

Assignment 3: Use Matplotlib or Seaborn to visualize data distributions and patterns in Generative AI datasets. Plot histograms, scatter plots, and heatmaps to analyze data characteristics.

Import Required Libraries

```
import pandas as pd      # pandas is used to create and handle tabular data using DataFrames
import matplotlib.pyplot as plt  #matplotlib.pyplot is used for customizing and showing
                                plots.
```

```
import seaborn as sns    #seaborn is built on top of matplotlib and makes beautiful statistical
                           charts easily.
```

Creating a small dataset of student scores

```
data = {
    'Student': ['A', 'B', 'C', 'D', 'E'],
    'Math': [85, 78, 92, 70, 88],
    'Science': [80, 75, 89, 65, 90],
    'English': [78, 82, 88, 72, 85]
}
df = pd.DataFrame(data)    # Convert to DataFrame
print(df)
```

Plot a Histogram- Histogram of Math scores

```
sns.histplot(df['Math'], kde=True, color='orange')  #sns.histplot(df['Math']):
                                                    Plots the distribution (frequency) of Math marks.,
                                                    kde=True: Adds a smooth curve (called Kernel
                                                    Density Estimate) to show the shape.,
```

```
plt.title("Histogram of Math Scores")
plt.xlabel("Math Score")
plt.ylabel("Number of Students")
plt.show()
```

Scatter Plot- Scatter plot between Math and Science

```
sns.scatterplot(x='Math', y='Science', data=df)    #In output Each dot represents
                                                    one student.
```

```
plt.title("Math vs Science Scores")
plt.xlabel("Math Score")
plt.ylabel("Science Score")
plt.show()
```

Heatmap (Correlation)- Correlation heatmap (only numeric data)

```
corr = df.drop('Student', axis=1).corr()
```

```
#df.drop('Student', axis=1): Removes the 'Student' column because it's not numeric.
corr(): Calculates correlation coefficients between subjects: Value +1: Strong positive
correlation, Value 0: No correlation, Value -1: Strong negative correlation.
```

```
sns.heatmap(corr, annot=True, cmap='YlGnBu')
```

```
#sns.heatmap(...): Creates a color-coded matrix: annot=True: Shows the correlation values in each cell, cmap='YlGnBu': Color style from Yellow-Green-Blue.
```

```
plt.title("Subject Score Correlation")
```

```
plt.show()
```