

PRESENTASI DAN ANALISIS DATA DENGAN



I Made Tirta

PRESENTASI DAN ANALISIS DATA DENGAN R

Diterbitkan oleh
UPT Penerbitan UNEJ
Jl. Kalimantan 37 Jember 68121
Telp. 0331-330224, Psw. 319, Fax. 0331-339029
E-mail: upt-penerbitan@unej.ac.id

Hak Cipta @ 2014
Cover : Noerkoentjoro W.D.

Perpustakaan Nasional RI – Katalog Dalam Terbitan

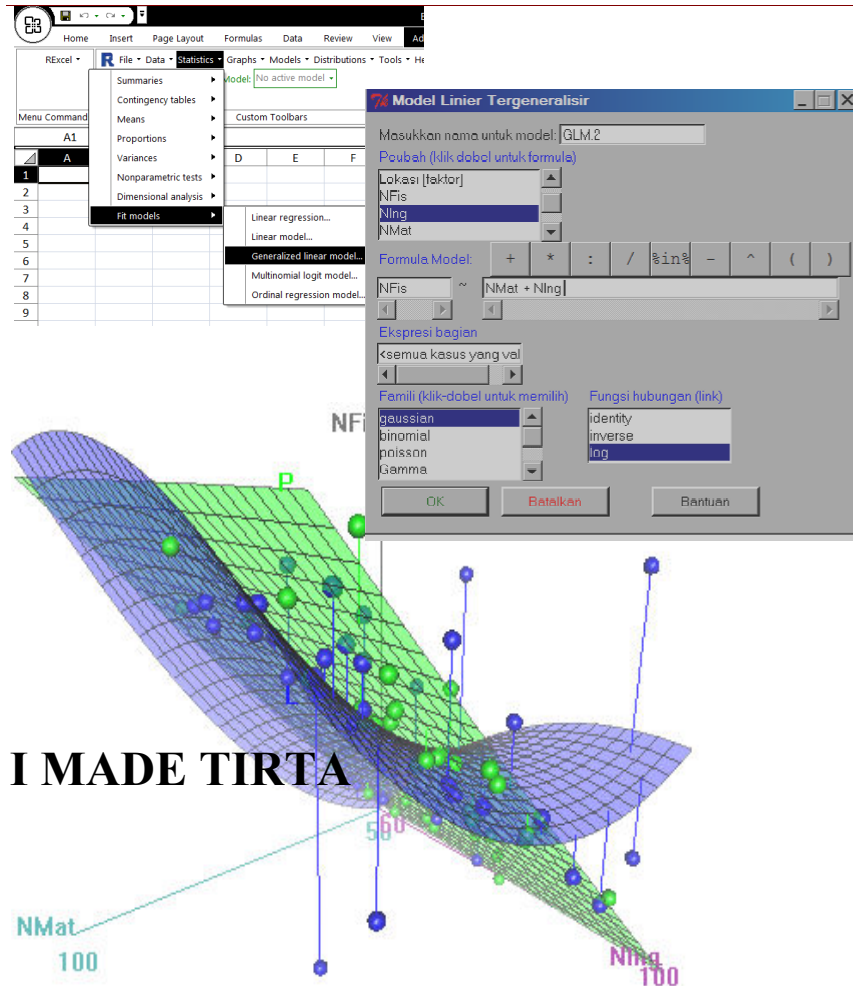
510.285
TIR TIRTA, I Made
p Presentasi dan Analisis Data dengan R/oleh I
Made Tirta.--Jember: Jember University Press,
2014
xxii, 320 hlm. ; 23 cm.

ISBN: 978-602-9030-12-9

1. PENGOLAHAN DAN ANALISA DATA
I. Judul

Hak Cipta dilindungi Undang-Undang. Dilarang memperbanyak tanpa
ijin tertulis dari penerbit, sebagian atau seluruhnya dalam bentuk
apapun, baik cetak, *photoprint*, maupun *microfilm*.

PRESENTASI DAN ANALISIS DATA DENGAN R



KONTRIBUSI PADA RCOMMANDER

*The R Commander comes with translations from English into several other languages. I am grateful to the following individuals for preparing these translations: Brazilian Portuguese, Adriano Azevedo Filho; Catalan, Manel Salamero; French, Philippe Grosjean; German: Gerhard Schoen; **Indonesian, I Made Tirta**; Italian, Stefano Calza; Japanese, Takaharu Araki; Romanian, Adrian Dusa; Russian, Alexey Shipunov; Slovenian, Jaro Lajovic; Spanish, Carlos Enrique Carleos Artime.*

Prof John Fox (jfox@mcmaster.ca)
Department of Sociology
McMaster University
Hamilton, Ontario, Canada L8S 4M4

PRAKATA

Pengalaman mengajar Statistika di berbagai jurusan, baik jurusan matematika (murni dan kependidikan), ataupun jurusan lain, ditemukan, bahwa oleh mahasiswa statistika dirasa sebagai salah satu matakuliah yang sulit dan membosankan. Banyak konsep-konsep abstrak yang terpaksa diterima mahasiswa walaupun belum dipahami dengan baik. Sementara itu bagi mahasiswa yang bukan jurusan matematika (atau statistika), ketika mereka membuat tugas akhir yang membutuhkan statistika, sebagian besar dari mereka hanya memaksakan analisis tertentu walaupun kondisi datanya tidak sesuai.

Persoalan lain yang dihadapi adalah kebutuhan akan penggunaan program komputer yang tidak bisa ditawar. Selain karena jumlah ataupun kompleksitas datanya, penggunaan program komputer dalam analisis data memungkinkan peneliti lebih fokus pada pemilihan metode atau model yang lebih baik dan tidak menghabiskan waktunya untuk melakukan perhitungan secara manual. Dalam hal ini kebanyakan program komputer (piranti lunak) yang dipakai adalah piranti lunak bajakan. Kondisi ini menempatkan Indonesia sebagai salah satu negara pembajak terkemuka (no 5 tahun 2005, no 3 tahun 2006)¹. Padahal, di

¹ Tempo Interaktif 29 Agustus 2006. *Indonesia Pembajak Software NO.3*.

sekitar kita tersedia piranti lunak yang legal dan gratis serta memiliki kemampuan yang tidak kalah dengan program komersial bahkan dengan menu sudah tersedia dalam Bahasa Indonesia. Sementara itu menu program komputer komersial yang berbahasa Inggris masih sulit dipahami bagi pengguna statistika pemula.

Matakuliah Metode statistika, umumnya ada hampir pada semua program studi di Perguruan Tinggi. Namun, mahasiswa Jurusan Matematika dan Pendidikan Matematika memiliki keunikan dibandingkan mahasiswa lain diantaranya:

1. Sebagai mahasiswa jurusan matematika, mereka diharapkan dan dituntut memiliki keterampilan dan pengetahuan yang lebih dalam hal pemanfaatan komputer. Oleh karena itu penguasaan program statistika berbasis *open source* sangat penting sehingga mereka dapat mengembangkan aplikasi metode statistika secara lebih leluasa dan fleksibel.
2. Sebagai mahasiswa jurusan pendidikan, mereka membutuhkan metode statistika beberapa dalam hal, misalnya seperti berikut ini.
 - a. Untuk menganalisis data dalam rangka penyelesaian tugas akhir. Sebagian besar tugas akhir (skripsi) mereka adalah tentang aplikasi pembelajaran matematika yang menggunakan data yang harus dianalisis dengan menggunakan metode statistika, baik dasar maupun tingkat lanjut.
 - b. Sebagai calon guru dalam bidang matematika, mereka harus memahami statistika baik secara praktis, maupun matematis (melalui penurunan dan pembuktian formal rumus-rumus statistika, yang biasanya dibahas pada matakuliah statistika matematika). Selain itu, mereka juga harus mampu menjelaskan konsep-konsep statistika yang abstrak melalui visualisasi/ilustrasi menggunakan komputer.

Kebutuhan akan butir 1, dapat dipenuhi dengan memanfaatkan program statistika berbasis *open source* R. Sebagai *open source* yang didukung oleh banyak statistisi di seluruh dunia, R memiliki rentang analisis statistika yang sangat luas mulai dari yang mendasar sampai tingkat lanjut, mulai dari berbasis stand alone, sampai berbasis web. Adanya dukungan beberapa penerbit besar (Springer, Chapman & Hall dan lain-lain) terhadap dokumentasi R juga menambah kemudahan dalam mengakses dan memanfaatkan kemampuan R.

Kebutuhan butir 2a, juga mudah dipenuhi dengan menggunakan program R, karena R memiliki kemampuan yang sangat baik dalam hal visualisasi grafis baik secara stand alone langsung pada komputer lokal, maupun melalui web yang bisa diakses dari jarak jauh.

Berdasarkan kondisi di atas, buku ini disusun secara lebih spesifik diantaranya:

1. Secara konten disusun dengan struktur yang memudahkan mahasiswa untuk mempelajari secara mandiri topik dan materi yang ada.
2. Secara dokumentasi, materi dan topik terkait disajikan dengan tetap mempertahankan dan merealisasikan link-link yang terkait, sehingga modul ini sangat sesuai untuk dijadikan e-modul yang memudahkan mahasiswa menelusuri hyperlink yang ada dalam naskah.
3. Dilengkapi dengan contoh ilustrasi/ visualiasi grafis. Sejauh ini ilustrasi ini hanya bisa dijalankan pada komputer stand alone. Pengembangan ke depan akan diusahakan adanya ilustrasi dan visualisasi yang berbasis web. Pada akhirnya nanti akan dihasilkan e-modul yang dilengkapi dengan visualisasi dinamik online berbasis web. Dengan demikian pembaca/mahasiswa akan langsung bisa mengeksplorasi contoh ilustrasi langsung pada bagian ilustrasi itu muncul, tanpa perlu keluar dari e-modul.

Pada kesempatan ini penulis menyampaikan terimakasih yang sebesar-besarnya, kepada semua pihak yang secara langsung maupun tidak langsung membantu terselesaikannya buku ini baik dari segi substansi maupun teknis terkait desain dan percetakan.

1. Bapak Dr. I Made Sumertajaya, S.Si, M.Si, dosen statistika Jurusan Statistika FMIPA Institut Pertanian Bogor yang banyak memberikan masukan dan koreksi menyangkut istilah, substansi, maupun sistematika penulisan naskah awal.
2. Bapak Ir Martinus M.Sc., Ph.D (dosen statistika Fakultas Pertanian UNEJ), Bapak Drs Suharto M.Kes (dosen statistika FKIP UNEJ), dan Ibu Dian Aggraini S.Si, M.Si (dosen statistika FMIPA UNEJ),

yang telah bersedia mencermati dengan seksama naskah pracetak sehingga mengurangi kesalahan eja, maupun ketidakjelasan rumus serta ilustrasi lainnya.

3. Tim editor UPT Penerbit Universitas Jember yang banyak memberi masukan dan koreksi dari segi penggunaan bahasa dan tatasaji.

Penulis masih tetap berharap adanya masukan dan kritik pembaca sehingga di kemudian hari buku ini bisa terus disempurnakan sehingga bisa memberi manfaat yang lebih baik bagi para pengguna.

Jember, Januari 2014

Penulis

I Made Tirta

DAFTAR ISI

PRAKATA.....	v
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xvii
DAFTAR GAMBAR.....	xviii
STRUKTUR BUKU.....	xxv
BAGIAN I PENGERTIAN UMUM.....	1
Bab 1. PENGERTIAN DAN LINGKUP STATISTIKA.....	3
Tujuan Belajar.....	3
Materi.....	4
1.1 Statistik dan Statistika.....	4
1.1.1 Ruang Lingkup Statistika.....	5
1.1.2 Manfaat dan Aplikasi Statistika.....	7
1.1.3 Penyalahgunaan Statistika.....	8
1.1.4 Peranan Komputer dalam Statistika.....	9

1.2	Notasi dan Istilah Penting dalam Statistika.....	10
1.2.1	Notasi Penting.....	10
1.2.2	Istilah Penting.....	15
1.3	Data.....	16
1.3.1	Data dan Informasi.....	16
1.3.2	Jenis dan Skala Data.....	17
1.3.3	Presentasi dan Analisis Data.....	20
1.4	Ringkasan dan Bacaan Lebih Lanjut.....	20
1.4.1	Ringkasan.....	20
1.4.2	Bacaan Lebih Lanjut.....	21
Bab 2.	SELAYANG PANDANG R.....	23
	Tujuan Belajar.....	23
	Materi.....	24
2.1	Fitur Umum R.....	24
2.2	Fitur RCommander.....	25
2.2.1	Tampilan RCommander Asli.....	29
2.2.2	Plugin: Kombinasi RCommander dengan Paket Lain.....	35
2.2.3	RCommander Termodifikasi.....	38
2.3	Visualisasi Statistika dengan Paket StatDemo.....	40
2.3.1	Grafik Kepadatan Distribusi Normal.....	41
2.3.2	Grafik Nilai-tengah dan Rata-rata.....	42
2.3.3	Grafik Kepadatan Distribusi Normal dan t.....	43
2.3.4	Grafik Nilai Kritis pada Distribusi t.....	44
2.3.5	Grafik Kepadatan Distribusi Normal.....	45

2.3.6	Grafik Kontur Distribusi Normal Bivariat.....	45
2.3.7	Ilustrasi Estimasi Selang.....	47
2.3.8	Ilustrasi Sabuk Keyakinan Regresi.....	47
2.3.9	Paket TeachingDemos	50
2.4	Visualisasi Statistika Berbasis Web dengan Paket Shiny	51
2.5	Rexcel: Menggabungkan R dengan Excel	53
2.5.1	Tampilan RCommader pada Excel.....	54
2.5.2	Komunikasi R dengan Excel	57
2.6	Ringkasan dan Bacaan Lebih Lanjut	60
2.6.1	Ringkasan	60
2.6.2	Bacaan lebih Lanjut.....	60
2.7	Soal-soal.....	61
BAGIAN II. TEORI PELUANG DAN DISTRIBUSI.....		63
Bab 3.	DASAR-DASAR TEORI PELUANG.....	65
	Tujuan Belajar.....	65
	Materi.....	66
3.1	Teori Peluang.....	66
3.1.1	Konsep Dasar Peluang.....	67
3.1.2	Peluang dan Frekuensi Relatif.....	67
3.1.3	Peristiwa Bersyarat.....	70
3.1.4	Peristiwa Saling Lepas dan Peristiwa Saling Bebas	71
3.2	Peubah Acak dan Fungsi Kepadatan Peluang.....	74
3.3	Distribusi Data	77

3.4	Distribusi dari Percobaan Bernoulli.....	79
3.5	Sampel Acak dan Karakteristiknya.....	82
3.5.1	Ukuran Pemusatan.....	82
3.5.2	Ukuran Penyebaran.....	85
3.6	Representasi Grafik Fungsi Kepadatan.....	88
3.7	Ringkasan dan Bacaan Lebih Lanjut.....	95
3.7.1	Ringkasan.....	95
3.7.2	Bacaan Lebih Lanjut.....	96
3.8	Soal-soal.....	96
Bab 4.	DISTRIBUSI NORMAL DAN APLIKASINYA.....	99
	Tujuan Belajar.....	99
	Materi.....	100
4.1	Fungsi Kepadatan Distribusi Normal.....	100
4.2	Menghitung Peluang pada Distribusi Normal.....	103
4.3	Aplikasi Distribusi Normal.....	109
4.4	Sampel Normal.....	111
4.4.1	Distribusi Sampel.....	111
4.4.2	Populasi, Sampel dan Teknik Sampling.....	114
4.4.3	Hipotesis, Kesalahan Tipe I dan II.....	115
4.5	Ringkasan Statistika Parametrik dengan R.....	118
4.5.1	Statistika Inferensial dengan Uji Hipotesis.....	119
4.5.2	Analisis Eksplorasi Data.....	123
4.6	Ringkasan dan Bacaan Lebih Lanjut.....	125
4.6.1	Ringkasan.....	125

4.6.2	Bacaan Lebih Lanjut.....	126
4.7	Soal-soal.....	126
BAGIAN III ANALISIS DATA DENGAN R.....		128
Bab 5.	EKSPLORASI DATA DENGAN R.....	129
	Tujuan Belajar.....	129
	Materi.....	130
5.1	Menyiapkan Data.....	130
5.1.1	Menyusun Dataset Baru.....	131
5.1.2	Mengaktifkan Data Internal R.....	132
5.1.3	Mengimpor Data.....	132
5.1.4	Membangkitkan data.....	136
5.2	Eksplorasi dan Presentasi Data Secara Numerik.....	137
5.2.1	Menampilkan Keseluruhan Data.....	139
5.2.2	Menghitung Ringkasan Statistik Data.....	140
5.3	Eksplorasi dan Presentasi Data DENGAN Grafik.....	145
5.3.1	Histogram.....	146
5.3.2	Boxplot.....	147
5.3.3	QQ-Plot.....	150
5.3.4	Plot Nilai-tengah.....	155
5.3.5	Diagram Pencar dan Matriks Diagram Pencar.....	157
5.3.6	Grafik 3D.....	161
5.4	Ringkasan dan Bacaan Lebih Lanjut.....	164
5.4.1	Ringkasan.....	164

5.4.2	Bacaan Lebih Lanjut.....	165
5.5	Soal-soal.....	165
Bab 6.	UJI BEDA NILAI TENGAH DAN PROPORSI.....	168
	Tujuan Belajar.....	169
	Materi.....	169
6.1	Estimasi Parameter.....	169
6.1.1	Penduga Titik.....	170
6.1.2	Penduga Selang Nilai Tengah.....	170
6.2	Uji Nilai-Tengah Satu dan Dua Kelompok.....	173
6.2.1	Estimasi dan Uji Nilai-tengah Satu Kelompok.....	173
6.2.2	Visualisasi Hubungan Uji-t satu kelompok dengan Estimasi Selang Nilai-Tengah	176
6.2.3	Uji Nilai-tengah Dua Kelompok Saling Bebas.....	181
6.2.4	Uji Nilai-tengah Dua Kelompok Berpasangan	187
6.3	ANAVA (Uji Beda Tiga Kelompok atau Lebih).....	192
6.3.1	Analisis Ragam Satu Arah.....	194
6.3.2	Analisis Ragam Multi Arah.....	201
6.4	Uji Proporsi.....	203
6.4.1	Uji Proporsi Sampel Tunggal	204
6.5	Ringkasan dan Bacaan Lebih Lanjut.....	207
6.5.1	Ringkasan	207
6.5.2	Bacaan Lebih Lanjut.....	208
6.6	Soal-soal.....	208
Bab 7.	ANALISIS KORELASI DAN REGRESI	210

Tujuan Belajar.....	210
Materi.....	211
7.1 Analisis Korelasi.....	211
7.1.1 Menghitung korelasi secara manual	212
7.1.2 Menghitung Korelasi dengan RCommander	214
7.2 Analisis Regresi Sederhana.....	217
7.2.1 Estimasi dan Uji Inferensi Parameter Regresi.....	218
7.2.2 Prediksi dalam Regresi	230
7.3 Model Linier: Regresi Linier Peubah Ganda.....	232
7.4 Pemilihan Model dengan Regresi Bertatar	236
7.5 Diagnostik Model Melalui Grafik.....	237
7.6 Ringkasan dan Bacaan Lebih Lanjut	246
7.6.1 Ringkasan	246
7.6.2 Bacaan Lebih Lanjut.....	246
7.6.3 Soal-soal	247
GLOSARIUM	249
DAFTAR PUSTAKA	258
LAMPIRAN.....	264
Daftar Referensi Terkait R.....	264
Mengunduh dan Menginstal R.....	274
Mengunduh (Men- <i>download</i>) R.....	274
Menginstal R	275
Menu Bahasa Indonesia dan Paket Tambahan.....	283

Struktur Menu RCommander.....	288
Struktur Menu RExcel	290
Daftar Persamaan	291
Daftar Contoh.....	295
INDEKS	297

DAFTAR TABEL

Tabel 2.1 Daftar Paket Plug-in RCmdr	36
Tabel 3.1 Tabel Kerja tabulasi peristiwa dalam toss uang logam	69
Tabel 3.2 Tabulasi Ruang Sampel pada Pelemparan Dua Dadu	73
Tabel 3.3 Perhitungan fungsi kepadatan jumlah mata dadu yang muncul pada pelemparan dua dadu	78
Tabel 6.1 Contoh Tabulasi Hasil Pengamatan Percobaan Satu Faktor	197

DAFTAR GAMBAR

Gambar 2.1 Tampilan Jendela RCommader.....	28
Gambar 2.2 Menu Langsung Paket Pengajaran Statistika Paket StatDemo	28
Gambar 2.3 Tampilan Menu Utama dan Panel Rcommander dalam Bahasa Indonesia	29
Gambar 2.4 Tampilan Menu File	31
Gambar 2.5 Tampilan Menu Edit	32
Gambar 2.6 Tampilan Menu Data dengan Submenunya.....	33
Gambar 2.7 Tampilan Menu Statistika dengan SubMenu Uji Nilai-tengah (Rerata).....	33
Gambar 2.8 Tampilan Menu Statistika dengan Submenu Pencocokan Model.....	34
Gambar 2.9 Tampilan Menu Grafik	34
Gambar 2.10 Tampilan Menu Distribusi dengan Submenu Distribusi Kontinu.....	35
Gambar 2.11 Tampilan Informasi Author dan Kontributor Rcommander	35

Gambar 2.12	Cara mengaktifkan paket <i>Plug-in</i>	37
Gambar 2.13	Menu RCommander dengan <i>Plug-in</i> Media Pembelajaran (StatDemo)	38
Gambar 2.14	Tampilan RCommander dengan BiodiversityR.....	39
Gambar 2.15	Tampilan RCommander dengan DAS (DAS+R)	40
Gambar 2.16.	Ilustrasi Grafik Distribusi Normal Baku dan Tidak Baku ..	41
Gambar 2.17.	Ilustrasi Hubungan antara Nilai-tengah dan Rata-rata dengan besarnya sampel	42
Gambar 2.18.	Ilustrasi Grafik Distribusi t dan Normal Baku (Z).....	43
Gambar 2.19.	Ilustrasi Titik kritis (nilai t Tabel) dari Distribusi t	44
Gambar 2.20.	Ilustrasi Nilai Kritis (Z Tabel) Distribusi Normal	46
Gambar 2.21.	Ilustrasi Interval Keyakinan Nilai-tengah.....	48
Gambar 2.22.	Ilustrasi Grafik Sabuk Keyakinan Regresi Linier.....	49
Gambar 2.23	Pilihan menu untuk toss uang logam atau toss dadu	51
Gambar 2.24	Tampilan web dengan Rshiny mengilustrasikan ringkasan data simulasi dan histogram beberapa jenis distribusi	52
Gambar 2.25	Tampilan web dengan Rshiny Mengilustrasikan grafik diagram pencar dan estimasi regresi	53
Gambar 2.26	Tampilan menu utama RCommander dalam Excel 2003	55
Gambar 2.27	Tampilan menu R Excel pada Excel 2003.....	55
Gambar 2.28	Tampilan Menu Rcommander pada Excel 2007.....	56
Gambar 2.29	Tampilan menu R Excel pada Excel 2007.....	56
Gambar 2.30	Contoh Tampilan Data Aktif R pada <i>Worksheet</i> Excel	57
Gambar 2.31	Contoh Tampilan luaran komputer pada worksheet Excel ..	58
Gambar 2.32	Tampilan luaran komputer pada jendela Rcommander	59

Gambar 2.33	Contoh Tampilan Menu Kontak R pada Worksheet Excel..	59
Gambar 3.1	Animasi toss uang logam atau toss dadu	69
Gambar 3.2	Tampilan ilustrasi Peluang dan Frekuensi Relatif.....	70
Gambar 3.3	Ilustrasi Contoh Grafik Tiga Distribusi (Normal, t dan Chi-sq)	89
Gambar 3.4	Ilustrasi Fungsi Distribusi Kumulatif dan Penentuan Nilai Median pada distribusi kontinu.	90
Gambar 3.5	Ilustrasi Fungsi Kumulatif untuk Distribusi Diskrit dan Penentuan Median	91
Gambar 3.6.	Posisi Nilai-Tengah, Modus dan Median untuk Distribusi Simetris dan Tidak Simetris	92
Gambar 3.7	Panel Slider untuk mengatur besarnya Nilai-tengah dan Deviasi Baku	93
Gambar 3.8	Grafik Data Acak berdasarkan Pemusatan dan penyebaran	94
Gambar 3.9	Menu Demos yang berasal dari pustaka TeachingDemos...	95
Gambar 4.1	Grafik kepadatan Distribusi Normal dengan berbagai Nilai tengah dan Ragam.	102
Gambar 4.2	Ilustrasi $P(x_1 < X < x_2)$ yang identik dengan luas daerah kurva normal yang dibatasi nilai x_1 (batas kiri) dan x_2 (batas kanan)	105
Gambar 4.3	Dialog untuk menghitung kuantil dan peluang distribusi normal.....	106
Gambar 4.4	Dialog untuk menggambar grafik peluang distribusi normal	107
Gambar 4.5	Membuat grafik distribusi Normal dengan rerata dan deviasi baku tertentu (Nilai-tengah 50, deviasi baku 5)	107
Gambar 4.6	Histogram menggambarkan distribusi data asli (atas) dan distribusi rata-rata sampel (bawah) untuk sampling sebanyak 100 kali, dengan masing-masing sampel berukuran 10.....	112

Gambar 5.1 Menu Data pada RCommander	131
Gambar 5.2 Tampilan Editor Data	131
Gambar 5.3 Mengaktifkan data dari paket pada R.....	132
Gambar 5.4 Dialog Baca Data dari Teks.....	134
Gambar 5.5 Dialog Nama File dan Direktori Impor	135
Gambar 5.6 Menu dan dialog untuk membangkitkan data dari distribusi normal.	137
Gambar 5.7 Nama Data Muncul pada Panel	138
Gambar 5.8 Informasi Data yang ada di Memori.....	138
Gambar 5.9. Tampilan Edit Data, Kolom atau Sel Dapat Di-edit	139
Gambar 5.10 Tampilan Lihat Data. Data tidak bisa diedit.....	140
Gambar 5.11 Opsi Menu Ringkasan Statistik	141
Gambar 5.12 Dialog Ringkasan Numerik Suatu Peubah	141
Gambar 5.13 Frekuensi distribusi dengan kecocokan chi kuadrat	143
Gambar 5.14. Dialog Korelasi dengan tiga macam pilihan korelasi	144
Gambar 5.15 Dialog Histogram	146
Gambar 5.16 Contoh Tampilan Histogram	147
Gambar 5.17 Deskripsi Boxplot.....	148
Gambar 5.18 Dialog Boxplot dengan Pilihan Kelompok.....	149
Gambar 5.19 Boxplot untuk Data dengan Kelompok (5 kelompok).....	149
Gambar 5.20 Dialog Perbandingan Kuantil	151
Gambar 5.21 Grafik Perbandingan Kuantil (QQ Plot) yang dilengkapi sabuk keyakinan. Gambar menunjukkan sebaran data masih dalam batas sabuk keyakinan.	152

Gambar 5.22	Grafik QQ-Norm untuk Distribusi Simetris	153
Gambar 5.23	Grafik QQ-Norm untuk Distribusi Tidak Simetris.....	154
Gambar 5.24	Grafik QQ-Norm untuk Distribusi Seragam dan Diskrit..	155
Gambar 5.25	Dialog Plot Nilai-tengah	156
Gambar 5.26	Tampilan Plot Nilai-tengah Satu Arah (Sekolah).....	157
Gambar 5.27	Tampilan Plot Nilai-tengah Dua Arah (Sekolah dan Jenis Kelamin).....	158
Gambar 5.28.	Tampilan Dialog Diagram Pencar dengan Plot Kelompok	158
Gambar 5.29.	Tampilan Dialog Matriks Diagram Pencar	159
Gambar 5.30.	Contoh Tampilan Diagram Pencar Sederhana Tanpa Memperhatikan Kelompok dan Tanpa Boxplot Marjin ...	159
Gambar 5.31.	Contoh Tampilan Diagram Pencar dengan Memperhatikan Kelompok dan Boxplot Marjin.....	160
Gambar 5.32.	Contoh Tampilan Matriks Diagram Pencar untuk empat peubah dengan Grafik Kuantil pada Diagonal	161
Gambar 5.33.	Contoh Tampilan Matriks Diagram Pencar untuk empat peubah dengan Histogram pada Diagonal	162
Gambar 5.34.	Menu dan Dialog Grafik 3D	163
Gambar 5.35.	Beberapa tampilan Grafik 3D dari Data yang sama dengan sudut pandang berbeda	164
Gambar 6.1.	Menu dan Dialog Uji t sampel tunggal	174
Gambar 6.2	Ilustrasi Uji t satu Sampel untuk H_0 ditolak.....	177
Gambar 6.3	Ilustrasi Uji t satu Sampel untuk H_0 diterima.....	178
Gambar 6.4	Dialog Uji-t Dua Kelompok Saling Bebas	182
Gambar 6.5	Panel Uji t Dua Sampel.....	184

Gambar 6.6 Ilustrasi Hubungan t Hitung, IK dan Nilai-p untuk H_0 Diterima	184
Gambar 6.7 Ilustrasi Hubungan t Hitung, IK dan p value untuk H_0 ditolak	185
Gambar 6.8 Dialog Uji t berpasangan	190
Gambar 6.9 Visualisasi distribusi data kelompok homogen dengan nilai-tengah sama	194
Gambar 6.10 Ilustrasi distribusi data kelompok homogen dengan nilai-tengah tidak sama	195
Gambar 6.11 Dialog Anava satu arah.....	195
Gambar 6.12 Ilustrasi Perbandingan Rata-rata Berpasangan	200
Gambar 6.13 Contoh Grafik Rata-rata Satu Arah	201
Gambar 6.14 Dialog Anava Multi Arah	202
Gambar 6.15 Contoh Grafik Rata-rata Dua Arah.....	204
Gambar 6.16 Menu dan Dialog Uji Proporsi Satu Sampel.....	206
Gambar 7.1 Diagram pencar data (X,Y) dengan berbagai nilai korelasi	213
Gambar 7.2 Menu dialog Matriks Korelasi untuk korelasi Pearson dan Spearman.....	215
Gambar 7.3 Menu dan Dialog Uji Korelasi Pearson dan Spearman	216
Gambar 7.4 Ilustrasi Garis regresi populasi, regresi penduga dengan sabuk keyakinannya.	219
Gambar 7.5 Diagram Pencar Data Asli (sebelah kiri) dan Sisa (sebelah kanan).....	224
Gambar 7.6 Dialog regresi linier sederhana untuk peubah NFis dan NMat	225
Gambar 7.7 Diagram Pencar Data NMat dan Nfis dengan koefisien determinasi relatif rendah.....	228

Gambar 7.8 Diagram Pencar X dan Y	229
Gambar 7.9 Ilustrasi Garis regresi populasi, regresi prediksi nilai Y dan rata-rata Y pada saat nilai tertentu dari X dengan interval keyakinannya.....	231
Gambar 7.10 Dialog Model Linier	234
Gambar 7.11 Dialog Regresi Bertatar	236
Gambar 7.12 Menu dan Dialog Diagnostik Grafik	238
Gambar 7.13 Grafik Diagnostik untuk Model Regresi yang relatif memenuhi syarat kenormalan dan bebas dari pencilan.	240
Gambar 7.14 Grafik Diagnostik untuk Model Regresi yang kurang memenuhi syarat kenormalan dan mengandung pencilan	241
Gambar 7.15 Diagram Pencar untuk Data Lengkap (kiri) dan untuk Data tanpa Pengamatan yang Dianggap Pencilan.....	245

STRUKTUR BUKU

Buku ini utamanya dimaksudkan sebagai salah satu referensi untuk belajar analisis statistika yang mendasar (untuk Diploma dan S1) yang lebih banyak menggunakan pendekatan aplikasi. Namun, secara praktis, dewasa ini penggunaan komputer dalam analisis data menggunakan statistika sudah merupakan kebutuhan yang tidak bisa dihindarkan. Ada banyak piranti lunak (*software*) yang tersedia untuk membantu analisis data mulai dari *open source* yang gratis dan mudah dimodifikasi, sampai yang bersifat komersial. Ada banyak alasan untuk memperkenalkan pemanfaatan *open source*. Oleh karena itu pada bab awal dibahas secara ringkas piranti lunak R, khususnya pada beberapa fitur dan paket yang banyak dipergunakan dalam pembahasan selanjutnya. Fitur yang dipilih adalah yang dapat dimanfaatkan oleh pengguna statistika yang tidak memiliki dasar-dasar pemrograman yang memadai sehingga lebih banyak menggunakan menu grafis.

Di sisi lain, teori peluang adalah dasar penalaran statistika. Seseorang juga perlu memiliki secukupnya pemahaman tentang teori peluang untuk dapat menghargai argumen yang mendasari pengembangan metode statistika. Lebih lanjut, teori distribusi menjadi dasar penurunan dan pengembangan uji-uji statistika (khususnya parametrik), sehingga juga perlu dibahas agar penggunaan statistika sebagai alat untuk

menganalisis data, lebih memperhatikan asumsi-asumsi yang harus dipenuhi.

Pembahasan dalam buku ini tidak dimaksudkan agar pembaca menguasai semua rumus-rumus matematika yang diberikan untuk diterapkan dalam perhitungan statistika, karena pada kenyataannya perhitungan dapat dilakukan dengan bantuan komputer. Aspek yang lebih penting yang perlu dikuasai adalah aspek logika dari rumus tersebut, yang merupakan benang merah yang menghubungkan antara kondisi data dengan jenis-jenis atau alternatif-alternatif uji yang dapat dipilih. Dengan demikian hasil analisis dapat memberi manfaat sebagaimana mestinya. Beberapa konsep statistika yang bersifat abstrak juga dicoba dibuat ilustrasi dan visualisasinya sehingga diharapkan konsep statistiknya menjadi lebih mudah dipahami. Pembahasan selebihnya dalam buku ini lebih menggunakan pendekatan praktis uji statistika dasar dengan memanfaatkan piranti lunak khususnya R.

Secara keseluruhan buku ini dibagi menjadi tiga bagian besar yaitu

1. Bagian pertama membahas pengertian mendasar statistika dan peran komputer dalam statistika. Bagian ini mengulas hal-hal fundamental untuk dikuasai agar dapat memanfaatkan R untuk analisis data. Penjelasan yang lebih detail terkait prosedur dan teknis mengunduh dan melengkapi R (termasuk penggabungannya dengan Microsoft Excel) disampaikan pada bagian lampiran.
2. Bagian kedua membahas dasar-dasar teori peluang yang mendasari teori statistika. Dalam bagian ini juga dibahas teori distribusi Normal yang menjadi dasar pengembangan sebagian besar metode analisis data tingkat dasar. Dalam kelas, intensitas dan kedalaman pemahaman teori ini perlu disesuaikan dengan kondisi pengguna statistika.
3. Bagian ketiga membahas aplikasi R untuk menganalisis data, khususnya analisis yang telah dibahas pada bab-bab sebelumnya. Pada bagian ini utamanya dibahas dasar teori dan asumsi dari metode-metode analisis, cara memasukkan input pada R serta menafsirkan hasil yang dikeluarkan.

Buku ini dilengkapi dengan CD program R versi 3.02 beserta beberapa paket mendasar yang dianggap penting, terutama yang terkait dengan pemanfaatan menu RCommander. Bagi pembaca yang ingin

memperoleh versi yang lebih baru disarankan untuk mengunduh sendiri di internet. Kepada pembaca juga diberikan rujukan lebih lanjut jika ingin mendalami topik lebih jauh terutama dengan memanfaatkan R.

BAGIAN I PENGERTIAN UMUM

Pada bagian pertama ini akan dibahas hal-hal yang sangat mendasar baik yang terkait dengan konsep dasar statistika, maupun piranti lunak statistika, khususnya R. Pembahasan ini diharapkan dapat membantu mempermudah pemahaman dalam pembahasan pada bab-bab berikutnya.

BAB 1. PENGERTIAN DAN LINGKUP STATISTIKA

Dewasa ini statistika telah berkembang menjadi salah satu metode yang banyak dipakai dalam menganalisis data. Metode statistika dimanfaatkan oleh banyak kalangan mulai dari guru, peneliti, baik yang profesional maupun tidak. Masyarakat juga banyak disugahi hasil analisis statistika baik yang sederhana (informasi tentang proporsi, rata-rata, tren) maupun yang rumit yang hanya bisa dipahami oleh orang-orang yang bergelut dalam bidang bersangkutan. Dalam bab ini akan dibahas hal-hal yang sangat mendasar terkait statistika yang dapat memberikan gambaran tentang hakekat dan ruang lingkup statistika.

TUJUAN BELAJAR

Setelah menyimak uraian pada bab ini pembaca diharapkan memiliki gambaran tentang statistika pada umumnya, menyebutkan manfaat statistika dalam berbagai bidang, menyebutkan manfaat komputer dalam statistika serta tumbuh sikap disiplin serta jujur dalam memanfaatkan statistika dalam bidangnya.

MATERI

1. Pengertian statistika
2. Ruang lingkup statistika
3. Manfaat dan aplikasi Statistika
4. Penyalahgunaan Statistika
5. Peranan komputer dalam statistika
6. Notasi dan istilah penting dalam statistika

1.1 STATISTIK DAN STATISTIKA

Statistik dan statistika adalah dua istilah yang saling berkaitan tetapi memiliki makna yang sedikit berbeda. Statistik berkaitan dengan angka-angka yang merupakan informasi penting dari suatu data, (misalnya maksimum, minimum, rata-rata dan sejenisnya). Statistika berkaitan dengan ilmu dan prosedur yang mempelajari pemanfaatan angka-angka statistik untuk menarik kesimpulan.

Istilah statistik berakar dari kata-kata yang berarti negara (dalam bahasa Latin: *statisticum collegium*=dewan negara, dalam bahasa Italia: *statista*=negarawan atau pilitikus, dalam bahasa Inggris *state*=negara), karena pada awalnya statisik dipergunakan untuk menggambarkan angka-angka yang terkait dengan kepentingan negara, seperti masalah ekonomi dan kependudukan. Penggunaan istilah statistika dalam arti pengumpulan dan pengolahan data pertama kali diperkenalkan pada abad 19 oleh Sir John Sinclair. Setelah itu statistika berkembang dan mulai banyak menggunakan matematika terutama teori peluang. Para ilmuwan yang banyak berkontribusi dalam teori peluang dan statistika diantaranya adalah Fisher (dasar-dasar statistika inferensial), Gosset (distribusi t) dan Pearson (regresi).

Untuk lebih memahami prinsip dasar statistika ada baiknya kita mengikuti beberapa definisi formal tentang statistika. Menurut *Webster's New Collegiate Dictionary* statistika didefinisikan

sebagai “cabang matematika yang berkaitan dengan pengumpulan, analisis, interpretasi, dan penyajian dari sejumlah data numerik”. Sedangkan menurut Mendenhall [26] statistika adalah suatu “bidang sains yang berkaitan dengan ekstraksi informasi dari data numerik dan menggunakannya untuk membuat keputusan tentang populasi dari mana data tersebut diperoleh.” Tujuan statistika adalah untuk membuat kesimpulan tentang suatu yang lebih luas (disebut populasi) berdasarkan keterangan yang ada pada contoh (disebut sampel) yang diambil dari populasi tersebut. *Teori statistika adalah suatu teori informasi yang berhubungan dengan pengangkaan informasi, menentukan percobaan atau prosedur untuk pengumpulan data, dengan biaya minimal, dari sejumlah informasi tertentu, dan menggunakan informasi ini untuk membuat kesimpulan-kesimpulan.* Pembuatan kesimpulan terhadap populasi yang tidak diketahui adalah prosedur yang terdiri atas dua langkah. Pertama, kita menentukan prosedur-prosedur penarikan kesimpulan yang cocok dari situasi yang dihadapi; dan kedua, kita mencari ukuran kecocokan dari kesimpulan yang dihasilkan.

Statistika dapat juga dikatakan sebagai ilmu yang mempelajari prosedur pengambilan keputusan dari sejumlah kecil informasi pada sebagian data (yang disebut sampel), untuk digeneralisasikan pada lingkungan yang lebih luas yaitu populasi. Prosedur ini meliputi, mengorganisir, meringkas, menyajikan data dan menguji dugaan. Informasi-informasi (yang umumnya berupa angka) yang diekstrak dari data disebut statistik. Nilai maksimum, minimum, rata-rata, korelasi, variansi sampel adalah sebagian dari nilai-nilai statistik.

1.1.1 RUANG LINGKUP STATISTIKA

Cakupan statistika sangat luas, berdasarkan cakupannya statistika dibedakan menjadi dua kelompok besar yaitu.

1. **Statistika Deskriptif** yaitu prosedur yang terkait dengan eksplorasi dan representasi data yang meliputi mengorganisir, meringkas dan menyajikan data, termasuk memeriksa distribusi data. Dengan berkembangnya komputer, statistika

deskriptif, terutama representasi data secara grafik, merupakan prosedur yang banyak dipakai sebelum melakukan pengolahan data lebih lanjut. Prosedur ini dikenal dengan EDA (*Exploratory Data Analysis*).

2. **Statistika Inferensial** yaitu prosedur lebih lanjut terkait dengan pengolahan data untuk pendugaan, peramalan atau pengujian dugaan. Sesuai dengan teori yang mendasari analisis statistika, jenis prosedur pengolahan data yang dapat diterapkan sangat bergantung pada kondisi data (misalnya jenis distribusi data). Berdasarkan asumsi yang dipergunakan, statistika inferensial dibedakan mejadi dua kelompok besar yaitu
 - a. **Statistika Parametrik**, yaitu prosedur yang didasarkan atas asumsi bahwa data diambil secara acak dan menyebar dengan distribusi tertentu (misalnya Normal, Poisson, Binomial dan sebagainya).
 - b. **Statistika nonparametrik**, yaitu prosedur yang hanya didasarkan atas asumsi bahwa data diambil secara acak tanpa ada asumsi khusus tentang distribusinya. Oleh karena itu, statistika nonparametrik sering juga disebut sebagai statistika bebas distribusi (*distribution-free statistics*).

Eksplorasi lebih jauh

1. Cari 3 sampai 4 referensi tentang metode statistika, baik dalam bentuk cetak di perpustakaan maupun dalam bentuk *online* di internet. Sebutkan dan simpulkan definisi statistika menurut referensi yang anda baca tersebut.
 2. Cari 3 ilmuan yang memiliki kontribusi terhadap perkembangan statistika dan uraikan secara ringkas kontribusinya.
-

-
3. Diskusikan dengan teman-teman anda apa yang dimaksud dengan istilah-istilah: data, statistika, statistik, statistika deskriptif dan statistika inferensial.
-

1.1.2 MANFAAT DAN APLIKASI STATISTIKA

Metode statistika dapat dijadikan sebagai salah satu alat untuk mengambil keputusan atau kebijakan yang didasarkan atas data. Data yang menjadi dasar pengambilan keputusan tidak harus berupa seluruh data yang ada di lapangan. Dengan prosedur tertentu kita dapat menggunakan hanya sebagian dari data (yang disebut sampel) untuk mengambil kesimpulan yang berlaku untuk keseluruhan (yang disebut populasi). Hal ini memungkinkan kita menghemat biaya dan waktu dalam pengumpulan data. Karena informasi hanya didasarkan atas sampel yang merupakan sebagian dari keseluruhan, statistika dilengkapi dengan informasi kesalahan ataupun tingkat kecocokan (*goodness of fit*) dari hasil yang diperoleh. Prosedur ini memungkinkan seseorang untuk mengambil kesimpulan atau keputusan dengan kesalahan yang dapat dikendalikan dan dengan biaya yang lebih murah, serta waktu yang lebih singkat. Sebagai contoh suatu lembaga penghitung pemilih dalam pemilihan umum dapat dengan cepat menggambarkan hasil pemilihan umum dengan tingkat akurasi yang diklaim sampai 99%.

Aplikasi metode statistika sangat luas, di antaranya dalam bidang politik, keamanan, ekonomi, kesehatan industri dan lain-lain. Dalam bidang politik dan keamanan, paling tidak pada setiap akhir tahun media masa selalu menyampaikan berbagai angka statistik (misalnya popularitas seorang politisi, jumlah kecelakaan di jalan raya, jumlah kejahatan dan trennya dari tahun ke tahun). Industri kosmetik dapat menggunakan statistika untuk menggambarkan persentase masyarakat yang memanfaatkan produk-produknya. Industri otomotif dapat memanfaatkan statistika khusus (kendali mutu/*quality control*) untuk menjaga kualitas produknya.

1.1.3 PENYALAHGUNAAN STATISTIKA

Sebagaimana telah dijelaskan sebelumnya bahwa statistika merupakan salah satu cara untuk mengolah data menjadi informasi yang selanjutnya dapat dijadikan sebagai dasar untuk memprediksi atau menyimpulkan sesuatu. Prosedur-prosedur statistika diturunkan secara deduktif (mengikuti kaidah-kaidah matematika dan logika), ini menjamin bahwa prosedur statistika adalah logis dan valid. Sementara itu pemanfaatan data sebagai dasar pengambilan keputusan juga memungkinkan statistika menghasilkan keputusan yang lebih objektif dan realistik.

Namun dalam kenyataannya, dalam prakteknya, sengaja atau tidak, tidak jarang terjadi seseorang menggunakan statistika tanpa mentaati persyaratan penting dalam penggunaan statistika. Ada beberapa penyalahgunaan yang mungkin terjadi diantaranya adalah seperti berikut ini.

1. Tidak mengambil data secara acak dan menganalisis hanya sebagian data yang dirasakan mendukung pendapat tertentu dan mengabaikan yang lainnya.
2. Hanya membandingkan atau menginformasikan nilai rata-rata dan menggunakannya untuk menarik kesimpulan tanpa menggunakan informasi tentang penyebaran atau kesalahan baku. Baird [1] menyatakan hal ini sebagai *naked statistics*. Sebagai ilustrasi, diketahui rata-rata kelompok A dan kelompok B masing-masing adalah 50 dan 60. Lalu, seseorang menyatakan bahwa kedua kelompok itu sangat berbeda. Sesungguhnya, hanya dengan informasi rata-rata saja, kita belum bisa menarik kesimpulan tentang populasi kelompok A dan B dengan menggunakan statistika.
3. Menampilkan grafik dengan skala yang tidak proporsional sedemikian sehingga memberikan gambaran (kesan) yang tidak sesuai dengan kenyataannya.

Hal yang perlu disadari adalah bahwa statistika sebagai suatu ilmu atau prosedur bersifat netral, tidak bisa berbohong, namun orang yang memanfaatkannya dapat dengan sengaja atau tidak

menyalahgunakannya untuk mendukung kepentingan tertentu. Mendenhall [26] menyatakan bahwa salah satu penyebab terjadinya penyalahgunaan statistika adalah karena praktek statistika tidak diatur regulasi dan setiap orang dapat menganggap dirinya sebagai statistikawan.

Eksplorasi lebih jauh

Buat kelompok dengan 2-3 anggota lalu diskusikan masalah berikut ini.

1. Eksplorasi di internet maupun di media masa contoh penyajian statistika (grafik atau tabel) yang dapat dikategorikan sebagai penggunaan statistika yang tidak semestinya
 2. Jelaskan adakah dampak penyalahgunaan statistika terhadap apresiasi masyarakat kepada metode statistika?
 3. Uraikan saran anda untuk meminimalkan penyalahgunaan statistika
-

1.1.4 PERANAN KOMPUTER DALAM STATISTIKA

Pada umumnya, dalam menganalisis data, kondisi data yang dihadapi sedemikian kompleks, baik dari segi distribusi, ukuran data, maupun banyaknya peubah yang terlibat. Oleh karena itu melakukan analisis secara manual (atau hanya dengan bantuan kalkulator) tidak lagi efektif. Dalam era sekarang, analisis data tidak bisa lagi dipisahkan dengan komputer. Dengan berkembangnya teknologi komputer dan khususnya piranti lunak statistika, konsentrasi pengolah data tidak lagi pada keterampilan menghitung atau sekedar menerapkan rumus yang ada. Konsentrasi pengguna statistika sekarang lebih pada pemahaman konsep, eksplorasi alternatif-alternatif prosedur atau metode yang lebih sesuai dengan kondisi data serta memberi interpretasi pada hasil yang diperoleh. Bahkan, untuk sampel yang relatif kecil dimana asumsi distribusi tidak bisa dipenuhi, analisis berbasis simulasi komputer (*CIS-Computer Intensive Statistics*) telah menjadi

alternatif analisis yang banyak dipilih selain menggunakan metode statistika nonparametrik. Paket komputer umumnya, dan khususnya paket analisis data, juga dilengkapi dengan kemampuan grafis yang memungkinkan seseorang memvisualisasikan data atau hasil analisisnya, atau bahkan memvisualisasikan konsep-konsep statistika yang abstrak, sehingga menjadi lebih mudah dipahami.

Ada banyak paket program statistika (misalnya S-Plus, Minitab, SPSS), namun sebagai ilustrasi, buku ini secara khusus memanfaatkan program statistika R yang berbasis *open source*. Ada beberapa alasan memilih program R diantaranya, sebagai *open source*, (i) R dapat diperoleh secara cuma-cuma, (ii) dapat dengan mudah dimodifikasi sesuai kebutuhan baik untuk pembelajaran maupun analisis data, (iii) telah tersedia menu dan dialog dalam Bahasa Indonesia, (iv) sangat mirip dengan S-Plus sehingga mudah bagi pembaca yang pernah menggunakan S-Plus. Untuk memudahkan pembahasan bab-bab selanjutnya, pada Bab 2 akan dibahas secara ringkas R dan beberapa paket terkait diantaranya RCommander dari Fox dan StatDemo dari Tirta [47] yang sudah di-*plugin* ke menu RCommander.

1.2 NOTASI DAN ISTILAH PENTING DALAM STATISTIKA

Pada bagian ini akan diuraikan beberapa notasi dan istilah yang banyak dipergunakan dalam statistika umumnya dan khususnya dalam buku ini.

1.2.1 NOTASI PENTING

Beberapa notasi penting yang banyak digunakan dalam metode statistika di antaranya adalah notasi jumlah, produk, dan rata-rata.

1. Jumlah (*sum*) beberapa peubah, misalnya x_1, x_2, \dots, x_n dinotasikan dengan notasi sigma besar dan didefinisikan seperti berikut ini.

$$D = x_1 + x_2 + \dots + x_n = \sum_{i=1}^n x_i$$

2. Rata-rata (*average*), sering juga disebut sebagai **rata-rata aritmatik**, dari peubah x_1, x_2, \dots, x_n didefinisikan dan dinotasikan seperti berikut ini.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \sum_{i=1}^n \frac{x_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

3. Hasil kali (*product*) peubah x_1, x_2, \dots, x_n dinotasikan dengan notasi pi besar dan didefinisikan seperti berikut ini.

$$P = x_1 \times x_2 \times \dots \times x_n = \prod_{i=1}^n x_i$$

4. **Rata-rata geometrik** peubah x_1, x_2, \dots, x_n didefinisikan sebagai

$$\sqrt[n]{x_1 \times x_2 \times \dots \times x_n}$$

Sifat-sifat

Jika k, k_1, k_2 adalah konstanta dan $x_i, y_i, i=1, 2, \dots, n$ adalah peubah maka berlaku

1. $\sum_{i=1}^n k = nk$

2.
$$\sum_{i=1}^n kx_i = k \sum_{i=1}^n x_i$$
3.
$$\sum_{i=1}^n (k_1x_i + k_2y_i) = k_1 \sum_{i=1}^n x_i + k_2 \sum_{i=1}^n y_i$$
4.
$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 = \sum_{i=1}^n (x_i^2 - \bar{x}^2)$$
5.
$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} = \sum_{i=1}^n x_i y_i - \frac{1}{n} \left(\sum_{i=1}^n x_i \sum_{i=1}^n y_i \right)$$
6.
$$\prod_{i=1}^n k_1 x_i = k_1^n \prod_{i=1}^n x_i$$

Berikut diuraikan bukti untuk no 4, selebihnya pembaca dipersilahkan membuktikan sendiri baik secara berkelompok maupun perorangan.

Buktikan bahwa

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 = \sum_{i=1}^n (x_i^2 - \bar{x}^2)$$

Bukti

$$\begin{aligned}
\sum_{i=1}^n (x_i - \bar{x})^2 &= \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) && \text{uraian bentuk kuadrat} \\
&= \sum_{i=1}^n (x_i^2) - 2\bar{x} \left(\sum_{i=1}^n x_i \right) + n\bar{x}^2 && \bar{x} \text{ sebagai konstanta} \\
&= \sum_{i=1}^n (x_i^2) - 2\bar{x}(n\bar{x}) + n\bar{x}^2 \\
&&& \text{hubungan antara } \bar{x} \text{ dengan } \sum_{i=1}^n x_i \\
&= \sum_{i=1}^n (x_i^2) - n\bar{x}^2 \\
&= \sum_{i=1}^n (x_i^2 - \bar{x}^2) \\
&= \sum_{i=1}^n (x_i^2) - n \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2 \\
&= \sum_{i=1}^n (x_i^2) - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2
\end{aligned}$$

Bertikut adalah beberapa contoh penggunaan notasi di atas.

Contoh 1.1.

Diketahui x_i untuk $i = 1, 2, \dots, 5$ dengan nilai seperti berikut: 10, 8,

4, 5, 9. Tentukan $\sum_{i=1}^n x_i$, $\sum_{i=1}^n x_i^2$, \bar{x} ,

Penyelesaian

$$1. \sum_{i=1}^n x_i = 10 + 8 + 4 + 5 + 9 = 36$$

$$2. \sum_{i=1}^n x_i^2 = 10^2 + 8^2 + 4^2 + 5^2 + 9^2 = 100 + 64 + 16 + 25 + 36 = 241$$

$$3. \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{36}{5} = 7\frac{1}{5}$$

Contoh 1.2

Dari soal pada contoh sebelumnya, tentukan $\prod_{i=1}^5 x_i$ dan $\overline{x^2}$

Jawab:

$$1. \prod_{i=1}^5 x_i = 10 \times 8 \times 4 \times 5 \times 9 = 14400$$

$$2. \overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2 = \frac{241}{5} = 48\frac{1}{5}$$

Untuk jumlah berpasangan, perhatikan, bahwa secara umum

$$\sum_{i=1}^n x_i y_i \neq \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)$$

Contoh 1.3

Sebagai ilustrasi perhatikan pasangan nilai x_i dan y_i seperti pada tabel berikut

I	x_i	y_i	$x_i y_i$
1	5	10	50
2	10	15	150
3	8	10	80
4	6	12	72
5	4	6	24
Total	33	53	376

$$= \sum_{i=1}^n \quad = \sum_{i=1}^5 x_i \quad = \sum_{i=1}^5 y_i \quad = \sum_{i=1}^5 x_i y_i$$

Dari nilai di atas terlihat bahwa $\sum_{i=1}^5 x_i y_i = 376$, tidak sama dengan $\sum_{i=1}^5 x_i = 33$ dikalikan dengan $\sum_{i=1}^5 y_i = 53$ (yaitu 1749).

1.2.2 ISTILAH PENTING

Ada beberapa istilah yang sering muncul dalam pembahasan metode statistika. Berikut adalah istilah-istilah yang akan sering muncul dan perlu dijelaskan sebelum membahas lebih jauh tentang metode statistika. Istilah-istilah lain yang lebih spesifik akan dijelaskan kemudian atau dapat dilihat pada bagian glosarium dari buku ini.

1. **Populasi** adalah seluruh data atau hasil observasi yang menjadi kepentingan. Menurut Mendenhall [26] populasi dalam statistika adalah kumpulan data bukan kumpulan orang pada wilayah tertentu.

2. **Sampel** adalah sebagian dari populasi yang diambil untuk diolah. Sampel diambil sedemikian rupa sehingga dapat mewakili populasi atau merupakan miniatur dari populasi.
3. **Sampel acak** adalah sampel yang diambil sedemikian sehingga setiap subjek dalam populasi sejenis memiliki peluang yang sama untuk dipilih.
4. **Metode/teknik sampling** adalah prosedur pengambilan sampel dari populasi sehingga diperoleh sampel yang mewakili populasi.
5. **Peubah** atau **variabel** adalah karakteristik dari suatu subjek yang bervariasi dari suatu subjek dengan subjek lainnya. Misalnya jika subjeknya mahasiswa, karakteristik yang dapat menjadi peubah di antaranya adalah jenis kelamin, tinggi badan, nilai ujian. Peubah secara umum dapat dibedakan menjadi peubah kualitatif (misalnya: jenis kelamin dan asal daerah) dan peubah kuantitatif (misalnya nilai ujian). Peubah kualitatif sering disebut sebagai faktor.
6. **Hipotesis** adalah pernyataan tentang parameter populasi yang akan diuji kebenarannya
7. **Data** adalah kumpulan hasil pengamatan yang terdiri atas beberapa peubah. Data dapat sekaligus terdiri atas peubah kuantitatif dan peubah kualitatif.
8. **Mean** adalah nilai yang menunjukkan rata-rata populasi. Istilah mean biasanya diterjemahkan dengan istilah rerata atau nilai tengah.
9. **Variansi** adalah kuadrat dari simpangan baku istilah lain dari variansi yang banyak dipergunakan adalah ragam. Perlu diperhatikan perbedaan rumus ragam populasi dan ragam sampel.

1.3 DATA

1.3.1 DATA DAN INFORMASI

Data umumnya diartikan sebagai kumpulan hasil pengamatan tentang atribut dari suatu subjek. Misalnya untuk subjek mahasiswa, atribut yang dapat menjadi perhatian di antaranya adalah tinggi badan, berat badan, jenis kelamin, nilai suatu matakuliah, IP (Indeks Prestasi) semester dan lain sebagainya. Karakteristik yang menjadi perhatian biasa disebut sebagai peubah. Hasil observasi dari masing-masing peubah dapat berupa angka (kuantitatif) atau berupa label pengelompokan, atau deskriptor lain (misalnya laki-perempuan, dalam kota dan luar kota). Jadi data dapat pula diartikan sebagai kumpulan dari peubah-peubah beserta deskriptornya.

Dari data, baik kualitatif maupun kuantitatif, selanjutnya diperoleh informasi yang diperlukan, dengan bantuan metode statistika. Beberapa informasi yang sederhana yang banyak dipergunakan di antaranya adalah tentang ringkasan statistik (minimum, maksimum, nilai-tengah, ragam dan lainn-lainnya), distribusi data baik secara keseluruhan maupun berdasarkan kelompok yang menjadi perhatian.

1.3.2 JENIS DAN SKALA DATA

Dilihat dari sumbernya, data dibedakan atas dua jenis yaitu data riil dan data simulasi/ data bangkitan. Data riil diperoleh dari hasil penelitian dengan mengikuti prosedur yang telah ditetapkan. Data riil ini ada yang dikumpulkan langsung oleh peneliti (primer), ada yang diperoleh dari instansi atau lembaga lain (sekunder). Data simulasi diperoleh dengan membangkitkannya melalui program komputer. Data simulasi sangat baik dipergunakan untuk mempelajari atau memahami prosedur statistika, karena besarnya, sifat-sifatnya (parameter dan jenis distribusinya) dapat dikendalikan.

Data yang disajikan dalam bentuk angka (numerik) pada umumnya dapat dibedakan menjadi dua kelompok besar, yaitu data hasil **pencacahan** (disebut **data enumerasi**) dan data dari hasil **pengukuran**. Pengukuran (*measurement*) menghasilkan data numerik yang disebut **data metrik**. Secara umum data ini dapat dikelompokkan kedalam beberapa jenis skala yaitu **nominal**,

ordinal, interval dan rasio yang dapat dijelaskan seperti berikut ini (Lihat Gravetter & Wallnau [10], Guilford & Fruchter [11]).

1. **Nominal.** Sering hasil pengukuran kategori disajikan dalam bentuk angka misalnya jenis kelamin, di label 0 untuk perempuan dan 1 untuk laki-laki. Dalam hal ini, angka 0 dan 1 hanyalah label atau angka dan bukan bilangan. Oleh karena itu angka-angka ini tidak dapat dimanipulasi secara aritmatik dan statistika dalam arti tidak dapat dijumlah maupun dirata-rata. Namun, kita bisa berbicara frekuensi atau proporsi masing-masing kategori.
2. **Ordinal.** Dalam beberapa hal, hasil pengukuran kontinu dikelompokkan menjadi beberapa kategori yang termasuk kategori kualitatif. Misalnya dalam hasil belajar, bukan perolehan angkanya yang diperhatikan tetapi rankingnya (misalnya sangat kurang, kurang, cukup, baik sangat baik). Walaupun ranking kualitatif ini ditransformasikan ke bentuk angka (misalnya 0-4), sesungguhnya angka-angka tersebut belum bisa dijumlah atau dirata-rata, dia hanya bisa diurut. Demikian juga sesungguhnya perubahan skor ujian dari skala asli (0-100) ke skala lima (0-4) sesungguhnya menghasilkan skala ordinal yang bisa dilabel E-A. Sesungguhnya angka-angka ini juga belum bisa dimanipulasi secara aritmatik (dalam arti dijumlah, dirata-rata dan sejenisnya). Skala ordinal belum memiliki unit pengukuran yang sama, yaitu jarak antara 0 ke 1 belum tentu sama dengan jarak 1 ke 2 dan seterusnya.
3. **Interval.** Pada skala interval, angka-angka sudah memiliki makna bilangan dan dapat dimanipulasi secara aritmatik (dijumlah, dikurangkan dan sebagainya). Contoh pengukuran dengan skala interval adalah skor hasil ujian. Skala ini sudah memiliki unit pengukuran yang sama (misalnya skor 65, 70, dan sebagainya). Skala ini juga sudah memiliki angka 0, namun belum mutlak. Artinya, seseorang yang mendapat skor 0 tidak berarti tidak memiliki pengetahuan sama sekali. Demikian juga suhu 0 derajat tidak berarti suatu benda tidak memiliki panas atau suhu sama sekali.

4. **Rasio.** Rasio merupakan skala yang paling sempurna. Skala ini telah memiliki 0 mutlak, dapat dibandingkan secara rasio. Contoh misalnya berat suatu benda. Dalam hal berat, bilangan 0 menunjukkan tidak ada berat yang diukur. Suatu objek yang berat 10 satuan dapat dikatakan beratnya dua kali berat objek yang beratnya 5 satuan. Rasio berat objek yang beratnya 75 unit dan 50 unit adalah 3:2 dan seterusnya.

Pencacahan umumnya dilakukan pada populasi yang bersifat kategorik, baik yang bersifat kualitatif maupun kuantitatif. Kategori kualitatif misalnya kelompok masyarakat kaya-miskin, daerah maju-tertinggal. Kategori kuantitatif biasanya aslinya merupakan distribusi kontinu yang dikelompokkan menjadi beberapa kategori misalnya hasil ujian dikelompokkan mejadi lulus-tidak lulus. Selain itu, ada kalanya data memuat informasi yang merupakan karakteristik kelompok yang bersifat alamiah. Kelompok ini tidak menunjukkan perbedaan tingkat kualitas satu dengan lainnya misalnya: jenis kelamin (Laki-Perempuan), jenis pekerjaan, kebangsaan dan sebagainya. Untuk data kategorik, hasil pencacahan dapat disajikan dalam bentuk frekuensi atau proporsi.

Dalam kenyataannya, suatu data sampel dapat terdiri atas beberapa peubah yang masing-masing memiliki skala berbeda, misalnya data mahasiswa dapat terdiri atas: (i) jenis kelamin dan agama yang bersifat nominal, (ii) nilai matakuliah yang bersifat ordinal, (A-E) atau interval (0-100), (iii) umur, tinggi dan berat badan yang bersifat rasio dan sebagainya. Data dapat berisi berbagai informasi yang dapat dikelompokkan sebagai karakteristik populasi yang menjadi perhatian atau kepentingan studi (misalnya, jenis kelamin, pendidikan, tinggi badan dan sebagainya)

Eksplorasi lebih jauh

1. Diskusikan dengan teman-teman anda (2-3 orang), beri contoh minimal 2 fenomena yang datanya harus dikumpulkan melalui pengukuran, dan minimal 2 fenomena yang datanya harus dikumpulkan melalui pencacahan.

2. Beri ilustrasi masing-masing minimal 2 contoh penggunaan angka yang merupakan skala nominal, ordinal, interval dan rasio.
-

1.3.3 PRESENTASI DAN ANALISIS DATA

Dengan bantuan komputer beserta piranti lunak yang sesuai, penyajian data dapat dilakukan dengan cara yang lebih menarik dan variatif. Hal ini akan memudahkan untuk melihat gambaran data secara umum. Selanjutnya peneliti akan mendapat gambaran metode/teknik statistika yang sesuai dengan keperluan dan kondisi data yang dihadapi. Pada Bab 5 akan dibahas beberapa variasi cara menyajikan data baik secara numerik maupun grafis. Namun, karena dalam melakukan analisis data sangat bergantung pada teori peluang dan statistika, maka pada sebelum membahas analisis data akan dibahas terlebih dahulu beberapa pengetahuan dan keterampilan yang dibutuhkan diantaranya terkait dengan pegerasian piranti lunak R dan hal mendasar dalam teori peluang.

1.4 RINGKASAN DAN BACAAN LEBIH LANJUT

1.4.1 RINGKASAN

1. Statistika dan statistik adalah dua istilah yang memiliki makna berbeda. Statistik berkaitan dengan angka-angka yang merupakan informasi penting dari suatu data, sedangkan statistika berkaitan dengan ilmu dan prosedur yang mempelajari pemanfaatan angka-angka statistik untuk menarik kesimpulan.
2. Metode statistika dibedakan menjadi dua kelompok besar yaitu statistika deskriptif yang hanya mendeskripsikan data dan statistika inferensial yang memanfaatkan informasi pada data untuk menarik kesimpulan untuk subjek yang lebih luas.

3. Metode statistika diturunkan secara deduktif-matematis atas dasar asumsi tertentu sehingga bersifat logis dan valid. Oleh karena itu penggunaan statistika perlu memperhatikan asumsi-asumsi yang mendasari agar hasilnya dapat lebih bermanfaat.
4. Metode statistika yang diturunkan dengan asumsi bahwa data memiliki distribusi tertentu dengan parameter yang tidak diketahui disebut metode statistika parametrik.
5. Dewasa ini metode statistika tidak bisa dipisahkan dengan komputer. Dengan bantuan aplikasi komputer di bidang statistika, pengguna statistika dapat mengeksplorasi data serta mencoba berbagai alternatif uji statistika yang lebih sesuai dengan kondisi data yang dihadapi.
6. Data adalah kumpulan hasil observasi dari peubah-peubah beserta deskriptornya.
7. Informasi dari data adalah karakteristik penting dari suatu data yang umumnya merupakan statistik dari data (ukuran pemusatan, ukuran penyebaran, jenis distribusi dan sebagainya)
8. Skalanya data metrik dapat dibedakan atas skala nominal, ordinal, interval dan rasio.

1.4.2 BACAAN LEBIH LANJUT

Buku ini tidak dimaksudkan untuk membahas semua tentang statistik secara detil. Dalam buku ini hal-hal mendasar berusaha dijelaskan dan diilustrasikan dengan menggunakan komputer. Pembaca disarankan agar juga membaca referensi lain tentang metode statistika terutama yang bersifat aplikatif. Referensi yang berbahasa Indonesia yang dapat dibaca diantaranya Hadi [12], Sudjana [40], sedangkan yang berbahasa Inggris diantaranya Guilford & Fruchter [11], Gravetter & Wallnau [10], Mendenhall [26]. Ada juga referensi elektronik yang tersedia di internet diantaranya NIST [31], Statsoft [32]. Referensi *on-line* terkait

dengan aplikasi R diantaranya Kuhnert & Venables [19], Maindonald [23], Vezalini [52], Zoonekyn [55], dan Faraway [7]. Khusus untuk regresi linier tergeneralisir (GLM) dapat dibaca McCullagh & Nelder [24].

BAB 2. SELAYANG PANDANG R

Komputer dan statistika dewasa ini sangat sulit dipisahkan. Paket komputer tidak saja membantu dalam menganalisis data yang membutuhkan perhitungan rumit, pemanfaatan program komputer juga sangat membantu dalam mengilustrasikan dan memvisualisasikan konsep-konsep statistika yang bersifat abstrak dan sulit dipahami. Pada bab ini akan dibahas secara umum program R. Selain membahas kemampuan R untuk analisis data, pembahasan juga diberikan terkait kemampuan R yang dapat dimanfaatkan sebagai alat bantu pembelajaran statistika. Dalam buku ini sengaja dipilih kemampuan R yang dapat dimanfaatkan oleh pembaca yang tidak cukup kuat kemampuan komputernya, yaitu yang lebih banyak menggunakan pendekatan menu grafis dari pada skrip program.

TUJUAN BELAJAR

Setelah menyimak uraian pada bab ini diharapkan pembaca dapat mengaktifkan dan mengeksplorasi menu Rcommander, R-excel serta Rshiny.

MATERI

1. Fitur Umum R
2. Fitur R-Commander asli dan Termodifikasi
3. Fitur Teaching Demos dan Stat Demo
4. Fitur RExcel

Pedoman untuk mengunduh dan menginstal R dengan paket-paket yang diperlukan dapat dilihat pada Lampiran.

2.1 FITUR UMUM R

R adalah paket analisis data dan presentasi grafik yang merupakan program *open source* dan termasuk keluarga bahasa S. Bahasa R mempunyai struktur bahasa yang sama dengan S-Plus, beda utamanya adalah S-Plus dikembangkan secara komersial, sedangkan R dikembangkan sebagai *open source*. Program R ini sudah dilengkapi banyak kemampuan internal untuk menganalisis data maupun menampilkan grafik. Bahkan, **kemampuan grafik merupakan keunggulan menonjol dari R maupun S-Plus dibanding paket lainnya**. Dewasa ini R populer dipergunakan baik dibidang akademik maupun industri. Sebagai *open source*, R bersifat *multi platform* (tersedia untuk sistem operasi **Windows, Mac dan Linux**) dan didukung oleh banyak programmer maupun statistisi yang tersebar di seluruh dunia. *Source* yang terkait dapat diperoleh di internet dan diijinkan untuk dimodifikasi sesuai keperluan.

Ada beberapa alasan R banyak dipilih oleh para statistisi di seluruh dunia diantaranya:

1. Gratis. Kondisi ini meyakinkan para statistisi/statistikawan, bahwa dimanapun mahasiswa berada dan bekerja nantinya, mereka tidak perlu khawatir akan tersedianya paket statistika

yang mereka pelajari. Sementara paket-paket komersial yang berharga mahal, walaupun juga bagus tetapi belum tentu perusahaan dan institusi tempat mereka bekerja akan mampu menyediakannya secara legal.

2. Fleksibel. Sebagai *open source*, mahasiswa yang memiliki keterampilan komputasi dapat secara fleksibel memenuhi kebutuhannya dengan memanfaatkan.
3. Secara spesifik memenuhi kebutuhan bidang, bidang tertentu seperti dinyatakan oleh McCullough [24] “*R is the best choice for teaching econometric/statistical computing to researchers who are not numerical analysts*”.

Pada dasarnya R berinteraksi secara CLI (*Command Line Interface*), yaitu dengan membaca perintah yang dikirim melalui Jendela R-Console. Untuk membantu pengguna yang biasa bekerja dengan menu, GUI (*Graphical User Interface*), berbagai macam menu telah dikembangkan oleh berbagai kelompok, salah satu diantaranya adalah RCommander (paket Rcmdr). Saat ini, menu RCommander telah tersedia dalam bahasa Indonesia. Selain itu ada juga GUI **SciViews/Tinn-R** bagi Programmer/Pengguna statistika khusus/lanjut. Berbagai kebutuhan analisis statistika yang banyak dipakai, telah diimplementasikan ke dalam menu GUI, salah satunya adalah RCommander.

Sesungguhnya R bukanlah program khusus untuk statistika, tetapi merupakan bahasa pemrograman yang didukung oleh banyak paket (*library*) yang terkait dengan analisis data dan ilustrasi grafik. Namun, sebagian besar analisis data menggunakan statistika lanjut hanya bisa dimanfaatkan melalui CLI.

2.2 FITUR RCOMMANDER

Sesungguhnya RCommander bukanlah satu-satunya GUI untuk R, tetapi RCommander dipilih dengan berbagai pertimbangan diantaranya:

1. memiliki tampilan yang sama baik pada versi windows maupun pada versi linux; pengembangan pada versi windows juga langsung dapat diaplikasikan pada versi linux;
2. telah tersedia menu dan dialog dalam bahasa Indonesia;
3. Fleksibel dikembangkan dengan mengaktifkan menu plug.in;
4. tersedia integrasi dengan Microsoft excel melalaui paket R-Excel.

Beberapa analisis tingkat dasar dan menengah sudah bisa dilakukan melalui menu **Rcommander** diantaranya seperti berikut ini.

1. **Impor/Ekspor Data** yang berasal dari SPSS, Minitab, Excel, maupun yang berupa ascii (txt).
2. **Eksplorasi dan Penyajian Data.** Data univariat seperti Ringkasan data, **Tabel Kontingensi**, berbagai diagram seperti **Histogram, Box-Plot, Normal Plot, QQ-Plot**. Eksplorasi data multivariat seperti **Tabel Korelasi, Diagram Pencar** (2 peubah) yang dilengkapi dengan Histogram atau Box-plot, **Matriks Diagram Pencar** (untuk lebih dari 2 peubah) yang digabung dengan Box-plot, QQ-Plot atau Histogram. Kita dapat mengeksplorasi data secara intensif sebelum memilih analisis yang sesuai.
3. **Analisis Data Parametrik. Uji-t:** Uji nilai-tengah dan Beda Nilai-tengah, **Uji-F:** untuk uji anava (baik satu maupun multi arah), Uji beda Proporsi. **Regresi biasa, Model Linier Tergeneralisir (GLM)** untuk regresi dengan data tidak harus berdistribusi normal dan bisa mengakomodasi data kuantitatif maupun kualitatif. Semua dapat dilengkapi dengan ilustrasi grafik yang memudahkan interpretasi hasil. Juga dilengkapi dengan diagnostik regresi atau model linier serta prosedur pemilihan model. Melalui kemampuan plug-in, telah dapat juga disertakan melalui menu beberapa analisis khusus seperti GEE (*Generalized Estimating Equation*) dan Analisis Time Series.

4. **Analisis Nonparametrik** seperti uji Wilcoxon dan Kruskal-Wallis.
5. Analisis lanjut yang bersifat khusus dapat dilakukan melalui CLI (*Command Line Intereface*).

Untuk mengaktifkan paket yang telah diinstal, secara umum dilakukan perintah berikut melalui jendela Rconsole:

```
library(nama_paket)
```

Khusus untuk RCommander dan StatDemo dilakukan perintah berikut:

```
library(Rcmdr)  
library(StatDemo)
```

Apabila paket Rcmdr telah diaktifkan, kita akan melihat tampilan seperti pada Gambar 2.1. Dari gambar tersebut terlihat bahwa jendela RCommander terbagi menjadi 3 jendela utama yaitu jendela skrip, jendela keluaran dan jendela pesan. Selain itu juga dilengkapi dengan Menu Utama dan Panel Grafis berbahasa Indonesia. StatDemo dapat diaktifkan dengan dua cara yaitu cara langsung seperti diuraikan di atas atau dengan cara mengaktifkan plugin melalui RCommander yang akan dibahas kemudian. Tampilan menu langsung StatDemo ditunjukkan pada Gambar 2.2. Tampilan paket ini sepenuhnya menggunakan Bahasa Indonesia. Pembahasan lebih detil dapat dilihat pada sesi berikutnya.

DAFTAR PUSTAKA

- [1] Baird JH.2003. *How Statistics Can Lie*. Green Section Reccoed. [Agustus 2008]
- [2] Chamber JM & Hastie TJ.1993. *StatisticalModels in S*.Chapman & Hall
- [3] Crawley, M.J. 2005. *Statistics: An Introduction using R*. Springer
- [4] Cryer, J.D. & Chan, K.-S. 2008. *Time Series Analysis With Applications in R*. Springer
- [5] Diggle PJ, Liang K-Y & Zeger SL, 1994. *Analysis of Longitudinal Data*. Oxford Science Publications.
- [6] Everitt, B. & Hothorn, T. 2006. *A Handbook of Statistical Analyses Using R*. Springer
- [7] Faraway JJ. 2004. *Linear Model with R*. Springer
- [8] Faraway JJ. 2006. *Extending Linear Models with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Springer

- [9] Garson DG. StatNotes: *Topics in Multivariate Analysis*. <http://StatNote/www2.chass.ncsu.edu/garson/pa765/statnote.htm> [Maret 2008]
- [10] Gravetter FJ & Wallnau LB. 2004. *Statistics for the Behavioral Sciences*. Thomson. International Student Edition (6th Edition).
- [11] Guilford JP & Fruchter. 1978. *Fundamental Statistics in Psychology and Education*. International Student Edition(6th Edition). McGraw-Hill.
- [12] Hadi S. 1982. *Statistika*. Andi Offset, Yogyakarta.
- [13] Hair JF, Black WC, Babin BC, Anderson RE & Tatham RL. 2006. *Multivariate Analysis*. Pearson, Prentice Hall.
- [14] Harrel F.E. 2001. *Regression Modeling Strategies, with Applications to Linear Models, Survival Analysis and Logistic Regression*. Springer
- [15] Hastie R.J & Tibshirani T. 1990. *Generalized Additive Models*, 5th edn, Chapman & Hall.
- [16] Heiberger, R.M. & Neuwirt, E. R *Through Excel, A spreadsheet Interfac for Statistics Data Analysis, and Graphics*. Springer
- [17] Hogg RV, McKean JW, Craig AT. 2005. *Introduction to Mathematical Statistics*. 6-th ed. London: Pearson Education International
- [18] Kindt R and Coe R. 2005. *Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies*. Nairobi: World Agroforestry Centre (ICRAF).
- [19] Kleiber, C. & Zeileis. 2008. *Applied Econometrics with R*. Springer

- [20] Kuhnert P & Venables WN. 2005. *An Introduction to R: Software for Statistical Modelling & Computing*. http://cran.r-project.org/doc/contrib/Kuhnert+Venables-R_Course_Notes.zip
- [21] Lewis P. MODULE: Basic Statistics. <http://philosophy.hku.hk/think/stat/> 5 Agustus 2008]
- [22] Liang K-Y. & Zeger.SL. 1986. Longitudinal data analysis using generalized linear models. *Biometrika*.**73**:13-22.
- [23] Maindonald JH. 2001. *Using R for Data Analysis and Graphics An Introduction*. <http://www.r-project.org>.
- [24] B. D. McCullough. 2010. H.D. Chapter 1 Econometric Computing with R. In Vinod (Editor). *Advances in Social Science Research Using R*. Springer.
- [25] McCullagh P & Nelder JA. 1989. *Generalized Linear Models*. Chapman & Hall
- [26] Mendenhall W. 1979. *Introduction to Probability and Statistics 5th ed.* Massachussets:Duxbury.
- [27] Mendenhall W. 1993. *Beginning Statistics A to Z*. Duxbury
- [28] Murrell, P. 2006. *R Graphics*. Chapman & Hall/CRC
- [29] Nelder J.A. & Wedderburn, Generalized Linear Models. *J.R.Statist.Soc*.**57**: 359-407.
- [30] Neter J. Wasserman, Kutner. 1985 *Applied Statistical Model*. Homewood, Illinois :Irwin.
- [31] NIST/SEMATECH *e-Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/>, [Januari 2008]
- [32] Reimann C., Fitzmoser, P., Garret RG., Dutter, R. 2008. *Statistical Data Analysis Explained Applied Environmental Statistics with R*. John Willey & Son

- [33] Rstudio. 2013. *shiny: Web Application Framework for R*. [R package version 0.6.0.99]
- [34] Sarkar. D. 2008. *Lattice: Multivariate Data Visualization with R*. Springer
- [35] Sheather, S. 2008. *A Modern Approach to Regression with R*. Springer
- [36] Shumway, R.H. & Stoffer, D.S. 2006. *Time Series Analysis and Its Applications With R Examples*. Springer
- [37] Snow, G. 2013. *TeachingDemos: Demonstrations for teaching and learning* [R package version 2.9]
- [38] Spector, P. 2008. *Data Manipulation with R*. Springer
- [39] StatSoft. 2006. *Electronic Statistics Textbook*. <http://www.statsoftinc.com/textbook/stathome.html>
- [40] Sudjana. 1996. *Metode Statistika*. Tarsito, Bandung.
- [41] Tirta.2005. Potensi dan Prospek Pemanfaatan OSS-R dalam Analisis Data dan Pengajaran Statistika. *Pancaran Pendidikan*. XVIII (61): 195-208
- [42] Tirta, IM.2005. *Panduan Program Statistika R*. Penerbit Universitas Jember
- [43] Tirta IM, Lestari B & Dewi YS. 2006. Estimasi Efek Tetap dan Acak pada Model Multiplikatif dengan Likelihood Bersama. *Jurnal ILMU DASAR*. 7:59-66.
- [44] Tirta IM. 2007. *Analisis Data dengan Respons tidak Saling Bebas dengan paket hglm dan gee pada OSS-R*. Makalah disampaikan pada Seminar Nasional Statistika VIII di ITS Surabaya. 3 Nopember 2007.
- [45] Tirta IM. 2007. *Pengembangan Piranti Lunak Statistika Berbasis OSS-R (Development of Statistical Software)*.

disajikan dalam seminar Nasional Statistika ke - 8 (SNS VIII) di FMIPA ITS surabaya pada tanggal 3 Nopember 2007

- [46] Tirta IM.2007. *R.GUI: Mendesain Paket Analisis dan Media Pembelajaran Statistika*. Penerbit Universitas Jember
- [47] Tirta IM. 2008. Paket *RcmdrPlugin.StatDemo*. <http://r.unej.ac.id>
- [48] Tirta IM. 2008. Paket *StatDemo*. <http://r.unej.ac.id>
- [49] Tirta IM. 2009. *Analisis Regresi dengan R*. UNEJ Press
- [50] Varmuza, K. & Filzmoser, P. 2009. *Introduction to Multivariate Statistical Analysis in Chemometrics*. Springer
- [51] Venables WN & Ripley BD. 1994. *Modern Applied Statistics with S-plus*. Springer.
- [52] Vealini J. 2002. *Using R for Introductory Statistics*. <http://www.r.project.org>.
- [53] Wackerly, DD., Mendenhall W., Scheafer RL., 1996. *Mathematical Statistics with Application*. Massachussets:Duxbury.
- [54] Wannacott TH & Wannacott RJ. 1990. *Introductory Statisticsfor Business and Economics*. Williey International Edition.
- [55] Zeger SL.& Liang KY. 1986. Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics*. **42**: 121-130
- [56] Zoonekyn V. 2005. *Statistics with R*. http://zoonek2.free.fr/UNIX/48_R/all.html

LAMPIRAN

DAFTAR REFERENSI TERKAIT R

Daftar Buku Teks Yang Membahas atau Menggunakan R Sebagai Ilustrasi.

No	Penulis	Tahun	Judul	Penerbit
Aktuaria/Finansial/Social Sciences				
1.	Boland, P.J.	2007	Statistical and Probabilistic Methods in Actuarial Science	Chapman & Hall
2.	Kleiber, C. & Zeileis, A.	2008	Applied Econometrics with R	Springer
3.	Vinod, H.D.	2008	Hands-on Intermediate Econometrics Using R: Templates for	World Scientific

No	Penulis	Tahun	Judul	Penerbit
			Extending Dozens of Practical Examples	
4.	Vinod, H.D. (Ed)	2010	Advances in Social Science Research Using R. Lecture Notes in Statistics	Springer
5.	David Ruppert	2010	Statistics and Data Analysis for Financial Engineering. Use R!	Springer
6.	Hard Pfaff	2012	Financial Risk Modelling and Portfolio Optimisation with R	Wiley
7.	Diethelm Würtz et al.	2009	Portfolio Optimization with R/Rmetrics	Rmetrics Associatio n & Finance Online
<i>Bayesian/MCMC</i>				
8.	Lynch, S.M.	2007	Introduction to Applied Bayesian Statistics and Estimation for Social Scientists	Springer
9.	Marin, J.- M. & Robert, C.P.	2007	Bayesian Core: A Practical Approach to Computational Bayesian Statistics	Springer
10.	Albert, J.	2009	Bayesian	Springer

No	Penulis	Tahun	Judul	Penerbit
			Computation with R	
11.	Hoff, P.D.	2009	A First Course in Bayesian Statistical Methods	Springer
12.	Robert, C. & Casella, G.	2010	Introducing Monte Carlo Methods with R	Springer
Biosains				
13.	Deonier, R.C., Tavaré, S. & Waterman, M.S.	2005	Computational Genome Analysis: An Introduction	Springer
14.	Robert Gentleman et al.	2005	Bioinformatics and Computational Biology Solutions Using R and Bioconductor	Springer
15.	Paradis, E.	2006	Analysis of Phylogenetics and Evolution with R	Springer
16.	Dudoit, S. & van der Laan, M.J.	2007	Multiple Testing Procedures and Applications to Genomics	Springer
17.	Siegmund, D. & Yakir, B.	2007	The Statistics of Gene Mapping	Springer
18.	Gentleman, R.	2008	Bioinformatics with R	Chapman & Hall/CRC
19.	Claude, J.	2008	Morphometrics	Springer

No	Penulis	Tahun	Judul	Penerbit
			with R	
20.	Peng, R.D. & Dominici, F.	2008	Statistical Methods for Environmental Epidemiology with R: A Case Study in Air Pollution and Health	Springer
21.	Reimann, C., et al.	2008	Statistical Data Analysis Explained: Applied Environmental Statistics with R	Wiley
22.	Gentleman, R.	2008	R Programming for Bioinformatics	Chapman & Hall/CRC
23.	Bolker, B.M.	2008	Ecological Models and Data in R	Springer
24.	Foulkes, A.S.	2009	Applied Statistical Genetics with R: For Population-Based Association Studies	Springer
25.	Stevens, M.H.H.	2009	A Primer of Ecology with R	Springer
26.	Steyerberg, E.W.	2009	Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating	Springer
Geosains				
27.	Diggle, P.J.	2006	Model-based	

No	Penulis	Tahun	Judul	Penerbit
	& Ribeiro, P.J.		Geostatistics	
28.	Bivand, R.S, et al.	2008	Applied Spatial Data Analysis with R	
Grafik/Visualisasi				
29.	Wickham, H.	2009	ggplot: Elegant Graphics for Data Analysis	Springer
30.	Hrishi Mittal	2011	R Graphs Cookbook	Packet Publishing
31.	D. Sarkar	2008	Lattice: Multivariate Data Visualization with R	Springer
32.	Dianne Cook and Deborah F. Swayne	2007	Interactive and Dynamic Graphics for Data Analysis.	Springer
Komputasi				
33.	Rizzo, M.L.	2008	Statistical Computing with R	
34.	Velten, K.	2009	Mathematical Modeling and Simulation: Introduction for Scientists and Engineers	
35.	Ramsay,JO, et al.	2009	Functional Data Analysis with R and Matlab	
36.	van Vasishth	2010	The Foundations	Springer

No	Penulis	Tahun	Judul	Penerbit
	and Michael Broe		of Statistics: A Simulation-based Approach	
Multivariat				
37.	Murtagh, F.	2005	Correspondence Analysis and Data Coding with JAVA and R	
38.	Everitt, B.S.	2005	An R and S-Plus Companion to Multivariate Analysis	
39.	Greenacre, M.	2007	Correspondence Analysis in Practice, Second Edition	
40.	Sarkar, D.	2008	Lattice: Multivariate Data Visualization with R	
41.	Francois Husson, Sébastien Lê, and Jérôme Pagès	2010	Exploratory Multivariate Analysis by Example Using R	Chapman & Hall/CRC
42.	Varmuza, K. & Filzmoser, P.	2009	Introduction to Multivariate Statistical Analysis in Chemometrics	
Program				
43.	Becker, R.A., et al.	1988	The New S Language	

No	Penulis	Tahun	Judul	Penerbit
44.	Chambers, J.M.	1998	Programming with Data	
45.	Venables, W.N. & Ripley, B.D.	2000	S Programming	
46.	Murrell, P.	2005	R Graphics	
47.	Chambers, J.M.	2008	Software for Data Analysis: Programming with R	
48.	Zuur, A.F., Ieno, E.N. & Meesters, E.	2009	A Beginner's Guide to R	
49.	Broman, K.W. & Sen, S.	2009	A Guide to QTL Mapping with R/qlt	
50.	Kabacoff, R.	2010	R in Action	
51.	Regresi Dasar			
52.	Chambers, J.M. & Hastie, T.J.	1992	Statistical Models in S	
53.	Fox, J.	2002	An R and S-Plus Companion to Applied Regression	
54.	Faraway, J.J.	2004	Linear Models with R	
Regresi Modern/ Tingkat Lanjut				
55.	Pinheiro, J.C. & Bates, D.M.	2000	Mixed-Effects Models in S and S-Plus	
56.	Therneau, J.M.	2000	Modeling Survival Data with S and S-Plus	

No	Penulis	Tahun	Judul	Penerbit
	T.M. & Grambsch, P.M.		Data: Extending the Cox Model	
57.	Harrell, F.E.	2001	Regression Modeling Strategies, with Applications to Linear Models, Survival Analysis and Logistic Regression	
58.	Venables, W.N. & Ripley, B.D.	2002	Modern Applied Statistics with S. Fourth Edition	
59.	Huet, S., Bouvier, A., Gruet, M.-A. & Jolivet, E.	2003	Statistical Tools for Nonlinear Regression	
60.	Faraway, J.J.	2006	Extending Linear Models with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models	
61.	Jureckova, J. & Picek, J.	2006	Robust Statistical Methods with R	
62.	Keele, L.	2008	Semiparametric Regression for the Social Sciences	
63.	Sheather, S.	2008	A Modern Approach to Regression with R	

No	Penulis	Tahun	Judul	Penerbit
64.	Ritz, C. & Streibig, J.C.	2009	Nonlinear Regression with R	
65.	Zuur, A., et al.	2009	Mixed Effects Models and Extensions in Ecology with R	
66.	Petris, G., Petrone, S. & Campagnoli, P.	2009	Dynamic Linear Models with R	
67.	Wright, D.B. & London, K.	2009	Modern Regression Techniques Using R: A Practical Guide	SAGE
68.	Wood, S.N.	2006	Generalized Additive Models: An Introduction with R	

Statistika Dasar

69.	Spector, P.	2008	Data Manipulation with R	
70.	Ihaka, R. & Gentleman, R.	1996	R: A Language for Data Analysis and Graphics	
71.	Dalgaard, P.	2002	Introductory Statistics with R	
72.	Heiberger, R.M. & Holland, B.	2004	Statistical Analysis and Data Display: An Intermediate Course with Examples in S-Plus, R, and SAS	

No	Penulis	Tahun	Judul	Penerbit
73.	Verzani, J.	2005	Using R for Introductory Statistics	
74.	Crawley, M.J.	2005	Statistics: An Introduction using R	
75.	Everitt, B. & Hothorn, T.	2006	A Handbook of Statistical Analyses Using R	
76.	Maindonald, J. & Braun, J.	2007	Data Analysis and Graphics Using R	
77.	Cook, D. & Swayne, D.F.	2007	Interactive and Dynamic Graphics for Data Analysis	
78.	Dalgaard, P.	2008	Introductory Statistics with R	
79.	Heiberger, R.M. & Neuwirth, E.	2009	R Through Excel	Springer

Time Series

80.	Pfaff, B.	2006	Analysis of Integrated and Cointegrated Time Series with R	
81.	Le, N.D. & Zidek, J.V.	2006	Statistical Analysis of Environmental Space-Time Processes	
82.	Shumway, R.H. & Stoffer, D.S.	2006	Time Series Analysis and Its Applications With R Examples	
83.	Pfaff, B.	2008	Analysis of	

No	Penulis	Tahun	Judul	Penerbit
			Integrated and Cointegrated Time Series with R, Second Edition	
84.	Cryer, J.D. & Chan, K.-S.	2008	Time Series Analysis With Applications in R	
85.	Cowpertwait, P.S.P. & Metcalfe, A.	2009	Introductory Time Series with R	
86.	Braun, W.J. & Murdoch, D.J.	2007	A First Course in Statistical Programming with R	
Wavelet				
87.	Nason, G.P.	2008	Wavelet Methods in Statistics with R	

MENGUNDUH DAN MENGINSTAL R

R adalah program *open source* yang dapat diunduh dari internet. Berikut diuraikan cara mengunduh dan menginstal R terutama pada sistem operasi Windows.

MENGUNDUH (MEN-DOWNLOAD) R

Ada beberapa cara yang dapat ditempuh untuk memperoleh program R beserta paket-paket pendukung (*library*)nya diantaranya adalah sebagai berikut.

1. Mengunjungi situs <http://www.cran.r-project.org>, lalu memilih situs bayangan (mirror) yang paling dekat dengan lokasi kita. Salah satu diantaranya adalah situs yang ada di Indonesia yaitu

<http://cran.repo.bppt.go.id/>. Pengguna dapat mengunduh (*download*) program secara cuma-cuma.

2. Menghubungi Laboratorium Statistika Jurusan Matematika FMIPA Universitas Jember. Kepada pengguna dikenakan biaya mengunduh (*download*) dan pengemasan yang tidak mahal;
3. Menghubungi UPT Penerbitan Universitas Jember dan membeli salah satu panduan atau manual tentang R yang berisi paket Program R, diantaranya adalah Tirta [42] [46];

Komponen program yang diperlukan untuk dapat menjalankan R dan R-Commander dengan baik adalah:

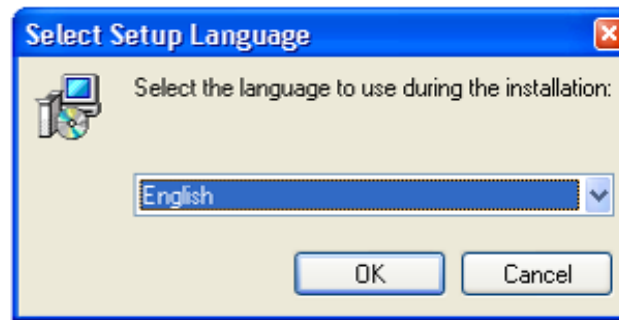
1. program R untuk windows yang biasanya diberi nama **R-Versi-win32.exe**;
2. semua paket pendukung/pustaka/library versi windows yang dibutuhkan seperti: **aaMI_versi.zip, ..., zoo_versi.zip**;
3. bagi yang ingin menggunakan R melalui skrip pemrograman diperlukan juga Tinn-R yaitu **Tinn-R_versi_setup.exe**.
4. Dalam notasi di atas, “versi” menunjukkan angka yang terkait dengan versi program bersangkutan, misalnya 2.5.0; 7.0.1 dan sebagainya.

MENGINSTAL R

Setelah memperoleh file “R-versi-win32.exe”, maka instalasi dilakukan sebagai berikut ini⁴.

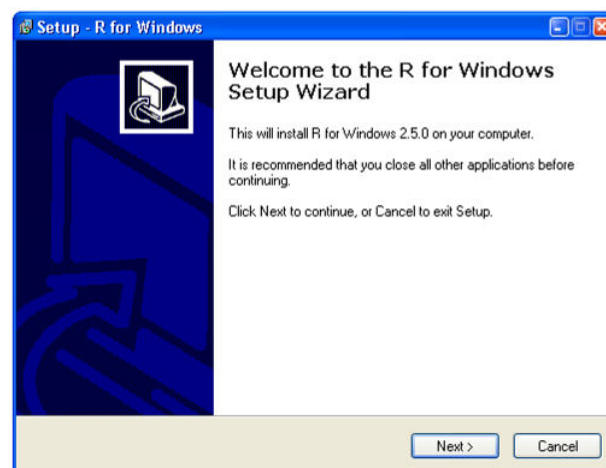
⁴ Versi yang telah beredar tahun 2012 adalah versi 3, namun prosedur ttatacara menginstal masih sama dengan versi sebelumnya, termasuk versi 2.5 yang diilustrasikan pada buku ini.

1. Klik “R-versi-win32.exe”, lalu akan muncul jendela dialog pemilihan bahasa. Saat ini belum ada dialog berbahasa Indonesia, untuk itu pilih English dan klik OK

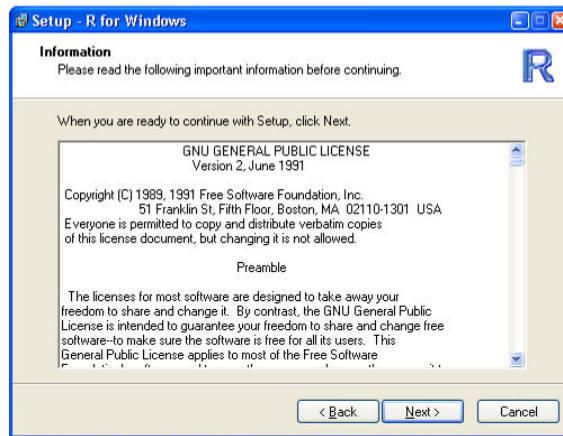


Ilustrasi 1 Tampilan Pemilihan Bahasa Instalasi

2. Selanjutnya akan muncul dialog pembukaan, seperti berikut ini, kita tinggal klik Next, lalu akan muncul dialog tentang informasi *Copy Right*. Sempatkan membaca ketentuan yang berlaku, selanjutnya klik next.

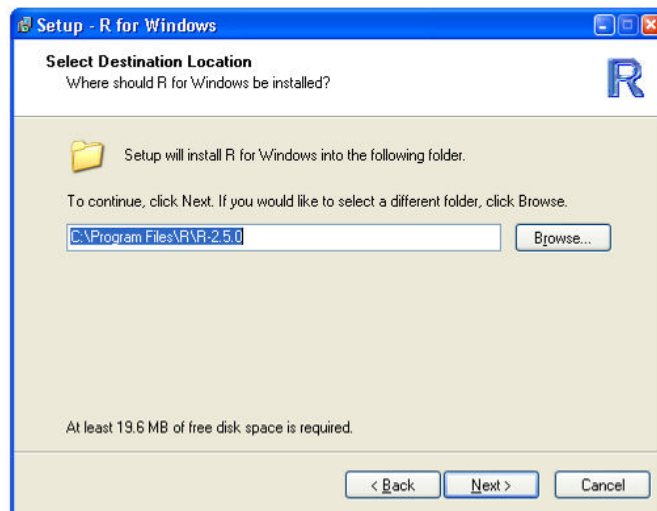


Ilustrasi 2 Tampilan Mulai Instalasi



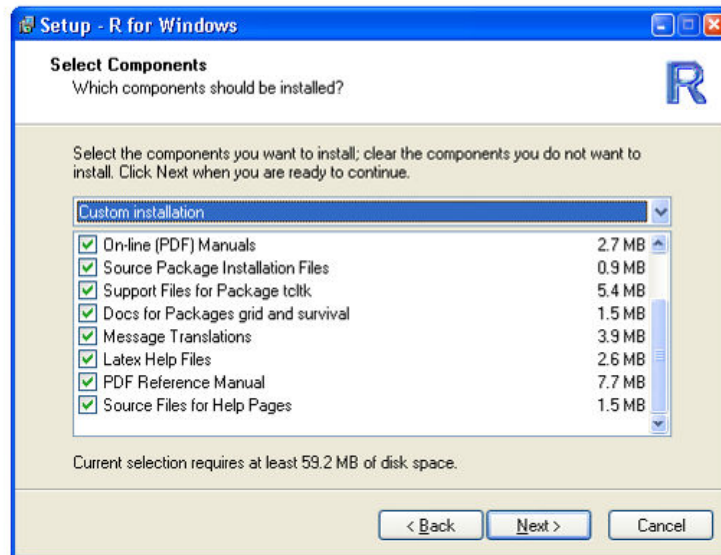
Ilustrasi 3 Tampilan Konfirmasi Lisensi GPL (*General Public License*)

3. Untuk pemilihan direktori sebaiknya ikuti sesuai yang ada pada dialog, misalnya C:\Program Files\R\... Kecuali, untuk yang menggunakan Windows Vista, sebaiknya R diinstal tidak di hardisk C, tetapi di D atau E.



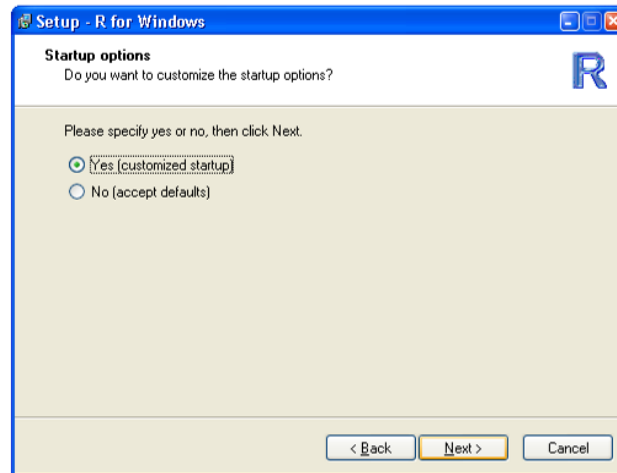
Ilustrasi 4 Konfirmasi Direktori

4. Untuk dokumentasi, jika ruang hardisk anda cukup longgar sebaiknya pilih semua dokumentasi terkait. Jika tidak, pilih yang penting diantaranya adalah: HTML files help, Online (pdf) manual, Support files for tcltk, Message Translation, PDF Reference manuals.

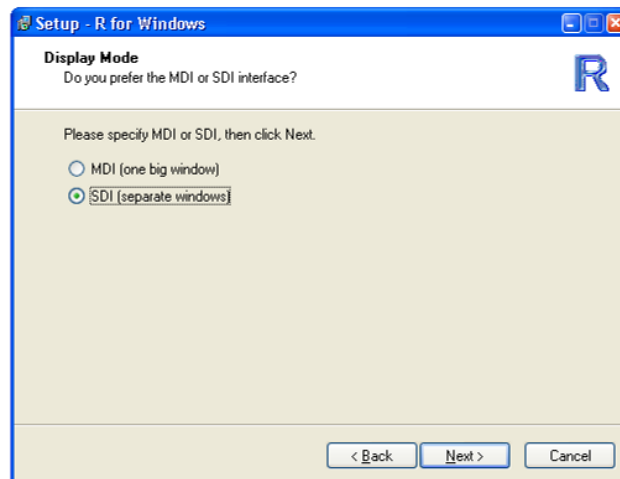


Ilustrasi 5 Tampilan Konfirmasi Komponen R

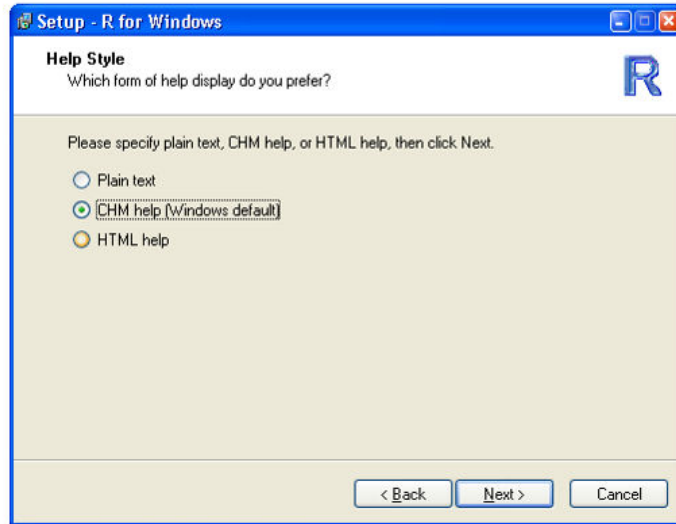
5. Kustomisasi atau pengaturan Opsi Startup, **pilih yes**. Dengan pilihan tersebut selanjutnya akan muncul beberapa dialog berikut. Untuk mengatur agar tampilan grafik dan jendela R terpisah, **pilih SDI**. Untuk dokumentasi bantuan/help bisa pilih **CHM atau HTML**. Untuk Startup ikuti pilihan yang telah disiapkan (*default*) R.



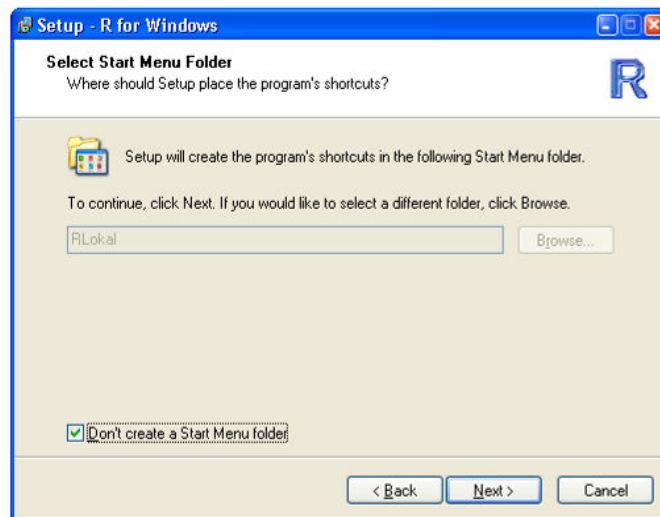
Ilustrasi 6 Tampilan Pemilihan Pengaturan *Startup* Aplikasi



Ilustrasi 7 Tampilan Pemilihan Jenis Tampilan

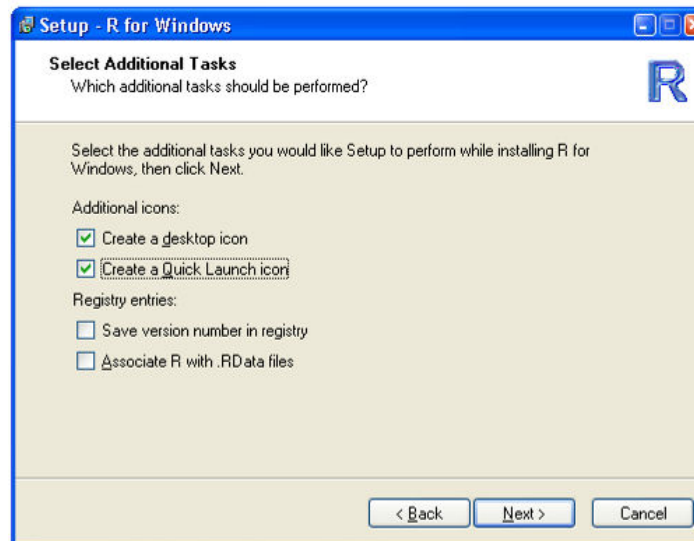


Ilustrasi 8 Tampilan Pemilihan Bahasa Instalasi



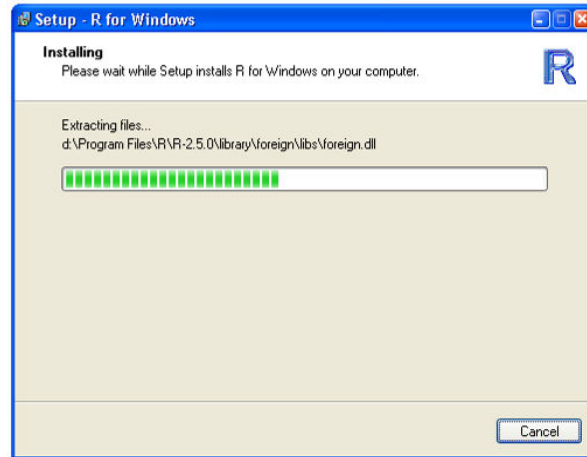
Ilustrasi 9 Tampilan Pemilihan Folder Startmenu

6. Untuk tugas tambahan (*Additional Tasks*), minimal bisa dipilih (di klik check), *Create desktop icon* dan *Create a Quick launch*



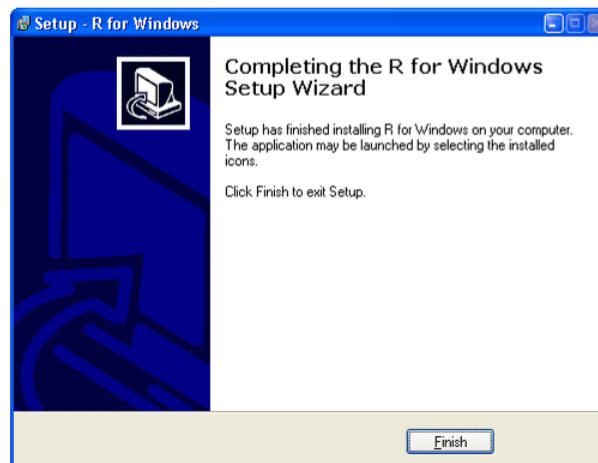
Ilustrasi 10 Tampilan Pemilihan *Task* Tambahan

7. Selanjutnya proses instalasi akan dimulai



Ilustrasi 11 Tampilan Proses Instalasi

8. Jika selesai akan muncul dialog terakhir berikut. Klik finish untuk menyelesaikan instalasi R



Ilustrasi 12 Tampilan Akhir Proses Instalasi

MENU BAHASA INDONESIA DAN PAKET TAMBAHAN

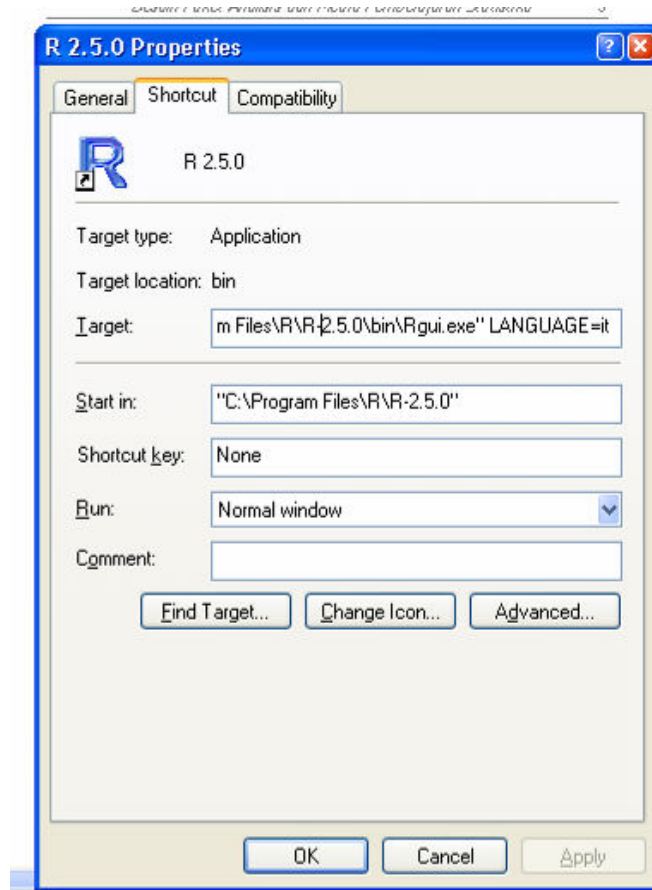
Untuk memanggil atau mengaktifkan R kita bisa klik icon R yang ada di desktop. Untuk memunculkan menu atau dialog dalam Bahasa Indonesia lakukan perubahan pada properties icon ini. Pada baris target ditulis

```
"C:\Program Files\R\R-2.5.0\bin\Rgui.exe"  
LANGUAGE=id
```

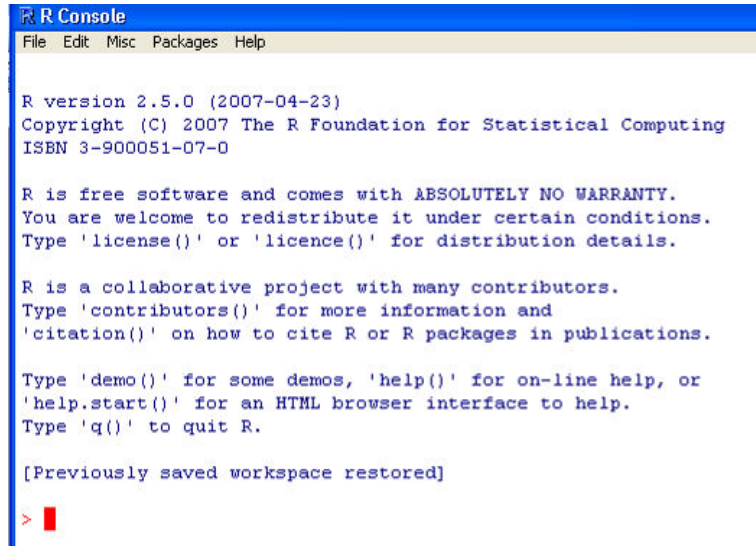
Perlu diketahui bahwa, sampai saat ini menu dan dialog dalam Bahasa Indonesia hanya tersedia pada paket RCommander, belum sampai pada menu RGUI (Rconsole). Untuk memulai R, kita dapat mengklik *short cut* yang ada pada desktop, atau melalui Start Menu. Setelah R dipanggil, kita akan memperoleh tampilan yang disebut Console R

Jika menggunakan Windowa 7 ke atas, maka pilihan bahasa pada R (jika tersedia) akan sama dengan pilihan bahasa pada Windows.

Bagi pembaca yang berminat menerjemahkan menudan dialog R ke dalam bahasa Indonesia dapat mengunjungi situs <http://translation.r-project.kr/>



Ilustrasi 13 Mengatur Pilihan Bahasa Menu dan Dialog R



R Console

File Edit Misc Packages Help

```
R version 2.5.0 (2007-04-23)
Copyright (C) 2007 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

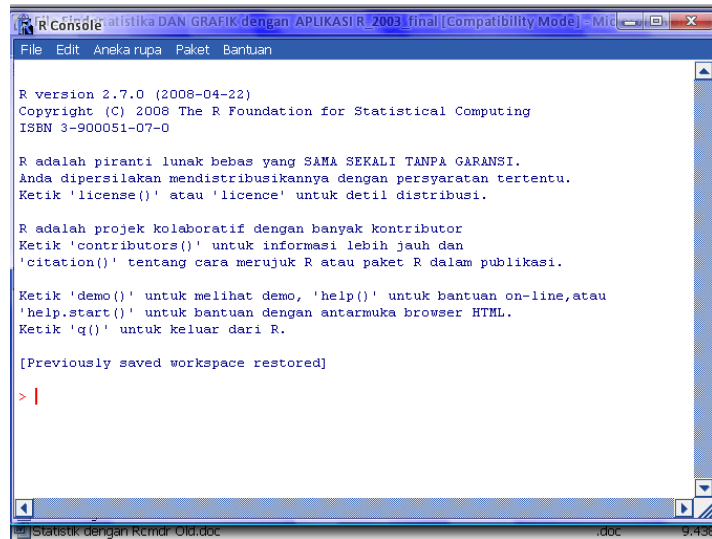
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> |
```

Ilustrasi 14 Tampilan Rconsole dalam Bahasa Inggris



R Console

File Edit Aneka rupa Paket Bantuan

```
R version 2.7.0 (2008-04-22)
Copyright (C) 2008 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R adalah piranti lunak bebas yang SAMA SEKALI TANPA GARANSI.
Anda dipersilakan mendistribusikannya dengan persyaratan tertentu.
Ketik 'license()' atau 'licence' untuk detail distribusi.

R adalah proyek kolaboratif dengan banyak kontributor
Ketik 'contributors()' untuk informasi lebih jauh dan
'citation()' tentang cara merujuk R atau paket R dalam publikasi.

Ketik 'demo()' untuk melihat demo, 'help()' untuk bantuan on-line, atau
'help.start()' untuk bantuan dengan antarmuka browser HTML.
Ketik 'q()' untuk keluar dari R.

[Previously saved workspace restored]

> |
```

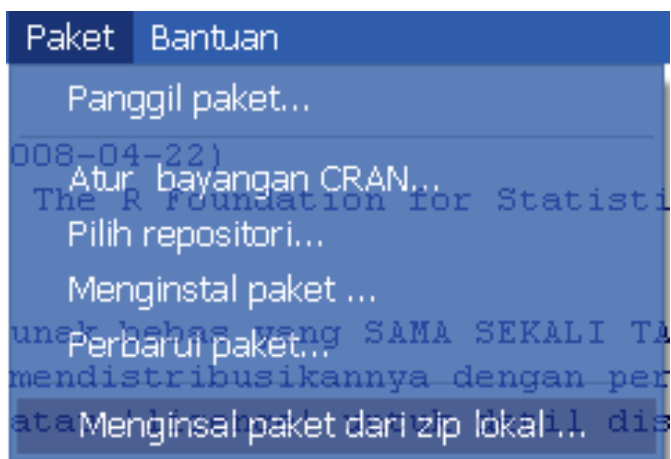
Ilustrasi 15 Tampilan Rconsole Dalam Bahasa Indonesia

Berikut adalah langkah-langkah untuk menginstal paket tambahan.

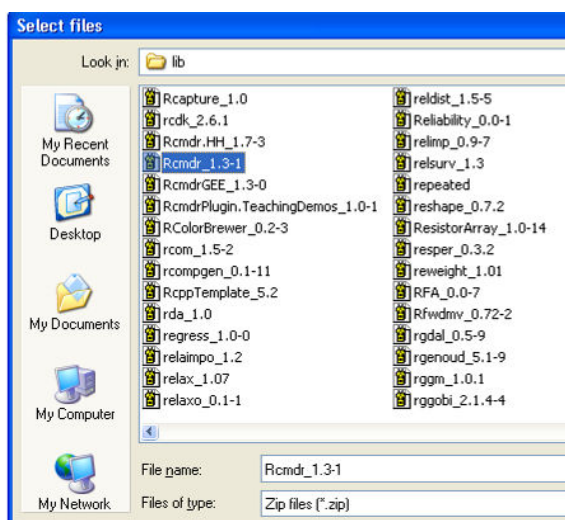
1. Aktifkan menu Packages seperti terlihat pada Ilustrasi 17 Memilih Paket Tambahan

Paket-> Menginstal paket dari Zip lokal.

2. Selanjutnya pilih direktori tempat anda menaruh paket-paket tambahan (*library*) dari R (lihat Ilustrasi 17). Ada beberapa paket penting yang harus diinstal untuk dapat mengikuti penjelasan dalam buku ini yaitu:
 - a. **Rcmdr** beserta paket-paket terkait (*car*, *dsb*), merupakan paket RGUI untuk analisis data yang dikembangkan oleh Fox dan dengan kontribusi menu bahasa Indonesia oleh Tirta.
 - b. **StatDemo_versi.zip** merupakan paket khusus bernahasa **Indonesia** untuk ilustrasi pembelajaran statistika yang dikembangkan oleh Tirta [48]
 - c. **RcmdrPlugin.StatDemo** sama dengan *statDemo* di atas, tetapi menunya bergabung dengan menu RCommander (Tirta [47])
 - d. **RcmdrPlugin.hglm** berisi paket *hglm* dan Menu GEE yang bergabung dengan menu RCommander. Paket *hglm* ditulis oleh Tirta *et al.* [43], sedangkan *gee* merupakan paket yang telah ada pada R, tetapi integrasi menu dalam bahasa Indonesia di RCommander dibuat oleh Tirta[45].



Ilustrasi 16 Menu Menginstal Paket Tambahan



Ilustrasi 17 Memilih Paket Tambahan


```

|--Uji t sp
berpasangan
|--Uji anava satu faktor
|--Uji anava multi faktor

|--Analisis -----|--Reliabilitas skala
dimensional |--Analisis Kom Utama
|--Analisis faktor
|--Analisis kluster

Grafik-----|--Grafik indeks
|--Histogram
|--Boxplot
|--QQplot
|--Diagram kuantil-kuantil
|--Diagram pencar
|--Matriks diagram Pencar
|--Grafik garis
|--Diagram rata-rata
|--Grafik batang
|--Grafik lingkaran
|--Grafik 3D

Distribusi-|--Distribusi Kontinu--|--Distribusi Normal
|--Distribusi t
|--Distribusi Chi-
kuad |--Distribusi Seragam
|-- ...
|--Distribusi Gumbel

-|--Distribusi Diskrit--|--Distribusi
Binomial |--Distribusi Poisson
|-- ...
|--Distribusi Hiperg

Alat -----|--Aktifkan paket
|--Aktifkan Plug-in
|--Pilihan

Bantuan ---|--Bantuan Commander
|--Pengantar RCommander
|--Bantuan data (jika ada)
|--Tentang Rcmdr

```

STRUKTUR MENU REXCEL

Menu RExcel

Run Code

```
Get R Value      ---- |--Array
                  |-- Dataframe
                  |-- Active Dataframe
```

```
Put Rvar         |-- Array
                  |-- Dataframe
```

Get R Output

Set R Working Directory

Load R file

Copy Code

Debug R

Error log

Option

Menu Kontak

Run Code in Rcmdr

Run Code

Get R Value

Put R Var

Get R Dataframe

Get Active Dataframe

Put R Dataframe

```
Rcmdr get      ---- |-- Last commands
                  |-- All commands
                  |-- Last outputs
                  |-- Complete outputs
                  |-- Last messages
                  |-- All messages
```

Get R Output

Inser current R plot

DAFTAR PERSAMAAN

Pers. 3.1	$P(A B) = \frac{P(A \cap B)}{P(B)}$	71
Pers. 3.2	$P(A \cup B) = P(A) + P(B)$, untuk $A \cap B = \emptyset$	71
Pers. 3.3	$P(A \cap B) = 0$, untuk $A \cap B = \emptyset$	72
Pers. 3.4	$A B \Leftrightarrow P(A) = P(A B)$	72
Pers. 3.5	$A B \Leftrightarrow P(A \cap B) = P(A) \times P(B)$	73
Pers.		3.6
	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$, untuk $A \cap B \neq \emptyset$	74
Pers. 3.7	$p(x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x}$, $x = 0, 1, 2, \dots, n$	80
Pers. 3.8	$p(x) = \pi (1 - \pi)^{x-1}$, $x = 1, 2, \dots$,	80
Pers. 3.9	$p(x) = \binom{x-1}{r-1} \pi^r (1 - \pi)^{x-r}$, $x = r, r+1, r+2, \dots$,	81
Pers. 3.10	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	82
Pers. 3.11	$s_x^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2}{n-1}$	85

$$\text{Pers. 3.12} \quad r = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \dots 86$$

Pers. 4.1

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right]; \quad -\infty < x < \infty \dots 100$$

$$\text{Pers. 4.2} \quad f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}; \quad -\infty < z < \infty \dots 101$$

$$\text{Pers. 4.3} \quad Z = \frac{X - \mu}{\sigma} \text{ sebaliknya } X = \mu + \sigma Z \dots 104$$

$$\text{Pers. 4.4} \quad Y = \sum_{i=1}^n a_i X_i \text{ berdistribusi } N \left(\mu \sum_{i=1}^n a_i, \sigma^2 \sum_{i=1}^n a_i^2 \right) \dots 111$$

$$\text{Pers. 4.5} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N(\mu, \sigma^2/n) \quad \text{atau}$$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \dots 111$$

$$\text{Pers. 6.1} \quad \bar{x} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \dots 170$$

$$\text{Pers. 6.2} \quad \sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \dots 170$$

Pers. 6.3 $\bar{x} - z_{\alpha/2} \times \sigma/\sqrt{n} < \mu < \bar{x} + z_{\alpha/2} \times \sigma/\sqrt{n}$ 171

Pers. 6.4 $\bar{x} - t_{n-1,\alpha/2} \times s/\sqrt{n} < \mu < \bar{x} + t_{n-1,\alpha/2} \times s/\sqrt{n}$ 172

Pers. 6.5 $z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$ dengan $\hat{p} = \frac{X}{n}$ 205

Pers. 7.1 $r = \frac{\sum_{i=1}^n x_i y_i - 1/n \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\sum_{i=1}^n x_i^2 - 1/n \left(\sum_{i=1}^n x_i\right)^2} \sqrt{\sum_{i=1}^n y_i^2 - 1/n \left(\sum_{i=1}^n y_i\right)^2}}$.. 211

Pers. 7.2 $s_{ee} = \sum_{i=1}^n \left[y - (\hat{\beta}_0 + \hat{\beta}_1 x_i) \right]^2$ 220

Pers. 7.3 $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$, 220

Pers. 7.4 $\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - 1/n \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - 1/n \left(\sum_{i=1}^n x_i\right)^2}$, 220

Pers. 7.5 $\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}}$, 220

Pers. 7.6 $s_{xy} = \sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}$, $s_{xx} = \sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}$.. 220

Pers. 7.7 $\hat{\beta}_1 = r_{xy} \frac{s_x}{s_y}$, dengan $s_x = \sqrt{s_{xx}}$; $s_y = \sqrt{s_{yy}}$ 220

Pers. 7.8 $\sigma^2 = s^2 = \frac{S_{ee}}{n-2}$ 221

Pers. 7.9 $s_{\beta_1}^2 = \frac{s^2}{S_{xx}}$, dan $s_{\beta_0}^2 = s^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)$ 221

Pers. 7.10 $s_{\hat{y}}^2 = s_{ee} \left[\frac{1}{n} + \frac{(x_p - \bar{x})^2}{S_{xx}} \right]$ dan

$\bar{y} = \hat{y} \pm t_{\alpha/2} s \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{S_{xx}}}$ 230

Pers. 7.11 $s_{\hat{y}}^2 = s_{ee} \left[1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{S_{xx}} \right]$ dan

$y = \hat{y} \pm t_{\alpha/2} s \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{S_{xx}}}$ 230

DAFTAR CONTOH

Contoh 1.1.....	13
Contoh 1.2.....	14
Contoh 1.3.....	15
Contoh 3.1.....	72
Contoh 3.2.....	75
Contoh 3.3.....	76
Contoh 3.4.....	81
Contoh 3.5.....	81
Contoh 3.6.....	82
Contoh 3.7.....	84
Contoh 3.8.....	86
Contoh 4.1.....	104
Contoh 4.2.....	108
Contoh 4.3.....	109
Contoh 6.1.....	172
Contoh 6.2.....	172
Contoh 7.1.....	213
Contoh 7.2.....	224
Contoh 7.3.....	225
Contoh 7.4.....	228
Contoh 7.5.....	242

INDEKS

Alternatif	204	Diagnostik .	236, 237, 239, 240
Anava	121, 168, 194, 201	Distribusi	30, 35, 41, 43, 44, 45, 46, 77, 78, 79, 84, 88, 89, 90, 91, 92, 93, 99, 100, 102, 103, 109, 125, 142, 249, 250, 251, 255, 288
Animasi	69	EDA.....	6, 145, 250
Berpasangan	186, 199	Eksponensial....	30, 78, 88, 118
Binomia	1	F	26, 79, 113, 120, 121, 150, 197, 201, 220, 223, 225, 226, 228, 242, 243, 287
.....	6, 30, 79, 118, 250, 288	Frekuensi ...	66, 67, 69, 70, 143
Box-plot.....	26	GUI.....	25, 60, 248, 261
chi-kwadrat.....	79, 143	HGLM.....	123
CIS.....	9		
Derajat	86, 210		
Deskriptif.....	5, 129, 254		
Determinasi	243		
Deviasi.....	.. 85, 93, 171, 198, 250, 251		

- hipotesis.....
 115, 116, 118, 119, 168,
 173, 174, 178, 179, 180,
 182, 185, 187, 188, 190,
 195, 222
- Hipotesis.....
 16, 115, 119, 173, 174, 181,
 187, 195, 204, 222, 251
- Histogram..... 26, 112, 146, 147,
 162, 251, 288
- Inferensial..... 6
- Interval
 18, 48, 174, 175, 185, 190,
 216, 251
- Keyakinan..... 47, 48, 49
- Korelasi
 26, 85, 122, 143, 144, 210,
 211, 213, 214, 215, 216,
 245
- Kuantil.....
 82, 83, 84, 151, 152, 161
- kumulatif 88, 89, 106
- Library..... 248
- Mean.....
 .. 26, 42, 48, 82, 85, 93, 197
- Median.....
 82, 83, 84, 85, 90, 91, 92,
 150, 224, 227, 242
- Modus..... 82, 84, 92
- Nominal..... 18, 254
- Normal..... 6, 26, 30, 41, 43, 45,
 78, 83, 84, 88, 100, 102,
 103, 106, 107, 109, 110,
 111, 112, 118, 122, 126,
 127, 150, 151, 173, 181,
 187, 195, 221, 222, 237,
 288
- Ordinal..... 18
- Parameter..... 40, 47, 168, 253
- Pemusatan..... 82, 93, 94
- pencar
 30, 146, 157, 211, 212, 226,
 228, 288
- Pencar
 26, 157, 158, 159, 160, 161,
 162, 223, 227, 228, 244,
 288
- Pencilan 243, 244
- Penyebaran 85
- plugin..... 10
- Plug-in 37, 38, 249, 288
- Populasi 15, 253
- Proporsi
 26, 122, 168, 202, 203, 204,
 205, 287
- p-value.....
 142, 174, 175, 178, 179,
 182, 185, 188, 189, 190,
 191, 195, 215, 216, 222,
 225, 226, 228, 242, 243
- Ragam..... 102, 192

Range.....	82, 85	simetris	
Rasio.....	19, 68	44, 45, 78, 83, 84, 88, 91,	
Rcmdr	25, 27, 36, 285, 288	100, 101, 125, 145, 150,	
RCommander		225	
10, 23, 25, 26, 27, 29, 31,		StatDemo.....	
35, 36, 37, 38, 50, 60, 105,		10, 23, 25, 27, 28, 36, 38,	
108, 119, 126, 129, 130,		40, 65, 69, 93, 109, 126,	
131, 133, 137, 140, 144,		175, 180, 182, 186, 191,	
145, 150, 168, 173, 180,		261, 285	
187, 193, 213, 224, 231,		Statistik.....	
237, 282, 287, 288	 4, 16, 20, 140, 141, 254	
Regresi.....		Statistika.....	
26, 47, 49, 122, 123, 157,		v, 4, 5, 6, 7, 8, 9, 10, 20, 28,	
216, 231, 239, 240, 243,		30, 33, 34, 36, 60, 66, 118,	
244, 253, 287		119, 129, 254, 258, 260,	
Rerata		261, 274, 287	
33, 107, 119, 155, 156, 157,		T 43, 44, 120, 121, 182, 183	
158, 168, 172, 180, 186,		Tinn-R	25, 274
198		Unimodal.....	255
Sampel.....		Variansi	
16, 66, 73, 82, 111, 172,		85, 155, 192, 193, 200, 255,	
176, 177, 183, 203, 204,		287	
205, 287		Wilcoxon.....	27, 287
SciViews.....	25	Z	
signifikansi		46, 100, 109, 112, 120, 121,	
.. 44, 45, 144, 172, 226, 249		259	

Filename: MetStatGRapR2014Maret.docx
Directory: C:\Users\netbook\Documents\DOC2013
Template: C:\Users\netbook\AppData\Roaming\Microsoft\Templates\Normal.d
otm
Title: ANALISIS DATA DENGAN
Subject:
Author: tirta
Keywords:
Comments:
Creation Date: 05/03/2014 11:46:00
Change Number: 41
Last Saved On: 10/09/2014 15:09:00
Last Saved By: netbook
Total Editing Time: 276 Minutes
Last Printed On: 13/10/2014 9:10:00
As of Last Complete Printing
Number of Pages: 327
Number of Words: 50.047 (approx.)
Number of Characters: 285.272 (approx.)