

Prediksi Keterlambatan Penerbangan Menggunakan Model Regresi Logistik, SES, dan SARIMA pada Data Maskapai AS

Flight Delay Prediction on U.S. Airline Data Using Logistic Regression SES, and SARIMA Models

Talitha Celin Widjaja, Evelin Alim Natadjaja*, Angeline Octavia Liemanto

Program Studi Sistem Informasi Bisnis, Universitas Ciputra Surabaya

**enatadjaja@student.ciputra.ac.id*

Abstrak: Keterlambatan penerbangan menjadi suatu permasalahan umum yang dialami oleh maskapai penerbangan di berbagai negara, termasuk Amerika Serikat. Penelitian ini menganalisis keterlambatan penerbangan menggunakan data performa maskapai penerbangan yang diambil dari Biro Statistik Transportasi Amerika Serikat. Data tersebut menyajikan informasi umum tentang penerbangan, bandar udara, jumlah penerbangan, serta faktor penyebab keterlambatan seperti cuaca, maskapai, sistem nasional (NAS), keamanan, dan keterlambatan pesawat sebelumnya. Menurut Fauziah (2022), untuk menjaga efisiensi operasional, meningkatkan layanan penerbangan, serta memastikan kepuasan penumpang dibutuhkan faktor kunci yaitu penjadwalan yang optimal dalam transportasi udara. Metode penelitian yang digunakan adalah model regresi logistik untuk melakukan analisis klasifikasi status keterlambatan dan model SES serta SARIMA untuk melakukan peramalan pada jumlah keterlambatan di bulan berikutnya. Hasil analisis menunjukkan bahwa model regresi logistik mampu melakukan klasifikasi penerbangan berdasarkan status keterlambatannya dengan akurasi 81%, presisi 83%, sensitivitas 90%, dan F1-score 86% dengan faktor penyebab yang paling berkontribusi adalah keterlambatan pesawat sebelumnya dan maskapai. Selain itu, model *Single Exponential Smoothing* (SES) digunakan untuk melakukan peramalan rata-rata keterlambatan selama 1 bulan ke depan. Hasilnya menunjukkan performa yang sangat baik dengan nilai MSE sebesar 47.044,64 dan MAPE sebesar 8,6% (0,086), yang mengindikasikan tingkat akurasi yang tinggi. Sementara itu, untuk peramalan jangka panjang selama 6 bulan ke depan, model SARIMA memberikan hasil yang lebih baik dengan MSE sebesar 44.303.781,25 dan MAPE sebesar 19,50% (0,1950), yang masih menunjukkan bahwa model mampu melakukan peramalan dengan baik. Temuan ini diharapkan dapat memberikan pengetahuan yang dapat digunakan untuk pengambilan keputusan dan penyusunan strategi yang lebih baik untuk meminimalisir keterlambatan penerbangan oleh maskapai penerbangan.

Kata Kunci: keterlambatan penerbangan, regresi logistik, SES, analisis, peramalan

Abstract: Flight delays are a common issue faced by airlines in many countries, including the United States. This study analyzes flight delays using airline performance data obtained from the U.S. Bureau of Transportation Statistics. The dataset provides general information about flights, airports, number of flights, and contributing delay factors such as weather, airline, National Aviation System (NAS), security, and late-arriving aircraft. According to Fauziah (2022), maintaining operational efficiency, improving airline services, and ensuring passenger satisfaction require a key factor: optimal scheduling in air transportation. The research methods used include logistic regression for classifying delay status and SES and SARIMA models for forecasting the number of delays in the following months. The analysis results show that the logistic regression model is capable of classifying flights based on delay status with an accuracy of 81%, precision of 83%, recall of 90%, and an F1-score of 86%, with the most contributing factors being late-arriving aircraft and airline delays. In addition, the Single

Exponential Smoothing (SES) model was used to forecast the average delay for the next month, showing excellent performance with an MSE of 47,044.64 and a MAPE of 8.6% (0.086), indicating a high level of accuracy. For long-term forecasting over the next six months, the SARIMA model produced even better results with an MSE of 44,303,781.25 and a MAPE of 19.50% (0.1950), demonstrating that the model can still provide reliable forecasts. These findings are expected to offer insights that can be used for decision-making and developing better strategies to minimize flight delays by airlines.

Keywords: flight delay, logistic regression, SES, analysis, forecasting

Naskah diterima 8 Juni 2025; direvisi 8 Juni 2025; dipublikasi 8 Juni 2025.
JUISI is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



1. Pendahuluan

Transportasi udara merupakan salah satu moda transportasi utama yang digunakan oleh masyarakat modern, khususnya untuk perjalanan jarak jauh antar kota maupun antar negara. Kecepatan, kenyamanan, dan ketepatan waktu menjadikannya pilihan unggulan dibanding moda transportasi lainnya. Selain mempercepat mobilitas manusia dan barang, transportasi udara juga berperan penting dalam mendukung pertumbuhan ekonomi, pengembangan pariwisata, serta memperluas konektivitas antar wilayah (Sembiring, 2020). Penjadwalan yang optimal dalam transportasi udara menjadi faktor kunci dalam menjaga efisiensi operasional, meningkatkan keandalan layanan penerbangan, serta memastikan kepuasan penumpang (Fauziah & Boni Wijaya, 2022). Namun demikian, fenomena keterlambatan penerbangan (*flight delay*) masih menjadi permasalahan yang sering terjadi, yang disebabkan oleh berbagai faktor seperti cuaca buruk, masalah operasional maskapai, kendala pada sistem lalu lintas udara nasional (NAS), isu keamanan, serta keterlambatan penerbangan sebelumnya (BTS | OT Delay, n.d.).

Keterlambatan penerbangan memberikan dampak yang luas, tidak hanya bagi penumpang sebagai konsumen, tetapi juga bagi maskapai penerbangan dan pengelola bandara. Dampak tersebut meliputi kerugian ekonomi yang signifikan, seperti meningkatnya biaya operasional akibat perubahan jadwal, terganggunya rotasi armada pesawat, serta menurunnya tingkat kepuasan dan kepercayaan pelanggan terhadap layanan yang diberikan oleh maskapai (Anupkumar, 2023). Oleh karena itu, diperlukan suatu model prediktif yang mampu mengidentifikasi potensi terjadinya keterlambatan penerbangan serta memperkirakan rata-rata durasi keterlambatan dalam periode tertentu.

Beberapa studi sebelumnya telah menerapkan metode regresi klasik dan analisis deret waktu dalam memprediksi keterlambatan penerbangan maupun variabel terkait lainnya. Misalnya, *Logistic Regression* sering digunakan dalam klasifikasi apakah penerbangan akan mengalami keterlambatan atau tidak, seperti pada penelitian oleh Tang (2021) yang mengembangkan model klasifikasi untuk prediksi keterlambatan penerbangan menggunakan data historis dan faktor cuaca dengan akurasi tinggi. Sedangkan untuk proyeksi durasi keterlambatan, analisis deret waktu seperti ARIMA dan model eksponensial smoothing telah menjadi pendekatan yang umum dipakai. Dalam penelitian yang dilakukan oleh Chai (2021), model ARIMA untuk memprediksi lalu lintas udara di Bandara Internasional Hong Kong, yang secara tidak langsung mencerminkan tren keterlambatan penerbangan, dengan hasil prediksi yang akurat dalam interval kepercayaan 95%.

Didukung oleh literatur terkait dan data historis dari *U.S. Department of Transportation – Bureau of Transportation Statistics* (BTS), penelitian ini mengintegrasikan metode *Logistic Regression* untuk memprediksi kemungkinan keterlambatan penerbangan, serta analisis deret waktu (*SES* dan *SARIMA*) untuk memperkirakan rata-rata durasi keterlambatan pada periode berikutnya. Pendekatan ini memungkinkan evaluasi yang komprehensif terhadap risiko dan dampak keterlambatan penerbangan secara praktis dan aplikatif.

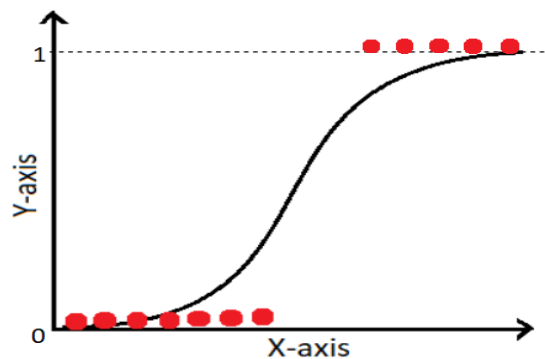
Untuk analisis deret waktu, fokus penelitian ini terbatas pada satu bandara tanpa membedakan maskapai, guna memberikan cakupan prediksi yang lebih luas dan aplikatif. Sebagai studi kasus untuk pendekatan deret waktu, penelitian ini menggunakan data dari Bandara Internasional *Hartsfield–Jackson Atlanta* (ATL), yang dikenal sebagai bandara tersibuk di dunia berdasarkan jumlah penumpang dan volume penerbangan tahunan (Freya Parsons, 2024). Pemilihan bandara ini didasarkan pada volume data yang besar, keragaman penerbangan, dan relevansinya dalam mewakili tantangan operasional bandara berskala global. Meskipun contoh dalam penelitian ini difokuskan pada Bandara Atlanta, model deret waktu yang dibangun bersifat fleksibel dan dapat diterapkan pada bandara lain sesuai kebutuhan pengguna.

Melalui analisis prediktif, penelitian ini bertujuan mengidentifikasi kontribusi berbagai faktor—seperti cuaca, jumlah penerbangan, dan penyebab keterlambatan lainnya—terhadap probabilitas dan perkiraan durasi keterlambatan penerbangan. Pemahaman mendalam terhadap faktor-faktor tersebut diharapkan dapat merumuskan strategi antisipatif yang lebih responsif dan efektif untuk meminimalkan dampak keterlambatan penerbangan di masa depan, meningkatkan efisiensi manajemen bandara, serta memperbaiki pengalaman penumpang secara menyeluruh.

2. Kajian Pustaka

2.1 Logistic Regression

Model analisis regresi merupakan metode statistik yang digunakan untuk mengevaluasi dan menggambarkan hubungan antara variabel, khususnya antara satu variabel dependen (terikat) dan satu atau lebih variabel independen (bebas), dalam suatu kerangka analisis multivariat. Salah satu bentuk regresi yang umum digunakan adalah regresi logistik, yang sangat efektif dalam menangani permasalahan klasifikasi biner, di mana variabel output hanya memiliki dua kemungkinan nilai, seperti "ya" atau "tidak", atau "terjadi" dan "tidak terjadi". Regresi logistik menggunakan fungsi sigmoid untuk mengubah nilai input menjadi probabilitas dengan rentang antara 0 dan 1. Nilai probabilitas inilah yang kemudian dijadikan dasar untuk menentukan kelas prediksi. Pendekatan ini tidak hanya mampu memprediksi hasil, tetapi juga memberikan interpretasi mengenai seberapa besar pengaruh variabel bebas terhadap kemungkinan terjadinya suatu peristiwa, sehingga menjadikannya alat analisis yang sangat berguna dalam penelitian berbasis data.



Gambar 1. Fungsi Sigmoid Regresi Logistik
Sumber: Khushwant Rai, 2020

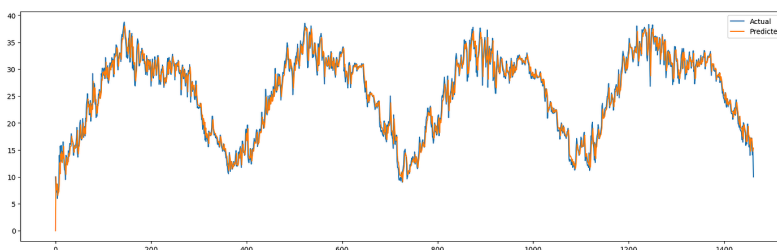
Menurut Google Developers dalam kursus *Machine Learning Crash Course*, regresi logistik menghasilkan probabilitas yang kemudian dikonversi menjadi prediksi kelas menggunakan ambang batas tertentu. Prediksi ini dibandingkan dengan label aktual untuk membentuk *confusion matrix* yang terdiri dari TP, TN, FP, dan FN. Matrix ini memungkinkan kita untuk mengevaluasi kinerja model klasifikasi biner secara menyeluruh.

	Actual positive	Actual negative
Predicted positive	True positive (TP): A spam email correctly classified as a spam email. These are the spam messages automatically sent to the spam folder.	False positive (FP): A not-spam email misclassified as spam. These are the legitimate emails that wind up in the spam folder.
Predicted negative	False negative (FN): A spam email misclassified as not-spam. These are spam emails that aren't caught by the spam filter and make their way into the inbox.	True negative (TN): A not-spam email correctly classified as not-spam. These are the legitimate emails that are sent directly to the inbox.

Gambar 2. Confusion Matrix
Sumber: <https://developers.google.com/>

2.2 SARIMA

Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) adalah pendekatan dalam analisis deret waktu yang dirancang khusus untuk menangani data dengan pola musiman. Model ini dinotasikan sebagai $ARIMA(p, d, q)(P, D, Q)^s$, dimana (p, d, q) merupakan bagian non-musiman dari model, sedangkan (P, D, Q) mewakili bagian musiman, dan s adalah jumlah periode per musim. SARIMA digunakan untuk menganalisis serta meramalkan suatu peristiwa yang terjadi dalam periode waktu tertentu secara berkala dengan mempertimbangkan baik komponen musiman maupun non-musiman dalam data. Berikut contoh grafik SARIMA yang menunjukkan pola musiman, terlihat dari data yang berulang pada grafik tersebut.



Gambar 3. Contoh Grafik SARIMA
Sumber: Tirtha Mutha, 2023

2.3 SES

Metode single exponential smoothing (SES) adalah suatu prosedur dengan mengulang perhitungan secara terus menerus menggunakan data observasi terbaru, setiap data yang digunakan dalam metode ini diberi bobot yang disimbolkan α (alpha) nilai alpha yang berkisar 0 sampai dengan 1 yang menghasilkan nilai tingkat kesalahan yang paling terkecil akan dipilih untuk digunakan dalam model peramalan (Dwi, 2016).

2.4 MAPE

Mean Absolute Percent Error merupakan perhitungan yang digunakan untuk menghitung rata-rata persentase kesalahan mutlak. MAPE memiliki rentang nilai tertentu yang dapat dijadikan tolok ukur dalam menilai seberapa baik kinerja suatu model peramalan. Nilai ini digunakan untuk mengukur tingkat kesalahan relatif antara hasil prediksi dan data aktual dalam bentuk persentase, sehingga memudahkan interpretasi. Semakin kecil nilai MAPE, maka semakin tinggi tingkat akurasi model, dan sebaliknya, nilai MAPE yang besar menunjukkan bahwa model tersebut kurang efektif dalam melakukan peramalan.

Tabel 1. *Range Nilai MAPE*

Nilai MAPE	Deskripsi
< 10%	Peramalan Model Sangat Baik
10 - 20%	Peramalan Model Baik
20 - 50%	Peramalan Model Layak
> 50%	Peramalan Model Buruk

Dalam rumus MAPE, terdapat beberapa komponen penting, yaitu n yang menyatakan jumlah pengamatan atau banyaknya periode waktu yang dianalisis, X_t yang merepresentasikan nilai aktual pada periode ke- t , serta F_t yang merupakan nilai hasil peramalan pada periode ke- t . Rumus MAPE sebagai berikut.

$$\text{MAPE} = \frac{\sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right|}{n} \times 100\%.$$

Gambar 4. Rumus MAPE

Sumber: Ngabidin, Z., Sanwidi, A., & Arini, E. R. (2023)

3. Metode Penelitian

Penelitian ini menggunakan pendekatan CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yang terdiri dari enam tahapan utama: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Pendekatan ini dipilih karena memberikan kerangka kerja yang sistematis dalam pengolahan dan analisis data, khususnya dalam konteks peramalan menggunakan metode SARIMA.

3.1 Business Understanding

Tujuan utama dari penelitian ini adalah untuk membangun model peramalan untuk memprediksi jumlah keterlambatan penerbangan berdasarkan data mengenai penyebab keterlambatan di penerbangan. Fokusnya adalah mengidentifikasi pola musiman dalam data keterlambatan penerbangan guna meningkatkan efisiensi operasional maskapai dan pengelolaan bandara kedepannya.

3.2 Data Understanding

Dataset ini mencakup informasi tentang penyebab keterlambatan penerbangan di Amerika Serikat sejak Januari 2022 hingga Februari 2025. Data diklasifikasikan ke dalam beberapa kategori penyebab keterlambatan, antara lain:

- *Carrier Delay*: Keterlambatan yang disebabkan oleh masalah internal maskapai, seperti pemeliharaan atau kru.
- *Weather Delay*: Keterlambatan akibat kondisi cuaca ekstrem.
- *National Aviation System (NAS) Delay*: Keterlambatan yang disebabkan oleh sistem navigasi udara nasional, termasuk lalu lintas udara dan operasi bandara.
- *Security Delay*: Keterlambatan yang disebabkan oleh masalah keamanan, seperti evakuasi terminal atau pemeriksaan keamanan yang memakan waktu lama.
- *Late Arriving Aircraft*: Keterlambatan yang terjadi karena pesawat tiba terlambat dari penerbangan sebelumnya.
- *Arrival Delay*: Selisih menit antara waktu kedatangan yang dijadwalkan dan waktu kedatangan yang sebenarnya. Kedatangan lebih awal menunjukkan angka negatif.

Data ini tersedia dalam format Excel dan dapat diunduh langsung dari situs BTS. Dalam tahap ini, dilakukan eksplorasi awal terhadap data untuk memahami struktur, kualitas, dan karakteristik data, termasuk identifikasi pola musiman, tren, outlier, dan missing value.

3.3 Data Preparation

Tahap *Data Preparation* merupakan proses penting dalam tahapan CRISP-DM yang bertujuan untuk memastikan data yang digunakan dalam pemodelan benar-benar bersih, valid, dan siap dianalisis. Pada penelitian ini, data yang digunakan berasal dari situs resmi *Bureau of Transportation Statistics* Amerika Serikat, yang memuat informasi bulanan terkait jumlah keterlambatan penerbangan berdasarkan penyebab keterlambatan seperti *Carrier Delay*, *Weather Delay*, *NAS Delay*, *Security Delay*, dan *Late Aircraft Delay*. Dikarenakan penelitian ini dibagi menggunakan dua metode yaitu *Time Series* dan *Logistic Regression*, tahap-tahap yang dilakukan untuk *data preparation* berbeda.

Langkah pertama yang dilakukan untuk kedua metode dalam tahap ini adalah melakukan identifikasi terhadap data yang hilang (*missing values*) maupun data yang tidak valid. Pemeriksaan dilakukan secara menyeluruh untuk mengetahui apakah terdapat nilai kosong, duplikat, atau nilai yang tidak sesuai. Apabila ditemukan data yang hilang, dilakukan penanganan dengan menghapus baris data yang mengandung nilai kosong jika persentasenya tergolong kecil dan tidak berpengaruh signifikan terhadap analisis. Kolom yang merepresentasikan informasi waktu, seperti bulan dan tahun, dikonversi menjadi format tanggal (*datetime*) agar dapat digunakan sebagai indeks dalam analisis deret waktu dalam proses *data cleaning*.

Setelah data dinyatakan bersih dan telah melalui tahap transformasi yang diperlukan, untuk metode Logistic Regression dilakukan proses *encoding* dan *feature scaling*. Lalu, dilakukan EDA atau yang biasa diketahui sebagai *Exploratory Data Analysis* untuk kedua metode, diantaranya ada menggunakan *boxplot* untuk melihat apakah ada *outlier*, *line chart*, dan juga *heatmap*. Untuk *training data* itu sendiri, data dibagi menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*). Untuk model regresi logistik dilakukan pembagian dengan proporsi 80% untuk pelatihan model dan 20% untuk pengujian. Untuk model *time series* dilakukan pembagian menjadi memprediksi bulan depan dan 6 bulan kedepan.

Kemudian untuk model SES dilakukan pembagian data dengan menggunakan data Januari 2022 hingga Januari 2025 sebagai training set dan Februari 2025 sebagai testing set untuk peramalan 1 bulan ke depan. Sedangkan untuk model SARIMA untuk prediksi jangka menengah (hingga 6 bulan ke depan), digunakan data Januari 2022 hingga Agustus 2024 untuk pelatihan dan data dari September 2024 hingga Februari 2025 untuk pengujian. Pembagian ini bertujuan agar model dapat dilatih menggunakan data historis dan dievaluasi terhadap data yang belum pernah dilihat sebelumnya untuk mengukur tingkat akurasinya.

Sebagai bagian dari persiapan akhir, dilakukan juga pemeriksaan terhadap stasioneritas data, mengingat metode SARIMA membutuhkan data yang bersifat stasioner. Jika ditemukan tren atau pola musiman yang membuat data tidak stasioner, maka dilakukan proses differencing sebanyak satu atau lebih kali hingga data memenuhi kriteria stasioner. Dengan tahapan *Data Preparation* ini, data keterlambatan penerbangan telah siap digunakan untuk tahap pemodelan dengan metode SARIMA.

3.4 Modeling

Pada tahap *Modeling*, data yang telah melalui proses *data preparation* digunakan untuk membangun model peramalan menggunakan metode Seasonal Autoregressive Integrated Moving Average (SARIMA). Metode SARIMA dipilih karena mampu menangani data deret waktu yang memiliki pola musiman, seperti halnya data keterlambatan penerbangan yang cenderung menunjukkan pola berulang pada periode tertentu, misalnya pada musim liburan atau musim dingin.

Setelah model terbentuk, dilakukan proses prediksi terhadap data uji untuk mengevaluasi performa model. Evaluasi dilakukan menggunakan metrik Mean Absolute Percentage Error (MAPE), yang memberikan gambaran mengenai tingkat kesalahan prediksi dalam bentuk persentase. Semakin rendah nilai MAPE, semakin baik model dalam melakukan prediksi terhadap data historis. Model SARIMA dipilih karena keunggulannya dalam menangkap pola musiman dan tren dalam data deret waktu, yang sangat relevan dalam konteks keterlambatan penerbangan yang biasanya dipengaruhi oleh faktor-faktor yang berulang sepanjang tahun.

3.5 Evaluation

Tahap evaluasi dalam kerangka CRISP-DM bertujuan untuk menilai kualitas dan kinerja model yang telah dibangun. Dalam penelitian ini, evaluasi dilakukan menggunakan metrik *Mean Absolute Percentage Error* (MAPE) untuk mengukur akurasi hasil peramalan. MAPE memberikan gambaran sejauh mana prediksi model mendekati nilai aktual dalam bentuk persentase kesalahan, sehingga dapat digunakan untuk menilai kelayakan model sebelum diterapkan lebih lanjut pada tahap deployment.

3.6 Deployment

Tahap deployment merupakan bagian akhir dari siklus CRISP-DM yang berfokus pada penerapan hasil analisis ke dalam konteks operasional atau pengambilan keputusan. Setelah model berhasil dibangun dan dievaluasi pada tahap sebelumnya, langkah ini bertujuan untuk menyajikan hasil peramalan secara informatif dan memastikan model dapat dimanfaatkan secara praktis.

Dalam penelitian ini, hasil peramalan disajikan dalam bentuk grafik visual yang membandingkan nilai aktual dan prediksi, sehingga memudahkan interpretasi tren keterlambatan penerbangan secara musiman. Visualisasi ini penting sebagai bentuk komunikasi hasil model kepada pemangku kepentingan non-teknis, seperti pengambil kebijakan di sektor transportasi udara.

Dengan demikian, tahap deployment tidak hanya berfungsi sebagai penyajian hasil, tetapi juga sebagai jembatan menuju implementasi model dalam skenario dunia nyata. Hasil peramalan ini dapat menjadi dasar bagi perencanaan operasional, mitigasi risiko keterlambatan, dan pengembangan sistem prediktif dalam domain transportasi udara. Tahapan ini melengkapi siklus CRISP-DM secara utuh, dari pemahaman bisnis hingga penerapan solusi berbasis data.

4. Hasil dan Pembahasan

Bagian ini membahas proses dan hasil analisis klasifikasi dan *time series* mengenai keterlambatan

penerbangan yang dilakukan dengan menggunakan model *logistic regression* dan SARIMA dengan bahasa pemrograman *Python*. Proses yang dilakukan adalah *data preparation*, *exploratory data analysis* (EDA), pemisahan *train* data dan *test* data, serta *model evaluation*.

1.1 Analisis Klasifikasi

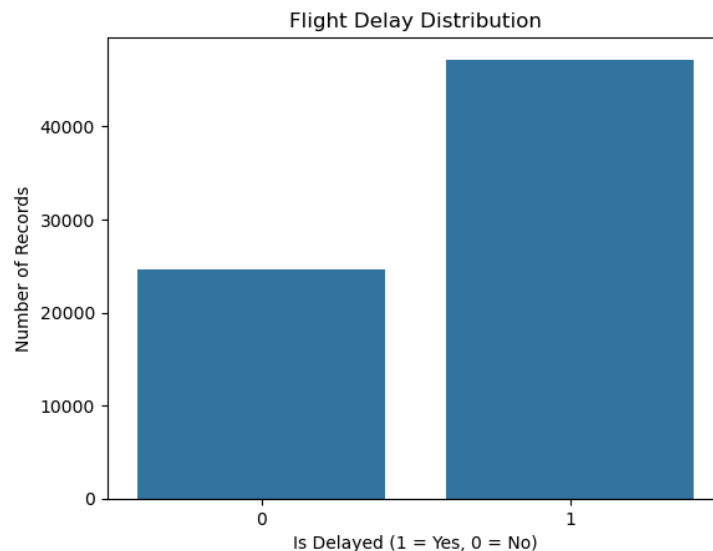
- *Data Preparation*

Sebelum membuat model, data mentah yang diambil dari sumber perlu disiapkan dahulu melalui beberapa langkah. Pertama dilakukan *loading* data dari *csv* yang dimasukkan ke dalam data frame. Setelah melakukan pengecekan struktur data, proses *data cleaning* dilakukan. Proses ini meliputi pemilihan kolom atau fitur yang akan digunakan untuk analisis yaitu kolom *year*, *month*, *carrier*, *airport*, *arr_flights*, *carrier_delay*, *weather_delay*, *nas_delay*, *security_delay*, *late aircraft delay*, dan *arr_del15*, lalu pengecekan baris yang kosong dan duplikat, serta membuat fitur baru yang akan menjadi target analisis yaitu fitur *isDelayed*.

Selanjutnya dilakukan proses *encoding* untuk fitur yang bersifat kategorikal dengan menggunakan *one-hot encoding*. Proses ini dilakukan karena model *logistic regression* hanya dapat menerima data numerik. Dengan proses tersebut, data kategorikal dapat diubah menjadi bentuk numerik agar bisa diproses oleh model. Setelah itu, variabel independen dan dependen didefinisikan untuk dilakukan *feature scaling* yang dapat menyeimbangkan skala antar fitur sehingga semua fitur berkontribusi secara seimbang dalam pelatihan model.

- *Exploratory Data Analysis*

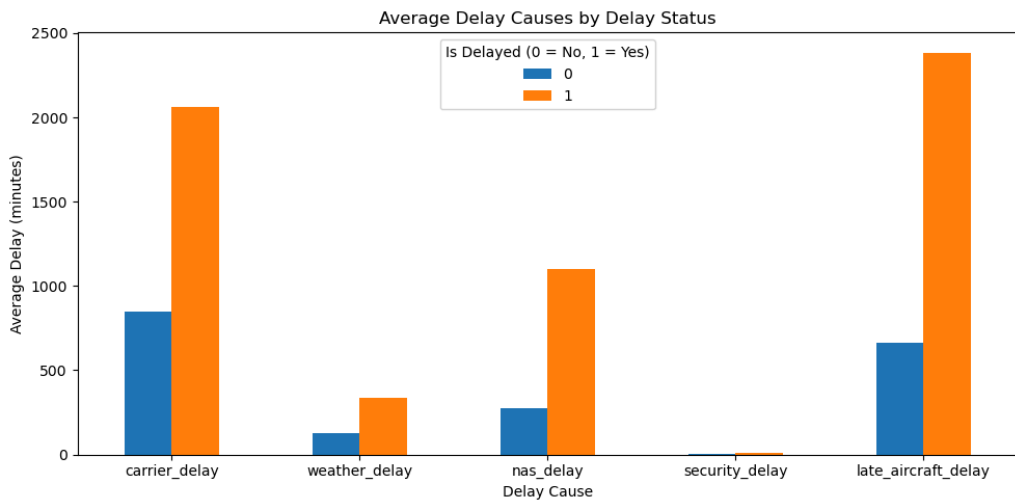
Pemahaman terhadap data menjadi kunci utama dalam membangun model sehingga perlu dilakukan *exploratory data* terlebih dahulu sebelum membuat model. Berikut ini adalah beberapa grafik yang memberikan gambaran terhadap data.



Gambar 5. Distribusi keterlambatan penerbangan

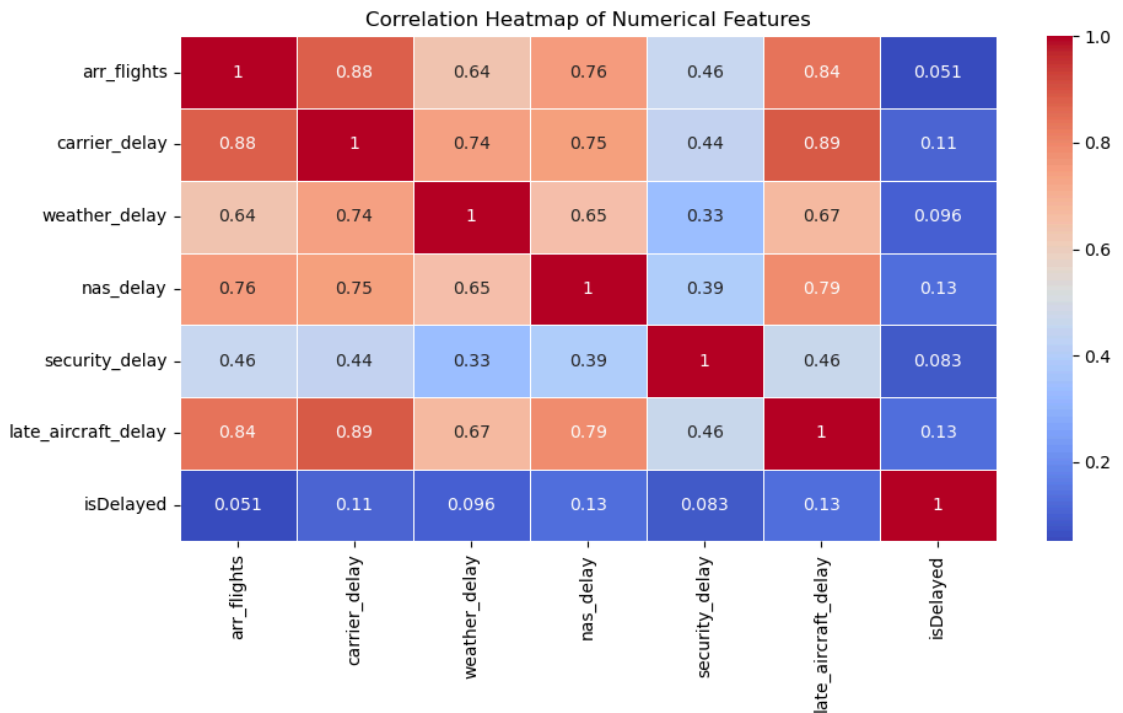
Pada Gambar 1 dilakukan visualisasi distribusi variabel target *isDelayed* yang menunjukkan apakah penerbangan mengalami keterlambatan lebih dari 15 menit. Visualisasi terhadap variabel target ini

dipilih untuk memahami keseimbangan kelas dalam dataset. Ketidakseimbangan kelas yang signifikan dapat menyebabkan model klasifikasi menjadi bias terhadap kelas mayoritas. Hasil visualisasi menunjukkan bahwa jumlah penerbangan yang mengalami keterlambatan (*isDelayed* = 1) lebih besar dibandingkan dengan penerbangan yang tidak mengalami keterlambatan (*isDelayed* = 0). Hal ini menunjukkan adanya ketidakseimbangan kelas, di mana kelas penerbangan terlambat menjadi mayoritas dalam dataset.



Gambar 6. Rata-rata Penyebab Keterlambatan

Visualisasi terhadap fitur penyebab keterlambatan pada Gambar 2 dilakukan untuk mengidentifikasi variabel yang paling berkontribusi terhadap keterlambatan penerbangan. Grafik batang menunjukkan rata-rata durasi (dalam menit) dari berbagai jenis keterlambatan pada penerbangan yang mengalami keterlambatan (*isDelayed* = 1) dibandingkan dengan yang tidak mengalami keterlambatan (*isDelayed* = 0). Hasil visualisasi ini mengonfirmasi bahwa keterlambatan akibat maskapai (*carrier delay*) dan keterlambatan pesawat sebelumnya (*late aircraft delay*) merupakan penyebab utama yang berkontribusi terhadap klasifikasi penerbangan sebagai terlambat.



Gambar 7. Heatmap Korelasi untuk Fitur Numerik

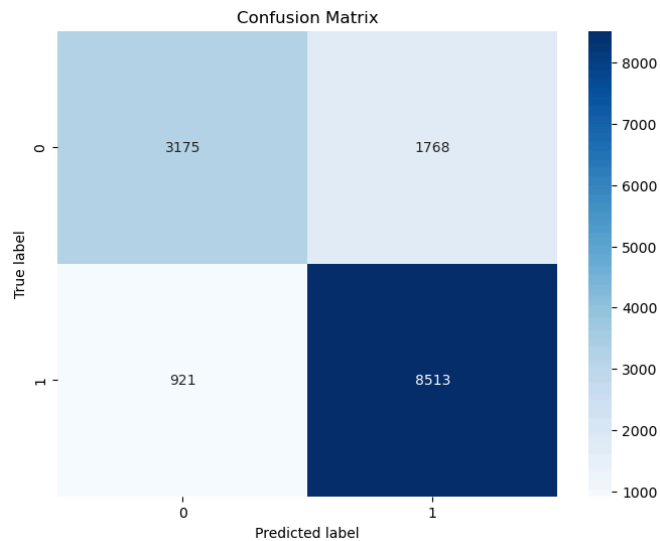
Gambar 3 menunjukkan *heatmap* korelasi yang memberikan ringkasan visual yang jelas mengenai hubungan antar fitur numerik serta hubungannya dengan variabel target *isDelayed*. Dengan menampilkan nilai korelasi antar variabel, visualisasi ini menyoroti fitur-fitur yang memiliki pengaruh lebih besar dalam memprediksi keterlambatan. Variabel-variabel ini divisualisasikan untuk mengidentifikasi prediktor yang memiliki korelasi tinggi, sehingga fitur yang kurang informatif dapat dipertimbangkan untuk dihilangkan.

- *Train Test Split*

Data yang sudah melalui proses persiapan dipisahkan menjadi dua bagian yaitu data *train* yang akan digunakan untuk pelatihan model dan data *test* yang digunakan untuk evaluasi model. Pada penelitian ini, data *train* diambil sebanyak 80% agar model dapat lebih akurat dan 20% sisanya akan digunakan untuk melakukan asesmen terhadap keakuratan model. Setelahnya, model *logistic regression* dibuat dengan menggunakan data *train* dan dilanjutkan dengan melakukan prediksi. Hasil prediksi kemudian dibandingkan dengan data *test*.

- *Model Evaluation*

Pada tahap ini, setelah melakukan prediksi dan perbandingan, dilakukan kalkulasi dari metrik akurasi, presisi, sensitivitas dan *F1-score* dengan menggunakan *confusion matrix* pada Gambar 4 dan *classification report*. Berikut adalah hasilnya



Gambar 8. *Confusion Matrix*

Berdasarkan *confusion matrix*:

- **True Positives (TP):** 8513 penerbangan berhasil diprediksi dengan benar sebagai terlambat.
- **True Negatives (TN):** 3175 penerbangan berhasil diprediksi dengan benar sebagai tidak terlambat.
- **False Positives (FP):** 1768 penerbangan diprediksi terlambat padahal sebenarnya tidak terlambat.
- **False Negatives (FN):** 921 penerbangan diprediksi tidak terlambat padahal sebenarnya terlambat.

Tabel 2. Classification Report

Report	<i>precision</i>	<i>recall</i>	f1-score	support
0	0.78	0.64	0.70	4943
1	0.83	0.90	0.86	9434
<i>accuracy</i>			0.81	14377
<i>macro avg</i>	0,80	0.77	0.78	14377
<i>weighted avg</i>	0.81	0.81	0.81	14377

Berikut adalah interpretasi dari hasil evaluasi model:

- **Akurasi** (0,81) menunjukkan bahwa model berhasil mengklasifikasikan 81% dari seluruh data uji dengan benar.
- **Presisi** (0,83) untuk kelas positif (*isDelayed* = 1) mengindikasikan bahwa 83% dari penerbangan yang diprediksi mengalami keterlambatan memang benar-benar terlambat, yang mencerminkan keandalan model dalam memprediksi keterlambatan.
- **Recall** (0,90) sangat tinggi, yang berarti model berhasil menangkap 90% dari seluruh penerbangan yang benar-benar mengalami keterlambatan. Hal ini penting dalam konteks prediksi

- keterlambatan, karena melewatkan keterlambatan aktual dapat berdampak besar.
- **F1-score** (0,86), sebagai rata-rata harmonis dari presisi dan recall, menunjukkan keseimbangan yang baik antara kemampuan model dalam mendeteksi keterlambatan sebenarnya dan meminimalkan prediksi keterlambatan yang salah.

1.2 Analisis *Time Series*

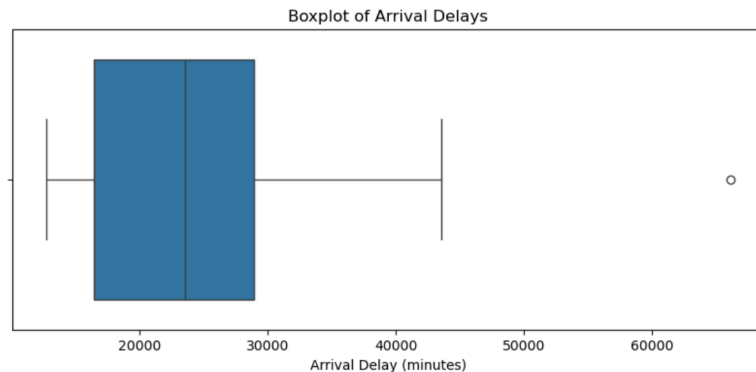
- *Data Preparation*

Sebelum melakukan pemodelan, data mentah yang diperoleh dari sumber perlu dipersiapkan melalui beberapa tahap. Pertama, data dimuat dari file CSV ke dalam data frame dan dilakukan pengecekan struktur serta tipe data untuk memastikan kelengkapan informasi. Selanjutnya, dilakukan *data cleaning* dengan pertama menyeleksi kolom yang relevan untuk analisis deret waktu, yakni kolom *year*, *month*, *airport*, dan *arr_delay*. Data kemudian di filter untuk fokus pada satu bandara tertentu, yaitu Bandara Internasional Hartsfield–Jackson Atlanta dengan kode *ATL*.

Tahapan berikutnya meliputi pemeriksaan nilai kosong dan duplikasi baris untuk menjaga kualitas data. Setelah itu, data diolah dengan mengelompokkan rata-rata keterlambatan kedatangan per bulan, diikuti dengan pembentukan kolom tanggal berbentuk *datetime* dari gabungan kolom *year* dan *month*. Kolom *datetime* ini kemudian dijadikan sebagai indeks waktu dalam data frame, sehingga siap untuk dianalisis menggunakan metode *time series*. Kolom *year* dan *month* yang sudah terintegrasi ke dalam indeks *datetime* kemudian dihapus agar tidak redundan dalam pemodelan. Terakhir dilakukan uji stasioneritas menggunakan metode Augmented Dickey-Fuller (ADF) untuk memastikan apakah data memerlukan proses differencing. Uji ADF menghasilkan p-value sebesar 0.0086 yang lebih kecil dari 0.05, sehingga data dapat dianggap stasioner dan tidak memerlukan proses differencing.

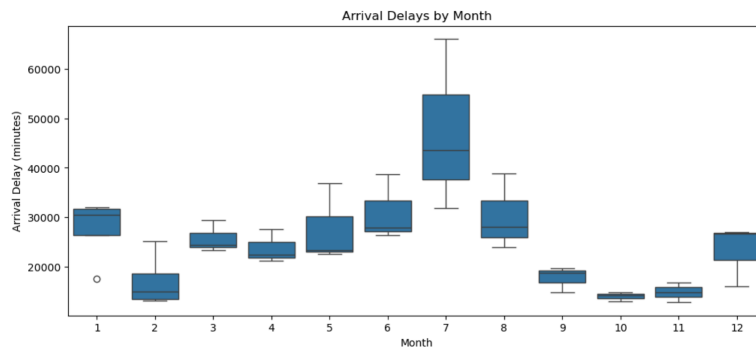
- *Exploratory Data Analysis*

Memahami data secara mendalam merupakan langkah krusial dalam pengembangan model, sehingga eksplorasi data secara menyeluruh perlu dilakukan sebelum membangun model. Visualisasi berikut menggunakan variabel target *arr_delay* (keterlambatan kedatangan dalam menit) karena variabel tersebut merupakan target utama untuk model peramalan deret waktu kami. Pemahaman terhadap distribusinya membantu kami mengevaluasi pola keterlambatan secara keseluruhan, mendeteksi pencilan, serta mengidentifikasi kemencengan atau ketidakteraturan dalam data yang mungkin mempengaruhi kinerja model. Berikut adalah beberapa visualisasi yang memberikan *insight* mengenai karakteristik target data.



Gambar 9. *Boxplot* Keterlambatan Kedatangan dalam Menit

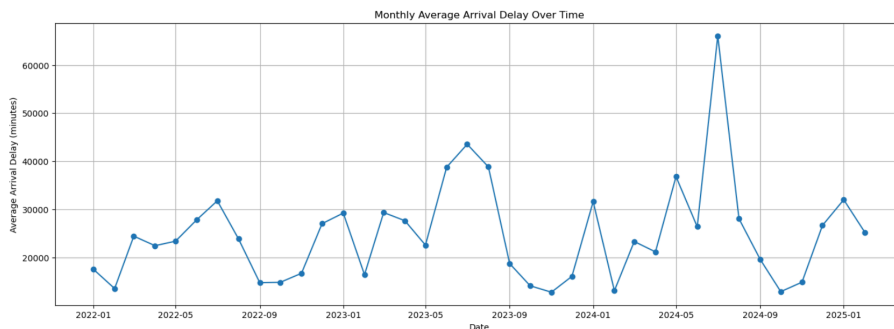
Pada Gambar 1. dilakukan visualisasi boxplot yang memperlihatkan berapa lama keterlambatan kedatangan dalam satuan menit. Sebagian besar keterlambatan berada di antara 15.000 hingga 35.000 menit, dengan nilai tengah sekitar 25.000 menit. Garis-garis menunjukkan rentang penuh keterlambatan, yaitu dari sekitar 10.000 hingga 45.000 menit. Ada juga satu keterlambatan yang sangat besar (sekitar 65.000 menit) yang ditampilkan sebagai titik, yang disebut outlier. Ini berarti nilainya jauh lebih tinggi dari yang lain dan tidak normal.



Gambar 10. *Boxplot* Keterlambatan Kedatangan Berdasarkan Bulan

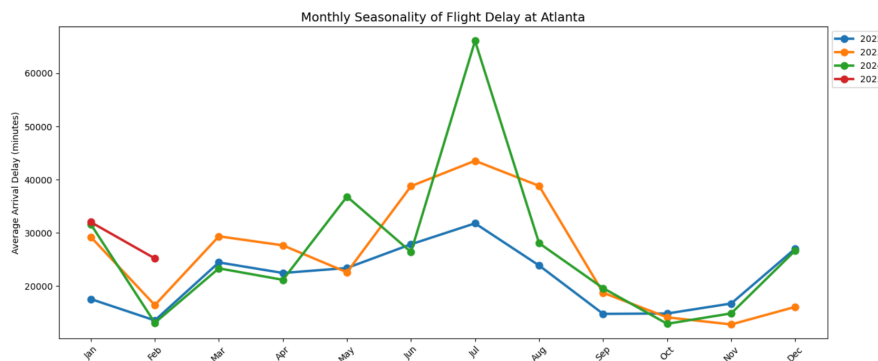
Pada Gambar 2. dilakukan visualisasi keterlambatan kedatangan berdasarkan bulan, terlihat variasi bulanan yang signifikan dalam keterlambatan kedatangan penerbangan. Juli mencatat keterlambatan median tertinggi dan variabilitas terbesar, mengindikasikan puncak dan tingkat keparahan keterlambatan selama bulan musim panas. Januari dan Desember juga menunjukkan keterlambatan median yang relatif tinggi, sesuai dengan gangguan perjalanan musim dingin yang umum terjadi.

Sebaliknya, bulan-bulan seperti Februari, September, Oktober, dan November menunjukkan keterlambatan median yang lebih rendah dengan distribusi yang lebih rapat, mengindikasikan operasi penerbangan yang lebih stabil selama periode ini. Meskipun demikian, Januari masih menunjukkan beberapa pencilan ekstrem, meskipun relatif terisolasi.



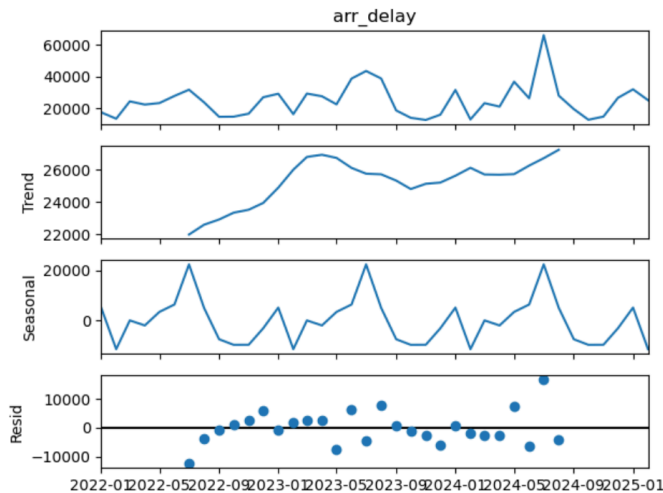
Gambar 11. Line Chart Keterlambatan Kedatangan Berdasarkan Bulan

Selanjutnya pada Gambar 3. dilakukan analisis *trend* dengan menggunakan *line chart* untuk mengamati fluktuasi yang nyata dalam rata-rata keterlambatan. Keterlambatan cenderung meningkat selama bulan-bulan pertengahan tahun, dengan lonjakan signifikan khususnya pada Juli 2024 dan Juli 2023, menunjukkan tekanan operasional atau gangguan selama musim perjalanan puncak musim panas. Periode seperti September 2022, Oktober 2023, dan Januari 2024 menunjukkan rata-rata keterlambatan yang relatif rendah, mengindikasikan kinerja ketepatan waktu yang lebih baik. Lonjakan dramatis pada Juli 2024, di mana rata-rata keterlambatan melebihi 65.000 menit, menonjol sebagai anomali besar, kemungkinan disebabkan oleh peristiwa tidak teratur atau ekstrem.



Gambar 12. Line Chart Keterlambatan dari 2022 hingga 2025

Pada Gambar 4. memperlihatkan evolusi pola keterlambatan dari tahun 2022 hingga 2025. Meskipun terdapat variasi antar-tahun, kita dapat mengidentifikasi tren musiman yang berulang—keterlambatan cenderung meningkat dari musim semi hingga musim panas, mencapai puncaknya terutama sekitar bulan Juli, kemudian menurun drastis di awal musim gugur sebelum sedikit naik kembali menjelang akhir tahun.



Gambar 13. Dekomposisi Deret Waktu Terhadap Data Keterlambatan Kedatangan

Pada Gambar 5. dapat disimpulkan bahwa data ini menunjukkan pola musiman (seasonality) yang sangat kuat disertai tren yang relatif stabil. Karakteristik seperti ini membuat data tersebut ideal untuk dipakai dalam peramalan berbasis deret waktu seperti SES dan SARIMA.

- *Train Test Split*

Dalam penelitian ini, data deret waktu keterlambatan kedatangan (*arr_delay*) dibagi berdasarkan horizon peramalan. Untuk prediksi 1 bulan ke depan, digunakan data Januari 2022–Januari 2025 sebagai training dan Februari 2025 sebagai testing. Sementara itu, prediksi 6 bulan menggunakan data training hingga Agustus 2024 dan testing September 2024–Februari 2025.

Selanjutnya, dibuat model Single Exponential Smoothing (SES) menggunakan data train untuk memprediksi 1 bulan. Sedangkan model SARIMA dibuat menggunakan data train untuk memprediksi 6 bulan. Hasil prediksi dari kedua model tersebut kemudian dibandingkan dengan data test.

- *Model Evaluation*

Pada tahap ini, setelah melakukan proses prediksi dan perbandingan, dilakukan kalkulasi nilai Mean Squared Error (MSE) dan Mean Absolute Percentage Error (MAPE) untuk mengevaluasi akurasi masing-masing model. Pada tabel berikut ditampilkan hasil perhitungan tersebut sebagai representasi kinerja model dalam memprediksi total delay.

Tabel 3. Mean Squared Error (MSE)

Mean Squared Error	1 Bulan	6 Bulan
SES	47044.64	90308099.07
SARIMA	92090677.26	44303781.25

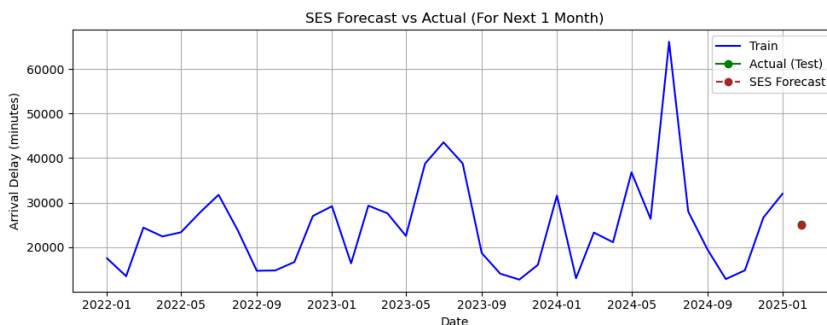
Untuk peramalan 1 bulan, model Simple Exponential Smoothing (SES) menghasilkan MSE rendah sebesar 47.044,64, menunjukkan prediksi yang cukup akurat, sedangkan SARIMA memiliki MSE jauh lebih besar, yaitu 90.308.099,07. Namun, untuk peramalan 6 bulan, SARIMA unggul dengan MSE 44.303.781,27 dibandingkan SES yang mencapai 92.090.677,26, menunjukkan kemampuan SARIMA dalam menangkap pola jangka panjang. Secara umum, MSE kedua model meningkat seiring bertambahnya horizon peramalan, mencerminkan peningkatan kesalahan prediksi pada periode yang lebih panjang.

Tabel 4. Mean Absolute Percentage Error (MSE)

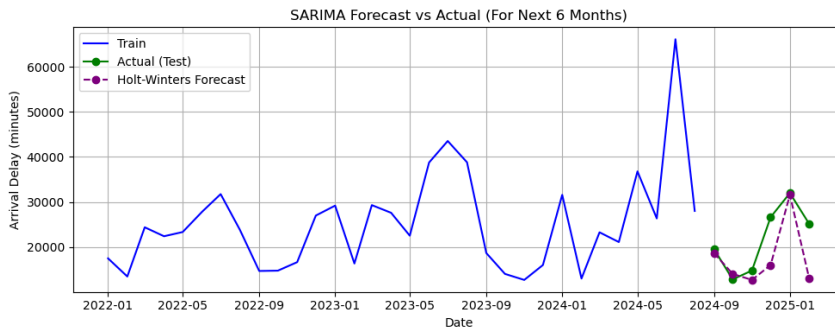
Mean Absolute Percentage Error	1 Bulan	6 Bulan
SES	0.0086	0.4858
SARIMA	0.3812	0.1950

Hasil MAPE menunjukkan pola yang sejalan dengan MSE, di mana model SES untuk prediksi 1 bulan memiliki tingkat kesalahan sangat rendah sebesar 8,6% (0.086), yang termasuk dalam kategori peramalan model sangat baik (<10%). Namun, untuk prediksi 6 bulan, MAPE SES meningkat signifikan menjadi 48,58% (0.4858), yang masuk kategori peramalan model layak mendekati buruk. Sementara itu, model SARIMA memiliki MAPE sebesar 38,12% (0.3812) untuk prediksi 1 bulan, sehingga termasuk peramalan model layak, dan menurun menjadi 19,50% (0.1950) untuk prediksi 6 bulan, yang termasuk peramalan model baik.

Dengan demikian, model SES sangat efektif untuk peramalan jangka pendek dengan akurasi tinggi, namun akurasinya menurun pada jangka panjang. Di sisi lain, SARIMA, meskipun kurang akurat pada horizon 1 bulan, menunjukkan kestabilan yang lebih baik pada prediksi jangka panjang, sehingga dapat menjadi pilihan lebih tepat untuk data dengan pola musiman dan tren yang kompleks. Berikut adalah gambar perbandingan antara hasil forecasting dengan data aktual menggunakan model SES untuk prediksi 1 bulan, serta perbandingan hasil forecasting dengan data aktual menggunakan model SARIMA untuk prediksi 6 bulan.



Gambar 14. Line Chart Hasil Forecasting Menggunakan SES untuk Prediksi 1 Bulan



Gambar 15. Line Chart Hasil Forecasting Menggunakan SARIMA untuk Prediksi 6 Bulan

5. Kesimpulan

Permasalahan keterlambatan penerbangan di Amerika Serikat telah dianalisis melalui pendekatan klasifikasi dan peramalan menggunakan model *Logistic Regression*, SES, dan SARIMA. Berdasarkan hasil analisis klasifikasi keterlambatan penerbangan menggunakan model *Logistic Regression* dapat disimpulkan bahwa model mampu mengklasifikasikan keterlambatan dengan baik setelah melalui tahapan *data preparation*, *exploratory data analysis* (EDA), *train-test split*, dan evaluasi model. Model menunjukkan kinerja yang baik dan efektif dalam mendeteksi penerbangan yang benar-benar terlambat sehingga meminimalkan kesalahan klasifikasi. Hasil ini menunjukkan bahwa *Logistic Regression* dapat menjadi pendekatan awal yang baik dalam membangun sistem prediksi keterlambatan penerbangan, meskipun ke depannya model yang lebih kompleks dapat dipertimbangkan untuk meningkatkan performa prediksi.

Untuk hasil analisis peramalan rata-rata keterlambatan penerbangan, model *Simple Exponential Smoothing* (SES) dan SARIMA menunjukkan performa yang berbeda tergantung pada jangka waktu peramalan. Untuk prediksi jangka pendek selama 1 bulan, model SES memberikan hasil MSE dan MAPE yang lebih baik daripada SARIMA yang hanya termasuk kategori peramalan layak. Namun, untuk prediksi jangka panjang selama 6 bulan, performa SARIMA lebih unggul dibandingkan SES yang mengalami penurunan akurasi yang mendekati kategori buruk. Secara umum, MSE kedua model meningkat seiring bertambahnya jangka peramalan, mencerminkan peningkatan tingkat kesalahan. Hasil ini menunjukkan bahwa SES sangat efektif untuk peramalan jangka pendek, sementara SARIMA lebih stabil dan cocok digunakan untuk peramalan jangka panjang, khususnya pada data yang memiliki pola musiman dan trend yang kompleks.

Temuan dalam studi ini dapat menjadi dasar bagi pengambil keputusan di industri penerbangan untuk mengembangkan sistem manajemen prediksi keterlambatan yang adaptif. Adapun rekomendasi yang dapat diberikan adalah (1) mengintegrasikan model prediksi ke dalam sistem operasional maskapai dan bandara guna meningkatkan efisiensi jadwal, (2) meningkatkan monitoring terhadap faktor keterlambatan utama seperti kesiapan maskapai, serta (3) melakukan evaluasi berkala terhadap performa model prediksi untuk disesuaikan dengan perubahan pola operasional atau musiman.

Keterbatasan dari studi ini adalah penggunaan data yang hanya mencakup periode Januari 2022 hingga Februari 2025. Penelitian selanjutnya disarankan untuk mempertimbangkan model prediksi berbasis *deep learning* atau *hybrid* untuk meningkatkan akurasi serta skalabilitas sistem prediksi keterlambatan penerbangan secara nasional maupun global.

Daftar Pustaka

- Aini, A. N., Ashar, A. U. O., Lestari, T. I., Nur, I. M., & Wasono, R. (2023). Pemodelan Tingkat Kemiskinan di Papua Barat dengan Pendekatan *Binary Logistic Regression*. *Square: Journal of Mathematics and Mathematics Education*, 5(2). <https://journal.walisongo.ac.id/index.php/square/article/view/17169>
- Anupkumar, A. (2023). "INVESTIGATING THE COSTS AND ECONOMIC IMPACT OF FLIGHT DELAYS IN THE AV" by Ashmith Anupkumar. <https://scholarworks.lib.csusb.edu/etd/1653/>
- BTS | OT Delay. (n.d.). Retrieved May 30, 2025, from https://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp?20=E
- Chai, S. (2021). *Hong Kong Air Traffic: Explanation and Prediction based on Sparse Seasonal ARIMA Model* (No. arXiv:2108.05817). arXiv. <https://doi.org/10.48550/arXiv.2108.05817>
- Fauziah, S., & Boni Wijaya, C. (2022). ANALISIS PENJADWALAN PENERBANGAN PESAWAT WINGS AIR DALAM MENGOPTIMASI WAKTU KEBERANGKATAN DAN KEDATANGAN PESAWAT DARI DAN MENUJU BANDAR UDARA SULTAN MUHAMMAD SALAHUDDIN BIMA. *Jurnal Manajemen Dirgantara*, 15(2), 301–308. <https://doi.org/10.56521/manajemen-dirgantara.v15i2.767>
- Freya Parsons. (2024, September 11). *World's busiest airport with 105million passengers a year has 5 runways & direct flights from the UK* | *The Sun*. <https://www.thesun.co.uk/travel/30376500/passenger-traffic-busiest-airport-flight>
- Google. (n.d.). *Thresholding*. Google Developers. <https://developers.google.com/machine-learning/crash-course/classification/thresholding>
- Hidayat, R., & Syarif, M. (2015). Peramalan jumlah penumpang kapal ferry menggunakan metode SARIMA. *Jurnal Matematika, Statistika dan Komputasi*, 11(2), 1–10.
- Katabba, Y. I. (2021). *Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) untuk memprediksi jumlah penumpang kereta api di Pulau Sumatera* (Skripsi, Universitas Jambi). Universitas Jambi Repository. <https://repository.unja.ac.id/23229/1/Skripsi%20Full%20Teks.pdf>
- Maulina, R., & Anggraeni, D. P. (2022). Metode Single Exponential Smoothing (SES) pada peramalan tingkat pengangguran terbuka di Indonesia. *E-Jurnal EVOS*. <https://ejournal.unwmataram.ac.id/evos/article/view/1335>
- Mutha, T. (2023). *Time series forecasting using SARIMA in Python*. Medium. <https://medium.com/@tirthamutha/time-series-forecasting-using-sarima-in-python-8b75cd3366f2>
- Ngabidin, Z., Sanwidi, A., & Arini, E. R. (2023). Implementasi metode double exponential smoothing Brown untuk meramalkan jumlah penduduk miskin. *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, 11(2), 328–338. <https://doi.org/10.37905/euler.v11i2.23054>
- Putra, A. W., & Suryawan, I. N. (n.d.). *Penerapan metode SARIMA dalam peramalan jumlah kunjungan wisatawan ke Bali*. *Jurnal Sistem Informasi STIKOM Bali*. <https://jsi.stikom-bali.ac.id/index.php/jsi/article/download/193/166>

- Rai, K. (2020, May 10). *The math behind logistic regression*. Medium.
<https://medium.com/analytics-vidhya/the-math-behind-logistic-regression-c2f04ca27bca>
- Sembiring, B. L. (2020). Analisa Kesepakatan ASEAN Open Skies Terhadap Lalu Lintas Penumpang dan Pesawat di Asia Tenggara. *WARTA ARDHIA*, 46(1), 18–25. <https://doi.org/10.25104/wa.v46i1.390.18-25>
- Tang, Y. (2021). Airline Flight Delay Prediction Using Machine Learning Models. *2021 5th International Conference on E-Business and Internet*, 151–154. <https://doi.org/10.1145/3497701.3497725>