```
import pandas as pd
```

```
from numpy.random import randn
from numpy.random import seed
from scipy.stats import pearsonr

seed(1)

data1 = randn(1000) + 100
data2 = data1 + (10*randn(1000) + 50)

corr,_ = pearsonr(data1,data2)
print('Pearsons Correlation:%.3f' % corr)
```

```
Pearsons Correlation:0.116
```

```
from math import *
def square_rooted(x):
  return round(sqrt(sum([a * a for a in x])))
def cosine_similarity(x,y):
  numerator = sum(a * b for a,b in zip(x,y))
  denominator = square_rooted(x) * square_rooted(y)
  return round(numerator / float(denominator),3)
print(cosine_similarity([3,45,7,2] , [2,54,13,15]))
```

```
0.958
```

```
from math import *
def jaccard_similarity(x,y):
  ic = len(set.union(*[set(x),set(y)]))
  uc = len(set.union(*[set(x),set(y)]))
  return ic/float(uc)
print(jaccard_similarity([0,1,2,5,6] , [0,2,3,5,7,9]))
```

```
1.0
```

```
from math import *
def euclidean_distance(x,y):
  return sqrt(sum(pow(a -b ,2) for a,b in zip(x,y)))
print(euclidean_distance([0,3,4,5] ,[7,6,3,-1]))
```

```
9.746794344808963
```

```
def manhattan_distance(x,y):
  return sum(abs(a -b) for a,b in zip(x,y))
print(manhattan_distance([10,20,10], [10,20,20]))
```

```
10
```