1. **Task Description**

This task focuses on performing **time series analysis** using Pandas. The key objectives include:

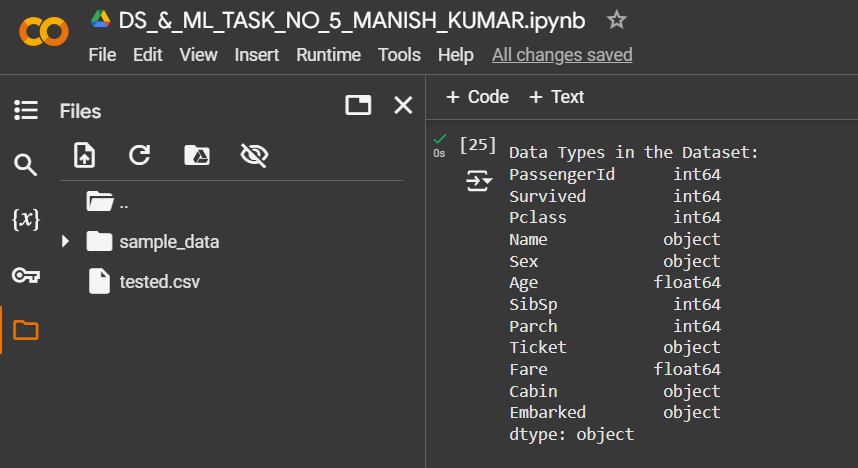
* Simulating a datetime column for time-based indexing and operations.
* Handling missing values in numeric columns.
* Resampling the dataset to calculate weekly averages.
* Computing rolling statistics like rolling mean and rolling standard deviation.
* Creating lagged features to identify time-based dependencies.
* Visualizing the time series using Matplotlib to understand trends, patterns, and variability.

1. **Attach Screenshot of Output**

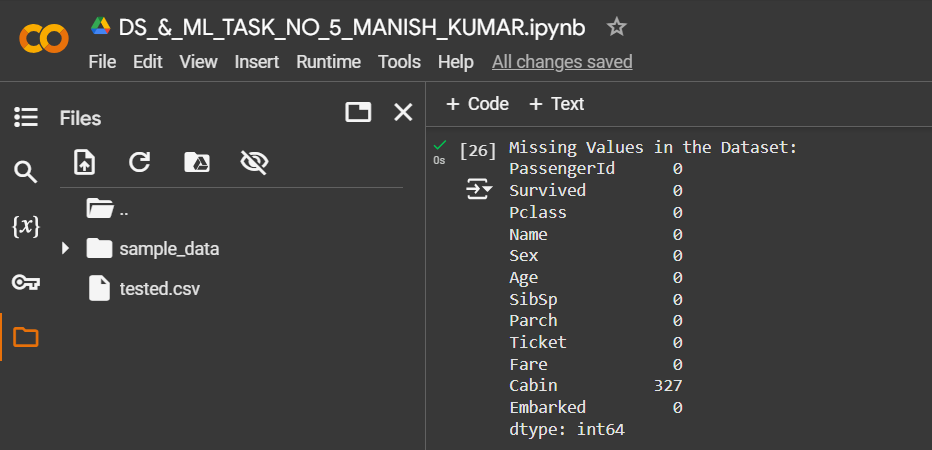
* **Dataset Preview:**



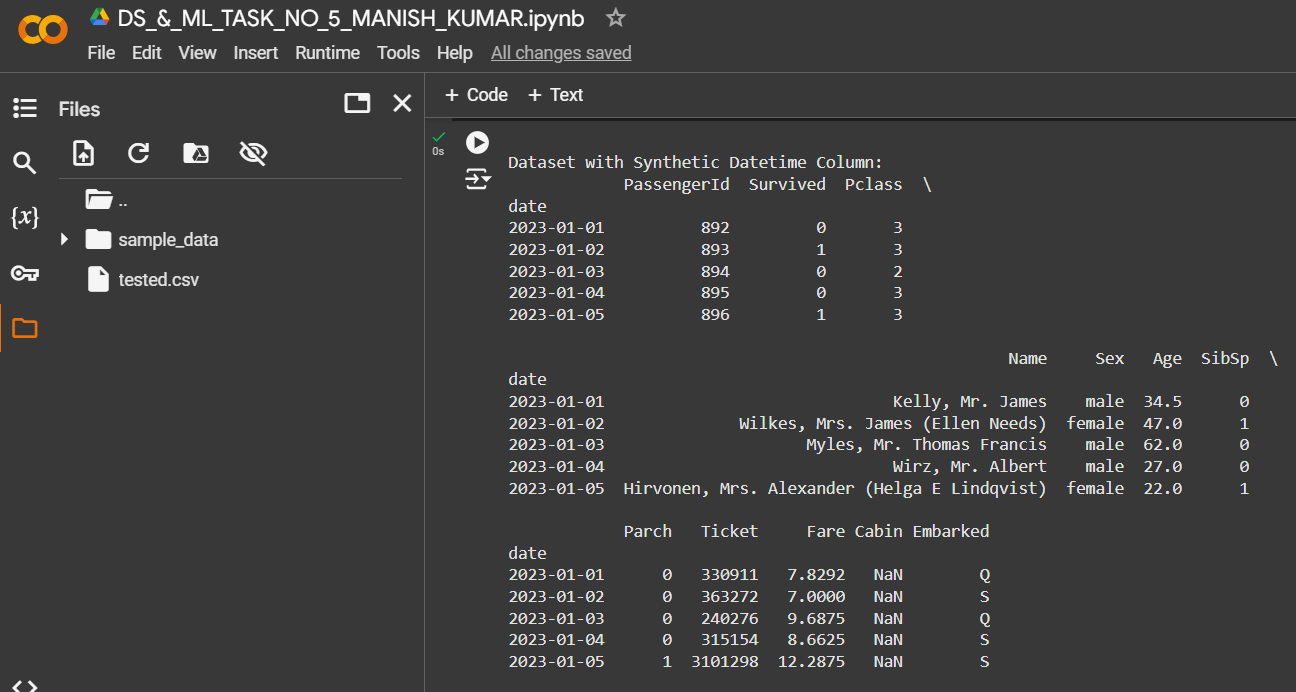
* **Data Types in the Dataset:**



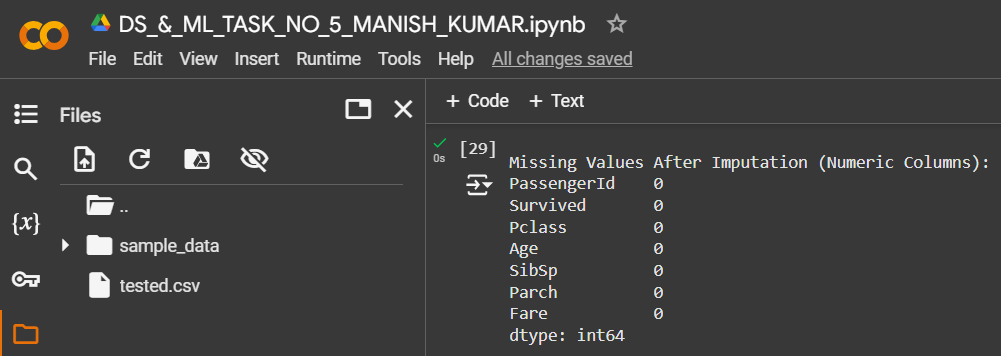
* **Missing values in the Dataset:**



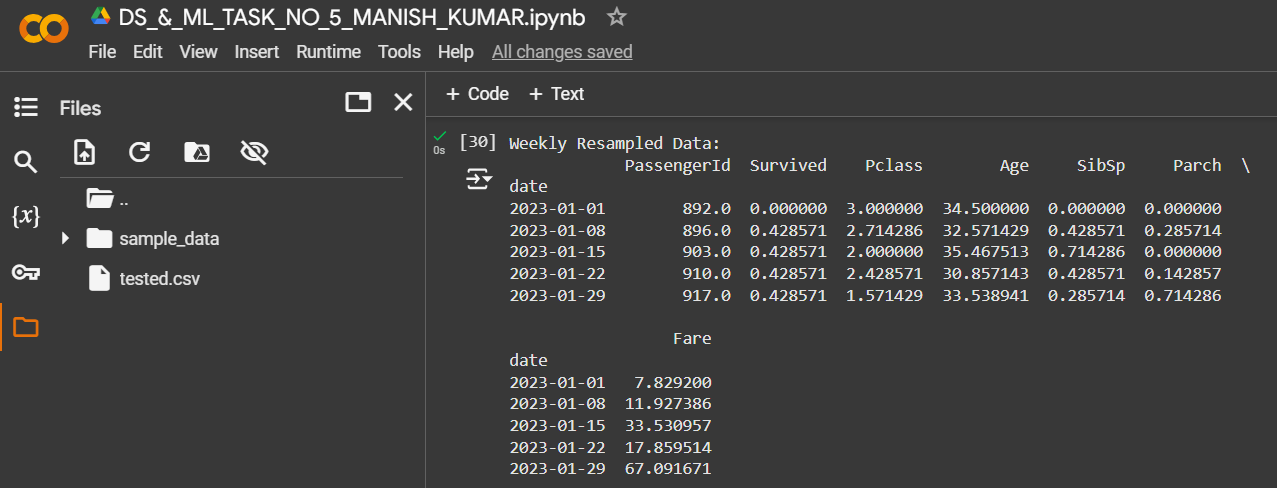
* **Dataset with Synthetic Datetime Column:**



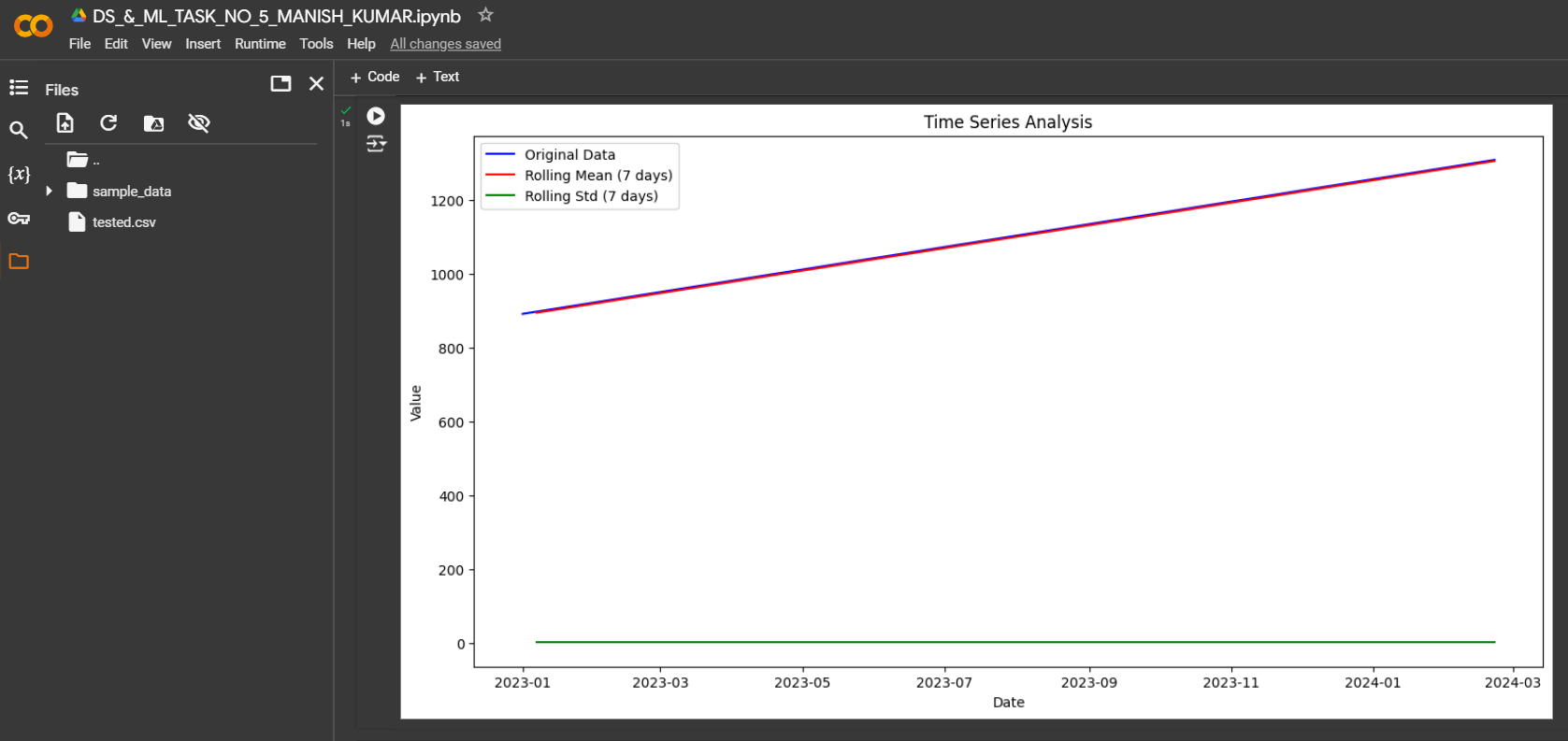
* **Missing values after Imputation:**



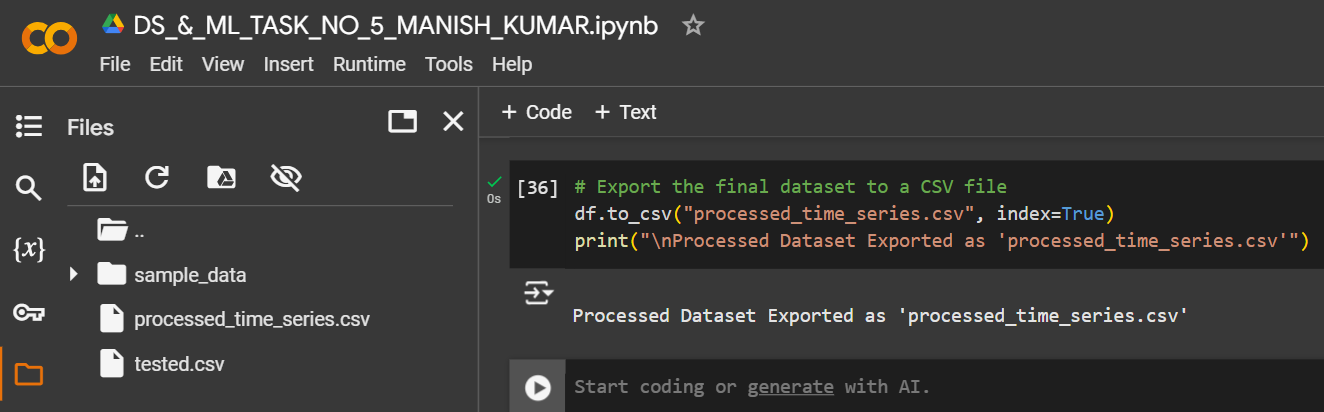
* **Weekly Resampled Data:**



* **Time Series Analysis:**



* **Processed Dataset Exported as 'processed\_time\_series.csv':**



1. **Describe Widget/Algorithm Used in Task**

**Algorithm Used:**

The task leverages the following algorithms and techniques:

* **Data Preprocessing**:
  + **Datetime Simulation**: A synthetic datetime column is added to mimic time-series data, and the column is set as the index to enable time-based operations.
  + **Handling Missing Values**: Missing values in numeric columns are replaced with the column mean for smoother analysis.
* **Resampling**:
  + The numeric data is resampled to a weekly frequency using Pandas' resample method, calculating weekly averages to understand long-term trends.
* **Rolling Statistics**:
  + **Rolling Mean**: A 7-day rolling window is applied to calculate the average value over the last 7 days, smoothing short-term fluctuations.
  + **Rolling Standard Deviation**: A 7-day rolling window is applied to measure variability in the data.
* **Lagged Features**:
  + **Lagging**: Shifted versions of the data are created (e.g., lagged by 1 day or 7 days) to capture time-dependent patterns.
* **Visualization**:
  + **Line Plots**: Rolling mean, rolling standard deviation, and the original data are plotted to visually compare trends and variability over time.

**Tools Used:**

* **Pandas**:

**\**

* + For data manipulation, resampling, rolling statistics, and handling missing values.
  + Used the resample, rolling, and shift functions to perform time-series-specific operations.
* **NumPy**:
  + Assisted in handling numerical operations while imputing missing values.
* **Matplotlib**:
  + For creating visualizations such as line plots to showcase trends, rolling statistics, and lagged features.
* **Jupyter/Colab Environment**:
  + Provided an interactive Python environment for implementing the time-series analysis and visualizing results.

**\*\*\* The End \*\*\***