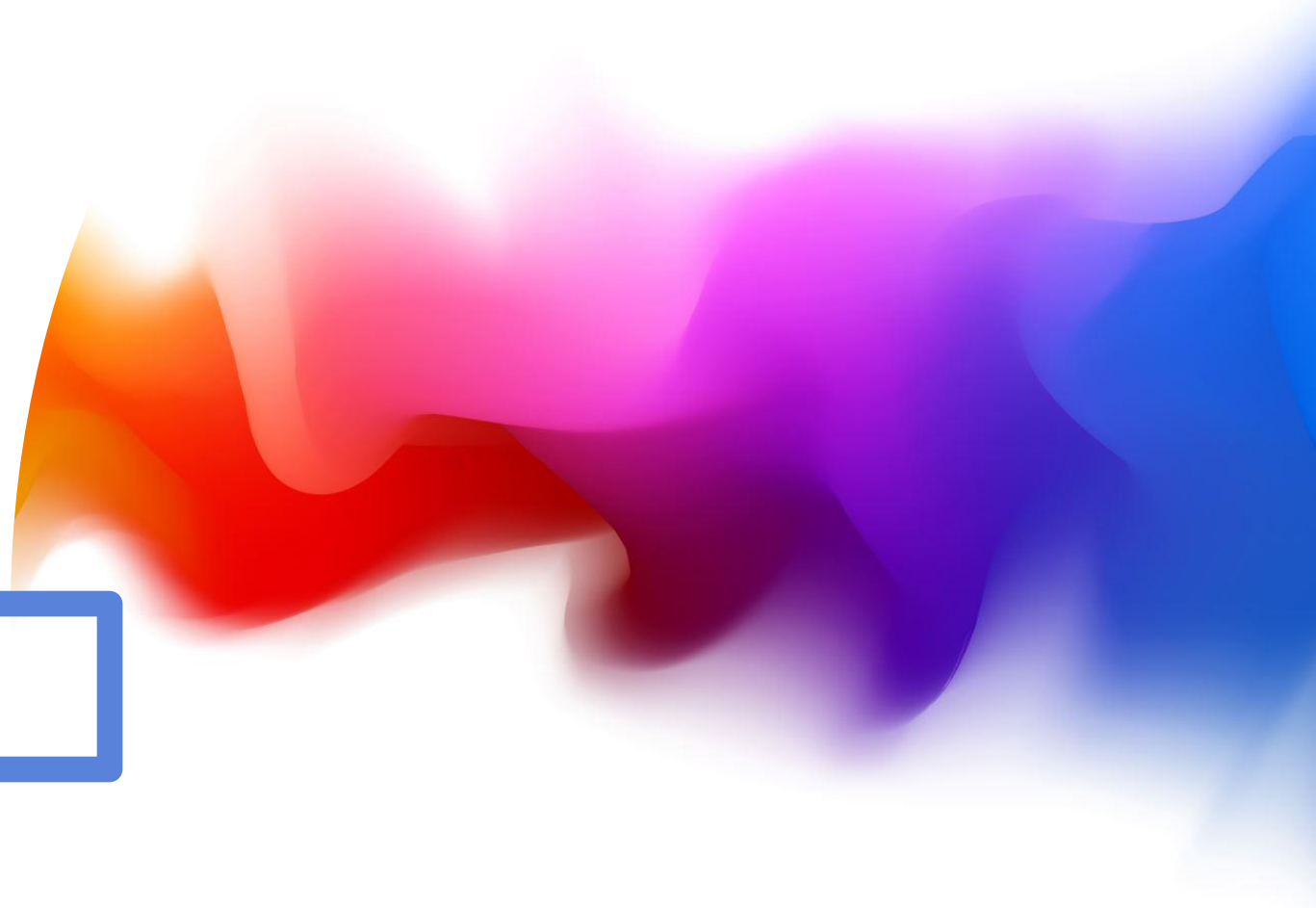
Global Power Analysis: Energy, Economics, and Emissions

Mid-Term Presentation

Kevin Shao – Mar 11, 2025



Project Overview & Research Questions

- Overview:
  - Integrate global power plant data with socioeconomic, emissions, and policy indicators.
- Research Questions:
  - How does a country's renewable power plant capacity share correlate with its CO_2 emissions (or carbon intensity) over time?
  - Do countries with stronger green growth policies show lower fossil-fuel capacity and higher renewables share?

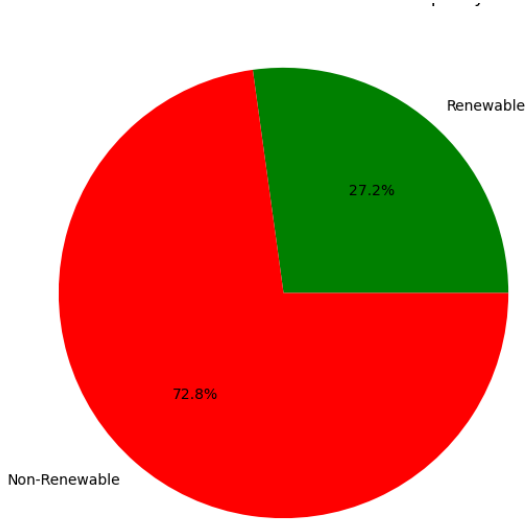
Data Sources & Integration

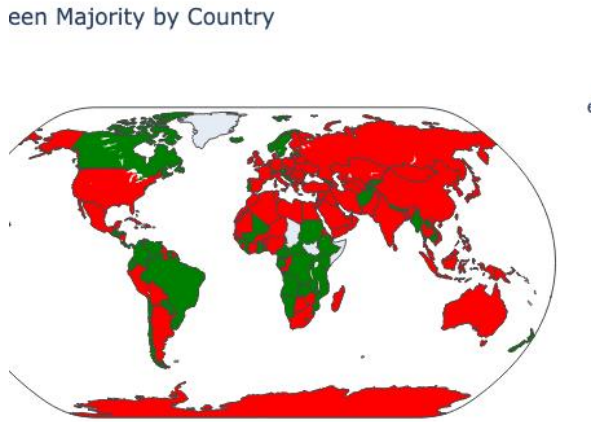
- Data Sources:
  - Global Power Plant Database
  - World Bank WDI (socioeconomic & emissions data)
  - IEA World Energy Balances
  - OECD Green Growth Indicators (policy)
- Integration Process:
  - Standardize country codes/names
  - Aggregate plant-level data by fuel type
  - Merge with country-level indicators

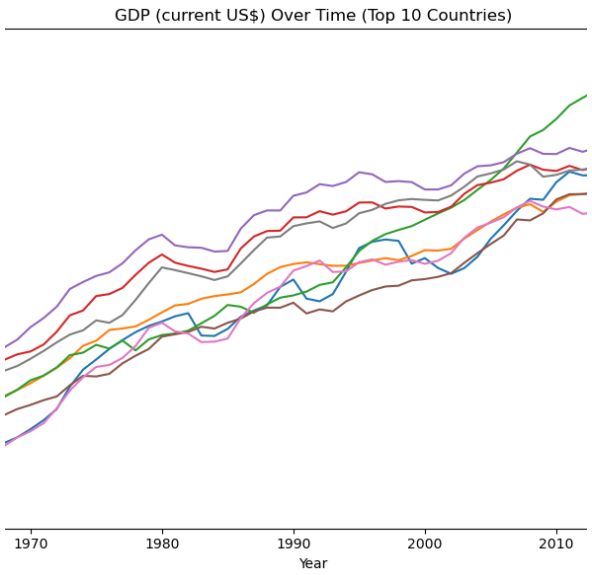
Table	Rows	Columns
world_bank	37,511	67
iea_balances	6,832	59
oecd_greengrowth	677,449	6
power_plants	34,936	13
country_capacity	167	3

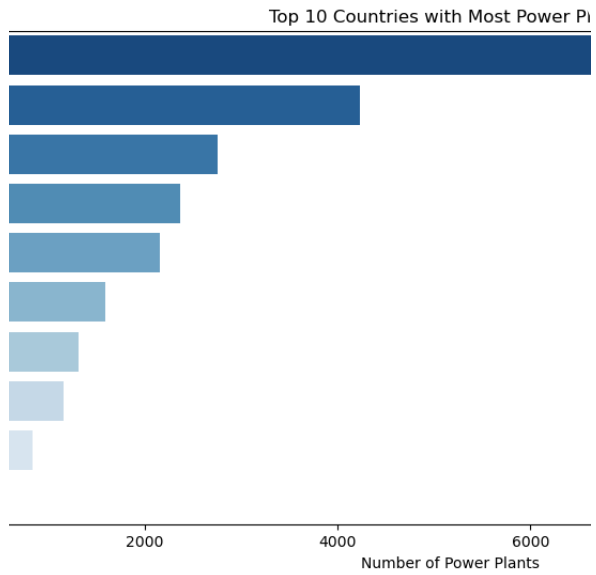
EDA

- Visualizations:
  - Bar charts of power plant counts by country
  - Capacity breakdown by fuel type
  - Geospatial mapping of power plant distribution
  - Time series of key World Bank indicators



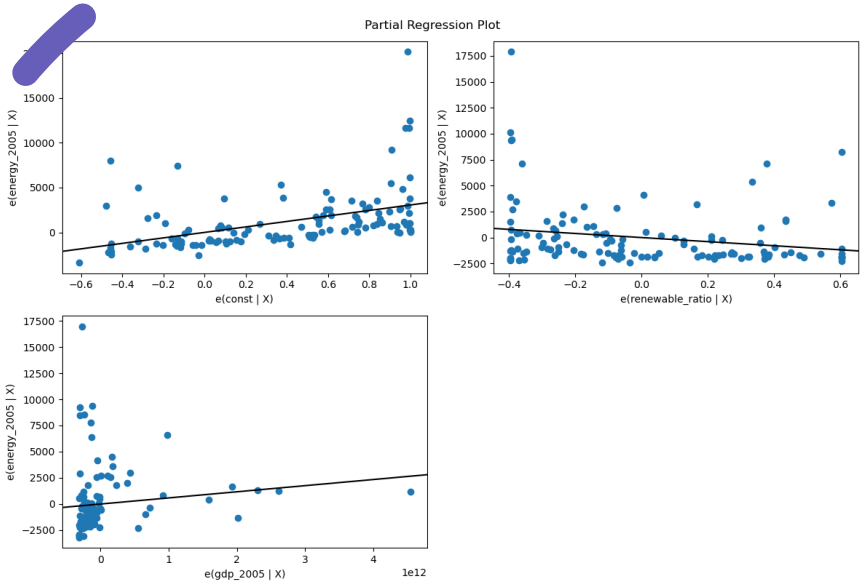






Simple Analysis & Findings

- Key Analysis:
  - Correlation analyses (e.g., renewable capacity ratio vs. energy use)
  - Basic regression results (simple OLS models)
- Findings:
  - Preliminary relationships are weak (low R^2 values)
  - Indications that fossil fuel share might influence energy use



Number of rows with all valid data: 125

OLS Regression Results

Dep. Variable:	energy_2005	R-squared:	0.063
Model:	OLS	Adj. R-squared:	0.048
Method:	Least Squares	F-statistic:	4.108
Date:	Sun, 23 Feb 2025	Prob (F-statistic):	0.0188
Time:	14:20:24	Log-Likelihood:	-1175.0
No. Observations:	125	AIC:	2356.
Df Residuals:	122	BIC:	2365.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	3056.8785	435.309	7.022	0.000	2195.141	3918.616
renewable_ratio	-1965.4141	834.444	-2.355	0.020	-3617.280	-313.548
gdp_2005	5.849e-10	4.14e-10	1.415	0.160	-2.34e-10	1.4e-09

Omnibus:	99.074	Durbin-Watson:	2.095
Prob(Omnibus):	0.000	Jarque-Bera (JB):	715.011
Skew:	2.832	Prob(JB):	5.46e-156
Kurtosis:	13.256	Cond. No.	2.35e+12

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 2.35e+12. This might indicate that there are strong multicollinearity or other numerical problems.

Advanced Panel Data Analysis

- Methodology:
  - Two-way Fixed Effects (FE) and Random Effects (RE) models
  - Hausman test to compare FE and RE specifications
- Key Results:
  - FE model (using varying variables like GDP and fossil_share)
  - RE model – note the challenges with time-invariant variables like renewable_ratio
  - Hausman test results (showing unobserved, time invariant factors matters)

=== Hausman Test Comparison ===

Model Comparison

	FE	RE
Dep. Variable	energy_use	energy_use
Estimator	PanelOLS	RandomEffects
No. Observations	4523	4523
Cov. Est.	Clustered	Unadjusted
R-squared	0.0057	0.0197
R-Squared (Within)	-0.0026	0.0081
R-Squared (Between)	-0.3263	0.2143
R-Squared (Overall)	-0.3100	0.2073
F-statistic	12.554	30.212
P-value (F-stat)	0.0000	0.0000
gdp	-3.187e-11	2.87e-10
	(-0.2695)	(8.5899)
fossil_share	-9.0111	6.8144
	(-0.9054)	(4.0822)
renewable_ratio		0.0080
		(1.873e-05)
Effects	Entity	
	Time	

RandomEffects Estimation Summary

Dep. Variable:	energy_use	R-squared:	0.0197
Estimator:	RandomEffects	R-squared (Between):	0.2143
No. Observations:	4523	R-squared (Within):	0.0081
Date:	Mon, Mar 10 2025	R-squared (Overall):	0.2073
Time:	23:29:45	Log-likelihood	-3.728e+04
Cov. Estimator:	Unadjusted		
		F-statistic:	30.212
Entities:	125	P-value	0.0000
Avg Obs:	36.184	Distribution:	F(3,4520)
Min Obs:	4.0000		
Max Obs:	46.000	F-statistic (robust):	30.212
		P-value	0.0000
Time periods:	46	Distribution:	F(3,4520)
Avg Obs:	98.326		
Min Obs:	22.000		
Max Obs:	124.00		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
gdp	2.87e-10	3.341e-11	8.5899	0.0000	2.215e-10	3.525e-10
fossil_share	6.8144	1.6693	4.0822	0.0000	3.5418	10.087
renewable_ratio	0.0080	427.48	1.873e-05	1.0000	-838.06	838.07

PanelOLS Estimation Summary

Dep. Variable:	energy_use	R-squared:	0.0057
Estimator:	PanelOLS	R-squared (Between):	-0.3263
No. Observations:	4523	R-squared (Within):	-0.0026
Date:	Mon, Mar 10 2025	R-squared (Overall):	-0.3100
Time:	23:29:45	Log-likelihood	-3.686e+04
Cov. Estimator:	Clustered		
		F-statistic:	12.554
Entities:	125	P-value	0.0000
Avg Obs:	36.184	Distribution:	F(2,4351)
Min Obs:	4.0000		
Max Obs:	46.000	F-statistic (robust):	0.4455
		P-value	0.6405
Time periods:	46	Distribution:	F(2,4351)
Avg Obs:	98.326		
Min Obs:	22.000		
Max Obs:	124.00		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
gdp	-3.187e-11	1.182e-10	-0.2695	0.7875	-2.636e-10	1.999e-10
fossil_share	-9.0111	9.9521	-0.9054	0.3653	-28.522	10.500

Key Insights & Policy Implications

- Insights:
  - Economic size (GDP) and fossil fuel share significantly relate to energy use.
  - Policy indicator (environment-related tax) shows promise for further exploration.
- Implications:
  - Findings can inform policymakers on how economic and energy mix factors affect emissions.
  - There is potential to identify best practices for reducing fossil fuel dependency.

Next Steps

- Next Steps in Analysis:
  - Refine panel regression models (including CO_2 emissions and other policy indicators).
  - Conduct additional robustness checks and diagnostics (e.g., tests for serial correlation).
  - Extend the analysis to answer the full research questions:
    - Explore the relationship between renewable capacity share and CO_2 emissions (total and per GDP).
    - Delve deeper into the impact of green growth policies on energy mix.
- Final Deliverables:
  - Cleaned merged dataset
  - Comprehensive analysis report
  - Interactive dashboards/visualizations for stakeholders

Project Status & Roadmap

- Current Status:
  - Data integration, EDA, simple analysis, and initial advanced panel regressions have been completed.
- Am I addressing my proposal questions?
  - Partially. The current work lays the groundwork by addressing economic and energy mix relationships.
  - The next phase will directly integrate CO_2 emission data to answer the key research questions on renewable share vs. emissions.
- Roadmap:
  - Finalize model specifications
  - Incorporate CO_2 and policy analysis in detail
  - Prepare final report and recommendations

1. Data Integration & EDA ✅
2. Simple Analysis ✅
3. Advanced Panel Analysis ✅
3a. Detailed Policy Analysis (NS)
4. Incorporate CO2 Emission Data (NS)
5. Model Refinement (Upcoming)
6. Deliverables (Upcoming)