Data Analyst Internship Assignment

This assignment is designed for candidates who are passionate about addressing climate change and are willing to dive deep into the domain of carbon markets. Before attempting this assignment, please familiarise yourself with key carbon market concepts and terminology. Some important terms you should understand include:

* Carbon Credit
* Issuance
* Retirement
* Methodology
* Project Developer
* Validation and Verification Body (VVB)
* Verified Carbon Unit (VCU)/Verified Emission Reduction(VER)
* Additionality
* Permanence
* Leakage

Please use AI tools (like ChatGPT, Claude) or conduct your own research to properly understand these terms and their significance in the carbon market context. A solid grasp of these concepts will be crucial for completing this assignment effectively.

**Task 1: Carbon Project Data and Analysis(34% of total score)**

1. Download projects data from at least two different carbon registries(e.g., Verra, Gold Standard)

* <https://registry.verra.org/app/search/VCS/All%20Projects>
* <https://registry.goldstandard.org/projects>
* <https://thereserve2.apx.com/myModule/rpt/myrpt.asp?r=111>

1. Clean and standardise the data from these sources, handling inconsistencies and missing values.
2. Create a unified dataset that includes key project information (e.g., project type, location, credits issued, vintage).
3. Perform basic analysis on issuances and retirements data, including:
   * Total issuances and retirements per year
   * Top project types by issuance volume
   * Geographical distribution of projects
   * Visualize key trends using appropriate charts or graphs.

**Task 2**: **Top Project Developer Profiling (33% of total score)**

1. Identify the top 20-30 project developers based on issuance volume from the dataset you built in previous task
2. Create a dataset of these developers, including:
   * Developer name
   * Website URL
   * Main project types
   * Associated projects on Verra & Gold Standard registries
   * Total credits issued, Total credits retired, Total credits remaining,
   * Any other relevant information you can find (This may involve manual data collection or web scraping, depending on data availability)

**Task 3: Credit Buyer Analysis (33% of total score)**

1. From the retirement data/notes in retired carbon credits, extract a list of organizations buying carbon credits.
2. Create a dataset of these buyers, including:
   * Buyer name
   * Total credits retired
   * Types of credits retired (if available)
   * Year(s) of retirement c) Provide a brief analysis of the top buyers and any notable trends.

**Deliverables**

1. Python scripts or Jupyter notebooks used for data analysis and any web scraping.
2. Three datasets in CSV format:
   * Cleaned and analyzed carbon project data
   * Top project developers profile
   * Carbon credit buyers
3. A brief report (max 2-3 pages) summarising your findings, including:
   * Key trends in carbon project issuances and retirements
   * Insights on top project developers
   * Notable patterns in credit retirement and buying behaviour

**Evaluation Criteria**

* Accuracy and completeness of data analysis
* Quality and relevance of the project developer and buyer datasets
* Ability to extract meaningful insights from the data
* Clear and concise presentation of findings
* Efficient use of Python for data manipulation and analysis

**Additional Notes**

* You may use AI tools like ChatGPT to assist with data interpretation or coding, but ensure you understand and can explain all aspects of your work. It’ll be asked during the interview if you’re shortlisted.
* If you encounter any challenges or make assumptions during your analysis, please document them in your report. Feel free to ask questions at [gautam@carbonmarketshq.com](mailto:gautam@carbonmarketshq.com) if something needs more clarification/information.

Email the complete artefacts & links to [gautam@carbonmarketshq.com](mailto:gautam@carbonmarketshq.com) with [shubham@carbonmarketshq.com](mailto:shubham@carbonmarketshq.com) in cc. Let us know how much time it took you to finish those tasks.