

Multi Regresi Linear Untuk Prediksi Kepadatan Penduduk Berdasarkan Variable Kependudukan Di Indonesia

Wisnu Arizky Kurniawan¹, Najmi Baihaqi², Isnan Ali³

Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya

Email : wisnuarizky01@gmail.com¹, najmibaihaqi@gmail.com², Isnanali072@gmail.com³

ABSTRAKSI

Penduduk adalah orang yang bertempat tinggal dan telah menetap atau berniat menetap dalam minimum 1 tahun pada wilayah tertentu. Masalah yang akan dibahas dalam penelitian ini adalah masalah kepadatan penduduk Indonesia yang semakin cepat dan mendesak. Dalam penelitian ini, variabel yang meliputi jumlah penduduk, penduduk yang lahir, dan penduduk yang meninggal dunia di Indonesia akan digunakan untuk memperkirakan jumlah penduduk yang akan terserap di masa mendatang. Tingkat kepadatan penduduk di Indonesia saat ini sulit diprediksi dengan cepat. Hal ini dikarenakan adanya pergerakan penduduk yang tidak terkendali dari Kawasan perdesaan ke perkotaan. Pergerakan ini dipengaruhi kondisi ekonomi penduduk yang menggantungkan hidupnya dari mencari pekerjaan di wilayah lain. Imbasnya, peningkatan dan kepadatan penduduk suatu wilayah di Indonesia seringkali berubah dengan cepat. Tujuan penelitian ini adalah untuk memprediksi jumlah penduduk di Indonesia menggunakan metode multi regresi linear sehingga dapat diidentifikasi kepadatan penduduk suatu wilayah. Penelitian ini menggunakan teknik multi linear regresi untuk mengestimasi jumlah penduduk yang berfokus pada pertumbuhan penduduk di Indonesia disetiap tahunnya. Data yang digunakan adalah data dari Badan Pusat Statistik (BPS) Indonesia periode 2010-2022. Variabel yang digunakan adalah jumlah penduduk, penduduk yang meninggal, dan penduduk yang lahir di Indonesia. Dengan menggunakan pendekatan multi regresi linear untuk mengembangkan model prediksi jumlah penduduk Indonesia. Metode ini digunakan untuk memprediksi jumlah penduduk berdasarkan data historis yang diperoleh dari Badan Pusat Statistik (BPS) Indonesia pada tahun 2010 sampai dengan tahun 2022 dengan tetap memperhatikan beberapa variabel seperti penduduk, penduduk yang meninggal, dan penduduk yang lahir. Prediksi yang dilakukan pada tahun 2023 menghasilkan angka kematian sebesar 1.926.300, kelahiran sebesar 4.638.200, dan penduduk sebesar 279.410.057. Dengan secara keseluruhan jumlah penduduk di Indonesia pada tahun 2023 sebesar 282.121.957.

Kata Kunci: kepadatan penduduk, variable kependudukan, multi regresi linear, prediksi, Indonesia

ABSTRACT

The population is defined as individuals who reside or intend to reside for a minimum of 1 year in a specific region. The issue to be addressed in this research is the rapid and pressing problem of population density in Indonesia. In this study, variables including population size, births, and deaths in Indonesia will be used to estimate the absorbed population in the future. The current population density in Indonesia is difficult to predict accurately due to uncontrolled migration from rural areas to urban areas. This migration is influenced by the economic conditions of the population, who seek livelihoods in other regions. As a result, the population increase and density in a particular area of Indonesia often change rapidly. The objective of this research is to predict the population size in Indonesia using the multiple linear regression method to identify the population density in a given area. This study employs the multiple linear regression technique to estimate the population size, focusing on population growth in Indonesia each year. The data used is from the Indonesian Central Bureau of Statistics (BPS) for the period 2010-2022. The variables used are population size, deaths, and births in Indonesia. By utilizing the multiple linear regression approach, a predictive model for the population size in Indonesia is developed based on historical data obtained from the Indonesian Central Bureau of Statistics (BPS) for the years 2010 to 2022, considering variables such as population, deaths, and births. The prediction for the year 2023 yields a death count of 1,926,300, a birth count of 4,638,200, and a population size of 279,410,057. Thus, the overall population in Indonesia for the year 2023 is projected to be 282,121,957.

Keywords: population density, population variables, multiple linear regression, prediction, Indonesia



Tanggal Submit : 14/07/2023
Tanggal Diterima : 29/03/2024
Tanggal Terbit : 30/03/2024

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0) International License (<https://creativecommons.org/licenses/by-nc-sa/4.0/>).

Publisher's Note: JPPM stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

I. PENDAHULUAN

Tingkat kepadatan penduduk di Indonesia saat ini sulit diprediksi dengan cepat. Hal ini dikarenakan adanya pergerakan penduduk yang tidak terkendali khususnya dari kawasan perdesaan ke perkotaan akibat tidak meratanya pembangunan ekonomi wilayah [1]. Pergerakan ini dipengaruhi kondisi ekonomi penduduk yang menggantungkan hidupnya dari mencari pekerjaan di wilayah lain [2]. Imbasnya, peningkatan dan kepadatan penduduk suatu wilayah di Indonesia seringkali berubah dengan cepat serta berpotensi menimbulkan permasalahan baru, seperti meningkatkan angka kemiskinan, pengangguran dan bahkan tingkat kejahatan.

Permasalahan kepadatan penduduk sangat mempengaruhi pengampu kebijakan dalam merencanakan dan menentukan arah pembangunan dimasa mendatang termasuk sulitnya pemerintah dan organisasi terkait untuk menghasilkan kebijakan dan strategi yang tepat untuk mengatasi pertumbuhan penduduk [3]. Oleh karenanya, sangat penting untuk dapat menentukan trend kepadatan penduduk suatu wilayah dengan cepat dan sesuai kondisi terkini. Kepadatan penduduk yang teridentifikasi dengan baik, mampu melahirkan kebijakan yang dapat mencegah terjadinya ketimpangan antara jumlah penduduk dan ketersediaan kebutuhan sehari-hari tertentu, seperti lapangan kerja, air, pangan, dan perumahan [1]. Ada banyak Teknik yang dapat digunakan untuk mengidentifikasi kepadatan penduduk, salah satunya menggunakan pendekatan prediksi berbasis multi linear regression (MLR).

MLR itu cara prediksi yang menggunakan garis lurus untuk menggambarkan hubungan diantara dua variabel atau lebih [4]. Pendekatan MLR banyak digunakan untuk menyelesaikan permasalahan prediksi, seperti prediksi pertumbuhan penduduk. Proyeksi penduduk merupakan prediksi jumlah penduduk yang menggunakan penghitungan ilmiah yang didasarkan variabel yang berpengaruh terhadap pertumbuhan penduduk dimasa yang akan datang. Variabel-variabel akan menentukan bagaimana besaran jumlah penduduk dan struktur penduduk. [1], [2], prediksi tentang penjualan [5], prediksi produksi padi [6], dan prediksi harga sewa [7]. Penggunaan teknik MLR tersebut menjadi dasar untuk digunakan dalam memprediksi kepadatan penduduk di dalam penelitian ini. Metode ini dipilih karena dapat menghasilkan estimasi yang

memungkinkan terbentuknya pola keterkaitan antar variabel [8], memungkinkan pengguna dengan mudah memasukkan beberapa variabel independent [4] dan mampu memprediksi suatu ketidakpastian dengan mengekstraksi pola data, dan proses perhitungan lebih singkat dengan perhitungan secara parallel [8].

Penelitian ini bertujuan untuk membangun model prediksi kepadatan penduduk Indonesia menggunakan metode linear multi-regresi. Hasil Penelitian diharapkan dapat memberikan masukan yang signifikan bagi pemerintah dan stakeholder lainnya dalam mempertimbangkan kebijakan dan strategi untuk mengatasi masalah kepadatan penduduk di Indonesia. Dengan demikian, hasil penelitian ini dapat digunakan oleh pemerintah untuk mengembangkan program dan kebijakan yang tepat untuk mengatasi masalah Pertumbuhan penduduk di Indonesia, seperti program keluarga berencana, peningkatan kualitas pendidikan dan kesehatan, dan pengembangan ekonomi. Hasil penelitian ini juga dapat memberikan kontribusi dalam pengembangan ilmu pengetahuan dan teknologi di Indonesia. Metode multi regresi linear yang diterapkan dalam penelitian ini juga dapat digunakan pada penelitian selanjutnya yang berkaitan dengan masalah serupa, sehingga dapat berkontribusi secara signifikan bagi kemajuan ilmu pengetahuan dan teknologi di Indonesia.

II. METODE PENELITIAN

2.1. Metode Regresi Linear

Dalam penelitian ini, metode yang digunakan adalah metode regresi linear berganda. Regresi linear berganda merupakan model regresi yang melibatkan lebih dari satu variabel independen. Analisis regresi linear berganda dilakukan untuk mengetahui arah dan seberapa besar pengaruh variabel independen terhadap variabel dependen. teknik statistik untuk mempelajari hubungan antara variabel dependen (y) dan dua atau lebih variabel independen (x_1, x_2, \dots, x_n). Seringkali regresi linier dapat digunakan untuk memprediksi nilai variabel dependen (y) berdasarkan nilai variabel independen yang diberikan (x_1, x_2, \dots, x_n).

Dalam regresi linier berganda, model matematis digunakan untuk menggambarkan hubungan antara variabel independen dan dependen. Model regresi linier berganda dapat dinyatakan sebagai berikut:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n + \varepsilon$$

Di mana:

- y adalah variabel dependen yang ingin Anda prediksi. x_1, x_2, \dots, x_n adalah variabel bebas yang digunakan untuk memprediksi nilai y .
- $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ adalah koefisien regresi atau parameter model yang menentukan besarnya pengaruh variabel independen terhadap variabel dependen.
- ε adalah kesalahan atau penyimpangan dari prediksi ke nilai sebenarnya.

Untuk membuat model regresi linier, seringkali diperlukan estimasi terlebih dahulu parameter $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ menggunakan metode kuadrat terkecil. Teknik ini memperkirakan parameter dengan meminimalkan jarak antara nilai y yang diamati dan nilai y yang diprediksi model. Setelah estimasi parameter selesai, model regresi linier berganda dapat digunakan untuk memprediksi nilai y baru berdasarkan nilai x_1, x_2, \dots, x_n yang diberikan. Metode ini sering digunakan dalam analisis data untuk memprediksi nilai variabel dependen yang kompleks dimana variabel independen dapat memberikan pengaruh yang berbeda terhadap variabel dependen.

2.2. Penentuan Atribut

Atribut-atribut yang digunakan untuk proses prediksi ini mengacu pada tujuan penelitian. Ada dua jenis variabel yang ditentukan, yaitu :

1) Variabel dependen (Y)

Variabel dependen (Y) merupakan variabel yang nilainya terikat, bisa disebut variabel terikat. Variabel Y yang digunakan yaitu penduduk

2) Variabel independen (X)

Variabel independen (X) merupakan variabel yang nilainya tidak tergantung pada nilai dari variabel lainnya atau bisa disebut sebagai variabel bebas. Variabel X yang digunakan adalah angka kematian dan kelahiran.

IV. HASIL DAN PEMBAHASAN

Data-data pendukung penelitian diperoleh dari situs resmi BPS (Badan Pusat Statistik) yang dimiliki oleh pemerintah yang mencakup data-data penduduk yang ada di Indonesia. Data yang diambil untuk penelitian ini adalah data kependudukan di seluruh Indonesia yang dihitung berdasarkan 3 variabel, diantaranya yaitu data kematian, kelahiran, dan penduduk.

Berdasarkan output regresi yang diberikan, kita dapat melihat bahwa model prediksi ini cukup akurat, dengan nilai R Square sebesar 0.9954 yang artinya sekitar 99.54% dari variasi data telah dijelaskan oleh model ini. Selain itu, kedua variabel independen (X Variable 1 dan X Variable 2) memiliki koefisien yang signifikan, dengan p-value yang sangat rendah (<0.001), sehingga kita dapat mengasumsikan bahwa kedua variabel tersebut memiliki pengaruh signifikan terhadap variabel dependen (yaitu jumlah penduduk).

Dari output regresi juga dapat dilihat bahwa persamaan regresi yang digunakan dalam model ini adalah:

$$y = 1723262741 - 137.6X \text{ Variable 1} - 254.15X \text{ Variable 2}$$

Dengan variabel dependen (y) adalah jumlah penduduk, X Variable 1 dan X Variable 2 adalah variabel independen yang masing-masing mewakili faktor-faktor tertentu yang mempengaruhi pertumbuhan penduduk. Berdasarkan hasil estimasi koefisien regresi, dapat dilihat bahwa variabel X Variable 1 memiliki koefisien sebesar -137.6006775 , yang menunjukkan bahwa setiap peningkatan satu satuan pada variabel tersebut akan berkontribusi terhadap penurunan sekitar 137.6006775 satuan pada variabel dependen (jumlah penduduk). Sementara itu, variabel X Variable 2 memiliki koefisien sebesar 254.1486998 , yang menunjukkan bahwa setiap peningkatan satu satuan pada variabel tersebut akan berkontribusi terhadap penurunan sekitar 254.1486998 satuan pada variabel dependen.

Dalam prediksi jumlah penduduk pada tahun-tahun berikutnya, model regresi tersebut dapat digunakan untuk memperkirakan jumlah penduduk di masa depan dengan memasukkan nilai variabel independen (X Variable 1 dan X Variable 2) yang relevan ke dalam persamaan regresi. Namun, perlu diingat bahwa hasil prediksi tersebut bersifat perkiraan dan dapat terpengaruh oleh faktor-faktor yang tidak dipertimbangkan dalam model regresi. Oleh karena itu, hasil prediksi perlu diinterpretasikan dengan hati-hati dan diuji keakuratannya dengan membandingkan dengan data aktual di masa depan. Data-data pendukung penelitian diperoleh dari situs resmi BPS (Badan Pusat Statistik) yang dimiliki oleh pemerintah yang mencakup data-data penduduk yang ada di Indonesia. Data yang diambil untuk penelitian ini adalah data kependudukan di seluruh Indonesia yang dihitung berdasarkan 3 variabel, diantaranya yaitu data kematian, kelahiran, dan penduduk.

Tabel 1. Data Penelitian

Tahun	Angka kematian	angka kelahiran	penduduk	Total Penduduk
2010	1.524.100	5.013.700	237.641.326	241.130.926
2011	1.541.500	4.997.100	241.993.000	245.448.600
2012	1.558.500	4.974.800	245.428.000	248.844.300
2013	1.577.500	4.949.000	248.818.000	252.189.500
2014	1.598.700	4.922.300	252.166.000	255.489.600
2015	1.622.200	4.894.600	254.821.000	258.093.400
2016	1.653.100	4.868.800	258.706.000	261.921.700
2017	1.684.000	4.839.700	261.891.000	265.046.700
2018	1.717.300	4.810.000	265.019.000	268.111.700
2019	1.753.200	4.779.800	268.077.000	271.103.600
2020	1.789.000	4.746.400	270.206.000	273.163.400
2021	1.833.000	4.712.500	272.685.000	275.564.500
2022	1.879.600	4.678.300	275.773.774	278.572.474
2023	1.926.300	4.638.200	279.410.057	282.121.957

Berdasarkan output regresi yang diberikan, kita dapat melihat bahwa model prediksi ini cukup akurat, dengan nilai R Square sebesar 0.9954 yang artinya sekitar 99.54% dari variasi data telah dijelaskan oleh model ini. Selain itu, kedua variabel independen (X

Variable 1 dan X Variable 2) memiliki koefisien yang signifikan, dengan p-value yang sangat rendah (<0.001), sehingga kita dapat mengasumsikan bahwa kedua variabel tersebut memiliki pengaruh signifikan terhadap variabel dependen (yaitu jumlah penduduk).

ANOVA					
	df	SS	MS	F	Significance F
Regression	2	1,795,897,294,399,480	897,948,647,199,740	1087.71238	2.00595E-12
Residual	10	8,255,386,841,102	825,538,684,110		
Total	12	1,804,152,681,240,580			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95,0%	Upper 95,0%
Intercept	1723262741	176054948.1	9.788209645	1.93345E-06	1330987871	2115537611	1330987871	2115537611
X Variable 1	-137.6006775	26.01583907	-5.289111651	0.000353004	-195.5675793	-79.63377574	-195.5675793	-79.63377574
X Variable 2	-254.1486998	27.29938084	-9.309687324	3.04966E-06	-314.9755108	-193.3218887	-314.9755108	-193.3218887

Dari output regresi juga dapat dilihat bahwa persamaan regresi yang digunakan dalam model ini adalah:

$$y = 1723262741 - 137.6X \text{ Variable 1} - 254.15X \text{ Variable 2}$$

Dengan variabel dependen (y) adalah jumlah penduduk, X Variable 1 dan X Variable 2 adalah variabel independen yang masing-masing mewakili faktor-faktor tertentu yang mempengaruhi pertumbuhan penduduk. Berdasarkan hasil estimasi koefisien regresi, dapat dilihat bahwa variabel X Variable 1 memiliki koefisien sebesar -137.6006775, yang menunjukkan bahwa setiap peningkatan satu satuan pada variabel tersebut akan berkontribusi terhadap penurunan sekitar 137,6006775 satuan pada variabel dependen (jumlah penduduk). Sementara itu, variabel X Variable 2 memiliki koefisien sebesar 254.1486998, yang menunjukkan bahwa setiap peningkatan satu satuan pada variabel tersebut akan berkontribusi terhadap penurunan sekitar 254,1486998 satuan pada variabel dependen.

Dalam prediksi jumlah penduduk pada tahun-tahun berikutnya, model regresi tersebut dapat digunakan untuk memperkirakan jumlah penduduk di masa depan dengan memasukkan nilai variabel independen (X Variable 1 dan X Variable 2) yang relevan ke dalam persamaan regresi. Namun, perlu diingat bahwa hasil prediksi tersebut bersifat perkiraan dan dapat terpengaruh oleh faktor-faktor yang tidak dipertimbangkan dalam model regresi. Oleh karena itu, hasil prediksi perlu diinterpretasikan dengan hati-hati dan diuji keakuratannya dengan membandingkan dengan data aktual di masa depan.

V. KESIMPULAN

Penelitian ini bertujuan untuk mengembangkan model prediksi perkembangan penduduk di Indonesia dengan menggunakan metode multi regresi linear. Metode regresi linier dipilih karena dapat menghasilkan estimasi berdasarkan data historis sehingga memungkinkan munculnya jaringan antar atribut yang mempengaruhi kesejahteraan penduduk di Indonesia. Data yang digunakan dalam penelitian ini berasal dari Badan Pusat Statistik (BPS) Indonesia periode 2010-

2022, yang mencakup variabel-variabel seperti jumlah penduduk, penduduk yang meninggal, dan penduduk yang lahir di Indonesia. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi penting bagi pemerintah dan stakeholder lainnya dalam merencanakan kebijakan dan strategi untuk mengatasi masalah pertumbuhan penduduk di Indonesia. Metode multi regresi linear yang digunakan dalam penelitian ini juga dapat diaplikasikan pada penelitian selanjutnya terkait masalah-masalah yang serupa.

Dalam hal prediksi jumlah penduduk pada tahun-tahun berikutnya, model regresi tersebut dapat digunakan untuk memperkirakan jumlah penduduk di masa depan dengan memasukkan nilai variabel independen (X Variable 1 dan X Variable 2) yang relevan ke dalam persamaan regresi. Namun, perlu diingat bahwa hasil prediksi tersebut bersifat perkiraan dan dapat terpengaruh oleh faktor-faktor yang tidak dipertimbangkan dalam model regresi. Oleh karena itu, hasil prediksi perlu diinterpretasikan dengan hati-hati dan diuji keakuratannya dengan membandingkan dengan data aktual di masa depan.

DAFTAR PUSTAKA

- [1] I. Indriani, D. Siregar, and A. P. Windarto, "Penerapan Metode Linear Regression dalam Mengestimasi Jumlah Penduduk," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 4, p. 1112, Aug. 2022, doi: 10.30865/jurikom.v9i4.4676.
- [2] C. Adi Rahmat and Y. Novianto, "Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM) Penerapan Metode Regresi Linier Berganda Untuk Mengestimasi Laju Pertumbuhan Penduduk Kabupaten Musi Banyuasin." [Online]. Available: <http://ejournal.unama.ac.id/index.php/jakakom>
- [3] H. Noviyanto and A. Fauzi, "PREDIKSI PERTUMBUHAN PENDUDUK DI INDONESIA MENGGUNAKAN METODE LEAST SQUARE PREDICTION OF POPULATION GROWTH IN INDONESIA USING THE LEAST SQUARE METHOD."
- [4] W. A. L. Sinaga, S. Sumarno, and I. P. Sari, "The Application of Multiple Linear Regression Method for Population Estimation Gunung Malela District," *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, vol. 1, no. 1, pp. 55–64, Mar. 2022, doi: 10.55123/jomlai.v1i1.143.
- [5] Y. Asohi, "IMPELEMENTASI ALGORITMA REGRESI LINIER BERGANDA UNTUK PREDIKSI PENJUALAN," 2020.
- [6] E. Triyanto, H. Sismoro, and A. D. Laksito, "IMPLEMENTASI ALGORITMA REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI PRODUKSI PADI DI KABUPATEN BANTUL," *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 4, no. 2, pp. 66–75, Jul. 2019, doi: 10.36341/rabit.v4i2.666.
- [7] M. Reza Fahlepi, A. Widjaja, and J. Surya Sumantri No, "Penerapan Metode Multiple Linear Regression Untuk Prediksi Harga Sewa Kamar Kost," 2019.
- [8] E. Pertumbuhan Penduduk di et al., "Estimation of Population Growth in Tasikmalaya District Using Multiple Linear Regression Methods."
- [9] T. Wilson, I. Grossman, M. Alexander, P. Rees, and J. Temple, "Methods for Small Area Population Forecasts: State-of-the-Art and Research Needs," *Popul Res Policy Rev*, vol. 41, no. 3, pp. 865–898, Jun. 2022, doi: 10.1007/s11113-021-09671-6.