# Capstone Project:

## Worldwide Electricity Generation

# Agenda

1. Introduction
2. The Dataset
3. Explorative Data Analysis
4. Conclusions
5. Next Steps

# Introduction

This project looks at data on **world electricity consumption**. Through multilinear regression analysis we aim to find insights into some of the factors that affect electricity generation.

We focus on **electricity generation per capita**.

Through our model, we aim to provide some suggestions for countries to increase their electricity generation, and some of the complications in doing this.

# The Dataset

The **World Energy Consumption dataset** was published on kaggle by Pralabh Poudel, and is a collection of key metrics maintained by **Our World in Data**.

This dataset has energy data for every country, with 129 variables and over 22,000 rows.

Before diving into the data, we decided to focus on data from **2001 onwards**. Unfortunately there was too much missing data between 2019 - 2022, so our data focuses on 2001 - 2018.

Countries with populations **smaller than 2 million** were excluded to emphasize trends within larger population groups.

# The Dataset

Our final cleaned dataset had around 2500 rows. And focused on:

- net electric imports
- GDP per capita
- coal production
- oil production
- greenhouse gas emissions per capita
- year
- percentage share of total energy by type:
  biofuel, coal gas, hydro, oil, nuclear, solar, wind

# The Approach

So how do we determine how much each of these variables affects electricity generation per capita?

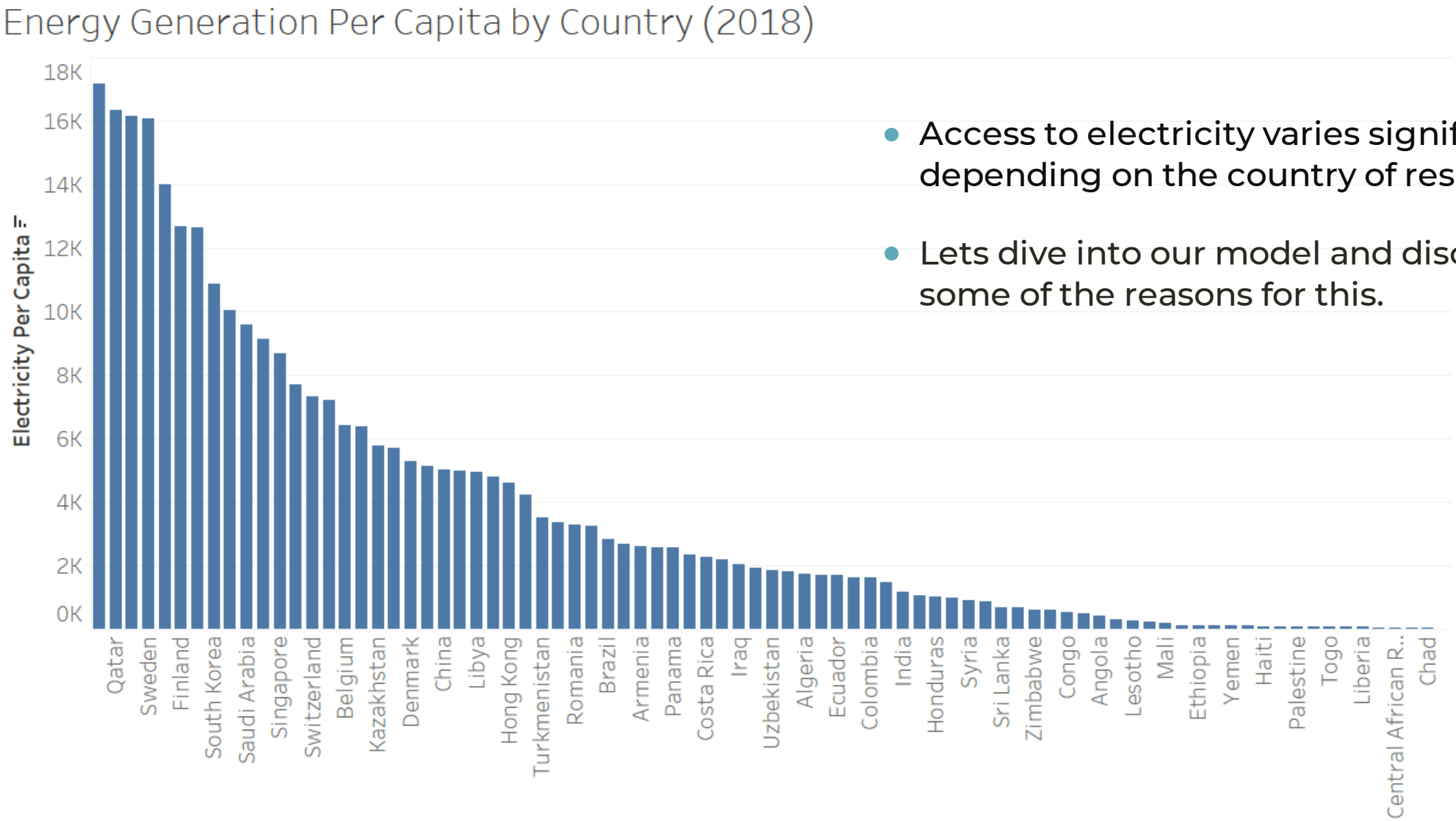
The answer is **multilinear regression analysis**.

Multilinear regression analysis is a method used to understand how multiple factors influence an outcome. It helps us identify which factors matter most and how much they affect the result. This information is valuable for home owners and potential buyers to make informed decisions.

Essentially we take all the data, and provide a line of best fit.

It took 3 iterations to produce a well fitted and robust model.

# Explorative Data Analysis




- Access to electricity varies significantly depending on the country of residence.
- Lets dive into our model and discuss some of the reasons for this.

# Conclusions

There are a number of interesting conclusions we can draw from our model:

- ***Greenhouse gas emissions per capita*** and **GDP per capita** are very highly correlated with electricity generated per capita.

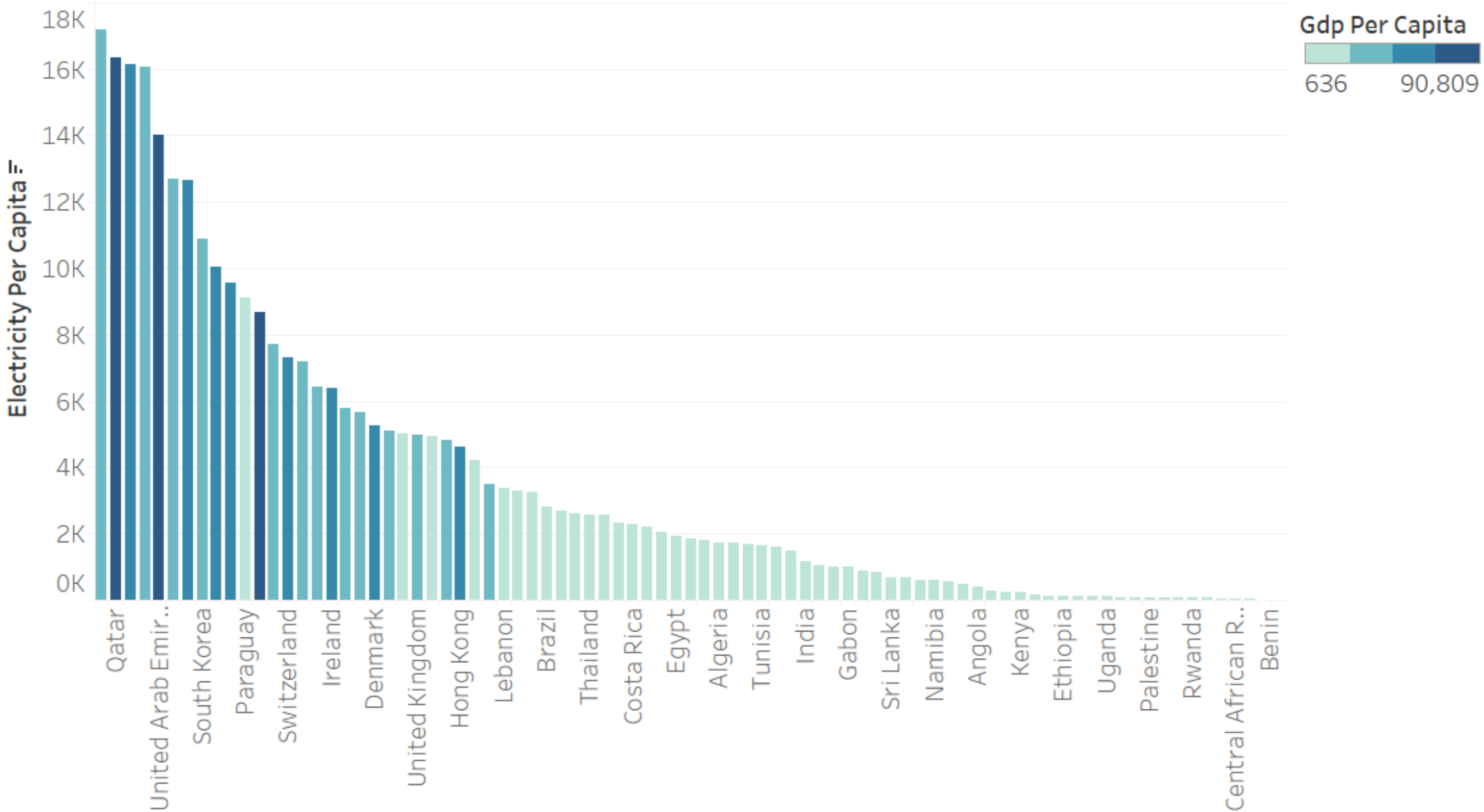
  Unfortunately our electricity is still highly reliant on carbon emitting resources. *The more electricity a country provides per person, the more CO2 is emitted.*

  Countries with low carbon emissions often lack access to electricity, this is one of the key problems we face in reducing our global emissions.

  Increasing electricity generation is an important step in increasing a countries' GDP.

  Can we raise low income countries out of poverty, without dramatically increasing their greenhouse gas emissions?

# Energy Generation Per Capita by Country (2018)



# Conclusions

- **Coal production** has a negative impact on electricity generation per capita

  The higher a country's per capita coal production, the lower its per capita electricity generation.

  This may be because coal production is an outdated form of electricity production, and more developed countries have moved away from it.

- **Oil production** has no impact on electricity generation per capita

  Oil is produced by both wealthy and poor countries, and is not exclusively used for electricity generation, this may be the reason it has no affect on our model.

# Conclusions

- **The year (2001 -2018) has no impact on electricity generation per person.**

  This is a surprise, as one would expect there to be an overall increase in energy per year. By the early 2000's, developed countries with a high GDP per capita, have reached a plateau in energy needed per person.

# Conclusions

Interestingly, the share of a countries' energy production by type seems to have a correlation on energy generation:

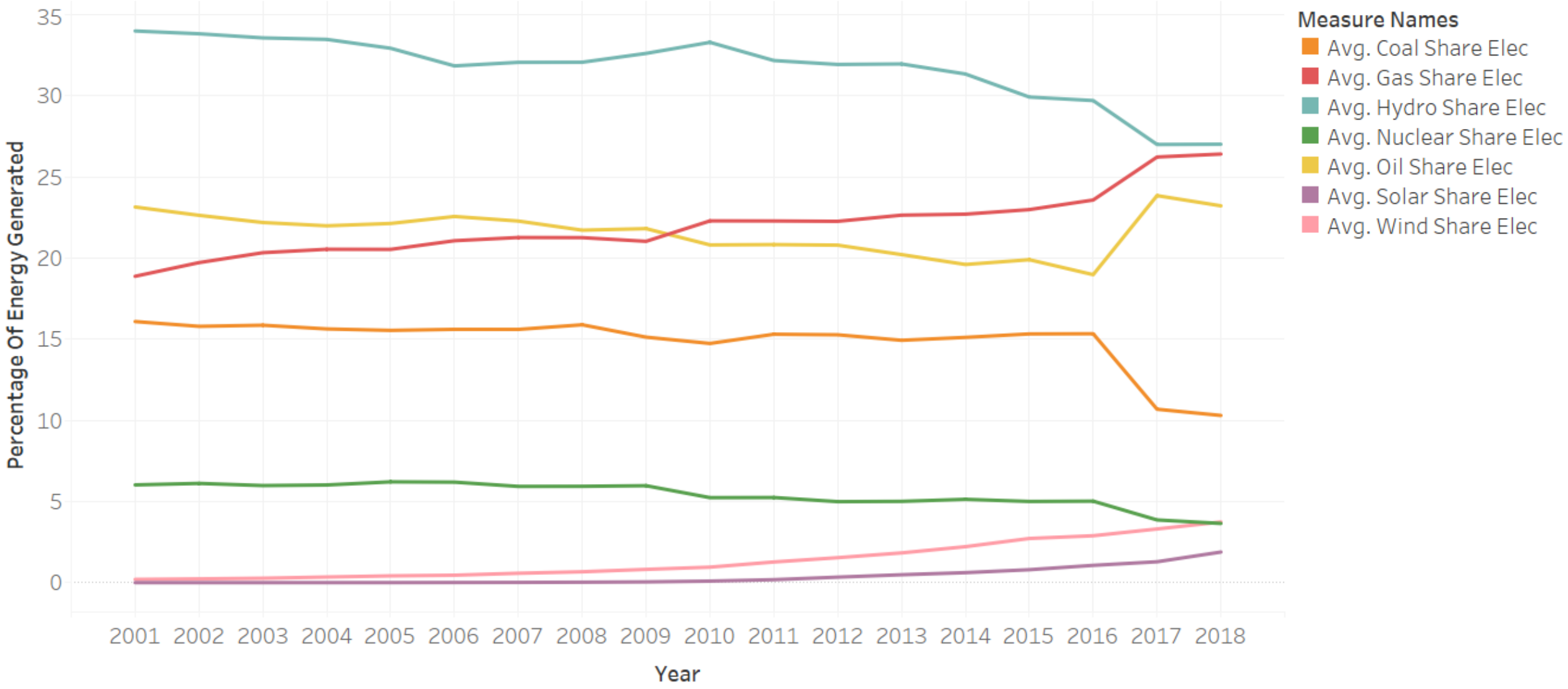
- The higher the percentage share of **hydro**, **nuclear** and **wind** generation, the **higher** the total electricity generation per capita

- The higher the percentage share of **oil**, **coal**, **biofuel** and **gas** generation, the **lower** the total energy per capita

  This could be because more developed countries, with a higher GDP per capita, are usually further along the transition to low carbon energy.

  Perhaps hydro, nuclear and wind all imply a greater diversity of energy types. Fossil fuels prices fluctuate with demand so diversifying how your energy is produced helps to limit the impact of this fluctuation.

# Average Country Share of Electricity Generated




- The average countries percentage of electricity produced by oil and gas is <u>increasing</u>


Next
Steps

Our regression model has produced some interesting insights into electricity production worldwide. There is a lot of data missing from 2019 onwards. Once the dataset has been updated, it would be great to update the model.

We have barely scratched the surface of what we can explore with this dataset. I would love to take a deeper dive into the different energy sources and the share of electricity produced.

Future analysis would look at the following questions:

- Does diversifying how you are producing energy increase supply?
- Which ways of producing energy improve affordability?
- Are there specific regions or countries where diversification of energy sources has led to significant improvements in electricity generation?
- Are there trade-offs between affordability and environmental sustainability?