In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
```

In [2]:
```
data = pd.read_csv('House Price India.csv')
```

In [8]:
```
data.head()
```

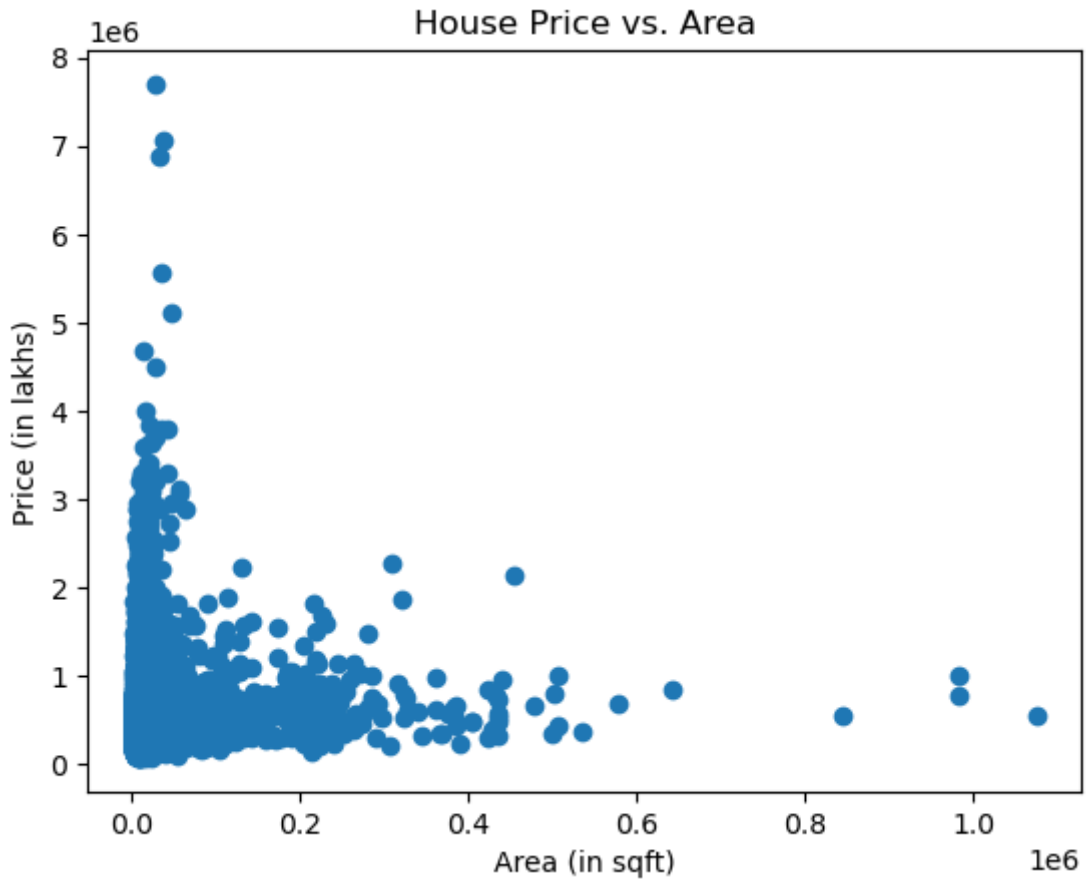
Out[8]:

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views	condit of ho
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4	
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0	
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0	
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0	
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0	

5 rows × 23 columns

In [4]:
```
price = data['Price']
area = data['lot area']
```

In [5]:
```
plt.scatter(area, price)
plt.title('House Price vs. Area')
plt.xlabel('Area (in sqft)')
plt.ylabel('Price (in lakhs)')
plt.show()
```



In [6]:
```
correlation = price.corr(area)

print('Correlation Coefficient:', correlation)
```

```
Correlation Coefficient: 0.08199199695424976
```

In [7]:

Out[7]:

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views	condit of ho
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4	
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0	
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0	
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0	
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0	

5 rows × 23 columns

In [ ]: