

BAB III

METODOLOGI PENELITIAN

A. Jenis Penelitian

Dalam hal ini penelitian yang akan digunakan adalah penelitian kuantitatif dengan analisis runtun waktu, yang didasarkan atas data yang relevan pada masa lalu, dengan menggunakan teknik dan metode-metode dalam penganalisaan data.

B. Tempat dan Jadwal Penelitian

Penelitian ini dilaksanakan di Badan Meteorologi, Klimatologi, Dan Geofisika yang beralamat di Jl. Gatot Subroto No. 20, Tambaksari, Sidanegara, Cilacap Tengah, Kabupaten Cilacap, Jawa Tengah 53272 atau berada di Balai Besar Wilayah II. Waktu penelitian mulai dilaksanakan pada Januari 2022. Berikut ini adalah tabel jadwal penelitian :

Tabel 1. Jadwal Penelitian

Kegiatan	2021/2022																			
	Des		Jan		Feb		Mar		Apr		Mei		Jun		Jul		Ags		Sep	
	Minggu ke																			
	1-2	3-4	1-2	3-4	1-2	3-4	1-2	3-4	1-2	3-4	1-2	3-4	1-2	3-4	1-2	3-4	1-2	3-4	1-2	3-4
Observasi																				
Pengambilan Data																				
Pengolahan Data																				
Penyusunan Laporan																				

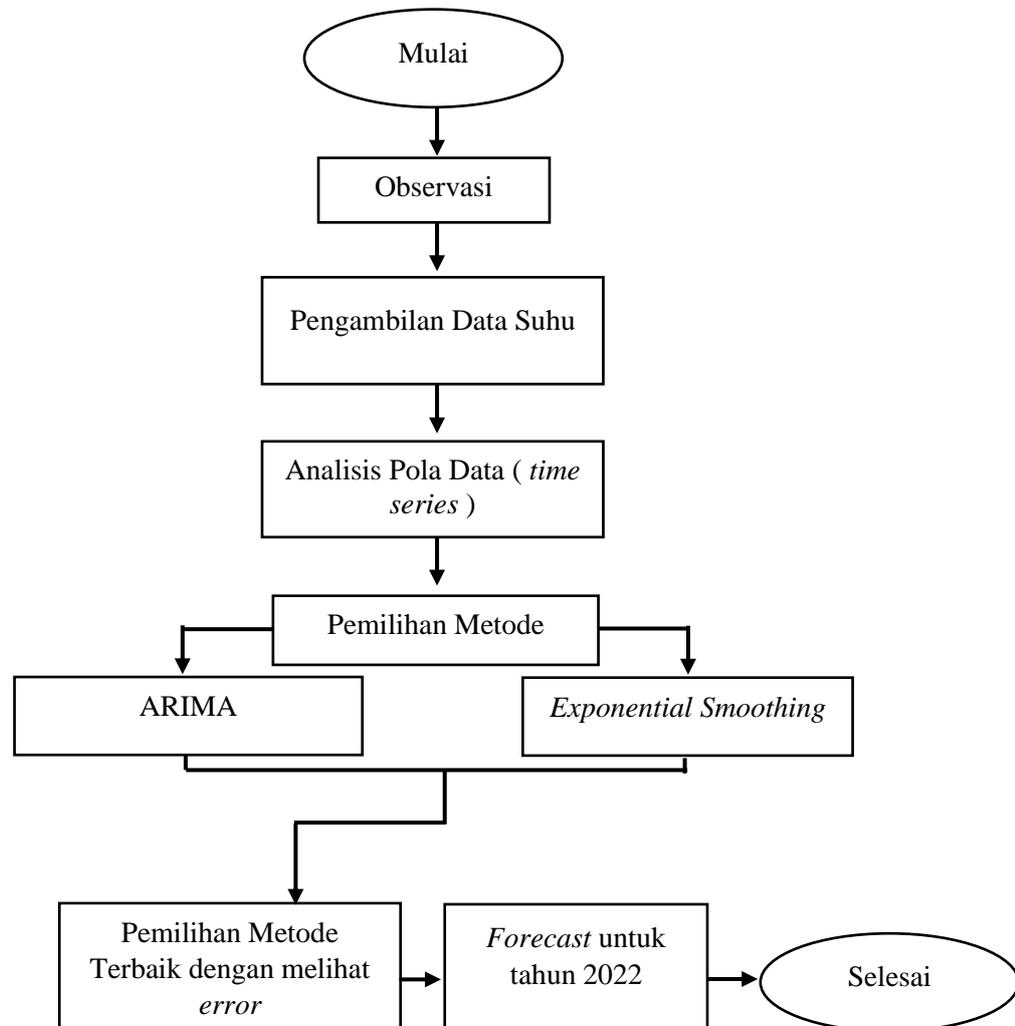
C. Sumber dan Jenis Data

Data yang digunakan merupakan data sekunder. Data sekunder adalah data yang mengacu pada informasi yang dikumpulkan dari sumber yang telah ada, sumber datanya yang diperoleh berupa data suhu yang diambil dari data *synop* yang terdapat dibagian penginputan data. Variabel yang digunakan untuk memprediksi yaitu data suhu maksimum dan

minimum harian yang merupakan data harian dari Januari 2016 sampai dengan Desember 2021.

D. Tahapan Penelitian

Berikut ini adalah alur tahapan (*flowchart*) penelitian :



Gambar 1. Tahapan Penelitian

langkah-langkah dalam meramalkan suhu udara maksimum dan minimum harian Kabupaten Cilacap tahun 2022 menggunakan metode ARIMA dan *Exponential Smoothing* adalah sebagai berikut :

1. Observasi

Obsevasi dilakukan penulis untuk mencari informasi dan mendapatkan data untuk penelitian.

2. Pengambilan Data

Sumber data yang digunakan penulis yaitu sumber data sekunder, yang diperoleh dari pengamatan suhu *max* dan *min* harian selama kurang lebih 5 tahun, yang terdapat di dalam file *synop*.

3. Menganalisis Pola Data *Time Series*

Data yang sudah diperoleh kemudian dilakukan proses mengubah data ke dalam bentuk *time series*, setelah itu dibentuk plot untuk menentukan data tersebut mengandung unsur *trend*, musiman, atau tidak keduanya, yang selanjutnya akan digunakan untuk menentukan model dari setiap metode dalam *forecasting*.

4. *Forecasting I*

Melakukan peramalan suhu maksimum dan minimum dengan model yang sudah dipilih dari setiap metode, yaitu dari Metode ARIMA dan Metode Exponential Smoothing dengan bantuan *Software Rstudio*.

5. Pemilihan Metode Terbaik

Setelah dilakukan proses peramalan dari setiap metode, diperoleh hasil dan nilai *error*. Kemudian dilihat dan dibandingkan dengan data *actual* yang sudah ada, dan dipilih metode terbaik dari dua metode tersebut.

6. *Forecasting II*

Melakukan peramalan suhu maksimum dan minimum dengan metode terbaik.

7. Kesimpulan

Menarik kesimpulan hasil peramalan suhu maksimum dan minimum harian Kabupaten Cilacap tahun 2022.

E. Alat Analisis Data

Alat yang digunakan untuk melakukan peramalan suhu maksimum dan minimum harian menggunakan ARIMA dan *Exponential Smoothing* sebagai berikut :

1. Laptop

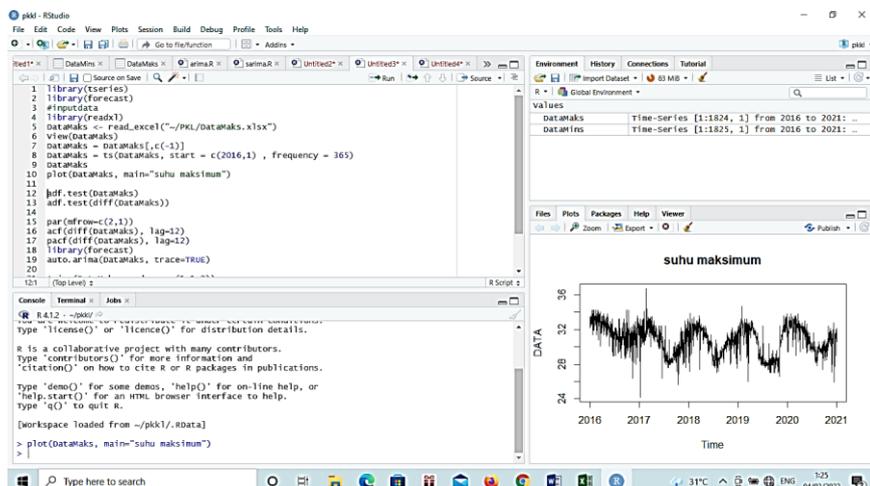
Laptop adalah komputer yang dapat dipindahkan dan dibawa dengan mudah, mayoritas laptop mempunyai fitur yang sama dengan komputer, yaitu mampu menjalankan perangkat lunak

dan mengelola berkas, dengan ukuran kecil dan dapat dibawah kemana-mana dengan sumber daya berasal dari baterai. Perangkat ini terdiri dari *processor*, LCD, *keyboard*, *touchpad*, RAM, *hardisk*, dll yang dikemas menjadi satu perangkat. Berbagai macam manfaat laptop untuk memudahkan penggunaanya, diantaranya, untuk pekerja kantor, dosen, mahasiswa, dll.

2. Software R

Software R adalah *software* untuk menganalisis data statistik, hampir sama seperti MINITAB, SPSS, dll. Sedangkan *Rstudio* adalah *Integrated Development Environment (IDE)* untuk *R* yang banyak digunakan hingga saat ini (Budiaji, 2019). Dapat dikatakan bahwa hampir semua pengguna *R* yang sudah mengetahui *RStudio* akan lebih memilih menggunakan *R* melalui *RStudio* dibandingkan dengan menggunakan *R*. Saat ini, *R* sudah dikenal luas sebagai salah satu *powerful software* untuk analisis data dan *Data Science*. Tentu saja selain *R* masih banyak software lain yang juga sering digunakan untuk analisis data, misalnya *Python*.

R dibuat dengan tujuan awal untuk komputasi statistika dan grafis. Dengan adanya aplikasi *Rstudio* ini, diharapkan mampu memudahkan peneliti untuk melakukan peramalan suhu udara maksimum dan minimum harian menggunakan ARIMA dan *Exponential Smoothing*. Berikut tampilan *Rstudio* :



Gambar 2. Tampilan Rstudio

Berikut contoh syntax data dalam *Rstudio* :

a. Membuat data menjadi data time series

```
library(readxl)
DataMaks <- read_excel("~/PKL/DataMaks.xlsx")
View(DataMaks)
DataMaks.ts = ts(DataMaks$DATA, start =
c(2016,1),end = c(2021,365), frequency = 365)
paste(DataMaks.ts)
```

b. Membuat Plot data :

```
plot.ts(DataMaks.ts)
plot(DataMaks)
```

c. Contoh Syntax Peramalan :

```
hwb.ka.add = HoltWinters(DataMaks.ts, alpha =
NULL, beta = NULL, gamma = NULL, seasonal =
"additive")
hwb.ka.add
plot(hwb.ka.add)
```

d. Contoh Syntax Menghitung Pengukuran Kesalahan :

```
mse.multi =
hwb.ka.multi$SSE/frequency(hwb.ka.multi$fitted)
rmse.multi = sqrt(mse.multi)
mape.multi = mean(abs(DataMaks.ts-
hwb.ka.multi$fitted[,1])/DataMaks.ts)*100
mse.multi
rmse.multi
mape.multi
```

e. Menghitung Prediksi :

```
pred.ka = predict(hwb.ka.add,30)
pred.ka
plot(pred.ka)
pred.ka = predict(hwb.ka.multi,30)
pred.ka
```

f. Menggambarkan Plot :

```
plot(DataMaks.ts, main = "suhu maksimum
cilacap", lwd= 2, col = "blue", xlim =
c(2016,2024), type="o",pch=15)
```

```

limitDate =
end(DataMaks.ts) [1]+(end(DataMaks.ts) [2]-
1)/frequency(DataMaks.ts)
abline(v=limitDate, lty=4)
lines(hwb.ka.add$fitted[,1], lwd=2, col="red",
type="o", pch=12)
lines(pred.ka, col = "green", type="o", pch=10)
legend("topright", legend = c("Data Aktual",
"Fitted Value", "Peramalan"), col =
c("blue","red","green"),lty = 1, pch =
c(15,12,10), inset = 0,05)

```

F. Analisis Data

Langkah – langkah dalam menganalisis data dalam penelitian ini adalah

1. Pengumpulan Data

Adapun data yang akan digunakan adalah data suhu udara maksimum dan minimum yang dimulai dari Bulan Januari 2016 sampai dengan Bulan Desember 2021. Data tersebut diperoleh dari data *synop*, yaitu data yang terdapat dibagian pengamatan yang di *input* manual oleh karyawan BMKG.

2. Pemodelan ARIMA

a. Kestasioneran Data

Menganalisis data yang diperoleh diubah kedalam bentuk *time series* dengan bantuan *software Rstudio*, yaitu dengan memanggil data yang sudah ada dalam excel ke *R* dan kemudian membuat plot ACF untuk mengetahui kestasioneran data. Untuk mengetahui kestasioneran data, yaitu dengan uji *Augmented Dickey-Fuller (Adf)*, jika data tidak stasioner dalam rata-rata maka dilakukan proses pembedaan (*differencing*), sampai data menjadi stasioner.

b. Identifikasi Model

Stasioneritas dari data dapat dilihat dari bentuk fungsi estimator fungsi autokorelasi sampel ACF (*Autocorrelation function*) dan estimator fungsi autokorelasi parsial (sampel PACF/*Partial ACF*), ataupun dengan melakukan uji unit *root*

terhadap data (Rosadi, 2013). Selanjutnya, jika telah dilakukan *preprocessing* terhadap data sehingga menghasilkan data yang stasioner, dapat di tentukan bentuk model ARMA (*Autoregressive Moving Average*) yang tepat dalam menggambarkan sifat-sifat data, dengan cara membandingkan plot sampel ACF/PACF dengan sifat-sifat fungsi ACF/PACF dari model ARMA. Untuk bisa mengamati dari kedua fungsi tersebut apabila fungsi parsial autokorelasi turun lambat, jika fungsi autokorelasi terjadi terputus pada lag-1 modelnya MA(1). Begitu dengan fungsi autokorelasi (ACF) turun lambat, jika fungsi parsial autokorelasi terputus di lag-1 modelnya AR(1).

Tabel 2. Proses ACF & PACF

Proses	Sampel ACF	Sampel PACF
<i>White Noise</i>	Tidak ada yang melewati batas interval pada lag > 0	Tidak ada yang melewati batas interval pada lag > 0
AR(p)	<i>Dies down</i> (turun cepat secara eksponensial/ sinusolida)	<i>Cut off after lag p</i> (terputus setelah lag p)
MA(q)	<i>Cut off after lag q</i> (terputus setelah lag q)	<i>Dies down</i> (turun cepat secara eksponensial/ sinusolida)
ARMA(p,q)	<i>Dies down after lag (q-p)</i> (turun cepat setelah lag (q-p))	<i>Dies down after lag (p-q)</i> (turun cepat setelah lag (p-q))

c. Estimasi

Nilai estimasi digunakan untuk menentukan nilai akhir peramalan. Untuk pengujian apakah koefisien hasil estimasi signifikan atau tidak (yakni uji hipotesis null koefisien bernilai 0 vs hipotesis alternatif koefisien tidak nol) dapat digunakan pengujian dengan statistik uji t yang akan berdistribusi *student-t* dengan derajat bebas $n-1$, n = banyaknya sampel.

d. *Diagnostic Check*

Langkah selanjutnya adalah melakukan *diagnostic check* dari model yang telah diestimasi, yakni melakukan verifikasi kesesuaian model dengan sifat-sifat data. Jika model merupakan model yang tepat, maka data yang dihitung dengan model (*fitted value*) akan memiliki sifat-sifat yang mirip dengan data asli.

Untuk melihat apakah residual bersifat *White Noise*, dapat dilakukan dengan dua cara, yakni pertama dengan melihat apakah plot sampel ACF/PACF residual yang terstandarisasi (residual dibagi estimasi standar deviasi residual) telah memenuhi sifat-sifat proses *White Noise* dengan mean 0 dan variansi 1. Cara kedua adalah dengan melakukan uji korelasi serial, yakni menguji hipotesis $H_0 : \rho_1 = \rho_2 = \dots = \rho_k, k < n$ (tidak terdapat korelasi serial dalam residual sampai lag- k , $k < n$). Uji ini dapat dilakukan dengan menggunakan statistik uji *Box-Pierce* $Q = \sum_{j=1}^k \rho(j)^2$, atau *Ljung Box* $Q = n(n+2) \sum_{j=1}^k \rho(j)^2 / (n-j)$ yang akan berdistribusi $\chi^2(k-(p+q))$, $k > (p+q)$. Disini $\rho(j)$ menunjuk kan nilai sampel ACF pada lag- j sedangkan p dan q menunjukkan order dari model ARMA(p,q) (Rosadi, 2013). Apabila hipotesis *diagnostic check* ditolak, maka model yang telah diidentifikasi di atas tidak dapat digunakan, dan selanjutnya dapat diidentifikasi kembali model yang mungkin sesuai untuk data. Dalam pemilihan metode terbaik kadangkalanya terdapat dua metode atau lebih yang lolos dari syarat-syarat yang ditentukan, jadi untuk pilihan metode terakhir dalam menentukan metode yang terbaik yakni dengan melihat nilai AIC (*Akaike Information Criterion*), AIC merupakan alat yang berguna dalam pemilihan model, AIC hanya dapat memberikan uji kualitas model yang relatif. Artinya, AIC tidak dan tidak dapat memberikan pengujian model yang menghasilkan informasi tentang kualitas model secara absolut. Jadi jika masing-masing model statistik yang diuji sama-sama tidak memuaskan atau tidak cocok untuk data, AIC tidak akan memberikan indikasi apa pun sejak awal, dengan rumus:

$$AIC = \ln(s_{m^2}) + 2m/T$$

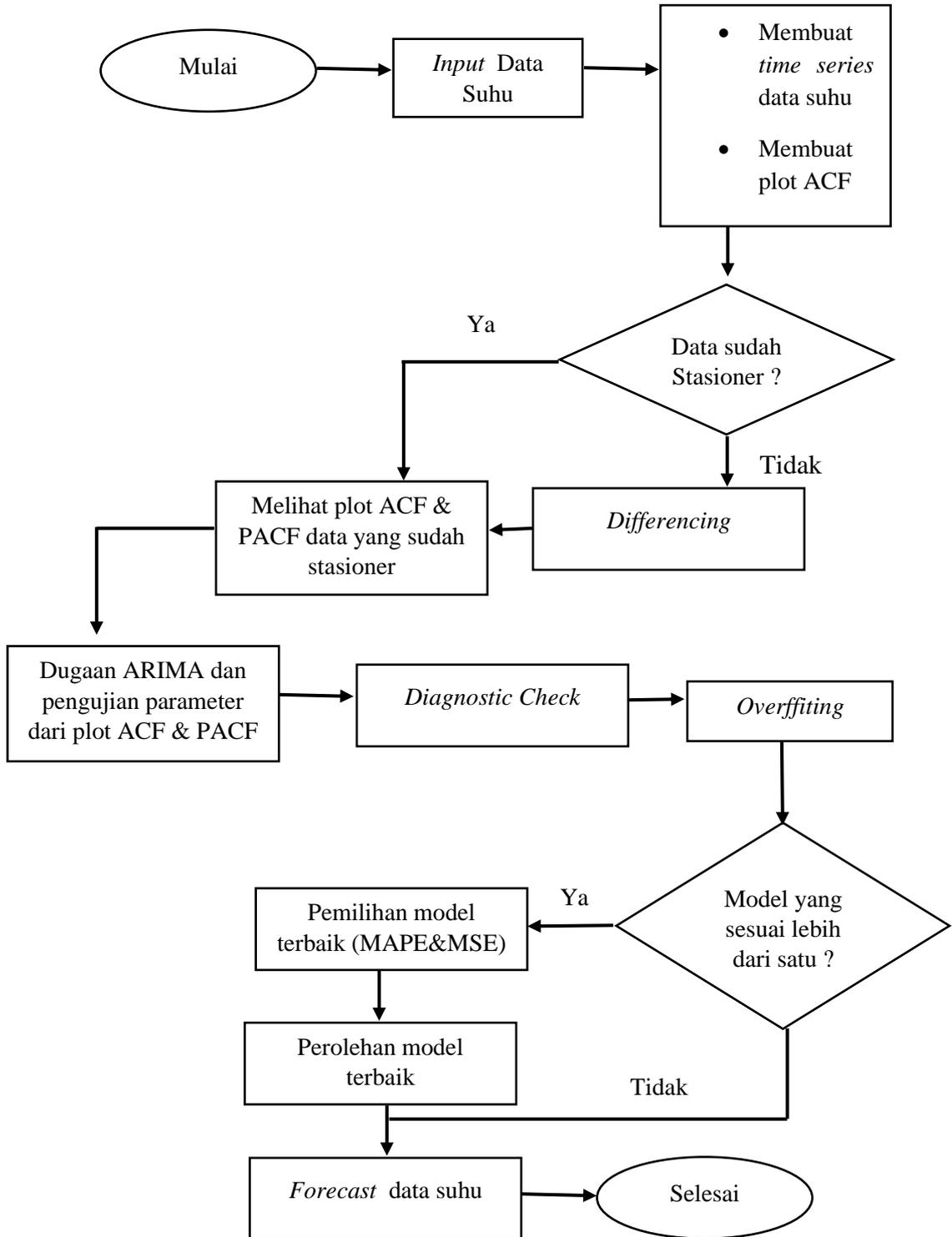
Dimana m adalah jumlah parameter dalam model, dan s_{m^2} (dalam contoh $AR(m)$) adalah *varians residual* yang diestimasi: $s_{m^2} = (\text{jumlah residu kuadrat untuk model } m)/T$. Itu adalah rata-rata sisa kuadrat untuk model m .

Kriteria dapat diminimalkan atas pilihan m untuk membentuk *trade-off* antara kecocokan model (yang menurunkan jumlah residu kuadrat) dan kompleksitas model, yang diukur dengan m . Jadi model $AR(m)$ versus $AR(m+1)$ dapat dibandingkan dengan kriteria ini untuk kumpulan data tertentu, dengan demikian model yang lebih disukai dalam hal kualitas relatif adalah model dengan nilai AIC minimum.

e. *Forecasting*

Setelah model terbaik diperoleh dari langkah-langkah pemodelan di model tersebut dapat digunakan untuk meramalkan sifat-sifat data di masa yang akan datang.

Berikut ini merupakan *flowchart* dari pemodelan ARIMA :



Gambar 3. Metode ARIMA untuk pembentukan model

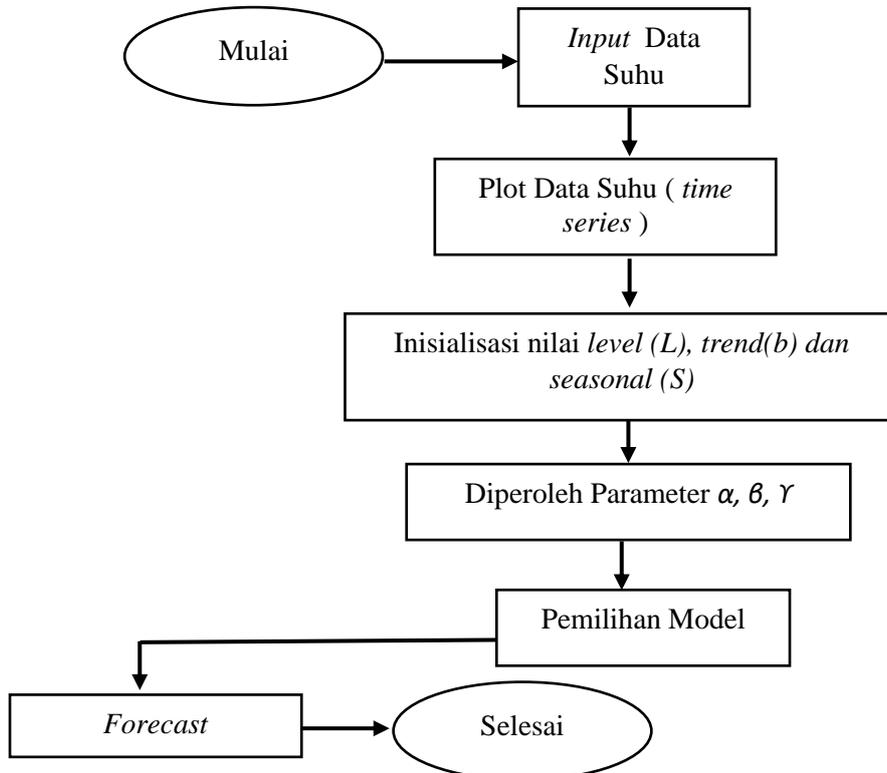
3. Pemodelan *Exponential Smoothing*

a. Pola Data

Menganalisis data yang diperoleh diubah kedalam bentuk *time series* dengan bantuan *software* Rstudio. Data yang telah diubah menjadi data *time series* kemudian dibuat plot data, kegunaan plot data ini untuk menentukan pola data berupa *trend*, musiman , atau tidak keduanya.

b. Pemilihan Model

Setelah plot data sudah diperoleh, yang dilakukan yaitu pemilihan salah satu model dari beberapa model yang terdapat dalam Metode *Exponential Smoothing*. Model yang telah didapatkan akan mendapatkan nilai parameter α , β , γ , kemudian melalui proses *forecasting* dan yang terakhir yaitu menentukan nilai *error* yang dihitung menggunakan MAPE & MSE. Berikut ini merupakan *flowchart* Pemodelan *Exponential Smoothing* :



Gambar 4. Metode *Exponential Smoothing* untuk pembentukan model

4. Dilakukan perbandingan masing-masing model dari dua metode tersebut, dengan melihat nilai *error* pada MAPE & MSE untuk menentukan metode terbaik.
5. *Forecasting* akhir untuk tahun 2022 dengan menggunakan metode terbaik, langkah terakhir yaitu menarik kesimpulan dari hasil peramalan tersebut.