

PENERAPAN SISTEM INTELIJENSIA BISNIS DAN K-MEANS CLUSTERING UNTUK MEMANTAU PRODUKSI TANAMAN OBAT

IMPLEMENTATION OF BUSINESS INTELLIGENCE SYSTEM AND K-MEANS CLUSTERING TO MONITOR THE MEDICINAL PLANTS PRODUCTION

Miwan Kurniawan Hidayat^{1,2)}, dan Rina Fitriana^{2,3)*}

¹⁾Prodi Teknik Industri Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, Senen, Jakarta

²⁾Magister Teknik Industri Universitas Trisakti, Jl. Kyai Tapa No.1, Grogol, Jakarta

³⁾Jurusan Teknik Industri Universitas Trisakti, Jl. Kyai Tapa No.1, Grogol, Jakarta

*Email: rinaf@trisakti.ac.id

Makalah: Diterima 08 Mei 2022; Diperbaiki 21 Juli 2022; Disetujui 20 Agustus 2022

ABSTRACT

Indonesia has biodiversity including medicinal plants. The medicinal plant business can be a profitable business prospect because it has high export opportunities. Based on the benefits obtained, the production of medicinal plants needs to be considered by monitoring and evaluating production results to increase productivity, especially for areas with low production levels. The Department of Food Crops and Horticulture of West Java Province through the website <https://opendata.jabarprov.go.id> has provided production datasets for each type of medicinal plant, but it has not yet become a dataset with various types of medicinal plants. The purpose of this study was to design an integrated data storage model in the form of a data warehouse, grouping medicinal plant production areas using data mining and designing data visualization in business intelligence systems. Business intelligence system design was carried out through several stages, namely system requirements analysis, identification of data and information needs, data warehouse design, data warehouse filling, data mining processes, data visualization, and system performance evaluation. The results of the research were the application of a data warehouse using a dimensional model with a star schema; a grouping of production areas using the K-Means algorithm with optimal $k=3$ and the number of elements produced in each cluster is 24 regions in cluster 0, 1 region in cluster 1, and 2 regions in cluster 2; a business intelligence system is implemented using a dashboard to show information on the amount of production and display the results of grouping potential areas for producing medicinal plants.

Keywords: business intelligence, data warehouse, dimensional models, visualization, medicinal plants

ABSTRAK

Indonesia memiliki keanekaragaman hayati termasuk tanaman obat. Bisnis tanaman obat bisa menjadi suatu prospek bisnis yang menguntungkan karena mempunyai peluang ekspor yang tinggi. Atas dasar manfaat yang didapat, maka produksi tanaman obat perlu diperhatikan dengan melakukan pemantauan dan evaluasi hasil produksi untuk meningkatkan produktivitas terutama untuk wilayah yang memiliki tingkat produksi rendah. Dinas Tanaman Pangan dan Hortikultura Provinsi Jawa Barat melalui situs web <https://opendata.jabarprov.go.id> telah menyediakan *dataset* produksi per jenis tanaman obat, namun belum menjadi satu dataset dengan berbagai jenis tanaman obat. Tujuan penelitian ini adalah merancang model penyimpanan data terintegrasi berupa *data warehouse*, pengelompokan wilayah produksi tanaman obat menggunakan *data mining* serta merancang visualisasi data pada sistem inteligensia bisnis. Perancangan sistem inteligensia bisnis dilakukan melalui beberapa tahapan yaitu analisis kebutuhan sistem, identifikasi kebutuhan data dan informasi, perancangan *data warehouse*, pengisian *data warehouse*, proses *data mining*, visualisasi data, dan evaluasi kinerja sistem. Hasil penelitian adalah penerapan *datawarehouse* menggunakan model dimensional dengan skema bintang; pengelompokan wilayah produksi menggunakan algoritma *K-Means* dengan k optimal = 3 dan jumlah elemen yang dihasilkan pada setiap *cluster* adalah 24 wilayah di *cluster* 0, 1 wilayah di *cluster* 1, dan 2 wilayah di *cluster* 2; sistem inteligensia bisnis diterapkan menggunakan *dashboard* untuk menunjukkan informasi jumlah produksi serta menampilkan hasil pengelompokan wilayah potensial penghasil tanaman obat.

Kata kunci: inteligensia bisnis, *data warehouse*, model dimensional, visualisasi, tanaman obat

PENDAHULUAN

Indonesia merupakan salah satu negara yang memiliki banyak keanekaragaman hayati termasuk tanaman yang memiliki khasiat sebagai obat (biofarmaka). Salah satu alternatif pengobatan di

masyarakat adalah dengan cara mengonsumsi tanaman obat, selain untuk menyembuhkan, tanaman obat bermanfaat untuk mencegah penyakit dan memulihkan kondisi kesehatan. Biofarmaka memiliki khasiat karena dalam tanaman tersebut terkandung zat aktif yang dapat mengobati suatu penyakit tertentu

atau jika tidak mengandung zat aktif tertentu namun mengandung efek resultan/sinergi dari beragam zat yang memiliki efek mengobati. Pemanfaatan biofarmaka sebagai obat yaitu dengan cara diminum, dihirup, atau ditempel (Sarno, 2019). Dalam penelitian tanaman obat menunjukkan sedikitnya terdapat 44 jenis tumbuhan obat dari 30 famili untuk kosmetik dan kesehatan wanita (Falah dan Hadiwibowo, 2017) serta 87 jenis tumbuhan yang termasuk dalam 42 famili yang dimanfaatkan sebagai obat oleh masyarakat (Rahayu dan Andini, 2019).

Keberadaan tanaman obat dapat juga dimanfaatkan untuk meningkatkan pendapatan dan kesejahteraan masyarakat. Bisnis tanaman obat bisa menjadi suatu prospek bisnis yang menguntungkan, hal ini disebabkan tanaman obat mempunyai peluang ekspor yang tinggi. Berdasarkan data dari Badan Pusat Statistik terdapat lebih dari 10 negara yang mengimpor tanaman obat, aromatik, dan rempah-rempah dari Indonesia. Data ekspor mulai dari tahun 2016 sampai 2020 dapat dilihat pada Tabel 1.

Provinsi Jawa Barat adalah salah satu wilayah penghasil tanaman obat. Wilayah produksi tanaman obat di provinsi Jawa Barat tersebar di beberapa kabupaten/kota. Atas dasar manfaat yang didapat dari tanaman obat yaitu untuk kesehatan dan kesejahteraan masyarakat, maka produksi tanaman obat perlu diperhatikan dengan melakukan pemantauan dan evaluasi hasil produksi untuk meningkatkan produktivitas terutama untuk wilayah yang memiliki tingkat produksi rendah. Dinas Tanaman Pangan dan Hortikultura Provinsi Jawa Barat melalui situs web <https://opendata.jabarprov.go.id> telah menyediakan *dataset* produksi per jenis tanaman obat, namun belum menjadi satu dataset dengan berbagai jenis tanaman obat. Berdasarkan kondisi yang ada dan menjadi permasalahan adalah data hasil produksi belum terintegrasi, belum ada pengelompokan data wilayah produksi, serta belum tersedia visualisasi hasil produksi untuk beragam jenis tanaman obat.

Tujuan penelitian ini adalah merancang model penyimpanan data terintegrasi berupa *data warehouse*, pengelompokan wilayah produksi tanaman obat menggunakan *data mining* serta merancang visualisasi data pada sistem intelijensia bisnis. Metode yang digunakan untuk *data mining* adalah *K-Means clustering* karena memiliki kemampuan dalam proses pengelompokan data yang berjumlah besar dengan kecepatan waktu komputasi yang relatif efisien (Ahmar *et al.*, 2018).

Visualisasi informasi bertujuan mewakili berbagai jenis data (geografis, numerik, teks, jaringan) secara visual untuk mengaktifkan dan memperkuat kognisi. Visualisasi informasi menawarkan cara intuitif untuk persepsi dan manipulasi informasi yang pada dasarnya memperkuat kinerja kognitif keseluruhan pemrosesan informasi, terutama untuk pengguna yang tidak memiliki keahlian (Po *et al.*, 2020). Visualisasi menunjukkan data stories dengan mengubah pola kuantitatif, relasional, atau spasial menjadi gambar (Dougherty dan Ilyankou, 2021). Sebuah visualisasi data harus menyampaikan data secara akurat dan tidak boleh menyesatkan atau mendistorsi. Jika satu angka dua kali lebih besar dari yang lain, tetapi dalam visualisasi terlihat hampir sama, maka visualisasinya salah. Pada saat yang sama, visualisasi data harus estetik (Wilke, 2019).

Visualisasi data sangat penting bagi organisasi karena mengurangi waktu yang dihabiskan di ruang rapat setidaknya 30%. Saat merepresentasikan teks dan angka dalam bentuk bagan dan tabel, maka dapat dengan mudah menyampaikan informasi kepada pihak lain. Beberapa teknik visualisasi yang dapat digunakan untuk mewakili kumpulan data adalah (Campbell, 2021):

- *Number Charts*

Dengan menggunakan *number charts* dapat secara efektif dan efisien menggambarkan informasi dalam format numerik.

Tabel 1. Ekspor tanaman obat, aromatik, dan Rempah-rempah (Badan Pusat Statistik, 2021)

No,	Negara Tujuan	2016	2017	2018	2019	2020
Berat bersih : Ton						
1	Pakistan	114 499,8	90 975,6	2 379,1	1 653,0	1 057,3
2	Thailand	60 117,4	68 299,9	101 664,0	162 812,3	64 560,9
3	Amerika Serikat	10 531,4	13 347,7	12 619,8	7 182,0	12 918,3
4	India	22 589,3	24 169,8	33 572,0	31 939,9	33 995,1
5	Vietnam	20 640,6	29 159,6	29 977,8	11 480,8	9 349,5
6	Singapura	11 218,0	11 179,9	10 170,5	5 007,0	4 276,3
7	Belanda	2 336,5	3 799,9	2 672,2	2 235,9	2 666,1
8	Tiongkok	15 900,4	5 496,1	4 413,2	10 297,8	18 950,5
9	Bangladesh	16 309,1	23 004,4	6 278,0	10 899,1	7 407,9
10	Jerman	897,8	1 305,9	1 357,1	1 341,3	1 515,3
11	Lainnya	41 123,6	55 053,4	130 989,5	73 291,2	118 597,9
Jumlah		316 163,9	325 792,2	336 093,2	318 140,3	275 295,1

- *Maps*
Salah satu keuntungan menggunakan *maps* untuk merepresentasikan data secara grafis yaitu menarik untuk dilihat, dapat dengan mudah merepresentasikan informasi dalam kumpulan data, dapat mewakili volume besar informasi kompleks mengenai berbagai topik menggunakan *maps* dan sangat mudah untuk mencerna informasi yang digambarkan melalui *maps*.
- *Pie Charts*
Pie charts digunakan sebagai cara tradisional untuk merepresentasikan informasi. *Pie charts* merupakan salah satu cara yang baik untuk memvisualisasikan data dalam kumpulan data.
- *Gauge Charts*
Gauge charts dapat merepresentasikan data dengan titik atau nilai data tunggal menggunakan grafik pengukur. Dapat digunakan pada laporan *dashboard* eksekutif yang dibagikan dengan pemangku kepentingan.

Penelitian terdahulu yang telah dilakukan terkait intelijensia bisnis serta pengelompokan data menggunakan proses *data mining* yaitu penelitian visualisasi harga dan pasokan beras di pasar induk beras Cipinang dengan membangun *dashboard* yang menampilkan berbagai data harga beras di tingkat grosir dan eceran serta data beras yang dipasok melalui Pasar Induk Beras Cipinang berbasis Google Data Studio (Sugiarto *et al.*, 2021). Penelitian penerapan *Sales Executive Dashboard* pada perusahaan *multistore* di Yogyakarta untuk memantau dan menganalisis kondisi penjualan. *Dashboard* menampilkan grafik yang memudahkan pengguna, khususnya manajer penjualan dan pemilik untuk mempelajari kondisi penjualan terkini (Santoso *et al.*, 2017). Penelitian tentang pemilihan teknik visualisasi yang paling tepat untuk analisis intelijensia bisnis bagi pengguna non ahli sesuai dengan kebutuhan informasi dan data dengan cara semi otomatis (Lavallo *et al.*, 2019). Penelitian Sistem Intelijensia Bisnis Dalam Manajemen Pengelolaan Pelanggan Dan Mutu Untuk Agroindustri Susu dengan menerapkan data mining dan OLAP Cube (Fitriana *et al.*, 2012). Penelitian perancangan model sistem intelijensia bisnis dengan menerapkan *data mining* untuk mengelompokan cacat produk ketika proses produksi di perusahaan roti (Fitriana *et al.*, 2017). Penelitian sistem intelijensia bisnis sebagai alat pendukung keputusan untuk meningkatkan kepuasan pelanggan atas layanan yang diberikan oleh perusahaan (Fitriana *et al.*, 2016). Penelitian sistem intelijensia bisnis sebagai alat untuk mendapatkan informasi yang bermanfaat dan akurat dalam pengambilan keputusan pada bagian pemasaran di pabrik roti dengan menerapkan data mining dan OLAP Cube (Fitriana *et al.*, 2018). Penelitian eksplorasi faktor-faktor yang terkait dengan penciptaan nilai bisnis pada intelijensia bisnis dan analitik di usaha manufaktur yang sedang menjalani transformasi Industri 4.0 (Bordeleau *et al.*, 2020).

Penelitian pengelompokan area rawan terjadi kecelakaan lalu lintas di jalan raya menggunakan *K-Means clustering* (Purwaningsih, 2019; Rahmadani *et al.*, 2021). Penerapana algoritma *clustering* data penjualan di rumah makan menggunakan *Fuzzy C-Means* (Agustini, 2017). Pemanfaatan *K-Means clustering* dalam upaya memetakan potensi pelanggan sebagai langkah strategis dalam menjalankan promosi pada suatu toko pakaian (Mardalius dan Christy, 2020). Penerapan teknik *clustering* dalam mengidentifikasi bahan makanan halal di kantin (Sucipto *et al.*, 2021).

Untuk mencapai keunggulan daya saing, sebuah organisasi akan berupaya mengubah bentuk data operasional ke bentuk intelijensia bisnis yang bermanfaat dalam memenuhi keperluan informasi dan mengumpulkan data dari berbagai sumber untuk mendapatkan informasi guna membantu dalam membuat keputusan yang tepat, hal ini dapat dengan mudah dilakukan menggunakan proses *data warehouse*. Sistem *data warehouse* dapat mendukung proses analisis yang dinamis dan menghasilkan informasi dengan cepat. Kelebihan penggunaan *query* dalam sistem *data warehouse* adalah terdapat konsistensi serta kualitas data, efisiensi dana, tepat waktu dalam pengaksesan data, produktivitas dan kinerja yang meningkat (MIT Database Notes, n.d.). Kumpulan data pada *data warehouse* berasal dari berbagai sumber, informasi yang diekstraksi kemudian digunakan untuk tujuan analisis, yang membantu bisnis membuat keputusan organisasi yang tepat (Campbell, 2021).

METODE PENELITIAN

Perancangan sistem intelijensia bisnis pada penelitian dilakukan melalui beberapa tahapan yaitu:

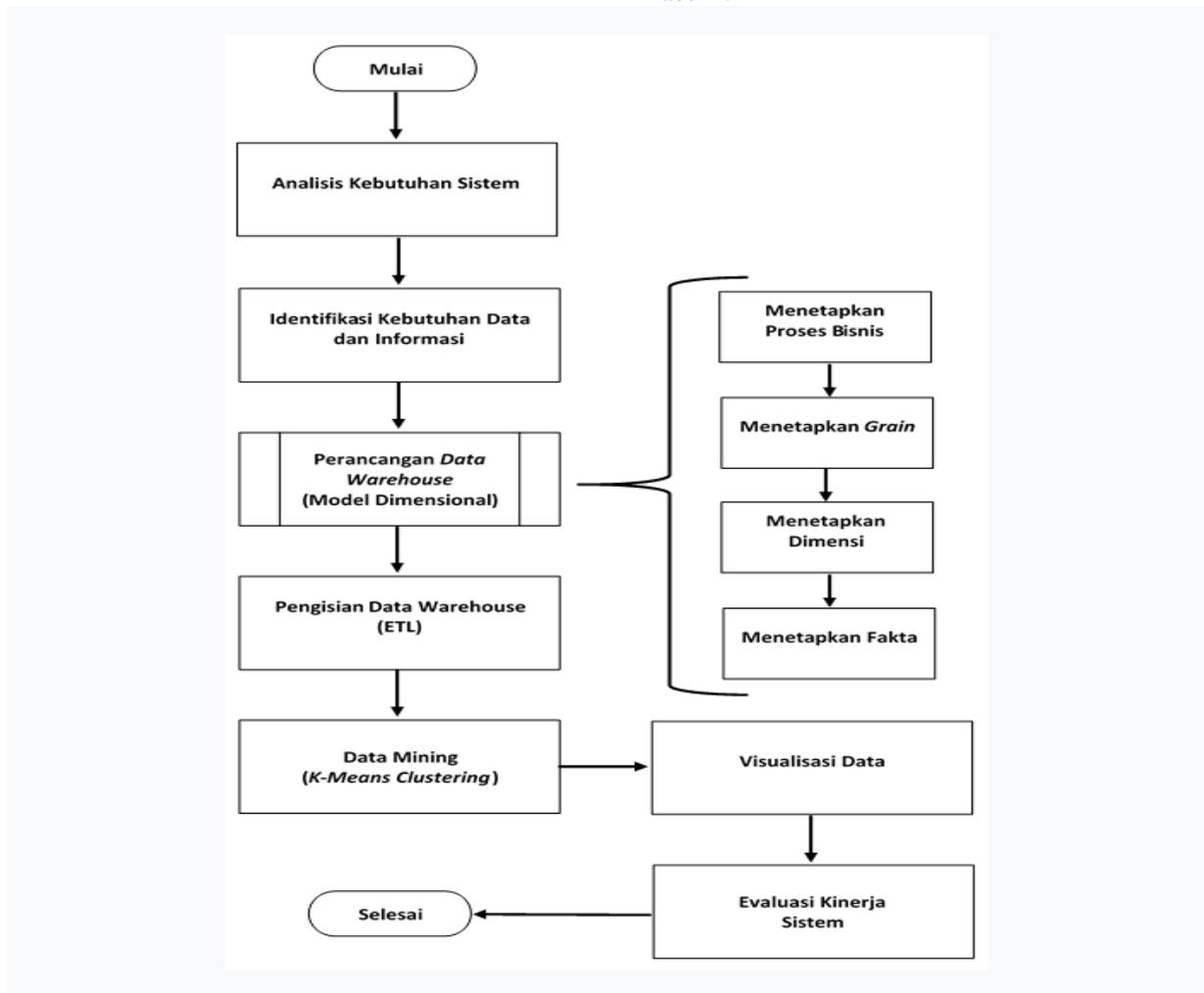
1. Analisis Kebutuhan Sistem
Analisis PIECES (*performance, information, economic, control, efficiency, service*) digunakan dalam melakukan analisis terhadap kebutuhan sistem intelijensia bisnis yang dibuat untuk memenuhi kebutuhan proses bisnis, sehingga penerapan sistem sesuai kebutuhan pengguna.
2. Identifikasi Kebutuhan Data dan Informasi.
Proses mengidentifikasi kebutuhan data dan informasi yang diperlukan pada penelitian yang dilakukan dengan teknik observasi.
3. Perancangan *Data Warehouse*.
Model dimensional digunakan pada perancangan *data warehouse* melalui empat tahap yang terdiri dari menetapkan proses bisnis, menetapkan *grain*, menetapkan dimensi dan menetapkan fakta.
4. Pengisian *Data Warehouse*
Data warehouse yang sudah dirancang diisi data melalui proses *Extract, Transform dan Load* (ETL). Pada proses ETL komputer akan melakukan ekstraksi, transformasi, dan memuat berbagai jenis data yang dihasilkan oleh komputer

- dari berbagai sumber data dan kemudian mengirimkan data ke *data warehouse*.
5. Proses *Data Mining*
K-Means clustering digunakan pada proses *data mining* untuk mengelompokkan wilayah potensial penghasil tanaman obat.
 6. Visualisasi Data
 Informasi yang dihasilkan dari *data warehouse* akan disajikan dalam bentuk visualisasi data melalui antarmuka *dashboard*.
 7. Evaluasi Kinerja Sistem
 Evaluasi kinerja sistem dilakukan untuk memastikan sistem yang dirancang berjalan sesuai dengan ketentuan sehingga kebutuhan pengguna terpenuhi. Evaluasi kinerja sistem dilakukan dengan cara uji verifikasi dan uji validasi. Uji verifikasi dilakukan untuk memastikan sistem berjalan sesuai dengan ketentuan dan tidak terjadi *bug*. Uji validasi dilakukan untuk memastikan sistem yang dirancang mampu menyampaikan informasi sesuai kebutuhan pengguna.

Diagram alir sebagai gambaran alur pengerjaan atau proses penelitian dapat dilihat pada Gambar 1.

HASIL DAN PEMBAHASAN

Pada proses bisnis pemantauan, jumlah hasil produksi dari semua tanaman obat di berbagai wilayah akan terekam dengan baik. Suatu sistem pemantauan diperlukan agar pemantauan berjalan baik didukung oleh ketersediaan data yang akurat. Dengan adanya sistem inteligensia bisnis dalam pemantauan hasil produksi, pengguna dapat menerima informasi hasil produksi tanaman obat untuk mendukung pengambilan keputusan. Informasi lengkap yang tersedia seperti informasi berbagai wilayah produksi, perbandingan jumlah produksi, dan pengelompokan wilayah produksi membantu pengambilan keputusan dalam waktu yang cepat. Kebutuhan sistem inteligensia bisnis yang dibuat untuk memenuhi kebutuhan proses bisnis berdasarkan hasil proses analisis dapat dilihat pada Tabel 2.



Gambar 1. Diagram alir metode penelitian

Tabel 2. Analisis PIECES

PIECES	Sistem Berjalan	Sistem Usulan
<i>Performance</i>	Memerlukan waktu proses lebih lama untuk mendapatkan informasi jumlah produksi dari beberapa jenis tanaman obat secara bersamaan dikarenakan informasi yang disajikan masih per jenis tanaman obat.	Penyampaian informasi melalui visualisasi dashboard untuk mengakses informasi jumlah produksi dari beberapa jenis tanaman obat secara bersamaan dilakukan dengan cepat.
<i>Information</i>	<ul style="list-style-type: none"> - Penghitungan jumlah produksi berbagai jenis tanaman obat pada suatu wilayah tertentu harus dilakukan secara manual yang dapat menyebabkan terjadi kesalahan hitung. - Tidak ada informasi perbandingan jumlah produksi. - Tidak ada pengelompokan (<i>clustering</i>) wilayah berdasarkan jumlah produksi. 	<ul style="list-style-type: none"> - Penghitungan jumlah produksi berbagai jenis tanaman obat dilakukan secara otomatis untuk mengurangi terjadi kesalahan hitung sehingga informasi lebih akurat. - Terdapat informasi perbandingan jumlah produksi berdasarkan wilayah, jenis produk dan tahun. - Terdapat informasi pengelompokan (<i>clustering</i>) wilayah berdasarkan jumlah produksi.
<i>Economic</i>	Penghitungan jumlah produksi berbagai jenis tanaman obat pada suatu wilayah tertentu harus dilakukan secara manual sehingga memerlukan biaya atas sumber daya manusia sebagai operator.	Penghitungan jumlah produksi berbagai jenis tanaman obat dilakukan secara otomatis sehingga tidak perlu ada biaya atas sumber daya manusia sebagai operator.
<i>Control</i>	Tidak ada kendali dalam mengintegrasikan data jumlah produksi dari beberapa jenis tanaman obat.	Data diintegrasikan dan disimpan dalam data warehouse sehingga data dapat dikendalikan.
<i>Efficiency</i>	Proses pencarian informasi dilakukan berulang untuk mendapatkan informasi jumlah produksi dari jenis tanaman obat yang berbeda.	Proses pencarian informasi dilakukan dengan cepat melalui filter informasi untuk jenis tanaman obat yang berbeda.
<i>Service</i>	Akses informasi melalui beberapa langkah berulang serta proses rekapitulasi manual yang memperlambat proses pengambilan keputusan.	Akses informasi dan pengambilan keputusan dapat dilakukan dengan cepat dan mudah melalui sistem intelijensia bisnis.

Arsitektur sistem intelijensia bisnis yang dirancang mencakup sistem dan perangkat lunak yang diperlukan untuk mengumpulkan, menyimpan, mengintegrasikan, menganalisis data serta menyajikan informasi tentang hasil produksi tanaman obat dapat dilihat pada Gambar 2. Informasi disampaikan kepada pihak-pihak yang memerlukan informasi jumlah produksi tanaman obat dalam menjalankan proses bisnis, diantaranya adalah:

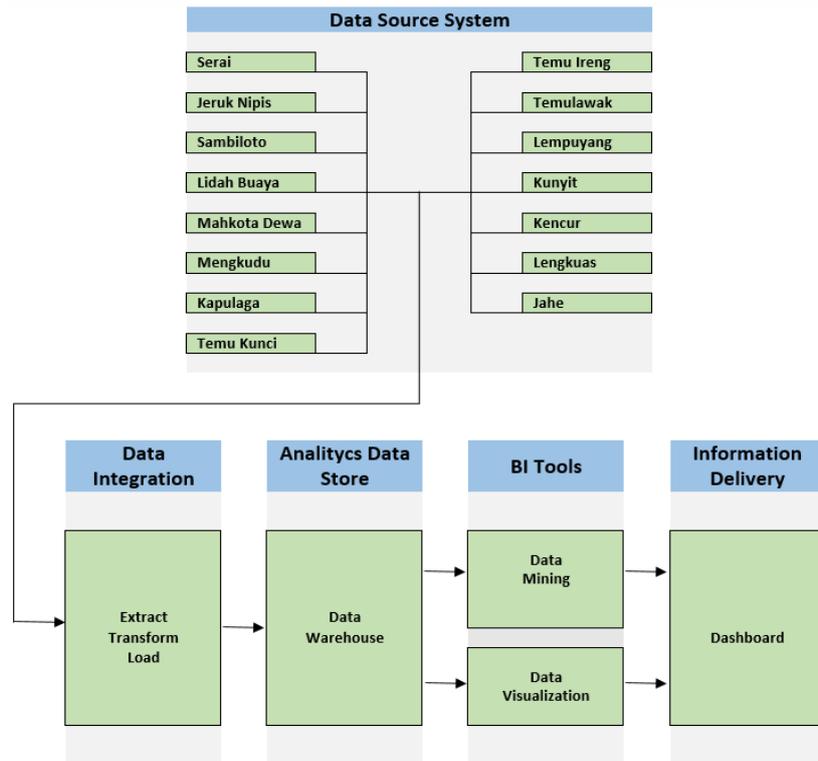
- Pemerintah
Informasi yang dihasilkan sistem intelijensia bisnis digunakan sebagai bahan evaluasi keberhasilan produksi di suatu wilayah; sebagai acuan dalam mengembangkan budidaya tanaman obat untuk meningkatkan hasil produksinya terutama untuk wilayah yang memiliki tingkat produksi rendah; serta sebagai acuan dalam mempromosikan potensi suatu wilayah.
- Produsen Obat
Sistem intelijensia bisnis menyajikan informasi sebagai acuan dalam pengambilan keputusan dalam menentukan wilayah sebagai sumber

pemasok bahan baku tanaman obat terutama industri obat tradisional.

- Produsen Pupuk
Informasi sistem intelijensia bisnis wilayah produksi tanaman obat dijadikan sebagai pertimbangan pada pengambilan keputusan dalam menentukan wilayah pemasaran dan distribusi pupuk.

Data pada visualisasi informasi bersumber dari data jenis produk, data wilayah produksi, data tahun produksi dan data jumlah produksi, sedangkan sumber data pada proses *data mining* untuk pengelompokan data wilayah produksi adalah data jumlah produksi seluruh jenis produk pada setiap wilayah.

Setelah data produksi berbagai tanaman obat terkumpul, dilanjutkan dengan proses *Extract, Transform, Load* (ETL). Keakuratan dan relevansi *Business Intelligence & Analytics* bergantung pada tingginya mutu data pada data warehouse yang bersumber dari internal dan eksternal menggunakan proses ETL (Souibgui *et al.*, 2019).



Gambar 2. Arsitektur Sistem Intelijensia Bisnis

Kettle ETL *tool* digunakan pada proses *Extract, Transform, Load*. Kettle adalah aplikasi ETL *open source* yang sering digunakan sebagai ETL *tool*. Kettle merupakan seperangkat alat dan aplikasi yang memungkinkan manipulasi data dari berbagai sumber.

Sistem manajemen basis data digunakan untuk menyimpan data keluaran pada sistem *data warehouse*. MySQL adalah sistem manajemen basis data yang digunakan pada penelitian. Untuk menambah, mengakses, dan mengolah data yang tersimpan dalam basis data komputer diperlukan sistem manajemen basis data yang memiliki peran penting dalam komputasi, sebagai penggunaan mandiri, atau sebagai bagian dari aplikasi lain.

Setelah semua komponen *data warehouse* dibuat dan proses ETL dijalankan sehingga *data warehouse* terisi data dengan baik, selanjutnya dilakukan proses *data mining* menggunakan algoritma *K-Means clustering* dan pembuatan *dashboard* sebagai antarmuka visualisasi data. Proses *data mining* menggunakan perangkat lunak RapidMiner Studio, sedangkan pembuatan *dashboard* pada penelitian menggunakan perangkat lunak Power BI yaitu perangkat lunak yang digunakan untuk membuat visualisasi data secara interaktif pada intelijensia bisnis.

Informasi hasil produksi tanaman obat yang tampil pada *dashboard* untuk pengambilan keputusan disajikan dalam bentuk grafik yang beragam disertai tabel. Fitur filter juga disediakan pada *dashboard* yang berguna untuk memilah data supaya lebih

mudah mencari serta mengelompokan data berdasarkan keinginan pengguna. Fitur filter yang disediakan pada *dashboard* yaitu filter berdasarkan nama kota, nama produk, dan tahun produksi. Ikhtisar dari informasi yang tampil juga terdapat pada *dashboard*.

Data hasil produksi berbagai tanaman obat merupakan hasil dari proses bisnis yang telah dilakukan pada sektor pertanian pada wilayah kabupaten dan kota di provinsi Jawa Barat, yaitu tanaman serai, jeruk nipis sambiloto, lidah buaya, mahkota dewa, mengkudu, kapulaga, temu kunci, temu ireng, temulawak, lempuyang, kunyit, kencur, lengkuas dan jahe. Data hasil produksi diperoleh dari situs web <https://opendata.jabarprov.go.id> yang bersumber dari Dinas Tanaman Pangan dan Hortikultura Provinsi Jawa Barat. Tampilan situs web sebagai sumber data dapat dilihat pada Gambar 3 dan Gambar 4.

Penetapan *grain* pada *data warehouse* berhubungan dengan hal yang akan diwakili oleh tabel fakta pada setiap *row*. *Grain* ditetapkan sebelum menentukan dimensi atau fakta, karena syarat dimensi atau fakta yaitu harus selalu konsisten dengan *grain*. Pada setiap dimensi terdapat berbagai data yang direlasikan terhadap *grain* yang ditetapkan untuk digunakan sebagai acuan dalam membuat struktur basis data. *Grain* adalah data dari kandidat fakta yang bisa dianalisis (Khotimah dan Sriyanto, 2016). *Grain* pada perancangan *data warehouse* ini adalah produksi tanaman obat sedangkan dimensi yang digunakan adalah tahun, kota, dan produk.

id	kode_provinsi	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	produksi_jahe	satuan	tahun
217	32	JAWA BARAT	3201	KABUPATEN BOGOR	566868	KILOGRAM	2021
218	32	JAWA BARAT	3202	KABUPATEN SUKABUMI	5204954	KILOGRAM	2021
219	32	JAWA BARAT	3203	KABUPATEN CIANJUR	15176108	KILOGRAM	2021
220	32	JAWA BARAT	3204	KABUPATEN BANDUNG	447602	KILOGRAM	2021
221	32	JAWA BARAT	3205	KABUPATEN GARUT	13115111	KILOGRAM	2021
222	32	JAWA BARAT	3206	KABUPATEN TASIKMALAYA	1854378	KILOGRAM	2021
223	32	JAWA BARAT	3207	KABUPATEN CIAMIS	542779	KILOGRAM	2021
224	32	JAWA BARAT	3208	KABUPATEN KUNINGAN	627135	KILOGRAM	2021
225	32	JAWA BARAT	3209	KABUPATEN CIREBON	1212	KILOGRAM	2021
226	32	JAWA BARAT	3210	KABUPATEN MAJALENGKA	778110	KILOGRAM	2021

Gambar 3. Tampilan situs web data produksi jahe di Provinsi Jawa Barat (Dinas Tanaman Pangan dan Hortikultura Provinsi Jawa Barat, 2022)

id	kode_provinsi	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	produksi_kapulaga	satuan	tahun
3	32	JAWA BARAT	3203	KABUPATEN CIANJUR	2369487	KILOGRAM	2013
30	32	JAWA BARAT	3203	KABUPATEN CIANJUR	2762226	KILOGRAM	2014
57	32	JAWA BARAT	3203	KABUPATEN CIANJUR	4298867	KILOGRAM	2015
84	32	JAWA BARAT	3203	KABUPATEN CIANJUR	6555416	KILOGRAM	2016
111	32	JAWA BARAT	3203	KABUPATEN CIANJUR	7061983	KILOGRAM	2017
138	32	JAWA BARAT	3203	KABUPATEN CIANJUR	10226621	KILOGRAM	2018
165	32	JAWA BARAT	3203	KABUPATEN CIANJUR	10559874	KILOGRAM	2019
192	32	JAWA BARAT	3203	KABUPATEN CIANJUR	21288837	KILOGRAM	2020
219	32	JAWA BARAT	3203	KABUPATEN CIANJUR	41269972	KILOGRAM	2021

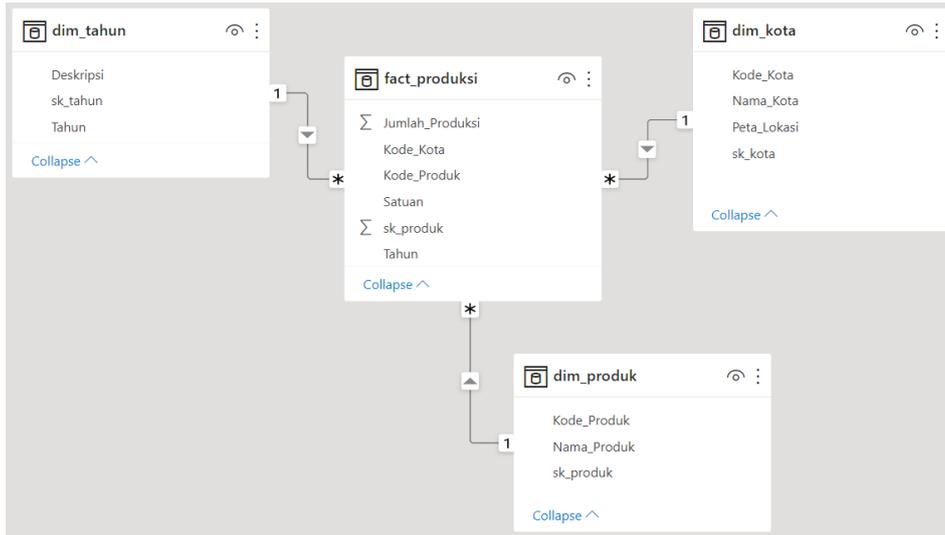
Gambar 4. Tampilan situs web data produksi kapulaga di kabupaten Cianjur (Dinas Tanaman Pangan dan Hortikultura Provinsi Jawa Barat, 2022)

Berdasarkan relasi dimensi dengan *grain* terbentuk tabel fakta. Tabel fakta dalam model dimensional menyimpan pengukuran kinerja yang dihasilkan dari peristiwa proses bisnis organisasi (Kimball dan Ross, 2013). Keterkaitan antara tabel fakta dengan tabel dimensi yang dirancang menggunakan *star schema* dapat dilihat pada Gambar 5.

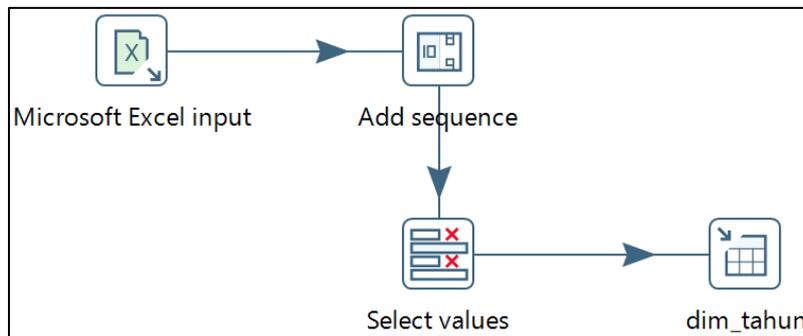
Setelah membuat rancangan tabel fakta dan tabel dimensi dilanjutkan dengan proses implementasi menggunakan Kettle *ETL tool* dan MySQL sebagai sistem manajemen basis data untuk penyimpanan data. Kettle *ETL tool* digunakan sebagai alat untuk menyusun rangkaian langkah transformasi proses *Extract, Transform* dan *Load* pada tabel fakta dan tabel dimensi.

Tabel *dim_tahun* terdiri dari 3 atribut yaitu *Sk_Tahun*, *Tahun* dan *Deskripsi*. Proses transformasi *dim_tahun* dimulai dari membaca *file* sumber yang berisi master periode dataset per tahun, dilanjutkan penambahan *sequence number*, pemilihan *fields* dan menulis data pada tabel *dim_tahun*. Alur transformasi *dim_tahun* dapat dilihat pada Gambar 6.

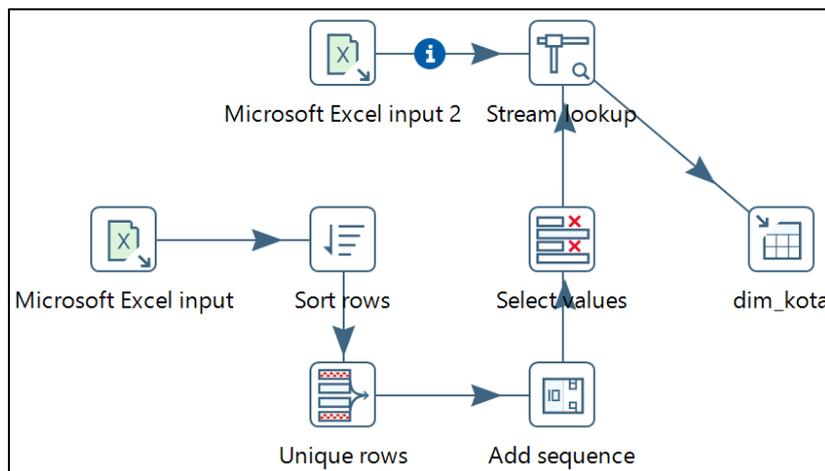
Tabel *dim_kota* terdiri dari 4 atribut yaitu *Sk_Kota*, *Kode_Kota*, *Nama_Kota* dan *Peta_Lokasi*. Proses transformasi *dim_kota* dimulai dari membaca *file* sumber yang berisi data kota dan data koordinat peta, dilanjutkan pembersihan duplikasi data, penambahan *sequence number*, pemilihan *fields*, penggabungan data kota dengan data koordinat peta serta menulis data pada tabel *dim_kota*. Alur transformasi *dim_kota* dapat dilihat pada Gambar 7.



Gambar 5. Diagram Bintang fact_ produksi



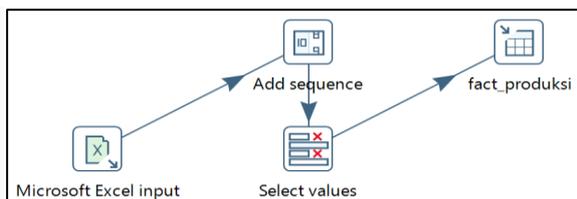
Gambar 6. Transformasi dim_tahun



Gambar 7. Transformasi dim_kota

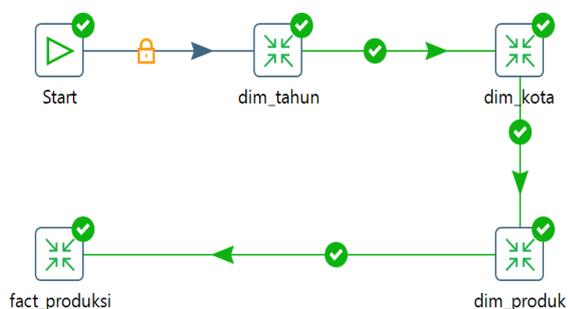
Tabel dim_produk terdiri dari 3 atribut yaitu Sk_Produk, Kode_Produk, dan Nama_Produk. Proses transformasi dim_produk dimulai dari membaca file sumber yang berisi data produk, dilanjutkan pembersihan duplikasi data, penambahan

sequence number, pemilihan fields, penggabungan data kota dengan data koordinat peta serta menulis data pada tabel dim_produk. Alur transformasi dim_produk dapat dilihat pada Gambar 8.



Gambar 8. Transformasi fact_produksi

Setelah seluruh tabel dimensi terbentuk dilanjutkan dengan transformasi tabel fact_produksi dengan file sumber yang berisi data produksi tanaman obat setiap wilayah kabupaten/kota di provinsi Jawa Barat per tahun. Alur transformasi fact_produksi dapat dilihat pada Gambar 9. Selain untuk menyusun langkah transformasi, Kettle ETL tool digunakan untuk menyusun Job yang menangani kendali atas aliran tugas yang terdiri atas beberapa transformasi. Setelah job produksi tanaman obat dijalankan, tabel fakta dan tabel dimensi akan terisi data hasil dari proses *Extract*, *Transform* dan *Load*. Informasi setiap langkah proses *Extract*, *Transform* dan *Load* yang dijalankan oleh Kettle ETL tool dapat dilihat pada tampilan *Execution Result*. Alur job transformasi dapat dilihat pada Gambar 9.



Gambar 9. Job Produksi Tanaman Obat

Terdapat permasalahan dalam pengembangan budidaya dan produksi tanaman obat, yaitu masalah Sumber Daya Manusia yang profesional, perhatian terhadap hasil-hasil penelitian ilmiah masih rendah, dan kurangnya dukungan pembiayaan untuk petani skala kecil (Salim dan Munadi, 2017). Pengelompokan wilayah produksi tanaman dilakukan agar proses pengembangan budidaya dan produksi tanaman obat tercapai secara optimal dengan efektif dan efisien.

K-Means adalah teknik pengelompokan yang banyak diterapkan. Dimulai dengan menentukan suatu titik pusat (*centroid*). Langkah berikutnya semua titik diarahkan menuju titik pusat terdekat dengan memperhatikan tingkat kedekatan. Setelah kelompok terbentuk, dilanjutkan dengan pembaruan

titik pusat pada setiap kelompok. Algoritma akan mengulangi langkah-langkah ini sampai tidak ada perubahan titik pusat (Aggarwal dan Reddy, 2014). Evaluasi kinerja algoritma pengelompokan *K-Means* terbaik berdasarkan Davies Bouldin Index (DBI) berdasarkan kuantitas dan kedekatan antar anggota cluster. Semakin kecil nilai DBI maka cluster yang dihasilkan semakin baik (Suarna *et al.*, 2021). Secara bertahap nilai *k* dimasukan dimulai dari 2, 3 dan 4 untuk menghitung nilai DBI. Hasil penghitungan DBI dapat dilihat pada Tabel 3.

Tabel 3. *Davies Bouldin Index* (DBI)

<i>K</i>	<i>Davies Bouldin Index</i>
2	0,570
3	0,255
4	0,493

Hasil penghitungan DBI menunjukkan bahwa nilai DBI terkecil adalah 0,255 pada *k* = 3 sebagai *k* optimal yang dipilih. Saat data produksi tanaman obat telah dimuat ke *data warehouse* melalui proses ETL, selanjutnya dilakukan proses *data mining* dengan algoritma *K-Means* clustering dan visualisasi data. Sebelum proses data mining dilaksanakan diawali dengan *data preprocessing* sebagai langkah untuk mempersiapkan data yang bersih dan siap digunakan pada proses selanjutnya. Hasil *data preprocessing* dimasukan kedalam RapidMiner Studio melalui proses *import data*. Sampel hasil proses *import data* dapat dilihat pada Tabel 4.

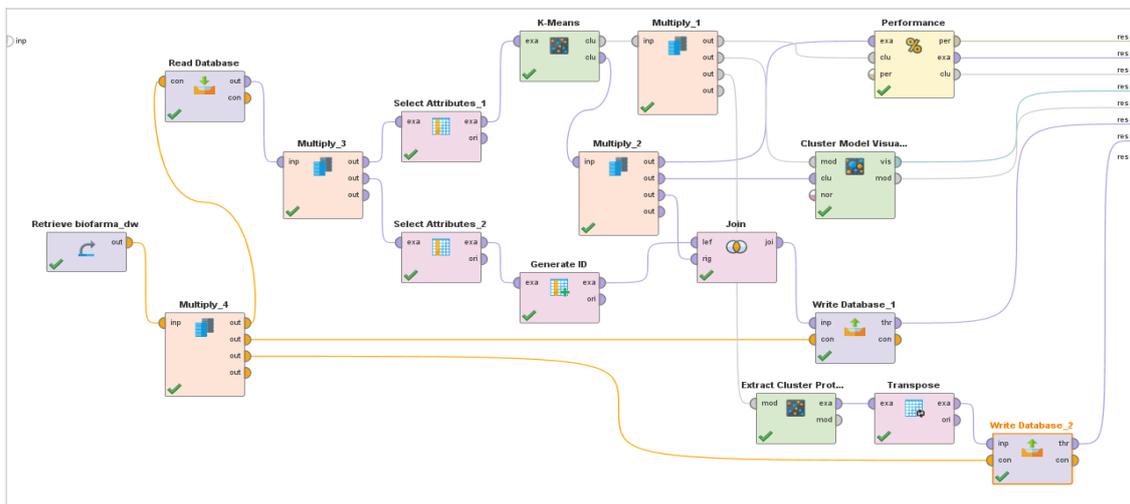
Pada perancangan proses *data mining* pada RapidMiner Studio menggunakan beberapa operator, diantaranya yaitu *Read/Write Database*, *Select Attributes*, *K-Means*, *Multiply*, *Cluster Distance Performance*, *Write Database*, *Cluster Model Visualizer* dan operator lainnya. Rangkaian operator pada RapidMiner Studio dapat dilihat pada Gambar 10.

Proses *K-Means clustering* pada *data mining* dilakukan untuk untuk mengelompokkan wilayah potensial penghasil tanaman obat dengan jumlah pengelompokan sebanyak 3 *cluster*. Proses *clustering* menghasilkan pengelompokan wilayah produksi tanaman obat yang tersebar di 3 *cluster*, jumlah wilayah pada masing-masing *cluster* adalah 24 wilayah berada pada *cluster* 0, 1 wilayah berada pada *cluster* 1, dan 2 wilayah berada pada *cluster* 2. Visualisasi hasil pengelompokan pada RapidMiner Studio dapat dilihat pada Gambar 11.

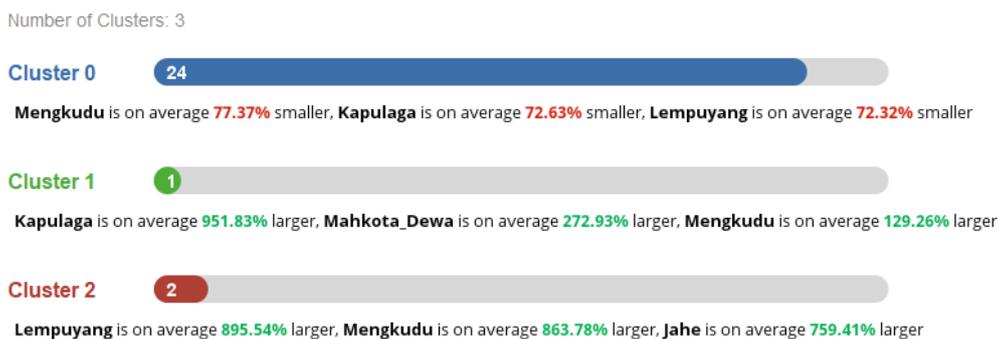
Pusat *cluster* akhir dihitung sebagai *mean* untuk setiap variabel dalam setiap *cluster* akhir. Pusat *cluster* akhir mencerminkan karakteristik untuk setiap *cluster* dapat dilihat pada Tabel 5.

Tabel 4. Sampel Data pada RapidMiner

Kode_Kota	Serai	Jeruk_Nipis	Sambiloto	Lidah_Buaya	Mahkota_Dewa	Mengkudu	Kapulaga	Temukunci
3201	2075340	221000	10128	722266	25812	5911	5591839	5037475
3202	405500	6000	4405	2092	277424	319302	7613519	35430
3203	278960	375300	20786	150518	9170838	20242554	106393283	684889
3204	36915	82000	9062	1658411	166983	99823	551965	39870
3205	0	0	0	3270	22440	75572	79030602	23220
3206	71193	965	277	26788	2572518	2416577	196748021	170
3207	143950	4415	0	28642	53552	522522	57584013	0
3208	0	1892	0	330	2476	5873	9490020	3390
3209	3000	0	1067	0	4620	8211	5783	1495
3210	0	0	0	200	14539	12770	461364	50610
3211	428800	145000	0	22000	1000	255700	321756	0
3212	1868	0	541	1352	1056	8877	0	4580
3213	2048500	10500	0	0	104	1857353	208509	515250
3214	2868	3150	9730	29718	400154	785708	9189143	549461
3215	6000	290	167108	908	5131910	302158	787018	5103225
3216	25300	0	0	2500	7500	19459	0	2032100
3217	2095440	266500	9470	63789	236691	203043	5841980	107852
3218	0	0	0	2349	249199	20368	24209246	0



Gambar 10. Rancangan Data Mining



Gambar 11. Visualisasi Jumlah Anggota Cluster

Tabel 5. Pusat *Cluster* Akhir

No	Jenis Tanaman	Cluster 0	Cluster 1	Cluster 2
1	Jahe	4900361.58	9365787.00	111125191.00
2	Jeruk Nipis	30898.54	965.00	187650.00
3	Kapulaga	5119569.00	196748021.00	92711942.50
4	Kencur	1567765.92	491276.00	6410899.50
5	Kunyit	2456002.13	1587981.00	28597458.50
6	Lempuyang	109575.33	176543.00	3940979.50
7	Lengkuas	3582231.51	1560903.00	9212757.50
8	Lidah Buaya	113969.83	26788.00	76894.00
9	Mahkota Dewa	285792.79	2572518.00	4596639.00
10	Mengkudu	238569.79	2416577.00	10159063.00
11	Sambiloto	11156.21	277.00	10393.00
12	Serai	303121.08	71193.00	139480.00
13	Temukunci	572270.42	170.00	354054.50
14	Temulawak	1104716.67	1123586.00	3305730.00
15	Temu Ireng	36731.17	10878.00	4335.00

Tabel 6. Anggota *Cluster*

No	Kota	Cluster	No	Kota	Cluster
1	Kab. Bogor	<i>cluster 0</i>	15	Kab. Karawang	<i>cluster 0</i>
2	Kab. Sukabumi	<i>cluster 0</i>	16	Kab. Bekasi	<i>cluster 0</i>
3	Kab. Cianjur	<i>cluster 2</i>	17	Kab. Bandung Barat	<i>cluster 0</i>
4	Kab. Bandung	<i>cluster 0</i>	18	Kab. Pangandaran	<i>cluster 0</i>
5	Kab. Garut	<i>cluster 2</i>	19	Kota Bogor	<i>cluster 0</i>
6	Kab. Tasikmalaya	<i>cluster 1</i>	20	Kota Sukabumi	<i>cluster 0</i>
7	Kab. Ciamis	<i>cluster 0</i>	21	Kota Bandung	<i>cluster 0</i>
8	Kab. Kuningan	<i>cluster 0</i>	22	Kota Cirebon	<i>cluster 0</i>
9	Kab. Cirebon	<i>cluster 0</i>	23	Kota Bekasi	<i>cluster 0</i>
10	Kab. Majalengka	<i>cluster 0</i>	24	Kota Depok	<i>cluster 0</i>
11	Kab. Sumedang	<i>cluster 0</i>	25	Kota Cimahi	<i>cluster 0</i>
12	Kab. Indramayu	<i>cluster 0</i>	26	Kota Tasikmalaya	<i>cluster 0</i>
13	Kab. Subang	<i>cluster 0</i>	27	Kota Banjar	<i>cluster 0</i>
14	Kab. Purwakarta	<i>cluster 0</i>			

Berdasarkan nilai pusat *cluster* menunjukkan hasil analisis pengelompokan menggunakan algoritma *K-Means* yaitu:

- *Cluster 0* berisi produksi mahkota dewa, mengkudu, kapulaga, temulawak, lempuyang, jahe yang rendah. Produksi jeruk nipis, kunyit, kencur, lengkuas yang sedang. Produksi serai, sambiloto, lidah buaya, temukunci, temu ireng yang tinggi.
- *Cluster 1* berisi produksi jeruk nipis, kunyit, kencur, lengkuas, serai, sambiloto, lidah buaya, temukunci yang rendah. Produksi mahkota dewa, mengkudu, temulawak, lempuyang, jahe, temu ireng yang sedang. Produksi kapulaga yang tinggi.

- *Cluster 2* berisi produksi temu ireng yang rendah. Produksi serai, sambiloto, lidah buaya, temukunci, kapulaga yang sedang. Produksi jeruk nipis, kunyit, kencur, lengkuas, mahkota dewa, mengkudu, temulawak, lempuyang, jahe yang tinggi.

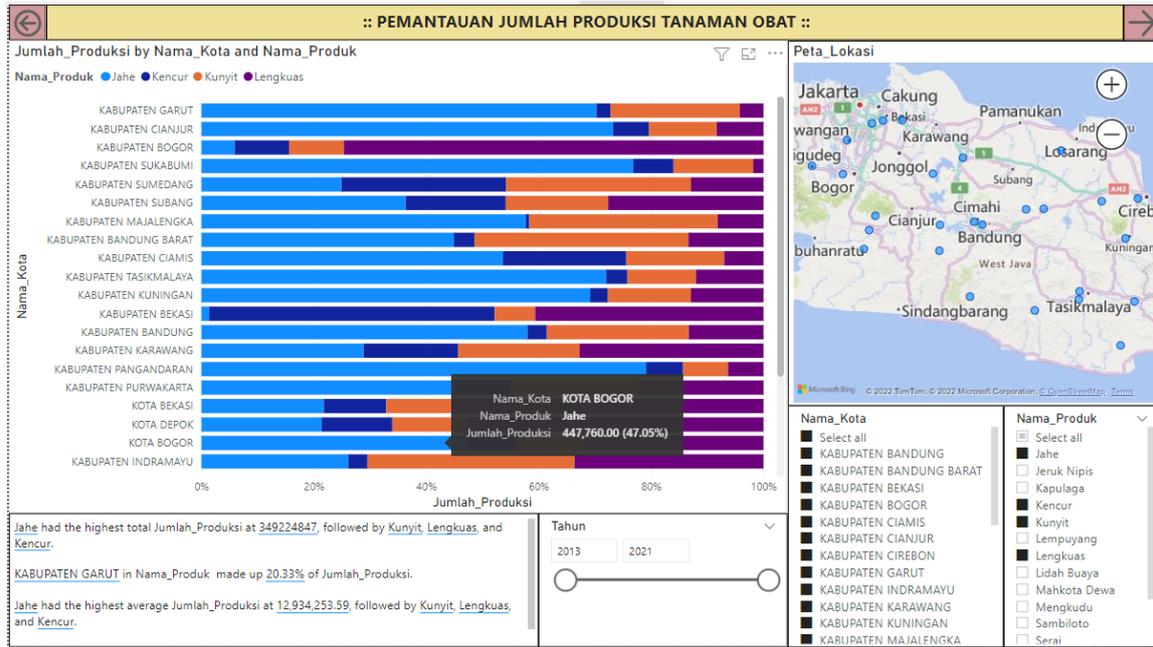
Anggota kelompok untuk masing-masing cluster yang terbentuk dapat dilihat pada Tabel 6. Antarmuka *dashboard* dirancang sesuai kebutuhan dan data pada *data warehouse* dibaca oleh Power BI. Informasi yang disajikan pada *dashboard* terdiri dari beberapa bagian yaitu informasi jumlah produksi berdasarkan nama kota dan nama produk dalam bentuk *stacked bar chart*; informasi jumlah produksi berdasarkan nama kota dan nama produk dalam

bentuk *clustered bar chart*; informasi jumlah produksi berdasarkan tahun dan nama produk dalam bentuk *donut chart*; informasi jumlah produksi berdasarkan tahun dan nama kota dalam bentuk *line chart*; informasi hasil proses pengelompokan wilayah produksi dengan menerapkan algoritma *K-Means* dalam bentuk *treemap* dan *line chart*.

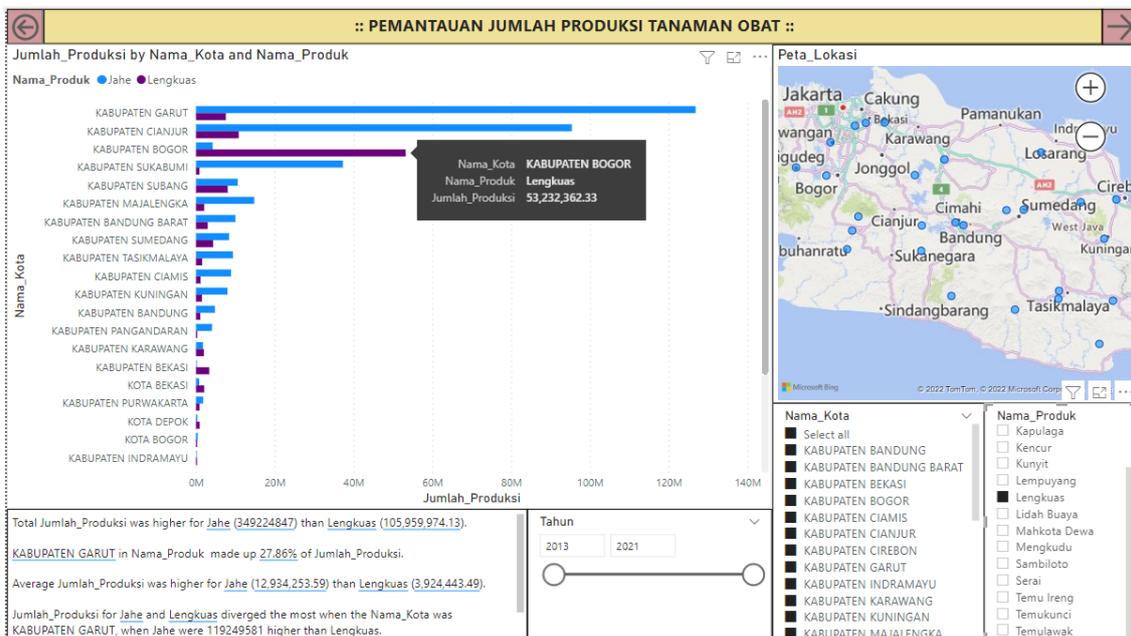
Selain menggunakan grafik, tampilan *dashboard* dilengkapi ikhtisar informasi, peta wilayah, fitur filter, dan disertai informasi yang rinci dalam bentuk tabel. Alasan penggunaan grafik dipilih

karena menampilkan data dalam bentuk grafik, lebih mudah bagi otak untuk memahami informasi dan memprosesnya (Campbell, 2021).

Tampilan *Stacked Bar Chart* pada Gambar 12 merupakan grafik dengan menggunakan batang untuk menggambarkan perbandingan antar kategori data, dengan membagi serta membandingkan seluruh bagian. Batang-batang pada grafik mewakili wilayah produksi, dan segmen pada batang mewakili jenis produk tanaman obat.



Gambar 12. *Stacked Bar Chart Dashboard*



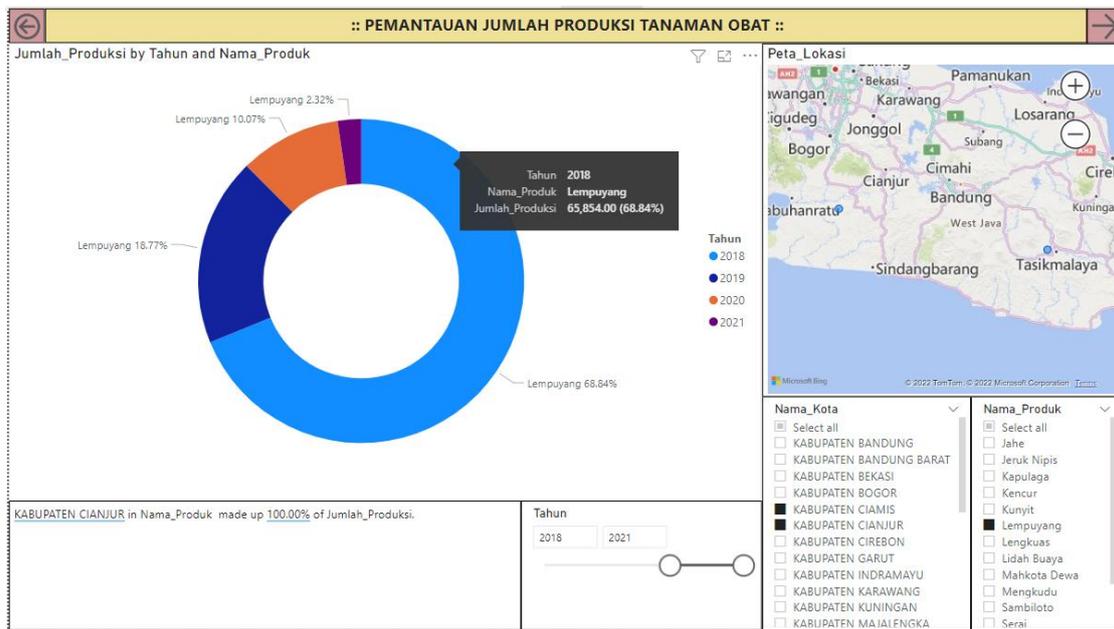
Gambar 14. *Clustered Bar Chart Dashboard*

Tampilan *Clustered Bar Chart* pada Gambar 13 merupakan grafik dengan menggunakan batang secara berkelompok yang berbeda ditempatkan berdampingan untuk membandingkan nilai pada semua kategori dengan menggunakan batang horizontal atau vertikal. Setiap kelompok batang pada grafik mewakili wilayah produksi, dan setiap batang pada kelompok mewakili jenis produk tanaman obat.

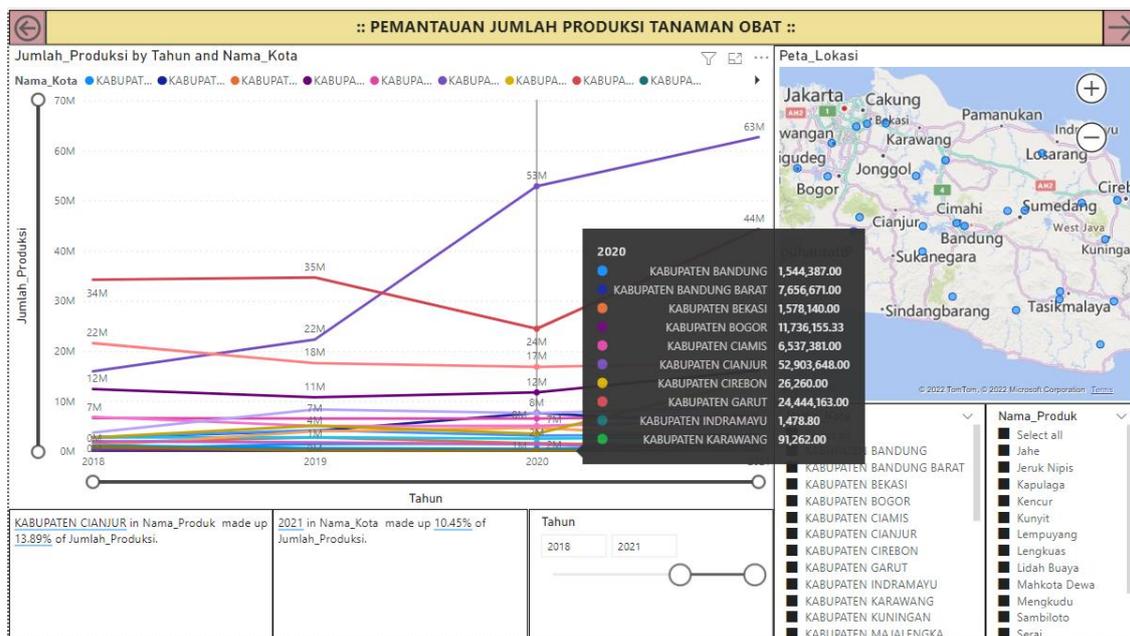
Tampilan *Donut Bar Chart* pada Gambar 14 merupakan grafik yang digunakan untuk menggambarkan perbandingan data berdasarkan kategori dengan ukuran setiap bagian merupakan

representasi proporsi masing-masing kategori. *Donut Bar Chart* digunakan untuk menunjukkan persentase jumlah produksi berdasarkan tahun dan jenis produk tanaman obat.

Tampilan *Line Chart* pada Gambar 15 merupakan grafik yang menggunakan garis untuk menyambungkan titik-titik data yang bernilai kuantitatif dalam rentang waktu tertentu sebagai representasi dari data historis. *Line Chart* digunakan untuk menampilkan jumlah produksi berdasarkan tahun dan wilayah produksi tanaman obat.



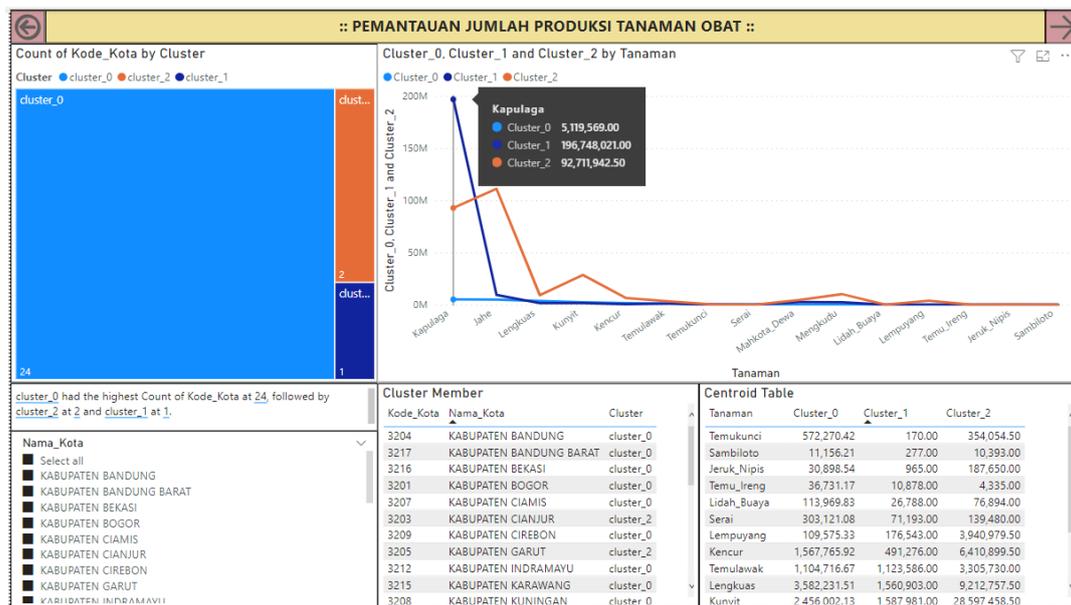
Gambar 15. Donut Chart Dashboard



Gambar 16. Line Chart Dashboard



Gambar 17. Table Dashboard



Gambar 18. Clustering Dashboard

Tampilan tabel pada Gambar 17 digunakan untuk menyajikan informasi dengan tampilan yang lebih rinci dalam penelusuran data. Hal ini memungkinkan untuk menggali informasi lebih dalam secara interaktif sesuai dengan kebutuhan pengguna yang belum tersajikan dalam grafik atau visualisasi data yang lain.

Hasil pengelompokan wilayah potensial penghasil tanaman obat menggunakan algoritma K-

Means pada proses data mining dapat diakses melalui Clustering Dashboard. Informasi yang ditampilkan pada dashboard adalah treemap dan tabel yang menunjukkan jumlah cluster dan anggotanya, serta line chart dan tabel yang menunjukkan nilai centroid setiap jenis produk di masing-masing cluster. Tampilan Clustering Dashboard dapat dilihat pada Gambar 18.

Evaluasi kinerja sistem dilakukan untuk memastikan sistem intelijensia bisnis yang dirancang

berjalan sesuai dengan ketentuan. Proses verifikasi dilakukan untuk memastikan proses pada sistem telah berjalan sesuai rancangan yang dibuat. Pada tahap ini diuji dengan dengan cara melakukan input pada sistem sehingga sistem menjalankan proses sampai sistem menghasilkan output, misal: input tahun produksi, pemilihan wilayah dan pemilihan jenis produk. Untuk memastikan sistem menghasilkan informasi yang sesuai dengan kebutuhan pengguna maka dilakukan proses validasi dengan menggunakan metode *black box*. Pada proses validasi dilakukan pengujian sistem intelijensia bisnis dengan cara menguji semua fitur sistem yang meliputi menu, grafik, ikhtisar, peta wilayah, fitur filter, tabel. Hasil pengujian menunjukkan sistem intelijensia bisnis berfungsi dengan baik dan menghasilkan informasi sesuai kebutuhan pengguna. Aspek kebutuhan sistem pada analisis PIECES dengan membandingkan fitur sistem yang berjalan dan hasil sistem usulan yang dirancang terpenuhi dengan baik.

KESIMPULAN DAN SARAN

Kesimpulan

Dataware house sebagai penyimpanan data yang terintegrasi dan menjadi sumber data pada sistem intelijensia bisnis telah berhasil diterapkan menggunakan model dimensional dengan skema bintang yang terdiri dari tiga tabel dimensi dan satu tabel fakta. Proses pengelompokan wilayah produksi tanaman obat menggunakan algoritma *K-Means clustering* berhasil dilakukan dengan k optimal = 3, jumlah elemen pada setiap *cluster* adalah 24 wilayah di *cluster* 0, 1 wilayah di *cluster* 1, dan 2 wilayah di *cluster* 2. Sistem intelijensia bisnis berhasil diterapkan dengan informasi pada *dashboard* untuk menunjukkan jumlah produksi berdasarkan wilayah dan jenis produk, menunjukkan persentase jumlah produksi berdasarkan tahun dan jenis produk, menampilkan jumlah produksi berdasarkan tahun dan wilayah produksi, serta menampilkan hasil pengelompokan wilayah potensial penghasil tanaman obat. Sistem intelijensia bisnis yang diterapkan telah melalui tahap evaluasi kinerja dan menghasilkan informasi secara ringkas, mudah, cepat dan akurat sesuai kebutuhan pengguna.

Saran

Penelitian lebih lanjut terkait dengan sistem intelijensia bisnis untuk memantau jumlah produksi tanaman obat dapat dikembangkan dengan merancang model arsitektur *data warehouse* yang disempurnakan melalui perbandingan hasil kinerja menggunakan metode lain serta memperkaya sumber data.

DAFTAR PUSTAKA

Aggarwal CC dan Reddy CK. 2014. *Data Clustering Algorithms and Applications*. CRC Press.

- Agustini F. 2017. Implementasi algoritma fuzzy C-means studi kasus penjualan di sushigroove restaurant. *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, 3(1), 127–132.
- Ahmar AS, Napitupulu D, Rahim R, Hidayat R, Sonatha Y, Azmi M. 2018. Using *K-Means clustering* to cluster provinces in Indonesia. *Journal of Physics: Conference Series*, 1028(1).
- Badan Pusat Statistik. 2021. *Ekspor Tanaman Obat, Aroma?k, dan Rempah-Rempah menurut Negara Tujuan Utama, 2012 - 2020*. <https://www.bps.go.id/statistable/2019/02/18/2019/ekspor-tanaman-obat-aromatik-dan-rempah-rempah-menurut-negara-tujuan-utama-2012-2020.html>
- Bordeleau FE, Mosconi E, dan de Santa-Eulalia LA. 2020. Business intelligence and analytics value creation in Industry 4.0: a multiple case study in manufacturing medium enterprises. *Production Planning and Control*. 31(2–3): 173–185.
- Campbell A. 2021. *Data Visualization Guide: Clear Introduction to Data Mining, Analysis, and Visualization*. Alex Campbell.
- Dinas Tanaman Pangan dan Hortikultura Provinsi Jawa Barat. 2022. *Data Produksi Tanaman Obat*. <https://opendata.jabarprov.go.id/>
- Dougherty J dan Ilyankou I. 2021. *Hands-on data visualization: interactive storytelling from spreadsheets to code*. O'Reilly Media, Inc.
- Falah F dan Hadiwibowo N. 2017. Species identification of traditional medicine plants for women's health in East Kalimantan: lesson learned from local wisdom. *Indonesian Journal of Forestry Research*, 4(1), 49–68.
- Fitriana R, Saragih J, dan Luthfiana N. 2017. Model business intelligence system design of quality products by using data mining in R Bakery Company. *IOP Conference Series: Materials Science and Engineering*. 277(1).
- Fitriana R, Saragih J, dan Firmansyah MA. 2016. Business intelligence system model proposals to improve the quality of service at PT GIA. *Proceeding of 9th International Seminar on Industrial Engineering and Management*, 2–10.
- Fitriana R, Saragih J, dan Hasyati BA. 2018. Perancangan model sistem intelijensia bisnis untuk menganalisis pemasaran produk roti di pabrik roti menggunakan metode data mining dan cube. *Jurnal Teknologi Industri Pertanian*, 28(1): 113–126.
- Fitriana R, Eriyatno, Djatna T, Kusmuljono BS. 2012. Peran sistem intelijensia bisnis dalam manajemen pengelolaan pelanggan dan mutu untuk agroindustri susu skala usaha menengah. *Jurnal Teknologi Industri Pertanian*. 22(3): 131–139.
- Khotimah K dan Sriyanto. 2016. Perancangan dan implementasi data warehouse untuk

- mendukung sistem akademik (Studi Kasus Pada STKIP Muhammadiyah Kotabumi). *Jurnal Teknologi Informasi Magister Darmajaya*. 2(01): 94–107.
- Kimball R dan Ross M. 2013. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition*. John Wiley & Sons, Inc.
- Lavalle A, Maté A, Trujillo J, Rizzi S. 2019. Visualization Requirements for business intelligence analytics: a goal-based, iterative framework. *Proceedings of the IEEE International Conference on Requirements Engineering*, 109–119.
- Mardalius dan Christy T. 2020. Mapping of Potential customers as a clothing promotion strategy using k-means clustering algorithm. *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, 6(1), 67–72. <https://doi.org/10.33480/jitk.v6i1.1414>
- MIT Database Notes. (n.d.). Department of Computer Science University of Cape Town. Retrieved May 5, 2022, from https://www.cs.uct.ac.za/mit_notes/database/pdfs/mit-database-notes.pdf
- Po L, Desimoni F, Bikakis N, Papastefanatos G. 2020. *Linked Data Visualization Techniques, Tools, and Big Data*. Morgan & Claypool.
- Purwaningsih E. 2019. Analisis kecelakaan berlalu lintas di kota jakarta dengan menggunakan metode K-Means. *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*. 5(1): 139–144. <https://doi.org/10.33480/jitk.v5i1.712>
- Rahayu SM dan Andini AS. 2019. Ethnobotanical study on medicinal plants in sesaot forest, narmada, West Lombok, Indonesia. *Biosaintifika: Journal of Biology & Biology Education*, 11(2), 234–242.
- Rahmadani N, Rahayu E, dan Lestari A. 2021. K-Means Clustering Areas Prone To Traffic Accidents in Asahan Regency. *Jurnal Ilmu Pengetahuan dan Teknologi Komputer*. 6(2):181–186. <https://doi.org/10.33480/jitk.v6i2.1519.K-MEANS>
- Salim Z, dan Munadi E. 2017. *Info Komoditi Tanaman Obat*. Badan Pengkajian dan Pengembangan Perdagangan Kementerian Perdagangan Republik Indonesia.
- Santoso HB, Sonia Putri PA, dan Oetomo BSD. 2017. Implementation of Sales executive dashboard for a multistore company in Yogyakarta. *International Journal of New Media Technology*, 4(1), 59–68. <https://doi.org/10.31937/ijnmt.v4i1.540>
- Sarno. 2019. Pemanfaatan tanaman obat (Biofarmaka) sebagai produk unggulan masyarakat desa Depok Banjarnegara. *Abdimas Unwahas*. 4(2): 73–78.
- Souibgui M, Atigui F, Zammali S, Cherfi S, Yahia, S. Ben. 2019. Data quality in ETL process: A preliminary study. *Procedia Computer Science*, 159, 676–687.
- Suarna N, Wijaya YA, Mulyawan, Hartati T, Suprapti T. 2021. Comparison K-Medoids algorithm and K-Means algorithm for clustering fish cooking menu from fish dataset. *IOP Conference Series: Materials Science and Engineering*, 1088(1). <https://doi.org/10.1088/1757-899X/1088/1/012034>
- Sucipto S, Al-Mubarak MS, dan Setiyawan DT. 2021. Hierarchical clustering bahan menu di kantin universitas untuk menunjang implementasi sistem jaminan halal. *Jurnal Teknologi Industri Pertanian*, 31(1), 103–124. <https://doi.org/10.1201/b19706>
- Sugiarto D, Mardianto I, Najih M, Adrian D, Pratama, DA. 2021. Perancangan dashboard untuk visualisasi harga dan pasokan beras di pasar induk beras cipinang. *Jurnal Teknologi Industri Pertanian*, 31(1), 12–19.
- Wilke CO. 2019. *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. 53 (9). O'Reilly Media, Inc.