

IMPLEMENTASI ARIMA DALAM *FORECASTING* PENGELUARAN BUKU BPKB DI POLRES SUKABUMI MENGGUNAKAN METODE CRISP-DM

Fahmi Imam Taufik^{1*}, Iwan Rizal Setiawan², Prajoko³

^{1,2}Universitas Muhammadiyah Kota Sukabumi, Jl.Syamsudin. SH No.50 S.H, Cikole, Kec. Cikole, Kota Sukabumi, Jawa Barat 43113, Telp. (0266) 218345

Keywords:

Forecasting, Time Series, ARIMA, CRISP-DM, BPKB.

Correspondent Email:

fahmiimamtaufik@ummi.ac.id



Copyright © 2026 The Author(s), Published by Jurnal Informatika dan Teknik Elektro Terapan. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Abstrak. Buku Pemilik Kendaraan Bermotor (BPKB) merupakan dokumen resmi yang diterbitkan oleh Kepolisian Negara Republik Indonesia sebagai bukti sah kepemilikan kendaraan bermotor. Pengelolaan data pengeluaran BPKB memiliki peran penting dalam mendukung perencanaan kebutuhan material BPKB. Namun, data pengeluaran BPKB di Polres Sukabumi masih lebih banyak dimanfaatkan sebagai laporan administrasi dan belum digunakan secara optimal untuk memprediksi kebutuhan pada periode mendatang. Penelitian ini bertujuan menerapkan metode *Autoregressive Integrated Moving Average* (ARIMA) untuk melakukan *Forecasting* pengeluaran BPKB berdasarkan data historis bulanan periode 2021–2025 dengan metodologi CRISP-DM yang meliputi *Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment*. Tahap pemodelan dilakukan melalui uji *stasioneritas Augmented Dickey-Fuller* (ADF), *differencing*, analisis ACF dan PACF, serta pemilihan model berdasarkan nilai AIC dan BIC. Hasil penelitian menunjukkan bahwa model ARIMA(2,1,0) merupakan model terbaik untuk memprediksi pengeluaran BPKB. Evaluasi menggunakan MAE, RMSE, dan MAPE menunjukkan model memiliki tingkat akurasi yang baik sehingga dapat digunakan untuk mendukung perencanaan kebutuhan BPKB dan pengelolaan persediaan secara lebih efektif dan efisien di Polres Sukabumi.

Abstract. *The Motor Vehicle Ownership Book (BPKB) is an official document issued by the Indonesian National Police as legal proof of vehicle ownership. Managing BPKB issuance data is essential for planning BPKB material requirements. However, at Sukabumi Regional Police (Polres Sukabumi), BPKB issuance data are primarily used for administrative reporting and have not been optimally utilized to predict future demand. This study aims to apply the Autoregressive Integrated Moving Average (ARIMA) method to Forecast monthly BPKB issuance based on historical data from 2021 to 2025 using the CRISP-DM methodology, which consists of Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. The Modelling process includes stationarity testing using the Augmented Dickey-Fuller (ADF) test, differencing, ACF and PACF analyses, and model selection based on AIC and BIC. The results show that ARIMA(2,1,0) is the best model for Forecasting BPKB issuance. Evaluation using MAE, RMSE, and MAPE indicates that the model achieves good Forecasting accuracy and can support BPKB requirement planning and inventory management more effectively and efficiently at Sukabumi Regional Police*

1. PENDAHULUAN

Buku Pemilik Kendaraan Bermotor (BPKB) adalah dokumen resmi yang

dikeluarkan oleh Polri sebagai tanda bukti kepemilikan kendaraan bermotor yang sah secara hukum. Dokumen ini memegang peran

strategis dalam sistem registrasi dan identifikasi kendaraan, baik sebagai bukti kepemilikan maupun penunjang administrasi. Proses penerbitan BPKB mencakup serangkaian tahapan mulai dari pendaftaran, verifikasi berkas, pemeriksaan fisik kendaraan, entri data, hingga penerbitan dokumen, sebagaimana diatur dalam Peraturan Kepolisian Nomor 7 Tahun 2021. Keberadaan sistem ini ditujukan untuk mendorong pelayanan publik yang lebih cepat, mudah, dan efisien [1].

Data BPKB yang dikumpulkan secara periodik memegang peran strategis dalam menunjang administrasi kendaraan bermotor. Pengelolaan data tersebut mencakup fungsi registrasi, identifikasi, dan pengawasan kendaraan sehingga pelayanan kepada masyarakat dapat berjalan lebih efektif dan efisien. Pertumbuhan jumlah kendaraan bermotor yang terus meningkat turut mendorong meningkatnya kebutuhan pengelolaan data administrasi BPKB [2].

Pemanfaatan data BPKB saat ini masih terbatas pada keperluan pelaporan administratif, sehingga potensinya untuk analisis maupun prediksi kebutuhan BPKB pada periode mendatang belum dioptimalkan. Kondisi ini mendorong perlunya pengembangan sistem pengolahan data yang lebih baik agar monitoring dan perencanaan kebutuhan BPKB dapat dilaksanakan secara lebih efektif dan efisien [3].

Forecasting merupakan pendekatan ilmiah untuk memperkirakan kondisi masa depan secara sistematis berdasarkan data historis, sekaligus menjadi dasar perencanaan, pengawasan, dan pengambilan keputusan [4]. Apabila data dikumpulkan secara berurutan dalam rentang waktu tertentu, pendekatan ini dikategorikan sebagai peramalan *Time Series*. Salah satu metode *Time Series* yang umum digunakan adalah ARIMA (*Autoregressive Integrated Moving Average*) [5].

Penelitian yang membandingkan ARIMA dan *Exponential Smoothing* untuk peramalan penjualan semen di PT. Sinar Abadi menunjukkan bahwa ARIMA lebih unggul karena menghasilkan nilai MSE yang lebih kecil [6]. Selain itu, meskipun metode Holt-Winter's menunjukkan performa baik pada data pelatihan, ARIMA terbukti lebih unggul

dalam memprediksi data uji dengan MAPE sebesar 4,9% dibandingkan Holt-Winter's sebesar 6,54% [7].

Penelitian yang membandingkan model ARIMA dan LSTM dalam peramalan harga saham menunjukkan bahwa ARIMA lebih baik dengan nilai rata-rata RMSE sebesar Rp198,62 dan MAPE sebesar 1,79% [8]. Selain itu, perbandingan metode FTS dan ARIMA dalam meramalkan inflasi Indonesia menunjukkan ARIMA(3,0,0) lebih unggul dengan MSE 0,009 dan MAPE 64,987% dibandingkan FTS dengan MSE 0,047 dan MAPE 132,548% [9].

ARIMA adalah metode analisis *Time Series* yang terdiri dari tiga komponen utama yaitu *Autoregressive* (AR), *Integrated* (I), dan *Moving Average* (MA). ARIMA dinilai efektif dalam menangani data yang bersifat fluktuatif dan tersusun secara periodik, serta fleksibel untuk berbagai jenis *Time Series* sehingga mampu menghasilkan prediksi jangka pendek maupun menengah dengan tingkat akurasi yang baik [10].

Berdasarkan karakteristik data pengeluaran BPKB di Polres Sukabumi yang menunjukkan fluktuasi dari waktu ke waktu, metode ARIMA dipilih karena kemampuannya menganalisis pola data historis. Penelitian ini mengangkat judul "*Implementasi ARIMA dalam Forecasting Pengeluaran Buku Pemilik Kendaraan Bermotor (BPKB) Menggunakan Metode CRISP-DM*" dengan harapan hasil peramalan dapat mendukung perencanaan dan pengambilan keputusan di Polres Sukabumi.

2. TINJAUAN PUSTAKA

2.1 *Forecasting*

Forecasting adalah metode yang digunakan untuk memperkirakan kejadian pada masa mendatang berdasarkan data historis yang tersedia. *Forecasting* memiliki peranan penting dalam membantu pengambilan keputusan, perencanaan kebutuhan, serta monitoring data pada berbagai bidang [5].

2.2 *Time Series*

Time Series atau deret waktu adalah kumpulan data yang disusun secara berurutan berdasarkan periode waktu tertentu, seperti harian, mingguan, bulanan, atau tahunan. Dalam analisis *Time Series*, data historis dikaji

untuk mengetahui hubungan antar periode sebagai landasan peramalan [9].

2.3 ARIMA

ARIMA atau metode Box-Jenkins merupakan salah satu metode analisis deret waktu yang banyak digunakan dalam peramalan karena kemampuannya menghasilkan prediksi yang akurat berdasarkan pola data historis. Model ARIMA dinotasikan dalam bentuk $ARIMA(p,d,q)$ di mana p adalah orde *Autoregressive* (AR), d adalah tingkat *differencing*, dan q adalah orde *Moving Average* (MA) [12].

2.4 Analisis ACF dan PACF

Dalam metode ARIMA, penentuan model dilakukan menggunakan analisis grafik ACF dan PACF. ACF digunakan untuk melihat korelasi antara data pada periode saat ini dengan data pada periode sebelumnya, sedangkan PACF digunakan untuk melihat hubungan murni antar data setelah pengaruh lag sebelumnya dihilangkan [16].

2.5 CRISP-DM

CRISP-DM (*Cross-Industry Standard Process for Data Mining*) merupakan kerangka kerja standar yang digunakan dalam proses *data mining*. Di Indonesia, CRISP-DM telah ditetapkan sebagai acuan Standar Kompetensi Kerja Nasional melalui Keputusan Menteri Ketenagakerjaan No. 299 Tahun 2020. Metodologi ini terdiri dari enam tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment* [13].

2.6 Penelitian Terdahulu

Berdasarkan beberapa penelitian terdahulu, metode *Autoregressive Integrated Moving Average* (ARIMA) terbukti efektif digunakan untuk melakukan *Forecasting* data *Time Series* melalui tahapan uji *stasioneritas*, *differencing*, analisis *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF), serta evaluasi menggunakan MAE, RMSE, dan MAPE [4]–[6], [9], [10], [12], [14]–[16]. Selain itu, metodologi CRISP-DM banyak digunakan sebagai kerangka kerja dalam pengembangan model prediksi karena menyediakan tahapan yang sistematis [13]. Namun, penerapan metode ARIMA pada data

pengeluaran Buku Pemilik Kendaraan Bermotor (BPKB) di lingkungan Kepolisian masih sangat terbatas. Oleh karena itu, penelitian ini menerapkan metode ARIMA dengan pendekatan CRISP-DM untuk melakukan *Forecasting* pengeluaran BPKB di Polres Sukabumi sebagai pendukung perencanaan kebutuhan material BPKB.

3. METODE PENELITIAN

Penelitian ini menggunakan metodologi CRISP-DM sebagai kerangka kerja penelitian. Dalam penelitian ini, metode ARIMA digunakan pada tahap pemodelan untuk melakukan *Forecasting* penggunaan BPKB di Polres Sukabumi berdasarkan data historis periode 2021–2025 [13].

Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Metode CRISP-DM

3.1 Business Understanding

Tahap *Business Understanding* bertujuan untuk mengidentifikasi permasalahan yang terjadi dalam pemilihan ukuran pakaian pada pembelian secara daring. Permasalahan yang dihadapi adalah belum tersedianya metode prediksi yang dapat memperkirakan kebutuhan BPKB pada periode mendatang. Data penggunaan BPKB selama ini hanya digunakan sebagai laporan administrasi berkala dan belum dimanfaatkan untuk prediksi kebutuhan, sehingga perencanaan masih dilakukan berdasarkan estimasi petugas dan berpotensi menimbulkan kekurangan atau kelebihan persediaan.

3.2 Data Understanding

Tahap *Data Understanding* dilakukan untuk memahami karakteristik dataset yang digunakan. Data yang digunakan merupakan data sekunder berupa catatan historis pengeluaran BPKB periode Januari 2021 hingga Desember 2025, yang diperoleh dari

bagian administrasi BPKB Polres Sukabumi. Data disusun dalam bentuk deret waktu bulanan berdasarkan rekapitulasi jumlah pengeluaran BPKB setiap bulan, dengan total 60 data observasi.

3.3 Data Preparation

Tahap *Data Preparation* dilakukan melalui proses pembersihan data, penghapusan data duplikat, Proses persiapan data dilakukan menggunakan *Python* dengan pustaka *Pandas* pada *Google Colaboratory*, meliputi pembagian data menjadi data training dan data testing, pengujian *stasioneritas* menggunakan *Augmented Dickey-Fuller* (ADF) Test, serta proses *differencing* apabila diperlukan. Data dibagi menjadi 80% data training (48 data, periode Januari 2021–Desember 2024) dan 20% data testing (12 data, periode Januari 2025–Desember 2025).

3.4 Modeling

Tahap *Modelling* dilakukan melalui langkah-langkah: (1) Uji *stasioneritas* menggunakan ADF; (2) *Differencing* jika diperlukan; (3) Analisis ACF dan PACF untuk menentukan dugaan awal nilai p dan q ; (4) Pencarian parameter optimal menggunakan *Auto ARIMA* serta pembentukan beberapa kandidat model $ARIMA(p,d,q)$; (5) Pemilihan model terbaik berdasarkan nilai AIC dan BIC terkecil. Pemodelan dilakukan menggunakan pustaka *Statsmodels* pada *Python* di lingkungan *Google Colaboratory*.

3.5 Evaluation

Evaluasi model dilakukan dengan menghitung nilai MAE, RMSE, dan MAPE antara hasil prediksi dan data aktual pada periode testing. Model dianggap baik apabila menghasilkan nilai kesalahan yang rendah. Hasil perhitungan MAPE diinterpretasikan berdasarkan kriteria: nilai MAPE < 10% (sangat baik), 10%–20% (baik), 20%–50% (cukup), dan > 50% (sangat buruk) [18].

3.6 Deployment

Tahap *Deployment* dilakukan dengan mengimplementasikan model *XGBoost* ke dalam aplikasi berbasis web menggunakan framework *Gradio*. Aplikasi kemudian dipublikasikan melalui *Hugging Face Spaces*, sehingga sistem dapat diakses melalui browser

dan memberikan rekomendasi ukuran pakaian secara *real-time* berdasarkan data antropometri yang dimasukkan oleh pengguna.

4. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi metode *ARIMA* untuk *Forecasting* pengeluaran BPKB di Polres Sukabumi menggunakan metodologi *CRISP-DM* yang mencakup enam tahapan: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*.

4.1 Business Understanding

Berdasarkan hasil observasi dan wawancara dengan bagian administrasi BPKB Polres Sukabumi, diketahui bahwa data pengeluaran BPKB selama ini hanya dimanfaatkan sebagai laporan administrasi dan belum digunakan untuk memprediksi kebutuhan periode berikutnya. Kondisi tersebut menyebabkan perencanaan kebutuhan material BPKB masih bergantung pada estimasi petugas. Oleh karena itu, penelitian ini menerapkan metode *Autoregressive Integrated Moving Average* (ARIMA) untuk melakukan *Forecasting* berdasarkan data bulanan pengeluaran BPKB periode 2021–2025. Kinerja model dievaluasi menggunakan *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE), sedangkan tingkat akurasi model berdasarkan nilai MAPE ditunjukkan pada Tabel 1.

Tabel 1. Kriteria Akurasi Model Berdasarkan Nilai MAPE

Rentang Nilai MAPE	Kategori Akurasi
< 10%	Sangat Baik
10% – 20%	Baik
> 20%	Perlu Perbaikan

4.2 Data Understanding

Data yang digunakan merupakan data sekunder berupa catatan historis pengeluaran BPKB periode Januari 2021 hingga Desember 2025 yang diperoleh dari bagian administrasi BPKB Polres Sukabumi. Data disusun dalam bentuk deret waktu bulanan berdasarkan

rekapitulasi jumlah pengeluaran BPKB setiap bulan, dengan total 60 data observasi. Informasi dataset disajikan pada Tabel 2.

Tabel 2. Informasi Dataset

Keterangan	Nilai
Periode Data	Januari 2021 – Desember 2025
Jenis Data	<i>Time Series</i>
Sumber Data	Administrasi BPKB Polres Sukabumi
Variabel Target	Pengeluaran/Penggunaan BPKB
Frekuensi Data	Bulanan
Jumlah Data	60 Data

Data yang digunakan dalam penelitian ini adalah data jumlah pengeluaran BPKB bulanan selama periode Januari 2021 sampai dengan Desember 2025, sehingga total terdapat 60 observasi. Statistik deskriptif dari data tersebut disajikan pada Tabel 3.

Tabel 3. Statistik Deskriptif Pengeluaran BPKB Tahun 2021–2025

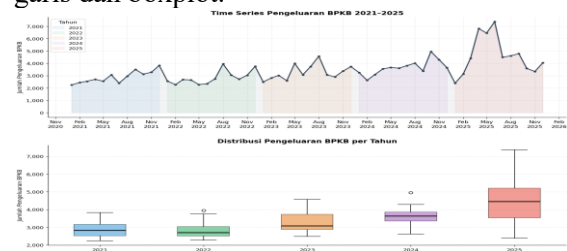
Statistik	Nilai
Jumlah Observasi	60 bulan
Nilai Minimum	2.244
Nilai Maksimum	7.369
Rata-rata	3.455,3
Standar Deviasi	1.054,2
Median	3.191,5

Statistik deskriptif per tahun disajikan pada Tabel 4 untuk melihat perkembangan pola pengeluaran BPKB dari tahun ke tahun.

Tabel 4. Statistik Deskriptif Pengeluaran BPKB per Tahun

Tahun	Total	Rata-rata	Min	Maks	Std. Dev
2021	34.632	2.886	2.244	3.827	488
2022	34.025	2.835	2.270	3.948	539
2023	39.376	3.281	2.490	4.577	620
2024	43.872	3.656	2.623	4.954	594
2025	55.412	4.618	2.389	7.369	1.540

Berdasarkan Tabel 4, rata-rata pengeluaran BPKB menunjukkan tren meningkat setiap tahun, dengan peningkatan paling signifikan pada tahun 2025. Nilai standar deviasi yang lebih tinggi pada tahun 2025 mengindikasikan fluktuasi data yang semakin besar. Pola deret waktu dan distribusi data per tahun ditampilkan pada Gambar 1 melalui grafik garis dan boxplot.



Gambar 1. *Time Series* dan Distribusi Pengeluaran BPKB Tahun 2021–2025

Dari Gambar 1 dapat diamati bahwa data menunjukkan pola tren yang meningkat, terutama pada tahun 2025, disertai dengan fluktuasi musiman dan variabilitas yang membesar pada periode pertengahan tahun 2025. Pola ini mengindikasikan bahwa data belum bersifat *stasioner* sehingga perlu dilakukan pengujian *stasioneritas* dan penanganan lebih lanjut pada tahap *Data Preparation*.

4.3 Data Preparation

Proses persiapan data dilakukan menggunakan *Python* dengan pustaka *Pandas* pada *Google Colab*. Data dibagi terlebih dahulu menjadi data training (80%) dan data testing (20%) sebagaimana disajikan pada Tabel 4, sehingga seluruh tahap identifikasi model hanya memanfaatkan data training dan data testing berperan sebagai data yang belum

pernah dilihat oleh model (*unseen data*). Selanjutnya dilakukan pengujian *stasioneritas* menggunakan *Augmented Dickey-Fuller* (ADF) Test terhadap data training. Hasil pengujian ditampilkan pada Tabel 5.

Tabel 5. Pembagian Data Training dan Testing

Selanjutnya Uji *stasioneritas* dilakukan menggunakan *Augmented Dickey-Fuller*

Keterangan	Jumlah	Periode
Total Data	60 bulan	Januari 2021 – Desember 2025
Data Training (80%)	48 bulan	Januari 2021 – Desember 2024
Data Testing (20%)	12 bulan	Januari 2025 – Desember 2025

(ADF) Test pada data training. Hipotesis pengujian adalah H_0 : data tidak *stasioner*, dan H_1 : data *stasioner*. Data dinyatakan *stasioner* apabila nilai $p\text{-value} < 0,05$. Hasil pengujian disajikan pada Tabel 6.

Tabel 1. Hasil Uji *Augmented Dickey-Fuller* (ADF) pada Data Training

Berdasarkan Tabel 4.6, data *training* memiliki nilai $p\text{-value}$ sebesar 0,0006 ($< 0,05$)

Data	ADF Statistic	$p\text{-value}$	Kesimpulan
Data Training (Asli)	-4,2406	0,0006	<i>Stasioner</i>
Data Training Differencing d=1	-8,9009	0,0000	<i>Stasioner</i>

sehingga telah memenuhi asumsi stasioner. Namun, karena masih terlihat adanya tren pada Gambar 4.1, dilakukan *differencing* satu kali ($d = 1$). Hasilnya, nilai $p\text{-value}$ menjadi 0,0000, sehingga data dinyatakan stasioner dan

digunakan pada pemodelan ARIMA selanjutnya.



Gambar 2. Perbandingan Data Training Sebelum dan Sesudah *Differencing*

Tahap berikutnya adalah Identifikasi orde *Autoregressive* (p) dan *Moving Average* (q) dilakukan dengan mengamati pola plot ACF dan PACF pada data training, baik sebelum maupun sesudah *differencing*. Hasil plot disajikan pada Gambar 3.



Gambar 3. Perbandingan Data Training Sebelum dan Sesudah *Differencing*

Berdasarkan Gambar 3, plot ACF pada data *training* asli menunjukkan pola *slowly decaying* yang mengindikasikan perlunya proses *differencing*. Setelah dilakukan *differencing* satu kali, pola ACF dan PACF digunakan untuk mengidentifikasi kandidat awal model ARIMA. Selanjutnya, beberapa kandidat model dievaluasi berdasarkan nilai AIC dan BIC untuk memperoleh model terbaik.

4.4 Modeling

Tahap Modeling bertujuan untuk menentukan kombinasi parameter ARIMA (p, d, q) terbaik menggunakan data *training*. Pencarian parameter optimal dilakukan dengan algoritma Auto ARIMA menggunakan rentang $p = 0-4$, $q = 0-4$, dan $d = 1$ sesuai hasil uji *stasioneritas*. Model terbaik dipilih

berdasarkan nilai Akaike Information Criterion (AIC) terendah. Hasil pencarian menunjukkan bahwa ARIMA(2,1,0) merupakan model terbaik dengan nilai AIC sebesar 733,7821 dan BIC sebesar 741,1827. Selanjutnya, untuk memastikan model yang dipilih benar-benar optimal, dilakukan perbandingan terhadap sembilan kandidat model ARIMA yang berasal dari hasil identifikasi ACF/PACF dan rekomendasi Auto ARIMA. Seluruh kandidat diestimasi menggunakan data *training* dan dievaluasi berdasarkan nilai AIC, BIC, serta Log-Likelihood, sebagaimana disajikan pada Tabel 7.

Tabel 7. Perbandingan Model Kandidat ARIMA

Model ARIMA	AIC	BIC	Log-Likelihood
(2,1,0)	732,8971	738,4476	-363,4486
(0,1,1)	733,3107	737,0110	-364,6553
(0,1,2)	734,1285	739,6789	-364,0642
(1,1,2)	734,2666	741,6672	-363,1333
(2,1,1)	734,4994	741,8999	-363,2497
(1,1,1)	734,6086	740,1591	-364,3043
(2,1,2)	735,7884	745,0391	-362,8942
(1,1,0)	741,8904	745,5907	-368,9452
(0,1,0)	745,6182	747,4683	-371,8091

Berdasarkan Tabel 7. model ARIMA(2,1,0) memiliki nilai AIC terendah sebesar 732,8971, sehingga dipilih sebagai model terbaik. Hasil ini konsisten dengan rekomendasi Auto ARIMA, yang menunjukkan bahwa model tersebut merupakan model paling optimal untuk data *training*. Selanjutnya, model ARIMA(2,1,0) diestimasi menggunakan data *training* sebanyak 48 bulan untuk memperoleh nilai parameter model. Ringkasan hasil estimasi parameter disajikan pada Tabel 8.

Tabel 8. Estimasi Parameter Model ARIMA(2,1,0) Data training

Parameter	Koefisien	Std. Error	z	P> z
AR(1)	-0,5141	0,147	-3,499	0,000
AR(2)	-0,4572	0,168	-2,716	0,007
sigma2	307.300	61.600	4,989	0,000

Tabel 4.8 menunjukkan bahwa kedua parameter *Autoregressive*, yaitu AR(1) dan AR(2), memiliki nilai *p-value* yang lebih kecil dari 0,05 (masing-masing 0,000 dan 0,007), sehingga kedua parameter tersebut signifikan secara statistik dan layak dipertahankan dalam model. Nilai AIC model pada data training adalah sebesar 732,90 dengan BIC sebesar 738,45.

4.5 Evaluation

Tahap Evaluation bertujuan untuk menilai kinerja model ARIMA(2,1,0) menggunakan data *testing* (*out-of-sample*) serta menguji kelayakan model melalui diagnostik *Residual*. Model yang telah diestimasi menggunakan data *training* kemudian digunakan untuk memprediksi data *testing* selama 12 bulan (Januari–Desember 2025). Hasil prediksi selanjutnya dibandingkan dengan data aktual untuk mengukur tingkat akurasi model, sebagaimana disajikan pada Tabel 9.

Periode	Aktual	Prediksi	Error	APE (%)
Jan 2025	2.389	4.284	-1.895,1	79,33
Feb 2025	3.130	4.251	-1.121,4	35,83
Mar 2025	4.414	3.980	434,3	9,84
Apr 2025	6.819	4.134	2.684,7	39,37
Mei 2025	6.450	4.179	2.271,0	35,21
Jun 2025	7.369	4.085	3.283,7	44,56
Jul 2025	4.491	4.113	378,0	8,42
Agt 2025	4.606	4.142	464,3	10,08
Sep 2025	4.785	4.114	670,7	14,02
Okt 2025	3.588	4.115	-527,3	14,70
Nov 2025	3.336	4.127	-791,3	23,72
Des 2025	4.035	4.121	-85,6	2,12

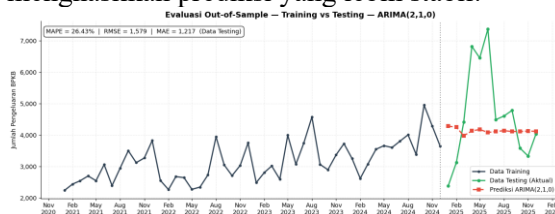
Tabel 9. Perbandingan Nilai Aktual dan prediksi pada data testing

Berdasarkan hasil pada Tabel 9, dihitung tiga ukuran akurasi peramalan, yaitu *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE), dan *Root Mean Squared Error* (RMSE), yang dirangkum pada Tabel 10.

Ukuran Akurasi	Nilai	Kategori
MAE	1.217,28	-
MAPE	26,43%	Perlu Perbaikan
RMSE	1.578,63	-

Tabel 10. Hasil Evaluasi Akurasi Model Pada Data Testing

Mengacu pada kriteria akurasi pada Tabel 1, nilai MAPE sebesar 26,43% termasuk dalam kategori "Perlu Perbaikan" karena melebihi batas 20%. Nilai tersebut dipengaruhi oleh lonjakan pengeluaran BPKB pada periode April–Juni 2025 yang berbeda secara signifikan dari pola historis, sehingga model ARIMA(2,1,0) kurang mampu mengikuti perubahan tersebut dan cenderung menghasilkan prediksi yang lebih stabil.



Gambar 4. Perbandingan Data Training, Data Testing dan hasil prediksi Model

Selanjutnya, dilakukan diagnostik *Residual* untuk memastikan bahwa model ARIMA(2,1,0) telah menangkap pola data dengan baik. Pengujian meliputi uji autokorelasi *Residual* menggunakan *Ljung-Box* serta uji normalitas *Residual* menggunakan Jarque-Bera dan Shapiro-Wilk pada *Residual* hasil *fitting* data *training*. Hasil uji *Ljung-Box* disajikan pada Tabel 11.

Lag	LB Statistic	<i>p-value</i>	Kesimpulan
6	3,4492	0,7507	<i>White noise</i>
12	5,1575	0,9525	<i>White noise</i>
18	11,2918	0,8816	<i>White noise</i>
24	16,6927	0,8615	<i>White noise</i>

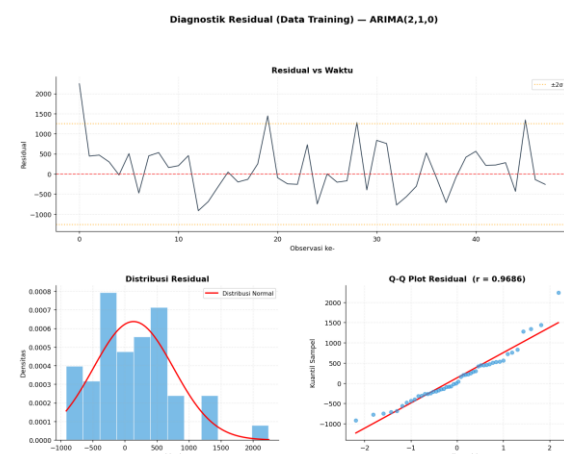
Tabel 11. Hasil Uji *Ljung-Box Residual* Model

Berdasarkan Tabel 11, seluruh nilai *p-value* pada berbagai *lag* lebih besar dari 0,05, sehingga *Residual* model ARIMA(2,1,0) memenuhi asumsi *white noise* dan tidak mengandung autokorelasi yang signifikan. Selanjutnya, normalitas *Residual* diuji menggunakan Jarque-Bera dan Shapiro-Wilk, sebagaimana disajikan pada Tabel 12.

Tabel 12. Hasil Uji Normalitas *Residual*

Uji Normalitas	Statistik	<i>p-value</i>	Kesimpulan
Jarque-Bera	11,4872	0,0032	Tidak Normal
Shapiro-Wilk	0,9434	0,0219	Tidak Normal

Berdasarkan Tabel 12, *Residual* model tidak berdistribusi normal karena nilai *p-value* pada uji Jarque-Bera dan Shapiro-Wilk lebih kecil dari 0,05. Kondisi ini dipengaruhi oleh adanya *outlier* akibat lonjakan pengeluaran BPKB pada pertengahan tahun 2025. Meskipun demikian, karena *Residual* telah memenuhi asumsi *white noise* berdasarkan uji *Ljung-Box*, model ARIMA(2,1,0) tetap layak digunakan untuk *Forecasting*. Visualisasi diagnostik *Residual* disajikan pada Gambar 5.

Gambar 5. Diagnostik *Residual* Model ARIMA(2,1,0): *Residual* terhadap Waktu, Histogram, dan Q-Q Plot

Gambar 5 menunjukkan adanya beberapa *Residual* ekstrem pada pertengahan tahun 2025. Selain itu, histogram *Residual* dan Q-Q Plot memperlihatkan penyimpangan dari

distribusi normal, sehingga memperkuat hasil uji normalitas pada Tabel 12.

4.6 Deployment

Tahap Deployment merupakan tahap akhir dalam metodologi CRISP-DM, yaitu menerapkan model yang telah divalidasi untuk menghasilkan *Forecasting* periode mendatang. Setelah model ARIMA(2,1,0) dinyatakan layak pada tahap evaluasi, model di-*refit* menggunakan seluruh 60 data historis agar estimasi parameter menjadi lebih optimal. Hasil *refit* menghasilkan nilai AIC sebesar 964,90 dan BIC sebesar 971,13. Selanjutnya, model digunakan untuk melakukan *Forecasting* pengeluaran BPKB selama 12 bulan (Januari–Desember 2026) beserta interval kepercayaan 95%, sebagaimana disajikan pada Tabel 13

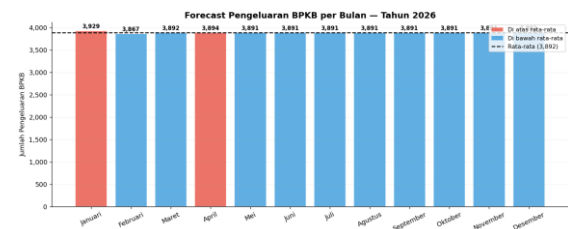
Bulan	Forecast	CI Bawah 95%	CI Atas 95%	Selisih vs Des-2025	% Perubahan
Januari 2026	3.929	2.302	5.556	-106	-2,63%
Februari 2026	3.867	1.778	5.956	-168	-4,16%
Maret 2026	3.892	1.492	6.291	-143	-3,54%
April 2026	3.894	1.192	6.597	-141	-3,49%
Mei 2026	3.891	915	6.867	-144	-3,57%
Juni 2026	3.891	668	7.114	-144	-3,57%
Juli 2026	3.891	438	7.345	-144	-3,57%
Agustus 2026	3.891	222	7.561	-144	-3,57%
September 2026	3.891	18	7.764	-144	-3,57%
Oktober 2026	3.891	-175	7.958	-144	-3,57%
November 2026	3.891	-360	8.143	-144	-3,57%
Desember 2026	3.891	-537	8.320	-144	-3,57%

Tabel 13. Hasil *Forecast* Pengeluaran BPKB Periode Jan-Des 2026

Berdasarkan Tabel 13, total proyeksi pengeluaran BPKB selama tahun 2026 sebesar 46.710 unit dengan rata-rata 3.892 unit per bulan. Proyeksi tertinggi terjadi pada Januari 2026 (3.929 unit) dan terendah pada Februari 2026 (3.867 unit). Secara umum, hasil *Forecasting* menunjukkan pola yang relatif stabil setelah lonjakan pengeluaran pada pertengahan tahun 2025. Visualisasi hasil *Forecasting* dan proyeksi bulanan disajikan pada Gambar 6 dan Gambar 7.



Gambar 7. Hasil *Forecasting* Pengeluaran BPKB Tahun 2026 terhadap Data Aktual 2021–2025



Gambar 8. Diagram Batang Proyeksi Pengeluaran BPKB per Bulan Tahun 2026

Secara keseluruhan, metode ARIMA(2,1,0) mampu memodelkan pola pengeluaran BPKB berdasarkan data historis periode 2021–2025 dan menghasilkan proyeksi pengeluaran tahun 2026. Meskipun evaluasi menunjukkan nilai MAPE sebesar 26,43% sehingga termasuk kategori *Perlu Perbaikan*, model telah memenuhi asumsi *Residual white noise* berdasarkan uji *Ljung-Box*. Nilai MAPE yang relatif tinggi dipengaruhi oleh lonjakan pengeluaran BPKB pada pertengahan tahun 2025 yang tidak mengikuti pola historis. Oleh karena itu, hasil *Forecasting* dapat digunakan sebagai acuan dalam perencanaan kebutuhan BPKB dengan tetap mempertimbangkan faktor

operasional yang dapat memengaruhi perubahan pengeluaran.

5. KESIMPULAN

Penelitian ini berhasil mengimplementasikan metode *Autoregressive Integrated Moving Average* (ARIMA) dengan pendekatan *Cross-Industry Standard Process for Data Mining* (CRISP-DM) untuk melakukan *Forecasting* pengeluaran BPKB di Polres Sukabumi berdasarkan data historis periode 2021–2025. Berdasarkan proses pemodelan, ARIMA(2,1,0) dipilih sebagai model terbaik dengan nilai AIC terendah. Hasil evaluasi pada data *testing* menghasilkan nilai MAE sebesar 1.217,28, RMSE sebesar 1.578,63, dan MAPE sebesar 26,43%. Meskipun nilai MAPE masih berada pada kategori *Perlu Perbaikan* akibat adanya lonjakan pengeluaran BPKB pada pertengahan tahun 2025, model telah memenuhi asumsi *Residual white noise* berdasarkan uji *Ljung-Box* sehingga layak digunakan untuk *Forecasting*. Model selanjutnya digunakan untuk memproyeksikan pengeluaran BPKB tahun 2026 dengan total estimasi sebesar 46.710 unit, sehingga dapat menjadi salah satu acuan dalam perencanaan kebutuhan dan pengelolaan persediaan BPKB di Polres Sukabumi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Allah SWT, dosen pembimbing, Polres Sukabumi yang telah memberikan izin dan menyediakan data penelitian, serta keluarga dan semua pihak yang telah memberikan dukungan, bimbingan, dan motivasi sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] Y. Pratama, Mulyadi, dan Djuniarsono, "Prosedur Peralihan Kepemilikan Buku Pemilik Kendaraan Bermotor Berdasarkan Peraturan Kepolisian Nomor 7 Tahun." *Karimah Tauhid*, vol. 4, no. 8, 2025. <https://doi.org/10.30997/karimahtauhid.v4i8.14770>
- [2] M. H. Lazuardi, "Kebijakan Pajak Kendaraan Bermotor, Dikaji dari Prinsip," *Jurnal Hukum Lingkungan Indonesia*, vol. 7, no. 2, 2021. <https://doi.org/10.38011/jhli.v7i2.317>
- [3] O. Veza dan Sayuti, "Perancangan Dashboard Informasi Target Pajak Kendaraan Bermotor pada Badan Pengelolaan Pajak dan Retribusi Daerah," *Engineering And Technology International Journal*, 2021.
- [4] S. Fauzani dan Rahmi, "Forecasting dengan Metode ARIMA," *Jurnal Ilmu Komputer*, 2023.
- [5] M. A. Rizqullah dan Suseno, "Analisis Forecasting Permintaan Produk Menggunakan Metode ARIMA," *Jurnal Multidisiplin Ilmu Akademik*, 2025. <https://doi.org/10.61722/jmia.v2i6.7312>
- [6] Y. D. Darma, N. R. Wiyanti, dan T. L. Sarumpaet, "Analisis Perbandingan Model ARIMA dan Exponential Smoothing dalam Memprediksi Persediaan," *Jurnal Bisnis Manajemen & Ekonomi*, vol. 19, no. 2, 2021. <https://doi.org/10.33197/jbme.vol19.iss2.2021.824>
- [7] D. R. Wibisono, "Perbandingan Metode Peramalan Berbasis Pola Tren Musiman untuk Prediksi Permintaan Obat dengan Model Holt-Winter's dan ARIMA," *TALENTA Conference Series: Energy and Engineering*, 2024. <https://doi.org/10.32734/ee.v7i1.2177>
- [8] A. D. Milniadi dan N. O. Adiwijaya, "Analisis Perbandingan Model ARIMA dan LSTM dalam Peramalan Harga Penutupan Saham," *SIBATIK Journal*, vol. 2, no. 6, 2023. <https://doi.org/10.54443/sibatik.v2i6.798>
- [9] A. Qalbi, W. Alwi, dan K. Nurfadilah, "Comparison of Fuzzy Time Series Methods and ARIMA for Inflation Data," *Eigen Mathematics Journal*, 2021. <https://doi.org/10.29303/emj.v4i2.122>
- [10] B. A. Wulandari, N. Latifah, dan F. Nugraha, "Implementasi Forecasting Permintaan Produk dengan Metode ARIMA," *Jurnal Sistem Informasi dan Teknologi*, 2025.
- [11] M. Hamdani, D. Primasari, dan H. Fajri, "Sistem Informasi Manajemen Pengarsipan BPKB di Polres Kabupaten Bogor Berbasis Web," *Jurnal INOVA-TIF*, 2021. <https://doi.org/10.32832/inova-tif.v1i2.5447>
- [12] Makkulau, A. T. Ampa, L. O. Saidi, Baharuddin, Amirullah, dan Hartini, "Penggunaan Metode ARIMA untuk Peramalan Data Inflasi di Indonesia," *Arus Jurnal Sains dan Teknologi*, 2024. <https://doi.org/10.57250/ajst.v2i2.978>
- [13] Y. Yudiana, A. Y. Agustina, dan N. Khofifah, "Prediksi Customer Churn Menggunakan Metode CRISP-DM," *IJIEB: Indonesian Journal of Islamic Economics and Business*, vol. 8, no. 1, 2023. <https://doi.org/10.30631/ijoieb.v8i1.1710>
- [14] L. Junaedi, N. Damastuti, Latipah, dan A. Widodo, "Penerapan Metode Seasonal

- ARIMA (SARIMA) untuk Peramalan Penjualan Barang," *JISEM*, vol. 01, no. 01, 2025.
<https://doi.org/10.33508/jisem.v1i01.7403>
- [15] F. D. Ulhaq, A. M. Aziz, M. Hasan, A. Rokhim, dan D. F. Hartatri, "Analisis *Forecasting* Harga Kopi Robusta, Arabika, dan Kakao Indonesia Periode 2024-2026 Model ARIMA," *Jurnal Penelitian Nusantara*, vol. 1, no. 3, 2025.
<https://doi.org/10.59435/menulis.v1i3.84>
- [16] F. Ariyanto dan M. B. Tamam, "Analisis *Time Series* dengan Metode ARIMA dan Aplikasinya," *Jurnal JATIM*, vol. 01, no. 02, 2020. <https://doi.org/10.31102/jatim.v1i2.972>
- [17] L. Hablinawati dan J. Nugraha, "Peramalan Nilai Tukar Petani di Daerah Istimewa Yogyakarta Menggunakan Metode ARIMA," *Emerging Statistics and Data Science Journal*, 2024.
<https://doi.org/10.20885/esds.vol2.iss.1.art9>
- [18] Q. A. Khairunnisa, N. D. Haryadi, dan N. Audyna, "Aplikasi Metode ARIMA dalam Meramalkan Rata-Rata Harga Beras di Indonesia," *Jurnal Agribisnis*, vol. 24, no. 2, 2022. <https://doi.org/10.31849/agr.v24i2.8683>