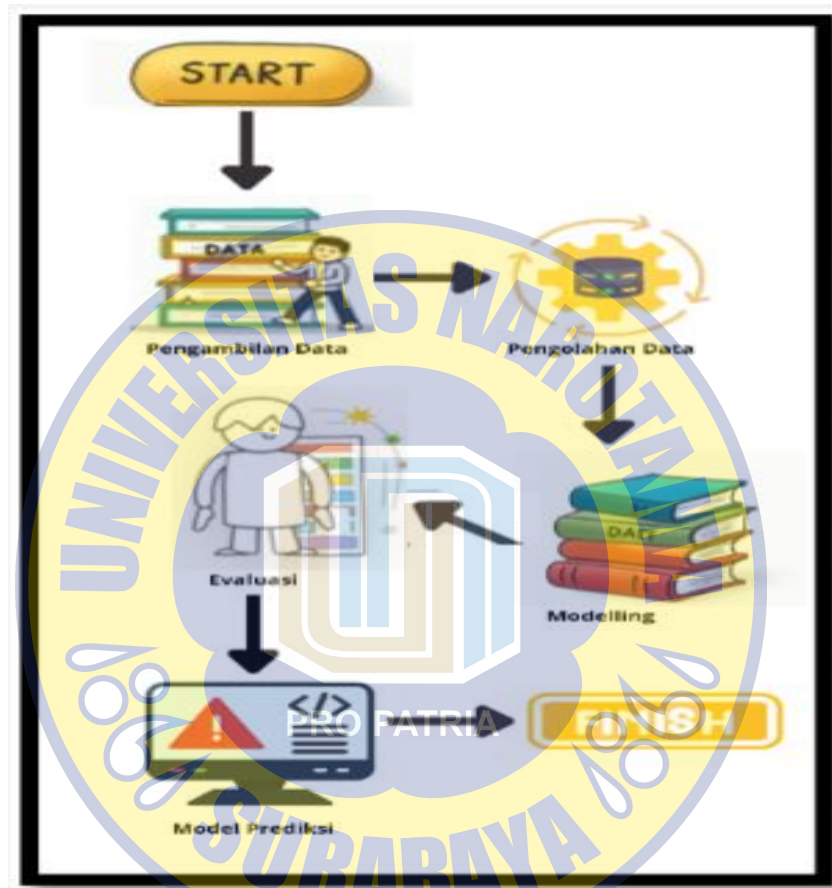


BAB III

METODOLOGI PENELITIAN



Gambar 3.1 Metodologi Penelitian

Metode Penelitian ini menggunakan **Gambar 3.1** sebagai ilustrasi alur proses yang dilakukan dalam penelitian. Langkah-langkahnya adalah sebagai berikut

1. Start

Langkah pertama adalah memulai seluruh rangkaian proses dalam penelitian ini. Pada fase ini, tim peneliti memastikan bahwa semua persiapan, termasuk pengumpulan data dan pemilihan metode, telah disiapkan dengan baik untuk memulai penelitian.

2. Pengambilan Data

Pada tahap ini, data yang diperlukan untuk penelitian dikumpulkan. Data yang digunakan mencakup berbagai informasi cuaca yang relevan, seperti curah hujan,. Pengambilan data ini sangat penting untuk memastikan kualitas data yang digunakan dalam analisis, yang nantinya akan mempengaruhi akurasi prediksi yang dihasilkan. Data yang digunakan sebagian besar bersumber dari Dinas Pekerjaan Umum Kota Surabaya.

3. Pengolahan Data

Setelah data terkumpul, langkah berikutnya adalah pengolahan data. Pada tahap ini, dilakukan pembersihan data untuk menghapus atau memperbaiki data yang tidak lengkap atau mengandung kesalahan. Selain itu, dilakukan juga normalisasi data agar lebih konsisten dan dapat diproses dengan baik oleh model analisis. Proses ini mencakup penanganan data yang hilang, penghapusan *outlier*, dan penyesuaian data agar siap untuk digunakan dalam analisis lebih lanjut.

4. Modelling

Pada tahap ini, model matematis atau algoritma analitik dibangun untuk menganalisis data yang telah diproses dan menghasilkan prediksi. Dalam penelitian ini, metode yang digunakan adalah Model ARIMA dan SARIMA, yang dipilih karena kemampuannya untuk menganalisis hubungan antar variabel dengan akurasi tinggi. Model ini memungkinkan peneliti untuk mengidentifikasi pola dan hubungan dalam data yang dapat digunakan untuk membuat prediksi yang lebih akurat mengenai curah hujan di Kota Surabaya.

5. Evaluasi

Setelah model dibangun, langkah selanjutnya adalah evaluasi untuk menilai kinerja model. Evaluasi dilakukan dengan membandingkan hasil prediksi model dengan data aktual yang tersedia. Metrik evaluasi yang digunakan meliputi (MAE), (RMSE). Metrik ini digunakan untuk mengukur seberapa akurat model dalam memprediksi curah hujan yang sebenarnya.

6. Model Prediksi

Setelah model dievaluasi dan dinilai baik, tahap berikutnya adalah menggunakan model untuk menghasilkan prediksi berdasarkan data baru. Pada fase ini, model yang telah terbukti akurat digunakan untuk meramalkan curah hujan pada

hari-hari mendatang, yang akan digunakan untuk perencanaan kegiatan MICE serta mitigasi risiko terkait cuaca ekstrem.

7. Finish

Setelah seluruh proses selesai, tahap penelitian mencapai titik akhir (*finish*). Pada titik ini, seluruh rangkaian alur kerja telah diselesaikan, dan hasil penelitian yang diperoleh dapat dilanjutkan ke tahap implementasi atau analisis lebih lanjut. Hasil dari penelitian ini akan digunakan untuk merancang sistem peringatan dini berbasis cuaca yang dapat membantu pihak terkait dalam merencanakan kegiatan MICE dengan lebih baik dan mengurangi dampak buruk yang disebabkan oleh cuaca buruk di Kota Surabaya.

3.1 Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif, yang bertujuan untuk memahami pola curah hujan dan meramalkan potensi dampaknya terhadap pelaksanaan kegiatan (MICE) di Kota Surabaya. Penelitian ini dirancang dengan metode deskriptif untuk memetakan pola curah hujan berdasarkan data historis, serta metode analitik untuk menerapkan teknik pemodelan deret waktu seperti ARIMA dan SARIMA dalam menghasilkan prediksi cuaca yang akurat. Penelitian ini juga memungkinkan evaluasi keakuratan prediksi menggunakan metrik seperti (MAE) Dan (RMSE) untuk memberikan hasil yang valid dan reliabel (Sugiyono 2019)

Penelitian ini juga bertujuan untuk mengintegrasikan hasil prediksi curah hujan ke dalam pengelolaan acara MICE. Prediksi ini penting untuk memitigasi risiko akibat cuaca ekstrem yang dapat memengaruhi kelancaran kegiatan MICE di lokasi strategis seperti *Grand City Surabaya*, *Stadion Gelora Bung Tomo*, *Tunjungan Plaza Convention Hall* dan *Gramedia Expl* (MICE Trends in Indonesia 2022)

Dalam rangka penelitian mengenai prediksi curah hujan sebagai penunjang kegiatan MICE di Kota Surabaya, Penelitian ini memutuskan untuk mengumpulkan data curah hujan harian yang akurat dan relevan. Data yang dibutuhkan mencakup catatan curah hujan harian selama 10 tahun, dari tahun 2013 hingga 2023. Pengumpulan data ini sangat penting, mengingat Surabaya adalah salah satu kota

besar di Indonesia yang sering mengadakan berbagai kegiatan MICE, dan cuaca dapat mempengaruhi kelancaran acara tersebut.

Salah satu sumber yang dipilih adalah Dinas (PU) Pengelolaan Sumber Daya Air (PSDA) Ngagel Surabaya, yang dikenal memiliki arsip data meteorologi yang komprehensif. saya percaya bahwa data dari PU PSDA Ngagel akan memberikan informasi yang valid dan dapat diandalkan untuk dianalisis.

Proses pengambilan data dimulai dengan mengajukan permohonan resmi kepada pihak PU PSDA Ngagel. Sebagai langkah awal, menghubungi pihak kampus Narotama untuk meminta surat pengantar yang diperlukan. Surat ini penting sebagai legitimasi atas permohonan dan menunjukkan bahwa penelitian ini merupakan bagian dari kegiatan akademik.

Setelah mendapatkan surat pengantar, segera mengirimkannya ke PU PSDA Ngagel Surabaya. Dalam surat tersebut menjelaskan tujuan penelitian ini dan jenis data yang saya butuhkan, yaitu data curah hujan harian dari tahun 2013 hingga 2023. juga menyertakan informasi kontak untuk memudahkan komunikasi lebih lanjut. Proses ini memerlukan ketelitian, karena ingin memastikan bahwa semua informasi yang diberikan jelas dan mudah dipahami oleh pihak PU PSDA.

Setelah menunggu selama dua hari, diterima konfirmasi dari pihak PU PSDA Ngagel bahwa permohonan telah disetujui. Pihak Pu meminta untuk datang langsung ke kantor mereka untuk mengambil data yang telah disiapkan. Pada hari yang ditentukan, oleh PU PSDA Ngagel dengan membawa flash disk, sebagai media penyimpanan untuk data yang akan diambil.

Sesampainya di PU PSDA Ngagel staf yang bertanggung jawab atas pengelolaan data curah hujan menjelaskan proses pengambilan data dan menunjukkan data yang telah disiapkan. Setelah melakukan verifikasi dan memastikan bahwa data tersebut sesuai dengan yang dibutuhkan, dan mentransfer data curah hujan harian dari tahun 2013 hingga 2023 ke dalam USB.

Selain itu, perlu mengambil data curah hujan dari (BMKG) sebagai sumber tambahan. Data yang diambil dari BMKG sangat penting karena mencakup informasi yang lebih luas dan dapat digunakan untuk membandingkan data yang diperoleh dari PU PSDA Ngagel. Diharapkan dengan menggabungkan data dari kedua sumber ini, analisis akan lebih komprehensif dan akurat.

Proses pengambilan data dari BMKG dilakukan melalui website resmi BMKG. Namun, menemukan bahwa data yang dapat diunduh dari situs tersebut terbatas, data hanya dapat diambil data curah hujan satu tahun dalam satu hari. Dengan demikian, untuk mengumpulkan data selama 10 tahun, memerlukan waktu

selama 10 hari berturut-turut. Meskipun sulit dan memakan waktu, komitmen tetap untuk mendapatkan data yang akurat dan lengkap.

Setiap hari, mengakses website BMKG, mencari data curah hujan harian untuk tahun yang berbeda, dan mengunduhnya satu per satu. Proses ini tidak selalu berjalan mulus; terkadang mengalami kesulitan dalam navigasi situs atau menghadapi masalah teknis. Ketekunan dan kerja keras dalam mencatat setiap langkah yang diambil. Setelah 10 hari, terkumpul data curah hujan harian dari tahun 2013 hingga 2022.

3.2.1 Sumber Data

Data yang digunakan dalam penelitian ini meliputi:

1. Data Utama
 - a) Data curah hujan bulanan Kota Surabaya untuk periode 2013–2022. Dinas Pekerjaan Umum (DPU) Kota Surabaya. (Dinas Pekerjaan Umum Kota Surabaya. (2022)
 - b) Data mencakup curah hujan harian dalam milimeter (mm) yang dikumpulkan dari beberapa stasiun pengamatan cuaca di wilayah strategis Surabaya.
2. Data Pendukung
 - a) Data lokasi MICE Lokasi strategis seperti hotel berbintang, pusat konvensi, dan tempat-tempat wisata yang sering digunakan untuk kegiatan MICE.

3.3 Pengolahan Data

3.3.1 Tahapan Pengolahan

1. Pengumpulan Data:

Data curah hujan diambil dari laporan bulanan Dinas PU Kota Surabaya dalam format *spreadsheet*. Data di impor ke perangkat lunak analitik seperti Python.

2. Pembersihan Data:

Data diperiksa untuk outlier, nilai hilang, atau anomali menggunakan metode statistik seperti:

- a) Teknik interpolasi untuk memperbaiki data hilang.

Interpolasi adalah proses matematis untuk memperkirakan nilai di antara dua titik data yang diketahui. Dengan menggunakan nilai-nilai yang ada, interpolasi dapat memberikan estimasi yang masuk akal untuk nilai yang hilang.

Data yang hilang dapat terjadi karena berbagai alasan, seperti kesalahan pengukuran, masalah teknis, atau ketidakhadiran data pada waktu tertentu. Mengisi data yang hilang penting untuk menjaga integritas analisis dan untuk memastikan bahwa model yang dibangun tidak bias atau tidak akurat.

Teknik Interpolasi yang Umum Digunakan

1. Interpolasi Linier

Metode paling sederhana yang mengasumsikan bahwa perubahan antara dua titik data adalah linier. Nilai yang hilang diestimasi dengan menarik garis lurus antara dua titik data yang diketahui.

$$y = y_1 + \frac{(x-x_1)(y_2-y_1)}{x_2-x_1}$$

Formula: Di mana $((x_1, y_1))$ dan $((x_2, y_2))$ adalah titik data yang diketahui, dan (x) adalah posisi di mana nilai hilang. (Ismail et al. 2023)

2. Interpolasi Polinomial

Menggunakan polinomial untuk memperkirakan nilai yang hilang. Ini lebih kompleks dan dapat memberikan hasil yang lebih akurat jika data memiliki pola yang tidak linier.

3. Interpolasi Spline

Menggunakan fungsi spline (serangkaian polinomial) untuk memperkirakan nilai. Spline cubic adalah salah satu metode yang umum digunakan, di mana fungsi spline dibangun untuk setiap interval antara titik data.

4. Interpolasi Berdasarkan Rata-rata

Mengisi nilai yang hilang dengan rata-rata dari nilai-nilai di sekitarnya. Ini adalah metode yang sederhana tetapi dapat mengurangi variabilitas data.

5. Interpolasi dengan Metode Lain

Metode lain seperti kriging, yang sering digunakan dalam geostatistik, atau metode berbasis machine learning, seperti regresi, juga dapat digunakan untuk mengisi data yang hilang.

- a) Transformasi data, seperti logaritma atau *square root*, diterapkan jika distribusi data tidak normal.

1. Analisis Eksplorasi Data (EDA)

- a) Visualisasi grafik menggunakan matplotlib di dalam python untuk memetakan pola musiman dan tren jangka panjang dalam data curah hujan.

Visualisasi grafik menggunakan Matplotlib untuk memetakan pola musiman dan tren jangka panjang dalam data curah hujan adalah proses yang melibatkan pembuatan grafik yang dapat membantu dalam memahami bagaimana curah hujan bervariasi sepanjang waktu. Ini sangat penting dalam analisis data deret waktu, terutama untuk data cuaca dan iklim.

1. Pola Musiman

Pola musiman merujuk pada fluktuasi yang terjadi secara teratur dalam data selama periode waktu tertentu, seperti bulanan atau tahunan. Dalam konteks curah hujan, ini bisa berarti peningkatan curah hujan selama musim hujan dan penurunan selama musim kemarau.

2. Tren Jangka Panjang

Tren jangka panjang adalah perubahan yang terjadi dalam data selama periode waktu yang lebih lama. Misalnya, analisis dapat menunjukkan apakah curah hujan secara keseluruhan meningkat atau menurun selama beberapa tahun terakhir.

- b) Analisis korelasi antara curah hujan dan bulan tertentu dilakukan untuk mengidentifikasi pola musiman.

3.4 Pemodelan dengan ARIMA dan SARIMA

Langkah-Langkah Pemodelan:

1. Pengujian Stasioneritas Data:

Augmented Dickey-Fuller (ADF) Test adalah uji statistik yang digunakan untuk memeriksa stasioneritas data deret waktu. Stasioneritas berarti bahwa statistik deskriptif dari data, seperti rata-rata dan variansi, tidak berubah seiring waktu. Dalam analisis deret waktu, sangat penting untuk memiliki data yang stasioner karena model

prediktif seperti ARIMA dan SARIMA memerlukan data yang tidak memiliki tren jangka panjang atau variansi yang meningkat.

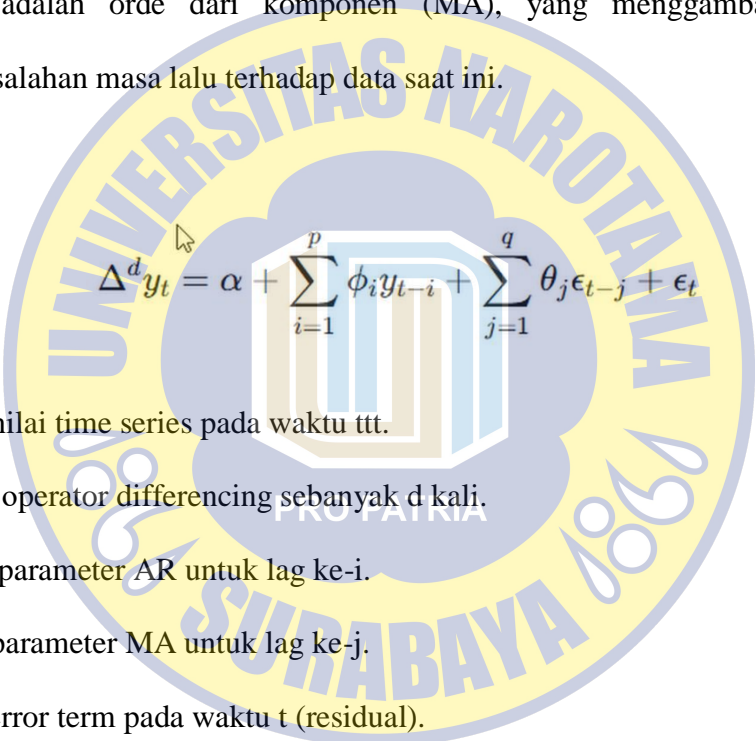
Uji ADF menguji hipotesis nol bahwa data memiliki akar unit (*non-stasioner*). Jika nilai p lebih besar dari level signifikansi (misalnya, 0,05), maka kita gagal menolak hipotesis nol dan data dianggap tidak stasioner. Dalam hal ini, diferensiasi (proses pengurangan data dari nilai sebelumnya) dilakukan untuk membuat data stasioner.

2. Identifikasi Parameter Model:

- b) ARIMA (p, d, q) dan SARIMA (P, D, Q, S) memiliki parameter yang perlu diidentifikasi.

3. ARIMA

- a) p adalah orde dari komponen (AR), yang menunjukkan jumlah periode lag yang digunakan dalam model.
- b) d adalah jumlah diferensiasi yang diperlukan untuk membuat data menjadi stasioner.
- c) q adalah orde dari komponen (MA), yang menggambarkan pengaruh kesalahan masa lalu terhadap data saat ini.


$$\Delta^d y_t = \alpha + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t$$

y_t adalah nilai time series pada waktu t .

Δ^d adalah operator differencing sebanyak d kali.

ϕ_i adalah parameter AR untuk lag ke- i .

θ_j adalah parameter MA untuk lag ke- j .

ϵ_t adalah error term pada waktu t (residual).

α adalah intercept atau konstanta.

4. SARIMA

Menambahkan komponen musiman untuk menangani data yang memiliki pola musiman yang jelas. Parameter musiman (P, D, Q, S) merujuk pada

komponen AR, I, MA, dan panjang periode musiman (S). (misalnya, S = 12 untuk data bulanan dengan musiman tahunan).

$$\Delta^d \Delta_s^D y_t = \alpha + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \sum_{k=1}^P \Phi_k y_{t-sk} + \sum_{l=1}^Q \Theta_l \epsilon_{t-sl} + \epsilon_t$$

Δ_s^D adalah operator differencing musiman sebanyak D kali pada musim dengan panjang S.

ϕ_i adalah parameter AR untuk lag ke-i.

θ_j adalah parameter MA untuk lag ke-j.

Φ_k adalah parameter AR musiman untuk lag ke-k.

Θ_l adalah parameter MA musiman untuk lag ke-l.

ϵ_t adalah error term pada waktu t.

α adalah intercept atau konstanta.”

5. Pembangunan Model

- a) Model ARIMA dan SARIMA dibangun menggunakan perangkat lunak seperti Python dengan pustaka statsmodels. Statsmodels menyediakan fungsi-fungsi untuk implementasi ARIMA dan SARIMA dengan parameter yang telah diidentifikasi sebelumnya.

6. Evaluasi Model

Setelah model dibangun, evaluasi dilakukan dengan mengukur performa prediksi model menggunakan beberapa metrik

- a) (MAE): Mengukur rata-rata dari selisih absolut antara nilai yang diprediksi dan nilai aktual.
- b) (RMSE): Mengukur akar kuadrat dari MSE, memberikan bobot lebih besar pada kesalahan yang lebih besar.

Model dengan nilai MAE atau RMSE terendah akan dianggap yang paling akurat, dan dengan demikian, lebih dapat diandalkan untuk prediksi curah hujan yang lebih tepat.

Perbandingan Model:

Hasil dari model ARIMA (non-musiman) akan dibandingkan dengan model SARIMA (musiman) untuk menentukan metode yang memberikan hasil terbaik dalam memprediksi curah hujan bulanan di Surabaya. (Box et al. 2015) (Shumway and Stoffer 2011)

3.5 Prediksi Curah Hujan

Model prediksi akan digunakan untuk

1. Memprediksi Curah Hujan Bulanan

Prediksi dilakukan untuk 12 bulan mendatang menggunakan model ARIMA dan SARIMA.

2. Analisis Musiman

Hasil prediksi SARIMA akan memberikan wawasan lebih mendalam mengenai pola musiman curah hujan di Surabaya.

3. Dampak pada Sektor MICE

Hasil prediksi akan diterjemahkan ke dalam rekomendasi operasional untuk sektor MICE, seperti:

- a) Waktu terbaik untuk mengadakan acara luar ruangan.
- b) Perencanaan transportasi dan akomodasi. (Hyndman and Athanasopoulos 2021)

3.6 Teknik Analisis Data

Langkah-Langkah Analisis:

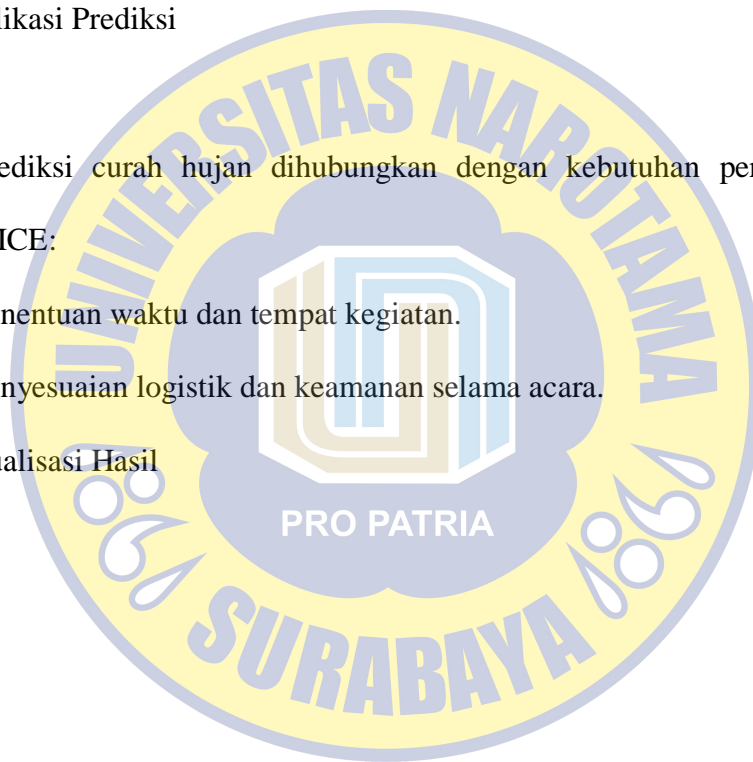
1. Validasi Model:

- a) Membandingkan hasil prediksi dengan data aktual menggunakan data uji (test set).
- b) Analisis perbandingan model ARIMA vs SARIMA dilakukan untuk memilih model terbaik.

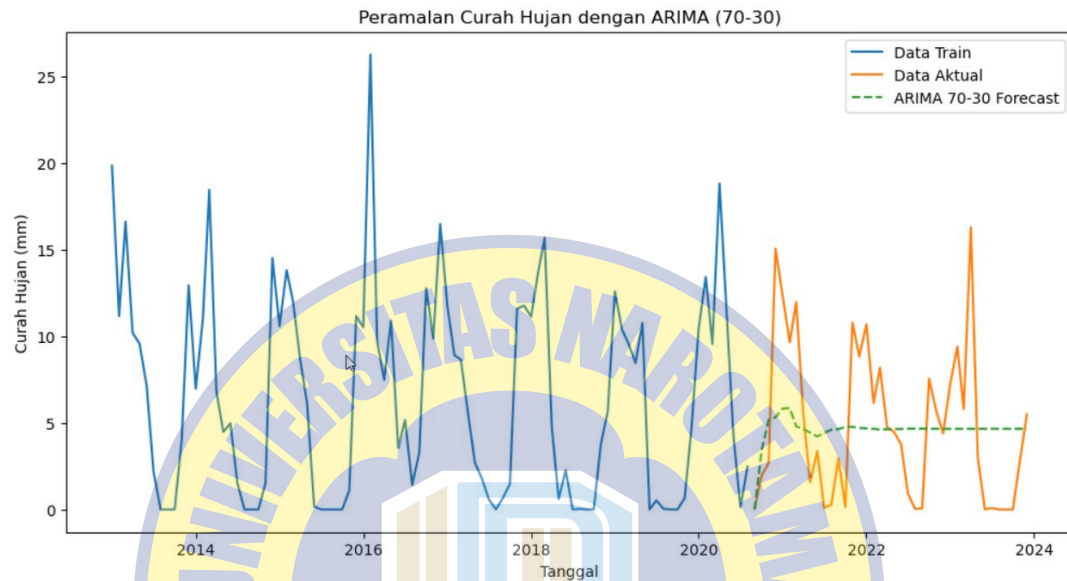
2. Implikasi Prediksi

- a) Prediksi curah hujan dihubungkan dengan kebutuhan perencanaan sektor MICE:
- b) Penentuan waktu dan tempat kegiatan.
- c) Penyesuaian logistik dan keamanan selama acara.

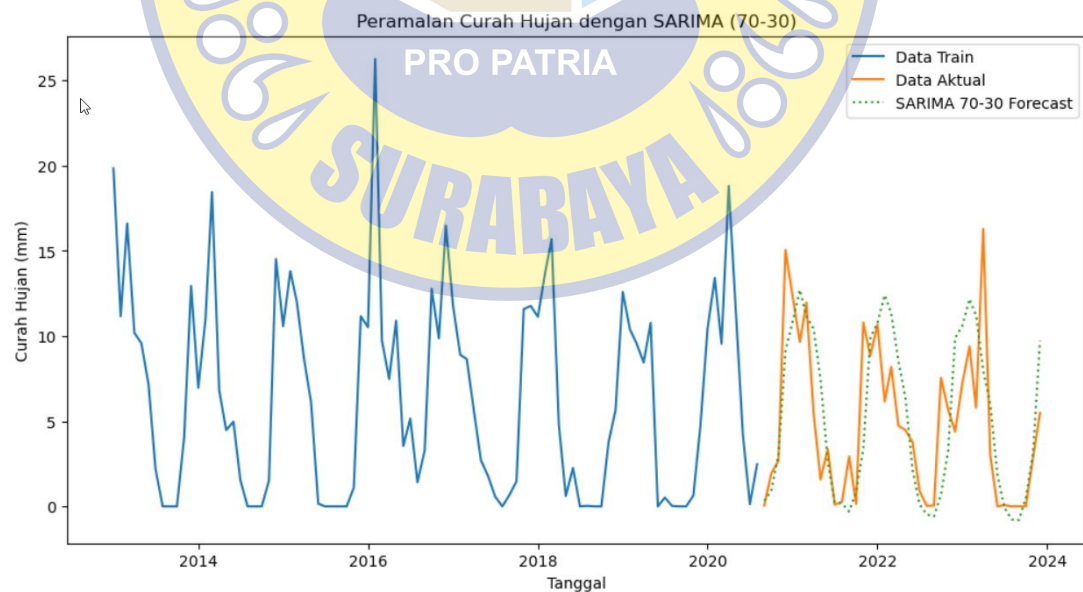
3. Visualisasi Hasil



- a) Hasil prediksi divisualisasikan melalui grafik deret waktu, yang menunjukkan data aktual, prediksi, dan rentang kepercayaan.



Gambar 3.2 Contoh Hasil Grafik Arima



Gambar 3.3 Contoh Hasil Grafik Sarim