

Deep Data Hackathon 2.0 - Round 1: ESG & Financial Performance Analysis

Team: CodeAxis

Team Leader: Shreyash Jha

Members: Abhishek Raj, Geetesh Mishra, Apurva Kumari

October 2, 2025

xAI Hackathon

Project Overview

- Objective: Analyze ESG and financial performance using a dataset to derive insights and build an interactive dashboard.
- Tools: Python, Jupyter Notebook, Streamlit, VS Code.
- Deliverables: EDA notebook, cleaned dataset, Streamlit dashboard, presentation.

Workflow: Initial Setup (Steps 1–4)

- **Step 1:** Installed Python 3.9+ and verified with `python --version`.
- **Step 2:** Installed VS Code extensions (Python, Jupyter, Pylance).
- **Step 3:** Created virtual environment (`venv`) and selected interpreter.
- **Step 4:** Installed libraries (`jupyter`, `pandas`, `numpy`, etc.) using `pip`.

Workflow: Notebook Creation and Data Loading (Steps 5–6)

- **Step 5:** Created `EDA.ipynb` in VS Code for Jupyter Notebook analysis.
- **Step 6:** Added initial cells for title, library imports (`pandas`, `plotly`, etc.), and loaded `company_esg_financial_dataset.csv`.

Workflow: Data Cleaning (Step 7)

- Checked for missing values (`df.isna().sum()`) and duplicates (`df.duplicated().sum()`).
- Filled missing GrowthRate with median by CompanyID.
- Removed duplicates to ensure data integrity.

Workflow: Feature Engineering (Step 8)

- Created new features: CarbonIntensity, WaterIntensity, EnergyIntensity.
- Computed average ESG score (ESG_Avg) and revenue growth (RevenueGrowth).
- Sorted data by CompanyID and Year.

Workflow: Exploratory Data Analysis (Step 9)

- Analyzed 10+ questions with visualizations (box plots, scatter plots, line plots).
- Example: ESG scores by industry, carbon emissions vs. ESG, revenue vs. ESG trends.
- Used `seaborn` and `plotly` for interactive plots.

Key Insights (Step 10)

- Tech industry leads in ESG scores; manufacturing lags.
- Carbon intensity negatively correlates with ESG performance.
- Europe shows steady ESG improvement (2015–2022).
- Governance pillar strongly linked to profitability.

Policy Recommendations (Step 11)

- Impose carbon-intensity targets for heavy industries.
- Implement water efficiency programs for agriculture and textiles.
- Offer tax benefits for companies improving ESG year-over-year.

Workflow: Save Cleaned Data (Step 12)

- Saved cleaned dataset as `cleaned_esg_dataset.csv` for dashboard use.
- Ensured all new features and cleaned data were preserved.

Workflow: Streamlit Dashboard (Step 13)

- Created `app.py` for an interactive dashboard using Streamlit.
- Features: Year and industry filters, box plot (ESG by industry), scatter plot (carbon vs. ESG).
- Ran locally with `streamlit run app.py`.

Workflow: Repository Structure (Step 14)

- Organized project:
 - `data/company_esg_financial_dataset.csv`
 - `data/cleaned_esg_dataset.csv`
 - `notebooks/EDA.ipynb`
 - `streamlit/app.py`
 - `slides/presentation.pdf`
 - `README.md`
- Pushed to repository using VS Code Git integration.

Conclusion

- Successfully analyzed ESG and financial data, uncovering key trends.
- Built an interactive Streamlit dashboard for stakeholder insights.
- Proposed actionable policies to enhance ESG performance.
- Team CodeAxis: Shreyash Jha (Leader), Abhishek Raj, Geetesh Mishra, Apurva Kumari.