DATA ANALYSIS AND VISUALIZATION

Week 8: Probability in Data Science

In-Class Activity: • Understand probability concepts relevant to data analysis.

Assignment: • Apply probability concepts to a given dataset, highlighting the significance in data analysis.

Answer:

```python
import numpy as np


#Importing pandas library
import pandas as pd


from scipy.stats import norm


# Uploading a dataset
df =pd.read_excel('/Exam records.xlsx')


# Calculate basic statistics
mean = df.mean()
std_dev = df.std()
median = df.median()


print("Mean:", mean)
print("Standard Deviation:", std_dev)
print("Median:", median)
```

```
    Mean: STUDENE ID     3.0
    EXAM SCORE    78.0
    dtype: float64
    Standard Deviation: STUDENE ID    1.581139
    EXAM SCORE    10.368221
    dtype: float64
    Median: STUDENE ID    3.0
    EXAM SCORE    80.0
    dtype: float64
```

```python
# Calculate the probability of a student scoring above a certain score
threshold_score = 80



print("Mean Score:", mean_score)
print("Standard Deviation:", std_dev)
print(f"Probability of scoring above {threshold_score}")
```

```
Mean Score: 78.0
Standard Deviation: 10.36822067666386
Probability of scoring above 80
```