**FINAL EXAM**

**ADVANCED BIG DATA ANALYTICS**

**IS 529-AL**

**NAMA : LIAN WIRA MANUEL MAHARAJA**

**NIM : 00000075938**

State the field of business you choose and give the company name **“Agricultural business/ EcoLianGrow ”** to answer the following questions.

Dengan Judul Artikel :

Link drive file Power BI : <https://drive.google.com/file/d/18pZJzKerKq935IxoH8zKZw7jgLC3Fz7/view?usp=sharing>

Link Report SAS Viya :

<https://vfl-072.engage.sas.com/links/resources/report?uri=%2Freports%2Freports%2F9165f025-a733-4250-b412-64a09e7e7f35>

Dataset yang dipilih berasal dari sumber [**https://katalog.data.go.id/**](https://katalog.data.go.id/) dengan judul **"Produksi Tanaman Pangan Menurut Kecamatan"**. Dataset ini berisi informasi tentang produksi tanaman pangan di Indonesia, khususnya per kecamatan, yang dapat digunakan untuk menganalisis pola produksi di wilayah yang berbeda.

**Judul artikel yang diangkat : Mengoptimalkan Produksi Pertanian melalui Business Intelligence: Pendekatan Analisis Big Data terhadap Produksi Tanaman Pangan Kawasan Tanggerang di EcoLianGrow**

### **Struktur dan Isi Dataset:**

Dataset ini berisi beberapa kolom utama yang dapat diolah lebih lanjut untuk analisis, antara lain:

1. **Kecamatan**: Nama kecamatan tempat produksi tanaman pangan dilakukan. Ini memberikan gambaran wilayah yang terlibat dalam produksi.
2. **Komoditas**: Jenis tanaman pangan yang diproduksi di setiap kecamatan, seperti padi, jagung, ubi jalar, dan komoditas pangan lainnya.
3. **Jumlah Produksi**: Jumlah total komoditas yang diproduksi (biasanya diukur dalam satuan ton atau kuintal). Kolom ini menjadi elemen utama dalam memahami kapasitas produksi tiap kecamatan.
4. **Tahun**: Tahun ketika data produksi tersebut dikumpulkan. Ini berguna untuk melakukan analisis tren produksi dari tahun ke tahun.
5. **Kode Wilayah** : Kode area pada suatu wilayah
6. **Satuan :** Satuan berat/bobot yang didapat

### **Kegunaan Dataset:**

Dataset ini sangat bermanfaat untuk bisnis pertanian seperti **EcoLianGrow**, terutama untuk:

1. **Analisis Kinerja Produksi**: Dengan melihat data jumlah produksi per kecamatan, EcoLianGrow dapat memahami daerah mana yang menghasilkan produksi tinggi dan mana yang perlu ditingkatkan.
2. **Perencanaan Strategis**: EcoLianGrow dapat memanfaatkan dataset ini untuk mengidentifikasi kecamatan yang memiliki potensi besar namun kurang dimanfaatkan, serta merencanakan distribusi sumber daya dengan lebih baik.
3. **Peluang Diversifikasi**: Dataset ini dapat membantu EcoLianGrow mengidentifikasi komoditas mana yang paling menguntungkan dan memberikan kesempatan untuk mengembangkan diversifikasi produk.
4. **Proyeksi Produksi Masa Depan**: Data historis dapat digunakan untuk membuat proyeksi produksi di masa depan dengan menggunakan metode analisis prediktif.

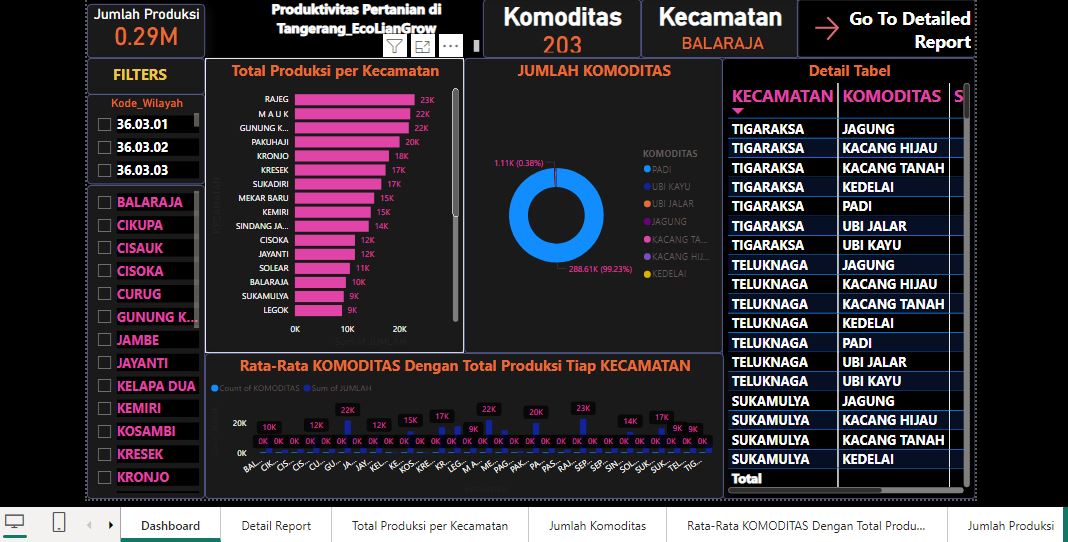
### **Manfaat Data untuk Business Intelligence:**

Dengan menggabungkan dataset ini ke dalam sistem **Business Intelligence** (BI), EcoLianGrow dapat mengoptimalkan hasil pertanian dengan cara:

* Melakukan **monitoring real-time** terhadap produksi.
* Mengidentifikasi **pola** dan **anomali** dalam produksi berdasarkan wilayah.
* **Membuat laporan** visual dan terperinci untuk memudahkan pengambilan keputusan.

1. **Create a representation of the Visual Analysis in the form of a visualization that contains some of the results of data analytics that you have implemented, by displaying the main page of your visual analytic menu so that other visual menus are visible on the main page.**

**Jawab :**

Dalam implementasi descriptive analytics di EcoLianGrow, terdapat beberapa pengukuran penting yang digunakan dan diterapkan. Pengukuran ini membantu dalam analisis dan evaluasi kinerja produksi perusahaan di berbagai kecamatan. Beberapa pengukuran yang digunakan antara lain:

1. Jumlah Produksi Total (Total Production)

Pengukuran ini menunjukkan jumlah total produksi di seluruh wilayah operasional EcoLianGrow, yang ditampilkan sebagai 0,29 juta ton. Pengukuran ini penting untuk memberikan gambaran umum tentang kapasitas produksi perusahaan secara keseluruhan.

1. Jumlah Komoditas (Number of Commodities)

Ini adalah pengukuran jumlah total komoditas yang diproduksi oleh EcoLianGrow. Saat ini, perusahaan memproduksi 203 komoditas, termasuk komoditas utama seperti padi, ubi kayu, dan kacang hijau. Pengukuran ini membantu dalam memahami diversifikasi produk perusahaan.

1. Produksi per Kecamatan (Production by District)

Pengukuran ini memperlihatkan jumlah produksi di setiap kecamatan, dengan kecamatan Rajeg dan M.A.U.K menjadi wilayah dengan hasil produksi tertinggi, masing-masing mencapai 22K ton. Pengukuran ini berguna untuk mengevaluasi performa setiap kecamatan dan menentukan wilayah yang perlu mendapatkan alokasi sumber daya tambahan.

1. Rata-Rata Produksi Komoditas per Kecamatan (Average Commodity Production by District)

Pengukuran ini menampilkan rata-rata produksi komoditas di setiap kecamatan, memungkinkan EcoLianGrow untuk menganalisis komoditas mana yang paling dominan di wilayah tertentu dan menyesuaikan strategi berdasarkan tren produksi.

1. Detail Report

Pengukuran ini memberikan rincian terperinci dari produksi setiap komoditas di masing-masing kecamatan. Data ini sangat penting untuk analisis lebih mendalam, membantu perusahaan mengidentifikasi komoditas mana yang memberikan hasil terbaik dan wilayah mana yang perlu ditingkatkan produksinya.

1. **Make a representation of visual analysis that can measure, monitor, and manage the business performance of your client, in accordance with the analysis of the preface section/ problem analysis from the research journal that you produce.**

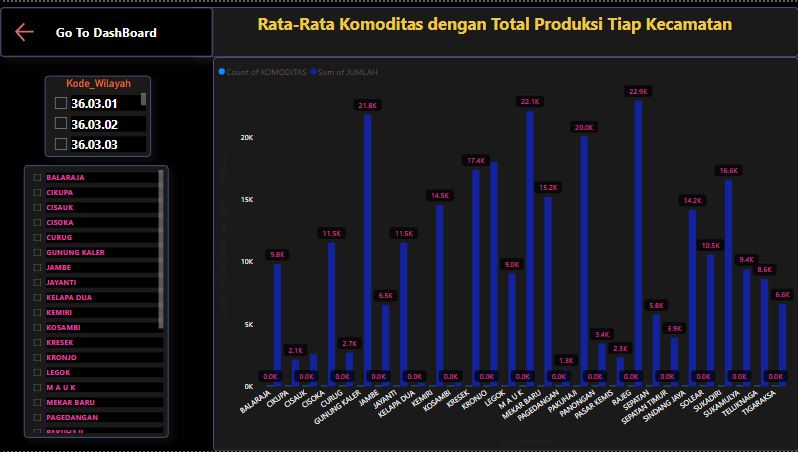
**Jawab:**

1. Menampilkan distribusi produksi yang merata di wilayah operasional perusahaan.BI memungkinkan EcoLianGrow mengukur efisiensi operasional berdasarkan data produksi yang terdistribusi di berbagai kecamatan, seperti yang terlihat dalam **Gambar tersebut** (Total Produksi per Kecamatan). Dengan BI, EcoLianGrow dapat memperbaiki strategi pengelolaan sumber daya dan meningkatkan produktivitas secara keseluruhan.



(Total Produksi per Kecamatan)

1. Menampilkan variasi produksi yang menunjukkan dampak cuaca terhadap hasil panen. **Gambar tersebut** (Rata-Rata Komoditas dengan Total Produksi Tiap Kecamatan) menunjukkan variasi hasil produksi yang dapat dipantau melalui BI. Dengan adanya prediksi berbasis BI, perusahaan bisa merencanakan langkah mitigasi, seperti memperkuat irigasi di area rawan kekeringan atau memindahkan produksi ke area yang lebih aman dari risiko cuaca.



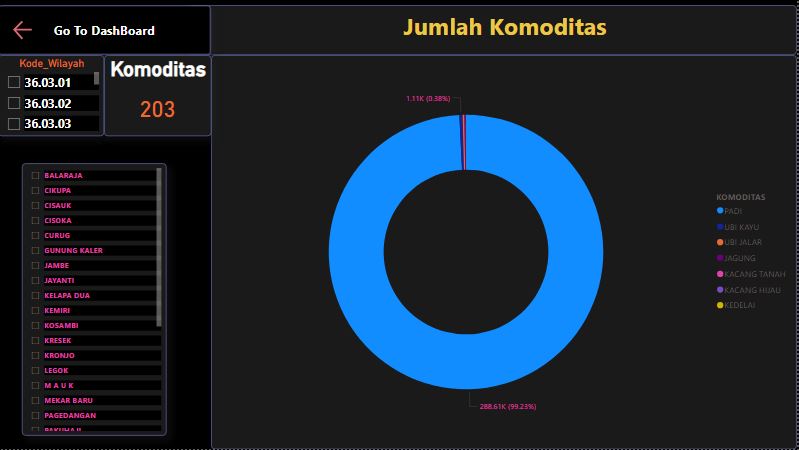
(Rata-Rata Komoditas dengan Total Produksi Tiap Kecamatan)

1. Menunjukkan fokus produksi di pasar domestik, dengan komoditas yang belum dipasarkan secara luas di luar negeri. **Gambar tersebut (Detail Report)** menunjukkan bahwa produksi EcoLianGrow masih berfokus pada pasar lokal. Dengan **BI**, perusahaan dapat menganalisis peluang ekspor dan memaksimalkan potensi untuk menjangkau pasar global.



(Detail Report)

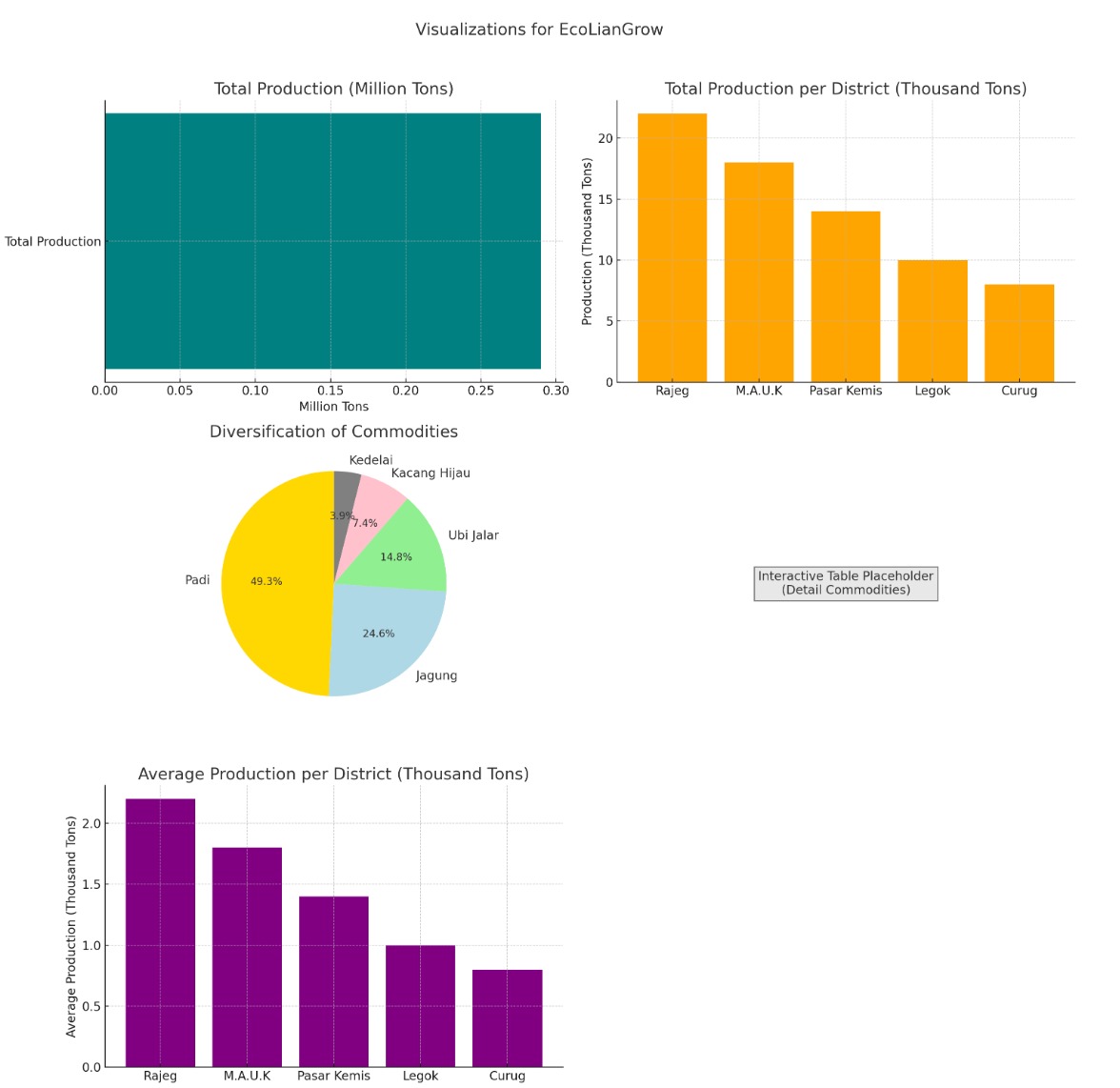
1. Menunjukkan potensi diversifikasi produk, termasuk produk organik. **Gambar tersebut**(Jumlah Komoditas) menunjukkan potensi diversifikasi produk ke arah produk organik. BI membantu EcoLianGrow memproyeksikan pertumbuhan pasar organik dan merencanakan strategi ekspansi yang optimal.

****

(Jumlah Komoditas)

1. **Create a representation of visual analysis that can provide decision making support for your client, in accordance with the analysis of the Literature Review and Methodology section from the research journal that you produce**

**Jawab :**

****

**Gambar Ilustrasi Visual yang dirancang dalam penelitian**

Fungsi analitik visual yang dirancang dalam penelitian ini bertujuan untuk memberikan dukungan pengambilan keputusan bagi EcoLianGrow melalui dashboard interaktif yang diimplementasikan menggunakan Power BI. Dashboard ini mencakup berbagai laporan visual yang dirancang untuk menganalisis data produksi pertanian secara mendalam, memberikan wawasan strategis, dan mendukung perencanaan operasional. Berikut adalah >3 jenis laporan yang dihasilkan:

1. **Laporan Jumlah Produksi Total**  
   Visualisasi berupa kartu (card) menampilkan jumlah produksi total EcoLianGrow, yaitu 0,29 juta ton. Laporan ini memberikan gambaran umum tentang kapasitas produksi perusahaan secara keseluruhan, membantu manajemen dalam memonitor kinerja secara real-time dan merencanakan strategi peningkatan produksi.

**Ilustrasi output yang diharapkan**



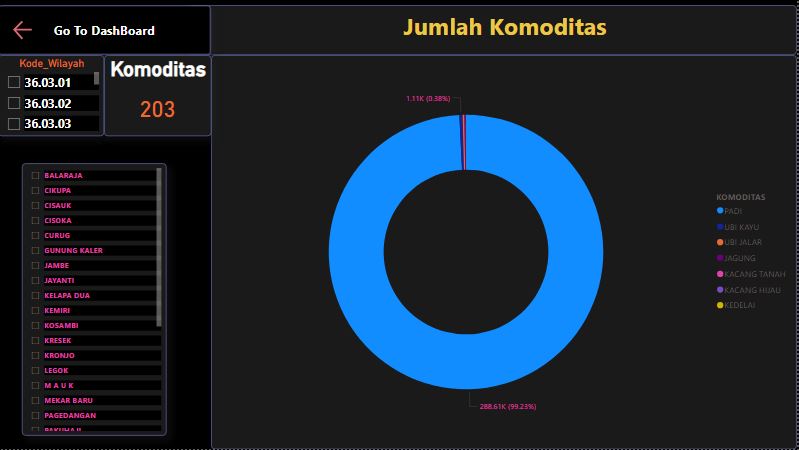
1. **Laporan Total Produksi per Kecamatan**  
   Visualisasi berupa diagram batang (bar chart) memperlihatkan distribusi produksi di setiap kecamatan, seperti Rajeg dan M.A.U.K yang mencatatkan produksi tertinggi. Laporan ini membantu perusahaan dalam menganalisis wilayah dengan performa terbaik dan mengidentifikasi kecamatan yang membutuhkan peningkatan sumber daya.

**Ilustrasi output yang diharapkan**



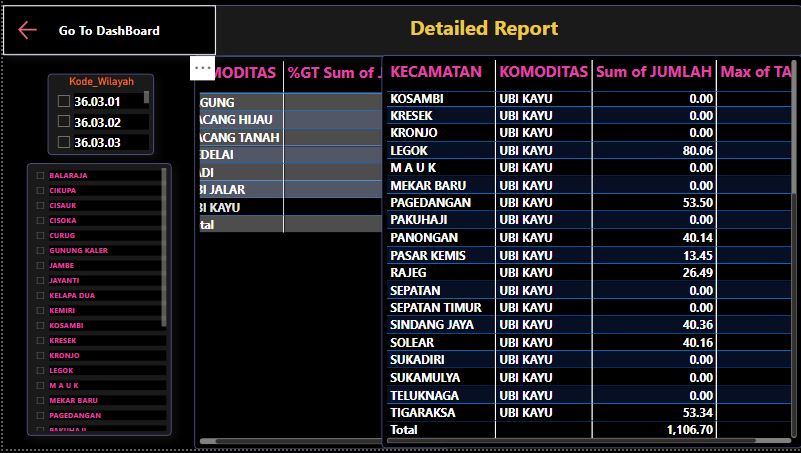
1. **Laporan Diversifikasi Komoditas**  
   Diagram donat (donut chart) menunjukkan jumlah dan proporsi berbagai jenis komoditas, seperti padi, jagung, dan ubi jalar. Dengan laporan ini, EcoLianGrow dapat memahami diversifikasi produknya dan mengidentifikasi peluang untuk meningkatkan diversifikasi sesuai tren permintaan pasar.

**Ilustrasi output yang diharapkan**



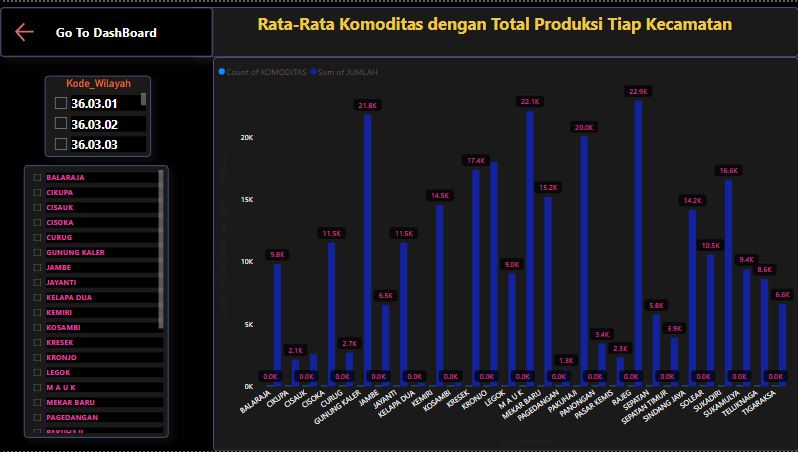
1. **Laporan Detail Komoditas**  
   Tabel interaktif (table visualization) menyajikan rincian data produksi setiap komoditas di masing-masing kecamatan. Laporan ini memungkinkan perusahaan untuk menganalisis performa komoditas tertentu secara mendalam, menentukan alokasi sumber daya, dan merencanakan strategi distribusi yang lebih efektif.

**Ilustrasi output yang diharapkan**



1. **Laporan Rata-Rata Produksi Komoditas per Kecamatan**  
   Diagram kolom (clustered column chart) menampilkan rata-rata produksi komoditas di berbagai kecamatan, memberikan wawasan tentang wilayah yang dominan dalam memproduksi komoditas tertentu. Laporan ini membantu perusahaan menyesuaikan strategi produksinya untuk memaksimalkan hasil berdasarkan tren lokal.

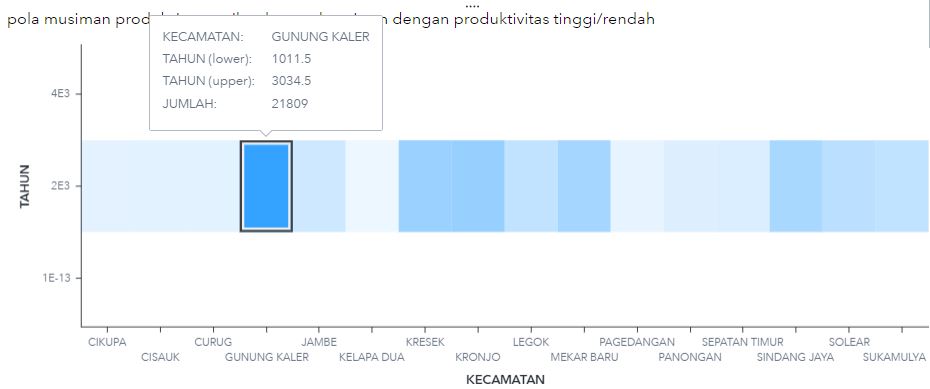
**Ilustrasi output yang diharapkan**



1. **Present a minimum of 3 types of visual analysis representations in the form of analytical results that you make have the following criteria:**
2. **Ability to uncover hidden patterns, which were not previously known or provide new insights/knowledge for the company**
3. **Interactive dashboard function, which can provide analytical reasoning facilitated by the interactive visual interface of the report presented. in accordance with Results and Discussion section of the research journal that you produce**

**Jawab :**

**Bagian 1 :** Heatmap Produksi Tanaman Pangan (Mengungkap Pola Tersembunyi di Wilayah dan Waktu).

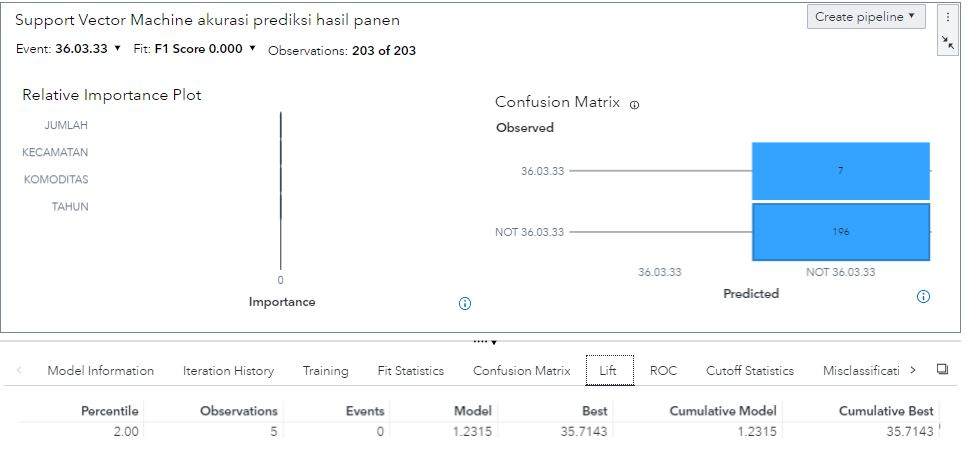
**Tujuan:**Untuk mengungkap pola produksi tanaman pangan berdasarkan wilayah (kecamatan) dan waktu (tahun).

#### **Langkah Pembuatan:**

1. **Buat Heatmap:**
   * **Sumbu X:** Tahun (waktu pengumpulan data).
   * **Sumbu Y:** Kecamatan.
   * **Warna:** Intensitas warna berdasarkan **Jumlah Produksi** (lebih gelap = produksi tinggi, lebih terang = produksi rendah).
2. **Tambahkan Filter:**
   * **Komoditas:** Filter interaktif untuk memilih jenis tanaman pangan (misalnya, padi, jagung, atau ubi jalar).
   * **Kode\_wilayah:** Filter wilayah tertentu.
3. **Fitur Interaktif:**
   * **Hover-over:** Menampilkan detail produksi untuk setiap kecamatan dan tahun.
   * **Drill-down:** Klik kecamatan tertentu untuk melihat data lebih detail pada tingkat komoditas.

#### **Insight yang Diberikan:** Heatmap dapat menunjukkan pola musiman produksi atau wilayah yang konsisten dengan produktivitas tinggi/rendah.

**Bagian 2** : SVM (



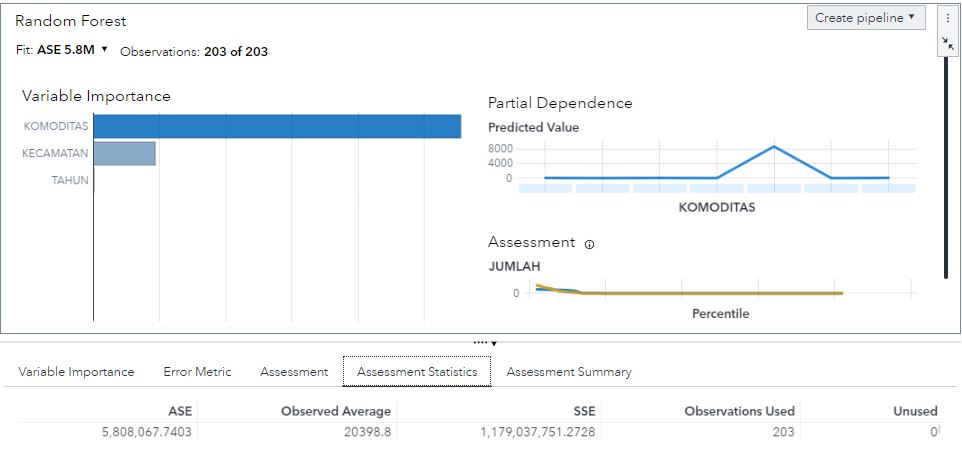
#### **Tujuan:** Untuk mengklasifikasikan tingkat produksi tanaman pangan di setiap kecamatan menjadi kategori seperti **Produksi Tinggi**, **Produksi Sedang**, dan **Produksi Rendah**.

#### **Langkah Pembuatan:**

1. **Bangun Model SVM:**
   * **Target:** Kategori produksi (misalnya, tinggi, sedang, rendah) berdasarkan **Jumlah Produksi**.
   * **Fitur:** Kode\_wilayah, Kecamatan, Komoditas, Tahun, Satuan.
2. **Visualisasi Hasil:**
   * Gunakan scatter plot untuk memvisualisasikan hasil klasifikasi.
   * **Sumbu X dan Y:** Gunakan dua fitur utama, misalnya, **Tahun** dan **Jumlah Produksi**.
   * **Warna:** Berikan warna berbeda untuk setiap kategori (misalnya, merah untuk rendah, biru untuk sedang, hijau untuk tinggi).
3. **Fitur Interaktif:**
   * Hover-over untuk melihat informasi kecamatan, komoditas, dan jumlah produksi di setiap titik.
   * Filter untuk memilih jenis komoditas atau wilayah tertentu.

#### **Insight yang Diberikan:**Mengidentifikasi kecamatan dengan produksi rendah, Memberikan informasi tentang kategori kecamatan berdasarkan kinerja produksinya.

Bagian 3 : Random Forest untuk Prediksi hasil panen



#### **Tujuan:** Untuk menganalisis faktor-faktor yang signifikan memengaruhi produksi tanaman pangan, seperti **Kode\_wilayah**, **Komoditas**, dan **Tahun**, menggunakan Random Forest.

#### **Langkah Pembuatan:**

1. **Bangun Model Random Forest:**
   * **Target:** Jumlah Produksi.
   * **Fitur:** Kode\_wilayah, Kecamatan, Komoditas, Tahun, Satuan.
2. **Visualisasi Pentingnya Fitur:**
   * Gunakan diagram batang horizontal untuk menampilkan tingkat kepentingan setiap fitur.
   * **X-Axis:** Pentingnya fitur.
   * **Y-Axis:** Nama fitur (urutan dari paling signifikan hingga paling rendah).
3. **Fitur Interaktif:**
   * Filter berdasarkan **Komoditas** untuk melihat bagaimana pentingnya fitur berubah untuk setiap tanaman pangan.
   * Filter berdasarkan **Kode\_wilayah** untuk fokus pada wilayah tertentu.

#### **Insight yang Diberikan:**

* Random Forest membantu mengidentifikasi fitur yang paling relevan untuk meningkatkan produksi.
* Misalnya, fitur **Komoditas** mungkin lebih signifikan di wilayah tertentu, sementara **Tahun** penting untuk pola musiman.

1. **Presenting visualization of analytic results that you produce as a recommendation for the business benefits/impact of the outcomes from the implementation of visual analyics, in accordance with conclusions section of the research journal that you produce**

**Jawab :**

****

**Heatmap Produksi Tanaman Pangan:**

* **Sumbu X:** Satuan
* **Sumbu Y:** Kecamatan
* **Warna:** Intensitas warna menunjukkan jumlah produksi / Frequency (lebih gelap = produksi tinggi, lebih terang = produksi rendah).

Dan kita melihat bahwa rata rata memiliki jumlah produksi yang hampir sama dalam setiap kecamatan / wilayah tersebut. Dan Penelitian ini menyoroti pentingnya pendekatan analitis berbasis data dalam meningkatkan efisiensi pengelolaan data besar dan akurasi prediksi hasil panen di sektor pertanian. Framework DCOVA terbukti efektif untuk pengelolaan data, dari pengumpulan hingga analisis, membantu identifikasi masalah dan strategi berbasis data (RQ1). Variabel "KOMODITAS" menjadi faktor utama yang memengaruhi produktivitas, diikuti "KECAMATAN," sementara "TAHUN" memiliki pengaruh lebih kecil (RQ2). Algoritma Random Forest unggul dalam identifikasi variabel signifikan, meskipun tingkat kesalahan masih tinggi, sedangkan SVM memiliki performa rendah karena ketidakseimbangan data (RQ3). Dashboard visual interaktif mempermudah pemangku kepentingan dalam menganalisis tren dan merancang strategi distribusi serta diversifikasi komoditas (RQ4). Pendekatan integratif ini terbukti meningkatkan efisiensi analisis dan pengambilan keputusan di sektor pertanian.

**REFERENSI :**

1. A. Sharma, B. Patel, dan C. D’Souza, “Diversification strategies for risk reduction in agriculture,” \*Journal of Agricultural Management\*, vol. 67, no. 2, pp. 203–215, Apr. 2020.
2. X. Li, dan H. Chen, “Big data analytics for agricultural efficiency improvement: Applications and challenges,” \*Smart Agriculture Journal\*, vol. 15, no. 1, pp. 95–110, May 2021.
3. L. Zhou, J. Wu, dan Y. Zhang, “The impact of extreme weather on agricultural production: Mitigation through technology,” \*Climate Risk Management\*, vol. 8, no. 3, pp. 121–133, Nov. 2019.
4. A. Kurniawan, “Challenges of accessing global markets for small agricultural enterprises,” \*Global Agricultural Economics Review\*, vol. 12, no. 3, pp. 145–158, Jul. 2019.
5. R. Johnson, “The rising demand for organic products in global markets,” \*Agricultural Economics Insights\*, vol. 19, no. 1, pp. 32–49, Jan. 2023.
6. L. Rahmawati, “Government policies for sustainable agriculture in Southeast Asia,” \*Journal of Environmental Policy Studies\*, vol. 23, no. 4, pp. 315–330, Dec. 2021.
7. R. Stevens, dan K. Lee, “Competitive pressures in agribusiness: How small players cope with large-scale competition,” \*Journal of Agribusiness Management\*, vol. 14, no. 2, pp. 65–78
8. A. Hartono, "Regulatory challenges in the agricultural sector and their impact on small enterprises," \*Journal of Agricultural Policy and Management\*, vol. 12, no. 3, pp. 145-160, Oct. 2019
9. A. Hartono, "Regulatory challenges in the agricultural sector and their impact on small enterprises," \*Journal of Agricultural Policy and Management\*, vol. 12, no. 3, pp. 145-160, Oct. 2019.
10. A. Patel dan R. Singh, "Global expansion strategies for organic products: A case study on small agribusinesses," \*Journal of Global Business Studies\*, vol. 25, no. 2, pp. 200-215, Apr. 2023.
11. T. Walker, "Smart agriculture: The role of sensors and irrigation technology in enhancing agricultural productivity," \*International Journal of Agricultural Technology\*, vol. 17, no. 1, pp. 85-100, Feb. 2022.
12. M. Nurdin, S. Rachman, dan H. Kartika, "Strategic partnerships for technological advancement in sustainable agriculture," \*Southeast Asia Agricultural Review\*, vol. 8, no. 2, pp. 45-60, May 2021.
13. L. Li, Z. Chen, dan H. Wang, "Diversification strategies for organic product markets: Profitability and competitiveness," \*Journal of Sustainable Agriculture\*, vol. 16, no. 3, pp. 95-115, Jun. 2020.