N of Cases = 30.0

N of Items = 32

Alpha = 0,899

**2. Hasil Uji Validitas**

Setelah item yang nilainya tidak valid ( $r$  hasil <  $r$  tabel) dikeluarkan satu persatu yaitu item A01, A07, A10, A11, A17, A23, A28 maka, kita lakukan proses uji reliabilitas dan validitas sama seperti di atas, didapatkan hasil uji validitas yang valid, yaitu:

● **TABEL 8.3** Output Uji Validitas dan Reliabilitas tentang Pengetahuan Rokok

**Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Cronbach's Alpha if Item Deleted
A02	40,60	100,110	,636	,903
A03	40,50	98,879	,614	,903
A04	39,87	101,499	,357	,907
A05	40,00	99,655	,413	,906
A06	39,77	96,737	,541	,903
A07	40,53	101,913	,362	,907
A08	40,07	101,306	,372	,906
A09	39,93	99,099	,512	,904
A12	40,63	102,447	,412	,906

**Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
A13	40,07	99,513	,516	,904
A14	39,97	96,378	,546	,903
A15	40,33	96,989	,548	,903
A16	39,80	98,855	,449	,905
A18	40,00	98,276	,402	,907
A19	40,03	96,378	,529	,904
A20	40,13	94,671	,819	,898
A21	39,57	99,909	,417	,906
A22	40,17	96,695	,516	,904
A24	40,27	98,271	,466	,905
A25	40,33	99,540	,385	,907
A26	40,57	97,633	,686	,901
A27	40,57	98,461	,616	,902
A29	40,20	99,752	,468	,905
A30	40,60	101,903	,447	,905
A31	40,27	97,306	,602	,902
A32	40,23	97,633	,580	,903

Dari hasil di atas didapatkan hasil bahwa nilai  $r$  hasil  $>$   $r$  tabel (0,351) maka pertanyaan di atas valid atau tepat untuk mengukur tingkat pengetahuan tentang rokok dan KTR.

**c. Hasil Uji Reliabilitas**

Perhatikan hasil reliabilitas pada uji tadi!

**Reliability Statistics**

Cronbach's Alpha	N of Items
,907	25

Dari hasil uji reliabilitas didapatkan bahwa  $r$  alpha (0,907)  $>$   $r$  tabel (0,351), maka semua pertanyaan tentang rokok dan kawasan tanpa rokok tersebut reliabel atau konsisten satu sama lainnya dalam mengukur tingkat pengetahuan rokok dan KTR.





# BAB 9

## ANALISIS DATA SURVEI (DATA SEKUNDER) PADA SPSS

### Kompetensi Dasar

Mampu menjelaskan aplikasi analisis data sekunder menggunakan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan aplikasi analisis data sekunder (SDKI 2012).
- Mampu menjelaskan proses pengolahan data SDKI 2012 menggunakan SPSS.

### Materi Pembelajaran

1. Aplikasi Analisis Data Sekunder (SDKI 2012).
2. Proses Pengolahan Data SDKI 2012 menggunakan SPSS.
3. Uji Regresi Linear Complex Sample (Complex Samples General Linear Model).

## APLIKASI ANALISIS DATA SEKUNDER (SDKI 2012)

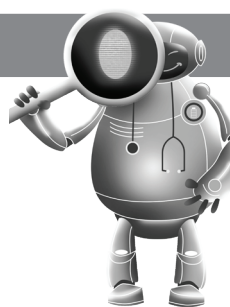
Analisis Data Sekunder (ADS) atau disebut juga *secondary data analysis* atau *existing statistic* memiliki karakter yang hampir sama dengan analisis isi (*content analysis*), peneliti memanfaatkan data berupa simbol-simbol dari media atau teks tertentu dan peneliti sendirilah yang mengolah data sedemikian rupa sehingga data tersebut dapat memiliki arti, sedangkan dalam ADS, peneliti cukup memanfaatkan data yang sudah matang yang dapat diperoleh dari instansi atau lembaga tertentu. Keuntungan menggunakan ADS, terutama jika peneliti menggunakan ADS dari survei besar dan terpercaya: (1) hemat waktu, tenaga, dan biaya penelitian karena tidak perlu mengumpulkan data primer di lapangan; (2) validitas dan reliabilitas data lebih terjamin karena survei yang dilakukan pada sampel yang besar dan telah dilakukan uji coba kuesioner penelitian; dan (3) bisa membandingkan tren atau data sekunder jika dilakukan pengumpulan data dari pihak lain secara reguler. Namun kelemahannya, variabel yang dibutuhkan pada penelitian kita akan terbatas dengan data yang telah ada pada data sekunder sebelumnya, dan tidak semua variabel-variabel yang dibutuhkan untuk topik penelitian kita dapat kumpulkan dari data sekunder. Peneliti dalam hal ini “tinggal menggunakan data yang telah ada”. Sumber data dalam ADS dapat diperoleh dari Riskesdas (Riset Kesehatan Dasar), SDKI (Survei Demografi dan Kesehatan Indonesia), Survei Terpadu Biologis dan Perilaku (STBP), BPS (Badan Pusat Statistik), IPM (Indeks Pembangunan Manusia), dan lain-lain. Pada bab ini akan dibahas pengolahan data sekunder SDKI 2012.

### STUDI KASUS 9.1

Peneliti membutuhkan data bobot, umur, dan pendidikan wanita usia subur data sekunder SDKI 2012.



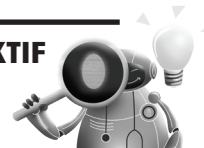
Buka data: IDI61FL.SAV -- (Data SDKI 2012 untuk Wanita usia subur)



### JAWABAN DETEKTIF

Berikut adalah langkah-langkah pemilihan variabel yang kita butuhkan.

1. Tentukan variabel apa saja yang akan diambil dalam penelitian atau tentukan definisi operasional dan kode pada masing-masing variabel. Misal: Tingkat Pendidikan WUS (V012) dan Umur WUS (V106)



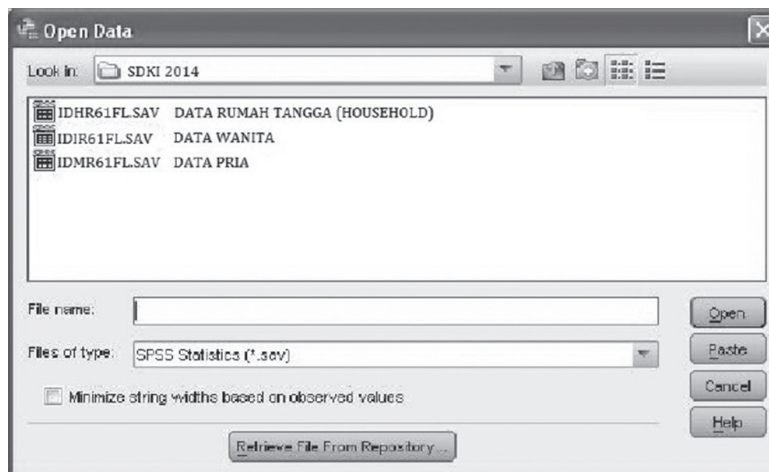
● **TABEL 9.1** Definisi Operasional Data SDKI 2012

No	Variabel	Definisi Operasional	Hasil Ukur	Cara Ukur	Nilai Ukur
1	Bobot ( <i>Sample weight</i> )	Nilai bobot responden wanita V005			Numerik
2	Umur WUS	Usia responden saat ini dihitung dari tanggal kelahiran V012	Usia responden dalam tahun	Kuesioner WUS nomor 103	Rasio
3	Pendidikan WUS	Tingkat pendidikan responden yang telah ditempuh saat ini V106	1. Tidak sekolah 2. SD/MI sederajat 3. SMA/SMK/MA sederajat 4. Akademi/D-1/D-2/D-3 5. Diploma/universitas	Kuesioner WUS, nomor 104-105	Ordinal

2. Tentukan kelompok target, *household* (rumah tangga)–IDHR61FL.SAV, wanita-IDIR61FL.SAV, atau pria-IDMR61FL.SAV

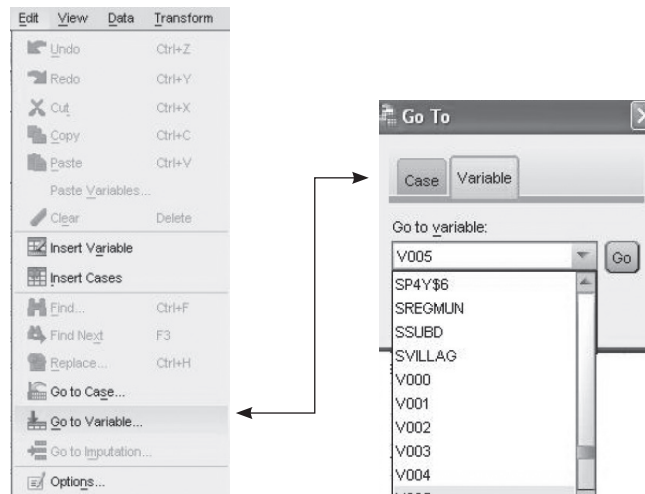
**File → Open → Data**

● **GAMBAR 9.1** Open Data ADS



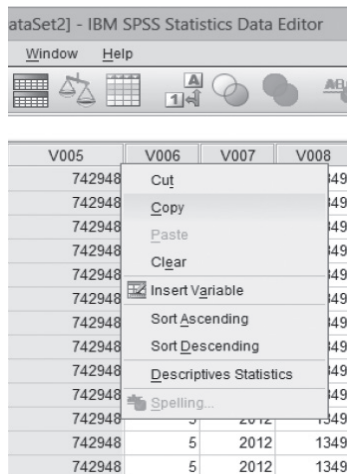
3. Setelah mengetahui kode pada masing-masing variabel yang dibutuhkan, cari variabel dengan **Edit → Go to Variable**

● **GAMBAR 9.2** Langkah Go to Variable ADS



4. *Copy* data yang ada di variabel V005 ke lembar kerja SPSS.

● **GAMBAR 9.3** Langkah Copy Data pada Kode V005



4. Lakukan hal yang sama hingga variabel yang kita butuhkan pada penelitian kita, misal umur WUS (V012) dan pendidikan WUS (V106) terpenuhi sebelum melakukan analisis sekunder lanjutan, lalu simpan file baru SPSS kita dengan nama baru, misal HIV AIDS SDKI.sav.

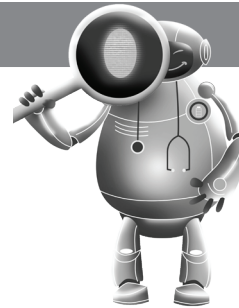


# PROSES PENGOLAHAN DATA SDKI 2012

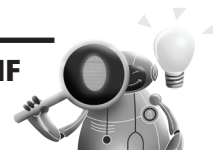
## STUDI KASUS 9.2

Lakukan proses pembobotan lalu analisis deskriptif variabel umur dan tingkat pendidikan ibu.

 Buka data: HIV AIDS \_NAJMAH\_SDKI.sav



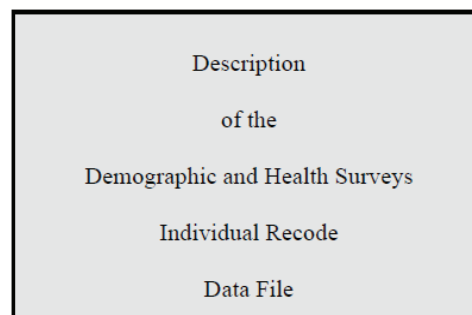
### JAWABAN DETEKTIF



#### LANGKAH 1: MEMBUAT NORMALISASI BOBOT

- Cari variabel bobot di buku panduan koding SDKI 2012 pada kelompok wanita → V005.

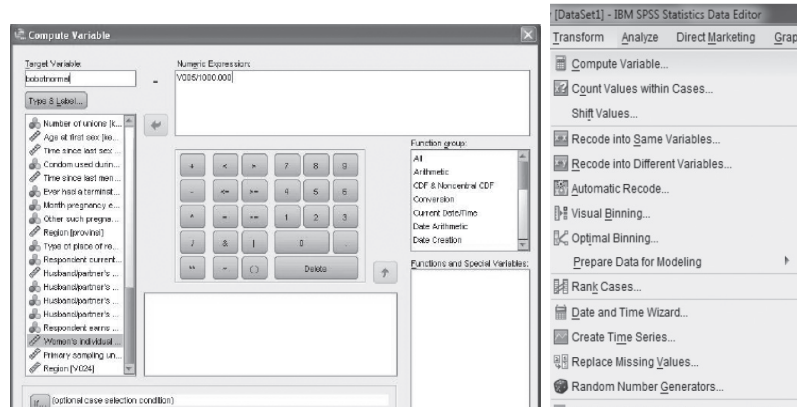
#### ● GAMBAR 9.4 Panduan Normalisasi Bobot



V005 Sample weight is an 8 digit variable with 6 implied decimal places. To use the sample weight divide it by 1000000 before applying the weighting factor. All sample weights are normalized such that the weighted number of cases is identical to the unweighted number of cases when using the full dataset with no selection. This variable should be used to weight all tabulations produced using the data file. For self-weighting samples this variable is equal to 1000000.

- Buat variabel baru “bobotnormal” dengan cara =  $V005/1000000$  (1juta) *Transform* → *Compute Variable*.

### ● GAMBAR 9.5 Langkah Compute Variabel



## LANGKAH 2: MEMBUAT PLAN UNTUK ANALISIS PEMBOBOTAN/SITE PLAN (COMPLEX SAMPLE)

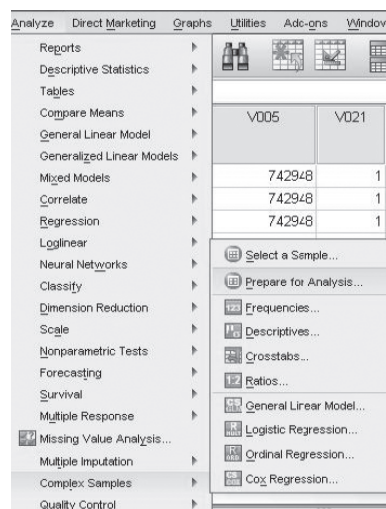
### ■ Cari Variabel **Strata (V024)** dan **cluster (V021)**

- V024** *De facto* region of residence. This is a copy of V101, added to this section to allow for analysis of completion rates by region.
- V021** Primary sampling unit is a number assigned to sample points to identify the primary sampling units for use in the calculation of sampling errors. This variable is usually the same as the cluster number and/or the ultimate area unit, but may differ if the sample design required a multistage selection process.

Sumber: Hal 10, Description of the Demographic and Health Survey-Individual recode, Data File, Measure DHS (DHS V) Version 1, 2012.

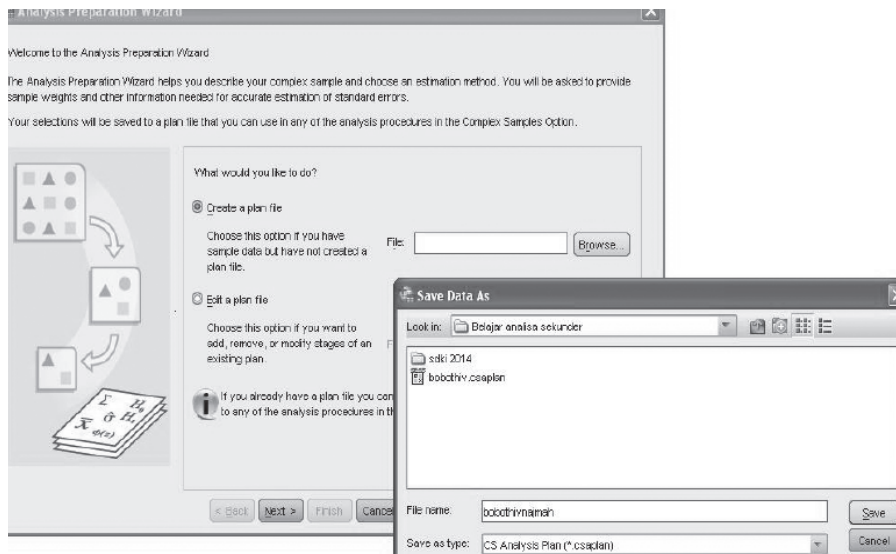
### ■ *Analyze* → *Complex Samples* → *Prepare for Analysis*

### ● GAMBAR 9.6 Langkah Analyze Site Plan (Pembobotan)



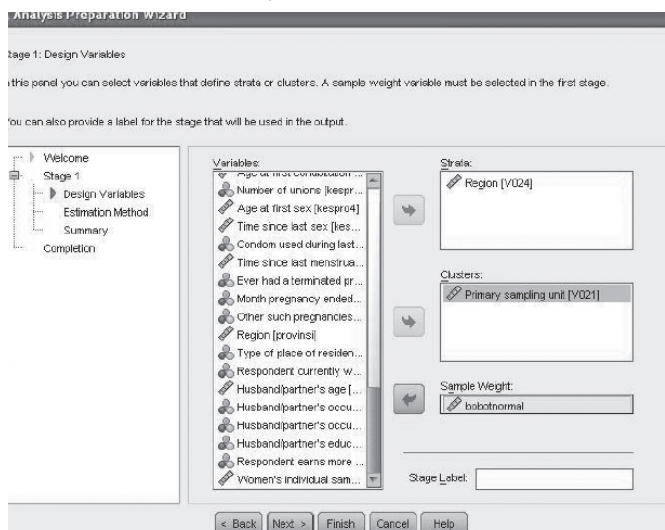
- Klik **Browse** → Buat nama File untuk plan, contoh: “bobothivnajmah” → **save** → **next**

● **GAMBAR 9.7** Langkah Create a New File untuk Plan



- Masukkan variabel Strata (V024), cluster (021) dan bobot normal. Klik **Next** → **Next** → **Next** → **Finish**, cek folder di mana data Anda analisis, maka akan terbentuk variabel **bobothivnajmah.csaplan** di folder data Anda.

● **GAMBAR 9.8** Input Variabel Strata, Cluster, dan Bobot Normal

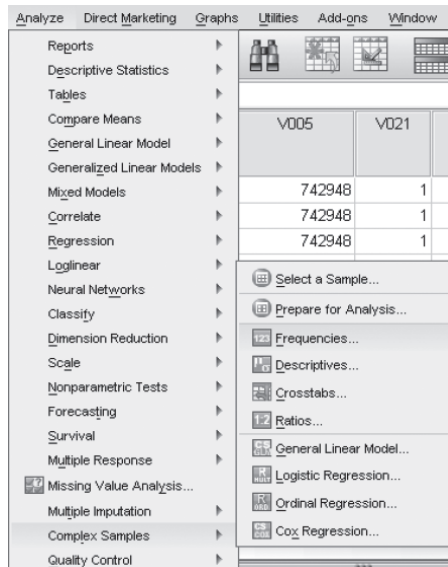


## LANGKAH 3: ANALISIS DENGAN PEMBOBOTAN

Misal: analisis jenis pekerjaan pada penelitian perilaku pencegahan penularan HIV-AIDS dan sikap terhadap ODHA pada wanita usia subur Indonesia.

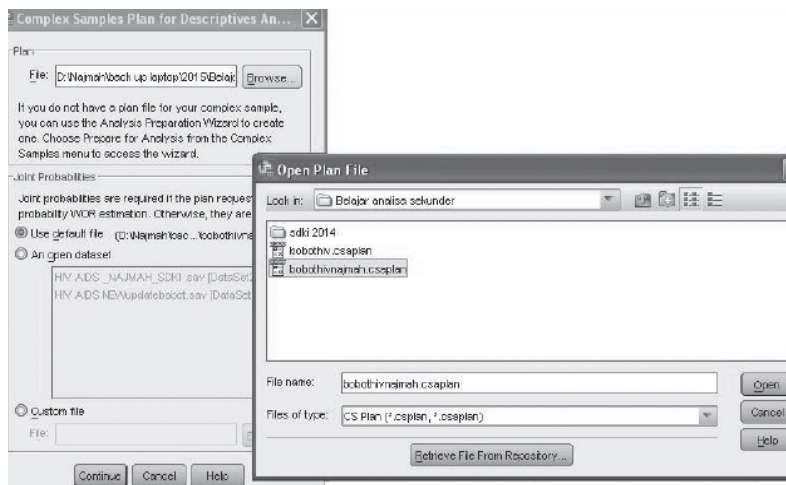
### 1. *Analyze* → *Complex Samples* → *Frequencies*.

#### ● **GAMBAR 9.9** Langkah *Analyze* *Frequencies* *Complex Samples*



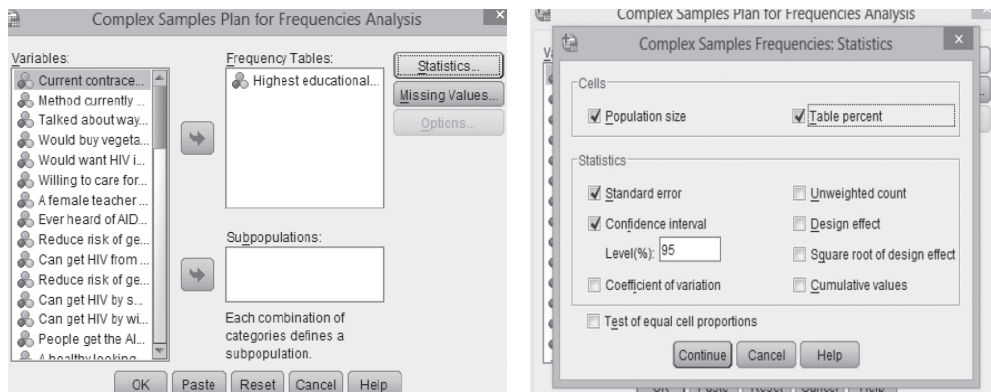
### 2. Pilih *Browse* → masukkan plan yang sudah dibuat “bobothivnajmah.csplan” → *Continue*.

#### ● **GAMBAR 9.10** Input Plan yang sudah dibuat pada *Analyze* *Frequencies*



- Masukkan variabel pendidikan ke kolom *Frequency Tables* → *Statistics* → *checkboxlist Standard error, confidence interval, dan table percent* → *Continue*.

● **GAMBAR 9.11** Complex Sample Frequencies Analysis dan Statistik



- Output Complex Samples

● **GAMBAR 9.12** Output Complex Samples

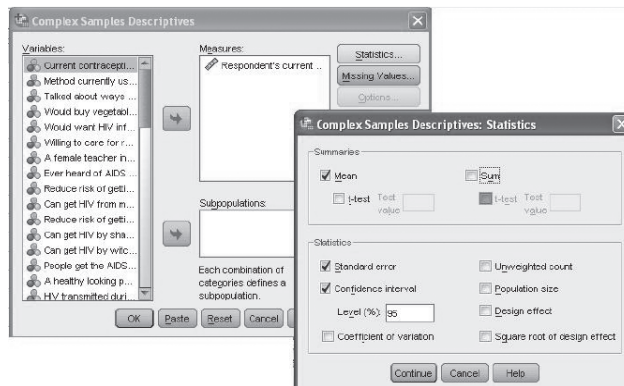
**Highest educational level**

		Estimate	Standard Error	95% Confidence Interval	
				Lower	Upper
Population Size	No education	1499,957	95,636	1312,388	1687,525
	Primary	15124,706	335,592	14466,515	15782,898
	Secondary	23430,217	352,940	22738,003	24122,431
	Higher	5552,120	244,035	5073,498	6030,741
	Total	45607,000	384,432	44853,019	46360,981
% of Total	No education	3,3%	0,2%	2,9%	3,7%
	Primary	33,2%	0,7%	31,8%	34,6%
	Secondary	51,4%	0,6%	50,2%	52,6%
	Higher	12,2%	0,5%	11,2%	13,2%
	Total	100,0%	0,0%	100,0%	100,0%

Dari 45.607 wanita usia subur di Indonesia sebanyak 1.500 (3.3%) wanita tidak bersekolah, 15.125 (33.2%) wanita tamat SD, 23.430 (51.4%) wanita tamat SMP dan 5.552 (12,2%) tamat SMA/PT. Dari data ini dapat dilihat bahwa mayoritas pendidikan wanita usia subur di Indonesia adalah tamat SMP.

- Masukkan variabel yang akan dianalisis misalnya umur responden (*respondent's current age*), pilih *statistic* → klik *mean, standar error, confidence interval, dan sum*.

### ● GAMBAR 9.13 Complex Sample Descriptives



#### 6. Output SPSS

##### Univariate Statistics

		Estimate	Standard Error	95% Confidence Interval	
				Lower	Upper
Mean	Respondent's current age	31,45	,072	31,31	31,59

7. Interpretasi: rata-rata umur responden adalah 31 tahun dengan rentang umur di populasi antara 31 hingga 32 tahun. Rentang derajat kepercayaan umur cukup presisi.

## UJI REGRESI LINEAR COMPLEX SAMPLE

Perbedaan uji regresi linear complex sample (*complex samples general linear model*—CSGLM) dengan analisis regresi linear biasa terletak pada metode sampling dan pembobotan. Pada *complex samples* metode yang digunakan adalah *cluster sampling* dan perlu memperhitungkan *weight/bobot*. Jika hanya menggunakan analisis regresi Linear biasa maka hasil kurang presisi karena jumlah populasi dan jumlah sampel dipilih pada setiap kelompok (*cluster*) yang menyebabkan probabilitas sampel untuk terpilih tidak sama.

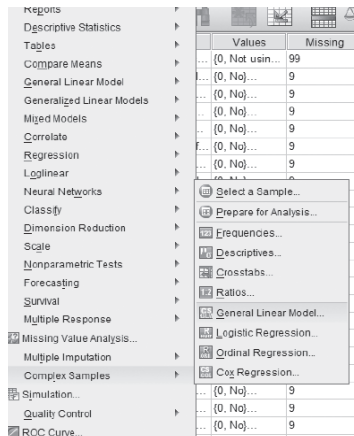


Buka data: HIV AIDS \_NAJMAH\_SDKI.sav

Berikut langkah-langkah *Complex Samples General Linear Model* (CSGLM).

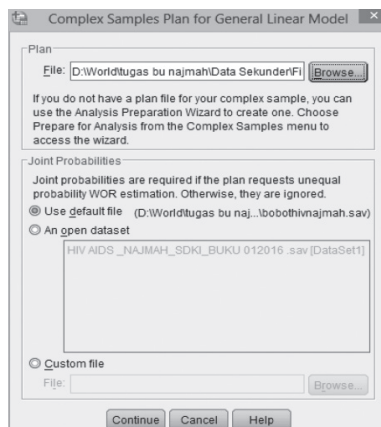
1. **Analyze** → **Complex Samples** → **General Linear Model**

● **GAMBAR 9.14** Langkah Complex Samples General Linear Model



2. Browse file plan bobothivnajmah.csaplan → **Continue**

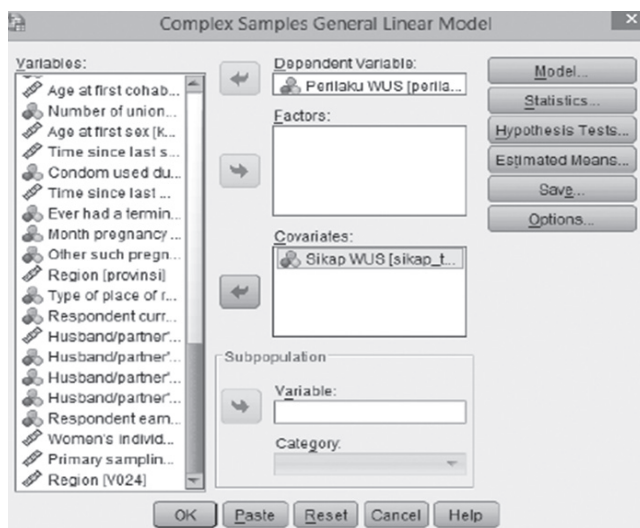
● **GAMBAR 9.15** Kotak Browse File Plan Complex Samples



3. Masukkan Variabel Dependen Perilaku Pencegahan Penularan HIV/AIDS pada WUS di kotak Variabel Dependen dan Variabel Independen Sikap WUS terhadap ODHA di kotak Covariates → **Ok**



● **GAMBAR 9.16** Kotak Input Variabel Dependen dan Independen CSGLM



#### 4. Output SPSS

Koefisien determinasi (R Square)

- ✓ Seberapa besar variasi variabel dependen dapat dijelaskan oleh variabel independen
- ✓ Seberapa besar variabel independen memengaruhi variabel dependen

R Square = 0–100%

##### Model Summary<sup>a</sup>

R Square	,000
----------	------

a. Model: Perilaku  
WUS = (Intercept) +  
sikap\_total

##### Tests of Model Effects<sup>a</sup>

Source	df1	df2	Wald F	Sig.
(Corrected Model)	1,000	1777,000	4,105	,043
(Intercept)	1,000	1777,000	3429,377	,000
sikap_total	1,000	1777,000	4,105	,043

a. Model: Perilaku WUS = (Intercept) + sikap\_total

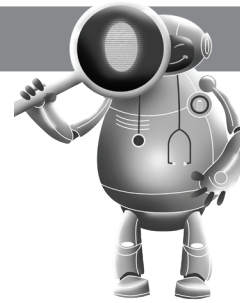
Hasil analisis diperoleh R square = 0,0005, artinya perilaku pencegahan HIV/AIDS pada WUS tidak memengaruhi sikap WUS terhadap ODHA (0,05%), sedangkan sisanya sebesar 95% dipengaruhi atau dijelaskan oleh variabel lain yang tidak dimasukkan dalam model penelitian. Nilai signifikansi 0.043 menunjukkan model regresi sikap WUS terhadap ODHA signifikan bisa memprediksi perilaku WUS dalam mencegah HIV atau ada pengaruh antara perilaku WUS dalam mencegah HIV dan sikap WUS terhadap ODHA.



### STUDI KASUS 9.3

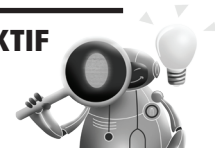
#### Pengaruh Pengetahuan dan Sikap WUS terhadap ODHA

 Buka data: HIV AIDS \_NAJMAH\_SDKI.sav

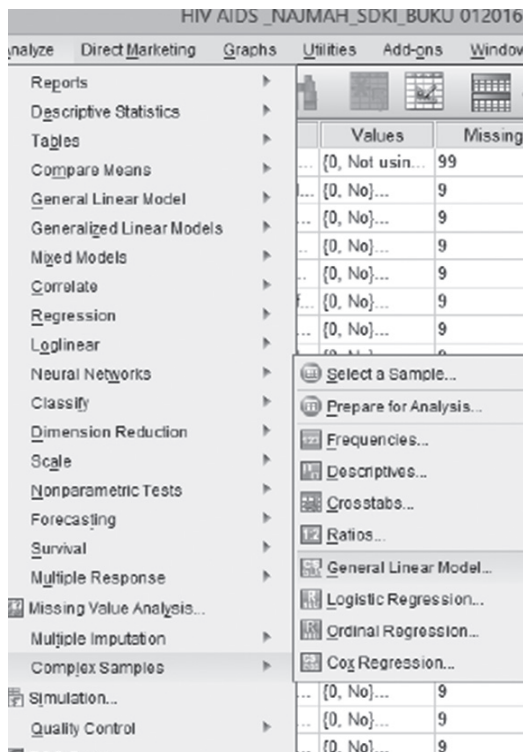


#### JAWABAN DETEKTIF

1. **Analyze → Complex Samples → General Linear Model**

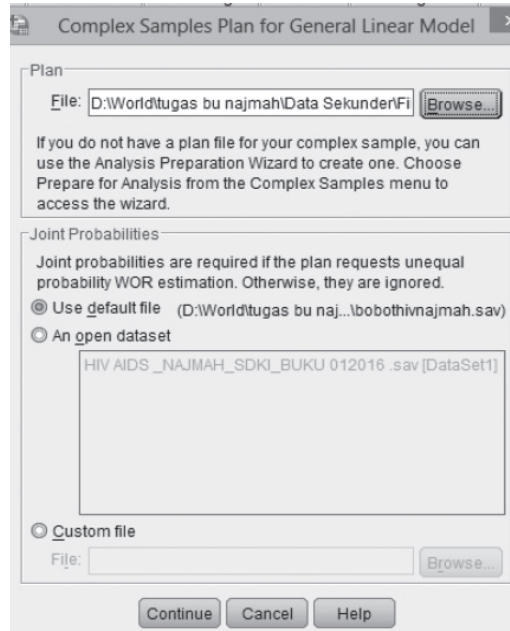


● **GAMBAR 9.17** Langkah Complex Samples General Linear Model



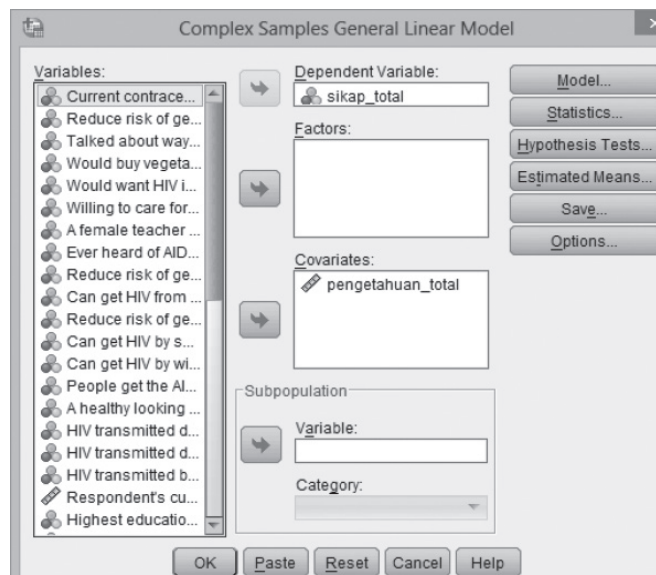
## 2. Browse file plan bobothivnajmah.csaplan → Continue

### ● GAMBAR 9.18 Kotak Browse File Plan Complex Samples



## 3. Masukkan Variabel Dependen Sikap di kotak Variabel Dependen dan Variabel Independen Pengetahuan di kotak Covariates → OK.

### ● GAMBAR 9.19 Kotak Input Variabel Dependen dan Independen CSLM



4. Output SPSS

Koefisien determinasi (R square)

- ✓ Seberapa besar variasi variabel dependen dapat dijelaskan oleh variabel independen
- ✓ Seberapa besar variabel independen memengaruhi variabel dependen

● **GAMBAR 9.20** Output Variabel Dependen dan Independen CSGLM

R square = 0 – 100%

**Model Summary<sup>a</sup>**

R Square	,131
----------	------

a. Model: sikap\_total  
= (Intercept) +  
pengetahuan\_total

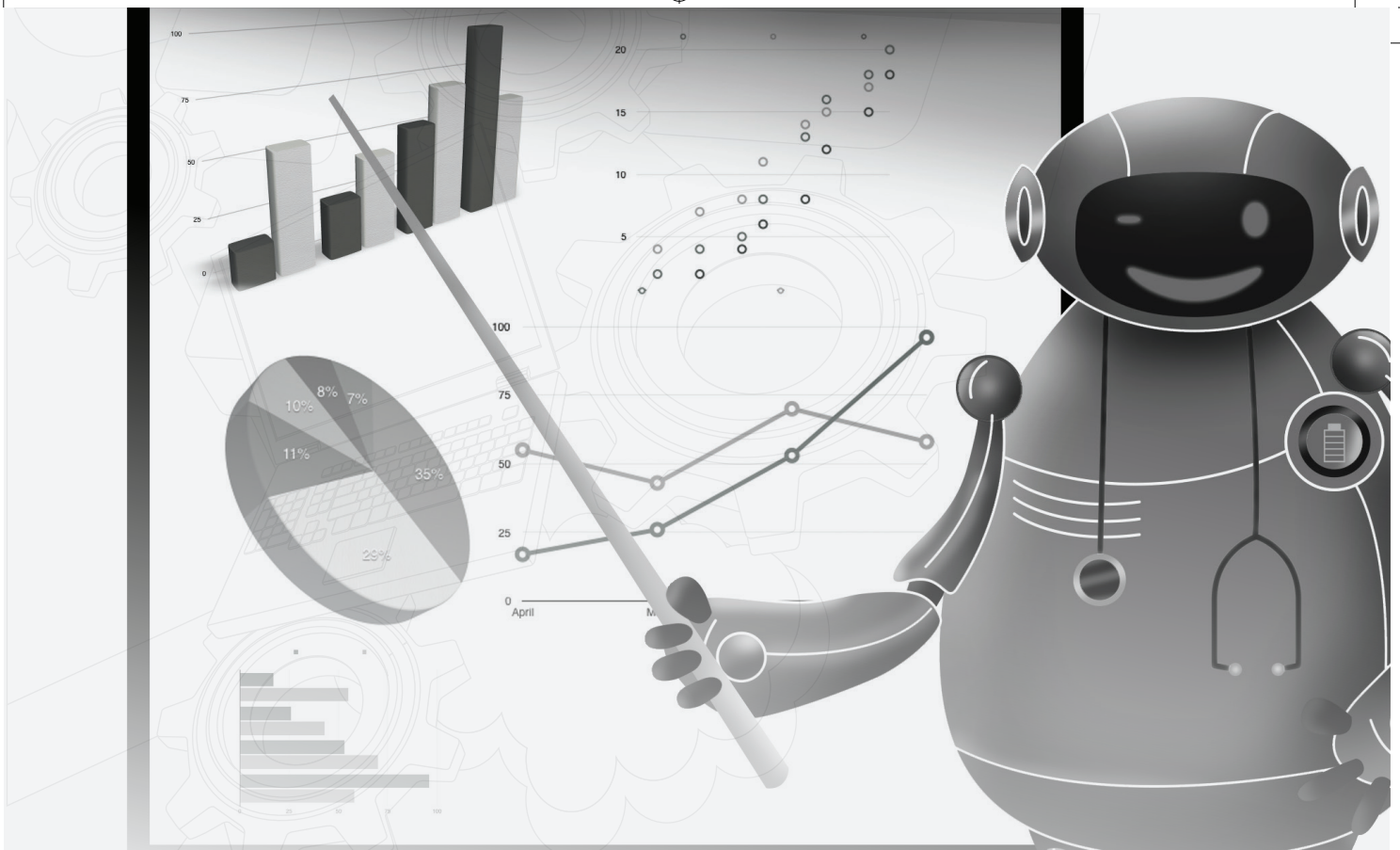
**Tests of Model Effects<sup>a</sup>**

Source	df1	df2	Wald F	Sig.
(Corrected Model)	1,000	1783,000	1432,753	,000
(Intercept)	1,000	1783,000	3970,196	,000
pengetahuan_total	1,000	1783,000	1432,753	,000

a. Model: sikap\_total = (Intercept) + pengetahuan\_total

Hasil analisis diperoleh R square = 0.131, artinya pengetahuan HIV/AIDS pada WUS memengaruhi sikap WUS terhadap ODHA sebesar 13.1%, sedangkan sisanya sebesar 86.9% dipengaruhi atau dijelaskan oleh variabel yang tidak dimasukkan dalam model regresi ini. Nilai signifikan, p value < 0.0001, menunjukkan model regresi ini bisa memprediksi sikap WUS terhadap ODHA atau ada pengaruh antara pengetahuan WUS tentang HIV/AIDS dan sikap WUS terhadap ODHA.





# BAB 10

## REGRESI LOGISTIK SEDERHANA & GANDA DENGAN SPSS

### Kompetensi Dasar

Mampu menjelaskan analisis regresi logistik sederhana menggunakan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan perbedaan Regresi Logistik Sederhana dan Berganda.
- Mampu menjelaskan aplikasi Regresi Logistik Sederhana Menggunakan SPSS.

### Materi Pembelajaran

1. Perbedaan Regresi Logistik Sederhana dan Berganda
2. Aplikasi Regresi Logistik Sederhana pada SPSS

## PERBEDAAN REGRESI LOGISTIK SEDERHANA DAN BERGANDA

Regresi logistik adalah metode yang paling banyak digunakan untuk menganalisis variabel *outcome* secara biner/dikotom. Analisis regresi logistik merupakan suatu pendekatan model yang matematis untuk menganalisis hubungan antara satu atau beberapa variabel independen (kategori dan numerik) dengan variabel dependen kategorik yang bersifat dikotom/biner. Variabel kategorik dikotom adalah variabel dengan dua nilai variasi atau kategori, misal status patah tulang (1 = patah tulang, 0 = tidak patah tulang), status BBLR (0 = BBLR, 1 = normal), dan status PJK (0 = tidak PJK; 1 = PJK) dan sebagainya. Regresi logistik sederhana untuk mengolah data dengan satu variabel independen atau faktor risiko, sedangkan regresi logistik berganda digunakan untuk mengolah data lebih dari satu variabel independen atau faktor risiko. Model regresi logistik dapat digunakan pada penelitian yang menggunakan metode potong lintang (*cross sectional*), kasus kontrol (*case-control*), maupun kohort (*cohort*).

Regresi logistik mengevaluasi efek dari satu faktor paparan atau lebih, dan biasa digunakan untuk:

1. Membandingkan variabel *outcome* di antara dua kategori dari variabel paparan atau perlakuan/perawatan.
2. Membandingkan lebih dari dua kelompok paparan atau faktor risiko.
3. Menganalisis efek variabel paparan baik kategori maupun kontinu (numerik).

**Pertama**, mari perhatikan rumus dasar *odds ratio* untuk memahami keterkaitan istilah *odds ratio* terhadap pemodelan hasil regresi logistik.

$$\text{Odds ratio variabel X} = \frac{\text{Odds pada kelompok terpapar}}{\text{Odds pada kelompok tidak terpapar}}$$

$$\text{Odds pada kelompok terpapar} = \text{Odds pada kelompok tidak terpapar} \times \text{OR variabel X}$$

**Kedua**, formula dasar pemodelan regresi logistik adalah sebagai berikut.

**Odd Ratio\* = Baseline × paparan (Exposure) .....(1 paparan)**  
**Odd Ratio \*= Baseline × paparan1 (Exposure1) × paparan2 (Exposure2) .....(2 paparan)**  
 \*OR *outcome*/penyakit/kondisi kesehatan

*Odds ratio* pada tiap kelompok terdiri atas dua model parameter:

- a. Baseline *odds*, nilai *odds* pada kelompok yang tidak terpapar.
- b. OR (*odds ratio*) *outcome*, menggambarkan efek paparan atau faktor risiko/variabel independen terhadap *odds ratio* penyakit.

● **TABEL 10.1** Model Parameter Tiap Odds Ratio

Grup paparan	Odds outcome	Odds outcome, pada regresi logistik
Terpapar (grup 1)	Baseline odds x OR faktor paparan	Baseline x faktor paparan/variabel independen
Tidak terpapar (grup 0)	Baseline Odds	Baseline

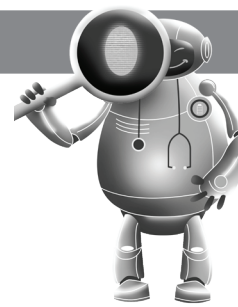
Sumber:

Sabri, L. and P.H. Sutanto, Modul Biostatistik dan Statistik Kesehatan 1999, Depok: Program Pascasarjana Program Studi IKM, Universitas Indonesia.


Betty, R. Kirkwood, and A. C. Jonathan. "Essential medical statistics." Kirkwood and Jonathan AC Sterne: Blackwell Science Ltd 414 (2003): 425.

# APLIKASI REGRESI LOGISTIK SEDERHANA (APLIKASI SPSS)

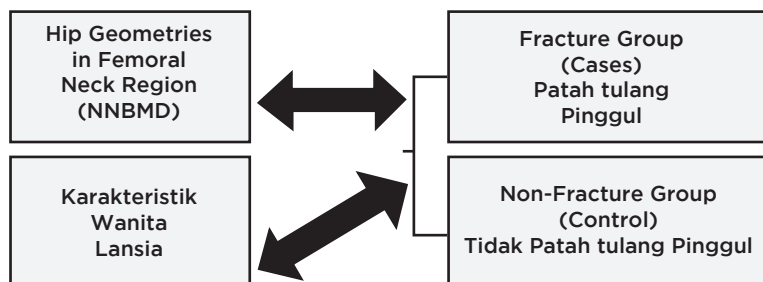
## STUDI KASUS 10.1



- Peneliti ingin mengetahui faktor-faktor yang memengaruhi kejadian patah tulang pinggul pada lansia di Geelong, Australia.
- Variabel independennya: (1) geometri pada tulang pinggul (Kepadatan tulang/Bone Mineral Density/BMD); (2) umur (*current age*); (3) berat badan (*weight*); (4) tinggi badan (*height*); (5) status minum alkohol (*drink status*); (6) status merokok (D1, *smoking*); (7) riwayat patah tulang pada keluarga (*Hip\_fracture\_parents\_spine\_hip*); (8) konsumsi kalsium dan vitamin D (*Calcium\_VitD2*); (9) status menopause (*F5xmenoadj*), (10) penggunaan steroid/kortisteroid (*Steroid2*)

 Buka Data: kasus\_iii\_hip\_fracture1.sav

Sumber: Najmah, Margaret Henry, Lyle Gurrin, Julie Pasco, 2009, Studi Geelong Osteoporosis Study/Gos, Melbourne, Australia, Thesis: 2009, University of Melbourne.



● **TABEL 10.2** Analisis Regresi Logistik Sederhana

Karakteristik/ <i>outcome</i> (patah tulang, n = 44, tidak patah tulang, n = 454)	OR	95% CI	P Value
Penggunaan hormon steroid/kortikosteroid, n(%) Ya			
Status merokok, n(%) Ya			
Status menopause, n(%) Ya			
Penggunaan kalsium/multivitamin D, n(%) Ya			
Terapi hormon, n(%)			
Physical Activity, n(%)			
Riwayat keluarga, n(%) Ya			
Status minum alkohol, n(%) Ya			
	<b>OR</b>	<b>95% CI</b>	<b>P Value</b>
Tinggi badan, cm			
Usia saat ini, tahun			
Berat badan, kg			
Kadar mineral tulang, g/cm <sup>2</sup>			

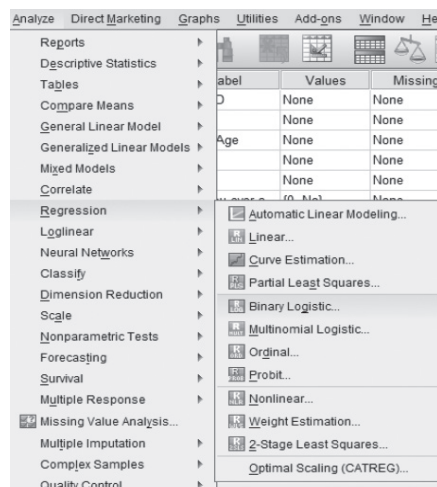
## JAWABAN DETEKTIF



### Langkah-Langkah Uji Regresi Linear Sederhana Aplikasi SPSS Variabel Status Merokok (D1)

Langkah 1: pada menu bar pilih *Analyze* → *Regression* → *Binary Logistic*.

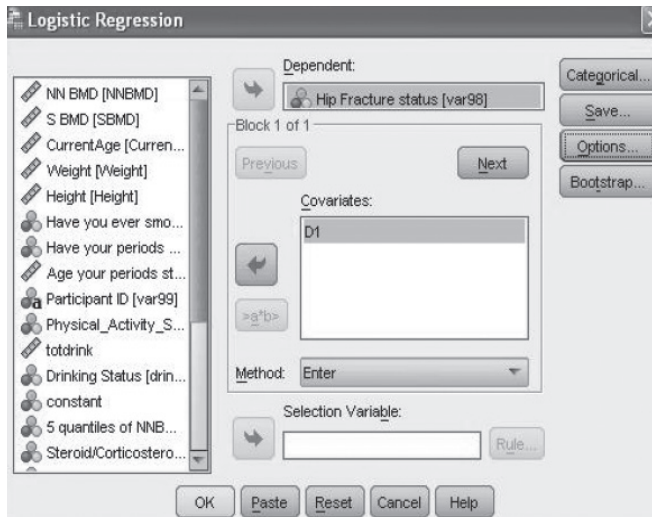
● **GAMBAR 10.1** Proses Analisis Regresi Logistik





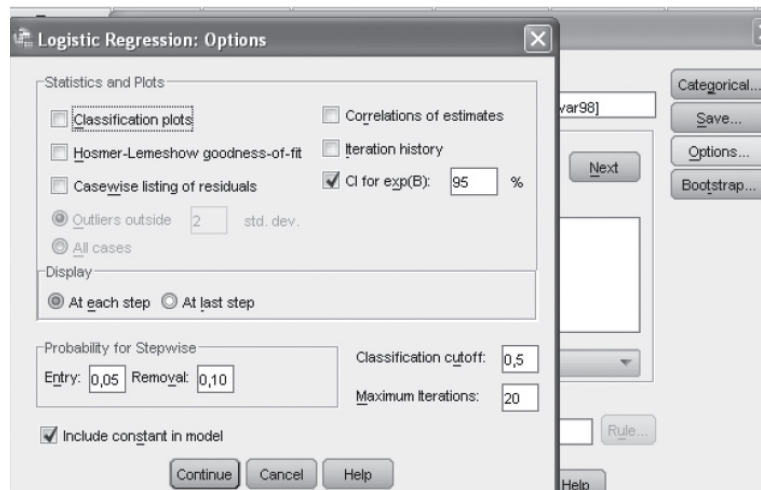
**Langkah 2:** masukkan variabel *hip fracture* status (status patah tulang) pada variabel *dependen*, dan masukkan status merokok (D1) variabel independen ke dalam *covariate*.

● **GAMBAR 10.2** Kotak Dialog Regresi Logistik



**Langkah 3:** klik **Option** dan ketik 95% Derajat Kepercayaan

● **GAMBAR 10.3** Kotak Dialog Options Regresi Logistik



**Langkah 4:** klik **continue** dan **OK**.

\*\*Identifikasi data yang hilang dan pemberian kode variabel dependen oleh SPSS

**Case Processing Summary**

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	499	100,0
	Missing Cases	0	,0
	Total	499	100,0
Unselected Cases		0	,0
Total		499	100,0

a. If weight is in effect, see classification table for the total number of cases.

**Dependent Variable Encoding**

Original Value	Internal Value
No	0
Yes	1

**Classification Table<sup>a,b</sup>**

Observed		Predicted		
		Hip Fracture status		Percentage Correct
		No	Yes	
Step 0	Hip Fracture status No	454	0	100,0
	Yes	45	0	,0
Overall Percentage				91,0

a. Constant is included in the model.

b. The cut value is ,500

**Variables in the Equation**

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-2,311	,156	218,741	1	,000	,099

**Variables not in the Equation**

	Score	df	Sig.
Step 0 Variables D1	1,430	1	,232
Overall Statistics	1,430	1	,232

**Omnibus Tests of Model Coefficients**

	Chi-square	df	Sig.
Step 1 Step	1,383	1	,240
Block	1,383	1	,240
Model	1,383	1	,240

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	300,966 <sup>a</sup>	,003	,006

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

\*\* Nilai OR, 95% Derajat Kepercayaan dan P value nilai variabel merokok

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 <sup>a</sup> D1	,382	,321	1,417	1	,234	1,465	,781	2,746
Constant	-2,447	,201	148,815	1	,000	,087		

a. Variable(s) entered on step 1: D1.

### Output Status Minum Alkohol

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 <sup>a</sup> drinkstatus	-1,226	,371	10,947	1	,001	,293	,142	,607
Constant	-1,883	,181	107,678	1	,000	,152		

a. Variable(s) entered on step 1: drinkstatus.

### Output SPSS BMD (*bone mineral density*)

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 <sup>a</sup> stdNNBMD	-,782	,191	16,737	1	,000	,458	,315	,666
Constant	-2,508	,189	175,468	1	,000	,081		

a. Variable(s) entered on step 1: stdNNBMD.

### Output SPSS Penggunaan Kortikosteroid/Steroid

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 <sup>a</sup> Steroid2	-,532	,616	,745	1	,388	,587	,175	1,967
Constant	-2,261	,162	194,490	1	,000	,104		

a. Variable(s) entered on step 1: Steroid2.

### LATIHAN:



Coba Anda lakukan latihan untuk variabel lainnya dengan langkah yang sama!  
 Variabel independennya: (1) geometri pada tulang pinggul (Kepadatan tulang/Bone Mineral Density/stdNNBMD); (2) **umur (current age)**; (3) **berat badan (weight)**; (4) **tinggi badan (height)**; (5) status minum alkohol (*drink status*); (6) status merokok (D1, *smoking*); (7) **riwayat patah tulang pada keluarga (Hip\_fracture\_parents\_spine\_hip)**; (8) **konsumsi kalsium dan vitamin D (Calcium\_VitD2)**; (9) **status menopause (F5xmenoadj)**; (10) penggunaan steroid/kortikosteroid (*Steroid2*).

**Langkah 5:** menuliskan laporan tabel dan interpretasi.

● **TABEL 10.3** Karakteristik Responden pada Kelompok Patah Tulang dan Tidak Patah Tulang

<b>Karakteristik/outcome (patah tulang, n = 44, tidak patah tulang, n = 454)</b>	<b>OR</b>	<b>95% CI</b>	<b>P Value</b>
Penggunaan hormon steroid/kortikosteroid, n (%) Ya	0,587	0,175-1,966	0,388
Status merokok, n (%) Ya	1,464	0,781-2,746	0,234
Penggunaan kalsium/multivitamin D, n (%) Ya	0,383	0,134-1,099	0,074
Aktivitas fisik, n (%) Ya	4,946	2,110-11,589	0,001
Riwayat keluarga, n (%) Ya	0,333	0,044-2,504	0,286
Status minum alkohol, n (%) Ya	0,293	0,141-0,606	0,001
	<b>OR</b>	<b>95% CI</b>	<b>P Value</b>
Tinggi badan, cm	1,009	0,960-1,060	0,715
Usia saat ini, tahun	1,076	1,028-1,127	0,002
Berat badan, kg	0,972	0,946-0,999	0,043
Kadar mineral tulang, g/cm <sup>2</sup>	0,457	0,314-0,665	0,001

**Interpretasi untuk hasil numerik:**

- Setiap kenaikan ketebalan 1 g/cm<sup>2</sup> kadar mineral tulang akan menurunkan risiko patah tulang pinggul sebesar 0,45 kali. Dengan derajat kepercayaan 95%, di populasi setiap kenaikan ketebalan 1 g/cm<sup>2</sup> kadar mineral tulang, akan menurunkan risiko patah tulang pinggul sebesar 0,31 kali hingga 0,67 kali. Nilai P 0,001 menunjukkan adanya hubungan antara kadar mineral tulang dan kejadian patah tulang pinggul pada wanita di Geelong, Australia.
- Perilaku tidak merokok dapat meningkatkan 1,5 kali risiko untuk tidak patah tulang pinggul dibandingkan dengan perilaku merokok. Di populasi, dengan derajat kepercayaan 95%, perilaku tidak merokok menurunkan 0,78 risiko tidak patah tulang pinggul dan meningkatkan risiko 2,7 untuk tidak patah tulang dibandingkan wanita perokok. Kesimpulan akhir dengan nilai signifikansi 0,23 menunjukkan tidak ada hubungan antara status merokok dan kejadian patah tulang pinggul pada wanita di Geelong Australia.
- Coba Anda interpretasikan hasil lainnya!



# BAB 11

## REGRESI LINEAR SEDERHANA & GANDA (SPSS & STATA)

### Kompetensi Dasar

Mampu menjelaskan aplikasi uji regresi linear berganda menggunakan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan analisis regresi linear ganda.
- Mampu menjelaskan fungsi regresi linear ganda.
- Mampu menjelaskan asumsi regresi linear.
- Mampu menjelaskan aplikasi analisis regresi linear ganda dalam SPSS.

### Materi Pembelajaran

1. Analisis regresi linear ganda
2. Regresi linear ganda
3. Asumsi regresi linear
4. Seleksi variabel regresi linear ganda

## REGRESI LINEAR GANDA

Jika persamaan garis regresi linear sederhana adalah  $y = a + b$ , maka pada regresi linear ganda adalah  $y = a + b_1 \times 1 + b_2 \times 2 + \dots$ . Pada regresi linear ganda yang menjadi variabel dependen adalah data yang bersifat *continue* atau numerik dan variabel independen bisa berupa data numerik maupun kategorik.

1. Regresi linear ganda
  - a. Variabel dependen: numerik
  - b. Variabel independen: numerik dan kategorik
2. Regresi logistik ganda
  - a. Variabel dependen: kategorik
  - b. Variabel independen: kategorik dan numerik

### FUNGSI REGRESI LINEAR GANDA

1. Menetapkan model matematis yang paling baik untuk menggambarkan hubungan variabel independen dan variabel dependen.
2. Menggambarkan hubungan kuantitatif antara variabel independen (X) dengan variabel dependen (Y) setelah dikontrol variabel lain.
3. Mengetahui variabel X mana yang penting/dominan dalam memprediksi variabel dependen.
4. Mengetahui adanya interaksi pada dua/lebih variabel independen terhadap variabel dependen.

### ASUMSI REGRESI LINEAR

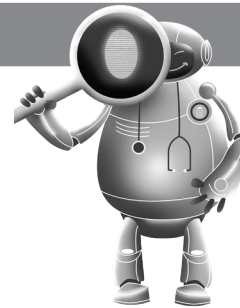
Agar prediksi yang dihasilkan valid maka harus memenuhi asumsi regresi linear, antara lain:

1. *Homoscedasticity*
  - a. Varian nilai variabel Y sama untuk semua nilai variabel X.
  - b. *Homoscedasticity*: plot residual membentuk tebaran merata di atas dan di bawah garis tengah nol.
  - c. *Heteroscedasticity*: tebaran residual mengelompok di bawah/di atas garis tengah nol.
2. Independensi/autokorelasi
  - a. Masing-masing variabel Y bebas satu sama lain, tidak boleh diukur dua kali. Bila penelitiannya *cross sectional* berarti terpenuhi asumsinya karena tidak diukur dua kali/*time series*.
  - b. Untuk menguji asumsi ini bisa juga diuji angka.
3. Linearitas
  - a. Nilai mean dari variabel Y untuk suatu kombinasi  $X_1, X_2, \dots$ , terletak pada garis linear yang dibentuk persamaan regresi.
  - b. Asumsi terpenuhi: hasil uji Anova regresi hasilnya signifikan  $< \alpha (0,05)$

4. Gauss/Normalitas
  - a. Variabel Y berdistribusi normal untuk setiap pengamatan variabel X. Asumsi terpenuhi bila grafik Normal P-P plot residual, titik tebarannya menyebar sekitar garis diagonal. Sebaliknya bila tebaran data menjauh garis diagonal maka asumsi tidak terpenuhi.
5. Diagnostik/Pengujian Kolinearitas
 

Antar-variabel independen terjadi hubungan yang lemah, apabila mempunyai  $r \leq 0,8$  atau nilai VIF  $< 10$ .

### STUDI KASUS 11.1



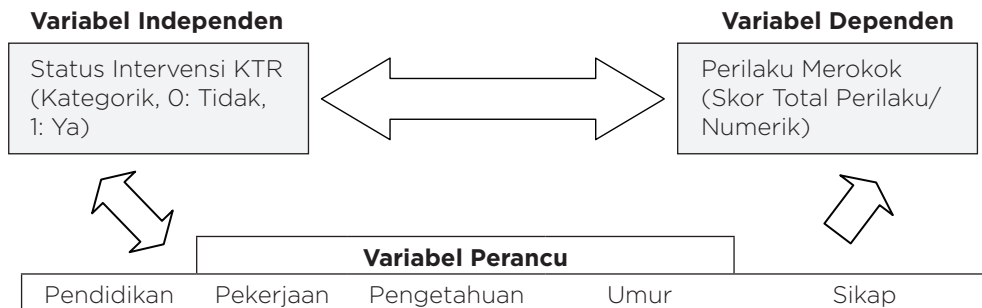
Studi Intervensi Klaster Kawasan Tanpa Rokok pada Tingkat Rumah Tangga



Buka Data: KASUS II\_KTR.sav

**Sumber data:** Najmah, Fenny Etrawati, Yeni, Feranita Utama., 2015, *Studi Intervensi Klaster Kawasan Tanpa Rokok pada Tingkat Rumah Tangga*. 2015: *Jurnal Ilmu Kesehatan Masyarakat Nasional: Universitas Indonesia*, Vol 9 No 4 Mei 2015. (<http://jurnalkesmas.ui.ac.id/index.php/kesmas/article/view/752>).

- Hipotesis: faktor apa yang dominan berpengaruh terhadap perubahan perilaku merokok pada desa yang mendapatkan intervensi terpadu kawasan tanpa rokok dan desa yang tidak mendapatkan intervensi terpadu di Kabupaten Ogan Ilir dikontrol oleh beberapa faktor perancu?



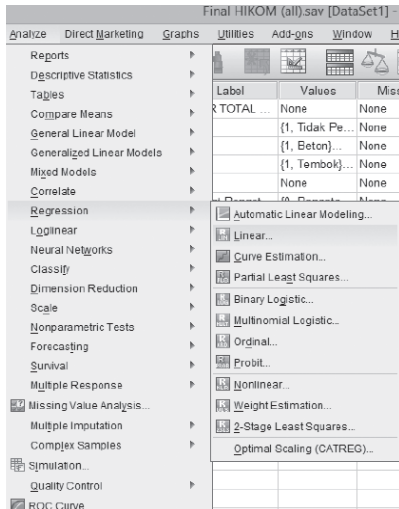
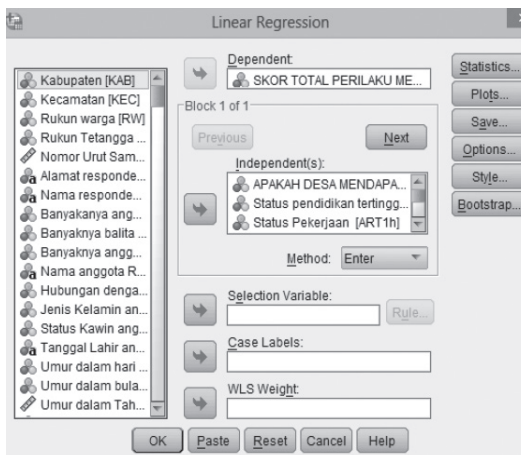
\*Informasi lebih lanjut mengenai Faktor Perancu (*Confounding factors*) di Buku Epidemiologi untuk mahasiswa kesehatan masyarakat, 2011: Rajagrafindo: Jakarta (hlm. 75-93) atau [download slides https://www.slideshare.net/najmahusman/bab-iv-faktor-perancu-part-1](https://www.slideshare.net/najmahusman/bab-iv-faktor-perancu-part-1)



## JAWABAN DETEKTIF



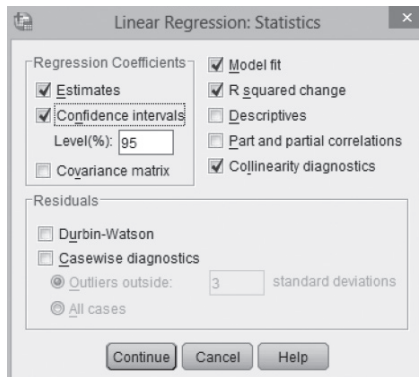
## Langkah-Langkah Pengolahan Data

1. *Analyze* → *Regression* → *Linear*● **GAMBAR 11.1** Langkah *Analyze* Regresi Linear Berganda2. Masukkan variabel dependen di kotak *Dependent* dan variabel *Independen* di kotak *Independen*.● **GAMBAR 11.2** Input Variabel *Dependent* dan *Independen*



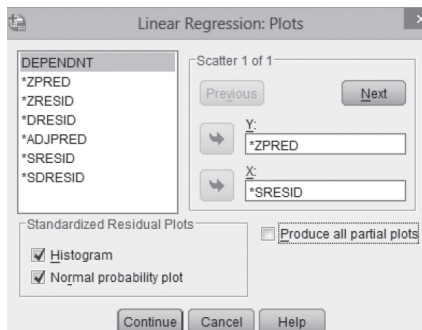
3. Pada bagian *Statistic checklist Confidence intervals, model fit, R squared change dan collinearity diagnostic* → *Continue*

● **GAMBAR 11.3** Regresi Linear: Statistics



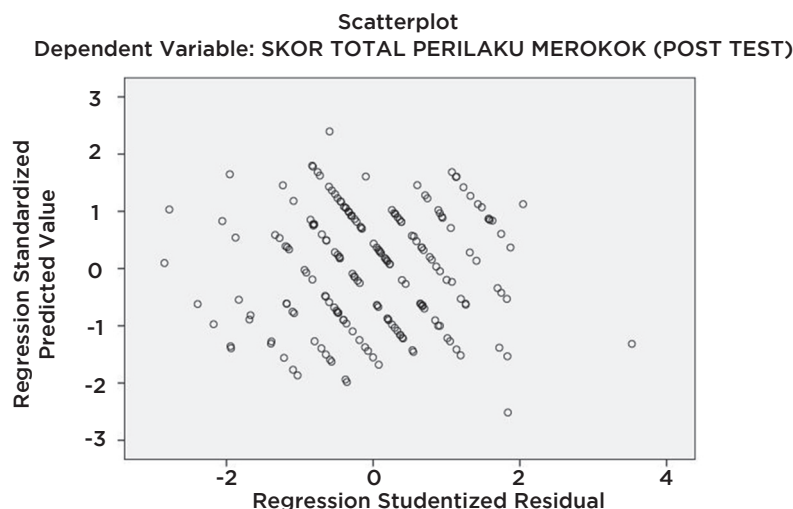
4. Pada bagian *Plots*, masukan *\*ZPRED* → *KOTAK Y*, *\*SRESID* → *KOTAK X*, *checklist Histogram dan normal probability plot* → *Continue* → *Ok*.

● **GAMBAR 11.4** Regresi Linear: Plots



5. Asumsi regresi linear
  - a. *Homoscedasticity*  
Asumsi *homoscedasticity* terpenuhi, terlihat dari tebaran data pada grafik *scatterplot* merata/serupa antara bagian atas dan bawah titik nol pada Gambar 11.5.
  - b. Asumsi Autokorelasi  
Asumsi autokorelasi sudah terpenuhi oleh karena pengumpulan datanya dilakukan satu kali/bukan *time series*.
  - c. Asumsi Linearitas  
Asumsi linearitas sudah terpenuhi oleh karena hasil uji Anova didapatkan *p-value* =  $0.0005 < \alpha (0.05)$ , artinya garis yang terbentuk dalam multivariat adalah linear.

● **GAMBAR 11.5** Asumsi Homoscedasticity Regresi Linear Ganda



● **GAMBAR 11.6** Asumsi Linearitas Regresi Linear Ganda

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	193,975	6	32,329	11,683	,000 <sup>b</sup>
	Residual	520,210	188	2,767		
	Total	714,185	194			

a. Dependent Variable: SKOR TOTAL PERILAKU MEROKOK [POST TEST]

b. Predictors: (Constant), skor\_tahu, Status Pekerjaan , APAKAH DESA MENDAPAT INTERVENSI, skor\_sikap, Umur dalam Tahun, Status pendidikan tertinggi

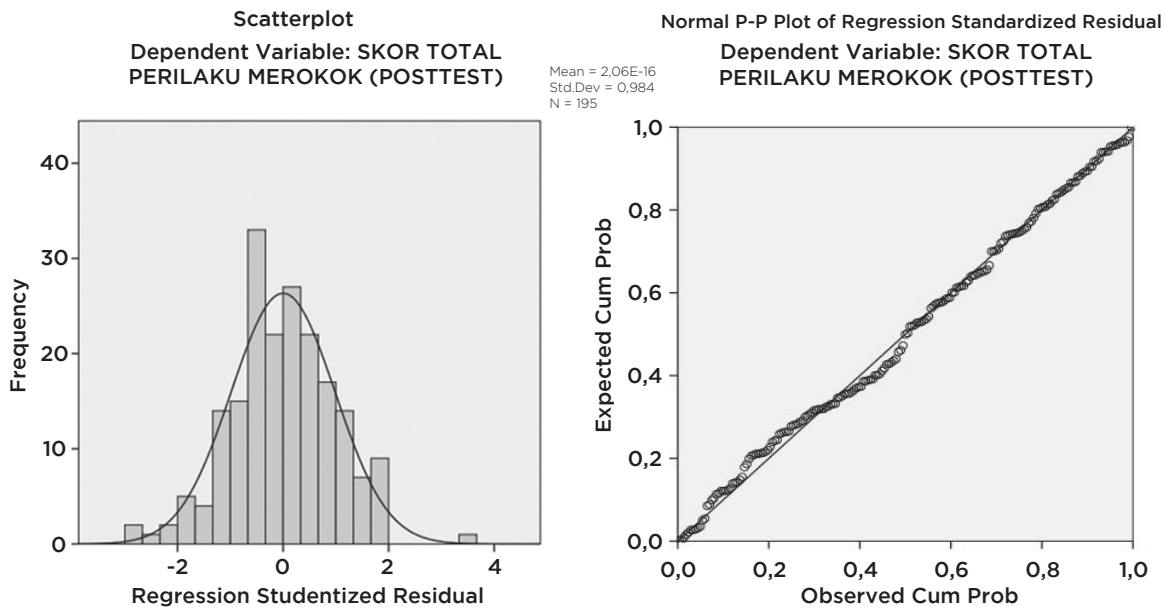
d. Asumsi Normalitas Data

Asumsi normalitas data terpenuhi karena hasil analisis dengan grafik histogram membentuk kurva normal dan grafik p-p plot menunjukkan tebaran data berimpit dengan garis diagonal.

e. Pengujian Kolinearitas

Hasil analisis menunjukkan tidak ada kolinearitas (hubungan yang sangat kuat) antara variabel intervensi, pendidikan, pekerjaan, umur, sikap, dan pengetahuan terbukti dari nilai VIF <10.

**GAMBAR 11.7** Asumsi Normalitas Data Regresi Linear Ganda



**GAMBAR 11.8** Pengujian Kolinearitas Regresi Linear Berganda

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	3,781	1,386		2,728	,007	1,047	6,516		
	APAKAH DESA MENDAPAT INTERVENSI	1,775	,245	,464	7,258	,000	1,293	2,258	,949	1,053
	Status pendidikan tertinggi	,192	,105	,125	1,835	,068	-,014	,399	,830	1,205
	Status Pekerjaan	-,080	,070	-,073	-1,149	,252	-,217	,057	,949	1,054
	Umur dalam Tahun	-,012	,010	-,075	-1,170	,244	-,031	,008	,931	1,071
	skor_sikap	,111	,030	,230	3,661	,000	,051	,170	,981	1,019
	skor_tahu	-,061	,071	-,055	-0,864	,389	-,201	,079	,939	1,065

a. Dependent Variable: SKOR TOTAL PERILAKU MEROKOK (POST TEST)

f. Kesimpulan Asumsi → Semua Asumsi dari Regresi Linear Terpenuhi

## SELEKSI VARIABEL

Coba perhatikan hasil bivariat dengan menggunakan Uji Korelasi di bawah ini, kita akan masuk ke tahap selanjutnya, yaitu aplikasi uji Regresi Linear Berganda, dengan syarat nilai signifikansi variabel ( $p\text{ value} < 0,25$ ).

● **TABEL 11.1** Hasil Uji Bivariabel antara Variabel Independen dan Variabel Dependen

Variabel	95% CI	p value
Intervensi	-2,141-(-1,159)	* <b>&lt;0,0001</b>
Pendidikan	-0,097-0,377	* <b>0,245</b>
Pekerjaan	-0,198-0,111	0,577
Umur (Tahun)	-0,036-0,007	<b>0,190</b>
Skor sikap	0,043-0,176	* <b>0,001</b>
Skor Pengetahuan	-0,202-0,111	0,556

\*\*Anda bisa melakukan uji Independent T-test atau regresi linear sederhana untuk menghasilkan tabel bivariat di atas.

Hasil analisis pada gambar di atas menunjukkan bahwa variabel independen yang dapat masuk ke dalam pemodelan adalah variabel intervensi, umur, dan sikap mengenai rokok ( $p$  value < 0,25). Sementara variabel pendidikan, pekerjaan dan skor pengetahuan tidak dapat masuk ke dalam pemodelan karena  $p$  value > 0,25. Sementara variabel pendidikan, pekerjaan dan skor pengetahuan tidak dapat masuk ke dalam pemodelan karena  $p$  value > 0,25. Secara substansi ketiga variabel tersebut dianggap penting sehingga tetap dimasukkan ke dalam pemodelan multivariat yang telah dijelaskan pada langkah-langkah sebelumnya. Langkah tersebut akan menghasilkan output SPSS sebagai berikut.

● **GAMBAR 11.9** Output Regresi Linear Ganda Full Model

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,523 <sup>a</sup>	<b>,273</b>	,250	1,662

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3,587	1,412		2,541	,012
	APAKAH DESA MENDAPAT INTERVENSI	1,769	,243	,462	7,275	,000
	PENDIDIKAN	,220	,113	,131	1,937	,054
	Status Pekerjaan	-,070	,068	-,064	-1,020	,309
	Umur dalam Tahun	-,011	,010	-,071	-1,094	,275
	skor_sikap	,109	,030	,227	3,610	,000
	skor_tahu	-,057	,071	-,052	-,808	,420

a. Dependent Variable: SKOR TOTAL PERILAKU MEROKOK

Tahap berikutnya, evaluasi seleksi variabel dengan batas  $p$  value < 0,05. Dari nilai  $p$  di atas, variabel status pekerjaan, umur, dan pengetahuan memiliki  $p$  value > 0,05 sehingga harus keluar dari model. Pengetahuan memiliki  $p$  value yang tertinggi, maka dikeluarkan terlebih dahulu. Kemudian lakukan analisis regresi linear ganda kembali terhadap variabel intervensi, pendidikan, pekerjaan, umur, sikap, dan pengetahuan. Dihasilkan output SPSS sebagai berikut.

1. Model Keseluruhan

● **GAMBAR 11.10a** Output Regresi Linear Ganda Model I

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,520 <sup>a</sup>	,271	,251	1,660

a. Predictors: (Constant), skor\_sikap, APAKAH DESA MENDAPAT INTERVENSI, Umur dalam Tahun, Status Pekerjaan , PENDIDIKAN

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2,846	1,072		2,655	,009
	APAKAH DESA MENDAPAT INTERVENSI	1,762	,243	,460	7,257	,000
	PENDIDIKAN	,237	,111	,141	2,126	,035
	Status Pekerjaan	-,070	,068	-,064	-1,018	,310
	Umur dalam Tahun	-,012	,010	-,076	-1,186	,237
	skor_sikap	,107	,030	,223	3,560	,000

a. Dependent Variable: SKOR TOTAL PERILAKU MEROKOK

Kita membandingkan OR sebelum dan setelah dikeluarkan variabel dengan *p value* tertinggi, skor pengetahuan. Setelah pengetahuan keluar terlihat *R square* berubah dan koefisien B untuk masing-masing juga tidak berubah lebih besar (> 10 %), sehingga variabel skor pengetahuan tetap dikeluarkan atau dengan kata lain skor pengetahuan bukan merupakan faktor perancu. Jika perubahan variabel Intervensi dan variabel lainnya lebih dari 10% (minimal satu variabel), maka variabel yang dikeluarkan tadi harus dimasukkan kembali. Perubahan koefisien B bisa dilihat pada Tabel 11.2

● **TABEL 11.2** Tabel Perubahan Koefisien B Sebelum dan Sesudah Variabel Pengetahuan Dikeluarkan

Variabel	Sebelum Pengetahuan Dikeluarkan	Setelah Pengetahuan Dikeluarkan	Perubahan Koefisien B
Intervensi	1.77	1.76	0.001%
Pendidikan	0.22	0.24	0.002%
Pekerjaan	-0.07	-0.07	0%
Umur	-0.01	-0.01	0%
Sikap	0.11	0.11	0%
Pengetahuan	-0.06	-	-

● **TABEL 11.3** Tahap Pemodelan Multivariabel

Variabel	Full Model (RR)	Model I (RR)	Model II (RR)	Model III (RR)
Intervensi	*0.464	*0.460	*0.460	*0.463
Pendidikan	0.125	0.141	<b>*0.132</b>	<b>*0.152</b>
Pekerjaan	-0.073	-0.064	-	-
Umur	-0.075	-0.076	-0.176	-
Sikap	*0.230	*0.223	*0.221	*0.261
Tahu	-0.055	-	-	-

Keterangan : (\*) = nilai sig (p Value) < alpha (5%)

Kemudian keluarkan variabel pekerjaan dan umur  $p\ value > 0.05$  lakukan satu persatu berdasarkan  $p\ value$  yang tertinggi, bandingkan koefisien sebelum dan sesudah apakah terjadi perubahan pada koefisien B. Selanjutnya akan didapatkan tahap pemodelan pada Tabel 11.3.

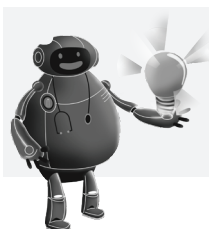
## 2. Laporan dan Interpretasi

Hasil analisis pada Tabel 11.3 menunjukkan bahwa secara bersama-sama variabel intervensi mempunyai pengaruh yang signifikan terhadap perubahan perilaku responden ( $p\ value < 0,05$ ). Hasil intervensi yang dilakukan mempunyai peluang 46% untuk mengurangi perilaku merokok responden (RR 0,46) setelah dikontrol oleh variabel pendidikan (RR 0,152) dan Sikap (RR 0,216).

● **TABEL 11.4** Laporan Hasil

Variabel	Crude RR	Adjusted RR
Intervensi	<b>*0.464</b>	<b>*0.463</b>
Pendidikan	0.125	<b>*0.152</b>
Pekerjaan	-0.073	-
Umur	-0.075	-
Sikap	<b>*0.230</b>	<b>*0.261</b>
Tahu	-0.055	-

**Sumber data:** Najmah, Fenny Etrawati, Yeni, Feranita Utama., 2015, Studi Intervensi Klaster Kawasan Tanpa Rokok pada Tingkat Rumah Tangga. 2015: Jurnal Ilmu Kesehatan Masyarakat Nasional: Universitas Indonesia, Vol 9 No. 4 Mei 2015. (<http://jurnalkesmas.ui.ac.id/index.php/kesmas/article/view/752>)



### CATATAN:

Anda cukup mengetik tahap awal untuk pemodelan. Selanjutnya, Anda tinggal delete variabel yang perlu dikeluarkan!

```

/MISSING LISTWISE
/STATISTICS COEFF OUTS R Anova
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT SKORPERILAKU
/METHOD=ENTER INTERVENSI PENDIDIKAN ART1h ART1f_tahun skor_sikap
skor_tahu.
    
```

```

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R Anova
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT SKORPERILAKU
/METHOD=ENTER INTERVENSI PENDIDIKAN ART1h ART1f_tahun skor_sikap.
    
```

```

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R Anova
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT SKORPERILAKU
/METHOD=ENTER INTERVENSI PENDIDIKAN ART1f_tahun skor_sikap.
    
```

```

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R Anova
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT SKORPERILAKU
/METHOD=ENTER INTERVENSI PENDIDIKAN skor_sikap.
    
```

#### 4. Output SPSS

##### a. Output Model I

● **GAMBAR 11.10b** Output Regresi Linear Ganda Model I

**Coefficients<sup>a</sup>**

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	2.846	1.072		2.655	.009
APAKAH DESA MENDAPAT INTERVENSI PENDIDIKAN	1.762	.243	.460	7.257	.000
Status Pekerjaan	.237	.111	.141	2.126	.035
Umur dalam Tahun	-.070	.068	-.064	-1.018	.310
skor_sikap	-.012	.010	-.076	-1.186	.237
	.107	.030	.223	3.560	.000

a. Dependent Variable: SKOR TOTAL PERILAKU MEROKOK [POST TEST]

## b. Output Model II

● **GAMBAR 11.11** Output Regresi Linear Ganda Model IICoefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.558	1.034		2.474	.014
	APAKAH DESA MENDAPAT INTERVENSI PENDIDIKAN	1.760	.243	.460	7.251	.000
	Umur dalam Tahun	.221	.110	.132	2.007	.046
	skor_sikap	-.012	.010	-.076	-1.185	.238
		.106	.030	.221	3.526	.001

a. Dependent Variable: SKOR TOTAL PERILAKU MEROKOK [POST TEST]

## c. Output Model III

● **GAMBAR 11.12** Output Regresi Linear Ganda Model IIICoefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.955	.901		2.169	.031
	APAKAH DESA MENDAPAT INTERVENSI PENDIDIKAN	1.772	.243	.463	7.294	.000
	skor_sikap	.255	.107	.152	2.385	.018
		.104	.030	.216	3.450	.001

a. Dependent Variable: SKOR TOTAL PERILAKU MEROKOK [POST TEST]





## BAB 12

# UJI CHI-SQUARE STATA & SPSS

### Kompetensi Dasar

Indikator Keberhasilan

Materi Pembelajaran

### Mampu menjelaskan analisis statistik uji *Chi-square* menggunakan Stata dan SPSS

- Mampu menjelaskan uji bivariat-uji hipotesis
  - Mampu menjelaskan uji Chi-square
  - Mampu menjelaskan uji Fisher Exact
  - Mampu menjelaskan aplikasi uji Chi-square menggunakan Stata dan SPSS
  - Mampu menjelaskan uji Chi-square dan Penggabungan Sel menggunakan Stata dan SPSS
1. Uji Bivariat-Uji Hipotesis
  2. Uji Chi-square
  3. Uji Fisher Exact
  4. Aplikasi Uji Chi-square
  5. Uji Chi-square dan Penggabungan Sel

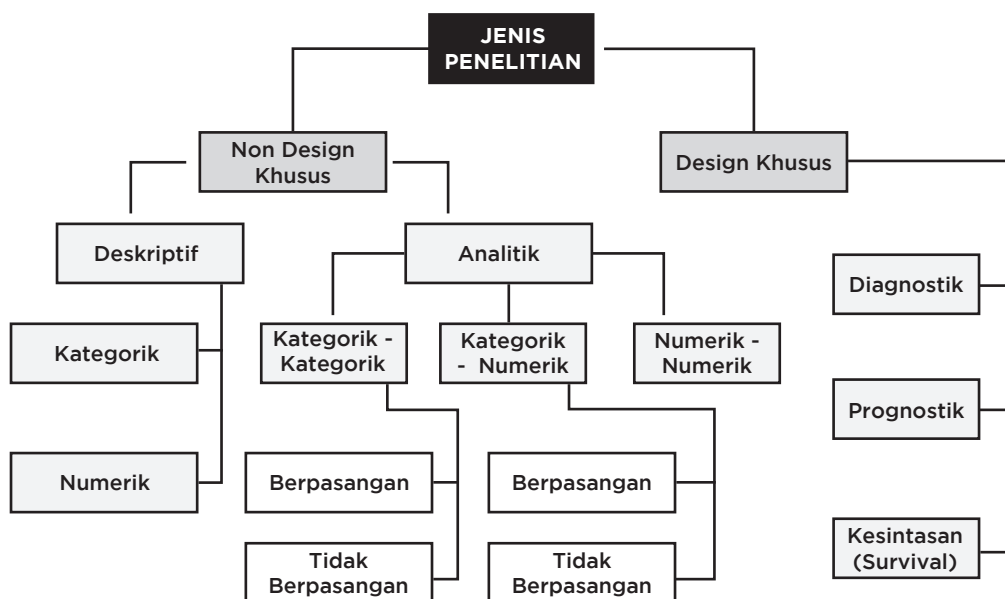
## UJI BIVARIAT–UJI HIPOTESIS

Prosedur Uji Hipotesis adalah sebagai berikut.

1. Menentukan  $H_0$  (*Null Hypothesis*) dan  $H_a$  (*Alternative Hypothesis*).
2. Menentukan tingkat kepercayaan, misal tingkat kepercayaan 95% atau tingkat signifikan (alfa) 5%.
3. Menentukan statistik hitung. Nilai statistik hitung tergantung pada metode statistik yang digunakan.
4. Mengambil keputusan. Keputusan terhadap hipotesis di atas ditentukan dengan membandingkan nilai statistik hitung dengan tingkat signifikan (alfa).

Untuk menentukan jenis uji yang akan kita gunakan dalam analisis statistik, alur pemikiran menuju hipotesis yang sesuai harus dipahami. Secara garis besar, analisis bivariat dalam penelitian ini adalah dengan menganalisis silang dua variabel, yaitu variabel independen dan variabel dependen. Bila nilai probabilitas (*p value*) kurang dari atau sama dengan alfa, berarti hasil perhitungan statistik bermakna (signifikan) dan apabila nilai *p value* lebih besar dari alfa; berarti hasil perhitungan statistik tidak bermakna (tidak signifikan). Berikut ini adalah berbagai uji statistik yang pada umumnya digunakan untuk analisis bivariat di bidang kesehatan.

● **GAMBAR 12.1** Jenis Penelitian secara Garis Besar



● **TABEL 12.1** Tabel Uji Statistik pada Analisis Bivariat-Uji Hipotesis yang Sering Digunakan di Bidang Kesehatan

VARIABEL I	VARIABEL II	UJI STATISTIK
Kategori	Kategori	Chi-square/Fisher Exact Regresi Logistik Sederhana
Kategori	Numerik	Uji T Anova
Numerik	Numerik	Korelasi Regresi Linear Sederhana

Independen	Variabel Dependen	
	Kategorik	Numerik
Kategorik	Chi-square/regresi logistik sederhana	1. T-test paired (sebelum-sesudah) 2. T-test independent (jika 2 kategori) 3. Anova (jika > 2 kategori)
Numerik	T-tes (jika 2 kategori) Anova (> 2 kategori)	Korelasi/regresi linear sederhana

Sumber:  
Besral. 2010. Analisis Data (Bivariat). Pelatihan Analisis Data Sekunder. Palembang, 29 November 2011. FKM UNSRI.  
Najmah. 2011. Manajemen dan Analisis Data Kesehatan, Kombinasi Teori dan Aplikasi SPSS. Nuha Medika:Yogyakarta.

## UJI CHI-SQUARE

Uji *Chi-square* atau yang juga disebut kaid kuadrat termasuk salah satu alat uji yang sering digunakan. Dalam statistik nonparametrik, uji Chi-Square satu sampel bisa digunakan untuk menguji apakah data sampel dapat menunjang hipotesis yang menyatakan bahwa populasi asal sampel mengikuti distribusi yang telah ditetapkan. Uji tersebut dinamakan *goodness of fit* (tes keselarasan), sebab pengujian dilakukan untuk mengetahui apakah sampel selaras dengan salah satu distribusi teoretis (distribusi normal, uniform, binomial, dan sebagainya). Namun pada praktiknya, uji ini tetap mengikuti prinsip pengujian Chi-Square, yakni membandingkan antara frekuensi harapan dengan frekuensi yang teramati. Uji Chi-Square hanya menyampaikan ada atau tidaknya hubungan antara variabel yang diteliti dan tidak memberikan informasi tentang besarnya tingkat kekuatan suatu hubungan.

Asumsi yang mendasari penggunaan Chi-Square selain pemilihan sampel secara acak adalah penggunaan sampel yang besar. Jika sampel yang digunakan ukurannya kecil akan menyebabkan frekuensi harapan menjadi sangat kecil, hal ini menyebabkan nilai perhitungan menjadi kurang efektif jika dibandingkan dengan nilai distribusi Chi-Square itu sendiri. Untuk uji Pearson Chi-Square mensyaratkan frekuensi harapan tidak boleh < 5 untuk tabel 2 × 2 dan < 10 untuk tabel lebih dari 2 × 2.

Syntax: tabulate [variabel], *chi2 row*

## UJI FISHER EXACT

Uji ini umumnya digunakan pada data yang menganalisis hubungan tabel  $2 \times 2$ , untuk kasus dengan jumlah data atau frekuensi sel yang sedikit akan memiliki nilai harapan  $< 5$  di mana distribusi akan berbeda dengan distribusi Chi-Square.

Syntax:

```
tabulate [var] [var], expected
```

```
tabulate [var] [var], exact
```

Uji Chi-square termasuk dalam uji hipotesis variabel kategorikal tidak berpasangan. Berikut ini merupakan diagram alur uji hipotesis variabel kategorikal dalam bentuk tabel silang  $B \times K$  untuk kelompok tidak berpasangan.

Catatan penting dari Gambar 12.2 adalah<sup>(3,5)</sup>:

1. Semua hipotesis untuk tabel  $B \times K$  tidak berpasangan **menggunakan Uji Chi-square** bila memenuhi syarat uji Chi-square!
2. Syarat uji *Chi-square* adalah:
  - Tidak ada sel yang nilai observed bernilai nol.
  - Sel yang mempunyai nilai *expected* kurang dari 5, maksimal 20% dari jumlah sel.
  - Nilai yang diambil "*continuity correction*".
3. Jika syarat uji *Chi-square* tidak terpenuhi, maka dipakai uji alternatifnya:
  - Alternatif uji *Chi-square* untuk tabel  $2 \times 2$  adalah **uji Fisher**
  - Alternatif uji *Chi-square* untuk tabel  $2 \times k$  adalah **uji Kolmogorov-Smirnov**
  - Penggabungan sel adalah langkah alternatif uji Chi-square untuk tabel selain  $2 \times 2$  dan  $2 \times k$  sehingga terbentuk suatu tabel  $B \times K$  yang baru. Setelah dilakukan **penggabungan sel**, uji hipotesis dipilih sesuai dengan tabel  $B \times K$  yang baru tersebut.

**Chi-Square Test**

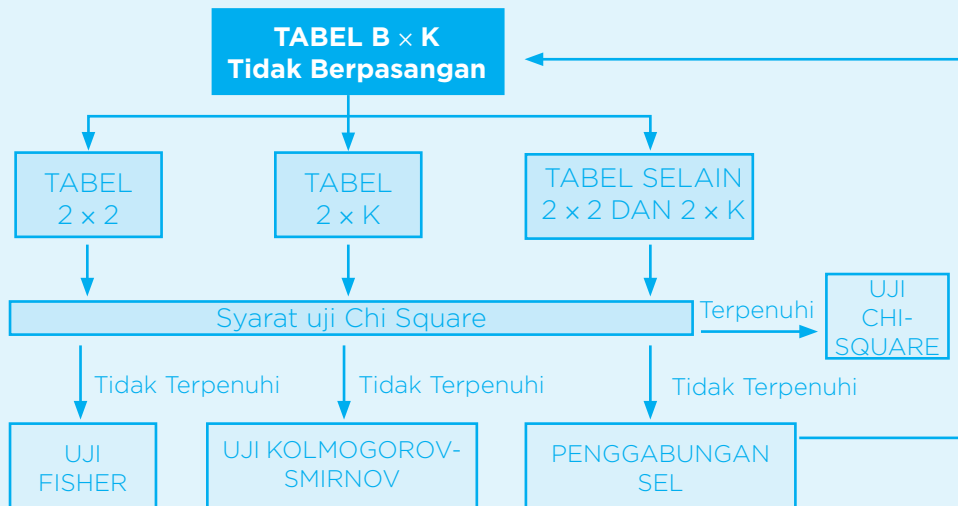
	Value	Df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Continuity Correction <sup>b</sup>					
Likelihood Ratio					
Fisher's Exact Test <sup>c</sup>					
Linear-by-Linear Association					
N of Valid Cases					

a. Pearson Chi Square, Nilai signifikansi (p value) untuk tabel  $2 \times k$ ,  $b \times k$  (selain tabel dari  $2 \times 2$ )

b. Continuity correction, Nilai signifikansi (p value) untuk tabel  $2 \times 2$ , Kai Kuadrat

c. Fisher's Exact test, Nilai signifikansi (p value) untuk tabel  $2 \times 2$ , Fisher Exact

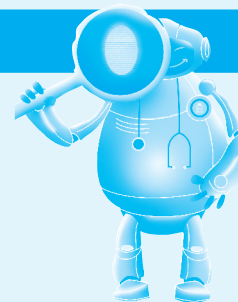
● **GAMBAR 12.2** Diagram Alur Uji Hipotesis Variabel Kategorikal Kelompok Tidak Berpasangan<sup>(3)</sup>



## APLIKASI Uji CHI-SQUARE

### STUDI KASUS 12.1

 Buka Data: WIDYAWATI 2015.sav



Sumber Data: Widyawati, 2015, Hubungan Pemberian Makanan Pendamping ASI dengan Status Gizi Kurang pada Anak Usia 12–24 Bulan di Wilayah Puskesmas Lesung Batu Kabupaten Empat Lawang Tahun 2015, Skripsi FKM Unsrj.

**Hipotesis:** Ada hubungan antara status gizi anak (1: kurus, 2: normal) (var no 11) dengan frekuensi MP ASI (1: jika kurang dari 3 kali, 2: sama dengan atau lebih dari 3 kali) (var no 22/C1) (*p value* <0.0001; OR 6,6)

Langkah-langkah untuk menentukan uji apakah yang mungkin digunakan untuk menjawab pertanyaan tersebut adalah sebagai berikut.

● **TABEL 12.3** Langkah-Langkah Penentuan Uji

No.	Langkah	Jawaban
1	Menentukan variabel yang diuji	Variabel yang diuji adalah Status Gizi Kurang (variabel dependen) dan Frekuensi MP ASI (variabel independen)

● **TABEL 12.3** Langkah-Langkah Penentuan Uji (Lanjutan)

No.	Langkah	Jawaban
2	Menentukan skala pengukuran variabel	Variabel Status Gizi Kurang merupakan variabel kategorikal (nominal) Variabel Frekuensi MP ASI merupakan variabel kategorikal (nominal)
3	Menentukan jenis hipotesis	Jenis hipotesis asosiatif
4	Menentukan jumlah kelompok	Jumlah kelompok yang diuji adalah 2 (Pemberian MP ASI Kurang dan Pemberian MP ASI Cukup)
5	Menentukan berpasangan atau tidak berpasangan	Pada kasus di atas, kedua kelompok tidak berpasangan
6	Menentukan Jenis Tabel	Jenis tabelnya adalah 2 x 2

### JAWABAN DETEKTIF



## LANGKAH-LANGKAH UJI CHI-SQUARE PADA STATA

Kita cukup memasukkan nilai tabel 2 x 2 di atas ke dalam command *syntax* cci (desain kasus kontrol), *csi* (*desain kohort* atau *experiment*).

. tab C1 kaskontrol

MPASI dalam sehari	status kasus kontrol		Total
	0 kasus (gi)		
kurang (<3 kali)	8	25	33
cukuo (>=3 kali)	32	15	47
Total	40	40	80

. cci 25 8 15 32

	Exposed	Unexposed	Total	Proportion Exposed
Cases	25	8	33	0.7576
Controls	15	32	47	0.3191
Total	40	40	80	0.5000
	Point estimate		[95% Conf. Interval]	
Odds ratio	6.666667		2.212315	20.89429 (exact)
Attr. frac. ex.	.85		.5479848	.95214 (exact)
Attr. frac. pop	.6439394			

chi2(1) = 14.91 Pr>chi2 = 0.0001

Atau memasukkan variabel langsung pada data:

1. Untuk mengetahui nilai *expected*, apakah ada di bawah 5 atau tidak. Berdasarkan hasil di bawah ini nilai *expected* memenuhi untuk uji *Chi-square*

**Command Syntax:** *tabulate outcome exposure, expected*

Tabulate outcome exposure, exact → jika uji *Chi-square* tidak terpenuhi, maka *fisher exact*

sebagai alternatif.

```
. tabulate statusgizi C1, expected
```

Key
<i>frequency</i>
<i>expected frequency</i>

status gizi sampel berdasarkan BB/TB	MPASi dalam sehari		Total
	kurang (<)	cukuo (>=)	
Normal	8 16.5	32 23.5	40 40.0
Kurus	25 16.5	15 23.5	40 40.0
Total	33 33.0	47 47.0	80 80.0

2. Melanjutkan Uji *Chi-square*

**Command Syntax:** *tabulate outcome eksposur, chi2 row*

```
. tabulate statusgizi C1, chi2 row
```

Key
<i>frequency</i>
<i>row percentage</i>

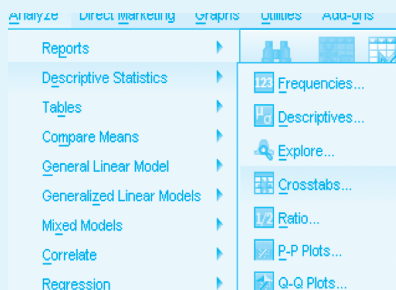
status gizi sampel berdasarkan BB/TB	MPASi dalam sehari		Total
	kurang (<)	cukuo (>=)	
Normal	8 20.00	32 80.00	40 100.00
Kurus	25 62.50	15 37.50	40 100.00
Total	33 41.25	47 58.75	80 100.00

Pearson chi2(1) = 14.9065 Pr = 0.000

LANGKAH-LANGKAH UJI CHI-SQUARE PADA PROGRAM SPSS

1. Klik *Analyze* → *Descriptive Statistics* → *Crosstabs*

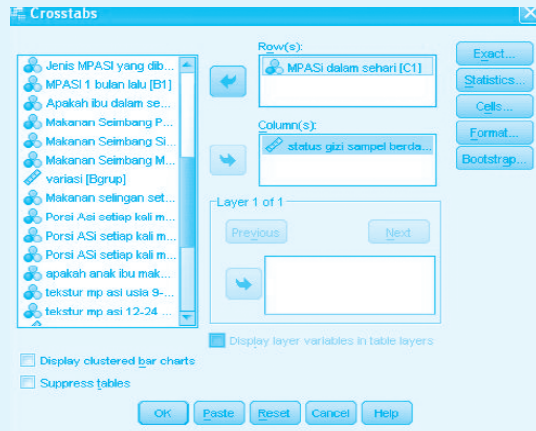
● **GAMBAR 12.3** Proses Analisis *Chi-Square*



2. Masukkan Variabel Independen (Frekuensi Pemberian MP ASI) ke dalam kolom 'ROWS'

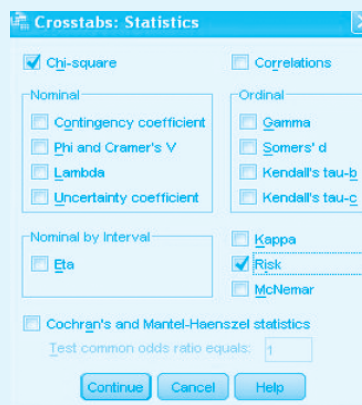
dan Variabel Dependen (Status Gizi Kurang) ke 'COLUMN'

### ● GAMBAR 12.4 Tampilan "Crosstabs"



3. Klik *Statistics* → *Chi-square* → *Continue* (untuk tabel  $2 \times 2$ , *risk ratio/odds ratio* bisa dihasilkan dengan klik *risk*)

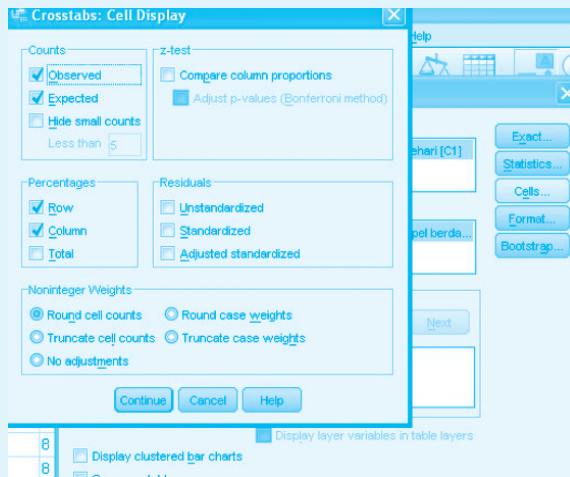
### ● GAMBAR 12.5 Tampilan Kolom *Statistics* pada "Crosstabs"



4. Klik *Cells*, lalu klik *Observed*, *Expected*, *Rows*, dan *Column* dan klik *Continue*.



● **GAMBAR 12.6** Tampilan Kolom Cells pada “Crosstabs”



5. Lalu klik OK.

**OUTPUT SPSS**

● **TABEL 12.4** Output Cross Tabulation MPASI dalam Sehari dan Status Gizi Sampel Berdasarkan BB/TB

**MPASI dalam sehari \* status gizi sampel berdasarkan BB/TB Crosstabulation**

		Status gizi sampel berdasarkan BB/TB		Total		
		Kurus	Normal			
<b>MPASI dalam sehari</b>	<b>kurang (&lt; 3 kali)</b>	Count	25	8	33	
		Expected Count	16,5	16,5	33,0	
		% within MPASi dalam sehari	75,8%	24,2%	100,0%	
		% within status gizi sampel berdasarkan BB/TB	62,5%	20,0%	41,2%	
	<b>cukup (≥ 3 kali)</b>	Count	15	32	47	
		Expected Count	23,5	23,5	47,0	
		% within MPASi dalam sehari	31,9%	68,1%	100,0%	
		% within status gizi sampel berdasarkan BB/TB	37,5%	80,0%	58,8%	
	<b>Total</b>		Count	40	40	80
			Expected Count	40,0	80,0	
		% within MPASi dalam sehari	50,0%	100,0%		
		% within status gizi sampel berdasarkan BB/TB	100,0%	100,0%		

● **TABEL 12.4** Output Cross Tabulation MPASI dalam Sehari dan Status Gizi Sampel Berdasarkan BB/TB (lanjutan)

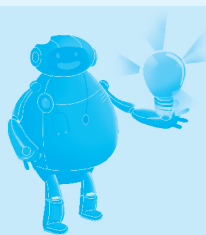
Chi-Square Tests					
	Value	Df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	14,907 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	13,204	1	,000		
Likelihood Ratio	15,484	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	14,720	1	,000		
N of Valid Cases	80				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 16.50.

b. Computed only for a 2x2 table

#### Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for MPASi dalam sehari (kurang (< 3 kali)/cukup (≥ 3 kali))	6,667	2,440	18,212
For cohort status gizi sampel berdasarkan BB/TB = Kurus	2,374	1,498	3,760
For cohort status gizi sampel berdasarkan BB/TB = Normal	,356	,189	,671
N of Valid Cases	80		



#### CATATAN:

*Hasil Asosiasi (Risk Estimate)*

*Kolom OR for MP ASI, hasil OR untuk studi desain kasus Kontrol*

*\* OR= Odds pada kasus (kurus)/Odds pada control(normal)=(25/8):*

*(15:32)=6,667 Kolom For Cohort status gizi sampel, hasil PR atau RR untuk studi desain kohort atau potong lintang*

- \* *kurus terhadap normal*
- \* *PR/RR= Risk pada kelompok terpapar/Risk pada kelompok tidak terpapar = (25/33) : (15/47) = 2,374 (MPASI kurang meningkatkan risiko untuk kurus sebesar 2,3 kali)*
- \* *normal terhadap kurus*
- \* *PR/RR= Risk pada kelompok terpapar/Risk pada kelompok tidak terpapar =(8/33) : (32/347) = 0,356 (MPASI kurang menurunkan risiko gizi normal sebesar 0,35 atau 65%)*

**LAPORAN HASIL**

● **TABEL 12.5** Laporan Hasil MPASI

Variabel	OR	IK 95%		p value
		Min	Maks	
<b>Frekuensi MP ASI</b>				
< 3 kali	<i>Ref</i>	2,4	18,2	< 0.0001
≥ 3 kali	6,6			

**INTERPRETASI**

**Hubungan antara Frekuensi Pemberian MP ASI dan Status Gizi Anak**

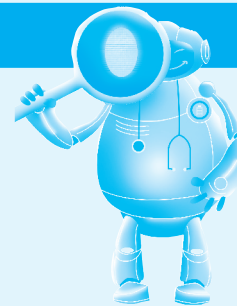
Anak usia 12–24 bulan dengan frekuensi pemberian MPASI kurang dari 3 kali meningkatkan risiko kejadian gizi kurang sebesar 6,6 kali dibandingkan dengan anak dengan frekuensi pemberian MP ASI lebih atau sama dengan 3 kali dalam sehari. Di populasi, anak dengan frekuensi pemberian MP ASI kurang 3 kali dalam 1 hari meningkatkan risiko kejadian gizi kurang berkisar 2,4 hingga 18,2 kali lebih tinggi dibandingkan dengan anak usia 12–24 bulan dengan frekuensi pemberian MP ASI 3 kali atau lebih dalam sehari (OR=6,6, 95% IK: 2,4–18,2). Kesimpulannya, berdasarkan nilai p value <0.0001 menunjukkan ada hubungan yang signifikan antara frekuensi pemberian MP ASI dan status gizi anak.

# CHI-SQUARE DAN PENGGABUNGAN SEL

UJI HIPOTESIS TABEL B × K SELAIN 2 × 2 DAN 2 × K

**STUDI KASUS 12.1**

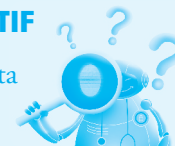
Anda ingin mengetahui hubungan antara tingkat pengetahuan (rendah, sedang, tinggi) dengan intake makanan (kurang, cukup, lebih). Anda membuat pertanyaan: “Apakah ada hubungan antara tingkat pengetahuan (rendah, sedang, tinggi) dengan intake makanan (kurang, cukup, lebih)?”



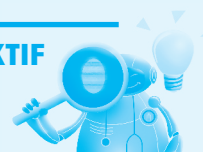
 Buka data dari bentuk data excel ke dalam lembar kerja SPSS (**Intake & Tahu\_Sopiyudin D.xls**)

**PERTANYAAN DETEKTIF**

Lakukan uji Chi-square lalu perhatikan nilai *expected*-nya, apakah bisa langsung kita gunakan signifikansi hasil Chi-square dari tabel di bawah ini? Lakukan analisis lebih lanjut!



## JAWABAN DETEKTIF



● **TABEL 12.6** Hubungan Tingkat Pengetahuan dan Intake Makanan

			Intake kalori			Total
			kurang	cukup	lebih	
Tingkat pengetahuan	rendah	Count	11	29	1	41
		Expected Count	9.8	19.7	11.5	41.0
		% within tingkat pengetahuan	26.8%	70.7%	2.4%	100.0%
	sedang	Count	12	19	27	58
		Expected Count	13.9	27.8	16.2	58.0
		% within tingkat pengetahuan	20.7%	32.8%	46.6%	100.0%
	tinggi	Count	1	0	0	1
		Expected Count	.2	.5	.3	1.0
		% within tingkat pengetahuan	100.0%	.0%	.0%	100.0%
Total		Count	24	48	28	100
		Expected Count	24.0	48.0	28.0	100.0
		% within tingkat pengetahuan	24.0%	48.0%	28.0%	100.0%

Sumber: Dahlan S. *Statistika untuk Kedokteran dan Kesehatan*. Jakarta: PT Arkas; 2004.

## Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	27.485	4	.000
Likelihood Ratio	32.283	4	.000
Linear-by-Linear Association	8.253	1	.004
N of Valid Cases	100		

a. 3 cells (33.3%) have expected count less than 5. The minimum expected count is .24.

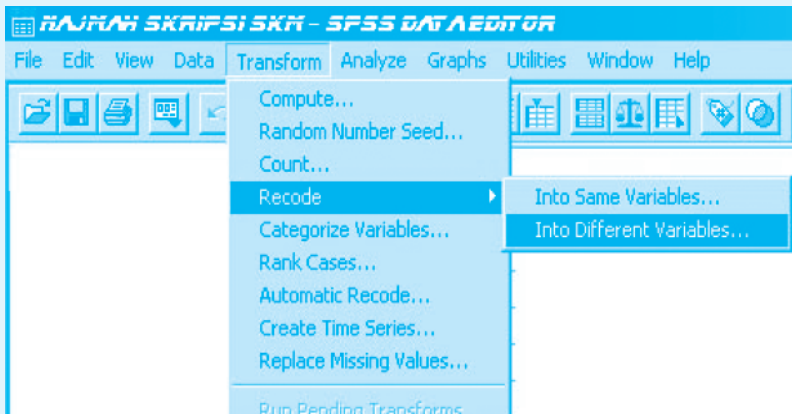
## Interpretasi

Tabel  $3 \times 3$  ini tidak layak untuk diuji dengan *chi square* karena sel yang nilai *expected* kurang dari lima ada 33,3% jumlah sel. Selain itu terdapat sel dengan nilai *observed* nol. Langkah selanjutnya adalah melakukan penggabungan sel. Anda memutuskan untuk menggabungkan kelompok pengetahuan tinggi dengan kelompok pengetahuan sedang. Alasan Anda menggabungkan kedua kelompok karena jumlah subjek yang termasuk ke dalam kelompok pengetahuan tinggi sedikit (satu subjek) sehingga digabung dengan kelompok subjek dengan pengetahuan sedang.

Lakukan transformasi data, langkah-langkahnya berikut ini:

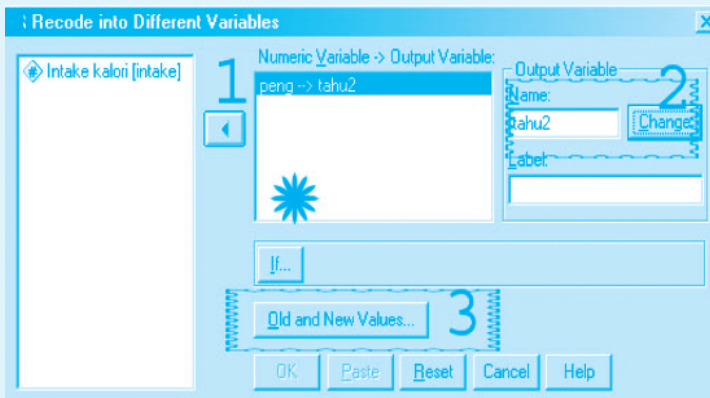
- Transform.....Recode.....*Recode into different variable*

● **GAMBAR 12.7** Proses Pengkodean Variabel Baru



- Masukkan tahu\_2 ke dalam output variabel

● **GAMBAR 12.8** Kotak Dialog “Recode into Different Variables”



- Klik kotak change
- Klik old and values
- Isilah kotak old value dan kotak new values (selanjutnya ikuti logika berpikir)

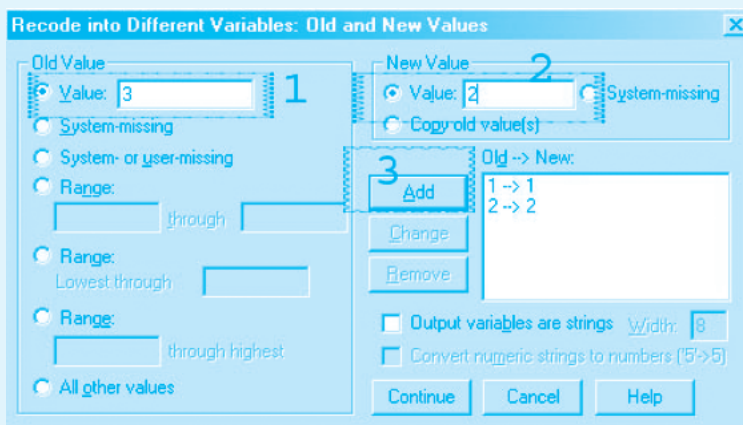
Logikanya adalah:

Kode 1 (old value), diubah menjadi kode 1 (new value)

Kode 2 (old value), diubah menjadi kode 2 (new value)

Kode 3 (old value), diubah menjadi kode 2 (new value)

● **GAMBAR 12.9** Kotak Dialog “Recode into Different Variables: Old and New Values”



- Sampai tahap ini, Anda akan memperoleh tampilan seperti Gambar 12.9.
  - Proses telah selesai, klik *continue*.
  - OK, dan lihat hasilnya ada variabel baru 'tahu2'.
  - Lakukan Uji Chi-Square kembali dengan variabel pengetahuan dengan dua kategori (tahu2) seperti latihan sebelumnya.
- Output hasil

TAHU2 \* Intake kalori Crosstabulation

		Intake kalori			Total	
		kurang	cukup	lebih		
TAHU2	rendah	Count	11	29	1	41
		Expected Count	9.8	19.7	11.5	41.0
		% within TAHU2	26.8%	70.7%	2.4%	100.0%
	sedang+	Count	13	19	27	59
		Expected Count	14.2	28.3	16.5	59.0
		% within TAHU2	22.0%	32.2%	45.8%	100.0%
Total	Count	24	48	28	100	
	Expected Count	24.0	48.0	28.0	100.0	
	% within TAHU2	24.0%	48.0%	28.0%	100.0%	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	23.928	2	.000
Likelihood Ratio	29.196	2	.000
Linear-by-Linear Association	10.696	1	.001
N of Valid Cases	100		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.84.

**Interpretasi:**

Tabel  $2 \times 3$  ini layak untuk diuji dengan Chi-Square karena tak ada sel yang bernilai 0, dan tidak ada nilai *expected* yang kurang dari 5. Nilai *significancy*-nya adalah  $< 0,001$ , menunjukkan ada hubungan antara pengetahuan dengan intake makanan atau *p value*  $< 0,001$  menunjukkan adanya bukti yang kuat untuk menolak hipotesis nul, tidak ada hubungan antara pengetahuan dengan intake makanan.

**Sumber:**

Dahlan S. 2004. *Statistika untuk Kedokteran dan Kesehatan*. Jakarta: PT Arkas.

Najmah. 2011. *Managemen dan Analisa Data Kesehatan. Kombinasi Teori dan Aplikasi SPSS*. Yogyakarta: Nuha Medika.







## BAB 13

# UJI INDEPENDENT STUDENT T-TEST STATA & SPSS

### Kompetensi Dasar

Mampu menjelaskan analisis statistik uji independent student T-Test menggunakan Stata dan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan uji Independent Student T-Test.
- Mampu menjelaskan aplikasi uji Independent Student T-Test menggunakan Stata dan SPSS.

### Materi Pembelajaran

1. Uji Beda Rata-Rata Tidak Berpasangan (*Independent Student T-Test*).
2. Aplikasi Uji Independent Student T-Test.

# UJI INDEPENDENT STUDENT T-TEST

## UJI BEDA RATA-RATA TIDAK BERPASANGAN (UJI T INDEPENDEN/INDEPENDENT SAMPLE T-TEST)

Standar deviasi diperoleh dari nilai varians gabungan dua kelompok sampel yang akan diuji. Ada dua kemungkinan, yakni varians sama dan varians yang berbeda. Dua kemungkinan nilai varians ini melahirkan dua jenis penghitungan nilai standar deviasi gabungan yang digunakan dalam pengujian dan dua jenis dari penghitungan *degree of freedom* (df) yang berbeda.

Untuk melakukan pengujian apakah varians sama atau berbeda, maka dilakukan uji rasio nilai varians dua kelompok tersebut. Hasil uji rasio dua nilai varians tersebut menyebar mengikuti distribusi F (Fisher).

Syntax :

ttest [var numerik], by [var kategorik] → varian homogen

ttest [var numerik], by [var kategorik] unequal → varian tidak homogen

Prosedur ini digunakan untuk membandingkan rata-rata sampel independen dengan menghitung *Student T-Test* dan menampilkan probabilitas dua arah selisih dua rata-rata <sup>(2)</sup>.

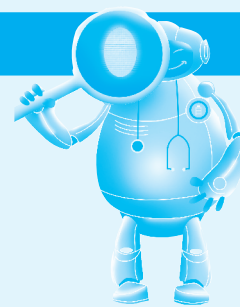
## APLIKASI UJI STUDENT (T TEST) TIDAK BERPASANGAN (INDEPENDENT T-TEST)

### STUDI KASUS 13.1

**Hipotesis 1:** Ada hubungan antara berat badan anak (kg) (BB anak) dengan status gizi anak (1: kurus, 2: normal) (var no 22/C1). Kita asumsikan berat badan anak berdistribusi normal.

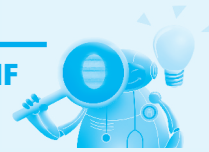


Buka Data: WIDYAWATI 2015.sav



Sumber Data: Widyawati, 2015. Hubungan Pemberian Makanan Pendamping ASI dengan Status Gizi Kurang pada Anak Usia 12–24 Bulan di Wilayah Puskesmas Lesung Batu Kabupaten Empat Lawang Tahun 2015, Skripsi FKM Unsri.

**JAWABAN DETEKTIF**



*Langkah-Langkah pada Aplikasi Stata*

**Command Syntax:**

ttest [var numerik], by [var kategorik] → varian homogen

ttest [var numerik], by [var kategorik] unequal → varian tidak homogen

- Masukkan variabel BB anak dan status gizi ke dalam *command syntax*

```
. ttest bbanak, by(statusgizi)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Normal	40	10.1225	.2625195	1.660319	9.591504	10.6535
Kurus	40	8.2725	.132529	.8381871	8.004435	8.540565
combined	80	9.1975	.1793799	1.604422	8.840453	9.554547
diff		1.85	.2940756		1.26454	2.43546

```

diff = mean(Normal) - mean(Kurus)          t = 6.2909
Ho: diff = 0                               degrees of freedom = 78

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 1.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 0.0000

. ttest bbanak, by(statusgizi) unequal

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Normal	40	10.1225	.2625195	1.660319	9.591504	10.6535
Kurus	40	8.2725	.132529	.8381871	8.004435	8.540565
combined	80	9.1975	.1793799	1.604422	8.840453	9.554547
diff		1.85	.2940756		1.261271	2.438729

```

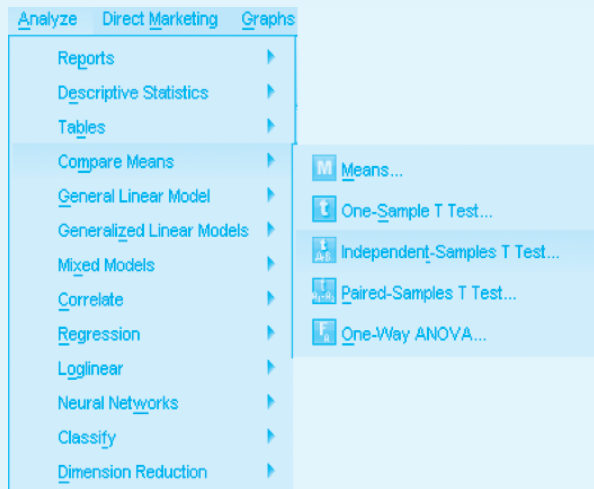
diff = mean(Normal) - mean(Kurus)          t = 6.2909
Ho: diff = 0                               Satterthwaite's degrees of freedom = 57.6665

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 1.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 0.0000
```

*Langkah-Langkah pada Aplikasi SPSS*

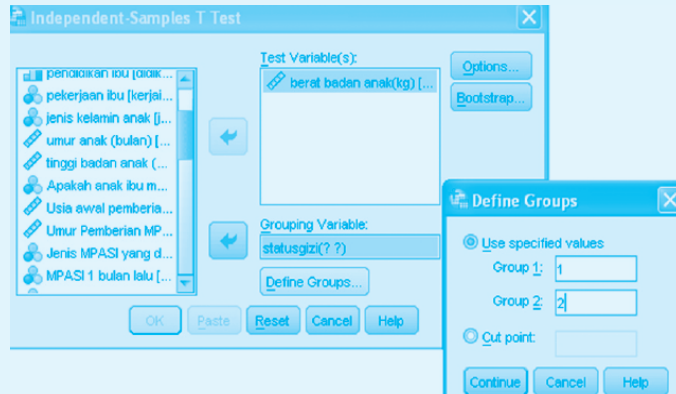
1. Klik *Analyze Compare Means Independent Sample T-Test*

**GAMBAR 13.1** Proses Analisis T-Test



- Masukkan variabel Berat Badan Anak ke dalam *Test Variable (s)* dan Status Gizi ke dalam *Grouping Variable Define Groups* masukkan kode Status Gizi yaitu 1 (Kurus) dan 2 (Normal) *Continue*.

● **GAMBAR 13.2** Proses Pemilihan Independent-Sample T-Test



● **GAMBAR 13.3** Output Data Independent Samples T-Test BB Anak dan Status Gizi

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
berat badan anak(kg)	Equal variances assumed	1,251	,267	-6,291	78	,000	-1,8500	,2941	-2,4355	-1,2645
	Equal variances not assumed			-6,291	57,667	,000	-1,8500	,2941	-2,4387	-1,2613

Group Statistics				
status gizi sampel berdasarkan BB/TB	N	Mean	Std. Deviation	Std. Error Mean
berat badan anak(kg) Kurus	40	8,273	,8382	,1325
Normal	40	10,122	1,6603	,2625

a. Menguji Varians

Pada kotak uji Lavene (nama uji hipotesis untuk menguji varians), nilai  $p = 0,267$ . Oleh karena nilai  $p > 0,05$ , maka varians data kedua kelompok sama (terima  $H_0$ ), tetapi yang perlu diingat adalah kesamaan varians tidak menjadi syarat mutlak untuk dua kelompok tidak berpasangan. Oleh karena varians sama, hasil uji T yang dilihat pada baris pertama (*equal variances assumed*).

*Uji Lavene*

$H_0$  : varians diasumsikan sama

$H_a$  : varians diasumsikan berbeda

b. Tampilan laporan

● **TABEL 13.1** Hubungan Berat Badan Anak dan Status Gizi Kurang

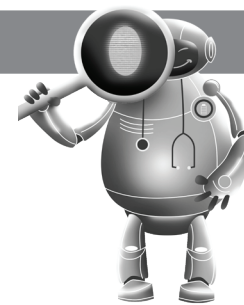
Gizi Kurang	Perbedaan Rata-rata	95% Derajat Kepercayaan	Nilai P
Berat badan anak (kg)	-1,85	(-2,43) hingga (-1,26)	<0,0001

c. Interpretasi

Perbedaan rata-rata (*mean difference*) = -1,85, mengindikasikan rata-rata berat badan anak pada kelompok kurus lebih rendah 1,85 kg dibandingkan dengan rata-rata berat badan anak pada kelompok normal. Di populasi umum, tingkat kepercayaan 95% mengindikasikan bahwa berat badan anak dapat menjadi faktor penyebab terjadinya status gizi kurang, dengan kisaran rata-rata berat badan 1,26–2,43 kg lebih rendah pada kelompok gizi anak kurang (kurus) dibandingkan status gizi normal. Kesimpulannya berdasarkan nilai signifikansi, *P value* = < 0,0001 (95% CI -2,43–1,26), menunjukkan kuatnya signifikansi untuk menolak hipotesis nul ( $H_0$ ), dengan kata lain “adanya perbedaan rata-rata berat badan anak pada status gizi anak yang kurus dan normal atau ada perbedaan signifikan antara berat badan anak dan status gizi kurang”.

### STUDI KASUS 13.2

**Hipotesis 2:** Ada hubungan antara usia pemberian MPASI pertama (bulan) dengan status gizi anak (1: kurus, 2: normal) (var no 22/C1) (*p value* <0.0001; OR 6,6). Kita asumsikan usia pemberian MPASI pertama berdistribusi normal.



 Buka Data: *wiwid\_kasuskontrol all1.sav*

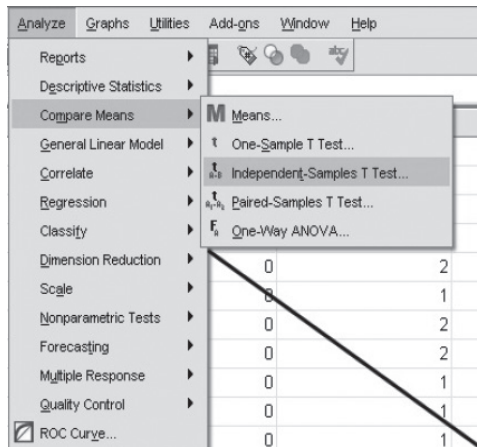
*Sumber Data: Widyawati. 2015. "Hubungan Pemberian Makanan Pendamping ASI dengan Status Gizi Kurang pada Anak Usia 12–24 Bulan di Wilayah Puskesmas Lesung Batu Kabupaten Empat Lawang Tahun 2015". Skripsi FKM Unsri.*

#### JAWABAN DETEKTIF

1. Lanjutkan uji Student T-test, Klik *Analyze Compare Means Independent Samples T-test*



● **GAMBAR 13.4** Proses Pemilihan Independent Samples T-Test



2. Masukkan variabel "*Usia Awal Pemberian MP ASI*" ke kolom "test variabel" dan variabel Status Gizi ke kolom *grouping* variabel, lalu klik *Define Groups*, masukkan kode Status Gizi yaitu 1 (Kurus) dan 2 (Normal)
3. Klik Ok

a. **Output SPSS**

● **GAMBAR 13.5** Output Data Independent Samples T-Test Usia Awal Pemberian MPASI dan Status Gizi

Group Statistics					
status gizi sampel berdasarkan BB/TB		N	Mean	Std. Deviation	Std. Error Mean
Usia awal pemberian MPASI (bulan)	Kurus	40	5.03	1.050	.166
	Normal	40	5.15	.949	.150

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Usia awal pemberian MPASI (bulan)	Equal variances assumed	1.245	.268	-.559	78	.578	-.125	.224	-570	.320
	Equal variances not assumed			-.559	77.214	.578	-.125	.224	-570	.320

b. **Laporan dan Interpretasi**

Lengkapi laporan tabel di bawah ini berdasarkan data *output* SPSS di atas!

● **TABEL 13.2** Hubungan Usia Anak dan Status Gizi Kurang

Gizi kurang	Perbedaan rata-rata	95% Derajat Kepercayaan	Nilai P
Usia badan anak (bulan)	.....	.....	.....

**c. Menguji Varians**

Pada kotak uji Levene (nama uji hipotesis untuk menguji varians), nilai  $p = 0,268$ .

● **GAMBAR 13.6** Uji Normalitas Kolmogorov-Smirnov Usia Awal Pemberian MPASI

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Usia awal pemberian MPASI (bulan)	.295	80	.000	.793	80	.000

a. Lilliefors Significance Correction

Oleh karena nilai  $p > 0,05$  maka varians data kedua kelompok sama (Terima  $H_0$ ). Tetapi hal yang perlu diingat, kesamaan varians tidak menjadi syarat mutlak untuk dua kelompok tidak berpasangan. Oleh karena varians sama, hasil uji T yang dilihat pada baris pertama (*equal variances assumed*).

Perbedaan rata-rata (*mean difference*) = -0,125, mengindikasikan rata-rata usia awal pemberian MPASI pada kelompok kurus lebih rendah 0,125 bulan dibandingkan rata-rata usia awal pemberian MPASI pada kelompok normal. Di populasi umum, tingkat kepercayaan 95% mengindikasikan bahwa perbedaan rata-rata Usia Awal Pemberian MPASI berada dalam rentang -0,57 (lebih rendah 0,57 bulan pada kelompok kurus) dan 0,320 (lebih tinggi 0,320 bulan pada kelompok kurus). Kesimpulan berdasarkan nilai signifikansinya, P value = 0,578 (95% CI -0,57; 0,320), menunjukkan lemahnya kekuatan signifikansi untuk menolak hipotesis nul ( $H_0$ ), tidak ada perbedaan Usia Awal Pemberian MPASI dengan Status Gizi.







## BAB 14

# UJI ANOVA STATA & SPSS

### Kompetensi Dasar

Mampu menjelaskan analisis statistik uji Anova menggunakan Stata dan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan Uji Anova.
- Mampu menjelaskan aplikasi Uji Anova menggunakan Stata dan SPSS.

### Materi Pembelajaran

1. Uji Anova
2. Aplikasi Uji Anova Pada Stata dan SPSS

## UJI ANOVA

Uji Anova adalah uji statistik inferensial parametrik yang digunakan peneliti untuk membandingkan dua atau lebih *mean* dari kelompok. Laporan hasil meliputi skor F dan tingkat probabilitas. Anova dan uji F memiliki tujuan yang sama yakni menguji rata-rata populasi, tetapi pada Anova, diuji lebih dari dua rata-rata populasi, sedangkan uji F bertujuan untuk menguji sama atau tidaknya varians.

Pengujian rata-rata pada lebih dari dua kelompok sampel digunakan One-Way Anova. Pada uji Anova memiliki syarat beberapa asumsi yang harus dipenuhi yakni normalitas data antarkelompok dan homogenitas varian. Uji normalitas data antarkelompok yang sering digunakan adalah uji Shapiro Wilk, sedangkan pada uji homogenitas varian One-Way Anova digunakan adalah uji Barlett's. Jika pada uji Anova hasilnya menolak  $H_0$  (nilai *p value* < 0,05), maka kita lanjutkan ke uji perbandingan berganda. Ada beberapa uji berganda salah satunya adalah uji Benferoni.

Asumsi yang digunakan pada uji Anova:

1. Populasi yang akan diuji memiliki distribusi normal.
2. Varians dari populasi tersebut sama.
3. Sampel tidak berhubungan satu sama lain.

## APLIKASI UJI ANOVA PADA SPSS

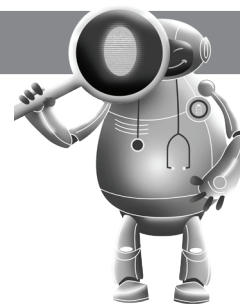
### STUDI KASUS 14.1

**Hipotesis:** Ada hubungan antara berat badan anak (kg) (var no 11) dengan tingkat pendidikan ibu (1: tidak tamat SD dan tamat SD, 2: SMP, 3 SMA/PT) (*p value* < 0.0001; OR 6,6).

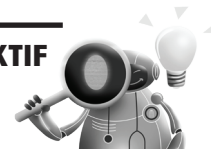
 Buka Data: WIDYAWATI 2015.sav

**Sumber Data:** Widyawati. 2015. "Hubungan Pemberian Makanan Pendamping ASI dengan Status Gizi Kurang pada Anak Usia 12–24 Bulan di Wilayah Puskesmas Lesung Batu Kabupaten Empat Lawang Tahun 2015". Skripsi FKM Unsri.

**Hipotesis alternatif:** ada perbedaan berat badan anak terhadap tingkat pendidikan ibu di wilayah Puskesmas Lesung Batu Kabupaten Empat Lawang Tahun 2015.



**JAWABAN DETEKTIF**



**LANGKAH-LANGKAH DENGAN STATA**

**Command syntax:**

Anova varnumerik varkategori

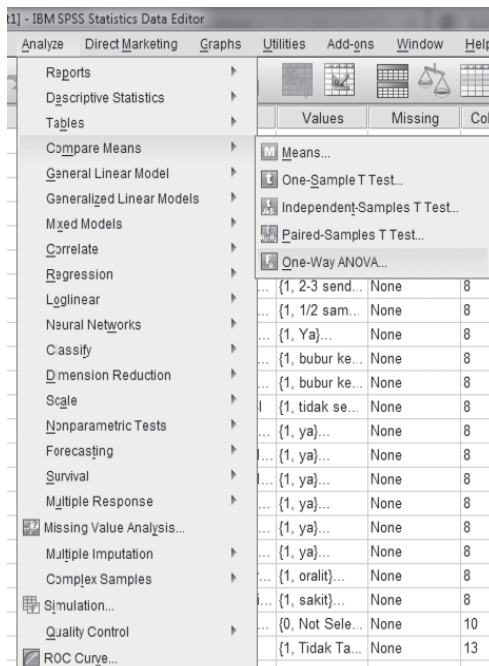
. anova bbanak didikibu

	Number of obs =	80	R-squared =	0.0056	
	Root MSE =	1.64206	Adj R-squared =	-0.0475	
Source	Partial SS	df	MS	F	Prob > F
Model	1.13335621	4	.283339052	0.11	0.9804
didikibu	1.13335621	4	.283339052	0.11	0.9804
Residual	202.226144	75	2.69634858		
Total	203.3595	79	2.57417089		

**LANGKAH-LANGKAH Uji ANOVA DENGAN SPSS**

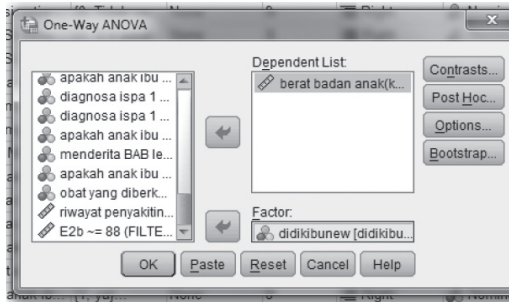
1. Lakukan Uji Normalitas terhadap variabel numerik, Berat Badan Anak dan Uji Varians. Kita asumsikan data yang kita miliki mempunyai distribusi yang normal dan varians data yang sama.
2. Lakukan Uji Anova  
*Analyze* → *Compare Means* → *One Way Anova*

**● GAMBAR 14.1** Proses Compare Means One Way Anova



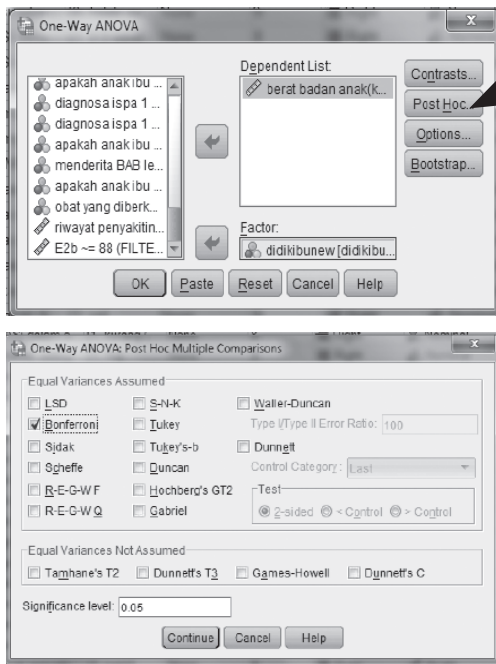
3. Kotak *Dependent List* diisi variabel numerik Berat Badan Anak dan Kotak *Factor* diisi variabel kategori Pendidikan Ibu.

● **GAMBAR 14.2** Kotak Dialog One Way Anova



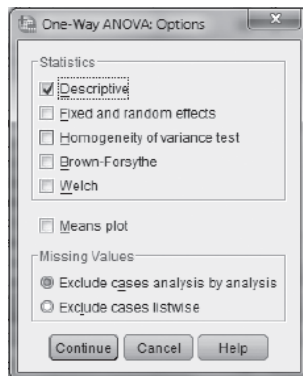
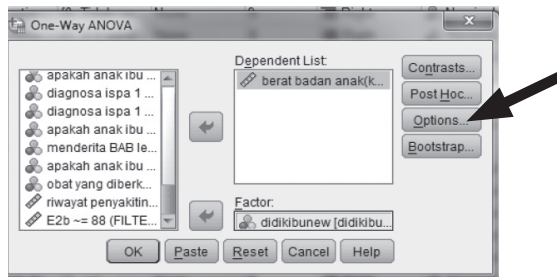
4. Klik *Post Hoc* → *Bonferroni* → *Continue*.

● **GAMBAR 14.3** Kotak Dialog One Way Anova Bonferroni



5. Klik *Options* → *Descriptive* → *Continue* → *OK*.

● **GAMBAR 14.4** Kotak Dialog One Way Anova Options



6. Output SPSS

● **GAMBAR 14.5** Output SPSS Uji Anova

**Descriptives**

berat badan anak(kg)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Tidak Tamat SD & Tamat SD	26	9.123	1.0923	.2142	8.682	9.564	7.2	11.5
Tamat SMP	21	9.333	2.3303	.5085	8.273	10.394	7.8	19.0
Tamat SMA/PT	33	9.170	1.4183	.2469	8.667	9.673	4.7	12.0
Total	80	9.198	1.6044	.1794	8.840	9.555	4.7	19.0

**ANOVA**

berat badan anak(kg)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.557	2	.278	.106	.900
Within Groups	202.803	77	2.634		
Total	203.359	79			

## Multiple Comparisons

Dependent Variable: berat badan anak(kg)

Bonferroni

(I) didikibunew	(J) didikibunew	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Tidak Tamat SD & Tamat SD	Tamat SMP	-.2103	.4762	1.000	-1.376	.955
	Tamat SMA/PT	-.0466	.4256	1.000	-1.088	.995
Tamat SMP	Tidak Tamat SD & Tamat SD	.2103	.4762	1.000	-.955	1.376
	Tamat SMA/PT	.1636	.4530	1.000	-.945	1.272
Tamat SMA/PT	Tidak Tamat SD & Tamat SD	.0466	.4256	1.000	-.995	1.088
	Tamat SMP	-.1636	.4530	1.000	-1.272	.945

## 7. Interpretasi

- a. Nilai Mean dan Standar Deviasi dari setiap kelompok. Rata-rata Berat Badan Anak dengan Pendidikan Ibu Tidak Tamat SD & Tamat SD adalah 9,123 kg dengan standar deviasi 1,0923 kg, pada Ibu dengan Pendidikan Tamat SMP adalah 9,333 kg dengan standar deviasi 2,3303 kg dan Ibu dengan Pendidikan Tamat SMA/PT adalah 9,170 kg dengan standar deviasi 1,4183 kg.
- b. Uji Anova
  - $p \text{ value (sig)} < \alpha = H_0$  ditolak, berarti ada perbedaan antara berat badan anak pada kelompok pendidikan ibu.
  - $p \text{ value (sig)} > \alpha = H_0$  diterima, berarti tidak ada perbedaan antara berat badan anak pada kelompok pendidikan ibu.
- c. Pada tabel (Gambar 14.4) diperoleh nilai Sig = 0,900 > alpha 0,05, berarti dapat disimpulkan bahwa ada bukti yang lemah untuk menolak hipotesis nul ( $H_0$ ) bahwa tidak ada perbedaan antara berat badan anak pada kelompok pendidikan ibu. Derajat kepercayaan 95% menunjukkan interval yang berkisar antara – dan +, berarti berat badan anak di populasi pada Pendidikan Ibu Tamat SMP dan Tamat SMA/PT dapat lebih rendah atau lebih tinggi dari kelompok ibu dengan pendidikan Tidak Tamat SD & Tamat SD.



## BAB 15

# PAIRED T-TEST

### Kompetensi Dasar

Mampu menjelaskan analisis statistik *Paired T-Test* menggunakan SPSS.

### Indikator Keberhasilan

- Mampu menjelaskan *Paired T-Test*.
- Mampu menjelaskan aplikasi *Paired T-Test* menggunakan SPSS.

### Materi Pembelajaran

Uji Beda Rata-Rata Berpasangan/*Paired T-Test* (Hipotesis Pre dan Post Test)



## UJI BEDA RATA-RATA BERPASANGAN/PAIRED T-TEST (HIPOTESIS PRE DAN POST TEST)

Peneliti ingin melakukan uji intervensi untuk mengurangi dampak buruk (*harm reduction*) asap rokok di ruangan ber-AC/tertutup di lingkungan Universitas Sriwijaya, peneliti ingin mengetahui:

- Apakah perbedaan tingkat pengetahuan bahaya rokok (numerik) dan kawasan tanpa rokok (KTR) sebelum dan setelah intervensi *harm reduction*?
- Apakah perbedaan sikap responden mengenai perilaku merokok di dalam ruangan ber-AC & tertutup (numerik) sebelum dan setelah intervensi *harm reduction*?

*Uji apa yang bisa kita lakukan?*

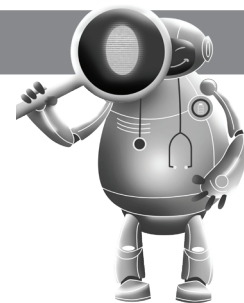
Di bidang kesehatan sering kali kita harus menarik kesimpulan apakah parameter dua populasi berbeda atau tidak. Uji statistik yang membandingkan perbedaan rata-rata (*mean difference*) dua kelompok data ini disebut uji beda dua mean. Pendekatan ujinya dapat menggunakan pendekatan distribusi T. Sebelum kita melakukan uji statistik dua kelompok data, kita perlu mengetahui apakah dua kelompok data tersebut berasal dari **dua kelompok yang independen** atau berasal dari **dua kelompok yang dependen/pasangan**.

Dikatakan kelompok independen bila data kelompok yang satu tidak tergantung dari kelompok kedua, misalnya membandingkan rata-rata kadar kolesterol orang desa dengan orang kota. Kadar kolesterol orang kota independen (tidak tergantung) dengan orang desa, dengan kata lain, sampel orang desa dan orang kota berbeda. Di lain pihak, kedua kelompok data dikatakan dependen/pasangan bila kelompok data yang dibandingkan datanya saling mempunyai ketergantungan, misalnya kadar kolesterol sebelum dan sesudah minum obat X di desa Y. Responden yang mengonsumsi obat X sebelum dan setelah intervensi adalah orang yang sama.

Bagaimana dengan contoh di atas? mari kita perhatikan contoh di bawah ini.

### STUDI KASUS 15.1

Pada bab ini kita akan membahas Uji T yang berasal dari dua kelompok yang dependen/pasangan, yaitu apakah ada perbedaan pengetahuan, sikap sebelum dan sesudah dilakukannya intervensi terpadu pengurangan dampak buruk (*harm reduction*) asap rokok pada ruangan tertutup/ber-AC di





lingkungan Universitas Sriwijaya. Responden pada intervensi ini adalah karyawan Unsri yang bersedia mengikuti intervensi selama satu bulan, lalu tingkat pengetahuan dan sikap mereka diukur sebelum dan setelah pelaksanaan intervensi.

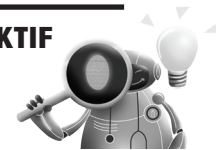
**Hipotesis 1:** Ada perbedaan **pengetahuan** tentang akibat rokok, bahaya rokok, penyakit akibat rokok dan kawasan tanpa rokok sebelum dan sesudah intervensi terpadu pengurangan dampak buruk (*harm reduction*) asap rokok pada ruangan tertutup/ber-AC di lingkungan Universitas Sriwijaya.

**Hipotesis 2:** Ada perbedaan **sikap** mengenai perilaku merokok di ruang kerja sebelum dan sesudah intervensi terpadu pengurangan dampak buruk (*harm reduction*) asap rokok pada ruangan tertutup/ber-AC di lingkungan Universitas Sriwijaya.

### UJI HIPOTESIS : PAIRED T-TEST ATAU UJI-T BERPASANGAN

 Buka data: PRE POST KTR.sav

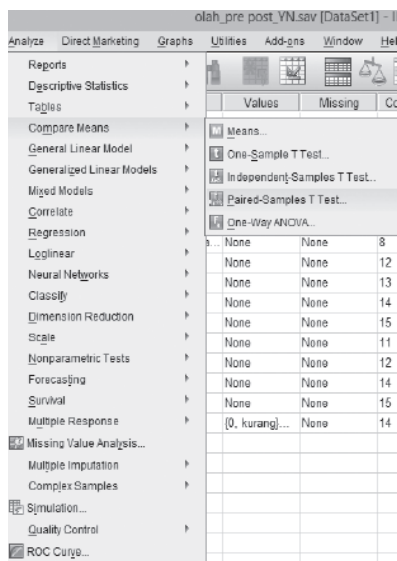
### JAWABAN DETEKTIF



#### Langkah-Langkah Pengolahan Data

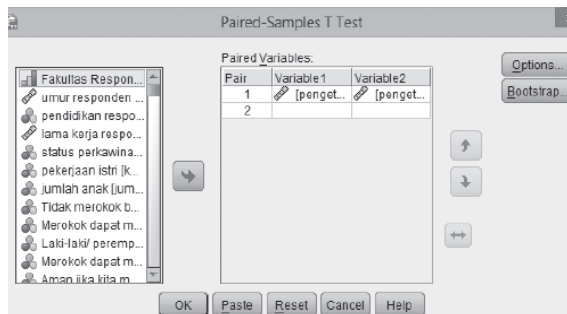
1. Analyze → Compare Means → Paired-Sample T-Test

● **GAMBAR 15.1** Langkah Analyze Paired-Sample T-Test



2. Masukkan Pengetahuan (Pre) ke kotak variabel 1 dan Pengetahuan (Post) ke kotak variabel 2 → Ok. Langkah yang sama juga dilakukan terhadap variabel sikap.

● **GAMBAR 15.2** Kotak Input Variabel Pengetahuan (Pre dan Post)



a. **Output SPSS**  
**Variabel Pengetahuan**

● **GAMBAR 15.3** Output SPSS Uji T Dependen Variabel Pengetahuan

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	penget_PRE - penget_POST	-2,150	5,639	,630	-3,405	-,895	-3,410	79	,001

**Variabel Sikap**

● **GAMBAR 15.4** Output SPSS Uji T Dependen Variabel Sikap

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	sikap_pre - sikap_post	-1,580	5,775	,642	-2,857	-,303	-2,463	80	,016

## b. Laporan Hasil

● **TABEL 15.1** Laporan Hasil Uji T Dependen Intervensi Terpadu Pengurangan Dampak Buruk (Harm Reduction) Asap Rokok pada Ruangan Tertutup/Ber-AC di Lingkungan Universitas Sriwijaya

Variabel	Kategori	N	Mean	Perbedaan Mean	Korelasi	CI 95%		P
						Lower	Upper	
Pengetahuan	Sebelum	81	42,13	-2,15	0,34	-3,4	-0,9	0,001
	Sesudah	81	44,28					
Sikap	Sebelum	81	25,25	-1,58	0,44	-2,8	-0,3	0,016
	Sesudah	81	26,83					

Sumber: Najmah, Fenny Etrawati, Yeni, Feranita Utama, 2016, *Intervensi Terpadu Pengurangan Dampak Buruk (Harm Reduction) Asap Rokok Pada Ruangan Berpengatur udara Di Lingkungan Universitas Sriwijaya*, *Bulletin Penelitian Kesehatan*, vol 44, no 4, Desember 2016: 219-226. <http://eprints.unsri.ac.id/7113>.

## c. Interpretasi

Untuk pengetahuan sebelum intervensi rata-rata skor 42,13 sedangkan setelah intervensi rata-rata skor 44,28. Hasil uji statistik didapatkan p-value  $0,001 < \alpha$  (0,05), artinya ada perbedaan yang signifikan pengetahuan tentang akibat rokok, bahaya rokok, penyakit akibat rokok dan kawasan tanpa rokok sebelum dan sesudah intervensi terpadu pengurangan dampak buruk (*harm reduction*) asap rokok pada ruangan tertutup/ber-AC di lingkungan Universitas Sriwijaya. Secara statistik diperoleh hasil bahwa ada korelasi yang lemah dan berpola positif antara pengetahuan sebelum dan sesudah intervensi.

Pada variabel sikap, skor rata-rata sebelum intervensi rata-rata 25,25 sedangkan setelah intervensi skor rata-rata 26,83. Hasil uji statistik didapatkan p-value  $0,016 < \alpha$  (0,05), artinya ada perbedaan yang signifikan mengenai sikap perilaku merokok di ruang kerja sebelum dan sesudah intervensi terpadu pengurangan dampak buruk (*harm reduction*) asap rokok pada ruangan tertutup/ber-AC di lingkungan Universitas Sriwijaya atau ada perbedaan rata-rata skor sikap sebelum intervensi dan setelah diintervensi. Secara statistik diperoleh hasil bahwa ada korelasi yang sedang dan berpola positif antara sikap sebelum dan sesudah intervensi.



# Glosarium Statistika Kesehatan

ISTILAH	DEFINISI
<b>Anova</b>	Uji F atau ANOVA digunakan untuk pengujian lebih dari dua kategori pada variabel independen. Asumsi yang digunakan pada pengujian ANOVA: <ul style="list-style-type: none"><li>· Populasi-populasi yang akan diuji berdistribusi normal</li><li>· Varians dari populasi-populasi tersebut adalah sama</li><li>· Sampel tidak berhubungan satu sama lain.</li></ul>
<b>Derajat Kepercayaan (Confidence Interval)</b>	Taksiran rentang nilai pada populasi yang dihitung dengan nilai yang diperoleh pada sampel.
<b>Faktor Perancu (confounding)</b>	Distorsi dalam memprediksi hubungan atau asosiasi antara faktor eksposur dan <i>outcome</i> (hasil) sehingga asosiasi sebenarnya tidak tampak atau ditutupi oleh faktor lainnya.
<b>Fisher Exact</b>	Alternatif uji kai kuadrat, jika syarat uji kai kuadrat tidak terpenuhi yaitu untuk tabel 2 x 2.
<b>Hipotesa</b>	Pernyataan sebagai jawaban sementara atas pertanyaan penelitian yang harus dijawab secara empiris. Hipotesis nol ( $H_0$ ) adalah hipotesis yang menunjukkan tidak ada perbedaan antar kelompok atau tidak ada hubungan antara variabel atau tidak ada korelasi antar variabel.
<b>Independen T test</b>	Digunakan untuk uji terhadap variabel independen yang terdiri dari lebih 2 kategori. Kategori dependen adalah variabel numerik, sedangkan variabel independennya adalah variabel kategori. Tujuan untuk mengetahui apakah ada perbedaan means (rata-rata) variabel numerik pada beberapa kategori variabel kategorik.
<b>Interval</b>	Pengukuran untuk membagi objek menjadi kelompok tertentu dan dapat diurutkan juga dapat ditentukan jarak dari urutan kelompok tersebut dan tidak mempunyai titik nol absolut. Misal: Suhu normal badan Andi biasanya 32 °C. Ketika dia menderita demam, suhu tubuhnya menjadi 37 °C. Berarti suhu Andi lebih panas 5°C daripada suhu normal. Nol derajat celcius bukan 0 absolut, artinya walaupun nilainya 0 bukan berarti suhu menjadi normal, tetapi tetap ada nilainya. Tetapi jika suhu tubuh dalam skala Kelvin (°K), termasuk dalam skala rasio karena memiliki 0 absolut/ mutlak.
<b>Kai Kuadrat (Chi Square)</b>	Uji hipotesis antara variabel dependen dan independen dimana bentuk variabel tersebut harus kategori dan jenis tabelnya 2xK atau BxK. Semua hipotesis untuk tabel B kali K tidak berpasangan. Syarat uji Kai Kuadrat adalah: <ul style="list-style-type: none"><li>· Tidak ada sel yang nilai observed yang bernilai nol</li><li>· Sel yang mempunyai nilai expected kurang dari 5, maksimal 20 % dari jumlah sel</li><li>· Nilai yang diambil '<i>continuity correction</i>'</li></ul>

G-1

ISTILAH	DEFINISI
<b>Kategorikal</b>	Berkaitan dengan gambaran karakteristik satu set data dengan skala pengukuran kategorikal, kita mengenal istilah jumlah atau frekuensi tiap kategori ( $n$ ) dan persentase tiap kategori (%), yang umumnya disajikan dalam bentuk tabel atau grafik. Skala pengukuran pada variabel kategorikal ada dua yaitu skala nominal dan skala ordinal.
<b>Nominal</b>	Pengukuran paling lemah tingkatannya, terjadi apabila bilangan atau lambang-lambang lain digunakan untuk mengkalsifikasikan obyek pengamatan. Misal : Jenis kelamin, hanya membedakan laki-laki dan perempuan tanpa melihat tingkatan atau urutan tertentu.
<b>Numerik</b>	Berkaitan dengan gambaran karakteristik satu set data dengan skala pengukuran numerik, dua parameter yang lazim digunakan yaitu parameter ukuran pemusatan dan parameter ukuran penyebaran. Beberapa parameter untuk ukuran pemusatan, yaitu <i>mean</i> , <i>median</i> , dan <i>modus</i> . Untuk parameter ukuran penyebaran, ada beberapa istila seperti; standar deviasi, varians, koefisien varians, interkuartil, range, dan minimum maksimum. Data variabel dengan skala pengukuran numerik umumnya disajikan dalam bentuk tabel dan grafik. Skala pengukuran pada variabel kategorikal ada dua yaitu skala interval dan ratio.
<b>Odds Rasio</b>	<p><i>Odds ratio</i> (OR) atau rasio odds merupakan perbandingan odds subyek sakit (kasus) dengan odds subyek tidak sakit (kontrol).</p> $\text{Odds Ratio (OR)} = \frac{\text{odds kasus } (\frac{a}{b})}{\text{odds kontrol } (\frac{c}{d})}$
<b>Ordinal</b>	Pengukuran ini tidak hanya membagi objek menjadi kelompok-kelompok yang tidak tumpang tindih, tetapi antara kelompok itu ada hubungan (rangking). Jadi dari kelompok yang sudah ditentukan dapat diurutkan menurut besar kecilnya. Dengan kata lain, data skala ordina mempunyai urutan kategori yang bermakna, tetapi tidak ada jarak yang terukur diantara kategori. Misal: Tingkat pendidikan.
<b>Populasi</b>	Keseluruhan unit di dalam pengamatan yang akan dilakukan.
<b>Prevalens Rasio</b>	Ukuran rasio untuk melihat hubungan antara paparan dan penyakit pada penelitian potong lintang ( <i>cross sectional</i> ) dengan menggunakan rumus odd rasio maupun risk rasio, hanya saja data yang digunakan bukan data kumulatif insidensi melainkan data prevalensi.
<b>Rasio</b>	Pengukuran untuk mengelompokkan data, kelompok itu pun dapat diurutkan dan jarak antara urutan pun dapat ditentukan. Selain itu, sifat lain untuk data dengan skala rasio kelompok tersebut dapat diperbandingkan ( <i>ratio</i> ). Hal ini disebabkan karena skala rasio mempunyai titik 'nol mutlak'. Misal : Usia Responden pada penelitian.

ISTILAH	DEFINISI
<b>Regresi Cox</b>	Regresi Cox atau regresi hazard proporsional ( <i>proportional hazards regression</i> ) berasumsi bahwa rasio hazard pada beberapa kelompok terpapar yang berbeda konstan setiap waktu. Regresi Cox sama seperti halnya risk ratio dengan membandingkan nilai subjek atau responden pada exposure dan outcome, tetapi pada regresi Cox, waktu dari setiap waktu <i>follow up</i> sangat dipertimbangkan, sehingga exposure pada subjek dengan outcome positif bisa dibandingkan dengan subjek dengan outcome yang masih diikuti dan outcome negatif.
<b>Regresi Logistik</b>	Suatu pendekatan model yang matematis untuk menganalisa hubungan antara satu atau beberapa variabel independen (kategori dan numerik) dengan variabel dependen kategorik yang bersifat dikotom/biner. Variabel kategorik dikotom adalah variabel dengan dua nilai variasi atau kategori,
<b>Regresi Linear</b>	Suatu pendekatan model untuk membuat prediksi.
<b>Reliabilitas</b>	Konsistensi suatu hasil pengukuran. Dalam penelitian ini reliabilitas kuesioner diukur dengan cara <i>one shot</i> . Disini pengukurannya hanya sekali dan hasilnya dibandingkan dengan pernyataan lain. (5)Pengujian reliabilitas dimulai dengan menguji validitas terlebih dahulu. Jadi jika sebuah pernyataan tidak valid, maka pernyataan tersebut dibuang. Pernyataan-pernyataan yang sudah valid kemudian baru secara bersama diukur reliabilitasnya. Untuk mengetahui reliabilitas suatu variabel (misal sikap) maka kita membandingkan nilai r tabel dengan nilai <b>r</b> hasil (nilai ALPHA pada output data). Ketentuannya bila r Alpha lebih besar daripada <b>r</b> tabel maka pertanyaan tersebut <i>reliable</i> dan sebaliknya.
<b>Risk Ratio</b>	<ul style="list-style-type: none"> <li>· Rasio dari risiko untuk terjadinya penyakit pada kelompok terpapar dibandingkan kelompok yang tidak terpapar.</li> <li>· Insidensi kumulatif kelompok terpapar merupakan proporsi kasus baru pada kelompok yang terpapar, sedangkan insidensi kumulatif kelompok tidak terpapar merupakan proporsi kasus baru pada kelompok yang tidak terpapar.</li> </ul> $\text{Risk Ratio} = \frac{\text{insidensi kumulatif kelompok terpapar } \left(\frac{a}{N1}\right)}{\text{insidensi kumulatif kelompok tidak terpapar } \left(\frac{b}{N0}\right)}$
<b>Sampel</b>	Sebagian dari populasi yang nilai/karakteristiknya diukur dan digunakan untuk menduga karakteristik dari populasi.

ISTILAH	DEFINISI
<b>Signifikansi (nilai P/P value)</b>	<p>Besarnya kemungkinan hasil yang diperoleh atau hasil yang lebih ekstrim diperoleh karena faktor peluang, bila hipotesis nol benar.</p> <ul style="list-style-type: none"> <li>• <b>p</b> value &lt;0.001; adanya bukti yang kuat untuk menolak hipotesa nul.</li> <li>• <b>p</b> value &lt;0.01 ; adanya bukti yang sedang untuk menolak hipotesa nul.</li> <li>• <b>p</b> value &gt;0.1; adanya bukti yang lemah untuk menolak hipotesa nul.</li> </ul>
<b>Student T test</b>	<p>Analisis perbandingan untuk dua sampel yang berpasangan, sampel yang berpasangan diartikan sebagai sebuah sampel dengan subjek yang sama namun mengalami dua perlakuan atau pengukuran yang berbeda. Digunakan untuk uji terhadap variabel independen yang terdiri dari lebih 2 kategori. Kategori dependen adalah variabel numerik, sedangkan variabel independennya adalah variabel kategori.</p>
<b>SPSS</b>	<p>SPSS (<i>Statistical Product and Service Solutions</i>) adalah software pengolahan data yang penggunaannya sangat tergantung dari penguasaan materi statistik sekaligus pemahaman perintah-perintah atau menu-menu di dalamnya. program komputer statistik yang mampu untuk memproses data statistik secara cepat dan tepat, menjadi berbagai output yang dikehendaki para pengambil keputusan.</p>
<b>STATA</b>	<p>Software statistik lengkap dan terintegrasi yang dapat memberikan apapun yang dibutuhkan dalam menganalisa data, manajemen data, dan grafik. Keunggulan STATA adalah cepat, akurat, dan mudah digunakan dengan aplikasi syntax. Proses analisa data dapat didokumentasikan dalam bentuk do file dan log file sehingga kita dapat melakukan analisa ulang pada data yang sama.</p>
<b>Uji Normalitas</b>	<p>Uji untuk mengetahui sebaran data normal atau tidak.</p>
<b>Validitas</b>	<p>Validitas mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam mengukur data. Untuk mengukur pengetahuan dan sikap, diperlukan alat ukur berupa kuesioner. Untuk mengukur validitas pernyataan yang berkaitan dengan pengetahuan dan sikap, dilakukan dengan cara melakukan korelasi antar skor masing-masing pernyataan terhadap skor total. Suatu pernyataan dikatakan valid bila skor pernyataan tersebut berkorelasi secara signifikan dengan skor totalnya. Keputusan uji, bila <b>r</b> hitung masing-masing pernyataan (dilihat pada output data) lebih besar dari <b>r</b> tabel maka <math>H_0</math> ditolak yang berarti valid dan jika <b>r</b> hitung lebih kecil dari <b>r</b> tabel maka <math>H_0</math> diterima yang berarti pernyataan tidak valid.</p>

**Sumber:**

1. Najmah. 2011. *Managemen dan Analisis Data Kesehatan*. Yogyakarta: Nuha Medika.
2. Najmah. 2015. *Epidemiologi untuk Mahasiswa Kesehatan Masyarakat*. Jakarta: Rajawali Pers.



# Indeks Statistika Kesehatan

## A

ADS 118–120  
analisis data sekunder 117  
analisis deskripsi sederhana 2  
analisis deskriptif 3, 25, 67, 85, 107, 121  
analisis multivariat 51  
analisis regresi sederhana 2  
analisis Student-T 2  
analisis survival 2, 53, 54, 56  
analisis univariat 25, 27, 85  
Anova/Manova 2  
autokorelasi 142, 145

## B

Besral 54, 155  
bivariat 2, 147, 148, 153, 154  
BPS 118

## C

*case-control* 26, 134  
Chi-Square 2, 155–159, 162, 164, 166, 167  
*cluster* 122, 123, 126  
*cluster sampling* 126  
*cohort* 134, 162  
*complex samples general linear model* 126  
*content analysis* 118  
*cross sectional* 12, 34, 40, 41, 134, 142  
CSGLM 126–128, 130, 131

## D

Dahlan 164, 167  
data numerik 7, 8, 28, 85, 87, 89, 95, 142  
*degree of freedom* 170  
derajat kepercayaan 2, 26, 36, 38, 41, 49, 51, 57,  
126, 140  
df 96, 112, 164, 166, 170

dikotom 134  
diskrit 12, 14, 16

## E

*existing statistic* 118

## F

FEV 12, 13

## H

hazard proporsional 54, 60, 62, 63, 64  
*hazard ratio* 53, 59, 63  
*homoscedasticity* 145

## J

jangkauan data 19  
Julie Pasco 135

## K

kecenderungan (*likelihood*) 51  
Kirkwood 54, 135  
Kleinbaum 54  
kohort 34, 54, 55, 134, 158, 162  
kolinearitas 146

## L

linearitas 145  
*log file* 4  
*loss of follow up* 54  
Lyle Gurrin 55, 135

## M

Margaret Henry 135  
*mean* 9, 12, 14, 18, 95, 97, 112, 125, 142, 173,  
175, 178, 184

median 12, 14, 19, 87, 95, 97  
metode Kaplan Meir 54

## N

Najmah 27, 35, 38, 40, 45, 71, 89, 90, 99, 108,  
109, 110, 135, 143, 150, 155, 167, 187  
*nearmiss* 41  
nilai *expected* 156, 158, 163, 164, 167  
nilai probabilitas 154  
nilai tengah 2, 95  
nominal 12, 14, 73, 76, 77, 86, 87, 88, 158

## O

*odds ratio* 2, 33, 34, 36, 40, 134, 160  
One-Way Anova 178  
OR 33, 34, 38, 39, 44, 45, 49, 50, 51, 134, 135,  
136, 138, 140, 149, 157, 162, 163, 173,  
178  
ordinal 12, 14, 72, 74, 86, 87, 88

## P

PR 33, 34, 40, 41, 162  
prevalensi rasio 34, 41

## R

*recode* 10, 30, 60, 122  
regresi hazard proporsional 54  
regresi linear 2, 126, 141, 142, 145, 148, 155  
regresi logistik berganda 134  
regresi logistik sederhana 43, 44, 133, 155  
reliabilitas 107–110, 112–115, 118  
*Respiratory Virus Study* 55  
*risk ratio* 2, 33, 34, 36, 38, 40, 41, 54, 160  
RR 33, 34, 37, 150, 162

## S

SDKI 108, 117–121, 126, 129  
*secondary data analysis* 118  
skala interval 87  
skala rasio 87, 88  
SPSS 67, 107, 117, 133, 141, 153, 169, 177, 183  
standar deviasi 2, 10, 12, 19, 87, 95, 112, 170,

182

standar *error* 87, 125  
studi eksperimental 34  
studi kasus kontrol 34, 38  
studi potong lintang 34  
*syntax* 1–5, 7, 19, 22, 23, 25, 26, 28, 30, 36, 43,  
45, 49, 89, 101, 158, 171, 179

## T

Tabachnick 54  
teknik bola salju (*snowball*) 26  
tes keselarasan 155

## U

uji Anova 142, 145, 177, 178  
uji Benferoni 178  
uji *Chi-square* 153, 156, 158, 159, 163  
uji *Fisher Exact* 153  
uji *independent student T-Test* 169  
uji Kolmogorov-Smirnov 156  
uji Lavene 172  
uji logistik berganda 43  
uji logistik sederhana 43  
uji proporsi 2  
uji regresi Cox 53  
uji *Shapiro Wilk* 178  
univariat 25, 27, 85

## V

*Vaccine Immunization Research Group* 55  
validitas 107–110, 112–114, 118  
variabel dependen 21, 54, 128, 131, 134, 137,  
138, 142, 144, 154, 157  
variabel independen 21, 128, 131, 134, 135,  
137, 142, 143, 148, 154, 157  
variabel kategorik 30, 62, 86  
variabel numerik 10, 30, 86, 87, 89, 96, 99, 179,  
180

## W

Widyawati 157, 170, 173, 178