```
In [2]:  import pandas as pd
         df = pd.read_csv("Week-5-Student-Scores 1.csv")
```

```
In [3]:  df.shape
```

```
Out[3]:  (121, 10)
```

```
In [4]:  df.dtypes
```

```
Out[4]:  Student_ID               object
         Gender                   object
         Subject                  object
         Study_Hours_per_Week    float64
         Attendance_Rate         float64
         Previous_Score          float64
         Quiz1                   float64
         Quiz2                   float64
         Quiz3                   float64
         Final_Score             float64
         dtype: object
```

```
In [5]:  df.isnull().sum()
```

```
Out[5]:  Student_ID              0
         Gender                 0
         Subject                0
         Study_Hours_per_Week   5
         Attendance_Rate        5
         Previous_Score         0
         Quiz1                  0
         Quiz2                  5
         Quiz3                  0
         Final_Score            0
         dtype: int64
```

```
In [6]:  df.duplicated().sum()
```

```
Out[6]:  1
```

```
In [7]:  diagnostics = pd.DataFrame({
             'Data Type': df.dtypes,
             'Missing Values': df.isnull().sum(),
             'Missing %': round((df.isnull().sum()/ len(df)) *100, 2) ,
             'Unique Values': df.nunique(),
             'Duplicates':[df.duplicated().sum()] + [''] * (len (df.columns) -1)
         })

         print(f"Shape of dataset: {df.shape}")
         diagnostics
```

```
         Shape of dataset: (121, 10)
```

	Data Type	Missing Values	Missing %	Unique Values	Duplicates
Student_ID	object	0	0.00	120	1
Gender	object	0	0.00	2	
Subject	object	0	0.00	3	
Study_Hours_per_Week	float64	5	4.13	76	
Attendance_Rate	float64	5	4.13	27	
Previous_Score	float64	0	0.00	45	
Quiz1	float64	0	0.00	45	
Quiz2	float64	5	4.13	51	
Quiz3	float64	0	0.00	43	
Final_Score	float64	0	0.00	39	

In [8]:
```python
# Get descriptive statistics for numeric columns
descriptive_stats = df.describe()
```

In [10]:
```python
# 3. Select Required Statistics and Rename (Corrected)
required_stats = ['mean', '50%', 'std', 'min', 'max', '25%', '75%']
descriptive_stats = descriptive_stats.loc[required_stats]
descriptive_stats = descriptive_stats.rename(index={'50%': 'median'})
```

In [11]:
```python
descriptive_stats = descriptive_stats.T
```

In [12]:
```python
print(descriptive_stats)
```

```
                          mean   median         std    min   max     25%     75%
Study_Hours_per_Week   8.050000     8.20    2.982806   0.00  19.6    6.00   10.00
Attendance_Rate        0.888017     0.89    0.067711   0.74   1.0    0.83    0.94
Previous_Score        59.991736    60.00   12.087525  34.00  98.0   51.00   69.00
Quiz1                 48.512397    49.00   12.518064  18.00  84.0   40.00   55.00
Quiz2                 49.750000    49.50   13.592117  15.00  81.0   42.00   58.00
Quiz3                 51.958678    52.00   11.236397  16.00  77.0   44.00   61.00
Final_Score           48.123967    47.00    9.460065  23.00  73.0   42.00   55.00
```

In [ ]:

Loading [MathJax]/extensions/Safe.js