

2023

# Estimasi Ketersediaan Padi Nasional sebagai Upaya Meningkatkan Ketahanan Pangan Menggunakan Algoritma XGBoost

Team 1 Group 4/ Swasembada Team





# Our team

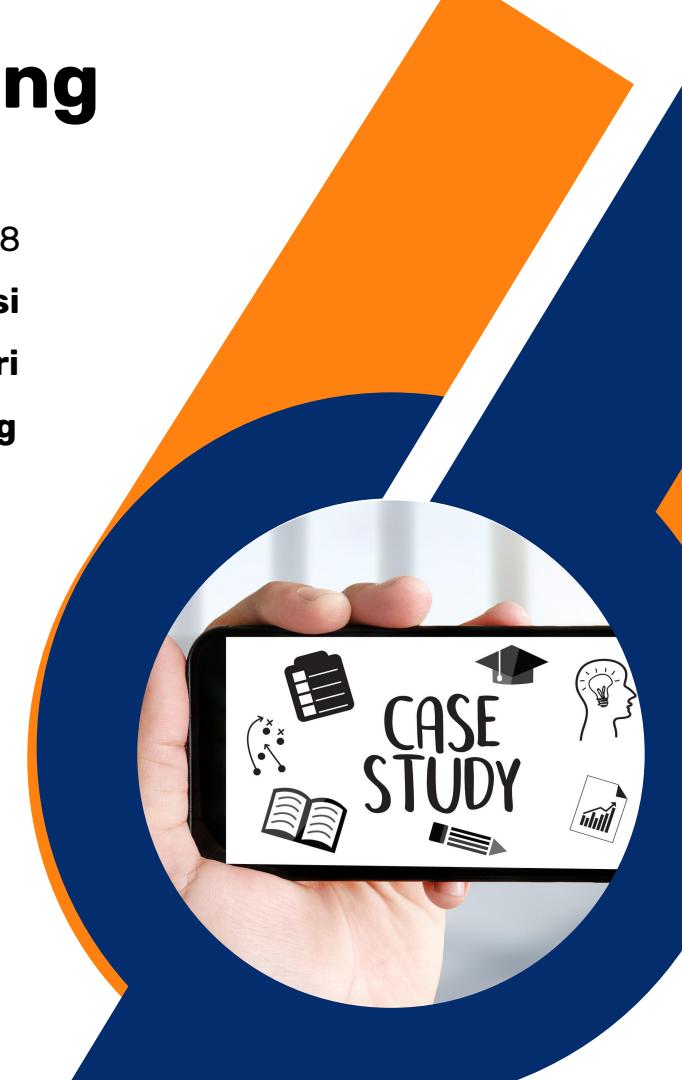


- Azka Fauzi A.P | Ketua Team Project
- Shintya Rahmawati | Data Visualisasi
- Kevin Caesar | Techinal Desain
- Zeni Fitri Adibatus | Data Visualisasi
- Nurulliza Camilla| Data Modeler
- Dede Risqi S | Data Engineer

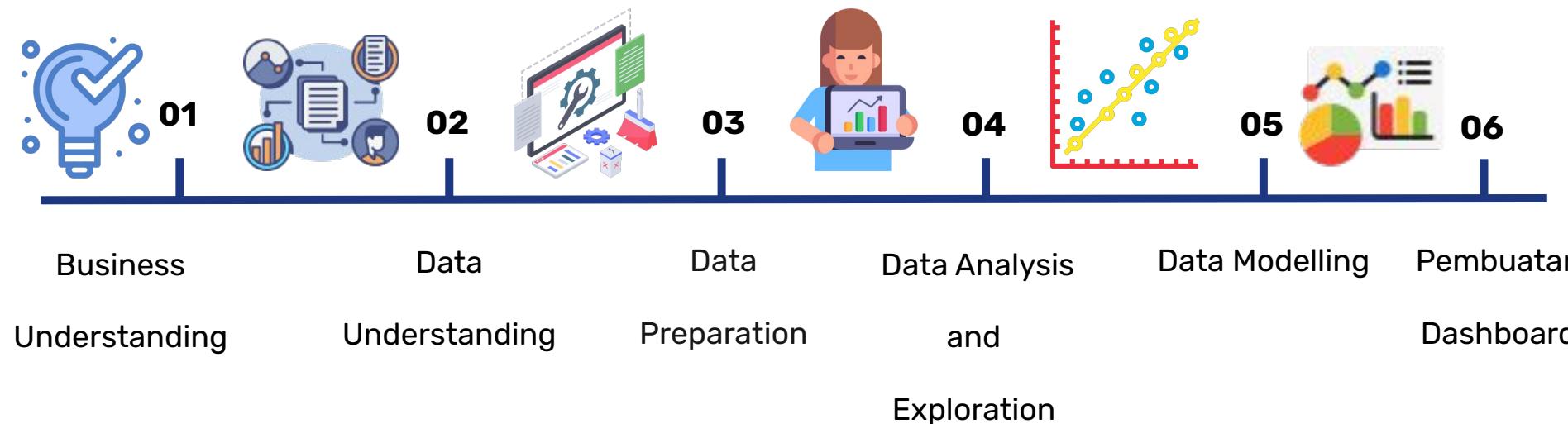
# Latar Belakang

Ketahanan pangan dalam Undang-Undang Republik Indonesia Nomor 18 tahun 2012 mengenai pangan didefinisikan sebagai **kondisi terpenuhinya pangan bagi rumah tangga yang tercermin dari tersedianya pangan yang cukup, baik jumlah maupun mutunya yang aman, merata, dan terjangkau.**

Adanya **upaya mewujudkan ketahanan pangan nasional** yang dihasilkan dari **ketersediaan, kontribusi yang merata, dan harga yang terjangkau terhadap pangan**. Maka, disini kami mencoba memberikan rekomendasi akan **Optimalisasi Kebijakan Produksi Padi**.



# Langkah-langkah





## Business Understanding

Upaya memastikan ketersedian pangan (padi) di Indonesia dan untuk mengatasi masalah defisitnya produksi beras di beberapa provinsi. Maka, dilakukan estimasi produksi padi pada tahun 2023 sebagai langkah awal untuk mengambil kebijakan dan tindakan preventif lainnya.

Faktor - faktor yang dinilai berpengaruh terhadap produksi yaitu luas panen, jumlah penduduk, konsumsi rumah tangga, suhu, harga beras dan penyinaran matahari.

# Our Goals



## Estimasi Produksi

---

Perkiraan jumlah produksi padi berdasarkan luas panen jika ditingkatkan 0,4%, jumlah penduduk, konsumsi rumah tangga, suhu, harga beras dan penyinaran matahari.

## Pengambilan Keputusan

---

Mengetahui provinsi mana saja yang berhasil surplus beras dan merekomendasikan saran terbaik kedepannya



## Data Understanding

- Data Collection

Pengumpulan data dilakukan dengan mengambil data dari website <https://www.bps.go.id/>, antara lain:

- Luas Panen, Produksi, dan Produktivitas Padi Menurut Provinsi 2018 - 2022
- Data konsumsi komoditas beras, perhitungan surplus dan defisit 2018 - 2022
- Suhu dan Penyinaran Matahari Menurut Provinsi 2018 - 2022



# Variable Data



Proyeksi Konsumsi  
Beras Rumah  
Tangga



Luas Panen dan  
Provinsi



Produksi Padi &  
produktivitas



Jumlah  
Penduduk



Produksi Beras



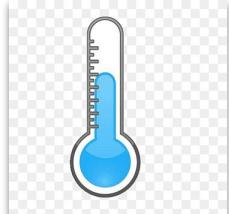
Surplus atau Defisit  
beras



Harga Beras



Lama Penyinaran  
Matahari



Suhu



## Data Preparation

Data berasal dari file terpisah, yang kemudian digabungkan menjadi gambar dibawah ini :

provinsi	kode_provinsi	tahun	luas panen (ha)	produktivitas (ku/ha)	produksi (ton)	jumlah penduduk	luas provinsi (ha)	produksi beras (ton)	proyeksi konsumsi rumah tangga(ton)	surplus/defisit(ton)	surplus/defisit(%)*	harga beras(Rp)	suhu	curah hujan (mm)	penyiraman matahari (%)	
0	ACEH	11	2018	329515.78	56.49	1861567.10	5243400	5795600	1066890.0	557210.000000	509680.000000	91.470002	10182	27.08	254.76	60.18
1	ACEH	11	2019	310012.46	55.30	1714437.60	5316300	5795600	982570.0	443379.000000	539191.000000	121.609503	11134	26.90	135.34	65.74

Total keseluruhan  
data yang ada

```
[ ] df.shape  
(170, 16)
```

## Data Cleaning

Menghilangkan missing value, mensinkronkan format data.





# Data Preparation

- Data Cleansing

Mengedit nama kolom, menghilangkan missing value, mensinkronkan format data.

```
[ ] new_columns = {'luas panen (ha)': 'luas_panen',
                 'produktivitas (ku/ha)': 'produktivitas',
                 'produksi (ton)': 'produksi',
                 'jumlah penduduk ': 'jumlah_penduduk',
                 'luas provinsi (ha)': 'luas_provinsi',
                 'produksi beras (ton)': 'produksi_beras',
                 'proyeksi konsumsi rumah tangga(ton)': 'konsumsi_rumah_tangga',
                 'surplus/ defisit(ton)': 'surplus_defisit',
                 'surplus/ defisit(%)*': 'presentase_surplus_defisit',
                 'harga beras(Rp)': 'harga_beras',
                 'curah hujan (mm)': 'curah_hujan',
                 'penyinaran matahari (%)': 'penyinaran_matahari'}
df = df.rename(columns=new_columns)
```



# Data Preparation

- Data Cleaning

Mengedit nama kolom, menghilangkan missing value, mensinkronkan format data.

```
[ ] df.isnull().sum()
```

```
provinsi          0
kode_provinsi    0
tahun            0
luas_panen       0
produktivitas     0
produksi          0
jumlah_penduduk  0
luas_provinsi    0
produksi_beras   0
konsumsi_rumah_tangga  0
surplus_defisit  0
presentase_surplus_defisit  0
harga_beras       0
suhu              0
curah_hujan      126
penyinaran_matahari 0
dtype: int64
```

```
[ ] df1 = df.drop(['provinsi', 'kode_provinsi', 'tahun', 'luas_provinsi', 'curah_hujan'], axis=1)
df1.head()
```

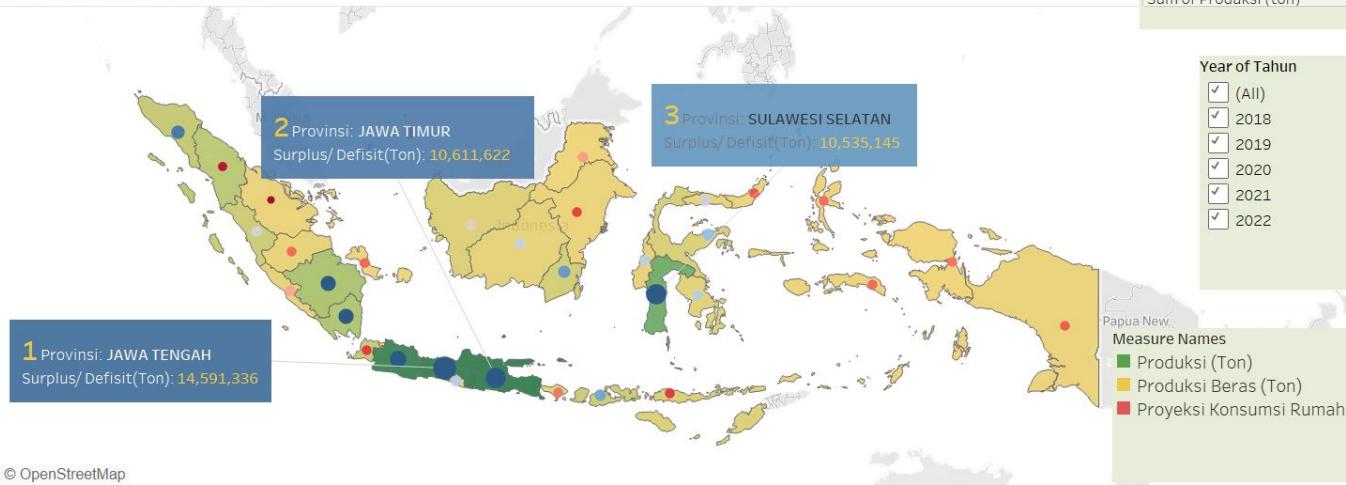
```
[ ] df1.isnull().sum()
```

```
luas_panen          0
produktivitas        0
produksi             0
jumlah_penduduk     0
produksi_beras       0
konsumsi_rumah_tangga  0
surplus_defisit      0
presentase_surplus_defisit  0
harga_beras          0
suhu                 0
penyinaran_matahari 0
dtype: int64
```

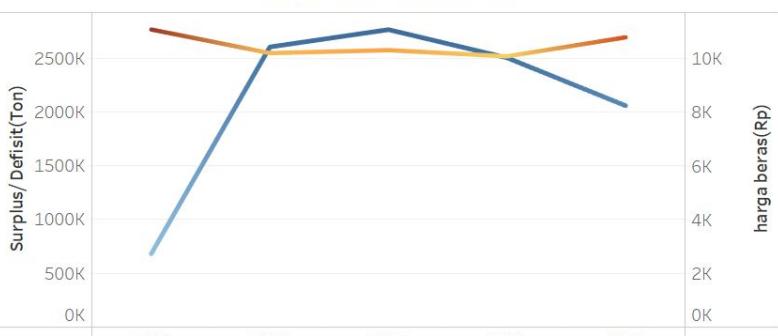
# Analisa Data

Produksi Padi Nasional dari tahun 2018 -2022

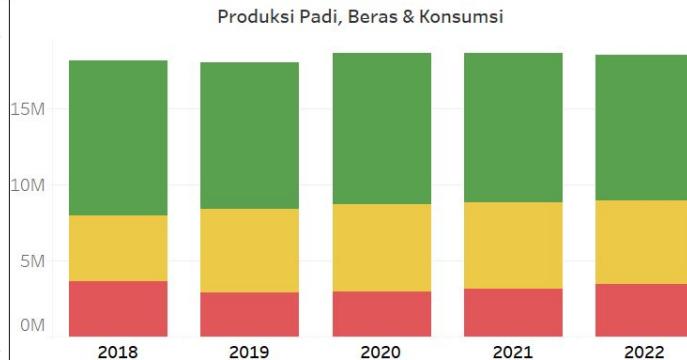
Sum of Produksi (ton) menurut Provinsi



Surplus Beras dan Harga Beras



Produksi Padi, Beras & Konsumsi

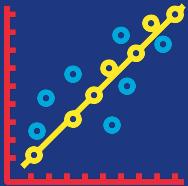


Provinsi
(All)
ACEH
BALI
BANTEN
BENGKULU
DI YOGYAKARTA
DKI JAKARTA
GORONTALO
JAMBI
JAWA BARAT
JAWA TENGAH
JAWA TIMUR
KALIMANTAN BARAT
KALIMANTAN SELATAN
KALIMANTAN TENGAH
KALIMANTAN TIMUR
KALIMANTAN UTARA
KEP. BANGKA BELITUNG
KEP. RIAU
LAMPUNG
MALUKU
MALUKU UTARA
NUSA TENGGARA BARAT
NUSA TENGGARA TIMUR
PAPUA
PAPUA BARAT
RIAU
SULAWESI BARAT
SULAWESI SELATAN
SULAWESI TENGAH
SULAWESI TENGGARA
SULAWESI UTARA
SUMATERA BARAT
SUMATERA SELATAN
SUMATERA UTARA



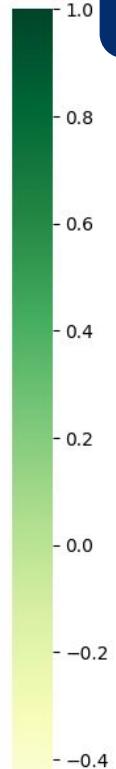
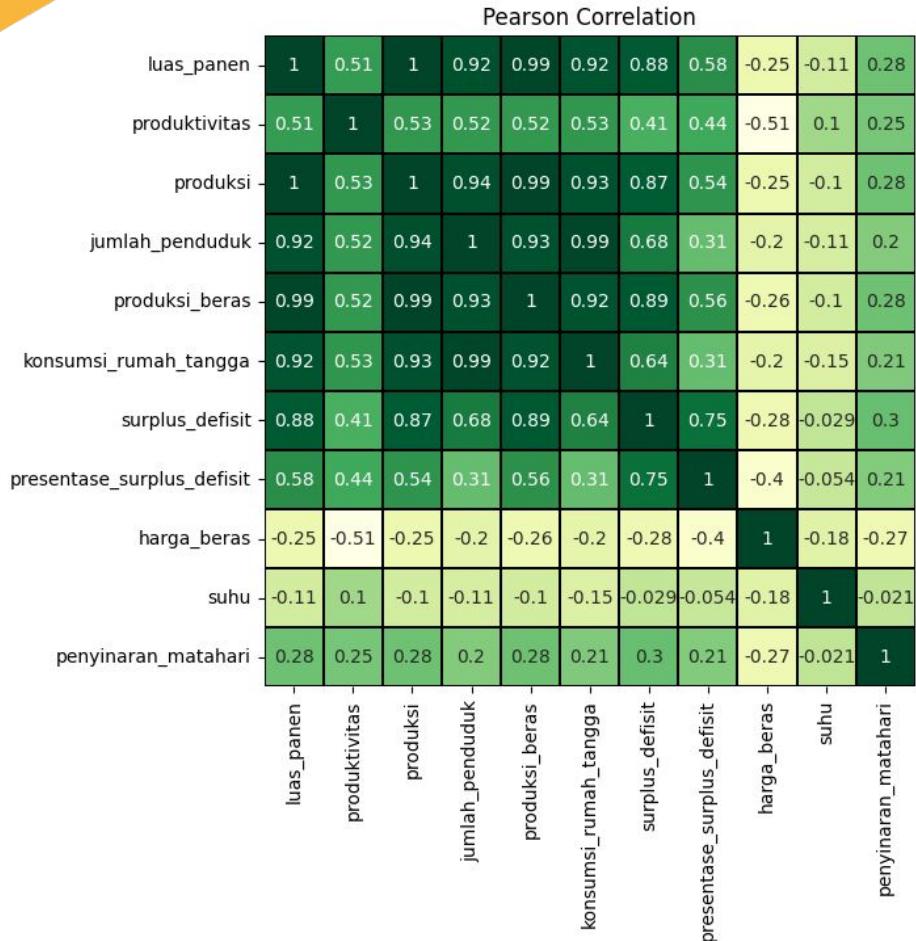
# Data Modelling

## Algoritma : XGboost



Algoritma yang dikembangkan dari gradient boosting. Gradient boosting merupakan algoritma yang dapat menemukan solusi yang optimal untuk berbagai masalah khususnya pada regresi, klasifikasi dan ranking. XGBoost menggunakan model yang lebih teratur untuk membangun struktur pohon regresi, sehingga dapat memberikan kinerja yang lebih baik dan mampu mengurangi kompleksitas model untuk menghindari overfitting.

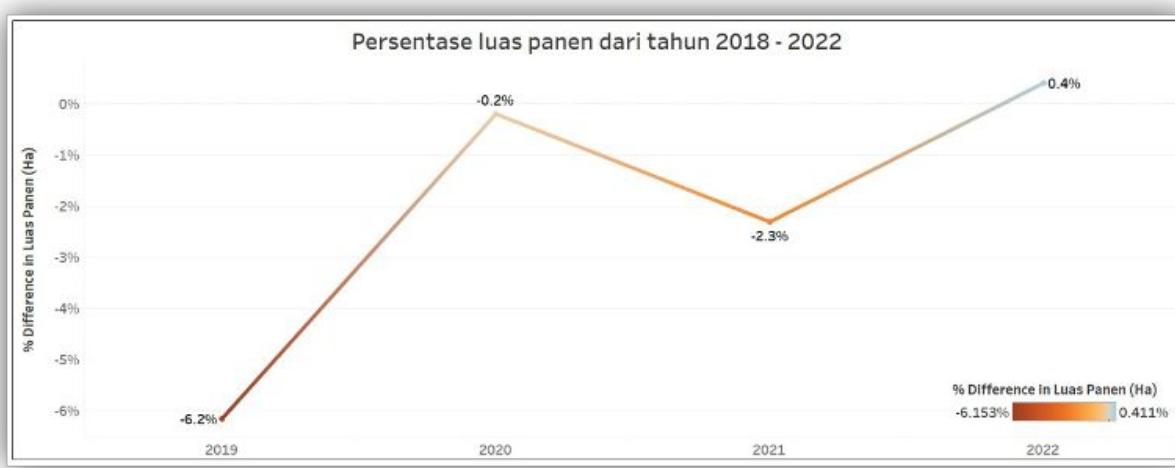
# Korelasi Variabel



## IMPLEMENTASI MODEL XGBOOST

### Data yang digunakan

Data yang digunakan untuk estimasi merupakan data asumsi, yang diperoleh dari peningkatan pada luas panen sebesar 0,4% dari tahun 2022. Peningkatan tersebut diambil berdasarkan grafik di bawah ini :





## R Square & MAPE

R Square : 0.909

MAPE : 14.00424551825613

Nilai R Square sebesar 0,909 atau 90.9% menunjukkan bahwa variabel x berpengaruh terhadap variabel y sebesar 90.9%. Sedangkan nilai MAPE sebesar 14,00 menunjukkan bahwa kemampuan model baik untuk melakukan peramalan.

## Hasil Estimasi

52758116.0

total_prod	
tahun	
2018	59200533.72
2019	54604033.34
2020	54649202.24
2021	54415294.22
2022	54748975.85

Diperoleh nilai estimasi untuk total produksi pada tahun 2023 yaitu sebesar 52.758.116 ton. Dimana nilai tersebut mengalami penurunan dari tahun sebelumnya.

## Prediksi Time - Series Model ARIMA

ARIMA atau Auto Regressive Integrated Moving Average adalah model yang banyak dipakai dalam peramalan data time series univariat. Sesuai namanya, model ARIMA terdiri 3 komponen yaitu Auto Regressive (AR), Integrated (I) dan Moving Average (MA) dan dinotasikan sebagai ARIMA(p, d, q). Hasil dari estimasi 2023 selanjutnya dilakukan pemodelan time - series menggunakan ARIMA untuk mengetahui produksi padi 5 tahun kedepan.

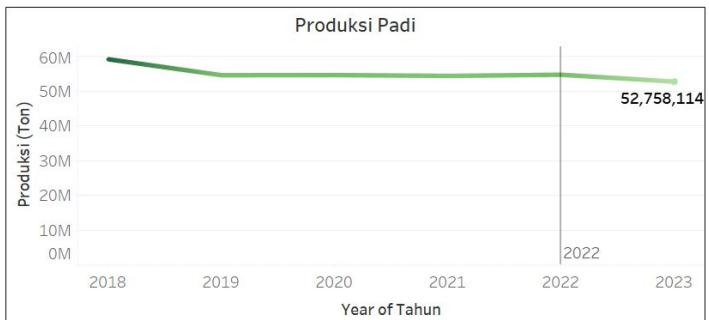
### Hasil Prediksi

ARIMA Predictions	
2024-12-31 00.00.00	55.599.362,01
2025-12-31 00.00.00	54.937.717,36
2026-12-31 00.00.00	55.091.795,31
2027-12-31 00.00.00	55.055.915
2028-12-31 00.00.00	55.064.270,49

Diperoleh prediksi data Total Produksi untuk Tahun 2024 - 2028 di Indonesia.

# Dashboard hasil prediksi

Hasil Prediksi untuk Produksi Padi



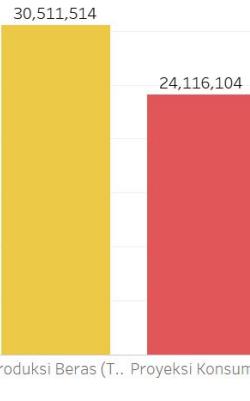
- Provinsi
- (All)
  - ACEH
  - BALI
  - BANTEN
  - BENGKULU
  - DI YOGYAKARTA
  - DKI JAKARTA
  - GORONTALO
  - JAMBI
  - JAWA BARAT
  - JAWA TENGAH
  - JAWA TIMUR
  - KALIMANTAN BARAT
  - KALIMANTAN SELATAN
  - KALIMANTAN TENGAH
  - KALIMANTAN TIMUR
  - KALIMANTAN UTARA
  - KEP. BANGKA BELITUNG
  - KEP. RIAU
  - LAMPUNG
  - MALUKU
  - MALUKU UTARA
  - NUSA TENGGARA BARAT
  - NUSA TENGGARA TIMUR
  - PAPUA
  - PAPUA BARAT
  - RIAU
  - SULAWESI BARAT
  - SULAWESI SELATAN
  - SULAWESI TENGAH
  - SULAWESI TENGGARA
  - SULAWESI UTARA
  - SUMATERA BARAT
  - SUMATERA SELATAN
  - SUMATERA UTARA

Produksi Padi & Surplus Beras Tahun 2023



produksi beras & konsumsi

Year of Tahun  
2023



# Kesimpulan

Pada tahun 2023 terdapat 16 provinsi yg bisa surplus:

JAWA TENGAH, JAWA TIMUR, JAWA BARAT, SUMATERA SELATAN, LAMPUNG, SULAWESI SELATAN, ACEH, SUMATERA BARAT, NTB, SULAWESI TENGAH, KALIMANTAN SELATAN, DI YOGYAKARTA, KALIMANTAN TENGAH, SULAWESI BARAT, SULAWESI TENGGARA, GORONTALO

Dari hasil prediksi untuk 5 tahun kedepan yang didapatkan, maka kami dapat menyarankan:

1. Provinsi dengan penghasil padi tinggi **dapat membantu memenuhi** kebutuhan padi pada provinsi dengan produksi rendah
2. **Peningkatan produktivitas padi di lahan yang tersedia** sehingga dapat membantu dalam pemenuhan produksi padi, dengan menerapkan beberapa teknik pertanian yang lebih optimal dan intensif seperti pertanian vertikal, hidroponik, atau sistem pertanian berkelanjutan
3. **Memastikan distribusi beras dapat berjalan lancar** pada provinsi yang surplus ke provinsi yang membutuhkan agar tidak terjadi kenaikan harga yg berlebih
4. **Konservasi Tanah dan Pengelolaan Air**, yaitu dengan mengurangi penggunaan bahan kimia yang berlebihan dan menggunakan teknologi irigasi yang efisien.
5. **Peran pemerintah** untuk mengadakan riset dan Inovasi di bidang pertanian, juga melakukan pelatihan bagi petani untuk mempelajari praktik terbaru dalam budidaya padi, pengelolaan lahan, teknologi pertanian, dan strategi yang dapat membantu Anda menghadapi tantangan penurunan luas lahan pertanian padi.



# Thank You

Swasembada Team.

DATA ACADEMY™