Corporate GHG Emissions Analysis Using Open Sustainability Data

1. Introduction and Objective

Greenhouse gas (GHG) emissions measurement is a critical component of corporate climate action, sustainability reporting, and climate risk management. Accurate accounting of emissions enables organizations to identify emission hotspots, assess climate-related risks, and design effective mitigation strategies.
The objective of this mini project is to analyse **real-world corporate GHG emissions data** sourced from open sustainability disclosures and to demonstrate a practical emissions accounting workflow aligned with internationally accepted standards. Specifically, the project aims to:

- Analyse **Scope 1, Scope 2 (market-based), and Scope 3 emissions**
- Quantify total and scope-wise emissions in **tCO_2e**
- Identify key emission patterns across scopes
- Develop an **Excel-based emissions model and dashboard**
- Align outputs with the **GHG Protocol** and climate disclosure expectations under **ISSB / IFRS S2**

2. Data Source

The analysis uses an **open-source corporate emissions dataset** made available through the Hugging Face platform. The dataset is compiled from publicly available corporate sustainability and annual reports and contains reported greenhouse gas emissions for multiple companies.

Key characteristics of the dataset include:

- Emissions reported in **tonnes of CO_2 equivalent (tCO_2e)**
- Coverage of **Scope 1, Scope 2, and Scope 3 emissions**
- Use of anonymized company identifiers (`Company_ID`)
- Multiple reporting years across companies

For this analysis, only the **latest available reporting year** was selected to ensure consistency and comparability.

3. Methodology

3.1 Emissions Accounting Framework

The analysis follows the **GHG Protocol Corporate Accounting and Reporting Standard**, which categorizes emissions into three scopes:

- **Scope 1:** Direct emissions from owned or controlled sources
- **Scope 2 (market-based):** Indirect emissions from purchased electricity, reflecting contractual sourcing choices
- **Scope 3:** Indirect emissions occurring across the value chain

The **market-based approach** was used for Scope 2 emissions as it is widely adopted in corporate reporting and reflects electricity procurement decisions.

3.2 Data Cleaning and Preparation

The following data preparation steps were applied using Microsoft Excel:

- Selection of relevant variables: Company ID, reporting year, Scope 1, Scope 2 (market-based), and Scope 3 emissions
- Replacement of missing values with zero to allow consistent aggregation
- Filtering to retain a single reporting year
- Removal of records where emissions for all three scopes were reported as zero

These steps ensured that only meaningful and comparable data points were included in the analysis.

3.3 Emissions Calculations

For each company:

- **Total emissions (tCO_2e)** were calculated as the sum of Scope 1, Scope 2, and Scope 3 emissions
- **Scope contribution percentages** were calculated to assess the relative importance of each scope

4. Results and Analysis

4.1 Scope-wise Emissions Profile

The aggregated results indicate that **Scope 3 emissions constitute the largest share of total corporate emissions**, followed by Scope 2 emissions, with Scope 1 emissions contributing the smallest share for most companies.

This distribution is consistent with real-world corporate emissions profiles, where value chain activities—such as purchased goods, transportation, and product use—typically dominate the overall carbon footprint.

4.2 Dashboard Overview

An Excel-based dashboard was developed to visually summarise the results, including:

- Total emissions by scope (tCO_2e)
- Relative contribution of Scope 1, Scope 2, and Scope 3 emissions
- Key summary indicators such as total emissions and average Scope 3 share

The dashboard enables quick interpretation of emissions structure and helps identify priority areas for mitigation.

5. Key Insights

- **Scope 3 emissions dominate corporate GHG footprints**, highlighting the importance of value chain engagement
- **Scope 2 emissions** represent a significant opportunity for reduction through renewable electricity sourcing
- **Scope 1 emissions**, while smaller in magnitude, are the most directly controllable
- Effective decarbonisation strategies must extend beyond direct operations to include suppliers and downstream activities

6. Disclosure Relevance (ISSB / IFRS S2)

This analysis aligns with key climate disclosure requirements under IFRS S2, including:

- Disclosure of **absolute GHG emissions**
- Coverage of **Scope 1, Scope 2, and Scope 3 emissions**
- Transparent description of **methodology and boundaries**
- Use of **tCO_2e** as a core climate metric

Such disclosures support climate risk assessment, transition planning, and comparability across organizations.

7. Limitations

- The analysis relies on **self-reported corporate data**, which may vary in completeness and quality
- Company identifiers are anonymized, limiting sector-specific or company-level benchmarking
- Detailed Scope 3 category analysis was not included to maintain simplicity

Despite these limitations, the project provides a realistic representation of corporate emissions accounting workflows.

8. Conclusion

This mini project demonstrates a practical and standards-aligned approach to analysing corporate GHG emissions using open sustainability data. By combining structured data cleaning, emissions calculations, and visual analytics, the project reflects real-world ESG and climate analytics practices.

The Excel model, dashboard, and summary report together form a strong foundation for understanding corporate climate disclosures and emissions management.