

Peramalan Produksi Padi Musiman di Jawa Barat: Analisis SARIMAX dengan Variabel Eksternal Klimatologis dan Pengelompokan K-means

Dosen Pengampu: Yuliagnis Transver Wijaya



Disusun Oleh:

Kelompok 5

Anggota Kelompok:

Rifqi Muhadzib Ahdan	222313350
Seto Haidar Yudhistira	222313375
Zidan Septian	222313447

PROGRAM STUDI D-IV KOMPUTASI STATISTIK

POLITEKNIK STATISTIKA STIS

TAHUN AJARAN 2024/2025

DAFTAR ISI

BAB I PENDAHULUAN.....	3
1.1 Latar Belakang	3
1.2 Rumusan Masalah.....	3
1.3 Tujuan	4
BAB II PEMBAHASAN.....	5
2.1 Pengumpulan Data	5
2.2 Pemrosesan Data	5
2.3 Pelaporan Hasil	7
2.4 Hasil Akhir.....	8
BAB III PENUTUP	18
3.1 Kesimpulan	18
DAFTAR PUSTAKA.....	19

BAB I

PENDAHULUAN

1.1 Latar Belakang

Padi merupakan komoditas utama dan makanan pokok bagi sebagian besar masyarakat Indonesia. Ketersediaan beras yang cukup menjadi faktor penting dalam menjaga ketahanan pangan nasional. Provinsi Jawa Barat sebagai salah satu dari lima provinsi penghasil padi terbesar di Indonesia memiliki peran strategis dalam pemenuhan kebutuhan pangan nasional. Menurut data Badan Pusat Statistik tahun (2019 - 2024), Jawa Barat konsisten menjadi penyumbang signifikan produksi padi nasional.

Namun, dalam beberapa tahun terakhir, perubahan iklim (climate change) telah menjadi ancaman serius bagi sektor pertanian, khususnya produksi padi. Perubahan iklim telah menyebabkan pergeseran musim tanam dan panen padi di berbagai wilayah Indonesia (Ruminta, 2016). Pergeseran ini tidak hanya mengubah pola produksi, tetapi juga meningkatkan risiko gagal panen akibat ketidaksesuaian waktu tanam dengan kondisi iklim optimal.

Produksi padi nasional di Indonesia menurun dari 65,76 juta ton pada 2011 menjadi 55,67 juta ton pada 2022, dengan total penurunan sebesar 10,09 juta ton (Galitan, 2024). Di beberapa kabupaten di Jawa Barat, penurunan ini berkaitan erat dengan perubahan iklim, seperti peningkatan suhu, perubahan pola curah hujan, fluktuasi kelembaban, dan penyinaran matahari, yang mempengaruhi produktivitas padi secara signifikan.

Mengingat pentingnya komoditas padi bagi ketahanan pangan nasional dan indikasi kuat pengaruh perubahan iklim terhadap produksinya, diperlukan analisis mendalam untuk memahami pola, tren, dan faktor-faktor yang mempengaruhi produksi padi di Jawa Barat. Analisis ini akan menjadi dasar untuk mengembangkan strategi adaptasi dan mitigasi yang efektif dalam menghadapi tantangan perubahan iklim pada sektor pertanian padi.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, dapat dirumuskan masalah sebagai berikut.

1. Adanya indikasi pergeseran musim tanam dan panen padi di Jawa Barat yang berpotensi mempengaruhi pola produksi dan ketersediaan pangan.
2. Peningkatan frekuensi gagal panen padi akibat perubahan kondisi iklim yang terus terjadi.

3. Kurangnya pemahaman komprehensif tentang korelasi spesifik antara variabel iklim (suhu, kelembapan, curah hujan, penyinaran matahari) dengan produksi padi di tingkat kabupaten/kota di Jawa Barat.
4. Terbatasnya model prediktif yang dapat memperkirakan produksi padi berdasarkan perubahan variabel iklim untuk perencanaan pertanian jangka panjang.

1.3 Tujuan

Berdasarkan latar belakang dan rumusan masalah tersebut, tujuan dari analisis ini adalah sebagai berikut.

1. Menganalisis pola produksi padi di Jawa Barat dan mengidentifikasi pergeseran musim panen yang mungkin terjadi akibat perubahan iklim dengan model SARIMAX
2. Mengelompokkan (clustering) kabupaten/kota di Jawa Barat berdasarkan karakteristik produksi padi untuk memahami pola spasial dan temporal.
3. Menganalisis korelasi antara variabel iklim (suhu, kelembapan, curah hujan, penyinaran matahari) dengan produksi padi untuk mengidentifikasi faktor-faktor signifikan yang mempengaruhi produktivitas.
4. Membangun model SARIMAX dengan variabel eksternal untuk memprediksi produksi padi di masa depan berdasarkan skenario perubahan iklim.
5. Membangun dashboard interaktif yang mengintegrasikan visualisasi data, analisis spasial, dan model prediktif sebagai alat pendukung keputusan bagi pemangku kepentingan.

BAB II

PEMBAHASAN

2.1 Pengumpulan Data

Untuk dapat menghasilkan analisis yang relevan dan akurat, data yang digunakan harus berasal dari sumber terpercaya, meliputi:

- a. Data Produksi Padi per kabupaten/kota di Jawa Barat (2019-2024)
<https://jabar.bps.go.id/id/statistics-table/2/NTIjMg==/produksi-padi-menurut-kabupaten-kota.html>
- b. Data Iklim (Curah Hujan, Kelembaban, Suhu, dan Penyinaran Matahari)
Data diambil dari : <https://earthengine.google.com/>
- c. Data administrasi kabupaten/kota di Jawa Barat
https://github.com/hitamcoklat/Jawa-Barat-Geo-JSON/blob/master/Jabar_By_Kec.geojson

Catatan : Kami menggunakan beberapa data dummy untuk keperluan analisis

2.2 Pemrosesan Data

- a. Cleaning dan Pre-Processing
Pembersihan data (menangani missing values, outliers), Agregasi data ke tingkat kabupaten/kota dan bulanan, dan Transformasi data sesuai kebutuhan analisis
- b. Eksplorasi Data
Analisis statistik deskriptif untuk memahami karakteristik data, Analisis korelasi antara variabel iklim dan produksi padi, Uji hipotesis untuk menentukan signifikansi hubungan antar variabel, Analisis dengan melibatkan variabel kategorik.
 - 1) Analisis Univariat (Eksplorasi Variabel Tunggal) Analisis ini berfokus pada karakteristik satu variabel independen pada satu waktu untuk memahami distribusi dan tendensinya. (Statistik Deskriptif, Histogram & Boxplot, Uji Normalitas)
 - 2) Analisis Bivariat (Hubungan Antar Variabel Numerik) Analisis ini mengeksplorasi hubungan antara variabel independen numerik dan variabel dependen numerik. (ScatterPlot, Analisis Korelasi, Uji Signifikansi)
 - 3) Analisis Perbandingan Kelompok (Hubungan Variabel Numerik & Kategorik) Analisis ini digunakan untuk membandingkan distribusi variabel numerik di berbagai kelompok atau kategori. (Boxplot & Violinplot per kategori, ANOVA)
- c. Pemodelan
 - 1) SARIMAX

Salah satu metode peramalan yang paling kuat dan fleksibel adalah SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous variables). Model ini dirancang secara khusus untuk menangani data yang memiliki tiga karakteristik utama: tren, pola musiman, dan pengaruh dari faktor-faktor eksternal. Kemampuannya untuk memodelkan ketiga aspek ini secara simultan menjadikan SARIMAX alat yang sangat efektif untuk menganalisis dan meramalkan data kompleks di dunia nyata, termasuk data produksi pertanian.

Dasar dari SARIMAX adalah model ARIMA. Secara umum, SARIMAX menggabungkan komponen autoregressive (AR), moving average (MA), integrasi (I) untuk membuat data stasioner, dan Komponen musiman (S), Serta pengaruh variabel luar (X) seperti suhu, curah hujan, dan faktor iklim lainnya.

Secara konseptual, formula SARIMAX dapat ditulis sebagai:

$$\phi_p(B)\Phi_P(B^m)(1-B)^d(1-B^m)^D y_t = \theta_q(B)\Theta_Q(B^m)\epsilon_t + \beta X_t$$

Dimana:

- y_t adalah nilai observasi pada waktu t
- $\phi_p, \Phi_P, \theta_q, \Theta_Q$ adalah polinomial untuk orde AR, SAR, MA, dan SMA
- B adalah operator *backshift* ($By_t = y_{t-1}$)
- $(1-B)^d$ adalah operator *differencing* non-musiman
- $(1-B^m)^D$ adalah operator *differencing* musiman
- ϵ_t adalah *white noise* atau *error term*
- βX_t merepresentasikan pengaruh dari variabel eksogen X_t

2) Clustering dengan K-Means

Clustering adalah sebuah pendekatan dalam unsupervised learning yang bertujuan untuk menemukan struktur atau pola tersembunyi dalam data tersebut. Tujuan utamanya adalah membagi populasi data menjadi beberapa kelompok (cluster) sedemikian rupa sehingga data dalam satu kelompok memiliki tingkat kesamaan yang tinggi, dan tingkat kesamaan yang rendah dengan data di kelompok lain.

K-Means adalah sebuah algoritma yang bertujuan untuk mempartisi (membagi) n observasi data ke dalam k buah cluster. Pembagian ini dilakukan dengan cara menetapkan setiap observasi ke cluster yang memiliki rata-rata (mean) terdekat. Rata-rata dari setiap cluster ini disebut sebagai titik pusat atau centroid, yang berfungsi sebagai prototipe atau representasi dari cluster tersebut.

Tujuan matematis dari K-Means adalah untuk meminimalkan Within-Cluster Sum of Squares (WCSS), yaitu total dari jarak kuadrat antara setiap titik data dengan centroid dari clusternya. Formula WCSS adalah sebagai berikut:

$$WCSS = \sum_{j=1}^K \sum_{i \in C_j} |x_i - \mu_j|^2$$

Dimana:

- K adalah jumlah cluster
- C_j adalah himpunan titik data dalam cluster ke-j
- x_i adalah titik data ke-i
- μ_j adalah centroid dari cluster ke-j

2.3 Pelaporan Hasil

a. Visualisasi Data

- Peta choropleth produksi padi per kabupaten/kota
- Grafik time series produksi padi dan variabel iklim
- Visualisasi hasil clustering dan analisis korelasi

b. Timeline Pembuatan Project

Pertemuan	Aktivitas	Output
7	Pengumpulan data produksi padi dan data iklim	Data mentah produksi padi dan data iklim
8	Finalisasi pengumpulan data dan penyusunan database	Database terintegrasi produksi padi dan data iklim
9	Pembersihan data dan penanganan missing values	Dataset bersih siap untuk analisis
10	Eksplorasi data	Hasil eksplorasi data
11	Penerapan algoritma clustering untuk kabupaten/kota	Visualisasi hasil clustering dengan integrasi di R Shiny

12	Pengembangan model SARIMAX	Model prediksi produksi padi awal dengan integrasi di R Shiny
13	Pengembangan Fitur dan Tampilan Dashboard	Visualisasi dan Fitur Dashboard
14	Fiksasi UI dan Fitur R Shiny	R Shiny final

c. Kontribusi

Setiap tahapan dan bagian dapat dikerjakan secara bersama. Namun, setiap anggota mendapat tanggung jawab untuk bagiannya masing-masing.

1) Rifqi Muhadzib Ahdan

Merancang tampilan, Menyusun script, Evaluasi seluruh dashboard

2) Seto Haidar Yudhistira

Mengolah Data, Menyusun script, Evaluasi seluruh dashboard

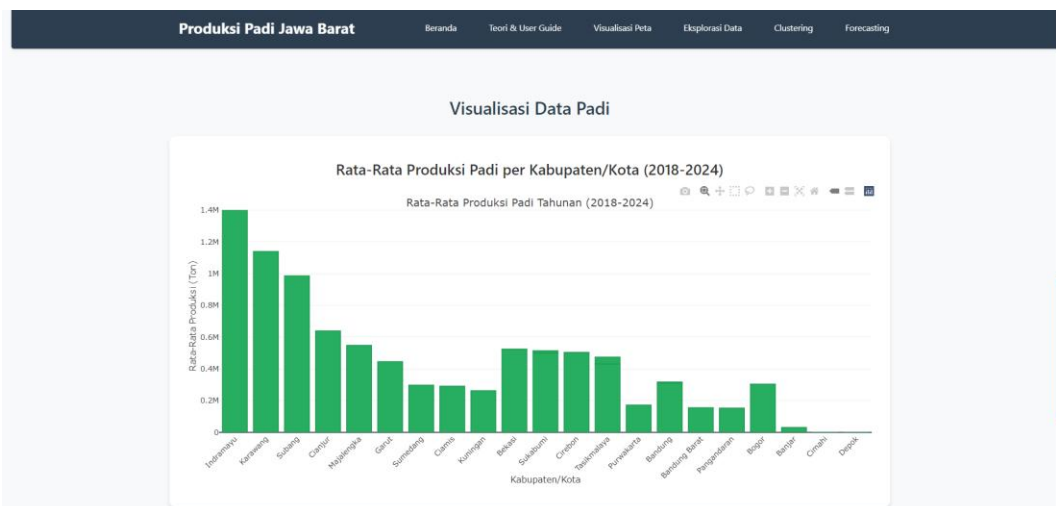
3) Zidan Septian

Membuat proposal, Menyusun script, Evaluasi seluruh dashboard

2.4 Hasil Akhir

a. Tampilan

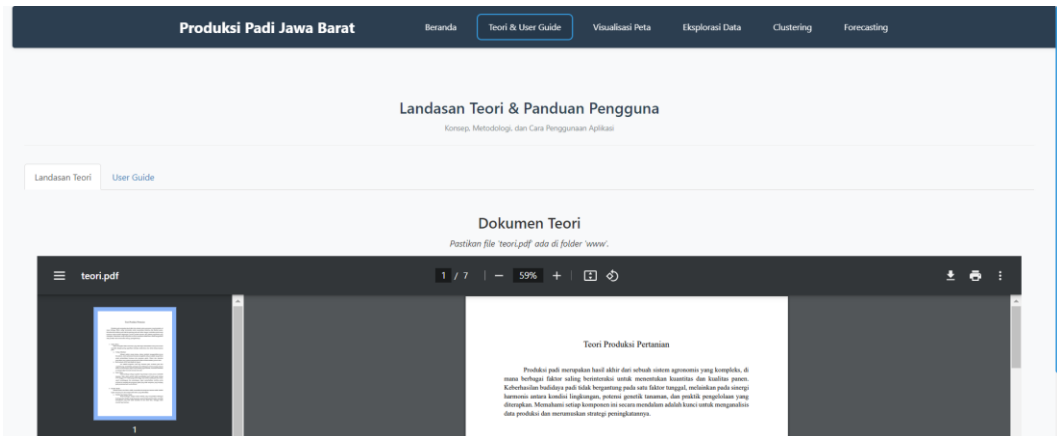
1) Beranda



Berisi tentang landasan teori yang berhubungan dengan produksi padi dan teori tentang peramalan menggunakan Model SARIMAX (Seasonal Autoregressive

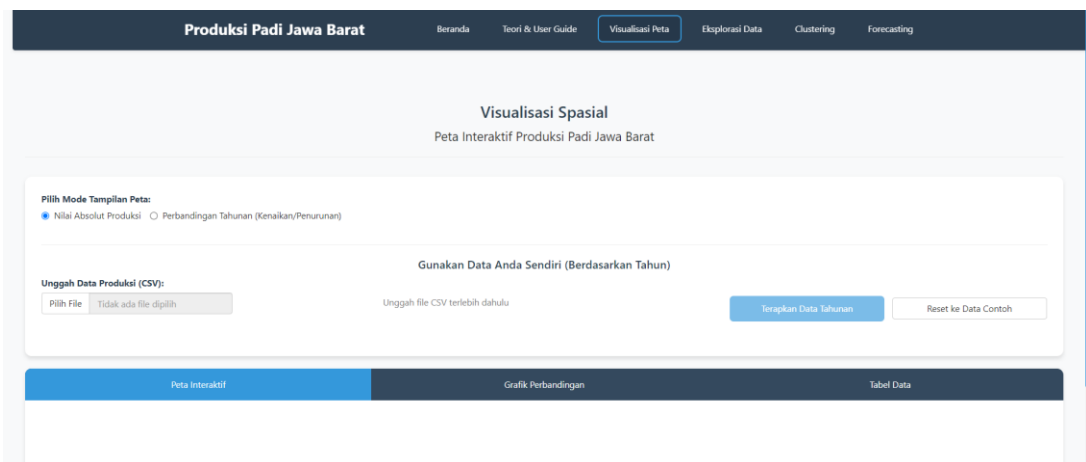
Integrated Moving Average with eXogenous variables). Halaman ini juga berisi tuntunan bagi pengguna dalam menjalankan aplikasi.

2) Teori & User Guide



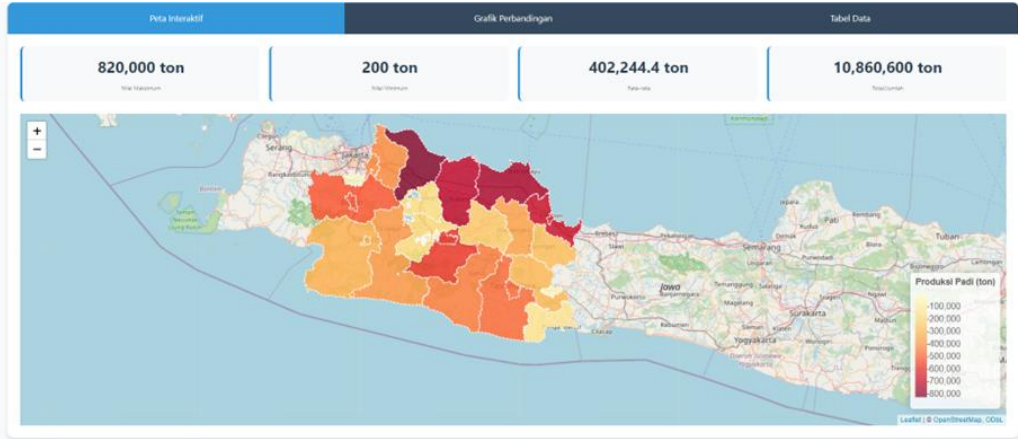
Berisi tentang landasan teori yang berhubungan dengan produksi padi dan teori tentang peramalan menggunakan Model SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous variables). Halaman ini juga berisi tuntunan bagi pengguna dalam menjalankan aplikasi.

3) Visualisasi Data

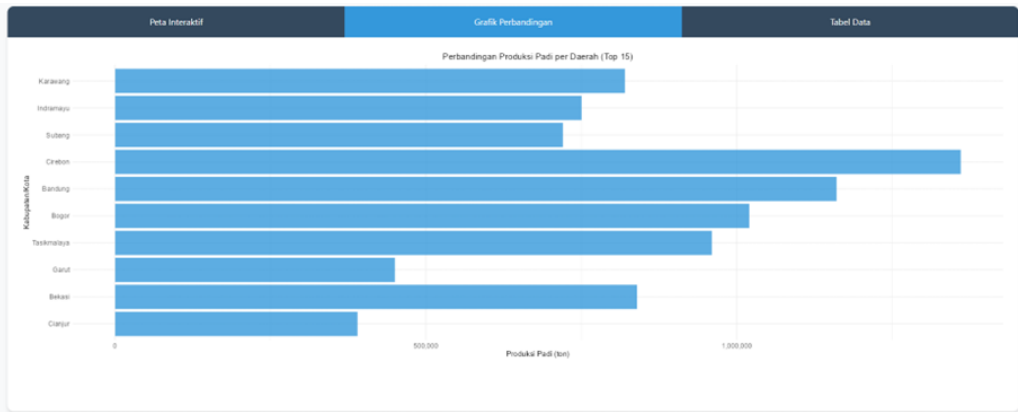


Tab untuk memvisualisasikan data yang ada menjadi grafik, peta, dan ringkasan tabel. Pengguna bisa mengunggah data .csv atau menggunakan data contoh yang telah disediakan. Akan ada 2 pilihan, yaitu data nilai absolut produksi padi atau perbandingan produksi padi antara tahun tertentu dengan tahun sebelumnya.

Peta Interaktif



Grafik Perbandingan



Tabel Data

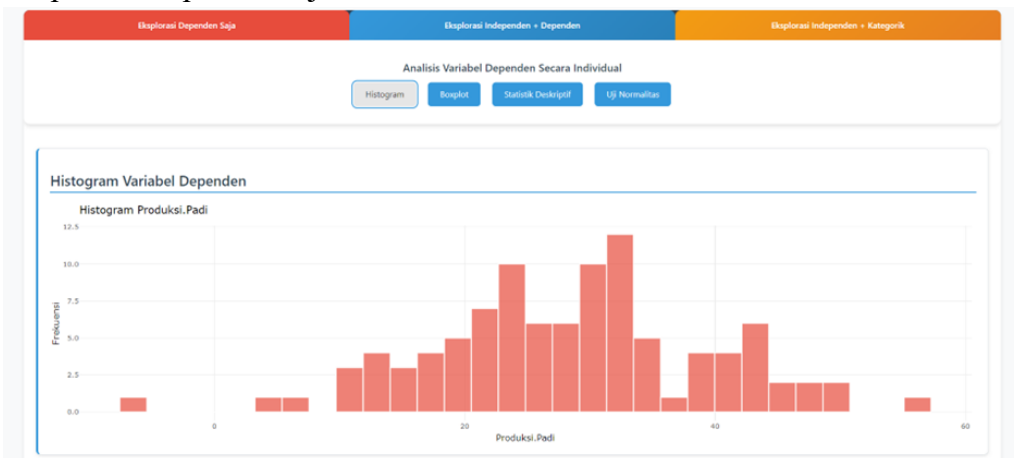
Peta Interaktif		Grafik Perbandingan		Tabel Data
Show 15 entries				Search:
Kabupaten	Produksi Padi (ton)			
Bogor	510,000			
Sukabumi	320,000			
Cianjur	390,000			
Bandung	580,000			
Garut	450,000			
Tasikmalaya	480,000			
Ciamis	300,000			
Pangandaran	150,000			
Kuningan	280,000			
Cirebon	680,000			
Majalengka	360,000			
Sumedang	250,000			
Indramayu	750,000			
Subang	720,000			
Purwakarta	180,000			
Showing 1 to 15 of 27 entries				Previous 1 2 Next

4) Eksplorasi Data

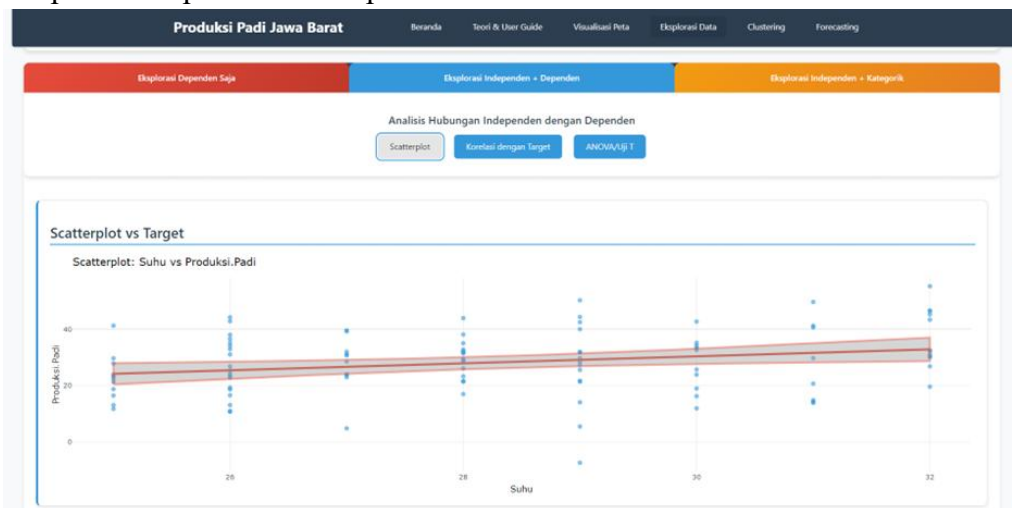
The screenshot shows the 'Eksplorasi Data' (Data Exploration) section of a web application titled 'Produksi Padi Jawa Barat'. The interface includes a navigation bar with links: Beranda, Teori & User Guide, Visualisasi Peta, Eksplorasi Data, Clustering, and Forecasting. Below the navigation bar, the main heading is 'Eksplorasi Data' with a subtitle 'Analisis Eksplorasi Data Interaktif dengan Drag & Drop'. The interface is divided into three main sections: 1. 'Upload Data untuk Eksplorasi': This section has a sub-heading 'Unggah File Data (CSV):' and a button 'Pilih File' next to a text input 'Data_Eksplorasi.csv'. Below this is a blue button 'Upload complete'. To the right, there is a blue button 'Report Data'. 2. 'Drag & Drop Variabel untuk Analisis': This section contains a 'Bank Variabel' with a sub-heading 'Drag variabel di bawah ini ke area yang sesuai:'. Below this are eight variable cards: 'Wilayah' (Nominal), 'Jenis.Padi' (Nominal), 'Luas.Lahan' (Nominal), 'Pupuk' (Nominal), 'Penggunaan.Air' (Nominal), 'Suhu' (Nominal), 'Curah.Hujan' (Nominal), and 'Produksi.Padi' (Nominal). Below the bank are three target areas: 'Variabel Dependen' (Drop variabel target di sini) with 'Produksi.Padi' inside; 'Variabel Independen' (Drop variabel analisis di sini (bisa multiple)) with 'Curah.Hujan' inside; and 'Variabel Kategorikal' (Drop variabel kategori di sini (opsional)) with 'Jenis.Padi' inside.

Pada tab ini, pengguna bisa mengunggah file berformat .csv atau menggunakan data contoh yang telah disediakan. Kemudian, pengguna bisa melakukan drag-and-drop setiap variabel dan memasukkannya atau mengelompokkannya ke Variabel Dependen, Variabel Independen, Variabel Kategorikal.

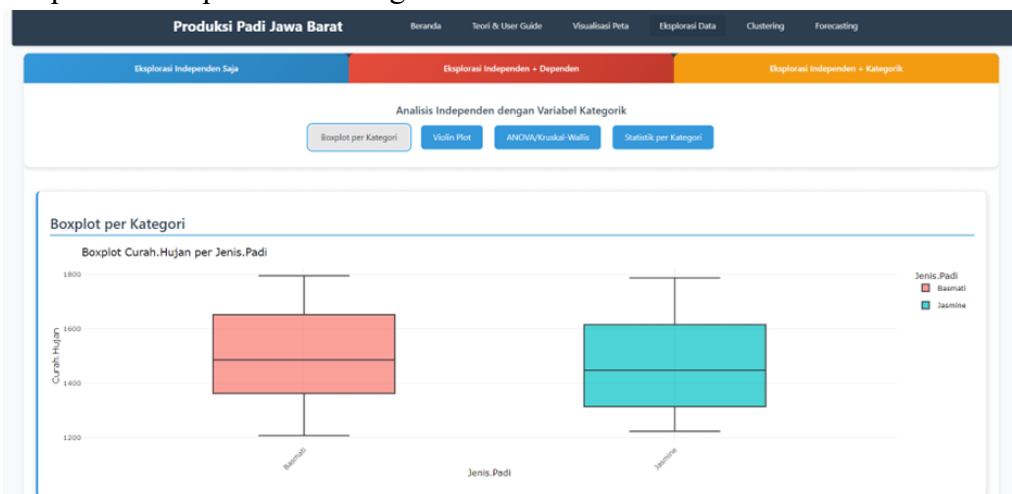
Eksplorasi Dependen saja



Eksplorasi Dependen + Independen



Eksplorasi Independen + Kategorik

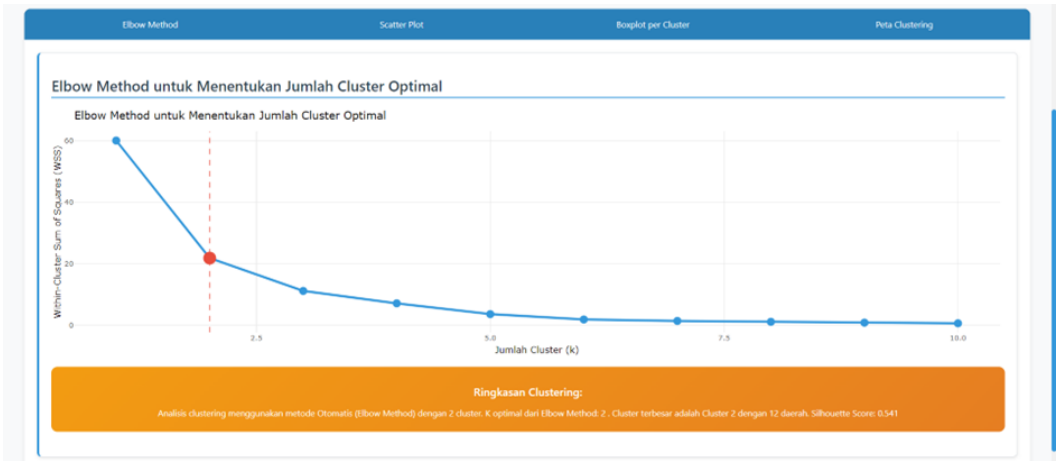


5) Clustering

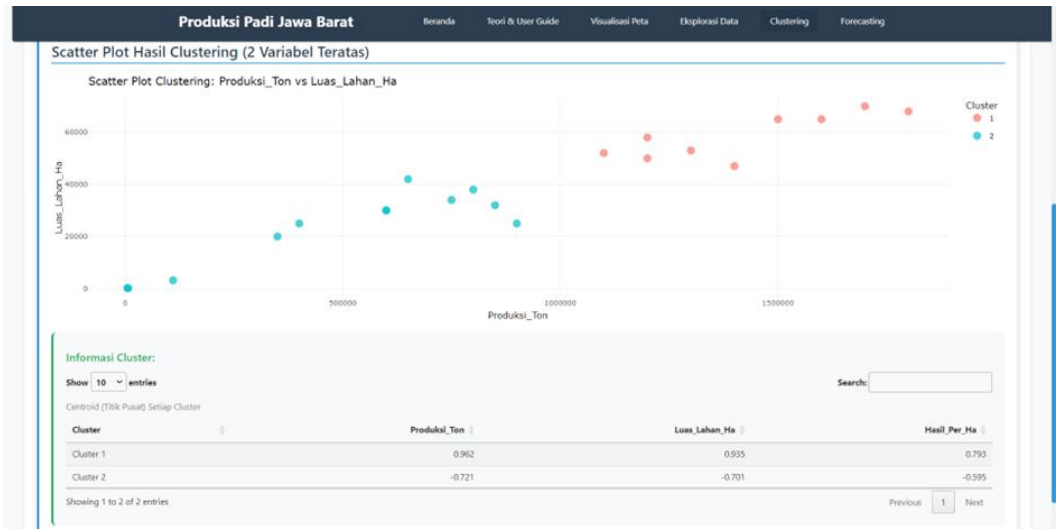
Tab ini berisikan analisis Clustering untuk pengelompokan daerah berdasarkan data padi. Pada tab ini, pengguna bisa mengunggah file berformat .csv atau menggunakan data contoh yang telah disediakan. Kemudian pengguna memilih variabel numerik untuk melakukan clustering dengan mencentang pilihan

yang ada dan selanjutnya memilih jumlah cluster (metode penentuan K). Jika pengguna memilih metode otomatis, akan digunakan Elbow Methods. Namun, pengguna juga bisa memilih metode manual (pengguna menentukan sendiri nilai K).

Elbow Method



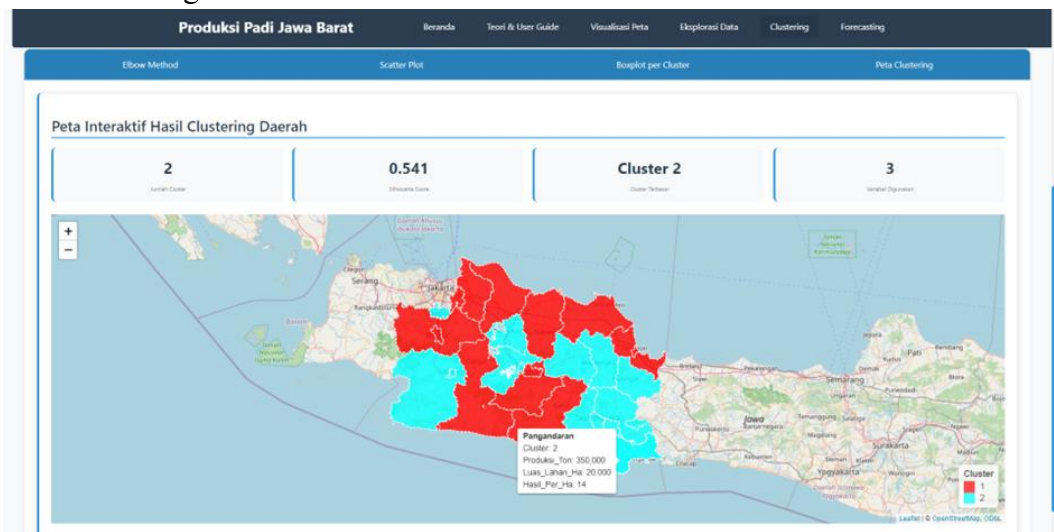
Scatter Plot



Boxplot per Cluster



Peta Clustering



6) Forecasting

Produksi Padi Jawa Barat

BerandaTeori & User GuideVisualisasi PetaEksplorasi DataClusteringForecasting

SARIMAX Forecasting

Model Peramalan Produksi Padi dengan Variabel Ekoogen

Input Data

Statistik Deskriptif

Matriks Korelasi

Pengecekan Asumsi

Forecasting & Tren

Format Data yang Diperlukan

Data harus dalam format CSV dengan kolom: No, Bulan, Tahun, Produksi Padi, Curah Hujan, Suhu, Kelembapan, Sinar Matahari

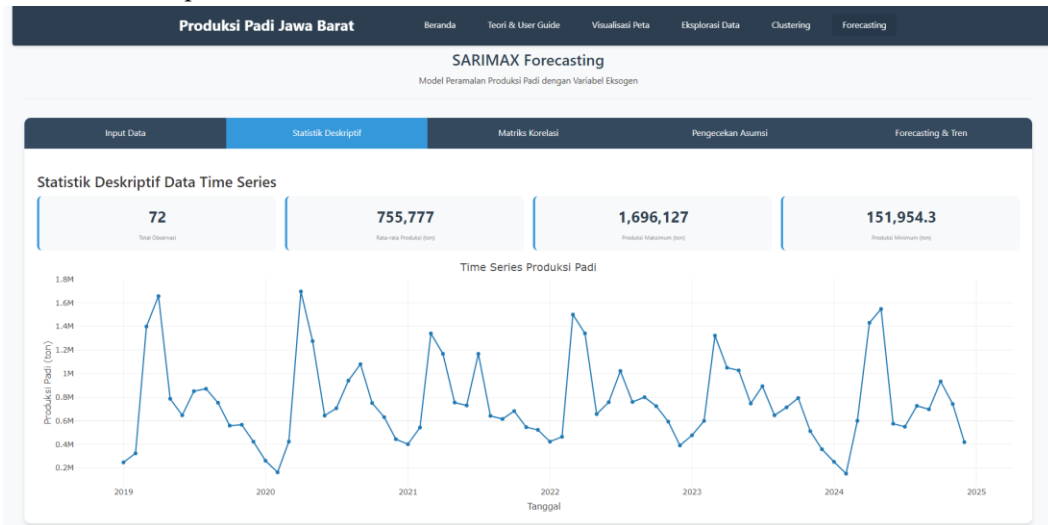
Pilih File CSV

Belum ada file dipilih

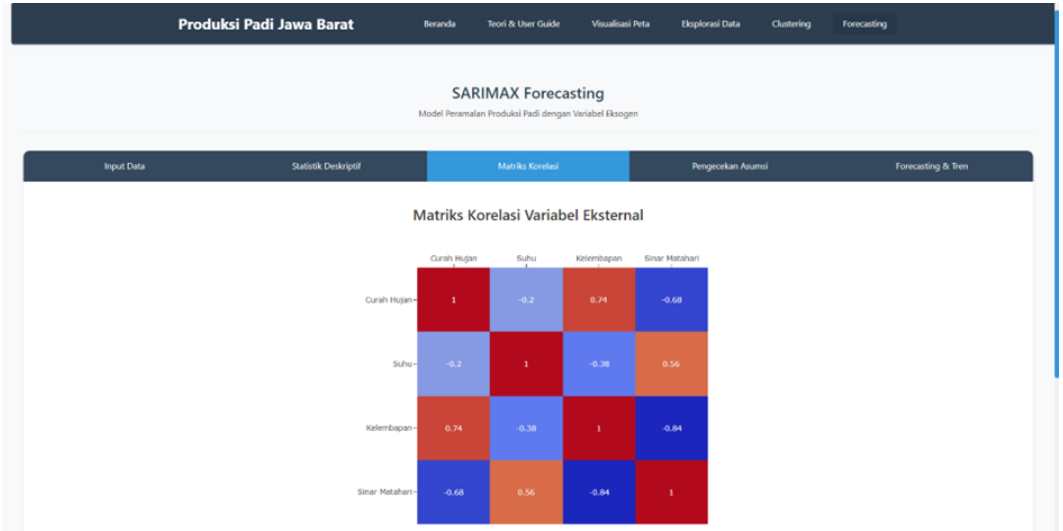
Gunakan Data Contoh

Pada tab ini, pengguna bisa mengunggah file berformat .csv yang sesuai dengan template atau menggunakan data contoh yang telah disediakan. Kemudian, pengguna bisa melakukan beberapa analisis dari mulai mengetahui Statistik deskriptif, melihat matriks korelasi, mengecek asumsi, hingga melihat tren dan peramalan dimasa mendatang.

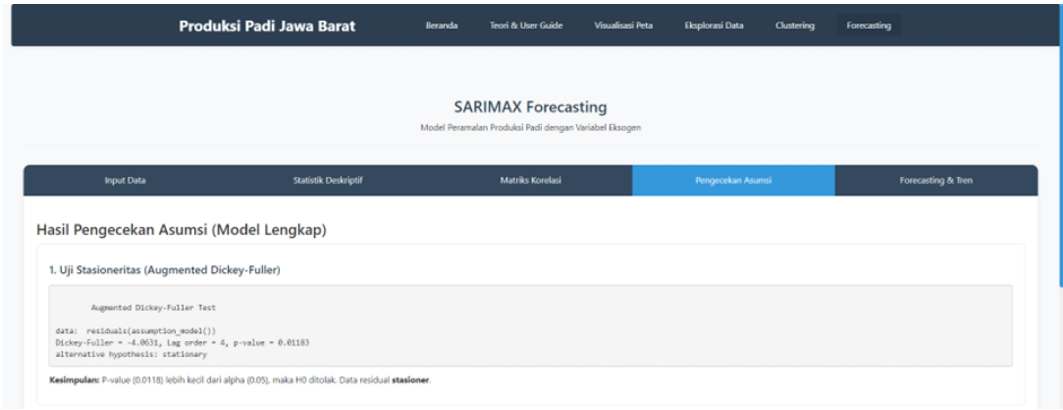
Statistik Deskriptif



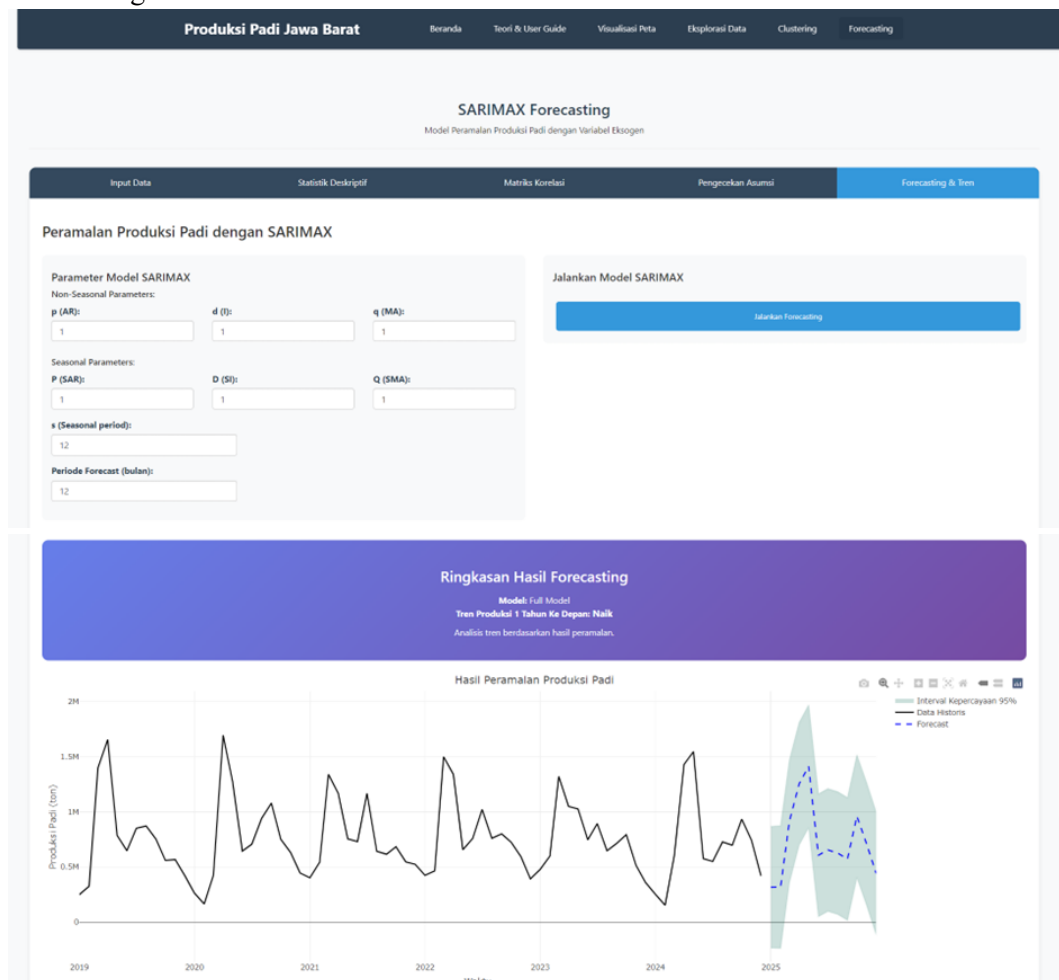
Matriks Korelasi



Pengecekan Asumsi



Forecasting & Tren



b. Tautan

1) Dashboard

<https://padi-jawabarat.shinyapps.io/padi-jawabarat/>

2) Github

<https://github.com/RifqiMuhA/produksipadi-jawabarat>

BAB III

PENUTUP

3.1 Kesimpulan

Perubahan iklim telah menimbulkan dampak signifikan terhadap pola produksi padi di Indonesia, termasuk di Jawa Barat sebagai salah satu lumbung padi nasional. Pergeseran musim tanam dan panen, peningkatan frekuensi gagal panen, dan fluktuasi produksi telah menjadi tantangan serius bagi ketahanan pangan nasional.

Penelitian ini diharapkan dapat memberikan pemahaman komprehensif tentang hubungan antara variabel iklim (suhu, kelembapan, curah hujan, penyinaran matahari) dengan produksi padi di Jawa Barat. Melalui pendekatan analisis data yang terintegrasi, penelitian ini akan menghasilkan model prediktif untuk memperkirakan produksi padi dan risiko gagal panen berdasarkan skenario perubahan iklim.

Dashboard interaktif yang dikembangkan akan menjadi alat pendukung keputusan bagi pemangku kepentingan dalam merencanakan strategi adaptasi dan mitigasi dampak perubahan iklim terhadap produksi padi. Hasil penelitian ini diharapkan dapat berkontribusi pada pengembangan kebijakan pertanian yang responsif terhadap perubahan iklim dan memperkuat ketahanan pangan nasional.

DAFTAR PUSTAKA

Data

Badan Pusat Statistik. (2020). Provinsi Jawa Barat Dalam Angka 2019. Diakses pada tanggal 27 April 2025 dari <https://jabar.bps.go.id/id/publication/2019/08/16/b0a6b953b76cbe2239dc912d/provinsi-jawa-barat-dalam-angka-2019.html>

Badan Pusat Statistik. (2021). Provinsi Jawa Barat Dalam Angka 2020. Diakses pada tanggal 27 April 2025 dari <https://jabar.bps.go.id/id/publication/2020/04/27/cfab9a400cf304f800182a5f/provinsi-jawa-barat-dalam-angka-2020.html>

Badan Pusat Statistik. (2022). Provinsi Jawa Barat Dalam Angka 2021. Diakses pada tanggal 27 April 2025 dari <https://jabar.bps.go.id/id/publication/2021/02/26/4d3f7ec6c519dda0b9785d45/provinsi-jawa-barat-dalam-angka-2021.html>

Badan Pusat Statistik. (2023). Provinsi Jawa Barat Dalam Angka 2022. Diakses pada tanggal 27 April 2025 dari <https://jabar.bps.go.id/id/publication/2022/02/25/0d261f828b581d8082bbc6c1/provinsi-jawa-barat-dalam-angka-2022.html>

Badan Pusat Statistik. (2024). Provinsi Jawa Barat Dalam Angka 2023. Diakses pada tanggal 27 April 2025 dari <https://jabar.bps.go.id/id/publication/2023/02/28/57231a828abbfdd50a21fe31/provinsi-jawa-barat-dalam-angka-2023.html>

Badan Pusat Statistik. (2025). Provinsi Jawa Barat Dalam Angka 2024. Diakses pada tanggal 27 April 2025 dari <https://jabar.bps.go.id/id/publication/2024/02/28/35ffe2d35104b39feb577e8f/provinsi-jawa-barat-dalam-angka-2024.html>

Jurnal

Amalina, T., Pramana, D., & Sari, B. (2022). Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food. Jurnal Ilmiah Wahana Pendidikan, 8(15),574-583

- Ananda Satria. (2020). *Visualisasi Data Iklim di Indonesia*. Diakses dari <https://anandasatria.github.io/visualisasi-data-iklim-di-indonesia/index.html#hero> pada 27 April 2025
- Galitan, J. H., Duko, F., & Hatim, F. (2024). Analisis Produksi Padi di Indonesia. *INNOVATIVE: Journal Of Social Science Research*, 4(2), 7281-7301.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts. Diakses dari <https://otexts.com/fpp2/index.html> pada 6 Juli 2025
- Mita, N. K. A., Siddiq, M.F., Laurant A., Erviana R., & Kurniawan, R. (2024). Optimalisasi Ketahanan Pangan: Perbandingan Metode Machine Learning dan Time Series dalam Memprediksi Produksi Padi di Jawa Tengah. *SENADA: Seminar Nasional Sains Data*.
- Ruminta. (2016). Indikasi Perubahan Iklim dan Dampaknya Terhadap Produksi Padi di Indonesia (Studi kasus: Sumatera Selatan dan Malang Raya). *Jurnal Agro*, 5(1), 48-60.