

Integrasi Model Analisis: Regresi Logistik dan FB Prophet pada Tingkat Pencemaran Udara di DKI Jakarta

Swipe





Profile Team



GreatEdu
Home for All Education Needs



Kampus
Merdeka
INDONESIA JAYA



**Ahmad Haical
Ramadhan**

Universitas Pamulang
Team Leader



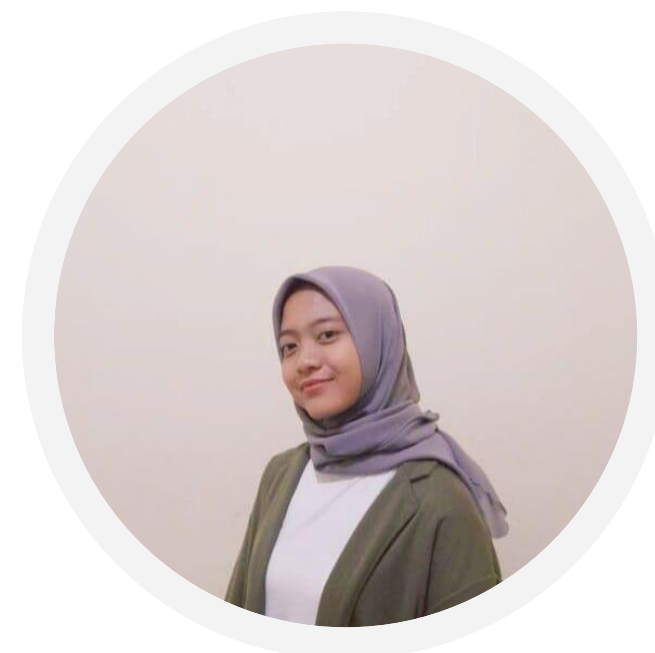
**I Made Wiraharja
Jaya Putra**

Universitas Udayana
Analysis



Wian Aliyanti

Institut Keguruan dan Ilmu
Pendidikan Siliwangi
Visualisasi



**Shinry Zhafirah
Achmad**

Universitas Negeri Semarang
Visualisasi

Project Overview



01 Business Understanding

02 Data Understanding

03 Data Preparation

04 Modelling

05 Evaluation

06 Deployment



GreatEdu
Home for All Education Needs

MSIB
magang dan studi independen bersertifikat

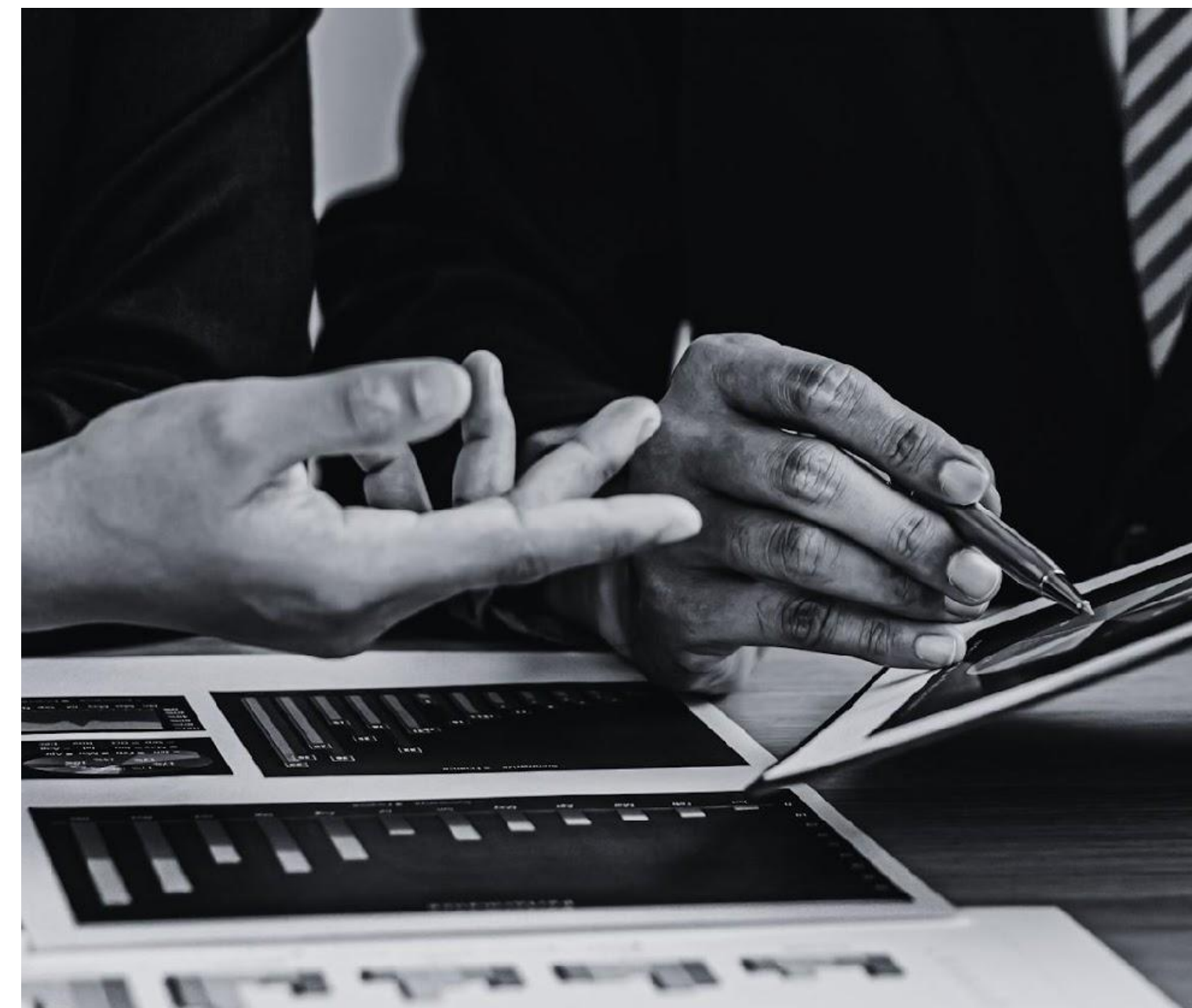
Kampus
Merdeka
INDONESIA JAYA



Business Understanding

Plant Heaven Nursery merupakan usaha yang bergerak dalam bidang pembudidayaan dan perawatan tanaman hias yang berada di DKI Jakarta. Pada saat ini Plant Heaven Nursery melakukan inovasi dengan menambah koleksi tanamannya yaitu Tanaman Sansevieria. Tanaman Sansevieria merupakan tanaman yang memiliki nilai estetika dan juga dikenal dapat menyaring polusi udara.

Kami tim data analyst akan memprediksikan daerah di DKI Jakarta untuk menentukan lokasi awal penjualan tanaman sansevieria berdasarkan daerah jakarta yang memiliki polusi tertinggi.





Tujuan

Dari dataset yang kami dapatkan di website Open Jakarta, kami tim data analyst dapat mengambil tindakan untuk menentukan lokasi penjualan Tanaman Sansevieria, berdasarkan variabel lokasi yang relevan.



Manfaat

Setelah tercapainya tujuan, diharapkan dapat meningkatkan penjualan tanaman sansevieria dan meminimalisir polusi di wilayah DKI Jakarta.



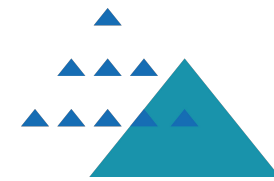
GreatEdu
Home for All Education Needs

MSIB
magang dan studi independen bersertifikat

Kampus
Merdeka
INDONESIA JAYA



Data Understanding



Menggunakan dataset Indeks Standar Pencemaran Udara (ISPU) yang diperoleh yang diukur dari 5 stasiun pemantau kualitas udara (SPKU) yang berlokasi di DKI Jakarta pada tahun 2021. Dataset ini diperoleh dari website <https://data.jakarta.go.id/dataset/indeks-standar-pencemaran-udara-ispu-tahun-2021>

Dataset ini terdiri dari 1825 baris dan 11 kolom. Berikut data yang digunakan :

Website

www.greatedu.co.id

SIB Cycle 5 | **Data Analyst**



Data Understanding



NO	Variable	Keterangan
1	tanggal	Tanggal pengukuran kualitas udara
2	stasiun	Lokasi pengukuran di stasiun
3	pm_{10}	Partikulat salah satu parameter yang diukur
4	$pm_{2,5}$	Partikulat salah satu parameter yang diukur
5	so_2	Sulfida (dalam bentuk SO_2) salah satu parameter yang diukur



Data Understanding



NO	Variable	Keterangan
6	co	Carbon Monoksida salah satu parameter yang diukur
7	o ₃	Ozon salah satu parameter yang diukur
8	no ₂	Nitrogen dioksida salah satu parameter yang diukur
9	max	Nilai ukur paling tinggi dari seluruh parameter yang diukur dalam waktu yang sama
10	critical	Parameter yang hasil pengukurannya paling tinggi
11	categorical	Kategori hasil perhitungan indeks standar pencemaran udara



GreatEdu
Home for All Education Needs

MSIB
Mangung dan Studi Independen Bersertifikat

Kampus
Merdeka
INDONESIA JAYA



Data Understanding

	tanggal	stasiun	pm10	pm2.5	so2	co	o3	no2	max	critical	categori
0	1/1/2021	DKI1 (Bunderan HI)	38.0	53.0	29.0	6.0	31.0	13.0	53	PM25	SEDANG
1	1/2/2021	DKI1 (Bunderan HI)	27.0	46.0	27.0	7.0	47.0	7.0	47	O3	BAIK
2	1/3/2021	DKI1 (Bunderan HI)	44.0	58.0	25.0	7.0	40.0	13.0	58	PM25	SEDANG
3	1/4/2021	DKI1 (Bunderan HI)	30.0	48.0	24.0	4.0	32.0	7.0	48	PM25	BAIK
4	1/5/2021	DKI1 (Bunderan HI)	38.0	53.0	24.0	6.0	31.0	9.0	53	PM25	SEDANG
...
1820	12/27/2021	DKI5 (Kebon Jeruk) Jakarta Barat	54.0	76.0	36.0	14.0	21.0	47.0	76	PM25	SEDANG
1821	12/28/2021	DKI5 (Kebon Jeruk) Jakarta Barat	44.0	68.0	20.0	11.0	21.0	33.0	68	PM25	SEDANG
1822	12/29/2021	DKI5 (Kebon Jeruk) Jakarta Barat	34.0	54.0	28.0	8.0	25.0	29.0	54	PM25	SEDANG
1823	12/30/2021	DKI5 (Kebon Jeruk) Jakarta Barat	53.0	75.0	25.0	15.0	23.0	44.0	75	PM25	SEDANG
1824	12/31/2021	DKI5 (Kebon Jeruk) Jakarta Barat	60.0	87.0	28.0	19.0	30.0	53.0	87	PM25	SEDANG

1825 rows x 11 columns





Data Preparation



Data yang digunakan tidak mencakup semua columns yang ada pada dataset. Beberapa variabel yang tidak digunakan di dropping dan ada juga variabel yang ditransformasi untuk memudahkan analisis. Maka dataset yang digunakan adalah dataset yang hanya meliputi variabel :

NO	Variable	Keterangan
1	tanggal	Tanggal pengukuran kualitas udara
2	kota	Lokasi dari stasiun pengukuran ISPU
3	pm ₁₀	Partikulat salah satu parameter yang diukur



Data Preparation



NO	Variable	Keterangan
4	$pm_{2,5}$	Partikulat salah satu parameter yang diukur
5	so_2	Sulfida (dalam bentuk SO2) salah satu parameter yang diukur
6	co	Carbon Monoksida salah satu parameter yang diukur
7	o_3	Ozon salah satu parameter yang diukur



Data Preparation



NO	Variable	Keterangan
8	no ₂	Nitrogen dioksida salah satu parameter yang diukur
9	max	Nilai ukur paling tinggi dari seluruh parameter yang diukur dalam waktu yang sama
10	stasiun	Lokasi titik pengukuran ISPU



GreatEdu
Home for All Education Needs

MSIB
Masyarakat Studi Independen Bersertifikat

Kampus
Merdeka
INDONESIA JAYA

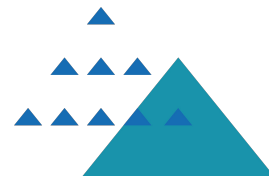


Data Preparation

	tanggal	Kota	pm10	pm2.5	so2	co	o3	no2	max	critical	categori	stasiun
0	1/1/2021	Jakarta Pusat	38.0	53.0	29.0	6.0	31.0	13.0	53	PM25	SEDANG	Bundaran HI
1	1/2/2021	Jakarta Pusat	27.0	46.0	27.0	7.0	47.0	7.0	47	O3	BAIK	Bundaran HI
2	1/3/2021	Jakarta Pusat	44.0	58.0	25.0	7.0	40.0	13.0	58	PM25	SEDANG	Bundaran HI
3	1/4/2021	Jakarta Pusat	30.0	48.0	24.0	4.0	32.0	7.0	48	PM25	BAIK	Bundaran HI
4	1/5/2021	Jakarta Pusat	38.0	53.0	24.0	6.0	31.0	9.0	53	PM25	SEDANG	Bundaran HI
...
1820	12/27/2021	Jakarta Barat	54.0	76.0	36.0	14.0	21.0	47.0	76	PM25	SEDANG	Kebon Jeruk
1821	12/28/2021	Jakarta Barat	44.0	68.0	20.0	11.0	21.0	33.0	68	PM25	SEDANG	Kebon Jeruk
1822	12/29/2021	Jakarta Barat	34.0	54.0	28.0	8.0	25.0	29.0	54	PM25	SEDANG	Kebon Jeruk
1823	12/30/2021	Jakarta Barat	53.0	75.0	25.0	15.0	23.0	44.0	75	PM25	SEDANG	Kebon Jeruk
1824	12/31/2021	Jakarta Barat	60.0	87.0	28.0	19.0	30.0	53.0	87	PM25	SEDANG	Kebon Jeruk

1825 rows x 12 columns





Modelling

Algoritma pertama yang digunakan adalah Regresi Logistic. Regresi Logistic adalah metode statistik yang digunakan dalam analisis data untuk memodelkan hubungan antara variabel dependen biner (dua kategori) dengan satu atau lebih variabel independen.

Dengan menggunakan Regresi Logistik, kami dapat menentukan variabel paling berpengaruh dengan memeriksa koefisien regresi logistik dari masing-masing variabel independen.

[Gcollab](#)

www.greatedu.co.id

SIB Cycle 5 | **Data Analyst**



Modelling

Algoritma kedua yang digunakan untuk forecasting adalah FB Prophet. FB Prophet merupakan algoritma yang digunakan untuk melakukan prediksi dengan data *time-series*.

[Gcollab](#)

www.greatedu.co.id

SIB Cycle 5 | **Data Analyst**



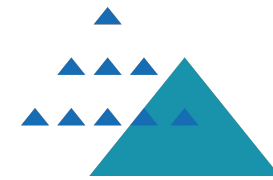
Evaluasi

```
1 # Melakukan prediksi pada data pengujian dan mengevaluasi performa model
2 y_pred = model.predict(X_test)
3 print("Accuracy:", accuracy_score(y_test, y_pred))
```

Accuracy: 0.9123287671232877

Bisa dilihat pada gambar disamping bahwa model Regresi Logistic ini memiliki Accuracy sebesar 0.91, yang menunjukkan bahwa model ini dapat memprediksi dengan benar sekitar 91% dari seluruh instance (baris) dalam data pengujian.



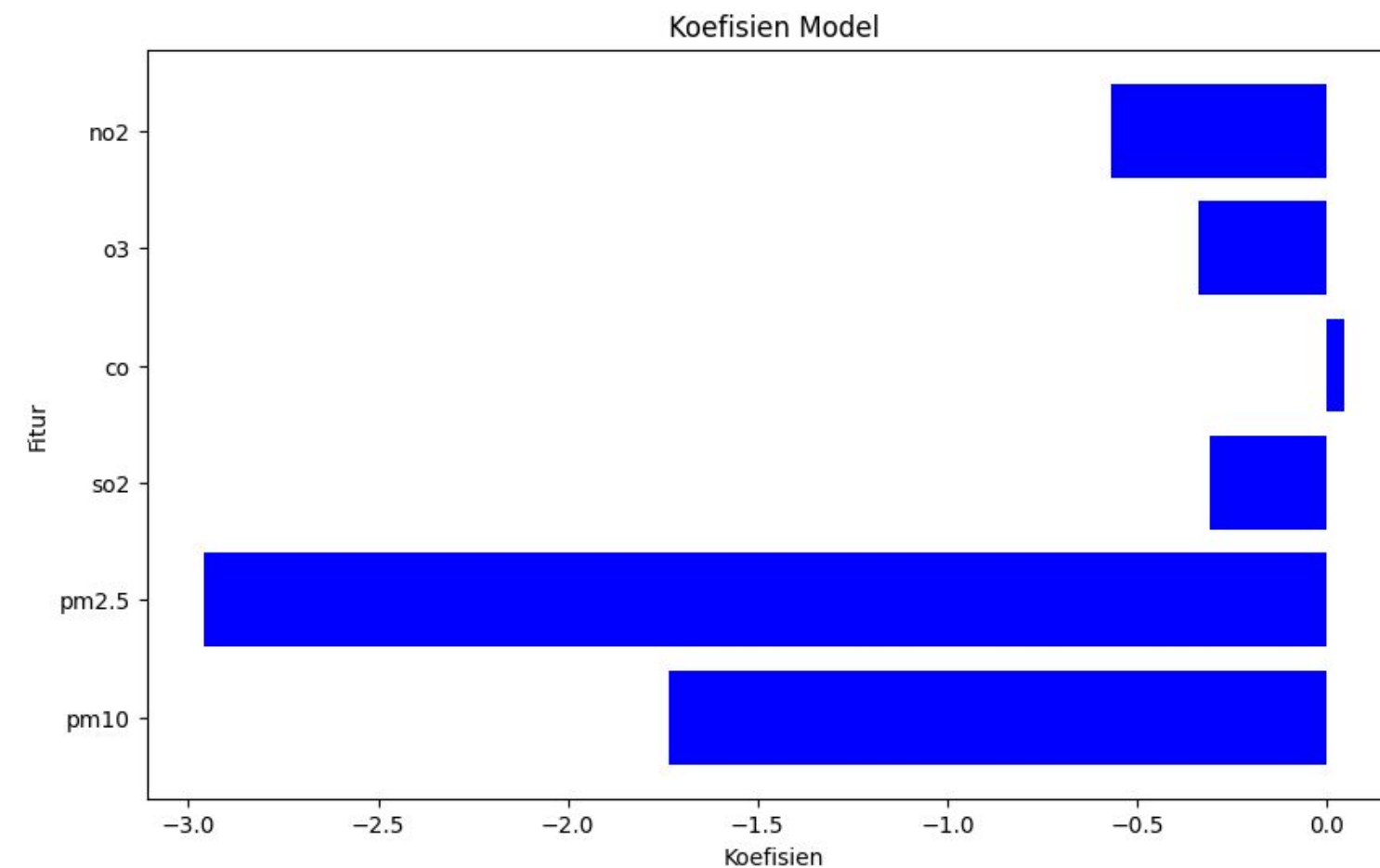


Evaluation

Berdasarkan hasil dari model regresi logistic, kami simpulkan bahwa variabel yang paling berpengaruh adalah variabel pm2.5.

```
1 # Mencetak koefisien model untuk memberikan interpretasi mengenai
2 # seberapa besar pengaruh setiap fitur terhadap hasil.
3 coef_df = pd.DataFrame({'Feature': features, 'Coefficient': model.coef_[0]})
4 print(coef_df)
5
6 # Visualisasi koefisien dengan plot batang horizontal mengarah ke kanan
7 plt.figure(figsize=(10, 6))
8 plt.barh(coef_df['Feature'], coef_df['Coefficient'], color='blue')
9 plt.xlabel('Koefisien')
10 plt.ylabel('Fitur')
11 plt.title('Koefisien Model')
12 plt.show()
```

	Feature	Coefficient
0	pm10	-1.734560
1	pm2.5	-2.959803
2	so2	-0.308645
3	co	0.046733
4	o3	-0.337020
5	no2	-0.570266

[Website](http://www.greatedu.co.id)www.greatedu.co.idSIB Cycle 5 | **Data Analyst**



GreatEdu
Home for All Education Needs

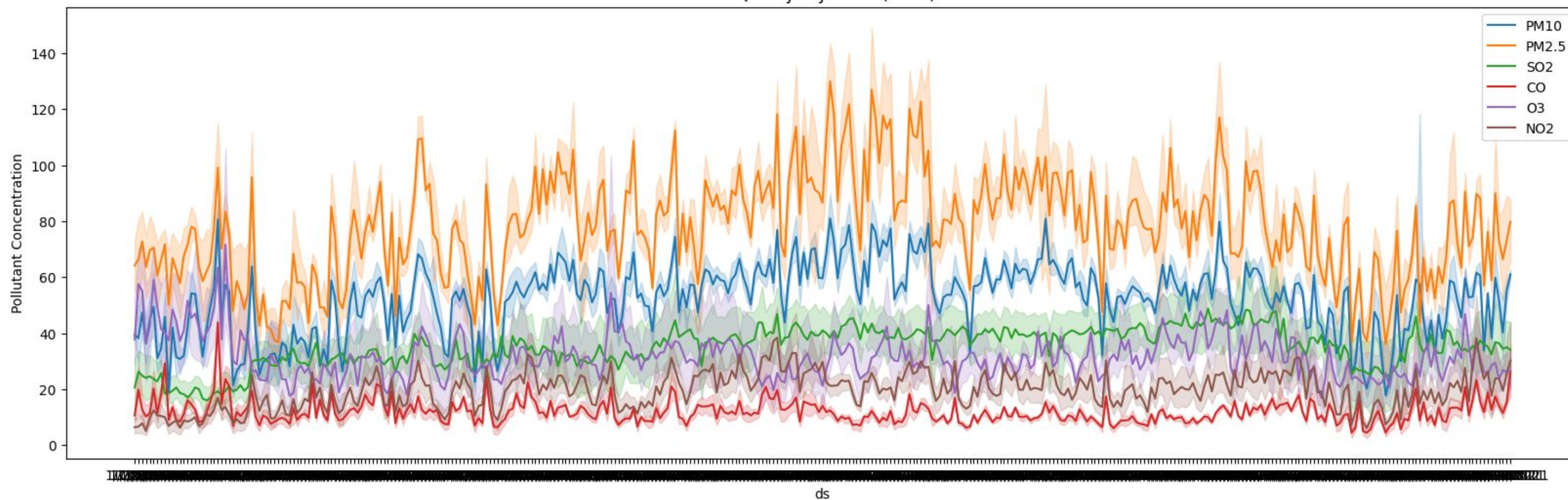
MSIB
magang dan studi independen bersertifikat

Kampus
Merdeka
INDONESIA JAYA



Evaluation

Air Quality in Jakarta (2021)





Evaluation



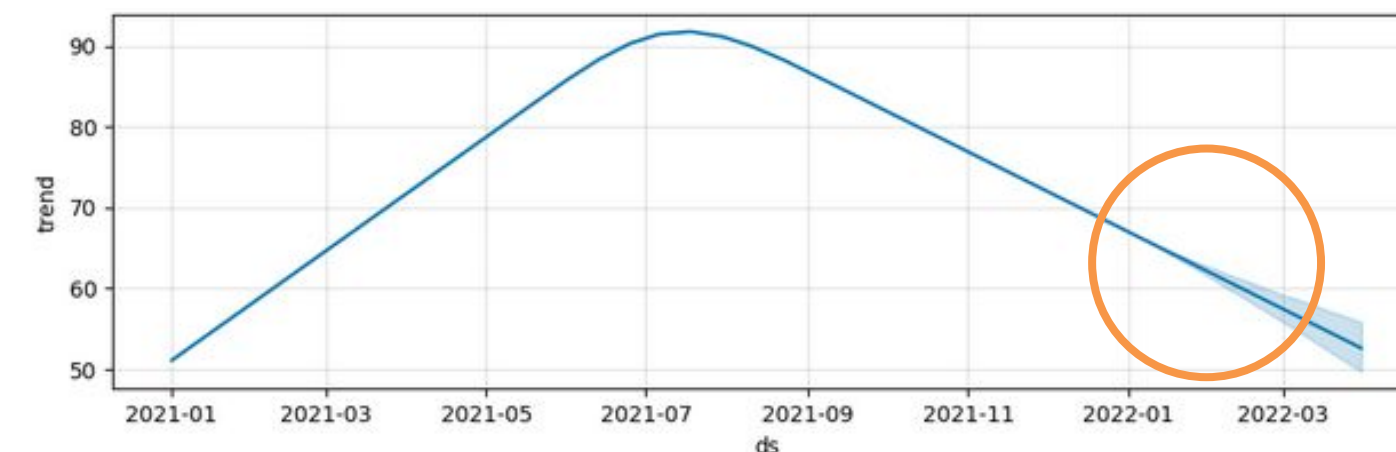
GreatEdu
Home for All Education Needs

MSIB
Magang dan Studi Independen Bersertifikat

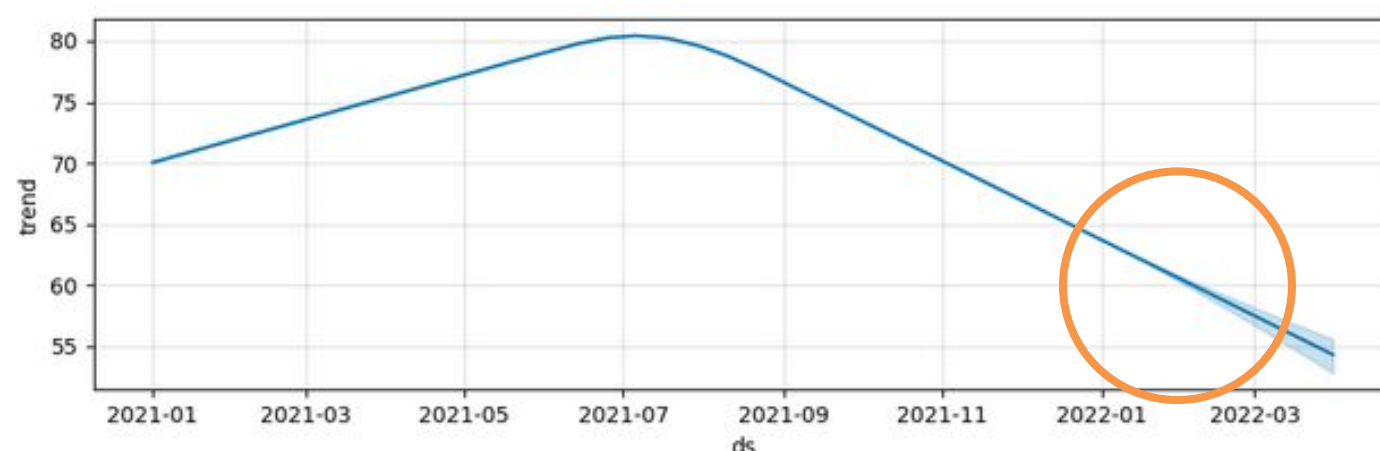
Kampus
Merdeka
INDONESIA JAYA

Jakarta Utara

- Prediksi nilai ISPU variabel pm 2.5 di **Jakarta Utara, Jakarta Selatan dan Jakarta Pusat** untuk 3 bulan awal pada tahun 2022 berturut-turut terdapat pada rentang **50-70**, **55-65**, **40-60** yang menunjukkan kategori pencemaran udara **sedang**

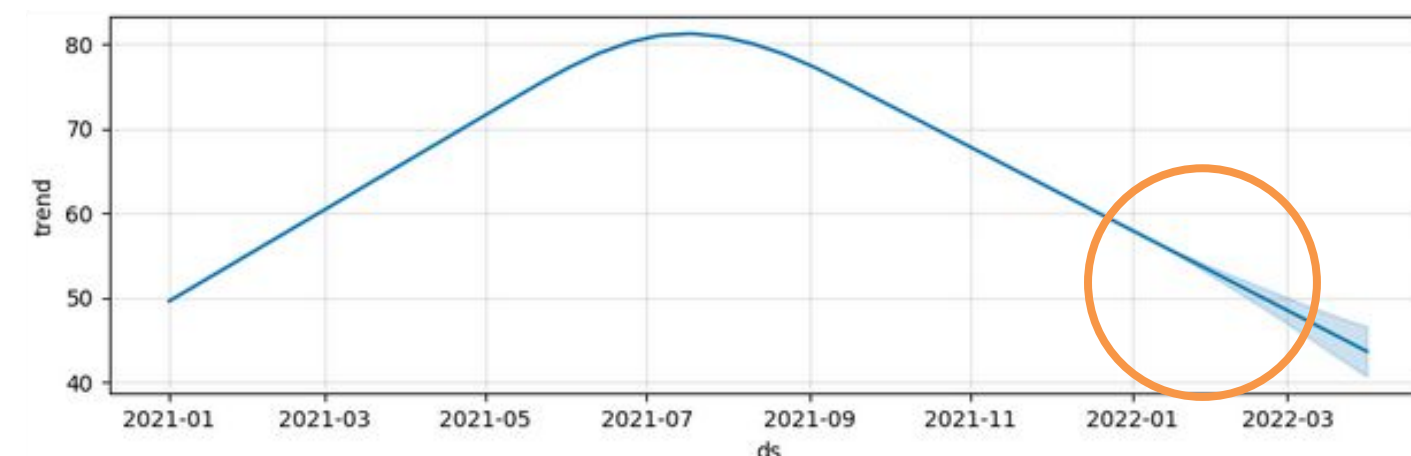


Jakarta Selatan



ISPU	Pencemaran Udara Level
0 - 50	Baik
51 - 100	Sedang
101 - 199	Tidak Sehat

Jakarta Pusat





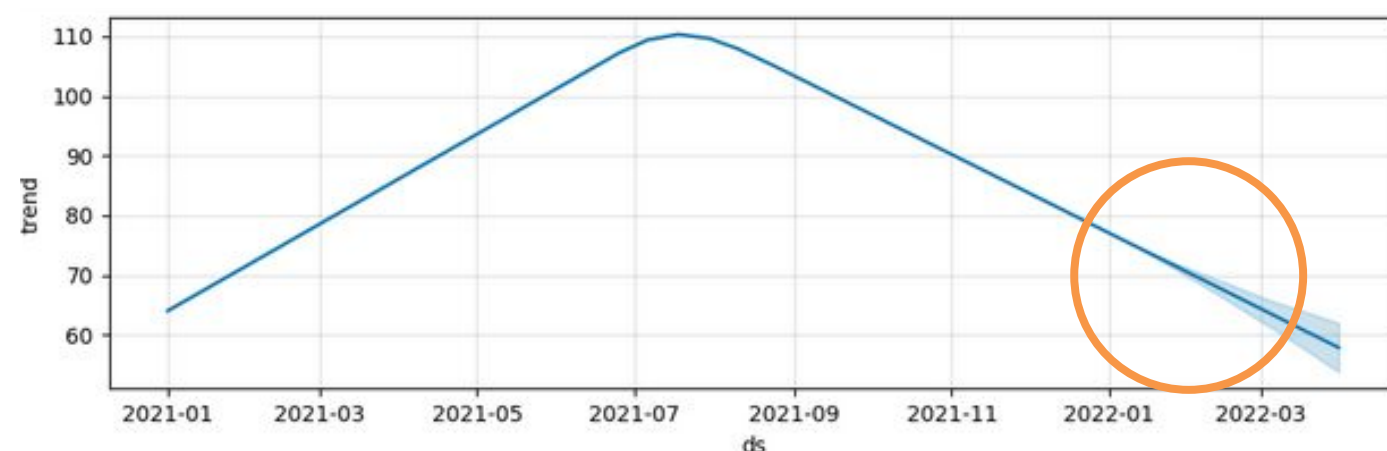
Evaluation



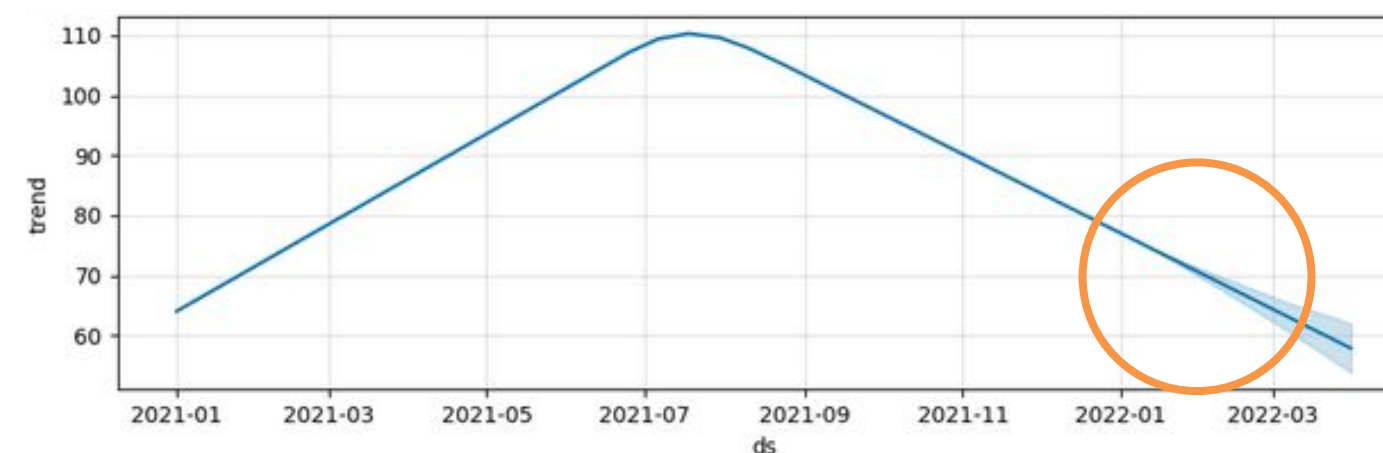
GreatEdu
Home for All Education Needs

MSIB
Magang dan Studi Independen Bersertifikat

Kampus
Merdeka
INDONESIA JAYA



Jakarta Timur



Jakarta Barat

ISPU	Pencemaran Udara Level	Dampak kesehatan
0 - 50	Baik	tidak memberikan dampak bagi kesehatan manusia atau hewan.
51 - 100	Sedang	tidak berpengaruh pada kesehatan manusia ataupun hewan tetapi berpengaruh pada tumbuhan yang peka.
101 - 199	Tidak Sehat	bersifat merugikan pada manusia ataupun kelompok hewan yang peka atau dapat menimbulkan kerusakan pada tumbuhan ataupun nilai estetika.
200 - 299	Sangat Tidak Sehat	kualitas udara yang dapat merugikan kesehatan pada sejumlah segmen populasi yang terpapar.
300 - 500	Berbahaya	kualitas udara berbahaya yang secara umum dapat merugikan kesehatan yang serius pada populasi (misalnya iritasi mata, batuk, dahak dan sakit tenggorokan).

- Prediksi nilai ISPU variabel pm 2.5 di **Jakarta Timur dan Jakarta Barat** untuk 3 bulan awal pada tahun 2022 terdapat pada rentang **60-80** yang menunjukkan kategori pencemaran udara **sedang**

Prediksi nilai ISPU tertinggi berada di **jakarta timur dan jakarta barat** yaitu dengan rentang **60-80** (pencemaran udara **sedang**)

Deployment



Dengan demikian, dapat kami simpulkan bahwa Jakarta Timur dan Barat adalah lokasi yang tepat untuk Plant Heaven Nursery dapat menjual produknya.

Dengan memfokuskan penjualannya di daerah Lubang Buaya dan Kebon Jeruk.

Website

www.greatedu.co.id

SIB Cycle 5 | **Data Analyst**



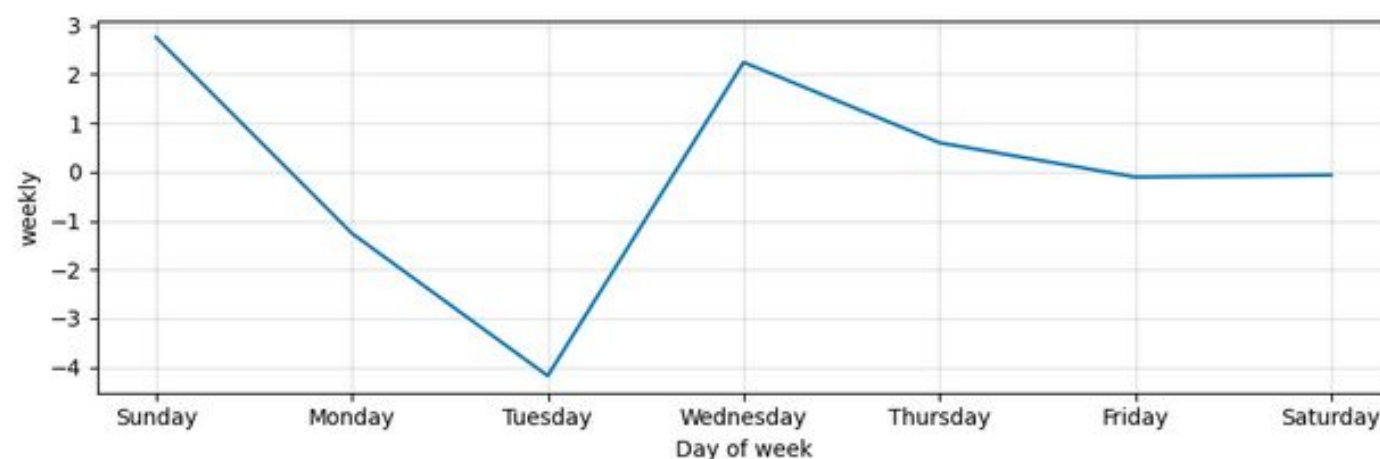
Deployment



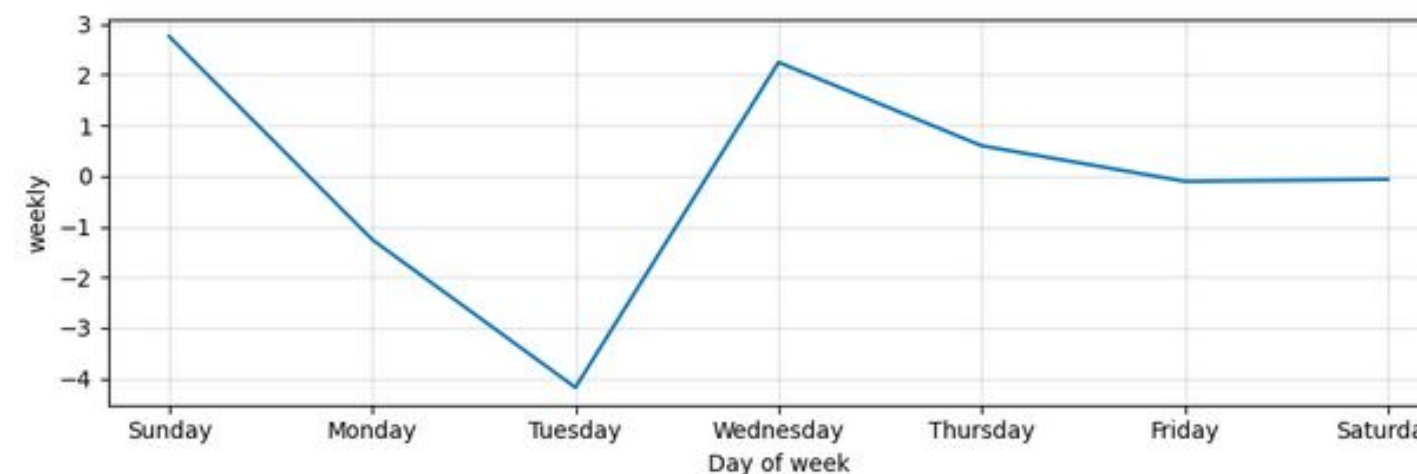
GreatEdu
Home for All Education Needs

MSIB
Magang dan Studi Independen Bersertifikat

Kampus
Merdeka
INDONESIA JAYA



Jakarta Timur



Jakarta Barat

- Pada gambar pertama dan kedua (berdasarkan hari), data cenderung tinggi di hari minggu (sunday) dan rabu (Wednesday). Lalu sangat rendah di hari Selasa (Tuesday)
- Saran dari tim kami untuk perusahaan Plant Heaven Nursery, penjualan di lokasi Jakarta Timur dan Barat sebaiknya lebih diperhatikan di hari minggu (sunday) dan Rabu (Wednesday).



Terima Kasih.



GreatEdu
Home for All Education Needs

MSIB
magang dan studi independen bersertifikat

Kampus
Merdeka
INDONESIA JAYA

Visualisasi Looker Studio



INDEKS STANDAR PENCEMAR UDARA DKI JAKARTA TAHUN 2021

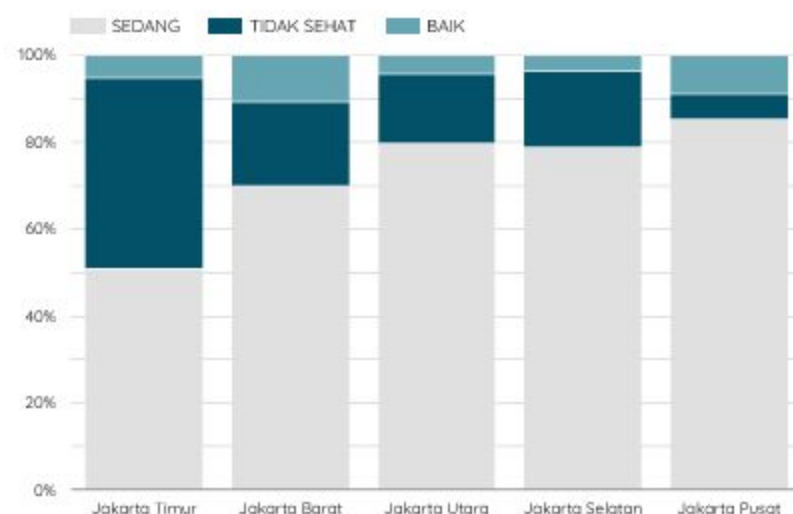
Pilih rentang tanggal

Kota

kategori

	tanggal	Kota	stasiun	kategori	critical
1.	1 Jan 2021	Jakarta P...	Bundaran HI	SEDANG	PM25
2.	1 Jan 2021	Jakarta S...	Jagakarsa	SEDANG	PM25
3.	1 Jan 2021	Jakarta TL	Lubang Bua...	BAIK	PM10
4.	1 Jan 2021	Jakarta Ut...	Kelapa Gad...	SEDANG	O3
5.	1 Jan 2021	Jakarta B...	Kebon Jeruk	BAIK	PM10
6.	2 Jan 2021	Jakarta P...	Bundaran HI	BAIK	O3
7.	2 Jan 2021	Jakarta S...	Jagakarsa	SEDANG	PM25
8.	2 Jan 2021	Jakarta TL	Lubang Bua...	SEDANG	CO
9.	2 Jan 2021	Jakarta Ut...	Kelapa Gad...	SEDANG	O3
10.	2 Jan 2021	Jakarta B...	Kebon Jeruk	SEDANG	CO
11.	3 Jan 2021	Jakarta S...	Jagakarsa	SEDANG	PM25
12.	3 Jan 2021	Jakarta TL	Lubang Bua...	SEDANG	CO
13.	3 Jan 2021	Jakarta P...	Bundaran HI	SEDANG	PM25

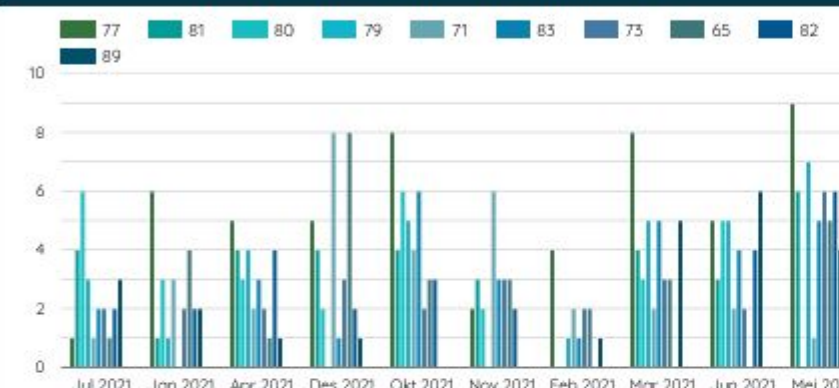
1 - 100 / 1825



INDEKS STANDAR PENCEMAR UDARA DKI JAKARTA TAHUN 2021 BERDASARKAN TIAP WILAYAH

Kota

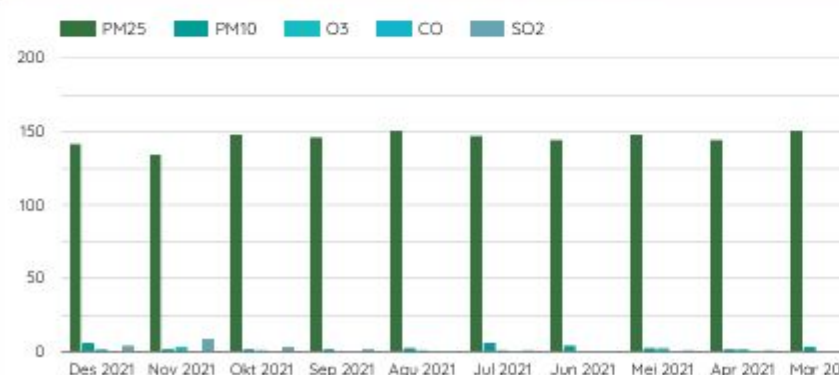
PENCEMARAN UDARA BERDASARKAN RENTANG NILAI



PENCEMARAN UDARA BERDASARKAN RENTANG NILAI



PENCEMARAN UDARA BERDASARKAN CRITICAL



PENCEMARAN UDARA BERDASARKAN CATEGORI





Flowchart

