**README**

**1. Installation Requirements:**

* Python (latest version)
* Apache Spark (latest version)
* Java (latest version)
* PySpark (latest version)
* Jupyter Notebook (latest version)

**2. Setup Instructions:**

1. Install Python from [Python's official website](https://www.python.org/downloads/).
2. Install Apache Spark from [Apache Spark's official website](https://spark.apache.org/downloads.html).
3. Install Java from [Java's official website](https://www.oracle.com/java/technologies/javase-jdk11-downloads.html).
4. Install PySpark using pip:

bash

pip install pyspark

1. Install Jupyter Notebook using pip:

bash

pip install notebook

**3. Execution Guide:**

1. Download the ZIP file containing the datasets (**ghg-emissions-by-gas.csv**, **GlobalLandTemperaturesByCountry.csv**) and the Jupyter Notebook file (X23223235\_updated\_DIA\_project\_final**.ipynb**) to your local machine.
2. Extract the content of the ZIP file to a folder on your computer.
3. Open a terminal or command prompt and go to the folder containing the extracted files.
4. Launch Jupyter Notebook by running the following command:

bash

jupyter notebook

1. In the Jupyter Notebook, go to the X23223235\_updated\_DIA\_project\_final**.ipynb** file and open it.
2. Execute the code in the notebook to reproduce the analysis and results.

**4. Troubleshooting Tips:**

* **Issue**: Unable to execute Jupyter Notebook.
* **Solution**: Ensure that Jupyter Notebook is installed and configured correctly. You can install Jupyter Notebook using pip:

bash

Copy code

pip install notebook