

School of Computer Science Engineering and Technology

Course- BTech
Course Code- CSET-214
Year- 2024-25
Date-

Type- Specialization Core
Course Name- Data Analysis using Python
Semester- Odd
Batch-

Lab # No. (9) Random variable, Bernoulli distribution, box and whisker plot, Stacked bar chart, Area Chart

CO Mapping

Lab No.	Name	CO1	CO2	CO3
9	Random variable, Bernoulli distribution, box and whisker plot, Stacked bar chart, Area Chart	↗	↗	↗

- 1.a) Write a Python program to simulate a random variable following a Bernoulli distribution. Generate 1000 random numbers, visualize them using a histogram, and calculate the mean, median, and mode of the dataset.
 - b) Write a Python program to simulate a random variable following a normal distribution. Generate 1000 random numbers, visualize them using a histogram, and calculate the mean, median, and mode of the dataset.
 - c) Write a Python program to Simulate rolling a fair 6-sided die 10000 times. Compute the mean, median, and mode of the outcomes and plot a bar chart showing the frequency of each outcome.
2. Write a Python program that detects outliers in a dataset using the median and interquartile range (IQR). Generate random data points (for e.g.- test scores). Add some outliers to the data and then identify outliers. Visualize the data using a boxplot.
3. a) Create a stacked bar chart between students and marks for the following data.

```
data = {  
    'Student': ['Student 1', 'Student 2', 'Student 3', 'Student 4', 'Student 5',  
               'Student 6', 'Student 7', 'Student 8', 'Student 9', 'Student 10'],  
    'Math': [85, 78, 92, 88, 76, 95, 89, 84, 91, 87],  
    'Physics': [79, 85, 88, 92, 80, 87, 90, 82, 85, 89],  
    'Chemistry': [91, 88, 83, 85, 79, 93, 89, 86, 87, 90] }
```
 - b) Show the Sales Trends by Product from Q1-2020 to Q2-2024 of following data using area chart.

School of Computer Science Engineering and Technology

```
quarters = [  
    'Q1 2020', 'Q2 2020', 'Q3 2020', 'Q4 2020',  
    'Q1 2021', 'Q2 2021', 'Q3 2021', 'Q4 2021',  
    'Q1 2022', 'Q2 2022', 'Q3 2022', 'Q4 2022',  
    'Q1 2023', 'Q2 2023', 'Q3 2023', 'Q4 2023',  
    'Q1 2024', 'Q2 2024'  
]  
  
# sales data for each product  
product_a_sales = [200, 180, 210, 230, 220, 240, 260, 280, 200, 180, 210, 230, 220, 240, 260, 280,  
300, 320]  
product_b_sales = [150, 160, 170, 190, 200, 210, 230, 250, 150, 160, 170, 190, 200, 210, 230, 250,  
270, 290]  
product_c_sales = [100, 120, 130, 140, 150, 160, 180, 190, 100, 120, 130, 140, 150, 160, 180, 190,  
200, 220]
```