Problem Statement -3

ESG Score Prediction and Sustainable Corporate Analysis

1. Introduction

You'll use a dataset on Environmental, Social, and Governance (ESG) scores for companies. This task is like how sustainability experts or investors review data to spot trends and suggest ways companies can improve their impact on the planet and society.

Your role: Act as a data advisor for a global investment firm focused on sustainable businesses (like those using ESG ratings to pick stocks). The firm wants insights from this dataset to help companies become eco-friendlier and more responsible. Stick to exploratory data analysis (EDA) and look at patterns and trends. This lets you focus on understanding the data and coming up with fresh ideas.

2. Dataset Overview

This dataset tracks company performance over time, including financial stats and ESG metrics. ESG stands for Environmental (e.g., pollution), Social (e.g., worker rights), and Governance (e.g., fair leadership)

Dataset Link - https://www.kaggle.com/datasets/shriyashjagtap/esg-and-financial-performance-dataset

The dataset may contain inconsistencies (e.g., outliers in temperatures or populations). treat this as part of real-world data challenge.

Problem Statement

Companies today face pressure to be sustainable - reducing pollution, treating people fairly, and running ethically while still making money. High ESG scores can attract investors and help the environment. Using this dataset, explore how financial performance (like revenue and growth) links to ESG factors and resource use (like carbon and water). Find patterns by industry, region, or year. Your goal: Provide insights to guide companies toward better practices, like cutting emissions to boost scores, for a greener business world.

Objectives

- Perform EDA to understand the data's structure, quality, and key patterns.
- Develop at least 10 targeted EDA questions (e.g., "Do companies with lower carbon emissions have higher ESG scores?") to explore the data.
- Pull out insights that could help companies improve sustainability, like tips on reducing energy use.
- Visualize findings in an interactive dashboard or visualizations for stakeholders.
- Present everything like you're advising company leaders, with creative ideas on how to use the insights.

Analysis Framework: Adapted EDA Roadmap

1. Structure your project like a professional analytics workflow. This provides a clear path while allowing flexibility for innovative ideas.
2. **Data Understanding**: Load the dataset, review its structure (e.g., data types, missing values), and summarize basic statistics (e.g., means, distributions). Identify any anomalies or cleaning needs or feature engineering if required.
3. **Data Preparation**: Handle issues like duplicates, outliers, or inconsistencies. For example, normalize units if needed or group data by regions/years for better analysis. Keep it lightweight - focus on enabling exploration.
4. **Question Formulation and Exploration**: Brainstorm and select at least 10 EDA questions. These should cover univariate (single variable), bivariate (relationships between two), and multivariate (multiple variables) analyses. Use statistical methods (e.g., correlations, trends) and visualizations (e.g., histograms, scatter plots, heatmaps) to answer them. Be creative.
5. **Insight Generation**: From your EDA, extract 5-7 key insights. Connect them to real advice, e.g., "Retail companies in Latin America cut water use over time - suggest sharing best practices."
6. **Policy Recommendations and Presentation:** Translate insights into 3-5 policy proposals. Prepare a presentation that tells a story or if using PowerBI or tableau, share the dashboard file or link highlighting how your analysis supports the solution.

Files to be submitted: GitHub link which contains the following files -

Python Analytics File: A Jupyter Notebook (.ipynb) or Python script (.py)

Dashboard: An interactive visualization tool:

- If Python-based: A dashboard file (e.g., Streamlit app) or a presentation file (ppt). Use visuals from you EDA analysis
- If Power BI/Tableau: Shareable file or link (e.g., published report).

Guidelines and Tips

Tools: Python for analysis (libraries like pandas, numpy, matplotlib, seaborn). For dashboards: Python options, Power BI, or Tableau.

Creativity Space: We won't specify exact questions or policies but use your judgment to explore underrepresented angles like how ESG changes with growth.

Ethics and Realism: Ensure insights are grounded in data; avoid overgeneralizing. Think like a business analyst: Prioritize high-impact, feasible recommendations.

Collaboration: Teams of 3-4 – divide the work like one for the coding, one for the Research about domain to know exactly how each feature is related to climate change analysis which will help in better insights, one for visualization and one to present those insights with presentation or dashboards.

We're here for questions during the event. Good Luck!