In [90]:
```python
%config IPCompleter.greedy=True
```

Using Python, create your own having columns plant name, sunlight exposure, plant height and answer the following questions:

- Is there a relationship between the number of hours of sunlight exposure and the height of the plants?
- Visualize the relationship between sunlight exposure and plant height using a scatterplot.
- Calculate the correlation coefficient between sunlight exposure and plant height. Is the correlation positive or negative? Is it strong or weak?
- Based on the correlation coefficient, can we conclude that there is a significant association between sunlight exposure and plant growth rate?

In [79]:
```python
import pandas as pd
import matplotlib.pyplot as plt
```

In [80]:
```python
plant_sunlight_data = {
    "name": ["pomegranate", "guava", "persimmon", "passionfruit", "apple",
"pea", "maize", "cotton", "peepal", "basil"],
    "exposure": [85, 60, 65, 90, 63, 32, 68, 101, 22, 49],
    "height": [12, 9, 8, 12, 7, 4, 9, 10, 3, 9]
}
```

In [81]:
```python
df = pd.DataFrame(plant_sunlight_data)
df
```

Out[81]:

	name	exposure	height
0	pomegranate	85	12
1	guava	60	9
2	persimmon	65	8
3	passionfruit	90	12
4	apple	63	7
5	pea	32	4
6	maize	68	9
7	cotton	101	10
8	peepal	22	3
9	basil	49	9

In [82]:
```python
r_df = df[["exposure", "height"]]
r_df
```

Out[82]:

	exposure	height
0	85	12
1	60	9
2	65	8
3	90	12
4	63	7
5	32	4
6	68	9
7	101	10
8	22	3
9	49	9

In [83]:
```python
corr_coeff = r_df["exposure"].corr(r_df["height"])
```

In [84]:
```python
if corr_coeff > 0.4:
    print("There is a relationship between the Hours of Sunlight Exposure
and the height of the plant.")
else:
    print("There is no sign of a relationship between the Hours of Sunlight
Exposure and the height of the plant.")
```

There is a relationship between the Hours of Sunlight Exposure and the height of the plant.

In [85]:
```python
plt.title("Hours of Sunlight Exposure and the Height of the plant (Scatter
plot)")
plt.xlabel("Plant Height (in cm)")
plt.ylabel("Sunlight Exposure (in hours)")
plt.grid()
plt.scatter(r_df["exposure"], r_df["height"], color="b", marker="d",
label="Data Point")
plt.legend()
```

Out[85]: <matplotlib.legend.Legend at 0x77b5e4223610>

Hours of Sunlight Exposure and the Height of the plant (Scatter plot)
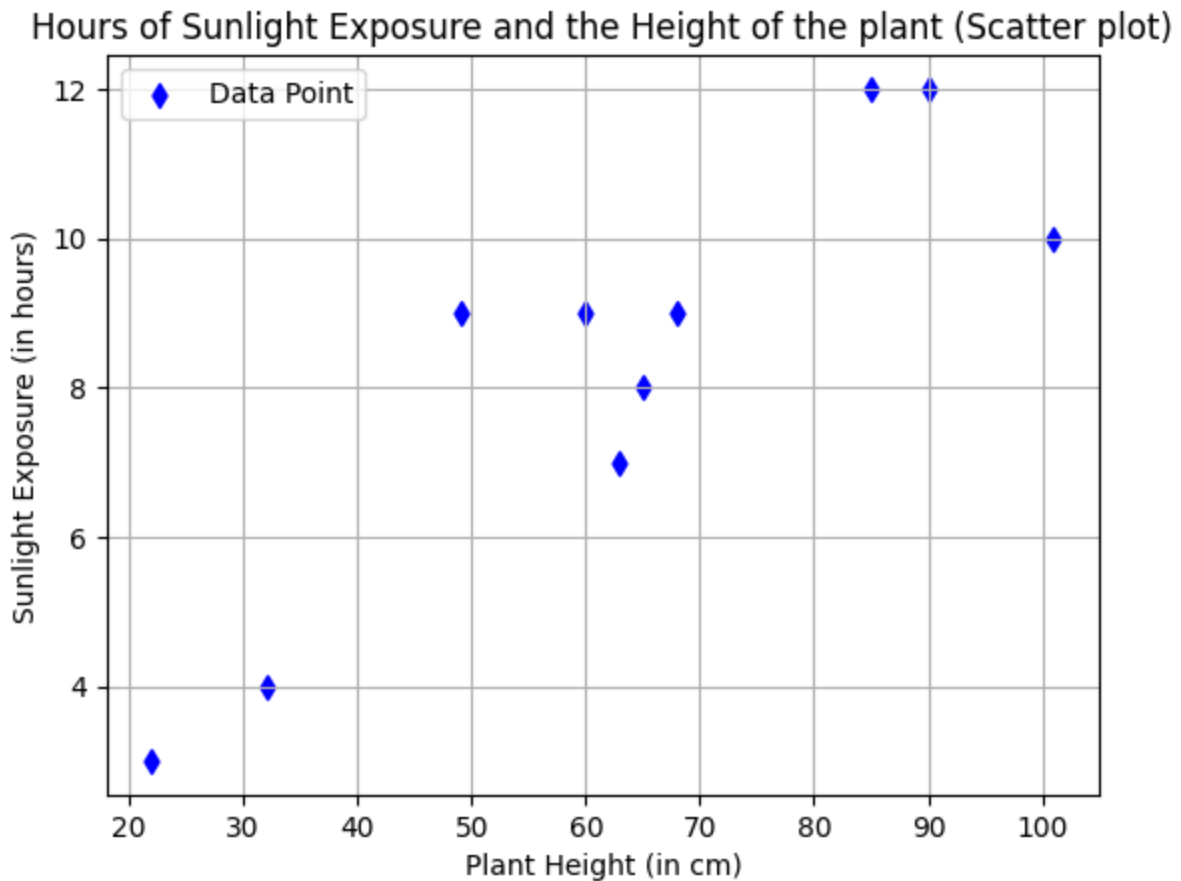
In [86]:
```python
if corr_coeff < 0:
    sign = "negative"
elif corr_coeff > 0:
    sign = "positive"
else:
    sign = "none"
```

In [87]:
```python
if abs(corr_coeff) >= 0.85:
    strength = "very strong"
elif abs(corr_coeff) < 0.85 and abs(corr_coeff) >= 0.5:
    strength = "fairly strong"
else:
    strength = "weak"
```

In [89]:
```python
print(f"The correlation coefficient is {corr_coeff}, and the correlation is
{sign} and {strength} in nature.")
if sign == "positive" and strength == "very strong":
    print("From the correlation coefficient, we can deduce that there is a
significant association between the parameters: Sunlight Exposure (in
hours) and Plant Height (in cm).")
```

The correlation coefficient is 0.8826924553887492, and the correlation is positive and very strong in nature.
From the correlation coefficient, we can deduce that there is a significant association between the parameters: Sunlight Exposure (in hours) and Plant Height (in cm).