

Combine Matplotlib with pandas to create advanced data visualizations with labeled data points and statistical overlays.

1. Add Task Description:

Objective:

Create advanced data visualizations by combining the capabilities of Matplotlib and pandas.

The task focuses on generating plots with:

Labeled data points for clarity in visualization.

Statistical overlays, such as trend lines, confidence intervals, or histograms.

Integration of pandas for data manipulation and seamless plotting.

2. Attach Screenshot Of Output.:



3. Describe Widget/Algorithm Used In Task:

pandas for Data Handling:

- **Purpose:** Prepare and process data efficiently.
- **Description:** pandas allows for cleaning, filtering, and grouping data before visualization.
- **Key Functions:**
 - `groupby()`: Aggregate data by categories.
 - `mean()`, `std()`: Calculate statistical summaries for error bars.

Matplotlib for Visualization:

- **Purpose:** Render detailed and customized plots.
- **Key Features Used:**
 - `scatter()`: Create scatter plots.
 - `annotate()`: Add labels to individual data points.
 - `errorbar()`: Include error bars for statistical overlays.
 - `plot()`: Add trend lines or regression curves.

Statistical Overlays:

- **Algorithm:** Use linear regression to fit a trend line.
- **Library:** Use numpy or scipy for regression.