


## INFORMASI PERUSAHAAN

<b>Nama</b>	Kaiture (Kaizo Agriculture)
<b>Bidang</b>	Agricultural Business
<b>Deskripsi</b>	<p>Perusahaan Kaiture, singkatan dari Kaizo Agriculture, adalah entitas yang berdedikasi untuk menciptakan kondisi agrikultur yang lebih baik di masa mendatang. Dengan fokus pada keberlanjutan lingkungan, peningkatan produktivitas, dan kesejahteraan petani, Kaiture berkomitmen untuk mengembangkan metode pertanian berkelanjutan, menerapkan teknologi canggih dalam pertanian, memberikan pelatihan dan pendidikan kepada petani, serta berkolaborasi dengan berbagai pihak untuk memperkuat infrastruktur pertanian dan kebijakan yang mendukung pertanian berkelanjutan. Sebagai pemimpin di bidang agrikultur, Kaiture menjunjung tinggi nilai-nilai inovasi, keberlanjutan, dan pemberdayaan petani dalam upaya mencapai tujuan mereka.</p>
<b>Visi</b>	Membawa inovasi berkelanjutan dalam agrikultur melalui teknologi pintar dan data-driven decision-making untuk meningkatkan hasil dan kualitas produksi tanaman di seluruh wilayah.
<b>Misi</b>	<ul style="list-style-type: none"> <li>• Memanfaatkan data analitik untuk meningkatkan efisiensi produksi tanaman.</li> <li>• Mempromosikan praktik pertanian ramah lingkungan dengan penggunaan pupuk dan pestisida yang tepat.</li> <li>• Menyediakan solusi berbasis teknologi untuk membantu petani meningkatkan hasil pertanian.</li> </ul>
<b>Logo Perusahaan</b>	

## DATA PREPARATION

### 1. Pembacaan Dataset

```
kaiture = pd.read_csv("CROP.csv")
kaiture
```

	Crop	Crop_Year	Season	State	Area	Production	Annual_Rainfall	Fertilizer	Pesticide	Yield	o
0	Maize	1997	Kharif	Assam	19216	14721	2051.4	1828786.72	5956.96	0.615652	30
1	Maize	1997	Kharif	Karnataka	502797	1391132	1266.7	47851190.49	155867.07	2.687778	30
2	Maize	1997	Kharif	Meghalaya	17175	24878	3818.2	1634544.75	5324.25	1.444286	30
3	Maize	1998	Kharif	Assam	19810	13965	2354.4	1957228.00	5744.90	0.608696	30
4	Maize	1998	Kharif	Karnataka	450472	1542607	1260.8	44509533.60	130636.88	3.369000	30
...	...	...	...	...	...	...	...	...	...	...	...
19884	Ragi	2018	Winter	Odisha	27944	19356	1635.9	4532516.80	9780.40	0.538889	30
19885	Rapeseed & Mustard	2018	Winter	Odisha	6019	1657	1635.9	976281.80	2106.65	0.271429	30
19886	Sesamum	2018	Winter	Odisha	11056	2563	1635.9	1763263.20	3869.60	0.212778	30
19887	Sugarcane	2018	Winter	Odisha	6778	417672	1635.9	10969391.60	2372.30	57.584545	30
19888	Urad	2018	Winter	Odisha	39560	13123	1635.9	6419632.00	13846.00	0.352759	30

Langkah pertama dimulai dengan pembacaan dataset “CROP.csv” dari folder jupyter dengan menggunakan library pandas.

### 2. Penghilangan Data Kosong

```
kaiture.isnull().sum()

Crop                0
Crop_Year           0
Season              0
State               0
Area                0
Production          0
Annual_Rainfall     0
Fertilizer          0
Pesticide           0
Yield               0
o                   0
dtype: int64
```

Langkah kedua dilakukan pengecekan data yang kosong atau null untuk menghindari error pada saat melakukan visualisasi. Karena tidak terdapat kolom yang memiliki nilai null, maka penghapusan data null tidak diperlukan.

### 3. Penghapusan Kolom

```
kaiture = kaiture.drop(columns=['o'])
kaiture
```

	Crop	Crop_Year	Season	State	Area	Production	Annual_Rainfall	Fertilizer	Pesticide	Yield
0	Maize	1997	Kharif	Assam	19216	14721	2051.4	1828786.72	5956.96	0.615652
1	Maize	1997	Kharif	Karnataka	502797	1391132	1266.7	47851190.49	155867.07	2.687778
2	Maize	1997	Kharif	Meghalaya	17175	24878	3818.2	1634544.75	5324.25	1.444286
3	Maize	1998	Kharif	Assam	19810	13965	2354.4	1957228.00	5744.90	0.608696
4	Maize	1998	Kharif	Karnataka	450472	1542607	1260.8	44509533.60	130636.88	3.369000
...	...	...	...	...	...	...	...	...	...	...
19884	Ragi	2018	Winter	Odisha	27944	19356	1635.9	4532516.80	9780.40	0.538889
19885	Rapeseed & Mustard	2018	Winter	Odisha	6019	1657	1635.9	976281.80	2106.65	0.271429
19886	Sesamum	2018	Winter	Odisha	11056	2563	1635.9	1763263.20	3869.60	0.212778
19887	Sugarcane	2018	Winter	Odisha	6778	417672	1635.9	10969391.60	2372.30	57.584545
19888	Urad	2018	Winter	Odisha	39560	13123	1635.9	6419632.00	13846.00	0.352759

Langkah ketiga dilakukan penghapusan kolom pada dataset karena terdapat kolom yang tidak diperlukan yaitu kolom ‘o’

### 4. Pengubahan Format Data

```
kaiture['Year'] = pd.to_datetime(kaiture['Crop_Year']).as
kaiture.head()
```

	Crop	Crop_Year	Season	State	Area	Production	Annual_Rainfall	Fertilizer	Pesticide	Yield
0	Maize	1997	Kharif	Assam	19216	14721	2051.4	1828786.72	5956.96	0.615652
1	Maize	1997	Kharif	Karnataka	502797	1391132	1266.7	47851190.49	155867.07	2.687778
2	Maize	1997	Kharif	Meghalaya	17175	24878	3818.2	1634544.75	5324.25	1.444286
3	Maize	1998	Kharif	Assam	19810	13965	2354.4	1957228.00	5744.90	0.608696
4	Maize	1998	Kharif	Karnataka	450472	1542607	1260.8	44509533.60	130636.88	3.369000

Langkah keempat dilakukan perubahan format kolom ‘Crop\_Year’ ke dalam kolom baru ‘Year’ dengan format baru yaitu date/time. Dengan begitu, dataset tersebut dapat dipakai untuk melakukan visualisasi tren atau visualisasi lainnya yang membutuhkan keterangan waktu.

### 5. Pengelompokkan Data 1

```
q1 = kaiture['Production'].quantile(0.25)
q2 = kaiture['Production'].quantile(0.58)

kaiture['Production_Level'] = kaiture['Production'].apply(
    lambda x: 'Rendah' if x <= q1 else 'Sedang' if x <= q2 else 'Tinggi'
)

kaiture.head()
```

	Crop	Crop_Year	Season	State	Area	Production	Annual_Rainfall	Fertilizer	Pesticide	Yield	Year	Production_Level
0	Maize	1997	Kharif	Assam	19216	14721	2051.4	1828786.72	5956.96	0.615652	1997-01-01	Tinggi
1	Maize	1997	Kharif	Karnataka	502797	1391132	1266.7	47851190.49	155867.07	2.687778	1997-01-01	Tinggi
2	Maize	1997	Kharif	Meghalaya	17175	24878	3818.2	1634544.75	5324.25	1.444286	1997-01-01	Tinggi
3	Maize	1998	Kharif	Assam	19810	13965	2354.4	1957228.00	5744.90	0.608696	1998-01-01	Tinggi
4	Maize	1998	Kharif	Karnataka	450472	1542607	1260.8	44509533.60	130636.88	3.369000	1998-01-01	Tinggi

Langkah kelima dilakukan proses binning yang pertama untuk membuat kolom baru bernama ‘Production\_Level’ yang memiliki nilai kategorik yang telah terkonversi dari data numerik kolom ‘Production’. Kolom ini memiliki 3 kategori dimana ketika data memiliki nilai pada kuartil 1 maka production tersebut dapat dikategorikan sebagai ‘Rendah’, ketika nilai berada diatas kuartil 1 tetapi dibawah kuartil 2 maka production dikategorikan sebagai ‘Sedang’, dan ketika nilai berada diatas kuartil 2 maka production dikategorikan sebagai ‘Tinggi’.

### 6. Pengelompokkan Data 2

```
kaiture['Rainfall_Level'] = kaiture['Annual_Rainfall'].apply(
    lambda x: 'Low' if x <= 1000 else 'Medium' if x <= 3000 else 'High'
)

kaiture.head()
```

	Crop	Crop_Year	Season	State	Area	Production	Annual_Rainfall	Fertilizer	Pesticide	Yield	Year	Production_Level	Rainfall_Level
0	Maize	1997	Kharif	Assam	19216	14721	2051.4	1828786.72	5956.96	0.615652	1997-01-01	Tinggi	Medium
1	Maize	1997	Kharif	Karnataka	502797	1391132	1266.7	47851190.49	155867.07	2.687778	1997-01-01	Tinggi	Medium
2	Maize	1997	Kharif	Meghalaya	17175	24878	3818.2	1634544.75	5324.25	1.444286	1997-01-01	Tinggi	High
3	Maize	1998	Kharif	Assam	19810	13965	2354.4	1957228.00	5744.90	0.608696	1998-01-01	Tinggi	Medium
4	Maize	1998	Kharif	Karnataka	450472	1542607	1260.8	44509533.60	130636.88	3.369000	1998-01-01	Tinggi	Medium

Langkah keenam dilakukan proses binning yang kedua untuk membuat kolom baru bernama ‘Rainfall\_Level’. Kolom tersebut akan berisi 3 kategori dari nilai kolom ‘Annual\_Rainfall’ yang merupakan curah hujan. Kategori tersebut didasarkan dengan value ‘Annual\_Rainfall’ dimana ketika value dibawah 1000 maka kategorinya adalah ‘Low’, ketika value diatas 1000 dan dibawah 2000 kategorinya menjadi ‘Medium’, dan ketika value diatas 2000 dan dibawah 3000 kategorinya menjadi ‘High’.

# Penerapan Business Intelligence untuk Optimalisasi Hasil Pertanian Kaiture

Nathan Vilbert Kosasih

Department of Information System, Multimedia Nusantara University, Indonesia

[nathan.vilbert@student.umn.ac.id](mailto:nathan.vilbert@student.umn.ac.id)

Accepted on Oct 25, 2024

Approved on Oct 25, 2024

**Abstract** – Perusahaan Kaiture telah meningkatkan efisiensi dan akurasi produksi tanaman melalui penerapan berbagai jenis analitik dan teknik machine learning. Penelitian ini mengeksplorasi empat jenis analitik—deskriptif, diagnostik, prediktif, dan preskriptif—untuk memberikan wawasan mendalam mengenai pola produksi, ancaman, dan peluang. Analitik deskriptif mengungkapkan 55 jenis tanaman yang diproduksi dengan total luas hasil panen 1,34 juta hektar. Analitik diagnostik mengidentifikasi faktor-faktor yang mempengaruhi hasil panen, sementara analitik prediktif menunjukkan tren penurunan rata-rata luas lahan terpakai dan stabilitas hasil panen hingga 2030. Dalam ranah preskriptif, analisis membantu menentukan penggunaan optimal pupuk dan pestisida serta memetakan wilayah produksi terbaik di India. Untuk memperkuat hasil analitik, perusahaan menerapkan model machine learning seperti Decision Tree, Gradient Boosting, dan Random Forest, di mana Gradient Boosting mencapai akurasi tertinggi dengan C-Statistic 0,979. Hasil penelitian ini memberikan kontribusi signifikan dalam perumusan strategi berbasis data untuk meningkatkan efisiensi operasional dan keberlanjutan produksi tanaman di Kaiture.

**Keywords** – SWOT-i, Strategy Plan, Agrikultur, Analitik, Machine learning, Decision Tree, Gradient Boosting, Random Forest, Efisiensi Produksi

## I. INTRODUCTION

Pertumbuhan sektor pertanian di Indonesia terus dianggap sebagai hal yang sangat penting dalam pembangunan ekonomi secara keseluruhan [1]. Bahkan, sektor pertanian telah menjadi penopang utama ekonomi nasional karena pertumbuhannya yang positif, sementara sektor lain cenderung mengalami penurunan [2]. Menurut (Santoso, 2006), ada beberapa alasan yang mendasari pentingnya sektor pertanian di Indonesia: (1) Potensi sumber daya pertanian yang luas dan beragam, (2) Kontribusinya yang signifikan terhadap pendapatan nasional, (3) Besarnya jumlah penduduk yang berating pada sektor ini untuk kehidupan mereka, dan (4) Perannya sebagai basis pertumbuhan di pedesaan. Sektor pertanian di Indonesia terus mengalami peningkatan yang signifikan [3]. Data menunjukkan bahwa produksi padi meningkat sebesar 15% dalam

periode 5 tahun terakhir, sementara nilai ekspor produk pertanian melonjak sebesar 20% dalam periode 3 tahun terakhir [4]. Dengan adanya banyak inisiatif untuk meningkatkan produktivitas, keberlanjutan, dan kesejahteraan petani, sektor ini menjadi salah satu pilar utama dalam perekonomian negara, berkontribusi sebesar 13% terhadap PDB [5].

Kaiture, singkatan dari Kaizo Agriculture, adalah perusahaan yang berdedikasi untuk menciptakan kondisi agrikultur yang lebih baik di masa mendatang. Dengan fokus pada keberlanjutan lingkungan, peningkatan produktivitas, dan kesejahteraan petani, Kaiture berkomitmen untuk mengembangkan metode pertanian berkelanjutan dan menerapkan teknologi canggih dalam pertanian. Selain itu, Kaiture memberikan pelatihan dan pendidikan kepada petani untuk memperkuat kapasitas mereka. Perusahaan ini ingin mendapatkan pemahaman yang akurat mengenai faktor-faktor yang mempengaruhi kesejahteraan petani dan produktivitas pertanian. Untuk itu, Kaiture mengadopsi pendekatan berbasis data, yang memungkinkan mereka untuk menganalisis data secara holistik dan mengidentifikasi pola serta tren yang relevan dalam sektor pertanian.

Kaiture menggunakan berbagai metode untuk mendukung sektor pertanian, termasuk penerapan teknologi canggih dan program pendidikan bagi petani. Dalam upaya ini, Kaiture memanfaatkan dataset Agricultural Futures untuk menganalisis tren pasar dan faktor-faktor yang mempengaruhi produktivitas pertanian. Model analisis ini dirancang untuk mengolah data historis perdagangan komoditas pertanian dengan menggunakan algoritma dan teknik analisis canggih. Tujuannya adalah untuk mendapatkan pemahaman yang lebih mendalam tentang dinamika pasar, sehingga perusahaan dapat mengidentifikasi tren harga, volume perdagangan, dan faktor-faktor lain yang mempengaruhi kesejahteraan petani. Dengan model analisis yang akurat, Kaiture dapat mengoptimalkan strategi mereka dalam mendukung keberlanjutan dan produktivitas pertanian.

Kaiture menggunakan berbagai metode untuk mendukung sektor pertanian, termasuk penerapan teknologi canggih dan program pendidikan bagi petani.

Dalam upaya ini, Kaiture memanfaatkan dataset Crop Yield untuk menganalisis tren pasar dan faktor-faktor yang mempengaruhi produktivitas pertanian. Model analisis ini dirancang untuk mengolah data historis perdagangan komoditas pertanian dengan menggunakan algoritma dan teknik analisis canggih. Tujuannya adalah untuk mendapatkan pemahaman yang lebih mendalam tentang dinamika pasar, sehingga perusahaan dapat mengidentifikasi tren harga, volume perdagangan, dan faktor-faktor lain yang mempengaruhi kesejahteraan petani. Dengan model analisis yang akurat, Kaiture dapat mengoptimalkan strategi mereka dalam mendukung keberlanjutan dan produktivitas pertanian.

Diharapkan bahwa model analisis ini dapat memberikan manfaat yang signifikan bagi Kaiture. Dengan menggunakan Power BI untuk memahami dinamika pasar dan faktor-faktor yang memengaruhi produktivitas pertanian, perusahaan dapat mengoptimalkan strategi dalam mendukung kesejahteraan petani, mengelola sumber daya secara efisien, dan meningkatkan keberlanjutan pertanian. Dalam jurnal ini, akan dijelaskan proses pembuatan, pemilihan, dan penggunaan model analisis, serta menyajikan analisis yang akurat dan berguna untuk Kaiture. Selain metode tradisional, analisis data juga dilakukan menggunakan Power BI untuk mendukung visualisasi dan pemrosesan data yang lebih mendalam.

## II. QUESTION I SUB-CLO-2

### A. SWOT-i Matrix



Gambar 1. SWOT Analysis

SWOT Matrix adalah alat strategis yang digunakan untuk menganalisis kekuatan (Strengths), kelemahan (Weaknesses), peluang (Opportunities), dan ancaman (Threats) suatu organisasi atau proyek. Dengan mengidentifikasi faktor-faktor ini, organisasi dapat merumuskan strategi yang memanfaatkan kekuatan dan peluang, serta mengatasi kelemahan dan ancaman. Dalam konteks Kaiture, SWOT Matrix membantu memahami posisi perusahaan dalam industri pertanian dan meningkatkan pengambilan keputusan berbasis data melalui analitik. Ini mendukung pengembangan strategi yang lebih efisien dan responsif terhadap kondisi pasar. Dalam konteks Kaiture, analisis SWOT mengidentifikasi faktor-faktor kunci yang memengaruhi keberhasilan perusahaan dalam industri pertanian.

- Strengths (S):
  - S1: Total produksi tanaman per tahun yang

memungkinkan analisis tren dan pengambilan keputusan berbasis data, memberikan keunggulan dalam memahami pola produksi.

- S2: Identifikasi musim terbaik untuk produksi tanaman, yang meningkatkan efisiensi penanaman dan hasil panen.
- S3: Rata-rata luas lahan digunakan sebagai tolok ukur untuk perencanaan sumber daya yang lebih baik, membantu alokasi sumber daya yang efisien.
- Weaknesses (W):
  - W1: Keterbatasan data pada hubungan penggunaan pestisida dan pupuk dengan hasil produksi, yang dapat membatasi pemahaman dalam analisis produktivitas.
  - W2: Kesulitan dalam memprediksi curah hujan tahunan yang akurat, memengaruhi perencanaan pertanian.
  - W3: Keterbatasan pemetaan luas lahan panen secara efektif, yang dapat mengganggu pengelolaan lahan.
- Opportunities (O):
  - O1: Penggunaan data produksi tahunan untuk meramalkan tren pasar, membantu perusahaan untuk menyesuaikan strategi pemasaran.
  - O2: Pemanfaatan musim terbaik untuk meningkatkan hasil panen, yang dapat meningkatkan profitabilitas.
  - O3: Perencanaan sumber daya yang lebih baik dengan informasi luas lahan, memungkinkan pengambilan keputusan yang lebih informasional.
- Threats (T):
  - T1: Risiko cuaca ekstrem yang berdampak pada produksi tanaman, yang dapat mengurangi hasil.
  - T2: Perubahan iklim yang memengaruhi pola musim, mengancam kestabilan hasil pertanian.
  - T3: Fluktuasi hasil panen akibat musim yang tidak dapat diprediksi, meningkatkan ketidakpastian dalam produksi dan pemasaran.

Dengan analisis SWOT tersebut, dibuatlah SWOT-i Matrix yang merupakan perkembangan dari analisis SWOT. Matrix ini mengintegrasikan informasi dan data lebih lanjut untuk mengevaluasi interaksi antara kekuatan, kelemahan, peluang, dan ancaman. Melalui SWOT-i Matrix, Kaiture dapat mengidentifikasi strategi yang lebih komprehensif, memungkinkan perusahaan untuk merespons secara proaktif terhadap perubahan pasar dan meningkatkan efektivitas keputusan berbasis data. Proses ini tidak hanya membantu dalam perencanaan strategis tetapi juga dalam pemantauan kinerja secara berkelanjutan.

Tabel 1. SWOT-i Matrix Kaiture

	Opportunities (O)	Threats (T)
--	-------------------	-------------

Strengths (S)	S1O1: Pemanfaatan total produksi tanaman per tahun untuk meramalkan tren pasar dan meningkatkan strategi pemasaran.	S1T1: Data produksi dapat digunakan untuk mitigasi risiko terhadap cuaca ekstrem dan dampaknya.
	S2O2: Identifikasi musim terbaik untuk produksi tanaman dapat meningkatkan efisiensi penanaman dan hasil panen.	S2T2: Analisis musim yang tepat dapat membantu mengurangi dampak negatif dari perubahan iklim.
	S3O3: Rata-rata luas lahan yang digunakan dapat memfasilitasi perencanaan sumber daya yang lebih baik dan efisien.	S3T3: Pemahaman tentang luas lahan dapat membantu mengurangi kerugian akibat fluktuasi hasil panen.
Weaknesses (W)	W1O1: Analisis hubungan penggunaan pestisida dan pupuk dapat membantu meningkatkan hasil produksi yang berkelanjutan.	W1T1: Keterbatasan data yang tersedia dapat menghambat analisis hubungan antara input dan hasil produksi.
	W2O2: Memahami curah hujan dan analisisnya dapat membantu dalam pengambilan keputusan terkait waktu penanaman.	W2T2: Ketidakpastian dalam curah hujan dapat meningkatkan risiko gagal panen di berbagai wilayah.
	W3O3: Pemetaan luas lahan hasil panen secara efektif dapat meningkatkan efisiensi dan produktivitas pertanian.	W3T3: Perubahan musim yang tidak terduga dapat mempengaruhi hasil panen secara signifikan dan merugikan petani.

## B. Rencana Strategis

Dalam menghadapi tantangan dan peluang yang diidentifikasi melalui analisis SWOT-i Matrix, Kaiture perlu mengembangkan strategi bisnis yang tepat untuk memastikan efisiensi operasional dan keberlanjutan pertumbuhan. Berdasarkan kekuatan, kelemahan, peluang, dan ancaman yang ada, strategi ini akan mengoptimalkan pemanfaatan data produksi, memitigasi risiko terkait cuaca, dan memperbaiki perencanaan sumber daya. Dengan pendekatan berbasis data, perusahaan dapat meningkatkan efisiensi produksi, mengatasi ketidakpastian iklim, dan merespons perubahan pasar dengan lebih efektif, sebagaimana dirangkum dalam tujuh rencana strategis berikut:

- 1) **Memanfaatkan Total Produksi untuk Analisis Pasar**  
Kaiture dapat memanfaatkan data produksi tahunan untuk menganalisis tren pasar yang berkembang. Langkah ini akan dilakukan dengan mengumpulkan dan menganalisis data dari tahun-tahun sebelumnya, kemudian memanfaatkan software Business Intelligence (BI) untuk memvisualisasikan pola-pola yang ada. Implementasi strategi ini akan membantu

perusahaan dalam memahami kebutuhan pasar secara lebih baik, sehingga bisa meningkatkan strategi pemasaran yang lebih tepat sasaran dan meningkatkan keuntungan secara signifikan.

- 2) **Optimalisasi Musim Produksi**  
Dengan mengidentifikasi musim terbaik untuk setiap jenis tanaman, Kaiture dapat meningkatkan efisiensi dalam penanaman dan hasil panen. Langkah ini melibatkan analisis data historis terkait musim dan tingkat produksi di berbagai wilayah. Setelah pola musim terbaik ditemukan, jadwal penanaman akan diatur untuk memaksimalkan produktivitas. Implementasi yang tepat akan memberikan dampak berupa peningkatan efisiensi operasional dan hasil panen yang lebih tinggi pada musim yang optimal.
- 3) **Rencana Penggunaan Lahan yang Efektif**  
Kaiture dapat menggunakan informasi mengenai rata-rata luas lahan yang digunakan untuk perencanaan sumber daya secara lebih efisien. Dengan memetakan distribusi lahan menggunakan teknologi GIS, perusahaan akan mampu melakukan alokasi sumber daya yang lebih tepat, mengurangi pemborosan, dan meningkatkan produktivitas. Implementasi ini akan berdampak positif pada efisiensi keseluruhan operasional perusahaan dan pengelolaan lahan yang lebih baik.
- 4) **Peningkatan Analisis Penggunaan Pestisida dan Pupuk**  
Untuk meningkatkan hasil produksi yang berkelanjutan, Kaiture perlu mengumpulkan data terkait penggunaan pestisida dan pupuk serta menganalisis hubungannya dengan hasil produksi. Melalui analisis korelasi dan regresi, perusahaan dapat menemukan pola-pola yang meningkatkan efisiensi penggunaan input pertanian ini. Hasil implementasi strategi ini akan meningkatkan produktivitas lahan secara berkelanjutan dengan mengurangi penggunaan input yang tidak efektif.
- 5) **Prediksi Curah Hujan yang Akurat**  
Dengan ketidakpastian cuaca yang meningkat, Kaiture perlu mengembangkan model prediksi curah hujan berbasis data historis untuk meminimalkan risiko gagal panen. Kerjasama dengan lembaga meteorologi serta penggunaan algoritma machine learning akan membantu perusahaan dalam menghasilkan prediksi yang lebih akurat. Implementasi strategi ini akan mengurangi risiko terkait perubahan cuaca yang tidak menentu, sekaligus memperkuat ketahanan produksi.
- 6) **Mitigasi Risiko Cuaca Ekstrem**  
Kaiture dapat menyusun rencana kontinjensi untuk menghadapi situasi cuaca ekstrem yang



dapat mengancam hasil produksi. Dengan melatih tim dalam penerapan rencana ini dan memastikan ketersediaan sumber daya yang memadai, perusahaan dapat lebih siap menghadapi ancaman dari cuaca yang tidak menentu. Dampaknya adalah pengurangan potensi kerugian yang disebabkan oleh gangguan cuaca ekstrem dan peningkatan keberlanjutan produksi.

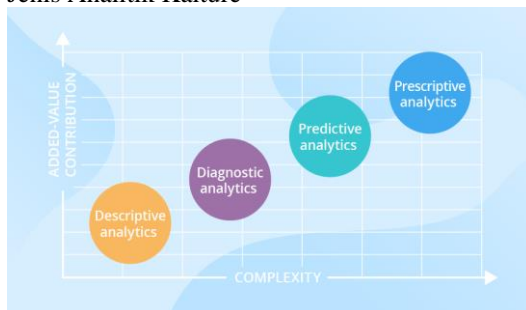
#### 7) Adaptasi terhadap Perubahan Iklim

Dalam menghadapi perubahan iklim yang terus berlangsung, Kaiture perlu melakukan analisis terhadap dampaknya pada pola produksi tanaman. Strategi adaptasi yang berbasis data akan dikembangkan untuk memastikan bahwa perusahaan dapat terus memproduksi meskipun di tengah perubahan lingkungan yang tidak dapat diprediksi. Implementasi langkah ini akan membantu Kaiture menjaga stabilitas dan keberlanjutan produksi dalam jangka panjang.

Dengan melaksanakan tujuh rencana strategis ini, Kaiture dapat meningkatkan efisiensi operasional dan respons terhadap perubahan pasar. Komitmen terhadap inovasi dan peningkatan kualitas juga akan memperkuat posisi Kaiture sebagai pemimpin di industri, memastikan pertumbuhan yang berkelanjutan dan daya saing yang tinggi ke depan.

### III. QUESTION II SUB-CLO-3

#### A. Jenis Analitik Kaiture



Gambar 2. Jenis-jenis Analitik

Analitik berperan penting dalam pengambilan keputusan di dunia pertanian yang semakin terdorong oleh teknologi dan data. Kaiture, sebagai perusahaan inovatif dalam agrikultur, menerapkan berbagai jenis analitik untuk memahami dan meningkatkan proses pertanian. Dengan memanfaatkan data dari berbagai sumber, Kaiture dapat mengidentifikasi tren, memahami faktor-faktor yang mempengaruhi hasil panen, serta memberikan rekomendasi yang tepat bagi petani. Dalam konteks ini, empat jenis analitik akan digunakan untuk membantu perusahaan mencapai tujuannya, yaitu analitik deskriptif, diagnostik, prediktif, dan preskriptif. Penerapan dari masing-masing jenis analitik pada perusahaan Kaiture dapat dijelaskan sebagai berikut:

#### 1) Analitik Deskriptif

Analitik deskriptif adalah jenis analitik yang digunakan untuk menggambarkan, meringkas, dan menjelaskan data yang telah dikumpulkan. Fokusnya adalah pada apa yang terjadi atau telah terjadi dalam suatu data. Dalam perusahaan Kaiture, analitik deskriptif digunakan untuk menyajikan dan menganalisis data historis terkait produksi tanaman, seperti total produksi per tahun, musim terbaik untuk penanaman, rata-rata luas lahan yang digunakan, dan distribusi geografis produksi tanaman. Tujuannya adalah untuk memberikan gambaran yang jelas tentang tren dan pola yang ada dalam data yang dikumpulkan. Pertanyaan dari analitik deskriptif ini bisa dijabarkan sebagai berikut:

- **Berapa total produksi tanaman per tahun untuk berbagai jenis tanaman?**

Tujuan: Menyajikan data jumlah produksi tahunan untuk berbagai jenis tanaman, sehingga dapat dianalisis tren produksi dari tahun ke tahun.

- **Musim apa yang paling banyak menghasilkan produksi tanaman?**

Tujuan: Mengidentifikasi musim utama di mana berbagai jenis tanaman diproduksi dalam jumlah tertinggi untuk membantu dalam perencanaan penanaman.

- **Berapa rata-rata luas lahan yang digunakan untuk penanaman berbagai jenis tanaman setiap tahunnya?**

Tujuan: Menampilkan distribusi rata-rata area lahan yang digunakan untuk penanaman berbagai jenis tanaman di setiap negara bagian.

- **Di wilayah mana saja sebuah tanaman diproduksi paling banyak?**

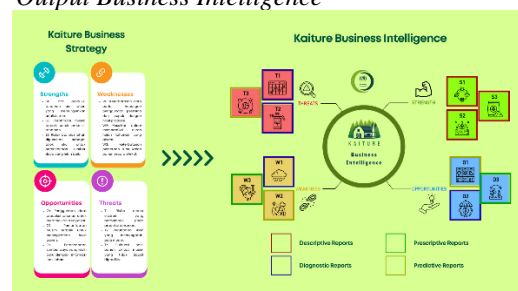
Tujuan: Menyajikan analisis distribusi geografis dari produksi berbagai jenis tanaman di seluruh negara bagian India.

#### 2) Analitik Diagnostik

Analitik diagnostik berfokus pada pemahaman alasan di balik suatu peristiwa atau hasil yang terjadi. Analitik ini melibatkan analisis data untuk menemukan pola, hubungan, atau penyebab dari masalah. Pada konteks perusahaan Kaiture, analitik diagnostik diterapkan untuk menganalisis hubungan dan korelasi antara berbagai faktor, seperti penggunaan pestisida dan pupuk, curah hujan tahunan, dan luas lahan dengan hasil produksi. Dengan melakukan analisis ini, Kaiture dapat memahami penyebab variabilitas dalam hasil produksi dan mengidentifikasi praktik terbaik yang dapat diterapkan. Pertanyaan dari analitik diagnostik dapat dijabarkan sebagai berikut:

- **Bagaimana hubungan antara penggunaan pestisida dan pupuk terhadap hasil produksi tanaman di wilayah tertentu?**  
Tujuan: Menganalisis korelasi antara penggunaan pestisida dan pupuk dengan hasil produksi tanaman untuk mengidentifikasi praktik terbaik.
  - **Apa hubungan antara tingkat curah hujan tahunan dan jumlah produksi tanaman di berbagai wilayah?**  
Tujuan: Memahami dampak curah hujan terhadap hasil produksi tanaman untuk menentukan pola yang dapat dijadikan acuan dalam pengambilan keputusan.
  - **Bagaimana luas lahan panen memengaruhi jumlah produksi tanaman di wilayah tertentu?**  
Tujuan: Mendiagnosis pengaruh luas lahan panen terhadap hasil produksi untuk menilai efisiensi penggunaan lahan.
  - **Sejauh mana musim mempengaruhi tingkat produksi tanaman di suatu wilayah?**  
Tujuan: Menganalisis pengaruh musim terhadap hasil produksi tanaman untuk membantu dalam perencanaan penanaman yang lebih baik.
- 3) **Analitik Prediktif**  
Analitik prediktif menggunakan teknik statistik dan algoritma pembelajaran mesin untuk memprediksi hasil di masa depan berdasarkan data historis. Analitik ini bertujuan untuk memberikan wawasan tentang apa yang mungkin terjadi. Pada perusahaan Kaiture, analitik prediktif digunakan untuk meramalkan tren masa depan, seperti hasil produksi tanaman, curah hujan, dan luas lahan yang akan digunakan. Analitik ini membantu Kaiture dalam perencanaan strategis dan pengambilan keputusan yang lebih baik berdasarkan data historis dan tren yang teridentifikasi.
- **Apa prediksi tren rata-rata hasil produksi tanaman untuk tahun berikutnya?**  
Tujuan: Menggunakan data historis untuk memperkirakan hasil rata-rata produksi tanaman di tahun mendatang, membantu perencanaan strategis.
  - **Bagaimana prediksi curah hujan di wilayah tertentu untuk tahun berikutnya?**  
Tujuan: Memprediksi pola curah hujan yang akan datang untuk menilai potensi dampaknya terhadap produksi tanaman.
  - **Apa prediksi luas lahan yang akan digunakan untuk panen di tahun berikutnya?**  
Tujuan: Memproyeksikan penggunaan lahan untuk panen di tahun mendatang untuk memfasilitasi perencanaan sumber daya.
  - **Bagaimana prediksi luas penggunaan lahan untuk tahun depan?**  
Tujuan: Mengestimasi perubahan dalam penggunaan lahan untuk membantu pengambilan keputusan terkait alokasi sumber daya.
- 4) **Analitik Preskriptif**  
Analitik preskriptif adalah jenis analitik yang memberikan rekomendasi atau tindakan yang perlu diambil untuk mencapai hasil yang diinginkan. Analitik ini menggunakan hasil analisis dari analitik deskriptif, diagnostik, dan prediktif untuk memberikan saran. Dalam konteks perusahaan Kaiture, analitik preskriptif digunakan untuk memberikan rekomendasi berdasarkan analisis yang dilakukan. Kaiture menggunakan pendekatan ini untuk menentukan negara bagian yang paling optimal untuk penanaman berbagai jenis tanaman, dengan mempertimbangkan faktor-faktor kesuburan tanah dan kondisi lingkungan lainnya. Pertanyaan dari analitik preskriptif dapat dijelaskan sebagai berikut:
- **Negara bagian mana yang paling optimal untuk penanaman berbagai jenis tanaman berdasarkan faktor-faktor kesuburan seperti curah hujan, luas lahan, jenis tanaman dominan, penggunaan pestisida dan pupuk, serta hasil produksi?**  
Tujuan: Memberikan rekomendasi lokasi optimal untuk penanaman tanaman berdasarkan analisis faktor kesuburan tanah dan kondisi lingkungan, serta menyarankan jenis tanaman yang sesuai untuk ditanam di wilayah tertentu.

#### B. Output Business Intelligence



Gambar 3. Output Business Intelligence Kaiture

## Kaiture Business Strategy



### Strengths

- S1: Total produksi tanaman per tahun yang memungkinkan analisis tren.
- S2: Identifikasi musim terbaik untuk produksi tanaman.
- S3: Rata-rata luas lahan digunakan sebagai tolak ukur untuk perencanaan sumber daya yang lebih baik.



### Weaknesses

- W1: Keterbatasan data pada hubungan penggunaan pestisida dan pupuk dengan hasil produksi.
- W2: Kesulitan dalam memprediksi curah hujan tahunan yang akurat.
- W3: Keterbatasan pemetaan luas lahan panen secara efektif.



### Opportunities

- O1: Penggunaan data produksi tahunan untuk meramalkan tren pasar.
- O2: Pemanfaatan musim terbaik untuk meningkatkan hasil panen.
- O3: Perencanaan sumber daya yang lebih baik dengan informasi luas lahan.

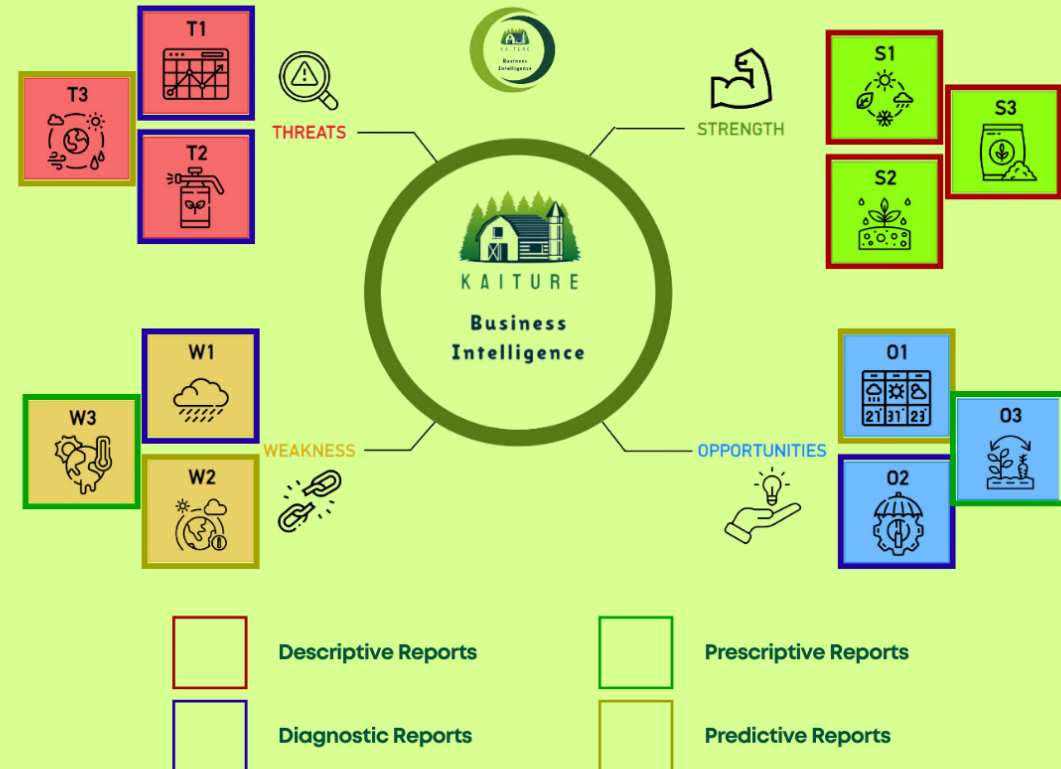


### Threats

- T1: Risiko cuaca ekstrem yang berdampak pada produksi tanaman.
- T2: Perubahan iklim yang memengaruhi pola musim.
- T3: Fluktuasi hasil panen akibat musim yang tidak dapat diprediksi.



## Kaiture Business Intelligence



PROSES PERUBAHAN BUSINESS STRATEGY MENJADI BUSINESS INTELLIGENCE PADA PERUSAHAAN KAITURE



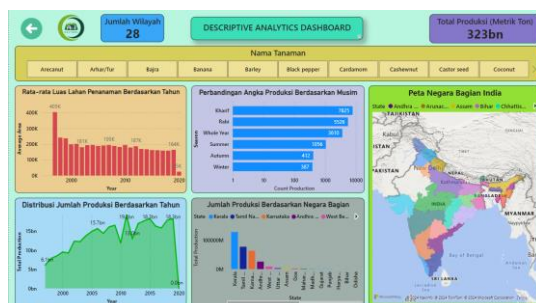
Kaiture Business Strategy menggunakan analisis SWOT (Strengths, Weaknesses, Opportunities, Threats) untuk menggambarkan strategi bisnis perusahaan. Bagian Strengths (Kekuatan), ditandai dengan warna biru muda, mencakup faktor-faktor seperti total produksi tahunan yang tinggi, standarisasi kualitas produk, dan ketersediaan sumber daya yang baik. Weaknesses (Kelemahan), yang ditandai dengan warna oranye, menggambarkan kelemahan seperti kesulitan dalam memprediksi curah hujan. Opportunities (Peluang), yang ditandai dengan warna merah muda, mencakup peluang seperti pemanfaatan musim terbaik untuk meningkatkan hasil panen. Terakhir, Threats (Ancaman), ditandai dengan warna ungu, menggambarkan ancaman seperti fluktuasi hasil panen akibat musim yang tidak dapat diprediksi.



Gambar 4. Kaiture Homepage

Kaiture Business Intelligence menunjukkan bagaimana elemen-elemen SWOT diterjemahkan ke dalam strategi intelijen bisnis melalui berbagai jenis laporan. Diagram ini mengelompokkan masing-masing elemen SWOT dengan berbagai jenis laporan untuk memberikan wawasan yang lebih mendalam. Berbagai jenis laporan tersebut dapat digambarkan sebagai berikut:

- 1) Laporan Deskriptif: Dalam gambar 3, laporan ini ditandai dengan batas kotak merah, digunakan untuk menggambarkan kekuatan. Visualisasi dashboard dari laporan deskriptif perusahaan Kaiture digambarkan sebagai berikut:



Gambar 5. Dashboard Analitik Deskriptif

Dashboard ini menampilkan visualisasi analitik deskriptif yang mencakup berbagai elemen, seperti rata-rata luas lahan penanaman, perbandingan

angka produksi, distribusi jumlah produksi, jumlah produksi per wilayah, peta negara bagian India, jumlah wilayah, serta total produksi. Slicer atau filter yang digunakan pada dashboard adalah berdasarkan jenis tanaman (crop), sehingga pengguna dapat menganalisis semua aspek tersebut sesuai dengan tanaman yang dipilih. Dengan adanya dashboard ini, pengguna dapat dengan mudah menganalisis:

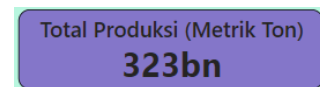
- a) Jumlah Wilayah yang Menanam Suatu Jenis Tanaman



Gambar 6. Jumlah Wilayah

Dalam menggambarkan jumlah wilayah yang menanam suatu jenis tanaman, digunakan card yang memakai variabel DistCount State yang menghitung berapa banyak negara bagian yang menanam sebuah jenis tanaman. Pada gambar 6, diperlihatkan bahwa jumlah negara bagian yang menanam semua jenis tanaman adalah 28 negara bagian.

- b) Total Produksi dari Suatu Jenis Tanaman



Gambar 7. Total Produksi (Metrik Ton)

Dalam menggambarkan total produksi sebuah tanaman dalam satuan metrik ton untuk berbagai jenis tanaman, digunakan card dengan memakai variabel measure Total Production yang menghitung jumlah produksi tanaman. Pada gambar 7, terlihat bahwa total produksi dari semua jenis tanaman dalam satuan metrik ton adalah sebesar 323 miliar tanaman.

- c) Rata-rata Luas Lahan Penanaman Suatu Jenis Tanaman berdasarkan Tahun Produksi

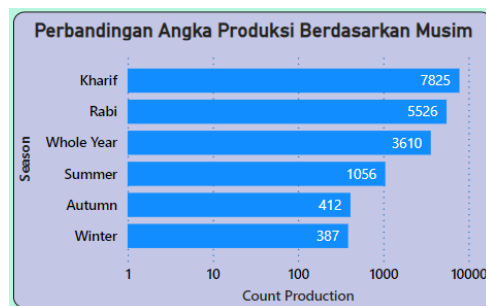


Gambar 8. Rata-rata Luas Lahan

Dalam menggambarkan rata-rata luas lahan yang dipakai untuk penanaman berbagai jenis

tanaman, digunakan stacked column chart dengan memakai variabel x-axis Year dan y-axis measure Average Area yang menghitung rata-rata dari luas lahan penanaman. Pada gambar 8, diperlihatkan perbandingan rata-rata luas lahan penanaman dari tahun 1997 sampai 2020. Berdasarkan grafik yang ditampilkan, terlihat bahwa rata-rata penggunaan luas lahan mencapai puncaknya pada tahun 1997, sebelum mengalami penurunan yang berkelanjutan hingga tahun 2020, di mana rata-rata luas lahan yang digunakan hanya mencapai sekitar 25 ribu hektar. Penurunan ini kemungkinan besar disebabkan oleh terbatasnya data yang tersedia pada tahun 2020, sehingga data terakhir yang dapat dijadikan acuan adalah tahun 2019, di mana rata-rata penggunaan luas lahan tercatat sebesar 164 ribu hektar.

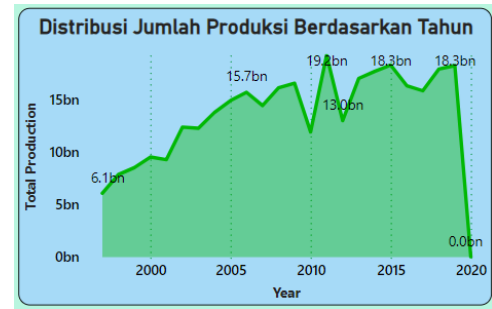
- d) Perbandingan Angka Produksi Suatu Jenis Tanaman berdasarkan Musim



Gambar 9. Perbandingan Angka Produksi

Untuk menggambarkan perbandingan angka produksi untuk berbagai jenis tanaman dalam musim yang berbeda, digunakan clustered bar chart yang memakai value y-axis season dan x-axis measure Count Production yang menghitung frekuensi produksi tanaman. Pada grafik dalam gambar 9, terlihat perbandingan angka produksi berdasarkan 6 musim yang berbeda. Angka produksi yang paling tinggi terdapat pada musim Kharif (hujan) yang memiliki frekuensi sebanyak 7,825 produksi, dan yang paling sedikit terjadi dalam musim winter (salju) yaitu sekitar 387 produksi.

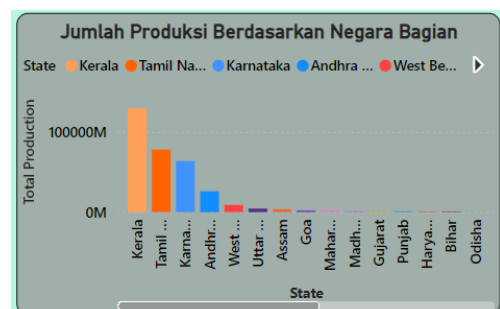
- e) Distribusi Jumlah Produksi Suatu Jenis Tanaman berdasarkan Tahun



Gambar 10. Distribusi Jumlah Produksi

Pada penggambaran distribusi jumlah produksi berbagai tanaman berdasarkan tahun, digunakan area chart yang memiliki x-axis year dan y-axis measure Total Production yang menghitung jumlah produksi secara keseluruhan. Pada gambar 10, diperlihatkan sebuah grafik yang memperlihatkan distribusi jumlah produksi semua tanaman dari tahun ke tahun dimana jumlah produksinya terus meningkat dari 1997 sampai pada tahun 2019. Penurunan drastis yang terlihat pada tahun 2020 dikarenakan oleh kurangnya data pada tahun tersebut.

- f) Perbedaan Jumlah Produksi Suatu Jenis Tanaman berdasarkan Negara Bagian



Gambar 11. Perbandingan Jumlah Produksi

Perbedaan jumlah produksi suatu jenis tanaman berdasarkan negara bagian yang berbeda dapat divisualisasikan menggunakan stacked column chart dengan x-axis state, y-axis measure Total Production, dan legend State untuk menandakan negara bagian yang berbeda. Pada gambar 11, terlihat stacked column chart yang memperlihatkan bahwa jumlah produksi terbanyak untuk semua jenis tanaman berada pada negara bagian Kerala yang jumlah produksinya melebihi 120 miliar.

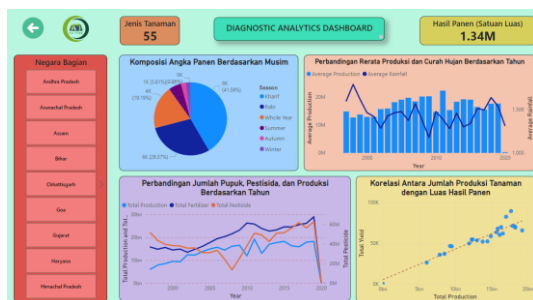
- g) Peta yang Menunjukkan Negara Bagian di India yang Memproduksi Jenis Tanaman Tertentu



Gambar 12. Peta Negara Bagian India

Untuk mempermudah visualisasi terhadap negara bagian di India yang memproduksi suatu jenis tanaman tertentu, digunakan map yang menunjukkan 28 negara bagian di India dengan warna yang berbeda-beda. Dalam gambar 12, terlihat peta yang menunjukkan warna-warna yang menandakan negara-negara bagian yang memproduksi seluruh tanaman yang berada dalam dataset.

- 2) Laporan Diagnostik: Dalam gambar 3, laporan ini ditandai dengan batas kotak biru, digunakan untuk mengidentifikasi ancaman, peluang, dan kelemahan. Visualisasi dashboard dari laporan diagnostik perusahaan Kaiture digambarkan sebagai berikut:



Gambar 13. Dashboard Analitik Diagnostik

Dashboard ini menampilkan visualisasi analitik diagnostik yang mencakup beberapa hal, seperti komposisi angka panen, jumlah jenis tanaman, hasil panen dalam satuan luas, perbandingan rata-rata produksi dan curah hujan, perbandingan jumlah pupuk, pestisida, dan produksi, serta korelasi antara jumlah produksi dengan luas hasil panen. Pada dashboard ini digunakan slicer atau filter yaitu negara bagian, agar pengguna bisa menganalisis aspek-aspek yang disebutkan berdasarkan negara bagian tertentu. Dengan adanya dashboard tersebut, pengguna dapat menganalisis:

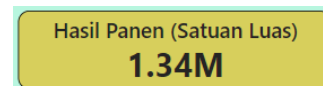
- a) Jumlah Jenis Tanaman yang Diproduksi pada Negara bagian Tertentu



Gambar 14. Jenis Tanaman

Untuk menggambarkan jumlah jenis tanaman yang berbeda pada berbagai negara bagian di India, digunakan card yang memiliki value measure DistCount Crop yang menghitung jumlah jenis tanaman tanpa duplikat. Pada gambar 14, terlihat bahwa jumlah jenis tanaman yang diproduksi pada semua negara bagian di India adalah berjumlah 55 jenis tanaman.

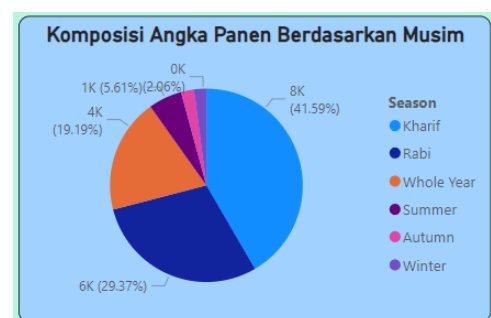
- b) Hasil Panen dalam Satuan Luas pada Negara Bagian Tertentu



Gambar 15. Hasil Panen (Satuan Luas)

Dalam memperlihatkan hasil panen dari semua tanaman dalam satuan luas pada negara bagian yang berbeda, digunakan visualisasi card dengan value measure Total Yield yang menghitung jumlah dari luas hasil panen. Dalam gambar 15, terlihat bahwa luas hasil panen dari seluruh negara bagian yang berbeda adalah sebesar 1.34 juta hektar.

- c) Komposisi Angka Panen berdasarkan Musim di Negara Bagian Tertentu

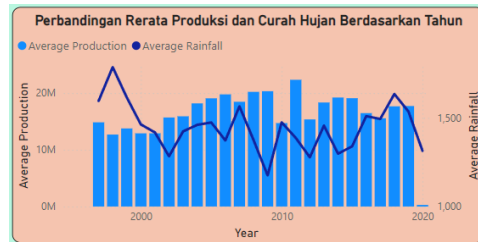


Gambar 16. Komposisi Angka Panen

Untuk memperlihatkan pengaruh musim terhadap frekuensi panen, dibuat visualisasi komposisi angka panen berdasarkan musim. Komposisi angka panen berdasarkan musim yang berbeda pada negara bagian tertentu digambarkan melalui pie chart yang menggunakan legend season (musim) dan values measure Count Yield yang menghitung angka frekuensi panen. Dari pie chart tersebut,

terlihat bahwa komposisi tertinggi dimiliki oleh musim Kharif dimana angka panen pada musim Kharif melebihi 8 ribu panen dengan persentase sebesar 41.59%. Di sisi lain, musim yang memiliki komposisi angka panen terkecil adalah musim winter (salju) yang memiliki angka panen dibawah seribu dan persentase sebesar 2.06%.

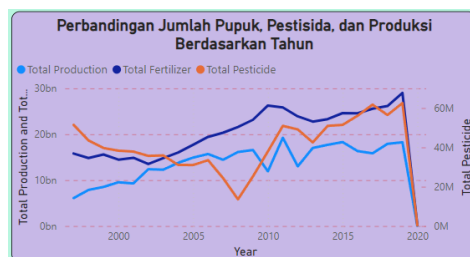
- d) Perbandingan Rata-rata Produksi dan Curah Hujan berdasarkan Tahun pada Negara Bagian



Gambar 17. Perbandingan Produksi dan Curah Hujan

Untuk memperlihatkan pengaruh curah hujan terhadap rata-rata produksi tanaman di sebuah negara bagian, digunakan line and clustered column chart yang menggunakan x-axis year, kolom y-axis measure Average Production, dan garis y-axis measure Average Rainfall. Pada gambar 17, dapat dilihat pengaruh curah hujan terhadap produksi dimana garis menggambarkan rata-rata curah hujan dan batang menggambarkan rata-rata produksi. Dari visualisasi tersebut, terlihat bahwa terdapat korelasi antara dua aspek tersebut walaupun tidak signifikan dimana ketika rata-rata curah hujannya menurun, maka rata-rata produksinya cenderung meningkat seperti pada tahun 2010. Dan ketika rata-rata curah hujannya meningkat, maka rata-rata produksinya menurun seperti pada tahun 1997 sampai tahun 2000.

- e) Pengaruh Jumlah Penggunaan Pupuk dan Pestisida Terhadap Jumlah Produksi Tanaman berdasarkan Tahun di Negara Bagian Tertentu

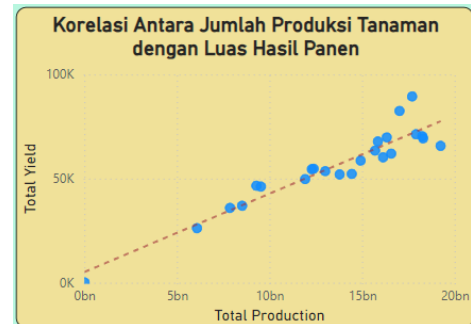


Gambar 18. Perbandingan Pupuk, Pestisida, dan Produksi

Untuk menggambarkan pengaruh penggunaan pupuk dan pestisida terhadap jumlah produksi

pada negara bagian yang berbeda, digunakan line chart dengan x-axis tahun dan 3 value y-axis yang berbeda yaitu measure Total Production, Total Fertilizer, dan Total Pesticide. Pada gambar 18, ditunjukkan visualisasi dari line chart yang dimaksud dimana terdapat 3 garis yang berbeda. Warna oranye mewakilkan pestisida, warna biru tua mewakilkan pupuk, dan warna biru muda mewakilkan produksi. Dari visualisasi tersebut, dapat dipahami bahwa jumlah penggunaan pupuk dan pestisida memiliki pengaruh, tetapi tidak signifikan terhadap jumlah produksi sebuah tanaman dimana garisnya memiliki pola yang mirip-mirip. Hampir ketiga garis tersebut memiliki puncak yang sama, yaitu pada tahun 2020.

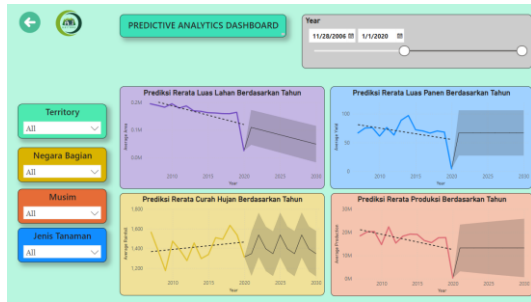
- f) Korelasi antara Jumlah Produksi Tanaman dengan Luas Hasil panen di Negara Bagian Tertentu



Gambar 19. Korelasi Produksi & Luas Panen

Untuk menggambarkan korelasi antara luas hasil panen terhadap produksi tanaman pada negara bagian yang berbeda, digunakan scatter plot dengan values Year (tahun), x-axis measure Total Production, dan y-axis measure Total Yield. Pada gambar 19, terlihat scatterplot yang menggambarkan titik yang mewakili setiap tahun dari 1997 sampai 2020. Dari visualisasi tersebut, terlihat bahwa terdapat korelasi yang signifikan antara luas hasil panen terhadap jumlah produksi dimana titik dari kedua aspek tersebut tersebar mengikuti garis tren yang menandakan terdapat hubungan positif yang kuat antara kedua hal tersebut.

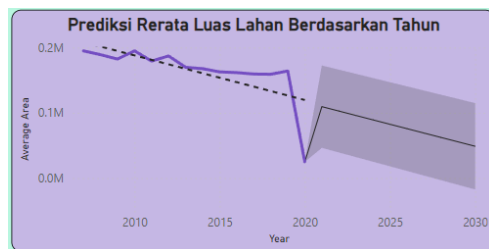
- 3) Laporan Prediktif: Dalam gambar 3, laporan ini ditandai dengan kotak kuning, digunakan untuk memprediksi peluang, kelemahan, dan ancaman yang akan datang. Visualisasi dashboard dari laporan prediktif perusahaan Kaiture digambarkan sebagai berikut:



Gambar 20. Dashboard Analitik Prediktif

Dashboard ini menampilkan visualisasi analitik prediktif yang memuat beberapa hal, seperti prediksi rata-rata luas lahan terpakai, prediksi rata-rata luas hasil panen, prediksi rata-rata curah hujan, dan prediksi rata-rata jumlah produksi. Keenam visualisasi tersebut diwakili dengan 5 slicer yaitu tahun, territory, negara, musim, dan jenis tanaman. Dengan 5 slicer tersebut, pengguna dapat memvisualisasikan prediksi luas lahan, luas hasil panen, curah hujan, dan jumlah produksi berdasarkan 5 aspek yang bisa mereka atur sebebasnya. Dashboard ini dapat membantu pengguna dalam menganalisis:

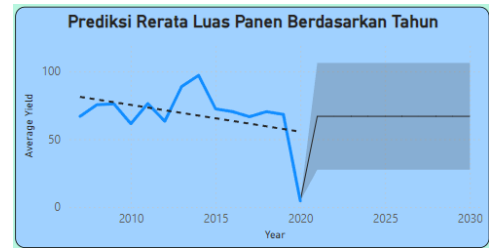
- Prediksi Rata-rata Luas Lahan Terpakai berdasarkan Tahun



Gambar 21. Prediksi Luas Lahan

Dalam memprediksi rata-rata luas lahan yang digunakan dalam menanam dari tahun ke tahun, digunakan line chart yang memiliki forecast dan trendline dengan value x-axis year dan y-axis measure Average Area. Gambar 21 diatas memperlihatkan visualisasi prediksi tersebut dengan menggunakan data dari 2006 hingga 2020 untuk memprediksi rata-rata luas lahan pada tahun 2021 hingga 2030 mendatang. Berdasarkan grafik tersebut, terlihat bahwa prediksi menunjukkan rata-rata penggunaan luas lahan akan menurun untuk tahun-tahun selanjutnya sampai tahun 2030.

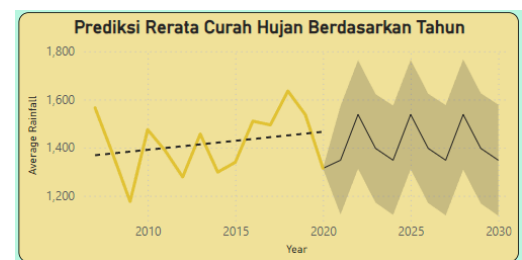
- Prediksi Rata-rata Luas Hasil Panen berdasarkan Tahun



Gambar 22. Prediksi Luas Hasil Panen

Untuk memprediksi luas hasil panen yang akan diperoleh pada masa mendatang, digunakan line chart dengan trend line dan forecast dengan value x-axis year dan y-axis measure Average Yield. Gambar 22 memperlihatkan prediksi luas hasil panen untuk tahun 2021 sampai 2030 menggunakan data yang diperoleh dari 2006 sampai 2020. Dari hasil visualisasi tersebut, terlihat bahwa prediksi menunjukkan luas hasil panen akan memiliki rata-rata yang stabil untuk tahun-tahun selanjutnya sampai tahun 2030.

- Prediksi Rata-rata Curah Hujan berdasarkan Tahun

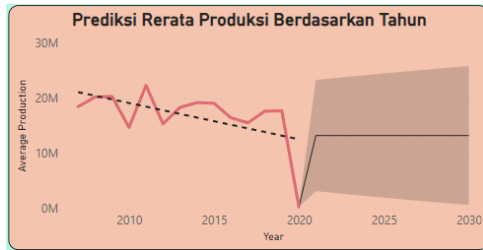


Gambar 23. Prediksi Curah Hujan

Untuk menunjukan prediksi rata-rata curah hujan, digunakan line chart yang memiliki forecast dan trendline dengan value x-axis year dan y-axis measure Average Rainfall. Gambar 23 merupakan visualisasi dari line chart tersebut dimana grafik menunjukkan prediksi curah hujan untuk tahun 2021 sampai tahun 2030 dengan menggunakan data dari 2006 sampai 2020. Berdasarkan hasil, prediksi menunjukkan bahwa akan ada fluktuasi berulang-ulang pada tahun 2021 sampai 2030 dimana curah hujannya akan naik turun seiring bergantinya tahun.

- Prediksi Rata-rata Jumlah Produksi Tanaman berdasarkan Tahun

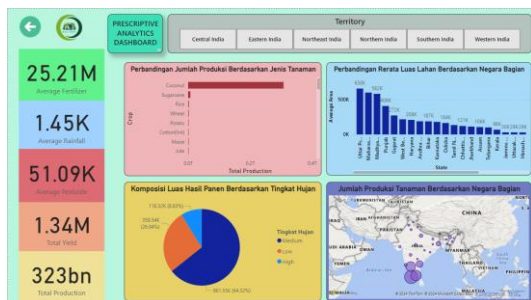




Gambar 24. Prediksi Produksi Tanaman

Dalam memprediksi rata-rata hasil produksi untuk masa mendatang, digunakan line chart dengan forecast dan trendline dengan value x-axis year dan y-axis measure Average Production. Gambar 23 menunjukkan line chart yang memiliki prediksi rata-rata produksi untuk tahun 2021 sampai 2030 dengan data dari 2006 sampai 2020. Dari hasil visualisasi tersebut, terlihat bahwa line chart menunjukkan prediksi *sideways* atau tidak ada fluktuasi dari rata-rata produksi dimana nilainya akan stabil untuk tahun 2021 sampai dengan 2030.

- 4) Laporan Preskriptif: Dalam gambar 3, laporan ini ditandai dengan batas kotak hijau terang, digunakan untuk menawarkan solusi terhadap kelemahan dan peluang. Visualisasi dashboard dari laporan preskriptif perusahaan Kaiture digambarkan sebagai berikut:



Gambar 25. Dashboard Analitik Preskriptif

Dashboard ini menampilkan visualisasi analitik preskriptif yang memuat beberapa hal, seperti perbandingan jumlah produksi, perbandingan rerata luas lahan, komposisi luas hasil panen, dan peta negara India yang menggambarkan jumlah produksi tanaman. Pada dashboard ini digunakan slicer atau filter dengan field Territory yang memuat 6 wilayah India yaitu Central India, Eastern India, Northeast India, Southern India, dan Western India. Slicer tersebut berguna agar pengguna dapat menganalisis aspek-aspek diatas berdasarkan 6 wilayah negara India yang berbeda-beda. Dengan dashboard ini, pengguna dapat dengan mudah melakukan analisis terhadap:

- a) Rata-rata Penggunaan Pupuk, Pestisida, dan Curah Hujan di Wilayah Tertentu



Gambar 26. Rata-rata Pestisida, Pupuk, Hujan

Untuk memberikan informasi terhadap rata-rata penggunaan pestisida, pupuk, dan juga curah hujan yang ada pada satu dari delapan wilayah di India, digunakan visualisasi card untuk setiap values yang berbeda-beda tersebut. Dari hasil visualisasi seperti pada Gambar 26, terlihat bahwa rata-rata penggunaan pupuk di seluruh wilayah adalah 25.21 milliar kilogram, sementara rata-rata penggunaan pestisida sebesar 51 ribu kilogram, dan untuk rata-rata curah hujan adalah 1,450 milimeter.

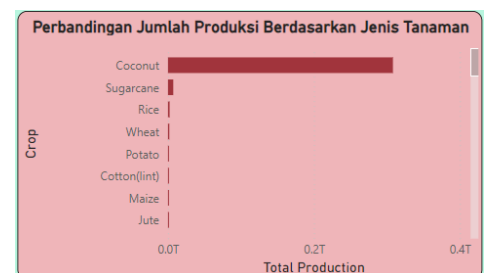
- b) Total Produksi Tanaman dan Luas Hasil Panen



Gambar 27. Total Produksi dan Luas Panen

Untuk menggambarkan total produksi dan luas panen yang dimiliki oleh sebuah wilayah di India, digunakan visualisasi jenis card dengan value measure Total Production dan Total Yield. Dari hasil visualisasi tersebut, terlihat bahwa total produksi dari seluruh tanaman di seluruh negara bagian adalah 323 milliar dan total luas hasil panen adalah 1.34 juta hektar. Hal ini dapat berubah bergantung pada wilayah yang dipilih.

- c) Perbandingan Jumlah Produksi berdasarkan Jenis Tanaman

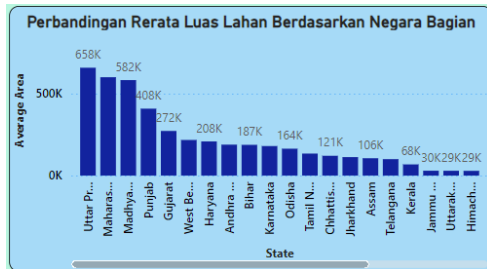


Gambar 28. Perbandingan Jumlah Produksi

Dalam mencari tanaman yang terbaik untuk diproduksi pada sebuah wilayah di India, digunakan visualisasi bar chart untuk memperlihatkan perbandingan jumlah produksi antara jenis tanaman. Bar chart tersebut dibuat dengan value y-axis jenis tanaman dan x-axis measure Total Production. Pada Gambar 28, terlihat perbandingan antara

seluruh jenis tanaman dimana jumlah produksi terbesar berada pada jenis tanaman coconut atau kelapa dimana jumlah produksi dari kelapa di seluruh wilayah di India mencapai lebih dari 310 miliar produksi dari tahun 1997 sampai 2020. Hal ini bisa berubah-ubah jika slicer wilayah India diganti-ganti oleh pengguna.

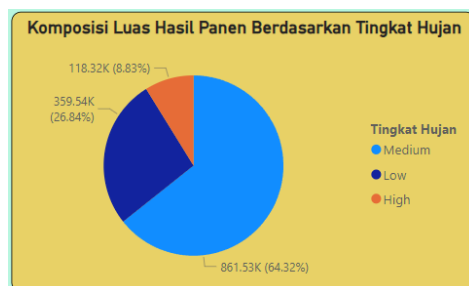
- d) Perbandingan Rata-rata Luas Lahan berdasarkan Negara Bagian



Gambar 29. Perbandingan Rerata Luas lahan

Dalam upaya menentukan negara bagian yang cocok dalam menanam sebuah jenis tanaman, dibutuhkan analisis terhadap rata-rata penggunaan luas lahan pada berbagai negara bagian di seluruh wilayah India. Karena itu dibuat visualisasi perbandingan rata-rata luas lahan berdasarkan negara bagian menggunakan stacked column chart dengan x-axis state dan y-axis measure Average Area. Dari visualisasi pada Gambar 29, terlihat bahwa negara bagian atau provinsi yang memiliki rata-rata penggunaan luas lahan terbesar di seluruh wilayah adalah provinsi Uttar Pradesh yang memiliki luas penanaman sebesar 658 ribu hektar.

- e) Komposisi Luas Hasil Panen berdasarkan Tingkat Hujan

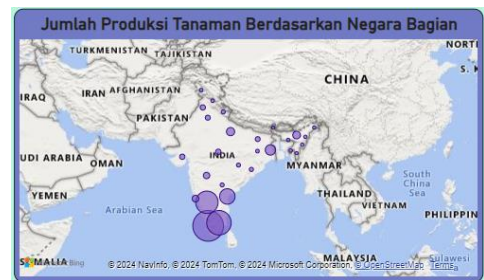


Gambar 30. Komposisi Luas Hasil Panen

Untuk menentukan kondisi cuaca yang optimal dalam proses penanaman, diperlukan analisis mendalam mengenai cuaca, khususnya terkait dengan tingkat hujan yang paling efektif untuk mendukung hasil panen yang maksimal. Dalam analisis ini, visualisasi yang digunakan adalah diagram lingkaran (pie

chart) yang menggambarkan komposisi luas hasil panen berdasarkan tiga tingkat atau level hujan yang berbeda, yaitu rendah, sedang, dan tinggi. Dari Gambar 30, terlihat bahwa hasil panen terbesar dalam satuan luas terjadi pada saat tingkat hujan sedang (medium). Sebaliknya, hasil panen dengan satuan luas terkecil berada pada saat tingkat hujan mencapai level tinggi (high).

- f) Peta yang Menggambarkan Jumlah Produksi Tanaman Berdasarkan Negara Bagian Di Wilayah Tertentu



Gambar 31. Peta Produksi Tanaman

Untuk mempermudah pengamatan terhadap wilayah yang baik untuk melakukan penanaman di India, digunakan peta yang menggambarkan jumlah produksi tanaman berdasarkan negara bagian di wilayah tertentu. Peta tersebut divisualisasikan dengan kolom state sebagai location dan kolom Total Production sebagai bubble size. Gelembung berwarna ungu menggambarkan daerah-daerah penanaman dan ukuran dari gelembung tersebut digunakan sebagai parameter atau tingkat jumlah produksi tanaman yang dimiliki oleh wilayah tersebut. Berdasarkan visualisasi pada Gambar 31, dapat dilihat bahwa wilayah yang memiliki jumlah produksi yang paling besar adalah wilayah Southern India, tepatnya pada negara bagian Kerala. Sebaliknya, wilayah tengah dari negara India atau Central India menjadi wilayah yang memiliki tingkat produksi paling sedikit

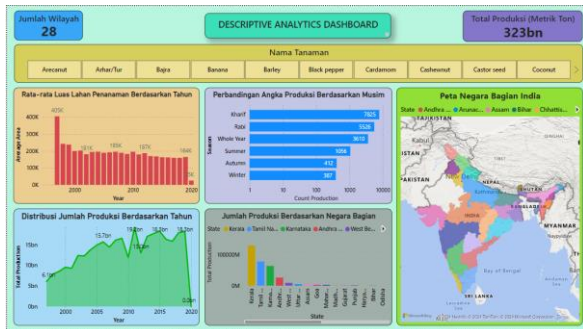
#### IV. QUESTION III SUB-CLO-5

##### A. Tujuan dan Manfaat Analitik Deskriptif

Analitik deskriptif yang diterapkan di Kaiture bertujuan untuk memahami dan menggambarkan situasi yang terjadi di lapangan terkait produksi pertanian. Melalui analitik ini, Kaiture dapat mengumpulkan, mengelompokkan, dan menganalisis data historis untuk memberikan gambaran yang jelas mengenai kondisi produksi, seperti total produksi tanaman per tahun, musim terbaik untuk penanaman, rata-rata luas lahan yang digunakan, serta distribusi

geografis produksi tanaman di berbagai wilayah.

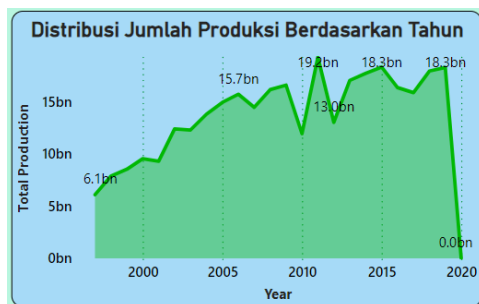
Maksud utama dari analitik deskriptif ini adalah untuk memberikan wawasan yang mudah dipahami tentang kinerja historis produksi pertanian. Dengan mengetahui pola-pola yang muncul dari data sebelumnya, Kaiture bisa memahami bagaimana proses produksi berlangsung, mengidentifikasi tren produksi, dan menentukan waktu atau lokasi yang tepat untuk menanam. Dalam implementasi analitik deskriptif, dibuat sebuah dashboard yang memuat visualisasi data-data historis yang akan memenuhi setiap pertanyaan yang termasuk analitik deskriptif.



Gambar 32. Dashboard Analitik Deskriptif

Dashboard ini menampilkan visualisasi analitik deskriptif yang mencakup berbagai elemen, seperti rata-rata luas lahan penanaman, perbandingan angka produksi, distribusi jumlah produksi, jumlah produksi per wilayah, peta negara bagian India, jumlah wilayah, serta total produksi. Slicer atau filter yang digunakan pada dashboard adalah berdasarkan jenis tanaman (crop), sehingga pengguna dapat menganalisis semua aspek tersebut sesuai dengan tanaman yang dipilih. Setiap visualisasi pada dashboard tersebut akan menjawab pertanyaan-pertanyaan berikut:

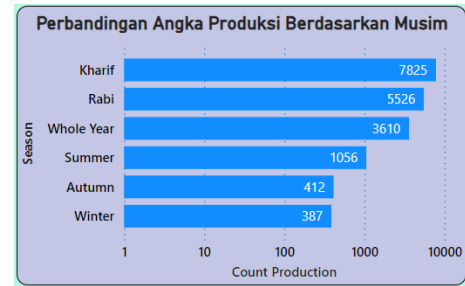
- 1) **Pertanyaan 1: Berapa total produksi tanaman per tahun untuk berbagai jenis tanaman?**



Gambar 33. Deskriptif Jumlah Produksi

Tujuan: Menyajikan data jumlah produksi tahunan untuk berbagai jenis tanaman, sehingga dapat dianalisis tren produksi dari tahun ke tahun.

- 2) **Pertanyaan 2: Musim apakah yang memiliki angka produksi yang paling banyak untuk berbagai tanaman?**



Gambar 34. Deskriptif Angka Produksi

Tujuan: Mengidentifikasi musim utama di mana berbagai jenis tanaman diproduksi dalam frekuensi tertinggi untuk membantu dalam perencanaan penanaman.

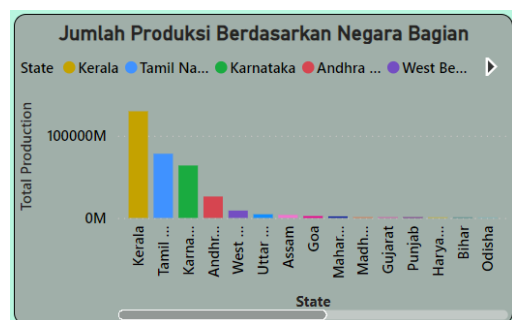
- 3) **Pertanyaan 3: Berapa rata-rata luas lahan yang digunakan untuk penanaman berbagai jenis tanaman setiap tahunnya?**



Gambar 35. Deskriptif Luas Lahan

Tujuan: Menampilkan distribusi rata-rata area lahan yang digunakan untuk penanaman berbagai jenis tanaman di setiap negara bagian.

- 4) **Pertanyaan 4: Di wilayah mana saja sebuah tanaman diproduksi paling banyak?**



Gambar 36. Deskriptif Produksi Negara

Tujuan: Menyajikan analisis distribusi geografis dari produksi berbagai jenis tanaman di seluruh negara bagian India.

- 5) **Pertanyaan 5: Dimana letak geografis yang baik untuk menanam sebuah tanaman?**



Gambar 37. Deskriptif Peta India

Tujuan: Menampilkan lokasi-lokasi utama di mana jenis tanaman tertentu biasanya ditanam atau diproduksi berdasarkan negara bagian.

#### B. Measurement Pada Kaiture

Terdapat beberapa measurement yang dibuat dalam data perusahaan Kaiture untuk membantu dalam analisis kinerja produksi dan efisiensi operasional. Measurement ini mencakup indikator penting seperti rata-rata penggunaan lahan, rata-rata penggunaan pupuk, rata-rata penggunaan pestisida, serta total hasil produksi dan hasil panen. Selain itu, terdapat juga pengukuran terkait curah hujan, yang berfungsi untuk mengkaji dampaknya terhadap hasil produksi. Measurement tersebut dapat dijabarkan sebagai berikut:

##### 1) Average Area

Menghitung rata-rata luas lahan yang digunakan untuk menanam berbagai jenis tanaman di seluruh wilayah. Dengan ini, perusahaan dapat memahami seberapa besar area yang dibutuhkan rata-rata untuk setiap tanaman, yang berguna dalam perencanaan alokasi sumber daya dan manajemen lahan.

$$\text{Average Area} = \text{CALCULATE}(\text{AVERAGE}(\text{crop\_yield}[\text{Area}])))$$

Gambar 38. Average Area

##### 2) Average Fertilizer

Mengukur rata-rata penggunaan pupuk untuk tanaman di berbagai wilayah. Data ini penting untuk menganalisis seberapa banyak pupuk yang digunakan dalam produksi tanaman dan untuk mengevaluasi apakah penggunaan pupuk sudah optimal atau bisa disesuaikan untuk meningkatkan hasil atau mengurangi biaya.

$$\text{Average Fertilizer} = \text{CALCULATE}(\text{AVERAGE}(\text{crop\_yield}[\text{Fertilizer}])))$$

Gambar 39. Average Fertilizer

##### 3) Average Pesticide

Menghitung rata-rata penggunaan pestisida untuk

setiap jenis tanaman di wilayah tertentu. Informasi ini berguna untuk mengevaluasi apakah penggunaan pestisida berada pada tingkat yang wajar atau apakah ada potensi untuk mengurangi penggunaannya sambil tetap mempertahankan produktivitas, terutama dalam rangka mendukung praktik pertanian berkelanjutan.

$$\text{Average Pesticide} = \text{CALCULATE}(\text{AVERAGE}(\text{crop\_yield}[\text{Pesticide}])))$$

Gambar 40. Average Pesticide

##### 4) Average Production

Menyajikan rata-rata hasil produksi tanaman di setiap wilayah. Data ini membantu dalam memahami tingkat produktivitas dari tahun ke tahun, memungkinkan perusahaan untuk mengevaluasi kinerja produksi dan membuat keputusan yang lebih baik terkait pengelolaan sumber daya dan peningkatan produksi.

$$\text{Average Production} = \text{CALCULATE}(\text{AVERAGE}(\text{crop\_yield}[\text{Production}])))$$

Gambar 41. Average Production

##### 5) Average Rainfall

Mengukur rata-rata curah hujan yang diterima di wilayah tertentu. Data curah hujan penting dalam menentukan kondisi pertumbuhan yang ideal bagi tanaman, membantu perusahaan memahami bagaimana curah hujan mempengaruhi hasil panen dan merencanakan kebutuhan irigasi yang tepat.

$$\text{Average Rainfall} = \text{CALCULATE}(\text{AVERAGE}(\text{crop\_yield}[\text{Annual\_Rainfall}])))$$

Gambar 42. Average Rainfall

##### 6) Average Yield

Menghitung rata-rata hasil per satuan luas lahan (misalnya, ton per hektar) untuk tanaman yang diproduksi. Measure ini penting untuk menganalisis efisiensi produksi tanaman, mengidentifikasi wilayah atau metode yang menghasilkan hasil tertinggi, dan mengoptimalkan praktik pertanian untuk meningkatkan produktivitas.

$$\text{Average Yield} = \text{CALCULATE}(\text{AVERAGE}(\text{crop\_yield}[\text{Yield}])))$$

Gambar 43. Average Yield

##### 7) Count Production

Menghitung jumlah entri atau catatan yang terkait dengan produksi tanaman. Measure ini memberikan gambaran tentang berapa banyak kali data produksi tanaman tercatat, membantu menganalisis volume produksi yang dilaporkan di seluruh wilayah atau periode waktu tertentu.

$$\text{Count Production} = \text{COUNT}(\text{crop\_yield}[\text{Production}]))$$

Gambar 44. Count Production

##### 8) Count Yield

Menghitung jumlah entri terkait hasil produksi per



satuan luas lahan. Dengan count ini, perusahaan dapat mengetahui berapa banyak kali hasil tanaman diukur atau dicatat dalam dataset, yang berguna dalam analisis agregat hasil di berbagai wilayah.

```
Count Yield = CALCULATE(COUNT(crop_yield[Yield]))
```

Gambar 45. Count Yield

#### 9) Count Season

Menghitung jumlah musim yang terkait dengan produksi tanaman. Data ini membantu mengidentifikasi musim yang paling sering dicatat dalam produksi tanaman dan bisa digunakan untuk analisis musiman terhadap hasil pertanian.

```
Count Season = CALCULATE(COUNT(crop_yield[Season]))
```

Gambar 46. Count Season

#### 10) Distinct Count Crop

Menghitung jumlah unik dari jenis tanaman yang ditanam atau diproduksi. Dengan measure ini, perusahaan dapat memahami keragaman tanaman yang ditanam di berbagai wilayah atau periode waktu, yang berguna dalam perencanaan rotasi tanaman atau analisis keberagaman produksi.

```
DistCount Crop = DISTINCTCOUNT(crop_yield[Crop])
```

Gambar 47. DistCount Crop

#### 11) Distinct Count State

Menghitung jumlah negara bagian yang unik atau wilayah geografis tempat tanaman diproduksi. Data ini membantu dalam menganalisis distribusi geografis produksi dan mengidentifikasi wilayah yang paling produktif atau strategis untuk produksi tanaman tertentu.

```
DistCount State = DISTINCTCOUNT(crop_yield[State])
```

Gambar 48. DistCount State

#### 12) Total Fertilizer

Menghitung total jumlah pupuk yang digunakan dalam produksi tanaman di seluruh wilayah. Data ini membantu mengevaluasi seberapa besar penggunaan pupuk secara keseluruhan dan memungkinkan analisis apakah ada peluang untuk mengoptimalkan penggunaan pupuk guna meningkatkan efisiensi atau mengurangi dampak lingkungan.

```
Total Fertilizer = CALCULATE(SUM(crop_yield[Fertilizer]))
```

Gambar 49. Total Fertilizer

#### 13) Total Pesticide

Menghitung total penggunaan pestisida di seluruh wilayah yang dicatat. Data ini berguna untuk memantau penggunaan pestisida dan menganalisis apakah ada kebutuhan untuk mengurangi penggunaannya, mendukung praktik pertanian yang

lebih ramah lingkungan, dan menjaga keseimbangan ekosistem.

```
Total Pesticide = CALCULATE(SUM(crop_yield[Pesticide]))
```

Gambar 50. Total Pesticide

#### 14) Total Production

Menghitung total hasil produksi tanaman di seluruh wilayah atau periode waktu tertentu. Data ini memberikan gambaran keseluruhan mengenai produktivitas dan membantu perusahaan dalam mengevaluasi kinerja produksi tanaman dari tahun ke tahun atau dari satu wilayah ke wilayah lain.

```
Total Production = CALCULATE(SUM(crop_yield[Production]))
```

Gambar 51. Total Production

#### 15) Total Yield

Menghitung total hasil per satuan luas lahan di seluruh wilayah. Data ini berguna untuk menilai produktivitas lahan secara keseluruhan dan memberikan wawasan mengenai efisiensi produksi. Total yield yang tinggi dapat menunjukkan bahwa sumber daya lahan digunakan secara optimal.

```
Total Yield = CALCULATE(SUM(crop_yield[Yield]))
```

Gambar 52. Total Yield

## V. QUESTION IV SUB-CLO-7

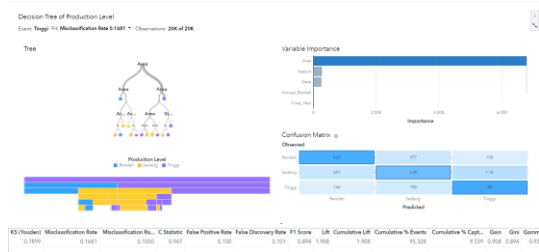
Dalam upaya meningkatkan akurasi prediksi dan pemahaman terhadap pola produksi tanaman, perusahaan Kaiture menggunakan berbagai teknik machine learning untuk mengklasifikasikan level produksi sebuah tanaman melalui target kolom Production Level. Tiga model utama yang digunakan dalam analisis ini adalah Decision Tree, Gradient Boosting, dan Random Forest. Setiap model ini menawarkan pendekatan yang unik dalam mengidentifikasi faktor-faktor penting yang mempengaruhi level produksi, dengan tujuan membantu perusahaan dalam merumuskan strategi yang lebih tepat dan berbasis data. Visualisasi dari masing-masing model ini akan memberikan gambaran yang jelas mengenai efektivitasnya dalam memprediksi dan mengklasifikasikan tingkat produksi berdasarkan variabel-variabel kunci seperti penggunaan lahan, curah hujan, dan input pertanian lainnya.

#### 1) Decision Tree

Decision Tree adalah algoritma machine learning yang digunakan untuk melakukan klasifikasi atau regresi dengan memecah dataset menjadi subset berdasarkan fitur tertentu, hingga mencapai keputusan akhir dalam bentuk "daun" yang merepresentasikan kelas atau nilai target. Setiap percabangan dalam pohon mewakili keputusan berdasarkan fitur data, yang diambil melalui



pemilihan fitur terbaik untuk memisahkan data. Dalam konteks perusahaan Kaiture, decision tree digunakan untuk mengklasifikasikan Production Level tanaman dengan memecah data berdasarkan variabel seperti luas lahan, curah hujan, penggunaan pupuk, dan pestisida.

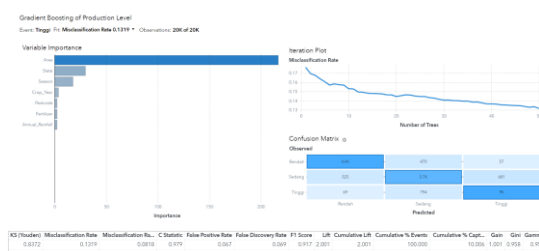


Gambar 53. Decision Tree Kaiture

Hasil Decision Tree untuk mengklasifikasikan Production Level pada dataset menunjukkan bahwa variabel yang paling berpengaruh dalam menentukan tingkat produksi adalah luas lahan (Area), diikuti oleh musim (Season) dan negara bagian (State). Visualisasi pohon keputusan menggambarkan percabangan berdasarkan variabel-variabel ini, dengan kategori Production Level dibagi menjadi Rendah, Sedang, dan Tinggi. Matriks kebingungan (Confusion Matrix) menunjukkan performa prediksi model, di mana tingkat kesalahan klasifikasi sebesar 0.1681. Sebagian besar prediksi akurat pada kelas Tinggi dengan 9K observasi yang diprediksi benar dari total 10K observasi. Model ini memiliki C-Statistic sebesar 0.947, yang menandakan model ini cukup baik dalam membedakan antara kelas produksi yang berbeda.

## 2) Gradient Boosting

Gradient Boosting adalah metode machine learning yang digunakan untuk meningkatkan akurasi prediksi dengan menggabungkan beberapa model sederhana (biasanya decision trees) secara bertahap. Setiap model baru berusaha memperbaiki kesalahan prediksi yang dibuat oleh model sebelumnya dengan meminimalkan fungsi loss secara bertahap. Teknik ini sangat efektif dalam menangani data yang kompleks dan memberikan prediksi yang akurat. Di Kaiture, Gradient Boosting dapat diterapkan untuk memprediksi Production Level berdasarkan berbagai faktor seperti luas lahan, curah hujan, dan musim.

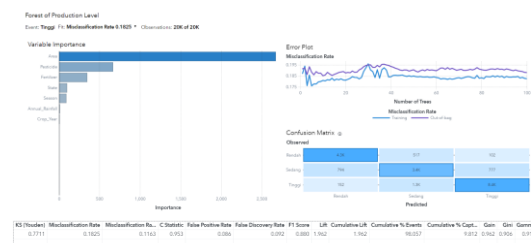


Gambar 54. Gradient Boosting Kaiture

Hasil Gradient Boosting untuk memprediksi Production Level menunjukkan performa yang baik dengan misclassification rate sebesar 0.1319, yang lebih rendah dibandingkan dengan metode lainnya, mengindikasikan model yang lebih akurat. Dari variable importance, faktor "Area" dan "State" merupakan variabel paling berpengaruh dalam menentukan level produksi, diikuti oleh "Season". Plot iterasi memperlihatkan bahwa dengan menambahkan jumlah pohon (trees), misclassification rate terus menurun hingga model mencapai stabilitas pada sekitar 50 pohon. Matriks kebingungan menunjukkan bahwa prediksi untuk kategori "Tinggi" memiliki presisi yang cukup baik dengan 9K observasi yang terklasifikasi dengan benar, sementara kategori "Rendah" dan "Sedang" juga menunjukkan prediksi yang cukup akurat. Nilai KS (0.8372) dan C-Statistic (0.979) mengindikasikan performa klasifikasi yang kuat, sementara F1 Score sebesar 0.917 menunjukkan keseimbangan yang baik antara presisi dan recall.

## 3) Forest Model

Forest model, atau lebih dikenal sebagai Random Forest, adalah metode ensemble learning yang menggabungkan banyak decision trees untuk membuat prediksi yang lebih akurat dan stabil. Dalam konteks Kaiture, Random Forest digunakan untuk mengklasifikasikan level produksi berdasarkan berbagai faktor seperti luas lahan (Area), musim (Season), dan curah hujan tahunan. Model ini berguna karena mampu menangani data yang kompleks dan tidak terlalu rentan terhadap overfitting, berkat pendekatan bagging yang digunakan, di mana banyak pohon keputusan dibangun dari subset data yang berbeda. Pada Kaiture, penggunaan Random Forest memberikan prediksi yang lebih robust dan dapat membantu perusahaan membuat keputusan yang lebih tepat dalam mengelola sumber daya dan meningkatkan efisiensi produksi berdasarkan pola historis yang ada.



Gambar 55. Forest Model Kaiture

Hasil model Random Forest untuk memprediksi Production Level menunjukkan performa yang baik dengan misclassification rate sebesar 0.1825. Meskipun tidak sebaik metode lain dengan misclassification rate yang lebih rendah, model ini tetap memberikan hasil yang cukup akurat.

Berdasarkan variable importance, faktor Area dan Pesticide merupakan variabel paling berpengaruh dalam menentukan level produksi, diikuti oleh Fertilizer. Plot iterasi memperlihatkan bahwa seiring bertambahnya jumlah pohon, misclassification rate stabil setelah sekitar 50 pohon, meskipun terdapat fluktuasi kecil. Matriks kebingungan menunjukkan bahwa prediksi untuk kategori Tinggi memiliki presisi yang cukup baik dengan 8.4K observasi terklasifikasi dengan benar, sementara kategori Rendah dan Sedang juga menunjukkan prediksi yang lumayan akurat. Nilai KS sebesar 0.7711 dan C-Statistic sebesar 0.953 mengindikasikan performa klasifikasi yang kuat, sementara F1 Score sebesar 0.880 menunjukkan keseimbangan yang cukup baik antara presisi dan recall.

Dari ketiga model yang digunakan oleh Kaiture, yaitu Decision Tree, Gradient Boosting, dan Random Forest, setiap model memberikan keunggulan yang berbeda dalam memprediksi Production Level tanaman. Gradient Boosting tampil sebagai model paling akurat dengan misclassification rate terendah sebesar 0.1319, sementara Decision Tree dan Random Forest masing-masing mencapai 0.1681 dan 0.1825. Faktor yang paling berpengaruh pada ketiga model adalah Area, diikuti oleh variabel lain seperti Season, State, dan Pesticide. Dari segi performa, Gradient Boosting memiliki C-Statistic tertinggi di 0.979, menunjukkan kemampuan prediksi yang superior, sementara Random Forest dan Decision Tree juga menunjukkan hasil yang baik dengan C-Statistic masing-masing 0.953 dan 0.947.

## VI. CONCLUSION

Perusahaan Kaiture telah memanfaatkan empat jenis analitik, yaitu analitik deskriptif, diagnostik, prediktif, dan preskriptif, serta teknik machine learning untuk meningkatkan efisiensi dan akurasi dalam pengelolaan produksi tanaman. Dari analitik deskriptif, perusahaan dapat memahami tren historis produksi tanaman yang menunjukkan total hasil panen sebesar 323 miliar ton di seluruh wilayah India. Analitik diagnostik memberikan insight mengenai variabel-variabel penting seperti luas lahan panen sebesar 1,34 juta hektar dan komposisi panen tertinggi pada musim Kharif dengan persentase 41,59%. Dengan analitik prediktif, Kaiture mampu memproyeksikan tren penggunaan lahan yang diprediksi menurun hingga 2030 dan memperkirakan stabilitas produksi tanaman pada angka rata-rata dari tahun 2021 hingga 2030. Melalui analitik preskriptif, Kaiture dapat menentukan strategi optimal berdasarkan perbandingan produksi antarwilayah, di mana Southern India (Kerala) mencatat produksi tertinggi. Di sisi machine learning, model Gradient Boosting menonjol dengan tingkat akurasi tertinggi, memiliki misclassification rate 0,1319 dan C-Statistic 0,979, memberikan prediksi produksi yang lebih akurat berdasarkan variabel utama seperti luas

lahan dan penggunaan pestisida. Keseluruhan analitik dan modelling ini membantu Kaiture dalam mengidentifikasi peluang peningkatan efisiensi produksi, mengurangi kesalahan prediksi, dan merumuskan strategi berbasis data untuk memaksimalkan hasil panen di berbagai wilayah India.

## ACKNOWLEDGMENT

Penulis mengucapkan terima kasih kepada Bapak Iwan Prasetyawan, S.Kom., M.M. dosen mata kuliah Advanced Big Data Analytics di Program Studi Sistem Informasi Universitas Multimedia Nusantara. Bapak Iwan Prasetyawan, S.Kom., M.M. telah dengan murah hati membagikan pengetahuan dan wawasan yang sangat berharga, yang sangat berkontribusi pada kemajuan penelitian ini. Selain itu, penulis mengakui kesiapan Bapak Iwan Prasetyawan, S.Kom., M.M. untuk memberikan masukan dan bimbingan sejak tahap perencanaan awal studi ini.

## REFERENCES

- [1] A. A. Adha and P. Andiny, "Pengaruh Tenaga Kerja dan Investasi di Sektor Pertanian terhadap Pertumbuhan Ekonomi Sektor Pertanian di Indonesia," *J. Samudra Ekonomika*, vol. 6, no. 1, pp. 40-49, 2022.
- [2] S. Maulidina, V. M. Zahara, and H. Sutjipto, "Analisis Faktor Yang Mempengaruhi Kemiskinan Pada Sektor Pertanian Di Indonesia Bagian Barat," in *National Conference on Applied Business, Education, & Technology (NCABET)*, vol. 2, no. 1, pp. 257-269, October 2022.
- [3] S. Rahayu, "Potensi Ekonomi Sektor Pertanian, Kehutanan Dan Perikanan Kabupaten Merangin," *Bussman Journal: Indonesian Journal of Business and Management*, vol. 2, no. 1, pp. 147-163, 2022.
- [4] P. Lasaksi, "Analisis Peran Sektor Pertanian Terhadap Perekonomian," *Lentera: Multidisciplinary Studies*, vol. 1, no. 3, pp. 165-171, 2023.
- [5] S. Febriani and H. Sulistiani, "Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5," *J. Teknol. Dan Sist. Inf.*, vol. 2, no. 4, pp. 89-95, 2021.
- [6] A. M. Argina, "Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes," *Indonesian Journal of Data and Science*, vol. 1, no. 2, pp. 29-33, 2020.
- [7] F. Pratama, M. Nasir, and S. Sauda, "Implementasi Metode Klasifikasi Dengan Algoritma Support Vector Machine Untuk Menentukan Stok Persediaan Barang Pada Koperasi Karyawan Pangan Utama," *Journal of Software Engineering Ampera*, vol. 1, no. 2, pp. 71-81, 2020.
- [8] A. I. Putri, Y. Syarif, P. Jayadi, F. Arrazak, and F. N. Salisah, "Implementasi Algoritma Decision Tree dan Support Vector Machine (SVM) untuk Prediksi Risiko Stunting pada Keluarga: Implementation of Decision Tree and Support Vector Machine (SVM) Algorithm for Stunting Risk Prediction," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 349-357, 2023.
- [9] S. Talib, S. Sudin, and M. D. Suratin, "PENERAPAN METODE SUPPORT VECTOR MACHINE (SVM) PADA KLASIFIKASI JENIS CENGKEH BERDASARKAN FITUR TEKSTUR DAUN," *PROSISKO: Jurnal Pengembangan Riset dan Observasi Sistem Komputer*, vol. 11, no. 1, pp. 26-34, 2024.
- [10] D. P. Sinambela, H. Naparin, M. Zulfadhilah, and N. Hidayah, "Implementasi Algoritma Decision Tree dan Random Forest dalam Prediksi Perdarahan Pascasalin," *Jurnal Informasi dan Teknologi*, pp. 58-64, 2023.