

PRESENTASI AKHIR

Peramalan Nilai Emisi CO₂ di Indonesia menggunakan Model ARIMA

oleh **Kelompok 3** – STK352 K2 Minor



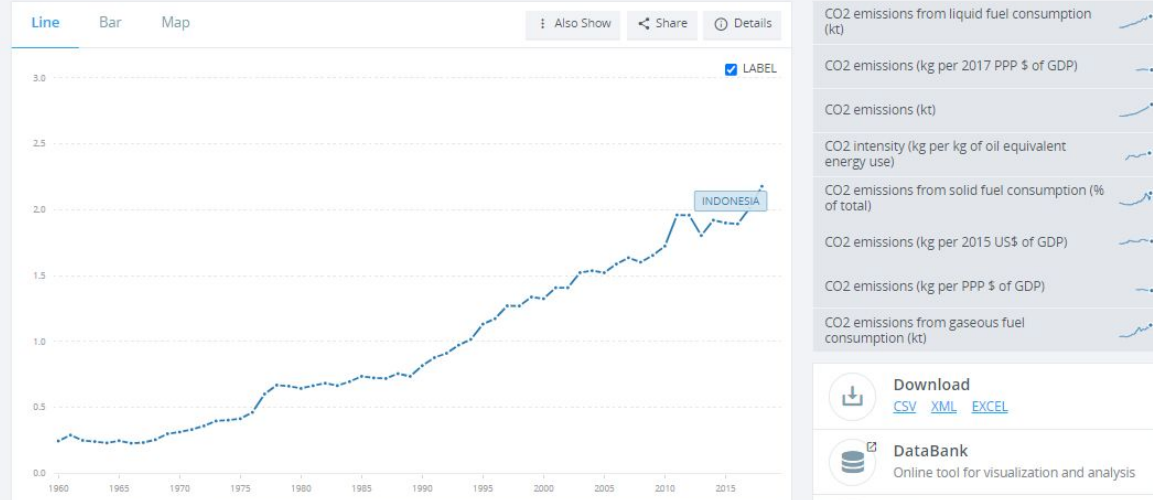
disusun oleh :

Nurul Fathi Qurrotul'aini	- G54180027
Ayudya Salsabila	- G54190060
Dimas Agus Putera	- G54190064
Laudza Muhammad Afif T.	- G64190052
M. Iqbal	- G64190074

CO2 emissions (metric tons per capita) - Indonesia

Data for up to 1990 are sourced from Carbon Dioxide Information Analysis Center, Environmental Sciences Division, Oak Ridge National Laboratory, Tennessee, United States. Data from 1990 are CAIT data: Climate Watch. 2020. GHG Emissions. Washington, DC: World Resources Institute. Available at: climatewatchdata.org/ghg-emissions. See SP.POP.TOTL for the denominator's source.

License: Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) [🔗](#)



Sumber Data

Data merupakan olahan dari Pusat Analisis Informasi Karbon Dioksida, Divisi Ilmu Lingkungan, Laboratorium Nasional Oak Ridge, Tennessee, Amerika Serikat.

Link : <https://data.worldbank.org/indicator/EN.ATM.CO2E.PC?locations=ID>

Eksplorasi Data

Data yang diambil dari sumber berupa berkas **.xlsx** yang menampilkan seluruh data nilai emisi CO₂ di **200 negara** dari tahun 1960–2018.

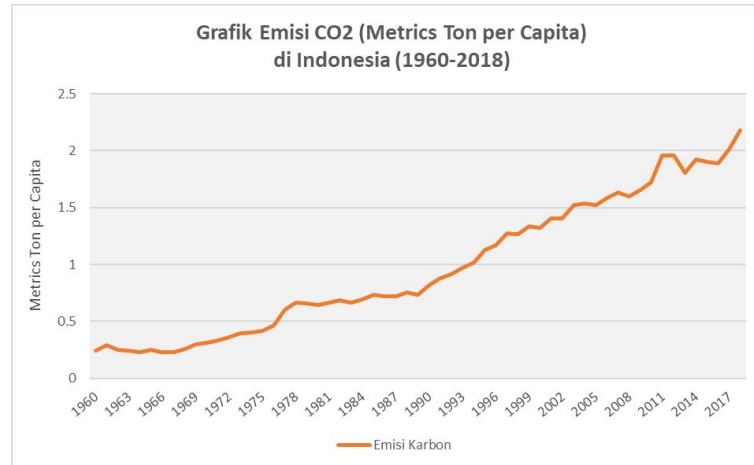
5	Country Name	Country C	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	1966
6	Aruba	ABW	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC							
7	Africa Eastern and Southern	AFE	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.90606	0.922474	0.930816	0.94057	0.996033	1.04728	1.033908
8	Afghanistan	AFG	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.046057	0.053589	0.073721	0.074161	0.086174	0.101285	0.107399
9	Africa Western and Central	AFW	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.09088	0.095283	0.096612	0.112376	0.133258	0.184803	0.193676
10	Angola	AGO	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.100835	0.082204	0.210533	0.202739	0.213562	0.205891	0.268937
11	Albania	ALB	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	1.258195	1.374186	1.439956	1.181681	1.111742	1.166099	1.333055
12	Andorra	AND	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC							
13	Arab World	ARB	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.609268	0.662618	0.727117	0.853116	0.972381	1.138674	1.251997
14	United Arab Emirates	ARE	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.119037	0.109136	0.163542	0.175833	0.132815	0.146822	0.160452
15	Argentina	ARG	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	2.383343	2.458551	2.538447	2.330685	2.553442	2.656466	2.806896
16	Armenia	ARM	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC							
17	American Samoa	ASM	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC							
18	Antigua and Barbuda	ATG	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.677418	0.866667	1.838457	1.487469	1.590448	2.561321	5.814611
19	Australia	AUS	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	8.582937	8.641569	8.835688	9.22644	9.759073	10.62232	10.32809
20	Austria	AUT	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	4.373319	4.496362	4.755362	5.155194	5.391004	5.252197	5.361725
21	Azerbaijan	AZE	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC							
22	Burundi	BDI	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC			0.015136	0.016081	0.015752	0.011851	0.015036
23	Belgium	BEL	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	9.941594	10.10387	10.64119	11.38693	11.05365	11.14166	11.04202
24	Benin	BEN	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.066354	0.052049	0.054209	0.047588	0.055304	0.057115	0.042383
25	Burkina Faso	BFA	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.009112	0.01873	0.017003	0.017504	0.021575	0.019841	0.019534
26	Bangladesh	BGD	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	0.294805	0.296924	0.317336	0.354674	0.354201	0.360641	0.355616
27	Bulgaria	BGR	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	2.833901	3.26992	3.835892	4.259781	5.262988	5.645652	5.905437
28	Bahrain	BHR	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	3.544435	10.54897	9.191856	6.710613	8.742028	6.554084	3.384305
29	Bahamas, The	BHS	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	3.749626	4.746245	5.995987	5.557806	8.118111	9.399207	7.465217

Pra-proses Data (I)

Dari keseluruhan data, dilakukan **reduksi** dan **transformasi** sehingga data hanya menampilkan nilai emisi CO₂ untuk negara **Indonesia** pada tahun 1960–2018.

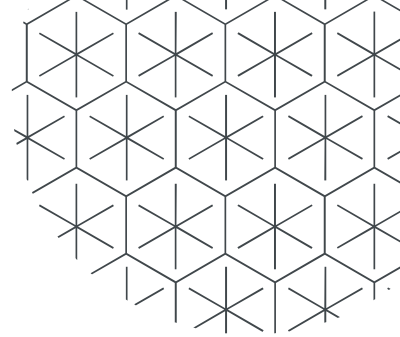
No.	Year	Emission
1	1960	0.243920444
2	1961	0.288847528
3	1962	0.248553409
4	1963	0.239783195
5	1964	0.229458195
6	1965	0.246241473
7	1966	0.227084338
8	1967	0.232007107
9	1968	0.253602314
10	1969	0.298784274
11	1970	0.312064909
12	1971	0.330738851
13	1972	0.358132823
14	1973	0.395605913

...



terlihat data memiliki **pola tren** positif.

Pra-proses Data (2)



Dari eksplorasi data sebelumnya, diperoleh informasi bahwa :

- Tidak ditemukan nilai Emisi CO₂ yang **kosong** (*NULL*) di setiap tahun.
- Melalui metode Visualisasi Data, Z-Scoring (threshold=3), dan Interquartile Range (IQR), tidak ditemukan **noise** maupun **outlier** pada data.
- Atribut yang akan diproses hanya **satu**, yakni Nilai Emisi CO₂.

Dapat disimpulkan bahwa data telah dapat dikatakan “bersih” dan tidak perlu dilakukan pra-proses data lebih lanjut.

Untuk membuat model **ARIMA** dan **Forecast** yang baik, maka data akan dipartisi dengan perbandingan antara Training dan Testing sebesar **8 : 2** karena ukuran data relatif kecil.

Peubah yang Diamati

Selain periode yang direpresentasikan sebagai “Tahun”, satu-satunya peubah yang lain yang ada pada data tersebut adalah **Nilai Emisi CO₂** dalam satuan **Metrics Ton per Capita**.

Apa maksud satuan pada peubah di atas?

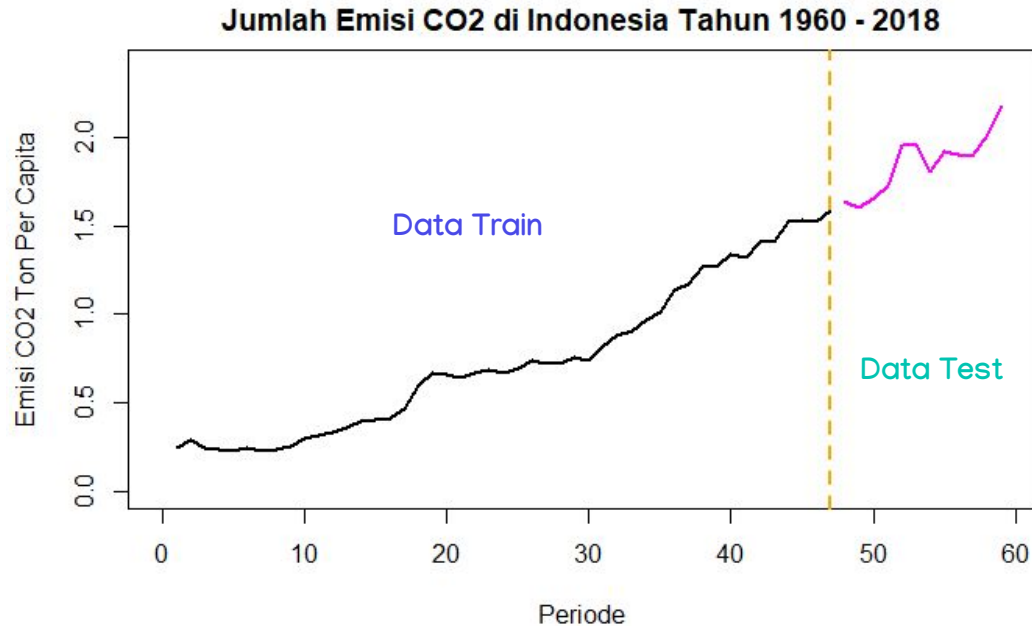
Dalam konteks Emisi CO₂, satuan tersebut berarti **kadar CO₂** sejumlah **1 Ton** (1000 kg) yang dihasilkan oleh suatu negara **per penduduk**-nya.

Peubah tersebut memiliki detail sebagai berikut.

Count	Mean	Median	Variance
59	0.974	0.755	0.349

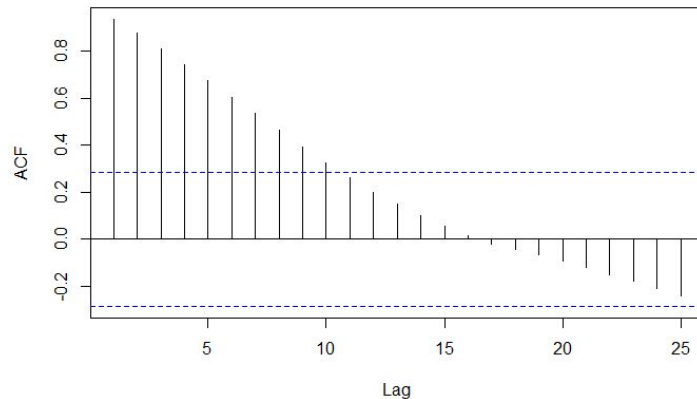
Visualisasi Data

Berikut adalah plot deret waktu dari data Emisi CO₂ yang kami gunakan.



Kestasioneran Deret Waktu

Berikut adalah visualisasi dari **Plot ACF** dan Output **uji Augmented Dickey-Fuller (ADF)** pada data train.



Pada plot dan output di samping/ bawah terlihat bahwa masih terjadi **tails-off secara perlahan pada sebaran lag** dan juga **p-value yang bernilai lebih dari 0.05**.

Kedua fakta tersebut mengindikasikan bahwa data deret waktu ini masih **belum stasioner**.

Augmented Dickey-Fuller Test

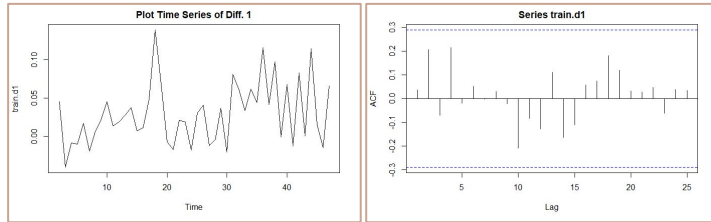
```
data: train
Dickey-Fuller = -1.6397, Lag order = 3, p-value = 0.7176
alternative hypothesis: stationary
```

Proses akan dilanjutkan ke **Differencing**.

Proses Differencing

Berikut adalah **Plot Deret Waktu Hasil Differencing, ACF**, dan hasil output dari uji **Augmented Dickey-Fuller Test** yang dilakukan terhadap data yang dihasilkan oleh proses **Differencing**.

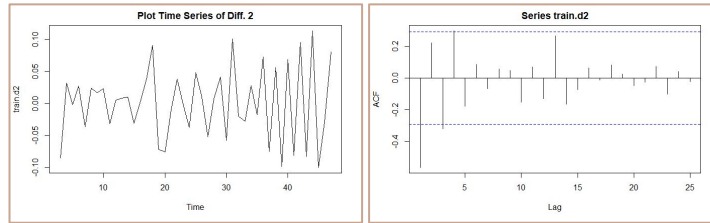
Diff. Orde 1



Augmented Dickey-Fuller Test

```
data: train.d1
Dickey-Fuller = -2.8125, Lag order = 3, p-value = 0.2509
alternative hypothesis: stationary
```

Diff. Orde 2



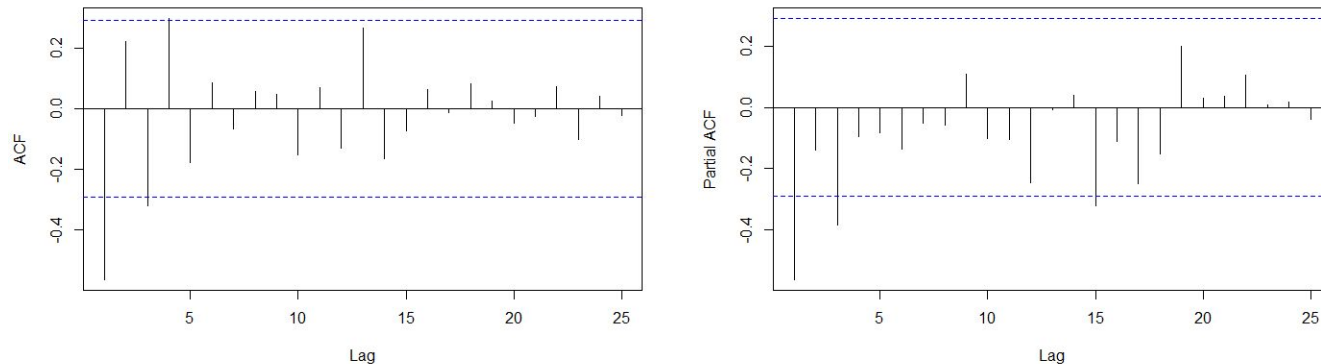
Augmented Dickey-Fuller Test

```
data: train.d2
Dickey-Fuller = -5.0048, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

Berdasarkan perbandingan kedua nilai orde pada proses **Differencing** di atas terlihat bahwa data deret waktu mulai dapat terindikasi stasioner saat nilai orde proses differencing bernilai 2. Oleh sebab itu, model dengan nilai **d=2** inilah yang akan digunakan dalam proses-proses berikutnya.

Analisis ACF dan PACF pada Model $d=2$

Berikut merupakan plot ACF dan PACF dari data deret waktu setelah dilakukan proses Differencing dengan orde yang bernilai 2 ($d=2$).



Berdasarkan dua plot di atas, dapat ditentukan model tentatif ARIMA sebagai berikut.

- $ARIMA(1,2,0)$
- $ARIMA(0,2,1)$

Identifikasi Model dengan EACF

Berikut merupakan plot EACF dari data deret waktu setelah dilakukan proses Differencing dengan orde yang bernilai 2 ($d=2$).

AR/MA		0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	o	x	o	o	o	o	o	o	o	o	o	o	o	o
1	o	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	o	o	x	o	o	o	o	o	o	o	o	o	o	o	o
4	x	o	x	o	o	o	o	o	o	o	o	o	o	o	o
5	o	x	x	o	o	o	o	o	o	o	o	o	o	o	o
6	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
7	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o

Berdasarkan plot EACF diatas, dapat disimpulkan bahwa model ARIMA tentatif yang didapatkan yaitu $ARIMA(1,2,2)$.

Model ARIMA Tentatif

Berdasarkan pengamatan plot ACF, PACF, dan juga ECF yang telah dilakukan sebelumnya, berikut ini merupakan **model ARIMA tentatif** yang akan dilakukan pendugaan parameter untuk dicari model yang terbaik.

- **ARIMA(1, 2, 0)**
- **ARIMA(0, 2, 1)**
- **ARIMA(1, 2, 2)**

Pada langkah selanjutnya, kelima model tentatif tersebut akan dilakukan proses **estimasi parameter** beserta **pengecekan nilai signifikansi**. Output yang diharapkan dari proses tersebut adalah model terbaik yang memiliki koefisien parameter (Log Likelihood dan AIC) yang signifikan.

Tabel Nilai Signifikansi Uji Parameter

Berikut adalah tabel nilai signifikansi untuk masing-masing model tentatif. Model yang diberi *highlight biru* menandakan bahwa model tersebut memiliki parameter yang *seluruhnya signifikan* dan dapat menjadi calon model terbaik yang akan dipilih nantinya.

Model Tentatif	Estimasi Parameter		Nilai Signifikansi
ARIMA(0, 2, 1)	MA1	-1	2.2e-16
ARIMA(1, 2, 0)	AR1	-0.8354	2.2e-16
ARIMA(1, 2, 2)	AR1	-0.5838	5.335e-06
	MA1	-1.9889	2.2e-16
	MA2	0.9999	2.2e-16

Tabel diatas menunjukkan bahwa ketiga model tentatif tersebut signifikan sehingga perlu pengecekan nilai AIC dan Log Likelihood.

Pemilihan Model ARIMA Terbaik (I)

Berdasarkan tabel signifikansi sebelumnya, didapat 3 model yang lolos dan akan dipilih sebagai model terbaik berdasarkan nilai **Log Likelihood** dan **AIC**-nya.

Model Tentatif	Log Likelihood	AIC
ARIMA(0, 2, 1)	36.27	-68.54
ARIMA(1, 2, 0)	38.33	-72.65
ARIMA(1, 2, 2)	64.35	-120.71*

Hasil perbandingan nilai Log Likelihood dan AIC pada ketiga model tentatif di atas menunjukkan bahwa **model terbaik** yang diperoleh adalah model **ARIMA(1 2, 2)**.

Pemilihan Model ARIMA Terbaik (2)

Selain menggunakan parameter nilai **Log Likelihood** dan **AIC** pada setiap model yang terpilih, pencarian model terbaik juga dicoba menggunakan fungsi `auto.arima()` yang disediakan oleh *library forecast* pada Rstudio.

Berikut adalah output yang diperoleh.

```
Series: train
ARIMA(0,2,1)

Coefficients:
      ma1
    -0.8893
s.e.    0.0838

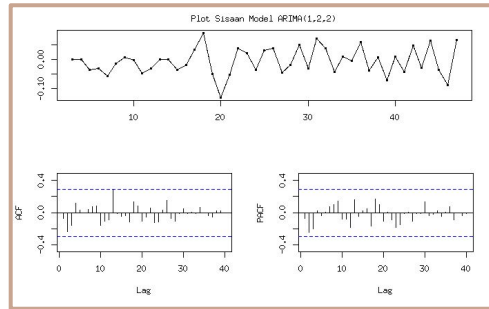
sigma^2 = 0.001695:  log likelihood = 79.42   AIC=-154.85   AICc=-154.56   BIC=-151.23
```

Terlihat bahwa pada penggunaan fungsi `auto.arima()` di atas menunjukkan bahwa **model terbaik** yang diperoleh adalah model **ARIMA(0, 2, 1)**.

Uji Kebebasan Sisaan pada Model

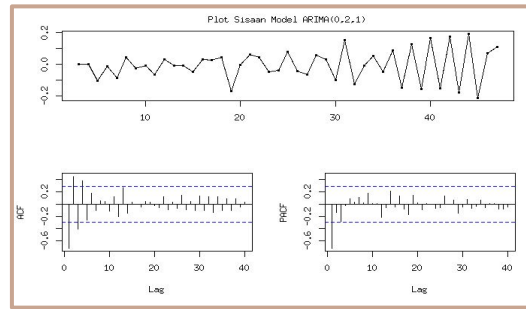
Karena diperoleh $ARIMA(0,2,1)$ dan $ARIMA(1,2,2)$ dari dua cara yang berbeda. Maka proses dilanjutkan dengan visualisasi plot sisaan beserta uji kebebasan sisaan pada model tersebut. Uji tersebut menggunakan **Ljung Box Test** yang tersedia dalam *library portes* di Rstudio. Berikut adalah perbandingan output-nya.

$ARIMA(1,2,2)$



lags	statistic	df	p-value
5	4.899970	5	0.4282097
10	7.180901	10	0.7082635
15	13.881321	15	0.5345485
20	17.900424	20	0.5939678
25	21.270399	25	0.6774392
30	26.132582	30	0.6683472

$ARIMA(0,2,1)$



lags	statistic	df	p-value
5	54.21591	5	1.892239e-10
10	57.58799	10	1.033454e-08
15	67.36213	15	1.312174e-08
20	67.83557	20	4.092801e-07
25	70.80584	25	2.919073e-06
30	78.07253	30	3.725572e-06

Pada $ARIMA(1,2,2)$

Setiap lag sisaan **Terima H_0** ($p\text{-value} > 0.05$), yang artinya **sisaan saling bebas**, yang artinya tidak terdapat autokorelasi pada sisaan. Dilakukan **Overfitting**.

Pada $ARIMA(0,2,1)$

Setiap lag sisaan **Tolak H_0** ($p\text{-value} < 0.05$), yang artinya **sisaan tidak saling bebas**, yang artinya terdapat autokorelasi pada sisaan.

Indikasi Overfitting

Dilakukan pengecekan *overfitting* terhadap model **ARIMA(1,2,2)** dengan menguji signifikansi model yang terdekat dengan model tersebut.

Model Tentatif	Estimasi Parameter		Nilai Signifikansi
ARIMA(2, 2, 2)	AR1	-0.662	3.77e-5
	AR2	-0.126	0.4285*
	MA1	-1.990	< 2.2e-16
	MA2	0.999	< 2.2e-16
ARIMA(1, 2, 3)	AR1	-0.189	0.391*
	MA1	-2.66	< 2.2e-16
	MA2	2.329	5.50e-5
	MA3	-0.66	0.014

Karena kedua model ARIMA di samping memiliki parameter yang **tidak signifikan**, artinya model yang dipilih tetap model

ARIMA(1,2,2).

Forecasting pada Data Train

Dengan menggunakan model **ARIMA(1, 2, 2)** dilakukan forecast terhadap periode setelah data train yang hasilnya nanti akan disandingkan dengan data aktual pada data test.

Hasil Forecast	Data Aktual (Test)	Error
1.4732	1.6352	-0.1620
1.6182	1.6018	0.0164
1.6117	1.6532	-0.0415
1.6937	1.7241	-0.0307
...
1.9783	1.8997	0.0859
2.0269	1.8924	0.0132
2.0767	2.1014	-0.1017

Mengukur Kebaikan Model

Berikut ditampilkan tabel nilai kebaikan model **ARIMA(1,2,2)** yang diterapkan pada forecasting data test (data aktual untuk periode tahun 2007 – 2018) menggunakan data train.

Nilai Kebaikan Model	
MAPE	4.2402 %
MSE	0.01134
MAD	0.07944

Nilai MAPE model **ARIMA(1,2,2)** adalah **4.2402%** dengan **MSE = 0.0113** dan **MAD = 0.07944**. Dapat disimpulkan bahwa model cukup baik dalam forecasting data.

Forecasting Periode ke Depan

Berikut ditampilkan tabel hasil forecast untuk 10 periode (tahun) ke depan.

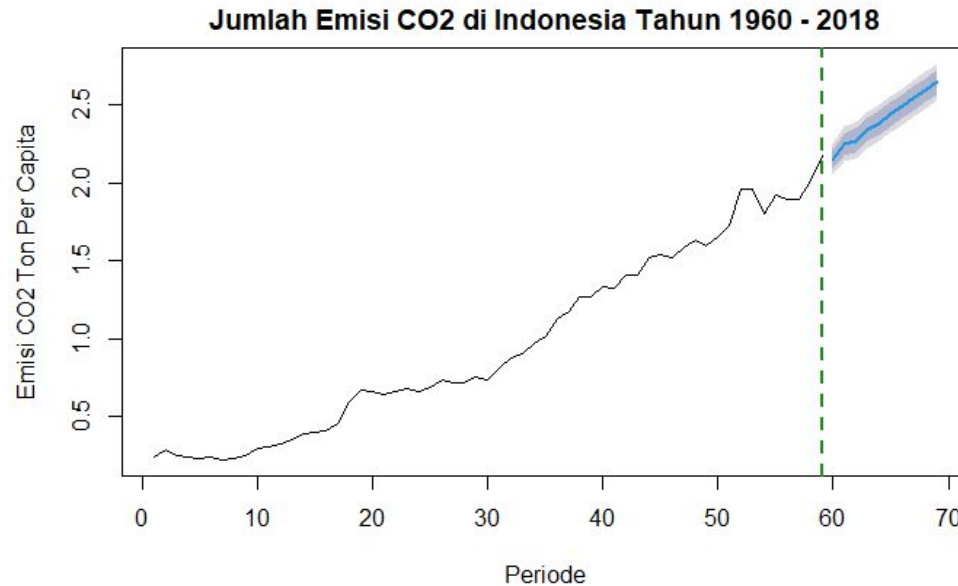
Tahun	Hasil Forecast
2019	2.1508
2020	2.2495
2021	2.2745
2022	2.3425
...	...
2026	2.5458
2027	2.5969
2028	2.6496

Dari pergerakan nilai forecast emisi CO₂ ton per kapita yang ditunjukkan pada tabel di samping, dapat dilihat bahwa data tersebut memiliki **tren yang positif** dari tahun ke tahun.

Hal ini tentu harus menjadi perhatian khusus bagi kita karena **jika lingkungan dibiarkan terus seperti ini, maka suatu saat bumi akan menjadi tidak layak huni.** Bahkan pada jangka waktu yang tidak lama lagi.

Forecasting Periode ke Depan

Berikut adalah visualisasi hasil forecast untuk 10 periode (tahun) ke depan.



Kesimpulan

- Dari sebanyak 6 model tentatif ARIMA yang diperoleh pada pengamatan plot ACF, PACF, dan EACF, model peramalan yang digunakan adalah model ARIMA(1,2,2) dengan MAPE sebesar 4.24%.
- Hasil peramalan banyaknya emisi CO₂ adalah emisi CO₂ akan terus meningkat dari tahun ke tahun. Hal ini berjalan sesuai dengan pola data itu sendiri, yakni Pola Tren Positif.

Implikasi dari meningkatnya jumlah CO₂ yang dilepas ke udara adalah efek rumah kaca. Apabila hal ini terus berlangsung dalam rentang waktu yang lama, maka akan terjadi kenaikan suhu permukaan bumi yang dapat mengakibatkan perubahan iklim.

Efek domino dari peristiwa tersebut adalah mencairnya es di kutub yang dapat mengakibatkan naiknya permukaan air laut. Menurunnya kualitas air, gelombang panas berkepanjangan, kekeringan, kebakaran hutan, cuaca yang sulit diprediksi, serta menurunnya intensitas panen bahan baku makanan.

Seluruh hal tersebut jelas mengancam kesejahteraan umat manusia di bumi.



Danke schön!

Is there any Question?



Data Emisi CO₂ di Indonesia (1960 - 2018)

Tahun	Nilai Emisi	Tahun	Nilai Emisi	Tahun	Nilai Emisi	Tahun	Nilai Emisi
1960	0.243920444	1975	0.412942076	1990	0.815761138	2005	1.521944445
1961	0.288847528	1976	0.461389594	1991	0.876744877	2006	1.587749693
1962	0.248553409	1977	0.600490434	1992	0.910355784	2007	1.635207076
1963	0.239783195	1978	0.667989729	1993	0.97185669	2008	1.601819308
1964	0.229458195	1979	0.660345907	1994	1.015535177	2009	1.653210477
1965	0.246241473	1980	0.642834916	1995	1.131545133	2010	1.724073581
1966	0.227084338	1981	0.663585649	1996	1.17297927	2011	1.960133257
1967	0.232007107	1982	0.682392176	1997	1.270396477	2012	1.959133194
1968	0.253602314	1983	0.664245133	1998	1.269366929	2013	1.804251041
1969	0.298784274	1984	0.694540628	1999	1.337342815	2014	1.920799967
1970	0.312064909	1985	0.734900509	2000	1.324594286	2015	1.899658692
1971	0.330738851	1986	0.723036368	2001	1.407609164	2016	1.89244089
1972	0.358132823	1987	0.718523982	2002	1.40809308	2017	2.01367107
1973	0.395605913	1988	0.755317201	2003	1.522267724	2018	2.178461553
1974	0.402292488	1989	0.734905627	2004	1.536865336		