# 6PAM2026 Data Science Project Module

## Initial Project Plan

**Student Full Name:** Riya Chandaria
**Student ID:** 21039580
**Course:** BSc Data Science Project

| | |
|---|---|
| Your career aims / industry you would like to get into. | |
| | |
| **Proposed project title (topic(s) that you are interested in).** | Predicting future CO2 and Greenhouse Gas Emissions by country/region and the breakdown of factors contributing to the emissions. |
| **Dataset(s) that you may use.** | https://github.com/owid/co2-data |
| **Data science/AI techniques that you may use.** | • Exploratory data analysis (i.e. data visualisation, correlation analysis)<br>• Machine learning techniques:<br>  ○ Regression models (i.e. linear/multivariate regression, ridge/lasso regression, logistic regression)<br>  ○ Classification models (i.e. random forests, gradient boosting, logistic regression, support vector machines)<br>  ○ Neural Networks<br>• Test and training data splits<br>• Evaluation metrics (i.e. mean absolute error, mean squared error, r-squared, confusion matrix)<br>• Plots and graphs |
| | |

## 1. Background and objectives

The world is becoming more concerned about the rate of increase in climate change, therefore now there is an urgent need for accurate insights into CO2 and greenhouse gas emission patterns. The major contributor to the greenhouse effect and global warming is CO2 emissions, with the majority coming from human-based sources such as deforestation and burning fossil fuels. There are plans in place like the Paris Agreement, which help monitor, predict and reduce future emissions.

As being able to accurately predict future CO2 and greenhouse gas emissions is critical to reducing climate change, my project aims to use historical data from countries all around the world between 1750 & 2022 and use machine learning techniques to predict future emissions and the different factors which will have the biggest impact. This will then provide insights into understanding the paths countries are on and where interventions are most needed to help reduce emissions.

## 2. Methodology

I will begin my project by preprocessing the historical data, from “Our World in Data”, such as handling any missing data, normalisation, etc. From here I can begin with the exploratory data analysis stage, where I will see if there are trends, correlations or outliers in the data. I will also use machine learning techniques to build predictive models by using regression models. Once the data has been split into test and training splits, I will train and test my models, to ensure they are robust. Finally, I will present my results visually and breakdown the contributing factors by country/region. This will offer insights into future emission trends and the key factors behind them.

## 3. References

- (Ritchie, et al., 2023) - https://ourworldindata.org/co2-and-greenhouse-gas-emissions
- (Ritchie, et al., 2024) - https://github.com/owid/co2-data
- (Jin, et al., 2024) - https://www.sciencedirect.com/science/article/pii/S0048969724024628#s0010
- (Bhatt, et al., 2023) - https://www.sciencedirect.com/science/article/pii/S2772782323000037#sec0004

## References


Bhatt, H. et al., 2023. Forecasting and mitigation of global environmental carbon dioxide emission using machine learning techniques. *Cleaner Chemical Engineering,* March.Volume 5.

Jin, Y. et al., 2024. Carbon emission prediction models: A review. *Science of The Total Environment,* 1 June.Volume 927.

Ritchie, H., Rosado, P. & Roser, M., 2023. *CO2 and Greenhouse Gas Emissions.* [Online]
Available at: https://ourworldindata.org/co2-and-greenhouse-gas-emissions

Ritchie, H., Rosado, P. & Roser, M., 2024. *Data on CO2 and Greenhouse Gas Emissions by Our World in Data.* [Online]
Available at: https://github.com/owid/co2-data


## 4. Timeline (PTO)

**Predicting future CO2 and Greenhouse Gas Emissions by country/region and the breakdown of factors contributing to the emissions**
PROJECT NAME

**Riya Chandaria**
STUDENT NAME

**October 2024**
START DATE

**April 2024**
END DATE

| Task ID | Task Name | Start Date | End Date | Duration (In Days) | 07/10/2024 | 14/10/2024 | 18/10/2024 | 21/10/2024 | 28/10/2024 | 04/11/2024 | 07/11/2024 | 11/11/2024 | 18/11/2024 | 25/11/2024 | 02/12/2024 | 09/12/2024 | 16/12/2024 | 23/12/2024 | 30/12/2024 | 06/01/2025 | 13/01/2025 | 20/01/2025 | 27/01/2025 | 03/02/2025 | 10/02/2025 | 17/02/2025 | 24/02/2025 | 03/03/2025 | 10/03/2025 | 17/03/2025 | 24/03/2025 | 28/03/2025 | 31/03/2025 | 07/04/2025 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| T01 | Initial Project Plan | 7th Oct | 18th Oct | 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T02 | Updating GitHub account | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T03 | Preparation for the ethics quiz | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T04 | Write up – background & motivation | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T05 | Write up - literature review | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T06 | Write up – data set | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T07 | Coding – model 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T08 | Coding – model 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T09 | Coding – mode 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| T10 | Write up - methodology | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T11 | Write up – analysis & results | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T12 | Write up - conclusions & future outlook | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T13 | Review | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T14 | Presentation preparation | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T15 | VIVA preparation | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |