

EKSPLORASI DATA WINE PUTIH DAN KARAKTERISTIK KIMIAWI MENGUNAKAN METABASE

Putu Satya Saputra^{*1}, Yogiswara Dharma Putra²

¹Politeknik Negeri Bali

²Universitas Udayana

Email: ¹satya@pnb.ac.id, ²yogiswaradharma Putra@unud.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 07 Mei 2025, diterima untuk diterbitkan: 26 Juli 2025)

Abstrak

Karakteristik kimiawi wine putih memiliki pengaruh penting terhadap persepsi mutu dan penilaian kualitas produk. Beragam unsur seperti kadar alkohol, keasaman, gula residu, dan senyawa sulfur berperan dalam membentuk rasa, aroma, serta stabilitas wine. Analisis terhadap data wine putih dilakukan dengan pendekatan visualisasi interaktif menggunakan platform Metabase untuk mengungkap pola distribusi nilai kualitas dan hubungan antarparameter kimia. Dataset yang digunakan terdiri atas 4898 sampel dengan 11 variabel kimia dan satu skor kualitas dalam skala ordinal. Proses analisis meliputi pembersihan data, transformasi ke basis data relasional, dan pembangunan dashboard visual. Visualisasi yang dihasilkan mencakup *bar chart*, *scatter plot*, *pie chart*, dan *area chart*, yang digunakan untuk mengevaluasi tren dan perbandingan antar kategori mutu. Hasil menunjukkan bahwa wine berkualitas tinggi cenderung memiliki kadar alkohol lebih tinggi, sementara mayoritas sampel berada pada kategori kualitas sedang. Visualisasi berbasis Metabase memberikan kemudahan interpretasi data secara menyeluruh dan mendukung pengambilan keputusan berbasis bukti. Pendekatan ini juga dapat menjadi dasar pengembangan pemantauan mutu yang lebih sistematis serta peluang penerapan model prediktif pada tahap penelitian lanjutan.

Kata kunci: *wine putih, visualisasi data, karakteristik kimia, metabase*

EXPLORING WHITE WINE DATA AND CHEMICAL CHARACTERISTICS USING METABASE

Abstract

Chemical characteristics play a crucial role in shaping the quality and sensory perception of white wine. Various components, such as alcohol content, acidity, residual sugar, and sulfur compounds, contribute to flavor, aroma, and product stability. This study analyzes white wine data through an interactive visualization approach using the Metabase platform to uncover quality distribution patterns and relationships among chemical parameters. The dataset consists of 4,898 samples with 11 chemical variables and one ordinal quality score. The analytical process includes data cleaning, transformation into a relational database, and the development of a visual dashboard. The resulting visualizations—bar charts, scatter plots, pie charts, and area charts—are used to evaluate trends and compare quality categories. Findings indicate that high-quality wines tend to have higher alcohol content, while the majority of samples fall within the medium-quality category. Visualization through Metabase facilitates a comprehensive interpretation of complex information and supports evidence-based decision-making. This approach provides a foundation for more systematic quality monitoring and opens opportunities for the future application of predictive modeling in wine quality assessment.

Keywords: *white wine, data visualization, chemical characteristics, metabase*

1. PENDAHULUAN

Kualitas wine telah menjadi subjek yang menarik dalam dunia penelitian ilmiah, bisnis minuman, serta eksplorasi data berbasis sains. Setiap botol wine menyimpan cerita kompleks mengenai asal-usul bahan, proses fermentasi, hingga parameter kimiawi yang memengaruhi rasa, aroma, dan nilai jualnya. Teknologi analitik modern mempermudah proses pemahaman terhadap faktor-faktor pembentuk mutu wine secara kuantitatif dan visual. Penggunaan platform analitik Metabase memungkinkan interpretasi data menjadi lebih intuitif, efisien, dan komunikatif (Lyronis et al., 2025).

Wine putih merupakan jenis wine yang diproduksi melalui fermentasi anggur berkulit terang (Hernández-Montes et al., 2021). Warna, keasaman, kadar alkohol, serta komposisi kimia lainnya membentuk profil sensorik yang khas pada wine putih. Beberapa penelitian terdahulu menyoroti pentingnya karakteristik kimia yaitu kadar alkohol, keasaman volatil, residu gula, serta senyawa sulfur dalam menentukan kualitas akhir produk wine (Shahrajabian and Sun, 2024).

Analisis visual memiliki keunggulan dalam menyederhanakan pemahaman terhadap pola, tren, dan hubungan antardimensi data yang tidak mudah ditangkap melalui tabel atau deskripsi numerik semata (Yuan et al., 2021). Visualisasi data pada wine putih memberikan kontribusi penting untuk mengidentifikasi parameter kimia yang secara signifikan berkorelasi terhadap persepsi kualitas.

Dataset wine putih yang digunakan dalam penelitian ini berasal dari *open source data kaggle* yang telah digunakan secara luas dalam penelitian eksplorasi, prediktif, dan pengujian algoritma pembelajaran mesin. Dataset ini memuat 4898 observasi dengan 11 variabel kimia dan satu skor kualitas yang diberikan dalam skala ordinal (antara 0 hingga 10). Variabel-variabel tersebut meliputi *fixed acidity*, *volatile acidity*, *citric acid*, *residual sugar*, *chlorides*, *free sulfur dioxide*, *total sulfur dioxide*, *density*, *pH*, *sulphates*, dan *alcohol*. Skor *quality* yang diberikan berasal dari penilaian panel ahli terhadap rasa, aroma, dan keseimbangan wine.

Keunggulan eksplorasi visual berbasis Metabase terletak pada pendekatannya yang *low-code* (Yavari et al., 2024), serta kemampuannya dalam menyajikan dashboard interaktif berbasis SQL dan integrasi database relasional. Platform Metabase digunakan untuk menggali informasi penting dari dataset wine putih, menampilkan pola distribusi kualitas, serta menyusun hubungan antar variabel melalui grafik bar, diagram lingkaran, dan visualisasi area. Pendekatan visual ini diharapkan mampu memberikan gambaran menyeluruh dan praktis mengenai bagaimana kombinasi unsur kimiawi dapat berperan dalam pembentukan kualitas wine.

Pemahaman terhadap kualitas wine tidak hanya penting untuk kebutuhan akademik, melainkan juga memiliki implikasi luas pada proses produksi,

pemasaran, serta preferensi konsumen. Industri wine global saat ini menghadapi tantangan dalam mempertahankan standar mutu sembari tetap mengikuti preferensi pasar yang terus berubah (Golicic, 2022). Beberapa studi menunjukkan bahwa konsumen modern lebih tertarik pada wine berkualitas tinggi namun tetap transparan dalam informasi komposisi dan metode produksinya (Sun et al., 2022).

Model analitik berbasis visual tidak hanya menghemat waktu, tetapi juga mendorong partisipasi lintas fungsi dalam pengambilan keputusan (Nandwani and Verma, 2021). Visualisasi interaktif memudahkan para pemangku kepentingan dari berbagai latar belakang untuk memahami data, menyampaikan insight, serta menyusun strategi berbasis bukti (Locke et al., 2023). Hasil visualisasi ini dapat digunakan oleh ahli kimia, teknolog pangan, dan analis kualitas untuk menyusun langkah-langkah peningkatan mutu produk secara sistematis.

Metode yang digunakan dalam penelitian ini bersifat kuantitatif deskriptif. Proses analitik melibatkan pemuatan dataset ke dalam database relasional, koneksi ke Metabase, dan pengembangan dashboard eksploratif. Setiap *insight* yang dihasilkan juga ditandai secara visual dan didokumentasikan untuk mendukung pengambilan keputusan dan interpretasi lanjutan.

Tingkat kualitas wine putih sering kali diasosiasikan secara langsung dengan kadar alkohol (Gutiérrez-Escobar, Aliaño-González and Cantos-Villar, 2021). Studi sebelumnya menunjukkan bahwa konsentrasi alkohol dapat memengaruhi persepsi konsumen terhadap kelezatan, kehangatan, dan kesan keseluruhan wine (Ivanova et al., 2022). Alkohol juga berperan sebagai pelarut bagi senyawa aromatik, yang berkontribusi terhadap pengalaman sensorik produk. Selain alkohol, kandungan *sulphates* dan *citric acid* juga menunjukkan kontribusi terhadap struktur rasa dan stabilitas wine. Penurunan *volatile acidity* diketahui dapat mengurangi aroma asam menyengat yang biasanya tidak diinginkan.

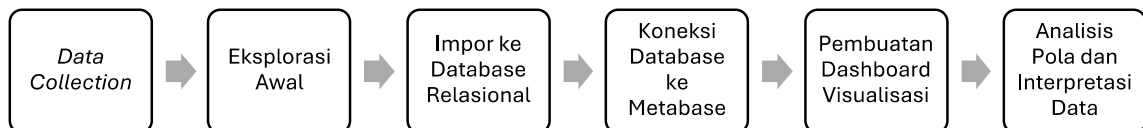
Visualisasi data memungkinkan pemetaan pola-pola tersebut dalam bentuk grafik yang mudah dimengerti, sekaligus dapat diakses secara *real time* oleh pengguna melalui dashboard. Kemampuan ini menjadikan Metabase sebagai alat yang ideal untuk kebutuhan eksplorasi awal dalam penelitian dan pengembangan mutu produk wine. Sebagai tambahan, dashboard visual juga dapat digunakan dalam pelatihan internal atau presentasi hasil evaluasi mutu kepada pemangku kepentingan non-teknis.

Eksplorasi data wine putih ini juga membuka peluang lanjutan untuk diterapkannya model prediktif berbasis *machine learning*. Kombinasi antara eksplorasi visual dan pemodelan prediktif dapat memperkuat proses *quality assurance* dan pengembangan produk baru dalam industri wine.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif yang difokuskan pada eksplorasi visual terhadap data wine putih menggunakan platform Metabase. Tujuannya untuk mengidentifikasi karakteristik kimiawi yang berkaitan dengan kualitas wine berdasarkan data terukur yang bersifat objektif. Analisis dilakukan dengan memanfaatkan metode visualisasi untuk memahami pola distribusi data.

Metode ini dipilih karena mampu menggambarkan informasi tersembunyi dalam data melalui representasi visual yang dapat dipahami lintas disiplin. Tahapan penelitian disusun secara sistematis mulai dari pemahaman dataset, pra-pemrosesan data, integrasi ke sistem database, hingga analisis visual berbasis Metabase.



Gambar 1. Alur Penelitian

2.1. Dataset Wine Putih

Dataset yang digunakan berasal dari platform Kaggle dan merupakan dataset terbuka (*open source*) yang berisi informasi kimia wine putih dan skor kualitasnya. Dataset ini telah digunakan dalam berbagai penelitian terdahulu untuk keperluan klasifikasi, regresi, maupun eksplorasi. File berformat CSV terdiri dari 4898 baris dan 12 kolom.

2.2. Eksplorasi Awal dan Pembersihan Data

Tahapan ini bertujuan untuk memahami struktur data secara umum, mengidentifikasi nilai-nilai ekstrem atau kosong, serta melakukan pembersihan data jika diperlukan. Proses ini dilakukan menggunakan Python (*pandas*) untuk mengecek tipe data, menghilangkan duplikasi, menangani nilai *null* serta menstandarkan nama kolom.

```
import pandas as pd
df = pd.read_csv('winequality-white.csv')
df.info()
df.dropna(inplace=True)
```

2.3. Impor ke Database Relasional

Data yang telah dibersihkan kemudian diimpor ke dalam sistem manajemen basis data relasional (SQLite). Tujuannya agar Metabase dapat melakukan *query* dan visualisasi secara efisien. Tabel dibuat dengan tipe data numerik untuk setiap kolom, dan kolom *quality* digunakan sebagai titik fokus analisis.

2.4. Koneksi Database ke Metabase

Metabase dikonfigurasi untuk terhubung ke database tersebut melalui panel administrasi. Setelah koneksi berhasil, tabel *white_wine* tersedia untuk eksplorasi visual. Proses ini bersifat *one-time setup* dan memungkinkan akses dinamis ke data terbaru.

2.5. Pembuatan Dashboard Visualisasi

Dashboard disusun untuk menjawab pertanyaan eksploratif mengenai kualitas wine putih berdasarkan karakteristik kimia.

Fokus utama diarahkan pada distribusi kualitas, hubungan antar fitur, dan perbedaan profil kimia berdasarkan kategori mutu. Visualisasi utama dalam dashboard mencakup:

- Bar Chart* untuk menggambarkan jumlah sampel pada setiap skor *quality*. Visualisasi ini digunakan untuk melihat distribusi umum kualitas wine dalam dataset.
- Scatter Plot* untuk memetakan hubungan antara *alcohol* (sumbu X) dan *quality* (sumbu Y). Digunakan untuk menangkap kecenderungan kadar alkohol terhadap kualitas wine.
- Pie Chart* untuk mengelompokkan wine ke dalam tiga kategori kualitas (*Low* ≤ 5 , *Medium* =6, *High* ≥ 7) berdasarkan skor *quality* menggunakan ekspresi CASE dalam SQL.
- Area Chart* untuk menunjukkan rata-rata kandungan kimia utama (*alcohol*, *residual sugar*, *sulphates*) per kategori kualitas wine. Digunakan untuk membandingkan tren fitur kimia berdasarkan klasifikasi mutu.

Setiap visualisasi dibangun melalui editor SQL maupun Query Builder di Metabase, dengan desain sederhana namun informatif agar mendukung proses eksplorasi yang efisien dan interaktif.

2.6. Analisis Pola dan Interpretasi Data

Analisis dilakukan terhadap setiap visualisasi yang ditampilkan di dashboard untuk mengidentifikasi pola-pola signifikan dalam data. Fokus utama diarahkan pada tren distribusi skor kualitas, kecenderungan meningkatnya kadar alkohol pada wine dengan skor tinggi, proporsi distribusi wine dalam kategori *Low*, *Medium*, dan *High Quality*, serta perbandingan rata-rata kandungan kimia utama antar kategori kualitas. Pola-pola tersebut dianalisis secara tematik untuk memberikan gambaran menyeluruh mengenai hubungan antar fitur kimia dan kontribusinya terhadap mutu wine putih secara keseluruhan. Hasil interpretasi kemudian disusun

dalam bentuk narasi ilmiah yang terstruktur pada bagian hasil dan pembahasan.

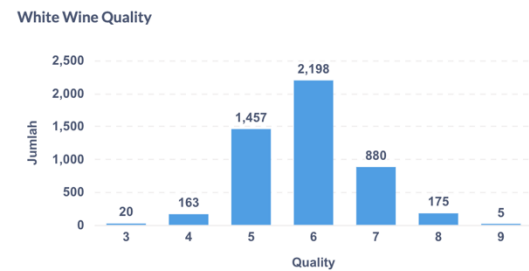
3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengeksplorasi karakteristik kimiawi wine putih dan kaitannya dengan kualitas produk melalui pendekatan visualisasi data berbasis Metabase. Tahapan eksploratif dilakukan untuk menggali pola umum, tren kandungan kimia, serta hubungan antar variabel penting terhadap skor kualitas. Hasil ini disajikan dalam dua bagian utama, yaitu deskripsi data awal dan interpretasi visualisasi dashboard.

Fixed Acidity	Volatile Acidity	Citric Acid	Residual Sugar	Chlorides	Free Sulfur Dioxide	Total Sulfur Dioxide	Density	pH	Sulphates	Alcohol	Quality
7	0.27	0.34	20.7	0.045	45	170	1	3	0.45	8.8	4
6.5	0.3	0.34	1.6	0.049	34	132	0.99	3.3	0.49	9.3	4
8.5	0.28	0.4	4.9	0.05	39	97	1	3.26	0.44	10.1	4
7.2	0.25	0.22	4.5	0.028	47	88	1	3.35	0.4	9.9	4
7.6	0.23	0.22	5.3	0.028	47	88	1	3.37	0.4	9.9	4
8.1	0.28	0.4	4.9	0.05	39	97	1	3.26	0.44	10.1	4
6.2	0.32	0.16	7	0.06	35	156	0.99	3.38	0.47	9.4	4
7	0.27	0.34	20.7	0.045	45	170	1	3	0.45	8.8	4
6.5	0.3	0.34	1.6	0.049	34	132	0.99	3.3	0.49	9.3	4
8.5	0.28	0.43	1.5	0.044	38	129	0.99	3.23	0.45	10	4

Gambar 2. Tampilan Dataset Wine Putih 10 Baris Pertama

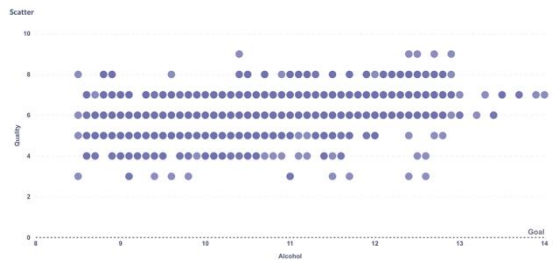
Tabel awal yang dianalisis terdiri dari 11 parameter kimiawi yang memengaruhi kualitas wine putih. Variabel *fixed acidity*, *volatile acidity*, *citric acid*, *residual sugar*, *chlorides*, *free sulfur dioxide*, *total sulfur dioxide*, *density*, *pH*, *sulphates*, dan *alcohol* diukur terhadap skor kualitas (*quality*) yang ditentukan oleh panel ahli. Sepuluh sampel pertama memperlihatkan bahwa sebagian besar wine memiliki skor kualitas 6, dengan nilai alkohol yang bervariasi antara 8.8 hingga 11.0. Kadar *volatile acidity* cenderung berada di rentang 0.22–0.32, sedangkan *sulphates* rata-rata mendekati 0.45. Data ini menunjukkan variasi komposisi yang cukup stabil, namun alkohol dan *residual sugar* memiliki fluktuasi yang lebih besar, mengindikasikan potensi kontribusi terhadap perbedaan persepsi mutu.



Gambar 3. Distribusi Skor Kualitas Wine (Bar Chart)

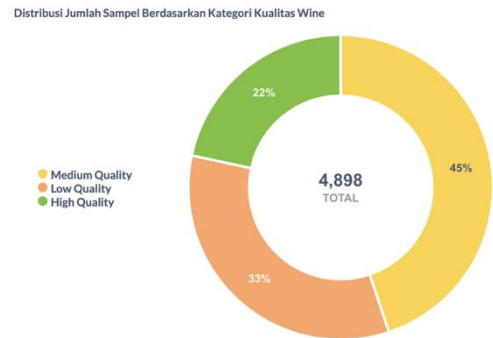
Visualisasi batang pada Gambar 3. menunjukkan distribusi jumlah sampel berdasarkan skor kualitas wine. Skor kualitas paling dominan adalah 6 dengan total 2.198 sampel, diikuti skor 5 (1.457 sampel) dan 7 (880 sampel). Distribusi ini membentuk pola lonceng tidak simetris dengan puncak pada skor 6, mengindikasikan bahwa sebagian besar wine putih berada dalam kategori mutu sedang. Skor kualitas 3 dan 9 hanya muncul dalam jumlah sangat kecil, masing-masing 20 dan 5 sampel, menunjukkan jarang wine dengan

kualitas sangat rendah maupun sangat tinggi dalam dataset ini. Pola ini mendukung hipotesis bahwa produksi wine secara umum cenderung menghasilkan kualitas menengah yang lebih konsisten.



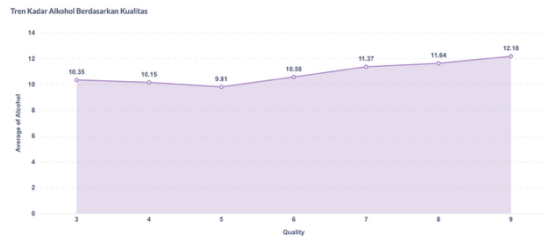
Gambar 4. Hubungan Alkohol dan Kualitas Wine (Scatter Plot)

Visualisasi pencar antara variabel *alcohol* dan *quality* memperlihatkan konsentrasi titik pada rentang alkohol antara 9.0 hingga 11.5. Sebagian besar wine dengan kualitas tinggi (7 hingga 9) memiliki kadar alkohol yang relatif lebih tinggi dibandingkan wine berkualitas rendah (3 hingga 5). Pola sebaran ini menunjukkan korelasi positif lemah antara kadar alkohol dan peningkatan skor kualitas. Hal ini menunjukkan bahwa alkohol dapat menjadi salah satu indikator kualitas dalam wine putih, meskipun bukan satu-satunya faktor yang menentukan.



Gambar 5. Distribusi Kategori Kualitas Wine (Pie Chart)

Visualisasi *pie chart* membagi seluruh dataset menjadi tiga kategori: *Low Quality* (33%), *Medium Quality* (45%), dan *High Quality* (22%). Proporsi terbesar berada pada kualitas sedang, mengonfirmasi hasil bar chart sebelumnya. Proporsi wine berkualitas tinggi relatif kecil, memperkuat anggapan bahwa mencapai mutu tinggi dalam produksi wine putih memerlukan kontrol kualitas yang lebih ketat. Presentasi distribusi ini juga membantu menyeimbangkan fokus analisis, agar tidak hanya bergantung pada rata-rata dataset secara keseluruhan, tetapi juga mempertimbangkan proporsi segmen unggulan dan tertinggal.



Gambar 6. Tren Rata-Rata Alkohol Berdasarkan Skor Kualitas (Area Chart)

Area chart menunjukkan tren rata-rata kadar alkohol pada masing-masing skor kualitas. Rata-rata alkohol pada wine dengan kualitas 3 hingga 5 menurun dari 10.35 ke 9.81, kemudian mengalami peningkatan konsisten mulai dari kualitas 6 (10.58) hingga kualitas 9 (12.18). Pola ini mempertegas hasil *scatter plot* sebelumnya, yaitu kecenderungan kadar alkohol yang meningkat seiring dengan naiknya skor kualitas. Kenaikan ini cukup signifikan, menandakan bahwa alkohol berperan sebagai salah satu penentu preferensi rasa dan persepsi mutu dalam wine putih. Grafik ini mendukung interpretasi bahwa kategori *high quality* memiliki profil alkohol lebih tinggi secara konsisten.

4. SIMPULAN

Eksplorasi data wine putih melalui visualisasi di Metabase memberikan gambaran yang jelas mengenai hubungan antara komposisi kimiawi dan kualitas produk. Mayoritas wine berada pada kualitas menengah, sementara kadar alkohol cenderung lebih tinggi pada wine dengan kualitas lebih baik. Pola ini menunjukkan bahwa alkohol dapat menjadi indikator penting dalam menentukan persepsi mutu, meskipun bukan satu-satunya faktor penentu. Visualisasi *bar chart*, *scatter plot*, *pie chart*, dan *area chart* berhasil menyederhanakan kompleksitas data menjadi wawasan yang mudah dipahami. Platform Metabase mempermudah proses ini melalui tampilan yang interaktif dan informatif. Temuan ini dapat dimanfaatkan dalam proses pengendalian mutu, strategi produksi, dan edukasi konsumen. Penelitian ini menegaskan bahwa di balik setiap angka terdapat cerita, dan visualisasi menjadi cara efektif untuk menceritakannya.

5. DAFTAR PUSTAKA

- GOLICIC, S.L., 2022. Changes in sustainability in the global wine industry. *International Journal of Wine Business Research*, 34(3), pp.392–409. <https://doi.org/10.1108/IJWBR-03-2021-0021>.
- GUTIÉRREZ-ESCOBAR, R., ALIAÑO-GONZÁLEZ, M.J. AND CANTOS-VILLAR, E., 2021. Wine Polyphenol Content and Its Influence on Wine Quality and Properties: A Review. *Molecules*, 26(3), p.718. <https://doi.org/10.3390/molecules26030718>
- HERNÁNDEZ-MONTES, E., ZHANG, Y., CHANG, B.-M., SHCHERBATYUK, N. and KELLER, M., 2021. Soft, Sweet, and Colorful: Stratified Sampling Reveals Sequence of Events at the Onset of Grape Ripening. *American Journal of Enology and Viticulture*, 72(2), pp.137–151. <https://doi.org/10.5344/ajev.2020.20050>.
- IVANOVA, N., YANG, Q., BASTIAN, S.E.P., WILKINSON, K.L. and FORD, R., 2022. Consumer understanding of beer and wine body: An exploratory study of an ill-defined concept. *Food Quality and Preference*, 98, p.104383. <https://doi.org/10.1016/j.foodqual.2021.104383>.
- LOCKE, J., CORBIN, C.M., COOK, C.R., EHRHART, M.G., DAVIS, C. and LYON, A.R., 2023. Using Stakeholder Input to Guide Data Visualization and Reporting to Promote Evidence-based Practice Use in Public Schools. *Global Implementation Research and Applications*, 3(2), pp.99–111. <https://doi.org/10.1007/s43477-023-00080-9>.
- LYRONIS, A., VAROUCHAKIS, E.A., ANYFANTI, I. and KARATZAS, G.P., 2025. A Unified Open-Source Metabase Design for Hydro-meteorological Data Descriptive Analysis. *Environmental Processes*, 12(2), p.18. <https://doi.org/10.1007/s40710-025-00758-2>.
- NANDWANI, P. and VERMA, R., 2021. A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), p.81. <https://doi.org/10.1007/s13278-021-00776-6>.
- SHAHRAJABIAN, M.H. and SUN, W., 2024. Assessment of Wine Quality, Traceability and Detection of Grapes Wine, Detection of Harmful Substances in Alcohol and Liquor Composition Analysis. *Letters in Drug Design & Discovery*, 21(8), pp.1377–1399. <https://doi.org/10.2174/1570180820666230228115450>.
- SUN, X., ZHANG, F., GUTIÉRREZ-GAMBOA, G., GE, Q., XU, P., ZHANG, Q., FANG, Y. and MA, T., 2022. Real wine or not? Protecting wine with traceability and authenticity for consumers: chemical and technical basis, technique applications, challenge, and perspectives. *Critical Reviews in Food Science and Nutrition*, 62(24), pp.6783–6808. <https://doi.org/10.1080/10408398.2021.1906624>.

- YAVARI, A., HARRISON, C.J., GORJI, S.A. and SHAFIEI, M., 2024. Hydrogen 4.0: A Cyber-Physical System for Renewable Hydrogen Energy Plants. *Sensors*, 24(10), p.3239. <https://doi.org/10.3390/s24103239>.
- YUAN, J., CHEN, C., YANG, W., LIU, M., XIA, J. and LIU, S., 2021. A survey of visual analytics techniques for machine learning. *Computational Visual Media*, 7(1), pp.3–36. <https://doi.org/10.1007/s41095-020-0191-7>.