

**DATA ANALISIS KECELAKAAN KERETA API DI INDONESIA
MENGUNAKAN METODE EXPLORATORY DATA ANALYSIS
Menggunakan Database Komite Nasional Keselamatan Transportasi
(KNKT) Tahun 2007-2022**



OLEH :

**ENDY AFRI WILDANU
09011281924076**

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2024**

**DATA ANALISIS KECELAKAAN KERETA API DI INDONESIA
MENGUNAKAN METODE EXPLORATORY DATA ANALYSIS
Menggunakan Database Komite Nasional Keselamatan Transportasi
(KNKT) Tahun 2007-2022**

SKRIPSI

**Diajukan Untuk Melengkapi Salah Satu
Syarat Memperoleh Gelar Sarjana Komputer**



**OLEH :
ENDY AFRI WILDANU
09011281924076**

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2024**

HALAMAN PENGESAHAN

SKRIPSI

**DATA ANALISIS KECELAKAAN KERETA API DI INDONESIA
MENGUNAKAN METODE EXPLORATORY DATA ANALYSIS**

Sebagai salah satu syarat untuk penyelesaian studi di
Program Studi SI Sistem Komputer

Oleh:

ENDY AFRI WILDANU

09011281924076

Pembimbing I : Dr. Rossi Passarella, M.Eng
NIP. 197806112010121004

Mengetahui

Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M.T.
196612032006041001

HALAMAN PERSETUJUAN

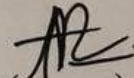
Telah diuji dan lulus pada :

Hari :

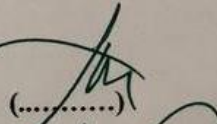
Tanggal :

Tim Penguji :


1. Ketua : Aditya Putra Perdana Prasetyo, S.kom., M.T.


(.....)

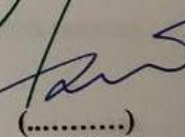
2. Sekretaris : Abdurahman, S.Kom., M.Han.


(.....)

3. Penguji : Huda Ubaya, M.T.


(.....)

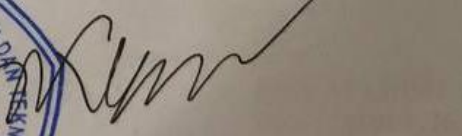
4. Pembimbing : Dr. Rossi Passarella, M.Eng.


(.....)

Mengetahui, 6/3/15

Ketua Jurusan Sistem Komputer




Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Endy Afri Wildanu
NIM : 09011281924076
Judul : Data Analisis Kecelakaan Kereta Api Di
Indonesia Menggunakan Metode Exploratory
Data Analysis

Hasil Pengecekan Software iThenticate/Turnitin : 10 %

Menyatakan bahwa laporan tugas akhir Saya merupakan hasil karya sendiri dan bukan penjiplakan / plagiat. Apabila ditemukan unsur penjiplakan / plagiat dalam laporan tugas akhir ini, maka Saya bersedia menerima sanksi akademik dari Universitas Sriwijaya.

Demikian, pernyataan ini Saya buat dalam keadaan sadar dan tanpa paksaan dari siapapun.



Indralaya, Desember 2024



Endy Afri Wildanu
09011281924076

HALAMAN PERSEMBAHAN

Skripsi ini saya dedikasikan untuk kedua Orang Tua saya yang sangat saya sayangi. Terima kasih sudah menjadi Orang Tua yang terbaik buat saya, terima kasih sudah menjadi motivasi terbaik buat saya, terima kasih untuk selalu ada buat saya, terima kasih untuk tidak menyerah mengingatkan saya untuk terus berjuang, terima kasih atas semua pengorbanan yang sudah kalian lakukan untuk saya. Izinkan saya membahagiakan dan membanggakan kalian setidaknya sekali dalam hidup saya. Terima kasih juga untuk diri saya sendiri sudah bisa berjuang sampai saat ini.

**“Tidak Ada Kata Terlambat, Yang Ada Hanya
Mau atau Tidak.”**

KATA PENGANTAR

Assalammu'alaikum Warrahmatullahi Wabarakatuh

Alhamdulillahirabbil'alamin. Puji dan syukur penulis panjatkan atas kehadiran Allah SWT yang telah memberikan karunia dan rahmat-Nya, sehingga penulis dapat menyelesaikan penulisan Proposal Tugas Akhir ini yang berjudul **“DATA ANALISIS KECELAKAAN KERETA API DI INDONESIA MENGGUNAKAN METODE EXPLORATORY DATA ANALYSIS”**.

Dalam proposal ini penulis menjelaskan mengenai data korban kecelakaan kereta api yang divisualisasikan dalam bentuk grafik. Dalam penelitian ini memvisualisasikan data yang ada didasarkan pada Final Report yang ada pada *website* Komite Nasional Keselamatan Transportasi (KNKT) menggunakan metode *Exploratory Data Analysis* (EDA) ditujukan untuk memperoleh informasi dengan lebih efisien, jelas, dan efektif. Penulis berharap agar tulisan ini dapat bermanfaat bagi banyak orang yang membutuhkan.

Pada kesempatan ini penulis ingin mengucapkan terima kasih kepada beberapa pihak atas ide dan saran serta bantuannya dalam menyelesaikan penulisan Proposal Tugas Akhir ini. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

1. Allah SWT, yang telah memberikan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan penulisan Proposal Tugas Akhir ini dengan baik dan lancar.
2. Orang tua saya tercinta untuk segala do'a, motivasi dan dukungannya baik moril, materil, maupun spiritual selama ini.
3. Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. Ir. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya serta selaku Pembimbing Akademik penulis di Jurusan Sistem Komputer.
5. Bapak Dr. Rossi Passarella, S.T. M.Eng., selaku Dosen Pembimbing Tugas Akhir yang telah berkenan meluangkan waktunya guna membimbing,
6. Memberikan saran dan motivasi serta bimbingan terbaik untuk penulis

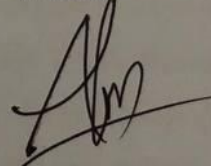
dalam menyelesaikan Tugas Akhir ini.

7. Admin Jurusan Sistem Komputer yang telah membantu mengurus seluruh pemberkasan.
8. Teman-teman yang sudah memberikan semangat dan motivasi kepada penulis dalam menyelesaikan Tugas Akhir ini.
9. Diri sendiri yang sudah mampu dan menyanggupi hingga akhir perkuliahan.
10. Dan semua pihak yang telah membantu.

Penulis menyadari bahwa laporan ini masih sangat jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangatlah diharapkan penulis agar dapat segera memperbaiki, sehingga laporan ini dapat dijadikan sebagai masukan ide dan pemikiran yang bermanfaat bagi semua pihak dan menjadi tambahan bahan bacaan bagi yang tertarik dalam visualisasi data. Akhir kata penulis berharap, semoga Proposal Tugas Akhir ini bermanfaat dan berguna bagi khalayak.

Wassalammu'alaikum Warahmatullahi Wabarakatuh

Indralaya, Desember 2024
Penulis



Endy Afri Wildanu
NIM. 09011281924076

DAFTAR ISI

	Halaman
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1. Pendahuluan.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Sistematika Penulisan.....	3
1.6.1. Bagian Awal Skripsi.....	4
1.6.2. Bagian Isi Skripsi.....	4
1.6.3. Bagian Akhir Skripsi.....	4
BAB II TINJAUAN PUSTAKA	5
2.1. Penelitian Terkait.....	5
2.2. Landasan Teori.....	6
2.2.1. Kecelakaan.....	6
2.2.2. Penyebab Kecelakaan Kereta Api.....	7
2.2.3. Karakteristik Kecelakaan Kerta Api.....	8
2.2.4. Website KNKT.....	10
2.2.5. Korelasi Antar Variabel.....	10
2.2.6. Data Preprocessing.....	10
2.2.7. Descriptive Statistics.....	14
2.2.8. EDA.....	21
BAB III METODE PENELITIAN	22
3.1. Pendahuluan.....	22

3.2. Kerangka Kerja.....	22
3.2.1. Studi Pustaka.....	23
3.2.2 Pengumpulan data.....	24
3.2.3. Data Preprocessing.....	24
3.2.4. Descriptive Statistics.....	26
3.2.5. Exploratory Data Analysis (EDA).....	30
3.2.6. Korelasi Antar Variabel.....	31
3.2.7. Kesimpulan.....	33
BAB IV HASIL DAN PEMBAHASAN.....	31
4.1. Pendahuluan.....	31
4.2. Data Preprocessing.....	31
4.3. Descriptive Statistics.....	31
4.3.1. Variabel Kuantitatif.....	32
4.3.2. Variabel Kualitatif.....	37
4.4. Exploratory Data Analysis (EDA).....	40
4.4.1. Variabel Tahun Berdasarkan Provinsi.....	40
4.4.2. Variabel Tahun Berdasarkan Penyebab Kecelakaan.....	44
4.4.3. Variabel Provinsi Berdasarkan Penyebab Kecelakaan.....	47
4.4.3. Variabel Provinsi Berdasarkan Penyebab Kecelakaan.....	47
4.5. Kesimpulan Hasil Visualisasi EBA.....	48
BAB V KESIMPULAN DAN SARAN.....	49
5.1. Kesimpulan.....	49
5.2. Saran.....	49

DAFTAR GAMBAR

	Halaman
Gambar 2.1. Contoh distribusi simetris dan asimetris.....	17
Gambar 2.2. Posisi relatif dalam distribusi frekuensi asimetris.....	18
Gambar 2.3. Dua distribusifrekuensi dengan varians yang berbeda.....	20
Gambar 3.1. Hasil <i>data preprocessing</i>	27
Gambar 3.2. Hasil <i>descriptive statistics</i>	28
Gambar 4.1. Frekuensi variabel tahun.....	34
Gambar 4.2. Visualisasi distribusi data menggunakan pengukuran pemusatan data	36
Gambar 4.3. Grafik persebaran data parameter <i>IQR</i>	37
Gambar 4.4. Visualisasi Provinsi.....	38
Gambar 4.5. Visualisasi penyebab kecelakaan.....	39
Gambar 4.6. Visualisasi penyebab kecelakaan dalam bentuk grafik lingkaran.....	40
Gambar 4.7. Visualisasi Provinsi Banten.....	41
Gambar 4.8. Visualisasi Provinsi D.I. Yogyakarta.....	41
Gambar 4.9. Visualisasi Provinsi DKI Jakarta.....	42
Gambar 4.10. Visualisasi Provinsi Jawa Barat.....	42
Gambar 4.11. Visualisasi Provinsi Jawa Timur.....	42
Gambar 4.12. Visualisasi Provinsi Jawa Tengah.....	43
Gambar 4.13. Visualisasi Provinsi Lampung.....	43
Gambar 4.14. Visualisasi Provinsi Sumatera Barat.....	43
Gambar 4.15. Visualisasi Provinsi Sumatera Utara.....	44
Gambar 4.16. Visualisasi Provinsi Sumatera Selatan.....	44
Gambar 4.17. Visualisasi Provinsi Tangerang.....	44
Gambar 4.18. Visualisasi Penyebab Kecelakaan Anjlok.....	45
Gambar 4.19. Visualisasi Penyebab Kecelakaan Tabrakan.....	46
Gambar 4.20. Visualisasi Penyebab Kecelakaan Gangguan Operasional.....	46
Gambar 4.21. Visualisasi Penyebab Kecelakaan Terbakar.....	46
Gambar 4.22. Visualisasi Penyebab Kecelakaan Patah As.....	47
Gambar 4.23. Visualisasi Penyebab Kecelakaan Lain-lain.....	47

	Halaman
Gambar 4.24. Visualisasi variabel provinsi berdasarkan penyebab kecelakaan.....	48

DAFTAR TABEL

	Halaman
Tabel 4.1. Hasil perhitungan parameter pengukuran pemusatan data dengan variabel tahun	35
Tabel 4.2. Hasil perhitungan parameter penyebaran data.....	36

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. <i>Import Library</i> yang digunakan.....	53
Lampiran 2. <i>Load dataset</i>	53
Lampiran 3. Cek informasi dataset.....	53
Lampiran 4. Cek gambaran dataset.....	53
Lampiran 5. Pemilihan operator yang digunakan.....	53
Lampiran 6. Menghapus variabel yang tidak digunakan.....	53
Lampiran 7. Menghapus nilai yang kosong.....	53
Lampiran 8. <i>Measures of frequency</i> dari variabel tahun.....	54
Lampiran 9. Membuat plot dari frekuensi tahun.....	54
Lampiran 10. <i>Measures of central tendency</i> pada variabel tahun.....	54
Lampiran 11. Plot variabel tahun menggunakan <i>maasures of central tendency</i> .	54
Lampiran 12. Plot diagram batang dari variabel operator.....	54
Lampiran 13. Plot diagram batang dari variabel faktor penyebab.....	55
Lampiran 14. Plot dari variabel faktor penyebab.....	55

BAB I

PENDAHULUAN

1.1. Pendahuluan

Transportasi kereta api merupakan salah satu jenis transportasi yang memiliki peran penting sebagai alat transportasi penduduk dan juga distribusi barang di Indonesia. Kereta api menjadi pilihan yang lebih efisien dan efektif dalam mengatasi masalah lalu lintas serta mempermudah transportasi antar kota dan pulau. Namun, sangat disayangkan bahwa kecelakaan kereta api masih terjadi dengan jumlah kecelakaan yang cukup tinggi di Indonesia. Kecelakaan kereta api tidak hanya menyebabkan korban jiwa dan luka-luka bagi penumpang dan kru, tetapi juga mengakibatkan kerusakan pada infrastruktur rel serta kendaraan kereta api itu sendiri. Oleh karena itu, sangat penting untuk melakukan analisis yang lebih dalam terhadap data kecelakaan kereta api guna mengidentifikasi serta menemukan faktor-faktor penyebabnya dan mengembangkan langkah-langkah pencegahan yang efektif untuk mengurangi kecelakaan yang terjadi.

Metode *Exploratory Data Analysis (EDA)* telah terbukti menjadi salah satu metode yang efektif dalam menganalisis data. *EDA* merupakan pendekatan statistik yang digunakan untuk menjelajahi dan menganalisis data dengan tujuan memahami karakteristik, pola, dan hubungan antar variabel dalam dataset. Dalam konteks kecelakaan kereta api, *EDA* dapat membantu peneliti untuk mendapatkan informasi lebih detail dan mudah, serta menggali informasi berguna yang mungkin tidak terlihat secara langsung.

Beberapa penelitian sebelumnya telah menggunakan metode *EDA* untuk menganalisis data pada penelitian yang mereka lakukan. Sebagai contoh, Feliks dan Loudry (2018) melakukan analisis terhadap kepadatan penumpang kereta rel listrik dan menemukan adanya hubungan yang kuat dan sangat kuat mengenai jadwal keberangkatan kereta dengan kepadatan penumpang pada setiap kereta KRL[1]. Hasil analisis *EDA* yang mereka lakukan membantu mengidentifikasi pola dan hubungan yang mungkin tidak terlihat secara kasat mata.

Terdapat juga studi lain yang dilakukan oleh Meita Aryani (2022) juga menggunakan metode *EDA* dalam menganalisis data kecelakaan penerbangan di

Indonesia yang menyatakan bahwa penggunaan metode *Exploratory Data Analysis* cukup efektif digunakan untuk mendeskripsikan serta melakukan analisis dari setiap variabel yang digunakan, baik variabel kuantitatif maupun variabel kealitatif[2]. Melalui analisis *EDA*, peneliti dapat mengungkapkan pola-pola menarik dalam dataset kecelakaan kereta api.

Selain itu Radhi *et al.*, (2021) juga menggunakan metode *EDA* untuk melakukan analisis big data. Studi mereka mengatakan bahwa metode *Exploratory Data Analysis* merupakan metode yang mampu untuk membantu proses analisis mereka[3]. Dengan menerapkan metode *EDA*, peneliti dapat memperoleh pemahaman yang lebih dalam tentang faktor-faktor yang berpengaruh terhadap kecelakaan kereta api.

Heri dan Linda (2019) melakukan penelitian mengenai analisis pelaksanaan keselamatan kerja dalam upaya meminimalkan kecelakaan kerja pada pegawai PT. Kereta Api Indonesia. Studi mereka membantu mengidentifikasi faktor-faktor yang mempengaruhi terjadinya kecelakaan dan memberikan rekomendasi untuk meningkatkan keselamatan dalam operasi kereta api [4]. Dengan penambahan metode *EDA* pada penelitian kali ini diharapkan dari penelitian tersebut dapat memberikan wawasan yang berharga dari data kecelakaan kereta api.

Berdasarkan penelitian-penelitian sebelumnya, dapat disimpulkan bahwa metode *EDA* merupakan pendekatan yang efektif dalam menganalisis data kecelakaan kereta api di Indonesia. Metode ini mampu membantu mengidentifikasi faktor-faktor penyebab, mengungkapkan pola, dan memahami hubungan antara variabel dalam dataset kecelakaan kereta api. Oleh karena itu, penelitian ini bertujuan untuk melakukan analisis data kecelakaan kereta api di Indonesia menggunakan metode *EDA*, dengan fokus pada identifikasi faktor-faktor penyebab kecelakaan serta pengembangan rekomendasi pencegahan yang lebih efektif.

1.2. Rumusan Masalah

Adapun rumusan masalah yang digunakan pada penelitian kali ini adalah sebagai berikut :

1. Apa saja faktor-faktor penyebab kecelakaan kereta api di Indonesia ?

2. Di Tahun berapakah sering terjadi kecelakaan kereta api di Indonesia ?
3. Di provinsi manakah sering terjadi kecelakaan kereta api ?

1.3. Batasan Masalah

Adapun batasan masalah yang dimiliki dalam penulisan penelitian kali ini adalah sebagai berikut :

1. Penelitian ini akan membatasi analisis pada data kecelakaan kereta api yang terjadi di wilayah Indonesia.
2. Penelitian ini akan menggunakan metode *Exploratory Data Analysis (EDA)* sebagai pendekatan utama untuk menganalisis data kecelakaan kereta api di Indonesia.
3. Penelitian ini akan berfokus pada penyebab, provinsi, tahun tertinggi pada kecelakaan kereta api di Indonesia .

1.4. Tujuan Penelitian

Adapun tujuan penelitian kali ini adalah sebagai berikut :

1. Mengidentifikasi faktor-faktor penyebab kecelakaan kereta api di Indonesia.
2. Mengungkap di tahun berapakah Indonesia mengalami kecelakaan kereta api terbanyak.
3. Mengetahui provinsi manakah yang sering terjadi kecelakaan kereta api.

1.5. Manfaat Penelitian

Adapun manfaat penelitian kali ini adalah sebagai berikut :

1. Peningkatan pemahaman tentang faktor-faktor penyebab kecelakaan kereta api di Indonesia.
2. Pengungkapan pola dan hubungan dalam dataset kecelakaan kereta api.
3. Pengembangan rekomendasi pencegahan yang lebih efektif.

1.6. Sistematika Penulisan

Sistematika penulisan berguna untuk memudahkan dalam memahami jalan pemikiran secara keseluruhan skripsi. Penulisan skripsi ini secara garis besar

dibagi menjadi tiga bagian, yaitu sebagai berikut :

1.6.1. Bagian Awal Skripsi

Bagian awal skripsi terdiri dari halaman judul, halaman pengesahan, halaman pernyataan, halaman motto dan halaman persembahan, abstrak, kata pengantar, daftar isi, daftar gambar, daftar tabel dan daftar lampiran.

1.6.2. Bagian Isi Skripsi

Bagian isi skripsi terdiri dari lima bab, yaitu sebagai berikut.

1. **BAB 1: PENDAHULUAN**

Bab ini berisi tentang latar belakang, rumusan masalah, batasan masalah tujuan dan manfaat penelitian serta sistematika penulisan skripsi.

2. **BAB 2: TINJAUAN PUSTAKA**

Bab ini berisi penjelasan yang berkaitan dengan penelitian terdahulu. Serta menjelaskan pemanfaatan *metode Exploratory Data Analysis (EDA)* dalam penelitian ini.

3. **BAB 3: METODE PENELITIAN**

Bab ini berisi penjelasan mengenai studi pendahuluan, tahap pemilihan dataset, perancangan algoritma, pengembangan sistem, analisis hasil dan penarikan kesimpulan.

4. **BAB 4: HASIL DAN PEMBAHASAN**

Bab ini berisi hasil penelitian berserta pembahasannya.

5. **BAB 5: PENUTUP**

Bab ini berisi simpulan dari penulisan skripsi dan saran yang diberikan penulis untuk mengembangkan skripsi ini.

1.6.3. Bagian Akhir Skripsi

Bagian akhir skripsi ini berisi daftar pustaka yang merupakan informasi mengenai buku-buku, sumber-sumber dan referensi yang digunakan penulis serta lampiran - lampiran yang mendukung dalam penulisan skripsi ini.

DAFTAR PUSTAKA

- [1] F. V. P. Samosir, L. P. Mustamu, E. D. Anggara, A. I. Wiyogo, and A. Widjaja, “*Exploratory Data Analysis* terhadap Kepadatan Penumpang Kereta Rel Listrik,” *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 2, pp. 449–467, 2021, doi: 10.28932/jutisi.v7i2.3700.
- [2] Aryani, Meita. (2022). *Implementasi Exploratory Data Analysis (Eda) Dalam Analisis Kecelakaan Penerbangan Di Indonesia Berdasarkan Tipe, Operator Dan Penyebab Kecelakaan.* (Skripsi Sarjana, Universitas Sriwijaya).
- [3] M. Radhi, A. Amalia, D. R. H. Sitompul, S. H. Sinurat, and E. Indra, “Analisis Big Data Dengan Metode *Exploratory Data Analysis (Eda)* Dan Metode Visualisasi Menggunakan *Jupyter Notebook*,” *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 4, no. 2, pp. 23–27, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2475.
- [4] H. Nugraha, “Analisis Pelaksanaan Program Keselamatan Dan Kesehatan Kerja Dalam Upaya Meminimalkan Kecelakaan Kerja Pada Pegawai Pt. Kereta Api Indonesia (Persero),” *Coopetition J. Ilm. Manaj.*, vol. 10, no. 2, pp. 93–102, 2019, doi: 10.32670/coopetition.v10i2.43.
- [5] A. Kusumaningsari, H. Suliantoro, and W. Budiawan, “Pengaruh Distraksi dan Kelelahan Kerja terhadap Tingkat Kewaspadaan Masinis dan Asisten Masinis Kereta Api Kaligung Mas (Studi Kasus di PT. KAI DAOP IV Semarang),” *Ind. Eng. Online J.*, vol. 3, no. 4, 2014.
- [6] H. Iridiastadi and E. Izazaya, “Kajian Taksonomi Kecelakaan Kereta Api di Indonesia Menggunakan *Human Factors Analysis and Classification System (HFACS)*,” *Pros. InSINas*, p. 51, 2012.
- [7] G. S. B. M. Mina Yumei Santi, “Karakteristik Kecelakaan Lalu Lintas Dan Lokasi *Black Spot* Di Kab. Cilacap,” *J. Tek. Sipil*, vol. 12, no. 4, pp. 259–266, 2016, doi: 10.24002/jts.v12i4.634.
- [8] T. Yulianto, “Analisis Penyebab Terjadinya Kecelakaan Kereta Api serta Usaha Prefentif yang Dapat Dilakukan,” 2010.
- [9] M. Kamber and J. Han, *Data Mining: Concepts and Techniques : Concepts and Techniques*. 2018.
- [10] D. Barapatre and A. Vijayalakshmi, “*Data preparation on large datasets for data science*,” *Asian J. Pharm. Clin. Res.*, vol. 10, pp. 485–488, 2017, doi: 10.22159/ajpcr.2017.v10s1.20526.

- [11] S. W. Chan *et al.*, “Jurnal Teknologi *Assessing Statistical Reasoning In Descriptive Statistics : A Qualitative Meta-*,” vol. 5, pp. 29–35, 2016.
- [12] S. Pérez-Vicente and M. Expósito Ruiz, “*Descriptive statistics*,” *Allergol. Immunopathol. (Madr)*., vol. 37, no. 6, pp. 314–320, 2009, doi:10.1016/j.aller.2009.10.005.
- [13] E. Allaj, “*Measuring variability and association for categorical data*,” *Fuzzy Sets Syst.*, vol. 421, pp. 29–43, 2021, doi: 10.1016/j.fss.2020.11.018.
- [14] A. P. King and R. J. Eckersley, “*Descriptive Statistics I: Univariate Statistics*,” *Stat. Biomed. Eng. Sci.*, pp. 1–21, 2019, doi:10.1016/b978-0-08-102939-8.00010-4.
- [15] V. M. Guerrero and C. Solis-Lemus, “*A generalized measure of dispersion*,” *Stat. Probab. Lett.*, vol. 164, p. 108806, 2020, doi: 10.1016/j.spl.2020.108806.
- [16] S. Manikandan, “*Measures of central tendency: Median and mode*,” *J. Pharmacol. Pharmacother.*, vol. 2, no. 3, pp. 214–215, 2011, doi: 10.4103/0976-500X.83300.
- [17] S. Manikandan, “*Measures of dispersion*,” *J. Pharmacol. Pharmacother.*, vol. 2, no. 4, pp. 315–316, 2011, doi: 10.4103/0976-500X.85931.
- [18] D. Pyle, *Data Preparation for Data Mining. San Francisco: The Morgan Kaufmann Series in Data Management Systems*. 1999.
- [19] R. Tinsley and J. Green, “*Data preparation*,” *Mark. Res. Heal. Soc. Care*, no. March, pp. 124–137, 2003, doi: 10.1017/9781107051386.003.
- [20] A. Hayes, “*Descriptive Statistics*,” *Investopedia*, 2021..
- [21] S. Chaudhari, “*Descriptive Statistics*,” 2018. <https://towardsdatascience.com/descriptive-statistics-f2beaf7a8df>(accessed Feb. 06, 2022).
- [22] J. Lee, *Statistics, Descriptive*, Second Edi., vol. 13. Elsevier, 2020.
- [23] I. C. Education, “*Exploratory Data Analysis*,” 2020. <https://www.ibm.com/cloud/learn/exploratory-data-analysis#toc-what-is-ex-ofRUduQ6> (accessed Feb. 06, 2022).