



# Prediksi Penjualan Bisnis Rumah Properti dengan Menggunakan Metode *Autoregressive Integrated Moving Average* (ARIMA)

Jefri Junifer Pangaribuan\*, Fanny, Okky Putra Barus, dan Romindo

Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pelita Harapan, Medan, Indonesia

Naskah Diterima : 21 Juni 2023; Diterima Publikasi : 1 Oktober 2023

DOI: 10.21456/vol13iss2pp154-161

## Abstract

Sales forecasting plays an important role in determining the company's strategy in the future because it allows control of planning and availability of home production according to consumer needs. Forecasting accuracy provides significant advantages for companies, including production cost savings and avoidance of unnecessary costs. Without accurate forecasting, a company will face difficulties in determining the quantity of house production, which can have a negative impact on the company's financial balance if the houses do not sell. This research implements the Autoregressive Integrated Moving Average (ARIMA) model to forecast property business house sales with a high level of accuracy to support future business decisions. The results of the research on the application of the Autoregressive Integrated Moving Average algorithm show that the ARIMA model (9,1,10) provides good forecasting results measured by the lowest AIC and BIC values compared to the other 4 models, namely ARIMA (10,1,9); ARIMA(8,1,9); ARIMA(10,1,10); and ARIMA (12,1,12) accompanied by an evaluation of measuring the accuracy of the model using RMSE, MSE, and MAPE with each value of 0.281409; 0.079191 and MAPE of 3.4% so that it can be said that sales forecasting provides a good level of accuracy.

**Keywords:** Data Mining; Autoregressive Integrated Moving Average; Forecasting; Sales

## Abstrak

Prediksi penjualan memegang peran penting dalam menentukan strategi perusahaan di masa depan karena memungkinkan pengendalian perencanaan dan ketersediaan produksi rumah sesuai dengan kebutuhan konsumen. Keakuratan prediksi memberikan keuntungan signifikan bagi perusahaan, termasuk penghematan biaya produksi dan menghindari biaya yang tidak perlu. Kesulitan dalam menentukan jumlah produksi rumah tanpa prediksi yang tepat dapat berdampak negatif pada keseimbangan keuangan perusahaan jika rumah tidak terjual. Penelitian ini mengimplementasikan model *Autoregressive Integrated Moving Average* untuk melakukan prediksi penjualan bisnis rumah properti dengan tingkat akurasi yang baik untuk dapat mendukung keputusan bisnis kedepannya. Hasil penelitian pada pengaplikasian algoritma *Autoregressive Integrated Moving Average* menunjukkan bahwa model ARIMA (9,1,10) memberikan hasil nilai prediksi yang baik diukur dari nilai AIC dan BIC yang paling rendah dibandingkan 4 model lainnya yaitu ARIMA (10,1,9); ARIMA (8,1,9); ARIMA (10,1,10); dan ARIMA (12,1,12) disertai evaluasi pengukuran keakuratan model dengan menggunakan RMSE dan MSE masing-masing yaitu 0.281409; 0.079191 serta MAPE sebesar 3.4% sehingga dapat dikatakan prediksi penjualan memberikan tingkat akurasi yang baik.

**Kata Kunci:** *Data Mining; Autoregressive Integrated Moving Average; Prediksi; Penjualan*

## 1. Pendahuluan

Kebutuhan pokok manusia dalam menjalani kehidupan sehari-hari mencakup kebutuhan papan, sandang, dan pangan (Ratningsih, 2017). Rumah tidak hanya dikatakan sebagai sebuah bangunan struktural yang kokoh, namun juga menjadi area berlandung, tempat beristirahat, tempat berbagi kebahagiaan bersama keluarga. Menteri Keuangan Sri Mulyani menyatakan bahwa generasi muda Republik Indonesia memiliki ancaman tidak dapat membeli rumah karena meningkatnya inflasi yang diikuti dengan kenaikan

suku bunga dari bank sentral (Catur Putri dan Junaedi, 2022). Penurunan permintaan akan berdampak sehingga perlu melakukan perencanaan strategis di masa yang akan datang. Salah satu hal krusial dalam suatu perusahaan yang dapat dilakukan analisis terdapat pada data penjualan rumah untuk dapat menentukan keuntungan yang akan diperoleh perusahaan di masa mendatang.

Prediksi penjualan dapat menjadi tolak ukur dalam menentukan strategi perusahaan di masa depan karena dapat mengendalikan perencanaan dan ketersediaan pada produksi rumah agar memenuhi kebutuhan dari

\*) Penulis korespondensi: [jefri.pangaribuan@uph.edu](mailto:jefri.pangaribuan@uph.edu)

konsumen (Ramadhan *et al.*, 2020). Prediksi yang akurat dapat memberikan keuntungan tersendiri bagi perusahaan berupa penghematan biaya dari sisi produksi rumah dan tidak mengeluarkan biaya yang berlebih untuk produksi. Jika prediksi tidak dilakukan maka perusahaan akan mengalami kesulitan dalam menentukan produksi rumah yang jika tidak terjual akan berdampak pada kerugian secara finansial bagi perusahaan (Žunić *et al.*, 2020).

Banyak metode yang dapat digunakan untuk melakukan prediksi, seperti metode ARIMA yang digunakan untuk prediksi jumlah kasus infeksi saluran pernafasan Akut (Saputra dan Rizky, 2019), metode *Naïve Bayes* untuk prediksi *video performance* akun *Youtube* (Leonie dan Pangaribuan, 2020), metode Apriori untuk prediksi barang pada distributor lampu (Damanik dan Poernomo, 2023), metode *Backpropagation* untuk prediksi harga saham (Barus dan Wijaya, 2021), dan metode *Extreme Learning Machine* untuk prediksi penyakit diabetes (Pangaribuan dan Suharjito, 2014). Pada penelitian sebelumnya menyatakan bahwa metode ARIMA dapat memberikan model serta hasil akurasi yang baik berdasarkan pengamatan data historis sedangkan pada metode *Prophet Facebook* cepat dan mudah digunakan tetapi memberikan hasil akurasi yang kurang baik (Menculini *et al.*, 2021). Pada penelitian ini akan mengimplementasikan metode ARIMA karena dapat memberikan hasil yang akurat dengan prediksi jangka pendek.

Rumusan masalah pada penelitian ini adalah bagaimana mengimplementasi metode ARIMA dalam melakukan prediksi penjualan pada bisnis properti dengan hasil yang akurat. Adapun tujuan penelitian ini yaitu dapat mengimplementasi metode ARIMA untuk melakukan prediksi pada penjualan bisnis properti dan melakukan analisis terhadap hasil keakuratan pada metode ARIMA.

## 2. Kerangka Teori

### 2.1. Bisnis Properti

Bisnis properti melibatkan pengembangan, pengelolaan, dan pemasaran properti. Prediksi penjualan penting dalam merencanakan strategi pemasaran dan pengembangan properti yang efektif. Bisnis properti memiliki potensi keuntungan tinggi karena nilai properti meningkat seiring pertumbuhan ekonomi dan permintaan pasar, namun juga melibatkan risiko fluktuasi pasar, persaingan, dan perubahan regulasi. Prediksi membantu pengambilan keputusan strategis dalam menghadapi risiko tersebut (Hassyddiqy dan Hasdiana, 2023).

### 2.2. Data Mining

*Data mining* adalah proses analisis data untuk mengidentifikasi pola dan hubungan yang membantu perusahaan memecahkan masalah bisnis kompleks (Vercellis, 2009). Metode pembelajaran induktif

digunakan untuk mengamati data masa lalu dan menghasilkan keputusan yang akurat. Metodologi *data mining* mencakup *Clustering*, *Classification*, *Regression*, *Association*, dan *Time Series Analysis* (Han *et al.*, 2012).

### 2.3. Time Series

*Time Series* merupakan serangkaian data yang diurutkan berdasarkan waktu. Data *time series* terdiri dari pengamatan yang diambil secara teratur dalam interval waktu yang tetap, seperti harian, bulanan, atau tahunan. Tujuan analisis *time series* adalah untuk mengidentifikasi pola, tren, dan fluktuasi dalam data seiring waktu. Dengan memahami pola dan karakteristik data *time series*, kita dapat membuat prediksi dan prediksi untuk nilai-nilai masa depan, mengidentifikasi faktor-faktor yang mempengaruhi perubahan dalam data, dan membuat keputusan yang lebih baik dalam konteks analisis berbasis waktu (Ramadhan *et al.*, 2020).

### 2.4. Prediksi

Prediksi adalah proses menilai masa depan berdasarkan data historis. Dalam bisnis, prediksi digunakan untuk merencanakan strategi berdasarkan pergerakan penjualan yang diharapkan (Indarwati *et al.*, 2018). Metode prediksi melibatkan analisis data historis untuk mengidentifikasi pola dan tren, seperti *time series analysis*, *regression analysis*, dan *artificial neural networks*. Prediksi digunakan dalam berbagai bidang seperti bisnis, ekonomi, keuangan, dan perencanaan produksi. Contohnya, memprediksi permintaan konsumen, mengoptimalkan persediaan barang, dan mendukung perencanaan dan strategi bisnis (Permatasari *et al.*, 2018).

### 2.5. ARIMA

ARIMA adalah model yang digunakan untuk prediksi deret waktu dengan mengandalkan data masa lalu dan saat ini (Wulandari dan Gernowo, 2019). Model ini memahami tren dan pola dalam data untuk memprediksi nilai masa depan dengan akurasi. ARIMA dapat diterapkan pada berbagai jenis data deret waktu, termasuk yang memiliki tren, musiman, dan variasi acak. ARIMA banyak digunakan dalam bidang ekonomi, keuangan, bisnis, dan industri untuk prediksi dan analisis deret waktu. Komponen dalam model ARIMA terdiri dari *autoregressive* (AR) yang menghubungkan nilai-nilai sebelumnya, *moving average* (MA) yang menghubungkan residual atau kesalahan, dan *integrated* (I) yang melibatkan *differencing* untuk mencapai stasioneritas.

### 2.6. Stasioneritas dan Nonstasioneritas

Stasioneritas dan nonstasioneritas adalah konsep penting dalam ARIMA. ARIMA membutuhkan data yang stasioner, dengan sifat statistik yang konstan dari waktu ke waktu. Ini memungkinkan model ARIMA untuk menangkap pola dasar dan membuat prakiraan

yang akurat. Nonstasioneritas merujuk pada deret waktu yang tidak menunjukkan sifat statistik stasioner. Hal ini sulit untuk membuat prakiraan yang akurat karena pola yang mendasari seringkali berubah seiring waktu. Model ARIMA dapat mengatasi nonstasioneritas dengan melakukan diferensiasi untuk menghilangkan tren atau komponen musiman, meninggalkan deret residual stasioner yang dapat dimodelkan menggunakan proses ARIMA. Terdapat berbagai metode untuk menguji stasioneritas, seperti inspeksi visual dan uji statistik seperti uji *Augmented Dickey-Fuller* (ADF) dan *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS).

### 2.7. Uji Augmented Dickey-Fuller

Uji *Augmented Dickey-Fuller* (ADF) adalah metode statistik yang digunakan untuk menguji keberadaan stasioneritas dalam suatu deret waktu. Uji ini memeriksa apakah deret waktu memiliki akar unit, yang menunjukkan keberadaan tren atau nonstasioneritas. Dengan melakukan uji ADF, kita dapat menentukan apakah perlu melakukan diferensiasi pada deret waktu untuk mencapai stasioneritas. Uji ADF menghasilkan statistik uji dan nilai *p-value*, di mana nilai *p-value* yang rendah menunjukkan bukti yang kuat untuk menolak hipotesis nol dan mengindikasikan keberadaan stasioneritas.

### 2.8. Plot ACF dan PACF

Plot ACF (*Autocorrelation Function*) dan PACF (*Partial Autocorrelation Function*) adalah grafik yang membantu mengidentifikasi model ARIMA yang sesuai untuk analisis deret waktu. Plot ACF menunjukkan korelasi antara nilai dalam deret waktu dengan nilai-nilai sebelumnya, sedangkan plot PACF menunjukkan korelasi yang tidak dapat dijelaskan oleh nilai-nilai di antara kedua titik waktu tersebut. Pola korelasi yang signifikan pada kedua plot ini mengindikasikan adanya pola autokorelasi dalam data. *P-value* yang tinggi menunjukkan kompleksitas model yang lebih besar. Dengan mempertimbangkan kedua plot ini, nilai *p*, *d*, dan *q* dapat ditentukan untuk model ARIMA yang akurat.

### 2.9. Pengukuran Akurasi Prediksi

Pengukuran akurasi prediksi adalah metode untuk evaluasi sejauh mana prediksi yang dibuat sesuai dengan data aktual. Hal ini digunakan untuk mengukur seberapa akurat prediksi tersebut dalam memprediksi nilai masa depan. Terdapat beberapa metrik umum yang digunakan, seperti *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE).

MSE adalah mengukur perbedaan kuadrat rata-rata antara nilai prediksi dan aktual dari variabel target dalam kumpulan data. Untuk perhitungan MSE dapat dilihat pada Persamaan (1).

$$MSE = \frac{1}{n} \sum (Y_i - \hat{Y}_i)^2 \quad (1)$$

RMSE mengukur selisih kuadrat antara nilai prediksi dan aktual untuk setiap pengamatan kemudian mengambil rata-rata dari perbedaan kuadrat ini dan mengambil akar kuadrat dari rata-rata. Adapun rumus RMSE terlihat pada Persamaan (2).

$$RMSE = \sqrt{\frac{1}{n} \sum (Y_i - \hat{Y}_i)^2} \quad (2)$$

MAPE mengukur rata-rata persentase kesalahan absolut dari nilai prediksi dari nilai sebenarnya yang dinyatakan sebagai persentase. Nilai MAPE diperoleh dari Persamaan (3).

$$MAPE = \frac{1}{n} \sum \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| * 100\% \quad (3)$$

Dalam rumus-rumus di atas, *Y* adalah nilai aktual,  $\hat{Y}_i$  adalah nilai prediksi, dan *n* adalah jumlah observasi dalam data. Dengan menggunakan metrik ini dapat mengukur tingkat akurasi prediksi dan mengevaluasi kualitas model dimana semakin kecil nilai MAE, MSE, dan RMSE, serta MAPE yang lebih rendah, semakin akurat prediksi tersebut.

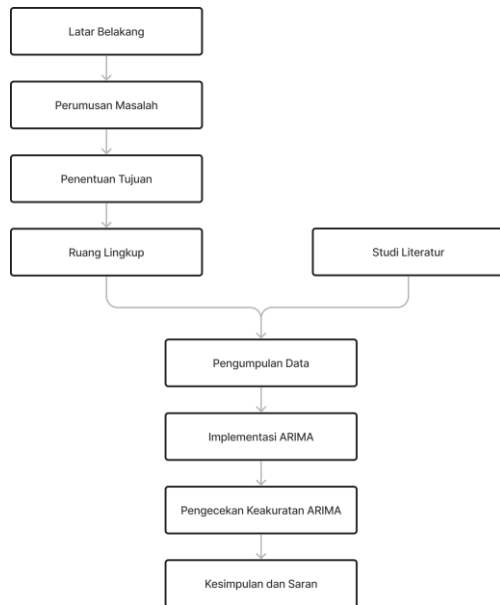
## 3. Metode

### 3.1. Metode Pengumpulan Data

Data yang digunakan diperoleh dari Kaggle pada *website* <https://www.kaggle.com/property-sales>, situs *web open source* yang menyediakan data yang dapat digunakan secara bebas. Dataset yang digunakan dalam penelitian ini adalah *House Property Sales Time Series* yang diperoleh dari perusahaan *HtAG Analytics*. Data ini mencakup penjualan properti rumah dari tahun 2007 hingga 2019 dengan berbagai jenis properti dan lokasi yang berbeda.

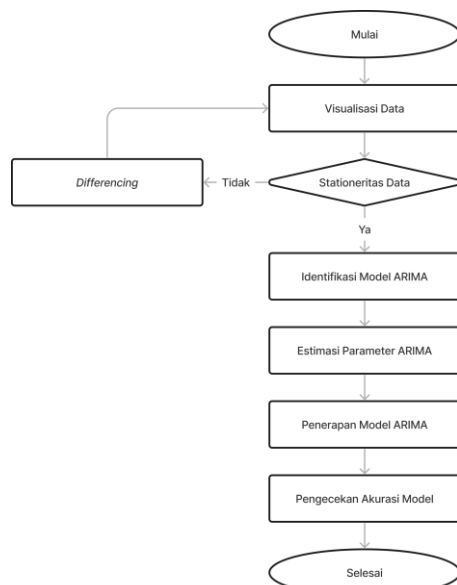
### 3.2. Tahapan Penelitian

Langkah-langkah dalam pembuatan model prediksi dengan ARIMA pada data penjualan rumah properti dapat ditunjukkan pada Gambar 1 meliputi pengumpulan data, implementasi metode ARIMA, analisis hasil prediksi, dan penarikan kesimpulan. Penelitian ini dimulai dengan menentukan latar belakang dan tujuan penelitian, kemudian melakukan studi literatur untuk memahami cara kerja metode ARIMA dan mengukur akurasi prediksi. Selanjutnya, data penjualan rumah properti dikumpulkan, dilakukan uji stasioneritas, dan dilakukan estimasi parameter dalam model ARIMA. Setelah itu, model ARIMA dapat diterapkan dan dapat melakukan pengecekan akurasi model. Langkah terakhir adalah menyimpulkan hasil dan memberikan saran.



Gambar 1. Tahapan Penelitian

Diagram alur penerapan metode ARIMA ditunjukkan pada Gambar 2 yang dimulai dengan memvisualisasikan data waktu untuk melihat pola. Selanjutnya, data dilakukan pengecekan untuk memastikan stationeritas dengan menggunakan metode seperti ADF atau plot autokorelasi dan kepadatan. Jika data tidak stationer, dilakukan *differencing* untuk membuatnya stationer. Setelah itu, plot ACF dan plot PACF digunakan untuk mengidentifikasi model ARIMA yang sesuai. Parameter model diestimasi menggunakan metode MLE. Setelah itu, model ARIMA diterapkan untuk melakukan prediksi menggunakan nilai historis. Di tahap akhir akan dilakukan pengecekan akurasi model dievaluasi dengan membandingkan prediksi dengan data aktual menggunakan metrik evaluasi seperti MSE, RMSE, dan MAPE.



Gambar 2. Diagram Alur Penerapan Metode ARIMA

### 3.3. Metode Penyelesaian

Metode penyelesaian analisis data melibatkan analisis kuantitatif dengan menggunakan metode prediksi ARIMA untuk data penjualan rumah properti. Tahapan penelitian yang dilakukan meliputi:

- 1) Visualisasi data untuk mengidentifikasi tren dan pola dalam data,
- 2) Uji stasioneritas untuk memastikan data stasioner,
- 3) *Differencing* untuk menghilangkan tren atau komponen non-stasioner dalam data,
- 4) Identifikasi model ARIMA menggunakan plot ACF dan PACF,
- 5) Estimasi parameter dalam model ARIMA dengan memilih nilai  $p$ ,  $d$ , dan  $q$  yang sesuai,
- 6) Penerapan model ARIMA pada data deret waktu,
- 7) Pengecekan akurasi model menggunakan metrik evaluasi seperti MSE, MAPE, dan RMSE.

Tahapan ini memberikan informasi tentang seberapa baik model dapat memprediksi nilai masa depan dari deret waktu.

## 4. Hasil dan Pembahasan

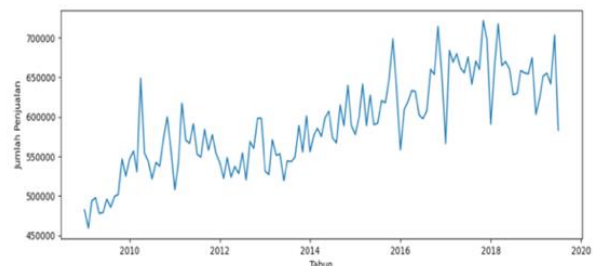
Pada analisa data yang dimulai pada visualisasi data, stationeritas data, *differencing*, melakukan estimasi parameter model ARIMA, melakukan penerapan model ARIMA, dan melakukan pengecekan akurasi model ARIMA yang telah dipilih.

### 4.1. Hasil Penelitian

Analisa data yang dimulai pada visualisasi data, stationeritas data, *differencing*, melakukan estimasi parameter model ARIMA, melakukan penerapan model ARIMA, dan melakukan pengecekan akurasi model ARIMA yang telah dipilih.

#### 1) Visualisasi Data

Dalam melakukan analisis data runtun waktu pada data penjualan rumah properti, tahap awal yang perlu dilakukan adalah melakukan plot data penjualan yang ditunjukkan pada Gambar 3.



Gambar 3. Plot Data Penjualan Rumah Properti

Pada hasil visualisasi data penjualan rumah properti dapat ditunjukkan pada Gambar 3 bahwa grafik menunjukkan bahwa penjualan memiliki tren yang cenderung naik dari waktu ke waktu dan ditunjukkan bahwa data tersebut memiliki flutuatif.

#### 2) Stasioneritas Data

Stasioneritas data pada ARIMA penting karena model ini menggunakan pola dan tren dari data

historis untuk memprediksi nilai masa depan. Jika data tidak stasioner, pola dan tren tersebut mungkin tidak konsisten dan sulit diprediksi. Oleh karena itu, langkah-langkah seperti diferensiasi atau transformasi data sering diperlukan sebelum menerapkan model ARIMA untuk mencapai stasioneritas. Untuk memeriksa stasioneritas data dengan menggunakan Uji *Augmented Dickey-Fuller* (ADF). Hasil dari uji *Dickey-Fuller* dapat ditunjukkan pada Tabel 1.

Tabel 1. *Dickey Fuller Test*

| Hasil Uji                   | Nilai Pengujian |
|-----------------------------|-----------------|
| <i>Test Statistic</i>       | -1.135268       |
| <i>p-value</i>              | 0.700805        |
| <i>Critical Value (1%)</i>  | -3.489590       |
| <i>Critical Value (5%)</i>  | -2.887477       |
| <i>Critical Value (10%)</i> | -2.580604       |

Hasil pengujian *Augmented Dickey-Fuller* pada Tabel 1 menunjukkan bahwa nilai *p-values* sebesar 0.700805. Nilai ini lebih besar dari  $\alpha$  yang bernilai 0.05. Oleh karena itu, hipotesis nol diterima dan hal ini menunjukkan bahwa data tidak bersifat stasioner sehingga data perlu dilakukan differencing karena tidak memiliki rata-rata dan varian yang konstan dari waktu ke waktu (Catur Putri & Junaedi, 2022).

### 3) Differencing

*Differencing* untuk mencapai stasioneritas data dengan menghilangkan pola atau tren non-stasioner dengan mengurangi setiap pengamatan pada waktu  $t$  dengan pengamatan pada waktu  $t-1$ . *Differencing* dapat dilakukan lebih dari satu kali. Pada penelitian ini, dilakukan *differencing* pertama karena data awal tidak stasioner. Setelah *differencing* pertama, dilakukan evaluasi stasioneritas data dengan menggunakan Uji *Augmented Dickey-Fuller* (ADF). Hasil dari uji *Dickey-Fuller* setelah *differencing* dapat ditunjukkan pada Tabel 2.

Tabel 2. *Dickey Fuller Test* setelah melakukan *differencing*

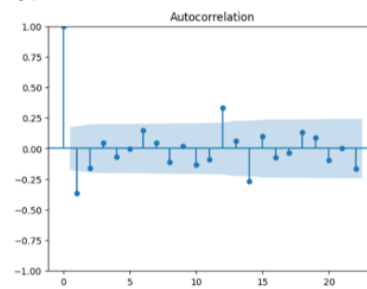
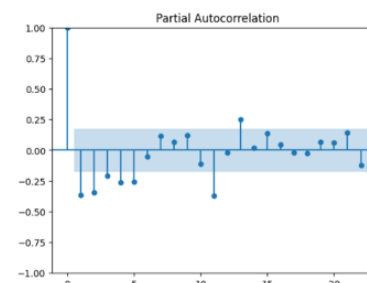
| Hasil Uji                   | Nilai Pengujian |
|-----------------------------|-----------------|
| <i>Test Statistic</i>       | -4.085974       |
| <i>p-value</i>              | 0.001022        |
| <i>Critical Value (1%)</i>  | -3.489590       |
| <i>Critical Value (5%)</i>  | -2.887477       |
| <i>Critical Value (10%)</i> | -2.580604       |

Pengujian ADF pada Tabel 2 menunjukkan bahwa *p-value* sebesar 0.001022, yang lebih kecil dari  $\alpha$  (0.05). Oleh karena itu, hipotesis nol ditolak, menunjukkan bahwa data penjualan rumah properti bersifat stasioner. Data stasioner memiliki statistik yang tetap seiring waktu, tanpa tren, pola musiman, atau fluktuasi yang konsisten. Karena data sudah stasioner, model ARIMA dapat diterapkan secara efektif. Data stasioner juga

cenderung memberikan hasil analisis dan prediksi yang lebih stabil dan akurat (Br Bangun, 2017).

### 4) Identifikasi Model ARIMA

Identifikasi model ARIMA melibatkan analisis plot ACF dan PACF untuk melihat pola korelasi dalam data. Informasi dari plot ini memberikan perkiraan awal nilai  $p$  dan  $q$  untuk model ARIMA. Namun, pemilihan model yang optimal memerlukan metode statistik atau kriteria seleksi model seperti AIC atau BIC. Hal ini penting untuk memastikan model yang dibangun akurat dan sesuai dengan data yang ada. Hasil plot ACF dan PACF dapat ditunjukkan pada Gambar 4 dan Gambar 5.

Gambar 4. *Plot ACF*Gambar 5. *Plot PACF*

Berdasarkan Gambar 4 dan Gambar 5 dapat dilakukan identifikasi model diawal dengan model ARIMA ( $p, d, q$ ) untuk menganalisis data tersebut, di mana  $p$  adalah jumlah lag pada plot PACF yang signifikan,  $q$  adalah jumlah lag pada plot ACF yang signifikan di atas garis putus-putus, dan  $d$  adalah tingkat diferensiasi pada data. Pada Gambar 4 dan Gambar 5 memberikan beberapa model ARIMA yang mungkin digunakan seperti ARIMA (8,1,11), ARIMA (8,1,9), ARIMA (9,1,10), ARIMA (10,1,9), dan ARIMA (10,1,10) sebagai model sementara.

### 5) Estimasi Model ARIMA

Estimasi model ARIMA menggunakan AIC dan BIC melibatkan perbandingan nilai-nilai kriteria tersebut dari berbagai model ARIMA yang diestimasi. AIC dan BIC adalah kriteria statistik yang menggabungkan informasi tentang kesesuaian dan kompleksitas model. AIC menekankan keseimbangan antara *goodness-of-fit* dan kompleksitas model, sedangkan BIC memberikan penalti yang lebih besar terhadap

model yang kompleks. Dengan memilih model yang memiliki nilai AIC atau BIC terendah, kita dapat memilih model ARIMA yang paling sesuai dengan data, dengan tingkat kompleksitas yang efisien. Pada Tabel 3 menyajikan hasil estimasi 5 model ARIMA, yang akan dipilih 1 model ARIMA terbaik berdasarkan perbandingan nilai AIC dan BIC.

Tabel 3. Pemilihan Estimasi Model ARIMA

| Model           | AIC      | BIC      |
|-----------------|----------|----------|
| ARIMA (8,1,11)  | 2956.215 | 3012.781 |
| ARIMA (8,1,9)   | 2959.487 | 3010.397 |
| ARIMA (9,1,10)  | 2955.897 | 3012.463 |
| ARIMA (10,1,9)  | 2955.964 | 3012.530 |
| ARIMA (10,1,10) | 2956.818 | 3016.212 |

Pada Tabel 3 menunjukkan bahwa model ARIMA (9,1,10) menunjukkan AIC dan BIC yang paling rendah dibandingkan dengan model lain sehingga prediksi ARIMA data penjualan akan menggunakan ARIMA (9,1,10) sebagai parameter dalam implementasi ARIMA.

6) Penerapan Model ARIMA

Dalam persiapan untuk implementasi model ARIMA, parameter yang dipilih berdasarkan nilai AIC dan BIC terendah adalah ARIMA (9,1,10). Selanjutnya, data akan dibagi menjadi data pelatihan dan data pengujian. Pembagian ini memungkinkan kita untuk menguji kehandalan model dan melihat sejauh mana model tersebut dapat memprediksi data yang belum pernah dilihat sebelumnya. Data pelatihan akan mencakup 80% dari keseluruhan data, sementara data pengujian akan mencakup 20%. Setelah pembagian data, kita dapat memulai proses fitting data menggunakan model ARIMA untuk menganalisis penjualan bisnis rumah properti dimana hasil penerapan model ARIMA dapat ditunjukkan pada tabel 4.

Tabel 4. Hasil Penerapan model ARIMA

| year_month | Forecast    |
|------------|-------------|
| 2017-06-01 | 651482.2628 |
| 2017-07-01 | 648199.2904 |
| 2017-08-01 | 635064.0927 |
| 2017-09-01 | 645570.5472 |
| 2017-10-01 | 662497.6103 |
| ...        | ...         |
| 2019-07-01 | 649505.5638 |

Pada Tabel 4 menunjukkan bahwa *year\_month* yaitu tahun dan bulan yang dilakukan *forecast* berarti hasil prediksi model ARIMA pada data penjualan rumah properti dengan metode prediksi metode ARIMA.

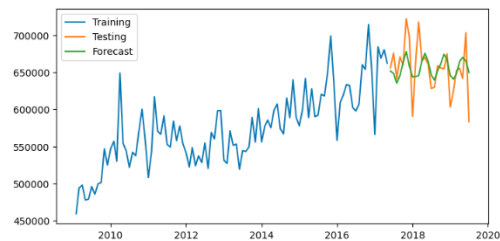
7) Pengecekan Akurasi Model

Setelah *fitting data* dengan model ARIMA, langkah selanjutnya adalah mengevaluasi akurasi model menggunakan beberapa metrik yang telah dipilih, yaitu RMSE, MSE, dan MAPE. Metrik-metrik ini digunakan untuk mengevaluasi seberapa akurat model ARIMA dalam menghasilkan

prediksi pada data penjualan bisnis rumah properti. Semakin kecil nilai-nilai metrik tersebut, semakin baik performa model dalam menghasilkan prediksi yang lebih akurat. Hasil pengecekan akurasi model dapat ditunjukkan bahwa nilai yang dihasilkan metrik RMSE, MSE dan MAPE masing-masing yaitu 0.281409; 0.079191 dan MAPE sebesar 3,4%. Hasil yang ditunjukkan pada MAPE kurang dari 10% yang dapat dianggap sebagai tingkat akurasi yang sangat baik (Swanson, 2015).

4.2. Pembahasan

Prediksi penjualan menggunakan algoritma ARIMA berhasil dilakukan pada *dataset* penjualan rumah bisnis properti dengan menggunakan model ARIMA (9,1,10) menunjukkan bahwa data tersebut dapat dikatakan baik karena berdasarkan grafik pada *timeseries* menunjukkan bahwa nilai peningkatan data aktual dan data *forecast* dapat dikatakan hampir sama yang ditunjukkan pada Gambar 6.



Gambar 6. Hasil Perbandingan Data Actual dan Data Forecast

Analisis keakuratan metode ARIMA dalam prediksi penjualan bisnis rumah properti memberikan informasi tentang sejauh mana model ini dapat memprediksi dengan akurat jumlah penjualan di masa depan. Evaluasi dilakukan dengan membandingkan nilai prediksi dengan data penjualan aktual menggunakan metrik evaluasi seperti MSE, RMSE, dan MAPE. Hasil analisis keakuratan mengindikasikan sejauh mana model ARIMA dapat memodelkan pola dan tren penjualan rumah. Nilai keakuratan yang rendah menunjukkan ketidakcocokan antara model dan data, sedangkan nilai yang tinggi menunjukkan prediksi yang akurat dan dapat diandalkan. Dalam penelitian ini, pengukuran akurasi model menggunakan MAPE menunjukkan tingkat *error* hanya 3%, yang menunjukkan tingkat akurasi yang baik. Tingkat *error* dari hasil penelitian ini lebih baik dari penelitian sebelumnya yang melakukan prediksi segmentasi pasar potensial menggunakan ARIMA dengan nilai tingkat *error* sebesar 5,22% (Andryana et al., 2023).

Model ARIMA dapat digunakan untuk penjualan properti yang lebih luas, termasuk apartemen dan properti komersial. Model ini dapat diadaptasi dengan menganalisis data *timeseries* dan memperhatikan tren, musiman, dan fluktuasi dalam data historis. Dengan demikian, model ARIMA dapat diterapkan pada



berbagai jenis data penjualan properti, termasuk penjualan apartemen, properti komersial, dan jenis properti lainnya. Penerapan model ARIMA dalam prediksi penjualan rumah properti memberikan dampak positif pada strategi bisnis perusahaan. Model ini membantu mengidentifikasi tren penjualan, menghindari ketidaksesuaian persediaan, mengurangi risiko finansial, dan memungkinkan pengambilan keputusan strategis yang lebih baik. Prediksi yang akurat membantu perusahaan menjadi lebih efisien, adaptif, dan berdaya saing di pasar properti yang dinamis.

## 5. Kesimpulan

Model ARIMA terbaik yang dapat digunakan untuk memprediksi penjualan rumah properti adalah model ARIMA (9,1,10) berdasarkan nilai kebaikan model dan pemenuhan asumsi-asumsi model. Berdasarkan model tersebut, hasil prediksi penjualan rumah properti selama Juni 2017 hingga pada Juli 2019 dengan pengukuran akurasi model dengan menggunakan MSE sebesar 0.079191; RMSE sebesar 0.281409 dan MAPE sebesar 3.4% sehingga dapat dinyatakan model yang diimplementasi memiliki tingkat akurasi yang baik. Saran untuk penelitian selanjutnya adalah mempertimbangkan penggunaan metode prediksi lain seperti metode *prophet* dengan variabel hari libur untuk meningkatkan akurasi serta eksplorasi penggunaan variabel tambahan dan perbandingannya dengan metode lain untuk menguji keunggulan metode yang digunakan (Menculini et al., 2021).

## Ucapan Terima Kasih

Artikel jurnal ini ditulis berdasarkan hasil penelitian yang dibiayai oleh LPPM (Lembaga Penelitian dan Pengabdian kepada Masyarakat) Universitas Pelita Harapan.

## Daftar Pustaka

- Andryana, S., Rahman, B., & Gunaryati, A., 2023. Predicting market segments from twitter data using ARIMA Time Series Analysis. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 9(1), 66-72. <https://doi.org/10.33480/jitk.v9i1.4275>
- Barus, O., & Wijaya, C., 2021. Implementasi Metode Neural Network Backpropagation dalam Prediksi Indeks Harga Saham Gabungan (IHSG). *Seminar Nasional (SEMINASTIKA)*, 3(1), 79-85. <https://doi.org/10.47002/seminastika.v3i1.252>
- Br Bangun, R. H., 2017. Penerapan Autoregressive Integrated Moving Average (ARIMA) pada peramalan produksi kedelai di Sumatera Utara. *JURNAL AGRICA*, 9(2), 90-100. <https://doi.org/10.31289/agrica.v9i2.484>
- Catur Putri, S. R., & Junaedi, L., 2022. Penerapan metode peramalan Autoregressive Integrated Moving Average pada sistem informasi pengendalian persediaan bahan baku. *Jurnal Ilmu Komputer Dan Bisnis*, 13(1), 164-173. <https://doi.org/10.47927/jikb.v13i1.293>
- Damanik, R. R., & Poernomo, M. H., 2023. Prediksi pembelian barang pada distributor lampu menggunakan metode apriori pada PT. XYZ. *JDMIS: Journal of Data Mining and Information System*, 1(1), 05-19. <https://doi.org/10.54259/jdmis.v1i1.1500>
- Han, J., Kamber, M., & Pei, J., 2012. *Data Mining: Concepts and Techniques*. In J. Han, M. Kamber, & J. Pei (Eds.), *Data Mining (Third Edition) (Third Edition)*. Elsevier
- Hassyddiqy, H., & Hasdiana, H., 2023. Analisis peramalan (forecasting) penjualan Dengan Metode ARIMA (Autoregressive Integrated Moving Average) Pada Huebee Indonesia. *Data Sciences Indonesia (DSI)*, 2(2), 92-100. <https://doi.org/10.47709/dsi.v2i2.2022>
- Indarwati, T., Irawati, T., & Rimawati, E., 2018. Penggunaan metode linear regression untuk prediksi penjualan smartphone. *Jurnal Teknologi Informasi Dan Komunikasi (TIKOMSiN)*, 6(2). <https://doi.org/10.30646/tikomsin.v6i2.369>
- Leonie, V., & Pangaribuan, J.J., 2020. prediksi video performance akun youtube buzzfeed menggunakan Metode Naïve Bayes. *Riset Dan E-Jurnal Manajemen Informatika Komputer*, 5(1).
- Menculini, L., Marini, A., Proietti, M., Garinei, A., Bozza, A., Moretti, C., & Marconi, M., 2021. Comparing Prophet and Deep Learning to ARIMA in Forecasting Wholesale Food Prices. *Forecasting*, 3(3), 644-662. <https://doi.org/10.3390/forecast3030040>
- Pangaribuan, J. J., & Suhartjito, 2014. Diagnosis of diabetes mellitus using extreme learning machine. *2014 International Conference on Information Technology Systems and Innovation, ICITSI 2014 - Proceedings*, 33-38. <https://doi.org/10.1109/ICITSI.2014.7048234>
- Permatasari, C. I., Sutopo, W., & Hisjam, M., 2018. Sales forecasting newspaper with ARIMA: A Case Study. *AIP Conference Proceedings*, 030017. <https://doi.org/10.1063/1.5024076>
- Ramadhan, M. R., Tursina, T., & Novriando, H., 2020. Implementasi fuzzy time series pada prediksi jumlah penjualan rumah. *Jurnal Sistem Dan Teknologi Informasi (Justin)*, 8(4), 418. <https://doi.org/10.26418/justin.v8i4.40186>
- Ratningsih, 2017. Forecasting Penjualan Rumah Dengan Menggunakan Metode Trend Moment pada PT. Rumakita Prima Karsa. *Perspektif*, 15(1), 40-48.
- Saputra, Moch. F. E., & Rizky, M. (2019). Peramalan jumlah kasus infeksi saluran pernafasan akut (ISPA) pada Laki-Laki Tahun 2019 dengan

- Metode ARIMA. *Jurnal Biometrika Dan Kependudukan*, 8(2), 138. <https://doi.org/10.20473/jbk.v8i2.2019.138-145>
- Swanson, D. A., 2015. On the Relationship among Values of the Same Summary Measure of Error when it is used across Multiple Characteristics at the Same Point in Time: An Examination of MALPE and MAPE.
- Vercellis, C., 2009. *Business Intelligence: Data Mining and Optimization for Decision Making (First Edition)*. John Wiley & Sons. <https://doi.org/10.1002/9780470753866>
- Wulandari, R. A., & Gernowo, R., 2019. Metode Autoregressive Integrated Movingaverage (ARIMA) dan Metode Adaptive Neuro Fuzzy Inference System (ANFIS) Dalam Analisis Curah Hujan. *BERKALA FISIKA: Jurnal Fisika Teori, Eksperimen, Dan Fisika Aplikasi*, 22(1), 41-48.
- Žunić, E., Korjenić, K., Hodžić, K., & Đonko, D., 2020. Application of facebook's prophet algorithm for successful sales forecasting based on real-world data. *International Journal of Computer Science and Information Technology*, 12(2), 23-36. <https://doi.org/10.5121/ijcsit.2020.12203>