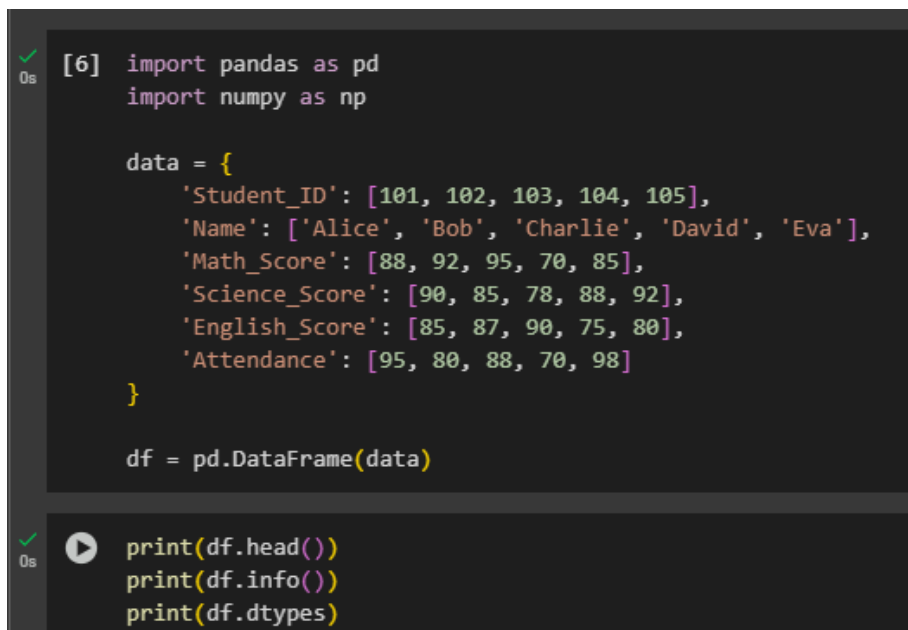


**AIM:**

To inspect and analyze datasets by viewing DataFrames, filtering and subsetting data using conditions, and calculating descriptive statistics including measures of central tendency and dispersion.

## 1. Viewing and inspecting DataFrames



```
[6] import pandas as pd
import numpy as np

data = {
    'Student_ID': [101, 102, 103, 104, 105],
    'Name': ['Alice', 'Bob', 'Charlie', 'David', 'Eva'],
    'Math_Score': [88, 92, 95, 70, 85],
    'Science_Score': [90, 85, 78, 88, 92],
    'English_Score': [85, 87, 90, 75, 80],
    'Attendance': [95, 80, 88, 70, 98]
}

df = pd.DataFrame(data)

print(df.head())
print(df.info())
print(df.dtypes)
```

OUTPUT:

```

Student_ID  Name  Math_Score  Science_Score  English_Score  Attendance
0         101   Alice         88             90             85             95
1         102    Bob         92             85             87             80
2         103  Charlie        95             78             90             88
3         104   David        70             88             75             70
4         105    Eva         85             92             80             98
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Student_ID      5 non-null     int64
1   Name            5 non-null     object
2   Math_Score      5 non-null     int64
3   Science_Score   5 non-null     int64
4   English_Score   5 non-null     int64
5   Attendance      5 non-null     int64
dtypes: int64(5), object(1)
memory usage: 372.0+ bytes
None
Student_ID      int64
Name            object
Math_Score      int64
Science_Score   int64
English_Score   int64
Attendance      int64
dtype: object

```

## 2. Filtering and subsetting data using conditions

```

[ ] =low_attendance = df[df['Attendance'] < 85]
    print("Students with low attendance:\n", low_attendance)

    =high_math = df[df['Math_Score'] > 90]
    print("Students with high Math scores:\n", high_math)

```

```

Students with low attendance:
  Student_ID  Name  Math_Score  Science_Score  English_Score  Attendance
1         102    Bob         92             85             87             80
3         104   David        70             88             75             70
Students with high Math scores:
  Student_ID  Name  Math_Score  Science_Score  English_Score  Attendance
1         102    Bob         92             85             87             80
2         103  Charlie        95             78             90             88

```

## 3. Descriptive statistics: measures of central tendency (mean, median, mode) and measures of dispersion (range, variance, standard deviation)

```

print("Mean:\n", df[['Math_Score', 'Science_Score', 'English_Score']].mean())
print("Median:\n", df[['Math_Score', 'Science_Score', 'English_Score']].median())
print("Mode:\n", df[['Math_Score', 'Science_Score', 'English_Score']].mode())

print("Range:\n", df[['Math_Score', 'Science_Score', 'English_Score']].max() - df[['Math_Score', 'Science_Score', 'English_Score']].min())
print("Variance:\n", df[['Math_Score', 'Science_Score', 'English_Score']].var())
print("Standard Deviation:\n", df[['Math_Score', 'Science_Score', 'English_Score']].std())

```



```
Mean:
  Math_Score      86.0
  Science_Score   86.6
  English_Score   83.4
dtype: float64
Median:
  Math_Score      88.0
  Science_Score   88.0
  English_Score   85.0
dtype: float64
Mode:
   Math_Score  Science_Score  English_Score
0           70             78             75
1           85             85             80
2           88             88             85
3           92             90             87
4           95             92             90
Range:
  Math_Score      25
  Science_Score   14
  English_Score   15
dtype: int64
Variance:
  Math_Score      94.5
  Science_Score   29.8
  English_Score   35.3
dtype: float64
Standard Deviation:
  Math_Score      9.721111
  Science_Score   5.458938
  English_Score   5.941380
dtype: float64
```