

4. PENGUJIAN

4.1. Pengujian Hasil Prediksi Penjualan

Data yang digunakan dalam proses prediksi :

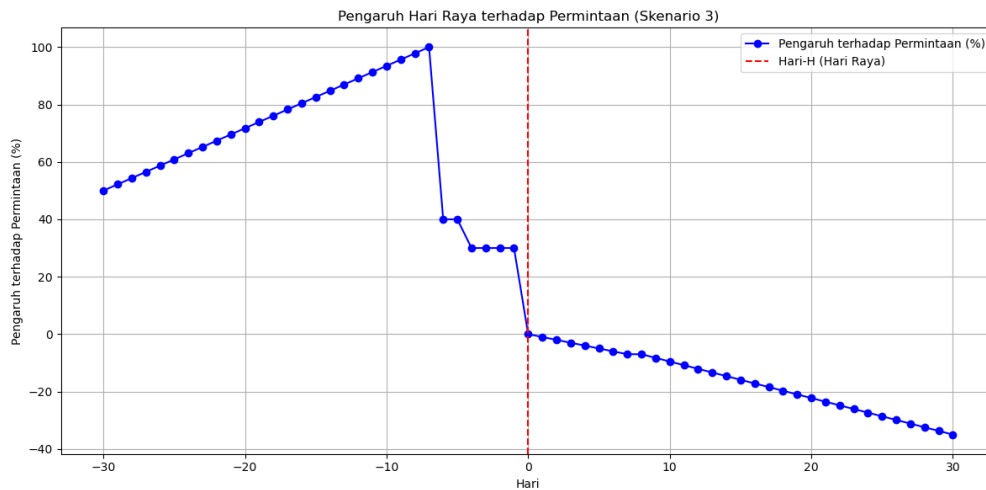
- Data utama: Data penjualan PT XYZ dari Januari 2021 hingga Desember 2024. Dengan data Januari 2021 hingga Desember 2023 sebagai *data training* dan Januari hingga Desember 2024 sebagai *data testing*.
- Data faktor eksogen: Hari Raya Idul Fitri dan Hari Libur Natal.

Prediksi dilaksanakan dengan membandingkan antara metode SARIMAX dan XGBoost dengan menggunakan *grid search* untuk mendapatkan parameter terbaik pada masing-masing produk. Pada model dilakukan beberapa pengujian pada subbab di bawah ini untuk mendapatkan hasil prediksi yang paling mendekati akurat.

4.1.1. Analisis Perhitungan Faktor Eksogen

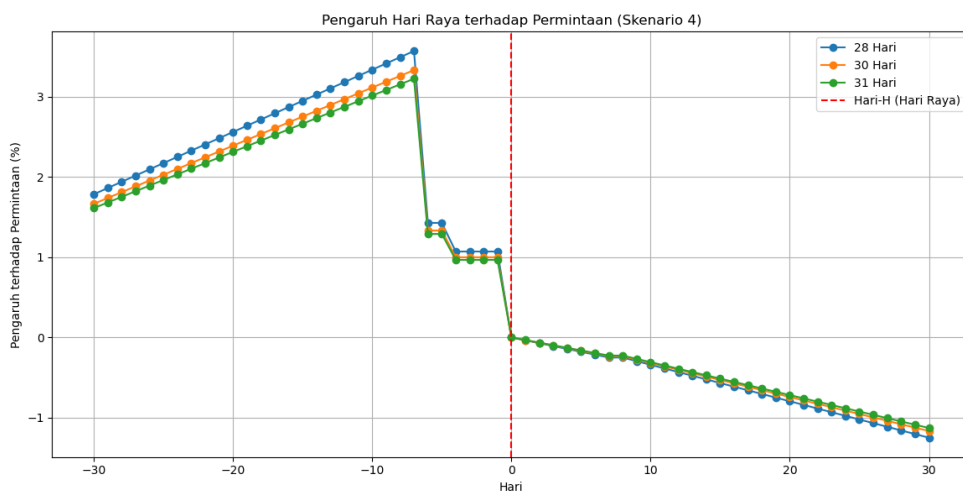
Beberapa cara perhitungan faktor eksogen yang berupa Hari Raya Idul Fitri dan Hari libur Natal dilakukan dengan :

- Skenario 1 adalah tanpa menggunakan faktor eksogen.
- Skenario 2 adalah menghitung berapa banyak hari dalam satu bulan yang termasuk dalam rentang 30 hari sebelum hari raya. Setiap hari yang berada dalam rentang tersebut dianggap memiliki pengaruh terhadap peningkatan permintaan dan diberi bobot 1. Jumlah hari yang memenuhi kriteria ini akan dijumlahkan sebagai representasi total efek hari raya pada bulan tersebut.
- Skenario 3 menghitung pengaruh hari raya terhadap permintaan berdasarkan jarak hari dari tanggal hari raya. 30 hingga 7 hari sebelum hari H, permintaan diperkirakan meningkat secara bertahap, dari 50% hingga mencapai puncaknya sebesar 100%. Setelah itu, permintaan mulai menurun, yaitu menjadi 40% pada 6 hingga 5 hari sebelum hari raya, lalu turun lagi menjadi 30% pada 3 hingga 1 hari sebelumnya, dan mencapai 0% tepat pada hari H. Setelah hari raya, dari H+1 hingga H+7, pengaruh terhadap permintaan menurun secara bertahap sebesar 1% setiap harinya. Kemudian, pada H+8 hingga H+30, pengaruh terus menurun hingga mencapai maksimal penurunan sebesar -35%. Hari-hari di luar rentang tersebut dianggap tidak memiliki pengaruh terhadap permintaan.



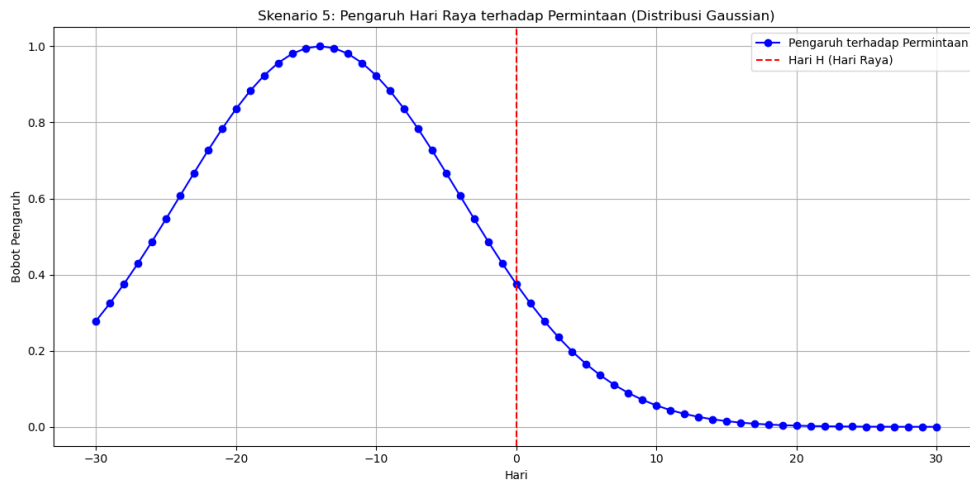
Gambar 4. 1 Grafik Bobot Pengaruh pada Skenario 3

- Skenario 4 menggunakan logika perhitungan dari Skenario 3 namun pada Skenario 4 total efek tersebut dibagi dengan jumlah hari dalam bulan untuk memperoleh nilai rata-rata harian. Pembagian ini bertujuan untuk melakukan normalisasi, sehingga membuat nilai akhir lebih proporsional dan sebanding antar bulan, terlepas dari panjang bulan atau jumlah hari yang terdampak oleh hari raya.



Gambar 4. 2 Grafik Bobot Pengaruh pada Skenario 4

- Skenario 5 menggunakan pendekatan matematis berupa distribusi Gaussian/ distribusi normal untuk menghitung pengaruh hari raya terhadap permintaan. Distribusi ini memiliki bentuk kurva lonceng dan memungkinkan untuk memberikan bobot tertinggi pada tanggal 14 hari sebelum hari raya, serta bobot yang menurun secara halus jika semakin jauh dari tanggal tersebut baik sebelum maupun sesudah.



Gambar 4. 3 Grafik Bobot Pengaruh pada Skenario 5

Hasil dari pengujian 10 produk pada masing-masing skenario dapat dilihat di Tabel 4.1.

Tabel 4. 1 Tabel Hasil Pengujian Pengaruh Perhitungan Hari Raya pada Produk Tertentu

Nama Produk	SKENARIO									
	1		2		3		4		5	
	RMSE SARIMAX	RMSE XGBoost	RMSE SARIMAX	RMSE XGBoost	RMSE SARIMAX	RMSE XGBoost	RMSE SARIMAX	RMSE XGBoost	RMSE SARIMAX	RMSE XGBoost
Dada Ayam Sliced	28,2493604	34,4509658	38,290305	33,4274975	38,4240604	34,5622086	38,3054557	34,561049	39,2153601	31,0557924
Brisket US	22,7339232	59,2142381	30,8565563	47,0027622	32,9949675	47,0566794	32,7040278	47,0566794	35,4671338	41,8784897
Bt Pot Biasa	168,874807	198,161508	91,5788895	175,214034	128,194915	166,038684	107,464514	166,038684	197,387796	235,764593
Tops Lok	113,730744	136,310969	68,1907855	70,9996364	85,9042451	86,3870435	87,8040022	86,3870435	75,8387855	84,2499868
Blade Lokal	281,753784	305,098242	154,626175	277,311471	80,5574671	237,396968	115,132144	235,602941	145,04797	278,66566
Tend Lok	188,848189	237,232524	132,057518	177,065167	111,164322	140,197927	113,577460	161,767616	111,54669	169,128474
Knk Lok	106,449646	134,295103	69,8755743	105,154472	64,770671	100,400976	64,6412145	100,400976	65,9185004	92,7373796
Sirlo Lok	139,600482	179,150833	97,3377822	184,01087	96,7053475	168,490408	83,0564757	95,9548192	105,648791	160,798021
Ayam Gendut Cut 12	18,4689559	24,4515006	20,3964179	35,4150403	21,3619712	28,2661638	21,6728313	33,1530578	20,7766792	17,2961982
Ayam Kampung	26,679684	27,1642756	17,970352	14,9351665	11,1632935	16,9064541	11,1081251	16,9666021	16,8431592	10,0553467

Tabel 4. 2 Tabel Hasil Statistik Pengujian Pengaruh Perhitungan Hari Raya pada Keseluruhan Produk Aktif

	Skenario 1		Skenario 2		Skenario 3		Skenario 4		Skenario 5	
RMSE	SARIMAX	XGBoost	SARIMAX	XGBoost	SARIMAX	XGBoost	SARIMAX	XGBoost	SARIMAX	XGBoost
<i>Average</i>	74,7	98,65	57,19	87,87	56,55	85,07	56,5	82,84	61,93	80,55
Std Dev	64,82	79,7	39,13	81,07	36,04	77,7	35,46	75,77	46,1	65,85
Jumlah Produk Menang	7	0	6	0	7	0	3	0	4	5

Analisa dilakukan berdasarkan Tabel 4.1 :

Skenario 1 menunjukkan performa terbaik untuk produk-produk yang tidak terpengaruh secara signifikan oleh musim atau momen tertentu seperti Hari Raya Idul Fitri dan Natal. Hal ini terlihat dari konsistensi nilai RMSE yang rendah pada beberapa produk yang umumnya memiliki pola permintaan stabil sepanjang tahun. Sebagai contoh, pada produk seperti Dada Ayam Sliced dan Brisket US model SARIMAX maupun XGBoost menghasilkan nilai RMSE yang lebih kecil dibandingkan skenario lainnya.

Skenario 2 menunjukkan beberapa produk seperti Bt Pot Biasa dan Tops Lok bahkan memperoleh hasil yang lebih optimal menggunakan skenario ini. Namun, untuk sebagian besar produk, pendekatan ini kurang tepat karena tidak mampu merepresentasikan variasi intensitas permintaan yang sebenarnya meningkat secara bertahap dan mencapai puncaknya mendekati hari H. Oleh karena itu, skenario ini kurang cocok diterapkan pada perusahaan dengan pola permintaan musiman yang dinamis seperti PT. XYZ, di mana efek musiman tidak bersifat linear maupun seragam.

Skenario 3 mencoba menggambarkan pola permintaan yang lebih realistis berdasarkan kondisi aktual di lapangan. Dalam skenario ini, pengaruh hari raya dihitung secara bertingkat, di mana permintaan meningkat bertahap menjelang hari H dan menurun setelahnya. Pendekatan ini mencerminkan pola konsumsi masyarakat yang lebih realistis, khususnya untuk produk-produk yang permintaannya sensitif terhadap momen tertentu seperti Blade Lokal, dan Tend Lok. Namun untuk produk yang tidak mengalami lonjakan musiman, skenario ini justru dapat menurunkan akurasi prediksi karena menambahkan variabel yang tidak relevan, sehingga meningkatkan risiko *overfitting* seperti pada Dada Ayam Sliced, Ayam Gendut Cut 12, dan Brisket US.

Skenario 4 merupakan penyempurnaan dari Skenario 3, dengan menambahkan proses normalisasi terhadap jumlah hari dalam bulan. Tujuan dari normalisasi ini adalah untuk menjaga agar efek musiman tetap proporsional dan tidak bersifat bias terhadap panjangnya bulan atau distribusi kalender. Pada skenario ini, produk dengan pola musiman tinggi pada hari raya seperti Sirlo Lok dan Knk Lok menunjukkan peningkatan performa yang signifikan. Meskipun Tend Lok dan Blade Lokal juga memiliki pola musiman yang kuat, nilai prediksi mereka mengalami sedikit penurunan dibandingkan Skenario 3, namun masih tetap lebih baik jika dibandingkan dengan skenario lainnya. Penurunan performa ini kemungkinan disebabkan oleh normalisasi yang membuat sinyal lonjakan permintaan yang sangat tajam menjadi sedikit kurang terlihat atau sedikit teredam.

Skenario 5 menggunakan pendekatan distribusi Gaussian untuk merepresentasikan efek hari raya. Skenario ini menunjukkan performa terbaik dalam peramalan produk ayam menggunakan model XGBoost. Hal ini dibuktikan dengan nilai RMSE XGBoost terendah yang diperoleh pada produk Ayam Gendut Cut 12 (17,30) dan Ayam Kampung (10,06). Selain itu, produk Dada Ayam Sliced juga mencatatkan hasil yang sangat baik dengan RMSE XGBoost sebesar 31,06, menjadikannya performa terbaik kedua untuk produk tersebut. Dengan demikian, Skenario 5 direkomendasikan untuk produk-produk unggas, khususnya saat menggunakan model XGBoost yang secara konsisten menunjukkan performa unggul pada skenario ini.

Berdasarkan Tabel 4.2, Skenario 4 menunjukkan performa prediksi yang lebih stabil dan konsisten di berbagai produk, baik pada model SARIMAX maupun XGBoost, ditunjukkan oleh nilai standar deviasi RMSE sebesar 35,46, yang merupakan yang terendah di antara semua skenario. Meskipun dari sisi jumlah kemenangan per produk, Skenario 1 dan 3 menunjukkan keunggulan numerik, Skenario 4 tetap unggul dari segi rata-rata akurasi keseluruhan, dengan nilai RMSE rata-rata paling rendah yaitu 56,5. Hal ini mengindikasikan bahwa Skenario 4 mampu mengurangi kesalahan prediksi secara keseluruhan, sehingga dapat dianggap sebagai pendekatan yang lebih umum dan baik untuk digunakan pada berbagai jenis produk.

Selain itu, pada setiap skenario, dapat diamati bahwa model SARIMAX secara konsisten menunjukkan kinerja yang lebih baik dibandingkan XGBoost, dengan nilai RMSE yang lebih rendah pada sebagian besar produk. Keunggulan ini mencerminkan kemampuan SARIMAX dalam menangani data *time-series* yang mengandung pola musiman yang sesuai dengan kebutuhan PT. XYZ.

4.1.2. Analisis Pengaruh Fitur Historis terhadap Performa Model XGBoost

Untuk meningkatkan akurasi model XGBoost, ditambahkan fitur-fitur historis seperti nilai permintaan di masa lalu dan rata-rata terakhir. Nilai yang hilang pada fitur ini diisi dengan rata-rata permintaan keseluruhan. Penambahan fitur ini bertujuan agar model dapat menangkap dinamika perubahan permintaan dari waktu ke waktu. Selain itu, informasi waktu berupa bulan dan tahun juga ditambahkan untuk membantu model mengenali pola musiman dalam data.

Berdasarkan Tabel 4.3, dilakukan evaluasi terhadap tiga variasi model. Model pertama merupakan model baseline yang tidak menggunakan fitur historis. Model kedua menggunakan fitur historis berupa permintaan satu dan dua bulan sebelumnya (*lag1* dan *lag2*) serta rata-rata permintaan tiga bulan terakhir (*rolling mean 3*). Model ketiga menggunakan tambahan fitur

historis yang lebih panjang, yaitu permintaan satu, dua, dan dua belas bulan sebelumnya (*lag1*, *lag2*, *lag12*) serta rata-rata permintaan tiga bulan dan dua belas bulan terakhir (*rolling mean 3* dan *rolling mean 12*). Hasil evaluasi menunjukkan bahwa model dengan fitur historis pendek memiliki nilai RMSE rata-rata sebesar 82,898, sedikit lebih tinggi dibandingkan model tanpa fitur historis yang memiliki RMSE sebesar 82,842, dengan selisih hanya 0,056. Meskipun perbedaannya kecil dalam hal akurasi, model dengan fitur historis pendek menunjukkan peningkatan signifikan dari sisi stabilitas, ditandai dengan penurunan standar deviasi dari 76,466 menjadi 62,880.

Model dengan fitur historis yang lebih panjang justru menunjukkan peningkatan RMSE rata-rata menjadi 84,570. Hal ini menunjukkan bahwa penambahan fitur historis yang terlalu jauh ke belakang, seperti *lag 12* dan *rolling mean 12*, tidak selalu memberikan keuntungan. Kemungkinan besar, informasi permintaan satu tahun sebelumnya tidak lagi relevan terhadap kondisi saat ini, atau justru memperkenalkan *noise* yang mengganggu proses pembelajaran model. Oleh karena itu, pemilihan fitur historis perlu disesuaikan dengan karakteristik data dan pola permintaan masing-masing produk.

Tabel 4. 3 Tabel Hasil Statistik Pengujian Pengaruh Fitur Historis pada XGBoost

RMSE XGBoost Skenario 4	<i>Average</i>	Std Dev
Tanpa Ditambahkan Fitur Historis	82,842	76,466
Dengan Ditambahkan Fitur Historis (<i>lag1 + lag2 + rolling mean 3</i>)	82,898	62,88
Dengan Ditambahkan Fitur Historis (<i>lag1 + lag2+ lag12 + rolling mean 3 + rolling mean 12</i>)	84,57	63,64

4.1.3. Analisis Pengaruh Penghapusan *Outlier* terhadap Performa Model

Outlier dalam data penjualan dihitung menggunakan metode Interquartile Range (IQR), di mana data yang berada di bawah batas bawah (*lower bound*) dan di atas batas atas (*upper bound*) dianggap sebagai *outlier* dan dihapus. Analisis ini dilakukan setelah penerapan Skenario

4, yaitu skenario normalisasi hari libur yang sebelumnya menunjukkan performa terbaik dan dengan menambahkan fitur historis untuk XGBoost.

Berdasarkan Tabel 4.4, produk yang tidak musiman seperti Dada Ayam Sliced dan Brisket US cenderung menunjukkan performa lebih baik dengan penghapusan *outlier* karena pola penjualannya relatif lebih stabil dan tidak banyak dipengaruhi oleh momen hari raya. Sebaliknya, peningkatan nilai RMSE secara keseluruhan kemungkinan disebabkan oleh penghapusan *outlier* yang mencakup lonjakan penjualan signifikan pada momen hari raya seperti Idul Fitri atau Natal. Lonjakan ini merupakan bagian penting dari pola musiman yang sebenarnya, sehingga menghilangkannya membuat model kehilangan informasi krusial untuk mengenali pola tersebut.

Tabel 4. 4 Tabel Hasil Perbandingan Dengan dan Tanpa Penghapusan Outlier pada Produk Tertentu

Nama Produk	Tanpa Penghapusan <i>Outlier</i>		Dengan Penghapusan <i>Outlier</i>	
	RMSE SARIMAX	RMSE XGBoost	RMSE SARIMAX	RMSE XGBoost
Dada Ayam Sliced	38,3054557	44,9005429	25,545066	38,120693
Brisket US	32,7040278	24,8090586	16,1184598	35,9548566
Bt Pot Biasa	107,4645147	163,0647509	270,781903	266,09023
Tops Lok	87,8040022	147,0815495	104,8629356	212,642752
Blade Lokal	115,1321448	196,1484363	150,8991356	255,8177368
Tend Lok	113,5774609	207,1070652	252,2342917	207,1070652
Knk Lok	64,6412145	121,0488245	96,9639201	120,0853976
Sirlo Lok	83,0564757	156,3834906	65,9372524	165,7862864
Ayam Gendut Cut 12	21,6728313	24,0523441	21,0522624	27,1306732
Ayam Kampung	11,1081251	17,1655343	55,1802113	55,2666512

Pada Tabel 4.5, perhitungan rata-rata RMSE dari seluruh produk aktif menunjukkan pola yang konsisten, yaitu peningkatan nilai RMSE setelah *outlier* dihapus pada kedua model, SARIMAX maupun XGBoost. Dengan demikian, penghapusan *outlier* justru dapat menghilangkan informasi penting yang dibutuhkan oleh model untuk mempelajari pola musiman yang khas

pada periode-periode tertentu. Namun baik dengan maupun tanpa penghapusan *outlier*, hasil pengujian tetap konsisten menunjukkan bahwa model SARIMAX memiliki performa terbaik dibandingkan XGBoost.

Tabel 4. 5 Tabel Hasil Statistik Pengujian Pengaruh Penghapusan *Outlier* pada Keseluruhan Produk Aktif

RMSE Skenario 4	<i>Average</i>	
	SARIMAX	XGBoost
Tanpa Penghapusan <i>Outlier</i>	57,03	84,343
Dengan Penghapusan <i>Outlier</i>	72,075	109,738
Selisih	15,045	25,395

4.1.4 Analisis Pengaruh Transformasi Yeo-Johnson terhadap Performa Model

Transformasi Yeo-Johnson dipilih karena mampu menangani data yang mengandung nilai nol serta bertujuan untuk mengurangi *skewness* dan meningkatkan kestasioneran data, yang diharapkan dapat meningkatkan akurasi model SARIMAX dan XGBoost. Analisis ini dilakukan setelah penerapan Skenario 4, yaitu skenario normalisasi hari libur yang sebelumnya menunjukkan performa terbaik dan dengan menambahkan fitur historis untuk XGBoost.

Berdasarkan Tabel 4.6, untuk penerapan transformasi Yeo-Johnson pada model SARIMAX memberikan hasil yang beragam. Beberapa produk seperti Brisket US, Bt Pot Biasa, Blade Lokal, dan Ayam Gendut Cut 12 menunjukkan penurunan RMSE yang cukup signifikan, menandakan peningkatan akurasi model. Namun, pada produk lain seperti Tend Lok, Tops Lok, dan Sirlo Lok, justru terjadi peningkatan RMSE, yang menunjukkan bahwa transformasi ini tidak selalu membawa perbaikan.

Sementara itu, untuk model XGBoost, efek transformasi Yeo-Johnson justru cenderung menurunkan performa. Hampir semua produk mengalami kenaikan RMSE setelah transformasi, bahkan secara ekstrem pada produk seperti Tend Lok (dari 207,11 menjadi 1530,24). Hal ini menunjukkan bahwa transformasi ini tidak cocok diterapkan pada XGBoost.

Tabel 4. 6 Tabel Hasil Perbandingan Dengan dan Tanpa transformasi Yeo-Johnson untuk Produk Tertentu

Nama Produk	Tanpa Transformasi Yeo-Johnson		Dengan Transformasi Yeo-Johnson	
	RMSE SARIMAX	RMSE XGBoost	RMSE SARIMAX	RMSE XGBoost
Dada Ayam Sliced	38,3054557	44,9005429	38,5004573	48,4239642
Brisket US	32,7040278	24,8090586	19,7206909	37,1059442
Bt Pot Biasa	107,4645147	163,0647509	97,4498422	144,6045627
Tops Lok	87,8040022	147,0815495	91,392954	167,1951237
Blade Lokal	115,1321448	196,1484363	62,4482177	218,4826928
Tend Lok	113,5774609	207,1070652	141,7711797	1530,241438
Knk Lok	64,6412145	121,0488245	65,4503906	132,5333056
Sirlo Lok	83,0564757	156,3834906	105,3176076	208,9528898
Ayam Gendut Cut 12	21,6728313	24,0523441	19,9291022	28,2392225
Ayam Kampung	11,1081251	17,1655343	11,9926702	17,7902912

Berdasarkan hasil pengujian pada Tabel 4.7, dapat dilihat bahwa penerapan transformasi Yeo-Johnson memberikan dampak yang berbeda terhadap performa model SARIMAX dan XGBoost. Model SARIMAX menunjukkan perbedaan nilai rata-rata RMSE yang sangat kecil, hanya sebesar 0,018 antara data tanpa dan dengan transformasi. Hal ini mengindikasikan bahwa transformasi Yeo-Johnson memberikan efek yang minimal pada akurasi model SARIMAX.

Sebaliknya, model XGBoost mengalami penurunan performa yang drastis dengan peningkatan rata-rata RMSE sebesar 47,992 setelah penerapan transformasi. Lonjakan ini mengindikasikan bahwa transformasi Yeo-Johnson justru membuat distribusi data menjadi kurang sesuai untuk XGBoost, sehingga akurasi model menurun secara signifikan.

Tabel 4. 7 Tabel Hasil Statistik Pengujian Pengaruh Transformasi Yeo-Johnson pada Keseluruhan Produk Aktif

RMSE Skenario 4	<i>Average</i>
-----------------	----------------

	SARIMAX	XGBoost
Tanpa Transformasi Yeo-Johnson	56,502	82,9
Dengan Transformasi Yeo-Johnson	56,484	130,892
Selisih	0,018	47,992

Oleh karena perbedaan RMSE pada SARIMAX dengan dan tanpa transformasi Yeo-Johnson relatif kecil, sedangkan pada XGBoost perbedaannya justru sangat besar dan cenderung memperburuk performa, maka diputuskan untuk menggunakan data tanpa transformasi. Selain itu, baik dengan maupun tanpa transformasi, model SARIMAX tetap menunjukkan performa yang lebih unggul dibandingkan XGBoost.

4.1.5 Perbandingan antara SARIMAX dan XGBoost

Berdasarkan hasil pengujian, konfigurasi terbaik diperoleh dengan menggunakan Skenario 4, yaitu skenario yang menerapkan normalisasi pengaruh hari raya sebagai variabel eksogen. Pengaruh hari raya dihitung berdasarkan jarak ± 30 hari dari tanggal perayaan, dengan permintaan meningkat sebelum dan menurun setelah hari raya. Konfigurasi ini digunakan tanpa penghapusan *outlier* dan tanpa transformasi Yeo-Johnson, serta dengan penambahan fitur historis untuk model XGBoost. Pendekatan ini terbukti lebih efektif dalam menangkap pola lonjakan musiman karena mampu merepresentasikan perubahan perilaku konsumen secara lebih realistis.

Dengan konfigurasi tersebut, sebagaimana terlihat pada Tabel 4.8, model SARIMAX menghasilkan nilai rata-rata *Root Mean Square Error* (RMSE) sebesar 56,502, sedangkan model XGBoost memiliki rata-rata RMSE sebesar 82,9. Perbedaan rata-rata error sebesar 26,398 poin, atau sekitar 31,86% lebih rendah pada SARIMAX, menunjukkan bahwa secara keseluruhan model ini mampu memberikan prediksi yang lebih akurat dibandingkan XGBoost. Hal ini mengindikasikan bahwa SARIMAX lebih efektif dalam menangkap pola musiman serta tren pada data penjualan historis.

Tabel 4. 8 Tabel Hasil Statistik Pengujian Akhir antara SARIMAX dan XGBoost pada Produk Keseluruhan

RMSE Skenario 4 + Tanpa Penghapusan <i>Outlier</i> + Tanpa Transformasi Yeo Johnson	<i>Average</i>
SARIMAX	56,502
XGBoost	82,9
Selisih	26,398

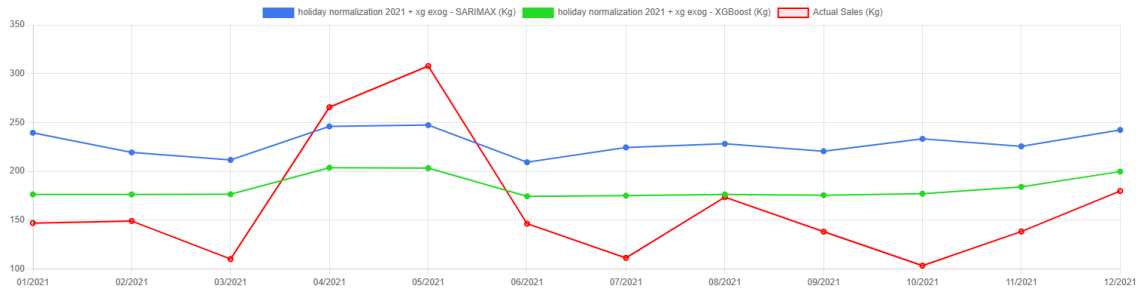
4.1.6 Visualisasi Hasil Prediksi Terbaik

Pada subbab ini ditampilkan visualisasi perbandingan antara data aktual dan hasil prediksi dari model SARIMAX serta XGBoost pada skenario terbaik untuk beberapa produk. Visualisasi ini bertujuan untuk menunjukkan sejauh mana masing-masing model mampu mengikuti pola data aktual serta menilai efektivitas prediksi yang dihasilkan. Garis berwarna merah mewakili data aktual, garis biru menunjukkan hasil prediksi dari model SARIMAX, dan garis hijau menunjukkan hasil prediksi dari model XGBoost.

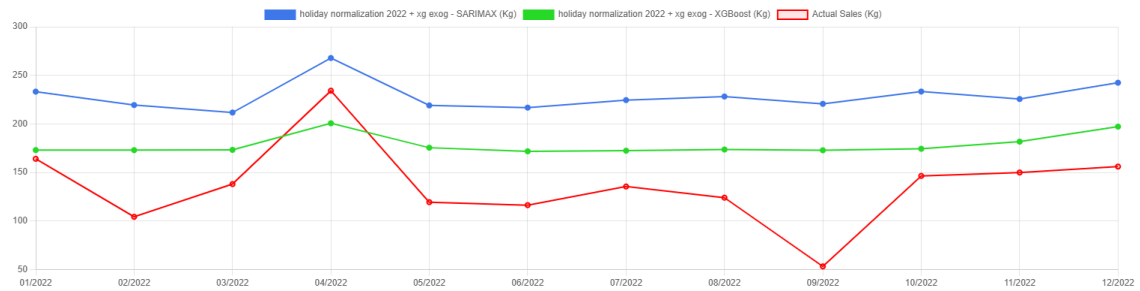
4.1.6.1 Produk Dada Ayam Sliced

Produk Dada Ayam Sliced merupakan contoh dari produk yang pola permintaan cenderung stabil tanpa lonjakan signifikan yang berkaitan dengan momen libur besar seperti Lebaran atau Natal. Hal ini dapat diamati pada grafik tahun 2021 hingga 2024 dimana penjualan aktual (*Actual Sales*) menunjukkan variasi yang tidak terlalu ekstrim serta tidak mengikuti pola musiman yang konsisten dari bulan ke bulan.

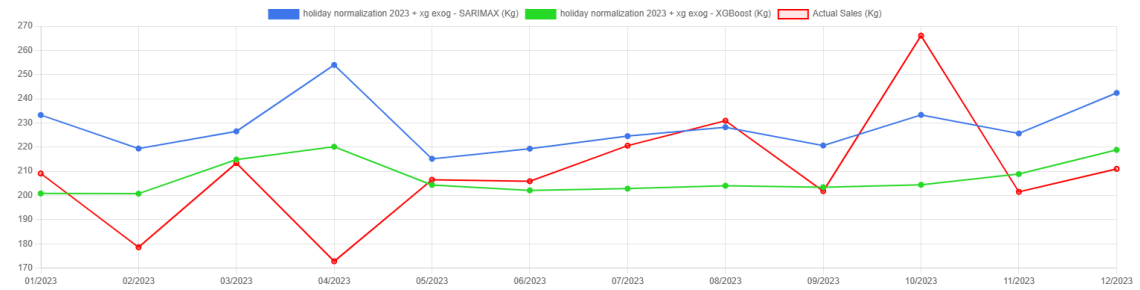
Dari keempat grafik Gambar 4.4 hingga 4.7, kedua model menunjukkan hasil prediksi yang kurang mampu mengikuti fluktuasi aktual, terutama pada bulan-bulan dengan perubahan tiba-tiba. Meskipun metode perhitungan hari raya skenario 4 diterapkan, hasil prediksi tetap tidak mengalami peningkatan signifikan karena memang tidak terdapat pola musiman yang dominan dalam data historis produk cenderung acak. Sehingga model prediksi seperti SARIMAX dan XGBoost kesulitan dalam mengikuti perubahan yang tidak berpola secara konsisten.



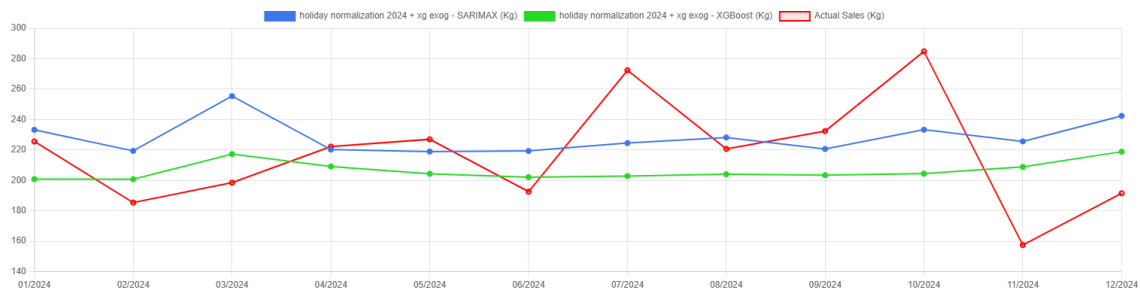
Gambar 4. 4 Visualisasi Grafik pada Produk Dada Ayam Sliced Tahun 2021



Gambar 4. 5 Visualisasi Grafik pada Produk Dada Ayam Sliced Tahun 2022



Gambar 4. 6 Visualisasi Grafik pada Produk Dada Ayam Sliced Tahun 2023

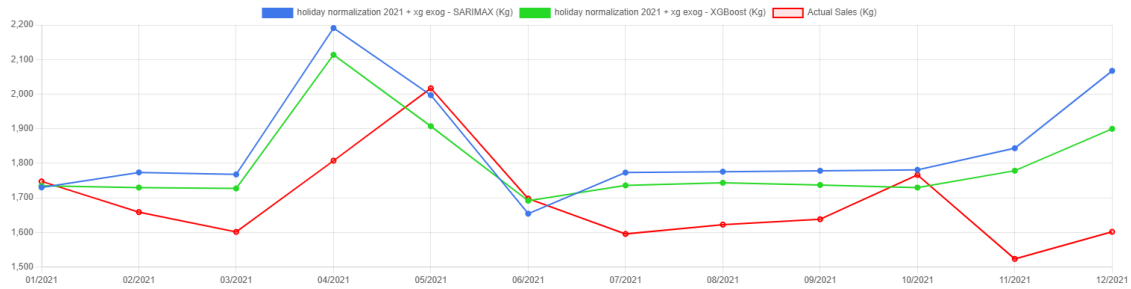


Gambar 4. 7 Visualisasi Grafik pada Produk Dada Ayam Sliced Tahun 2024

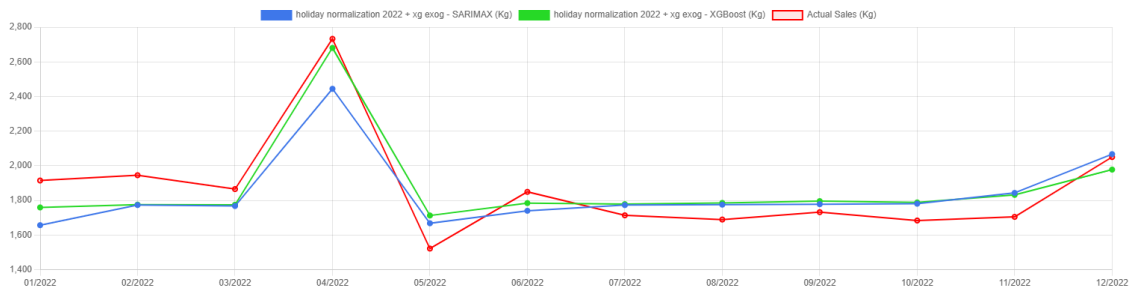
4.1.6.2 Produk Tend Lok

Produk Tend Lok merupakan produk musiman yang menunjukkan pola lonjakan permintaan yang signifikan menjelang momen-momen besar seperti Idul Fitri dan Natal. Hal ini

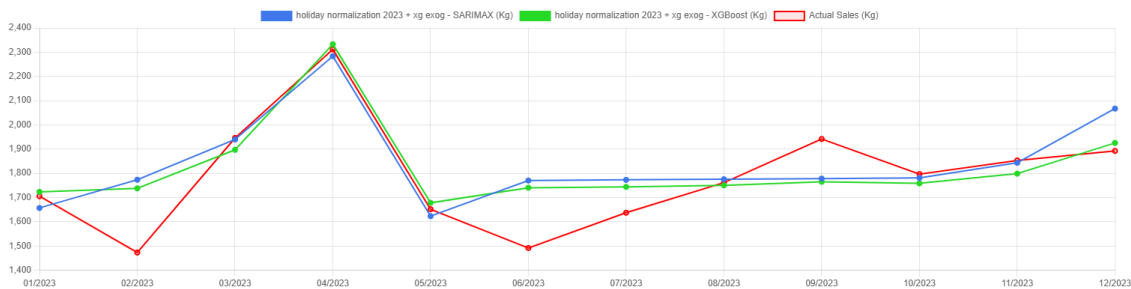
terlihat jelas dari puncak penjualan yang konsisten muncul sekitar bulan Maret/ April/ Mei (menjelang Lebaran) dan Desember (menjelang Natal) pada setiap tahun pengamatan.



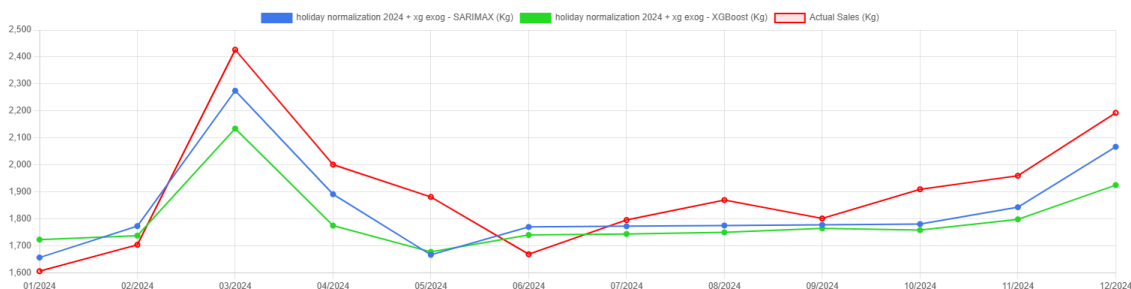
Gambar 4. 8 Visualisasi Grafik pada Produk Tend Lok Tahun 2021



Gambar 4. 9 Visualisasi Grafik pada Produk Tend Lok Tahun 2022



Gambar 4. 10 Visualisasi Grafik pada Produk Tend Lok Tahun 2023



Gambar 4. 11 Visualisasi Grafik pada Produk Tend Lok Tahun 2024

Berdasarkan Gambar 4.8 hingga 4.11, dapat dilihat bahwa kedua model mampu menangkap pola kenaikan dan penurunan yang terjadi akibat momen Hari Raya, terutama Natal dan Idul Fitri. Namun, di luar periode tersebut, data aktual cenderung tidak stasioner dan fluktuatif, sehingga sulit diprediksi secara akurat oleh model. Kedua model cenderung menghasilkan prediksi yang lebih landai. Jika dibandingkan, baik SARIMAX maupun XGBoost sama-sama mampu mengenali pola musiman, tetapi SARIMAX menunjukkan hasil yang lebih mendekati data aktual.

4.2. Perbandingan dengan Metode Prediksi Manual di PT. XYZ

Untuk membandingkan akurasi prediksi sistem dengan metode manual yang biasa digunakan di PT. XYZ, dilakukan pengujian dengan melibatkan dua orang karyawan yang mewakili divisi gudang dan divisi pembelian. Kedua karyawan tersebut diminta untuk mengisi sebuah formulir prediksi yang berisi daftar nama produk beserta estimasi jumlah penjualan untuk bulan Mei 2025, berdasarkan pengalaman dan metode manual yang selama ini mereka gunakan.

Setelah prediksi manual dari kedua karyawan dikumpulkan, hasilnya dibandingkan dengan data penjualan aktual bulan Mei 2025. Perbandingan dilakukan dengan menghitung nilai Root Mean Square Error (RMSE) dari masing-masing prediksi manual serta dari prediksi sistem yang dihasilkan oleh model SARIMAX dan XGBoost. Melalui perbandingan nilai RMSE ini, dapat dievaluasi seberapa besar perbedaan tingkat akurasi antara prediksi manual karyawan dengan prediksi berbasis model *time series* dan *machine learning*. Sehingga dapat menunjukkan potensi sistem dalam meningkatkan efisiensi dan ketepatan peramalan penjualan dibandingkan metode konvensional yang selama ini digunakan oleh PT. XYZ.

Tabel 4. 9 Tabel Perbandingan Hasil Pengujian Prediksi antara Metode Manual dan Sistem

	Orang Gudang	Orang Pembelian	Hasil Prediksi	Nilai Aktual
Total kuantitas penjualan	12.920	13.185	11.736,694	11.989,13
Selisih dengan aktual	930,87	1.195,87	252.436	-
Rata-rata RMSE	90,42	104,12	27,37	-

Tabel 4.9 menunjukkan bahwa prediksi sistem memberikan hasil yang paling mendekati nilai aktual, dengan total kuantitas penjualan sebesar 11.736,69, dibandingkan nilai aktual sebesar 11.989,13. Sementara itu, prediksi manual dari divisi gudang dan pembelian masing-masing menunjukkan selisih yang lebih besar, yaitu 930,87 dan 1.195,87. Nilai RMSE dari sistem prediksi sebesar 27,37 juga jauh lebih rendah dibandingkan metode manual, yang masing-masing memiliki RMSE sebesar 90,42 dan 104,12.

Berdasarkan hasil ini, dapat disimpulkan bahwa sistem mampu mengurangi tingkat kesalahan prediksi secara signifikan. Penurunan rata-rata RMSE hingga 71,87% dibandingkan metode manual menunjukkan bahwa sistem prediksi memberikan akurasi yang jauh lebih tinggi. Hal ini membuktikan bahwa sistem yang dikembangkan lebih efektif dalam mendukung pengambilan keputusan pengadaan stok di PT. XYZ dibandingkan metode konvensional yang selama ini digunakan.

4.3. Pengujian Sistem kepada Pengguna

Sistem telah dapat diakses secara daring sejak tanggal 7 Mei 2025 untuk keperluan uji coba oleh pengguna internal. Pada tanggal 4 Juni 2025, dilakukan proses evaluasi sistem dengan melibatkan empat perwakilan pengguna dari masing-masing divisi, yaitu penjualan, pembelian, gudang, dan keuangan.

Pengujian dilakukan melalui dua pendekatan. Pendekatan pertama adalah *User Acceptance Testing (UAT)*, yang dilakukan berdasarkan pengamatan langsung terhadap interaksi pengguna dengan sistem. Tujuannya adalah untuk memastikan bahwa seluruh fitur yang disediakan telah berfungsi sesuai kebutuhan pengguna serta sesuai dengan alur kerja (*workflow*) bisnis yang telah berjalan sebelumnya. Pendekatan kedua adalah *System Usability Scale (SUS)*, yang dilakukan melalui pengisian kuesioner digital menggunakan Google Form. Metode ini digunakan untuk mengukur tingkat kemudahan penggunaan serta kepuasan pengguna terhadap antarmuka dan fungsionalitas sistem.

4.2.1 User Acceptance Testing (UAT)

Pada proses *User Acceptance Testing (UAT)*, setiap skenario pengujian dijalankan sebanyak lima kali untuk memastikan sistem berfungsi sesuai kebutuhan pengguna akhir. Hasil dari tiap percobaan dicatat dan dibandingkan dengan harapan hasil yang telah ditentukan sebelumnya. Pengujian ini bertujuan untuk memvalidasi bahwa sistem sudah siap digunakan dalam lingkungan nyata, dengan mencerminkan kondisi operasional sesungguhnya dan memastikan tidak ada kendala fungsional yang signifikan.

4.2.1.1 Modul Penjualan

Pada Tabel 4.10 menunjukkan hasil pengujian *User Acceptance Testing* (UAT) untuk modul penjualan, yang mencakup enam skenario utama. Setiap skenario diuji sebanyak lima kali dan seluruhnya berhasil sesuai dengan harapan, menunjukkan bahwa sistem telah berjalan dengan baik dan mendukung proses penjualan secara lengkap dan akurat.

Tabel 4. 10 Tabel Pengujian UAT Modul Penjualan

No	Skenario	Harapan Hasil	Hasil Aktual	Percobaan
1	Pencatatan data pelanggan	Sistem harus menyimpan data pelanggan secara lengkap ke dalam basis data dan menampilkannya pada daftar pelanggan untuk transaksi berikutnya.	Data pelanggan bisa tersimpan dengan baik, meskipun kolom <i>contact person</i> dikosongkan.	5/5 berhasil
2	Pembuatan <i>sales order</i> berdasarkan pemesanan pelanggan	Sistem harus membentuk dokumen <i>sales order</i> yang berisi informasi pelanggan, daftar produk, jumlah dan harga.	<i>Sales order</i> berhasil dibuat, baik untuk produk yang ada maupun yang masih <i>pending</i> .	5/5 berhasil
3	Pembuatan dokumen <i>sales</i> yang merepresentasikan transaksi penjualan resmi	Sistem dapat mengonversi <i>sales order</i> yang telah disetujui menjadi dokumen <i>sales</i> secara otomatis, dengan seluruh data yang relevan ikut tertransfer.	Dokumen <i>sales</i> bisa langsung dibuat dari <i>sales order</i> yang sudah disetujui, namun terkadang membutuhkan waktu agar nomor <i>sales order</i> muncul.	5/5 berhasil
4	Pembuatan <i>delivery order</i>	Sistem harus menghasilkan <i>delivery</i>	<i>Delivery order</i> dapat dibuat dengan	5/5 berhasil

		<i>order</i> berdasarkan dokumen <i>sales</i> .	mudah berdasarkan <i>sales</i> dengan tanggal pengiriman terdekat.	
5	Pembuatan dokumen <i>sales return</i> jika ada barang rusak atau dikembalikan	Sistem harus mencatat pengembalian barang dengan membuat dokumen <i>sales return</i> berdasarkan <i>delivery order</i> terkait.	<i>Sales return</i> bisa dibuat dengan memilih <i>delivery order</i> yang sesuai, dan prosesnya lancar.	5/5 berhasil
6	Pembuatan <i>invoice</i>	Sistem harus menghasilkan <i>invoice</i> final yang mencerminkan total nilai transaksi, termasuk pengurangan akibat retur, serta menyediakan opsi unduh nota, dan memungkinkan verifikasi setelah pembayaran pelanggan.	<i>Invoice</i> berhasil diunduh tanpa masalah dan mencerminkan hasil perhitungan nilai transaksi secara akurat. Verifikasi pembayaran dapat dilakukan tanpa kendala.	5/5 berhasil

4.2.1.2 Modul Pembelian

Pada Tabel 4.11 menunjukkan hasil pengujian *User Acceptance Testing* (UAT) untuk modul pembelian, yang mencakup lima skenario utama. Setiap skenario diuji sebanyak lima kali dan seluruhnya berhasil sesuai dengan harapan, menunjukkan bahwa sistem telah berjalan dengan baik dan mendukung proses pembelian secara lengkap dan akurat.

Tabel 4. 11 Tabel Pengujian UAT Modul Pembelian

No	Skenario	Harapan Hasil	Hasil Aktual	Percobaan
1	Pencatatan data <i>supplier</i>	Sistem harus menyimpan data <i>supplier</i> secara lengkap ke	Data <i>supplier</i> bisa tersimpan dengan	5/5 berhasil

		dalam basis data dan menampilkannya pada daftar <i>supplier</i> untuk transaksi berikutnya.	baik, meskipun kolom <i>contact person</i> dikosongkan.	
2	Pembuatan <i>Purchase order</i> berdasarkan detail barang yang dipesan dari <i>supplier</i>	Sistem harus membentuk dokumen <i>purchase order</i> yang berisi informasi <i>supplier</i> , daftar produk, jumlah dan harga.	<i>Purchase order</i> berhasil dibuat lengkap dengan detail produk dan harga yang sesuai pesanan.	5/5 berhasil
3	Pembuatan <i>receiving goods</i>	Sistem harus dapat menghasilkan dokumen <i>receiving goods</i> berdasarkan <i>purchase order</i> (PO) sebagai bagian dari proses pencatatan kedatangan barang.	<i>Receiving goods</i> dapat dibuat langsung dari <i>Purchase Order</i> tanpa kendala.	5/5 berhasil
4	Pembuatan dokumen <i>purchase return</i> jika ada barang rusak atau perlu dikembalikan.	Sistem harus mencatat pengembalian barang dengan membuat dokumen <i>purchase return</i> yang merujuk pada <i>receiving goods</i> terkait, serta mendukung pembuatan nota retur sebagai dokumen pendukung transaksi pengembalian.	<i>Purchase return</i> berhasil dibuat dengan memilih <i>receiving goods</i> yang sesuai, dan nota retur yang dihasilkan sesuai.	5/5 berhasil
5	Pencatatan <i>invoice</i>	Sistem harus menghasilkan <i>invoice</i> final yang mencerminkan total nilai transaksi sesuai barang yang diterima, termasuk	<i>Invoice</i> final tercatat dengan akurat tanpa masalah. Proses verifikasi	5/5 berhasil

		pengurangan akibat retur. Selain itu juga, memungkinkan verifikasi setelah melakukan pembayaran.	pembayaran berjalan lancar tanpa error.	
--	--	--	---	--

4.2.1.3 Modul Stok

Pada Tabel 4.12 menunjukkan hasil pengujian *User Acceptance Testing* (UAT) untuk modul stok, yang mencakup lima skenario utama. Setiap skenario diuji sebanyak lima kali dan seluruhnya berhasil sesuai dengan harapan, menunjukkan bahwa sistem telah berjalan dengan baik dan mendukung proses pengelolaan stok secara lengkap dan akurat.

Tabel 4. 12 Tabel Pengujian UAT Modul Stok

No	Skenario	Harapan Hasil	Hasil Aktual	Percobaan
1	Pencatatan stok masuk	Sistem harus mencatat penambahan stok berdasarkan dokumen <i>sales return</i> dan <i>receiving goods</i> yang telah disetujui, sehingga stok produk langsung diperbarui di sistem.	Penambahan stok tercatat dengan benar setelah input dokumen <i>sales return</i> dan <i>receiving goods</i> , stok langsung terupdate di sistem.	5/5 berhasil
2	Pencatatan stok keluar	Sistem harus mencatat pengurangan stok berdasarkan transaksi penjualan dan memberikan rekomendasi <i>batch</i> yang harus dikeluarkan berdasarkan tanggal kedaluwarsa terdekat. Selain itu, sistem juga harus mendukung pembuatan surat jalan sebagai dokumen resmi pengiriman barang.	Pengurangan stok berjalan lancar, dan sistem menampilkan rekomendasi <i>batch</i> sesuai tanggal kedaluwarsa yang tepat. Selain itu, surat jalan dapat diunduh tanpa kendala.	5/5 berhasil

3	Penyesuaian stok	Sistem harus mencatat perubahan stok akibat penyusutan atau <i>trimming</i> dengan membuat dokumen <i>stock adjustment</i> .	<i>Stock adjustment</i> dapat dibuat dengan mudah, perubahan stok tercatat akurat sesuai input penyusutan dan <i>trimming</i> .	5/5 berhasil
4	Pencatatan stok rusak	Sistem harus mencatat pengurangan stok untuk barang rusak berdasarkan laporan kerusakan, sehingga stok tercermin secara akurat.	Laporan kerusakan berhasil diinput dan stok otomatis berkurang sesuai jumlah barang rusak yang dilaporkan.	5/5 berhasil
5	Menampilkan kartu stok	Sistem harus menampilkan kartu stok yang mencerminkan jumlah stok kuantitas dan nilai setiap produk secara <i>real-time</i> dan akurat.	Kartu stok tampil dengan data yang lengkap dan <i>up-to-date</i> .	5/5 berhasil

4.2.1.4 Modul Prediksi

Pada Tabel 4.13 menunjukkan hasil pengujian *User Acceptance Testing* (UAT) untuk modul prediksi, yang mencakup dua skenario utama. Setiap skenario diuji sebanyak lima kali dan seluruhnya berhasil sesuai dengan harapan, menunjukkan bahwa sistem telah berjalan dengan baik dan mendukung proses pengadaan stok secara lengkap dan akurat.

Tabel 4. 13 Tabel Pengujian UAT Modul Prediksi

No	Skenario	Harapan Hasil	Hasil Aktual	Percobaan
1	Melakukan prediksi permintaan	Sistem menampilkan hasil prediksi dengan RMSE terendah dan memilih model terbaik secara otomatis.	Hasil prediksi yang dipilih benar dan sistem menampilkan laporan prediksi secara detail.	5/5 berhasil

2	Perhitungan <i>reorder point</i>	Sistem menampilkan daftar produk yang berada di bawah <i>reorder point</i> berdasarkan hasil prediksi.	Produk yang perlu dilakukan pemesanan ulang ditampilkan secara jelas dan akurat, serta diberikan rekomendasi jumlah pembelian berdasarkan hasil prediksi untuk membantu perencanaan stok.	5/5 berhasil
---	----------------------------------	--	---	-----------------

4.2.2 System Usability Scale (SUS)

Metode SUS digunakan untuk memberikan nilai kuantitatif terhadap pengalaman pengguna dalam menggunakan sistem. Daftar pertanyaan kuesioner dapat dilihat pada Tabel 4.14.

Tabel 4. 14 Daftar Pertanyaan SUS

No.	Pertanyaan
1	Saya rasa saya ingin sering menggunakan sistem ini.
2	Saya merasa sistem ini terlalu rumit.
3	Saya merasa sistem ini mudah digunakan.
4	Saya rasa saya memerlukan bantuan dari seseorang yang memiliki pengetahuan teknis untuk dapat menggunakan sistem ini.
5	Saya merasa berbagai fungsi dalam sistem ini terintegrasi dengan baik.
6	Saya merasa sistem ini memiliki terlalu banyak inkonsistensi.
7	Saya merasa sebagian besar orang akan bisa belajar menggunakan sistem ini dengan cepat.

8	Saya merasa sistem ini sangat membingungkan.
9	Saya merasa percaya diri saat menggunakan sistem ini.
10	Saya perlu belajar banyak sebelum bisa menggunakan sistem ini.

Masing-masing pertanyaan memiliki jawaban dengan *rating* 1 sampai 5, di mana 1 artinya sangat tidak setuju, 2 tidak setuju, 3 netral, 4 setuju, dan 5 sangat setuju. Untuk menghitung skor akhir, pada pertanyaan ganjil, nilai *rating* dikurangi 1, sedangkan pada pertanyaan genap, nilai 5 dikurangi *rating* yang diberikan. Setelah itu, semua hasil dijumlahkan, lalu dikalikan dengan 2,5 untuk mendapatkan skor SUS akhir.

Tabel 4. 15 Jawaban Responden

Nama Responden	Divisi	Jawaban Responden									
		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Lani	Penjualan	5	2	4	3	5	1	3	2	4	4
Clementine	Pembelian	5	1	4	2	5	1	2	4	5	2
Sila	Keuangan	5	1	5	2	5	1	4	1	5	2
Andik	Gudang	5	2	5	2	4	1	4	1	4	2

Tabel 4. 16 Skor Akhir Perhitungan SUS

Hasil Perhitungan										Jumlah	Nilai SUS
Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10		
4	3	3	2	4	4	2	3	3	1	29	72.5
4	4	3	3	4	4	1	1	4	3	31	77.5
4	4	4	3	4	4	3	4	4	3	37	92.5

4	3	4	3	3	4	3	4	3	3	34	85
Hasil Akhir											81,875

Hasil akhir yang didapat dari hasil pengujian dengan perhitungan SUS adalah 81,875. Hal ini membuktikan bahwa sistem yang dibuat dapat dikatakan baik.