

POLITEKNIK KEUANGAN NEGARA STAN
AKUNTANSI SEKTOR PUBLIK
MATA KULIAH: ANALITIKA DATA KEUANGAN SEKTOR PUBLIK (ADKSP)

**LAPORAN PENERAPAN TIME SERIES FORECASTING DENGAN ORANGE DATA
MINING PADA DATA FISKAL DAN MAKROEKONOMI: KASUS DATA
PENERIMAAN KEPABEANAN DAN CUKAI**



Disusun oleh

Kelompok 3 Kelas 6 Sistem Informasi-4 dengan Anggota:

- | | |
|--------------------------|------------|
| 1. Andrey Satrio Pramono | 4131230218 |
| 2. Fauzi Abdillah | 4131230058 |
| 3. Yopam Yanandra Nabel | 4131230264 |

TANGERANG SELATAN

2026

DAFTAR ISI

DAFTAR ISI	2
BAB I PENDAHULUAN DAN PEMAHAMAN KONTEKS	3
1.1 Latar Belakang Instrumen Fiskal yang Dianalisis	3
1.2 Tujuan Peramalan	3
1.3 Signifikansi bagi Manajemen Keuangan Sektor Publik.....	3
BAB II EKSPLORASI DATA DAN IDENTIFIKASI KOMPONEN DERET WAKTU	4
2.1 Deskripsi Dataset	4
2.2 Statistik Deskriptif	4
2.3 Line Chart Data Historis	4
2.4 Identifikasi Komponen.....	5
2.5 Grafik Correlogram Awal.....	5
BAB III ANALISIS STASIONERITAS DAN PRA-PEMROSESAN DATA	7
3.1 Evaluasi Stasioneritas Rata-Rata dan Varians.....	7
3.2 Pembuktian Kebutuhan Differencing.....	7
3.3 Inspeksi Grafik ACF Pasca-Differencing.....	7
BAB IV PEMBANGUNAN, EVALUASI, DAN PROYEKSI MODEL	9
4.1 Spesifikasi Konfigurasi Model yang Diuji.....	9
4.2 Perbandingan Metrik Evaluasi dan Pemilihan Model Terbaik.....	9
4.3 Analisis Residual White Noise	10
4.4 Peramalan Masa Depan.....	10
BAB V IMPLIKASI MANAJERIAL DAN REKOMENDASI KEBIJAKAN	11
LAMPIRAN	13

BAB I

PENDAHULUAN DAN PEMAHAMAN KONTEKS

1.1 Latar Belakang Instrumen Fiskal yang Dianalisis

Penerimaan kepabeanan dan cukai merupakan salah satu sumber pendapatan negara yang signifikan dalam Anggaran Pendapatan dan Belanja Negara (APBN). Cukai Hasil Tembakau (CHT) khususnya telah menjadi instrumen strategis dalam kebijakan fiskal Indonesia, tidak hanya sebagai sumber penerimaan tetapi juga sebagai alat kesehatan masyarakat untuk mengurangi konsumsi produk tembakau.

Karakteristik unik dari data penerimaan cukai adalah terdapatnya distorsi musiman dan pola pergeseran bertingkat (step-shift) yang diakibatkan oleh penyesuaian tarif yang ditetapkan pemerintah secara berkala. Sebelum tarif cukai baru diberlakukan, terjadi fenomena penimbunan pita cukai (forestalling effect) oleh produsen dan distributor, yang memicu lonjakan penerimaan signifikan pada periode sebelum kenaikan tarif dan penurunan tajam setelahnya. Fenomena ini menciptakan volatilitas dan distorsi dalam data historis yang memerlukan pemahaman mendalam untuk keperluan peramalan dan perencanaan fiskal yang akurat.

Dengan memanfaatkan teknik peramalan deret waktu (time series forecasting) menggunakan metode Autoregressive Integrated Moving Average (ARIMA) melalui platform Orange Data Mining, laporan ini bertujuan untuk mengidentifikasi komponen struktural, mengevaluasi stasioneritas, dan membangun model prediktif yang robust untuk mendukung pengambilan keputusan kebijakan fiskal di Kementerian Keuangan.

1.2 Tujuan Peramalan

Tujuan utama penulisan laporan ini adalah:

1. Mengidentifikasi dan menjelaskan komponen struktural deret waktu penerimaan cukai, termasuk tren, musiman, siklik, dan anomali.
2. Mengevaluasi stasioneritas data historis (Januari 2014 – Desember 2025) melalui analisis grafis dan Correlogram ACF/PACF.
3. Menerapkan teknik differencing secara sistematis untuk menstabilkan rata-rata dan varians deret waktu.
4. Membangun, membandingkan, dan mengevaluasi minimal tiga arsitektur model ARIMA dengan kriteria RMSE, MAE, dan MAPE.
5. Melakukan proyeksi penerimaan cukai untuk 12 bulan ke depan dengan tingkat keyakinan yang terdokumentasi.
6. Merumuskan rekomendasi kebijakan fiskal yang aplikatif dan data-driven berdasarkan hasil analisis.

1.3 Signifikansi bagi Manajemen Keuangan Sektor Publik

Hasil laporan ini diharapkan memberikan manfaat bagi:

1. Direktorat Jenderal Bea dan Cukai: Untuk perencanaan penerimaan, penetapan target, dan strategi pengawasan kepatuhan.
2. Kementerian Keuangan: Untuk penyusunan APBN dan perhitungan asumsi dasar fiskal yang lebih akurat.
3. Tim Pengendalian Inflasi (TPIP): Untuk mengantisipasi dampak lanjutan dari penyesuaian tarif cukai terhadap tingkat harga.
4. Regulator dan Pemangku Kepentingan: Untuk pemahaman mendalam tentang pola perilaku industri terkait forestalling effect.

BAB II

EKSPLORASI DATA DAN IDENTIFIKASI KOMPONEN DERET WAKTU

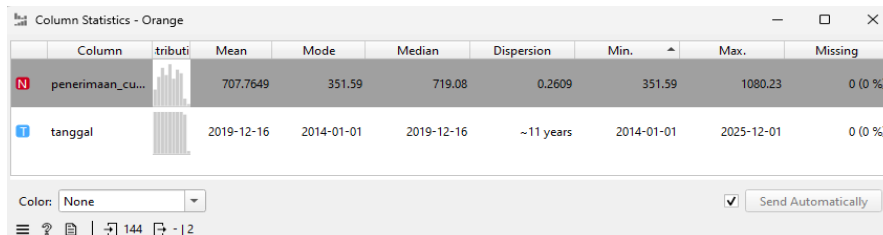
2.1 Deskripsi Dataset

Dataset yang digunakan dalam analisis ini adalah data penerimaan kepabeanan dan cukai bulanan dari Direktorat Jenderal Bea dan Cukai Kementerian Keuangan Republik Indonesia.

Atribut	Keterangan
Sumber Data	Direktorat Jenderal Bea dan Cukai
Periode Observasi	Januari 2014 – Desember 2025 (144 observasi bulanan)
Variabel Utama	Penerimaan_Cukai (dalam miliar Rupiah)
Struktur File	Kolom tanggal (YYYY-MM-DD) dan penerimaan_cukai_miliar_rp (numerik)
Frekuensi	Bulanan (12 observasi per tahun)

2.2 Statistik Deskriptif

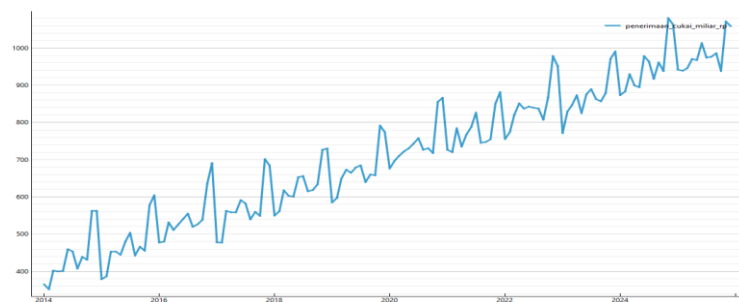
Statistik deskriptif memberikan gambaran awal tentang karakteristik data penerimaan cukai historis:



Column	tributi	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
penerimaan_cukai_miliar_rp		707.7649	351.59	719.08	0.2609	351.59	1080.23	0 (0 %)
tanggal		2019-12-16	2014-01-01	2019-12-16	~11 years	2014-01-01	2025-12-01	0 (0 %)

Berdasarkan widget Column Statistics di atas, didapat hasil bahwa rata-rata penerimaan bea cukai adalah Rp707,7649 miliar dan mediannya berada di Rp719,08 miliar dengan koefisien dispersi 0,2609 yang berarti variasi data cukup besar sebesar 26,09% dari rata-rata. Median yang lebih besar dari mean menunjukkan bahwa distribusi data lebih condong ke kiri.

2.3 Line Chart Data Historis



Data historis menunjukkan tren naik sepanjang 2014 sampai 2025 dari kisaran 350–400 di awal periode ke kisaran 950–1080 di akhir periode. Hal itu mengindikasikan pertumbuhan penerimaan cukai secara struktural dan sejalan dengan kenaikan tarif CHT dan volume konsumsi rokok legal dari tahun ke tahun. Di sisi lain, terlihat pola berulang dengan periode 12 bulan yang cukup konsisten. Setiap tahun terdapat lonjakan tajam di sekitar akhir tahun dan diikuti penurunan tajam di awal tahun berikutnya. Pola tersebut berulang di setiap pergantian tahun dan membentuk semacam tangga naik (*step-shift*). Selain siklus tahunan yang konsisten, ada indikasi fluktuasi jangka menengah, misalnya dapat dilihat

pada grafik di atas bahwa terjadi lonjakan yang makin sering dan makin tinggi sejak tahun 2020 ke atas. Itu dimungkinkan karena ada perubahan kebijakan tarif CHT yang lebih agresif pasca 2020. Tidak ada outlier dari pola. Lonjakan yang ada bukanlah anomali statistik karena selalu terjadi di periode yang sama tiap tahun.

2.4 Identifikasi Komponen

1. Tren

Data penerimaan cukai menunjukkan tren naik (uptrend) jangka panjang dari tahun 2014 hingga 2022, mencerminkan pertumbuhan ekonomi dan peningkatan basis pajak. Namun, sejak 2023 terlihat stabilisasi dengan fluktuasi di sekitar level tertinggi (plateau). Tren naik ini mencerminkan efektivitas penetapan tarif cukai yang lebih tinggi dan pertumbuhan volume konsumsi produk cukai di Indonesia.

2. Musiman

Pola musiman yang teratur dan berulang setiap tahun terlihat dalam data. Umumnya penerimaan cukai cenderung lebih rendah pada kuartal pertama (Januari–Maret) dan meningkat menjelang akhir tahun (September–Desember). Pola ini disebabkan oleh perilaku konsumen yang meningkat pada musim liburan, terutama di sekitar perayaan Hari Raya dan tahun baru.

3. Siklikal

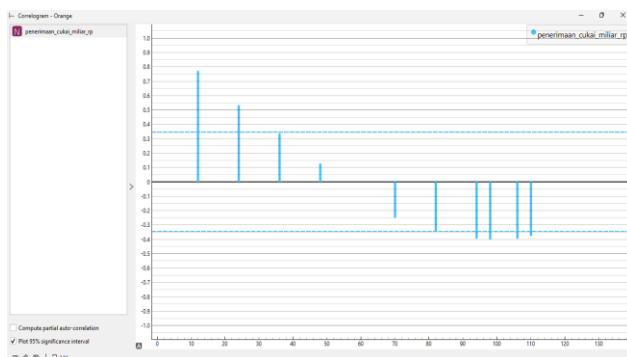
Terdapat beberapa titik perubahan struktural (structural break) yang terlihat pada grafik data historis. Lonjakan tajam terjadi sebelum tanggal efektif berlakunya penyesuaian tarif cukai baru, diikuti dengan penurunan setelahnya. Fenomena ini disebabkan oleh forestalling effect, dimana produsen dan distributor mempercepat pembelian pita cukai sebelum tarif naik untuk menghindari biaya lebih tinggi. Periode-periode kritis terlihat pada sekitar:

- Juli 2019 (Kenaikan CHT Fase I)
- Januari 2020 (Kenaikan CHT Fase II)
- Mei 2020 (Penyesuaian tarif lanjutan)

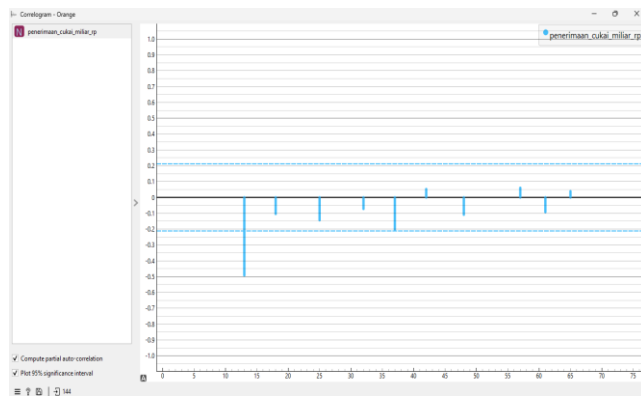
4. Anomali

Beberapa nilai ekstrem terlihat pada periode-periode sebelum implementasi kenaikan tarif cukai, mencerminkan spike penerimaan yang signifikan akibat forestalling effect. Nilai-nilai ini bukan merupakan kesalahan pengukuran tetapi representasi dari perilaku rasional pelaku industri dalam merespons perubahan kebijakan tarif.

2.5 Grafik Correlogram Awal



Nilai Autocorrelation Function (ACF) yang signifikan (melewati confidence interval) muncul di lag kelipatan 12 bulan, yaitu lag 12 sekitar 0,78, lag 24 sekitar 0,53, lag 36 sekitar 0,33, dan terus turun (decay) secara bertahap. Pada lag sekitar 70 sampai 110, nilai ACF berbalik negatif dan juga signifikan. Pola yang turun secara bertahap ini merupakan karakteristik dari data yang belum stasioner.



Nilai Partial Autocorrelation Function (PACF) signifikan hanya di lag 12 sekitar -0,5, sedangkan nilai di lag lain berada di dalam confidence interval. Pola yang naik (atau turun) tajam setelah lag 12 ini merupakan indikasi adanya struktur autoregresif musiman dengan periode 12 bulan, bahkan sebelum data di-*differencing*.

BAB III

ANALISIS STASIONERITAS DAN PRA-PEMROSESAN DATA

3.1 Evaluasi Stasioneritas Rata-Rata dan Varians

Dari Line Chart data historis, terlihat bahwa grafik mengalami tren naik secara konsisten dari tahun ke tahun. Oleh karena itu, rata-rata penerimaan kepabeanaan dan cukai tiap tahun naik tanpa ada indikasi kembali ke rata-rata awal. Tidak ada bagian data yang berfluktuasi secara stabil di sekitar nilai tertentu. Jadi, data historis tidak stasioner terhadap mean.

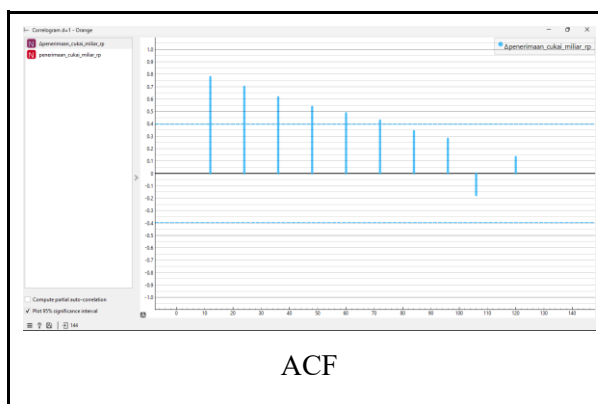
Di sisi lain, terlihat bahwa nilai puncak dan lembah tiap tahun relatif stabil dan tidak ada data per tahun yang menunjukkan selisih nilai maksimum dan minimum yang tinggi secara ekstrem. Pola naik tajam lalu turun tajam (*step-shift* musiman akhir tahun ke awal tahun) tetap berulang di setiap mean yang berbeda dan tidak beresilasi di mean yang sama karena mean tidak konsisten sepanjang waktu. Artinya varians data tidak bisa dianggap konstan dalam kerangka model karena bergantung pada tren mean yang terus bergeser naik. Jadi, data historis tidak stasioner terhadap varians.

3.2 Pembuktian Kebutuhan Differencing

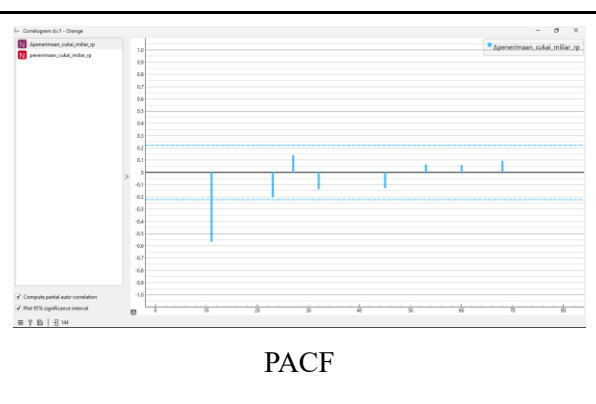
Telah dibuktikan bahwa mean dan varians data adalah konstan, maka data historis penerimaan kepabeanaan dan cukai tidak stasioner. Karena data historis tidak stasioner, diperlukan *differencing* pada data historis untuk menghilangkan tren mean yang selanjutnya akan berdampak pada kekonstanan varians. Perlu diketahui bahwa *differencing* (Δy) merupakan metode penghitungan untuk membentuk data baru dengan formula $\Delta y = y_t - y_{t-1}$.

3.3 Inspeksi Grafik ACF Pasca-Differencing

1. Differencing d-1



ACF



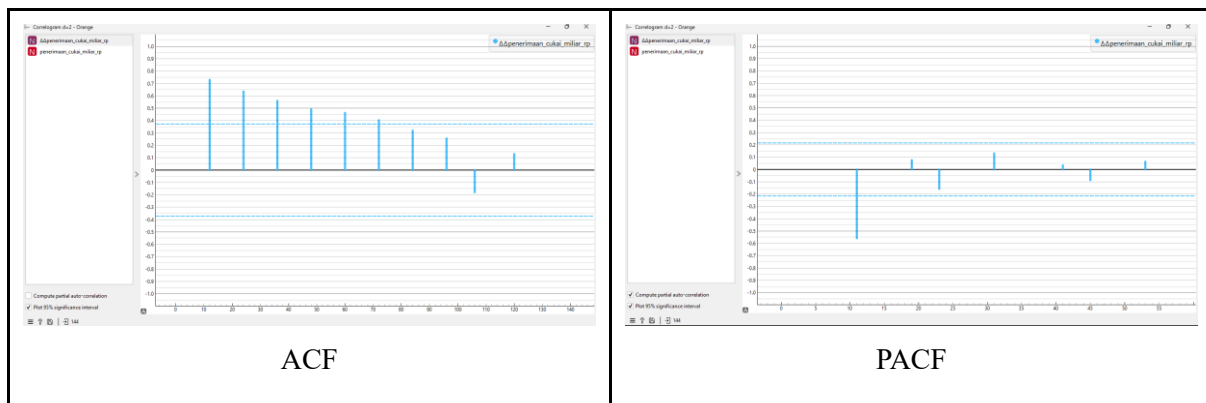
PACF

Setelah dilakukan differencing dengan $d=1$, data berfluktuasi stabil di sekitar nilai 0 sepanjang 2014 sampai 2025 tanpa tren naik ataupun turun. Ini menandakan bahwa *differencing* orde 1 dapat menghilangkan tren dan menstabilkan mean. Amplitudo fluktuasi juga relatif konsisten di seluruh periode, meski ada beberapa lonjakan lebih ekstrem di 2015, 2017, dan 2023 (mencapai -180 sampai -

210). Lonjakan ekstrem ini konsisten muncul di titik pergantian tahun sesuai dengan pola yang sudah dijelaskan sebelumnya dan bukan merupakan noise acak.

Nilai ACF masih signifikan di lag kelipatan 12 (lag 12 sekitar 0,78, terus turun bertahap sampai lag 96 sekitar 0,13, lalu berbalik negatif di lag 108). Pola *decay* bertahap ini menunjukkan *differencing* $d=1$ sudah cukup menstabilkan tren non-musiman, tapi struktur musiman tahunan (lag 12) masih tersisa dan belum hilang. Ini artinya $d=1$ saja menyelesaikan masalah tren, tetapi tidak menyelesaikan masalah musiman. Di sisi lain, nilai PACF signifikan hanya di lag 12, sementara semua lag lain berada di dalam confidence interval.

2. Differencing $d=2$



Dapat dilihat dari Line Chart untuk *differencing* $d=2$ bahwa fluktuasi tetap berputar di sekitar 0, tapi amplitudonya lebih besar dibanding $d=1$. Ini artinya differencing kedua memperbesar amplitudo fluktuasi. Di sisi lain, pola nilai ACF dan PACF juga sama persis dengan yang $d=1$ sehingga *differencing* untuk $d=2$ tidak menghasilkan informasi baru. Jadi, *differencing* kedua ini tidak memberikan perbaikan struktural, tetapi justru memperbesar amplitudo fluktuasi data. Dengan kata lain, *differencing* $d=2$ menyebabkan data menjadi overdifferentencing. Oleh karena itu, dipilih $d=1$.

BAB IV PEMBANGUNAN, EVALUASI, DAN PROYEKSI MODEL

4.1 Spesifikasi Konfigurasi Model yang Diuji

Dibentuk tiga model, yaitu Model Baseline ARIMA(0,0,0), Model Integrated ARIMA(1,1,1) dan ARIMA(1,2,1), dan Model Seasonal ARIMA (12,1,0). Model Baseline ARIMA (0,0,0) tidak memiliki komponen AR, *differencing*, maupun MA ($p=0$, $d=0$, $q=0$). Model ini hanya memprediksi berdasarkan rata-rata konstan dan fungsinya sebagai pembanding minimum serta untuk membuktikan bahwa model lain memang lebih baik dari sekadar menebak rata-rata. Model Integrated ARIMA(1,1,1) dan ARIMA(1,2,1) menggunakan komponen AR orde 1 dan MA orde 1, dengan *differencing* orde 1 dan orde 2 secara berturut-turut. Model ini diuji untuk melihat apakah pada AR dan MA jangka pendek cukup menjelaskan pola data, sekaligus membandingkan efek *differencing* orde 1 melawan orde 2 terhadap akurasi model. Selanjutnya, Model Seasonal ARIMA(12,1,0) menggunakan $p=12$, $d=1$, $q=0$. Nilai $p=12$ dipilih karena data bersifat bulanan dengan pola musiman tahunan dan dari correlogram PACF $d=1$, cuma lag 12 yang signifikan melewati batas confidence interval, sementara lag lain berada di dalam interval. Ini menandakan bahwa struktur autoregresif musiman berada tepat di lag 12, bukan di lag lain. Nilai $q=0$ dipilih karena pada correlogram ACF $d=1$, nilai di lag 12 dan kelipatannya turun secara bertahap, tanpa penurunan atau kenaikan tajam. Karena tidak ada pola *cut-off* yang tajam pada ACF, komponen MA tidak diperlukan.

4.2 Perbandingan Metrik Evaluasi dan Pemilihan Model Terbaik

Evaluation Parameters	RMSE	MAE	MAPE	POCID	R ²	AIC	BIC
ARIMA(1,1,1)	63.3	37.4	0.054	40.7	0.523	907.9	915.2
ARIMA(1,1,1) (in-sample)	62.0	28.2	0.059	31.5	0.887	1554.1	1563.0
ARIMA(1,2,1)	82.8	51.3	0.075	52.5	0.184	917.7	924.9
ARIMA(1,2,1) (in-sample)	70.4	22.1	0.063	32.9	0.855	1580.7	1589.6
ARMA(0,0,0)	253.0	237.1	0.277	27.1	-6.633	1043.5	1048.4
ARMA(0,0,0) (in-sample)	184.7	149.4	0.222	0.00	-2.22e-16	1915.6	1921.6
ARIMA(12,1,0)	35.7	23.7	0.033	72.9	0.848	811.6	843.0
ARIMA(12,1,0) (in-sample)	43.4	18.5	0.036	55.9	0.945	1405.9	1444.4

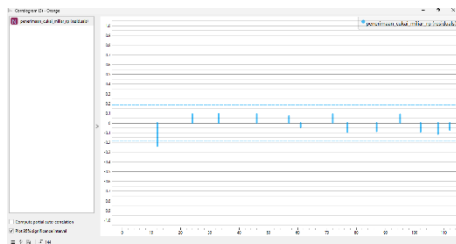
Model ARIMA(12,1,0) mencatat RMSE terendah, 35,7, jauh di bawah ARIMA(1,1,1) sebesar 63,3, ARIMA(1,2,1) sebesar 82,8, dan ARMA(0,0,0) sebesar 253,0. Ini membuktikan bahwa error prediksi model musiman jauh lebih kecil dibanding model lain.

Selanjutnya, ARIMA(12,1,0) mencatat MAE 23,7, dibanding ARIMA(1,1,1) sebesar 37,4, ARIMA(1,2,1) sebesar 51,3, dan ARMA(0,0,0) sebesar 237,1. Jika dilihat selisih rata rata error absolut ARIMA(12,1,0) terhadap model kedua terbaik, selisihnya bahkan mencapai hampir 14 miliar rupiah per periode, angka yang signifikan dalam konteks penerimaan cukai bulanan.

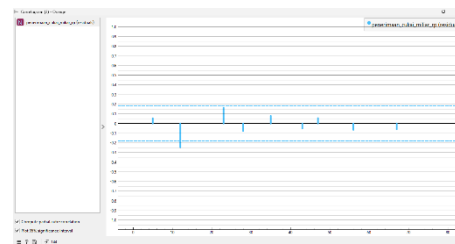
Berikutnya, ARIMA(12,1,0) mencatat MAPE 0,033 atau setara 3,3%, jauh lebih akurat dibanding ARIMA(1,1,1) sebesar 5,4%, ARIMA(1,2,1) sebesar 7,5%, dan ARMA(0,0,0) sebesar 27,7%. Tingkat kesalahan di bawah 5 persen umumnya dianggap sangat akurat dalam peramalan deret waktu.

Dari ketiga metrik tersebut, Model ARIMA(12,1,0) dapat menghasilkan prediksi jauh lebih akurat dibanding model yang lainnya. Lebih lanjut, diperhatikan bahwa model yang baik ditandai dengan residual yang berperilaku seperti white noise, artinya tidak ada pola atau autokorelasi yang tersisa di dalamnya. Jika masih terdapat struktur pada residual, berarti ada informasi dalam data yang belum bisa ditangkap model.

4.3 Analisis Residual White Noise



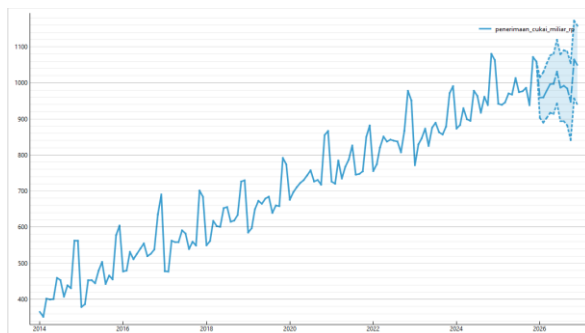
ACF Residual ARIMA(12,1,0)



PACF Residual ARIMA(12,1,0)

Berdasarkan gambar ACF Residual, hampir semua lag berada di dalam confidence interval, tidak ada pola decay atau struktur berulang seperti yang terlihat di correlogram data mentah. Cuma satu titik yang sedikit melewati batas, yaitu lag 12, dengan nilai sekitar -0,22. Pola serupa terlihat pada grafik PACF Residual. Hampir seluruh lag berada di dalam confidence interval. Hanya lag 12 yang menyentuh tepi batas bawah di rentang -0,2 sampai -0,3 dekat dengan batas signifikansi. Karena hampir seluruh lag di kedua Correlogram berada di dalam batas signifikansi, residual model ARIMA(12,1,0) secara umum sudah mendekati white noise. Model berhasil menangkap sebagian besar struktur data, termasuk pola tren dan pola musiman. Namun, satu lag yang melewati batas di kedua Correlogram tersebut menunjukkan bahwa masih ada sisa kecil struktur musiman tahunan yang belum sepenuhnya tertangkap model, meski nilainya kecil. Hal itu bukan berarti model gagal, tetapi indikasi model sudah mendekati optimal.

4.4 Peramalan Masa Depan



	penerimaan_cukai_miliar_rp (forecast)	penerimaan_cukai_miliar_rp (95%CI low)	penerimaan_cukai_miliar_rp (95%CI high)
1	958.907	902.623	1015.19
2	959.503	889.637	1029.37
3	978.341	902.869	1053.81
4	996.27	916.651	1075.89
5	997.123	913.734	1080.51
6	1030.94	942.925	1118.96
7	986.369	893.499	1079.24
8	991.859	893.562	1090.16
9	984.127	881.745	1086.51
10	947.779	841.415	1054.14
11	1064.41	956.438	1172.37
12	1048.1	938.537	1157.66

Tabel di atas adalah hasil proyeksi 12 bulan kedepan (Januari–Desember 2026) untuk penerimaan kepabeanan dan cukai dengan menggunakan Model ARIMA(12,1,0) yang melanjutkan pola historis secara konsisten. *Confidence interval* melebar seiring bertambahnya horizon waktu. Hal itu merupakan sesuatu yang wajar karena ketidakpastian proyeksi bertambah seiring menjauhnya jarak dari data historis yang terakhir.

Model ini tetap perlu dikombinasikan dengan informasi kebijakan tarif CHT terkini dari otoritas fiskal karena mengingat angka aktual berpotensi melampaui batas atas interval kepercayaan pada bulan bulan tertentu, khususnya menjelang akhir tahun, jika terjadi lonjakan forestalling akibat kebijakan tarif baru.

BAB V

IMPLIKASI MANAJERIAL DAN REKOMENDASI KEBIJAKAN

1. Bagaimana pola pergeseran bertingkat (step-shift) terlihat pada grafik penerimaan cukai historis?

Bukti angka Rp350–400 miliar pada awal periode dibandingkan Rp950–1.080 miliar pada akhir periode, ditambah pola lonjakan-penurunan tahunan, menunjukkan bahwa penerimaan mengalami pergeseran level secara bertahap atau *step-shift*.

kenaikan yang terjadi tidak berbentuk garis lurus. Grafik menunjukkan pola seperti tangga, yaitu terjadi kenaikan tajam pada akhir tahun, kemudian penurunan kembali pada awal tahun berikutnya, sebelum akhirnya bergerak menuju level yang lebih tinggi.

Hal ini juga didukung oleh statistik deskriptif. Rata-rata penerimaan selama periode pengamatan adalah Rp707,7649 miliar, sedangkan median sebesar Rp719,08 miliar. Koefisien dispersi sebesar 0,2609 atau 26,09% menunjukkan bahwa terdapat variasi yang cukup besar dalam data penerimaan.

Selain itu, laporan mengidentifikasi adanya titik perubahan struktural yang berkaitan dengan kebijakan cukai, terutama sekitar:

- Juli 2019 – Kenaikan CHT Fase I
- Januari 2020 – Kenaikan CHT Fase II
- Mei 2020 – Penyesuaian tarif lanjutan

2. Apakah penerapan differencing $d = 1$ mampu menstabilkan lonjakan tarif berkala tersebut?

$d = 1$ berhasil menstabilkan tren, tetapi belum sepenuhnya menghilangkan lonjakan musiman berkala. Bukti utamanya adalah ACF lag 12 yang masih sebesar sekitar 0,78. Oleh karena itu, diperlukan model yang mampu menangkap pola musiman tersebut.

Berdasarkan hasil Line Chart setelah differencing $d = 1$, differencing orde pertama terbukti mampu menghilangkan tren jangka panjang. Sebelum differencing, data menunjukkan tren naik yang kuat dari tahun 2014 hingga 2025. Setelah dilakukan differencing $d = 1$, data berubah menjadi fluktuasi yang bergerak di sekitar nilai 0.

Hal ini menunjukkan bahwa tren kenaikan jangka panjang berhasil dihilangkan, mean data menjadi lebih stabil, dan amplitudo fluktuasi relatif lebih konsisten.

Namun, differencing $d = 1$ belum sepenuhnya menghilangkan pola lonjakan berkala. Buktinya terlihat pada hasil ACF setelah differencing. Nilai ACF masih signifikan pada pola kelipatan 12 bulan:

- Lag 12 $\approx 0,78$
- kemudian terus mengalami penurunan secara bertahap;
- hingga sekitar lag 96 $\approx 0,13$
- dan baru berbalik negatif sekitar lag 108.

Nilai ACF sebesar 0,78 pada lag 12 merupakan indikasi yang cukup kuat bahwa masih terdapat hubungan antarperiode dengan jarak satu tahun. Dengan kata lain, meskipun tren sudah berhasil dihilangkan, pola musiman tahunan masih bertahan.

PACF juga memperkuat hasil tersebut karena hanya menunjukkan signifikansi utama pada lag 12.

Ketika dicoba differencing $d = 2$, hasilnya justru kurang baik. Fluktuasi menjadi lebih besar, dengan rentang sekitar -200 hingga 150 dan bahkan disebut lebih liar dibandingkan $d = 1$. Kondisi ini menunjukkan indikasi *overdifferencing*.

3. Bandingkan model Moving Average $MA(q)$ dengan Autoregressive $AR(p)$. Apakah galat masa lalu (error shocks) lebih berpengaruh dibanding nilai historis penerimaan?

Berdasarkan hasil model, nilai historis penerimaan lebih berpengaruh dibandingkan galat masa lalu. Hal ini terutama terlihat dari kuatnya hubungan pada lag 12 dan terpilihnya ARIMA(12,1,0) sebagai model dengan RMSE, MAE, dan MAPE terbaik

Pada Correlogram awal, ACF menunjukkan korelasi yang kuat pada lag kelipatan 12:

- Lag 12 $\approx 0,78$
- Lag 24 $\approx 0,53$
- Lag 36 $\approx 0,33$

Nilai tersebut kemudian mengalami *decay* secara bertahap. Sementara itu, PACF menunjukkan signifikansi utama pada lag 12 sekitar -0,5. Hasil tersebut menunjukkan bahwa nilai historis penerimaan memiliki pola hubungan yang kuat, khususnya dengan penerimaan satu tahun sebelumnya.

Temuan ini kemudian diperkuat oleh hasil pemilihan model. Laporan membandingkan Baseline ARIMA(0,0,0), Integrated ARIMA(1,1,1), Integrated ARIMA(1,2,1), dan Seasonal ARIMA(12,1,0). Dari perbandingan metrik RMSE, MAE, dan MAPE, model yang dipilih sebagai Model Champion adalah ARIMA(12,1,0) karena memiliki nilai terendah pada ketiga metrik tersebut.

4. Bagaimana analisis fiskal memisahkan pertumbuhan penerimaan organik dari distorsi penimbunan pita cukai akhir tahun (forestalling effect)?

Analisis fiskal tidak cukup hanya melihat angka penerimaan, tetapi juga durasi dan pola kenaikan. Kenaikan jangka panjang dari Rp350–400 miliar menjadi Rp950–1.080 miliar mencerminkan pertumbuhan struktural, sedangkan lonjakan sesaat yang langsung turun kembali lebih tepat dikategorikan sebagai efek sementara seperti *forestalling effect*.

Berdasarkan grafik historis, tidak semua kenaikan penerimaan dapat langsung dianggap sebagai pertumbuhan penerimaan yang sebenarnya (organic growth). Laporan menunjukkan bahwa terdapat pola musiman:

- Awal tahun (Jan–Mar): penerimaan cenderung rendah.
- Akhir tahun (Sep–Des): penerimaan meningkat.
- Menjelang perubahan tarif: terjadi lonjakan tajam, biasanya diikuti penurunan.

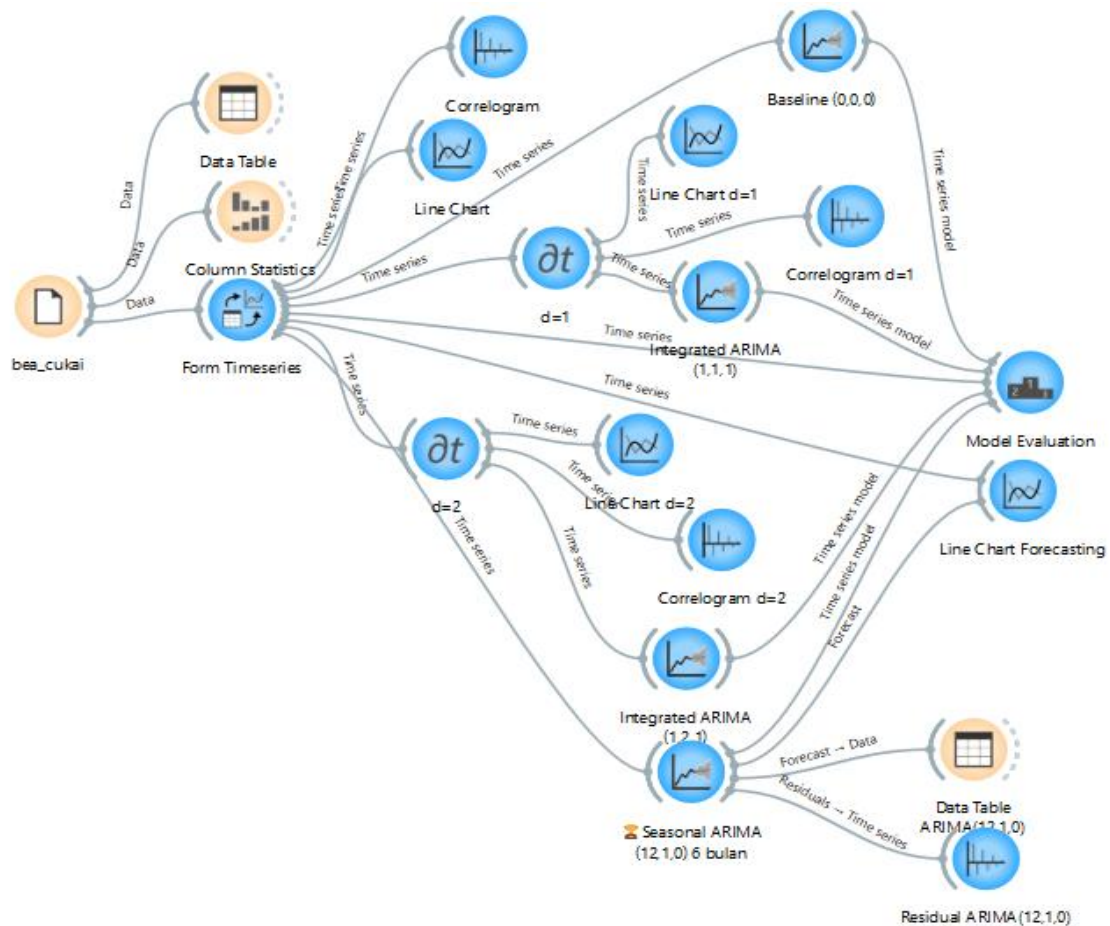
Untuk memisahkannya, analisis fiskal dapat menggunakan hasil analisis time series sebagai berikut.

Pendekatan Analisis Time Series

1. Tren jangka panjang Pertumbuhan organik terlihat dari kenaikan konsisten level penerimaan, misalnya dari Rp350–400 miliar ke Rp950–1.080 miliar. Kenaikan berkelanjutan ini lebih mencerminkan pertumbuhan nyata dibanding lonjakan singkat.
2. Identifikasi lonjakan berulang Lonjakan menjelang akhir tahun atau sebelum tarif naik harus dikategorikan sebagai policy-driven spike, bukan pertumbuhan organik. Contoh: Juli 2019, Januari 2020, dan Mei 2020, di mana terjadi *forestalling effect* akibat percepatan pembelian pita cukai.
3. Perbandingan dengan baseline ARIMA(12,1,0) Model ARIMA digunakan untuk memetakan pola normal penerimaan. Jika data aktual jauh di atas baseline, terutama menjelang perubahan tarif, selisih tersebut dapat ditafsirkan sebagai distorsi forestalling, bukan pertumbuhan normal.

LAMPIRAN

Workflow Orange



Keterangan koneksi antar-widget:

Widget File menjadi sumber data, terhubung ke tiga widget, yaitu Data Table (menampilkan data mentah dalam bentuk tabel), Column Statistics (menghasilkan ringkasan statistik deskriptif), dan Form Timeseries (mengubah data tabel biasa menjadi objek deret waktu yang bisa diproses widget widget time series lainnya). Dari Form Timeseries, data dimasukkan ke Line Chart (visualisasi tren historis) dan Correlogram (menampilkan ACF dan PACF data mentah sebelum differencing). Form Timeseries juga terhubung ke widget Difference (d=1), yang hasilnya diteruskan ke dua widget, yaitu Line Chart d=1 (visualisasi data pasca differencing pertama) dan Correlogram d=1 (ACF dan PACF pasca differencing pertama). Form Timeseries juga terhubung langsung ke widget Difference (d=2), dengan hasil diteruskan ke Line Chart d=2 dan Correlogram d=2. Terdapat empat model ARIMA dari input yang berbeda, yaitu ARIMA (0,0,0) yang menerima input langsung dari Form Timeseries (data belum didifferencing), ARIMA (1,1,1) yang menerima input dari hasil Difference (d=1), ARIMA (1,2,1) yang menerima input dari hasil Difference (d=2), dan ARIMA (12,1,0), model terbaik, yang menerima input langsung dari Form Timeseries. Output keempat model tersebut dimasukkan ke widget Model Evaluation yang juga menerima input langsung dari Form Timeseries sebagai data uji. Sebagai model terbaik, ARIMA (12,1,0) dihubungkan ke Line Chart dan Correlogram dengan input berupa residual dari model tersebut.