Stats Exp 1

Experiment Number: 1

AIM: Import all the required Python libraries and perform statistical functions.

Software Used: Python with NumPy and SciPy libraries (presumably in a Jupyter Notebook or similar environment)

Theory:

This experiment explores fundamental statistical concepts and their implementation in Python. Key concepts include:

- **Measures of Central Tendency:**
  - **Mean:** The average of a dataset, calculated by summing all values and dividing by the number of values.
  - **Median:** The middle value in a sorted dataset, representing the 50th percentile.
  - **Mode:** The most frequent value in a dataset.
- **Measures of Dispersion:**
  - **Variance:** A measure of how spread out the data is around the mean. It's calculated as the average of the squared differences from the mean.
  - **Standard Deviation:** The square root of the variance. It provides a more interpretable measure of data spread in the same units as the original data.

Dataset Information:

The experiment uses a sample dataset: `data = [10, 20, 20, 30, 40, 40, 40, 50, 60]`

This dataset appears to be a small, arbitrary set of numbers for demonstration purposes. No further information about the data's origin or meaning is provided.

Code:

Python

```
import numpy as np
from scipy import stats

# Sample data
data = [10, 20, 20, 30, 40, 40, 40, 50, 60]

# Mean
mean_value = np.mean(data)
```

```
print(f"Mean: {mean_value}")

# Median
median_value = np.median(data)
print(f"Median: {median_value}")

# Mode
mode_value = stats.mode(data)
print(f"Mode: {mode_value.mode[0]} (appears {mode_value.count[0]} times)")

# Variance
variance_value = np.var(data, ddof=1)  # ddof=1 for sample variance
print(f"Variance: {variance_value}")

# Standard Deviation
std_deviation_value = np.std(data, ddof=1)  # ddof=1 for sample standard
deviation
print(f"Standard Deviation: {std_deviation_value}")
```

Output:

```
Mean: 34.444444444444444
Median: 40.0
Mode: 40 (appears 3 times)
Variance: 302.77777777777777
Standard Deviation: 17.40313619770067
```

Result:

The code successfully calculated the mean, median, mode, variance, and standard deviation of the sample dataset. The results show that the data is centered around 40 (median and mode), with a moderate degree of dispersion (standard deviation of 17.4).

Learning Outcomes:

- Learned how to import and utilize NumPy and SciPy libraries for statistical analysis in Python.
- Gained practical experience in calculating essential descriptive statistics, including measures of central tendency and dispersion.
- Understood the importance of these measures in summarizing and interpreting data.
- Developed skills in writing and executing Python code for basic statistical tasks.