

# The Relationship between Prosperity Levels and Inflation, Investment, and Gender Equality in the World Before the COVID-19 Pandemic

## A. Pengantar

Dalam riset ini akan membahas mengenai hubungan antara tingkat kemakmuran suatu negara yang diwakili dengan nilai Gross Domestic Product Percapita (GDPP) dalam bentuk USD dengan tingkat inflasi, Foreign Direct Investment (FDI) dan Kesetaraan Gender sebelum pandemi COVID-19.

GDP Percapita adalah salah satu indikator ekonomi yang mengukur rata-rata nilai dari seluruh barang dan jasa yang dihasilkan oleh suatu negara, GDP Percapita digunakan sebagai ukuran standar untuk mengukur tingkat kemakmuran ekonomi suatu negara dan membandingkan kesejahteraan penduduknya dengan negara lain. Semakin tinggi GDP Percapita suatu negara, semakin tinggi juga pendapatan rata-rata per individu di negara tersebut, dan ini biasanya dikaitkan dengan tingkat hidup yang lebih baik.

Foreign Direct Investment (FDI) atau Investasi Langsung Asing adalah suatu bentuk investasi yang dilakukan oleh individu, perusahaan, atau entitas ekonomi asing untuk mendirikan atau mengakuisisi bisnis atau aset produktif di negara lain.

Tingkat inflasi mengacu pada tingkat kenaikan harga-harga secara umum dalam suatu ekonomi dalam periode waktu tertentu. Dalam kata lain, inflasi adalah laju atau persentase perubahan harga-harga barang dan jasa secara keseluruhan dalam suatu negara atau wilayah tertentu selama jangka waktu tertentu.

Gender inequality adalah ketidaksetaraan atau ketidakadilan yang terjadi antara laki-laki dan perempuan dalam berbagai aspek kehidupan, termasuk akses terhadap kesempatan, sumber daya, dan hak-hak sosial, ekonomi, dan politik. Ketidaksetaraan gender dapat muncul dalam berbagai bentuk, termasuk pendapatan, kesempatan pendidikan, partisipasi politik, dan akses terhadap layanan kesehatan.

## B. Tujuan

Riset ini akan mencoba menguji Proporsi Dua Sampel, Uji proporsi dua sampel adalah metode statistik yang digunakan untuk membandingkan dua proporsi dari dua kelompok yang berbeda yakni sampel yang memiliki label `gdpp_level growth` dan `gdpp_level high`.

Riset ini turut menggunakan metode statistik linear regression dengan multiple predictor untuk mencari seberapa besar hubungan antara variabel `y` Gross Domestic Product Percapita (GDPP) dengan variabel `x` (tingkat inflasi, Foreign Direct Investment (FDI), Kesetaraan Gender).

## C. Data dan Sumber Data

Data yang digunakan merupakan data yang dikumpulkan melalui sumber data yakni Bank Dunia (World Bank), UNDP, dan macrotrends pada tahun 2019. Sample negara yang digunakan yaitu 118 negara (baris) dan 5 kolom :

1. `country_name`, merupakan negara-negara yang akan diobservasi
2. `gdpp`, GDP Percapita adalah salah satu indikator ekonomi yang mengukur rata-rata nilai dari seluruh barang dan jasa yang dihasilkan oleh suatu negara dalam USD.

3. fdi, Investasi Langsung Asing adalah suatu bentuk investasi yang dilakukan oleh individu, perusahaan, atau entitas ekonomi asing untuk mendirikan atau mengakuisisi bisnis atau aset produktif di negara lain dalam Miliar USD.
4. inflation, inflasi atau tingkat kenaikan harga-harga secara umum dalam suatu ekonomi dalam periode waktu tertentu dalam bentuk persentase yang dikonversi menjadi desimal
5. gender\_inequality, ketidaksetaraan atau ketidakadilan yang terjadi antara laki-laki dan perempuan dalam berbagai aspek kehidupan dengan Skor 0-1.

#### D. Proses Eksplorasi dan Uji Statistik

Kita ingin melihat prediksi gdp dengan menggunakan tingkat inflasi, Foreign Direct Investment (FDI), dan Kesetaraan Gender kita dapat menggunakan Linear Regression dengan Multiple Predictor.

##### 1. Data Preparation

- a. Data Collection
- b. Mengubah tipe data yang tidak sesuai
- c. Standarisasi nilai dari fdi menjadi dalam miliar rupiah
- d. Hanya mengambil nilai fdi yang positif atau diatas nol saja
- e. Hanya mengambil nilai inflasi diatas 0 dan dibawah 15% atau 0.15% saja

##### 2. Exploratory Data Analysis

- a. melihat 10 Teratas dan terbawah negara pada masing masing variabel.
- b. Melakukan labeling pada data jika gdp > 20.000 maka akan dilabeli high dan dibawah itu dilabeli growth. Hal ini sesuai dengan laporan bank dunia pada tahun 2021.
- c. Melihat Proporsi data setelah dilabeli
- d. Melihat Distribusi Data
- e. Cek Korelasi Data
- f. Melihat Outlier dari Data

##### 3. Statistical Inference

- a. Uji Proporsi Dua Sampel, Uji proporsi dua sampel adalah metode statistik yang digunakan untuk membandingkan dua proporsi dari dua kelompok yang berbeda yakni sampel yang memiliki label gdp\_level growth dan gdp\_level high. Untuk menghitung dua sampel uji z proporsi, kita dapat menggunakan stats.ttest\_ind.

##### 4. Linear Regression

- a. Cek Linearitas Data dengan menggunakan Scatterplot
- b. Transformasi Log
- c. Linear Regression dengan menggunakan multiple predictor yakni log\_fdi, log\_inflation, log\_gender\_inequality dengan outcome log\_gdp beserta interpretasi koefisien, rsquare, prob f-stats, prob t-stats, dan Prob(JB).

- d. Menguji model dengan uji asumsi seperti normalitas residual, homokedastisitas, linearitas, linearitas variabel prediktor.

## E. Statistical Inference

### 1. FDI

Null hypothesis ( $H_0$ ), alternative hypothesis ( $H_1$ ), and significance level

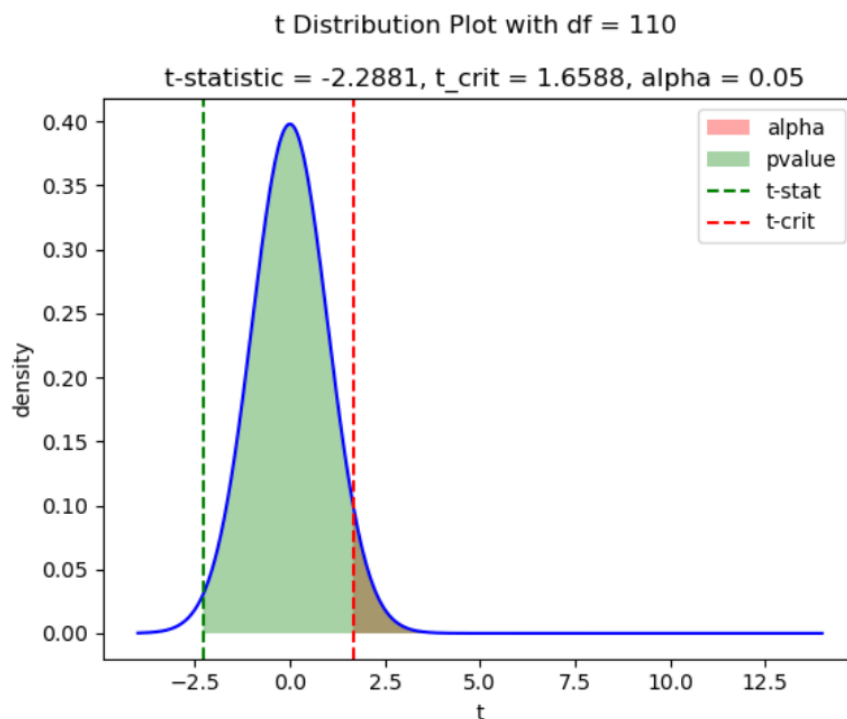
$H_0$  : Nilai FDI dinegara yang memiliki level gdpp growth dan high sama.

$$H_0: \mu A = \mu B$$

$H_1$  : Nilai FDI dinegara yang memiliki level gdpp growth dan high tidak sama.

$$H_1: \mu A \geq \mu B$$

Significance level = 0.05



Conclusion :

Nilai  $p > \alpha$  (0.05), maka gagal tolak  $H_0$ . Dengan demikian, dapat disimpulkan bahwa jumlah Nilai FDI dinegara yang memiliki level gdpp growth dan high sama.

Karena interval kepercayaan terletak antara -41.736 dan -10.502, tidak terdapat perbedaan antara FDI dinegara yang memiliki level gdpp growth dan high.

## 2. Inflation

Null hypothesis ( $H_0$ ), alternative hypothesis ( $H_1$ ), and significance level

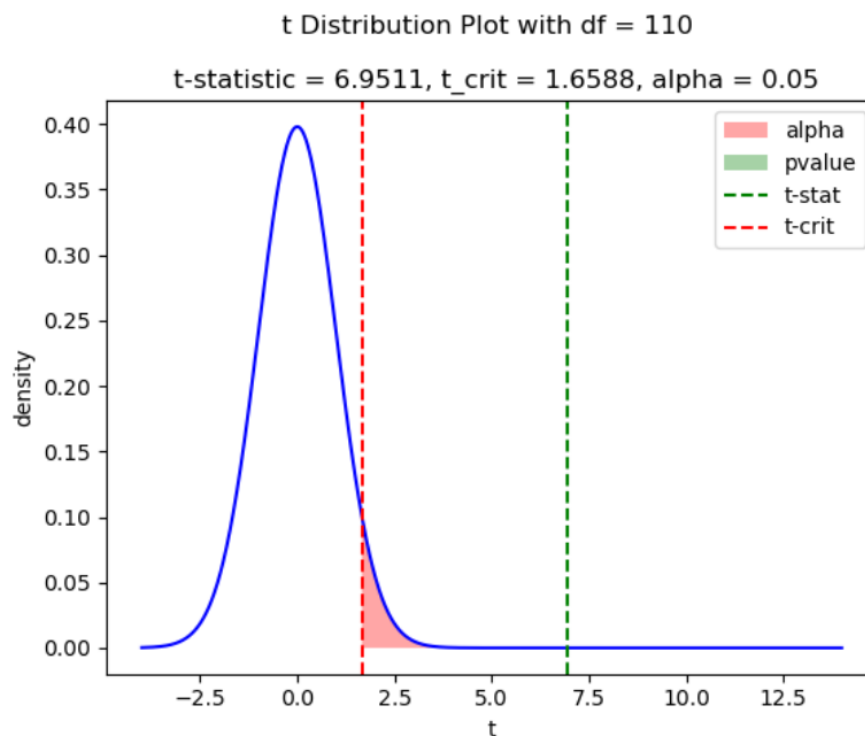
$H_0$  : Inflasi dinegara yang memiliki level gdpp growth dan high sama.

$$H_0: \mu A = \mu B$$

$H_1$ : Inflasi dinegara yang memiliki level gdpp growth dan high tidak sama.

$$H_1: \mu A \geq \mu B$$

Significance level = 0.05



### Conclusion

Nilai  $p < \alpha$  (0.05), maka tolak  $H_0$ . Dengan demikian, dapat disimpulkan bahwa inflasi dinegara yang memiliki level gdpp growth dan high tidak sama.

Karena interval kepercayaan terletak antara 0.0157 dan 0.028, terdapat perbedaan antara inflasi dinegara yang memiliki level gdpp growth dan high. Dapat dikatakan bahwa inflasi dinegara yang memiliki level gdpp growth lebih besar daripada gdpp high. Dengan demikian, jika negara yang memiliki gdpp level growth ingin bertransformasi menjadi gdpp level high perlu menjaga tingkat inflasi agar stabil.

## Gender Inequality

Null hypothesis ( $H_0$ ), alternative hypothesis ( $H_1$ ), and significance level

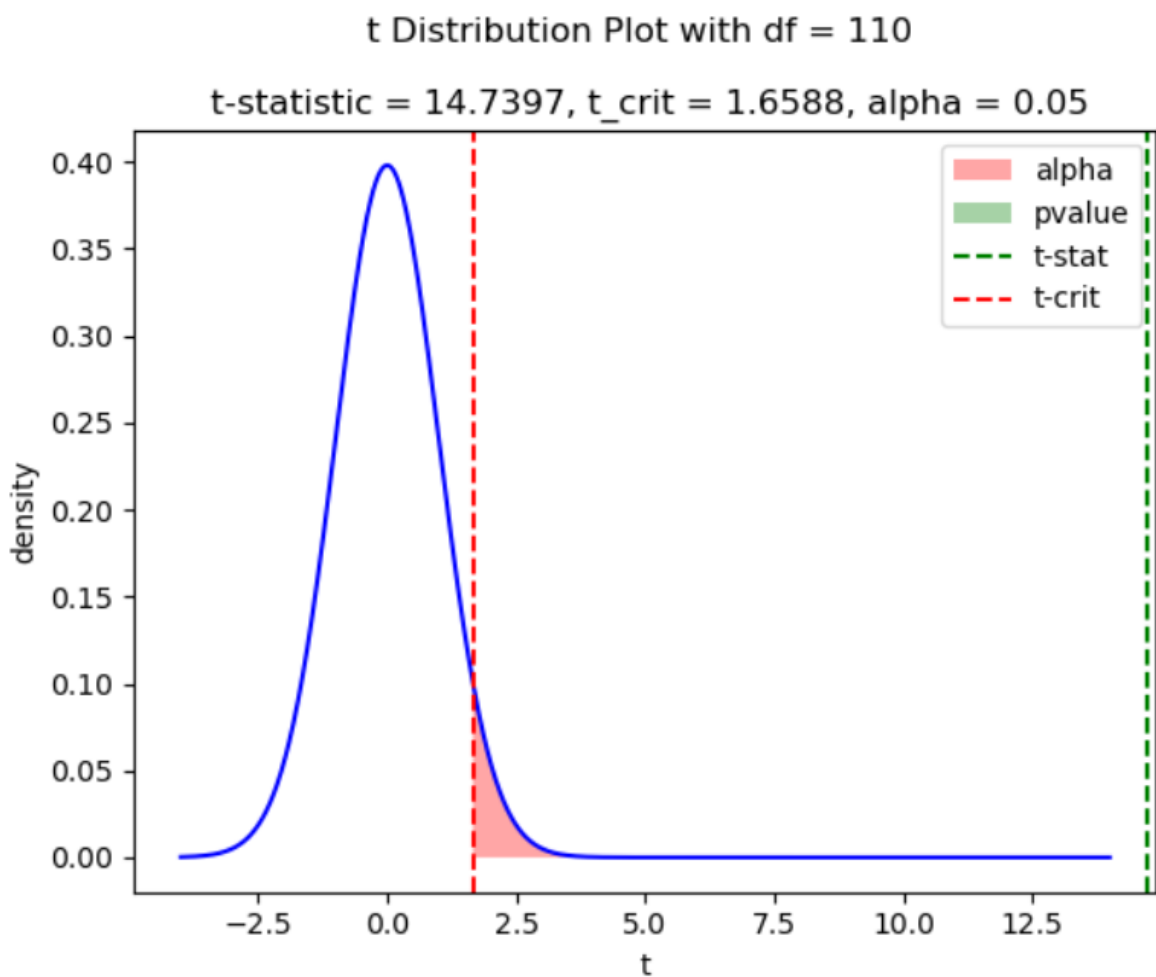
$H_0$ : Kesetaraan Gender dinegara yang memiliki level gdpp growth dan high sama.

$$H_0: \mu A = \mu B$$

$H_1$  : Kesetaraan Gender dinegara yang memiliki level gdpp growth dan high tidak sama.

$$H_1: \mu A \geq \mu B$$

Significance level = 0.05



## Conclusion

Nilai  $p < \alpha$  (0.05), maka tolak  $H_0$ . Dengan demikian, dapat disimpulkan bahwa tingkat kesetaraan gender dinegara yang memiliki level gdpp growth dan high tidak sama.

Karena interval kepercayaan terletak antara 0.243 dan 0.319, terdapat perbedaan antara tingkat kesetaraan gender dinegara yang memiliki level gdpp growth dan high. Dapat dikatakan bahwa tingkat kesetaraan gender atau ketidaksetaraan dan ketidakadilan yang terjadi antara laki-laki dan perempuan dinegara yang memiliki level gdpp growth lebih besar daripada gdpp high. Dengan demikian, jika negara yang memiliki gdpp level growth ingin bertransformasi menjadi gdpp level high perlu menyetarakan gender dinegaranya.

Linear Regression

Transformation With Log

	coef	std err
<b>Intercept</b>	6.579698	0.255887
<b>log_fdi</b>	0.161948	0.036015
<b>log_inflation</b>	-0.187034	0.070810
<b>log_gender_inequality</b>	-0.996690	0.091710

$$\log\_gdpp = 6.58 + 0.16 \times \log\_fdi - 0.19 \times \log\_inflation - 0.99 \times \log\_gender\_inequality + e$$

Conclusion

Koefisien 6.58: Ini adalah koefisien intercept atau nilai konstanta dalam model. Ini menunjukkan nilai perkiraan  $\log\_gdpp$  ketika semua variabel independen ( $\log\_fdi$ ,  $\log\_inflation$ , dan  $\log\_gender\_inequality$ ) memiliki nilai nol.

Koefisien 0.16 untuk  $\log\_fdi$ : Ini menunjukkan hubungan linier antara  $\log\_fdi$  dengan  $\log\_gdpp$ . Dalam konteks ini, peningkatan satu unit atau satu persen dalam  $\log\_fdi$  dalam bentuk miliar (dengan mempertahankan variabel lainnya tetap) dihubungkan dengan peningkatan sebesar 0.16 dalam  $\log\_gdpp$ .

Koefisien -0.19 untuk  $\log\_inflation$ : Ini menunjukkan hubungan linier antara  $\log\_inflation$  dengan  $\log\_gdpp$ . Dalam konteks ini, peningkatan satu unit atau satu persen dalam  $\log\_inflation$  atau harga secara keseluruhan makin naik (dengan mempertahankan variabel lainnya tetap) dihubungkan dengan penurunan sebesar 0.19 dalam  $\log\_gdpp$ .

Koefisien -0.99 untuk  $\log\_gender\_inequality$ : Ini menunjukkan hubungan linier antara  $\log\_gender\_inequality$  dengan  $\log\_gdpp$ . Dalam konteks ini, peningkatan satu unit atau satu

persen dalam log\_gender\_inequality atau perempuan dan laki laki makin tidak setara (dengan mempertahankan variabel lainnya tetap) dihubungkan dengan penurunan sebesar 0.99 dalam log\_gdpp.

```
results.rsquared
0.748650272703284
```

#### Conclusion

0.748650272703284, atau sekitar 74.87%, dari variasi dalam variabel dependen "log\_gdpp" dapat dijelaskan oleh variabel independen (log\_fdi, log\_inflation, dan log\_gender\_inequality) yang ada dalam model regresi linier. Dengan kata lain, model tersebut berhasil menjelaskan sekitar 74.87% dari variasi dalam log\_gdpp.

```

=====
                        OLS Regression Results
=====
Dep. Variable:          log_gdpp      R-squared:                0.749
Model:                  OLS           Adj. R-squared:           0.742
Method:                 Least Squares   F-statistic:              107.2
Date:                  Thu, 06 Jul 2023   Prob (F-statistic):       2.99e-32
Time:                  22:06:46          Log-Likelihood:           -110.80
No. Observations:      112              AIC:                     229.6
Df Residuals:          108              BIC:                     240.5
Df Model:               3
Covariance Type:       nonrobust
=====
                        coef      std err          t      P>|t|      [0.025      0.975]
-----
Intercept              6.5797      0.256       25.713     0.000        6.072        7.087
log_fdi                0.1619      0.036        4.497     0.000        0.091        0.233
log_inflation          -0.1870      0.071       -2.641     0.009       -0.327       -0.047
log_gender_inequality  -0.9967      0.092     -10.868     0.000       -1.178       -0.815
=====
Omnibus:               0.616      Durbin-Watson:           1.925
Prob(Omnibus):         0.735      Jarque-Bera (JB):        0.538
Skew:                  0.168      Prob(JB):                0.764
Kurtosis:              2.948      Cond. No.                 19.3
=====

```

#### Conclusion

Adj. R-squared : 0.742, atau sekitar 74.2%, dari variasi dalam variabel dependen "log\_gdpp" dapat dijelaskan oleh kombinasi variabel independen (log\_fdi, log\_inflation, dan log\_gender\_inequality) dalam model regresi linier.

Prob F-Statistics

$H_0$  : FDI, Inflasi dan Kesenjangan Gender secara bersamaan tidak memiliki pengaruh signifikan terhadap  $gdpp$ .

$H_1$ : FDI, Inflasi dan Kesenjangan Gender secara bersamaan memiliki pengaruh signifikan terhadap  $gdpp$ .

tingkat signifikansi 0.05

Prob (F-statistic) : memiliki value 2.99e-32 Dengan nilai p-value yang sangat kecil seperti ini (jauh di bawah tingkat signifikansi yang digunakan seperti 0.05), kita dapat menyimpulkan bahwa terdapat cukup bukti statistik yang kuat untuk menolak hipotesis nol. Dalam konteks ini, itu berarti FDI, Inflasi dan Kesenjangan Gender secara bersamaan memiliki pengaruh atau kontribusi signifikan terhadap  $gdpp$  dalam regresi linear.

Prob T Stats

FDI

$\sim H_0$  : FDI tidak memiliki pengaruh signifikan terhadap  $gdpp$ .

$\sim H_1$  : FDI memiliki pengaruh signifikan terhadap  $gdpp$ .

Inflasi

$\sim H_0$ : Inflasi tidak memiliki pengaruh signifikan terhadap  $gdpp$ .

$\sim H_1$ : Inflasi memiliki pengaruh signifikan terhadap  $gdpp$ .

Kesenjangan Gender

$\sim H_0$ : Kesenjangan Gender tidak memiliki pengaruh signifikan terhadap  $gdpp$ .

$\sim H_1$ : Kesenjangan Gender memiliki pengaruh signifikan terhadap  $gdpp$ .

tingkat signifikansi : 0.05

Kesimpulan :

$\log\_fdi$ : P-value sebesar 0.000 menunjukkan bahwa variabel  $\log\_fdi$  memiliki hubungan yang signifikan secara statistik dengan variabel dependen ( $\log\_gdpp$ ) dalam model regresi linier. Karena nilai p-value sangat kecil (lebih kecil dari tingkat signifikansi 0.05), kita dapat menolak hipotesis nol dan menyimpulkan bahwa  $\log\_fdi$  berkontribusi secara signifikan terhadap perubahan dalam  $\log\_gdpp$ .



log\_inflation: P-value sebesar 0.009 menunjukkan bahwa variabel log\_inflation juga memiliki hubungan yang signifikan secara statistik dengan variabel dependen (log\_gdpp) dalam model regresi linier. Dalam hal ini, kita juga dapat menolak hipotesis nol dan menyimpulkan bahwa log\_inflation berkontribusi secara signifikan terhadap perubahan dalam log\_gdpp.

log\_gender\_inequality: P-value sebesar 0.000 menunjukkan bahwa variabel log\_gender\_inequality memiliki hubungan yang signifikan secara statistik dengan variabel dependen (log\_gdpp) dalam model regresi linier. Seperti pada kasus sebelumnya, kita dapat menolak hipotesis nol dan menyimpulkan bahwa log\_gender\_inequality berkontribusi secara signifikan terhadap perubahan dalam log\_gdpp.

Prob(JB)

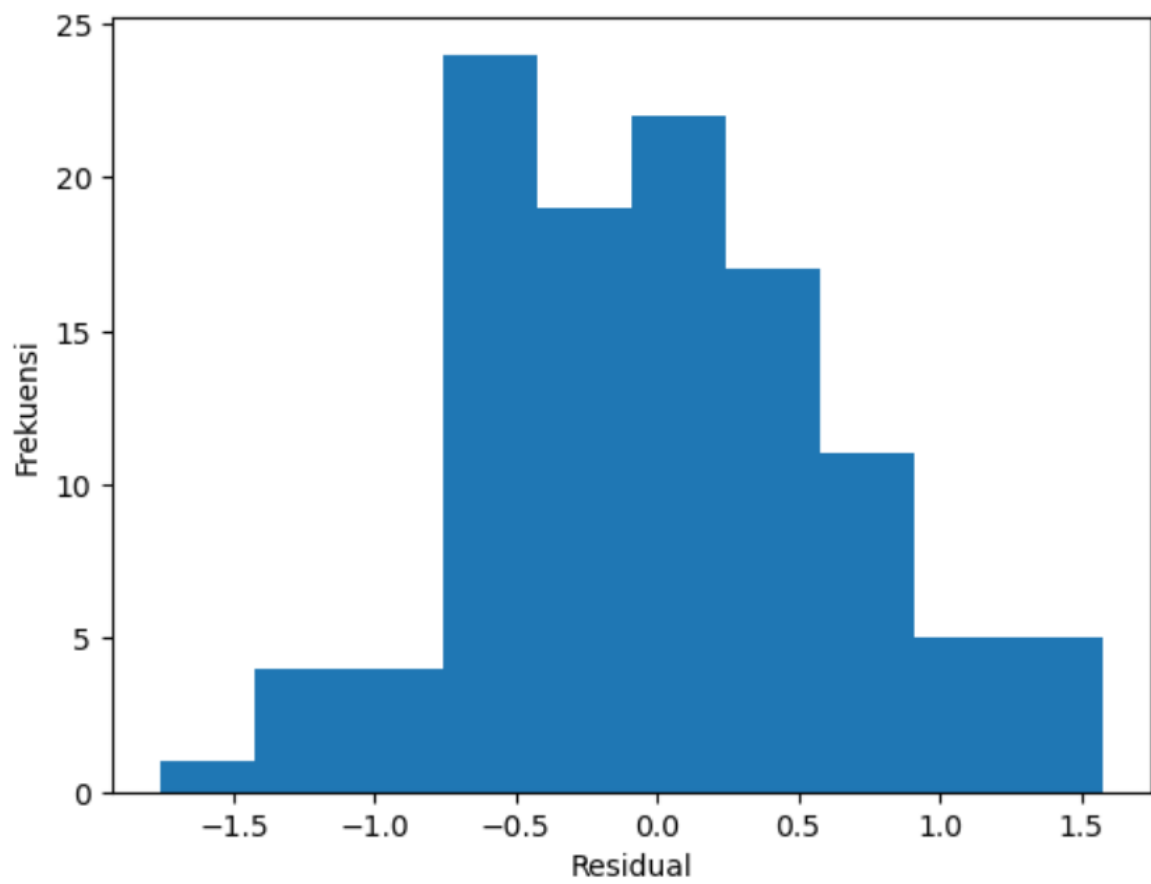
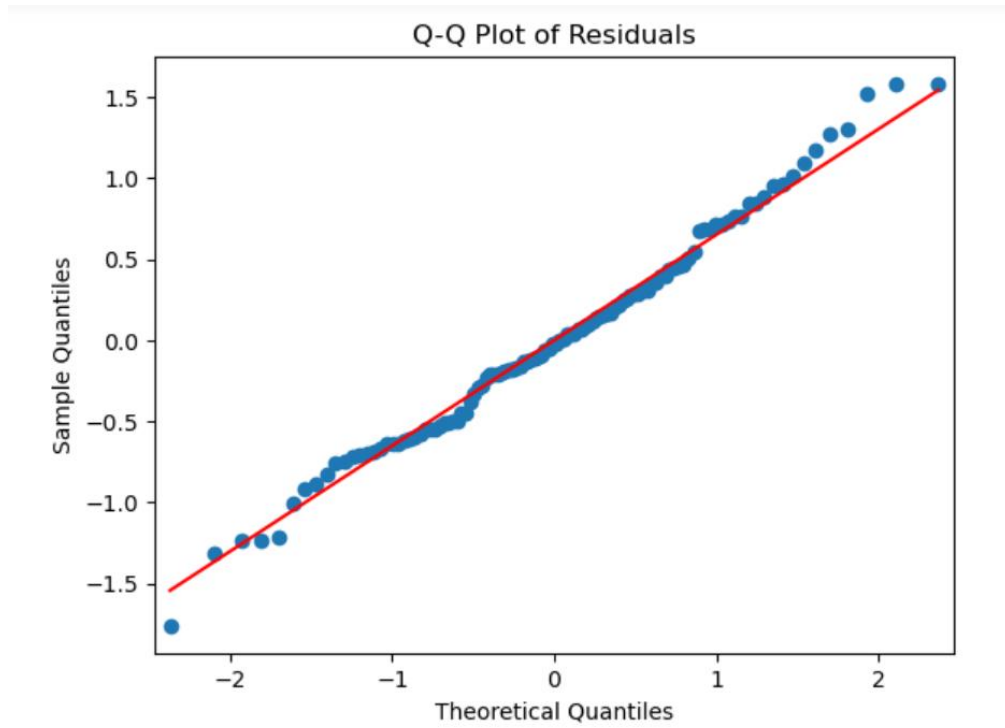
$\sim H_0$  : residual dalam model regresi linier tidak mengikuti distribusi normal.

$\sim H_1$  : residual dalam model regresi linier mengikuti distribusi normal.

Dengan nilai Prob(JB) sebesar 0.764, probabilitas tersebut cukup tinggi (lebih besar dari tingkat signifikansi yang umumnya digunakan seperti 0.05). Ini menunjukkan bahwa tidak ada cukup bukti statistik yang signifikan untuk menolak hipotesis nol. Dalam konteks ini, hipotesis nol menyatakan bahwa residual dalam model regresi linier mengikuti distribusi normal.

Uji Asumsi

Normalitas Residual



Shapiro-Wilk Test:

Statistic = 0.9906477332115173, p-value = 0.6425069570541382

Conclusion

Nilai p (p-value): Nilai p yang dihasilkan dari uji Shapiro-Wilk. Nilai p ini menunjukkan probabilitas bahwa distribusi data Anda sebenarnya adalah distribusi normal, jika nilai p cukup tinggi (lebih besar dari tingkat signifikansi yang telah ditentukan, biasanya 0.05 atau 0.01). Pada kasus Anda, nilai p-nya adalah 0.6425069570541382.

Homokedastisitas

Breusch-Pagan Test (Homoskedasticity):

p-value = 0.14117514171106357

Conclusion

tingkat signifikansi adalah 0.05, karena nilai p-value (0.14117514171106357) lebih besar dari 0.05, kita tidak memiliki cukup bukti statistik yang signifikan untuk menolak hipotesis nol. Dalam konteks ini, hipotesis nol menyatakan bahwa tidak ada heteroskedastisitas (variansi residual konstan) dalam model regresi linier.

Dengan demikian, berdasarkan nilai p-value, tidak ada cukup bukti statistik yang signifikan untuk menolak asumsi homoskedastisitas dalam model regresi linier. Ini menunjukkan bahwa varian residual dalam model regresi linier dapat dianggap konstan di semua tingkat nilai prediktor (variabel independen).

Linearitas

Rainbow Test (Linearity):

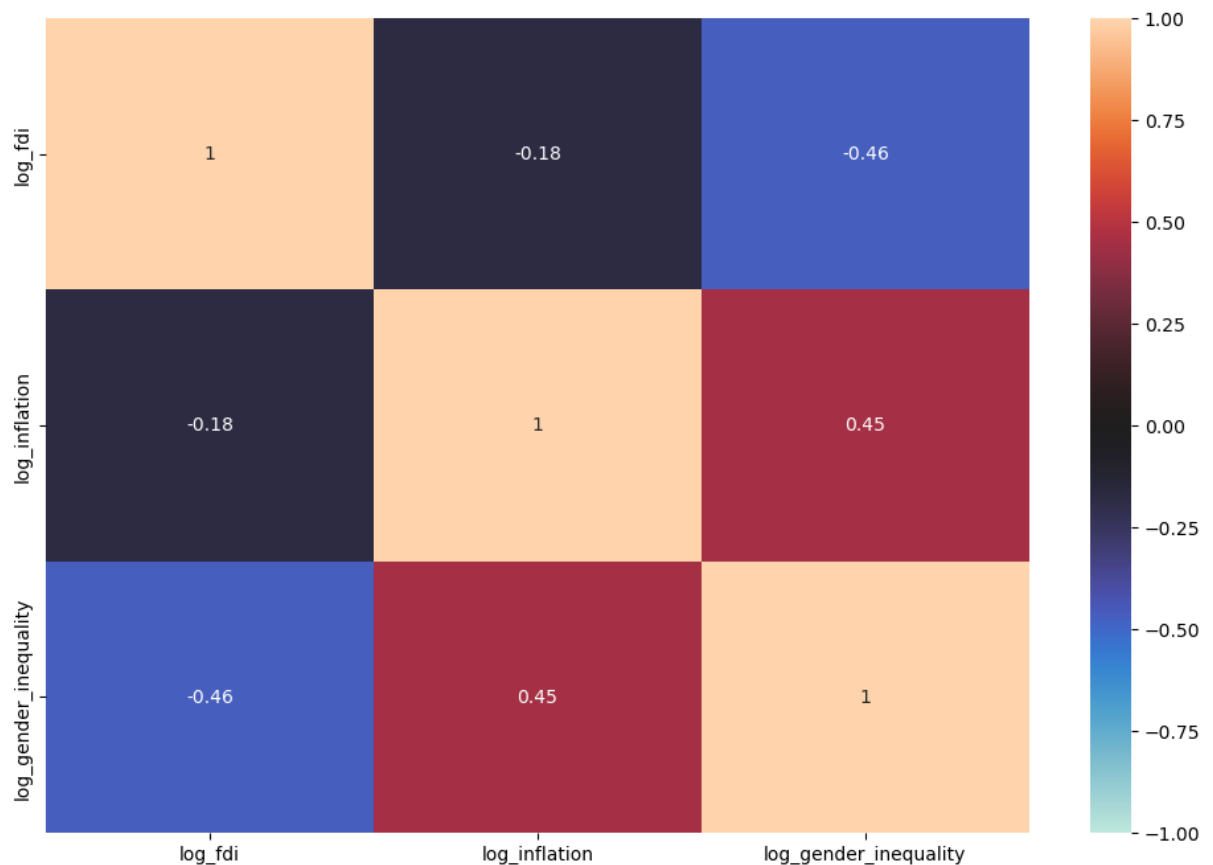
p-value = 0.35643032951327275

Conclusion

tingkat signifikansi adalah 0.05, karena nilai p-value (0.35643032951327275) lebih besar dari 0.05, kita tidak memiliki cukup bukti statistik yang signifikan untuk menolak hipotesis nol. Dalam konteks ini, hipotesis nol menyatakan bahwa hubungan antara variabel independen dan variabel dependen dalam model regresi linier bersifat linier.

Dengan demikian, berdasarkan nilai p-value, tidak ada cukup bukti statistik yang signifikan untuk menolak asumsi linearitas dalam model regresi linier. Ini menunjukkan bahwa hubungan antara variabel independen dan variabel dependen dapat dianggap bersifat linier dalam model tersebut.

#### Linearitas Antar Variabel Independen



#### Conclusion

Jika kita lihat dari value korelasi, tidak ada yang memiliki nilai lebih dari 0.8, maka dapat disimpulkan bahwa tidak terdapat korelasi antar variabel independen atau x.

#### Kesimpulan Keseluruhan.

Data yang digunakan pada pemodelan ini ialah data inflasi, fdi, dan gender inequality. Data bersumber dari Macrotrends, Worldbank dan UNDP.

Data yang digunakan yakni dengan nilai inflasi yang diatas 0 dan dibawah 0.05 dan fdi diatas 0.

saat uji statistika inference, dengan uji Proporsi Dua Sampel menyatakan jumlah FDI dinegara yang memiliki level gdpp growth dan high sama, inflasi dinegara yang memiliki level gdpp growth dan high tidak sama dan tingkat kesetaraan gender dinegara yang memiliki level gdpp growth dan high tidak sama.

saat melakukan linear regression menyatakan Koefisien senilai 6.58 Ini adalah koefisien intercept atau nilai konstanta dalam model. Ini menunjukkan nilai perkiraan log\_gdpp ketika semua variabel independen (log\_fdi, log\_inflation, dan log\_gender\_inequality) memiliki nilai nol. log\_fdi bernilai 0.16 menunjukkan hubungan linier antara log\_fdi dengan log\_gdpp. Dalam konteks ini, peningkatan satu unit atau satu persen dalam log\_fdi dalam bentuk miliar (dengan mempertahankan variabel lainnya tetap) dihubungkan dengan peningkatan sebesar 0.16 dalam log\_gdpp. log\_inflation bernilai -0.19 menunjukkan hubungan linier antara log\_inflation dengan log\_gdpp. Dalam konteks ini, peningkatan satu unit atau satu persen dalam log\_inflation atau harga secara keseluruhan makin naik (dengan mempertahankan variabel lainnya tetap) dihubungkan dengan penurunan sebesar 0.19 dalam log\_gdpp. log\_gender\_inequality bernilai -0.99 menunjukkan hubungan linier antara log\_gender\_inequality dengan log\_gdpp. Dalam konteks ini, peningkatan satu unit atau satu persen dalam log\_gender\_inequality atau perempuan dan laki laki makin tidak setara (dengan mempertahankan variabel lainnya tetap) dihubungkan dengan penurunan sebesar 0.99 dalam log\_gdpp. dan secara uji prob t statistik masing masing variabel independen berpengaruh terhadap variabel dependen baik secara parsial maupun simultan. sekitar 74.87%, dari variasi dalam variabel dependen "log\_gdpp" dapat dijelaskan oleh variabel independen (log\_fdi, log\_inflation, dan log\_gender\_inequality) yang ada dalam model regresi linier. Dengan kata lain, model tersebut berhasil menjelaskan sekitar 74.87% dari variasi dalam log\_gdpp. dan pada uji asumsi seperti normalitas residual, homokedastisitas, linearitas, linearitas variabel prediktor menyatakan hasil yang cukup baik untuk kebaikan model.

#### Saran Kebijakan

Jika negara ingin melakukan transformasi dari yang sebelumnya gdpp growth menjadi negara yang memiliki gdpp high maka perlu menurunkan nilai inflasi dan menurunkan tingkat kesetaraan gender (read : semakin setara) berdasarkan uji proporsi dua sampel yang dilakukan diatas

Jika negara ingin melakukan transformasi dari yang sebelumnya gdpp growth menjadi negara yang memiliki gdpp high berdasarkan linear regression maka perlu meningkatkan nilai fdi, menurunkan nilai inflasi dan menurunkan tingkat kesetaraan gender (read : semakin setara), karena berdasarkan uji statistik terbukti secara signifikan memengaruhi nilai dari gdpp.

#### Further Research

Jika memiliki waktu lebih banyak, lakukan tinjauan teoritis dari masing masing variabel berdasarkan penelitian terdahulu.

Dapat pula menggunakan metode regresi logistik berdasarkan negara yang memiliki gdpp growth dan gdpp high.

#### Link Youtube

<https://youtu.be/IPdzTCIb8Kk>

Link Medium : <https://medium.com/@prasetyow938/the-relationship-between-prosperity-levels-and-inflation-investment-and-gender-equality-in-the-dad2b220deb>

## Reference

Materi Pacmann

<https://www.pythoncharts.com/matplotlib/pie-chart-matplotlib/>

<https://python.plainenglish.io/how-to-make-a-beautiful-donut-chart-and-nested-donut-chart-in-matplotlib-92040c8bbeea>

<https://www.macrotrends.net/countries/ranking/foreign-direct-investment>

<https://hdr.undp.org/data-center/thematic-composite-indices/gender-inequality-index#/indicies/GII>

<https://data.worldbank.org/>