Forecasting Global Environmental Trends (2000 – 2023)

Presented by: Piyusha Gurung
Raju Manthena Cholla
Srija Sama

Research Objective

Research Question

How can we forecast USA temperature trends using historical environmental data?

Rationale

- Understand long-term climate patterns
- Identify key environmental drivers of temperature change
- Apply forecasting models to predict future trends

Is This a Time Series Dataset?

Annual measurements (2000–2024)

Consistent, longitudinal data per country

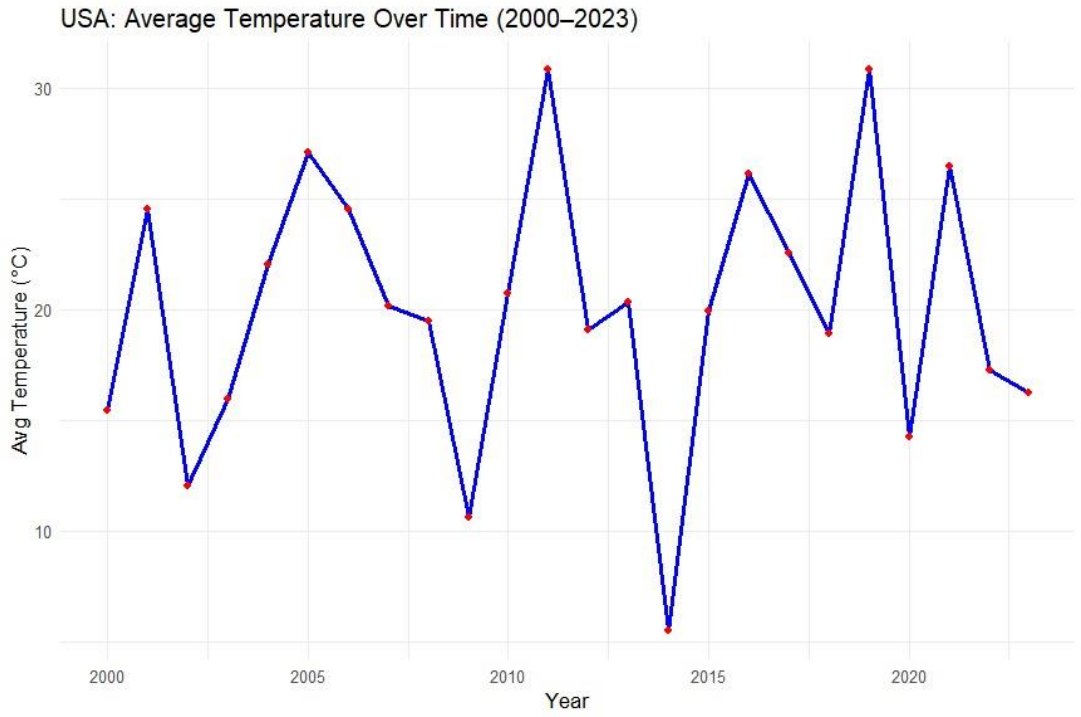
Suitable for forecasting models like ARIMA, ETS, and Linear Regression

**** Models chosen based on trend, seasonality, and data properties ****

Dataset Overview & Insights

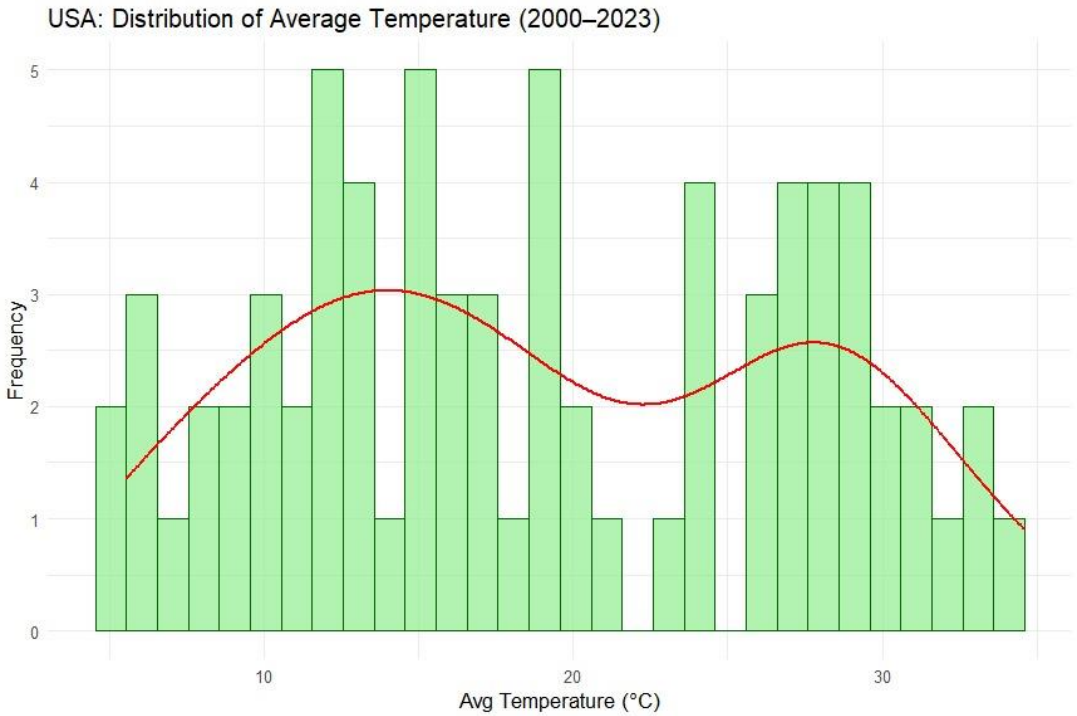
- **Dataset Title:** Global Environmental Trends (2000–2024)
- **Source**: Kaggle – temperature.csv
- **Frequency:** Annual (2000–2023)
- **Countries**: 15
- **Focus Country: USA**
- **Key Variable:** Average Temperature (°C)
- **Other Variables:** CO_2 Emissions, Sea Level Rise, Rainfall, Renewable Energy, Extreme Weather Events

Average Temperature Over Time (2000-2023)



- Year-to-year temperature variability is high
- Peaks in 2011 and 2019 show extreme values
- Pattern is irregular and non-seasonal

Distribution of Average Temperature Over Time (2000-2023)



- Multimodal distribution (multiple peaks)
- Frequent temps around 10 -20°C and 25 - 30°C
- High variability supports ARIMA & ETS models

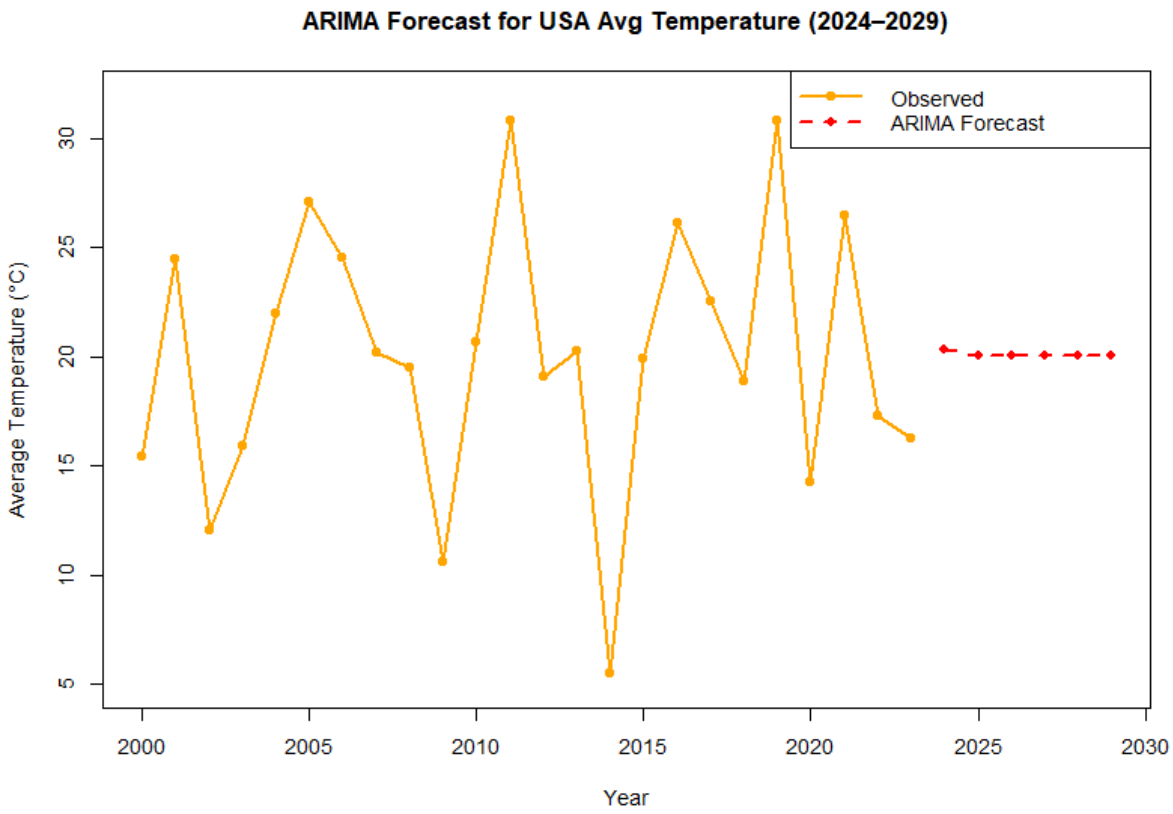
Forecasts from both models were plotted and compared:

- ARIMA and ETS both show upward trends
- Visual comparison shows model agreement on rising temperature
- Forecasts labeled through 2029 with legends

Forecasting Models

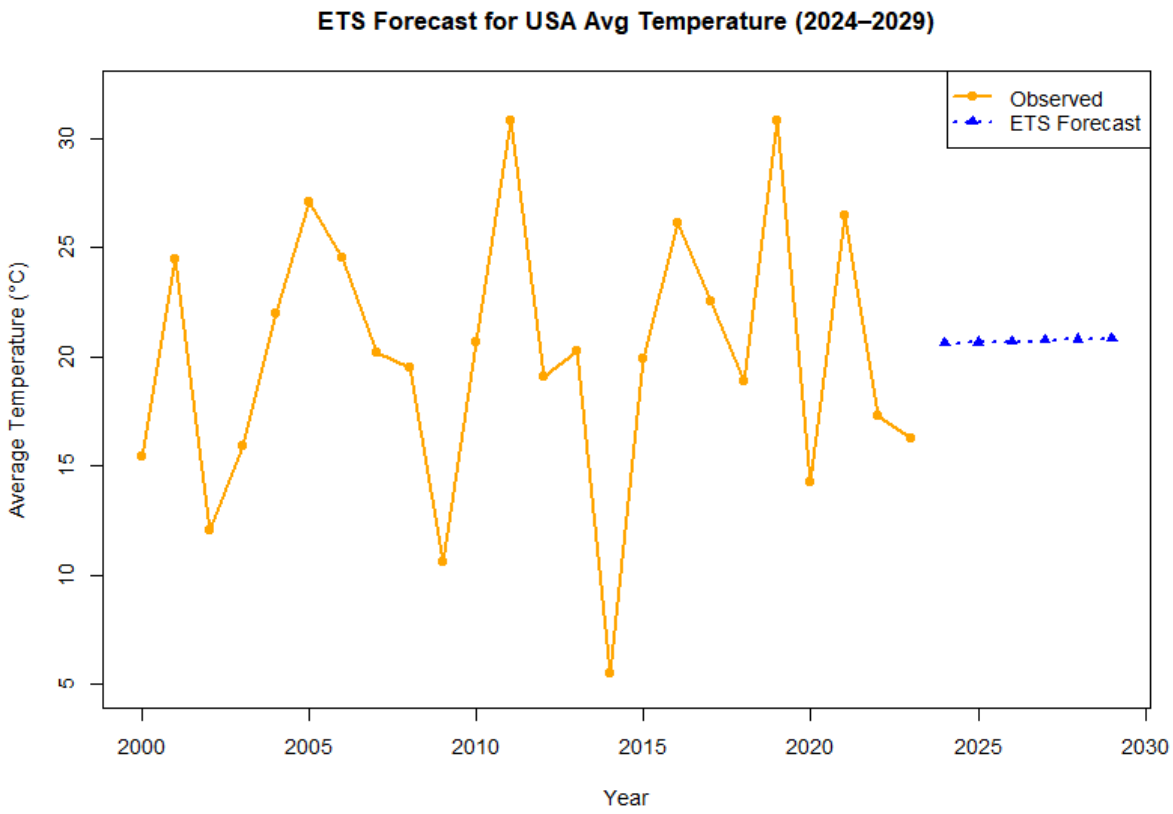
(ARIMA & ETS)

ARIMA FORECAST (2024 – 2029)



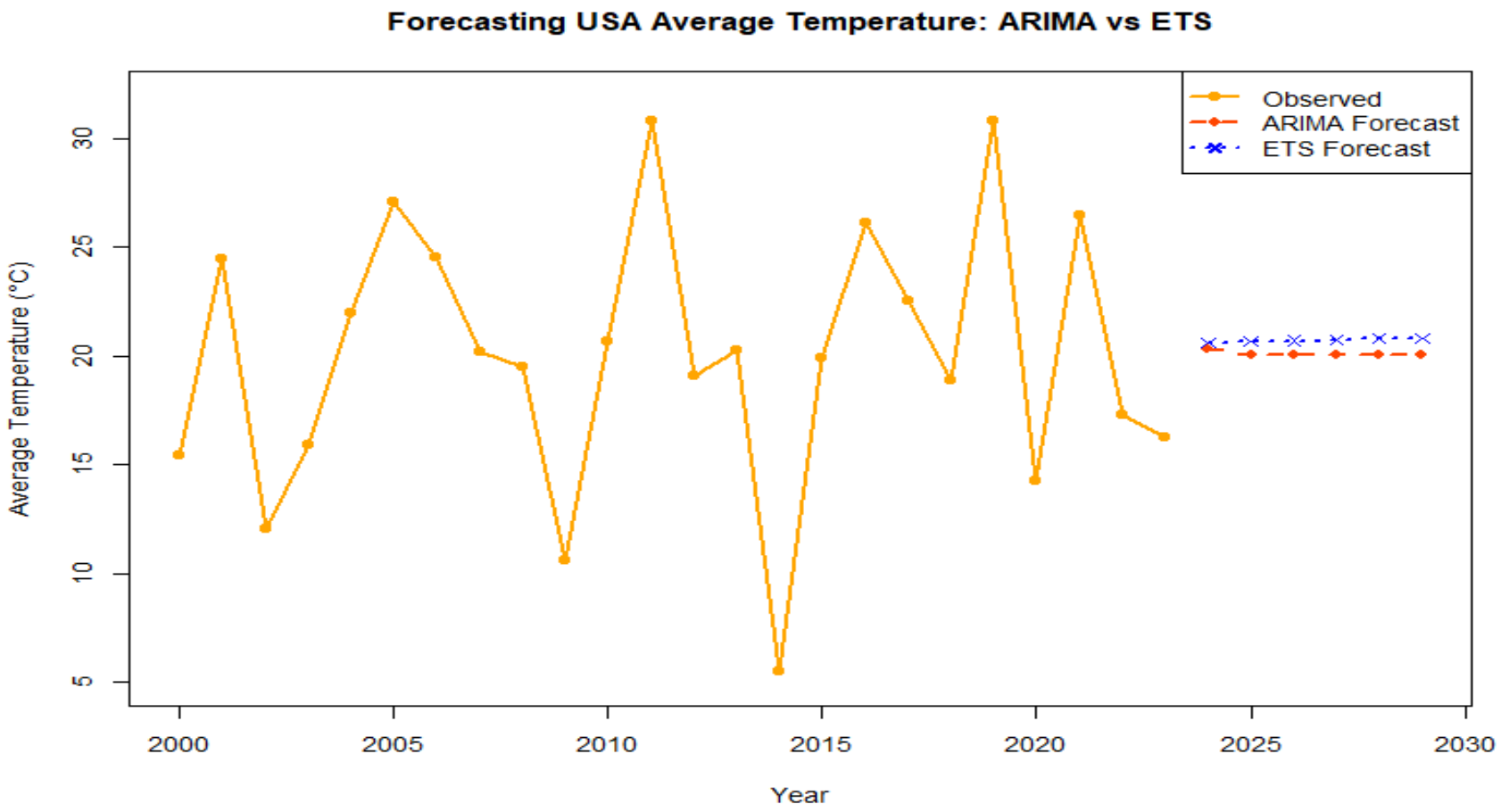
- ARIMA forecast predicts stable average temperatures around 20°C from 2024 to 2029.
- Reflects a flattening trend after years of historical fluctuations.
- Suggests no significant short-term rise or drop, based on past patterns.

ETS FORECAST (2024 – 2029)

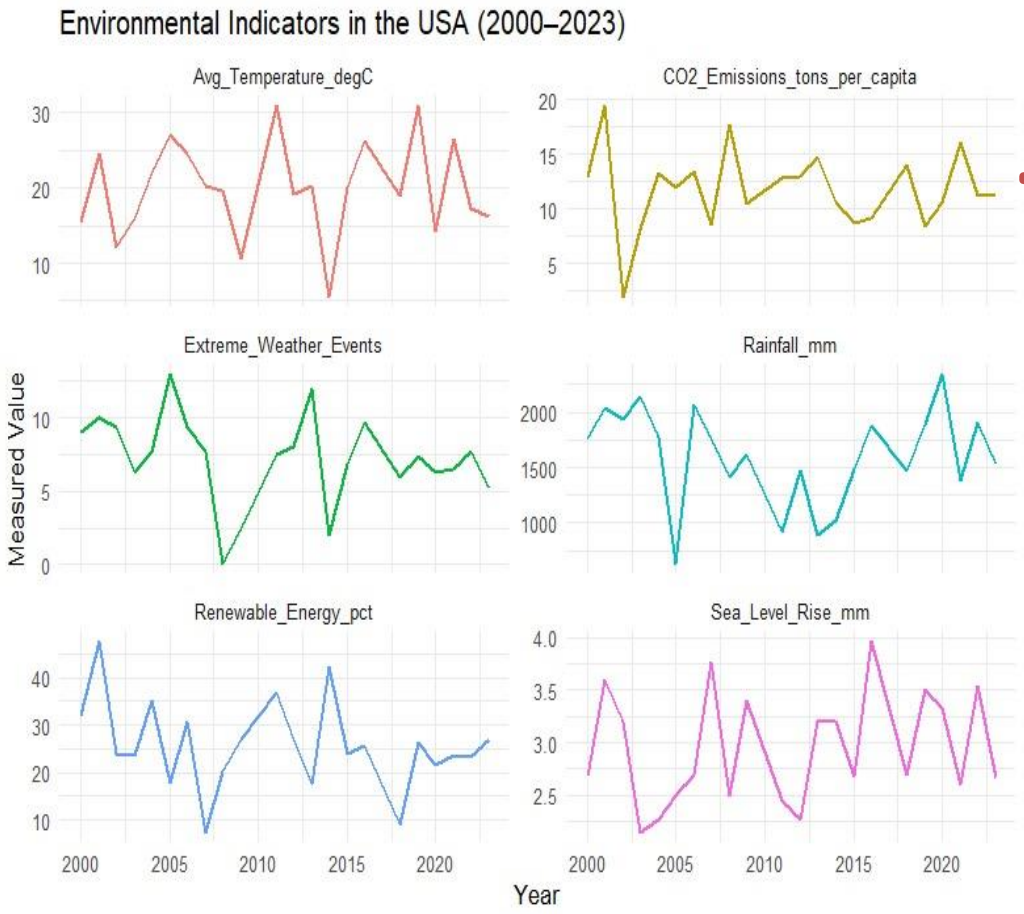


- ETS model forecasts a steady upward trend from 2024 to 2029 around 21°C.
- Captures long-term patterns using error, trend, and seasonality components.
- Suggests moderate but consistent warming, aligned with environmental trends.

ARIMA VS ETS


Forecasting USA Average Temperature: ARIMA vs ETS
Observed
ARIMA Forecast
ETS Forecast
Average Temperature (°C)
5
10
15
20
25
30
2000
2005
2010
2015
2020
2025
2030
Year

Environmental Indicators



- Key indicators like CO_2 emissions, rainfall, and sea level rise shown over time
- CO_2 dropped in 2008, likely tied to the economic recession
- Renewable energy shows no clear upward trend, while sea level rise is persistent
- Helps understand factors that might influence temperature changes

a. Regression Analysis (lm()):

- Predicts average temperature based on CO_2 emissions, rainfall, renewable energy %, population, and forest area
- summary(reg_model) gives coefficients, p-values, and R^2 for interpretation

b. Multivariate Visualization (ggpairs):

- Visualizes correlations and pairwise relationships between variables
- Helps detect multicollinearity or weak associations before modeling

Forecasting Models

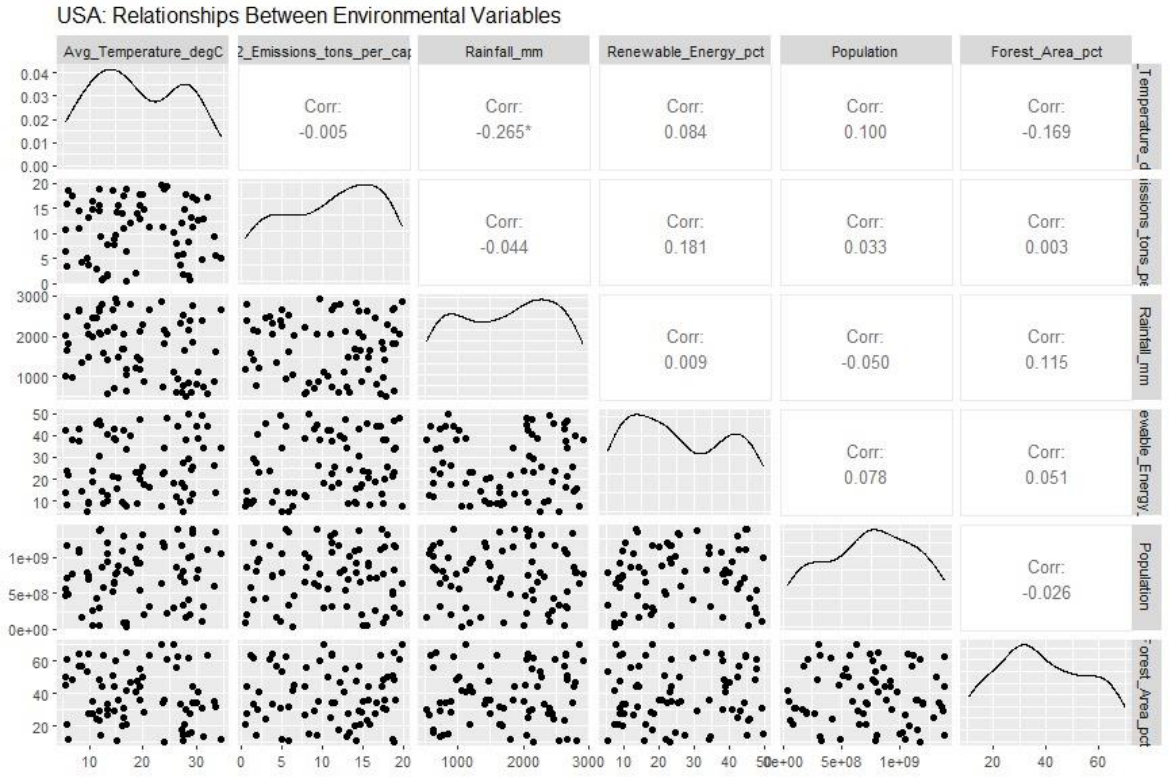
(Regression & Multivariate Visualization)

Regression Analysis

```
Call:
lm(formula = Avg_Temperature_degC ~ ., data = regression_data)

Residuals:
    Min      1Q  Median      3Q     Max
-15.369  -6.232  -0.614   7.295  15.044

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2.444e+01  4.284e+00   5.705 2.87e-07
CO2_Emissions_tons_per_capita -4.925e-02  1.667e-01  -0.295   0.7685
Rainfall_mm                  -2.784e-03  1.317e-03  -2.115   0.0382
Renewable_Energy_pct          5.675e-02  7.115e-02   0.798   0.4279
Population                    1.642e-09  2.453e-09   0.670   0.5055
Forest_Area_pct              -7.012e-02  5.690e-02  -1.232   0.2222

(Intercept)                  ***
CO2_Emissions_tons_per_capita
Rainfall_mm                  *
Renewable_Energy_pct
Population
Forest_Area_pct
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.178 on 67 degrees of freedom
Multiple R-squared:  0.1054,    Adjusted R-squared:  0.03863
F-statistic: 1.579 on 5 and 67 DF,  p-value: 0.1781
```

- Rainfall was the only moderately significant predictor ($p \approx 0.038$)
- Other variables like CO_2, renewable energy, and forest area were not statistically significant
- Model had low explanatory power (Adjusted $R^2 \approx 3.9\%$)

Multivariate Regression & Correlation Analysis



- Rainfall shows a modest negative correlation with temperature (Corr: -0.265)
- Other variables (e.g., CO_2, population, forest area) show weak or no significant correlation
- Suggests that temperature trends are influenced by multiple interacting factors, not just one variable

Conclusion

- ✓ analyzed average temperature trends in the USA from 2000 to 2023 using time series data
- ✓ forecasted short-term trends with ARIMA and ETS models, both indicating temperature stability
- ✓ found that climate patterns are shaped by multiple environmental factors, not a single variable

Thank You