KARNATAKA'S OZONE STORY:

TRENDS AND FORECASTS

MEET THE TEAM

SATHWIK NAG C V

GRADUATE STUDENT

MASTER'S IN DECISION ANALYTICS VCU , VA

SATYANARAYAN NALDIGA

GRADUATE STUDENT

MASTER'S IN DECISION ANALYTICS VCU , VA

SHREYA MISHRA

GRADUATE STUDENT

MASTER'S IN DECISION ANALYTICS VCU , VA

MADHUMITHA POLAMREDDY

GRADUATE STUDENT

MASTER'S IN DECISION ANALYTICS VCU , VA

Overview of the Dataset and Data Dictionary

Source: Karnataka State Pollution Control Board (KSPCB)
Dataset: Karnataka Air Quality Data

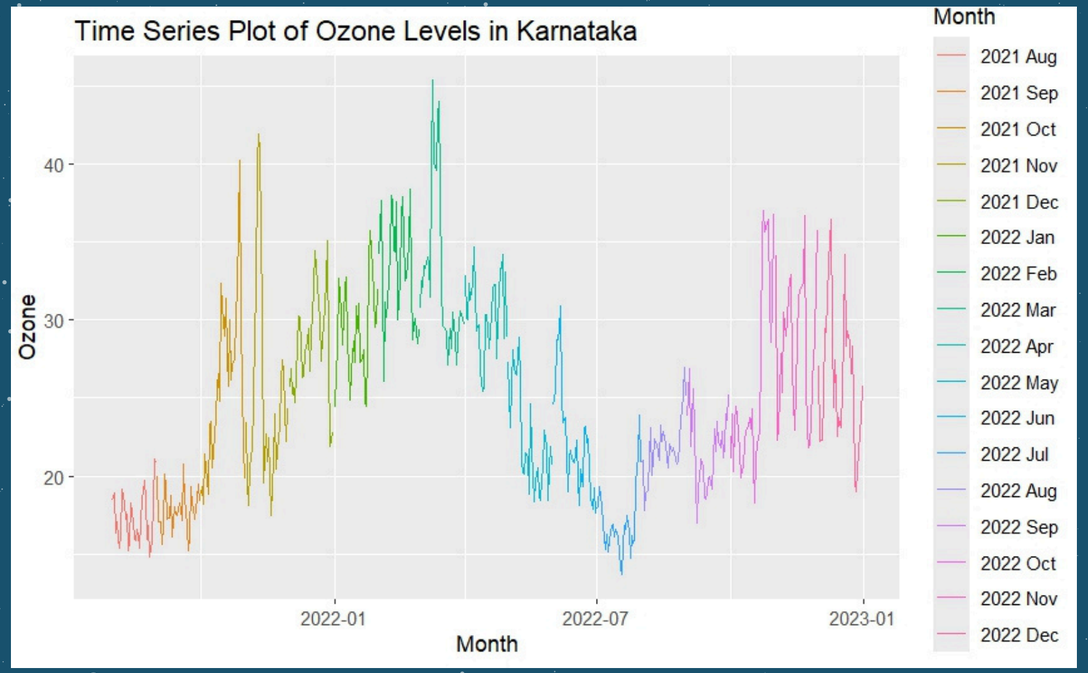
Timeframe: August 2021 - August 2023

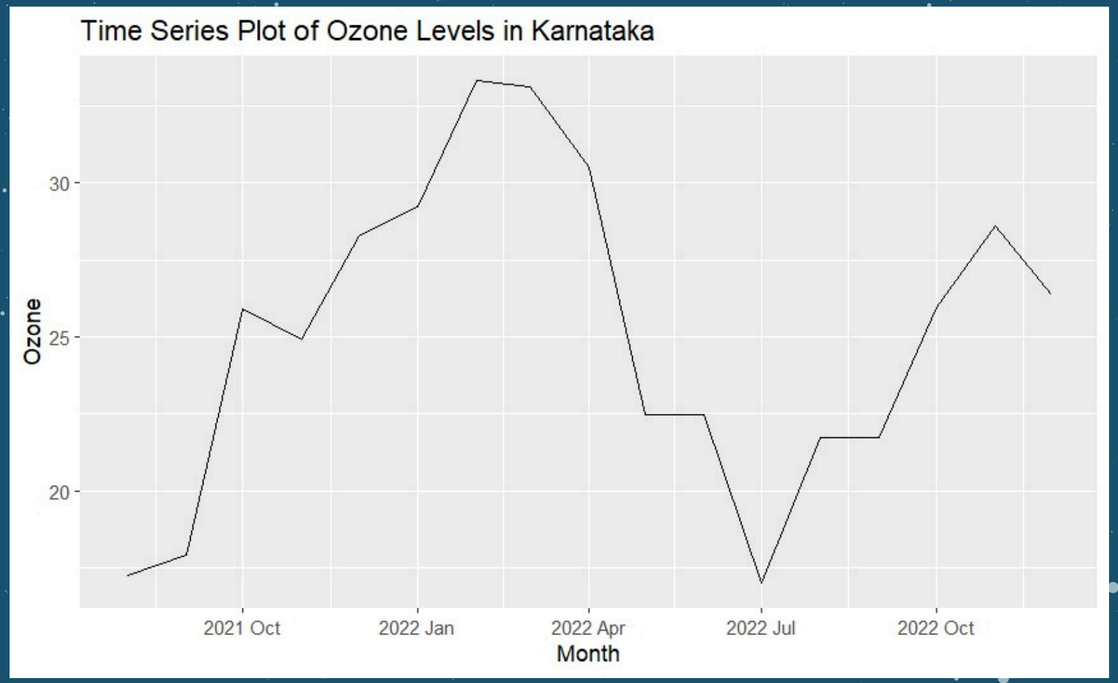
Relevance: The Karnataka air quality dataset is ideal for forecasting pollutant levels using techniques.

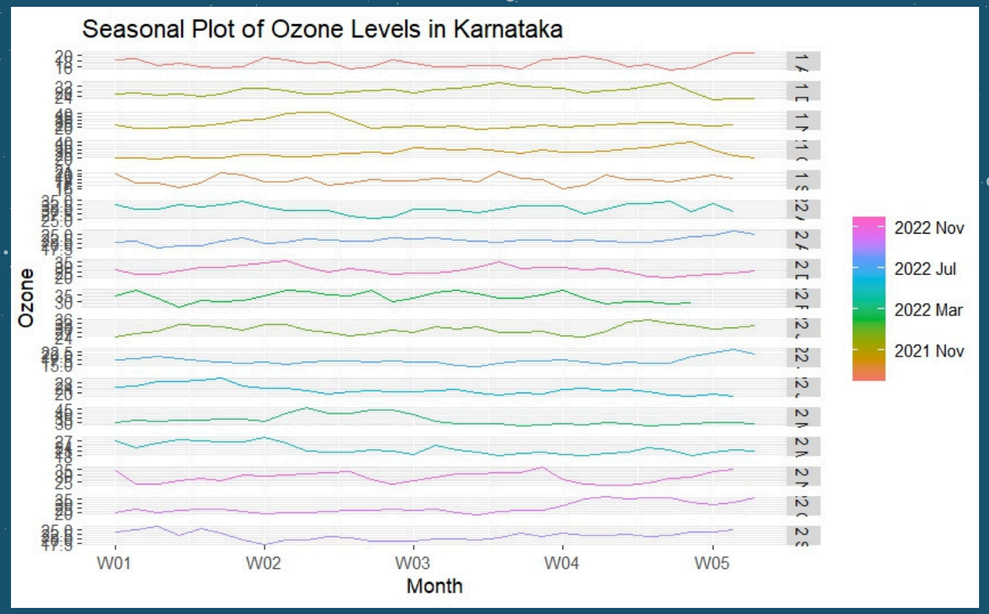
Cleaning : The dataset was cleaned by handling missing values.
No outliers were found, as the dataset covers a short duration and is well-structured.

Column Name	Description	Data Type
From Date	Start date of the observation period	String
To Date	End date of the observation period	String
Ozone	Level of Ozone (µg/m³)	Float
CO	Level of Carbon Monoxide (mg/m³)	Float
SO2	Level of Sulfur Dioxide (µg/m³)	String (needs cleaning)
NO2	Level of Nitrogen Dioxide (µg/m³)	Float
PM10	Particulate Matter (PM10) levels (µg/m³)	Float
PM2.5	Particulate Matter (PM2.5) levels (µg/m³)	Float
State	State name	String
City	City name	String
Station	Monitoring station name	String

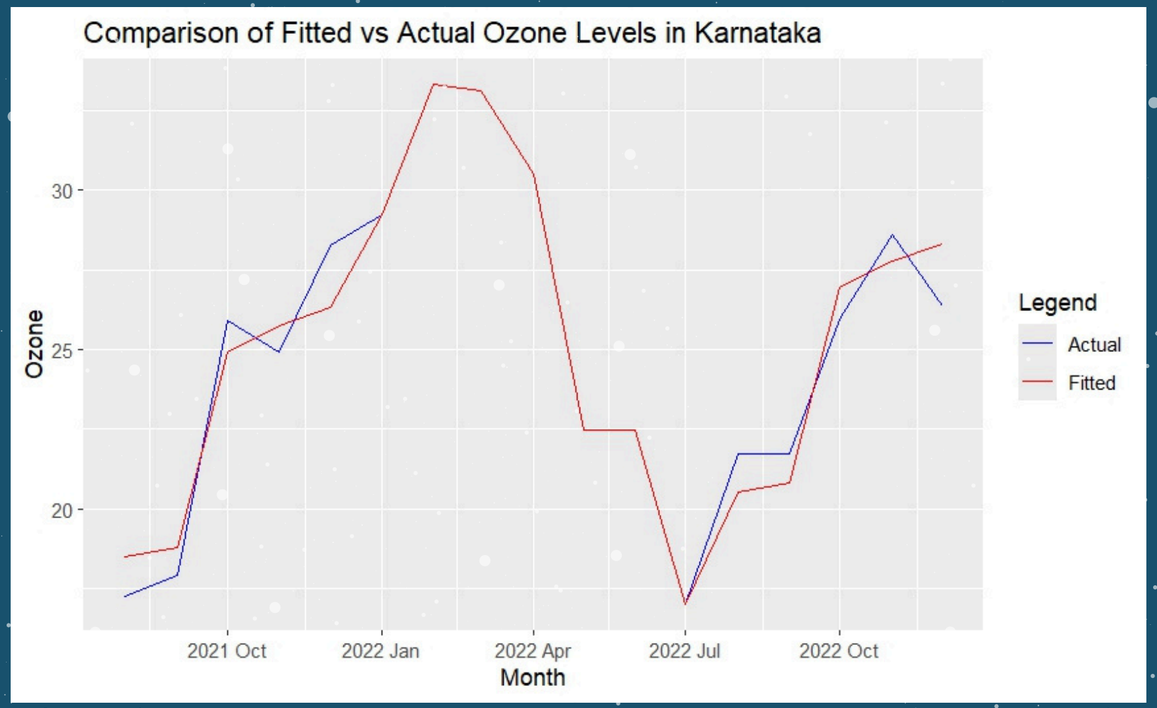
Time Series Regression: Unlocking Patterns Over Time

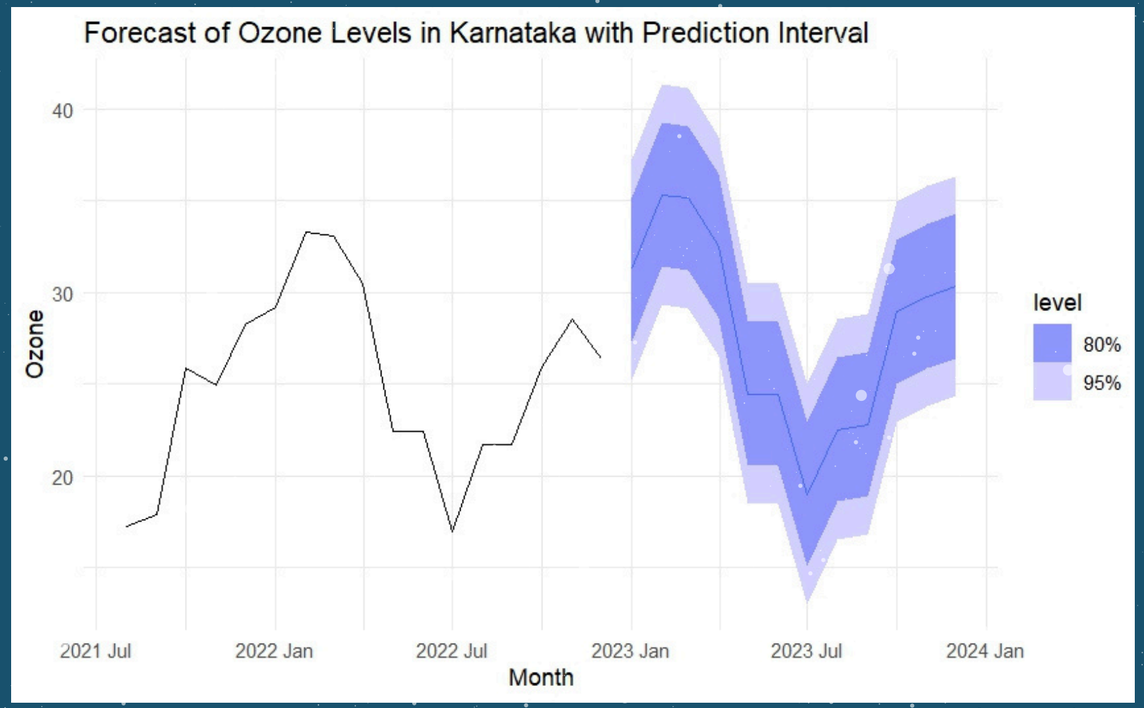






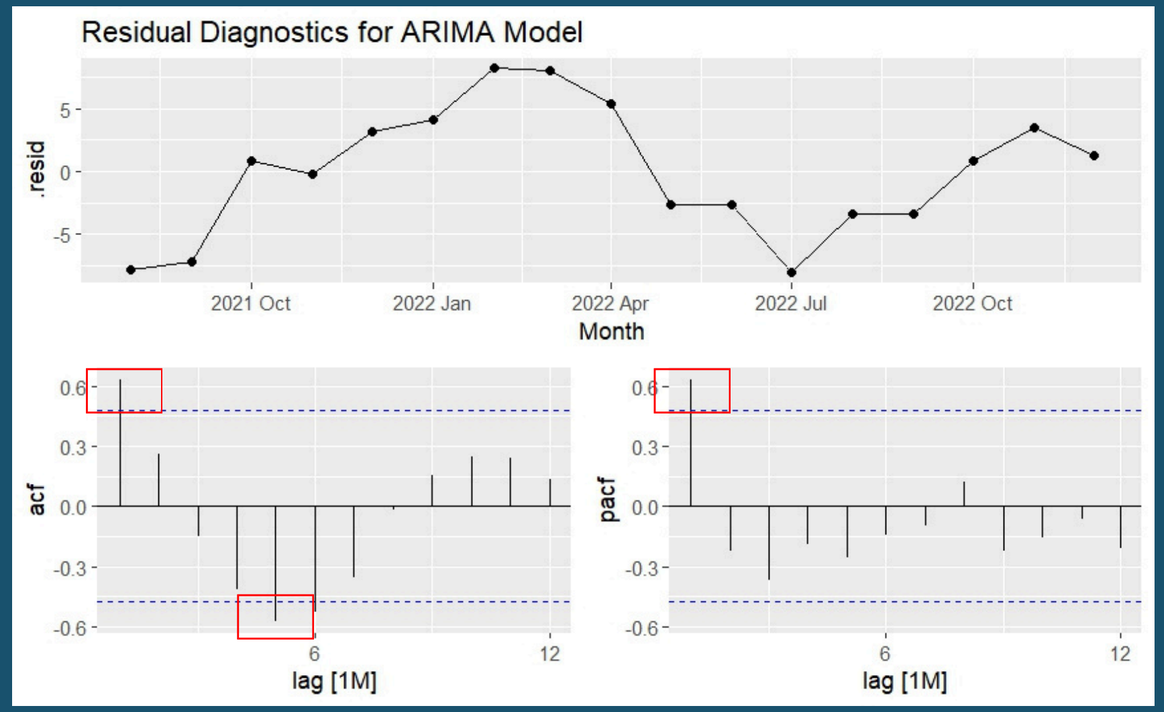
Analysis and Forecasting of Ozone Levels in Karnataka: Trends, Predictions, and Confidence Intervals

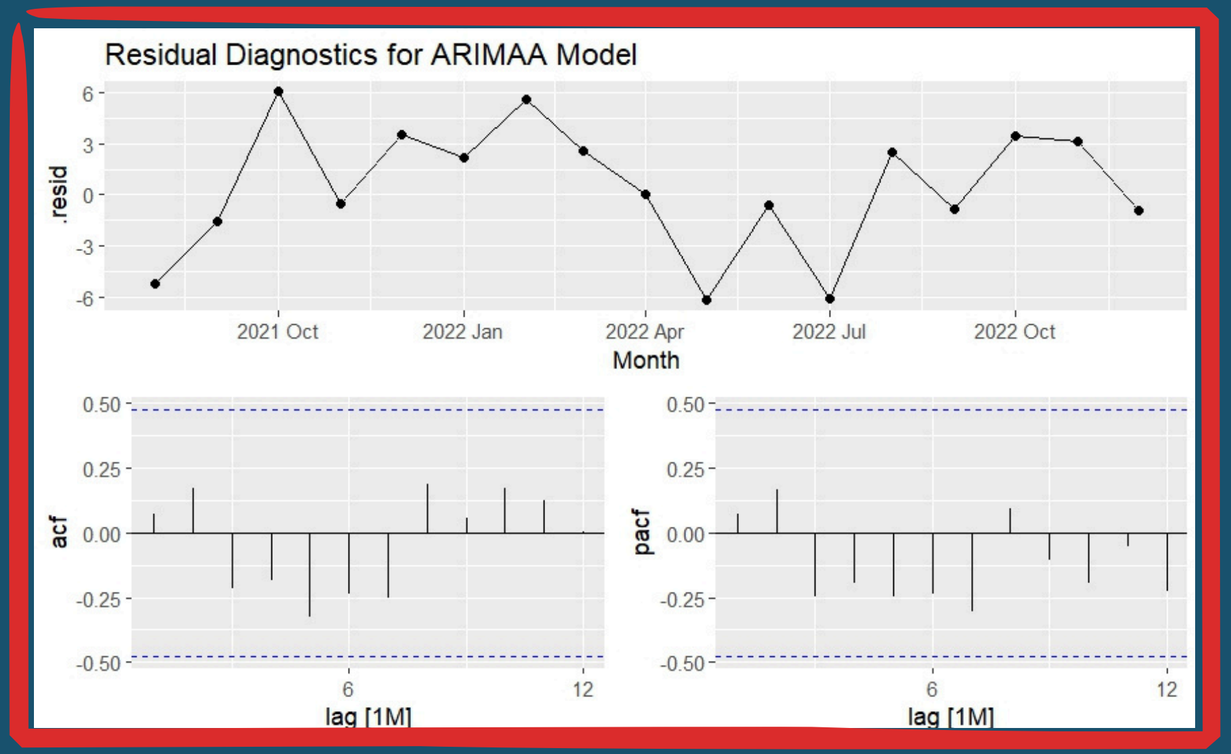




Adjusted R-squared: 0.8513

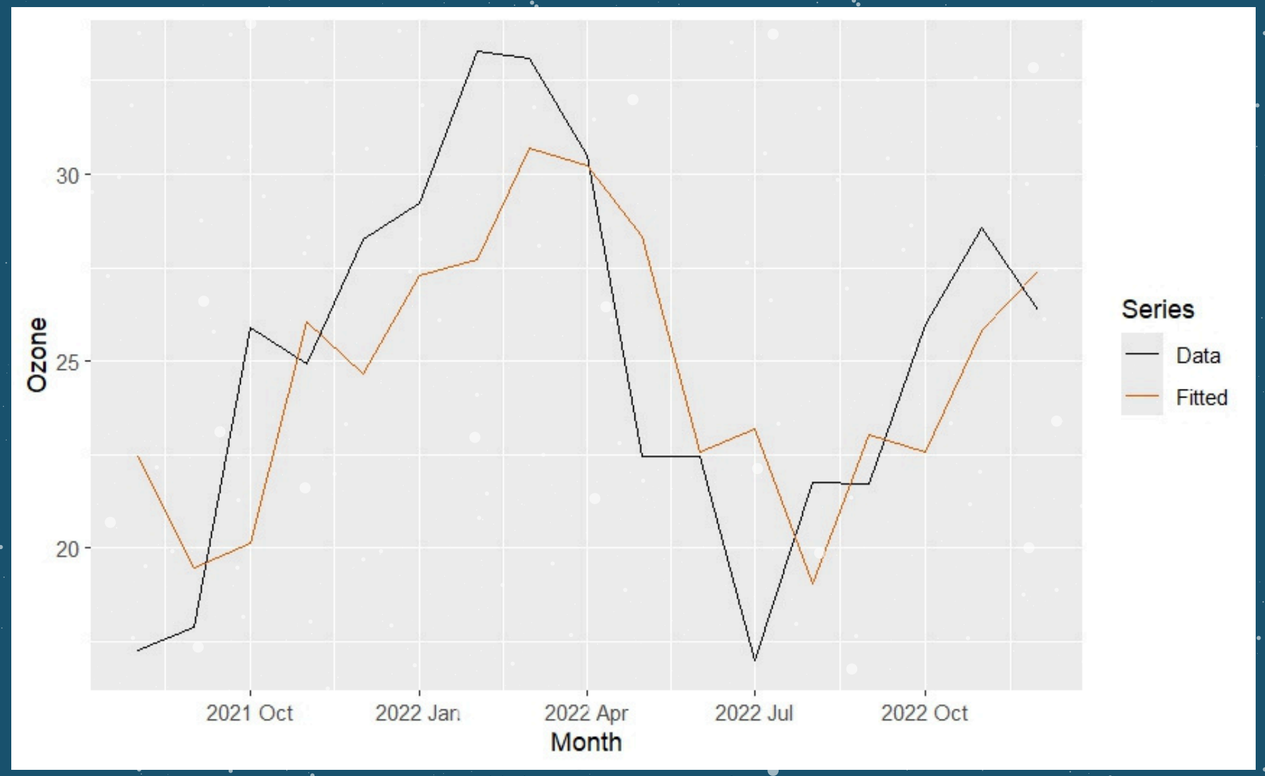
Residual Analysis and Model Fitting for ARIMA Forecasting

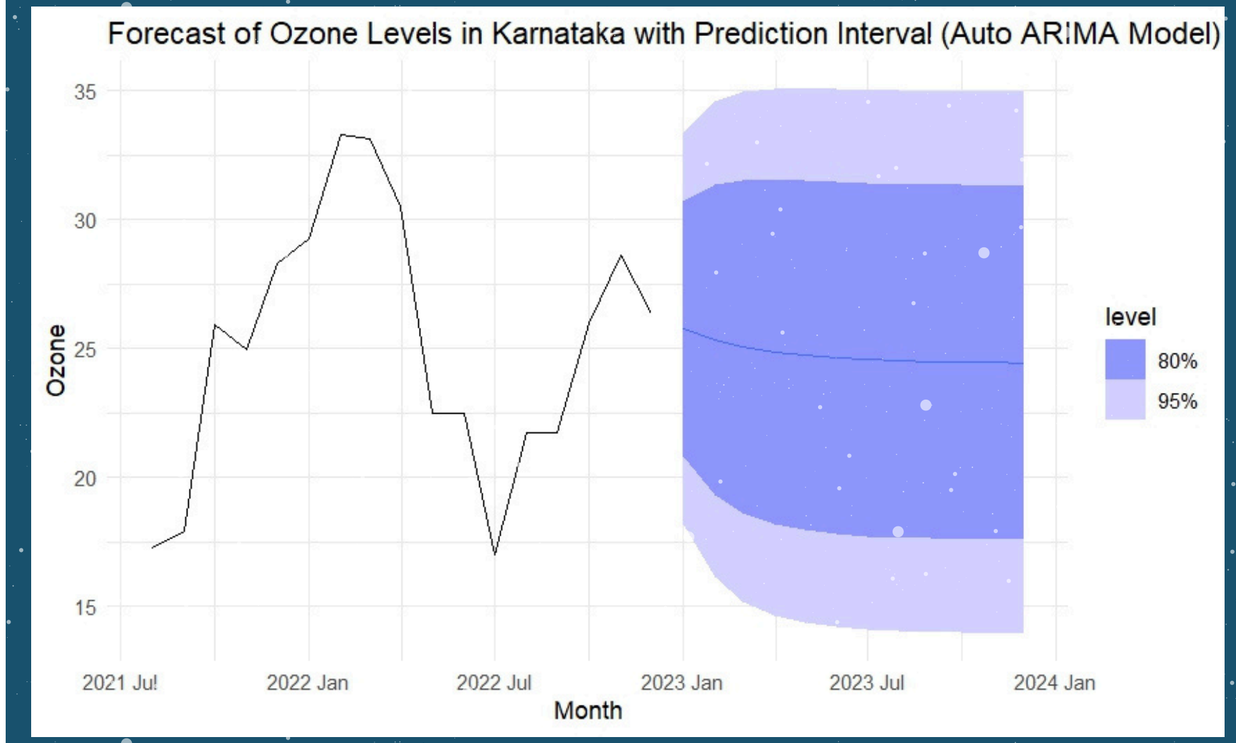




- pdq(0, 0, 0) + PDQ(0, 0, 0)), #AIC=106.71
- pdq(1, 0, 0) + PDQ(0, 0, 0) #AIC=98.71

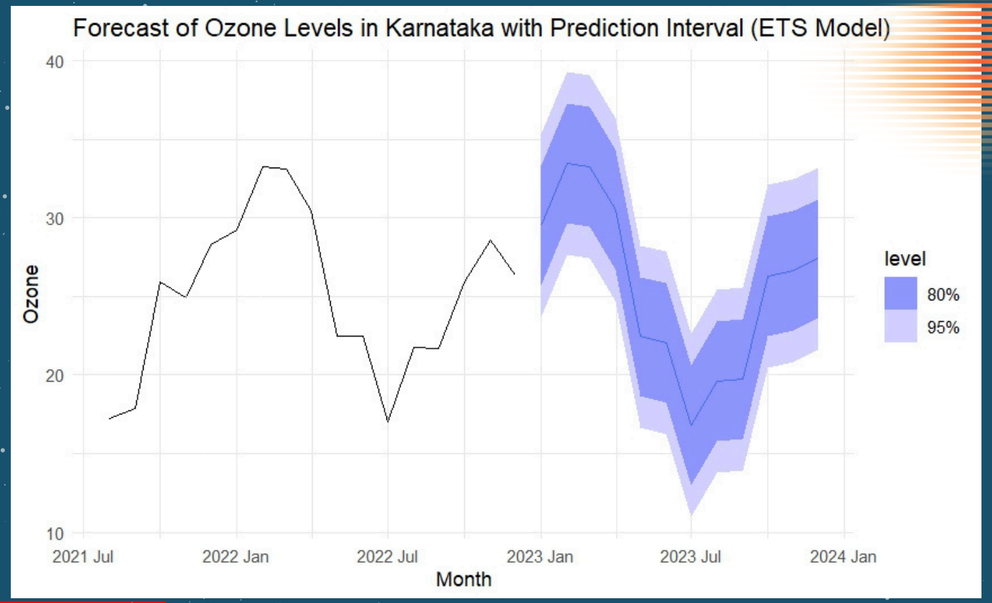
ARIMA Forecasting- AUTO





- pdq(1, 0, 0) PDQ(0,0,0) #AIC=98.71

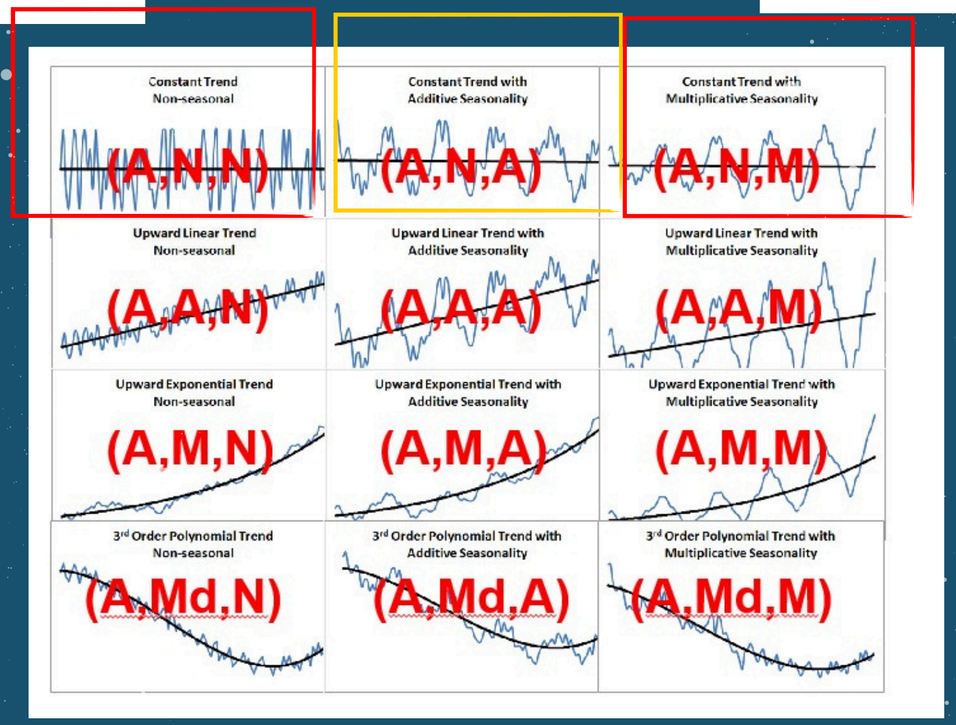
Exponential smoothing ETS- MANUAL



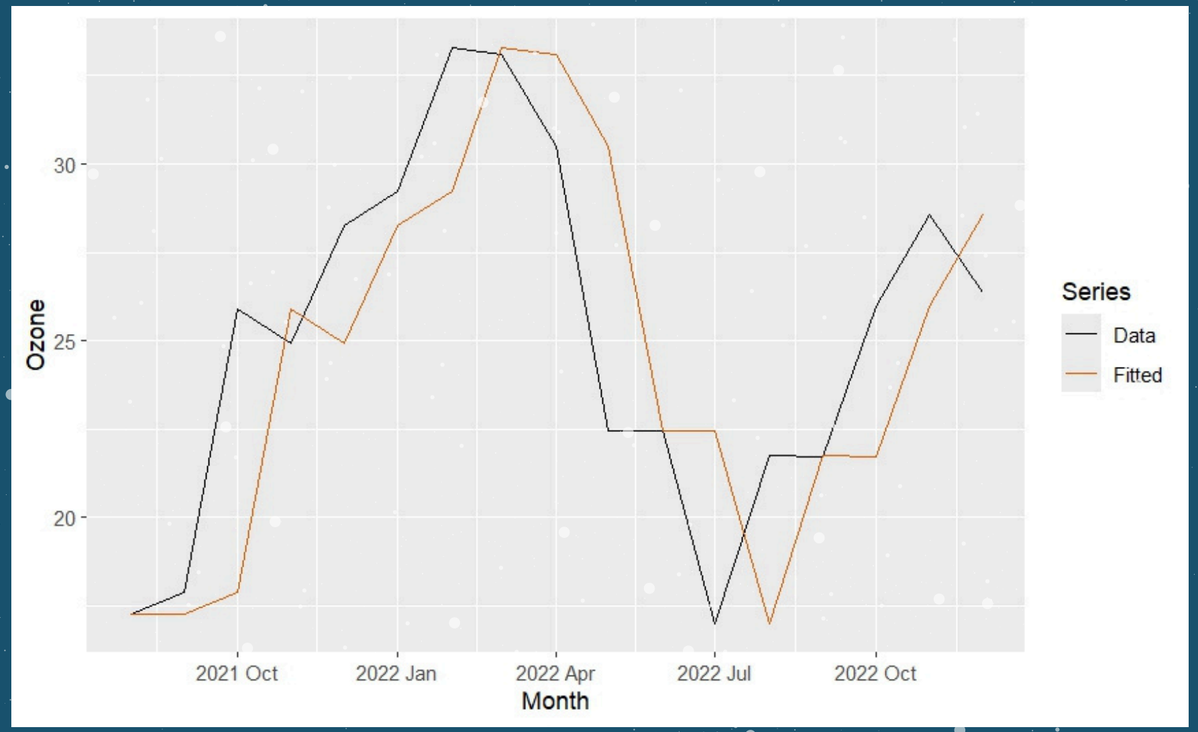
A,N, N= AIC 99.2

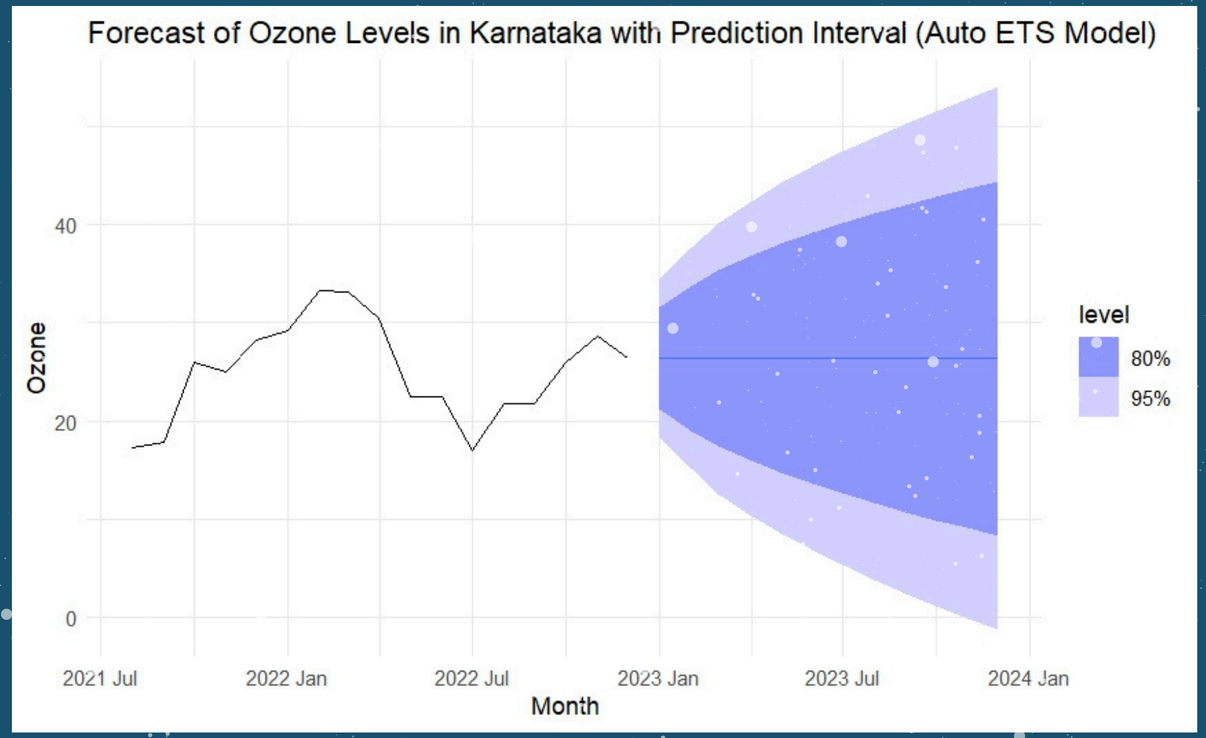
A ,N, A = AIC 85.6

A, N, M = AIC 86.3



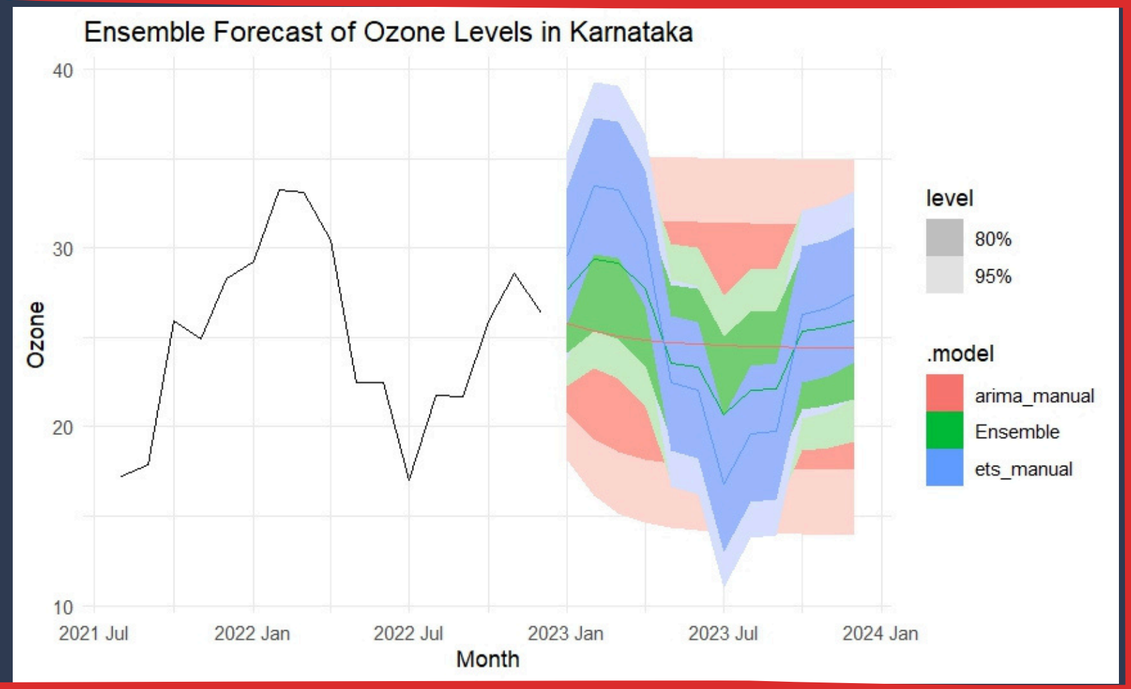
Exponential smoothing ETS- AUTO





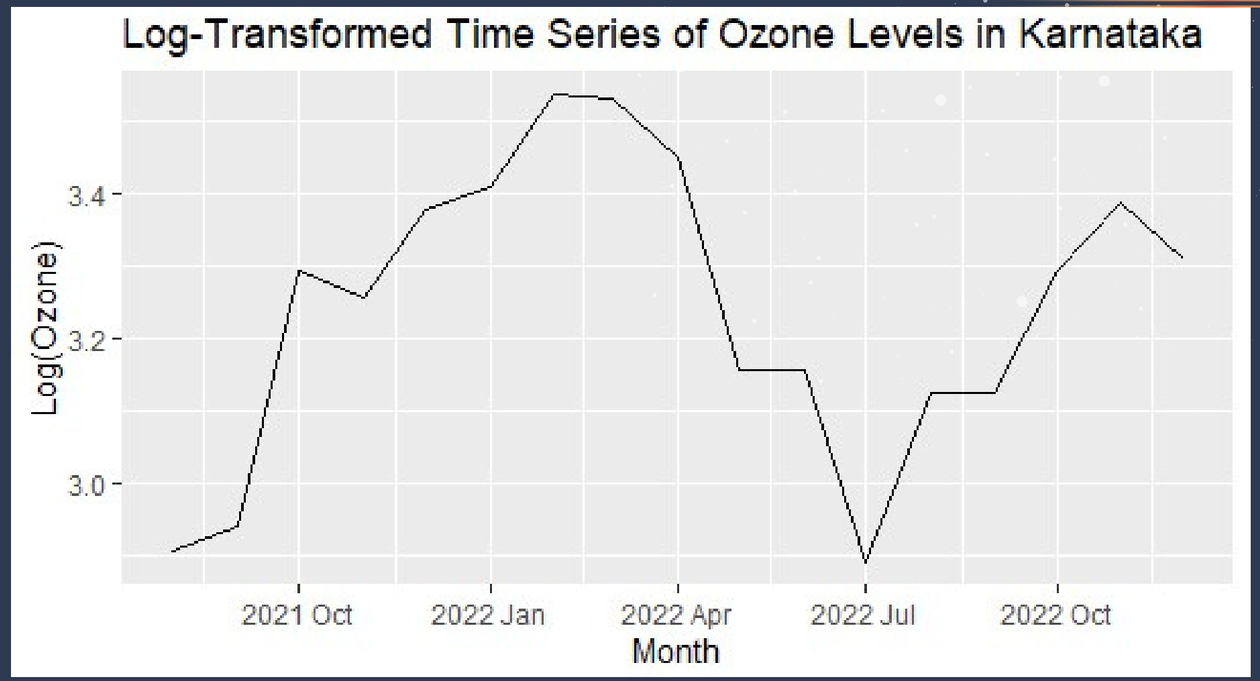
A ,N, N - AIC 99.7

Ensemble Forecasting



- The ensemble forecast combines ARIMA, ETS, regression, and ensemble models for more robust predictions.

Log Transformed Time Series



- The log transformation stabilizes variance and reduces the impact of extreme values in ozone levels.
- The graph shows a clear seasonal pattern with peaks around early 2022 and a sharp dip in mid-2022.

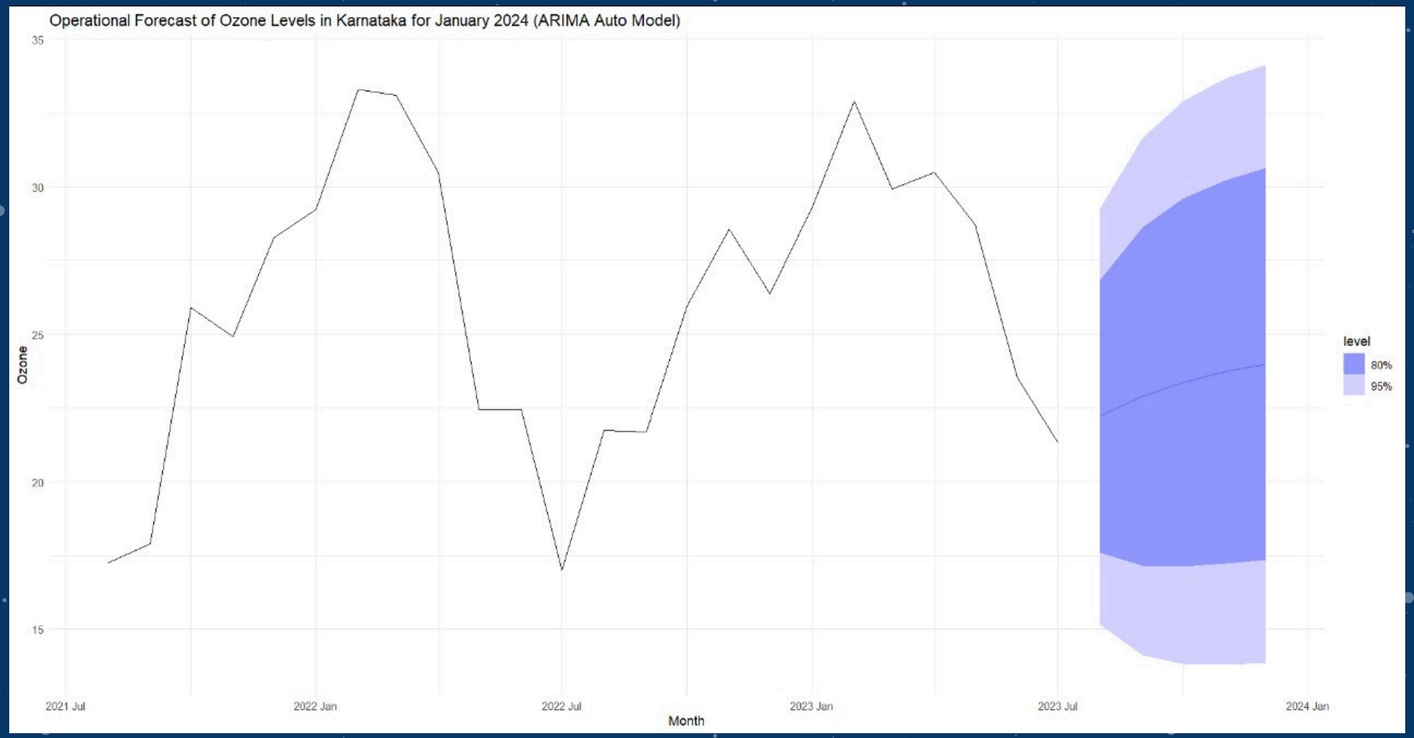
CROSS VALIDATION

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSE_adj	ACF1
arima_auto	Test	1.27	16.5	11.7	-56.7	83.1	1.45	1.24	0.91
arima_manual	Test	2.17	16.5	11.7	-56.7	83.1	1.45	1.24	0.91
naive	Test	-0.328	16.5	11.7	-56.7	83.1	1.45	1.24	0.91
ensemble (EQ)	Test	0.827	16.7	12	-59.5	85.6	1.48	1.26	0.912
combination (in-var)	Test	0.48	17.2	12.4	-61.6	88.6	1.54	1.29	0.916
ets_manual	Test	0.387	17.3	12	-62.2	89.6	1.56	1.3	0.918
ets_auto	Test	-0.328	16.5	11.7	-66.8	90.7	1.49	1.24	0.91
snaive	Test	0.0052	17.2	12.1	-64.6	90.7	1.55	1.29	0.917
ts_reg	Test	-2.01	17.4	13.1	-77.4	99.3	1.62	1.3	0.917

Why did we Choose ARIMA as the best model ?

We chose ARIMA as both manual and auto methods gave consistent pdq(1, 0, 0) + PDQ(0, 0, 0) with strong accuracy.

OPERATIONAL FORECAST




```
> # Print the forecasted values for January 2024
> print(arima_auto_forecast_january_2024)
# A fable: 5 x 4 [1M]
# Key:     .model [1]
  .model          Month     Ozone .mean
  <chr>           <mth>    <dist> <dbl>
1 ARIMA(Ozone) 2023 Aug N(22, 13)  22.2
2 ARIMA(Ozone) 2023 Sep N(23, 20)  22.9
3 ARIMA(Ozone) 2023 Oct N(23, 24)  23.4
4 ARIMA(Ozone) 2023 Nov N(24, 26)  23.7
5 ARIMA(Ozone) 2023 Dec N(24, 27)  24.0
>
```

- Ozone levels show significant variability with recurring peaks and troughs, indicating possible seasonal patterns or external influences.
- By late 2023, the levels have stabilized, suggesting less fluctuation in recent months.
- The mean forecast value suggests a slight upward trend in ozone levels, continuing the pattern from late 2023.

RECOMMENDATIONS

1 Public Health Advisory

2 Pollution Control Measures

3 Enhanced Monitoring and Predictive Capability

4 Urban Planning and Vegetation

5 Mitigation Planning for Critical Periods

Thank you for LISTENING!