



DØLab

# TETRIS PROGRAM

---

[Analisa Produksi Kopi di Indonesia]

[Rapika Rosalia]

[Rapika.Rosalia@gmail.com]

**#StackYourSkill**

# TETRIS PROGRAM

## Step 1 - Data Collection & Data Integration

Data untuk analisis produksi kopi di Indonesia diambil dari situs resmi Satuan Data Pertanian ([satudata.pertanian.go.id](http://satudata.pertanian.go.id)). Sumber data ini dipilih karena merupakan sumber resmi yang disediakan oleh Kementerian Pertanian Republik Indonesia, sehingga dapat diandalkan untuk analisis yang akurat dan terpercaya.

1. Langkah-langkah Pengumpulan Data
  - Mencari Topik yang ingin di Analisa, pada tahap ini di pilih adalah produksi kopi di Indonesia hasil panen dari Perkebunan kopi di indonesia
  - mencari sumber data dan data di ambil dari [satudata.pertanian.go.id](http://satudata.pertanian.go.id)
  - Mengambil Data- Seleksi Data
2. Data yang digunakan merupakan External Data- Data berasal dari website <https://satudata.pertanian.go.id/>

Lampiran Screenshot untuk mendukung proses ini.

Collapsible

Show  entries

Search:

No	Judul	Abstraksi	Tanggal Rilis	Owner	Lihat Rincian
16	Luas Areal Kopi Menurut Provinsi di Indonesia, 2014-2021	Luas Areal Kopi Menurut Provinsi di Indonesia, 2014-2021	Selasa, 10 Agustus 2021	KEMENTERIAN PERTANIAN	<a href="#">Lihat Rincian</a>
32	Produktivitas Kopi Menurut Provinsi di Indonesia, 2014-2021	Produktivitas Kopi Menurut Provinsi di Indonesia, 2014-2021	Selasa, 10 Agustus 2021	KEMENTERIAN PERTANIAN	<a href="#">Lihat Rincian</a>
47	Produksi Kopi Menurut Provinsi di Indonesia, 2014-2021	Produksi Kopi Menurut Provinsi di Indonesia, 2014-2021	Selasa, 10 Agustus 2021	KEMENTERIAN PERTANIAN	<a href="#">Lihat Rincian</a>

Showing 1 to 3 of 3 entries (filtered from 156 total entries)

Previous **1** Next

# TETRIS PROGRAM



## Step 2 - Data Cleansing

1. Jelaskan langkah-langkah yang dilakukan untuk membersihkan data
  - mengambil setiap data pertahun(tahun 2014-2019) dan mengurutkannya menjadi satu kolom
  - Mengabungkan semua data menjadi satu table yaitu table data kopi dan menyimpan data dalam csv
  - Mengunaka MYSQL membuat nama table baru dengan nama totalproduksikopi yang akan di gunakan untuk Analisa
  - Menghubungkan SQL dengan phyton
  - Load Dataset
  - **Memeriksa Kelengkapan Data:** Langkah pertama dalam pembersihan data adalah memeriksa kelengkapan data untuk memastikan tidak ada nilai yang hilang atau data yang tidak lengkap.

### **Pembersihan Data:**

1. **Identifikasi dan Penanganan Nilai-nilai yang Hilang:** Setelah memeriksa kelengkapan data, langkah selanjutnya adalah mengidentifikasi dan menangani nilai-nilai yang hilang. Jika ada baris data yang memiliki nilai-nilai yang hilang, baris tersebut akan dihapus atau nilai yang hilang akan diisi dengan metode yang sesuai.
2. **Identifikasi dan Penanganan Data Duplikat:** Data duplikat dapat mengganggu analisis, oleh karena itu, langkah selanjutnya adalah mengidentifikasi dan menghapus data duplikat dari dataset.

### **Proses Manipulasi:**

1. **Menentukan Karakteristik Data:** Setelah membersihkan data, langkah berikutnya adalah menentukan karakteristik data seperti rata-rata, median, dan deviasi standar.
2. **Menangani Outlier:** Outlier dapat mempengaruhi hasil analisis, oleh karena itu, teknik winsorization digunakan untuk menangani outlier dalam dataset.

### **Deskripsi Statistik dan Distribusi Nilai:**

1. **Deskripsi Statistik:** Menampilkan deskripsi statistik dari data seperti mean, median, dan quartil untuk memahami distribusi data secara umum.
2. **Distribusi Nilai:** Melakukan visualisasi distribusi nilai dari kolom Produktivitas\_Kopi dan Luas\_Areal\_Kopi untuk memahami pola distribusi data.

#StackYourSkill

# TETRIS PROGRAM

mengambil setiap data pertahun(tahun 2014-2019) dan mengurutkannya menjadi satu kolom  
Menggabungkan semua data menjadi satu table yaitu table data kopi dan menyimpan data dalam csv

Produktivitas Kopi Menurut Provinsi di Indonesia, 2014-2021						
Coffee Yield by Province in Indonesia, 2014-2021						
No.	Provinsi/Province	2014	2015	2016	2017	2018
1	Aceh	772	698	717	754	756
2	Sumatera Utara	1.034	1.050	1.141	1.072	1.084
3	Sumatera Barat	960	962	748	763	812
4	Riau	838	900	897	923	945
5	Kepulauan Riau	63	-	-	-	-
6	Jambi	827	816	825	874	934
7	Sumatera Selatan	651	532	578	889	912
8	Kepulauan Bangka Belitung	259	155	375	282	437
9	Bengkulu	749	745	746	787	794
10	Lampung	692	788	824	773	798
11	DKI Jakarta	-	-	-	-	-
12	Jawa Barat	880	866	856	746	844
13	Banten	484	492	365	536	519
14	Jawa Tengah	850	754	626	567	699
15	DI. Yogyakarta	507	498	514	431	495
16	Jawa Timur	754	836	786	791	793
17	Bali	596	632	622	493	538
18	Nusa Tenggara Barat	576	570	568	600	605
19	Nusa Tenggara Timur	501	521	535	514	530
20	Kalimantan Barat	518	528	534	533	529
21	Kalimantan Tengah	647	615	641	620	608
22	Kalimantan Selatan	574	635	640	623	634
23	Kalimantan Timur	229	215	242	255	244

Produksi Kopi Menurut Provinsi di Indonesia, 2014-2021								
Coffee Production by Province in Indonesia, 2014-2021								
No.	Provinsi/Province	2014	2015	2016	2017	2018	2019	2020
1	Aceh	49.823	47.444	65.231	68.493	70.774	72.652	70.774

Luas Areal Kopi Menurut Provinsi di Indonesia, 2014-2021								
Coffee Area by Province in Indonesia, 2014-2021								
No.	Provinsi/Province	2014	2015	2016	2017	2018	2019	2020
1	Aceh	120.667	121.272	123.443	123.807	124.236	125.331	126.045
2	Sumatera Utara	81.644	82.156	85.412	90.241	93.095	95.374	95.477
3	Sumatera Barat	42.893	42.925	38.365	33.276	31.622	26.696	25.244

Provinsi	Produktivitas_Kopi	Luas_Areal_Kopi	Produksi_Kopi	Tahun
Aceh	772,09	120666,9	49822,64	2014
Sumatera Utara	1033,76	81644	58951	2014
Sumatera Barat	960,23	42893,47	33556,27	2014
Riau	838,06	4712,41	2465,04	2014
Kepulauan Riau	62,5	47	1	2014
Jambi	827,4	26094	13582,64	2014
Sumatera Selatan	651,04	249381,41	135287,3	2014
Kepulauan Bangka Belitung	259,42	24,65	4,06	2014
Bengkulu	749,05	90883	56459,6	2014
Lampung	691,62	155007	92111	2014
DKI Jakarta	0	0	0	2014
Jawa Barat	880,03	32486,47	17531,68	2014
Banten	484,04	6576,55	2511	2014
Jawa Tengah	850,22	40675	24883	2014
DI. Yogyakarta	507,17	1741,1	440,92	2014
Jawa Timur	753,8	102213,38	58135,27	2014
Bali	596,34	36078	15920,42	2014
Nusa Tenggara Barat	576,24	12397,34	4817,91	2014
Nusa Tenggara Timur	501,25	72208	21508	2014
Kalimantan Barat	517,89	11763	3720	2014
Kalimantan Tengah	646,67	1752,9	371,01	2014
Kalimantan Selatan	574,5	4043	1284	2014
Kalimantan Timur	228,83	4823	562	2014
Kalimantan Utara	362,56	2177	306	2014
Sulawesi Utara	527,93	7714,14	2992,15	2014
Gorontalo	875,17	1755,9	953,94	2014
Sulawesi Tengah	912,35	8372,7	3572	2014
Sulawesi Selatan	630,39	72584	30107	2014
Sulawesi Barat	623,53	15786,73	6361,12	2014
Sulawesi Tenggara	439,91	9482,3	2921,37	2014
Maluku	539,72	1448,93	432,14	2014
Maluku Utara	220,06	2344	384	2014
Papua	407,97	10214	1771	2014
Papua Barat	486,36	504	128,4	2014

#StackYourSkill

# TETRIS PROGRAM

Mengunaka MYSQL membuat nama table baru dengan nama totalproduksikopi yang akan di gunakan untuk Analisa

```
1 • use kopi_indonesia;
2
3 • CREATE TABLE totalproduksikopi (
4     Provinsi VARCHAR(50),
5     Produktivitas_Kopi DECIMAL(10,2),
6     Luas_Areal_Kopi DECIMAL(15,2),
7     Produksi_Kopi DECIMAL(15,2),
8     Tahun YEAR(4)
9 );
10
11 • LOAD DATA INFILE '/path/to/data.txt'
12 INTO TABLE totalproduksikopi
13 FIELDS TERMINATED BY ';'
14 ENCLOSED BY ''
15 LINES TERMINATED BY '\n';
16
17 • SELECT * FROM totalproduksikopi INTO OUTFILE '/path/to/data.csv'
18 FIELDS TERMINATED BY ','
19 ENCLOSED BY ''
20 LINES TERMINATED BY '\n';
```

#StackYourSkill



# TETRIS PROGRAM

## Load Dataset dan Memeriksa Kelengkapan Data

```
dtypes: float64(3), int64(1), object(1)
memory usage: 8.1+ KB
C:\Users\ASUS\analisa_kopi\analysakopi.py:29: UserWarning:
ion) or database string URI or sqlite3 DBAPI2 connection
QLAlchemy.
    df = pd.read_sql(query, connection)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 204 entries, 0 to 203
Data columns (total 5 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Provinsi               204 non-null    object
1   Produktivitas_Kopi     204 non-null    float64
2   Luas_Areal_Kopi        204 non-null    float64
3   Produksi_Kopi          204 non-null    float64
4   Tahun                  204 non-null    int64
dtypes: float64(3), int64(1), object(1)
memory usage: 8.1+ KB
█
```

#StackYourSkill

# TETRIS PROGRAM



## Pembersihan Data dan Proses Manipulasi:

```
# Identifikasi dan tangani nilai-nilai yang hilang
data.dropna(inplace=True) # Menghapus baris yang memiliki nilai-nilai yang hilang

# Identifikasi dan tangani data duplikat
data.drop_duplicates(inplace=True) # Menghapus data duplikat

# 2. Proses Manipulasi

# Menentukan data karakteristik
data = data[['Provinsi', 'Produktivitas_Kopi', 'Luas_Areal_Kopi', 'Produksi_Kopi', 'Tahun']]

# Menangani outlier (misalnya, menggunakan teknik winsorization)
from scipy.stats.mstats import winsorize
data['Produksi_Kopi'] = winsorize(data['Produksi_Kopi'], limits=[0.05, 0.05]) # Winsorization pada produksi kopi

# Menampilkan setiap hasil pra-pemrosesan data
print("Data Setelah Pembersihan:")
print(data.head())
```

Data Setelah Pembersihan:

	Provinsi	Produktivitas_Kopi	Luas_Areal_Kopi	Produksi_Kopi	Tahun
0	Aceh	772.09	120666.90	49822.64	2014
1	Sumatera Utara	945.16	81644.00	58951.00	2014
2	Sumatera Barat	945.16	42893.47	33556.27	2014
3	Riau	838.06	4712.41	2465.04	2014
4	Kepulauan Riau	62.50	47.00	1.00	2014

#StackYourSkill

# TETRIS PROGRAM

## Deskripsi Statistik

### Deskripsi Statistik dan Distribusi Nilai:

1. Deskripsi Statistik:  
Menampilkan deskripsi statistik dari data seperti mean, median, dan kuartil untuk memahami distribusi data secara umum.
2. Distribusi Nilai:  
Melakukan visualisasi distribusi nilai dari kolom Produktivitas\_Kopi dan Luas\_Areal\_Kopi untuk memahami pola distribusi data.

```
Deskripsi Statistik Data:
C:\Users\ASUS\.virtualenvs\ASUS-mVsQ382S\Lib\site-packages\numpy\lib\function_base.py:4824: UserWarning: Warning: 'partition' will ignore the 'mask' of the MaskedArray.
  arr.partition(
C:\Users\ASUS\.virtualenvs\ASUS-mVsQ382S\Lib\site-packages\numpy\lib\function_base.py:4824: UserWarning: Warning: 'partition' will ignore the 'mask' of the MaskedArray.
  arr.partition(
C:\Users\ASUS\.virtualenvs\ASUS-mVsQ382S\Lib\site-packages\numpy\lib\function_base.py:4824: UserWarning: Warning: 'partition' will ignore the 'mask' of the MaskedArray.
  arr.partition(

```

	Produktivitas_Kopi	Luas_Areal_Kopi	Produksi_Kopi	Tahun
count	204.000000	204.000000	204.000000	204.000000
mean	574.799216	33638.769608	18906.906765	2016.500000
std	240.042002	45577.647634	29560.998758	1.712026
min	62.500000	7.000000	0.440000	2014.000000
25%	466.442500	1755.150000	411.112500	2015.000000
50%	607.255000	9818.500000	3057.000000	2016.500000
75%	754.107500	43534.695000	22443.987500	2018.000000
max	945.160000	156918.500000	107219.000000	2019.000000

```

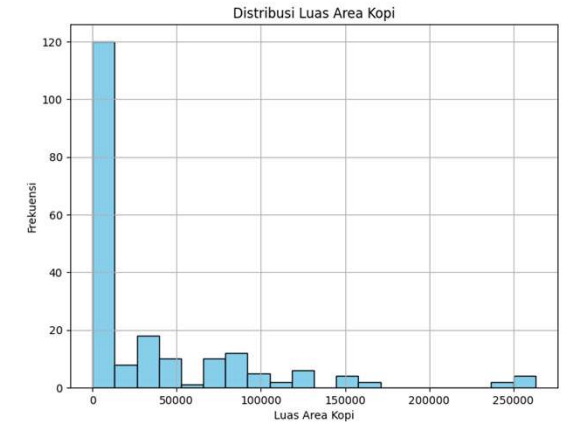
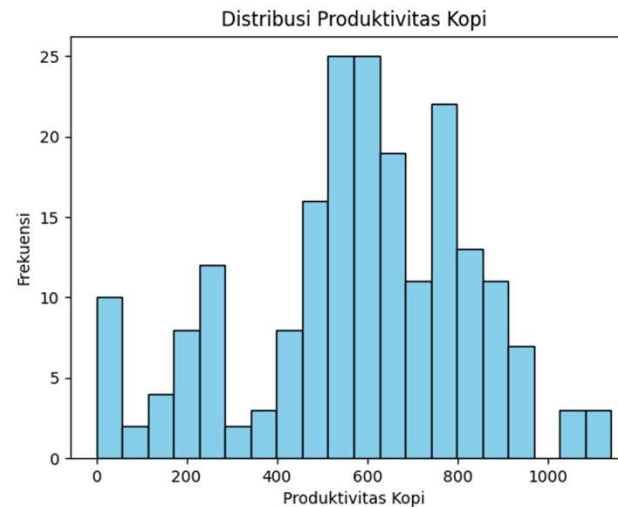
0 / 0  0 0  Share Code Link  Explain Code  Comment Code  Find Bugs  Code Chat  Blackbox  Search Error  Python 3.11.4 ('ASUS-mVsQ382S': F
```



## Deskripsi Statistik

### Deskripsi Statistik dan Distribusi Nilai:

1. Distribusi Nilai:  
Melakukan visualisasi distribusi nilai dari kolom Produktivitas\_Kopi dan Luas\_Areal\_Kopi untuk memahami pola distribusi data.
- Dan pada grafik dapat di Tarik kesimpulan bahwa distribusi produktivitas kopi berada di antara 400-800 kg/hektar dan, Distribusi luas area kopi terkonsentrasi pada area-area kecil. Hal ini menunjukkan bahwa banyak petani kopi memiliki kebun kopi yang tidak terlalu luas.





## Step 3 - Data Exploration & Data Visualisation

1. langkah-langkah yang dilakukan untuk melakukan eksplorasi data dan menghasilkan visualisasi data

Pada tahap ini proses dan tools yang di gunakan adalah phyton dengan probel yang di Analisa adalah

- Menghitung total dari produksi kopi,luas area kopi, produktivitas kopi pertahun dan membandingkan dengan tahun sebelumnya, dengan tujuan untuk melihat pertumbuhan peroduksi kopi tersebut
- Membuat visualisasi distribusi produksi kopi per provinsi tujuan untuk melihat perovinsi mana yang memiliki produksi kopi paling tinggi
- Membuat visualisasi pengaruh luas area dengan produksi kopi tujuan melihat apakah luas area berpengaruh terhadap produksi kopi
- Membuat visualisassi kontribusi provinsi terhadap produksi kopi nasional tujuan, melihat provinsi mana yang paling banyak berkontiri busi

1. Lampirkan screenshot untuk mendukung proses ini

# TETRIS PROGRAM



## Step 3 - Data Exploration & Data Visualisation

Menghitung total dari produksi kopi, luas area kopi, produktivitas kopi pertahun dan membandingkan dengan tahun sebelumnya, dengan tujuan untuk melihat pertumbuhan perproduksi kopi tersebut

Pilih Tahun

2015



Total Produksi

633,181

↑ kg

Luas Areal

1,230,000

↑ ha

Produktivitas

51.48

↑ kg/ha

Perubahan Total Produksi (2014 - 2015):

17,394 kg

Perubahan Luas Areal (2014 - 2015):

-495 ha

Perubahan Produktivitas (2014 - 2015):

1.43 %

Pada tahun **2015**, total produksi kopi di Indonesia mencapai **633,181 kg** dengan total luas areal **1,230,000 ha**. Produktivitas kopi rata-rata adalah **51.48 kg/ha**. Perubahan:\*\* Dibandingkan dengan tahun **2014**, terdapat peningkatan dalam total produksi kopi sebesar **17,394 kg**. Luas areal kopi mengalami penurunan sebesar **495 ha**, dan produktivitas kopi mengalami peningkatan sebesar **1.43%**.

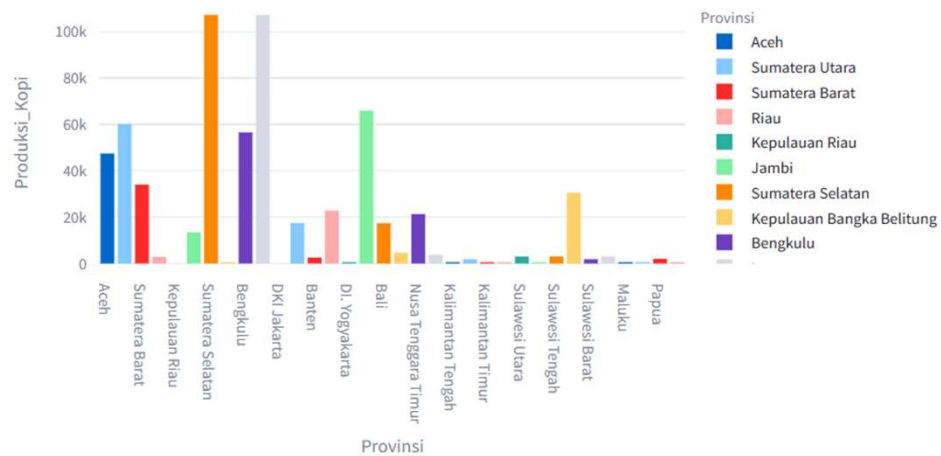
#StackYourSkill

# TETRIS PROGRAM

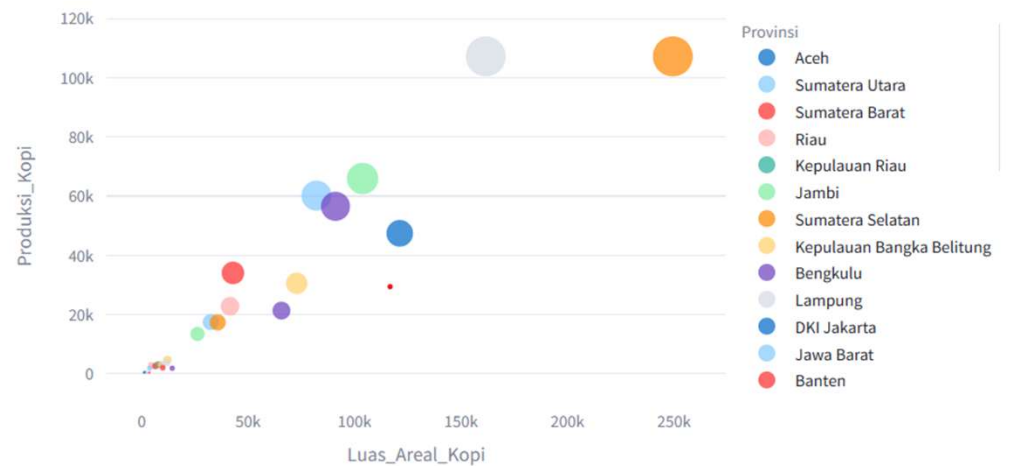
## Step 3 - Data Exploration & Data Visualisation

- Membuat visualisasi distribusi produksi kopi per provinsi tujuan untuk melihat perovinsi mana yang memiliki produksi kopi paling tinggi
- Membuat visualisasi pengaruh luas area dengan produksi kopi

Distribusi Produksi Kopi per Provinsi



Pengaruh Luas Areal terhadap Produksi Kopi

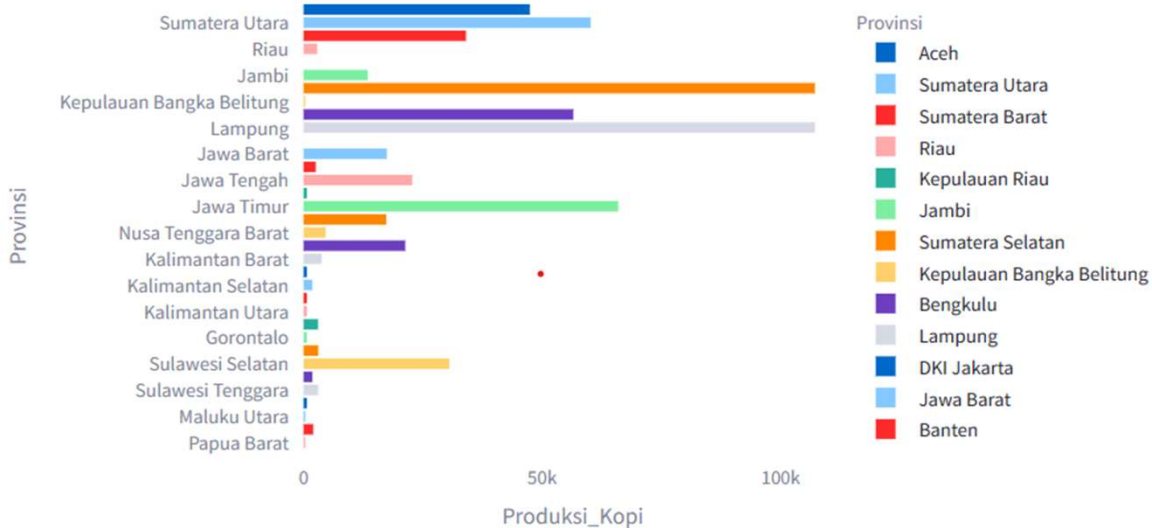


# TETRIS PROGRAM

## Step 3 - Data Exploration & Data Visualisation

- Membuat visualisasi kontribusi provinsi terhadap produksi kopi nasional tujuan, melihat provinsi mana yang paling banyak berkontribusi

Kontribusi Provinsi terhadap Produksi Nasional





# TETRIS PROGRAM



## Step 3 - Insight Analysis

- Setelah menjelajahi berbagai aspek produksi kopi di Indonesia, kita dapat menyimpulkan bahwa Sumatera Selatan, Jambi, dan Lampung menjadi tiga provinsi dengan produksi kopi tertinggi, sementara Sumatera Utara menjadi yang tertinggi dalam produktivitas kopi. Penemuan ini menggarisbawahi pentingnya keragaman geografis dalam kontribusi terhadap industri kopi nasional. Selain itu, analisis juga menegaskan adanya hubungan positif antara luas areal dan produksi kopi, memperkuat pentingnya pengembangan dan investasi dalam infrastruktur pertanian.
- Sebagai langkah yang lebih berkelanjutan dan produktif, diperlukan upaya bersama antara pemerintah, petani, dan pemangku kepentingan lainnya. Pemerintah harus melanjutkan dan meningkatkan upaya untuk meningkatkan produksi kopi nasional dengan strategi yang inklusif dan berkelanjutan. Sementara itu, petani kopi di seluruh negeri perlu didukung dalam mengadopsi teknologi modern dan praktik pertanian terbaik guna meningkatkan produktivitas dan kualitas kopi mereka. Dengan kerja sama dan komitmen bersama, kita dapat memastikan bahwa kopi Indonesia terus menjadi andalan global dan sumber kebanggaan bagi bangsa.

### Note

- perlu melakukan analisa terhadap pengaruh perubahan cuaca terhadap produksi kopi  
pengaruh harga dan hama/penyakit terhadap produksi kopi.

#StackYourSkill

DØLab

# **AYO #STACKYOURSKILL SEKARANG**

dan Persiapkan Diri Menjadi Praktisi Data!

---

