

ANALISIS DATA EKSPLORATIF
Menggunakan *Software Open Source R*

ugmpress.ugm.ac.id

ugmpress.ugm.ac.id

ANALISIS DATA EKSPLORATIF
Menggunakan *Software Open Source R*

Sri Haryatmi Kartiko

GADJAH MADA UNIVERSITY PRESS

**ANALISIS DATA EKSPLORATIF MENGGUNAKAN SOFTWARE
OPEN SOURCE R**

Penulis:

Sri Haryatmi Kartiko

Korektor:

Dim'z

Desain sampul:

Pram's

Tata letak isi:

Ari

Penerbit:

Gajah Mada University Press
Anggota IKAPI

Ukuran : 15,5 × 23 cm; xiv + 211 hlm

ISBN : 978-979-420-814-4

Font : Times New Roman
1901030-B5E-100(3).

Redaksi:

Jl. Grafika No. 1, Bulaksumur
Yogyakarta, 55281

Telp./Fax.: (0274) 561037

www.ugmpress.ugm.ac.id | gmupress@ugm.ac.id

Cetakan Pertama: November 2013

Cetakan Kedua: Juli 2016

2780.021.02.19

Hak Penerbitan © 2013 Gajah Mada University Press

Dilarang mengutip dan memperbanyak tanpa izin tertulis dari penerbit, sebagian atau seluruhnya dalam bentuk apa pun, baik cetak, photoprint, microfilm, dan sebagainya.

KATA PENGANTAR

Alhamdulillah, segala puji dan syukur selalu dipanjatkan kehadirat Allah SWT, yang senantiasa memberikan kemudahan dalam menyelesaikan segala urusan, karena hanya dengan rahmat-Nya, buku yang berjudul *Analisis Data Eksploratif Menggunakan Software Open Source R* dapat direalisasikan dengan baik, walaupun penulisan buku ini sempat tertunda beberapa waktu. Buku ini disusun untuk mahasiswa FMIPA Matematika dan Statistika di hampir seluruh universitas di Indonesia, sebagai pegangan untuk mata kuliah wajib *Analisis Data Eksploratif*. Selain itu, buku ini juga dapat digunakan untuk para peneliti yang akan menganalisa data yang dimiliki. Buku ini memuat secara garis besar teknik penyajian data yang merupakan eksplorasi data melalui bentuk visual. Selain itu, buku ini juga memaparkan ihwal teknik lain dalam menganalisa data, seperti pengujian hipotesis beberapa parameter populasi dan analisis regresi yang banyak sekali digunakan untuk menganalisa data pasangan variabel independen dan dependen. Setelah menyelesaikan proses membaca, memahami, dan mengerjakan latihan dalam buku ini, para pembaca yang budiman diharapkan mampu menyajikan data secara visual, serta mampu melakukan analisis data secara eksploratif, yang meliputi perbandingan populasi secara visual maupun konfirmasi, sampai dengan memodelkan hubungan statistik antara dua variabel dalam bentuk regresi.

Selanjutnya, tidak lupa kami haturkan terima kasih yang sedalam-dalamnya kepada Jurusan Matematika FMIPA dan Majelis Guru Besar Universitas Gadjah Mada Yogyakarta atas kerjasamanya, sehingga proses penerbitan buku *Analisis Data Eksploratif* ini dapat terlaksana dengan baik. Ucapan terima kasih juga penulis sampaikan untuk keluarga, Kamas Kartiko, anak dan cucu, Itsari Wigati, Joko Sendang, Iqbal

Wicaksana, dan Arya Suta atas semangat dan dorongannya. Tak lupa terima kasih untuk mas M Yunus Anis yang telah membantu sebagai editor bahasa dan mas Ari Dwi Hartanto sebagai editor (\LaTeX), sehingga isi dari buku ini dapat tersampaikan ke para pembaca dengan baik.

Akhir kata, kami mengucapkan selamat membaca dan berpandang mesra dengan rumus dan teknik yang kami sajikan. Semoga buku sederhana ini dapat bermanfaat untuk para mahasiswa, khususnya, dan para penganalisa data pada umumnya. Kami menyadari bahwa buku ini masih jauh dari kata sempurna, untuk itu kami mengharapkan kritik dan saran yang membangun untuk penyempurnaannya.

Yogyakarta, 23 Oktober 2013

Penulis

Prof. Dr. Sri Haryatmi Kartiko, M.Sc.

DAFTAR ISI

KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xiii
1 PENDAHULUAN	
DIAGRAM BATANG DAN DAUN	1
1.1. PENDAHULUAN	1
1.2. MENYEDERHANAKAN ANGKA DAN DAFTAR <i>TALLY</i>	3
1.3. DIAGRAM BATANG DAN DAUN	9
1.4. SINTAKS PROGRAM R	21
1.5. LATIHAN	25
2 RINGKASAN NUMERIK: UKURAN PUSAT	27
2.1. RATA-RATA	28
2.2. MEDIAN	33
2.3. MODUS	36
2.4. TRIRATA	37
2.5. RATA-RATA TENGAH	39
2.6. SINTAKS PROGRAM R	40
2.7. LATIHAN	41
3 RINGKASAN NUMERIK: UKURAN SEBARAN	42
3.1. <i>RANGE</i>	43
3.2. SEBARAN TENGAH	45
3.3. MEAN DEVIASI	46
3.4. VARIANSI DAN STANDAR DEVIASI	49
3.5. SINTAKS PROGRAM R	53
3.6. LATIHAN	53

4	PENGGUNAN RINGKASAN NUMERIK	55
4.1.	DIAGRAM KOTAK DAN TITIK	56
4.2.	STANDARDISASI	62
4.3.	SINTAKS PROGRAM R	70
4.4.	LATIHAN	70
5	TRANSFORMASI ANGKATAN	73
5.1.	MEMILIH TRANSFORMASI YANG UNGGUL	73
5.2.	TRANSFORMASI UNTUK SATU ANGKATAN	79
5.3.	TRANSFORMASI UNTUK BEBERAPA ANGKATAN	82
5.4.	SINTAKS PROGRAM R	87
5.5.	LATIHAN	88
6	SAMPEL RANDOM, DISTRIBUSI NORMAL, t, DAN F	89
6.1.	POPULASI DAN SAMPEL	90
6.2.	DISTRIBUSI NORMAL, t, DAN F	99
6.2.1.	Distribusi Normal	99
6.2.2.	Distribusi t	110
6.2.3.	Distribusi F	111
6.3.	SINTAKS PROGRAM R	112
6.4.	LATIHAN	113
7	STATISTIKA KONFIRMASI	116
7.1.	KONSEP DASAR UJI HIPOTESIS	116
7.2.	UJI HIPOTESIS RATA-RATA POPULASI	119
7.3.	UJI HIPOTESIS RATA-RATA BEBERAPA POPULASI	127
7.4.	SINTAKS PROGRAM R	139
7.5.	LATIHAN	140
8	REGRESI EKSPLORASI DAN KONFIRMASI	142
8.1.	ANALISIS REGRESI EKSPLORASI	142
8.1.1.	Estimasi Koefisien Regresi	143
8.1.2.	Meluruskan	150

8.2.	ANALISIS REGRESI KONFIRMASI	156
8.2.1.	Estimasi Koefisien Regresi	157
8.2.2.	Koefisien Korelasi r^2	159
8.3.	SINTAKS PROGRAM	163
8.4.	LATIHAN	165
9	ANALISIS DATA KATEGORIK	168
9.1.	DATA CACAH TABEL KATEGORIK 2×2	168
9.1.1.	Uji Homogenitas	170
9.1.2.	Uji Independensi	173
9.2.	DATA CACAH TABEL KATEGORIK $p \times q$	180
9.2.1.	Uji Homogenitas	181
9.2.2.	Uji Independensi	188
9.3.	SINTAKS PROGRAM	196
9.4.	LATIHAN	200
	INDEKS	205
	DAFTAR PUSTAKA	207
A	TABEL DISTRIBUSI NORMAL	208
B	TABEL DISTRIBUSI t	209
C	TABEL DISTRIBUSI F	210
D	TABEL DISTRIBUSI CHI-SQUARE	211

DAFTAR TABEL

1.1	Angka Tingkat Kematian Karena Penyakit Jantung, per 100.000	2
1.2	Tingkat Kematian Karena Penyakit Jantung: Laki-laki Usia 20-29 dan Usia 30-39 Tahun	4
1.3	Tingkat Kematian Karena Penyakit Jantung, Laki-Laki Usia 20-29 Tahun	5
1.4	Daftar <i>Tally</i> dari Tabel 1.2 Dibulatkan	6
1.5	Tingkat Kematian Karena Penyakit Jantung: Laki-laki Usia 30-39 dan Usia 40-49 Tahun	8
1.6	Daftar <i>Tally</i> dari Tabel 1.5, Dibulatkan	8
1.7	Diagram Batang dan Daun Usia 30-39 Tahun	10
1.8	Diagram Batang dan Daun Berdampingan	12
1.9	Batang dan Daun Bertolak Belakang	13
1.10	Diagram Batang dan Daun Ganda	14
1.11	Diagram Batang dan Daun Setengah Batang	14
1.12	Diagram Batang dan Daun Usia 30-39 Tahun	16
1.13	Diagram Batang dan Daun Bertolak Belakang Bernama	17
1.14	Distribusi Frekuensi Kelompok Laki-Laki Usia 20-29 Tahun	19
1.15	Distribusi Frekuensi Kelompok Laki-Laki Usia 30-39 Tahun	19
1.16	Distribusi Frekuensi Kelompok Laki-Laki Usia 40-49 Tahun	20
1.17	Jumlah Pengangguran di Beberapa Kota Jawa Tengah dalam Ribuan	26
4.1	Diagram Batang dan Daun Sebelum dan Sesudah Pusat Dikeluarkan	64
4.2	Pengamatan Curah Hujan Selama 8 Hari	67

4.3	Ringkasan Numerik Data Tabel 1.1	68
5.1	Penduduk (dalam jutaan) 1990-2012 dan Pertumbuhannya	73
5.2	Nisbah Menurut Tahun Sensus Berurutan Untuk DIY dan Jawa Tengah	75
5.3	Diagram Batang dan Daun Nisbah DIY dan Jawa Tengah	76
5.4	Logaritma Pertumbuhan Penduduk	77
5.5	Tangga Transformasi	80
5.6	Hasil Panen Padi di Provinsi DIY Per-Kabupaten	83
5.7	Transformasi yang Sesuai dengan Nisbah	87
6.1	Pengambilan Sampel dengan Tabel Bilangan Random .	93
6.2	Penduduk Desa Swasembada	93
6.3	Sampel Penduduk Desa Swasembada	94
6.4	Sampel Berukuran 2, 3, dan 4	95
6.5	Diagram Batang dan Daun Rata-rata	96
6.6	Diagram Batang dan Daun	100
6.7	Populasi Hipotetis	114
7.1	Kemungkinan Keputusan Pengujian	118
7.2	Kemungkinan Keputusan Pengadilan	118
7.3	Frekuensi Kata Positif dan Lawan Katanya	125
7.4	Distribusi Gabungan	130
7.5	Distribusi Terpisah	131
7.6	Tabel Hasil Observasi k Perlakuan	133
7.7	Tabel Anava	134
7.8	Kritikan Menurut IQ (Ditransformasikan dengan Mengambil Akarnya)	135
7.9	Tabel Anava Data Kritikan	136
8.1	Data Pasangan X dan Y	143
8.2	$Sisa = Y - \hat{Y}$	148
8.3	Transformasi $X' = \frac{-1000}{X}$	149

8.4	Penduduk Surakarta, 1990 - 2012 (Dalam Jutaan) . . .	151
8.5	Transformasi Yang Sesuai Dengan Berbagai Kurva . .	154
8.6	Pasangan harga X dan Y	158
8.7	Data Penyakit Diabetes	165
8.8	Pendapatan dan Pengeluaran	165
9.1	Tabel Kategorik 2×2 , Kategori A dan B	169
9.2	Tabel Kategorik 2×2 , Banyaknya Sukses dan Gagal .	170
9.3	Tabel Kategorik 2×2	172
9.4	Tabel Kategorik 2×2	173
9.5	Tabel Kategorik 2×2	174
9.6	Tabel Kategorik 2×2	176
9.7	Tabel Kategorik 2×2	178
9.8	Tabel Kategorik $p \times q$	180
9.9	Tabel Kategorik $p \times q$	182
9.10	Tabel Kategorik 3×2	184
9.11	Tabel Harga $\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$	185
9.12	Tabel Kategorik 3×3	186
9.13	Tabel Harga $\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$	187
9.14	Tabel Kategorik $p \times q$	188
9.15	Tabel Kategorik 3×3	191
9.16	Tabel Kategorik 3×3	192
9.17	Tabel Harga E_{ij}	193
9.18	Tabel Harga E_{ij}	193
9.19	Tabel Kategorik 2×4	194
9.20	Tabel Kategorik 2×2	200
9.21	Tabel Kategorik 2×2	201
9.22	Tabel Kategorik 5×4	202
9.23	Tabel Kategorik 3×3	202
9.24	Tabel Kategorik 4×3	204

DAFTAR GAMBAR

1.1	Histogram Kelompok Usia 20-29 Tahun.	19
1.2	Histogram Kelompok Usia 30-39 Tahun.	20
1.3	Histogram Kelompok Usia 40-49 Tahun.	20
4.1	Diagram Kotak dan Titik	58
4.2	Diagram Kotak dan Titik dengan Harga Ekstrem	60
4.3	Diagram Kotak dan Titik Sebelum dan Sesudah Pusat Dikeluarkan	64
4.4	Diagram Kotak dan Titik Angkatan Asli	68
4.5	Diagram Kotak dan Titik Pusat Dikeluarkan	69
4.6	Diagram Kotak dan Titik Hasil Standardisasi	70
5.1	Diagram Kotak dan Titik Jumlah Penduduk DIY dan Jawa Tengah	74
5.2	Diagram Kotak dan Titik Logaritma Penduduk DIY dan Jawa Tengah	77
5.3	Diagram Kotak dan Titik Log dan $-1/\text{Akar}$	79
5.4	Diagram Kotak dan Titik Jumlah Penduduk DIY, Akar, Log dan $-1/\text{Akar}$	81
5.5	Diagram Kotak dan Titik	84
5.6	Diagram Kotak dan Titik Logaritma	84
5.7	Diagram Kotak dan Titik Akar	85
5.8	Diagram Kotak dan Titik Minus Kebalikan	85
5.9	Diagram Kotak dan Titik Minus Kebalikan Kuadrat	86
6.1	Distribusi Normal	99
6.2	Distribusi Normal Dengan Variansi Berbeda	101
6.3	Distribusi t	110
6.4	Distribusi F	112

7.1	Diagram Kotak dan Titik Data Kritikan	137
8.1	Diagram <i>Skater</i> Data	144
8.2	Pertigaan Bawah, Tengah, dan Atas	144
8.3	Pusat Pertigaan Bawah, Tengah, dan Atas	150
8.4	Diagram Skater Penduduk Surakarta	152
8.5	Skater Diagram $\log Y = Y'$ versus X	155

ugmpress.ugm.ac.id