



## Penerapan Data Mining Untuk Prediksi Jumlah Kunjungan Wisatawan di Kebumen Menggunakan Metode Regresi Linier Sederhana

Rohmatulloh Muhamad Ikhsanuddin<sup>1</sup>, Shona Chayy Bilqisth<sup>2</sup>, Helmi Bahar Alim<sup>3</sup>  
Universitas Putra Bangsa<sup>1</sup>, Universitas Negeri Semarang<sup>2</sup>  
email: ikhsanuddin@fst.universitaspurabangsa.ac.id

Page | 1

### ABSTRAK

Kebumen memiliki potensi wisata alam di pantai selatan Jawa Tengah sebagai alternatif berwisata dengan unggulan destinasi pantai seperti Pantai Menganti, Pantai Karang Bolong, Goa Jatijajar dan Waduk Sempor. Pariwisata memiliki peranan dalam peningkatan pendapatan daerah sehingga perlu didukung pemerintah melalui investasi penunjang pariwisata. Prediksi jumlah wisatawan di Kebumen sebagai faktor yang mempengaruhi investor dalam membuka usaha tentunya diperlukan dalam pengambilan keputusan. Berdasarkan data BPS jumlah kunjungan di Kebumen pada Oktober 2024 sebanyak 257.745 sehingga diharapkan prediksi jumlah kunjungan wisatawan tahun berikutnya dapat menjadikan minat investor di Kebumen. Hasil perhitungan menunjukkan bahwa metode regresi linier menghasilkan akurasi sebesar 94,60% dapat membantu pemerintah untuk meyakinkan investor pendukung pariwisata di Kebumen.

**Kata Kunci:** prediksi; pariwisata; metode regresi linier

### ABSTRACT

*Kebumen has natural tourism potential on the south coast of Central Java as an alternative for tourism with superior beach destinations such as Menganti Beach, Karang Bolong Beach, Jatijajar Cave and Sempor Reservoir. Tourism has a role in increasing regional income so it needs to be supported by the government through tourism supporting investments. Predicting the number of tourists in Kebumen as a factor influencing investors in opening a business is certainly needed in decision making. Based on BPS data, the number of visits to Kebumen in October 2024 was 257.745, so it is hoped that predictions of the number of tourist visits next year can generate investor interest in Kebumen. The calculation results show that the linear regression method produces an accuracy of 94,60% which can help the government to convince investors to support tourism in Kebumen.*

**Keywords:** prediction; tourism; linear regression method

### PENDAHULUAN

Pariwisata merupakan salah satu sektor penunjang pendapatan daerah sehingga pemerintah daerah harus dapat memanfaatkan. Pemerintah daerah memiliki kebijakan otonomi daerah sehingga memiliki otoritas penuh bagi daerahnya untuk memberdayakan potensi daerah yang ada. Pariwisata sebagai salah satu kebijakannya yaitu penarikan pajak dan retribusi dengan didukung pelayanan publik yang baik dari pemerintah daerah (Alwi dkk, 2021).

Kebumen sebagai salah satu kabupaten yang berada di pesisir pantai selatan dari Provinsi Jawa Tengah memiliki potensi wisata. Potensi wisata yang dimiliki Kebumen paling banyak adalah wisata alam mulai pesona wisata pantai, waduk, goa dan pemandian. Potensi wisata di Kebumen yang paling banyak diminati antara lain Pantai Menganti, Pantai Pandan Kuning, Pantai Karang Bolong, Waduk Sempor, Goa Jatijajar, dan Pemandian Air Panas Krakal. Potensi wisata yang dimiliki inilah menjadi daya tarik alternatif lokasi wisata di pantai selatan Jawa Tengah sehingga perlu didukung oleh pemerintah daerah. Wilayah Kebumen saat ini sudah didukung oleh jalur transportasi seperti jalur kereta api, bus lintas selatan Jawa, serta bandara internasional Yogyakarta.

Menurut data BPS per 2 Desember 2024 bahwa jumlah perjalanan wisatawan nusantara dengan tujuan Kabupaten Kebumen pada bulan Oktober 2024 tercatat sebanyak 257.745 pengunjung. Jumlah kunjungan wisatawan di Kebumen inilah yang perlu dilakukan untuk memprediksi jumlah kunjungan wisatawan pada periode waktu berikutnya sehingga dapat dimanfaatkan oleh pemerintah. Pemerintah daerah diharapkan dapat menggunakan data prediksi jumlah kunjungan wisatawan untuk meningkatkan minat pengusaha berinvestasi di sektor pariwisata Kebumen. Prediksi jumlah wisatawan yang berkunjung di Kebumen tentu dapat meyakinkan investor dalam mendirikan usaha pendukung dibidang pariwisata seperti pengusaha jasa penginapan, jasa sewa kendaraan, jasa usaha makan, jasa pusat oleh-oleh dan kerajinan.

Penelitian ini bertujuan untuk mengetahui prediksi jumlah pengunjung wisatawan di Kabupaten Kebumen dengan menggunakan metode regresi linier sederhana. Prediksi merupakan dugaan atau prediksi mengenai terjadinya suatu kejadian atau peristiwa diwaktu yang akan mendatang (Baihaqi dkk, 2019). Fungsi dari prediksi yaitu membuat suatu rencana kebutuhan yang harus dibuat yang dinyatakan dalam kuantitas atau jumlah sebagai fungsi dari waktu. Metode regresi linier sederhana adalah teknik data mining untuk menentukan bahwa ada hubungan antara variable yang diprediksi dengan variabel lainnya (Maulana dkk, 2024).

Penelitian terdahulu yang sudah dilakukan untuk prediksi jumlah mahasiswa baru dengan pengujian akurasi dengan MAPE diperoleh hasil untuk prodi matematika, ilmu kelautan, biologi, sistem informasi, arsitektur, dan teknik lingkungan dengan masing-masing sebesar 92,8%, 91,24%, 94,16%, 93,54%, 92,02%, dan 92,48% (Almumtazah dkk, 2021). Selanjutnya pada penelitian berkaitan dengan hasil prediksi jumlah mahasiswa baru di Universitas Bina Bangsa prodi manajemen hasil pengujian tingkat akurasi sebesar 96,556% (Muhartini dkk, 2021). Penelitian yang lain pada prediksi kunjungan pasien pada RSUD Cilacap hasil akurasi pada 26 model menghasilkan tingkat akurasi sebesar 80% atau termasuk dalam kategori sangat baik dan baik (Baihaqi dkk, 2019).

## METODE

Penelitian ini menggunakan metode regresi linier sederhana dengan langkah awal melakukan identifikasi kebutuhan pada permasalahan minat investor dibidang pendukung pariwisata di Kebumen. Untuk meyakinkan minat diperlukan data prediksi jumlah kunjungan wisatawan pada tahun berikutnya sehingga pengumpulan data dilakukan dari data BPS Kebumen. Setelah data kunjungan wisata sudah diperoleh tahapan berikutnya dilakukan analisis metode regresi linier untuk menentukan prediksi jumlah kunjungan wisatawan di Kebumen. Persamaan regresi linier sudah ditentukan selanjutnya dilakukan pengujian validasi nilai error dari metode regresi linier sederhana dengan menggunakan nilai MAPE. Tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

Penelitian ini menggunakan data BPS Kebumen pada laman [kebumenkab.bps.go.id](http://kebumenkab.bps.go.id) dengan periode Januari 2022 sampai dengan Oktober 2024. Data kunjungan wisata di Kebumen sebagai bahan proses prediksi pada tahun berikutnya dapat dilihat pada Tabel 1.

Tabel 1. Data Kunjungan Wisata di Kebumen

BULAN	TAHUN		
	2022	2023	2024
Januari	549,171	367,050	434,447
Februari	456,765	271,408	278,082
Maret	570,630	365,239	350,477
April	309,898	528,165	692,685
Mei	468,381	448,342	318,930
Juni	235,154	354,968	305,496
Juli	285,518	379,742	276,411
Agustus	216,310	257,342	283,920
September	255,647	272,564	315,094
Oktober	285,985	298,057	257,745
November	228,085	280,899	-
Desember	288,098	285,033	-

Metode regresi linier merupakan salah satu perhitungan time series metode kuantitatif dimana waktu yang digunakan sebagai dasar prediksi. Tahapan menggunakan metode regresi linier yaitu :

- a. Menentukan nilai intercept dengan persamaan sebagai berikut :

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

- b. Menentukan nilai koefisien dengan persamaan sebagai berikut :

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

- c. Menentukan persamaan regresi dengan persamaan sebagai berikut :

$$Y = a + bX$$

dimana :

Y = Variabel terikat

a = Intercept

b = Koefisien variabel X

X = Variabel bebas

- d. Menghitung nilai eror dengan menggunakan MAPE (Mean Absolute Percentage Error)

$$MAPE = \frac{\sum \frac{|Y - Y'|}{Y} x 100\%}{n}$$

dimana :

Y = Data aktual

Y' = Data prediksi

n = Jumlah data

**HASIL DAN PEMBAHASAN**

Tahapan awal dalam penelitian ini yaitu menentukan variabel faktor penyebab (X) yaitu periode waktu dan variabel faktor akibat (Y) yaitu jumlah wisatawan yang berkunjung di Kebumen. Metode regresi linier sederhana digunakan untuk membantu menentukan prediksi jumlah kunjungan wisatawan di Kebumen pada periode bulan berikutnya. Berikut tahapan metode regresi linier sederhana sebagai berikut :

1. Menghitung nilai XY dan X<sup>2</sup>

Tabel 2. Perhitungan XY dan X<sup>2</sup>

TAHUN	BULAN	PERIODE BULAN (X)	JUMLAH WISATAWAN (Y)	X.Y	X <sup>2</sup>
2022	Januari	1	549,171	549,171	1
	Februari	2	456,765	913,530	4
	Maret	3	570,630	1,711,890	9
	April	4	309,898	1,239,592	16
	Mei	5	468,381	2,341,905	25
	Juni	6	235,154	1,410,924	36
	Juli	7	285,518	1,998,626	49
	Agustus	8	216,310	1,730,480	64
	September	9	255,647	2,300,823	81
	Oktober	10	285,985	2,859,850	100
	November	11	228,085	2,508,935	121
	Desember	12	288,098	3,457,176	144
2023	Januari	13	367,050	4,771,650	169
	Februari	14	271,408	3,799,712	196
	Maret	15	365,239	5,478,585	225
	April	16	528,165	8,450,640	256
	Mei	17	448,342	7,621,814	289
	Juni	18	354,968	6,389,424	324
	Juli	19	379,742	7,215,098	361
	Agustus	20	257,342	5,146,840	400
<b>JUMLAH</b>		<b>210</b>	<b>7,121,898</b>	<b>71,896,665</b>	<b>2,870</b>

2. Mengitung nilai intercept

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$a = \frac{(7.121.898)(2.870) - (210)(71.896.665)}{20(2.870) - (210)^2} = 401.620,1$$

3. Menghitung nilai koefisien

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{20(71.896.665) - (210)(7.121.898)}{20(2.870) - (210)^2} = - 4.335,735$$

4. Menentukan persamaan regresi  
 $Y = a + bX$ , sehingga diperoleh  
 $Y = 401.620,1 - 4.335,735 X$
5. Menentukan prediksi  
 Setelah persamaan regresi telah ditentukan, langkah berikutnya menentukan nilai prediksi sesuai dengan persamaan  $Y = 401.620,1 - 4.335,735 X$ . Nilai X diperoleh dari periode bulan pada nilai X 21 – 30 bulan sehingga diperoleh hasil pada Tabel 3.

Tabel 3. Hasil Nilai Y Prediksi

TAHUN	BULAN	PERIODE BULAN (X)	Y Prediksi
2023	September	21	310,569.7
	Oktober	22	306,233.9
	November	23	301,898.2
	Desember	24	297,562.5
2024	Januari	25	293,226.7
	Februari	26	288,891
	Maret	27	284,555.3
	April	28	280,219.5
	Mei	29	275,883.8
	Juni	30	271,548.1
<b>JUMLAH</b>		<b>255</b>	<b>2,910,589</b>

6. Menghitung nilai error  
 Tahapan menghitung nilai error dihitung berdasarkan perbandingan nilai Y aktual dengan nilai Y prediksi sehingga diperoleh hasil pada Tabel 4.

Tabel 4. Hasil Perhitungan Nilai Error

TAHUN	BULAN	PERIODE BULAN (X)	Y Aktual	Y Prediksi	Y-Y'	Y-Y'	MAPE 100
2023	September	21	272,564	310,569.68	-38,005.68	38,005.68	13.94
	Oktober	22	298,057	306,233.94	-8,176.94	8,176.94	2.74
	November	23	280,899	301,898.21	-20,999.21	20,999.21	7.48
	Desember	24	285,033	297,562.47	-12,529.47	12,529.47	4.40
2024	Januari	25	434,447	293,226.74	141,220.26	141,220.26	32.51
	Februari	26	278,082	288,891.00	-10,809.00	10,809.00	3.89
	Maret	27	350,477	284,555.27	65,921.73	65,921.73	18.81
	April	28	692,685	280,219.53	412,465.47	412,465.47	59.55
	Mei	29	318,930	275,883.80	43,046.20	43,046.20	13.50
	Juni	30	305,496	271,548.06	33,947.94	33,947.94	11.11
<b>JUMLAH</b>		<b>255</b>	<b>3,516,670</b>	<b>2,910,588.70</b>	<b>606,081.30</b>	<b>787,121.91</b>	<b>167.92</b>

Berdasarkan hasil perhitungan di atas dapat disimpulkan bahwa  $MAPE = \frac{\sum \frac{|Y-Y'|}{Y} \times 100\%}{n}$  yaitu tingkat kesalahan sebesar  $\frac{167,92\%}{255} = 5,60\%$  sehingga nilai tingkat keakurasian dapat dihitung menggunakan rumus sebagai berikut :

$$Akurasi = 100 \% - error$$

$$Akurasi = 100 \% - 5,60\%$$

$$Akurasi = 94,60\%$$

Nilai akurasi menggunakan metode regresi linier sebesar 94,60% untuk prediksi jumlah kunjungan wisatawan di Kebumen sehingga dapat digunakan untuk memprediksi pada tahun berikutnya untuk meyakinkan investor disektor pariwisata oleh pemerintah daerah.

## SIMPULAN

Berdasarkan hasil penelitian yang dilakukan menggunakan metode regresi linier sederhana untuk memprediksi jumlah kunjungan wisatawan di Kebumen diperoleh hasil tingkat akurasi sebesar 94,60%. Hasil prediksi kunjungan wisatawan di Kebumen dengan nilai b negatif maka jumlah kunjungan wisatawan mengalami penurunan meskipun pada aktualnya data kunjungan wisatawan mengalami kenaikan.

## REFERENSI

- Alumtazah, N., Azizah, N., Putri, Y. L., & Novitasari, D. C. R. (2021). *Prediksi Jumlah Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana*. *Jurnal Ilmiah Matematika Dan Terapan*, 18(1), 31–40.
- Alwi, M. I., Sasana, H., Jalunggono, G., & Ekonomi, F. (2019). *Analysis Of The Effect Of Hotel Occupancy Rates, Number Of Tourists, And Number Of Tourist Objects On The Income Of The Tourism Sector In Kebumen District*. *Directory Journal of Economic*, 1(3), 294–306.
- Baihaqi, W. M., Dianingrum, M. R., & Kurnia A. N. (2019). *Regresi Linier Sederhana Untuk Memprediksi Kunjungan Pasien Di Rumah Sakit Berdasarkan Jenis Layanan Dan Umur Pasien*. *Jurnal SIMETRIS*, 10(2), 671–680.
- Maulana, A., Martanto, M., & Ali, I. (2024). *Prediksi Hasil Produksi Panen Bawang Merah Menggunakan Metode Regresi Linier Sederhana*. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(4), 2884–2888.
- Muhartini, A. A., Sahroni, O., Dwi Rahmawati, S., Febrianti, T., Mahuda, I. (2021). *Analisis Peramalan Jumlah Penerimaan Mahasiswa Baru Dengan Menggunakan Metode Regresi Linear Sederhana*. *Jurnal Bayesian*, 1(1), 17–23.