



PREDIKSI TINGKAT INFLASI SEBAGAI LANGKAH PEMULIHAN EKONOMI BERDASARKAN KONDISI PANDEMI COVID-19

Calculus Clan

Data Science Academy COMPFEST 2023





Problem Understanding - BACKGROUND

Pandemi COVID-19 telah mempengaruhi berbagai macam aspek dalam kehidupan, tak terkecuali dalam bidang ekonomi. Untuk mendukung pemulihan ekonomi, penting bagi kita untuk dapat mengambil keputusan terbaik yang berdasarkan data. Salah satu faktor yang dapat mempengaruhi pemulihan ekonomi adalah pengaruh inflasi. Inflasi yang terkendali dapat berkontribusi untuk memulihkan ekonomi. Pengaruh inflasi terhadap pemulihan ekonomi negara juga dapat bervariasi tergantung pada berbagai faktor dan kondisi ekonomi tertentu.

Dengan adanya kontribusi dari pengaruh inflasi terhadap pemulihan ekonomi ini, kami ingin membuat model untuk melakukan prediksi tingkat inflasi berdasarkan kondisi pandemi COVID-19. Model ini dibangun dengan asumsi bahwa data kasus Covid-19 akan berpengaruh terhadap mobilitas yang dapat mempengaruhi tingkat inflasi.

Dengan menggunakan dataset Kasus Covid Bulanan, *Mobility Report*, Nilai Neraca Perdagangan, dan Indeks Harga Konsumen, kami yakin bahwa dataset tersebut akan menghasilkan insight yang berguna untuk membangun model prediksi tingkat inflasi.

OUR OBJECTIVES

Goals

- Mencari variabel yang berpengaruh besar terhadap tingkat inflasi
- 2) Memprediksi tingkat inflasi berdasarkan variabel yang paling berpengaruh
- 3) Membantu para pengambil keputusan kebijakan ekonomi dengan menggunakan model prediksi yang akan dibangun

Hipotesis



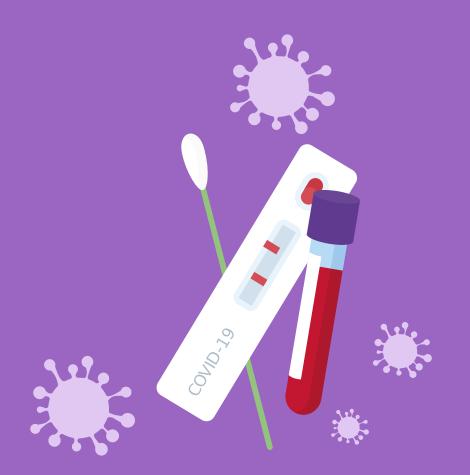


Berdasarkan pengecekan korelasi awal, 3 variabel independen yang memiliki korelasi tinggi dengan target variabel inflasi adalah 'IHK', 'Perubahan persentase mobilitas di rekreasi dan retail', dan 'perubahan persentase mobilitas di toko kebutuhan dan apotek'.



Possible Issues

 Data yang digunakan pada saat kondisi pandemi hanya tersedia dari bulan Agustus 2020. Sehingga model yang akan dibuat hanya dapat memprediksi sesuai kondisi pandemi dari bulan Agustus 2020 saja.



01 Initial Data Understanding

Setelah dilakukan Preprocessing data, kami akan melakukan pemahaman terhadap data terlebih dahulu dengan melakukan EDA.

LET'S START!

DATASET CHARACTERISTIC

```
Int64Index: 28 entries, 0 to 27
Data columns (total 11 columns):
    Column
                                                                    Non-Null Count Dtype
    Bulan
                                                                    28 non-null
                                                                                    object
    Kasus Positif Covid
                                                                                    float64
                                                                    28 non-null
    Inflasi
                                                                                    float64
                                                                    28 non-null
    Perubahan Persentase Mobilitas di Rekreasi dan Retail
                                                                    28 non-null
                                                                                    float64
    Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek
                                                                    28 non-null
                                                                                    float64
    Perubahan Persentase Mobilitas Masyarakat di Taman
                                                                    28 non-null
                                                                                    float64
    Perubahan Persentase Mobilitas Masyarakat di Stasiun
                                                                    28 non-null
                                                                                    float64
    Perubahan Persentase Mobilitas Masyarakat di Kantor
                                                                    28 non-null
                                                                                    float64
    Perubahan Persentase Mobilitas Masyarakat di Residential Area
                                                                                    float64
                                                                    28 non-null
    Data Neraca Perdagangan
                                                                    28 non-null
                                                                                    float64
    Data Index Harga Konsumen
                                                                    28 non-null
                                                                                    float64
dtypes: float64(10), object(1)
```

Dataset tersebut mengandung beberapa tipe data, termasuk numerik dan kategorik. Kami mengubah tipe data untuk variabel Neraca Perdagangan dan Index Harga Konsumen dari object menjadi float agar bisa dihitung correlationnya.



DESC STATS FOR ALL VARIABLE

	Kasus Positif Covid	Inflasi	Perubahan Persentase Mobilitas di Rekreasi dan Retail	Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek	Perubahan Persentase Mobilitas Masyarakat di Taman	Perubahan Persentase Mobilitas Masyarakat di Stasiun	Perubahan Persentase Mobilitas Masyarakat di Kantor	Perubahan Persentase Mobilitas Masyarakat di Residential Area	Data Neraca Perdagangan	Data Index Harga Konsumen
count	2.800000e+01	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000
mean	2.341596e+05	1.964643	-1.943716	18.623228	9.859853	-21.062582	-10.038101	10.043683	3.479464	107.960357
std	3.156067e+05	1.235330	13.688839	14.633296	27.496832	12.399537	15.766596	3.786173	1.686520	2.633074
min	6.311000e+03	-0.100000	-24.322581	-8.612903	-24.964286	-45.258065	-28.935484	4.580645	0.963000	104.850000
25%	6.517875e+04	1.120000	-15.010484	3.774194	-10.941667	-30.653226	-23.075000	6.583333	2.239750	106.127500
50%	1.270490e+05	1.605000	0.977381	24.035484	0.903226	-23.355952	-14.111290	9.854839	3.208000	106.645000
75%	2.173162e+05	3.170000	10.325806	30.533333	35.702688	-9.650269	5.425000	13.533333	4.828750	110.090000
max	1.231386e+06	3.850000	18.645161	39.645161	76.967742	-0.709677	18.766667	16.933333	7.565000	112.870000



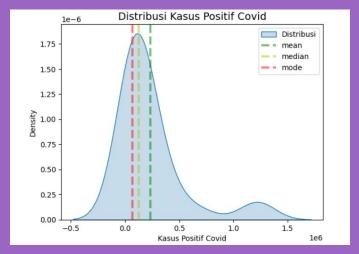


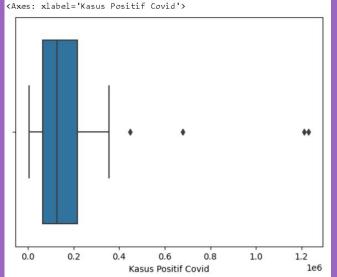
KASUS POSITIF COVID-19

count 2.800000e+01 2.341596e+05 mean std 3.156067e+05 min 6.311000e+03 25% 6.517875e+04 50% 1.270490e+05 75% 2.173162e+05 1.231386e+06 max

Name: Kasus Positif Covid, dtype: float64

Dari tabel di atas, didapat bahwa mean untuk Kasus Covid-19 berada di angka 234,159.6 lalu untuk standard deviasinya berada di angka 315,606.7 dan kasus paling sedikit yang ditemukan dalam rentang waktu Agustus 2020 hingga November 2022 adalah sebanyak 6,311 kasus.



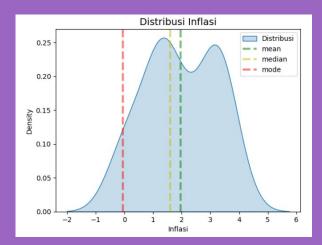


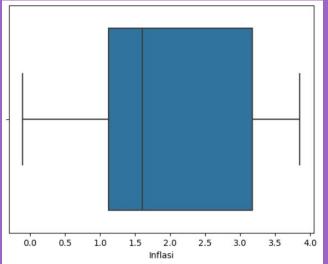


DATA INFLASI

count	28.000000
mean	1.964643
std	1.235330
min	-0.100000
25%	1.120000
50%	1.605000
75%	3.170000
max	3.850000
Name:	Inflasi, dtype: float64

Dari tabel di atas, didapat bahwa mean untuk inflasi berada di angka 1.964643 lalu untuk standard deviasinya berada di angka 1.235330 dan didapat inflasi terendah berada di angka -o.1 yang mana mengartikan deflasi.



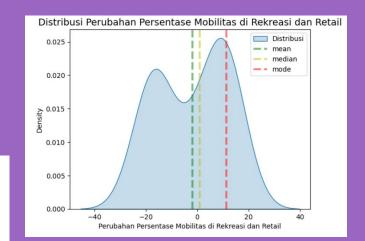


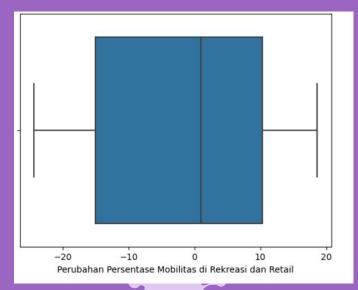
PERUBAHAN PERSENTASE MOBILITAS DI REKREASI DAN RETAIL

28.000000 count -1.943716 mean std 13.688839 min -24.322581 25% -15.010484 50% 0.977381 75% 10.325806 18.645161 max

Name: Perubahan Persentase Mobilitas di Rekreasi dan Retail

Dari tabel di atas, didapat bahwa mean untuk perubahan persentase mobilitas di rekreasi dan retail berada di angka -1.943716% lalu untuk standard deviasinya berada di angka 13.688839% dan didapat perubahan persentase terendah berada di angka -24.322581% sedangkan tertinggi 18.645161%. Hal ini menunjukkan bahwa perubahan persentase mobilitas di tempat rekreasi dan retail menunjukkan angka yang cukup signifikan.



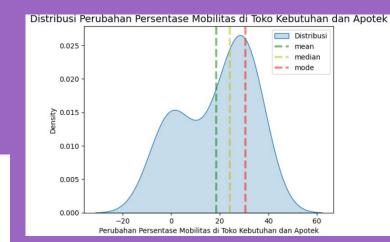


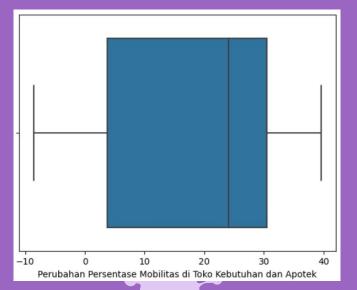
PERUBAHAN PERSENTASE MOBILITAS DI TOKO KEBUTUHAN DAN APOTEK

28.000000 count 18,623228 mean std 14.633296 min -8.612903 25% 3.774194 50% 24.035484 75% 30.533333 39.645161 max

Name: Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek

Dari tabel di atas, didapat bahwa mean untuk perubahan persentase mobilitas di toko kebutuhan dan apotek berada di angka 18.623228% lalu didapat perubahan persentase terendah berada di angka -8.612903%, sedangkan tertinggi 39.645161%. Dari karakteristik ini didapatkan bahwa selama pandemi covid, mobilitas di toko kebutuhan dan apotek cukup tinggi.



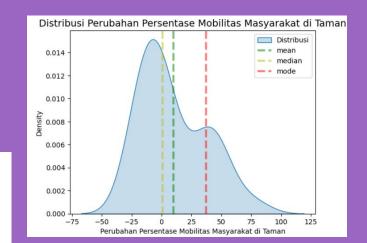


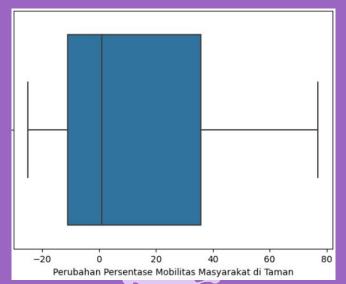
PERUBAHAN PERSENTASE MOBILITAS DI TAMAN

count	28.000000
mean	9.859853
std	27.496832
min	-24.964286
25%	-10.941667
50%	0.903226
75%	35.702688
max	76.967742
	32 30

Name: Perubahan Persentase Mobilitas Masyarakat di Taman

Dari tabel di atas, didapat bahwa mean untuk perubahan persentase mobilitas di taman berada di angka 9.859853% lalu didapat perubahan persentase terendah berada di angka -24.964286%, sedangkan tertinggi 76.967742%. Dari karakteristik ini didapatkan bahwa selama pandemi covid, mobilitas di taman mengalami perubahan yang sangat tinggi dilihat dari persentase perubahan maksimal yang dicapai adalah hampir sebesar 77%.



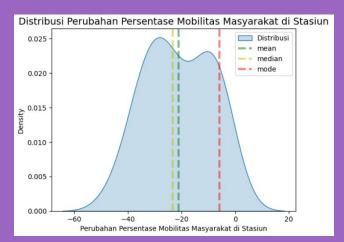


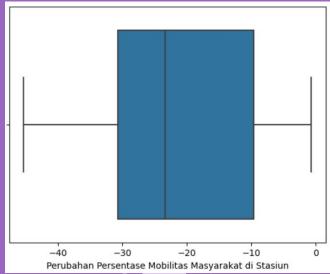
PERUBAHAN PERSENTASE MOBILITAS DI STASIUN

28.000000 count -21.062582 mean std 12.399537 min -45.258065 25% -30.653226 50% -23.355952 75% -9.650269 -0.709677 max

Name: Perubahan Persentase Mobilitas Masyarakat di Stasiun,

Dari tabel di atas, didapat bahwa mean untuk perubahan persentase mobilitas di stasiun berada di angka-21.062582% lalu didapat perubahan persentase terendah berada di angka-45.258065%, sedangkan tertinggi -0.709677%. Dari karakteristik ini didapatkan bahwa selama pandemi covid, mobilitas di stasiun mengalami perubahan yang besar, cenderung mengalami perubahan persentase yang minus dimana mengartikan penurunan atau pengurangan mobilitas.



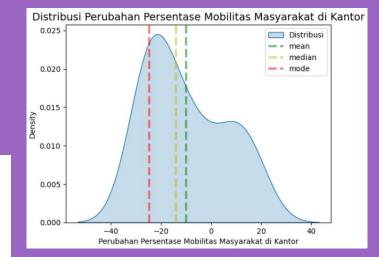


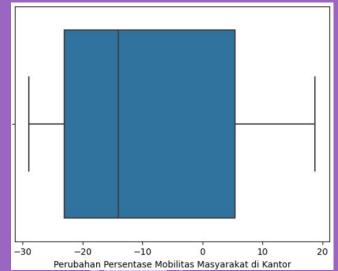
PERUBAHAN PERSENTASE MOBILITAS DI KANTOR

count 28.000000 -10.038101 mean std 15.766596 min -28.935484 25% -23.075000 50% -14,111290 75% 5.425000 18.766667 max

Name: Perubahan Persentase Mobilitas Masyarakat di Kantor

Dari tabel di atas, didapat bahwa mean untuk perubahan persentase mobilitas di kantor berada di angka -10..038101% lalu didapat perubahan persentase terendah berada di angka -28.935484%, sedangkan tertinggi 18.76667%. Dari karakteristik ini didapatkan bahwa selama pandemi covid, mobilitas di kantor mengalami perubahan yang cukup signifikan. Jika dilihat dari perubahan persentase meannya, rata-rata mobilitas di kantor mengalami pengurangan.



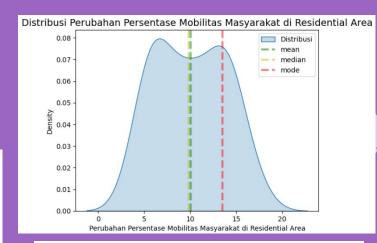


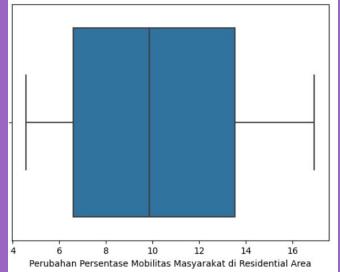
PERUBAHAN PERSENTASE MOBILITAS DI RESIDENTIAL AREA

28.000000
10.043683
3.786173
4.580645
6.583333
9.854839
13.533333
16.933333

Name: Perubahan Persentase Mobilitas Masyarakat di Residential Area,

Dari tabel di atas, didapat bahwa mean untuk perubahan persentase mobilitas di Residential Area berada di angka 10.043683% lalu didapat perubahan persentase terendah berada di angka 4.580645%, sedangkan tertinggi 16.93333%. Dari karakteristik ini didapatkan bahwa selama pandemi covid, mobilitas di Residential Area mengalami perubahan mobilitas yang tidak terlalu tinggi. Perubahan yang didapatkan masih antara kisaran 4-17%.

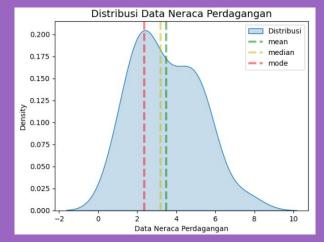


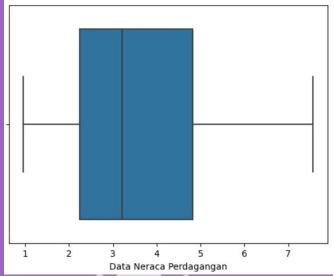


DATA NERACA PERDAGANGAN

count	28.000000			
mean	3.479464			
std	1.686520			
min	0.963000			
25%	2.239750			
50%	3.208000			
75%	4.828750			
max	7.565000			
Name:	Data Neraca Perdagangan,	dtype:	float64	

Dari tabel di atas, didapat bahwa mean untuk neraca perdagangan berada di angka 3.479464 lalu didapat perubahan persentase terendah berada di angka 0.963, sedangkan tertinggi 7.565. Dari karakteristik ini menunjukkan nilai minimum dari neraca perdagangan tidak mencapai nilai minus yang dimana menunjukkan tidak terjadi defisit selama pandemi.

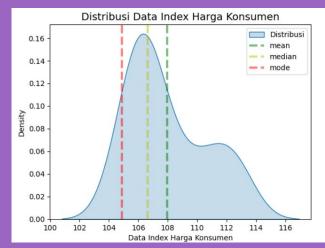


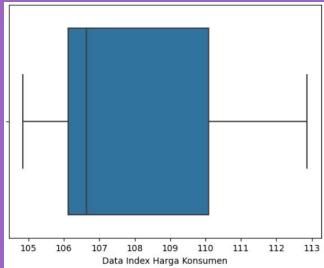


DATA INDEX HARGA KONSUMEN

count	28.000000			
mean	107.960357			
std	2.633074			
min	104.850000			
25%	106.127500			
50%	106.645000			
75%	110.090000			
max	112.870000			
Name:	Data Index Harga	Konsumen,	dtype:	float64

Dari tabel di atas, didapat bahwa mean untuk neraca perdagangan berada di angka 107.960357 lalu didapat perubahan persentase terendah berada di angka 104.85 sedangkan tertinggi 112.87. Dari karakteristik ini menunjukkan rata-rata indeks harga konsumen mencapai hingga angka 108, yang dimana mengartikan perubahan rata-rata harga sekelompok barang dan jasa yang biasa dikonsumsi oleh rumah tangga mengalami perubahan besar hingga lebih dari 2 kali lipat.







KORELASI DENGAN INFLASI

Dapat kita lihat dari hasil korelasi di samping, bahwa inflasi memiliki hubungan dengan yang kuat variabel Index Harga Konsumen, perubahan persentase mobilitas di rekreasi dan retail. perubahan persentase mobilitas di toko kebutuhan dan apotek. Namun hasil korelasi ini adalah hasil sebelum dilakukan feature scalling.



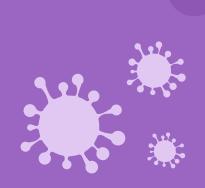




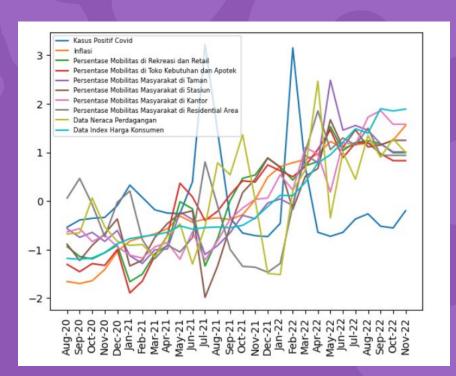
Dilakukan feature scaling karena perubahan persentase itu dalam satuan persen, sementara data neraca perdagangan itu dalam milyar rupiah. Singkatnya adalah memiliki satuan dasar yang berbeda.







SETELAH DILAKUKAN FEATURE SCALLING, DIDAPATKAN



standard deviasinya:	
Kasus Positif Covid	1.01835
Inflasi	1.01835
Data Neraca Perdagangan	1.01835
Data Index Harga Konsumen	1.01835
Perubahan Persentase Mobilitas di Rekreasi dan Retail	1.01835
Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek	1.01835
Perubahan Persentase Mobilitas Masyarakat di Taman	1.01835
Perubahan Persentase Mobilitas Masyarakat di Stasiun	1.01835
Perubahan Persentase Mobilitas Masyarakat di Kantor	1.01835
Perubahan Persentase Mobilitas Masyarakat di Residential Area	1.01835

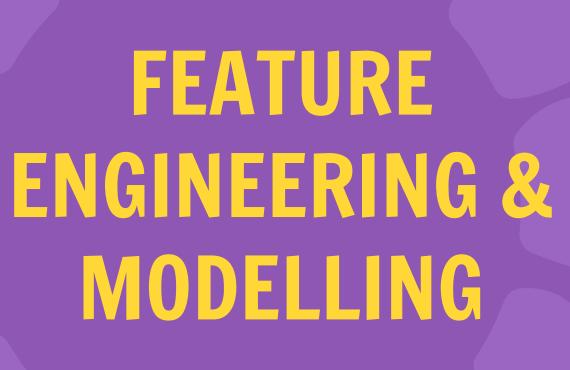


SETELAH DILAKUKAN FEATURE SCALLING, DIDAPATKAN





Dapat dilihat bahwa setelah dilakukan Feature Scalling, feature yang berkorelasi paling kuat adalah data IHK, data neraca perdagangan, dan perubahan persentase mobilitas masyarakat di residential area.









Dilakukan pemodelan dengan Multiple Linear Regression

```
#multiple Linear Regression
y = data standard.Inflasi
X = data standard.drop(columns = ['Inflasi', 'Bulan'])
#Pake validation train-test split
X train, X test, y train, y_test = train_test_split(X, y, test_size=2/5, random_state=0)
#Define model
multiple linear regressor = LinearRegression()
#Fit model
model = multiple linear regressor.fit(X train, y train)
prediksi_inflasi = model.predict(X_test)
print("MAE: ", mean absolute error(y test, prediksi inflasi))
print("MSE: ", mean squared error(y test, prediksi inflasi))
print('RMSE:', np.sqrt(mean squared error(y test, prediksi inflasi)))
print("R Squared scores: ", model.score(X train, y train))
print()
hasil = pd.DataFrame(prediksi inflasi)
y test = pd.DataFrame(y test)
v test = v test.reset index()
hasil = y test.join(hasil)
hasil = hasil.drop(columns = ['index'])
hasil.rename(columns = {0 : 'Prediksi'}, inplace = True)
hasil
```

Didapatkan

MAE: 0.2618604996710057 MSE: 0.11489109430731353 RMSE: 0.3389558884387665

R Squared scores: 0.9748008296195679

Inflasi Prediksi

	11111451	ri conkan
0	-1.644293	-1.305066
1	1.092564	0.919721
2	-0.284108	-0.105342
3	0.705118	0.413163
4	-0.828182	-1.372702
5	-0.374787	-0.379843
6	1.183243	1.166534
7	-0.383030	-0.441382
8	0.960668	0.218829
9	1.150269	1.486554
10	0.490785	0.106509
11	-1.701998	-1.774496



Dilakukan Crossing Val Score

```
scores_mae = -1 * cross_val_score(multiple_linear_regressor, X_train, y_train, cv = 5, scoring = 'neg_mean_absolute_error')
print("MAE scores : \n", scores mae)
print("Average MAE Score (Across experminets): ")
print(scores mae.mean())
print()
scores_r2 = cross_val_score(multiple_linear_regressor, X_train, y_train, scoring='r2', cv=5)
print("R squared scores: \n", scores r2)
print("Average R squared score (Across experiments): ")
print(scores r2.mean())
```

Didapatkan...

```
MAE scores :
 [0.16412024 0.32584232 0.54755034 0.47609977 1.27355268]
Average MAE Score (Across experminets):
0.5574330710230251
R squared scores:
 [ 0.95464127  0.77033308 -0.16538134  0.80824102 -3.02719487]
Average R squared score (Across experiments):
-0.13187217020555558
```





Akan dilakukan feature selection menggunakan metode Backward Elimination untuk menyeleksi feature yang lebih berpengaruh terhadap model

1. Pertama

Model:	OLS	Adj. R-squared:	0.941
Dependent Variable	: Inflasi	AIC:	8.9824
Date:	2023-07-06 13:41	BIC:	22.3044
No. Observations:	28	Log-Likelihood:	5.5088
Df Model:	9	F-statistic:	48.63
Df Residuals:	18	Prob (F-statistic):	8.36e-11
R-squared:	0.960	Scale:	0.061450

Intercept Kasus Positif Covid Data Neraca Perdagangan Data Index Harga Konsumen

Perubahan Persentase Mobilitas di Rekreasi dan Retail Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek

Perubahan Persentase Mobilitas Masyarakat di Taman Perubahan Persentase Mobilitas Masyarakat di Stasiun Perubahan Persentase Mobilitas Masyarakat di Kantor Perubahan Persentase Mobilitas Masyarakat di Residential

 Omnibus:
 2.459 Durbin-Watson:
 1.112

 Prob(Omnibus):
 0.292 Jarque-Bera (JB):
 1.707

 Skew:
 0.400 Prob(JB):
 0.426

 Kurtosis:
 2.092 Condition No.:
 38

Coef.	Std.Err.	t	P> t	[0.025	0.975]
0.0000	0.0468	0.0000	1.0000	-0.0984	0.0984
0.1914	0.0845	2.2644	0.0361	0.0138	0.3690
0.3049	0.5817	0.5242	0.6065	-0.9172	1.5271
0.1997	0.3291	0.6066	0.5517	-0.4918	0.8912
-0.2934	0.3001	-0.9774	0.3413	-0.9239	0.3372
0.2980	0.3018	0.9874	0.3365	-0.3361	0.9322
-0.1753	0.2270	-0.7724	0.4499	-0.6521	0.3015
0.0529	0.1748	0.3026	0.7656	-0.3144	0.4202
-0.0439	0.0834	-0.5268	0.6047	-0.2191	0.1313
0.7009	0.1776	3.9470	0.0009	0.3278	1.0739





2. Kedua

Model:	OLS	Adj. R-squared:	0.944
Dependent Variable	: Inflasi	AIC:	7.1245
Date:	2023-07-06 13:41	BIC:	19.1143
No. Observations:	28	Log-Likelihood:	5.4378
Df Model:	8	F-statistic:	57.44
Df Residuals:	19	Prob (F-statistic):	1.10e-11
R-squared:	0.960	Scale:	0.058512

Intercept Kasus Positif Covid Data Neraca Perdagangan Data Index Harga Konsumen

Perubahan Persentase Mobilitas di Rekreasi dan Retail Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek

Perubahan Persentase Mobilitas Masyarakat di Taman Perubahan Persentase Mobilitas Masyarakat di Kantor Perubahan Persentase Mobilitas Masyarakat di Residential Area

 Omnibus:
 2.203 Durbin-Watson:
 1.079

 Prob(Omnibus):
 0.332 Jarque-Bera (JB):
 1.663

 Skew:
 0.419 Prob(JB):
 0.435

 Kurtosis:
 2.149 Condition No.:
 31

Coef.	Std.Err.	t	P> t	[0.025	0.975]
0.0000	0.0457	0.0000	1.0000	-0.0957	0.0957
0.1957	0.0813	2.4069	0.0264	0.0255	0.3659
0.2029	0.4625	0.4386	0.6659	-0.7651	1.1709
0.2331	0.3026	0.7703	0.4506	-0.4002	0.8664
-0.2116	0.1274	-1.6601	0.1133	-0.4783	0.0552
0.2790	0.2881	0.9685	0.3450	-0.3239	0.8819
-0.1359	0.1813	-0.7493	0.4629	-0.5154	0.2437
-0.0341	0.0750	-0.4551	0.6542	-0.1910	0.1228
0.6966	0.1727	4.0330	0.0007	0.3351	1.0581





Akan dilakukan feature selection menggunakan metode Backward Elimination untuk menyeleksi feature yang lebih berpengaruh terhadap model

3. Ketiga

Model:	OLS	Adj. R-squared:	0.946
Dependent Variable	e: Inflasi	AIC:	5.4066
Date:	2023-07-06 13:41	BIC:	16.0642
No. Observations:	28	Log-Likelihood:	5.2967
Df Model:	7	F-statistic:	68.38
Df Residuals:	20	Prob (F-statistic):	1.39e-12
R-squared:	0.960	Scale:	0.056149

Intercept

Kasus Positif Covid

Data Index Harga Konsumen

Perubahan Persentase Mobilitas di Rekreasi dan Retail Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek

Perubahan Persentase Mobilitas Masyarakat di Taman Perubahan Persentase Mobilitas Masyarakat di Kantor Perubahan Persentase Mobilitas Masyarakat di Residential Area

 Omnibus:
 2.098 Durbin-Watson:
 1.047

 Prob(Omnibus):
 0.350 Jarque-Bera (JB):
 1.547

 Skew:
 0.380 Prob(JB):
 0.461

 Kurtosis:
 2.135 Condition No.:
 13

Coef. Std.Err. t P>|t| [0.025] 0.975] 0.0000 0.0448 0.0000 1.0000 -0.0934 0.0934 0.1954 0.0796 2.4533 0.0235 -0.0293 0.3616 0.3524 0.1296 2.7205 0.0132 -0.822 0.6227 -0.2041 0.1237 -1.6498 0.1146 -0.4623 0.0540 0.3578 0.2206 1.6220 0.1205 -0.1023 0.8179 -0.0873 0.1407 -0.6206 0.5419 -0.3808 0.2062 -0.0362 0.0733 -0.4933 0.6272 -0.1891 0.1167 0.6534 0.1390 4.7008 0.0001 0.3634 0.9433





4. Keempat

Model:	OLS	Adj. R-squared:	0.948
Dependent Variable	e: Inflasi	AIC:	3.7452
Date:	2023-07-06 13:41	I BIC:	13.0706
No. Observations:	28	Log-Likelihood:	5.1274
Df Model:	6	F-statistic:	82.72
Df Residuals:	21	Prob (F-statistic)	: 1.63e-13
R-squared:	0.959	Scale:	0.054126

Intercept

Kasus Positif Covid

Data Index Harga Konsumen

Perubahan Persentase Mobilitas di Rekreasi dan Retail Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek

Perubahan Persentase Mobilitas Masyarakat di Taman Perubahan Persentase Mobilitas Masyarakat di Residential Area

 Omnibus:
 2.561 Durbin-Watson:
 1.127

 Prob(Omnibus):
 0.278 Jarque-Bera (JB):
 1.791

 Skew:
 0.423 Prob(JB):
 0.408

 Kurtosis:
 2.094 Condition No.:
 10

 Coef.
 Std.Err.
 t
 P>|t|
 [0.025]
 0.975]

 0.0000
 0.0440
 0.0000
 1.0000
 -0.0914
 0.0914

 0.2115
 0.0714
 2.9630
 0.0074
 0.0630
 0.3599

 0.3222
 0.1121
 2.8753
 0.0091
 0.0892
 0.5552

 -0.2152
 0.1195
 -1.8006
 0.0861
 -0.4637
 0.0333

 0.4211
 0.1761
 2.3904
 0.0263
 0.0548
 0.7874

 -0.1201
 0.1218
 -0.9861
 0.3353
 -0.3733
 0.1332

 0.6449
 0.1354
 4.7623
 0.0001
 0.3633
 0.9264





5. Kelima

Model:	OLS	Adj. R-squared:	0.948
Dependent Variable	: Inflasi	AIC:	3.0125
Date:	2023-07-06 13:41	BIC:	11.0058
No. Observations:	28	Log-Likelihood:	4.4937
Df Model:	5	F-statistic:	99.19
Df Residuals:	22	Prob (F-statistic):	2.47e-14
R-squared:	0.958	Scale:	0.054058

Intercept

Kasus Positif Covid

Data Index Harga Konsumen

Perubahan Persentase Mobilitas di Rekreasi dan Retail Perubahan Persentase Mobilitas di Toko Kebutuhan dan

Apotek

Perubahan Persentase Mobilitas Masyarakat di Residential Area

 Omnibus:
 2.323 Durbin-Watson:
 0.960

 Prob(Omnibus):
 0.313 Jarque-Bera (JB):
 1.660

 Skew:
 0.398 Prob(JB):
 0.436

 Kurtosis:
 2.112 Condition No.:
 9

 Coef.
 Std.Err.
 t
 P>|t|
 [0.025
 0.975]

 0.0000
 0.0439
 0.0000
 1.0000 -0.0911
 0.0911
 0.0911

 0.2127
 0.0713
 2.9833
 0.0069
 0.0649
 0.3606

 0.3258
 0.1119
 2.9107
 0.0081
 0.0937
 0.5579

 -0.2125
 0.1194
 -1.7796
 0.0890
 -0.4601
 0.0351

 0.3862
 0.1725
 2.2395
 0.0356
 0.0286
 0.7439

 0.5587
 0.1034
 5.4033
 0.0000
 0.3443
 0.7732





Akan dilakukan feature selection menggunakan metode **Backward Elimination untuk menyeleksi feature yang lebih** berpengaruh terhadap model

6. Ke-enam

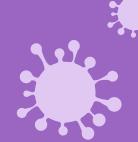
Model: OLS Adj. R-squared: 0.943 Dependent Variable: Inflasi AIC: 4.7781 Date: 2023-07-06 13:41 BIC: 11.4392 No Observations: Log-Likelihood: 2 6 1 0 9 112.6 Df Model: F-statistic: Df Residuals: Prob (F-statistic): 9.38e-15 R-squared: 0.951 Scale: 0.059151

> P>|t| [0.025 0.975] Intercept 0.0000 0.0460 | 0.0000 1.0000 -0.0951 0.0951 Kasus Positif Covid 0.2058 0.0745 2.7627 0.0111 0.0517 0.3599 0.3062 0.1165 2.6280 0.0150 0.0652 0.5472 Data Index Harga Konsumen

Coef. Std.Err.

Perubahan Persentase Mobilitas di Toko Kebutuhan dan Apotek 0.2656 0.1659 | 1.6012 0.1230 -0.0775 0.6088 Perubahan Persentase Mobilitas Masyarakat di Residential Area 0.4900 0.1003 4.8835 0.0001 0.2824 0.6975

Omnibus: 1 458 Durbin-Watson: 0.784 Prob(Omnibus): 0.482 Jarque-Bera (JB): 1.150 Skew: 0.276 Prob(JB): 0.563 2 175 Condition No : Kurtosis:



7. Ke-tujuh

Kurtosis:

Model:	OLS	Adj. R-squared:	0.939
Dependent Variable:	Inflasi	AIC:	5.7373
Date:	2023-07-06 13:41	BIC:	11.0661
No. Observations:	28	Log-Likelihood:	1.1313
Df Model:	3	F-statistic:	140.1
Df Residuals:	24	Prob (F-statistic):	2.42e-15
R-squared:	0.946	Scale:	0.063005

2.023 Condition No.:

Intercept		0.0000 0.0474	0.0000	1.0000 -0.0979	0.0979
Kasus Positif C	ovid	0.1127 0.0481	2.3441	0.0277 0.0135	0.2120
Data Index Harga Ko	onsumen	0.4435 0.0814	5.4500	0.0000 0.2756	0.6115
rsentase Mobilitas Mas	yarakat di Residential Area	0.5883 0.0819	7.1826	0.0000 0.4192	0.7573
2.627 Durbin-Watson:	0.554				
0.269 Jarque-Bera (JB):	1.580				
0.316 Prob(JB):	0.454				
	Kasus Positif C Data Index Harga Korsentase Mobilitas Mas 2.627 Durbin-Watson: 0.269 Jarque-Bera (JB):	Kasus Positif Covid Data Index Harga Konsumen rsentase Mobilitas Masyarakat di Residential Area 2.627 Durbin-Watson: 0.554 0.269 Jarque-Bera (JB): 1.580	Kasus Positif Covid 0.1127 0.0481 Data Index Harga Konsumen 0.4435 0.0814 rsentase Mobilitas Masyarakat di Residential Area 0.5883 0.0819 2.627 Durbin-Watson: 0.554 0.269 Jarque-Bera (JB): 1.580	Kasus Positif Covid 0.1127 0.0481 2.3441 Data Index Harga Konsumen 0.4435 0.0814 5.4500 rsentase Mobilitas Masyarakat di Residential Area 0.5883 0.0819 7.1826 2.627 Durbin-Watson: 0.554 0.269 Jarque-Bera (JB): 1.580	Kasus Positif Covid 0.1127 0.0481 2.3441 0.0277 0.0135 Data Index Harga Konsumen 0.4435 0.0814 5.4500 0.0000 0.2756 rsentase Mobilitas Masyarakat di Residential Area 0.5883 0.0819 7.1826 0.0000 0.4192 2.627 Durbin-Watson: 0.554 0.269 Jarque-Bera (JB): 1.580

• R square score 0.946 berarti 94.6% dari dependent variable dapat dijelaskan dengan independent variable yang kami pilih.

Coef. Std.Err.

P>Itl [0.025 0.975]

- Karena f_pvalue kami (Prob (F-statistic): 2.42e-15) lebih rendah dari 0.05, maka dapat disimpulkan bahwa model kami lebih baik daripada model lain yang lebih sederhana.
- semua independent variable kami, 'Kasus positif covid', 'data indeks harga konsumen', 'persentase mobilitas masyarakat di residential area', memiliki p-value kurang dari 0.05 yang mana menunjukkan bahwa terdapat bukti yang cukup untuk 'Kasus positif covid', 'data indeks harga konsumen', 'persentase mobilitas masyarakat residential area' affects 'Inflasi'.



MAE: 0.2514643490426728 MSE: 0.08757232148245976 RMSE: 0.29592620952267773

R Squared scores: 0.9633581278774068

Inflasi Prediksi

0	-1.644293	-1.297454
1	1.092564	1.020674
2	-0.284108	-0.275465
3	0.705118	0.236725
4	-0.828182	-1.150478
5	-0.374787	0.020928
6	1.183243	1.318087
7	-0.383030	-0.364060
8	0.960668	0.804627
9	1.150269	1.520248
10	0.490785	0.086178
11	-1.701998	-1.382644

	Inflasi	Prediksi	Residual
0	-1.644293	-1.297454	-0.323666
1	1.092564	1.020674	-0.303121
2	-0.284108	-0.275465	-0.339309
3	0.705118	0.236725	-0.160153
4	-0.828182	-1.150478	-0.068438

Dapat dilihat juga, Mae, MSE, RMSE, dan skor cross-validation MAE, setelah dilakukan BE, menurun, sedangkan r-square dan skor CV r-square, setelah dilakukan BE, naik, maka model membaik setelah dilakukan feature selection BE

```
MAE scores :
```

[0.17789709 0.19990428 0.18713513 0.33053271 0.33699331]

Average MAE Score (Across experminets):

0.246492504787378

R squared scores:

[0.95812258 0.87518479 0.8998104 0.92659453 0.72917434]

Average R squared score (Across experiments): 0.8777773261919519





Linearity Assumption

```
sns.lmplot(x='Inflasi', y='Prediksi', data=hasil, fit reg=False)
line coords = np.arange(hasil[['Inflasi', 'Prediksi']].min().min(),
                        hasil[['Inflasi', 'Prediksi']].max().max())
plt.plot(line coords, line coords,
         color='darkorange', linestyle='--')
plt.ylabel('Prediksi Inflasi', fontsize=14)
plt.xlabel('Inflasi', fontsize=14)
plt.title('Linearity Assumption', fontsize=16)
plt.show()
                     Linearity Assumption
     1.5
     1.0
     0.5
Prediksi Inflasi
     -0.5
    -1.0
    -1.5
             -1.5
                     -1.0
                             -0.5
                                     0.0
                                              0.5
                                                      1.0
```

Inflasi



Dari code ini dapat dilihat bahwa titik titik residual berada disekitar garis diagonal dan terlihat linear dan searah dengan garis. Hal ini menunjukkan bahwa terdapat hubungan linear antara variable independen dengan variable dependen kami.

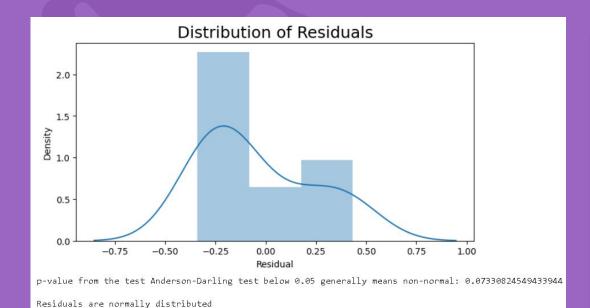
Asumsi Linearitas terpenuhi







Distributions of Residuals



Dari code ini dapat dilihat bahwa kami mendapatkan p-value = 0.073308 yang berarti residual berdistribusi normal (walaupun terlihat miring ke kiri, hal ini dikarenakan sampel yang cukup kecil yaitu 28).

Asumsi Normalitas terpenuhi







Multicollinearity Assumption

```
from statsmodels.stats.outliers influence import variance inflation factor
# the independent variables set
X = data standard[['Kasus Positif Covid', 'Data Index Harga Konsumen', 'Perubahah Persentase Mobilitas Masyarakat di Residential Area']]
# VIE dataframe
vif data = pd.DataFrame()
vif_data["feature"] = X.columns
# calculating VIF for each feature
vif_data["VIF"] = [variance_inflation_factor(X.values, i)
                          for i in range(len(X.columns))]
print(vif_data)
                                             feature
                                                           VIE
                                 Kasus Positif Covid 1.027842
                          Data Index Harga Konsumen 2.943444
  Perubahan Persentase Mobilitas Masyarakat di R... 2.981062
```

Dapat dilihat bahwa VIF semua prediktor <5. Maka terbukti bahwa model tidak memiliki multikolonieritas.

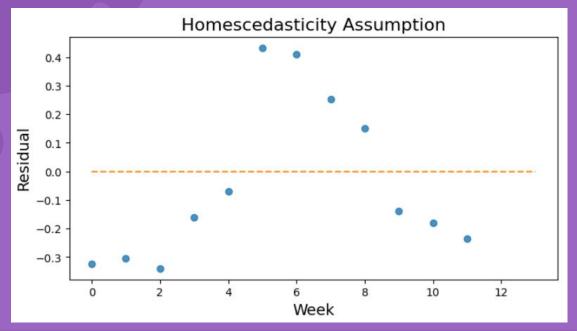
Asumsi Multikolinearitas terpenuhi





Homescedasticity Assumption





Sebaran residualnya tidak membentuk pola tertentu (menyebar secara merata).

Asumsi Homoskedastisitas terpenuhi

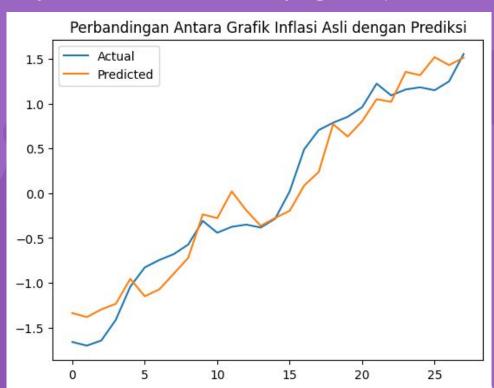






Berdasarkan beberapa analisis residual tersebut...

Dapat dibuktikan bahwa model yang kami ajukan memenuhi syarat persamaan regresi linear



Dengan grafik dapat dilihat bahwa telah terdapat kemiripan antara data inflasi aktual dengan data inflasi yang telah kami prediksi menggunakan model regresi linear kami

Conclusions

Model kami memenuhi semua pengujian dalam langkah-langkah validasi model, sehingga kami dapat menyimpulkan bahwa model kami dapat bekerja dengan baik untuk memprediksi Inflasi di masa mendatang dengan menggunakan tiga variabel independen 'kasus positif Covid', 'data indeks harga konsumen', 'persentase orang mobilitas di residential area'. Model kami menunjukkan skor r-kuadrat yang baik sebesar 0,946 berarti 94,6% dari variabel dependen kami dapat dijelaskan menggunakan variabel independen kami.





Recommendations

Ada baiknya jika data yang digunakan adalah data yang paling update dengan kondisi terkini. Hal ini juga meningkatkan kualitas model jika dihubungkan dengan data yang paling terbaru. Model yang menggunakan data paling terbaru dapat menghasilkan prediksi yang sesuai dengan perubahan-perubahan yang terjadi pada kondisi aktual.

