

# Projectopdracht - CO2-uitstoot

Student: Noof Mousa

In dit rapport worden voor de finale projectopdracht van de cursus 'Data Analytics with Python' de volgende vragen beantwoord:

1. What is the biggest predictor of a large CO2 output per capita of a country (*Wat is de grootste voorspeller van een hoog CO2-uitstoot per hoofd (persoon) van de bevolking van een land*)?
2. Which countries are making the biggest strides in decreasing CO2 output (*Welke landen hebben de grootste stappen gezet in het terugdringen van de CO2-uitstoot*)?
3. Which non-fossil fuel energy technology will have the best price in the future (*Welke niet-fossiele energietechnologie zal in de toekomst de beste prijs hebben*)?

Voor het beantwoorden van deze vragen is de volgende data gebruikt van 'Our World in Data':

- <https://raw.githubusercontent.com/owid/co2-data/master/owid-co2-data.csv>
- <https://ourworldindata.org/grapher/levelized-cost-of-energy.csv?v=1&csvType=full&useColumnShortNames=true>, storage\_options = {'User-Agent': 'Our World In Data data fetch/1.0'}

Dit rapport is in het Nederlands geschreven, maar de notebook in Google Colab is in het Engels gemaakt, omdat de data en de cursus ook in het Engels zijn.

## Algemeen

Voor het beantwoorden van de vragen zijn de Python bibliotheken in Google Colab Notebook geïmporteerd. Vervolgens is de data geïmporteerd. Voor het gebruik van de data is de dataframe gecontroleerd en waar nodig opgeschoond, waarbij ook een kopie van de originele data is gemaakt.

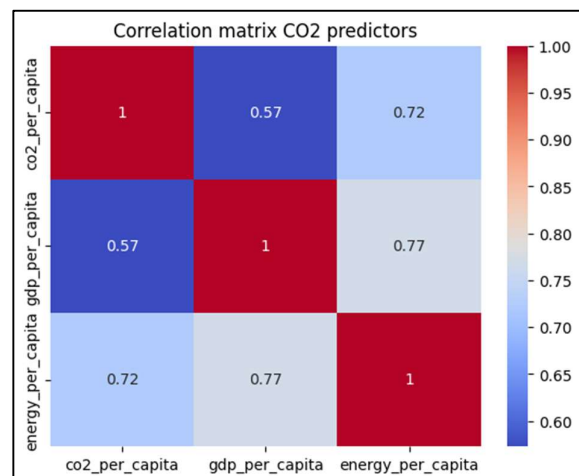
## Wat is de grootste voorspeller van een hoog CO2-uitstoot per hoofd (persoon) van de bevolking van een land

Voor het beantwoorden van deze vraag is de data gebruikt: <https://raw.githubusercontent.com/owid/co2-data/master/owid-co2-data.csv>.

Als eerst is een nieuwe kolom berekend en toegevoegd aan deze dataframe; gdp\_per\_capita. In dit geval staat gdp voor gross domestic product. Vervolgens zijn uit de dataframe de volgende kolommen gebruikt door deze kolommen te filteren in een kopie van de dataframe:

- Country (Land)
- Year (Jaar)
- co2\_per\_capita (CO2-uitstoot per persoon)
- gdp\_per\_capita (GDP per persoon)
- population (bevolking)
- energy\_per\_capita (energiegebruik per persoon)

Om te analyseren of de GDP per persoon en het energiegebruik per persoon voorspellers zijn voor een hoog CO2-uitstoot per persoon is de correlatie berekend tussen deze drie factoren. Dit is zichtbaar in figuur 1: Correlation matrix. Uit deze matrix kan geconcludeerd worden dat de CO2-uitstoot per persoon met zowel het energiegebruik per persoon als de GDP per persoon een positieve correlatie heeft. Daarnaast kan ook geconcludeerd worden dat het energiegebruik per persoon de grootste voorspeller is voor een hoog CO2-uitstoot per persoon. Deze correlatie betekent echter niet dat er sprake is van een oorzakelijk verband (causation).



Figuur 1: Correlation Matrix

Andere factoren zoals transport en levensstijlkeuzes kunnen ook de CO2-uitstoot beïnvloeden. Om een oorzakelijk verband (causation) vast te stellen is meer onderzoek nodig.

In de bijlage zijn 2 grafieken toegevoegd waaruit ook blijkt dat de CO2-uitstoot per persoon met zowel het energiegebruik per persoon als de GDP per persoon een positieve correlatie heeft.

Figuur 4 laat per land de correlatie zien tussen de CO2-uitstoot per persoon en het energiegebruik per persoon voor tien jaar in de periode 2013-2022.

Figuur 5 laat per land de correlatie zien tussen de CO2-uitstoot per persoon en GDP per persoon voor tien jaar in de periode 2013-2022.

### **Welke landen hebben de grootste stappen gezet in het terugdringen van de CO2-uitstoot**

Voor het beantwoorden van deze vraag is ook de data gebruikt: <https://raw.githubusercontent.com/owid/co2-data/master/owid-co2-data.csv>

De volgende stappen zijn uitgevoerd om te analyseren welke landen de grootste stappen hebben gezet in het terugdringen van de CO2-uitstoot:

1. Er is een nieuwe dataframe gemaakt, waarbij de 1<sup>e</sup> opschonende dataframe (df) wordt gegroepeerd op de kolom 'country' (land).
2. Voor elke land wordt vervolgens de eerste en laatste waarde van de kolom 'CO2\_per\_capita' (CO2-uitstoot per persoon) opgehaald.
3. Daarna worden twee nieuwe kolommen aangemaakt voor de eerste en laatste waarde van de CO2-uitstoot per persoon, waarbij uit de kolom 'country' de landen worden verwijderd waarvan zowel de eerste als de laatste waarde van de CO2-uitstoot per persoon ontbreekt.
4. Als laatste wordt de relatieve verandering in percentage tussen de laatste en eerste waarde van de CO2-uitstoot per persoon berekend. Dit resulteert in de volgende top 5 landen met de grootste daling van de CO2-uitstoot per persoon:

Country	First	Last	% change
Liberia	1.056	0.149	-85.890152
Ukraine	14.546	3.438	-76.364636
Gabon	8.852	2.327	-73.712155
Luxembourg	36.981	11.138	-69.881831
Kuwait	67.905	21.235	-68.72837

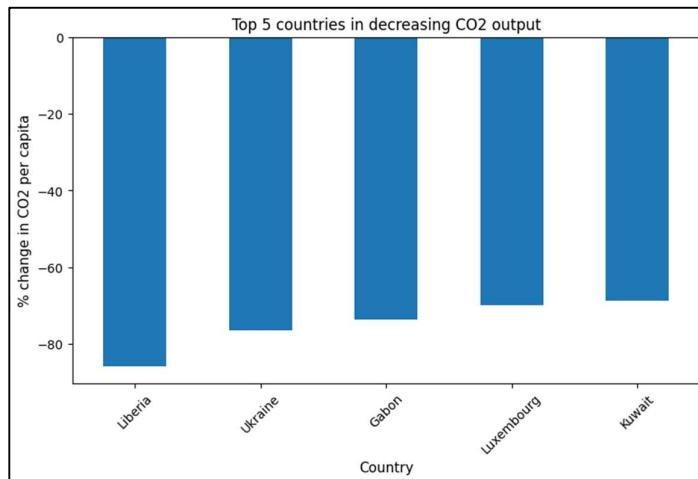


Figure 2: Top 5 countries in decreasing CO2 output

### Welke niet-fossiele energietechnologie zal in de toekomst de beste prijs hebben

Voor deze vraag is de data uit webpagina van 'Our World in Data' gedownload via de downloadopties onder de grafiek.

Er was geen data te vinden met specifieke informatie over de prijzen per energietechnologie. Hierdoor is voor het beantwoorden van deze vraag de aanname gemaakt dat de prijs per energietechnologie gelijk is aan de kosten per energietechnologie.

Om de toekomstige prijs per energietechnologie te berekenen is gebruikt gemaakt van lineaire regressie.

De toekomstige prijs wordt hierbij berekend op basis van historische data over de prijs per energietechnologie.

Voor de analyse is gekozen voor het jaar 2030, vijf jaar in de toekomst vanaf dit moment (december 2024).

In het regressiemodel worden X (jaren) als onafhankelijke variabele en Y (kosten voor de specifieke energietechnologie) als afhankelijke variabele gebruikt.

De onafhankelijke variabele is de variabele die in deze analyse wordt controleert of gemanipuleerd. In dit geval vertegenwoordigt X de jaren.

De afhankelijke variabele is de uitkomst die met deze analyse wordt voorspeld. In dit geval vertegenwoordigt Y de prijs per energietechnologie.

In deze analyse wordt het lineaire regressie algoritme gebruikt om de best passende lijn te vinden die beschrijft hoe de afhankelijke variabele (prijs) verandert ten opzichte van de onafhankelijke variabele (jaren). Dit is weergegeven in figuur 3 hieronder.

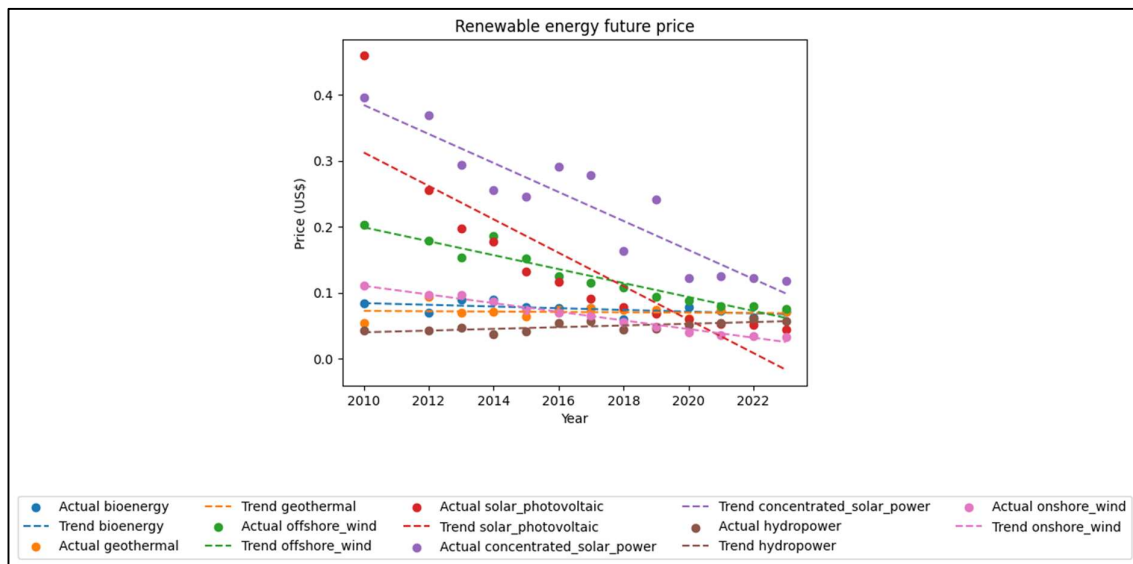


Figure 3: Renewable energy future price

Op basis van de uitgevoerde analyse zijn de volgende energietechnologieën de top 3 met naar verwachting in de beste prijs in 2030:

1. Solar\_photovoltaic
2. Concentrated\_solar\_power
3. Offshore\_wind

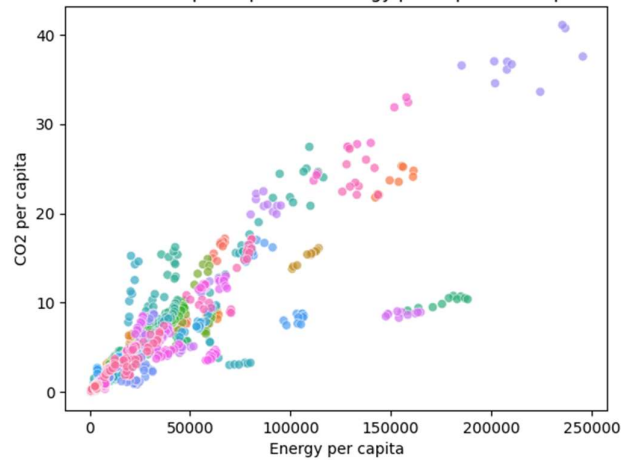
In deze analyse wordt een poging gedaan om de prijs per energietechnologie te voorspellen, waarbij Solar Photovoltaic de grootste daling laten zien. Echter, als we een lijn proberen te trekken van de realisatie van Solar Photovoltaic, lijkt de lijn op exponentiële afname. Bij een exponentiële afname daalt de lijn aan het begin snel en daarna daalt de lijn langzamer per jaar. Hierdoor wordt aanbevolen om voor het trekken van harde conclusies om basis van lineaire regressie een analyse met betrekking tot exponentiële afname voor Solar Photovoltaic uit te voeren.

### **Slot**

Deze projectopdracht benadrukt het belang van het begrijpen van de factoren die van invloed zijn op de CO<sub>2</sub>-uitstoot en het potentieel van hernieuwbare energietechnologieën. Overheidsinstanties, bedrijven en individuen zouden samen moeten werken om duurzame toepassingen te bevorderen en te investeren in schone energietechnologieën om een gezonde leefomgeving voor toekomstige generaties te creëren.

## Bijlage

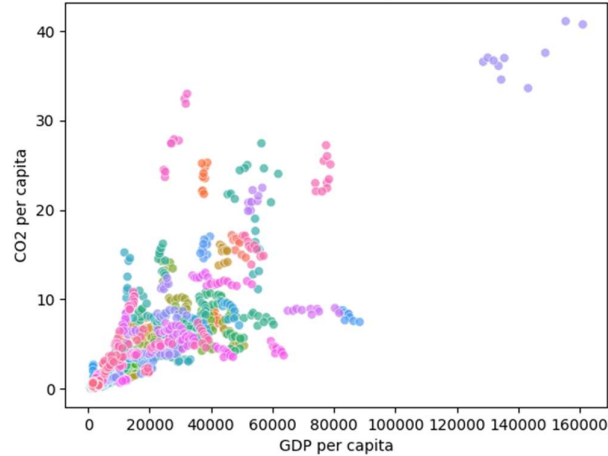
Correlation between CO2 per capita and Energy per capita in the period 2013-2022



● Afghanistan	● Costa Rica	● Iceland	● Montenegro	● Sierra Leone
● Albania	● Cote d'Ivoire	● India	● Morocco	● Singapore
● Algeria	● Croatia	● Indonesia	● Mozambique	● Slovakia
● Angola	● Cuba	● Iran	● Myanmar	● Slovenia
● Argentina	● Cyprus	● Iraq	● Namibia	● South Africa
● Armenia	● Czechia	● Ireland	● Nepal	● South Korea
● Australia	● DR Congo	● Israel	● Netherlands	● Spain
● Austria	● Denmark	● Italy	● New Zealand	● Sri Lanka
● Azerbaijan	● Djibouti	● Jamaica	● Nicaragua	● Sweden
● Bahrain	● Dominica	● Japan	● Niger	● Switzerland
● Bangladesh	● Dominican Republic	● Jordan	● Nigeria	● Syria
● Barbados	● Ecuador	● Kazakhstan	● North Korea	● Taiwan
● Belarus	● Egypt	● Kenya	● North Macedonia	● Tajikistan
● Belgium	● El Salvador	● Kuwait	● Norway	● Tanzania
● Benin	● Equatorial Guinea	● Kyrgyzstan	● Oman	● Thailand
● Bolivia	● Estonia	● Laos	● Pakistan	● Togo
● Bosnia and Herzegovina	● Eswatini	● Latvia	● Palestine	● Trinidad and Tobago
● Botswana	● Ethiopia	● Lebanon	● Panama	● Tunisia
● Brazil	● Finland	● Lesotho	● Paraguay	● Turkey
● Bulgaria	● France	● Liberia	● Peru	● Turkmenistan
● Burkina Faso	● Gabon	● Libya	● Philippines	● Uganda
● Burundi	● Gambia	● Lithuania	● Poland	● Ukraine
● Cambodia	● Georgia	● Luxembourg	● Portugal	● United Arab Emirates
● Cameroon	● Germany	● Madagascar	● Qatar	● United Kingdom
● Canada	● Ghana	● Malawi	● Romania	● United States
● Cape Verde	● Greece	● Malaysia	● Russia	● Uruguay
● Central African Republic	● Guatemala	● Mali	● Rwanda	● Uzbekistan
● Chad	● Guinea	● Malta	● Saint Lucia	● Venezuela
● Chile	● Guinea-Bissau	● Mauritania	● Sao Tome and Principe	● Vietnam
● China	● Haiti	● Mauritius	● Saudi Arabia	● Yemen
● Colombia	● Honduras	● Mexico	● Senegal	● Zambia
● Comoros	● Hong Kong	● Moldova	● Serbia	● Zimbabwe
● Congo	● Hungary	● Mongolia	● Seychelles	

Figuur 4: Correlation between CO2 per capita and Energy per capita in the period 2013-2022

Correlation between CO2 per capita and GDP per capita in the period 2013-2022



● Afghanistan	● Costa Rica	● Iceland	● Montenegro	● Sierra Leone
● Albania	● Cote d'Ivoire	● India	● Morocco	● Singapore
● Algeria	● Croatia	● Indonesia	● Mozambique	● Slovakia
● Angola	● Cuba	● Iran	● Myanmar	● Slovenia
● Argentina	● Cyprus	● Iraq	● Namibia	● South Africa
● Armenia	● Czechia	● Ireland	● Nepal	● South Korea
● Australia	● DR Congo	● Israel	● Netherlands	● Spain
● Austria	● Denmark	● Italy	● New Zealand	● Sri Lanka
● Azerbaijan	● Djibouti	● Jamaica	● Nicaragua	● Sweden
● Bahrain	● Dominica	● Japan	● Niger	● Switzerland
● Bangladesh	● Dominican Republic	● Jordan	● Nigeria	● Syria
● Barbados	● Ecuador	● Kazakhstan	● North Korea	● Taiwan
● Belarus	● Egypt	● Kenya	● North Macedonia	● Tajikistan
● Belgium	● El Salvador	● Kuwait	● Norway	● Tanzania
● Benin	● Equatorial Guinea	● Kyrgyzstan	● Oman	● Thailand
● Bolivia	● Estonia	● Laos	● Pakistan	● Togo
● Bosnia and Herzegovina	● Eswatini	● Latvia	● Palestine	● Trinidad and Tobago
● Botswana	● Ethiopia	● Lebanon	● Panama	● Tunisia
● Brazil	● Finland	● Lesotho	● Paraguay	● Turkey
● Bulgaria	● France	● Liberia	● Peru	● Turkmenistan
● Burkina Faso	● Gabon	● Libya	● Philippines	● Uganda
● Burundi	● Gambia	● Lithuania	● Poland	● Ukraine
● Cambodia	● Georgia	● Luxembourg	● Portugal	● United Arab Emirates
● Cameroon	● Germany	● Madagascar	● Qatar	● United Kingdom
● Canada	● Ghana	● Malawi	● Romania	● United States
● Cape Verde	● Greece	● Malaysia	● Russia	● Uruguay
● Central African Republic	● Guatemala	● Mali	● Rwanda	● Uzbekistan
● Chad	● Guinea	● Malta	● Saint Lucia	● Venezuela
● Chile	● Guinea-Bissau	● Mauritania	● Sao Tome and Principe	● Vietnam
● China	● Haiti	● Mauritius	● Saudi Arabia	● Yemen
● Colombia	● Honduras	● Mexico	● Senegal	● Zambia
● Comoros	● Hong Kong	● Moldova	● Serbia	● Zimbabwe
● Congo	● Hungary	● Mongolia	● Seychelles	

Figuur 5: Correlation between CO2 per capita and GDP per capita in the period 2013-2022