

*Penerbit*  
**TOHAR MEDIA**

Editor :  
Wahyuddin S  
Nurdjanah Hamid

# PENGANTAR DATASCIENCE

M. Syauqi Haris, Qomarotun Nurlaila, Ratnadewi, John Friadi,  
Satya Arisena Hendrawan, Dwiny Meidelfi, Rofiq Harun, Defni,  
Garcia, A. Ratna Sari Dewi, Aisyah Mutia Dawis

# **PENGANTAR DATA SCIENCE**

## **Penulis**

M. Syauqi Haris, Qomarotun Nurlaila, Ratnadewi, John Friadi,  
Satya Arisena Hendrawan, Dwiny Meidelfi, Rofiq Harun,  
Defni, Garcia, A. Ratna Sari Dewi, Aisyah Mutia Dawis

## **Editor**

Wahyuddin S  
Nurdjanah Hamid

Penerbit

# **TOHAR MEDIA**

## PENGANTAR DATA SCIENCE

### Penulis :

M. Syauqi Haris, Qomarotun Nurlaila, Ratnadewi, John Friadi,  
Satya Arisena Hendrawan, Dwiny Meidelfi, Rofiq Harun, Defni,  
Garcia, A. Ratna Sari Dewi, Aisyah Mutia Dawis

**Editor :** Wahyuddin S, Nurdjanah Hamid

**ISBN :** 978-623-8148-23-3

### Desain Sampul dan Tata Letak

Ai Siti Khairunisa

### Penerbit

CV. Tohar Media

**Anggota IKAPI No. 022/SSL/2019**

### Redaksi :

JL. Rappocini Raya Lr 11 No 13 Makassar

JL. Hamzah dg. Tompo. Perumahan Nayla Regency Blok D No.25

Gowa Telp. 0852-9999-3635/0852-4352-7215

Email : [toharmedia@yahoo.com](mailto:toharmedia@yahoo.com)

Website : <https://toharmedia.co.id>

**Cetakan Pertama Februari 2023**

**Hak Cipta dilindungi undang-undang.** Dilarang memperbanyak sebagian atau seluruh isi buku ini dalam bentuk apapun, baik secara elektronik maupun mekanik termasuk memfotocopy, merekam atau dengan menggunakan sistem penyimpanan lainnya, tanpa izin tertulis dari penerbit.

#### Undang-undang Nomor 19 Tahun 2002 Tentang Hak Cipta

1. Barang siapa dengan sengaja dan tanpa hak mengumumkan atau memperbanyak suatu ciptaan atau memberi izin untuk itu, dipidana dengan pidana penjara paling lama 7 (Tujuh) tahun dan/atau denda paling banyak **Rp. 5.000.000.000,00 (Lima Miliar Rupiah)**
2. Barang siapa dengan sengaja menyiarkan, memamerkan, mengedarkan, atau menjual kepada umum suatu ciptaan atau barang hasil pelanggaran hak cipta atau hak terkait sebagaimana dimaksud pada ayat 1, dipidana paling lama 5 (lima tahun) dan/atau denda paling banyak **Rp. 500.000.000,00 (Lima Ratus Juta Rupiah)**

## Kata Pengantar

Pujian dan limpahan syukur selalu kami panjatkan kehadiran Allah SWT yang telah memberikan bermacam nikmat dan karunia-Nya sehingga penulis dapat menyelesaikan buku dengan judul “**PENGANTAR DATA SCIENCE**” dengan tepat waktu tanpa ada kendala yang berarti. Adapun tujuan dari penulisan ini adalah untuk memudahkan orang banyak mengetahui apa manfaat dari rumput laut dan pengelolaan hingga sampai ke tahap konsumsi.

Kesuksesan dalam penyusunan buku ini tentunya bukan semata atas usaha dari penulis saja. Ada banyak pihak yang turut membantu dalam tercapainya kesuksesan yang sekarang ini dapat dirasakan. Baik dari pihak yang membantu dalam bentuk dukungan moril maupun materil. Penulis buku ini pun dari berbagai universitas yang ada di Indonesia, dengan itu ada banyak pemahaman yang bisa kita ambil dari buku ini. Oleh karena itu sang penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh pihak yang turut berpartisipasi dalam tercapainya kesuksesan dari buku ini.

Buku yang ada dihadapan para pembaca sekalian ini pasti memiliki sangat banyak kekurangan dan masih jauh dari kata sempurna. Sehingga penulis banyak berharap kepada para pembaca sekalian untuk memberikan kritik dan sarannya agar buku ini dapat menjadi buku yang lebih sempurna dan lengkap.

Bandung, 14 Desember 2022

Penulis

## DAFTAR ISI

<b>Halaman Depan</b>	<b>_i</b>
<b>Halaman Penerbit</b>	<b>_ii</b>
<b>Kata Pengantar</b>	<b>_iii</b>
<b>Daftar Isi</b>	<b>_iv</b>
<b>Bab 1. Konsep Data Science</b>	<b>_1</b>
1.1. Data Science Saat Ini	_1
1.2. Pengertian Data Science	_3
1.3. Bagaimana Data Science Bekerja	_5
1.4. Apa Yang Harus Dipelajari	_7
<b>Bab 2. Konsep Statistika untuk Data Science</b>	<b>_10</b>
2.1. Pendahuluan	_10
2.2. Model Data Science	_14
2.3. Statistika dalam Data Science	_14
2.4. Jenis data Statistika	_17
2.5. Pengukuran dan Perbandingan Data	_18
2.6. Probabilitas	_23
2.7. Penerapan Statistik Untuk Data Science	_26
<b>Bab 3. Hierarki Data Clustering</b>	
3.1. Pengantar	_29
3.2. Metode Hierarki Clustering	_32
3.3. Single-Linkage Clustering	_34
3.4. Complete-Linkage Clustering	_35
3.5. k-Means Clustering	_38
3.6. Contoh k-Means Clustering pada dunia kerja	_40
3.7. Perilaku MSB, MSE, dan Pseudo-F sebagai pemroses Algoritma k-Means	_45
<b>Bab 4. SQL Basis Data</b>	<b>_51</b>
4.1. Pengantar	_51
4.2. Sintak SQL	_53
4.3. Sintak SQL AND, OR dan NOT	_57
4.4. Sintak SQL ORDER BY	_59
4.5. Sintak SQL INSERT INTO	_60
4.6. Sintak SQL SELECT	_60

<b>Bab 5. R For Data Scientist</b>	<b>_63</b>
5.1. Pengenalan R Programmin	_63
5.2. Dasar Pemrograman R	_69
5.3. Fungsi dan Paket Library pada R	_77
5.4. Logika Loop dan IF pada R	_79
5.5. <i>Import dan Cleaning</i> Data	_82
5.6. Visualisasi Data	_87
5.7. Penutup	_88
<b>Bab 6. Python For Data Scientist</b>	<b>_91</b>
6.1. Pengantar	_91
6.2. Modelling	_91
6.3. Learning	_98
6.4. Analisis Eksplorasi	_102
<b>Bab 7. Data Penginderaan Jauh Satelit</b>	<b>_115</b>
7.1. Pengantar	_115
7.2. Pengertian Penginderaan Jauh Menurut Para Ahli	_115
7.3. Sejarah Penginderaan Jauh	_117
7.4. Komponen-komponen Penginderaan Jauh	_118
7.5. Keunggulan, Keterbatasan dan Kelemahan Penginderaan Jauh	_122
<b>Bab 8. Memahami Visualisasi Data</b>	<b>_127</b>
8.1. Pengantar	_127
8.2. Tujuan, Fungsi, dan Jenis	_128
8.3. Penutup	_134
<b>Bab 9. Quantitative Mini Research : Analisis Regresi Berorientasi Kasus Kesehatan, Ketimpangan, dan Kemiskinan</b>	<b>_137</b>
9.1. Latar Belakang	_137
9.2. Analisis Makro-Komparatif dan Analisis Regresi Berganda	_138
9.3. Ketimpangan dan Kesehatan	_140
9.4. Analisis Regresi OLS	_141
9.5. Membalikkan Regresi Luar Dalam	_144
9.6. Mempelajari dari Kasus	_155
9.7. Penutup	_159

**Bab 10. Learning \_161**

10.1. Pengantar Machine Learning \_161

10.2. AI Learning Model : Knowledge-Based  
Classification \_162

10.3. AI Learning Models: Feedback-Based  
Classification \_162

10.4. Mamfaat dari *Machine Learning* \_182

10.5. Bagaimana machine Learning Bekerja \_183

10.6. Penutup \_184

**Bab 11. Teknologi Big Data \_185**

11.1. Sejarah Big Data \_185

11.2. Potensi Big Data \_190

11.3. Penerapan Big Data \_191

11.4. Penutup \_195

**Datar Pustaka \_196**