**CarbonX: Code Description**

This Python code analyzes historical Bitcoin Greenhouse Gas (GHG) emission data and performs predictions for future emissions. It uses the following libraries:

* pandas: For data manipulation and analysis.
* numpy: For numerical computations.
* matplotlib.pyplot: For creating plots and visualizations.
* seaborn: For creating more sophisticated visualizations.
* xgboost: For training an XGBoost regression model.
* sklearn.metrics.mean\_squared\_error: For evaluating the model performance using mean squared error.

**The code performs the following tasks:**

1. Data Loading and Preprocessing:
   * Loads the historical Bitcoin GHG emission data from a CSV file.
   * Prints the shape and column datatypes of the loaded data.
   * Renames the columns and modifies the date format.
   * Prepares the dataset for analyzing GHG emissions.
2. Data Visualization:
   * Plots the GHG emissions over time.
   * Splits the data into training and test sets and visualizes the split.
   * Displays a specific week of GHG emissions.
   * Creates box plots to visualize emissions by year and month.
3. Feature Engineering:
   * Defines a function to create time series features based on the index.
   * Applies feature engineering to the dataset.
4. Model Training:
   * Divides the data into training and test sets for model training.
   * Defines a regression model using XGBoost.
   * Trains the model and evaluates its feature importances.
   * Displays the feature importances as a bar plot.
5. Prediction and Evaluation:
   * Makes predictions on the test set and merges them with the dataset.
   * Plots the predicted values against the ground truth values.
   * Displays the estimated and predicted GHG emissions for comparison.
   * Calculates the root mean squared error (RMSE) score on the test set.
   * Calculates and displays the average prediction error by date.
6. Future Predictions:
   * Creates a new dataframe for the next year.
   * Generates time series features for the next year's data.
   * Uses the trained model to predict GHG emissions for the next year.
   * Plots the predicted values along with the historical data.

**Instructions**

**To run this code, follow these steps:**

1. Install the required libraries: pandas, numpy, matplotlib, seaborn, and xgboost.
2. Prepare the data: Provide the path to the GHG Emission.csv file in the **pd.read\_csv()** function.
3. Run the code in a Python environment that supports the required libraries.
4. Examine the output plots and printed information to analyze the GHG emissions data.
5. Review the model performance by checking the RMSE score on the test set and the average prediction error by date.
6. Use the trained model to make predictions for future GHG emissions by running the relevant code section.
7. Examine the predicted values and compare them with the historical data.

Note: Some code sections may be commented out. Uncomment these sections if you want to visualize additional plots or perform specific analyses.

Feel free to modify the code and experiment with different features, models, or data sources to gain further insights into Bitcoin GHG emissions.