```
In [1]:  import csv
         l1 = ['a', 'b', 'c','d', 'e', 'a', 'b']

         l2 = []

         print(l1)
         len1 = len(l1)
         print('list length = ', len1)

         l2 = set(l1)  #Copies the only distinct elements from the source
         print(l2)
         len2 = len(l2)
         print('list length = ', len2)
```

```
['a', 'b', 'c', 'd', 'e', 'a', 'b']
list length =  7
{'b', 'c', 'd', 'a', 'e'}
list length =  5
```
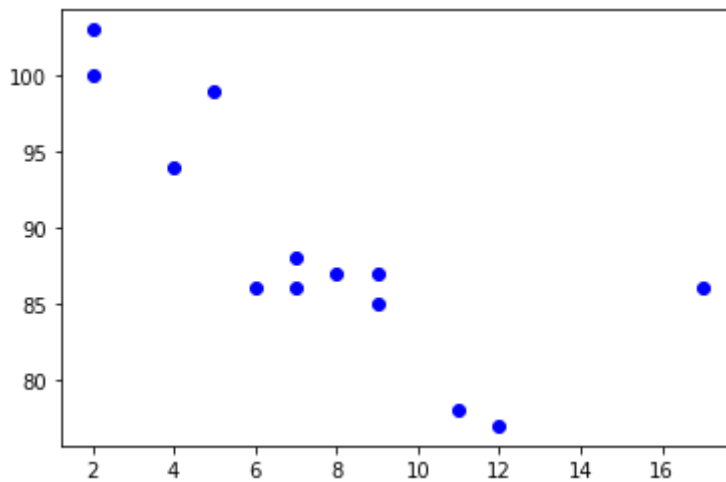
```
In [2]:  import matplotlib.pyplot as plt

         x =[5, 7, 8, 7, 2, 17, 2, 9, 4, 11, 12, 9, 6]

         y =[99, 86, 87, 88, 100, 86, 103, 87, 94, 78, 77, 85, 86]

         plt.scatter(x, y, c ="blue")
         # To show the plot

         plt.show()
```



```
In [10]:  import matplotlib.pyplot as plt

          x =[5, 7, 8, 7, 2, 17, 2, 9, 4, 11, 12, 9, 6]

          y =[99, 86, 87, 88, 100, 86, 103, 87, 94, 78, 77, 85, 86]

          plt.scatter(x, y, c ="red")

          # To show the plot
          plt.xlabel('Hours of study')

          plt.ylabel('ExamGrade')

          plt.title('Regression Graph')

          plt.show()
```
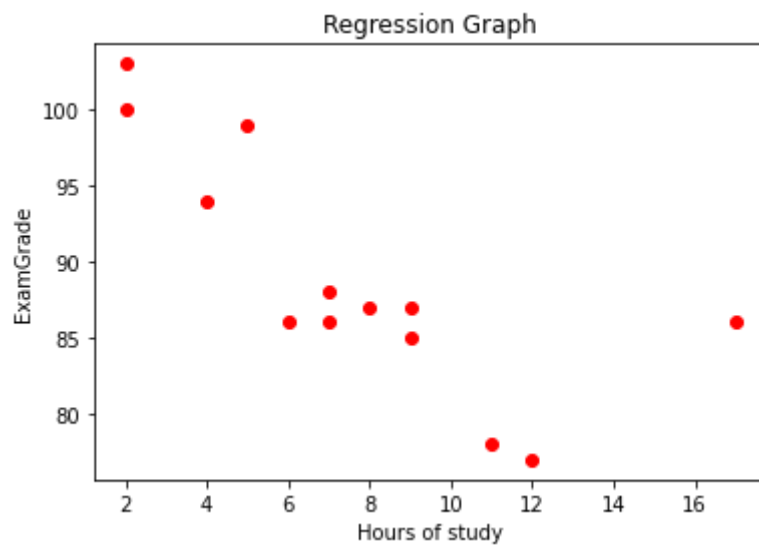
Regression Graph

In [11]:
```python
import matplotlib.pyplot as plt
import math
import numpy as np

x =np.arange(0, math.pi * 2, 0.05)

y =np.sin(x)

#plt.scatter(x, y, c ="red")

plt.scatter(x, y)
# To show the plot
plt.xlabel('angle')

plt.ylabel('sine of x')

plt.title('Sine Curve')

plt.show()
```
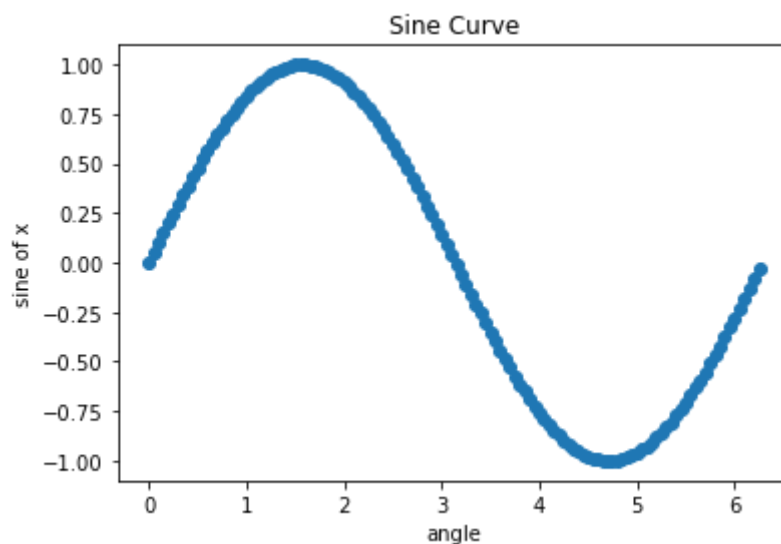
Sine Curve

In [14]:
```python
import matplotlib.pyplot as plt
import math
import numpy as np

x =np.arange(0, math.pi * 2, 0.05)

y =np.sin(x)
```
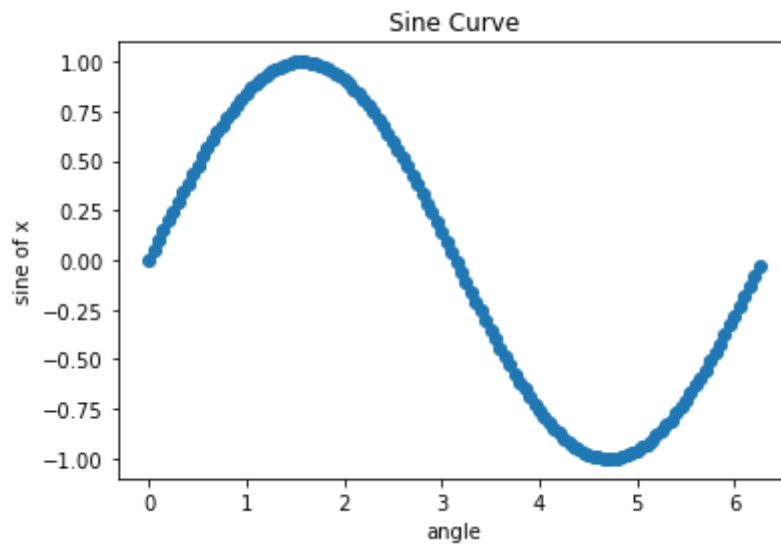
```python
#plt.scatter(x, y, c ="red")

plt.scatter(x, y)
# To show the plot
plt.xlabel('angle')

plt.ylabel('sine of x')

plt.title('Sine Curve')

plt.show()

%matplotlib inline
```



```python
In [16]:  import pandas as pd
          df1 = pd.read_csv('D:/csvsample/sample1.csv')
```

```python
In [21]:  df1.head(1)
```

Out[21]:

	Name	Branch	Year	CGPA
0	Nikhil	COE	2	9.0

```python
In [22]:  print(df1.columns)
```

```
Index(['Name', 'Branch', 'Year', 'CGPA'], dtype='object')
```

```python
In [23]:  type(df1)
```

Out[23]: pandas.core.frame.DataFrame

```python
In [25]:  print(df1.columns[1])
```

```
Branch
```

```python
In [26]:  df1.keys()
```

Out[26]: Index(['Name', 'Branch', 'Year', 'CGPA'], dtype='object')

```python
In [28]:  df1.shape[0]
```

Out[28]: 6

```python
In [29]:  df1.describe()
```

Out[29]:

	Year	CGPA
count	6.000000	6.000000
mean	2.000000	8.966667
std	0.632456	0.598888
min	1.000000	7.800000
25%	2.000000	9.025000
50%	2.000000	9.100000
75%	2.000000	9.250000
max	3.000000	9.500000

In [30]:
```python
df1.shape
```

Out[30]: (6, 4)

In [31]:
```python
print(df1.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 4 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Name    6 non-null      object
 1   Branch  6 non-null      object
 2   Year    6 non-null      int64
 3   CGPA    6 non-null      float64
dtypes: float64(1), int64(1), object(2)
memory usage: 320.0+ bytes
None
```

In [38]:
```python
import pandas as pd

df = pd.DataFrame({'RegNo': [100, 101, 102, 103, 104],
                   'Name': ['abc', 'def', 'ghi', 'jkl', 'mno'],
                   'M1': [25, 35, 45, 55, 65] })

print('The created data frame:\n')
display(df)

df.describe()

df.shape
```

The created data frame:

	RegNo	Name	M1
0	100	abc	25
1	101	def	35
2	102	ghi	45
3	103	jkl	55
4	104	mno	65

Out[38]: (5, 3)