

### ### Project Analysis Framework

Goal :

Use clear and concise data to reveal differences in how industries perform in the low-carbon economy. Provide data-driven insights to guide businesses and policymakers in developing effective low-carbon industrial policies and encouraging sustainable green investment strategies.

### ### Data Preparation and Preprocessing

#### 1. Loading Data (15 points)

- Datasets:

- Load the Public Profit and Emission Database and OWID CO2 Dataset for firm-level analysis of profitability and emissions.

- Load the CarbonPlan CDR Database to include data on carbon dioxide removal (CDR) costs.

- Task:

- Use at least two datasets.

- Perform all data processing (e.g., merging, reshaping) directly within the `.qmd`` file.

#### 2. Handling Missing Data and Cleaning (15 points)

- Steps:

- Identify missing values for each dataset and conduct preprocessing to address these gaps.

- Analyze variables with substantial missing values and investigate possible causes.

- Outcome:

- Categorize each variable using Altair's syntax (e.g., Quantitative, Nominal) after addressing missing data.

### 3. Data Merging and Structuring (10 points)

- Task:
  - Merge profitability and emissions data to create a unified analysis dataset.
  - Identify and resolve issues such as duplicate entries or data inconsistencies during the merging process.

### ### Visualization and Analysis

#### Question 2: Comparing Carbon Emission Efficiencies Across Industries

Goal: Explore how industries vary in terms of carbon emissions efficiency and profitability.

#### 1. Static Plots (25 points)

- Create at least two static visualizations using Altair or GeoPandas:
  - Scatter Plot:
    - X-axis: Profitability
    - Y-axis: Carbon emissions
    - Industries differentiated by color.
    - Expected Insights:
      - Positively Correlated Industries: Energy or heavy industries may show higher emissions with higher profitability.
      - Negatively Correlated or Neutral Industries: Technology or service sectors may exhibit low emissions but achieve high profitability.
  - Bar Chart:
    - Compare carbon emission efficiencies across industries.
    - Industries with higher bars represent lower efficiency (more emissions per dollar of profit).

- Expected Insights: Highlight sectors needing improvement in emissions management while maintaining profitability.

- Additional options:

- Visualize emission distributions across industries to identify sectors with the highest contributions.

## 2. Correlation Coefficient Analysis (15 points)

- Task:

- Calculate the correlation coefficient to quantify the linear relationship between profitability and carbon emissions.

- Expected Insights:

- Strong positive correlation: Industries should adopt strategies for emission reduction while exploring low-carbon profitability models.

- Weak or no correlation: Sectors with low emissions yet high profitability are prime targets for policy support.

## ### Question 3: Shiny Dashboard Development (25 points)

Design a Shiny dashboard (2020–2022 data) for dynamic exploration of carbon emissions by industry or company.

- Features:

- Include a dynamic chart to display differences in emissions across industries or the correlation between profitability and emissions.

- Allow users to filter data by industry or firm.

- Analyze carbon emission efficiency of industries and its relationship with profitability.

## Modeling Carbon Emission Intensities (20 points)

Task:

Use the provided carbon emissions data to calculate emission intensity (carbon emissions per unit of profit) for firms and industries.

Compare emission intensities across industries using Pandas and Altair to identify patterns and outliers.

Simulate scenarios to explore how reducing emission intensities could impact overall profitability.

### ### Additional Components and Writeup

#### 1. Reproducibility (10 points)

- Requirements:

- Share all code on GitHub.
- Include dataset links and `.qmd` file.`
- Ensure reproducibility by enabling TAs to clone the repository, knit the `.qmd` file,` and obtain identical results.

#### 2. Git Usage (10 points)

- Requirements:

- Use multiple branches for different project phases.
- Final repository must have a single main branch with a commit history showing equal contributions from group members.

#### 3. Extra Credit: Text Analysis (up to 10 points)

- Task:

- Apply natural language processing (NLP) to analyze corporate sustainability reports

or climate policy discussions.

- Example: Examine New York Times article "Climate Summit, in Early Days, Is Already on a 'Knife Edge'" to explore profitability and emissions in climate policies.

<https://www.nytimes.com/2024/11/13/climate/cop24-climate-finance.html>

Expected Outcome:

Provide comprehensive insights for policymakers and businesses to align profitability with sustainability in the low-carbon economy.