

## Analisis Regresi dan Korelasi untuk Proyeksi Produksi Minyak Bumi dan Gas Alam Indonesia menggunakan Bahasa Pemrograman Python

Kasliono<sup>1)</sup>, Edi Suharmono<sup>2)</sup>, Povi<sup>3)</sup>, Risca Meriani<sup>4)</sup>, Niken Candraningrum<sup>5)</sup>

<sup>1)</sup>Program Studi Rekayasa Sistem Komputer, Fakultas MIPA, Universitas Tanjungpura

<sup>2)3)4)</sup>Program Studi Kimia, Fakultas MIPA, Universitas Tanjungpura

<sup>5)</sup>Program Studi Teknik Informatika, Fakultas Teknik, Universitas Tanjungpura

\*Correspondence Author: [kasliono@siskom.untan.ac.id](mailto:kasliono@siskom.untan.ac.id), Pontianak, Indonesia

DOI: <https://doi.org/10.37012/jtik.v9i2.1756>

### Abstrak

Penelitian ini memiliki tujuan untuk mengetahui suatu cara pendekatan analisis data menggunakan bahasa pemrograman *Python* yang dapat diterapkan dalam industri minyak bumi dan gas alam serta memprediksi hasil produksi minyak bumi dan gas alam sampai pada tahun 2030. Penggunaan metode dalam penelitian ini yaitu dengan pendekatan deskriptif kuantitatif. Tujuan digunakannya metode ini yaitu untuk menguraikan secara sistematis peristiwa atau kejadian yang terjadi melalui penggunaan angka-angka dalam menganalisis data penelitian ini. Data tersebut kemudian diolah dengan bahasa pemrograman *Python* menggunakan *library* seperti *Pandas*, *NumPy*, *Matplotlib*, dan *Scikit-Learn*. Dalam penelitian ini data diolah dengan cara analisis regresi dan korelasi. Hasil penelitian yang diperoleh yaitu terjadinya penurunan yang cukup signifikan dari hasil prediksi produksi minyak bumi dan gas alam setiap tahunnya. Hasil prediksi produksi minyak bumi dan gas alam yang paling besar terjadi pada tahun 2022 yang menghasilkan minyak bumi sebesar 210.218,41 (000 barel) dan gas alam sebesar 2.709.176 (MMscf). Sedangkan hasil prediksi produksi minyak bumi dan gas alam pada tahun 2030 yaitu sebesar 116.827,69 (000 barel) dan 2.597.292 (MMscf). Minyak bumi dan gas alam dalam penelitian ini mempunyai keterkaitan yang lemah dengan nilai korelasi positif. Nilai korelasi sebesar 0.387558 menunjukkan bahwa adanya kecenderungan ketika produksi minyak bumi meningkat, produksi gas alam juga cenderung meningkat, begitupun sebaliknya.

**Kata Kunci:** Regresi Linear, Korelasi, Minyak Bumi, Gas Alam

### Abstract

*This research aims to find out a way to approach data analysis using the Python programming language which can be applied in the petroleum and natural gas industry as well as predicting petroleum and natural gas production results until 2030. The method used in this research is a quantitative descriptive approach. The purpose of using this method is to systematically describe events or occurrences that occur through the use of numbers in analyzing this research data. The data is then processed in the Python programming language using libraries such as Pandas, NumPy, Matplotlib, and Scikit-Learn. In this research, the data was processed using regression and correlation analysis. The research results obtained were that there was a significant decline in the predicted results of oil and natural gas production every year. The predicted results for the largest production of petroleum and natural gas will occur in 2022, which will produce petroleum of 210,218.41 (000 barrels) and natural gas of 2,709,176 (MMscf). Meanwhile, the predicted results for oil and natural gas production in 2030 are 116,827.69 (000 barrels) and 2,597,292 (MMscf). Petroleum and natural gas in this study have a weak relationship with positive correlation values. The correlation value of 0.387558 shows that there is a tendency when oil production increases, natural gas production also tends to increase, and vice versa.*

**Keywords:** Linear Regression, Correlation, Petroleum, Natural Gas

## PENDAHULUAN

Minyak bumi dan gas alam adalah dua jenis aset alam yang tidak dapat diperbaharui. Minyak bumi terdiri dari suatu campuran yang kompleks dan sebagian besar mengandung zat hidrokarbon. Beberapa contoh zat hidrokarbon yang ada dalam minyak bumi meliputi alkana, sikloalkana dan komponen-komponen lainnya (Lendeng, dkk., 2021). Gas bumi ataupun gas alam merupakan salah satu jenis energi fosil yang ada di ladang-ladang gas alam, minyak, serta tambang batubara. Komponen utama dari gas alam yaitu metana ( $\text{CH}_4$ ). Gas alam termasuk ke dalam kategori sumber daya alam dengan cadangan terbesar ketiga di dunia setelah batu bara dan minyak bumi (Pratomo dan Fajar., 2022). Hingga sekarang minyak bumi dan gas alam tetap menjadi barang dagangan yang terus diekspor dan diimpor di Indonesia karena kedua sumber daya alam tersebut sangat penting dalam memenuhi kebutuhan energi dunia (Dewanto., 2023). Kedua sumber daya alam tersebut menjadi sumber utama bahan bakar fosil yang digunakan dalam berbagai sektor seperti transportasi, industri, dan rumah tangga. Selain itu, minyak bumi dan gas alam juga termasuk dalam sumber energi yang memiliki nilai penting bagi aktivitas ekonomi di tingkat nasional (Nonci, dkk., 2020). Namun, hasil produksi minyak bumi di dalam negeri saat ini masih terus mengalami penurunan. Hal tersebut terjadi disebabkan oleh faktor geopolitik dan keterbatasan sumber daya alam yang semakin memburuk (Dewanto., 2023). Selain itu, penyebab lain penurunan minyak bumi saat ini yaitu dikarenakan industri minyak bumi di tingkat nasional yang telah mencapai usia yang lebih dari 100 tahun, sehingga hasil produksinya pun menurun.

Dalam sejarah Republik Indonesia, puncak produksi minyak bumi tertinggi hanya terjadi dua kali yaitu pada tahun 1977 dan 1995. Pada tahun 1977 produksi minyak bumi mencapai 1,68 juta bpd dan pada tahun 1995 produksi mencapai 1,62 juta bpd. Setelah tahun 1995 produksi minyak bumi di Indonesia cenderung mengalami penurunan dengan laju penurunan alami yaitu sekitar 12% setiap tahunnya. Akan tetapi, sejak tahun 2004 penurunan produksi minyak bumi berhasil ditahan menjadi sekitar 3% setiap tahunnya.

Produksi minyak bumi dan gas alam secara efisien dan berkelanjutan menjadi sangat penting dalam menjaga kestabilan ekonomi dan keberlanjutan lingkungan (Restra dan Kesdm, 2015; Nonci, dkk., 2020). Maka dari itu, data produksi minyak bumi dan gas alam sangat dibutuhkan dalam operasional industri minyak bumi dan gas alam untuk menjadi

---

elemen kunci dalam mengidentifikasi tren dan memahami karakteristik lapangan minyak dan gas, serta mengoptimalkan proses produksi. Analisis data produksi tersebut menjadi kritis dalam mengambil keputusan strategis bagi perusahaan minyak dan gas, regulator pemerintah, serta lembaga penelitian di bidang energi. Oleh karena itu, untuk memprediksi hasil dari minyak bumi dan gas alam setiap tahunnya diperlukan pendekatan analisis yang lebih canggih dan efisien agar mendapatkan data dari produksi minyak bumi dan gas alam tanpa harus turun ke lapangan. Data tersebut dapat diprediksi dengan menggunakan sistem pemrograman *Python*.

*Python* merupakan salah satu bahasa pemrograman yang sangat canggih, yang diciptakan oleh Guido Van Rossum dan diluncurkan pada tahun 1991. Selain itu, *Python* juga termasuk ke dalam bahasa pemrograman yang sangat diminati akhir-akhir ini dan banyak dipakai oleh perusahaan-perusahaan besar (Alfarizi, dkk., 2023). *Python* telah menjadi bahasa pemrograman pilihan di berbagai bidang, termasuk *Machine Learning*, *Deep Learning*, ilmu data dan analisis statistik. *Python* memiliki beberapa kelebihan yaitu, sintaksis yang mudah dipahami, memiliki banyak *library* yang dapat digunakan untuk menganalisis data seperti *Pandas*, *NumPy*, dan *Matplotlib*, serta beragam alat analisis yang terus berkembang seperti *SciPy* dan *Scikit-Learn* (Alfarizi, dkk., 2023).

Proses penggunaan *Python* dalam menganalisis data produksi minyak bumi dan gas alam yakni dengan analisis yang mendalam, dan visualisasi yang informatif sehingga dapat memberikan pengolahan data secara cepat. Selain itu, dengan menggunakan bahasa pemrograman *Python*, para peneliti dan praktisi di industri minyak dan gas alam dapat mengembangkan model prediksi produksi, melakukan analisis deplesi lapangan minyak dan gas, serta melakukan simulasi skenario untuk mengantisipasi dampak kebijakan atau kejadian alam yang berbeda. Maka dari itu, peneliti ingin memberikan penjelasan tentang bagaimana cara pendekatan analisis data menggunakan bahasa pemrograman *Python* yang dapat diterapkan dalam industri minyak dan gas alam, memprediksi hasil produksi minyak bumi dan gas alam untuk beberapa tahun yang akan datang serta menyampaikan informasi mengenai hubungan minyak bumi dan gas alam.

## METODE

Penelitian ini dilaksanakan pada akhir bulan Juli 2023 hingga akhir Agustus 2023. Tempat pelaksanaannya beralokasi di Universitas Tanjungpura Pontianak. Metode dari penelitian ini yaitu menggunakan pendekatan deskriptif kuantitatif. Adapun tujuan digunakannya metode ini yaitu untuk menguraikan secara sistematis peristiwa atau kejadian yang terjadi melalui penggunaan angka-angka dalam menganalisis data penelitian ini. Penelitian kuantitatif ini mampu menganalisis aspek dari suatu situasi peristiwa ataupun kejadian yang telah terjadi sebelumnya. Fokus dari penelitian kuantitatif ini yaitu untuk menggambarkan karakteristik apa adanya. Sumber data yang dijadikan acuan dalam penelitian ini yaitu data sekunder (Dewanto., 2023). Data sekunder tersebut diambil dari data instansi Badan Pusat Statistik (BPS) pada produksi minyak bumi dan gas alam Indonesia tahun 1996-2021. Data tersebut telah disusun dengan baik oleh instansi Badan Pusat Statistik (BPS) sehingga data tersebut siap untuk diolah. Berikut ini data-data yang digunakan dalam penelitian ini:

**Tabel 1.** Data Produksi Minyak Bumi dan Gas Alam Indonesia Tahun 1996-2021

Tahun	Minyak Mentah dan Kondensat (000 barel)	Gas Alam (MMscf)
1996	548.648,30	3.164.016,20
1997	543.752,60	3.166.034,90
1998	534.892,00	2.978.851,90
1999	494.643,00	3.068.349,10
2000	484.393,30	2.845.532,90
2001	480.116,10	3.762.828,50
2002	397.308,50	2.279.373,90
2003	383.700,00	2.142.605,00
2004	404.992,90	3.026.069,30
2005	387.653,50	2.985.341,00
2006	357.477,40	2.948.021,60
2007	348.348,00	2.805.540,30
2008	358.718,70	2.790.988,00
2009	346.313,00	2.887.892,20
2010	344.888,00	3.407.592,30
2011	329.249,30	3.256.378,90
2012	314.665,90	2.982.753,50
2013	301.191,90	2.433.364,00
2014	287.902,20	2.433.364,00
2015	286.814,20	2.433.364,00
2017	292.373,80	2.433.364,00
2018	281.826,61	2.433.364,00
2019	273.494,80	2.433.364,00
2020	259.246,80	2.433.364,00
2021 <sup>x</sup>	240.324,50	2.433.364,00

Ket: x = angka sementara

Dalam penelitian ini data tersebut diolah dengan menggunakan bahasa pemrograman *Python* dengan cara analisis regresi dan korelasi. Analisis regresi merupakan suatu pendekatan yang dapat memeriksa korelasi antara sebuah variabel yang disebut sebagai variabel terikat, dengan satu atau dua variabel yang berperan sebagai variabel penjelas. Variabel terikat ini juga dikenal dengan variabel respon, sedangkan variabel penjelas disebut variabel *independen* (Gujarati, 2003; Syilfi, dkk., 2012).

Melalui penggunaan analisis regresi *linear*, peneliti dapat memprediksi nilai variabel menjadi lebih akurat. Dalam analisis regresi, variabel terbagi menjadi dua jenis yakni variabel terikat dan variabel bebas. Variabel terikat disimbolkan dengan huruf Y. Variabel terikat adalah variabel yang dapat terpengaruh oleh kondisi variabel lain. Variabel bebas atau variabel prediktor disimbolkan dengan huruf X. Variabel bebas merupakan variabel yang tidak terpengaruh oleh kondisi variabel lain (Shaputra dan Syarif., 2021). Analisis regresi juga dapat dimanfaatkan untuk mengidentifikasi korelasi sebab-akibat antara variabel X dan variabel Y berdasarkan situasi yang aktual. Dalam penelitian ini, persamaan regresi dimanfaatkan untuk mendapatkan garis regresi yang cocok dengan *scatter plot* data. Garis regresi ini diperoleh melalui metode kuadrat yang paling kecil, sehingga dihasilkan bentuk persamaan regresinya seperti yang tertera dibawah ini:

$$Y = \alpha + \beta X$$

Keterangan:

Y = Nilai estimasi/ taksiran untuk variabel terikat (tak bebas Y)

$\alpha$  = Titik potong garis regresi pada sumbu Y

$\beta$  = Gradien garis regresi

X = Nilai variabel bebas

(Wibowo dan Andriyatna., 2020)

Selain itu, analisis korelasi dalam penelitian ini dimanfaatkan untuk mengidentifikasi tingkat kedekatan keterkaitan antara dua atau lebih variabel yang berbeda, dan diukur melalui koefisien korelasi. Korelasi juga berguna untuk menentukan ukuran kekuatan hubungan antara variabel-variabel tersebut dengan rentang nilai 0 sampai 1. Koefisien korelasi memiliki nilai dengan rentang dari -1 sampai 1, di mana -1 menunjukkan adanya hubungan negatif yang sempurna (terbalik), 0 menunjukkan ketiadaan hubungan, dan 1

menunjukkan hubungan positif yang sempurna. Sehingga koefisien korelasi memiliki hubungan yang erat dengan persamaan regresi, sebab persamaan regresi merepresentasikan keterikatan antara dua atau lebih variabel (Wibowo dan Andriyatna., 2020).

**Tabel 2.** Interpretasi terhadap Koefisien Korelasi (Sumber: Sanny dan Rina., 2020)

No	Interval	Tingkat Hubungan
1.	0,00-0,199	Sangat Lemah
2.	0,20-0,399	Lemah
3.	0,40-0,599	Sedang
4.	0,60-0,799	Kuat
5.	0,80-1,000	Sangat Kuat

## HASIL DAN PEMBAHASAN

Pengambilan data dalam penelitian ini dilaksanakan dengan mengambil informasi dari data BPS nasional mengenai capaian produksi minyak bumi dan gas alam tahun 1996-2021. Proses dilakukan dalam mengolah data tersebut yaitu dengan menggunakan bahasa pemrograman *Python* yang terdiri dari *Pandas*, *NumPy*, *Matplotlib*, dan *Scikit-Learn*. Rumus *library* dari *Python* tersebut penulis gunakan untuk menganalisis tren perkembangan migas di Indonesia dari tahun 2022-2030. Berdasarkan hasil penelitian didapatkan data prediksi dari produksi minyak bumi dan gas alam Indonesia sebagai berikut:

**Tabel 3.** Data Prediksi Migas Indonesia Tahun 2022-2030

Tahun	Minyak Mentah dan Kondensat (000 barel)	Gas Alam (MMscf)
2022	210.218,41	2.709.176
2023	198.544,57	2.695.190
2024	186.870,73	2.681.205
2025	175.196,89	2.667.219
2026	163.523,05	2.653.234
2027	151.849,21	2.639.249
2028	140.175,37	2.625.263
2029	128.501,53	2.611.278
2030	116.827,69	2.597.292

Dari data yang tercantum pada tabel diatas, dapat dilihat bahwa terjadi penurunan yang cukup signifikan dari hasil proyeksi produksi migas setiap tahunnya. Berdasarkan dari data penelitian tersebut hasil dari proyeksi produksi minyak bumi yang paling besar terjadi pada tahun 2022 yang menghasilkan minyak bumi sebesar 210.218,41 (000 barel) dan hasil proyeksi produksi gas alam yang paling besar terjadi pada tahun 2022 yang menghasilkan

gas alam sebesar 2.709.176 (MMscf). Adapun hasil proyeksi dari penelitian yang diperoleh yaitu tingkat keakuratan data prediksi produksi minyak bumi Indonesia tahun 2022-2030 sebesar 91%. Sedangkan tingkat keakuratan data prediksi produksi gas alam Indonesia tahun 2022-2030 hanya sebesar 10%.

Dalam memperoleh data prediksi minyak bumi dan gas alam Indonesia tahun 2022-2030 dilakukan dengan menggunakan bahasa pemrograman *Python* yaitu melalui analisis regresi dan korelasi. Berikut ini beberapa cara yang dilakukan dalam menganalisis data menggunakan bahasa pemrograman *Python* untuk memprediksi produksi minyak bumi dan gas alam Indonesia tahun 2022-2030:

Langkah awal yang dilakukan yaitu meng-*import library* dengan memasukkan perintah seperti pada Gambar 1.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

**Gambar 1.** *Import library*

Pandas merupakan *library* yang digunakan untuk mengakses dan membaca data yang berbentuk file seperti csv. Numpy merupakan *library* yang digunakan untuk kebutuhan *scientific* dan matematis, dalam hal ini memproses data pada analisis regresi linear sederhana dengan Python berupa array. *Matplotlib* merupakan *library* yang digunakan untuk visualisasi data menjadi grafik plot. *Sklearn* merupakan *library* yang digunakan untuk membuat *machine learning*. Dalam *library Sklearn* perintah yang dilakukan untuk analisis regresi *linear* sederhana yang terdapat pada *scikit learn* diperlukan perintah `from sklearn.linear_model import LinearRegression`.

Selanjutnya, tahap yang dilakukan untuk mengkoneksikan *google colab* ke *google drive* yaitu dengan melakukan perintah seperti pada Gambar 2 berikut ini.

```
from google.colab import drive
drive.mount('/content/drive')
```

↳ Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force\_remount=True).

**Gambar 2.** Mengkoneksikan Google Colab ke Google Drive

Langkah selanjutnya yaitu mengakses dan menampilkan data dengan perintah seperti pada Gambar 3.

```
[3] pd.read_csv('/content/drive/MyDrive/Colab Notebooks/prod_minyak_bumi_gas_alam.csv')
```

	Tahun	Minyak Bumi	Gas Alam
0	1996	548648300	3164016.20
1	1997	543752600	3166034.90
2	1998	534892000	2978851.90
3	1999	494643000	3068349.10

**Gambar 3.** Tampilan Data Csv

Kemudian tahap yang dilakukan selanjutnya yaitu membuat fungsi untuk mendefinisikan perintah yang berada pada Gambar 3 menjadi seperti Gambar 4. Lalu, perintah dimasukkan dengan menggunakan perintah `prod_minyak_gas.shape` dengan tujuan untuk mengetahui ukuran data. Ukuran data csv yang ditampilkan yaitu 25 baris dan 3 kolom.

```
[ ] prod_minyak_gas = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/prod_minyak_bumi_gas_alam.csv')
print(prod_minyak_gas.head())
prod_minyak_gas.shape
```

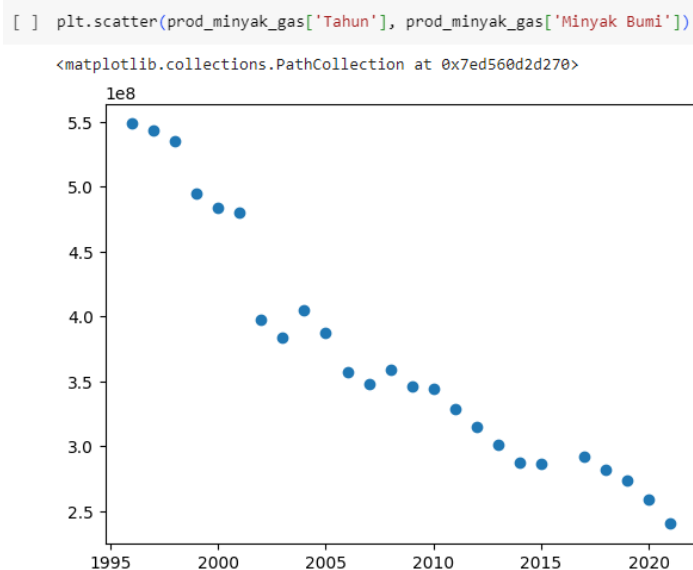
	Tahun	Minyak Bumi	Gas Alam
0	1996	548648300	3164016.2
1	1997	543752600	3166034.9
2	1998	534892000	2978851.9
3	1999	494643000	3068349.1
4	2000	484393300	2845532.9

(25, 3)

**Gambar 4.** Membuat Fungsi dan Melihat Ukuran Data

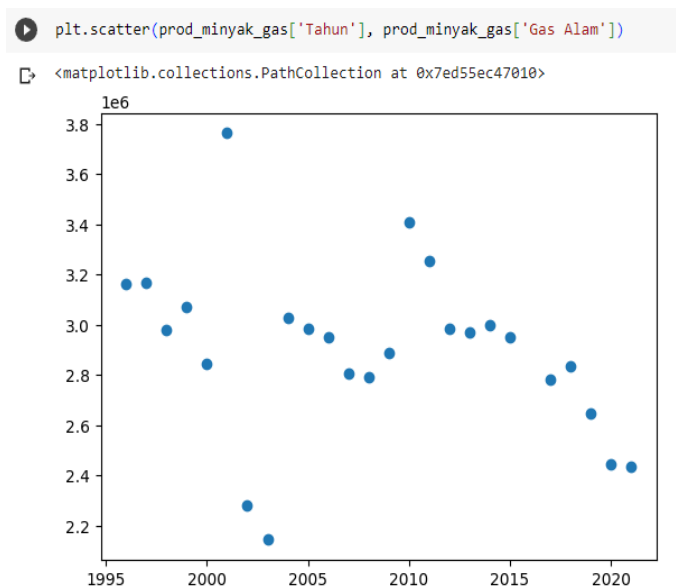
Proses yang dilakukan berikutnya yaitu membentuk pola data antara tahun dan minyak bumi melalui *scatter plot* seperti pada Gambar 5 berikut ini.





**Gambar 5.** Pola Data Antara Tahun dan Minyak Bumi

Berdasarkan grafik yang ditampilkan pada Gambar 5 tersebut, data antara tahun dan minyak bumi membentuk pola *linear*. Selanjutnya, membentuk pola data antara tahun dan gas alam melalui *scatter plot* seperti pada Gambar 6 berikut ini.



**Gambar 6.** Pola Data Antara Tahun dan Gas Alam

Grafik yang ditampilkan pada Gambar 6 menunjukkan pola *linear* antara data tahun dan gas alam. Hasil dari pola *linear* tersebut menunjukkan bahwa adanya hubungan antara tahun dan produksi minyak bumi di Indonesia. Artinya, produksi minyak bumi cenderung meningkat atau menurun secara proporsional dengan perubahan tahunnya. Berikutnya yaitu

melihat nilai korelasi. Analisis korelasi berfungsi untuk melihat hubungan antara variabel sehingga dapat diketahui hubungan sebab dan akibat dari variabel yang ada (Sholeh., 2022). Apabila keterkaitan antara dua variabel tidak *linear*, maka koefisien dari korelasi tidak dapat mengindikasikan tingkat keakuratan antara dua variabel. Nilai koefisien korelasi terletak pada rentang  $-1 < 0 < 1$ . Apabila  $r = -1$  maka nilai korelasinya menjadi negatif sempurna. Hal ini menunjukkan bahwa hubungan antara dua variabel yaitu variabel X terhadap variabel Y menjadi sangat lemah. Apabila  $r = 1$  maka nilai korelasinya dapat dinyatakan sempurna, yang artinya bahwa hubungan antara variabel X dan variabel Y sangat kuat. Simbol “r” artinya korelasi jika diukur dalam sampel (Sudjana, 2005). Koefisien korelasi dengan angka 0 mengindikasikan bahwa tidak ada keterkaitan antara dua variabel yang sedang diuji. Apabila terbentuk garis lurus pada sebaran data maka hubungan antara kedua variabel merupakan linear sempurna. Namun, pada praktiknya sangat jarang ditemukan data yang dapat menghasilkan pola garis lurus yang sempurna (Safitri, 2016). Berikut nilai korelasi dari penelitian ini yang dapat dilihat pada Gambar 7 dibawah ini.

```
[ ] prod_minyak_gas.corr()
```

	Tahun	Minyak Bumi	Gas Alam
Tahun	1.000000	-0.955151	-0.310688
Minyak Bumi	-0.955151	1.000000	0.387558
Gas Alam	-0.310688	0.387558	1.000000

**Gambar 7.** Nilai Korelasi Data

Berdasarkan nilai korelasi yang ditampilkan pada Gambar 7, dapat disimpulkan bahwa korelasi antara "Tahun" dan "Minyak Bumi" yaitu sebesar -0.955151. Hal tersebut mengindikasikan bahwa adanya keterkaitan yang kuat secara negatif antara tahun dan produksi minyak bumi. Artinya, semakin tinggi tahunnya, semakin rendah produksi minyak bumi, begitu juga sebaliknya. Adapun korelasi antara "Tahun" dan "Gas Alam" yaitu -0.310688. Hal tersebut mengindikasikan bahwa adanya keterkaitan negatif yang lebih lemah antara tahun dan produksi gas alam. Artinya, ada kecenderungan bahwa produksi gas alam menurun seiring bertambahnya tahun, tetapi hubungannya tidak sekuat hubungan antara tahun dan produksi minyak bumi. Serta korelasi antara "Minyak Bumi" dan "Gas Alam"

yaitu 0.387558. Hal ini mengindikasikan bahwa adanya keterkaitan positif yang lemah antara produksi minyak bumi dan produksi gas alam. Artinya, ketika produksi minyak bumi meningkat, ada kecenderungan bahwa produksi gas alam juga akan meningkat, tetapi hubungannya tidak sekuat hubungan antara tahun dan produksi minyak bumi. Berdasarkan hasil dari data korelasi tersebut maka, semakin mendekati nilai 1 atau -1, maka korelasi menjadi semakin kuat dan apabila semakin mendekati nilai 0, maka korelasi menjadi semakin lemah atau hampir tidak ada hubungan antara variabel tersebut.

Selanjutnya, yaitu melakukan permodelan data variabel bebas dan terikat. Permodelan ini dilakukan dengan tujuan untuk mengetahui pengaruh satu variabel *independent* (variabel bebas) terhadap variabel *independent* (variabel terikat) (Wahyudin, dkk., 2020). Bentuk dari variabel bebas dalam tipe *feature* matriks dan variabel terikat dalam vektor target. Dalam penelitian ini variabel bebas menggunakan kolom tahun sementara variabel terikat menggunakan kolom minyak bumi dan gas alam. Berikut ini merupakan perintah untuk permodelan antara data tahun dan minyak bumi yang dapat dilihat pada Gambar 8.

```
[ ] x_Tahun = prod_minyak_gas['Tahun']
X_Tahun = x_Tahun[:, np.newaxis]
y_prodminyak = prod_minyak_gas['Minyak Bumi'].values

regressor = LinearRegression()
model = regressor.fit(X_Tahun, y_prodminyak)

<ipython-input-8-21759fbbb8d3>:2: FutureWarning: Support for multi-dimensional indexing (e.g. `obj[:, None]`) is deprecated and will be removed
X_Tahun = x_Tahun[:, np.newaxis]
```

**Gambar 8.** Permodelan Data

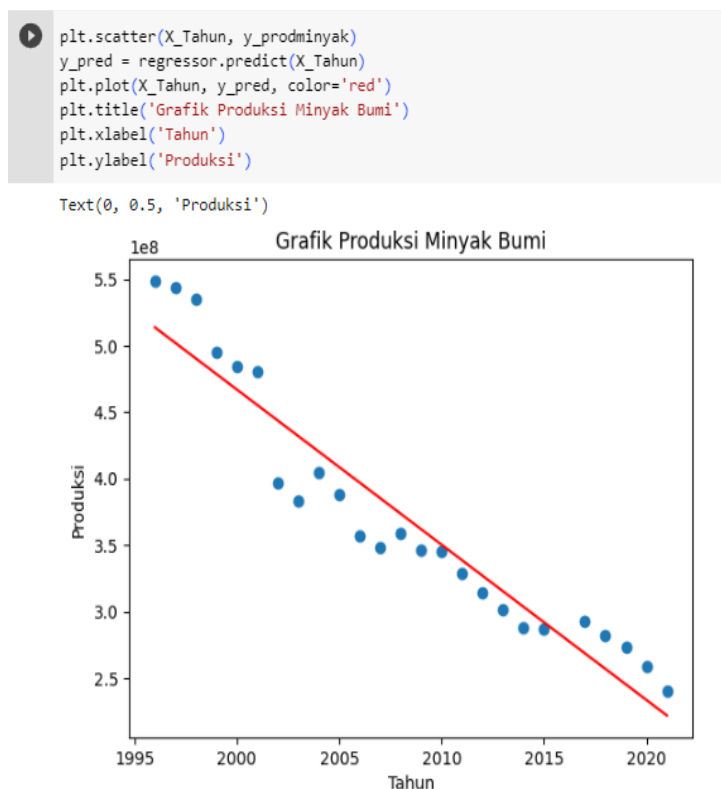
Langkah yang dilakukan selanjutnya yaitu mencari nilai koefisien dan *intercept* dari data tahun dengan minyak bumi seperti yang ditunjukkan pada Gambar 9 dibawah ini.

```
[9] regressor.coef_
array([-11673840.07264957])

[10] regressor.intercept_
23814723046.294872
```

**Gambar 9.** Nilai Koefisien dan *Intercept* dari Data Tahun dengan Minyak Bumi

Berdasarkan langkah sebelumnya diperoleh nilai koefisien sebagai  $\beta_1$  yaitu 11673840,073 dan nilai *intercept* sebagai  $\beta_0$  yaitu 23814723046,295. Persamaan garis yang digunakan dalam analisis *linear* sederhana penelitian ini yaitu  $y = \beta_0 + \beta_1 X$ , sehingga diperoleh nilai  $y = 23814723046,295 - 11673840,073X$ . Lalu untuk melihat garis regresi *linear* dapat dilakukan dengan perintah seperti pada Gambar 10 berikut.



**Gambar 10.** Grafik Garis Regresi Linear

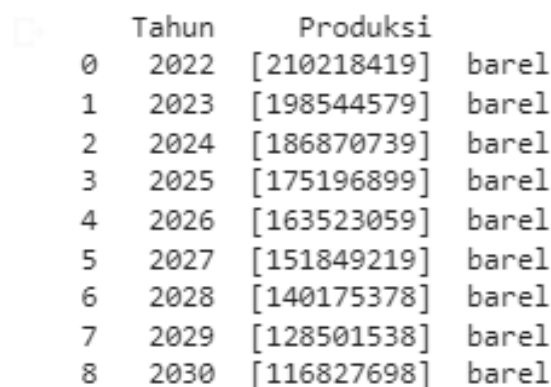
Kemudian proses yang dilakukan yaitu mencari tahu data prediksi produksi minyak bumi sampai tahun 2030 dengan perintah yang dapat dilihat pada Gambar 11 berikut.

```
[12] masuk = int(input('Tahun yang akan diprediksi :'))
X_baru = np.array([[masuk]])
y_pred = regressor.predict(X_baru)
y_pred = y_pred.astype(int)
print(y_pred, 'barel')
```

Tahun yang akan diprediksi :2022  
[210218419] barel

**Gambar 11.** Prediksi Produksi Minyak Bumi

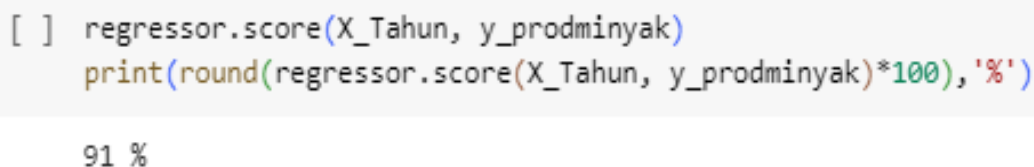
Selanjutnya yaitu menyusun data prediksi produksi minyak bumi hingga tahun 2030 menjadi bentuk list, dapat dilihat pada Gambar 12.



	Tahun	Produksi
0	2022	[210218419] barel
1	2023	[198544579] barel
2	2024	[186870739] barel
3	2025	[175196899] barel
4	2026	[163523059] barel
5	2027	[151849219] barel
6	2028	[140175378] barel
7	2029	[128501538] barel
8	2030	[116827698] barel

**Gambar 12.** Data Prediksi Minyak Bumi Hingga Tahun 2030

Kemudian untuk melihat keakuratan data prediksi yang telah didapatkan dapat dilakukan dengan perintah seperti pada Gambar 13 dibawah ini.



```
[ ] regressor.score(X_Tahun, y_prodminyak)
print(round(regressor.score(X_Tahun, y_prodminyak)*100), '%')

91 %
```

**Gambar 13.** Akurasi Prediksi Produksi Minyak Bumi

Hasil eksekusi dari perintah tersebut menghasilkan nilai 91% yang berarti keakuratan data proyeksi produksi minyak bumi sebesar 91%.

Langkah selanjutnya yaitu melakukan permodelan pada data variabel bebas dan terikat menggunakan data tahun dan gas alam. Berikut ini perintah yang dilakukan untuk permodelan antara data tahun dan gas alam, seperti pada Gambar 14.



```
x_Tahun = prod_minyak_gas['Tahun']
X_Tahun = x_Tahun[:, np.newaxis]
y_prodgas = prod_minyak_gas['Gas Alam'].values

regressor = LinearRegression()
model = regressor.fit(X_Tahun, y_prodgas)

<ipython-input-40-483209b0580f>:2: FutureWarning: Support for multi-dimensional indexing (e.g. `obj[:, None]`) is deprecated and will
X_Tahun = x_Tahun[:, np.newaxis]
```

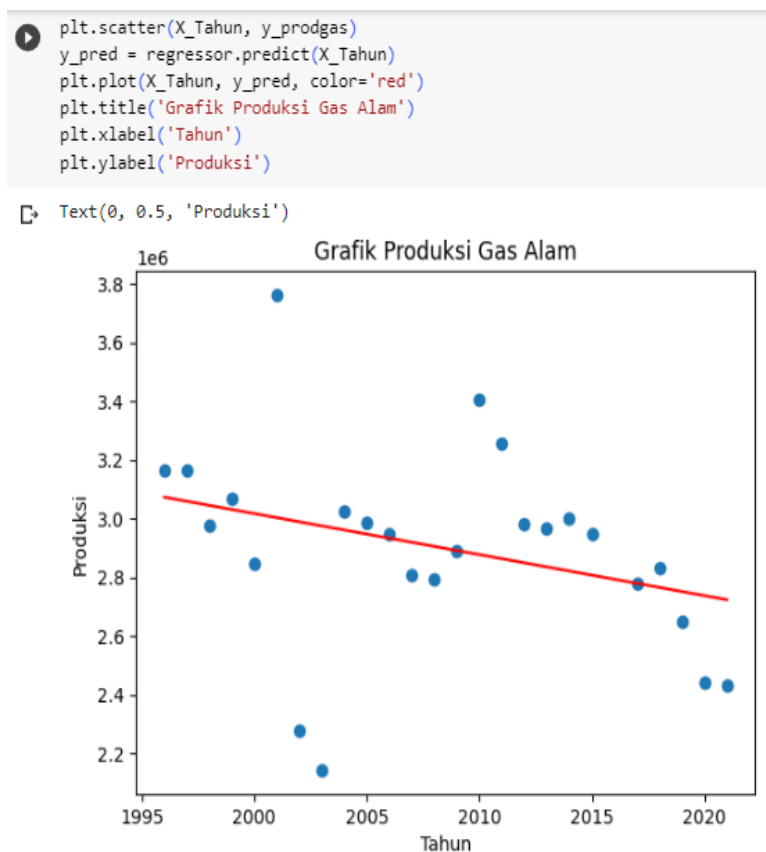
**Gambar 14.** Permodelan Data

Selanjutnya, yaitu mencari nilai koefisien dan *intercept* dari data tahun dengan gas alam seperti pada Gambar 15 berikut.

```
[41] regressor.coef_  
  
array([-13985.45357692])  
  
regressor.intercept_  
  
30987763.41757694
```

**Gambar 15.** Nilai Koefisien dan *Intercept* dari Data Tahun dengan Gas Alam

Berdasarkan langkah sebelumnya diperoleh nilai koefisien sebagai  $\beta_1$  yaitu -13985,454 dan nilai *intercept* sebagai  $\beta_0$  yaitu 30987763,418, sehingga diperoleh nilai  $y = 30987763,418 - 13985,454X$ . Lalu untuk melihat garis regresi linear dapat dilakuakn dengan perintah, seperti yang ada pada Gambar 16 berikut.



**Gambar 16.** Grafik Garis Regresi Linear

Langkah selanjutnya yaitu mencari tahu data prediksi produksi gas alam hingga tahun 2030 dengan perintah seperti yang terdapat pada Gambar 17 berikut.

```
masuk = int(input('Tahun yang akan diprediksi :'))  
X_baru = np.array([[masuk]])  
y_pred = regressor.predict(X_baru)  
y_pred = y_pred.astype(int)  
print(y_pred, 'MMscf')  
  
Tahun yang akan diprediksi :2022  
[2709176] MMscf
```

**Gambar 17.** Prediksi Produksi Gas Alam

Kemudian proses yang dilakukan berikutnya yaitu menyusun data prediksi produksi gas alam hingga tahun 2030 menjadi bentuk list seperti pada Gambar 18 berikut.

	Tahun	Produksi	
0	2022	[2709176]	MMscf
1	2023	[2695190]	MMscf
2	2024	[2681205]	MMscf
3	2025	[2667219]	MMscf
4	2026	[2653234]	MMscf
5	2027	[2639249]	MMscf
6	2028	[2625263]	MMscf
7	2029	[2611278]	MMscf
8	2030	[2597292]	MMscf

**Gambar 18.** Data Prediksi Gas Alam Hingga Tahun 2030

Selanjutnya yaitu melihat keakuratan data prediksi yang telah didapatkan dengan melakukan perintah seperti pada Gambar 19 berikut.

```
regressor.score(X_Tahun, y_prodgas)  
print(round(regressor.score(X_Tahun, y_prodgas)*100), '%')  
  
10 %
```

**Gambar 19.** Akurasi Prediksi Produksi Gas Alam

Hasil eksekusi dari perintah tersebut menghasilkan nilai 10% yang berarti keakuratan data prediksi produksi gas alam yaitu sebesar 10%.

---

## KESIMPULAN DAN REKOMENDASI

Kesimpulan yang dapat diambil dari hasil penelitian ini sebagai berikut:

1. Untuk melakukan analisis prediksi produksi minyak bumi dan gas alam dalam beberapa tahun yang akan datang dapat dilakukan dengan bahasa pemrograman *Python*, menggunakan beberapa *library* seperti *Pandas*, *NumPy*, *Matplotlib*, dan *Scikit-Learn*. Tingkat akurasi keakuratan dari hasil analisis prediksi minyak bumi ini cukup tinggi yaitu sebesar 91%. Namun, pada gas alam tingkat akurasi keakuratannya cukup rendah yaitu hanya sebesar 10%.
2. Hasil penelitian yang diperoleh yaitu terjadinya penurunan yang cukup signifikan dari hasil prediksi produksi minyak bumi dan gas alam setiap tahunnya. Hasil prediksi produksi minyak bumi dan gas alam yang paling besar terjadi pada tahun 2022 yang menghasilkan minyak bumi sebesar 210.218,41 (000 barel) dan gas alam sebesar 2.709.176 (MMscf). Sedangkan hasil prediksi produksi minyak bumi dan gas alam pada tahun 2030 yaitu sebesar 116.827,69 (000 barel) dan 2.597.292 (MMscf).
3. Minyak bumi dan gas alam dalam penelitian ini mempunyai keterkaitan yang lemah dengan nilai korelasi positif. Nilai korelasi sebesar 0.387558 menunjukkan bahwa adanya kecenderungan ketika produksi minyak bumi meningkat, produksi gas alam juga cenderung meningkat, begitupun sebaliknya.

## REFERENSI

- Alfarizi, M. R. S., Al-farish, M. Z., Taufiqurrahman, M., Ardiansah, G., & Elgar, M. (2023). Penggunaan Python Sebagai Bahasa Pemrograman untuk Machine Learning dan Deep Learning. *Karimah Tauhid*, 2(1), 1–6.
- Dewanto, M. E. (2023). Proyeksi Produksi Migas Indonesia Sampai dengan Tahun 2045. *Jurnal Ilmu Ekonomi*, 7(02), 195–210.
- Gujarati, D. (2003). *Ekonometrika Dasar*. Zain, S, penerjemah. Terjemahan dari: *Basic Econometric*. Erlangga: Jakarta.
- Lendeng, L. C., Sugiarto, B. A., & Rumagit, A. M. (2021). Interactive Learning based on Animation in Petroleum Subject for Grade XI Senior High. *J. Tek. Elektro Dn Komput*, 16(2), 183–192.



- Nonci, R. (2020). Analisa Deskripsi Minyak dan Gas (Study Kasus Lapangan “X”). *Equilibrium: Jurnal Penelitian Pendidikan Dan Ekonomi*, 17(02), 44–50.
- Pratomo, L. B., & TK, B. F. (2022). Tinjauan Singkat Optimalisasi Penggunaan Gas Bumi pada Sektor Rumah Tangga. *Eksergi*, 18(1), 1–11.
- Restra, & Kesdm. (2015). Rencana Strategis Kementerian Energi dan Sumber Daya Mineral. Kementerian Energi dan Sumber Daya Mineral.
- Safitri, W. R., (2016). Analisis Korelasi Pearson dalam Menentukan Hubungan Antara Kejadian Demam Berdarah dengan Kepadatan Penduduk di Kota Surabaya pada Tahun 2012-2014. *Jurnal Ilmiah Keperawatan*. 2(2).
- Sanny, B. I., & Rina, K. D. (2020). Pengaruh Net Interest Margin (NIM) terhadap Return on Asset (ROA) pada PT Bank Pembangunan Daerah Jawa Barat dan Banten Tbk Periode 2013-2017. *Jurnal E-Bis (Ekonomi-Bisnis)*. 4(1), 82.
- Shaputra, R. D., & Syarif, H. (2021). Implementasi Regresi Linier untuk Prediksi Penjualan dan Cash Flow pada Aplikasi Point of Sales Restoran. *Jurnal Automata*. 2(1).
- Sholeh, M. (2022). Penerapan Regresi Linear Ganda Untuk Memprediksi Hasil Nilai Kuesioner Mahasiswa Dengan Menggunakan Python. *Jurnal Dinamika Informatika*, 11(1), 13–24.
- Sudjana., (2005). *Metode Statistika*. Bandung: Tarsito.
- Syilfi., Dwi, I., & Diah, S. (2012). Analisis Regresi Linier Piecewise Dua Segmen. *Jurnal Gaussian*, 1(1), 220.
- Wahyudin, A. A.-F. N., Primajaya, A., & Irawan, A. S. Y. (2020). Penerapan Algoritma Regresi Linear Berganda pada Estimasi Penjualan Mobil Astra Isuzu. *Techno. Com*, 19(4), 364–374.
- Wibowo, R. A., & Andriyatna, A. K. (2020). Analisis Korelasi dalam Penentuan Arah Antar Faktor pada Pelayanan Angkutan Umum di Kota Magelang. *Theta Omega: Journal of Electrical Engineering, Computer and Information Technology*, 1(2), 2745-6412.