



EVALUASI PENGGUNA *JUPYTER NOTEBOOK* PADA *PYTHON* DALAM PEMBELAJARAN *DATA SCIENCE* (STUDI KASUS : KAPAL TITANIC)

Aryanto Nur¹, Pramesti Refita Cahyani²

^{1,2}Fakultas Teknik Informatika, Universitas Binasarana Informatika

¹aryantonur@gmail.com, ²pramestirefitacahyani@gmail.com

ABSTRAK

Penelitian ini berjudul "Evaluasi Penggunaan *Jupyter Notebook* pada *Python* dalam Pembelajaran *Data Science* (Studi Kasus: Kapal Titanic)" berfokus pada tragedi tenggelamnya RMS Titanic, yang terjadi pada 15 April 1912. Kasus ini dipilih karena Titanic merupakan salah satu bencana pelayaran paling terkenal dalam sejarah, di mana 1.514 dari 2.224 penumpang kehilangan nyawa setelah kapal menabrak gunung es. Salah satu faktor penyebab tingginya angka kematian adalah kurangnya jumlah sekoci penyelamat. Penelitian ini menganalisis data untuk mengeksplorasi peluang bertahan hidup penumpang berdasarkan kelompok atau golongan tertentu, memberikan wawasan tentang faktor-faktor yang mempengaruhi keselamatan. Dengan menggunakan *Python* dan *Jupyter Notebook*, penelitian ini bertujuan untuk menganalisis data historis Titanic dan menerapkan konsep data science dalam pembelajaran. *Python*, sebagai bahasa pemrograman yang kaya pustaka, memungkinkan analisis data yang efisien, sementara *Jupyter Notebook* menyediakan antarmuka interaktif yang memudahkan visualisasi dan interpretasi hasil. Melalui analisis data ini, diharapkan dapat dihasilkan temuan yang relevan terkait pola bertahan hidup penumpang, serta pemahaman yang lebih mendalam tentang aplikasi analisis data historis untuk kejadian penting dalam sejarah. Penelitian ini tidak hanya menambah wawasan tentang tragedi Titanic, tetapi juga menunjukkan potensi penggunaan alat analisis modern dalam konteks pendidikan data science, sehingga dapat memberikan kontribusi yang signifikan dalam pengajaran dan pemahaman analisis data di era digital.

Kata kunci: *Python, Jupyter Notebook, data science.*

ABSTRACT

This research is titled "Evaluation of the Use of Jupyter Notebook on Python in Data Science Learning (Case Study: The Titanic Shipwreck)" and focuses on the tragedy of the RMS Titanic sinking, which occurred on April 15, 1912. This case was chosen because the Titanic is one of the most famous maritime disasters in history, where 1,514 out of 2,224 passengers lost their lives after the ship struck an iceberg. One of the factors contributing to the high death toll was the lack of lifeboats. This study analyzes data to explore the survival chances of passengers based on specific

Article History

Received: Oktober 2024

Reviewed: Oktober 2024

Published: Oktober 2024

Plagiarism Checker No 234

Prefix DOI :

10.8734/Kohesi.v1i2.365

Copyright : Author

Publish by : Kohesi



This work is licensed under
a [Creative Commons
Attribution-NonCommercial
4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



groups, providing insights into the factors influencing safety. Using Python and Jupyter Notebook, this research aims to analyze historical Titanic data and apply data science concepts in learning. Python, as a rich library programming language, enables efficient data analysis, while Jupyter Notebook provides an interactive interface that facilitates the visualization and interpretation of results. Through this data analysis, it is expected to produce relevant findings regarding passenger survival patterns and a deeper understanding of the application of historical data analysis for significant historical events. This research not only enhances insights into the Titanic tragedy but also demonstrates the potential of modern analytical tools in the context of data science education, contributing significantly to teaching and understanding data analysis in the digital age.

Keywords: Python, Jupyter Notebook, data science

PENDAHULUAN

RMS Titanic adalah sebuah kapal laut yang tenggelam pada tanggal 15 April 1912, setelah menabrak gunung es di Samudra Atlantik Utara pada tanggal 15 April 1912, kapal laut RMS Titanic terbalik. Dibangun oleh White Star Line, sebuah perusahaan galangan kapal Inggris, Titanic merupakan salah satu kapal penumpang termewah dan terbesar pada zamannya dan dianggap sebagai pencapaian teknik yang luar biasa pada saat itu. Dalam pelayaran perdananya, RMS Titanic secara luas dianggap “tidak dapat tenggelam.” Namun, kapal ini terbenam setelah menabrak sebuah gunung es pada 14 April 1912, sekitar pukul 23:40 (waktu kapal), saat sedang melaju dengan kecepatan 41 km/h dari Southampton, Inggris, menuju New York City. Tragedi ini berlangsung dengan sangat cepat. Meskipun beberapa sekoci berhasil diluncurkan dan operasi penyelamatan dilakukan, namun jumlah sekoci yang tersedia tidak mencukupi. Kapal ini hanya dilengkapi dengan 16 sekoci penyelamat yang, dalam keadaan darurat, tidak cukup untuk menampung semua penumpang dan kru. Hal ini menjadi faktor utama yang memperburuk situasi saat Titanic mulai tenggelam. Dari total 2.224 orang penumpang dan awak, sebanyak 1.514 orang kehilangan nyawa, menjadikannya sebagai bencana maritim paling mematikan dalam sejarah masa damai.

Titanic merupakan kapal terbesar di dunia pada pelayaran perdananya, dengan panjang 882 kaki 9 inci (269,06 meter), lebar maksimum 92 kaki 6 inci (28,19 meter), tinggi keseluruhan 104 kaki (32 meter) dari bagian bawah lunas ke bagian atas jembatan, dan berat keseluruhan 46.328 ton, Titanic merupakan kapal terbesar di dunia pada pelayaran pertamanya. Kapal ini dilengkapi dengan berbagai fasilitas mewah untuk penumpangnya, seperti restoran kelas satu yang elegan, kamar tidur mewah, kolam renang, dan lapangan tenis, sehingga menciptakan citra yang glamor di kalangan masyarakat.

Proses pembuatan Titanic bersamaan dengan kapal lain, Olympus, yang juga memiliki kesamaan dengan Titanic, menunjukkan ambisi industri perkapalan pada masa itu untuk menciptakan kapal yang lebih besar dan lebih megah. Selain reputasinya sebagai kapal yang “tidak bisa tenggelam”, ukuran dan keanggunan Titanic yang memukau menarik banyak perhatian pada pelayaran perdananya. Meskipun dilengkapi dengan fitur keselamatan canggih seperti kompartemen kedap air dan pintu kedap air, desain dan implementasi keselamatan yang ada ternyata tidak cukup untuk melindungi penumpang dari bencana yang mengerikan ini.



Kompas. (2021) menyatakan bahwa Titanic berlayar dari Southampton pada tanggal 10 April 1912. Setelah itu, kapal tersebut mendarat keesokan harinya di Queenstown dan keesokan harinya di Cherbourg. Titanic seharusnya tiba di New York pada tanggal 17 April 1912, tetapi tidak pernah sampai di sana. Bencana Titanic dianggap oleh Beesly Lawrence, (2016: 1) sebagai salah satu kejadian paling mengerikan, terutama mengingat kapal tersebut dipasarkan sebagai “tidak dapat tenggelam” dan tenggelam segera setelah pelayaran perdananya.

Dalam konteks penelitian ini, kumpulan data menjadi penekanan utama. Dataset, menurut Kamus Oxford, adalah sekelompok data yang terdiri dari potongan-potongan yang berbeda. Menurut Byjus, dataset adalah sekelompok data yang sering ditampilkan secara tabular dan menjadi subjek penelitian ilmu data. Menurut Campbell. (2018:12), data science menganalisis variabel dalam kumpulan data dengan menggunakan kombinasi alat, algoritma, dan bahasa komputer. Konsep ini selaras dengan pendapat Grus. (2019:10) yang menyatakan bahwa ilmu data mencakup keterampilan *hacking*, pengetahuan matematika dan statistik, serta pemahaman spesifik domain.

Visualisasi data merupakan aspek penting dalam proses analisis. Visualisasi data, seperti yang didefinisikan oleh Sudipa (2023), adalah tindakan mengubah data numerik menjadi representasi grafis atau visual, seperti bagan, peta, dan diagram, untuk membuat data lebih mudah dipahami dan dianalisis oleh konsumen. Menurut Moertini, Veronica S., dkk. (2020:20), tujuan akhir dari data science adalah untuk menghasilkan kesimpulan dari data setelah dilakukan pengolahan dan analisis.

Campbell. (2018:12) juga menyatakan bahwa terdapat dua faktor utama dalam ilmu data: *data science* dan *data analyst*. *Data science* menggunakan data untuk mengekstrak informasi dari kumpulan data dan menerapkan *algoritme* untuk menentukan kemungkinan terjadinya sesuatu. Sebaliknya, *data analyst* menggunakan data historis untuk menjelaskan kondisi saat ini. Dengan kata lain, *data science* berfokus pada pemecahan masalah melalui pembangunan model, sementara *data analyst* berfokus pada interpretasi data untuk pengambilan keputusan.

Data mining, yang digunakan dalam penelitian ini, berfungsi untuk memprediksi dan memberikan wawasan berdasarkan dataset, sehingga pembaca dapat mengambil keputusan dan tindakan yang tepat. Membuat deskripsi dan prediksi adalah bagian dari peran data mining dalam penelitian ini, menurut Muflikhah, Lalil, dkk. (2018). Hasil dari analisis data terbagi menjadi tiga kategori, menurut Retnowardhani, Astari, dkk. (2019): preskriptif, prediktif, dan deskriptif. Meskipun analisis deskriptif menawarkan ringkasan dari kumpulan data yang dianalisis, analisis ini tidak memberikan penjelasan atas pola yang diamati atau alasan terjadinya suatu peristiwa. Analisis preskriptif, di sisi lain, menawarkan ide untuk situasi potensial di masa depan dan menjelaskan penyebab terjadinya suatu kejadian, sedangkan analisis prediktif membuat prediksi tentang apa yang akan terjadi di masa depan.

Tujuan dari penelitian ini adalah untuk memahami elemen-elemen yang memengaruhi tingkat keselamatan penumpang selama bencana Titanic melalui investigasi menyeluruh terhadap statistik yang relevan. Dengan memanfaatkan teknik data mining dan visualisasi data, kami berharap dapat menghasilkan wawasan berharga yang tidak hanya menggambarkan tragedi Titanic tetapi juga memberikan kontribusi terhadap pengetahuan keselamatan laut di masa depan.

Dalam kegiatan penelitian ini, perangkat lunak yang digunakan untuk proses data science menggunakan *Jupyter Notebook* pada *Visual Studio Code*. *Jupyter Notebook* sebelumnya dikenal sebagai *IPython Notebook* yang berbasis bahasa pemrograman *Python*, Fisika, J., & Sriwijaya, U. (2019). *Jupyter Notebook* merupakan aplikasi web terbuka, aplikasi ini menghasilkan dokumentasi (notebook) yang menggabungkan input (kode) dan output menjadi



satu berkas. *Jupyter Notebook* juga mencakup berbagai aspek dalam analisis data dan pengembangan model. Penggunaannya meliputi: visualisasi data, persiapan data, analisis data eksplorasi, pemodelan statistik, pembelajaran mesin, dan banyak lagi.

1. Bahasa pemrograman *Python* Bahasa pemrograman *Python* digunakan dalam penelitian ini untuk menganalisis dan memvisualisasikan dataset. Guido van Rossum menciptakan bahasa pemrograman *Python* pada akhir tahun 1980-an, Kurnia, Rahmi, (2022). *Python* adalah bahasa populer yang digunakan oleh ahli bahasa untuk pekerjaan yang berkaitan dengan pemrosesan bahasa alami, pengembang web menggunakan untuk membuat situs web dinamis, dan administrator sistem menggunakan untuk tujuan lain. *Python* adalah bahasa pemrograman serba guna yang dapat digunakan untuk menangani masalah numerik. Meskipun demikian, *Python* dapat secara efektif menangani masalah numerik dan visualisasi data ketika dipasangkan dengan pustaka seperti *NumPy*, *Seaborn*, *Matplotlib*, dan *Pandas*. *Python* adalah salah satu bahasa pemrograman terbaik untuk visualisasi data.
2. Dataset adalah kumpulan data yang tersusun yang memudahkan analisis. Dataset ini dapat digunakan untuk mengembangkan model, mengantisipasi hasil, dan memberikan jawaban atas pertanyaan. Pembelajaran mesin dan ilmu data sangat bergantung pada kualitas dataset Flach, Peter, (2012). Dataset yang dipilih adalah dataset multivariat. Dataset dengan tiga atau lebih variabel disebut multivariat Byjus, (2019). Kita dapat melihat data sampel dari penumpang yang berada di kapal Titanic pada saat bencana terjadi berdasarkan dataset yang diberikan kepada kita.
3. Visualisasi Informasi Visualisasi data adalah teknik yang bertujuan untuk menggambarkan data dengan lebih jelas dalam bentuk diagram yang prosesnya adalah untuk memperoleh, memahami, dan membandingkan, menurut Galatlapang, Yanuangga, dkk. (2020: 59). Dalam konteks pemrosesan informasi dan analisis data, visualisasi data sangat penting. Memahami tren dan pola data dalam visualisasi menjadi lebih sederhana karena kecenderungan otak manusia untuk menganalisis pola dan hubungan yang terlihat dalam visualisasi. Jenis variabel dalam data harus dipertimbangkan saat membuat visualisasi data. Penyajian data kualitatif harus berbeda dengan data kuantitatif. Baik data kuantitatif maupun kualitatif dianalisis dalam penelitian ini. Grafik batang dapat digunakan untuk memvisualisasikan dua variabel data: satu variabel kualitatif dan variabel lainnya kuantitatif. Di sisi lain, penulis memvisualisasikan data menggunakan estimasi distribusi kernel dan histogram jika kedua variabel data bersifat kuantitatif.
4. Hipotesis dari penelitian ini adalah bahwa peluang seseorang untuk selamat dari tragedi Titanic dipengaruhi oleh beberapa elemen.

Dimulai dengan konteks isu-isu yang disebutkan di atas, penulis makalah ini menyajikan contoh kapal Titanic sebagai berikut:

1. Apa faktor-faktor yang berkontribusi terhadap tenggelamnya RMS Titanic, dan bagaimana data historis dapat digunakan untuk menganalisis kejadian tersebut?
2. Bagaimana data dapat diolah dan divisualisasikan menggunakan *Jupyter Notebook* untuk memberikan wawasan yang lebih dalam mengenai kejadian tenggelamnya *Titanic*?
3. Bagaimana pengaruh desain dan fitur keselamatan Titanic, seperti kompartemen kedap air dan jumlah sekoci penyelamat, terhadap tingkat keselamatan penumpang pada saat kecelakaan?
4. Bagaimana data mining dapat digunakan untuk menganalisis dan memahami factor-faktor yang berkontribusi terhadap kecelakaan Titanic?



Fokus penelitian ini adalah menerapkan teknik penelitian kuantitatif. Tujuan dari penelitian ini adalah untuk menguji variabel-variabel yang mempengaruhi peluang penumpang RMS Titanic untuk selamat. Dengan menggunakan pendekatan kuantitatif, Penelitian ini hanya akan berfokus pada analisis kuantitatif dan tidak mencakup analisis kualitatif. Data yang digunakan terbatas pada dataset yang tersedia di Kaggle, yang mungkin tidak mencakup semua variabel yang relevan. Dataset ini mencakup berbagai informasi tentang penumpang Titanic, seperti:

1. *Passengerid*: Id penumpang
2. *Survived*: Kondisi keselamatan penumpang (1=*Survived*, 0=*Not survived*)
3. *Pclass*: Kelas/ golongan tiket penumpang (1st *Upper*, 2nd *Middle*, 3rd *Lower*)
4. *Name*: Nama penumpang
5. *Sex*: Jenis kelamin penumpang (*male* dan *female*)
6. *Age*: Umur penumpang
7. *SibSp*: *Siblings* atau keterangan bahwa penumpang mempunyai saudara di kapal (1=Memiliki saudara di kapal, 0=Tidak memiliki saudara di kapal)
8. *Parch*: *Parents or children* atau keterangan bahwa penumpang mempunyai orang tua atau anak di kapal (1=Memiliki orangtua atau anak-anak di kapal, 0=Tidak memiliki orangtua atau anak-anak di kapal)
9. *Ticket*: Nomor tiket penumpang
10. *Fare*: Tarif tiket penumpang
11. *Cabin*: No kabin penumpang
12. *Embarked*: Asal penumpang (C=*Cherbourg*, Q=*Queenstown*, S=*Southampton*).

Dari rumusan masalah di atas, tujuan penelitian ini adalah sebagai berikut:

1. Mengidentifikasi dan menganalisis faktor-faktor yang berkontribusi terhadap tenggelamnya RMS Titanic, termasuk desain kapal, kondisi cuaca, dan keputusan navigasi.
2. Mengaplikasikan teknik data mining untuk menganalisis data historis terkait Titanic, dengan tujuan menemukan pola dan hubungan yang relevan yang mungkin telah diabaikan.
3. Meningkatkan kesadaran dan pendidikan tentang pentingnya keselamatan laut dan peran data dalam pengambilan keputusan yang lebih baik di industri pelayaran.
4. Mengembangkan rekomendasi untuk praktik keselamatan pelayaran di masa depan berdasarkan pelajaran yang dipetik dari tragedi Titanic dan penerapan prinsip-prinsip ilmu data.

TINJAUAN PUSTAKA

Python adalah bahasa pemrograman tingkat tinggi yang sering digunakan dalam sejumlah pengembangan perangkat lunak dan domain aplikasi.

Jupyter Notebook, sebuah IDE (*Integrated Development Environment*) interaktif, Anda dapat menulis dan mendistribusikan dokumen dengan teks naratif, kode, visual, dan hasil eksekusi kode. Jupyter sering digunakan untuk pembelajaran mesin, simulasi numerik, visualisasi data, pemrosesan data, dan eksplorasi.

Visual Studio Code adalah editor kode Microsoft yang tersedia untuk diunduh secara gratis di semua komputer desktop. Pilihan terbaik untuk pengembang adalah editor kode ini karena fitur dan add-on-nya yang luas. Hampir semua sistem operasi, termasuk Windows, Mac OS, Linux, dan lainnya, didukung oleh *Visual Studio Code*. Para developer profesional lebih memilih *Visual Studio Code* sebagai editor pilihan mereka, menurut studi Stack Overflow. Dengan persentase pengguna sebesar 71,07%, *Visual Studio Code* menduduki peringkat pertama dari 21 program editor teks yang bersaing. Untuk mengurangi kebutuhan akan perangkat



berspesifikasi tinggi, *Visual Studio Code* didesain sederhana dan semudah mungkin. Anda dapat menggunakan aplikasi ini untuk menulis atau memodifikasi kode sumber dalam sejumlah bahasa pemrograman yang berbeda. seperti *JavaScript*, *TypeScript*, *Node.js*, dan banyak lagi.

Selain itu, *Visual Studio Code* menyediakan ekosistem dan ekstensi yang sangat besar. Hal ini membuatnya sangat kompatibel dengan *Python*, *PHP*, *.NET*, *Java*, dan bahasa atau lingkungan runtime lainnya. Perangkat lunak yang disebut *Visual Studio code* digunakan untuk mengembangkan sintaks untuk aplikasi. Anda harus terbiasa dengan empat versi *Visual Studio code* berikut ini: IDE, App Center, Azure DevOps, dan *Visual Studio Code*. 1. *Visual Studio IDE*: Dengan komponen penyelesaian kode di tingkat sumber dan mesin, program ini membantu pengguna membuat program dengan cepat.

Mayoritas programmer yang menggunakan *Visual Studio code* IDE untuk proyek pemrograman yang dipublikasikan adalah pengembang tunggal, tim, atau organisasi TI. 2. Pusat Aplikasi *Visual Studio code*: Tujuan utama dari program ini adalah untuk mengotomatiskan siklus hidup program yang beroperasi pada Windows, Mac, iOS, dan Android. 3. Azure DevOps - Azure DevOps menyederhanakan proses pengembangan aplikasi untuk pelanggan dan merupakan pengembangan dari *Visual Studio Code Team Service (VSTS)*.

Banyak sumber daya tambahan tersedia untuk memaksimalkan inisiatif. Misalnya, Azure Test Plan untuk pengujian aplikasi, Azure Repos untuk manajemen file, Azure Boards untuk perencanaan, dan sebagainya. 4. *Visual Studio Code*: Dewaweb telah memberikan penjelasan yang sangat rinci tentang program ini di masa lalu. Sederhananya, *Visual Studio Code* adalah editor kode gratis dan open-source yang dapat digunakan untuk membuat aplikasi di berbagai sistem operasi.

Seaborn adalah paket visualisasi data *Python* berbasis *matplotlib*. Antarmuka tingkat tinggi untuk membuat grafik statistik yang menarik secara visual dan edukatif ditawarkan oleh paket ini.

Matplotlib adalah *library Python* yang fokus pada visualisasi data seperti membuat *plot* grafik.

Pandas adalah *library Python* yang fokus untuk proses analisis data seperti manipulasi data, persiapan data, dan pembersihan data.

Numpy (*Numerical Python*) adalah *library python* yang berfokus pada scientific computing.

Data Science Studi tentang sistem komputer dan matematika dikenal sebagai ilmu data. Kerumitan bidang keilmuan yang diteliti membuat ilmu data menjadi alat yang sangat ampuh untuk manajemen data yang sangat besar. Semua aspek manajemen data, seperti mengumpulkan, memodifikasi, dan menganalisis data menggunakan teknik pemodelan yang berbeda, termasuk dalam *data science*. Informasi dari pengumpulan data ini akan digunakan untuk mendukung pilihan, wawasan, dan rencana bisnis jangka panjang perusahaan.

Data Visualisasi adalah proses menempatkan informasi ke dalam cara yang dapat dimengerti. Lebih tepatnya, ini adalah sekumpulan data yang telah direpresentasikan secara visual. Diperkirakan bahwa menggunakan visualisasi data dapat membantu peserta menyerap informasi dengan lebih cepat. Menurut penelitian menunjukkan bahwa 80% lebih seseorang akan mencerna sesuatu dan memahami lebih cepat dari apa yang dilihat secara visual.

Sedangkan sisanya melalui tulisan yang dibaca. Intinya menjelaskan otak manusia lebih mampu mengenali gambar lebih cepat dari jenis konten yang lainnya. visualisasi data bisa menjadi metode untuk memberikan pemahaman informasi lebih mudah kepada pesertanya. Selain itu juga memiliki beberapa macam fungsi. 1. Salah satu penggunaan visualisasi data yang paling populer adalah membuat data mudah diinterpretasikan.



Hal ini menarik perhatian pada data yang penting di antara banyak hasil yang diperiksa. 2. Meningkatkan pemahaman Meningkatkan pemahaman tentang proses perusahaan adalah tujuan kedua dari visualisasi data. 3. Meningkatkan nilai barang atau jasa Nilai barang dan jasa dapat meningkat dengan penggunaan visualisasi data. Ketersediaan visualisasi data dapat memberikan kenyamanan dan kemudahan bagi pengguna. Memberikan konsumen akses ke analisis mutakhir Dasbor interaktif dengan GIS internal, prakiraan sekali klik, dan analisis aliran digunakan oleh banyak organisasi. Hasilnya, perusahaan atau organisasi dapat memperoleh pemahaman yang lebih mendalam, menemukan tren yang tidak disadari, dan memanfaatkan prospek komersial yang menguntungkan.

METODE PENELITIAN

Pendekatan kuantitatif adalah pendekatan yang digunakan dalam penelitian ini. Meiryani. (2021) mendefinisikan analisis kuantitatif sebagai kegiatan mengumpulkan dan menganalisis data yang memiliki nilai numerik dan dapat dikuantifikasikan. Berbagai teknik, termasuk analisis komparatif, regresi linier, dan analisis deskriptif, dapat digunakan dalam analisis data kuantitatif. Tujuan dari teknik analisis kuantitatif adalah menggunakan analisis statistik dan matematika untuk mendukung pertanyaan atau hipotesis penelitian. Dengan mendeskripsikan dan menginterpretasikan data non-numerik, analisis kualitatif, di sisi lain menekankan pada pemahaman yang lebih mendalam tentang suatu peristiwa.

Menurut Sugiyono. (2020: 361), populasi yang diteliti terdiri dari 2.240 penumpang yang berada di atas kapal Titanic pada saat tragedi terjadi. Kumpulan data penelitian ini terdiri dari 890 orang, atau sekitar 40% dari seluruh populasi. Dataset tersebut dipilih dan diolah melalui beberapa tahap, termasuk pengimporan pustaka-pustaka pemrograman *Python* yang membantu dalam proses pengolahan data. Sebagaimana diungkapkan oleh Will Cukierski. (2012), dataset yang sesuai dihubungkan ke dalam dataframe untuk dianalisis. Proses pembersihan data atau "*data cleaning*" dilakukan terhadap dataset ini untuk menghilangkan data-data yang kosong atau tidak berisi nilai, sesuai dengan pendapat Hidayati dan Kahirina. (2020).

Hal ini menyoroti betapa pentingnya prosedur ini karena data yang baru diekstrak sering kali memiliki sifat yang tidak terstruktur. Setelah pembersihan data, penulis melakukan analisis lebih lanjut untuk mengidentifikasi korelasi antara variabel-variabel yang termasuk dalam kumpulan data. Memilih variabel mana yang akan divisualisasikan membutuhkan pertimbangan dalam tahap ini. Selanjutnya, dengan menggunakan library Python yang diimpor termasuk *Pandas*, *NumPy*, *Matplotlib*, dan *Seaborn*, visualisasi data dilakukan. Visualisasi ini bertujuan untuk menyajikan hubungan antar variabel dengan cara yang mudah dipahami oleh pembaca. Menurut O. Wilke. (2019:2), visualisasi data yang baik tidak hanya jelas dan informatif tetapi juga estetik. Oleh karena itu, penulis memastikan bahwa visualisasi yang disajikan tidak hanya informatif tetapi juga disusun dengan baik agar tidak menimbulkan kesan "*ugly*" atau kurang menarik.

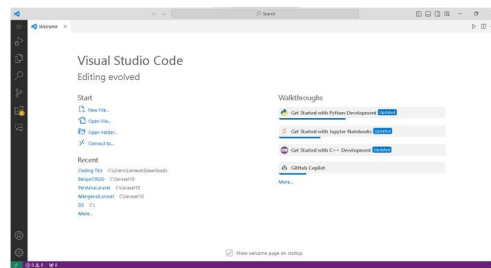
Dalam konteks penelitian ini, yaitu "Evaluasi Penggunaan *Jupyter Notebook* pada *Python* dalam Pembelajaran *Data Science* (Studi Kasus: Kapal Titanic)" penulis memfokuskan studi pada tenggelamnya RMS Titanic. Kasus ini diambil sebagai studi kasus karena tenggelamnya Titanic pada 15 April 1912 merupakan salah satu tragedi pelayaran paling terkenal dalam sejarah. Titanic, yang sedang melakukan pelayaran perdananya, tenggelam setelah menabrak gunung es, menyebabkan 1.514 dari 2.224 penumpang kehilangan nyawa. Salah satu faktor yang diduga menjadi penyebab tingginya angka kematian adalah kurangnya jumlah sekoci penyelamat. Selain itu, analisis lebih lanjut menunjukkan bahwa kelompok atau golongan tertentu memiliki peluang lebih besar untuk bertahan hidup dibandingkan dengan kelompok lainnya, sehingga

data ini memberikan peluang untuk mempelajari faktor-faktor yang mempengaruhi tingkat keselamatan penumpang. Penelitian ini juga diharapkan dapat memberikan wawasan lebih dalam tentang bagaimana analisis data historis dapat diterapkan untuk memahami kejadian-kejadian penting dalam sejarah seperti tenggelamnya Titanic.

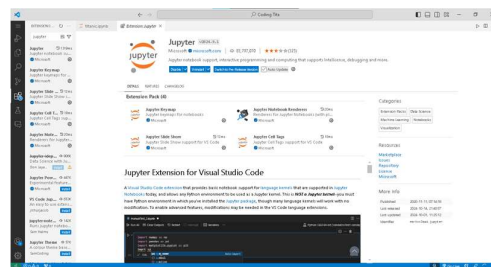
HASIL DAN PEMBAHASAN

Salah satu kapal karam paling terkenal sepanjang masa adalah Titanic, dan studi kasus ini menggunakan set data untuk memperjelas konsep ilmu data. Kapal Titanic menabrak gunung es pada tanggal 15 April 1912, dalam pelayaran pertamanya, dan tenggelam. Sekitar 1.514 dari 2.224 penumpang dan awak kapal tewas akibat kejadian ini. Kurangnya sekoci disebut-sebut sebagai salah satu alasan utama tingginya angka kematian. Ada beberapa kelompok sosial tertentu yang lebih mungkin untuk bertahan hidup daripada yang lain. Saya akan mencoba menjawab pertanyaan berikut dalam jurnal ini:

- Bagaimana data dapat diolah dan divisualisasikan menggunakan *Jupyter Notebook* untuk memberikan wawasan yang lebih dalam mengenai kejadian tenggelamnya *Titanic*?



Gambar 1.1. *Visual Studio Code (VS Code).*



Gambar 1.2. *Software Jupyter Notebook.*

“Pertama-tama dataset yang kita gunakan dapat unduh di link <https://www.kaggle.com/c/titanic/data>. Kebutuhan Software: *Python, Pandas, Numpy, Seaborn, Matplotlib.*”

1. Langkah 1 : “Import Modules and Libraries

Sebelum memulai pertama-tama kita harus mengimport *library* dan *modules* yang akan kita gunakan:”



```
titanic.ipynb X
Titanic > titanic.ipynb > import pandas as pd
+ Code + Markdown | Run All Restart Clear All Out

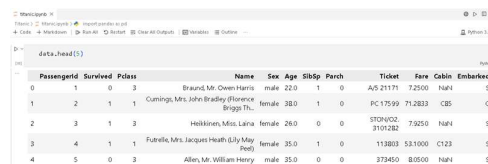
import pandas as pd
data=pd.read_csv("train.csv")

[20]
```

Gambar 1.3 Import Modules and Libraries.

2. Langkah 2 : Data Preparation

Selanjutnya kita mencoba melihat 5 data teratas dan 5 data kebawah dengan perintah ``data.tail(5)`` untuk 5 data terbawah, dan ``data.head(5)`` untuk 5 data teratas.



	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th.)	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

Gambar 1.4 Data Teratas.

Ada 890 catatan dalam dataset ini yang menawarkan rincian lengkap mengenai penumpang kapal Titanic, seperti nama, jenis kelamin, usia, jumlah orang tua, jumlah saudara kandung, nomor tiket, harga, dan nomor kabin. Kelas satu, kelas dua, dan kelas tiga adalah tiga kelas kabin penumpang yang sesuai dengan berbagai tingkat kenyamanan dan harga tiket. Kelas satu umumnya terdiri dari penumpang yang lebih kaya dan memiliki akses ke fasilitas yang lebih baik, sedangkan kelas dua dan kelas tiga diisi oleh penumpang dengan status ekonomi yang lebih rendah.

Status kehidupan penumpang dalam dataset ini dinyatakan dengan angka nol dan satu, di mana angka nol menunjukkan bahwa penumpang tersebut tidak selamat dari tragedi tersebut, sementara angka satu menandakan bahwa mereka berhasil bertahan hidup. Hal ini memberikan informasi penting untuk menganalisis faktor-faktor yang memengaruhi keselamatan penumpang selama kecelakaan Titanic.

Selain itu, tempat keberangkatan penumpang dibagi menjadi tiga lokasi: Cherbourg, Queenstown, dan Southampton, yang dilambangkan dengan huruf C, Q, dan S. Lokasi ini memberikan konteks geografis yang berharga untuk menganalisis pola keberangkatan penumpang dan bagaimana faktor lokasi mungkin berhubungan dengan kemungkinan keselamatan. Misalnya, penumpang yang berangkat dari lokasi tertentu mungkin memiliki akses yang berbeda terhadap sekoci penyelamat atau menghadapi tantangan yang berbeda dalam situasi darurat.

Secara keseluruhan, dataset ini menawarkan peluang yang luas untuk menganalisis interaksi antara berbagai variabel, seperti kelas kabin, jenis kelamin, dan lokasi keberangkatan, serta bagaimana variabel-variabel ini berkontribusi terhadap hasil keselamatan penumpang dalam tragedi yang sangat terkenal ini. Kumpulan data ini, jika dianalisis dengan benar, dapat



memberikan wawasan yang lebih dalam mengenai dinamika yang terjadi selama kejadian bencana dan membantu dalam memahami variabel-variabel yang memengaruhi peluang seseorang untuk selamat dari krisis.

```
data.head()
```

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
886	0	2	Montali, Rev. Jouda	male	27.0	0	0	211536	13.00	NaN	S
887	1	1	Graham, Miss Margaret Edm	female	19.0	0	0	112052	30.00	B42	S
888	0	3	Johansen, Miss Catherine Helen "Cassie"	female	NaN	1	2	Ward-8607	23.45	NaN	S
889	1	1	Bahr, Mr. Karl Howell	male	26.0	0	0	111369	30.00	C14B	C
890	0	3	Dooly, Mr. Patrick	male	32.0	0	0	370376	7.75	NaN	Q

Gambar 1.5 Data Terbahah.

Selanjutnya, kita akan melakukan analisis terhadap fungsi record yang tidak memiliki nilai dalam dataset. Untuk melakukan ini, kita dapat menggunakan perintah `data.isnull()` yang akan mengembalikan dataframe boolean dengan nilai True untuk setiap entri yang kosong. Ini membantu kita mengidentifikasi variabel mana yang memiliki data hilang.

Setelah itu, kita juga akan menggunakan perintah `data.isnull().sum()` untuk menghitung jumlah total record yang hilang di setiap kolom. Dengan informasi ini, kita dapat memahami sejauh mana data kita terpengaruh oleh nilai yang hilang dan mengidentifikasi langkah-langkah yang perlu diambil untuk menangani masalah ini.

Analisis terhadap record yang hilang sangat penting dalam konteks analisis data, karena data yang tidak lengkap dapat memengaruhi hasil analisis dan kesimpulan yang diambil. Dengan mengetahui kolom mana yang memiliki data yang hilang, kita dapat memutuskan apakah akan menghapus record tersebut, mengisi nilai yang hilang dengan metode tertentu, atau melakukan imputasi untuk memperbaiki dataset. Hal ini akan memastikan bahwa analisis yang dilakukan selanjutnya akan lebih akurat dan representatif dari keseluruhan populasi data.

Dengan langkah ini, kita dapat memastikan bahwa data yang digunakan dalam analisis lebih dapat diandalkan dan memberikan hasil yang valid untuk menjawab pertanyaan penelitian terkait dengan keselamatan penumpang Titanic.

```
data.isnull()
```

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	False	False	False	False	False	False	False	False	False	True	False
1	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	True	False
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	True	False
...
886	False	False	False	False	False	False	False	False	False	True	False
887	False	False	False	False	False	False	False	False	False	False	False
888	False	False	False	False	False	True	False	False	False	True	False
889	False	False	False	False	False	False	False	False	False	False	False
890	False	False	False	False	False	False	False	False	False	True	False

891 rows x 12 columns

Gambar 1.6 Fungsi record yang tidak memiliki nilai

```
data.isnull().sum()
```

PassengerId	0
Survived	0
Pclass	0
Name	0
Sex	0
Age	177
SibSp	0
Parch	0
Ticket	0
Fare	0
Cabin	687
Embarked	2
dtype:	int64

Gambar 1.7 Fungsi jumlah record yang hilang.

3. Langkah 3 : Analisis Data Visualisasi

Selanjutnya kita mencoba Menampilkan siapa saja penumpang di kapal Titanic (Berdasarkan umur, jenis kelamin dan kelas tiket).

a. Kelas Tiket



Gambar 1.8. Peluang keselamatan penumpang Titanic saat kecelakaan berbeda-beda tergantung pada kelas tiket yang mereka miliki.

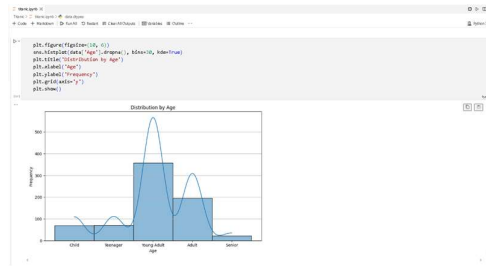
Dalam analisis Gambar 1.8, komposisi penumpang di kapal Titanic menunjukkan dinamika yang menarik di antara tiga kelas yang berbeda. Kelas satu, yang umumnya terdiri dari penumpang yang lebih kaya, memiliki proporsi besar dari orang dewasa yang mungkin memiliki akses lebih baik terhadap sekoci dan keselamatan. Penumpang kelas satu sering kali memiliki fasilitas dan ruang yang lebih nyaman, sehingga memungkinkan mereka untuk merespons situasi darurat dengan lebih cepat dan efisien. Hal ini berpotensi meningkatkan peluang keselamatan mereka selama kecelakaan.

Sementara itu, kelas dua, seperti yang telah dicatat, menampung banyak penumpang anak-anak di bawah usia 16 tahun, baik laki-laki maupun perempuan. Keberadaan anak-anak ini menunjukkan bahwa kelas dua menjadi pilihan bagi keluarga yang melakukan perjalanan. Namun, meskipun mereka memiliki kesempatan yang lebih baik dibandingkan dengan penumpang kelas tiga, perlindungan anak-anak dalam situasi darurat tetap menjadi perhatian utama, mengingat mereka membutuhkan bimbingan dan bantuan.

Di kelas tiga, analisis menunjukkan bahwa kelompok penumpang yang lebih tua lebih mendominasi. Kelas ini mungkin terdiri dari individu yang lebih rentan, baik dari segi fisik maupun sosial. Mereka yang berasal dari kelas tiga mungkin memiliki akses terbatas terhadap sekoci dan dukungan dalam situasi darurat, yang memperbesar risiko mereka selama kecelakaan. Kondisi sosial dan ekonomi yang mungkin mereka hadapi dapat membatasi kemampuan mereka untuk merespons secara efektif dalam keadaan darurat.

Secara keseluruhan, distribusi umur dan jenis kelamin di antara ketiga kelas ini memberikan wawasan penting tentang bagaimana stratifikasi sosial mempengaruhi pengalaman penumpang selama tragedi Titanic. Penumpang kelas satu lebih mungkin untuk selamat karena mereka memiliki akses yang lebih besar dan sumber daya yang lebih banyak. Penumpang yang bepergian di kelas dua dan tiga, yang memiliki ciri-ciri demografis yang berbeda, memiliki hambatan dan kesenjangan yang mencolok dalam hal keselamatan. Analisis ini menggambarkan pentingnya konteks sosial dalam menentukan hasil selama situasi krisis, serta bagaimana demografi penumpang berkontribusi pada perbedaan dalam tingkat keselamatan di antara kelompok-kelompok tersebut.

b. Umur



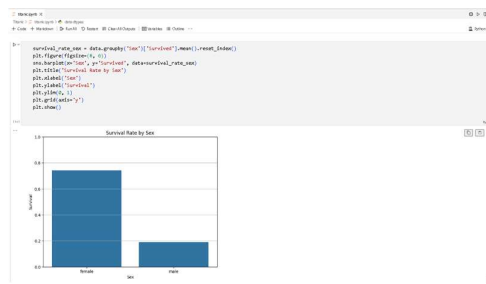
Gambar 1.9. Peluang penumpang selamat dari kecelakaan Titanic berdasarkan umur

Dalam analisis Gambar 1.9 menunjukkan peluang selamat dari kecelakaan Titanic berdasarkan kategori umur. Dalam analisis ini, penumpang yang termasuk dalam kategori anak-anak dan remaja menunjukkan angka keselamatan yang seimbang, mencerminkan bahwa kedua kelompok ini memiliki peluang yang relatif setara untuk bertahan hidup dalam situasi darurat. Sementara itu, kelompok anak muda dewasa menunjukkan peluang selamat yang jauh lebih tinggi dibandingkan kelompok lainnya. Hal ini menunjukkan bahwa anak muda memiliki kapasitas untuk merespons dengan cepat dalam keadaan darurat, sehingga meningkatkan peluang mereka untuk selamat.

Di sisi lain, kategori dewasa mencatat jumlah selamat yang menunjukkan bahwa mereka menghadapi risiko lebih besar dalam situasi ini. Meskipun mereka memiliki angka yang lebih baik dibandingkan anak-anak dan remaja, risiko yang lebih tinggi dalam keadaan darurat membuat peluang mereka lebih kecil dibandingkan dengan anak muda. Sebaliknya, kategori orang tua menunjukkan angka selamat yang sangat rendah. Ini menandakan tantangan besar yang dihadapi kelompok ini dalam situasi penyelamatan. Keterbatasan fisik dan kemampuan untuk merespons dengan cepat dapat menjadi faktor yang memengaruhi peluang keselamatan mereka.

Analisis frekuensi ini menyoroti bahwa usia dan kategori demografis sangat memengaruhi peluang keselamatan dalam tragedi Titanic. Meskipun anak-anak dan remaja mendapatkan perhatian, anak muda tetap menjadi kategori dengan peluang bertahan hidup terbaik, sementara kelompok orang tua menghadapi tantangan signifikan yang menekankan pentingnya mempertimbangkan faktor usia dalam merespons situasi darurat

c. Jenis Kelamin



Gambar 1.10. Peluang penumpang selamat dari kecelakaan Titanic berdasarkan jenis kelamin



Dalam Analisis Gambar 1.10 menampilkan peluang selamat dari kecelakaan Titanic berdasarkan jenis kelamin. Dalam analisis ini, terlihat bahwa penumpang perempuan memiliki peluang yang lebih tinggi untuk bertahan hidup dibandingkan penumpang pria. Fenomena ini dapat dipahami melalui beberapa faktor yang memengaruhi hasil keselamatan selama tragedi tersebut.

Pertama, terdapat kemungkinan bahwa prioritas penyelamatan ditujukan kepada perempuan dan anak-anak, yang merupakan praktik umum dalam situasi darurat. Hal ini berakar pada norma sosial yang mengedepankan perlindungan terhadap wanita dan anak-anak dalam situasi berbahaya. Seiring dengan upaya penyelamatan yang dilakukan oleh awak kapal, perempuan mungkin diberikan kesempatan lebih untuk naik ke sekoci penyelamat, sehingga meningkatkan peluang mereka untuk selamat. Kedua, perbedaan dalam perilaku dan respons terhadap situasi darurat juga mungkin berkontribusi pada hasil ini. Penumpang pria, yang cenderung memiliki ekspektasi untuk melindungi orang lain, mungkin lebih berfokus pada membantu anggota keluarga atau penumpang lainnya, sehingga mengorbankan kesempatan mereka sendiri untuk menyelamatkan diri. Sebaliknya, perempuan mungkin lebih terfokus pada keselamatan diri mereka dan anak-anak, yang membuat mereka lebih efektif dalam merespons situasi kritis ini.

Selanjutnya, pengaruh struktur kelas juga tidak dapat diabaikan. Dalam konteks Titanic, penumpang perempuan dari kelas atas mungkin memiliki akses yang lebih baik ke sekoci penyelamat dibandingkan penumpang pria dari kelas yang sama. Ketidaksetaraan dalam akses ini mungkin memperkuat hasil keselamatan yang lebih baik bagi perempuan secara keseluruhan.

Analisis ini menunjukkan bahwa jenis kelamin berperan penting dalam menentukan peluang selamat selama tragedi Titanic. Dengan mempertimbangkan norma sosial, perilaku individu, dan faktor kelas, kita dapat memahami mengapa perempuan memiliki peluang lebih besar untuk bertahan hidup dibandingkan pria. Hasil ini menyoroti kompleksitas situasi penyelamatan dan bagaimana berbagai faktor dapat saling berinteraksi dalam menentukan hasil yang berbeda bagi kelompok demografis yang berbeda.

KESIMPULAN

Analisis mengenai tenggelamnya Kapal Titanic menunjukkan bahwa peristiwa tragis ini tidak hanya dipengaruhi oleh faktor teknis, tetapi juga oleh aspek sosial dan demografis yang kompleks. Data yang diolah melalui Jupyter Notebook menggunakan Python berhasil mengungkap bahwa stratifikasi sosial memainkan peran kunci dalam menentukan peluang bertahan hidup penumpang. Penumpang kelas satu, yang memiliki akses lebih baik terhadap sekoci penyelamat dan fasilitas lainnya, menunjukkan tingkat keselamatan yang jauh lebih tinggi dibandingkan penumpang kelas dua dan tiga, yang sering kali kurang beruntung dalam situasi darurat. Selain itu, analisis berdasarkan usia menunjukkan bahwa anak-anak dan remaja memiliki peluang bertahan hidup yang lebih baik, sedangkan orang dewasa dan lansia cenderung menghadapi risiko lebih besar, sering kali karena keterbatasan fisik dan kemampuan untuk merespons cepat terhadap situasi darurat. Perbedaan gender juga berkontribusi pada hasil keselamatan; perempuan memiliki peluang lebih tinggi untuk bertahan hidup, yang dapat dijelaskan oleh norma sosial yang mendorong perlindungan terhadap wanita dan anak-anak dalam situasi berbahaya. Kesimpulannya, tragedi Titanic memberikan wawasan yang mendalam tentang bagaimana faktor-faktor sosial, demografis, dan perilaku manusia berinteraksi dalam menentukan hasil keselamatan. Pemahaman ini penting tidak hanya untuk menganalisis kejadian masa lalu, tetapi juga untuk merumuskan kebijakan keselamatan di masa depan.



Berdasarkan temuan penelitian ini, disarankan agar penelitian lebih lanjut dilakukan untuk menggali lebih dalam berbagai faktor yang mempengaruhi keselamatan penumpang dalam kejadian serupa. Analisis lebih lanjut dapat mencakup variabel-variabel lain, seperti faktor psikologis penumpang dalam situasi darurat, serta studi perbandingan dengan kecelakaan kapal lainnya. Penggunaan metode analisis data modern, seperti *Python* dan *Jupyter Notebook*, tidak hanya akan meningkatkan pemahaman kita terhadap data historis, tetapi juga dapat membantu dalam memprediksi dan mencegah potensi risiko di masa depan. Lebih jauh lagi, penting untuk meningkatkan kesadaran akan akses yang setara terhadap fasilitas penyelamatan dalam setiap perencanaan dan pengoperasian kapal, serta merumuskan kebijakan keselamatan yang lebih inklusif dan responsif terhadap kebutuhan semua penumpang, tanpa memandang status sosial, usia, atau gender. Dengan langkah-langkah ini, diharapkan bahwa tragedi serupa dapat diminimalkan, dan keselamatan penumpang dapat ditingkatkan secara signifikan.

UCAPAN TERIMA KASIH

Saya mengucapkan terima kasih kepada Tuhan Yang Maha Esa, orang-orang yang saya cintai, teman-teman, dan semua orang yang telah membantu dalam melaksanakan dan menyusun penelitian ini. Secara khusus, saya ingin mengucapkan terima kasih kepada para dosen pengajar yang telah memberikan nasihat, saran, dan pengeditan yang mendalam. Hasil penelitian ini tidak akan memadai tanpa bimbingan dan bantuan Anda. Diharapkan bahwa penemuan ini akan berdampak besar pada ilmu data, khususnya dalam memahami tragedi tenggelamnya Titanic.

DAFTAR PUSTAKA

1. Abdul Hafeez, Ali Hassan Sial. (2021). Comparative Analysis of Data Visualization Libraries Matplotlib and Seaborn in Python [HEC Y Cat] (diakses 2 Oktober 2024).
2. Few, S. (2008). Visual Business Intelligence - Data Visualization and Analysis - BI's Blind Spots. <https://www.perceptualedge.com/blog/?p=367> (diakses 29 September 2024).
3. Fisika, J., & Sriwijaya, U. (2019). Penggunaan Piranti Lunak Jupyter Notebook dalam Upaya Mensosialisasikan Open Science Dedi Setiabudidaya. 2–5 (diakses 1 Oktober 2024).
4. Galahartlambang, Y., Khotiah, T., & Jumain, J. (2021). Visualisasi Data Dari Dataset COVID-19 Menggunakan Pemrograman Python. Jurnal Ilmiah Intech: Information Technology Journal of UMUS, 3(01), 58–64. <https://doi.org/10.46772/intech.v3i01.417> (diakses 2 Oktober 2024).
5. Hidayati, F., & Kahirina. (2020). Tak Hanya agar Rapi, Pahami Arti dan Pentingnya Data Cleansing, Yuk! <https://glints.com/id/lowongan/data-cleansing-cleaning/> (diakses 2 Oktober 2024).
6. Huda, N. (2022). Mengenal Visual Studio Code. <https://www.dewaweb.com/blog/mengenal-visual-studio-code/> (diakses 9 Oktober 2024).
7. Meiryani. (2021). Memahami Perbedaan Analisis Kualitatif dan Analisis Kuantitatif dalam Penelitian Ilmiah. <https://accounting.binus.ac.id/2021/08/12/memahami-perbedaan-analisis-kualitatif-dan-analisis-kuantitatif-dalam-penelitian-ilmiah/> (diakses 2 Oktober 2024).
8. Moertini, V. S., dkk. (2020). Pengantar Data Science dan Aplikasinya bagi Pemula. Bandung, Indonesia: Unpar Press. Hlm. 20.



9. Novanra, F., Rusharijanto, R., & Rasyifa, dkk. (2019). Analisis Film Titanic Sebagai Media Kampanye Mitigasi Bencana Kapal Tenggelam. SEMIOTIKA: Jurnal Komunikasi, 13(02), 1–8.
<https://journal.ubm.ac.id/index.php/semiotika/article/download/1942/1584> (diakses 2 Oktober 2024).
10. Retnowardhani, A., & Ramdani, T. (2019). Jenis-Jenis Data Analytics.
<https://mmsi.binus.ac.id/2019/11/26/jenis-jenis-data-analytics/> (diakses 2 Oktober 2024).
11. Rohman, Y. A. (2019). Pengenalan NumPy, Pandas, Matplotlib | by Yasir Abdur Rohman | Medium (diakses 9 Oktober 2024).
12. Shanika, W. (2021). Jupyter Notebook for Data Analytics: A Beginner's Guide, 9 Agustus 2021.
<https://www.bmc.com/blogs/installing-jupyter-for-big-data-and-analytics/> (diakses 1 Oktober 2024).
13. Sugiono. (2020). Metode Penelitian Kualitatif & Kuantitatif. Yogyakarta: Penerbit Pustaka Ilmu. Hlm. 361–362. <https://anyflip.com/xobw/rfpq/basic> (diakses 2 Oktober 2024).
14. Tenggelamnya RMS Titanic.
https://id.wikipedia.org/wiki/Tenggelamnya_RMS_Titanic (diakses 1 Oktober 2024).
15. Tim Byjus. (2019). Data Sets.
<https://byjus.com/maths/data-sets/> (diakses 2 Oktober 2024).
16. Tim Byjus. (2019). Multivariate Datasets When the Dataset Contains Three or More Than Three Variables.
<https://byjus.com/maths/data-Sets/#:~:text=Multivariate%20Datasets&text=When%20the%20dataset%20contains%20three,or%20more%20than%20three%20variables> (diakses 2 Oktober 2024).
17. Titanic - Machine Learning from Disaster. <https://www.kaggle.com/c/titanic/overview> (diakses 27 September 2024).
18. RMS Titanic - Wikipedia.
https://id.wikipedia.org/wiki/RMS_Titanic (diakses 1 Oktober 2024).
19. Harri, P. (2023). [Python] — Mengenal Bahasa Pemrograman Python via Jupyter Notebook.
<https://medium.com/@puterahari> (diakses 9 Oktober 2024).