**Assignment -4**

**1. Create a Simple Dataset, Calculate Basic Statistics, and Visualize with Bar, Pie, and Box Plot**

**Step 1: Create Dataset**

data = randi(100, 1, 10); % Random dataset of 10 integers between 1 and 100

**Step 2: Calculate Basic Statistics**

mean\_data = mean(data);

median\_data = median(data);

std\_data = std(data);

**Step 3: Visualize**

% Bar Plot

figure;

bar(data);

title('Bar Plot');

xlabel('Data Index');

ylabel('Value');

% Pie Chart

figure;

pie(data);

title('Pie Chart');

% Box Plot

figure;

boxplot(data);

title('Box Plot');

**2. Apply Log and Square Root Transformations and Visualize Data**

**Step 1: Create Dataset**

data = randi(100, 1, 10); % Random dataset of 10 integers between 1 and 100

**Step 2: Apply Transformations**

log\_data = log(data);

sqrt\_data = sqrt(data);

**Step 3: Visualize**

% Original Data

figure;

subplot(3,1,1);

bar(data);

title('Original Data');

% Log Transformed Data

subplot(3,1,2);

bar(log\_data);

title('Log Transformed Data');

% Square Root Transformed Data

subplot(3,1,3);

bar(sqrt\_data);

title('Square Root Transformed Data');

**3. Calculate Correlation Coefficient and Visualize with Scatter Plot**

**Step 1: Create Dataset**

x = randi(100, 1, 10); % Random dataset for x

y = randi(100, 1, 10); % Random dataset for y

**Step 2: Calculate Correlation Coefficient**

correlation = corrcoef(x, y);

**Step 3: Visualize**

figure;

scatter(x, y);

title('Scatter Plot');

xlabel('X Data');

ylabel('Y Data');

**4. Import Data from CSV, Display Basic Information, and Visualize Features**

**Step 1: Import Data**

data = readtable('yourfile.csv'); % Replace 'yourfile.csv' with your actual file path

**Step 2: Display Basic Information**

summary(data);

**Step 3: Visualize Features**

figure;

histogram(data.YourColumnName); % Replace 'YourColumnName' with actual column name

title('Histogram of YourColumnName');

**5. Handle Missing Values, Normalize Data, and Visualize Distribution**

**Step 1: Create Dataset with Missing Values**

data = randi(100, 1, 10);

data(randi([1, 10], 1, 2)) = NaN; % Introduce some NaNs

**Step 2: Handle Missing Values**

data = fillmissing(data, 'linear'); % Replace NaNs with linear interpolation

**Step 3: Normalize Data**

normalized\_data = (data - min(data)) / (max(data) - min(data)); % Min-Max normalization

**Step 4: Visualize Distribution**

figure;

histogram(normalized\_data);

title('Normalized Data Distribution');