

4395792
3852575
4637457
4567356
3967945
2126548
1312638
8674424

DATA SCIENCE

TIM PENULIS:

Yunita Ardilla, Guntoro, Surya Afnarius, Agustinus Budi Santoso, Rezania Agramanisti Azdy,
Rona Putra, Nur Fitrianiingsih Hasan, Eka Prasetya Adhy Sugara, Yohanssen Pratama,
Hidayatul Fitri, Ashari Wicaksono, Arnita

DATA SCIENCE

TIM PENULIS:

Yunita Ardilla, Guntoro, Surya Afnarius, Agustinus Budi Santoso, Rezanía Agramanisti Azdy,
Rona Putra, Nur Fitrianiingsih Hasan, Eka Prasetya Adhy Sugara, Yohanssen Pratama,
Hidayatul Fitri, Ashari Wicaksono, Arnita



DATA SCIENCE

Tim Penulis:

**Yunita Ardilla, Guntoro, Surya Afnarius, Agustinus Budi Santoso, Rezania Agramanisti
Azdy, Rona Putra, Nur Fitrianiingsih Hasan, Eka Prasetya Adhy Sugara, Yohanssen
Pratama, Hidayatul Fitri, Ashari Wicaksono, Arnita**

Desain Cover:

Helmaria Ulfa

Tata Letak:

Handarini Rohana

Editor:

Evi Damayanti

ISBN:

978-623-459-158-3

Cetakan Pertama:

Agustus, 2022

Hak Cipta 2022, Pada Penulis

Hak Cipta Dilindungi Oleh Undang-Undang

Copyright © 2022

by Penerbit Widina Bhakti Persada Bandung

All Right Reserved

Dilarang keras menerjemahkan, memfotokopi, atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari Penerbit.

PENERBIT:

WIDINA BHAKTI PERSADA BANDUNG

(Grup CV. Widina Media Utama)

Komplek Puri Melia Asri Blok C3 No. 17 Desa Bojong Emas
Kec. Solokan Jeruk Kabupaten Bandung, Provinsi Jawa Barat

Anggota IKAPI No. 360/JBA/2020

Website: www.penerbitwidina.com

Instagram: @penerbitwidina

KATA PENGANTAR

Rasa syukur yang teramat dalam dan tiada kata lain yang patut kami ucapkan selain mengucap rasa syukur. Karena berkat rahmat dan karunia Tuhan Yang Maha Esa, buku yang berjudul “Data Science” telah selesai disusun dan berhasil diterbitkan, semoga buku ini dapat memberikan sumbangsih keilmuan dan penambah wawasan bagi siapa saja yang memiliki minat terhadap pembahasan tentang Data Science.

Akan tetapi pada akhirnya kami mengakui bahwa tulisan ini terdapat beberapa kekurangan dan jauh dari kata sempurna, sebagaimana pepatah menyebutkan *“tiada gading yang tidak retak”* dan sejatinya kesempurnaan hanyalah milik Tuhan semata. Maka dari itu, kami dengan senang hati secara terbuka untuk menerima berbagai kritik dan saran dari para pembaca sekalian, hal tersebut tentu sangat diperlukan sebagai bagian dari upaya kami untuk terus melakukan perbaikan dan penyempurnaan karya selanjutnya di masa yang akan datang.

Terakhir, ucapan terima kasih kami sampaikan kepada seluruh pihak yang telah mendukung dan turut andil dalam seluruh rangkaian proses penyusunan dan penerbitan buku ini, sehingga buku ini bisa hadir di hadapan sidang pembaca. Semoga buku ini bermanfaat bagi semua pihak dan dapat memberikan kontribusi bagi pembangunan ilmu pengetahuan di Indonesia.

Agustus, 2022

Penulis

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI	iv
BAB 1 INTRODUCTION TO DATA SCIENCE	1
A. Pendahuluan	2
B. Manfaat dan Kegunaan <i>Data Science</i> dan <i>Big Data</i>	4
C. Aspek Data	5
D. Proses <i>Data Science</i>	9
E. Rangkuman Materi	11
BAB 2 INTRODUCTION TO STATISTICAL FOR DATA SCIENCE	15
A. Pendahuluan	16
B. Keuntungan Statistik Dalam <i>Data Science</i>	18
C. Jenis Data Statistik	20
D. Probabilitas	22
E. Populasi dan Sampel	23
F. Statistik Deskriptif	24
G. Penerapan Statistik Untuk <i>Data Science</i>	28
H. Rangkuman Materi	28
BAB 3 INTRODUCTION SQL FOR DATA SCIENCE	31
A. Pendahuluan	32
B. Kenapa Keterampilan SQL Diperlukan Oleh <i>Data Scientist</i> ?	34
C. Kategori SQL	35
D. Topik-Topik Penting SQL Untuk <i>Data Science</i>	39
E. Rangkuman Materi	41
BAB 4 SQL BASIC QUERY	45
A. Pendahuluan	46
B. Pengenalan SQL	47
C. Software Pendukung SQL Terhadap DBMS	48
D. Sintaks Dasar SQL	51
E. Rangkuman Materi	60

BAB 5 SQL INTERMEDIATE QUERY	63
A. Pendahuluan.....	64
B. <i>Operator SQL</i>	65
C. <i>Group By</i>	70
D. <i>Having</i>	71
E. <i>Order By</i>	72
F. <i>Join</i>	73
G. Rangkuman Materi	79
BAB 6 SQL UPPER INTERMEDIATE – ADVANCED QUERY	83
A. Pendahuluan.....	84
B. <i>Statemen Sub Query</i>	85
C. Logika Penggabungan	87
D. Logika Percabangan.....	88
E. Logika Perulangan.....	90
F. <i>Statemen Cursor</i>	90
G. <i>Function</i>	92
H. <i>Store Procedure</i>	93
I. <i>View</i>	95
J. <i>Triggers</i>	97
K. Contoh Kasus	98
L. Rangkuman Materi	101
BAB 7 PYTHON FOR DATA SCIENCE.....	105
A. Pendahuluan.....	106
B. Python Untuk <i>Data Science</i>	107
C. Library Python Untuk <i>Data Science</i>	108
D. <i>Data Preprocessing</i>	112
E. <i>Data Eksplorasi & Analyze</i>	119
F. Rangkuman Materi	123
BAB 8 DATA VISUALIZATION	127
A. Pendahuluan.....	128
B. Data dan Konteks Data	130
C. Media Visualisasi Data.....	131
D. Persiapan Data Dengan <i>Google Sheets</i>	136
E. Transformasi Data Ke Bentuk Visual	144
F. Beberapa Kesalahan Dalam Visualisasi Data	149

G. Rangkuman Materi	151
BAB 9 CREATE DASHBOARD WITH DATA STUDIO	155
A. Pendahuluan	156
B. Google Data Studio	156
C. Bagaimana Cara Menggunakan Data Studio?	157
D. Google Data Studio: Sumber Data dan Konektor	159
E. Sebelum Membuat Dashboard Pada Google Data Studio	160
F. Membuat Dashboard di Google Data Studio	163
G. Cara Menambahkan Sumber Data <i>Google Spreadsheet</i> di Data Studio	164
H. Cara Membuat Laporan Kosong Pertama di Google Data Studio	167
I. Alat Edit Laporan	169
J. Panel Tema dan Tata Letak	171
K. Rangkuman Materi	172
BAB 10 INTRODUCTION MACHINE LEARNING	175
A. Pendahuluan	176
B. Metode <i>Machine Learning</i>	178
C. Rangkuman Materi	184
BAB 11 CLASSIFICATION	187
A. Pendahuluan	188
B. Materi	188
C. Jenis-Jenis Algoritma Klasifikasi	194
D. Rangkuman Materi	205
BAB 12 REGRESSION	207
A. Pendahuluan	208
B. Regresi Linear Sederhana	208
C. Regresi Linear Berganda	212
D. Regresi Non Linear	217
E. Rangkuman Materi	218
GLOSARIUM	223
PROFIL PENULIS	227



DATA SCIENCE

BAB 1: *INTRODUCTION TO DATA SCIENCE*

Yunita Ardilla, S.Kom., M.M.T

UIN Sunan Ampel Surabaya

BAB 1

INTRODUCTION TO DATA SCIENCE

A. PENDAHULUAN

Data Science memungkinkan suatu bisnis dapat memproses sejumlah data yang besar baik berupa data terstruktur maupun tidak terstruktur guna mendeteksi pola-pola dalam data tersebut. Hal ini mampu memberikan keuntungan bagi perusahaan atau bisnis dalam peningkatan efisiensi, identifikasi peluang pasar baru, meningkatkan keunggulan bersaing dan sebagai salah satu pendukung dalam pembuatan keputusan.

Ketika berbicara mengenai data science maka hal ini juga akan bersinggungan dengan istilah big data, yaitu merupakan kumpulan-kumpulan data yang begitu besar dan kompleks sehingga sulit untuk dilakukan proses analisisnya jika hanya menggunakan teknik manajemen data tradisional seperti RDBMS (Sistem Manajemen Basis Data Relasional). Data science melibatkan penggunaan perpaduan beberapa metode untuk menganalisis data yang berjumlah besar dan kompleks untuk mengekstrak pengetahuan yang tersembunyi di dalamnya. Hubungan antara big data dan data science ibarat seperti hubungan antara minyak mentah dan kilang minyak. Dimana data science dan big data telah berevolusi dari statistik dan manajemen data tradisional berubah menjadi disiplin ilmu yang berbeda.

Adapun karakteristik dari big data meliputi:

1. Volume: Berapa banyak data yang ada?
Maksudnya adalah jumlah atau kuantitas data yang dihasilkan dari banyak transaksi serta volume data yang disimpan. Pada umumnya big data biasanya menggunakan skala Terabytes (per 1000 Gigabytes)
2. Variasi: seberapa beragam jenis data yang berbeda?
Maksudnya adalah apakah data bersifat terstruktur atau tidak terstruktur. Data terstruktur merupakan data yang mempunyai elemen-elemen yang dapat diakses seperti keys (*primary keys*,

DAFTAR PUSTAKA

- Jamil, M. Rosihan. Fuad, A. (2016). *Cloud Computing Teori dan Aplikasi*. Deepublisher.
- Ardilla, Y. (2019) '*Short Term Forecasting of Electricity Load: A Comparison of Methods to Paiton Subsystem East Java & Bali*', *Journal of Information Technology and Computer Science*, 4(3), p. 284. doi: 10.25126/jitecs.201943159.
- Kebutuhan, S., Lunak, P., Pembangunan, I. F., Lunak, P., Zulen, A. A., & Bachtiar, A. M. (2011). *Jurusan Teknik Informatika - Institut Teknologi Bandung GL01-ATM*.
- Marasco, J. (2007). *Software requirements*. In *Dr. Dobb's Journal* (Vol. 32, Issue 9). <https://doi.org/10.1201/b12149-5>
- Barcelos, L. V., & Penteado, R. D. (2017). *Elaboration of software requirements documents by means of patterns instantiation*. *Journal of Software Engineering Research and Development*, 5(1). <https://doi.org/10.1186/s40411-017-0038-9>



DATA SCIENCE

BAB 2: *INTRODUCTION TO STATISTICAL FOR DATA SCIENCE*

Guntoro, S.T., M.Kom

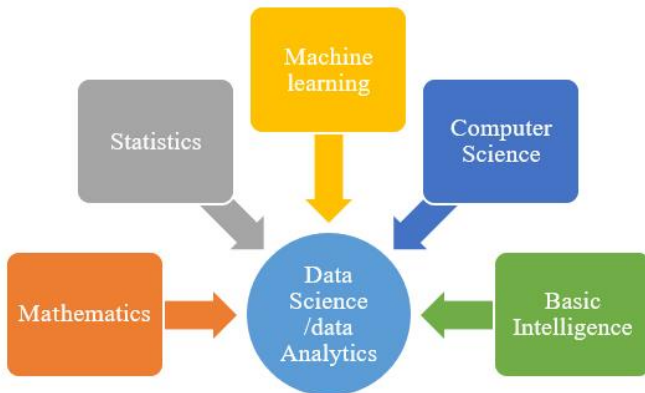
Universitas Lancang Kuning

BAB 2

INTRODUCTION TO STATISTICAL FOR DATA SCIENCE

A. PENDAHULUAN

Sebelum kita mempelajari ilmu statistik dalam penerapan *data science*, kita akan mengenal apa itu statistik, statistika dan *data science*. Seperti yang terlihat pada gambar 1. *Data science* mencakup berbagai keilmuan yaitu matematika, *statistic*, *machine learning*, ilmu komputer dan *basic intelligence*.



Gambar 1. Keilmuan yang Terlibat dalam *Data Science*

1. Statistika

Statistika adalah salah satu bidang ilmu yang berfokus pada data numerik. Di sisi lain, statistik adalah data mentah yang dikumpulkan untuk mempelajari lebih lanjut tentang masalah tertentu. Oleh karena itu, statistika merupakan bidang studi yang berfokus pada data. Dimulai dengan pengumpulan data, analisis data, dan penyajian data.

DAFTAR PUSTAKA

- Efron, B., & Hastie, T. (2021). *Computer Age Statistical Inference, Student Edition: Algorithms, Evidence, and Data Science (Vol. 6)*. Cambridge University Press.
- Ott, R. L., & Longnecker, M. T. (2015). *An introduction to statistical methods and data analysis*. Cengage Learning.
- Meloun, M., & Militký, J. (2011). *Statistical data analysis: A practical guide*. Woodhead Publishing Limited.
- Efron, B., & Tibshirani, R. (1991). *Statistical data analysis in the computer age*. *Science*, 253(5018), 390-395.
- Diggle, P. J. (2015). *Statistics: a data science for the 21st century*. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(4), 793-813.



DATA SCIENCE

BAB 3: *INTRODUCTION SQL FOR DATA SCIENCE*

Ir. Surya Afnarius, M.Sc, Ph.D

Departeman Sistem Informasi FTI Universitas Andalas

BAB 3

INTRODUCTION SQL FOR DATA SCIENCE

A. PENDAHULUAN

SQL adalah konsep dan alat analisis data yang penting untuk dipelajari dan digunakan dalam *Data Science* (Lateef, 2022; Teate, 2021; Tiwari, 2021). SQL ini dipakai untuk melakukan query dan mengelola data yang disimpan di dalam database relasional, seperti MySQL, MsSQL, Oracle dan PostgreSQL (Lateef, 2022; Teate, 2021; Tiwari, 2021; Ippolito, 2019). Database relasional itu dibentuk oleh kumpulan tabel dua dimensi yang terdiri dari sejumlah kolom yang tetap dan sejumlah baris yang dapat berubah-ubah banyaknya.

Menurut Teate (2021), Tiwari (2021) dan Ste (2019), SQL bersama dengan Python dan R, telah dianggap sebagai tiga keterampilan yang paling banyak diminta di dalam Data Science (lihat Gambar 3.1 dan Gambar 3.2). Salah satu keuntungan utama dalam menggunakan SQL adalah ketika melakukan operasi terhadap data di dalam database, data tersebut diakses secara langsung tanpa perlu menyalinnya. Kemampuan SQL ini dapat mempercepat eksekusi alur kerja.

DAFTAR PUSTAKA

- GeeksforGeeks. (2021). SQL | DDL, DQL, DML, DCL and TCL Commands. Diakses 5 Juni 2022 dari <https://www.geeksforgeeks.org/sql-ddl-dql-dml-dcl-tcl-commands/>
- Ippolito, P. P. (2019). *SQL For Data Science*. Diakses 5 Juni 2022 dari <https://towardsdatascience.com/sql-for-data-science-f9a42ae66bbb>
- Jidge, A. (2020). *Everything That Needs to Master SQL Skills For Data Science*. Diakses 5 Juni 2022 dari <https://towardsdatascience.com/everything-that-needs-to-master-sql-skills-for-data-science-dca851995b8d>
- Kuncoro, A. A. (2021). Manfaat Memahami Bahasa Query Dalam Pemrograman. Diakses 5 Juni 2022 dari <http://teknik-informatika-s1.stekom.ac.id/informasi/baca/Manfaat-Memahami-Bahasa-Query-Dalam-Pemrograman/925d1b50cf96fe5447879b4b57ef57c168e4fc34>
- Lateef, Z. (2022). *SQL For Data Science: One stop Solution for Beginners*. Diakses 5 Juni 2022 dari <https://www.edureka.co/blog/sql-for-data-science/>
- Metwalli, S. A. (2020). *Databases 101: Introduction to Databases for Data Scientists*. Diakses 5 Juni 2022 dari <https://towardsdatascience.com/databases-101-introduction-to-databases-for-data-scientists-ee18c9f0785d>
- Ste, A. (2019). *How to Become More Marketable as a Data Scientist*. Diakses 5 Juni 2022 dari <https://www.kdnuggets.com/2019/08/marketable-data-scientist.html>
- Tanimura, C. (2021). *SQL for Data Analysis*. Sebastopol: O'Reilly Media, Inc.
- Teate, R.M.P. (2021). *SQL for Data Scientists*. Hoboken, John Wiley & Sons, Inc.
- Techopedia. (2014). *Data Manipulation Language (DML)*. Diakses 5 Juni 2022 dari <https://www.techopedia.com/definition/1179/data-manipulation-language-dml>

- Techopedia. (2020). *Data Definition Language (DDL)*. Diakses 5 Juni 2022 dari <https://www.techopedia.com/definition/1175/data-definition-language-ddl>
- Tiwari, N. (2021). *SQL For Data Science: A Beginner's Guide!* Diakses 5 Juni 2022 dari <https://www.analyticsvidhya.com/blog/2021/06/sql-for-data-science-a-beginners-guide/>



DATA SCIENCE

BAB 4: *SQL BASIC QUERY*

Agustinus Budi Santoso, S.S.T., M.Cs

Universitas Sains dan Teknologi Komputer

BAB 4

SQL BASIC QUERY

A. PENDAHULUAN

SQL dari bahasa lengkapnya *Structured Query Language* adalah *Query* merupakan bentuk inti dari penggunaan bahasa dalam database. Penggunaan bahasa standar SQL merupakan bentuk bahasa pemrograman yang memiliki alur sejarah perkembangan dari awal mulanya yang panjang sampai saat ini (Dewaweb, n.d.). Basis data atau database merupakan sesuatu hal yang dapat digunakan dalam penerapan SQL yang sangat penting dan perlu diperhatikan dalam struktur dalam Query Dasar atau Basic Query. MySQL merupakan software database open source yang sering digunakan untuk mengolah basis data yang menggunakan bahasa SQL (Sitinjau Daniel Dido Jantce TJ & Suwita, 2020). MySQL sendiri merupakan bentuk salah satu aplikasi yang menggunakan DBMS dan sangat banyak digunakan di berbagai pemrogram aplikasi desktop, web maupun mobile.

Bahasa *Query* dalam DBMS sangat menjadi dasar dalam setiap pemrograman dan terbagi menjadi dua. DDL (*Data Definition Language*) biasanya digunakan untuk menjelaskan tabel dalam database dan satunya lagi adalah DML (*Data Manipulation Language*) yang berguna atau di alokasikan untuk membentuk manipulasi data dengan *query* atau menampilkan informasi sesuai permintaan di dalam database sehingga menampilkan informasi yang dapat dimanfaatkan dan dibaca, *Sub-Query* juga merupakan bagian DML yang diartikan adalah *query* di dalam *query*. Sehingga dalam mendapatkan suatu informasi pada isi tabel ataupun dalam view atau relasi antar tabel sangat kompleks, kita bisa menggunakan berbagai kondisi atau klausa dengan memberikan permintaan perintah tertentu.

DAFTAR PUSTAKA

- Dewaweb. (n.d.). *Pengertian SQL, Fungsi, Beserta Perintah Dasarnya*. <https://www.dewaweb.com/blog/sql-pengertian-fungsi-beserta-perintah-dasarnya/>.
- Meidi, D. P. P., & Ken, K. P. (2019). Penerjemah Teks Bahasa Indonesia Ke Dalam Dml (Data Manipulation Language) dengan Sub-Query. *Jurnal Ilmiah Komputer Dan Informatika*, 2(9), 1–8.
- nesamedia. (2022). *Pengertian DDL dan DML*. <https://www.nesamedia.com/pengertian-ddl-dan-dml/>.
- Sitinjak Daniel Dido Jantce TJ, M., & Suwita, J. (2020). Analisa Dan Perancangan Sistem Informasi Administrasi Kursus Bahasa Inggris Pada Intensive English Course Di Ciledug Tangerang. *Ipsikom*, 8(1).



DATA SCIENCE

BAB 5: *SQL INTERMEDIATE QUERY*

Rezania Agramanisti Azdy, S.Kom., M.Cs

Institut Teknologi dan Bisnis PalComTech

BAB 5

SQL INTERMEDIATE QUERY

A. PENDAHULUAN

SQL merupakan bahasa pemrograman yang dapat digunakan untuk mengelola sebuah basis data, yaitu dengan cara mengakses, mengubah, maupun memanipulasi data yang terdapat pada sebuah basis data. Definisi lain mengenai SQL adalah bahwa SQL merupakan bahasa pemrograman yang dapat digunakan untuk melakukan request yang melekat pada sebuah basis data atau SMBD tertentu seperti MySQL, PostgreSQL, Interbase, dan Oracle (Nugroho, 2008). MySQL sendiri merupakan perangkat lunak Sistem Manajemen Basis Data yang multithread, multiuser, dan menggunakan SQL sebagai bahasanya (Andrianto dan Nursikuwagus, 2017).

Pada bahasan sebelumnya, telah diberikan *syntax* maupun *query* dasar SQL yang dikategorikan ke dalam Data Definition Language (DDL) maupun Data Manipulation Language (DML). Pada Bab ini akan dibahas *query SQL* lanjutan atau yang dikategorikan ke dalam *SQL Intermediate Query* sekaligus menjembatani pembahasan pada materi sebelumnya dengan materi *query SQL advanced* yang terdapat pada Bab berikutnya.

SQL intermediate query yang akan dibahas pada Bab ini yaitu:

- ✓ *Operator SQL:*
 - Operator Aritmatika
 - Operator Perbandingan
 - Operator Logika
- ✓ *Group By*
- ✓ *Having*
- ✓ *Order By*
- ✓ *Join*
 - *Inner Join*
 - *Left Join*

DAFTAR PUSTAKA

- Andrianto, P., & Nursikuwagus, A. (2017). Sistem Informasi Pelayanan Kesehatan Berbasis Web di Puskesmas. Prosiding Seminar Nasional Komputer dan Informatika (SENASKI) 2017.
- Nugroho, B. (2008). Panduan lengkap menguasai perintah SQL. Mediakita. tutorialspoint. *SQL - Operators*. Diakses pada 22 Juni 2022, dari <https://www.tutorialspoint.com/sql/sql-operators.htm>
- tutorialspoint. *SQL ORDER BY Keyword*. Diakses pada 23 Juni 2022, dari https://www.w3schools.com/sql/sql_orderby.asp
- W3Schools. *SQL Operators*. Diakses pada 22 Juni 2022, dari https://www.w3schools.com/sql/sql_operators.asp
- W3School. *SQL GROUP BY Statement*. Diakses pada 23 Juni 2022, dari https://www.w3schools.com/sql/sql_groupby.asp
- W3Schools. *SQL HAVING Clause*. Diakses pada 23 Juni 2022, dari https://www.w3schools.com/sql/sql_having.asp
- W3Schools. *SQL Joins*. Diakses pada 25 Juni 2022, dari https://www.w3schools.com/sql/sql_join.asp



DATA SCIENCE

BAB 6: *SQL UPPER INTERMEDIATE - ADVANCED QUERY*

Rona Putra, M.Kom

Institut Agama Islam Negeri (IAIN) Curup

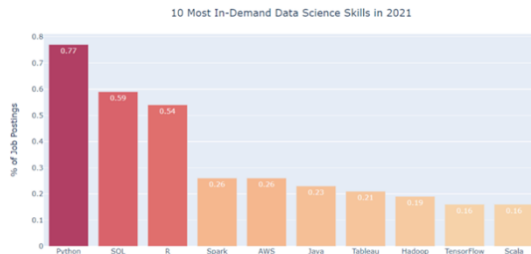
BAB 6

SQL UPPER INTERMEDIATE- ADVANCED QUERY

A. PENDAHULUAN

Pemanfaatan SQL atau *Structured Query Language* untuk mengelola visualisasi data pada bidang *Data Science* mulai dirasakan seiring berkembangnya volume data dan kebutuhan akan informasi dari hasil pengelolaan data. Statemen SQL memungkinkan pengguna dapat membuat, menghapus, memanipulasi dan memvisualisasi data yang ada pada basis data untuk kebutuhan analisis *data science* dengan perintah yang sederhana. Peningkatan fungsionalitas dari statemen SQL dan kebutuhan akan visualisasi yang diinginkan mengharuskan pengguna untuk mendalami lebih jauh statemen-statement yang ada di SQL.

Penggunaan SQL sekarang sangat populer, salah satunya karena kemudahan dalam memahami statemen-statement dalam SQL dan dianggap cukup kompatibel untuk dimanfaatkan dalam manipulasi dan visualisasi data. Lebih jauh, sekarang SQL banyak dimanfaatkan dalam analisis *Data Science*, Berikut statistik kebutuhan skill SQL bagi *Data Science* (Neelu Tiwari, 2021):



DAFTAR PUSTAKA

- Banerjee, S. (2019). *SQL Functions*. Diambil kembali dari www.janbasktraining.com:
<https://www.janbasktraining.com/blog/sql-functions/>
- Beaulieu, A. (2009). *Learning SQL, Second Edition*. Dalam A. Beaulieu, *Learning SQL* (hal. 172). Sebastopol: O'Reilly Media, Inc.
- Department of Computer Science University of Cape Town. (2017). *Database system*. Cape Town, Afrika Selatan.
- EMC Education Services. (2015). *Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Indiana: John Wiley & Sons, Inc. .
- Fungsi SQL: Bagaimana menulis Fungsi dalam SQL?* (t.thn.). Diambil kembali dari ms.ichlese.at: <https://ms.ichlese.at/sql-functions-how-write-function-sql>
- Linoff, G. S. (2016). *Data Analysis Using SQL and Excel*. Indiana: John Wiley & Sons, Inc.
- MySQL Triggers*. (2020). Diambil kembali dari www.w3resource.com:
<https://www.w3resource.com/mysql/mysql-triggers.php>
- MySQL Views*. (t.thn.). Diambil kembali dari mysqltutorial.org:
<https://www.mysqltutorial.org/mysql-views-tutorial.aspx>
- Neelu Tiwari*. (2021, Juni). Diambil kembali dari Analytics Vidhya:
<https://www.analyticsvidhya.com/blog/2021/06/sql-for-data-science-a-beginners-guide/>
- Simplilearn*. (2022, mei). Diambil kembali dari Simplilearn.com:
<https://www.simplilearn.com/tutorials/sql-tutorial/stored-procedure-in-sql>
- SQL Functions*. (t.thn.). Diambil kembali dari www.sqltutorial.org:
<https://www.sqltutorial.org/sql-functions/>



DATA SCIENCE

BAB 7: *PYTHON FOR DATA SCIENCE*

Nur Fitrianiingsih Hasan, M.Kom

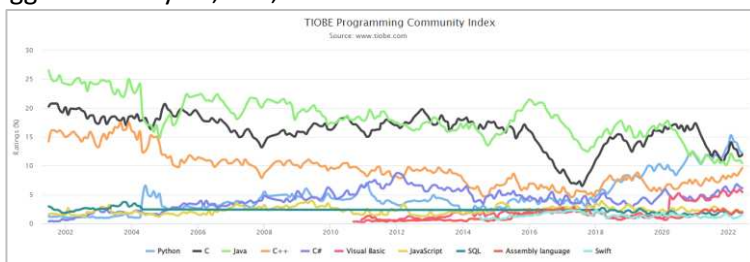
Universitas Muhammadiyah Papua

BAB 7

PYTHON FOR DATA SCIENCE

A. PENDAHULUAN

Pertanyaan yang sering muncul saat mulai mempelajari tentang data science adalah “Apa kaitannya *Python* dan *Data Science*?” Jawabannya adalah karena *python* adalah jenis bahasa tingkat tinggi yang mendekati bahasa manusia dan bersifat open source maka banyak pengembang membuat *library python* ditujukan untuk data science, memudahkan pengguna dalam melakukan analisis dan membuat model prediktif dalam *data science*. Python dapat dikatakan sebagai bahasa pemrograman dengan perkembangan tercepat. Perkembangan ini didukung oleh library python yang melimpah, *powerfull* dan mendekati bahasa manusia (Rangkuti et al., 2021), hal ini membuat python menjadi bahasa pemrograman terpopuler dalam satu dekade terakhir, paling tidak masuk dalam TOP 10 bahasa pemrograman terpopuler di dunia. Kondisi ini dibuktikan dari hasil survey sejak 2012 oleh TIOBE Index terkait bahasa pemrograman yang sering dicari melalui google, data menunjukkan per bulan Juni 2022 python berada pada peringkat 1 dengan rating 12,20% mengungguli seniorinya C, C++, dan Java.



Gambar 1. Survey 10 bahasa pemrograman terpopuler (TIOBE, 2022)

- Gunakan dataset pada <https://fitriahasan.sagootech.com/dataset/>, kemudian lakukan analisis pada file **who_data.csv** untuk menjawab negara mana yang memiliki jumlah populasi terbesar dan terkecil dengan pendekatan data science menggunakan *Python*.

DAFTAR PUSTAKA

- Afikah, P., Avorizano, A., Afandi, I. R., & Hasan, F. N. (2022). Implementasi Business Intelligence Untuk Menganalisis Data Kasus Virus Corona di Indonesia Menggunakan Platform Tableau. *Pseudocode*, 9(1), 25–32. <https://ejournal.unib.ac.id/index.php/pseudocode/article/view/20726>
- Harahap, M., Rozi, F., Yennimar, Y., & Siregar, S. D. (2021). Analisis Wawasan Penjualan Supermarket dengan Data Science. *Data Sciences Indonesia (DSI)*, 1(1), 1–7. <https://doi.org/10.47709/dsi.v1i1.1173>
- Hasan, N. F. (2020). *Penerapan Jaringan Syaraf Tiruan dengan Algoritma Backpropagation untuk Peramalan Penjualan Air Minum Dalam Kemasan* [Universitas Amikom Yogyakarta]. <https://eprints.amikom.ac.id//20/1/18.51.1063> - Nur Fitrianiingsih Hasan.pdf
- Hasan, N. F., & Fatta, H. Al. (2019). Peramalan Jumlah Penjualan Menggunakan Jaringan Syaraf Tiruan Backpropagation Pada Perusahaan Air Minum Dalam Kemasan. *Jurnal Teknik Informatika Dan Sistem Informasi (JuTISI)*, 5(2), 179–188. <http://114.7.153.31/index.php/jutisi/article/view/1607>
- Hasan, N. F., & Wati, V. (2021). Deteksi Cyberbullying pada Facebook Menggunakan Algoritma K-Nearest Neighbor. *Journal of Smart System*, 1(1), 35–44. <https://doi.org/10.36728/jss.v1i1.1605>
- Kurniawan, D. (2022). *Pengenalan Machine Learning dengan Python*. Elex Media Komputindo.
- Rangkuti, Y. M., Al Idrus, S. I., & Tarigan, D. D. (2021). *Pengantar Pemrograman Python* (R. R. Rerung (ed.)). Media Sains Indonesia.

- Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information (Switzerland)*, 11(4). <https://doi.org/10.3390/info11040193>
- Saepuloh, D. (2020). VISUALISASI DATA COVID 19 PROVINSI DKI JAKARTA MENGGUNAKAN TABLE DATA VISUALIZATION OF COVID 19 PROVINCE DKI JAKARTA USING TABLEAU bernama Severe Acute Respiratory Syndrome Perbedaan Tableau Desktop Tableau Public Open Source Berbayar (bukan open sourc. *Jurnal Riset Jakarta*, 13(2), 55–64.
- TIOBE. (2022). *TIOBE Index for June 2022*. <https://www.tiobe.com/tiobe-index/>



DATA SCIENCE

BAB 8: *DATA VISUALIZATION*

Eka Prasetya Adhy Sugara, S.T., M.Kom

Institut Teknologi dan Bisnis PalComTech

BAB 8

DATA VISUALIZATION

A. PENDAHULUAN

Istilah visualisasi data bukanlah hal baru dan juga bukan merupakan sebuah cabang ilmu baru, namun merupakan bagian dari ilmu data atau Data Science. Namun apa sebenarnya visualisasi data tersebut?. Visualisasi data merupakan sebuah cara dalam menyajikan data atau memberikan informasi dalam bentuk grafik, diagram, atau representasi visual lainnya (Tara J. Brigham, 2061). Tujuan utama dari visualisasi data adalah kemampuannya untuk memvisualisasikan data, mengkomunikasikan informasi secara jelas dan efektif (Vitaly Friedman, 2008). Mengapa perlu menggunakan visualisasi data?. Sebagai contoh, dosen memiliki data jumlah mahasiswa yang hadir per mata kuliah per hari dalam satu minggu (Tabel 1). Misalnya Ketua Program Studi (Kaprodi) hanya ingin mengetahui berapa persentase rata-rata kehadiran mahasiswa per minggu untuk semua mata kuliah saja, tentu hal ini akan menyulitkan dosen untuk memberikan penjelasan dengan efisien. Sebagai perbandingan, perhatikan Gambar 1 yang memberikan visualisasi data Tabel 1. Berdasarkan visualisasi pada Gambar 1, informasi yang dibutuhkan Kaprodi didapatkan dengan cepat dan lebih efisien.

DAFTAR PUSTAKA

- Steven Few (2004), *Selecting the Right Graph for Your Message*, diakses tanggal 23 Juni 2022 dari https://docs.google.com/viewer?docex=1&url=http://www.perceptualedge.com/articles/ie/the_right_graph.pdf
- Tara J. Brigham (2016) *Feast for the Eyes: An Introduction to Data Visualization*, *Medical Reference Services Quarterly*, 35:2, 215-223, DOI:10.1080/02763869.2016.1152146
- Vitaly Friedman (2008) "*Data Visualization and Infographics*" in: *Graphics, Monday Inspiration, January 14th, 2008*.
<https://support.google.com/docs/answer/190718?hl=en#zippy=https://support.google.com/docs/table/25273?hl=id>
<https://tekno.idsatunews.com/2020/05/25/kesalahan-umum-dalam-visualisasi-data/>
<https://www.dicoding.com/academies/177/tutorials/8132>
<https://www.abs.gov.au/websitedbs/D3310114.nsf/Home/Statistical+Language+-+what+are+data>



DATA SCIENCE

BAB 9: *CREATE DASHBOARD WITH DATA STUDIO*

Yohanssen Pratama, S.Si., M.T.

Institut Teknologi Del

BAB 9

CREATE DASHBOARD WITH DATA STUDIO

A. PENDAHULUAN

Google Data Studio adalah alat pelaporan dan visualisasi data canggih yang memungkinkan kita untuk dapat membuat dashboard menarik hanya dengan beberapa klik. Selain itu Google Data Studio sangat mudah digunakan, akan tetapi untuk mempelajarinya bisa jadi menantang apabila anda baru memulai menggunakannya. Sehingga diharapkan tutorial pada buku ini mengenai Google Data Studio dapat membantu anda untuk mampu memahaminya. Bab ini akan membawa anda untuk dapat memahami beberapa dasar-dasar Google Data Studio – yaitu cara membuat dasbor dalam waktu singkat dan cara membagikan laporan anda.

B. GOOGLE DATA STUDIO

Data Studio adalah pendatang baru dalam dunia visualisasi data dan alat untuk pelaporan. Google Data Studio dibuat oleh Google pada tahun 2016 dan sejak itu telah mendapatkan banyak pengguna, terutama di kalangan pemasaran dan data analyst.

Untuk penggunaannya Data Studio sepenuhnya gratis dan tidak ada versi berbayarnya. Kita dapat menggunakannya sebagai alternatif dari alat pelaporan lain yang berbayar seperti Tableau dan Power BI.

Selain itu untuk dapat menggunakan Data Studio tidak memerlukan pembayaran di awal, kita dapat mempelajari dan mengakses semua fiturnya secara gratis. Bahkan kita dapat menggunakannya untuk tujuan komersil (bisnis) tanpa perlu melakukan pembayaran sepeser pun.

Walaupun kita tidak akan mendapatkan fitur super canggih, tetapi hal ini tidak akan menghentikan kita untuk dapat membuat sebuah dashboard yang profesional. Selain itu, dengan Data Studio kita akan dimudahkan untuk dapat membagikan dashboard yang dimiliki dan melakukan kolaborasi dengan orang lain.

DAFTAR PUSTAKA

G. Kemp, G. White, “ Google Data Studio for Beginner,” Apress, 2021
blog.coupler.io [diakses 17 Juli 2022]



DATA SCIENCE

BAB 10: *INTRODUCTION MACHINE LEARNING*

Hidayatul Fitri, S.T., M.Eng

STAIN Teungku Dirundeng Meulaboh, Aceh Barat

BAB 10

INTRODUCTION MACHINE LEARNING

A. PENDAHULUAN

Memasuki era big data, menjadi tantangan sendiri bagi manusia. Dimana, manusia akan hidup secara berdampingan dengan mesin dan beberapa aktivitas membutuhkan mesin. Perkembangan data setiap harinya semakin bertumbuh, meningkat tajam tanpa mengenal waktu. Contohnya, dalam setiap menit banyak berita yang di publish pada website, foto yang di upload pada sosial media, video yang di unggah pada youtube, kutipan yang di share pada twitter dan masih banyak informasi-informasi lainnya. Sehingga informasi ini menjadi data yang tersimpan dan memiliki manfaat. Pemanfaatan data pun semakin bijak dilakukan oleh para ilmuwan dan para peneliti.

Machine learning merupakan bagian dari *Artificial Intelligence* (AI) atau kecerdasan buatan. Proses kerja dari machine learning yaitu mempelajari perilaku atau pola yang diajarkan, machine learning mengikuti cara belajarnya manusia dengan mempelajari dari masalah sebelumnya dan mengikuti contoh yang ada untuk menjawab persoalan-persoalan yang akan datang. Perbedaan mesin dan manusia, sebagaimana manusia belajar dari pengalaman-pengalaman yang pernah terjadi dan di alaminya, seperti belajar bagaimana dia bersepeda, belajar bagaimana cara berjalan, dan belajar menggunakan smartphone. Sedangkan mesin mengikuti instruksi yang diberikan untuk melakukan sesuatu oleh programmer. Maka muncullah pertanyaan baru, bagaimana mesin bisa melakukan sesuatu seperti manusia, dengan belajar dari pengalaman tidak hanya mengikuti instruksi yang diberikan saja. Oleh karena itu, hadirilah bidang ilmu *machine learning* untuk menjawab pertanyaan tersebut, dimana mesin dapat belajar dari pengalaman seperti manusia tepatnya belajar dari data. Sistem pada *machine learning* akan lebih pintar dari waktu ke waktu tanpa

DAFTAR PUSTAKA

- Arulkumar, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). *Deep Reinforcement Learning: A brief survey. IEEE Signal Processing Magazine*, 34(6), 26–38. <https://doi.org/10.1109/MSP.2017.2743240>
- C A Padmanabha Reddy, Y., Viswanath, P., & Eswara Reddy, B. (2018). Semi-supervised learning: a brief review. *International Journal of Engineering & Technology*, 7(1.8), 81. <https://doi.org/10.14419/ijet.v7i1.8.9977>
- Daqiqil Id, I. (2021). *Machine Learning: Teori, Studi Kasus dan Implementasi Menggunakan Python*. Riau: UR PRESS.
- Ghahramani, Z. (2004). Unsupervised Learning. In *Machine Learning* (pp. 72–112). https://doi.org/10.1007/978-3-540-28650-9_5
- Iriondo, R., & Shukla, P. (2022). Learn what is machine Learning (ML), how it works, and its importance in five minutes. Retrieved 18 June 2022, from Towards AI website: [https://towardsai.net/p/machine-learning/what-is-machine-learning-ml-b58162f97ec7#:~:text=Mitchell%5B2%5D,through experience%5B2%5D."](https://towardsai.net/p/machine-learning/what-is-machine-learning-ml-b58162f97ec7#:~:text=Mitchell%5B2%5D,through%5B2%5D,experience%5B2%5D.)
- Moser, G. (1990). In Memoriam, Arthur Samuel: Pioneer in Machine Learning. *AI Magazine (AAAI)*, 11(3), 10–11.
- Swamy, M. K., & Puligilla, T. (2020). Unsupervised Learning. In *Machine Learning and Big Data* (pp. 233–249). <https://doi.org/10.1002/9781119654834.ch9>



DATA SCIENCE

BAB 11: *CLASSIFICATION*

Ashari Wicaksono, S.Kel., M.Si., M.Sc

Program Studi Ilmu Kelautan, Universitas Trunojoyo Madura

BAB 11

CLASSIFICATION

A. PENDAHULUAN

Dalam bab ini akan menjelaskan pengertian dari klasifikasi dan tipe klasifikasi yang umum digunakan dalam bidang *Data Science*, dan *Machine Learning*. Bab ini akan lebih banyak menggunakan contoh pada Bahasa pemrograman *python* dengan *package scikit-learn* (<https://scikit-learn.org/>) dalam studi kasusnya.

B. MATERI

Metode dalam melakukan analisis suatu data seperti klasifikasi (*classification*) sekilas memiliki kesamaan dengan teknik klasterisasi (*clustering*), namun pada penerapannya terutama di bidang *data science* maupun *machine learning* hal tersebut merupakan metode pembelajaran (*learning*) yang berbeda. Beberapa ahli menjelaskan pengertian dari algoritma klasifikasi yang dapat dilihat pada pernyataan berikut:

1. *Classification is a supervised approach to grouping together items of interest and discriminant analysis and machine learning methods are amongst the approaches that have been used.*¹
2. *The problem of classification (also called discriminant analysis) involves classifying objects into classes when there is already information about the nature of the classes.*²
3. *Supervised learning techniques attempt to discern results and learn by trying to find in a labeled data set. Human interaction is required to label the data.*³

¹ Maharaj, E. A. et al. *Time Series Clustering and Classification*. (CRC Press. 2019). Hal. 2.

² Bouveyron, C. et al. *Model-based Clustering and Classification for Data Science with Applications in R*. (Cambridge University Press. 2019). Hal. 14.

³ Cielen, D. et al. *Introducing Data Science, Big Data, Machine Learning, and More, Using Python Tools*. (Manning Shelter Island. 2016). Hal. 65.

DAFTAR PUSTAKA

- Boschetti, A., L. Massaron. 2016. *Python Data Science Essentials Second Edition*. Birmingham: Packt Publishing.
- Bouveyron, C., G. Celeux, T. B. Murphy, and A. E. Raftery. 2019. *Model-based Clustering and Classification for Data Science with Applications in R*. Cambridge University Press.
- Cielen, D., A. D. B. Meysman, M. Ali. 2016. *Introducing Data Science, Big Data, Machine Learning, and More, Using Python Tools*. Manning Publications. New York.
- Górecki, T., M. Krzysko, and W. Wolynski. 2016. *Variable Selection for Classification of Multivariate Functional Data*. Springer.
- Maharaj, E. A., P. D'Urso, J. Caiado. 2019. *Time Series Clustering and Classification*. CRC Press.
- Mehta, D. P., and S. Sahni. 2004. *Handbook of Data Structures and Applications*. London. CRC Press.
- Moertini, V. S., M. T. Adithia. 2020. *Pengantar Data Science dan Aplikasinya bagi Pemula*. Bandung: Unpar Press.
- Ozdemir, S. 2016. *Principle of Data Science*. Birmingham. Packt Publishing.
- Ratner, B. 2011. *Statistical and Machine-Learning Data Mining Technique for Better Predictive Modeling and Analysis of Big Data Second Edition*. CRC Press.
- VanderPlas, J. 2017. *Python Data Science Handbook*. California. O'Reilly Media.
- Walpole, R. E. 1995. *Pengantar Statistika Edisi ke-3*. Jakarta. Gramedia Pustaka Utama.



DATA SCIENCE

BAB 12: *REGRESSION*

Dr. Arnita, S.Si, M.Si

Program Studi Ilmu Komputer Universitas Negeri Medan

BAB 12

REGRESSION

A. PENDAHULUAN

Analisis regresi adalah salah satu teknik yang digunakan untuk mengidentifikasi hubungan antara sepasang variabel atau lebih. Regresi digunakan untuk menemukan suatu fungsi yang memodelkan data dengan meminimalkan *error* atau selisih antara nilai prediksi dengan nilai sebenarnya. Pada *Machine Learning*, regresi termasuk pada mesin pembelajaran *supervised learning*. *Supervised learning* adalah mesin pembelajaran yang proses pembelajarannya di bawah pengawasan supervisor. Algoritma ini memerlukan label untuk membangun model dengan tingkat akurasi yang berubah-ubah dari waktu ke waktu. Semakin banyak input model yang digunakan maka keakurasiannya akan semakin tinggi.

Pada Bab ini akan dibahas tentang macam-macam regresi berdasarkan jumlah dan bentuk variabel yang terlibat dalam pemodelan. Secara garis besar regresi dapat dibedakan menjadi regresi sederhana dan regresi multiple. Regresi sederhana adalah pemodelan yang melibatkan sepasang variabel yaitu variabel bebas/*independent* dan variabel dependent. Jika dilihat dari bentuk hubungannya, analisis regresi dapat di bedakan menjadi regresi linear dan regresi non linear. Namun, jika dilihat dari keterlibatan variabel dalam membentuk model maka model, maka analisis regresi dapat dibedakan menjadi regresi sederhana dan regresi berganda.

B. REGRESI LINEAR SEDERHANA

Analisis regresi linear sederhana adalah analisis yang menjelaskan hubungan linear antara satu variabel independent (X) dengan variabel dependent (Y). Pada analisis regresi linear sederhana, variabel yang mempengaruhi disebut variabel bebas/*independent*. Sedangkan variabel yang dipengaruhi disebut variabel terikat/*dependent*. Model regresi linear

DAFTAR PUSTAKA

- Ajija, Shochrul Rohmatul, dkk. 2011. Cara Cerdas Menguasai Eviews. Jakarta: Salemba Empat.
- Ghozali, I. (2016) Aplikasi Analisis Multivariete Dengan Program IBM SPSS 23. Edisi 8. Semarang: Badan Penerbit Universitas Diponegoro.
- Hosmer, David W. and Stanley Lemeshow. 2000. Applied Logistic Regression Second Edition. New York: John Wiley & Sons, Inc
- Hasan, M.I. (1999). Pokok-pokok Statistika. Jakarta : Erlangga
- Juliandi, A, Irfan, Manurung. S. 2014. Metode Penelitian Bisnis, Konsep dan Aplikasi. UMSU press
- Montgomery, D.C., Peck, E.A., & Vining, G.G. (1992). Introduction to Linear Regression Analysis. Toronto: John Wiley & Sons.
- Sudijono, Anas. 1996. Pengantar Statistik Pendidikan. Jakarta: Rajawali
- Yuliara, I.M. 2016. Modul Perkuliahan. Jurusan Fisika Fakultas Matematika dan Pengetahuan Alam. Univeristas Udayana.

PROFIL PENULIS

Yunita Ardilla, S.Kom., M.M.T



Penulis lahir di Kota Sidoarjo pada tanggal 04 Juni 1992. Riwayat pendidikannya dimulai sejak lulus dari program sarjana Teknik Informatika pada tahun 2014 dari Institut Teknologi Sepuluh Nopember, yang kemudian dilanjutkan sampai jenjang pasca sarjana (S2) Manajemen Teknologi Informasi di Institut Teknologi Sepuluh Nopember.

Ia juga menjadi dosen tetap (*faculty member*) Fakultas Dakwah dan Komunikasi di Universitas Islam Negeri Sunan Ampel Surabaya sebagai dosen sistem informasi manajemen. Informasi lengkap dapat dilihat di web pribadi www.yunita.net |surel: yunita.ardilla@uinsby.ac.id.

Guntoro, S.T., M.Kom



Penulis lahir di Bangkinang, 12 Januari 1988. Lulus S1 di Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau tahun 2011. Lulus S2 di Program Studi Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) IPB University tahun 2015. Saat ini adalah Dosen tetap di Program Studi

Teknik Informatika Fakultas Ilmu Komputer, Universitas Lancang Kuning. Mengampu mata kuliah Forensika Digital, Sistem Operasi, Teknik Kompilasi, Metode Numerik dan Manajemen Proyek Perangkat Lunak. Saat ini di amanahkan sebagai Kepala Bagian Penelitian LPPM Universitas Lancang Kuning. Sebagai Editor in Chief Jurnal Dinamisia: Jurnal Pengabdian Kepada Masyarakat, Jurnal JITACS: *Journal of Information Technology and Computer Science* Universitas Lancang Kuning. Pernah menulis buku Pengantar Teknologi Informasi, Sistem Informasi Manajemen, Data Mining dan Aplikasinya, buku Praktis Desain Dokumen dan Informasi dengan Canva Android: Membuat Dokumen dan Informasi Berbasis Digital dan Cetak. Ia bisa ditemui di www.guntoro.web.id

Ir. Surya Afnarius, M.Sc, Ph.D



Penulis lahir 9 April 1964 di Kota Bukittinggi, Sumatera Barat. SD, SMP dan SMA ditamatkan di Kota Padang. Pada 1983 terdaftar sebagai mahasiswa ITB. Lulus S-1 di Program Studi Teknik Informatika ITB pada Maret 1988 di Bandung. Lulus S-2 Magister Sains Geoinformatika Universiti Teknologi Malaysia tahun 1998 di Johor Bahru, Malaysia. Lulus S-3 Doktor Falsafah Geoinformatika Universiti Teknologi Malaysia tahun 2003 di Johor Bahru, Malaysia. Selepas lulus dari ITB, bekerja beberapa tahun di PT Semen Padang dan Daiwabo Information System, Jepang. Kemudian menjadi Dosen tetap di Teknik Elektro Unand pada Desember 1995. Mulai Agustus 2010 hingga kini menjadi dosen tetap di Prodi Sistem Informasi Fakultas Teknologi Informasi Universitas Andalas. Sehari-hari mengasuh perkuliahan tentang 'Data', yaitu: Manajemen database, Perancangan database, Analisis Data Spasial, Data warehouse dan Basis data lanjut.

Agustinus Budi Santoso, S.S.T, M.Cs



Penulis bertempat lahir pada Semarang 03 September 1990. Penulis merupakan anak ke empat dari empat bersaudara. Memiliki riwayat sekolah pada SMA Sint Louis Semarang dan melanjutkan studi pada Universitas STEKOM di Semarang dan UKSW Salatiga. Dengan dedikasi dan motivasi yang tinggi untuk pendidikan maka saat ini menjadi salah satu tenaga pengajar di Universitas STEKOM. Penulis juga memiliki beberapa kegemaran seperti memancing dan touring bersama rekan kerja dan semasa sekolah untuk menambah wawasan dan rekan sahabat pada kehidupannya. Ia memiliki keahlian khusus dalam bidang logika dan algoritma pada Sistem Informasi yang menjadi unggulan untuk seorang analisis dan pengujian data algoritma suatu aplikasi dan sistem pada suatu perusahaan yang telah bekerjasama sejak 2015. Beberapa pengembangan algoritma sistem yang telah dipakai pada pelayanan masyarakat seperti Puskesmas, Toko Retail. Dan

pengembangan dalam bidang sistem website. Penulis juga merupakan seorang pendiri pengembang usaha di bidang IT yang mengembangkan bersama tim IT bernama SantosoWeb.com. Bersama tim produksi membuat produksi dalam bidang Jasa pembuatan software website, Android, dan web hosting. Sampai sekarang masih aktif dalam bidang pengembangan teknologi yang bertujuan memajukan teknologi dan perekonomian segala sektor baik perdagangan, layanan jasa dan pendidikan.

Rezania Agramanisti Azdy, S.Kom., M.Cs



Penulis kelahiran Yogyakarta 15 November 1986 ini merupakan seorang dosen yang ditugaskan pada ITBIS PalComTech (Program Studi S1 Informatika). Pendidikan yang telah ditempuh adalah pendidikan jenjang S1 pada jurusan Ilmu Komputer Universitas Gajah Mada, kemudian dilanjutkan ke jenjang S2 pada Universitas dan Jurusan yang sama.

Rona Putra, M.Kom



Penulis merupakan Dosen di Institut Agama Islam Negeri Curup, Bengkulu program studi Ilmu Perpustakaan dan Informasi Islam (IPII), menempuh pendidikan S2 Sistem Informasi pada Universitas Dinamika Bangsa Jambi. Sebelum terjun sebagai akademisi, pernah bekerja sebagai Programmer dan IT Support di RSUD Raden Mattaher Jambi, Klinik Mata Kambang Jambi, Ditlantas Polda Jambi, PT. Andalan Putra Jambi dan Jambi Media Teknologi. Pengalaman dalam mengelola basis data rumah sakit yang menghimpun data transaksi yang cukup besar dan kebutuhan akan laporan transaksi yang bervariasi dan beberapa proyek sistem informasi lainnya menjadi modal utama untuk memulai mendalami bidang Data Science. Meskipun sekarang tidak berada dalam naungan keilmuan bidang komputer namun masih sangat

relevan dan bersinggungan terutama dari sisi Ilmu Informasi. Bidang Ilmu Perustakaan dan Informasi masih mempelajari Pangkalan Data (Basis data), Pengindeksan, Pengelolaan Perpustakaan Digital dan Digitalisasi Arsip yang yang memanfaatkan teknik-teknik yang terdapat pada *data science*.

Nur Fitrianiingsih Hasan, M.Kom



Penulis lahir di Papua tahun 1994. Menyelesaikan pendidikan S1 teknik informatika konsentrasi pemrograman web dan S2 teknik informatika konsentrasi *Business Intelligence* di Universitas Amikom Yogyakarta. Saat ini merupakan dosen tetap di Prodi Ilmu Komputer Universitas Muhammadiyah Papua (UM Papua) dan sebagai kepala laboratorium UM Papua. Penulis tertarik pada topik riset *artificial intelligence & data science*. Selain aktif sebagai akademisi, penulis juga merupakan praktisi *technopreneurship* di SAGOO TECH Solutions sebuah software house & IT konsultan. Penulis dapat ditemui pada laman pribadinya di fitriahasan.sagootech.com

Eka Prasetya Adhy Sugara, S.T., M.Kom



Penulis saat ini berprofesi sebagai Dosen Tetap di Institut Teknologi dan Bisnis PalComTech Program Studi D3 Desain Komunikasi Visual. Mata kuliah yang diampu antara lain adalah Animasi 2 Dimensi, Animasi 3 Dimensi dan Desain Multimedia. Penulis memiliki ketertarikan di bidang Desain Grafis dan Aplikasi Multimedia serta menguasai aplikasi *Adobe Photoshop*, *Adobe Illustrator*, *Adobe After Effect*, *Adobe Premiere*, *Adobe Audition*, *Blender* dan *Corel Draw*. Buku yang telah diterbitkan penulis antara lain *Penganggaran untuk Taman Pendidikan Al-Qur'an* (2019) dan *Cara Praktis Membangun Aplikasi Mobile Studi Kasus: M-Commerce Sale and Ordering* (2021).

Yohanssen Pratama, S.Si., M.T.



Penulis mendapatkan gelar Sarjana Sains di bidang Fisika dari Institut Teknologi Bandung pada tahun 2011. Lalu kembali melanjutkan studi di kampus yang sama dan selesai pada 2015 dengan gelar Magister Teknik di bidang Informatika. Mulai tahun 2022 melanjutkan studi Doktoralnya di Nara Institut of Science and Technology, Jepang. Sejak 2015 sampai dengan sekarang, ia mendidik dan mempersiapkan calon penerus bangsa di Program Studi DIV Teknologi Rekayasa Perangkat Lunak, Fakultas Informatika dan Teknik Elektro, Institut Teknologi Del. Ia mendedikasikan dirinya untuk mempersiapkan generasi yang mahir di bidang teknologi rekayasa perangkat lunak dari daerah Toba, Sumatera Utara. Sehingga diharapkan kedepannya terkhusus untuk daerah Toba dapat menjadi pusat IT terkemuka di Indonesia. *Email:* yohanssen@outlook.com

Hidayatul Fitri, S.T., M.Eng



Penulis berasal dari pulau Sumatera, Aceh tepatnya di Aceh Barat, Meulaboh. Suatu daerah yang merupakan tempat asalnya pahlawan nasional Teuku Umar. Penulis menekuni bidang keilmuan Teknologi Informasi dan menyelesaikan pendidikan Strata-1 Teknik Informatika di Universitas Muhammadiyah Yogyakarta dan melanjutkan Strata-2 Magister Teknologi Informasi di Universitas Gadjah Mada (UGM). Penulis aktif dalam beberapa kegiatan dan penelitian keilmuan berkaitan dengan Teknologi Informasi. Beberapa penelitian penulis berkaitan dengan *machine learning* dan *artificial intelligent*. Saat ini penulis bekerja di sebuah kampus Islam Negeri di Aceh Barat, STAIN Teungku Dirundeng. Penulis menerima kritikan dan saran dan penulis dapat dihubungi melalui email: fitri.hidayatul64@gmail.com

Ashari Wicaksono, S.Kel., M.Si., M.Sc



Penulis merupakan Dosen di Program Studi Ilmu Kelautan, Universitas Trunojoyo Madura. Pendidikan S1 diselesaikan di Program Studi yang sama pada tahun 2010, dan menyelesaikan pendidikan S2 *Double Degree* Ilmu Kelautan dan Ilmu Kebumihan pada tahun 2017. Tahun 2015 pada saat studi magister penulis berkesempatan mempelajari Bahasa pemrograman yaitu Fortran. Pada tahun 2021 penulis berkesempatan mengikuti program pelatihan dari Kementerian Komunikasi dan Informatika untuk Dosen pada bidang *Artificial Intelligent*. Pada awal tahun 2022 penulis mengikuti sertifikasi internasional pada bidang *Artificial Intelligent*. *Book chapter* ini merupakan pengalaman pertama yang penulis dapatkan, dimana sebelumnya penulis telah menerbitkan buku ber-ISBN dengan tema WebGIS. Saat ini penulis terus mendalami kegiatan dalam data science dengan Bahasa pemrograman lainnya dalam mendukung kompetensi sebagai seorang pendidik.

Dr. Arnita, S.Si, M.Si



Penulis lahir di Pangkalan Berandan, 21 Juni 1976. Memperoleh gelar sarjana dari Jurusan Matematika Universitas Sumatera Utara pada Tahun 1999. Memperoleh Gelar Master dari Program Studi Statistika Institut Pertanian Bogor pada tahun 2005 dan Memperoleh gelar Doktor dari program Studi Matematika Universitas Sumatera Utara. Saat ini penulis tercatat sebagai dosen di prodi Ilmu Komputer Universitas Negeri Medan sejak Tahun 2008 hingga sekarang. Beberapa karya buku yang sudah pernah dihasilkan diantaranya adalah Pengantar Statistika tahun 2013, Teknik Peraman Data Time Series tahun 2020, Pengantar Statistika Terapan tahun 2020, Pengenalan Data Mining tahun 2021 dan Teori Graph dan Aplikasinya tahun 2021. Beberapa artikel hasil penelitian terkait data mining dan statistika sudah dihasilkan dan dipublikasi baik di jurnal serta prosiding nasional juga internasional bereputasi.

Data science adalah ilmu yang saat ini sangat populer dan dapat diterapkan di berbagai industri. Data science merupakan ilmu yang menggabungkan sebuah kemahiran di bidang ilmu tertentu dengan keahlian pemrograman, matematika, dan statistik. Tentu saja, kebutuhan akan orang-orang yang mahir di bidang ini semakin dibutuhkan oleh banyak bisnis, termasuk perusahaan-perusahaan besar. Tujuannya adalah untuk mengekstrak sebuah pengetahuan atau informasi dari data. Biasanya, orang-orang yang mahir dalam bidang data science menggunakan algoritma machine learning atau pemelajaran mesin. Hal ini berguna dalam mengolah teks, gambar, video, audio, dan lain-lain untuk menghasilkan sistem kecerdasan buatan. Sistem kecerdasan buatan ini dapat dirancang untuk melakukan berbagai tugas yang terlalu sulit untuk kecerdasan manusia. Hasil data yang diolah sistem kecerdasan buatan akan dapat dimanfaatkan oleh analis dan pengguna dalam bisnis untuk merancang strategi yang tepat untuk menyelesaikan suatu masalah atau mencapai sebuah tujuan. Jika perusahaan tidak mampu mengolah data, bisnis tidak akan memiliki pedoman untuk strategi operasi yang efektif dan efisien untuk mendapat keuntungan. Saat ini, perusahaan-perusahaan di seluruh dunia semakin menyadari pentingnya data science, kecerdasan buatan, dan machine learning. Jika sebuah bisnis ingin berkompetisi dan tetap relevan, ia harus mampu mengimplementasi data science. Kemampuan mengambil, memahami, memproses, dan menyaring nilai dari suatu data serta memvisualisasikannya adalah keahlian yang semakin penting di dekade yang akan datang. Tentu saja, orang yang memiliki pemahaman data science yang baik akan menjadi berharga dan banyak dicari.