**Note:** The task originally mentioned the Titanic dataset, which doesn't contain time-based data. Time series can only be implemented on the datasets that contain time related datapoints. Therefore, the analysis was performed on the **Daily Climate Dataset**, which includes time-based data points, making it suitable for time series analysis.

**Objective:**

The goal of this task was to perform time series analysis using **Pandas**, focusing on techniques like **resampling**, **rolling statistics**, and **lagging**. These methods are essential for understanding temporal patterns in time series data and can provide valuable insights into trends, seasonality, and volatility.

**Dataset:**

The dataset used in this analysis contains daily climate data for Delhi, including features such as temperature (meantemp), humidity, wind speed, and pressure. The date column is used as the index, which is essential for performing time series operations.

**Approach:**

1. **Importing Libraries and Dataset:**
   * The first step was importing the necessary libraries (pandas, numpy, matplotlib) and loading the dataset. The date column was set as the index to facilitate time-based operations.
2. **Exploratory Data Analysis (EDA):**
   * Basic exploration of the dataset was done to understand its structure and check for missing values. This was followed by generating summary statistics for the key columns in the dataset.
3. **Data Visualization:**
   * Initial visualizations were created for the meantemp (mean temperature) column to get an understanding of how the temperature fluctuates over time. This helped in identifying any obvious trends or seasonal patterns.
4. **Resampling:**
   * **Resampling** is a technique used to change the frequency of the time series data. In this case, we resampled the data from daily frequency to monthly frequency. The **monthly mean temperature** was calculated to help us understand the long-term trends and smooth out the day-to-day fluctuations.
   * This technique helped identify broader patterns in temperature over time, reducing the noise from short-term variations.
5. **Rolling Statistics:**
   * **Rolling statistics** were applied to the meantemp column to calculate the **7-day rolling mean** and **7-day rolling standard deviation**.
   * The **rolling mean** helped us observe the smoothed temperature trend over a 7-day window, while the **rolling standard deviation** highlighted the volatility of the temperature, helping us understand periods of higher or lower variability.
   * These rolling statistics are especially useful for identifying seasonality or short-term fluctuations in time series data.
6. **Lagging:**
   * **Lagging** refers to creating new features by shifting the time series data by a certain number of periods. In this case, we created **lag features** by shifting the meantemp column by 1 and 2 days. This allowed us to analyze how past temperatures relate to the current temperature, which is valuable for time series forecasting models.
   * Lagging is particularly useful in predictive modeling, as it captures temporal dependencies between past and present values.
7. **Data Preprocessing for Prediction:**
   * The meantemp column was selected as the target variable for prediction. The data was split into training and testing sets (80% for training and 20% for testing).
   * A **sliding window approach** was used to create features for the model, where the last 60 days of temperature data were used to predict the next day's temperature. This method helps the model learn the patterns and relationships in the time series.
8. **Multivariate Time Series Prediction:**
   * The dataset was extended to include multiple features (temperature, humidity, wind speed, and pressure). By including more features, we aimed to capture more complex relationships in the data for better prediction accuracy.
   * The same data preprocessing steps were applied to this multivariate data.
9. **Forecasting Future Values:**
   * A function was implemented to forecast the next **30 days** of climate data. The model used the most recent data to make predictions for the next month, allowing us to visualize the expected trends in temperature, humidity, wind speed, and pressure.
   * This process involved using the trained model to iteratively predict future values, updating the input data with each new prediction.

**Results:**

* **Resampling**: By resampling the data to monthly frequency, we were able to observe the overall temperature trend while smoothing out short-term fluctuations. This helped in identifying broader climate patterns.
* **Rolling Statistics**: The rolling mean provided insights into the long-term trend, while the rolling standard deviation helped identify periods of higher volatility. These techniques are crucial for understanding the short-term dynamics in the data.
* **Lagging**: The lag features allowed us to explore the relationship between current and past temperature values. These lag features are often used in time series forecasting to capture dependencies between time periods.

Visual result of predictions:

