Air Quality Index Analysis and Correlation Effect

BY NUR IMAM MASRI

List of contents

- Introduction
- Destination
- Business Approach
- Dataset (Data Preparation, EDA, Data Preprocessing)
- Insight Analysis
- Modeling
- Conclusion

INTRODUCTION

Kualitas udara bisa menjadi penentu kesehatan suatu lingkungan dan mahluk hidup di dalamnya. Udara yang sehat pastinya berdampak baik bagi kehidupan di dalamnya, begitupula sebaliknya.

Pencemaran udara adalah masuknya atau dimasukannya zat, energi, dan atau komponen lain kedalam udara oleh kegiatan manusia, sehingga mutu udara turun sampai ke tingkat tertentu yang menyebabkan atau mempengaruhi kesehatan manusia.

Destination

Setelah melihat permasalahan sebelumnya, saya berencana untuk membuat sistem yang bisa mendesteksi kualitas udara dengan tujuan:

- *Supaya kita bisa mengetahui apakah kualitas udara yang kita hirup ini tergolong baik, tidak sehat, atau mungkin berbahaya bagi kesehatan.*

- *Jika kita mengetahui kualitas udara disekitar kita, kita bisa lebih mempersiapkan diri sebelum keluar rumah. Misalkan, jika kita tahu bahwa kualitas udara sedang tidak sehat. Maka kita akan mempersiapkan diri terlebih dahulu sebelum keluar rumah, seperti menyiapkan beberapa masker yang dipakai agar bisa menyaring udara yang kita hirup supaya tidak terkena penyakit gangguan saluran pernafasan, dll.*

- Berdasarkan peraturan yang telah ditetapkan pemerintah (Peraturan Menteri LHK RI P.14/MENLHK/SETJEN/KUM.1/7/2020) mengenai kategori kualitas udara, dengan machine learning ini kita dapat mengetahui ***apakah*** kualitas udara tersebut (Good, Moderate, Unhealthy, Very Unhealthy, Dangerous).
    -
- Pada kasus yang lebih luas, kualitas udara sangat berpengaruh terhadap perekonomian negara. Sebab siapa yang ingin berinvestasi pada suatu produk dinegara yang kualitas udara buruk, berpolusi, dll. Akan menjadi pertanyaan apakah hasil produksi tersebut baik untuk dikonsumsi, apakah produk tersebut tidak terkontaminasi oleh udara yang tidak baik & apakah produk tersebut layak untuk diperjualkan.
    -
- Setelah kita tahu kualitas udaranya, kita juga akan tahu componen polusi udara apa yang membuat kualitas udara memburuk. Setelah kita tahu komponennya, barulah kita bisa meminimalisir komponen polusi tersebut

There are many different types of air pollutants, such as :

- Gases (including ammonia, carbon monoxide, sulfur dioxide, nitrous oxides, methane, carbon dioxide and chlorofluorocarbons),
    - PM10 - Particulate Matter
    - SO2 - Sulfur Dioxide
    - CO - Carbon Monoxide
    - O3 - Ozone
    - NO2 - Natrium Dioxide
    - CFC - Chlorofluorocarbon
    - HC - Hidrokarbon
    - Pb - Timah
    - CO2 - Carbon Diaoksida
- Particulates (both organic and inorganic), and
- Biological Molecules

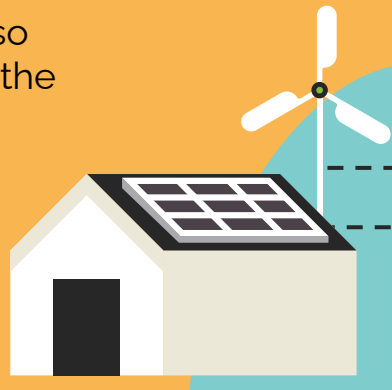
Air pollution can cause diseases, allergies, and even death to humans; It can also cause harm to other living organisms such as animals and food crops, and may damage the natural environment

CO2 Emissions and Economic

Fossil fuel has been the world's primary source of electricity.

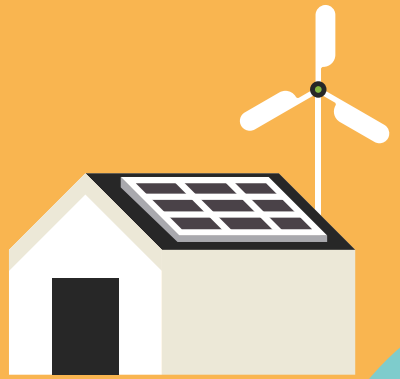
The majority or 50% of the installed capacity of power plants in Indonesia still comes from fossil energy. The need for electrical energy from year to year continues to increase in line with the increasing rate of economic growth, population, and the development of the industrial sector.

The use of fossil energy will produce waste in the form of CO2 gas which causes infrared radiation from the earth to return to the earth's surface so that it can cause global warming. In addition, the use of fossil energy as the main source of power generation can also result in depletion of natural resource reserves, such as oil, coal, and gas.



Carbon Tax

Carbon tax is a tax levied on the burning of carbon-based fuels such as coal, oil and gas. The carbon tax is a core policy created to reduce and eliminate the use of fossil fuels whose burning can damage the climate.



Datasets

1. ***AIR QUALITY INDEX (by cities)***
2. ***AIR QUALITY INDEX (top countries)***
3. ***Pollutant Standards Index Jogja 2020***
4. ***GDP Per Capita***
5. ***CO2 Emissions Per Capita***
6. ***Electricity Generated Year***
7. ***Annual CO Emissions by Region***
8. ***Perkembangan Jumlah Kendaraan Bermotor Menurut Jenis (Unit), 2018-2020***
9. ***Jumlah Kendaraan Bermotor Menurut Provinsi dan Jenis Kendaraan (unit)***
10. ***Jumlah Penduduk Hasil Proyeksi Menurut Provinsi dan Jenis Kelamin (Ribu Jiwa), 2018-2020***

Datasets

1. ***AIR QUALITY INDEX (by cities)***
2. ***AIR QUALITY INDEX (top countries)***

Most Polluted Cities and Countries (IQAir Index)

Data | Code (1) | Discussion (0) | Metadata

52 | New Notebook | Download (216 kB)

AIR QUALITY INDEX (by cities) - IQAir.csv (603.43 kB)

Detail | Compact | Column | 10 of 19 columns

About this file

Dataset covers almost 6k cities where which were most polluted as per IQ Air index.

Data Explorer

Version 2 (608.44 kB)

- AIR QUALITY INDEX (by citi...
- AIR QUALITY INDEX- top co...

# Rank	A City	# 2021	A JAN(2021)	A FEB(2021)	A MAR(2
Rank	city	2021	january	feb	march
1 – 6475	6216 unique values	1.5 – 106	- 10% 8.5 1% Other (5746) 89%	- 9% 8.8 1% Other (5846) 90%	- 5.3 Other (62
1	Bhiwadi, India	106.2	145.8	129.8	120.2
2	Ghaziabad, India	102	199.9	172.2	97.8
3	Hotan, China	101.5	-	-	158
4	Delhi, India	96.4	183.7	142.2	80.5
5	Jaunpur, India	95.3	182.2	143.5	91

Datasets

1. ***AIR QUALITY INDEX (by cities)***
2. ***AIR QUALITY INDEX (top countries)***

	Rank	City	2021	JAN(2021)	FEB(2021)	MAR(2021)	APR(2021)	MAY(2021)	JUN(2021)	JUL(2021)	AUG(2021)	SEP(2021)	OCT(2021)	NOV(2021)	DEC(2021)	2020	2019	2018	2017
0	1	Bhiwadi, India	106.2	145.8	129.8	120.2	125.7	86.5	95.9	55.6	55.4	37.1	91.1	188.6	136.6	95.5	83.4	125.4	-
1	2	Ghaziabad, India	102.0	199.9	172.2	97.8	86.3	52.9	47.2	35.3	37.6	30.8	89.7	218.3	163	106.6	110.2	135.2	144.6
2	3	Hotan, China	101.5	-	-	158	91.1	167.4	57.4	70.9	93.2	79.3	126.1	111.5	62.6	110.2	110.1	116	91.9
3	4	Delhi, India	96.4	183.7	142.2	80.5	72.9	47.4	47.1	35.6	36.9	30.2	73.7	224.1	186.4	84.1	98.6	113.5	108.2
4	5	Jaunpur, India	95.3	182.2	143.5	91	70	51.1	40.7	33.5	34.2	36.8	75.7	196	195.7	-	-	-	-

	Rank	Country/Region	2021	2020	2019	2018	Population
0	1	Bangladesh	76.9	77.1	83.3	97.1	164,689,383
1	2	Chad	75.9	-	-	-	16,425,859
2	3	Pakistan	66.8	59	65.8	74.3	220,892,331
3	4	Tajikistan	59.4	30.9	-	-	9,537,642
4	5	India	58.1	51.9	58.1	72.5	1,380,004,385

Datasets

3. ***Pollutant Standards Index Jogja 2020***

Dataset yang saya ambil bersumber dari kaggle berjudul **"Air Quality in Yogyakarta, Indonesia (2020)".** Dataset ini berisi pengukuran polusi udara yang terjadi di yogyakarta pada tahun 2020.

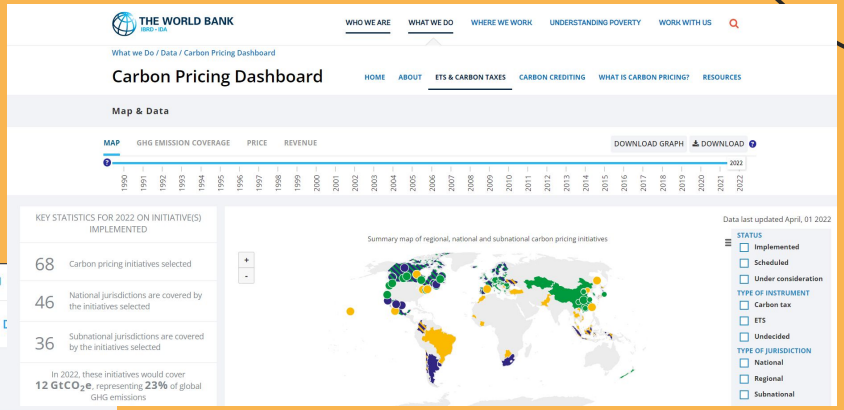
Attribute Information:

- **Date** - Tanggal pengukuran
- **PM10** - Pengukuran Partikulat Matter
- **SO2** - Pengukuran Sulfur Dioksida
- **CO** - Pengukuran Karbon Monoksida
- **O3** - Pengukuran ozon
- **NO2** - Pengukuran Natrium DIoksida
- **Max** - Nilai pengukuran tertinggi
- **Critical Component** - Komponen yang memiliki nilai pengukuran tertinggi
- **Category** - Kategori pencemaran udara, baik atau tidak

	Date	PM10	SO2	CO	O3	NO2	Max	Critical Component	Category
0	1/1/2020	30	2	69	19	0	69	CO	Moderate
1	1/2/2020	16	2	58	33	0	58	CO	Moderate
2	1/3/2020	12	2	46	18	0	46	CO	Good
3	1/4/2020	8	2	84	29	0	84	CO	Moderate
4	1/5/2020	8	3	50	0	0	50	CO	Good
5	1/6/2020	12	2	97	18	0	97	CO	Moderate

Datasets

4. ***GDP Per Capita***
5. ***CO2 Emissions Per Capita***
6. ***Electricity Generated Year***
7. ***Annual CO Emissions by Region***



Datasets

4. ***GDP Per Capita***
5. ***CO2 Emissions Per Capita***
6. ***Electricity Generated Year***
7. ***Annual CO Emissions by Region***

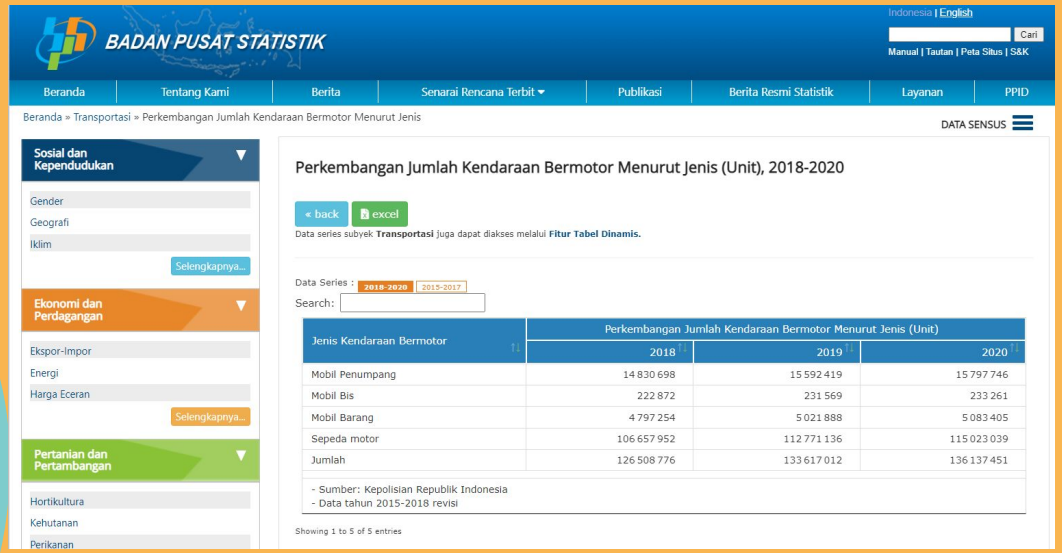
	Country Name	Year	GDP
0	Aruba	1960	0.0
1	Aruba	1961	0.0
2	Aruba	1962	0.0
3	Aruba	1963	0.0
4	Aruba	1964	0.0

	Entity	Code	Year	Annual CO2 emissions (zero filled)
0	Afghanistan	AFG	1750	0
1	Afghanistan	AFG	1751	0
2	Afghanistan	AFG	1752	0
3	Afghanistan	AFG	1753	0
4	Afghanistan	AFG	1754	0

	Entity	Year	Generated_Electricity	Fossil_Energy	Nuclear_Energy	Renewabl
0	Afghanistan	2000	0.467	0.155	0.0	
1	Afghanistan	2001	0.592	0.094	0.0	
2	Afghanistan	2002	0.687	0.132	0.0	
3	Afghanistan	2003	0.939	0.309	0.0	
4	Afghanistan	2004	0.894	0.329	0.0	

Datasets

8. Perkembangan Jumlah Kendaraan Bermotor Menurut Jenis (Unit), 2018-2020
9. Jumlah Kendaraan Bermotor Menurut Provinsi dan Jenis Kendaraan (unit)
10. Jumlah Penduduk Hasil Proyeksi Menurut Provinsi dan Jenis Kelamin (Ribu Jiwa), 2018-2020



Datasets

	Year	Mobil Penumpang	Mobil Bis	Mobil Barang	Sepeda motor	Jumlah
0	2015	12304221	196309	4145857	88656931	105303318
1	2016	13142958	204512	4326731	94531510	112205711
2	2017	13968202	213359	4540902	100200245	118922708
3	2018	14830698	222872	4797254	106657952	126508776
4	2019	15592419	231569	5021888	112771136	133617012
5	2020	15797746	233261	5083405	115023039	136137451

	Provinsi	2018	2019	2020
0	ACEH	5243.4	5316.3	5388.1
1	SUMATERA UTARA	14476.0	14639.4	14798.4
2	SUMATERA BARAT	5411.8	5479.5	5545.7
3	RIAU	6717.6	6835.1	6951.2
4	JAMBI	3527.1	3566.2	3604.2

	Year	Province	Mobil Penumpang	Bus	Truk	Sepeda Motor	Jumlah
0	2021	Aceh	166570	1103	67817	2089319	2324809
1	2021	Sumatera Utara	690543	5893	271352	6062939	7030727
2	2021	Sumatera Barat	278705	4239	135086	2118305	2536335
3	2021	Riau	367515	6060	207390	3485246	4066211
4	2021	Jambi	296892	35646	182830	2520112	3035480

Data Preparation

- Missing Values
- Column Type
- Duplicate Value
- Value Format
- Change Column Name
- Change Transformation Table

Data Preparation

https://github.com/nurimammasri/Air-Quality-Index-Analysis-and-Correlation-Effect

01. Most Polluted Cities and Countries (IQAir Index).ipynb

https://bit.ly/AQI_01

02. Air Quality in Yogyakarta, Indonesia.ipynb

https://bit.ly/AQI_02

03. Co2 Emissions and Economic.ipynb

https://bit.ly/AQI_03

04. Additional Data.ipynb

https://bit.ly/AQI_04

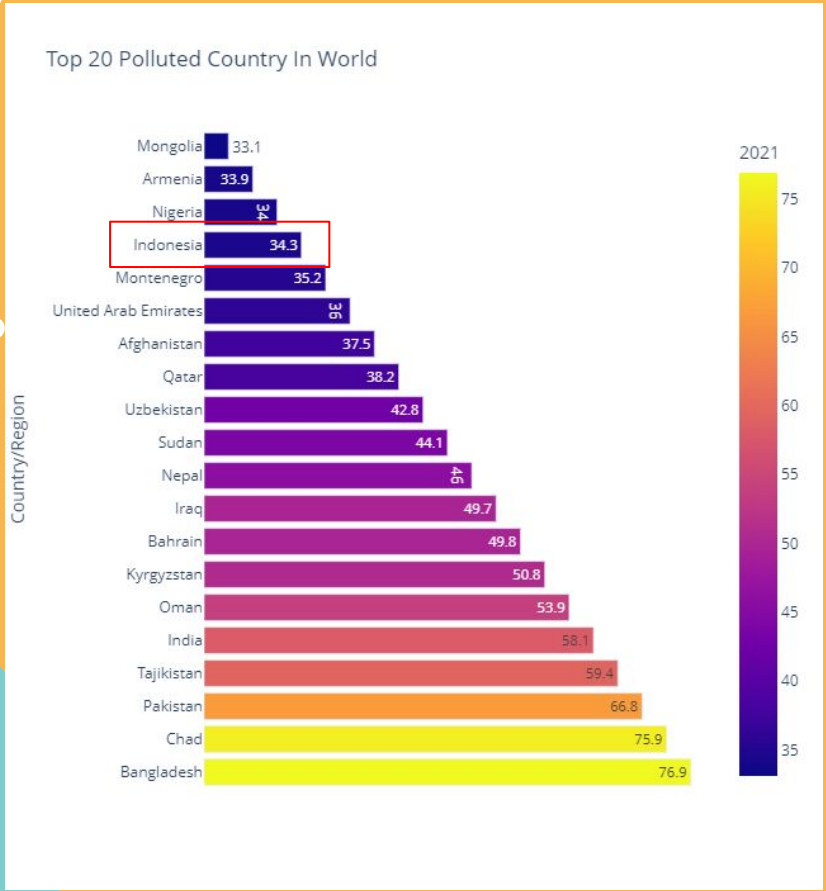
Insight Analysis

KUALITAS UDARA

Seperti yang sudah disinggung sebelumnya mengenai kualitas udara, bagaimana cara menilai bahwa udara tersebut baik atau tidaknya. Berikut adalah Kategori Indeks Standar Pencemar Udara (ISPU) yang dikeluarkan oleh Kementrian Lingkungan Hidup dan Kehutanan

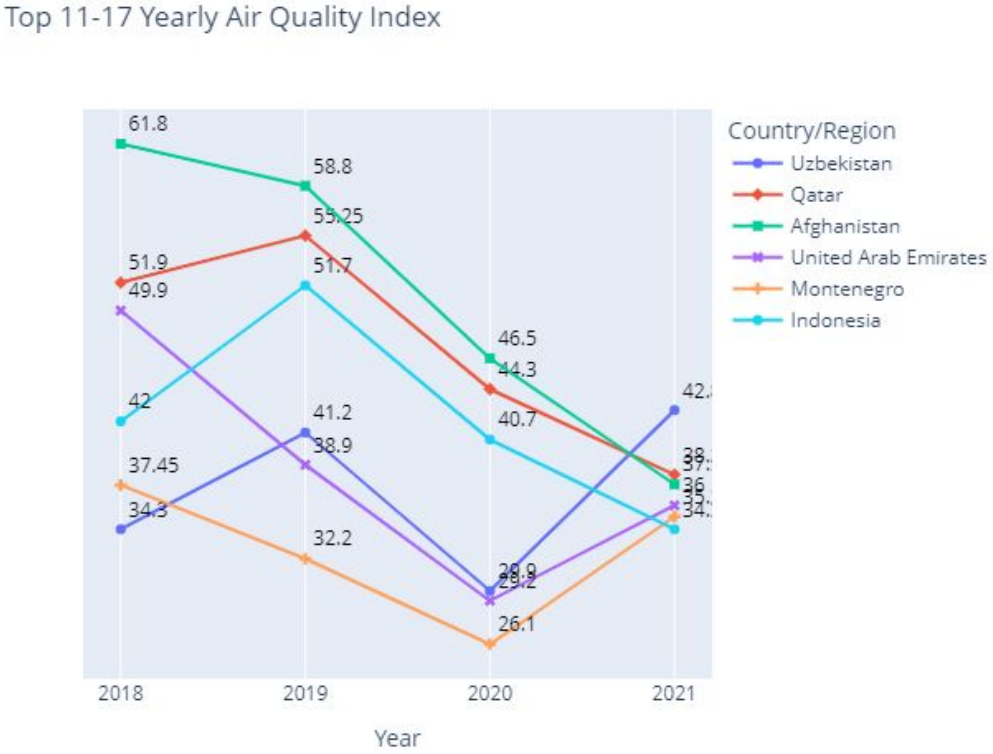
Rentang	Kategori	Penjelasan
1-50	Baik	Tingat mutu udara yang sangat baik, tidak memberikan efek negatif terhadap manusia, hewan dan tumbuhan
51-100	Sedang	Tingkat mutu udara masih dapat diterima pada kesehatan manusia, hewan dan tumbuhan.
101-200	Tidak Sehat	Tingkat mutu udara yang bersifat merugikan pada manusia, hewan dan tumbuhan.
201-300	Sangat Tidak Sehat	Tingkat mutu udara yang dapat meningkatkan resiko kesehatan pada sejumlah segmen populasi yang terpapar.
301+	Berbahaya	Tingkat mutu udara yang dapat merugikan kesehatan serius pada populasi dan perlu penanganan cepat.

Most Polluted Cities and Countries (IQAir Index)



It can be seen that Indonesia has a 17th Rating with an AQI (Air Quality Index) value of 34.3 so it is considered Good.

Most Polluted Cities and Countries (IQAir Index)

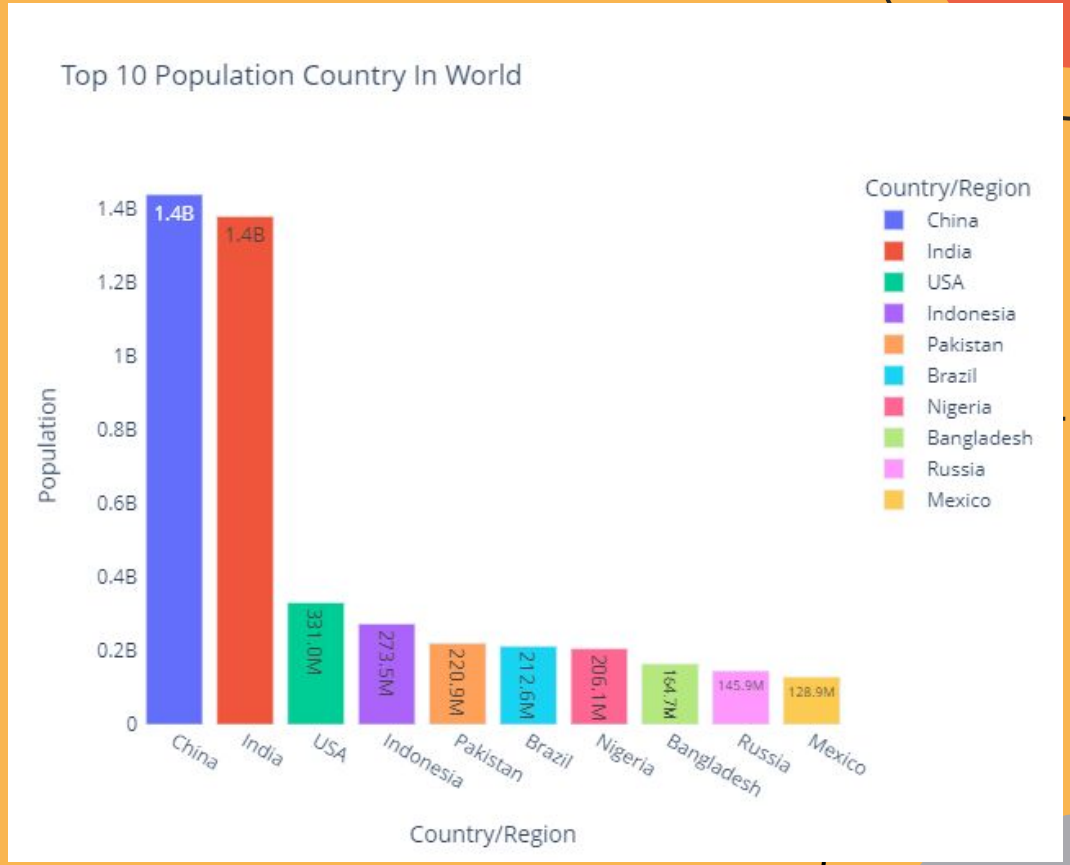


The value of AQI (Air Quality Index) in Indonesia in **2018 was 42**
Then it rose 9.7 in 2019 by 51.7
And down in 2020 and 2021 maybe because of the number of vehicles and the process of activity during Covid-19.

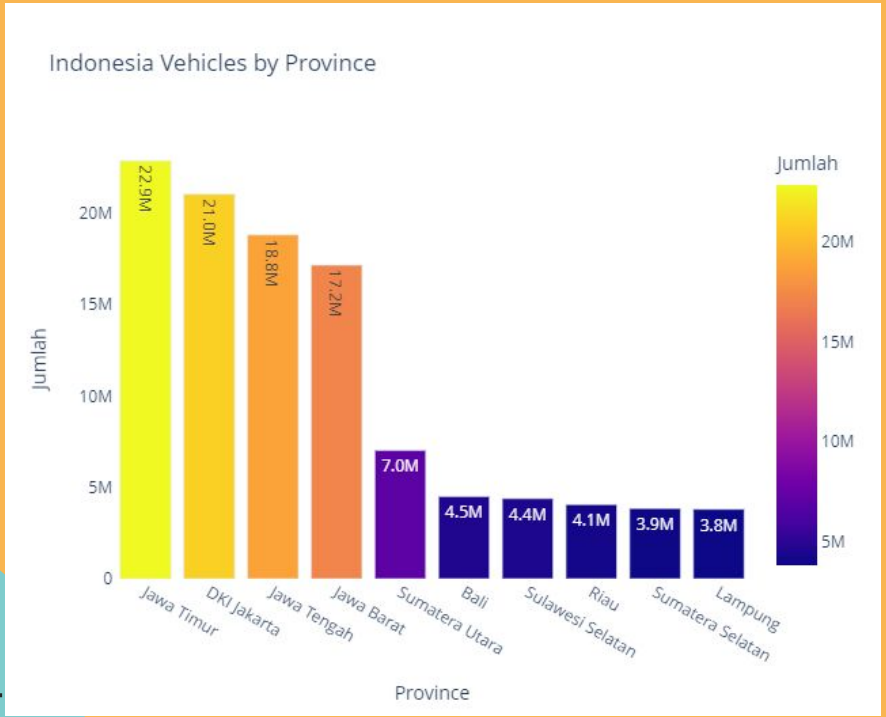
Most Polluted Cities and Countries (IQAir Index)

The total population in Indonesia ranks 4th:
China = 1.4B
India = 1.4B
USA = 331M
Indonesia 273.5M

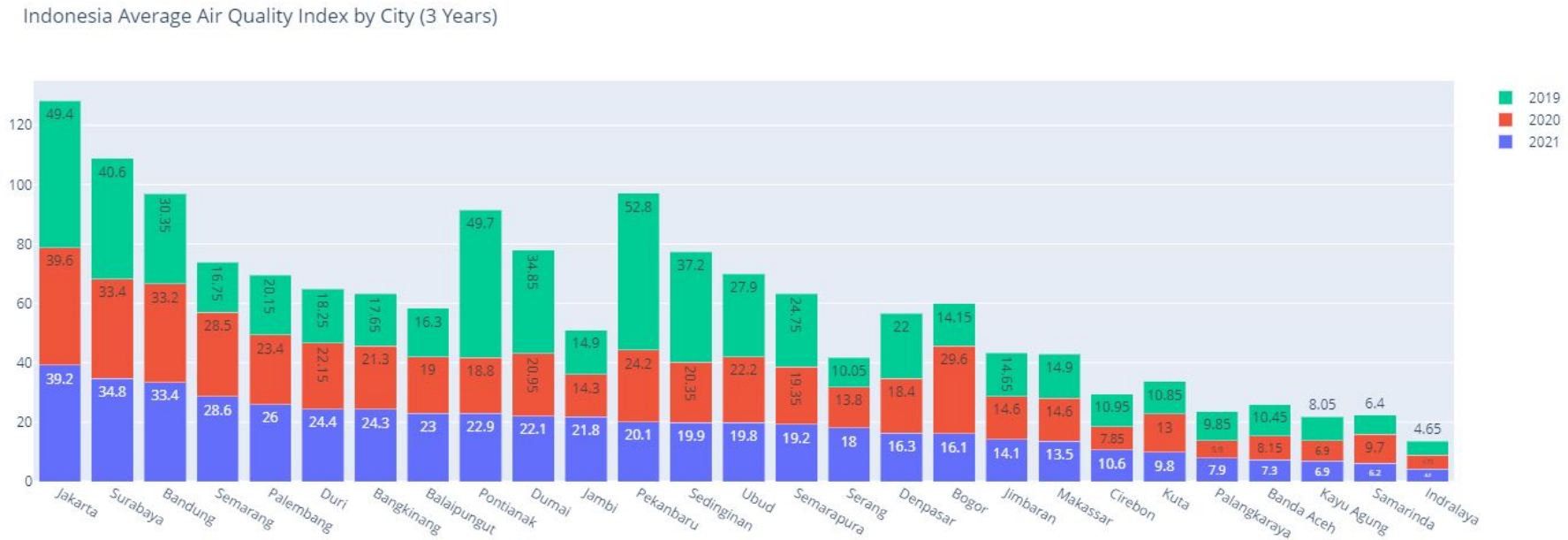
The population of Indonesia has increased from time to time. The increase in population has a negative impact on the environment. the availability of green land as a source of clean air in urban areas is also reduced due to the many existing green lands being converted as settlements. Thus it can be said that an increase in population can lead to reduced availability of clean air. The reduced availability of clean air can also be caused by air pollution due to motor vehicle fumes.



Most Polluted Cities and Countries (IQAir Index)



Most Polluted Cities and Countries (IQAir Index)

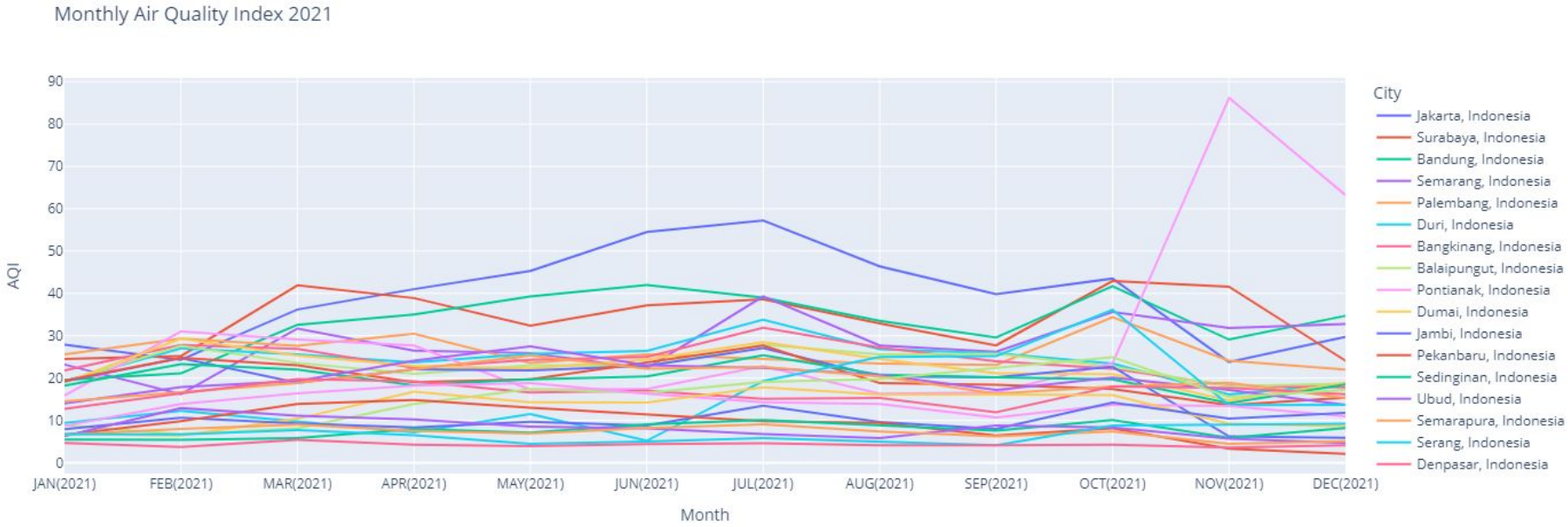


Most Polluted Cities and Countries (IQAir Index)

Keypoints

- Overall, air pollution has gone down in the last 4 years being year 2021 to be the lowest
- Jakarta has highest average AQI score with 39.2
- Indralaya in South Sumatra has lowest score with 4.2
- 4 out of 5 highest polluted cities is in Java
- Lowest 5 polluted cities located in Sumatra and Kalimantan
  Many cities have around 15-25 AQI score

Most Polluted Cities and Countries (IQAir Index)

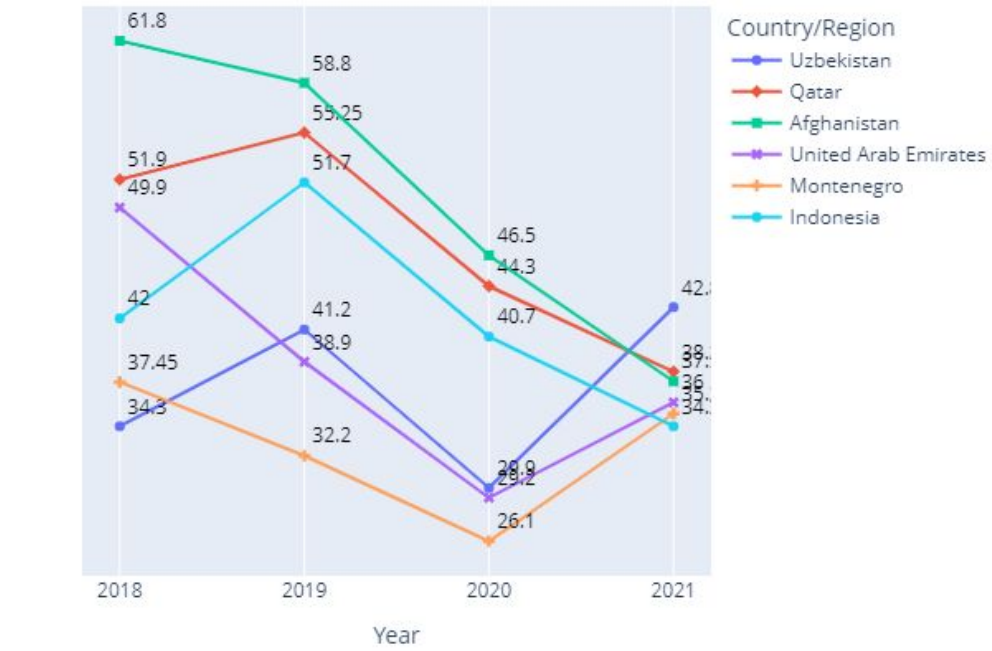


Most Polluted Cities and Countries (IQAir Index)

Keypoints

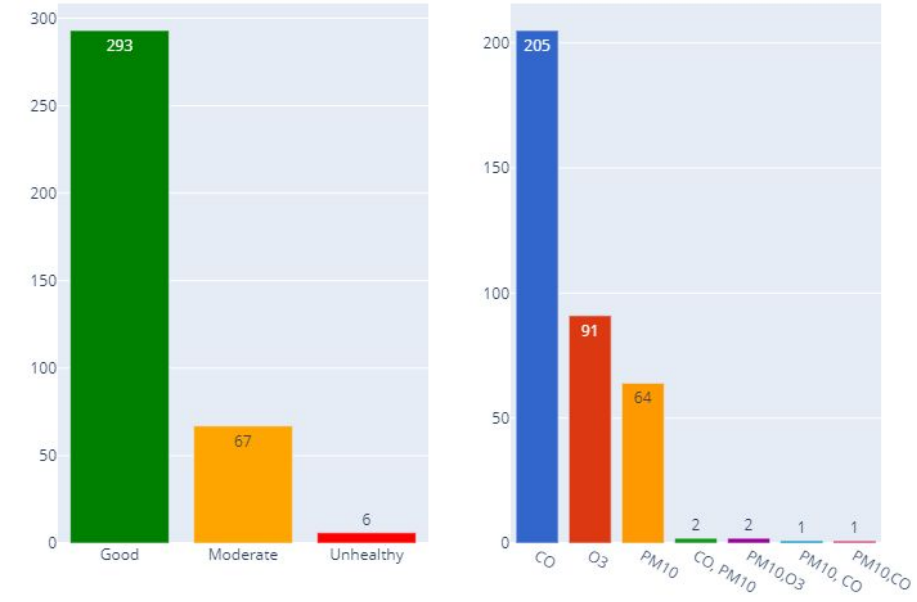
- **Jakarta have overall high AQI with highest on July with 57.2**
- **Pontianak has staggering rise of 86.2 AQI on November yet followed by inverse effect in cities like Bandung, Jakarta, Serang, and Jambi**
- **Indralaya experienced healthy fall of AQI after March by almost 40 points**

Most Polluted Cities and Countries (IQAir Index)



The value of AQI (Air Quality Index) in Indonesia in **2018 was 42**
Then it rose 9.7 in 2019 by 51.7
And down in 2020 and 2021 maybe because of the number of vehicles and the process of activity during Covid-19.

Air Quality Category and Critical Component



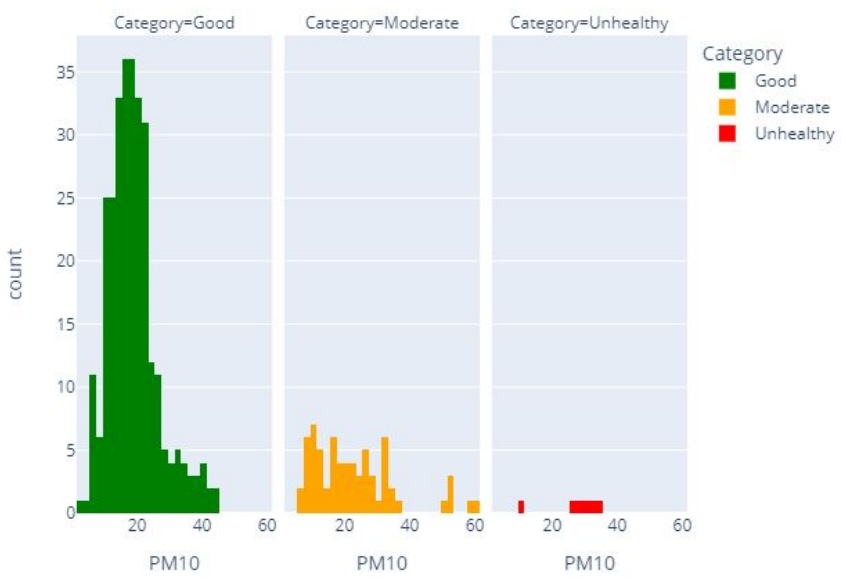
Air Quality Category and Critical Component

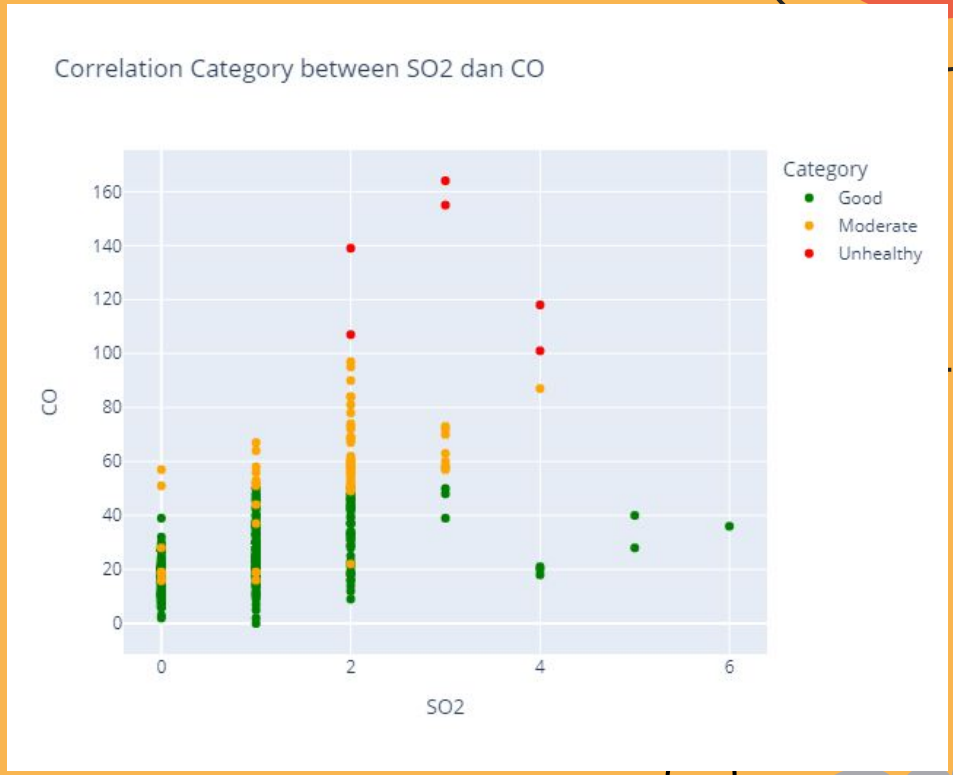
Observation:

- There are 3 air qualities in the data
- Good air quality is the highest air quality detected in 2020 in Yogyakarta at 80%
- Seen less/unhealthy air quality, too small in 2020
- CO, O3 and PM10 dominate the category in terms of critical value

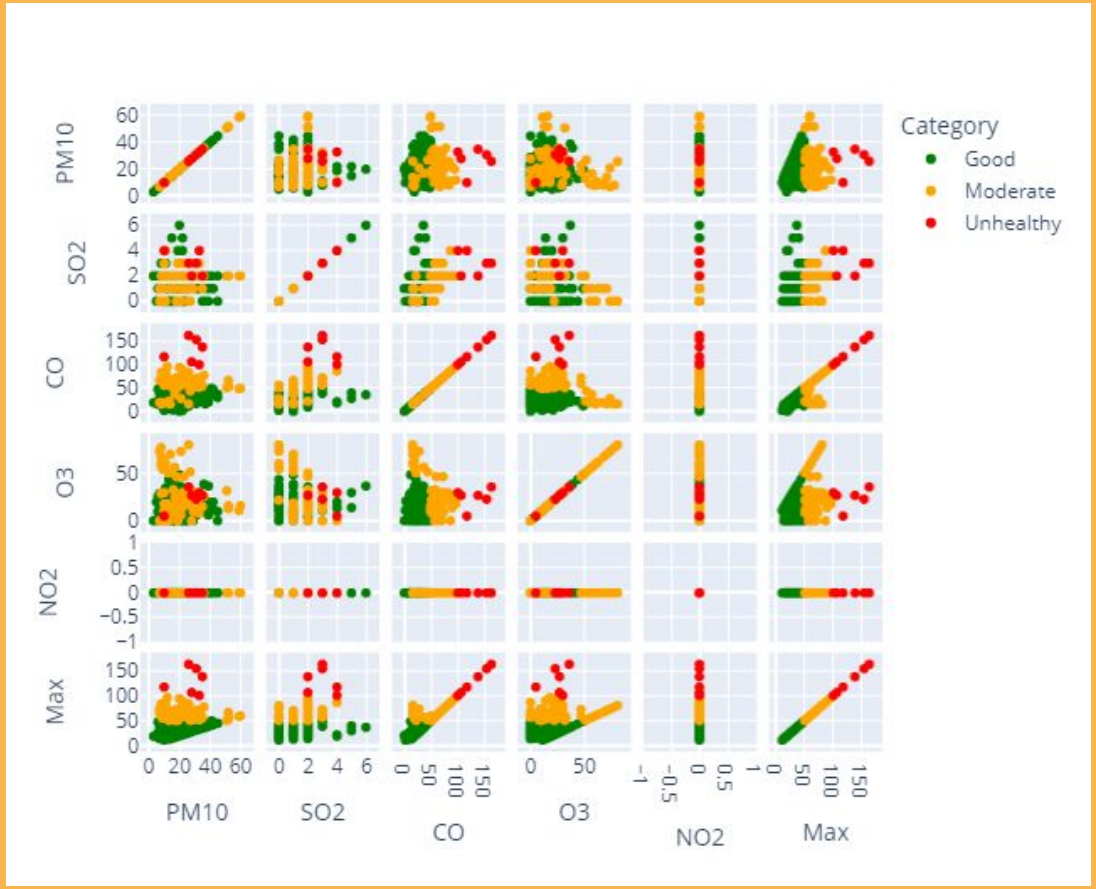
- It can be seen that there are outliers detected in some columns
- Except for column NO2, because it contains only 0 (zero) data
- However, this oulier data will not be discarded because we want to know the quality of the air produced

Air Quality Category and Critical Component

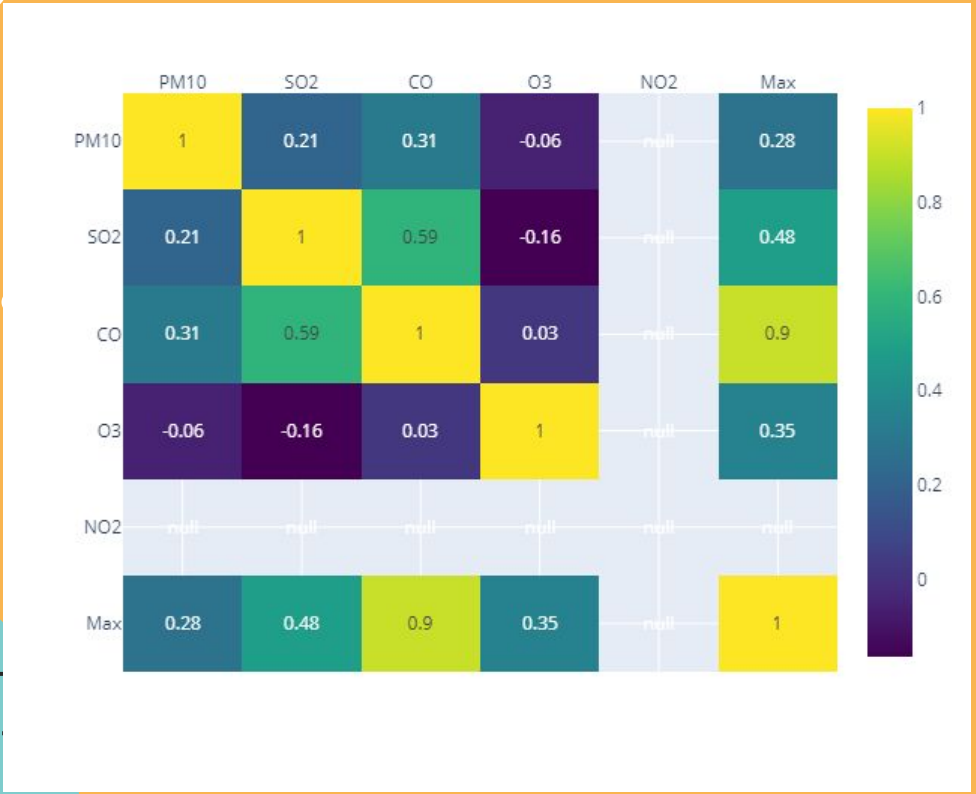


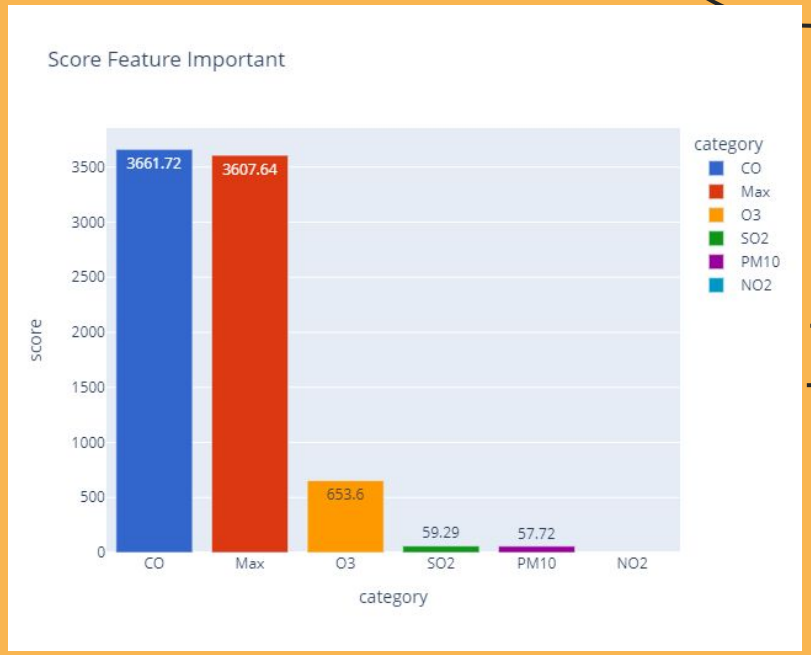


Air Quality Category and Critical Component



Air Quality Category and Critical Component

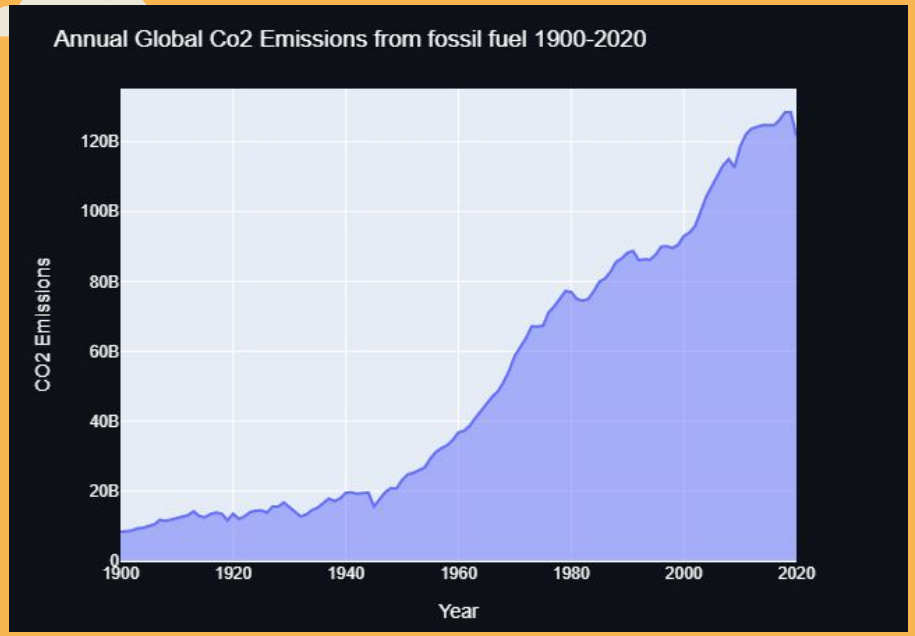


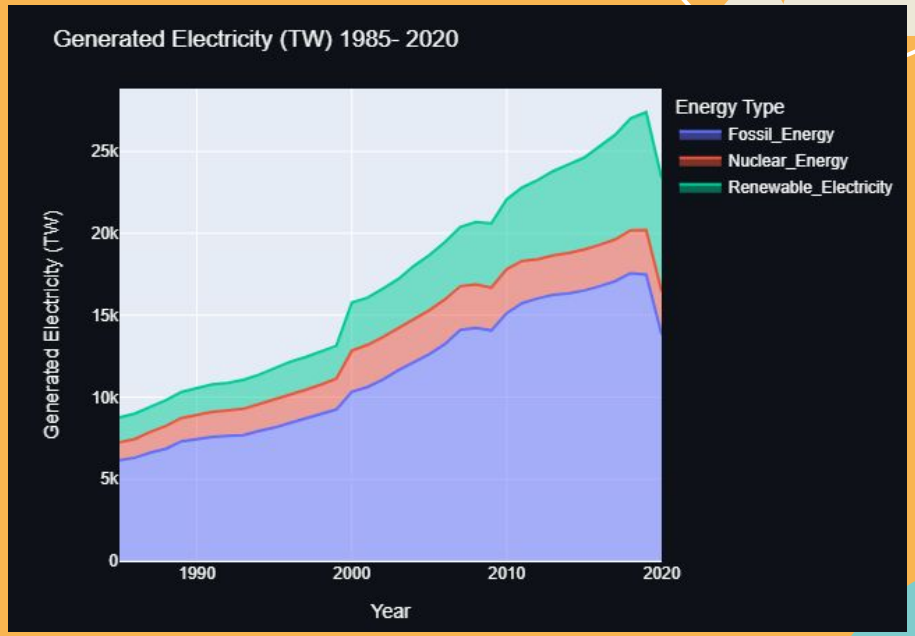


Air Quality Category and Critical Component

Setelah dicek pada tabel multicollinearity, Kolom CO dan Max mempunyai nilai yang tinggi dan kedua kolom tersebut berhubungan dengan target (Category). Saya memilih Max untuk didrop karena kolom CO diperlukan dalam modeling nanti dan kolom Max juga tidak berpengaruh, Karena isi dari kolom Max sudah diwakilkan oleh kolom-kolom polusi udara.

CO2 Emissions and Economic



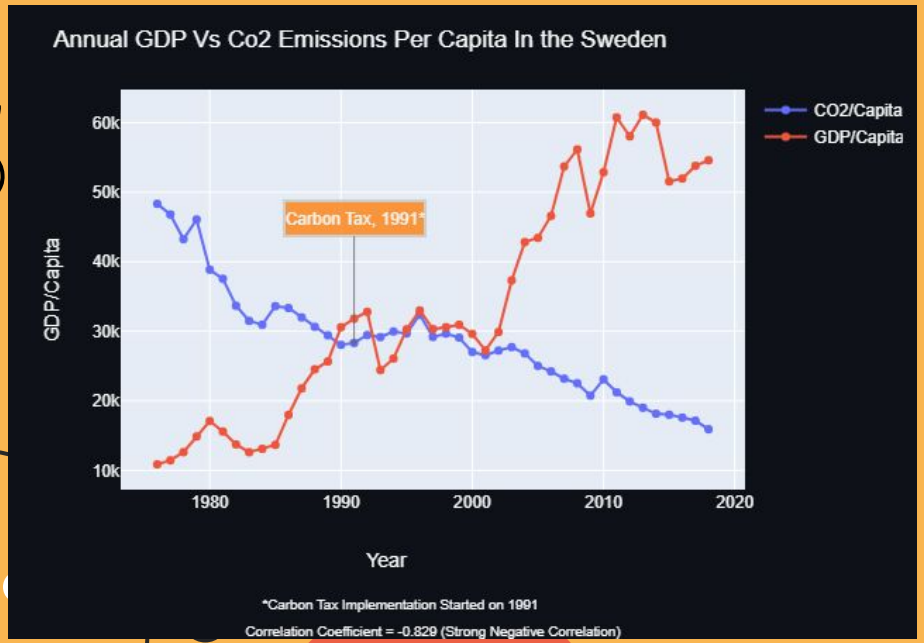


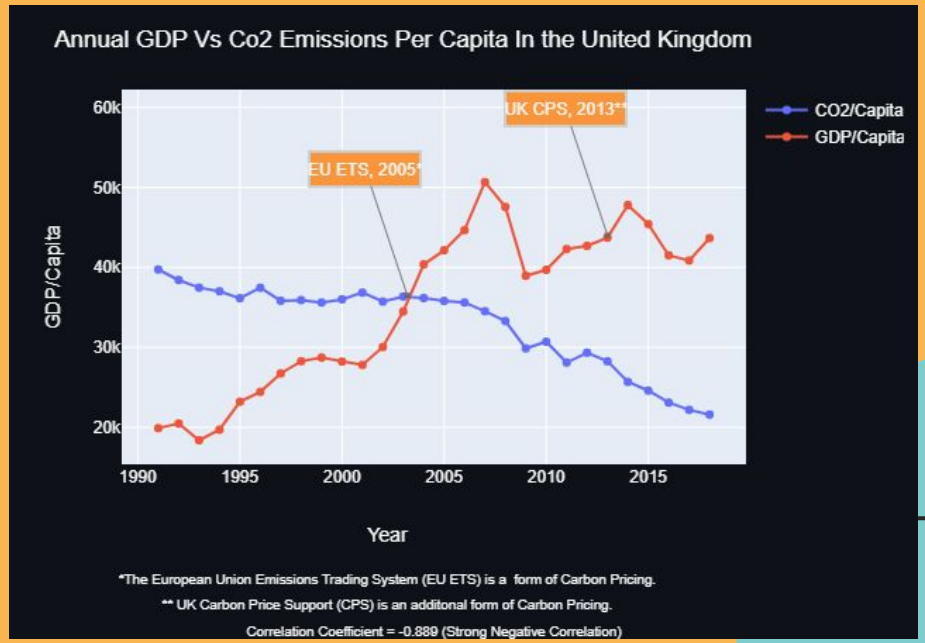
Fossil fuel has been the world's primary source of electricity. The majority or 50% of the installed capacity of power plants in Indonesia still comes from fossil energy.

CO2 Emissions and Economic

Carbon Tax

Carbon tax is a tax levied on the burning of carbon-based fuels such as coal, oil and gas. The carbon tax is a core policy created to reduce and eliminate the use of fossil fuels whose burning can damage the climate.





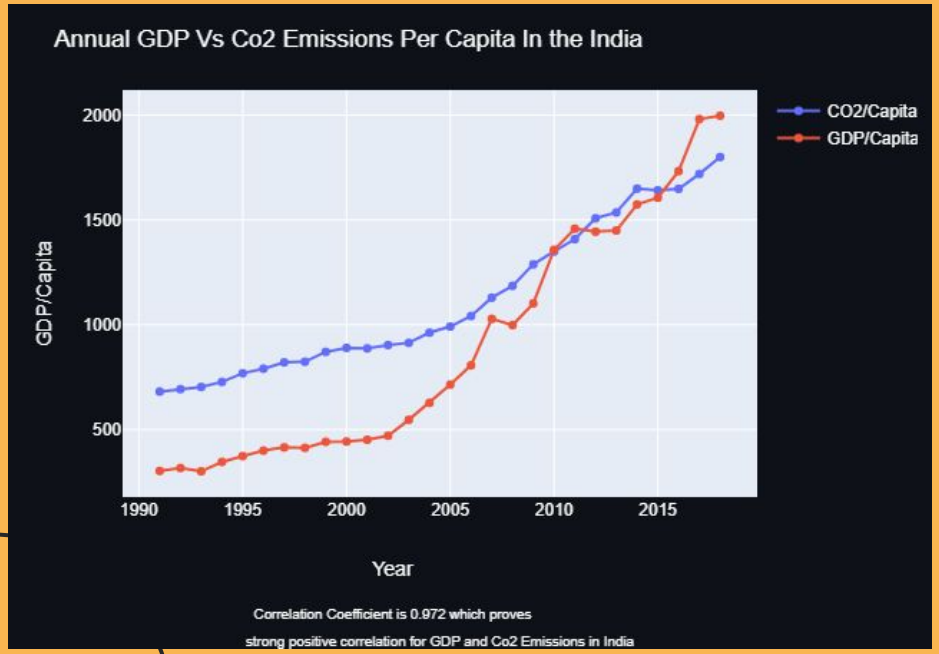
CO2 Emissions and Economic

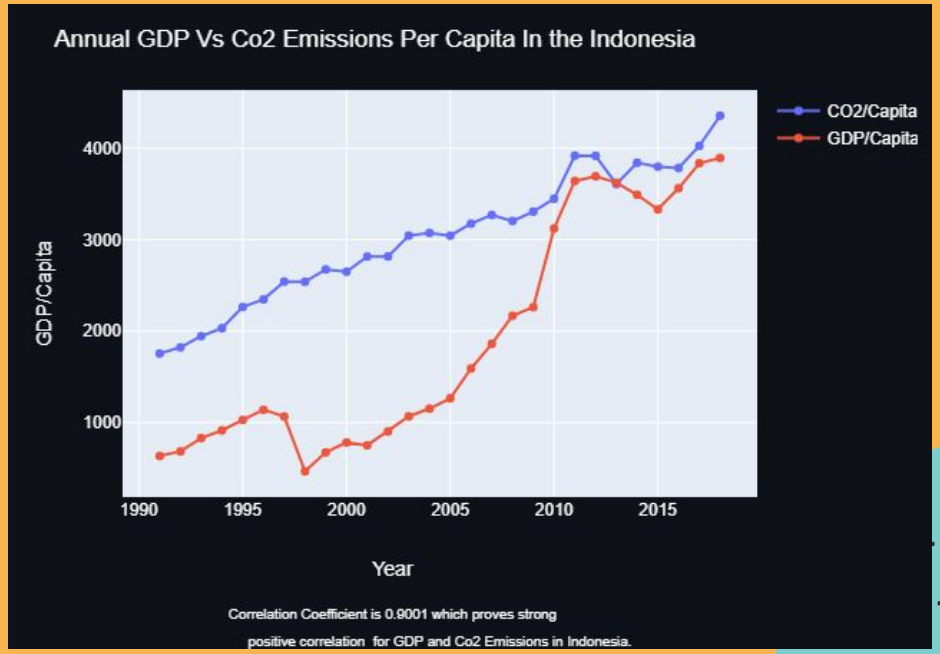
Summary : After Implementation of Carbon Tax, UK and Sweden Experience Increased Economic Activity (GDP) 👍 and Decreased Carbon Emissions 🔻

Some 40 countries and more than 20 cities, states and provinces already use carbon pricing mechanisms, with more planning to implement them in the future.

CO2 Emissions and Economic

NO carbon Tax





Summary : After seeing the increase in the value of Indonesia and India, Compared to the UK and Sweden which implemented the Carbon Tax, India and Indonesia's Carbon Emissions increased also related to the increase Putting a price on carbon can encourage low-carbon growth and lower greenhouse gas emissions. Putting a Price Tag on Carbon Reduces Carbon Emission and Supports Economic Growth.

Thanks