Summary Statistics Python Notebok for Individual Project 1

Step 1: Import relevant Python functions from library.py.

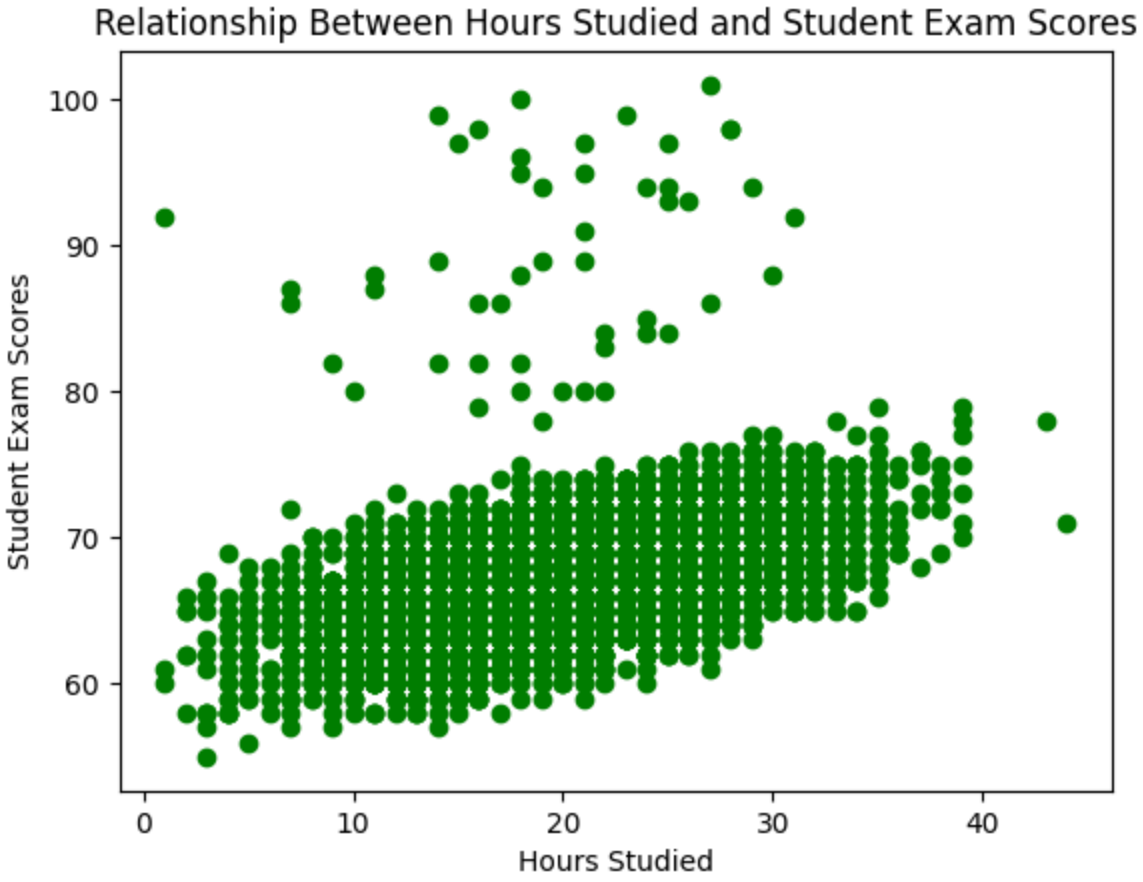
Step 2: Define functions to produce summary statistics and visualizations using generate_summary_stats, grab_median, generate_study_hours_viz, and generate_sleep_viz functions from library.py.

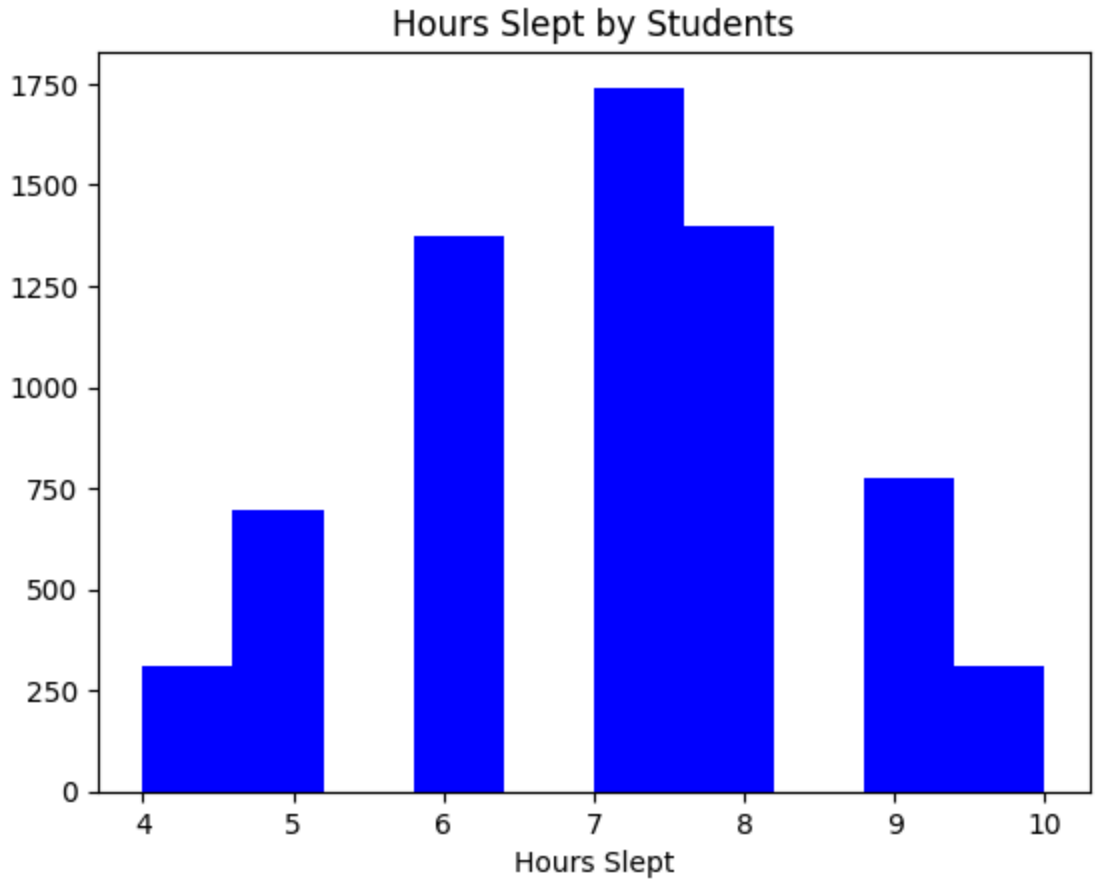
In [8]:
```
# NBVAL_IGNORE_OUTPUT
from lib.library import (
    generate_summary_stats,
    grab_median,
    generate_sleep_viz,
    generate_study_hours_viz,
)


def summarize():
    """Using the Student Performance csv and summary statistics functions fr
    library.py, this function produces summary statistics (mean, median,
    mode, standard deviation, percentiles, max, and min) for each column of
    """
    basic_summary = generate_summary_stats()
    median = grab_median()
    return basic_summary, median


def create_visualizations():
    """This function generates scatterplot and histogram visualizations of t
    the respective functions from library.py."""
    generate_study_hours_viz()
    generate_sleep_viz()
```

Step 3: Validate the outputs of the summary stats & visualizations.

In [9]:
```
summary = summarize()
describe_stats = summary[0]
medians = summary[1]
assert describe_stats.loc["mean"]["Previous_Scores"] == 75.07053125472983
assert medians["Sleep_Hours"] == 7.0
assert describe_stats.loc["std"]["Physical_Activity"] == 1.0312310926271286
```

In [10]:
```
create_visualizations()
```


Relationship Between Hours Studied and Student Exam Scores
Student Exam Scores
100
90
80
70
60
0
10
20
30
40
Hours Studied


Hours Slept by Students
1750
1500
1250
1000
750
500
250
0
4
5
6
7
8
9
10
Hours Slept