R Project

Temperature Trends and Electricity Generation

Submitted By:
Anjana Sankar
Majida Nasrin M P
Rishan P

Dataset

- Source: data.gov.in
- Annual & seasonal mean temperatures (1980–2021)
- Yearly electricity generation by source (Coal, Wind, Solar, Hydro, etc.)
- Cleaned for missing values, formatted for time-series analysis
- Final dataset links climate patterns with energy trends

Introduction

- Climate change is raising global temperatures
- Energy demand and supply are closely linked to climate
- Studying both helps understand future risks and needs
- Goal: analyze temperature trends and electricity generation

Raw Dataset (Temperature)

	YEAR	ANNUAL	JAN.FEB	MAR.MAY	JUN.SEP	OCT.DEC
	<int>	<chr>	<chr>	<chr>	<chr>	<chr>
1	1901	25.42	20.11	27.64	28.16	23.1
2	1902	25.42	20.88	27.96	27.98	22.51
3	1903	25.01	19.99	27.02	28	22.33
4	1904	24.93	19.76	27.23	27.57	22.56
5	1905	24.84	18.36	26.38	28.2	23.13
6	1906	25.18	19.75	27.39	27.78	23.11

Data Cleaning & Processing

- Converted YEAR to Date → enables time-series plots.
- Converted columns to numeric → allows calculations.
- Dropped NAs → ensures no missing values in analysis.
- Filtered 1980–2021 → focused on reliable, recent data.
- Saved cleaned datasets for reproducibility

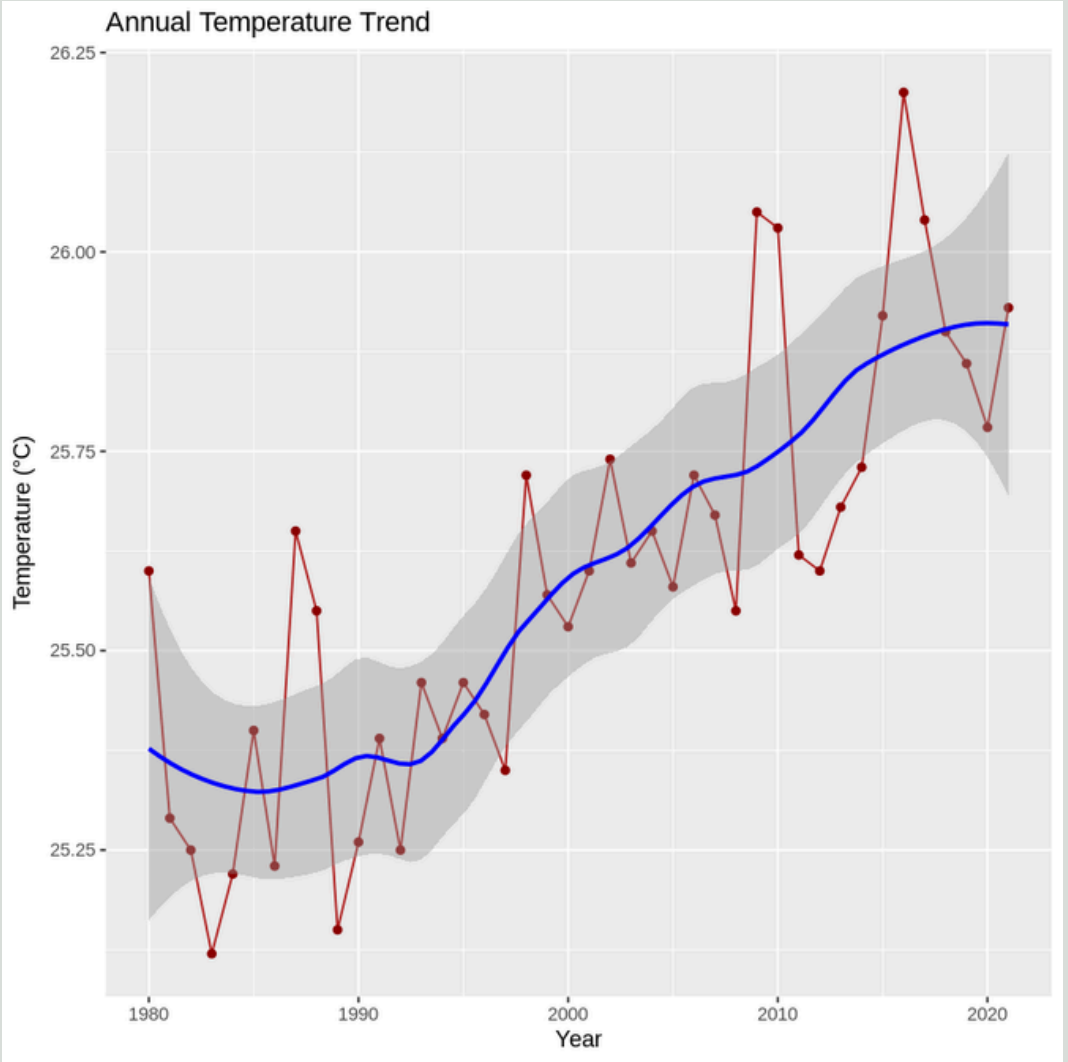
Cleaned Dataset (Temperature)

	YEAR	ANNUAL	JAN.FEB	MAR.MAY	JUN.SEP	OCT.DEC
	<date>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	1980-01-01	25.60	20.82	28.00	27.92	23.31
2	1981-01-01	25.29	20.50	27.39	27.96	22.84
3	1982-01-01	25.25	20.41	26.81	28.07	23.18
4	1983-01-01	25.12	20.09	26.93	28.12	22.68
5	1984-01-01	25.22	19.94	27.86	27.65	22.87
6	1985-01-01	25.40	20.62	28.01	27.71	22.89

Annual Temperature Trend

Annual mean temperature plot with LOESS smoothing

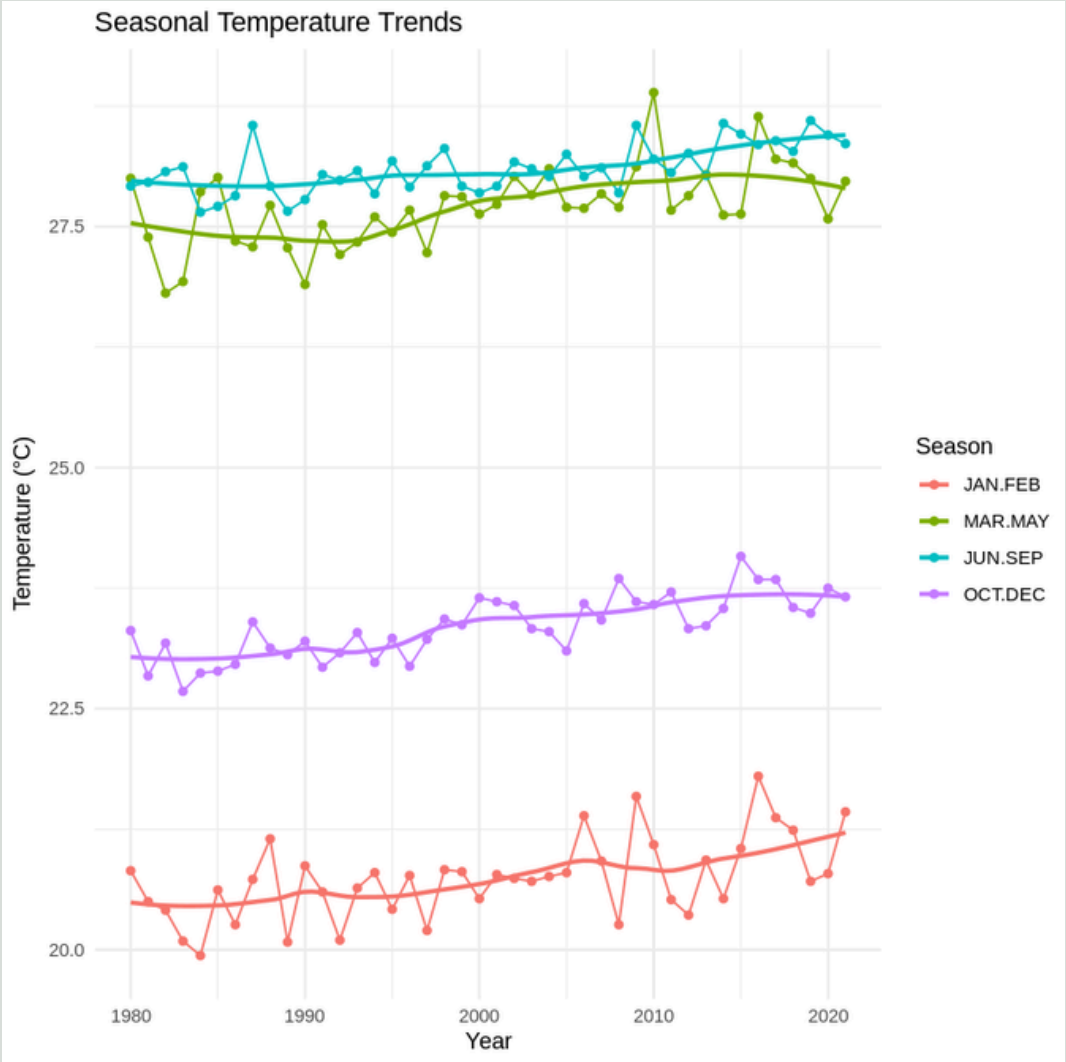
- Plotted the annual mean temperature over the years.
- Observed a gradual warming trend despite natural variability.
- The trend highlights that recent decades are generally warmer than earlier ones(Gradual warming trend since 1980).
- LOESS curve highlights fluctuations & El Niño spikes



Seasonal Temperature Trends

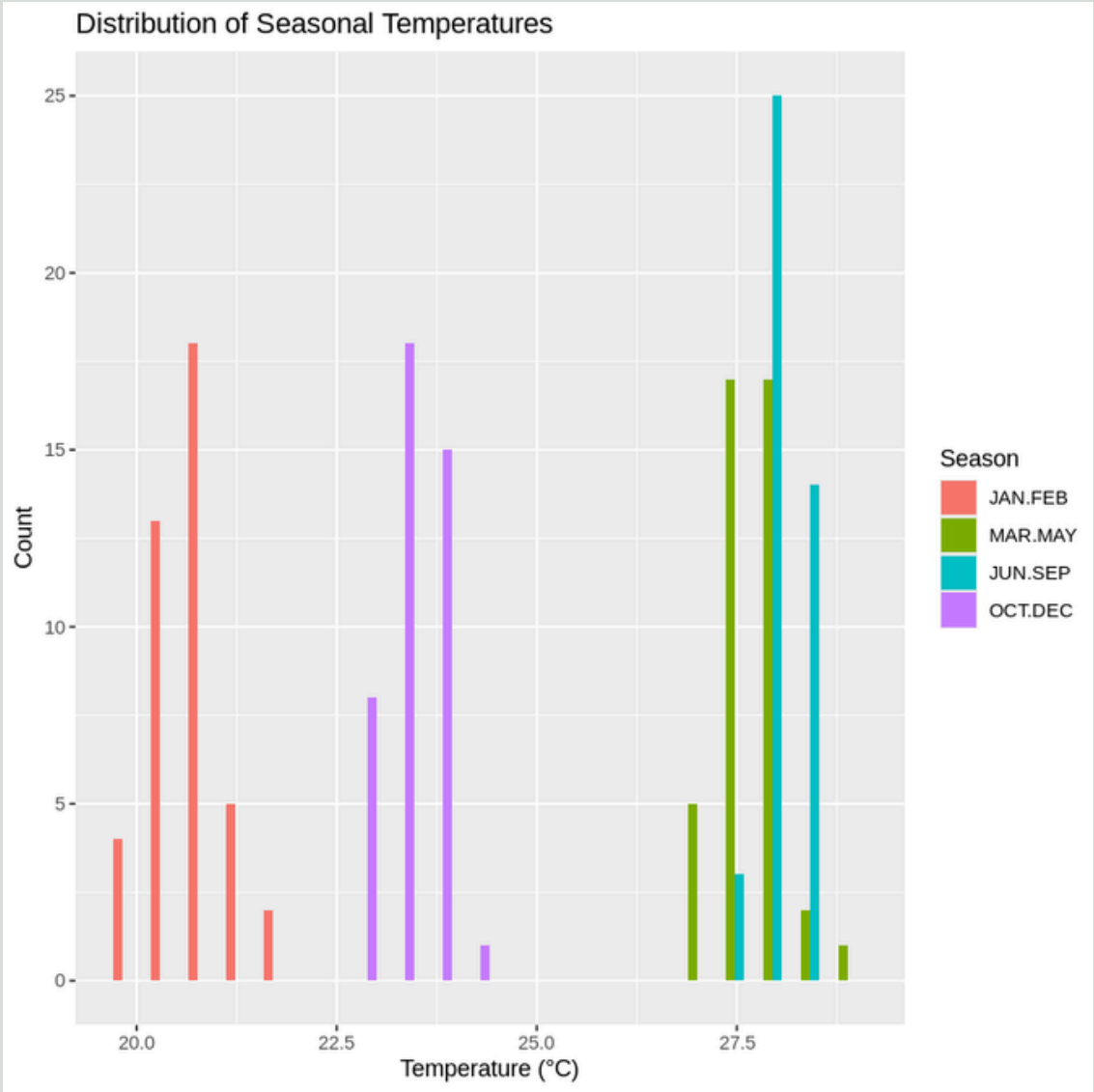
Seasonal plots (JAN–FEB, MAR–MAY, JUN–SEP, OCT–DEC)

- Seasonal plots show that some seasons (e.g., summer months) have higher variability, while winters are comparatively stable.
- Jan–Feb: stable
- Mar–May & Jun–Sep: higher variability
- Oct–Dec: moderate variation
- Insight: Seasonal differences may affect energy demand patterns and natural events.



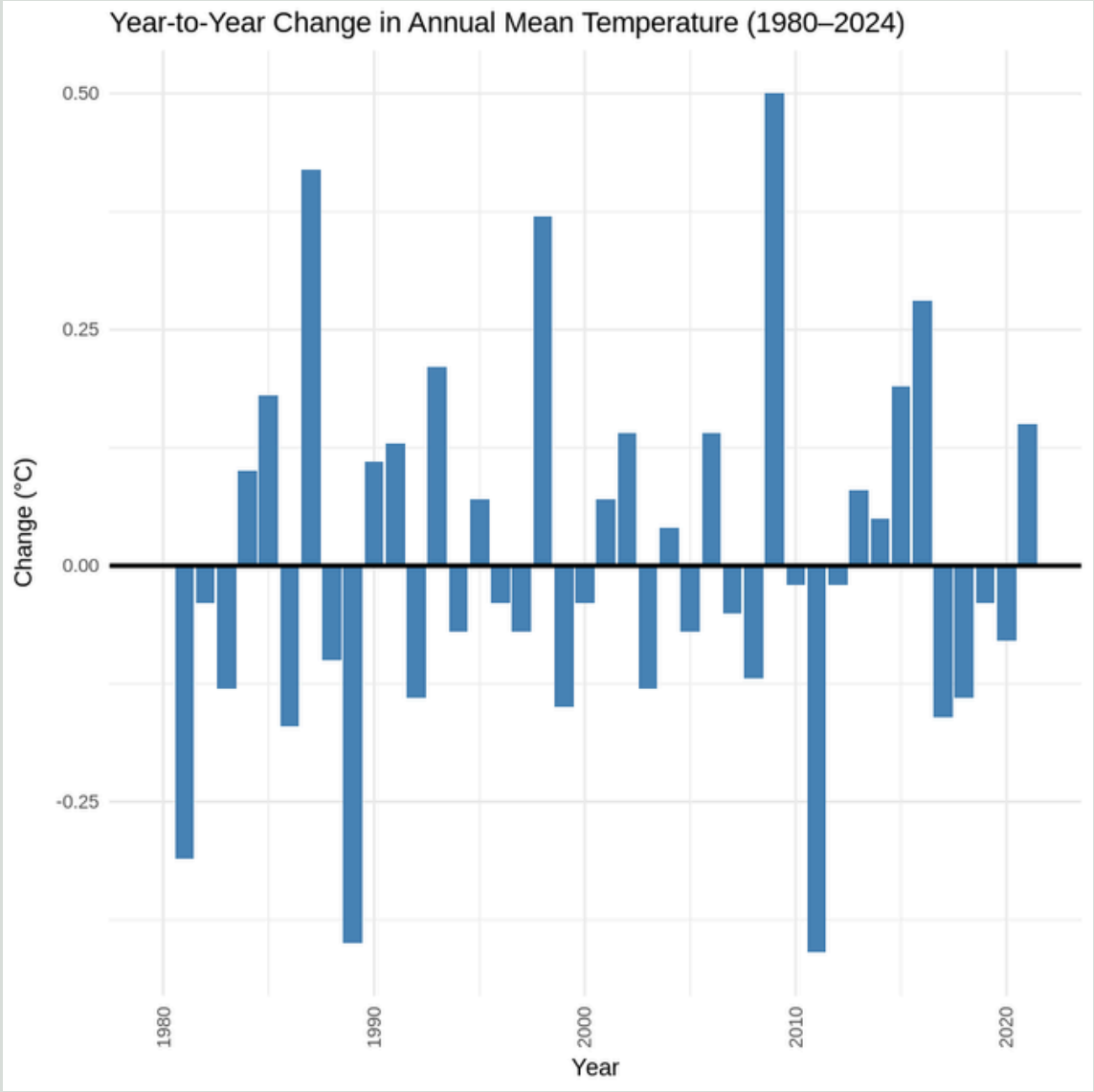
Distribution of Seasonal Temperatures

- Used histograms and density plots to visualize seasonal temperature spreads.
- Jun–Sep widest spread, prone to extremes
- Winters concentrated near median
- Insight: High variability seasons may correspond to extreme weather events.



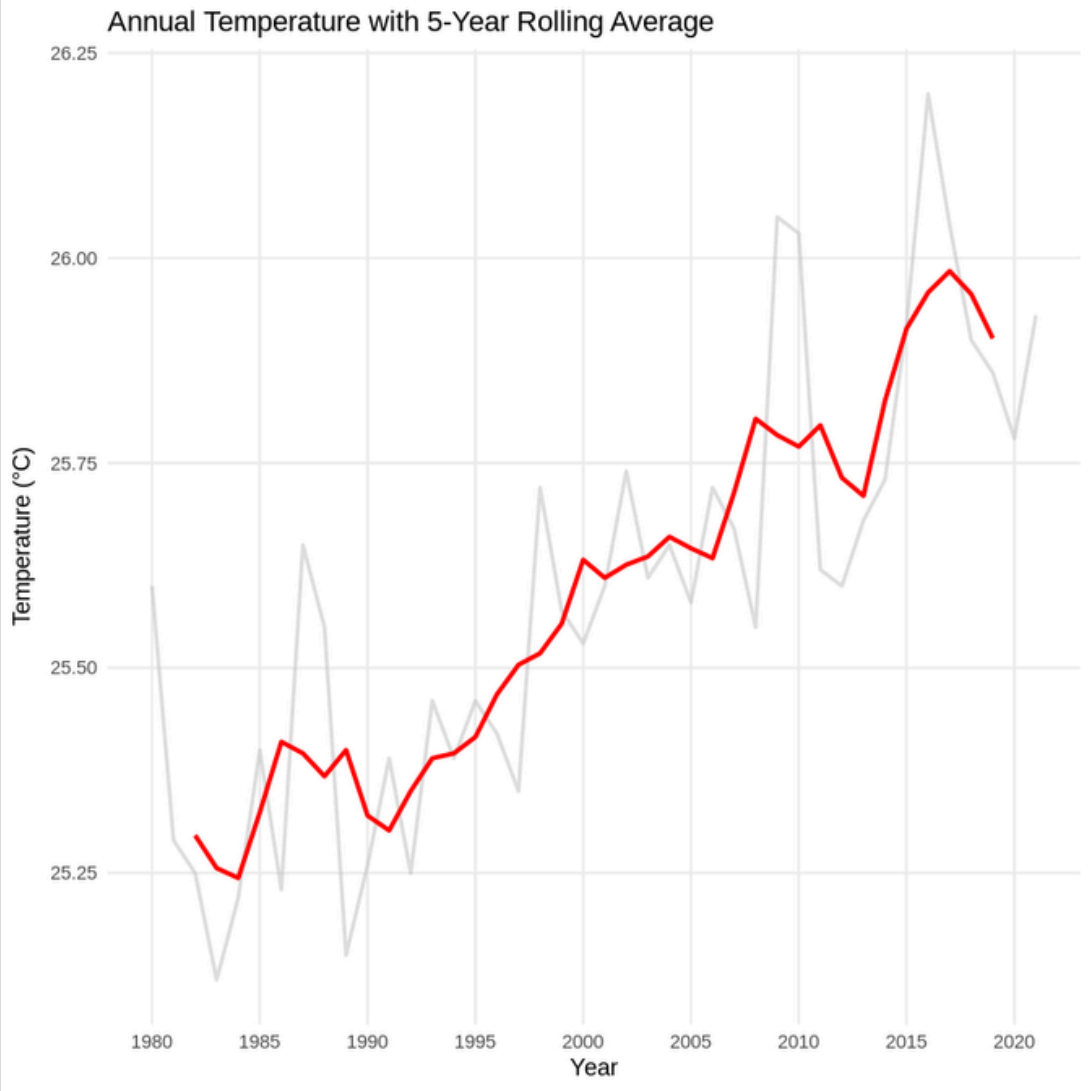
Year-to-Year Change in Annual Temperature

- Calculated the difference between consecutive years.
- Bar plot highlights years with unusually high or low changes.
- Insight: Some years experienced sharp increases, likely due to climate phenomena like El Niño.



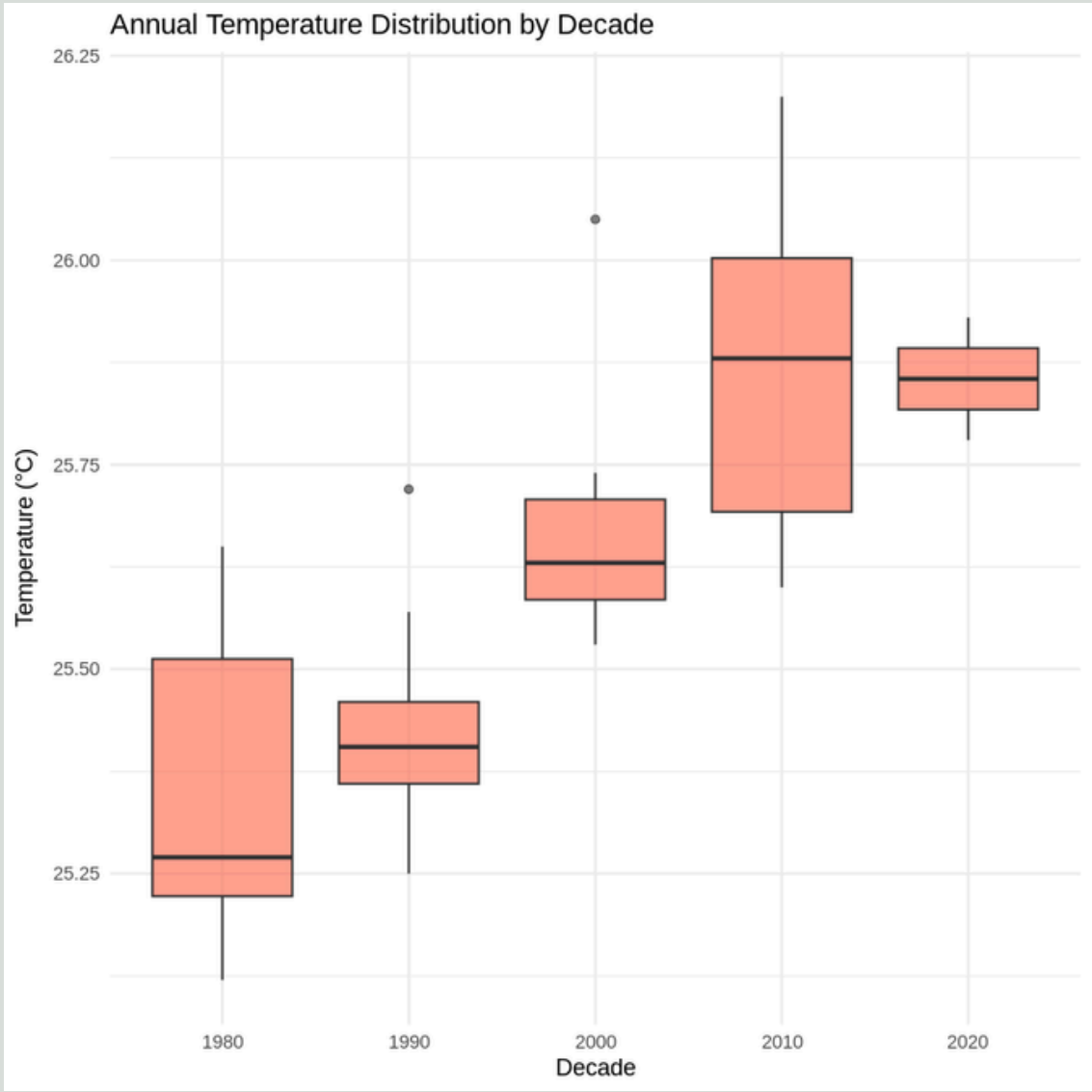
Rolling Average Trend

- Smoothed annual temperatures using a 5-year rolling mean.
- Helps reduce noise from natural variability and highlights the long-term warming trend.
- Insight: Rolling average confirms a consistent upward trend in temperatures over the past four decades



Annual Temperature Distribution by Decade

- Grouped years by decade and plotted boxplots.
- Insight: Recent decades (2010s–2020s) show higher median temperatures and slightly more spread, indicating warming and variability over time.



Hottest Years

YEAR	ANNUAL	JAN.FEB	MAR.MAY	JUN.SEP	OCT.DEC	Rolling_5yr	Decade
<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1983	25.12	20.09	26.93	28.12	22.68	25.256	1980
2016	26.20	21.80	28.64	28.35	23.84	25.958	2010

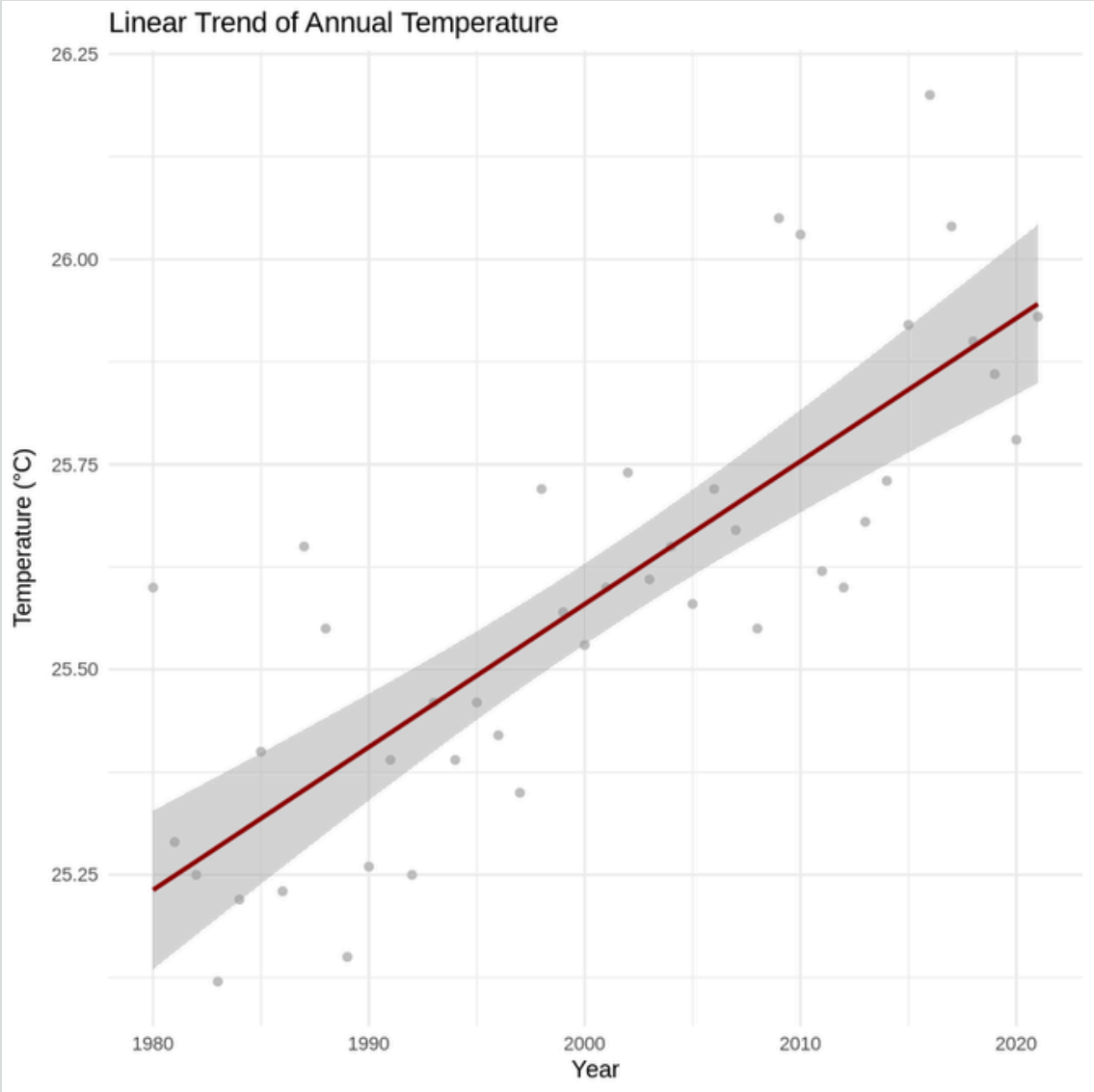
A data.frame: 5 × 9

	YEAR	ANNUAL	JAN.FEB	MAR.MAY	JUN.SEP	OCT.DEC	Rolling_5yr	Decade	Change
	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2009	26.05	21.59	28.12	28.55	23.61	25.784	2000	0.50
2	1987	25.65	20.73	27.29	28.55	23.40	25.396	1980	0.42
3	2011	25.62	20.52	27.67	28.06	23.71	25.796	2010	-0.41
4	1989	25.15	20.08	27.28	27.66	23.06	25.400	1980	-0.40
5	1998	25.72	20.83	27.82	28.31	23.43	25.518	1990	0.37

- Identified years with the highest annual mean temperatures.
- Insight: These years can be used as reference points for studying impacts on ecosystems, agriculture, and energy demand.

Linear Regression

- Fitted a linear model on annual temperatures to quantify the overall warming trend.
- Insight: The slope of the regression line confirms that temperature is increasing over time, reinforcing the findings from rolling averages and seasonal analysis.



Inference: Temperature Data

- Gradual warming trend since 1980
- Annual increase of ~0.017 °C per year
- Seasonal variations: Summer shows higher variability
- Forecasts indicate ~26 °C by 2030

Raw Dataset (Electricity)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Parameter	2015-04	2015-05	2015-06	2015-07	2015-08	2015-09	2015-10	2015-11	2015-12	2016-01	2016-02	2016-03	2016-04	2016-05	2016-06	2016-07	2016-08
2	Coal - Ger	69395.78	74322.05	68102.08	69686.27	69025.51	74409.06	81209.3	71269.75	75315.98	78146.02	75210.15	80507.25	83037.5	80984.36	76709.43	70535.53	67767.45
3	Oil & Gas	3246.28	3242.43	3796.41	3662.57	3854.07	4372.24	4156.93	4038.93	4263.18	3946.32	3943.46	4599.28	4173.18	4321.59	4069.68	4027.92	4439.72
4	Nuclear - (	3506.1	3457.01	3449.28	3022.21	2834.19	2806.7	2876.22	2868.49	2950.27	3056.56	2998.87	3587.72	3285.31	2948.61	2866.87	3258.16	3549.8
5	Hydro - Ge	9421.19	12854.8	12509.46	15094.48	17779.11	12299.09	9025.41	6821.86	6339.87	6159.08	6087.2	6985.17	7812.97	10722.43	12295.89	15082.8	17667.61
6	Solar - Ger	512.31	560.91	515.11	455.52	510.73	600.51	726.49	578.11	651.61	535.11	743.56	954.35	1027.64	1025.41	911.02	851.91	929.36
7	Wind - Ge	1832.3	2673.76	3946.56	6591.88	5113.45	3221.01	2023.15	1530.58	1498.03	1056.32	1649.17	1893.26	2231.34	4132.08	5854.24	6977.52	6818.77
8	Small-Hyd	519.15	685.93	775.37	1069.77	1166.68	943.72	747.36	604.45	528.58	449.36	372.61	496.8	526.52	657.7	644.99	1118.24	1286.59
9	Bio Power	1688.88	1222.24	853.03	742.6	726.06	707.68	844.47	1416.99	2287.2	2325.4	2217.97	1911.34	1303.38	1025.84	696.03	599.29	598.56
10	Total	90121.99	99019.13	93947.3	100325.3	101009.8	99360.01	101609.3	89129.16	93834.72	95674.17	93222.99	100935.2	103397.8	105818	104048.2	102451.4	103057.9
11	Growth Ra																	
12																		

Data Cleaning (Electricity)

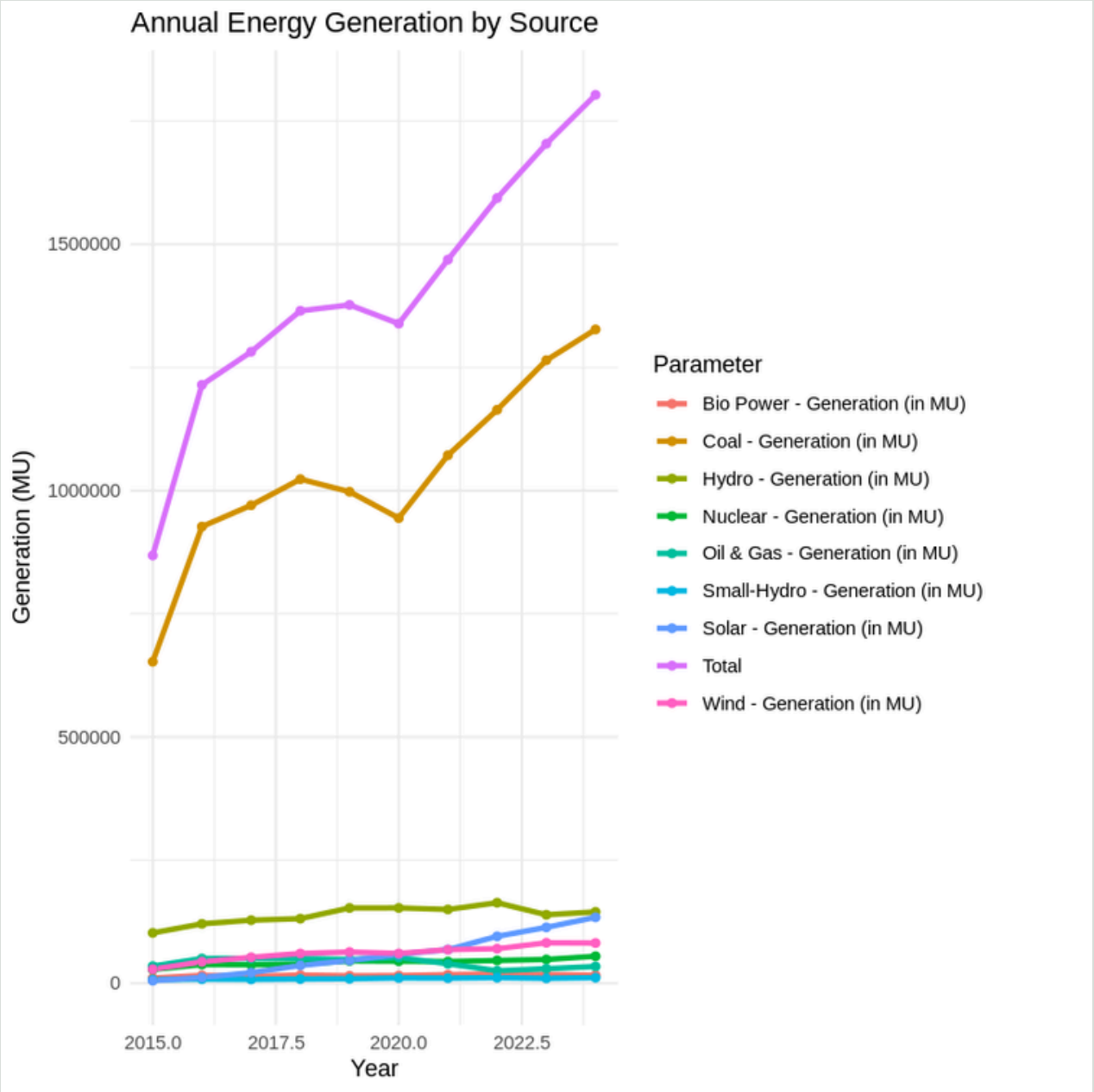
- Converted wide → long format → tidy structure.
- Removed commas → numeric values.
- Parsed dates & extracted year → proper time handling.
- Aggregated yearly by source → yearly totals for each energy source.

Cleaned Dataset (Electricity)

Year	Bio Power - Generation (in MU)	Coal - Generation (in MU)	Hydro - Generation (in MU)	Nuclear - Generation (in MU)	Oil & Gas - Generation (in MU)	Small-Hydro - Generation (in MU)	Solar - Generation (in MU)	Total	Wind - Generation (in MU)
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
2015	10489.15	652735.8	102145.3	27770.47	34633.04	7041.01	5111.30	868356.7	28430.72
2016	15871.26	926913.0	120510.2	37897.56	50211.70	8278.84	11565.31	1214708.2	43460.23
2017	14025.43	970293.4	127928.6	37430.36	50256.04	7792.49	21475.42	1281830.8	52629.01
2018	16573.71	1023220.0	131020.1	39050.71	50257.34	8509.89	36037.90	1364980.9	60311.28
2019	15125.12	997670.4	152751.9	45162.88	47980.39	8985.33	46003.28	1376993.6	63314.31
2020	15514.09	944209.4	152835.8	44613.11	52170.75	10529.11	58683.59	1338983.3	60427.48

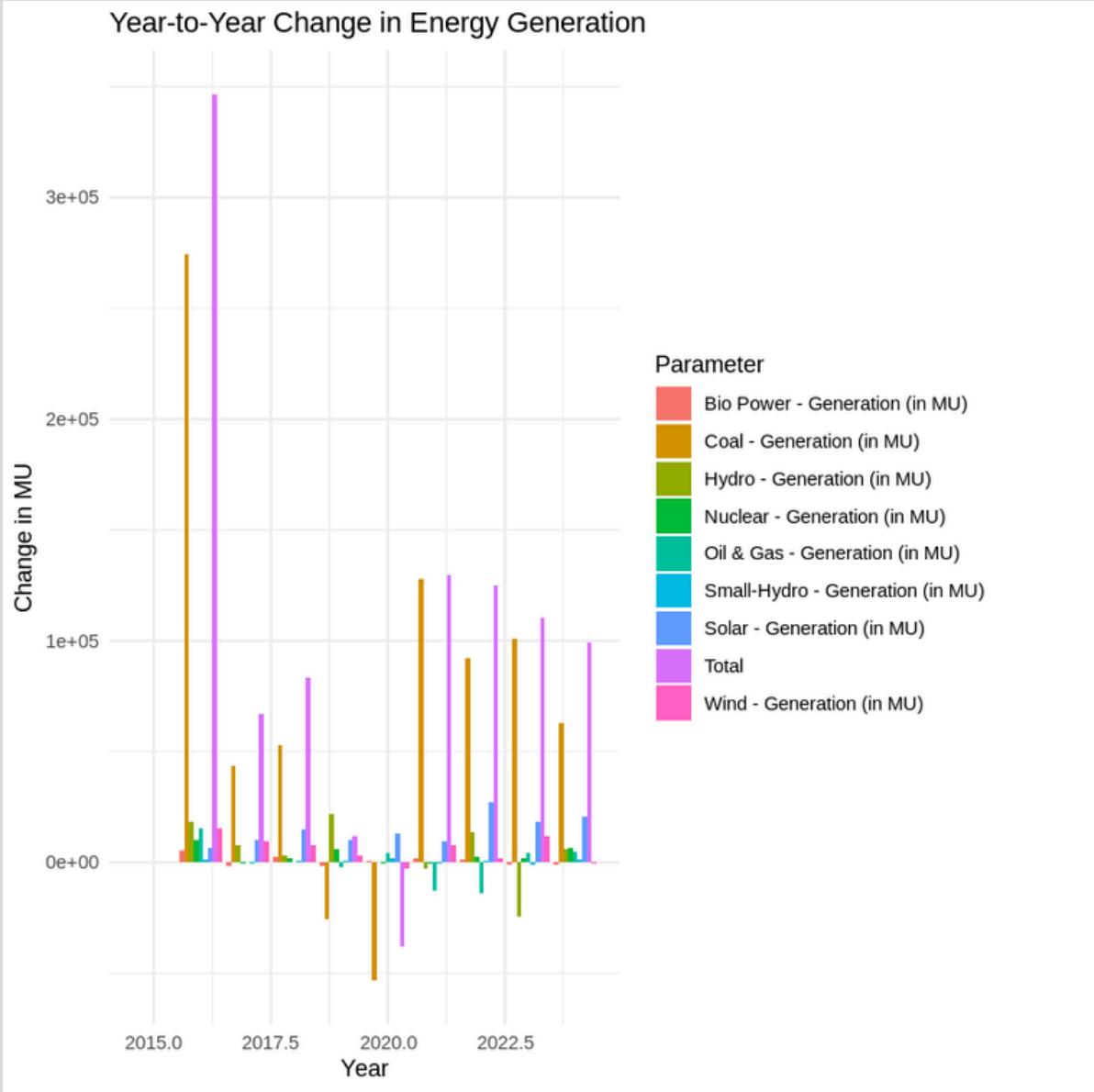
Annual Energy Generation by Source

- Total generation shows strong, continuous upward growth.
- Coal is the volume leader, but Wind is showing the steepest growth and is catching up fast.
- Hydro and Nuclear generation is flat and contributes only a small fraction of the total.



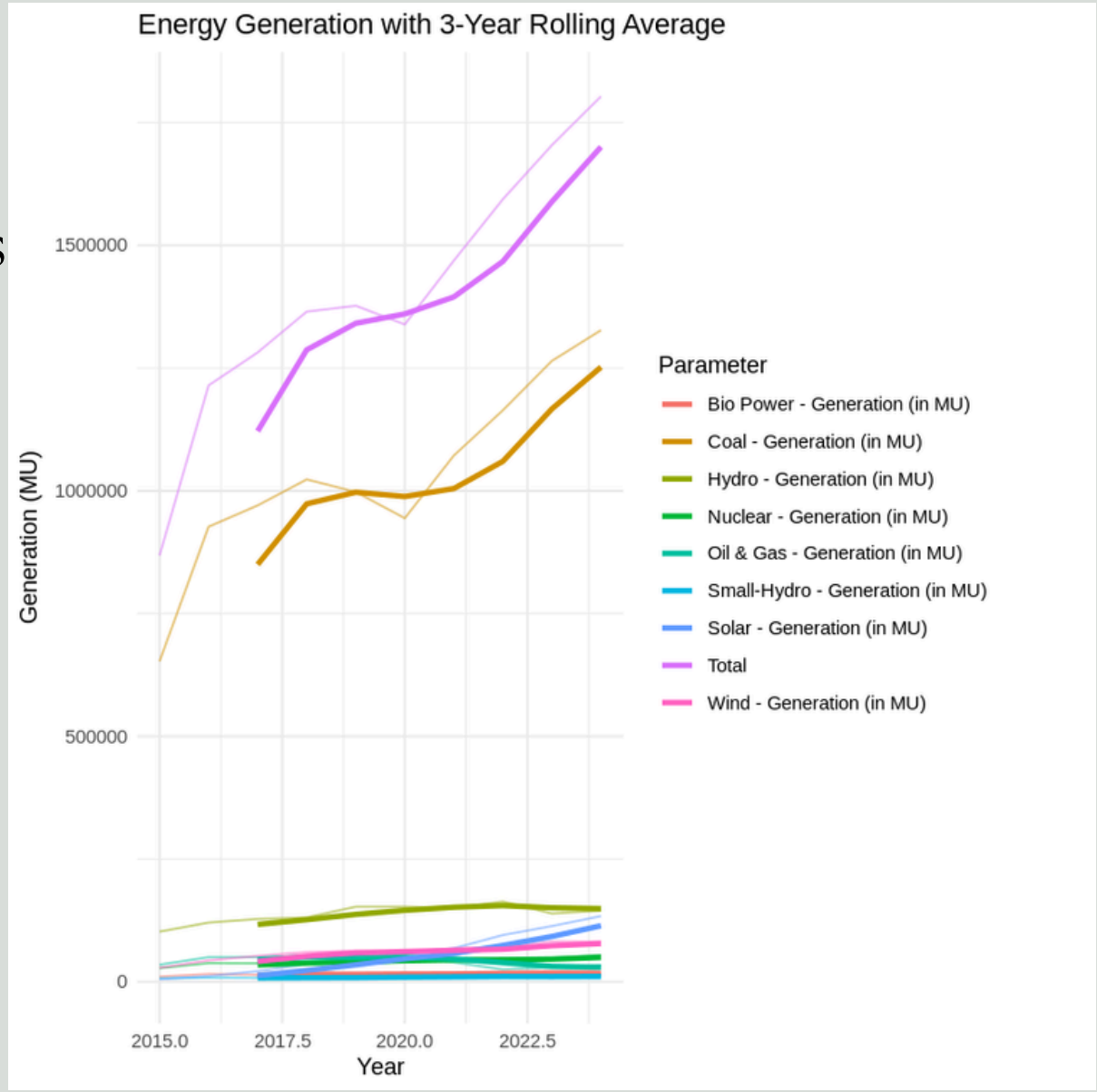
Year-to-Year Change in Energy Generation

- Total change is the most significant yearly change.
- Coal's massive surges and drops dominate individual source volatility.
- Wind generation consistently drives positive growth
- The largest Total change occurred around 2016.0.
- Hydro, Nuclear, and Solar changes are minimal and stable.



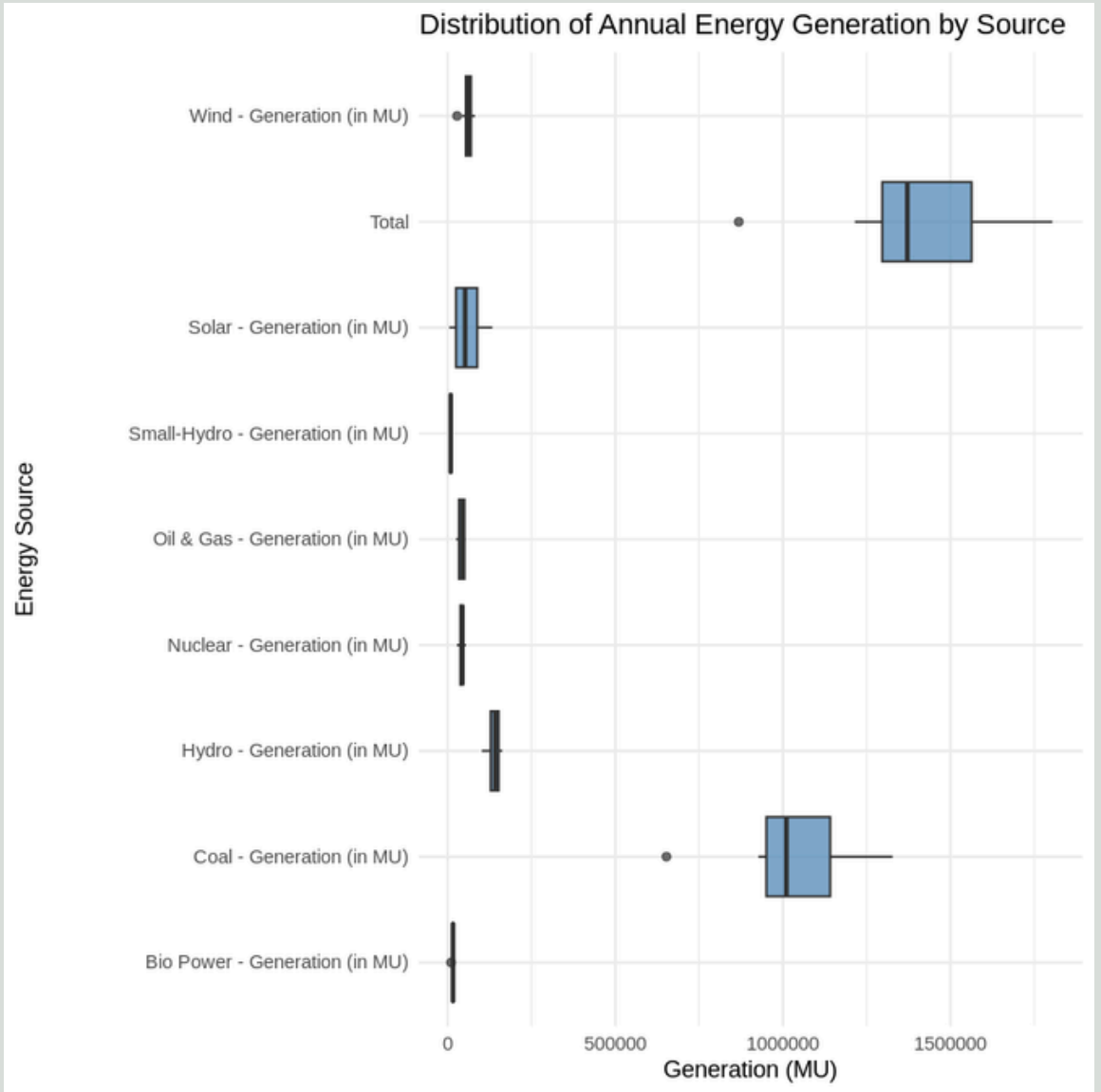
3-Year Rolling Average

- Total generation is consistently rising.
- Wind generation is now the largest source and is growing fast.
- Coal remains a major, growing power source.
- Most other sources (Hydro, Solar, Nuclear, etc.) are minor contributors.
- Hydro and Nuclear generation levels are relatively stable (flat).
-



Distribution of Annual Energy Generation by Source

- Total generation is the largest in scale and range.
- Coal is the single dominant energy source.
- All other sources are minimal in volume and variability.



Annual Energy Generation Top Years

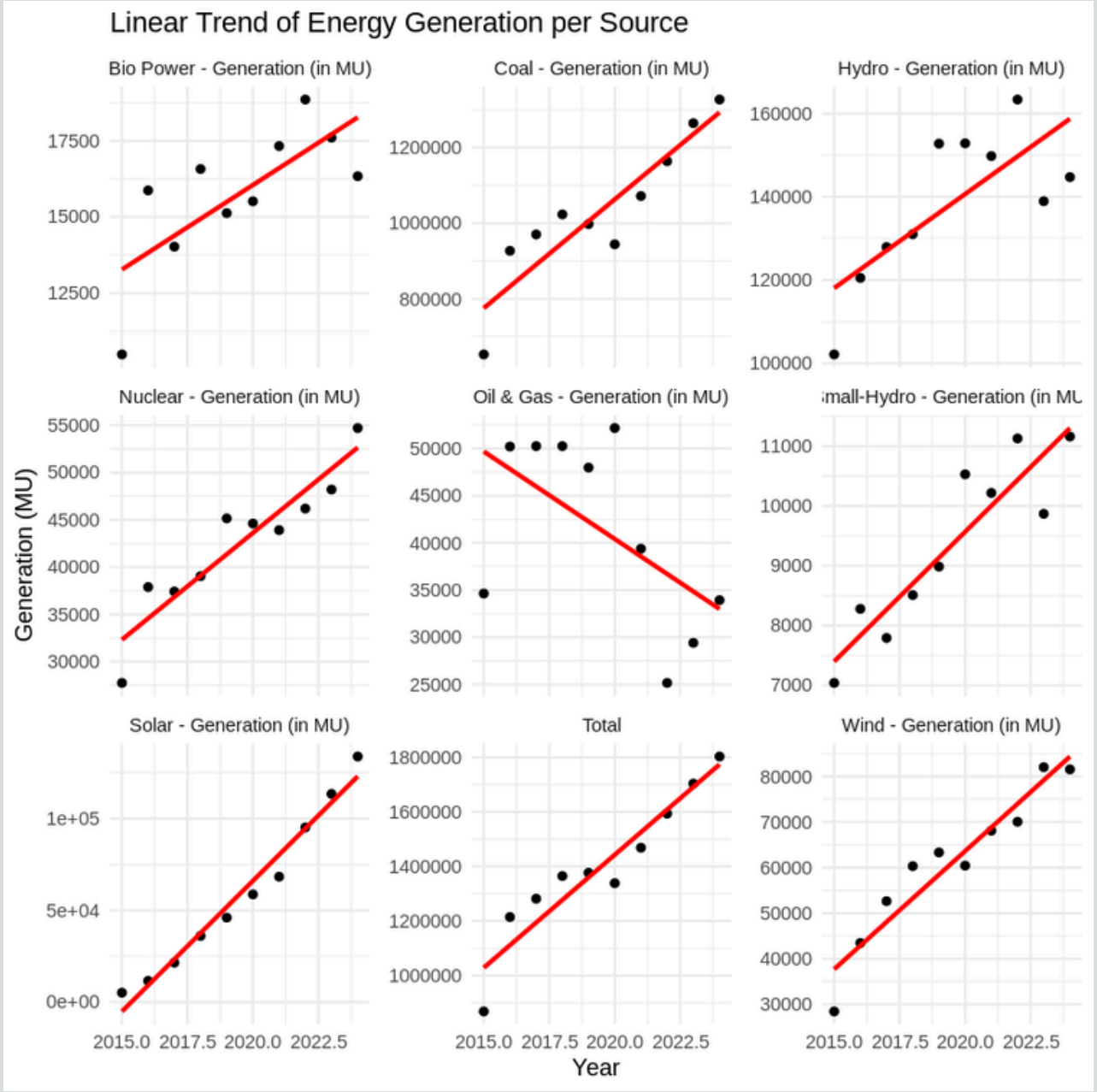
- Total generation peaks in the latest year shown (2024), confirming continuous growth.
- Coal is the largest overall contributor, consistently producing over 1.1 Million MU.
- Solar shows the most aggressive recent growth.
- Hydro generation remains stable..
- Wind generation hit its peak in 2023, with a slight dip in 2024.

A grouped_df: 27 × 5

Parameter	Year	Generation	YearlyChange	Rolling3yr
<chr>	<dbl>	<dbl>	<dbl>	<dbl>
Bio Power - Generation (in MU)	2022	18852.72	1525.74	17231.263
Bio Power - Generation (in MU)	2023	17603.55	-1249.17	17927.750
Bio Power - Generation (in MU)	2021	17326.98	1812.89	15988.730
Coal - Generation (in MU)	2024	1327203.42	62516.60	1251982.247
Coal - Generation (in MU)	2023	1264686.82	100630.32	1166870.593
Coal - Generation (in MU)	2022	1164056.50	92188.04	1060044.777
Hydro - Generation (in MU)	2022	163344.77	13577.26	155316.013
Hydro - Generation (in MU)	2020	152835.76	83.84	145535.923
Hydro - Generation (in MU)	2019	152751.92	21731.83	137233.530
Nuclear - Generation (in MU)	2024	54704.06	6499.63	49701.057
Nuclear - Generation (in MU)	2023	48204.43	2009.75	46105.680
Nuclear - Generation (in MU)	2022	46194.68	2276.75	44908.573
Oil & Gas - Generation (in MU)	2020	52170.75	4190.36	50136.160
Oil & Gas - Generation (in MU)	2018	50257.34	1.30	50241.693
Oil & Gas - Generation (in MU)	2017	50256.04	44.34	45033.593

Linear Trends in Annual Energy Generation

- Total generation shows a strong, consistent positive linear trend.
- The majority of sources are increasing, including Coal, Hydro, Solar, and Wind.
- Oil & Gas is the only source with a clear negative linear trend (declining generation).



Inference: Electricity Data

- Total electricity generation is steadily increasing
- Coal remains the largest source
- Wind and Solar show fastest growth in recent years
- Hydro and Nuclear remain relatively stable

Relationship between temperature and Energy consumption

- Objective: Explore the relationship between temperature and electricity generation.
- Data Sources:
  1. Annual temperature data (temp_cleaned.csv)
  2. Annual electricity generation data (annual_energy.csv)
- Methods: Data cleaning, correlation analysis, regression modeling, forecasting

Data Cleaning

- Removed empty and "Total" rows in energy data.
- Reshaped energy data from wide → long format.
- Converted Year and numeric columns to numeric type.
- Merged temperature and energy datasets on Year.
- Resulting dataset: temp_energy_combined.csv

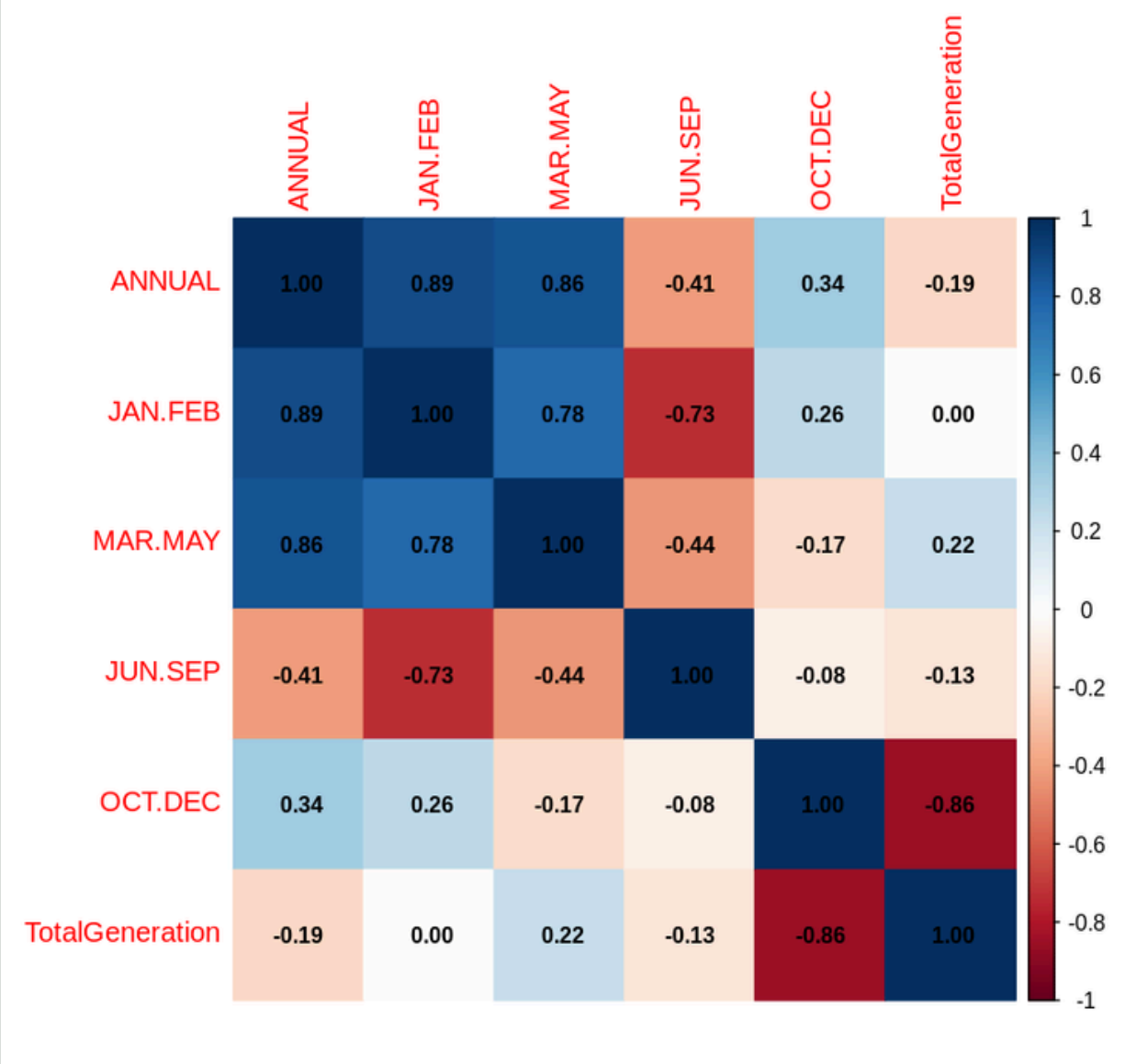
Data Overview

	Year	YEAR	ANNUAL	JAN.FEB	MAR.MAY	JUN.SEP	OCT.DEC	TotalGeneration
	<dbl>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2015	2015	25.92	21.05	27.63	28.46	24.08	868356.7
2	2016	2016	26.20	21.80	28.64	28.35	23.84	1214708.2
3	2017	2017	26.04	21.37	28.20	28.39	23.84	1281830.8
4	2018	2018	25.90	21.24	28.16	28.28	23.55	1364980.9
5	2019	2019	25.86	20.71	28.00	28.60	23.49	1376993.6
6	2020	2020	25.78	20.79	27.58	28.45	23.75	1338983.3

- Covers 1980–2024 (temperature and electricity generation)
- ~40+ years of climate & energy records
- Variables: Year, Annual/Seasonal Temperature, Energy by Source
- Used for trend analysis, correlation & forecasting

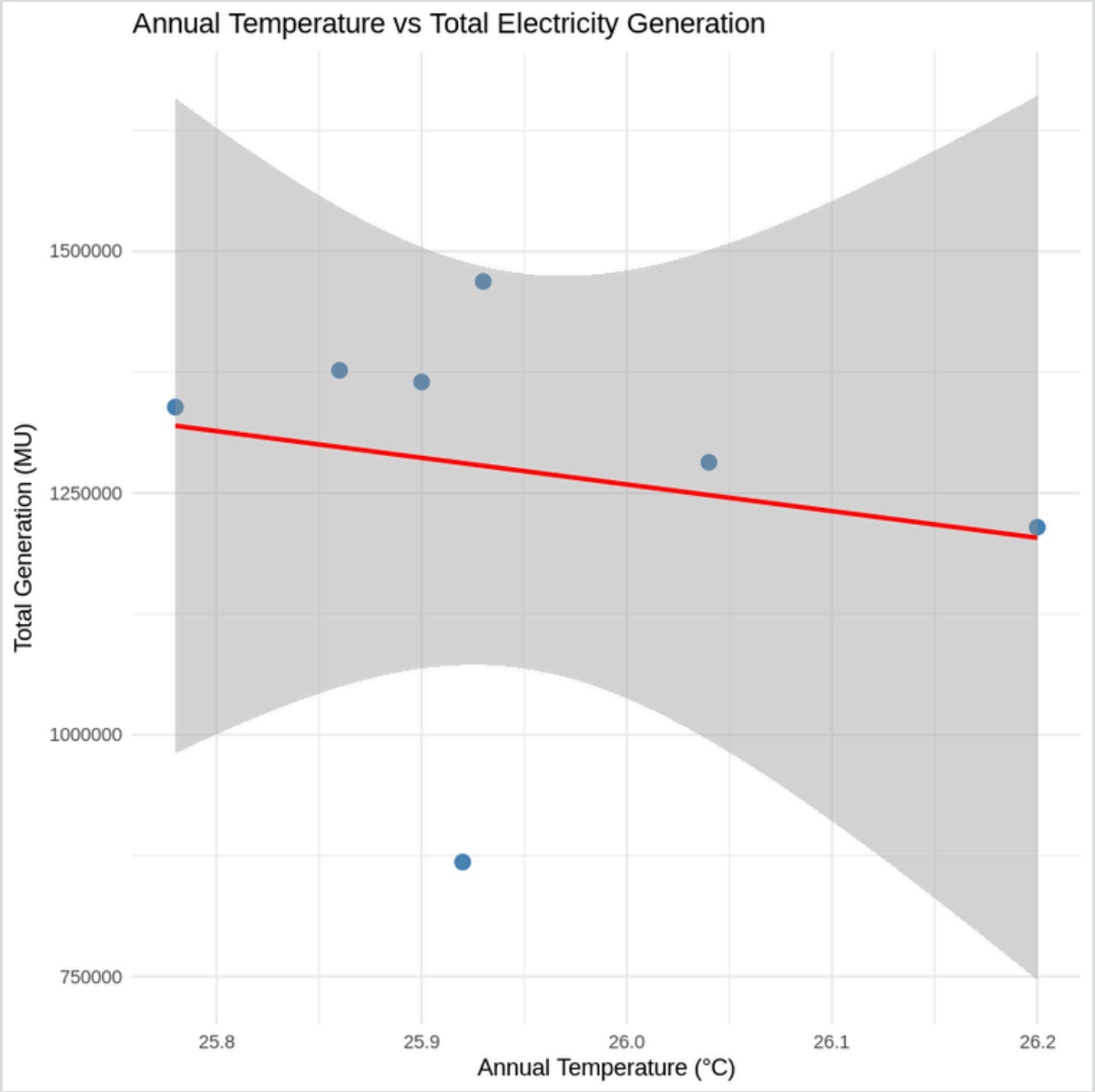
Correlation Analysis

- Correlation Analysis of Annual, Seasonal, and Total Generation values.
- Seasonal generation patterns vary significantly.
- Winter and spring (Jan–May) align closely with annual trends.
- Autumn (Oct–Dec) shows an inverse relationship with total generation.
- Summer (Jun–Sep) tends to contrast with early-year months.



Scatter plot of Annual Temperature vs Total Generation.

- Negative trend: as temperature rises, total generation tends to fall.
- The red regression line shows a downward slope, but correlation is weak.
- Wide confidence band indicates high uncertainty in the relationship.



Prediction of Temperature

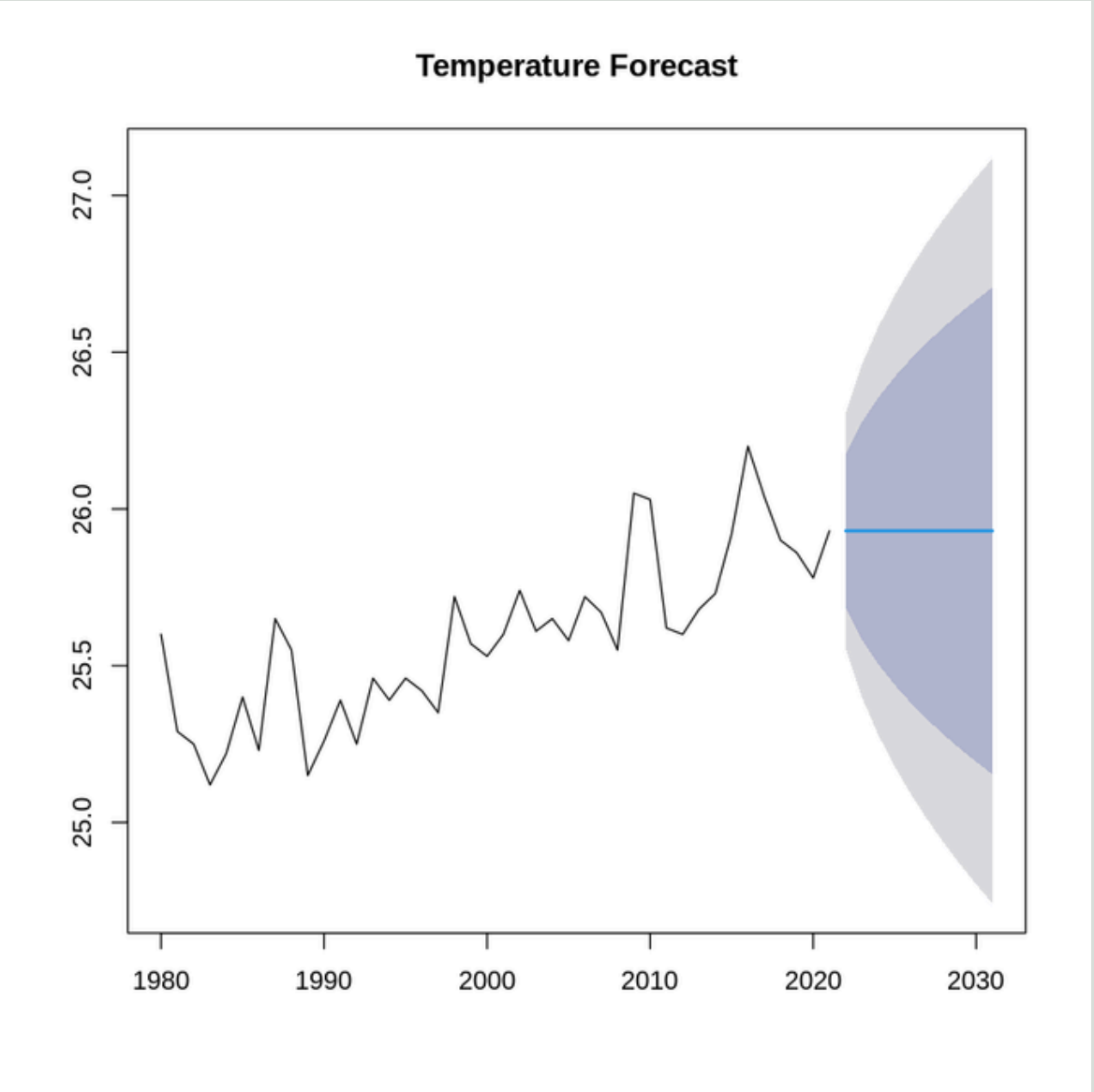
- Model: Annual temperature regressed on Year.
- Significant upward trend: coefficient of Year = +0.0174 °C/year ($p < 0.001$).
- Good fit: $R^2 = 0.65$, meaning ~65% of variation explained by time.
- Prediction: Temperature expected to rise steadily from 26.0°C in 2025 to 26.1°C in 2030.
- Interpretation: Strong evidence of warming trend over years.

```
Call:
lm(formula = ANNUAL ~ Year, data = temp_cleaned)

Residuals:
     Min       1Q   Median       3Q      Max
-0.23833 -0.10298 -0.02692  0.06946  0.36837

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -9.243893   4.011803  -2.304   0.0265 *
Year         0.017412   0.002005   8.683 9.62e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1575 on 40 degrees of freedom
Multiple R-squared:  0.6533,Adjusted R-squared:  0.6447
F-statistic: 75.39 on 1 and 40 DF,  p-value: 9.621e-11
  Year PredictedTemp
1 2025      26.01516
2 2026      26.03257
3 2027      26.04999
4 2028      26.06740
5 2029      26.08481
6 2030      26.10222
```

Annual temperature shows a significant warming trend, rising by ~0.017 °C per year.

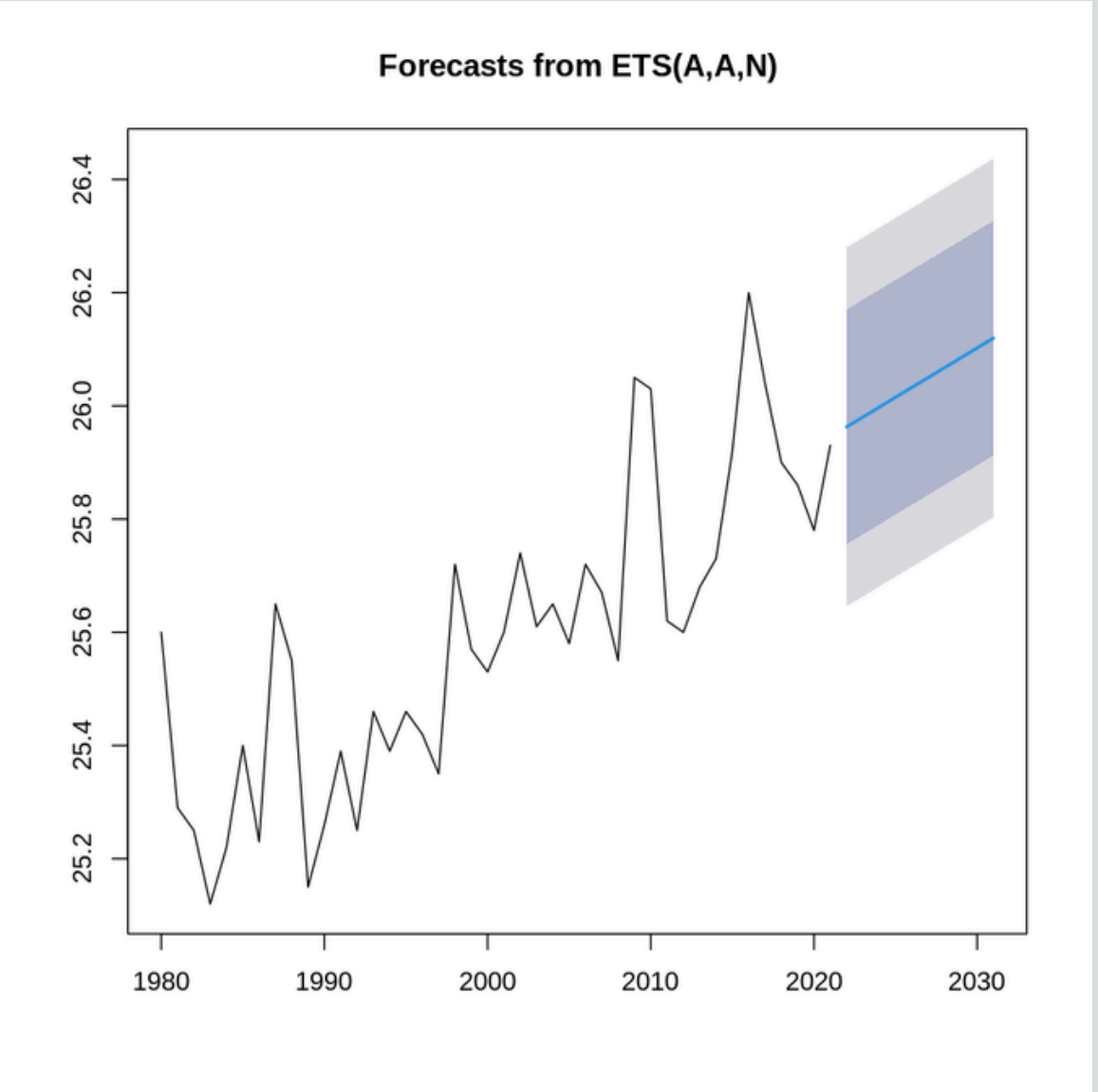
ARIMA Forecast

- Best-fit ARIMA model selected by auto.arima()
- Gradual warming observed since 1980
- Projected rise to ~26 °C by 2030
- Confidence intervals widen into the future
- Continued warming trend expected



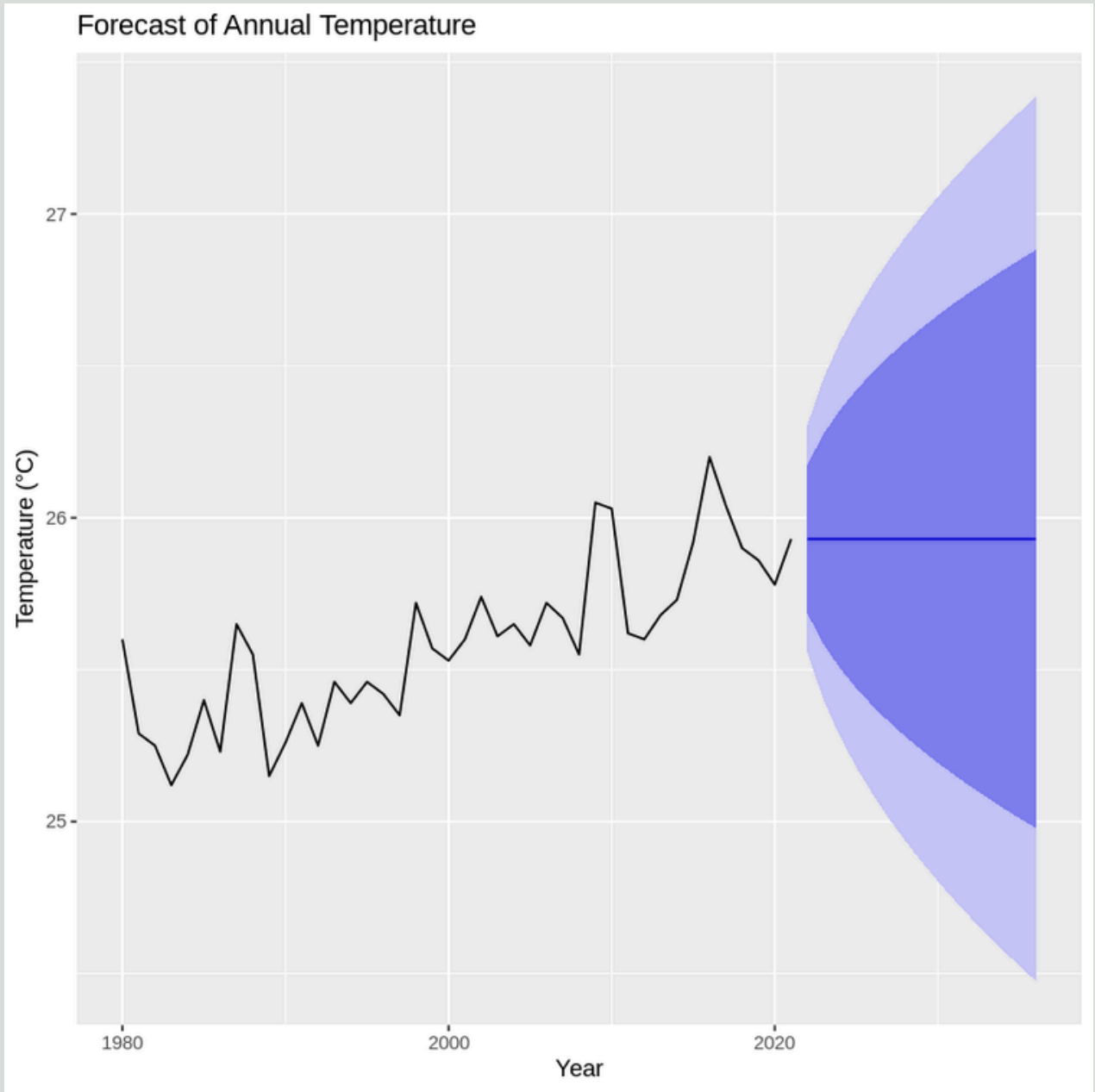
ETS Forecast

- Exponential smoothing model (ETS)
- Forecast trend similar to ARIMA
- Confirms steady warming projection
- Adds robustness to prediction results



Forecast Comparison

- Regression shows a steady linear rise
- ARIMA is flexible, captures time-series behavior
- ETS is consistent, supports warming outlook
- All models predict warming toward ~26 °C by 2030



Linking Climate & Energy

- Rising temperatures influence electricity demand & supply
- Seasonal variations impact generation patterns
- Weak negative link: higher temps → lower total generation
- Climate change adds uncertainty to future energy planning

Conclusion

- Long-term analysis confirms a significant warming trend since 1980, at ~0.017 °C/year.
- Forecasting methods (Regression, ARIMA, ETS) all project a rise toward 26 °C by 2030.
- Electricity generation is growing strongly, with Coal dominant and Wind rising fastest.
- Linking climate and energy reveals both risks (reduced generation, rising demand) and opportunities (renewable expansion, smarter planning).
- Overall, evidence highlights the need for integrated climate–energy strategies to address future challenges.

References

Code & Data: GitHub Repository